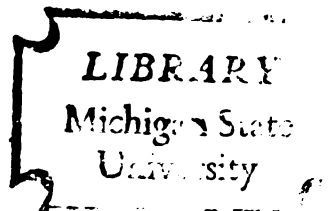


EFFICIENT COMPUTATIONAL PROCEDURES  
FOR OBTAINING OPTIMAL FEEDBACK  
CONTROL OF DISTRIBUTED PARAMETER  
SYSTEMS

Thesis for the Degree of Ph. D.  
MICHIGAN STATE UNIVERSITY  
TUMMALA RAMAMOCHAN LAL  
1970



This is to certify that the

thesis entitled

**Efficient Computational Procedures for  
Obtaining Optimal feedback Control of  
Distributed Parameter Systems**

presented by

Tummala, Ramamohan Lal

has been accepted towards fulfillment  
of the requirements for

Ph. D. degree in Electrical Engineering &  
Systems Science

*Robert O. Barr*  
*John B. Kreer*  
 Major professor

Date May 20, 1970.



## ABSTRACT

### EFFICIENT COMPUTATIONAL PROCEDURES FOR OBTAINING OPTIMAL FEEDBACK CONTROL OF DISTRIBUTED PARAMETER SYSTEMS

By

Tummala Ramamohan Lal

In this thesis techniques have been developed to synthesize the sub-optimal feedback controls for a class of distributed parameter systems. The original system, characterized by partial differential equations is reduced to a set of ordinary differential equations by means of a consistent approximation along the spatial domain. The technique uses no prior information about the optimal open-loop control. The feedback parameters are obtained by solving a parameter optimization problem with differential constraints using a hybrid computer.

The difficulty of solving these problems on the hybrid computer is the large set of differential equations that result due to fine spatial discretization. The number of integrators available on any analog computer is limited, so a decomposition principle is used to decompose a large set of differential equations system into lower order independent subsystems. An iterative method is used to obtain the solution. The convergence theorems are stated and proved. With this treatment a larger system (a finer spatial discretization) can be treated than otherwise would be feasible.

The second method uses the a priori information available about the optimal open-loop control to obtain the time-varying feedback gains. The hybrid computer implementation of this method is simple and straightforward. The timevarying gains are obtained by sequentially solving the parameter optimization problems on a smaller interval than given in the problem. The number of parameters to be determined in the parameter optimization problem is equal to the number of state functions in the given problem. The method terminates when the desired performance is obtained.

EFFICIENT COMPUTATIONAL PROCEDURES FOR  
OBTAINING OPTIMAL FEEDBACK CONTROL OF  
DISTRIBUTED PARAMETER SYSTEMS

By

Tummala Ramamohan Lal

A THESIS

Submitted to  
Michigan State University  
in partial fulfillment of the requirements  
for the degree of

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering and Systems Science

1970

Dedicated to  
Lord Sri Venkateswaraswami

## ACKNOWLEDGMENTS

The author wishes to express his gratitude to Dr. John B. Kreer, for his supervision, constant encouragement and his many helpful comments and suggestions. He also expresses his sincere appreciation to Dr. Robert O. Barr for his interest and constant guidance throughout the period devoted to this thesis. The author also expresses his sincere appreciation to Dr. J.S. Frame for giving many helpful comments. Thanks are also due to Dr. G.L. Park and Dr. R.C. Dubes for their encouragement and advice during the writing of this thesis.

Finally the author owes a great debt of love and gratitude to his wife, Droupathi, for her patience and understanding during his graduate study and her constant help during the preparation of this thesis.

## TABLE OF CONTENTS

| Chapter |                                                                            | Page |
|---------|----------------------------------------------------------------------------|------|
| I       | INTRODUCTION .....                                                         | 1    |
|         | 1.1 Introduction .....                                                     | 1    |
|         | 1.2 Literature Survey Pertinent to the Study<br>of this Dissertation ..... | 1    |
|         | 1.3 Contribution of The Dissertation .....                                 | 5    |
|         | 1.4 Problem Formulation .....                                              | 7    |
|         | 1.4.1 Mathematical model .....                                             | 7    |
|         | 1.4.2 Constraints .....                                                    | 9    |
|         | 1.4.3 Performance Indices .....                                            | 10   |
|         | 1.5 Controllability .....                                                  | 11   |
|         | 1.6 Observability .....                                                    | 14   |
|         | 1.6.1 Definitions .....                                                    | 14   |
| II      | ANALYTICAL SOLUTIONS FOR THE OPTIMAL ENDPOINT<br>CONTROL PROBLEMS .....    | 18   |
|         | 2.1 Introduction .....                                                     | 18   |
|         | 2.2 Necessary Conditions for Optimality .....                              | 18   |
|         | 2.3 Example .....                                                          | 24   |
|         | 2.4 Application of Functional Analysis .....                               | 26   |
| III     | EXISTING METHODS OF COMPUTING OPTIMUM CONTROL ...                          | 29   |
|         | 3.1 Introduction .....                                                     | 29   |
|         | 3.2 Existing Methods for Optimal Open-Loop<br>Control .....                | 29   |
|         | 3.3.1 Sakawa's method .....                                                | 31   |
|         | 3.3.2 Direct search on performance index ..                                | 34   |
|         | 3.4 Existing Methods for Obtaining Optimal<br>Feedback Control .....       | 36   |
|         | 3.4.1 Seinfeld and Kumar's method .....                                    | 37   |
|         | 3.4.2 Koivo and Kruh's method .....                                        | 39   |

| Chapter |                                                                                                         | Page |
|---------|---------------------------------------------------------------------------------------------------------|------|
| IV      | A DECOMPOSITION PRINCIPLE .....                                                                         | 42   |
|         | 4.1 Introduction.....                                                                                   | 42   |
|         | 4.2 Definitions and Theorems .....                                                                      | 43   |
|         | 4.2.1 Spectral radius of a matrix .....                                                                 | 43   |
|         | 4.2.2 Spectral norm of a matrix .....                                                                   | 43   |
|         | 4.2.3 Convergence of a matrix .....                                                                     | 45   |
|         | 4.2.4 Bounds for the spectral radius of a<br>matrix .....                                               | 47   |
|         | 4.2.5 Conditions for the existence of an<br>inverse of $(I-M)$ when $M$ is an<br>arbitrary matrix ..... | 49   |
|         | 4.3 A Decomposition Principle .....                                                                     | 50   |
|         | 4.3.1 Algorithm .....                                                                                   | 51   |
|         | 4.3.2 Convergence theorems .....                                                                        | 56   |
|         | 4.3.3 Average rate of convergence .....                                                                 | 63   |
|         | 4.4 Examples .....                                                                                      | 64   |
|         | 4.4.1 Analytical example .....                                                                          | 64   |
|         | 4.4.2 Computer results .....                                                                            | 65   |
| V       | DEVELOPMENT OF ALGORITHM - I .....                                                                      | 79   |
|         | 5.1 Introduction .....                                                                                  | 79   |
|         | 5.2 Problem Formulation .....                                                                           | 80   |
|         | 5.3 Algorithm - I .....                                                                                 | 80   |
|         | 5.4 Computer Results .....                                                                              | 88   |
|         | 5.5 Sensitivity Considerations .....                                                                    | 99   |
|         | 5.5.1 Sensitivity coefficients .....                                                                    | 99   |
|         | 5.5.2 Computational algorithm .....                                                                     | 102  |
| VI      | DEVELOPMENT OF ALGORITHM - II .....                                                                     | 105  |
|         | 6.1 Introduction .....                                                                                  | 105  |
|         | 6.2 Problem Formulation .....                                                                           | 105  |
|         | 6.3 Algorithm - II .....                                                                                | 106  |
|         | 6.4 Example .....                                                                                       | 109  |
|         | 6.5 Computer Results .....                                                                              | 119  |
| VII     | CONCLUSIONS .....                                                                                       | 123  |
|         | 7.1 Conclusions .....                                                                                   | 123  |
|         | 7.2 Possible Extensions .....                                                                           | 124  |
|         | REFERENCES .....                                                                                        | 126  |

# LIST OF FIGURES

| Figure |                                                                                                   | Page |
|--------|---------------------------------------------------------------------------------------------------|------|
| 1.1    | Complete Null Controllability at Time $t_0$ in $\Gamma' \subseteq \Gamma$ .....                   | 16   |
| 1.2    | Complete Null $\delta$ -Controllability at $t_0$ in $\Gamma' \subseteq \Gamma$ .                  | 16   |
| 1.3    | Complete Controllability in $\Gamma'$ .....                                                       | 16   |
| 1.4    | Finite Tube of System Trajectories .....                                                          | 17   |
| 1.5    | Output Trajectory Corresponding to Output Transformation $M$ in the Transformed Space .....       | 17   |
| 4.1    | Flow Chart of Algorithm .....                                                                     | 53   |
| 4.2    | Initial Guess, Exact Solution and Iterative Solutions for $u_s(t)$ in Example I .....             | 67   |
| 4.3    | Convergence of the Algorithm for $u_2(t)$ in Example I .....                                      | 68   |
| 4.4    | Initial Guesses for Figure 4.5(a) and Figure 4.5(b) .....                                         | 70   |
| 4.5a   | Convergence of the Algorithm for $u_2(t)$ in Example II for Initial Guess A .....                 | 71   |
| 4.5b   | Convergence of the Method for $u_2(t)$ in Example II for Initial Guess B .....                    | 72   |
| 4.6    | Convergence of the Method for $u_4(t)$ in Example III                                             | 74   |
| 4.7    | Initial Guess, Exact Solution, and the Iterative Solution for $u_4(t)$ in Example III .....       | 75   |
| 4.8    | Initial Guess, Exact Solution, and the Iterative Solution for $u_2(t)$ in Example IV .....        | 77   |
| 4.9    | Convergence of the Method for Three Partitions in Example IV for a Sample Function $u_2(t)$ ..... | 78   |
| 5.1    | Illustration of the Discussion in Section 5.3 .....                                               | 84   |

| Figure |                                                  | Page |
|--------|--------------------------------------------------|------|
| 5.2    | Flow Chart of the Algorithm .....                | 89   |
| 5.3    | Verification of Sakawa's Results .....           | 95   |
| 5.4    | Unconstrained Case .....                         | 96   |
| 5.5    | Constrained Case .....                           | 97   |
| 6.1    | Optimal Open Loop Control by Direct Search ..... | 110  |
| 6.2a   | Plot of $2 \sinh p_1/p_1$ .....                  | 114  |
| 6.2b   | Plot of $-2 \sinh p_2/p_2$ .....                 | 114  |
| 6.2c   | Plot of $\tan y = 1/By$ .....                    | 114  |
| 6.2d   | Plot of $\coth X = -BX$ (for $B > 0$ ) .....     | 116  |
| 6.3    | Hybrid Computer Set Up for the Example .....     | 121  |

## CHAPTER I

### INTRODUCTION

#### 1.1 Introduction.

Recent contributions to the theory of optimal control have been concerned primarily with systems whose behaviour can be described by ordinary differential equations. While many physical systems have a spatial energy distribution sufficiently aggregated during the course of motion to be described by ordinary differential equations, many others require formulation by partial differential equations. As a result, the development of optimal control theory for distributed systems is of increasing interest from both theoretical and practical points of view.

#### 1.2 Literature Survey Pertinent to the Study of this Dissertation

Research on optimal control of distributed parameter systems was initiated by Butkovskii and Lerner (B-5), who attempted to define certain types of control problems that might arise. Butkovskii (B-4, B-6, B-7, B-8) subsequently considered the optimal control for a class of systems describable by a set of non-linear integral equations, which can be derived from linear partial differential equations. He derived a maximum principle (in the sense of Pontryagin) embodying the

necessary conditions for optimality of such systems. However, Butkovskii's result requires the explicit solution of system equations, thus restricting the results to linear systems. In addition, Butkovskii's maximum principle results in an optimal control in the form of a solution to a non-linear integral equation involving multiple integrals. Such an integral equation is not solvable in all cases.

The above deficiency was removed by Katz (K-2) who formulated a general maximum principle which could be applied to first order hyperbolic systems and parabolic systems as well as lumped parameter systems and did not depend on the prior representation of the system by integral equations. Egorov (E-1, E-2) presented necessary conditions for second-order hyperbolic systems and parabolic systems. The most complete definition of the control problem was given by Wang (W-2), and Wang and Tung (W-1), who introduced the concepts of controllability and observability and derived necessary conditions similar to those of Katz (K-2) and Egorov (E-1) based on Dynamic programming.

Several authors have recently considered necessary conditions for specific systems. This is because the multiplicity of possible control problems that can be conceived for distributed parameter systems is many orders of magnitude greater than for lumped parameter systems. Some of the reasons for this are:

1. The boundary control has no analog in the lumped parameter case,
2. The distributed and boundary controls are, in general, functions of spatial variables as well as of time,
3. There are many different ways of specifying the admissible controls,
4. The state depends on space as well as on time.

The different ways of specifying "what an optimum control is" may be elaborated as follows. The fixed terminal state problem for lumped parameter system means that the state has to take on  $N$  specified numerical values. For the distributed parameter case, the value of the state at the terminal time can be specified at every point in the space, only at certain points, or only in certain regions. Similarly there are many possible cost functionals or performance criteria. Finally there exists a very significant difference between treating ordinary differential equations and partial differential equations. With ordinary differential equations, a very nice uniformly applicable theory exists for treating an  $n$ th order differential equation as an initial value problem. With partial differential equations different classes of equations, even of the same order, have very different characteristics and must be treated differently in each case.

Brogan (B-1, B-2) extended Butkovskii's maximum principle to systems with non-homogeneous boundary conditions.

Axelband (A-1) obtained the eigenfunction expansion for the control of linear distributed systems. McCausland (M-1) used a Fourier series representation of the temperature distribution in a slab to select the input heating to bring the spatial distribution harmonics of the error distribution in a slab to zero. Linear and non-linear programming schemes were proposed by Sakawa (S-1, S-2) to solve approximately an integral equation resulting from Butkovskii's maximum principle.

The studies cited above have been based in general on the linearity of the system or the ability to solve the system equation analytically. Very little work is reported in the area of non-linear distributed parameter systems. Denn (D-1) studied a non-linear distributed control problem using variational methods and showed the linear system as a special case.

As in the lumped parameter systems the variational calculus often yields the form of optimal control rather easily, but the complete synthesis of optimal controls is a major problem. Seinfeld and Lapidus (S-7) applied direct search and steepest ascent methods for solving a class of systems described by first order hyperbolic and parabolic equations. Wismer (W-3) applied multilevel optimization techniques to a diffusion system and stated that general convergence theorems are difficult to prove. Sage and Chaudhuri (S-3) discussed the spatial and time discretization schemes for approximately solving the problems in distributed systems by the known techniques of lumped systems.

Thus far we have considered only the case of obtaining an open loop control law. Very little has been reported in the synthesis of feedback controls for distributed parameter systems. Seinfeld and Kumar (S-6) first obtained the sub-optimal feedback controls for a class of distributed parameter systems. Their method of obtaining the sub-optimal feedback controls is based on the existence of the optimal open loop solution. The feedback parameters are determined by choosing a criterion that yields system performance that in some manner approximates the optimal open loop behaviour. Koivo and Kruh (K-5) used the same criterion for the design of feedback controller but deviated from the above, by using a gradient technique in the parameter space to determine the optimal feedback parameters. Both used discretization of the space variables for computational purposes.

The disadvantage in all the above cases is the complexity of the computations because of the increased dimensionality inherent in these systems. The dimensionality is increased as the spatial discretization step becomes smaller. Either we can discretize the necessary conditions or we can discretize the original partial differential equations. Wang raises the question of relative merit between these two types of discretizations.

### 1.3 Contribution of the Dissertation

In this dissertation, efficient computational procedures for obtaining optimal feedback control of distributed parameter systems are given. The first method uses no prior information

about the optimal open-loop control. A finite difference scheme is used to approximate the infinite dimensional system by a finite dimensional system. Then a multidimensional parameter optimization technique is used to obtain the constant feedback gains. (Chapter 5)

One of the difficulties in parameter optimization is the large dimensionality of the approximate differential system. The number of integrators available on any analog computer installation is limited by the complexities involved in maintenance. So a decomposition principle, which divides the large set of differential equations due to the above approximation, into lower order independent subsystems is stated. The solution in this case is obtained by an iterative technique. The convergence theorems are proved and the theory is illustrated with several examples. (Chapter 4)

In the second method, a priori information about the optimal open loop control is used to obtain time varying gains in contrast to the fixed gains. The implementation of this method on the hybrid computer is straightforward. The time varying gains are obtained by sequentially solving parameter optimization problems with differential constraints. The number of parameters to be determined is equivalent to the number of state functions in the distributed parameter system. The method terminates when the desirable performance is obtained. (Chapter 6)

#### 1.4 Problem Formulation

The main prerequisites for the analytical design of an optimum control system consists of:

- i) establishing an adequate mathematical model of the physical systems to be controlled,
- ii) determining the constraints imposed by physical limitations and design specifications, and then expressing them in terms of the pertinent physical variables,
- iii) selecting a realistic performance index.

##### 1.4.1 Mathematical model

The dynamical behaviour of distributed parameter systems can be described by a system of partial differential equations or a set of non-linear integral equations, which result, in general, from the solution of linear partial differential equations. This thesis considers only distributed parameter systems described by the partial differential equations of the form:

$$\frac{\partial Q(x,t)}{\partial t} = G(Q(x,t), m(x,t), x, t) \quad (1.1)$$

$$Q(x, t_0) = Q_0(x) \quad , \quad x \in \Omega \quad (1.2)$$

$$S_b Q(x_b, t) = u(x_b, t) \quad , \quad x_b \in \Omega_b \quad (1.3)$$

where  $G$  has continuous first order derivatives with respect to  $x$  and  $t$  and is twice continuously differentiable with respect to the remaining arguments. In the above equation,

the following symbols are used:

$Q(x,t) = Q(x_1, x_2, x_3, \dots, x_n, t)$ , a  $p$ -dimensional state variable

$m(x,t) = m(x_1, x_2, x_3, \dots, x_n, t)$ , a  $q$ -dimensional control variable

$u(x_b, t) = u(t)$ , an  $r$ -dimensional boundary control variable

independent of the space variable  $x$ .

$\Omega$  = a given finite (connected) region in Euclidean  $n$ -space;

and  $\Omega_b$ , the boundary of  $\Omega$ .

$S_b$  = a linear operator.

It can be seen from above that the state variable is not only a function of time, but also function of spatial domain. Thus the state of the dynamic system at any fixed time  $t$  can be generally specified by a set of functions  $\{Q_i(x,t), i = 1, \dots, p\}$ , defined for all  $x \in \Omega$ . The set of all possible functions of  $x$  defined on  $\Omega$ , that  $Q_i(x,t)$  can be any time  $t$ , will be called the state component function space  $\Gamma_i$ , and the product space  $\Gamma = \Gamma_1 \times \Gamma_2 \times \Gamma_3 \times \dots \times \Gamma_p$  will be called the state function space. This definition is similar to the state space in the case of lumped parameter systems.

The possible control variables can be placed conveniently in two classes.

a) Distributed controls  $m(x,t)$ , a  $q$ -dimensional control variable,

b) Boundary controls  $u(t) = (u_1(t), \dots, u_r(t))$ , where  $m(x,t)$  and  $u(t)$  are piecewise continuous functions of their arguments and are allowed to assume values from bounded convex regions  $V$  and  $W$  respectively. Any control that belongs to

these convex regions is called an admissible control.

Finally, we will assume that all the problems considered in this study are "well-posed" and thus possess the following properties.

- a) The solutions to Equations 1-1, 1-2 and 1-3 exist.
- b) The solutions are uniquely determined.
- c) The solution depends continuously on the initial data. This says that small changes in the initial data will cause correspondingly small changes in the solution  $Q(x,t)$ .

#### 1.4.2 Constraints

In distributed parameter systems, the constraints may be related to dynamic variables defined on certain subsets or all of the spatial domain  $\bar{\Omega}$ . They are essentially equality and inequality constraints.

The class of equality constraints is of the form,

$$Z[x,t,Q(x,t),m(x,t)] = 0$$

where  $Z$  is a vector functional of its arguments defined on certain subsets or all of  $\bar{\Omega}$ ; where  $\bar{\Omega}$  is the closure of  $\Omega$ . Typical examples are:

- 1) Boundary conditions which represent certain interactions between the dynamic system and its environment,
- 2) Physical quantities which are expressible as functionals of the system dynamic variables that may be required to remain invariant during the course of motion. A possible form of this

constraint is:

$$\int_{\Omega} F[x, Q(x, t), m(x, t)] d\Omega = \text{constant}$$

where  $F$  is a specified function of its arguments.

The class of inequality constraints is of the form:

$$g_l \leq R[x, t, Q(x, t), m(x, t)] \leq g_u$$

where  $g_l$  and  $g_u$  may be either functions of time,  $t$ , and/or the spatial variable  $x$  or constants. Typical examples are bounded state functions of the form:

$$\max_{x \in \Omega} |Q_i(x, t)| \leq M_i = \text{constant},$$

and bounded control variables of the form:

$$|m(x, t)| \leq g_i(x) \quad \text{almost everywhere on } \bar{\Omega}$$

or

$$|u(t)| \leq M = \text{constant}.$$

#### 1.4.3 Performance Indices

A generalized integral performance index for distributed parameter systems with fixed terminal time  $T$  can be written in the form:

$$C_I = \int_0^T \int_{\Omega} [P_1(t, x, Q_d(x, t), Q(x, t), m(x, t))] d\Omega dt \quad (1.4)$$

For terminal control where the final time  $T$  is fixed, a performance index can be defined in the form of a spatial integral:

$$C_T = \int_{\Omega} P_o(Q(x, t), T, x) d\Omega \quad (1.5)$$

The problem of minimizing (maximizing) a performance index in the form of Eq. (1.4) can be reduced to a terminal control problem by defining

$$Q_1(x, t) = \int_0^t [P_1(\tau, x, Q_d(x, \tau), Q(x, \tau), m(x, \tau))] d\tau \quad (1.6)$$

$$\text{and then } Q_1(x, T) = \int_0^T P_1 d\tau \quad (1.7)$$

where  $Q_d(x, t)$  is the desired state. Thus Eq. (1.4) is transformed into the form of (1.5), that is

$$C_I = \int_{\Omega} Q_1(x, T) d\Omega \quad (1.8)$$

In other words  $C_I$  represents the optimal transfer of the initial spatial distribution to a final desired distribution in a specified time. We have seen that with  $T$  fixed,  $C_I$  could be transformed to  $C_T$ , and in case  $T$  is free, we seek the first time when the state lies in some given  $\epsilon$ -neighborhood of the desired state. If it is necessary and possible to choose an admissible control such that the phase trajectory in function space exactly coincides with a desired distribution at  $T$ , then the trajectory connecting the initial and desired states is unique and hence optimal. This brings us to the concept of controllability.

### 1.5 Controllability

In any control problem it is important to consider the question "Can any initial state of a given system be transferred to any desired state in a finite period of time by admissible

control action?" We follow the definitions given by Wang and Tung (W-1).

Let  $\phi(t, x, Q(x, t_0), t_0)$  be a solution of (1.1) with specified input functions and boundary conditions given in (1.2) and (1.3). Then  $\phi$  satisfies the following:

- i)  $\phi[t_0, x, Q(x, t_0), t_0] = Q(x, t_0)$
- ii)  $\frac{\partial \phi}{\partial t} = G[\phi(x, t, Q(x, t_0), t_0), x, t, m(x, t)]$
- iii)  $\phi[t, x_b, Q(x, t_0), t_0] = u(t, x_b)$

The initial state of a distributed system  $Q(x, t_0)$  is said to be null controllable at time  $t_0$ , if there exist admissible controls (see Section 1.4.1)  $m(x, t)$  and  $u(t)$  that will transfer  $Q(x, t_0)$  to the null state in a finite time  $T$ ; that is, the solution

$$\phi[t_0 + T, x, Q(x, t_0), t_0] = 0 \text{ almost everywhere in } \Omega.$$

In general  $T$  depends upon both  $t_0$ , and  $Q(x, t_0)$ . The initial state is null  $\delta$ -controllable at time  $t_0$ , if

$$\|\phi[t_0 + T, x, Q(x, t_0), t_0]\| \leq \delta$$

where the norm is a spatial norm and a typical spatial norm is

$$\|\phi\| = \left[ \int_{\Omega} \phi^t B \phi \, d\Omega \right]^{\frac{1}{2}}$$

Obviously a null controllable state is also null  $\delta$ -controllable.

However, the converse is not necessarily true.

In many systems, only the states belonging to  $\Gamma'$  (a subset of the state function space  $\Gamma$  containing the null state) are null controllable. This fact leads to the following definitions:

A distributed parameter system is said to be completely null controllable at time  $t_0$  in  $\Gamma'$ ,  $\Gamma' \subseteq \Gamma$ , if there exist admissible input functions which will transfer every state in  $\Gamma'$  to the null state in finite time. (See Figure 1.1).

Similarly we can define complete null  $\delta$ -controllability in  $\Gamma'$ . Here, the null state must be an interior point of  $\Gamma'$ . (See Figure 1.2).

By imposing the condition that the terminal state is an arbitrary element in  $\Gamma'$ , we have the stronger types of controllability namely: Complete controllability in  $\Gamma'$  and Complete  $\delta$ -controllability in  $\Gamma'$ . (See Figure 1.3).

The notion of  $\delta$ -controllability is useful when dealing with approximate systems. For example the following result is true.

If a convergent approximate system is completely controllable, then the exact system is completely  $\delta$ -controllable. This follows directly from the definition of the convergent approximation, that is, for a given level of discretization, the solutions of the approximate systems are within some  $\epsilon$ -neighborhood of the exact solution.

Finally if the distributed system is asymptotically stable about the null state for all initial states in  $\Gamma'$ , then the system is completely null  $\delta$ -controllable in  $\Gamma'$ .

## 1.6 Observability

The notion of observability of a dynamical system is associated with processing of data obtained from observations on the system. Thus the basic question is:

Given a mathematical model of a free dynamical system (Control  $m(x,t) = 0$  and  $u(t) = 0$ ) and the output transformation  $\mathcal{M}$ , is it possible to determine the system state at any time  $t$  by observing the output over a finite time interval,  $(t, t+T)$ , where  $T$  may depend on the system properties and the output transformation  $\mathcal{M}$ ?

### 1.6.1 Definitions

Let  $\theta_S(T)$  be a finite tube of system trajectories (with no distributed or boundary control)  $\phi[t, x, Q(x, t_0), t_0]$  defined on time interval  $(t_0, t_0+T)$  and with  $Q(x, t_0) \in \Gamma'(t_0)$  the initial section of  $\theta_S(T)$  (a subset of the state function space) (See Figure 1.4). Let  $\theta_O(T)$  be the tube of output trajectories corresponding to a given continuous output transformation  $\mathcal{M}$  of all the trajectories in  $\theta_S(T)$  (See Figure 1.5).

A distributed parameter system is said to be completely observable in  $\Gamma'(t_0)$  at time  $t_0$ , if there exists a finite time  $T$  and a one to one continuous mapping from  $\theta_O(T)$  to  $\Gamma'(t_0)$ . If in addition to the above conditions,  $\Gamma'(t_0) = \Gamma$ , then the system is said to be completely observable.

In contrast to the lumped systems, there are no precise mathematical conditions to test the controllability and

observability properties for a general class of distributed parameter systems. The controllability is associated with the ability of steering one system state to another in a finite amount of time by means of certain admissible controls. The lack of general methods to test this property justifies the consideration of optimal end-point control problems. The study of feedback control requires that the system be observable, that is, it is possible to determine the system state completely at any time from a finite amount of observed output data.

