OPTIMAL CONTROL FOR A CLASS OF HEREDITARY SYSTEMS

by

Nicolaos D. Georganas

Submitted to the Department of Electrical Engineering, University of Ottawa, in partial fulfillment of the requirements for the degree of Doctor of Philosophy.

Department of Electrical Engineering
Faculty of Pure and Applied Science
University of Ottawa
Ottawa-Canada

July 1970

ABSTRACT

Optimal control problems for hereditary systems have been given a great deal of attention in recent years.

In this thesis, a class of nonlinear systems with finite heredity in the state and the control variables is considered. Using Pontryagin's geometric approach, necessary conditions for the optimal control, optimal initial control, and optimal initial state function are derived, extending Pontryagin's and Friedman's [2] Maximum Principles.

The introduction of optimal initial data, in addition to the optimal control, is considered to be very important for the optimization of several physical systems whose evolution depends on past data.

AKNOWLEDGEMENTS

I am very grateful to my advisor , Professor G. S. Glinski, for his encouragement , understanding and guidance .

No thanks can express my gratitude to Dr. N. U. Ahmed , for the kind words , stimulating discussions and critical comments , without which this thesis would not have been possible .

Thanks are also due to the members of the staff and the graduate students of the Electrical Engineering Department , for their friendship throughout the years of this research .

The financial assistance of the National Research Council of Canada , through an N. R. C. Postgraduate Scholarship, is gratefully aknowledged .

TABLE OF CONTENTS

CHAPTER 1

INTRODUCTION

In the mathematical formulation of a physical process the simplest method is to assume that the future behavior of the system depends only upon the present state and not at all upon its past history, and, furthermore, that the influence of the present state is instantaneous. This assumption leads to a system of ordinary differential equations by which a great variety of physical processes can be satisfactorily described.

For some physical systems, however, the above assumption is not justifiable. Their future behavior is dependent not only on the present state but also in some measure of their past and, indeed, in some cases on their immediate future. Systems of that type appear in the fields of economics, population studies, biophysics, physiology, traffic flow, etc. In [11] and [14] an excellent bibliography for such systems is found.

In case where the system's evolution depends not only on its present state but also on its past history, we call the system hereditary in the language of Emile Picard [17] and Vito Volterra [18]. The first problem where we find a process which will be called today hereditary, is, it appears [20], the problem of " Elastische Nachwirkungen " treated in 1874 by L. Boltzmann [19]. However, E. Picard is the one who explicitly indicated the distinction between hereditary and non-hereditary mechanics in a paper dated 1907 [17]. The mathematical theory of hereditary phaenomena was founded by Vito Volterra [18] and the Italian school of mathematicians of the early 20^{th} century.

The mathematical formulation of such processes by a system of ordinary differential equations is not any more possible. Instead, they may be described by a system of difference-differential, integro-differential, integro-difference-differential, or, in general, functional-differential equations.

Control problems for hereditary systems have been given a great deal of attention in recent years. An excellent bibliography can be found in [4] . An example of a control process with time-delay in the state and the control variables is the following 'bipropellant gas-pressurized liquid rocket system' [16] shown in Fig. 1.1 .

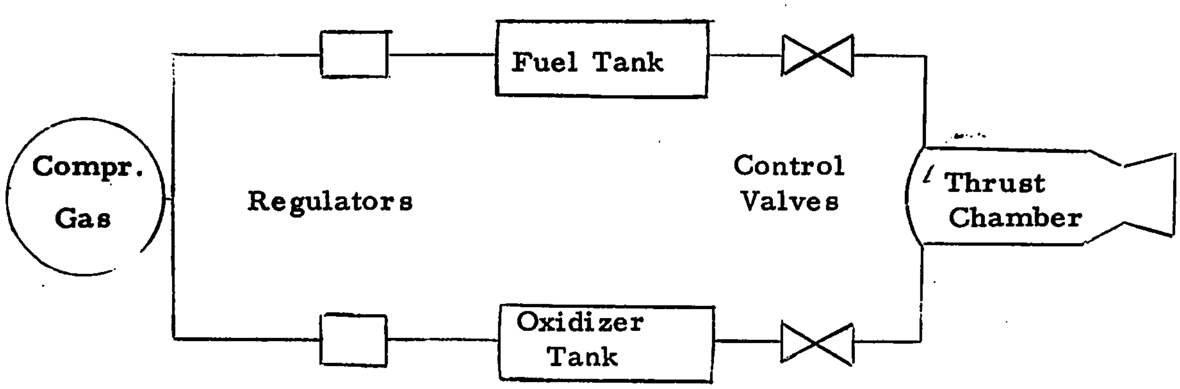

Fig. 1.1

Here, two distinguishable mechanisms contribute to the problem of stability and control : dynamic behavior of the feed system and combustion phaenomena in the thrust chamber . The plant shown in Fig. 1.1 can be represented mathematically by a linear vector difference-differential equation of the form :

(1.1) $\dot{x}(t) = Ax(t) + Bx(t-\beta) + Cu(t) + Du(t-\beta)$

where A and B are constant n xn matrices, C and D are constant n x m matrices, x(t) is an n- vector representing the state of the system, u(t) is an m- vector representing the control, and $\beta > 0$ is a constant time delay. The delayed terms in (1.1) arise from substituting algebraic equations for the control valves into the differential equations for the feed system and combustion chamber . Physically, this means that the present rate of change of certain of the state variables (e. g. ,feed system pressures and flows) depends on present values of the control, whereas the present rate of change of other state variables (e. g. , Chamber pressure) depends on past values of the control .

In particular, optimal control problems for hereditary systems have been treated extensively in the last decade . Kharatishvili [3]has extended Pontryagin's [1] Maximum Principle for systems with discrete delays in the state and the control variables . Friedman [2]considered the case of continuous heredity in the state and the control and has given necessary conditions, existence and uniqueness theorems for the optimal control and some results in controllability. Oǧuztöreli [4] considered the control problem of general hereditary systems and has given, among several important results , existence theorems for the optimal control. Halanay [5] has extended Pontryagin's Maximum principle for a very general class of time-lag systems. In his work he used the abstract multipliers rule of Hestenes [12] and he considered the initial control vector as given . Chyung and Lee [21] considered several interesting problems with delays in the state variables only and Lee [6] gave necessary, sufficient and existence conditions for the optimal control and the optimal initial control, for a simple linear system with one delay in the state and the control .

In this thesis we consider a fairly general class of nonlinear systems with both discrete and continuous heredity in the state and the control variables.

In Chapter 2, preliminary material, necessary for the subsequent chapters, is presented.

In Chapter 3, we derive necessary conditions (extension to Pontryagin's Maximum Principle) for the optimal control and the optimal initial control, with the complete proof given not in this chapter but in Appendix A, to facilitate reading. The initial state function is assumed to be given beforehand and several cases are examined. In the linear time-optimal case, a 'bang-bang' principle and a uniqueness theorem for the optimal control and the optimal initial control are obtained. Examples demonstrate the effect of the optimal choice of the initial control.

In Chapter 4, we derive an extension of the Maximum Principle of Chapter 3 for the case where the initial state function is not specified beforehand but chosen in an optimal fashion; that is, necessary conditions for the optimality of the pair $\{u, \phi\}$ are obtained. An example shows the effect of the optimal choice of the initial state function on the cost.

Finally, conclusions and suggestions for further study are given.

CHAPTER 2

PRELIMINARIES

2.1 NOTATION AND DEFINITIONS

The following notation is going to be used throughout the thesis. The real Euclidean n- space will be denoted by E^n. An element $x \in E^n$ is a column vector, e.g.,

$$x = \begin{bmatrix} x^1 \\ x^2 \\ . \\ .. \\ x^n \end{bmatrix}$$

Capital letters will denote matrices, e.g. A, B, C, etc. The transpose operation will be denoted by T, e.g. A^T, x^T.

If $x, y \in E^n$ then we define the inner product of x and y as

$$(x, y) = \sum_{i=1}^{n} x^i y^i$$

and the norm as

$$|x| = \sum_{i=1}^{n} |x^i|$$

The orthogonality condition $(x, y) = 0$ will be denoted be $x \perp y$

Definition 2.1.1 Compactness in E^n

A set $U \subset E^n$ is called compact if it is closed and bounded.

Definition 2.1.2 Measurable functions

A function $u : I \rightarrow E^n$ is called (Lebesgue) measurable if $u^{-1}(B) = \{ t \in I : u(t) \in B \}$ is a (L-) measurable set for any Borel set $B \subset E^r$.

Definition 2.1.3. Regular (or Lebesgue) points

Let $u(t)$, $t \in I = [t_0, t_1]$, be a measurable function taking values in $U \subset E^r$. We shall say that the point $\tau \in I$ is a regular point for the function u , if

$$\lim_{\mu(I_\tau) \rightarrow 0} \frac{\mu(A_\tau \cap I_\tau)}{\mu(I_\tau)} = 1$$

for every $I_\tau \subset I$ containing τ , where $A_\tau = \{t \in I : u(t) \in N(u(\tau))\} = u^{-1}(N(u(\tau)))$, N being a neihborhood of $u(\tau)$ in U , and μ is the Lebesgue measure on the real line .

It is to be noted that

(i) almost all points in I are regular points for u , i.e. , the set of non-regular points has L- measure zero [26] .

(ii) points of discontinuity of u are not regular points .

Definition 2.1.4 Continuous functions

Let x(t) be an n- vector valued function defined on [a, b] . If for any given $\epsilon > 0$, there exists a $\delta > 0$, such that

$$|x(t) - x(t_0)| < \epsilon$$

whenever

$$|t - t_0| < \delta$$

x(t) is said to be continuous at the point t_0 . If this happens everywhere on [a, b], then x(t) is said to be continuous on [a, b].

The space of all continuous functions defined on [a, b] and taking values in E^n will be denoted by $C([a,b], E^n)$.

Definition 2.1.5 Absolutely continuous functions

Let x(t) be an n- vector valued function defined on [a,b]. Suppose that $(t_1, t_1+h_1), (t_2, t_2+h_2), \ldots, (t_k, t_k+h_k)$ are nonoverlapping intervals in (a,b). x(t) is said to be absolutely continuous on [a,b] if, for any given $\epsilon>0$, there exists a $\delta>0$, such that

$$\sum_{r=1}^{k} |x(t_r+h_r) - x(t_r)| < \epsilon$$

for all choices of intervals with

$$\sum_{r=1}^{k} h_r < \delta$$

It is well known that

(i) an absolutely continuous function is continuous.

(ii) an absolutely continuous function on [a,b] has a derivative almost everywhere on [a,b].

(iii) if a function f is integrable on [a,b] then the function $F(t) = \int_a^t f(x)dx$ is absolutely continuous on [a,b].

The space of all absolutely continuous functions defined on [a,b] and taking values in E^n will be denoted by $AC([a,b], E^n)$.

2.2 FUNCTIONAL-DIFFERENTIAL EQUATIONS

The general form of a functional-differential equation is

$$\dot{x}(t) = V(t, x(\cdot)), \quad \text{for } t > t_0$$

where V is a functional of Volterra type and $x(\cdot)$ represents the function x on the interval $a \le s \le t$.

2.2.1 Examples of Functional-Differential Equations [4]

(i) Differential equations with a lagging argument.

Let $a_0(t), a_1(t), \ldots, a_m(t)$ be given continuous functions for $t > a$ such that

$$a \le a_i(t) \le t \quad \text{for } t \geq a \qquad (i = 0, 1, \ldots, m)$$

An equation of the form

$$\dot{x}(t) = V(t, x(a_0(t)), x(a_1(t)), \ldots, x(a_m(t))), \qquad t > t_0 \tag{2.2.1.1}$$

where $V(t, y_0, y_1, \ldots, y_m)$ is a given function of $t, y_0, y_1, \ldots, y_m$, is called a differential equation with a lagging argument. It is frequently written

$$a_i(t) = t - h_i(t) \qquad (0 \le h_i(t) \le t - a)$$

so that Eqn. (2.2.1.1) can be written in the form

$$\dot{x}(t) = V(t, x(t-h_0(t)), x(t-h_1(t)), \ldots, x(t-h_m(t))) \tag{2.2.1.2}$$

(ii) Difference-differential equations with retarded argument.

If in Eqn. (2.2.1.2) the time lags $h_0(t), h_1(t), \ldots, h_m(t)$, reduce to constants $h_0 = 0, h_1, h_2, h_3, \ldots, h_m$ (>0), respectively, we obtain

$$\dot{x}(t) = V(t, x(t), x(t-h_1), \ldots, x(t-h_m)) \tag{2.2.1.3}$$

which is called a difference-differential equation with retarded argument; the constants $h_0, h_1, \ldots, h_m$ are called spans or retardations.

(iii) Integro-differential equations.

An equation of the form

$$\dot{x}(t) = V\left(t, x(t), \int_a^t F(t, s, x(s))\, ds\right) \tag{2.2.1.4}$$

where V and F are functions of all their arguments , is called an integro-differential equation .

(iv) Integro-difference-differential equations .

An equation of the form

$$(2.2.1.5) \qquad \dot{x}(t) = V\left(t,\ x(t),\ x(t-h_1), \ldots,\ x(t-h_m),\ \int_a^t F(t,\ s,\ x(s))ds \right)$$

where V and F are known functions in all their arguments, is called an integro-difference-differential equation . Equations of this type contain difference-differential and integro-differential equations as special cases .

(v) Differential equations with retarded argument .

An equation of the form

$$(2.2.1.6) \qquad \dot{x}(t) = \int_0^{H(t)} d_s R(s,t)\ x(t-s) \quad , \quad (H(t) \geq 0)$$

where $R(s,t)$ is a given $n \times n$ matrix valued function of bounded variation in s for each $t \geq s \geq a$, and the integral is in Stieltjes sense , is called a differential equation with retarded argument . The subscript s denotes that the integration is performed w.r.t. s at a fixed t .

(vi) Equations of Krasovskii - Razumikhin type .

An equation of the fcrm

$$(2.2.1.7) \qquad \dot{x}(t) = V(t,\ x(s);\ \ t-h \leq s \leq t), \text{ for } t > t_0$$

where $h > 0$ is a given constant and V a given functional of Volterra type, is called a Krasovskii - Razumikhin equation .

(vii) Equations of the form :

$$\dot{x}(t) = V(t, x(t - h_1(t, x(t))), \ldots, x(t - h_m(t, x(t)))), \quad t \geq t_0 \tag{2.2.1.8}$$

where

$$0 \leq h_i(t, x(t)) \leq t - \alpha \qquad \text{for } t > t_0, \quad i = 1, \ldots, m .$$

2.3 OPTIMIZATION PROBLEMS FOR HEREDITARY SYSTEMS

2.3.1 Dynamic Characterization of a Control System with Heredity .

Consider the system described by the following set of functional-differential equations ,

$$\dot{x}^i(t) = V_i(t, x(s), u(s) ; \alpha \leq s \leq t), \quad (i = 1, 2, \ldots, n) \tag{2.3.1.1}$$

or, for brevity ,

$$\dot{x}(t) = V(t, x(\cdot), u(\cdot)), \quad \text{a.e. (almost everywhere)} \tag{2.3.1.2}$$

on $I = [t_0, t_1] \subset [t_0, \infty)$, with the initial data

$$x(t_0) = x_0, \quad x(t) = \phi(t) \quad \text{for } t \in [\alpha, t_0] \subset (-\infty, t_0), \tag{2.3.1.3}$$

where

i) $x(t) \in E^n$ for $t \in I$ (State vector)

ii) $u(t) \in E^r$ for $t \in [\alpha, t_1]$ (Control vector)

iii) V is a vector valued Volterra functional of $x(\cdot)$, $u(\cdot)$ and an ordinary vector valued function of t .

Definition 2.3.1.1

The part of u corresponding to the initial interval $[\alpha, t_0)$, will be called initial control .

A vector valued function $x(t)$, $t \in [a, t_1]$, is called a solution of (2.3.1.2), corresponding to a given set of initial data (2.3.1.3) and control u, if it is absolutely continuous on I, satisfies (2.3.1.2) a.e. on I and coincides with $\phi(t)$ for $t \in [a, t_0)$. The solution is a continuation of the initial state function in the forward direction along the t axis.

Because of physical limitations of any control process, control vectors and initial functions cannot assume arbitrary values; i.e., they are, in general, subjected to certain constraints.

Let Φ be the set of all admissible initial state functions ϕ and Δ be the set of all admissible control vectors. Let, also, $G \subset E^n$ and $U \subset E^r$ be the ranges of the admissible initial state functions and control functions, respectively, i.e., let

(2.3.1.4)
$$\phi(t) \in G, \quad t \in [a, t_0), \quad \text{for all} \quad \phi \in \Phi$$
$$u(t) \in U, \quad t \in [a, t_1], \quad \text{for all} \quad u \in \Delta$$

The choice of the sets Φ, G, Δ, U describes specific features of the control process given by (2.3.1.2).

The set U will be called the control restraint set and the set G will be called the initial state function restraint set. A pair $\{u, \phi\} \in \Delta x \Phi$ will be called an admissible pair or admissible policy. The set

(2.3.1.5)
$$P = \Delta x \Phi$$

will be called the policy space of the system (2.3.1.2).

By the nature of a control process we are free to choose an admissible pair from the policy space P. From this freedom of selection of an admissible pair $\{u, \phi\}$, there naturally arises the problem of finding the best, i.e., the optimal policy, in some sense.

In modern control systems design one usually introduces as criteria of optimality some functionals (which generally depend on time, state vector, control vector and initial data) called performance indices or performance criteria or cost functionals , etc . An admissible pair $\{u^*, \phi^*\}$ is said to be optimal, if it minimizes (or maximizes) the cost functionals subject to certain physical constraints .

2. 3. 2 Existence and Uniqueness of Solutions

Definition 2. 3. 2. 1

The functional $V(t, x(\cdot), u(\cdot))$ is called continuous in t , if $V(t, x(\cdot), u(\cdot))$ is a continuous function of t , for $t \in I$, the remaining arguments being fixed .

Definition 2. 3. 2. 2

The functional $V(t, x(\cdot), u(\cdot))$ is said to satisfy a (global) Lipschitz condition with respect to $x(\cdot)$, if there exists a function $k_1 \in L_1(I)$ such that

$$\left| V(t, x(\cdot), u(\cdot)) - V(t, y(\cdot), u(\cdot)) \right| \leq |k_1(t)| \, \|x - y\|_t$$

a. e. on I and for all functions $u \in \Delta$, where $|V| = \sum_{i=1}^{n} |V_i|$ and $\|x\|_t = \sup_{t_0 \leq s \leq t} |x(s)|$

Definition 2. 3. 2. 3

The functional $V(t, x(\cdot), u(\cdot))$ is said to be continuous in t, x and u , if for any $\epsilon > 0$ there exists a $\delta > 0$, such that the relations

$$|t - s| < \delta , \qquad t, s \in [t_0, \gamma), \quad \gamma \in [t_0, t_1]$$
$$\|x - y\|_\gamma < \delta , \qquad x, y \in AC(I, E^n)$$
$$\|u - v\|_\gamma < \delta, \qquad u, v \in \Delta$$

imply

$$\left| V(t, x(\cdot), u(\cdot)) - V(s, y(\cdot), v(\cdot)) \right| < \epsilon$$

<u>THEOREM 2.3.2.1</u> (Existence and Uniqueness of solutions)

Let the functional $V(t, x(\cdot), u(\cdot))$ be continuous on $I \times AC \times \Delta$, and suppose that there exists a function $k_1 \in L_1(I)$ such that

$$\left| V(t, x(\cdot), u(\cdot)) - V(t, y(\cdot), u(\cdot)) \right| \leq |k_1(t)| \, \|x - y\|_t$$

a.e. on I and for all functions $u \in \Delta$.* Let $\phi \in \Phi \subset C([\alpha, t_0), G)$; $x(t_0) = x_0 \in F$, where G, F are compact subsets of E^n , and $u \in \Delta$. Then, there exists a unique solution of Eqn. (2.3.1.2) corresponding to the given ϕ , x_0, and u .

<u>Proof</u> :

The proof follows the same lines as in [4 , pp. 24-30]

2.3.3 <u>Performance Criteria</u>

Consider an admissible pair $\{u, \phi\} \in P$. Let x(t) be the trajectory of system (2.3.1.2) corresponding to the above pair and the initial condition $x(t_0) = x_0$. To measure the performance of the policy $\{u, \phi\}$ we associate with the system (2.3.1.2) a functional of the form

$$J[u, \phi] = \int_{t_0}^{t_1} k^0(s, t_1) \, f^0(x(s), u(s), s) \, ds \qquad (2.3.3.1)$$

where k^0 and f^0 are scalar valued functions, which are well defined and sufficiently smooth with respect to all their arguments .

The functional (2.3.3.1) is the <u>performance criterion</u> for the system (2.3.1.2), which we shall call the <u>cost</u> of the admissible policy $\{u, \phi\}$.

<u>Examples of performance criteria</u> :

If $k^0 \equiv 1$, $f^0 \equiv 1$, then (2.3.3.1) reduces to

$$J[u, \phi] = t_1 - t_0 \qquad (2.3.3.2)$$

* Δ is defined in section 2.3.4

and we have the time-Optimal problem .

If $k^0 \equiv 1$, and $f^0(x(t), u(t), t) = x^T(t)W(t)x(t) + u^T(t)U(t)u(t)$ where W, U are symmetric, positive (semi -)definite matrices, then (2.3.3.1) reduces to the usual quadratic cost criterion

(2.3.3.3) $$J[u, \phi] = \int_{t_0}^{t_1} [\, x^T(s)W(s)x(s) + u^T(s)U(s)u(s)\,]ds$$

2.3.4 Formulation of the Optimization Problem.

Consider system (2.3.1.2) with the initial data (2.3.1.3). Let $x(t)$, $t \in I$, be the trajectory corresponding tothe admissible pair $\{u, \phi\} \in P$ and the initial state $x(t_0) = x_0$. Assume that the performance criterion (2.3.3.1) is considered and that

i) G, the initial state function restraint set, is a compact subset of E^n
ii) U, the control restraint set, is a compact subset of E^r
iii) Φ, the set of admissible initial state functions, is the class $C([\alpha, t_0), G)$
iv) Δ, the set of admissible control functions, is the class of all bounded, measurable functions defined on $[\alpha, t_1]$ and takingvalues in U .

Let also S_0 and S_1 be two smooth manifolds in E^n.

Then, we can formulate the optimization problem as follows : Find an admissible policy $\{u, \phi\} \in \Delta \times \Phi$, so that the control u transfers the initial state $x(t_0) = x_0 \in S_0$ of system (2.3.1.2) to a terminal state $x(t_1) = x_1 \in S_1$ with minimum cost.

Thus, $\{u^*, \phi^*\}$ is an optimal policy, if

(i) $\{u^*, \phi^*\}$ is admissible, i.e., $u^* \in \Delta$, $\phi^* \in \Phi$
(ii) u^* transfers ϕ^* and $x_0 \in S_0$ to $x_1 \in S_1$
(iii) $J[u^*, \phi^*] \leq J[u, \phi]$ for all $\{u, \phi\} \in P$ which transfer $x_0 \in S_0$ to $x_1 \in S_1$

2.3.5 Questions related to the Optimization Problem

The optimization problem formulated in 2.3.4 gives rise to the following questions :

(i) Does there exist an admissible pair $\{u, \phi\}$ such that the control u transfers ϕ and $x_0 \in S_0$ to the target $x_1 \in S_1$?

(ii) Assuming that at least one such pair exists, does there exist an optimal pair $\{u^*, \phi^*\}$?

(iii) If there exists an optimal pair, what conditions must it satisfy ?

(iv) If there exists an optimal pair , is it unique ?

(v) How can we synthesize an optimal pair ?

Question (i) (controllability) is very important for the whole problem, but equally difficult to answer . We shall assume that there is a positive answer to that question,throughout the thesis, without further mention . The existence of an optimal pair (question (ii)) is a problem which we will also not consider in this thesis . We shall assume that there exists an optimal pair . We shall concentrate mainly in answering question (iii), i.e., in finding necessary conditions that an optimal pair $\{u^*, \phi^*\}$ must satisfy . These necessary conditions give actually all the pairs that are 'candidates' for optimality , thus,becoming a very powerful tool for the solution of the optimization problem . Question (iv) (uniqueness of the optimal pair) will be answered partially . No attempt will be made in this thesis to answer question (v) (synthesis problem) which is most difficult even for simple systems without heredity .

Before proceeding to the derivation of the necessary conditions for optimality of the hereditary system , we shall briefly present the celebrated Maximum Principle of L. S. Pontryagin :

2.4 PONTRYAGIN'S MAXIMUM PRINCIPLE

Consider the system described by the following vector differential equation

(2.4.1) $$\dot{x}(t) = f\,(\,x(t),\ u(t)\,)$$

where

i) $x(t) \in E^n$, $t \in I = [t_0, t_1]$, is the system's state

ii) $u(t) \in U \subset E^r$, $t \in I$, is the control vector

iii) $f(x, u)$ is an n- vector valued function with components $f^1(x, u)$, $f^2(x, u), \ldots, f^n(x, u)$. The functions f^i, $i=1, 2, \ldots, n$, are defined and are assumed to be continuous in $E^n \times U$. Moreover, they are assumed to be continuously differentiable w.r.t. x .

