

# Penalization and regularisation of quadratic cost controllability problems in a finite dimensional space

G. Chen, W.H. Mills

► **To cite this version:**

G. Chen, W.H. Mills. Penalization and regularisation of quadratic cost controllability problems in a finite dimensional space. [Research Report] RR-0010, INRIA. 1980. inria-00076551

**HAL Id: inria-00076551**

**<https://hal.inria.fr/inria-00076551>**

Submitted on 24 May 2006

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

**IRIA**

Rapports de Recherche

N° 10

**PENALIZATION  
AND  
REGULARIZATION  
OF QUADRATIC COST  
CONTROLLABILITY PROBLEMS  
IN A FINITE  
DIMENSIONAL SPACE**

**Goong CHEN  
Wendell H. MILLS, Jr.**

Institut National  
de Recherche  
en Informatique  
et en Automatique

Domaine de Voluceau  
Rocquencourt  
BP 105 78150 Le Chesnay  
France  
Tél. 954 90 20

Mars 1980

PENALIZATION AND REGULARIZATION  
OF QUADRATIC COST CONTROLLABILITY PROBLEMS  
IN A FINITE DIMENSIONAL SPACE

Goong CHEN\* and Wendell H. MILLS, Jr.\*\*

Résumé : Nous étudions les problèmes de contrôlabilité avec un coût quadratique dans un espace de dimension finis par les méthodes de pénalisation et de régularisation. Les feedback-synthèses associées à ces problèmes sont aussi étudiés. Nous calculons des approximations par la méthode des éléments finis du contrôle optimale sur un exemple numérique pour illustrer notre méthode.

Abstract : We discuss controllability problems with a quadratic cost in a finite dimensional space by the methods of penalization and regularization. Feedback synthesis derived from these problems is also studied. Finite element approximations to the optimal control are computed with a numerical example illustrating our methods.

---

\* INRIA, Domaine de Voluceau, Rocquencourt, 78150 LE CHESNAY, France. and Department of Mathematics, Pennsylvania States University, University Park, PA 16802, U.S.A.

\*\* Department of Mathematics, Pennsylvania States University, University Park, PA 16802, U.S.A.

§ 0. INTRODUCTION

Given a control system

$$(IC) \left\{ \begin{array}{l} \frac{dx(t;u)}{dt} = A(t) x(t;u) + f(t) + B(t) u(t), \quad 0 \leq t \leq T \\ x(0;u) = x_0 \end{array} \right.$$

where

$x(t)$  : the state of the system at time  $t$

$x_0$  : initial state ( $\in \mathbb{R}^n$ )

$u \in U_{ad}$  is an admissible control

$A(t) \equiv (a_{ij}(t))_{n \times n} \in C([0, T]; \mathbb{R}^{n \times n})$

$B(t) \equiv (b_{pq}(t))_{n \times m} \in L^2([0, T]; \mathbb{R}^{n \times m}), \quad m \leq n$

$f(t) \equiv (f_i(t)) \in L^2([0, T]; \mathbb{R}^n)$

the controllability problem is : for some prescribed final state  $x_1$ , find an admissible control  $u$  such that the state of the system at  $t=T$  is  $x_1$ , i.e.,

$$(0.1) \quad x(T;u) \equiv x_1$$

Examples of such controllability problems are easy to find. Projectiles being used by military forces are one such instance. The control of a projectile's flight—its pitching, rolling and yawing—is achieved through maneuvering the three principal control surfaces : elevators, ailerons and rudder. We want to achieve a certain target  $x_1$ . This becomes a controllability problem, with  $u(t)$  steering the system from  $x_0$  to  $x_1$ .

Of course, the above is just a simplified model because we have not taken many other factors (guidance, gust disturbances and physical constraints) into account. Also, the target may be moving.

Controllability problems have been undergoing extensive study by mathematicians and engineers. For example, the classical rank condition criterion :  $\text{rank} (B, AB, \dots, A^{n-1}B) = n$  for controllability of an autonomous system.

$$\left\{ \begin{array}{l} \frac{dx(t)}{dt} = Ax(t) + Bu(t) \\ A, B \text{ are constant } n \times n, n \times m \text{ matrices} \end{array} \right.$$

is well-known, and there are some generalizations of this to partial differential equations (e.g. [9] ). Basically, this approach differs very much from that of optimal control in the sense that no optimization is needed in the procedures.

The controllability terminal condition (0.1) actually can be regarded as a constraint thereby enabling one to apply the optimal control technique to study the controllability problem. This is clearly indicated in Lee and Markus' far reaching paper [4], where optimal control problems for systems governed by nonlinear ordinary differential equations are studied in great length and under a very general setting. In particular, the target set can be moving.

An important way to handle physical constraints is by the method of penalization. In [2], [3], [6], this method was used to attack the optimal control problem with constraints : instead of attempting to solve directly the constrained phase optimal control problem, one considers an unconstrained problem wherein the original cost functional is augmented by a nonnegative penalty term. This sharply increases the cost associated with trajectories which violate the phase constraints. In many cases the constrained phase solution of the original optimization problem may be approximated to any desired degree of accuracy by optimal solutions to a sequence of augmented-cost unconstrained problems. In [5], the penalization method is also applied to tackle the optimal control problem, where the governing equation itself is considered as a constraint.

In comparison with the existing literature, our treatment in this paper has the following features :

- (i) The problem we are considering is a controllability problem associated with a quadratic cost functional.
- (ii) The methods of penalization and regularization are used. In addition to the convergence results, optimization procedures by feedback synthesis are concerned.
- (iii) Finite element approximations -a constructive aspect- are included.

In this paper, we will only study a simple model of unconstrained linear (time-varying) systems governed by ordinary differential equations. Systems with constraints and those governed by partial differential equations will be treated in a forthcoming paper.

