ON THE DETERMINATION OF TIME OPTIMAL CONTROLS
FOR LINEAR STATIONARY SYSTEMS

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
Norbert Bernard Hemesath
1964

THESIS

This is to certify that the

thesis entitled

ON THE DETERMINATION OF TIME
OPTIMAL CONTROLS FOR LINEAR
STATIONARY SYSTEMS

presented by

Norbert B. Hemesath

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in Electrical Engineering

Herman E. Koenig
Major professor

Date May 15, 1964

O-169

ABSTRACT

ON THE DETERMINATION OF TIME OPTIMAL CONTROLS FOR LINEAR STATIONARY SYSTEMS

by Norbert Bernard Hemesath

The last decade has witnessed an intense interest in time optimal control, i.e., the minimum time transfer of the n^{th} order system

$$\dot{X} = AX + BU(t)$$

from an arbitrary initial state to an arbitrary terminal state, subject only to the constraint that the components of the control vector, U(t), be bounded and measurable. The optimal control, when it exists, is known to have components which are piecewise continuous and assume only their extreme values. Furthermore, when A has real, distinct, non-positive eigenvalues, each control component has (n-1) or less discontinuous points. The optimal control is uniquely determined when the (n-1) discontinuous points or "switching times" and the initial sign of each of its components are known.

This thesis develops nonlinear equations which the optimal control satisfies and some techniques for solving these equations. The special case in which U(t) is a scalar is analyzed separately, and it is shown that the optimal control is the solution of an n^{th} order transcendental set in the (n-1) switching times and the minimum control time, t_n.

ON THE DETERMINATION OF TIME OPTIMAL CONTROLS FOR LINEAR STATIONARY SYSTEMS

By

Norbert Bernard Hemesath

A THESIS

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

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1964

ACKNOWLEDGEMENTS

I wish to thank Dr. H. E. Koenig for his help and guidance during the writing of this thesis and to acknowledge the support of the National Science Foundation during my years of graduate study. I am grateful to Suzanne for her patience and encouragement through some difficult periods.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF APPENDICES

I. INTRODUCTION

Any realistic design of a physical system imposes constraints on some if not all of the system parameters. Voltages, currents, forces, displacements, torques, velocities are all limited in magnitude. System designs ignoring such restrictions are often trivial and physically meaningless. A realistic design involves choosing components and parameters from within a given set of constraints such that "acceptable" performance is achieved as measured by some predetermined criterion. Optimization theory is concerned with the question: "Given a performance index, what is the 'best' design within the framework of the constraints imposed on the parameters?" If the physical system is described by the state model*

$$\dot{X} = F\,[X(t),U(t)] \tag{1-1}$$

then an important class of optimal problems involves choosing the vector U such that the scalar integral

---

*In this thesis upper case letters indicate vectors and matrices, lower case letters are scalars.

$$m = \int_{t_o}^{t_i} h[X(t),U(t)]dt \qquad (1\text{-}2)$$

is minimum, where $X(t_o)$ and $X(t_i)$ are the initial and terminal states, respectively, of (1-1), and U is an r^{th} order excitation vector constrained to lie within a compact region R of the r-dimensional space.

Typical problems in this class are: "What control function, subject to magnitude constraints, will transfer a system from one state to another in minimum time?" "What control will transfer a system from one state to another with minimum energy expenditure?" These and many similar questions have been investigated in recent years.

Time optimal control, the problem of taking a system from one state to another in minimum time, represents one of the most widely discussed optimization problems. It had its genesis in the efforts of researchers to determine when a relay controller should be switched to simultaneously reduce the error and its derivative to zero in minimum time following a step input.

The relay controller (also called bang-bang because of its on-off nature) is a simple, economical approach to closed loop control and is therefore attractive in many applications [1],[2]. Figure 1.1 below is a simple position servo using a DC motor whose armature voltage is applied discontinuously (off-full on) by the relay.

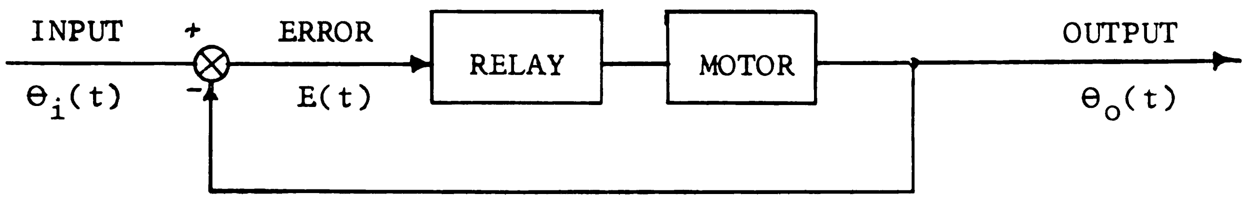


Figure 1.1 Relay controller.

The operation may be briefly described as follows: If the input and output positions differ, the relay applies armature voltage of the proper polarity to the motor to reduce the error. When the error changes sign the relay reverses the armature voltage. The system response finally depends upon the characteristics of the motor and its load. For no damping the system oscillates with constant amplitude; a damped system oscillates with decreasing amplitude and increasing frequency [1].

Reversing the relay before the error reaches zero reduces the "hunting" and shortens the settling time. System design procedures incorporating linear anticipatory switching, wherein the relay is reversed as some linear function of the error and its derivative before the error reaches zero, have been developed in the phase plane for second order systems such as that of Figure 1.1 [3],[4].

The concept of optimum performance is a natural outgrowth of the more sophisticated relay controller with anticipatory switching. Hopkin in 1950 defined optimum performance as "that behavior in which the system returns to rest with zero error in the shortest time following a step input" [5]. Hopkin and McDonald both demonstrated with heuristic proofs in the phase plane that a second order system with real characteristic roots and a bounded forcing function achieves zero error in minimum time when the forcing function assumes only its extreme values and reverses sign at a critical boundary which is a nonlinear function of error and error rate [5],[6].

In 1954 Bogner and Kazda, considering higher order systems with real roots, attempted to extend the phase plane concepts to a phase space [4]. In the case where the number of relay reversals is one less than the order of the system, their results indicate that a unique path exists from an arbitrary initial point in the phase space to the origin. Bushaw in 1952 discussed the second order system with complex roots as an abstract mathematical problem [7].

Bellman, Glicksberg and Gross in 1956 "imbedded" the optimum relay controller problem in a more general, precisely stated mathematical problem, and gave the first rigorous proof that for a rather general class of systems there exists an optimal controller and it is "bang-bang" in character [8]. LaSalle further generalized the theory to include time varying linear systems [9]. Concurrently several Russian authors, working independently, developed similar results. The approach of Bellman and LaSalle is topological; Desoer arrived at many of their conclusions using variational calculus [10]. Finally, the Russian mathematician, L. S. Pontryagin, devised a "maximum principle" which is applicable to a very broad class of systems and problems [11]. It too can be used to derive properties of the time optimal controller.

Thus, a distinguished body of theory relevant to the bang-bang control problem has been developed with the mathematical form of the control firmly established for systems of arbitrary order with both real and complex eigenvalues.

Although the form (bang-bang) of the control function is well known, the problem which has not been satisfactorily solved is this: "Given a system with an initial state, X_o, for which a time optimal control exists, how is that control found?" If this question has no reasonable answer, then optimal control remains a mathematician's game.

The object of this dissertation is to derive sets of conditions which the time optimal control for linear, constant coefficient systems <u>necessarily</u> satisfies, and to show that the equations representing these conditions can be solved by numerical techniques.

The body of this thesis contains six sections followed by three appendices. Section two is devoted exclusively to developing the mathematical theory of the time optimal problem considered in this thesis. Important properties of the controlled system and of the control itself are discussed.

Section three deals with the scalar control problem and develops the equations which the scalar control must satisfy. Section four introduces the vector control concept and develops the necessary optimization equations. A modified Lagrange multiplier method is used to obtain a solution to the resulting equations of optimization. Two sections on examples and conclusions complete the main body of the thesis.

Appendix A includes a new method for solving nonlinear equations by transforming the algebraic equations into differential equations whose solution at one endpoint represents the solution to the nonlinear algebraic equations. Appendix B contains standard material on numerical techniques and is included primarily for continuity. Appendix C gives some computer programs used to solve the optimization equations set down in sections three and four.

II. MATHEMATICAL THEORY

Statement of the Problem

The physical systems considered in this thesis can be described by a system of linear, constant coefficient, ordinary differential equations

$$\dot{x}_i = \sum_{j=1}^{n} a_{ij}x_j + \sum_{k=1}^{r} b_{ik}u_k(t) \qquad (2\text{-}1)$$

$$i = 1,\ 2,\ ---,\ n$$

where x_1, ---, x_n are the state variables which completely define the system, and $\dot{x}_i$ indicates differentiation of x_i with respect to the independent variable, t. Equation (2-1) may be written in the vector notation

$$\dot{X} = AX + BU(t) \qquad (2\text{-}2)$$

where X is an n-component column vector, A is an nxn constant matrix, B is an nxr constant matrix, U is an r-component column vector.

The trajectory X(t) of (2-2), as a function of t, is uniquely determined on an interval $0 \leq t \leq t_1$ when the control, U(t), and the initial condition $X(0) = X_o$, are specified. The ability to control the system lies in the freedom to choose U(t), the entries of which are assumed to satisfy the inequality

$$|u_k(t)| \leq 1 \qquad k = 1, ---, r. \tag{2-3}$$

Suppose also that the controls, $u_k(t)$, are piecewise continuous, i.e., continuous for all t, $0 \leq t \leq t_1$, except at a finite set of points t_i at which the controls may have finite discontinuities. Any control, U(t), whose components satisfy these two conditions is called an <u>admissible</u> control.

The time optimal control problem may now be stated as follows:

> <u>Given two points X_o and X_1 in the state space, among all admissible controls, U(t), which transfer the state point from X_o to X_1 (if such controls exist), find one which minimizes the time, $t_1 - t_o$</u> [11].

Here $X(t_o) = X_o$ and $X(t_1) = X_1$, and X(t) is the solution to (2-2) corresponding to control U(t).

The maximum principle as given by Pontryagin, can be used to establish some of the mathematical properties of the control which is the solution to the problem posed above.*

---

*The development of this chapter is not intended to present any new material, and therefore theorem proofs, readily available from such sources as Bellman, LaSalle, and Pontryagin, are omitted [8],[9],[10],[11].

Pontryagin's Maximum Principle

The maximum principle developed by Pontryagin and his associates states a necessary condition which an optimal control and the associated optimal trajectory of a system must satisfy. The process to be controlled is assumed to have a state model of the form

$$\dot{X} = F\,(X,U) \tag{2-4}$$

where: X is an n-component vector (state vector)

U is an r-component vector

F and its partials with respect to x_i, $i = 1, 2, ---, n$, are continuous on the direct product of the control space and the state space.

The performance is to be measured by the functional

$$J = \int_{t_o}^{t_1} f_o[X(t),U(t)]dt \tag{2-5}$$

where $f_o(X,U)$ together with its partial derivatives is defined and continuous on the direct product of the control and state space. Then the fundamental problem of optimal control is stated as follows:

Given any two points X_o and X_1 in the phase space, select from among all admissible controls, U(t), which transfer the phase point from X_o to X_1 (if such controls exist), the one which minimizes the functional, J,

where $X(t_o) = X_o$ and $X(t_1) = X_1$, and X(t) is the solution to (2-4) associated with control U(t).

Application of the maximum principle requires:

1. Augmentation of system (2-4) with the equation

$$\dot{x}_o = f_o(X,U) \tag{2-6}$$

2. Introduction of the adjoint set of equations

$$\dot{\bar{\psi}} = -J^T \bar{\psi} \tag{2-7}$$

where $\bar{\psi} = (\psi_o, \psi_1, ---, \psi_n)$

3. A scalar function H relating the augmented system and the adjoint system.

The augmented system is

$$\dot{\bar{X}} = \bar{F}(X,U) \tag{2-8}$$

where

$$\bar{F} = (f_o, f_1, ---, f_n) = (f_o, F)$$

$$\bar{X} = (x_o, x_1, ---, x_n) = (x_o, X).$$

The matrix J of (2-7) is the Jacobian of $\bar{F}(X,U)$, i.e.,

$$J = (A_{ij}) \quad i = 0, 1, ---, n$$
$$j = 0, 1, ---, n$$

$$A_{ij} = \partial f_i(X,U)/\partial x_j. \tag{2-9}$$

Observe that $\bar{\psi}$, $\bar{F}$, and $\bar{X}$ are all (n+1)-component vectors and that $\bar{F}(X,U)$ is not a function of x_o. The scalar function H relating (2-7) and (2-8) is defined

$$H = \overline{\psi}^T \, \overline{F}(X,U). \tag{2-10}$$

Systems (2-7) and (2-8) are obtainable from (2-10) as

$$\dot{x}_i = \partial H / \partial \psi_i \qquad i = 0, 1, ---, n$$

$$\dot{\psi}_i = -\partial H / \partial x_i \qquad i = 0, 1, ---, n.$$

Note that for constant values of X and $\overline{\psi}$ the function H depends only upon the vector parameter U. The maximum principle may now be stated as follows.

Maximum Principle

Let U(t), $t_o \leq t \leq t_1$, be an admissible control such that the corresponding trajectory, X(t), beginning at X_o at time t_o passes through X_1 at time t_1. If X(t) and U(t) are optimal it is necessary that:

1. There exist a non-zero continuous vector function $\overline{\psi}(t) = \psi_o(t), \psi_1(t), ---, \psi_n(t)$ corresponding to U(t) and X(t).
2. For every t, $t_o \leq t \leq t_1$, the function H of the variable U attains its maximum at U = U(t).

The maximum principle stated above is a necessary condition for optimality, but the fact that it has been satisfied does not assure the existence of an optimal control. Mathematical questions concerning the existence and

<u>uniqueness</u> of optimal controls are very important and difficult. The following sections deal with some of these questions in the particular case of time optimal, linear, stationary systems.

Properties of the Optimal Control

For the time optimal problem stated in the first section of this chapter the functional, J, is

$$J = \int_{t_o}^{t_1} f_o(X,U)dt = t_1 - t_o \tag{2-11}$$

which implies

$$f_o(X,U) = 1. \tag{2-12}$$

The augmented system (2-7) is

$$\begin{aligned} \dot{x}_o &= 1 \\ \dot{X} &= AX + BU \end{aligned} \tag{2-13}$$

and the adjoint system (2-7) defined in terms of $\psi = \psi_1, \psi_2, ---, \psi_n$ is

$$\begin{aligned} \dot{\psi}_o &= 0 \\ \dot{\psi} &= -A^T\psi. \end{aligned} \tag{2-14}$$

The function H can be written as

$$H(\overline{\psi},X,U) = \psi_o + \psi^T AX + \psi^T BU \qquad (2\text{-}15)$$

The maximum principle states that if U(t) is optimal then H, considered as a function of U alone, assumes its maximum at U = U(t). Since (2-15) is linear in U this implies that each component of U assumes its greatest magnitude and the sign of its coefficient. Since from (2-3) $|u_i| \leq 1$, the control U(t) may be written

$$U(t) = \operatorname{sgn} \psi^T B \qquad (2\text{-}16)$$

where for r-dimensional vectors A and B, A = sgn B means that a_i = sign of b_i, i = 1, ---, r. This result may be formally stated as

<u>Theorem 2-1</u>

If an optimal control function for the time optimal problem exists it is of the form

$$U(t) = \operatorname{sgn} \psi^T B$$

where $\psi(t)$ is a non-zero solution of the adjoint* system

$$\dot{\psi} = -A^T \psi.$$

---

*The system $\dot{Y} = -A^T Y$ is called the adjoint to $\dot{X} = AX$.

Therefore the form of the optimal control, if it exists, is established as piecewise constant or "bang-bang", and each component of the control switches sign at the zeros of the corresponding component of $\psi^T B$.

Perhaps a logical question at this point is: "Under what conditions does an optimal control exist, and is it unique?" A simple example may provide some insight into this problem. Consider the scalar equation

$$\dot{x} = ax + bu \tag{2-17}$$

By the maximum principle $u = \pm 1$ and never switches sign since the adjoint solution, $y(t) = e^{-at}y_o$, has no zeros. Thus the solution to (2-17) is

$$x(t) = e^{at}(x_o + \frac{bu}{a}) - \frac{bu}{a}$$

and if the desired terminal state is $x(t) = 0$, the optimal solution, if it exists, must satisfy

$$\left(\frac{1}{\frac{ax_o}{bu} + 1}\right) = e^{at} \qquad t \geq 0 \tag{2-18}$$

Whether or not (2-18) has a solution depends upon the values of the parameters, a and b, and the sign of u. There are three cases of interest.

