ON THE COMPUTATIONAL ALGORITHMS FOR OPTIMAL CONTROL PROBLEMS WITH GENERAL CONSTRAINTS

Keiichi Kaji

A thesis submitted to the Faculty of Science, University of the Witwatersrand, Johannesburg, in fulfilment of the requirements for the degree of Doctor of Philosophy

Johannesburg 1992

DECLARATION

I hereby declare that this thesis is entirely my own work. It is being submitted for the degree of Doctor of Philosophy in the University of the Witwatersrand, Johannesburg. It has not been submitted before for any degree or examination in any other University.

Keiichi Kaji

Keiichi Kaji

25th day of February, 1992.

To my sister Takae Kaji.

ACKNOWLEDGEMENTS

The author expresses his deep gratitude to his supervisor, Dr. K. H. Wong for initiating the problem and for his encouragement, understanding, sincere and continued guidance throughout this work without which this thesis would not have been possible.

He also thanks the members of the Department of Computational and Applied Mathematics at the University of the Witwatersrand, South Africa for their friendship and assistance.

Thanks also due to Mrs Tammy Main for typing the thesis.

Last but by no means least, he thanks his sister and parents for their encouragement, support and understanding throughout the period of his study in South Africa.

ABSTRACT

In this thesis, we study the following four types of optimal control problems:

(i) Problems governed by systems of ordinary differential equations;

(ii) Problems governed by systems of ordinary differential equations with time–delayed arguments appearing in both the state and the control variables;

(iii) Problems governed by linear systems subject to sudden jumps in parameter values;

(iv) A chemical reactor problem governed by a couple of nonlinear diffusion equations.

The aim of this thesis is to devise computational algorithms for solving the optimal control problems under consideration. However, our main emphasis are on the mathematical theory underlying the techniques, the convergence properties of the algorithms and the efficiency of the algorithms.

Chapters II and III deal with problems of the first type, Chapters IV and V deal with problems of the second type, and Chapters VI and VII deal with problems of the third and fourth type respectively. A few numerical problems have been included in each of these Chapters to demonstrate the efficiency of the algorithms involved.

The class of optimal control problems considered in Chapter II consists of a nonlinear system, a nonlinear cost functional, initial equality constraints, and terminal equality constraints. A Sequential Gradient–Restoration Algorithm is used to devise an iterative algorithm for solving this class of problems. The convergence properties of the algorithm are investigated.

The class of optimal control problems considered in Chapter III consists of a nonlinear system, a nonlinear cost functional, and terminal as well as interior points equality constraints. The technique of control parameterization and Liapunov concepts are used to solve this class of problems.

A computational algorithm for solving a class of optimal control problems involving terminal and continuous state constraints of inequality type was developed by Ref. 103 in 1989. In Chapter IV, we extend the results of Ref. 103 to a more general class of constrained time–delayed optimal control problems, which involves terminal state equality constraints, as well as terminal state inequality constraints and continuous state inequality constraints.

In Ref. 104, a computational scheme using the technique of control parameterization was developed for solving a class of optimal control problems in which the cost functional includes the full variation of control. Chapter V is a straightforward extension of Ref. 104 to the time–delayed case. However the main contribution of this chapter is that many numerical examples have been solved.

In Chapter VI, a class of linear systems subject to sudden jumps in parameter values is considered. To solve this class of stochastic control problem, we try to seek for the best feedback control law depending only on the measurable output. Based on this idea, we convert the original problem into an approximate constrained deterministic optimization problem, which can be easily solved by any existing nonlinear programming technique.

In Chapter VII, a chemical reactor problem and its control to achieve a desired output temperature is considered. A finite element Galerkin method is used to convert the original distributed optimal control problem into a quadratic programming problem with linear constraints, which can be solved by any standard quadratic programming software.

CONTENTS

Page

LIST OF TABLES AND FIGURES

CHAPTER I

INTRODUCTION

Optimal control theory has been of considerable importance and has many successful applications in engineering, natural science and social science. The theory has been developed and formalized as a generalized extension of the calculus of variations. It has been used to solve a diverse collection of problems. However, real–world problems are often too complex to be solved analytically. Therefore, computational algorithms are unavoidable in solving these real–world problems.

Our main concern in this thesis is to develop efficient algorithms for solving the following four types of optimal control problems:

(i) Problems governed by systems of ordinary differential equations;

(ii) Problems governed by systems of ordinary differential equations with time–delayed arguments appearing in both the state and the control variables;

(iii) Problems governed by linear systems subject to sudden jumps in parameter values;

(iv) A chemical reactor problem governed by a couple of nonlinear diffusion equations.

1.1 Ordinary Differential Equations

There are many efficient gradient type algorithms in the literature for solving optimal control problems governed by ordinary differential equations. As early as in 1968, it was used in conjunction with the method of penalty functions in Refs. 21, 44. In 1970, the gradient information together with the approach of Bryson–Denham was used in Ref. 16.

However, both of these methods did not ensure the satisfaction of all the constraints and the boundary conditions at the end of each iteration. Thus, the comparison of the cost functional values between each iteration was not possible. To avoid this drawback, Kelley in Ref. 35 proposed to have a restoration phase at the end of the gradient phase. He superimposed to the change of the control associated with the gradient phase a perturbation obtained by combining linearly some arbitrarily prescribed functions $f_1(t)$, $f_2(t), \cdots$. The constants of the combination were determined iteratively in order to satisfy the terminal conditions. Unfortunately, there is no clear-cut way to choose the functions $f_1(t), f_2(t), \cdots$.

In the late 1960's, the Sequential Gradient-Restoration Algorithm (SGRA) was developed for solving equality constrained nonlinear optimization problems by Miele and his co-workers in Refs. 58-59. The algorithm consists of a sequence of two major phases: a gradient-type minimization in a subspace tangent to the constraint surface, and a feasibility restoration procedure. It has been used with great success for many large-scale problems, mainly in the area of engineering optimization.

However, convergence properties of the SGRA for nonlinear programming problems were studied only recently by Rom and Avriel in Ref. 76-77, where global convergence was successfully proven, and an asymptotic rate of convergence was established.

In Ref. 60, Miele et al extended the use of SGRA to optimal control problems. Nevertheless, no convergence properties of the SGRA were included.

In Chapter II we develop the SGRA for solving a class of optimal control problems governed by ordinary differential equations, which involves both initial and terminal equality constraints. This class of optimal control problems is more general than that of

Ref. 60, because the state variable x is fixed at the time $t = 0$ in Ref. 60. The problem is summarized as follows:

Minimize a function J which depends on the state variable $x(t)$, the control variable $u(t)$, and the parameter π, where J is a scalar, x is an n–vector, u an m–vector and π a p–vector, and t is the normalized time (cf. Ref. 43) with $0 \leq t \leq 1$. $x(t)$, $u(t)$ and π are required to satisfy n scalar equality differential equations. Furthermore, $x(t)$ and π are required to satisfy the initial condition (r scalar equality equations) as well as the final conditions (s scalar equality equations).

The SGRA developed in Chapter II is very similar to the SGRA developed in Refs. 76–77 for solving nonlinear programming. Moreover, the proof of convergence result of the SGRA for solving optimal control problems is established for the first time in history. The proof of this convergence result is strongly motivated by that of Ref. 65. More precisely, we have extended the convergence result for the mathematical programming problems reported in Ref. 65 to optimal control problems.

An important feature of our algorithm is that a scheme is incorporated in such a way that at points far from optimum, only a rough estimate of feasibility is required; and the feasibility requirements are gradually tightened as the optimal solution is approached. This scheme can save a great deal of computational time.

In Chapter III, we present a new innovative approach for solving a class of optimal control problems governed by ordinary differential equations which involves both terminal and interior points constraints. The problem is summarized as follows:

Minimize a functional J which depends on the state variable $x(t)$, the control variable $u(t)$ and the parameter π, where J is a scalar, x is a n–vector, u is a r–vector and π is a p–vector, t is the normalized time with $0 \leq t \leq 1$. $x(t)$, $u(t)$ and π are required to satisfy q terminal or interior points constraints.

Using the technique of control parameterization described in Refs. 15, 28, 32, 86, 88, 89, 95, 97, 102, 103, 108, 110, 112, 125, 126, the optimal control problem can be approximated by a constrained nonlinear programming problem. There are many existing techniques for solving this problem, such as the constrained Quasi–Newton method described in the Appendix of Ref. 112. We shall however, resort to another rather unconventional, somewhat brute force manner for solving this problem namely a method based on the Liapunov Concepts which was originally suggested in Ref. 74.

After having converted the original control problem into a constrained nonlinear problem, we convert this constrained problem into an unconstrained problem via Liapunov form. Based on the fact that the stationary points of the Lagrangian function can be found by setting all the derivations equal to zero, we define a suitable positive–definite Liapunov Function (L) with terms consisting of all the derivatives of the Lagrangian function. Finally, the problem of finding the minimum of (L) can be easily solved by any existing unconstrained nonlinear programming technique, such as the conjugate gradient method.

From the illustrative examples solved in Section 3.7, the method of control parameterization in conjunction with Liapunov Concept appears to work very efficiently for the class of problems described in Chapter III.

1.2 Ordinary Delay–Differential Equations

In recent years, significant emphasis has been given to time–delayed differential equations, because of many applications in engineering, natural science and social science. Excellent bibliographies for such practical problems can be found in Refs. 18, 120

In Ref. 35, Kharatispvili extended Pontryagin's Maximum principle to systems with discrete delays in both the state and the control variables.

Results concerning the existence of optimal controls for systems governed by ordinary delay–differential equations have been investigated by Ahmed (Refs. 3–4), Angell (Refs. 6–8), Chyung and Lee (Ref. 19), Georganas and Ahmed (Ref. 24), Lee (Refs. 37–38), Nababan and Teo (Ref. 67), Oguztöreli (Ref. 69,pp. 181–192), Teo, Wu and Clements (Ref. 116), Tsuzakan (Ref. 117) and Wang (Ref. 118).

It appears that very few results (Refs. 66, 96, 116, and 122–126) on computational algorithms for hereditary optimal control problems are available in the open literature.

In both Chapter IV and V of this thesis, we consider optimal control problems governed by the following delay–differential equation

$$\frac{dx}{dt} = f(t, x(t), x(t-h), u(t), u(t-h)), \quad t \in [0,1] \tag{1.1}$$

together with some given initial conditions.

In Chapter IV, a computational scheme using the technique of control parameterization and a new constraint transcription is presented for solving the nonlinear time-delayed optimal control problem governed by (1.1), together with terminal state equality constraints, terminal state inequality constraints, and continuous state constraints. Moreover, the cost functional consists of a terminal cost and an integral cost. The traditional method of constraint transcription is to convert the continuous state constraints of the form

$$g(t,x(t)) \leq 0, \qquad t \in [0,T] \tag{1.2}$$

into the form

$$\int_0^T \max\{0,\ g(t,x(t))\}\ dt = 0 \tag{1.3}$$

However, as mentioned in Ref. 103, the equality constraint (1.3) so obtained is non-smooth and do not satisfy the usual constraint qualification. Thus, convergence is not guaranteed and some oscillation may exist in the numerical computation.

Thus, in Chapter IV, we use the idea of Ref. 103 to obtain a new constraint transcription. By using this new constraint transcription, we can overcome the two disadvantages mentioned above.

The above technique has already been adopted by Teo and Jenning in Ref. 103 for solving a class of optimal control problems involving terminal state inequality constraints and continuous state constraints. In Ref. 110, Teo and Wong extend all the results of Ref. 103 to a more general class of optimal control problems involving additional terminal equality constraints, together with the constraints reported in Ref. 103. The contribution of Chapter IV of this thesis is to extend the result of Ref. 110 to a general class of time–delayed control problems.

Although the convergence result of this paper is exactly the same as that of Ref. 110, the assumptions that we use in obtaining these convergence result are much less restrictive than those of Ref. 110. In fact, only one additional assumption apart from those given in Ref. 103 has been used.

The problem considered in Chapter V is very similar to that considered in Chapter IV, with the exception that the cost functional in Chapter V is the sum of not only the terminal cost and the integral cost, but also the full variation of the control. Moreover, for the sake of simplicity, we have omitted equality constraints in our discussion in Chapter V.

In Ref. 14, Blatt initiated the idea of including a cost associated with switching control from one discrete value to another into the usual cost functional in an optimal control problem.

In Ref. 68, Noussair extended this concept to the case where the controls are piecewise continuous functions defined on $[0,T]$ with values in some compact subset of R^n. The full variation of a control is defined as the sum of variations of all its components.

In Ref. 50, Matula derived the necessary condition in an integral form and a local necessary condition in the case of monotonic controls.

The optimal control problem considered in Ref. 104 was very similar to that considered in Chapter V of this thesis, with the exception that the system considered in Ref. 104 had no time–delayed arguments in both the state variables and the control variables. There, by using the control parameterization technique presented in Refs. 15, 32, 86, 88, 89 and the enforced smoothing technique presented in Ref. 100, the optimal control problem can be approximated by a sequence of standard optimal parameter selection problems, which can be solved by any standard mathematical programming technique, such as the standard constrained Quasi–Newton method. It is interesting to note that by using the control parameterization technique, the term involving full variation of control appearing in the cost functional simply becomes a non–smooth function in the form of ℓ_1–norm. This non–smooth function is then smoothed by the enforced smoothing technique mentioned above. Rigorous convergence results have been obtained in Ref. 104.

The main contribution of Chapter V is that we extend all the convergence results of Ref. 104 to the time–delayed system. Although this extension to time–delayed system is quite straightforward, we have solved a few numerical examples in Chapter V of this thesis to illustrate that the technique used in Ref. 104 works equally well for time–delayed problems.

1.3 Linear Systems Subject to Sudden Jumps in Parameter Values

In Refs. 22, 37, Florentin, Krasovskii and Lidskii initiated the study of optimal control for a class of linear systems with random parameters. Some of their research focused on the case of jump linear quadratic (JLQ) systems where the parameters are Markov jump processes with finitely countable states and where an optimal control law is sought with respect to the mathematical expectation of a quadratic cost. In Refs. 91, 128, 129, Sworder and Wonham established that, in this case, the optimal control law consisted of a full state feedback regulator. In Ref. 40, Levine and Athans considered the case without jumps and proposed the optimal control law for the LQ problem. In Ref. 53, McLane did the similar work for the Linear Quadratic Gaussian (LQG) problem which takes account of the continuously acting disturbances modelled by Gaussian white noise. In Ref. 48, Mariton and Bertrand established the best feedback control law depending only on measurable output rather than on the state of the system for the JLQ problem. This result was further improved when Mariton designed JLQG regulator in Ref. 47 where only part of the system state, the output, is measured.

In Chapter VI, we deal with exactly the same problem as that considered in Ref 48. By seeking for the best feedback control law depending only on the measurable output, we convert the original problem into an equivalent standard constrained deterministic optimization problem, which can be easily solved by any nonlinear programming technique. This method is better than those considered in Refs. 40, 48 which are based on the necessary condition of optimality, because their algorithms could not guarantee that the system would be stable in every mode during each iteration of the algorithm. Same as in Ref. 9, the gradients of both the performance index and the constraints were derived via the introduction of the Hamiltonian and co–state matrices. Finally, an example has been solved to illustrate the efficiency of our method.

1.4 Nonlinear Diffusion Equation for the Chemical Reaction

In Chapter VII, a chemical reactor problem and its control to achieve a desired output temperature is considered. The variables to be controlled are the temperature and the concentration of oxygen, which are functions of both position and time, and are described by a couple of nonlinear diffusion equations. The available controls are input temperature and input oxygen concentration. The objective function to be minimized is the mean square error in the actual output temperature compared to a desired output temperature. By linearizing the differential equations around a nominal solution and then applying a finite element Galerkin Scheme to the resulting distributed system, the original problem can be converted into a sequence of quadratic programming problems with linear constraints which can be easily solved by any standard quadratic programming software. The approach of first converting the original distributed optimal control problem into a sequence of approximate problem involving only lumped parameter systems and then developing computational algorithm for solving the discretized problem has been widely used in Refs. 11, 73, 115 (Chapter VI, Section 7).

CHAPTER II

CONVERGENCE PROPERTIES OF THE SEQUENTIAL GRADIENT–RESTORATION ALGORITHM FOR A CLASS OF OPTIMAL CONTROL PROBLEMS INVOLVING INITIAL AND TERMINAL EQUALITY CONSTRAINTS

2.1 Introduction

The class of optimal control problems considered in this chapter, which is more general than that of Ref. 60 , involves a nonlinear dynamical system together with terminal equality constraints and initial equality constraints.

Our main aim is to use the Sequential Gradient–Restoration Algorithm (SGRA) for solving the optimal control problem iteratively and to investigate the convergence properties of the algorithm.

In Section 2.2, we introduce certain notations, which will be used throughout this thesis.

In Section 2.3 and Section 2.4, we describe the optimal control problem $(2.P)$, and impose certain assumptions.

In Section 2.5, we derive first–order necessary conditions for optimality together with a constraint qualification for the problem.

In Section 2.6, we describe an approximate method for solving the problem $(2.P)$. This approximate method involves minimizing the cumulative errors in both the constraints and the optimum conditions.

In Section 2.7, the restoration phase is discussed. The restoration phase is designed to restore feasibility. More precisely, in the restoration phase, we use nominal functions $\tilde{x}(t)$, $\tilde{u}(t)$, $\tilde{\pi}$ to generate functions $\tilde{x}(t) + \Delta\tilde{x}(t)$, $\tilde{u}(t) + \Delta\tilde{u}(t)$, $\tilde{\pi} + \Delta\tilde{\pi}$ so that the new functions so generated will satisfy both the differential equations and the boundary conditions.

In Section 2.8, the gradient phase is first discussed. The gradient phase is designed to decrease the functional in a subspace tangent to the constraint surface. More precisely, in the gradient phase, nominal functions $x(t)$, $u(t)$, π satisfying the differential equations and boundary conditions are assumed. We then try to seek functions $x(t) + \Delta x(t)$, $u(t) + \Delta x(t)$, $\pi + \Delta\pi$ so that the value of the functional is decreased.

These variations $\Delta x(t)$, $\Delta u(t)$, $\Delta\pi$ are obtained by minimizing the first–order change of the functional subject to the linearized differential equations, the linearized boundary conditions and a quadratic constraint on the variations of the control and the parameter.

After having discussed the gradient phase, we present the SGRA in Section 2.8. Each iteration of the SGRA consists of a sequence of gradient–type minimization and feasibility restoration. The convergence properties of the SGRA are discussed in Section 2.8.

In Section 2.9, one numerical example is solved to demonstrate the efficiency of the SGRA.

2.2 Notation

Let $|\cdot|$ denote the usual norm in an Euclidean space. For any matrix $M \equiv (M_{ij})_{\substack{i=1,\dots,n_1, \\ j=1,\dots,n_2}}$

let

$$|M| = \Big[\sum_{i=1,\ldots,n_1} \sum_{j=1,\ldots,n_2} (M_{ij})^2 \Big]^{\frac{1}{2}}$$

for any $G : R^{n_1} \longrightarrow R^{n_2}$, let

$$\frac{\partial G}{\partial y} = \begin{bmatrix} \partial G_1/\partial y_1 & \cdots & \partial G_1/\partial y_{n_1} \\ & \vdots & \\ \partial G_{n_2}/\partial y_1 & \cdots & \partial G_{n_2}/\partial y_{n_1} \end{bmatrix}$$

2.3 Statement of the problem

The purpose of this chapter is to study the minimization of the functional

$$J = \int_0^1 f(t,x(t),u(t),\pi)\,dt + [g(x,\pi)]_0 + [h(x,\pi)]_1 \tag{2.1}$$

with respect to the functions $x(t)$, $u(t)$ together with the parameter π which satisfy the differential equation

$$\dot{x}(t) = \phi(t,x(t),u(t),\pi) \tag{2.2}$$

together with the initial condition

$$[\omega(x,\pi)]_0 = 0 \tag{2.3}$$

and the final condition

$$[\psi(x,\pi)]_1 = 0 \qquad (2.4)$$

where $x : [0,1] \to R^n$ is an absolutely continuous function such that $\dot{x}(t)$ is essentially bounded on [0,1]; $u : [0,1] \to R^m$ is a bounded measurable function; π is a vector of R^p; and $\phi : [0,1] \times R^n \times R^m \times R^p \to R^n$, $f : [0,1] \times R^n \times R^m \times R^p \to R$, $g : R^n \times R^p \to R$, $h : R^n \times R^p \to R$, $\omega : R^n \times R^p \to R^r$, $\psi : R^n \times R^p \to R^s$ are given functions.

For convenience, the above problem is referred to as Problem (2.P).

2.4 Basic Definitions and Assumptions

Let L_∞^n denote the Banach space $L_\infty([0,1], R^n)$ of all essentially bounded measurable functions from [0,1] into R^n. Its norm is defined by

$$\| x \|_\infty \equiv \operatorname*{ess\,sup}_{t \in [0,1]} |x(t)| \qquad (2.5)$$

Let $\mathscr{U}$ be the set defined by

$$\mathscr{U} = \{z = (x,u,\pi) : x : [0,1] \to R^n \text{ is an absolutely continuous function such that } \dot{x} \in L_\infty^n ;\ u \in L_\infty^m ;\ \pi \in R^p\} \qquad (2.6)$$

For any $z \in \mathscr{U}$, let

$$\| z \| = \| x \|_\infty + \| \dot{x} \|_\infty + \| u \|_\infty + |\pi| + |x|_0 + |x|_1 \qquad (2.7)$$

Let $\mathscr{F}$ be a subset of $\mathscr{U}$ defined by

$$\mathscr{F} = \{z = (x,u,\pi) \in \mathscr{U} : (x,u,\pi) \text{ satisfies constraints (2.2), (2.3) and (2.4)}\}$$

We assume that the following conditions are satisfied.

Assumption (2.4A)

2.4.A1. The functions f and ϕ are continuous on $[0,1]$ for each $(x,u,\pi) \in R^n \times R^m \times R^p$.

2.4.A2. There exists positive constants K_1 and K_2 such that

$$|\hat{F}_\beta(t,x^1,u^1,\pi^1)| \le K_1, \tag{2.8}$$

$$|\hat{G}_\beta(x^1,\pi^1)| \le K_1, \tag{2.9}$$

and

$$|\hat{F}_\beta(t,x^1,u^1,\pi^1) - \hat{F}_\beta(t,x^2,u^2,\pi^2)| \le K_2\,(|x^1-x^2| + |u^1-u^2| + |\pi^1-\pi^2|), \tag{2.10}$$

$$|\hat{G}_\beta(x^1,\pi^1) - \hat{G}_\beta(x^2,\pi^2)| \le K_2\,(|x^1-x^2| + |\pi^1-\pi^2|), \tag{2.11}$$

for any $t \in [0,1]$, $x^1, x^2 \in R^n$, $u^1, u^2 \in R^m$, $\pi^1, \pi^2 \in R^p$, where $\hat{F}$ (respectively, $\hat{G}$) denotes any of the functions f and ϕ (respectively, g, h, ω and ψ); and $\hat{F}_\beta$ (respectively, $\hat{G}_\beta$) denotes the gradient of the function $\hat{F}$ (respectively, $\hat{G}$) with respect to β. Here, β represents any of the vectors x, u and π.

2.4.A3. The functions $\partial^2 \hat{F}/\partial\beta_1\partial\beta_2$ and $\partial^2 \hat{G}/\partial\beta_1\partial\beta_2$ are continuous, where for each $i = 1,2$, β_i represents any of the variables x_i, i = 1,...,n; u_i, i = 1,...,m, and π_i, i = 1,...,p.

2.5 First–Order Necessary Conditions and Constraint Qualification

In this section, we shall derive first-order necessary conditions for optimality together with a constraint qualification for the problem (2.P). The main function of the constraint qualification in the mathematical programming is to ensure that the algorithm does not jam up at undesirable points.

For brevity, we write

$$\hat{F} = \hat{F}(t,x,u,\pi) \tag{2.12}$$

$$\hat{G} = \hat{G}(x,\pi) \tag{2.13}$$

$$\hat{F}_\beta = \hat{F}_\beta(t,x,u,\pi) \tag{2.14}$$

and

$$\hat{G}_\beta = \hat{G}_\beta(x,\pi) \tag{2.15}$$

where $\hat{F}$, $\hat{G}$, $\hat{F}_\beta$, and $\hat{G}_\beta$ are as defined before.

Remark 2.5.1 From the standard existence and uniqueness theorem in the theory of ordinary differential equations, it follows that the trajectory $x(t)$ that satisfies (2.2) is uniquely determined once $(x(0), u(t), \pi)$ is specified.

Let $z = (x(t), u(t), \pi) \in \mathcal{F}$ be given. By adjoining the differential equation (2.2) to J with the Lagrange multiplier function $\lambda^0(t)$, we get

$$J = [g]_0 + [h]_1 + \int_0^1 \{f + (\lambda^0)^T (\dot{x} - \phi)\}\, dt \tag{2.16}$$

Integrating by parts the last term on the right hand side of (2.16), we get

$$J = [g]_0 + [h]_1 + \int_0^1 [f - (\lambda^0)^T \phi - (\dot{\lambda}^0)^T x]\, dt$$

$$+ [(\lambda^0)^T x]_1 - [(\lambda^0)^T x]_0 \tag{2.17}$$

In view of Remark 2.5.1, z is completely determined by $(x(0), u(t), \pi)$. Thus, by considering the variation in J due to the variation in $(x(0), u(t), \pi)$, we obtain

$$\Delta J = [g_x - (\lambda^0)^T]_0 [\Delta x]_0 + [h_x + (\lambda^0)^T]_1 [\Delta x]_1$$

$$+ \int_0^1 \{f_x - (\lambda^0)^T \phi_x - (\dot{\lambda}^0)^T\} \Delta x \, dt + \int_0^1 \{f_u - (\lambda^0)^T \phi_u\} \Delta u \, dt$$

$$+ \int_0^1 \{f_\pi - (\lambda^0)^T \phi_\pi\} \Delta \pi \, dt + \{[g_\pi]_0 + [h_\pi]_1\} \Delta \pi \tag{2.18}$$

Choose λ^0 to satisfy the following system of differential equations:

$$(\dot{\lambda}^0)^T = f_x - (\lambda^0)^T \phi_x \tag{2.19a}$$

$$[\lambda^0]_1^T = -[h_x]_1 \tag{2.19b}$$

Then, ΔJ simplifies to

$$\Delta J = [g_x - (\lambda^0)^T]_0 [\Delta x]_0 + \int_0^1 [f_u - (\lambda^0)^T \phi_u] \Delta u \, dt$$

$$+ \{\int_0^1 [f_\pi - (\lambda^0)^T \phi_\pi] \, dt\} \Delta \pi + \{[g_\pi]_0 + [h_\pi]_1\} \Delta \pi \tag{2.20}$$

Consider the constraint functions:

$$[\omega_i]_0, \qquad i = 1,\ldots,r \tag{2.21}$$

where $[\omega_i]_0$ denotes the i^{th} component of $[\omega]_0$. In this case, we have

$$[\Delta\omega_i]_0 = [\omega_{i,x} \Delta x + \omega_{i,\pi} \Delta \pi]_0, \qquad i = 1,\ldots,r. \tag{2.22}$$

We now consider the constraint functions:

$$[\psi_i]_1, \qquad i = 1,\ldots,s \tag{2.23}$$

where $[\psi_i]_1$ denotes the i^{th} component of $[\psi]_1$. Then, by using a similar argument as that given for ΔJ, we can show that

$$[\Delta\psi_i]_1 = -([\lambda^i]_0)^T [\Delta x]_0 - \int_0^1 (\lambda^i)^T \phi_u \Delta u \, dt$$

$$+ \{[\psi_{i,\pi}]_1 - \int_0^1 (\lambda^i)^T \phi_\pi \, dt\} \Delta\pi, \qquad i = 1,\ldots,s \tag{2.24}$$

where

$$(\dot{\lambda}^i)^T = -(\lambda^i)^T \phi_x \tag{2.25a}$$

$$(\lambda^i)_1^T = -[\psi_{i,x}]_1 \tag{2.25b}$$

Let us now construct $(\Delta x\ (0),\ \Delta u\ (t),\ \Delta\pi)$ such that $\Delta J < 0$ and that the following constraints are satisfied.

$$[\Delta\omega_i]_0 = 0, \qquad i = 1,\ldots,r \tag{2.26}$$

and

$$[\Delta\psi_i]_1 = 0, \qquad i = 1,\ldots,s \tag{2.27}$$

Multiply the i^{th} equation in (2.22) by μ_i and the i^{th} equation in (2.24) by $\hat{\mu}_i$. Sum them together. Then, by adding the resulting equation to (2.20), we obtain

$$\Delta J + \sum_{i=1}^{r} \mu_i [\Delta\omega_i]_0 + \sum_{i=1}^{s} \hat{\mu}_i [\Delta\psi_i]_1 = \int_0^1 K^u \Delta u \, dt + K^\pi \Delta\pi + K^x [\Delta x]_0 \tag{2.28}$$

where

$$K^{u}(t) = f_u - (\lambda^0)^T \phi_u - \sum_{i=1}^{s} \hat{\mu}_i (\lambda^i)^T \phi_u \tag{2.29a}$$

$$K^{\pi} = \int_0^1 \{f_\pi - (\lambda^0)^T \phi_\pi - \sum_{i=1}^{s} \hat{\mu}_i (\lambda^i)^T \phi_\pi\}\, dt$$

$$+ [g_\pi + \sum_{i=1}^{r} \mu_i\, \omega_{i,\pi}]_0 + [h_\pi + \sum_{i=1}^{s} \hat{\mu}_i\, \psi_{i,\pi}]_1 \tag{2.29b}$$

and

$$K^{x} = [g_x - (\lambda^0)^T + \sum_{i=1}^{r} \mu_i\, \omega_{i,x} - \sum_{i=1}^{s} \hat{\mu}_i (\lambda^i)^T]_0 \tag{2.29c}$$

Now choose

$$\Delta u\,(t) = -\kappa\,(K^{u}(t))^T \tag{2.30}$$

$$\Delta \pi = -\kappa\,(K^{\pi})^T \tag{2.31}$$

$$(\Delta x)_0 = -\kappa\,(K^{x})^T \tag{2.32}$$

where $\kappa > 0$. Then,

$$\Delta J + \sum_{i=1}^{r} \mu_i\,[\Delta\omega_i]_0 + \sum_{i=1}^{s} \hat{\mu}_i\,[\Delta\psi_i]_1$$

$$= -\kappa\,(\int_0^1 |K^{u}(t)|^2\, dt + |K^{\pi}|^2 + |K^{x}|^2) \tag{2.33}$$

which is negative unless $K^u(t) \equiv 0 \; \forall \; t \in [0,1]$, $K^\pi = 0$, and $K^x = 0$.

Next, we need to determine μ_i, i = 1,...,r, and $\hat{\mu}_i$, i = 1,...,s, so that the constraints (2.26) and (2.27) are satisfied. In view of (2.22), (2.24), (2.30)–(2.32) and (2.29), the matrix equation for μ and $\hat{\mu}$ takes the form

$$\begin{bmatrix} M^1 & M^2 \\ M^3 & M^4 \end{bmatrix} \begin{bmatrix} \mu \\ \hat{\mu} \end{bmatrix} = \begin{bmatrix} W^1 \\ W^2 \end{bmatrix} \tag{2.34}$$

where $M^1 \in R^{r\times r}$, $M^2 \in R^{r\times s}$, $M^3 \in R^{s\times r}$, $M^4 \in R^{s\times s}$, $W^1 \in R^r$, $W^2 \in R^s$, and

$$M^1_{i,j} = -[\omega_{i,x}(\omega_{j,x})^T]_0 - [\omega_{i,\pi}(\omega_{j,\pi})^T]_0 \tag{2.35a}$$

$$M^2_{i,j} = [\omega_{i,x}\lambda^j]_0 - [\omega_{i,\pi}]_0 \{([\psi_{j,\pi}]_1)^T - \int_0^1 (\phi_\pi)^T \lambda^j \, dt\} \tag{2.35b}$$

$$M^3_{i,j} = ([\lambda^i]_0)^T ([\omega_{j,x}]_0)^T - \{[\psi_{i,\pi}]_1 - \int_0^1 (\lambda^i)^T \phi_\pi \, dt\} ([\omega_{j,\pi}]_0)^T \tag{2.35c}$$

$$\begin{aligned} M^4_{i,j} = &-([\lambda^i]_0)^T [\lambda^j]_0 + \int_0^1 (\lambda^i)^T \phi_u (\phi_u)^T \lambda^j \, dt \\ &+ \{[\psi_{i,\pi}]_1 - \int_0^1 (\lambda^i)^T \phi_\pi \, dt\} \{\int_0^1 (\phi_\pi)^T \lambda^j \, dt - ([\psi_{j,\pi}]_1)^T\} \end{aligned} \tag{2.35d}$$

$$\begin{aligned} W^1_i = &[\omega_{i,x}]_0 [(g_x)^T - \lambda^0]_0 + [\omega_{i,\pi}]_0 [\int_0^1 \{(f_\pi)^T - (\phi_\pi)^T \lambda^0\} \, dt \\ &+ ([g_\pi]_0)^T + ([h_\pi]_1)^T] \end{aligned} \tag{2.36a}$$

$$W_i^2 = -([\lambda^i]_0)^T [(g_x)^T - \lambda^0]_0 - \int_0^1 (\lambda^i)^T \phi_u [(f_u)^T - (\phi_u)^T \lambda^0]\, dt$$
$$+ \{[\psi_{i,\pi}]_1 - \int_0^1 (\lambda^i)^T \phi_\pi\, dt\} \{\int_0^1 \{(f_\pi)^T - (\phi_\pi)^T \lambda^0\}\, dt$$
$$+ ([g_\pi]_0)^T + ([h_\pi]_1)^T] \quad (2.36b)$$

Let

$$M(z) = \begin{bmatrix} M^1(z) & M^2(z) \\ M^3(z) & M^4(z) \end{bmatrix} \quad (2.37)$$

Then, for each $z \in \mathscr{U}$, $\mu(z)$ and $\hat{\mu}(z)$ exist if and only if the matrix $M(z)$ is nonsingular. In other words, if $(M(z))^{-1}$ does not exist, it is not possible to control $(\Delta x(0), \Delta u(t), \Delta \pi)$ to satisfy one or more of the constraints (2.26) and (2.27).

Definition 2.5.1 Any $z \in \mathscr{U}$ (not necessarily feasible) is said to satisfy the constraint qualification of the problem (2.P) if $(M(z))^{-1}$ exists.

Note that we have allowed non–feasible z to be included in the definition of constraint qualification. This is done for the convenience of future developmen.

At this stage, we have constructed $(\Delta x(0), \Delta u(t), \Delta \pi)$ that decreases J and satisfies the constraints (2.22) and (2.24). From (2.33), the only case in which we cannot decrease J is when

$$K^u = 0 \ \forall\ t \in [0,1],\ K^\pi = 0, \qquad K^x = 0 \quad (2.38)$$

Now, let

$$\lambda(t) = \lambda^{0}(t) + \sum_{i=1}^{s} \hat{\mu}_i \, \lambda^{i}(t) \tag{2.39}$$

Then, by (2.19) and (2.25), we have

$$(\dot{\lambda}(t))^{T} = f_x - \lambda^{T} \phi_x \tag{2.40a}$$

$$[\lambda^{T} + h_x + (\hat{\mu})^{T} \psi_x]_1 = 0 \tag{2.40b}$$

Next, by combining (2.29) and (2.39), the conditions (2.38) become

$$f_u - \lambda^{T} \phi_u = 0 \tag{2.41a}$$

$$\int_0^1 \{f_\pi - \lambda^{T} \phi_\pi\} \, dt + [g_\pi + \mu^{T} \omega_\pi]_0 + [h_\pi + \hat{\mu}^{T} \psi_\pi]_1 = 0 \tag{2.41b}$$

$$[-\lambda^{T} + g_x + \mu^{T} \omega_x]_0 = 0 \tag{2.41c}$$

Thus, under the constraint qualification condition, we seek a $z \in \mathscr{F}$ together with $\lambda(t)$, μ, and $\hat{\mu}$, such that (2.40) and (2.41) are satisfied.

Note that the technique used in this section is motivated by that of Ref. 17.

2.6 Approximate Method

For each $z = (x,u,\pi) \in \mathscr{U}$, $\lambda(t) \in R^{n}$, $\mu \in R^{r}$, $\hat{\mu} \in R^{s}$, let P, $\hat{P}$ and Q be defined by

$$P(z) = \hat{P}(x,\dot{x},u,\pi)$$

$$= \int_0^1 |\dot{x} - \phi|^2 \, dt + |[\omega(x,\pi)]_0|^2 + |[\psi(x,\pi)]_1|^2 \tag{2.42}$$

and

$$Q(z,\lambda,\dot{\lambda},\mu,\hat{\mu})$$
$$= \int_0^1 \{ |\dot{\lambda}^T - f_x + \lambda^T \phi_x|^2 + |f_u - \lambda^T \phi_u|^2 \} \, dt$$
$$+ \left| \int_0^1 \{f_\pi - \lambda^T \phi_\pi\} \, dt + [g_\pi + \mu^T \omega_\pi]_0 + [h_\pi + \hat{\mu}^T \psi_\pi]_1 \right|^2$$
$$+ |[-\lambda^T + g_x + \mu^T \omega_x]_0|^2 + |[\lambda^T + h_x + \hat{\mu}^T \psi_x]_1|^2 \tag{2.43}$$

Here, P measures the cumulative errors in the constraints, while Q measures the cumulative errors in the optimum condition.

Remark 2.6.1 In view of Assumption (2.4.A2), we see that the function $P(z)$ is continuous in z with respect to the norm defined by (2.7). Thus, we conclude that for any $z \in \mathscr{U}$, $P(z) \neq \infty$.

Define

$$T_\epsilon = \left\{ z \in \mathscr{U} : \sqrt{P(z)} \leq \epsilon \right\} \tag{2.44}$$

and

$$\Delta_\epsilon = \{ z \in \mathscr{U} : \exists\, \lambda(t) \in R^n,\ \mu \in R^r \text{ and } \hat{\mu} \in R^s$$
$$\text{such that } \sqrt{Q(z,\lambda,\dot{\lambda},\mu,\hat{\mu})} \leq \epsilon \}, \tag{2.45}$$

where $\lambda : [0,1] \longrightarrow R^n$ is an absolutely continuous function.

The approximate method is to seek a z such that

$$z \in \mathscr{U} \cap T_{\epsilon_1} \cap \Delta_{\epsilon_2},$$

where ϵ_1 and ϵ_2 are small, preselected numbers.

2.7 Restoration Phase

For any nominal function $\tilde{z} \in (\tilde{x},\tilde{u},\tilde{\pi}) \in \mathscr{U} - T_0$, we seek a function:
$\tilde{z} + \Delta\tilde{z} = (\tilde{x} + \Delta\tilde{x}, \tilde{u} + \Delta\tilde{u}, \tilde{\pi} + \Delta\tilde{\pi}) \in \mathscr{U} \cap T_0$.

If quasilinearization is employed, equations (2.2) – (2.4) are approximated by

$$\Delta\dot{\tilde{x}} - \tilde{\phi}_x\Delta\tilde{x} - \tilde{\phi}_u\Delta\tilde{u} - \hat{\phi}_\pi\Delta\tilde{\pi} + (\dot{\tilde{x}} - \tilde{\phi}) = 0 \tag{2.46a}$$

$$[\tilde{\omega} + \tilde{\omega}_x\Delta\tilde{x} + \tilde{\omega}_\pi\Delta\tilde{\pi}]_0 = 0 \tag{2.46b}$$

$$[\tilde{\psi} + \tilde{\psi}_x\Delta\tilde{x} + \tilde{\psi}_\pi\Delta\tilde{\pi}]_1 = 0, \tag{2.46c}$$

where the tilde imposed on the functions ϕ, ω, ψ and their derivatives denotes evaluation of the functions at the beginning of the restoration phase. In order to prevent $\Delta\tilde{x}(t)$, $\Delta\tilde{u}(t)$, and $\Delta\tilde{\pi}$ from becoming too large, we imbed (2.46) in a one–parameter family:

$$\Delta\dot{\tilde{x}} - \tilde{\phi}_x\Delta\tilde{x} - \tilde{\phi}_u\Delta\tilde{u} - \tilde{\phi}_\pi\Delta\tilde{\pi} + \tilde{\alpha}(\dot{\tilde{x}} - \tilde{\phi}) = 0 \tag{2.47a}$$

$$[\tilde{\alpha}\tilde{\omega} + \tilde{\omega}_x\Delta\tilde{x} + \tilde{\omega}_\pi\Delta\tilde{\pi}]_0 = 0 \tag{2.47b}$$

$$[\tilde{\alpha}\tilde{\psi} + \tilde{\psi}_x\Delta\tilde{x} + \tilde{\psi}_\pi\Delta\tilde{\pi}]_1 = 0 \tag{2.47c}$$

where $\tilde{\alpha}$ is a constant. For each prescribed $\tilde{\alpha}$, we seek the minimum of the cost functional

$$\tilde{J} = \frac{1}{2\tilde{\alpha}}\left[\int_0^1 |\Delta\tilde{u}|^2\, dt + |\Delta\tilde{\pi}|^2\right] \tag{2.48}$$

subject to (2.47)

Let the above problem be referred to as $(2.\tilde{P})$.

Remark 2.7.1 Since system (2.46) (and hence system (2.47)) is obtained from equations (2.2) – (2.4) by quasilinearization, it is clear that the constraint qualification for the problem $(2.\tilde{P})$ is exactly the same as that for the problem $(2.P)$. Thus, by following exactly the same argument as that given in Section 2.5, we can show that if $\tilde{z}$ satisfy the constraint qualification of the problem $(2.P)$ (i.e., $(M(\tilde{z}))^{-1}$ exist), then there exist $\Delta\tilde{z} = (\Delta\tilde{x}, \Delta\tilde{u}, \Delta\tilde{\pi})$ and $\tilde{\lambda}(t) \in R^n$, $\tilde{\mu} \in R^r$, $\tilde{\hat{\mu}} \in R^s$ such that (2.47) together with the following equations are satisfied:

$$\dot{\tilde{\lambda}}(t) = -\tilde{\lambda}^T \tilde{\phi}_x \tag{2.49a}$$

$$\Delta\tilde{u}^T / \tilde{\alpha} - \tilde{\lambda}^T \tilde{\phi}_u = 0 \tag{2.49b}$$

$$-\int_0^1 \{\tilde{\lambda}^T \tilde{\phi}_\pi\}\, dt + [\tilde{\mu}^T \tilde{\omega}_\pi]_0 + [\tilde{\hat{\mu}}^T \tilde{\psi}_\pi]_1 + \frac{\Delta\tilde{\pi}^T}{\tilde{\alpha}} = 0 \tag{2.49c}$$

$$[-\tilde{\lambda}^T + \tilde{\mu}^T \tilde{\omega}_x]_0 = 0 \tag{2.49d}$$

$$[\tilde{\lambda}^T + \tilde{\hat{\mu}}^T \tilde{\psi}_x]_1 = 0 \tag{2.49e}$$

2.7.1 Coordinate Transformation

Let

$$\tilde{A}(t) = \frac{\Delta\tilde{x}(t)}{\tilde{\alpha}}, \tag{2.50}$$

$$\tilde{B}(t) = \frac{\Delta\tilde{u}(t)}{\tilde{\alpha}}, \tag{2.51}$$

$$\tilde{C} = \frac{\Delta\tilde{\pi}}{\tilde{\alpha}}. \tag{2.52}$$

Then, (2.47a), and (2.49a) – (2.49c) become:

$$\dot{\tilde{A}} = \tilde{\phi}_x \tilde{A} + \tilde{\phi}_u \tilde{B} + \tilde{\phi}_\pi \tilde{C} - (\dot{\tilde{x}} - \tilde{\phi}) \tag{2.53}$$

$$\dot{\tilde{\lambda}}^T = -\tilde{\lambda}^T \tilde{\phi}_x \tag{2.54}$$

$$\tilde{B}^T = \tilde{\lambda}^T \tilde{\phi}_u \tag{2.55}$$

$$\tilde{C}^T = \int_0^1 \tilde{\lambda}^T \tilde{\phi}_\pi \, dt - [\tilde{\mu}^T \tilde{\omega}_\pi]_0 - [\tilde{\hat{\mu}}^T \tilde{\psi}_\pi]_1 \tag{2.56}$$

and the boundary conditions (2.47b) – (2.47c) and (2.49d) – (2.49e) become:

$$[\tilde{\omega} + \tilde{\omega}_x \tilde{A} + \tilde{\omega}_\pi \tilde{C}]_0 = 0 \tag{2.57}$$

$$[\tilde{\psi} + \tilde{\psi}_x \tilde{A} + \tilde{\psi}_\pi \tilde{C}]_1 = 0 \tag{2.58}$$

$$[-\tilde{\lambda}^T + \tilde{\mu}^T \tilde{\omega}_x]_0 = 0 \tag{2.59}$$

$$[\tilde{\lambda}^T + \tilde{\hat{\mu}}^T \tilde{\psi}_x]_1 = 0 \tag{2.60}$$

2.7.2 Integration Technique

The aim of this section is to describe a method for solving the restoration corrections $\tilde{A}(t)$, $\tilde{B}(t)$, and $\tilde{C}$.

From (2.54) and (2.60), we obtain

$$\tilde{\lambda}(t) = -(N(1,t))^T ([\tilde{\psi}_x]_1)^T \tilde{\hat{\mu}} \tag{2.61}$$

where $N(t,\tau)$ satisfies the following matrix equation:

$$\frac{\partial N(t,\tau)}{\partial t} = \tilde{\phi}_x(t) \, N(t,\tau) \tag{2.62a}$$

$$N(\tau,\tau) = \mathrm{I} \tag{2.62b}$$

$$N(t,\tau) = 0 \qquad \text{if } \tau > t, \tag{2.62c}$$

where I denotes the identity matrix of appropriate dimension.

Using equation (2.61), it follows from (2.55), (2.56) and (2.59) that

$$\tilde{B}(t) = \tilde{\kappa}_1(t)\,\tilde{\hat{\mu}} \tag{2.63}$$

$$\tilde{C} = \tilde{\kappa}_2\,\tilde{\mu} + \tilde{\kappa}_3\,\tilde{\hat{\mu}} \tag{2.64}$$

and

$$\tilde{\kappa}_4\,\tilde{\mu} + \tilde{\kappa}_5\,\tilde{\hat{\mu}} = 0 \tag{2.65}$$

where

$$\tilde{\kappa}_1(t) = -(\tilde{\phi}_n(t))^{\mathrm{T}}\,(N(1,t))^{\mathrm{T}}\,([\tilde{\psi}_x]_1)^{\mathrm{T}} \tag{2.66a}$$

$$\tilde{\kappa}_2 = -([\tilde{\omega}_n]_0)^{\mathrm{T}} \tag{2.66b}$$

$$\tilde{\kappa}_3 = -\int_0^1 (\tilde{\phi}_x)^{\mathrm{T}}\,(N(1,t))^{\mathrm{T}}\,([\tilde{\psi}_x]_1)^{\mathrm{T}}\,dt - ([\tilde{\psi}_n]_1)^{\mathrm{T}} \tag{2.66c}$$

$$\tilde{\kappa}_4 = ([\tilde{\omega}_x]_0)^{\mathrm{T}} \tag{2.66d}$$

and

$$\tilde{\kappa}_5 = (N(1,0))^{\mathrm{T}}\,([\tilde{\psi}_x]_1)^{\mathrm{T}} \tag{2.66e}$$

Thus, from (2.53), (2.63) and (2.64), we obtain

$$\tilde{A}(t) = N(t,0)\,\tilde{A}(0) + \tilde{\kappa}_6(t)\tilde{\mu} + \tilde{\kappa}_7(t)\tilde{\hat{\mu}} + \tilde{\kappa}_8(t) \tag{2.67}$$

where

$$\tilde{\kappa}_6(t) = [\int_0^t N(t,\tau)\,\tilde{\phi}_n(\tau)\,d\tau]\,\tilde{\kappa}_2 \tag{2.68a}$$

$$\tilde{\kappa}_7(t) = \int_0^t N(t,\tau)\,\tilde{\phi}_u(\tau)\,\tilde{\kappa}_1(\tau)d\tau + [\int_0^t N(t,\tau)\,\tilde{\phi}_\pi(\tau)d\tau]\,\tilde{\kappa}_3 \tag{2.68b}$$

$$\tilde{\kappa}_8(t) = -\int_0^t N(t,\tau)\,[\dot{\tilde{x}}(\tau) - \dot{\tilde{\phi}}(\tau)]\,d\tau \tag{2.68c}$$

Then, it follows from (2.65), (2.57), (2.58), (2.67), and (2.64) that

$$\tilde{\kappa}_4\,\tilde{\mu} + \tilde{\kappa}_5\,\hat{\tilde{\mu}} = 0 \tag{2.69a}$$

$$[\tilde{\omega}_x]_0\,\tilde{A}(0) + [\tilde{\omega}_\pi\tilde{\kappa}_2]_0\,\tilde{\mu} + [\tilde{\omega}_\pi\tilde{\kappa}_3]_0\,\hat{\tilde{\mu}} = -[\tilde{\omega}]_0 \tag{2.69b}$$

$$[\tilde{\psi}_x]_1\,N(1,0)\,\tilde{A}(0) + [\tilde{\psi}_x\tilde{\kappa}_6 + \tilde{\psi}_\pi\tilde{\kappa}_2]_1\,\tilde{\mu} + [\tilde{\psi}_x\tilde{\kappa}_7 + \tilde{\psi}_\pi\tilde{\kappa}_3]_1\,\hat{\tilde{\mu}}$$
$$= -[\tilde{\psi} + \tilde{\psi}_x\tilde{\kappa}_8]_1 \tag{2.69c}$$

Remark 2.7.2.1 System (2.69) constitutes a system of $(n+r+s)$ equations in $(n+r+s)$ unknowns (i.e., $\tilde{A}(0)$, $\tilde{\mu}$ and $\hat{\tilde{\mu}}$). It has a unique solution if and only if the nominal function $\tilde{z}$ satisfies the constraint qualification (cf. Remark 2.6.1).

Remark 2.7.2.2 Note that $\tilde{A}(t)$, $\tilde{B}(t)$ and $\tilde{C}$ can be easily obtained from (2.67), (2.55) and (2.56), respectively, once $\tilde{A}_0$, $\tilde{\mu}$ and $\hat{\tilde{\mu}}$ are determined from (2.69). Then, $\Delta\tilde{x}(t)$, $\Delta\tilde{u}$, $\Delta\tilde{\pi}$ can be calculated from (2.50) – (2.52) if $\tilde{\alpha}$ is known.

The method for determining the step-size $\tilde{\alpha}$ is discussed in the next section.

2.7.3 The Restoration Algorithm

Algorithm 2.7.3 **(Restoration Algorithm)**

Step 0 Let $\tilde{\gamma} \in (0,\frac{1}{2})$, $\tilde{\beta} \in (\frac{1}{2},1)$ be given constants. Choose $\tilde{z}^0 = (\tilde{x}^0,\tilde{u}^0,\tilde{\pi}^0) \in \mathscr{Z}$. Set i = 0.

Step 1 If $P(\tilde{z}^i) = 0$, let $\tilde{z}^{i+j} = \tilde{z}^i$ $\forall$ $j \geq 0$ and stop; otherwise go to Step 2.

Step 2 Let the ominal function be $\tilde{z}^i$. Then, calculate $\tilde{A}(\tilde{z}^i)(t)$, $\tilde{B}(\tilde{z}^i)(t)$ and $\tilde{C}(\tilde{z}^i)$ by the method described in Section 6.2.

Step 3 Choose $\tilde{\alpha}_i$ to be the first element in the sequence $1, \tilde{\beta}, (\tilde{\beta})^2, \ldots$ such that

$$P(\tilde{z}^i + \tilde{\alpha}_i \tilde{D}(\tilde{z}^i)) - P(\tilde{z}^i) \leq -2\,\tilde{\gamma}\,\tilde{\alpha}_i\,P(\tilde{z}^i) \,, \tag{2.70}$$

where

$$\tilde{D}(\tilde{z}^i) = (\tilde{A}(\tilde{z}^i), \tilde{B}(\tilde{z}^i), \tilde{C}(\tilde{z}^i)) \tag{2.71}$$

Step 4 Set

$$\tilde{z}^{i+1} = \tilde{z}^i + \tilde{\alpha}_i \tilde{D}(\tilde{z}^i), \tag{2.72}$$

and go to Step 1.

Definition 2.7.3 Let $r(\tilde{z},i)$ be the function generated at the i^{th} iteration of Algorithm 2.7.3, starting from the nominal function $r(\tilde{z},0) = \tilde{z}$.

To continue, we assume that the following assumptions are satisfied.

Assumption 2.7.3A

(2.7.3A1) Assumption (2.3.A) is satisfied.

(2.7.3A2) Let $\{r(\tilde{z},i)\}$ be a sequence generated by Algorithm 2.7.3. Then, each of its elements satisfies the constraint qualification. Hence, from Remark 2.7.2.1, for each $\tilde{z}$ and i,

$$\text{Det}\,[\tilde{M}(r(\tilde{z},i))] \neq 0 \tag{2.73}$$

where $\tilde{M}$ is the coefficient matrix of system (2.69). If $\{r(\tilde{z},i)\}$ is an infinite sequence, then

$$\lim_{i \to \infty} \text{Det}\,[\tilde{M}(r(\tilde{z},i))] \neq 0 \tag{2.74}$$

Lemma 2.7.3.1 Let Assumption (2.7.3A) be satisfied. If $\{r(\tilde{z},i)\}$ is a sequence generated by Algorithm 2.7.3, then $\{r(\tilde{z},i)\}$ has the following properties:

(i) There exists a constant $\ell_1 > 0$, independent of both $\tilde{z}$ and i, such that

$$\|\tilde{D}(r(\tilde{z},i))\| \leq \ell_1 [P(r(\tilde{z},i))]^{\frac{1}{2}} \tag{2.75}$$

(ii) $\tilde{D}(r(\tilde{z},i)) \in \mathcal{U}$ for any $r(\tilde{z},i) \in \mathcal{U}$. (2.76)

Proof For brevity, $(\tilde{z},i)$ is used to denote $r(\tilde{z},i)$ in the proof.

Proof of Part (i) From (2.62), we have

$$N(\tilde{z},i)(t,\tau) = I + \int_\tau^t \tilde{\phi}_x(\tilde{z},i)(s)\, N(\tilde{z},i)(s,\tau)\, ds \tag{2.77}$$

Thus, from Assumption (2.3.A2), we obtain

$$|N(\tilde{z},i)(t,\tau)| \leq \sqrt{n} + \int_\tau^t K_1\, |N(\tilde{z},i)(s,\tau)|\, ds \tag{2.78}$$

By Gronwall's lemma (cf. Ref. 5, p.62, Lemma 2.3.5), it follows that

$$\begin{aligned} |N(\tilde{z},i)(t,\tau)| &\leq \exp[(t-\tau)K_1] \\ &\leq \exp(K_1) \end{aligned} \tag{2.79}$$

Thus, by (2.79), Assumption (2.4.A2), the Cauchy–Schwarz inequality, and (2.42), we deduce from (2.66) and (2.68) that there exists a constant $\bar{K}_1 > 0$, independent of both $\tilde{z}$ and i, such that

$$|\tilde{\kappa}_j(\tilde{z},i)| \le \hat{K}_1, \qquad j = 2,\dots,5 \tag{2.80}$$

$$|\tilde{\kappa}_j(\tilde{z},i)(t)| \le \hat{K}_1, \quad \forall\, t \in [0,1], \qquad j = 1,6,7 \tag{2.81}$$

and

$$|\tilde{\kappa}_8(\tilde{z},i)(t)| \le \hat{K}_1\, [P(\tilde{z},i)]^{\frac{1}{2}} \quad ,\forall\, t \in [0,1], \tag{2.82}$$

where $\tilde{\kappa}_i$, i = 1,...,8, are as defined in (2.66a) – (2.66e) and (2.68a) – (2.68c).

From (2.69), Assumption (2.7.3A2), we have

$$\begin{bmatrix} [\tilde{A}(\tilde{z},i)]_0 \\ \tilde{\mu}(\tilde{z},i) \\ \tilde{\hat{\mu}}(\tilde{z},i) \end{bmatrix} = [\tilde{M}(\tilde{z},i)]^{-1} \begin{bmatrix} 0 \\ -[\tilde{\omega}(\tilde{z},i)]_0 \\ -[\tilde{\psi}(\tilde{z},i)]_1 + [\tilde{\psi}_x(\tilde{z},i)\,\tilde{\kappa}_8(\tilde{z},i)]_1 \end{bmatrix} \tag{2.83}$$

where $\tilde{M}$ is the coefficient matrix of the system (2.69).

In view of (2.79) – (2.81) and Assumption (2.4.A2), it is clear that all the coefficients of the matrix $\tilde{M}(\tilde{z},i)$ are bounded. Thus, we deduce from (2.75) that there exists a constant $\hat{K}_2 > 0$, independent of both $\tilde{z}$ and i, such that

$$|[\tilde{M}(\tilde{z},i)]^{-1}| \le \hat{K}_2 \tag{2.84}$$

By (2.83), (2.84), (2.42), (2.82) and Assumption (2.4.A2), it follows that

$$|[\tilde{A}(\tilde{z},i)]_0|$$
$$\leq \hat{K}_2 [P(\tilde{z},i)]^{\frac{1}{2}} [1 + (1 + \hat{K}_1 K_1)^2]^{\frac{1}{2}} = \hat{K}_3 [P(\tilde{z},i)]^{\frac{1}{2}} \tag{2.85a}$$

$$|[\tilde{\mu}(\tilde{z},i)]| \leq \hat{K}_3 [P(\tilde{z},i)]^{\frac{1}{2}} \tag{2.85b}$$

and

$$|[\tilde{\hat{\mu}}(\tilde{z},i)]| \leq \hat{K}_3 [P(\tilde{z},i)]^{\frac{1}{2}} \tag{2.85c}$$

where

$$\hat{K}_3 = \hat{K}_2 [1 + (1 + \tilde{K}_1 K_1)^2]^{\frac{1}{2}} \tag{2.86}$$

Thus, by virtue of (2.79), (2.81) – (2.82), (2.85) and (2.80), we deduce from the formulae of $\tilde{A}(t)$, $\tilde{B}(t)$ and $\tilde{C}$ (defined by (2.67), (2.63) and (2.64), respectively) that there exists a constant $\hat{K}_4 > 0$, independent of both $\tilde{z}$ and i, such that

$$|[\tilde{A}(\tilde{z},i)]_1| \leq \hat{K}_4 [P(\tilde{z},i)]^{\frac{1}{2}} \tag{2.87a}$$

$$\|\tilde{A}(\tilde{z},i)\|_\infty \leq \hat{K}_4 [P(\tilde{z},i)]^{\frac{1}{2}} \tag{2.87b}$$

$$\|\tilde{B}(\tilde{z},i)\|_\infty \leq \hat{K}_4 [P(\tilde{z},i)]^{\frac{1}{2}} \tag{2.87c}$$

and

$$|\tilde{C}(\tilde{z},i)| \leq \hat{K}_4 [P(\tilde{z},i)]^{\frac{1}{2}} \tag{2.87d}$$

Hence, from (2.53), (2.87b) – (2.87d), Assumption (2.4.A2) and (2.42), it follows that there exists a constant $\hat{K}_5 > 0$, independent of both $\tilde{z}$ and i, such that

$$\|\dot{\tilde{A}}(\tilde{z},i)\|_\infty \leq \hat{K}_5 [P(\tilde{z},i)]^{\frac{1}{2}} \tag{2.88}$$

Thus, from (2.7), (2.87b), (2.88), (2.87c) – (2.87d), (2.85a), and (2.87a), we obtain

$$
\begin{aligned}
&\|\tilde{D}(\tilde{z},t)\| \\
&= \|\tilde{A}(\tilde{z},t)\|_\infty + \|\dot{\tilde{A}}(\tilde{z},t)\|_\infty + \|\tilde{B}(\tilde{z},t)\|_\infty + |\tilde{C}(\tilde{z},t)| + |[\tilde{A}(\tilde{z},t)]_0| \\
&+ |[\tilde{A}(\tilde{z},t)]_1| \leq \ell_1 [P(\tilde{z},t)]^{\frac{1}{2}}
\end{aligned}
\tag{2.89}
$$

where

$$
\ell_1 = \hat{K}_3 + 4\hat{K}_4 + \hat{K}_5 \tag{2.90}
$$

Proof of Part (ii) The proof follows easily from (2.89) and Remark 2.6.1.