## §I. CONTROLLABILITY AND PENALTY

Throughout this paper, we assume that

$$Uad \equiv L^2(0, T; R^m)$$

Suppose that  $\phi(t, s)$  is the fundamental matrix solution of  $A(t)$ , i.e.,  $\phi(t, s)$  satisfies

$$(1.1) \quad \begin{cases} \frac{d}{dt} \phi(t, s) = A(t)\phi(t, s) & 0 \leq s \leq t \leq T \\ \phi(s, s) = I_{n \times n} \end{cases}$$

In case  $s = 0$ , we simply write  $\phi(t)$  instead of  $\phi(t, 0)$ . Then we have

$$\phi(t, s) = \phi(t)\phi(s)^{-1}$$

The solution  $x(t;u)$  to (LC) becomes

$$(1.2) \quad \begin{aligned} x(t;u) &= \Phi(t) x_0 + \int_0^t \Phi(t,s) [F(s) + B(s) u(s)] ds \\ &= \Phi(t) x_0 + F(t) + \int_0^t \Phi(t,s) B(s) u(s) ds, \quad F(t) \equiv \int_0^t \Phi(t,s) f(s) ds \end{aligned}$$

For any  $x(t;u)$  satisfying (LC), suppose we have a quadratic functional

$$J(x,u) \equiv \int_0^T [ \|C x(t;u) - z(t)\|^2 + \langle Nu(t), u(t) \rangle ] dt$$

where

$$\begin{aligned} C : L^2(0,T;R^n) &\rightarrow L^2(0,T;R^p) \text{ is linear} \\ N : L^2(0,T;R^m) &\rightarrow L^2(0,T;R^m) \text{ is linear such that} \end{aligned}$$

$$\langle Nu, u \rangle_{L_m^2(0,T)} \geq \nu \|u\|_{L_m^2(0,T)}^2 \quad \nu > 0$$

and  $N = N^*$

$$z \in L^2(0,T;R^p)$$

We are now in a position to pose the Optimal Controllability Problem (OCP)

Given some  $x_1 \in R^n$ , find a control  $u \in U_{ad}$  such that the solution  $x(t;u)$  of (LC) satisfies the terminal condition  $x(T;u) = x_1$  and the control cost  $J(x,u)$  is minimal.

If there is a  $u$  which steers the system from  $x_0$  to  $x_1$  (with a minimal cost  $J(x,u)$ ), we say that (LC) is controllable from  $x_0$  to  $x_1$ . If (LC) is controllable for arbitrary  $x_0$  and  $x_1$ , we say that (LC) is (globally) controllable.

Without requiring the minimality of the control cost, the concept of controllability is dual to that of observability. The following duality theorem can be found in [7].

Theorem 1.1 The system

$$(1.3) \quad \frac{dx(t;u)}{dt} = A(t) x(t;u) + B(t) u(t)$$

is controllable for any initial and final states  $x_0, x_1$  if and only if the system

$$(10) \quad \begin{cases} \frac{dy(t)}{dt} = -A^*(t) y(t) \\ w(t) = B(t)^* y(t) \end{cases}$$

is observable. (10) is observable if and only if

$$(1.4) \quad Z(T) \equiv \Phi(T) \left[ \int_0^T \Phi(s)^{-1} B(s) B^*(s) (\Phi(s)^{-1})^* ds \right] \Phi(T)^*$$

is a positive definite matrix. Furthermore, for any  $x_0, x_1$ , the control

$$\hat{u}(t) = B^*(t) \Phi^*(T, t) Z(T)^{-1} [x_1 - \Phi(T) x_0] \quad 0 \leq t \leq T$$

steers (1.3) from  $x_0$  to  $x_1$  such that  $\|\hat{u}\|_{L_m^2(u, T)}$  is minimal.

Here we prove a slightly generalized version of the above theorem.

Theorem 1.2. The system (1C) is controllable for any initial states  $x_0$  and  $x_1$  if and only if the system (10) is observable. (10) is observable if and only if

$$(1.5) \quad Z_N(T) \equiv \int_0^T \Phi(T, s) B(s) N(s)^{-1} B^*(s) \Phi^*(T, s) ds$$



is a positive definite matrix. For any  $x_0, x_1$  the control

$$(1.6) \quad \hat{u}(t) \equiv N(t)^{-1} B^*(t) \Phi^*(T, t) Z_N^{-1}(T) [x_1 - \Phi(T)x_0 - F(T)], \quad 0 \leq t \leq T$$

steers (LC) from  $x_0$  to  $x_1$  such that  $\langle N\hat{u}, \hat{u} \rangle_{L_m^2(0, T)}$  is minimal.

Proof : We first note that Theorem 1.1 remains true with the presence of  $f(t) \neq 0$  in (1.3) because only a translation  $F(T)$  is resulted from  $f$  at  $t = T$ .

Now, we show that  $\hat{u}$  steers (LC) from  $x_0$  to  $x_1$ .

$$\begin{aligned} x(T) &= \Phi(T)x_0 + \int_0^T \Phi(T, s) [B(s)\hat{u}(s) + f(s)] ds \\ &= \Phi(T)x_0 + F(T) + \int_0^T \Phi(T, s) B(s) N^{-1}(s) B^*(s) \Phi^*(T, s) ds \{ Z_N^{-1}(T) [x_1 - \Phi(T)x_0 - F(T)] \} \\ &= \Phi(T)x_0 + F(T) + Z_N^{-1}(T) \{ Z_N^{-1}(T) [x_1 - \Phi(T)x_0 - F(T)] \} \\ &= x_1 \end{aligned}$$

Note that there is no question about the existence of  $Z_N^{-1}(T)$  because  $Z(T)$  in (1.4) is positive and  $N$  is invertible.

Next we prove that  $\hat{u}$  has the minimal cost  $\langle Nu, u \rangle$ . Let  $u$  also steer  $x_0$  to  $x_1$ . Then

$$x_1 - \Phi(T)x_0 - F(T) = \int_0^T \Phi(T, s) B(s) u(s) ds = \int_0^T \Phi(T, s) B(s) \hat{u}(s) ds$$

so

$$\int_0^T \Phi(T, s) B(s) [u(s) - \hat{u}(s)] ds = 0$$

$$\int_0^T [x_1 - \phi(T)x_0 - F(T)]^* (Z_N^{-1}(T))^* \phi(T,s) B(s) [u(s) - \hat{u}(s)] ds = 0 \text{ (as a scalar)}$$

Hence  $\int_0^T \hat{u}^*(s) N(s) [u(s) - \hat{u}(s)] ds = 0$  i.e.,  $\langle \hat{u}, N(u-\hat{u}) \rangle_{L_m^2(0,T)} = 0$

with  $N^* = N$ , we derive

$$\langle Nu, u \rangle - \langle N\hat{u}, \hat{u} \rangle = \langle N(u - \hat{u}), u - \hat{u} \rangle \geq \nu \|u - \hat{u}\|_{L_m^2(0,T)}^2 \geq 0$$

The above is equal to 0 if and only if  $u = \hat{u}$ . Q.E.D.

We have some more to say about this theorem in §III.

Now return to the (OCP). Instead of directly minimizing  $J(x, u)$ , we consider

$$(1.7) \quad J_\epsilon(x, u) \equiv J(x, u) + \psi(\epsilon) |x(T; u) - x_1|^2 \quad \epsilon > 0$$

where  $\psi(\epsilon)$  is a positive continuous function from  $\mathbb{R}^+$  into  $\mathbb{R}^+$ , strictly decreasing on  $(0, \epsilon_0)$  for some  $\epsilon_0 > 0$  and such that

$$(1.8) \quad \psi(\epsilon) \rightarrow +\infty \quad \text{as} \quad \epsilon \downarrow 0$$

The simplest such  $\psi(\epsilon) = 1/\epsilon$ . In general,  $\psi(\epsilon)$  is a function of the mesh size  $\epsilon$  in finite element approximations. Here in (1.7), we regard the controllability terminal condition  $x(T; u) = x_1$  as a constraint. If  $x(T; u) = x_1$  is not satisfied, the penalty  $\psi(\epsilon) |x(T; u) - x_1|^2$  is imposed.

