NASA CONTRACTOR REPORT

NASA CR-71

c. 1

NASA CR-715

STUDY OF OPTIMAL AND ADAPTIVE CONTROL THEORY

by C. D. Johnson

Prepared by
UNIVERSITY OF ALABAMA RESEARCH INSTITUTE
Huntsville, Ala.
for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION • WASHINGTON, D. C. • APRIL 1967

NASA CR-715

STUDY OF OPTIMAL AND ADAPTIVE CONTROL THEORY

By C. D. Johnson

Prepared under Contract No. NAS 8-11231 by
UNIVERSITY OF ALABAMA RESEARCH INSTITUTE
Huntsville, Ala.

for George C. Marshall Space Flight Center

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION

For sale by the Clearinghouse for Federal Scientific and Technical Information
Springfield, Virginia 22151 – CFSTI price $3.00

TABLE OF CONTENTS

I. Introduction

This report is the final report on National Aeronautics and Space Administration Contract No. NAS8-11231 (including Modifications No. 1 and No. 2) entitled "Study of Optimal and Adaptive Control Theory".

During the study period, which began in May 1964 and continued through June 1966, the principal investigator conducted investigations into a variety of theoretical questions which arise naturally in the study of modern optimal and adaptive control techniques for large launch vehicles. The particular topics chosen for investigation were selected through consultations with the staff of the Aero-Astrodynamics Laboratory of the George C. Marshall Space Flight Center, Huntsville Alabama. The principal investigator is especially grateful to Mr. Clyde Baker, Mr. Judson Lovingood, Dr. David Ford* and Mr. Tommy Carter, of the Aero-Astrodynamics Laboratory, for their many stimulating and informative discussions during this study.

The material in this report is arranged in chapters, with each chapter representing a self-contained exposition of a particular topic. Reference and figure citations in the individual chapters refer only to the list of references and collection of figures given in that particular chapter.

The subject of optimal control with a "Minimax-type" performance index received particular attention during this study because of its' potential application in the design of load-minimizing control systems for large launch vehicles. Some of the methods which have previously been used to solve such problems are summarized in Chapter II. In addition, Chapter II contains a detailed account of an essentially "new" method for effectively solving such problems. Application of this proposed new method is illustrated by five examples which are worked in detail.

The study of "worst-case" optimal control problems with unknown disturbances leads naturally to the study of variational problems with "competing controls". This class of problems can be cast as continuous games in function spaces and was first studied as such by R. Isaacs in the early 1950's. A general study of this class of problems was initiated by the principal investigator in the early period of the contract. However, the idea of a general study was subsequently abandoned with the publication of Isaac's

---

* Now with the Department of Mathematics, Emory University.

highly original, and now well-known, 1965 treatise on this subject. Instead, several specific examples were studied in detail by the principal investigator and those examples are presented in Chapter III.

The so-called "phase-variable canonical form" for single-input linear dynamical systems has found many applications in the area of modern linear control theory. Two of the most interesting theoretical aspects of this topic are the matrix theoretic structure of and computational algorithms for the required transformation matrix. The results of studies of these two topics are described in Chapters IV and V.

In many practical applications of modern control theory the concept of invariant hyperplanes in the system state space plays an important role. For instance, this concept forms the theoretical foundation for N.A.S.A.'s "Drift-Minimum" control principle. In the course of the present study the principal investigator studied the algebraic theory of invariant hyperplanes for linear dynamical systems and, by this means, was able to show connections between that topic and the important and related subjects of controllability and observability. These results are described in Chapter VI.

The application of optimal control techniques in the design of large launch vehicle control systems has centered around the well-known, and almost completely solved, linear optimal regulator problem. During the present study, the principal investigator considered several variations on the usual formulation of the linear optimal regulator problem. In particular, the fixed-time regulator problem with time-invariant feedback control and the unspecified-time regulator problem with bounded control were studied in detail. The results obtained for these two problems are described in Chapters VII and VIII.

II. Optimal Control with Chebyshev Minimax Performance Index[1]

C. D. Johnson

Summary

The optimal control of dynamical systems with conventional Mayer, Lagrange, and Bolza type performance indices has been studied in some detail [1],[2] [2], [3]. In the present work the optimal control of dynamical systems with a certain minimax type performance index, which cannot be expressed in the Mayer, Lagrange, or Bolza format, is studied. The form of the optimal control is described and certain geometric properties of the solution are discussed. Several examples are worked in detail to illustrate application of the proposed method of solution.

1. Introduction

Consider the class of dynamical systems described by

$$\dot{\underset{\sim}{x}} = \underset{\sim}{F}(\underset{\sim}{x}, t, u(t)) \qquad \cdot = d/dt$$

where $\underset{\sim}{x}$ is the system state vector and $u(t)$ is the scalar input or control. The optimal control of this class of dynamical systems has been studied in detail for the three particular cases in which $J[u]$, the performance index functional to be minimized, is:

(i) a scalar function G of the initial and/or terminal state (the Mayer type problem)

$$J[u] = G(\underset{\sim}{x}(t_o), \underset{\sim}{x}(T), t_o, T),$$

(ii) a time integral of a scalar function L evaluated along the state trajectory between the initial and terminal states (the Lagrange type problem)

$$J[u] = \int_{t_o}^{T} L(\underset{\sim}{x}(t), u(t), t)\, dt,$$

---

1. This research was conducted at the University of Alabama Research Institute and was supported by the National Aeronautics and Space Administration under Contract NAS8-11231 and Grant No. NsG-381. This paper was presented at the 1966 Joint Automatic Control Conference, Seattle, Washington, August 1966 and will appear in the A.S.M.E. Transactions, Journal of Basic Engineering, March 1967.

2. Numbers in brackets designate References at end of paper.

(iii) a sum of (i) and (ii) (the Bolza type problem)

$$J[u] = G(\underset{\sim}{x}(t_o), \underset{\sim}{x}(T), t_o, T) + \int_{t_o}^{T} L(\underset{\sim}{x}(t), u(t), t)\, dt.$$

The theoretical work of Pontryagin, Bellman, LaSalle, Kalman, Berkovitz, and others has led to the development of a relatively complete mathematical theory for this particular class of optimal control problems.

In many practical applications, however, the actual physical performance criterion cannot be expressed as a Mayer, Lagrange, or Bolza type performance index. For example, in the case of regulator type control systems, the actual performance index may be expressed as per-cent overshoot due to a step change in load with a specified upper bound on settling time. In other practical applications the performance index may be expressed as the maximum or peak value of velocity, acceleration, force, torque, stress, temperature, etc., which occurs during some specified interval of control. Performance indices of this type fall into the general category of what we will call Chebyshev minimax performance indices in which the objective is to minimize the maximum value attained by a scalar function $C(\underset{\sim}{x}(t))$, evaluated along the state trajectory, over some specified closed interval of time.[3]

Optimal control problems with minimax performance indices of the type described above belong to a broad class of extremal problems which have their origin in the highly original researches of the Russian mathematician P. L. Chebyshev [Tchebycheff]. In his 1854 studies [15] of mechanical linkages which generate approximate straight line motion, Chebyshev introduced the important idea of characterizing the quality of the approximation in terms of the maximum deviation from the desired straight line. Those investigations led to the formulation of more general mathematical problems of minimax approximation involving functions least deviating from zero and ultimately led to the development of the well-known Chebyshev

---

3. It is remarked that the term "minimax" has also been used (in a basically different sense) to describe a variety of conflict-type optimization problems which arise in game theory [4], statistical communication theory [5], [6], and optimal control theory [7]-[14]. In those problems, the performance index is usually of the conventional Mayer, Lagrange, or Bolza type and the minimum and maximum operations are taken with respect to the policies or control actions of two conflicting elements.

polynomials [16] and the more general theory of Chebyshev approximation [17], [18]. The extension of Chebyshev's minimax approximation ideas to problems in approximating solutions to ordinary differential equations was considered in a 1907 paper by Young [19] and more recently by Lanczos [20] and Carter [21].

In 1956, Bellman, Glicksburg, and Gross, [22] gave one of the first accounts of the application of Chebyshev minimax performance indices to problems of optimal control of dynamical systems. In subsequent investigations, Bellman [23], [24], [25], Sevin [26], and Bellman, Glicksburg and Gross [27] studied a variety of particular examples from this class of problems. Some of the more recent researches in this class of problems are described in [28]-[35].