In addition to the system (2.4.1) there are given :

1) the smooth manifolds , S_0 and S_1 , in E^n

2) a real valued cost functional

$$J\,[u] = \int_{t_0}^{t_1} f^0(\,x(s),\ u(s)\,)ds$$

where f^0 is defined and is continuous , together with f^0_x , on all of $E^n \times U$.

3) a compact control restraint set U in E^r .

4) a class of admissible control functions :

$\Delta = \{\,u$: u is a bounded,measurable function defined on I and taking values in $U\,\}$

The <u>optimization problem</u> can be stated now as follows : Among all the admissible controls $u \in \Delta$, which tranfer the state of system (2.4.1) from some point $x_o \in S_o$ to some point $x_1 \in S_1$,(if such controls exist) , find one for which the functional $J\,[u]$ takes on the least possible value .

Definition 2.4.1 The Augmented System

The system

$$\dot{\hat{x}}(t) = \hat{f}\,(\hat{x}(t),\ u(t)\,) \tag{2.4.2}$$

where $\hat{x} = [x^0,\ x\,]^T$, $x^0(t) \triangleq \int_{t_0}^{t} f^0(x(s),\ u(s)\,)ds$, $\hat{f} = [f^0, f]^T$,
is called the augmented system .

Definition 2.4.2 The Adjoint System

The system

$$\dot{\hat{\psi}}(t) = -\,\hat{f}_{\hat{x}}\,(\hat{x}(t), u(t))\ \hat{\psi}(t) \tag{2.4.3}$$

where $\hat{\psi}$ is an n+1 -vector and $\hat{f}_{\hat{x}}$ is the Jacobian of $\hat{f}$ (w.r.t. $\hat{x}$), is called the (augmented) adjoint system corresponding to system (2.4.2) .

Definition 2.4.3 The Hamiltonian Function

The function

$$\hat{H}\ :\ E^{n+1} \times U \times E^{n+1} \longrightarrow E^1$$

defined by

$$\hat{H}(\,\hat{x}, u, \hat{\psi}\,) = (\,\hat{\psi},\ \hat{f}\,) = \sum_{i=0}^{n} \psi_i\, f^i\,(x,\ u) \tag{2.4.4}$$

is the (augmented) Hamiltonian function corresponding to system (2.4.2).

With the above definitions we can now state Pontryagin's Maximum Principle [1] :

THEOREM 2.4.1

Let $u(t)$, $t \in I$, be an admissible control which transfers the state of system (2.4.1) from some position $x_0 \in S_0$ to some position $x_1 \in S_1$, and let $\hat{x}(t)$ be the corresponding augmented trajectory (starting at the point $\hat{x}_0 = [0, x_0]^T$) . In order that $u(t)$ and $\hat{x}(t)$ be optimal it is necessary that there exists a nontrivial , absolutely continuous, vector valued function $\hat{\psi}(t) = [\psi_0(t), \psi_1(t), \ldots, \psi_n(t)]^T$ corresponding to the functions $u(t)$ and $\hat{x}(t)$, such that :

(i) the function $\hat{H}(\hat{x}(t), v, \hat{\psi}(t))$ of the variable $v \in U$ attains its maximum at the point $v = u(t)$, almost everywhere on the interval $[t_0, t_1]$, i.e.,

$$\hat{H}(\hat{x}(t), u(t), \hat{\psi}(t)) = \max_{v \in U} \hat{H}(\hat{x}(t), v, \hat{\psi}(t)) \qquad \text{a.e. on } I$$

(ii) at the terminal time $t = t_1$, the relations

$$\psi_0(t_1) \leq 0$$

$$\hat{M}(\hat{x}(t_1), \hat{\psi}(t_1)) \triangleq \max_{v \in U} \hat{H}(\hat{x}(t), v, \hat{\psi}(t))\Big|_{t=t_1} = 0$$

are satisfied . Furthermore , it turns out that if $\hat{\psi}(t)$, $\hat{x}(t)$, $u(t)$ satisfy systems (2.4.3), (2.4.2) and condition (i) , the time functions $\psi_0(t)$, and $\hat{M}(\hat{x}(t), \hat{\psi}(t))$ are constant on I . Thus condition (ii) may be verified at any time $t \in [t_0, t_1]$, and not just at t_1 .

(iii) the following transversality conditions are satisfied at both endpoints :

$$\psi(t_0) \perp S_0 \text{ (at } x_0) \quad , \qquad \psi(t_1) \perp S_1 \text{ (at } x_1) \quad .$$

Remark 2.4.1

The requirement that the control restraint set U be compact assures the existence of the maximum of the Hamiltonian, for given x(t) and $\psi(t)$. This fact is based on the following fundamental Lemma :

Lemma 2.4.1

A continuous function defined on a compact set attains its maximum and minimum in this set.

CHAPTER 3

OPTIMAL CONTROL PROBLEMS FOR A CLASS OF HEREDITARY SYSTEMS WITH SPECIFIED INITIAL STATE FUNCTION

3.1 INTRODUCTION

In this Chapter we consider a class of hereditary systems, which belongs to the general class (2. 3.1. 2) . The policy space is taken to be

(3.1.1) $$P = \Delta \times \{\phi\}$$

where ϕ is a given initial state function . Necessary conditions for the optimal control and the optimal initial control are derived ,for different cases . In the linear time - optimal case, theorems concerning the form and the uniqueness of the optimal control and the optimal initial control are obtained . Examples demonstrate the effect of the optimal choice of the initial control .

3.2 STATEMENT OF THE PROBLEM

Let a controlled system be given, which is described by the functional - differential equation

(3.2.1) $$\dot{x}(t) = \int_{[\alpha,t]} d_s G(t-s)\, f(x(s), u(s), s)$$

for $t \in I = [t_0, t_1] \subset [t_0, \infty)$, with the initial data

$$x(t_0) = x_0 \quad , \quad x(t) = \phi(t) \quad \text{for} \quad t \in [\alpha, t_0) \subset (-\infty, t_0)$$

where

i) $x(t) \in E^n$ for $t \in I$

ii) $u(t) \in U \subset E^r$ for $t \in [a, t_1]$

iii) f, $f_t \in C(E^{n+1} \times U, E^n)$, $f_x \in C$

iv) $G(s)$ is an $n \times n$ matrix valued function of bounded variation on $[0, t-a]$, $t \in [a, t_1]$

v) the integration is considered in the Stieltjes sense, with respect to s, for fixed t.

We assume that

1) $\phi \in C([a, t_0), G)$, where G is a bounded subset of E^n.

2) U, the control restraint set, is a compact subset of E^r.

3) Δ, the set of all admissible control functions, is the class of all bounded, measurable functions, defined on $[a, t_1]$ and taking values in U.

4) $$G(t) = S(t) + K(t)$$

where $K(t)$ is a matrix valued function, absolutely continuous in t, and $S(t)$ is the " saltus " part of G, i.e., a matrix valued function, whose components $s_{ij}(t)$ are piecewise continuous functions with discontinuities, of the first kind, at the points (assumed to be finite the number) :

$$0 = h_0 < h_1 < \dots < h_m \qquad (a < h_0)$$

and let $A_i(t)$ be the matrices of the jumps of $S(t-s)$ at the points $s = t - h_i$, $(i = 0, 1, \dots, m)$.

In addition, two smooth manifolds [1] S_0 and S_1 in E^n, with dimensions r_0 and r_1 $(\leq n)$, respectively, are given.

From assumption 4) and the usual rules of integration for Stieltjes integrals [4, p. 79 ; 25, p. 168] we find :

$$\int_a^t d_s\, S(t-s)\, f(x(s),u(s),s) = \sum_{i=0}^{m} A_i(t)\, f(x(t-h_i),u(t-h_i),t-h_i)$$

for $t > h_m$, and

$$\int_a^t d_s\, K(t-s)\, f(x(s),u(s),s) = \int_a^t L(t-s)\, f(x(s),u(s),s)\, ds$$

where $L(t) = -\frac{dK(t)}{dt}$

Thus, system (3.2.1) reduces to the equivalent form (for $t > h_m$)

(3.2.2) $$\dot{x}(t) = \sum_{i=0}^{m} A_i(t)\, f(x(t-h_i),u(t-h_i),t-h_i) + \int_a^t L(t-s)\, f(x(s),u(s),s)ds$$

In the sequel we shall consider and study the controlled process (3.2.2) for $t \in I = [t_0, t_1]$.

The cost functional considered here is

(3.2.3) $$J[u] = \int_{t_0}^{t_1} k^0(t_1-s)\, f^0(x(s),u(s),s)\, ds$$

where $k^0 \in C^1(I)$ and $f^0, f^0_t \in C(E^{n+1} \times U, E^1)$, $f^0_x \in C(E^{n+1} \times U, E^n)$

We shall concentrate on the following <u>optimization problem</u> : Find necessary conditions for a control $u \in \Delta$ to be <u>optimal</u>, i.e., minimizing the cost functional (3.2.3) while transfering the state of system (3.2.2) from $x_0 \in S_0$ to $x_1 \in S_1$.

3.3 BASIC DEFINITIONS

3.3.1 The Variational System

Let $u(t)$, $t \in [a, t_1]$, be a control in Δ and let $x(t)$ be the solution of system (3.2.2) corresponding to this control, with initial condition $x(t_0) = x_0$ and initial state function $\phi(t)$.

Let $y(t)$ be the solution corresponding to the same control $u(t)$ and initial state function $\phi(t)$, and beginning, at the same time t_0, at the point :

$$y_0 = x_0 + \epsilon \xi_0 + O(\epsilon)$$

where ξ_0 is a constant vector in E^n and $\lim_{\epsilon \to 0} \frac{O(\epsilon)}{\epsilon} = 0$

The solution $y(t)$ has the form :

$$y(t) = x(t) + \epsilon \delta x(t) + O(\epsilon)$$

where $\delta x(t)$ is a vector valued function, independent of ϵ and satisfying the system ,

$$(3.3.1.1) \quad (\delta \dot{x})(t) = \sum_{i=0}^{m} A_i(t) F_x(t-h_i) \delta x(t-h_i) + \int_a^t L(t-s) F_x(s) \delta x(s)\, ds$$

for $t \in I$, with $\delta x(t_0) = \xi_0$ and $\delta x(t) = 0$ for $t \in [a, t_0)$, where $F_x(t) \triangleq f_x(x(t), u(t), t)$

System (3.3.1.1) is called the Variational System corresponding to system (3.2.2).

3.2.2 The Adjoint System

Consider the adjoint to (3.3.1.1) system :

$$(3.3.2.1) \quad \dot{\psi}(t) = - F_x^T(t) \left\{ \sum_{i=o}^{m} A_i^T (t+h_i) \psi(t+h_i) + \int_t^{t_1} L^T(s-t)\, \psi(s)\, ds \right\}$$

on I, with $\psi(t_1) = \psi_1$ and $\psi(t) = 0$ for $t > t_1$.

A vector $\psi(t)$ is a solution of (3.3.2.1) if it is absolutely continuous, on $[t_0, t_1)$, and it satisfies (3.3.2.1) a.e. on I, with $\psi(t_1) = \psi_1$ and $\psi(t) = 0$ for $t > t_1$.

$\psi(t)$ is called the adjoint response and (3.3.2.1) the Adjoint System.

3.3.3 The Hamiltonian Functionals

We shall now introduce the following functionals :

$$H(x(t), \psi(\cdot), v, t) = \sum_{i=0}^{m} (A_i(t+h_i) f(x(t), v, t), \psi(t+h_i)) + \int_t^{t_1} (L(s-t) f(x(t), v, t), \psi(s)) \, ds, \quad t \in I$$

(3.3.3.1)

$$H_p(\phi(t), \psi(\cdot), v, t) = \sum_{i=0}^{m} \theta_i(t) (A_i(t+h_i) f(\phi(t), v, t), \psi(t+h_i)) + \int_{t_0}^{t_1} (L(s-t) f(\phi(t), v, t), \psi(s)) \, ds, \quad t \in [a, t_0)$$

where

$$\theta_i(t) = \begin{cases} 1 & \text{if } t \in [t_0 - h_i, t_1 - h_i) \\ 0 & \text{if } t \notin [t_0 - h_i, t_1 - h_i) \end{cases} \qquad (i=0,1,\dots,m)$$

We shall call H and H_p the Hamiltonian functionals. The functional character is due to the presence of $\psi(\cdot)$ in H and H_p.

We also introduce the functional

(3.3.2.2) $$R(t) \triangleq R(x(\cdot), \psi(t), u(\cdot), t) = \left(\psi(t), \sum_{i=0}^{m} A_i(t) f(x(t-h_i), u(t-h_i), t-h_i) + \int_a^t L(t-s) f(x(s), u(s), s) \, ds \right)$$

3.3.4 Augmented Variables and Systems

We increase the dimension of the state space by introducing a new state variable x^0 defined by

$$x^0(t) = \int_{t_0}^{t} k^o(t-s)\, f^0(x(s),u(s),s)\,ds \qquad (3.3.4.1)$$

We define now the following augmented quantities

$$\hat{x}(t) = \begin{bmatrix} x^0(t) \\ x(t) \end{bmatrix}, \quad \hat{\phi}(t) = \begin{bmatrix} 0 \\ \phi(t) \end{bmatrix}, \quad \hat{f} = \begin{bmatrix} f^0 \\ f \end{bmatrix}, \quad \hat{\psi}(t) = \begin{bmatrix} \psi_0(t) \\ \psi(t) \end{bmatrix},$$

$$\hat{A}_0(t) = \left[\begin{array}{c|c} k^0(0) & 0 \\ \hline 0 & A_0(t) \end{array}\right], \quad \hat{A}_i(t) = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & A_i(t) \end{array}\right], \ (i=1,2,\ldots,m)$$

$$\hat{L}(t-s) = \left[\begin{array}{c|c} \frac{dk^0(t-s)}{dt}\, e(s-t_0) & 0 \\ \hline 0 & L(t-s) \end{array}\right]$$

where $e(s-t_0)$ is the unit step function .

$\hat{x}(t)$ is now the response of the augmented system

$$\dot{\hat{x}}(t) = \sum_{i=0}^{m} \hat{A}_i(t)\, \hat{f}(\hat{x}(t-h_i),u(t-h_i),t-h_i) + \int_{a}^{t} \hat{L}(t-s)\, \hat{f}(\hat{x}(s),u(s),s)\,ds \qquad (3.3.4.2)$$

with $\hat{x}(t) = \phi(t)$, $t \in [a,t_0)$, and $\hat{x}(t_0) = \hat{x}_0 = [0,x_0]^T$

The augmented variational system is now defined as :

$$(\dot{\widehat{\delta x}})(t) = \sum_{i=0}^{m} \hat{A}_i(t)\, \hat{F}_{\hat{x}}(t-h_i)\widehat{\delta x}(t-h_i) + \int_{a}^{t} \hat{L}(t-s)\, \hat{F}_{\hat{x}}(s)\widehat{\delta x}(s)\,ds \qquad (3.3.4.3)$$

with $\hat{\delta x}(t_0) = [0, \xi_0]^T$, $\hat{\delta x}(t) = 0$ for $t \in [a, t_0)$

The augmented adjoint response $\hat{\psi}(t)$ is the solution of the augmented adjoint system :

(3.3.4.4) $$\dot{\hat{\psi}}(t) = -\hat{F}_{\hat{x}}^T(t) \{ \sum_{i=0}^{m} \hat{A}_i^T(t+h_i)\hat{\psi}(t+h_i) + \int_t^{t_1} \hat{L}^T(s-t)\hat{\psi}(s)\, ds \}$$

on I , with $\hat{\psi}(t_1) = \hat{\psi}_1 = [\psi_0(t_1), \psi_1]^T$ and $\hat{\psi}(t) = 0$ for $t > t_1$.

The augmented functionals $\hat{H}$, $\hat{H}_p$, and $\hat{R}$ are now defined as

(3.3.4.5) $$\hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t) = \sum_{i=0}^{m} (\hat{A}_i(t+h_i)\, \hat{f}(\hat{x}(t), v, t) \;,\; \hat{\psi}(t+h_i)) + \int_t^{t_1} (\hat{L}(s-t)\, \hat{f}(\hat{x}(t), v, t) \;,\; \hat{\psi}(s))ds \;,\; t \in I$$

$$\hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t) = \sum_{i=0}^{m} \theta_i(t)\, (\hat{A}_i(t+h_i)\hat{f}(\hat{\phi}(t), v, t) \;,\; \hat{\psi}(t+h_i)) + \int_{t_0}^{t_1} (\hat{L}(s-t)\, \hat{f}(\hat{\phi}(t), v, t) \;,\; \hat{\psi}(s))ds \;,\; t \in [a, t_0)$$

(3.3.4.6) $$\hat{R}(t) = (\hat{\psi}(t) \;,\; \sum_{i=0}^{m} \hat{A}_i(t)\hat{f}(\hat{x}(t-h_i), u(t-h_i), t-h_i) + \int_a^t \hat{L}(t-s)\, \hat{f}(\hat{x}(s), u(s), s)ds) \;,\; t \in I \;.$$

3.4 NECESSARY CONDITIONS FOR OPTIMALITY

We consider first the case where the function f does not depend explicitly on the time t .

The following Theorem gives necessary conditions for an admissible control to be optimal .

THEOREM 3.4.1

Consider the control process in E^n :

$$\dot{x}(t) = \sum_{i=0}^{m} A_i(t) f(x(t-h_i), u(t-h_i)) + \int_a^t L(t-s) f(x(s), u(s)) ds \tag{3.4.1}$$

with bounded , measurable controls $u(t)$, $t \in [a, t_1]$, in the compact restraint set $U \subset E^r$. Let Δ be the class of all admissible controls.

For each $u \in \Delta$ with response $x(t)$, let the cost functional be

$$J[u] = \int_{t_0}^{t_1} k^0(t_1 - s) f^0(x(s), u(s)) ds \tag{3.4.2}$$

If u is an optimal control in Δ with augmented response $\hat{x}(t)$, $t \in I$, then there exists a nontrivial adjoint response ψ (and augmented response $\hat{\psi}(t)$) , absolutely continuous on $[t_0, t_1)$, such that :

(i)

$$\hat{H}(\hat{x}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } [t_0, t_1]$$

$$\hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } [a, t_0)$$

(ii) at the terminal time $t = t_1$:

$$\psi_0(t_1) \le 0 \qquad \text{and} \qquad \hat{R}(t_1) = 0$$

Moreover, $\psi_0(t) = \text{const.} \le 0$ for all $t \in I$

(iii) the following transversality conditions are satisfied at both endpoints of the trajectory $x(t)$,

$$\psi(t_0) \perp S_o \quad (\text{at } x_0) , \qquad \psi(t_1) \perp S_1 \quad (\text{at } x_1)$$

<u>Proof</u> :

The proof follows the steps of Pontryagin's [1] approach, as utilized by Friedman [2], modified adequately with the use of results of Oǧuztöreli [4]. We present this proof, which is fundamental for this thesis, in Appendix A and not here.

Now we can consider the more general case where f is also a function of time. However, the optimal problem can be reduced to the previous one, as shown below :

<u>COROLLARY 3.4.1</u>

Consider the control process in E^n :

$$(3.4.3) \quad \dot{x}(t) = \sum_{i=0}^{m} A_i(t) f(x(t-h_i), u(t-h_i), t-h_i) + \int_a^t L(t-s) f(x(s), u(s), s)\, ds$$

on $I = [t_0, t_1]$, with $u \in \Delta$, $u(t) \in U$ for $t \in [a, t_1]$.

For each $u \in \Delta$ with response $x(t)$, let the cost functional be :

$$(3.4.4) \qquad J[u] = \int_{t_0}^{t_1} k^0(t_1 - s)\, f^0(x(s), u(s), s)\, ds$$

If u is an optimal control in Δ with augmented response $\hat{x}(t)$, $t \in I$, then there exists a nontrivial adjoint response $\psi(t)$ (and augmented response $\hat{\psi}(t)$), absolutely continuous on $[t_0, t_1)$, such that :

(i)

$$\hat{H}(\hat{x}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } [t_0, t_1]$$

$$\hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } [a, t_0)$$

(ii) at the terminal time $t = t_1$:

$$\psi_0(t_1) \leq 0 \qquad \text{and} \qquad \hat{R}(t_1) = 0$$

Moreover , $\psi_0(t) = \text{const} \leq 0$ for all $t \in I$

(iii) the following transversality conditions are satisfied at both endpoints :

$$\psi(t_0) \perp S_0 \ (\text{at } x_0) \quad , \quad \psi(t_1) \perp S_1 \ (\text{at } x_1)$$

Proof :

We define a new state variable , x^{n+1} , given by :

(3.4.5) $x^{n+1} = t$, for $t \in [a, t_1]$

Consider now the system in E^{n+1}

(3.4.6)
$$\dot{x}(t) = \sum_{i=0}^{m} A_i(t) f(x(t-h_i), u(t-h_i), x^{n+1}(t-h_i)) + \int_a^t L(t-s) f(x(s), u(s), x^{n+1}(s)) ds$$
$$\dot{x}^{n+1}(t) = 1$$

for $t \in I$, with $x(t) = \phi(t)$, $t \in [a, t_0)$, $x(t_0) = x_0$, $x^{n+1}(t) = t$ for $t \in [a, t_0]$ and the cost functional

(3.4.7)
$$J[u] = \int_{t_0}^{t_1} k^0(t_1 - s) f^0(x(s), u(s), x^{n+1}(s)) ds$$

Now , the initial and target manifolds for system (3.4.6) in E^{n+1}, are the sets $S_0 \times \{t_0\}$ and $S_1 \times R$.

The system (3.4.6) with state $[x, x^{n+1}]^T$ and the cost functional (3.4.7) are exactly of the forms (3.4.1) and (3.4.2), respectively ; thus , we can apply Theorem 3.4.1 .

First , let us introduce the following time augmented quantities :

$$\hat{\hat{x}} = \begin{bmatrix} \hat{x} \\ x^{n+1} \end{bmatrix}, \quad \hat{\hat{\phi}}(t) = \begin{bmatrix} \hat{\phi}(t) \\ t \end{bmatrix}, \quad \hat{\hat{f}} = \begin{bmatrix} \hat{f} \\ 1 \end{bmatrix}, \quad \hat{\hat{\psi}} = \begin{bmatrix} \hat{\psi} \\ \psi_{n+1} \end{bmatrix}$$

$$\hat{\hat{A}}_0(t) = \left[\begin{array}{c|c} \hat{A}_0(t) & 0 \\ \hline 0 & 1 \end{array}\right], \quad \hat{\hat{A}}_i(t) = \left[\begin{array}{c|c} \hat{A}_i(t) & 0 \\ \hline 0 & 0 \end{array}\right], \qquad (i = 1, 2, \ldots, m)$$

$$\hat{\hat{L}} = \left[\begin{array}{c|c} \hat{L} & 0 \\ \hline 0 & 0 \end{array}\right]$$

The time-augmented system is defined now as :

$$(3.4.8) \quad \dot{\hat{\hat{x}}}(t) = \sum_{i=0}^{m} \hat{\hat{A}}_i(t)\hat{\hat{f}}(\hat{\hat{x}}(t-h_i), u(t-h_i)) + \int_a^t \hat{\hat{L}}(t-s)\hat{\hat{f}}(\hat{\hat{x}}(s), u(s))ds$$

with $\hat{\hat{x}}(t_0) = [0, x_0, t_0]^T$ and $\hat{\hat{x}}(t) = \hat{\hat{\phi}}(t)$, $t \in [a, t_0)$,

and the time-augmented adjoint system , based on $u(t)$ and $\hat{\hat{x}}(t)$, is :

$$(3.4.9) \quad \dot{\hat{\hat{\psi}}}(t) = -\hat{\hat{F}}^T_{\hat{\hat{x}}}(t)\{ \sum_{i=0}^{m} \hat{\hat{A}}^T_i(t+h_i)\hat{\hat{\psi}}(t+h_i) + \int_t^{t_1} \hat{\hat{L}}^T(s-t)\hat{\hat{\psi}}(s)ds \}$$

or ,

$$(3.4.10) \quad \begin{aligned} &\dot{\psi}_0(t) = 0 \\ &\dot{\psi}(t) = -F^T_x(t)\{ \sum_{i=0}^{m} A^T_i(t+h_i)\psi(t+h_i) + \int_t^{t_1} L^T(s-t)\psi(s)ds \} \\ &\dot{\psi}_{n+1}(t) = -\hat{F}^T_t(t)\{ \sum_{i=0}^{m} \hat{A}^T_i(t+h_i)\hat{\psi}(t+h_i) + \int_t^{t_1} \hat{L}^T(s-t)\hat{\psi}(s)\, ds \} \end{aligned}$$

where $\hat{F}_t(t) \triangleq \hat{f}_t(x(t), u(t), t)$.