Theorem 1.3 For every  $\epsilon > 0$ , there is a unique  $\hat{u}_\epsilon \in L^2(0, T; \mathbb{R}^m)$  which minimizes  $J_\epsilon(x, u)$ .

The proof is standard : find the weak limit of a subsequence of a mini-

zing sequence and use the lower semicontinuity of  $J_\epsilon(x, u)$  to show minimality. The uniqueness follows from strict convexity of  $J_\epsilon$ . Q.E.D.

So for every  $\epsilon > 0$ , let  $\hat{u}_\epsilon$  be the unique control which minimizes  $J_\epsilon(x, u)$ . For  $0 < \epsilon_2 < \epsilon_1 (< \epsilon_0)$ ,

$$\begin{aligned} J_{\epsilon_2}(x, \hat{u}_{\epsilon_2}) &= J(x, \hat{u}_{\epsilon_2}) + \psi(\epsilon_2) |x(T; \hat{u}_{\epsilon_2}) - x_1|^2 \\ &= J(x, \hat{u}_{\epsilon_2}) + \psi(\epsilon_1) |x(T; \hat{u}_{\epsilon_2}) - x_1|^2 + (\psi(\epsilon_2) - \psi(\epsilon_1)) |x(T; \hat{u}_{\epsilon_2}) - x_1|^2 \\ &\geq J_{\epsilon_1}(x, \hat{u}_{\epsilon_1}) + (\psi(\epsilon_2) - \psi(\epsilon_1)) |x(T; \hat{u}_{\epsilon_2}) - x_1|^2 \\ &\geq J_{\epsilon_1}(x, \hat{u}_{\epsilon_1}) \end{aligned}$$

Accordingly,  $J_\epsilon(x, \hat{u}_\epsilon)$  is increasing as  $\epsilon$  is decreasing to 0. This means that cost gets higher as more controllability is achieved.

The following is the fundamental theorem.

Theorem 1.4 The system (LC) is controllable from  $x_0$  to  $x_1$ , if and only if  $J_\epsilon(x, \hat{u}_\epsilon)$  is bounded from above by some  $M > 0$  (depending on  $x_0, x_1$ ). If (LC) is controllable from  $x_0$  to  $x_1$ , then  $\hat{u}_\epsilon$  converges strongly to a control  $\hat{u}$  (as  $\epsilon \rightarrow 0$ ) which solves the (OPC).

Proof : If (LC) is controllable from  $x_0$  to  $x_1$ , then there is a control  $\bar{u}$  which steers (LC) from  $x_0$  to  $x_1$ , so

$$(1.9) \quad J(x, \bar{u}) = J(x, \bar{u}) + \psi(\epsilon) |x(T; \bar{u}) - x_1|^2 \geq \min_u J_\epsilon(x, u) = J_\epsilon(x, \hat{u}_\epsilon)$$

Thus  $J_\varepsilon(x, \hat{u}_\varepsilon)$  is bounded from above for all  $0 < \varepsilon < \varepsilon_0$  by  $M \equiv J(x, \bar{u})$ .

Conversely, assume that all  $J_\varepsilon(x, \bar{u}_\varepsilon)$  is bounded from above by some  $M > 0$ . Then

$$J(x, \hat{u}_\varepsilon) + \psi(\varepsilon) |x(T; \hat{u}_\varepsilon) - x_1|^2 \leq M$$

Hence

$$(1.10) \quad |x(T; \hat{u}_\varepsilon) - x_1|^2 \leq M/\psi(\varepsilon) \rightarrow 0 \text{ as } \varepsilon \rightarrow 0$$

Because  $\{\hat{u}_\varepsilon\}$  is bounded in  $L_m^2(0, T)$ , it contains a subsequence  $\hat{u}_{\varepsilon_n} \rightarrow \hat{u}$  weakly convergent. From (1.2) and (1.10), we see that

$$\begin{aligned} x(T; \hat{u}) &= \lim_{\varepsilon_n \downarrow 0} \left\{ \Phi(T) x_0 + F(T) + \int_0^T \Phi(T, s) B(s) \hat{u}_{\varepsilon_n}(s) ds \right\} \\ &= \lim_{\varepsilon_n \downarrow 0} x(T; \hat{u}_{\varepsilon_n}) = x_1 \end{aligned}$$

So  $\hat{u}$  accomplishes controllability. (LC) is controllable from  $x_0$  to  $x_1$ . From (1.9) we have

$$(1.11) \quad J(x, \hat{u}) \geq J_\varepsilon(x, \hat{u}_\varepsilon) \geq J(x, \hat{u}_\varepsilon).$$

On the other hand, weak consequence  $\hat{u}_{\varepsilon_n} \rightarrow \hat{u}$  implies the strong convergence of  $x(t; \hat{u}_{\varepsilon_n})$  to  $x(t; \hat{u})$ . We have

$$(1.12) \quad J(x, \hat{u}) \leq \liminf_{\varepsilon_n \downarrow 0} J(x, \hat{u}_{\varepsilon_n})$$

Combining (1.11) and (1.12) and the remark before Theorem 1.4., we conclude

$$J(x, \hat{u}) = \lim_{\varepsilon_n \downarrow 0} J(x, \hat{u}_{\varepsilon_n})$$

This implies that  $\hat{u}_{\varepsilon_n}$  converges strongly to  $\hat{u}$  in  $L_m^2(0, T)$  since  $\langle Nv, v \rangle^{1/2}$  is an equivalent norm in  $L_m^2(0, T)$ . We also remark that the whole sequence  $\{\hat{u}_{\varepsilon_n}\}$  converges strongly to  $\hat{u}$ , since every subsequence  $\{\hat{u}_{\varepsilon_n}\}$  contains a subsequence converging strongly to  $\hat{u}$ .

Now for any  $\bar{u}$  steering  $x_0$  to  $x_1$ , we have

$$J(x, \bar{u}) \leq \liminf_{\varepsilon_n} J_{\varepsilon_n}(\hat{u}_{\varepsilon_n})$$

for a subsequence  $\hat{u}_{\varepsilon_n}$ . From (.19) we deduce that

$$J(\bar{u}) \geq J(\hat{u})$$

Hence  $\hat{u}$  solves the (OCP).

