工
4 55

A NUMERICAL APPROACH TO THE SOLUTION OF NONLINEAR OPTIMAL CONTROL PROBLEMS

MASASHI OKUDAIRA

A Numerical Approach to the Solution of Nonlinear Optimal Control Problems

MASASHI OKUDAIRA

MARCH 1979

Preface

This dissertation develops several algorithms for the optimal control of nonlinear dynamical systems. The systems under consideration are described by nonlinear differential equations and the objective is to find a control function with or without constraint which steers the state of the system from a manifold to another manifold so as to minimize an associated performance index of integral type. The fundamental attitude of the dissertation is to reduce the problem to a sequence of linear two-point boundary-value problems (TPBVP's).

The dissertation consists of four chapters, including the introductory one. Chapter 2 develops a time-decomposition algorithm for solving stiff linear TPBVP's, that is, linear TPBVP's with rapidly convergent and rapidly divergent particular solutions. The algorithm belongs to the multipoint approach and succeeds in reducing the numerical error in applying the superposition principle. The chapter that follows discusses the solution of nonlinear optimal control problems without control constraint. The time-decomposition algorithm is applied to the problem in combination with linearization methods. The combined algorithm can effectively be applied also to multiple-target problems, that is, problems containing discontinuities and additional boundary conditions in the intermediate points of control duration. Chapter 4, the last chapter, deals with control-constrained optimization problems. A direct method is modified to treat problems with specified terminal condition. The problem is reduced to a sequence of linear TPBVP's by introducing an artificial variable. Various typical examples attached in each chapter illustrate several features of the proposed algorithms.

Acknowledgements

The author wishes to express his sincere gratitude to Dr. Yoshikazu Nishikawa, Professor of Kyoto University, who has suggested the field of reseach of the present dissertation and has given him a constant and generous guidance and encouragement in promoting this work. In the preparations of the present disseration, the author has been given many aids and useful advices by Dr. Takeo Ojika, Associate Professor of Osaka Kyoiku University. Acknowledgement must also be made to Dr. Nobuo Sannomiya, Associate Professor of Kyoto University, and Dr. Hidekiyo Itakura, Research Associate of Kyoto University, for stimulating and invaluable discussions of all kind. The author is likewise grateful to the members of Professor Nishikawa's Laboratory for their excellent cooperation.

All of the numerical calculations were made on FACOM M-190 at the Data Processing Center of Kyoto University.

Contents

Chapter 1

Introduction

1.1. Optimal Control Problems in Nonlinear Systems

With the advance in the field of digital computers, enormous efforts have been expended on the development of numerical solution techniques for optimal control problems in nonlinear systems within last two decades [1, 6, 8, 12, 17, 33, 43, 46, 66, 67, 78]. These techniques can be classified into two categories. The one is so called direct methods [8-11, 15, 34, 35, 46, 48, 50, 65] and the other indirect methods [4, 6, 16, 17, 41, 42, 55, 63, 73]. The direct method generates a sequence of control functions so that the given performance index may successively be minimized or maximized. On the other hand, the indirect method transforms the problem into a two-point boundary-value problem (TPBVP) by applying the minimum principle [68] or the variational principle [28] and the optimal control is determined by its solution.

The optimal control problem is formulated as follows. The differential equation which governs the dynamical system is expressed as

$$\dot{x} = f(t, x, u), \tag{1.1}$$

where $x(t)$ is an n-dimensional state vector and $u(t)$ is an m-dimensional control vector. The objective is to find the control function u which transfers the state x from a manifold to another manifold so as to minimize the following performance index:

$$J = \int_{t_0}^{t_f} L(t, x, u)dt. \tag{1.2}$$

In a practical sense, a certain constraint is imposed on the control function such as

$$g(u) \leqq 0. \tag{1.3}$$

The functional g is often of the saturation type, that is, the set U of the admissible control functions often takes the form of

$$U = \{u|\ |u_i(t)| \leqq M_i,\ t \in [t_0, t_f],\ i = 1, 2, \ldots, m\}. \tag{1.4}$$

Such a formulation of optimal control problems requires the complete information about the system to be controlled and this is one of reasons why the state-space approach is less made use of than the transfer function approach in many engineering systems [29]. However, there are many problems to which the state-space approach can effectively be applied. For example, to the problem in the field of aerospace engineering, the approach is suitable by nature of the problem [46, 54, 60]. Also in other fields, many workers have made efforts in the application of the approach [53].

Since the direct method adjusts the control function directly at each iteration of the calculation, it is not so difficult to take (1.3) or (1.4) into consideration, while it is rather elaborate to adjust the control function so that the state may satisfy the imposed manifold condition. The indirect method, on the contrary, can easily deal with specified terminal point problems, though it is difficult to apply it to the control-constrained problem except for the linear problem of small order.

This dissertation is concerned with the development of computational schemes in view of the indirect method. In applying the indirect method to the optimal control problems, there are at least two main difficulties to be resolved: The one is that the derived TPBVP is often difficult to solve because of the numerical error. The other is that it is rather difficult or even impossible to treat problems with control constraint as (1.4).

Generally the derived TPBVP is nonlinear. There are two approaches to solve nonlinear TPBVP's, that is to say, the shooting method [2, 40, 64, 70, 72] and the linearization method [4, 42, 55, 63]. The shooting method solves the given nonlinear differential equation iteratively with a sequence of estimated initial conditions of the missing ones until the solution satisfies the manifold condition. The method suffers from the critical sensitivity of boundary values to any small change of initial conditions and therefore often fails to obtain the solution to the non-linear problem, though it has the advantage that only initial conditions in the preceding iteration need be available for the present calculation.

On the other hand, the linearization method reduces the nonlinear TPBVP to a sequence of linear TPBVP's, whose solutions satisfy the manifold condition and are to converge to the solution to the original nonlinear problem. The difficulty of solving nonlinear TPBVP's by the method, which is adopted in the text, lies in the fact that the linearized differential equation often has a stiff structure, that is, the solution has both rapidly convergent and rapidly divergent components [23]. The stiffness causes serious numerical error in applying the superposition principle [42] to the linearized TPBVP. Therefore, one must develop an efficient computational algorithm to solve such a differential equation.

Several algorithms so called multipoint approach have been proposed for such a TPBVP [22, 47, 51, 52, 61, 62, 71]. The fundamental idea lying in these algorithms is to divide the overall integration interval into several subintervals, accordingly, the idea is to divide the overall TPBVP into several sub-TPBVP's [62] or initial value problems [22, 47, 51, 71] with shorter integration intervals, since the stiffness does not cause so serious numerical error when the integration interval is not so long. The solution to the original problem is obtained by adjusting the provisionally assumed boundary values so as to make the solutions of subproblems continuous at the boundaries. In the text, a time-decomposition algorithm for a linear TPBVP, which was first presented as a two-subinterval algorithm [59, 61, 62], is extended to multi-subinterval one in theory as well as in numerical experiments. The nonlinear optimal control problems, including a problem with discontinuities and additional boundary conditions in the intermediate points of control duration, are solved by the algorithm in combination with the interaction-coordination algorithm [55, 63] and the quasilinearization method [42].

For the numerical solution of control-constrained problems, the direct method has mainly been employed. This is chiefly because the TPBVP derived by the minimum principle usually contains very strong nonlinearities which result from necessary conditions for optimality and therefore is practically very difficult or even impossible to solve. The direct method, however, becomes less effective when the terminal condition is specified. For example, the steepest-descent method [35] suffers from poor accuracy of the solution and the method proposed by Bryson and Denham [10] involves additional integration procedures, including the integration of matrix differential equation, for determining the Lagrange multipliers. Further the solutions obtained by these methods do not satisfy the specified terminal condition until the optimum is attained. Therefore, it is desirable to develop a new algorithm which remedies the defects the both methods have. The requirement for the new algorithm is to have the property that the control function is easily adjusted to satisfy the control constraint (1.3) or (1.4) and, at the same time, to make the state of the system satisfy the imposed manifold condition. Recently, Miele et al. have proposed a sequential gradient-restoration algorithm which satisfies the requirement [20, 24-26, 48, 50]. The algorithm is composed of the gradient phase and the restoration phase.. The gradient phase aims at the minimization of the values of the performance index while the restoration phase, composed of several iterations, aims at making the solution be consistent with all the constraint. The algorithm has a merit that each solution at the end of the restoration phase is a feasible one. However, turning inside out, excessive restoration phase must be carried out to obtain the optimal solution.

In this text, another algorithm is developed which also satisfies the above requirement [58]. The steepest-descent method is modified to treat the problem with specified terminal condition. The basic idea is to reduce the problem into a sequence of linear TPBVP's which contain explicitly the

the control variable term. To this end, an artificial variable is introduced in the system equation. The aritificial variable and the control variable are corrected iteratively, using the solution to the linear TPBVP, so as to attain the optimality.

1.2. Description of Contents

This text consists of four chapters, including the present introductory chapter.

Chapter 2 discusses the time-decomposition algorithm to solve a stiff linear TPBVP. The algorithm was first presented as a two-subinterval algorithm [59, 61, 62]. In the text, it is extended to multi-subinterval one in theory as well as in numerical experiments. Moreover, an error analysis is made through an example by comparing the algorithm with a basic superposition principle.

In a stiff problem, since some particular solutions of the system equation increase and others decrease rapidly as the independent variable changes, the integration of the system equation suffers from a serious numerical error. In the time-decomposition algorithm, first, the overall interval of integration is divided into several subintervals at several intermediate points. These points are called 'torn times.' Then, in each subinterval, sub-TPBVP with arbitrarily chosen boundary conditions is solved. Second, the exact boundary values which guarantee the continuity of the solutions at the torn times are determined algebraically. Owing to the division of the integration interval, the numerical error is effectively reduced in spite of the stiffness. The effectiveness of the method is demonstrated by solving two illustrative examples.

Though it has been pointed out so far that such a division is much effective for stiff problems, no analysis has been made how the numerical error is reduced by the division. It is also shown in this text how it is, through examining an example.

In Chapter 3, the time-decomposition algorithm is applied to nonlinear optimal control problems. The performance index to be minimized is of a quadratic type in state and control. No constraint is imposed on the control function, that is, M_i in (1.4) is assumed to be infinity.

First, we consider the system described by a nonlinear differential equation without any discontinuity on the overall interval. The initial time, the terminal time, and the initial conditions of the state are specified, and the terminal conditions may be or may not be specified. The problem is reduced to a nonlinear TPBVP by applying the minimum principle and further reduced to a sequence of linear ones by the interaction-coordination algorithm [55, 63] and the quasilinearization method [4, 42]. Theoretically speaking, the linear TPBVP's can be solved by the superposition principle. However, as have been mentioned, the principle suffers from numerical errors, since the derived TPBVP's are in themselves more or less stiff. Especially when the interaction-coordination algorithm is employed to solve the nonlinear TPBVP, linear TPBVP's to be solved are quite stiff, since it is empirically known that the convergence property of the algorithm is much improved by modifying the original TPBVP's stiffer.

Two physical problems are solved by the time-decomposition algorithm in combination with the linearization methods to show the effectiveness of the combined algorithm.

Second, the multiple-target problem [30, 32, 56, 57, 74-76] is considered in the latter half of Chapter 3. The system equation contains discontinuities at intermediate points of the overall interval. These points are called 'corner times.' The boundary conditions are specified at several corner times as well as the initial and the terminal times. By specifying provisional corner times, the problem is reduced to a nonlinear multipoint boundary-value problem (MPBPV) due to the minimum principle. This nonlinear MPBVP is further reduced to a sequence of linear ones by use of the linearization methods mentioned above. The linear MPBVP is solved by the time-decomposition algorithm for a discontinuous problem. The problem is decomposed into several subinterval TPBVP's. The missing boundary conditions of these TPBVP's are determined by using the solutions obtained with arbitrarily chosen boundary conditions. Since, different from the continuous case, discontinuities of variables may occur at the corner times, it is impossible to integrate the differential equation in series. Therefore, decomposition of the overall interval plays an essential role to solve such a problem. After solving the nonlinear MPBVP, the provisionally specified corner times are corrected by a gradient method. The correction procedure is iterated until the optimum is attained. The solution in each iteration satisfies the boundary conditions exactly.

To illustrate how the combined algorithm works, two linear problems are solved and their solutios are compared with the analytical ones. An example of nonlinear problem is also attached.

In Chapter 4, a modified direct method is developed to solve optimal control problems in nonlinear systems where the control function is subject to the constraint of (1.4) [58]. The basic idea of the method is to modify a steepest-descent method, which is often adopted for such a problem, by introducing the interaction-coordination technique in order to make problems with the terminal manifold specified easily treatable. The steepest-descent method generates a sequence of control functions which successively reduce the value of the performance index. The control function is corrected iteratively by using the solution to the adjoint equation. The present idea is to modify the system equation by introducing an interaction variable, which acts as an additional control fuction, to make the state satisfy the terminal condition. By the method, the problem is reduced to a sequence of linear TPBVP's with the control function appearing explicitly. By use of their solutions, the control function is iteratively optimized and the interaction variable is iteratively corrected so that the solution to the modified equation coincides with that of the original equation.

In order to verify the effectiveness of the proposed algorithm, several examples are presented, including a state- and control-constrained problem and an on-off type problem.

Chapter 2

A Time-Decomposition Algorithm for the Solution of a Stiff Linear Two-Point Boundary-Value Problem

2.1. Introduction

In the last ten years, several researchers have been employed in the study of numerical solution of stiff systems for an ordinary differential equation(ODE) [22, 23, 38, 47, 51, 69, 71, 81]. The word 'stiff' originally means that the solution of an ODE contains both the 'much fast' and the 'much slowly' convergent components. For example, when the system matrix of the linear differential equation has both the large and the small negative eigenvalues, the system is stiff. Initial value problems for such an ODE appear in the analysis of electronic circuits [7, 13]. The difficulty of solving such a problem lies in the fact that one must choose integration step size small enough to follow the rapidly changing components while the integration interval requires to be long enough in order to solve the differential equation until the much slowly convergent components reach to steady state. Thus, the computational time and the stability of the integration routine become a matter of importance. For such a problem, many algorithms have been proposed, e.g., Adams-Bashforth, Adams-Moulton [27], and Gear [19].

In the text, the word 'stiff' is used in an extended sense [23], that is to say, we consider two-point boundary-value problems (TPBVP's) such as follows:

$$\frac{d}{dt}\begin{bmatrix} x \\ p \end{bmatrix} = D(t)\begin{bmatrix} x \\ p \end{bmatrix} + h(t), \tag{2.1}$$

with the boundary condition

$$x(t_0) = \pi_0,\ x(t_f) = \pi_f, \tag{2.2}$$

where x and p are variables of same order. The task is to find the missing initial conditions. Assume that the system matrix $D(t)$ be a constant matrix and have both the large positive and the large negative eigenvalues. Basically, the missing initial conditions can be determined by the superposition principle [42]. The obtained values, however, contain some numerical errors. Since the differential equation (2.1) is unstable both forward and backward, a few numerical errors at the initial point lead to a serious error at the end of the integration interval, whichever direction the equation may be integrated to. Different from initial value problems, such a difficulty can not be overcome only by improving the integration routine.

Two-point boundary-value problems appear in various areas of mathematical physics, for example, optimal control problems [e.g., 12, 18, 67], boundary-layer problems [64], phase-locked-loop design [77], and so on. In many cases, the TPBVP has a stiff structure. For example, the TPBVP derived from optimal control problems of regulator type, which will be treated in the following chapter, inevitably has both positive and negative eigenvalues.

Several algorithms due to multipoint approach have been proposed for such TPBVP's [22, 47, 51, 52, 61, 62, 71]. The fundamental idea lying in these algorithms is to divide the overall integration interval into several subintervals. Accordingly the idea is to divide the overall TPBVP into several sub-TPBVP's [61, 62] or initial value problems [22, 47, 51, 71] with shorter integration intervals. The intermediate points of division are called 'torn times' in what follows. The solution to the original problem is obtained by adjusting the provisionally given boundary values so as to ensure the continuity of the solutions of subproblems at the torn times.

In this chapter, a time-decomposition algorithm with two subintervals [61, 62] is extended to the one with multi-subintervals in theory as well as in numerical experiments. Moreover, an error analysis is made by comparing the algorithm with the basic superposition principle in application to an example.

The time-decomposition algorithm, which belongs to the methods due to multipoint approach, divides the overall TPBVP into sub-TPBVP's by assuming provisional boundary values at the torn times. Then, making use of their solutions, the boundary values are corrected by an algebraic means so as to ensure the continuity of the solutions at the torn times. This correction can be done by only one calculation. The solution to the overall problem is obtained by solving sub-TPBVP's with boundary conditions thus obtained.

In the above discussion, only the *linear* TPBVP is dealt with by the proposed algorithm. This may seem a serious limitation to the algorithm.

However, since *nonlinear* TPBVP's can be reduced to linear ones by making use of the well-known quasilinearization method [4, 42] or, by the interaction-coordination algorithm [55, 61, 63], the limitation of linearity is not so serious.

Techniques similar to the present algorithm have been developed in Refs. 22, 47, 51, and 71, which have dealt with nonlinear TPBVP's by dividing the overall interval into L-subintervals. In Refs. 22 and 71, the values of all the elements of the variable are estimated at each boundary, and the nonlinear differential equation is solved in each subinterval with these estimated initial values. Then, the estimated values are corrected by the quasilinearization method so that the estimated values at the i-th boundary may coincide with the calculated terminal values in the $(i-1)$-th subinterval. The solution does not satisfy the continuity condition until the iteration is terminated. In Refs. 47 and 51, at the outset the nonlinear TPBVP is reduced to a sequence of linear ones by the quasilinearization method. Then, the overall interval is divided into L-subintervals. All the values at each boundary are determined by the method of particular solutions [45].

These algorithms determine all the boundary values at once, so that they must take the inverse of an $n(L-1)\times n(L-1)$-dimensional matrix when the order of the differential equation is n.

On the other hand, our time-decomposition algorithm reduces the problem to L-sub-TPBVP's. Therefore, in order to make the solutions continuous at the torn times, only $n(L-1)/2$ elements are to be determined at once. That is to say, the algorithm needs only to take the inverse of an $\{n(L-1)/2\}\times\{n(L-1)/2\}$-dimensional matrix, the existence of which is also discussed in the text.

Baumann [3] has proposed a trajectory-decomposition algorithm for a control problem with discontinuity, which is also applicable to the problem considered in this chapter. However, since the subinterval boundary values are corrected by a gradient method, it often takes much computing time to attain exact values.

The organization of this chapter is as follows. In Section 2.2, the principle of superposition is briefly summarized, on which the time-decomposition algorithm is based. Section 2.3 discusses the time-decomposition algorithm in detail. This section includes several theorems for the algorithm. The algorithm is summarized in Section 2.4 and is applied to two illustrative examples with stiff structure in Section 2.5. In Section 2.6, by examining an example, it is analyzed how the numerical error is reduced by the proposed algorithm.

2.2. The Principle of Superposition

Let us consider the TPBVP of (2.1) and (2.2). Both $x(t)$ and $p(t)$ are n-dimensional vectors, $D(t)$ is a $2n\times 2n$-dimensional matrix, and $h(t)$ is a $2n$-dimensional vector function. Both $D(t)$ and $h(t)$ are continuous in t. Though, for simplicity, the boundary condition is assumed to be given as (2.2) in the following, more general case can be treated with a slight modification of the argument.

Let $\Phi(t, t_0)$ be the $2n\times 2n$-dimensional transition matrix corresponding to the homogeneous part of (2.1) with the initial condition $\Phi(t_0, t_0) = I_{2n}$, where I_{2n} is the $2n\times 2n$-dimensional identity matrix. Then the solution to

(2.1) subject to a set of initial conditions $[x'(t_0), p'(t_0)]'$ can be written as

$$x(t) = \Phi_{11}(t, t_0)x(t_0) + \Phi_{12}(t, t_0)p(t_0) + v_1(t, t_0), \quad (2.3)$$

$$p(t) = \Phi_{21}(t, t_0)x(t_0) + \Phi_{22}(t, t_0)p(t_0) + v_2(t, t_0), \quad (2.4)$$

where $\Phi_{ij}(t, t_0)$ is the $n \times n$-dimensional element matrix of $\Phi(t, t_0)$ ($i, j =$ 1, 2) with

$$\begin{bmatrix} \Phi_{11}(t, t_0), & \Phi_{12}(t, t_0) \\ \Phi_{21}(t, t_0), & \Phi_{22}(t, t_0) \end{bmatrix} \triangleq \Phi(t, t_0)$$

and

$$v(t, t_0) = \begin{bmatrix} v_1(t, t_0) \\ v_2(t, t_0) \end{bmatrix} \triangleq \int_{t_0}^{t} \Phi(t, \tau)h(\tau)d\tau .$$

Let $\bar{p}_0$ be an initial approximation to the initial condition of p. Then, from (2.3) subject to $x(t_0) = \pi_0$ and $p(t_0) = \bar{p}_0$, the corresponding terminal value $\bar{x}_f \triangleq x(t_f)$ is given by

$$\bar{x}_f = \Phi_{11}(t_f, t_0)\pi_0 + \Phi_{12}(t_f, t_0)\bar{p}_0 + v_1(t_f, t_0). \quad (2.5)$$

Similarly, let p_0 be the exact initial condition of p which satisfies the given terminal condition $x(t_f) = \pi_f$. Then, from (2.3) with $x(t_0) = \pi_0$ and $p(t_0) = p_0$, we have

$$\pi_f = \Phi_{11}(t_f, t_0)\pi_0 + \Phi_{12}(t_f, t_0)p_0 + v_1(t_f, t_0). \quad (2.6)$$

Subtracting (2.5) from (2.6) gives

$$\Phi_{12}(t_f, t_0)(p_0 - \bar{p}_0) = \pi_f - \bar{x}_f . \tag{2.7}$$

The result obtained above is summarized in the following lemma.

Lemma 2.1.

Let $\bar{x}_f$ be the terminal value of x obtained from (2.3) with a set of initial conditions $[\pi_0', \bar{p}_0']'$, where $\bar{p}_0$ is an arbitrary n-dimensional vector. If the matrix $\Phi_{12}(t_f, t_0)$ is nonsingular, then, from (2.7), the exact initial condition of p is given by

$$p_0 = \bar{p}_0 + \Phi_{12}^{-1}(t_f, t_0)(\pi_f - \bar{x}_f) . \tag{2.8}$$

2.3. Time-Decomposition Algorithm

In this section, a computational technique termed a time decomposition is proposed in order to overcome the difficulty caused by numerical errors in applying the superposition principle [42, 57, 59, 62]. The time decomposition is to decompose the overall interval into several subintervals, say, L-subintervals.

Suppose that $\overline{x(t_i)}$ be an estimated value of p at the torn time t_i ($i = 1, 2, \ldots, L-1$). Then, from (2.3) and (2.4), it is seen that the solution to (2.1) in the subinterval $[t_i, t_{i+1}]$ with the boundary conditions $x(t_i) = \overline{x(t_i)}$ and $x(t_{i+1}) = \overline{x(t_{i+1})}$ satisfies

$$\overline{x(t_{i+1})} = \phi_{11}(i+1, i)\overline{x(t_i)} + \phi_{12}(i+1, i)\overline{p(t_i^+)} + v_1(t_{i+1}, t_i), \tag{2.9}$$

$$\overline{p(t_{i+1}^-)} = \phi_{21}(i+1, i)\overline{x(t_i)} + \phi_{22}(i+1, i)\overline{p(t_i^+)} + v_2(t_{i+1}, t_i) \tag{2.10}$$

$$(i = 0, 1, \ldots, L-1;\ t_L = t_f),$$

where $\phi_{ij}(\lambda, \mu)$ denotes $\Phi_{ij}(t_\lambda, t_\mu)$ and $p(t_i^+)$ and $p(t_{i+1}^-)$ denote values of the variable p at t_i and t_{i+1}, respectively, in the i-th subinterval. For the estimated boundary condition $\overline{x(t_i)}$, the continuity condition for p at $t = t_i$, i.e. $p(t_i^-) = p(t_i^+)$, is not always satisfied.

By calculating the difference $\overline{p(t_i^-)} - \overline{p(t_i^+)}$, let us now determine the exact value of x at $t = t_i$, with which both the given boundary conditions and the continuity condition of p at $t = t_i$ are satisfied. If $\phi_{12}(i+1, i)$ is nonsingular, we can rewrite (2.9) as follows:

$$\overline{p(t_i^+)} = \phi_{12}^{-1}(i+1, i)[\overline{x(t_{i+1})} - \phi_{11}(i+1, i)\overline{x(t_i)} - v_1(t_{i+1}, t_i)]. \quad (2.11)$$

Substitution of (2.11) into (2.10) yields

$$\overline{p(t_{i+1}^-)} = \phi_{21}(i+1, i)\overline{x(t_i)} + \phi_{22}(i+1, i)\phi_{12}^{-1}(i+1, i)[\overline{x(t_{i+1})} - \phi_{11}(i+1, i)\overline{x(t_i)} - v_1(t_{i+1}, t_i)] + v_2(t_{i+1}, t_i). \quad (2.12)$$

By replacing i in (2.11) by $i+1$ and subtracting it from (2.12), we obtain

$$\overline{p(t_{i+1}^-)} - \overline{p(t_{i+1}^+)} = S_i \overline{x(t_i)} + T_i \overline{x(t_{i+1})} + U_i \overline{x(t_{i+2})} + V_i, \quad (2.13)$$

where

$$\left.\begin{aligned} S_i &= \phi_{21}(i+1, i) - \phi_{22}(i+1, i)\phi_{12}^{-1}(i+1, i)\phi_{11}(i+1, i), \\ T_i &= \phi_{22}(i+1, i)\phi_{12}^{-1}(i+1, i) + \phi_{12}^{-1}(i+2, i+1)\phi_{11}(i+2, i+1), \\ U_i &= -\phi_{12}^{-1}(i+2, i+1), \\ V_i &= -\phi_{22}(i+1, i)\phi_{12}^{-1}(i+1, i)v_1(t_{i+1}, t_i) + v_2(t_{i+1}, t_i) + \\ &\quad \phi_{12}^{-1}(i+2, i+1)v_1(t_{i+2}, t_{i+1}) \\ &\qquad (i = 0, 1, \ldots, L-2). \end{aligned}\right\} \quad (2.14)$$

Taking (2.13) and (2.14) for all i $(i=0, 1, \ldots, L-2)$ into account, we can establish the following theorem.

Theorem 2.1.

Let p be the solution to (2.1) in the interval $[t_i, t_{i+1}]$ with the boundary conditions $x(t_i)=\overline{x(t_i)}$ and $x(t_{i+1})=\overline{x(t_{i+1})}$ $(i=0, 1, \ldots, L-1;$ $\overline{x(t_0)}=\pi_0$, $\overline{x(t_f)}=\pi_f)$. Let $\overline{p(t_i^+)}$ and $\overline{p(t_{i+1}^-)}$ be the value of p at $t=t_i$ and $t=t_{i+1}$, respectively, in the i-th subinterval. Then, the following relation holds:

$$\overline{P} = \Gamma \overline{X} + V \qquad (2.15)$$

where

$$\Gamma = \begin{bmatrix} T_0, U_0, & & & \\ S_1, T_1, U_1, & & \mathbf{0} & \\ & \ddots & \ddots & \ddots \\ & & S_{L-3}, T_{L-3}, U_{L-3}, & \\ \mathbf{0} & & & S_{L-2}, T_{L-2} \end{bmatrix}, \qquad (2.16)$$

$$\overline{X} = [\overline{x'(t_1)}, \overline{x'(t_2)}, \ldots, \overline{x'(t_{L-1})}]', \qquad (2.17)$$

$$\overline{P} = [(\overline{p(t_1^-)} - \overline{p(t_1^+)})', \ldots, (\overline{p(t_{L-1}^-)} - \overline{p(t_{L-1}^+)})']', \qquad (2.18)$$

$$V = [(S_0\pi_0 + V_0)', V_1', \ldots, V_{L-3}', (V_{L-2} + U_{L-2}\pi_f)']'. \qquad (2.19)$$

Note that Γ and V are independent of the choice of the boundary conditions $\overline{x(t_i)}$ $(i=1, 2, \ldots, L-1)$. Hence we have the following corollary.

Corollary 2.1.

Suppose Γ of (2.16) be nonsingular. Let $X=[x'(t_1), x'(t_2), \ldots, x'(t_{L-1})]'$, where $x(t_i)$ is the value of the exact solution to (2.1) with the boundary conditions $x(t_0) = \pi_0$ and $x(t_f) = \pi_f$ at $t = t_i$ ($i = 1, 2, \ldots, L-1$). Then X and $\overline{X}$ of (2.17) are related by the following algebraic equation.

$$X = \overline{X} - \Gamma^{-1}\overline{P}. \tag{2.20}$$

Proof.

For the exact solution $x(t_i)$, obviously p is continuous in $t \in [t_0, t_f]$. Hence,

$$0 = \Gamma X + V. \tag{2.21}$$

Subtract (2.15) from (2.21). Then the nonsingularity of Γ proves the validity of (2.20).

Q.E.D.

Remark 2.1.

Corollary 2.1 means that the solution to the given TPBVP can be obtained by solving several numbers of the subinterval TPBVP's. Hence, it is suggested that the time-decomposition algorithm is also applicable to the problem, having discontinuities in the system equation, which will be discussed in the following chapter.

Now let us consider the nonsingularity of Γ.

Theorem 2.2.

Suppose that $\phi_{12}(\lambda, 0)$ ($\lambda = 1, 2, \ldots, L$) and $\phi_{12}(\lambda+1, \lambda)$ ($\lambda = 1, 2, \ldots, L-1$) be nonsingular. Then, Γ is nonsingular.

Before proceeding to the proof of Theorem 2.2, we prove the following two lemmas.

Lemma 2.2.

For arbitrary λ, μ, and ν, the following relation holds:

$$\phi_{ij}(\lambda, \nu) = \sum_{k=1}^{2} \phi_{ik}(\lambda, \mu)\phi_{kj}(\mu, \nu) \qquad (i, j = 1, 2). \tag{2.22}$$

Proof.

From the transition property of Φ,

$$\Phi(t_\lambda, t_\nu) = \Phi(t_\lambda, t_\mu)\Phi(t_\mu, t_\nu). \tag{2.23}$$

Hence, expansion of (2.23) proves (2.22).

Q.E.D.

Lemma 2.3.

Assert the hypothesis of Theorem 2.2. Then, the following sequence of matrices $\overline{T}_i$ is well-defined:

$$\overline{T}_i = -S_i \overline{T}_{i-1}^{-1} U_{i-1} + T_i \qquad (i = 0, 1, \ldots, L-2), \tag{2.24}$$

where $S_0 \triangleq 0$, and $\overline{T}_{-1}^{-1}$ and U_{-1} are arbitrary matrices. Furthermore $\overline{T}_i$ is given by

$$\overline{T}_i = \phi_{12}^{-1}(i+2, i+1)\phi_{12}(i+2, 0)\phi_{12}^{-1}(i+1, 0). \tag{2.25}$$

Proof.

Clearly, it suffices to prove (2.25). We prove (2.25) inductively. First, by the definition of $\overline{T}_i$ in (2.14) and Lemma 2.2, we have

$$\overline{T}_0 = T_0 = \phi_{22}(1, 0)\phi_{12}^{-1}(1, 0) + \phi_{12}^{-1}(2, 1)\phi_{11}(2, 1)$$

$$= \phi_{12}^{-1}(2, 1)[\phi_{12}(2, 1)\phi_{22}(1, 0) + \phi_{11}(2, 1)\phi_{12}(1, 0)]\phi_{12}^{-1}(1, 0)$$

$$= \phi_{12}^{-1}(2, 1)\phi_{12}(2, 0)\phi_{12}^{-1}(1, 0). \tag{2.26}$$

(2.26) shows that (2.25) holds for $i = 0$.

Second, we show that the relation (2.25) holds for $i = k+1$ if it holds for $i = k$. From (2.24) and (2.25),

$$\begin{aligned}\overline{T}_{k+1} - T_{k+1} &= -S_{k+1}\overline{T}_k^{-1}U_k \\ &= [\phi_{21}(k+2, k+1) - \phi_{22}(k+2, k+1)\phi_{12}^{-1}(k+2, k+1)\cdot \\ &\quad \phi_{11}(k+2, k+1)]\phi_{12}(k+1, 0)\phi_{12}^{-1}(k+2, 0).\end{aligned} \tag{2.27}$$

Substituting

$$\phi_{21}(k+2, k+1)\phi_{12}(k+1, 0) = \phi_{22}(k+2, 0) - \phi_{22}(k+2, k+1)\phi_{22}(k+1, 0) \tag{2.28}$$

into (2.27), we obtain

$$\begin{aligned}\overline{T}_{k+1} &= T_{k+1} + \phi_{22}(k+2, 0)\phi_{12}^{-1}(k+2, 0) - \phi_{22}(k+2, k+1)\cdot[\\ &\quad \phi_{22}(k+1, 0) + \phi_{12}^{-1}(k+2, k+1)\phi_{11}(k+2, k+1)\phi_{12}(k+1, 0)]\cdot \\ &\quad \phi_{12}^{-1}(k+2, 0) \\ &= T_{k+1} + \phi_{22}(k+2, 0)\phi_{12}^{-1}(k+2, 0) - \phi_{22}(k+2, k+1)\phi_{12}^{-1}(k+2, k+1) \\ &= \phi_{12}^{-1}(k+3, k+2)\phi_{11}(k+3, k+2) + \phi_{22}(k+2, 0)\phi_{12}^{-1}(k+2, 0) \\ &= \phi_{12}^{-1}(k+3, k+2)\phi_{12}(k+3, 0)\phi_{12}^{-1}(k+2, 0).\end{aligned} \tag{2.29}$$

(2.29) implies that (2.25) holds for $i = k+1$. Thus, the proof is completed.

Q.E.D.

Now we can proceed to prove Theorem 2.2.

Proof of Theorem 2.2.