The time-augmented functionals are :

$$\hat{\hat{H}}(\hat{\hat{x}}(t), \hat{\hat{\psi}}(\cdot), v, t) = \sum_{i=0}^{m} \psi_{n+1}(t+h_i) + \hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t)$$

(3.4.11) $$\hat{\hat{H}}_p(\hat{\hat{\phi}}(t), \hat{\hat{\psi}}(\cdot), v, t) = \sum_{i=0}^{m} \theta_i(t)\psi_{n+1}(t+h_i) + \hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t)$$

$$\hat{\hat{R}}(\hat{\hat{x}}(\cdot), \hat{\hat{\psi}}(t), u(\cdot), t) = \psi_{n+1}(t) + \hat{R}(\hat{x}(\cdot), \hat{\psi}(t), u(\cdot), t)$$

Applying now Theorem 3.4.1 to system (3.4.6) we obtain :

(3.4.12)
$$\hat{\hat{H}}(\hat{\hat{x}}(t), \hat{\hat{\psi}}(\cdot), u(t), t) = \max_{v \in U} \hat{\hat{H}}(\hat{\hat{x}}(t), \hat{\hat{\psi}}(\cdot), v, t) \quad \text{a.e. on } I$$
$$\hat{\hat{H}}_p(\hat{\hat{\phi}}(t), \hat{\hat{\psi}}(\cdot), u(t), t) = \max_{v \in U} \hat{\hat{H}}_p(\hat{\hat{\phi}}(t), \hat{\hat{\psi}}(\cdot), v, t) \quad \text{a.e. on } [\alpha, t_0)$$

From (3.4.11), (3.4.12) and the fact that ψ_{n+1} is not a function of $v \in U$, we obtain

(3.4.13) $$\hat{H}(\hat{x}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } I$$

(3.4.14) $$\hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } [\alpha, t_0)$$

Let us now consider a fixed endpoint problem for system (3.4.3), i.e., $x(t_0) = x_0$ and $x(t_1) = x_1$. Then, the optimal problem for system (3.4.6) is a free-right-endpoint problem with target manifold $\{x_1\} \times R$. Condition (iii) of Theorem 3.4.1 asserts that

(3.4.15) $$[\psi(t_1), \psi_{n+1}(t_1)]^T \perp \{x_1\} \times R$$

This condition implies that :

(3.4.16) $$\psi_{n+1}(t_1) = 0$$

From (3.4.11) and (3.4.16) we obtain

$$(3.4.17) \qquad \hat{\hat{R}}(t_1) = \psi_{n+1}(t_1) + \hat{\hat{R}}(t_1) = \hat{R}(t_1)$$

From condition (ii) of Theorem 3.4.1, as applied to system (3.4.6), and (3.4.17) we finally obtain

$$(3.4.18) \qquad \hat{R}(t_1) = \hat{\hat{R}}(t_1) = 0$$

The proof of the condition $\psi_0(t) = \text{const} \leq 0$, for $t \in I$, follows from Theorem 3.4.1 .

Now, let us suppose that the initial and terminal points of system (3.4.3) belong to the smooth manifolds S_0 and S_1 . Then, the corresponding points of system (3.4.6) belong to the sets $S_0 \times \{t_0\}$ and $S_1 \times R$. Applying the transversality conditions of Theorem 3.4.1, we obtain

$$(3.4.19) \qquad \begin{aligned} &[\psi(t_0), \psi_{n+1}(t_0)]^T \perp S_0 \times \{t_0\}, \quad (\text{at } [x_0, t_0]^T) \\ &[\psi(t_1), \psi_{n+1}(t_1)]^T \perp S_1 \times R \qquad (\text{at } [x_1, t_1]^T) \end{aligned}$$

from which it follows

$$(3.4.20) \qquad \psi(t_0) \perp S_0 \ (\text{at } x_0), \qquad \psi(t_1) \perp S_1 \ (\text{at } x_1)$$

This completes the proof of Corollary 3.4.1 .-

We consider now the optimal problem where the criterion of performance is the time of transition from x_0 to x_1 . A control $u \in \Delta$ will be called time-optimal if it steers the system from x_o to x_1 in the shortest possible time . The system considered here is (3.4.21) $\equiv$ (3.4.3) $\equiv$ (3.2.2) with cost functional $J[u] = t_1 - t_0$. Since this is a particular case of (3.4.4) , when $f^0 \equiv 1$ and $k^0 \equiv 1$, the time-optimal problem is reduced to the previous one, with the following result :

COROLLARY 3.4.2

Consider the control process in E^n :

$$\dot{x}(t) = \sum_{i=0}^{m} A_i(t) f(x(t-h_i), u(t-h_i), t-h_i) + \int_a^t L(t-s) f(x(s), u(s), s) ds \tag{3.4.21}$$

with $u \in \Delta$, $u(t) \in U$ for $t \in [a, t_1]$, and with cost functional

$$J[u] = t_1 - t_0 \tag{3.4.22}$$

If u is a time-optimal control in Δ , with response $x(t)$, $t \in I$, then there exists a nontrivial adjoint response $\psi(t)$ (and augmented response $\hat{\psi}(t)$) , absolutely continuous on $[t_0, t_1)$, such that :

(i)

$$H(x(t), \psi(\cdot), u(t), t) = \max_{v \in U} H(x(t), \psi(\cdot), v, t) \quad \text{a.e. on } I$$

$$H_p(\phi(t), \psi(\cdot), u(t), t) = \max_{v \in U} H_p(\phi(t), \psi(\cdot), v, t) \quad \text{a.e. on } [a, t_0)$$

(ii) at the terminal time $t = t_1$:

$$\psi_0(t_1) \leq 0 \qquad \text{and} \qquad R(t_1) \geq 0$$

Moreover , $\psi_0(t) = c. \leq 0$, for all $t \in I$.

(iii) the following transversality conditions are satisfied at both endpoints :

$$\psi(t_0) \perp S_0 \ (\text{at } x_0) \quad , \quad \psi(t_1) \perp S_1 \ (\text{at } x_1)$$

Proof :

We need only prove part (i) and the second condition of part (ii), since the rest is covered from Corollary 3.4.1, which holds for the present case, as well.

Now,

$$\hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t) = \sum_{i=0}^{m} \psi_0(t+h_i) + H(x(t), \psi(\cdot), v, t)$$

$$(3.4.23) \quad \hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t) = \sum_{i=0}^{m} \theta_i(t)\psi_0(t+h_i) + H_p(\phi(t), \psi(\cdot), v, t)$$

$$\hat{R}(\hat{x}(.), \hat{\psi}(t), u(.), t) = \psi_0(t) + R(x(\cdot), \psi(t), u(\cdot), t)$$

Applying now Corollary 3.4.1 and using (3.4.23) we obtain:

$$(3.4.24) \quad \begin{aligned} &H(x(t), \psi(\cdot), u(t), t) = \max_{v \in U} H(x(t), \psi(\cdot), v, t) \quad \text{a.e. on } I \\ &H_p(\phi(t), \psi(\cdot), u(t), t) = \max_{v \in U} H_p(\phi(t), \psi(\cdot), v, t) \quad \text{a.e. on } [a, t_0) \end{aligned}$$

and

$$(3.4.25) \quad \hat{R}(t_1) = \psi_0(t_1) + R(t_1) = 0$$

from which

$$(3.4.26) \quad R(t_1) = -\psi_0(t_1) \geq 0$$

Q.E.D.

Remark 3.4.1 : The results of Theorem 3.4.1, Corollary 3.4.1 and Corollary 3.4.2 remain valid for hereditary processes of the form :

$$(3.4.27) \quad \dot{x}(t) = \sum_{i=0}^{m} A_i(t) f_{(i)}(x(t-h_i), u(t-h_i), t-h_i) + \int_a^t L(t-s) f_{(m+1)}(x(s), u(s), s) ds$$

where $f_{(i)}$, ($i=0,1,\ldots,m+1$), are vector valued functions, not necessarily identical, and the following changes are made to the adjoint system and the functionals H, H_p and R :

$$(3.4.28) \quad \dot{\psi}(t) = -\sum_{i=0}^{m} F^T_{x,(i)}(t) A_i^T(t+h_i)\psi(t+h_i) - F^T_{x,(m+1)}(t)\int_t^{t_1} L^T(s-t)\psi(s)ds$$

$$(3.4.29) \quad H(x(t), \psi(\cdot), v, t) = \sum_{i=0}^{m} (A_i(t+h_i) f_{(i)}(x(t), v, t) \ , \ \psi(t+h_i)) + \int_t^{t_1} (L(s-t) f_{(m+1)}(x(t), v, t) \ , \ \psi(s)) ds$$

$$(3.4.30) \quad H_p(\phi(t), \psi(\cdot), v, t) = \sum_{i=0}^{m} \theta_i(t) (A_i(t+h_i) f_{(i)}(\phi(t), v, t), \psi(t+h_i)) + \int_{t_0}^{t_1} (L(s-t) f_{(m+1)}(\phi(t), v, t) \ , \ \psi(s)) ds$$

$$(3.4.31) \quad R(t) = (\ \psi(t) \ , \ \sum_{i=0}^{m} A_i(t) f_{(i)}(x(t-h_i), u(t-h_i), t-h_i) + \int_a^t L(t-s) f_{(m+1)}(x(s), u(s), s) ds \ )$$

<u>Remark 3.4.2</u> : <u>Moving Target Set</u>

Let S_0 be a fixed manifold in E^n but $S_1(t)$ a moving r_1-dimensional manifold, whose dependence on t is differentiable . If u is a control in Δ which transfers the state of system (3.2.2) $\equiv$(3.4.3) from some point $x_0 \in S_0$ to the target set $S_1(t_1)$, at time t_1, and which minimizes the cost functional (3.4.4) , then all statements of

Corollary 3.4.1 remain valid for the present case, except the second condition of part (ii), which is modified .

Indeed, let us consider a parametric curve $(z^1(\theta), z^2(\theta), \ldots, z^n(\theta))$, lying entirely on the smooth manifold $S_1(t_1)$ and passing through the the point $x(t_1) \in S_1(t_1)$. Then, the curve $(z^1(\theta), z^2(\theta), \ldots, z^n(\theta), \theta)$ lies entirely on $S_1(t_1) \times R$ and passes through the point $(x(t_1), t_1)^T$. The tangent vector $(q^1, q^2, \ldots, q^n, 1)$, ($q^i = \frac{dz^i}{d\theta}$), of this curve at $(x(t_1), t_1)^T$ is also a tangent vector of the manifold $S_1(t_1) \times R$ at the point $(x(t_1), t_1)^T$ [15]. From part (iii) Of Corollary 3.4.1, we obtain

$$(3.4.32) \qquad \sum_{i=1}^{n} \psi_i(t_1)\, q^i + \psi_{n+1}(t_1) \cdot 1 = 0$$

and

$$(3.4.33) \qquad \hat{\hat{R}}(t_1) = \psi_{n+1}(t_1) + \hat{R}(t_1) = 0$$

From (3.4.32) and (3.4.33) we conclude that ,if $(q^1, q^2, \ldots, q^n)$ is a vector tangent at $S_1(t_1)$ at the point $x(t_1)$, then

$$(3.4.34) \qquad \hat{R}(t_1) = \sum_{i=1}^{n} \psi_i(t_1)\, q^i$$

<u>Remark 3.4.3</u> : <u>A Tracking Problem</u>

Let $z = z(t)$ be a moving particle in E^n. The optimal problem now is to find a control u in Δ such that the corresponding trajectory of system (3.2.2) ≡ (3.4.3) intersects the curve $z = z(t)$ at some time $t = t_1$, while the cost functional (3.4.4) achieves its minimum.

Since $z(t)$ can be considered as a one dimensional manifold in E^n, the results in Remark 3.4.2 applyhere, as well. Thus, if u is an optimal control in Δ, such that $x(t_1) = z(t_1)$, and if $(q^1, q^2, \ldots, q^n)$ is a tangent vector of the curve $z = z(t)$, at $t = t_1$, then

$$\hat{R}(t_1) = \sum_{i=1}^{n} \psi_i(t_1)\, q^i \tag{3.4.35}$$

with the rest of the assertions of Corollary 3.4.1 remaining unchanged. (when applicable to the tracking problem.)

Remark 3.4.4 : Free-endpoint Problem

If the target set S_1 is all of E^n, then the transversality condition of Corollary 3.4.1, at the right-hand endpoint, implies that

$$\psi(t_1) = 0 \tag{3.4.36}$$

All other statements of Corollary 3.4.1 remain unchanged.

In all previous problems, we considered fixed initial time t_0, but free terminal time t_1. If we consider now the case where t_1 is also fixed, we can state the following

COROLLARY 3.4.3

Consider the control process in E^n

$$\dot{x}(t) = \sum_{i=0}^{m} A_i(t) f(x(t-h_i), u(t-h_i), t-h_i) + \int_{a}^{t} L(t-s) f(x(s), u(s), s)\, ds \tag{3.4.37}$$

on $I = [t_0, t_1]$, where t_0 and t_1 are fixed, with $u \in \Delta$, $u(t) \in U$ for $t \in [a, t_1]$, and with cost functional

$$J[u] = \int_{t_0}^{t_1} k^0(t_1 - s)\, f^0(x(s), u(s), s)\, ds \tag{3.4.38}$$

If u is an optimal control in Δ, with augmented response $\hat{x}(t)$, $t \in I$, then there exists a nontrivial adjoint response $\psi(t)$ (and augmented response $\hat{\psi}(t)$), absolutely continuous on $[t_0, t_1)$, such that :

(i)

$$\hat{H}(\hat{x}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } I$$

$$\hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } [a, t_0)$$

(ii) $\psi_0(t) = \text{const} \leq 0$ for all $t \in I$

(iii) the following transversality conditions are satisfied

$$\psi(t_0) \perp S_0 \ (\text{at } x_0) \ , \qquad \psi(t_1) \perp S_1 \ (\text{at } x_1)$$

Proof :

The proof is exactly the same as in [1, pp. 66-69] .

We observe that Corollary 3.4.3 differs from Corollary 3.4.1 only in the statement concerning $R(t_1)$, which in 3.4.3 is missing . This is due to the fact that, in the fixed time problem, time perturbations of the form $t_1 + \epsilon \delta t$ (s. Appendix A) are not allowed .

Remark 3.4.5 :

In all previous problems , we considered that $u(t) \in U \subset E^r$, for all $t \in [a, t_1]$. The results obtained will not change , if the initial control $u(t) \in U_1$, $t \in [a, t_0)$, and the control $u(t) \in U_2$, $t \in [t_0, t_1]$, where U_1 and U_2 are compact subsets of E^r, not necessarily identical .

From the proof of Theorem 3.4.1 (s. Appendix A) we also observe that the Theorem and its Corollaries remain valid , even if U_1 or U_2 or both , are all of E^r . (Wherever the R.H.S. of conditions (i) exist) .

The following example considers a control system with distinct control restraint sets U_1 and U_2 and fixed time .

EXAMPLE 3.4.1 : An energy-optimal problem

Consider the system in $E^1 = R$.

$$\dot{x}(t) = -x(t-1) + u(t) + u(t-2) \quad , \quad t \in [0, 2] \tag{3.4.39}$$

with $x(t) = \phi(t) = 1, \quad t \in [-1, 0) \quad , \quad x(0) = 2$

and the cost functional

$$J[u] = \int_0^2 u^2(t)dt \tag{3.4.40}$$

with $u \in \Delta$, $u(t) \in U_1 = [-1,1], t \in [-2,0)$, and $u(t) \in U_2 = R$, for $t \in [0,2]$.

We shall consider the following optimization problem with fixed time : Find a control u in Δ, which transfers the state of system (3.4.39) from the initial point $x(0) = 2$ to the terminal point $x(2) = -2$ and minimizes the cost functional (3.4.40) .

The solution of system (3.4.39) is given by (s. Appendix B) :

$$x(t) = N(t)x(0) + \int_{-1}^{0} N(t-s-1)\phi(s)ds + \int_0^t N(t-s)\{u(s) + u(s-2)\}ds \tag{3.4.41}$$

where N(t) satisfies :

$$\begin{aligned} \dot{N}(t) &= -N(t-1) \quad , \quad t > 0 \\ N(t) &= 0 \; , \; t \in [-2,0) \quad , \quad N(0) = 1 \end{aligned} \tag{3.4.42}$$

Using the "method of steps" [8], we find the solution of (3.4.42) in the interval $[0,2]$ as

$$N(t) = \begin{cases} 1 & , \text{ for } t \in [0,1] \\ 2-t & , \text{ for } t \in [1,2] \end{cases} \tag{3.4.43}$$

The "characteristic " function N(t) is plotted in Fig. 3.4.1 :

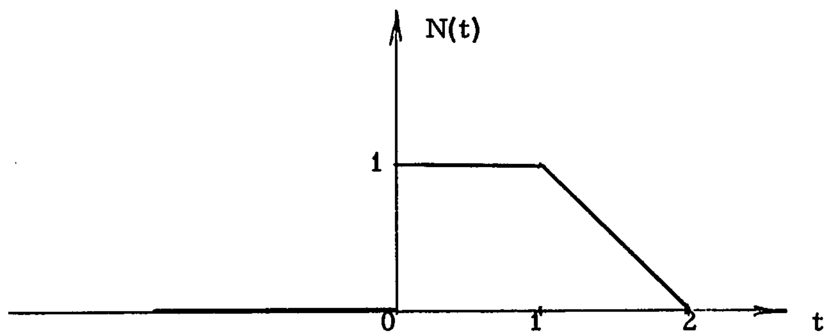

Fig . 3.4.1

We shall now apply the results of Remarks 3.4.1 and 3.4.5, observing that system (3.4.39) is of the form (3.4.27) , with $A_0(t) = A_1(t) = A_2(t) \equiv 1$, $f_{(0)}(x,u,t) = u$, $f_{(1)}(x,u,t) = -x$, $f_{(2)}(x,u,t) = u$ and $L(t) \equiv 0$.

The adjoint system is now :

(3.4.44) $$\dot{\psi}(t) = +\,\psi(t+1) \quad , \quad t \in [0,2]$$
with $\psi(t) = 0$ for $t > 2$

From (3.4.29) and (3.4.30) we obtain the Hamiltonian functionals

(3.4.45)
$$\hat{H} = \sum_{i=0}^{2} (\hat{A}_i(t+h_i)\hat{f}_{(i)}(\hat{x}(t),u,t) \ , \ \hat{\psi}(t+h_i)) \quad , \quad t \in [0,2]$$
$$\hat{H}_p = \sum_{i=0}^{2} \theta_i(t)\,(\hat{A}_i(t+h_i)\hat{f}_{(i)}(\hat{\phi}(t),u,t) \ , \ \hat{\psi}(t+h_i)), \ t \in [-2,0)$$

where the augmented quantities are defined as :

$\hat{x} = [\,x^0, x\,]^T$, ($\dot{x}^0(t) \triangleq f^0(x,u,t) = u^2(t)$) , $\hat{f}_{(0)} = [f^0, f_{(0)}]^T$, $\hat{f}_{(i)} = [0, f_{(i)}]^T$, (i=1,2) , $\hat{\psi} = [\,\psi_0, \psi\,]^T$

$$\hat{A}_0 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} , \quad \hat{A}_i = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix} , \quad (\, i = 1,2 \,)$$

and $h_0 = 0$, $h_1 = 1$, $h_2 = 2$,

$$\theta_0(t) = \begin{cases} 1 , & t \in [0,2) \\ 0 , & t \notin [0,2) \end{cases} , \quad \theta_1(t) = \begin{cases} 1 , & t \in [-1,1) \\ 0 , & t \notin [-1,1) \end{cases} , \quad \theta_2(t) = \begin{cases} 1 , & t \in [-2,0) \\ 0 , & t \notin [-2,0) \end{cases}$$

Thus, from (3.4.45),

$$\hat{H} = \psi_0 u^2(t) + \psi(t)u(t) - x(t)\psi(t+1) + \psi(t+2)u(t) \quad , \quad t \in [0,2]$$

(3.4.46)

$$\hat{H}_p = \psi(t+2)\, u(t) \quad , \quad t \in [-2,-1)$$

$$= -\phi(t)\, \psi(t+1) + \psi(t+2)u(t) \quad , \quad t \in [-1,0)$$

Since the restraint set for the control $u(t)$, $t \in [0,2]$, is the whole space $E^1 = R$, the maximum principle implies that the optimal control will be merely an extremal of $\hat{H}$, i.e., $\frac{\partial \hat{H}}{\partial u} = 0$

or ,

(3.4.47) $$2\psi_0 u(t) + \psi(t) + \psi(t+2) = 0$$

or,

(3.4.48) $$u(t) = -\frac{\psi(t) + \psi(t+2)}{2\psi_0} \quad , \quad t \in [0,2]$$

($\psi_0 \neq 0$, since otherwise the optimization problem has no solution)

The solution of the adjoint system (3.4.44) is (s. Appendix B) :

(3.4.49) $$\psi(t) = N(2-t)\, \psi(2)$$

Thus, from (3.4.48) and (3.4.49) ,

(3.4.50) $$u(t) = k\, \{ N(2-t) + N(-t) \} \quad , \quad t \in [0,2]$$

where $k = -\frac{\psi(2)}{2\psi_0}$, or

(3.4.51) $$u(t) = \begin{cases} k & , \ t = 0 \\ kt & , \ t \in (0,1] \\ k & , \ t \in [1,2] \end{cases}$$

From the maximum principle for the optimal initial control, whose restraint set is $U_1 = [-1,1]$, and $\hat{H}_p$ from (3.4.46), we obtain the form of the optimal initial control :

(3.4.52) $$u(t) = \operatorname{sgn}\{\psi(t+2)\}, \qquad t \in [-2,0)$$

or,

(3.4.53) $$u(t) = \operatorname{sgn}\{N(-t)\psi(2)\} = \operatorname{sgn}\{N(-t)\}\operatorname{sgn}\{\psi(2)\} = \operatorname{sgn}\{\psi(2)\} = a = \text{const.}$$

for $t \in [-2,0)$.

We shall determine now the two unknown constants k and a from the condition

(3.4.54) $$x(2) = -2$$

From (3.4.41), (3.4.51) and (3.4.53) we obtain

(3.4.55)
$$\begin{aligned} -2 = x(2) &= N(2)x(0) + \int_{-1}^{0} N(2-s-1)\phi(s)ds + \\ &+ \int_0^2 N(2-s)\{u(s)+u(s-2)\}ds = \int_1^2 N(s)ds + \int_0^1 N(2-s)ksds + \\ &+ \int_1^2 N(2-s)kds + a\int_0^2 N(2-s)ds = 0.5 + k\int_1^2 (s-2)^2 ds + k + 1.5a = \\ &= 0.5 + \frac{k}{3} + k + 1.5a \end{aligned}$$

or,

(3.4.56) $$\frac{4k}{3} + 1.5a = -2.5$$

Now, $a = \operatorname{sgn}\{\psi(2)\} = \operatorname{sgn}\{-2\psi_0 k\} = \operatorname{sgn} k$

If $k > 0$ then $a = +1$ and $\frac{4k}{3} + 1.5a > -2.5$

If $k < 0$ then $a = -1$ and, from (3.4.56) we obtain

(3.4.57) $$k = -0.75$$

Thus, $k = -0.75$ and $a = -1$ determine a unique solution of the optimization problem.

The optimal control is now ,

$$(3.4.58) \quad u(t) = \begin{cases} -1 & , \; t \in [-2,0) \\ -0.75 & , \; t = 0 \\ -0.75t & , \; t \in (0,1] \\ -0.75 & , \; t \in [1,2] \end{cases} \qquad \text{(optimal initial control)}$$

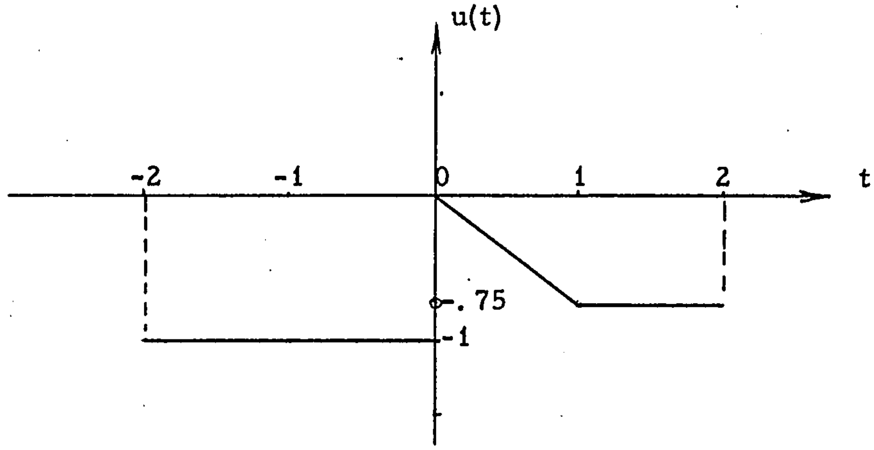

Fig. 3.4.2

The corresponding to (3.4.58) minimum cost is

$$(3.4.59) \qquad J[u] = \int_0^2 u^2(t)dt = 0.75$$

If instead of the optimal initial control $u(t) = -1$, for $t \in [-2,0)$, we use any other constant control $u(t) = v$, $1 \geq v > -1$, $t \in [-2,0)$, then, from (3.4.55), (3.4.56) and (3.4.51), we can easily find that the cost is going to be

$$(3.4.60) \qquad 0.75 < J[u] \leq 12$$

Remark 3.4.6 : In the above example, since $N(2) = 0$, we can see from (3.4.55) that every point $x(0) = x_0 \in E^1$ can be transferred to $x(2) = -2$, in the same time $t = 2$, optimally, with the use of the control (3.4.58).