The rate of convergence (1.10) can be improved to

$$\psi(\varepsilon)^{1/2} |x(T; \hat{u}_{\varepsilon}) - x_1|^2 \rightarrow 0 \quad \text{as } \varepsilon \downarrow 0$$

because  $J(x, \hat{u}_{\varepsilon})$  converges to  $J(x, \hat{u})$ .

Q.E.D.

Theorem 1.4 formulates an equivalent condition for controllability depending on initial and terminal conditions only. In the next theorem, we will formulate an equivalent condition for global controllability.

Theorem 1.5. (Global uniform bounds)

The system (LC) is controllable from  $x_0$  to  $x_1$  for arbitrary given  $x_0, x_1$  if and only if there exists a positive constants  $M$  independent of  $x_0, x_1$  such that

$$J_{\epsilon}(x, \hat{u}_{\epsilon}) \leq M |x_0 - \Phi(T)x_1 - F(T)|^2 + 4 \| |C| |^2 \| \Phi(\cdot)x_0 + F(\cdot) \|_{L_n^2(0,T)}^2 + \\ + 2 \| |z| \|_{L_p^2(0,T)}^2$$

Proof: By theorem 1.2, the control

$$\bar{u}(t) \equiv N(t)^{-1} B^*(t) \Phi^*(T,t) Z_N^{-1} [x_1 - \Phi(T)x_0 - F(T)]$$

steers (LC) from  $x_0$  to  $x_1$ . Therefore

$$J(x, \bar{u}) \geq J_{\epsilon}(x, \hat{u}_{\epsilon}) \quad \text{for all } 0 < \epsilon < \epsilon.$$

But

$$J(x, \bar{u}) = \int_0^T (|Cx(t; \bar{u}) - z(t)|^2 + \langle N(t) \bar{u}(t), \bar{u}(t) \rangle) dt \\ \leq \int_0^T [2|Cx(t; \bar{u})|^2 + 2|z(t)|^2 + \langle N(t) \bar{u}(t), \bar{u}(t) \rangle] dt$$

It is clear that

$$\langle N\bar{u}, \bar{u} \rangle \leq K_1 |x_0 - \Phi(T)x_0 - F(T)|^2$$

with

$$K_1 \equiv \left\| \int_0^T Z_N^{-1*} \Phi(T,t) B(t) B^*(t) \Phi(T,t)^* Z_N^{-1} dt \right\|$$

Also from (1.2)

$$2 \int_0^T |Cx(t, \bar{u})|^2 dt \leq 4 \left\{ \int_0^T [ |C(\Phi(t)x_0 + F(t))|^2 + |C \int_0^T \Phi(t,s) B(s) \\ N^{-1}(s) B^*(s) \Phi^*(t,s) \cdot Z_N^{-1} (x_1 - \Phi(T)x_0 - F(T))|^2 ds] dt \right\}$$

$$\leq 4 \|C\|^2 \|\Phi(\cdot) x_0 + F(\cdot)\|_{L_m^2(0,T)}^2 + K_2 \|x_0 - \Phi(T)x_0 - F(T)\|^2$$

where  $K_2$  can be similarly defined.

Let  $M \equiv K_1 + K_2$ , the proof is complete.

Q.E.D.

§ II. REGULARIZATION AND PENALIZATION

In this section we adopt an approach [5] which provides us with a sequence of regularized controls approximating optimal control. Let

$$U_1 \equiv H_m^1(0, T) = \{u | u, u' \in L_m^2(0, T)\}$$

and consider minimizing

$$(2.1) \quad J_{\epsilon_1, \epsilon_2}(x, u) \equiv J(x, u) + \psi(\epsilon_1) |x(T; u) - x_1|^2 + \epsilon_2 \int_0^T \langle u'(t), u'(t) \rangle dt$$

for all  $u \in H_m^1(0, T)$ . Using the same argument as in Theorem 1.3, one easily proves that there is a unique element  $\hat{u}_{\epsilon_1, \epsilon_2}$  in  $U_1$  which minimizes  $J_{\epsilon_1, \epsilon_2}$ .

Theorem 2.1. Assume the system is controllable from  $x_0$  to  $x_1$ . As  $\epsilon_1, \epsilon_2 \rightarrow 0$   $\hat{u}_{\epsilon_1, \epsilon_2}$  converges strongly in  $L_m^2(0, T)$  to some  $\hat{u}$  which is the unique solution for the (OCP). For each fixed  $\epsilon_1 > 0$ ,  $\hat{u}_{\epsilon_1, \epsilon_2}$  converges strongly in  $L_m^2(0, T)$  to  $\hat{u}_{\epsilon_1}$  which is the unique solution in  $U_1$  Theorem 1.4.

Proof : Let  $\hat{u}_{\epsilon_1, \epsilon_2}$ ,  $\hat{u}_{\epsilon_1}$  minimize  $J_{\epsilon_1, \epsilon_2}(x, u)$  and  $J_{\epsilon_1}(x, u)$ , respectively.

Since  $H_m^1(0, T)$  is dense in  $L_m^2(0, T)$ , there exists an  $H_m^1$ -sequence  $w_n^{\epsilon_1} \rightarrow \hat{u}_{\epsilon_1}$  strongly in  $L_m^2(0, T)$  such that

$$J(x, w_n^{\epsilon_1}) + \psi(\epsilon_1) |x(T; w_n^{\epsilon_1}) - x_1|^2 \leq J(x, \hat{u}_{\epsilon_1}) + \psi(\epsilon_1) |x(T; \hat{u}_{\epsilon_1}) - x_1|^2 + \eta/2$$

for any  $\eta > 0$  provided  $n$  is large enough.

On the other hand, we have

$$J_{\varepsilon_1, \varepsilon_2}(x, \hat{u}_{\varepsilon_1, \varepsilon_2}) \leq J_{\varepsilon_1, \varepsilon_2}(x, w_n^{\varepsilon_1}) \leq J_{\varepsilon_1}(x, \hat{u}_{\varepsilon_1}) + \eta/2 + \varepsilon_2 \|w_n^{\varepsilon_1}\|_{L^2}$$

We choose  $\varepsilon_2$  so small that  $\varepsilon_2 \|w_n^{\varepsilon_1}\| \leq \eta/2$ . Thus for any  $\eta > 0$ , there exists  $\varepsilon(\eta)$  such that for  $\varepsilon_2 \leq \varepsilon(\eta)$ ,

$$J_{\varepsilon_1, \varepsilon_2}(x, \hat{u}_{\varepsilon_1, \varepsilon_2}) \leq J_{\varepsilon_1}(\hat{u}_{\varepsilon_1}) + \eta$$

so

$$\overline{\lim}_{\varepsilon_2 \downarrow 0} J_{\varepsilon_1, \varepsilon_2}(x, \hat{u}_{\varepsilon_1, \varepsilon_2}) \leq J_{\varepsilon_1}(\hat{u}_{\varepsilon_1})$$