Lemma 2.7.3.2 Let Assumption (2.7.3A) be satisfied. Then,

$$
\lim_{\tilde{\alpha}\to\infty} \frac{P(r(\tilde{z},t) + \tilde{\alpha}\tilde{D}(r(\tilde{z},t))) - P(r(\tilde{z},t))}{\tilde{\alpha}} = -2P(r(\tilde{z},t)) \tag{2.91}
$$

Proof For brevity, define

$$
r(\tilde{z},t) - \tilde{z} = (\tilde{x},\tilde{u},\tilde{\pi}) \tag{2.92}
$$

from (2.42), we have

$$
\begin{aligned}
&P(\tilde{z} + \tilde{\alpha}\tilde{D}(\tilde{z})) - P(\tilde{z}) \\
&= \int_0^1 \{|\dot{\tilde{x}}(t) + \tilde{\alpha}\dot{\tilde{A}}(\tilde{z})(t) - \phi(t,\tilde{x}(t) + \tilde{\alpha}\tilde{A}(\tilde{z})(t), \tilde{u}(t) + \tilde{\alpha}\tilde{B}(\tilde{z})(t), \tilde{\pi} + \tilde{\alpha}\tilde{C}(\tilde{z}))|^2 \\
&\qquad - |\dot{\tilde{x}}(t) - \phi(t,\tilde{x}(t), \tilde{u}(t), \tilde{\pi})|^2\}\, dt \\
&+ |[\omega(\tilde{x} + \tilde{\alpha}\tilde{A}(\tilde{z}), \tilde{\pi} + \tilde{\alpha}\tilde{C}(\tilde{z}))]_0|^2 - |[\omega(\tilde{x},\tilde{\pi})]_0|^2 \\
&+ |[\psi(\tilde{x} + \tilde{\alpha}\tilde{A}(\tilde{z}), \tilde{\pi} + \tilde{\alpha}\tilde{C}(\tilde{z}))]_1|^2 - |[\psi(\tilde{x},\tilde{\pi})]_1|^2
\end{aligned}
\tag{2.93}
$$

By the mean valued theorem, we obtain

$$
\begin{aligned}
&|\dot{\tilde{x}}(t) + \tilde{\alpha}\, \dot{\tilde{A}}(\tilde{z})(t) - \phi(t, \tilde{x}(t) + \tilde{\alpha}\, \tilde{A}(\tilde{z})(t), \tilde{u}(t) + \tilde{\alpha}\, \tilde{B}(\tilde{z})(t), \tilde{\pi} + \tilde{\alpha}\, \tilde{C}(\tilde{z}))|^2 \\
&- |\dot{\tilde{x}}(t) - \phi(t, \tilde{x}(t), \tilde{u}(t), \tilde{\pi})|^2 \\
&= |\dot{\tilde{x}}(t) + \tilde{\alpha}\, \dot{\tilde{A}}(\tilde{z})(t) - \phi(t, \tilde{x}(t), \tilde{u}(t), \tilde{\pi}) - \tilde{\alpha}\, L_1(\tilde{z})(t)|^2 \\
&- |\dot{\tilde{x}}(t) - \phi(t, \tilde{x}(t), \tilde{u}(t), \tilde{\pi})|^2 \\
&= \tilde{\alpha}^2\, |\dot{\tilde{A}}(\tilde{z})(t) - L_1(\tilde{z})(t)|^2 \\
&+ 2\, \tilde{\alpha}\, [\dot{\tilde{x}}(t) - \phi(t, \tilde{x}(t), \tilde{u}(t), \tilde{\pi})]^T [\dot{\tilde{A}}(\tilde{z})(t) - L_1(\tilde{z})(t)] \\
&= \tilde{\alpha}^2\, |\dot{\tilde{A}}(\tilde{z})(t) - L_2(\tilde{z})(t)|^2 \\
&+ 2\, \tilde{\alpha}^2\, [\dot{\tilde{A}}(\tilde{z})(t) - L_2(\tilde{z})(t)]^T [L_2(\tilde{z})(t) - L_1(\tilde{z})(t)] \\
&+ \tilde{\alpha}^2\, |L_2(\tilde{z})(t) - L_1(\tilde{z})(t)|^2 \\
&+ 2\, \tilde{\alpha}\, [\dot{\tilde{x}}(t) - \phi(t, \tilde{x}(t), \tilde{u}(t), \tilde{\pi})]^T [\dot{\tilde{A}}(\tilde{z})(t) - L_3(\tilde{z})(t)] \\
&+ 2\, \tilde{\alpha}\, [\dot{\tilde{x}}(t) - \phi(t, \tilde{x}(t), \tilde{u}(t), \tilde{\pi})]^T [L_2(\tilde{z})(t) - L_1(\tilde{z})(t)]
\end{aligned}
\tag{2.94}
$$

where

$$
\begin{aligned}
&L_1(\tilde{z})(t) \\
&= \phi_x(t, \tilde{x}(t) + \hat{\alpha}\tilde{A}(\tilde{z})(t), \tilde{u}(t) + \hat{\alpha}\tilde{B}(\tilde{z})(t), \tilde{\pi} + \hat{\alpha}\tilde{C}(\tilde{z}))\tilde{A}(\tilde{z})(t) \\
&+ \phi_u(t, \tilde{x}(t) + \hat{\alpha}\tilde{A}(\tilde{z})(t), \tilde{u}(t) + \hat{\alpha}\tilde{B}(\tilde{z})(t), \tilde{\pi} + \hat{\alpha}\tilde{C}(\tilde{z}))\tilde{B}(\tilde{z})(t) \\
&+ \phi_\pi(t, \tilde{x}(t) + \hat{\alpha}\tilde{A}(\tilde{z})(t), \tilde{u}(t) + \hat{\alpha}\tilde{B}(\tilde{z})(t), \tilde{\pi} + \hat{\alpha}\tilde{C}(\tilde{z}))\tilde{C}(\tilde{z}),
\end{aligned}
\tag{2.95}
$$

$$
\begin{aligned}
&L_2(\tilde{z})(t) \\
&= \phi_x(t, \tilde{x}(t), \tilde{u}(t), \tilde{\pi})\, \tilde{A}(\tilde{z})(t) + \phi_u(t, \tilde{x}(t), \tilde{u}(t), \tilde{\pi})\, \tilde{B}(\tilde{z})(t) \\
&+ \phi_\pi(t, \tilde{x}(t), \tilde{u}(t), \tilde{\pi})\, \tilde{C}(\tilde{z}),
\end{aligned}
\tag{2.96}
$$

and $\hat{\alpha}$ is an intermediate value satisfying

$$0 \leq \hat{\alpha} \leq \tilde{\alpha} \tag{2.97}$$

From (2.94), (2.96), and (? 53), it follows that

$$\begin{aligned}
&|\dot{\tilde{x}}(t) + \tilde{\alpha}\,\dot{\tilde{A}}(\tilde{z})(t) - \phi(t,\tilde{x}(t) + \tilde{\alpha}\,\tilde{A}(\tilde{z})(t), \tilde{u}(t) + \tilde{\alpha}\,\tilde{B}(\tilde{z})(t), \tilde{\pi} + \tilde{\alpha}\,\tilde{C}(\tilde{z}))|^2 \\
&- |\dot{\tilde{x}}(t) - \phi(t,\tilde{x}(t), \tilde{u}(t), \tilde{\pi})|^2 \\
&= (\tilde{\alpha}^2 - 2\,\tilde{\alpha})\,|\dot{\tilde{x}}(t) - \phi(t,\tilde{x}(t), \tilde{u}(t), \tilde{\pi})|^2 \\
&+ (2\,\tilde{\alpha} - 2\,\tilde{\alpha}^2)\,[\dot{\tilde{x}}(t) - \phi(t,\tilde{x}(t), \tilde{u}(t), \tilde{\pi})]^T\,[L_2(\tilde{z})(t) - L_1(\tilde{z})(t)] \\
&+ \tilde{\alpha}^2\,|L_2(\tilde{z})(t) - L_1(\tilde{z})(t)|^2
\end{aligned} \tag{2.98}$$

From (2.95), (2.96), Assumption (2.4.A2), (2.89) and (2.97), we obtain

$$\|L_2(\tilde{z})(t) - L_1(\tilde{z})(t)\|_\infty \leq K_2\,\ell_1^2\,\hat{\alpha}\,P(\tilde{z}) \leq K_2\,\ell_1^2\,\tilde{\alpha}\,P(\tilde{z}) \tag{2.99}$$

Thus, from Cauchy–Schwartz inequality, (2.99) and (2.42), it follows that

$$\begin{aligned}
&\left|\int_0^1 [\dot{\tilde{x}}(t) - \phi(t,\tilde{x}(t), \tilde{u}(t), \tilde{\pi})]^T\,[L_2(\tilde{z})(t) - L_1(\tilde{z})(t)]\,dt\right| \\
&= K_2\,\ell_1^2\,\tilde{\alpha}\,[P(\tilde{z})]^{\frac{3}{2}}
\end{aligned} \tag{2.100}$$

Now, by (2.98), (2.100), and (2.99), we deduce that

$$\begin{aligned}
\Big|\lim_{\tilde{\alpha}\to 0} \frac{1}{\tilde{\alpha}} \int_0^1 [\,|\dot{\tilde{x}}(t) + \tilde{\alpha}\,\dot{\tilde{A}}(\tilde{z})(t) \\
- \phi(t,\tilde{x}(t) + \tilde{\alpha}\,\tilde{A}(\tilde{z})(t), \tilde{u}(t) + \tilde{\alpha}\,\tilde{B}(\tilde{z})(t), \tilde{\pi} + \tilde{\alpha}\,\tilde{C}(\tilde{z}))|^2 \\
- |\dot{\tilde{x}}(t) - \phi(t,\tilde{x}(t), \tilde{u}(t), \tilde{\pi})|^2]\,dt
\end{aligned}$$

$$+ 2\int_0^1 [|\dot{\tilde{x}}(t) - \phi(t,\tilde{x}(t), \tilde{u}(t), \tilde{\pi})|^2]\, dt\Big|$$

$$\leq \lim_{\tilde{\alpha}\to 0} \{\tilde{\alpha}\, P(\tilde{z}) + [2\tilde{\alpha} - 2\tilde{\alpha}^2]\, K_2\, \ell_1^2\, [P(\tilde{z})]^{\frac{3}{2}} + \tilde{\alpha}^3 K_2^2\, \ell_1^4\, [P(\tilde{z})]^{\frac{1}{2}}\} \tag{2.101}$$

Thus, it follows from (2.101) and Remark 2.5.1 that

$$\lim_{\tilde{\alpha}\to 0} \frac{1}{\tilde{\alpha}} \int_0^1 [|\dot{\tilde{x}}(t) + \tilde{\alpha}\, \dot{\tilde{A}}(\tilde{z})(t)$$

$$- \phi(t,\tilde{x}(t) + \tilde{\alpha}\, \tilde{A}(\tilde{z})(t), \tilde{u}(t) + \tilde{\alpha}\, \tilde{B}(\tilde{z})(t), \tilde{\pi} + \tilde{\alpha}\, \tilde{C}(\tilde{z}))|^2$$

$$- |\dot{\tilde{x}}(t) - \phi(t,\tilde{x}(t), \tilde{u}(t), \tilde{\pi})|^2]\, dt$$

$$= -2\int_0^1 [|\dot{\tilde{x}}(t) - \phi(t,\tilde{x}(t), \tilde{u}(t), \tilde{\pi})|^2]\, dt \tag{2.102}$$

Next, by following a similar argument as that used to obtain (2.102) from (2.94), with the exception that (2.53) is being replaced by both (2.57) and (2.58), we can readily show that

$$\lim_{\tilde{\alpha}\to 0} \{|[\omega(\tilde{x} + \tilde{\alpha}\, \tilde{A}(\tilde{z}), \tilde{\pi} + \tilde{\alpha}\, \tilde{C}(\tilde{z}))]_0|^2 - |[\omega(\tilde{x},\tilde{\pi})]_0|^2\}/\tilde{\alpha}$$
$$= -2\, |[\omega(\tilde{x}, \tilde{\pi})]_0|^2 \tag{2.103}$$

and

$$\lim_{\tilde{\alpha}\to 0} \{|[\psi(\tilde{x} + \tilde{\alpha}\, \tilde{A}(\tilde{z}), \tilde{\pi} + \tilde{\alpha}\, \tilde{C}(\tilde{z}))]_1|^2 - |[\psi(\tilde{x},\tilde{\pi})]_1|^2\}/\tilde{\alpha}$$
$$= -2\, |[\psi(\tilde{x}, \tilde{\pi})]_1|^2 \tag{2.104}$$

Combining (2.93), (2.102) – (2.104), we obtain

$$\lim_{\tilde{\alpha}\to \infty} \frac{P(\tilde{z} + \tilde{\alpha}\, \tilde{D}(\tilde{z})) - P(\tilde{z})}{\tilde{\alpha}} = -2\, P(\tilde{z}) \tag{2.105}$$

This completes the proof of the lemma.

Remark 2.7.3.2 Consider Step 3 of Algorithm 2.7.3. Since $\tilde{\gamma} \in (0,\cdot'2)$ and $P(r(\tilde{z},i)) > 0$, it is clear from Lemma 2.7.3.2 that the inequality (2.70) will be satisfied when $\tilde{\alpha}_i$ is sufficiently small. Thus, Step 3 of Algorithm 2.7.3 is well-defined.

Remark 2.7.3.3 Let $\{\tilde{z}^i\}$ be a sequence in $\mathscr{U}$ such that $\lim_{i\to\infty} P(\tilde{z}^i) = a$, where a is a finite number. Furthermore, let $\{\alpha_i\}$ be a sequence of real numbers such that $\lim_{i\to\infty} \alpha_i = 0$. Then, by following a similar argument as that given in the proof of Lemma 2.7.3.2, we can show that

$$\lim_{i\to\infty} \frac{P(\tilde{z}^i + \alpha_i \tilde{D}(\tilde{z}^i)) - P(\tilde{z}^i)}{\alpha_i} = -2a$$

Theorem 2.7.3 Let $\{\tilde{z}^i\}$ be a sequence generated by Algorithm 2.7.3. Then the sequence $\{P(\tilde{z}^i)\}$ converges to zero at a quadratic rate.

Proof Firstly, we want to prove that

$$\lim_{i\to\infty} P(\tilde{z}^i) = 0$$

Since $\{P(\tilde{z}^i)\}$ is a non–increasing sequence of real numbers, which has a lower bound zero, it is convergent. Suppose that

$$\lim_{i\to\infty} P(\tilde{z}^i) = a > 0 \qquad (2.106)$$

From (2.70) and (2.72), we have

$$2\,\tilde{\gamma}\tilde{\alpha}_i\,P(\tilde{z}^i) \leq P(\tilde{z}^i) - P(\tilde{z}^{i+1}) \tag{2.107}$$

for all i. Hence, it follows from (2.107) and (2.106) that

$$\lim_{i\to\infty} \tilde{\alpha}_i = 0 \tag{2.108}$$

Now, let N be a positive integer such that $\tilde{\alpha}_i < 1$ for all $i \geq N$. Since $\tilde{\alpha}_i$ is the first element in the sequence $1, \tilde{\beta}, \tilde{\beta}^2, \ldots$ such that (2.70) holds. Furthermore, $\tilde{\alpha}_i/\tilde{\beta}$ is the element which just preceeds $\tilde{\alpha}_i$ in the above sequence. Thus, (2.70) does not hold if we replace $\tilde{\alpha}_i$ by $\tilde{\alpha}_i/\tilde{\beta}$ in (2.70). On this basis, it follows that

$$P(\tilde{z}^i + (\tilde{\alpha}_i/\tilde{\beta})\,\tilde{D}(\tilde{z}^i)) - P(\tilde{z}^i) > -(2\,\tilde{\alpha}_i/\tilde{\beta})\,\tilde{\gamma}\,P(\tilde{z}^i) \tag{2.109}$$

for all $i \geq N$.

From (2.106) and (2.108), we note that $\lim_{i\to\infty} P(\tilde{z}^i) = a$ and $\lim_{i\to\infty} \{\tilde{\alpha}_i/\tilde{\beta}\} = 0$, respectively. Thus, by substituting $\alpha_i = \tilde{\alpha}_i/\tilde{\beta}$ in Remark 2.7.3.3, we have

$$\lim_{i\to\infty} \{P(\tilde{z}^i + (\tilde{\alpha}_i/\tilde{\beta})\,\tilde{D}(\tilde{z}^i)) - P(\tilde{z}^i)\}/2(\tilde{\alpha}_i/\tilde{\beta}) = -a.$$

On the other hand, it is clear from (2.109) that

$$\lim_{i\to\infty} \{P(\tilde{z}^i + (\tilde{\alpha}_i/\tilde{\beta})\,\tilde{D}(\tilde{z}^i)) - P(\tilde{z}^i)\}/2(\tilde{\alpha}_i/\tilde{\beta}) \geq -\tilde{\gamma}\,a.$$

Combining them together, we obtain

$$\lim_{i \to \infty} (1 - \tilde{\gamma}) P(\tilde{z}^i) \leq 0.$$

This is, however, a contradiction. Therefore, we conclude that

$$\lim_{i \to \infty} P(\tilde{z}^i) = 0. \tag{2.110}$$

Next, we want to prove that $\{P(\tilde{z}^i)\}$ converges to zero at a quadratic rate.

By letting $\tilde{\alpha} = 1$ in (2.98), we obtain from (2.98) and (2.99) that

$$|\dot{\tilde{x}}^i(t) + \dot{\tilde{A}}(\tilde{z}^i)(t) - \phi(t, \tilde{x}^i(t) + \tilde{A}(\tilde{z}^i)(t), \tilde{u}^i(t) + \tilde{B}(\tilde{z}^i)(t), \tilde{\pi}^i + \tilde{C}(\tilde{z}^i))|^2$$
$$\leq K_2^2 \ell_1^4 [P(\tilde{z}_i)]^2 \tag{2.111}$$

In view of the mean value theorem, (2.57), Assumption (2.4.A2) and (2.75), we have

$$|[\omega(\tilde{x}^i + \tilde{A}(\tilde{z}^i), \tilde{\pi}^i + \tilde{C}(\tilde{z}^i))]_0|^2$$
$$= |[\omega(\tilde{x}^i, \tilde{\pi}^i) + \tilde{\omega}_x(\tilde{x}^i + \hat{\alpha}\tilde{A}(\tilde{z}^i), \tilde{\pi}^i + \hat{\alpha}\tilde{C}(\tilde{z}^i))\tilde{A}(\tilde{z}^i)$$
$$+ \tilde{\omega}_\pi(\tilde{x}^i + \hat{\alpha}\tilde{A}(\tilde{z}^i), \tilde{\pi}^i + \hat{\alpha}\tilde{C}(\tilde{z}^i))\tilde{C}(\tilde{z}^i)]_0|^2$$
$$= |\{[\omega_x(\tilde{x}^i + \hat{\alpha}\tilde{A}(\tilde{z}^i), \tilde{\pi}^i + \hat{\alpha}\tilde{C}(\tilde{z}^i)) - \omega_x(\tilde{x}^i + \tilde{A}(\tilde{z}^i), \tilde{\pi}^i + \tilde{C}(\tilde{z}^i))]\tilde{A}(\tilde{z}^i)$$
$$+ [\omega_\pi(\tilde{x}^i + \hat{\alpha}\tilde{A}(\tilde{z}^i), \tilde{\pi}^i + \hat{\alpha}\tilde{C}(\tilde{z}^i)) - \omega_\pi(\tilde{x}^i + \tilde{A}(\tilde{z}^i), \tilde{\pi}^i + \tilde{C}(\tilde{z}^i))\tilde{C}(\tilde{z}^i)\}_0|$$

$$\leq (1 - \hat{\alpha})^2 K_2^2 |[\tilde{A}(\tilde{z}^i) + \tilde{C}(\tilde{z}^i)]_0|^4$$
$$\leq K_2^2 \|\tilde{D}(\tilde{z}^i)\|^4$$
$$\leq K_2^2 \ell_1^4 [P(\tilde{z}^i)]^2, \tag{2.112}$$

where $\hat{\alpha}$ is an intermediate value satisfying

$0 \le \hat{\alpha} \le 1$

Similarly, we can prove

$$\|[\psi(\tilde{x}^i + \tilde{A}(\tilde{z}^i), \tilde{\pi}^i + \tilde{C}(\tilde{z}^i)]_1\|^2 \le K_2^2 \ell_1^4 [P(\tilde{z}^i)]^2 . \tag{2.113}$$

From (2.111) – (2.113), we get

$$P(\tilde{z}^i + \tilde{D}(\tilde{z}^i)) \le 3\ K_2^2 \ell_1^4 [P(\tilde{z}^i)]^2 \tag{2.114}$$

Now, in view of (2.110), there exists an integer $i_0 > 0$ such that whenever $i \ge i_0$, we have

$$P(\tilde{z}^i) \le \frac{1 - 2\tilde{\gamma}}{3\ K_2^2\ \ell_1^4} \tag{2.115}$$

From (2.114) and (2.115), we get

$$P(\tilde{z}^i + \tilde{D}(\tilde{z}^i)) \le [1 - 2\tilde{\gamma}]\ P(\tilde{z}^i)$$

whenever $i \ge i_0$.

Thus, inequality (2.70) can be satisfied by choosing $\tilde{\alpha}_i = 1$ for all $i \ge i_0$. Hence,

$$\tilde{z}^{i+1} = \tilde{z}^i + \tilde{D}(\omega) \tag{2.116}$$

for all $i \ge i_0$.

From (2.110), (2.114) and (2.116), we conclude that $\{P(\tilde{z}^i)\}$ converges to zero at a quadratic rate.

Remark 2.7.3.4 From Theorem 2.7.3, it is clear that under Assumption (2.7.3A), Algorithm 2.7.3 generates in a finite number, $i(\epsilon, \tilde{\tilde{z}})$, of iterations, a function $r(\tilde{\tilde{z}}, i(\epsilon, \tilde{\tilde{z}}))$ such that $r(\tilde{\tilde{z}}, i(\epsilon, \tilde{\tilde{z}})) \in T_\epsilon$.

Lemma 2.7.3.3 Suppose that Assumption (2.7.A3) is satisfied. Let $\tilde{\tilde{z}} \in \mathscr{U}$ be such that $P(\tilde{\tilde{z}}) > 0$, and let $r(\tilde{\tilde{z}}, i)$ be as defined in Definition 2.7.3. Then, there exists a constant $\ell_2 > 0$, independent of both $\tilde{\tilde{z}}$ and i, such that

$$\|r(\tilde{\tilde{z}}, i) - \tilde{\tilde{z}}\| \leq \ell_2 [P(\tilde{\tilde{z}})]^{\frac{1}{2}}$$

Proof For brevity, let

$$r(\tilde{\tilde{z}}, i) = \tilde{\tilde{z}}^i = (\tilde{\tilde{x}}^i, \tilde{\tilde{u}}^i, \tilde{\tilde{\pi}}^i)$$

Let $\tilde{\tilde{\alpha}}(\tilde{\tilde{z}}, i)$ be the value of $\tilde{\tilde{\alpha}}$ obtained in the i^{th} iteration of Step 3 of Algorithm 2.7.3. Without loss of generality, we can assume that $\{\tilde{\tilde{\alpha}}(\tilde{\tilde{z}}, i)\}$ is an infinite sequence.

From the proof of Theorem 2.7.3, it is clear that for each $\tilde{\tilde{z}} \in \mathscr{U} \cap T_\epsilon$ and for all i

$$\tilde{\tilde{\alpha}}(\tilde{\tilde{z}}, i) > 0 \tag{2.117}$$

and

$$\lim_{i \to +\infty} \tilde{\tilde{\alpha}}(\tilde{\tilde{z}}, i) = 1 \neq 0 \tag{2.118}$$

Now, it is clear from (2.117) and (2.118) that there exists an $\hat{\alpha}$, $0 < \hat{\alpha} \leq 1$, independent of both $\tilde{\tilde{z}}$ and i, such that

$$1 \geq \tilde{\tilde{\alpha}}(\tilde{\tilde{z}}, i) \geq \hat{\alpha} \tag{2.119}$$

From (2.70) and (2.72), we have

$$P(\tilde{z}^{i+1}) \leq [1 - 2\,\tilde{\gamma}\,\hat{\alpha}]\, P(\tilde{z}^i) = \hat{\sigma}\, P(\tilde{z}^i)$$

where

$$0 < \hat{\sigma} = 1 - 2\,\tilde{\gamma}\,\hat{\alpha} < 1 \tag{2.120}$$

Thus,

$$P(\tilde{z}^i) \leq (\hat{\sigma})^i\, P(\tilde{z}) \tag{2.121}$$

By (2.75) and (2.121), we have

$$\|\tilde{D}(\tilde{z}^i)\| \leq \ell_1 [(\hat{\sigma})^i\, P(\tilde{z})]^{\frac{1}{2}} \tag{2.122}$$

But

$$\tilde{z}^i = \tilde{z} + \sum_{j=0}^{i-1} \tilde{\alpha}(\tilde{z}, j)\, \tilde{D}(\tilde{z}^j) \tag{2.123}$$

Thus, from (2.123), (2.119), (2.122) and (2.120), it follows that

$$\|\tilde{z}^i - \tilde{z}\|$$

$$\leq \ell_1 [P(\tilde{z})]^{\frac{1}{2}} \sum_{j=0}^{i-1} [(\hat{\sigma})^j]^{\frac{1}{2}}$$

$$\leq \ell_1 [P(\tilde{z})]^{\frac{1}{2}} / \{1 - (\hat{\sigma})^{\frac{1}{2}}\}$$

$$= \ell_2 [P(\tilde{z})]^{\frac{1}{2}},$$

where

$$\ell_2 = \ell_1 / [1 - (\hat{\sigma})^{\frac{1}{2}}]$$

This completes the proof.

2.8 Sequential Gradient–Restoration Algorithm

2.8.1 Gradient Phase

Let $z = (x,u,\pi) \in T_\epsilon$ be any nominal function for the gradient phase. Define

$$z + \Delta z = (x + \Delta x,\ u + \Delta u,\ \pi + \Delta \pi).$$

Then, the first variation of

$$\{J(z + \Delta z) - J(z)\}$$

can be expressed as:

$$\Delta J = \int_0^1 \{f_x \Delta x + f_u \Delta u + f_\pi \Delta \pi\}\, dt + [g_x \Delta x + g_\pi \Delta \pi]_0 + [h_x \Delta x + h_\pi \Delta \pi]_1 \tag{2.124}$$

If $\epsilon > 0$ is sufficiently small, then the differential equations (2.2) together with the boundary conditions (2.3) – (2.4) can be approximated up to the first order by

$$\Delta \dot{x} = \phi_x \Delta x + \phi_u \Delta u + \phi_\pi \Delta \pi \tag{2.125a}$$

$$[\omega_x \Delta x + \omega_\pi \Delta \pi]_0 = 0 \tag{2.125b}$$

$$[\psi_x \Delta x + \psi_\pi \Delta \pi]_1 = 0 \tag{2.125c}$$

To the first order, the minimum of J is achieved if the first variation (2.124) is minimized subject to (2.125a) – (2.125c). To make the problem meaningful, we want the variations $\Delta u(t)$, $\Delta\pi$ to satisfy the isoperimetric constraint

$$\int_0^1 |\Delta u|^2\, dt + |\Delta\pi|^2 = K \tag{2.126}$$

Let the above problem be referred to as $(2.\hat{P})$

Remark 2.8.1 Since system (2.125) is obtained from equations (2.2) – (2.4) by quasilinearization, it is clear that the constraint qualification for the problem $(2.\hat{P})$ is exactly the same as that for the problem $(2.P)$ (cf. Remark 2.6.1). Thus, by following exactly the same argument as that given in Section 2.5, we can show that if z satisfies the constraint qualification of the problem $(2.P)$ (i.e., $(M(z))^{-1}$ exists, where M is defined by (2.37)), then there exist $\Delta z = (\Delta x, \Delta u, \Delta\pi)$ and $\lambda(t) \in R^n$, $\mu \in R^r$, $\hat{\mu} \in R^s$, $\alpha \in R$ ($1/2\alpha$ is the Lagrange multiplier of the isoperimetric constraint (2.126)) such that (2.125) – (2.126) together with the following equations are satisifed.

$$\dot{\lambda}^T = f_x - \lambda^T \phi_x \tag{2.127a}$$

$$\Delta u^T/\alpha - \lambda^T \phi_u + f_u = 0 \tag{2.127b}$$

$$\int_0^1 (f_\pi - \lambda^T \phi_\pi)\, dt + [g_\pi + \mu^T \omega_\pi]_0 + [h_\pi + \hat{\mu}^T \psi_\pi]_1 + \Delta\pi^T/\alpha = 0 \tag{2.127c}$$

$$[-\lambda^T + g_x + \mu^T \omega_x]_0 = 0 \tag{2.127d}$$

$$[\lambda^T + h_x + \hat{\mu}^T \psi_x]_1 = 0 \tag{2.127e}$$

2.8.2 Coordinate Transformation

Let

$$A = \Delta x \,/\, \alpha \tag{2.128a}$$

$$B = \Delta u \,/\, \alpha \tag{2.128b}$$

$$C = \Delta \pi \,/\, \alpha \tag{2.128c}$$

Then, equations (2.125a) and (2.127a) – (2.127c) become:

$$\dot{A} = \phi_x A + \phi_u B + \phi_\pi C \tag{2.129a}$$

$$\dot{\lambda}^T = f_x - \lambda^T \phi_x \tag{2.129b}$$

$$B^T = -f_u + \lambda^T \phi_u \tag{2.129c}$$

$$C^T = \int_0^1 \{-f_\pi + \lambda^T \phi_\pi\}\, dt - [g_\pi + \mu^T \omega_\pi]_0 - [h_\pi + \hat{\mu}^T \psi_\pi]_1 \tag{2.129d}$$

and the boundary conditions (2.125b) – (2.125c) and (2.127d) – (2.127e) become:

$$[\omega_x A + \omega_\pi C]_0 = 0 \tag{2.130a}$$

$$[\psi_x A + \psi_\pi C]_1 = 0 \tag{2.130b}$$

$$[-\lambda^T + g_x + \mu^T \omega_x]_0 = 0 \tag{2.130c}$$

$$[\lambda^T + h_x + \hat{\mu}^T \psi_x]_1 = 0 \tag{2.130d}$$

Moreover, the isoperimetric constraint (2.126) becomes:

$$\alpha^2 \left[\int_0^1 B^T B dt + C^T C \right] = K \tag{2.131}$$

Since we do not know what value of K to use in (2.131), we shall consider α as the steplength in the equations

$$\Delta x(t) = \alpha A(t) \tag{2.132a}$$

$$\Delta u(t) = \alpha B(t) \tag{2.132b}$$

$$\Delta \pi = \alpha C \tag{2.132c}$$

in order to perform a line search on the objective function J. The method for determining the step–size α is given in Section 2.8.4.

2.8.3 Integration Technique

In this section, we shall describe a method for solving the gradient corrections $A(t)$, $B(t)$ and C.

From (2.129b), (2.130d) and (2.62), we obtain

$$\lambda(t) = -(N(1,t))^{\mathrm{T}} [([h_{\mathrm{x}}]_1)^{\mathrm{T}} + ([\psi_{\mathrm{x}}]_1)^{\mathrm{T}} \hat{\mu}] - \int_t^1 (N(\tau,t))^{\mathrm{T}} (f_{\mathrm{x}}(\tau))^{\mathrm{T}} \, d\tau \tag{2.133}$$

Using (2.133), we obtain from (2.129c), (2.129d) and (2.130c) that

$$B(t) = \not{k}_1(t)\, \hat{\mu} + \not{k}_2(t) \tag{2.134a}$$

$$C = \not{k}_3\, \mu + \not{k}_4\, \hat{\mu} + \not{k}_5 \tag{2.134b}$$

and

$$\not{k}_6\, \mu + \not{k}_7\, \hat{\mu} = \not{k}_8 \tag{2.135a}$$

$$\not{k}_1(t) = -(\phi_{\mathrm{u}}(t))^{\mathrm{T}} (N(1,t))^{\mathrm{T}} ([\psi_{\mathrm{x}}]_1)^{\mathrm{T}} \tag{2.135b}$$

$$\mathcal{k}_2(t) = -(f_u(t))^T - (\phi_u(t))^T [(N(1,t))^T ([h_x]_1)^T + \int_t^1 N(\tau,t) ((f_x(\tau))^T d\tau] \quad (2.135c)$$

$$\mathcal{k}_3 = -([\omega_\pi]_o)^T \quad (2.135d)$$

$$\mathcal{k}_4 = -([\psi_\pi]_1)^T - [\int_0^1 (\phi_\pi(t))^T (N(1,t))^T dt]([\psi_x]_1)^T \quad (2.135e)$$

$$\mathcal{k}_5 = -\int_0^1 (f_\pi(t))^T d\tau - [\int_0^1 (\phi_\pi(t))^T [(N(1,t))^T ([h_x]_1)^T + \int_t^1 N(t,\tau) (f_x(\tau))^T d\tau] dt - [(g_\pi)_o + (h_\pi)_1]^T \quad (2.135e)$$

$$\mathcal{k}_6 = ([\omega_x]_o)^T \quad (2.135f)$$

$$\mathcal{k}_7 = (N(1,0))^T ([\psi_x]_1)^T \quad (2.135g)$$

$$\mathcal{k}_8 = -([g_x]_o)^T - (N(1,0))^T ([h_x]_1)^T - \int_0^1 N(\tau,0) (f_x(\tau))^T d\tau \quad (2.135h)$$

Thus, from (2.129a), (2.134a) and (2.134b), we obtain

$$A(t) = N(t,0) A(0) + \mathcal{k}_9(t) \mu + \mathcal{k}_{10}(t) \hat{\mu} + \mathcal{k}_{11}(t) \quad (2.136)$$

where

$$\mathcal{k}_9(t) = \mathcal{k}_3 \int_0^t N(t,\tau) \phi_\pi(\tau) d\tau \quad (2.137a)$$

$$\mathcal{k}_{10}(t) = \int_0^t N(t,\tau) [\phi_u(\tau) \mathcal{k}_1(\tau) + \phi_\pi(\tau) \mathcal{k}_4] d\tau \quad (2.137b)$$

$$\mathcal{k}_{11}(t) = \int_0^t N(t,\tau) [\phi_u(\tau) \mathcal{k}_2(\tau) + \phi_\pi(\tau) \mathcal{k}_5] d\tau \quad (2.137c)$$

From (2.135a), (2.130a), (2.130b), (2.136) and (2.134b), we have

$$\mathscr{K}_6 \mu + \mathscr{K}_7 \hat{\mu} = \mathscr{K}_8 \tag{2.138a}$$

$$[\omega_x]_0 A(0) + [\omega_\pi \mathscr{K}_3]_0 \mu + [\omega_\pi \mathscr{K}_4]_0 \hat{\mu} = -[\omega_\pi \mathscr{K}_5]_0 \tag{2.138b}$$

$$[\psi_x]_1 N(1,0) A(0) + [\psi_x \mathscr{K}_9 + \psi_\pi \mathscr{K}_3]_1 \mu + [\psi_x \mathscr{K}_{10} + \psi_\pi \mathscr{K}_4]_1 \hat{\mu}$$
$$= -[\psi_x \mathscr{K}_{11} + \psi_\pi \mathscr{K}_5]_1 \tag{2.138c}$$

Remark 2.8.3.1 Similar to the system (2.69), the system (2.138) has a unique solution in $A(0)$, μ and $\hat{\mu}$ if and only if the nominal function z satisfies the constraint qualification (cf. Remark 2.8.1).

Remark 2.8.3.2 Note also that $A(t)$, $B(t)$ and C can be easily calculated from (2.136), (2.134a) and (2.134b), respectively, once $A(0)$, μ and $\hat{\mu}$ are determined from (2.138). Thus, $\Delta x(t)$, $\Delta u(t)$, $\Delta \pi$ can be obtained from (2.128a) – (2.128c) if α is known.

2.8.4 An Iteration of the Sequential Gradient–Restoration Algorithm

Let $\gamma_\ell \in (0, \frac{1}{2})$, $\beta \in (\frac{1}{2}, 1)$ be arbitrary but fixed. Then, we define

$$\Lambda(\epsilon) = \left\{(\beta)^\ell : (\beta)^\ell \geq \gamma_\ell \, \epsilon, \;\; \ell \in \{0, 1, 2, ...\}\right\}, \tag{2.139}$$

i.e., the set, $\Lambda(\epsilon)$, of step-size corrections contains all the values of the form 1, β, $(\beta)^2$, $(\beta)^3$,... which are equal to or greater than $\gamma_\ell \, \epsilon$.

We now define an iteration of the Sequential Gradient Restoration Algorithm as follows:

Algorithm 2.8.4

Step 0 Let γ, $\gamma_\ell \in (0, \frac{1}{2})$; $\beta \in (\frac{1}{2}, 1)$ be given constants.

Step 1 From a nominal function $z = (x, u, \pi) \in T_\epsilon \cap \mathscr{U}$, obtain $D(z) = (A(z), B(z), C(z)) \in \mathscr{U}$ together with $\lambda(z)$, $\mu(z)$, $\hat{\mu}(z)$ by solving system (2.129) – (2.130).

Step 2 Compute the step-size $\overline{\alpha}(z,\epsilon)$ as follows:

$$\overline{\alpha}(z,\epsilon) = \begin{cases} 0 & \text{if } \tilde{Q}(z) < \gamma_\ell \, \epsilon \\ 0 & \text{if } J(r(z+\alpha D(z), i(\epsilon, z+\alpha D(z)))) - J(z) > -\gamma\alpha\tilde{Q}(z) \\ & \quad \forall\ \alpha \in \Lambda(\epsilon) \\ \max\{\alpha \in \Lambda(\epsilon) : J(r(z+\alpha D(z), i(\epsilon, z+\alpha D(z)))) - J(z) \le -\gamma\alpha\tilde{Q}(z)\} \end{cases} \tag{2.140}$$

where

$$\tilde{Q}(z) = Q(z, \lambda(z), \dot{\lambda}(z), \mu(z), \hat{\mu}(z)) \tag{2.141}$$

and $i(\epsilon,z)$ denotes the minimum number of iterations of Algorithm 2.7.3 required for obtaining a function belonging to the set T_ϵ, starting from the nominal function z.

Step 3 Let

$$GR(z,\epsilon) = z + r(z + \overline{\alpha}(z,\epsilon)\, D(z),\ i(\epsilon, z + \overline{\alpha}(z,\epsilon)\, D(z))) \tag{2.142}$$

In order to establish some important properties concerning Algorithm 2.8.3 and to prove a convergence result for the sequential gradient–restoration algorithm given in the next section, we need the following assumptions:

Assumption (2.8.4A)

2.8.4A1 The constraint qualification is satisfied for any nominal function z in Step 1 of Algorithm 2.8.4.

2.8.4A2 Assumptions (2.7.3A1) and (2.7.3A2) are satisfied.

2.8.4A3 The set $\Delta_0 \cap T_0$ is non-empty, i.e., there exists a $z \in \mathcal{F}$ such that the necessary condition for optimality is satisfied.

Lemma 2.8.4.1 Suppose that Assumption (2.8.4A) is satisfied. Then, there exists a constant $\ell_3 > 0$ such that

$$\|D(z)\| \le \ell_3 \tag{2.143}$$

for any nominal function z defined in Step 1 of Algorithm 2.8.4.

Proof The proof can be easily obtained by using a similar approach as that given in the proof of Lemma 2.7.3.1.

Lemma 2.8.4.2 Suppose that Assumption (2.8.4A) is satisfied. Then, there exists a constant $\ell_4 > 0$ such that

$$\|D(z^2) - D(z^1)\| \le \ell_4 \|z^2 - z^1\| \tag{2.144}$$

and

$$|\tilde{Q}(z^2) - \tilde{Q}(z^1)| \le \ell_4 \|z^2 - z^1\| \tag{2.145}$$

for any nominal function z^1, z^2 defined in Step 1 of Algorithm 2.8.4.

Proof Let z^1, z^2 be any nominal functions defined in Step 1 of Algorithm 2.8.4. Using (2.62) and Assumption (2.4.A2), we get

$$\begin{aligned}
&|N(z^2)(t,\tau) - N(z^1)(t,\tau)| \\
&\le \int_\tau^t |\phi_x(z^2)(s)\, N(z^2)(s,\tau) - \phi_x(z^1)(s)\, N(z^1)(s,\tau)|\; ds \\
&\le \int_\tau^t \{|\phi_x(z^2)(s)\, N(z^2)(s,\tau) - \phi_x(z^2)(s)\, N(z^1)(s,\tau)| \\
&\qquad + |\phi_x(z^2)(s)\, N(z^1)(s,\tau) - \phi_x(z^1)(s)\, N(z^1)(s,\tau)|\}\; ds
\end{aligned}$$

$$\leq \int_\tau^t K_1 \, |N(z^2)(s,\tau) - N(z^1)(s,\tau)| \; ds$$

$$+ \int_\tau^t K_2 \, |N(z^1)(s,\tau)| \; \|z^2 - z^1\| \; ds \tag{2.146}$$

By using a similar approach as that used to obtain (2.79) from (2.62), it follows that

$$|N(z)(s,\tau)| \leq \exp(K_1) \tag{2.147}$$

for all nominal functions z defined in Step 1 of Algorithm 2.8.4.

Thus, from (2.146) and (2.147), we get

$$|N(z^2)(t,\tau) - N(z^1)(t,\tau)|$$

$$\leq \int_\tau^t K_1 \, |N(z^2)(s,\tau) - N(z^1)(s,\tau)| \; ds + K_2 \exp(K_1) \, \|z^2 - z^1\| \tag{2.148}$$

Applying Gronwall's lemma, it follows that

$$|N(z^2)(t,\tau) - N(z^1)(t,\tau)|$$

$$\leq K_2 \exp(K_1) \exp(K_1(t-\tau)) \, \|z^2 - z^1\|$$

$$\leq K_2 \, [\exp(K_1)]^2 \, \|z^2 - z^1\| \leq \hat{\mathcal{K}}_1 \, \|z^2 - z^1\| \tag{2.149}$$

where

$$\hat{\mathcal{K}}_1 = K_2 \, [\exp(K_1)]^2 \tag{2.150}$$

Using (2.149), Assumption (2.4.A2), and (2.147), and then following a similar approach as that used in obtaining (2.149) from (2.62), we deduce from (2.135) and (2.137) that there exists a constant $\hat{\mathcal{K}}_2 > 0$ such that for all $t \in [0,1]$, we have

$$|\mathscr{k}_j(z^2)(t) - \mathscr{k}_j(z^1)(t)| \le \hat{\mathscr{K}}_1 \|z^2 - z^1\|,\ j = 1,2,9,10,11 \tag{2.151}$$

and

$$|\mathscr{k}_j(z^2) - \mathscr{k}_j(z^1)| \le \mathscr{K}_2 \|z^2 - z^1\|,\ j = 3,4,5,6,7,8 \tag{2.152}$$

where $\mathscr{k}_j$, $j = 1,\ldots,11$, are as defined in (2.135) and (2.137).

From (2.138), Assumption (2.8.4.A1) and Remark 2.8.3.1, we have

$$\begin{bmatrix} [A\ (z^2)]_0 \\ \mu(z^2) \\ \hat{\mu}(z^2) \end{bmatrix} - \begin{bmatrix} [A\ (z^1)]_0 \\ \mu(z^1) \\ \hat{\mu}(z^1) \end{bmatrix}$$

$$= [\overline{M}\,(z^2)]^{-1}\, b\,(z^2) - [\overline{M}\,(z^1)]^{-1}\, b\,(z^1) \tag{2.153}$$

where

$$b\,(z) = \begin{bmatrix} \mathscr{k}_8\,(z) \\ -\,[(\omega_\pi(z)\ \mathscr{k}_5(z)]_0 \\ -\,[\psi_x(z)\ \mathscr{k}_{11}(z) + \psi_\pi(z)\ \mathscr{k}_5(z)]_1 \end{bmatrix} \tag{2.154}$$

and $\overline{M}$ is the coefficient matrix of the system (2.138).

Using Assumption (2.4.A2), (2.152), (2.151), (2.149) and (2.147), we deduce from (2.138) that there exists a constant $\hat{\mathscr{K}}_4 > 0$ such that

$$|b\,(z^2) - b\,(z^1)| \le \hat{\mathscr{K}}_4 \|z^2 - z^1\| \tag{2.155}$$

and

$$|\overline{M}_{ij}(z^2) - \overline{M}_{ij}(z^1)| \le \hat{\mathscr{K}}_4 \|z^2 - z^1\|,\ i,j = 1,\ldots,n+r+s \tag{2.156}$$

where $\overline{M}_{ij}$ is the coefficient of the matrix $\overline{M}$ of system (2.138). By Assumption (2.8.4A1) and Remark 2.8.3.1, it follows that there exists a constant $\hat{\mathscr{K}}_5 > 0$ such that

$$|\text{Det}\,[\overline{M}(z)]| \geq \hat{\mathscr{K}}_5 \tag{2.157}$$

for any nominal function z defined in Step 1 of Algorithm 2.8.4.

Thus, we deduce from (2.156) and (2.157) that there exists a constant $\hat{\mathscr{K}}_6 > 0$ such that

$$|(\overline{M}(z^2))^{-1} - (\overline{M}(z^1))^{-1}| \leq \hat{\mathscr{K}}_6 \,\|z^2 - z^1\| \tag{2.158}$$

Thus, from (2.153), (2.158) and (2.155), we can show that there exists a constant $\hat{\mathscr{K}}_7 > 0$ such that

$$|[A(z^2)]_0 - [A(z^1)]_0| \leq \hat{\mathscr{K}}_7 \,\|z^2 - z^1\| \tag{2.159a}$$

$$|\mu(z^2) - \mu(z^1)| \leq \hat{\mathscr{K}}_7 \,\|z^2 - z^1\| \tag{2.159b}$$

and

$$|\hat{\mu}(z^2) - \hat{\mu}(z^1)| \leq \hat{\mathscr{K}}_7 \,\|z^2 - z^1\| \tag{2.159c}$$

In view of (2.149), (2.151), (2.159) and (2.152), we deduce from the formulae of $A(t)$, $B(t)$ and C (defined in (2.136), (2.134a) and (2.134b), respectively) that there exists a constant $\hat{\mathscr{K}}_8 > 0$ such that

$$|[A(z^2)]_1 - [A(z^1)]_1| \leq \hat{\mathscr{K}}_8 \,\|z^2 - z^1\| \tag{2.160a}$$

$$\|A(z^2) - A(z^1)\|_\infty \leq \hat{\mathscr{K}}_8 \,\|z^2 - z^1\| \tag{2.160b}$$

$$\|B(z^2) - B(z^1)\|_\infty \le \hat{\mathscr{K}}_8 \|z^2 - z^1\| \tag{2.160c}$$

and

$$|C(z^2) - C(z^1)| \le \hat{\mathscr{K}}_8 \|z^2 - z^1\| \tag{2.160d}$$

Thus, from (2.129a), Assumption (2.4.A2) and (2.160b) – (2.160d), it follows that there exists a constant $\hat{\mathscr{K}}_9 > 0$ such that

$$\|\dot{A}(z^2) - \dot{A}(z^1)\|_\infty \le \hat{\mathscr{K}}_9 \|z^2 - z^1\| \tag{2.161}$$

Thus, from (2.7), (2.160b), (2.161), (2.160c), (2.160d), (2.159a) and (2.160a), we get

$$\begin{aligned} &\|D(z^2) - D(z^1)\| \\ &= \|A(z^2) - A(z^1)\|_\infty + \|\dot{A}(z^2) - \dot{A}(z^1)\|_\infty + \|B(z^2) - B(z^1)\|_\infty \\ &+ \|C(z^2) - C(z^1)\|_\infty + |[A(z^2)]_0 - [A(z^1)]_0| + |[A(z^2)]_1 - [A(z^1)]_1| \\ &\le \hat{\mathscr{K}}_{10} \|z^2 - z^1\| \end{aligned} \tag{2.162}$$

where

$$\hat{\mathscr{K}}_{10} = \hat{\mathscr{K}}_7 + 4\,\hat{\mathscr{K}}_8 + \hat{\mathscr{K}}_9 \tag{2.163}$$

From the definition of $\tilde{Q}(z)$ and $Q(z)$ (defined in (2.141) and (2.43), respectively), together with the definition of $\lambda(z)$, $\mu(z)$ and $\hat{\mu}(z)$ (obtained by solving the system (2.129) and (2.130)), it follows that

$$\begin{aligned} &\tilde{Q}(z^2) - \tilde{Q}(z^1) \\ &= \int_0^1 [|B(z^2)(t)|^2 - |B(z^1)(t)|^2]\,dt + |C(z^2)|^2 - |C(z^1)|^2 \end{aligned} \tag{2.164}$$

Thus, from (2.164), (2.160c), (2.160d), Lemma 2.8.4.1 and the definition of $D(z)$, we obtain

$$|\tilde{Q}(z^2) - \tilde{Q}(z^1)| \leq 2\, \ell_3\, \hat{\mathscr{K}}_8 \, \|z^2 - z^1\| \tag{2.165}$$

Let

$$\ell_4 = \max \{ \hat{\mathscr{K}}_{10}, 2\, \ell_3\, \hat{\mathscr{K}}_8 \} \tag{2.166}$$

Then, the rest of the proof follows easily from (2.162) and (2.165).

Remark 2.8.4.1 In view of Lemma 2.8.4.1, it is clear that for any $z \in \mathscr{U}$, the set

$\{z + \alpha D(z) : \alpha \in [0,1]\}$

is compact with respect to the norm defined by (2.7).

Lemma 2.8.4.3 Suppose that Assumption (2.8.4A) is satisfied. Then,

$$\lim_{\alpha \to 0} \frac{J(z + \alpha D(z)) - J(z)}{\alpha} = -\tilde{Q}(z) \tag{2.167}$$

Proof Define

$z = (x, u, \pi)$.

From (2.1), (2.128) and the mean value theorem, we obtain

$$\begin{aligned} & J(z + \alpha D(z)) - J(z) \\ & = \alpha \int_0^1 [\bar{L}_1(z)(t) + \bar{L}_2(z)(t) - \bar{L}_1(z)(t)]\, dt \\ & \qquad + \alpha [\bar{L}_3(z) + \bar{L}_4(z) - \bar{L}_3(z)], \end{aligned} \tag{2.169}$$

where

$$\overline{L}_1(z)(t) = f_x(t,x(t), u(t), \pi)\, A(z)(t) + f_u(t,x(t), u(t), \pi)\, B(z)(t) + f_\pi(t,x(t), u(t), \pi)\, C(z), \tag{2.170}$$

$$\begin{aligned}\overline{L}_2(z)(t) = {} & f_x(t,x(t) + \overline{\alpha}_1 A(z)(t), u(t) + \overline{\alpha}_1 B(z)(t), \pi + \overline{\alpha}_1 C(z))\, A(z)(t) \\ & + f_u(t,x(t) + \overline{\alpha}_1 A(z)(t), u(t) + \overline{\alpha}_1 B(z)(t), \pi + \overline{\alpha}_1 C(z))\, B(z)(t) \\ & + f_\pi(t,x(t) + \overline{\alpha}_1 A(z)(t), u(t) + \overline{\alpha}_1 B(z)(t), \pi + \overline{\alpha}_1 C(z))\, C(z),\end{aligned} \tag{2.171}$$

$$\overline{L}_3(z) = [g_x(x,\pi)\, A(z)]_0 + [g_\pi(x,\pi)\, C(z)]_0 + [h_x(x,\pi)\, A(z)]_1 + [h_\pi(x,\pi)\, C(z)]_1\ , \tag{2.172}$$

$$\begin{aligned}\overline{L}_4(z) = {} & [g_x(x + \overline{\alpha}_2 A(z)(t), \pi + \overline{\alpha}_2 C(z))\, A(z)]_0 \\ & + [g_\pi(x + \overline{\alpha}_2 A(z)(t), \pi + \overline{\alpha}_2 C(z))\, C(z)]_0 \\ & + [h_x(x + \overline{\alpha}_3 A(z)(t), \pi + \overline{\alpha}_3 C(z))\, A(z)]_1 \\ & + [h_\pi(x + \overline{\alpha}_3 A(z)(t), \pi + \overline{\alpha}_3 C(z))\, C(z)]_1\end{aligned} \tag{2.173}$$

and $\overline{\alpha}_i$, $i = 1,2,3$, are appropriate intermediate values satisfying

$$0 \leq \overline{\alpha}_i \leq \alpha. \tag{2.174}$$

From (2.170), (2.171), Assumption (2.3.A2), (2.143), and (2.174), we deduce that

$$|\overline{L}_1(z)(t) - \overline{L}_2(z)(t)| \leq K_2\, \overline{\alpha}_1\, \ell_3^2 \leq K_2\, \alpha\, \ell_3^2 \tag{2.175}$$

Similarly, by (2.172), (2.173), Assumption (2.3.A2), we obtain

$$|\overline{L}_3(z)(t) - \overline{L}_4(z)(t)| \leq K_2\, \ell_3^2\, (\overline{\alpha}_2 + \overline{\alpha}_3) \leq 2\, K_2\, \alpha\, \ell_3^2 \tag{2.176}$$

Now, combining (2.169), (2.175), (2.176), it follows that

$$\lim_{\alpha \to 0} \frac{J(z + \alpha D(z)) - J(z)}{\alpha} = \int_0^1 \bar{L}_1(z)(t)\, dt + \bar{L}_3(z), \tag{2.177}$$

From (2.170), (2.129b) – (2.129d), (2.129a), (2.130d), (2.130c), (2.172), (2.130b) and (2.130a), we can show that

$$\begin{aligned}
&\int_0^1 \bar{L}_1(z)(t) \\
&\quad = \int_0^1 (\dot{\lambda}(z)(t))^T A(z)(t) + (\lambda(z)(t))^T \\
&\qquad \{\phi_x(t, x(t), u(t), \pi)\, A(z)(t) + \phi_u(t, x(t), u(t), \pi)\, B(z)(t) \\
&\qquad\quad + \phi_\pi(t, x(t), u(t), \pi)\, C(z)\}\, dt \\
&\quad - \int_0^1 (B(z)(t))^T B(z)(t)\, dt - (C(z))^T C(z) \\
&\quad - [\{g_\pi(x,\pi) + (\mu(z))^T \omega_\pi(x,\pi)\}\, C(z)]_0 \\
&\quad - [\{h_\pi(x,\pi) + (\hat{\mu}(z))^T \psi_\pi(x,\pi)\}\, C(z)]_1 \\
&\quad = [(\lambda(z))^T A(z)]_1 - [(\lambda(z))^T A(z)]_0 \\
&\quad - \int_0^1 (B(z)(t))^T B(z)(t)\, dt - (C(z))^T C(z) \\
&\quad - [\{g_\pi(x,\pi) + (\mu(z))^T \omega_\pi(x,\pi)\}\, C(z)]_0 \\
&\quad - [\{h_\pi(x,\pi) + (\hat{\mu}(z))^T \psi_\pi(x,\pi)\}\, C(z)]_1 \\
&\quad = -[\{h_x(x,\pi) + (\hat{\mu}(z))^T \psi_x(x,\pi)\}\, A(z)]_1 \\
&\quad - [\{g_x(x,\pi) + (\mu(z))^T \omega_x(x,\pi)\}\, A(z)]_0 \\
&\quad - \int_0^1 (B(z)(t))^T B(z)(t)\, dt - (C(z))^T C(z) \\
&\quad - [\{g_\pi(x,\pi) + (\mu(z))^T \omega_\pi(x,\pi)\}\, C(z)]_0 \\
&\quad - [\{h_\pi(x,\pi) + (\hat{\mu}(z))^T \psi_\pi(x,\pi)\}\, C(z)]_1 \\
&\quad = -\bar{L}_3(z) - \int_0^1 (B(z)(t))^T B(z)(t)\, dt - (C(z))^T C(z)
\end{aligned}$$

(2.178)

Of the other hand, from the definitions of $\tilde{Q}(z)$ and $Q(z)$ (defined in (2.141) and (2.43), respectively) together with the definitions of $\lambda(z)$, $\mu(z)$ and $\hat{\mu}(z)$ (obtained by solving the system (2.129) and (2.130)), it follows that

$$\tilde{Q}(z) = \int_0^1 (B(z)(t))^T B(z)(t)\, dt + (C(z))^T C(z) \tag{2.179}$$

Combining (2.177), (2.178) and (2.179), the conclusion of the lemma follows readily

Definition 2.8.4 For any $\bar{z} \in \mathscr{U}$ and $\epsilon > 0$, let

$$B(\bar{z},\epsilon) := \{z \in \mathscr{U} : \|\bar{z} - z\| \le \epsilon\} \tag{2.180}$$

Lemma 2.8.4.4 Let Assumption (2.8.4A) be satisfied, and let $\bar{z} \in T_0$ be such that $\bar{z} \notin \Delta_0$. Then, there exists a constant $\rho^{\bar{z}} > 0$ and an integer $\ell^{\bar{z}} > 0$ such that for any $z \in B(\bar{z},\rho^{\bar{z}})$ and for all $i \in \{0,1,2,...\}$

$$J(r(z + (\beta)^{\ell^{\bar{z}}} D(z)), i) - J(z) \le -\gamma(\beta)^{\ell^{\bar{z}}} \tilde{Q}(z) \tag{2.181}$$

Proof For each $z \in \mathscr{U}$, let

$$z + \alpha D(z) = \hat{z} = (\hat{x},\hat{u},\hat{\pi}) \tag{2.182}$$

and

$$r(z + \alpha D(z), i) - [z + \alpha D(z)] = \hat{D} = (\hat{A},\hat{B},\hat{C}) \tag{2.183}$$

Thus, by (2.1), the mean value theorem, (2.4.A2), definition of $\|\hat{D}\|$, (2.183), and Theorem 2.7.3.3, it follows that

$$
\begin{aligned}
& J(r(z + \alpha D(z)), \hat{\pi}) - J(z + \alpha D(z)) \\
&= \int_0^1 \{f_x(\hat{z} + \alpha_1 \hat{D})(t)\, \hat{A}(t) + f_u(\hat{z} + \alpha_1 \hat{D})(t)\, \hat{B}(t) + f_\pi(\hat{z} + \alpha_1 \hat{D})(t)\, \hat{C}\}\, dt \\
&+ [\{g_x(\hat{x} + \alpha_2 \hat{A}, \hat{\pi} + \alpha_2 \hat{C})\, \hat{A} - \{g_\pi(\hat{x} + \alpha_2 \hat{A}, \hat{\pi} + \alpha_2 \hat{C})\, \hat{C}]_0 \\
&+ [\{h_x(\hat{x} + \alpha_3 \hat{A}, \hat{\pi} + \alpha_3 \hat{C})\, \hat{A} - \{h_\pi(\hat{x} + \alpha_3 \hat{A}, \hat{\pi} + \alpha_3 \hat{C})\, \hat{C}]_1 \\
&\leq 2\, K_1 \|\hat{D}\| \leq 2\, K_1\, \ell_2\, [P(z + \alpha D(z))]^{\frac{1}{2}}
\end{aligned}
\tag{2.184}
$$

where $\hat{\alpha}_i$, $i = 1,2,3$ are intermediate values satisfying

$$
0 \leq \hat{\alpha}_i \leq 1 \tag{2.185}
$$

Now, from (2.169), (2.178), (2.179), (2.175) and (2.176), it is clear that

$$
\left| \frac{J(z + \alpha D(z)) - J(z)}{\alpha} + \tilde{Q}(z) \right| \leq 3\, K_2\, \alpha\, \ell_3^2 \tag{2.186}
$$

for all $z \in \mathscr{U}$ and for all $\alpha \in [0,1]$.

Thus,

$$
\left| \frac{J(z + \alpha D(z)) - J(z)}{\alpha} + \tilde{Q}(z) \right| \leq 0.01\, \tilde{Q}(\bar{z}) \tag{2.187}
$$

for all $z \in \mathscr{U}$ and for all $\alpha \in [0, \alpha_4]$, where

$$
\alpha_4 = \max \left\{ \frac{0.01\, \tilde{Q}(\bar{z})}{3\, K_2\, \ell_3^2}, 1 \right\} \tag{2.188}
$$

Combining (2.184) and (2.185), we have

$$J(r(z+\alpha D(z)), i) - J(z) + \gamma \alpha \tilde{Q}(z)$$
$$\leq 2 K_1 \ell_2 \sqrt{P(z + \alpha D(z))} - \alpha(1-\gamma) \tilde{Q}(z)$$
$$+ 0.01 \alpha \tilde{Q}(\bar{z}) \qquad (2.189)$$

for all $z \in \mathscr{U}$ and for all $\alpha \in [0, \alpha_4]$.

In view of (2.189) and Lemma 2.8.4.2, we have

$$J(r(z+\alpha D(z)), i) - J(z) + \gamma \alpha \tilde{Q}(z)$$
$$\leq 2 K_1 \ell_2 \sqrt{P(z + \alpha D(z))} - \alpha(0.99-\gamma) \tilde{Q}(z)$$
$$+ 0.01 \alpha \ell_4 \|z - \tilde{\bar{z}}\| \qquad (2.190)$$

for all $z \in \mathscr{U}$ and for all $\alpha \in [0, \alpha_4]$.