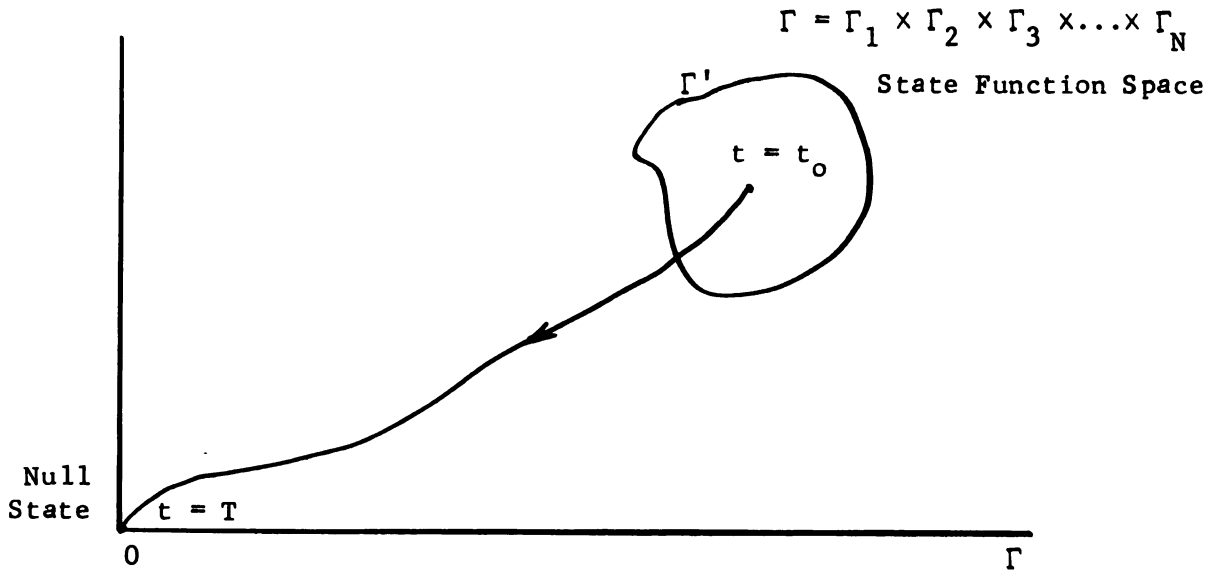


Figure 1.1 Complete Null Controllability at Time  $t_0$  in  $\Gamma' \subseteq \Gamma$

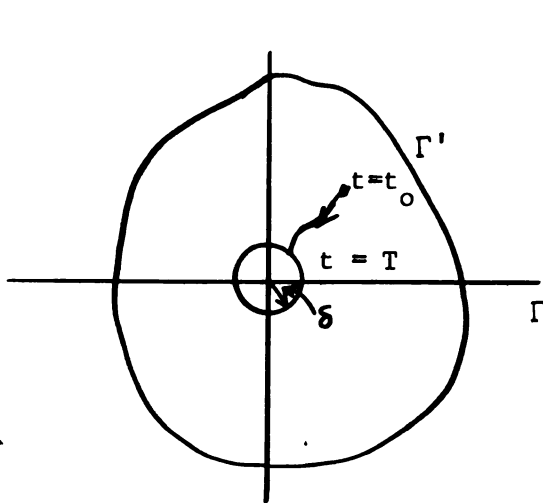


Figure 1.2 Complete Null  $\delta$ -Controlability at  $t_0$  in  $\Gamma' \subseteq \Gamma$

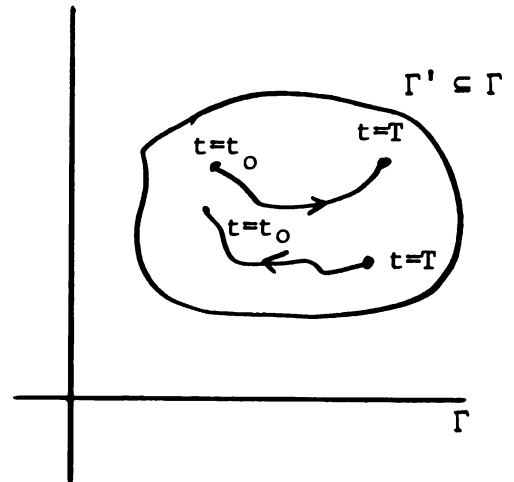


Figure 1.3 Complete Controllability in  $\Gamma'$

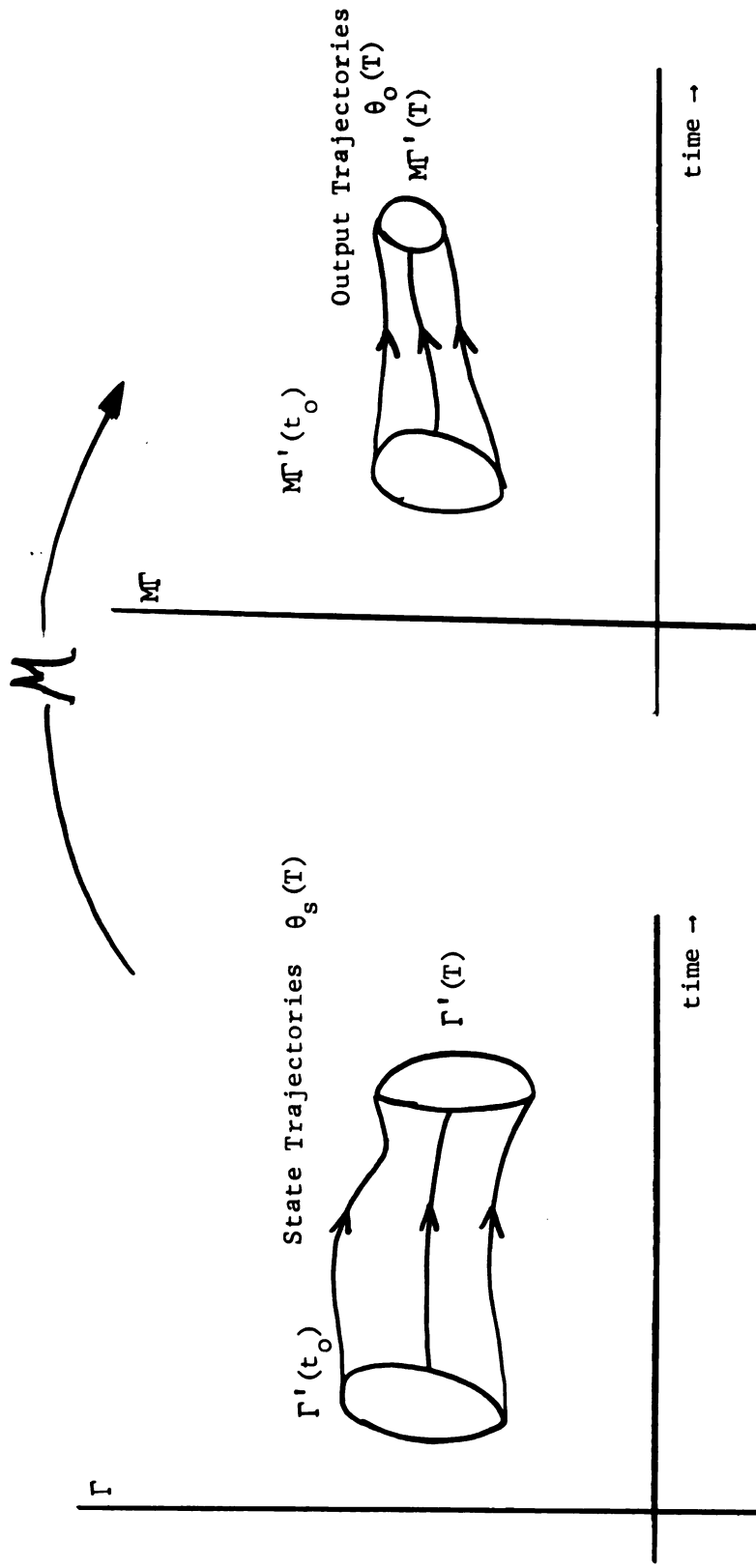


Figure 1.4 Finite Tube of System Trajectories

Figure 1.5 Output Trajectories Corresponding to Output Transformation  $M$  in the Transformed Space

## CHAPTER II

### ANALYTICAL SOLUTIONS FOR THE OPTIMAL ENDPOINT CONTROL PROBLEMS

#### 2.1 Introduction

In this chapter we obtain the necessary conditions and analytical solutions for fixed time, free terminal state problems, where the differential constraints are either in the form of linear partial differential equations or non-linear partial differential equations with proper boundary conditions.

#### 2.2 Necessary Conditions for Optimality

Before proceeding to obtain analytical solutions for these problems, the necessary conditions are derived by using the dynamic programming approach. This approach was used by Wang and Tung (W-1) where integral constraints were considered. Brogan (B-1) also applied the dynamic programming approach but instead of integral constraints, he considered the differential constraints in the form of linear partial differential equations. The following derivation closely follows the derivation of Brogan (B-2). Let us consider the cost functional

$$C \triangleq \int_{\Omega} P_0(Q(x, t_f), t_f) d\Omega + \int_{t_0}^{t_f} \int_{\Omega} P_1(Q(x, t), m(x, t), u(t), t) d\Omega dt \quad (2.1)$$

subject to the linear partial differential equations with side conditions

$$\partial Q / \partial t = S_x Q(x, t) + D(x, t) m(x, t) \quad (2.2)$$

$$Q(x, t_0) = Q_0(x) \quad (2.3)$$

$$S_b Q(x_b, t) = u(t) \quad (2.4)$$

The optimal control problem is to find admissible controls  $m(x, t)$  and/or  $u(t)$  so that the performance index is optimized (minimized).

Let the minimum of (2.1) be denoted by

$$\min_{F \in \mathcal{A}} C \triangleq \Pi(Q(x, t_0), t_f - t_0) \quad (2.5)$$

where  $F$  denotes the general forcing function, and  $\mathcal{A}$  is the set of admissible controls. (Eq. 2.2 can be obtained in that form by using the extended operator and thus converting the non-homogeneous boundary conditions to homogeneous boundary conditions, Brogan (B-2)).

Now (2.5) becomes

$$\Pi(Q(x, t_0), t_f - t_0) = \min_{F \in \mathcal{A}} \left\{ \int_{\Omega} P_0 d\Omega + \int_{t_0}^{t_f} \int_{\Omega} P_1 d\Omega dt \right\} \quad (2.6)$$

$$\begin{aligned} &= \min_{F \in \mathcal{A}} \left\{ \int_{\Omega} P_0(Q(x, t), t_f) d\Omega \right. \\ &\quad \left. + \int_{t_0+\epsilon}^{t_f} \int_{\Omega} P_1 d\Omega dt + \int_{t_0}^{t_0+\epsilon} \int_{\Omega} P_1 d\Omega dt \right\} \end{aligned} \quad (2.7)$$

where

$$P_0 \triangleq P_0(Q(x, t), t_f)$$

$$P_1 \triangleq P_1(Q(x, t), x, t)$$

where  $F(x, t)$  denotes a general control variable.

Using Bellman's principle of optimality (B-9), if the cost  $C$  is to be a minimum during the total period  $(t_o, Q(x, t_o))$  to  $(t_f, Q(x, t_f))$ , then it is necessary that the cost incurred during the shorter interval  $(t_o + \epsilon, Q(x, t_o + \epsilon))$  to  $(t_f, Q(x, t_f))$  be minimum also. The cost during this later interval is equal to the sum of the first two integrals in (2.7) so that

$$\begin{aligned} \Pi(Q(x, t_o), t_f - t_o) = \min_{F \in \mathcal{A}} \left\{ \int_{t_o}^{t_o + \epsilon} \int_{\Omega} P_1 d\Omega dt \right. \\ \left. + \Pi(Q(x, t_o + \epsilon), t_f - t_o - \epsilon) \right\} \quad (2.8) \end{aligned}$$

The minimization in (2.8) is to be performed by optimizing the first increment of the control  $F(x, t)$ . After some manipulations Brogan (B-2) has shown that the necessary condition for optimality is,

$$\frac{\partial \Pi}{\partial T}(Q(x, t), T) = \min_{F \in \mathcal{A}} \int_{\Omega} \left\{ P_1(Q(x, t), F(x, t), t) + \left( \frac{\delta \Pi}{\delta Q} \right)^t \cdot \frac{\partial Q}{\partial t} \right\} d\Omega \quad (2.9)$$

In (2.9),  $T = t_f - t$  has the meaning of time to go, or time remaining to apply control to the system, and  $\delta \Pi / \delta Q$  is the functional derivative (G-1). In view of the definition in (2.5) for  $\Pi$ , (2.1) gives the initial condition for the differential system in (2.9)

$$\Pi[Q(x, t_f), T = 0] = \int_{\Omega} P_o(Q(x, t_f), t_f) d\Omega \quad (2.10)$$

To simplify the notation, let  $Y(x, t) = \delta \Pi / \delta Q$ . The vector  $Y(x, t)$  has  $N$  components, the same as  $Q(x, t)$ . Now the

(N+1)th component can be added to Y and Q as is done in the lumped parameter systems. Let

$$U(t,x) \triangleq \begin{bmatrix} \partial Q / \partial t \\ P_1 \end{bmatrix} \quad (2.11)$$

$$P(x,t) = \begin{bmatrix} Y(x,t) \\ 1 \end{bmatrix}$$

Now (2.9) can be written as

$$\frac{\partial \Pi}{\partial t} = \min_{F \in \mathcal{A}} \int_{\Omega} P^t(x,t) U(x,t) d\Omega. \quad (2.12)$$

When P and U are members of a Hilbert space, the inner product notation can be used to define the Pre-Hamiltonian H

$$H(Q,P,F,t) \triangleq \int_{\Omega} P^t U d\Omega = \langle P, U \rangle_{\Omega} \quad (2.13)$$

The Hamiltonian  $H^0$  is defined by

$$H^0(Q,P,t) = \min_{F \in \mathcal{A}} H(Q,P,F,t) \quad (2.14)$$

Thus Eq. (2.14) represents the minimum principle, which obviously could be rewritten as maximum principle by a change of sign in the definition of P. Equation (2.12) can now be written in the form of the Hamilton-Jacobi equation

$$\frac{\partial \Pi}{\partial T} = H^0(Q,P,t) \quad (2.15)$$

with the initial conditions given by (2.10). A pair of partial differential equations analogous to Hamilton's canonical equations can be found which are equivalent to the Hamilton-Jacobi

equation.

From the definition of  $Y(x,t)$ ,

$$\frac{\partial Y}{\partial t} = \frac{\partial}{\partial t} \left( \frac{\delta \Pi}{\delta Q} \right) = \frac{\delta}{\delta Q} \left( \frac{\partial \Pi}{\partial t} \right) \quad (2.16)$$

Making use of (2.15) and noting that  $T = t_f - t$ , we obtain

$$\frac{\partial \Pi}{\partial t} = \frac{\partial \Pi}{\partial T} \cdot \frac{\partial T}{\partial t} = - \frac{\partial \Pi}{\partial T} = -H^0 \quad (2.17)$$

Therefore,

$$\frac{\partial Y}{\partial t} = - \delta H^0 / \delta Q \quad (2.18)$$

Directly from the definition of  $H$ , it is seen that

$$\frac{\delta H}{\delta Y} = \partial Q / \partial t \quad (2.19)$$

The above equation can be shown as follows:

For fixed  $x \in \Omega$ ,  $Y(x,t)$ ,  $Q(x,t)$  are vector functions of time. Now the first variation of  $H$  with respect to  $Y$  can be obtained as follows.

For nth order canonical equations, (2.13) gives

$$H \triangleq \int_{\Omega} P^t U \, d\Omega = \int_{\Omega} Y^t \frac{\partial Q}{\partial t} \, d\Omega$$

Now,

$$\begin{aligned} H(Y + \delta y) &= \int_{\Omega} (Y + \delta y)^t \cdot \frac{\partial Q}{\partial t} \, d\Omega \\ &= H(Y) + \int_{\Omega} \delta Y^t \cdot \frac{\partial Q}{\partial t} \, d\Omega \end{aligned}$$

As  $\|\delta y\| \rightarrow 0$ , we get for fixed  $x$ ,  $\delta H / \delta Y = \partial Q / \partial t$

Since  $x$  is any element of  $\Omega$ , we get (2.19). Kalman (K-7) showed that if the solution to (2.15) is analytic, then

$$\delta H / \delta Y = \delta H^0 / \delta Y \quad (2.20)$$

so that

$$\partial Q / \partial t = \delta H^0 / \delta Y \quad (2.21)$$

Thus the pair of  $n$ th order canonical equations, similar to lumped parameter case, are

$$\begin{aligned} \partial Q / \partial t &= \delta H^0 / \delta Y \\ \partial Y / \partial t &= -\delta H^0 / \delta Q \end{aligned} \quad (2.22)$$

The initial conditions for  $Q$  are  $Q(x, t_0) = Q_0(x)$ . The second set of conditions, for free terminal state problems, are

$$Y(x, t_f) = \frac{\delta \Pi}{\delta Q} (Q(x, t_f), T = 0) = \frac{\partial P_0}{\partial Q} \quad (2.23)$$

and if the terminal state is fixed then the other condition would be,

$$Q(x, t_f) = Q_d(x) \quad (2.24)$$

If the  $(n+1)$ th order canonical pair is desired, the  $(n+1)$  components are obtained from (2.11) as

$$\frac{\partial Q_{n+1}}{\partial t} = P_1 ; \quad \frac{\partial Y_{n+1}}{\partial t} = 0 \quad (2.25)$$

Since we are concerned with fixed time and free endpoint problems in this dissertation, we apply these necessary conditions

to a specific problem of this class.

### 2.3 Example

The problem considered here is to drive the temperature distribution in a one dimensional conducting body from its initial zero state to as near as possible to the desired distribution  $q_d(x)$  at a fixed time  $t$ , by forcing the temperature at one end of the body to have an optimal time history  $u(t)$ . The control  $u(t)$  is required to satisfy

$$|u(t)| \leq 1 \quad \text{for all } t \quad (2.26)$$

The cost function is

$$C = \int_0^1 (q_d(x) - q(t_1, x))^2 dx \quad (2.27)$$

The system equation is

$$\partial q / \partial t = \partial^2 q / \partial x^2 \quad (2.28)$$

with the initial and boundary conditions

$$q(x, t_0) = 0, \quad q(0, t) = 0, \quad q(1, t) = u(t) \quad (2.29)$$

Now applying the extended definition of the operator, Brogan (B-2), the non-homogeneous boundary conditions are converted to homogeneous boundary conditions to the system. Thus (2.28) and (2.29) reduce to

$$\begin{aligned} \partial q / \partial t &= \partial^2 q / \partial x^2 + \delta'(x - 1)u(t) \\ q(x, t_0) &= 0, \quad q(0, t) = q(1, t) = 0 \end{aligned} \quad (2.30)$$

where  $\delta(x - \ell)$  is an impulse function.

The Pre-Hamiltonian as defined in (2.13) is

$$\begin{aligned}
 H &= \int_0^{\ell} \frac{\delta \Pi}{\delta q} \left[ \frac{\partial^2 q}{\partial x^2} + \delta'(\xi - \ell) u(t) \right] d\xi \\
 &= \int_0^{\ell} \frac{\delta \Pi}{\delta q} \left[ \frac{\partial^2 q}{\partial x^2} \right] d\xi - \int_0^{\ell} \frac{\delta \Pi}{\delta q} (\delta'(\xi - \ell) u(t)) d\xi \\
 &= \int_0^{\ell} \frac{\delta \Pi}{\delta q} \left( \frac{\partial^2 q}{\partial x^2} \right) d\xi - \frac{d}{d\xi} \left( \frac{\delta \Pi}{\delta q} \right) \Big|_{\xi=\ell} u(t)
 \end{aligned} \tag{2.31}$$

The control  $u^0(t)$  which minimizes (2.31) subject to the constraint of (2.26) is, excluding the possibility of the singular control,

$$u^0(t) = \text{sgn} \left[ \frac{\partial}{\partial \xi} \left( \frac{\delta \Pi}{\delta q} \right) \Big|_{\xi=\ell} \right] \tag{2.32}$$

Now let  $Y(x, t) \triangleq (\delta \Pi / \delta q)$ . Then the necessary conditions yield

$$\partial Y / \partial t = -\delta H^0 / \delta q \tag{2.33}$$

subject to the condition at  $t_1$  that

$$Y(x, t) = \partial P / \partial q^0 = -2(q_d(x) - q(x, t_1)) \tag{2.34}$$

Integration by parts within (2.31), so that  $q(x, t)$  appears in undifferentiated form, facilitates finding that

$$\delta H^0 / \delta q = \partial^2 Y / \partial x^2 \tag{2.35}$$

Substituting  $T = t_1 - t$  into (2.33) we obtain,

$$\partial Y / \partial t = -\partial Y / \partial T = -\delta H^0 / \delta q = -\partial^2 Y / \partial x^2$$

Therefore

$$\partial Y / \partial T = \partial^2 Y / \partial x^2 \quad (2.36)$$

Thus (2.34) and (2.36) form an initial value problem for the diffusion equation. Thus the necessary conditions yield a two point boundary value problem which sometimes can be converted to initial value problem. Because of the complexity of the equations, the analytical solutions are very difficult to obtain, and thus computational methods are used to obtain the solutions.

Similar results can be obtained for this class of problems by using functional analysis.

#### 2.4 Application of Functional Analysis

Let the state space at a given time be denoted by  $H_2$  (Hilbert space with  $L_2$  norm) and let the control variable space be  $H_1$ . Then the solution to (2.2) with zero initial conditions and homogeneous boundary conditions represent a mapping of elements from  $H_1$  into  $H_2$ , and at time  $t_1$  can be written as,

$$Q(x, t_1) = L_{t_1} F \quad (2.37)$$

The cost function to be minimized is

$$\begin{aligned} C &= \int_{\Omega} [Q_d(x) - Q(x, t_1)]^t [Q_d(x) - Q(x, t_1)] d\Omega \\ &= \|Q_d(x) - Q(x, t_1)\|_{H_2}^2 \\ &= \|Q_d(x) - L_{t_1} F\|_{H_2}^2 \end{aligned}$$

$$\begin{aligned}
&= \langle L_{t_1} F, L_{t_1} F \rangle_{H_2} + \langle Q_d, Q_d \rangle_{H_2} - 2 \langle Q_d, L_{t_1} F \rangle_{H_2} \\
&= \langle F, L_{t_1}^* L_{t_1} F \rangle_{H_1} + \langle Q_d, Q_d \rangle_{H_2} - 2 \langle L_{t_1}^* Q_d, F \rangle_{H_1} \\
&= \langle F, L_{t_1}^* L_{t_1} F - 2 L_{t_1}^* Q_d \rangle_{H_1} + \langle Q_d, Q_d \rangle_{H_2} \quad (2.38)
\end{aligned}$$

where  $L_{t_1}^*$  is an adjoint of  $L_{t_1}$ .

Now the cost  $C$  will be minimized if the term depending on  $F$  is minimized, since  $Q_d$  is fixed. If  $F^0$  is the optimal control, then any other control  $F = F^0 + \epsilon \bar{F}$  will satisfy

$$\|Q_d - L_{t_1}(F^0 + \epsilon \bar{F})\|_{H_2}^2 \geq \|Q_d - L_{t_1} F^0\|_{H_2}^2$$

or

$$\langle F^0 + \epsilon \bar{F}, L_{t_1}^* L_{t_1} (F^0 + \epsilon \bar{F}) - 2 L_{t_1}^* Q_d \rangle_{H_1} \geq \langle F^0, L_{t_1}^* L_{t_1} F^0 - 2 L_{t_1}^* Q_d \rangle_{H_1}$$

or

$$\epsilon \langle \bar{F}, 2 L_{t_1}^* L_{t_1} F^0 - 2 L_{t_1}^* Q_d \rangle + \epsilon^2 \langle \bar{F}, L_{t_1}^* L_{t_1} \bar{F} \rangle \geq 0 \quad (2.39)$$

But since

$$\langle \bar{F}, L_{t_1}^* L_{t_1} \bar{F} \rangle = \langle L_{t_1} \bar{F}, L_{t_1} \bar{F} \rangle = \|L_{t_1} \bar{F}\|^2 \geq 0$$

Equation (2.39) requires that

$$2\epsilon \langle \bar{F}, L_{t_1}^* L_{t_1} F^0 - L_{t_1}^* Q_d \rangle \geq 0 \quad (2.40)$$

for arbitrary  $\epsilon$ , and so

$$L_{t_1}^* L_{t_1} F^0 = L_{t_1}^* Q_d \quad (2.41)$$

is the necessary condition for the optimality which is in the form of an integral equation and this is the same as the solution of the two point boundary value problem expressed in the integral form.

## CHAPTER III

### EXISTING METHODS OF COMPUTING OPTIMAL CONTROL

#### 3.1 Introduction

In this chapter, we will develop some of the existing computational techniques for obtaining optimal open loop and closed loop controls for distributed parameter systems. These computational methods have been developed because of the difficulties encountered in solving these problems analytically as shown in Chapter II.

#### 3.2 Existing Methods for Optimal Open Loop Control

As a consequence of the necessary conditions for optimality discussed in Chapter II, two point boundary value problems in terms of partial differential equations are obtained. This is similar to the lumped parameter case when Brogan's (B-2) extended operator method is used to reduce the multiple boundary value problems into initial value problems. Sage and Chaudhuri (S-4) spatially discretized the necessary conditions and applied the gradient and quasilinearization techniques available for lumped parameter systems. The gradient method is based on iteration on an assumed control trajectory to improve continuously the performance index. The quasilinearization technique linearizes the state equations to generate a sequence of

convergent approximations to the actual trajectory while retaining the boundary conditions (for example, the Newton-Raphson method in function space). There are some other methods called shooting methods where the boundary conditions are iterated upon, while the actual state and adjoint equations are retained. Seinfeld and Lapidus (S-7) developed two methods called Direct Search technique and Steepest Descent method for boundary value problems described by partial differential equations. The steepest descent method is an extension of Bryson's method of steepest descent for optimal control problems in lumped parameter systems. It is a gradient method based on small perturbations about a nominal trajectory. Sakawa (S-1) converted the optimal control problem into a non-linear programming problem. Khatri and Goodson (K-1) discussed approximate methods of solving a class of optimal control problems using calculus of variations. Their approximation consists of harmonic truncation in the S-domain.

In the next section, the methods given by Sakawa (S-1) and Seinfeld and Lapidus (S-7) are discussed in detail, since some of the results obtained from these methods are utilized in the computation of feedback control for example problems in Chapter VI. As in the case of lumped parameter systems, each of the above methods have their advantages and disadvantages and no one method would serve as the best choice for all the types of problems involved.

### 3.3.1 Sakawa's Method

Sakawa's method of obtaining the optimal control can best be illustrated by an example. The process of one sided heating of a metal in a furnace is described by a diffusion equation.

$$\frac{\partial q}{\partial t} = \frac{\partial^2 q}{\partial x^2} \quad (3.1)$$

with the boundary conditions

$$q(x,0) = 0$$

$$\left. \frac{\partial q}{\partial x} \right|_{x=0} = \alpha \{ q(0,t) - v(t) \} \quad (3.2)$$

$$\left. \frac{\partial q}{\partial x} \right|_{x=1} = 0$$

and the temperature  $v(t)$  is controlled by the fuel flow  $u(t)$  and satisfies the following differential equation:

$$r \frac{dv}{dt} + v(t) = u(t) \quad (3.3)$$

and

$$0 \leq u(t) \leq 1 \quad (3.4)$$

where  $r$  is the time constant of the furnace and  $u(t)$  is normalized properly. The performance is:

$$J[u(t)] = \int_0^1 \{ q^*(x) - q(x,T) \}^2 dx \quad (3.5)$$

where  $q^*(x)$  is desired distribution and  $q(x,T)$  is the actual distribution at time  $t = T$ . Equation (3.1) along with the boundary conditions (3.2) can be converted to an integral

equation,

$$q(x, T) = \int_0^T g(x, T - t) u(t) dt \quad (3.6)$$

where

$$g(x, t) = \frac{k^2 \cos k(1-x)}{\cos k - \frac{k}{\alpha} \sin k} e^{-k^2 t} + 2k^2 \sum_{i=1}^{\infty} \frac{\cos (1-x)\beta_i}{(k^2 - \beta_i^2) \left( \frac{1}{\alpha} + \frac{1+\alpha}{2\beta_i} \right) \cos \beta_i} \quad (3.7)$$

where

$$k = \frac{1}{\sqrt{r}} \quad \text{and} \quad \beta_i \quad \text{are the roots of} \quad \beta \tan \beta = \alpha$$

Thus the optimal control problem can be stated as follows:

Given (3.6) and the constraint  $0 \leq u(t) \leq 1$  on the interval  $0 \leq t \leq T$ , find  $u(t)$  such that the performance index given in (3.5) is minimized.

Now the conversion of this problem into a non-linear programming problem is given. After applying numerical integration formula to (3.5), the approximate performance index  $DJ(u)$  is expressed as

$$J[u] \approx DJ[u] = \sum_{i=0}^n C_i \{q^*(x_i) - q(x_i, T)\}^2 \quad (3.8)$$

where  $C_i$ 's are the weights assigned to the values of integrand at the point  $x_i$ . The values of  $x_i$  and the weights  $C_i$  are known for each integration formula. As an example, if the Simpson's rule is used, the values of  $x_i$ 's and  $C_i$  are given from standard tables as

$$\begin{aligned}
x_i &= i/n \quad (i = 0, 1, 2, \dots, n) \\
C_0 &= C_n = 1/3n \\
C_1 &= C_3 = \dots = C_{n-1} = 4/3n \\
C_2 &= C_4 = \dots = C_{n-2} = 2/3n
\end{aligned} \tag{3.9}$$

where  $n$  is an even number.

Applying the same integration formula to (3.6), the approximate value of  $q(x_i, T)$  is given by,

$$q(x_i, T) \approx \bar{q}(x_i, T) = T \sum_{j=0}^n C_j g(x_i, T - \tau_j) u(\tau_j) \tag{3.10}$$

where

$$\tau_j = jT/n \quad (j = 0, 1, \dots, n)$$

Putting

$$\begin{aligned}
TC_j g(x_i, T - \tau_j) &= a_{ij} \\
u(\tau_j) &= u_j \\
q^*(x_i) &= q_i^*
\end{aligned} \tag{3.11}$$

and substituting (3.10) into (3.8) yields,

$$DJ[u] \approx F[u] = \sum_{i=0}^n C_i (q_i^* - \sum_{j=0}^n a_{ij} u_j)^2 \tag{3.12}$$

The constraint in (3.4) is written as

$$0 \leq u_j \leq 1 \quad (j = 0, 1, \dots, n) \tag{3.13}$$

Consequently, the minimization problem of the functional in Eq. (3.5) is approximately reduced to a minimization of the function in (3.12) of  $n+1$  variables  $u_j$ 's subject to the

constraints of (3.13). Thus the optimal control problem is reduced to a quadratic programming problem in this case and the solution can be obtained with the known methods.

### 3.3.2 Direct Search on the Performance Index

Sakawa's algorithm essentially gives solutions for linear problems that could be transformed to integral equations. Seinfeld and Lapidus (S-7) extended the Direct Search method of lumped parameter systems. The ease of handling non-linearities and control constraints as well as the success in handling singular problems make the method attractive. Let us discuss this method briefly, and note the advantages and disadvantages of this method over the others.

Assume that the interval  $(0, t_f)$  is divided into  $L$ -segments and  $(0, 1)$  is divided into  $N$ -segments. We select  $u_k(t)$ ,  $k = 1, 2, \dots, q$ . The direct search algorithm can be outlined as follows:

1) Guess  $u_1^0(t), u_2^0(t), \dots, u_q^0(t)$ , the starting control functions, where  $u_k(t)$ ,  $k = 1, 2, \dots, q$  are the boundary controls.

2) The system equations are integrated over the given domain with these starting control functions, to obtain the value of the performance index  $P^0$ .

3) Now allow  $u_1(t)$  to vary and find  $u_1(t)$  which minimizes the performance index,

$$P[q(x, t_f), t_f] = F[q(x, t_f), q_d(x, t_f), u_1, u_2, \dots, u_q(t)]$$

In other words, fix  $u_2(t), u_3(t), \dots, u_q(t)$  at the assumed

7

values  $u_2^0(t), u_3^0(t), \dots, u_q^0(t)$  and vary only  $u_1(t)$  until that value is found which minimizes the performance index.

We call the resulting index  $P^1$  with the control vector

$$u^1(t) = (u_1^1(t), u_2^0(t), \dots, u_q^0(t))$$

At this point note that

$$P^1 \leq P^0 \quad (3.14)$$

since at worst the control  $u_1^1(t) = u_1^0(t)$  is obtained.

4) Repeat the procedure for  $u_k(t)$ ,  $k = 2, \dots, q$ .

5) Return to  $k = 1$  and repeat the steps (1) through (4) to obtain any improvement in the performance index so that the consecutive values are within a prescribed error bound.

The direct search on the performance index offers the following advantages for the distributed parameter systems.

1) Minimum storage capacity is required since only the last  $P^j$  and the last control function

$$u^j(t) = (u_1^j(t), \dots, u_q^j(t))$$

has to be retained.

2) Control constraints are handled simply.

3) Knowledge of the variational formulation and the two point boundary value problem is not required to use this method.

4) Non-linear systems are handled in the same way as linear systems.

The disadvantage of this method is the excessive amount of computation time because of the large number of integrations to be performed. As is the case in all the other methods, the convergence to a global extremum has not been proven in the general case.

In contrast to the open-loop control, there are not many studies on the computation of feedback controls for distributed parameter systems. In the next section, the existing methods for obtaining the feedback control for terminal optimal control problems is presented.

### 3.4 Existing Methods for Obtaining Optimal Feedback Control

The techniques developed in section 3.3 result in a control which is a function of the independent variables and the initial conditions, a so called open-loop control. From an engineering point of view, it is desirable to have the optimal control as a function of the state, and possibly time, such a control is usually called a feedback control law.

Seinfeld and Kumar (S-6) first obtained the sub-optimal feedback controls, for a class of distributed parameter systems. Their method requires the existence of the optimal open-loop control. The feedback parameters are chosen by minimizing a system performance which in some manner approximates the optimal behavior. Koivo and Kruh (K-5) used the same criterion for the design of feedback controller but deviated from (S-6) in the actual design procedure. They used the gradient technique in parameter space. This method requires the transformation

of the syst

two methods

3.4

Cor

defined on

be denoted

Co

$\frac{\partial Q}{\partial t}$

$\frac{\partial Q}{\partial t}$

along with

where  $Q(\cdot)$

distribut

vector,  $a$

condition

and  $m(x)$

and  $v$ .

followin

function

traject

of the system into corresponding integral equation form. These two methods are discussed next.