Case 1: $a < 0$, $b \neq 0$

An optimal solution exists for arbitrary x_o since the left hand side is always less than one if $u = \text{sgn}\left(\frac{a\ x_o}{b}\right)$, while the right side approaches zero as t grows large.

Case 2: $a > 0$, $b \neq 0$

No solution exists for values of x_o such that $\left|\frac{a\ x_o}{b}\right| > 2$. Under this condition the left hand side is always less than one while the right hand side exceeds unity.

Case 3: a arbitrary, $b = 0$

No solution exists since $x(t) = e^{at}x_o$ never reaches the origin.

The behavior in case three above is related to the concept of <u>controllability</u> as introduced by R. E. Kalman [12]. He defines a system to be <u>completely controllable</u> if for arbitrary states X_o and X_1, and times t_o and t_1, there exists a control which transfers the system from state X_o at time t_o to state X_1 at time t_1. Kalman has also stated the following [13]

<u>Theorem 2-2</u>

The system described by (2-2) is completely controllable if and only if the matrix $[A^{n-1}B, A^{n-2}B, ---, AB, B]$ has maximum rank.

The system described in case three does not satisfy this theorem, and is, therefore, not controllable. It is apparent that a system which is to be optimally transferred from an arbitrary initial state to the origin must be <u>necessarily</u> completely controllable.

However complete controllability is <u>not sufficient</u> for optimal control of a system with arbitrary initial state. Consider case two; the system is completely controllable by theorem 2-2, yet for certain initial states there is no optimal control.

On the other hand the system of case one is completely controllable and always has an optimal control. This leads to the final factor affecting the existence of optimal controls, <u>stability</u>. System one has a stable characteristic root while system two has an unstable characteristic root, where a stable root is defined as an eigenvalue of the matrix A in (2-2) with non-positive real part. The intuitive evidence might lead one to expect that which the following theorem states [11]

<u>Theorem 2-3</u>

If the matrix A has stable eigenvalues and if the system (2-2) is completely controllable, then there exists an optimal control which transfers an arbitrary initial phase point, X_0, to the origin.

The previous theorem establishes the existence of an optimal control under certain conditions. Another very important property of the control is its uniqueness which is assured by

<u>Theorem 2-4</u>

Let $U_1(t)$ and $U_2(t)$ be two optimal controls (defined on the intervals $t_o \leq t \leq t_1$ and $t_o \leq t \leq t_2$ respectively) which transfer the phase point X_o to X_1. Then these controls coincide, i.e., $t_1 = t_2$ and $U_1(t) = U_2(t)$.

In systems which have real eigenvalues the optimal control has an especially significant property wherein the number of switchings of each control component is related to the order of the system.

<u>Theorem 2-5</u>

If the matrix A of (2-2) has real, non-positive roots, then each component, u_i, i = 1, ---, r, of the optimal control will switch not more than (n-1) times where n is the order of the system.

The property stated in theorem 2-5 above is exploited in the following chapters of this thesis to develop sets of equations which the optimal control must satisfy. Henceforth, unless otherwise stated, the system under consideration is that characterized by (2-2) with the additional restrictions that the eigenvalues of A are real, non-positive and simple (non-repeated), and that the system is completely controllable, i.e., all subsequent development is concerned with linear, stationary systems for which a unique, time optimal control always exists.

III. SCALAR CONTROL PROBLEM

The time optimal control of physical systems governed by a single control variable is perhaps the most significant of the class of time optimal problems in applications. Indeed, all single input, single output, linear control systems currently designed with s-domain techniques fall into this class. For these systems usually (but not necessarily) the error and its first (n-1) derivatives are reduced to zero. Because of its importance and its mathematical tractability, the scalar control problem is considered first as a special case.

The Switching Equations

Many of the early efforts to determine time optimal controls were based on the phase space, a generalization of the phase plane which is so useful for second-order systems [4],[16]. The approach is that of establishing switching surfaces in the phase space, i.e., surfaces at which the control function changes sign. However, as the order of the system increases the problem of eliminating the time variable from the system equations becomes very difficult. Recently several authors have suggested techniques for determining the control as a function of switching times rather than as a function of the system state [17],[18]. This approach will be used here.

Consider the completely controllable system

$$\dot{X} = AX + Bu(t) \qquad (3\text{-}1)$$

where A is a square matrix with simple, real, non-positive eigenvalues, B is a column vector and u(t) is a scalar control function. The theorems of the previous section guarantee the existence of a unique control function which will transfer the solution of (3-1) from an arbitrary initial state to the origin in minimum time.

The equations of optimal control are simplified if the system (3-1) is reduced to principal coordinates, i.e., A is diagonalized. Defining the nonsingular linear transformation

$$X = PY \qquad (3\text{-}2)$$

equation (3-1) becomes

$$\dot{Y} = (P^{-1}AP)Y + P^{-1}Bu(t) \qquad (3\text{-}3)$$

where [14],[15]

$$P^{-1}AP = D = \mathrm{diag}(\lambda_1, \lambda_2, ---, \lambda_n) \qquad (3\text{-}4)$$

and the λ_i are the eigenvalues of A. If we let $z_i = y_i/(P^{-1}B)_i$ where $(P^{-1}B)_i$ is the i^{th} component of the vector $P^{-1}B$ then (3-3) reduces to

$$\dot{Z} = DZ + Iu(t) \tag{3-5}$$

where I is a column vector of ones.

The solution to (3-5) is [19]

$$Z(t) = e^{Dt}Z_o + \int_o^t e^{D(t-r)} Iu(r)dr \tag{3-9}$$

If t_n represents the time required for an optimal control to transfer Z_o to the origin, then the optimal solution is

$$Z(t_n) = 0 = e^{Dt_n}Z_o + \int_o^{t_n} e^{D(t_n-r)} Iu(r)dr. \tag{3-10}$$

Multiplying both sides of (3-10) by $(e^{Dt_n})^{-1} = e^{-Dt_n}$ gives

$$-Z_o = \int_o^{t_n} e^{-Dr} Iu(r)dr. \tag{3-11}$$

Since u(t) is piecewise constant and reverses sign at most (n-1) times on the interval $[0,t_n]$, (3-11) may be written as the sum of n integrals of alternating sign

$$-Z_o = u\left\{\int_o^{t_1} e^{-Dr} Idr - \int_{t_1}^{t_2} e^{-Dr} Idr + --- + (-1)^{(n-1)} \int_{t_{n-1}}^{t_n} e^{-Dr} Idr\right\} \tag{3-12}$$

where $u = \pm 1$ and t_1, ---, t_{n-1} are the (n-1) switching times which must satisfy the ordering constraints

$$0 \leq t_1 \leq t_2 \leq --- \leq t_{n-1} \leq t_n. \qquad (3\text{-}13)$$

If none of the $\lambda_i = 0$ integration of (3-12) gives n equations

$$-\lambda_i z_{io} = u\left[1-2e^{-\lambda_i t_1}+2e^{-\lambda_i t_2} - --- -(-1)^n 2e^{-\lambda_i t_{n-1}}+(-1)^n e^{-\lambda_i t_n}\right]$$

$$i = 1, ---, n \qquad (3\text{-}14)$$

where z_{io} is the i^{th} element of the vector Z_o. The modification required where some $\lambda_i = 0$ is obvious.

To find the unique optimal control it is necessary to find the ordered set

$$0 \leq t_1 \leq t_2 \leq --- \leq t_{n-1} \leq t_n$$

with minimum t_n which satisfies the n transcendental equations in (3-14) and to find the correct sign for u.

Bounds on the Control Time

Systems of transcendental equations, in general, cannot be solved analytically and the solutions are not unique. However, any solution of (3-14) which simultaneously satisfies (3-13) is unique [4]. Since numerical methods are required to solve (3-14), a "good" first approximation to the desired solution is necessary if an iterative procedure is to prove successful. Initial estimates must be made for t_1, t_2, ---, t_n as well as the sign of u. The switching times, of course, must be positive and must satisfy the ordering constraints (3-13). However, the ordering constraints contain no information about the magnitudes of these switching times. Intuitively one might expect that the control time required to bring a system to the origin is a function of the system initial state and the eigenvalues. The following theorem shows that this is, indeed, true, and provides a useful basis for choosing the initial vector with which to begin an iterative solution technique.

<u>Theorem 3-1</u>

If the eigenvalues are all distinct from zero, the minimum time T_o required to transfer the normalized system

$$\dot{z}_i = \lambda_i z_i + u(t) \qquad i = 1, 2, ---, n \tag{3-15}$$

from an arbitrary initial state, Z_o, to the origin satisfies the inequality

$$T_o \geq \max t_i$$

where:

$$t_i = \frac{1}{|\lambda_i|} \log(|\lambda_i z_{io}| + 1) \qquad i = 1, ---, n \tag{3-16}$$

Proof:

Since the minimum time solution requires the simultaneous transfer to the origin of all states, the minimum time solution, T_o, must equal or exceed the maximum of the set

t_1, t_2, ---, t_n, where t_i is the minimum time required to transfer the initial state, z_{io}, to zero. From (3-14) the control which transfers z_{io} to the origin in minimum time is the solution to

$$-\lambda_i z_{io} = u(1-e^{-\lambda_i t_i}), \quad t_i > 0. \qquad (3\text{-}17)$$

The solution to (3-17) is

$$t_i = -\frac{1}{\lambda_i} \log(1 + \frac{\lambda_i z_{io}}{u}).$$

and since $t_i > 0$ it follows that $(1 + \frac{\lambda_i z_{io}}{u}) > 1$

and

$$u = \text{sgn}\ (\lambda_i z_{io}). \qquad (3\text{-}18)$$

It follows, therefore, that

$$t_i = \frac{1}{|\lambda_i|} \log\ (1 + |\lambda_i z_{io}|) \qquad (3\text{-}19)$$

and the theorem is proved.

In case one of the $\lambda_i = 0$, say $\lambda_1 = 0$, the solution corresponding to (3-17) is

$$0 = ut + z_{10} \qquad (3\text{-}20)$$

and the corresponding t_i is

$$t_1 = |z_{10}| \ . \qquad (3\text{-}21)$$

Hence for a system with $\lambda_1 = 0$

$$T_o \geq \max \left[\left|z_{10}\right| , t_i \right] \qquad (3\text{-}22)$$

where $t_i = \dfrac{1}{|\lambda_i|} \log \left(1+\left|\lambda_i z_{io}\right|\right).$

$$i = 2,\ 3,\ ---,\ n$$

IV. THE VECTOR CONTROL PROBLEM

Consider the linear system

$$\dot{X} = AX + BU(t) \qquad (4\text{-}1)$$

where:

X is the n component state vector

A is an nxn matrix with real, non-positive, simple eigenvalues

B is an nxr matrix

U is the r-component control vector

$|u_i| \leq 1 \quad i = 1, 2, ---, r \qquad 2 \leq r \leq n.$

From the results of the previous section and other investigators the control vector, U_o, which transfers an arbitrary initial state, X_o, to the origin in minimum time is known to be unique. Further, each component of U_o assumes only its extreme values, is piecewise constant, and switches not more than (n-1) times. The set of equations which establish the switching times on the components of U_o is obtained by extending the development of the previous section.

The Switching Equations

Let the system in (4-1) be transformed to principal coordinates by the linear transformation X = PY so that

$$\dot{Y} = DY + CU(t) \qquad (4\text{-}2)$$

where

$$D = \text{diag}\ (\lambda_1 --- \lambda_n) = P^{-1}AP$$
$$C = P^{-1}B$$

Since the linear, constant coefficient differential equation (4-2) has the vector solution

$$Y(t) = e^{Dt}Y_o + \int_0^t e^{D(t-\tau)}CU(\tau)d\tau \qquad (4\text{-}3)$$

the optimal control vector, U(t), must satisfy the relation

$$-e^{Dt_n}Y_o = \int_o^{t_n} e^{D(t_n-\tau)}CU(\tau)d\tau \qquad (4\text{-}4)$$

or

$$-Y_o = \int_o^{t_n} e^{-D\tau}CU(\tau)d\tau \qquad (4\text{-}5)$$

$$-Y_o = \int_o^{t_n} e^{-D\tau}C_1u_1(\tau)d\tau + --- + \int_o^{t_n} e^{-D\tau}C_ru_r(\tau)d\tau \qquad (4\text{-}6)$$

where the C_i, i = 1, 2, ---, r, represent the columns of C.

The fact that $|u_i| = 1$ for all t on the interval $(0, t_n)$ with (n-1) switchings or less, makes the integration of (4-6) elementary when it is written as n scalar equations

$$-y_{io} = \int_{o}^{t_n} e^{-\lambda_i \tau} c_{i1} u_1 d\tau + \text{---} + \int_{o}^{t_n} e^{-\lambda_i \tau} c_{ir} u_r d\tau \;. \qquad (4\text{-}7)$$

$$i = 1,\ 2,\ \text{---},\ n$$

Now each of the r integrals in (4-7) can be written as n integrals so that (4-7) becomes

$$-y_{io} = c_{i1}u_1 \left\{ \int_{o}^{t_{11}} e^{-\lambda_i \tau} d\tau - \int_{t_{11}}^{t_{12}} e^{-\lambda_i \tau} d\tau + \text{---} \right.$$

$$\left. + (-1)^{(n-1)} \int_{t_{1,n-1}}^{t_n} e^{-\lambda_i \tau} d\tau \right\} + \text{---} + c_{ir}u_r \left\{ \int_{o}^{t_{r1}} e^{-\lambda_i \tau} d\tau \right.$$

$$\left. - \int_{t_{r1}}^{t_{r2}} e^{-\lambda_i \tau} d\tau + \text{---} + (-1)^{(n-1)} \int_{t_{r,n-1}}^{t_n} e^{-\lambda_i \tau} d\tau \right\}$$

$$i = 1,\ 2,\ \text{---},\ n \qquad (4\text{-}8)$$

where the t_{ij}, $i = 1, \text{---}, r$, $j = 1, \text{---}, (n-1)$ are the (n-1) switching times associated with u_i and satisfy the inequalities

$$0 \leq t_{i1} \leq t_{i2} \leq \text{---} \leq t_{i,n-1} \leq t_n \quad i = 1, \text{---}, r. \qquad (4\text{-}9)$$

Assuming $\lambda_i \neq 0$ (4-8) reduces to

$$
\begin{aligned}
-y_{io} = {} & \frac{c_{i1}u_1}{\lambda_i}\left[1 - 2e^{-\lambda_i t_{11}} + 2e^{-\lambda_i t_{12}} - \cdots + (-1)^n e^{-\lambda_i t_n}\right] \\
& + \frac{c_{i2}u_2}{\lambda_i}\left[1 - 2e^{-\lambda_i t_{21}} + \cdots + (-1)^n e^{-\lambda_i t_n}\right] \\
& \vdots \\
& + \frac{c_{ir}u_r}{\lambda_i}\left[1 - 2e^{-\lambda_i t_{ri}} + 2e^{-\lambda_i t_{r2}} - \cdots + (-1)^n e^{-\lambda_i t_n}\right]
\end{aligned}
$$

$$i = 1, 2, \cdots, n \qquad (4\text{-}10)$$

The result for a particular $\lambda_i = 0$ calls for a trivial modification.

The optimal control vector U_o is completely specified by the solution to (4-10) for minimum t_n subject to the constraints in (4-9). Since u_i may be ± 1 (4-10) is actually 2^r distinct sets of n equations, and each of these sets involves $r(n-1) + 1$ variables. Since $2 \leq r \leq n$ the number of unknowns exceeds the number of equations. If such a system has one solution it has an infinity of solutions each of which is obtained by arbitrarily specifying $r(n-1) + 1 - n$ of the variables and solving the resulting normal* set. However the problem of finding the unique time optimal control remains.

---

*A normal set is one with the same number of equations as unknowns.

Constrained Minimum Problem

Finding the time optimal control from set (4-10) may be viewed as a constrained minimization problem [17]. Since t_n appears in each equation of (4-10) only once, in a term of the form, $K_i e^{-\lambda_i t_n}$, it is possible to solve explicitly for t_n in terms of the remaining r(n-1) switching times, thus

$$t_n = f(t_{11}, t_{12}, ---, t_{r,n-1}). \tag{4-11}$$

Substitution of (4-11) into the remaining (n-1) equations of (4-10) gives a set of n-1 equations which is independent of t_n

$$g_i(t_{11}, t_{12}, ---, t_{r,n-1}) = 0 \quad i = 1,2,---,(n-1). \tag{4-12}$$

The optimal time solution is now obtained by minimizing $f(t_{11}, ---, t_{r,n-1})$ subject to the constraint equations of the form given in (4-12), and the ordering constraints, (4-9).