Let Λ_i be the nonsingular matrix defined by

$$\Lambda_i = \begin{array}{l} n\cdot i \left\{ \vphantom{\begin{matrix}a\\a\\a\end{matrix}} \right. \\ n(L-1) \left\{ \vphantom{\begin{matrix}a\\a\\a\\a\end{matrix}} \right. \end{array} \left[\begin{array}{ccc} I_n & & 0 \\ & \ddots & \\ & & -\bar{T}_{i-1}^{-1} U_{i-1} \\ \hdashline & & I_n \qquad \\ & 0 & \qquad \ddots \\ & & \qquad\qquad I_n \end{array}\right] \quad (i = 1, 2, \ldots, L-2). \qquad (2.30)$$

Multiplying Γ on the right by Λ_i $(i = 1, 2, \ldots, L-2)$ transforms Γ into

$$\begin{bmatrix} \bar{T}_0 & & 0 \\ & \ddots & \\ * & & \bar{T}_{L-2} \end{bmatrix} \qquad (2.31)$$

Due to Lemma 2.3, the matrix (2.31) is nonsingular. Hence, Theorem 2.2 is proved.

Q.E.D.

Remark 2.2.

The nonsingularity of $\Phi_{12}(t_f, t_0)$ is necessary and sufficient for the existence of a unique solution to the linear TPBVP (2.1) and (2.2). Similarly, the nonsingularity of $\Phi_{12}(t_{i+1}, t_i)$ is necessary and sufficient for the unique existence of i-th subinterval solution. Thus, the necessary and sufficient condition for the applicabilty of the time-decomposition algorithm to the TPBVP is that there exists a unique solution in each subinterval defined by arbitrary two torn times, including initial and terminal times.

Once the value of the transition matrix is obtained, Γ^{-1} could be calculated directly from (2.16). In practice, however, the following procedure of calculation is more efficient in reducing the numerical error: (i) for the linear TPBVP of (2.1) and (2.2), set $h(t) = 0$, $\pi_0 = \pi_f = 0$, and $\overline{X}$ equal to the ν-th unit vector ($\nu = 1, 2, \ldots, n(L-1)$), (ii) obtain $\overline{P}$ by solving the TPBVP with the corresponding boundary condition in each subinterval, (iii) let $\overline{P}$ be the ν-th column of Γ ($\nu = 1, 2, \ldots, n(L-1)$); then, calculate Γ^{-1}.

Figure 2.1 illustrates the time-decomposition algorithm in the case $L = 2$. After calculating Γ^{-1} as mentioned above, we estimate $\overline{x(t_1)}$, the value of x at $t = t_1$, and solve sub-TPBVP for each subinterval. Then, using the difference $\overline{p(t_1^-)} - \overline{p(t_1^+)}$, the exact value $x(t_1)$ is determined by (2.20) and sub-TPBVP's are again solved with boundary conditions thus obtained. Each solution forms a part of the overall solution.

The overall computational algorithm will be summarized in the following section.

2.4. Summary of the Algorithm

We now summarize the above discussion.

Step 0-1: Set $h(t) = 0$, $t \in [t_0, t_f]$ in (2.1).

Step 0-2: Compute (2.1) in each subinterval with the boundary conditions

$$x(t_0) = 0, \quad [x'(t_1), x'(t_2), \ldots, x'(t_{L-1})]' = \overline{X} = e_\nu, \quad x(t_f) = 0, \qquad (2.32)$$

where e_ν is the ν-th unit vector ($\nu = 1, 2, \ldots, L-1$). Then, let the obtained difference $\overline{P} = [(\overline{p(t_1^-)} - \overline{p(t_1^+)})', (\overline{p(t_2^-)} - \overline{p(t_2^+)})', \ldots, (\overline{p(t_{L-1}^-)} -$

$p(t^{+}_{L-1}))']'$ be the ν-th column of the matrix Γ.

Step 0-3: Calculate Γ^{-1}.

Step 1: Estimate $\overline{x(t_i)}$ $(i = 1, 2, \ldots, L-1)$ arbitrarily.

Step 2: Compute the subinterval solution $\overline{p(t_i^{\pm})}$ to the problem (2.1) with the boundary conditions

$$x(t_0) = \pi_0, \quad x(t_i) = \overline{x(t_i)} \quad (i = 1, 2, \ldots, L-1), \quad x(t_f) = \pi_f. \qquad (2.33)$$

Step 3: From (2.20) obtain the exact value $x(t_i)$ at the torn time t_i $(i = 1, 2, \ldots, L-1)$.

Step 4: Solve (2.1) in each subinterval $[t_i, t_{i+1}]$ $(i = 0, 1, \ldots, L-1)$ with the boundary conditions obtained at *Step* 3 and $x(t_0) = \pi_0$ and $x(t_f) = \pi_f$.

2.5. Illustrative Examples

To show the effectiveness of the proposed time-decomposition algorithm, the following examples are solved by both the basic superposition principle and the time-decomposition algorithm, and the solutions are compared with the analytical ones. For the numerical integration of the differential equations, the fourth-order Runge-Kutta-Gill method is employed in double precision arithmetic with integration step size of 0.001.

Example 2.1 [62].

Let us consider first the following two-dimensional stiff linear TPBVP:

$$\frac{d}{dt}\begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} 1, & 1 \\ \kappa^2, & 1 \end{bmatrix}\begin{bmatrix} x \\ p \end{bmatrix}, \quad x(0) = 1, \quad p(t_f) = 0, \qquad (2.34)$$

where κ and t_f are positive constants. The eigenvalues of the system matrix are $1 \pm \kappa$. Therefore, if κ is large, the TPBVP has a stiff structure.

To the problem of (2.34), the analytical solution can be obtained as follows:

$$x(t) = \frac{1}{2}\{\exp[(1+\kappa)t] + \exp[(1-\kappa)t]\} - \frac{1}{2}\{\exp[(1+\kappa)t] - \exp[(1-\kappa)t]\}\cdot \{\exp[(1+\kappa)t_f] - \exp[(1-\kappa)t_f]\}/\{\exp[(1+\kappa)t_f] + \exp[(1-\kappa)t_f]\}, \quad (2.35)$$

$$p(t) = \frac{\kappa}{2}\{\exp[(1+\kappa)t] - \exp[(1-\kappa)t]\} - \frac{\kappa}{2}\{\exp[(1+\kappa)t] + \exp[(1-\kappa)t]\}\cdot \{\exp[(1+\kappa)t_f] - \exp[(1-\kappa)t_f]\}/\{\exp[(1+\kappa)t_f] + \exp[(1-\kappa)t_f]\}. \quad (2.36)$$

For comparison, the TPBVP with $t_f = 5.0$ and $\kappa = 5.0$ is solved by the two methods. In applying the time-decomposition algorithm, the interval is divided at $t = t_1 = 2.5$ into two subintervals. The results are shown in Table 1. Table 1 shows that the time-decomposition algorithm reduces the numerical error in the latter half of the integration interval while the basic superposition principle (No Time Decomposition) fails to follow the exact solution. In the following section, an analysis is made on this example.

Example 2.2 [51].

Second, let us consider the following TPBVP:

$$\frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ c_1 x - c_2 t^2 + 1 \end{bmatrix}, \quad x(0) = 0, \quad x(1) = 0.5, \quad (2.37)$$

where the constants c_1 and c_2 are related to each other by $c_1 = 2c_2$. Then, the analytical solution of (2.37) is given by

$$x(t) = 0.5t^2, \quad y(t) = t. \quad (2.38)$$

Table 2.1. The solutions of Example 2.1 ($t_f = 5.0, \kappa = 5.0$).

	No Time Decomposition		Time Decomposition ($L = 2$, torn time = 2.5)		Analytical Solution	
t	x	p	x	p	x	p
0.0	1.0	-0.50000E1	1.0	-0.50000E1	1.0	-5.0
0.5	0.13534E0	-0.67668E0	0.13534E0	-0.67668E0	0.13534E0	-0.67668E0
1.0	0.18316E-1	-0.91578E-1	0.18316E-1	-0.91578E-1	0.18316E-1	-0.91578E-1
1.5	0.24788E-2	-0.12394E-1	0.24788E-2	-0.12394E-1	0.24788E-2	-0.12394E-1
2.0	0.33546E-3	-0.16773E-2	0.33546E-3	-0.16773E-2	0.33546E-3	-0.16773E-2
2.5	0.45408E-4	-0.22696E-3	0.45396E-4	-0.22702E-3	0.45400E-4	-0.22700E-3
3.0	0.62981E-5	-0.29951E-4	0.61437E-5	-0.30718E-4	0.61442E-5	-0.30721E-4
3.5	0.39231E-5	0.11300E-4	0.83146E-6	-0.41573E-5	0.83153E-6	-0.41576E-5
4.0	0.62209E-4	0.30992E-3	0.11253E-6	-0.56260E-6	0.11254E-6	-0.56265E-6
4.5	0.12473E-2	0.62362E-2	0.15331E-7	-0.75630E-7	0.15333E-7	-0.75637E-7
5.0	0.25052E-1	0.12529E0	0.41216E-8	-0.1869E-11	0.41223E-8	0.0

The problem is solved for various values of c_1 and c_2. In Table 2.2, solutions obtained by the basic superposition principle and the time-decomposition algorithm with two-subintervals, and also the analytical solution are shown for $c_1 = 2,000$ and $c_2 = 1,000$. It is seen that the numerical error is much reduced by the time-decomposition algorithm. When $c_1 = 1,000$ and $c_2 = 500$, the superposition principle is effectively applied to the problem, though the result is not listed. However, when $c_1 = 2,000$ and $c_2 = 1,000$, the numerical error is accumulated and therefore the soltion gradually becomes less accurate as the independent variable, i.e., time, increases. When $c_1 = 3,000$ and $c_2 = 1,500$, even the time-decomposition algorithm with L (number of subintervals) = 2 suffers from the numerical error. Table 2.3 shows the results by the algorithm with $L \doteq 2$ and $L = 4$. Four-subinterval algorithm succeeds in obtaining a satisfactory solution. Finally, the algorithm is applied to the problem with $c_1 = 10,000$ and $c_2 = 5,000$, the eigenvalues of the system matrix are ± 100. The problem is successfully solved by the ten-subinterval algorithm. The result is shown in Table 2.4 together with the result by the four-subinterval algorithm.

2.6. Error Analysis

In this section, it is examined how the numerical error is reduced by the time-decomposition algorithm, by referring the problem of Example 2.1.

Let us reconsider the TPBVP (2.34). The transition matrix of (2.34) is given by

Table 2.2. The solutions of Example 2.2 ($c_1 = 2{,}000$ and $c_2 = 1{,}000$).

t	No·Time Decomposition		Time Decomposition ($L = 2$, torn time = 0.5)		Analytical Solution	
	x	y	x	y	x	y
0.0	0.0	0.9110169192E-9	0.0	0.9110170109E-9	0.0	0.0
0.1	0.5000000020E-2	0.9999999999E-1	0.5000000021E-2	0.9999999999E-1	0.005	0.1
0.2	0.2000000002E-1	0.2000000000E0	0.2000000002E-1	0.2000000000E0	0.02	0.2
0.3	0.4500000002E-1	0.3000000000E0	0.4500000002E-1	0.3000000000E0	0.045	0.3
0.4	0.8000000002E-1	0.4000000000E0	0.7999999997E-1	0.3999999976E0	0.08	0.4
0.5^-	0.1250000001E0	0.5000000045E0	0.1249999953E0	0.4999997895E0	0.125	0.5
0.5^+			0.1250000000E0	0.5000000000E0		
0.6	0.1800000088E0	0.6000003924E0	0.1800000000E0	0.6000000000E0	0.18	0.6
0.7	0.2450007683E0	0.7000343569E0	0.2450000000E0	0.6999999995E0	0.245	0.7
0.8	0.3200672549E0	0.8030077275E0	0.3199999999E0	0.7999999559E0	0.32	0.8
0.9	0.4108877234E0	0.1163306995E+1	0.4049999137E0	0.8999961379E0	0.405	0.9
1.0	0.1015431913E+1	0.2405081591E+2	0.4999924399E0	0.9996618993E0	0.5	1.0

Table 2.3. The solutions of Example 2.2 ($c_1 = 3{,}000$ and $c_2 = 1{,}500$).

t	Time Decomposition ($L = 2$, torn time = 0.5)		Time Decomposition ($L = 4$, torn times = 0.25, 0.5, 0.75)	
	x	y	x	y
0.0	0.0	0.1665186039E-8	0.0	0.1665186046E-8
0.1	0.5000000031E-2	0.9999999996E-1	0.5000000031E-2	0.9999999996E-1
0.2	0.2000000003E-1	0.2000000000E0	0.2000000003E-1	0.2000000000E0
0.25^-	0.3125000003E-1	0.2499999999E0	0.3125000003E-1	0.2499999999E0
0.25^+			0.3125000003E-1	0.2500000000E0
0.3	0.4500000003E-1	0.2999999997E0	0.4500000003E-1	0.3000000000E0
0.4	0.7999999890E-1	0.3999999380E0	0.8000000003E-1	0.3999999999E0
0.5^-	0.1249997294E0	0.4999851785E0	0.1249999998E0	0.4999999861E0
0.5^+	0.1250000000E0	0.5000000000E0	0.1250000000E0	0.4999999992E0
0.6	0.1800000000E0	0.5999999999E0	0.1800000000E0	0.5999999999E0
0.7	0.2449999999E0	0.6999999953E0	0.2449999999E0	0.6999999955E0
0.75^-	0.2812499987E0	0.7499999273E0	0.2812499988E0	0.7499999304E0
0.75^+			0.2812500000E0	0.7500000011E0
0.8	0.3199999795E0	0.7999988769E0	0.3200000000E0	0.8000000000E0
0.9	0.4049950985E0	0.8997313851E0	0.4050000000E0	0.8999999992E0
1.0	0.4988269998E0	0.9357521309E0	0.4999999968E0	0.9999998251E0

Table 2.4. The solutions of Example 2.2 ($c_1 = 10{,}000$ and $c_2 = 5{,}000$).

t	Time Decomposition ($L=4$, torn times = 0.25, 0.5, 0.75) x	y	Time Decomposition ($L=10$, torn times = 0.1, 0.2, ..., 0.9) x	y
0.0	0.0	0.9904520444E-8	0.0	0.9904520431E-8
0.1^-	0.5000000104E-2	0.9999999948E-1	0.5000000104E-2	0.9999999948E-1
0.1^+			0.5000000104E-2	0.9999999948E-1
0.2^-	0.2000000010E-1	0.1499999992E0	0.2000000010E-1	0.1999999994E0
0.2^+			0.2000000010E-1	0.1999999995E0
0.25^-	0.3124999974E-1	0.2499999631E0	0.3125000010E-1	0.2499999995E0
0.25^+	0.3125000010E-1	0.2499999995E0		
0.3^-	0.4500000010E-1	0.2999999995E0	0.4500000010E-1	0.2999999991E0
0.3^+			0.4500000010E-1	0.2999999995E0
0.4^-	0.7999999905E-1	0.3999998943E0	0.8000000009E-1	0.3999999985E0
0.4^+			0.8000000010E-1	0.3999999995E0
0.5^-	0.1249768380E0	0.4976837863E0	0.1250000001E0	0.4999999974E0
0.5^+	0.1249999238E0	0.5000076343E0	0.1250000001E0	0.4999999995E0
0.6^-	0.1800000001E0	0.5999999966E0	0.1800000001E0	0.5999999962E0
0.6^+			0.1800000001E0	0.5999999996E0
0.7^-	0.2449992945E0	0.6999294429E0	0.2450000001E0	0.6999999949E0
0.7^+			0.2450000001E0	0.6999999996E0
0.75^-	0.2811452852E0	0.7396285118E0	0.2812500001E0	0.7499999994E0
0.75^+	0.2812506867E0	0.7499313405E0		
0.8^-	0.3200000047E0	0.7999995368E0	0.3200000000E0	0.7999999934E0
0.8^+			0.3200000001E0	0.7999999997E0
0.9^-	0.4049999893E0	0.8999989213E0	0.4050000000E0	0.8999999916E0
0.9^+			0.4050000001E0	0.8999999997E0
1.0	0.4997625140E0	0.9762513929E0	0.4999999999E0	0.9999999791E0

$$\underline{\Phi}(t, 0) = \begin{bmatrix} \underline{\phi}_{11}(t, 0), & \underline{\phi}_{12}(t, 0) \\ \underline{\phi}_{21}(t, 0), & \underline{\phi}_{22}(t, 0) \end{bmatrix}$$

$$= \frac{1}{2}\begin{bmatrix} \{\exp[(1+\kappa)t] + \exp[(1-\kappa)t]\}, & \frac{1}{\kappa}\{\exp[(1+\kappa)t] - \exp[(1-\kappa)t]\} \\ \kappa\{\exp[(1+\kappa)t] - \exp[(1-\kappa)t]\}, & \{\exp[(1+\kappa)t] + \exp[(1-\kappa)t]\} \end{bmatrix}. \quad (2.39)$$

In the following, the quantities without numerical errors are presented with the underline. For example, $\underline{\Phi}(t, 0)$ denotes the analytical transition matrix while $\Phi(t, 0)$ denotes the one obtained by a numerical procedure. By using (2.39), the general solution can be written as

$$x(t) = \underline{\phi}_{11}(t, 0)x(0) + \underline{\phi}_{12}(t, 0)p(0), \quad (2.40)$$

$$p(t) = \underline{\phi}_{21}(t, 0)x(0) + \underline{\phi}_{22}(t, 0)p(0). \quad (2.41)$$

First, we consider the numerical error in the superposition principle. Let ${}^1p_f$ be the terminal value of p when (2.34) is solved with $x(0) = \pi_0$ and $p(0) = {}^1p_0$. Then, by Lemma 2.1, the exact initial value $\underline{p}_0$ of p is given by

$$\underline{p}_0 = {}^1p_0 - \underline{\phi}_{22}^{-1}(t_f, 0)\,{}^1p_f. \quad (2.42)$$

The exact value of $\underline{p}_0$ can be determined by (2.42). However, a certain error is contained in the numerical solution procedure, such as round-off and trancation errors. The numerical error may occur in the following three procedures: (i) The procedure of the integration of (2.34), that is, so called formula error of the integration routine, (ii) the procedure of calculating (2.42), and (iii) the procedure of taking the inverse of $\underline{\phi}_{22}(t_f, 0)$. In this section, we assume that the numerical error is due only to (iii).

Denote the numerically obtained transition matrix by $\Phi(t, 0)$ and put

$$\phi_{22}^{-1}(t_f, 0) = \underline{\phi}_{22}^{-1}(t_f, 0) + \varepsilon_1. \qquad (2.43)$$

Then, the initial value ${}^2p_0$ obtained by the superposition principle is

$$\begin{aligned} {}^2p_0 &= {}^1p_0 - \phi_{22}^{-1}(t_f, 0)\,{}^1p_f \\ &= \underline{p}_0 - \varepsilon_1\,{}^1p_f. \end{aligned} \qquad (2.44)$$

Define $E(t)$ by

$$E(t) = \{[x(t) - \underline{x}(t)]^2 + [p(t) - \underline{p}(t)]^2\}^{1/2} \qquad (2.45)$$

where $\underline{x}(t)$ and $\underline{p}(t)$ are the exact solutions and $x(t)$ and $p(t)$ are the numerical solutions with the initial conditions $x(0) = \pi_0$ and $p(0) = {}^2p_0$. Then,

$$\begin{aligned} E^2(t) &= [\underline{\phi}_{12}(t, 0)\varepsilon_1\,{}^1p_f]^2 + [\underline{\phi}_{22}(t, 0)\varepsilon_1\,{}^1p_f]^2 \\ &= \Big[\frac{1}{4\kappa^2}\{\exp[2(1+\kappa)t] + \exp[2(1-\kappa)t] - 2\exp(2t)\} + \\ &\quad \frac{1}{4}\{\exp[2(1+\kappa)t] + \exp[2(1-\kappa)t] + 2\exp(2t)\}\Big]\varepsilon_1^2\,{}^1p_f^2 \\ &\cong \frac{(\kappa^2+1)}{4\kappa^2}\varepsilon_1^2\,{}^1p_f^2\exp[2(1+\kappa)t]. \end{aligned} \qquad (2.46)$$

For a special case, let ${}^1p_0 = 0$. Then,

$${}^1p_f = \underline{\phi}_{21}(t_f, 0)\pi_0$$

and

$$E(t) \cong \frac{\sqrt{\kappa^2+1}}{4}|\varepsilon_1||\pi_0|\exp[(1+\kappa)(t_f+t)]. \qquad (2.47)$$

Second, let us consider the numerical error in the time-decomposition algorithm. We consider the case where the overall interval is divided into two subintervals at $t = t_1 = t_f/2$. We write

$$\phi_{12}^{-1}(t_1, 0) = \underline{\phi}_{12}^{-1}(t_1, 0) + \varepsilon_2,$$

$$\phi_{22}^{-1}(t_f, t_1) = \underline{\phi}_{22}^{-1}(t_f, t_1) + \varepsilon_3, \tag{2.48}$$

$$\Gamma^{-1} = \underline{\Gamma}^{-1} + \varepsilon_4.$$

Now we follow the time-decomposition algorithm.

1) We solve (2.34) with $x(0) = \pi_0$, $x(t_1) = \overline{x(t_1)} = 0$ in the subinterval I: $[0, t_1]$, and with $x(t_1) = \overline{x(t_1)}, p(t_f) = 0$ in the subinterval II: $[t_1, t_f]$.

Subinterval I:

Let ${}^1p_0 = 0$, then $x(t_1) = \underline{\phi}_{11}(t_1, 0)\pi_0$. Let $\underline{p}_1$ be the initial value with which the solution to (2.34) is to hit $x(t_1) = \overline{x(t_1)}$. Then,

$$\bar{p}_1 = \underline{\bar{p}}_1 - \varepsilon_2\underline{\phi}_{11}(t_1, 0)\pi_0$$

and

$$\overline{p(t_1^-)} = \overline{\underline{p}(t_1^-)} - \varepsilon_2\underline{\phi}_{22}(t_1, 0)\underline{\phi}_{11}(t_1, 0)\pi_0 \tag{2.49}$$

Subinterval II:

Similarly,

$$\overline{p(t_1^+)} = \overline{\underline{p}(t_1^+)} = 0 \tag{2.50}$$

2) Next, we determine the boundary value $x(t_1)$ of x by (2.20) of Corollary 2.1, that is,

$$x_1 \triangleq x(t_1) = \overline{x(t_1)} - \Gamma^{-1}[\underline{\bar{P}} - \varepsilon_2\underline{\phi}_{22}(t_1, 0)\underline{\phi}_{11}(t_1, 0)\pi_0] = \underline{x}_1 - \delta \tag{2.51}$$

where $\underline{P} = \overline{p(t_1^-)} - \overline{p(t_1^+)}$ and

$$\delta = (\underline{\Gamma}^{-1} + \varepsilon_4)[-\varepsilon_2 \underline{\Phi}_{22}(t_1, 0)\underline{\Phi}_{11}(t_1, 0)\pi_0] + \varepsilon_4 \underline{P}. \tag{2.52}$$

Substituting

$$\begin{aligned} \underline{P} &= \underline{\Phi}_{21}(t_1, 0)\pi_0 - \underline{\Phi}_{22}(t_1, 0)\underline{\Phi}_{12}^{-1}(t_1, 0)\underline{\Phi}_{11}(t_1, 0)\pi_0 , \\ \underline{\Gamma}^{-1} &= \underline{\Phi}_{12}(t_1, 0)\underline{\Phi}_{22}^{-1}(t_f, 0)\underline{\Phi}_{22}(t_f, t_1) \end{aligned} \tag{2.53}$$

into (2.52), we obtain

$$\begin{aligned} \delta = \{&-\varepsilon_2[\underline{\Phi}_{12}(t_1, 0)\underline{\Phi}_{22}^{-1}(t_f, 0)\underline{\Phi}_{22}(t_f, t_1) + \varepsilon_4]\underline{\Phi}_{22}(t_1, 0)\underline{\Phi}_{11}(t_1, 0) + \\ &\varepsilon_4[\underline{\Phi}_{21}(t_1, 0) - \underline{\Phi}_{22}(t_1, 0)\underline{\Phi}_{12}^{-1}(t_1, 0)\underline{\Phi}_{11}(t_1, 0)]\}\pi_0 . \end{aligned} \tag{2.54}$$

3) Finally, we solve (2.34) with $x(0) = \pi_0$, $x(t_1) = x_1$ in the subinterval I: $[0, t_1]$, and with $x(t_1) = x_1$, $p(t_f) = 0$ in the subinterval II: $[t_1, t_f]$. The initial estimate of the missing condition in each subinterval is set to zero. The analytical value $\underline{x}_1$ is given by

$$\underline{x}_1 = \underline{\Phi}_{11}(t_1, 0)\pi_0 - \underline{\Phi}_{12}(t_1, 0)\underline{\Phi}_{22}^{-1}(t_f, 0)\underline{\Phi}_{21}(t_f, 0)\pi_0 . \tag{2.55}$$

I) The calculated value p_0 of $p(0)$ with which the solution to (2.34) is to hit x_1 is

$$\begin{aligned} p_0 &= \Phi_{12}^{-1}(t_1, 0)[x_1 - \underline{\Phi}_{11}(t_1, 0)\pi_0] \\ &= [\underline{\Phi}_{12}^{-1}(t_1, 0) + \varepsilon_2][\underline{x}_1 - \delta - \underline{\Phi}_{11}(t_1, 0)\pi_0] \\ &= \underline{p}_0 + \varepsilon_2[\underline{x}_1 - \underline{\Phi}_{11}(t_1, 0)\pi_0 - \delta] - \delta\underline{\Phi}_{12}^{-1}(t_1, 0) \\ &= \underline{p}_0 - \delta_1 , \end{aligned} \tag{2.56}$$

where

$$\delta_1 = \varepsilon_2[\underline{\phi}_{12}(t_1, 0)\underline{\phi}_{22}^{-1}(t_f, 0)\underline{\phi}_{21}(t_f, 0)\pi_0 + \delta] + \delta\underline{\phi}_{12}^{-1}(t_1, 0) . \quad (2.57)$$

II) Similarly,

$$\begin{aligned} p(t_1^+) &= -\phi_{22}^{-1}(t_f, t_1)\underline{\phi}_{21}(t_f, t_1)x_1 \\ &= -[\underline{\phi}_{22}^{-1}(t_f, t_1) + \varepsilon_3]\underline{\phi}_{21}(t_f, t_1)(\underline{x}_1 - \delta) \\ &= \underline{p}(t_1^+) - \varepsilon_3\underline{\phi}_{21}(t_f, t_1)\underline{x}_1 + [\underline{\phi}_{22}^{-1}(t_f, t_1) + \varepsilon_3]\underline{\phi}_{21}(t_f, t_1)\delta \\ &= \underline{p}(t_1^+) - \delta_2 , \end{aligned} \quad (2.58)$$

where

$$\begin{aligned} \delta_2 &= \varepsilon_3\underline{\phi}_{21}(t_f, t_1)\underline{x}_1 - [\underline{\phi}_{22}^{-1}(t_f, t_1) + \varepsilon_3]\underline{\phi}_{21}(t_f, t_1)\delta \\ &= \varepsilon_3\underline{\phi}_{21}(t_f, t_1)[\underline{\phi}_{11}(t_1, 0) - \underline{\phi}_{12}(t_1, 0)\underline{\phi}_{22}^{-1}(t_f, 0)\underline{\phi}_{21}(t_f, 0)]\pi_0 - \\ &\quad [\underline{\phi}_{22}^{-1}(t_f, t_1) + \varepsilon_3]\underline{\phi}_{21}(t_f, t_1)\delta . \end{aligned} \quad (2.59)$$

Now we evaluate (2.54), (2.57), and (2.59). Since, for sufficiently large κ and t,

$$\Phi(t, 0) \cong \frac{1}{2} \exp[(1+\kappa)t] \begin{bmatrix} 1, & \frac{1}{\kappa} \\ \kappa, & 1 \end{bmatrix} \quad (2.60)$$

the following relations are obtained:

$$\delta \cong \Big[-\varepsilon_2(\frac{1}{2\kappa} + \varepsilon_4)\frac{1}{4}\exp[(2(1+\kappa)t_1] + \varepsilon_4\{\frac{1}{2}\exp[(1+\kappa)t_1]$$

$$- \frac{1}{2}\exp[(1+\kappa)t_1]2\kappa\exp[-(1+\kappa)t_1]\frac{1}{2}\exp[(1+\kappa)t_1]\}\Big]\pi_0$$

$$\cong -\frac{\varepsilon_2}{8\kappa}\exp[2(1+\kappa)t_1]\pi_0\,, \tag{2.61}$$

$$\delta_1 \cong \varepsilon_2\frac{1}{2\kappa}\exp[(1+\kappa)t_1]\,2\exp[-(1+\kappa)t_f]\frac{\kappa}{2}\exp[(1+\kappa)t_f]\pi_0 +$$

$$\{\varepsilon_2 + 2\kappa\exp[-(1+\kappa)t_1]\}\delta$$

$$= \frac{\varepsilon_2}{4}\exp[(1+\kappa)t_1]\pi_0 - \frac{\varepsilon_2^2}{8\kappa}\exp[2(1+\kappa)t_1]\pi_0$$

$$\cong \frac{\varepsilon_2}{4}\exp[(1+\kappa)t_1]\pi_0\,, \tag{2.62}$$

$$\delta_2 \cong \Big[\varepsilon_3\frac{\kappa}{2}\exp[(1+\kappa)(t_f - t_1)]\{\frac{1}{2}\exp[(1+\kappa)t_1] - \frac{1}{2\kappa}\exp[(1+\kappa)t_1]\cdot$$

$$2\exp[-(1+\kappa)t_f]\frac{\kappa}{2}\exp[(1+\kappa)t_f]\} + \{2\exp[-(1+\kappa)(t_f - t_1)] + \varepsilon_3\}\cdot$$

$$\frac{\kappa}{2}\exp[(1+\kappa)(t_f - t_1)]\frac{\varepsilon_2}{8\kappa}\exp[2(1+\kappa)t_1]\Big]\pi_0$$

$$= \{2\exp[-(1+\kappa)(t_f - t_1)] + \varepsilon_3\}\frac{\varepsilon_2}{16}\exp[(1+\kappa)(t_f + t_1)]\pi_0$$

$$\cong \frac{\varepsilon_2}{8}\exp[2(1+\kappa)t_1]\pi_0\,. \tag{2.63}$$

We define $E_{TD}(t)$ by (2.45) for the error estimation of the time-decomposition algorithm. Then,

I) For $0 < t \leq t_1$,

$$E_{TD}^2(t) = [\Phi_{12}^2(t, 0) + \Phi_{22}^2(t, 0)]\delta_1^2$$

$$\cong \frac{\kappa^2+1}{4\kappa^2}\exp[2(1+\kappa)t]\frac{\varepsilon_2^2}{16}\exp[2(1+\kappa)t_1]\pi_0^2 \tag{2.64}$$

and therefore

$$E_{TD}(t) \cong \frac{\sqrt{\kappa^2+1}}{8\kappa}|\varepsilon_2||\pi_0|\exp[(1+\kappa)(t + t_1)]\,. \tag{2.65}$$

II) For $t_1 \leqq t \leqq t_f$,

$$E_{TD}^2(t) = \{[\Phi_{11}(t, t_1)\delta + \Phi_{12}(t, t_1)\delta_2]^2 + [\Phi_{21}(t, t_1)\delta + \Phi_{22}(t, t_1)\delta_2]^2\}$$

$$= [\frac{\kappa^2+1}{4}\delta^2 + \frac{\kappa^2+1}{4\kappa^2}\delta_2^2 + 2\frac{\kappa^2+1}{4\kappa}\delta\delta_2]\exp[2(1+\kappa)(t - t_1)]$$

$$+ [\frac{1-\kappa^2}{2}\delta^2 - \frac{1-\kappa^2}{2\kappa^2}\delta_2^2] + [\frac{\kappa^2+1}{4}\delta^2 + \frac{\kappa^2+1}{4\kappa^2}\delta^2 - 2\frac{\kappa^2+1}{4\kappa}\delta\delta_2]\cdot$$

$$\exp[2(1-\kappa)(t - t_1)]. \qquad (2.66)$$

Since, from (2.61) and (2.63),

$$\frac{\kappa^2+1}{4}\delta^2 + \frac{\kappa^2+1}{4\kappa^2}\delta_2^2 \pm \frac{\kappa^2+1}{2\kappa}\delta\delta_2 = \frac{\kappa^2+1}{4\kappa^2}(\kappa\delta \pm \delta_2)^2$$

$$= \begin{cases} 0 \\ \frac{\kappa^2+1}{64\kappa^2}\varepsilon_2^2\exp[4(1+\kappa)t_1] \end{cases} \qquad (2.67)$$

and

$$\frac{1-\kappa^2}{2}\delta^2 - \frac{1-\kappa^2}{2\kappa^2}\delta_2^2 = \frac{1-\kappa^2}{2\kappa^2}[(\kappa\delta)^2 - \delta_2^2] = 0, \qquad (2.68)$$

we obtain

$$E_{TD}^2(t) = \frac{\kappa^2+1}{64\kappa^2}\varepsilon_2^2\pi_0^2\exp[4(1+\kappa)t_1]\exp[2(1-\kappa)(t - t_1)], \qquad (2.69)$$

thus,

$$E_{TD}(t) = \frac{\sqrt{\kappa^2+1}}{8\kappa}|\varepsilon_2||\pi_0|\exp[(1-\kappa)t + (1+3\kappa)t_1]. \qquad (2.70)$$

Now we examine the above mentioned result by using the result of Example 2.1. The exact values $\underline{x}_1$ and $\underline{p}(t_1^+)$, the numerical values x_1 and $p(t_1^+)$, and δ and δ_2, are

$$\underline{x}_1 = 0.4539992976312E-4\,, \quad \underline{p}(t_1^+) = 0.2269996488093E-3\,,$$

$$x_1 = 0.4539603683907E-4\,, \quad p(t_1^+) = 0.2269801841891E-3\,, \quad (2.71)$$

$$\delta = -0.389292405E-8\,, \quad \delta_2 = 0.194646202E-7.$$

(2.71) shows that the relation $\delta_2 = -\kappa\delta$ is nearly satisfied by the numerical solution. Substituting (2.71) into (2.61), we obtain

$$\varepsilon_2 = -0.1457141E-19 \quad (2.72)$$

Figure 2.2 shows the results discussed above. In the figure, the dot represents the error norm of the solution by the superposition principle and the cross represents the one by the time-decomposition algorithm. The exact value of $E_{TD}(t)$ is also shown by the solid line. It is seen that the numerical error is much reduced by the time-decomposition algorithm and that the numerical error is analyzed effectively by the above discussion.