#### 3.4.1 Seinfeld and Kumar's method

Consider a parabolic or first order hyperbolic system defined on a fixed spatial domain  $\Omega$ . Let the boundary of  $\Omega$  be denoted by  $\Omega_b$ .

Consider the system described by

$$\frac{\partial Q}{\partial t} = G[t, x, Q(x, t), Q_x(x, t), Q_{xx}(x, t), u(t), m(x, t)] \quad (3.15)$$

along with the boundary conditions

$$Q(x, t_0) = Q_0(x) \quad x \in \Omega, \quad t \in [0, t_f] \quad (3.16)$$

$$S_b Q(x_b, t) = u(x_b, t) \quad x \in \Omega_b, \quad t \in [0, t_f] \quad (3.17)$$

where  $Q(x, t)$  is the  $p$ -dimensional state vector,  $m(x, t)$  the distributed control,  $u(t)$  the boundary control an  $m$ -dimensional vector, and (3.16) and (3.17) represent the initial and boundary conditions respectively. In addition, we may constrain  $u(t)$  and  $m(x, t)$  to assume values from bounded convex regions  $W$  and  $V$ .

The open-loop optimal control problem is posed in the following manner. Determine  $u(t) \in W$  to minimize a scalar functional of the state, desired state, and the control trajectories.

$$C = \int_{\Omega} [Q_d(x, T) - Q(x, T)]^T [Q_d(x, T) - Q(x, T)] d\Omega \quad (3.18)$$

Let us assume that the open loop control is computed using any one of the suitable techniques described in sections 3.2 and

3.3. Let us represent the open loop control (Optimal) as:

$$u^*(t) = \theta(t, Q(x, o)) \quad (3.19)$$

to stress the implicit dependence on the initial state.

The present problem is to determine the closed loop control laws, denoted by  $u_c(t)$ , that yield system performance, that in some manner approximates the optimal behavior. Thus we require a criterion to compare the open loop and closed loop system performance. One of the following criteria can be used.

$$a) \quad \min_{u_c(t)} \int_0^{t_f} \|\theta(t, Q(x, o)) - u_c(t)\|^2 dt \quad (3.20)$$

or

$$b) \quad \min_{u_c(t)} \int_{\Omega} \|Q^*(x, T) - Q_c(x, T)\|^2 d\Omega \quad (3.21)$$

where  $Q^*(x, T)$  is the optimal state trajectory obtained by the application of the optimal open loop control law and  $Q_c(x, T)$  is the state resulting from the application of the closed loop control law  $u_c(t)$ .

To carry out either of the above minimizations it is necessary to assume functional forms for the feedback laws  $u_c(t)$  which include the adjustable parameters that can be determined by the minimizations. Let us assume for convenience  $m = p = 1$ , and

where  $k_w$

either of t

several way

feedback la

Let

the sub-opt

min

$k_w$

Expanding t

min  $\int$

$k_w$  o

Differenti

we get

k

It is evi

function

dependen

ditions

given pr

$$u_c(t) = k_w h \left[ \int_{\Omega} q(x,t) d\Omega \right] \quad (3.22)$$

where  $k_w$  is the parameter to be determined by minimizing either of the two equations (3.20) and (3.21). There are several ways of choosing the functional relationship for the feedback laws, and the form given in (3.22) is not unique.

Let us carry out the minimizations for  $k_w$  by using the sub-optimality criterion (a) given in (3.20), i.e.

$$\min_{k_w} \int_0^{t_f} \{u^*(t) - k_w h \left[ \int_{\Omega} q^*(x,t) d\Omega \right]\}^2 dt$$

Expanding the terms under the integral sign, we desire

$$\min_{k_w} \int_0^{t_f} \{u^{*2}(t) + k_w^2 h^2 \left[ \int_{\Omega} q^*(x,t) d\Omega \right] - 2k_w h \left[ \int_{\Omega} q^*(x,t) d\Omega \right]\} dt \quad (3.23)$$

Differentiating with respect to  $k_w$  and equating to zero, we get

$$k_w = \frac{\int_0^{t_f} \{u^*(t) h \left[ \int_{\Omega} q^*(x,t) d\Omega \right]\} dt}{\int_0^{t_f} h^2 \left[ \int_{\Omega} q^*(x,t) d\Omega \right] dt} \quad (3.24)$$

It is evident that the value of  $k_w$  obtained is an implicit function of the initial condition,  $q(x,0)$ , because of the dependency of the open loop control laws on the initial conditions of the system.

#### 3.4.2 Koivo and Kruh's method

Let  $u^*(t)$  be the optimal open loop control for the given problem, and the corresponding optimal state  $q^*(x,t_f)$

at  $t = t_f$ . Then the feedback control is assumed in the form:

$$u_s(q_s, h, t) = F[q_s(x^d, h, t), h, t] \quad (3.25)$$

where

$$x^d = \text{col } [x^1, x^2, \dots, x^M]$$

$$h = \text{col } [h^1, h^2, \dots, h^M]$$

and  $q_s(x, t)$  denotes the solution of the given system when the feedback control  $u_s(t)$  is applied,  $x^d$  denotes the  $M$  sensor locations,  $h$  represent the feedback constants of the controller to be determined. The purpose of the design is to obtain the feedback coefficients  $h$ , so as to minimize,

$$DJ = \int_{\Omega} |q_s(x, x^d, h, t_f) - q^*(x, t_f)| d\Omega \quad (3.26)$$

The method of obtaining the parameters is as follows:

Let us assume that the system is described by the integral equation, which in the linear time invariant, constant coefficient case can be transformed into this form by the use of Laplace transform techniques. Thus,

$$q(x, t) = \int_0^t g(x, t-\tau) u(\tau) d\tau \quad (3.27)$$

where  $g(x, t-\tau)$  is the known characteristic of the system, and let  $T = t_f$ .

The first differential of the (3.26) with respect to  $h$  can be written as,

$$\Delta DJ[x, h; \Delta h] = \sum_{m=1}^M \rho^m \Delta h^m \quad (3.28)$$

where

$$\rho^m = \int_{\Omega} \text{sgn} [q_s(x, x^d, h, T) - q^*(x, T)] \frac{\partial q_s}{\partial h^m} dx \quad (3.29)$$

$$\frac{\partial q_s}{\partial h^m} = \int_0^T g(x, T-\tau) \frac{\partial F[q_s(x^d, h, t), h, t]}{\partial h^m} d\tau \quad (3.30)$$

Requiring a constant step size

$$\|\Delta h\|^2 = \sum_{m=1}^M (\Delta h^m)^2 = \text{constant} \quad (3.31)$$

we have the following algorithm:

- 1) Compute the optimal trajectory  $q^*(x, t)$  and the corresponding performance index.
- 2) For each  $x^m$ , approximate initially the value of  $h^m$ ,  $m = 1, 2, \dots, M$ .
- 3) Compute the approximately optimum trajectory  $q_s(x, x^d, h, t)$  and DJ from (3.26).
- 4) Compute  $\rho^m$ ,  $m = 1, 2, \dots, M$
- 5) Change  $h^m$  to  $h^m + \Delta h^m$ ,  $m = 1, 2, \dots, M$  so as to decrease DJ.

The  $\Delta h^m$  used in reference (K-5) is

$$\Delta h^m = - \frac{\rho^m \|\Delta h\|}{\left( \sum_{m=1}^M (\rho^m)^2 \right)^{\frac{1}{2}}}, \quad m = 1, 2, \dots, M$$

- 6) Repeat from (3) until the minimum of DJ is obtained.

In the next chapter, a decomposition principle which decomposes a large differential system into smaller order independent subsystems is stated, and the convergence theorems are proved.

#### 4.1 Intro

A

equation of

approach".

finite diff

continuous

that the ar

quantization

systems are

and the con

approximat

scheme has

In this ch

this thesi

composing

discretiz

systems.

complete

tion in d

## CHAPTER IV

### A DECOMPOSITION PRINCIPLE

#### 4.1 Introduction

A classical way of solving the partial differential equation on analog computers is the so called "parallel approach". The method replaces the space derivative by a finite difference scheme while keeping the time derivative continuous. Intrinsic to such an approach is the problem that the amount of equipment required grows larger with finer quantization of the space variable. Distributed parameter systems are characterized by partial differential equations, and the computational techniques require some kind of approximation. The approximation by a finite difference scheme has the disadvantage of demanding large computers. In this chapter, a method which is an original contribution of this thesis is proposed to circumvent this difficulty by decomposing the large set of equations resulting from the space discretization into a set of lower order independent subsystems. This requires an iterative technique to obtain the complete solution. Of course the price paid for this reduction in dimensionality is increased computer time.

## 4.2 Definitions and Theorems

Before stating the central theorems of this chapter with proofs, some of the necessary concepts are developed in the following sections.

### 4.2.1 Spectral radius of a matrix

Let  $A = (a_{ij})$  be an  $n \times n$  complex matrix with eigenvalues  $\lambda_i$ ,  $1 \leq i \leq n$ . Then

$$\rho(A) = \max_{1 \leq i \leq n} |\lambda_i| \quad (4.1)$$

is the spectral radius of the matrix  $A$ .

### 4.2.2 Spectral norm of a matrix $A$

Let  $A = (a_{ij})$  be an  $n \times n$  complex matrix. Then

$$\|A\| = \sup_{x \neq 0} \frac{\|Ax\|}{\|x\|} \quad (4.2)$$

is the spectral norm of the matrix  $A$ .

Theorem 1. If  $A$  and  $B$  are two  $n \times n$  matrices, and  $\alpha$  is any scalar, then  $\|A\| > 0$ , unless  $A \equiv 0$ ,

$$\begin{aligned} \text{(i)} \quad & \|\alpha A\| = |\alpha| \cdot \|A\| \\ \text{(ii)} \quad & \|A + B\| \leq \|A\| + \|B\| \\ \text{(iii)} \quad & \|A \cdot B\| \leq \|A\| \cdot \|B\| \\ \text{(iv)} \quad & \|Ax\| \leq \|A\| \|x\| \quad \text{for all vectors of } x \end{aligned} \quad (4.3)$$

where  $\|x\|$  is an Euclidean norm.

Proof: The proof is given in Varge (V-1).

Corollary: For an arbitrary  $n \times n$  complex matrix  $A$ ,

$$\|A\| \geq \rho(A) \quad (4.4)$$

Proof: If  $\lambda$  is any eigenvalue of  $A$ , and  $x$  is any eigenvector associated with the eigenvalue  $\lambda$ , then  $Ax = \lambda x$ . Thus

$$|\lambda| \cdot \|x\| = \|\lambda x\| = \|Ax\| \leq \|A\| \cdot \|x\| \quad (4.5)$$

from which we conclude

$$\|A\| \geq |\lambda| \quad \text{for all eigenvalues of } A,$$

which proves (4.4).

Theorem 2. Varga (V-1). If  $A = (a_{ij})$  is an  $n \times n$  complex matrix, then

$$\|A\| = [\rho(A^*A)]^{\frac{1}{2}}$$

where  $A^*$  is the conjugate transpose of  $A$ .

Corollary: If  $A$  is an  $n \times n$  Hermitian matrix, then

$$\|A\| = \rho(A) \quad (4.6)$$

Moreover, if  $g_m(x)$  is any real polynomial of degree  $m$  in  $x$ , then,

$$\|g_m(A)\| = \rho(g_m(A)) \quad (4.7)$$

Proof: If  $A$  is Hermitian, then  $A = A^*$ , and thus

$$\|A\|^2 = \rho(A^*A) = \rho(A^2) = \rho^2(A)$$

$$\therefore \|A\| = \rho(A)$$

Now since

$g_m(A)$  is a

4.2

Let

convergent

converges to

Theorem 3.

convergent

Proof: For

$n \times n$  matrix

normal form

SAS

where each

Now since  $g_m(x)$  is a real polynomial in the variable  $x$ ,  $g_m(A)$  is also Hermitian, and (4.7) is proven, because

$$\|g_m(A)\|^2 = \rho(g_m(A) \cdot g_m^*(A)) = \rho^2(g_m(A)).$$

Q.E.D.

#### 4.2.3 Convergence of a matrix A

Let  $A$  be an  $n \times n$  complex matrix. Then  $A$  is convergent (to zero) if the sequence of matrices  $A, A^2, A^3, \dots$  converges to the null matrix, and is divergent otherwise.

Theorem 3. If  $A$  is an  $n \times n$  complex matrix, then  $A$  is convergent if and only if  $\rho(A) < 1$ .

Proof: For the given matrix  $A$ , there exists a non-singular  $n \times n$  matrix  $S$ , which reduces the matrix  $A$  to its Jordan normal form, i.e.

$$SAS^{-1} = \tilde{A} \equiv \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_r \end{bmatrix} \quad (4.8)$$

where each of the  $n_\ell \times n_\ell$  submatrices  $J_\ell$  has the form

$$J_\ell = \begin{bmatrix} \lambda_\ell & 1 & & & \\ & \lambda_\ell & 1 & & \\ & & \lambda_\ell & 1 & \\ & & & \ddots & 1 \\ & & & & \lambda_\ell \end{bmatrix}, \quad 1 \leq \ell \leq r \quad (4.9)$$

Since each submatrix  $J_\ell$  is upper triangular, so is  $\tilde{A}$ .

Thus the set  $\{\lambda_\ell\}_{\ell=1}^r$  includes all the distinct eigenvalues of the matrices  $A$  and  $\tilde{A}$ , which are similar matrices from (4.8). By direct computation with (4.8), we get

$$(\tilde{A})^m = \begin{bmatrix} J_1^m & & & \\ & J_2^m & & \\ & & \ddots & \\ & & & J_r^m \end{bmatrix}, \quad m \geq 1 \quad (4.10)$$

The entries of the powers of the matrix  $J_\ell$  are determined as follows:

$$J_\ell^2 = \begin{bmatrix} \lambda_\ell & 1 & 0 & \dots & 0 \\ 0 & \lambda_\ell & 1 & 0 & \dots & 0 \\ 0 & 0 & \lambda_\ell & 1 & 0 & \dots \\ & & & \ddots & \ddots & \\ \bigcirc & & & & \lambda_\ell & 1 \\ & & & & & \lambda_\ell \end{bmatrix} \begin{bmatrix} \lambda_\ell & 1 & 0 & \dots & 0 \\ 0 & \lambda_\ell & 1 & 0 & \dots & 0 \\ & & & & & \bigcirc \\ & & & & & 1 \\ & & & & & \lambda_\ell \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_\ell^2 & 2\lambda_\ell & 1 & 0 & \dots & 0 \\ & & & & & 1 \\ & & & & & 2\lambda_\ell \\ \bigcirc & & & & & \lambda_\ell^2 \end{bmatrix}$$

and in general if we define

$$J_\ell^m = (d_{ij}^{(m)}(\ell)) \quad 1 \leq i, j \leq n_\ell$$

then

$$d_{ij}^{(m)}(\ell) = \begin{cases} 0 & j < i \\ \binom{m}{j-i} \lambda_\ell^{m-j+i} & \text{for } i \leq j \leq \min(n_\ell, m+i) \\ 0 & m+i < j \leq n_\ell \end{cases} \quad (4.11)$$

where

$$\binom{m}{k} = \frac{m!}{k! (m-k)!}$$

Now if  $A$  is convergent, then by definition in 4.2.3  $A^m \rightarrow 0$  as  $m \rightarrow \infty$ . But  $(\tilde{A})^m = S A^m S^{-1}$ . So it follows that  $\tilde{A}^m \rightarrow 0$  as  $m \rightarrow \infty$ . Consequently each  $J_\ell^{(m)} \rightarrow 0$  as  $m \rightarrow \infty$  so that the diagonal entries  $\lambda_\ell$  of  $J_\ell$  must satisfy  $|\lambda_\ell| < 1$  for all  $1 \leq \ell \leq r$ . Clearly

$$\rho(A) = \rho(\tilde{A}) = \max_{1 \leq \ell \leq r} |\lambda_\ell| < 1$$

which proves the first part. On the other hand if

$\rho(A) = \rho(\tilde{A}) < 1$ , then  $|\lambda_\ell| < 1$  for all  $1 \leq \ell \leq r$ . Then by making direct use of (4.11) and the fact that  $|\lambda_\ell| < 1$ , it follows that

$$\lim_{m \rightarrow \infty} d_{ij}^{(m)}(\ell) = 0 \quad \text{for all } i \geq 1; j \geq n_\ell$$

Thus each  $J_\ell$  is convergent, and  $\tilde{A}$  is convergent. Finally,

$$A^m = S^{-1} \tilde{A}^m S$$

This proves that the matrix  $A$  is convergent.

#### 4.2.4 Bounds for the spectral radius of a matrix

It is generally difficult to determine precisely the spectral radius of a given matrix. Nevertheless, upper bounds

can be easily found from the following theorem:

Theorem 4. Let  $A = (a_{ij})$  be an arbitrary  $n \times n$  complex matrix, and let

$$\Lambda_i \equiv \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \quad 1 \leq i \leq n$$

Then all the eigenvalues  $\lambda$  of  $A$  lie in the union of the disks,

$$|z - a_{ii}| \leq \Lambda_i \quad 1 \leq i \leq n \quad (4.12)$$

Proof: Let  $\lambda$  be any eigenvalue of the matrix  $A$ , and let  $x$  be an eigenvector of  $A$  corresponding to  $\lambda$ . We normalize the vector  $x$  so that its largest component in modulus is unity.

By definition,

$$(\lambda - a_{ii})x_i = \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij}x_j \quad 1 \leq i \leq n$$

In particular, if  $|x_r| = 1$ , then

$$|\lambda - a_{rr}| \leq \sum_{\substack{j=1 \\ j \neq r}}^n |a_{rj}| \cdot |x_j| \leq \sum_{\substack{j=1 \\ j \neq r}}^n |a_{rj}| = \Lambda_r$$

Thus, the eigenvalue  $\lambda$  lies in the disk  $|z - a_{rr}| \leq \Lambda_r$ .

But since  $\lambda$  was an arbitrary eigenvalue of  $A$ , it follows that all the eigenvalues of the matrix  $A$  lie in the union of disks  $|z - a_{ii}| \leq \Lambda_i$ ,  $1 \leq i \leq n$  completing the proof.

Corollary. If  $A = (a_{ij})$  is an arbitrary  $n \times n$  complex matrix and

$$\mu \equiv \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad (4.13)$$

then  $\rho(A) \leq \mu$ .

Thus the maximum of the row sums of the moduli of the entries of the matrix  $A$  gives a simple upper bound. Since  $A$  and  $A^t$  have the same eigenvalues,

$$\mu' \equiv \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \quad (4.14)$$

then  $\rho(A) \leq \mu'$ .

#### 4.2.5 Conditions for the existence of an inverse of (I-M) when M is an arbitrary matrix

Theorem 5. If  $M$  is an arbitrary complex matrix with  $\rho(M) < 1$  then  $I-M$  is nonsingular, and

$$(I - M)^{-1} = I + M + M^2 + \dots \quad (4.15)$$

where the series on the right converges. Conversely, if the series on the right converges, then  $\rho(M) < 1$ .

Proof: First assume that  $\rho(M) < 1$ . If  $\mu$  is an eigenvalue of  $M$ , then  $1 - \mu$  is an associated eigenvalue of  $I-M$ , and, as  $\rho(M) < 1$ ,  $I-M$  is nonsingular. From the identity,

$$1 + \mu + \mu^2 + \dots + \mu^r = \frac{1 - \mu^{r+1}}{1 - \mu}$$

and substituting  $\mu = M$ , we get

$$I - (I-M)(I + M + M^2 + \dots + M^r) = M^{r+1}$$

we have, upon premultiplying by  $(I-M)^{-1}$ , that

$$(I-M)^{-1} - (I + M + M^2 + \dots + M^r) = (I-M)^{-1}M^{r+1}$$

Thus,

$$\|(I-M)^{-1} - (I + M + M^2 + \dots + M^r)\| \leq \|(I-M)^{-1}\| \cdot \|M^{r+1}\|$$

for all  $r \geq 0$ . As  $M$  is convergent, it follows that

$\|M^{r+1}\| \rightarrow 0$  as  $r \rightarrow \infty$ . Thus the series in (4.15) converges and is equal to  $(I-M)^{-1}$ . Conversely, if the series converges, let  $\mu$  be the eigenvalue of  $M$ , corresponding to an eigenvector  $x$ , then

$$(I + M + M^2 + \dots)x = (1 + \mu + \mu^2 + \dots)x$$

Thus the convergence of the matrix series implies the convergence of the series  $1 + \mu + \mu^2 + \dots$  for any eigenvalue  $\mu$  of  $M$ . However, as is well known, for this series of complex numbers to converge, it is necessary that  $|\mu| < 1$  for all eigenvalues of  $M$ , and thus  $\rho(M) < 1$ , completing the proof.

#### 4.3 A Decomposition Principle

Many of the physical systems described by partial differential equations involve at most three dimensions in the space domain. To solve such problems in three space, a spatial discretization is used which yields a set of ordinary differential equations. The number of ordinary differential equations in this set increases rapidly as finer and finer spatial discretizations are used.

One method of solving this set of differential equations is by the use of an analog computer. The difficulty is that the number of integrators on any analog computer is limited and the cost of the equipment increases markedly when additional sophisticated integrators are added to the available facility. The second method of solving this set of differential equations is by the use of digital computers. The disadvantage in this case is the large amount of memory required and the increase of computation time with the number of equations. Hence a hybrid computer solution obtains the advantages of both the analog computer where the differential equations are solved in parallel and the digital computer is used for the logical and control functions.

Thus the necessity of an algorithm for obtaining at least an approximate solution of the given system of equations with smaller number of integrators and small amount of core memory is established. In the next sections, an algorithm is stated and proved, which increases the capabilities of the analog and digital computers and thus obtains an approximate solution to a larger set of equations with fewer integrators and less core memory.

#### 4.3.1 Algorithm

Many of the systems characterized by the partial differential equations yield tridiagonal matrices when discretized. If the discretized matrices are not in the tridiagonal form, they can be reduced to this form without computing the eigenvalues

by using several existing methods, such as in Ellsworth (E-4).

Now, given the equation

$$\dot{X} = A X, X(t_0) = X_0 \quad (4.16)$$

where  $A$  is a tridiagonal  $n \times n$  real matrix,  $X$  is  $1 \times n$  vector, and  $X_0$  is the initial condition vector. We define a partition of  $X$  as follows:

$$X = (X_1, X_2, X_3, \dots, X_N)'$$

where  $X_j$  has dimension  $n_j$ . The decomposed problem thus also partitions the matrix  $A$  into  $N^2$  blocks  $A_{ij}$  such that,

$$\begin{aligned} \dot{X}_1 &= A_{11} X_1 + A_{12} X_2 \\ \dot{X}_2 &= A_{21} X_1 + A_{22} X_2 + A_{23} X_3 \\ &\dots \\ &\dots \\ \dot{X}_{N-1} &= A_{N-1,N-2} X_{N-2} + A_{N-1,N-1} X_{N-1} + A_{N-1,N} X_N \\ \dot{X}_N &= A_{N,N-1} X_{N-1} + A_{N,N} X_N \end{aligned} \quad (4.17)$$

Then the iterative method of obtaining the solution of the above equations is given in the flow chart in Figure 4.1.

This method is like the Gauss-Seidel method for the solution of linear equations, since the values obtained for the other variables are utilized immediately. The initial vector  $P$ , which is the same for all the variables, is used to reduce the storage in the digital computer. The Jacobi method can also be used, but it increases the storage

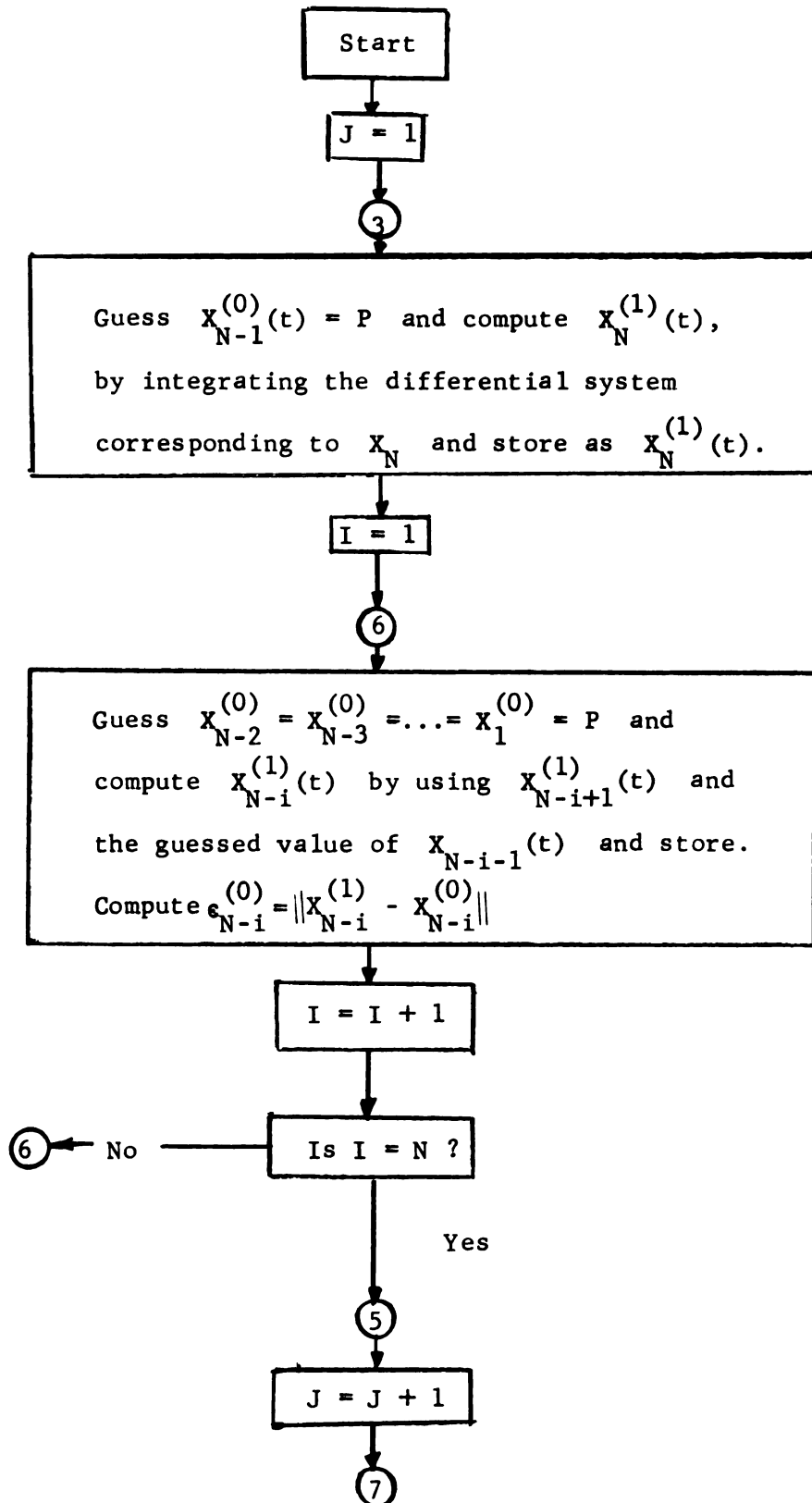


Figure 4.1 Flow Chart of the Algorithm

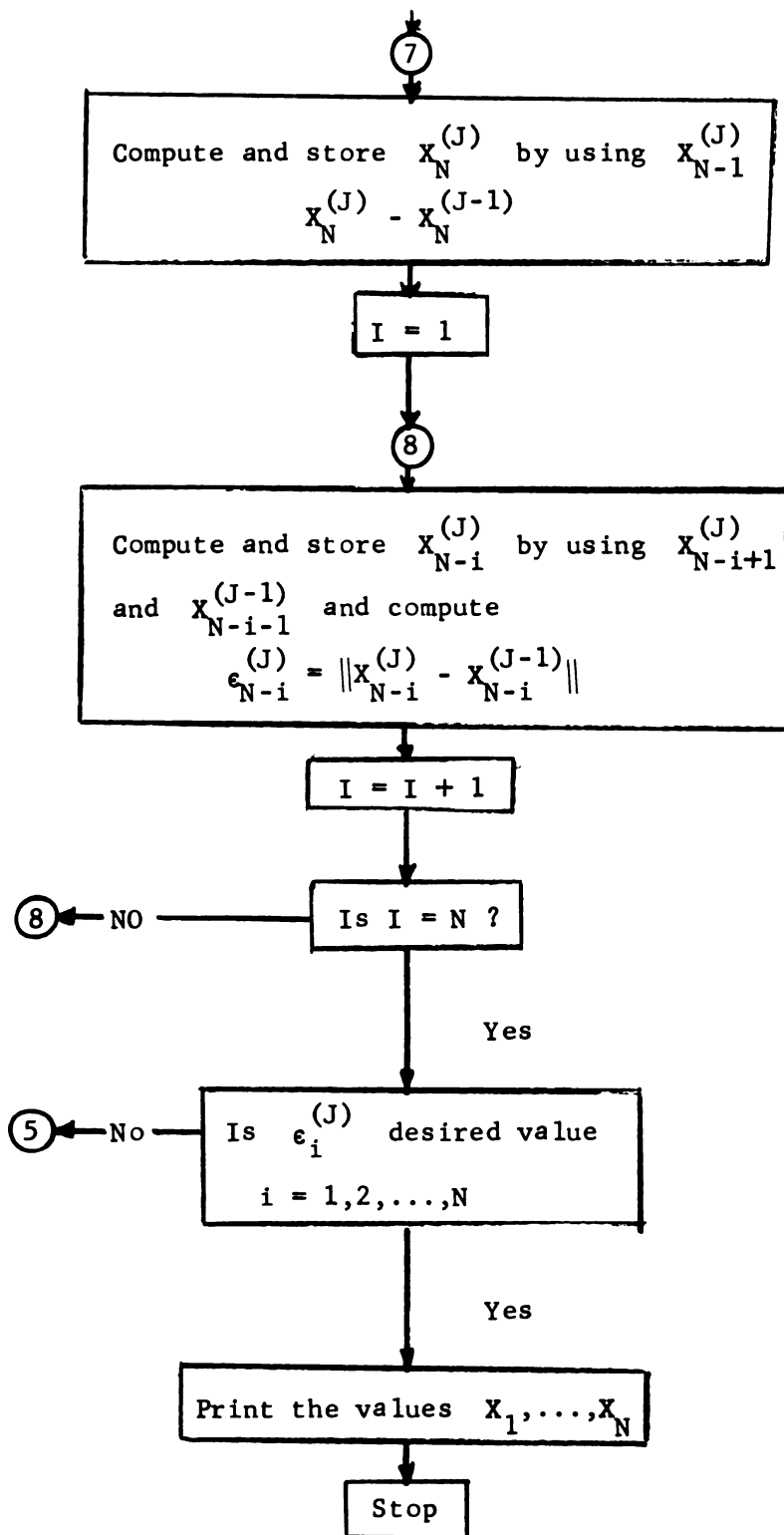


Figure 4.1 Flow Chart of the Algorithm (contd.)

required because the values of the variables corresponding to two consecutive iterations are to be stored.