In the next section we are going to apply the results of Section 3.4 to the time-optimal study of linear hereditary systems.

3.5 LINEAR TIME-OPTIMAL PROBLEM

In this section we shall apply the results of Corollary 3.4.2 to linear hereditary processes .

Suppose that :

$$f(x, u, t) = B'x + C'u + d'(t) \tag{3.5.1}$$

where B' and C' are constant matrices of dimension $n \times n$ and $n \times r$, respectively , and $d'(t)$ is an $n \times 1$ vector valued function , continuous in t.

We shall assume that the control restraint set U is a compact and convex polyhedron in E^r, and to be more specific let

$$U = \{(u_1, u_2, \ldots, u_r) \in E^r : |u_i| \leq 1 , i=1,2,\ldots,r\} \tag{3.5.2}$$

System (3.2.2) reduces now to the linear hereditary system :

$$\begin{aligned} \dot{x}(t) = \sum_{i=0}^{m} B_i(t)x(t-h_i) + \int_a^t L_1(t-s)x(s)ds + \sum_{i=0}^{m} C_i(t)u(t-h_i) + \\ + \int_a^t L_2(t-s)u(s)ds + d(t) \qquad , \quad t \in I \end{aligned} \tag{3.5.3}$$

with

$$x(t) = \phi(t) \text{ for } t \in [a, t_0) , \quad x(t_0) = x_0 \tag{3.5.4}$$

where

$$\begin{aligned} B_i(t) &= A_i(t)B' , \quad L_1(t) = L(t)B' \\ C_i(t) &= A_i(t)C' , \quad L_2(t) = L(t)C' \\ d(t) &= \sum_{i=0}^{m} A_i(t)d'(t-h_i) + \int_a^t L(t-s)d'(s)ds \end{aligned} \tag{3.5.5}$$

From (B.1.11) of Appendix B we obtain the following solution of system (3.5.3) :

$$x(t) = N(t_0, t)x(t_0) + \sum_{i=0}^{m} \int_{t_0-h_i}^{t_0} N(s+h_i, t)B_i(s+h_i)\phi(s)ds +$$

$$(3.5.6) \qquad + \int_{a}^{t_0} \{ \int_{t_0}^{t} N(s,t)L_1(s-\tau)ds \} \phi(\tau)d\tau +$$

$$+ \int_{t_0}^{t} N(s,t) \{ \sum_{i=0}^{m} C_i(s)u(s-h_i) + \int_{a}^{s} L_2(s-\tau)u(\tau)d\tau + d(s) \} ds$$

with $x(t) = \phi(t)$ for $t \in [a, t_0)$, and $x(t_0) = x_0$

where $N(s,t)$ (<u>Kernel Matrix of the Second Kind</u> , as given by Oǧuztöreli [4 , pp. 91-92]) satisfies the system

$$(3.5.7) \qquad \frac{\partial N(s,t)}{\partial t} = \sum_{i=0}^{m} B_i(t)N(s, t-h_i) + \int_{t_0}^{t} L_1(t-\tau)N(s,\tau)d\tau$$

with $N(t,t) = I$, $N(s,t) = 0$ for $s>t$, $t \in [t_0, t_1]$, $s \in [t_0, t]$

In [4, pp. 91-92] another matrix $M(s,t)$ (<u>Kernel Matrix of the First Kind</u>) is introduced , which satisfies the system :

$$(3.5.8) \qquad \frac{\partial M(s,t)}{\partial t} = \sum_{i=0}^{m} B_i(t)M(s, t-h_i) + \int_{t_0}^{t} L_1(t-\tau)M(s,\tau)d\tau \quad , t \in I$$

with $Ms,t) = \delta(t-s)I$ for $t \in [a, t_0]$, $s \in [a, t_0]$

and by use of which the solution of (3. 5. 3) is given by :

$$(3.5.9) \qquad x(t) = \int_{a}^{t_0} M(s,t)\phi(s)ds + \int_{t_0}^{t} N(s,t) \{ \sum_{i=0}^{m} C_i(s)u(s-h_i) + \int_{a}^{s} L_2(s-\tau)u(\tau)d\tau +$$

$$+ d(s) \} ds \quad , \quad t \in [a, t_1] \setminus \{t_0\}$$

Expression (3. 5. 9) is simpler than (3. 5. 6) , but the computation of $M(s,t)$ introduces an extra difficulty to the problem . We therefore prefer using expression (3. 5. 6), which requires the knowledge of one kernel matrix only .

The adjoint system in the present linear case becomes :

(3.5.10) $$\dot{\psi}(t) = -\sum_{i=0}^{m} B_i^T(t+h_i)\psi(t+h_i) - \int_t^{t_1} L_1^T(s-t)\psi(s)ds$$

with $\psi(t_1) = \psi_1$ and $\psi(t) = 0$ for $t > t_1$.

The solution of (3.5.10) is given, from (B.2.5) of Appendix B, as

(3.5.11) $$\psi(t) = N^T(t, t_1)\psi(t_1)$$

where the matrix $N(s,t)$ is given by (3.5.7) .

We shall now derive an extension of a Theorem of J. P. LaSalle [13].

THEOREM 3.5.1

All time-optimal control functions u of the linear system (3.5.3) are of the form :

(3.5.12) $$u(t) = \begin{cases} \operatorname{sgn}\{ [\sum_{i=0}^{m} C_i^T(t+h_i)N^T(t+h_i, t_1) + \int_t^{t_1} L_2^T(s-t)N^T(s,t_1)ds]\psi(t_1) \} & t \in [t_0, t_1] \\ \operatorname{sgn}\{ [\sum_{i=0}^{m} \theta_i(t)C_i^T(t+h_i)N^T(t+h_i, t_1) + \int_{t_0}^{t_1} L_2^T(s-t)N^T(s,t_1)ds]\psi(t_1) \} & t \in [\alpha, t_0) \end{cases}$$

Proof :

From Corollary 3.4.2 , we can write the Maximum Principle for system (3.5.3) as :

(3.5.13) $$\begin{aligned} H'(\psi(\cdot), u(t), t) &= \max_{v \in U} H'(\psi(\cdot), v, t) \quad \text{a.e. on } [t_0, t_1] \\ H_p'(\psi(\cdot), u(t), t) &= \max_{v \in U} H_p'(\psi(\cdot), v, t) \quad \text{a.e. on } [\alpha, t_0) \end{aligned}$$

where

$$H'(\psi(\cdot),u(t),t) \triangleq \sum_{i=0}^{m} (C_i(t+h_i)u(t), \psi(t+h_i)) + \int_t^{t_1} (L_2(s-t)u(t), \psi(s))ds \qquad t \in I$$

(3.5.14)

$$H_p'(\psi(\cdot),u(t),t) \triangleq \sum_{i=0}^{m} \theta_i(t)(C_i(t+h_i)u(t), \psi(t+h_i)) + \int_{t_0}^{t_1} (L_2(s-t)u(t), \psi(s))ds \qquad t \in [a, t_0)$$

Since these Hamiltonia functionals are linear functions of the control ,they achieve their maximum on the boundary of the compact and convex restraint set U .

From (3.5.14) ,

$$\begin{aligned} H'(\psi(\cdot),u(t),t) &= (u(t), \sum_{i=0}^{m} C_i^T(t+h_i)\psi(t+h_i)) + \int_t^{t_1} (u(t), L_2^T(s-t)\psi(s))ds = \\ (3.5.15) \qquad &= (u(t) , \sum_{i=0}^{m} C_i^T(t+h_i)\psi(t+h_i) + \int_t^{t_1} L_2^T(s-t)\psi(s)ds) = \\ &= (u(t) , [\sum_{i=0}^{m} C_i^T(t+h_i)N^T(t+h_i, t_1) + \int_t^{t_1} L_2^T(s-t)N^T(s, t_1)ds] \psi(t_1)) \end{aligned}$$

and , similarly ,

$$(3.5.16) \qquad H_p'(\psi(\cdot),u(t),t) = (u(t), [\sum_{i=0}^{m} \theta_i(t) C_i^T(t+h_i)N^T(t+h_i, t_1) + \int_{t_0}^{t_1} L_2^T(s-t)N^T(s, t_1)ds] \psi(t_1))$$

From (3.5.13) , (3.5.15) and (3.5.16) , and the fact that U is the hypercube (3.5.2) , we obtain easily (3.5.12) .

Q.E.D.

<u>Definition 3.5.1</u>

The control system (3.5.3) is said to be <u>normal</u> , if no component of the vectors

$$(3.5.17)\quad V_1(t)\psi(t_1) = \{\sum_{i=0}^{m} C_i^T(t+h_i)N^T(t+h_i,t_1) + \int_t^{t_1} L_2^T(s-t)N^T(s,t_1)ds\}\psi(t_1)$$

$$t\in[t_0,t_1]$$

and

$$(3.5.18)\quad V_2(t)\psi(t_1) = \{\sum_{i=0}^{m}\theta_i(t)C_i^T(t+h_i)N^T(t+h_i,t_1) + \int_{t_0}^{t_1} L_2^T(s-t)N^T(s,t_1)ds\}\psi(t_1)$$

$$t\in[a,t_0)$$

is identically zero on an interval of positive length .

In other words the system (3.5.3) is normal, if the columns of the matrices $V_1(t)$ and $V_2(t)$ are linearly independent on every interval of positive length .

Remark 3.5.1 : From Theorem 3.5.1 and for a normal system, we see that , for a given nontrivial adjoint function ψ , the maximum principle (3,5,13) determines uniquely (a.e.) the control u .

We can now state the following uniqueness theorem for the optimal control of a normal, linear hereditary system .

THEOREM 3.5.2

Let $u_1(t)$ and $u_2(t)$ be two optimal controls (given on the intervals $[a,t_1]$ and $[a,t_2]$, respectively) , which transfer the state of system (3.5.3) from the point x_0 to the same point x_1, with the same initial state function ϕ . Then, if the system is normal , these controls coincide a.e. , i.e. , $t_1 = t_2$ and $u_1(t) \equiv u_2(t)$ a.e. on $[a,t_1]$.

Proof :

First, it is clear that $t_1 = t_2$, since if $t_1 > t_2$ or $t_1 < t_2$ then u_2 or u_1 , respectively , is not a time-optimal control .

From (3.5.6) and for $t=t_1$, we have :

(3.5.19)

$$x_1 = x(t_1) = N(t_0,t_1)x_0 + \sum_{i=0}^{m} \int_{t_0-h_i}^{t_0} N(s+h_i,t_1)B_i(s+h_i)\phi(s)ds + \int_{a}^{t_0} \{\int_{t_0}^{t_1} N(s,t_1)L_1(s-\tau)ds\}\phi(\tau)d\tau + \int_{t_0}^{t_1} N(s,t_1)\{\sum_{i=0}^{m} C_i(s)u_1(s-h_i) + \int_{a}^{s} L_2(s-\tau)u_1(\tau)d\tau + d(s)\}ds =$$

$$= N(t_0,t_1)x_0 + \sum_{i=0}^{m} \int_{t_0-h_i}^{t_0} N(s+h_i,t_1)B_i(s+h_i)\phi(s)ds + \int_{a}^{t_0} \{\int_{t_0}^{t_1} N(s,t_1)L_1(s-\tau)ds\}\phi(\tau)d\tau + \int_{t_0}^{t_1} N(s,t_1)\{\sum_{i=0}^{m} C_i(s)u_2(s-h_i) + \int_{a}^{s} L_2(s-\tau)u_2(\tau)d\tau + d(s)\}ds$$

from which ,

(3.5.20)

$$\int_{t_0}^{t_1} N(s,t_1)\{\sum_{i=0}^{m} C_i(s)u_1(s-h_i) + \int_{a}^{s} L_2(s-\tau)u_1(\tau)d\tau\}ds = \int_{t_0}^{t_1} N(s,t_1)\{\sum_{i=0}^{m} C_i(s)u_2(s-h_i) + \int_{a}^{s} L_2(s-\tau)u_2(\tau)d\tau\}ds$$

Now, from (3.5.20) and the ψ function whose existence is asserted from Corollary 3.4.2 , we obtain :

(3.5.21)

$$\left(\int_{t_0}^{t_1} N(s,t_1)\{\sum_{i=0}^{m} C_i(s)u_1(s-h_i) + \int_{a}^{s} L_2(s-\tau)u_1(\tau)d\tau\}ds,\ \psi(t_1)\right) = \left(\int_{t_0}^{t_1} N(s,t_1)\{\sum_{i=0}^{m} C_i(s)u_2(s-h_i) + \int_{a}^{s} L_2(s-\tau)u_2(\tau)d\tau\}ds,\ \psi(t_1)\right)$$

Now,

$$(\int_{t_0}^{t_1} N(s,t_1)\{\sum_{i=0}^{m} C_i(s)u_k(s-h_i) + \int_{a}^{s} L_2(s-\tau)u_k(\tau)d\tau\}ds\ ,\ \psi(t_1)\ ) =$$

$$= \int_{t_0}^{t_1}(N(s,t_1)\{\sum_{i=0}^{m} C_i(s)u_k(s-h_i)\}, \psi(t_1)\ )ds\ + \int_{t_0}^{t_1}\int_{a}^{s}(N(s,t_1)L_2(s-\tau)u_k(\tau), \psi(t_1))d\tau ds$$

(3.5.22)

$$= \sum_{i=0}^{m}\int_{t_0}^{t_1}(\ N^T(s,t_1)\psi(t_1)\ \ ,\ \ C_i(s)u_k(s-h_i)\ )\ ds\ +$$

$$+ \int_{t_0}^{t_1}\int_{a}^{s}(\ N^T(s,t_1)\psi(t_1)\ \ ,\ \ L_2(s-\tau)u_k(\tau)\ )d\tau ds\ =$$

$$= \sum_{i=0}^{m}\int_{t_0}^{t_1}(\ \psi(s)\ \ ,C_i(s)u_k(s-h_i)\ )ds + \int_{t_0}^{t_1}\int_{a}^{s}(\psi(s)\ ,\ L_2(s-\tau)u_k(\tau))d\tau ds$$

where $k = 1, 2$.

From (3.5.21) and (3.5.22), we obtain :

$$\sum_{i=0}^{m}\int_{t_0}^{t_1}(\ C_i(s)u_1(s-h_i)\ ,\ \psi(s)\ )ds + \int_{t_0}^{t_1}\int_{a}^{s}(\ L_2(s-\tau)u_1(\tau)\ ,\ \psi(s))d\tau ds =$$

(3.5.23)

$$= \sum_{i=0}^{m}\int_{t_0}^{t_1}(\ C_i(s)u_2(s-h_i)\ ,\ \ \psi(s)\ )ds + \int_{t_0}^{t_1}\int_{a}^{s}(L_2(s-\tau)u_2(\tau)\ ,\psi(s))d\tau ds$$

Now, from the Hamiltonians (3.5.14) , we obtain :

$$\int_{t_0}^{t_1} H'(\ \psi(\cdot), u_k(s), s\ )ds = \int_{t_0}^{t_1}\sum_{i=0}^{m}(C_i(s+h_i)u_k(s),\ \psi(s+h_i)\ )ds\ +$$

$$+ \int_{t_0}^{t_1}\int_{s}^{t_1}(L_2(\tau-s)u_k(s)\ ,\ \psi(\tau)\ )d\tau ds =$$

(3.5.24)

$$= \sum_{i=0}^{m}\int_{t_0+h_i}^{t_1}(C_i(s)u_k(s-h_i)\ ,\psi(s)\ )ds\ \ +$$

$$+ \int_{t_0}^{t_1}\int_{t_0}^{s}(L_2(s-\tau)u_k(\tau)\ ,\ \psi(s)\ )d\tau ds$$

$$\int_a^{t_0} H'_p(\psi(\cdot),u_k(s),s)ds = \int_a^{t_0}\sum_{i=0}^{m}\theta_i(s)(C_i(s+h_i)u_k(s),\ \psi(t+h_i))ds +$$

$$+\int_a^{t_0}\int_{t_0}^{t_1}(L_2(\tau-s)u_k(s),\ \psi(\tau))d\tau ds =$$

(3.5.25)

$$= \sum_{i=0}^{m}\int_{t_0}^{t_0+h_i}(C_i(s)u_k(s-h_i),\ \psi(s))ds +$$

$$+\int_{t_0}^{t_1}\int_a^{t_0}(L_2(s-\tau)u_k(\tau),\ \psi(s))dsd\tau$$

where $k = 1, 2$.

From (3.5.24) and (3.5.25) we obtain :

$$\int_{t_0}^{t_1} H'(\psi(\cdot),u_k(s),s)\,ds + \int_a^{t_0} H'_p(\psi(\cdot),u_k(s),s)ds =$$

(3.5.26)

$$= \sum_{i=0}^{m}\int_{t_0}^{t_1}(C_i(s)u_k(s-h_i),\psi(s))ds + \int_{t_0}^{t_1}\int_a^{s}(L_2(s-\tau)u_k(\tau),\psi(s))d\tau ds$$

and using (3.5.23) ,

$$\int_{t_0}^{t_1} H'(\psi(\cdot),u_1(s),s)\,ds + \int_a^{t_0} H'_p(\psi(\cdot),u_1(s),s)\,ds =$$

(3.5.27)

$$= \int_{t_0}^{t_1} H'(\psi(\cdot),u_2(s),s)ds + \int_a^{t_0} H'_p(\psi(\cdot),u_2(s),s)\,ds$$

or,

$$\int_{t_0}^{t_1}\{H'(\psi(\cdot),u_1(s),s) - H'(\psi(\cdot),u_2(s),s)\}ds +$$

(3.5.28)

$$+\int_a^{t_0}\{H'_p(\psi(\cdot),u_1(s),s) - H'_p(\psi(\cdot),u_2(s),s)\}ds = 0$$

Now, since u_1 is an optimal control ,

$$(3.5.29)\quad \begin{aligned} H'(\psi(\cdot),u_1(s),s) &\geq H'(\psi(\cdot),u_2(s),s) , \quad \text{a.e. on } I \\ H_p'(\psi(\cdot),u_1(s),s) &\geq H_p'(\psi(\cdot),u_2(s),s) , \quad \text{a.e. on}[\alpha,t_0) \end{aligned}$$

From (3.5.28) and (3.5.29) we conclude that ,

$$(3.5.30)\quad \begin{aligned} H'(\psi(\cdot),u_1(t),t) &= H'(\psi(\cdot),u_2(t),t) , \quad \text{a.e. on } I \\ H_p'(\psi(\cdot),u_1(t),t) &= H_p'(\psi(\cdot),u_2(t),t) , \quad \text{a.e. on } [\alpha,t_0) \end{aligned}$$

i.e. , u_1 and u_2 satisfy the maximum principle (3.5.13) ,for the same nontrivial ψ . Hence , it follows from the Remark 3.5.1 that

$$(3.5.31)\qquad u_1(t) \equiv u_2(t) \quad , \quad \text{a.e.} \quad \text{on} \quad [\alpha,t_1]$$

Q.E.D.

COROLLARY 3.5.1

For normal control systems of the type (3.5.3) there is at most one optimal control function and if there is one , it is of the "bang-bang" type (3.5.12) .

Thus , for a normal control system,the only way to reach the target in minimum time is by "bang-bang" controls of the form (3.5.12) .

Definition 3.5.2

We shall call a control $u(t)$, $t \in [\alpha,t_1]$, an extremal , if it satisfies the maximum principle (3.5.13) , where ψ is some nontrivial solution of system (3.5.10) .

To find the optimal control which transfers the state from x_0 to x_1 , one may first find all the extremal controls which realize this transfer , and then pick up the unique (by virtue of Theorem 3.5.2) extremal control which realizes the transition in the shortest time .

The following example gives an application of the results of Section 3.5 , and demonstrates the effect of the initial control in the optimal time of a normal , linear, time -delay system .

EXAMPLE 3.5.1 : A time-optimal problem

Consider the system (3.5.32) in $E^1=R$:

$$(3.5.32) \qquad \dot{x}(t) = x(t) + x(t-\theta) + u(t) + u(t-\theta) \quad , \qquad t \in [0,T]$$

with $x(t) = \phi(t) = \text{const.} = c$, for $t \in [-\theta, 0)$, $x(0) = c$

and $u \in \Delta$, $u(t) = U = [-1,1]$.

The optimization problem can be stated as follows :

Find an admissible control in Δ, which transfers the state of system (3.5.32) from the initial point $x(0) = c$ to the origin $x(T) = 0$, in minimum time T .

From (3.5.6), we obtain the solution of system (3.5.32) as :

$$(3.5.33) \qquad x(t) = N(t)x(0) + \int_{-\theta}^{0} N(t-s-\theta)\phi(s)ds + \int_{0}^{t} N(t-s)\{u(s) + u(s-\theta)\}ds$$

where $N(t)$ satisfies the system (s. Appendix B) :

$$(3.5.34) \qquad \dot{N}(t) = N(t) + N(t-\theta)$$

with $N(t) = 0$ for $t < 0$, and $N(0) = 1$

The solution of (3.5.34) is obtained by the "method of steps" [8] and it is

(3.5.35) $$N(t) = \sum_{n=0}^{[\frac{t}{\theta}]} \frac{(t-n\theta)^n}{n!} e^{t-n\theta} \quad , t \in [0, T]$$

where $[r]$ is the "integral part" of the real number r.

The optimal controls are obtained from (3.5.12):

(3.5.36) $$u(t) = \begin{cases} \operatorname{sgn}\{N(T-t-\theta)\}\operatorname{sgn}\{\psi(T)\} & , t \in [-\theta, 0) \\ \operatorname{sgn}\{N(T-t-\theta)+N(T-t)\}\operatorname{sgn}\{\psi(T)\} & , t \in [0, T-\theta) \\ \operatorname{sgn}\{N(T-t)\}\operatorname{sgn}\{\psi(T)\} & , t \in [T-\theta, T] \end{cases}$$

(if $T < \theta$ ommit the second expression and consider the third for $t \in [0, T]$)

From (3.5.35) we observe that $N(t) > 0$ for all $t \geq 0$. Thus, (3.5.36) reduces to

(3.5.37) $$u(t) = \operatorname{sgn}\{\psi(T)\} = \text{const.} \quad \text{for } t \in [-\theta, T]$$

This means that in the elementary example which we consider, the 'switching set' of the optimal control is empty. In general, the switching times would have been found as solutions of the equations:

(3.5.38) $$\begin{aligned} N(T-t-\theta) &= 0 & , t \in [-\theta, 0) \\ N(T-t-\theta) + N(T-t) &= 0 & , t \in [0, T-\theta) \\ N(T-t) &= 0 & , t \in [T-\theta, T] \end{aligned}$$

provided that the optimal time T is known [7].

Let $c = 0.5$. From (3.5.37) we see that the $\operatorname{sgn}\{\psi(T)\}$ determines the value of the optimal control. Thus, we shall try to find the value of

(3.5.39) $$a \doteq \operatorname{sgn}\{\psi(T)\}$$

for which the initial state $x_0 = 0.5$ can be transfered to the origin of the state space .

From (3.5.33) and for $\phi(t) = c = 0.5$, $u(t) = a$, we obtain :

$$x(t) = cN(t) + c\int_{-\theta}^{0} N(t-s-\theta)ds + 2a\int_{0}^{t} N(t-s)ds =$$

(3.5.40)

$$= cN(t) + c\int_{t-\theta}^{t} N(s)ds + 2a\int_{0}^{t} N(s)ds$$

To determine now the optimal time T , we have to solve the equation $x(t) = 0$, and take $T = \inf \{ t : x(t) = 0 \}$.

From (3.5.40) ,

$$\dot{x}(t) = c\dot{N}(t) + cN(t) - cN(t-\theta) + 2aN(t) =$$

(3.5.41) $$= cN(t) + cN(t-\theta) + cN(t) - cN(t-\theta) + 2aN(t) =$$

$$= 2(a + c)N(t)$$

If $\dot{x}(t) \geq 0$ for all $t > 0$, then the system never comes to the origin $x = 0$.

If , however , $\dot{x}(t) < 0$ for all $t > 0$, then there is a time $T = t_{opt}$ for which $x(T) = 0$, while $x(t) > 0$ for all $t \in [0, T)$.