2.7. Concluding Remarks

In this chapter, the time-decomposition algorithm is proposed for the solution of a stiff linear TPBVP. The algorithm belongs to the multipoint approach method. The algorithm first divides the prescribed overall interval into several subintervals. Then, assuming the values of some of variables at torn times, sub-TPBVP's are solved in each subinterval. Finally, the assumed values are corrected so as to ensure the continuity of the solutions at the torn times. It is shown that the algorithm can be applied to the problem, so long as it has a unique solution in each subinterval.

The algorithm succeeds in overcoming the difficulty which the superposition principle encounters. That is to say, the missolution caused by the stiffness of the given problem and/or the length of the integration interval is much reduced by the proposed algorithm.

In the discussion of this chapter, only the linear TPBVP is dealt with. This may seem a serious limitation of the time-decomposition algorithm. However, since the nonlinear TPBVP can be reduced to a sequence of linear ones by the interaction-coordination algorithm and the quasilinearization method, the limitation is not so serious. The combined algorithm will be used in the subsequent chapter.

In Ref. 71, Roberts and Shipman have pointed out important questions to be resolved for the multipoint approach method: that is, (i) How many multipoints? (ii) Where shall the multipoints be specified? (iii) How shall the initial values at the multipoint be selected? They have suggested that (i) the number of multipoints should be as few as possible to make the size of matrix to be inverted as small as possible, (ii) the points should be specified near the region of numerical instability if the problem has round-off or instability difficulties, and (iii) the easiest way to choose the initial trial values for the internal points is to select the missing initial conditions at t_0 and integrate forward until t_f. They have treated the nonlinear TPBVP as it is and they must solve nonlinear sub-problems by the shooting method. Therefore, the selection of the multipoints and the initial conditions has significant influence on the convergence of the algorithm.

In use of the proposed time-decomposition algorithm, we may answer to the above question as follows: (i) The number of multipoints (torn times) had better be as few as possible. However, since the size of matrix to be inverted is half as large as that of Roberts and Shipman, more points may be specified. For the determination of the maximum length Δt of subintervals, Meile et al. [47] have proposed to make $\exp[\lambda_{max}\Delta t]$ sufficiently small, where λ_{max} is the maximum eigenvalue of the system matrix. If Δt is too large, the solution obtained may contain serious discontinuities at torn times. (ii) and (iii) Since in each subinterval the superposition principle is applied, the selection of the multipoints and the initial conditions is not a so serious matter, so long as the length of each subinterval is suitable. When the initial condition is chosen to be near the exact solution, of course, the numerical error is much reduced as we have seen in Section 2.6. Thus, the proposed algorithm is applied much easily to TPBVP's than the algorithm of Roberts and Shipman.

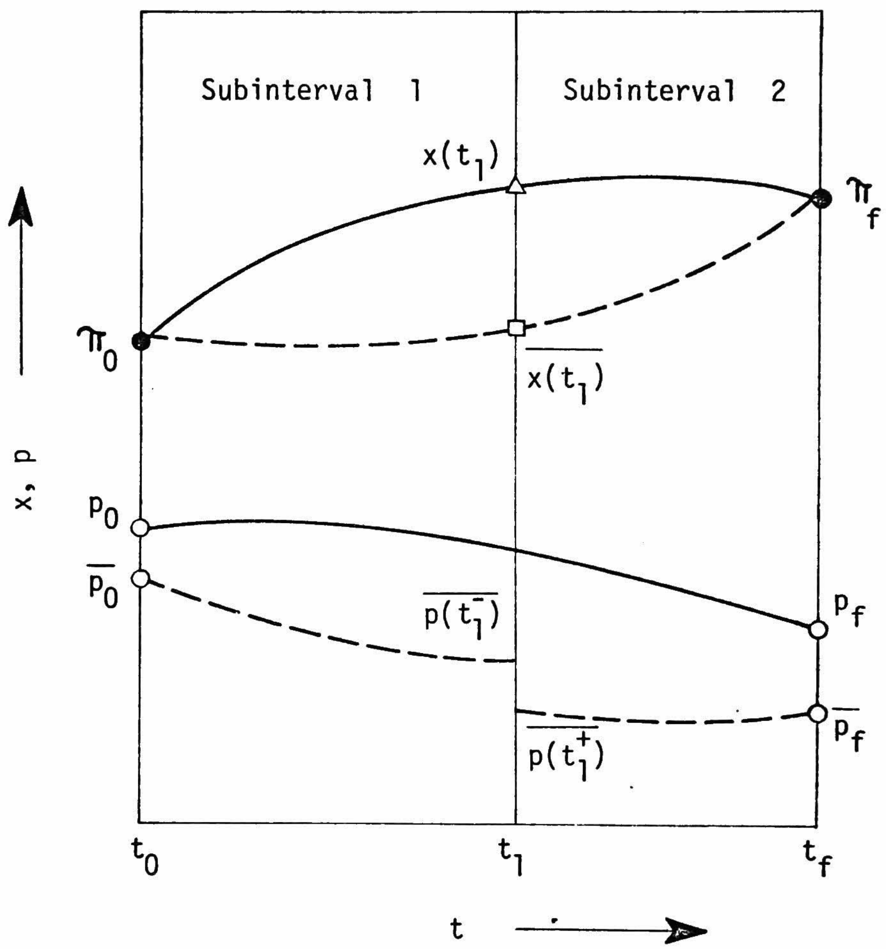

Fig. 2.1. Solutions by time decomposition.

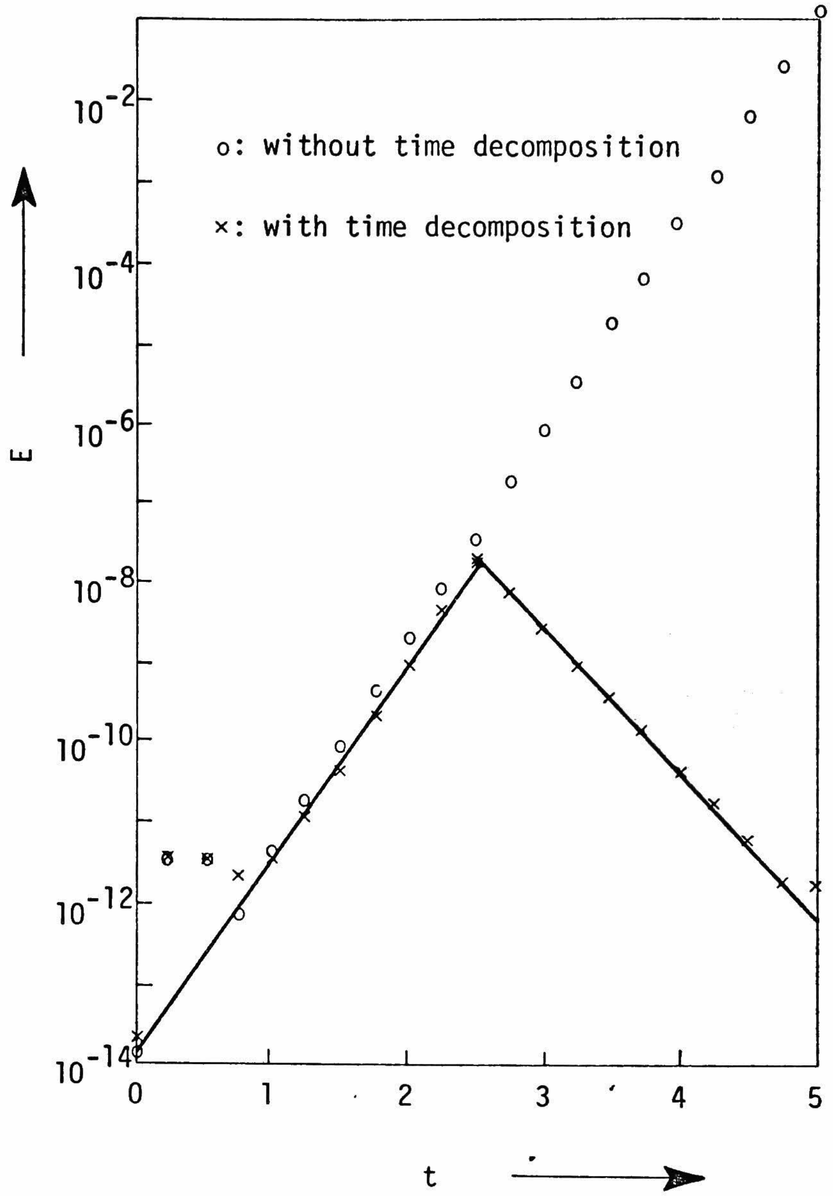

Fig. 2.2. Dependence of error norms of the numerical solutions on time t.

Chapter 3

Solution of Nonlinear Optimal Control Problems by Use of the Time-Decomposition Algorithm

3.1. Introduction

The optimal control of physical systems has been a matter of central concern in the control problem and many workers have developed algorithms for the solution of the problem of various types. Since, except for a minority of current engineering problems, it is difficult to solve such a problem by pure analytical methods [12, 36, 39, 44], the practical interest has been layed on the development of numerical techniques which enable us to solve such a problem on a digital computer.

This chapter is concerned with showing the application of the time-decomposition algorithm developed in the preceding chapter to the solution of nonlinear optimal control problems in conjunction with the linearization method such as the quasilinearization method [4, 42] and the interaction-coordination algorithm [55, 61, 63].

The system considered in this chapter is described by a set of non-linear differential equations and the objective is to minimize a performance index of a quadratic type in state and control. For such an approach based on the state-space formulation, the perfect information about the system to be controlled must be available and this is one of reasons why some researchers have claimded that the transfer function method which can be used only with a partial information about the system is superior to the state-space approach [29]. There are, however, many problems to which the state-space approach can effectively be applied. Especially in the field of aerospace engineering, the approach is effective by nature of the problem. Also in other fields, many workers have made efforts in the application of the approach [53].

We consider in this chapter the following nonlinear optimal control problem. The dynamical system is governed by the differential equation:

$$\dot{x} = A(t)x + B(t)u + f(t, x) , \tag{3.1}$$

where $x(t)$ is an n-dimensional state vector, $u(t)$ is an m-dimensional control vector, and $A(t)$ and $B(t)$ are matrices of compatible order. f is an n-dimensional nonlinear function of C^2 with respect to x. Both cases are considered that A, B, and f are continuous in t and that they are piecewise continuous. The objective is to transfer the state of the system from a certain manifold to another manifold so as to minimize an associated performance index:

$$J = \frac{1}{2}\int_{t_0}^{t_f} [x'Q(t)x + u'R(t)u]dt , \tag{3.2}$$

where $Q(t)$ and $R(t)$ are positive semidefinite and positive definite matrices, respectively, of compatible order. The control function u is assumed not to be subject to any constraint such as (1.4).

For the solution of the problem, many algorithms have been proposed. They can be classified into two categories: (i) the direct method [8-11, 15, 34, 35, 46, 48, 50, 65] and (ii) the indirect method [4, 6, 16, 17, 41, 42, 55, 63, 73]. The direct method generates a sequence of control functions so that the performance index may successively be optimized. The control function is generated by correcting the preceding control function by using the solution of the adjoint equation. For this approach, main questions are how to choose the initial approximation of the control function and how to correct it.

On the other hand, the indirect method reduces the problem into a two-point boundary-value problem (TPBVP) by applying the minimum principle or the variational principle and the optimal control is determined by its solution. Therefore, for the indirect method, the main question is how to solve nonlinear TPBVP's.

For the solution of nonlinear TPBVP's, many algorithms have been proposed, such as the shooting method [2, 40, 64, 70, 72], the invariant imbedding [5, 23], the quasilinearization method (or the generalized Newton-Raphson method)[4, 42, 49], and the interaction-coordination algorithm [55, 61, 63]. The former two methods treat the nonlinear problem as it is, while the latter two methods reduces the nonlinear problem into a sequence of linear problems. The TPBVP derived by the minimum principle often has a stiff structure. The numerical error affects much the convergence characteristics of the linearization method for such a problem. Therefore, in this chapter, the nonlinear TPBVP is solved by the method, i.e., the quasilinearization method and the interaction-coordination algorithm, with the additional use of the time-decomposition algorithm.

In the former half of this chapter, the solution of a nonlinear regulator problem with continuous quantities is considered. That is to say, matrices A and B, and the vector function f are assumed to be continuous in t. The derived TPBVP is linearized by the quasilinearization method or the interaction-coordination algorithm. In Refs. 55 and 63, we have shown that even the TPBVP with strong nonlinearities can be solved by the interaction-coordination algorithm, by choosing appropriate values of parameters, called weights and that by experience one must choose a large value for one of the weights to attain fast convergence for such a problem. However, since the reduced TPBVP's have a stiff structure with large values of weights, there is a limitation to the available values of weights and therefore, to the convergence range of the algorithm. The quasilinearization method also fails to obtain the solution of the problem with Jacobian matrix characterized by large positive and large negative eigenvalues [47]. We show in this chapter that the convergence characteristics of these methods are far improved by the additional use of the time-decomposition algorithm.

In the latter half of this chapter, the case is considered that matrices A and B, and the vector function f contain discontinuities at several intermediate points of the overall control interval and moreover some of the state variables may be specified at these points. These points are called 'corner times.' Such a problem is called an optimal control problem with discontinuities [30, 32, 75] or a multiple-target problem [56, 57, 74, 76] because of the additional conditions on the state.

By applying the minimum principle, we obtain a nonlinear multipoint boundary-value problem (MPBVP) as the necessary condition for optimality. Since the corresponding solution to the adjoint equation is discontinuous at the points where the state variables are specified, it is impossible to integrate the given differential equations in series. Therefore, decomposition of the overall interval into several subintervals at several points is essential to the solution of such a problem. These points are called 'torn times.' The set of torn times includes all of the corner times.

The time-decomposition algorithm decomposes the overall interval into several subintervals. The boundary values at torn times are initially chosen arbitrarily and then corrected so as to make the solutions compatible at torn times. Then, the algorithm is effectively applied to the MPBVP in conjunction with the linearization method. The specified values of the state variables at the corner times are chosen as the initial estimates and the correction is made except for these variables. Even to the case that the corner times are also the parameters to be optimized, the time-decomposition algorithm can be applied. The corner times are first assumed arbitrarily and the algorithm is applied together with the linearization method. Then, their optimal corrections are made by a gradient method.

In Section 3.2, a nonlinear regulator problem of continuous type is discussed. In Section 3.2.1, the problem is formulated and the TPBVP is derived to obtain the optimal control. In Section 3.2.2, two linearization methods, the interaction-coordination algorithm and the quasilinearization method are briefly sketched and the applicability of the time-decomposition algorithm to the TPBVP linearized by the former algorithm is discussed. The combined algorithm is applied to two physical problems in Section 3.2.3.

Following the continuous case, the discontinuous case, the multiple-target problem is discussed in Section 3.3. The problem considered is stated and the necessary condition for optimality is derived in Section 3.3.1. Then, in Section 3.3.2, the time-decomposition algorithm of Chapter 2 is modified to solve the derived MPBVP. In Section 3.3.3, two linear problems and a nonlinear problem are solved by the proposed algorithm and the solutions to linear problems are compared with the analytical one.

3.2. A Nonlinear Optimal Control Problem

3.2.1. Problem Statement

Consider the nonlinear control system defined by (3.1). The objective is to control the system so as to minimize the performance index, starting the initial state $x(t_0) = \pi_0$, where π_0 is a prescribed constant vector. The initial time t_0 and the terminal time t_f are assumed to be specified. The terminal condition on x may be or may not be specified, but, for simplicity, we assume it is specified as $x(t_f) = \pi_f$ in the following formalization. Both cases are treated in examples.

Define the Hamiltonian H of (3.1) and (3.2) as follows:

$$H = \frac{1}{2}(x'Qx + u'Ru) + p'(Ax + Bu + f), \tag{3.3}$$

where p is an n-dimensional costate vector. Then, due to the minimum principle, a necessary condition for optimality is derived as follows:

$$\frac{dx}{dt} = (\frac{\partial H}{\partial p})' = A(t)x + B(t)u + f(t, x) , \tag{3.4.1}$$

$$\frac{dp}{dt} = -(\frac{\partial H}{\partial x})' = -Q(t)x - A'(t)p - (\frac{\partial f}{\partial x})'p , \tag{3.4.2}$$

$$(\frac{\partial H}{\partial u})' = R(t)u + B'(t)p = 0, \tag{3.4.3}$$

$$x(t_0) = \pi_0 , \quad x(t_f) = \pi_f . \tag{3.5}$$

Therefore, substituting $u = -R^{-1}B'p$ into (3.4.1), we obtain the following nonlinear two-point boundary-value problem:

$$\frac{d}{dt}\begin{bmatrix} x \\ p \end{bmatrix} = D(t)\begin{bmatrix} x \\ p \end{bmatrix} + h(t, x, p) \tag{3.6}$$

with the boundary condition (3.5), where,

$$D(t) = \begin{bmatrix} A(t), & -B(t)R^{-1}(t)B'(t) \\ -Q(t), & -A'(t) \end{bmatrix} , \tag{3.7}$$

$$h(t, x, p) = \begin{bmatrix} h_1(t, x, p) \\ h_2(t, x, p) \end{bmatrix} = \begin{bmatrix} f(t, x) \\ -(\frac{\partial f}{\partial x})'p \end{bmatrix} , \tag{3.8}$$

and $h_i(t, x, p)$ is an n-dimensional vector function ($i = 1, 2$).

3.2.2. Linearization of the Nonlinear TPBVP

Since it is usually impossible to solve the nonlinear TPBVP analytically, one must resort to numerical techniques for the solution. In this section, two linearization methods, the quasilinearization method [42] and the interaction-coordination alforithm [55, 63] are briefly sketched in view of the additional use of the time-decomposition algorithm.

(a) the quasilinearization method

Let $({}^{k-1}x', {}^{k-1}p')'$ be $(k-1)$-th approximation to the solution of (3.5) and (3.6). Then, first-order Taylor series expansion of (3.6) about it is

$$\frac{d}{dt}\begin{bmatrix} {}^{k}x \\ {}^{k}p \end{bmatrix} = D(t)\begin{bmatrix} {}^{k}x \\ {}^{k}p \end{bmatrix} + h(t, {}^{k}x, {}^{k}p)$$

$$= [D(t) + [\frac{\partial h(t, {}^{k-1}x, {}^{k-1}p)}{\partial x}, \frac{\partial h(t, {}^{k-1}x, {}^{k-1}p)}{\partial p}]]\begin{bmatrix} {}^{k}x - {}^{k-1}x \\ {}^{k}p - {}^{k-1}p \end{bmatrix} +$$

$$D(t)\begin{bmatrix} {}^{k-1}x \\ {}^{k-1}p \end{bmatrix} + h(t, {}^{k-1}x, {}^{k-1}p), \qquad (3.9)$$

$${}^{k}x(t_0) = \pi_0, \quad {}^{k}x(t_f) = \pi_f.$$

where $({}^{k}x', {}^{k}p')'$ is the value of the k-th iteration. Therefore, the solution of (3.5) and (3.6) is obtained by the following algorithm:

Step 1: Assume nominal values ${}^{1}x$ and ${}^{1}p$ of x and p, respectively, which satisfy $x(t_0) = \pi_0$ and $x(t_f) = \pi_f$. Set $k = 2$.

Step 2: Compute ${}^{k}x$ and ${}^{k}p$ by solving (3.9).

Step 3: If ${}^{k}x$ and ${}^{k}p$ are sufficiently close to ${}^{k-1}x$ and ${}^{k-1}p$, respectively, that is, if ${}^{k}G$ defined by

$${}^{k}G = \{\int_{t_0}^{t_f} [({}^{k}x - {}^{k-1}x)^2 + ({}^{k}p - {}^{k-1}p)^2]dt\}^{1/2} \qquad (3.10)$$

is sufficiently small, the calculation is terminated. Otherwise, replace ${}^{k-1}x$ and ${}^{k-1}p$ by ${}^{k}x$ and ${}^{k}p$, respectively, and replace k by k+1. Then, return to *Step* 2.

The above algorithm is called the quasilinearization method or the generalized Newton-Raphson method [4, 42].

(b) the interaction-coordination algorithm

Next, we consider another linearization method. We introduce scalar parameters β and κ, called weights, and vector variables $y(t)$ and $q(t)$, called interaction variables. Using these, we modify (3.6) as follows:

$$\frac{d}{dt}\begin{bmatrix} x \\ p \end{bmatrix} = D^*(t)\begin{bmatrix} x \\ p \end{bmatrix} + [D(t) - D^*(t)]\begin{bmatrix} y \\ q \end{bmatrix} + h(t, y, q) \ , \qquad (3.11)$$

where

$$D^*(t) = \begin{bmatrix} A(t), & -\beta B(t)R^{-1}(t)B'(t) \\ -\kappa Q(t), & -A'(t) \end{bmatrix}. \qquad (3.12)$$

Then, for the solutions x and p of (3.11) with the boundary condition (3.5) to be optimal, the additional constraints:

$$x = y, \quad p = q \qquad (3.13)$$

must hold.

Since (3.11) is linear in x and p, we can easily solve the TPBVP, once the values of y and q are provided. If the solutions x and p satisfy (3.13), they are the solutions to the TPBVP of (3.5) and (3.6). However, generally, it is not the case, that is to say, the interaction balance:

$$\dot{r}(t) \triangleq \begin{bmatrix} x(t) - y(t) \\ p(t) - q(t) \end{bmatrix} = 0, \qquad t \in [t_0, t_f] \qquad (3.14)$$

is not satisfied. The interaction-coordination algorithm adjusts y and q as follows:

$$\begin{bmatrix} {}^{k+1}y \\ {}^{k+1}q \end{bmatrix} = \begin{bmatrix} {}^{k}y \\ {}^{k}q \end{bmatrix} + \alpha\, {}^{k}r \quad , \tag{3.15}$$

until

$${}^{k}G \triangleq \left[\frac{1}{2n(t_f - t_0)} \int_{t_0}^{t_f} {}^{k}r'(t)\, {}^{k}r(t)\, dt\right]^{1/2} \tag{3.16}$$

is reduced to zero or sufficietly small value, where k is the iteration number and α is a positive constant step size.

By experience, we know that a large value must be chosen for one of the weigths to obtain fast convergence. However, the larger the value is, the stiffer (3.11) is. Therefore, there is a limitation on the availble value for weights in view of the numerical error. As we mentioned in Chapter 2, the time-decomposition algorithm is effective in reducing the numerical error, the additional use of it remarkably is expected to improve the convergence characteristics of the interaction-coordination algorithm.

Remark 3.1.

The transition matrices of linear TPBVP's derived by the interaction-coordination algorithm are the same to each other. Therefore, we need to calculate the transition matrix, consequently Γ of (2.16), only once in the whole iterations. On the other hand, those of linear TPBVP's derived by the quasilinearization method are different from each other and therefore we must calculate them and Γ at every iteration. Thus, the additional use of the time-decomposition algorithm is a less burden to the interaction-coordination algorithm than to the quasilinearization method.

In the rest of this subsection, we show the applicability of the time-decomposition algorithm to the linear TPBVP of (3.5) and (3.11) in the case that A, B, Q, and R are constant matrices. To this end, it suffices to prove the following theorem.

Theorem 3.1.

Assume that A, B, Q, and R in (3.11) are constant matrices. Let Φ be the transition matrix of

$$\frac{d}{dt}\begin{bmatrix} x \\ p \end{bmatrix} = D^*(t)\begin{bmatrix} x \\ p \end{bmatrix} . \qquad (3.17)$$

Then, $\phi_{12}(t, 0)$, the element matrix of Φ, is nonsingular for arbitrary t $(t \neq 0)$, if the pair (A, B) is controllable.

To prove the theorem, we begin with the following lemma, due to Wonham [79].

Lemma 3.1.

Assume that the pair (A, B) is stabilizable† and the pair (C, A) is detectable††. Then, for arbitrary Q (≥ 0) and R (> 0), the matrix Riccati equation

$$P\bar{A} + \bar{A}'P - P\bar{B}R^{-1}\bar{B}'P + \bar{C}'Q\bar{C} = 0 \qquad (3.18)$$

has a unique nonnegative solution P.

From Lemma 3.1, we see that for A, B, Q, and R in the hypothesis of Theorem 3.1, there exists a unique nonnegative solution P to (3.18).

---

†: there exists a matrix K such that $A + BK$ is stable.

††: there exists a matrix K such that $A + KC$ is stable.

Now we prove the theorem.

Proof of Theorem 3.1.

First, we prove the theorem for $Q \equiv 0$. In this case, from (3.17), $p(t)$ is given by

$$p(t) = \exp[-A't]p_0 , \qquad (3.19)$$

then,

$$\dot{x} = Ax - \beta BR^{-1}B'\exp[-A't]p_0 \qquad (3.20)$$

and

$$x(t) = \exp[At]x_0 - \beta\int_0^t \exp[A(t-\tau)]BR^{-1}B'\exp[A'(t-\tau)]d\tau \cdot \exp[-A't]p_0 , \qquad (3.21)$$

where x_0 and p_0 are the initial conditions of x and p, respectively. Therefore,

$$\phi_{12}(t, 0) = -\beta\int_0^t \exp[A(t-\tau)]BR^{-1}B'\exp[A'(t-\tau)]d\tau\exp[-A't] . \qquad (3.22)$$

Since $\exp[-A't]$ is nonsingular, we consider the nonsigularity of

$$\tilde{\phi}_{12}(t, 0) = \int_0^t \exp[A(t-\tau)]BR^{-1}B'\exp[A'(t-\tau)]d\tau \qquad (t \neq 0) . \qquad (3.23)$$

Assume, to the contrary, that $\tilde{\phi}_{12}(t_1, 0)$ is not nonsingular. Then, there exists a non-zero vector p_1 such that

$$\tilde{\phi}_{12}(t_1, 0)p_1 = 0 . \qquad (3.24)$$

Then,

$$p_1'\tilde{\phi}_{12}(t_1, 0)p_1 = \int_0^{t_1} p_1'\exp[A(t_1 - \tau)]BR^{-1}B'\exp[A'(t_1 - \tau)]p_1 d\tau$$

$$= \int_0^{t_1} ||B'\exp[A'(t_1 - \tau)]p_1||_{R^{-1}} d\tau = 0. \qquad (3.25)$$

Since R is positive definite, (3.25) means that

$$p_1'\exp[At]B = 0, \qquad \forall t \in [0, t_1]. \qquad (3.26)$$

Differentiating (3.26) with respect to t, we obtain

$$p_1'[B, AB, \dots, A^{n-1}B] = 0, \qquad (3.27)$$

which contradicts the controllability of the pair (A, B). Thus, $\phi_{12}(t, 0)$ $(t \neq 0)$ is nonsingular.

Next, we consider the case $Q > 0$. Define a $2n \times 2n$-dimensional matrix S as

$$S = \begin{bmatrix} I_n, & 0_n \\ P, & -I_n \end{bmatrix}, \qquad (3.28)$$

where P is the above mentioned solution to (3.18). Then S^{-1} is identical to S. Let

$$\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = S \begin{bmatrix} x \\ p \end{bmatrix}, \qquad (3.29)$$

then, from (3.17)

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = SD^*S^{-1}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix}. \qquad (3.30)$$

Here,

$$SD^*S^{-1} = \begin{bmatrix} A - BR^{-1}B'P, & BR^{-1}B' \\ 0, & -(A - BR^{-1}B'P)' \end{bmatrix}. \quad (3.31)$$

Thus, this case is reduced to the case $Q \equiv 0$. Assume that $\phi_{12}(t, 0)$ is not nonsingular, then we deduce that the pair $(A - BR^{-1}B'p, B)$ is not controllable, wgich contradicts the controllability of the pair (A, B) [80]. Thus, the theorem is proved.

Q.E.D.

3.2.3. Illustrative Examples

In this section, we examine the numerical solution of physical problems in order to illustrate the application of the algorithm described in the preceding subsection. For the numerical integration of the differential equations, the fourth-order Runge-Kutta-Gill method is employed, where use is made of one hundred grid points. The criterion for convergence is set equal to 5.0×10^{-5}.

Example 3.1 [54].

The following equations approximately describe a three-axis attitude-control system of an orbiting space vehicle.

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= \varepsilon x_4 + \varepsilon x_4 x_6 + \varepsilon x_3 u_3 + u_1, \\ \dot{x}_3 &= x_4, \\ \dot{x}_4 &= -\varepsilon x_2 - \varepsilon x_2 x_6 - \varepsilon x_1 u_3 + u_2, \end{aligned} \quad (3.32)$$

$$\dot{x}_5 = x_6 ,$$

$$\dot{x}_6 = \varepsilon x_2 x_6 + \varepsilon x_1 u_2 + u_3 ,$$

where ε is a parameter introduced for convenience. The performance index is taken to be

$$J = \frac{1}{2}\int_0^5 \left(\sum_{i=1}^{6} x_i^2 + \sum_{i=1}^{3} u_i^2 \right) dt . \tag{3.33}$$

The initial condition is $x(0) = [1, 0, 1, 0, 1, 0]$ and the terminal condition is not specified.

To solve this problem by the interaction-coordination algorithm, the overall system is decomposed into the three subsystem of $[x_1, x_2]$, $[x_3, x_4]$, and $[x_5, x_6]$ [55, 63]. For solving one of the subsystems, the variables of the other subsystems and those in the nonlinear terms are replaced by interaction vectors. The necessary condition for Subsystem 1 is described by

$$\begin{aligned}
\dot{x}_1 &= x_2 , \\
\dot{x}_2 &= -\beta p_2 + (\beta - 1) q_2 + \varepsilon [y_4(1+y_6) - y_3(\varepsilon y_3 q_2 - \varepsilon y_1 q_4 + q_6)], \\
\dot{p}_1 &= -\kappa x_1 + (\kappa - 1) y_1 + \varepsilon^2[-y_3 q_2 q_4 + y_1(q_4^2 + q_6^2)] , \\
\dot{p}_2 &= -\kappa x_2 - p_1 + (\kappa - 1) y_2 + \varepsilon [q_4(1+y_6) - y_4 q_6] .
\end{aligned} \tag{3.34}$$

Similar problems obtained for Subsystems 2 and 3 are omitted here.

As an example, let $\varepsilon = 6$. The terminal condition is given by $p(t_f) = 0$. Figure 3.1 shows variations of G defined by (3.16) with the computing time T when the problem is solved by the interaction-coordination algorithm with the initial estimates ${}^1y$ and ${}^1q$ being solutions x and p, respectively of the homogeneous part of the reduced linear TPBVP. In the figure, L denotes the number of subintervals. For $L=2$, convergence rate is very poor since the value of weights is limited small because of the numerical error. When $L=3$ and $\beta=20$, convergence is obtained but for $\beta=30$, G is oscillatory. When $L=4$ and $\beta=30$, G converges quite rapidly without oscillation. In this case it takes 46 iterations to converge. In the figure, we show the best data for α.

Figure 3.2 shows variations of G defined by (3.10) of the quasilinearization method with the computing time T for $\varepsilon=3.2$. The initial estimates are taken to be the same as that of the interaction-coordination algorithm with $\beta=\kappa=1$. When $L=1$, that is, the time decomposition is not applied, G is divergent. When $L=2$, the variation of G is oscillatory because of the numerical error. The value of

$$\{\sum_{i=1}^{6}[p_i(2.5+0) - p_i(2.5-0)]^2\}^{1/2}$$

at the end of the first iteration is 0.031 which is to be zero. When $L = 4$, convergence is obtained with the above value 0.002. It takes 12 iterations to converge. For $\varepsilon \geq 3.2$, the method diverges even with $L=10$.

Figures 3.3 and 3.4 show the optimal trajectory and time history of the optimal control, respectively, for $\varepsilon=6$.

Example 3.2 [60].

Next we consider the following minimum-energy transfer of a low-thrust propulsion vehicle between circular orbits:

$$
\begin{aligned}
\dot{x}_1 &= x_2 , \\
\dot{x}_2 &= x_1 - \varepsilon[2x_4 + x_1/(x_1^2 + x_3^2 + x_5^2)^{3/2}] + u_1 , \\
\dot{x}_3 &= x_4 , \\
\dot{x}_4 &= x_3 + \varepsilon[2x_2 - x_3/(x_1^2 + x_3^2 + x_5^2)^{3/2}] + u_2 , \\
\dot{x}_5 &= x_6 , \\
\dot{x}_6 &= -\varepsilon x_5/(x_1^2 + x_3^2 + x_5^2)^{3/2} + u_3 ,
\end{aligned}
\tag{3.35}
$$

where ε is a parameter introduced for conveniece. The objective is to transfer the state of the system from $x(0) = \pi_0$ to $x(\pi) = \pi_f$ so as to minimize the performance index:

$$
J = \frac{1}{2}\int_0^{\pi} (u_1^2 + u_2^2 + u_3^2)dt . \tag{3.36}
$$

To solve the problem by the interaction-coordination algorithm, the overall system is decomposed into the three subsystems in the same manner as Example 3.1. The derived equations for Subsystem 1 are as follows:

$$
\begin{aligned}
\dot{x}_1 &= x_2 , \\
\dot{x}_2 &= x_1 - \beta p_2 + (\beta - 1)q_2 - 2\varepsilon y_4 - \varepsilon y_1/(y_1^2 + y_3^2 + y_5^2)^{3/2} , \\
\dot{p}_1 &= -\kappa x_1 - p_2 + \kappa y_1 + \varepsilon[q_2(-2y_1^2 + y_3^2 + y_5^2) - 3y_1(y_3 q_4 + y_5 q_6)]/ \\
&\quad (y_1^2 + y_3^2 + y_5^2)^{5/2} , \\
\dot{p}_2 &= -\kappa x_2 - p_1 + \kappa y_2 - 2\varepsilon q_4 .
\end{aligned}
\tag{3.37}
$$

where $-\kappa x_1 + \kappa y_1$ and $-\kappa x_2 + \kappa y_2$ of the third and the fourth equations, respectively, are artificially added terms to accelerate the convergence [55]. Similar equations for Subsystems 2 and 3 are omitted here.

By way of example, let $\pi_0 = [0, 0, 1, 0, 0, 0]$ and $\pi = [*, -0.75, 1.5, 0, 0, \pi/5]$, that is to say, $x_1(\pi)$ is set free. Then, $p_1(\pi) = 0$ is obtained from the necessary condition for optimality. Figure 3.5 shows variations of G with the computing time T for $\varepsilon = 4$. When $L = 1$, that is, when the time decomposition is not applied, G is divergent. The reason is that the values of the weights are to be bounded small because of the limitation of numerical accuracy. On the other hand, the variation of with $L = 2$ is oscillatory. Since in this case larger values of the weights can be used than in the previous case, the divergent tendency is suppressed, but convergence is not attained. When $L = 4$, since the algorithm can use much larger values of the weights without loss of numerical accuracy, convergence is obtained.