In this paper, we formulate a particular class of optimal control problems with Chebyshev minimax [C-minimax] performance index and describe a method of solution which is essentially different from those proposed in [22]-[35]. Some geometric properties of the optimal trajectory are discussed and a practical technique for computing the optimal control is described. Several examples are worked in detail to illustrate application of the theory.

2. Statement of the Problem

The problem is to find, in the class of piecewise continuous functions, a scalar control $u = u(t)$ which minimizes the functional

$$J[u] = \max_{t_o \le t \le T} C(\underset{\sim}{x}(t)) \tag{1}$$

subject to the following conditions[4]

$$\dot{\underset{\sim}{x}} = \underset{\sim}{F}(\underset{\sim}{x}, u(t)) \qquad (\cdot = d/dt) \tag{2}$$

$$\underset{\sim}{x}(t_o) = \underset{\sim}{x}_o \qquad \underset{\sim}{x}_o \in D \tag{3}$$

---

4. The case in which the independent variable t (T) appears as an explicit argument in one or more of the functions C, F, ($\mathcal{T}$) can be case into the form of (1)-(5) by introducing an auxiliary state variable x_{n+1} defined by

$$\dot{x}_{n+1} = 1$$

$$x_{n+1}(t_o) = t_o$$

$$\mathcal{J}(\underset{\sim}{x}(T)) = 0 \qquad T \geq t_o \tag{4}$$

$$u(t) \in U \qquad t_o \leq t \leq T \tag{5}$$

In (1), $\underset{\sim}{x} = (x_1, \ldots, x_n)$ is an n-vector: the system state vector, and $C(\underset{\sim}{x})$ is the performance index: a real, single valued, scalar function of $\underset{\sim}{x}$ defined throughout a set D of the n-dimensional euclidean state space E^n. In (2), $\underset{\sim}{F}$ is a vector function continuous in u and continuously differentiable with respect to $\underset{\sim}{x} \in D$. Equation (4) defines the terminal manifold, $\mathcal{J} \subset D$, an m-dimensional ($m<n$) hypersurface of admissible terminal states $\underset{\sim}{x}(T)$. The terminal time T is specified implicitly, by (4), as the first time $t \geq t_o$ which satisfies $\mathcal{J}(\underset{\sim}{x}(t)) = 0$. Problems involving more explicit restrictions on T [for example, such as $T = T^*$ or $T \leq T^*$ where T^* is some specified constant] can be accommodated in this formulation by the technique described in footnote 4 below. It is assumed that $C(\underset{\sim}{x})$ and $\mathcal{J}(\underset{\sim}{x})$ are once continuously differentiable and $\mathcal{J}$ is connected.

A piecewise continuous real valued function u(t) with values belonging to the closed, convex, and bounded set U is called an admissible control. An admissible control $u = u^o(t)$ which yields an absolute minimum of the functional (1), subject to the restrictions (2)-(4), is called optimal.[5] An optimal control of the form $u^o(t) = u^o(\underset{\sim}{x}(t))$ is an optimal control law. An integral curve of (2) corresponding to an optimal control, is an optimal trajectory. The set $D \subset E^n$ is taken as the set of all states $\underset{\sim}{x}$ which are controllable to $\mathcal{J}$. That is, for each initial state $\underset{\sim}{x}_o \in D$ there exists at least one admissible control u(t) such that the corresponding solution of (2) satisfies (3) and (4). Hereafter, we assume that D is non-void and $u^o(\underset{\sim}{x})$ exists for all $\underset{\sim}{x} \in D$.

3. Form of the Solution

Let $u^o(\underset{\sim}{x})$ be an optimal control law, and let

$$J[u^o(\underset{\sim}{x}); \underset{\sim}{x}_o] = V(\underset{\sim}{x}_o), \qquad \underset{\sim}{x}_o \in D \tag{6}$$

---

5. It may be noted that an optimal control for the functional (1) is also optimal for every functional of the form $J[u] = \max_{t_o \leq t \leq T} M[C(\underset{\sim}{x}(t))]$ where M(C) is any monotonically increasing continuous function of C. For this reason, the previously stated assumption concerning the differentiability of the performance index can usually be realized, even when the original function $C(\underset{\sim}{x})$ is not continuously differentiable, by proper choice of an alternative performance index $M[C(\underset{\sim}{x})]$.

From (1), it is clear that

$$V(\underset{\sim}{x}) \geq C(\underset{\sim}{x}) \qquad \forall\ \underset{\sim}{x} = \underset{\sim}{x}_o \in D \tag{7}$$

Thus, any admissible control $u(t)$ is optimal if the corresponding solution of (2) satisfies (3), (4) and the condition

$$C(\underset{\sim}{x}(t)) \leq C(\underset{\sim}{x}_o) \qquad t_o \leq t \leq T \tag{8}$$

On the terminal manifold (4)

$$V(\underset{\sim}{x}) = C(\underset{\sim}{x}) \qquad \forall\ \underset{\sim}{x} \in \mathcal{T} \tag{9}$$

It is assumed hereafter that $V(\underset{\sim}{x})$ is <u>continuous</u> at each state $\underset{\sim}{x}$ in the interior of D.

Let $R_o \supset \mathcal{T}$ be the set of all states $\underset{\sim}{x} \in D$ with the following properties. For each $\underset{\sim}{x}_o \in R_o$ an admissible control $u = \phi(t; \underset{\sim}{x}_o)$ exists such that (8) is satisfied everywhere along the corresponding solution of (2) and, in addition

$$\mathcal{J}(\underset{\sim}{x}(T)) = 0 \qquad \text{for some } T \geq t_o \tag{10}$$

$$\underset{\sim}{x}(t) \in R_o \qquad \forall\ t_o \leq t \leq T \tag{11}$$

Clearly, the set R_o is connected and closed relative to D. It is remarked that $\phi(t; \underset{\sim}{x}_o)$ is not unique, in general. Moreover, the set $R_o - \mathcal{T}$ might be empty. Let ∂R_o denote the boundary of the set R_o and suppose that ∂R_o is defined by $B(\underset{\sim}{x}) = 0$, $\underset{\sim}{x} \in D$. Suppose also that $\underset{\sim}{\nabla} B(\underset{\sim}{x})$ exists at $\underset{\sim}{x}$ and let $\underset{\sim}{\nu}(\underset{\sim}{x})$ be the <u>outward</u> pointing normal to the boundary ∂R_o. Thus, $\underset{\sim}{\nu}(\underset{\sim}{x}) = \pm \underset{\sim}{\nabla} B(\underset{\sim}{x})$ depending on the choice of $B(\underset{\sim}{x})$. If $U'(\underset{\sim}{x}) \subset U$ is the particular <u>open</u> set of admissible values of the control u defined by[6]

$$U'(\underset{\sim}{x}) = \{u \mid u \in U;\ \langle \underset{\sim}{\nu}(\underset{\sim}{x}),\ \underset{\sim}{F}(\underset{\sim}{x}, u) \rangle < 0\}, \qquad \underset{\sim}{x} \in \partial R_o \tag{12}$$

then it follows from the definition of R_o that states $\underset{\sim}{x} \in \partial R_o \subset D$, which do not lie

---

6. $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle$ denotes the inner product of $\underset{\sim}{x}$ and $\underset{\sim}{y}$.

on the terminal manifold (4), have the following properties wherever $\underset{\sim}{\nu}(\underset{\sim}{x})$ exists:[7]

(i) If the set $U'(\underset{\sim}{x})$ is not empty then

$$\inf_{u \in U'(\underset{\sim}{x})} \langle \nabla C(\underset{\sim}{x}), \underset{\sim}{F}(\underset{\sim}{x}, u) \rangle = 0 \qquad \underset{\sim}{x} \in \partial R_o \tag{13}$$

(ii) If the set $U'(\underset{\sim}{x})$ is empty then

$$\min_{u \in U} \langle \underset{\sim}{\nu}(\underset{\sim}{x}), \underset{\sim}{F}(\underset{\sim}{x}, u) \rangle = 0 \qquad \underset{\sim}{x} \in \partial R_o \tag{14}$$

Moreover, for each control $u^* = u^*(\underset{\sim}{x})$ which satisfies (14)

$$\langle \nabla C(\underset{\sim}{x}), \underset{\sim}{F}(\underset{\sim}{x}, u^*(\underset{\sim}{x})) \rangle \leq 0 \qquad \underset{\sim}{x} \in \partial R_o \tag{15}$$

It also follows, from the definition of R_o, that the regions of ∂R_o at which (14) and (15) are satisfied (i.e.: $U'(\underset{\sim}{x})$ is empty) are built up from integral manifolds of optimal trajectories which belong to R_o.