Define

$$\theta(z,\alpha) = 2 K_1 \ell_2 \sqrt{P(z + \alpha D(z))} - \alpha(0.99-\gamma) \tilde{Q}(z)$$
$$+ 0.01 \alpha \ell_4 \|z - \tilde{\bar{z}}\| \qquad (2.191)$$

Now, in view of the fact that $\bar{z} \in T_0$ satisfy the system (2.2) – (2.4) exactly, the mean value theorem and (2.129a), we get

$$\phi(t, \bar{x}(t) + \alpha A(\bar{z})(t), \bar{u}(t) + \alpha B(\bar{z})(t), \bar{\pi} + \alpha C(\bar{z})) - [\dot{\bar{x}}(t) + \alpha \dot{A}(\bar{z})(t)]$$
$$= [\phi(t, \bar{x}(t) + \alpha A(\bar{z})(t), \bar{u}(t) + \alpha B(\bar{z})(t), \bar{\pi} + \alpha C(\bar{z}))$$
$$- \phi(t, \bar{x}(t), \bar{u}(t), \bar{\pi})] - \alpha \dot{A}(\bar{z})(t)$$

$$= \frac{1}{2}\alpha^2 \times \Big[\sum_{j=1}^{n} \sum_{i=1}^{n} \frac{\partial^2 \phi(x(t))}{\partial x_i \, \partial x_j} \times A_i(\overline{z})(t)$$

$$\times A_j(\overline{z})(t) + \sum_{j=1}^{m} \sum_{i=1}^{m} \frac{\partial^2 \phi(x(t))}{\partial u_i \, \partial u_j} \times B_i(\overline{z})(t)$$

$$\times B_j(\overline{z})(t) + \sum_{j=1}^{p} \sum_{i=1}^{p} \frac{\partial^2 \phi(x(t))}{\partial \pi_i \, \partial \pi_j} \times C_i(\overline{z}) \cdot C_j(\overline{z})$$

$$+ \sum_{j=1}^{m} \sum_{i=1}^{n} 2 \times \frac{\partial^2 \phi(x(t))}{\partial x_i \, \partial u_j} \times A_i(\overline{z})(t) \times B_j(\overline{z})(t)$$

$$+ \sum_{j=1}^{p} \sum_{i=1}^{n} 2 \times \frac{\partial^2 \phi(x(t))}{\partial x_i \, \partial \pi_j} \times A_i(\overline{z})(t) \times C_j(\overline{z})$$

$$+ \sum_{j=1}^{p} \sum_{i=1}^{m} 2 \times \frac{\partial^2 \phi(x(t))}{\partial u_i \, \partial \pi_j} \times B_i(\overline{z})(t) \times C_j(\overline{z}) \Big], \tag{2.192}$$

where

$$x(t) = (\overline{x}(t) + \tilde{\alpha}\, A(\overline{z})(t), \overline{u}(t) + \tilde{\alpha}\, B(\overline{z})(t), \overline{\pi} + \tilde{\alpha}\, C(\overline{z}))$$
$$= \overline{z}(t) + \tilde{\alpha}\, D(\overline{z})(t) \tag{2.193}$$

and $\tilde{\alpha}$ is an intermediate value satisfying

$$0 \leq \tilde{\alpha} \leq \alpha \tag{2.194}$$

Hence, from (2.192), Assumption (2.4.A3), Remark 2.8.4.1 and Lemma 2.7.4.1, it is clear that there exists a constant $K_1(\overline{z}) > 0$ such that

$$|\phi(t, \overline{x}(t) + \alpha A(\overline{z})(t), \overline{u}(t) + \alpha B(\overline{z})(t), \overline{\pi} + \alpha C(\overline{z}))$$
$$- [\dot{\overline{x}}(t) + \alpha \dot{A}(\overline{z})(t)]| \leq K_1(\overline{z}) \times \alpha^2 \tag{2.195}$$

for almost all $t \in [0,1]$.

Now, by using a similar argument as that used to obtain (2.195), with the exception that (2.129a) is now being replaced by both (2.130a) and (2.130b), we can readily show that there exists a constant $K_2(\overline{z}) > 0$ such that

$$|[\omega(\overline{x} + \alpha A(\overline{z}), \overline{\pi} + \alpha C(\overline{z}))]_0| \leq K_2(\overline{z}) \times \alpha^2 \tag{2.196}$$

and

$$|[\psi(\overline{x} + \alpha A(\overline{z}), \overline{\pi} + \alpha C(\overline{z}))]_1| \leq K_2(\overline{z}) \times \alpha^2 \tag{2.197}$$

Thus, from (2.42) and (2.195) – (2.197), it is clear that

$$\sqrt{P(\overline{z}(t) + \alpha D(\overline{z})(t))} \leq K(\overline{z}) \times \alpha^2, \tag{2.198}$$

where

$$K(\overline{z}) = [K_1^2(\overline{z}) + 2 K_2^2(\overline{z})]^{\frac{1}{2}} \tag{2.199}$$

Let

$$\alpha_5 = \min\left[\alpha_4, \frac{(0.99 - \gamma)\ \tilde{Q}(\overline{z})}{K(\overline{z}) \times 4\ K_1\ \ell_2}\right] \tag{2.200}$$

Thus, from (2.191), (2.199) and (2.200), it is clear that

$$\theta(\overline{z},\alpha) \leq -\frac{\alpha}{2}(0.99 - \gamma)\ \tilde{Q}(\overline{z}) \tag{2.201}$$

for all $\alpha \in [0,\alpha_5]$.

In view of Remark 2.6.1 and Lemma 2.8.4.2, it follows that for each $\alpha \geq 0$, $\theta(\cdot,\alpha)$ is continuous with respect to the norm defined by (2.7). Now, let $\alpha = \beta^{i^{\bar{z}}}$, where $i^{\bar{z}}$ is a positive integer such that $\beta^{i^{\bar{z}}} < \alpha_5$. In view of the above fact together with (2.201), it is clear that there exists a positive constant $\rho^{\bar{z}}$ such that for all $z \in B(\bar{z},\rho^{\bar{z}})$, we have

$$\theta(z, \beta^{i^{\bar{z}}}) - \theta(\bar{z}, \beta^{i^{\bar{z}}}) \leq \frac{\beta^{i^{\bar{z}}}}{4}(0.99-\gamma)\,\tilde{Q}(\bar{z})$$

$$\Longrightarrow \theta(z, \beta^{i^{\bar{z}}}) \leq -\frac{\beta^{i^{\bar{z}}}}{4}(0.99-\gamma)\,\tilde{Q}(\bar{z}) \leq 0 \qquad (2.202)$$

Thus from (2.190), (2.191) and (2.202), we get

$$J(r(z+\beta^{i^{\bar{z}}}D(z)), i) - J(z) + \gamma\beta^{i^{\bar{z}}}\tilde{Q}(z)$$

$$\leq \theta(z, \beta^{i^{\bar{z}}})$$

$$\leq 0 \qquad (2.203)$$

for all $z \in B(\bar{z}, \rho^{\bar{z}})$.

This completes the proof of this lemma.

Theorem 2.8.4 Let Assumptions (2.8.4A) be satisfied, and let $\bar{z} \in T_0$ be such that $\bar{z} \notin \Delta_0$. Then, there exists $\rho(\bar{z}) > 0$, $\epsilon(\bar{z}) > 0$ and $\delta(\bar{z}) < 0$ such that for all $\epsilon \in [0,\epsilon(\bar{z})]$,

$$J(GR(z,\epsilon)) - J(z) \leq \delta(\bar{z}),\ \forall\, z \in B(\bar{z},\rho(\bar{z})) \cap T_\epsilon \qquad (2.204)$$

where $GR(z,\epsilon)$ is the function generated from $z \in T_\epsilon$ by Algorithm 2.8.4.

Proof The proof is exactly the same as that given for Proposition 4 of Ref. 65, except that Assumption 2(i) and Assumption 2(iii) in the proof of Proposition 4 of Ref. 65, are being replaced by Lemma 2.8.4.2 and Lemma 2.8.4.4.

2.8.5 The Main Algorithm

In this subsection, we present a computational algorithm, called the Sequential Gradient Restoration Algorithm (SGRA), for solving the problem (2.P). Its convergence properties will also be investigated. This algorithm has two main features. Firstly, sufficient amount of decrease in the objective function J within each combined minimization – restoration step is guaranteed in order to ensure convergence of the algorithm. Secondly, in order to prevent the disadvantage of requiring a huge amount of iterations in the restoration algorithm (Algorithm 2.7.3), we develop a scheme in which the value of ϵ is reduced as the algorithm approaches the optimal solution. Thus, at points far from the optimum, we require only a rough estimate of feasibility, whereas by approaching the optimal solution, the feasibility requirements are gradually tightened.

Algorithm 2.8.5 (SGRA)

Step 0 Let $\bar{\gamma} \in (0, \frac{1}{2})$, $\bar{\beta} \in (\frac{1}{2}, 1)$, $\epsilon_0 > 0$ be given constants. Choose a function $z^0 \in \mathcal{U}$. Set $k = 0$, $\ell = 0$, $\epsilon = \epsilon_0$.

Step 1 Use Algorithm 2.7.3 to generate a function $r(z^k, i(\epsilon, z^k)) \in T_\epsilon$. Set

$$z = r(z^k, i(\epsilon, z^k)) \qquad (2.205)$$

Step 2 Use Algorithm 2.8.4 to compute $GR(z, \epsilon)$.

Step 3 If

$$J(GR(z,\epsilon)) - J(z) \leq -\bar{\gamma}\,\epsilon, \qquad (2.206)$$

go to Step 4; else set $y^\ell = z$, $\epsilon = \bar{\beta}\,\epsilon$, $\ell = \ell + 1$ and go to Step 1.

Step 4 Set $z^{k+1} = GR(z, \epsilon)$, $k = k + 1$ and go to Step 1.

Theorem 2.8.5 Let Assumption (2.8.4A) be satisfied. Suppose that the sequence $\{z^k\}$ generated by Algorithm 2.8.5 is contained in a compact set (with respect to the norm defined by (2.7)). Then, one of the following two terminations can occur:

(i) If Algorithm 2.8.5 jams at a particular point $\hat{z}$ and generates an infinite sequence $\{y^\ell\}_{\ell=0}^{\infty}$, then any accumulation point y^* (with respect to the norm defined by (2.7)) satisfies

$$y^* \in T_0 \cap \Delta_0 \tag{2.207}$$

(ii) If Algorithm 2.8.5 generates an infinite sequence $\{z^k\}$, then any accumulation point z^* of this sequence (with respect to the norm defined by (2.7)) satisfies

$$z^* \in T_0 \cap \Delta_0 \tag{2.208}$$

Proof The proof of (i) follows easily from Theorem 2.8.4 and Theorem 2.6.3. The proof of the fact that $z^* \in T_0$ in part (ii) is exactly the same as that given for Proposition 2 of Ref. 65. The proof of the fact that $z^* \in \Delta_0$ is exactly the same as that given for Proposition 3 of Ref. 65, except that Assumption (2.3.3) in the proof of Proposition 3 of Ref. 65 is being replaced by Theorem 2.8.4.

2.9 An Illustrative Example

Consider the problem (cf. Ref. 35, pp. 245–247) of minimizing the cost functional

$$J = -x_2(1) \tag{2.209}$$

subject to the differential constraints

$$\dot{x}_1 = -(u + \beta u^2)\, x_1 \tag{2.210a}$$

$$\dot{x}_2 = u x_1 \tag{2.210b}$$

and the boundary conditions

$$\gamma\, x_1(1) + (1-\gamma) = x_1(0) \tag{2.211a}$$

$$\gamma\, x_2(1) = x_2(0) \tag{2.211b}$$

By letting $x_1(0) = \pi_1$ and $x_2(0) = \pi_2$, (2.211a) and (2.211b) are transformed to

$$x_1(0) - \pi_1 = 0 \tag{2.212a}$$

$$x_2(0) - \pi_2 = 0 \tag{2.212b}$$

$$\gamma\, x_1(1) + (1-\gamma) - \pi_1 = 0 \tag{2.212c}$$

$$\gamma\, x_2(1) - \pi_2 = 0, \tag{2.212d}$$

Thus, this problem belongs to the same class of problem as that described in Section 2.3.

Choose $\gamma = 0.1$ and $\beta = 0.5$. Furthermore, let

$$z^0 = (x^0, u^0, \pi^0),$$

where

$$x^0(t) = \begin{bmatrix} 1 \\ 1 \end{bmatrix},\ u^0(t) = 0,\ \pi^0 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} \tag{2.213}$$

as the nominal function. Then, Algorithm 2.8.5 is employed to solve the problem iteratively. From the computed result tabulated in Table 2.9.1 it is clear that at the 40th iteration, both $P(z)$ and $\tilde{Q}(z)$ are extremely close to zero. The optimal control is plotted in Figure 2.9.1a. The optimal states are plotted in Figure 2.9.1b. The optimal parameters are $\pi_1 = 0.05$ and $\pi_2 = 0.91$.

No. of iteration	$J(z^k)$	$P(z^k)$	$\tilde{Q}(z^k)$
0	–0.5350	6.48×10^{-6}	0.91046
2	–0.5516	6.26×10^{-5}	0.00701
4	–0.5573	8.54×10^{-5}	0.00323
6	–0.5610	9.38×10^{-5}	0.00106
8	–0.5644	1.53×10^{-5}	0.00121
10	–0.5662	3.98×10^{-5}	0.00085
12	–0.5675	2.26×10^{-5}	0.00063
14	–0.5685	1.78×10^{-5}	0.00048
16	–0.5692	1.59×10^{-5}	0.00038
18	–0.5697	1.55×10^{-5}	0.00031
20	–0.5703	2.00×10^{-5}	0.00026
22	–0.5707	1.64×10^{-5}	0.00022
24	–0.5710	1.61×10^{-5}	0.00018
26	–0.5714	1.60×10^{-5}	0.00016
28	–0.5716	1.56×10^{-5}	0.00014
30	–0.5719	1.60×10^{-5}	0.00012
32	–0.5721	2.23×10^{-5}	0.00011
34	–0.5722	1.95×10^{-5}	0.00010
36	–0.5723	3.04×10^{-5}	0.00009
38	–0.5725	1.83×10^{-5}	0.00008
40	–0.5727	1.96×10^{-5}	0.00007

Table 2.9.1 The Numerical Results for Example 2.9.1

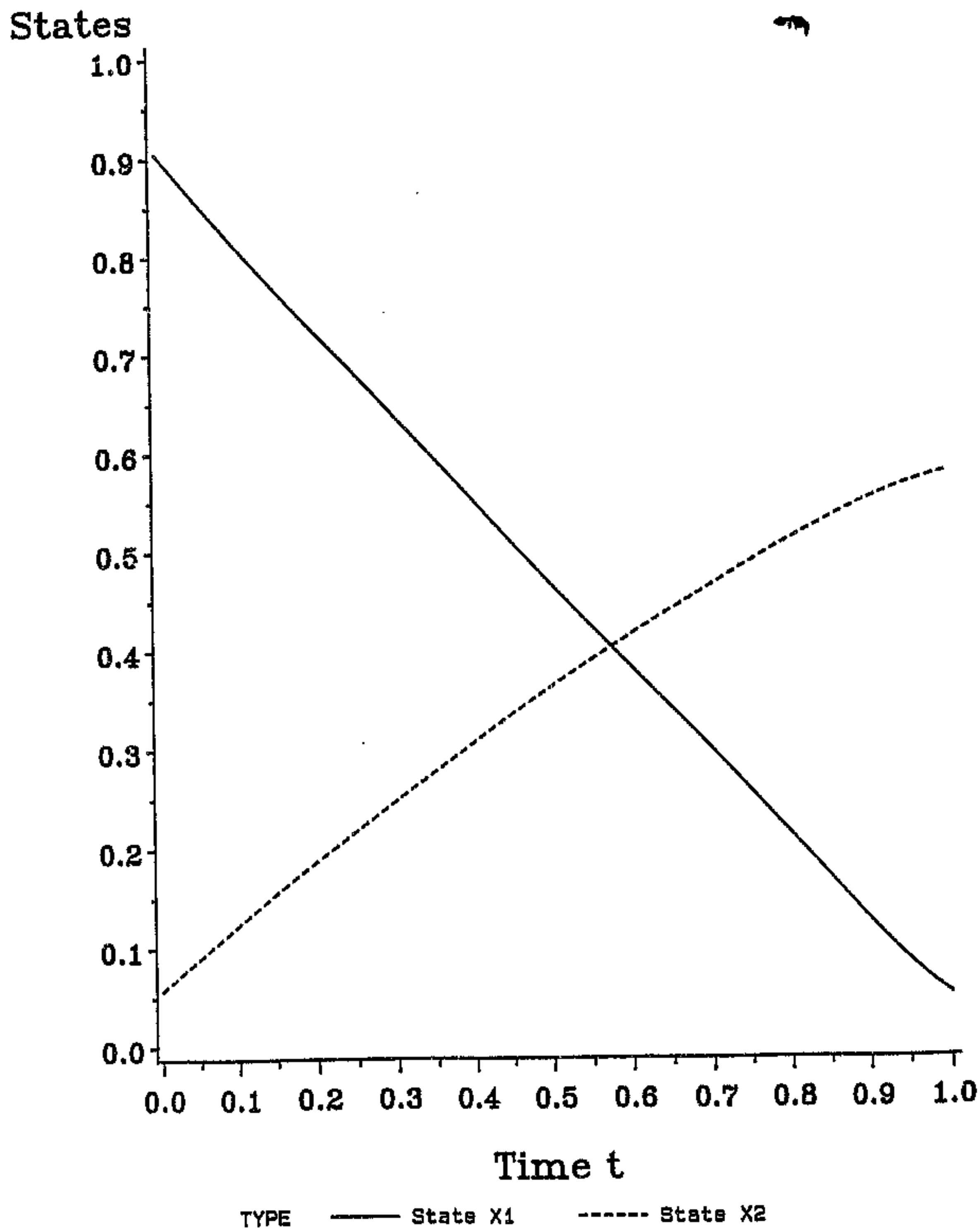

Figure 2.9.1a Computed Optimal States for Example 2.9.1

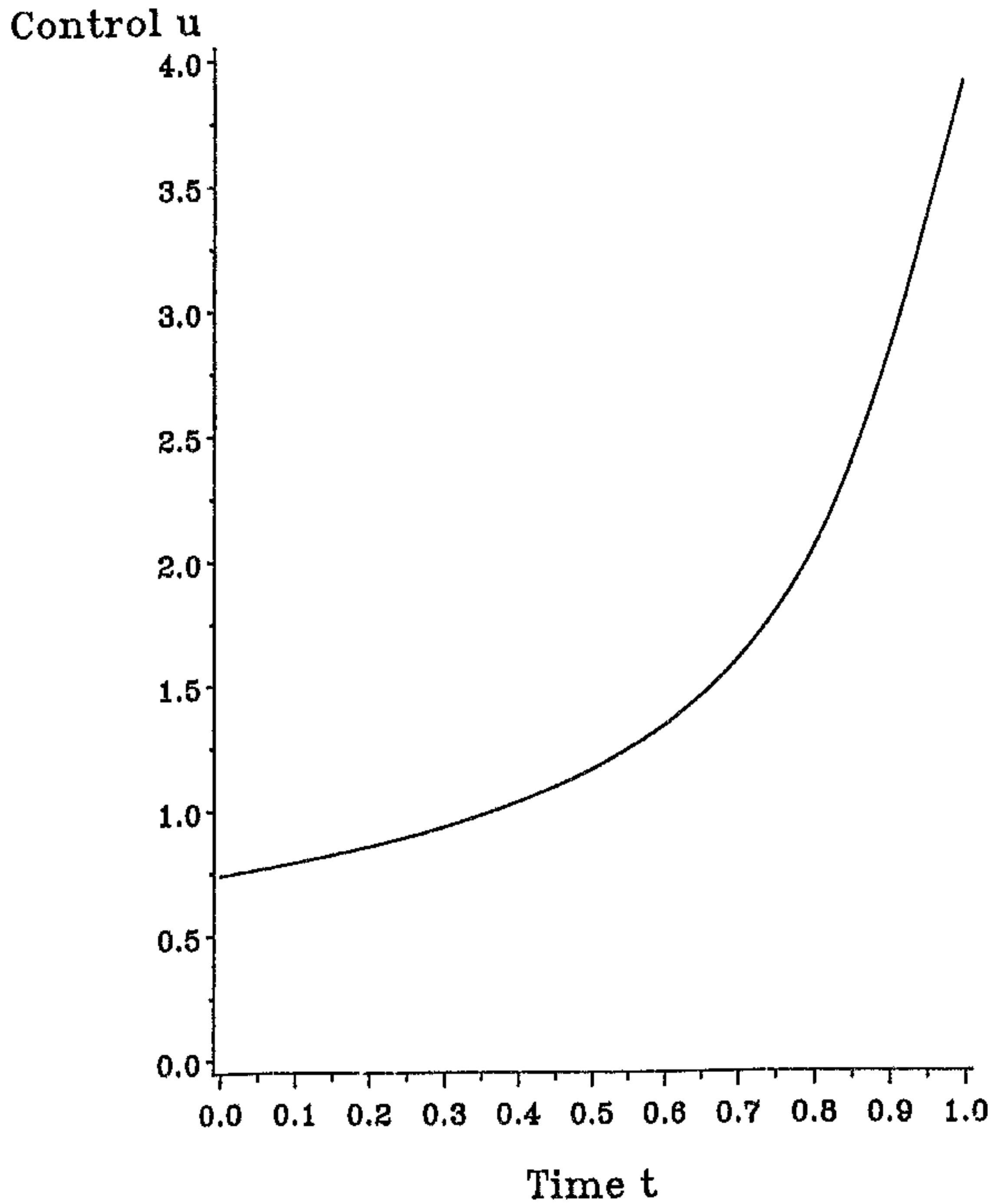

Figure 2.9.1b Computed Optimal Control for Example 2.9.1.

CHAPTER III

A COMPUTATIONAL METHOD FOR OPTIMAL CONTROL PROBLEM WITH TERMINAL AND INTERIOR POINTS CONSTRAINTS

3.1 Introduction

The class of optimal control problems considered in this Chapter involves terminal equality constraints and interior points equality constraints. In this chapter, we developed a new innovative computational scheme using the technique of control parameterization and Liapunov concepts for solving this class of optimal control problems.

In Section 3.2, we describe the optimal control problem $(3.P_1)$ and assume certain conditions.

In Section 3.3, using the technique of control parameterization, the problem $(3.P_1)$ is converted into an optimal parameter selection problem $(3.P_2)$.

In Section 3.4, using the technique of Liapunov concepts which was originally suggested in Ref. 74, the problem $(3.P_2)$ can be converted into an unconstrained problem in Lagrangian form $(3.P_3)$. Based on the fact that the stationary points of the Lagrangian function can be found by setting all the derivations equal to zero, we define a suitable positive – definite Liapunov Function (L) with terms consisting of all the derivatives of the Lagrangian function. Finally, the problem $(3.P_4)$ of finding the minimum of (L) can be easily solved by any existing unconstrained nonlinear programming technique, such as the conjugate gradient method.

However, to find the gradients of the function (L), we need to calculate both the gradients and Hessian of the objective function and the constraints of the problem $(3.P_2)$. The formula for these gradients and Hessians are derived in Section 3.5 and Section 3.6 respectively.

Finally, two examples are solved numerically in Section 3.7.

3.2 Statement of the Problem

Consider the following differential equation on the fixed time interval [0,1]

$$\dot{x}(t) = f(t, x(t), u(t), \pi), \tag{3.1a}$$

where $x \in R^n$, $u \in R^r$, $\pi \in R^p$ are respectively, the state, the control and the parameter vector; $f \equiv [f_1, \ldots, f_n]^T \in R^n$; and the superscript T denotes the transpose. The initial function for the differential equation (3.1a) is

$$x(0) = x^0, \tag{3.1b}$$

where x^0 is a given vector in R^n. The terminal and interior point conditions are

$$g_k(x(t_k), \pi) = 0, \qquad k = 1, \ldots, q, \tag{3.2}$$

where $0 < t_k \leq 1$ with at least one t_k equal to 1 to constitute the terminal constraint.

Let $\mathcal{U}$ be the class of all admissible controls defined by
$\mathcal{U} = \{u : [u_1, \ldots, u_r]^T$ is a bounded measurable function from [0,1] into $R^r\}$.

Define

$$\mathscr{V} = \{z = (u,\pi) : u \in \mathscr{U}, \pi \in R^p\}. \tag{3.2d}$$

For each $z \in \mathscr{V}$, let $x(\cdot|z)$ be the corresponding vector–valued function which is absolutely continuous on (0,1] and satisfies the differential equation (3.1a) almost everywhere on (0,1], together with the initial condition (3.1b). The function is called the solution of the system (3.1) corresponding to $z \in \mathscr{V}$. Let $\mathscr{F}$ be the set of all feasible z defined by

$$\mathscr{F} = \{z \in \mathscr{V} : g_k(x(t_k|z), \pi) = 0, \quad k = 1,\ldots,q\}$$

We may now state our optimal control problem as follows:

Problem (3.P_1) Subject to system (3.1), find a $z \in \mathscr{F}$ such that the cost functional

$$J(z) = \phi(x(1|z), \pi) + \int_0^1 f_0(t, x(t|z), u(t), \pi)\, dt \tag{3.3}$$

is minimized over $\mathscr{F}$, where ϕ and f_0 are given real–valued functions.

Let L_∞^s be the Banach space as defined in Section 2.4.

We assume throughout that the following conditions are satisfied.

(3.A1)
$f : [0,1] \times R^n \times R^r \times R^p \longrightarrow R^n$
$\phi : R^n \times R^p \longrightarrow R$
$f_0 : [0,1] \times R^n \times R^r \times R^p \longrightarrow R,$
$g_k : R^n \times R^p \longrightarrow R, \qquad k = 1,\ldots,q$

(3.A2) There exist positive constants K_1 and K_2 such that

$$|F_\beta(t,x^1,u^1,\pi^1)| \leq K_1$$
$$|G_\beta(x^1,\pi^1)| \leq K_1$$

and

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

for any $t \in [0,1]$, x^1, $x^2 \in R^n$, u^1, $u^2 \in R^r$ and π^1, $\pi^2 \in R^p$, where F (respectively G) denot any of the functions f and f_0 (respectively, ϕ and g_k, $k = 1,\ldots,q$); F_β (respectively G_β) denotes the gradient of F (respectively G) with respect to β. Here β represents any of the vectors x, u and π.

(3.A3) There exists positive constant K_3 such that

$$\frac{\partial^2 f_i(t,x,u,\pi)}{\partial\alpha_1\,\partial\alpha_2} \leq K_3, \qquad i = 0,\ldots,n, \tag{3.4}$$

$$\frac{\partial^2 \phi(x,\pi)}{\partial\alpha_1\,\partial\alpha_2} \leq K_3 \tag{3.5}$$

and

$$\frac{\partial^2 g_i(x,\pi)}{\partial\alpha_1\,\partial\alpha_2} \leq K_3, \qquad i = 1,\ldots,q, \tag{3.6}$$

for any $t \in [0,1]$, $x \in R^n$, $u \in R^r$ and $\pi \in R^p$. Here α_1 and α_2 represents any of the variables $x_1,\ldots,x_n$, $u_1,\ldots,u_r$, $\pi_1,\ldots,\pi_p$.

3.3 Control Parameterization

The control parameterization method will be used to convert the problem $(3.P_1)$ into an optimal parameter selection problem $(3.P_2)$.

Let $\mathcal{I}$ be a partition of the interval [0,1] such that $\mathcal{I}$ consists of m elements defined by

$$\mathcal{I} = \{I_j\}_{j=1}^{m}$$

where

$$I_j = [t_{j-1}, t_j),$$

and

$$t_{j+1} - t_j = \frac{1}{m}, \qquad j = 0,...,m-1$$
$$0 = t_0 < t_1 < ... < t_m = 1.$$

Let $\overline{\mathcal{U}}$ be the set consisting of all those elements from $\mathcal{U}$ which are piecewise constant and consistent with the partition $\mathcal{I}$. It is clear that each $\overline{u} \in \overline{\mathcal{U}}$ can be written as

$$\overline{u}(t) = \sum_{k=1}^{m} \sigma^k \chi_{I_k}(t), \qquad t \in [0,1) \tag{3.7}$$

where $\sigma^k \in R^r$ and χ_I denote the indicator function of I defined by

$$\chi_I(t) = \begin{cases} 1, & t \in I \\ 0, & \text{otherwise} \end{cases} \tag{3.8}$$

This means that each control $\overline{u} \in \overline{\mathcal{U}}$ can be identified uniquely by a control parameter σ and vice versa, where

$$\sigma = [(\sigma^1)^T, \ldots, (\sigma^m)^T] \in R^{mr}.$$

Thus, when no confusion can arise, we shall interchangeably refer to $\bar{u} \in \bar{\mathcal{U}}$ and $\sigma \in R^r$.

Define

$$\bar{\mathcal{V}} = \{\xi = [\sigma^T, \pi^T]^T \in R^{mr+p} : \sigma \in R^{mr}, \pi \in R^p\},$$

where ξ is the overall decision pa[illegible]meter vector.

Thus, for each $\xi \in \bar{\mathcal{V}}$, the differential equation (3.1) takes the form

$$\dot{x}(t) = \bar{f}(t, x(t), \xi) \tag{3.9a}$$

with the initial condition

$$x(0) = x^0, \tag{3.9b}$$

where

$$\bar{f}(t, x(t), \xi) = f(t, x(t), \sum_{k=1}^{m} \sigma^k \chi_{I_k}(t), \pi). \tag{3.9c}$$

Let $x(\cdot \mid \xi)$ be the solution of system (3.9) corresponding to each $\xi \in \bar{\mathcal{V}}$. Let

$$\bar{g}(\xi) = [g_1(x(t_1 \mid \xi), \pi), \ldots, g_q(x(t_k \mid \xi), \pi)]^T. \tag{3.10}$$

Let $\mathcal{F}$ be the set of all feasible ξ defined by

$$\mathcal{F} = \{\xi \in \mathcal{V} : \bar{g}(\xi) = 0\}.$$

We may now specify the approximate problem (3.P_2) as follows:

Problem (3.P_2) Subject to system (3.9), find an overall decision parameter vector $\xi \in \mathcal{F}$ such that the cost functional

$$\mathcal{J}(\xi) = \phi(x(1|\xi), \pi) + \int_0^1 \bar{f}_0 \, (t, x(t|\xi), \xi) \, dt \tag{3.11}$$

is minimized over $\mathcal{F}$, where $\bar{f}_0$ is obtained from f_0 in an obvious manner.

3.4 Liapunov Concept

The problem (3.P_2) can be converted into a non–constrained problem in lagrangian form as follows:

Problem (3.P_3) Minimize

$$Q(\xi, \lambda) = \mathcal{J}(\xi) + \lambda^T \bar{g}(\xi), \tag{3.12}$$

where $\lambda \in R^q$ is a constant Lagrange multiplier.

Using the gradient of the function (3.12), we can introduce a positive–definite Liapunov Function (L). For example, let

$$L(w) = \sum_{i=1}^{N} K_i \, [h_i(w)]^2, \tag{3.13}$$

where

$$N = mr + p + q, \tag{3.14}$$

$$w = (\xi^T, \lambda^T)^T \in R^N, \tag{3.15}$$

$$h_i(w) = \begin{cases} \dfrac{\partial Q(\xi,\lambda)}{\partial \xi_i} = \dfrac{\partial J(\xi)}{\partial \xi_i} + \sum_{k=1}^{q} \lambda_k \dfrac{\partial \bar{g}_k(\xi)}{\partial \xi_i}, \\ \qquad i = 1, \ldots, mr + p \\ \dfrac{\partial Q(\xi,\lambda)}{\partial \lambda_{i-mr-p}} = \bar{g}_{i-mr-p}(\xi) \\ \qquad i = mr + p + 1, \ldots, N, \end{cases} \tag{3.16}$$

and K_i, i = 1,...,N are arbitrary chosen positive weighting factors. Obviously, L becomes zero only when all the gradients of $Q(\xi,\lambda)$ becomes zero – a necessary condition for optimality for the problem $(3.P_3)$. Thus, we do know that (ξ^*,λ^*) satisfy the necessary condition of optimality of the problem $(3.P_3)$ when $L = 0$ in (3.13). Accordingly, we pose the following optimization problem:

Problem $(3.P_4)$ Find a vector $w \in R^N$ which minimizes L defined in (3.13).

Remark 3.4.1 L may also be defined as follows:

(a) $$L(w) = \sum_{i=1}^{N} K_i \operatorname{Abs}[h_i(w)] \tag{3.17}$$

(b) $$L(w) = \sum_{i=1}^{N} K_i [h_i(w)]^{2m_i} \tag{3.18}$$

where $m_1, m_2, \ldots, m_N$ are real integers.

(c) $$L(w) = [h_1(w), \ldots, h_N(w)]^T S [h_1(w), \ldots, h_N(w)], \tag{3.19}$$

where S is a positive definite N × N matrix.

Problem (3.P_4) is an unconstrained mathematical programming problem, which can be solved using one of the existing optimization techniques such as conjugate gradient method.

Most of these techniques require the computation of the gradient of L and hence both the gradient and Hessian of $\bar{J}$ and $\bar{g}_k$, k = 1,...,q, because

$$\frac{\partial L(w)}{\partial \xi_j}$$

$$= 2 \sum_{i=1}^{mr+p} \left[K_i \left[\frac{\partial \bar{J}(\xi)}{\partial \xi_i} + \sum_{k=1}^{q} \lambda_k \frac{\partial \bar{g}_k(\xi)}{\partial \xi_i} \right] \times \left[\frac{\partial^2 \bar{J}(\xi)}{\partial \xi_i \partial \xi_j} + \sum_{k=1}^{q} \lambda_k \frac{\partial^2 \bar{g}_k(\xi)}{\partial \xi_i \partial \xi_j} \right] \right]$$

$$+ 2 \sum_{i=mr+p+1}^{N} \left[K_i \, \bar{g}_{i-mr-p}(\xi) \times \frac{\partial \bar{g}_{i-mr-p}(\xi)}{\partial \xi_j} \right], \; j = 1,\ldots,mr+p, \tag{3.20}$$

and

$$\frac{\partial L(w)}{\partial \lambda_j} = 2 \sum_{i=1}^{mr+p} \left[K_i \left[\frac{\partial \bar{J}(\xi)}{\partial \xi_i} + \sum_{k=1}^{q} \lambda_k \frac{\partial \bar{g}_k(\xi)}{\partial \xi_i} \right] \right] \times \frac{\partial \bar{g}_j(\xi)}{\partial \xi_i}, \; j = 1,\ldots,q \tag{3.21}$$

Since $\bar{J}$ and $\bar{g}_k$ are not straightforward explicit functions of ξ, the gradients $\partial\bar{J}/\partial\xi_i$, $\partial\bar{g}_k/\partial\xi_i$ and the Hessians $\partial^2\bar{J}/\partial\xi_i\partial\xi_j$ and $\partial^2\bar{g}_k/\partial\xi_i\partial\xi_j$ need to be computed in a round about way – by forming artificial optimal control problems and taking Fréchet derivatives. We shall discuss the method for computing these gradients and Hessians in the next two sections.

3.5 The Gradients of $\bar{J}$ and $\bar{g}_k$

The methods for computing the gradients of $\bar{J}$ and $\bar{g}_k$ are well known in the literature of combined optimal control and optimal parameter selection problem. (For example, see Ref. 95). We shall summarize the results as follows:

The gradient of $\bar{J}$ is given by

$$\frac{d\bar{J}}{d\xi} = \int_0^1 \frac{\partial H(t,x(t|\xi),\xi,\psi^0(t|\xi))}{\partial\xi}\,dt, \tag{3.22}$$

where $x\,(t|\xi)$ is the solution of system (3.9); $H : [0,1] \times R^n \times R^{mr+p} \times R^n \longrightarrow R$ is the Hamiltonian function for $\bar{J}$ defined by

$$H\,(t,x,\xi,\psi) = \bar{f}_0\;(t,x,\xi) + (\psi)^T\,\bar{f}\;(t,x,\xi) \tag{3.23}$$

and $\psi^0\,(t|\xi) \in R^n$ satisfy the following adjoint system

$$[\dot{\psi}^0\,(t)]^T = -\frac{\partial H(t,x(t|\xi),\xi,\psi^0(t))}{\partial x\,(t|\xi)},\;\; t \in [0,1] \tag{3.24a}$$

$$[\psi^0\,(1)]^T = \frac{\partial\phi(x(1|\xi),\pi)}{\partial x\,(1|\xi)} = \beta^0\,(x\,(1|\xi),\,\pi) \tag{3.24b}$$

Remark 3.5.1 When we solve $(3.P_4)$ by an iterative method, such as the conjugate gradient method, the value of ξ is known at every iteration. Thus, the solution of system (3.9) and (3.24) do not constitute a two point boundary value problem.

The gradient of $\bar{g}_k$ is given by

$$\frac{\partial \bar{g}_k}{\partial \xi} = \int_0^{t_k} \frac{\partial \tilde{H}(t, x(t|\xi), \xi, \psi^k(t|\xi))}{\partial \xi}\, dt, \qquad k = 1,\ldots,q \tag{3.25}$$

where $x(t|\xi)$ is the solution of system (3.9); $\tilde{H} : [0,1] \times R^n \times R^{mr+p} \times R^n \longrightarrow R$ is the Hamiltonian function for each of the constraint $\bar{g}_k$, $k = 1,\ldots,q$, defined by

$$\tilde{H}(t,x,\xi,\psi) = (\psi)^T \, \bar{f}\,(t,x,\xi) \tag{3.26}$$

and $\psi^k(t|\xi) \in R^n$, $k = 1,\ldots,q$, satisfy the following adjoint systems

$$[\dot{\psi}^k(t)]^T = -\frac{\partial \tilde{H}(t, x(t|\xi), \xi, \psi^k(t))}{\partial x(t|\xi)}, \qquad t \in [0, t_k] \tag{3.27a}$$

$$[\psi^k(t_k)]^T = \frac{\partial g_k(x(t_k|\xi), \pi)}{\partial x(t_k|\xi)} = \beta^k(x(t_k|\xi), \pi) \tag{3.27b}$$

Lemma 3.5.1 There exists a positive constant L_1, independent of $\xi \in R^{mr+p}$, such that

$$\|\psi^k(\xi)\|_\infty \leq L_1, \qquad k = 0,\ldots,q. \tag{3.28}$$

Proof The proof is similar to that given for Lemma 3.2 of Ref. 126.

3.6 Hessian of $\bar{J}$ and $\bar{g}_k$

The method for computing the Hessian of $\bar{J}$ and $\bar{g}_k$ is much more complicated. Firstly, we shall derive the formula of $\partial^2 \bar{J}/\partial\xi_i\, \partial\xi_j$, i,j = 1,...,mr + p. For each i = 1,...,mr + p, we consider the following associated optimal control problem

Minimize

$$\hat{J}^i(\xi) \equiv \frac{\partial \bar{J}}{\partial \xi_i} = \int_0^1 \frac{\partial H(t,x(t|\xi),\xi,\psi^0(t|\xi))}{\partial \xi_i}\, dt. \tag{3.29}$$

For each $\xi \in \bar{\mathcal{V}}$, let

$$\begin{bmatrix} x(t|\xi) \\ \psi^0(t|\xi) \end{bmatrix} \in R^{2n} \tag{3.30}$$

be considered as the new state variable, where $x(t|\xi)$ and $\psi^0(t|\xi)$ are solutions of (3.9) and (3.24) respectively. For each i = 1,...,mr + p, let

$$\mathcal{H}^i : [0,1] \times R^n \times R^n \times R^{mr+p} \times R^n \times R^n \longrightarrow R$$

be the new Hamiltonian function defined by

$$\mathcal{H}^i(t,x,\psi,\xi,\Psi^0,\hat{\Psi})$$

$$= \frac{\partial H(t,x,\xi,\psi)}{\partial \xi_i} + (\Psi)^T \bar{f}(t,x,\xi) - (\hat{\Psi})^T \left[\frac{\partial H(t,x,\xi,\psi)}{\partial x}\right]^T$$

$$= \frac{\partial \overline{f}_0(t,x,\xi)}{\partial \xi_i} + (\psi)^T \frac{\partial \overline{f}(t,x,\xi)}{\partial \xi_i} + (\Psi)^T \overline{f}(t,x,\xi) - (\hat{\Psi})^T \left[\frac{\partial \overline{f}_0(t,x,\xi)}{\partial x} + (\psi)^T \frac{\partial \overline{f}(t,x,\xi)}{\partial x}\right]^T. \tag{3.31}$$

Now, let $\Psi^{oi}(t|\xi) \in R^n$ and $\hat{\Psi}^{oi}(t|\xi) \in R^n$ be the co–state variable corresponding to the state variable $x(t|\xi)$ and $\psi^o(t|\xi)$ respectively. Then $\Psi^{oi}(t|\xi)$ and $\hat{\Psi}^{oi}(t|\xi)$ need to satisfy the following adjoint system:

$$\{\dot{\Psi}^{oi}(t)]^T$$

$$= -\frac{\partial \mathscr{H}^i(t,x(t|\xi),\psi^o(t|\xi),\xi,\Psi^{oi}(t|\xi),\hat{\Psi}^{oi}(t|\xi))}{\partial x(t|\xi)}$$

$$= -\frac{\partial}{\partial x(t|\xi)}\left[\frac{\partial \overline{f}_0(t,x(t|\xi),\xi)}{\partial \xi_i}\right]$$

$$-[\psi^o(t|\xi)]^T \frac{\partial}{\partial x(t|\xi)}\left[\frac{\partial \overline{f}(t,x(t|\xi),\xi)}{\partial \xi_i}\right] - [\Psi^{oi}(t)]^T \frac{\partial \overline{f}(t,x(t|\xi),\xi)}{\partial x(t|\xi)}$$

$$+ [\hat{\Psi}^{oi}(t|\xi)]^T \left\{\frac{\partial}{\partial x(t|\xi)}\left[\frac{\partial f_0(t,x(t|\xi),\xi)}{\partial x(t|\xi)}\right]^T + \frac{\partial}{\partial x(t|\xi)}\left[\left[\frac{\partial \overline{f}(t,x(t|\xi),\xi)}{\partial x(t|\xi)}\right]^T \psi^o(t|\xi)\right]\right\} \tag{3.32a}$$

with the final condition

$$\Psi^{oi}(1) = 0 \tag{3.32b}$$

and

$[\dot{\hat{\Psi}}^{oi}(t)]^T$

$$= -\frac{\partial \mathscr{H}'(t, x(t|\xi), \psi^o(t|\xi), \xi, \Psi^{oi}(t|\xi), \hat{\Psi}^{oi}(t))}{\partial \psi^o(t|\xi)}$$

$$= -\left[\frac{\partial \mathscr{f}\ (t, x(t|\xi), \xi)}{\partial \xi_i}\right]^T + [\hat{\Psi}^{oi}(t)]^T \frac{\partial \mathscr{f}\ (t, x(t|\xi), \xi)}{\partial x(t|\xi)} \tag{3.33a}$$

with the initial condition

$$\hat{\Psi}^{oi}(0) = 0. \tag{3.33b}$$

Remark 3.6.1 During every iteration of any iterative method, the value of ξ is known. Thus, we first solve systems (3.9) and (3.24) to obtain $x(t|\xi)$ and $\psi^o(t|\xi)$ respectively. Once the value of $x(t|\xi)$ is known, we then solve system (3.33) to obtain $\hat{\Psi}^{oi}(t|\xi)$. Finally, $\Psi^{oi}(t|\xi)$ can be obtained by solving system (3.32) after all the values of $x(t|\xi), \psi^o(t|\xi)$ and $\hat{\Psi}^{oi}(t|\xi)$ have been obtained.

Lemma 3.6.1 There exists a positive constant L_2, independent of ξ, such that

$$|\hat{\Psi}^{oi}(t|\xi)| \le L_2 \tag{3.34}$$

and

$$|\Psi^{oi}(t|\xi)| \le L_2 \tag{3.35}$$

for all $i = 1,\ldots,mr + p$ and for all $t \in [0,1]$.

Proof From (3.32), we get

$\hat{\Psi}^{oi}(t|\xi)$

$$= \int_0^t \left\{ \frac{\partial \overline{f}}{\partial \xi_i}(\tau, x(\tau|\xi), \xi) + \left[\frac{\partial \overline{f}}{\partial x(\tau|\xi)}(\tau, x(\tau|\xi), \xi) \right]^T \hat{\Psi}^{oi}(\tau|\xi) \right\} d\tau. \tag{3.36}$$

Thus, from (3.36), the Cauchy–Schwartz Inequality and Assumption (3.A2), we get

$|\hat{\Psi}^{oi}(t|\xi)|$

$$\leq \int_0^t K_1(1 + |\hat{\Psi}^{oi}(\tau|\xi)|\ d\tau$$

$$\leq K_1 + \int_0^t K_1 |\hat{\Psi}^{oi}(\tau)|\ d\tau. \tag{3.37}$$

By Gronwall's lemma (cf. Ref 15,p.62, Lemma 2.3.5), it follows that

$|\hat{\Psi}^{oi}(t|\xi)|$
$\leq K_1 \exp(K_1 t)$
$\leq K_1 \exp(K_1)$
$= k_1,$ (3.38)

where

$$k_1 \equiv K_1 \exp(K_1)\ . \tag{3.39}$$

From (3.32), we get

$$[\Psi^{oi}(t|\xi)]^T$$

$$= \int_t^1 \left\{ \frac{\partial}{\partial x(\tau|\xi)} \left[\frac{\partial \overline{f}\ (\tau, x(\tau|\xi), \xi)}{\partial \xi_i} \right] + [\psi^o(\tau|\xi)]^T \frac{\partial}{\partial x(\tau|\xi)} \left[\frac{\partial \overline{f}\ (\tau, x(\tau|\xi), \xi)}{\partial \xi_i} \right] \right.$$

$$+ [\Psi^{oi}(\tau|\xi)]^T \frac{\partial \overline{f}\ (\tau, x(\tau|\xi), \xi)}{\partial x(\tau|\xi)} - [\hat{\Psi}^{oi}(\tau|\xi)]^T \left\{ \frac{\partial}{\partial x(\tau|\xi)} \left[\frac{\partial \overline{f}_i\ (\tau, x(\tau|\xi), \xi)}{\partial x(\tau|\xi)} \right]^T \right.$$

$$\left. - \frac{\partial}{\partial x(\tau|\xi)} \left[\left[\frac{\partial \overline{f}\ (\tau, x(\tau|\xi), \xi)}{\partial x(\tau|\xi)} \right]^T \psi^o(\tau|\xi) \right] \right\} d\tau. \tag{3.40}$$

Thus, from (3.40), the Cauchy–Schwartz inequality, Assumption (3.A3), Lemma 3.5.1, (3.38) and Assumption (3.A2), we get

$$|\Psi^{oi}(t|\xi)| \leq k_2 + \int_t^1 k_3 |\Psi^{oi}(\tau|\xi)|\, d\tau, \tag{3.41}$$

where

$$k_2 \equiv K_3(n + nL_1 + n^2 k_1 + n^3 L_1) \tag{3.42}$$

and

$$k_3 \equiv n^2 K_1 \tag{3.43}$$

Thus, by Gronwall's lemma, it follows that

$$\begin{aligned} &|\Psi^{oi}(t|\xi)| \\ &\leq k_2 \exp[k_3(1-t)] \\ &\leq k_2 \exp(k_3). \end{aligned} \tag{3.44}$$

By choosing

$$L_2 = \max\{k_1, k_2 \exp(k_3)\},$$

the conclusion of the lemma follows easily from (3.37) and (3.44).

Theorem 3.6.1 The Hessian matrix of $\bar{J}$ can be computed as follows:

$$\frac{\partial^2 \bar{J}}{\partial\xi_i \partial\xi_j}$$

$$= -[\hat{\Psi}^{0i}(1|\xi)]^T \left[\frac{\partial\beta^0(x(1|\xi),\pi)}{\partial\xi_j} + \frac{\partial\beta^0(x(1|\xi),\pi)}{\partial x(1|\xi)} \times \int_0^1 N(1,t) \times \frac{\partial \bar{f}(t,x(t|\xi),\xi)}{\partial\xi_j} dt\right]$$

$$+ \int_0^1 \frac{\partial \mathscr{H}^i(t,x(t|\xi),\psi^0(t|\xi),\xi,\Psi^{0i}(t|\xi),\hat{\Psi}^{0i}(t|\xi))}{\partial\xi_j} dt, \quad i,j = 1,\ldots,mr + p, \tag{3.45}$$

where β^0 is as defined in (3.24b) and $N(t,\tau)$ satisfy the following system

$$\frac{\partial N(t,\tau)}{\partial t} = \frac{\partial \bar{f}(t,x(t|\xi),\xi)}{\partial x(t|\xi)} N(t,\tau), \qquad 0 \leq \tau \leq t \leq 1 \tag{3.46a}$$

$$N(\tau,\tau) = I \text{ (identity matrix)}. \tag{3.46b}$$

Proof Let $\xi \in R^{mr+p}$ be given and let $\rho \in R^{mr+p}$ be arbitrary but fixed. Let

$$\xi(\epsilon) = \xi + \epsilon\rho, \tag{3.47}$$

where ϵ is an arbitrarily small real number. For brevity, let $x(\cdot,\epsilon)$ and $\psi^0(\cdot,\epsilon)$ denote, respectively, the solution of (3.9) and (3.24) corresponding to $\xi(\epsilon)$. Clearly, from (3.9), we have

$$\Delta x(t|\xi)$$

$$= \frac{dx(t|\epsilon)}{d\epsilon}\Big|_{\epsilon=0}$$

$$= \int_0^t \left[\frac{\partial \bar{f}(\tau, x(\tau|\xi), \xi)}{\partial x(\tau|\xi)} \Delta x(\tau|\xi) + \frac{\partial \bar{f}(\tau, x(\tau|\xi), \xi)}{\partial \xi} \rho\right] d\tau . \tag{3.48}$$

Thus,

$$\Delta \dot{x}(t|\xi)$$

$$= \frac{\partial \bar{f}(t, x(t|\xi), \xi)}{\partial x(t|\xi)} \Delta x(t|\xi) + \frac{\partial \bar{f}(t, x(t|\xi), \xi)}{\partial \xi} \rho$$

$$\equiv \Delta \bar{f}(t, x(t|\xi), \xi) \tag{3.49a}$$

with the initial condition

$$\Delta x(0|\xi) = 0. \tag{3.49b}$$

In view of (3.48) and Assumption (3.A2), it can be proved, by using a similar argument as that used in the proof of Lemma 3.6.1, that

$$|\Delta x(t|\xi)| \leq k_1|\rho|, \tag{3.50}$$

where

$$k_1 = K_1 \exp(K_1), \tag{3.51}$$

and K_1 is as defined in (3.7). Moreover, from (3.49), we get

$$\Delta x(t|\xi) = \left[\int_0^t N(t,\tau) \frac{\partial \bar{f}(\tau, x(\tau|\xi), \xi)}{\partial \xi} d\tau\right] \rho \tag{3.52}$$

where $N(t,\tau)$ is as defined in (3.46).

Now, in view of (3.24), it can be proved, using similar argument as that used to obtain (3.49) from (3.9), that

$$\Delta \dot{\psi}^0(t|\xi)$$

$$= -\frac{\partial}{\partial x(t|\xi)} \left[\frac{\partial H(t, x(t|\xi), \xi, \psi^0(t|\xi))}{\partial x(t|\xi)}\right]^T \Delta x(t|\xi)$$

$$-\frac{\partial}{\partial \xi} \left[\frac{\partial H(t, x(t|\xi), \xi, \psi^0(t|\xi))}{\partial x(t|\xi)}\right]^T \rho$$

$$-\frac{\partial}{\partial \psi^0(t|\xi)} \left[\frac{\partial H(t, x(t|\xi), \mathrm{i}, \psi^0(t|\xi))}{\partial x(t|\xi)}\right]^T \Delta \psi^0(t|\xi)$$

$$= \Delta\sigma\,(t,x\,(t|\,\xi),\,\xi,\,\psi^0(t|\,\xi)) \tag{3.53a}$$

with the final condition

$$\Delta\psi^0(1|\,\xi) = \frac{\partial[\beta^0(x(1\,|\,\xi),\pi)]^T}{\partial\xi}\,\rho + \frac{\partial[\beta^0(x(1\,|\,\xi),\pi)]^T}{\partial x\,(1\,|\,\xi)}\,\Delta x\,(1\,|\,\xi), \tag{3.53b}$$

where $\sigma : [0,1] \times R^n \times R^{mr+p} \times R^n \longrightarrow R^n$ is defined by

$$\sigma(t,x,\xi,\psi) = -\left[\frac{\partial H(t,x,\xi,\psi)}{\partial x}\right]^T. \tag{3.53c}$$

In view of (3.53), (3.28), Assumption (3.A2), (3.50), Assumption (3.A2) and (3.24b), it is clear that

$$|\,\Delta\dot{\psi}^0(t|\,\xi)\,|$$

$$\leq K_3|\rho|[n^2k_1 + n^2L_1k_1 + n\,(mr + p) + n(mr + p)\,L_1] + K_1|\,\Delta\psi^0(t|\,\xi)\,| \tag{3.54}$$

and

$$|\,\Delta\psi^0(1\,|\,\xi)\,| \leq K_3|\,\rho\,|[n(mr + p) + n^2k_1]. \tag{3.55}$$

From (3.54) and (3.55), it can be proved, using a similar argument as that used in the proof of Lemma 3.6.1, that

$$\Delta\psi^0(t|\,\xi) \leq k_2|\,\rho\,|, \tag{3.56}$$

where

$$k_2 = K_3[n^2k_1 + n^2L_1k_1 + n(mr + p) + n(mr + p)L_1 + n(mr + p) + n^2k_1] \exp (K_1) \tag{3.57}$$

Now, in view of (3.29), (3.31) and (3.53c), it is clear that for each $i = 1,\ldots,mr + p$, $\hat{J}^i(\xi(\epsilon))$ can be expressed as

$$\hat{J}^i(\xi(\epsilon))$$

$$= \int_0^1 \{ \mathscr{H}^i(t,x(t;\epsilon),\ \psi^o(t;\epsilon),\ \xi(\epsilon),\ \Psi^{oi}(t|\xi),\ \hat{\Psi}^{oi}(t|\xi)) - [\Psi^{oi}(t|\xi)]^T f(t,x(t;\epsilon),\ \xi(\epsilon))$$

$$- [\hat{\Psi}^{oi}(t|\xi)]^T \sigma(t, x(t;\epsilon),\ \xi(\epsilon),\ \psi^o(t;\epsilon))\}\ dt. \tag{3.58}$$

Thus, from (3.58), (3.49a) and (3.53a), we get

$$\Delta \hat{J}^i(\xi)$$

$$= \frac{d\hat{J}^i(\xi(\epsilon))}{\partial \epsilon}\bigg|_{\epsilon=0}$$

$$= \frac{d\hat{J}^i(\xi)}{\partial \xi}\rho$$

$$= \int_0^1 \{ \frac{\partial \mathscr{H}^i(t,x(t|\xi),\psi^o(t|\xi),\xi,\Psi^{oi}(t|\xi),\hat{\Psi}^{oi}(t|\xi))}{\partial x(t|\xi)} \Delta x(t|\xi)$$

$$+ \frac{\partial \mathscr{H}^i(t,x(t|\xi),\psi^o(t|\xi),\xi,\Psi^{oi}(t|\xi),\hat{\Psi}^{oi}(t|\xi))}{\partial \psi^o(t|\xi)} \Delta \psi^o(t|\xi)$$

$$+ \frac{\partial \mathscr{H}^i(t,x(t|\xi),\psi^o(t|\xi),\xi,\Psi^{oi}(t|\xi),\hat{\Psi}^{oi}(t|\xi))}{\partial \xi}\rho$$

$$- [\Psi^{oi}(t|\xi)]^T \Delta \dot{x}(t|\xi) - [\hat{\Psi}^{oi}(t|\xi)]^T \Delta \dot{\psi}^o(t|\xi)\Big\} \, dt. \tag{3.59}$$

Thus, from (3.59), (3.32a) and (3.33). we get

$$\frac{d\hat{J}^i(\xi)}{\partial \xi}\rho$$

$$= \int_0^1 \Big\{ -\frac{d}{dt}[(\Psi^{oi}(t|\xi))^T \Delta x(t|\xi) + (\hat{\Psi}^{oi}(t|\xi))^T \Delta \psi^o(t|\xi)]$$

$$+ \frac{\partial \mathscr{H}^i(t, x(t|\xi), \psi^o(t|\xi), \xi, \Psi^{oi}(t|\xi), \hat{\Psi}^{oi}(t|\xi))}{\partial \xi}\rho \Big\} \, dt$$

$$= (\Psi^{oi}(0|\xi))^T \Delta x(0|\xi) - (\Psi^{oi}(1|\xi))^T \Delta x(1|\xi) + (\hat{\Psi}^{oi}(0|\xi))^T \Delta \psi^o(0|\xi) - (\hat{\Psi}^{oi}(1|\xi))^T \Delta \psi^o(1|\xi)$$

$$+ \int_0^1 \frac{\partial \mathscr{H}^i(t, x(t|\xi), \psi^o(t|\xi), \xi, \Psi^{oi}(t|\xi), \hat{\Psi}^{oi}(t|\xi))}{\partial \xi}\rho \, dt. \tag{3.60}$$

However, from (3.49b), Lemma 3.6.1, (3.32b), (3.33b), (3.50) and (3.56), we obtain

$$(\Psi^{oi}(0|\xi))^T \Delta x(0|\xi) = 0 \tag{3.61}$$

$$(\Psi^{oi}(1|\xi))^T \Delta x(1|\xi) = 0 \tag{3.62}$$

and

$$(\hat{\Psi}^{oi}(0|\xi))^T \Delta \Psi^o(0|\xi) = 0. \tag{3.63}$$

Thus, in view of (3.61)--(3.63), (3.53b) and (3.52), equation (3.60) can be simplified as

$$\frac{d\hat{J}^i(\xi)}{d\xi}\rho$$

$$= \Big\{-(\hat{\Psi}^{oi}(1|\xi))^T \Big[\frac{\partial\beta^o(x(1|\xi),\pi)}{\partial\xi} + \frac{\partial[\beta^o(x(1|\xi),\pi)]^T}{\partial x(1|\xi)}\int_o^1 N(1,t) \times \frac{\partial\bar{f}\ (t,x(t|\xi),\xi)}{\partial\xi}\,dt\Big]$$

$$+ \int_o^1 \frac{\partial\mathscr{H}^{\,i}(t,x(t|\xi),\psi^o(t|\xi),\xi,\Psi^{oi}(t|\xi),\hat{\Psi}^{oi}(t|\xi))}{\partial\xi}\,dt\Big\}. \tag{3.64}$$

Since ρ is arbitrary, the conclusion of the theorem follows easily from (3.64) and the definition of $\hat{J}^i(\xi)$.

Our next aim is to derive the formula for $\partial^2\,\tilde{g}_k(\xi)/\partial\xi_i\partial\xi_j$. for each k = 1,...,q and i,j = 1,...,mr + p. For each k = 1,...,mr + p. For each k = 1,...,q and i = 1,...,mr + p, we consider the following associated optimal control problem.

Minimize

$$\hat{g}^{k,i}(\xi) = \frac{\partial\tilde{g}^k(\xi)}{\partial\xi_i} = \int_o^{t_k} \frac{\partial\tilde{H}\ (t,x(t|\xi),\xi,\psi^k(t|\xi))}{\partial\xi_i}\,dt. \tag{3.65}$$

For each i = 1,...,mr + p, let $\hat{\mathscr{H}}^i : [0,1] \times R^n \times R^n \times R^{mr+p} \times R^n \times R^n \longrightarrow R$ be the new Hamiltonian function defined by

$$\tilde{\mathscr{H}}^{i}(t,x,\psi^{k},\xi,\Psi,\hat{\Psi}) = \frac{\partial \tilde{H}\ (t,x,\xi,\psi)}{\partial \xi_i} + (\Psi)^{T} \bar{f}\ (t,x,\xi) - (\hat{\Psi})^{T} \left[\frac{\partial \hat{H}\ (t,x,\xi,\psi)}{\partial x}\right]^{T} \tag{3.66}$$

Let $\Psi^{ki}(t|\xi) \in R^{n}$ and $\hat{\Psi}^{ki}(t|\xi)$ be the co–state variable corresponding to $x(t|\xi)$ and $\psi^{k}(t|\xi)$ respectively. Then, they need to satisfy the following adjoint system:

$$[\dot{\Psi}^{ki}(t)]^{T} = -\frac{\partial \tilde{\mathscr{H}}^{i}(t,x(t|\xi),\psi^{k}(t|\xi),\xi,\Psi^{ki}(t|\xi),\hat{\Psi}^{ki}(t|\xi))}{\partial x(t|\xi)} \tag{3.67a}$$

with the final condition

$$\Psi^{ki}(t_{k}) = 0 \tag{3.67b}$$

and

$$[\dot{\hat{\Psi}}^{ki}(t)]^{T} = -\frac{\partial \hat{\mathscr{H}}^{i}(t,x(t|\xi),\psi^{k}(t|\xi),\xi,\Psi^{ki}(t|\xi),\hat{\Psi}^{ki}(t|\xi))}{\partial \psi^{k}(t|\xi)} \tag{3.68a}$$

with the initial condition

$$\hat{\Psi}^{ki}(0) = 0. \tag{3.68b}$$

Lemma 3.6.2 There exists a positive constant L_3, indepe ' ξ, such that

$$|\hat{\Psi}^{ki}(t|\xi)| \leq L_3 \tag{3.69}$$

and

$$|\Psi^{ki}(t|\xi)| \leq L_3 \tag{3.70}$$

for all $k = 1,\ldots,q$, $i = 1,\ldots,mr + p$ and for all $t \in [0,t_k]$.