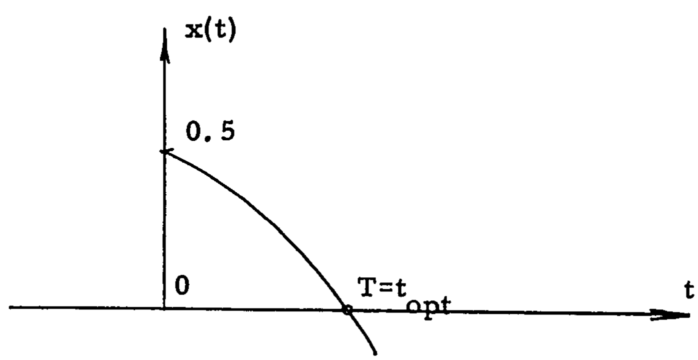

Fig. 3.5.1

Thus, from (3.5.41) we see that the initial state $x_0=c=0.5$ can be transferred to the origin, if

$$a + c < 0$$

or,

$$a < -0.5$$

Therefore, $a = -1$ and the optimal control is

$$(3.5.42) \qquad u(t) = \begin{cases} -1 & , t \in [-\theta, 0) \quad \text{(optimal initial control)} \\ -1 & , t \in [0, T] \end{cases}$$

Now, we can solve the equation

$$(3.5.43) \qquad x(t) = 0.5N(t) + 0.5\int_{t-\theta}^{t} N(s)ds - 2\int_{0}^{t} N(s)ds = 0$$

using the Newton - Raphson scheme :

$$(3.5.44) \qquad t^{k+1} = t^k - \frac{x(t^k)}{\dot{x}(t^k)}$$

where $\dot{x}(t^k) = -N(t^k)$

The results are shown in Table 3.5.1, for different values of the delay :

θ	0.1	0.2	0.3
$T=t_{opt}$	0.374194	0.391520	0.401763

Table 3.5.1

We observe that the optimal time increases with the delay θ.

Now, let $\theta = 0.1$ and suppose that we vary the initial control $u(t) = v \in [-1,1]$, $t \in [-\theta, 0)$. It can be easily shown that the optimal control is again $u(t) = -1$, $t \in [0, T]$.

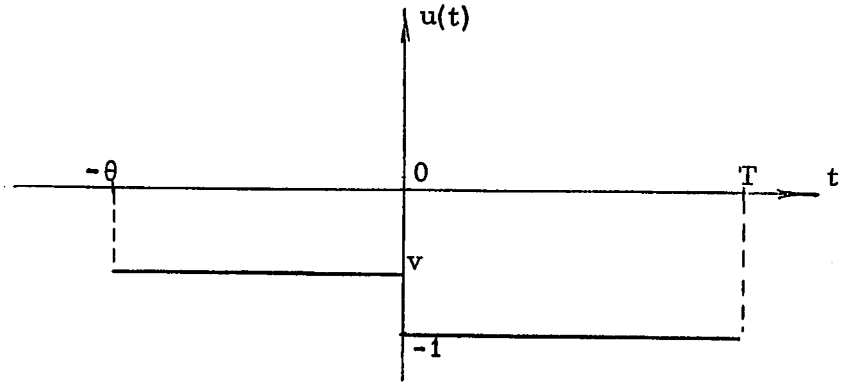

Fig. 3.5.2

Then, the optimal time varies with the initial control, as shown in Table 3.5.2 :

v	-1	-0.8	-0.6	-0.4	-0.2	0
T	0.374194	0.392762	0.412017	0.431969	0.452635	0.474154

Table 3.5.2

We observe that indeed the optimal time increases as the initial control v goes away from its optimal value, found above .

Remark 3.5.2 : The necessary condition $R(T) \geq 0$ (s. Corollary 3.4.2) is satisfied . Indeed

(3.5.45) $R(T) = \psi(T)\{x(T) + x(T-\theta) + u(T) + u(T-\theta)\} = \psi(T)\{x(T-\theta) - 2\} > 0$

since $\psi(T) < 0$ and $0 \leq x(t) \leq 0.5$ for $-\theta \leq t \leq T$.

3.6 DISCUSSIONS

In the study of hereditary systems, terms like "present" and "past" are often met. These concepts are quite relevant and one has to define them clearly.

In this Chapter we considered the hereditary system (3.2.2) on the time interval $[t_0, t_1]$. The initial time t_0 is specified and it is exactly the time when the system is "switched on" and starts following the dynamic equation (3.2.2). The "past" of the system, relative to that time, is specified by the initial data $u(t), \phi(t), t \in [\alpha, t_0)$. Since t_0 is a 'time of interest', and it may very well be a future instant, the initial data, which in a sense "prepare" the system's evolution, can play a very significant role in the overall performance of the system.

As a step towards the estimation of a pair of optimal initial data, we derived in this Chapter necessary conditions that an optimal control and an optimal initial control must satisfy, with the initial state function considered as given. The method used was Pontryagin's geometric approach, modified properly to handle problems introduced by the heredity of the system. The Oğuztöreli's kernel matrices of the first and second kind, were extremely useful in the proof of the Theorem 3.4.1. The Maximum Principle for the optimal initial control was introduced with the use of a second Hamiltonian functional. An example of an energy optimal problem shows the effect of the optimal choice of the initial control in the cost of the system.

In the linear time-optimal case, it is found that the optimal control of a normal system is, at most, one. That is, if there exists an optimal control, then it is unique. An example shows the improvement of the optimal time of a simple time-delay system, caused by the optimization of the initial control.

CHAPTER 4

OPTIMAL CONTROL ON THE POLICY SPACE $P = \Delta \times \Phi$

4.1 INTRODUCTION

In the previous Chapter we considered several optimization problems for a class of hereditary systems, with the initial state function specified, and we derived necessary conditions that an optimal control must satisfy.

In this Chapter, we consider the same class of systems and we derive necessary conditions that an optimal pair $\{u, \phi\}$ must satisfy. In other words our policy space is now

$$P = \Delta \times \Phi \tag{4.1.1}$$

An example shows the improvement achieved in the optimal time by the use of an optimal initial state function.

4.2 STATEMENT OF THE PROBLEM

Consider the system (4.2.1) $\equiv$ (3.2.2) in E^n:

$$\dot{x}(t) = \sum_{i=0}^{m} A_i(t) f(x(t-h_i), u(t-h_i), t-h_i) + \int_{a}^{t} L(t-s) f(x(s), u(s), s)\, ds \tag{4.2.1}$$

for $t \in I = [t_0, t_1] \subset [t_0, \infty)$, and with the initial data $x(t_0) = x_0$, $x(t) = \phi(t)$ for $t \in [a, t_0) \subset (-\infty, t_0)$

where

i) $x(t) \in E^n$ for $t \in I$

ii) $u(t) \in U \subset E^r$ for $t \in [a, t_1]$

iii) f , $f_t \in C(E^{n+1} x\, U, E^n)$, $f_x \in C$

iv) $A_i(t)$, ($i=0,1,\ldots,m$) , are $n \times n$ matrix valued functions defined on I and continuous in t .

v) $L(t)$ is an integrable $n \times n$ matrix valued function , defined on $[a, t_1]$.

We assume that

1) $\phi \in \Phi = C([a, t_0), G)$, where G is a compact subset of E^n .

2) U , the control restraint set , is a compact subset of E^r .

3) Δ , the set of admissible control functions , is the class of all bounded, measurable functions , defined on $[a, t_1]$ and taking values in U .

The cost functional considered here is

$$J[u, \phi] = \int_{t_0}^{t_1} k^0(t_1 - s)\, f^0(x(s), u(s), s)\, ds \tag{4.2.2}$$

where $k^0 \in C^1(I)$, f^0 , $f^0_t \in C(E^{n+1} x\, U, E^1)$, $f^0_x \in C(E^{n+1} x\, U, E^n)$

In addition , two smooth manifolds S_0 and S_1 in E^n, are given.

The optimization problem can be stated now as follows : Find necessary conditions for an admissible pair $\{u, \phi\} \in P = \Delta x \Phi$ to be optimal ; i.e., minimizing the cost functional (4.2.2), while transfering the state of system (4.2.1) from some point $x(t_0) = x_0 \in S_0$ to a point $x(t_1) = x_1 \in S_1$.

4.3 BASIC DEFINITIONS

The following definitions are exactly the same with the ones in Section 3.3 of Chapter 3 :

4.3.1 The Variational System

(4.3.1.1) $$(\dot{\delta x})(t) = \sum_{i=0}^{m} A_i(t)F_x(t-h_i)\delta x(t-h_i) + \int_a^t L(t-s)F_x(s)\delta x(s)ds$$

for $t \in I$, with $\delta x(t_0) = \xi_0$ and $\delta x(t) = 0$ for $t \in [a, t_0)$.

4.3.2 The Adjoint System

(4.3.2.1) $$\dot{\psi}(t) = - F_x^T(t) \{ \sum_{i=0}^{m} A_i^T(t+h_i)\psi(t+h_i) + \int_t^{t_1} L^T(s-t)\psi(s)ds \}$$

for $t \in I$, with $\psi(t_1) = \psi_1$ and $\psi(t) = 0$ for $t > t_1$.

4.3.3 The Hamiltonian Functionals

(4.3.3.1)
$$H(x(t), \psi(\cdot), v, t) = \sum_{i=0}^{m} (A_i(t+h_i)f(x(t), v, t), \psi(t+h_i)) + \int_t^{t_1} (L(s-t)f(x(t), v, t), \psi(s))ds$$

$$H_p(\dot{\phi}(t), \psi(\cdot), v, t) = \sum_{i=0}^{m} \theta_i(t) (A_i(t+h_i)f(\phi(t), v, t), \psi(t+h_i)) + \int_{t_0}^{t_1} (L(t-s)f(\phi(t), v, t), \psi(s))ds$$

We also introduce the functional

(4.3.3.2) $$R(t) \triangleq R(x(\cdot), \psi(t), u(\cdot), t) =$$
$$= (\psi(t), \sum_{i=0}^{m} A_i(t)f(x(t-h_i), u(t-h_i), t-h_i) + \int_a^t L(t-s)f(x(s), u(s), s)ds)$$

A new addition to the definitions of Section 3. 3 will be the following one :

4. 3. 4 The Functional $K(\phi(\cdot), u(\cdot), \psi(\cdot))$

$$K(\phi(\cdot),u(\cdot),\psi(\cdot)) = \sum_{i=0}^{m} \int_{t_0-h_i}^{t_0} (\phi(t), F_x^T(t) A_i^T(t+h_i)\psi(t+h_i))dt + \tag{4.3.4.1}$$
$$+ \int_{\alpha}^{t_0} (\phi(t), F_x^T(t) \int_{t_0}^{t_1} L^T(s-t)\psi(s)ds)dt$$

4. 3. 5 Augmented Variables and Systems

The augmented variables and systems are defined exactly as in subsection 3. 3. 4 of Chapter 3 .

4. 4 A MAXIMUM PRINCIPLE

The following Theorem is an extension of Corollary 3. 4. 1 of Chapter 3 .

THEOREM 4. 4. 1

Consider the control process in E^n :

$$\dot{x}(t) = \sum_{i=0}^{m} A_i(t)f(x(t-h_i), u(t-h_i), t-h_i) + \int_{\alpha}^{t} L(t-s)f(x(s), u(s), s)ds \tag{4.4.1}$$

on I , with $u \in \Delta$, $u(t) \in U$ for $t \in [\alpha, t_1]$.

For each admissible pair $\{u, \phi\} \in \Delta \times \Phi$ with response $x(t)$, let the cost functional be :

$$J[u, \phi] = \int_{t_0}^{t_1} k^0(t_1-s)f^0(x(s), u(s), s)ds \tag{4.4.2}$$

If $\{u, \phi\}$ is an optimal pair in $\Delta \times \Phi$ with augmented response $\hat{x}(t)$, $t \in I$, then there exists a nontrivial adjoint response $\psi(t)$ (and augmented response $\hat{\psi}(t)$), absolutely continuous on $[t_0, t_1)$, such that :

$$\hat{H}(\hat{x}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } I$$

(i) $$\hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } [a, t_0)$$

$$K(\phi(\cdot), u(\cdot), \psi(\cdot)) = \sup_{\overset{*}{\phi} \in \Phi} K(\overset{*}{\phi}(\cdot), u(\cdot), \psi(\cdot))$$

(ii) at the terminal time $t = t_1$:

$\psi_0(t_1) \leq 0$ and $\hat{R}(t_1) = 0$

Moreover , $\psi_0(t) = \text{const.} \leq 0$ for all $t \in I$

(iii) the following transversality conditions are satisfied at both endpoints of the trajectory $x(t)$:

$$\psi(t_0) \perp S_0 \ (\text{at } x_0) \quad , \quad \psi(t_1) \perp S_1 \ (\text{at } x_1)$$

<u>Proof</u> :

The present case can be reduced to the one where f is independent of time, exactly as in the proof of Corollary 3.4.1 . Thus, we need only consider the case where f does not depend explicitly on time, and derive an extension of Theorem 3.4.1 .

Having already the proof of Theorem 3.4.1 (s. Appendix A), we shall simply indicate the changes that have to be made in the different steps of this proof .

We consider first the fixed -endpoints problem, i.e., $x(t_0)=x_0$, and $x(t_1)=x_1$, where t_0 is a fixed time and t_1 is free.

Let $\{u, \phi\}$ be an optimal pair in $\Delta \times \Phi$, with augmented response $\hat{x}(t)$, $t \in I$.

Step 1 : Perturbation of the control u, on $[a, t_1]$, and of the initial state function ϕ, on $[a, t_0)$.

The perturbed control u^* is defined as in (A. 3).

The perturbed initial state function is defined as :

$$(4.4.3) \qquad \phi^*(t) = \phi(t) + \epsilon \delta\phi(t)$$

where $\delta\phi \in C([a, t_0), E^n)$ and $\epsilon > 0$ is small enough so that $\phi^* \in \Phi = C([a, t_0), G)$.

Remark 4.4.1: If the set $A = \{t \in [a, t_0) : \phi(t) \in \partial G\}$ is not empty, then we shall consider only those perturbations $\delta\phi \in C([a, t_0), E^n)$ for which $\delta\phi(t) = 0$ for $t \in A$.

Step 2 : The basic variational formula.

Exactly as in Step 2 of the proof of Theorem 3.4.1 (s. Appendix A) we establish, by direct evaluation, the variational formula

$$(4.4.4) \qquad \hat{x}^*(t_1 + \epsilon\delta t) = \hat{x}(t_1) + \epsilon \{\hat{\Gamma} x(t_1) + \int_a^{t_0} M(s, t_1)\delta\hat{\phi}(s)ds\} + O(\epsilon)$$

where $\hat{\Gamma} x(t_1)$ is the vector given by (A. 6) and $M(s,t)$ is the Oǧuztöreli's Kernel Matrix of the First Kind [4] associated with the system :

$$\hat{z}(t) = \sum_{i=0}^{m} \int_{t_0}^{t} \hat{A}_i(s)\hat{F}_{\hat{x}}(s-h_i)\hat{z}(s-h_i)ds + \int_{t_0}^{t}\int_{a}^{\tau}\hat{L}(\tau-s)\hat{F}_{\hat{x}}(s)\hat{z}(s)ds\,d\tau +$$

(4.4.5)
$$+ \epsilon\{\varphi(t) + g(t)\} + O(\epsilon) \quad , \quad t \in [t_0, t_1]$$

$$\hat{z}(t_0) = 0$$
$$\hat{z}(t) = \epsilon\delta\hat{\phi}(t) \quad , \quad t \in [a, t_0)$$

which is analogous to system (A.19) in Appendix A . The functions $\varphi(t)$ and $g(t)$ are given by (A.13) and (A.17) .

Now , we define the complex perturbation :

(4.4.6) $$b = \{ \delta\hat{\phi}(\cdot) , \tau_i , v_i , \delta t_i , \delta t ; \quad i=1,2,\dots,s \}$$

By defining the binary operation :

(4.4.7) $$\lambda' b' + \lambda'' b'' = \{ \lambda'\delta\hat{\phi}'(\cdot) + \lambda''\delta\hat{\phi}''(\cdot), \tau_i, v_i, \lambda'\delta t_i' + \lambda''\delta t_i'', \lambda'\delta t' + \lambda''\delta t''; i=1,\dots,s\}$$

for $\lambda', \lambda'' \in R^+ = [0, \infty)$ and b', b'' the complex perturbations :

(4.4.8)
$$b' = \{ \delta\hat{\phi}'(\cdot), \tau_i, v_i , \delta t_i' , \delta t' ; \quad i=1,2,\dots,s \}$$
$$b'' = \{ \delta\hat{\phi}''(\cdot), \tau_i, v_i , \delta t_i'', \delta t''; \; i=1,2,\dots,s \}$$

we form the space :

(4.4.9) $$B = \{ b : b \text{ is a complex perturbation} \}$$

Let ,

(4.4.10) $$(\hat{Z}x)(t) \triangleq (\hat{\Gamma}x)(t) + \int_{a}^{t_0} M(s,t)\delta\hat{\phi}(s)ds$$

To emphasize the dependence on b , we write

(4.4.11) $$(\hat{Z}x)_b = (\hat{Z}x)(t_1)$$

The vectors $(\hat{Z}x)_b$, $b \in B$, emanating from the point $x(t_1)$, form a <u>convex cone</u> in E^{n+1} :

(4.4.12) $\tilde{K}_{t_1} = \{ \hat{Z}x_b : b \in B \}$

<u>Step 3</u> : <u>Fundamental Lemmas</u> .

Lemma A.1 remains valid, in the present case . The following two Lemmas are completely analogous to Lemmas A.2 and A.3 in Appendix A .

<u>Lemma 4.4.1</u>

If Λ is a smooth curve which begins at $\hat{x}(t_1) = \hat{x}_1$ and which has a tangent ray L at this point, pointing into the interior of $\tilde{K}_{t_1}$, then there exist a control u_* in Δ and an initial state function ϕ_* in Φ, such that the corresponding trajectory $\hat{x}_*(t)$, $t \in I$, starting at $\hat{x}_0$, passes through a point of Λ distinct from $\hat{x}_1$.

<u>Proof</u> :

The proof is exactly the same as of Lemma 3 in [1, pp. 94-98] in connection with (4.4.4) . Now a perturbation b and an $\epsilon > 0$ can be found , so that $\hat{x}_{*b}(t_1 + \epsilon \delta t_b) \in \Lambda$. If we denote by u_* and ϕ_* the functions which correspond to this perturbation b , then Lemma 4.4.1 is proven .

<u>Lemma 4.4.2</u>

If the pair $\{u, \phi\}$ is optimal , then the ray L starting at the point $\hat{x}(t_1)$ and pointing in the direction of the negative x^0 axis , does not belong to the interior of the cone $\tilde{K}_{t_1}$.

Proof ;

Using Lemma 4.4.1 , a proof by contradiction follows in exactly the same way as in Lemma A.3 . Namely , suppose that $L \subset \overset{\circ}{\tilde{K}}_{t_1}$. Then , by Lemma 4.4.1 and for $\Lambda \equiv L$, there exist a control u_* in Δ and an initial state function ϕ_* in Φ , such that the corresponding trajectory $\hat{x}_*(t)$, $t \in I$, starting at $\hat{x}_0$, passes through a point of L distinct from $\hat{x}_1$, i.e. , for some time t_1' ,

(4.4.13) $$x_*^i(t_1') = x^i(t_1) = x_1^i \quad , \quad (i=1,2,\dots,n)$$

and

(4.4.14) $$x_*^0(t_1') < x^0(t_1) = J[u,\phi]$$

But this contradicts the optimality of the pair $\{u,\phi\}$. Therefore $L \not\subset \overset{\circ}{\tilde{K}}_{t_1}$.

Q.E.D.

Corollary 4.4.1

There exists a hyperplane of support to $\tilde{K}_{t_1}$ at its vertex , i.e., a hyperplane such that $\tilde{K}_{t_1}$ lies entirely in one of the two closed half-spaces defined by it .

Consequently , there exists a non-zero vector $\alpha = [\alpha^0, \alpha^1, \dots, \alpha^n]^T$ such that

(4.4.15) $$(\alpha, \hat{Zx}_b) \leq 0 \quad , \text{ for all } \hat{Zx}_b \in \tilde{K}_{t_1}$$

Let again $\hat{\psi}(t)$ be the solution of the augmented adjoint system , with final value $\hat{\psi}(t_1) = \alpha$. Then ,

(4.4.16) $$(\hat{\psi}(t_1) , \hat{Zx}_b) \leq 0 \text{ , for all } \hat{Zx}_b \in \tilde{K}_{t_1}$$

Similarly to (A.42) and (A.44), we obtain

$$\psi_0(t_1) \le 0 \tag{4.4.17}$$

and, furthermore,

$$\psi_0(t) = \text{const.} \le 0 \,, \quad \text{for all} \quad t \in I \,. \tag{4.4.18}$$

Step 4 : The Maximum Principle

Considering the elementary perturbation

$$b_1 = \{ \hat{\delta\phi}(\cdot), \tau_1, v_1, \delta t_1, \delta t \} = \{ 0, \tau_1, v_1, 1, 0 \} \tag{4.4.19}$$

we reduce the problem to the previous one with fixed initial state function.

Thus, the Maximum Principle for the optimal control, obtained there, remains valid for the present case. Namely :

$$\hat{H}(\hat{x}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } I \tag{4.4.20}$$

$$\hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } [\alpha, t_0) \tag{4.4.21}$$

Similarly, considering the elementary perturbation

$$b_2 = \{ \delta\phi(\cdot), \tau_1, v_1, \delta t_1, \delta t \} = \{ 0, \tau_1, v_1, 0, \delta t \} \tag{4.4.22}$$

we obtain again, as in (A.72),

$$\hat{R}(t_1) = 0 \tag{4.4.23}$$

Now, let us consider an elementary perturbation of the form

$$b_3 = \{ \hat{\delta\phi}(\cdot), \tau_1, v_1, \delta t_1, \delta t \} = \{ \hat{\delta\phi}(\cdot), \tau_1, v_1, 0, 0 \} \tag{4.4.24}$$

From (4.4.5) and (4.4.10),

(4.4.25) $$(\hat{Z}x)(t) = \lim_{\epsilon \to 0} \frac{\hat{z}(t)}{\epsilon}$$

where $\hat{z}(t)$ is given by (4.4.5) with $\varphi(t) = g(t) = 0$, $t \in I$.

Thus, $\hat{Z}x(t)$ satisfies the system :

(4.4.26)
$$\hat{Z}x(t) = \sum_{i=0}^{m} \int_{t_0}^{t} \hat{A}_i(s)\hat{F}_{\hat{x}}(s-h_i)\,\hat{Z}x(s-h_i)\,ds + \int_{t_0}^{t}\int_{a}^{\tau} \hat{L}(\tau-s)\hat{F}_{\hat{x}}(s)\,\hat{Z}x(s)\,ds\,d\tau \qquad t \in I$$
$$\hat{Z}x(t_0) = 0$$
$$\hat{Z}x(t) = \delta\hat{\phi}(t) \quad \text{for } t \in [a, t_0)$$

or,

(4.4.27)
$$\dot{\hat{Z}x}(t) = \sum_{i=0}^{m} \hat{A}_i(t)\hat{F}_{\hat{x}}(t-h_i)\,\hat{Z}x(t-h_i) + \int_{t_0}^{t} \hat{L}(t-s)\hat{F}_{\hat{x}}(s)\,\hat{Z}x(s)\,ds + $$
$$+ \int_{a}^{t_0} \hat{L}(t-s)\hat{F}_{\hat{x}}(s)\,\delta\hat{\phi}(s)\,ds$$
$$\hat{Z}x(t_0) = 0$$
$$\hat{Z}x(t) = \delta\hat{\phi}(t) \quad , \text{for } t \in [a, t_0)$$

Now, from system (4.4.27) and the augmented adjoint system we obtain :

(4.4.28)
$$(\hat{Z}x(t_1), \hat{\psi}(t_1)) - (\hat{Z}x(t_0), \hat{\psi}(t_0)) = \int_{t_0}^{t_1} \frac{d}{dt}(\hat{Z}x(t), \hat{\psi}(t))\,dt =$$
$$= \int_{t_0}^{t_1} (\dot{\hat{Z}x}(t), \hat{\psi}(t))\,dt + \int_{t_0}^{t_1} (\hat{Z}x(t), \dot{\hat{\psi}}(t))\,dt =$$
$$= \sum_{i=0}^{m} \int_{t_0}^{t_1} (\hat{A}_i(t)\hat{F}_{\hat{x}}(t-h_i)\,\hat{Z}x(t-h_i), \hat{\psi}(t))\,dt +$$

$$+\int_{t_0}^{t_1}\left(\int_{t_0}^{t}\hat{L}(t-s)\hat{F}_{\hat{x}}(s)\,\hat{Z}x(s)ds\ ,\ \hat{\psi}(t)\right)dt\ +$$

$$+\int_{t_0}^{t_1}\left(\int_{a}^{t_0}\hat{L}(t-s)\hat{F}_{\hat{x}}(s)\delta\hat{\phi}(s)ds\ ,\ \hat{\psi}(t)\right)dt\ -$$

$$-\sum_{i=0}^{m}\int_{t_0}^{t_1}\left(\hat{Z}x(t)\ ,\ \hat{F}_{\hat{x}}^{T}(t)\hat{A}_i^{T}(t+h_i)\hat{\psi}(t+h_i)\right)dt\ -$$

$$-\int_{t_0}^{t_1}\left(\hat{Z}x(t)\ ,\ \hat{F}_{\hat{x}}^{T}(t)\int_{t}^{t_1}\hat{L}^{T}(s-t)\hat{\psi}(s)ds\right)dt\ =$$

$$=\sum_{i=0}^{m}\int_{t_0-h_i}^{t_0}\left(\hat{A}_i(t+h_i)\hat{F}_{\hat{x}}(t)\delta\hat{\phi}(t)\ ,\ \hat{\psi}(t+h_i)\right)dt\ +$$

$$+\int_{t_o}^{t_1}\left(\int_{a}^{t_0}\hat{L}(t-s)\hat{F}_{\hat{x}}(s)\delta\hat{\phi}(s)ds\ ,\ \hat{\psi}(t)\right)dt\ =$$

$$=\sum_{i=0}^{m}\int_{t_0-h_i}^{t_0}\left(\delta\phi(t)\ ,\ F_x^{T}(t)A_i^{T}(t+h_i)\psi(t+h_i)\right)dt\ +$$

$$+\int_{a}^{t_0}\left(\delta\phi(t)\ ,\ F_x^{T}(t)\int_{t_0}^{t_1}L^{T}(s-t)\psi(s)ds\right)dt\ \leq\ 0$$

by (4.4.16) .