Thus loosely speaking, the decomposition principle starts with the last partition and sequentially proceeds to the first and cycles from last to first to last etc., until the error is less than or equal to the stipulated value.

Assume that  $A$  is a  $24 \times 24$  tridiagonal matrix. Let  $Z$  be divided into two partitions and let each partition contain a  $12 \times 12$  matrix, i.e.,

$$A = \left[ \begin{array}{c|c} \begin{array}{c} 12 \times 12 \\ \text{matrix} \end{array} & \bigcirc \\ \hline \bigcirc & \begin{array}{c} 12 \times 12 \\ \text{matrix} \end{array} \end{array} \right] \begin{array}{c} a \\ b \end{array}$$

where  $a$  and  $b$  are non-zero elements.

Then a hybrid computer that has twelve integrators, could be used to solve this  $24 \times 24$  matrix with the storage of only two functions. If the matrix contains different numerical values, these can be adjusted by means of the digital computer. Now, if we suppose that each function is sampled at 50 points in the interval of interest, then two functions require only 200 words of core memory. In the next section, the convergence theorem for this decomposition principle is given.

### 4.3.2 Convergence theorems

First a theorem for  $N = 2$  is proved and then it is generalized.

Theorem 6. Given the equation

$$\dot{X} = A X, \quad X(0) = X_0 \quad \text{on } t \in (0, t_f)$$

Let  $A$  be partitioned into two partitions, i.e.,  $N = 2$  in (4.17) and be rewritten as

$$\begin{aligned} \dot{X}_1 &= A_1 X_1 + B_1 X_2 \\ \dot{X}_2 &= A_2 X_2 + B_2 X_1 \end{aligned} \quad (4.18)$$

while requiring that  $A_1$  and  $A_2$  are stable. Then a proper choice of  $\Delta t$  such that  $\rho(A_1 \Delta t) < 1$ ,  $\rho(A_2 \Delta t) < 1$ ,

1) guarantees the convergence of the algorithm in section 4.3.1

2)  $\rho(A_1 \Delta t), \rho(A_2 \Delta t)$  controls the rate of convergence.

Proof: Let  $(0, t_f)$  be partitioned into  $M$  subintervals and  $t_i \in (0, t_f)$ . Then (4.18) can be approximated at  $t_i$  by using the midpoint approximation

$$\frac{X_1(t_i) - X_1(t_{i-1})}{\Delta t} = \frac{A}{2} (X_1(t_i) + X_1(t_{i-1})) + \frac{B_1}{2} (X_2(t_i) + X_2(t_{i-1}))$$

or

$$\begin{aligned} X_1(t_i) = X_1(t_{i-1}) &+ \left[ \frac{A_1}{2} X_1(t_i) + \frac{B_1}{2} X_2(t_i) + \frac{A_1}{2} X_1(t_{i-1}) \right. \\ &\left. + \frac{B_1}{2} X_2(t_{i-1}) \right] \Delta t + \theta(\Delta t^2) \end{aligned} \quad (4.19)$$

Similarly  $x_2(t_i)$  can be written as

$$x_2(t_i) = x_2(t_{i-1}) + \left[ \frac{A_2}{2} x_2(t_i) + \frac{B_2}{2} x_1(t_i) + \frac{A_2}{2} x_2(t_{i-1}) + \frac{B_2}{2} x_1(t_{i-1}) \right] \Delta t + o(\Delta t^2) \quad (4.20)$$

Equations (4.19) and (4.20) could be written in the matrix forms

$$\begin{bmatrix} x_1(t_i) \\ x_2(t_i) \end{bmatrix} = \Delta t \begin{bmatrix} \frac{A_1}{2} & \frac{B_1}{2} \\ \frac{B_2}{2} & \frac{A_2}{2} \end{bmatrix} \begin{bmatrix} x_1(t_i) \\ x_2(t_i) \end{bmatrix} + \begin{bmatrix} I + A_1 \frac{\Delta t}{2} & B_1 \frac{\Delta t}{2} \\ B_2 \frac{\Delta t}{2} & I + A_2 \frac{\Delta t}{2} \end{bmatrix} \begin{bmatrix} x_1(t_{i-1}) \\ x_2(t_{i-1}) \end{bmatrix} + o(\Delta t^2)$$

or

$$\begin{bmatrix} x_1(t_i) \\ x_2(t_i) \end{bmatrix} = \begin{bmatrix} A_1 & B_1 \\ B_2 & A_2 \end{bmatrix} \frac{\Delta t}{2} \begin{bmatrix} x_1(t_i) \\ x_2(t_i) \end{bmatrix} + \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} \begin{bmatrix} x_1(t_{i-1}) \\ x_2(t_{i-1}) \end{bmatrix} + o(\Delta t^2) \quad (4.21)$$

Substituting

$$x_1(t_i) = (I - A_1 \frac{\Delta t}{2})^{-1} \left[ B_1 \frac{\Delta t}{2} x_2(t_i) + m_{11} x_1(t_{i-1}) + m_{12} x_2(t_{i-1}) \right] + o(\Delta t^2)$$

$$x_2(t_i) = (I - A_2 \frac{\Delta t}{2})^{-1} \left[ B_2 \frac{\Delta t}{2} x_1(t_i) + m_{21} x_1(t_{i-1}) + m_{22} x_2(t_{i-1}) \right] + o(\Delta t^2)$$

we get

$$\begin{bmatrix} x_1(t_i) \\ x_2(t_i) \end{bmatrix} = \frac{\Delta t}{2} \begin{bmatrix} A+B_1(I-A_2 \frac{\Delta t}{2})^{-1} B_2 \frac{\Delta t}{2} & 0 \\ 0 & A_2+B_2(I-A_1 \frac{\Delta t}{2})^{-1} B_1 \frac{\Delta t}{2} \end{bmatrix} \begin{bmatrix} x_1(t_i) \\ x_2(t_i) \end{bmatrix} + N \begin{bmatrix} x_1(t_{i-1}) \\ x_2(t_{i-1}) \end{bmatrix} + o(\Delta t^2)$$

where

$$N \triangleq \begin{bmatrix} (I - A_2 \frac{\Delta t}{2})^{-1} \frac{\Delta t}{2} m_{21} + m_{11} & (I - A_2 \frac{\Delta t}{2})^{-1} \frac{\Delta t}{2} m_{22} + m_{12} \\ (I - A_1 \frac{\Delta t}{2})^{-1} \frac{\Delta t}{2} m_{11} + m_{21} & (I - A_1 \frac{\Delta t}{2})^{-1} \frac{\Delta t}{2} m_{12} + m_{22} \end{bmatrix} \quad (4.22)$$

Now let  $\Delta t^2$  be sufficiently small that all the terms multiplied by  $\Delta t^2$  can be neglected. In the above the inverses  $(I - A_1 \Delta t)^{-1}$  and  $(I - A_2 \Delta t)^{-1}$  exists if and only if the matrices  $A_1$  and  $A_2$  are such that  $\rho(A_1 \Delta t)$  and  $\rho(A_2 \Delta t) < 1$  by theorem 5, in section 4.2.5. The series then are

$$(I - A_1 \frac{\Delta t}{2})^{-1} = I + A_1 \frac{\Delta t}{2} + \theta(\Delta t^2)$$

$$(I - A_2 \frac{\Delta t}{2})^{-1} \approx I + A_2 \frac{\Delta t}{2} + \theta(\Delta t^2)$$

and the approximation

$$(I - A_1 \frac{\Delta t}{2})^{-1} \approx I + A_1 \frac{\Delta t}{2}$$

$$(I - A_2 \frac{\Delta t}{2})^{-1} \approx I + A_2 \frac{\Delta t}{2}$$

The neglected terms are of the order of  $\Delta t^2/4$ . The approximations of the various terms in the matrix in (4.22) are as follows:

$$\begin{aligned}
A_1 \frac{\Delta t}{2} + B_1 (I - A_2 \frac{\Delta t}{2})^{-1} B_2 \frac{\Delta t^2}{2} &\approx A_1 \frac{\Delta t}{2} \\
A_2 \frac{\Delta t}{2} + B_2 (I - A_1 \frac{\Delta t}{2})^{-1} B_1 \frac{\Delta t^2}{2} &\approx A_2 \frac{\Delta t}{2} \\
(I - A_2 \frac{\Delta t}{2})^{-1} \frac{\Delta t}{2} m_{21} + m_{11} &\approx m_{11} = I + A_1 \frac{\Delta t}{2} \\
(I - A_1 \frac{\Delta t}{2})^{-1} \frac{\Delta t}{2} m_{12} + m_{22} &\approx m_{22} = I + A_2 \frac{\Delta t}{2} \\
(I - A_2 \frac{\Delta t}{2})^{-1} \frac{\Delta t}{2} m_{22} + m_{12} &\approx m_{12} + I \frac{\Delta t}{2} = B_1 \frac{\Delta t}{2} + I \frac{\Delta t}{2} \\
(I - A_1 \frac{\Delta t}{2})^{-1} \frac{\Delta t}{2} m_{11} + m_{21} &\approx m_{21} + I \frac{\Delta t}{2} = B_2 \frac{\Delta t}{2} + I \frac{\Delta t}{2}
\end{aligned} \tag{4.23}$$

Hence (4.22) can be approximated as

$$\begin{bmatrix} x_1(t_i) \\ x_2(t_i) \end{bmatrix} = \begin{bmatrix} A_1 \frac{\Delta t}{2} & 0 \\ 0 & A_2 \frac{\Delta t}{2} \end{bmatrix} \begin{bmatrix} x_1(t_{i-1}) \\ x_2(t_{i-1}) \end{bmatrix} + \begin{bmatrix} I + A_1 \frac{\Delta t}{2} & \frac{B_1}{2} \Delta t + I \frac{\Delta t}{2} \\ B_2 \frac{\Delta t}{2} + I \frac{\Delta t}{2} & I + A_2 \frac{\Delta t}{2} \end{bmatrix} \begin{bmatrix} x_1(t_{i-1}) \\ x_2(t_{i-1}) \end{bmatrix}$$

Thus we can write for known values of  $x_1(t_{i-1})$ ,  $x_2(t_{i-1})$  and  $\Delta t$ ,

$$X(t_i) = P X(t_{i-1}) + G \tag{4.24}$$

where  $G$  is a constant matrix.

Let us assume  $X^*$  is the unique solution of the above equation.

Then writing

$$X^m - X^* = \epsilon^m$$

where  $\epsilon^m$  is the error at the  $m$ th iteration,

$$\begin{aligned}
\epsilon^m &= P X^{m-1} - P X^* \\
&= P(X^{m-1} - X^*) \\
&= P \epsilon^{m-1} \\
&= \dots \\
&= \dots \\
&= P^m \epsilon^0
\end{aligned} \tag{4.25}$$

where

$$P \triangleq \begin{bmatrix} A_1 \frac{\Delta t}{2} & 0 \\ 0 & A_2 \frac{\Delta t}{2} \end{bmatrix}$$

Thus the error vectors tend to zero, if and only if the spectral radius  $\rho(P)$  is less than one. This can be done by proper choice of  $\Delta t$ . This also requires that the decomposed systems be stable since any perturbations of the solution should not make the systems unstable.

The approximations made in (4.23) are accurate because  $B_1$  and  $B_2$  are of the form

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & & & & \\ \cdot & & & & \\ \cdot & & \bigcirc & & \\ * & & & & \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 & 0 & 0 & * \\ 0 & & & & \\ 0 & & & & \\ \cdot & & \bigcirc & & \\ 0 & & & & \end{bmatrix} \quad \text{respectively,}$$

where  $*$  represents the non-zero element.

Q.E.D.

Now let there be  $N = 3$  diagonal blocks  $A_{ii}$ , such that

$$\dot{x}_1 = A_{11} x_1 + A_{12} x_2$$

$$\dot{x}_2 = A_{21} x_1 + A_{22} x_2 + A_{23} x_3$$

$$\dot{x}_3 = A_{32} x_2 + A_{33} x_3$$

then by similar computations mentioned in the above,

$$\begin{bmatrix} x_1(t_i) \\ x_2(t_i) \\ x_3(t_i) \end{bmatrix} = \frac{\Delta t}{2} \begin{bmatrix} A_{11} & A_{12} & 0 \\ A_{21} & A_{22} & A_{23} \\ 0 & A_{32} & A_{33} \end{bmatrix} \begin{bmatrix} x_1(t_i) \\ x_2(t_i) \\ x_3(t_i) \end{bmatrix} + \begin{bmatrix} I + A_{11} \frac{\Delta t}{2} & A_{12} \frac{\Delta t}{2} & 0 \\ A_{21} \frac{\Delta t}{2} & I + A_{22} \frac{\Delta t}{2} & A_{23} \frac{\Delta t}{2} \\ 0 & A_{32} \frac{\Delta t}{2} & I + A_{33} \frac{\Delta t}{2} \end{bmatrix} \begin{bmatrix} x_1(t_{i-1}) \\ x_2(t_{i-1}) \\ x_3(t_{i-1}) \end{bmatrix}$$

and by suitable approximations that neglect  $\frac{\Delta t^2}{4}$  terms,

$$x_j(t_i) = A_{jj} \frac{\Delta t}{2} x_j(t_i) + G_j, \quad j = 1, 2, 3 \quad (4.26)$$

where  $G$  is a constant matrix for a given  $x_j(t_{i-1})$ ,  $j = 1, 2, 3$  and  $\Delta t$ .

Similarly for  $N$  partitions,

$$\begin{bmatrix} x_1(t_i) \\ x_2(t_i) \\ x_3(t_i) \\ \vdots \\ x_N(t_i) \end{bmatrix} = \frac{\Delta t}{2} \begin{bmatrix} A_{11} & A_{12} & 0 & & 0 \\ A_{21} & A_{22} & A_{23} & 0 & 0 \\ 0 & A_{32} & A_{33} & A_{34} & 0 \dots 0 \\ & & & \ddots & \\ \bigcirc & & & & \end{bmatrix} \begin{bmatrix} x_1(t_i) \\ x_2(t_i) \\ x_3(t_i) \\ \vdots \\ x_N(t_i) \end{bmatrix} + M \begin{bmatrix} x_1(t_{i-1}) \\ x_2(t_{i-1}) \\ x_3(t_{i-1}) \\ \vdots \\ x_N(t_{i-1}) \end{bmatrix}$$

where  $M$  is a known matrix of the form

$$M \triangleq \begin{bmatrix} I + A_{11} \frac{\Delta t}{2} & A_{12} \frac{\Delta t}{2} & 0 & \dots & 0 \\ A_{21} \frac{\Delta t}{2} & I + A_{22} \frac{\Delta t}{2} & A_{23} \frac{\Delta t}{2} & 0 & 0 \end{bmatrix} \quad (4.27)$$

Since  $\Delta t$  is fixed and  $X_1(t_{i-1}), X_2(t_{i-1}) \dots$  are known, the matrix  $M$  is completely known. Hence we get,

$$X_j(t_i) = A_{jj} \frac{\Delta t}{2} X_j(t_{i-1}) + G_j \quad (4.28)$$

where  $j = 1, 2, \dots, N$

and thus the convergence is assured if  $\rho(A_{jj} \frac{\Delta t}{2})$  is less than unity. Thus the following theorem is established.

Theorem 7. Given

$$\dot{X} = A X, \quad X(0) = X_0 \quad \text{on } t \in (0, t_f)$$

where  $A$  is a time invariant tridiagonal matrix partitioned such that

$$\begin{aligned} \dot{X}_1 &= A_{11} X_1 + A_{12} X_2 \\ \dot{X}_2 &= A_{21} X_1 + A_{22} X_2 + A_{23} X_3 \\ &\dots \\ \dot{X}_{N-1} &= A_{N-1,N-2} X_{N-2} + A_{N-1,N-1} X_{N-1} + A_{N-1,N} X_N \\ \dot{X}_N &= A_{N,N-1} X_{N-1} + A_{N,N} X_N \end{aligned} \quad (4.29)$$

the algorithm converges if the matrices  $A_{11}, A_{22}, \dots, A_{NN}$  are stable and if the spectral radius of the matrices  $A_{11} \Delta t/2$

$A_{22} \frac{\Delta t}{2}, \dots, A_{NN} \frac{\Delta t}{2}$  is less than unity for a suitable choice of  $\Delta t$ .

#### 4.3.3 Average of rate of convergence

From Equation (4.25) we have,

$$\epsilon^m = P^m \epsilon^0.$$

Using the matrix norm defined in section 4.2.2 and the vector norm, we have

$$\|\epsilon^m\| \leq \|P^m\| \cdot \|\epsilon^0\| \quad (4.30)$$

assuming  $\|\epsilon^0\|$  is not a norm of a null vector then  $\|P^m\|$

gives some idea about the rate of convergence. Thus if

$\|P^m\| < 1$ , then

$$R(P^m) = -\ln [\|P^m\|^{1/m}] = \frac{-\ln \|P^m\|}{m} \quad (4.31)$$

is the average rate of convergence for  $m$  iterations. Eq.

(4.31) can be written as

$$\|P^m\|^{1/m} = e^{-R(P^m)}$$

Therefore,

$$\frac{\|\epsilon^m\|}{\|\epsilon^0\|} \leq e^{-R(P^m)}$$

Let  $\sigma = \|\epsilon^m\|/\|\epsilon^0\|$  and  $[R(P^m)]^{-1} = N_m$ , then

$$\sigma^N \leq \frac{1}{e} \quad (4.32)$$

so that  $N_m$  is the measure of the number of iterations required to reduce the norm of the initial error vector by a factor  $e$ .

In the next section, some examples are presented illustrating the above theory. The examples include an analytical example as well as computer examples.

#### 4.4 Examples

##### 4.4.1 Analytical example

Consider the  $2 \times 2$  matrix equation

$$\begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}, \quad \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

for which the unique solution is  $u_1(t) = u_2(t) = 0$  for all  $t$ ; and let the number of partitions be 2, i.e.,

$$A = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$$

$$\dot{u}_1 = -2 u_1 + u_2 \quad u_1(0) = 0 \quad (4.33)$$

$$\dot{u}_2 = -2 u_2 + u_1 \quad u_2(0) = 0 \quad (4.34)$$

Equations (4.33) and (4.34) can be rewritten as

$$\dot{u}_1(t) = e^{-2t} \int_0^t e^{2\tau} u_2(\tau) d\tau \quad (4.35)$$

$$u_2(t) = e^{-2t} \int_0^t e^{2\tau} u_1(\tau) d\tau \quad (4.36)$$

assume  $u_1^0(t) = \text{constant} = 1$ , then (4.35) and (4.36) gives

$$u_2^{(1)}(t) = e^{-2t} \int_0^t e^{2\tau} d\tau = \frac{1-e^{-2t}}{2}, \quad t \geq 0$$

$$\begin{aligned} u_1^{(1)}(t) &= e^{-2t} \int_0^t e^{2\tau} u_2^{(1)}(\tau) d\tau \\ &= e^{-2t} \int_0^t \frac{e^{2\tau}}{2} (1 - e^{-2\tau}) d\tau \\ &= \frac{1}{4} - \frac{e^{-2t}}{4} - \frac{te^{-2t}}{2}, \quad t \geq 0 \end{aligned}$$

Note that  $u_1^{(1)}(t)$  is always positive; so now to simplify computation, use,

$$u_1^{(1)}(t) = \max_{0 \leq t < \infty} u_1^{(1)}(t) = 1/4$$

Then

$$u_2^{(2)}(t) = e^{-2t} \int_0^t e^{2\tau} \cdot \frac{1}{4} d\tau = \frac{1}{4} \left( \frac{1-e^{-2\tau}}{2} \right)$$

$$\begin{aligned} u_1^{(2)}(t) &= e^{-2t} \int_0^t e^{2\tau} u_2^{(2)}(\tau) d\tau \\ &= \frac{1}{4} \left[ \frac{1}{4} - \frac{e^{-2t}}{4} - \frac{te^{-2t}}{2} \right], \quad t \geq 0 \end{aligned}$$

$$\max_{t \geq 0} u_1^{(2)}(t) = \frac{1}{16}$$

Thus it is obvious that,

$$u_1^{(i+1)}(t) < u_1^{(i)}(t)$$

$$u_2^{(i+1)}(t) < u_2^{(i)}(t)$$

and as  $i \rightarrow \infty$ , the solution  $u_1(t) = u_2(t) = 0$  is obtained.

#### 4.4.2 Computer results

Example 1. Consider a matrix differential equation

$$\dot{X} = A X, \quad X(0) = 4.95$$

where,

Then part

with  $u_i$  (

The initi

same figu

tion afte

convergen

gives the

defined as

The oscil

.005 is d

By

of the ma

are given

where,

$$A = \begin{bmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -2 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix}$$

Then partition A into two partitions as follows.

$$\begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_3 \quad (4.37)$$

$$\begin{bmatrix} \dot{u}_3 \\ \dot{u}_4 \end{bmatrix} = \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} u_3 \\ u_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_2 \quad (4.38)$$

with  $u_i(0) = 4.95$ ,  $i = 1, \dots, 4$ .

The initial guess of  $u_2(t)$  is given in Figure 4.2. In the same figure the exact solution  $u_2(t)$  and the iterative solution after iteration 1 and iteration 2 are given. Note the convergence of the solution to the exact solution. Figure 4.3 gives the convergence of the algorithm. The error norm is defined as,

$$\|e^{(i)}\| = |u_2^{(i+1)}(t) - u_2^{(i)}(t)| \quad (4.39)$$

The oscillation of the error from .005 to .008 and back to .005 is due to the limitation of the analog accuracy.

By the corollary to Theorem 4, the spectral radius of the matrix given in (4.37) and the matrix given in (4.38) are given as follows.

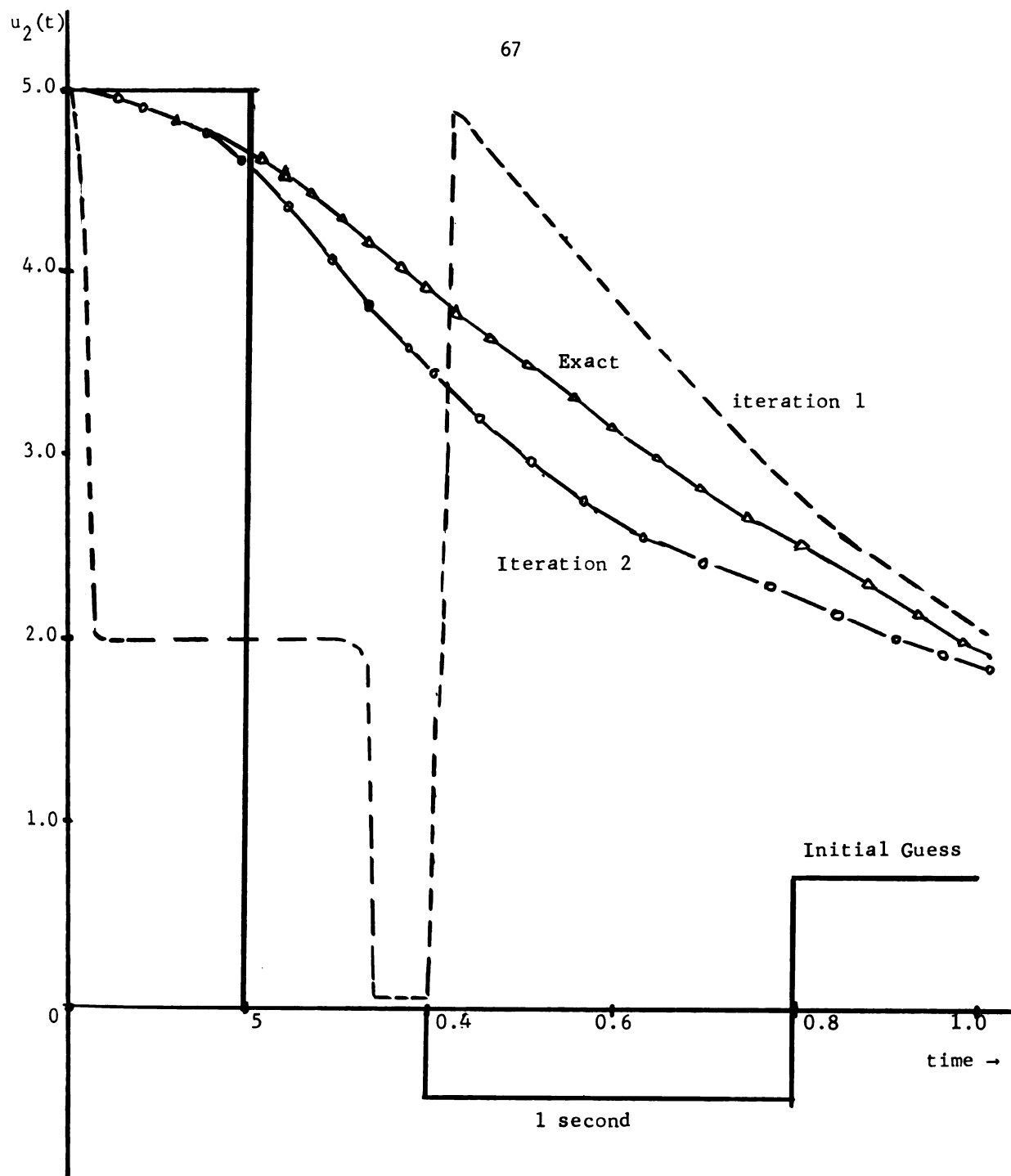
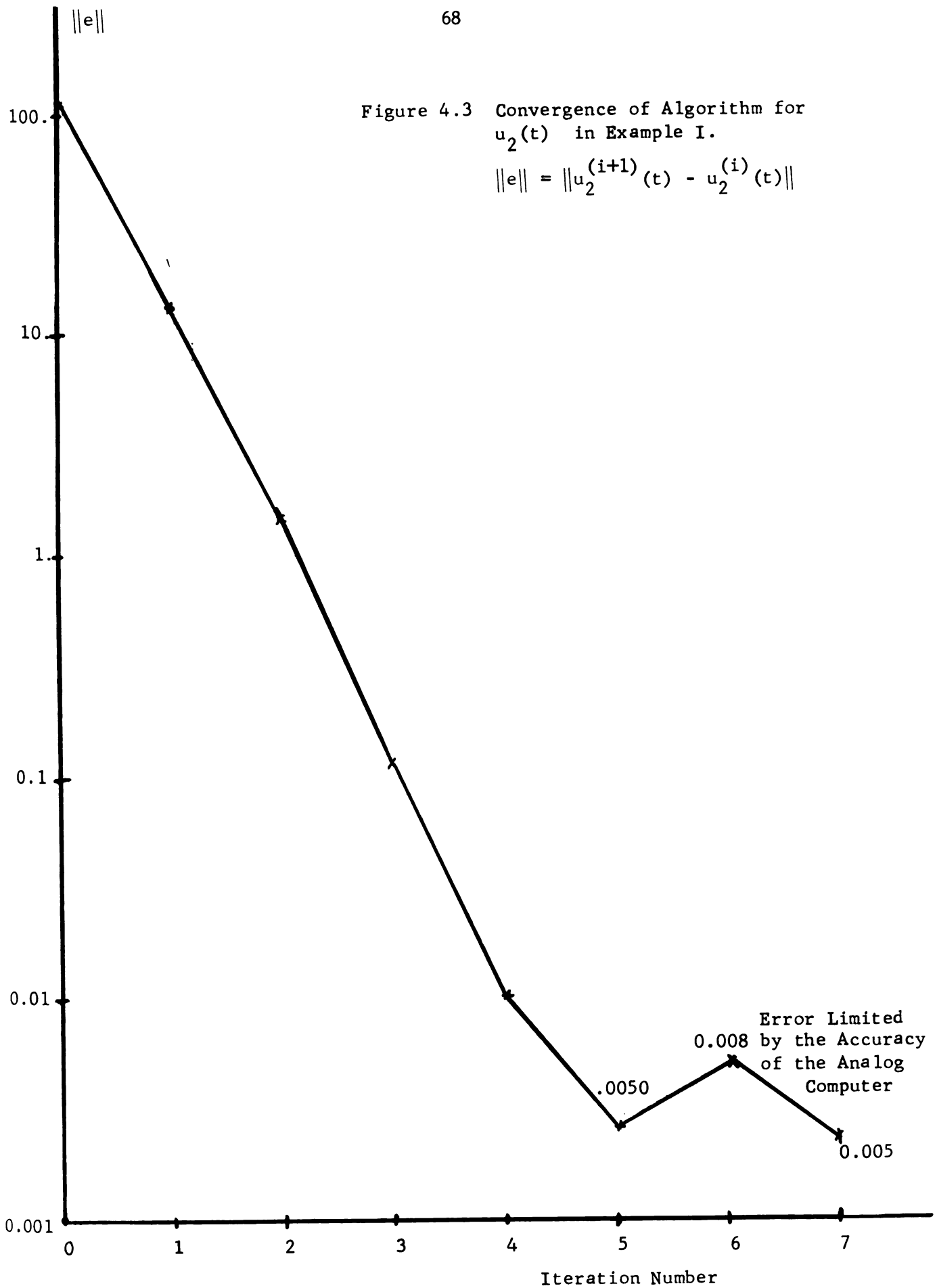


Figure 4.2 Initial Guess, Exact Solution and Iterative Solutions for  $u_2(t)$  in Example I.

Figure 4.3 Convergence of Algorithm for  $u_2(t)$  in Example I.

$$\|e\| = \|u_2^{(i+1)}(t) - u_2^{(i)}(t)\|$$



$$\rho(A_{11}) \leq 3$$

$$\text{and } \Delta t = 1/50$$

$$\rho(A_{22}) \leq 3$$

$$\therefore \rho(A_{11} \Delta t) \leq \frac{3}{50}, \quad \rho(A_{22} \Delta t) \leq \frac{3}{50}$$

and hence the algorithm converges.

Example II. Consider a  $4 \times 4$  matrix and divide it into two partitions, i.e.,  $N = 2$ . The matrix is given as follows.

$$A = \begin{bmatrix} -2 & 1 & 0 & 0 \\ 1 & -2 & 1 & 0 \\ 0 & 1 & -3 & 1 \\ 0 & 0 & 1 & -3 \end{bmatrix} \quad \text{where } \dot{X} = A X, \quad X(0) = 4.95$$

and thus,

$$\begin{aligned} \begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \end{bmatrix} &= \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_3 & \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix} &= \begin{bmatrix} 4.95 \\ 4.95 \end{bmatrix} \\ \begin{bmatrix} \dot{u}_3 \\ \dot{u}_4 \end{bmatrix} &= \begin{bmatrix} -3 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} u_3 \\ u_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_2 & \begin{bmatrix} u_3(0) \\ u_4(0) \end{bmatrix} &= \begin{bmatrix} 4.95 \\ 4.95 \end{bmatrix} \end{aligned}$$

Figure 4.4 shows the initial guesses given for example II.