Now as  $\varepsilon_1, \varepsilon_2 \rightarrow 0$ ,  $\hat{u}_{\varepsilon_1, \varepsilon_2}$  is bounded in  $L_m^2(0, T)$ , so it contains a subsequence converging weakly to  $\hat{u}$  in  $L_m^2$  for some  $\hat{u}$ . According to the proof of Theorem 1.4., this  $\hat{u}$  steers (IC) from  $x_0$  to  $x_1$ . Also,

$$J(x, u) \leq \lim_{\varepsilon_1, \varepsilon_2 \downarrow 0} J(x, \hat{u}_{\varepsilon_1, \varepsilon_2}) \leq \lim_{\varepsilon_1 \downarrow 0} J_{\varepsilon_1}(x, \hat{u}_{\varepsilon_1}) = J(x, \hat{u})$$

Hence for a subsequence  $\varepsilon_1^{(n)}, \varepsilon_2^{(n)}$ , we have

$$\lim_{\varepsilon_1^{(n)}, \varepsilon_2^{(n)} \downarrow 0} \langle N\hat{u}_{\varepsilon_1^{(n)}, \varepsilon_2^{(n)}}; \hat{u}_{\varepsilon_1^{(n)}, \varepsilon_2^{(n)}} \rangle = \langle N\hat{u}, \hat{u} \rangle$$

This implies the strong convergence of  $\hat{u}_{\varepsilon_1^{(n)}, \varepsilon_2^{(n)}} \rightarrow \hat{u}$  in  $L_m^2$ .

For fixed  $\varepsilon_1$ , one easily sees from the proof of [5, Theorem 1.5, p. 361] that  $\hat{u}_{\varepsilon_1, \varepsilon_2}$  converges strongly to  $\hat{u}_{\varepsilon_1}$  as  $\varepsilon_2 \rightarrow 0$ . Q.E.D.



### § III. FEEDBACK SYNTHESIS

The optimal control  $\hat{u}_\epsilon$  minimizing  $J_\epsilon$  can be characterized [5] by the variational equation

$$\begin{aligned} & \langle Cx(o; \hat{u}) - z, Cx(o; v) - Cx(t; \hat{u}) \rangle_{L_p^2(o,T)} + \langle N\hat{u}, v - \hat{u} \rangle_{L_m^2(o,T)} + \\ & + \psi(\epsilon) \langle x(T; \hat{u}) - x_1, x(T; v) - x(T; \hat{u}) \rangle_{R^n} = 0 \end{aligned}$$

Introducing the adjoint state  $p_\epsilon(t)$

$$\begin{cases} \frac{dp_\epsilon(t)}{dt} = -A^* p_\epsilon(t) - C^* Cx(t; \hat{u}_\epsilon) + C^* z(t) \\ p_\epsilon(T) = \psi(\epsilon) [x(T; \hat{u}_\epsilon) - x_1] \end{cases}$$

we obtain

$$\int_0^T \langle N(t) \hat{u}_\epsilon(t) + B^*(t) p_\epsilon(t), v(t) - \hat{u}_\epsilon(t) \rangle_{R^m} = 0 \quad \forall v \in U_{ad}$$

Hence

$$(3.1) \quad \hat{u}_\epsilon = -N^{-1} B^* p_\epsilon$$

Thus  $x(t; \hat{u}_\epsilon)$  and  $p_\epsilon(t)$  are coupled through

$$(3.2) \quad \begin{cases} \frac{dx}{dt} - Ax + BN^{-1} B^* p_\epsilon = f \\ x(o; \hat{u}_\epsilon) = x_0 \\ \frac{dp_\epsilon}{dt} + A^* p_\epsilon + C^* Cx = C^* z \\ p_\epsilon(T) = \psi(\epsilon) [x(T; \hat{u}_\epsilon) - x_1] \end{cases}$$

on  $[0, T]$ .

Consider the above (3.2) beginning at some intermediate time  $s \in [0, T]$ .

We write

$$(3.3) \quad \left\{ \begin{array}{l} \frac{dx}{dt} - Ax + BN^{-1} B^* p_\epsilon = f \\ \frac{dp_\epsilon}{dt} + A^* p_\epsilon + C^* Cx = C^* z \\ x(s; \hat{u}_\epsilon) = h \\ p_\epsilon(T) = \psi(\epsilon) [x(T; \hat{u}_\epsilon) - x_1] \end{array} \right.$$

We decompose  $x, p_\epsilon$  into  $x = y_1 + y_2, p_\epsilon = P_1 + P_2$ , where  $y_1, P_1$  satisfy

$$(3.4) \quad \left\{ \begin{array}{l} \frac{dy_1}{dt} - Ay_1 + BN^{-1} B^* P_1 = 0 \\ -\frac{dP_1}{dt} + A^* P_1 + C^* C y_1 = 0 \\ y_1(s) = h \\ P_1(T) = \psi(\epsilon) y_1(T) \end{array} \right.$$

and  $y_2, P_2$  satisfy

$$(3.5) \quad \left\{ \begin{array}{l} \frac{dy_2}{dt} - Ay_2 + BN^{-1} B^* P_2 = f \\ -\frac{dP_2}{dt} + A^* P_2 + C^* C y_2 = C^* z \\ y_2(s) = 0 \\ P_2(s) = \psi(\epsilon) y_2(T) - \psi(\epsilon) x_1 \end{array} \right.$$

From (3.3) we see that  $P_1(t)$  ( $s \leq t \leq T$ ) is a linear function of  $h$ , in particular, at  $t = s$ . So we write  $P_1(s) = P(s)h$ .  $P_2(t)$  is independent of  $h$ , so we denote it by  $\gamma_\epsilon(t)$ . Thus we have the affine relation

$$(3.6) \quad p_\epsilon(t) = P_\epsilon(t)h + \gamma_\epsilon(t) = P_\epsilon(t) x(t; \hat{u}_\epsilon) + \gamma_\epsilon(t)$$

thereby achieving the decoupling

$$(3.7) \quad \left\{ \begin{array}{l} P'_\epsilon + P_\epsilon A + A^* P_\epsilon - P_\epsilon B N^{-1} B^* P_\epsilon = -C^* C \\ P_\epsilon(T) = \psi(\epsilon) \quad \text{Inxn} \end{array} \right.$$

$$(3.8) \quad \left\{ \begin{array}{l} P_\epsilon(T) = \psi(\epsilon) \quad \text{Inxn} \end{array} \right.$$

and  $\gamma_\epsilon(t)$  satisfies

$$(3.9) \quad \left\{ \begin{array}{l} \gamma'_\epsilon + A^* \gamma_\epsilon - P_\epsilon B N^{-1} B^* \gamma_\epsilon = -P_\epsilon f + C^* z \\ \gamma_\epsilon(T) = -\psi(\epsilon) x_1 \end{array} \right.$$

$$(3.10) \quad \left\{ \begin{array}{l} \gamma_\epsilon(T) = -\psi(\epsilon) x_1 \end{array} \right.$$