Proof The proof is similar to that given for Lemma 3.6.1.

Theorem 3.6.2 The Hessian matrix of $\tilde{g}_k$, $k = 1,\ldots,q$ can be computed as follows:

$$\frac{\partial^2 \tilde{g}_k}{\partial \xi_i \partial \xi_j}$$

$$= -(\hat{\Psi}^{ki}(t_k|\xi))^T \left[\frac{\partial \beta^k(x(t_k|\xi),\pi)}{\partial \xi_j} + \frac{\partial \beta^k(x(t_k|\xi),\pi)}{\partial x(t_k|\xi)} \times \int_0^{t_k} N(t_k,t) \times \frac{\partial \overline{f}\,(t,x(t|\xi),\xi)}{\partial \xi_j}\, dt\right]$$

$$+ \int_0^{t_k} \frac{\partial\ \tilde{\mathscr{H}}^{i}(t,x(t|\xi),\psi^k(t|\xi),\xi,\Psi^{ki}(t|\xi),\hat{\Psi}^{ki}(t|\xi))}{\partial \xi_j}\, dt, \quad i,j = 1,\ldots,mr + p \tag{3.71}$$

Proof The proof is similar to that given to Theorem 3.6.1, except that Lemma 3.6.1 is being replaced by Lemma 3.6.2 in the proof.

3.7 An Illustrative Example

In this section, we consider two examples. For each of these two examples, the fourth order Runge Kutta integration scheme is used to integrate the system forward in time and the adjoint system backward in time. To solve the corresponding approximate problem $(3.P_4)$, we used the nonlinear programming software NLQPL described in Ref. 83.

Example 3.7.1 Optimal design of suspended cable (cf. Ref 94).

Consider a cable under self–weight and distributed load along its span. After appropriate statistical analysis and normalization, the total (non–dimensional) weight of the cable is given by

$$J = \int_0^1 \frac{1}{\beta} \sqrt{(1 + s^2)(1 + (dy(x)/dx)^2)}\, dx \tag{3.72}$$

where

$$\frac{d^2y(x)}{dx^2} = \alpha \sqrt{1 + (dy(x)/dx)^2}\sqrt{1 + s^2} + \beta \tag{3.73a}$$

$$y(0) = 0, \qquad \frac{dy(0)}{dx} = 0 \tag{3.73b}$$

$$\frac{dy(1)}{dx} = s, \tag{3.73c}$$

where

$\alpha =$ given constant which relates the specific weight and the maximum permissible stress;

$\beta =$ an adjustable parameter representing the ratio of total loading to horizontal tension in the cable; and

$s =$ maximum slope of the cable which is also adjustable.

The optimal design problem is to determine β and s such that J is minimized.

Let

$$y_1(x) \equiv y(x) \tag{3.74}$$

$$y_2(x) \equiv dy(x)/dx. \tag{3.75}$$

The problem is equivalent to

$$\min \quad J = (s-\beta)\ /\ \beta \tag{3.76}$$

subject to

$$\frac{dy_2(x)}{dx} = \alpha\sqrt{1\ +\ (y_2(x))^2}\sqrt{1\ +\ s^2} + \beta \tag{3.77a}$$

$$y_2(0) = 0 \tag{3.77b}$$

and the constraint

$$g_1(y_2(1)) = y_2(1) - s = 0. \tag{3.77c}$$

The approximate problem

$$\min \quad L\,(\beta,s,\lambda)$$

has been solved using the weighting factors $K_1 = 1$, $K_2 = 1$ and $K_3 = 10$, where L is as defined in (3.13). Using (1,1,1) as the initial guess for the parameter (β,s,λ), NLQPL generated the optimal parameter for 5 different values of α as listed in Table 3.7.1. It appears that the method works very well for this problem, since all the numerical results are extremely close to the optimal solutions obtained in Ref. 94. At the computed optimal parameter, the values of L and g_1 are very close to zero.

a	Optimal parameter			Number of iterations	L	J	g_1
	β	s	λ				
0.1	1.1711	1.3858	9.5639	44	0.5860×10^{-8}	1.8333	$-0.$[illegible]88×10^{-4}
0.3	0.7689	1.4252	6.7807	13	0.4363×10^{-9}	2.8455	-0.1910×10^{-6}
0.4	0.5625	1.4467	9.0825	18	0.1292×10^{-10}	3.9296	0.5[illegible]$ \times 10^{-6}$
0.45	0.4578	1.4579	11.9096	18	0.2248×10^{-9}	4.8538	0.3036×10^{-5}
0.5	0.3521	1.4693	17.7048	26	0.2091×10^{-8}	5.3461	-0.9790×10^{-5}

Table 3.7.1 The Numerical Results for Example 3.7.1

Example 3.7.2

Consider the optimal control problem of minimizing

$$J = \int_0^1 [1 + x_1^2(t) + x_2^2(t) + u^2(t)]\, dt \tag{3.78}$$

subject to the differential equations

$$\dot{x}_1(t) = u(t) - x_2^2(t) + 1 \tag{3.79a}$$

$$\dot{x}_2(t) = u(t) - x_1(t)\, x_2(t) \tag{3.79b}$$

with the initial condition

$$x_1(0) = 0, \qquad\qquad x_2(0) = 1 \tag{3.79c}$$

and the terminal equality constraints

$$g_1(x_1(1)) = x_1(1) - 1 = 0 \tag{3.79d}$$

$$g_2(x_2(1)) = x_2(1) - 2 = 0. \tag{3.79e}$$

Using the control parameterization method with the number of partition m equal to 10, the approximate problem

$$\min \quad L(\sigma_1, \sigma_2, \ldots, \sigma_{10}, \lambda_1, \lambda_2)$$

is solved using the weighting factors $K_i = 0.4$, $i = 1,\ldots,10$; $K_{11} = 1.5$ and $K_{12} = 1.0$, where L is as defined in (3.13). Using 0.5 as the initial guess for all the parameters, NLQPL generated final convergence after 66 iterations. The corresponding value of L is

0.23807×10^{-5}, which is extremely close to zero. At the computed optimal control, the value of J is 7.94450 and the values of g_1 and g_2 are 0.39044×10^{-7} and -0.86640×10^{-7} respectively, which are extremely close to zero. The optimal control in plotted in Figure 3.7.2a.

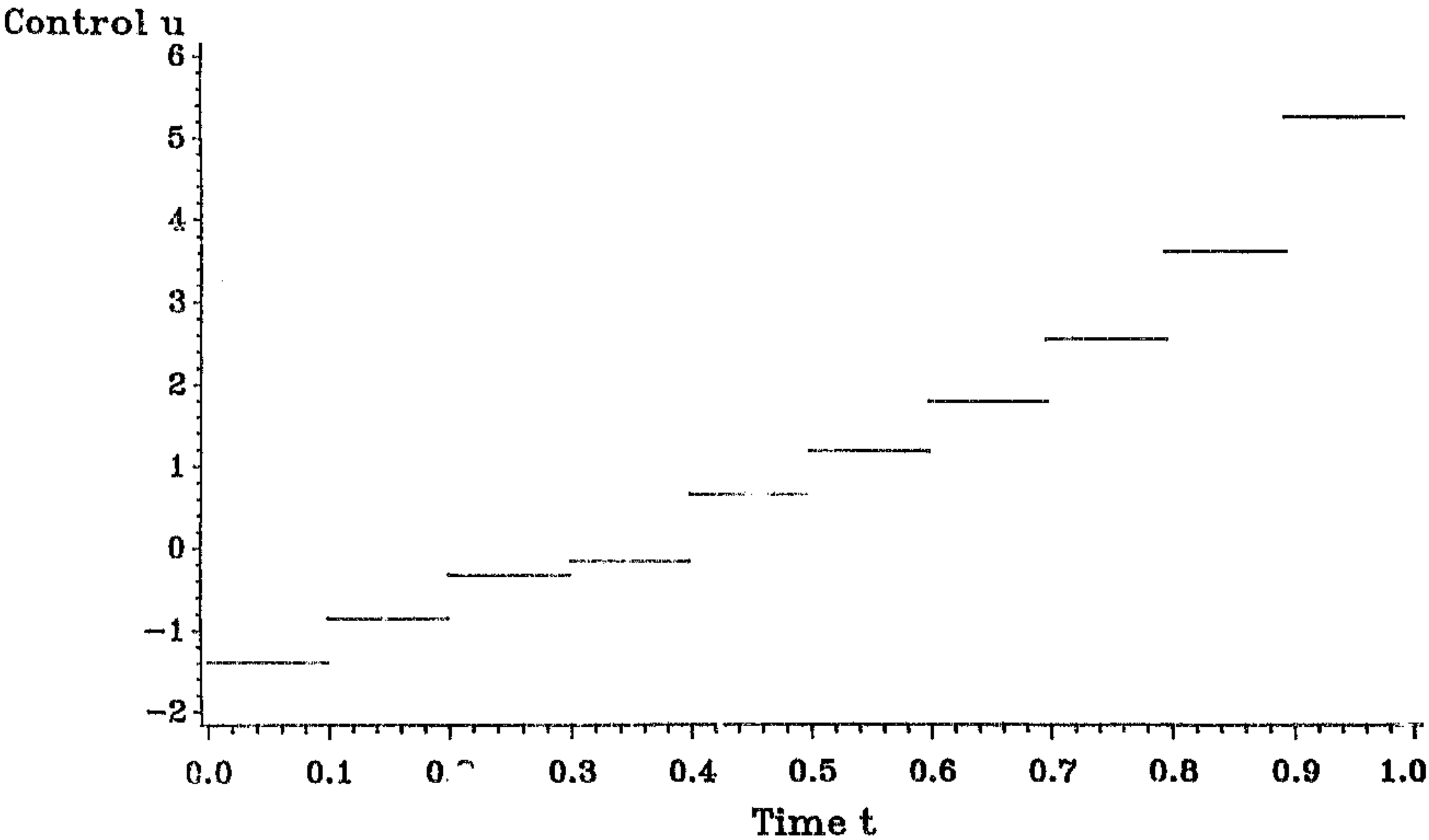

Figure 3.7.2a Computed Optimal Control for Example 3.7.2.

CHAPTER IV

A CONTROL PARAMETERIZATION ALGORITHM FOR TIME–DELAYED OPTIMAL CONTROL PROBLEMS WITH TERMINAL STATE CONSTRAINTS AND CONTINUOUS STATE CONSTRAINTS

4.1 Introduction

In this chapter, we extend the results of Ref. 110 to a more general class of constrained time–delayed optimal control problems, which involves terminal state equality constraints, as well as terminal state inequality constraints and continuous state constraints.

In Section 4.2, we describe the optimal control problem (4.P), and assume certain conditions.

In Section 4.3, we generate a sequence of approximate problems (4.P (P)) by the technique of control parameterization.

In Section 4.4, by using a constraint transcription similar to that used in Refs. 103, 110, we replace the state continuous inequality constraints by integral constraints consisting of very smooth functions. The integral constraints so constructed can satisfy the usual constraint qualification. For more detail of the properties of this constraint transcription, please refer to the Introduction Section in Chapter 1. The new approximate problem with the state continuous inequality constraints in (4.P (P)) being replaced by the integral constraints can be easily solved by the Algorithm $A2$ of Ref.103.

In Section 4.5, we investigate the convergence properties of the sequence of optimal controls of the approximate problems to the true optimal control. The main convergence results is given in Theorems 4.5.1 and 4.5.2.

In Section 4.6, two examples have been solved to illustrate the efficiency of the method discussed in this chapter.

4.2 Statement of the Problem

Consider a process described by the following system of delay–differential equations on the fixed time interval $(0, T]$,

$$\dot{x}(t) = f(t, x(t), x(t-h), u(t), u(t-h)), \tag{4.1a}$$

where $x = [x_1, \ldots, x_r]^T \in R^n$, $u = [u_1, \ldots, u_r]^T \in R^r$ are, respectively, the state and control vectors; $f = [f_1, \ldots, f_n]^T \in R^n$; the superscript T denotes the transpose and h is the time–delay satisfying

$$0 < h < T. \tag{4.1b}$$

For the sake of simplicity, we have confined our discussion in this paper to the case of a single time delay. However, all the results can be extended in a straightforward manner to the case of multiple time delays. The initial function for the differential equation (4.1a) is

$$x(t) = \zeta(t), \; t \in [-h, 0); \qquad x(0) = x^0, \tag{4.1c}$$

where $\zeta = [\zeta_1,...,\zeta_n]^T \in R^n$ is a given, piecewise continuous function on $[-h,0)$ and $x^0 \in R^n$ is a given vector. Define

$$U = \{v = [v_1,...,v_r]^T \in R^r : \alpha_i \leq v_i \leq \beta_i, \quad i = 1,...,r\}, \tag{4.2}$$

where α_i and β_i, $i = 1,...,r$ are real numbers. Clearly, U is a compact and convex subset of R^r. Let $\mathscr{U}$ be the class of all admissible controls defined by

$$\mathscr{U} = \{u : u = [u_1,...,u_r]^T \text{ is a bounded measurable function defined on } [-h,T] \text{ such that } u(t) \in U \text{ for all } t \in [0,T] \text{ and } u(t) = \beta(t) \text{ for all } t \in [-h,0)\}, \tag{4.3}$$

where β is a given piecewise continuous function defined on $[-h,0)$.

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

$$\|u\|_\infty = \text{ess} \sup_{t\in[-h,T]} \left[\sum_{i=1}^{r} (u_i(t))^2 \right]^{\frac{1}{2}}. \tag{4.4}$$

For each $u \in L^r_\infty$, let $x(\cdot|u)$ be an absolutely continuous function defined on $[0,T]$ which satisfies the differential equation (4.1a) almost everywhere on $[0,T]$ and the initial function (4.1c) everywhere on $[-h,0]$. This function $x(\cdot|u)$ is called the solution of system (4.1) corresponding to the control $u \in L^r_\infty$.

The terminal state inequality and equality constraints and continuous state inequality constraints are specified as follows:

$$\phi_i(x(T|u)) \geq 0, \qquad i = 1,\dots,N_I \qquad (4.5a)$$

$$\psi_i(x(T|u)) = 0, \qquad i = 1,\dots,N_E \qquad (4.5b)$$

$$g_i(t,x(t|u), x(t-h|u)) \geq 0,\ t \in [0,T], \qquad i = 1,\dots,N, \qquad (4.5c)$$

where ϕ_i, $i = 1,\dots,N_I$ and ψ_i, $i = 1,\dots,N_E$ are real–valued functions defined on R^n; and g_i, $i = 1,\dots,N$ are real–valued functions defined on $[0,T] \times R^{2n}$. Let

$$\Theta = \{u \in \mathscr{U}: \phi_i(x(T|u)) \geq 0, \qquad i = 1,\dots,N_I;$$
$$\min_{t\in[0,T]} g_i(t,x(t|u), x(t-h|u)) \geq 0, \qquad i = 1,\dots,N\}. \qquad (4.6)$$

Let

$$\mathscr{F} = \{u \in \Theta : \psi_i(x(T|u)) = 0, i = 1,\dots,N_E\}. \qquad (4.7)$$

Elements from $\mathscr{F}$ are called feasible controls and $\mathscr{F}$ is called the class of feasible controls.

We may now state the optimal control problem as follows.

Problem (4.P) subject to the system (4.1), find a control $u \in \mathscr{F}$ such that the functional

$$J(u) = \Phi_0(x(T|u)) + \int_0^T \mathscr{L}_0(t,x(t|u), x(t-h|u), u(t), u(t-h))\, dt \qquad (4.8)$$

is minimized over $\mathscr{F}$, where Φ_0 and $\mathscr{L}_0$ are given real–valued functions.

We assume that the following conditions are satisfied throughout.

(4.A1) $f : [0,T] \times R^{2n} \times R^{2r} \longrightarrow R^n$ is piecewise continuous on $[0,T]$ for each $(x,y,u,v) \in R^{2n} \times R^{2r}$ and are continuously differentiable with respect to each of the components of x,y,u and v for each $t \in [0,T]$. Furthermore, for each compact set $\Omega \in R^{2r}$, there exists a constant $K > 0$ such that

$$|f(t,x,y,u,v)| \leq K(1 + |x| + |y|) \tag{4.9}$$

for all $(x,y,u,v) \in R^{2n} \times \Omega$, where $|\cdot|$ denotes the usual Euclidean norm;

(4.A2) for each $i = 1,\ldots,N_I$, $\phi_i : R^n \longrightarrow R$ is continuously differentiable;

(4.A3) for each $i = 1,\ldots,N_E$, $\psi_i : R^n \longrightarrow R$ is continuously differentiable;

(4.A4) for each $i = 1,\ldots,N$, $g_i : [0,T] \times R^{2n} \longrightarrow R$ is continuously differentiable;

(4.A5) $\Phi_0 : R^n \longrightarrow R$ is continuously differentiable;

(4.A6) $\mathscr{L}_0 : [0,T] \times R^{2n} \times R^{2r} \longrightarrow R$ is piecewise continuous on $[0,T]$ for each $(x,y,u,v) \in R^{2n} \times R^{2r}$ and is continuously differentiable with respect to each of the components of x,y,u and v for each $t \in [0,T]$.

Remark 4.2.1 For each $u \in \mathscr{U}$, there exists a unique absolutely vector–valued function $x(\cdot|u)$ which satisfies the system (4.1). This is done as follows: We subdivide the interval $[0,T]$ into $[0,h]$, $[\ell h, (\ell+1)h]$, $\ell = 1,\ldots,k$, $[(k+1)h, T]$. Thus, finding unique solution of system (4.1) is the same as finding unique solution (4.1) on each of these subintervals successively with appropriate boundary conditions. Thus, the conclusion follows from repeated applications of well–known results on ordinary differential equations.

Lemma 4.2.1 There exists a bounded subset X of R^n such that $\{\|x(\cdot|u)\|_\infty : u \in \mathscr{U}\} \subset X$.

Proof Since U is bounded, the proof is similar to that given for Lemma 4.2 of Ref. 125.

4.3 Control Parameterization

As in Chapter 3, control parameterization method will be used to solve the problem (4.P). However the method used in this chapter is not exactly the same as that used in Chapter 3. To be more precise, let $\{\mathcal{J}_p\}_{p=1}^{\infty}$ be a sequence of partitions of the interval $[0,T]$ such that $\mathcal{J}_p$ has n_p elements; $\mathcal{J}_{p+1}$ is a refinement of $\mathcal{J}_p$ and $\|\mathcal{J}_p\| \longrightarrow 0$ as $p \longrightarrow \infty$, where $\|\mathcal{J}_p\|$ is the length of the largest interval in the partition $\mathcal{J}_p$. In this chapter, we assume that

$$\mathcal{J}_p = \{I_j^p\}_{j=1}^{n_p}, \tag{4.10}$$

where

$$I_j^p = [t_{j-1}^p, t_j^p),\ 0 = t_0^p < t_1^p < \ldots < t_{n_p}^p = T. \tag{4.11}$$

Let $\mathcal{U}^p$ be the subset of all admissible controls which are piecewise constant and consistent with the partition $\mathcal{J}_p$. Then, each $u^p \in \mathcal{U}^p$ can be written as

$$u^p(t) = \begin{cases} \beta(t), & t \in [-h,0) \\ \sum_{k=1}^{n_p} \sigma_k^p \chi_k^p(t), & t \in [0,T] \end{cases} \tag{4.12}$$

where $\sigma_k^p \in R^r$ and χ_k^p denote the characteristic function of I_k^p, i.e.

$$\chi_k^p(t) = \begin{cases} 1, & \text{if } t \in I_k^p \\ 0, & \textit{otherwise} \end{cases} \tag{4.13}$$

Thus, each control $u^p \in \mathscr{U}^p$ can be uniquely identified with a control parameter vector σ^p and vice versa, where

$$\sigma^p = [(\sigma_1^p)^T, \ldots, (\sigma_{n_p}^p)^T]. \tag{4.14}$$

Thus, we shall interchangeably refer to $u^p \in \mathscr{U}^p$ and $\sigma^p \in \mathscr{U}^p$.

Given a control parameter $\sigma^p \in \mathscr{U}^p$, let $x(\cdot \mid \sigma^p)$ be the solution of the differential equation

$$\dot{x}(t) = \hat{f}(t, x(t), x(t-h), \sigma^p) \tag{4.15a}$$

with the initial conditions

$$x(t) = \phi(t),\ t \in [-h, 0), \qquad x(0) = x^0, \tag{4.15b}$$

where

$$\hat{f}(t, x(t), x(t-h), \sigma^p) = f(t, x(t), x(t-h), u^p(t), u^p(t-h)). \tag{4.15c}$$

Let Θ^p and $\mathscr{F}^p$ be defined, respectively, from Θ and $\mathscr{F}$ by replacing u by σ^p in the definitions of Θ and $\mathscr{F}$.

We define the approximate problem (4.P(p)) as follows:

Problem (4.P(p)) Find a control vector $\sigma^p \in \mathscr{F}^p$ such that the cost functional

$$J(\sigma^p) = \Phi_0(x(T \mid \sigma^p)) + \int_0^T \hat{\mathscr{L}}_0(t, x(t \mid \sigma^p), x(t-h \mid \sigma^p), \sigma^p)\, dt \tag{4.16}$$

is minimized over $\mathscr{F}^p$, where

$$\hat{\mathscr{L}}_0(t,x(t|\sigma^p), x(t-h|\sigma^p), \sigma^p) = \mathscr{L}_0(t,x(t|u^p), x(t-h|u^p), u^p(t), u^p(t-h)). \tag{4.17}$$

The following theorem is a consequence of Lem na 4.3 and 4.4 of Ref 125.

Lemma 4.3.1 Let $\{u^p\}_{p=1}^{\infty}$ be a sequence of admissible control, which converge to $\bar{u}$ almost everywhere on $[0,T]$. Then

(i) $$\lim_{p\to\infty} \|x(u^p) - x(\bar{u})\|_{\infty} = 0 \tag{4.18}$$

(ii) $$\lim_{p\to\infty} J(u^p) = J(\bar{u}) \tag{4.19}$$

4.4 Constraint Approximation

For each $i = 1,\ldots,N$, the corresponding state inequality constraint (4.5e) when restricted to the space $\mathscr{U}^p$ can be written as

$$G_i(\sigma^p) = \int_0^T \min\{0, g_i(t,x(t|\sigma^p), x(t-h|\sigma^p))\}\, dt = 0 \tag{4.20}$$

which is, however, non–smooth.

As in Ref. 103, we replace the non–smooth functions $\min\{0, g_i(t,x(t|\sigma^p), x(t-h|\sigma^p))\}$ by the smooth ones given by $g_{i,\epsilon}(t,x(t|\sigma^p), x(t-h|\sigma^p))$, where

$$g_{i,\epsilon}(t,x(t|\sigma^p),\, x(t-h|\sigma^p))$$

$$=\begin{cases} g_i(t,x(t|\sigma^p),\, x(t-h|\sigma^p)), & \text{if } g_i(t,x(t|\sigma^p),\, x(t-h|\sigma^p)) < -\epsilon \\ -[g_i(t,x(t|\sigma^p),\, x(t-h|\sigma^p)) - \epsilon]^2/4\epsilon, & \text{if } |g_i(t,x(t|\sigma^p),\, x(t-h|\sigma^p))| \le \epsilon \\ 0 & \text{if } g_i(t,x(t|\sigma^p),\, x(t-h|\sigma^p)) > \epsilon. \end{cases} \tag{4.21}$$

For each $i = 1,\ldots,N$, define

$$G_{i,\epsilon}(\sigma^p) = \int_0^T g_{i,\epsilon}(t,x(t|\sigma^p),\, x(t-h|\sigma^p))\, dt. \tag{4.22}$$

We define two approximate problems, which will be referred to as $(4.P_\epsilon(\text{p}))$ and $(4.P_{\epsilon,\gamma}(\text{p}))$.

Problem $(4.P_\epsilon(\text{p}))$ Problem $(4.P(\text{p}))$ with the continuous state inequality constraint (4.5c) replaced by

$$G_{i,\epsilon}(\sigma^p) = 0, \qquad i = 1,\ldots,N. \tag{4.23}$$

Let $\mathcal{F}_\epsilon^p$ be the feasible region of $(4.P_\epsilon(\text{p}))$, defined by

$$\begin{aligned} \mathcal{F}_\epsilon^p = \{\sigma^p \in \mathcal{U}^p : G_{i,\epsilon}(\sigma^p) = 0, &\quad i = 1,\ldots,N; \\ \phi_i(x(T|\sigma^p)) \ge 0, &\quad i = 1,\ldots,N_I; \\ \psi_i(x(T|\sigma^p)) = 0, &\quad i = 1,\ldots,N_E; \end{aligned} \tag{4.24}$$

Then, it is clear that

$$\mathcal{F}_\epsilon^p \subset \mathcal{F}^p. \tag{4.25}$$

Note that the equality constraints (4.24) fail to satisfy the usual constraint qualification. Thus, we may encounter numerical difficulty. To overcome this difficulty, we consider the second approximation problem as follows:

Problem ($4.P_{\epsilon,\gamma}$(p)) Problem ($4.P$ (p)) with (4.5c) being replaced by

$$G_{i,\epsilon}(\sigma^p) \geq -\gamma, \qquad i = 1,\ldots,N. \tag{4.26}$$

In order to study the relationship between the optimal solution of the problems ($4.P$ (p)) and ($4.P_\epsilon$(p)) as $\epsilon \to 0$, we need the following additional assumptions:

(4.A7) For each $\sigma^p \in \mathscr{F}^p$, there exists a sequence $\{\sigma^{p,\epsilon}\} \in \mathscr{F}^p_\epsilon$ such that $\{\sigma^{p,\epsilon}\}$ converges to σ^p in the Euclidean norm whenever $\epsilon \to 0$.

Lemma 4.4.1 For each $p > 0$, let $\{\sigma^{p,*}_\epsilon\}$ be a sequence in ϵ of optimal controls of ($4.P_\epsilon$(p)) and let $\sigma^{p,*}$ be an optimal control of the problem ($4.P$ (p)). Then

(i) the sequence $\{\sigma^{p,*}_\epsilon\}$ has a subsequence which converges to an optimal control of the problem ($4.P$ (p)) in the Euclidean norm.

(ii) $$\lim_{\epsilon \to 0} J(\sigma^{p,*}_\epsilon) = J(\sigma^{p,*}) \tag{4.27}$$

Proof (i) From the definition of $\mathscr{U}^{p}$, the sequence of vectors $\{\sigma_{\epsilon}^{p,*}\}$ is obviously bounded in the Euclidean norm. Therefore, there exists a subsequence, denoted by $\{\sigma_{\bar{\epsilon}}^{p,*}\}$, such that $\{\sigma_{\bar{\epsilon}}^{p,*}\}$ converges to $\bar{\sigma}^{p}$ in the Euclidean norm. Since each $\sigma_{\bar{\epsilon}}^{p,*} \in \mathscr{F}_{\bar{\epsilon}}^{p} \subset \mathscr{F}^{p}$, it is clear from Lemma 4.3.1 (i) and Assumption (4.A2) – (4.A4) that $\bar{\sigma}^{p}$ is also a feasible control of (4.P (p)). We now wish to prove that $\bar{\sigma}^{p}$ is an optimal control of the problem (4.P (p)).

In view of Assumption (4.A7), there exists a sequence $\{\sigma^{p,*,\bar{\epsilon}}\} \in \mathscr{F}_{\bar{\epsilon}}^{h}$ such that $\sigma^{p,*,\bar{\epsilon}}$ converges to $\sigma^{p,*}$ in the Euclidean norm as $\bar{\epsilon} \to 0$. Since $\sigma_{\bar{\epsilon}}^{p,*}$ is the optimal control of $P_{\bar{\epsilon}}(p)$, we have

$$J(\sigma_{\bar{\epsilon}}^{p,*}) \leq J(\sigma^{p,*,\bar{\epsilon}}) \tag{4.28}$$

Using Lemma 4.3.1, we obtain by taking limit as $\bar{\epsilon} \to 0$ in (4.28) that

$$J(\bar{\sigma}^{p}) \leq J(\sigma^{p,*}). \tag{4.29}$$

Thus, from (4.29) and the definition of $\sigma^{p,*}$, we conclude that $\bar{\sigma}^{p}$ is also an optimal control of the problem (4.P (p)).

(ii) Since the sequence $\{J(\sigma_{\epsilon}^{p,*})\}$ is a non–increasing sequence of ϵ and has a lower bound $J(\sigma^{p,*})$, the whole sequence $\{J(\sigma_{\epsilon}^{p,*})\}$ is convergent.

Now, from part (i) and Lemma 4.3.1, we have

$$\lim_{\bar{\epsilon}\to 0} J(\sigma^{p,*}_{\bar{\epsilon}}) = J(\sigma^{p,*}), \tag{4.30}$$

where $\{\sigma^{p,*}_{\bar{\epsilon}}\}$ is the subsequence of $\{\sigma^{p,*}_{\epsilon}\}$ as described in part (i) of the proof of this lemma. Thus, we have

$$\lim_{\epsilon\to 0} J(\sigma^{p,*}_{\epsilon}) = \lim_{\bar{\epsilon}\to 0} J(\sigma^{p,*}_{\bar{\epsilon}}) = J(\sigma^{p,*}) \tag{4.31}$$

Lemma 4.4.2 There exists a $\bar{\gamma}(\epsilon) > 0$ such that for all γ, $0 < \gamma < \bar{\gamma}(\epsilon)$, any feasible control vector $\sigma^{p}_{\epsilon,\gamma}$ of the problem $(4.P_{\epsilon,\gamma}(p))$, i.e.

$$G_{i,\epsilon}(\sigma^{p}_{\epsilon,\gamma}) \geq -\gamma, \qquad i = 1,\ldots,N \tag{4.32a}$$

$$\phi_i(x(T|\sigma^{p}_{\epsilon,\gamma})) \geq 0, \qquad i = 1,\ldots,N_I \tag{4.32b}$$

$$\psi_i(x(T|\sigma^{p}_{\epsilon,\gamma})) = 0, \qquad i = 1,\ldots,N_E \tag{4.32c}$$

is also a feasible control parameter vector of the problem $(4.P(p))$.

Proof Let

$$y(t) = x(t-h). \tag{4.33}$$

Then, for each $i = 1,\ldots,N$ and any $\sigma^p \in \mathcal{U}^p$, we have

$$\frac{\partial g_i(t,x(t|\sigma^p),y(t|\sigma^p))}{\partial t} = \sum_{j=1}^{N} \frac{\partial g_i(t,x(t|\sigma^p),y(t|\sigma^p))}{\partial x_j(t|\sigma^p)} f_j(t,x(t|\sigma^p),y(t|\sigma^p))$$

$$+ \sum_{j=1}^{N} \frac{\partial g_i(t,x(t|\sigma^p),y(t|\sigma^p))}{\partial y_j(t|\sigma^p)} \frac{\partial y_j(t|\sigma^p)}{\partial t}$$

$$+ \frac{\partial g_1(t,x(t|\sigma^p),x(t-h|\sigma^p))}{\partial t} . \qquad (4.34)$$

Now, by (4.A4) and Lemma 4.2.1, there exists a positive constant m_1 such that, for all $\sigma^p \in \mathscr{U}^p$,

$$\left|\frac{\partial g_1(t,x(t|\sigma^p),y(t|\sigma^p))}{\partial t}\right| \leq m_1, \forall\, t \in [0,T] . \qquad (4.35)$$

The remainder of the proof of this lemma is very similar to that given in Lemma 3.3 of Ref. 103.

In view of Lemma 4.4.2, we can compute a sequence of suboptimal control parameters $\{\sigma^{p,*}_{\epsilon,\gamma}\}$ to Problem $(4.P\,(p))$ with each of them in $\mathscr{F}^p$. The following algorithm is exactly the same as Algorithm A2 of Ref. 103.

Algorithm 4.4.1

Step 0. Choose initial values of $\epsilon(= 10^{-2})$ and $\gamma(= \frac{\epsilon T}{4})$, say.

Step 1. Solve $(4.P_{\epsilon,\gamma}(p))$ to give $\sigma^{p,*}_{\epsilon,\gamma}$.

Step 2. If $\sigma^{p,*}_{\epsilon,\gamma}$ is not feasible, then set $\gamma = \gamma/2$ and go to Step 1; else, if $\epsilon > \bar{\epsilon}$, (e.g. say $\bar{\epsilon} = 10^{-6}$), then $\epsilon := \epsilon/10$ and $\gamma := \gamma/10$ and go to Step 1; else, stop.

Remark 4.4.1 From Remark 4.2 of Ref. 105, the problem $(4.P_{\epsilon,\gamma}(p))$ is essentially a nonlinear mathematical programming problem in control parameters, which can be solved by any standard software package such as NLPQL (see Ref. 83). However, in order to solve the problem $(4.P_{\epsilon,\gamma}(p))$, we need to know the gradient of the cost functional and that of the constraints with respect to the vector σ^p.

Remark 4.4.2 For each σ^p, let the function $\lambda(t|\sigma^p) : [-h,\ T+h] \longrightarrow R^n$ be the solution of the adjoint system

$$\begin{aligned}\dot{\lambda}(t) = &- f_x(t, x(t|\sigma^p), x(t-h|\sigma^p), \sigma^p)\ \lambda(t) \\ &- f_{\bar{x}}(t+h, x(t+h|\sigma^p), x(t|\sigma^p), \sigma^p)\ \lambda(t+h) \\ &- \mathscr{L}_{0x}(t, x(t|\sigma^p), x(t-h|\sigma^p), \sigma^p),\ t \in [0, T] \end{aligned} \tag{4.36a}$$

with the final conditions

$$\lambda(T) = \Phi_{0x}(x(T|\sigma^p)) \tag{4.36b}$$

$$\lambda(t) = 0, \qquad t \in [T,\ T+h] \tag{4.36c}$$

Following a similar approach as that used in the proof of Theorem 2.1 of Ref. 125, it can be shown that the gradient of the cost functional J is given by

$$\begin{aligned}\frac{\partial J}{\partial \sigma^p} = \int_0^T &[f_\sigma(t, x(t|\sigma^p), x(t-h|\sigma^p), \sigma^p)\ \lambda(t) \\ &+ f_\sigma(t+h, x(t+h|\sigma^p), x(t|\sigma^p), \sigma^p)\ \lambda(t+h) \\ &+ \mathscr{L}_{0\sigma}(t, x(t|\sigma^p), x(t-h|\sigma^p), \sigma^p)]\ dt \end{aligned} \tag{4.37}$$

Remark 4.4.3 Let $\hat{\lambda}(t|\sigma^p) : [-h,\ T+h] \longrightarrow R^n$ be defined from $\lambda(t|\sigma^p)$ by replacing the terms $\mathscr{L}_0$ and Φ_0 in the definition of λ by 0 and ϕ_i respectively. Then the gradient of the constraint function ϕ_i is given by

$$\begin{aligned}\frac{\partial \phi_i}{\partial \sigma^p} = \int_0^T &[f_\sigma(t, x(t|\sigma^p), x(t-h|\sigma^p), \sigma^p)\ \hat{\lambda}(t) \\ &+ f_\sigma(t+h, x(t+h|\sigma^p), x(t|\sigma^p), \sigma^p)\ \hat{\lambda}(t+h)]\ dt \end{aligned} \tag{4.38}$$

Remark 4.4.4 The method for calculating the gradient of the constraint functions ψ_i is exactly the same as that given for the gradient of ϕ_i.

Remark 4.4.5 Let $\bar{\lambda}(t|\sigma^p) : [-h, T+h] \to R^n$ be defined from $\lambda(t|\sigma^p)$ by replacing the terms $\mathscr{L}_0$ and Φ_0 in the definition of $\bar{\lambda}$ by $g_{i,\epsilon}$ and 0 respectively. Then the gradient of $G_{i,\epsilon}$ is given by

$$\begin{aligned}\frac{\partial G_{i,\epsilon}}{\partial \sigma^p} = \int_0^T & [f_\sigma(t, x(t|\sigma^p), x(t-h|\sigma^p), \sigma^p)\, \lambda(t) \\ & + f_\sigma(t+h, x(t+h|\sigma^p), x(t|\sigma^p), \sigma^p)\, \lambda(t+h) \\ & + g_{i,\epsilon_\sigma}(t, x(t|\sigma^p), x(t-h|\sigma^p), \sigma^p))]\, dt \end{aligned} \tag{4.39}$$

Theorem 4.4.1 Let $\{\sigma^{p,*}_{\epsilon,\gamma}\}$ be a sequence in ϵ of control parameter vectors produced by Algorithm 4.4.1. Then

$$J(\sigma^{p,*}_{\epsilon,\gamma}) \to J(\sigma^{p,*}), \tag{4.40}$$

where $\sigma^{p,*}$ is an optimal control of the problem (4.P (p)). Furthermore, any accumulation point of $\{\sigma^{p,*}_{\epsilon,\gamma}\}$ is also an optimal solution of Problem (4.P (p)).

Proof Clearly

$$J(\sigma^{p,*}) \le J(\sigma^{p,*}_{\epsilon,\gamma}) \le J(\sigma^{p,*}_{\epsilon}) \tag{4.41}$$

where $\sigma^{p,*}_{\epsilon}$ is as defined for Lemma 4.4.1.

Thus, by Lemma 4.4.1 and (4.41), we have

$$\lim_{\epsilon \to 0} J(\sigma^{p,*}_{\epsilon,\gamma}) = J(\sigma^{p,*}). \tag{4.42}$$

To prove the second part of the theorem, we note that the sequence $\{\sigma_{\epsilon,\gamma}^{p,*}\}$ ir : belongs to a compact subset of $R^{r \times n_p}$. Thus, there exists a subsequence of the sequence $\{\sigma_{\epsilon,\gamma}^{p,*}\}$, which is again denoted by the original sequence and a $\bar{\sigma}^p$ such that

$$\lim_{\epsilon \to 0} |\sigma_{\epsilon,\gamma}^{p,*} - \bar{\sigma}^p| = 0 \tag{4.43}$$

Thus, from (4.43), Lemma 4.3.1, and (4.4.2), we have

$$\lim_{\epsilon \to 0} J(\sigma_{\epsilon,\gamma}^{p,*}) = J(\bar{\sigma}^p) = J(\sigma^{p,*}) \tag{4.44}$$

Now, from Assumption (4.A2) – (4.A4) and part (i) of Lemma 4.3.1, it is clear that $\bar{\sigma}^p$ is also a feasible control of the problem $(4.P(p))$. Thus, the proof is complete.

4.5 Some Convergence Results

In this section, we shall investigate some convergence properties of the sequence of approximate optimal controls to the true optimal control.

Let $\{\sigma^{p,*}\}_{p=1}^{\infty}$ be a sequence of optimal control vectors of the finite-dimensional problem $(4.P(p))$ and let $\{u^{p,*}\}_{p=1}^{\infty}$ be the corresponding sequence of controls in $\mathscr{U}$. Clearly

$$J(u^{p+1,*}) \leq J(u^{p,*}) .$$

As in Refs. 28, 97, 103, we shall investigate the following two questions:

(i) Does $J(u^{p,*})$ converge to the true optimal cost?

(ii) Does $u^{p,*}$ converge to the true optimal control in some sense?

For the above two purposes, we first need to define the δ–tolerated version of the approximate problem (4.P (p)) and the original problem (4.P).

Definition 4.5.1 A control vector $\sigma^p \in \Theta^p$ (respectively, control $u \in \Theta$) is said to be δ–tolerated if it satisfied the constraints

$$|\psi_i(x(T|\sigma^p))| \leq \delta, \qquad i = 1,...,N_E. \tag{4.45}$$

(respectively),

$$|\psi_i(x(T|u))| \leq \delta, \qquad i = 1,...,N_E. \tag{4.46}$$

Let $\mathcal{F}^{p,\delta}$ (respectively, $\mathcal{F}^\delta$) be the subset of Θ^p (respectively, Θ) such that the δ–tolerated constraints (4.42) (respectively, (4.43)) are satisfied. We can now define the δ–tolerated version of the approximate problem (4.P (p)) and the original problem (4.P) as follows:

Problem (4.P^δ(p)) Find a control vector $\sigma^p \in \mathcal{F}^{p,\delta}$ such that the cost functional (4.16) is minimized over $\mathcal{F}^{p,\delta}$.

Problem (4.P^δ) Find a control $u \in \mathcal{F}^\delta$ such that the cost functional (4.8) is minimized over $\mathcal{F}^\delta$.

Lemma 4.5.1 For each p > 0, let $\{\sigma^{p,\delta,*}\}$ be a decreasing sequence in δ of optimal controls of (4.P^δ(p)) and let $\sigma^{p,*}$ be an optimal control of the problem (4.P (p)). Then

(i) the sequence $\{\sigma^{p,\delta,*}\}$ has a subsequence which converges to an optimal control of the problem (4.P (p)) in the Euclidean norm.

(ii) $$\lim_{\delta \to 0} J(\sigma^{p,\delta,*}) = J(\sigma^{p,*}). \tag{4.47}$$

Proof (i) From the definition of $\mathcal{U}^p$, the sequence of vectors $\{\sigma^{p,\delta,*}\}$ is obviously bounded in the Euclidean norm. Therefore, there exists a subsequence, denoted by $\{\sigma^{p,\bar{\delta},*}\}$, such that $\{\sigma^{p,\bar{\delta},*}\}$ converges to $\bar{\sigma}^p$ in the Euclidean norm. By the continuity of the constraint (4.42) with respect to δ, Assumption 4.A3, and part (i) of Lemma 4.3.1, it is clear that $\bar{\sigma}^p \in \mathcal{F}^p$. We want to prove that $\bar{\sigma}^p$ is also an optimal control of the problem (4.P(p)).

Since $\mathcal{F}^p \subset \mathcal{F}^{p,\bar{\delta}}$ for all $\bar{\delta} > 0$, we have $\sigma^{p,*} \in \mathcal{F}^{p,\bar{\delta}}$ for all $\bar{\delta} > 0$. Thus

$$J(\sigma^{p,\bar{\delta},*}) \leq J(\sigma^{p,*}). \tag{4.48}$$

Using part (ii) of Lemma 4.3.1, we obtain by taking limit as $\bar{\delta} \to 0$ in (4.48) that

$$J(\bar{\sigma}^p) \leq J(\sigma^{p,*}). \tag{4.49}$$

Thus, from (4.49) and the definition of $\sigma^{p,*}$, we conclude that $\bar{\sigma}^p$ is also an optimal control of the problem (4.P(p)).

Proof (ii) For each $p > 0$, the sequence $\{J(\sigma^{p,\delta,*})\}$ is a non–decreasing sequence of δ and has an upper bound $J(\sigma^{p,*})$, thus the whole sequence $\{J(\sigma^{p,\delta,*})\}$ is convergent. Now, from the result of part (i) of this lemma we have

$$\lim_{\bar{\delta} \to 0} J(\sigma^{p,\bar{\delta},*}) = J(\sigma^{p,*}) \tag{4.50}$$

where $\{\sigma^{p,\bar{\delta},*}\}$ is the subsequence of $\{\sigma^{p,\delta,*}\}$ as described in part (i) of the proof of this lemma.

Thus, we have,

$$\lim_{\delta \to 0} J(\sigma^{p,\delta,*}) = \lim_{\bar{\delta} \to 0} J(\sigma^{p,\bar{\delta},*}) = J(\sigma^{p,*}).$$

Before we proceed further, we need the following additional assumption:

(4.A8) $$\lim_{\delta \to 0} J(u^{\delta,*}) = J(u^*),$$

where $u^{\delta,*}$ and u^* are optimal control of the problem (4.P^δ) and (4.P) respectively.

Remark 4.5.1 The result of (4.A8) (for bounded measurable controls) <u>cannot</u> proved by using an argument similar to that given for Lemma 4.4.1 (for control vectors). This is due to the fact that the sequence $\{u^{\delta,*}\}$ does not necessarily possess any accumulation point, either in the a.e. topology or L_∞ topology. However, the result of (4.A8) remain valid in almost all real–life problems. In fact, our formulation of the problem may be wrong if (4.A8) is not satisfied.

Remark 4.5.2 The problems (4.P^δ(p)) and (4.P^δ) involve only continuous state inequality constraints and terminal state inequality constraints. Hence they have the same structure as the problem (4.P (p)) and (4.P) in Ref. 103 respectively. Thus, the convergence result of Theorem 4.2 of Ref. 103 can be applied to the problem (4.P^δ(p)) as $p \to \infty$. In other words, for each $\delta > 0$, we have

$$\lim_{p \to \infty} J(u^{p,\delta,*}) = J(u^{\delta,*}), \tag{4.51}$$

where $u^{p,\delta,*}$ and $u^{\delta,*}$ are optimal controls of the Problem $(4.P^{\delta}(p))$ and $(4.P^{\delta})$ respectively.

Theorem 4.5.1 Let $u^{p,\delta,*}$, $u^{\delta,*}$ and $u^{p,*}$ be the optimal control of the approximate problem $(4.P^{\delta}(p))$, $(4.P^{\delta})$ and $(4.P(p))$ respectively. Let u^{*} be an optimal control of the Problem $(4.P)$. Then

$$\lim_{p \to \infty} J(u^{p,*}) = J(u^{*}). \tag{4.52}$$

Proof From part (ii) of Lemma 4.5.1 and Assumption (4.A8), it is clear that for any given integer $p > 0$, $J(u^{p,\delta,*}) - J(u^{\delta,*})$, considered as a function of δ, is continuous at some interval $[0,\delta_0]$, where $\delta_0 > 0$. From Remark 4.5.2 it is clear that for any $\delta \in [0,\delta_0]$, $\{J(u^{p,\delta,*}) - J(u^{\delta,*})\}$ considered as a sequence of p, converges monotonically to zero as $p \to \infty$. Hence by Dini Theorem (cf. Ref. 80), $\{J(u^{p,\delta,*}) - J(u^{\delta,*})\}$ converges to zero uniformly with respect to δ, $\delta \in [0,\delta_0]$. In other words, for any given $\varepsilon > 0$, there exists $p_0 > 0$ such that

$$|J(u^{p,\delta,*}) - J(u^{\delta,*})| \leq \varepsilon/3 \tag{4.53}$$

for all $p \geq p_0$ and $0 < \delta \leq \delta_0$.

Again by part (ii) of Lemma 4.5.1 and Assumption (4.A8), for each $p \geq p_0$, there exists $\delta_1(p)$, $0 < \delta_1(p) \leq \delta_0$, such that

$$|J(u^{p,\delta_1(p),*}) - J(u^{p,*})| \leq \varepsilon/3 \tag{4.54}$$

and

$$|J(u^{\delta_1(p),*}) - J(u^*)| \leq \epsilon/3 \tag{4.55}$$

for all $0 < \delta \leq \delta_1(p)$. Hence by substituting $\delta = \delta_1(p)$ in (4.53), we deduce from (4.53), (4.54) and (4.55) that

$$|J(u^{p,*}) - J(u^*)| \leq \epsilon \tag{4.56}$$

for all $p \geq p_0$. This completes the proof of this theorem.

Theorem 4.5.2 Let u^* be an optimal control of the problem (4.P). Suppose that the sequence $\{u^{p,*}\}_{p=1}^{\infty}$ of optimal controls of the approximate problem $\{(4.P(p))\}_{p=1}^{\infty}$ converge almost everywhere on $[0,T]$ to a control $\bar{u}$. Then $\bar{u}$ is also an optimal control of the problem (4.P).

Proof Since $u^{p,*} \to \bar{u}$ a.e. on $[0,T]$, it follows from part (ii) of Lemma 4.3.1 that

$$\lim_{p \to \infty} J(u^{p,*}) = J(\bar{u}). \tag{4.57}$$

From part (i) of Lemma 4.3.1 and Assumptions (4.A2) – (4.A4), it is clear that $\bar{u}$ is also a feasible control of the problem (4.P). However from (4.57) and Theorem 4.5.1, we have

$$J(\bar{u}) = J(u^*). \tag{4.58}$$

Thus, the proof of this theorem is complete.

4.6 Illustrative Examples

Example 4.6.1 Consider the problem of minimizing

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

subject to the delayed differential equation

$$\dot{x}(t) = x(t) \sin(x(t)) + x(t-1) + u(t), \qquad t \in [0,2]$$
$$x(t) = 1.0, \qquad t \in [-1,0]$$

together with the terminal state equality constraint

$g_1(x(2)) = -119.854 + 22x(2) - x^2(2) = 0$

and the continuous state inequality constraint

$g_2(x(t)) = 114.6 - 8t - x^2(t) \geq 0,\ t \in [0,2]$.

The problem has been computed using the constraint transcription of Section 4.4 to handle the continuous state inequality constraint (4.5c). Let $g_{2,\epsilon}$ be constructed from g_2 according to (4.22). We can then present two measures of constraint violation. The results obtained by using Algorithm 4.4.1 are tabulated in Table 4.6.1. The constrained state and the optimal control produced are given in Figures 4.6.1a and 4.6.1b, respectively.

In Figure 4.6.1a, the graph of $x(t)$ lies below the curve $\sqrt{114.6 - 8t}$ for all $t \in [0,1]$. This implies that $x(t) \leq \sqrt{114.6 - 8t}$ for all $t \in [0,1]$. Hence it is clear that $g_2(x(t)) \geq 0$ for all $t \in [0,1]$.

ϵ	γ	$J(u)$	$g_1(x(2))$	$\int_0^2 g_{2,\epsilon}\,dt$	$\int_0^2 \min[g_2,0]\,dt$
10^{-3}	10^{-4}	164.584	-0.20×10^{-4}	-0.1×10^{-3}	0
10^{-4}	10^{-5}	164.510	-0.83×10^{-4}	-0.1×10^{-4}	0
10^{-5}	10^{-6}	164.306	-0.46×10^{-5}	-0.1×10^{-5}	0
10^{-6}	10^{-7}	164.306	-0.11×10^{-5}	-0.1×10^{-6}	0
10^{-7}	10^{-8}	164.306	-0.11×10^{-5}	-0.1×10^{-7}	0

Table 4.6.1 Results generated by Algorithm 4.4.1 for Example 4.6.1.

Example 4.6.2 Consider the problem of minimizing

$$J(u) = \int_0^1 [x_1^2(t) + x_2^2(t) + 0.005\, u^2(t)]\, dt$$

subject to the delayed differential equations

$$\dot{x}_1(t) = x_2(t - \tfrac{1}{3})$$
$$\dot{x}_2(t) = -x_2(t) + u(t)$$
$$x_1(0) = 0$$
$$x_2(0) = -1,\ t \in [-\tfrac{1}{3}, 0]$$

together with the terminal state inequality constraint

$$g_1(x(1)) = -x_2(1) \geq 0$$

and the continuous state inequality constraint

$$g_2(x(\cdot)) = -x_2^2(t) + 8(t-0.5)^2 - 0.5 \geq 0, \qquad t \in [0,1].$$

The problem has been solved by using exactly the same method as that described for Example 4.6.1. The results obtained are tabulated in Table 4.6.2. The constrained state $x_2(t)$ and the optimal control produced are given in Figures 4.6.2a and 4.6.2b, respectively.

In Figure 4.6.2a, the graph of $x_2(t)$ also lies below the curve $\sqrt{8(t-0.5)^2 - 0.5}$ for all $t \in [0,1]$. Thus $x_2(t) \leq \sqrt{8(t-0.5)^2 - 0.5}$ for all $t \in [0,1]$. Hence $g_2(x(t)) \geq 0$ for all $t \in [0,1]$.

ϵ	γ	$J(u)$	$g_1(x(2))$	$\int_0^1 g_{2,\epsilon}\,dt$	$\int_0^1 \min[g_2, 0]\,dt$
10^{-3}	10^{-4}	0.5047	0.30×10^{-5}	-0.1×10^{-3}	0
10^{-4}	10^{-5}	0.5029	0.53×10^{-5}	-0.1×10^{-4}	0
10^{-5}	10^{-6}	0.5055	0.35×10^{-2}	-0.1×10^{-5}	0
10^{-6}	10^{-7}	0.5055	0.35×10^{-2}	-0.1×10^{-6}	0

Table 4.6.2 Results generated by Algorithm 4.4.1 for Example 4.6.2.

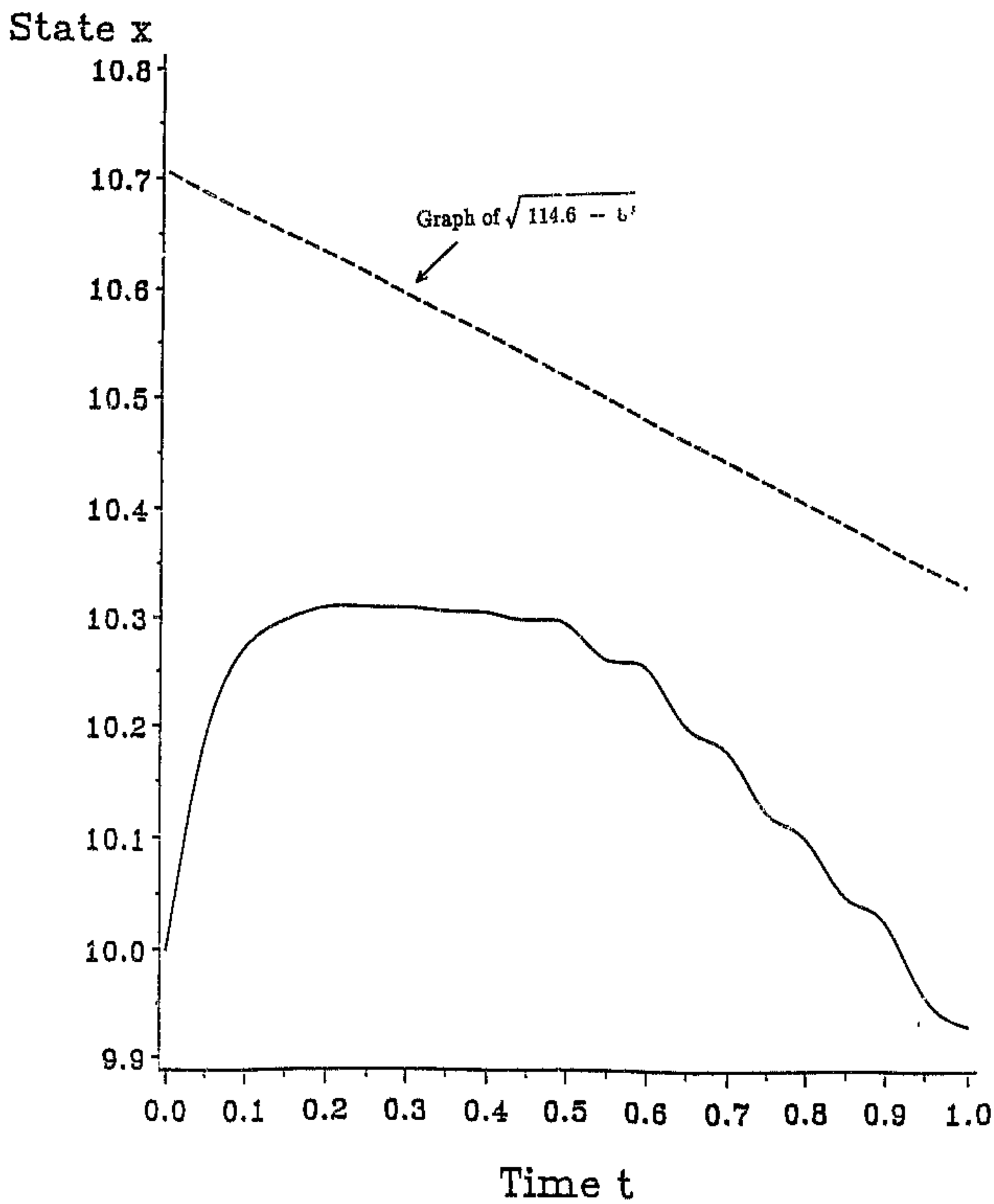

Figure 4.6.1a Computed Optimal State for Example 4.6.1.

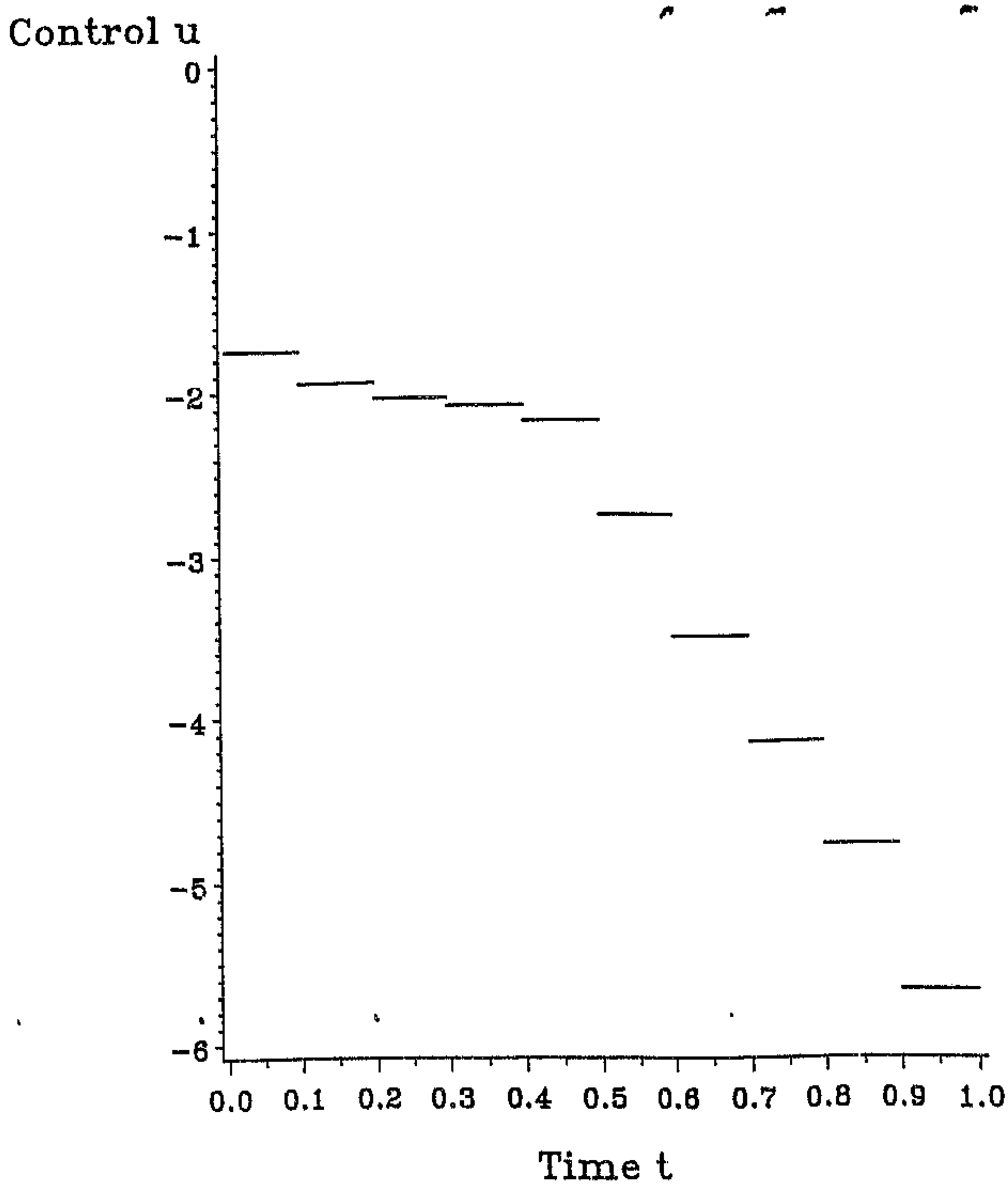

Figure 4.6.1b Computed Optimal Control for Example 4.6.1.