Multiplying by $\epsilon>0$ and recalling that $\epsilon\delta\phi(t)=\phi^*(t)-\phi(t)$, we obtain

$$\sum_{i=0}^{m}\int_{t_0-h_i}^{t_0}\left(\phi^*(t)\ ,\ F_x^{T}(t)A_i^{T}(t+h_i)\psi(t+h_i)\right)dt\ +$$

$$+\int_{a}^{t_0}\left(\phi^*(t)\ ,\ F_x^{T}(t)\int_{t_0}^{t_1}L^{T}(s-t)\psi(s)ds\right)dt\ \leq \tag{4.4.29}$$

$$\leq\sum_{i=0}^{m}\int_{t_0-h_i}^{t_0}\left(\phi(t)\ ,\ F_x^{T}(t)A_i^{T}(t+h_i)\psi(t+h_i)\right)dt\ +$$

$$+\int_{a}^{t_0}\left(\phi(t)\ ,\ F_x^{T}(t)\int_{t_0}^{t_1}L^{T}(s-t)\psi(s)ds\right)dt$$

From (4.3.4.1) and (4.4.29), we obtain

(4.4.30) $$K(\phi^*(\cdot),u(\cdot),\psi(\cdot)) \leq K(\phi(\cdot),u(\cdot),\psi(\cdot))$$

for all $\phi^* \in \Phi$. Therefore,

(4.4.31) $$K(\phi(\cdot),u(\cdot),\psi(\cdot)) = \sup_{\phi^* \in \Phi} K(\phi^*(\cdot),u(\cdot),\psi(\cdot))$$

<u>Step 5</u> : <u>The transversality conditions</u>.

By modifying Lemmas A.4 and A.5 in the same way as Lemmas A.2 and A.3, the proof of the transversality conditions will be the same as in Step 5 of Appendix A.

Q.E.D.

<u>Remark 4.4.1</u> : All Corollaries and Remarks in Section 3.4 of Chapter 3, remain valid in the present case, with the additional necessary condition (4.4.31), where K is given by (4.3.4.1), except in the case considered in Remark 3.4.1, where we define

(4.4.32) $$K(\phi(\cdot),u(.),\psi(.)) = \sum_{i=0}^{m} \int_{t_0-h_i}^{t_0} (\phi(t), F^T_{(i)x}(t)A_i^T(t+h_i)\psi(t+h_i))dt + \int_{a}^{t_0} (\phi(t), F^T_{(m+1)x}(t)\int_{t_0}^{t_1} L^T(s-t)\psi(s)ds)dt$$

<u>Remark 4.4.2</u>: Since the class Φ is not compact, the functional $K(\phi(\cdot),u(\cdot),\psi(\cdot))$ may not take its supremum in Φ.

4.5 LINEAR TIME-OPTIMAL PROBLEM

Consider again the linear hereditary system (4.5.1) ≡ (3.5.3)

$$\dot{x}(t) = \sum_{i=0}^{m} B_i(t)x(t-h_i) + \int_a^t L_1(t-s)x(s)ds + \sum_{i=0}^{m} C_i(t)u(t-h_i) + \int_a^t L_2(t-s)u(s)ds + d(t) \tag{4.5.1}$$

on I , with $x(t_0) = x_0$, $x(t) = \phi(t)$ for $t \in [a, t_0)$

The solution of (4.5.1) is given by (3.5.6) . The adjoint system corresponding to (4.5.1) is given by (3.5.10) and its solution by (3.5.11) . The Hamiltonian functionalsreduce to (3.5.14) and the maximum principle (3.5.13), for the optimal control, indicates that the optimal control is not an explicit function of the optimal initial state function . Thus Theorem 3.5.1 ("Bang-bang" principle) still remains valid .

Now , from (4.3.4.1) and (4.5.1) we obtain

$$K(\phi(\cdot), \psi(\cdot)) = \sum_{i=0}^{m} \int_{t_0-h_i}^{t_0} (\phi(t) , B_i^T(t+h_i)\psi(t+h_i))dt + \int_a^{t_0} (\phi(t) , \int_{t_0}^{t_1} L_1^T(s-t)\psi(s)ds)dt \tag{4.5.2}$$

From part (i) of Theorem 4.4.1 and relation (4.5.2), we conclude that the optimal initial state function is also not an explicit function of the optimal control . However , since they both depend on the adjoint function ψ , given by Theorem 4.4.1 , they are not really independent functions , and they have to be determined simultaneously .

The following example demonstrates the improvement in the optimal time, achieved by the use of an optimal initial state function :

EXAMPLE 4.5.1 : A time-optimal problem , with varying initial state function .

We consider again the previous example 3.5.1 , but with the initial state function not fixed , i.e., we consider the system

$$(4.5.3) \qquad \dot{x}(t) = x(t) + x(t-\theta) + u(t) + u(t-\theta) \quad , \quad t \in [0,T]$$

with $x(t) = \phi(t)$, $t \in [-\theta, 0)$, $x(0) = x_0 = c$ and $u \in \Delta$, $u(t) \in U = [-1,1]$, $\phi \in \Phi = C([-\theta,0),G)$ where $G=[-1,1]$.

The optimization problem can be stated as follows :
Find an admissible pair $\{u, \phi\} \in \Delta x \Phi$, which transfers $x(0)=x_0$ to $x(T) = 0$ in minimum time T .

The solution of system (4.5.3) is :

$$(4.5.4) \quad x(t) = N(t)x(0) + \int_{-\theta}^{0} N(t-s-\theta)\phi(s)ds + \int_{0}^{t} N(t-s)\{u(s) + u(s-\theta)\}ds$$

where $N(t)$ was found previously to be

$$(4.5.5) \qquad N(t) = \sum_{n=0}^{[\frac{t}{\theta}]} \frac{(t-n\theta)^n}{n!} e^{t-n\theta} \quad , \quad t \in [0,T]$$
$$\qquad = 0 \quad , \quad t \in [-\theta, 0)$$

The optimal control was found to be of the form

$$(4.5.6) \qquad u(t) = \text{sgn}\,\{\psi(T)\} = \alpha \quad , \quad t \in [-\theta, T]$$

From (4.5.2) and (4.5.3) we obtain ,

$$(4.5.7) \qquad K(\phi(\cdot), \psi(\cdot)) = \int_{-\theta}^{0} \phi(t)\psi(t+\theta)dt$$

Since $N(t) > 0$, $t \in [0,T]$, it follows that $\psi(t) = N(T-t)\psi(T) \neq 0$ for $\psi(T) \neq 0$ and for all $t \in [0,T]$. Thus, the continuous function ϕ

which maximizes the functional (4.5.7), will be of the form

$$\phi(t) = \operatorname{sgn}\{\psi(t+\theta)\} = \operatorname{sgn}\{N(T-t-\theta)\psi(T)\} = \operatorname{sgn}\{\psi(T)\} = a \qquad (4.5.8)$$

Now, from (4.5.4), (4.5.6) and (4,5,7), we obtain

$$\begin{aligned} x(t) &= cN(t) + a\int_{-\theta}^{0} N(t-s-\theta)ds + 2a\int_{0}^{t} N(s)ds = \\ &= cN(t) + a\int_{t-\theta}^{t} N(s)ds + 2a\int_{0}^{t} N(s)ds \end{aligned} \qquad (4.5.9)$$

and ,

$$\begin{aligned} \dot{x}(t) &= c\dot{N}(t) + aN(t) - aN(t-\theta) + 2aN(t) = \\ &= cN(t) + cN(t-\theta) + 3aN(t) - aN(t-\theta) = \\ &= (c+3a)N(t) + (c-a)N(t-\theta) \end{aligned} \qquad (4.5.10)$$

Let again $c = 0.5$, and consider the equation $x(t) = 0$. Then, $T = \inf\{ t \geq 0 \; ; \; x(t) = 0 \}$. Consider now the following two cases :

(i) $\underline{a = 1}$

From (4.5.10),

$$\dot{x}(t) = 3.5N(t) - 0.5N(t-\theta) > 0 \qquad (4.5.11)$$

since $N(t)$ is monotonically increasing .

Fig. 4.5.1

Thus, in this case, the state $x_0 = 0.5$ cannot be brought to the origin .

(ii) $\underline{a = -1}$

In this case ,

(4.5.12) $\dot{x}(t) = -2.5N(t) + 1.5N(t-\theta) = -0.5(5N(t)-3N(t-\theta)) < 0 \quad , t > 0$

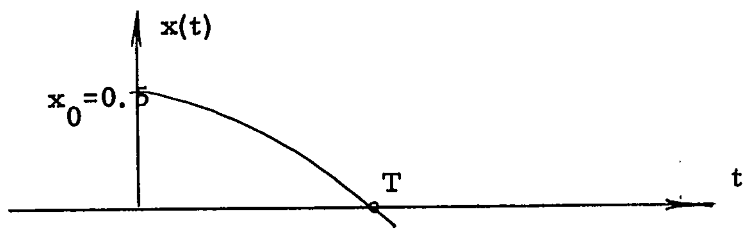

Fig. 4.5.2

Hence , the initial state $x_0 = 0.5$ is controllable for

$$u(t) = -1 \quad , t \in [0, T] , \quad \phi(t) = -1 \ , t \in [-\theta, 0)$$

and these are the optimal functions for the problem .

Now , we shall compare the optimal times of Example 3.5.1 , where $\phi(t) = 0.5$, $t \in [-\theta, 0)$, and the present one , where $\phi(t) = \phi_{opt}(t) = -1$, $t \in [-\theta , 0)$. In both cases , the optimal control is $u(t) = -1$, $t \in [-\theta, T]$. Let $x_{0.5}(t)$ and $x_{-1}(t)$ be the corresponding trajectories , starting from the same point $x_0 = 0.5$. Then ,

(4.5.13) $$x_{0.5}(t) = 0.5N(t) + 0.5\int_{t-\theta}^{t} N(s)ds - 2\int_{0}^{t} N(s)ds$$

(4.5.14) $$x_{-1}(t) = 0.5N(t) - \int_{t-\theta}^{t} N(s)ds - 2\int_{0}^{t} N(s)ds$$

and we observe that

(4.5.15) $$x_{-1}(t) < x_{0.5}(t) \quad , \text{ for all } t > 0 \text{ and } \theta > 0$$

From (4.5.14) and the following Fig. 4.5.3, it follows that

(4.5.16) $$T_{-1} < T_{0.5}$$

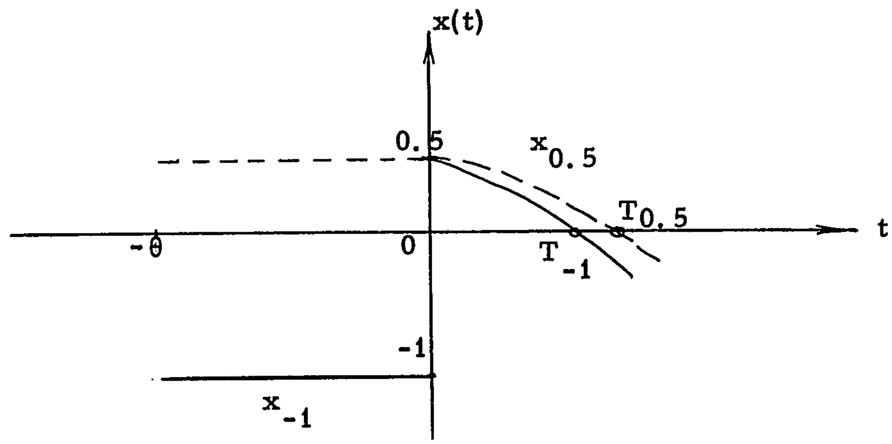

Fig. 4.5.3

Hence , an improvement of the optimal time is achieved by the use of an optimal initial state function .

The optimal time is computed as before , with the use of the Newton-Raphson method , and the results , for different values of θ, are shown in Table 4.5.1 , where the results from Table 3.5.1 are also included , for comparison purposes .

θ	0.1	0.2	0.3
$T_{0.5}$	0.374194	0.391520	0.401763
T_{-1}	0.251882	0.182322	0.182339

Table 4.5.1

From Table 4.5.1 we observe that , in the present case , the optimal time does not increase with the delay .

From (4.5.13) and (4.5.14) we see that for $t \leq \theta$, $x_{0.5}(t)$ and $x_{-1}(t)$ are independent of θ . Thus , there exist "critical" values of the delay θ , different for the two cases , such that :

$\theta_{crit.} = T \quad$, and $T = T(\theta) = T_{\theta_{crit.}} = \text{const.}$, for $\theta \geq \theta_{crit.}$.

In Fig. 4.5.4 , the functions $T_{0.5} = T_{0.5}(\theta)$ and $T_{-1} = T_{-1}(\theta)$ are plotted, with the use of Table 4.5.1 .

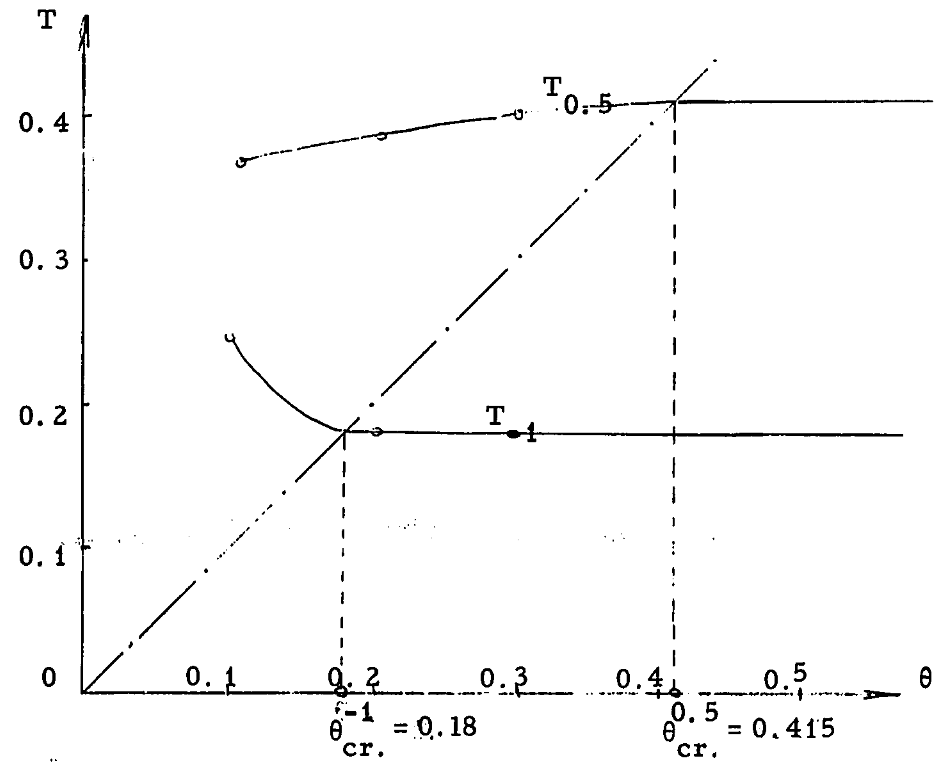

Fig. 4.5.4

4.6 DISCUSSIONS

Theorem 4.4.1 establishes necessary conditions that an optimal pair $\{ u , \phi \}$ must satisfy . However , since the class Φ of admissible initial state functions is not compact , a maximizing element for the functional K may not exist in Φ . It is left for further study to consider a compact class of admissible initial state functions , since, in that case , the proof by Pontryagin's method meets certain technical difficulties . The example 4.5.1 indicates that it would be of greater practical interest to consider a class Φ of piecewise continuous functions, instead of continuous . Indeed , this can be done , without affecting the results of Theorem 4.4.1 .

CONCLUSIONS AND SUGGESTIONS FOR FURTHER STUDY

In this thesis, a class of hereditary systems, described by a set of nonlinear integro-difference-differential equations, was considered. Following Pontryagin's geometric approach, as used also by Friedman[2], necessary conditions for the optimal control, optimal initial control, and optimal initial state function were derived. These conditions extend Pontryagin's and Friedman's Maximum Principles. It appears to the author that Pontryagin's method is very powerful, but some modifications are necessary for more complex hereditary systems.

As discussed at the end of Chapter 3, the introduction of optimal initial data $u(t)$ and $\phi(t)$, for $t \in [\alpha, t_0)$, is very important for the optimization of several systems whose initial data can be chosen from a certain class. Examples given, demonstrated the effect of the optimal choice of the initial data in the energy cost and the optimal time of simple time-delay systems. It was also found that the optimal time, as a function of the delay θ, presents a "saturation" phaenomenon for a "critical" value of the delay. We believe that this observation can be generalised, in the case of more complex linear hereditary systems. This is left for further study.

In the case of normal, linear, hereditary systems, a 'bang-bang' principle (Theorem 3.5.1) and a Uniqueness Theorem for the optimal control, and optimal initial control, were established, for a given initial state function.

The Oǧuztöreli's Kernel Matrices of the First and Second Kind, and particularly the latter one, were proven to be very useful, not only for the solution of linear hereditary systems described by integro-difference-differential equations, but also for their optimal control. We believe that an extensive study of these matrices can reveal many important features of control processes described by equations of the above type . This is another problem proposed for further study.

The problem of existence of an optimal pair , which was not considered in this thesis , may be solved by a direct extension of the results in [4] and [24] . This is also left for further study .

REFERENCES

[1] L. S. Pontryagin , V. G. Boltyanskii , R. V. Gamkrelidze and E. F. Mishchenko , " The Mathematical Theory of Optimal Processes", Interscience , New York , 1962 .

[2] A. Friedman , " Optimal Control for Hereditary Processes " , Arch. Rat. Mech. Anal. , 15, 1964 , pp. 396-419 .

[3] G. L. Kharatishvili , " A Maximum Principle in Extremal Problems with Delays " , Proc. Conf. on Math. Th. of Control, Univ. of South. Calif. , 1967 , A. V. Balakrishnan and L. W. Neustadt, eds., Acad. Press , New York, 1967 , pp. 26-34 .

[4] M. N. Oǧuztöreli , " Time - Lag Control Systems ", Acad. Press, New York , 1966 .

[5] A. Halanay , " Optimal Controls for Systems with Time-Lag " , J. S. I. A. M. in Control , Vol. 6, 1968 , pp 215-235 .

[6] E. B. Lee , " Variational Problems for Systems Having Delay in the Control Action " , I. E. E. E. Trans. on Aut. Control , Dec. 1968 , pp. 697-699 .

[7] W. R. Wells and Y. Kashiwagi , " A Time - Optimal Control Study of a Second Order Linear System with Delay " , Proc. J. A. C. C. , 1968 , pp. 644-655 .

[8] L. E. El'sgol'ts , " Introduction to the Theory of Differential Equations with Deviating Arguments " , Holden-Day, San Francisco, 1966

[9] E. B. Lee and L. Markus , " Foundations of Optimal Control Theory ", New York , Wiley , 1967 .

[10] A. Halanay , " Differential Equations - Stability, Oscillations, Time Lags. " , Acad. Press , New York , 1966 .

[11] N. H. Choksy , " Time-Delay Systems - a Bibliography " , I. R. E. Trans. Aut. Control , Vol. AC-5, pp. 66-70 , Jan. 1960 .

[12] M. R. Hestenes , " On Variational Theory and Optimal Control Theory " , J. S. I. A. M. in Control , 3, 1965 , pp. 23-48 .

[13] J. P. LaSalle , " The Time Optimum Control Problem " , Contr. Theory Nonl. Oscillations , Princeton Univ. Press , 5, 1-24, 1961 .

[14] R. Weiss , " Transportation Lag - An Annotated Bibliography " , I. R. E. Trans. Aut. Control , Vol. AC-4, pp. 56-68 , May 1959 .

[15] N. U. Ahmed , " Optimal and Adaptive Systems " , Classnotes of course EE 5401 , University of Ottawa , 1969 .

[16] K. S. Day and T. C. Hsia , " Optimal Control of Linear Time-Lag Systems " , Proc. J. A. C. C. , 1968 , pp. 1046-1055 .

[17] E. Picard , " La Mechanique Classique et ses Approximations Successives " , Rivista di Scienza, 1, 1907 , pp. 4-15 .

[18] V. Volterra , " Sur la Théorie Mathematique des Phénomènes Héréditaires " , J. Math. Pures et Appl. , 7, 1928, pp. 249-298 .

[19] L. Boltzmann , " Zur Theorie der Elastischen Nachwirkungen " , Wien. Ber. , 70-72 , 1874 , pp. 275-300 .

[20] T. Vogel , " Théorie des Systèmes Évolutifs " , Gauthier -Villars, Paris, 1965 .

[21] D. H. Chyung and E. B. Lee , " Optimal Systems with Time Delays " , Proc. 3d I. F. A. C. Conf. , London , 1966 .

[22] G. S. Glinski and N. D. Georganas , " Optimal Control for a Class of Nonlinear Systems with Finite Heredity " , Proc. 4th Annual Princeton Conference on Information Sciences and Systems , March 1970.

[23] G. S. Glinski and N. D. Georganas, " Optimal Control for a Class of Hereditary Systems ", T. R. No. 70-5, El. Eng. Dept. , U. of Ottawa, 1970.

[24] N. U. Ahmed , " Optimal Control for Hereditary Dynamical Systems " , T. R. No. 70-8 , Electr. Eng. Dept. , Univ. of Ottawa , May 1970 .

[25] S. Hartman and J. Mikusinski , " The Theory of Lebesgue Measure and Integration " , Pergamon Press ,New York, 1961.

[26] F. Riesz and B. V. Sz. -Nagy , " Functional Analysis " , translated from the 2d French edition by L. F. Boron, F. Ungar ,1955, p. 13 .

APPENDIX A

PROOF OF THEOREM 3.4.1

The proof follows the same steps as in [1 , pp. 75-114] and [2 , pp. 397-406] :

Let $u(t)$, $t \in [\alpha, t_1]$, be an optimal control in Δ , with augmented response $\hat{x}(t)$, $t \in I$.

Step 1 : Perturbation of the control function on $[\alpha, t_1]$

Let $\tau_1, \tau_2, \tau_3, \ldots, \tau_s$ be regular points of $u(t)$, $t \in [\alpha, t_1]$, such that

(A.1) $$\alpha < \tau_1 < \tau_2 < \ldots < \tau_s < t_1$$

Consider now arbitrary non-negative numbers $\delta t_1, \delta t_2, \ldots, \delta t_s$, an arbitrary real number δt and arbitrary points $v_1, v_2, \ldots, v_s$ in U.

Consider also the intervals :

(A.2) $$I_j = \{ t \in [\alpha, t_1] : \tau_j - \epsilon \delta t_j < t \leq \tau_j \}, \quad (j=1,2,\ldots,s)$$

For sufficiently small ϵ (>0) , the above intervals are mutually disjoint and belong to the basic interval $[\alpha, t_1]$.

Let us define now a control u^* , as follows :

(A.3) $$u^*(t) = \begin{cases} u(t) & , \text{ if } t \notin I_j \text{ for all } j \\ v_j & , \text{ if } t \in I_j \end{cases}$$

We shall say that the control u^* is obtained by perturbing the control u . $u^* \in \Delta$ (for ϵ small enough) .