Using Lagrange multipliers to find the minimum, form the scalar function [20],[21],[25]

$$\begin{aligned} H(t_{11},---,t_{r,n-1},\pi_1,---,\pi_{n-1}) &= f(t_{11},---,t_{r,n-1}) \\ &+ \pi_1 g_1(t_{11},---,t_{r,n-1}) + --- + \pi_{n-1} g_{n-1}(t_{11},---,t_{r,n-1}) \end{aligned} \tag{4-13}$$

where the π_i are constant multipliers. Consider now the problem

is that

other

ments*

original

[illegible]

problem of locating the extrema of H. A necessary condition is that all its partial derivatives vanish

$$\frac{\partial H}{\partial t_{ij}} = 0 \quad i = 1, ---, r \quad j = 1, ---, n-1$$

$$\frac{\partial H}{\partial \pi_k} = 0 \quad k = 1, ---, n-1 \tag{4-14}$$

Set (4-14) contains (r+1)(n-1) equations in the same number of unknowns, and its solutions are the stationary points* of H. Among its solutions are the minima of the original function, f, subject to the constraints, $g_i = 0$.

Two difficulties arise:

1. The solutions obtained are minima satisfying the constraint set, $g_i = 0$, yet they may not satisfy the inequality constraints, (4-9), which order the switching times.
2. The desired minimum time solution may be on the boundary of the closed constraint set defined by inequalities (4-9), in which case it is not necessarily a stationary point of H, i.e.,

$$\left.\frac{\partial H}{\partial t_{ij}}\right|_{T_o} \neq 0 \quad \text{for some } t_{ij} \tag{4-15}$$

where T_o represents the time optimal solution vector.

---

*A stationary point of a function is one at which all its partial vanish.

As a simple example illustrating the second point, consider the extreme values of $f(x) = (x-1)^3$, $0 \leq x \leq 2$. Figure 4.1 below shows the minimum at x = 0 and the maximum at x = 2. The only stationary point is x = 1.

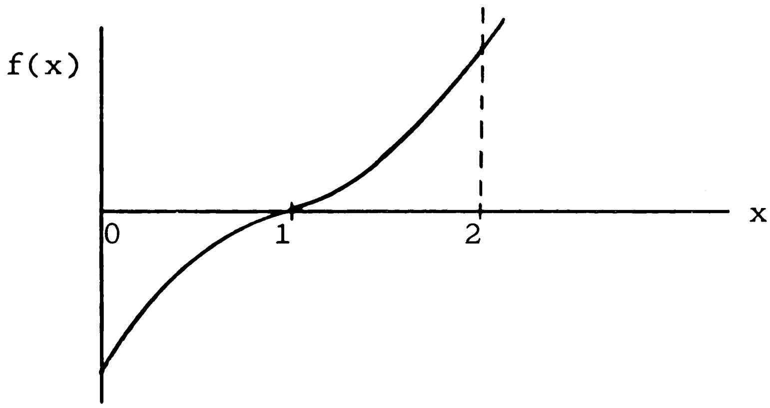


Figure 4.1 $f(x) = (x-1)^3$

The two difficulties discussed above impair the usefulness of the classical Lagrange technique; in the first case undesired solutions are obtained while in the second the minimum time solution cannot be found because it is not a stationary point. A method which permits use of the constraints, $g_i = 0$, as well as the ordering constraints in (4-9) is given next.

Modified Lagrange Technique

Several authors have extended the Lagrange multiplier technique so that inequality as well as equality constraints may be handled [22],[23],[24]. This is done by observing that an inequality can be transformed into an equality by introducing a variable parameter s such that

$$f(x) \geq 0 \text{ implies } f(x) = s^2. \qquad (4\text{-}16)$$

Thus the inequality constraints in (4-9) can be re-written as

$$\begin{aligned} t_{i1} &\geq 0 \\ t_{i2} &\geq t_{i1} \\ &\vdots \\ t_n &\geq t_{i,n-1} \end{aligned} \qquad i = 1,---,r \qquad (4\text{-}17)$$

and replaced by the following set of equalities

$$\begin{aligned} t_{i1} &= s_{i1}^2 \\ t_{i2} &= t_{i1}+s_{i2}^2 \\ &\vdots \\ t_n &= t_{i,n-1}+s_{in}^2 \end{aligned} \qquad i = 1,---,r \qquad (4\text{-}18)$$

Sucessive substitution of each member of (4-18) into the following one yields each t_{ij} defined in terms of the s_{ij} alone

$$\begin{aligned} t_{i1} &= s_{i1}^2 \\ t_{i2} &= s_{i1}^2+s_{i2}^2 \\ &\ \vdots \\ t_n &= s_{i1}^2+s_{i2}^2+\cdots+s_{in}^2 \end{aligned} \qquad i = 1,2,\cdots,r \qquad (4\text{-}19)$$

Now finding the minimum time solution requires minimizing

$$t_n = f(t_{11},\cdots,t_{r,n-1})$$

subject to the constraints

$$g_i(t_{11},\cdots,t_{r,n-1}) = 0 \quad i = 1\cdots,n-1$$

$$\begin{aligned} t_{i1} &= s_{i1}^2 \\ t_{i2} &= s_{i1}^2+s_{i2}^2 \\ &\ \vdots \\ t_n &= s_{i1}^2+s_{i2}^2+\cdots+s_{in}^2 \ . \end{aligned} \qquad i = 1,2,\cdots,r \qquad (4\text{-}20)$$

Observe that by substituting into f and g_i the relations from (4-20) defining the t_{ij} the remaining constraints in (4-20) are r in number. Thus, the problem becomes one of minimizing

$$t_n = f(s_{ij}) \qquad \begin{aligned} i &= 1,\cdots,r \\ j &= 1,\cdots,n-1 \end{aligned} \qquad (4\text{-}21)$$

subject to the constraints

$$g_j(s_{ij}) = 0 \quad i = 1,---,r \quad j = 1,---,n-1 \qquad (4\text{-}22)$$

$$h_k = -f(s_{ij})+s_{k1}^2+s_{k2}^2+---+s_{kn}^2 = 0$$

$$k = 1,---,r \qquad (4\text{-}23)$$
$$i = 1,---,r$$
$$j = 1,---,n-1$$

Now $f(s_{ij})$ may be minimized by the usual Lagrange technique of forming

$$H(s_{ij},\pi_k,u_i) = f+\pi_k g_k+u_i h_i$$

$$i = 1,---,r \qquad (4\text{-}24)$$
$$j = 1,---,n$$
$$k = 1,---,n-1$$

and setting partials with respect to all variables equal to zero

$$\frac{\partial H}{\partial s_{ij}} = \frac{\partial f}{\partial s_{ij}} + \sum_{k=1}^{n-1} \pi_k \frac{\partial g_k}{\partial s_{ij}} + \sum_{m=1}^{r} u_m \frac{\partial h_m}{\partial s_{ij}} = 0$$

$$\frac{\partial H}{\partial \pi_k} = g_k = 0 \qquad \begin{array}{l} i = 1,---,r \\ j = 1,---,n \\ k = 1,---,n-1 \end{array} \qquad (4\text{-}25)$$

$$\frac{\partial H}{\partial u_i} = h_i = 0$$

This is a set of nr + n-1 + r = (r+1) (n+1)-2 equations in the same number of variables. However, in practice, r of these equations may be eliminated immediately. None of the functions, f and g_i, contain variables s_{1n}, s_{2n} ---, s_{rn}, since t_n does not appear in them and s_{1n}, ---, s_{rn} do not appear in the equations defining the t_{ij} (see (4-20)). Therefore

$$\frac{\partial H}{\partial s_{in}} = \frac{\partial f}{\partial s_{in}} + \sum_{k=1}^{n-1} \pi_k \frac{\partial g_k}{\partial s_{in}} + \sum_{m=1}^{r} u_m \frac{\partial h_m}{\partial s_{in}} = 0$$

$$i = 1,---,r$$

$$= 0 + 0 + 2u_i s_{in} = 0$$

$$\frac{\partial H}{\partial s_{in}} = 2u_i s_{in} = 0 \qquad i = 1,2,---,r \qquad (4\text{-}26)$$

and it is necessary that either $u_i = 0$ or $s_{in} = 0$. In the usual case where each control component switches (n-1) distinct times, all $s_{ij} \neq 0$. The conclusion is that $u_1 = u_2 = \cdots = u_r = 0$, thus reducing the number of equations and the number of unknowns in set (4-25) by r, and leaving n(r+1)-1 equations in n(r+1)-1 variables. The cases in which some or all of the s_{in} are zero can also be handled by the "either-or" rule. Therefore, the system to be solved is always the following normal one of dimension n(r+1)-1

$$\frac{\partial H}{\partial s_{ij}} = \frac{\partial f}{\partial s_{ij}} + \sum_{k=1}^{n-1} \pi_k \frac{\partial g_k}{\partial s_{ij}} + \sum_{m=1}^{r} u_m \frac{\partial h_m}{\partial s_{ij}} = 0$$

$$\frac{\partial H}{\partial \pi_k} = g_k = 0 \qquad \begin{array}{l} i = 1,\cdots,r \\ j = 1,\cdots,n-1 \\ k = 1,\cdots,n-1 \end{array} \qquad (4\text{-}27)$$

$$\frac{\partial H}{\partial u_i} = h_i = 0$$

Among the solutions to this set will be the unique time optimal solution. Furthermore, the two difficulties encountered in the classical Lagrange development have been eliminated:

1. Every solution (only real solutions are considered) is a realizable control since, by virtue of (4-20), all the switching times are positive and ordered.
2. The minimum will always occur at a stationary

point of H since the variables s_{ij}, π_k, u_m are all unrestricted in range.

Bounds of the Control Time

The system of equations in (4-27) is highly nonlinear in the variables, s_{ij}, and therefore not amenable to analytic solution. More often than not the convergence of numerical methods of solution is dependent critically upon a "good" first approximation to the solution. The motivation for the following two theorems is the establishment of upper and lower bounds on the optimal control time, t_n, as an aid in choosing an approximate solution vector.

Theorem 4-1

The time, t_n, representing the time optimal solution to (4-1) with the origin as terminal state satisfies

$$t_n \geq \max P_i$$

where

$$P_i = \frac{1}{|\lambda_i|} \log \left[\frac{(-\operatorname{sgn} x_{io})(|b_{i1}| + \cdots + |b_{ir}|) + \lambda_i x_{io}}{(-\operatorname{sgn} x_{io})(|b_{i1}| + \cdots + |b_{in}|)} \right] \qquad (4\text{-}28)$$

$$i = 1, \cdots, n$$

Proof: Completely analogous to the proof of theorem 3-1.

If each entry u_2 --- u_r of the control vector, U, is chosen to be $\pm u_1$, i.e.,

$$u_i = \pm u_1 \quad i = 2,---,r$$

then the vector control problem becomes a scalar control problem, for a typical equation from the set (4-1) is

$$\dot{x}_i = \lambda_i x_i + b_{i1}u_1 \pm b_{i2}u_1 \pm b_{i3}u_1 \pm --- \pm b_{ir}u_1$$

or

$$\dot{x}_i = \lambda_i x_i + (b_{i1} \pm b_{i2} \pm --- \pm b_{ir})u_1 \qquad (4\text{-}29)$$

$$i = 1,---,n.$$

Evidently the set (4-29) can be written in $2^{(r-1)}$ ways since each of the last (r-1) columns of B may assume either the plus or minus sign. Thus (4-29) represents $2^{(r-1)}$ different scalar control problems. Let

$$T_k \quad k = 1, 2, ---, 2^{(r-1)}$$

be the optimal control time associated with the k^{th} control problem, and consider the minimum of the set T_k, say T_o. T_o is the time required to reduce an arbitrary initial state, X_o, to the origin under the influence of a particular control vector, U_o, in which each entry is some specific choice of $\pm u_1$. Either U_o is the optimal control vector or it isn't; in either case

$$t_n \leq T_o$$

and we have:

Theorem 4-2

1. Let t_n be the optimal solution to the vector control problem

$$\dot{X} = AX + BU$$

2. Define $2^{(r-1)}$ scalar control problems by

$$u_i = \pm u_1 \quad i = 2,\text{---},r$$

3. Define the set T_k, $k = 1,\text{---},2^{(r-1)}$, where T_k is the optimal solution, if it exists, to the k^{th} scalar control problem.

4. Define $T_o = \min \; T_k$

then

$$t_n \leq T_o \; .$$

Theorems 4-1 and 4-2, taken together, establish both a lower and an upper bound on the optimal control time, t_n. These bounds are very useful for choosing an approximate solution vector with which to initiate an iterative numerical scheme. In most cases the bounds are reasonably "sharp"; the lower bound always exists, while in certain cases the upper one may not exist. Consider the example

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} \tag{4-30}$$

This is a well-defined second-order vector control problem. Since r = 2, there are two derived scalar problems.

First Scalar Problem, $u_2 = u_1$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ u_1 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 2 \\ 0 \end{bmatrix} u_1$$

Second Scalar Problem, $u_2 = -u_1$

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} u_1 \\ -u_1 \end{bmatrix}$$

$$= \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} u_1$$

Neither of the two problems has a time optimal control since neither system is controllable. Therefore, theorem 4-2 cannot be used to establish an upper bound on the optimal solution to the vector problem. However, this difficulty arises if and only if each of the $2^{(r-1)}$ derived scalar systems is not controllable.

Most of the effort in this section has been directed toward deriving a set of equations, (4-27), which the time optimal solution to the vector control problem necessarily satisfies. A generalization of Lagrange's multiplier technique was used to handle the inequality constraints on the switching times. Two theorems bounding the control time, t_n, above and below were stated.

Finally, the results of this section hold also for the scalar problem (r=1), but, in practice, the n^{th} order set, (3-14), of the previous section is easier to use since its dimension is (n-1) less than that of (4-27).

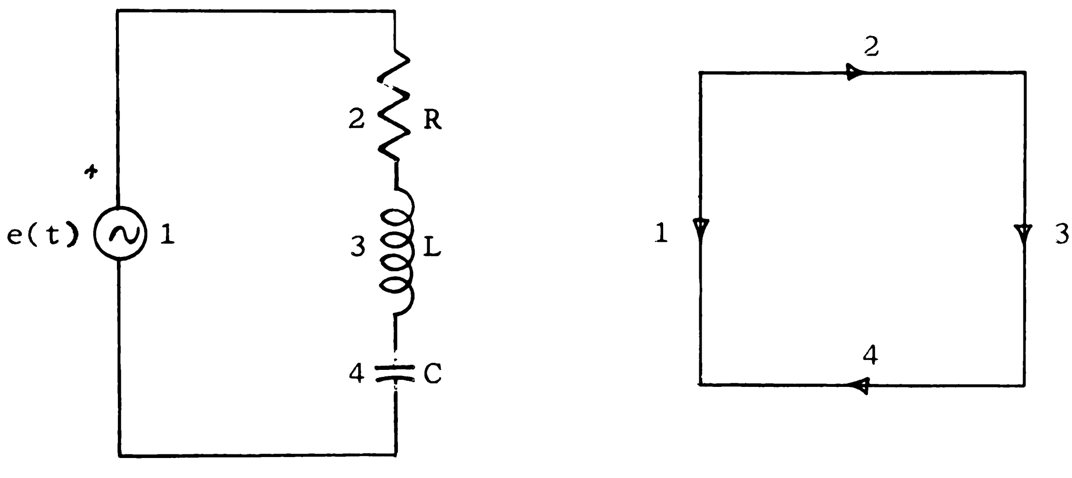
V. EXAMPLES

The previous sections have been devoted to developing the theory of time optimal control. Equations for which the optimal control is a solution have been derived. In this section application of the theory developed above to several simple, physical systems is considered.

Scalar Control Problems

<u>Example 1</u>

Consider the simple R-L-C system shown in Figure 5.1(a).



(a) (b)

Figure 5.1 Simple R-L-C System.