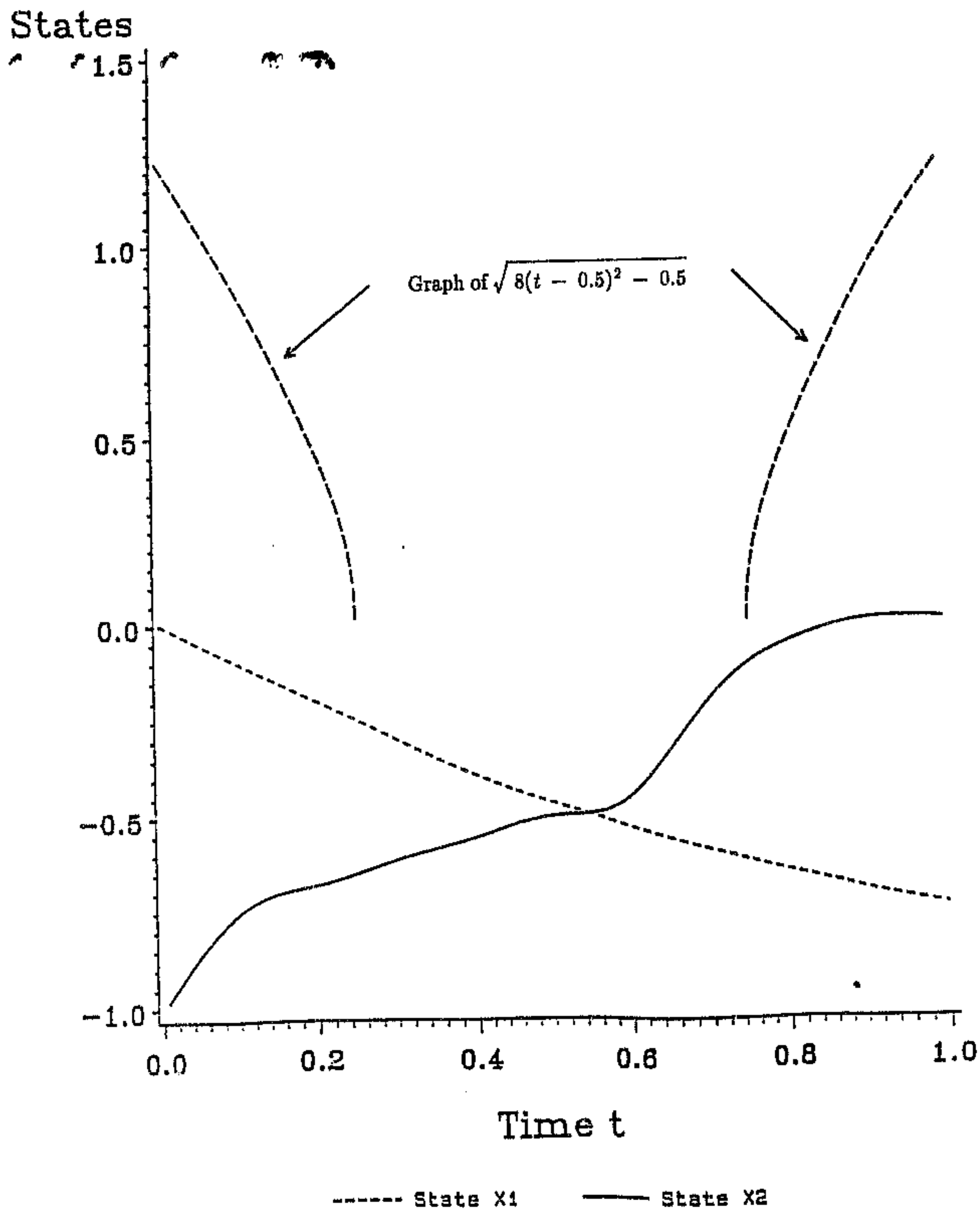

Figure 4.6.2a Computed Optimal States for Example 4.6.2.

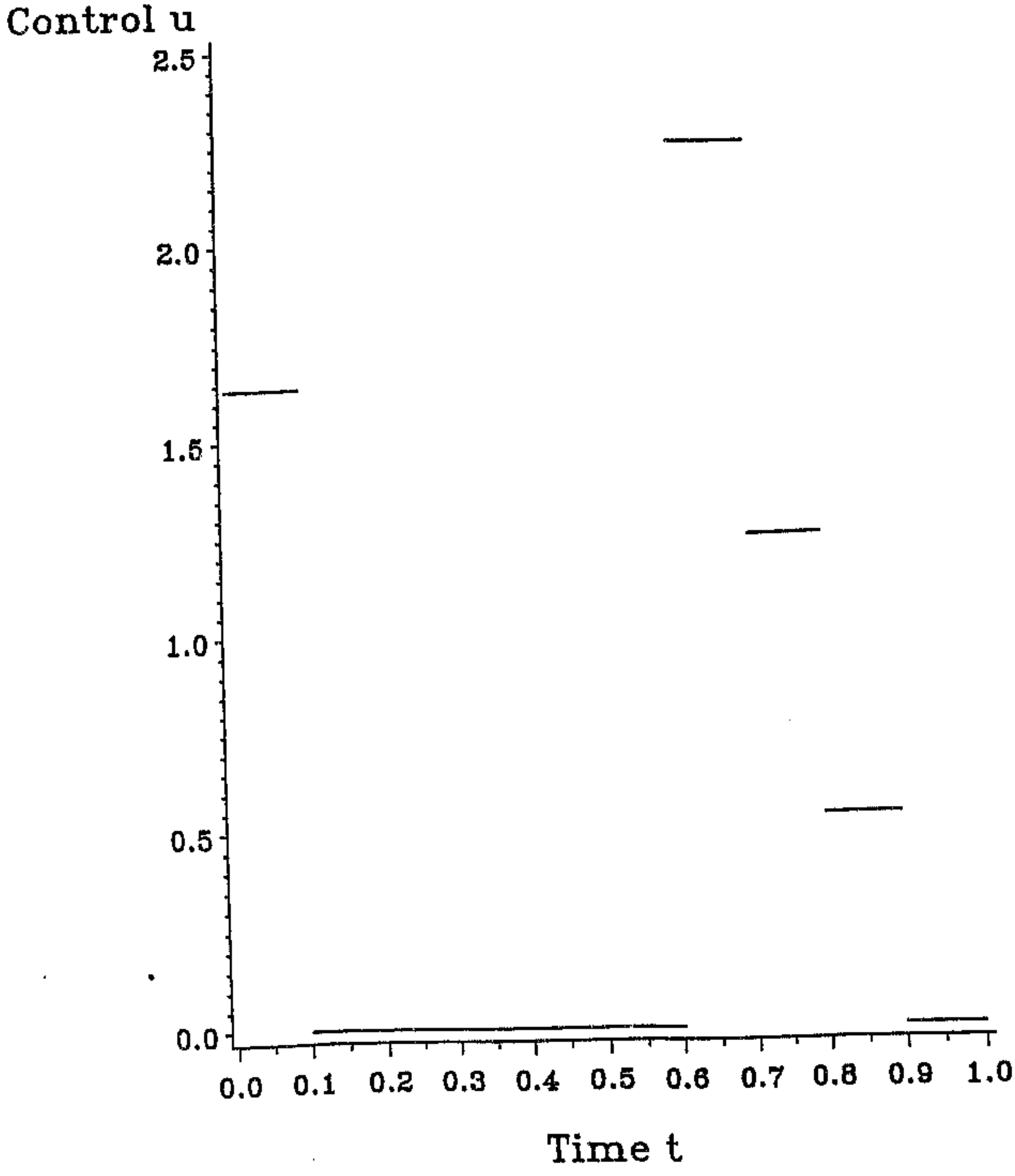

Figure 4.6.2b Computed Optimal Control for Example 4.6.2.

CHAPTER V

CONTROL PARAMETERIZATION METHOD FOR A CLASS OF NONLINEAR TIME–DELAYED OPTIMAL CONTROL PROBLEMS WITH A COST OF CHANGING CONTROL

5.1 Introduction

Similar to Chapter IV, the problem considered in this chapter also consists of a nonlinear time–delayed system, together with terminal state inequality constraints and continuous state inequality constraints. However, the cost functional of this Chapter is the sum of not only the integral cost and the terminal cost, but also the full variation of the control. Moreover, for the sake of simplicity, we have omitted terminal state equality constraints in this chapter.

The main contribution of this chapter is that we have extended all the results of Ref. 104 to the time–delayed system.

In Section 5.2, we describe the optimal control problem and assume certain conditions.

In Section 5.3, the technique of control parameterization is used to obtain a sequence of approximate problems (5.P (p)).

The aim of Section 5.4 is to 'smooth' the term involving the full variation of the control of the problem (5.P (p)), since this term is not a smooth function in the ℓ_1–norm.

In Section 5.5, we use the constraint transcription similar to those used in Refs. 103, 104 and Chapter IV of this thesis to convert the continuous state inequality constraints of the problem (5.P (p)) into integral constraints, so that the resulting integral constraints can satisfy the usual constraint qualification. Similar to Chapter IV, the modified problem obtained in such a way can be easily solved by Algorithm A2 of Ref. 103.

In Section 5.6, the convergence properties of the sequence of approximate optimal controls of the problem (5.P (p)) to the true optimal control is discussed.

In Section 5.7, we solve two numerical examples to illustrate that the technique developed in Ref. 103 works equally well for time-delayed systems.

5.2 Statement of the Problem

Let $x(t)$ satisfy the same differential equation and initial condition as defined in (4.1). Let U be as defined in (4.2). As mentioned in Section 4.2, U is a compact and convex subset of R^r.

The full variation of a function $v = [v_1,\ldots,v_r]^T : [0, T] \longrightarrow R^r$ is defined as:

$$\overset{T}{\underset{0}{V}} v(t) = \sum_{i=1}^{r} \overset{T}{\underset{0}{V}} v_i(t), \tag{5.1}$$

where $\overset{T}{\underset{0}{V}} v_i(t)$ denotes the total variation of the real-valued function $v_i(\cdot)$ on $[0, T]$.

If $\overset{T}{\underset{0}{V}} v(t) < \infty$, then the function $v(\cdot)$ is said to be of bounded variation.

Define

$\mathscr{V} \equiv \{u : u = [u_1,\ldots,u_r]^T$ is a function of bounded variation defined on $[-h, T]$ such that $u(t) \in U$ for all $t \in [0,T]$, and $u(t) = \beta(t)$ for all $t \in [-h,0)\}$

where β is a given piecewise continuous function defined on $[-h,0)$. For each $u \in \mathscr{V}$, let $x(\cdot|u)$ be the solution of the system (4.1) corresponding to the control $u \in \mathscr{V}$. The inequality terminal state constraints and inequality continuous state constraints are specified as follows:

$$\phi_i(x(T|u)) \geq 0, \qquad i = 1,\ldots,N_I \tag{5.2}$$

where ϕ_i, $i = 1,\ldots,N_I$, are given real valued functions defined on R^n, and

$$g_i(t,x(t|u), x(t-h|u)) \geq 0, \forall t \in [0,T], i = 1,\ldots,N, \tag{5.3}$$

where g_i, $i = 1,\ldots,N$, are given real valued functions defined on $[0,T] \times R^{2n}$.

Let $\mathscr{G}$ be the class of all those control functions in $\mathscr{V}$ such that the constraints (5.2) and (5.3) are satisfied.

We may now state our optimal control problem as follows:

Problem (5.Q) Given the system (4.1), find a control $u \in \mathscr{G}$ such that the cost functional

$$J(u) = \Phi_0(x(T|u)) + \int_0^T \mathscr{L}_0(t,x(t|u), x(t-h|u), u(t), u(t-h))\, dt + c \bigvee_0^T u(t) \tag{5.4}$$

is minimized over $\mathscr{G}$, where c is a positive weighting factor, and Φ_0, $\mathscr{L}_0$ are given real–valued functions.

The following conditions are assumed throughout:

(5.A1) f satisfies the conditions appearing in (4.A1);

(5.A2) for each $i = 1,...,N_I$, $\phi_i : R^n \longrightarrow R$ is continuously differentiable;

(5.A3) for each $i = 1,...,N_S$, $g_i : [0,T] \times R^{2n} \longrightarrow R$ is continuously differentiable;

(5.A4) $\Phi_0 : R^n \longrightarrow R$ is continuously differentiable;

(5.A5) $\mathscr{L}_0$ satisfies the conditions appearing in (4.A6).

Define

$$q_0(u) = \Phi_0(x(T|u)) + \int_0^T \mathscr{L}_0(t, x(t|u), x(t-h|u), u(t), u(t-h))\, dt \tag{5.5}$$

from (5.4) and (5.5), we have

$$J(u) = q_0(u) + c \overset{T}{\underset{0}{\bigvee}} u(t) \tag{5.6}$$

If $u, v \in \mathscr{V}$ are such that

$$u(t) = v(t), \text{ a.e. in } [0,T] \tag{5.7}$$

it is clear from Lemma 4.3.1 that

$$q_0(u) = q_0(v) \tag{5.8}$$

However, it is not necessarily true that

$$\bigvee_0^T u(t) = \bigvee_0^T v(t).$$

Consider a control $v \equiv [v_1,\ldots,v_r]^T \in \mathscr{V}$. Let

$$t_{i,1}, t_{i,2}, \ldots$$

be the discontinuity points of the i^{th} component v_i of the control v in $(0,T)$. Let

$$A_i = \bigcup_{k=1}^{\infty} \{t_{i,k}\}$$

From theorem 2.9.3 of Ref. 104, we note that A_i and hence $\bigcup_{i=1}^{r} A_i$; contains at most a countable number of points.

We now construct a function $u = [u_1,\ldots,u_r]^T$ from v as follows:

$$u_i(t) = \begin{cases} v_i(t) & \text{if } t \in [(0,T) \cup (-h,0)] \backslash A_i \\ v_i(t-0) \text{ if } t \in A_i \backslash [\{0\} \cup \{T\}] \\ v_i(t+0) \text{ if } t = 0 \\ v_i(t-0) \text{ if } t = T \end{cases} \tag{5.9}$$

Clearly, $u \in \mathscr{V}$ and is such that

(i) $u(t) = v(t)$ a.e. in $[-h,T]$ (5.10)

and

(ii) $\bigvee_0^T u(t) \leq \bigvee_0^T v(t)$ (5.11)

We call u a minimal bounded variation function of the set $\{v\}$ of functions where $v(t) = u(t)$ a.e. in $[-h, T]$. Clearly,

$$g_0(u) \leq g_0(v). \tag{5.12}$$

Let $\mathscr{U}$ be the set which consists of all such minimal bounded variation functions of $\mathscr{V}$. Elements of $\mathscr{U}$ are called admissible controls and $\mathscr{U}$ is called the class of admissible controls.

Remark 5.2.2 Note that our aim in problem $(5.Q)$ is to minimize the cost functional (5.4). Thus, by (5.12) it suffices to confine our attention to the set $\mathscr{U}$ rather than the set $\mathscr{V}$ in the minimization process.

Let $\mathscr{F}$ be a set which consists of all those elements in $\mathscr{U}$ such that (5.2) and (5.3) are satisfied.

For convenience, problem $(5.Q)$ with $\mathscr{G}$ replaced by $\mathscr{F}$ will be referred to as problem $(5.P)$. Clearly, an optimal control problem to problem $(5.P)$ is also an optimal control to problem $(5.Q)$, and vice versa.

5.3 Control Parameterization

As in Chapter 3, the control parameterization method will be used to solve the problem $(5.P)$. Let $\mathscr{I}_p$, I_j^p, $\mathscr{U}^p$, σ^p and $x(\cdot \mid \sigma^p)$ be as defined in Section 4.3.

Then, the total variation term appearing in the cost functional (5.4) takes the following simple form:

$$\sum_{i=1}^{r} \sum_{k=1}^{n_p-1} \left| \sigma_i^{p,k+1} - \sigma_i^{p,k} \right| \quad . \tag{5.13}$$

Define

$$\Lambda^p = \{\sigma^p \in \mathscr{U}^p : \Phi_i(x(T|\sigma^p)) \geq 0,\ i = 1,\ldots,N_I\} \tag{5.14}$$

and

$$\mathscr{F}^p = \{\sigma^p \in \mathscr{U}^p : g_i(t, x(t|\sigma^p), x(t-h|\sigma^p)) \geq 0, \forall\, t \in [0,T], i = 1,\ldots,N\} \tag{5.15}$$

We may now specify the approximate problem (5.P (p)) as follows:

Problem (5.P (p)) Find a control parameter vector $\sigma^p \in \Omega^p$ such that the cost functional

$$G_0(\sigma^p) = \Phi_0(x(T|\sigma^p)) + \int_0^T \hat{\mathscr{L}}_0(t, x(t|\sigma^p), x(t-h|\sigma^p), \sigma^p)\, dt + c \sum_{i=1}^{r} \sum_{k=1}^{n_p-1} \left| \sigma_i^{p,k+1} - \sigma_i^{p,k} \right| \tag{5.21}$$

is minimized over Ω^p, where $\hat{\mathscr{L}}_0$ is obtained from $\mathscr{L}_0$ in an obvious manner.

5.4 Smoothing the Cost Functional

To solve the problem (5.P) via the control parameterization technique, we need to solve a sequence of finite dimensional optimization problems (5.P (p)). However, the last term in the cost functional of the problem (5.P (p)) is non–differentiable. The aim of this section is to use the idea of Ref. 100 to "smooth" the cost functional. Same as in Ref. 100, we define

$$S_\rho(y) = \begin{cases} |y| & \text{if } |y| > \rho \\ (y^2 + \rho^2)/2\rho & \text{if } |y| \leq \rho \end{cases} \tag{5.22}$$

We now consider the following approximate problem.

Problem (5.P_ρ(p)) The problem (5.P(p)) with the cost functional (5.21) replaced by

$$\begin{aligned} G^p_0(\sigma^p) &= \Phi_0\,(x\,(T|\,\sigma^p)) + \int_0^T \hat{\mathscr{L}}_0\,(t,x\,(t|\,\sigma^p),\, x\,(t-h|\,\sigma^p),\, \sigma^p) \\ &\quad + c\sum_{i=1}^{r}\ \sum_{k=1}^{n_p-1} S_\rho\,(\sigma_i^{p,k+1} - \sigma_i^{p,k}) \\ &= Q_0(\sigma^p) + c\sum_{i=1}^{r}\ \sum_{k=1}^{n_p-1} S_\rho\,(\sigma_i^{p,k+1} - \sigma_i^{p,k}) \end{aligned} \tag{5.23}$$

Theorem 5.4.1 let $\sigma^{p,*}$ and $\sigma^{p,\rho,*}$ be, respectively, optimal solutions to the problem (5.P (p)) and problem (5.P_ρ(p)). Then

$$0 \leq G_0(\sigma^{p,\rho,*}) - G_0(\sigma^{p,*}) \leq c\,[r\,(n_p - 1)]\rho/2. \tag{5.24}$$

Proof The proof is exactly the same as that given for Theorem 4.1 of Ref. 104.

5.5 Constraint Approximation

For each $\sigma^p \in \mathscr{U}^p$, let $G_{i,\epsilon}(\sigma^p)$ be as defined in (4.22). Same as in Section 4.4, we define the following related approximate problem $(5.P_{\rho,\epsilon}(p))$. The problem $(5.P_\rho(p))$ with (5.15) replaced by

$$G_{i,\epsilon}(\sigma^p) = 0, \qquad i = 1,\ldots,N \tag{5.25}$$

Problem $(5.P_{\rho,\epsilon,\gamma}(p))$ The problem $(5.P_\rho(p))$ with (5.15) replaced by

$$\gamma + G_{i,\epsilon}(\sigma^p) \geq 0, \qquad i = 1,\ldots,N_S \tag{5.26}$$

Now, let $\mathscr{F}_\epsilon^p$ be the feasible region of $(5.P_{\rho,\epsilon}(p))$, defined by

$$\mathscr{F}_\epsilon^p = \{\sigma^p \in \Lambda^p : G_{i,\epsilon}(\sigma^p) = 0, \qquad i = 1,\ldots,N\} \tag{5.27}$$

Now, we need to make the following assumption:

(5.A6) Exactly the same assumption as that given for (4.A7).

Remark 5.5.1 By using exactly the same assumption as (4.A7), it is clear that for each fixed $\rho > 0$, all the convergence results given in Lemma 4.4.1 will still remain valid for the problem $P_{\rho,\epsilon}(p)$ as $\epsilon \to 0$.

Remark 5.5.2 In view of Lemma 4.4.2, it is clear that for any given $\epsilon > 0$, there exists a $\overline{\gamma}(\epsilon) > 0$ such that for all γ, $0 < \gamma < \overline{\gamma}(\epsilon)$, any feasible control vector $\sigma^p_{\rho,\epsilon,\gamma}$ of problem $(5.P_{\rho,\epsilon,\gamma}(p))$, i.e.

$$\gamma + G_{i,\epsilon}(\sigma^{p}_{\rho,\epsilon,\gamma}) \geq 0, \qquad i = 1,\ldots,N \tag{5.28a}$$

$$\phi_i\,(x\,(T|\,\sigma^{p}_{\rho,\epsilon,\gamma})) \geq 0, \qquad i = 1,\ldots,N_I \tag{5.28b}$$

is also a feasible control vector of the problem $(5.P_\rho(p))$. Thus, Algorithm 4.4.1 can be used to generate a sequence of sc' 'ons to problem $(5.P_\rho(p))$.

Theorem 5.5.1 Let $\{\sigma^{p,\rho,*}_{\epsilon,\gamma(\epsilon)}\}$ be a sequence in ϵ of the sub–optimal control vectors generated by Algorithm 4.4.1. Then

$$\lim_{\epsilon\to 0} G^{\rho}_{0}(\sigma^{p,\rho,*}_{\epsilon,\gamma(\epsilon)}) = G^{\rho}_{0}(\sigma^{p,\rho,*}),$$

where $\sigma^{p,\rho,*}$ is an optimal control vector of the problem $(5.P_\rho(p))$.

Furthermore, any accumulation point of $\{\sigma^{p,\rho,*}_{\epsilon,\gamma(\epsilon)}\}$ is a solution of the problem $(5.P_\rho(p))$.

Proof The proof is similar to that given for Theorem 4.4.1, except that Lemma 4.4.1 is being replaced by Remark 5 5.1 in the proof of this theorem.

Remark 5.5.3 The method for calculating the gradient of the cost functional and that of the constraints of the problem $(5.P_{\rho,\epsilon,\gamma}(p))$ is similar to that given for the problem $(4.P_{\epsilon,\gamma}(p))$ (cf. Remark 4.4.2 – Remark 4.4.5).

5.6 Some Convergence Results

Similar to Section 4.5, we shall now investigate some convergence properties of the optimal control of the approximate problem $\{5.P(p)\}$ to the true optimal control.

To continue, let $\mathring{\mathscr{F}}$ be the interior of the set $\mathscr{F}$ in the sense that the constraints (5.2) and (5.3) are satisfied as strict inequalities.

Similar to Section 6 of Ref. 104, we assume that the following condition is satisfied:

(5.A7) Let $u^* \in \mathscr{F}$ be an optimal control of problem $(5.P)$. Then, for any $\epsilon > 0$, there exists a control $u \in \mathring{\mathscr{F}}$ such that

$$0 \leq J(u) - J_0(u^*) \leq \epsilon.$$

The next theorem is a direct consequence of Theorem 6.2 of Ref. 104.

Theorem 5.6.1 Let $\{u^{p,*}\}$ be a sequence of optimal control of the approximate problem $(5.P\,(p))$. Then $\{u^{p,*}\}$ has a subsequence $\{u^{p(\ell),*}\}$ which converges to a control $\bar{u}$ a.e. on $[-h, T]$, and $\bar{u}$ is also an optimal control of the problem $(5.P)$.

Remark 5.6.1 Note that the above convergence result is much stronger than that given for Theorem 4.5.2.

5.7 Illustrative Examples

For illustration, we solved the approximate problems ($5.P_{10}$-4(10)) for each of the two examples below. The first example is a simple linear time–lag optimal control problem, where an additional cost is introduced for the variation of the control. The second example is a nonlinear time–lag optimal control problem, where the computation can be regularized by a small penalty on a cost of changing control.

Example 5.7.1 Consider the problem of minimizing

$$J(u) = \int_0^1 \{-6\, x_1(t) - 12\, x_2(t-0.2) + 3\, u_1(t) + u_2(t)\}\, dt$$
$$+ c\,\{\bigvee_0^1 u_1(t) + \bigvee_0^1 u_2(t)\}$$

subject to

$$\dot{x}_1(t) = u_2(t), \qquad 0 \leq t \leq 1$$
$$\dot{x}_2(t) = -\, x_1(t-0.2) + u_1(t), \qquad 0 \leq t \leq 1$$

with initial conditions:

$$x_1(t) = 1, \qquad -0.2 \leq t \leq 0$$
$$x_2(t) = 0, \qquad -0.2 \leq t \leq 0$$

and constraints:

$$|\, u_i(t)| \leq 10, \qquad i = 1,2.$$

Figures 5.7.1a to 5.7.1i show graphs of the two optimal control functions and the two optimal state functions for three different values of the penalty parameter ($c = 0.01$, 0.1, 1).

This example is very similar to that considered in Example 5.7.1 of Ref. 104, with the exception of the term involving time–lag. In Example 5.7.1 of Ref. 104, the linear optimal control changes from a bang–bang control to a control with no variation as the penalty term changes from zero to a large number. This interesting property also holds for this example. That is, as the penalty increases, the control u_1 changes from the nearly bang–bang solution to an almost constant solution (see Figures 5.7.1a to 5.7.1c) and the control u_2 changes from the bang–bang solution to smoother functions (see Figures 5.7.1d to 5.7.1f). Furthermore, the state functions are also smoothed and vary less when the penalty terms get larger (see Figures 5.7.1g to 5.7.1i).

Example 5.7.2 Consider the problem of minimizing

$$J(u) = \int_0^1 e^{-t}\,[1.4\{1 + 0.25\,(1 - e^{-6t})\}\,x(t) - 0.1923\,u(t - 0.1)]\,dt + c \mathop{\bigvee}_0^1 u(t),$$

subject to

$$\dot{x}(t) = 0.5\,\{1 - u(t - 0.1)\,x(t) - x^2(t)\}, \qquad 0 \le t \le 1$$

with initial condition:

$$x(0) = 0.45$$

and constraints:

$x(t) \geq 0.44$

$0 \leq u(t) \leq u_{max}$.

Graphs of the optimal control and the optimal state functions for three different values of the penalty parameter ($c = 0$, 0.001, 0.01) are presented in Figures 5.7.2a to 5.7.2f (where the control is bounded by $u_{max} = 2$) and 5.7.2g to 5.7.2ℓ (where the control is essentially allowed to be unbounded by setting $u_{max} = 10$).

With no variation term ($c = 0$), the control do not vary smoothly, however. even with a small amount of penalty term $c = 0.001$ and 0.01, the control functions are smoothed (see Figures 5.7.2a to 5.7.2c and 5.7.2g to 5.7.2i). Like the previous example, the state functions can be further smoothed if we increase the value of the penalty parameters (see Figures 5.7.2d to 5.7.2f and 5.7.2j to 5.7.2ℓ).

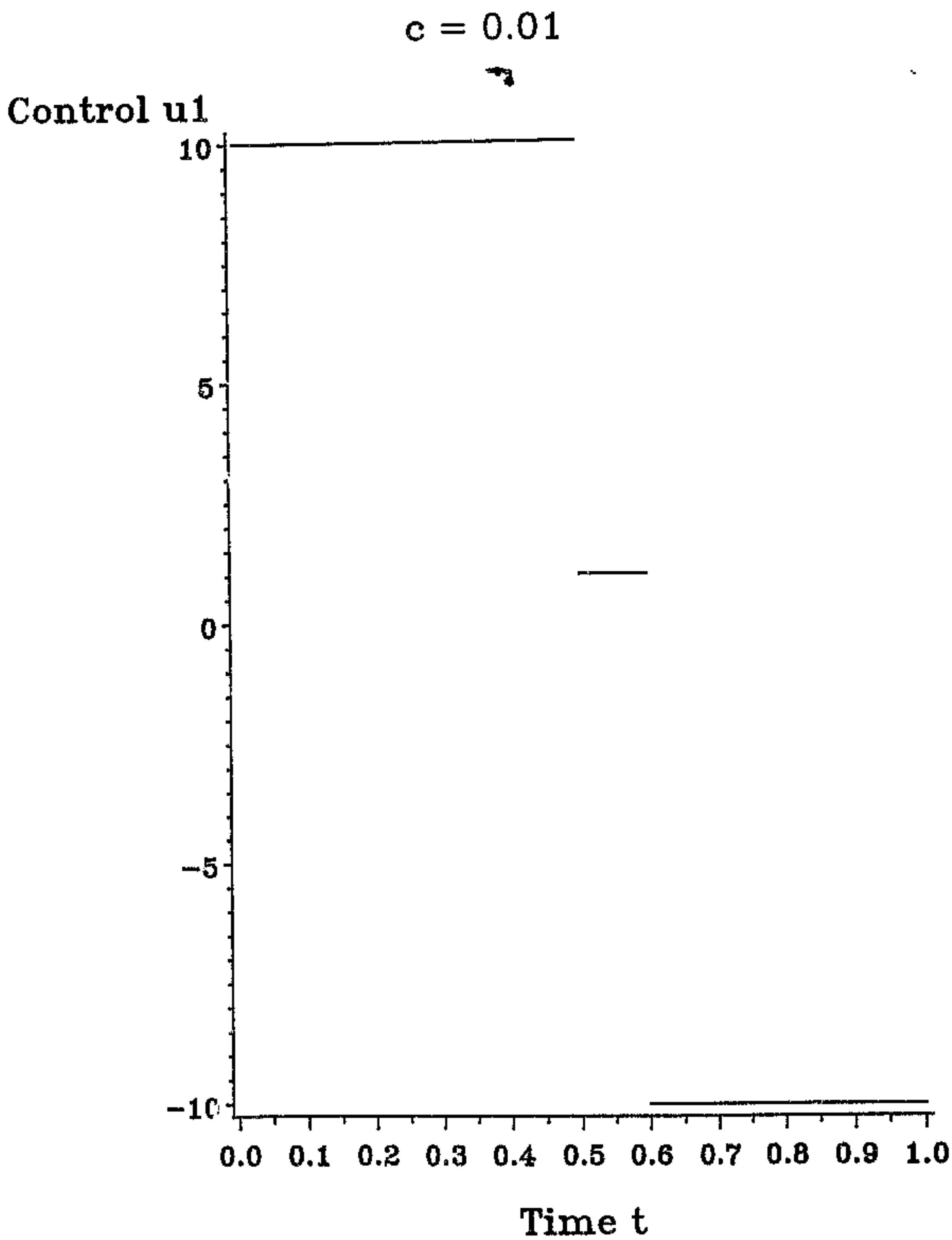

Figure 5.7.1a Computed Optimal Control (u 1) for Example 5.7.1.

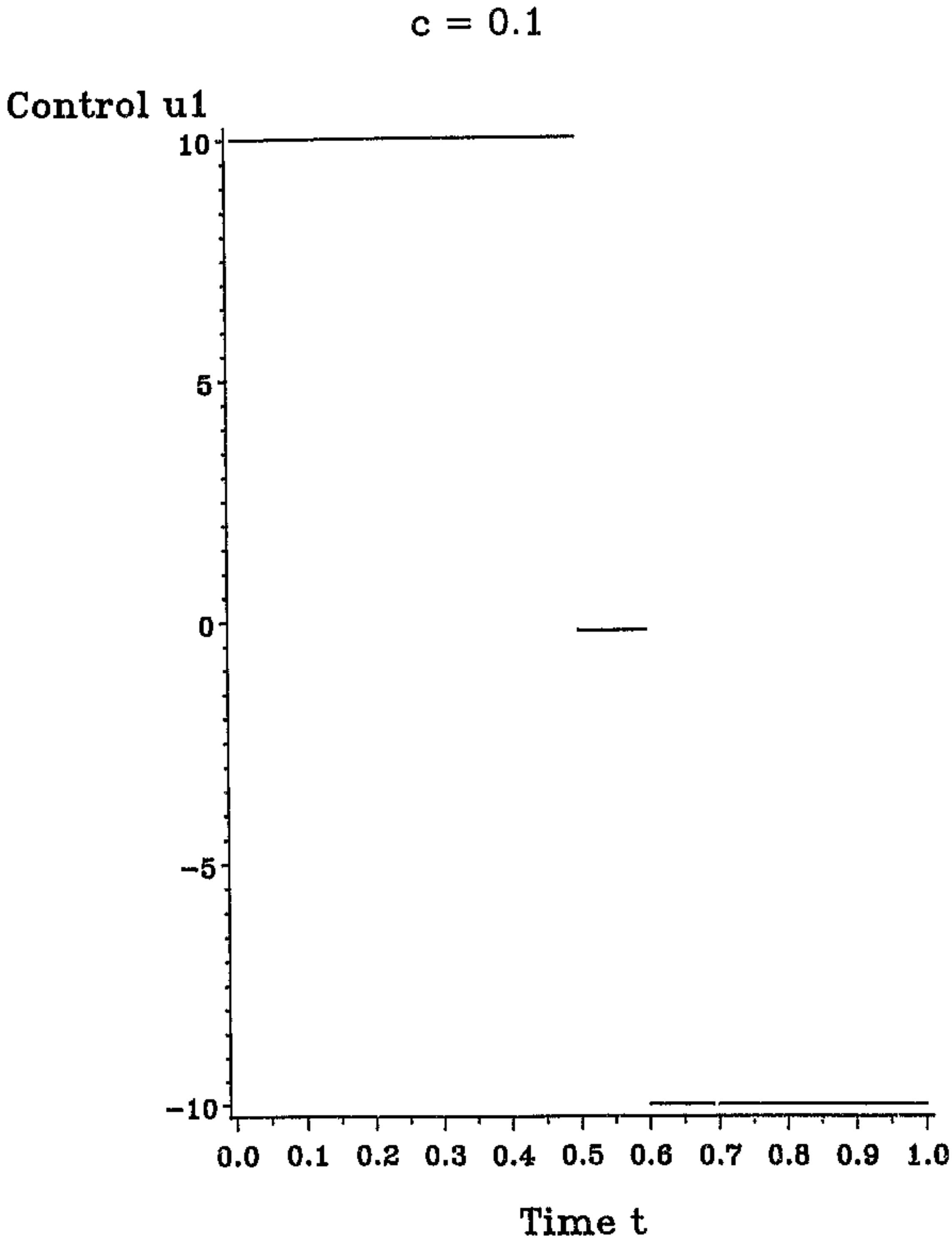

Figure 5.7.1b Computed Optimal Control (u 1) for Example 5.7.1

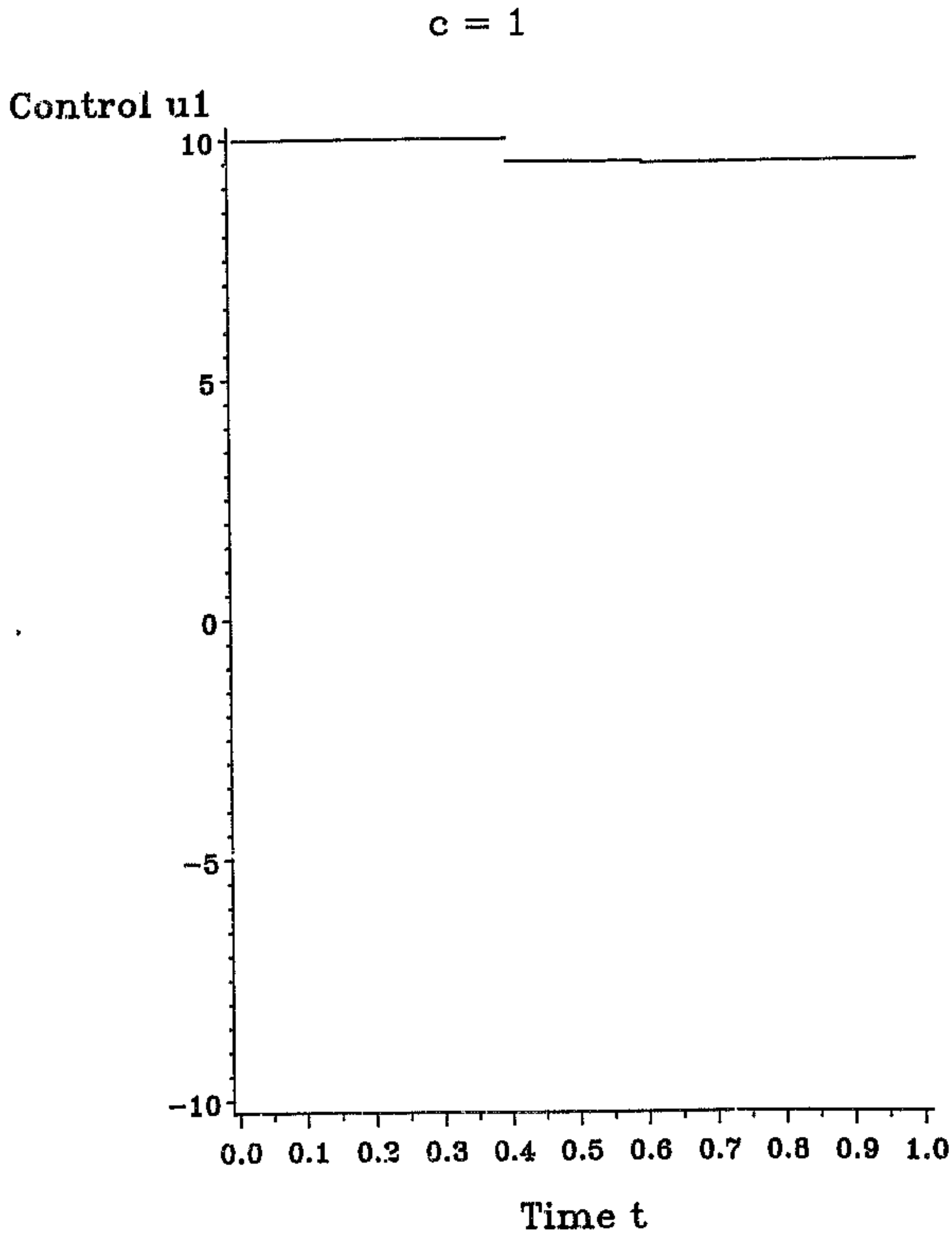

Figure 5.7.1c Computed Optimal Control (u 1) for Example 5.7.1.

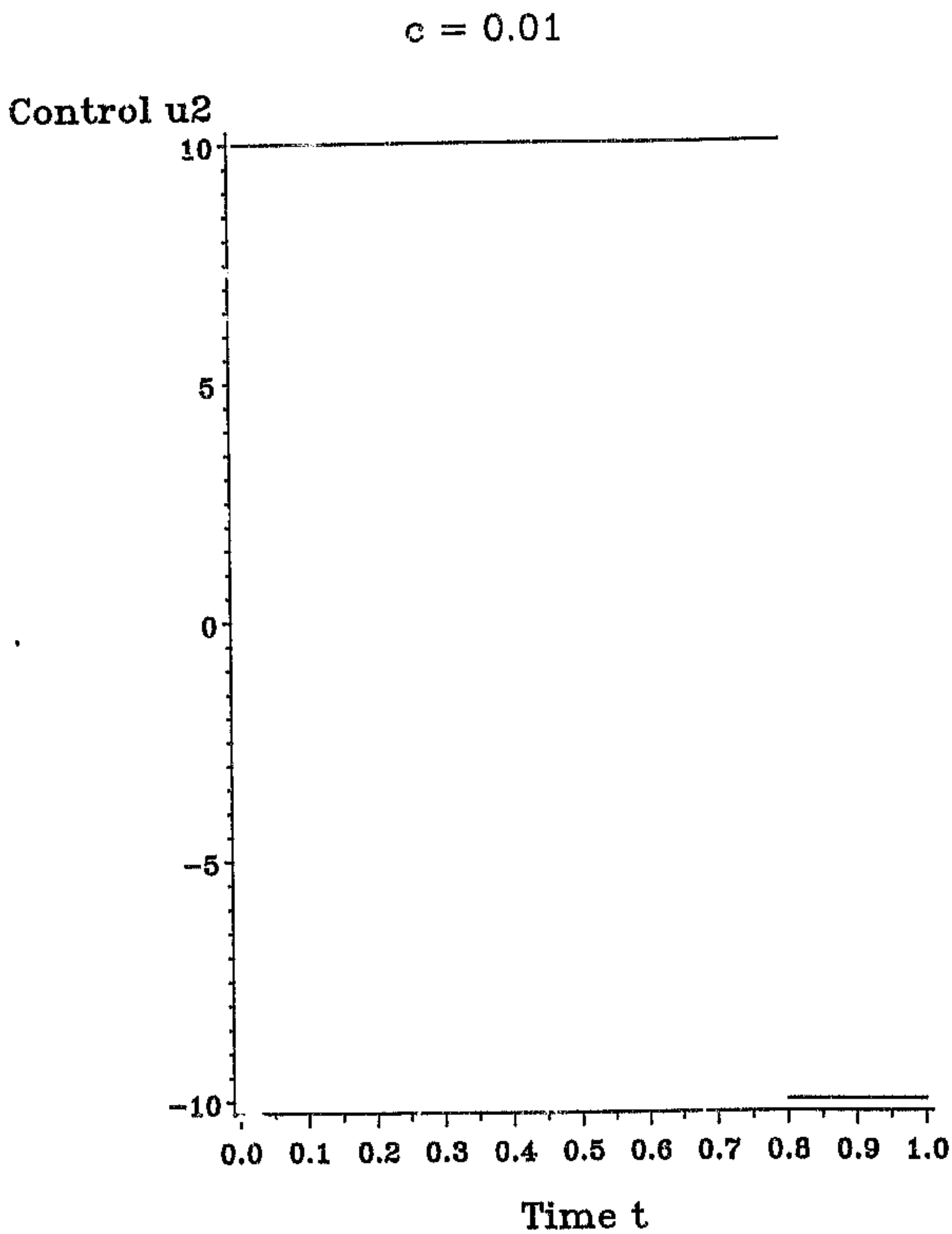

Figure 5.7.1d Computed Optimal Control ($u2$) for Example 5.7.1.

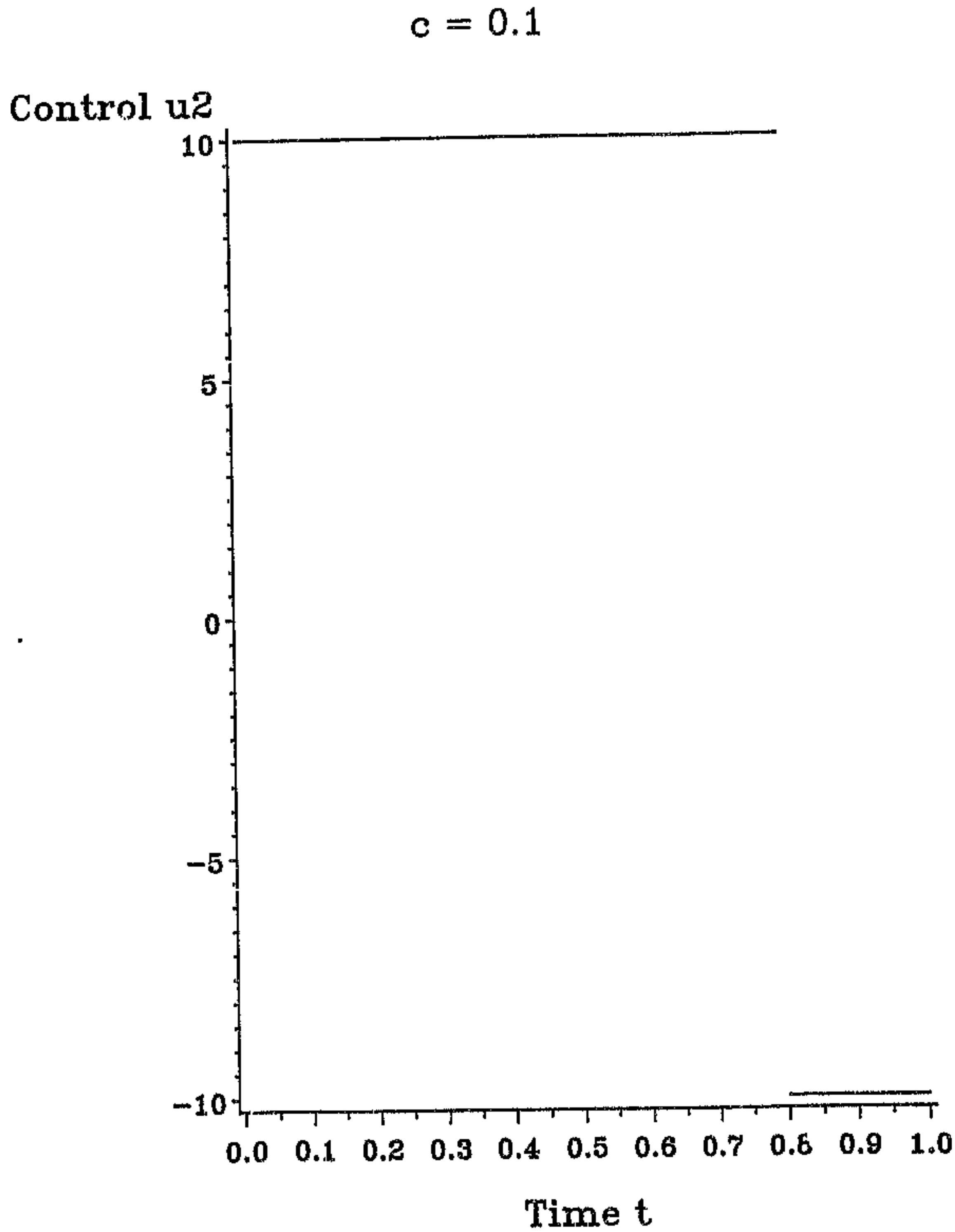

Figure 5.7.1c Computed Optimal Control (u 2) for Example 5.7.1.

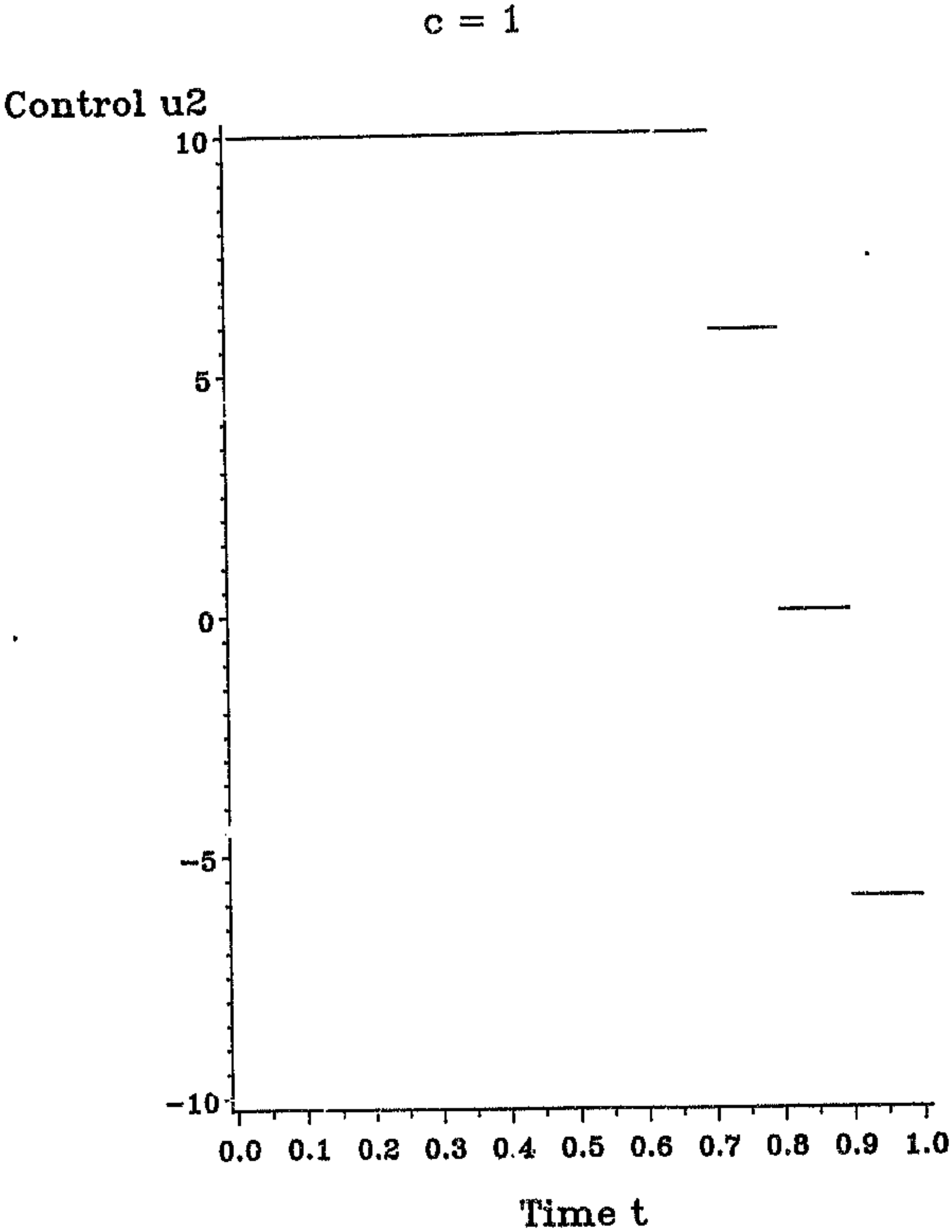

Figure 5.7.1f Computed Optimal Control ($u2$) for Example 5.7.1.

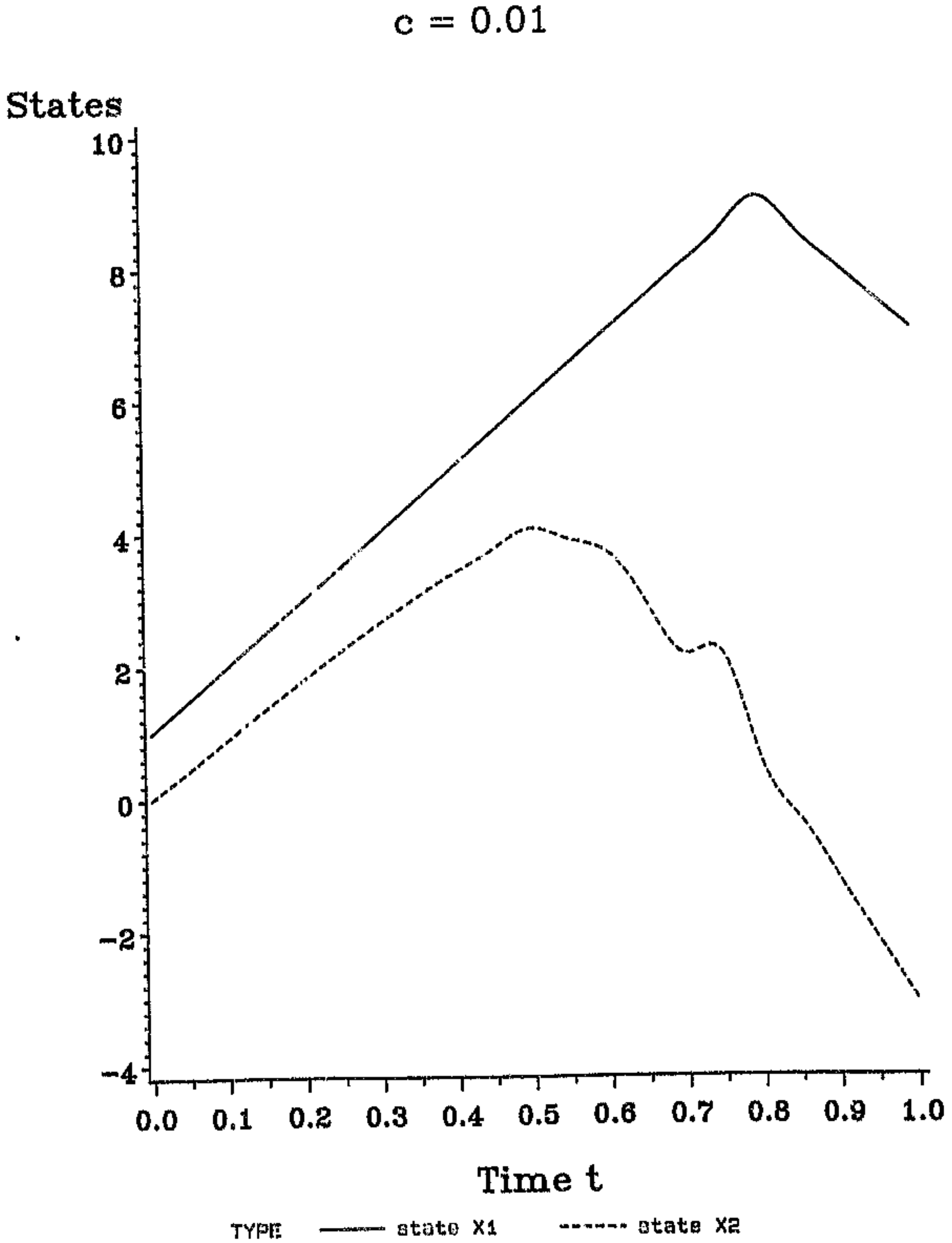

Figure 5.7.1g Computed Optimal States for Example 5.7.1.

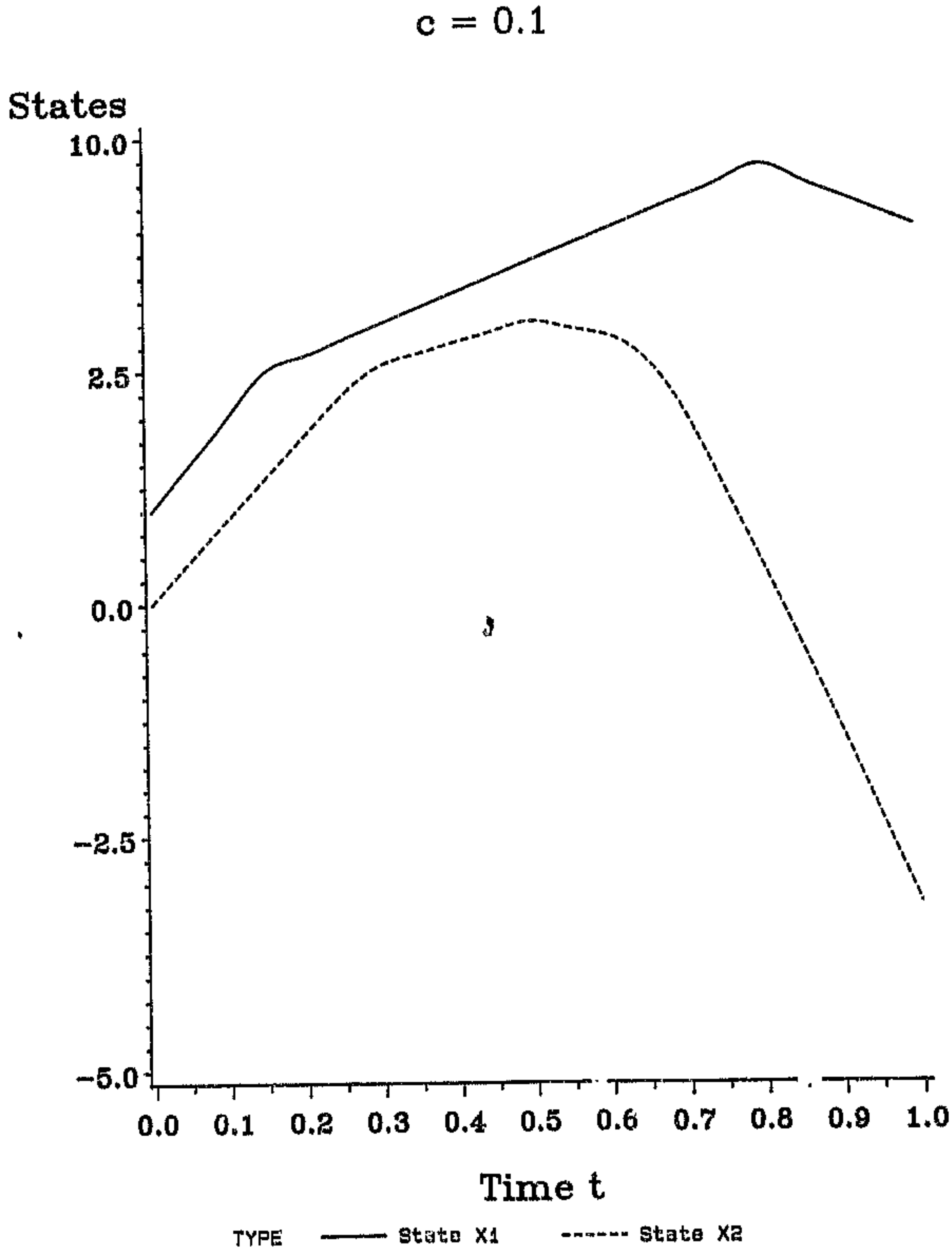

Figure 5.7.1h Computed Optimal States for Example 5.7.1.

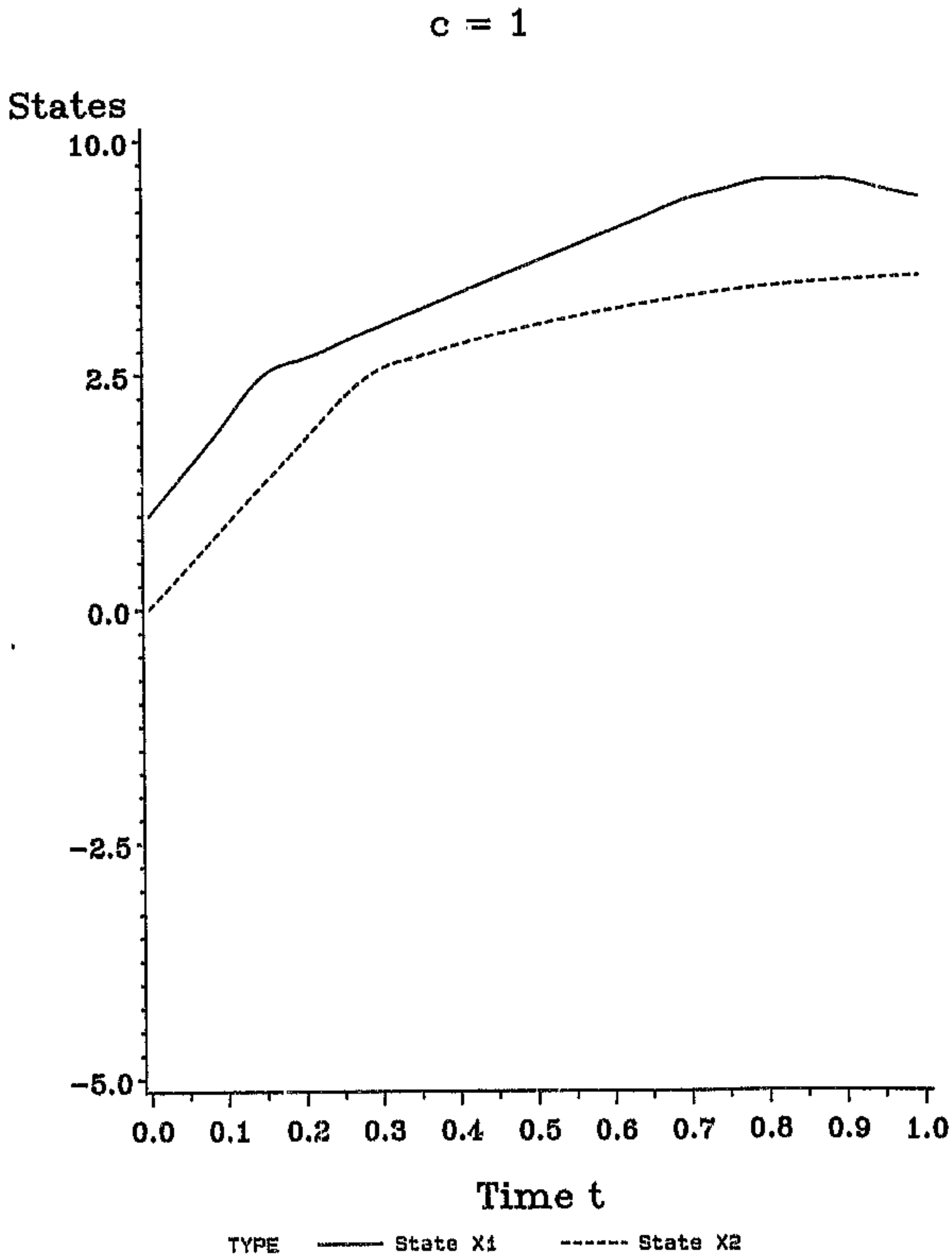

Figure 5.7.1i Computed Optimal States for Example 5.7.1.