The problem is solved also by the quasilinearization method. Figure 3.6 shows variations of G with the computing time T for $\varepsilon = 4$. We see from the figure that, since the TPBVP derived by the quasilinearization method is not so stiff, the appication of the time-decomposition algorithm has only disadvantage of consuming time.

In Figures 3.7 and 3.8, the optimal trajectory and time history of the optimal control, respectively, are shown for $\varepsilon = 4$.

Remark 3.2.

It should be noted that it has no advantage to apply the time-decomposition algorithm to the TPBVP to which the conventional superposition principle can offer a satisfactorily accurate solution, but, to the contrary, has the disadvantage of consuming more computing time.

3.3. A Multiple-Target Problem

Next, we consider the case that the system equation contains discontinuities at corner times and that some elements of the state variable are specified there [56, 57].

3.3.1. Problem Statement

In this section, we discuss a solution of the following multiple-target problem. The system equation is described by

$$\dot{x} = A_i(t)x + B_i(t)u + f_i(t, x) , \quad t_{i-1} \leqq t < t_i \quad (i = 1, 2, \ldots, N), \tag{3.38}$$

where $x(t)$ is an n-dimensional state vector, $u(t)$ an m-dimensional control vector. A_i and B_i are $n\times n$- and $n\times m$-dimensional matrices, respectively, and f_i is an n-dimensional vector function of the class C^2 with respect to x and these are continuous in $t \in [t_{i-1}, t_i]$. The boundary conditions are given by

$$L_i x(t_i) = \pi_i \quad (i = 0, 1, \ldots, N), \tag{3.39}$$

where π_i is an r_i-dimensional prescribed vector and L_i is an $r_i\times n$-dimensional matrix containing only one nonzero element in each row, and it is assumed that t_i is not specified and is a parameter to be determined ($i = 1, 2, \ldots, N-1$).

The objective is to minimize the following performance index of the quadratic type:

$$J = \frac{1}{2}\int_{t_0}^{t_N} [x'Q(t)x + u'R(t)u]dt \tag{3.40}$$

with respect to u and t_i ($i=1, 2, \ldots, N-1$), where $Q(t)$ is an $n\times n$-dimensional symmetric positive semidefinite matrix and $R(t)$ an $m\times m$-dimensional symmetric positive definite matrix.

Now define the Hamiltonian H of (3.38) and (3.40) as

$$H = \frac{1}{2}(x'Qx + u'Ru) + \sum_{i=1}^{N} p'(A_i x + B_i u + f_i) . \tag{3.41}$$

Then, according to the variational principle, the necessary conditions for optimality are obtained as follows [1, 12].

$$\left.\begin{aligned} \dot{x} &= \left(\frac{\partial H^{(i)}}{\partial p}\right)' = A_i(t)x + B_i(t)u + f_i(t, x) , & (3.42)\\ \dot{p} &= -\left(\frac{\partial H^{(i)}}{\partial x}\right)' = -Q(t)x - A_i'(t)p - \left(\frac{\partial f_i}{\partial x}\right)'p , & (3.43)\\ \left(\frac{\partial H^{(i)}}{\partial u}\right)' &= R(t)u + B_i'(t)p = 0 , & (3.44) \end{aligned}\right\} \begin{aligned} & t \in [t_{i-1}, t_i) \\ & (i=1, 2, \ldots, N) \end{aligned}$$

$$g_{t_i} \triangleq H(t_i^-) - H(t_i^+) = 0 \qquad (i=1, 2, \ldots, N-1) \tag{3.45}$$

with the boundary conditions

$$\begin{aligned} & L_i x(t_i) = \pi_i \quad (i=0, 1, \ldots, N) , \\ & p_j(t_i^-) = p_j(t_i^+) = \nu_{ji} \ (\text{if } x_j(t_i) \text{ is not specified}; \ i=1, 2, \ldots, N-1) , \end{aligned} \tag{3.46}$$

where $p_j(t_i)$ denotes the j-th element of $p(t_i)$ and ν_{ji} is a Lagrange multiplier which is to be determined so as to make x_j continuous at $t=t_i$. Substitution of $u=-R^{-1}B_i'p$ into (3.42) yields the following multipoint boundary-value problem (MPBVP):

$$\frac{d}{dt}\begin{bmatrix} x \\ p \end{bmatrix} = \begin{bmatrix} A_i(t), & -E_i(t) \\ -Q(t), & -A_i'(t) \end{bmatrix}\begin{bmatrix} x \\ p \end{bmatrix} + \begin{bmatrix} h_{1i}(t, x) \\ h_{2i}(t, x, p) \end{bmatrix} \quad (i = 1, 2, \ldots, N) \quad (3.47)$$

constrained by the boundary conditions (3.46) and the optimality condition for t_i (3.45), where $E_i = B_i R^{-1} B_i'$, $h_{1i}(t, x) = f_i(t, x)$, and $h_{2i}(t, x, p) = -(\frac{\partial f_i}{\partial x})' p$.

Once the t_i's are assumed, the problem is reduced to solving the MPBVP of (3.46) and (3.47) with discontinuities at the specified corner times. The solution to the MPBVP, however, does not necessarily satisfy the optimality condition (3.45). Then, the assumed values of the t_i's are to be corrected by an appropriate algorithm. In our algorithm, they are corrected by

$$^{\ell+1}t_i = {}^{\ell}t_i - \eta \, \mathrm{sign}[{}^{\ell}g_{t_i}] \quad \left(\begin{matrix} i = 1, 2, \ldots, N-1 \\ \ell = 1, 2, \ldots \end{matrix} \right), \quad (3.48)$$

until (3.45) is attained, where ℓ denotes the iteration number and η is a positive constant step size which is reduced according to the change of the sign of g_{t_i}.

Thus, the multiple-target problem is solved by a three-level algorithm. The objective of the highest level is the optimal correction of the corner times, that is, t_i's are corrected according to (3.48), using the solution to the MPBVP with specified corner times obtained in the lower levels. The nonlinear MPBVP is reduced to a sequence of linear MPBVP's at the intermediate level and they are solved at the lowest level by a discontinuous version of the time-decomposition algorithm. In the following, the interaction-coordination algorithm is adopted for the linearization,

3.3.2. Modification of the Time-Decomposition Algorithm

In this section we modify the time-decomposition algorithm of Chapter 2 to solve linear MPBVP's. Hereafter, for simplicity, the discontinuity is assumed to occur only once during the overall control duration. Then, let $N = 2$ and let the corner time $t = t_1$. Also, for simplicity, we restrict our discussion to the case where the boundary conditions are given in the form:

$$\psi_i[x(t_i^-), x(t_i^+)] = \begin{bmatrix} x_{1(i)}(t_i^-) - x_{1(i)}(t_i^+) \\ x_{2(i)}(t_i) - \pi_i \\ x_{3(i)}(t_i^-) - x_{3(i)}(t_i^+) \end{bmatrix} = 0 \quad (i = 0, 1, 2), \qquad (3.49)$$

where $x_{j(i)}$ is an $r_{j(i)}$-dimensional vector with $r_{1(i)} + r_{2(i)} + r_{3(i)} = n$ $(i = 0, 1, 2)$. (3.49) means that some elements of the state variable x are specified at $t = t_1$ and that x is continuous in t.

In this case, the necessary conditions for optimality (3.42)∿(3.46) can be written as:

Subinterval 1: $t \in [t_0, t_1]$

$$\left.\begin{aligned} \dot{x} &= A_1 x + B_1 u + f_1(t, x), \\ \dot{p} &= -Qx - A_1' p - \left(\frac{\partial f_1}{\partial x}\right)' p, \\ \left(\frac{\partial H^{(1)}}{\partial u}\right)' &= Ru + B_1' p = 0, \end{aligned}\right\} \quad (3.50.1)$$

Subinterval 2: $t \in [t_1, t_2]$

$$\left.\begin{aligned} \dot{x} &= A_2 x + B_2 u + f_2(t, x), \\ \dot{p} &= -Qx - A_2' p - \left(\frac{\partial f_2}{\partial x}\right)' p, \\ \left(\frac{\partial H^{(2)}}{\partial x}\right)' &= Ru + B_2' p = 0, \end{aligned}\right\} \quad (3.50.2)$$

$$H^{(i)} = \frac{1}{2}(x'Qx + u'Ru) + p'(A_i x + B_i u + f_i) \quad (i = 1, 2),$$

$$
\left.\begin{aligned}
x(t_0) &= \pi_0, \\
x_2(t_1) &= \pi_1, \\
p_1(t_1) &= \nu_{11}, \\
p_3(t_1) &= \nu_{31},
\end{aligned}\right\} (3.51.1)
\qquad
\left.\begin{aligned}
x_2(t_1) &= \pi_1, \\
x(t_2) &= \pi_2, \\
p_1(t_1) &= \nu_{11}, \\
p_3(t_1) &= \nu_{31},
\end{aligned}\right\} (3.51.2)
$$

$$
g_{t_1} \triangleq H^{(1)}(t_1) - H^{(2)}(t_1) = 0. \qquad (3.52)
$$

Therefore, the problem is reduced to finding the boundary conditions $x_1(t_1)$ and $x_3(t_1)$ which guarantee the continuity of $p_1(t)$ and $p_3(t)$ at $t = t_1$, and also to finding the optimal corner time t_1 which satisfies (3.51).

Once t_1 is assumed, the MPBVP of (3.50) and (3.51) can be solved with use of both the interaction-coordination algorithm and the time-decomposition algorithm. Suppose that (3.50) is linearized as (3.47). Let $\widetilde{p_1(t_1^{\pm})}$ and $\widetilde{p_3(t_1^{\pm})}$ be the values of the solutions p_1 and p_3 at $t = t_1^{\pm}$, respectively, to the TPBVP's with the boundary conditions $x(t_0) = \pi_0$, $x_1(t_1) = \widetilde{x_1(t_1)}$, $x_2(t_1) = \pi_1$, $x_3(t_1) = \widetilde{x_3(t_1)}$, and $x(t_2) = \pi_2$. Then, from Theorem 2.1,

$$
\begin{bmatrix} \widetilde{p_1(t_1^-)} - \widetilde{p_1(t_1^+)} \\ \widetilde{p_3(t_1^-)} - \widetilde{p_3(t_1^+)} \end{bmatrix} = \tilde{\Gamma} \begin{bmatrix} \widetilde{x_1(t_1)} \\ \widetilde{x_3(t_1)} \end{bmatrix} + \begin{bmatrix} \Gamma_{12}\pi_1 \\ \Gamma_{32}\pi_1 \end{bmatrix} + \begin{bmatrix} W_1 \\ W_3 \end{bmatrix}, \qquad (3.53)
$$

where Γ_{ij} is an $r_{i(1)} \times r_{j(1)}$-dimensional matrix, W_i is an $r_{i(1)}$-dimensional vector with

$$\begin{bmatrix} \Gamma_{11}, \Gamma_{12}, \Gamma_{13} \\ \Gamma_{21}, \Gamma_{22}, \Gamma_{23} \\ \Gamma_{31}, \Gamma_{32}, \Gamma_{33} \end{bmatrix} = \Gamma, \qquad \tilde{\Gamma} = \begin{bmatrix} \Gamma_{11}, \Gamma_{13} \\ \Gamma_{31}, \Gamma_{33} \end{bmatrix}, \qquad [W_1', W_2', W_3']' = V . \tag{3.54}$$

Hence, similarly to Corollary 2.1, the exact solution $X = [x_1'(t_1), x_3'(t_1)]'$ is given by

$$X = \tilde{X} - \tilde{\Gamma}^{-1}\tilde{P}, \tag{3.55}$$

where $\tilde{X} = [\widehat{x_1'(t_1)}, \widehat{x_3'(t_1)}]'$, $\tilde{P} = [(\widehat{p_1(t_1^-)} - \widehat{p_1(t_1^+)})', (\widehat{p_3(t_1^-)} - \widehat{p_3(t_1^+)})']'$. Then, the linear TPBVP's are solved again with the exact boundary conditions. The interation-coordination algorithm is iterated until the solution reduces (3.16) to zero or a sufficient small value.

After the lowest and the intermediate level calculations, the corner time t_1 is corrected by (3.48). The highest level calculation is carried out until g_{t_1} is reduced to zero or a sufficient small value.

3.3.3. Summary of the Algorithm

In this section, we summarize the results obtained above into the form of an algorithm.

Step 1: Set $\ell = 1$, $k = 1$. Assume ${}^1t_1$. Let ${}^1\widehat{x_1({}^1t_1)} = {}^1\widehat{x_3({}^1t_1)} = 0$, ${}^1y(t) = {}^1q(t) = 0, t \in [t_0, t_2]$. Choose appropriate values of β, κ, α, and η.

Step 2: Solve the homogeneous part of (3.50) with the boundary conditions $\tilde{X} = e_\nu$, the ν-th unit vector, $x(t_0) = 0$, $x_2({}^1t_1) = 0$, and $x(t_2) = 0$. Then, the difference $\tilde{P}$ represents the ν-th column of Γ. Calculate Γ^{-1}.

Step 3: Solve each linear TPBVP of the subintervals with the boundary conditions

Subinterval 1:

$$x(t_0) = \pi_0,\quad x_1(t_1) = {}^k\widetilde{x_1(t_1)},\quad x_2(t_1) = \pi_1,\quad x_3(t_1) = {}^k\widetilde{x_3(t_1)}\,, \tag{3.56}$$

Subinterval 2:

$$x_1(t_1) = {}^k\widetilde{x_1(t_1)},\quad x_2(t_1) = \pi_1,\quad x_3(t_1) = {}^k\widetilde{x_3(t_1)},\quad x(t_2) = \pi_2\,.$$

Let us denote the solutions p_1 and p_3 at $t = t_1^{\pm}$ as $\widetilde{p_1(t_1^{\pm})}$ and $\widetilde{p_3(t_1^{\pm})}$, respectively.

Step 4: By (3.55), determine the exact solution $x(t_1)$. Solve the subinterval TPBVP's again with use of $x(t_1)$ and π_0 and π_2. Let us denote the solution as ${}^kx$ and ${}^kp$.

Step 5: If ${}^kG$ of (3.16) is small enough, proceed to *Step* 6. Otherwise, correct ${}^ky$ and ${}^kq$ by (3.15), replace ${}^k\widetilde{x_1(t_1)}$ and ${}^k\widetilde{x_3(t_1)}$ by ${}^kx_1(t_1)$ and ${}^kx_3(t_1)$, respectively, and replace k by k+1. Then, return to *Step* 3.

Step 6: Compute ${}^\ell g_{t_1}$ by (3.52). If ${}^\ell g_{t_1}$ is small enough, the optimum is attained and the calculation is terminated. Otherwise, correct ${}^\ell t_1$ by (3.48), replace ℓ by ℓ+1, and return to *Step* 2.

3.3.4 Illustrative Examples

Three physical problems are examined to illustrate the applications of the present algorithm. For the numerical integration of the differential equations, the fourth-order Runge-Kutta-Gill scheme is employed, where use is made of one hundred grid-points in the overall interval.

Example 3.3 [76].

Let us consider the problem of minimizing the functional:

$$J[u] = \int_0^2 u^2 dt \tag{3.57}$$

with respect to the control u and the corner time t_1. The state equations governing the system are

$$\left.\begin{aligned} \dot{x}_1 &= x_2 , \\ \dot{x}_2 &= u , \end{aligned}\right\} t \in [0, t_1) \quad (3.58.1) \qquad \left.\begin{aligned} \dot{x}_1 &= x_2 , \\ \dot{x}_2 &= 2u , \end{aligned}\right\} t \in [t_1, 2] \quad (3.58.2)$$

and the boundary conditions are

$$\left.\begin{aligned} x_1(0) &= 1, \\ x_2(0) &= 1, \end{aligned}\right\} (3.59.1) \qquad \left.\begin{aligned} x_1(2) &= 0, \\ x_2(2) &= 0, \end{aligned}\right\} (3.59.2)$$

$$x_1(t_1) = 0.5. \tag{3.60}$$

(3.58) implies that the mass of the article is reduced by a half at the corner time t_1.

For this problem, the necessary conditions for optimality are written as follows:

$$\left.\begin{aligned} \dot{x}_1 &= x_2 , \\ \dot{x}_2 &= -0.5p_2 , \\ \dot{p}_1 &= 0 , \\ \dot{p}_2 &= -p_1 , \end{aligned}\right\} t \in [0, t_1) \quad (3.61.1) \qquad \left.\begin{aligned} \dot{x}_1 &= x_2 , \\ \dot{x}_2 &= -2p_2 , \\ \dot{p}_1 &= 0 , \\ \dot{p}_2 &= -p_1 , \end{aligned}\right\} t \in [t_1, 2] \quad (3.61.2)$$

$$x_1(0) = x_2(0) = 1 , \quad (3.62.1) \qquad x_1(2) = x_2(2) = 0 , \quad (3.62.2)$$

$$x_1(t_1) = 0.5\,, \tag{3.63}$$

$$p_2(t_1^-) = p_2(t_1^+)\,, \tag{3.64}$$

$$u = -0.5p_2\,, \quad (3.65.1) \qquad\qquad u = -p_2\,, \quad (3.65.2)$$

$$g_{t_1} = u(t_1^-)^2 + p_1(t_1^-)x_2(t_1) + p_2(t_1^-)u(t_1^-) - u(t_1^+)^2 - p_1(t_1^+)x_2(t_1) - 2p_2(t_1^+)u(t_1^+) = 0\,. \tag{3.66}$$

It is easily shown that the general solutions to (3.61) are given by

$$\left.\begin{aligned} x_1(t) &= [c_1t^3 - 3c_2t^2 + 12c_3t + 12c_4]/12\,, \\ x_2(t) &= [c_1t^2 - 2c_2t + 4c_3]/4\,, \\ p_1(t) &= c_1\,, \\ p_2(t) &= -c_1t + c_2\,, \end{aligned}\right\} \tag{3.67.1}$$

$$\left.\begin{aligned} x_1(t) &= [d_1t^3 - 3c_2t^2 + 3d_3t + 3d_4]/3\,, \\ x_2(t) &= d_1t^2 - 2d_2t + d_3\,, \\ p_1(t) &= d_1\,, \\ p_2(t) &= -d_1t + d_2\,, \end{aligned}\right\} \tag{3.67.2}$$

where $c_1 \sim c_4$ and $d_1 \sim d_4$ are constants.

Now let us follow the algorithm of Section 3.3.3.

Step 1: Let ${}^1t_1 = t_1$. Since the problem is linear, it is unnecessary to utilize the interaction-coordination algorithm.

Step 2: $\tilde{\Gamma}$ is obtained as follows: Solve (3.61) with the boundary conditions $x_1(0) = x_2(0) = x_1(t_1) = x_1(2) = x_2(2) = 0$, and $x_2(t_1) = 1$. The coefficients c_i and d_i are obtained as follows:

$$c_1 = 12/t_1^2, \quad c_2 = 4/t_1, \quad c_3 = c_4 = 0,$$
$$d_1 = 3/(t_1 - 2)^2, \quad d_2 = (t_1 + 4)/(t_1 - 2)^2, \tag{3.68}$$
$$d_3 = 4(t_1 + 1)/(t_1 - 2)^2, \quad d_4 = -4t_1/(t_1 - 2)^2.$$

Hence,

$$\tilde{\Gamma} = \widetilde{p_2(t_1^-)} - \widetilde{p_2(t_1^+)} = -c_1 t_1 + c_2 + d_1 t_1 - d_2$$
$$= -2(3t_1 - 8)/t_1(c_1 - 2). \tag{3.69}$$

Step 3: Let $\widetilde{x_2(t_1)} = 0$. The solutions to (3.61) with the boundary conditions $x_1(0) = x_2(0) = 1$, $x_1(t_1) = 0.5$, $x_2(t_1) = \widetilde{x_2(t_1)}$, and $x_1(2) = x_2(2) = 0$ yield

$$c_1 = 12(t_1 + 1)/t_1^3, \quad c_2 = 2(4t_1 + 3)/t_1^2, \quad c_3 = c_4 = 1,$$
$$d_1 = -3/(t_1 - 2)^3, \quad d_2 = -3(t_1 + 2)/2(t_1 - 2)^3, \tag{3.70}$$
$$d_3 = -6t_1/(t_1 - 2)^3, \quad d_4 = (6t_1 - 4)/(t_1 - 2)^3.$$

Hence,

$$\widetilde{p_2(t_1^-)} - \widetilde{p_2(t_1^+)} = -c_1 t_1 + c_2 + d_1 t_1 - d_2$$
$$= (-8t_1^3 + 17t_1^2 + 16t_1 - 48)/2t_1^2(t_1 - 2)^2. \tag{3.71}$$

Step 4: Substitution of (3.69) and (3.71) into (3.55) yields

$$x_2(t_1) = \widetilde{x_2(t_1)} - \tilde{\Gamma}^{-1}[\widetilde{p_2(t_1^-)} - \widetilde{p_2(t_1^+)}]$$

$$= (-8t_1^3 + 17t_1^2 + 16t_1 - 48)/4t_1(t_1 - 2)(3t_1 - 8) . \qquad (3.72)$$

Again solve the TPBVP's (3.61) with the boundary conditions thus obtained. Then, the constants are obtained to be as follows:

$$c_1 = (12t_1^3 - 81t_1^2 + 72t_1 + 48)/t_1^3(t_1 - 2)(3t_1 - 8) ,$$

$$c_2 = (16t_1^3 - 77t_1^2 + 60t_1 + 48)/t_1^2(t_1 - 2)(3t_1 - 8) ,$$

$$c_3 = c_4 = 1 ,$$

$$d_1 = 3(-8t_1^3 + 5t_1^2 + 48t_1 - 48)/4t_1(t_1 - 2)^3(3t_1 - 8) , \qquad (3.73)$$

$$d_2 = (-8t_1^4 - 33t_1^3 + 96t_1^2 + 112t_1 - 192)/4t_1(t_1 - 2)^3(3t_1 - 8) ,$$

$$d_3 = (-8t_1^4 - 9t_1^3 + 81t_1^2 - 32t_1 - 48)/t_1(t_1 - 2)^3(3t_1 - 8) ,$$

$$d_4 = (8t_1^3 + t_1^2 - 76t_1 + 8)/(t_1 - 2)^3(3t_1 - 8) .$$

Step 5: This step is skipped since the interaction-coordination algorithm is not utilized.

Step 6: The gradient g_{t_1} is given by

$$g_{t_1} = H^{(1)}(t_1) - H^{(2)}(t_1) = \frac{3}{4}[p_2(t_1)]^2 + [p_1(t_1^-) - p_1(t_1^+)]x_2(t_1)$$

$$= \frac{3}{4}(-c_1t_1 + c_2)^2 + (c_1 - d_1)x_2(t_1) \qquad (3.74)$$

with (3.72) and (3.73). (3.74) is a rational function of t_1. It is not difficult to solve $g_{t_1} = 0$ with the aid of a digital computer. Figure 3.9 shows the dependences of g_{t_1} and J on t_1. It can be seen from the figure that the optimal t_1 is 1.377⋯.

The algorithm is also carried out on a digital computer. The initial estimate of t_1 is 1.0 and the step size η in (3.48) is set to 0.1. When the sign of g_{t_1} is changed, η is reduced by a fifth. Figure 3.10 shows the variations of t_1, g_{t_1}, and J with the computing time T. The criterion for convergence is set at $g_{t_1} = 1.0\times10^{-3}$. After 16 iterations the algorithm converged. As seen from Figure 3.10, too severe criterion for the optimality with respect to t_1 contributes only to the computing time.

Figures 3.11 and 3.12 show the optimal trajectory on the x_1-x_2 plane and time history of the optimal control u, respectively.

Example 3.4 [75].

Next, we consider the optimal control of the system with Coulomb friction. The system equation is described by

$$\left.\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= u + \mu,\end{aligned}\right\} x_2 > 0 \quad (3.75.1) \qquad \left.\begin{aligned}\dot{x}_1 &= x_2, \\ \dot{x}_2 &= u - \mu,\end{aligned}\right\} x_2 \leq 0 \quad (3.75.2)$$

A constraint on u of saturation type is contained in the original problem but is omitted here. The objective is to transfer the state of the system from $[a, b]$ to the origin so as to minimize

$$J = \int_0^{t_f} u^2 dt . \qquad (3.76)$$

The corner time t_1 at which the system equation turns from (3.75.1) to (3.75.2) is determined by

$$x_2(t_1) = 0. \qquad (3.77)$$

The analytical solutions can be obtained in a similar way to

Example 3.3 as follows:

$$\left.\begin{aligned} x_1(t) &= [c_1 t^3 - 3(c_2 - 2\mu)t^2 + 12c_3 t + 12c_4]/12 , \\ x_2(t) &= [c_1 t^2 - 2(c_2 - 2\mu)t + 4c_3]/4 , \\ p_1(t) &= c_1 , \\ p_2(t) &= -c_1 t + c_2 , \end{aligned}\right\} t \in [0, t_1) \quad (3.78.1)$$

$$\left.\begin{aligned} x_1(t) &= [d_1 t^3 - 3(d_2 + 2\mu)t^2 + 12d_3 t + 12d_4]/12 , \\ x_2(t) &= [d_1 t^2 - 2(d_2 + 2\mu)t + 4d_3]/4 , \\ p_1(t) &= d_1 , \\ p_2(t) &= -d_1 t + d_2 , \end{aligned}\right\} t \in [t_1, t_f] \quad (3.78.2)$$

where

$$\left.\begin{aligned} c_1 &= 12(bt_1 + 2a - 2\pi_1)/t_1^3 , \\ c_2 &= [12(a - \pi_1) + 8bt_1 + 2\mu t_1^2]/t_1^2 , \\ c_3 &= b , \quad c_4 = a , \end{aligned}\right\} \quad (3.79.1)$$

$$\left.\begin{aligned} d_1 &= 24\pi_1/(t_f - t_1)^3, \quad d_2 = [12(t_f + t_1)\pi_1/(t_f - t_1)^3] - 2\mu , \\ d_3 &= 6t_1 t_f \pi_1/(t_f - t_1)^3, \quad d_4 = t_f^2 (t_f - 3t_1)\pi_1/(t_f - t_1)^2, \end{aligned}\right\} \quad (3.79.2)$$

and

$$\pi_1 = x_1(t_1) = 0.5t_1^3[2a + b(t_f - t_1)]/[t_1^3 + (t_f - t_1)^3] . \quad (3.80)$$

As an example, let $[a, b] = [0.5, 1.5]$, $t_f = 3.0$, $\mu = -0.2$. Figure 3.13

shows the dependences of g_{t_1} and J on t_1. The optimal corner time $t_1 = 0.8181$. The problem is solved by the proposed algorithm where the initial estimate of t_1 is set to 1.5, η is set to 0.1, and the integration step size is set to 0.015. Figure 3.14 shows the variations of t_1, g_{t_1}, and J with the computing time T. Figures 3.15 and 3.16 show the optimal trajectory on the $x_1 - x_2$ plane and time history of the optimal control u, respectively.

Example 3.5.

Finally, we consider the problem whose system equation contains non-linearities. The problem is discontinuous version of the three-axis attitude-control problem. The sytem equations are described by (3.32).

Suppose that the parameter ε changes discontinuously from 1.0 to 2.0 at $t = t_1$ at which $x_1(t_1) = 0.1$, $x_3(t_1) = 0.15$, and $x_5(t_1) = 0.005$ are to be satisfied. The objective is to find the control u and the corner time t_1 which minimize (3.33), starting from $x(0) = [1, 0, 1, 0, 1, 0]'$.

Figure 3.17 shows the variations of t_1, g_{t_1}, and J with the iteration number ℓ for correcting the corner time t_1. It takes 4.8 seconds to attain convergence. The parameters chosen are $\beta = \kappa = 1$, $\alpha = 1.0$, $\eta = 0.5$, ${}^1t_1 = 2.5$. η is reduced by a half when the sign of g_{t_1} is changed. Optimal corner time t_1 is obtained as 3.25. The optimal trajectory and time history of the optimal control are shown in Figures 3.18 and 3.19, respectively.

The examples could be solved by the direct method such as a steepest-descent method which will be summarized in the following chapter. Then, however, we must adjust both the corner time t_1 and the control function at the same time or sequentially, which makes the convergence characteristics of the method worse. Moreover, to make the solution satisfy the boundary conditions, the idea of penalty function must be employed, which makes the convergence characteristics still worse. Therefore, for the problem without control constraint, the indirect method such as the proposed method is more effective than the direct method.

3.4. Concluding Remarks

In this chapter, nonlinear optimal control problems are solved by use of the time-decomposition algorithm in conjunction with the linearization method.

In the former half of the chapter, the optimal control problem of systems described by a differential equation without discontinuity is considered. The time-decomposition algorithm is additionally employed to reduce the numerical error in solving linear TPBVP's derived by the linearization method, i.e., the quasilinearization method and the interaction-coordination algorithm.

Additional use of the time-decomposition algorithm with the quasilinearization method has the advantage that the numerical error is reduced and therefore the convergence region is widened nearly to the theoretical one. However, the widened region is at most the theoretical one. The combined algorithm has the disadvantage of consuming computing time, because the transition matrix and the matrix for correcting boundary values must be recalculated at each iteration.

On the other hand, when the time-decomposition algorithm is used with the interaction-coordination algorithm, the transition matrix and the correction matrix need be calculated only once. It is empirically known that the convergence characteristics of the interaction-coordination algorithm is much improved by modifying the TPBVP stiffer. Thus, the additional use of the time-decomposition algorithm is much more suitable to the interaction-coordination algorithm than to the quasilinearization method.

In the latter half of the chapter, the solution of a multiple-target problem is discussed. The system equation contains discontinuities at several unspecified intermediate points called 'corner times' and some elements of the state variable are specified at corner times as well as at the initial and the terminal times. The problem is reduced to a nonlinear MPBVP with unspecified corner times. Assuming the values of corner times, the nonlinear MPBVP is solved by the linearization method with the additional use of a discontinuous version of the time-decomposition algorithm. The optimal correction of corner times is made by a gradient method. Since, different from the algorithm of Ref. 75, the proposed algorithm does not employ the idea of penalty function, the solution satisfies the specified boundary condition exactly. The algorithm can offer analytical solutions to linear problems as well as the method of Ref.76. Moreover, the algorithm can treat nonlinear problems on a digital computer.

The proposed algorithm can be applied to the state-constrained problem [15, 21, 26, 31, 41] by assuming the number of corner times.

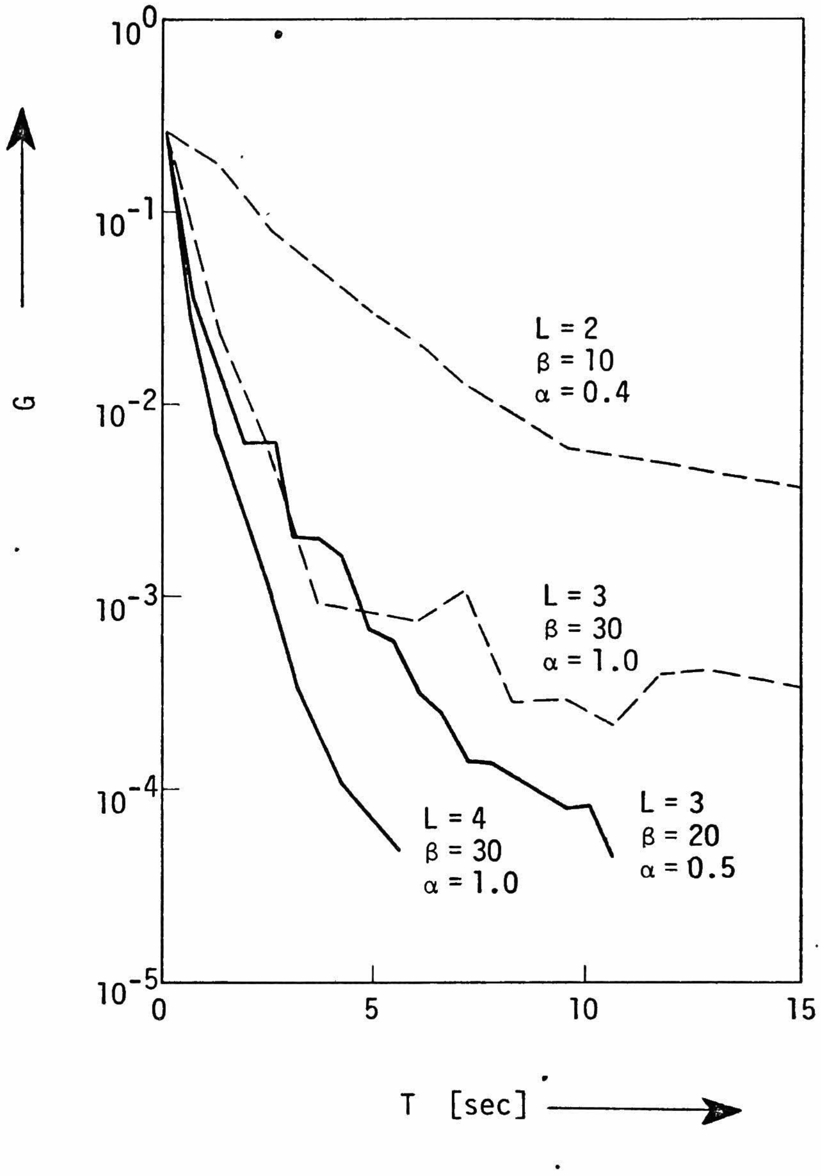

Fig. 3.1. Variations of G in case of the interaction-coordination algorithm with the computing time T for $\varepsilon = 6$ (Example 3.1).