It is clear that $u = \phi(t; \underset{\sim}{x}_o)$ is an optimal control when $\underset{\sim}{x}_o \in R_o$. Moreover, an optimal control law $u^o(\underset{\sim}{x})$, $\underset{\sim}{x} \in R_o$, satisfies,[8] in addition to (8), (10), (11), the condition

$$\langle \nabla C(\underset{\sim}{x}), \underset{\sim}{F}(\underset{\sim}{x}, u^o(\underset{\sim}{x})) \rangle \leq 0 \tag{16}$$

for every state $\underset{\sim}{x}$ interior to R_o and satisfies the appropriate condition (13) or (14), (15) whenever $\underset{\sim}{x} \in \partial R_o$ ($x \notin \mathcal{T}$). From (6) it follows that

$$V(\underset{\sim}{x}) = C(\underset{\sim}{x}) \qquad \forall\, \underset{\sim}{x} \in R_o \tag{17}$$

---

7. Equations (13)-(15) remain valid at points on ∂R_o where $B(\underset{\sim}{x})$ is not continuously differentiable provided that the set $U'(\underset{\sim}{x})$ is interpreted as the set of values $u \in U$ which "point" the local velocity vector $\underset{\sim}{F}(\underset{\sim}{x}, u)$, $\underset{\sim}{x} \in \partial R_o$, into the interior of the set R_o.

8. It should be stressed that $C(\underset{\sim}{x}(t))$ need not be monotonic non-increasing along every optimal trajectory which passes through a given initial state $\underset{\sim}{x}_o \in R_o$. However, it is evident from (8) and the definition of R_o that through each initial state $\underset{\sim}{x}_o \in R_o$ there passes at least one optimal trajectory along which $C(\underset{\sim}{x}(t))$ is monotonic non-increasing. In R_o, optimal trajectories of this latter type are the only ones which possess the Markovian property required for application of the Principle of Optimality [36]. It follows that an optimal control law can be defined in R_o only for optimal trajectories of this latter type.

It is remarked that the set R_o has a convenient interpretation in terms of Liapunov stability theory. In particular, if the performance index $C(\underset{\sim}{x})$ is considered as a generalized Liapunov function for the system $\dot{\underset{\sim}{x}} = \underset{\sim}{F}(\underset{\sim}{x}, u^o(\underset{\sim}{x}))$, in the sense of LaSalle [37], then the interior of R_o is the corresponding estimate of the domain D of asymptotic stability with respect to the terminal manifold $\mathcal{T}$.

Let $R_m \subset (D - R_o)$ denote the largest (not necessarily connected) set of states $\underset{\sim}{x}$ with the following properties. For each $\underset{\sim}{x}_o \in R_m$, an admissible control $u = \gamma(t; \underset{\sim}{x}_o)$ and a time t_1 exist such that the following conditions are satisfied along the corresponding solution of (2).

$$\text{(i)} \quad \underset{\sim}{x}(t) \in R_m \quad , \quad \forall\, t_o \le t < t_1 \tag{18}$$

$$\text{(ii)} \quad \underset{\sim}{x}(t_1) \in \partial R_o \tag{19}$$

$$\text{(iii)} \quad C(\underset{\sim}{x}(t)) \le C(\underset{\sim}{x}(t_1)), \qquad \forall\, t_o \le t \le t_1 \tag{20}$$

and

$$\text{(iv)} \quad \langle \underset{\sim}{p}(t), \underset{\sim}{F}(\underset{\sim}{x}(t), \gamma(t)) \rangle = \max_{u \in U} \langle \underset{\sim}{p}(t), \underset{\sim}{F}(\underset{\sim}{x}(t), u(t)) \rangle \equiv 0, \quad t_o \le t \le t_1 \tag{21}$$

where $\underset{\sim}{p}(t) = (p_1(t), \ldots, p_n(t))$ is a real, continuous n-vector which satisfies the differential equations

$$\dot{p}_i = -\sum_{j=1}^{n} p_j \frac{\partial F_j(\underset{\sim}{x}, \gamma)}{\partial x_i} \quad , \qquad \begin{matrix} (\cdot = d/dt) \\ i = 1, \ldots, n \end{matrix} \tag{22}$$

and the boundary conditions

$$\underset{\sim}{p}(t_1) + \nabla \underset{\sim}{C}(\underset{\sim}{x}(t_1)) = \begin{cases} 0, & \text{if } \underset{\sim}{x}(t_1) \in (\partial R_o - \mathcal{T}) \\ \text{normal to } \mathcal{T}, & \text{if } \underset{\sim}{x}(t_1) \in \mathcal{T} \end{cases} \tag{23}$$

In other words, R_m is the largest set of initial states $\underset{\sim}{x}_o \in (D - R_o)$ for which the condition (20) is satisfied naturally along solutions of the following, Mayer type, variational

problem: Find an admissible control $u(t)$, $t_o \le t \le t_1$, which minimizes $C(\underset{\sim}{x}(t_1))$ subject to the restrictions[9]

$$\begin{aligned} &\dot{\underset{\sim}{x}} = \underset{\sim}{F}(\underset{\sim}{x}, u(t)) \\ &\underset{\sim}{x}(t) \in R_m && \forall\ t_o \le t < t_1 \\ &\underset{\sim}{x}(t_1) \in \partial R_o && (t_1 \text{ is unrestricted}) \end{aligned} \tag{24}$$

For this Mayer type variational problem, let $\underset{\sim}{x}^*(\underset{\sim}{x}_o) = \underset{\sim}{x}^*(t_1(\underset{\sim}{x}_o); \underset{\sim}{x}_o)$ denote a minimizing "terminal" state corresponding to an initial condition $\underset{\sim}{x}_o \in R_m$. Then it follows, from (7), (20) and the minimizing property of $\underset{\sim}{x}^*(\underset{\sim}{x}_o)$, that for each initial state $\underset{\sim}{x}_o \in R_m$ an optimal control for the original problem (1)-(5) is given by

$$u^o(t) = \begin{cases} \gamma(t; \underset{\sim}{x}_o) & t_o \le t \le t_1 \\ \phi(t; \underset{\sim}{x}^*(\underset{\sim}{x}_o)) & t_1 < t \le T \end{cases} \tag{25}$$

The function (6), in this case, is therefore given by

$$V(\underset{\sim}{x}_o) = V(\underset{\sim}{x}^*(\underset{\sim}{x}_o)) = C(\underset{\sim}{x}^*(\underset{\sim}{x}_o)), \qquad \begin{matrix} \underset{\sim}{x}_o \in R_m \\ \underset{\sim}{x}^*(\underset{\sim}{x}_o) \in \partial R_o \end{matrix} \tag{26}$$

It is evident from (26) that optimal trajectories in the set R_m lie on hypersurfaces of constant $V(\underset{\sim}{x})$. Moreover, since $C(\underset{\sim}{x})$ is assumed continuous, the function $V(\underset{\sim}{x})$ defined by (17) and (26) must be continuous[10] at points $\underset{\sim}{x}$ where optimal trajectories from R_m cross over common boundaries of R_o and R_m. Thus, the equation defining the locus of such boundary points can be obtained by equating the two expressions (17) and (26).

---

9. Equations (21) and (22) are, respectively, the Hamiltonian function and the canonical equations which arise from application of Pontryagin's maximum principle [1] to the Mayer type variational problem (24). Equation (23) is the corresponding transversality condition for $\underset{\sim}{p}(t_1)$ and is a consequence of combining the usual transversality condition [i.e. $\underset{\sim}{p}(t_1) + \underset{\sim}{\nabla} C(\underset{\sim}{x}(t_1))$ should be normal to the terminal manifold] with the natural boundary condition (13)

10. This continuity property of $V(\underset{\sim}{x})$, together with (13) and (17), shows that $dV(\underset{\sim}{x}(t))/dt$ along optimal trajectories in R_m is also zero at points $\underset{\sim}{x} \in \partial R_o$ where optimal trajectories from R_m cross over ∂R_o.

It is remarked that, in certain instances, the condition (21) may fail to yield a well-defined control $\gamma(t) = \gamma(\underset{\sim}{x}(t), \underset{\sim}{p}(t))$ during some positive interval of time. In such cases, the possibility of singular solutions [38], [39] of the variational problem (24) should be investigated.