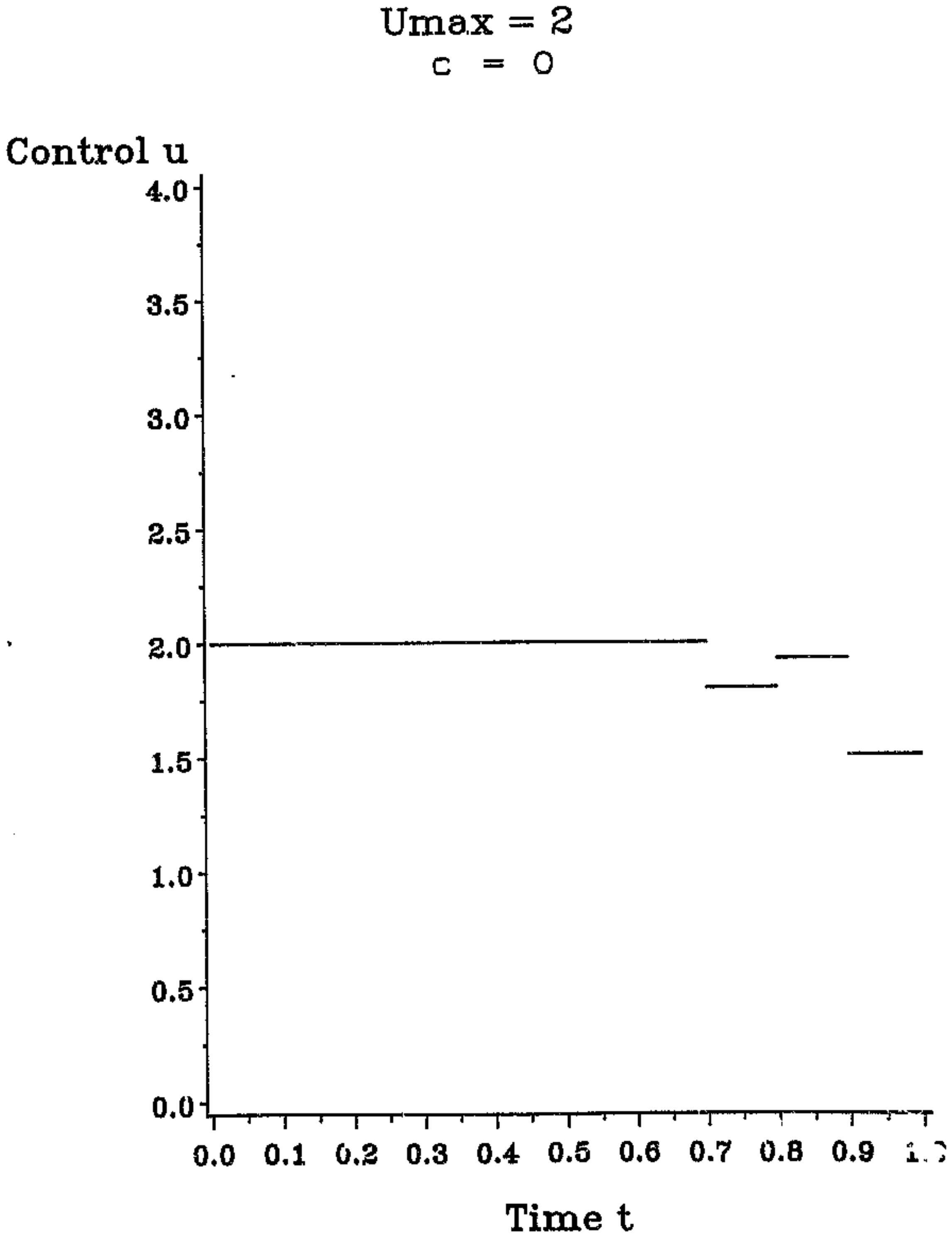

Figure 5.7.2a Computed Optimal Control for Example 5.7.2.

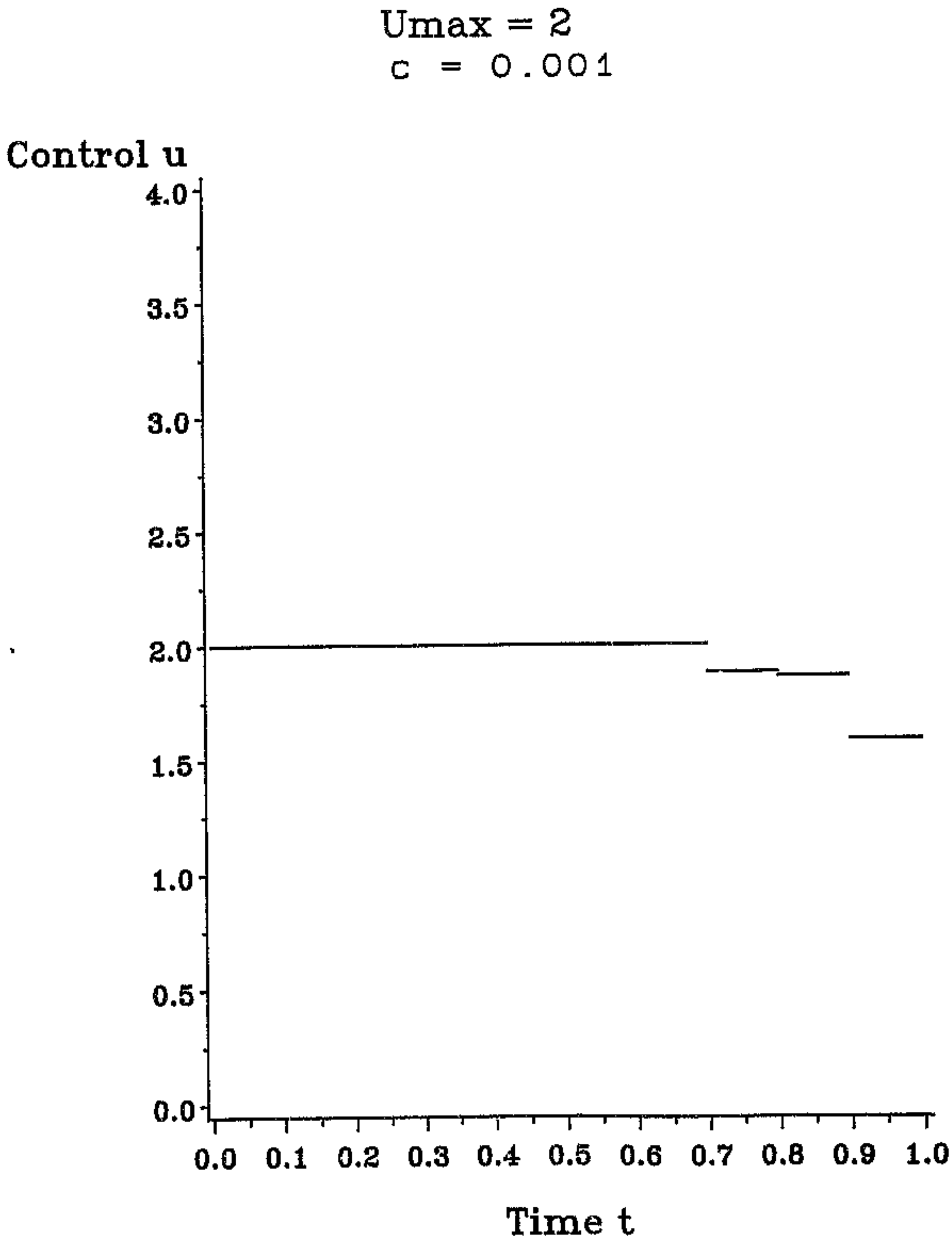

Figure 5.7.2b Computed Optimal Control for Example 5.7.2.

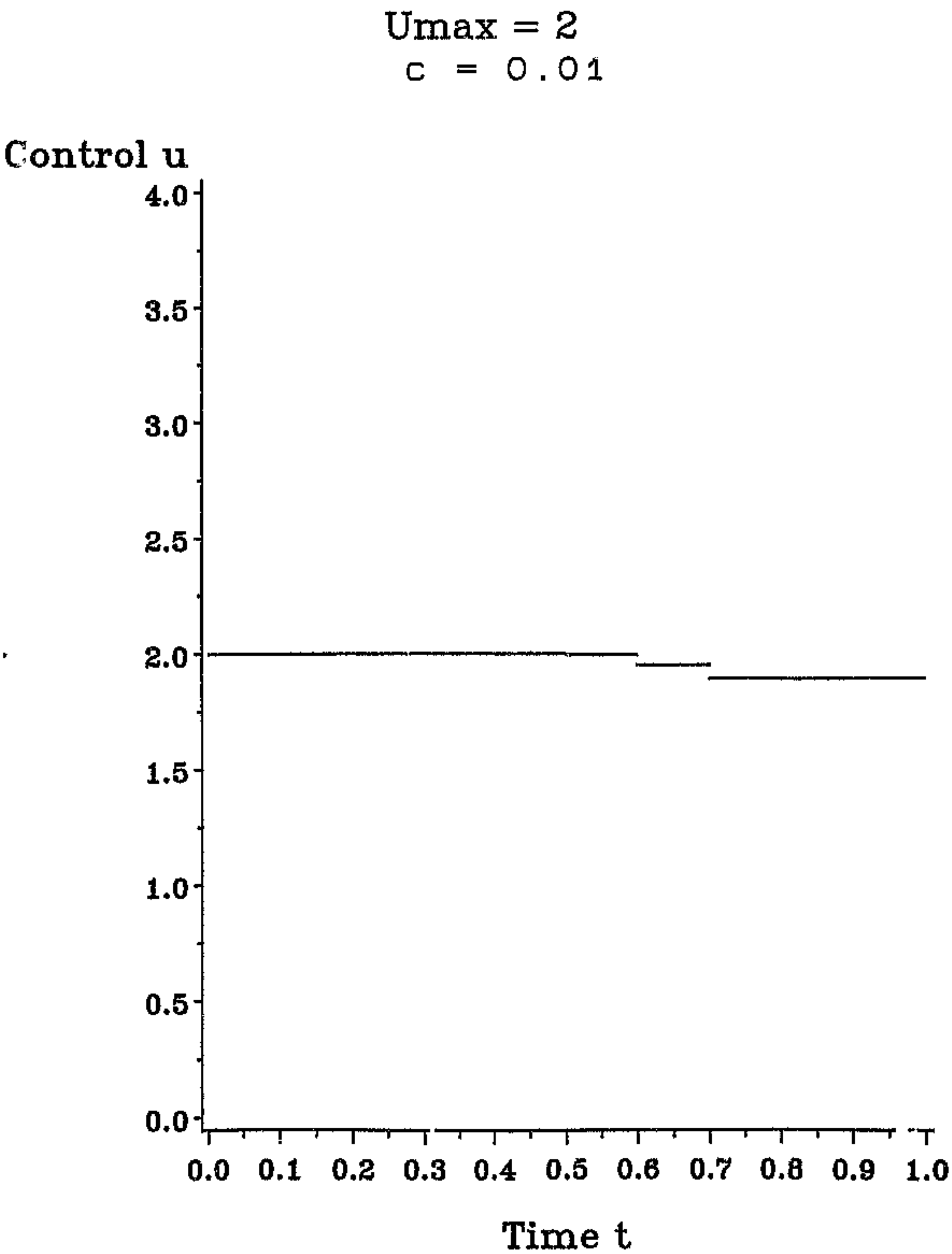

Figure 5.7.2c Computed Optimal Control for Example 5.7.2.

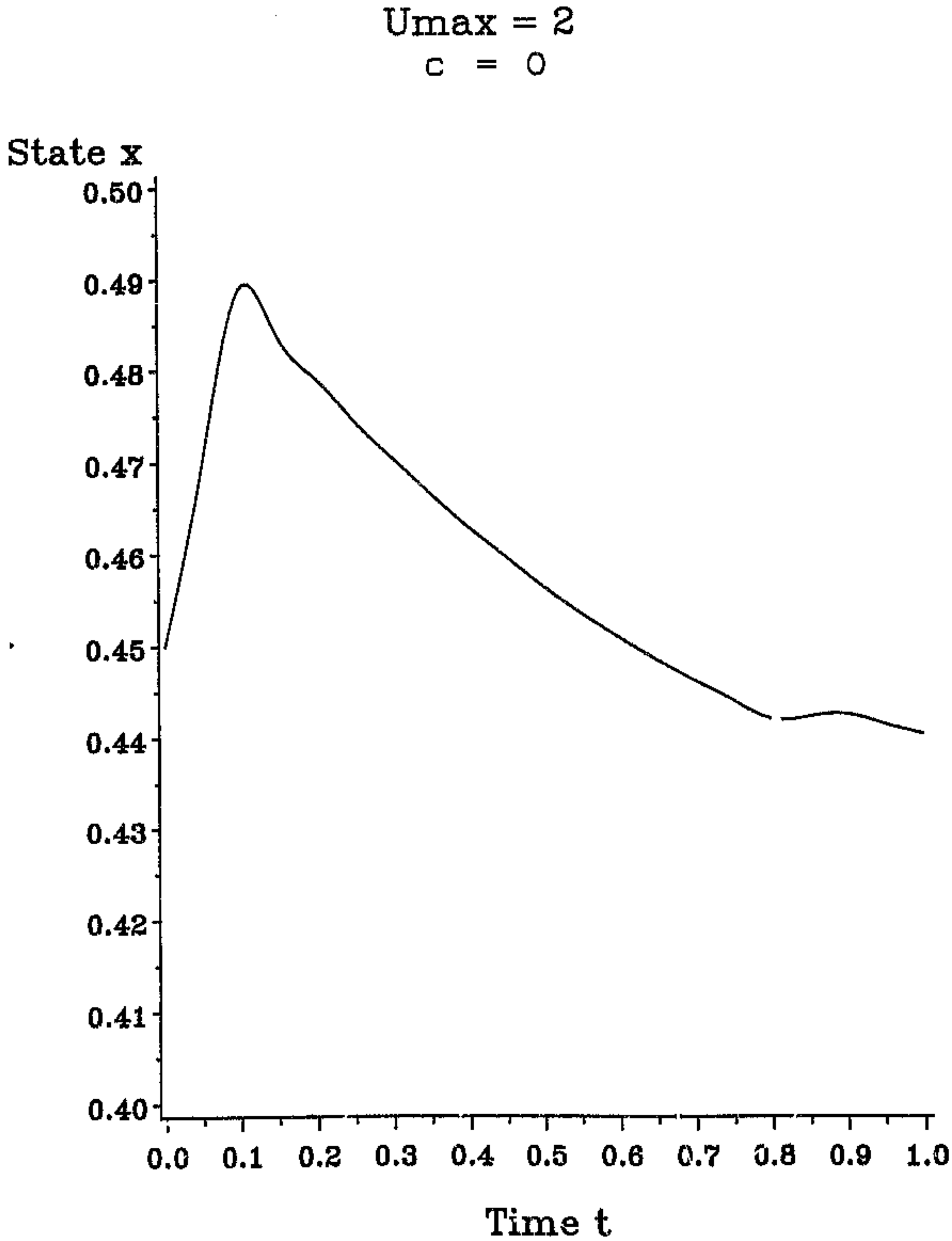

Figure 5.7.2d Computed Optimal State for Example 5.7.2.

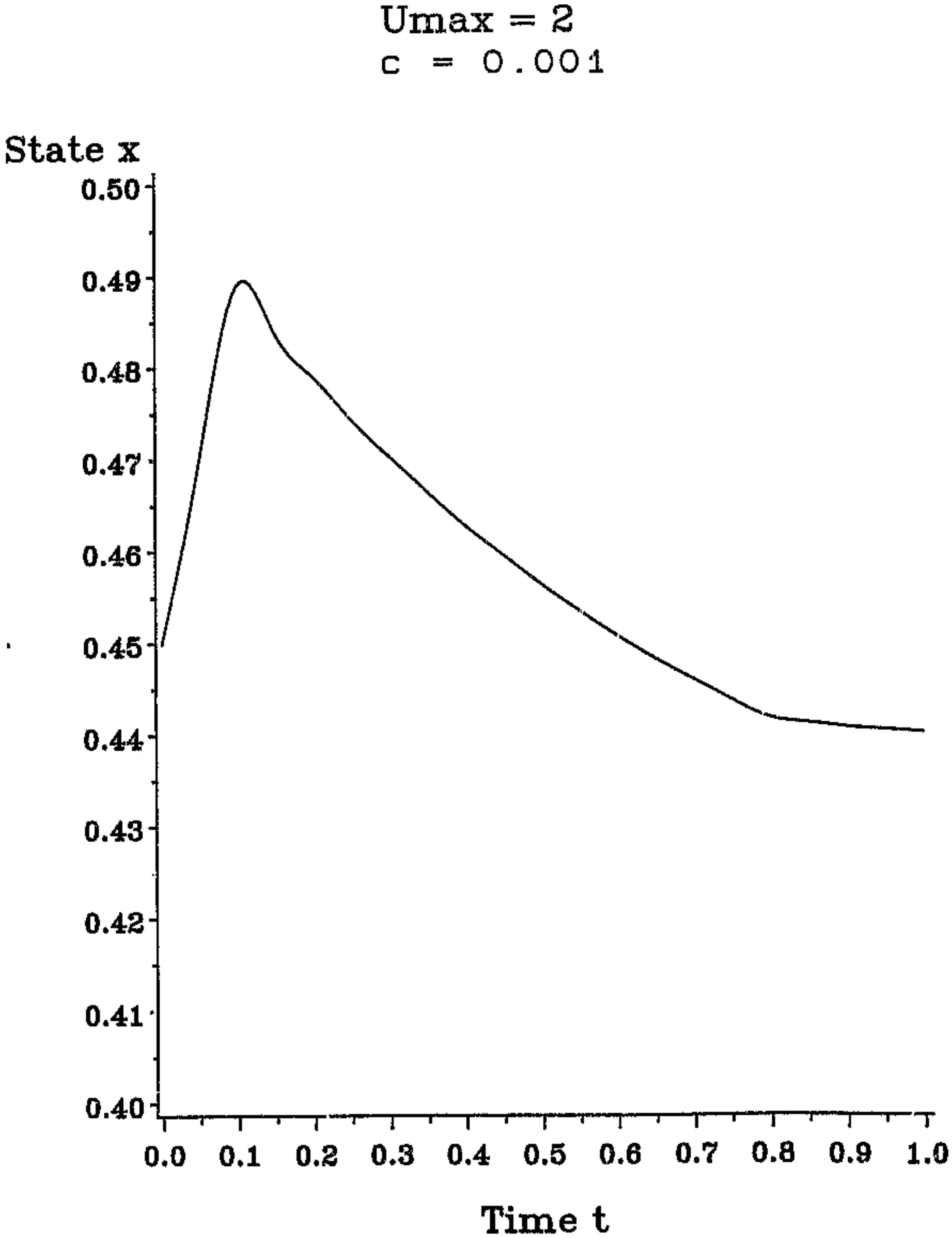

Figure 5 7.2e Computed Optimal State for Example 5.7.2.

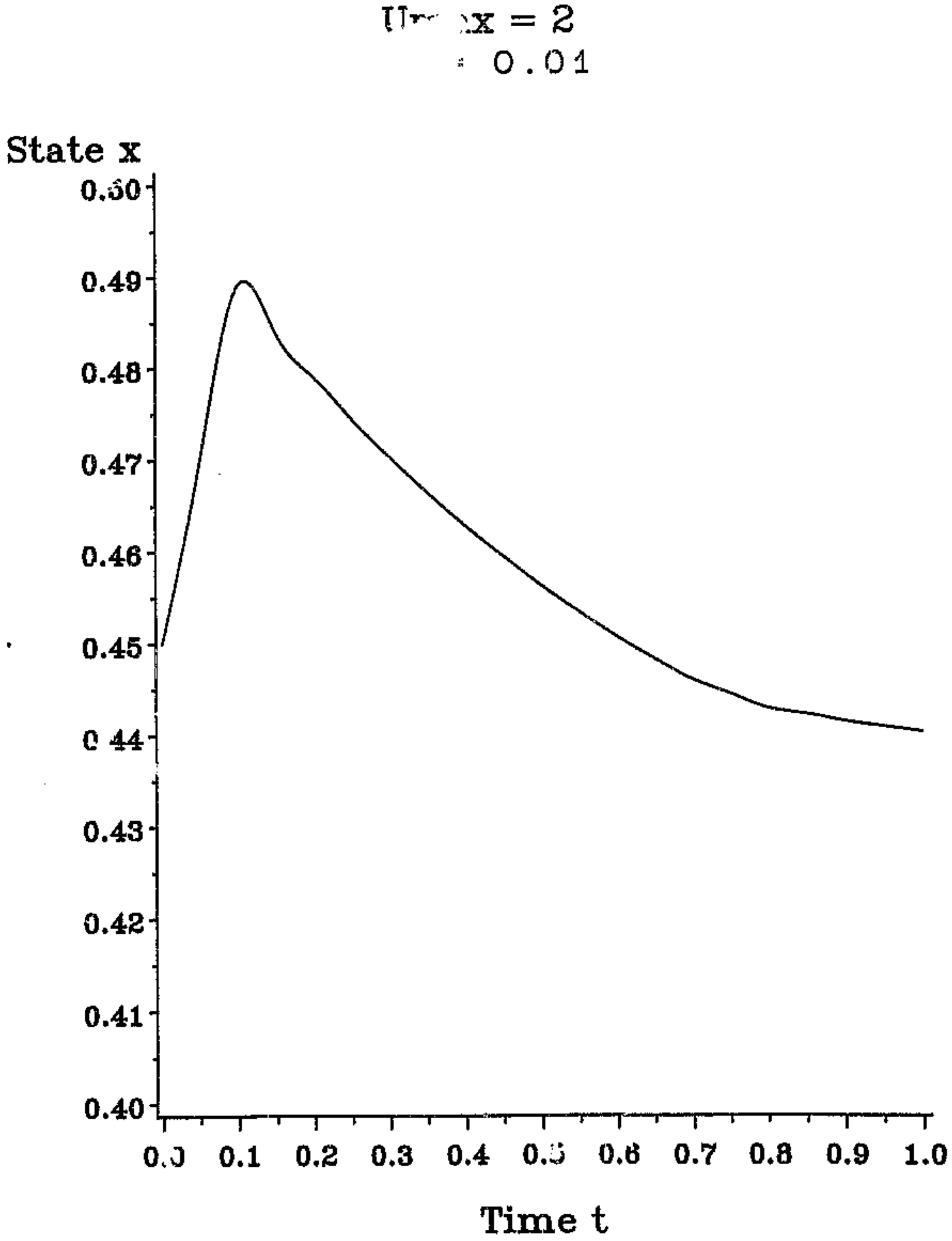

Figure 5.7.2f Computed Optimal State for Example 5.7.2.

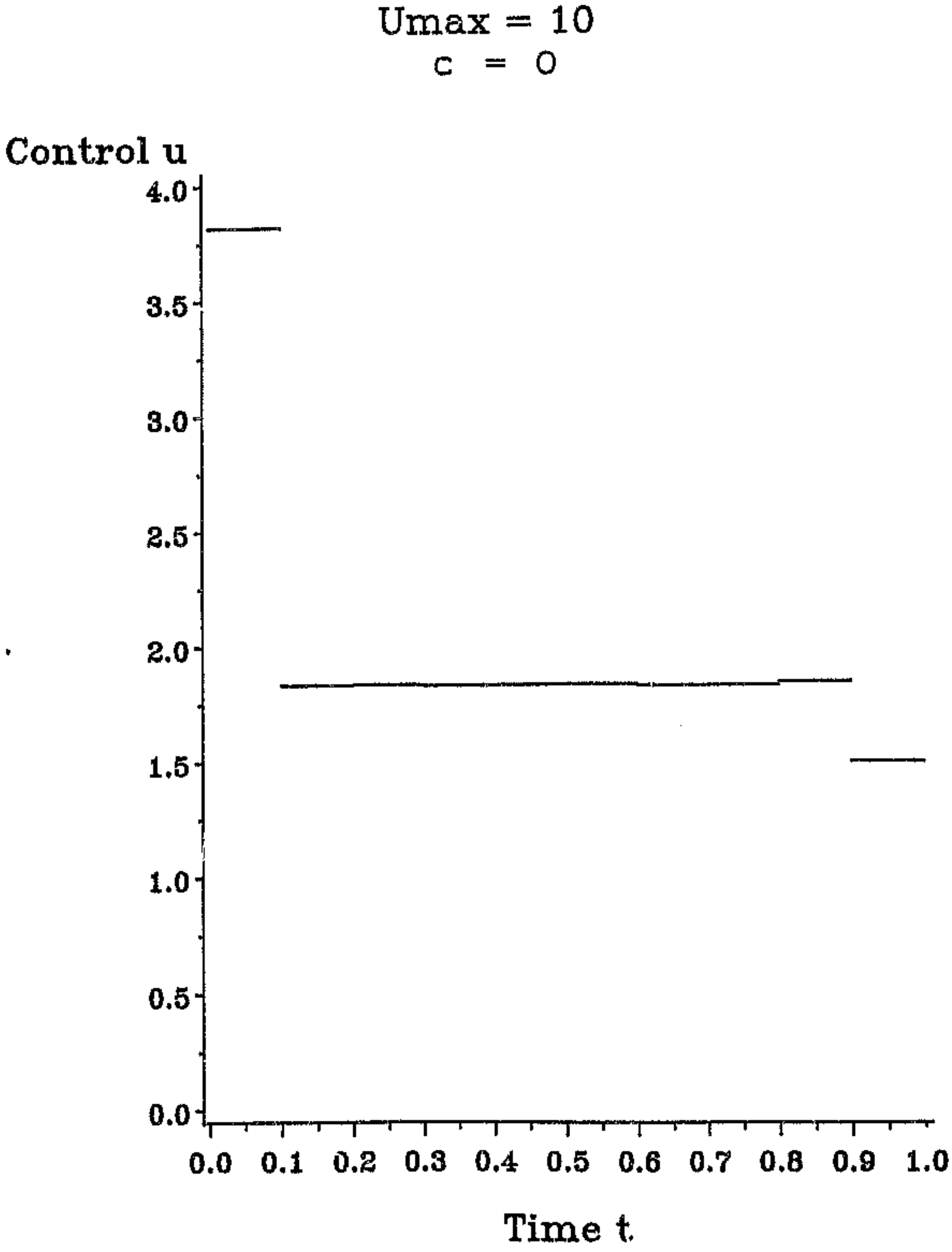

Figure 5.7.2g Computed Optimal Control for Example 5.7.2.

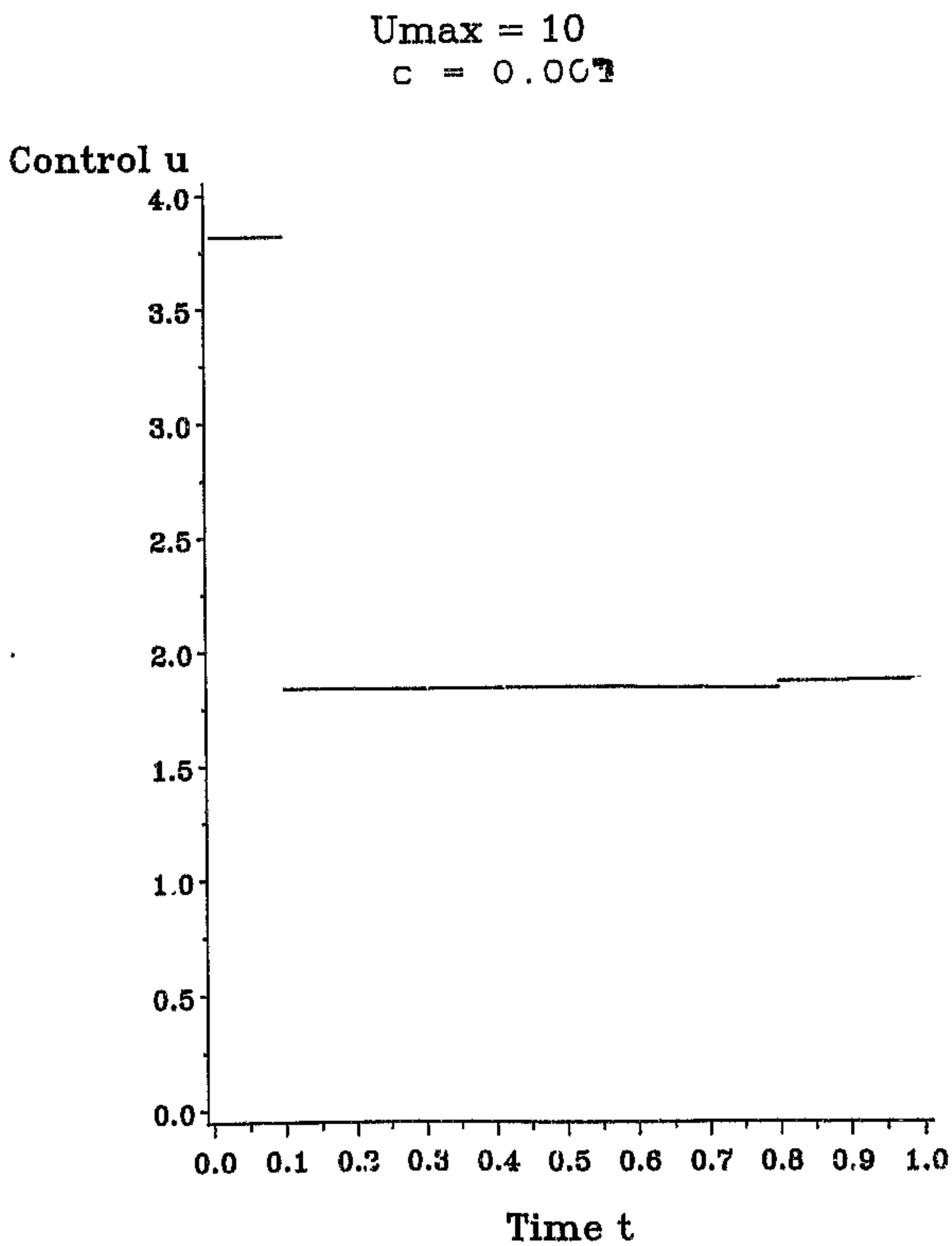

Figure 5.7.2h Computed Optimal Control for Example 5.7.2.

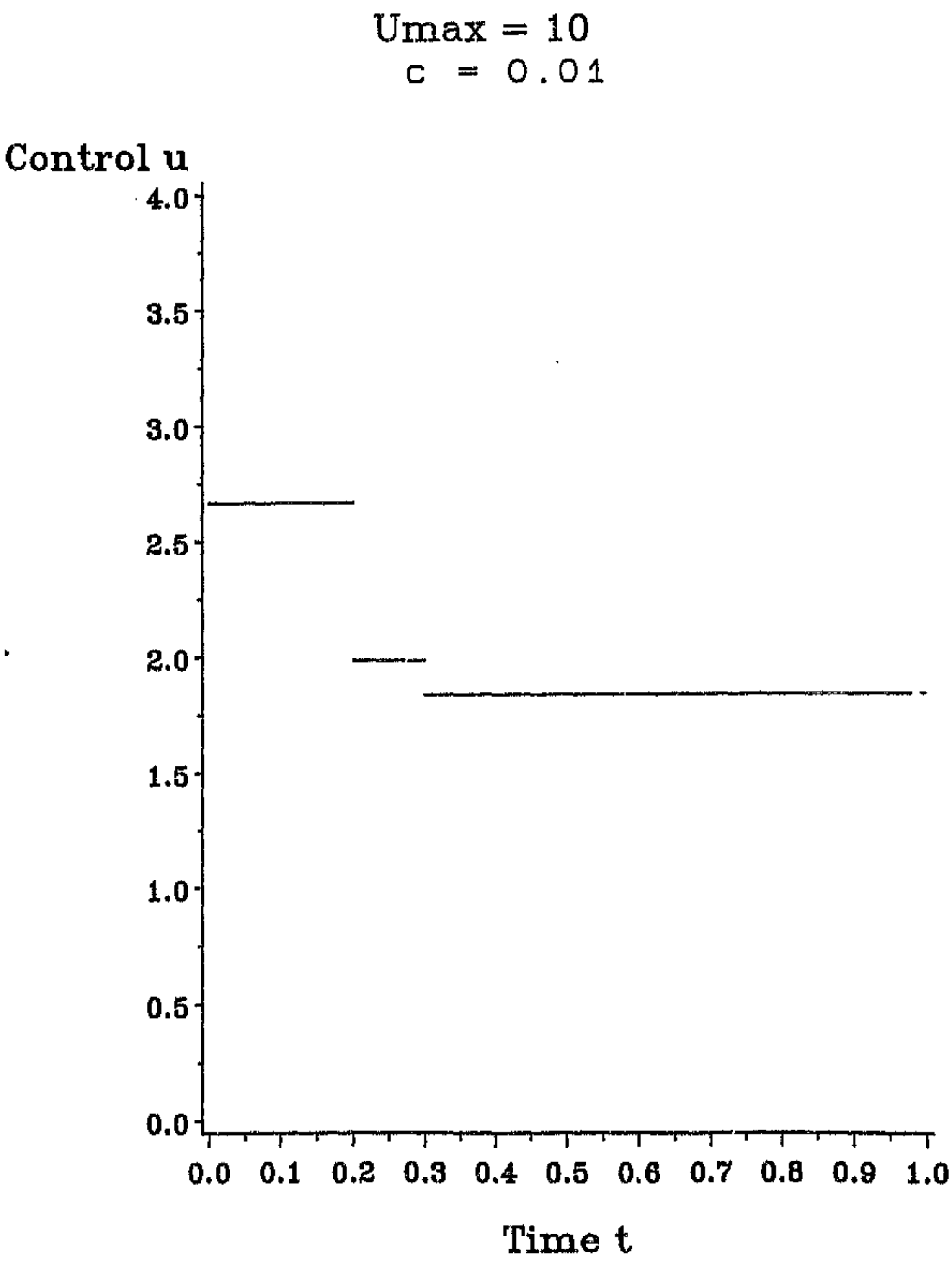

Figure 5.7.2i Computed Optimal Control for Example 5.7.2.

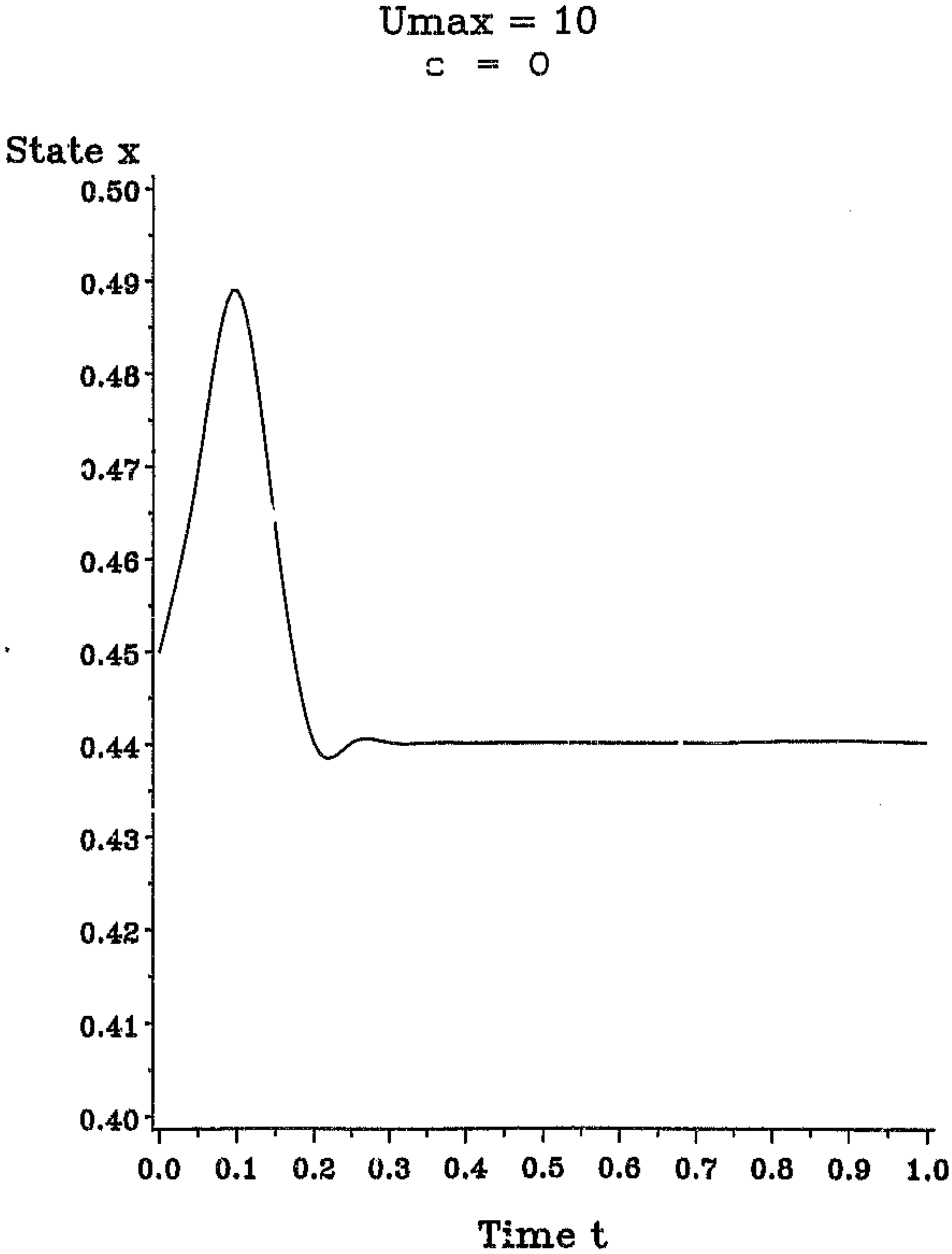

Figure 5.7.2j Computed Optimal State for Example 5.7.2.

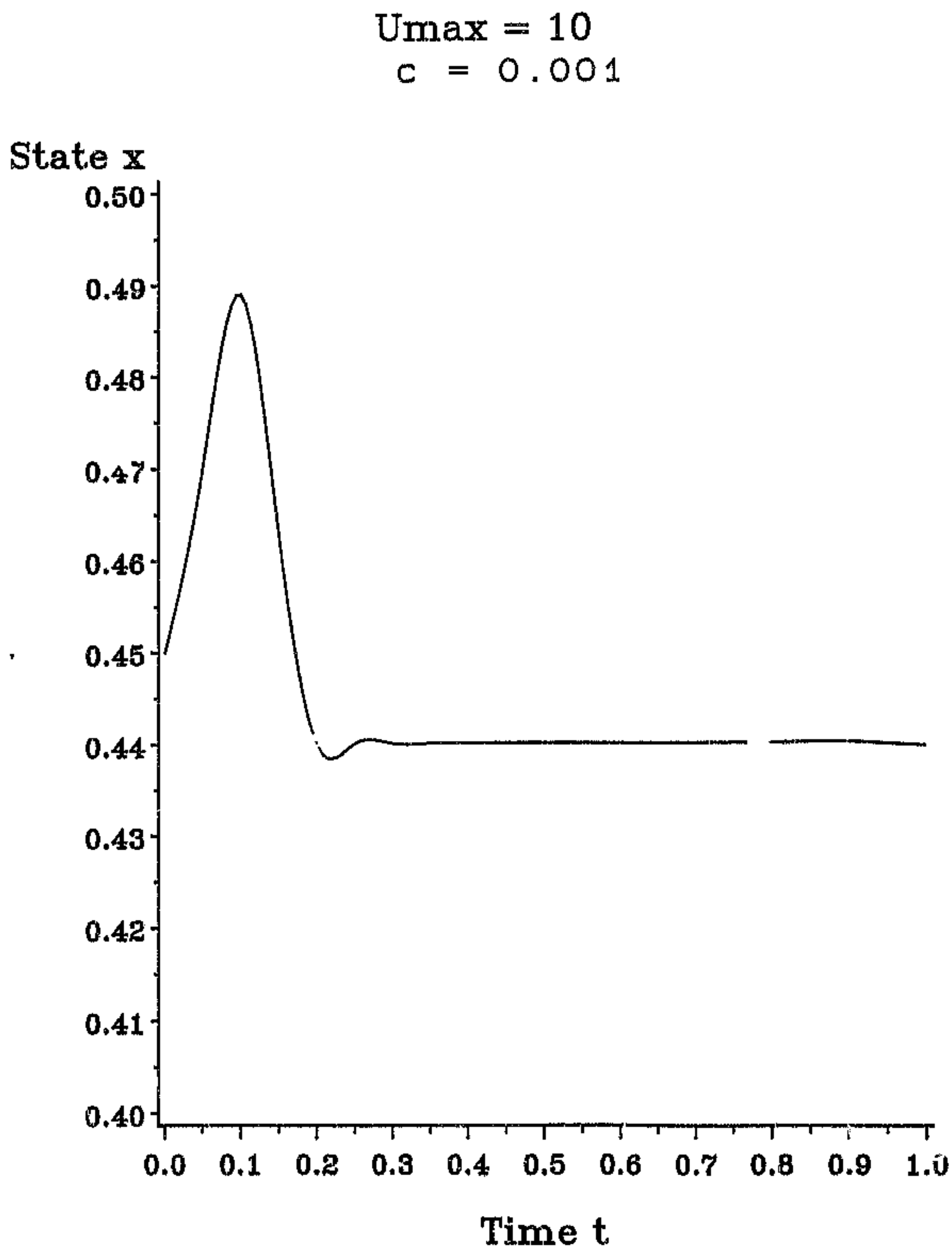

Figure 5.7.2k Computed Optimal State for Example 5.7.2.

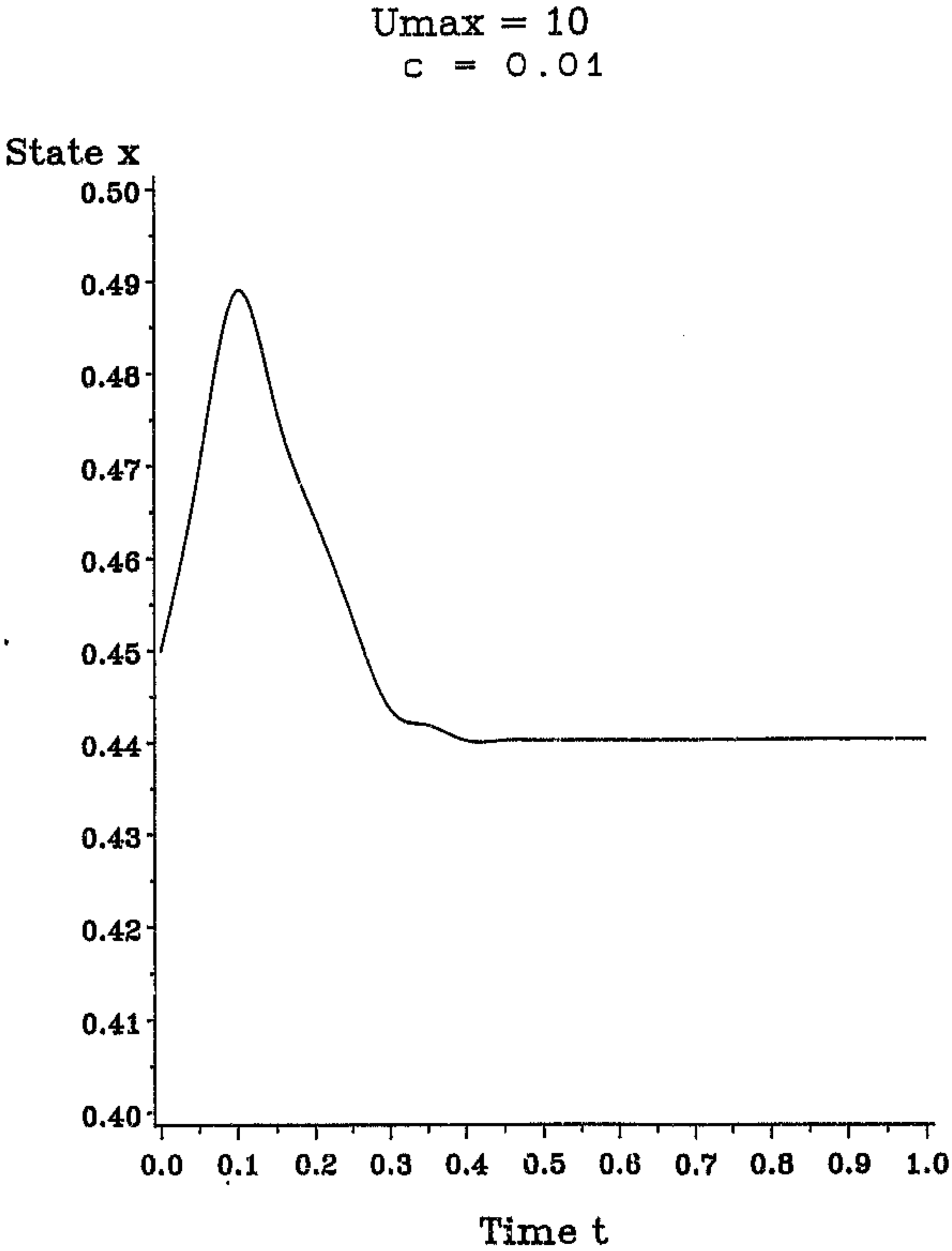

Figure 5.7.2ℓ Computed Optimal State for Example 5.7.2.

CHAPTER VI

OUTPUT FEEDBACK FOR A CLASS OF LINEAR SYSTEMS WITH STOCHASTIC PARAMETERS

6.1 Introduction

The class of stochastic optimal control problems considered in this Chapter, which overlaps with that of Ref. 48 involves a linear system where the parameters are Markov jump processes with finitely countable states, and where an optimal control law is sought with respect to the mathematical expectation of a quadratic cost. Based on the idea of seeking for the best feedback control law, depending only on the measurable output, we convert the original problem into an approximate constrained deterministic optimization problem.

In Section 6.2, we describe the optimal control problem and recall certain results relating to this problem.

In Section 6.3, by restricting the control to have the form of linear output feedback mode, we obtain an approximate constrained deterministic optimization problem (3.P), which can be easily solved by any existing nonlinear programming technique.

In Section 6.4, we present the method for calculating the gradient of the cost functional and the gradients of the constraints of the approximated problem (3.P).

In Section 6.5, an example of Ref. 48 is resolved using the method discussed in this Chapter. The results obtained are as good as those obtained in Ref. 48.

6.2 Statement of the Problem

The system under study is described by its state – space representation. The plant state $x(t) \in R^n$ satisfied a linear vector differential equation

$$\begin{cases} \dot{x}(t) = A(r(t))\, x(t) + B(r(t))\, u(t) & (6.1a) \\ y(t) = C(r(t))\, x(t) & (6.1b) \end{cases}$$

where $u(t) \in R^m$ is the control vector and $y(t) \in R^r$ is the output vector. The entries of the matrices $A(r(t))$, $B(r(t))$ and $C(r(t))$ are random, since they depend on the plant mode $r(t)$, which is a stochiastic jump process with a finite valuation set $S = \{1,...,N\}$. The dynamics of the plant mode $r(t)$ are then described in terms of known transition probabilities

$$\text{Prob}\{r(t+\Delta) = j \mid r(t) = i\}$$

$$= \begin{cases} \pi_{ij}\Delta + 0(\Delta) & \text{if } i \neq j \\ 1 + \pi_{ii}\Delta + 0(\Delta) & \text{if } i = j \end{cases} \qquad (6.2)$$

The matrices $[A(t), B(t), C(t)]$ will be denoted by $[A_i, B_i, C_i]$ when the system operates in the i^{th} mode ($r(t) = i$). A distribution P_0 of the initial mode $r(t_0)$ completes the process description.

For certain purposes, a performance criterion is defined as follows:

$$J(u, t_0, x(t_0), r(t_0))$$

$$= E\left\{\frac{1}{2}\int_{t_0=0}^{\infty} [x(\tau)^T Q\, x(\tau) + u(\tau)^T R\, u(\tau)]\, d\tau \;\middle|\; t_0, x(t_0), r(t_0)\right\} \tag{6.3}$$

where E stands for the mathematical expectation and Q and R are, respectively, a semi–positive and a positive symmetric weighting matrix.

From Refs. 91, 128, 129, it is clear that the optimal control u^* can be expressed in a closed–loop form

$$u^*(t) = U(x(t), r(t)) \tag{6.4}$$

with gains involving a set of N coupled Riccati equations. However, in practice all the required state and mode variables are often not available to the designer. As mentioned in Ref. 48, the more realistic control structures, if it exists, is the best linear output feedback mode dependent control law, i.e. the law

$$u^*(t) = -F(r(t))\, y(t) \tag{6.5}$$

which would minimize (6.3).

6.3 An Approximate Problem

Firstly, we restrict the control to have the form of linear output feedback mode

$$u(t) = -F_i\, y(t) \qquad \text{when } r(t) = i, \tag{6.6}$$

where F_i (i = 1,...,N) $\in R^{m \times r}$ is the gain matrices when the system is operating in mode i.

Theorem 6.3.1 Assume that the system starts in mode i_0 and the initial state is x_0. Then the performance criteria (6.3) can be reduced to

$$\overline{J} = \frac{1}{2} x_0^T K_{i_0} x_0 \tag{6.7}$$

where $K_i \in R^{n \times n}$ (when $r(t) = i$) satisfy the following equation

$$K_i \tilde{A}_i + \tilde{A}_i^T K_i + Q + C_i^T F_i^T R F_i C_i + \sum_{j=1}^{N} \pi_{ij} K_j = 0, \qquad i = 1,\ldots,N \tag{6.8}$$

where

$$\tilde{A}_i = A_i - B_i F_i C_i, \qquad i = 1,\ldots,N, \tag{6.9}$$

provided that the matrices

$$\overline{A}_i = (\tilde{A}_i + \frac{1}{2} \pi_{ii} I), \qquad i = 1,\ldots,N \tag{6.10}$$

are stable.

Although the result of Theorem 6.3.1 was taken from Ref. 48, no proof has been given in Ref. 48. Thus, for the sake of convenience of the reader, we shall provide the proof of Theorem 6.3.1.

Proof of Theorem 6.3.1 From (6.3), (6.5), (6.1b), we get

$$J \triangleq E\left\{\frac{1}{2}\int_{t_0=0}^{\infty} x(\tau)^T \left[Q + C(r(\tau))^T F(r(\tau))^T R\, F(r(\tau))\, C(r(\tau))\right] x(\tau)\, d\tau \;\middle|\; t_0, x(t_0), r(t_0)\right\}. \tag{6.11}$$

In view of (6.8), (6.11) can be simplified to

$$J = E\left\{\frac{1}{2}\int_{t_0=0}^{\infty} - x(\tau)^T \left[\tilde{A}(r(\tau))^T K(r(\tau)) + K(r(\tau))\,\tilde{A}(r(\tau)) + \sum_{j=1}^{N} \pi_{r(\tau),j}\, K(r(\tau))\right] x(\tau)\, d\tau \;\middle|\; t_0, x(t_0), r(t_0)\right\}. \tag{6.12}$$

Now let

$$V(t) = x(t)^T K(r(t))\, x(t) \tag{6.13}$$

Then, from (6.1a), (6.5), (6.1b) and (6.9), we have

$$\dot{x}(t) = \tilde{A}\,(\;(t))\, x(t). \tag{6.14}$$

Thus, by differentiating (6.13), we get

$$\dot{V}(t) = x(t)^T \left[\tilde{A}(r(\tau))^T K(r(\tau)) + K(r(\tau))\,\tilde{A}(r(\tau)) + \dot{K}(r(\tau))\right] x(\tau). \tag{6.15}$$

On the other hand,

$$E\left\{\dot{K}(r(\tau))\Big|\, r(t)=\mathrm{i}\right\} = \lim_{\Delta\to 0+} \frac{E\left\{K(r(t+\Delta))\Big|\, r(t) = \mathrm{i}\right\} - E\left\{K(r(t))\Big|\, r(t) = \mathrm{i}\right\}}{\Delta} \tag{6.16}$$

But,

$$\lim_{\Delta\to 0+} E\left\{K(r(t+\Delta))\Big|\, r(t)=\mathrm{i}\right\}$$
$$= K_\mathrm{i} + \Delta \sum_{\mathrm{j}=1}^{N} \pi_{\mathrm{ij}} K_\mathrm{j} + 0(\Delta), \tag{6.17}$$

so that (6.16) becomes

$$E\left\{\dot{K}(r(t))\Big|\, r(t)=\mathrm{i}\right\} = \sum_{\mathrm{j}=1}^{N} \pi_{\mathrm{ij}} K_\mathrm{j}. \tag{6.18}$$

Thus, from (6.15), (6.18), we get

$$\dot{V}(t) = x(t)^{\mathrm{T}} \left[\tilde{A}(r(\tau))^{\mathrm{T}} K(r(\tau)) + K(r(\tau))\, \tilde{A}(r(\tau)) + \sum_{\mathrm{j}}^{N} \pi_{\mathrm{ij}} K_\mathrm{j}\right] x(t). \tag{6.19}$$

From (6.12) and (6.19), we get

$$\mathcal{J} = E\left\{\frac{1}{2}\int_{t_0=0}^{\infty} -\dot{V}(t)\,dt \;\Big|\; t_0, x(t_0), r(t_0)\right\}$$

$$= \frac{1}{2}\left[x_0^{\mathrm{T}} K_{\mathrm{i}_0} x_0 - E\left\{V(\infty) \;\Big|\; t_0, x(t_0), r(t_0)\right\}\right] \tag{6.20}$$

Since $\tilde{A}_\mathrm{i} + \frac{1}{2}\pi_{\mathrm{ii}} I$ is stable for all i, we have

$$E\left\{V(\infty)\;\middle|\; t_0, x(t_0), r(t_0)\right\} = 0. \tag{6.21}$$

The conclusion of this theorem follows easily from (6.20) and (6.21).

Furthermore, by assuming that the initial state x_0 is a random vector uniformly distributed on the surface of the n-dimension unit sphere, the performance (6.7) can be reduced to

$$\bar{J} = \frac{1}{2n}\,\text{trace}\,(K_{i_0}). \tag{6.22}$$

The performance degradation of $\bar{J}$ with respect to the optimal performance under full state feedback J^* can be measured by the degree of suboptimality μ

$$\mu = \frac{\bar{J} - J^*}{J^*} \tag{6.23}$$

Equation (6.22) emphasizes the dependence of the performance on the initial plant mode $r(t_0)$. As it was done for the initial plant state, it is possible to average out this dependence by assuming a uniform distribution for $r(t_0)$. More precisely, if,

$$P_{0i} = \text{Prob}\,(r(t_0) = i) = \frac{1}{N} \quad , \tag{6.24}$$

an average suboptimal criterion is defined by

$$\bar{J} = \frac{1}{2Nn}\sum_{i=1}^{N} \text{trace}\,(K_i) \tag{6.25}$$

Furthermore, a sufficient condition for $\bar{A}_i$, i = 1,...,N, to be stable is

$$\Lambda_i = \frac{1}{2}\left[-\bar{A}_i - \bar{A}_i^T\right] \tag{6.26}$$

to be positive definite. Thus, by Sylvester's criterion, it is necessary and sufficient to ensure that for each i = 1,...,N, the determinants of all the principle minors of Λ_i are greater than or equal to $\epsilon > 0$, where ϵ is some small positive number. Let the determinants of the n principle minors of Λ_i be denoted by $g_{i,1},\ldots,g_{i,n}$ respectively. Then we have

$$g_{i,j} \geq \epsilon \qquad j = 1,\ldots,n \tag{6.27}$$

Thus, we obtain a deterministic optimization problem (6.P) as follows:

$$(6.P) \qquad \begin{cases} \min \bar{J} = \dfrac{1}{2Nn} \displaystyle\sum_{i=1}^{N} \text{trace}\,(K_i) & (6.28) \\ \text{subject to} & \\ \quad g_{i,j} \geq \epsilon, \quad i = 1,\ldots,N;\ j = 1,\ldots,n & (6.29) \end{cases}$$

where K_i are the solutions of (6.8). The problem (6.P) can be easily solved by any existing nonlinear programming software, such as NLQPL.

In order to solve the problem (6.P), we need to calculate the gradient of $\bar{J}$ and $g_{i,j}$ with respect to the gain matrix F_i. The method for calculating these gradients are given in Section 6.4.

6.4 Gradient of the cost Functional and the constraints

To derive a formula for the gradient of the cost functional $\bar{\mathcal{J}}$ and the constraints g_{ij}, we need to regard equation (6.8) as the state equation in optimal control problem. We first introduce the Hamiltonian $\mathcal{H}$ and the co–state matrices λ_k (k = 1,...,N) for $\bar{\mathcal{J}}$. Let

$$\mathcal{H} = \frac{1}{2Nn}\sum_{i=1}^{N}\text{trace}\,\{K_i\} + \sum_{i=1}^{N}\text{trace}\left\{\lambda_i\left[K_i\tilde{A}_i + \tilde{A}_i^T K_i + Q + C_i^T F_i^T R\, F_i C_i + \sum_{j=1}^{N}\pi_{ij}K_j\right]\right\} \tag{6.30}$$

where the co–state matrices λ_k (k = 1,...,N) $\in R^{n\times n}$ are symmetric matrices satisfying the following adjoint system:

$$\frac{\partial\mathcal{H}}{\partial K_k} = 0$$

$$\Rightarrow \frac{1}{2Nn} I + \lambda_k\tilde{A}_k^T + \tilde{A}_k\lambda_k + \sum_{i=1}^{N}\pi_{ik}^T\,\lambda_i = 0 \tag{6.31}$$

Theorem 6.4.1 The gradient of $\bar{\mathcal{J}}$ is given by

$$\frac{\partial\bar{\mathcal{J}}}{\partial F_i} = \frac{\partial\mathcal{H}}{\partial F_i} = 2\,R\,F_iC_i\lambda_iC_i^T \qquad (i = 1,...,N) \tag{6.32}$$

Proof Let F_i be any parameter matrix when $r\,(t) = i$ and let ΔF_i be any perturbation about F_i. Define

$$F_i(\epsilon) = F_i + \epsilon\Delta F_i \tag{6.33}$$

For brevity, let $K_k(\cdot)$ and $K_k(\cdot,\epsilon)$ (k = 1,...,N) denote, respectively, the solutions of the system (6.8) corresponding to parameter matrices $(F_1,...,F_i,...,F_N)$ and the perturbed parameter matrices $(F_1,...,F_{i-1},F_i + \epsilon\Delta F_i,F_{i+1},...,F_N)$ respectively. Let

$$\Delta K_k = \frac{dK_k(\cdot,\epsilon)}{d\epsilon}\Big|_{\epsilon=0} , \qquad (k = 1,...,N) \tag{6.34}$$

Then, from (6.28), (6.30) and (6.8), we get

$$\Delta \mathcal{J} = \frac{d\ \mathcal{J}(F_i(\epsilon))}{d\epsilon}\Big|_{\epsilon=0}$$

$$= \Delta\ \mathcal{H}$$

$$= \left\langle\frac{\partial \mathcal{H}}{\partial F_i}, \Delta F_i\right\rangle + \sum_{k=1}^{N}\left\langle\frac{\partial \mathcal{H}}{\partial K_k}, \Delta K_k\right\rangle , \tag{6.35}$$

where the inner produce of two matrices

$$A = (a_{kl})_{\substack{k=1,\cdots,r_1 \\ l=1,\cdots,r_2}} \qquad \text{and} \qquad B = (b_{kl})_{\substack{k=1,\cdots,r_1 \\ l=1,\cdots,r_2}}$$

is defined by

$$\langle A,B\rangle = \sum_{l=1}^{r_2}\sum_{k=1}^{r_1} a_{kl}\, b_{kl} \tag{6.36}$$

Hence from (6.31) and (6.35), we get

$$\Delta \mathcal{J} = \left\langle\frac{\partial \mathcal{H}}{\partial F_i}, \Delta F_i\right\rangle \tag{6.37}$$

Since ΔF_i is arbitrary, the conclusion of the theorem follows readily from (6.37).

Remark 6.4.1 To calculate $\frac{\partial \overline{J}}{\partial F_i}$, we need to solve the system comprising of (6.8) and (6.30). This system is known as system of Liapunov equations, which can be solved by existing software.

Remark 6.4.2 The method for calculating the gradients of the constraints is exactly the same as that for calculating the gradients of $\overline{J}$.

6.5 Illustrative Example

Example 6.5.1 (Same as the example given in p.899 of Ref. 48)

Let $x \in R^2$, $u \in R$, $y \in R$. There are two modes, with the transition matrix $\Pi = \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix}$.

In mode 1, the plant is governed by

$$A_1 = \begin{bmatrix} -1 & 0.5 \\ 0 & 1 \end{bmatrix}, \qquad B_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad C_1 = (0,1).$$

In mode 2, the plant is governed by

$$A_2 = \begin{bmatrix} 0.25 & 0.5 \\ 0 & 0.25 \end{bmatrix}, \qquad B_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \qquad C_2 = (0,1).$$

The criterion (6.3) is considered with unity weighting matrices Q and R.