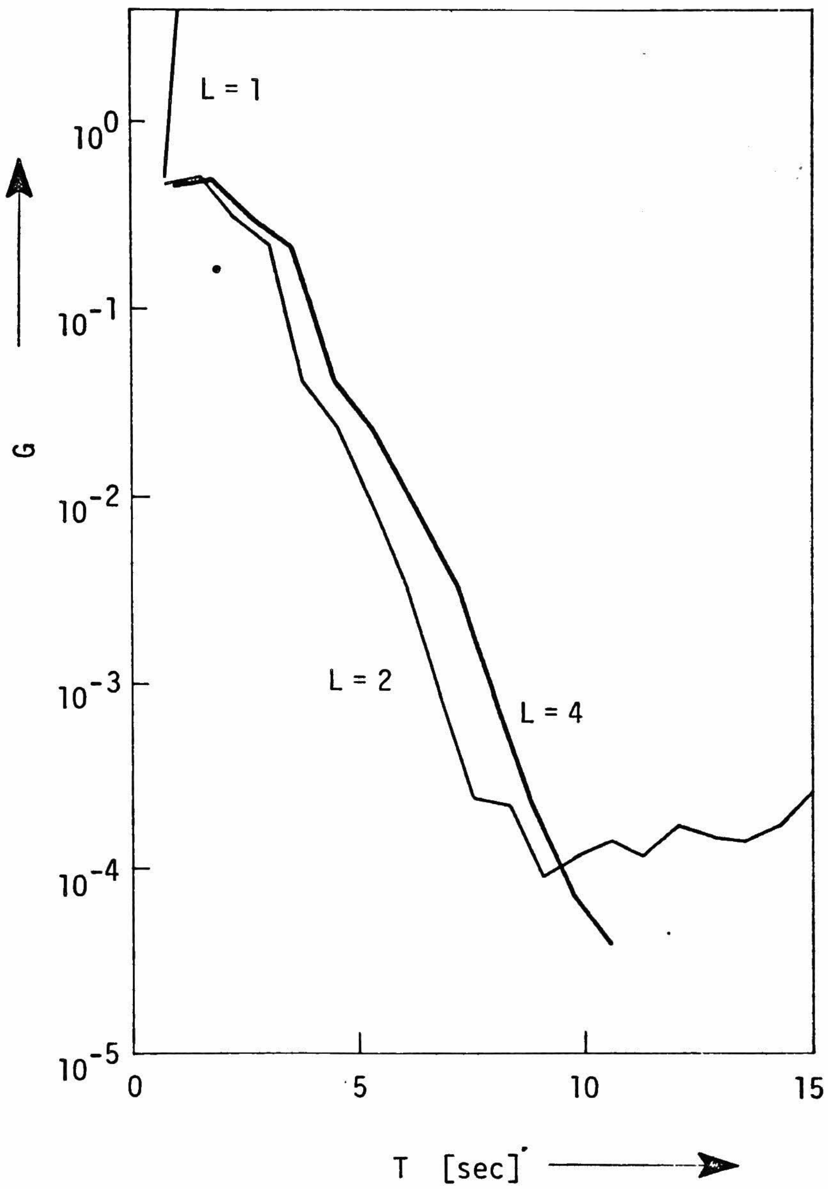

Fig. 3.2. Variations of G in case of the quasi-linearization method with the computing time T for $\varepsilon = 3.2$ (Example 3.1).

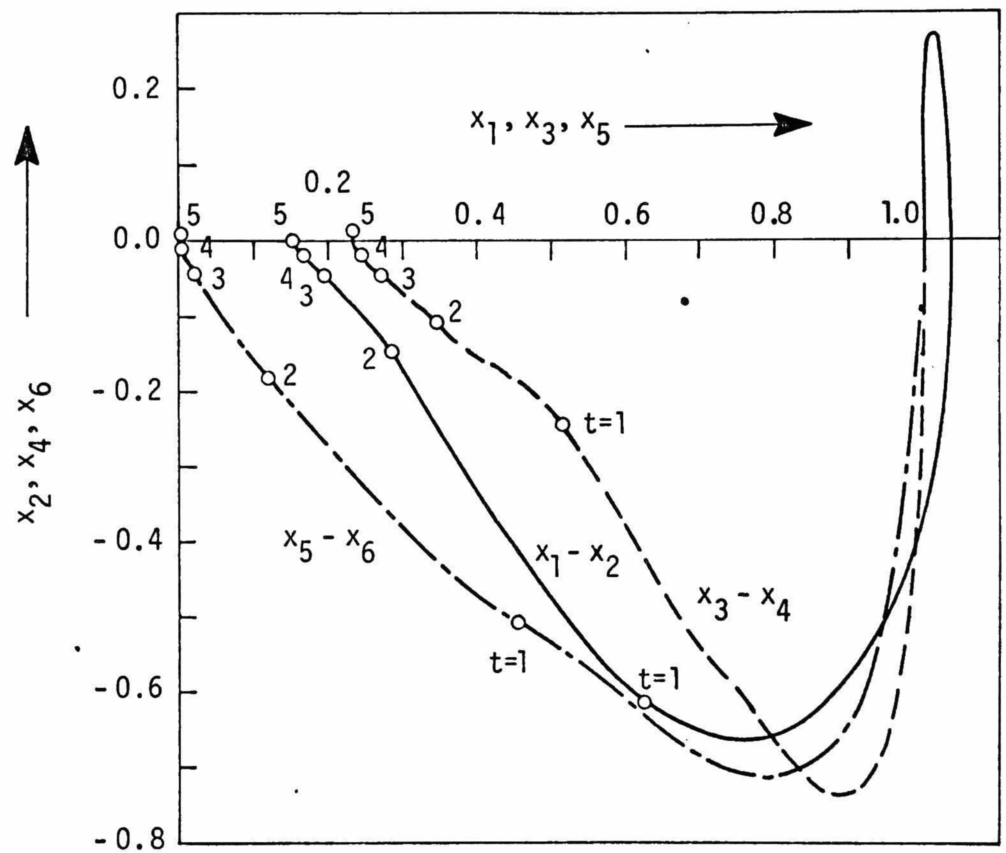

Fig. 3.3. The optimal trajectory on the $x_i - x_{i+1}$ plane ($i = 1, 3, 5$: $\varepsilon = 6$; Example 3.1).

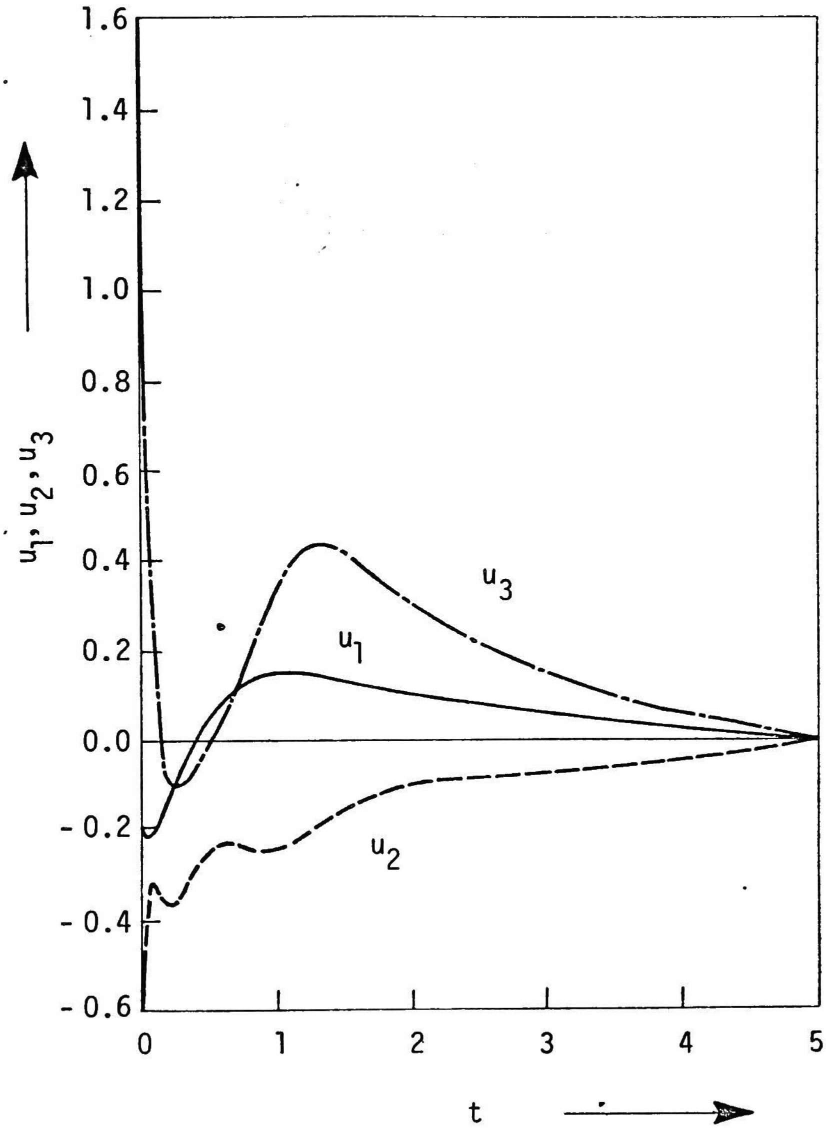

Fig. 3.4. Time history of the optimal control u for $\varepsilon = 6$ (Example 3.1).

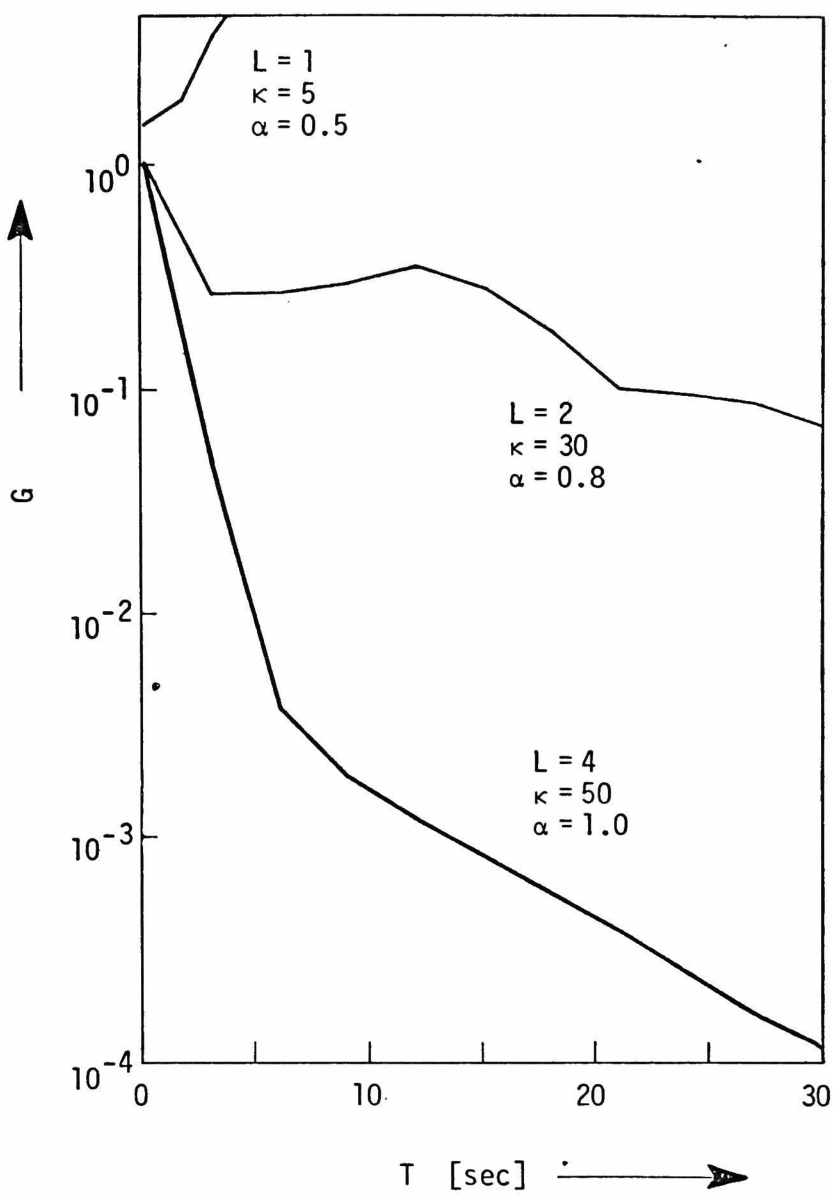

Fig. 3.5. Variations of G in case of the interaction-coordination algorithm with the computing time T for $\varepsilon = 4$ and $\beta = 1$ (Example 3.2).

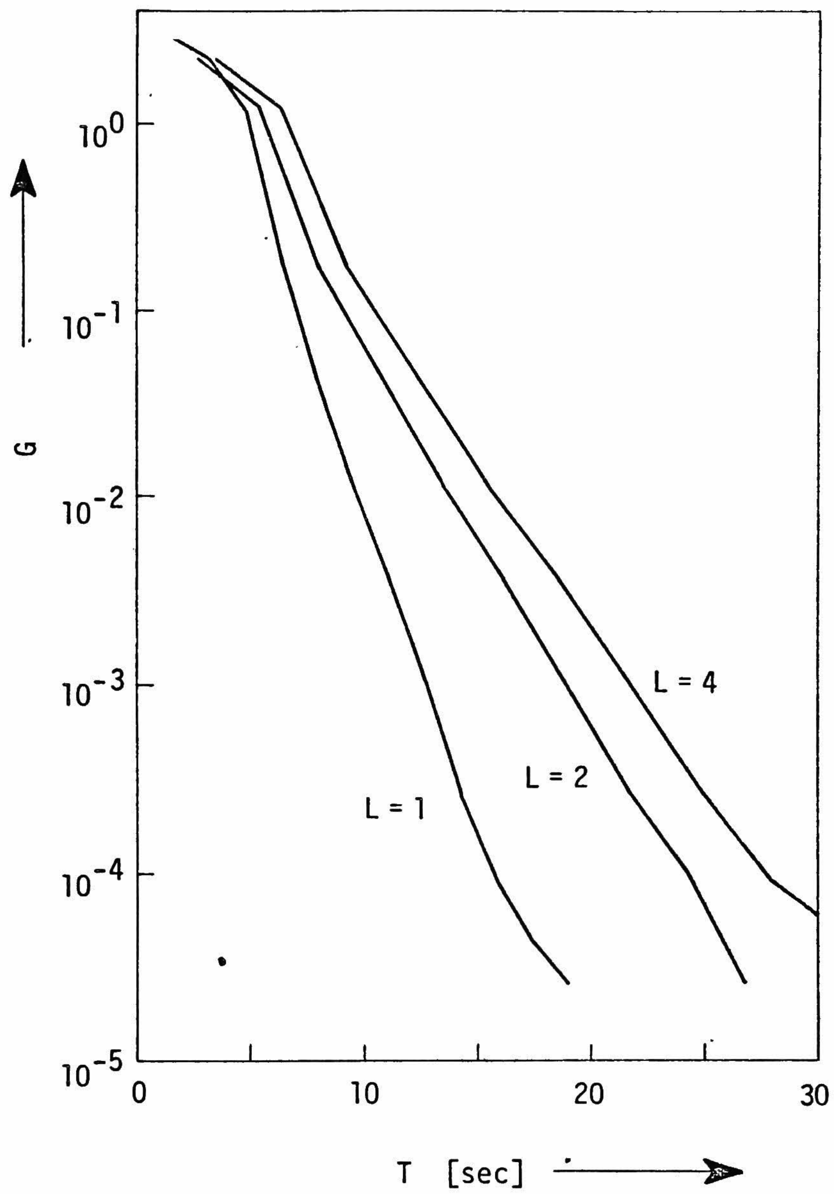

Fig. 3.6. Variations of G in case of the quasi-linearization method with the computing time T for $\varepsilon = 4$ (Example 3.2).

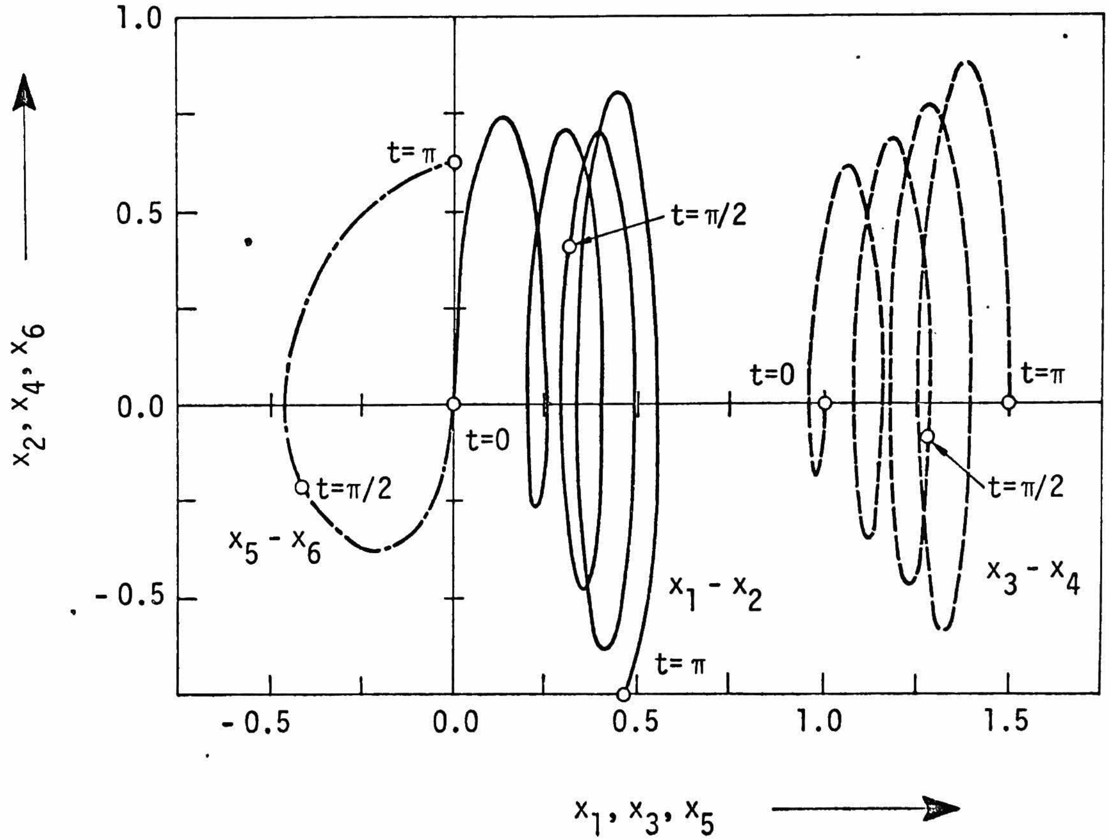

Fig. 3.7. The optimal trajectory on the $x_i - x_{i+1}$ plane ($i = 1, 3, 5$: $\varepsilon = 4$; Example 3.2).

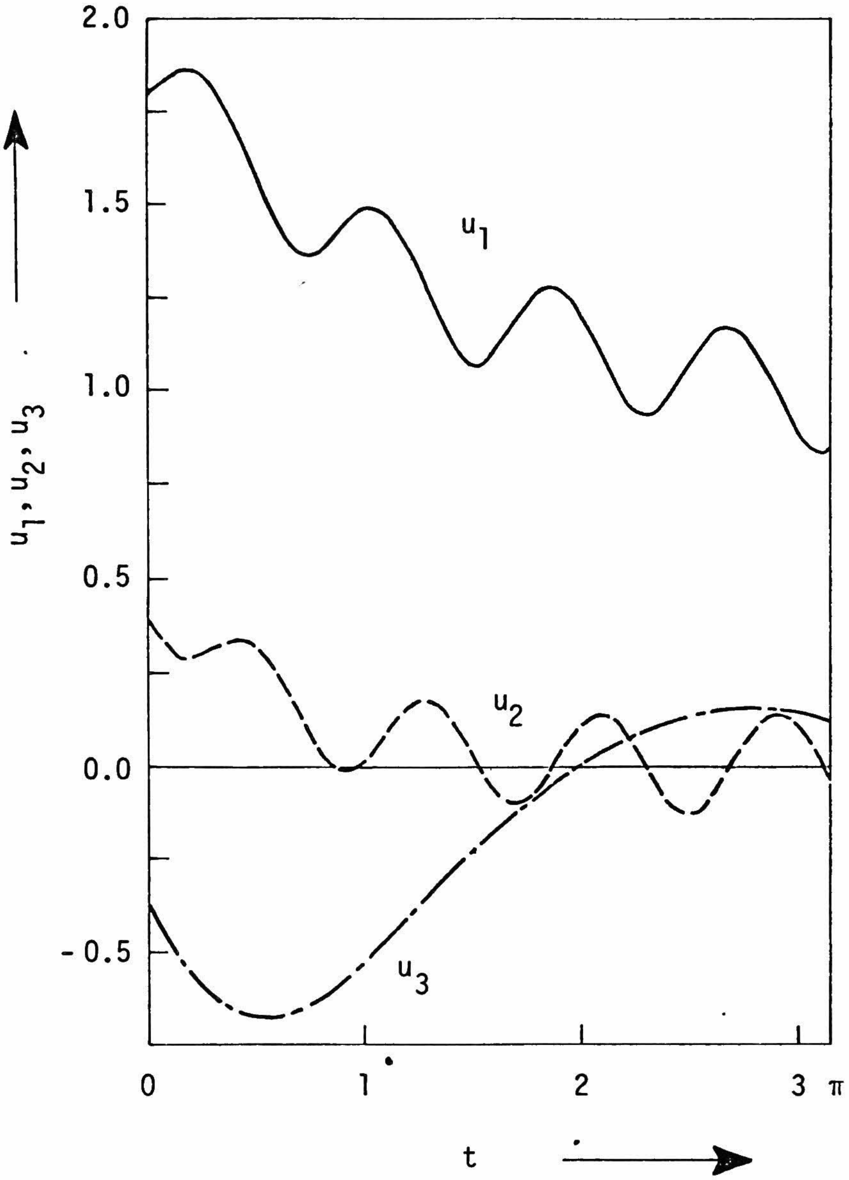

Fig. 3.8. Time history of the optimal control u for $\varepsilon = 4$ (Example 3.2).

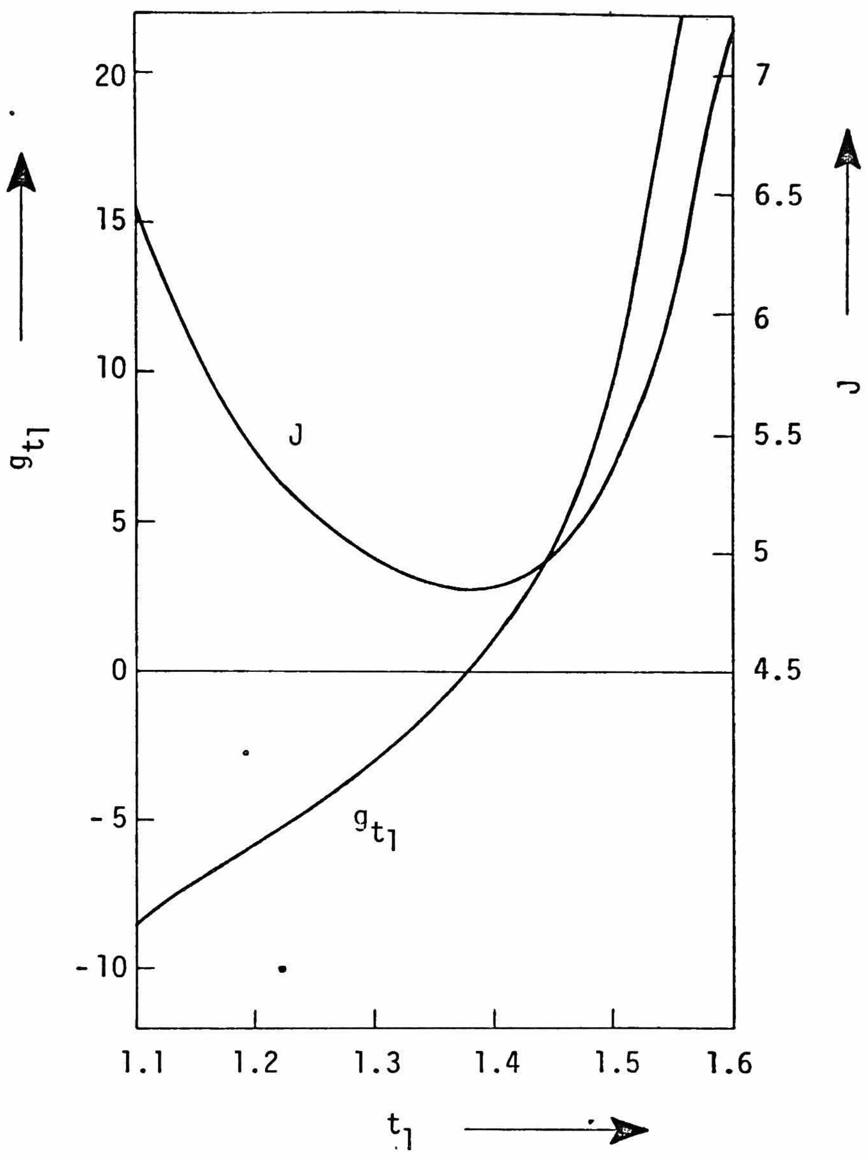

Fig. 3.9. Dependences of g_{t_1} and J on the corner time t_1 (Example 3.3).

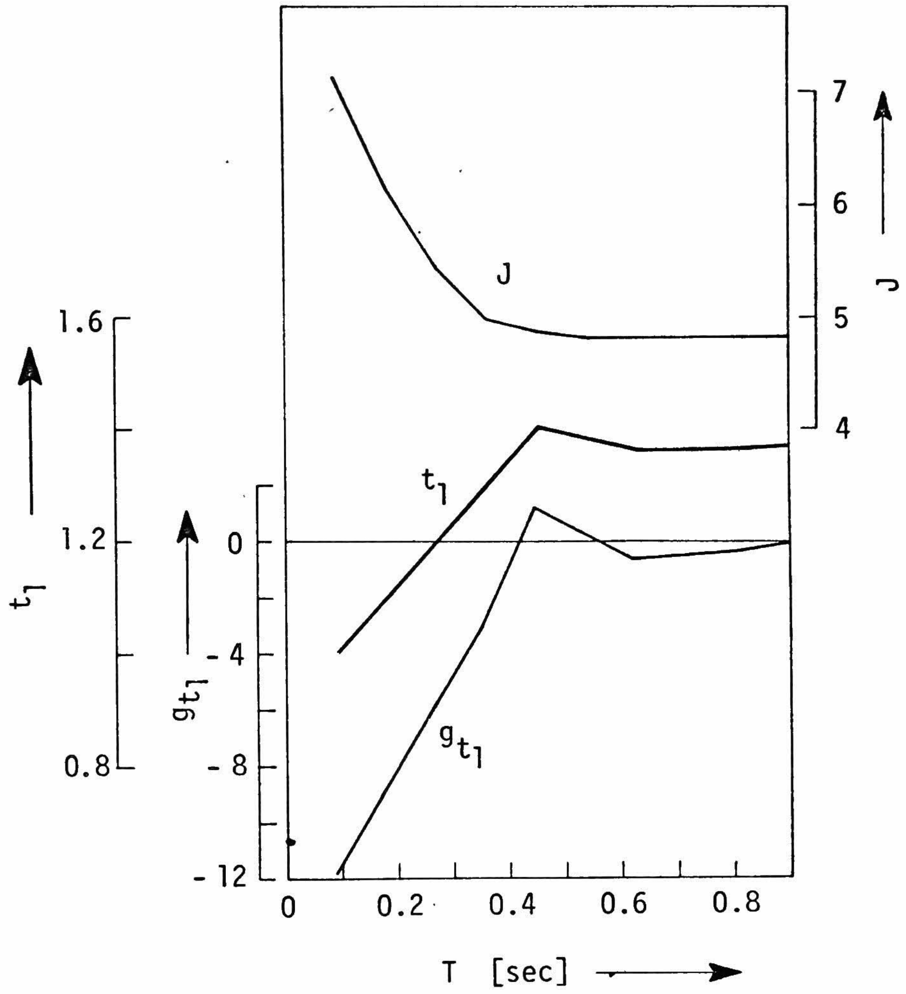

Fig. 3.10. Variations of t_1, g_{t_1}, and J with the computing time T (Example 3.3)

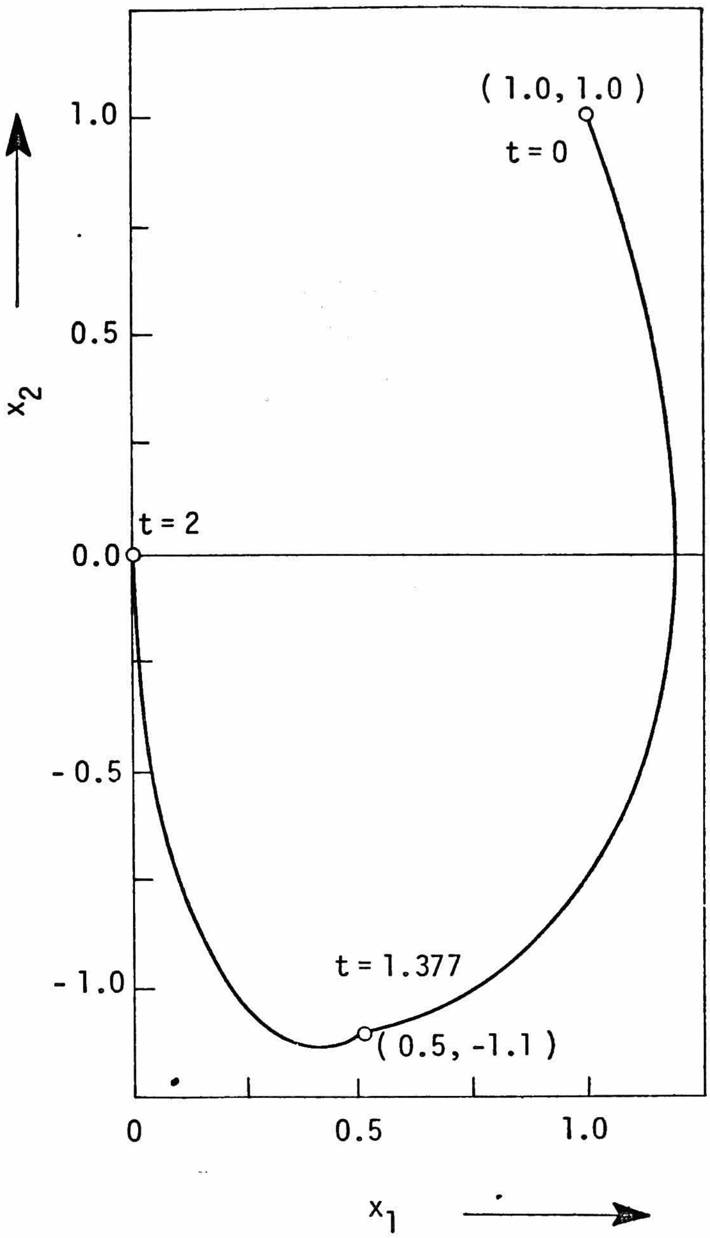

Fig. 3.11. The optimal trajectory on the $x_1 - x_2$ plane (Example 3.3).

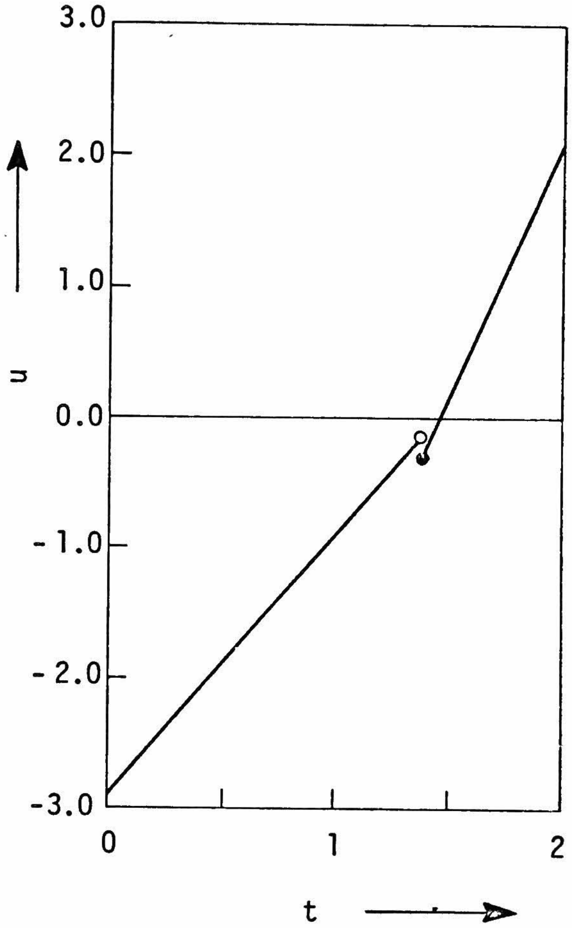

Fig. 3.12. Time history of the optimal control u (Example 3.3).

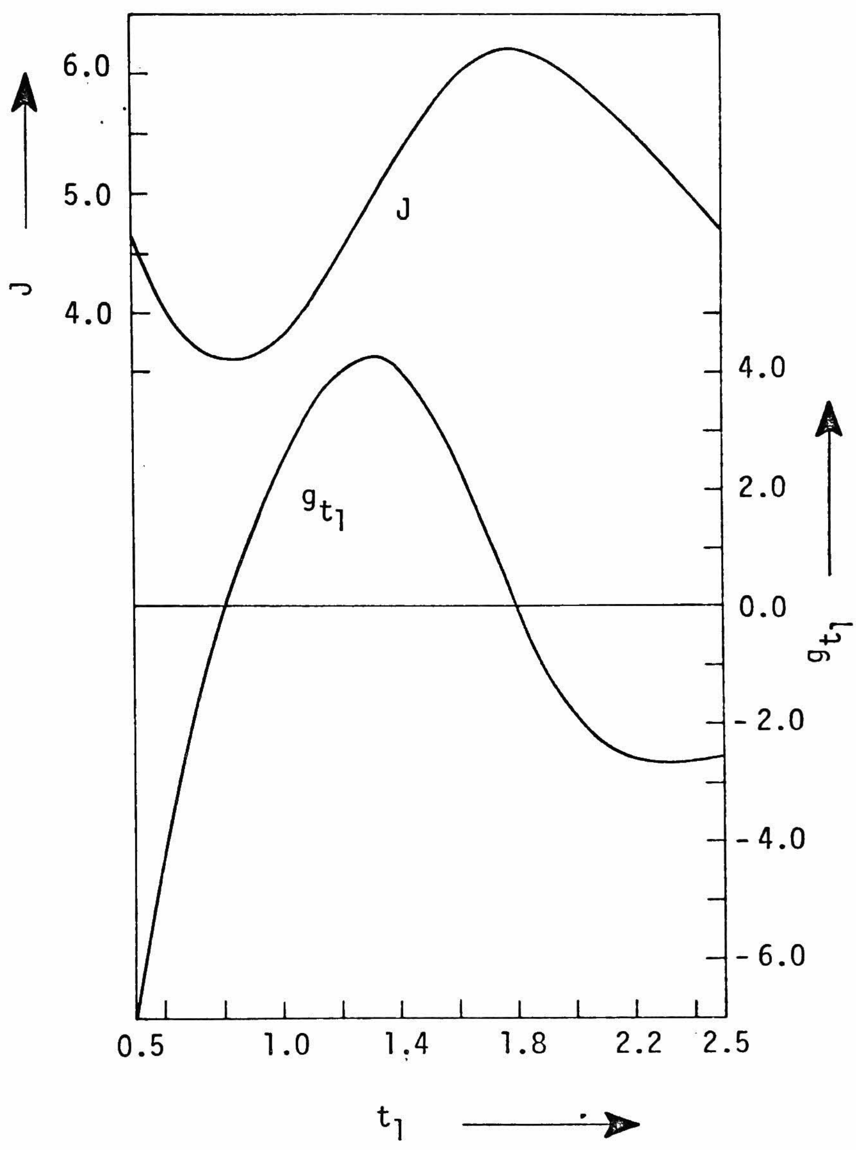

Fig. 3.13. Dependences of g_{t_1} and J on the corner time t_1 (Example 3.4).