A state model of the system based on the linear graph of Figure 5.1(b) is [26]

$$\frac{d}{dt}\begin{bmatrix} v_4(t) \\ \\ i_3(t) \end{bmatrix} = \begin{bmatrix} 0 & \frac{1}{C} \\ \\ -\frac{1}{L} & -\frac{R}{L} \end{bmatrix}\begin{bmatrix} v_4(t) \\ \\ i_3(t) \end{bmatrix} + \begin{bmatrix} 0 \\ \\ \frac{e(t)}{L} \end{bmatrix}. \tag{5-1}$$

The second-order system in (5-1) has eigenvalues

$$\lambda_{1,2} = -\frac{R}{2L} \pm \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} \tag{5-2}$$

If R, L, and C are positive, and if

$$\frac{R^2}{4L^2} > \frac{1}{LC}$$

the theorems of Chapter II establish the existence of a unique, bang-bang control which switches once, and reduces both the initial voltage, $v_4(0)$, and the initial current, $i_3(0)$, to zero in minimum time. Since the equations which must be solved to yield the optimal control are derived from a state model written in principal coordinates, the coeffi- cient matrix in (5-1) must be diagonalized. The nonsingular linear transformation which does this is

$$\begin{bmatrix} v_4 \\ \\ i_3 \end{bmatrix} = \begin{bmatrix} \frac{1}{C} & \frac{1}{C} \\ \\ \left(-\frac{R}{2L} + \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}\right) & \left(-\frac{R}{2L} - \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}\right) \end{bmatrix} \begin{bmatrix} y_1 \\ \\ y_2 \end{bmatrix}. \tag{5-3}$$

and the transformed system is

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -\frac{R}{2L} + \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} & 0 \\ 0 & -\frac{R}{2L} - \sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}} \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

$$+ \begin{bmatrix} \dfrac{e(t)}{2L\sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}} \\ \dfrac{-e(t)}{2L\sqrt{\frac{R^2}{4L^2} - \frac{1}{LC}}} \end{bmatrix} \tag{5-4}$$

For the particular case where $R=3$, $L=1$, $C=\frac{1}{2}$, (5-4) becomes

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} e(t) \tag{5-5}$$

One more linear transformation

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \tag{5-6}$$

reduces (5-5) to the normalized form of (3-8)

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \end{bmatrix} e(t) \qquad (5\text{-}7)$$

The nonlinear switching equations corresponding to (3-14) from which the optimal solution is obtained are

$$\begin{aligned} -\lambda_1 z_{10} &= e(t)(1-2e^{t_1}+e^{t_2}) \\ -\lambda_2 z_{20} &= e(t)(1-2e^{2t_1}+e^{2t_2}) \end{aligned} \qquad (5\text{-}8)$$

where z_{10} and z_{20} are initial conditions related to $v_4(0)$ and $i_3(0)$, respectively, by the product of the linear transformations (5-6) and (5-3). Since e(t) assumes only the values +1 and -1, (5-8) may be solved for both cases, and the results in Chapter III indicate the optimal solution is the one which satisfies $0 \leq t_1 \leq t_2$. Indeed, the control is now uniquely specified: (1) the sign of e(t) is the sign of the control for $0 \leq t \leq t_1$, (2) t_1 is the time at which the control switches, (3) t_2 is the time at which the control is removed and at which the system state is zero.

Table 1 below lists the optimal solution of equations (5-8) for eight different sets of initial conditions.

Table 1. Optimal Solutions for R-L-C System

$z_1(0)$	$z_2(0)$	$v_4(0)$	$i_3(0)$	$e(t)$	t_1	t_2
2	3	-2	4	-1	1.3863	1.6094
3	2	2	1	-1	1.8477	2.1622
-5	9	-28	23	1	2.4112	2.7909
37	25	24	13	-1	4.1650	4.5085
-20	-12	-16	-4	1	3.7739	4.2212
-12	-20	16	-28	1	3.0445	3.3673
5	87	-164	169	1	1.7442	2.7371
-75	17	-184	109	1	4.8667	5.2138

The above data were obtained using Program I of Appendix C. This program solves the second order set (5-8), using the technique of Appendix A; i.e., the nonlinear algebraic set is transformed into an equivalent differential set which is solved by a Runge-Kutta method. Theorems (3-1) and (3-2) guide the choice of an approximate solution vector with which to begin the numerical process. The results obtained indicate that it is rather easy to choose an initial approximation which will converge to the desired solution.

Example 2

A higher order system shown schematically in Figure 5.2(a) consists of two masses interconnected with springs and dashpots and excited by the force driver, f(t). The springs and dashpots are described by linear terminal relations.

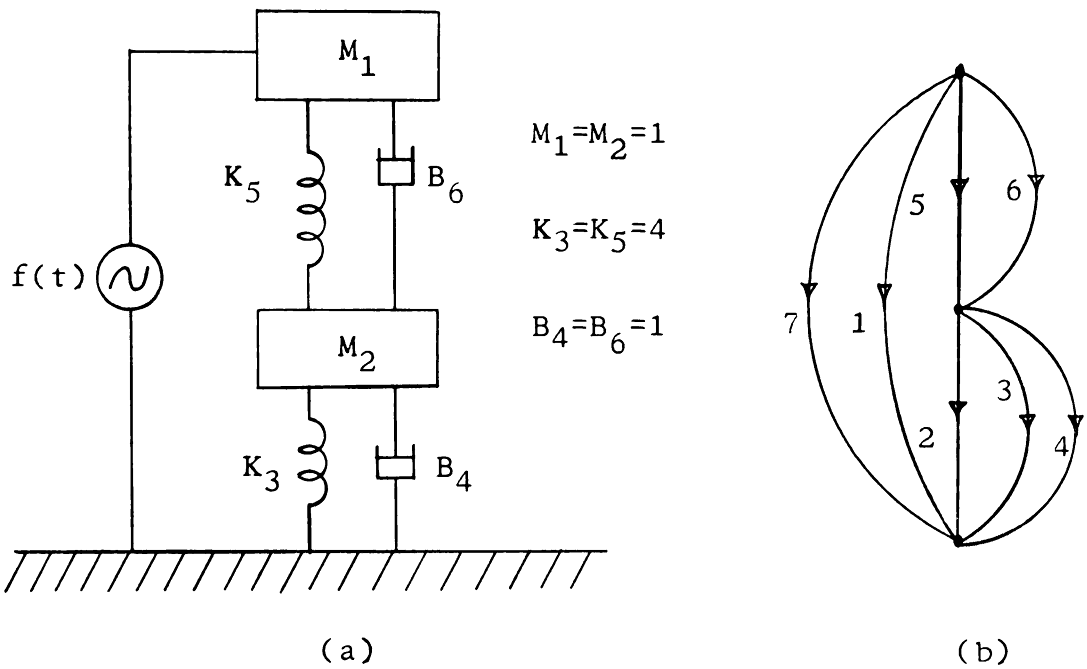


Figure 5.2 Mechanical System

A state model of this system based on the linear graph of Figure 5.2(b) is

$$\frac{d}{dt}\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ x_1\\ x_2\end{bmatrix} = \begin{bmatrix}-8 & 4 & 2 & 1\\ 4 & -4 & 1 & -1\\ 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\end{bmatrix}\begin{bmatrix}\dot{x}_1\\ \dot{x}_2\\ x_1\\ x_2\end{bmatrix} + \begin{bmatrix}0\\ -f(t)\\ 0\\ 0\end{bmatrix} \qquad (5\text{-}9)$$

where x_1 and x_2 are the displacements of masses M_1 and M_2, respectively, from the equilibrium position while $\dot{x}_1$ and $\dot{x}_2$ are the corresponding velocities.

The coefficient matrix in (5-9) has real, simple, negative eigenvalues, and the linear transformation which diagonalizes the system is

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0.1315 & 0.3103 & -0.6319 & 0.8673 \\ -0.0813 & 0.5021 & -1.0224 & -0.5360 \\ -0.5131 & -0.9854 & 0.5209 & -0.0849 \\ 0.3171 & -1.5943 & 0.8429 & 0.0525 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} \tag{5-10}$$

And (5-9) written in principal coordinates becomes

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} -0.2563 & 0 & 0 & 0 \\ 0 & -0.3149 & 0 & 0 \\ 0 & 0 & -1.2130 & 0 \\ 0 & 0 & 0 & -10.2159 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} + \begin{bmatrix} -0.0875 \\ 0.5054 \\ 0.9559 \\ 0.5289 \end{bmatrix} f(t) \tag{5-11}$$

The linear transformation

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \end{bmatrix} = \begin{bmatrix} -0.0875 & 0 & 0 & 0 \\ 0 & 0.5054 & 0 & 0 \\ 0 & 0 & 0.9559 & 0 \\ 0 & 0 & 0 & 0.5289 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} \qquad (5\text{-}13)$$

reduces (5-11) to the normalized form

$$\frac{d}{dt}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} = \begin{bmatrix} -0.2563 & 0 & 0 & 0 \\ 0 & -0.3149 & 0 & 0 \\ 0 & 0 & -1.2130 & 0 \\ 0 & 0 & 0 & -10.2159 \end{bmatrix}\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ z_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} f(t) . \qquad (5\text{-}13)$$

From (3-14) the nonlinear switching equations from which the optimal solution is obtained are

$$\begin{aligned} -\lambda_1 z_{10} &= f(t)(1-2e^{-\lambda_1 t_1}+2e^{-\lambda_1 t_2}-2e^{-\lambda_1 t_3}+e^{-\lambda_1 t_4}) \\ -\lambda_2 z_{20} &= f(t)(1-2e^{-\lambda_2 t_1}+2e^{-\lambda_2 t_2}-2e^{-\lambda_2 t_3}+e^{-\lambda_2 t_4}) \\ -\lambda_3 z_{30} &= f(t)(1-2e^{-\lambda_3 t_1}+2e^{-\lambda_3 t_2}-2e^{-\lambda_3 t_3}+e^{-\lambda_3 t_4}) \\ -\lambda_4 z_{40} &= f(t)(1-2e^{-\lambda_4 t_1}+2e^{-\lambda_4 t_2}-2e^{-\lambda_4 t_3}+e^{-\lambda_4 t_4}) \end{aligned} \qquad (5\text{-}14)$$

where the λ_i are the diagonal entries of (5-13).

Table 2 below lists the optimal time solution of equations (5-14) for two different sets of initial displacements and initial velocities of the masses, M_1 and M_2.

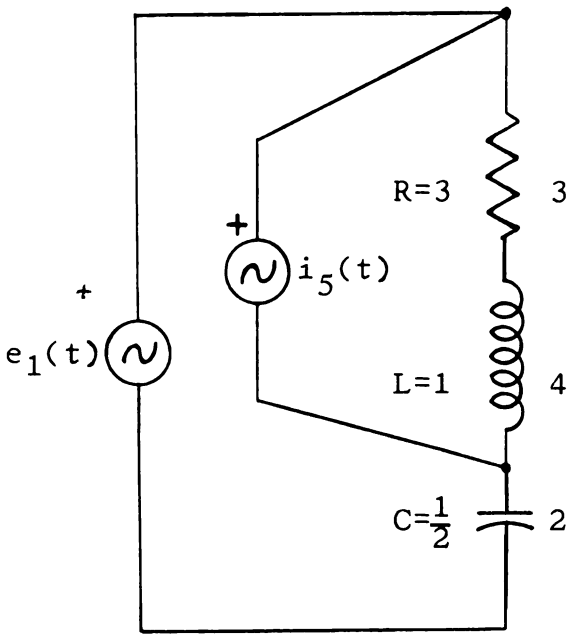
Table 2. Optimal Solutions for Mechanical System

$\dot{x}_1$	$\dot{x}_2$	x_1	x_2	Sign	t_1	t_2	t_3	t_4
1.533	-2.596	-0.633	-0.722	1	2.6299	5.4552	6.0723	6.1399
1.700	-4.405	0.229	0.971	1	3.1660	5.7931	6.4022	6.4699

The column labeled "sign" indicates the value of f(t), $0 \leq t \leq t_1$, and the t_i are the successive times at which f(t) reverses sign. Program II of Appendix C was used to solve (5-14) for the optimal controls of Table 2. The choice of the initial, approximate solution was based upon theorems (3-1) and (3-2). As one might expect, the choice of an initial approximation which converges to the desired solution is somewhat more critical for the fourth-order system than it is for second order-systems.

Vector Control Problem

A simple example of a physical system whose mathematical model fits the structure of the vector control problem is the R-L-C system of Figure 5.3(a).



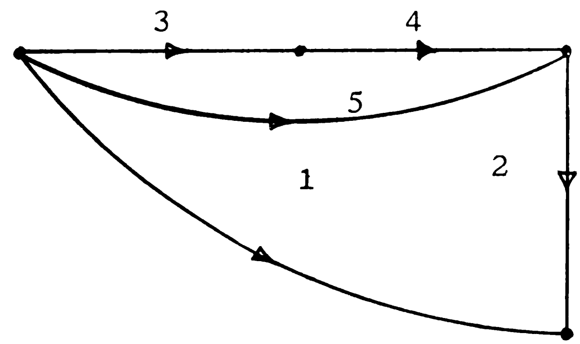


Figure 5.3 Two driver R-L-C system.

A state model of the system, formulated from the linear graph 5.3(b) is [26]

$$\frac{d}{dt}\begin{bmatrix} v_2 \\ i_4 \end{bmatrix} = \begin{bmatrix} 0 & 2 \\ -1 & -3 \end{bmatrix}\begin{bmatrix} v_2 \\ i_4 \end{bmatrix} + \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}\begin{bmatrix} i_5(t) \\ e_i(t) \end{bmatrix} . \qquad (5\text{-}15)$$

The diagonalized system becomes

$$\frac{d}{dt}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} + \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}\begin{bmatrix} i_5(t) \\ e_i(t) \end{bmatrix} \qquad (5\text{-}16)$$

where

$$\begin{bmatrix} v_2 \\ i_4 \end{bmatrix} = \begin{bmatrix} 1 & -1 \\ -\frac{1}{2} & 1 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} . \tag{5-17}$$

Application of the extended Lagrange method to (5-15) leads to the following fifth-order system corresponding to (4-27), which must be solved to obtain the optimal control

$$\frac{\partial f}{\partial s_{11}} + Z \frac{\partial g}{\partial s_{11}} = 0$$

$$\frac{\partial f}{\partial s_{21}} + Z \frac{\partial g}{\partial s_{21}} = 0$$

$$g = 0 \tag{5-18}$$

$$f - s_{11}^2 - s_{12}^2 = 0$$

$$f - s_{21}^2 - s_{22}^2 = 0$$

where

$$f(s_{11},s_{21}) = \log \left[\frac{-Y_{10}+4u_1(1-2e^{s_{11}^2})+2u_2(1-2e^{s_{21}^2})}{-4u_1-2u_2}\right]$$

$$g(s_{11},s_{21}) = -2y_{20}+2u_1(1-2e^{2s_{11}^2})+2u_2(1-2e^{2s_{21}^2})$$

$$+(2u_1+2u_2)\left[\frac{-y_{10}+4u_1(1-2e^{s_{11}^2})+2u_2(1-2e^{s_{21}^2})}{-4u_1-2u_2}\right]^2$$

If t_{11} and t_{21} represent, respectively, the times at which the first and the second controls switch and t_o is the time at which control is removed, then the following equations relate the t_{ij} and the s_{ij}

$$\begin{aligned} t_{11} &= s_{11}^2 \\ t_{21} &= s_{21}^2 \\ t_o &= s_{11}^2+s_{12}^2 \\ t_o &= s_{21}^2+s_{22}^2 \end{aligned} \qquad (5\text{-}19)$$

Program III of Appendix C was used to solve (5-18) for the variables, S_{ij} and Z. Equation (5-19) was used to determine the switching times of the optimal control. Table 3 below lists the solutions for several different sets of initial conditions.

Table 3. Optimal Solutions for Two Control R-L-C System

$v_2(0)$	$i_4(0)$	i_5	e_1	t_{11}	t_{21}	t_o
-14	11.5	-1	1	0.8765	1.1616	2.1810
12	6.5	1	-1	0	1.3013	2.6498
-8	-2.0	-1	1	0	0.9729	2.0423
-82	84.5	1	1	1.6468	1.2030	2.1943

The columns labeled i_5 and e_1 indicate the sign of i_5 and e_1 on the intervals, $0 \leq t \leq t_{11}$, $0 \leq t \leq t_{21}$, respectively.

VI. CONCLUSION

The time optimal control of physical systems described by a set of first-order linear, constant coefficient differential equations has been extensively discussed in the literature during the past few years. Most researchers have been concerned with establishing the salient mathematical features of the optimal control, a.e., existence and uniqueness, while only a handful have studied techniques for finding the control. This thesis has developed and extended techniques for determining the optimal control for that class of systems which is completely controllable and which has simple, real eigenvalues.

The introduction traces the history of the time optimal problem from its genesis in the relay controller up through the rigorous analysis in a precise mathematical form. Section two contains a precise mathematical statement of the time optimal problem. Pontryagin's maximum principle is used to show the bang-bang nature of the optimal control, and a theorem relating the number of switchings of each control component to the order of the system is stated.

The scalar control problem is discussed extensively, and a transcendental set of equations in the switching times which the optimal control must satisfy is developed.

Theorems bounding the optimal time are stated and proved.