From the properties of the sets R_o, R_m *it follows that an optimal trajectory* $\underset{\sim}{x}(t)$ which passes through an initial state $\underset{\sim}{x}_o \notin R_o \cup R_m$ must necessarily[11] enter the set $R_o \cup R_m$ at some time t_2 $(t_o < t_2 < T)$. However, on the boundary of the set $R_o \cup R_m$, (6) is known and is given by

$$V(\underset{\sim}{x}) = \begin{cases} C(\underset{\sim}{x}), & \underset{\sim}{x} \in \partial R_o \\ \\ C(\underset{\sim}{x}^*(\underset{\sim}{x})), & \underset{\sim}{x} \in \partial R_m \end{cases} \tag{27}$$

where ∂R_m denotes the boundary of the set R_m. Thus, the boundary of the set $R_o \cup R_m$ can be treated as a new terminal manifold $\mathcal{J}^2$. In this way, the process described above for constructing the sets R_o, R_m can be repeated for the new terminal manifold $\mathcal{J}^2$ and sets R_o^2 and R_m^2 (analogous to R_o, R_m) can be constructed. Continuing with this process the region D may be completely partitioned into the two families of sets $\{R_o^i\} = \{R_o, R_o^2, R_o^3, \ldots\}$; and $\{R_m^i\} = \{R_m, R_m^2, R_m^3, \ldots\}$. When the partitioning of D into the sets $\{R_o^i\}, \{R_m^i\}$ is completed, the optimal control for the original problem (1)-(5) is completed, the optimal control for the original problem (1)-(5) is known. Suppose, for example, that the initial state $\underset{\sim}{x}_o$ belongs to a set $R_o^k \subset \{R_o^i\}$, $k \geq 2$. The optimal control, during the time interval, $t_o \leq t \leq t_k$, when $\underset{\sim}{x}(t) \in R_o^k$, can be chosen as any admissible control for which $C(\underset{\sim}{x}(t)) \leq C(\underset{\sim}{x}_o)$ and $\underset{\sim}{x}(t_k) \in \partial R_m^i$ for some $R_m^i \subset \{R_m^i\}$. The existence of at least one such control follows from the definition of the R_o type sets. Upon entering the neighboring set R_m^i, the continuation of the optimal control is determined by solving the appropriate, Mayer type, variational

---

11. It has been tacitly assumed that the two sets R_o, R_m can, in fact, be constructed in the manner described. This constructive procedure will fail, for example, if there exists some neighborhood $N \supset \mathcal{J}$ such that $C(\underset{\sim}{x}) > \max_{\underset{\sim}{x} \in \mathcal{J}} C(\underset{\sim}{x})$ $\forall$ $\underset{\sim}{x} \in (N - \mathcal{J})$ and $dC(\underset{\sim}{x}(t))/dt$ is sign indefinite along every admissible trajectory $\underset{\sim}{x}(t) \in N$ which satisfies (4). Such cases are degenerate from the point of view of the present theory. This degeneracy can usually be removed, however, by properly re-defining the terminal manifold (4).

problem (24) where the "terminal manifold" is taken as the boundaries of the immediately adjoining sets of the R_o type. In this way, the state $\underset{\sim}{x}(t)$ progresses alternately[12] and optimally through the sets of the R_o and R_m type and eventually reaches the original terminal manifold .

4. Minimax Points

The function $V(\underset{\sim}{x})$, defined in (6), associates a characteristic number with each initial state $\underset{\sim}{x}_o = \underset{\sim}{x} \in D$. This number represents the maximum value of the scalar function $C(\underset{\sim}{x}(t))$ which occurs along the particular optimal trajectory which starts at $\underset{\sim}{x}_o$. It some applications, it may be desirable to identify the actual state (or states) $\underset{\sim}{x} = \underset{\sim}{\xi}$, along the optimal trajectory which starts at $\underset{\sim}{x}_o$, at which the maximum of $C(\underset{\sim}{x}(t))$ occurs. We shall call a charateristic state $\underset{\sim}{\xi} = \underset{\sim}{\xi}(\underset{\sim}{x}_o)$ a minimax point for the state $\underset{\sim}{x}_o$. Every state $\underset{\sim}{x}_o \in D$ has an associated minimax point. However, the minimax point $\underset{\sim}{\xi}(\underset{\sim}{x}_o)$ associated with a given state $\underset{\sim}{x}_o$ is not unique, in general, owing to the presence of the equality signs in (8) and (20).

From the definition of the sets $\{R_o^i\}$ it follows that

$$\underset{\sim}{\xi}(\underset{\sim}{x}_o) = \underset{\sim}{x}_o \qquad \forall\, \underset{\sim}{x}_o \in \{R_o^i\} \tag{28}$$

defines at least one of the minimax points associated with each state $\underset{\sim}{x}_o \in \{R_o^i\}$. Moreover, for each initial state $\underset{\sim}{x}_o \in \{R_m^i\}$, at least one of the associated minimax points is the state[13] $\underset{\sim}{x} \in \partial\{R_o^i\}$ at which the optimal trajectory, starting at $x_o \in \{R_m^i\}$, first enters one of the sets $\{R_o^i\}$.

5. A Constructive Procedure for Identifying the Sets $\{R_o^i\}$, $\{R_m^i\}$

The set R_o can be identified numerically by means of a backward-time flooding technique provided that, for every $\underset{\sim}{x}_o \in R_o$, there exists at least one optimal control

---

12. Some of the sets $\{R_o^i\}$ might share common boundaries of the type described by (14)-(15). In such cases, there may exist (non-unique) optimal trajectories which do not progress alternately through the sets $\{R_o^i\}$, $\{R_m^i\}$. Some examples of this type are illustrated in Section 8 below.

13. Here, $\partial\{R_o^i\}$ denotes the boundary of the union of the sets $\{R_o^i\}$.

law such that $T < \infty$ for the corresponding optimal trajectory.[14] For this purpose, we set $\tau = T - t$ $(\tau \geq 0)$ in (2) and consider the reverse-time solutions $\underset{\sim}{x}(\tau)$ of (2) corresponding to various choices of admissible control functions $u(\tau)$. At each state $\underset{\sim}{x} \in R_o$ the condition

$$\min_{u \in U} \langle \nabla C(\underset{\sim}{x}), \underset{\sim}{F}(\underset{\sim}{x}, u) \rangle \leq 0 \qquad \underset{\sim}{x} \in R_o \tag{29}$$

is satisfied. Thus, in backward time one can always find, at each state $\underset{\sim}{x} \in R_o$, at least one admissible control value which yields

$$\frac{dC(\underset{\sim}{x}(\tau))}{d\tau} \geq 0 \qquad \underset{\sim}{x} \in R_o \tag{30}$$

provided the set $(R_o - \mathcal{T})$ is not empty.

Consider a particular (admissible) reverse-time solution $\underset{\sim}{x}(\tau)$ with initial condition satisfying $\underset{\sim}{x}(\tau=0) \in \mathcal{T}$ and along which the condition (30) is always satisfied. It is clear that each state $\underset{\sim}{x} \in D$ which can be "reached" by such a solution is contained in the set R_o. Moreover, it follows from the definition of R_o that each state $\underset{\sim}{x} \in R_o$ must be reachable by at least one such solution. Thus, the set R_o is the set of all states $\underset{\sim}{x}$ which can be reached by admissible reverse-time solutions of (2) with initial conditions satisfying $\underset{\sim}{x}(\tau=0) \in \mathcal{T}$ and along which the condition (30) is always satisfied. It is recalled that the appropriate condition (13) or (14), (15) is satisfied at non-terminal states $\underset{\sim}{x}$ on the boundary of R_o.

When the boundary of R_o is known, the backward-time technique can be used to establish optimal trajectories in R_m by integrating equations (2) and (22), in reverse time, starting on the boundary[15] of R_o. In this case, the "initial conditions" for (22) are given by (23) and the optimal control $\gamma(\tau) = \gamma(\underset{\sim}{x}(\tau), \underset{\sim}{p}(\tau))$ is determined from (21), provided the solution does not contain singular sub-arcs. It is clear from the definition of R_o that as (2) and (22) are integrated through R_m in reverse time, the value of

---

14. Example 3, in Section 8 below, illustrates a case in which this condition is not satisfied at each $\underset{\sim}{x}_o \in R_o$.

15. It is remarked that, in general, not all states $\underset{\sim}{x} \in \mathcal{T}$ are necessarily terminal states for an optimal trajectory originating outside $\mathcal{T}$. In particular, if ∂R_o contains points $\underset{\sim}{x} \in \mathcal{T}$ those points may, or may not, serve as terminal states $\underset{\sim}{x}(T)$ for optimal trajectories originating in R_m.