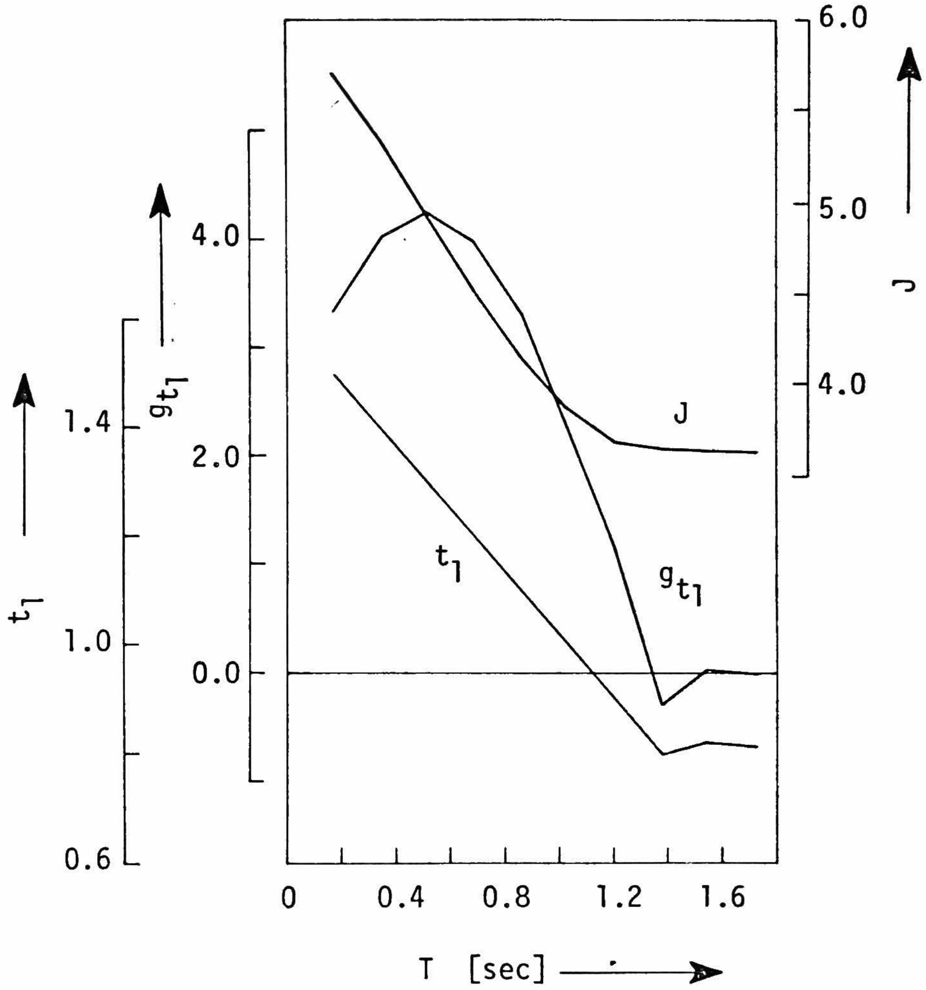

Fig. 3.14. Variations of t_1, g_{t_1}, and J with the computing time T (Example 3.4).

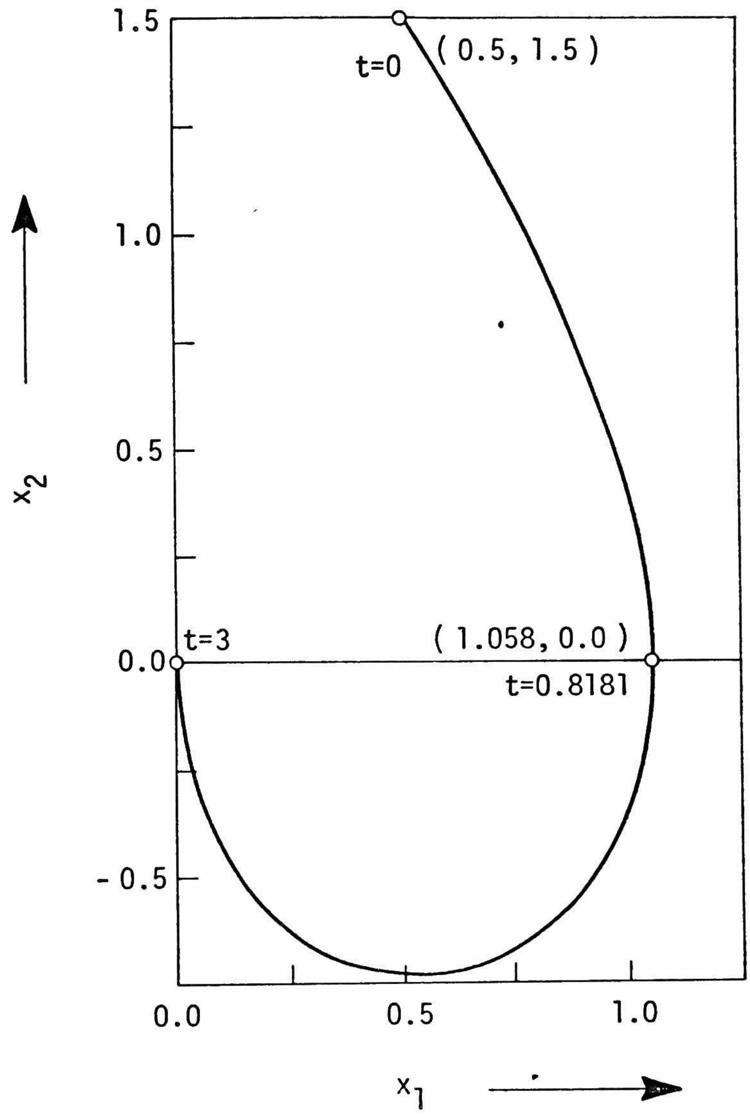

Fig. 3.15. The optimal trajectory on the $x_1 - x_2$ plane (Example 3.4).

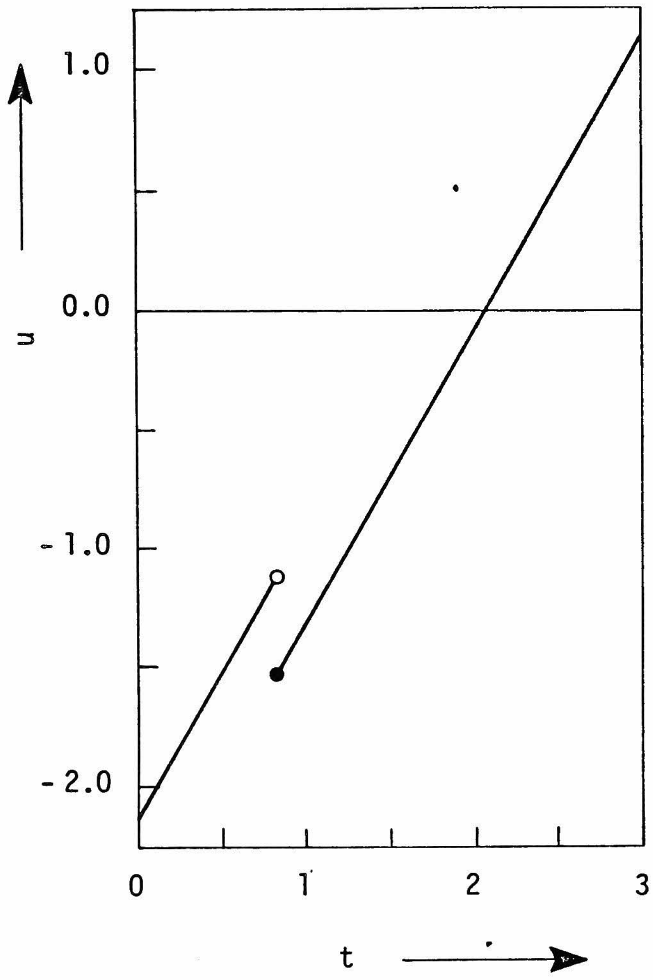

Fig. 3.16. Time history of the optimal control u (Example 3.4).

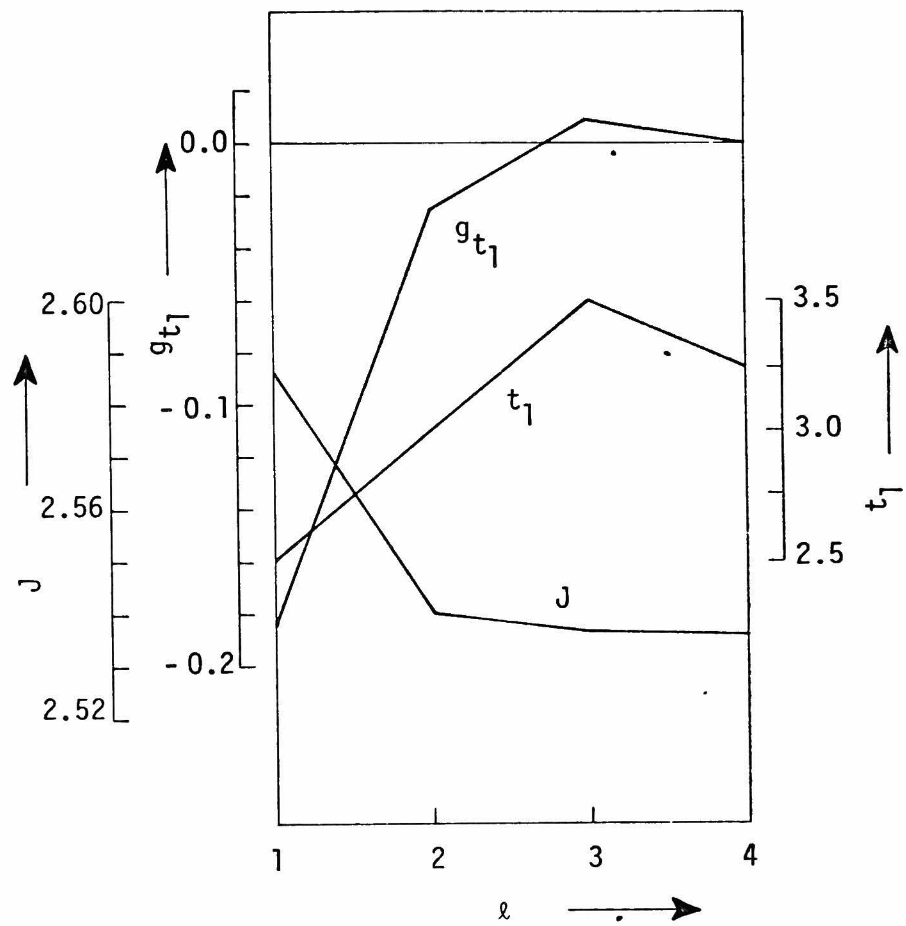

Fig. 3.17. Variations of t_1, g_{t_1}, and J with the iteration number ℓ (Example 3.5).

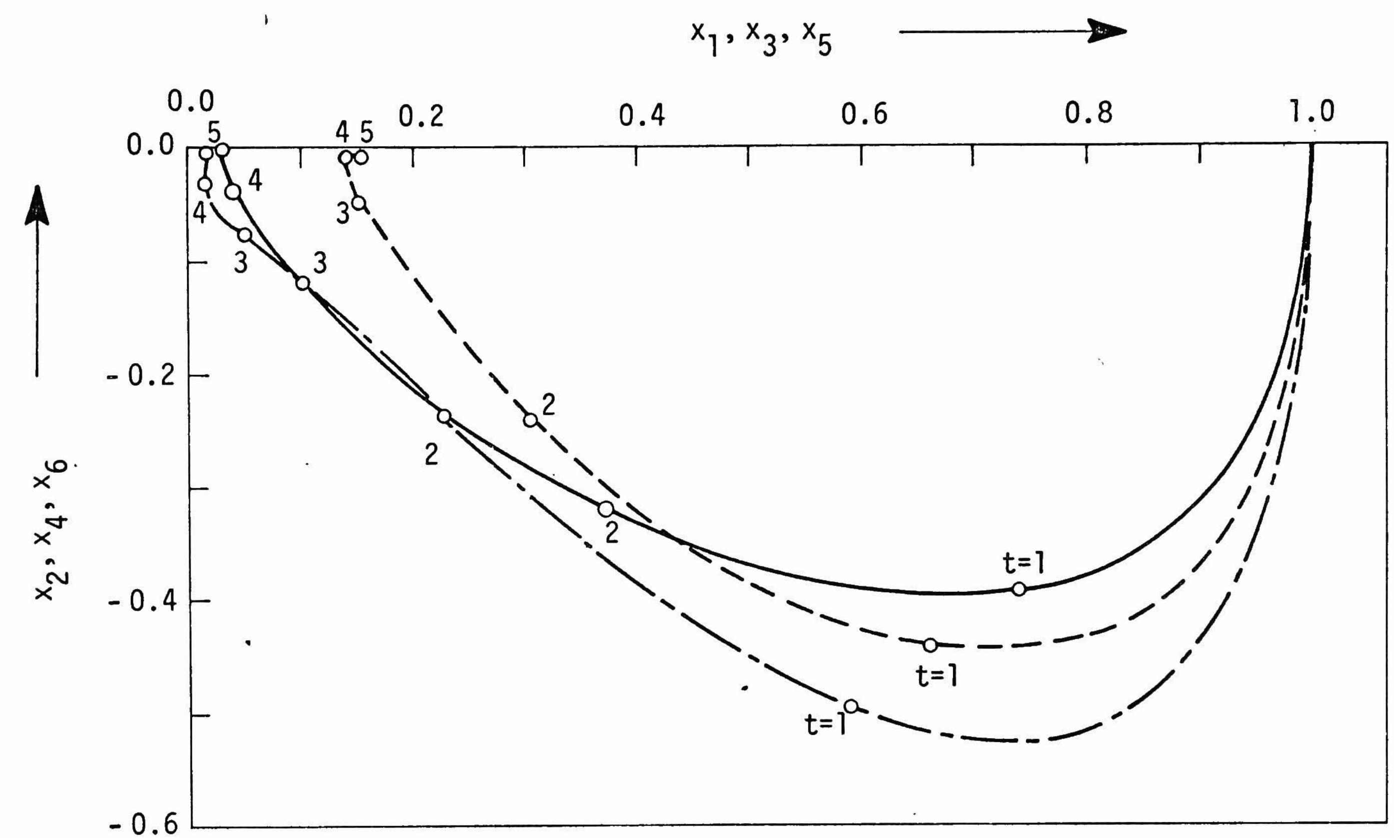

Fig. 3.18. The optimal trajectory on the $x_i - x_{i+1}$ plane ($i = 1, 3, 5$; Example 3.5).

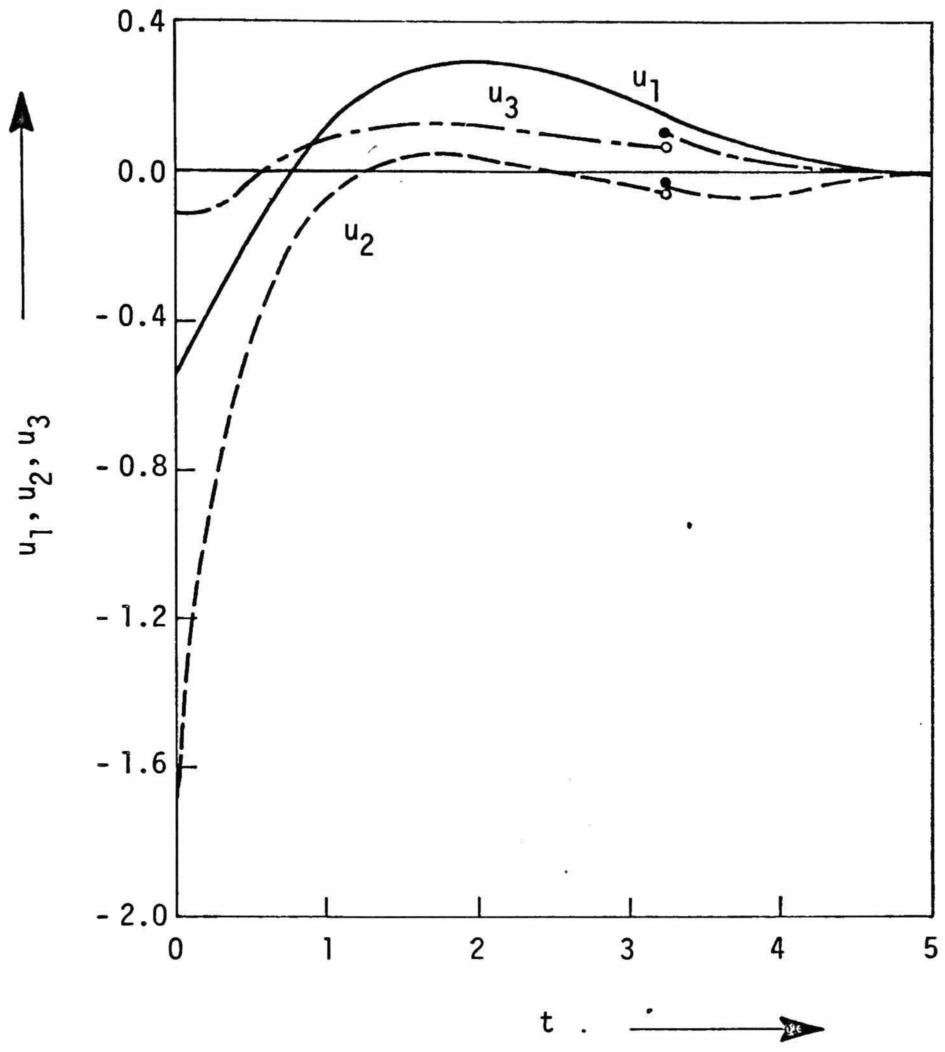

Fig. 3.19. Time history of the optimal control *u* (Example 3.5).

Chapter 4

A Modified Direct Method for Solving Nonlinear Optimal Control Problems with Control Constraint

4.1. Introduction

In the preceding chapter, we have considered nonlinear optimal control problems where no constraint is imposed on the control function. Such a problem is reduced to the nonlinear multipoint boundary-value problem and then we have shown that the algorithm combining the time-decomposition algorithm and the linearization method is effective for its solution.

In a practical sense, however, a certain constraint is often imposed on the control function, for example, a constraint on the magnitude, or a constraint on the total energy. Theoretically speaking, the indirect method, the approach of reducing the problem into two-point (or multi-point) boundary-value problem (TPBVP or MPBVP) by the direct use of the minimum principle, can be employed for the determination of the optimal control. However, the derived TPBVP generally contains very strong non-linearities which result from necessary conditions for optimality and is often practically impossible to solve. Therefore, the indirect method is generally not employed except for the linear problem of small order.

The direct method, on the contrary, can easily be applied to such a problem since the method adjust the control function directly. However, when the terminal condition is specified, the method becomes less effective. For example, the steepest-descent method [35] suffers from poor accuracy of the solution and the method proposed by Bryson and Denham [10] involves additional integration procedures, including the integration of matrix differential equation, for determining the Lagrange multipliers. The solutions obtained by these methods do not satisfy the specified terminal condition until the optimum is attained. Thus, both the direct and the indirect methods as they are have defects to be settled and it is desirable to develop a new algorithm which remedies the defects.

The requirement for the new algorithm is to have the property that the control function is easily adjusted to satisfy the control constraint and, at the same time, to make the state of the system satisfy the specified boundary conditions. Recently, Miele et al. have proposed a sequential gradient-restoration algorithm [20, 24-26, 48, 50]. The algorithm is composed of the alternate succession of gradient phases and restoration phases. In the gradient phase, nominal functions x and u satisfying the differential equation and the boundary condition are varied so as to reduce the value of performance index. In the restoration phases, the obtained functions which satisfy the constraints to the first order are corrected so as to be consistent with all the differential equation and the boundary condition. At the end of restoration phase which involves several iterations, the solution is feasible one and this is one of major merits of the algorithm. Turning inside out, excessive restoration phases must be carried out to obtain the optimal solution.

In this chapter, another algorithm is proposed which has the property mentioned above. The basic idea of the algorithm is to combine the steepest-descent method with the interaction-coordination algorithm discussed in the preceding chapter. The system equation for the steepest-descent is modified to be a linear TPBVP with respect to the state and the costate variables by introducing interaction variables. In the modified equation, the control variable is explicitly contained. Once values of the control variable and the interaction variables are provided, the TPBVP is easily solved. Then, by using the obtained solution, these variables are corrected so that the control variable may satisfy the optimality condition and the interaction variables may agree with the corresponding state or costate variables. Contrary to the sequential gradient-restoration algorithm, the solution at each iteration is infeasible except for the final one. This is the defect that all the indirect methods have. But the excessive corrections as the sequential gradient-restoration algorithm has can be avoided to obtain the optimal solution.

The proposed algorithm, of course, can be applied to problems without control constraint.

In Section 4.2, a modified direct method is proposed for solving control-constrained problems. In Section 4.2.1, the problem is formulated. Combining the steepest-descent method with the interaction-coordination algorithm, the proposed algorithm is constructed in Section 4.2.2. The algorithm is summarized in Section 4.3.2. Section 4.2.4 discusses a sufficient condition for convergence.

In Section 4.3, the proposed algorithm is applied to several examples. The first and the second examples are the same examples as in Chapter 3 except for the additional control constraint. In the third example, a problem with additional state constraint is examined. The performance indices of these examples are taken to be quadratic in state and control. In the fourth, i.e., the final example, the problem whose performance index is given by the total fuel of control is considered.

4.2 A Modofied Direct method

4.2.1. Problem Statement

In this chapter, an optimal control problem formulated in state-space form is dealt with. The problem is to find the control function, constrained in magnitude, which steers the system from the initial state to the specified terminal state so as to minimize an associated performance index.

Consider a dynamical system governed by the following equation.

$$\dot{x} = A(t)x + B(t)u + f(t, x) , \tag{4.1}$$

where $x(t)$ is an n-dimensional state vector, $u(t)$ is an m-dimensional control vector. A and B are $n{\times}n$- and $n{\times}m$-dimensional matrices, respectively, continuous in time t. f is an n-dimensional nonlinear vector function of the class C^2 with respect to x and continous in t. The objective is to find the control u which transfers the state from $x(t_0) = \pi_0$ to $x(t_f) = \pi_f$ subject to (4.1) under the control constraint:

$$|u_i(t)| \leq M_i , \quad t \in [t_0, t_f] \qquad (i = 1, 2, \dots, m) \tag{4.2}$$

and, in so doing, minimizes the performance index:

$$J = \frac{1}{2}\int_{t_0}^{t_f} [x'Q(t)x + u'R(t)u]dt\,. \tag{4.3}$$

In the above, π_0 and π_f are both n-dimensional vectors prescribed, and the symmetric matrices $Q(t)$ and $R(t)$, both continuous in t, are positive semidefinite and positive definite, respectively. The initial time t_0 and the terminal time t_f are assumed to be fixed.

A necessary condition for optimality of the problem is derived by using the minimum principle. So long as (4.1) is satisfied, (4.3) is identical to

$$\hat{J} = \int_{t}^{t_f} \{\frac{1}{2}[x'Q(t)x + u'R(t)u] + p'(t)[A(t)x + B(t)u + f(t,x) - \dot{x}]\}dt \tag{4.4}$$

with an arbitrary n-dimensional costate vector $p(t)$. Define the Hamiltonian H as :

$$H = \frac{1}{2}(x'Qx + u'Ru) + p'(Ax + Bu + f)\,. \tag{4.5}$$

Then, the first variation $\delta\hat{J}$ of $\hat{J}$ is given by [12]:

$$\delta\hat{J} = p'(t_0)\delta x(t_0) - p'(t_f)\delta x(t_f) + \int_{t_0}^{t_f} \{(\frac{\partial H}{\partial x} + \dot{p}')\delta x + \frac{\partial H}{\partial u}\delta u\}dt\,. \tag{4.6}$$

Now let p satisfy

$$\dot{p} = -(\frac{\partial H}{\partial x})'\,, \tag{4.7}$$

then, since $x(t_0)$ and $x(t_f)$ are specified, $\delta\hat{J}$ is rewritten into

$$\delta\hat{J} = \int_{t_0}^{t_f} \frac{\partial H}{\partial u}\delta u dt\,. \tag{4.8}$$

Thus, the necessary condition for optimality is obtained as follows:

$$\dot{x} = Ax + Bu^* + f\,, \tag{4.9}$$

$$\dot{p} = -Qx - A'p - (\frac{\partial f}{\partial x})'p \tag{4.10}$$

with the boundary condition:

$$x(t_0) = \pi_0\,, \qquad x(t_f) = \pi_f\,, \tag{4.11}$$

where u^* satisfies

$$\text{icomp}[\frac{\partial H}{\partial u}\bigg|_{u^*(t)}] \begin{cases} > 0\,, & u_i^*(t) = -M_i \\ = 0\,, & |u_i^*(t)| < M_i \\ < 0\,, & u_i^*(t) = M_i \end{cases} \tag{4.12}$$

and icomp[·] denotes the i-th component of the argument vector.

When the terminal point $x(t_f)$ is not specified, the boundary condition at $t = t_f$ in (4.11) is replaced by $p(t_f) = 0$. In this situation, a common means to find such a control variable u^* is the steepest-descent method which adjusts the control variable directly until the optimum is attained.

However, when $x(t_f)$ is specified, the steepest-descent method can not be applied to the problem without certain modifications, as (4.1) becomes overdetermined. One typical technique is to reduce the problem to the one with unspecified terminal condition by introducing an appropriate terminal cost in the performance index, and another is the one proposed by Bryson and Denham [10]. However, the former often suffers from poor accuracy of the obtained solution and the latter involves additional integration procedure, including the integration of matrix differential equation, for determining the Lagrange multipliers.

4.2.2. Solution Procedure

As mentioned above, the direct method, as it is, is not effective to the problem with specified terminal condition. On the other hand, if the control variable is not constrained, the indirect method can effectively be applied to such a problem. Therefore, in this section, an idea of the indirect method is introduced to overcome the difficulty caused by the direct method. To begin with, let us summarize the steepest-descent method adopted in our method in the form of an algorithm for a problem with unspecified terminal condition.

Step 1: Set $k=1$ and assume ${}^{1}u(t)$ $t \in [t_0, t_f]$.

Step 2: Solve (4.9) with the initial condition $x(t_0) = \pi_0$ and $u^* = {}^{k}u$.

Step 3: Using the solution ${}^{k}x$ and ${}^{k}u$, solve (4.10) backward from t_f to t_0 with the initial condition $p(t_f) = 0$.

Step 4: Calculate $\frac{\partial H}{\partial u}$. If $\frac{\partial H}{\partial u}$ satisfies (4.12), ${}^{k}u$ is the optimal control. Then, the iteration is terminated. Otherwise proceed to *Step* 5.

Step 5: Correct the control variable as:

$$ {}^{k+1}u(t) = {}^{k}u(t) - \eta\, {}^{k}g(t)\,, \qquad t \in [t_0, t_f]\,, \tag{4.13} $$

where k denotes the iteration number, η is a positive step size and ${}^{k}g(t)$ is the function defined by

$$\text{icomp}[{}^k g(t)] = \begin{cases} 0, & \text{icomp}[\frac{\partial H}{\partial u}\Big|_{{}^k u(t)}] > 0 \text{ and } {}^k u_i(t) = -M_i, \text{ or} \\ & \text{icomp}[\frac{\partial H}{\partial u}\Big|_{{}^k u(t)}] < 0 \text{ and } {}^k u_i(t) = M_i, \\ \text{icomp}[\frac{\partial H}{\partial u}\Big|_{{}^k u(t)}], & \text{otherwise.} \end{cases} \quad (4.14)$$

When ${}^{k+1}u_i(t)$ thus obtained does not satisfy (4.2), it is replaced by the boundary value M_i or $-M_i$ in accordance with its sign. Replace k by $k+1$ and return to *Step* 2.

When the terminal point $x(t_f)$ is specified as $x(t_f) = \pi_f$, the above algorithm can not be applied as it is. In this case the problem is reduced to a problem with unspecified terminal condition by introducing a terminal cost into the performance index as follows:

$$\bar{J} = J + \zeta\{x(t_f) - \pi_f\}^2 , \quad (4.15)$$

where ζ is a penalty parameter. Then the terminal condition of p is given by $p(t_f) = 2\zeta\{x(t_f) - \pi_f\}$ and the above mentioned algorithm is utilized.

Now we construct an algorithm for solving the problem given by (4.1)∿(4.3) by combining the steepest-descent method with the interaction-coordination algorithm. Let us rewrite (4.9) and (4.10), using the interaction variables y and q, as follows:

$$\dot{x} = Ax - E(p - q) + f(t, y) + Bu , \quad (4.16)$$

$$\dot{p} = -\kappa Qx - A'p - (\frac{\partial f(t, y)}{\partial y})'q + (\kappa - 1)Qy , \quad (4.17)$$

where E is an arbitrary $n \times n$-dimensional matrix. We adopt $\beta BR^{-1}B'$ for it corresponding to (3.11). Once the control variable u and the interaction variables y and q are provided, (4.16) and (4.17) with the boundary condition (4.11) is a linear TPBVP. When the solution does not satisfy the interaction balance (3.14) and/or the optimality condition (4.12), the interaction variables are corrected according to (3.15) and the control vector $u(t)$ is also corrected according to (4.13).

Remark 4.1.

Due to Theorem 3.1, we see that the linear TPBVP of (4.11), (4.16), and (4.17) with $E = \beta BR^{-1}B'$ has a unique solution, provided that the pair (A, B) is controllable.

4.2.3. Summary of the Algorithm

We now summarize the result of the preceding subsection in the form of an algorithm.

Step 1: Set $k = 1$, and prescribe the parameters α, η, β, and κ. Assume the initial function ${}^1u$, ${}^1y$, and ${}^1q$.

Step 2: Solve the linear TPBVP as given by (4.11), (4.16), and (4.17).

Step 3: Upon use of the solution ${}^kx$ and ${}^kp$, compute

$$ {}^kG = \{\frac{1}{(t_f - t_0)}[\frac{1}{2n}\int_{t_0}^{t_f} {}^kr'(t)\,{}^kr(t)dt + \frac{1}{m}\int_{t_0}^{t_f} {}^kg'(t)\,{}^kg(t)dt]\}^{1/2}, \quad (4.18) $$

where ${}^kr$ and ${}^kg$ are defined by (3.14) and (4.14), respectively. To obtain ${}^kg$, $\frac{\partial H}{\partial u} = Ru + B'p$ is calculated with $u = {}^ku$ and $p = {}^{k+1}q$. If ${}^kG \leqq \sigma$ (σ : a small positive number prescribed), the calculation is terminated. Otherwise, proceed to *Step* 4.

Step 4: Correct ${}^ky$ and ${}^kq$ according to (3.15), ${}^ku$ according to (4.13).

Replace k by $k+1$ and return to *Step* 2.

Remark 4.2.

If the nonlinearity is weak, the solution to the unperturbed equation, that is, the system equation without nonlinearity, is close to that to the perturbed equation. Therefore, as the initial estimates ${}^1y$ and ${}^1q$, it is reasonable to choose the solutions x and p, respectively, to the linear TPBVP derived by the optimality condition for the unperturbed system. The initial estimate of u is then calculated by (4.13) and (4.14) with ${}^0u=0$, $\eta=1.0$, and $k=0$, using ${}^1q$.

4.2.4 Convergence Proof

In this section, we consider a sufficient condition for the convergence of the proposed algorithm in case that the control function is not constrained. The case that it is constrained remains unsolved.

Let $z=[x', p']'$ and $w=[y', q']'$, then (3.15), (4.11), (4.13), (4.16), and (4.17) can be rewritten as follows:

$$ {}^k\dot{z} = D_1{}^kz + D_2{}^kw + D_3{}^ku + h(t, {}^kw) , \tag{4.19} $$

$$ [I_n, 0]z(t_0) = \pi_0, \quad [I_n, 0]z(t_f) = \pi_f , \tag{4.20} $$

$$ {}^{k+1}w = {}^kw + \alpha{}^kr_1 , \tag{4.21} $$

$$ {}^{k+1}u = {}^ku - \alpha{}^kr_2 , \tag{4.22} $$

where

$$ {}^kr_1 = {}^kz - {}^kw , \tag{4.23} $$

$$ {}^{k}r_2 = {}^{k}u + D_4 {}^{k+1}w , \tag{4.24} $$

and

$$ D_1 = \begin{bmatrix} A, & -\beta BB' \\ -\kappa Q, & -A' \end{bmatrix} , \quad D_2 = \begin{bmatrix} 0, & \beta BB' \\ (\kappa-1)Q, & 0 \end{bmatrix}, \quad D_3 = \begin{bmatrix} B \\ 0 \end{bmatrix}, \quad D_4 = [0, B'] , \tag{4.25} $$

$$ h(t, w) = [f'(t, y), -q'(\frac{\partial f(t, y)}{\partial y})]' . $$

In (4.19) to (4.25), we assume $R(t)$ to be identity matrix without loss of generality and we set η of (4.13) equal to α.