The vector control problem is considered separately since it is considerably more complex than the scalar case. The equations which the vector control must satisfy contain more unknowns than equations, and consequently the set has infinitely many solutions. The problem is reformulated as a minimization of the final control time, t_n, subject to a set of equality constraints and a set of inequality constraints. Simple extensions of the Lagrange multipliers permit handling the inequalities, and a normal set which the optimal vector satisfies is derived. Again, theorems bounding the optimal control time are given.

In Appendix A a numerical technique for solving nonlinear algebraic equations is developed. This procedure, based upon some of Wirth's work [26], transforms the algebraic set into a differential set whose solution at one endpoint of the interval, [0, 1], is a root of the algebraic set. Other methods applicable to the solution of nonlinear equations are included in Appendix B.

Several examples of physical systems are analyzed. The state models and the transcendental equations in the switching times are developed. The results of numerical solutions carried out on a digital computer are listed in tabular form for both the scalar and the vector control.

APPENDIX A

NUMERICAL SOLUTION OF NONLINEAR ALGEBRAIC EQUATIONS

The analysis of many engineering problems has been hindered by systems of nonlinear algebraic equations. Very little is known about the properties of the solutions of such systems. Indeed, the very question of the existence of solutions to such a set can be fully answered only for very special subclasses such as polynomials.

Wirth gives sufficient conditions for the existence of a unique solution to such a set and devised an algorithm for obtaining the solution [26]. His theorem is stated below.

Theorem

Let $G(T,X) = 0$ be an n-dimensional vector function of the n-dimensional vector, T, and the r-dimensional parameter vector, X. For every X such that:

1. The entries of $\frac{\partial F}{\partial T}$ exist, are bounded for all T and satisfy a Lipschitz condition on T for all T.
2. $\left|\det \frac{\partial F}{\partial T}\right| \geq k > 0$ for all T, k a constant.

Then there exists a unique T such that $G(T,X) = 0$.

The hypotheses of this theorem are necessarily quite restrictive since a unique solution is required. Thus, for example, a single polynomial equation of degree two or higher will not satisfy the hypotheses.

Therefore, while the theorem is extremely useful for a narrow class of systems, it is not applicable to a broad class of problems of engineering interest.

In practice, existence alone may be the significant property of the system of equations, i.e., even though a set may have many, even infinitely many, solutions one particular solution may provide an acceptable result.

Algebraic Systems

The theorem stated and proved below is an extension of Wirth's work. It lists sufficient conditions for the existence of a solution to a nonlinear equation set, $G(T,X) = 0$, in some compact region, R. Perhaps more significant is the fact that an algorithm for obtaining a solution is contained in the proof.

Theorem A

Given:

1. $G(T,X) = 0$, an n-dimensional vector function of the n-dimensional vector, T, and the r-dimensional vector, X.
2. A compact region, R, in the $(n+r)$ space defined

by:

$$\|T - T_o\| \leq C$$

$$\|X - X_o\| \leq D_1$$

3. For all $T \varepsilon R$, $X \varepsilon R$ the entries of $\frac{\partial G}{\partial T}$ exist, are bounded and satisfy a Lipschitz condition with respect to T.

4. $\left|\det \frac{\partial G}{\partial T}\right| \geq k > 0$ in R

5. $\|G(T_o,\bar{X})\| < \frac{C}{M}$ where $M = \max_R \left\|\frac{\partial G(T,\bar{X})}{\partial T}^{-1}\right\|$ and $\bar{X}$ is a particular $X \varepsilon R$

then:

1. There exists $\bar{T} \varepsilon R$ such that $G(\bar{T},\bar{X}) = 0$
2. $\bar{T} = T(0)$

   where

$$\frac{dT}{dt} = \left[\frac{\partial G(T,\bar{X})}{\partial T}\right]^{-1} G(T_o,\bar{X}) \qquad T(1)=T_o \qquad \text{(A-1)}$$

Proof:

1. Consider T as a function of a scalar independent variable t so that we have

$$G(T(t),\bar{X} = 0$$

2. Define a function $H(T(t),t) = G(T(t),\bar{X})-tG(T_o,\bar{X})$ and differentiate with respect to t:

$$\frac{dH}{dt} = \frac{\partial G(T(t),\bar{X})}{\partial T} \cdot \frac{dT}{dt} - G(T_o,\bar{X}) \tag{A-2}$$

Note that $\frac{\partial G}{\partial T}$ exists and is nonsignular everywhere in R by hypothesis.

3. Now assume $\frac{dH}{dt} = 0$ which implies that

$$\frac{dT}{dt} = \left(\frac{\partial G(T(t),\bar{X})}{\partial T}\right)^{-1} G(T_o,\bar{X}) \tag{A-3}$$

4. Each entry in $\left[\frac{\partial G}{\partial T}\right]^{-1}$ satisfies a Lipschitz condition with respect to T. This holds because $\left|\det \frac{\partial G}{\partial T}\right| \geq k$ and sums and products of Lipschitz functions are again Lipschitz functions from Lemma 2.

5. The right hand side of (A-3) also satisfies

$$\max_T \left\|\frac{\partial G}{\partial T}^{-1} G(T_o\bar{X})\right\| \leq \left\|\frac{\partial G}{\partial T}^{-1}\right\| \left\|G(T_o,\bar{X})\right\| < M \cdot \frac{C}{M} = C \tag{A-4}$$

from the hypothesis.

6. Thus the differential equation (A-3) satisfies all the hypotheses of Lemma 1; therefore, it has a unique solution, T(t), such that T(t) is interior to R for $t \epsilon [0,1]$ and $T(1) = T_o$.

7. Substituting (A-3) into (A-2) shows that

$$\frac{dH}{dt} \equiv 0$$

Integrating we have

$$\int_{1}^{0} \frac{d}{dt} H(T(t),\overline{X},t)dt \equiv 0=H(T(t),\overline{X},t)\Big|_{1}^{0}$$

Therefore

$$H(T(0),\overline{X},0)-H(T(1),\overline{X},1) \equiv 0$$

and

$$G(T(0),\overline{X})-G(T(1),\overline{X})+G(T_o,\overline{X}) \equiv 0 \qquad \text{(A-6)}$$

8. Since the initial condition for the differential equation (A-3) is chosen such that $T(1) = T_o$, (A-6) above becomes

$$G(T(0),\overline{X}) \equiv 0 \qquad \text{(A-7)}$$

and $T(0)$, the solution of (A-3) evaluated at $t = 0$, is a solution to the algebraic set, $G(T,\overline{X}) = 0$.

Mixed Algebraic and Differential Systems

In formulating nonlinear mathematical models of physical systems using linear graph techniques, the final representation is often of the form

$$\dot{X} = F(X,Y,E(t)) \qquad X(0) = C \qquad \text{(A-8)}$$

$$G(X,Y) = 0$$

where $E(t)$ is a known function of time. If $G(X,Y) = 0$ can be solved analytically for $Y = H(X)$, the solution can be substituted in the differential set to yield

$$\dot{X} = F(X,H(X),E(t)) \qquad X(0) = C \tag{A-9}$$

and this equation can then be solved for $X(t)$. Hence the statement that the system has a unique solution. Wirth, in the theorem previously mentioned, stated conditions sufficient for the existence of a unique solution to $G(X,Y) = 0$, and, therefore, also to the whole system (A-8). However, this result includes a rather limited subclass of the totality of functions, $G(X,Y) = 0$. For example, the polynomial $g(x,y) = y^2-x^2 = 0$ has two well-behaved solutions $y_1(x) = x$, $y_2(x) = -x$. It does not satisfy the hypotheses of Wirth's theorem because it does not have a unique solution; nevertheless in this case a non-unique complete solution to (A-8) can be obtained.

Therefore, it appears that in some cases sufficient conditions for the existence of a complete solution to (A-8) are desirable even though that solution is non-unique. The following theorem is addressed to this problem.

Theorem B

Given:

1. $$\begin{aligned} \dot{X} &= F(X,Y,E(t)) \qquad X(0) = X_o \\ G(X,Y) &= 0 \end{aligned} \tag{A-10}$$

2. F continuous in X, Y and E for X, Y ε R defined by $\|X-X_o\| \leq D_1$, $\|Y-Y_o\| \leq C$.
3. G satisfies all hypotheses of theorem A, and in addition $\left[\frac{\partial G}{\partial Y}\right]^{-1}$ satisfies a Lispschitz condition with respect to X as well as Y for all X, Y in R.
4. E(t) is piecewise continuous.

Then there exists a continuous solution, X(t), to (A-10) defined on $0 \leq t \leq t_2$ satisfying $X(0) = X_o$ where

$$t_2 = \min\left[t_o, t_1\right]$$

t_1 is first t such that $\|X(t_1) - X_o\| = D_1$

t_o is first t such that $\|Y(t_o) - Y_o\| = C$.

<u>Proof</u>

1. By <u>Theorem A</u> there exists Y_o such that $G(Y_o,X_o) = 0$ and Y_o is interior to R.
2. Y is determined as the solution of the differential equation

$$\frac{dY}{dS} = \frac{\partial G(Y,X)}{\partial Y}^{-1} G(\underline{Y},X) \qquad Y(1) = \underline{Y} \qquad \text{(A-11)}$$

evaluated at S = O, Y(0). But since the right hand side of (A-11) satisfies a Lipschitz condition with respect to both X and Y, by Lemma 4 the solution vector Y(S,X) is a continuous function of the parameter vector, X, so that

$$\|Y(0,X_o)-Y(0,X_1)\| < \varepsilon \text{ if only } \|X_1-X_o\| < \delta \qquad \text{(A-12)}$$

3. Thus applying a numerical solution technique to

$$\dot{X} = F(X,Y,E)$$

gives

$$X_1 = X_o+F_1(X_o,Y_o,E) \text{ where } \|F_1(X_o,Y_o,E)\| < \delta \quad \text{(A-13)}$$

by taking a suitably small step size.

4. Using X_1 and Y_o in (A-11) a new solution Y_1 is determined, which by (A-12) satisfies

$$\|Y_1 - Y_o\| < \varepsilon$$

5. Repeating the above procedure a sequence

$$Y(X_o), Y_1(X_1), ---Y_n(X_n), ---$$

is generated where

$$\|Y_i - Y_{i-1}\| < \varepsilon \text{ for all } i$$

6. Using this sequence of Y_i's a numerical solution to (A-10) is obtained over range

$$\|X_i - X_o\| \leq D_1$$

$$\|Y_i - Y_o\| \leq C$$

Lemma 1

Given:

1. $\dot{Z} = f(Z) \qquad Z(1) = C \qquad$ (A-14)

2. A closed region, R, defined by $\|Z - C\| \leq C_1$

3. There exists C_3 such that for all Z_1, Z_2 in R:

$$\|f(Z_2)-f(Z_1)\| \leq C_3 \|Z_2-Z_1\|$$

4. $C_2 = \max_R \|f(Z)\|$, $C_2 < C_1$

Then there exists a unique solution, $Z(t)$, for $t \in [0,1]$ such that $Z(t) \in R$ and $Z(1) = C$.

Proof (Existence)

1. The following integral equation is completely equivalent to the initial value problem (A-14):

$$Z(t) = C - \int_t^1 f[Z(t^1)]dt^1 \qquad \text{(A-15)}$$

2. Define a sequence of functions $\left\{Z_n(t)\right\}$ as follows:

$$Z_o = C$$

$$\vdots$$

$$Z_{n+1} = C - \int_t^1 f[Z_n(t^1)]dt^1$$

Note that we know about the existence of the above integral only if $Z_n \in R$ for all n.

3. <u>Assertion</u>: $Z_n \in R$ for all n, all $t \in [0,1]$

By induction

a) $Z_o = C \in R$

b) Assume $Z_n(t) \in R$

c) Then:

$$Z_{n+1}(t) = C - \int_t^1 f[Z_n(t^1)]dt^1$$

$$\|Z_{n+1} - C\| \leq \int_t^1 \|f(Z_n)\| dt^1 \leq C_2(1-t) \leq C_2 \text{ for } t \in [0,1]$$

Hence:

$$\|Z_{n+1} - C\| \leq C_2 < C_1 \text{ for } t \in [0,1], \text{ for all } n \quad \text{(A-16)}$$

from hypothesis. Hence we restrict t, $0 \leq t \leq 1$. Note the strict inequality implies Z_n is an interior point of R.

4. Now consider:

$$Z_1 - Z_o = - \int_t^1 f(Z_o)dt^1$$

$$\|Z_1 - Z_o\| \leq \int_t^1 \|f(Z_o)\| dt^1 \leq C_2(1-t)$$

$$\|Z_2 - Z_1\| \leq \int_t^1 \|f(Z_1) - f(Z_o)\| dt^1$$

And by hypothesis (Lipschitz condition):

$$\left\|Z_2 - Z_1\right\| \leq C_3 \int_t^1 \left\|Z_1 - Z_o\right\| dt^1 \leq C_2 C_3 \int_t^1 (1-t^1) dt^1$$

$$= \frac{C_2 C_3 (1-t)^2}{2}$$

By induction for any n:

$$\left\|Z_{n+1} - Z_n\right\| \leq C_2 C_3^{\,n} \frac{(1-t)^{n+1}}{(n+1)!} \qquad \text{(A-17)}$$

5. Hence it follows that $\left[Z_n\right]$ converges uniformly for any interval $[0, t_o]$, $t_o \leq 1$, to a limit function,

$$Z(t) = C - \int_t^1 f[Z(t^1)] dt^1 \qquad \text{(A-18)}$$

which satisfies the integral equation and hence the differential equation (A-14).

<u>Uniqueness</u>:

6. It remains to show that the solution, Z(t), derived by successive approximations is, under our hypotheses, the only solution of

$$\dot{Z} = f(Z) \qquad Z(1) = C$$

on the interval $0 \leq t \leq 1$.

7. Assume there exists a second solution, $Y(t)$, such that $Y(1) = C$. Since Y is continuous and in R at time $t = 1$, it is in R for $0 \leq t_1 \leq t \leq 1$. Let $t_2 = \max\,[0, t_1]$, then we have for $t_2 \leq t \leq 1$:

$$Y(t) - Z(t) = -\int_t^1 \left[f(Y) - f(Z)\right]\, dt^1$$

$$\|Y(t)-Z(t)\| \leq C_3 \int_t^1 \|Y-Z\| dt^1 < \varepsilon + C_3 \int_t^1 \|Y-Z\| dt^1 \qquad \text{(A-19)}$$

where ε is constant > 0

By Gronwall's Lemma [28]

$$\|Y(t)-Z(t)\| \leq \varepsilon e^{C_3(1-t)} \qquad \text{(A-20)}$$

Hence $Y(t) = Z(t)$, $t_2 \leq t \leq 1$, since ε is arbitrarily small.

If $t_2 = 0$ the proof is complete; if $t_2 = t_1 > 0$ we have a contradiction since t_1 was chosen so that $Y(t_1)$ was on the boundary of R. But

$$Y(t_1) = Z(t_1) \qquad \text{(A-21)}$$

which implies $Z(t_1)$ is on the boundary of R also. However by equation (A-16) $Z(t_1)$ is an interior point of R. Hence for $0 \leq t \leq 1$, $Y(t) = Z(t)$.

Lemma 2

Given:

1. $\left|f(x)-f(y)\right| \leq K_1\left|x-y\right|$

$$\left|g(x)-g(y)\right| \leq K_2\left|x-y\right|$$

2. $x, y \in R$, some compact region

Then:

1. $\left|f(x)+g(x)-f(y)-g(y)\right| \leq K_3\left|x-y\right|$

2. $\left|f(x)g(x)-f(y)g(y)\right| \leq K_4\left|x-y\right|$

3. $\left|\frac{1}{f(x)} - \frac{1}{f(y)}\right| \leq K_5\left|x-y\right|$ if $f(Z) \geq k > 0$, $Z \in R$

Proof

1. $\left|f(x)+g(x)-f(y)-g(y)\right| \leq \left|f(x)-f(y)\right|$

$$+\left|g(x)-g(y)\right| \leq k_1\left|x-y\right| + k_2\left|x-y\right|$$

hence:

$$\left|f(x)+g(x)-f(y)-g(y)\right| \leq k_3\left|x-y\right| \quad , \; k_3 = k_1+k_2$$

hence assertion 1.

2. $\left|f(x)g(x)-f(y)g(y)\right| = \left|f(x)g(x)-f(x)g(y)+f(x)g(y)\right.$

$$\left.-f(y)g(y)\right| \leq \left|f(x)g(x)-f(x)g(y)\right| + \left|-f(y)g(y)\right.$$

$$\left.+f(x)g(y)\right| \leq \left|f(x)\right|\left|g(x)-g(y)\right| + \left|g(y)\right|\left|f(x)-f(y)\right|$$

But since f(x) and g(x) are continuous on a compact region they have maxima, M_1 and M_2, respectively in R so that

$$\left|f(x)g(x)-f(y)g(y)\right| \leq M_1K_2\left|x-y\right| + M_2K_1\left|x-y\right| \leq K_4\left|x-y\right|$$

$$\left|f(x)g(x)-f(y)g(y)\right| \leq K_4\left|x-y\right|$$

hence assertion 2.