Figure 4.5 shows the convergence of  $u_2$  with the guesses given in Figure 4.4. The spectral radii are estimated by using the corollary of theorem 4 and  $\Delta t$  is chosen.

$$\rho(A_{11}) \leq 3$$

$$\rho(A_{22}) \leq 4$$

$$\Delta t = \frac{1}{50}, \quad \rho(A_{11} \Delta t) \leq \frac{3}{50},$$

$$\rho(A_{22} \Delta t) \leq \frac{4}{50}.$$

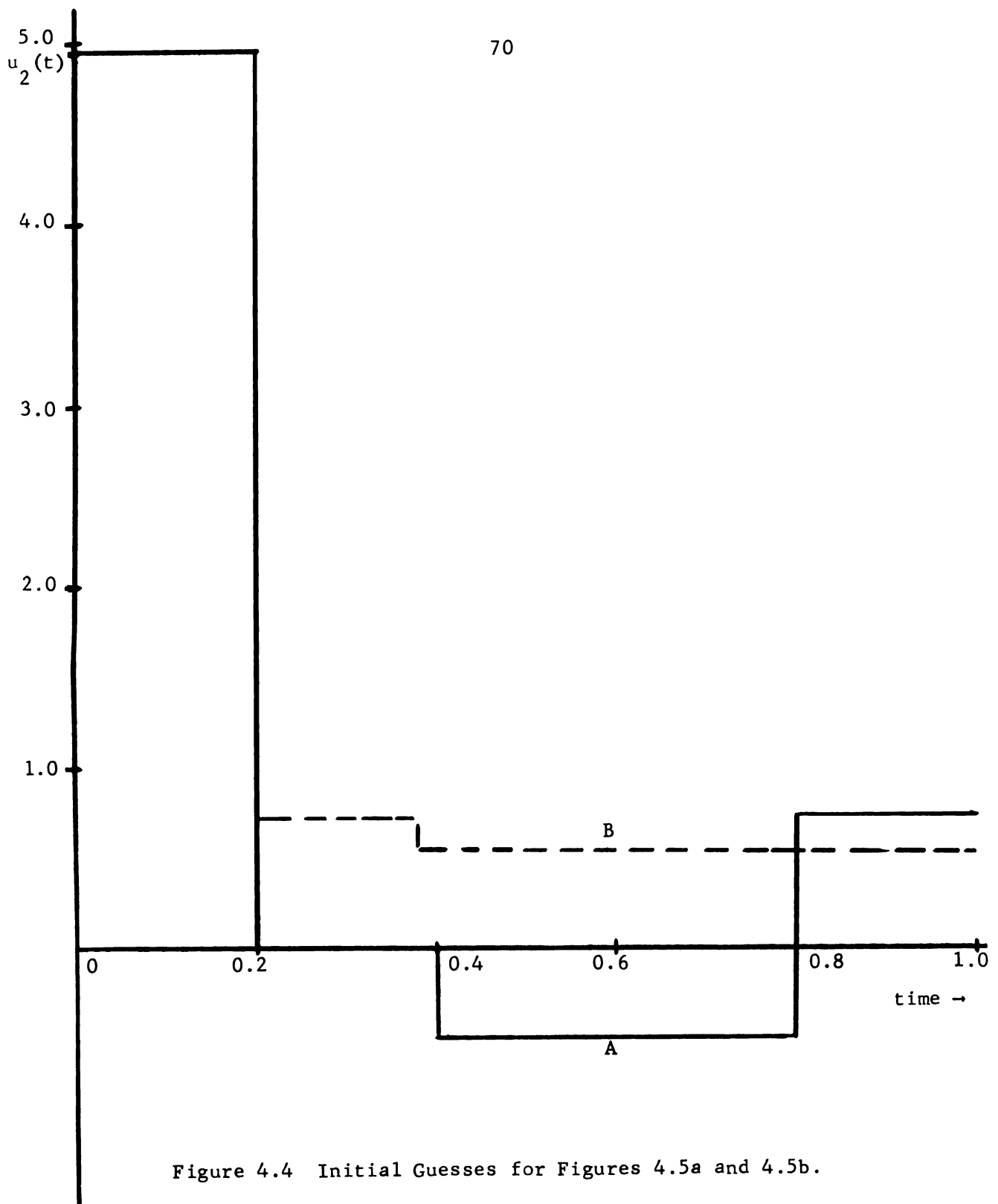


Figure 4.4 Initial Guesses for Figures 4.5a and 4.5b.



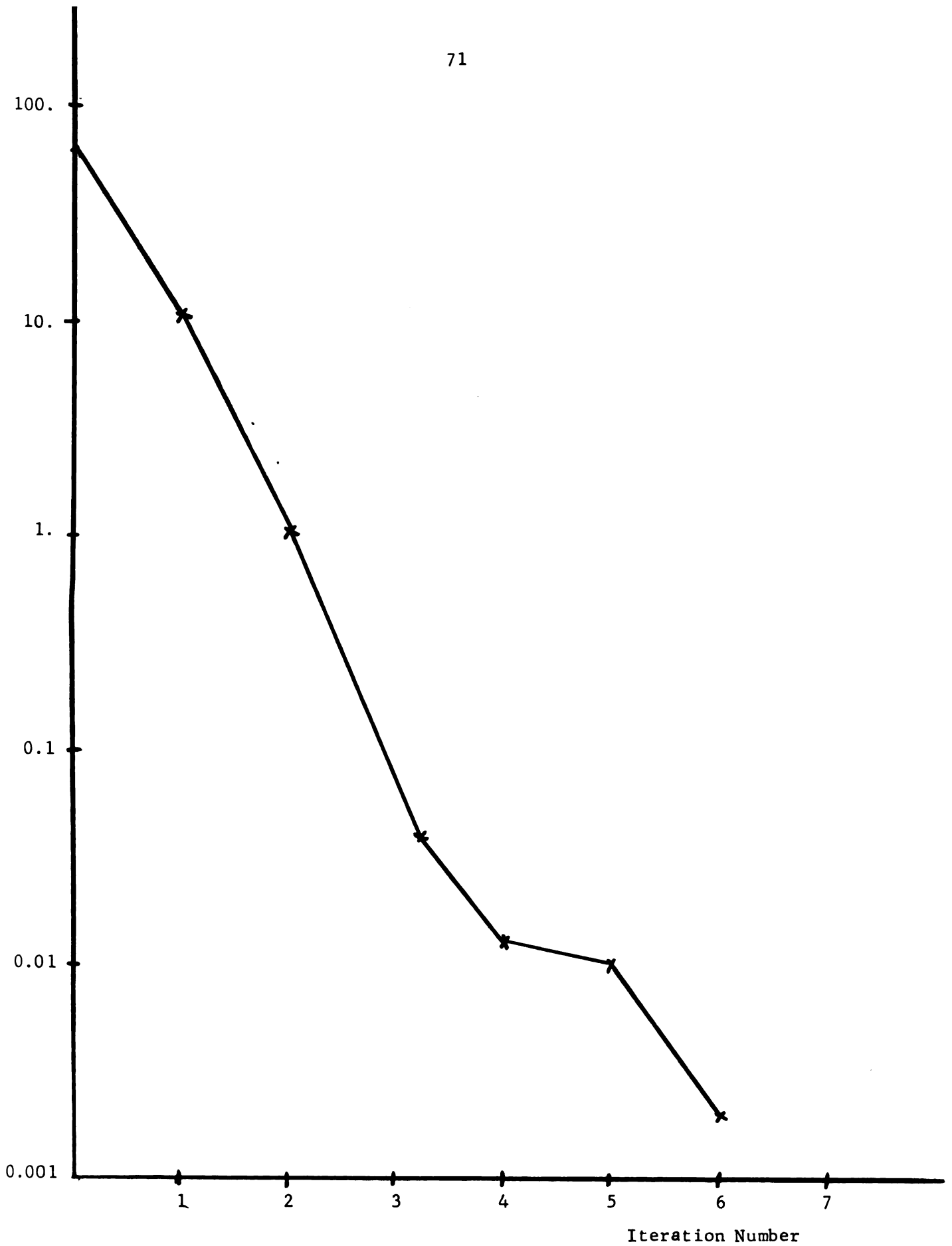
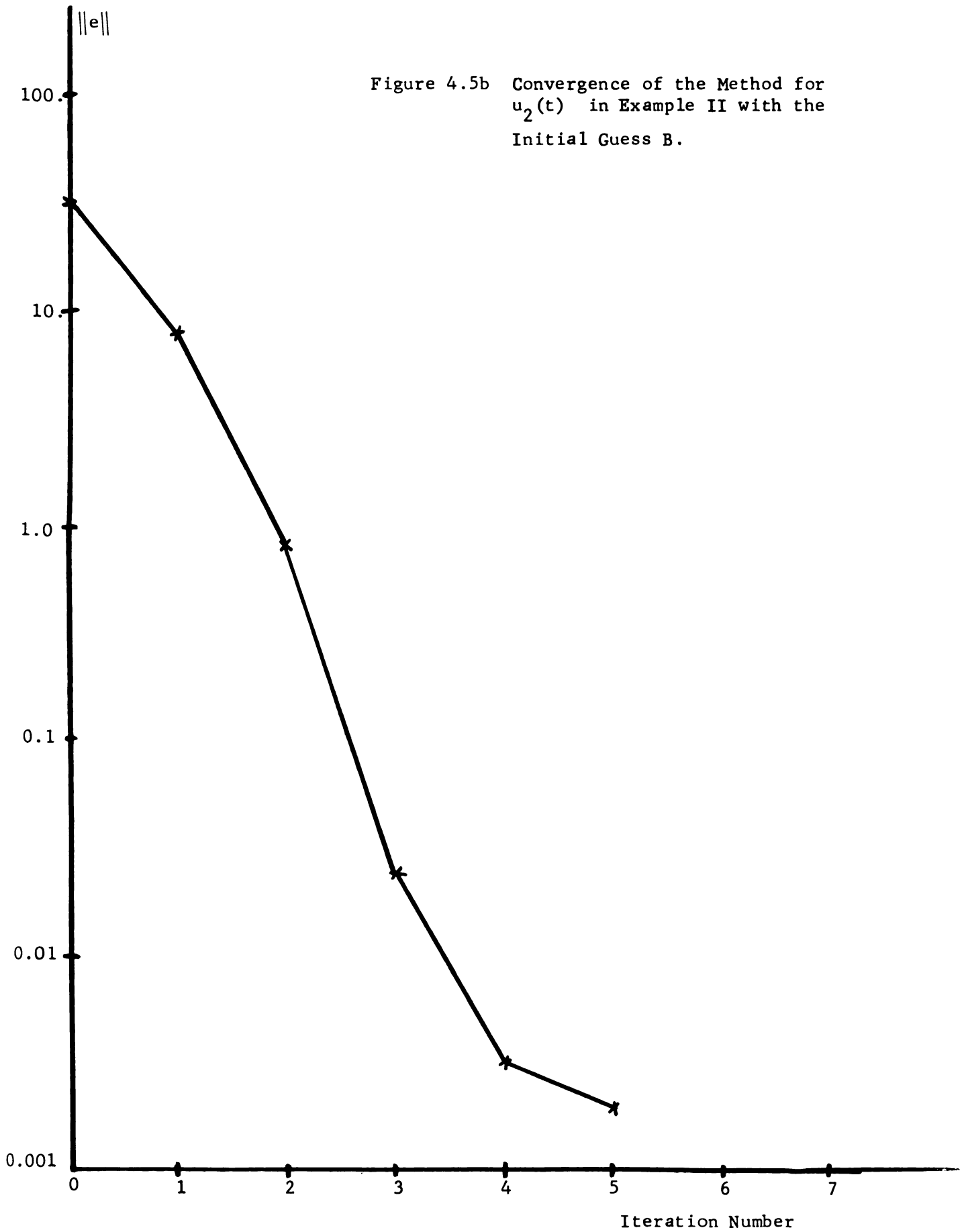


Figure 4.5a Convergence of the Algorithm for  $u_2(t)$  in Example II. for the Initial Estimate A.

Figure 4.5b Convergence of the Method for  $u_2(t)$  in Example II with the Initial Guess B.



Example III. Given an  $8 \times 8$  matrix of the following type.

The exact eigenvalues are known to be  $-4 \sin^2 \pi/18$

$$A = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & -2 \end{bmatrix}$$

and the matrix equation

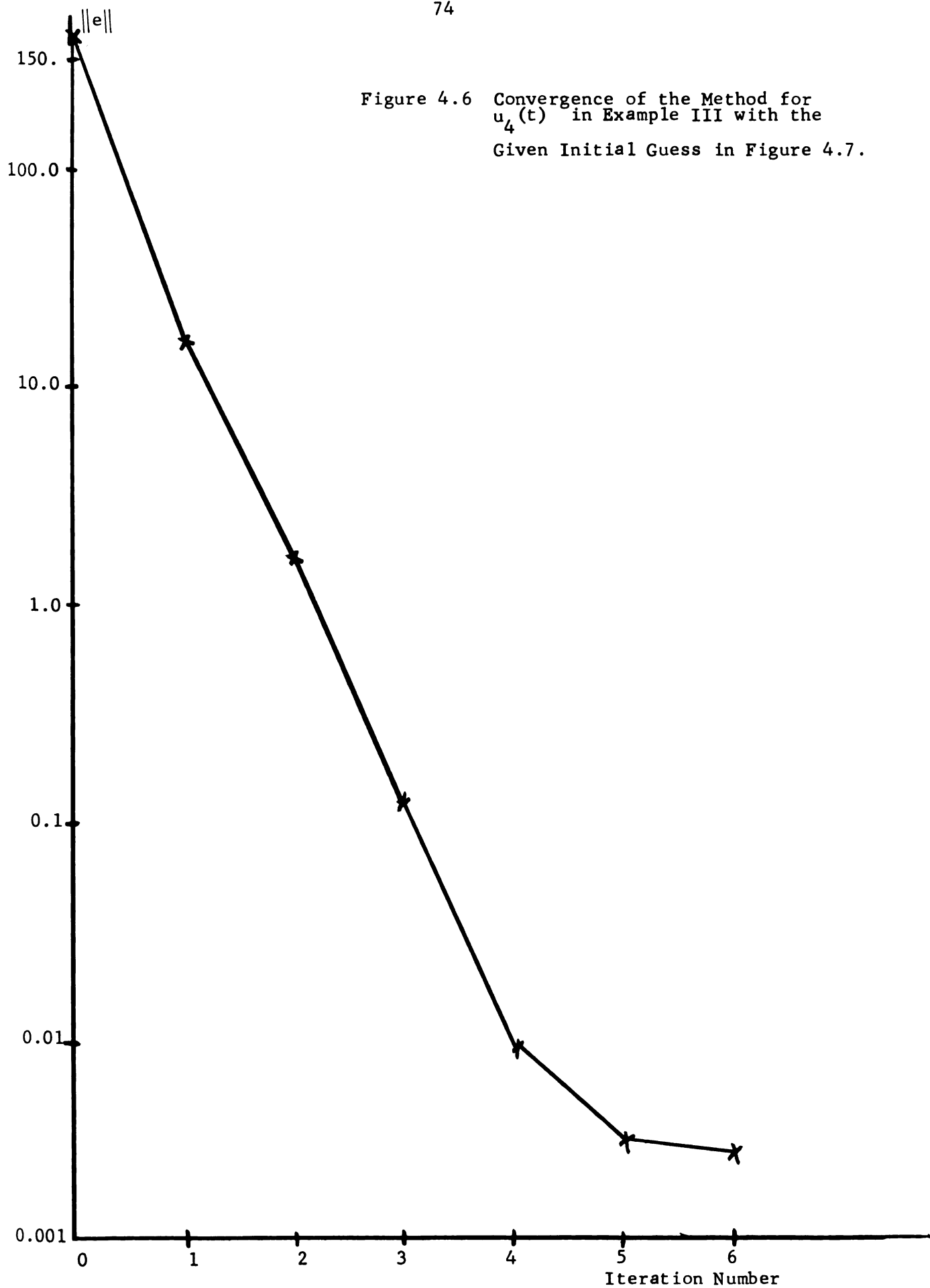
$$\dot{X} = A X ; X(0) = 4.95.$$

The number of partitions are two, i.e.  $N = 2$ . Thus we have two independent  $4 \times 4$  matrices. Figure 4.6 shows the convergence of  $u_4(t)$  and Figure 4.7 gives the initial guess and the value of  $u_4(t)$  after 5 iterations and the exact  $u_4(t)$ . The spectral radii of these matrices are found in the same way as in the previous examples. The value of  $\Delta t$  is chosen as  $1/50$ , and the interval of interest is  $(0,1)$ .

Example IV. Finally a  $6 \times 6$  matrix is considered and the number of partitions are 3, i.e.  $N = 3$ . The partitioned matrices and the given matrix are given below.

$$A = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{bmatrix}$$

Figure 4.6 Convergence of the Method for  $u_4(t)$  in Example III with the Given Initial Guess in Figure 4.7.



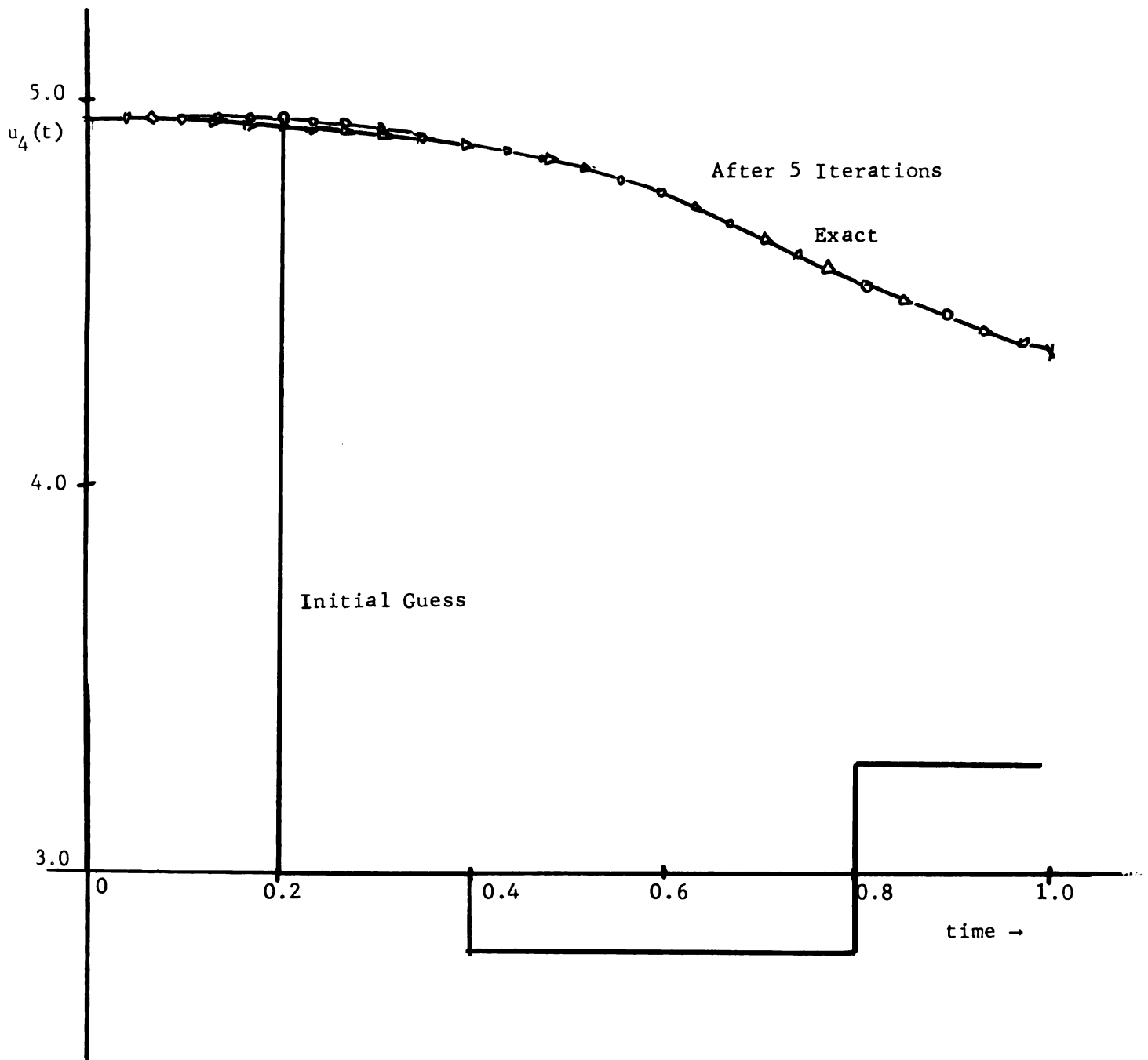


Figure 4.7 Initial Guess, Exact Solution, and the Iterative Solution for  $u_4(t)$  in Example III.

and

Figure 4

initial g

are diff

gives th

The spect

is chosen

because

The spect

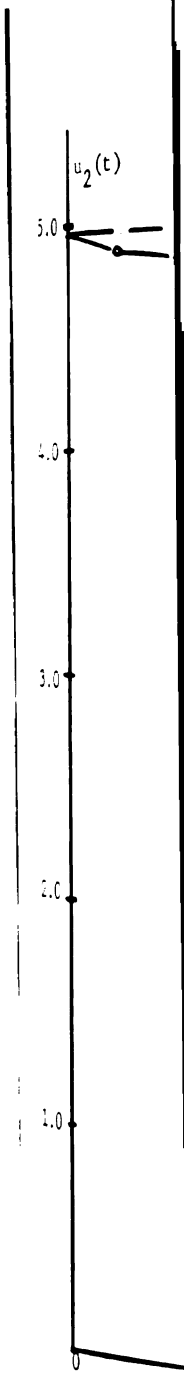
which end

and

$$\begin{aligned} \begin{bmatrix} \dot{u}_1 \\ \dot{u}_2 \end{bmatrix} &= \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u_3, \quad \begin{bmatrix} u_1(0) \\ u_2(0) \end{bmatrix} = \begin{bmatrix} 4.95 \\ 4.95 \end{bmatrix} \\ \begin{bmatrix} \dot{u}_3 \\ \dot{u}_4 \end{bmatrix} &= \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} u_3 \\ u_4 \end{bmatrix} + \begin{bmatrix} u_2 \\ u_5 \end{bmatrix}, \quad \begin{bmatrix} u_3(0) \\ u_4(0) \end{bmatrix} = \begin{bmatrix} 4.95 \\ 4.95 \end{bmatrix} \\ \begin{bmatrix} \dot{u}_5 \\ \dot{u}_6 \end{bmatrix} &= \begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} u_5 \\ u_6 \end{bmatrix} + \begin{bmatrix} 1 \\ 0 \end{bmatrix} u_4, \quad \begin{bmatrix} u_5(0) \\ u_6(0) \end{bmatrix} = \begin{bmatrix} 4.95 \\ 4.95 \end{bmatrix} \end{aligned}$$

Figure 4.8 shows a sample function  $u_2(t)$  along with given initial guess. The exact solution and the iterated solutions are difficult to distinguish after 5 iterations. Figure 4.9 gives the convergence of  $u_2(t)$  with the norm given in (4.39). The spectral radii are same as given in example I, and  $\Delta t$  is chosen to be  $1/50$ . The interval of interest is  $(0,1)$ .

Thus these examples show the usefulness of this method because of the fast convergence exhibited by these examples. The spectral radius can be obtained by using the theorems given which enables one to choose  $\Delta t$ .



Figure

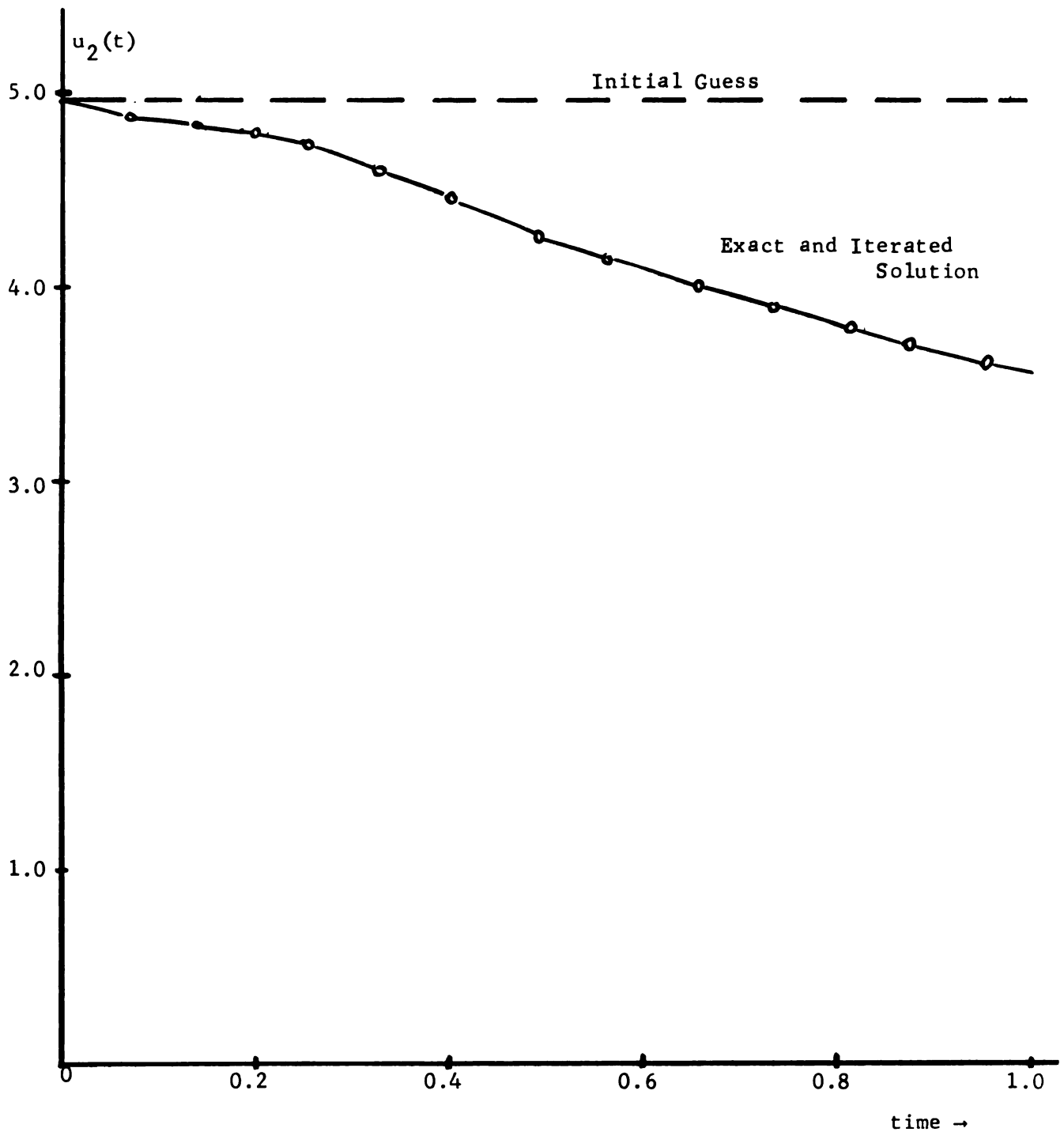


Figure 4.8 Initial Guess, Exact Solution, and the Iterative Solution for  $u_2(t)$  in Example IV.

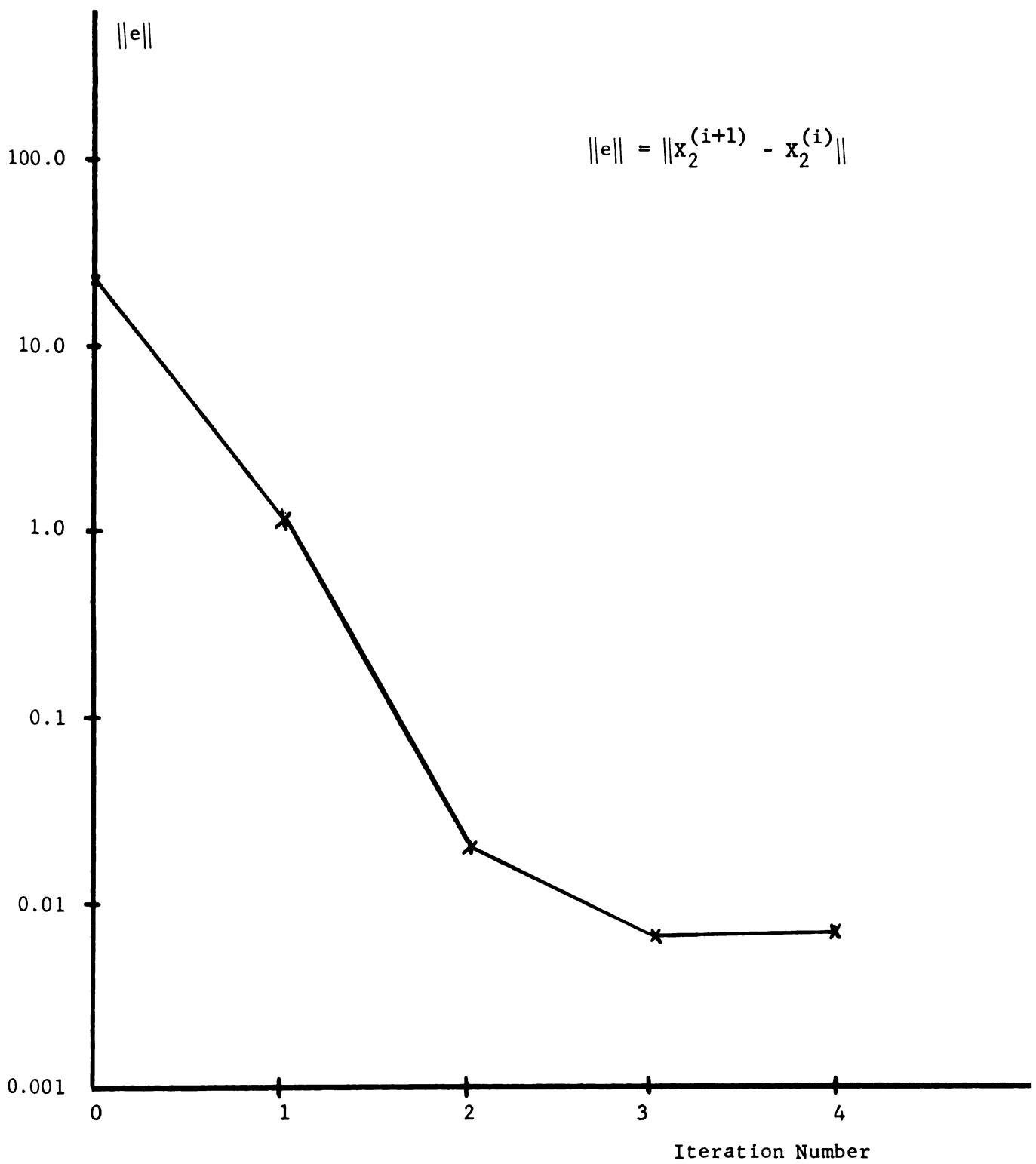


Figure 4.9 Convergence of the Method for Three Partitions in Example IV for a Sample Function  $u_2(t)$ .