We begin with some definitions.

Definition 4.1. The norm $\|x\|$ of vector function $x(t)$ is defined by $\|x\| \triangleq \max_{t\varepsilon[t_0, t_f]} [x'(t)x(t)]^{1/2}$ and the norm $\|A\|$ of matrix function $A(t, \tau)$ by $\|A\| \triangleq \max_{t,\tau \varepsilon [t_0, t_f]} [\text{trace } A(t, \tau)A'(t, \tau)]^{1/2}$.

Definition 4.2. $\Phi(t, \tau)$ denotes the transition matrix of the homogeneous part of (4.19).

Definition 4.3. Scalar quantities a_1 and a_2 denote $\|\Phi\|$ and $\max[\|D_2\|, \|D_3\|]$, respectively.

Definition 4.4. We define the closed region Ω and U as

$$ \begin{aligned} \Omega &= \{w \mid \|w - w^*\| \leq 2\delta\} , \\ U &= \{u \mid \|u - u^*\| \leq 2\delta\} , \end{aligned} \tag{4.26} $$

where w^* and u^* are the optimal solutions to the problem and δ is a positive constant.

On the function $h(t, w)$, we make the following assumption.

Assumption 4.1. For arbitrary w_1 and w_2 belonging to Ω, there exists a scalar constant b_1 such that

$$\|h(t, w_1) - h(t, w_2)\| \leqq b_1\|w_1 - w_2\| . \tag{4.27}$$

Therefore, for arbitrary u_1 and u_2, the following relation holds.

$$\|h(t, w_1) - h(t, w_2)\| \leqq b_1[\|w_1 - w_2\| + \|u_1 - u_2\|] . \tag{4.28}$$

Let us now consider the convergence of the algorithm.

The solution to (4.19) and (4.20) can be written as

$$z(t) = \Phi(t, t_0)z_0({}^k w, {}^k u) + \int_{t_0}^{t} \Phi(t, \tau)[D_2{}^k w + D_3{}^k u + h(\tau, {}^k w)]d\tau . \tag{4.29}$$

Taking (4.19) to (4.22), Definitions 4.1 to 4.4, and Assumption 4.1 into account, we see that there exists a constant b_2 such that

$$\|z_0(w_1, u_1) - z_0(w_2, u_2)\| \leqq b_2(t_f - t_0)[\|w_1 - w_2\| + \|u_1 - u_2\|] , \tag{4.30}$$

$$w_i \in \Omega , \quad u_i \in U \quad (i = 1, 2) .$$

Therefore, since

$$\begin{aligned} {}^{k+1}z - {}^k z = {} & \Phi(t, t_0)[z_0({}^{k+1}w, {}^{k+1}u) - z_0({}^k w, {}^k u)] + \\ & \int_{t_0}^{t} \Phi(t, \tau)[D_2({}^{k+1}w - {}^k w) + D_3({}^{k+1}u - {}^k u) + h(\tau, {}^{k+1}w) - h(\tau, {}^k w)]d\tau , \end{aligned} \tag{4.31}$$

we obtain

$$\|{}^{k+1}z - {}^{k}z\| \leq \|\Phi\|\cdot\|z_0({}^{k+1}w, {}^{k+1}u) - z_0({}^{k}w, {}^{k}u)\| +$$

$$\|\Phi\|(t_f - t_0)[\|D_2\|\cdot\|{}^{k+1}w - {}^{k}w\| + \|D_3\|\cdot\|{}^{k+1}u - {}^{k}u\| +$$

$$\|h(t, {}^{k+1}w) - h(t, {}^{k}w)\|]$$

$$\leq a_1(t_f - t_0)(a_2 + b_1 + b_2)[\|{}^{k+1}w - {}^{k}w\| + \|{}^{k+1}u - {}^{k}u\|]$$

$$= \alpha a_1(t_f - t_0)(a_2 + b_1 + b_2)[\|{}^{k}r_1\| + \|{}^{k}r_2\|]$$

$$\leq 2\alpha\mu{}^{k}\bar{r}\,, \tag{4.32}$$

where $\mu \triangleq a_1(t_f - t_0)(a_2 + b_1 + b_2)$, ${}^{k}\bar{r} \triangleq \max\,[\|{}^{k}r_1\|, \|{}^{k}r_2\|]$. By the definition,

$${}^{k}r_1 = {}^{k}z - {}^{k}w$$

$$= {}^{k}z - {}^{k-1}z + (1-\alpha){}^{k-1}r_1 \tag{4.33}$$

and

$${}^{k}r_2 = {}^{k}u + D_4{}^{k+1}w$$

$$= {}^{k-1}u - \alpha{}^{k-1}r_2 + D_4({}^{k}w + \alpha{}^{k}r_1)$$

$$= {}^{k-1}u + D_4{}^{k}w - \alpha{}^{k-1}r_2 + \alpha D_4{}^{k}r_1$$

$$= (1-\alpha){}^{k-1}r_2 + \alpha D_4{}^{k}r_1\,, \tag{4.34}$$

then,

$$\|{}^{k}r_1\| \leq 2\mu\alpha{}^{k-1}\bar{r} + |1-\alpha|\cdot\|{}^{k-1}\bar{r}_1\|$$

$$\leq [|1-\alpha| + 2\mu\alpha]{}^{k-1}\bar{r} \tag{4.35}$$

and

$$\|{}^{k}r_2\| \leqq |1-\alpha|\,\|{}^{k-1}r_2\| + \alpha\|D_4\|\,\|{}^{k}r_1\|$$

$$\leqq |1-\alpha|^{k-1}\bar{r} + \alpha a_2[\,|1-\alpha| + 2\mu\alpha]^{k-1}\bar{r}$$

$$= \{|1-\alpha| + \alpha a_2[\,|1-\alpha| + 2\mu\alpha]\}^{k-1}\bar{r}, \qquad (4.36)$$

therefore, we obtain

$${}^{k}\bar{r} = \max[\,\|{}^{k}r_1\|,\ \|{}^{k}r_2\|\,]$$

$$= \max\{|1-\alpha| + 2\mu\alpha,\ |1-\alpha| + \alpha a_2[\,|1-\alpha| + 2\mu\alpha]\}^{k-1}\bar{r}. \qquad (4.37)$$

Let

$$\phi(\alpha) = \max\{|1-\alpha| + 2\mu\alpha,\ |1-\alpha| + \alpha a_2[\,|1-\alpha| + 2\mu\alpha]\}, \qquad (4.38)$$

then, from the above discussion, we see that a sufficient condition for the mapping (4.19) to (4.22) to be a contraction mapping is that

($S1$): There exists an α ($0 < \alpha < 1$) such that $0 < \phi(\alpha) < 1$

and

($S2$): If ${}^{k}w \in \Omega$ and ${}^{k}u \in U$, then ${}^{k+1}w \in \Omega$ and ${}^{k+1}u \in U$

are to hold at the same time.

Now we consider the conditions ($S1$) and ($S2$). For ($S1$), we have the following lemma.

Lemma 4.1.

Assume that $\mu < \frac{1}{2}$ and $a_2 < \frac{1}{2\mu}$. Then ($S1$) holds.

Proof.

Assume that $\phi(\alpha) < 1$ holds for $\alpha = \alpha^*$ ($0 < \alpha^* < 1$). Then

$$1 - \alpha^* + 2\mu\alpha^* < 1\,, \tag{4.39}$$

$$1 - \alpha^* + a_2\alpha^*(1 - \alpha^* + 2\mu\alpha^*) < 1\,. \tag{4.40}$$

Therefore,

$$(1-2\mu)\alpha^* > 0\,, \tag{4.41}$$

$$\alpha^*\{a_2 - 1 - a_2(1-2\mu)\alpha^*\} < 0\,. \tag{4.42}$$

Since $0 < \alpha^* < 1$, (4.41) and (4.42) imply

$$\mu < \frac{1}{2}\,, \tag{4.43}$$

$$\frac{a_2 - 1}{a_2(1-2\mu)} < \alpha^* < 1\,. \tag{4.44}$$

From (4.44), we obtain

$$a_2 < \frac{1}{2\mu}\,. \tag{4.45}$$

Thus, the condition ($S1$) holds if μ and a_2 satisfy (4.43) and (4.45).

Q.E.D.

Next, we consider the condition ($S2$). In the following, we make additional assumptions:

Assumption 4.2. $\mu < \frac{1}{2}$ and $a_2 < \frac{1}{2\mu}$.

Assumption 4.3. $\|{}^1w - w^*\| \leqq \delta$ and $\|{}^1u - u^*\| \leqq \delta$.

Assumption 4.4. $0 < \phi(\alpha) < 1$.

Assumption 4.5. $\nu \triangleq \max[\|{}^1r_1\|, \|{}^1r_2\|] \leqq [1-\phi(\alpha)]\delta/\alpha$.

Then, the following lemma holds.

Lemma 4.2.

On the Assumptions 4.1 to 4.5, following relations hold.

$$ {}^{k}w \in \Omega , \quad {}^{k}u \in U , \tag{4.46} $$

$$ \|{}^{k+1}z - {}^{k}z\| \leqq 2\alpha\nu\mu[\phi(\alpha)]^{k-1}, \qquad (k = 1, 2, \ldots). \tag{4.47} $$

$$ {}^{k}\overline{r} \leqq \nu[\phi(\alpha)]^{k-1} \tag{4.48} $$

Proof.

We prove the lemma inductively.

First, by the assumptions,

$$ {}^{1}w \in \Omega , \quad {}^{1}u \in U , \quad {}^{1}\overline{r} = \max[\|{}^{1}r_1\|, \|{}^{1}r_2\|] = \nu \tag{4.49} $$

and from (4.32),

$$ \|{}^{2}z - {}^{1}z\| \leqq 2\alpha\mu{}^{1}\overline{r} \leqq 2\alpha\nu\mu , \tag{4.50} $$

thus, (4.46) to (4.48) hold for $k = 1$.

Second, we assume (4.46) to (4.48) hold up to $k = i$. Then,

$$ \begin{aligned} \|{}^{i+1}w - w^*\| &\leqq \|{}^{i+1}w - {}^{i}w\| + \|{}^{i}w - {}^{i-1}w\| + \cdots + \|{}^{1}w - w^*\| \\ &\leqq \alpha\sum_{\ell=1}^{i} \|{}^{\ell}r_1\| + \delta \leqq \frac{\alpha\nu}{1 - \phi(\alpha)} + \delta \leqq 2\delta , \end{aligned} \tag{4.51} $$

$$ \begin{aligned} \|{}^{i+1}u - u^*\| &\leqq \|{}^{i+1}u - {}^{i}u\| + \|{}^{i}u - {}^{i-1}u\| + \cdots + \|{}^{1}u - u^*\| \\ &\leqq \alpha\sum_{\ell=1}^{i} \|{}^{\ell}r_2\| + \delta \leqq \frac{\alpha\nu}{1 - \phi(\alpha)} + \delta \leqq 2\delta , \end{aligned} \tag{4.52} $$

and from (4.32) and (4.37),

$$ {}^{i+1}\bar{r} \leqq \phi(\alpha)\,{}^{i}\bar{r} \leqq \phi(\alpha)\nu[\phi(\alpha)]^{i-1} = \nu[\phi(\alpha)]^{i}, \tag{4.53} $$

$$ \|{}^{i+2}z - {}^{i+1}z\| \leqq 2\alpha\mu\,{}^{i+1}\bar{r} \leqq 2\alpha\nu\mu[\phi(\alpha)]^{i} . \tag{4.54} $$

(4.51) to (4.54) show that relations (4.46) to (4.48) hold for $k = i+1$. Thus, the lemma is proved.

Q.E.D.

We are now ready to establish the following theorem.

Theorem 4.1.

On the Assumptions 4.1 to 4.5, the sequences $\{{}^{k}z\}$ and $\{{}^{k}w\}$ converge to the limit function w^* uniformly, and the sequence $\{{}^{k}u\}$ to u^* as $k \to \infty$.

Proof.

From (4.48) of Lemma 4.2 and Assumption 4.4,

$$ \lim_{k\to\infty} \|{}^{k}z - {}^{k}w\| = \alpha \lim_{k\to\infty} \|{}^{k}r_1\| = 0 , \tag{4.55} $$

therefore, the sequence $\{{}^{k}z\}$ agrees to $\{{}^{k}w\}$ in the limit. Further, since

$$ \begin{aligned} \|{}^{k+m}w - {}^{k}w\| &\leqq \|{}^{k+m}w - {}^{k+m-1}w\| + \|{}^{k+m-1}w - {}^{k+m-2}w\| + \cdots + \\ &\quad \|{}^{k+1}w - {}^{k}w\| \\ &\leqq \alpha \sum_{\ell=k}^{k+m-1} \|{}^{\ell}r_1\| \leqq \frac{\alpha\nu[\phi(\alpha)]^{k-1}}{1-\phi(\alpha)} \end{aligned} \tag{4.56} $$

and

$$\| {}^{k+m}z - {}^{k}z \| \leq \| {}^{k+m}z - {}^{k+m-1}z \| + \| {}^{k+m-1}z - {}^{k+m-2}z \| + \cdots +$$

$$\| {}^{k+1}z - {}^{k}z \|$$

$$\leq 2\alpha\mu \sum_{i=k-1}^{k+m-2} [\phi(\alpha)]^{i} \leq \frac{2\alpha\mu[\phi(\alpha)]^{k-1}}{1-\phi(\alpha)} , \qquad (4.57)$$

the sequences $\{{}^{k}z\}$ and $\{{}^{k}w\}$ are Cauchy sequences. As ${}^{k}z$ and ${}^{k}w$ are continuous in t and the space of the function z and w with the definition of the norm is complete. Thus, the sequences $\{{}^{k}z\}$ and $\{{}^{k}w\}$ have w^* as a limit. From (4.22), then, $\{{}^{k}u\}$ has a limit u^*.

Q.E.D.

From the above theorem, it is seen that if the nonlinearity is weak and the control duration $(t_f - t_0)$ is sufficiently short, the algorithm converges with the initial estimates y, q, and u as recommended in Remark 4.2.

Remark 4.3.

When the control is constrained, the convergence is not ensured by the theorem. However, then, better convergence characteristics might be expected, since the norm of the control function is bounded smaller.

4.3. Illustrative Examples

In this section, some numerical problems are presented to show the effectiveness of the present algorithm. For the numerical integration, a fourth-order Runge-Kutta-Gill scheme is employed with the step size of the integration routine $\Delta t = 0.025$ (Example 4.1), $\pi/100$ (Example 4.2), 0.005 (Example 4.3), and 0.001 (Example 4.4). Example 4.4 is solved in double precision arithmetic. The initial estimates of y, q, and u are determined according to Remark 4.2.

Example 4.1 (Control-Constrained and Unspecified Terminal-Condition Problem).

We consider again Example 3.1 where the additional control constraint $|u_i(t)| \leqq M_i$ is imposed ($i=1, 2, 3$). Corresponding to (4.16) and (4.17), following equations are obtained for Subsystem 1:

$$\begin{aligned}
\dot{x}_1 &= x_2 , \\
\dot{x}_2 &= -\beta(p_2 - q_2) + \varepsilon[y_4(1+y_6) - y_3(\varepsilon y_3 q_2 - \varepsilon y_1 q_4 + q_6)] + u_1 , \\
\dot{p}_1 &= -\kappa(x_1 - y_1) - y_1 + \varepsilon^2[-y_3 q_2 q_4 + y_1(q_4^2 + q_6^2)] , \\
\dot{p}_2 &= -\kappa(x_2 - y_2) - p_1 - y_2 + \varepsilon[q_4(1+y_6) - y_4 q_6],
\end{aligned} \tag{4.58}$$

where $\beta(p_2 - q_2)$ in the right hand side of the second equation corresponds to $E(p_2 - q_2)$ of (4.16). Similar equations are obtained for Subsystems 2 and 3 but are omitted here.

By way of example, let $x_1(0) = x_3(0) = x_5(0) = 1$ and $x_2(0) = x_4(0) = x_6(0) = 0$ and the terminal condition $x(5)$ be not specified as same as Example 3.1. The bound M_i is set equal to 0.8 ($i=1, 2, 3$). The weights β and κ and the step size α of (3.15) and η of (4.13) are set equal to unity.

Variations of G defined by (4.18) with the computing time T is shown in Figure 4.1 for $\varepsilon = 0.1$. It takes 6 itrations for the proposed algorithm to attain convergence with the convergence criterion 5×10^{-5}. In order to check the effectiveness of the proposed algorithm, the problem is solved also by the steepest-descent method with the initial estimate of u being zero. The step size η is set to 0.01. For calculating G of the steepest-descent method, the first term in the right hand side of (4.18) is omitted. The proposed algorithm converges much faster than the steepest-descent method. In the figure, the variation of G by our method is also depicted where $\varepsilon = 1$, $\beta = 5$, and $\kappa = 1$. For this case, the steepest-descent method fails to converge. Figure 4.2 shows time histories of the optimal control u_2 in case of M_i being 0.8 and ∞ ($i = 1, 2, 3$) for $\varepsilon = 0.1$.

Example 4.2 (Control-Constrained and Specified Terminal-Condition Problem).

Next, we consider the problem of Example 3.2 with additional control constraint $|u_i(t)| \leqq M_i$ ($i = 1, 2, 3$). In this case the derived TPBVP for Subsystem 1 can be written as

$$\dot{x}_1 = x_2,$$

$$\dot{x}_2 = x_1 - \beta(p_2 - q_2) - 2\varepsilon y_4 - \varepsilon y_1/(y_1^2 + y_3^2 + y_5^2)^{3/2} + u_1,$$

$$\dot{p}_1 = -p_2 + \varepsilon[q_2(-2y_1^2 + y_3^2 + y_5^2) - 3y_1(y_3 q_4 + y_5 q_6)]/ \qquad (4.59)$$

$$(y_1^2 + y_3^2 + y_5^2)^{5/2} + \kappa(x_1 - y_1),$$

$$\dot{p}_2 = -p_1 - 2\varepsilon q_4 + \kappa(x_2 - y_2),$$

where $\kappa(x_1 - y_1)$ and $\kappa(x_2 - y_2)$ in the right hand side of the third and the fourth eqution, respectively, are artificially added terms to accelerate the convergence rate. Similar problems for Subsystems 2 and 3 are omitted here.

By way of example, let $x_1(0) = x_2(0) = x_4(0) = x_5(0) = x_6(0) = 0$, $x_3(0) = 1$, $x_2(\pi) = -0.75$, $x_3(\pi) = 1.5$, $x_4(\pi) = x_5(\pi) = 0$, $x_6(\pi) = \pi/5$, and $x_1(\pi)$ be free. The bound M_i on the control is set equal to 0.5 ($i = 1, 2, 3$).

The weight β and the step size α of (3.15) and η of (4.13) are set equal to unity.

In order to check the effectivenss of the present algorithm, the problem is solved also by the steepest-descent method with the initial estimate of u being zero. Since the steepest-descent method as it is can not be applied to specified terminal-condition problems, the problem is reduced to a free terminal-condition problem by adding a terminal cost in the performance index (3.36):

$$\bar{J} = J + \zeta\{[x_2(\pi)+0.75]^2 + [x_3(\pi)-1.5]^2 + x_4(\pi)^2 + x_5(\pi)^2 + [x_6(\pi)-\pi/5]^2\}, \qquad (4.60)$$

where ζ is a positive scalar parameter. Then, the algorithm of Section 4.2.2 is applied.

Figure 4.3 shows variations of G defined by (4.18) with the computing time T for $\varepsilon=1$ with $\kappa=0$ and $\kappa=1$. For calculating G of the steepest-descent method, the first term in the right hand side of (4.18) is omitted. It takes 61 ($\kappa=0$) and 55 ($\kappa=1$) iterations for the present algorithm to attain convergence with the criterion σ being 5×10^{-5}, while 722 iterations are needed for the steepest-descent method with $\eta=0.01$. For larger values of η, the steepest-descent method diverged.

For the steepest-descent method, ζ is set equal to unity and the resultant terminal values of x_2, x_3, x_4, x_5, and x_6 are -0.774, 1.435, 0.026, 0.037, and 0.51, respectively, which are far from the desired solutions. The greater ζ is, the more iterations it takes to converge.

Time histories of the optimal control u are shown in Figure 4.4

in case of M_i ($i = 1, 2, 3$) equal to 0.5 and ∞.

Example 4.3 (State- and Control-Constrained Problem) [31].

Now we consider a problem with constraints in both the state and the control is examined. The system dynamics is given by

$$\begin{aligned} \dot{x}_1 &= x_2, & x_1(0) &= 0, \\ \dot{x}_2 &= -x_2 + u, & x_2(0) &= -1, \end{aligned} \tag{4.61}$$

and the performance index is taken to be

$$J = \int_0^1 (x_1^2 + x_2^2 + 0.005u^2)dt . \tag{4.62}$$

In addition to the control constraint $|u(t)| \leqq M$, the following state constraint is imposed:

$$S(t, x_2) = x_2(t) - 8(t - 0.5)^2 + 0.5 \leqq 0 . \tag{4.63}$$

The problem without control constraint is a well known test problem to the algorithms for solving state-constrained problems [31].

By using a well known penalty function method [14, 41], the problem is reduced to an ordinary control-constrained problem. To begin with, define a penalty function L by:

$$L(S) = \begin{cases} 0, & (S \leqq 0) \\ P_e \cdot S^2, & (S > 0) \end{cases} \tag{4.64}$$

where P_e is a penalty parameter. Then, the follwoing TPBVP is obtained:

$$
\begin{aligned}
\dot{x}_1 &= x_2, & x_1(0) &= 0, \\
\dot{x}_2 &= -x_2 - \beta(p_2 - q_2)/0.01 + u, & x_2(0) &= -1, \\
\dot{p}_1 &= -2\kappa x_1 + 2(\kappa-1)y_1, & p_1(1) &= 0, \\
\dot{p}_2 &= -2\kappa x_2 - p_1 + p_2 + 2(\kappa-1)y_2 - \left.\frac{\partial L}{\partial x_2}\right|_{x_2=y_2}, & p_2(1) &= 0.
\end{aligned} \qquad (4.65)
$$

When the penalty parameter P_e tends to infinity, the solution to (4.65) satisfying the interaction balance (3.14) and the optimality condition (4.12) tends to that of the given constrained problem [37, 41].

As an example, let M be 5.0, $P_e = 125$, $\beta = 0.01$, $\kappa = 15$, $\alpha = \eta = 0.4$ in our method, and $\eta = 0.001$ in the steepest-descent method. Figure 4.5 shows variations of G with the computing time T. When the step size of the steepest-descent method is taken to be 0.005, it does not converge. The optimal trajectory x_2 in this case is shown in Figure 4.6, which satisfies the state constraint with sufficient accuracy. Figure 4.7 shows time histories of the optimal control u in case of M equal to 5.0 and ∞.

Example 4.4 (Minimum-Fuel Problem) [54].

Finally, let us examine a problem with another type of performance index. The objective is to minimize the following functional:

$$J = \int_0^{t_f} \sum_{i=1}^{2} |u_i(t)|\,dt \qquad (4.66)$$

subject to

$$\dot{x} = Ax + Bu, \qquad (4.67)$$

$$x(0) = \pi_0, \quad x(t_f) = 0, \qquad (4.68)$$

$$|u_i(t)| \leqq 1 \quad (i = 1, 2), \tag{4.69}$$

where $x(t)$ is a four-dimensional state vector and $u(t)$ is a two-dimensional control vector. The matrices A and B are given by

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & c \\ 0 & 0 & 0 & 1 \\ 0 & -c & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix}. \tag{4.70}$$

This problem is derived by linearizing the problem of a minimum-fuel control in a reaction gas jet system [54].

Subject to (4.16) and (4.17), we obtain the following TPBVP:

$$\begin{aligned} \dot{x}_1 &= x_2, \\ \dot{x}_2 &= cx_4 - \beta(p_2 - q_2) + u_1, \\ \dot{x}_3 &= x_4, \\ \dot{x}_4 &= -cx_2 - \beta(p_4 - q_4) + u_2, \\ \dot{p}_1 &= 0, \\ \dot{p}_2 &= -p_1 + cp_4, \\ \dot{p}_3 &= 0, \\ \dot{p}_4 &= -cp_2 - p_3. \end{aligned} \tag{4.71}$$

Here, the matrix E in (4.16) is taken to be $\beta BB'$. Due to the minimum principle, the necessary condition for optimality on the control variable is given by

$$u_i(t) = \begin{cases} \mathrm{sign}[p_i(t)], & |p_i(t)| \geqq 1 \\ 0, & |p_i(t)| < 1 \end{cases} \quad (i = 1, 2), \tag{4.72}$$

and therefore (4.13) is modified as

$$^{k+1}u_i(t) = \begin{cases} \mathrm{sign}[^{k+1}q_i(t)], & |^{k+1}q_i(t)| \geqq 1 \\ 0, & |^{k+1}q_i(t)| < 1 \end{cases} \quad (i = 1, 2), \tag{4.73}$$

where q_i is an interaction variable corresponding to p_i $(i = 1, 2)$.

As an example, let $\pi_0 = [0.1, 0.0, 0.15, 0.0]$, $t_f = 1$, $c = 0.1$. The weight β and the step size α are taken to be unity. After 8 iterations with 7.856[sec] computing time, the optimal switcing sheme is obtained as shown in Table 4.1.

Table 4.1. The switching scheme of the optimal control.

	$0 \leqq t < t_1$	$t_1 \leqq t < t_2$	$t_2 \leqq t < t_3$	$t_3 \leqq t < t_4$	$t_4 \leqq t \leqq 1$
u_1	-1	0	0	0	1
u_2	-1	-1	0	1	1

$(t_1 = 0.106,\ t_2 = 0.190,\ t_3 = 0.822,\ t_4 = 0.880)$

The value J of the performance index is 0.594. The switching times of the exact solution are given by $t_1 = 0.105$, $t_2 = 0.189$, $t_3 = 0.821$, and $t_4 = 0.880$ with $J = 0.592$. In calculating analytical values, the terms of order higher than the first in c are neglected [54]. Figure 4.8 shows optimal trajectory on the $x_1 - x_2$ plane.

4.4. Concluding Remarks

An algorithm combining the steepest-descent method with the interaction-coordination algorithm is proposed to solve the constrained optimal control problems in nonlinear systems. By introducing the interaction variables, the problem is reduced to a sequence of linear TPBVP's preserving explicitly the terms of the control variable in the system equations. Then, the interaction variables and the control variable are adjusted so as to attain the optimality, using their solutions.

Since the proposed algorithm adjusts the control variable directly, it is not difficult to treat a control-constrained problem. Moreover, unlike other direct methods, it can easily deal with terminal constraints, since at each iteration stage, the problem is reduced to a linear TPBVP. Thus, the algorithm is successfully applicable to the control-constrained problems with specified terminal condition as well as to those with unspecified terminal condition.

A sufficient condition for the convergence of the algorithm is derived for unconstrained problems. The convergence proof for constrained problems remains unsolved.

Illustrative examples show that the present algorithm converges much faster than the steepest-descent method, and that it can also deal with problems with constraints both in state and control by addition of a penalty function and with on-off type problems.

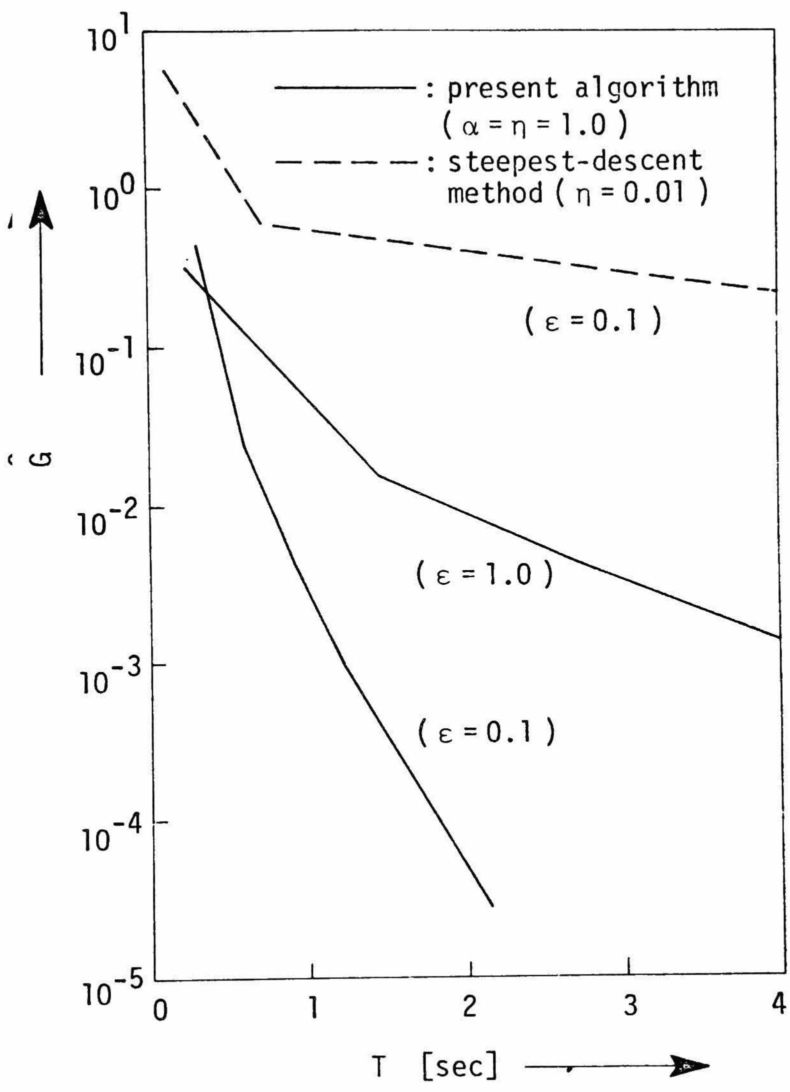

Fig. 4.1. Variations of G with the computing time T (Example 4.1).

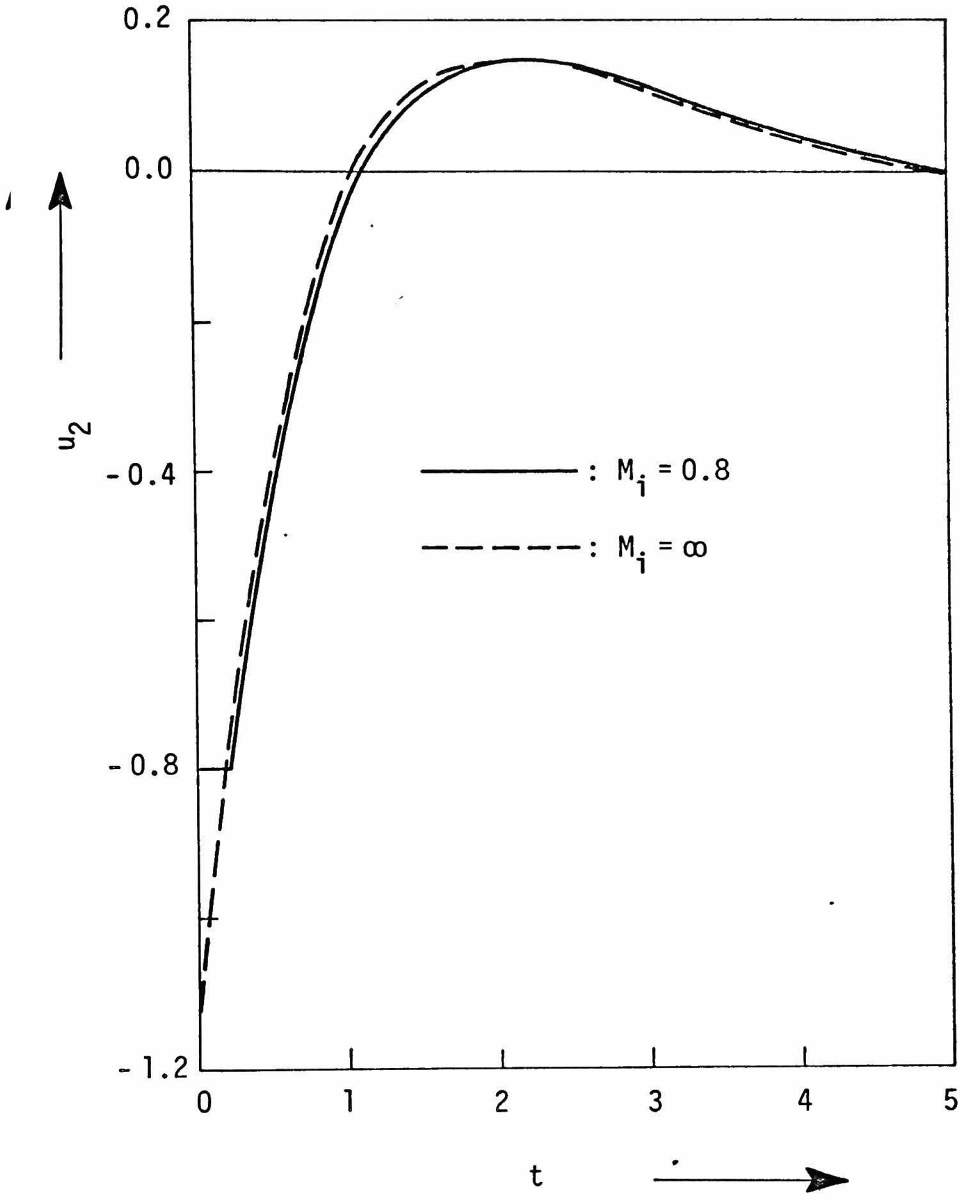

Fig. 4.2. Time histories of the optimal control u_2 (Example 4.1).