3. $$\left|\frac{1}{f(x)} - \frac{1}{f(y)}\right| = \left|\frac{1}{f(x)}\,\frac{1}{f(y)}\,[f(y)-f(x)]\right|$$

$$\left|\frac{1}{f(x)} - \frac{1}{f(y)}\right| \leq \left|\frac{1}{f(x)}\right|\left|\frac{1}{f(y)}\right| K_1 \left|x-y\right|$$

$$\left|\frac{1}{f(x)} - \frac{1}{f(y)}\right| \leq \frac{K_1}{k^2} \left|x-y\right| = K_5\left|x-y\right|$$

hence assertion 3.

<u>Lemma 3 (Gronwall's Lemma)</u>

Given:

1. $u, v \geq 0$, C_1, a positive constant

2. $$u \leq C_1 + \int_o^t u\, v\, dt_1 \qquad \text{(A-22)}$$

Then:

$$u \leq C_1\, e^{\int_o^t v\, dt_1} \qquad \text{(A-23)}$$

Proof

1. From (1):

$$\frac{u}{C_1 + \int_0^t u\ v\ dt_1} \leq 1$$

$$\frac{u\ v}{C_1 + \int_0^t u\ v\ dt_1} \leq v \qquad \text{(A-24)}$$

2. Integrate over 0 to t:

$$\int_0^t \frac{u(S)v(S)\quad dS}{C_1 + \int_0^S u(S_1)v(S_1)dS_1} \leq \int_0^t v(S)dS \qquad \text{(A-25)}$$

note:

$$u(S)v(S) = \frac{d}{dS}\left[C_1 + \int_0^S u(S_1)v(S_1)\ dS_1\right]$$

Thus:

$$\log\left[C_1 + \int_0^S u(S_1)v(S_1)dS_1\right]\Bigg|_0^t \leq \int_0^t v(S)dS$$

$$\log\left\{\frac{C_1 + \int_0^t u(S_1)v(S_1)dS_1}{C_1}\right\} \leq \int_0^t v(S)dS \qquad \text{(A-26)}$$

And:

$$C_1 + \int_o^t u(S_1)v(S_1)dS_1 \le C_1\, e^{\int_o^t v(S)dS}$$

3. Finally using the second hypothesis:

$$u \le C_1 + \int_o^t u(S)v(S)dS \le C_1\, e^{\int_o^t v(S)dS}$$

$$u \le C_1 e^{\int_o^t v(S)dS}$$

<u>Lemma 4</u>

Given:

1. $\dot{X} = F(X,\alpha) \qquad X(0) = X_o$ (A-27)

   where X is an n-vector, α an r-vector

2. $F(X,\alpha)$ satisfies a Lipschitz condition with respect to both X and α in some compact (n+r) -space, R.

Then:

the solution, $X(t,\alpha)$, given by:

$$X(t,\alpha) = X_o + \int_o^t F[X(t_1),\alpha]dt_1$$

is a continuous function of the vector parameter, α.

Proof

1. Define:

$$\overline{X} = \begin{bmatrix} X \\ \alpha \end{bmatrix} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \\ a_1 \\ \vdots \\ a_r \end{bmatrix} \text{ then: } \quad \dot{\overline{X}} = \begin{bmatrix} \dot{x}_1 \\ \dot{x}_n \\ 0 \\ \\ 0 \end{bmatrix} = \begin{bmatrix} f_1(\overline{X}) \\ f_n(\overline{X}) \\ 0 \\ \\ 0 \end{bmatrix} = G(\overline{X}) \quad \text{(A-28)}$$

where:

$$\alpha = \begin{bmatrix} a_1 \\ \vdots \\ a_r \end{bmatrix} \qquad \overline{X}(0) = \begin{bmatrix} x_1(0) \\ \vdots \\ x_n(0) \\ a_1 \\ \vdots \\ a_r \end{bmatrix} = \begin{bmatrix} X_o \\ \alpha \end{bmatrix}$$

2. Now we have the system:

$$\dot{\overline{X}} = G(\overline{X}) \qquad \overline{X}(0) = \begin{bmatrix} X_o \\ \alpha \end{bmatrix} \qquad \text{(A-29)}$$

where $G(\overline{X})$ now satisfies a Lipschitz condition with respect to $\overline{X}$ in R.

3. Solution to (2) is:

$$\overline{X}(t) = \overline{X}(0) + \int_o^t G(\overline{X})dt^1 \qquad \text{(A-30)}$$

4. Consider two solutions at time t, $\overline{X}_1(t)$ and $\overline{X}_2(t)$, with different initial conditions

$$\overline{X}_2(t)-\overline{X}_1(t) = \overline{X}_2(0)-\overline{X}_1(0) + \int_0^t [G(\overline{X}_2(t))-G(\overline{X}_1(t))]dt^1$$

$$\left\|\overline{X}_2(t)-\overline{X}_1(t)\right\| \leq \left\|\overline{X}_2(0)-\overline{X}_1(0)\right\| + \int_0^t \left\|G(\overline{X}_2)-G(\overline{X}_1)\right\| dt^1$$

And by Lipschitz condition:

$$\left\|\overline{X}_2(t)-\overline{X}_1(t)\right\| \leq \left\|\overline{X}_2(0)-\overline{X}_1(0)\right\| + K\int_0^t \left\|\overline{X}_2(t)-\overline{X}_1(t)\right\| dt^1$$

Now by Lemma 3, Gronwall's Lemma:

$$\left\|\overline{X}_2(t)-\overline{X}_1(t)\right\| \leq \left\|\overline{X}_2(0)-\overline{X}_1(0)\right\| e^{Kt}$$

If $\left\|\overline{X}_2(0)-\overline{X}_1(0)\right\| < \delta$ (A-31)

Then:

$$\left\|\overline{X}_2(t)-\overline{X}_1(t)\right\| < \delta e^{Kt} = \varepsilon \qquad \text{(A-32)}$$

Hence

$$\left\|\overline{X}_2(t)-\overline{X}_1(t)\right\| < \varepsilon \text{ if only } \left\|\overline{X}_2(0)-\overline{X}_1(0)\right\| < \delta \qquad \text{(A-33)}$$

and the conclusion that $\overline{X}(t)$ is a continuous function of its initial conditions.

5. In particular if $X_1(0) = X_2(0)$ then

$$\bar{X}_2(0) - \bar{X}_1(0) = \alpha_2 - \alpha_1 \tag{A-34}$$

From the definition of $\bar{X}$ in (A-28)

$$\left\| X_2(t, \alpha_2) - X_1(t, \alpha_1) \right\| \leq \left\| \bar{X}_2(t) - \bar{X}_1(t) \right\| \tag{A-35}$$

Using (A-33), (A-34), (A-35)

$$\left\| X_2(t, \alpha_2) - X_1(t, \alpha_1) \right\| < \varepsilon \text{ if only } \left\| \alpha_2 - \alpha_1 \right\| < \delta \tag{A-36}$$

But (A-36) is precisely the statement that $X(t, \alpha)$ is a continuous function of the parameter, α.

APPENDIX B

NUMERICAL METHODS

This section briefly describes two of the standard techniques, Newton's method and the gradient method, which may be employed to generate numerical solutions to any set of n equations in n unknowns. In particular, these algorithms are applicable to the time optimal control equations of Chapter III, (3-14), and Chapter IV, (4-27).

Newton's Method

The n equations

$$f_i(y_1, y_2, \text{---}, y_n) = 0 \qquad i = 1, 2, \text{---}, n$$

in the n unknowns $y_1, y_2, \text{---}, y_n$ can be written in vector form as

$$F(Y) = 0 \tag{B-1}$$

If each of the functions, $f_i(y_i) = 0$, can be expanded in a Taylor series in n variables about some n-dimensional point, $y_1^{(0)}\ y_2^{(0)}\ \text{---}\ y_n^{(0)}$, we have

$$f_i(y_1, y_2, ---, y_n) = f_i(y_1^{(0)} y_2^{(0)} --- y_n^{(0)})$$

$$+ \sum_{j=1}^{n} (y_j - y_j^{(0)}) \frac{\partial f_i(y_1^{(0)} --- y_n^{(0)})}{\partial y_j}$$

$$+ \frac{1}{2} \sum_{k=1}^{n} \sum_{j=1}^{n} (y_k - y_k^{(0)})(y_j - y_j^{(0)}) \frac{\partial^2}{\partial y_j \partial y_k} f_i(y_1^{(0)} --- y_n^{(0)})$$

$$+ --- \qquad i = 1, 2, ---, n \qquad \text{(B-2)}$$

Discarding the second and higher order terms $f_i(Y)$ can be approximated by

$$f_i(Y) = f_i(Y^{(0)}) + \sum_{j=1}^{n} (y_j - y_j^{(0)}) \frac{\partial}{\partial y_j} f_i(Y^{(0)}) \qquad \text{(B-3)}$$

$$i = 1, 2, ---, n$$

Rewriting the system (B-3) in vector form we have

$$F(Y) = F(Y^{(0)}) + J(Y^{(0)})(Y - Y^{(0)}) \qquad \text{(B-4)}$$

where $J(Y)$ is the usual Jacobian matrix defined by

$$J(Y) = (A_{ij}) \qquad A_{ij} = \frac{\partial f_i(Y)}{\partial y_j}$$

If the vector, $Y^{(0)}$, is an approximate solution to (B-1), and if the matrix $J(Y^{(0)})$ is nonsingular, then we would expect the vector

$$Y = Y^{(0)} - J^{-1}(Y^{(0)})F(Y^{(0)}) \tag{B-5}$$

to be a more accurate approximate solution. This concept leads naturally to the following algorithm by which, hopefully, we are able to generate a sequence of successively closer approximations, $Y^{(n)}$, given by

$$Y^{(n)} = Y^{(n-1)} - J^{-1}(Y^{(n-1)})F(Y^{(n-1)}) \tag{B-6}$$

provided that $J^{-1}(Y^{(k)})$ exists for each k = 0, 1, 2, ---. This is Newton's Method for solution of the system, (B-1) [31],[33].

The sequence $Y^{(0)}, Y^{(1)}, ---, Y^{(n)}, ---$ beginning with an arbitrary $Y^{(0)}$ may not always converge to a solution of (B-1). Sufficient conditions for the convergence of Newton's Method to a solution were published by L. V. Kantorovich in 1937 [30],[31]. His theorem gives conditions under which the sequence, $Y^{(0)}, Y^{(1)}, ---, Y^{(n)}, ---$, converges to a solution, without assuming the existence of a solution a priori. His theorem is stated below.

Kantorovich's Theorem

Given:

1. The normal set, $F(Y) = 0$, where for $Y = Y^{(0)}$ the Jacobian inverse, $J^{-1}(Y^{(0)})$, exists and satisfies

$\left\| J^{-1}(Y^{(0)}) \right\| \leq B_o$, B_o a positive constant.

2. $\left\| J^{-1}(Y^{(0)})F(Y^{(0)}) \right\| \leq D_o$, D_o a positive constant.

3. In the region defined by (B-7) below, F(Y) is twice continuously differentiable with respect to the components of Y and

$$\sum_{j,k=1}^{n} \left| \frac{\partial^2 f_i}{\partial y_j \, \partial y_k} \right| \leq K \quad i = 1, 2, ---, n$$

4. The constants B_o, D_o, K satisfy

$$h_o = B_o D_o K \leq 1/2$$

Then:

1. $F(Y) = 0$ has a solution $\overline{Y}$ which is located in the region

$$\left\| Y - Y^{(0)} \right\| \leq N(h_o) D_o = \left[\frac{1 - \sqrt{1-2h_o}}{b_o} \right] D_o. \qquad \text{(B-7)}$$

2. The successive approximations, $Y^{(n)}$, defined by

$$Y^{(n)} = Y^{(n-1)} - J^{-1}(Y^{(n-1)})F(Y^{(n-1)})$$

exist and converge to $\overline{Y}$, and the speed of convergence may be estimated by the inequality

$$\left\| Y^{(n)} - \overline{Y} \right\| \leq \frac{1}{2(n-1)} (2h_o)^{(2^n - 1)} D_o.$$

Gradient Method

The gradient method (also called steepest descent) generates solutions to the n^{th} order normal set, $F(Y) = 0$, by minimizing the scalar function

$$\Omega = F^T M F \tag{B-8}$$

where M is a positive definite matrix, usually but not necessarily the unit matrix. Since M is positive definite Ω takes its absolute minimum if and only if $F = 0$. Thus every solution of $F(Y) = 0$ is an absolute minimum of Ω, and every absolute minimum of Ω corresponds to a solution of $F(Y) = 0$.

The technique used to find the minimum of Ω depends upon its geometry in hyperspace [33],[34]. For any particular value of K

$$\Omega = K \tag{B-9}$$

is an n-dimensional surface in hyperspace. For the two-dimensional case a contour map such as Figure 1 may be found. Here C is the point at which $\Omega = 0$, and the contour lines are constant values of Ω.

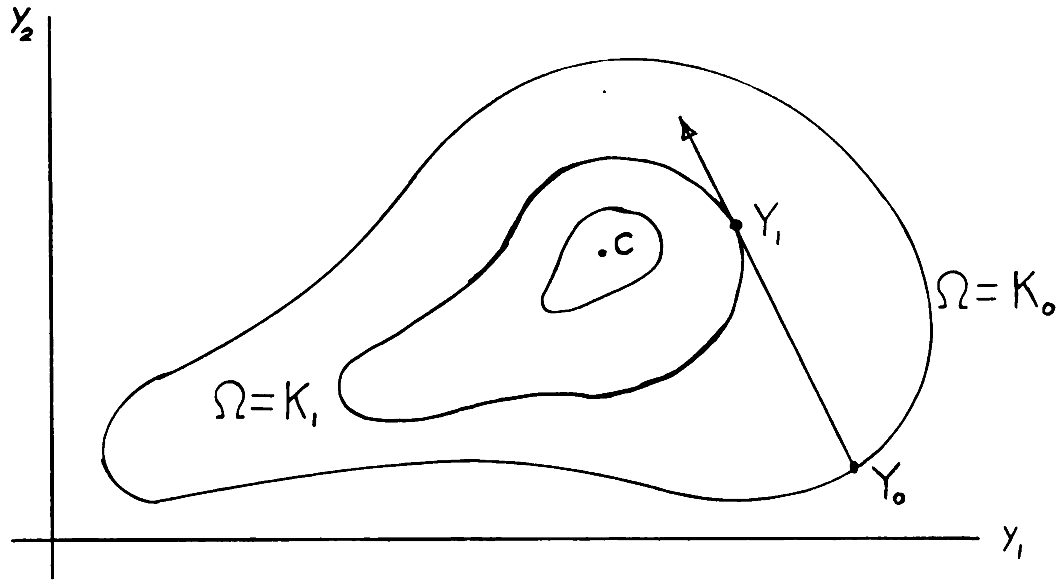


Figure B.1 Contour map of Ω= constant.

We seek a method of progressing from a vector point $Y_o = (Y_{10}, Y_{20}, ---, Y_{no})$ to another vector point, $Y_1 = Y_o + kZ$, where k is a scalar and Z is an arbitrary vector, such that

$$\Omega(Y_o + kZ) < \Omega(Y_o). \tag{B-10}$$

Thus we minimize the function

$$\Psi(k) = \Omega(Y_o + kZ) \tag{B-11}$$

of the single variable k by setting its derivative equal to zero:

$$\frac{d}{dk}\Psi(k) = Z^T \Omega_Y(Y_o + kZ) = 0 \tag{B-12}$$

Since condition (B-10) requires $k \neq 0$ we exclude this possibility by further specifying that

$$Z^T \Omega_Y(Y_o) \neq 0 \tag{B-13}$$

where Ω_Y, commonly called the gradient, is the column vector whose entries are

$$\Omega_{y_i} = \frac{\partial \Omega}{\partial y_i} . \tag{B-14}$$

Condition (B-13) states that the line in the Z direction through Y_o may <u>not</u> be orthogonal to the gradient at that point, and hence not tangent to the surface

$$\Omega(Y_o) = K_o .$$

Similarly, (B-12) says the line in the Z direction through Y_o is tangent to the surface

$$\Omega(Y_1) = K_1 \qquad K_1 < K_o$$

at Y_1 as shown in Figure B.1.