From Ref. 48, we know that the optimal solution with full state feedback is given by:

$$K_1 = \begin{bmatrix} 0.4949 & 0.1010 \\ 0.1010 & 2.4495 \end{bmatrix} \text{ and a gain matrix } \quad F_1 = (0.1010 \quad 2.4495);$$

$$K_2 = \begin{bmatrix} 2.3548 & 0.5635 \\ 0.5635 & 1.7688 \end{bmatrix} \text{ and a gain matrix } \quad F_2 = (0.5635 \quad 1.7688);$$

The technique employed by Ref. 48 which is based on the necessary condition of optimality, yields the following results:

$$K_1 = \begin{bmatrix} 0.5 & 0.0995 \\ 0.0995 & 2.4503 \end{bmatrix} \text{ and a gain matrix } \quad F_1 = 2.5135;$$

$$K_2 = \begin{bmatrix} 3.0 & 0.6577 \\ 0.6577 & 1.7967 \end{bmatrix} \text{ and a gain matrix } \quad F_2 = 1.9319.$$

The results obtained by solving the approximate problem $(6.P)$ is as follows:

$K_1 = \begin{bmatrix} 0.5 & 0.1010 \\ 0.1010 & 2.4497 \end{bmatrix}$ and a gain matrix $F_1 = 2.47$[illegible]9;

$$K_2 = \begin{bmatrix} 3.0 & 0.6583 \\ 0.6583 & 1.7967 \end{bmatrix} \text{ and a gain matrix } \quad F_2 = 1.9319.$$

Let J^*, J^0 and J^1 be the value of the objective function which corresponds to the full state feedback solution, the optimal output feedback obtained by Mariton and Bertrand (Ref. 48) and the optimal output feedback obtained by our method respectively. Let μ^0 and μ^1 be the value of μ obtained by substituting $J = J^0$ and $J = J^1$ in (6.23) respectively. The performance can be summed up in Table 6.5.1, depending on the initial mode distribution

$$P_0 = \begin{bmatrix} P_{01} \\ 1 - P_{01} \end{bmatrix}$$

From the values of μ^0 and μ^1 in Table 6.5.1, it is clear that the expected performances obtained by our method is slightly better than those obtained by Mariton and Bertrand (Ref. 48). In both cases, the worse performance ($P_{01} = 0$, $r(t_0) = 2$) using output feedback (6.5) is only about 10% worse than the optimal expected performance using full state feedback.

P_{G1}	1	0.5	0
$8 \times J^*$	2.9444	3.6309	4.3168
$8 \times J^0$	2.9503	3.8735	4.7967
$8 \times J^1$	2.9497	3.8732	4.7967
μ^0(%)	0.20	6.68	11.12
μ^1(%)	0.18	6.67	11.12

Table 6.5.1 Comparison of Optimal Solution

CHAPTER VII

OPTIMAL CONTROL OF A CHEMICAL REACTOR

7.1 Introduction

In this chapter, we consider a chemical reactor problem involving a couple of nonlinear diffusion equations and boundary conditions. The variables to be controlled are the temperature and concentration of oxygen, which are functions of both position and time. The objective is to choose the best input temperature and input oxygen concentration in order to achieve a desired output temperature as much as possible.

In Section 7.2, we describe the differential equations governing the chemical reaction, together with the boundary conditions of the chemical reaction. Then, the optimal control problem is defined.

In Section 7.3, a finite element Galerkin method is used to convert the distributed optimal control problem into a sequence of finite dimensional quadratic programming problem (7.P), which can be solved by any standard quadratic programming technique.

In Section 7.4, specific computational aspect for implementing the method for solving (7.P) is discussed. Most significantly, we describe a method for finding the exponential of an extremely ill-conditioned matrix, whose eigenvalues and inverse cannot be easily evaluated.

In Section 7.5, we first solve numerically the chemical reactor problem when there is no constraint on the control, that is, no constraint on the input temperature and the input oxygen concentration. Secondly impose certain upper and lower bounds on the input oxygen concentration and resolve the problem. Thirdly, we impost certain upper and lower bounds on the input temperature and resolve the problems. In all cases, the output temperature obtained is extremely close to the desired output temperature. In general, an increase in the number of partitions of the linear spline function used over the time interval [0,1] will lead to the convergence of the optimal control of the approximate problem to the true optimal control.

7.2 Nonlinear Diffusion Equations for the Chemical Reactor

The differential equations governing the chemical reaction are given by

$$\frac{\partial y}{\partial t} = -f\frac{\partial y}{\partial z} + \frac{1}{Pe_h}\frac{\partial^2 y}{\partial z^2} + \beta\, D_m\, R\,(y,x) \tag{7.1}$$

$$0 = -f\frac{\partial x}{\partial z} + \frac{1}{Pe_m}\frac{\partial^2 x}{\partial z^2} - D_m\, R\,(y,x), \tag{7.2}$$

where

$$R\,(y,x) = x\exp\,\{\gamma\,(1-[y\,(1+\alpha\, T_r\, y)]^{-1})\}, \tag{7.3}$$

where f is a given constant representing the velocity parameter of the reaction; Pe_h, β, D_m, Pe_m, γ and α are also given constants. Values of these constants taken from a chemistry laboratory are as follows:

$f = 0.01$, $Pe_h = 95$, $\beta = 0.38812$, $D_m = 0.89173$, $Pe_m = 235$,

$\gamma = 12.297$ $\alpha = 2.5696 \times 10^{-4}$.

The boundary conditions for the problem are

$$\frac{\partial y\ (z_f,t)}{\partial z} = 0 \qquad \frac{\partial x\ (z_f,t)}{\partial z} = 0 \tag{7.4}$$

$$y\,(0,t) = y_I(t), \qquad x\,(0,t) = x_I(t) \tag{7.5}$$

$$\frac{\partial y\ (0,t)}{\partial z} = 0 \qquad \frac{\partial x\ (0,t)}{\partial z} = 0 \tag{7.6}$$

Here, $x\ (z,t)$ and $y\ (z,t)$ are normalized variables representing, respectively, the oxygen concentration C and the temperature T at position z and at time t, where

$$x\,(z,t) = C\,(z,t)/C_r \tag{7.7}$$

and

$$y\,(z,t) = T\,(z,t)/T_r \tag{7.8}$$

Values of the constants C_r and T_r taken from the chemistry laboratory are as follows:

$C_r = 0.01$ mole fraction oxygen (1 mole %)

$T_r = 434.27$ Kelvin.

$x_I(t)$ and $y_I(t)$, the inlet oxygen concentration and the inlet temperature respectively, are the controls for our problem. Sometimes, the gradient of the inlet oxygen concentration and the gradient of the inlet temperature can also be regarded as controls, in which case equation (7.6) can be replaced by

$$\frac{\partial y\ (0,t)}{\partial z} = \bar{y}_I(t), \qquad \frac{\partial x\ (0,t)}{\partial z} = \bar{x}_I(t), \tag{7.9}$$

where $\bar{y}_I(t)$ and $\bar{x}_I(t)$ are now additional controls.

The objective is to have the outlet temperature $y\,(z_f,\,t)$ to approach a desired temperature distribution $y_d(t)$ as closely as possible. This can be formulated as

$$\text{minimize } J = \int_0^1 [y\,(z_f,t) - y_d(t)]^2\, dt \tag{7.10}$$

Here, the time interval $t \in [0,1]$ is used. Any other time interval $t' \in [0,T]$ can be converted into the time interval $t \in [0,1]$ by using the transformation

$$t = \frac{t'}{T} \tag{7.11}$$

7.3 Method of Solution

A finite element Galerkin method will be used to convert the distributed optimal control problem into a finite dimensional mathematical programming problem with the values of the initial temperature and input oxygen concentration together with their respective gradients at every grid point being the variables. Let the temperature and concentration be approximated as follows:

$$y\,(z,t) \approx \sum_{i=1}^{N+1} B_i(t)\; Y_i(z) \tag{7.12}$$

and

$$x\,(z,t) \approx \sum_{i=1}^{N+1} B_i(t)\; X_i(z), \tag{7.13}$$

where the integer N and the functions $B_i(t)$ are subject to choice substituting (7.12) and (7.13) into (7.1), we get

$$\sum_{i=1}^{N+1} \dot{B}_i(t)\, Y_i(z) = -f \sum_{i=1}^{N+1} B_i(t)\, Y_i'(z) + \frac{1}{Pe_h} \sum_{i=1}^{N+1} B_i(t)\, Y_i''(z) + \beta\, D_m\, R\,(y,x) \qquad (7.14)$$

Multiplying (7.14) by $B_j(t)$ and integrating over [0,1], we get

$$\begin{aligned} \sum_{i=1}^{N+1} Y_i(z) \int_0^1 B_j(t)\, \dot{B}_i(t)\, dt &= -f \sum_{i=1}^{N+1} Y_i'(z) \int_0^1 B_j(t)\, B_i(t)\, dt \\ &+ \frac{1}{Pe_h} \sum_{i=1}^{N+1} Y_i''(z) \int_0^1 B_j(t)\, B_i(t)\, dt \\ &+ \beta\, D_m \int_0^1 R\,(y,x)\, B_j(t)\, dt \end{aligned} \qquad (7.15)$$

To convert (7.15) into a linear differential equation involving $Y_i(z)$ and $X_i(z)$, the term $R\,(y,x)$ must be linearized around a suitable steady state solution $(\bar{y}\,(z),\, \bar{x}\,(z))$ as follows:

$$R\,(y,x) = R\,(\bar{y},\bar{x}) + (y - \bar{y})\, R_y\,(\bar{y},\bar{x}) + (x - \bar{x})\, R_x\,(\bar{y},\bar{x}) \qquad (7.16)$$

However, from (7.3)

$$R\,(\bar{y},\bar{x}) = \bar{x}\, R_x\,(\bar{y},\bar{x}) \qquad (7.17)$$

Combining (7.16) and (7.17), we get

$$R\,(y,x) = (y - \bar{y})\, R_y\,(\bar{y},\bar{x}) + x\, R_x\,(\bar{y},\bar{x}) \qquad (7.18)$$

a suitable choice for the steady state solution is $\overline{y} = 0.8175$ and $\overline{x} = 0.25$. This solution is independent of z and corresponds to a concentration of oxygen of 0.0025 and a temperature of 82° C.

In view of (7.18), (7.12) and (7.13), equation (7.15) can be written as

$$\sum_{i=1}^{N+1} Y_i(z) \int_0^1 B_j(t)\, \dot{B}_i(t)\, dt = -f \sum_{i=1}^{N+1} Y_i'(z) \int_0^1 B_j(t)\, B_i(t)\, dt$$

$$+ \frac{1}{Pe_h} \sum_{i=1}^{N+1} Y_i''(z) \int_0^1 B_j(t)\, B_i(t)\, dt$$

$$- \beta\, D_m\, \overline{y}\, R_y\,(\overline{y},\overline{x}) \int_0^1 B_j(t)\, dt$$

$$+ \beta\, D_m\, R_y\,(\overline{y},\overline{x}) \sum_{i=1}^{N+1} Y_i(z) \int_0^1 B_j(t)\, B_i(t)\, dt$$

$$+ \beta\, D_m\, R_x\,(\overline{y},\overline{x}) \sum_{i=1}^{N+1} X_i(z) \int_0^1 B_j(t)\, B_i(t)\, dt \qquad (7.19)$$

By using exactly the same argument as that used to obtain (7.19) from (7.1), we can easily deduce from (7.2) that

$$0 = -f \sum_{i=1}^{N+1} X_i'(z) \int_0^1 B_j(t)\, B_i(t)\, dt$$

$$+ \frac{1}{Pe_m} \sum_{i=1}^{N+1} X_i''(z) \int_0^1 B_j(t)\, B_i(t)\, dt$$

$$+ D_m \bar{y} R_y (\bar{y},\bar{x}) \int_0^1 B_j(t)\, dt$$

$$- D_m R_y (\bar{y},\bar{x}) \sum_{i=1}^{N+1} Y_i(z) \int_0^1 B_j(t)\, B_i(t)\, dt$$

$$- D_m R_x (\bar{y},\bar{x}) \sum_{i=1}^{N+1} X_i(z) \int_0^1 B_j(t)\, B_i(t)\, dt \qquad (7.20)$$

Now let

$$\underset{\sim}{X}(z) = [X_1(z),\ X_2(z),\ldots,X_{N+1}(z)]^T \qquad (7.21)$$

$$\bar{\underset{\sim}{X}}(z) = \underset{\sim}{X}'(z) \qquad (7.22)$$

$$\underset{\sim}{Y}(z) = [Y_1(z),\ Y_2(z),\ldots,Y_{N+1}(z)]^T \qquad (7.23)$$

$$\bar{\underset{\sim}{Y}}(z) = \underset{\sim}{Y}'(z) \qquad (7.24)$$

Let the matrices A and C and the vector $\underset{\sim}{b}$ be defined as

$$A_{ji} = \int_0^1 B_j(t)\, B_i(t)\, dt\,,\ \ i,j = 1,\ldots,N+1, \qquad (7.25)$$

$$C_{ji} = \int_0^1 B_j(t)\, \dot{B}_i(t)\, dt\,,\ \ i,j = 1,\ldots,N+1, \qquad (7.26)$$

$$b_j = \int_0^1 B_j(t)\, dt \quad , \qquad j = 1,\ldots,N+1 \qquad (7.27)$$

Then the equations (7.19) and (7.20) can be written as

$$\underset{\sim}{X}'(z) = \overline{\underset{\sim}{X}}(z) \tag{7.28}$$

$$\overline{\underset{\sim}{X}}'(z) = f\,Pe_m\,\overline{\underset{\sim}{X}}(z) + Pe_m\,D_m\,R_y\,(\overline{y},\overline{x})\,\underset{\sim}{Y}(z)$$

$$+ Pe_m\,D_m\,R_x\,(\overline{y},\overline{x})\,\underset{\sim}{X}(z) - Pe_m\,D_m\,\overline{y}\,R_y\,(\overline{y},\overline{x})\,A^{-1}\,\underset{\sim}{b} \tag{7.29}$$

$$\underset{\sim}{Y}'(z) = \overline{\underset{\sim}{Y}}(z) \tag{7.30}$$

$$\overline{\underset{\sim}{Y}}'(z) = f\,Pe_h\,\overline{\underset{\sim}{Y}}(z) + (Pe_h\,A^{-1}\,C - \beta\,Pe_h\,D_m\,R_y\,(\overline{y},\overline{x})\,I_{N+1})\,\underset{\sim}{Y}(z)$$

$$-\,Pe_h\,\beta\,D_m\,R_x\,(\overline{y},\overline{x})\,\underset{\sim}{X}(z) + Pe_h\,\beta\,D_m\,\overline{y}\,R_y\,(\overline{y},\overline{x})\,A^{-1}\,\underset{\sim}{b} \tag{7.31}$$

Let

$$\underset{\sim}{\chi}(z) = [\underset{\sim}{X}^T(z),\ \overline{\underset{\sim}{X}}^T(z),\ \underset{\sim}{Y}^T(z),\ \overline{\underset{\sim}{Y}}^T(z)]^T \tag{7.32}$$

Then

$$\underset{\sim}{\chi}'(z) = \Gamma\,\underset{\sim}{\chi}(z) + \underset{\sim}{\hat{\Gamma}}, \tag{7.33a}$$

where

$$\Gamma = \begin{bmatrix} 0 & I_{N+1} & 0 & 0 \\ K_1\, I_{N+1} & K_2\, I_{N+1} & K_3\, I_{N+1} & 0 \\ 0 & 0 & 0 & I_{N+1} \\ K_4\, I_{N+1} & 0 & \hat{A} & K_5\, I_{N+1} \end{bmatrix} \qquad (7.33b)$$

$$\hat{\underset{\sim}{\Gamma}} = \begin{bmatrix} \underset{\sim}{0} \\ \underset{\sim}{K_6} \\ \underset{\sim}{0} \\ \underset{\sim}{K_7} \end{bmatrix}, \qquad (7.33c)$$

where

$$K_1 = Pe_m\, D_m\, R_x\, (\bar{y},\bar{x}), \qquad (7.33d)$$

$$K_2 = f\, Pe_m, \qquad (7.33e)$$

$$K_3 = Pe_m\, D_m\, R_y\, (\bar{y},\bar{x}), \qquad (7.33f)$$

$$K_4 = -\, Pe_h\, \beta\, D_m\, R_x\, (\bar{y},\bar{x}), \qquad (7.33g)$$

$$K_5 = f\, Pe_h, \qquad (7.33h)$$

$$\underset{\sim}{K_6} = -\, Pe_m\, D_m\, \bar{y}\, R_y\, (\bar{y},\bar{x})\, A^{-1}\, \underset{\sim}{b} \qquad (7.33i)$$

$$\underset{\sim}{K_7} = Pe_h\, \beta\, D_m\, \bar{y}\, R_y\, (\bar{y},\bar{x})\, A^{-1}\, \underset{\sim}{b} \qquad (7.33j)$$

$$\hat{A} = Pe_h\, A^{-1}\, C - \beta\, Pe_h\, D_m\, R_y\, (\bar{y},\bar{x})\, I_{N+1} \qquad (7.33k)$$

0 is the N + 1 by N + 1 zero matrix, $\underset{\sim}{0}$ is the N + 1 zero vector, and I_{N+1} is the N + 1 by N + 1 identity matrix.

So far, the functions $B_i(t)$ have not been specified. A suitable choice of $B_i(t)$ is the linear B splines function defined on the time interval [0,1] with N equal subintervals. That is, let

$$B_i(t) = S_1(N\,t - i), \tag{7.34}$$

where

$$S_1(t) = \begin{cases} 1 + t, & t \in [-1,0] \\ 1 - t, & t \in [0,1] \\ 0, & \text{otherwise} \end{cases} \tag{7.35}$$

Then it can be easily verified that the elements of A, C and $\underset{\sim}{b}$ are given by

$$A_{1,1} = A_{N+1,N+1} = \frac{1}{3N} \qquad A_{i,i} = \frac{2}{3N}, \qquad i = 2,\ldots,N$$

$$A_{i,i+1} = \frac{1}{6N},\ i = 1,\ldots,N \qquad A_{i,i-1} = \frac{1}{6N}, \qquad i = 2,\ldots,N+1 \tag{7.36}$$

and $A_{i,j} = 0$ otherwise

$$C_{1,1} = -\frac{1}{2} \qquad C_{N+1,N+1} = \frac{1}{2}$$

$$C_{i,i+1} = \frac{1}{2},\ i = 1,\ldots,N \qquad C_{i,i-1} = -\frac{1}{2}, \qquad i = 2,\ldots,N+1 \tag{7.37}$$

and $C_{i,j} = 0$, otherwise

$$\underset{\sim}{b}_1 = \underset{\sim}{b}_{N+1} = \frac{1}{2N} \qquad \underset{\sim}{b}_i = \frac{1}{N}, \qquad i = 2,\ldots,N \tag{7.38}$$

From (7.32), it is clear that

$$\underset{\sim}{\chi}(z) = e^{\Gamma z}\underset{\sim}{\chi}(0) + \int_0^z e^{\Gamma(z-\tau)}\hat{\underset{\sim}{\Gamma}}\, d\tau$$

$$= e^{\Gamma z}\underset{\sim}{\chi}(0) + \underset{\sim}{r}(z) \ , \tag{7.39a}$$

where

$$\underset{\sim}{r}(z) = (e^{\Gamma z} - I_{4N+4})\,\Gamma^{-1}\hat{\underset{\sim}{\Gamma}} \tag{7.39b}$$

Now, let $y_d(t) = y_d$ be a constant in (7.10). Then, in view of (7.12), (7.25) and (7.33), the objective function (7.10) can be written as

$$\text{minimize } J = \int_0^1 \left[\sum_{i=1}^{N+1} B_i(t)\ Y_i(z_f) - y_d\right]^2 dt$$

$$= \int_0^1 \left[\sum_{i=1}^{N+1} B_i^2(t)\ Y_i^2(z_f) + 2\sum_{i=1}^{N} B_i(t)\ B_{i+1}(t)\ Y_i(z_f)\ Y_{i+1}(z_f)\right.$$

$$\left. - 2\,y_d \sum_{i=1}^{N+1} B_i(t)\ Y_i(z_f) + y_d^2\right] dt$$

$$= \underset{\sim}{Y}^T(z_f)\ \hat{H}\ \underset{\sim}{Y}(z_f) + \hat{\underset{\sim}{g}}^T\ \underset{\sim}{Y}(z_f) + y_d^2, \tag{7.40}$$

where the matrix $\hat{H}$ and the vector $\hat{\underset{\sim}{g}}$ are given by

$\hat{H}_{1,1} = \hat{H}_{N+1,N+1} = \frac{1}{3N}$ $\qquad$ $\hat{H}_{i,j} = \frac{2}{3N}$ $\qquad$ $i = 2,\ldots,N$

$\hat{H}_{i,i+1} = \frac{1}{6N}$, $\qquad$ $i = 1,\ldots,N$ $\qquad$ $\hat{H}_{i,i-1} = \frac{1}{6N}$, $\qquad$ $i = 2,\ldots,N+1$

$\hat{H}_{i,j} = 0$, otherwise $\qquad$ (7.41)

$$\hat{g}_i = -\frac{y_d}{N}, \qquad i = 1,\dots,N+1$$

$$\hat{g}_i = -\frac{2y_d}{N}, \qquad i = 2,\dots,N \tag{7.42}$$

In view of (7.32), the objective function (7.40) can be written as

$$\text{minimize } J = \chi^T(z_f)\, H^*\, \chi(z_f) + g^{*T}\, \chi(z_f) + y_d^2 \tag{7.43}$$

where

$$H^* = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{H} & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{7.44}$$

$$g^* = \begin{bmatrix} \underset{\sim}{0} \\ \underset{\sim}{0} \\ \hat{g} \\ \underset{\sim}{0} \end{bmatrix}, \tag{7.45}$$

Substitute (7.39a) into (7.43), we get

$$\text{minimize } J = \underset{\sim}{\chi}(0)^T\, \hat{H}\, \underset{\sim}{\chi}(0) + \hat{g}^T\, \chi(0) + \hat{k}, \tag{7.46}$$

where

$$\hat{H} = (e^{\Gamma z_f})^T\, H^*\, e^{\Gamma z_f} \tag{7.47}$$

$$\hat{g} = (e^{\Gamma z_f})^T [(H + H^{*T}) \underset{\sim}{r}(z_f) + \underset{\sim}{g}^*] \tag{7.48}$$

$$\hat{k} = \underset{\sim}{r}^T(z_f) H^* \underset{\sim}{r}(z_f) + (\underset{\sim}{g}^*)^T \underset{\sim}{r}(z_f) + y_d^2 \tag{7.49}$$

Now, in view of (7.12), (7.13), (7.24), (7.22), (7.32), the boundary condition (7.4) can be written as

$$\sum_{i=1}^{N+1} B_i(t)\, Y_i'(z_f) = \sum_{i=1}^{N+1} B_i(t)\, X_i'(z_f) = 0$$

$$\Longrightarrow \quad \sum_{i=1}^{N+1} B_i(t)\, \overline{Y}_i(z_f) = \sum_{i=1}^{N+1} B_i(t)\, \overline{X}_i(z_f) = 0$$

$$\Longrightarrow \quad \sum_{i=1}^{N+1} B_i(t)\, \chi_{3N+3+i}(z_f) = \sum_{i=1}^{N+1} B_i(t)\, \chi_{N+1+i}(z_f) = 0 \tag{7.50}$$

By letting t to be equal to the grid points $0, \frac{1}{N}, \frac{2}{N}, \ldots, 1$, we easily obtain from (7.50) and (7.39a) that

$$\left\{ e^{\Gamma z_f} \chi(0) + \underset{\sim}{r}(z_f) \right\}_i = 0 \ , \tag{7.51}$$

where $i = \begin{array}{l} N+2,\ldots,2N+2 \text{ and} \\ 3N+4,\ldots,4N+4. \end{array}$

Similarly, the boundary condition (7.6) can be written as

$$\overline{\underset{\sim}{Y}}(0) = \overline{\underset{\sim}{X}}(0) = 0$$

$$\Longrightarrow \quad \{\chi(0)\}_i = 0, \tag{7.52}$$

$$i = \begin{matrix} N+2,\ldots,2N+2 \text{ and} \\ 3N+4,\ldots,4N+4. \end{matrix}$$

Now, let

$$\underset{\sim}{y} \equiv \chi(0) \tag{7.53}$$

Then, in view of (7.46), (7.51) and (7.52), the problem with the gradient of the inlet temperature and the gradient of the inlet oxygen concentration equal to zero can be written as

$$\underset{\underset{\sim}{y} \in R^{4N+4}}{\text{minimize}} \; J = \frac{1}{2}\underset{\sim}{y}^T \hat{H} \underset{\sim}{y} + \hat{\underset{\sim}{g}}^T \underset{\sim}{y} + \hat{k}$$

subject to

$$\left\{e^{\Gamma z_f} \underset{\sim}{y}\right\}_i = -\Gamma_i(z_f)$$

(7.*P*1) $\qquad i = N+2,\ldots,2N+2 \text{ and } 3N+4,\ldots,4N+4$

and

$$\underset{\sim}{y}_i = 0$$

$$i = N+2,\ldots,2N+2 \text{ and } 3N+4,\ldots,4N+4$$

Similarly, the problem with the gradient of the inlet temperature and the gradient of the inlet oxygen concentration being regarded as additional controls can be written as

(7.*P*2) minimizes $J = \frac{1}{2} \underset{\sim}{y}^T \hat{H} \underset{\sim}{y} + \hat{\underset{\sim}{g}}^T \underset{\sim}{y} + \hat{k}$
$\underset{\sim}{y} \in R^{4N+4}$

subject to

$$\left\{ e^{\Gamma z_f} \underset{\sim}{y} \right\}_i = -\Gamma_i (z_f)$$

$i = N+2, \ldots, 2N+2$ and $3N+4, \ldots, 4N+4$

Both (7.*P*1) and (7.*P*2) are quadratic programming problems with linear constraints, which can be easily solved by any standard quadratic programming software.

Once the optimal value of $\underset{\sim}{y}$ is found, the optimal control can be easily found by the formula

$$x_I(t) = x(0,t)$$

$$= \sum_{i=1}^{N+1} B_i(t)\, X_i(0)$$

$$= \sum_{i=1}^{N+1} B_i(t)\, \underset{\sim}{y}_i(0) \qquad (7.53)$$

$$y_I(t) = y(0,t)$$

$$= \sum_{i=1}^{N+1} B_i(t)\, Y_i(0)$$

$$= \sum_{i=1}^{N+1} B_i(t)\, \underset{\sim}{\chi}_{2N+2+i}(0) \qquad (7.54)$$

$$\bar{x}_I(t) = \frac{\partial x \ (0,t)}{\partial z}$$

$$= \sum_{i=1}^{N+1} B_i(t)\, X_i'(0)$$

$$= \sum_{i=1}^{N+1} B_i(t)\, u_{N+1+i} \tag{7.55}$$

$$\bar{y}_I(t) = - \cdot \frac{\partial y \ (0,t)}{\partial z}$$

$$= \sum_{i=1}^{N+1} B_i(t)\, Y_i'(0)$$

$$= \sum_{i=1}^{N+1} B_i(t)\, u_{3N+3+i} \tag{7.56}$$

The temperature and concentration and their respective gradient at any time t and at any position z can be easily obtained by the following formulae

$$x\,(z,t) = \sum_{i=1}^{N+1} B_i(t)\, \chi_i(z) \tag{7.57}$$

$$y\,(z,t) = \sum_{i=1}^{N+1} B_i(t)\, \chi_{2N+2+i}(z) \tag{7.58}$$

$$x'\,(z,t) = \sum_{i=1}^{N+1} B_i(t)\, \chi_{N+1+i}(z) \tag{7.59}$$

$$y'(z,t) = \sum_{i=1}^{N+1} B_i(t)\, \chi_{2N+2+i}(z), \tag{7.60}$$

where $\underset{\sim}{\chi}(z)$ can be found from (7.40) and (7.53).

7.4 Computational Aspects and Implementation

Only two aspects of implementing the above method need discussion. They are (i) evaluation of $\underset{\sim}{\Gamma}(z)$ and (ii) calculation of the exponential matrix $e^{\Gamma z}$.

From (7.39b), the evaluation of $\underset{\sim}{\gamma}(z)$ involves the calculation of the inverse of Γ^{-1}. However, Γ is extremely ill–conditioned – the reciprocal of Γ's condition number is of order 10^{-18}. Thus, any attempt to calculate the inverse of Γ numerically failed. Thus $\underset{\sim}{\gamma}(z)$ cannot be calculated directly from (7.39b). Instead we use the Gaussian quadrature rule of order 5 to evaluate the integral

$$\underset{\sim}{\gamma}(z) = \int_0^z e^{\Gamma(z-\tau)}\, \hat{\underset{\sim}{\Gamma}}\, d\tau$$

The ill–conditioning of Γ also affected the choice of method to calculate the exponential matrix $e^{\Gamma z}$. Numerical methods for calculating the eigenvalues of Γ all failed. Thus, we need to find some methods which does not involve the eigenvalues of Γ. Two approaches can be adopted.

Firstly, we consider the series expansion of the exponential matrix, viz

$$e^{\Gamma z} = \sum_{i=0}^{\infty} \frac{\Gamma^i z^i}{i!} \tag{7.61}$$

The series is considered to have converged when the difference between two successive intermediate matrices is less than 10^{-10} in both the L_1 and L_∞ matrix norms. These norms are calculated by the IMSL routines DNR1RR and DNRIRR respectively. However, the series expansion method to calculate the exponential matrix can sometimes create two types of errors, namely, roundoff and cancellation errors, which can render the method unusable. Thus, a second method is used in order to check the accuracy of the result that we obtained in the series expansion method. This method involves scaling and the Padé approximation (cf. Ref. 119). It has been found that the rate of convergence of the second method was faster than that of the first method, but there was no significant difference between the results produced by the two methods.

A further difficulty in calculating the exponential matrix is encountered in the choice of z_f. For $z_f > 0.1$, the exponential matrix consists of some elements which is so large that they cannot be handled by the computer. As a compromise, we choose z_f to be 0.1.

Finally, the IMSL routine DQPROG, based on an algorithm by Goldfarb is used to solve the quadratic programming problem (7.$P1$) and (7.$P2$).

7.5 Illustrative Examples

Example 7.5.1 We consider the problem (7.$P2$) with the following desired outlet temperature y_d = 0.7, 0.63, 0.5 and 0.8, which corresponds to 31° C, 0° C, – 55° C and 74° C respectively.

For y_d = 0.7, the problem has been solved with N = 4, 8 and 12. For y_d = 0.63 and 0.5, the problem has been solved with N = 8 only. For y_d = 0.8, the problem has been solved with N = 4 only.

From the results tabulated in Tables 7.5.1–7.5.6, it is clear that all the optimal solutions have the following same properties:

(i) The optimal value of J is extremely close to zero, which shows that the actual output temperature is almost the same as the desired output temperature.

(ii) The boundary condition (7.4) is satisfied with the value of $x'(z_f, t)$ and $y'(z_f, t)$ being of order 10^{-14} or less.

(iii) The controls $x(0,t)$ and $y(0,t)$ are almost constant functions of t, with the input temperature $y(0,t)$ almost the same as the desired output temperature for all $t \in [0,1]$. The variations of the controls $x'(0,t)$ and $y'(0,t)$ are also very small as t varies from 0 to 1.

(iv) For each fixed z, the oxygen concentration $x(z,t)$ and the temperature $y(z,t)$ remain almost constant with time.

(v) For each fixed t, the oxygen concentration $x(z,t)$ increases along the reactor, while the temperature $y(z,t)$ decreases along the reactor.

Moreover, from the results tabulated in Tables 7.5.1–7.5.3 only, we can also observe the following convergence properties of the optimal control with respect to the number of partitions

when y_d is fixed at 0.7, the optimal controls $x(0,t)$ and $y(0,t)$ will converge to 0.5 and 0.71 respectively for all $t \in [0,1]$, as the number of partitions N increases. However, this convergence property is not so prominent in the controls $x'(0,t)$ and $y'(0,t)$.

Graphs of the optimal controls for $y_d = 0.7$ and N = 12 are plotted in Figure 7.5.1.

Example 7.5.2 We consider the Problem (7.*P*2) with the desired outlet temperature $y_d = 0.7$ and the following additional constraint

$$x(0,t) \leq 0.3 \tag{7.62}$$

The problem has been solved with N = 4. From the results tabulated in Table 7.5.7, it is clear that the optimal solutions obtained in this example have exactly the same properties as those obtained in Example 7.5.1.

Example 7.5.3 Same as Example 7.5.2 except that the constraint (7.62) in Example 7.5.2 is now being replaced by

$$x(0,t) \geq 0.6 \tag{7.63}$$

The problem has been solved with N = 4. From the results tabulated in Table 7.5.8, it is clear that the properties of the optimal solutions obtained in these examples are almost the same as those obtained in Examples 7.5.1 and 7.5.2, with the exception that the last property obtained in both Examples, 7.5.1 and 7.5.2, is now being reversed, i.e. for each fixed t, the oxygen concentration $x(z,t)$ decreases along the reactor, while the temperature $y(z,t)$ increases along the reactor.

Example 7.5.4 Same as Example 7.5.2, except that the constraint (7.62) in Example 7.5.2 is now being replaced by

$$y(0,t) \leq 0.6 \tag{7.64}$$

The problem has been solved with N = 4. From the results tabulated in Table 7.5.9, it is clear that the optimal solutions obtained in this example have exactly the same properties as those obtained in Example 7.5.3.

Example 7.5.5 We consider the problem (7.P2) with the desired outlet temperature $y_d = 0.7$ and the following additional constraints

$$0.0 \le x(0, 0.0) \le 0.2 \tag{7.65}$$

$$0.2 \le x(0, 0.25) \le 0.4 \tag{7.66}$$

$$0.4 \le x(0, 0.50) \le 0.6 \tag{7.67}$$

$$0.6 \le x(0, 0.75) \le 0.8 \tag{7.68}$$

$$0.8 \le x(0, 1.00) \le 1.0 \tag{7.69}$$

The problem has been solved with N = 4. From the results tabulated in Table 7.5.10, it is clear that the optimal solution have the following properties:

(i) Exactly the same as the first property obtained in the optimal solution of Example 7.5.1.

(ii) Exactly the same as the second property obtained in the optimal solution of Example 7.5.1.

(iii) The input temperature $y(0,t)$ still remains almost the same as the desired output temperature for all $t \in [0,1]$. However, all the other controls $x(0,t)$, $x'(0,t)$ and $y'(0,t)$ are not constant functions of t any more. In fact, $x(0,t)$

switches from its upper bounds when $t = 0.2$ and 0.4 to its lower bounds when $t = 0.8$ and 1.0. (See Figure 7.5.2).

(iv) For each fixed z, the temperature $y(z,t)$ still remains almost constant with time, but not the oxygen concentration $x(z,t)$.

(v) For certain values of t, $x(z,t)$ and $y(z,t)$ may increase along the chemical reactor, while for other values of t, $x(z,t)$ and $y(z,t)$ may decrease only the chemical reactor.

Graph of optimal control $x(0,t)$, together with its upper and lower bounds, are plotted in Figure 7.5.2.

Example 7.5.6 We consider the problem (7.P2) with the desired outlet temperature $y_d = 0.7$ and the following additional constraints

$$0.0 \leq y(0, 0.0) \leq 0.2 \tag{7.70}$$

$$0.2 \leq y(0, 0.25) \leq 0.4 \tag{7.71}$$

$$0.4 \leq y(0, 0.50) \leq 0.6 \tag{7.72}$$

$$0.6 \leq y(0, 0.75) \leq 0.8 \tag{7.73}$$

$$0.8 \leq y(0, 1.00) \leq 1.0 \tag{7.74}$$

The problem has been solved with N = 4. From the results tabulated in Table 7.5.11, it is clear that the optimal solution have the following properties:

(i) Exactly the same as the first property obtained in the optimal solution of Example 7.5.1.

(ii) Exactly the same as the second property obtained in the optimal solution of Example 7.5.1.

(iii) None of the controls $x(0,t)$, $y(0,t)$, $x'(0,t)$ and $y'(0,t)$ is a constant function of t. In fact, $y(0,t)$ switches from its lower bounds when $t = 0.2$ and 0.4 to its upper bounds when $t = 0.6$ and back to its lower bounds when $t = 0.8$ and 1.0. (See Figure 7.5.3.)

(iv) For every $z \neq z_f$, both the temperature $x(z,t)$ and oxygen concentration $y(z,t)$ are not constant functions of time t.

(v) Same as the last property obtained in the optimal solution of Example 7.5.5.

Graph of the optimal control $y(0,t)$, together with its upper and lower bounds, are plotted in Figure 7.5.3

7.6 Conclusion

From the numerical results obtained in Section 7.5, it appears that the method described in this paper works very well for the problem (7.*P*2) described in Section 7.3. In all the numerical examples solved, the actual output temperature obtained is almost the same as the desired output temperature. In general, an increase in the number of partitions N over the time interval [0,1] will lead to the convergence of the optimal control of the approximate problems to the true optimal control.

By imposing various restrictions on the control, the nature of the optimal solutions will change considerably.

Unfortunately, we have not successfully solved any numerical examples for the problem (7.*P*1) described in Section 7.3 at this moment.

Table 7.5.1 Computed Results ($y_d = 0.7$ $N = 4$)

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.00	4.97516E-01	1.38732E-01	7.01047E-01	-1.92872E-02
0.25	5.05336E-01	1.10186E-01	7.01053E-01	-2.34300E-02
0.50	4.86080E-01	1.78394E-01	7.01533E-01	-3.05826E-02
0.75	4.86434E-01	1.42181E-01	7.01164E-01	-2.27175E-02
1.00	4.84256E-01	1.84058E-01	7.01796E-01	-4.11886E-02

Objective Value = -0.6938893903E-16

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	5.00029E-01	5.02012E-01	5.03453E-01	5.04334E-01	5.04633E-01
0.25	5.07337E-01	5.08922E-01	5.10075E-01	5.10781E-01	5.11021E-01
0.50	4.89316E-01	4.91874E-01	4.93735E-01	4.94874E-01	4.95260E-01
0.75	4.99011E-01	5.01048E-01	5.02528E-01	5.03434E-01	5.03742E-01
1.00	4.87600E-01	4.90249E-01	4.92177E-01	4.93356E-01	4.93757E-01

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	1.12484E-01	8.57263E-02	5.82128E-02	2.87105E-02	-3.90799E-14
0.25	8.97797E-02	6.85881E-02	4.66176E-02	2.37976E-02	-3.55271E-14
0.50	1.45023E-01	1.10661E-01	7.52014E-02	3.83864E-02	-2.84217E-14
0.75	1.15464E-01	8.80710E-02	5.98250E-02	3.05358E-02	-3.90799E-14
1.00	1.50037E-01	1.14640E-01	7.79208E-02	3.97778E-02	-3.19744E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	7.00693E-01	7.00403E-01	7.00165E-01	7.00047E-01	7.00000E-01
0.25	7.00646E-01	7.00353E-01	7.00154E-01	7.00038E-01	7.00000E-01
0.50	7.00981E-01	7.00551E-01	7.00245E-01	7.00061E-01	7.00000E-01
0.75	7.00753E-01	7.00429E-01	7.00193E-01	7.00049E-01	7.00000E-01
1.00	7.01093E-01	7.00593E-01	7.00258E-01	7.00064E-01	7.00000E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	-1.61320E-02	-1.27979E-02	-8.98203E-03	-4.67415E-03	3.77476E-15
0.25	-1.73344E-02	-1.22226E-02	-7.81147E-03	-3.82223E-03	4.44089E-16
0.50	-2.45792E-02	-1.84049E-02	-1.22382E-02	-6.12209E-03	4.44089E-16
0.75	-1.83770E-02	-1.40340E-02	-9.53239E-03	-4.84[illegible]6E-03	-1.11022E-15
1.00	-2.96715E-02	-2.06298E-02	-1.30880E-02	-6.39300E-03	8.88176E-16

Table 7.5.2 Computed Results ($y_d = 0.7$ N = 8)

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.000	4.97517E-01	1.35067E-01	7.01116E-01	-2.16404E-02
0.125	4.99949E-01	1.29570E-01	7.01115E-01	-2.28259E-02
0.250	4.97644E-01	1.37607E-01	7.01200E-01	-2.44272E-02
0.375	4.96782E-01	1.40818E-01	7.01183E-01	-2.33038E-02
0.500	4.96869E-01	1.40495E-01	7.01185E-01	-2.34990E-02
0.675	4.97177E-01	1.39504E-01	7.01151E-01	-2.22720E-02
0.750	4.98842E-01	1.33626E-01	7.01107E-01	-2.17522E-02
0.875	4.99363E-01	1.31670E-01	7.01121E-01	-2.75614E-02
1.000	4.98809E-01	1.33601E-01	7.01142E-01	-2.29424E-02

Objective Value = -0.1332267629E-14

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	5.00120E-01	5.02097E-01	5.03534E-01	5.04413E-01	5.04711E-01
0.125	5.02299E-01	5.04157E-01	5.05508E-01	5.06335E-01	5.06616E-01
0.250	5.00140E-01	5.02115E-01	5.03551E-01	5.04429E-01	5.04728E-01
0 375	4.95335E-01	5.01354E-01	5.02821E-01	5.03719E-01	5.04024E-01
0.500	4.99417E-01	5.01431E-01	5.02896E-01	5.03791E-01	5.04096E-01
0.675	4.99706E-01	5.01708E-01	5.03158E-01	5.04047E-01	5.04349E-01
0.750	5.01265E-01	5.03179E-01	5.04571E-01	5.05423E-01	5.05712E-01
0.875	5.01751E-01	5.03639E-01	5.05012E-01	5.05852E-01	5.06137E-01
1.000	5.01232E-01	5.03148E-01	5.04541E-01	5.05393E-01	5.05683E-01

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	1.12094E-01	8.64885E-02	5.80673E-02	2.96382E-02	-6.03961E-14
0.125	1.05330E-01	8.03801E-02	5.46102E-02	2.78751E-02	-7.10543E-14
0.250	1.11898E-01	8.54097E-02	5.80327E-02	2.96228E-02	-5.32907E-14
0.375	1.14420E-01	8.73025E-02	5.93107E-02	3.02743E-02	-6.75016E-14
0.500	1.14167E-01	8.71106E-02	5.91808E-02	3.02079E-02	-5.68434E-14
0.675	1.13311E-01	8.64414E-02	5.87220E-02	2.99733E-02	-6.39488E-14
0.750	1.08544E-01	8.28043E-02	5.62508E-02	2.87117E-02	-6.75016E-14
0.875	1.07015E-01	8.16599E-02	5.54786E-02	2.83183E-02	-5.68434E-14
1.000	1.08594E-01	8.28695E-02	5.63020E-02	2.87388E-02	-6.03951E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	7.00724E-01	7.00414E-01	7.00186E-01	7.00047E-01	7.00000E-01
0.125	7.00709E-01	7.00398E-01	7.00177E-01	7.00044E-01	7.00000E-01
0.250	7.00763E-01	7.00427E-01	7.00189E-01	7.00047E-01	7.00000E-01
0.375	7.00761E-01	7.00430E-01	7.00192E-01	7.00048E-01	7.00000E-01
0.500	7.00761E-01	7.00430E-01	7.00192E-01	7.00048E-01	7.00000E-01
0.675	7.00745E-01	7.00423E-01	7.00190E-01	7.00048E-01	7.00000E-01
0.750	7.00714E-01	7.00406E-01	7.00182E-01	7.00046E-01	7.00000E-01
0.875	7.00717E-01	7.00404E-01	7.00180E-01	7.00045E-01	7.00000E-01
1.000	7.00730E-01	7.00411E-01	7.00183E-01	7.00046E-01	7.00000E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	-1.75657E-02	-1.34773E-02	-9.20156E-03	-4.69348E-03	4.44089E-15
0.125	-1.78728E-02	-1.32694E-02	-8.83458E-03	-4.43721E-03	0.00000E+00
0.250	-1.93082E-02	-1.43135E-02	-9.46890E-03	-4.72685E-03	9.99201E-15
0.375	-1.88365E-02	-1.42440E-02	-9.56294E-03	-4.81652E-03	-2.22045E-15
0.500	-1.88878E-02	-1.42346E-02	-8.54425E-03	-4.80590E-03	5.55112E-15
0.675	-1.82360E-02	-1.39235E-02	-9.41306E-03	-4.78172E-03	5.21725E-15
0.750	-1.75568E-02	-1.33342E-02	-9.00806E-03	-4.55950E-03	2.22045E-15
0.875	-1.79161E-02	-1.34103E-02	-8.96181E-03	-4.50671E-03	1.33227E-15
1.000	-1.82709E-02	-1.36689E-02	-9.11728E-03	-4.57680E-03	6.66134E-15

Table 7.5.3 Computed Results ($y_d = C$ ' N = 12)

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.000	4.97534E-01	1.38358E-01	7.01120E-01	-2.18285E-02
0.083	4.99239E-01	1.32099E-01	7.01134E-01	-2.34726E-02
0.167	4.98308E-01	1.35236E-01	7.01178E-01	-2.30746E-02
0.250	4.97238E-01	1.39698E-01	7.01044E-01	-1.80655E-02
0.333	5.03055E-01	1.18884E-01	7.01023E-01	-2.21050E-02
0.417	4.99938E-01	1.29434E-01	7.01140E-01	-2.21652E-02
0.500	5.00688E-01	1.27564E-01	7.00929E-01	-1.57087E-02
0.583	5.04938E-01	1.11922E-01	7.01001E-01	-2.25596E-02
0.667	5.02008E-01	1.21969E-01	7.01128E-01	-2.32861E-02
0.750	4.99706E-01	1.30664E-01	7.01049E-01	-1.92315E-02
0.833	5.01571E-01	1.23964E-01	7.01053E-01	-2.24149E-02
0.917	5.00453E-01	1.27096E-01	7.01289E-01	-3.09874E-02
1.000	4.92021E-01	1.56359E-01	7.01645E-01	-8.93462E-02

Objective Value = 0.1595945597E-14

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	5.00041E-01	5.02023E-01	5.03463E-01	5.04344E-01	5.04643E-01
0.083	5.01635E-01	5.03530E-01	5.04907E-01	5.05750E-01	5.06036E-01
0.167	5.00761E-01	5.02702E-01	5.04113E-01	5.04977E-01	5.05270E-01
0.250	4.99768E-01	5.01765E-01	5.03216E-01	5.04103E-01	5.04404E-01
0.3?1	5.05208E-01	5.06910E-01	5.08147E-01	5.08904E-01	5.09161E-01
0.41;	5.02286E-01	5.04144E-01	5.05496E-01	5.06323E-01	5.06604E-01
0.500	5.02980E-01	5.04804E-01	5.06129E-01	5.06939E-01	5.07214E-01
0.583	5.06968E-01	5.08575E-01	5.09743E-01	5.10457E-01	5.10700E-01
0.667	5.04222E-01	5.05875E-01	5.07251E-01	5.08031E-01	5.08296E-01
0.750	5.02074E-01	5.03945E-01	5.05306E-01	5.06138E-01	5.06420E-01
0.833	5.03819E-01	5.05596E-01	5.06888E-01	5.07678E-01	5.07946E-01
0.917	5.02763E-01	5.04594E-01	5.05927E-01	5.06742E-01	5.07019E-01
1.000	4.94864E-01	4.97120E-01	4.98762E-01	4.99766E-01	5.00109E-01

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	1.12331E-01	8.56702E-02	5.81912E-02	2.97015E-02	-9.23706E-14
0.083	1.07377E-01	8.19340E-02	5.56627E-02	2.84119E-02	-6.75016E-14
0.167	1.09975E-01	8.3957?3E-02	5.70474E-02	2.91205E-02	-8.52651E-14
0.250	1.13255E-01	8.63249E-02	5.86256E-02	2.99222E-02	-7.46070E-14
0.333	9.64722E-02	7.36018E-02	4.99975E-02	2.55186E-02	-6.88178E-14
0.417	1.05786E-01	8.03878E-02	5.46287E-02	2.78863E-02	-8.17124E-14
0.500	1.03448E-01	7.88307E-02	5.35312E-02	2.73215E-02	-7.81597E-14
0.583	9.10452E-02	6.94831E-02	4.72045E-02	2.40945E-02	-8.52651E-14
0.667	9.93138E-02	7.58589E-02	5.15578E-02	2.63194E-02	-8.17124E-14
0.750	1.06079E-01	8.09134E-02	5.49660E-02	2.80563E-02	-8.88178E-14
0.833	1.00733E-01	7.68453E-02	5.22002E-02	2.66439E-02	-6.88178E-14
0.917	1.03684E-01	7.92530E-02	5.38723E-02	2.75015E-02	-9.59233E-14
1.000	1.27694E-01	9.76541E-02	6.63874E-02	3.38977E-02	-7.81597E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	7.00726E-01	7.00415E-01	7.00187E-01	7.00047E-01	7.00000E-01
0.083	7.00719E-01	7.00404E-01	7.00180E-01	7.00045E-01	7.00000E-01
0.167	7.00755E-01	7.00422E-01	7.00187E-01	7.00047E-01	7.00000E-01
0.250	7.00702E-01	7.00410E-01	7.00187E-01	7.00046E-01	7.00000E-01
0.333	7.00641E-01	7.00360E-01	7.00161E-01	7.00041E-01	7.00000E-01
0.417	7.00730E-01	7.00407E-01	7.00179E-01	7.00045E-01	7.00000E-01
0.500	7.00630E-01	7.00371E-01	7.00170E-01	7.00043E-01	7.00000E-01
0.583	7.00618E-01	7.00343E-01	7.00;53E-01	7.00038E-01	7.00000E-01
0.667	7.00707E-01	7.00389E-01	7.00170E-01	7.00042E-01	7.00000E-01
0.750	7.00680E-01	7.00395E-01	7.00178E-01	7.00045E-01	7.00000E-01
0.833	7.00665E-01	7.00375E-01	7.00168E-01	7.00042E-01	7.00000E-01
0.917	7.00770E-01	7.00413E-01	7.00179E-01	7.00044E-01	7.00000E-01
1.000	7.00960E-01	7.00621E-01	7.00223E-01	7.00055E-01	7.00000E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	-1.76051E-02	-1.35103E-02	-9.22867E-03	-4.70523E-03	-6.66134E-16
0.083	-1.81641E-02	-1.34258E-02	-8.85616E-03	-4.51601E-03	3.55271E-15
0.167	-1.89907E-02	-1.42112E-02	-9.37347E-03	-4.65632E-03	6.66134E-15
0.250	-1.60078E-02	-1.30198E-02	-9.15702E-03	-4.72545E-03	5.32907E-15
0.333	-1.63682E-02	-1.19250E-02	-7.87249E-03	-4.04403E-03	1.11022E-14
0.417	-1.84610E-02	-1.37804E-02	-9.03522E-03	-4.46681E-03	5.99520E-15
0.500	-1.41043E-02	-1.16561E-02	-8.28152E-03	-4.30605E-03	1.53211E-14
0.583	-1.61682E-02	-1.15156E-02	-7.59511E-03	-3.82616E-03	9.10383E-15
0.667	-1.84609E-02	-1.33743E-02	-8.62023E-03	-4.22614E-03	1.28786E-14
0.750	-1.64697E-02	-1.29081E-02	-8.80947E-03	-4.45882E-03	1.24345E-14
0.833	-1.67310E-02	-1.23604E-02	-8.31957E-03	-4.22380E-03	6.66134E-16
0.917	-2.14752E-02	-1.45340E-02	-9.10466E-03	-4.42348E-03	-4.44089E-16
1.000	-2.75430E-02	-1.86084E-02	-1.14703E-02	-5.46656E-03	1.70974E-14

Table 7.5.4 Computed Results ($y_d = 0.63 \quad N = 8$)

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.000	7.93973E-01	2.20535E-01	6.31812E-01	-3.54607E-02
0.125	7.96331E-01	2.11951E-01	6.31808E-01	-3.65787E-02
0.250	7.94175E-01	2.18479E-01	6.31885E-01	-3.78848E-02
0.375	7.93643E-01	2.21542E-01	6.31856E-01	-3.66160E-02
0.500	7.94120E-01	2.19740E-01	6.31872E-01	-3.74561E-02
0.675	7.93137E-01	2.23399E-01	6.31851E-01	-3.60537E-02
0.750	7.95996E-01	2.13193E-01	6.31803E-01	-3.60768E-02
0.875	7.94184E-01	2.18631E-01	6.31834E-01	-3.56374E-02
1.000	7.98270E-01	2.05729E-01	6.31584E-01	-2.78000E-02

Objective Value = -0.7771561172E-15

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	7.97971E-01	8.01130E-01	8.03427E-01	8.04832E-01	8.05309E-01
0.125	8.00175E-01	8.03214E-01	8.05424E-01	8.06776E-01	8.07235E-01
0.250	7.98155E-01	8.01303E-01	8.03583E-01	8.04993E-01	8.05469E-01
0.375	7.97660E-01	8.00836E-01	8.03145E-01	8.04557E-01	8.05037E-01
0.500	7.98105E-01	8.01256E-01	8.03547E-01	8.04949E-01	8.05425E-01
0.675	7.97188E-01	8.00389E-01	8.02716E-01	8.04140E-01	8.04623E-01
0.750	7.99862E-01	8.02918E-01	8.05141E-01	8.06500E-01	8.06962E-01
0.875	7.98176E-01	8.01323E-01	8.03612E-01	8.05012E-01	8.05487E-01
1.000	8.01997E-01	8.04840E-01	8.07079E-01	8.02387E-01	8.08032E-01

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	1.79109E-01	1.36622E-01	9.28063E-02	4.73702E-02	-4.26326E-14
0.125	1.72271E-01	1.31456E-01	8.93096E-02	4.55869E-02	-6.68434E-14
0.250	1.78418E-01	1.36162E-01	9.25120E-02	4.72222E-02	-4.97380E-14
0.375	1.80000E-01	1.37353E-01	9.32580E-02	4.76223E-02	-6.68434E-14
0.500	1.78597E-01	1.36286E-01	9.25922E-02	4.72627E-02	-5.32907E-14
0.675	1.81470E-01	1.38442E-01	9.40481E-02	4.80048E-02	-4.61853E-14
0.750	1.73248E-01	1.32191E-01	8.98063E-02	4.58402E-02	-4.97380E-14
0.875	1.78416E-01	1.36128E-01	9.24805E-02	4.72052E-02	-4.26326E-14
1.000	1.66884E-01	1.27239E-01	8.64205E-02	4.41096E-02	-4.61853E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	6.31171E-01	6.30686E-01	6.30299E-01	6.30075E-01	6.30000E-01
0.125	6.31155E-01	6.30650E-01	6.30290E-01	6.30073E-01	6.30000E-01
0.250	6.31204E-01	6.30677E-01	6.30301E-01	6.30075E-01	6.30000E-01
0.375	6.31194E-01	6.30676E-01	6.30302E-01	6.30076E-01	6.30000E-01
0.500	6.31198E-01	6.30675E-01	6.30301E-01	6.30075E-01	6.30000E-01
0.875	6.31196E-01	6.30679E-01	6.30304E-01	6.30077E-01	6.30000E-01
0.750	6.31155E-01	6.30652E-01	6.30291E-01	6.30073E-01	6.30000E-01
0.875	6.31185E-01	6.30671E-01	6.30300E-01	6.30075E-01	6.30000E-01
1.000	6.31056E-01	6.30611E-01	6.30277E-01	6.30070E-01	6.30000E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	-2.88293E-02	-2.18348E-02	-1.48047E-02	-7.51464E-03	8.88178E-16
0.125	-2.89802E-02	-2.16061E-02	-1.44272E-02	-7.25455E-03	4.44089E-15
0.250	-3.02088E-02	-2.25707E-02	-1.50208E-02	-7.52590E-03	4.88498E-15
0.375	-2.95211E-02	-2.23287E-02	-1.50121E-02	-7.57212E-03	6.66134E-15
0.500	-2.98202E-02	-2.24302E-02	-1.49804E-02	-7.52520E-03	-2.88658E-15
0.675	-2.93596E-02	-2.23525E-02	-1.50904E-02	-7.62788E-03	1.11022E-15
0.750	-2.87561E-02	-2.15989E-02	-1.44737E-02	-7.29099E-03	-2.22045E-15
0.875	-2.31807E-02	-2.21653E-02	-1.49029E-02	-7.50840E-03	6.88338E-15
1.000	-2.46059E-02	-1.96468E-02	-1.36228E-02	-6.97896E-03	1.22125E-14

Table 7.5.5 Computed Results ($y_d = 0.5 \quad N = 8$)

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.000	1.34440E+00	3.73442E-01	5.03120E-01	-6.15321E-02
0.125	1.34601E+00	3.67644E-01	5.03103E-01	-6.18348E-02
0.250	1.34548E+00	3.69404E-01	5.03139E-01	-6.27095E-02
0.375	1.34469E+00	3.72282E-01	5.03142E-01	-6.24904E-02
0.500	1.34495E+00	3.71273E-01	5.03155E-01	-6.28610E-02
0.675	1.34359E+00	3.76348E-01	5.03123E-01	-6.12287E-02
0.750	1.34762E+00	3.61781E-01	5.03098E-01	-6.25948E-02
0.875	1.34260E+00	3.79737E-01	5.03172E-01	-6.13786E-02
1.000	1.35063E+00	3.52463E-01	5.02665E-01	-4.55320E-02

Objective Value = -0.6522560269E-15

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	1.35117E+00	1.35652E+00	1.36041E+00	1.36279E+00	1.36360E+00
0.125	1.35267E+00	1.35794E+00	1.36178E+00	1.36412E+00	1.36492E+00
0.250	1.35218E+00	1.35746E+00	1.36133E+00	1.36368E+00	1.36449E+00
0.375	1.35144E+00	1.35678E+00	1.36066E+00	1.36303E+00	1.362[illegible]
0.500	1.35168E+00	1.35701E+00	1.36088E+00	1.36324E+00	1.364[illegible]
0.675	1.35041E+00	1.35581E+00	1.35973E+00	1.36213E+00	1.3625[illegible]
0.750	1.35419E+00	1.35927E+00	1.36315E+00	1.36546E+00	1.36624E+00
0.875	1.34846E+00	1.35492E+00	1.35886E+00	1.36130E+00	1.36212E+00
1.000	1.35701E+00	1.36205E+00	1.36572E+00	1.36795E+00	1.36872E+00

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	3.03398E-01	2.31473E-01	1.57250E-01	8.02650E-02	-3.90798E-14
0.125	2.98750E-01	2.27940E-01	1.54861E-01	7.90465E-02	-3.19744E-14
0.250	3.00224E-01	2.29092E-01	1.55643E-01	7.94463E-02	-3.55271E-14
0.375	3.02519E-01	2.30826E-01	1.56816E-01	8.00445E-02	-2.84217E-14
0.500	3.01744E-01	2.30255E-01	1.56434E-01	7.98501E-02	-3.90799E-14
0.675	3.05724E-01	2.33236E-01	1.58445E-01	8.08749E-02	-3.90799E-14
0.750	2.94072E-01	2.24409E-01	1.52463E-01	7.78232E-02	-3.55271E-14
0.875	3.08523E-01	2.35400E-01	1.59925E-01	8.16312E-02	-3.55271E-14
1.000	2.85822E-01	2.17890E-01	1.47983E-01	7.55307E-02	-3.55271E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	5.02016E-01	5.01137E-01	5.00508E-01	5.00128E-01	5.00000E-01
0.125	5.01990E-01	5.01124E-01	5.00502E-01	5.00126E-01	5.00000E-01
0.250	5.02011E-01	5.01133E-01	5.00505E-01	5.00127E-01	5.00000E-01
0.375	5.02016E-01	5.01138E-01	5.00508E-01	5.00129E-01	5.00000E-01
0.500	5.02022E-01	5.01139E-01	5.00508E-01	5.00127E-01	5.00000E-01
0.675	5.02017E-01	5.01144E-01	5.00512E-01	5.00129E-01	5.00000E-01
0.750	5.01975E-01	5.01111E-01	5.00495E-01	5.00124E-01	5.00000E-01
0.875	5.02050E-01	5.01161E-01	5.00518E-01	5.00130E-01	5.00000E-01
1.000	5.01790E-01	5.01041E-01	5.00473E-01	5.00120E-01	5.00000E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.000	-4.95856E-02	-3.75286E-02	-2.52849E-02	-1.27572E-02	2.66454E-15
0.125	-4.94708E-02	-3.72230E-02	-2.49574E-02	-1.25728E-02	6.43929E-15
0.250	-5.01399E-02	-3.76384E-02	-2.51616E-02	-1.26471E-02	8.88178E-15
0.375	-5.00850E-02	-3.77092E-02	-2.52823E-02	-1.27332E-02	2.44249E-15
0.500	-5.03984E-02	-3.78563E-02	-2.53033E-02	-1.27135E-02	2.22045E-15
0.675	-4.95985E-02	-3.76799E-02	-2.54227E-02	-1.28503E-02	2.98650E-15
0.750	-4.95876E-02	-3.69949E-02	-2.46683E-02	-1.23900E-02	6.88338E-15
0.875	-5.04747E-02	-3.83807E-02	-2.57958E-02	-1.29896E-02	3.10862E-15
1.000	-4.15810E-02	-3.32777E-02	-2.32270E-02	-1.19402E-02	7.10543E-15