$C(\underset{\sim}{x}(\tau))$ must first decrease below its initial value. The reverse-time integration through R_m is continued until a point $\underset{\sim}{x}$ is reached where any further continuation of the reverse-time integration will result in the value of $C(\underset{\sim}{x}(\tau))$ exceeding its initial value. Each point $\underset{\sim}{x}$ determined in this manner is a boundary point[16] of R_m.

When the boundaries of R_o and R_m are known, the backward-time procedure described above may be used to identify the sets R_o^2, R_m^2; R_o^3, R_m^3, etc. in an analogous way.

The procedure described above suggests the possibility of using an analog or digital computer and self-organizing system techniques to search out and determine points on the boundaries of the sets $\{R_o^i\}$, $\{R_m^i\}$ by completely automatic machine solution. This is an interesting area for further research.

6. Secondary Performance Indices

The optimal control law $u^o(\underset{\sim}{x})$ is, in general, not unique[17] in the sets $\{R_o^i\}$. For this reason, the design and instrumentation of C-minimax optimal control laws affords a degree of flexibility which is not usually associated with optimal control laws for other performance indices. For example, in the sets $\{R_o^i\}$ it is not unusual to find that the same optimal performance is obtained when the controller is expressed as either (i) a bang-bang control law, (ii) a linear, continuous, control law, (iii) a non-linear, continuous, control law, or (iv) a combination of (i), (ii) and (iii).

The non-uniqueness of the C-minimax optimal control law in the sets $\{R_o^i\}$ suggests the possibility of introducing a secondary performance index for those sets. Suppose, for example, that $\hat{u}(\underset{\sim}{x})$ is an optimal control law for (2)-(5) with a certain Mayer, Lagrange, or Bolza type performance index. Then, the control law $\hat{u}(\underset{\sim}{x})$ can be used as the C-minimax optimal control law for (2)-(5) in a set $R_o^k \subset \{R_o^i\}$ provided that

$$\langle \nabla C(\underset{\sim}{x}), \underset{\sim}{F}(\underset{\sim}{x}, \hat{u}(\underset{\sim}{x})) \rangle \leq 0 \qquad \forall\, \underset{\sim}{x} \in R_o^k \tag{31}$$

---

16. Not all points of ∂R_m have this property, in general. In particular, some subsets of ∂R_m may be defined by integral manifolds of optimal trajectories which belong to R_m. For instance, see Example 4, Section 8.

17. The optimal control law in the sets $\{R_m^i\}$ can be non-unique in certain exceptional cases. See Example 2, Section 8.

is satisfied together with the appropriate boundary conditions. Some applications of this technique are illustrated in Section 8 below.

7. Some Alternative Methods of Solution

If the performance index $C(\underset{\sim}{x})$ is a non-negative definite function it can be shown that the functional (1) can be written as[18]

$$\max_{t_o \le t \le T} C(\underset{\sim}{x}(t)) = \lim_{\mu \to \infty} \left[\int_{t_o}^{T} [C(\underset{\sim}{x}(t))]^{\mu} \, dt \right]^{1/\mu}, \quad C(\underset{\sim}{x}(t)) \ge 0 \quad (32)$$

where $(\cdot)^{1/\mu}$ denotes the real and positive μ^{th} root of $(\cdot)$. [An elementary proof of (32) is given in the Appendix.] Thus, for the special case when $C(\underset{\sim}{x})$ is non-negative definite, a solution to the original C-minimax optimal control problem (1)-(5) can be obtained, through a limiting process, by minimizing instead the Lagrange-type performance index

$$J[u] = \lim_{\mu \to \infty} \int_{t_o}^{T} [C(\underset{\sim}{x}(t))]^{\mu} \, dt \quad (33)$$

subject to the same conditions (2)-(5). It is interesting to note that, although the control obtained by minimizing the performance index (33), as $\mu \to \infty$, does coincide with one of the optimal controls which minimizes (1), it does not exhibit the same degree of non-uniquesness, in general. An application of this alternative method of solution is illustrated in Example 1, Section 8. A further discussion of this topic may be found in [35].

Another alternative method of solution consists in replacing the original C-minimax optimal control problem (1)-(5) by the following problem of controllability in restricted state space: Find an admissible control $u = u^c(t)$ which transfers the state of the dynamical system (2) from the initial state $\underset{\sim}{x}(t_o) = \underset{\sim}{x}_o \in D$ to the terminal

---

18. Equation (32) may be recognized as the definition of the norm in a $\mathcal{L}_\infty [t_o, T]$ Banach space [40] with elements $C(\underset{\sim}{x}(t)) \ge 0$.

manifold (4), in some time interval $[t_o, T]$, subject to the state space inequality constraint

$$[C(\underset{\sim}{x}(t)) - \beta] \leq 0 \qquad t_o \leq t \leq T \tag{34}$$

where β is a specified real scalar constant. The solution to this controllability problem, if it exists, is not unique, in general. Let $\Phi(\underset{\sim}{x}_o, \beta)$, $\underset{\sim}{x}_o \in D$, be the set of all controls $u^c(t) = \phi^c(t; \underset{\sim}{x}_o, \beta)$ which are solutions and let $B(\underset{\sim}{x}_o)$ be the set of all β for which $\Phi(\underset{\sim}{x}_o, \beta)$ is non-empty. It is evident that

$$C(\underset{\sim}{x}_o) \leq \max_{t_o \leq t \leq T} C(\underset{\sim}{x}(t)) \leq \beta \tag{35}$$

along the solution of

$$\dot{\underset{\sim}{x}} = \underset{\sim}{F}(\underset{\sim}{x}, \phi^c(t; \underset{\sim}{x}_o, \beta)), \qquad \phi^c \in \Phi(\underset{\sim}{x}_o, \beta), \qquad \beta \in B(\underset{\sim}{x}_o) \tag{36}$$

Clearly, $B(\underset{\sim}{x}_o)$ is bounded from below by $C(\underset{\sim}{x}_o)$. Let $\beta^*(\underset{\sim}{x}_o)$ be the greatest lower bound of $B(\underset{\sim}{x}_o)$. Then for each $\underset{\sim}{x}_o \in D$, and over the union of all $\Phi(\underset{\sim}{x}_o, \beta)$, $\beta \in B(\underset{\sim}{x}_o)$, we have[19]

$$\inf \max_{t_o \leq t \leq T} C(\underset{\sim}{x}(t)) = \beta^*(\underset{\sim}{x}_o) \tag{37}$$

along the solutions of (36). Therefore, if $\beta^*(\underset{\sim}{x}_o) \in B(\underset{\sim}{x}_o)$, a control $\phi^c(t; \underset{\sim}{x}_o, \beta^*(\underset{\sim}{x}_o)) \in \Phi(\underset{\sim}{x}_o, \beta^*(\underset{\sim}{x}_o))$ is optimal with respect to the original C-minimax performance index (1). In this case $\beta^*(\underset{\sim}{x}) = V(\underset{\sim}{x})$ in the sense of (6).

Warga [30], [31] has developed an elegant alternative method of solution by using the result (37) to convert the original problem (1)-(5) into a special Mayer problem in restricted state space and then applying a comprehensive set of necessary conditions, (developed by Warga), for solutions of Mayer-type optimal control

---

19. Suppose, for example, that there exists a $\phi^c \in \bigcup_{\beta \in B(\underset{\sim}{x}_o)} \Phi(\underset{\sim}{x}_o, \beta)$ such that $\max_{t_o \leq t \leq T} C(\underset{\sim}{x}(t)) < \beta^*(\underset{\sim}{x}_o)$. Then, there must exist a $\tilde{\beta} < \beta^*(\underset{\sim}{x}_o)$ such that $\max_{t_o \leq t \leq T} C(\underset{\sim}{x}(t)) \leq \tilde{\beta}$. This latter result implies that $\tilde{\beta} \in B(\underset{\sim}{x}_o)$ which is a contradiction.

problems in restricted state space. In an independent study, Dubovitskii and Milyutin [28], [29] used function space techniques involving Stieltjes integrals to develop a set of necessary conditions (for C-minimax control problems) which closely resemble some of the results given in [30]. It is remarked that the possibility of using Stieltjes integrals in the study of C-minimax control problems was pointed out by Bellman et al. in [22].

The functional equation technique of dynamic programming has also been used to study discrete versions of several special cases of the problem (1)-(5). Some of the results are described in [23], [24] and [25].

8. Examples

The following examples illustrate application of the method of solution proposed in Section 3. It will be noted that the simplicity of these examples permits the validity of the solutions obtained to be readily verified by inspection.