Let $\hat{x}(t)$, $t \in I$, be the trajectory starting at $\hat{x}_0$, which corresponds to $u(t)$ (and the fixed initial state function $\hat{\phi}(t)$) and let $\hat{x}^*(t)$ be the trajectory starting at $\hat{x}_0$ and which corresponds to $u^*(t)$ (and $\hat{\phi}(t)$) .

In the following weshall make extensive use of the relation

$$\text{(A.4)} \qquad \int_{\tau+p\epsilon}^{\tau+q\epsilon} f(x(s),u(s),s)\,ds = \epsilon(q-p)f(x(\tau),u(\tau),\tau) + O(\epsilon)$$

holding for any regular point τ of u , and arbitrary real numbers p , q .

Step 2 : The basic variational formula .

We shall now show that ,

$$\text{(A.5)} \qquad \hat{x}^*(t_1 + \epsilon\,\delta t) = \hat{x}(t_1) + \epsilon\,\widehat{\Gamma x}(t_1) + O(\epsilon)$$

where $\widehat{\Gamma x}(t_1)$ is a vector which does not depend on ϵ and which is a linear function of δt , $\delta t_1, \delta t_2, \ldots, \delta t_s$:

$$\text{(A.6)} \qquad \begin{aligned} \widehat{\Gamma x}(t_1) = {} & \hat{W}(t_1)\delta t + \sum_{i=0}^{m}\sum_{k=1}^{s} N(\tau_k+h_i, t_1)\hat{A}_i(\tau_k+h_i)\{\hat{f}(\hat{x}(\tau_k), v_k) - \hat{f}(\hat{x}(\tau_k), u(\tau_k))\}\delta t_k + \\ & + \sum_{a\le\tau_k\le t_0} \left(\int_{t_0}^{t_1} N(s,t_1)ds\right)\hat{L}(t_1-\tau_k)\{\hat{f}(\hat{x}(\tau_k), v_k) - \hat{f}(\hat{x}(\tau_k), u(\tau_k))\}\delta t_k + \\ & + \sum_{t_0<\tau_k<t_1} \left(\int_{\tau_k}^{t_1} N(s,t_1)ds\right)\hat{L}(t_1-\tau_k)\{\hat{f}(\hat{x}(\tau_k), v_k) - \hat{f}(\hat{x}(\tau_k), u(\tau_k))\}\delta t_k \end{aligned}$$

where $N(s,t)$ is given by (A.21) and

$$\text{(A.7)} \qquad \hat{W}(t_1) = \sum_{i=0}^{m} \hat{A}_i(t_1)\hat{f}(\hat{x}(t_1-h_i), u(t_1-h_i)) + \int_{a}^{t_1} \hat{L}(t_1-s)\hat{f}(\hat{x}(s), u(s))\,ds$$

We shall prove the variational formula (A.5) by direct evaluation.

Indeed ,

$$\hat{x}^*(t_1+\epsilon\delta t) - \hat{x}^*(t_1) = \sum_{i=0}^{m}\int_{t_1}^{t_1+\epsilon\delta t}\hat{A}_i(s)\hat{f}(\hat{x}^*(s-h_i), u^*(s-h_i))ds \; +$$

$$+\int_{t_1}^{t_1+\epsilon\delta t}\int_{a}^{\tau}\hat{L}(\tau-s)\hat{f}(\hat{x}^*(s), u^*(s))dsd\tau \; =$$

$$= \sum_{i=0}^{m}\int_{t_1}^{t_1+\epsilon\delta t}\hat{A}_i(s)\hat{f}(\hat{x}(s-h_i), u(s-h_i))ds \; +$$

(A. 8)

$$+\int_{t_1}^{t_1+\epsilon\delta t}\int_{a}^{\tau}\hat{L}(\tau-s)\hat{f}(\hat{x}(s), u(s))dsd\tau \; + O(\epsilon) =$$

$$= \epsilon\delta t\{\sum_{i=0}^{m}\hat{A}_i(t_1)\hat{f}(\hat{x}(t_1-h_i), u(t_1-h_i)) \; +$$

$$+\int_{a}^{t_1}\hat{L}(t_1-s)\hat{f}(\hat{x}(s), u(s))ds \;\} \; + O(\epsilon) =$$

$$= \epsilon\hat{W}(t_1)\delta t \; + \; O(\epsilon)$$

In the above evaluation we made the assumption that t_1-h_i, ($i=0,1,\ldots,m$), are regular points of u , and we made use of the continuity of f in x and u .

Now ,

(A.9)

$$\hat{x}^*(t) - \hat{x}(t) = \sum_{i=0}^{m}\int_{t_0}^{t}\hat{A}_i(s)\{\hat{f}(\hat{x}^*(s-h_i), u^*(s-h_i)) - \hat{f}(\hat{x}(s-h_i), u(s-h_i))\}ds \; +$$

$$+\int_{t_0}^{t}\int_{a}^{\tau}\hat{L}(\tau-s)\{\hat{f}(\hat{x}^*(s), u^*(s)) - \hat{f}(\hat{x}(s), u(s))\}dsd\tau \; = \; J_1 + J_2$$

(A.10) $$J_1 = \sum_{i=0}^{m}\int_{t_0}^{t}\hat{A}_i(s)\{\hat{f}(\hat{x}^*(s-h_i), u^*(s-h_i)) - \hat{f}(\hat{x}(s-h_i), u^*(s-h_i))\}ds \; +$$

$$+\sum_{i=0}^{m}\int_{t_0}^{t}\hat{A}_i(s)\{\hat{f}(\hat{x}(s-h_i),u^*(s-h_i)) - \hat{f}(\hat{x}(s-h_i),u(s-h_i))\}ds = J_{11}+J_{12}$$

$$J_{11} = \sum_{i=0}^{m}\int_{t_0}^{t}\hat{A}_i(s)\{\hat{f}(\hat{x}^*(s-h_i),u^*(s-h_i))-\hat{f}(\hat{x}(s-h_i),u^*(s-h_i))\}ds =$$

(A.11)
$$= \sum_{i=0}^{m}\int_{t_0}^{t}\hat{A}_i(s)\hat{f}_{\hat{x}}(\hat{x}(s-h_i),u^*(s-h_i))\{\hat{x}^*(s-h_i)-\hat{x}(s-h_i)\}ds + O(\epsilon) =$$

(since f_x is continuous in both arguments)

$$= \sum_{i=0}^{m}\int_{t_0}^{t}\hat{A}_i(s)\hat{f}_{\hat{x}}(x(s-h_i),u(s-h_i))\{\hat{x}^*(s-h_i) - \hat{x}(s-h_i)\}ds + O(\epsilon) =$$

$$= \sum_{i=0}^{m}\int_{t_0}^{t}\hat{A}_i(s)\hat{F}_{\hat{x}}(s-h_i)\{\hat{x}^*(s-h_i) - \hat{x}(s-h_i)\}ds + O(\epsilon)$$

$$J_{12} = \sum_{i=0}^{m}\int_{t_0}^{t}\hat{A}_i(s)\{\hat{f}(\hat{x}(s-h_i),u^*(s-h_i)) - \hat{f}(\hat{x}(s-h_i),u(s-h_i))\}ds =$$

(A.12)
$$= \epsilon\sum_{i=0}^{m}\sum_{t_0-h_i\le\tau_k<t-h_i}\hat{A}_i(\tau_k+h_i)\{\hat{f}(\hat{x}(\tau_k),v_k) - \hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k + O(\epsilon) =$$

$$= \epsilon\varphi(t) + O(\epsilon)$$

where

(A.13)
$$\varphi(t) = \sum_{i=0}^{m}\sum_{t_0-h_i\le\tau_k<t-h_i}\hat{A}_i(\tau_k+h_i)\{\hat{f}(\hat{x}(\tau_k),v_k)-\hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k$$

$$J_2 = \int_{t_0}^{t}\int_{a}^{\tau}\hat{L}(\tau-s)\{\hat{f}(x^*(s),u^*(s))-\hat{f}(\hat{x}(s),u^*(s))\}ds\,d\tau \quad +$$

(A.14)
$$+\int_{t_0}^{t}\int_{a}^{\tau}\hat{L}(\tau-s)\{\hat{f}(\hat{x}(s),u^*(s))- \hat{f}(\hat{x}(s),u(s))\}ds\,d\tau = J_{21}+J_{22}$$

$$J_{21} = \int_{t_0}^{t} \int_{\alpha}^{\tau} \hat{L}(\tau-s)\{\hat{f}(x^*(s),u^*(s)) - \hat{f}(\hat{x}(s),\overset{*}{u}(s))\}\,ds\,d\tau =$$

(A.15)

$$= \int_{t_0}^{t} \int_{\alpha}^{\tau} \hat{L}(\tau-s)\hat{F}_{\hat{x}}(s)\{\hat{x}^*(s) - \hat{x}(s)\}\,ds\,d\tau + O(\epsilon)$$

$$J_{22} = \int_{t_0}^{t} \int_{\alpha}^{\tau} \hat{L}(\tau-s)\{\hat{f}(\hat{x}(s),u^*(s)) - \hat{f}(\hat{x}(s),u(s))\}\,ds\,d\tau =$$

(A.16) $$= \epsilon \int_{t_0}^{t} \sum_{\alpha \le \tau_k < \tau} \hat{L}(\tau-\tau_k)\{\hat{f}(x(\tau_k),v_k) - \hat{f}(x(\tau_k),u(\tau_k))\}\,\delta t_k\,d\tau + O(\epsilon) =$$

$$= \epsilon g(t) + O(\epsilon)$$

where

(A.17) $$g(t) = \int_{t_0}^{t} \sum_{\alpha \le \tau_k < \tau} \hat{L}(\tau-\tau_k)\{\hat{f}(\hat{x}(\tau_k),v_k) - \hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\,\delta t_k\,d\tau$$

Now, from (A.9) ,

(A.18) $$\hat{x}^*(t) - \hat{x}(t) = J_{11} + J_{12} + J_{21} + J_{22}$$

and by setting $\hat{z}(t) \triangleq \hat{x}^*(t) - \hat{x}(t)$,

$$\hat{z}(t) = \sum_{i=0}^{m} \int_{t_0}^{t} \hat{A}_i(s)\hat{F}_{\hat{x}}(s-h_i)\hat{z}(s-h_i)\,ds + \int_{t_0}^{t} \int_{\alpha}^{\tau} \hat{L}(\tau-s)\hat{F}_{\hat{x}}(s)\hat{z}(s)\,ds\,d\tau +$$

(A.19)

$$+ \epsilon\{\varphi(t) + g(t)\} + O(\epsilon) \quad , \quad t \in I$$

$$\hat{z}(t) = 0 \quad , \; t \in [\alpha, t_0]$$

The solution of (A.19) is given from (B.1.11) as :

(A.20) $$\hat{z}(t) = \epsilon\int_{t_0}^{t} N(s,t)\ \{\dot{\varphi}(s) + \dot{g}(s)\}\, ds \ + O(\epsilon)$$

where N(s,t) is Oǧuztöreli's [4] Kernel Matrix of the Second Kind, satisfying the system

(A.21) $$\frac{\partial N(s,t)}{\partial t} = \sum_{i=0}^{m} \hat{A}_i(t)\hat{F}_{\hat{x}}(t-h_i)N(s,t-h_i) + \int_{t_0}^{t} \hat{L}(t-\tau)\hat{F}_{\hat{x}}(\tau)N(s,\tau)d\tau$$

$$N(s,t) = 0 \ \text{ for } \ t< s \ , \text{ and } \ N(t,t) = I$$

We define now the vector

(A.22) $$\hat{\Delta x}(t) = \int_{t_0}^{t} N(s,t)\{\dot{\varphi}(s) + \dot{g}(s)\}ds$$

From (A.17) :

(A.23) $$\dot{g}(t) = \sum_{a\leq \tau_k < t} \hat{L}(t-\tau_k)\{\hat{f}(\hat{x}(\tau_k),v_k) - \hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k$$

and from (A.13)

(A.24) $$\varphi(t) = \sum_{i=0}^{m} \varphi_i(t)$$

(A.25) $$\varphi_i(t) = \sum_{t_0-h_i\leq \tau_k < t-h_i} \hat{A}_i(\tau_k+h_i)\{\hat{f}(\hat{x}(\tau_k),v_k) - \hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k =$$

$$= \sum_{k=1}^{s} H_i(\tau_k,t)\{\hat{f}(\hat{x}(\tau_k),v_k) - \hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k$$

where

(A.26) $$H_i(\tau_k,t) = \begin{cases} 1 & \text{if } \ \tau_k \in [t_0-h_i, t-h_i) \\ 0 & \text{if } \ \tau_k \notin [t_0-h_i, t-h_i) \end{cases}, \quad t\in[t_0,\infty)$$

Therefore

$$\text{(A. 27)} \qquad \dot{\varphi}_i(t) = \sum_{k=1}^{s} \delta(t-(\tau_k+h_i))\hat{A}_i(\tau_k+h_i)\{\hat{f}(\hat{x}(\tau_k),v_k)-\hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k$$

and from (A. 22), (A. 23) and (A. 27), we obtain

$$\hat{\Delta}x(t) = \sum_{i=0}^{m}\sum_{k=1}^{s} N(\tau_k+h_i,t)\hat{A}_i(\tau_k+h_i)\{\hat{f}(\hat{x}(\tau_k),v_k)-\hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k +$$

$$+\int_{t_0}^{t} \sum_{a\le \tau_k<\tau} N(\tau,t)\hat{L}(t-\tau_k)\{\hat{f}(\hat{x}(\tau_k),v_k)-\hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k d\tau =$$

$$\text{(A. 28)} \qquad = \sum_{i=0}^{m}\sum_{k=1}^{s} N(\tau_k+h_i,t)\hat{A}_i(\tau_k+h_i)\{\hat{f}(\hat{x}(\tau_k),v_k)-\hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k +$$

$$+\sum_{a\le\tau_k\le t_0}\left(\int_{t_0}^{t} N(\tau,t)d\tau\right)\hat{L}(t-\tau_k)\{\hat{f}(\hat{x}(\tau_k),v_k)-\hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k +$$

$$+\sum_{t_0<\tau_k<t}\left(\int_{\tau_k}^{t} N(\tau,t)d\tau\right)\hat{L}(t-\tau_k)\{\hat{f}(\hat{x}(\tau_k),v_k)-\hat{f}(\hat{x}(\tau_k),u(\tau_k))\}\delta t_k$$

We observe that $\hat{\Delta}x(t)$ does not depend on ϵ and that it is a linear function of $\delta t_1, \delta t_2, \ldots, \delta t_s$.

Defining now ,

$$\text{(A. 29)} \qquad \hat{\Gamma}x(t_1) = \hat{W}(t_1)\delta t + \hat{\Delta}x(t_1)$$

we obtain from (A. 8) , (A. 20) , (A. 28) and (A. 29)

$$\text{(A. 30)} \qquad \hat{x}^*(t_1+\epsilon\delta t) - \hat{x}(t_1) = \epsilon\hat{\Gamma}x(t_1) + O(\epsilon)$$

and the basic variational formula is proven .

<u>Remark A.1</u> : For any other choice of regular points $\tau_1, \tau_2, \ldots, \tau_s$ on $[\alpha, t_1]$, the variational formula remains valid with minor modifications in the proof .

The proof now follows exactly the steps of Pontryagin's approach [1, p.92-]. Namely :

(i) We define the <u>complex perturbation</u>

(A.31) $$a = \{ \tau_i, v_i, \delta t_i, \delta t \quad ; i=1,2,\ldots,s \}$$

and write $\hat{\Gamma x}_a = \hat{\Gamma x}(t_1)$ to emphasize its dependence on a .

(ii) If we consider two complex perturbations

(A.32) $$\begin{aligned} a' &= \{ \tau_i, v_i, \delta t'_i, \delta t' \quad ; i=1,2,\ldots,s\} \\ a'' &= \{ \tau_i, v_i, \delta t''_i, \delta t'' \quad ; i=1,2,\ldots,s\} \end{aligned}$$

we define the linear combinations

(A.33) $$\lambda'a' + \lambda''a'' = \{ \tau_i, v_i, \lambda'\delta t'_i + \lambda''\delta t''_i, \lambda'\delta t' + \lambda''\delta t'' \; ; i=1,2,\ldots,s \}$$

for $\lambda', \lambda'' \in R^+ = [0, \infty)$, and construct the space of perturbations

(A.34) $$A = \{ a : a \text{ is a complex perturbation } \}$$

(iii) The vectors $\hat{\Gamma x}_a$, $a \in A$, emanating from the point $\hat{x}(t_1)$ form the set

(A.35) $$\hat{K}_{t_1} = \{ \hat{\Gamma x}_a : a \in A \}$$

(iv) $\hat{K}_{t_1}$ is a convex cone in E^{n+1}. (Cone of attainability)

Step 3 : Fundamental Lemmas

Lemma A.1

Let $\hat{\psi}(t)$ be an arbitrary, nontrivial, solution of the augmented adjoint system (3.3.4.4) and let $\hat{\delta x}(t)$ be an arbitrary solution of the augmented variational system (3.3.4.3) . Then

(A.36) $$(\hat{\psi}(t_1), \hat{\delta x}(t_1)) = (\hat{\psi}(t_0), \hat{\delta x}(t_0))$$

where $(\hat{\psi}(t), \hat{\delta x}(t)) = \sum_{i=0}^{n} \psi_i(t)\delta x^i(t)$

Proof :

The systems (3.3.4.3) and (3.3.4.4) are adjoint by construction . Thus :

$$(\hat{\psi}(t_1), \hat{\delta x}(t_1)) - (\hat{\psi}(t_0), \hat{\delta x}(t_0)) = \int_{t_0}^{t_1} \frac{d}{dt}(\hat{\psi}(t), \hat{\delta x}(t))dt =$$

$$= \int_{t_0}^{t_1} (\dot{\hat{\psi}}(t), \hat{\delta x}(t))dt + \int_{t_0}^{t_1} (\hat{\psi}(t), \dot{\hat{\delta x}}(t))dt = 0$$

Q.E.D.

Remark A.2 : For hereditary processes, the inner product $(\hat{\psi}(t), \hat{\delta x}(t))$ is not constant on $[t_0, t_1]$, in contrast to simple dynamical systems (s. Lemma 1 in [1, p. 86]).

Lemma A.2

If Λ is any smooth curve which begins at $\hat{x}(t_1)=\hat{x}_1$ and which has a tangent ray L at this point, pointing into the interior of the cone $\hat{K}_{t_1}$, then there exists a control u_* in Δ such that the corresponding trajectory $\hat{x}_*(t)$, $t \in I$, starting at $\hat{x}_0$, passes through a point of Λ distinct from $\hat{x}_1$.

Proof :

With the use of the variational formula (A. 5), the proof is exactly the same as of Lemma 3 in [1 , pp. 94-98] .

Lemma A. 3

If u is an optimal control in Δ with augmented response $\hat{x}(t)$, then the ray L, in E^{n+1}, starting at the point $\hat{x}(t_1)$ and pointing in the direction of the negative x^0 axis, does not belong to the interior of the cone $\hat{K}_{t_1}$.

Proof :

Suppose that $L \subset \overset{o}{\hat{K}}_{t_1}$ (interior of $\hat{K}_{t_1}$). Then, by lemma A. 2 and for $\Lambda \equiv L$, there exists a control u_* in Δ, such that the corresponding trajectory $\hat{x}_*(t)$ passes through a point of L different from $\hat{x}(t_1)$, i.e.,

$$\text{(A. 38)} \qquad \begin{aligned} x_*^i(t'_1) &= x^i(t_1) , \quad (i=1,2,\dots,n) \\ x_*^0(t'_1) &< x^0(t_1) = J[u] \end{aligned}$$

But this contradicts the assumption that u is optimal. Thus $L \not\subset \overset{o}{\hat{K}}_{t_1}$.

Q. E. D.

From Lemma A. 3 , we obtain the following

<u>Corollary A. 1</u>

There exists a hyperplane of support to $\hat{K}_{t_1}$ at its vertex, i. e., a hyperplane such that $\hat{K}_{t_1}$ lies entirely in one of the two closed half-spaces defined by it .

Consequently, there exists a non-zerovector $\alpha=(\alpha^0, \alpha^1, \ldots, \alpha^n)$ such that

(A. 39) $(\alpha, L) \geq 0$ and

(A. 40) $(\alpha, \hat{\Gamma x}_a) \leq 0$ for all $\hat{\Gamma x}_a \in \hat{K}_{t_1}$

Let $\hat{\psi}(t)$ be the solution of the augmented adjoint system (3. 3. 4. 4) with terminal value $\hat{\psi}(t_1) = \psi_1 = \alpha$. Then

(A. 41) $(\hat{\psi}(t_1), \hat{\Gamma x}_a) \leq 0$ for all $\hat{\Gamma x}_a \in \hat{K}_{t_1}$

Since we may take $L = (-1, 0, 0, \ldots, 0)$, we obtain from (A. 39)

(A. 42) $\psi_0(t_1) \leq 0$

Moreover , since f does not depend explicitly on x^0 , $\frac{\partial f^i}{\partial x^0} = 0$, for all $i=0,1,\ldots,n$. Thus ,

(A. 43) $\frac{d\psi_0}{dt} = 0$

and , from (A. 42), (A. 43) ,

(A. 44) $\psi_0(t) = \text{const.} \leq 0$, for all $t \in [t_0, t_1]$

Step 4 : The Maximum Principle .