Table 7.5.6 Computed Results ($y_d = 0.8 \quad N = 4$)

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.00	-7.77659E-30	2.81749E-01	8.02379E-01	-4.72896E-02
0.25	-1.25217E-27	2.81749E-01	8.02379E-01	-4.72896E-02
0.50	1.83494E-27	2.81749E-01	8.02379E-01	-4.72896E-02
0.75	-6.48288E-32	2.81749E-01	8.02379E-01	-4.72896E-02
1.00	2.20369E-28	2.81749E-01	8.02379E-01	-4.72896E-02

Objective Value = -0.9714451465E-16

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	5.10924E-03	9.14841E-03	1.20854E-02	1.38819E-02	1.44921E-02
0.25	5.10924E-03	9.14841E-03	1.20854E-02	1.38819E-02	1.44921E-02
0.50	5.10924E-03	9.14841E-03	1.20854E-02	1.38819E-02	1.44921E-02
0.75	5.10924E-03	9.14841E-03	1.20854E-02	1.38819E-02	1.44921E-02
1.00	5.10924E-03	9.14841E-03	1.20854E-02	1.38819E-02	1.44921E-02

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	2.28954E-01	1.74697E-01	1.18684E-01	6.05807E-02	-2.48690E-14
0.25	2.28954E-01	1.74697E-01	1.18684E-01	6.05807E-02	-2.48690E-14
0.50	2.28954E-01	1.74697E-01	1.18684E-01	6.05807E-02	-2.48690E-14
0.75	2.28954E-01	1.74697E-01	1.18684E-01	6.05807E-02	-1.77636E-14
1.00	2.28954E-01	1.74697E-01	1.18684E-01	6.05807E-02	-1.77636E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	8.01527E-01	8.00862E-01	8.00385E-01	8.00097E-01	8.00000E-01
0.25	8.01527E-01	8.00862E-01	8.00385E-01	8.00097E-01	8.00000E-01
0.50	8.01527E-01	8.00862E-01	8.00385E-01	8.00097E-01	8.00000E-01
0.75	8.01527E-01	8.00862E-01	8.00385E-01	8.00097E-01	8.00000E-01
1.00	8.01527E-01	8.00862E-01	.00385E-01	8.00097E-01	8.00000E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	-3.79319E-02	-2.85614E-02	-1.91435E-02	-9.63829E-03	-8.88178E-16
0.25	-3.79319E-02	-2.85614E-02	-1.91435E-02	-9.63829E-03	4.21885E-15
0.50	-3.79319E-02	-2.85614E-02	-1.91435E-02	-9.63829E-03	0.00000E+00
0.75	-3.79319E-02	-2.85614E-02	-1.91435E-02	-9.63829E-03	3.10862E-15
1.00	-3.79319E-02	-2.85614E-02	-1.91435E-02	-9.63829E-03	5.55112E-15

Table 7.5.7 Computed Results ($y_d = 0.7$ $N = 4$ $x(0,t) \le 0.3$)

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.00	3.00000E-01	8.34370E-01	7.07045E-01	-1.40043E-01
0.25	3.00000E-01	8.34370E-01	7.07045E-01	-1.40043E-01
0.50	3.00000E-01	8.34370E-01	7.07045E-01	-1.40043E-01
0.75	3.00000E-01	8.34370E-01	7.07045E-01	-1.40043E-01
1.00	3.00000E-01	8.34370E-01	7.07045E-01	-1.40043E-01

Objective Value = -0.416333033E-16

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	3.15130E-01	3.27092E-01	3.35790E-01	3.41110E-01	3.42917E-01
0.25	3.15130E-01	3.27092E-01	3.35790E-01	3.41110E-01	3.42917E-01
0.50	3.15130E-01	3.27092E-01	3.35790E-01	3.41110E-01	3.42917E-01
0.75	3.15130E-01	3.27092E-01	3.35790E-01	3.41110E-01	3.42917E-01
1.00	3.15130E-01	3.27092E-01	3.35790E-01	3.41110E-01	3.42917E-01

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	6.78024E-01	5.17347E-01	3.51472E-01	1.79404E-01	-3.55271E-14
0.25	6.78024E-01	5.17347E-01	3.51472E-01	1.79404E-01	-3.55271E-14
0.50	6.78024E-01	5.17347E-01	3.51472E-01	1.79404E-01	-2.48690E-14
0.75	6.78024E-01	5.17347E-01	3.51472E-01	1.79404E-01	-4.26326E-14
1.00	6.78024E-01	5.17347E-01	3.51472E-01	1.79404E-01	-4.26326E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	7.04521E-01	7.02552E-01	7.01139E-01	7.00286E-01	7.00000E-01
0.25	7.04521E-01	7.02552E-01	7.01139E-01	7.00286E-01	7.00000E-01
0.50	7.04521E-01	7.02552E-01	7.01139E-01	7.00286E-01	7.00000E-01
0.75	7.04521E-01	7.02552E-01	7.01139E-01	7.00286E-01	7.00000E-01
1.00	7.04521E-01	7.02552E-01	7.01139E-01	7.00286E-01	

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	-1.12332E-01	-8.45818E-02	-5.66916E-02	-2.85428E-02	-4.44089E-16
0.25	-1.12332E-01	-8.45818E-02	-5.66916E-02	-2.85428E-02	5.32907E-15
0.50	-1.12332E-01	-8.45818E-02	-5.66916E-02	-2.85428E-02	2.88658E-15
0.75	-1.12332E-01	-8.45818E-02	-5.66916E-02	-2.85428E-02	2.88658E-15
1.00	-1.12332E-01	-8.45818E-02	-5.66916E-02	-2.85428E-02	6.43929E-15

Table 7.5.8 Computed Results ($y_d = 0.7$ $N = 4$ $x(0,t) \geq 0.6$)

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.00	6.00000E-01	-2.23000E-01	6.98117E-01	3.74291E-02
0.25	6.00000E-01	-2.23000E-01	6.98117E-01	3.74291E-02
0.50	6.00000E-01	-2.23000E-01	6.98117E-01	3.74291E-02
0.75	6.00000E-01	-2.23000E-01	6.98117E-01	3.74291E-02
1.00	6.00000E-01	-2.23000E-01	6.98117E-01	3.74291E-02

Objective Value = 0.124900090E-15

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	5.95956E-01	5.92759E-01	5.90435E-01	5.89013E-01	5.88530E-01
0.25	5.95956E-01	5.92759E-01	5.90435E-01	5.89013E-01	5.88530E-01
0.50	5.95956E-01	5.92759E-01	5.90435E-01	5.89013E-01	5.88530E-01
0.75	5.95956E-01	5.92759E-01	5.90435E-01	5.89013E-01	5.88530E-01
1.00	5.95956E-01	5.92759E-01	5.90435E-01	5.89013E-01	5.88530E-01

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	-1.81214E-01	-1.38270E-01	-9.39371E-02	-4.79488E-02	-3.19744E-14
0.25	-1.81214E-01	-1.38270E-01	-9.39371E-02	-4.79488E-02	-3.55271E-14
0.50	-1.81214E-01	-1.38270E-01	-9.39371E-02	-4.79488E-02	-3.19744E-14
0.75	-1.81214E-01	-1.38270E-01	-9.39371E-02	-4.79488E-02	-3.19744E-14
1.00	-1.81214E-01	-1.38270E-01	-9.39371E-02	-4.79488E-02	-3.55271E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	6.98792E-01	6.99318E-01	6.99696E-01	6.99924E-01	7.00000E-01
0.25	6.98792E-01	6.99318E-01	6.99696E-01	6.99924E-01	7.00000E-01
0.50	6.98792E-01	6.99318E-01	6.99696E-01	6.99924E-01	7.00000E-01
0.75	6.98792E-01	6.99318E-01	6.99696E-01	6.99924E-01	7.00000E-01
1.00	6.98792E-01	6.99318E-01	6.99696E-01	6.99924E-01	7.00000E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	3.00226E-02	2.26060E-02	1.51518E-02	7.62857E-03	0.00000E+00
0.25	3.00226E-02	2.26060E-02	1.51518E-02	7.62857E-03	1.11022E-15
0.50	3.00226E-02	2.26060E-02	1.51518E-02	7.62857E-03	2.22045E-15
0.75	3.00226E-02	2.26060E-02	1.51518E-02	7.62857E-03	1.99840E-15
1.00	3.00226E-02	2.26060E-02	1.51518E-02	7.62857E-03	3.33067E-15

Table 7.5.9 Computed Results ($y_d = 0.7$ N = 4 $y(0,t) \le 0.6$)

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.00	3.89690E+00	-1.18432E+01	6.00000E-01	1.98780E+00
0.25	3.89690E+00	-1.18432E+01	6.00000E-01	1.98780E+00
0.50	3.89690E+00	-1.18432E+01	6.00000E-01	1.98780E+00
0.75	3.89690E+00	-1.18432E+01	6.00000E-01	1.98780E+00
1.00	3.89690E+00	-1.18432E+01	6.00000E-01	1.98780E+00

Objective Value = -0.360822483E-15

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	3.68214E+00	3.51236E+00	3.38890E+00	3.31339E+00	3.28773E+00
0.25	3.68214E+00	3.51236E+00	3.38890E+00	3.31339E+00	3.28773E+00
0.50	3.68214E+00	3.51236E+00	3.38890E+00	3.31339E+00	3.28773E+00
0.75	3.68214E+00	3.51236E+00	3.38890E+00	3.31339E+00	3.28773E+00
1.00	3.68214E+00	3.51236E+00	3.38890E+00	3.31339E+00	3.28773E+00

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	-9.62396E+00	-7.34329E+00	-4.98884E+00	-2.54646E+00	-2.84217E-14
0.25	-9.62396E+00	-7.34329E+00	-4.98884E+00	-2.54646E+00	-3.19744E-14
0.50	-9.62396E+00	-7.34329E+00	-4.98884E+00	-2.54646E+00	-2.84217E-14
0.75	-9.62396E+00	-7.34329E+00	-4.98884E+00	-2.54646E+00	-3.19744E-14
1.00	-9.62396E+00	-7.34329E+00	-4.98884E+00	-2.54646E+00	-2.48690E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	6.35822E-01	6.63774E-01	6.83832E-01	6.95938E-01	7.00000E-01
0.25	6.35822E-01	6.63774E-01	6.83832E-01	6.95938E-01	7.00000E-01
0.50	6.35822E-01	6.63774E-01	6.83832E-01	6.95938E-01	7.00000E-01
0.75	6.35822E-01	6.63774E-01	6.83832E-01	6.95938E-01	7.00000E-01
1.00	6.35822E-01	6.63774E-01	6.83832E-01	6.95938E-01	7.00000E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	1.59445E+00	1.20056E+00	8.04688E-01	4.05141E-01	4.44089E-16
0.25	1.59445E+00	1.20056E+00	8.04688E-01	4.05141E-01	3.99680E-15
0.50	1.59445E+00	1.20056E+00	8.04688E-01	4.05141E-01	4.44089E-15
0.75	1.59445E+00	1.20056E+00	8.04688E-01	4.05141E-01	0.00000E+00
1.00	1.59445E+00	1.20056E+00	8.04688E-01	4.05141E-01	5.10703E-15

Table 7.5.10 Computed Results ($y_d = 0.7$ N = 4 $x(0,t)$ bounded)

0.0 ≤ x(0,0.00) ≤ 0.2
0.2 ≤ x(0,0.25) ≤ 0.4
0.4 ≤ x(0,0.50) ≤ 0.6
0.6 ≤ x(0,0.75) ≤ 0.8
0.8 ≤ x(0,1.00) ≤ 1.0

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.00	2.00000E-01	1.19499E+00	7.08000E-01	-1.21655E-01
0.25	4.00000E-01	4.86679E-01	7.02901E-01	-3.69198E-02
0.50	4.86714E-01	1.79123E-01	7.00765E-01	-2.33895E-04
0.75	6.00000E-01	-2.16754E-01	6.96519E-01	1.02584E-01
1.00	8.00000E-01	-9.19738E-01	6.80102E-01	2.37741E-01

Objective Value = -0.416822463E-16

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	2.21020E-01	2.38660E-01	2.51033E-01	2.58597E-01	2.61167E-01
0.25	4.08797E-01	4.15720E-01	4.20745E-01	4.23816E-01	4.24859E-01
0.50	4.89945E-01	4.92481E-01	4.94318E-01	4.95441E-01	4.95823E-01
0.75	5.96030E-01	5.92352E-01	5.90526E-01	5.89104E-01	5.88620E-01
1.00	7.83271E-01	7.68893E-01	7.60321E-01	7.54402E-01	7.52391E-0

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	9.68933E-01	7.36272E-01	4.99818E-01	2.55079E-01	-2.84217E-14
0.25	3.93084E-01	2.99039E-01	2.02931E-01	1.03856E-01	-2.84217E-14
0.50	1.44095E-01	1.09429E-01	7.42142E-02	3.78666E-02	-2.48690E-14
0.75	-1.79369E-01	-1.37999E-01	-9.40330E-02	-4.80301E-02	-3.19744E-14
1.00	-7.51586E-01	-5.74980E-01	-3.91001E-01	-1.99624E-01	-2.84217E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	7.05578E-01	7.03353E-01	7.01564E-01	7.00403E-01	7.00000E-01
0.25	7.02112E-01	7.01310E-01	7.00624E-01	7.00163E-01	7.00000E-01
0.50	7.00661E-01	7.00446E-01	7.00222E-01	7.00059E-01	7.00000E-01
0.75	6.98146E-01	6.98115E-01	6.99655E-01	6.99921E-01	7.00000E-01
1.00	6.94122E-01	6.96891E-01	6.98679E-01	6.99672E-01	7.00000E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	-1.18310E-01	-1.02177E-01	-7.51012E-02	-3.98807E-02	2.22045E-16
0.25	-4.09032E-02	-3.61972E-02	-2.54501E-02	-1.60549E-02	3.33067E-15
0.50	-6.97058E-03	-1.16846E-02	-1.01393E-02	-5.79507E-03	4.88498E-15
0.75	6.26578E-02	3.51646E-02	1.91448E-02	8.13652E-03	-6.66134E-16
1.00	1.67152E-01	1.11987E-01	6.84062E-02	3.24304E-02	1.99840E-15

Table 7.5.11 Computed Results ($y_d = 0.7$ N = 4 $y(0,t)$ bounded)

0.0 ≤ y(0,0.00) ≤ 0.2
0.2 ≤ y(0,0.25) ≤ 0.4
0.4 ≤ y(0,0.50) ≤ 0.6
0.6 ≤ y(0,0.75) ≤ 0.8
0.8 ≤ y(0,1.00) ≤ 1.0

The Optimal Controls

t	x(0,t)	x'(0,t)	y(0,t)	y'(0,t)
0.00	2.44365E+01	-8.42607E+01	0.00000E+00	1.33419E+01
0.25	2.10581E+01	-7.27655E+01	2.00000E-01	7.71500E+00
0.50	6.43685E+00	-2.11110E+01	6.00000E-01	6.89152E-01
0.75	4.18470E+00	-1.28708E+01	6.00000E-01	1.78182E+00
1.00	1.27349E-27	1.50057E+00	8.00000E-01	-3.49600E+00

Objective Value = 0.825745925E-06

Solutions for x(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	2.29088E+01	2.17012E+01	2.08231E+01	2.02860E+01	2.01035E+01
0.25	1.97412E+01	1.87031E+01	1.79181E+01	1.74881E+01	1.73315E+01
0.50	6.05587E+00	5.78667E+00	5.53374E+00	5.40717E+00	5.36714E+00
0.75	3.95155E+00	3.76751E+00	3.63378E+00	3.55200E+00	3.52422E+00
1.00	2.91736E-02	5.43546E-02	7.33933E-02	8.52358E-02	8.92780E-02

Solutions for x'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	-6.84525E+01	-5.22288E+01	-3.54840E+01	-1.81127E+01	-2.84217E-14
0.25	-5.89054E+01	-4.48652E+01	-3.04598E+01	-1.55450E+01	-3.90799E-14
0.50	-1.70006E+01	-1.29149E+01	-8.7600[illegible]+00	-4.47007E+00	-2.13163E-14
0.75	-1.04379E+01	-7.95621E+00	-5.40299E+00	-2.75756E+00	-3.19744E-14
1.00	1.35322E+00	1.11780E+00	7.78208E-01	4.00816E-01	-2.84217E-14

Solutions for y(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	2.45379E-01	4.42084E-01	5.84779E-01	6.71184E-01	7.00185E-01
0.25	3.53515E-01	4.93249E-01	6.04078E-01	6.75140E-01	6.99378E-01
0.50	6.19950E-01	6.47311E-01	6.74246E-01	6.93716E-01	7.00816E-01
0.75	6.33148E-01	6.60487E-01	6.80993E-01	6.93776E-01	6.98141E-01
1.00	7.48091E-01	7.21519E-01	7.09202E-01	7.04236E-01	7.02981E-01

Solutions for y'(z,t)

t	z=0.02	z=0.04	z=0.06	z=0.08	z=0.10
0.00	1.11127E+01	8.51356E+00	5.73991E+00	2.89241E+00	1.35447E-14
0.25	7.47520E+00	6.37211E+00	4.62030E+00	2.42936E+00	1.55431E-15
0.50	1.24721E+00	1.42203E+00	1.21173E+00	6.33701E-01	2.22045E-16
0.75	1.52160E+00	1.20402E+00	8.39058E-01	4.33335E-01	2.44249E-15
1.00	-1.84553E+00	-8.97879E-01	-3.89863E-01	-1.36748E-01	3.33067E-15

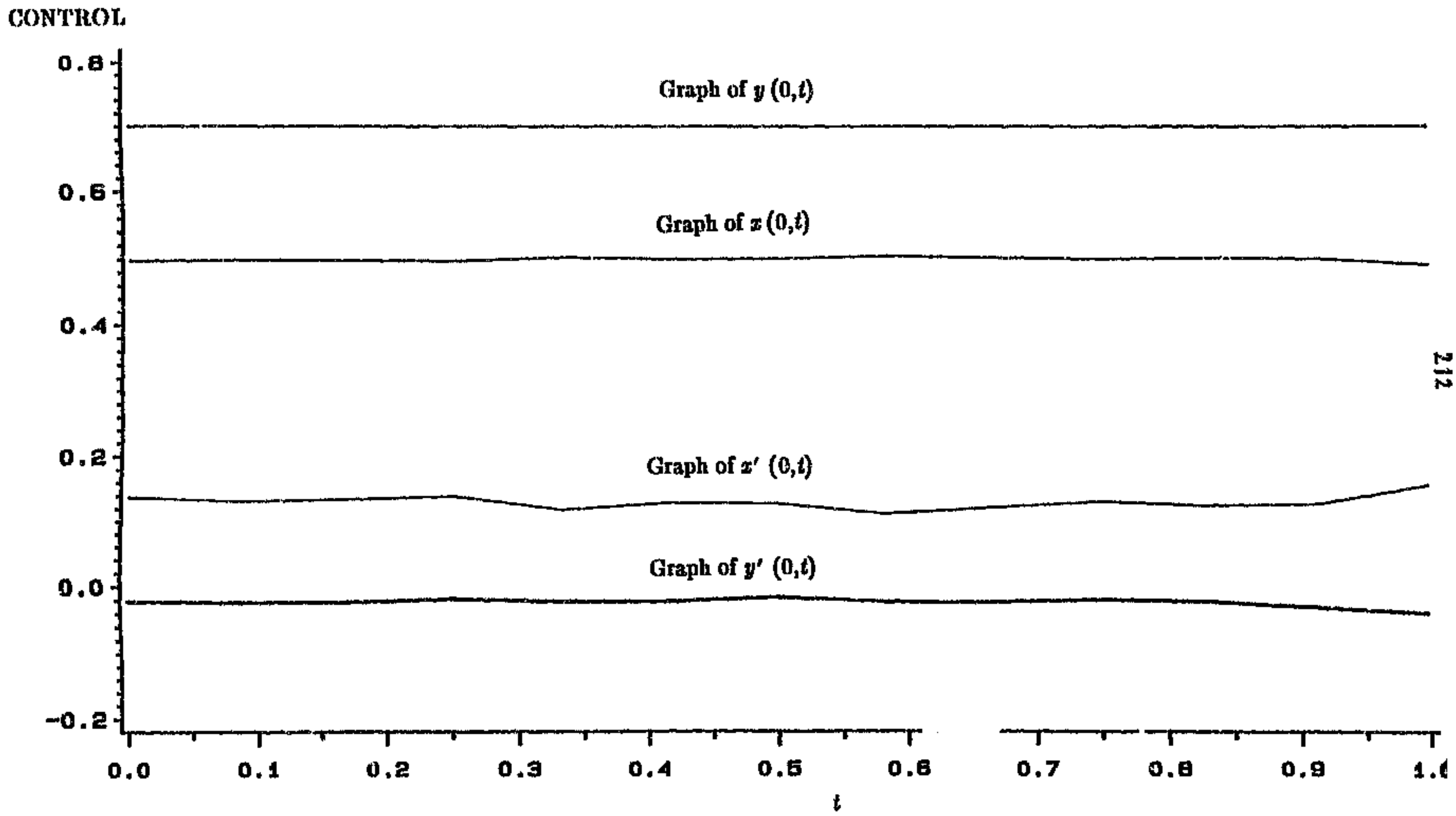

Figure 7.5.1 The Optimal Controls ($y_d = 0.7$ $N = 12$)

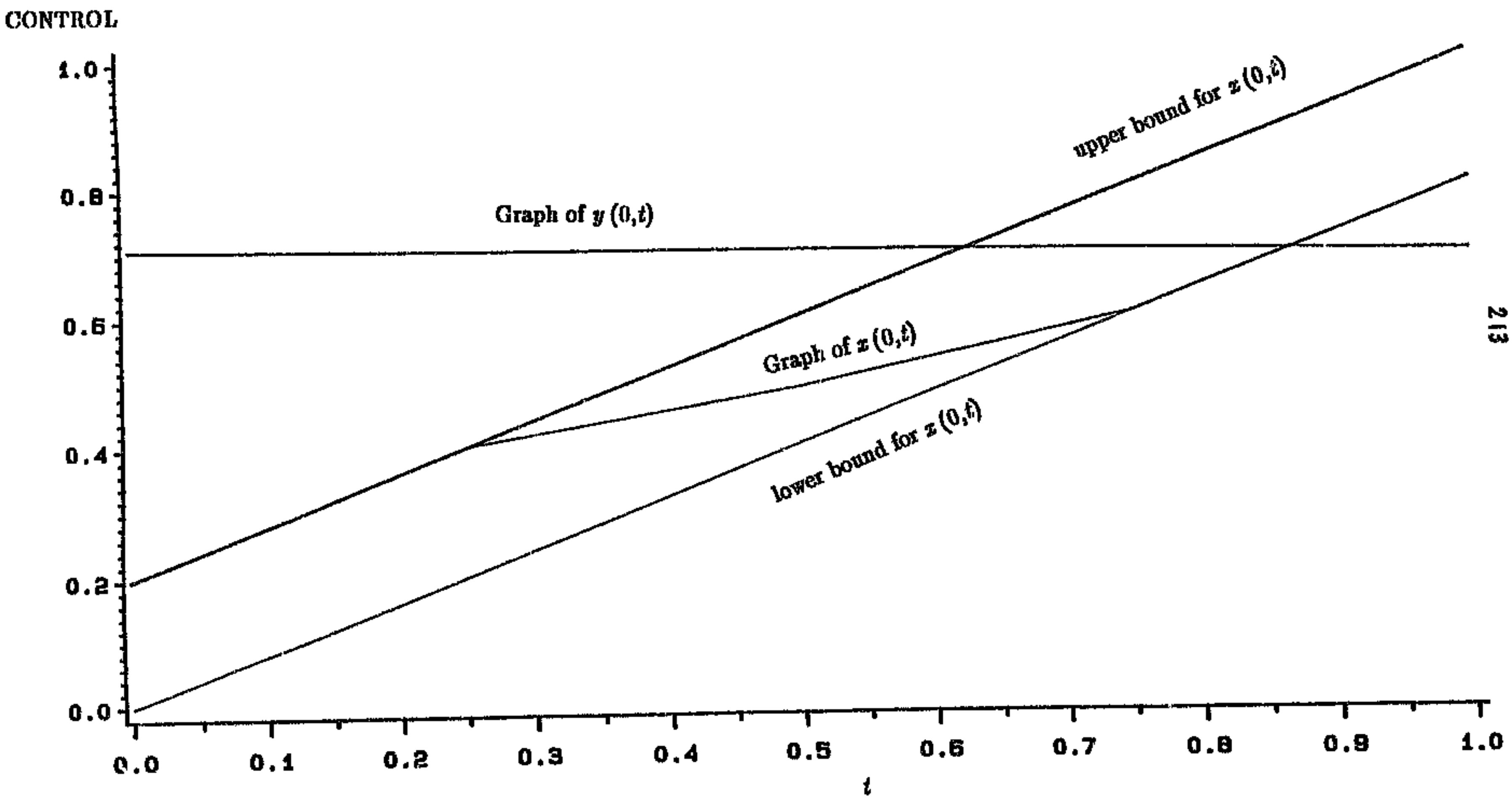

Figure 7.5.2 The Optimal Controls ($y_d = 0.7$ N = 4 $z(0,t)$ bounded)

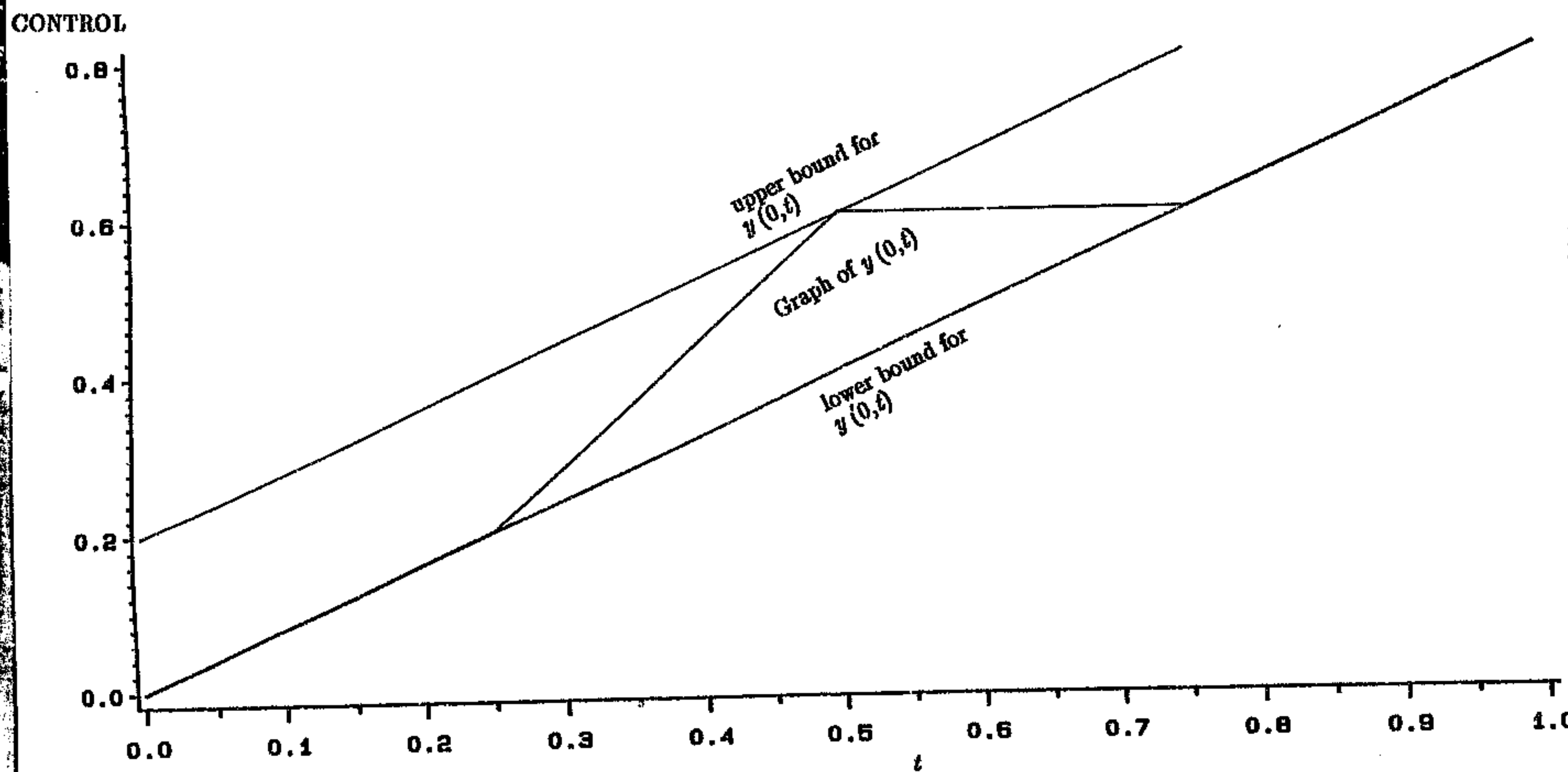

Figure 7.5.3 The Optimal Control ($y_d = 0.7$ $N = 4$ $y(0,t)$ bounded)

CHAPTER VIII

CONCLUSION AND SUGGESTIONS FOR FURTHER STUDY

The main concern of this thesis has been to devise efficient algorithms for solving the four types of optimal control problems mentioned in Chapter I. A few numerical problems were included in each chapter to demonstrate the efficiency of the algorithms used. For the case of deterministic control problems governed either by ordinary differential equations or ordinary differential equations with time-delayed arguments, convergence properties have also been established for the algorithms used in Chapters II, IV and V. In the near future, we would like to establish convergence results for the method used in Chapter III also, which involved control parameterization technique in conjunction with Liapunov concept.

For the sequential gradient-restoration algorithm discussed in Chapter II, the rate of convergence is usually very fast during the first few iterations, but becomes much slower in the vicinity of the optimal solution. Thus, it appears that a method for determining any norm between the sequence of controls generated by the algorithm and the true optimal control is very desirable.

Also, it will be a very interesting problem to extend the convergence results obtained in Chapter II to a more general class of nonlinear problems, which involve not only the initial constraints and terminal constraints, but also continuous constraints. The proof of convergence result of such a class of optimal control problems will be much harder than that given for the problem discussed in Chapter II, owing to the existence of continuous constraints.

For the method discussed in Chapter III, it can be very time–consuming when the total number of parameters in the approximate problem is large.

Another disadvantage of the algorithm is that there is no clear–cut w to choose the weighting factors in the approximate problem. At this moment, we nee determine these factors by trial and error. However, in spite of these drawbacks, we still believe that it is a very good method, since very accurate results have been obtained f the two numerical examples that we solved in Chapter III. As mentioned in the first paragraph of this section, some proof of convergence result for the algorithm used would be very desirable.

In Chapter IV, a computational algorithm similar to that developed in Ref. 103 was used for solving a class of time–delayed optimal control problems which involved a nonlinear functional, terminal state inequality and equality constraints, together with continuous state constraints. This method involved both the technique of control parameterization and constraint transcription.

At the present moment, the constraint transcription does not allow the con ous state constraints to be of their full generality, because we need to exclude the control variable from these continuous state constraints. It will be very useful if we can extend the application of the method developed in Ref. 103 to the case where the continuous constraints can have the following form

$$g_i\,(t,\, x\,(t),\, x\,(t-h),\, u\,(t),\, u\,(t-h)) \geq 0,\ i = 1,\ldots,N,\quad \forall\, t \in [0, T] \tag{8.1}$$

As mentioned in the Introduction Section in Chapter I, the problem considered in Chapter V differed from that considered in Chapter IV only in the way that the cost functional of the problem in Chapter V consisted of not only terminal cost and integral cost, but also the full variation of contro' Since this problem was actually formed from the problem considered in Ref. 104 b addition of time–delayed arguments in both the state variables and the control variables, all the convergence results of Ref. 104 were extended to the problem described in Chapter V in a straightforward manner. However, the author does not know if the method developed in Ref. 104 can also be applied to the class of optimal control problem, where the cost functional involve not only the full variation of control, but also the full variation of the state variable. This is an open problem.

In Chapter VI, we considered a stochastic linear quadratic problem with Markov jump processes in the parameters. By seeking for the best feedback control law depending only on the measurable output, we converted the original stochastic optimal control problem into a standard constrained deterministic optimization problem. A very challenging job for future research will be to use this idea to solve the following time–delayed stochastic control problem

$$\min E\{\tfrac{1}{2}\int_0^\infty [x(\tau)^T Q\, x(\tau) + u(\tau)^T R\, u(\tau)$$

$$+ x(\tau-h)^T \overline{Q}\, x(\tau-h) + u(\tau-h)^T \overline{R}\, u(\tau-h)]\, d\tau\} \tag{8.2}$$

subject to

$$\dot{x}(t) = A(r(t))\, x(t) + B(r(t))\, u(t)$$

$$+ \overline{A}(r(t-h))\, x(t-h) + \overline{B}(r(t-h))\, x(t-h) \tag{8.3a}$$

$$y(t) = C(r(t))\, x(t) \tag{8.3b}$$

$$x(t) = \phi(t), \qquad t \in [-h,0] \tag{8.3c}$$

$$u(t) = \beta(t), \qquad t \in [-h,0] \tag{8.3d}$$

and

$$\text{Prob}\{r(t+\Delta) = j \mid r(t) = i\}$$

$$= \begin{cases} \pi_{ij}\,\Delta + 0\,(\Delta) & \text{if } i \neq j \quad \forall\, t \in [-h,\infty) \\ 1 + \pi_{ij}\,\Delta + 0\,(\Delta) & \text{if } i = j\,. \end{cases} \tag{8.4}$$

The chemical reactor problem considered in Chapter VII involved a couple of nonlinear diffusion equations. The available control variables were the input temperature and the input oxygen concentration, which were both functions of time. By linearizing the differential equations around a nominal solution and then applying a finite element Galerkin Scheme to the resulting distributed system, the original problem were converted to a sequence of quadratic programming problems with linear constraints.

From the numerical result, it appeared that the above method worked very efficiently for the problem defined in Chapter VII. However, if we modify the above problem a little bit by taking the velocity parameter $f(t)$ of the chemical reactor to be another control variable, then the above method will not work for this modified problem any more, since we shall not be able to convert the original problem into a sequence of approximate problems. A more sophisticated algorithm needs to be developed for the modified problem.

Finally, for the problem defined in Chapter VII, we have not yet established any convergence properties of the sequence of optimal control of the approximate problem to the true optimal control. This question is yet to be answered, also.

REFERENCES

1. Abadie, J., and Carpentier, J., *Generalization of the Wolfe Reduced Gradients for the Case of Nonlinear Constraints*, Optimization, edited by R. Fletcher, Academic Press, New York, New York, 1969.

2. Abadie, J., and Guidou, J., *Numerical Experiments with the GRG Method,* Integer and Nonlinear Programming, edited by J. Abadie, North–Holland, Amsterdam, 1970.

3. Ahmed, N. U., *Optimal Control for Hereditary Dynamical Systems,* Technical Report, University of Ottawa, No. 70–8, 1970.

4. Ahmed, N. U., *Optimal Control of a Class of Hereditary Systems with Lagging Control,* International Journal of Systems Science, Vol. 5, pp. 278–293, 1974.

5. Ahmed, N. U., *Elements of Finite–Dimensional Systems and Control Theory,* Longman Scientific and Technical, Essex, England, 1988.

6. Angell, T. S., *Existence of Optimal Control without Convexity and a Bang–Bang Theorem for Linear Volterra Equations,* Journal of Optimization Theory and Applications, Vol. 19, pp. 63–79, 1976.

7. Angell, T. S., *Existence Theorems for Optimal Control Problems Involving Functional Differential Equations,* Journal of Optimization Theory and Applications, Vol. 7, pp. 149–169, 1971.

8. Angell, T. S., *On the Optimal Control of Systems Governed by Nonlinear Volterra Equations*, Journal of Optimization Theory and Applications, Vol. 19, pp. 29–45, 1976.

9. Athans, M., *The Matrix Maximum Principle*, Information and Control, Vol. 11, No. 5 and 6, pp. 592–606, 1967.

10. Athans, M., and Levine, W. S., *On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems*, IEEE Transactions on Automatic Control, Vol. AC–15, No. 1, pp. 44–48, 1970.

11. Barnes, E. R., *An Extension of Gilbert's Algorithm for Computing Optimal Controls*, Journal of Optimization Theory and Applications, Vol. 7, pp. 420–443, 1971.

12. Barnett, S., *Matrices in Control Theory*, London : Van Nostrand Reinhold, 1971.

13. Barnett, S., and Cameron, R. G., *Introduction to Mathematical Control Theory*, Second Edition, Oxford : Clarendon Press, 1985.

14. Blatt, J. M., *Optimal Control with a Cost of Switching Control*, Journal of Australian Mathematical Society, Series B, pp. 316–332, 1976.

15. Bosarge, W. E. Jr., and Johnson, O. G., *Direct Method Approximation to the State Regulator Control Problem Using a Ritz–Trefftz Suboptimal Control*, IEEE Transactions on Automatic Control, Vol. AC–15, pp. 627–631, 1970.

16. Bryson, A. E., and Denham, W. F., *A Steepest–Ascent Method for Solving Optimum Programming Problems*, Journal of Applied Mechanics, Vol. 84, 1962.

17. Bryson, A. E., and Ho, Y. C., *Applied Optimal Control – Optimization*, Estimation and Control, John–Wiley & Sons, 1975.

18. Choksy, N. H., *Time Lag Systems*, A Bibliography, I.R.E. Transaction on Automatic Control, Vol. AC–5, pp.66–70, 1960.

19. Chyung, D. H., and Lee, E. B., *Delayed Action Control Problems*, IFAC. IV, Warsaw, pp. 3–15, 1969 or Automatica, Vol. 6, pp. 395–400, 1970.

20. Craven, B. D., *Mathematical Programming and Control Theory*, Chapman and Hall, London, 1978.

21. Fiacco, A. V., and McCormick, G. P., *Nonlinear Programming : Sequential Unconstrained Minimization Technique*, John Wiley, New York, New York, 1968.

22. Florent[illegible] J. J., *Optimal Control of Continuous [illegible]ime, Markov, Stochastic Systems*, Journal of Electronics and Control, Vol. 10, No. 6, pp. 473–488, 1961.

23. Gabay, D., and Luenberger, D. G., *Efficiently Converging Minimization Methods Based on the Reduced Gradients*, SIAM Journal on Control and Optimization, Vol. 14, pp. 42–61, 1976.

24. Georganas, N. D., and Ahmed, N. U., *Optimal Initial Data and Control for a Class of Linear Hereditary Systems*, Information and Control, Vol. 22, pp. 394–402, 1973.

25. Goh, C. J., and Teo, K. L., *A Computational Method for Stochastic Optimal Parameter Selection Problems with Probability Constraints,* Bulletin of the Australian Mathematical Society, Vol. 41, pp. 393–405, 1990.

26. Goh, C. J., and Teo, K. L., *MISER : A FORTRAN Program for Solving Optimal Control Problems,* Advances in Engineering Software, Vol. 10, pp. 90–99, 1988.

27. Goh, C. J., and Teo, K. L., *MISER : An Optimal Control Software,* Theory and User Manual, Applied Research Corporation, Singapore, 1987.

28. Goh, C. J., and Teo, K. L., *Control Parameterization : A Unified Approach to Optimal Control Problems with General Constraints,* Automatica, Vol. 24, pp. 3–18, 1988.

29. Goldstein, A. A., *Convex Programming in Hilbert Space,* Bulletin of American Mathematical Society, Vol. 70, pp. 709–710, 1964.

30. Golub, G. H., and Van Loan, C. F., *Matrix Computations,* Oxford : Norch Oxford Academic, 1983.

31. Gonzalez, S., and Miele, A., *Sequential Gradient–Restoration Algorithm for Optimal Control Problems with General Boundary Conditions,* Journal of Optimization Theory and Applications, Vol. 26, pp. 395–425, 1978.

32. Hicks, G. A., and Ray, W. H., *Approximation Methods for Optimal Control Synthesis,* Canadian Journal of Chemical Engineering, Vol. 49, pp. 522–528, 1971.

33. Jennings, L. S., Fisher, M. E., Teo, K. L., and Goh, C. J., *MISER3, Version 1.0 : An Optimal Control Software,* Theory and User Manual, EMCOSS Ltd, Perth, 1990.

34. Jennings, L. S., and Teo, K. L., *Computational Algorithm for Functional Inequality Constrained Optimization Problems,* Automatica, Vol. 26, pp. 371–376, 1990.

35. Kelley, H. J., *Gradient Theory of Optimal Flight Paths,* ARS Journal, Vol. 30, 1960.

36. Kharatishvili, G. L., *A Maximum Principle in External Problems with Delays,* Mathematical Theory of Control, Academic Press, New York, pp. 26–34, 1967.

37. Krasovskii, N. N., and Lidskii, E. A., *Analytic Design of Controllers in Systems with Random Attributes,* Automatic Remote Control, Vol. 22, pp. 1021–1025, pp. 1141–1146, pp. 1289–1294, 1961.

38. Lee, E. B., *Optimal Control of Systems with Time Delays,* Joint Automatic Control Conference, pp. 451–452, 1969.

39. Lee, E. B., *Variational Problems for Systems Having Delays in the Control Action,* IEEE Transactions on Automatic Control, Vol. AC–13, pp. 697–699, 1969.

40. Levine, W. S., and Athans, M., *On the Determination of the Optimal Constant Output Feedback Gains for Linear Multivariable Systems,* IEEE Transactions on Automatic Control, Vol. AC–15, No. 1, pp. 44–48, 1970.

41. Levitin, E. S., and Polyak, B. T., *Constrained Minimization Methods,* USSR Computational Mathematics and Mathematical Physics, Vol. 6, pp. 1–50, 1966.

42. Lim, C. C., and Teo, K. L., *A Stochastic Optimal Control Approach to a Mathematical Drug Administration Model,* Mathematical and Computer Modelling, Vol. 12, pp. 1009–1015, 1989.

43. Long, R. S., *Newton–Raphson Operator ; Problems with Undetermined End Points,* AIAA Journal, Vol. 3, 1965.

44. Lootsma, F. A., *Boundary Properties of Penalty Functions for Constrained Minimization,* Ph.D. Thesis, Technische Hogeschool Eindhoven, The Netherlands, 1970.

45. Luenberger, D. G., *Introduction to Linear and Nonlinear Programming,* Addison–Wesley, Reading, MA, 1973.

46. Luenberger, D. G., *The Gradient Projection Method along Geodesics,* Management Science, Vol. 18, pp. 620–631, 1972.

47. Mariton, M., *On the Influence of Noise on Jump Linear Systems,* IEEE Transactions on Automatic Control, Vol. AC–32, No. 12, pp. 1094–1097, 1987.

48. Mariton, M., and Bertrand, P., *Output Feedback for a Class of Linear System with Stochastic Jump Parameters,* IEEE Transactions on Automatic Control, Vol. AC–30, No. 9, pp. 898–900, 1985.

49. Martin, R. B., and Teo, K. L., *A Worst–Case Optimal Parameter Selection Model of Cancer Chemotherapy,* IEEE Transactions on Biomedical Engineering, 1991.

50. Matula, J., *On an Extremum Problem,* Journal of Australian Mathematical Society, Series B, pp. 376–392, 1987.

51. Mayne, D. Q., and Polak, E., *An Exact Penalty Function Algorithm for Optimal Control Problems with Control and Terminal Equality Constraints,* Parts 1 and 2, Journal of Optimization Theory and Applications, Vol. 32, pp. 211–246, 345–364, 1980.

52. McCormick, G. P., and Tapia, R. A., *The Gradient Projection Method under Mild Differentiability Conditions,* SIAM Journal on Control, Vol. 10, pp. 93–98, 1972.

53. McLane, P. J., *Optimal Stochastic Control of Linear Systems with State and Control Dependent Disturbances,* IEEE Transactions on Automatic Control, Vol. AC–16, No. 6, pp. 793–798, 1971.

54. Miele, A., *Gradient Algorithms for the Optimization of Dynamic Systems,* in C. T. Leondes (ed.), Control and Dynamic Systems : Advances in Theory and Applications, Academic Press, New York, Vol. 16, pp. 1–52, 1980.

55. Miele, A., *Primal–Dual Sequential Gradient–Restoration Algorithms for Optimal Control Problems and Their Application to Flight in a Windshear,* in Proceedings of the International Conference on Optimization Techniques and Applications, (edited by K. L. Teo et al.), National University of Singapore, Kent Ridge, Singapore, pp. 53–94, 1987.

56. Miele, A., *Recent Advances in Gradient Algorithms for Optimal Control Problems*, Journal of Optimization Theory and Applications, Vol. 17, pp. 361–430, 1975.

57. Miele, A., Damoulakis, J. N., Cloutier, J. R., and Tietze, J. L., *Sequential Gradient–Restoration Algorithm for Optimal Control Problems with Nondifferentiable Constraints*, Journal of Optimization Theory and Applications, Vol. 13, pp. 218–255, 1974.

58. Miele, A., Huang, H. Y., and Heideman, J. C., *Sequential Gradient–Restoration Algorithm for the Minimization of Constrained Functions*, Ordinary and Conjugate Gradient Version, Journal of Optimization Theory and Applications, Vol. 4, pp. 213–246, 1969.

59. Miele, A., Levy, A., and Cragg, E. E., *Modifications and Extensions of the Conjugate Gradient–Restoration Algorithm for Mathematical Programming Problems*, Journal of Optimization Theory and Applications, Vol. 7, pp. 450–472, 1971.

60. Miele, A., Pritchard, R. E., and Damoulakis, J. N., *Sequential Gradient–Restoration Algorithm for Optimal Control Problems*, Journal of Optimization Theory and Applications, Vol. 5, pp. 235–282, 1970.

61. Miele, A., and Wang, T., *Dual-Properties of Sequential Gradient–Restoration Algorithms for Optimal Control Problems*, in R. Conti, E. De Giorgi and F. Giannessi (eds.), Optimization and Related Fields, Springer–Verlag, pp. 331–357, 1986.

62. Miele, A., and Wang, T., *Primal–Dual Properties of Sequential Gradient–Restoration Algorithms for Optimal Control Problems*, Part 1, Basic Problems, in F. R. Payne et al. (ed.), Integral Methods in Science and Engineering, Hemisphere Publishing Corporation, Washington, DC, pp. 577–607, 1986.

63. Miele, A., and Wang, T., *Primal–Dual Properties of Sequential Gradient–Restoration Algorithms for Optimal Control Problems*, Part 2, General Problem, Journal of Mathematical Analysis and Applications, Vol. 119, pp. 21–54, 1986.

64. Miele, A., Want, T., and Basapur, V. K., *Primal and Dual Formulations of Sequential Gradient–Restoration Algorithms for Trajectory Optimization Problems*, Acta Astronautica, Vol. 13, pp. 491–505, 1986.

65. Mukai, H., and Polak, E., *On the Use of Approximations in Algorithms for Optimization Problems with Equality and Inequality Constraints*, SIAM Journal of Numerical Analysis, Vol. 15, pp. 674–693, 1978.

66. Murray, J. M., and Teo, K. L., *On a Computational Algorithm for a Class of Optimal Control Problems Involving Discrete Time Delayed Arguments*, Journal of Australian Mathematical Society, Series B, Vol. 20, pp. 315–343, 1978.

67. Nababan, S.M., and Teo, K. L., *An Existence Theorem for Optimal Controls of Non–Linear Vector Delay–Differential Equations with Time–Lag Controls*, International Journal of Systems Science, Vol. 8, No. 10, pp. 1163–1172, 1977.

68. Noussair, E. S., *On the Existence of Piecewise Continuous Optimal Controls*, Journal of Australian Mathematical Society, Series B, pp. 31–37, 1977.

69. Oguztöreli, M. N., *Time Lag Control Systems*, Academic Press, 1966.

70. Polak, E., and Mayne, D. Q., *A Feasible Directions Algorithm for Optimal Control Problems with Control and Terminal Inequality Constraints*, IEEE Transactions on Automatic Control, Vol. AC–22, pp. 741–751, 1977.

71. Polak, E., and Mayne, D. Q., *First–Order Strong Variation Algorithms for Optimal Control Problems with Terminal Inequality Constraints*, Journal of Optimization Theory and Applications, Vol. 16, pp. 303–325, 1975.

72. Powell, M. J. D., *A Fast Algorithm for Nonlinearly Constrained Optimization Calculations*, in G. A. Matson (ed.), Numerical Analysis : Lecture Notes in Mathematics, Springer–Verlag, Vol. 630, 1978.

73. Rao, K. C., Prabhu, S. S., and Mehta, S. C., *Optimal Control of Linear Distributed Parameters Systems by Finite–Element Galerkin's Technique*, Optimal Control Applications and Methods, Vol. 3, pp. 66–77, 1982.

74. Rao, P. V., Lau, C. K., and Quek, L. C., *Optimization Technique Based on Liapunov Concepts*, Computer and Electric Engineering, Vol. 8, pp. 193–206, 1981.

75. Ray, W. H., and Szekely, J., *Process Optimization*, John–Wiley & Sons, New York, New York, 1973.

76. Rom, M., and Avriel, M., *Properties of the Sequential Gradient–Restoration Algorithm* (*SGRA*), Part 1 : Introduction and Comparison with Related Methods, Journal of Optimization Theory and Applications, Vol. 62, pp. 77–98, 1989.

77. Rom, M., and Avriel, M., *Properties of the Sequential Gradient–Restoration Algorithm (SGRA)*, Part 2 : Convergence Analysis, Journal of Optimization Theory and Applications, Vol. 62, pp. 99–125, 1989.

78. Rosen, J. B., *The Gradient Projection Method for Nonlinear Programming.* Part 1. Linear Constraints, Journal of Industrial and Applied Mathematics Society, Vol. 8, pp. 181–217, 1960.

79. Rosen, J. B., *The Gradient Projection Method for Nonlinear Programming.* Part 2. Nonlinear Constraints, Journal of Industrial and Applied Mathematics Society, Vol. 9, pp. 514–532, 1961.

80. Royden, H. L., *Real Analysis,* Second Edition, The MacMillan Company, New York, 1968.

81. Sakawa, Y., *On Local Convergence of an Algorithm for Optimal Control,* Numcerical Functional Analysis and Optimization, Vol. 3, pp. 301–319, 1981.

82. Sakawa, Y., and Shindo, Y., *On Global Convergence of an Algorithm for Optimal Control,* IEEE Transaction on Automatic Control, Vol. AC–25, pp. 1149–1153, 1980.

83. Schittkowski, K., *NLPQL : A FORTRAN Subroutine Solving Constrained Nonlinear Programming Problems,* Operations Research Annals, Vol. 5 pp. 485–500, 1985.

84. Schittkowski, K., *On the Convergence of a Sequential Quadratic Programming Method with an Augumented Lagrangian Line Search Function,* Math. Operationsforschung and Statistik, Ser. Optimization, Vol. 14, pp. 197–216, 1983.

85. Shindo, Y., and Sakawa, Y., *Local Convergence of an Algorithm for Solving Optimal Control Problems,* Journal of Optimization Theory and Applications, Vol. 46, pp. 265–293, 1985.

86. Sirisena, H. R., *Computation of Optimal Controls Using a Piecewise Polynomial Parametrization,* IEEE Transactions on Automatic Control, Vol. AC–18, pp. 409–411, 1973.

87. Sirisena, H. R., and Chou, F. S., *An Efficient Algorithm for Solving Optimal Control Problems with Linear Terminal Constraints,* IEEE Transactions on Automatic Control, Vol. 4 AC–214, pp. 275–277, 1976.

88. Sirisena, H. R., and Chou, F. S., *Convergence of Control Parametrization Ritz Method for Nonlinear Optimal Control Problems,* Journal of Optimization Theory and Applications, Vol. 29, pp. 369–382, 1979.

89. Sirisena, H. R., and Tan, K. S., *Computation of Constrained Optimal Controls Using Parametrization Techniques,* IEEE Transactions on Automatic Control, Vol. AC–19, pp. 431–433, 1974.

90. Styczen, K., *Characterization of the Smoothness of Optimal Periodic Control Problem Solution,* International Journal of Control, Vol. 46, pp. 753–767, 1987.

91. Sworder, D. D., *Feedback Control of a Class of Linear Systems with Jump Parameters*, IEEE Transactions on Automatic Control, Vol. AC–14, No. 1, pp. 9–14, 1969.

92. Teo, K. L., *Convergence Analysis for a Computational Method for Optimal Control*, Problems of Control and Information, Vol. 16, pp. 315–328, 1987.

93. Teo, K. L., Ahmed, N. U., and Fisher, M. E., *Optimal Feedback Control for Linear Stochastic Systems Driven by Counting Processes*, Journal of Engineering Optimization, Vol. 36, pp. 479–486, 1989.

94. Teo, K. L., Ang, B. W., and Wang, C. M., *Least Weight Cables : Optimal Parameter Selection Approach*, Engineering Optimization, Vol. 9, pp. 249–264, 1985.

95. Teo, K. L., and Clements, D. J., *Control Parametrization Algorithm for Convex Optimal Control Problems with Linear Constraints*, Numerical Functional Analysis and Optimization, Vol. 8, pp. 515–540, 1985–1986.

96. Teo, K. L., and Craven, B. D., *On a Computational Algorithm for Time–Lag Optimal Control Problems with Restricted Phase Coordinates*, Journal of Australian Mathematical Society, Vol. B. 21, pp. 385–397, 1980.

97. Teo, K. L., and Goh, C. J., *A Computational Method for Combined Optimal Parameter Selection and Optimal Control Problems with General Constraints*, Journal of Australian Mathematical Society, Series B, Vol. 30, pp. 350–364, 1989.

98. Teo, K. L., and Goh, C. J., *A Simple Computational Procedure for Optimization Problems with Functional Inequality Constraints,* IEEE Transactions on Automatic Control, Vol. AC–32, pp. 940–941, 1987.

99. Teo, K. L., and Goh, C. J., *A Unified Computational Method for Several Stochastic Optimal Control Problems,* in International Series of Numerical Mathematics, Birkhauser Verlag, Vol. 86, pp. 467–476, 1989.

100. Teo, K. L., and Goh, C. J., *On Constrained Optimization Problems with Non–Smooth Cost Functionals,* Applied Mathematics and Optimization, Vol. 17, pp. 181–190, 1988.

101. Teo, K. L., Goh, C. J., and Lim, C. C., *A Computational Method for a Class of Dynamical Optimization Problems in which the Terminal Time is Conditionally Free,* IMA Journal of Mathematical Control and Information, Vol. 6, pp. 81–95, 1989.

102. Teo, K. L., Goh, C. J., and Wong, K. H., *A Unified Computational Approach to Optimal Control Problems,* Longman Scientific and Technical, 1991.

103. Teo, K. L., and Jennings, L, S., *Nonlinear Optimal Control Problems with Continuous State Inequality Constraints,* Journal of Optimization Theory and Applications, Vol. 63, pp. 1–22, 1989.

104. Teo, K. L., and Jennings, L. S., *Optimal Control with a Cost on Changing Control,* Journal of Optimization Theory and Applications, Vol. 68, No. 2, 1991.

105. Teo, K. L., Jepps, G., Moore, E. J., and Hayes, S., *A Computational Method for Free Time Optimal Control Problems, with Application to Maximizing the Range of an Aircraft–like Projectile,* Journal of Australian Mathematical Society, Series B, Vol. 28, pp. 393–413, 1987.

106. Teo, K. L., Leong, K. K., and Goh, C. J., *Nonlinearly Constrained Optimal Control Problems Involving Piecewise Smooth Controls,* Journal of Australian Mathematical Society, Series B, Vol. 32, pp. 151–179, 1990.

107. Teo, K. L., Liu, Y., and Goh, C. J., *Nonlinearly Constrained Discrete Time Optimal Control Problems,* Applied Mathematics and Computation, Vol. 38, pp. 227–248, 1990.

108. Teo, K. L., and Womersley, R. S., *A Control Parametrization Algorithm for Optimal Control Problems Involving Linear Systems and Linear Terminal Inequality Constraints,* Numerical Functional Analysis and Optimization, Vol. 6, pp. 291–313, 1983.

109. Teo, K. L., and Wong, K. H., *A Computational Method for Time–Lag Control Problems with Control and Terminal Inequality Constraints,* Optimal Control Applications and Methods, Vol. 8, pp. 377–396, 1987.

110. Teo, K. L., and Wong, K. H., *Nonlinearly Constrained Optimal Control Problems,* to appear in Australian Mathematical Society, Series B.

111. Teo, K. L., Wong, K, H., and Clements, D. J., *A Feasible Directions Algorithm for Time–Lag Optimal Control Problems with Control and Terminal Inequality Constraints*, Journal of Optimization Theory and Applications, Vol. 46, pp. 295–317, 1985.

112. Teo, K. L., Wong, K. H., and Clements, D. J., *Optimal Control Computation for Linear Time–Lag Systems with Linear Terminal Constraints*, Journal of Optimization Theory and Applications, Vol. 44, pp. 509–526, 1984.

113. Teo, K. L., Wong, K. H., and Goh, C. J., *Optimal Maintenance of a System of Machines with Weakest Link Dependence Performance*, Optimal Control Methods and Applications, Vol. 10, pp. 113–127, 1989.

114. Teo, K. L., Wong, K. H., and Wu, Z. S., *An Optimal Control Problem Involving a Class of Linear Time–Lag Systems*, Journal of Australian Mathematical Society, Series B, Vol. 28, pp. 93–113, 1986.

115. Teo, K. L., and Wu, Z. S., *Computational Methods for Optimizing Distributed Systems*, Academic Press, Orlando, 1984.

116. Teo, K. L., Wu, A. S., and Clements, D. J., *A Computational Method for Convex Optimal Control Problems Involving Linear Hereditary Systems*, International Journal of Systems Science, Vol. 12, No. 9, pp. 1045–1060, 1981.

117. Tsuzakan, T., *Existence of Optimal Controls for Problems Involving Systems with Delays,* Ph.D. Thesis, Brown University, 1975.

118. Wang, P. K. C., *Time–Optimal Control of Time–Lag Systems with Time–Lag Controls,* Journal of Mathematical Analysis and Applications, Vol. 52, pp. 366–378, 1975.

119. Ward, R. C., *Numerical Computation of the Matrix Exponential with an Accuracy Estimate,* SIAM Journal of Numerical Analysis, Vol. 14, No. 4, pp. 600–610, 1977.

120. Weiss, R., *Transportation Lag–An Annotated Bibliography,* IRE on Automatic Control, Vol. AC–4, pp.56–68, 1959.

121. Wolfe, P., *Methods for Nonlinear Constraints,* Nonlinear Programming, edited by J Abadie, North–Holland, Amsterdam, 1967.

122. Wong, K. H., *A Control Parameterization Algorithm for Nonlinear Time–Lag Optimal Control Problems,* Opsearch, Vol. 25, pp. 177–184, 1988.

123. Wong, K. H., *Convergence Analysis of a Computational Method for Time–Lag Optimal Control Problems,* International Journal of Systems Science, Vol. 19, pp. 1437–1450, 1988.

124. Wong, K. H., *On the Computational Algorithms for Time–Lag Optimal Control Problems,* Ph.D. Thesis, University of New South Wales, 1984.

125. Wong, K. H., Clements, D. J., and Teo, K. L., *Optimal Control Computation for Nonlinear Time–Lag Systems*, Journal of Optimization Theory and Applications, Vol. 47, pp. 91–107, 1985.

126. Wong, K. H., and Teo, K. L., *A Conditional Gradient Method for a Class of Time–Lag Optimal Control Problems*, Journal of Australian Mathematical Society, Series B, Vol. 25, pp. 518–537, 1984.

127. Wong, K. H., Goh, C. J., and Teo, K. L., *Optimal Maintenance of Stochastically Deteriorating Systems*, International Journal of Systems Science, Vol. 22, No 2, pp. 375–381, 1991.

128. Wonham, W. M., *On a Matrix Riccati Equation of Stochastic Control*, SIAM Journal of Control, Vol. 6, No. 4, pp. 681–697, 1968.

129. Wonham, W. M., *Random Differential Equations in Control Theory*, Probabilistic methods in Applied Mathematics, Vol. 2, A. T. Bharucha–Reid, New York, Academic, pp. 131–212 1970.