## CHAPTER V

### DEVELOPMENT OF ALGORITHM - I

#### 5.1 Introduction

In this chapter, a description of Algorithm-I for obtaining an optimal feedback control for a class of distributed parameter systems is given. This method is different from the existing methods in the following way:

- i) No a priori information of the existence of the optimal open-loop control is necessary.
- ii) The disadvantage of computing optimal open-loop control, whenever there is a change in the initial distribution is removed.
- iii) The computational method for obtaining feedback parameters is more efficient than the existing methods in the sense that more accurate solutions could be obtained. This is because of the extended capabilities of obtaining solutions for larger differential system using the decomposition algorithm described in Chapter IV.

The problem is formulated in the following sections and the algorithm is developed. The algorithm is illustrated by an example with different constraints on the control.

## 5.2 Problem Formulation

Consider a distributed system characterized by a vector partial differential equation

$$\frac{\partial Q}{\partial t} = G[t, x, Q(t, x), Q_x(x, t), Q_{xx}(x, t), u(t), m(x, t)] \quad (5.1)$$

along with the boundary condition

$$Q(x, t_0) = Q_0(x) \quad x \in \Omega, \quad t \in [0, t_f] \quad (5.2)$$

$$S_b Q(x_b, t) = u(x_b, t) \quad x_b \in \Omega_b, \quad t \in [0, t_f] \quad (5.3)$$

where the symbols are explained as follows.

$\Omega$  : a given finite (connected) region in Euclidean n-space and  $\Omega_b$ , the boundary of  $\Omega$ .

$G$  : spatially varying differential operator on  $Q$  which may include parameters which are linear functions of  $Q, m, x$  or  $t$ .

$Q_0(x)$  : initial state vector, i.e., at  $t = 0$ ,  $Q(x, t) = Q_0(x)$

$u(t)$  : boundary control.

## 5.3 Algorithm-I

The algorithm considered here involves forming a semi-discrete approximation of (5.1) through (5.3) by placing a grid on the spatial domain.

The spatial variables are discretized by defining a vector,

$$X_i = (i_1(\Delta x_1), i_2(\Delta x_2), \dots, i_j(\Delta x_j), \dots, i_n(\Delta x_n))'$$

which in effect places the grid on the region  $\Omega$ . The prime denotes the transpose. Here the elements of  $i = [i_1, i_2, i_3, \dots, i_n]$  are integers defined by  $i_j = 0, 1, \dots, N_j$  where

$$N_j = \frac{(x_j)_{\max} - (x_j)_{\min}}{\Delta x_j} \quad (5.4)$$

Assuming that the operator  $G$  is at most second order in  $X$ , it can be approximated as follows:

$$G(Q(x_i, t), m(x_i, t), x_i, t) \approx$$

$$G_i[Q_i(t), Q_{i+I_1}(t), Q_{i+I_2}(t), \dots, Q_{i+I_n}(t), m_i(t), u(t), t] \quad (5.5)$$

where  $I_k = \{i/i = 0 \text{ except for the } k\text{th element which equals to } 1\}$

$i$  ranges over all the interior mesh points, and the functions  $G_i$  are assumed to be real valued and class  $C^2$ . As an example, consider a rectangular mesh in  $E^2$  and using the above notation,  $i = (i_1, i_2)$  and consider the mesh point  $(1,1)$ . Then  $I_1 = (1,0)$  and  $I_2 = (0,1)$ . Therefore the points that will be considered are  $(1,1)$ ,  $(2,1)$ ,  $(0,1)$ ,  $(1,2)$ ,  $(1,0)$ .

Thus following the above notation, the discretized vector partial differential equation in (5.1) can be rewritten along with (5.2) and (5.3) as follows.

$$\frac{\partial Q_i}{\partial t} = G_i[Q_i(t), Q_{i+I_1}(t), Q_{i+I_2}(t), \dots, Q_{i+I_n}(t), m_i(t), t, u(t)]$$

$$X_i \in \Omega, \quad t \geq 0 \quad (5.6)$$

along with the discretized versions of the boundary conditions,

$$Q_i(t = 0) = Q_{i0} \quad x_i \in \Omega, t \geq 0 \quad (5.7)$$

$$S_{bi} Q_i = u_i(t) \quad x_i \in \Omega_b, t \geq 0 \quad (5.8)$$

Now, before stating the optimal feedback control problem, a brief discussion to motivate the material in this chapter is given.

In general, there are several methods of obtaining open-loop control for distributed parameter systems. Some of the methods are discussed in Chapter III. But in practice, it is desirable to have a closed-loop control such as optimal control as a function of state and possibly of time. Thus in the case of distributed systems, feedback methods similar to lumped parameter systems can be discussed. Though it is difficult to obtain analytically the controller for a large class of problems, assuming that it is possible, the implementation of this control law is difficult. This difficulty arises because of the infinite dimensional character of the state vector which is a function of spatial domain as well as time. So some kind of approximation is necessary so that it is possible to reconstruct the state function by a finite number of measurements along the spatial domain, while keeping the time continuous. Then a polynomial fit can be used to get the complete state function. The coefficients of the approximating polynomial will vary with time. Thus it can be visualized as a black box containing a device which has as its inputs, the values of the state measured at finite number of points along the spatial domain and as its output the coefficients of the

specified degree best fit polynomial in some given sense (as an example Chebychev fit of  $n$ th degree). This is illustrated in Figure 5.1, for fixed  $t$ ,

$$Q(x,t) \approx a_n(t)q^n(t) + a_{n-1}(t)q^{n-1}(t) + \dots + a_1(t)q(t) + a_0(t) \quad (5.9)$$

Since the coefficients are time varying, the polynomial fit is very difficult to perform. Another method could be to fit the polynomial at each time. Since the coefficients are different at each time, this requires a large computation time at each time interval and if the state variable is of higher dimension than one, the polynomial fit is difficult to perform partially because the theory of polynomial approximation is not very well developed in higher dimensions.

In light of the above discussion, it is desirable to obtain a feedback control in terms of measurements made at a finite number of points in the spatial domain. Now let  $u_c(t)$  be denoted by the feedback control law which is written as,

$$u_c(t) = \sum_{r=1}^n K_r F(Q_r(t), Q_{dr}(t)) \quad (5.10)$$

where  $K_r$  is either zero or an unknown matrix and  $F$  is a suitably chosen function of the state vector and the desired state  $Q_{dr}(t)$  discretized by a finite difference scheme as discussed earlier. Thus, given

$$\frac{\partial Q_i}{\partial t} = G_i(Q_i(t), Q_{i+1}(t), \dots, Q_{i+I_n}(t), t), \quad i = 1, \dots, n$$

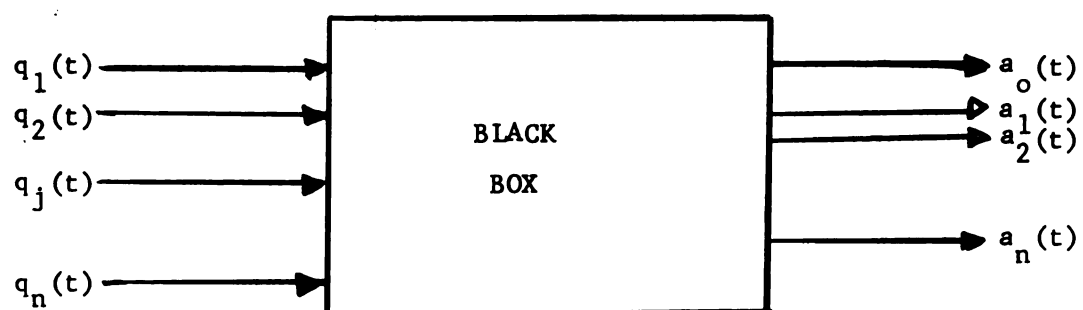


Figure 5.1 Illustration for the Discussion in Section 5.3.

$$Q(x, t) \approx a_n(t) q^n(t) + a_{n-1} q^{n-1}(t) + \dots + a_1(t) q(t) + a_0(t)$$

(t fixed)

along with the boundary conditions

$$Q_i(t = 0) = Q_{i0} \quad x_i \in \Omega, t \geq 0$$

$$S_{b_i} Q_i(t) = u_i(t) \quad x_i \in \Omega_b, t \geq 0$$

$$|u_i(t)| \leq 1$$

to find the optimal control of the form given in (5.10) such that the following performance index is minimized.

$$C = \int_{\Omega} [Q_d(x, t_f) - Q(x, t_f)]^T [Q_d(x, t_f) - Q(x, t_f)] d\Omega$$

which by the same discretization scheme, becomes,

$$C \approx \bar{C} = \sum_{i=1}^n [Q_d(x_i, t_f) - Q(x_i, t_f)]^T [Q_d(x_i, t_f) - Q(x_i, t_f)] \quad (5.11)$$

$x_i \in \Omega$

There are in general two ways of obtaining the feedback parameter  $K_r$ , given in (5.10). One of the methods is to obtain  $K_r$  such that (5.11) is minimized. The second method is to obtain  $K_r$  such that

$$\int_0^{t_f} [u^*(t) - u_c(t)]^T [u^*(t) - u_c(t)] dt \quad (5.12)$$

is minimized, where  $u^*(t)$  is the optimal open-loop control obtained by one of the existing methods. The first method is discussed in this chapter.

To simplify the subsequent derivations, a linear diffusion system will be considered. Thus given the system,

$$\frac{\partial g}{\partial t}$$

with the b

$$\frac{\partial g}{\partial x}$$

$$\frac{\partial g}{\partial x}$$

$$q(x)$$

The perform

Discretize

(5.4), the

where

where prime

are  $n \times n$

tridiagonal

Now assume

$$u(t)$$

Then incorp

$$\frac{\partial q}{\partial t} = \frac{\partial^2 q}{\partial x^2}, \quad \Omega = [0,1]$$

with the boundary conditions

$$\left. \frac{\partial q}{\partial x} \right|_{x=0} = \alpha \{ q(0,t) - u(t) \} \quad (5.13)$$

$$\left. \frac{\partial q}{\partial x} \right|_{x=1} = 0$$

$q(x,0) = 0$  , and with the constraint,

$$0 \leq u(t) \leq 1$$

The performance index to be minimized is

$$J = \int_0^1 [q_d(x) - q(x,t_f)]^2 dx \quad (5.14)$$

Discretize (5.13) and (5.14) using the scheme explained in (5.4), the discretized system can be written as,

$$\dot{X} = A X + b u(t) \quad X(0) = X_0 \quad (5.15)$$

where

$$X = [q_1(t), q_2(t), \dots, q_n(t)]'$$

where prime indicates the transpose. The matrices  $A$  and  $b$  are  $n \times n$  and  $n \times 1$  respectively. The matrix  $A$  is of the tridiagonal form.

Now assume

$$u(t) = u_c(t) = \sum_{i=1}^n K_i q_i(t) = \langle K, X(t) \rangle \quad (5.16)$$

Then incorporating (5.16) into (5.15), yields

$$\dot{X} = \bar{A} X, \quad X(0) = X_0 \quad (5.17)$$

where  $\bar{A}$  contains the unknown feedback parameters to be determined. The solution of (5.17) with the given initial condition is,

$$X(t) = e^{\bar{A}t} X_0 \quad (5.18)$$

assuming that the eigenvalues of  $\bar{A}$  in terms of the unknown parameters are known. Expand (5.18) into a constituent matrix expansion to get,

$$X(t) = F(K_1, K_2, K_3, \dots, K_r, t) X_0 = f(K_1, K_2, \dots, K_r) \quad (5.19)$$

Similar use of the discretization scheme for the performance index in (5.14) gives,

$$\Delta J = \sum_{i=1}^n [q_d(x_i) - q(x_i, t_f)]^2 \quad (5.20)$$

Substituting (5.19) into (5.20), we obtain

$$\Delta J = \sum_{i=1}^n [q_d(x_i) - f(K_1, K_2, \dots, K_r)]^2 \quad (5.21)$$

Since  $u_c(t)$  is constrained to be in the limits

$$0 \leq u_c \leq 1 \quad (5.21a)$$

Now (5.21) is minimized with a given constraint (5.21a), and the desired distribution  $q_d$  to obtain the parameters  $k_1, k_2, \dots, k_r$ . Thus the problem is reduced into a parameter optimization problem in a parameter space of n-dimensions.

Since the matrix  $A$  in (5.15) contains unknown constants  $k_1, k_2, \dots, k_r$ , the eigenvalues are very difficult to obtain in terms of  $k_1, k_2, \dots, k_r$ , for large matrices. An

alternative method is to implement this method on the hybrid computer. Several automatic parameter optimization schemes with differential constraints are discussed in the literature. Some of these methods are discussed in Bekey and Karplus (B-10).

In most of these methods, the differential equations are simulated on the analog computer and the computation of the gradient and adjustment of the parameters is done by the digital computer. Since the number of integrators available on an analog computer are limited, this method limits the order of spatial discretization. So the decomposition principle discussed in Chapter IV is used to increase the capabilities of the analog computer. This enables one to solve a higher order differential equation than that is usually possible with the available integrators. The flow chart of this algorithm is given in Figure 5.2.

#### 5.4 Computer Results

As an example, the following problem is considered.

Consider

$$\frac{\partial q}{\partial t} = \frac{\partial^2 q}{\partial x^2} \quad (5.22)$$

where  $q(x,t)$  is the temperature distribution in the metal in dependence on the space coordinate  $x$ , ( $0 \leq x \leq 1$ ) and time  $t$  ( $0 \leq t \leq T$ ). The space coordinate  $x$  is normalized with respect to the thickness of the metal and  $t$  is normalized so that the coefficients corresponding to the thermal diffusivity is unity. The initial and boundary conditions are given by

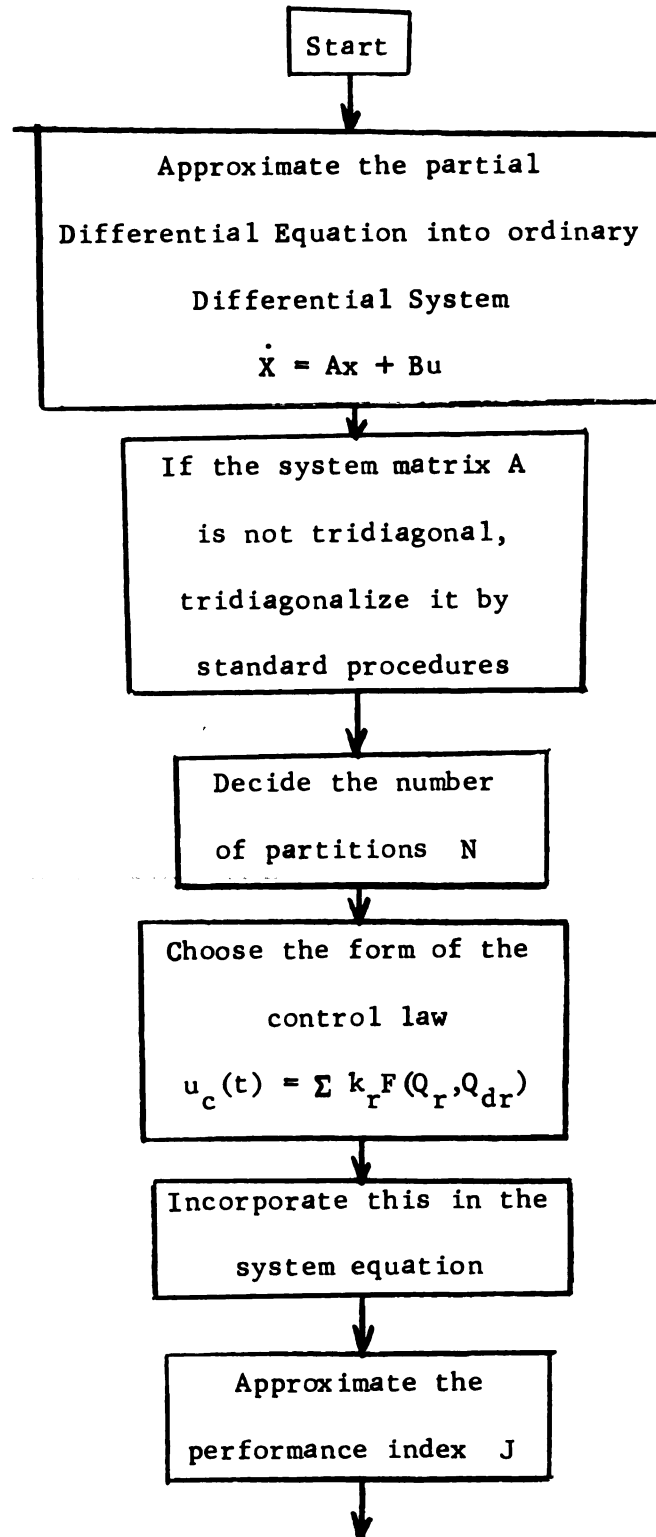


Figure 5.2 Flow Chart for Algorithm - I

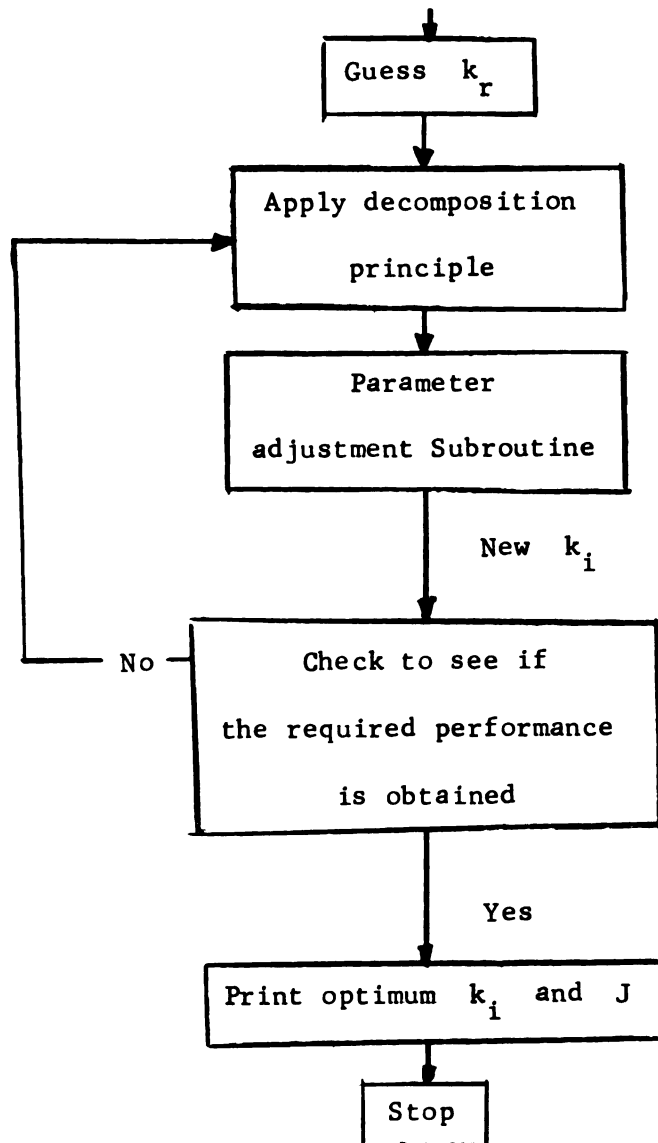


Figure 5.2 Flow Chart for Algorithm - I

$$q(x,0) = 0$$

$$\left. \frac{\partial q}{\partial x} \right|_{x=0} = \alpha \{ q(0,t) - v(t) \} \quad (5.23)$$

$$\left. \frac{\partial q}{\partial x} \right|_{x=1} = 0$$

where  $\alpha$ , the heat transfer coefficient assumed to be constant,  
 $v(t)$ , the temperature of the gas medium is controlled by the  
 fuel flow  $u(t)$  and they are related by

$$r \frac{dv}{dt} + v(t) = u(t) \quad (5.24)$$

where  $r$  is the time constant of the furnace, and  $u(t)$  is  
 normalized properly.

The problem here is to obtain  $u(t)$  ( $0 \leq t \leq T$ ) such that

$$I[u(t)] = \int_0^1 [q^*(x) - q(x,T)]^2 dx \quad (5.25)$$

is minimized.

Furthermore  $u(t)$  is constrained to

$$0 \leq u(t) \leq 1 \quad (5.26)$$

The various constants in the above problem are:

$$\begin{array}{ll} \alpha = 10 & 0 \leq x \leq 1 \\ r = 0.04 & , \quad q^*(x) = 0.2 \end{array}$$

Now we discretize the system using the scheme mentioned in  
 section 5.2, to obtain

$$\frac{d}{dt} \begin{bmatrix} v(t) \\ q_1(t) \\ q_2(t) \\ \vdots \\ q_9(t) \end{bmatrix} = 100 \begin{bmatrix} -0.25 & 0 & 0 & 0 & \dots & 0 \\ 0.5 & -1.5 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & & 0 \\ & & & \ddots & & \\ 0 & & & & 0 & 0 & 1 & -2 & 1 \\ 0 & & & & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} v(t) \\ q_1(t) \\ q_2(t) \\ \vdots \\ q_9(t) \end{bmatrix} + \begin{bmatrix} 25 \\ 0 \\ \vdots \\ 0 \end{bmatrix} u(t) \quad (5.27)$$

$$u(t) = u_c(t) = \sum_{i=1}^9 k_r q(x_r, t)$$

where  $k_r$  is either a constant gain or zero, depending on the number of points to be sampled. Suppose we sample at all nine points, then (5.27) becomes,

$$d/dt (X(t)) = A(k_1, k_2, \dots, k_9) X(t) \quad (5.28)$$

$$X(0) = X_0 = 0$$

where  $X(t) = q(t)/100$ , and  $A$  is given as follows.

$$A(k_1, k_2, \dots, k_9) = \begin{bmatrix} -0.25 & k_1 & k_2 & k_3 & \dots & k_9 \\ 0.5 & -1.5 & 1 & 0 & \dots & 0 \\ 0 & 1 & -2 & 1 & \dots & 0 \\ & & & \ddots & & \\ & & & & 1 & -2 & 1 \\ & & & & 0 & 1 & -1 \end{bmatrix} \quad (5.29)$$

Now to apply the decomposition principle, the matrix is partitioned as follows.

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad (5.30)$$

where  $A_{11}$ ,  $A_{12}$ ,  $A_{21}$ , and  $A_{22}$  are  $5 \times 5$  matrices. Also let

$$X = (Y_1', Y_2')$$

where prime denotes transpose. Then (5.28) can be written as

$$\begin{aligned}\dot{Y}_1 &= A_{11} Y_1 + A_{12} Y_2 \\ \dot{Y}_2 &= A_{21} Y_1 + A_{22} Y_2\end{aligned}\tag{5.31}$$

From the matrix given in (5.29), the decomposition principle requires the storage of all the functions in the vector  $Y_2$ , but a little modification using the superposition principle avoids this difficulty.

Let

$$v(t) = v_1(t) + v_2(t)$$

where

$$\frac{dv(t)}{dt} = a v(t) + k_1 q_1(t) + k_2 q_2(t) + \dots + k_9 q_9(t)$$

Then writing

$$\frac{d}{dt} v_1(t) = a v_1(t) + k_1 q_1(t) + k_2 q_2(t) + \dots + k_4 q_4(t)$$

$$\frac{d}{dt} v_2(t) = a v_2(t) + k_5 q_5(t) + \dots + k_9 q_9(t)$$

two different partitions of the matrix in (5.29) are obtained as follows.

$$\dot{Y}_1 = \begin{bmatrix} -0.25 & k_1 & k_2 & k_3 & k_4 \\ 0.5 & -1.5 & 1 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 1 & -2 \end{bmatrix} Y_1 + \begin{bmatrix} 0 \\ 0.5 v_2 \\ 0 \\ 0 \\ x_5 \end{bmatrix} \quad (5.32)$$

$$A_{22} = \begin{bmatrix} -2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 1 & -1 & 0 \\ k_5 & k_6 & k_7 & k_8 & k_9 & -0.25 \end{bmatrix}$$

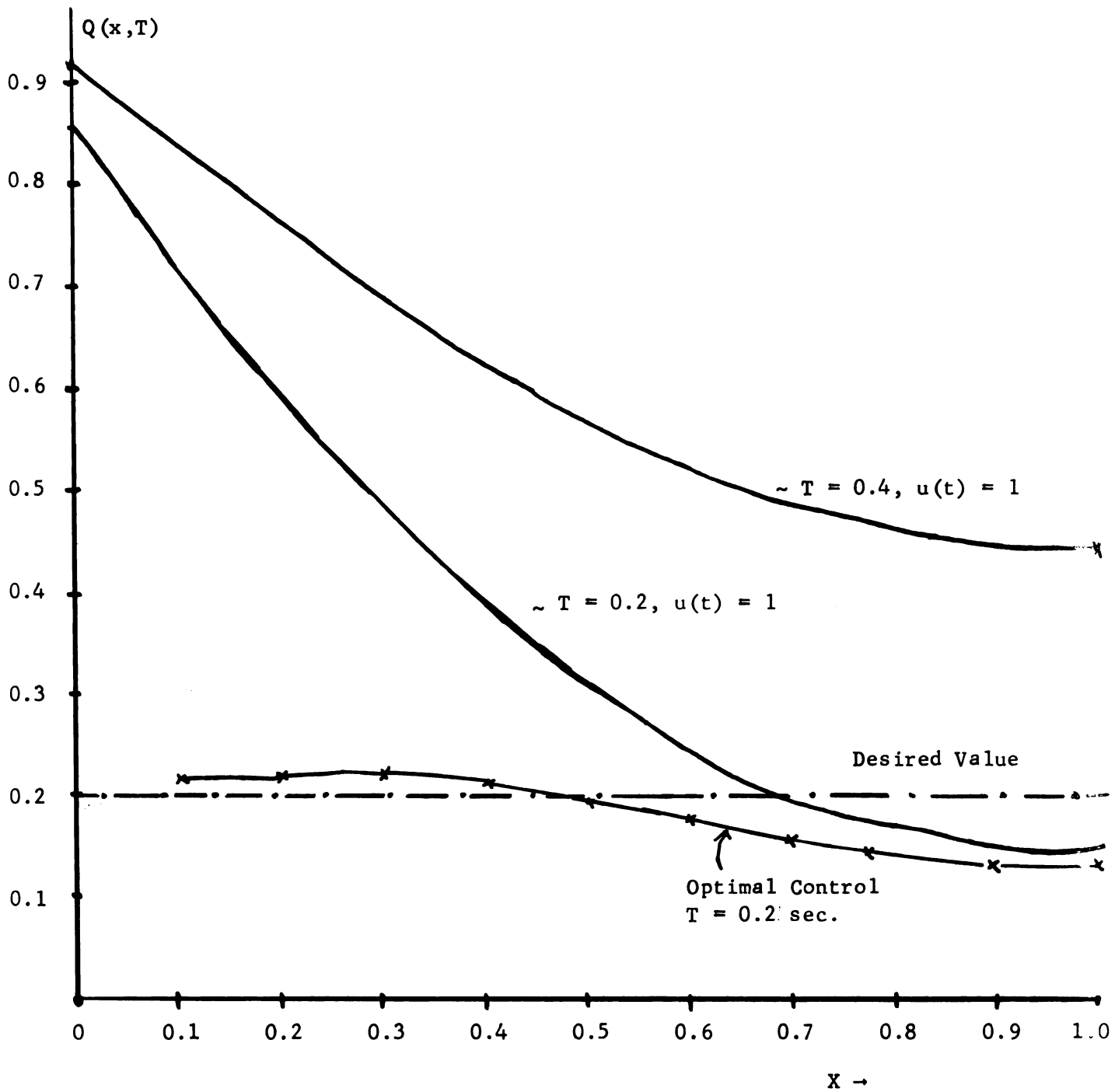
$$\text{and } b = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$\text{and } \dot{Y}_2 = A_{22} Y_2 + b x_4 \quad (5.33)$$

and thus storage of only three functions are necessary. The vectors  $Y_1$  and  $Y_2$  are as follows.

$$Y_1 = \begin{bmatrix} v_1(t) \\ q_1(t) \\ q_2(t) \\ q_3(t) \\ q_4(t) \end{bmatrix} \quad Y_2 = \begin{bmatrix} q_5(t) \\ q_6(t) \\ q_7(t) \\ q_8(t) \\ q_9(t) \\ v_2(t) \end{bmatrix} \quad (5.34)$$

The results are summarized below. Figure 5.3 represents the verification of Sakawa's results using the hybrid computer. The optimal control obtained by Sakawa for this problem is applied and the resulting state is verified. Figure 5.4 and Figure 5.5 represent the state functions obtained with two feedback



$\Delta X = 0.1$  ,  $0 \leq u(t) \leq 1$

Figure 5.3 Verification of Sakawa's Results.