Let now τ_1 be a regular point of $u(t)$, $t \in [a, t_1]$, and let v_1 be an arbitrary point in U . Consider the elementary perturbation

(A.45) $$a_1 = \{ \tau_1, v_1, \delta t_1, \delta t \} = \{ \tau_1, v_1, 1, 0 \}$$

Then , from (A.29) ,

(A.46) $$\hat{\Gamma x}_a = \hat{\Delta x}_a$$

Since $z(t)$ is the unique solution of (A.19) and

(A.47) $$\hat{\Delta x}(t) = \lim_{\epsilon \to 0} \frac{\hat{z}(t)}{\epsilon}$$

$\hat{\Delta x}(t)$ satisfies the system :

(A.48)
$$\hat{\Delta x}(t) = \sum_{i=0}^{m} \int_{t_0}^{t} \hat{A}_i(s) \hat{F}_{\hat{x}}(s-h_i) \hat{\Delta x}(s-h_i) ds + \int_{t_0}^{t} \int_{a}^{\tau} \hat{L}(\tau - s) \hat{F}_{\hat{x}}(s) \hat{\Delta x}(s) ds d\tau + \{ \varphi(t) + g(t) \} , \quad t \in I$$

$$\hat{\Delta x}(t) = 0 , \quad t \in [a, t_0]$$

or,

(A.49)
$$\dot{\hat{\Delta x}}(t) = \sum_{i=0}^{m} \hat{A}_i(t) \hat{F}_{\hat{x}}(t-h_i) \hat{\Delta x}(t-h_i) + \int_{a}^{t} \hat{L}(t-s) \hat{F}_{\hat{x}}(s) \hat{\Delta x}(s) ds + \dot{\varphi}(t) + \dot{g}(t) , \quad t \in I$$

$$\hat{\Delta x}(t) = 0 , \quad t \in [a, t_0]$$

We consider now two cases :

(i) $\tau_1 \in [t_0, t_1]$

Now ,

(A.50) $g(t) = 0$, for $t \in [t_0, \tau_1]$

(A.51) $\varphi(t) = 0$, for $t \in [t_0, \tau_1]$

Thus, from (A.48) ,

(A.52) $\hat{\Delta x}(t) = 0$, for $t \in [\alpha, \tau_1]$

Now, from (A.23), (A.24) and (A.27) ,

(A.53) $$\dot{\varphi}(t) = \sum_{i=0}^{m} \delta(t-(\tau_1+h_i))\hat{A}_i(\tau_1+h_i)\{\hat{f}(\hat{x}(\tau_1), v_1) - \hat{f}(\hat{x}(\tau_1), u(\tau_1))\}$$

and

(A.54) $$\dot{g}(t) = \hat{L}(t-\tau_1)\{\hat{f}(\hat{x}(\tau_1), v_1) - \hat{f}(\hat{x}(\tau_1), u(\tau_1))\} \qquad , t \in I$$

From the augmented adjoint system (3.3: 4.4) , with $\hat{\psi}_1 = \alpha$, and (A.49) we obtain :

$$(\hat{\Delta x}(t_1), \hat{\psi}(t_1)) - (\hat{\Delta x}(\tau_1), \hat{\psi}(\tau_1)) = \int_{\tau_1}^{t_1} \frac{d}{dt}(\hat{\Delta x}(t), \hat{\psi}(t))dt =$$

(A.55) $$= \int_{\tau_1}^{t_1} (\dot{\hat{\Delta x}}(t), \hat{\psi}(t))dt + \int_{\tau_1}^{t_1} (\hat{\Delta x}(t), \dot{\hat{\psi}}(t))dt =$$

$$= \int_{\tau_1}^{t_1} (\dot{\varphi}(t), \hat{\psi}(t))dt + \int_{\tau_1}^{t_1} (\dot{g}(t), \hat{\psi}(t))dt$$

From (A.41), (A.46) , (A.52) and (A.55) , we obtain now :

$$(A.56) \qquad \int_{\tau_1}^{t_1} (\dot{\varphi}(t), \hat{\psi}(t))dt + \int_{\tau_1}^{t_1} (\dot{g}(t), \hat{\psi}(t))dt \leq 0$$

and , by using (A.53), (A.54),

$$(A.57) \qquad \sum_{i=0}^{m} (\hat{A}_i(\tau_1+h_i)\hat{f}(\hat{x}(\tau_1), v_1), \hat{\psi}(\tau_1+h_i)) + \int_{\tau_1}^{t_1} (\hat{L}(s-\tau_1)\hat{f}(\hat{x}(\tau_1), v_1), \hat{\psi}(s))ds \leq$$

$$\leq \sum_{i=0}^{m} (\hat{A}_i(\tau_1+h_i)\hat{f}(\hat{x}(\tau_1), u(\tau_1)), \hat{\psi}(\tau_1+h_i)) + \int_{\tau_1}^{t_1} (\hat{L}(s-\tau_1)\hat{f}(\hat{x}(\tau_1), u(\tau_1)), \hat{\psi}(s))ds$$

or, by (3.3.4.5),

$$(A.58) \qquad \hat{H}(\hat{x}(\tau_1), \hat{\psi}(\cdot), v_1, \tau_1) \leq \hat{H}(\hat{x}(\tau_1), \hat{\psi}(\cdot), u(\tau_1), \tau_1)$$

This relation holds for arbitrary $v_1 \in U$. Thus,

$$(A.59) \qquad \hat{H}(\hat{x}(\tau_1), \hat{\psi}(\cdot), u(\tau_1), \tau_1) = \max_{v \in U} \hat{H}(\hat{x}(\tau_1), \hat{\psi}(\cdot), v, \tau_1)$$

But τ_1 is an arbitrary regular point of u in $[t_0, t_1]$, and almost allpoints in $[t_0, t_1]$ are regular points of u. Thus,

$$(A.60) \qquad \hat{H}(\hat{x}(t), \hat{\psi}(\cdot), u(t), t) = \max_{v \in U} \hat{H}(\hat{x}(t), \hat{\psi}(\cdot), v, t) \qquad \text{a.e. on } [t_0, t_1]$$

(ii) $\tau_1 \in [a, t_0)$

Now, from (A.53), (A.54), (3.3.4.4) and (A.49), we obtain,

$$(\hat{\Delta x}(t_1), \hat{\psi}(t_1)) - (\hat{\Delta x}(t_0), \hat{\psi}(t_0)) = \int_{t_0}^{t_1} \frac{d}{dt} (\hat{\Delta x}(t), \hat{\psi}(t)) dt =$$

$$= \int_{t_0}^{t_1} (\dot{\hat{\Delta x}}(t), \hat{\psi}(t)) dt + \int_{t_0}^{t_1} (\hat{\Delta x}(t), \dot{\hat{\psi}}(t)) dt = \tag{A.61}$$

$$= \int_{t_0}^{t_1} ((\dot{\varphi}(t) + \dot{g}(t)) , \hat{\psi}(t)) dt$$

Using the same arguments as in case (i) , we obtain

$$\int_{t_0}^{t_1} ((\dot{\varphi}(t) + \dot{g}(t)) , \hat{\psi}(t)) dt \leq 0 \tag{A.62}$$

Since ,

$$\int_{t_0}^{t_1} \delta(t - (\tau_1 + h_i)) dt = H_i(\tau_1, t_1) = \theta_i(\tau_1) \tag{A.63}$$

where H_i is given by (A.26) and $\theta_i(t)$ is given in subsection 3.3.3, Chapter 3, we finally obtain :

$$\begin{aligned} &\sum_{i=0}^{m} \theta_i(\tau_1) (\hat{A}_i(\tau_1 + h_i) \hat{f}(\hat{x}(\tau_1), v_1), \hat{\psi}(\tau_1 + h_i)) + \int_{t_0}^{t_1} (\hat{L}(s - \tau_1) \hat{f}(\hat{x}(\tau_1), v_1), \hat{\psi}(s)) ds \leq \\ &\leq \sum_{i=0}^{m} \theta_i(\tau_1) (\hat{A}_i(\tau_1 + h_i) \hat{f}(\hat{x}(\tau_1), u(\tau_1)), \hat{\psi}(\tau_1 + h_i)) + \int_{t_0}^{t_1} (\hat{L}(s - \tau_1) \hat{f}(\hat{x}(\tau_1), u(\tau_1)), \hat{\psi}(s)) ds \end{aligned} \tag{A.64}$$

or, by (3.3.4.5) ,

$$\hat{H}_p(\hat{\phi}(\tau_1), \hat{\psi}(\cdot), v_1, \tau_1) \leq \hat{H}_p(\hat{\phi}(\tau_1), \hat{\psi}(\cdot), u(\tau_1), \tau_1) \tag{A.65}$$

for any regular point $\tau_1 \in [a, t_0)$, and any $v_1 \in U$. Thus ,

$$\hat{H}_p(\hat{\phi}(t), \hat{\psi}(.), u(t), t) = \max_{v \in U} \hat{H}_p(\hat{\phi}(t), \hat{\psi}(\cdot), v, t) \quad \text{a.e. on } [a, t_0) \tag{A.66}$$

Consider now the elementary perturbation

(A.67) $$a_2 = \{ \tau_1, v_1, \delta t_1, \delta t \} = \{ \tau_1, v_1, 0, \delta t \}$$

Then,

(A.68) $$\hat{\Delta x}(t_1) = 0$$

From (A.41), we obtain

(A.69) $$(\hat{\psi}(t_1) , \hat{\Delta x}(t_1) + \hat{W}(t_1)\delta t) \leq 0$$

or,

(A.70) $$(\hat{\psi}(t_1) , \hat{W}(t_1))\delta t \leq 0$$

But δt is an arbitrary real number. Thus,

(A.71) $$(\hat{\psi}(t_1), \hat{W}(t_1)) = 0$$

and, by (3.3.4.6) and (A.7):

(A.72) $$\hat{R}(t_1) = 0$$

<u>Step 5</u> : <u>The transversality conditions.</u>

If, in addition to the control perturbation on $[a, t_1]$, we consider a perturbation of the initial state:

(A.73) $$x^*(t_0) = x(t_0) + \epsilon \xi + O(\epsilon)$$

where ξ is a constant vector in E^n, then, following the same procedure as in Step 1, we derive the analogous to (A.5) variational formula

(A.74) $$\hat{x}^*(t_1 + \epsilon \delta t) = \hat{x}(t_1) + \epsilon \{ \hat{\Gamma x}(t_1) + \hat{\mu}(t_1) \} + O(\epsilon)$$

where the vector $\hat{\mu}(t)$ is the solution of the system:

(A.75)

$$\dot{\hat{\mu}}(t) = \sum_{i=0}^{m} \hat{A}_i(t)\hat{F}_{\hat{x}}(t-h_i)\hat{\mu}(t-h_i) + \int_a^t \hat{L}(t-s)\hat{F}_{\hat{x}}(s)\hat{\mu}(s)ds$$

$$\hat{\mu}(t_0) = \hat{\xi} = (0, \xi)^T$$

$$\hat{\mu}(t) = 0 \quad , \text{ for } t \in [a, t_0)$$

(Observe that system (A.75) and the augmented variational system (3.3.4.3) are identical .)

Indeed , if we proceed as in Step 1 , we derive a system analogous to (A.19) :

(A.76)

$$\hat{z}(t) = \sum_{i=0}^{m} \int_{t_0}^{t} \hat{A}_i(s)\hat{F}_{\hat{x}}(s-h_i)\hat{z}(s-h_i)ds + \int_{t_0}^{t}\int_{a}^{\tau} \hat{L}(\tau-s)\hat{F}_{\hat{x}}(s)\hat{z}(s)ds\,d\tau +$$

$$+ \in \{\hat{\xi} + \varphi(t) + g(t)\} + O(\in) \quad , \quad t \in I$$

$$\hat{z}(t) = 0 \, , \quad \text{for } t \in [a, t_0)$$

from which (A.75) and (A.74) are derived in the same way as (A.29) and (A.5) from (A.19) .

Now , let S_0 be a smooth r_0-dimensional ($r_0 < n$) manifold in E^n , which passes through the point x_0 , and let T_0 be the tangent plane of S_0 at that point . Let $\hat{T}_0$ be the hyperplane in E^{n+1} consisting of all points of the form $(0,x)^T$ with $x \in T_0$. It is obvious that $\hat{T}_0$ passes through $\hat{x}_0$.

If $N(s,t)$ is the Kernel Matrix of the Second Kind associated with the system (A.75) then , from (B.1.11) , the solution of (A.75) is

(A.77)
$$\hat{\mu}(t) = N(t_0, t)\,\hat{\xi} \qquad , \; t \in I$$

and the set

(A.78) $$\hat{T}_0^{\,1} = \{ \hat{\mu}(t_1) : \hat{\xi} \in \hat{T}_0 \}$$

is a hyperplane passing through $\hat{x}(t_1)$.

Let $\hat{K}_{t_1}'$ be the convex hull of the set $\hat{T}_0^{\,1} \cup \hat{K}_{t_1}$. Obviously $\hat{K}'_{t_1}$ is a convex cone with vertex at $\hat{x}(t_1)$.

The following Lemma is an extension of Lemma A.2 , and its proof is completely analogous to the one of Lemma 10 in [1, p. 108] :

Lemma A.4

Let t_1 be a regular point for the control u(t) , $t \in [a, t_1]$, and let $\hat{x}(t)$ be the trajectory which corresponds to u(t) and starts at the point $\hat{x}_0$. Furthermore , let Λ be a manifold in E^{n+1} , whose dimension does not exceed n . Suppose that Λ has an edge , and that $\hat{x}(t_1)$ lies on this edge . Let M be the half-plane tangent to Λ at $\hat{x}(t_1)$. If the cones $\hat{K}'_{t_1}$ and M (which have a common vertex at $\hat{x}(t_1)$) are not separated , there exists a control $u_*(t)$ and a point $x_0^* \in S_0$, such that the trajectory $\hat{x}_*(t)$ which corresponds to u_* and $\hat{x}_0^* = (0, x_0^*)^T$ passes through a point of Λ not on the edge of Λ .

Now, let $u(t), t \in [a, t_1]$, be an optimal control in Δ and let $\hat{x}(t)$ be the corresponding optimal trajectory , which yields the solution of the problem with variable endpoints . Let $\hat{x}(t_0) = \hat{x}_0$ and $\hat{x}(t_1) = \hat{x}_1$. Let T_1 be the tangent plane of the r_1-dimensional manifold S_1 at the point x_1, and let $\hat{T}_1$ be the r_1-dimensional plane in E^{n+1} consisting of all points of the form $(x^0(t_1), x)$, with $x \in T_1$. Let Q be the point set generated by all the rays passing through each point of $\hat{T}_1$ in the direction of the negative x^0 axis . The set Q is an (r_1+1)-dimensional half-plane , and its boundary points belong to $\hat{T}_1$.

<u>Lemma A.5</u>

The cones $\hat{K}'_{t_1}$ and Q, which have a common vertex at $\hat{x}(t_1)$, are separated.

<u>Proof</u> :

Let Λ_{t_1} be the manifold with an edge, consisting of all points $(x^0, x) \in E^{n+1}$ for which $x^0 \leq x^0(t_1)$ and $x \in S_1$. Then the tangent half-plane of Λ_{t_1} at $\hat{x}(t_1)$ coincides with Q.

Suppose now that $\hat{K}'_{t_1}$ and Q are not separated. By Lemma A.4 there exists a control $u_*(t)$ such that the corresponding trajectory $\hat{x}_*(t)$ starting at the point $\hat{x}^*_0 = (0, x^*_0)$ (where $x^*_0 \in S_0$), passes through a point of Λ_{t_1} not on its edge. In other words, there exists a $t' > t_0$ and a point $\hat{\eta} \in \Lambda_{t_1}$, not on the edge of Λ_{t_1}, such that :

(A.79) $$\hat{x}_*(t') = \hat{\eta}$$

But $\hat{\eta}$ has the form $\hat{\eta} = (\eta^0, \eta)$ where $\eta \in S_1$. In addition, since $\hat{\eta}$ is not on the edge of Λ_{t_1}, $\eta^0 < x^0(t_1)$, i.e.,

(A.80) $$\begin{aligned} x_*^i(t') &= \eta^i, \quad (i=1,2,\ldots,n) \\ x_*^0(t') &= \eta^0 < x^0(t_1) = J[u] \end{aligned}$$

This is a contradiction to the optimality of u and $\hat{x}$, and the Lemma is proven.

Q.E.D.

Since the cones $\hat{K}'_{t_1}$ and Q are separated, there exists a nonzero vector

(A.81) $$\hat{c} = (c^0, c^1, \ldots, c^n)$$

such that $\hat{K}'_{t_1}$ (and hence $\hat{K}_{t_1} \subset \hat{K}'_{t_1}$) is in the half-space :

(A.82) $\quad (\hat{c}, \hat{y}) \leq 0$

where $\hat{y}$ are vectors emanating from $\hat{x}(t_1)$.

In particular, the ray L_{t_1} (in Q) is in the half-space $(\hat{c}, \hat{y}) \geq 0$.

Thus, the solution of the augmented adjoint system (3.3.4.4) with $\hat{\psi}(t_1) = \hat{\psi}_1 = \hat{c}$, satisfies the conditions (i),(ii), in Theorem 3.4.1.

We shall show now that $\psi(t)$ satisfies the transversality conditions at both endpoints of $x(t)$.

The plane $\hat{T}_1 \subset Q$ lies entirely in the half-space $(\hat{c}, \hat{y}) \geq 0$ and, consequently, in the hyperplane $(\hat{c}, \hat{y}) = 0$ or, equivalently

(A.83) $\quad (\hat{\psi}(t_1), \hat{y}) = 0$

Now, if $\eta = (\eta^1, \eta^2, \ldots, \eta^n)$ is an arbitrary tangent vector of S_1 at $x(t_1)$, i.e., $\eta \in T_1$, then the vector $\hat{\eta} = (0, \eta) \in \hat{T}_1$ and consequently

(A.84) $\quad (\hat{\psi}(t_1), \hat{\eta}) = 0$

This leads to

(A.85) $\quad (\psi(t_1), \eta) = 0$

which proves the transversality condition at the right-hand endpoint.

Furthermore, since $\hat{T}_0^1 \subset \hat{K}'_{t_1}$, the plane $\hat{T}_0^1$ lies entirely in the half-space $(\hat{c}, \hat{y}) \leq 0$ and consequently in the hyperplane $(\hat{c}, \hat{y}) = 0$

(A.86) $\quad (\hat{\psi}(t_1), \hat{y}) = 0$

That is, for every vector $\hat{\xi} \in \hat{T}_0$, the vector $\hat{\mu}(t_1) = N(t_0, t_1)\hat{\xi}$ lies on the hyperplane (A.86) :

(A.87) $\quad (\hat{\psi}(t_1), \hat{\mu}(t_1)) = 0$

But system (A. 75) is identical to the augmented variational system (3. 3. 4. 3) for which Lemma A. 1 holds . Thus :

(A. 88) $$(\hat{\psi}(t_0) , \hat{\mu}(t_0)) = (\hat{\psi}(t_1) , \hat{\mu}(t_1)) = 0$$

or,

(A. 89) $$(\hat{\psi}(t_0) , \hat{\xi}) = 0$$

from which

(A. 90) $$(\psi(t_0) , \xi) = 0$$

and the transversality condition at the left-hand endpoint is proven .

This completes the proof of Theorem 3. 4. 1 .

APPENDIX B

B.1 SOLUTION OF A LINEAR HEREDITARY SYSTEM

Consider the system in E^n

(B.1.1) $$\dot{x}(t) = \sum_{i=0}^{m} B_i(t)x(t-h_i) + \int_{a}^{t} L_1(t-s)x(s)ds + h(t)$$

On $[t_0, t_1]$, with $x(t) = \phi(t)$, $t \in [a, t_0)$, $x(t_0) = x_0$

Let $Y(s,t)$ be a nxn matrix valued function, undetermined as yet, with the properties

(B.1.2) $$Y(t,t) = I \quad , \quad Y(s,t) = 0 \quad \text{for} \quad t < s$$

Multiplying system (B.1.1), from the left, by $Y(s,t)$ and integrating we obtain, as in [10, pp. 359-371], :

(B.1.3) $$\int_{\gamma}^{t} Y(s,t)\dot{x}(s)ds = \sum_{i=0}^{m} \int_{\gamma}^{t} Y(s,t)B_i(s)x(s-h_i)ds + \int_{\gamma}^{t} Y(s,t)\int_{a}^{s} L_1(s-\tau)x(\tau)d\tau ds + \int_{\gamma}^{t} Y(s,t)h(s)ds \quad , \quad \gamma \in [t_0, t_1]$$

Integrating by parts,

$$Y(t,t)x(t) - Y(\gamma,t)x(\gamma) - \int_{\gamma}^{t} \frac{\partial Y(s,t)}{\partial s} x(s)ds =$$

$$= \sum_{i=0}^{m} \int_{\gamma-h_i}^{\gamma} Y(s+h_i, t)B_i(s+h_i)x(s)ds \quad +$$

$$+\sum_{i=0}^{m}\int_{\gamma}^{t-h_i} Y(s+h_i,t)B_i(s+h_i)x(s)ds + \int_{\gamma}^{t} Y(s,t)\int_{\alpha}^{\gamma} L_1(s-\tau)x(\tau)d\tau ds +$$

$$+\int_{\gamma}^{t} Y(s,t)\int_{\gamma}^{s} L_1(s-\tau)x(\tau)d\tau ds + \int_{\gamma}^{t} Y(s,t)h(s)ds =$$

(B.1.4)

$$= \sum_{i=0}^{m}\int_{\gamma-h_i}^{\gamma} Y(s+h_i,t)B_i(s+h_i)x(s)ds + \int_{\alpha}^{\gamma}\{\int_{\gamma}^{t} Y(s,t)L_1(s-\tau)ds\}x(\tau)d\tau +$$

$$+\sum_{i=0}^{m}\int_{\gamma}^{t} Y(s+h_i,t)B_i(s+h_i)x(s)ds + \int_{\gamma}^{t}\int_{s}^{t} Y(\tau,t)L_1(\tau-s)d\tau ds +$$

$$+\int_{\gamma}^{t} Y(s,t)h(s)ds$$

or ,

$$x(t) = Y(\gamma,t)x(\gamma) + \sum_{i=0}^{m}\int_{\gamma-h_i}^{\gamma} Y(s+h_i,t)B_i(s+h_i)x(s)ds +$$

(B.1.5)

$$+\int_{\alpha}^{\gamma}\{\int_{\gamma}^{t} Y(s,t)L_1(s-\tau)ds\}x(\tau)d\tau + \int_{\gamma}^{t} Y(s,t)h(s)ds +$$

$$+\int_{\gamma}^{t}\{\frac{\partial Y(s,t)}{\partial s} + \sum_{i=0}^{m} Y(s+h_i,t)B_i(s+h_i) + \int_{s}^{t} Y(\tau,t)L_1(\tau-s)d\tau\}x(s)ds$$

If we ask now that $Y(s,t)$ satisfies the system :

$$\text{(B.1.6)} \quad \frac{\partial Y(s,t)}{\partial s} + \sum_{i=0}^{m} Y(s+h_i,t)B_i(s+h_i) + \int_{s}^{t} Y(\tau,t)L_1(\tau-s)d\tau = 0 \quad , s\in[\gamma,t]$$

with $Y(t,t) = I$, $Y(s,t) = 0$ for $s > t$

then the solution of (B.1.1) is going to be :

(B.1.7) $$x(t) = Y(\gamma,t)x(\gamma) + \sum_{i=0}^{m} \int_{\gamma-h_i}^{\gamma} Y(s+h_i,t)B_i(s+h_i)x(s)ds \; +$$

$$+ \int_{a}^{\gamma} \{\int_{\gamma}^{t} Y(s,t)L_1(s-\tau)ds\} x(\tau)d\tau + \int_{\gamma}^{t} Y(s,t)h(s)ds$$

for $\gamma = t_0$.

If now $N(s,t)$ is the "characterestic matrix" of system (B.1.1), i.e., if $N(s,t)$ satisfies the system :

(B.1.8) $$\frac{\partial N(s,t)}{\partial t} = \sum_{i=0}^{m} B_i(t)N(s,t-h_i) + \int_{a}^{t} L_1(t-\tau)N(s,\tau)d\tau \;,\; t\in[\gamma,t_1]$$

with $N(s,s) = I$, $N(s,t) = 0$ for $s > t$,

then , by substituting $N(\gamma,t)$ in (B.1.7) (for $h \equiv 0$) , we obtain :

(B.1.9) $$N(\gamma,t) = Y(\gamma,t)N(\gamma,\gamma) + \sum_{i=0}^{m} \int_{\gamma-h_i}^{\gamma} Y(s+h_i,t)B_i(s+h_i)N(\gamma,s)ds \; +$$

$$+ \int_{a}^{\gamma} \{\int_{\gamma}^{t} Y(s,t)L_1(s-\tau)ds\} N(\gamma,\tau)d\tau$$

or ,

(B.1.10) $$N(\gamma,t) = Y(\gamma,t) \;,\; \text{for } t,\gamma \in [t_0,t_1]$$

Thus , the solution of (B.1.1) can be also written as :

(B.1.11) $$x(t) = N(\gamma,t)x(\gamma) + \sum_{i=0}^{m} \int_{\gamma-h_i}^{\gamma} N(s+h_i,t)B_i(s+h_i)x(s)ds \; +$$

$$+ \int_{a}^{\gamma} \{\int_{\gamma}^{t} N(s,t)L_1(s-\tau)ds\} x(\tau)d\tau + \int_{\gamma}^{t} N(s,t)h(s)ds$$

for $\gamma = t_0$, where $N(s,t)$ (<u>Kernel Matrix of the Second Kind</u> of [4]) satisfies (B.1.8) .

If $B_i(t) = B_i$ =const. , it can be shown [4, p. 104] that

(B.1.12) $$N(s,t) = N(t-s)$$

B.2 SOLUTION OF THE 'ADJOINT' SYSTEM

The adjoint system corresponding to (B.1.1) (for $h \equiv 0$) is :

$$\dot{\psi}(t) = -\sum_{i=0}^{m} B_i^T(t+h_i)\psi(t+h_i) - \int_t^{t_1} L_1^T(s-t)\psi(s)ds \quad , t \in [t_0, t_1] \tag{B.2.1}$$

with $\psi(t_1) = \psi_1$, $\psi(t) = 0$ for $t > t_1$

Let $N(s,t)$ be the matrix given by (B.1.8). Then

$$\int_t^{t_1} N^T(t,s)\dot{\psi}(s)ds = -\sum_{i=0}^{m} \int_t^{t_1} N^T(t,s)B_i^T(s+h_i)\psi(s+h_i)ds - \int_t^{t_1} N^T(t,s) \int_s^{t_1} L_1^T(\tau-s)\psi(\tau)d\tau ds \tag{B.2.2}$$

or,

$$N^T(t,t_1)\psi(t_1) - N^T(t,t)\psi(t) - \int_t^{t_1} \frac{\partial N^T(t,s)}{\partial s} \psi(s)ds =$$

$$= -\sum_{i=0}^{m} \int_{t+h_i}^{t_1} N^T(t,s-h_i)B^T_i(s)\psi(s)ds - \int_t^{t_1}\left(\int_t^{\tau} N^T(t,s)L_1^T(\tau-s)ds\right)\psi(\tau)d\tau = \tag{B.2.3}$$

$$= -\sum_{i=0}^{m} \int_t^{t_1} N^T(t,s-h_i)B^T_i(s)\psi(s)ds - \int_t^{t_1}\left(\int_t^{s} N^T(t,\tau)L_1^T(s-\tau)d\tau\right)\psi(s)ds$$

or,

$$\psi(t) = N^T(t,t_1)\psi(t_1) - \int_t^{t_1} \left\{\frac{\partial N(t,s)}{\partial s} - \sum_{i=0}^{m} B_i(s)N(t,s-h_i) - \int_t^{s} L_1(s-\tau)N(t,\tau)d\tau\right\}^T \psi(s)ds \tag{B.2.4}$$

or, finally

$$\psi(t) = N^T(t,t_1)\psi(t_1) \quad , \quad t \in [t_0, t_1] \tag{B.2.5}$$

VITA

NAME	:	Nicolaos D. Georganas
DATE OF BIRTH	:	June 15 , 1943
PLACE OF BIRTH	:	Athens , Greece

EDUCATION :

I) HIGHSCHOOL	:	3d Lyceum , Athens-Greece .
II) UNIVERSITY	:	National Technical University of Athens , Athens-Greece .
DEGREE	:	DIPL. ENG. (1966)

END OF REEL

ARC-30