Example 1. As a special case of (1)-(5), let

$$J[u] = \max_{t_o \le t \le T} x_1^2(t) \tag{38}$$

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u \end{aligned} \tag{39}$$

$$\underset{\sim}{x}(t_o) = \underset{\sim}{x}_o \tag{40}$$

$$\underset{\sim}{x}(T) = \underset{\sim}{0} \qquad T \text{ is unrestricted} \tag{41}$$

$$|u(t)| \le 1 \tag{42}$$

For this problem, it is readily verified by inspection that the set R_o consists of the closed set of states $\underset{\sim}{x}$ bounded by the curves

$$\partial R_o : \begin{cases} x_2 = 0 & (43) \\ x_1 + \frac{1}{2} |x_2| x_2 = 0 & (44) \end{cases}$$

and lying in the second and fourth quadrants of the x_1, x_2 plane. Equation (13) is

satisfied along the boundary segment defined by (43) and (14)-(15) are satisfied along the boundary segment defined by (44). The optimal control $\phi(t; \underset{\sim}{x}_o)$, $\underset{\sim}{x}_o \in R_o$, can be chosen as any admissible control which satisfies (8), (11) and (41). An optimal control law must satisfy the additional requirement

$$\frac{d}{dt}(x_1^2) \leq 0 \qquad t_o \leq t \leq T \tag{45}$$

along the corresponding solution of (39). One control law satisfying these requirements is given by

$$u^o(\underset{\sim}{x}) = -\text{sgn}\left[x_1 + \frac{1}{2}|x_2| x_2\right] \qquad \underset{\sim}{x} \in R_o \tag{46}$$

where, in this particular instance,

$$\text{sgn}[0] = \begin{cases} +1 & \text{if } x_1 < 0 \\ \\ -1 & \text{if } x_1 > 0 \end{cases} \tag{47}$$

The control law (47) may be recognized [1] as the time-optimal control law for the dynamical system described by (39)-(42).

The set R_m consists of the largest set of states $\underset{\sim}{x} \in (E^2 - R_o)$ for which the condition

$$x_1^2(t) \leq x_1^2(t_1) \qquad t_o \leq t \leq t_1 \tag{48}$$

is satisfied naturally along solutions of the, Mayer type, variational problem (24). The necessary conditions satisfied by such solutions are, from (21)-(23),

$$\max_{|u| \leq 1} [p_1(t)\, x_2(t) + p_2(t)\, u(t)] \equiv 0 \qquad t_o \leq t \leq t_1 \tag{49}$$

$$\dot{p}_1 = 0 \tag{50}$$

$$\dot{p}_2 = -p_1 \tag{51}$$

$$p_1(t_1) = -2x_1(t_1) \tag{52}$$

$$p_2(t_1) = 0 \tag{53}$$

From (49) it follows that

$$u^{o}(t) = \text{sgn } p_2(t) \qquad \underset{\sim}{x}(t) \in R_m \tag{54}$$

Moreover, it is readily verified that in order to satisfy (49)-(53) it is necessary to choose

$$p_2(t) = -2\,|x_2(t)|\,x_1(t_1) \tag{55}$$

Thus, (54) can be written in the form

$$u^{o}(t) = -\text{sgn}\,[x_1(t_1)] \qquad \underset{\sim}{x} \in R_m \tag{56}$$

or, since $\text{sgn } x_1(t_1) = \text{sgn } x_2(t)$, $(x_2(t) \neq 0)$, (56) can be written in the control law form

$$u^{o}(\underset{\sim}{x}(t)) = -\text{sgn}\,[x_2(t)] \qquad \underset{\sim}{x} \in R_m \tag{57}$$

From (39) and (57) the optimal trajectories in the set R_m are obtained as the one parameter family of curves defined by

$$x_1 + \frac{1}{2}\,|x_2|\,x_2 = k \qquad k = \text{real, scalar constant} \tag{58}$$

Since $\dot{x}_1(t)$ does not change sign along the optimal trajectories in R_m, it follows from (58) that (48) is satisfied in the region $(E^2 - R_o)$ where

$$\begin{cases} x_1 + \frac{1}{4}\,x_2^2 \geq 0 & x_2 > 0 \\ x_1 - \frac{1}{4}\,x_2^2 \leq 0 & x_2 < 0 \end{cases} \tag{59}$$

Thus, the set R_m is bounded by the curves

$$\partial R_m : \begin{cases} x_2 = 0 \\ x_1 + \frac{1}{4}\,|x_2|\,x_2 = 0 \end{cases} \tag{60}$$

In the complement of $R_o \cup R_m$ every admissible control law has the property that if $\underset{\sim}{x}(t_o) \in E^2 - (R_o \cup R_m)$ then $\underset{\sim}{x}(t_2)$ satisfies the second of (60) for some t_2 and, in addition, $\underset{\sim}{x}(t) \in E^2 - (R_o \cup R_m)$ and $d/dt(x_1^2(t)) \leq 0$, $t_o \leq t \leq t_2$. It follows that $E^2 - (R_o \cup R_m)$ is a set R_o^2 of the R_o type and <u>every</u> admissible control law is an optimal control law for $\underset{\sim}{x} \in R_o^2$! In particular, one can choose

$$u^o(\underset{\sim}{x}(t)) = -\text{sgn}\,[x_2(t)] \qquad \underset{\sim}{x} \in R_o^2 \tag{61}$$

The x_1, x_2 plane is now completely partitioned into sets of the R_o and R_m type. One choice for the corresponding set of optimal control laws is given by (46), (57) and (61). It may be noted that (46) gives a correct optimal control law for <u>all</u> three sets R_o, R_m, R_o^2. Thus, for this example, the minimum-time control law also minimizes (38).

In the sets R_o, R_o^2, (6) is given by

$$V(\underset{\sim}{x}) = x_1^2, \qquad \underset{\sim}{x} \in R_o \cup R_o^2 \tag{62}$$

In the set R_m, the $V(\underset{\sim}{x})$ = constant contours correspond to the trajectories (58). Thus, in R_m

$$V(\underset{\sim}{x}) = [x_1 + \tfrac{1}{2}\,|x_2|\,x_2]^2 \qquad \underset{\sim}{x} \in R_m \tag{63}$$

The sets R_o, R_m, R_o^2 are shown in Fig. 1, together with some representative $V(\underset{\sim}{x})$ = constant contours and a typical optimal trajectory corresponding to the optimal control law (46).

In this particular example, it turns out that the sets R_o and R_o^2 share a common boundary defined by (44). For this reason, the previously selected optimal control law for the set $R_o \cup R_o^2$ [i.e. (46) and (61)] can be replaced by any admissible control law such that (45) is satisfied. One such alternative optimal control law, which in view of (57) happens to be optimal for the set R_m as well, is given by

$$u^o(\underset{\sim}{x}) = -\text{sgn}\,[x_1 + \tfrac{1}{4}\,|x_2|\,x_2] \qquad \underset{\sim}{x} \in E^2 \tag{64}$$