One easily verifies that  $P$  has the following properties ([8]) :

1.  $P_\epsilon^*(t) = P_\epsilon(t)$  for all  $t \in [0, T]$
2.  $P_\epsilon(t) \geq 0$ , i.e.,  $\langle P_\epsilon(t)y, y \rangle_{\mathbb{R}^n} \geq 0$  for  $y \in \mathbb{R}^n$
3.  $P_\epsilon(t_1) \leq P_\epsilon(t_2)$  if  $t_1 \geq t_2$
4.  $P_{\epsilon_2}(t) \geq P_{\epsilon_1}(t)$  if  $0 < \epsilon_2 < \epsilon_1 < \epsilon_0$

The terminal conditions (3.8) and (3.10) indicate that

$$\|P_\epsilon(T)\| \rightarrow \infty, |\gamma_\epsilon(T)| \rightarrow \infty \quad \text{as } \epsilon \downarrow 0$$

This seems to be a great disadvantage for Riccati equations synthesis.

Now consider the special case of  $C \equiv 0$ . Then

$$J(x, u) = \int_0^T (|z(t)|^2 + \langle N(t) u(t), u(t) \rangle) dt$$

so  $J(x,u)$  is minimal if and only if  $\langle Nu, u \rangle$  is minimal. The optimal control  $\hat{u}$  on  $[0, T]$  is known in Theorem 1.2. By Bellman's dynamic programming, the optimal control  $\hat{u}_\varepsilon(t)$  which minimizes  $\langle Nu, u \rangle_{L_m^2(s, T)}$

with initial and terminal states  $x(s; \hat{u}_\varepsilon)$  and  $x_1$  (resp.) is

$$(3.11) \quad \hat{u}(s, t) = N^{-1}(t) B^*(t) \phi^*(T, t) Z_N^{-1}(T, s) [x_1 - \phi(T, s) x(s; \hat{u}_\varepsilon) - F(T, s)] \quad (s \leq t \leq T)$$

where

$$Z_N(T, s) \equiv \int_s^T \phi(T, \zeta) B(\zeta) N(\zeta)^{-1} B^*(\zeta) \phi^*(T, \zeta) d\zeta$$

$$F(T, s) = \int_s^T \phi(T, \zeta) f(\zeta) d\zeta$$

In the context of (3.3),  $x(s; \hat{u}_\varepsilon)$  is h. Therefore, comparing (3.1), (3.6), (3.11) and letting  $t = s$ , we get

Theorem 3.1. Assume that  $C \equiv 0$  and  $B$  is nonsingular (a priori,  $m = n$ ). Then

$$P_\varepsilon(\cdot) \rightarrow \phi^*(T, \cdot) Z_N^{-1}(T, \cdot) \phi(T, \cdot)$$

$$\gamma_\varepsilon(\cdot) \rightarrow \phi^*(T, \cdot) Z_N^{-1}(T, \cdot) [x_1 - F(T, \cdot)]$$

as  $\varepsilon \rightarrow 0$ .

The above theorem explores to some extent the structure of solutions to Riccati equations. One can see that as  $s \rightarrow T$ , the (operator) norm of  $\phi^*(T, s) Z_N^{-1}(T, s) \phi(T, s)$  becomes unbounded. This coincides with our earlier observation that  $\|P_\varepsilon(s)\| \rightarrow \infty$  as  $s \rightarrow T$ . Similar observation also holds for  $\gamma_\varepsilon(s)$  as  $s \rightarrow T$ . Nevertheless, we do not know what the rate of convergence is as related to  $\varepsilon$ .

In general  $m < n$  so  $B$  can not be nonsingular, we tend to think that theorem 3.1. remains true. We have yet no proof to confirm this.

As for the cost functional  $J_{\varepsilon_1, \varepsilon_2}$  in (2.1), the optimal control  $\hat{u}_{\varepsilon_1, \varepsilon_2}$  is determined by

$$\left\{ \begin{array}{l} x' = Ax + Bu + f \\ x(0) = x_0 \\ p' = A^*p - C^*Cx + C^*z \\ p(T) = \psi(\varepsilon_1) [x(T) - x_1] \\ - \varepsilon_2 u'' + Nu + B^*p = 0 \\ u'(0) = u'(T) = 0 \\ 0 \leq t \leq T \end{array} \right.$$

This is a coupled system of initial, terminal and boundary value problems.

#### § IV. NUMERICAL APPROXIMATIONS AND EXAMPLE

We choose  $S^h$  to be the space of continuous precewise cubics on  $[0, T]$  on a mesh of length  $h$  and let

$$\hat{S}^h \equiv \prod_{i=1}^m S^h$$

Let  $\{\phi_i\}$  be a basis for  $S^h$ . If  $[0, T]$  is divided into  $k$  intervals, then the number of elements in  $\{\phi_i\}$  is  $3k + 1$ . Consequently, the set of elements

$$\hat{\phi}_{ij} \equiv \phi_i \hat{e}_j, \quad \hat{e}_j : \text{the } j\text{-th unit vector, } 1 \leq j \leq m$$

becomes a basis for  $\hat{S}^h$  consisting of  $NS \equiv (3k+1) \cdot m$  elements.

We want to approximate the optimal control  $\hat{u}$  by elements in the space  $\hat{S}^h$ . According to the theory we have established in § II,  $\hat{u}$  can be approximated to any desired accuracy by  $\hat{u}_\varepsilon$ . Therefore, we consider the minimization problem

$$\begin{aligned} \text{Min}_{v \in \hat{S}^h} J(x, v) = \{ & \int_0^T [ \|C \underline{x}(t; v) - \underline{z}(t)\|^2 + \langle \underline{N}(t) v(t); v(t) \rangle ] dt + \\ & + \psi(h) |x(T; v) - x_1|^2 \} \end{aligned}$$

where  $\underline{x}$ ,  $\underline{z}$ ,  $\underline{N}$  are finite element representations of these functions. The variational formulation of this problem is :

$$(4.1) \quad \left\{ \begin{array}{l} \text{Find } \hat{u}_h \in \hat{S}^h \text{ such that} \\ \int_0^T [ \langle C \underline{x}(t; v) - \underline{z}(t), C \underline{x}(t; v) - C \underline{x}(t; \hat{u}_h) \rangle + \langle \underline{N}(t) v(t), \\ v(t) - \hat{u}_h(t) \rangle ] dt + \psi(h) \langle \underline{x}(T; v) - x_1, \underline{x}(T; v) - \underline{x}(T; \hat{u}_h) \rangle = \\ = 0 \quad \quad \quad v \in \hat{S}^h \end{array} \right.$$

We state the following fundamental theorem here. Its proof will be given in detail in a forthcoming paper.