At Y_1 we choose a new direction, Z, and proceed as before. In this way a monotonically decreasing sequence, $\Omega(Y_i)$, is obtained which is bounded below by the minimum of $\Omega(Y)$ and therefore has a limit which is the minimum.

In the method of steepest descent the vector, Z, is always chosen as the gradient

$$Z = \Omega_Y \tag{B-15}$$

since Ω_Y is the direction of most rapid variation of Ω. However, each step involves many calculations, and a variety of simpler choices of z can be used.

Since solving (B-12) for the minimizing k may be exceedingly difficult, an approximate solution which simplifies the calculation is highly desirable. Consider the Maclaurin series expansion of $\Psi(k)$, the function to be minimized:

$$\Psi(k) = \Psi(0) + \Psi'(0)k + \frac{\Psi''(0)}{2!} k^2 + --- \qquad \text{(B-16)}$$

If we assume the second order approximation and minimize we have:

$$\begin{aligned} \Psi'(k) &= \Psi'(0) + \Psi''(0)k = 0 \\ k &= \frac{-\Psi'(0)}{\Psi''(0)} \end{aligned} \qquad \text{(B-17)}$$

And from (B-12)

$$\begin{aligned} \Psi'(0) &= Z^T \Omega_Y(Y_o) \\ \Psi''(0) &= Z^T J(Y_o) Z \end{aligned} \qquad \text{(B-18)}$$

where

$$J(Y_o) = (A_{ij}) = \frac{\partial^2}{\partial y_i \partial y_j} \Omega(Y) \Big|_{Y = Y_o}$$

Therefore, in the steepest descent method where

$$Z = \Omega_Y(Y_o)$$

an approximate solution for the scalar, k, at the n^{th} step of the process is

$$k_n = \frac{-\Omega_Y^T(Y_n)\ \Omega_Y(Y_n)}{\Omega_Y^T(Y_n)J(Y_n)\Omega_Y(Y_n)} \tag{B-19}$$

and the iterates are defined by

$$Y_{n+1} = Y_n + k_n\Omega_Y(Y_n). \tag{B-20}$$

The gradient method, by its very nature, always converges, yet there are some difficulties. These stem from the fact that the method converges to a stationary point of the surface, Ω. However, in general, the surface may have maxima, minima or saddle points all of which are stationary points. Therefore each "solution" obtained must be carefully examined to see if it is a true minimum. When a machine is used for solution this is readily accomplished by substituting the "solution" in the set, $F(Y) = 0$, to see if it checks.

In the case where $J(Y_i)$, $i = 1, 2, ---$, is positive definite we see from (B-19) that k_n is always negative. Therefore, the descent direction is that of the negative gradient and the sequence (B-20) always converges to a minimum though it may be a relative minimum.

APPENDIX C

This section contains a brief description of the digital computer programs used to obtain the numerical solutions tabulated in the examples. A complete listing of the programs is included after the program descriptions. All programs are written in the Fortran language.

Program I--Kurung

This program determines the unique time optimal control for the second order system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_1 \\ b_2 \end{bmatrix} u(t) \qquad \text{(C-1)}$$

It may be used exactly as listed provided a data deck, as outlined below, is inserted at the end of the program.

The first data card contains a single integer N, the number of initial conditions for which optimal solutions are desired, located on the card by statement 12. The second card contains the parameters P, Q, b_1, and b_2 in the format of statement 10. The N remaining data cards contain the initial conditions and the associated approximate switching times for which optimal solutions are desired. The initial conditions are x_{10} and x_{20} while t_1 is the estimated switching time and t_2 is the final control time.

The output consists of ten numbers: x_{10}, x_{20}, P, Q, U, S_{10}, S_{20}, t_1, t_2, ERROR. The value of U(t), $t_1 \leq t \leq t_2$, is given by U while t_1 and t_2 are the switching time and the final control time, respectively. The ERROR is the norm of the state vector at time $t = t_2$. The approximate switching times, S_{10} and S_{20}, are included for reference.

This program uses the numerical solution technique of Appendix A, and the differential set is solved by the fourth-order Runge-Kutta method.

Program II--Optima

This program solves the optimization equations which yield the time optimal control for the fourth-order system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} r_1 & 0 & 0 & 0 \\ 0 & r_2 & 0 & 0 \\ 0 & 0 & r_3 & 0 \\ 0 & 0 & 0 & r_4 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} U(t) \qquad \text{(C-2)}$$

It may be used as tabulated with the addition of a data deck.

The first data card, in the format of statement 25, contains a single integer N, the number of points for which optimal solutions are desired. The second card contains the four eigenvalues in order of increasing magnitude in the format of statement 15. The N remaining cards each contain the coordinates of an initial point and the approximate switching times associated with that point in the order: x_{10}, x_{20}, x_{30}, x_{40}, t_1, t_2, t_3, t_4 where $0 < t_1 < t_2 < t_3 < t_4$. Format statement 20 describes the field on the card.

The output consists of the initial conditions x_{10}, x_{20}, x_{30}, x_{40}, the value of U(t), $0 \le t \le t_1$, the switching times t_1, t_2, t_3, t_4, and the norm of the state vector at time $t = t_4$, E.

This program also uses the numerical technique of Appendix A employing the fourth-order Runge-Kutta method for solving the differential equations. The step size may be changed by replacing the statement H = -0.005 by one which reads the desired size. If this is done the 200 in the statement, DO 75, must be replaced with an integer K such such that HK = -1.0.

Program III--Inequa

This program determines the time optimal control vector for the system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} \begin{bmatrix} U_1(t) \\ U_2(t) \end{bmatrix} . \qquad \text{(C-3)}$$

If a data deck is supplied it may be used as listed.

The first data card, in the format of statement 14, contains an integer N which is the number of initial points for which solutions are desired. Card two, in the format of statement 12, contains the numbers P, Q, b_{11}, b_{12}, b_{21}, b_{22} in that order. The N remaining cards each contain a set of numbers x_{10}, x_{20}, S_{11}, S_{21}, S_{12}, S_{22} in the format of statement 10. The numbers x_{10} and x_{20} are initial conditions while the S_{ij} are related to the switching times t_1, t_2 and t_o by

$$
\begin{aligned}
t_1 &= S_{11}^2 \\
t_2 &= S_{21}^2 \\
t_o &= S_{11}^2 + S_{12}^2 \\
t_o &= S_{21}^2 + S_{22}^2
\end{aligned}
\tag{C-4}
$$

where t_1 and t_2 are the switching times of the first and second controls, respectively, and t_o is the final control time. The initial choices of the S_{ij} are made by assuming initial values for the switching times and then solving equations (C-4).

The output consists of U_1 and U_2, the values of the controls $U_1(t)$ and $U_2(t)$ initially and the switching times t_1, t_2, t_o. The norm of the state vector at $t = t_o$ is given by E while F is a measure of the error in the solution of the set corresponding to (4-27). A desired solution is one in which both F and E approach zero. The quantity D is the determinant of the Jacobian matrix which must be inverted at each iteration. Its value is significant only if it approaches zero, in which case the solution obtained is of questionable value.

Newton's method is used to solve the minimization equations in this program. In the event that the successive iterates do not converge to the desired solution a message to that effect is printed out.

Program Kurung

```
      PROGRAM KURUNG
    5 FORMAT(1H1,34HN B HEMESATH  RUNGE KUTTA SOLUTION)
   10 FORMAT(4(F10.2))
   12 FORMAT(I3)
   15 FORMAT(1H0,3HX10,7X,3HX20,7X,1HP,9X,1HQ,9X,1HU,9X,3HS10,7X,3HS20,
     17X,2HT1,8X,2HT2,8X,5HERROR)
   20 FORMAT(1H0,7(F10.2),3(E15.8))
      PRINT 5
      PRINT 15
      READ 12 M1
      READ 10 P,Q,B1,B2
      DO 25 I=1,M1
   35 READ 10 X10,X20,S10,S20
      H=-0.04
      N=25
      DO 25 I=1,2
      Y1=S10
      Y2=S20
      U=(-1.0)**I
      F1=((X10*P*U)/B1-1.0)+2.0*EXPF(-P*Y1)-EXPF(-P*Y2)
      F2=((X20*Q*U)/B2-1.0)+2.0*EXPF(-Q*Y1)-EXPF(-Q*Y2)
      DO 30 J=1,N
      D=2.0*P*Q*EXPF(-P*Y1-Q*Y2)*(EXPF((Q-P)*(Y2-Y1))-1.0)
      P11=(Q*EXPF(-Q*Y2))/D
      P12=(-P*EXPF(-P*Y2))/D
      P21=(2.0*Q*EXPF(-Q*Y1))/D
      P22=(-2.0*P*EXPF(-P*Y1))/D
      G1=P11*F1+P12*F2
      G2=P21*F1+P22*F2
      R11=H*G1
      R12=H*G2
      T1=Y1+0.5*R11
      T2=Y2+0.5*R12
      D=2.0*P*Q*EXPF(-P*T1-Q*T2)*(EXPF((Q-P)*(T2-T1))-1.0)
      R21=H*(Q*EXPF(-Q*T2)*F1-P*EXPF(-P*T2)*F2)/D
      R22=H*(2.0*Q*EXPF(-Q*T1)*F1-2.0*P*EXPF(-P*T1)*F2)/D
      T1=Y1+0.5*R21
      T2=Y2+0.5*R22
      D=2.0*P*Q*EXPF(-P*T1-Q*T2)*(EXPF((Q-P)*(T2-T1))-1.0)
      R31=H*(Q*EXPF(-Q*T2)*F1-P*EXPF(-P*T2)*F2)/D
      R32=H*(2.0*Q*EXPF(-Q*T1)*F1-2.0*P*EXPF(-P*T1)*F2)/D
      T1=Y1+R31
      T2=Y2+R32
      D=2.0*P*Q*EXPF(-P*T1-Q*T2)*(EXPF((Q-P)*(T2-T1))-1.0)
      R41=H*(Q*EXPF(-Q*T2)*F1-P*EXPF(-P*T2)*F2)/D
      R42=H*(2.0*Q*EXPF(-Q*T1)*F1-2.0*P*EXPF(-P*T1)*F2)/D
      Y1=Y1+(R11+2.0*R21+2.0*R31+R41)/6.0
      Y2=Y2+(R12+2.0*R22+2.0*R32+R42)/6.0
      IF(Y1*Y1+Y2*Y2-10.0**5.0) 30,45,45
```

```
30 CONTINUE
45 CONTINUE
   X1=(X10-(B1/P)*U)*EXPF(P*Y2)+(2.0*EXPF(P*(Y2-Y1))-1.0)*(B1/P)*U
   X2=(X20-(B2/Q)*U)*EXPF(Q*Y2)+(2.0*EXPF(Q*(Y2-Y1))-1.0)*(B2/Q)*U
   ERROR=SQRTF(X1*X1+X2*X2)
25 PRINT 20, X10,X20,P,Q,U,S10,S20,T1,T2,ERROR
   STOP
   END
   END
```

Progam Optima

```
   PROGRAM OPTIMA
 5 FORMAT(1H1,65HTIME OPTIMAL SOLUTION FOR FOURTH ORDER SYSTEM WITH S
  1CALAR CONTROL)
10 FORMAT(1H0,3X,3HX10,5X,3HX20,5X,3HX30,5X,3HX40,4X,1HU,9X,2HT1,13X,
  12HT2,13X,2HT3,13X,2HT4,13X,1HE)
11 FORMAT(1H0,67HJACOBIAN DETERMINANT APPROACHING ZERO. TRY NEW APPRO
  1XIMATE SOLUTION)
12 FORMAT(1H0,39HSOME SWITCHING TIME HAS BECOME NEGATIVE)
13 FORMAT(1H0,33HPRODUCT OF R4 AND T4 IS TOO LARGE)
15 FORMAT(4F10.2)
17 FORMAT(1H0,4F8.2,F5.1,5E15.8)
20 FORMAT(8F10.2)
25 FORMAT(I3)
   PRINT 5
   PRINT 10
   DIMENSION A(4,4),R(4),P(4,4),T(4),Y(4),F(4),G(4),S(4)
   H=-0.005
   N1=4
   D=1.0
   E=0.0
   J1=1
   READ 25 M1
   READ 15 (R(K),K=1,4)
   DO 100 I=1,M1
   READ 20 X10,X20,X30,X40, (S(K),K=1,4)
   DO 100 J=1,2
   U=(-1.0)**(J+1)
   G(1)=U*R(1)*X10+1.0-2.0*EXPF(-R(1)*S(1))+2.0*EXPF(-R(1)*S(2))
  1-2.0*EXPF(-R(1)*S(3))+EXPF(-R(1)*S(4))
   G(2)=U*R(2)*X20+1.0-2.0*EXPF(-R(2)*S(1))+2.0*EXPF(-R(2)*S(2))
  1-2.0*EXPF(-R(2)*S(3))+EXPF(-R(2)*S(4))
   G(3)=U*R(3)*X30+1.0-2.0*EXPF(-R(3)*S(1))+2.0*EXPF(-R(3)*S(2))
  1-2.0*EXPF(-R(3)*S(3))+EXPF(-R(3)*S(4))
   G(4)=U*R(4)*X40+1.0-2.0*EXPF(-R(4)*S(1))+2.0*EXPF(-R(4)*S(2))
```

```
   1-2.0*EXPF(-R(4)*S(3))+EXPF(-R(4)*S(4))
    DO 30 K=1,4
 30 Y(K)=S(K)
    DO 75 L=1,200
    DO 35 K=1,4
 35 T(K)=Y(K)
    DO 70 M=1,4
    DO 40 K=1,4
    DO 40 N=1,4
    IF(N-4) 37,38,38
 37 A(K,N)=(-1.0**(N-1))*2.0*R(K)*EXPF(-R(K)*T(N))
    GO TO 40
 38 A(K,N)=(-1.0**(N-1))*R(K)*EXPF(-R(K)*T(N))
 40 CONTINUE
    CALL INVERT (A,N1,D,J1)
    IF(D*D-(10.0**(-40))) 42,42,41
 41 CONTINUE
    DO 45 K=1,4
    F(K)=0.0
    DO 45 N=1,4
 45 F(K)=F(K)+A(K,N)*G(N)
    DO 50 K=1,4
 50 P(K,M)=H*F(K)
    IF(M-3) 55,65,65
 55 DO 70 K=1,4
    T(K)=Y(K)+P(K,M)/2.0
    GO TO 70
 65 DO 70 N=1,4
    T(N)=Y(N)+P(N,M)
 70 CONTINUE
    DO 80 K=1,4
 80 Y(K)=Y(K)+(1.0/6.0)*(P(K,1)+2.0*P(K,2)+2.0*P(K,3)+P(K,4))
    IF(Y(1)) 43,67,67
 67 IF(Y(2)) 43,68,68
 68 IF(Y(3)) 43,69,69
 69 IF(Y(4)) 43,72,72
 72 IF(-R(4)*T(4)-100.0) 75,44,44
 75 CONTINUE
    G(1)=EXPF(R(1)*Y(4))*(U*R(1)*X10+1.0-2.0*EXPF(-R(1)*Y(1))+2.0*EXPF
   1(-R(1)*Y(2))-2.0*EXPF(-R(1)*Y(3))+EXPF(-R(1)*Y(4)))
    G(2)=EXPF(R(2)*Y(4))*(U*R(2)*X20+1.0-2.0*EXPF(-R(2)*Y(1))+2.0*EXPF
   1(-R(2)*Y(2))-2.0*EXPF(-R(2)*Y(3))+EXPF(-R(2)*Y(4)))
    G(3)=EXPF(R(3)*Y(4))*(U*R(3)*X30+1.0-2.0*EXPF(-R(3)*Y(1))+2.0*EXPF
   1(-R(3)*Y(2))-2.0*EXPF(-R(3)*Y(3))+EXPF(-R(3)*Y(4)))
    G(4)=EXPF(R(4)*Y(4))*(U*R(4)*X40+1.0-2.0*EXPF(-R(4)*Y(1))+2.0*EXPF
   1(-R(4)*Y(2))-2.0*EXPF(-R(4)*Y(3))+EXPF(-R(4)*Y(4)))
    E1=G(1)*G(1)+G(2)*G(2)+G(3)*G(3)+G(4)*G(4)
    E=SQRTF(E1)
    PRINT 17 X10,X20,X30,X40,U,(Y(N),N=1,4),E
    GO TO 100
 42 PRINT 11
    GO TO 100
```

```
   43 PRINT 12
      GO TO 100
   44 PRINT 13
  100 CONTINUE
      STOP
      END
      SUBROUTINE INVERT (B,N,DET,J)
C     THIS SUBROUTINE INVERTS AN ARBITRARY MATRIX BY THE GAUSS
C     ELIMINATION METHOD
      DIMENSION B(4,4),A(4,8)
      IF DIVIDE CHECK 26,26
   26 IF ACCUMULATOR OVERFLOW 25,25
   25 DET=1.0
      N2=N+N
      L=N-1
      DO 10 I=1,N
      DO 10 J=1,N
   10 A(I,J)=B(I,J)
      K=N+1
      DO 11 I=1,N
      DO 12 J=K,N2
   12 A(I,J)=0.0
      M=N+I
   11 A(I,M)=1.0
      DO 6 J=1,L
      M=J+1
      IF(A(J,J)) 1,2,1
    2 DO 3 I=M,N
      IF(A(I,J)) 4,3,4
    3 CONTINUE
   24 DET=0.0
      IF ACCUMULATOR OVERFLOW 28,30
   30 IF DIVIDE CHECK 28,27
    4 DO 5 K=1,N2
    5 A(J,K)=A(J,K)+A(I,K)
    1 DO 6 I=M,N
      IF(A(I,J))8,6,8
    8 X=A(I,J)/A(J,J)
      DO 9 K=J,N2
    9 A(I,K)=A(I,K)-X*A(J,K)
    6 CONTINUE
      IF(A(N,N)) 23,24,23
   23 DO 13 KX=1,L
      M=N-KX
      J=M+1
      DO 13 IX=1,M
      I=M-IX+1
      IF(A(I,J)) 15,13,15
   15 X=A(I,J)/A(J,J)
      DO 14 K=I,N2
   14 A(I,K)=A(I,K)-X*A(J,K)
   13 CONTINUE
```

```
      DO 16 I=1,N
      X=A(I,I)
      DET=DET*X
      DO 16 J=1,N2
   16 A(I,J)=A(I,J)/X
      DO 18 I=1,N
      DO 18 J=1,N
      L=J+N
   18 B(I,J)=A(I,L)
      IF ACCUMULATOR OVERFLOW 28,29
   29 IF DIVIDE CHECK 28,27
   27 J=1
      RETURN
   28 J=-1
      RETURN
      END
      END
```