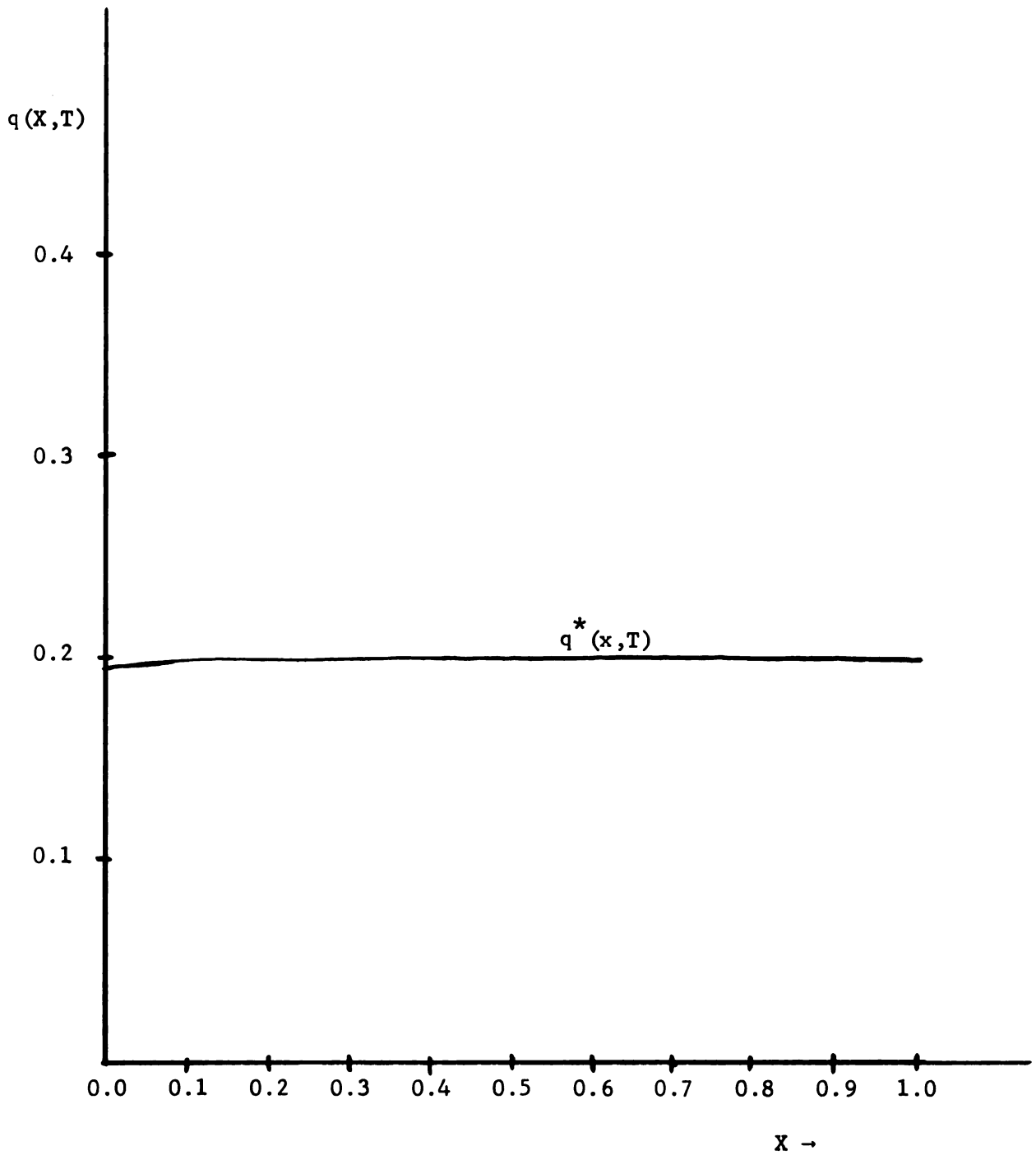


Figure 5.4 Unconstrained Case

$$q^*(x, T) = 0.2 \quad \text{and}$$

$$q(x, 0) = 0.$$

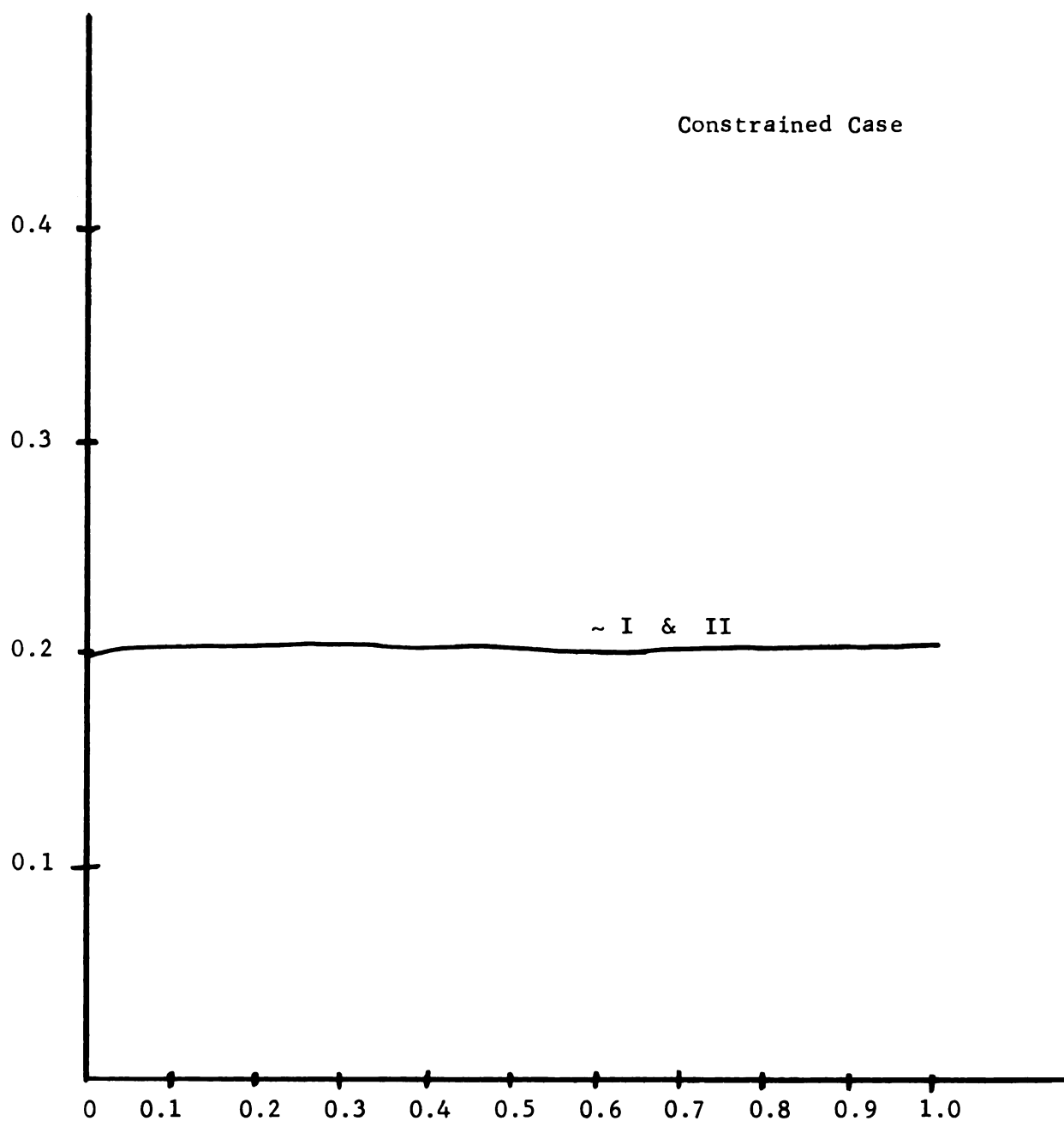


Figure 5.5 Constrained Case

$$0 \leq u_c \leq 10 \quad (\text{I})$$

$$0 \leq u_c \leq 1 \quad (\text{II})$$

parameters and the state in these cases is measured at  $x = 0.1$  and  $x = 1.0$  respectively. The different cases considered are,

- i) Unconstrained case, that is the control is not constrained. The time interval is  $(0, .4)$  and the resulting state is shown in Figure 5.4. The resulting performance is 0.0105. The corresponding feedback coefficients are,

$$k_1 = 55.16$$

$$k_2 = 54.65$$

and the corresponding control is

$$\begin{aligned} u_c(t) = & k_1(q_d(0.1,t) - q(0.1,t)) \\ & + k_2(q_d(1.0,t) - q(1.0,t)) \end{aligned}$$

- ii) Constrained case I, that is the control is constrained to be within some prescribed limits. The time interval of interest is  $(0,0.4)$  and the resulting state is shown in Figure 5.5. The form of the control is,

$$\begin{aligned} u_c(t) = & k_1(q_d(0.1,t) - q(0.1,t)) \\ & + k_2(q_d(1.0,t) - q(1.0,t)) \end{aligned}$$

and the performance obtained is 0.011. The corresponding feedback coefficients are

$$k_1 = 54.89$$

$$k_2 = 64.03$$

The control is constrained to be

$$0 \leq u_c(t) \leq 10$$

- iii) Constrained Case II. The form of the feedback control is as follows.

$$u_c(t) = k_1(q_d(0.1,t) - q(0.1,t)) \\ + k_2(q_d(1.0,t) - q(1.0,t))$$

and the time interval of interest is (0,0.4) and hence  $T = 0.4$ . The control is constrained as

$$0 \leq u_c(t) \leq 1$$

The performance obtained is 0.011, and the corresponding feedback coefficients are,

$$k_1 = 59.34$$

$$k_2 = 52.47$$

The resulting state is shown in Figure 5.5.

## 5.5 Sensitivity Considerations

In the above, the feedback coefficients are assumed to be constants. In general the coefficients are time varying. So the performance index is sensitive with respect to the initial conditions and final time for constant gains. This is illustrated by the following example. A method of obtaining these sensitivity coefficients is given, and the extension of the method to the general case is straightforward.

### 5.5.1 Sensitivity coefficients

Let us consider a one dimensional diffusion equation,

$$\frac{\partial q}{\partial t} = \frac{\partial^2 q}{\partial x^2}, \quad \Omega = [0,1] \quad (5.35)$$

and the boundary conditions

$$\begin{aligned}\frac{\partial q}{\partial x}\bigg|_{x=0} &= \alpha\{q(0,t) - v(t)\} \\ \frac{\partial q}{\partial x}\bigg|_{x=1} &= 0 \\ q(x,0) &= 0\end{aligned}\tag{5.23}$$

and  $v(t)$  is given by

$$\begin{aligned}r \frac{dv}{dt} + v(t) &= u(t) \\ 0 \leq u(t) &\leq 1\end{aligned}\tag{5.24}$$

and it is required to minimize

$$I[u(t)] = \int_0^1 (q^*(x) - q(x,T))^2 dx\tag{5.25}$$

Discretizing (5.22) through (5.26) using the scheme explained in (5.14), the discretized system can be written as

$$X = A X + b u(t), \quad X(0) = C$$

Using (5.16)  $u_c(t) = u(t) = \langle K, X(t) \rangle$  then

$$X = A X + b \langle K, X \rangle\tag{5.35}$$

where  $\langle \cdot \cdot \rangle$  is the scalar product.

For given  $k_i$ , the performance index is sensitive to both the final time and the initial conditions. So,

$$I[u(t)] = \int_0^1 [q^*(x) - q(x,T)]^2 dx\tag{5.36}$$

or

$$\Delta J[u(t)] = \sum_{i=1}^n (q^*(x_i) - X_i(T))^2$$

Now (5.35) gives the solution

$$X(T) = e^{BT} C \quad (5.37)$$

where  $B = (A + bK^t)$ .

Substituting (5.37) into (5.36) and writing  $q(x_i, T) = X_i(T)$

$$\Delta J[C, T] = \sum_{i=1}^n (q^*(x_i) - q(x_i, T))^2$$

where  $\Delta J = \Delta J(C, T)$  to emphasize the dependence of  $\Delta J$  on  $C$  and  $T$ . Now the sensitive coefficients of  $\Delta J$  with respect to  $C$  and  $T$  are given by

$$\frac{\partial \Delta J}{\partial C_k} = 2 \sum_{i=1}^n (q^*(x_i) - q(x_i, T)) \frac{\partial q}{\partial C_k} \quad \begin{matrix} k = 1, \dots, n \\ i = 1, \dots, n \end{matrix}$$

$$\frac{\partial \Delta J}{\partial T} = 2 \sum_{i=1}^n (q^*(x_i) - q(x_i, T)) \frac{\partial q_i(x_i, T)}{\partial T} \quad \begin{matrix} k = 1, 2, \dots, n \\ i = 1, 2, \dots, n \end{matrix}$$

Thus writing

$$e_i = q^*(x_i) - q(x_i, T), \quad i = 1, 2, \dots, n$$

then

$$E = (e_1, e_2, \dots, e_n) \quad (5.38)$$

$$X = (q_1, q_2, \dots, q_n)$$

we have

$$\frac{\partial \Delta J}{\partial C_k} = 2 E' \frac{\partial X}{\partial C_k}, \quad \frac{\partial X}{\partial C_k} = \begin{bmatrix} \frac{\partial X_1}{\partial C_k} \\ \frac{\partial X_2}{\partial C_k} \\ \vdots \\ \frac{\partial X_n}{\partial C_k} \end{bmatrix} \quad (5.39)$$

$$\frac{\partial \Delta J}{\partial T} = 2 E' \frac{\partial X}{\partial T}, \quad \frac{\partial X}{\partial T} = \begin{bmatrix} \frac{\partial X_1}{\partial T} \\ \vdots \\ \frac{\partial X_n}{\partial T} \end{bmatrix} \quad (5.40)$$

Also from Eq. (5.37)

$$X(C, T) = e^{BT} \cdot C$$

Therefore

$$\frac{\partial X}{\partial T} = B \cdot e^{BT} \cdot C = B e^{BT} C = B X(T) \quad (5.41)$$

$$\frac{\partial X}{\partial C} = \begin{bmatrix} \frac{\partial X_1}{\partial C_1} & \cdots & \frac{\partial X_1}{\partial C_n} \\ \vdots & & \vdots \\ \frac{\partial X_n}{\partial C_1} & \cdots & \frac{\partial X_n}{\partial C_n} \end{bmatrix} = e^{BT} \quad (5.42)$$

So (5.39) and (5.40) become,

$$\frac{\partial \Delta J}{\partial C} = 2 E' \cdot e^{BT} \quad (5.43)$$

$$\frac{\partial \Delta J}{\partial T} = 2 E' \cdot B X(T) \quad (5.44)$$

and represent the sensitivity of the performance functional with respect to the initial conditions and the final time  $T$ .

### 5.5.2 Computational algorithm

The sensitivity coefficients given in (5.43) and (5.44) can be obtained by known  $e^{BT}$  and  $X(T)$ . But computation of

$e^{BT}$  by using the series

$$e^{BT} = (AT)^j / j!$$

or finding the constituent idempotents is quite cumbersome, due to the necessity of finding the eigenvalues. However, the following algorithm can be used. Now write (5.43),

$$\frac{\partial \Delta J}{\partial C} = \begin{bmatrix} \frac{\partial \Delta J}{\partial C_1} \\ \frac{\partial \Delta J}{\partial C_2} \\ \vdots \\ \frac{\partial \Delta J}{\partial C_n} \end{bmatrix} = 2 E' \begin{bmatrix} \frac{\partial X_1}{\partial C_1} & \frac{\partial X_1}{\partial C_2} & \cdots & \frac{\partial X_1}{\partial C_n} \\ \frac{\partial X_2}{\partial C_1} & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{\partial X_n}{\partial C_1} & \frac{\partial X_n}{\partial C_2} & & \frac{\partial X_n}{\partial C_n} \end{bmatrix}$$

where  $e^{BT} = \frac{\partial X}{\partial C}$

Now  $X(T) = e^{BT} \cdot C$

and  $C = (C_1, C_2, \dots, C_n)$  and hence we can write the following

$$X(T) = \begin{bmatrix} \frac{\partial X_1}{\partial C_1} & \frac{\partial X_1}{\partial C_2} & \cdots & \frac{\partial X_1}{\partial C_n} \\ \frac{\partial X_2}{\partial C_1} & \frac{\partial X_2}{\partial C_2} & \cdots & \frac{\partial X_2}{\partial C_n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \frac{\partial X_n}{\partial C_1} & \cdot & \cdots & \frac{\partial X_n}{\partial C_n} \end{bmatrix} \begin{bmatrix} C_1 \\ C_2 \\ \cdot \\ \cdot \\ \cdot \\ C_n \end{bmatrix}$$

Now sequentially setting the initial condition

$$c^1 = (c_1, 0, 0, \dots, 0)$$

$$c^2 = (0, c_2, 0, \dots, 0)$$

.....

$$c^n = (0, 0, \dots, c_n)$$

and integrating the system equations (5.35) for given values of  $k_1, k_2, \dots$  to get  $x^1(T), x^2(T), \dots$ , where

$$x^1(T) = \begin{bmatrix} \frac{\partial x_1}{\partial c_1} \\ \vdots \\ \frac{\partial x_n}{\partial c_1} \end{bmatrix}, \quad x^2(T) = \begin{bmatrix} \frac{\partial x_1}{\partial c_2} \\ \vdots \\ \frac{\partial x_n}{\partial c_2} \end{bmatrix} \quad \text{etc.}$$

Thus the sensitivity coefficients given in (5.43) are then

$$\frac{\partial \Delta J}{\partial c} = 2 E' (x^1(T), x^2(T), \dots, x^n(T))$$

Finally it can be seen that (5.43) and (5.44) for sensitivity coefficients depend on the error at the final time and hence conclude that if the state is reachable within the specified time  $T$ , the coefficients obtained for this  $T$  are approximately valid for all  $T$ . The approximation will be good because the error vector  $E$  will be close to zero.

## CHAPTER VI

### DEVELOPMENT OF ALGORITHM - II

#### 6.1 Introduction

In this chapter a description of Algorithm-II for obtaining an optimal feedback-control for a class of distributed systems is given. This differs from Algorithm-I where the feedback coefficients are assumed to be constant with respect to time in contrast to time varying gains obtained in Algorithm-II. In this algorithm, the optimal open-loop control is assumed to be available.

#### 6.2 Problem Formulation

Consider a distributed system characterized by a vector partial differential equation

$$\frac{\partial Q}{\partial t} = G[t, x, Q(x, t), Q_x(x, t), Q_{xx}(x, t), u(t), m(x, t)]$$

along with the boundary conditions

$$\begin{aligned} Q(x, t_0) &= Q_0(x) & x \in \Omega, t \in [0, t_f] \\ S_b Q(x_b, t) &= u(x_b, t) & t \in [0, t_f], x_b \in \Omega_b \end{aligned} \quad (6.1)$$

where  $u(t)$  is the boundary control,  $Q(x, t)$ , the vector of state functions and the performance functional to be minimized is

$$C \triangleq \int_{\Omega} [Q_d(x, t_f) - Q(x, t_f)]^T [Q_d(x, t_f) - Q(x, t_f)] d\Omega$$

with  $u(t)$  is constrained such that  $a \leq u(t) \leq b$ .

### 6.3 Algorithm - II

Loosely speaking, the algorithm involves measuring the state vector at any finite number of points in the spatial domain and then obtaining time varying feedback coefficients such that the feedback control so obtained is close to the optimal open loop control obtained in some given sense. Since the closed-loop control obtained results in a degradation of the optimum performance this control is called sub-optimal feedback control.

Let  $Q(\beta, t)$  be the state vector measures at  $x = \beta$ ,  $x \in \Omega \subseteq \mathbb{R}^n$ . Then let the feedback control be represented by

$$u_c(t) = F(Q(\beta, t), Q_d(\beta, t), K(t), t)$$

where  $u_c(t)$  is an  $r$ -dimensional vector,

$K(t)$  is a  $r \times n$  matrix with an off-diagonal term zero,

$Q(\beta, t)$  is the state vector of  $n \times 1$ .

For the following discussion, assume  $u_c(t)$  in the following form:

$$u_c(t) = K^t(t)[Q(\beta, t) - Q_d(\beta, t)] \quad (6.2)$$

Then the time varying coefficients can be obtained by minimizing one of the following

$$C \triangleq \int_{\Omega} [Q_d(x, t_f) - Q(x, t_f)]^t [Q_d(x, t_f) - Q(x, t_f)] d\Omega \quad (6.3)$$

$$I \triangleq \int_0^{t_f} [u^*(t) - u_c(t)] [u^*(t) - u_c(t)] dt \quad (6.4)$$

where  $Q_d(x, t_f)$  is the desired distribution and  $u^*(t)$  is the optimal open loop control. Here the second type of functional is used since the aim of this algorithm is to obtain the feedback controls utilizing the open-loop control.

Several methods exist for obtaining the open loop control for different classes of problems. Two methods called quadratic programming and direct search method are discussed in section 3.3 of Chapter III. Having obtained the open-loop control by one of the above methods, the method of obtaining feedback control is discussed.

First the system characterized by the partial differential equation is discretized in the spatial domain either by truncating the higher order terms of the corresponding integral equation using integral transform techniques or by a finite difference method. For parabolic systems s-domain approximation is very effective. The finite difference approximation is discussed in section 5.3. Now to determine the coefficients, we write

$$K_{jj}(t) = \sum_{i=1}^N \alpha_{ij} L(t_{i-1}, t_i), \quad j = 1, \dots, r \quad (6.5)$$

where

$$L(t_{i-1}, t_i) = \begin{cases} 1 & t_{i-1} \leq t \leq t_i \\ 0 & \text{otherwise} \end{cases} \quad (6.6)$$

The feedback coefficients are approximated by piecewise constant functions where  $N$  is the number of subdivisions of

the time interval  $(0, t_f)$ .

Hence

$$u_c^{(\ell)}(t) = k_{\ell\ell}(t) e^{(\ell)}(t) \quad \ell = 1, \dots, r$$

where

$$e^{(\ell)}(t) = [q^{(\ell)}(\beta, t) - q_d^{(\ell)}(\beta, t)] \quad (6.7)$$

Substituting (6.5) into (6.7),

$$u_c^{(\ell)}(t) = \begin{cases} \sum_{i=1}^N \alpha_{i\ell} e^{(\ell)}(t) & t_{i-1} \leq t \leq t_i \\ 0 & \text{otherwise} \end{cases} \quad (6.8)$$

for all  $i = 1, 2, \dots, N$ ;  $\ell = 1, \dots, r$ .

Substituting (6.8) into the given performance functional in (6.4)

$$I = \int_0^{t_f} [u^*(t) - u_c(t)]^T [u^*(t) - u_c(t)] dt \quad \text{is written as}$$

$$\begin{aligned} \tilde{I} &= \int_0^{t_f} \sum_{\ell=1}^r (u^{*(\ell)}(t) - \sum_{i=1}^N \alpha_{i\ell} L(t_{i-1}, t_i) e^{(\ell)}(t))^2 dt \\ &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \left\{ \sum_{\ell=1}^r (u^{*(\ell)}(t) - \alpha_{i\ell} e^{(\ell)}(t))^2 \right\} dt \\ &= \sum_{i=1}^N \int_{t_{i-1}}^{t_i} \left\{ \sum_{\ell=1}^r (u^{*(\ell)}(t) - \alpha_{i\ell} e^{(\ell)}(t))^2 \right\} dt \end{aligned}$$

Thus the problem is divided into  $N$ -independent subsystems

where the parameter optimization is performed and the number of parameters are equivalent to the number of state functions, at each stage. The initial condition for the  $i$ th stage is the final value of the state vector of the  $(i-1)$ th stage.

#### 6.4 Example

Consider a slab of material bounded by the planes  $x = 0$  and  $x = 1$  which is in contact with a heat transfer medium of temperature  $u(t)$  at  $x = 0$  and is perfectly insulated at  $x = 1$ . The dimensionless slab temperature  $q(x,t)$  is governed by the one-dimensional heat equation,

$$\frac{\partial q}{\partial t} = \frac{\partial^2 q}{\partial x^2} \quad (6.10)$$

$$q(x,0) = 0$$

$$\left. \frac{\partial q}{\partial x} \right|_{x=0} = \alpha (q(0,t) - u(t)) \quad (6.11)$$

$$\left. \frac{\partial q}{\partial x} \right|_{x=1} = 0$$

The optimal control problem consists of determining  $u(t)$ ,  $0 \leq t \leq t_f$ ,  $t_f$  specified, to minimize the integral average deviation of the temperature at  $t = t_f$ , from a desired distribution  $q_d(x)$ , namely, to minimize,

$$P = \int_0^1 [q_d(x) - q(x, t_f)]^2 dx \quad (6.12)$$

with constraint on  $u(t)$

$$0 \leq u(t) \leq 1 \quad (6.13)$$

Here  $q_d(x) = 0.2$ ,  $t_f = 0.2$ ,  $\alpha = 10$ .

The optimal open-loop control for the above problem is obtained by using direct search on the performance index (see section 3.3.2), and is shown in Figure 6.1. The feedback-control is

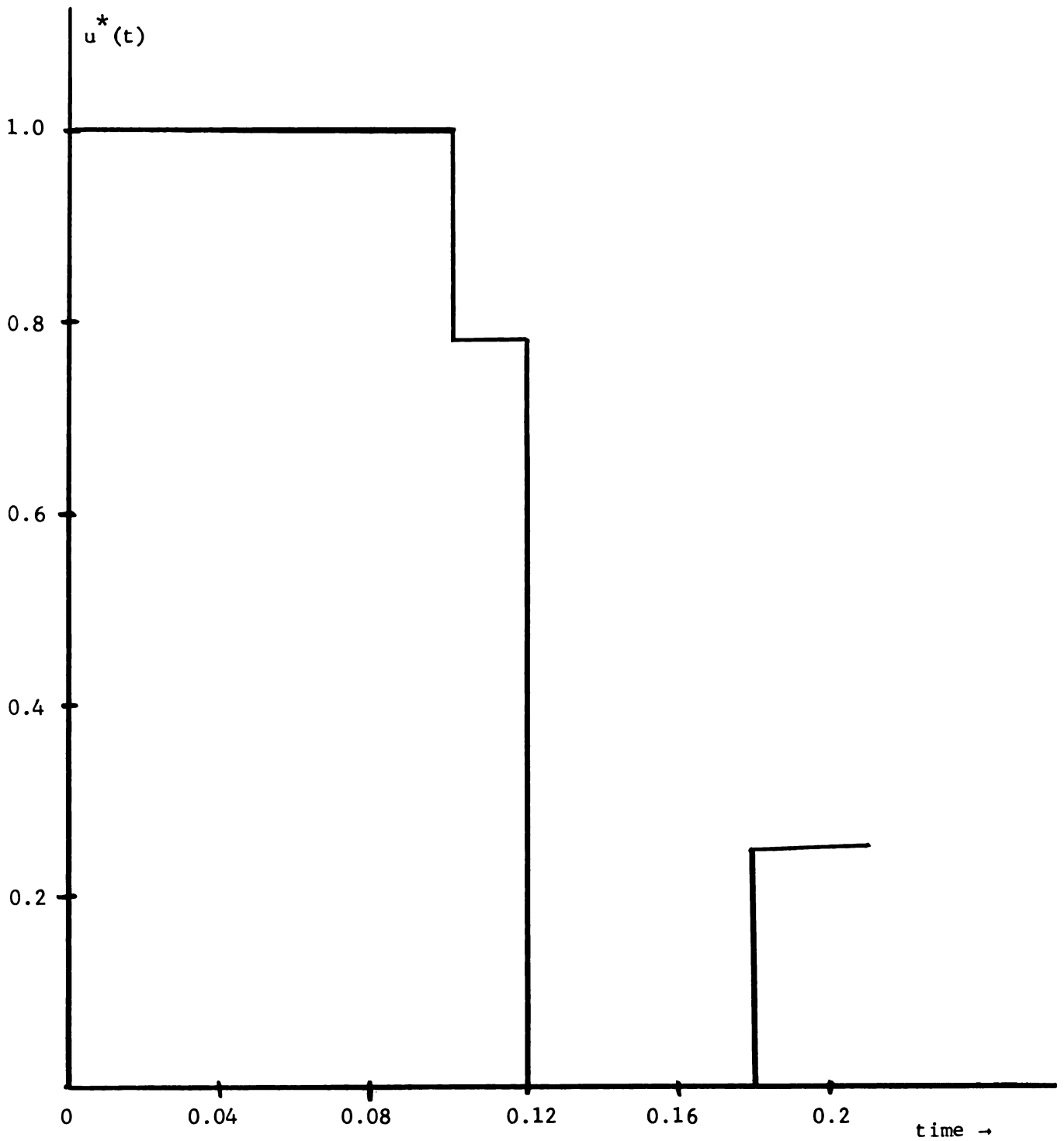


Figure 6.1 Optimal Open Loop Control by Direct Search.

$$0 \leq u(t) \leq 1$$

obtained as follows. Applying the Laplace transform to (6.10) subject to the initial condition yields,

$$\frac{\partial^2 Q}{\partial x^2} = s Q(x, s) \quad (6.14)$$

where  $Q(x, s) = \mathcal{L}(q(x, t))$ . Similar transformation of the boundary conditions yield,

$$\frac{\partial Q(x, s)}{\partial x} \Big|_{x=0} = \alpha \{Q(0, s) - u(s)\} \quad (6.15)$$

$$\frac{\partial Q(x, s)}{\partial x} \Big|_{x=1} = 0 \quad (6.16)$$

where  $u(s) = \mathcal{L}(u(t))$ .

The general solution of (6.14) is

$$Q(x, s) = C_1(s) \sinh(\sqrt{s} x) + C_2(s) \cosh(\sqrt{s} x) \quad (6.17)$$

where  $C_1(s)$  and  $C_2(s)$  are arbitrary functions of  $s$ . They are determined such that the general solution (6.17) satisfies the boundary conditions (6.15) and (6.16). Thus

$$Q(0, s) = C_2(s)$$

$$\frac{\partial Q}{\partial x} = C_1 \sqrt{s} \cosh \sqrt{s} x + C_2 \sqrt{s} \sinh \sqrt{s} x$$

Therefore

$$\frac{\partial Q}{\partial x} \Big|_{x=0} = C_1 \sqrt{s} = \alpha \{C_2 - u(s)\} \quad (6.18)$$

$$\frac{\partial Q}{\partial x} \Big|_{x=1} = C_1 \sqrt{s} \cosh \sqrt{s} + C_2 \sqrt{s} \sinh \sqrt{s} = 0 \Rightarrow$$

$$C_1 \sqrt{s} = -C_2 \sqrt{s} \tanh \sqrt{s} \quad (6.19)$$

Equations (6.18) and (6.19) yield

$$C_1(s) = \frac{-\alpha u(s) \sinh \sqrt{s}}{\sqrt{s} \sinh \sqrt{s} + \alpha \cosh \sqrt{s}} \quad (6.20)$$

$$C_2(s) = \frac{\alpha u(s) \cosh \sqrt{s}}{\sqrt{s} \sinh \sqrt{s} + \alpha \cosh \sqrt{s}} \quad (6.21)$$

Thus (6.17) gives,

$$Q(x,s) = u(s) \frac{\alpha \cosh(1-x)\sqrt{s}}{\sqrt{s} \sinh \sqrt{s} + \alpha \cosh \sqrt{s}} \quad (6.22)$$

or

$$\frac{Q(x,s)}{u(s)} = \frac{\cosh(1-x)\sqrt{s}}{\frac{\sqrt{s}}{\alpha} \sinh \sqrt{s} + \cosh \sqrt{s}} \quad (6.23)$$

Now before proceeding further, the following lemmas are proved.

Lemma 1. The equation

$$\cosh z + Bz \sinh z = 0 \quad (6.24)$$

has only imaginary roots and if  $z = x + iy$ , the roots are

the solutions of the equation

$$y \tan y = 1/B \quad (6.25)$$

Proof. Given

$$\cosh z + Bz \sinh z = 0 \quad (6.26)$$

implies  $\cosh z = -Bz \sinh z$  and now consider different cases.

Case 1. We know  $z = x + iy$  and  $x \neq 0$ ,  $y \neq 0$ , then (6.26)

can be written as

$$\coth z = -Bz$$

If  $z = x + iy$ , then

$$\coth (x + iy) = \frac{\sinh 2x - i \sin 2y}{\cosh 2x - \cos 2y} = -B(x + iy) \quad (6.27)$$

Equating the real and imaginary parts on both sides,

$$-\frac{1}{x} \frac{\sinh 2x}{\cosh 2x - \cos 2y} = B \quad (6.28)$$

$$\frac{1}{y} \frac{\sin 2y}{\cosh 2x - \cos 2y} = B \quad (6.29)$$

Equations (6.28) and (6.29) when equated yield,

$$-(1/x)(\sinh 2x) = (1/y)(\sin 2y)$$

i.e. letting  $p_1 = 2x$  and  $p_2 = 2y$ , we have

$$-(2 \sinh p_1)/p_1 = (2 \sin p_2)/p_2$$

Both the left hand side and right hand side functions of (6.30) are even functions of  $p_1$  and  $p_2$  respectively. They are plotted in Figure 6.2 (a) and Figure 6.2 (b). It is clear that there is no  $p_1$  and  $p_2$  to satisfy the above equation. Hence there are no roots with  $x = 0$  and  $y = 0$ .