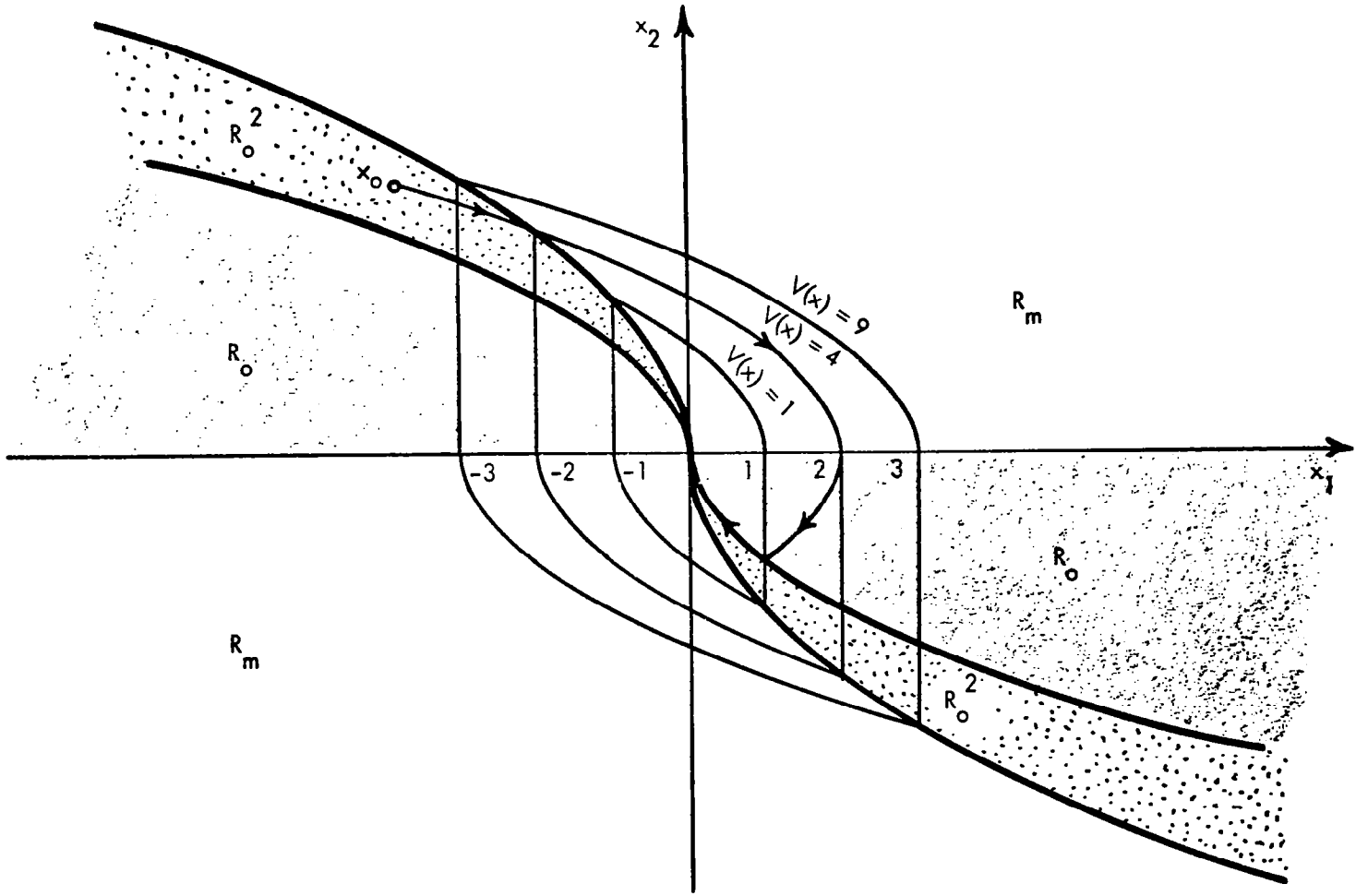

FIGURE 1

Any control law which minimizes (38) also minimizes

$$J[u] = \max_{t_o \le t \le T} |x_1(t)| \tag{65}$$

and vice versa. Moreover, from (32), (33) it follows that the limiting control law obtained by minimizing the integral

$$J[u] = \int_{t_o}^{T} |x_1(t)|^{\mu} \, dt, \tag{66}$$

as $\mu \to \infty$, must coincide with one of the control laws which minimize (38) and (65). Fuller [41] has shown that the control law $u(\underset{\sim}{x}, \mu)$ which minimizes (66), subject to the conditions (39)-(42), can be written as

$$u(\underset{\sim}{x}, \mu) = -\operatorname{sgn} \left[x_1 + \frac{1}{2} \frac{1 - k(\mu)}{1 + k(\mu)} |x_2| \, x_2 \right] \tag{67}$$

where the parameter $k(\mu)$ is determined from a certain auxiliary algebraic equation and has the particular values $k(1) \approx 0.01433$, $k(2) \approx 0.05862$, and $\lim_{\mu \to \infty} k(\mu) = 1/3$. Thus, the control law which minimizes (66), as $\mu \to \infty$, coincides exactly with the alternative C-minimax optimal control law (64).

Example 2. The following example illustrates a case in which the performance index $C(\underset{\sim}{x})$ is not sign definite and the optimal trajectories in a subset of R_m are not unique.

Suppose the performance index of the problem (1)-(5) has the form

$$J[u] = \max_{t_o \le t \le T} x_1(t) \tag{68}$$

with (2)-(5) the same as (39)-(42) in Example 1. Then, following the same procedure as in the previous example, it may be verified that the set R_o is the closed set of points in the fourth quadrant of the x_1, x_2-plane bounded by the curves

$$\partial R_o : \begin{cases} x_2 = 0 & x_1 \geq 0 \\ x_1 - \frac{1}{2} x_2^2 = 0 & x_2 \leq 0 \end{cases} \tag{69}$$

The set R_m is likewise found to be the set of all points in the first, second, and third quadrants of the x_1, x_2-plane with the exception of the points on the positive x_1-axis. In the particular subset $R_{m_1} \subset R_m$ defined by

$$R_{m_1} = \{ \underset{\sim}{x} \mid x_1 + \frac{1}{2} x_2^2 \geq 0, \quad x_2 > 0 \} \tag{70}$$

the optimal control $u^o(t)$ is unique and can be written in the control law form

$$u^o(\underset{\sim}{x}) = -\text{sgn}\,(x_2) \qquad x_2 \in R_{m_1} \tag{71}$$

In the set $R_{m_2} = R_m - R_{m_1}$, however, the optimal control, obtained by solving the appropriate Mayer problem (24), is non-unique. This is due to the fact that the function $C(\underset{\sim}{x}) = x_1$, evaluated on the "terminal manifold" ∂R_o, attains its minimum value at a state ($\underset{\sim}{x} = \underset{\sim}{0}$) which can be reached, using an admissible control satisfying (20), from every initial state $\underset{\sim}{x}_o \in R_{m_2}$. That is, the functional $J[u]$ defined by (68) turns out to be "independent of path" for every trajectory $\underset{\sim}{x}(t) \in R_{m_2}$, $t_o \leq t < T$, which satisfies (41). It follows that

$$V(\underset{\sim}{x}) = \text{constant} = 0 \qquad \forall\ \underset{\sim}{x} \in R_{m_2} \tag{72}$$

It may be verified that the set $E^2 - R_o \cup R_m$ is a set R_o^2 of the R_o type. The sets R_o, R_m and R_o^2, together with some representative $V(\underset{\sim}{x}) =$ constant contours and a typical optimal trajectory, are illustrated in Fig. 2.