Program Inequa

```
      PROGRAM INEQUA
    5 FORMAT(1H1,47HN B HEMESATH  LAGRANGE MULTIPLIER SOLUTION WITH,
     123H INEQUALITY CONSTRAINTS)
   10 FORMAT(6(F10.2))
   12 FORMAT(6F12.6)
   14 FORMAT(I3)
   15 FORMAT(1H0,2F8.2,4E15.8,4E10.3)
   20 FORMAT(1H0,5X,2HU1,6X,2HU2,7X,2HT1,13X,2HT2,13X,2HT0,14X,1HT,12X,
     11HZ,10X,1HF,10X,1HE,10X,1HD)
   35 FORMAT(1H0,25HLOGF ARGUMENT IS NEGATIVE,5X,3HT0=E15.8)
      PRINT 5
      PRINT 20
      DIMENSION A(5,5)
      N=5
      D=1.0
      J=1
      READ 14 M1
      READ 12 P,Q,B11,B12,B21,B22
      DO 50 I=1,M1
      READ 10 X10,X20,T11,T21,T12,T22
      DO 50 K=1,2
      U1=(-1.0)**K
      DO 50 L=1,2
      U2=(-1.0)**L
      M=0
      S11=T11
```

```
      S21=T21
      S12=T12
      S22=T22
      U=1.0/(-B11*U1-B12*U2)
      B=B21*U1+B22*U2
      D1=P*X10+B11*U1*(1.0-2.0*EXPF(-P*S11*S11))
     1+B12*U2*(1.0-2.0*EXPF(-P*S21*S21))
      B1=4.0*U1*P*B11*EXPF(-P*S11*S11)
      B2=4.0*U2*P*B12*S21*EXPF(-P*S21*S21)
      B3=4.0*U1*Q*B21*S11*EXPF(-Q*S11*S11)
      B4=4.0*U2*Q*B22*S21*EXPF(-Q*S21*S21)
      IF(D1*U) 40,40,30
   30 H1=(-1.0/P)*(B1/D1)
      H2=(-1.0/P)*(B2/D1)
      G1=B3+(Q/P)*B*B1*U*((D1*U)**(Q/P-1.0))
      G2=B4+(Q/P)*B*B2*U*((D1*U)**(Q/P-1.0))
      Z=(-H1*G1-H2*G2)/(G1*G1+G2*G2)
   25 D1=P*X10+B11*U1*(1.0-2.0*EXPF(-P*S11*S11))
     1+B12*U2*(1.0-2.0*EXPF(-P*S21*S21))
      B1=4.0*U1*P*B11*EXPF(-P*S11*S11)
      B2=4.0*U2*P*B12*S21*EXPF(-P*S21*S21)
      B3=4.0*U1*Q*B21*S11*EXPF(-Q*S11*S11)
      B4=4.0*U2*Q*B22*S21*EXPF(-Q*S21*S21)
      IF(D1*U) 40,40,32
   32 H1=(-1.0/P)*(B1/D1)
      H2=(-1.0/P)*(B2/D1)
      G1=B3+(Q/P)*B*B1*U*((D1*U)**(Q/P-1.0))
      G2=B4+(Q/P)*B*B2*U*((D1*U)**(Q/P-1.0))
      H11=(-1.0/P)*((D1*(-2.0*P*S11*B1+B1/S11)-B1*B1)/(D1*D ))
      H12=(-1.0/P)*((-B1*B2)/(D1*D1))
      H22=(-1.0/P)*((D1*(-2.0*P*S21*B2+B2/S21)-B2*B2)/(D1*D1))
      H21=H12
      G11=-2.0*Q*S11*B3+B3/S11+(Q/P)*B*U*(((D1*U)**(Q/P-1.0))*(-2.0*P*
     1S11*B1+B1/S11)+(Q/P-1.0)*U*B1*B1*((D1*U)**(Q/P-2.0)))
      G12=(Q/P)*B*U*(B1*B2*U*(Q/P-1.0)*((D1*U)**(Q/P-2.0)))
      G21=G12
      G22=-2.0*Q*S21*B4+B4/S21+(Q/P)*B*U*(((D1*U)**(Q/P-1.0))*(-2.0*P*
     1S21*B2+B2/S21)+(Q/P-1.0)*U*B2*B2*((D1*U)**(Q/P-2.0)))
      A(1,1)=H11+Z*G11
      A(1,2)=0.0
      A(1,3)=H12+Z*G12
      A(1,4)=0.0
      A(1,5)=G1
      A(2,1)=H21+Z*G21
      A(2,2)=0.0
      A(2,3)=H22+Z*G22
      A(2,4)=0.0
      A(2,5)=G2
      A(3,1)=G1
      A(3,2)=0.0
      A(3,3)=G2
      A(3,4)=0.0
```

```
      A(3,5)=0.0
      A(4,1)=H1-2.0*S11
      A(4,2)=-2.0*S12
      A(4,3)=H2
      A(4,4)=0.0
      A(4,5)=0.0
      A(5,1)=H1
      A(5,2)=0.0
      A(5,3)=H2-2.0*S21
      A(5,4)=-2.0*S22
      A(5,5)=0.0
      F1=H1+Z*G1
      F2=H2+Z*G2
      F3=Q*X20+B21*U1*(1.0-2.0*EXPF(-Q*S11*S11))+B22*U2*(1.0-2.0*
     1EXPF(-Q*S21*S21))+B*((D1*U)**(Q/P))
      F4=(-1.0/P)*LOGF(D1*U)-S11*S11-S12*S12
      F5=(-1.0/P)*LOGF(D1*U)-S21*S21-S22*S22
      M=M+1
      CALL INVERT (A,N,D,J)
      S11=S11-A(1,1)*F1-A(1,2)*F2-A(1,3)*F3-A(1,4)*F4-A(1,5)*F5
      S12=S12-A(2,1)*F1-A(2,2)*F2-A(2,3)*F3-A(2,4)*F4-A(2,5)*F5
      S21=S21-A(3,1)*F1-A(3,2)*F2-A(3,3)*F3-A(3,4)*F4-A(3,5D*F5
      S22=S22-A(4,1)*F1-A(4,2)*F2-A(4,3)*F3-A(4,4)*F4-A(4,5D*F5
      Z=Z-A(5,1)*F1-A(5,2)*F2-A(5,3)*F3-A(5,4)*F4-A(5,5)*F5
      T1=S11*S11
      T2=S21*S21
      TO=T1+S12*S12
      T=T2+S22*S22
      IF(-Q*TO-40.0) 55,40,40
   55 IF(-Q*T-40.0) 60,40,40
   60 E1=EXPF(P*TO)*(X10-(B11*U1/P)*(2.0*EXPF(-P*T1)-EXPF(-P*TO)-1.0)
     1-(B12*U2/P)*(2.0*EXPF(-P*T2)-EXPF(-P*TO)-1.0))
      E2=EXPF(Q*TO)*(X20-(B21*U1/Q)*(2.0*EXPF(-Q*T1)-EXPF(-Q*TO)-1.0)
     1-(B22*U2/Q)*(2.0*EXPF(-Q*T2)-EXPF(-Q*TO)-1.0))
      E=SQRTF(E1*E1+E2*E2)
      F=F1*F1+F2*F2+F3*F3+F4*F4+F5*F5
      PRINT 15 U1,U2,T1,T2,TO,T,Z,F,E,D
      IF(M-30) 45,45,50
   45 IF(F-0.0001) 50,50,25
   40 PRINT 35
   50 CONTINUE
      STOP
      END
      SUBROUTINE INVERT (B,N,DET,J)
C     THIS SUBROUTINE INVERTS AN ARBITRARY MATRIX BY THE GAOSS
C     ELIMINATION METHOD
      DIMENSION B(5,5),A(5,10)
      IF DIVIDE CHECK 26,26
   26 IF ACCUMULATOR OVERFLOW 25,25
   25 DET=1.0
      N2=N+N
      L=N-1
```

```
      DO 10 I=1,N
      DO 10 J=1,N
   10 A(I,J)=B(I,J)
      K=N+1
      DO 11 I=1,N
      DO 12 J=K,N2
   12 A(I,J)=0.0
      M=N+I
   11 A(I,M)=1.0
      DO 6 J=1,L
      M=J+1
      IF(A(J,J)) 1,2,1
    2 DO 3 I=M,N
      IF(A(I,J)) 4,3,4
    3 CONTINUE
   24 DET=0.0
      IF ACCUMULATOR OVERFLOW 28,30
   30 IF DIVIDE CHECK 28,27
    4 DO 5 K=1,N2
    5 A(J,K)=A(J,K)+A(I,K)
    1 DO 6 I=M,N
      IF(A(I,J))8,6,8
    8 X=A(I,J)/A(J,J)
      DO 9 K=J,N2
    9 A(I,K)=A(I,K)-X*A(J,K)
    6 CONTINUE
      IF(A(N,N)) 23,24,23
   23 DO 13 KX=1,L
      M=N-KX
      J=M+1
      DO 13 IX=1,M
      I=M-IX+1
      IF(A(I,J)) 15,13,15
   15 X=A(I,J)/A(J,J)
      DO 14 K=I,N2
   14 A(I,K)=A(I,K)-X*A(J,K)
   13 CONTINUE
      DO 16 I=1,N
      X=A(I,I)
      DET=DET*X
      DO 16 J=1,N2
   16 A(I,J)=A(I,J)/X
      DO 18 I=1,N
      DO 18 J=1,N
      L=J+N
   18 B(I,J)=A(I,L)
      IF ACCUMULATOR OVERFLOW 28,29
   29 IF DIVIDE CHECK 28,27
   27 J=1
      RETURN
   28 J=-1
      RETURN
      END
      END
```

REFERENCES

1. Thaler, G. J., and Brown, R. G. Analysis and Design of Feedback Control Systems. New York: McGraw-Hill Book Company, Inc., 1960.

2. Graham, D., and McRuer, D. Analysis of Nonlinear Control Systems. New York: John Wiley and Sons, Inc., 1961.

3. Weiss, H. K. "Analysis of Relay Servomechanisms," Journal of the Aeronautical Sciences, Vol. 13, No. 7 (1946), pp. 364-376.

4. Bogner, I., and Kazda, L. F. "An Investigation of the Switching Criteria for Higher Order Contactor Servomechanisms," AIEE Transactions, Vol. 73, Pt. II (1954), pp. 118-127.

5. Hopkin, A. M. "A Phase Plane Approach to the Compensation of Saturating Servomechanisms," AIEE Transactions, Vol. 70, Pt. I (1951), pp. 631-639.

6. McDonald, D. Nonlinear Techniques for Improving Servo Performance. Proceedings of the National Electronics Conference, Chicago, Illinois, Vol. 6, 1950, pp. 400-421.

7. Bushaw, D. Optimal Discontinuous Forcing Terms. Contributions to the Theory of Nonlinear Oscillations, Vol. IV, Princeton University Press, 1958.

8. Bellman, R., Glicksberg, I., and Gross, O. "On the Bang-Bang Control Problem," Quarterly Journal of Applied Mathematics, Vol. 14 (1956), pp. 11-18.

9. LaSalle, J. P. Time Optimal Control Systems. Proceedings of the National Academy of Sciences, Vol. 45 (1959), pp. 573-577.

10. Desoer, C. A. "The Bang-Bang Servo Problem Treated by Variational Techniques," Information and Control, Vol. 2 (1959), pp. 333-348.

11. Pontryagin, L. S., Boltyanskii, V. G., Gamkrelidze, R. V., and Mishchenko, E. F. The Mathematical Theory of Optimal Processes. New York: Interscience Publishers, 1962.

12. Kalman, R. E. On the General Theory of Control Systems. Proceedings of the First International Congress of the International Federation of Automatic Control, Vol. 1 (1960), pp. 481-492.

13. ________. "Contributions to the Theory of Optimal Control," Boletin de la Matematica Mexicana, 1960.

14. Birkhoff, G., and MacLane, S. A Survey of Modern Algebra. New York: The Macmillan Company, 1953.

15. Hohn, F. E. Elementary Matrix Algebra. New York: The Macmillan Company, 1958.

16. Athanassiades, M., and Smith, O. J. M. "Theory and Design of High-Order Bang-Bang Control Systems," IRE Professional Group on Automatic Control Transactions, Vol. AC-6 (May 1961), pp. 125-134.

17. Lee, E. B. "Mathematical Aspects of the Synthesis of Linear Minimum Response-Time Controllers," IRE Professional Group on Automatic Control Transactions, Vol. AC-5 (September 1960), pp. 283-289.

18. Smith, F. B., Jr. "Time Optimal Control of Higher-Order Systems," IRE Professional Group on Automatic Control Transactions, Vol. AC-6 (February 1961), pp. 16-21.

19. Coddington, E., and Levinson, N. Theory of Ordinary Differential Equations. New York: McGraw-Hill Book Company, 1955.

20. Taylor, A. E. Advanced Calculus. Boston: Ginn and Company, 1955.

21. Olmsted, J. M. H. Advanced Calculus. New York: Appleton-Century-Crofts, Inc., 1961.

22. Dorn, W. S. "On Lagrange Multipliers and Inequalities," Journal of the Operations Research Society of America, Vol. 9 (1961), pp. 91-104.

23. John, F. Extremum Problems With Inequalities as Subsidiary Conditions. Studies and Essays, Courant Anniversary Volume. New York: Interscience Publishers, 1948, pp. 187-204.

24. Tucker, A. W. "Linear and Non-Linear Programming," Journal of the Operations Research Society of America, Vol. 5 (1957), pp. 244-257.

25. Courant, R., and Hilbert, D. Methods of Mathematical Physics. New York: Interscience Publishers, 1953.

26. Wirth, J. L. "Time Domain Models of Physical Systems and Existence of Solutions." Unpublished thesis, Michigan State University, East Lansing, Michigan, 1962.

27. Bellman, R. Stability Theory of Differential Equations. New York: McGraw-Hill Book Company, Inc., 1953.

28. Struble, R. A. Nonlinear Differential Equations. New York: McGraw-Hill Book Company, Inc., 1962.

29. Kantorovich, L. V. Approximate Methods of Higher Analysis. New York: Interscience Publishers, 1958.

30. Henrici, P. Discrete Variable Methods in Ordinary Differential Equations. New York: John Wiley and Sons, Inc., 1962.

31. Zaguskin, V. L. Handbook of Numerical Methods for the Solution of Algebraic and Transcendental Equations. London: Pergammon Press, 1961.

32. Hildebrand, F. B. Introduction to Numerical Analysis. New York: McGraw Hill Book Company, Inc., 1956.

33. Booth, A. D. Numerical Methods. London: Butterworths Scientific Publications, 1957.

34. Householder, A. S. Principles of Numerical Analysis. New York: McGraw-Hill Book Company, Inc., 1953.