Theorem 4.1. As sume that (LC) is controllable from  $x_0$  to  $x_1$  and let  $\hat{u}$  be the solution of the (OCP). Let  $\hat{u}_h$  denote the unique solution of (4.1). Then

$$\hat{u}_h \rightarrow \hat{u} \quad \text{in } L^2(0, T) \quad \text{as } h \rightarrow 0$$

Setting

$$\hat{u}_h = \sum_{i=1}^{NS} \alpha_i \psi_i$$

and letting  $v$  run through all the basis elements  $\{ \psi_j \}_1^{NS}$ , the variational equation (4.1) is reduced to the following  $NS \times 1$  quadratic system :

$$(4.2) \quad \left( I_{NS \times NS} - \begin{bmatrix} 1 \\ 1 \\ \vdots \\ \vdots \\ \vdots \\ 1 \end{bmatrix}_{NS \times 1} \vec{\alpha}^* \right) \{ [\psi(h)^{-1} (G+D) + M ] \vec{\alpha} + [\psi(h)^{-1} \beta + \zeta] \} = 0$$

where

$$\vec{\alpha}^* = [\alpha_1, \alpha_2, \dots, \alpha_{NS}]$$

$$G = (\gamma_{ij})_{NS \times NS}, \quad \gamma_{ij} \equiv \int_0^T \langle C \int_0^t \Phi(t,s) B(s) \psi_i(s) ds, C \int_0^t \Phi(t,s) B(s) \psi_j(s) ds \rangle dt$$

$$D = (\delta_{ij})_{NS \times NS}, \quad \delta_{ij} \equiv \int_0^T \langle N(s) \psi_i(s), \psi_j(s) \rangle ds$$

$$M = (\mu_{ij})_{NS \times NS}, \quad \mu_{ij} \equiv \langle \int_0^T \Phi(T,s) B(s) \psi_i(s) ds, \int_0^T \Phi(T,s) B(s) \psi_j(s) ds \rangle$$

$$\vec{\beta}^* = [\beta_1, \beta_2, \dots, \beta_{NS}], \quad \beta_j \equiv \int_0^T \langle C(\Phi(t)x_0 + F(t)) - z(t), C \int_0^t \Phi(t,s) B(s) \psi_j(s) ds \rangle dt$$

$$\vec{\zeta}^* = [\zeta_1, \zeta_2, \dots, \zeta_{NS}], \quad \zeta_j \equiv \langle \Phi(T)x_0 + F(T) - x_1, \int_0^T \Phi(T,s) B(s) \psi_j(s) ds \rangle$$

System 4.2. is very ill-conditioned. The standard non-linear system solvers will not work for (4.2) with cubics. This is typical of known penalization methods for linear partial differential equations [1]. So, instead, we solve the linearized system

$$(4.3) \quad [\psi(h)^{-1} (G+D) + M ] \vec{\alpha} + [\psi(h)^{-1} \beta + \zeta] = 0$$

The matrix  $\psi(h)^{-1} (G+D) + M$  is also very ill-conditioned for large  $\psi(h)$ .

But we can use some standard stiff linear system solvers and compute

$$\alpha = - [\psi(h)^{-1}(G + D) + M]^{-1} [\psi(h)^{-1} \beta + \zeta]$$

It is clear that if  $\alpha$  is a solution of (4.3), then  $\alpha$  is also a solution of (4.2). In fact, (4.2) corresponds to the linear variational problem

$$(4.4) \left\{ \begin{array}{l} \text{Find } \hat{u}_h \in \tilde{S}^h \text{ such that} \\ \int_0^T [\langle C\underline{x}(t; \hat{u}_h) - \underline{z}(t), C(\underline{x}(t; v) - \Phi(t)x_0 - \underline{F}(t)) \rangle + \langle N(t)\hat{u}_h(t), \\ v(t) \rangle] dt + \psi(h) \langle \underline{x}(T; \hat{u}_h) - \underline{x}_1, \underline{x}(T; v) - \Phi(T)x_0 - \underline{F}(T) \rangle = 0 \end{array} \right.$$

Thus a solution of (4.4) is a solution of (4.1).

We end this paper with a simple example.

Example :  $n = m = 2$  ,  $T = \pi/2$

$$\left\{ \begin{array}{l} \frac{d}{dt} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix} \\ \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \end{bmatrix}, \quad \begin{bmatrix} x_1(T) \\ x_2(T) \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \end{array} \right.$$

$$\Phi(t, s) = \begin{bmatrix} \cos(t-s) & \sin(t-s) \\ -\sin(t-s) & \cos(t-s) \end{bmatrix}$$