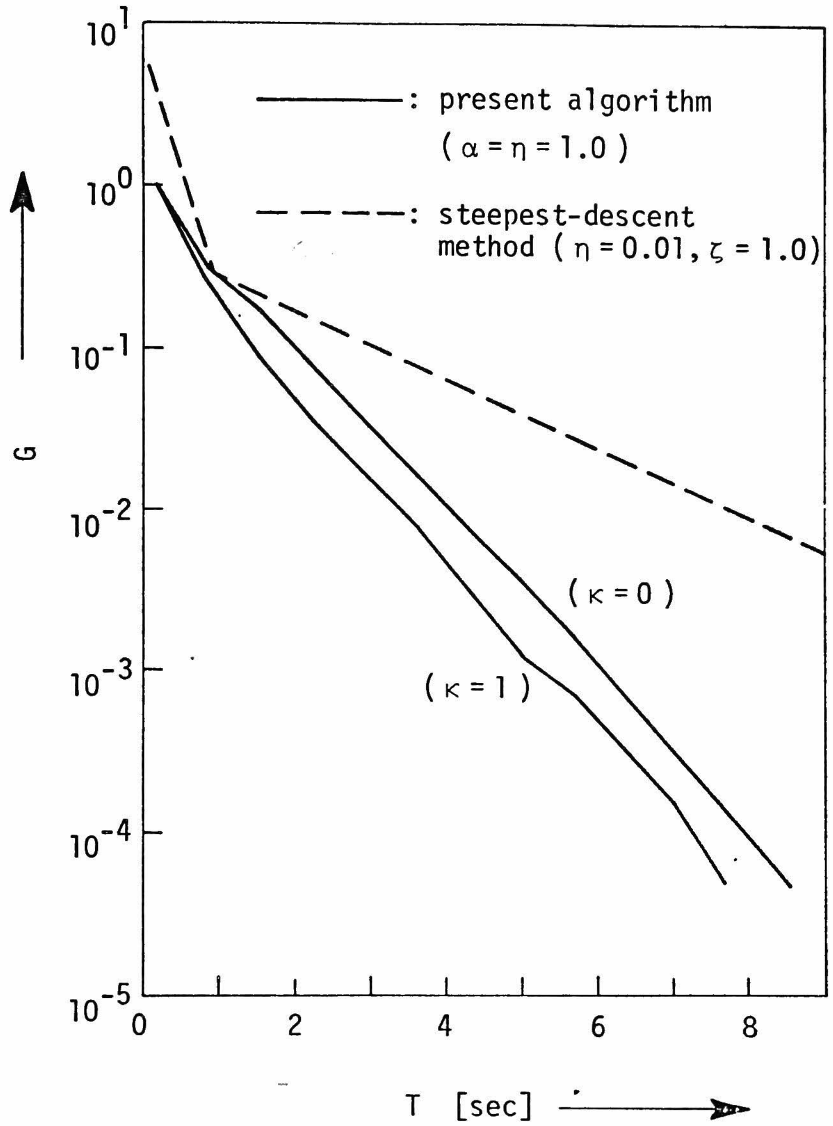

Fig. 4.3. Variations of G with the computing time T (Example 4.2).

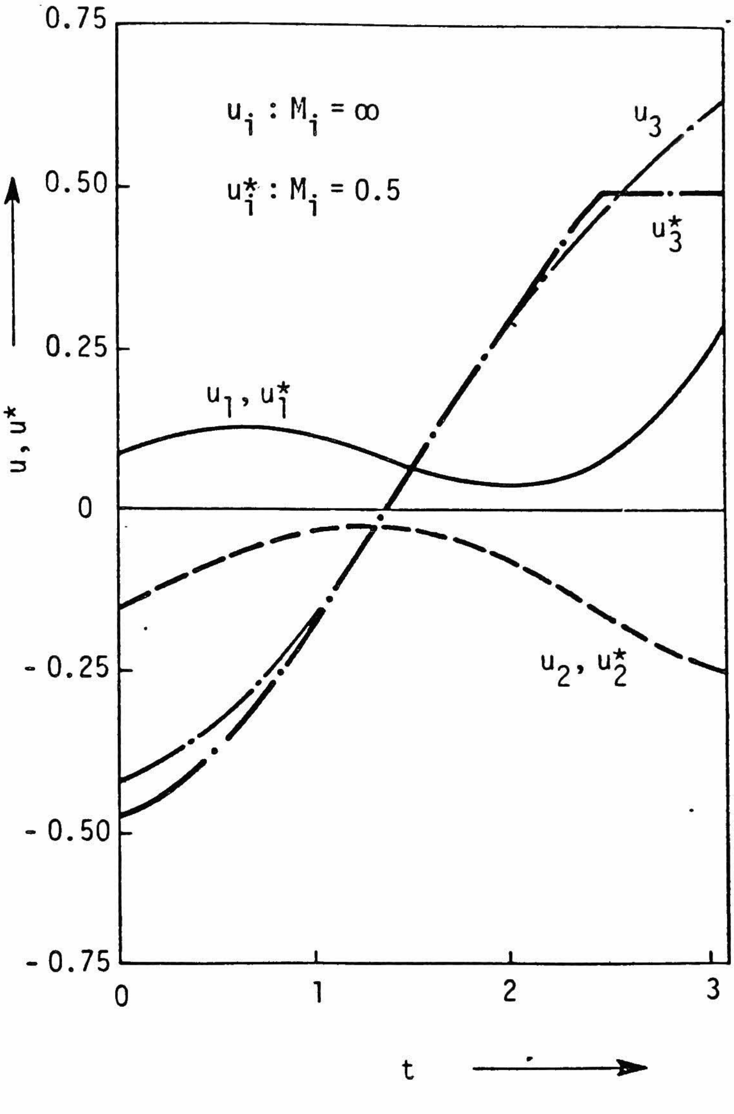

Fig. 4.4. Time hitories of the optimal control u (Example 4.2).

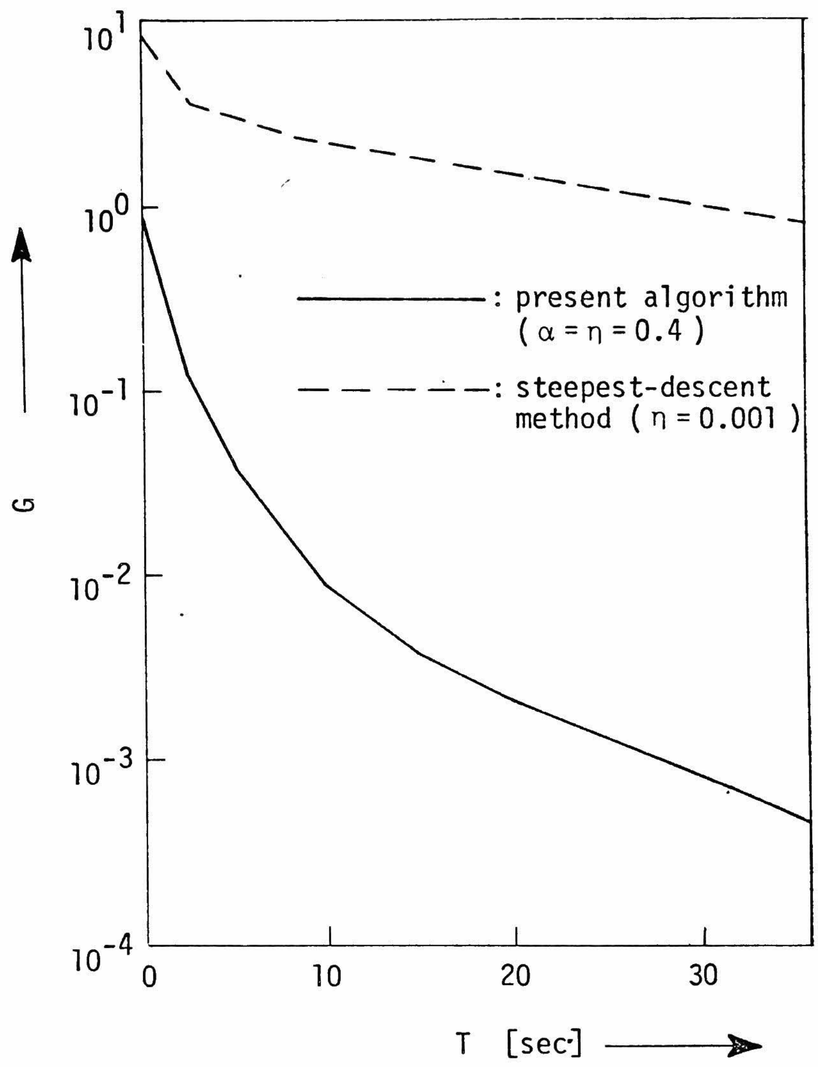

Fig. 4.5. Variations of G with the computing time T (Example 4.3).

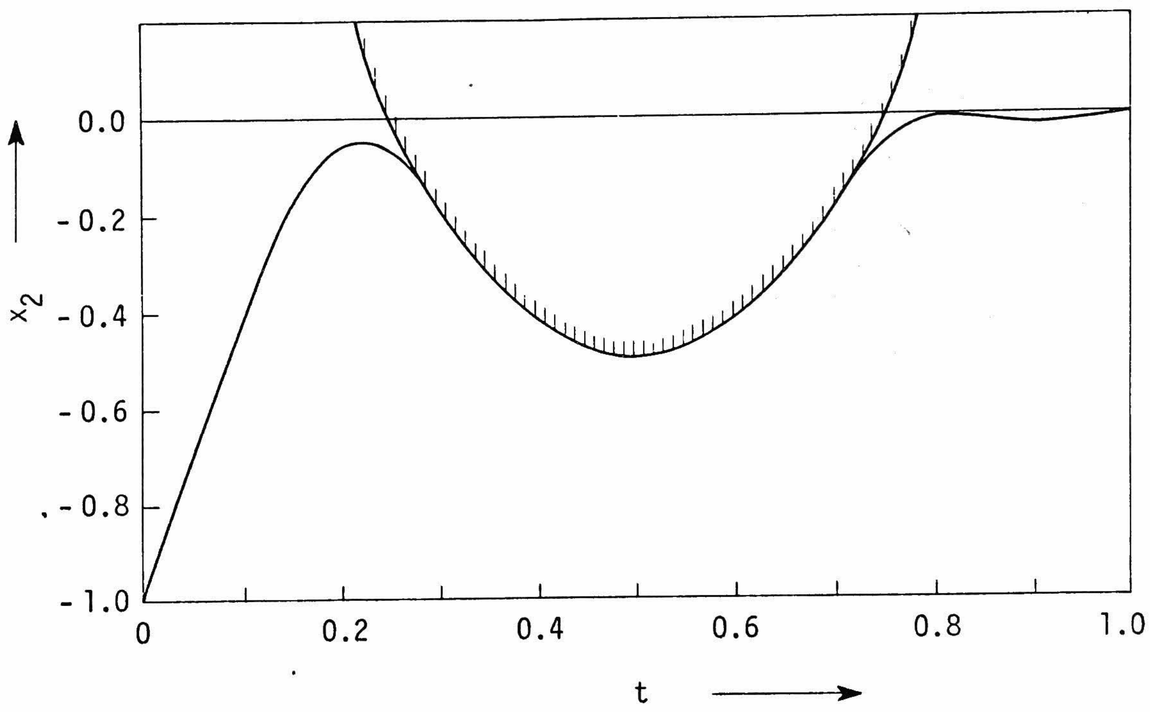

Fig. 4.6. The optimal trajectory of x_2 (Example 4.3).

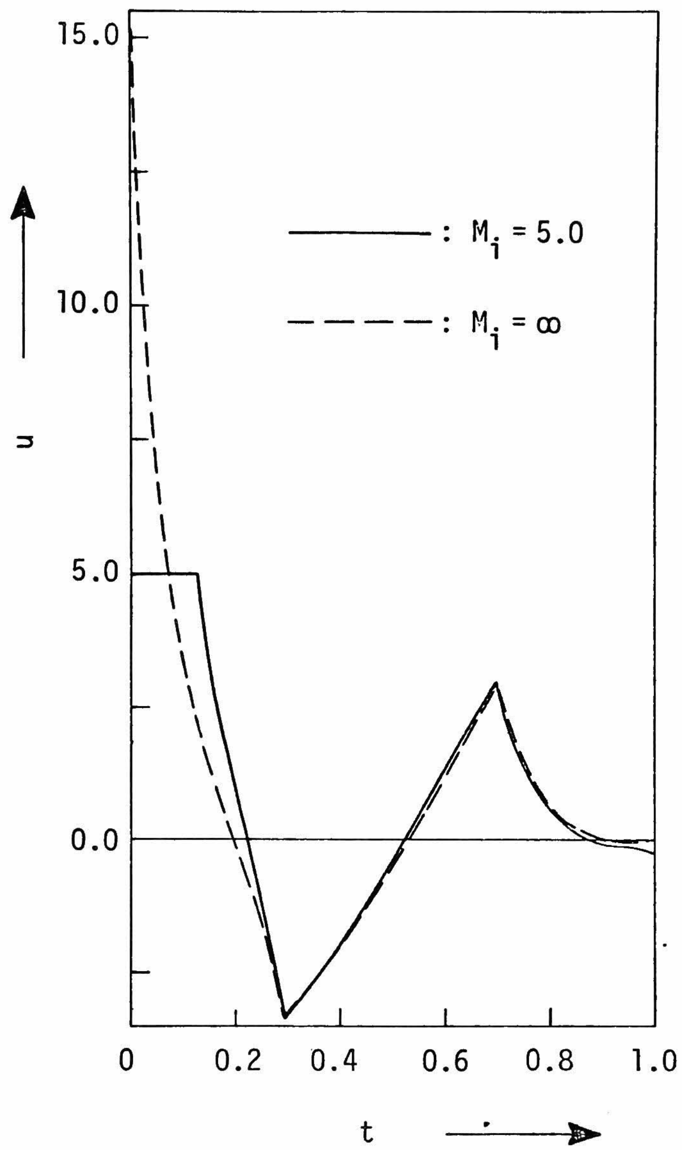

Fig. 4.7. Time histories of the optimal control u (Example 4.3).

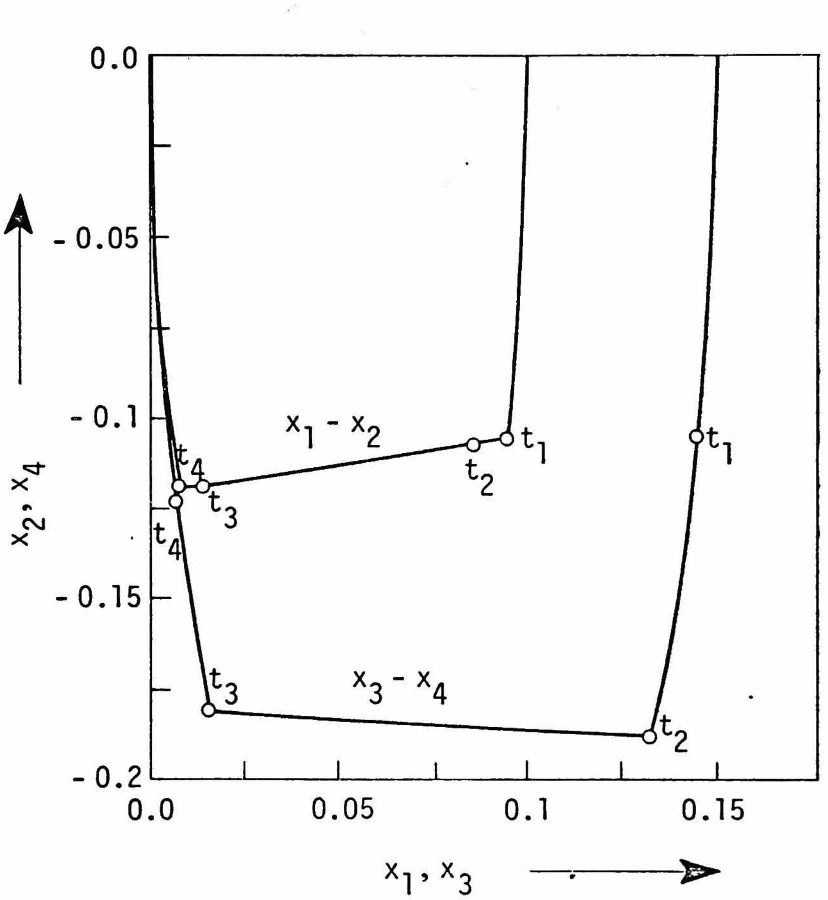

Fig. 4.8. The optimal trajectory on the $x_i - x_{i+1}$ plane ($i = 1, 2$; Example 4.4).

General Conclusion

In this text, we have considered the numerical solution of optmization problems in nonlinear systems. The developed algorithms cover many problems formulated in state-space form: with relatively strong nonlinearities and/or long control duration, with discontinuous parameters in system equations, and with control constraint.

In Chapter 2, the time-decomposition algorithm with multi-subintervals is developed for the solution of stiff linear TPBVP's. It is shown that the necessary and sufficient condition for the algorithm to be applicable is that the TPBVP has a unique solution in each subinterval and then in the overall interval. The algorithm is successfully applied to two illustrative examples. Based upon the hypothesis that the missolution of the problem by the superposition principle is due to the numerical error in taking the inverse of a transition matrix, we have made an error analysis through the example and shown that the time-decomposition algorithm reduces the error norm, i.e., the distance between the exact solution and the numerical solution, in the latter half of the integration interval, when the overall interval is divided into two subintervals.

In Chapter 3, nonlinear optimal control problems without control constraint are solved by the time-decomposition algorithm in conjunction with the linearization methods, i.e., the quasilinearization method and the interaction-coordination algorithm. In the former half of the chapter, problems with continuous parameters are considered. The problem is reduced to a nonlinear TPBVP by the minimum principle and further to a sequence of linear TPBVP's by the linearization method. It is shown that the time-decomposition algorithm can be applied to the problem so long as the linear part of the system equation is controllable. The effectiveness of the combined algorithm is illustrated through examining typical examples. The latter half of the chapter deals with problems with discontinuous parameters. Some of the state variables are also specified at the corner times at which the discontinuities occur. By dividing the overall interval at these corner times, the time-decomposition algorithm is applied. The optimal selection of the corner times is attained by the steepest-descent method, using thus obtained solution. Since the idea of penalty function is not employed, the solution satisfies the boundary conditions exactly and convergence is rapid.

The last chapter develops an algorithm for the solution of optimization problems with control constraint of saturation type. Adding the costate variable to and subtracting the corresponding interaction variable from the state equation, the problem is reduced to the solution of a sequence of linear TPBVP's. The interaction variables and the control variable are corrected so as to attain the optimality condition. Control constraint is easily made to be satisfied since the control variable is directly adjusted, and the specified boundary conditions are exactly satisfied. The algorithm is applied to several examples including a problem with additional state constraint and an on-off type problem, and satisfactory results are obtained.

As seen in the text, the optimal control is calculated based upon the perfect information of the system considered. However, it is generally impossible to identify practical systems perfectly, except for some systems in the aerospace engineering. Moreover, the optimal control is a critical control in the sense that a little identifying error of the parameters of the system may cause a serious effect to the action of the system.

This is a reason why recently many researchers have shifted their practical interest from the optmization problem as discussed in the text to the stabilization and the pole-assignment problems using state-feed back control.

Thus, to implement the optimal control theory, further reseaches must be made in the field of system identification theory and so called robust control theory.

References

1. Athans, M. and Falb, P. L.: "Optimal Control," McGraw-Hill Book Co., New York, 1966.

2. Baird, C. A., Jr.: Modified Quasilinearization Technique for the Solution of Boundary Value Problems for Ordinary Differential Equations, Journal of Optimization Theory and Applications, vol. 3, pp. 227-242 (1969).

3. Bauman, E. J.: Multilevel Optimization Techniques with Applications to Trajectory Decomposition, *in* Advances in Control Systems, vol. 6, Edited by C. T. Leondes, pp. 159-220, Academic Press, New York, 1968.

4. Bellman, R. E. and Kalaba, R. E.: "Quasilinearization and Nonlinear Boundary-Value Problems," American Elsevier Publishing Co., New York, 1965.

5. Bellman, R. E., Kalaba, R. E., and Prestrud, M. C.: "Invariant Embedding and Radiative Transfer in Slabs of Finite Thickness," American Elsevier Publishing Co., New York, 1963.

6. Beltrami, E. J.: "An Algorithmic Approach to Nonlinear Analysis and Optimization," Academic Press, New York, 1970.

7. Brayton, R., Gustavson, F., and Liniger, W.: A Numerical Analysis of the Transient Behavior of a Transistor Circuit, IBM Journal of Research and Development, vol. 10, pp. 292-299 (1966).

8. Breakwell, J. V.: The Optimization of Trajectories, Journal of SIAM, vol. 7, pp. 215-247 (1959).

9. Breakwell, J. V., Speyer, J. L., and Bryson, A. E.: Optimization and Control of Nonlinear Systems Using the Second Variation, Journal of SIAM on Control, Ser. A, vol. 1, pp. 193-223 (1963).

10. Bryson, A. E. and Denham, W. F.: A Steepest-Ascent Method for Solving Optimum Programming Problems, Journal of ASME, Ser. E, vol. 29, pp. 247-257 (1962).

11. Bryson, A. E., Denham W. F., and Dreyfus, S. E.: Optimal Programming Problems with Inequality Constraints I: Necessary Conditions for Extremal Solutions, AIAA Journal, vol. 1, pp. 2544-2550 (1963).

12. Bryson, A. E. and Ho, Y. C.: "Applied Optimal Control," Blaisdell Publishing Co., Waltham, Massachusetts, 1969.

13. Chua, L. O. and Lin, P. M.: "Computer-Aided Analysis of Electronic Circuits: Algorithm and Computational Technique," pp. 517-534, Prentice-Hall, Englewood Cliffs, N.J., 1975.

14. Courant, R.: Calculus of Variations and Supplementary Notes and Exercises (1945-1946), Revised and Amended by Moser, J., New York University Institute of Mathematical Science, New York, pp. 270-280 (1956-1957).

15. Denham, W. F. and Bryson, A. E.: Optimal Programming Problems with Inequality Constraints II: Solution by Steepest-Ascent, AIAA Journal, vol. 2, pp. 25-34 (1964).

16. Dennis, J. E., Jr.: On Newton-Like Methods, Numerische Mathematik, vol. 11, pp. 324-330 (1968).

17. Dyer, P. and McReynolds, S. R.: "The Computation and Theory of Optimal Control," Academic Press, New York, 1970.

18. Falb, P. L. and deJong, J. L.: "Some Successive Approximation Methods in Control and Oscillation Theory," Academic Press, New York, 1969.

19. Gear, C. W.: "Numerical Initial Value Problems in Ordinary Differential Equations," Prentice-Hall, Englewood Cliffs, N.J., 1971.

20. Gonzalez, S. and Miele, A.: Sequential Gradient-Restoration Algorithm for Optimal Control Problems with Nondifferential Constraints and General Boundary Conditions, Rice University, Aero-Astronautics Report No. 143 (1978).

21. Graham, R. G. and Leondes, C. T.: An Extended Quasilinearization Algorithm, Journal of Optimization Theory and Applications, vol. 12, pp. 268-284 (1973).

22. Graney, L.: The Numerical Solution of Unstable Two-Point Boundary-Value Problems, Journal of Computational and Applied Mathematics, vol. 3, pp. 173-180 (1977).

23. Guderley, K. G.: Some Remarks about Two-Point Boundary-Value Problems and Invariant Embedding, Journal of Optimization Theory and Applications, vol. 17, pp. 493-509 (1975).

24. Heideman, J. C. and Levy, A. V.: Sequential Gradient-Restoration Algorithm for Optimal Control Problems, Part 1, Theory, Journal of Optimization Theory and Applications, vol. 15, pp. 203-222 (1975).

25. Heideman, J. C. and Levy, A. V.: Sequential Gradient-Restoration Algorithm for Optimal Control Problems, Part 2, Examples, Journal of Optimization Theory and Applications, vol. 15, pp. 223-244 (1975).

26. Henning, G. R. and Miele, A.: Sequential Gradient-Restoration Algorithm for Optimal Control Problems with Bounded State, Journal of Optimization Theory and Applications, vol. 12, pp. 61-98 (1973).

27. Henrich, P.: "Discrete Variable Methods in Ordinary Differential Equations," John Wiley & Sons, New York, 1962.

28. Hestenes, M. R.: "Calculus of Variations and Optimal Control Theory," John Wiley & Sons, New York, 1966.

29. Horowitz, I. M. and Shaked, U.: Superiority of Transfer Function Over State-Variable Methods in Linear Time-Invariant Feedback System Design, IEEE Trans. on AC, vol. AC-20, pp. 84-97 (1975).

30. Ichikawa, K. and Tamura, K.: Optimal Problems with Discontinuity, Trans. Japan Soc. Mech. Engs., vol. 37, pp. 286-293 (1971)(in Japanese).

31. Jacobson, D. H. and Lele, M. M.: A Transformation Technique for Optimal Control Problems with a State Variable Inequality Constraint, IEEE Trans. on AC, vol. AC-14, pp. 457-464 (1969).

32. Kano, H.: Discontinuous Optimal Control Problems, Journal of Japan Soc. Mech. Engs., vol. 72, pp. 1034-1042 (1969)(in Japanese).

33. Keller, H. B.: "Numerical Methods for Two-Point Boundary-Value Problems," Blaisdell Publishing Co., Waltham, Massachusetts, 1968.

34. Kelley, H. J.: Gradient Theory of Optimal Flight Paths, Journal of the American Rocket Society, vol. 30, pp. 612-619 (1960).

35. Kelley, H. J.: Methods of Gradient, Chapter 6 of Optimization Techniques, Edited by G. Leitmann, Academic Press, New York, 1962.

36. Leitmann, G.: "An Introduction to Optimal Control," McGraw-Hill Book Co., New York, 1966.

37. Lele, M. M. and Jacobson, D. H.: A Proof of the Convergence of the Kelly-Bryson-Penalty Function Technique for State-Constrained Control Problems, Journal of Mathematical Analysis and Applications, vol. 26, pp. 103-109 (1969).

38. Liniger, W. and Willoughby, R. A.: Efficient Integration Methods for Stiff System of Ordinary Differential Equations, SIAM Journal on Numerical Analysis, vol. 7, pp. 47-66 (1970).

39. Marchal, C.: Second-Order Test in Optimization Theories, Journal of Optimization Theory and Applications, vol. 15, pp. 633-666 (1975).

40. Mataušek, M. R.: Direct Shooting Method for the Solution of Boundary-Value Problems, Journal of Optimization Theory and Applications, vol. 12, pp. 152-172 (1973).

41. McGill, R.: Optimal Control, Inequality State Constraints, and the Generalized Newton-Raphson Algorithm, Journal of SIAM on Control, Ser. A, vol. 3, pp. 291-298 (1965).

42. McGill, R. and Kenneth, P.: Solutions of Variational Problems by Means of a Generalized Newton-Raphson Operator, AIAA Journal, vol. 2, pp. 1761-1766 (1964).

43. Meyer, G. H.: "Initial Value Methods for Boundary Value Problems," Academic Press, New York, 1973.

44. Miele, A., Editor: "Theory of Optimum Aero-Dynamic Shapes," Academic Press, New York, 1965.

45. Miele, A.: Method of Particular Solutions for Linear Two-Point Boundary-Value Problems, Journal of Optimization Theory and Applications, vol. 2, pp. 260-275 (1968).

46. Miele, A.: Recent Advances in Gradient Algorithms for Optimal Control Problems, Journal of Optimization Theory and Applications, vol. 17, pp. 361-430 (1975).

47. Miele, A., Aggarwal, A. K., and Tietze, J. L.: Solutions of Two-Point Boundary-Value Problems with Jacobian Matrix Characterized by Large Positive Eigenvalues, Journal of Computational Physics, vol. 15, pp. 117-133 (1974).

48. Miele, A., Damoulakis, J. N., Cloutier, J. R., and Tietze, J. L.: Sequential Gradient-Restoration Algorithm for Optimal Control Problems with Nondifferential Constraints, Journal of Optimization Theory and Applications, vol. 13, pp. 218-255 (1974).

49. Miele, A. and Iyer, R. R.: Modified Quasilinearization Method for Solving Nonlinear Two-Point Boundary-Value Problems, Journal of Mathematical Analysis and Applications, vol. 36, pp. 674-692 (1971).

50. 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. 235-281 (1970).

51. Miele, A., Well, K. H., and Tietze, J. L.: Multipoint Approach to the Two-Point Boundary-Value Problem, Journal of Mathematical Analysis and Applications, vol. 44, pp. 625-642 (1973).

52. Morrison, D. D., Riley, J. D., and Zancanaro, J. F.: Multiple Solution Method for Two-Point Boundary-Value Problems, Communications of the Association for Computing Machinery, vol. 5, pp. 613-614 (1962).

53. Niem, A., Editor: "A Link between Science and Applications of Automatic Control," Preprints of the Seventh Triennial World Congress of the IFAC, Helsinki, Finland, June, 1978.

54. Nishikawa, Y., Hayashi, C., and Sannomiya, N.: Fuel and Energy Minimization in Three Dimensional Attitude Control of an Orbiting Satellite, Proc. IFAC Symposium on Peaceful Uses of Automation in Outer Space, Stravanger, Norway, pp. 287-298, 1965.

55. Nishikawa, Y., Ojika, T., Shimazutsu, H., and Okudaira, M.: An Interaction-Coordination Algorithm with Modified Performance Index for Nonlinear Optimal Control Problems, Journal of Optimization Theory and Applications, vol. 25, pp. 229-245 (1978).

56. Nishikawa, Y. and Okudaira, M.: A Method for Solving Multiple-Target Problems by Use of a Time-Decomposition Algorithm, Preprints of 16th Convention of Soc. Instrument and Control Engineers, pp. 483-484, 1977 (in Japanese).

57. Nishikawa, Y. and Okudaira, M.: A Time-Decomposition Algorithm for the Solution of Multiple-Target Problems, Memoirs of the Faculty of Engineering, Kyoto University, vol. XL, Part 2, pp. 62-77 (1978).

58. Nishikawa, Y., Okudaira, M., and Ojika, T.: A Numerical Method for Solving Optimal Control Problems with Control Constraint, Preprints of 18th Joint Convention of Automatic Control, pp. 27-28, 1975 (in Japanese).

59. Nishikawa, Y., Okudaira, M., and Ojika, T.: An Interaction-Coordination Algorithm with Time- and State-Decompositions for Nonlinear Optimal Control Problems, Memoirs of the Faculty of Engineering, Kyoto University, vol. XL, Part 3, pp. 169-181 (1978).

60. Nishikawa, Y. and Sannomiya, N.: Optimal Variable-Thrust Transfer between Neghboring Elliptic Orbits, Proc. 6th International Symposium on Space Technology and Science, Tokyo, pp. 311-322, 1965.

61. Ojika, T.: Decomposition and Coordination Algorithms for the Optimal Control of Nonlinear Dynamical Systems, Kyoto University, Department of Electrical Engineering, Doctoral Thesis, March, 1975.

62. Ojika, T., Nishikawa, Y., and Okudaira, M.: A Time-Decomposition Algorithm for a Stiff Linear Two-Point Boundary-Value Problem and Its Application to a Nonlinear Optimal Control Problem, Journal of Optimization Theory and Applications, vol. 27, pp. 241-258 (1979).

63. Ojika, T., Nishikawa, Y., Shimazutsu, H., and Okudaira, M.: An Interaction-Coordination Algorithm with Modified Performance Index for Optimal Control of Coupled Nonlinear Systems, Trans. Soc. Instrument

and Control Engineers, vol. 13, pp. 463-469 (1977)(in Japanese).

64. Osborne, M. R.: On Shooting Methods for Boundary-Value Problems, Journal of Mathematical Analysis and Applications, vol. 27, pp. 417-433 (1969).

65. Padmanabhan, L. and Bankoff, S. G.: A Direct Second-Variational Methods for Free-End Point, Free-Time Control Problems by Newton's Method in Function Space, International Journal of Control, vol. 9, pp. 167-177 (1969).

66. Pierre, D. A.: "Optimization Theory with Applications," John Wiley & Sons, New York, 1969.

67. Polak, E.: "Computational Methods in Optimization," Academic Press, New York, 1969.

68. Pontryagin, L. S., Boltyanskii, V., Gamkrelidze, R., and Mishchenko, E.: "The Mathematical Theory of Optimal Processes," Interscience Publishers, John Wiley & Sons, New York, 1962.

69. Roberts, S. M.: Simultaneous and Sequential Solution of a Stiff Differential Equation System, Journal of Optimization Theory and Applications, vol. 25, pp. 207-218 (1978).

70. Roberts, S. M. and Shipman, J. S.: Continuation in Shooting Methods for Two-Point Boundary-Value Problems, Journal of Mathematical Analysis and Applications, vol. 18, pp. 45-58 (1967).

71. Roberts, S. M. and Shipman, J. S.: Multipoint Solution of the Two-Point Boundary-Value Problems, Journal of Optimization Theory and Applications, vol. 7, pp. 301-318 (1971).

72. Roberts, S. M. and Shipman, J. S.: "Two-Point Boundary-Value Problems: Shooting Methods," American Elsevier, New York, 1972.

73. Schley, C. H., Jr. and Lee, I.: Optimal Control Computation by the Newton-Raphson Method and the Riccati Transformation, IEEE Trans. on AC, vol. AC-12, pp. 139-144 (1967).

74. Swiger, J. M.: Applications of the Theory of Minimum Normed Operator to Optimum Control System Problems, *in* Advances in Control Systems, vol. 3, Edited by C. T. Leondes, pp. 151-218, Academic Press, New York, 1966.

75. Tamura, K.: A Numerical Approach to the Optimum Control Problem with Discontinuity, Trans. Soc. Instrument and Control Engineers, vol. 8, pp. 600-605 (1972)(in Japanese).

76. Tomizuka, M. and Tsujioka, Y.: Formulation of a Multiple Target Problem and Its Analysis, Trans. Soc. Instrument and Control Engineers, vol. 6, pp. 228-233 (1970)(in Japanese).

77. Viterbi, A. J.: Acquisition and Tracking Behavior of Phase-Locked Loops, Proc. Symposium on Active Networks and Feedback Systems, Polytechnic Institute of Brooklyn, New York (1960).

78. Wismer, D. A., Editor: "Optimization Methods for Large-Scale Systems, with Applications," McGraw-Hill Book Co., New York, 1971.

79. Wonham, W. M.: On a Matrix Riccati Equation of Stochastic Control, SIAM Journal on Control, vol. 6, pp. 681-697 (1968).

80. Wonham, W. M.: On Pole Assignment in Multi-Input Controllable Linear Systems, IEEE Trans. on AC, vol. AC-12, pp. 660-665 (1967).

81. Ziv, A. and Amdursky, V.: On the Numerical Solution of Stiff Linear Systems of the Oscillatory Type, SIAM Journal on Applied Mathematics, vol. 33, pp. 593-606 (1977).