Case 2.  $x = 0$  and  $y \neq 0$ , then (6.26) becomes

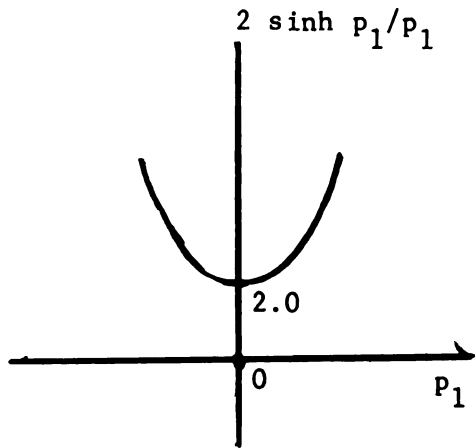
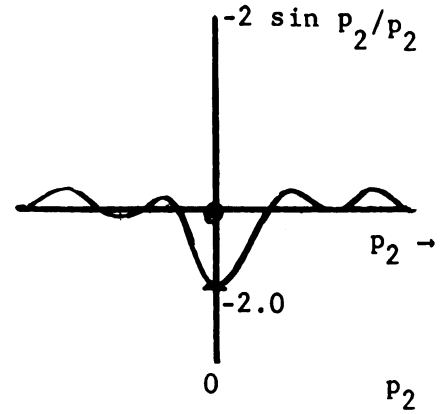
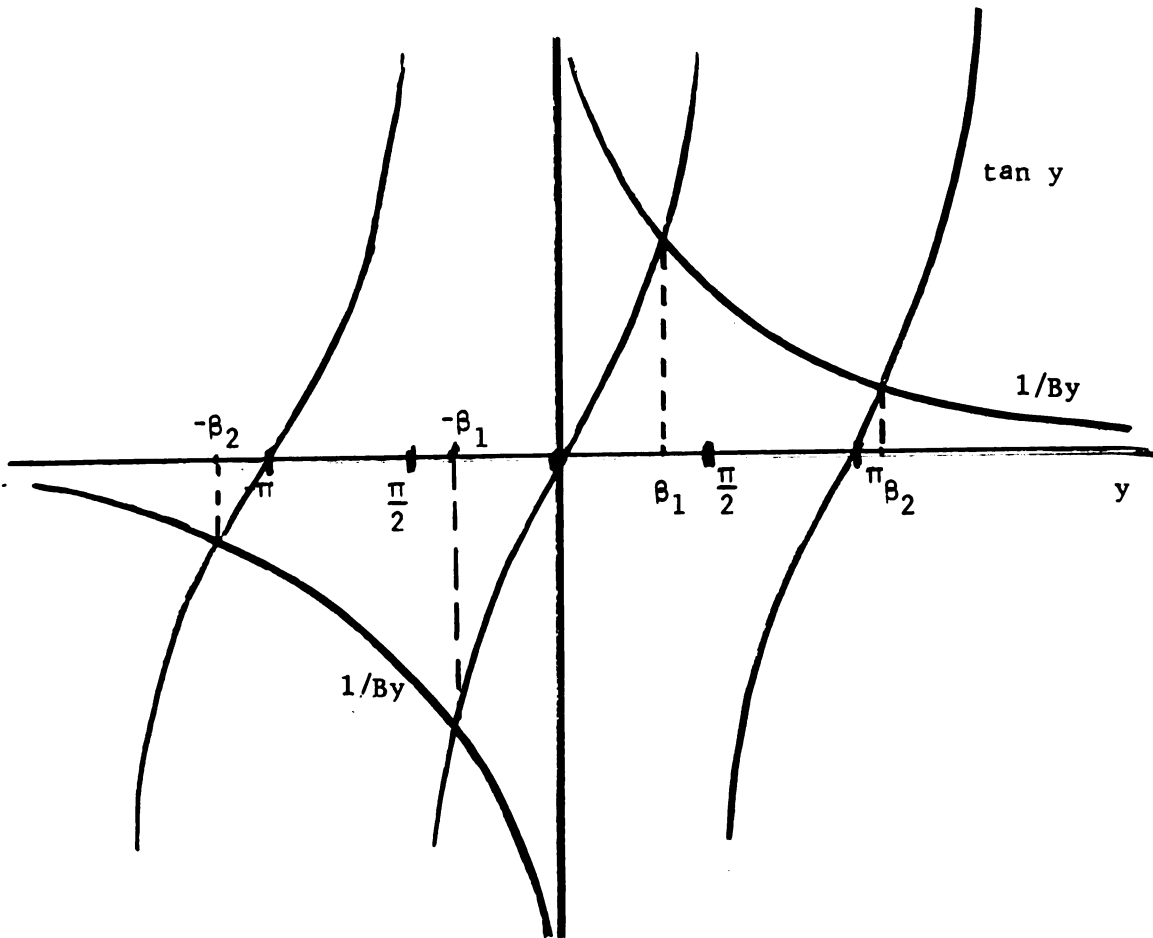
$$\cosh iy + i By \sinh iy = 0$$

$$\text{or} \quad \cosh iy + i By(i \sin y) = 0$$

$$\text{or} \quad \cos y = By \sin y$$

$$\text{or} \quad y \tan y = 1/B$$

which is exactly equation (6.25). The plot of  $\tan y = 1/By$  is shown in Figure 6.2 (c).

Figure 6.2a Plot of  $2 \sinh p_1/p_1$ Figure 6.2b Plot of  $-2 \sin p_2/p_2$ Figure 6.2c Plot of  $\tan y = 1/By$

Case 3.  $x \neq 0$ ,  $y = 0$ , then (6.26) becomes

$$\cosh x + Bx \sinh x = 0$$

or 
$$\cosh x = -Bx \sinh x$$

and since  $x \neq 0$ ,

$$\coth x = -Bx$$

For positive  $B$ , there are no intersection points and hence no roots, for  $x \neq 0$  and  $y = 0$ . (See Figure 6.2 (d)).

Thus the only roots of (6.24) are imaginary and they are the solutions of

$$y \tan y = 1/B$$

Q.E.D.

Lemma 2. The roots of  $\cosh Bz = 0$ ,  $B \neq 0$ , are completely imaginary and are given by

$$y_i = \pm(2i - 1)\frac{\pi}{2B}, \quad i = 1, 2, 3, \dots$$

Proof. Let  $z = x + iy$ , then

$\cosh Bz = 0$  can be expanded in the following way,

$$0 = \cosh Bz = \cosh Bx \cosh iBy + \sinh Bx \sinh iBy$$

or

$$\cosh Bx \cos By + i \sinh Bx \sin By = 0$$

Equating the real and imaginary parts, we have the real part,

$$\cosh Bx \cos By = 0 \quad \text{implies} \quad \cos By = 0.$$

Therefore 
$$y_i = \pm 1/B(2i - 1)\frac{\pi}{2} \quad i = 1, 2, 3, \dots \quad (6.31)$$

and the imaginary part,

$$\sinh Bx \sin By = 0 \quad \text{implies from (6.31) that}$$

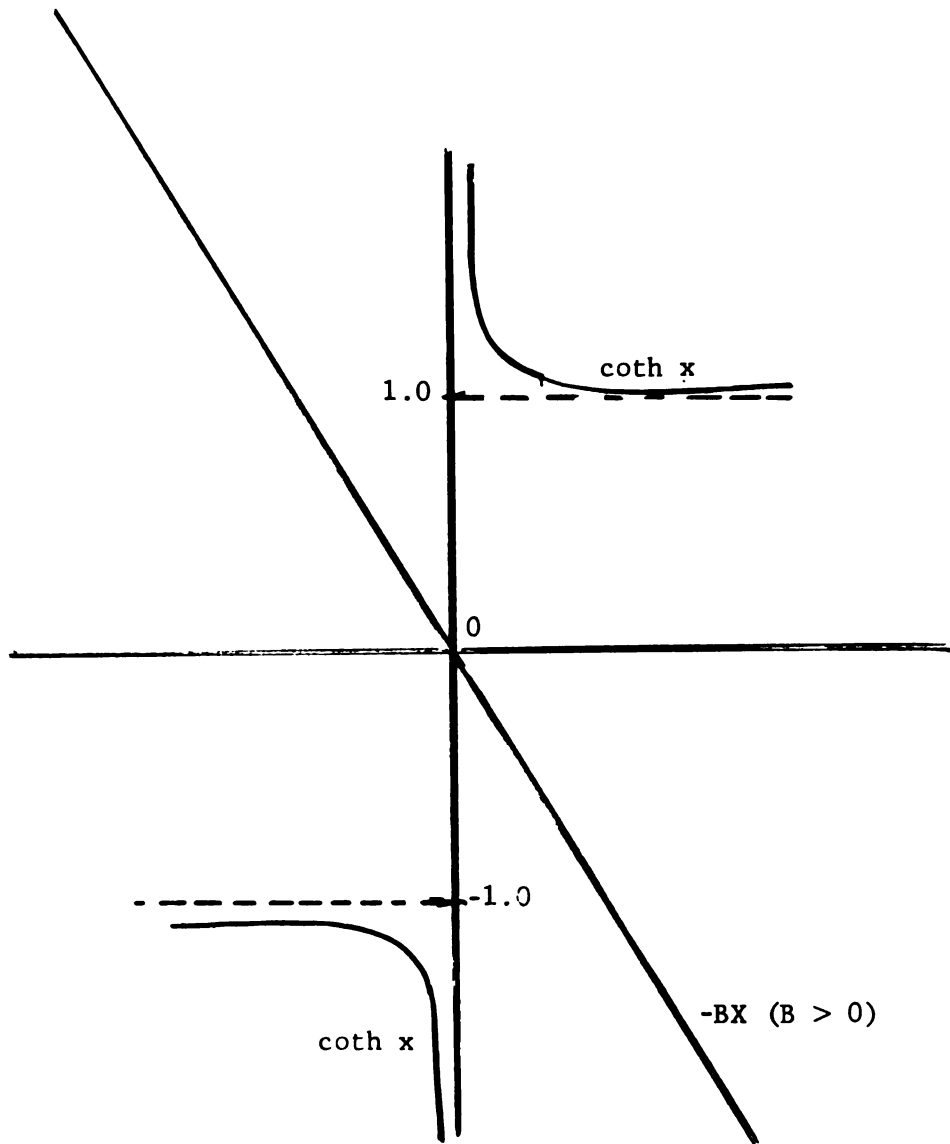


Figure 6.2d Plot of  $\coth x = -BX$  (for  $B > 0$ ).

$\sinh Bx = 0$  implies  $Bx = 0$  and hence  $x = 0$  since  $B \neq 0$  by the hypothesis. Q.E.D.

Now using the above lemmas the equation (6.23) can be written at  $x = \sigma$ , as

$$\frac{Q(\sigma, s)}{u(s)} = \frac{\prod_{i=1}^{\infty} (1 + s/y_i^2)}{\prod_{i=1}^{\infty} (1 + s/\beta_i^2)}, \quad 0 \leq \sigma \leq 1$$

where  $y_i = \pm \frac{1}{(1-\sigma)} (2i - 1)\pi/2$ ,  $i = 1, 2, \dots$

$\beta_i$  are the roots of  
 $\beta \tan \beta = \alpha.$

As can be seen from Figure 6.2 (c), the  $\beta_i$ 's increase very fast and since their squares occur in the denominator, the infinite product can be approximated by a finite product. Similarly by choosing the point  $\sigma$  as close as possible to the end point, the roots  $y_i$ 's can be made very large, and thus approximating the numerator by a finite product. Thus,

$$\frac{Q(\sigma, s)}{u(s)} = \frac{(1 + \frac{s}{y_1^2})}{(1 + \frac{s}{\beta_1^2})(1 + \frac{s}{\beta_2^2})} \quad (6.32)$$

is quite a good approximation, i.e.

$$(1 + \frac{s}{\beta_1^2})(1 + \frac{s}{\beta_2^2})Q(\sigma, s) = (1 + \frac{s}{y_1^2})u(s)$$

Transform this into time domain, and let  $v_1(t) = Q(\sigma, t)$

$$\frac{\ddot{v}_1}{\beta_1 \beta_2} + \frac{\dot{v}_1}{\beta_1 \beta_2} (\beta_1^2 + \beta_2^2) + v_1(t) = u(t) + \frac{1}{2} \frac{\dot{u}(t)}{y_1}$$

or

$$\ddot{v}_1 + (\beta_1^2 + \beta_2^2) \dot{v}_1 + \beta_1^2 \beta_2^2 v_1(t) = \beta_1^2 \beta_2^2 u(t) + \frac{\beta_1^2 \beta_2^2}{2} \frac{\dot{u}(t)}{y_1}$$

where dot denotes the differentiation with respect to time.

Using the following transformation

$$\begin{aligned} z_1 &= v_1(t) \\ z_2 &= \dot{v}_1 - \frac{\beta_1^2 \beta_2^2}{2} u(t), \end{aligned}$$

the equations given in (6.33) will be transformed into

$$\begin{bmatrix} \dot{z}_1 \\ \dot{z}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -a_0 & -a_1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(t)$$

where

$$\begin{aligned} a_0 &= \beta_1^2 \beta_2^2 \\ a_1 &= (\beta_1^2 + \beta_2^2) \\ b_1 &= \frac{\beta_1^2 \beta_2^2}{2} \\ b_2 &= \frac{\beta_1^2 \beta_2^2}{2} (\beta_1^2 + \beta_2^2) + \beta_1^2 \beta_2^2 \end{aligned}$$

Now assuming,

$$u(t) = u_c(t) = k(t)[Q(\sigma, t) - Q_d(\sigma, t)]$$

$$k(t) = \sum_{i=1}^N \alpha_i(t) L(t_{i-1}, t_i), \quad [t_{i-1}, t_i] \in [0, t_f]$$

where

$$L(t_{i-1}, t_i) = \begin{cases} 1 & t_{i-1} \leq t \leq t_i \\ 0 & \text{otherwise} \end{cases}$$

then

$$u_c(t) = \begin{cases} \sum_{i=1}^N \alpha_i e(t), & t_{i-1} \leq t \leq t_i \\ 0 & \text{otherwise} \end{cases}$$

Thus (6.9) becomes

$$\tilde{I} = \sum_{i=1}^N \int_{t_{i-1}}^{t_i} (u^*(t) - \alpha_i e(t))^2 dt.$$

Thus each  $\alpha_i$  can be obtained independent of the other  $\alpha$ .

Thus the feedback gain  $k(t)$  is approximated by a piecewise constant on sub-intervals and  $N$  can be chosen in an iterative manner until the desired performance measure is obtained.

The next section gives the computer set up and the results.

### 6.5 Computer Results

The sample value is taken at  $x = 1$ . Hence the system model becomes,

$$\begin{aligned} \begin{bmatrix} \dot{v}_1 \\ \dot{v}_2 \end{bmatrix} &= \begin{bmatrix} 0 & 1 \\ -\beta_1^2 \beta_2^2 & -(\beta_1^2 + \beta_2^2) \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} 0 \\ \beta_1^2 \beta_2^2 \end{bmatrix} u(t) \\ &= \begin{bmatrix} 0 & 1 \\ -52.98 & -21.3976 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 52.98 \end{bmatrix} u(t) \end{aligned}$$

where  $v(0) = 0$ .

The time interval of interest is  $(0, 0.2)$  and the analog computer set up is shown in Figure 6.4.

We start by dividing the interval  $(0, 0.2)$  into four sub-intervals as follows.

$\alpha_1$  belongs to  $t \in (0, 0.1)$

$\alpha_2$  belongs to  $t \in (0.1, 0.12)$

$\alpha_3$  belongs to  $t \in (0.12, 0.18)$

$\alpha_4$  belongs to  $t \in (0.18), 0.2)$

Then the algorithm follows by setting the initial conditions on the integrators 200, 201, and 241 to zero. Then the constant  $\alpha_1$  is obtained by one dimensional search by keeping the analog computer in the repetitive mode. Having obtained the constant the initial conditions are set up by using the digital computer which are the final conditions for the first stage. Then a one dimensional search yields the second constant  $\alpha_2$ . Similarly  $\alpha_3$ , etc. are obtained by following the above procedure. The constants are as follows.

$\alpha_1 = 5.76$  ,  $t \in (0, 0.1)$

$\alpha_2 = 13.57$  ,  $t \in (0.1, 0.12)$

$\alpha_3 = 0.0$  ,  $t \in (0.12, 0.18)$

$\alpha_4 = 51.49$  ,  $t \in (0.18, 0.2)$

and  $\tilde{I} = 0.0872$ .

Then in the next step the interval is divided into six subintervals. The subintervals and the constants are,

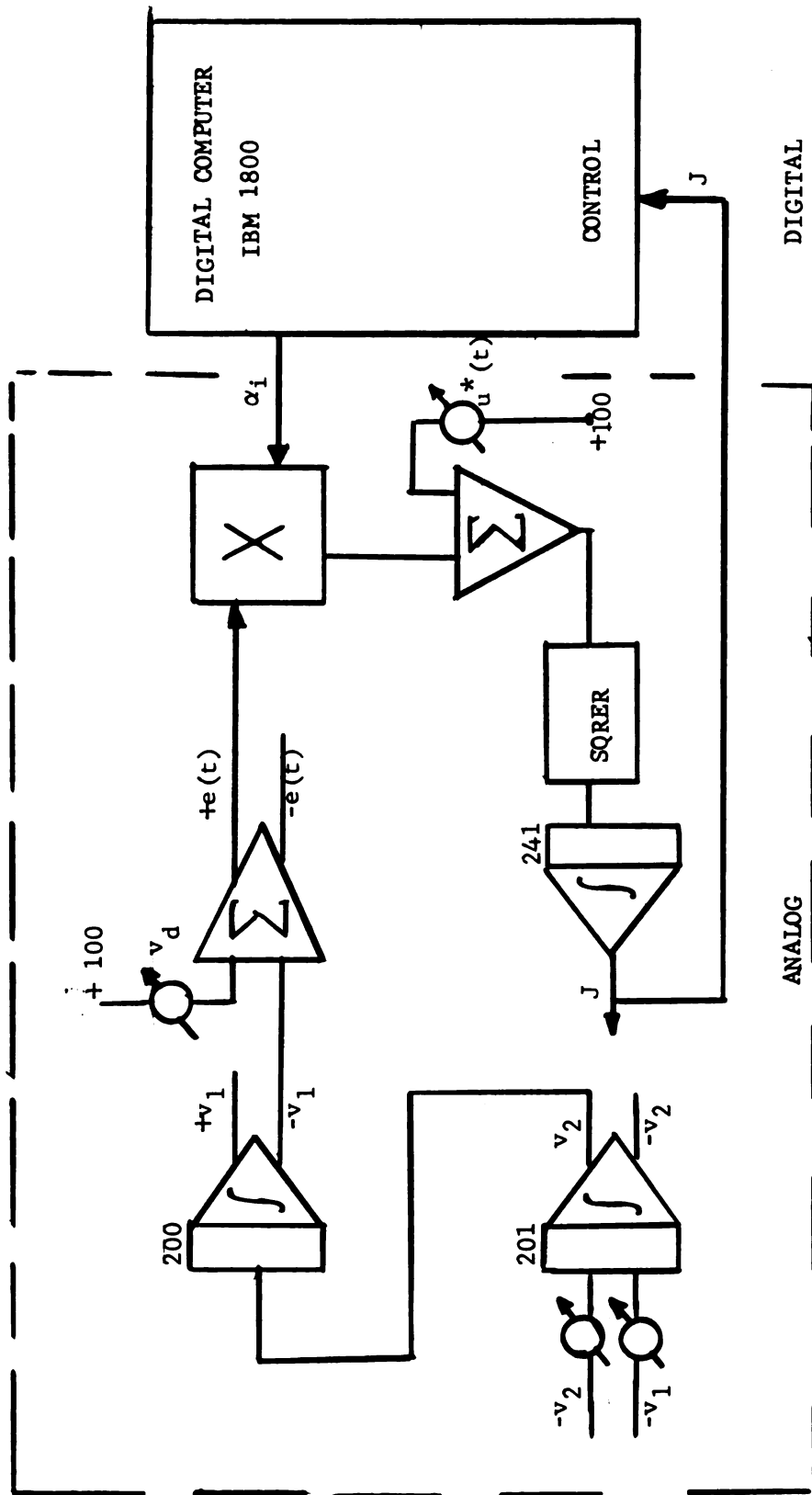


Figure 6.3 Hybrid Computer Set Up for the Example.

$$\alpha_1 = 5.52 \quad , \quad t \in (0, 0.05)$$

$$\alpha_2 = 7.09 \quad , \quad t \in (0.05, 0.07)$$

$$\alpha_3 = 10.38 \quad , \quad t \in (0.07, 0.1)$$

$$\alpha_4 = 16.52 \quad , \quad t \in (0.1, 0.12)$$

$$\alpha_5 = 0.0 \quad , \quad t \in (0.12, 0.18)$$

$$\alpha_6 = -12.05 \quad , \quad t \in (0.18, 0.2)$$

and  $\tilde{I} = 0.0215$ .

Thus by increasing the number of intervals, the performance index  $I$  can be made smaller and smaller, till it satisfies the required performance.

## CHAPTER VII

### CONCLUSIONS

#### 7.1 Conclusions

In this thesis, computational methods have been developed to obtain the optimal feedback controls for a class of distributed parameter systems. The class of problems considered in this these are "well posed" (see Section 1.4.1) and possess the following properties.

a) Solutions to the system equations

$$\frac{\partial Q}{\partial t} = G[Q(x,t), m(x,t), u(t), x, t]$$

$$Q(x, t_0) = Q_0(x) \quad ; \quad x \in \Omega \subseteq R^n$$

$$S_b Q(x_b, t) = u(t) \quad ; \quad x_b \in \Omega_b$$

with the given boundary conditions exist.

b) The solutions are uniquely determined.

c) The solution depends continuously on the initial data. This says that small changes in the initial data will cause correspondingly small changes in the solution,  $Q(x,t)$ .

The algorithms for obtaining the sub-optimal feedback control are discussed in Chapters V and VI. This is called sub-optimal feedback control because the application of these feedback control laws result in a degradation of optimum performance. In both the algorithms, the problem is reduced to a parameter

optimization problem with differential constraints. The a priori information available about the optimal open loop control is used in the second method to obtain the time varying feedback gains. The methods are easily implemented on the hybrid computer. The hybrid system available in the Hybrid Simulation and Control Laboratory (IBM 1800-AD-4) was used to obtain solutions for the examples in the thesis.

The capability of obtaining a solution for a differential system on an analog computer is limited by the number of integrators available on a given facility. A decomposition principle, which decomposes a large set of differential system equations into lower order independent subsystems which are solved iteratively is described in Chapter IV. The convergence theorems are stated and proved. With this treatment, a larger system (a finer spatial discretization) can be considered which would not be feasible otherwise.

Thus a significant contribution is made in this thesis in the area of distributed parameter systems by developing some efficient computer algorithms for obtaining feedback-controls and solving some of the problems encountered in the actual implementation on the computer.

## 7.2 Possible Extensions

In this thesis linear systems were considered, but the methods can be extended to non-linear systems. These non-linear problems must be well-posed. The verification of these conditions for non-linear systems are very difficult. It may be possible

to apply linearization techniques about a nominal trajectory in applying the above methods to some non-linear systems. These yield approximate results. Another possible extension is to find a class of problems where the state can be approximated by small order polynomial fit so that the results available in the lumped case could be applied.

This thesis emphasizes the fact that the results obtained in the case of lumped parameter systems cannot be applied directly for distributed parameter systems and thus new results obtained in this thesis are necessary. In these lines the thesis can be extended by changing the performance index such that the number and location of the measuring instruments along the spatial domain are optimized while penalizing the system for using large number of sensors. For solving these systems detailed comparison of the results, if possible, obtained by using approximate techniques are desirable. A listing of the best approximations for reducing several of the infinite dimensional systems which are common to finite dimensional systems will be very helpful.

## REFERENCES

## REFERENCES

- A-1 Axelband, Elliott, I., "An Approximation Technique for the Control of Linear Distributed Parameter Systems with Bounded Inputs", IEEE Transactions of Automatic Control, Vol. Ac-11, pp. 42-45, January 1966.
- A-2 Athans, M. and Falb, P.L., "Optimal Control", McGraw-Hill Book Co., New York, N.Y., 1966.
- B-1 Brogan, W.L., "Dynamic Programming and a Distributed Maximum Principle", Proc. JACC, 1967.
- B-2 Brogan, W.L., "Optimal Control Theory Applied to Systems Described by Partial Differential Equations", Advances in Control Systems, Vol. 6, 1968.
- B-3 Berg, P.W. and McGregor, J.L., "Elementary Partial Differential Equations", Holden-Day publications, San Francisco, 1966.
- B-4 Butkovskii, A.G., "Optimum Processes in Systems with Distributed Parameters", Automation and Remote Control, Vol. 21, pp. 13-21, 1961.
- B-5 Butkovskii, A.G. and Larner, A.Y., "The Optimum Control of Systems with Distributed Parameters", Automation and Remote Control, Vol. 21, pp. 472-477, 1960.
- B-6 Butkovskii, A.G., "The Maximum Principle for Optimum Systems with Distributed Parameters", Automation and Remote Control, Vol. 22, pp. 1156-1169, 1962.
- B-7 Butkovskii, A.G., "The Broadened Principle of the Maximum for Optimal Control Problems", Automation and Remote Control, Vol. 24, pp. 292-304, 1963.
- B-8 Butkovskii, A.G., "Some Approximate Methods for Solving Problems of Optimal Control of Distributed Parameter Systems", Automation and Remote Control, Vol. 22, pp. 1429-1438, 1961.
- B-9 Bellman, R., "Dynamic Programming", Princeton University Press, Princeton, New Jersey, 1957.

- B-10 Bekey, G.A., Karplus, W., "Hybrid Computation", McGraw Hill, 1969.
- D-1 Denn, M.M., "Optimal Boundary Control for a Non-Linear Distributed System", Int. J. Control, pp. 167, 1966.
- E-1 Egorov, L., "Optimal Control by Processes in Certain Systems with Distributed Parameters", Automation and Remote Control (English Translation), Vol. 25, p. 613, 1964.
- E-2 Egorov, I., "Optimal Processes in Systems Containing Distributed Parameters", Automation and Remote Control, p. 977, 1965.
- E-3 Eveleigh, V.W., "Adaptive Control and Optimization Techniques", McGraw-Hill Book Co., New York, N.Y., 1967.
- E-4 Ellsworth, W.C., "Hybrid Computer Solution of Linear State Models", Ph.D. Thesis, Department of Electrical Engineering, Michigan State University, East Lansing, Michigan, 1969.
- F-1 Fernando, L. Alvarado, R. Mukundan, "An Optimization Problem in Distributed Parameter Systems", Int. J. Control, Vol. 9, p. 665, 1969.
- G-1 Gelfand, I.M. and Fomin, S.V., "Calculus of Variations", Prentice Hall Incorporated, Englewood Cliffs, N.J., 1963.
- K-1 Khatri and Goodson, "Optimal Control of Systems with Distributed Parameters", JACC, p. 390, 1965.
- K-2 Katz, S., "A General Minimum Principle for End Point Control Problems", J. Electronics and Control, p. 189, 1964.
- K-3 Kim and Erzberger, "On the Design of Optimum Distributed Parameter Systems with Boundary Control Functions", IEEE Transactions on Automatic Control, p. 22, 1967.
- K-4 Kim and Gajwani, "A Variational Approach to Optimum Distributed Systems", IEEE Transactions on Automatic Control, p. 191, 1968.
- K-5 Koivo, A.J. and Kruh, P., "On the Design of Approximately Optimal Feedback Controllers for a Distributed Parameter System", Int. J. Control, Vol. 10, p. 53, 1969.
- K-6 Kreindler, B. and Athans, M., "Optimal Control with Piece-wise Constant Gains", IEEE Transactions on Automatic Control, August 1968.

- K-7 Kalman, R.E., "The Theory of Optimal Control and Calculus of Variations", RIAS Rept. 61-3, Res. Inst. for Advan. Studied, Baltimore, Maryland, 1961.
- L-1 Leitmann, "Topics in Optimization", Academic Press.
- L-2 Lapidus, L. and Luus, R., "Optimal Control of Engineering Processes", Blaisdell Publishing Company, 1967.
- M-1 McCausland, I.J., "On Optimum Control of Temperature Distribution in a Solid", J. Electronics and Control, Vol. 14, No. 6, pp. 655-68, 1963.
- R-1 Russell, D.L., "Optimal Regulation of Linear Symmetric Hyperbolic Systems with Finite Dimensional Controls", SIAM J. on Control, Vol. 4, p. 276, 1966.
- S-1 Sakawa, Y., "Solution of an Optimum Control Problem in Distributed Parameter Systems", IEEE Transactions on Automatic Control, Vol. AC-9, p. 420, 1964.
- S-2 Sakawa, Y., "Optimal Control of Certain Types of Linear Distributed Parameter Systems", IEEE Transactions on Automatic Control, Vol. AC-11, p. 35, 1966.
- S-3 Sage, A.P. and Chaudhuri, S.P., "Discretization Schemes and the Optimum Control of Distributed Parameter Systems", Proc. Asimilar Conference on Circuits and Systems, Monterey, California, p. 191, 1967.
- S-4 Sage, A.P. and Chaudhuri, S.P., "Gradient and Quasi-linearization Computational Techniques for Distributed Systems", Int. J. on Control, p. 81, 1967.
- S-5 Sage, A.P., "Optimum Systems Control", Prentice Hall Inc., New Jersey, 1968.
- S-6 Seinfeld, J.H. and Kumar, K.S.P., "Synthesis of Sub-Optimal Feedback Controls for a Class of Distributed Parameter Systems", Int. J. Control, Vol. 1, p. 417, 1968.
- S-7 Seinfeld, J.H. and Lapidus, L., "Computational Aspects of the Optimal Control of Distributed Parameter Systems", Chemical Engineering Science, Vol. 23, p. 1461, 1968.
- S-8 Seinfeld, J.H. and Lapidus, L., "Singular Solutions in the Optimal Control of Lumped and Distributed Parameter Systems", Chemical Engineering Science, Vol. 23, p. 1485, 1968.
- V-1 Varga, R., "Matrix Iterative Analysis", Prentice Hall Inc., Englewood Cliffs, New Jersey, 1962.

- W-1 Wang, P.K.C. and Tung, F., "Optimum Control of Distributed Parameter Systems", J. Basic Engineering, Trans. of ASME, Vol. 86D, p. 67, 1964.
- W-2 Wang, P.K.C., "Control of Distributed Parameter Systems", Advances in Control, Vol. 1, 1963.
- W-3 Wismer, D.A., "An Efficient Computational Procedure for the Optimization of a Class of Distributed Parameter Systems", J. Basic Engineering, Trans. of ASME, p. 190, June 1969.
- W-4 Wendroff, B., "Theoretical Numerical Analysis", Academic Press, 1966.
- W-5 Wilde, D.L. and Beightler, C.S., "Foundations of Optimization", Prentice Hall Inc., New Jersey, 1967

MICHIGAN STATE UNIVERSITY LIBRARIES



3 1293 03085 5534