$$C = 0_{2 \times 2}, \quad z = 0_{2 \times 1}, \quad N \equiv I_{2 \times 2}$$

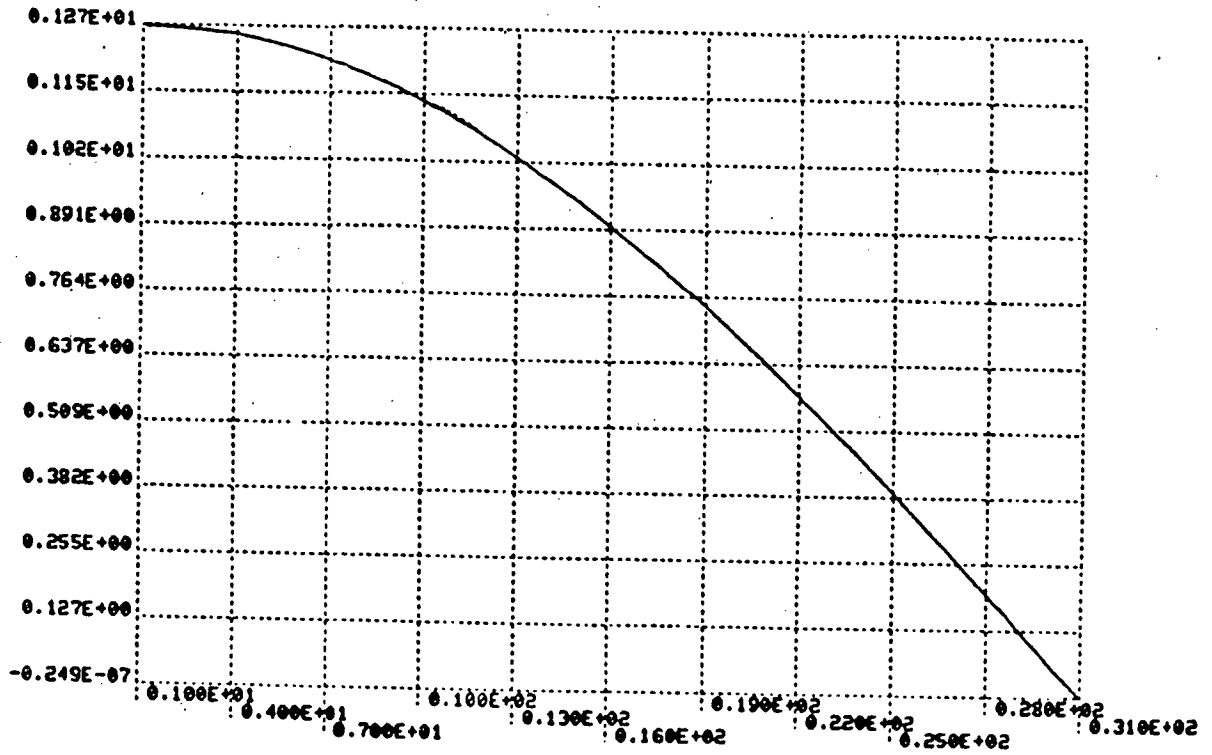
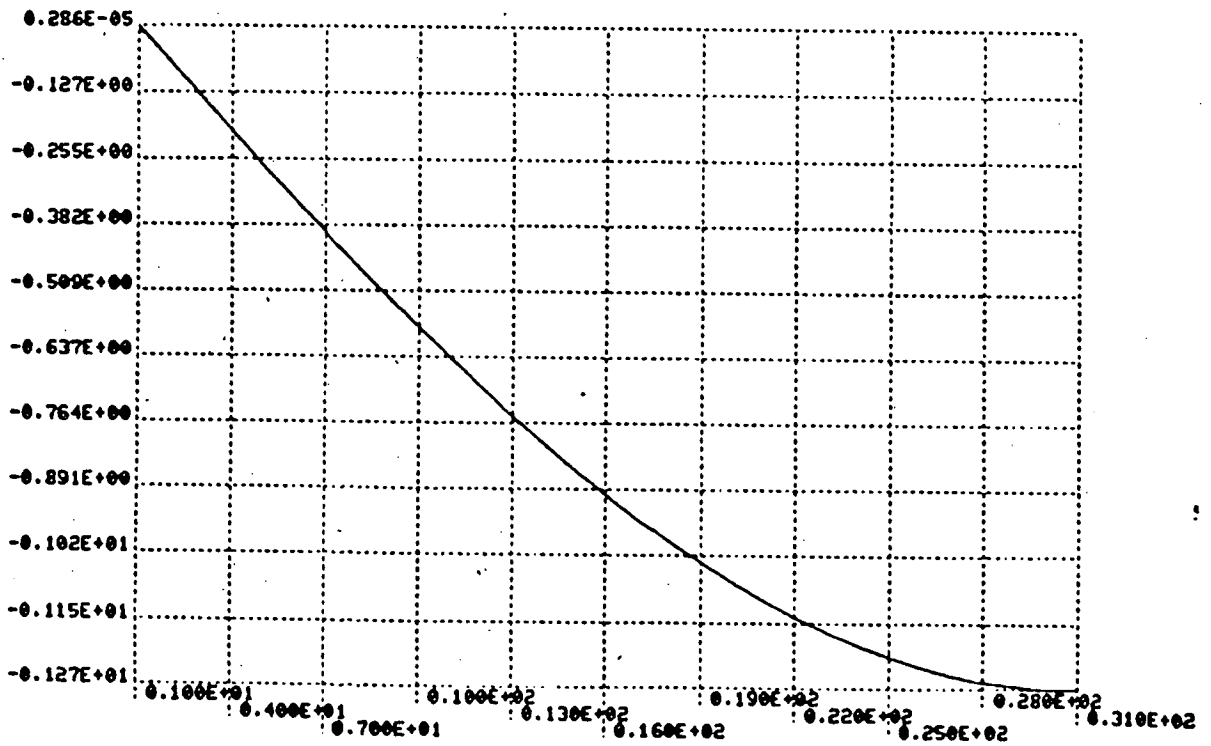


The exact solution and the optimal control are, respectively,

$$x(t) = \begin{bmatrix} \left(\frac{4}{\pi} t - 1\right) \cos t + \sin t \\ - \left(\frac{4}{\pi} t - 1\right) \sin t + \cos t \end{bmatrix}$$

$$u(t) = \begin{bmatrix} \frac{4}{\pi} \cos t \\ - \frac{4}{\pi} \sin t \end{bmatrix} = \begin{bmatrix} u_1(t) \\ u_2(t) \end{bmatrix}$$

For convenience, we input the exact  $\Phi(t,s)$  in our computation of (4.3). In the following two graphs, one can see that the broken line (representing the finite element solution coincides almost exactly with the dark line (exact solution)). The error is  $\leq 10^{-4}$ .

(i)  $u_1(t)$ (ii)  $u_2(t)$

REFERENCES

- 1 Babuska and A.K. Aziz, The mathematical Foundations of the Finite Element Method with Applications to Partial Differential Equations, A.K. Aziz, ed., Academic Press, New York, 1972, pp. 5-359.
- 2 S.S.L. Chang, Optimal Control in Bounded Phase Space, Automatica, Vol. 1, Pergamon Press, New York, 1962, pp. 55-67.
- 3 S.S.L. Chang, An extension of Ascoli's theorem and its applications to the theory of optimal control, AFOSR Report 1962.
- 4 E.B. Lee and L. Markus, Optimal Control for Nonlinear Processes, Arch. Rat. Mech. Anal., 8 (1961), pp. 36-58.
- 5 J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, English Translation, Springer Verlag, New York 1971.
- 6 D.L. Russell, Penalty Functions and Bounded Phase Coordinate Control, Journal SIAM Control, Ser. A, 2 (1965), n° 3, pp. 409-422.
- 7 D.L. Russell, Control Theory, Book in press.
- 8 L.Tartar, Linear Control Theory and Riccati Equations, MRC Technical Summary Report #1555, University of Wisconsin-Madison, Feb. 1976.
- 9 R. Triggiani, Extension of Rank Conditions for Controllability and observability to Banach spaces and Unbounded Operation, SIAM Journal Control Optimization, 14 (1976), pp. 313-338.

ACKNOWLEDGEMENT : The first author would like to thank IRIA and Pennsylvania State University for supporting his research. He especially thanks Professor D.C. Rung, Chairman of the Department of Mathematics, for many personal encouragements. Both Authors thank IRIA for typing and preparing this report.