Example 3. As another special case of (1)-(5), let

$$J[u] = \max_{t_o \leq t \leq T} [x_1(t) + x_2(t)]^2 \tag{73}$$

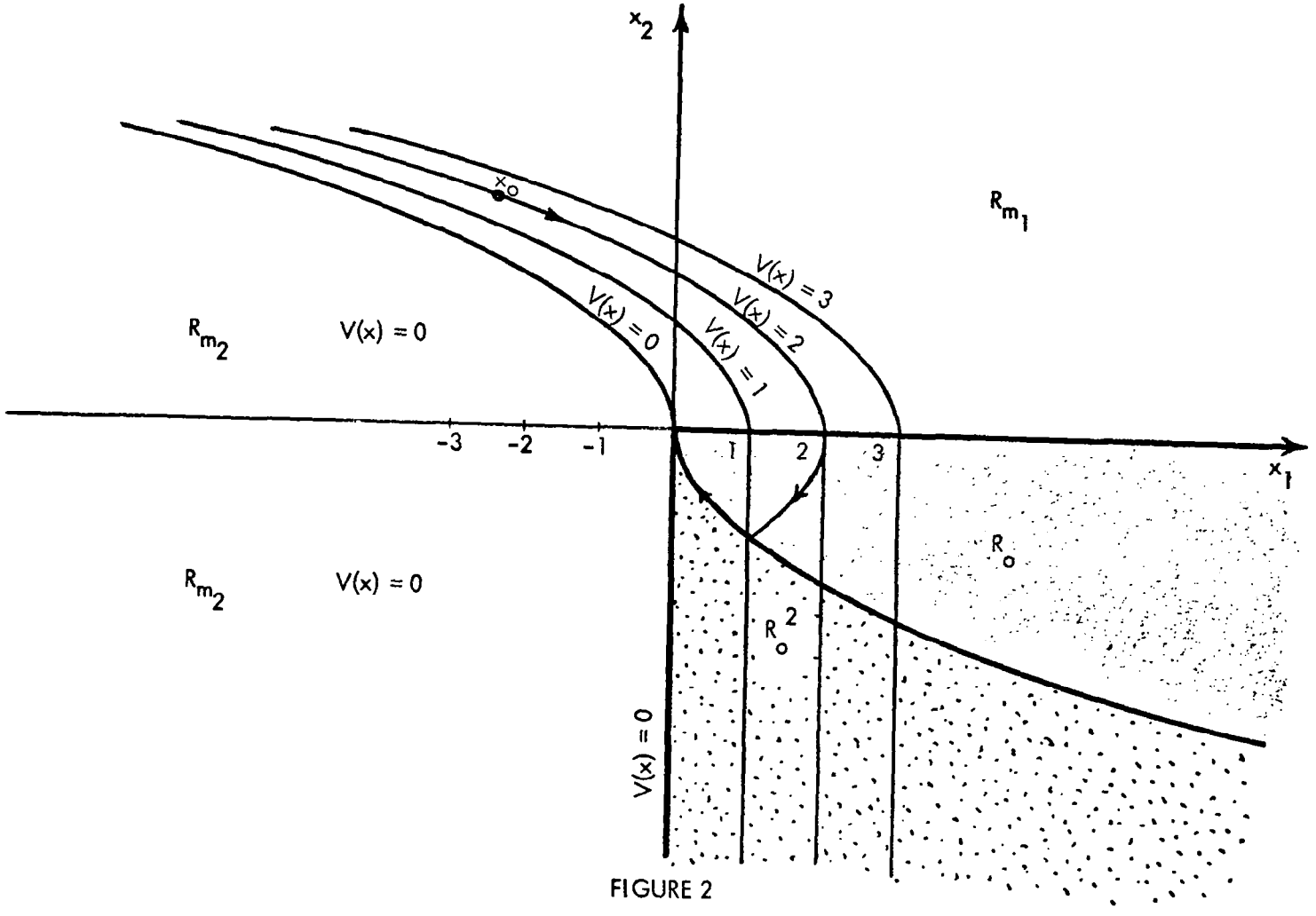

FIGURE 2

with (2)-(5) the same as (39)-(42) in Example 1. The absolute minimum of the functional (73) occurs on the line

$$x_1 + x_2 = 0 \tag{74}$$

which is an integral curve of (39) when u is chosen appropriately. Therefore, the construction of the set R_o begins by looking for a segment of the line (74) which is an integral curve of (39), for some admissible control, and along which the terminal condition (41) is satisfied. It is readily verified that such a segment exists and is defined by

$$x_1 + x_2 = 0 \qquad |x_1| \leq 1 \tag{75}$$

It is observed that the terminal condition (41) is satisfied along the trajectory (75) only as $t \to \infty$. The remainder of the set R_o consists of the set of all states $\underset{\sim}{x}$ which can be joined to the segment (75) by admissible integral curves of (39) along which the performance index $[x_1(t) + x_2(t)]^2$ is identically non-increasing. This set is bounded in part by the lines $|x_2| = 1$ and in part by the curved segments defined by

$$x_2 = \begin{cases} + \sqrt{-2x_1 - 1} & x_1 \leq -1 \\ \\ - \sqrt{2x_1 - 1} & x_1 \geq 1 \end{cases} \tag{76}$$

It is observed that ∂R_o is not everywhere differentiable.

The optimal control law in the set R_o can be constructed in a variety of ways since the optimal control in that region is non-unique. One possible optimal control law is constructed as follows. Equations (75) and (76) together define a continuous curve Ω in the x_1, x_2 state plane. In the subset of R_o which lies to the right of the curve Ω, the optimal control law can be chosen as

$$u^o(\underset{\sim}{x}) = \begin{cases} -\text{sgn } \psi_1(\underset{\sim}{x}), \ \psi_1(\underset{\sim}{x}) \neq 0, & x_1 > 1 \\ +1, \ \psi_1(\underset{\sim}{x}) = 0, & x_1 > 1 \\ -\text{sgn } \zeta(\underset{\sim}{x}), \ \zeta(\underset{\sim}{x}) \neq 0, & |x_1| \leq 1 \\ -x_2 - k\,\zeta(\underset{\sim}{x}), \ \zeta(\underset{\sim}{x}) = 0, & |x_1| \leq 1, \ k \geq 0 \end{cases} \tag{77}$$

where $\psi_1(\underset{\sim}{x}) = \sqrt{2x_1 - 1} + x_2$ and $\zeta(\underset{\sim}{x}) = x_1 + x_2$. Likewise, in the subset of R_o which lies to the left of the curve Ω, the optimal control law can be chosen as

$$u^o(\underset{\sim}{x}) = \begin{cases} -\text{sgn } \psi_2(\underset{\sim}{x}), & \psi_2(\underset{\sim}{x}) \neq 0, \quad x_1 < -1 \\ -1, & \psi_2(\underset{\sim}{x}) = 0 \quad x_1 < -1 \\ -\text{sgn } \zeta(\underset{\sim}{x}), & \zeta(\underset{\sim}{x}) \neq 0, \quad |x_1| \leq 1 \\ -x_2 - k\,\zeta(\underset{\sim}{x}), & \zeta(\underset{\sim}{x}) = 0, \quad |x_1| \leq 1, \quad k \geq 0 \end{cases} \tag{78}$$

where $\psi_2(\underset{\sim}{x}) = -\sqrt{-2x_1 - 1} + x_2$. The arbitrary scalar constant k in (77) and (78) is introduced as a technical device to permit stabilization [39] of the integral curve (75). Altematively, the control law in the last members of (77) and (78) can be replaced by the expression

$$u^o(\underset{\sim}{x}) = -\text{sgn } (x_1 + x_2), \quad |x_1| \leq 1 \tag{79}$$

in which case the representative point $\underset{\sim}{x}(t)$ will move along the curve (75) in a sliding (chattering) mode [42].

The set R_m is determined, as before, by solving the appropriate Mayer type variational problem (24) using the boundary of R_o as the terminal manifold. In this way, it is found that the boundary of R_m is defined, in part, by the lines $|x_2| = 1$ and in part, by the curved segments

$$x_2 = \begin{cases} -1 + 2\sqrt{-x_1} & x_1 \leq -1 \\ \\ 1 - 2\sqrt{x_1} & x_1 \geq 1 \end{cases} \tag{80}$$

The optimal control law in the set R_m can be written as

$$u^o(\underset{\sim}{x}) = -\text{sgn } x_2 \qquad \underset{\sim}{x} \in R_m \tag{81}$$

The complement of $R_o \cup R_m$ is a set of the R_o type (and is therefore denoted by R_o^2) because for each initial state $\underset{\sim}{x}_o \in R_o^2$ there exists an admissible control law and a time $t_2 > t_o$ such that $\underset{\sim}{x}(t) \in R_o^2$ and $d/dt[x_1(t) + x_2(t)]^2 \leq 0$ for all $t_o \leq t \leq t_2$

and $\underset{\sim}{x}(t_2) \in \partial R_m$. For example, choose the control law

$$u^o(\underset{\sim}{x}) = -\text{sgn } x_2 \qquad \underset{\sim}{x} \in R_o^2 \tag{82}$$

In the sets R_o and R_o^2, (6) is given by

$$V(\underset{\sim}{x}) = (x_1 + x_2)^2 \qquad \underset{\sim}{x} \in R_o \cup R_o^2 \tag{83}$$

In the set R_m, $V(\underset{\sim}{x})$ = constant along optimal trajectories and (6) is given by

$$V(\underset{\sim}{x}) = [x_1 \text{ sgn } x_2 + \tfrac{1}{2}(1 + x_2^2)]^2, \qquad \underset{\sim}{x} \in R_m \tag{84}$$

Expressions (80), which define the common boundary segments between R_m and R_o^2 can be obtained, alternatively, by equating expressions (83) and (84).

The sets R_o, R_m, and R_o^2, together with some representative $V(\underset{\sim}{x})$ = constant contours and a typical optimal trajectory, are illustrated in Fig. 3.

It is interesting to compare the optimal control law (77), (78), (81), (82), for the present example with the optimal control law for the problem of minimizing the Lagrange, integral type, performance index

$$J[u] = \int_{t_o}^{T} [x_1(t) + x_2(t)]^2 \, dt, \qquad (T \text{ - unspecified}) \tag{85}$$

subject to the same restrictions (39)-(42). The solution to this latter problem has been described in [39], [43], and [44] and consists of both a bang-bang mode and a singular mode. The singular mode trajectory is identical with the line segment (75) and is joined, at $|x_1| = 1$, to two bang-bang switching curves. A comparison of these two solutions is shown in Fig. 4. It may be verified, from Fig. 4 and the fact that the sets R_o, R_o^2 share a common boundary, that the control law which minimizes (85) also minimizes (73) for initial conditions $\underset{\sim}{x}_o$ sufficiently near the origin.

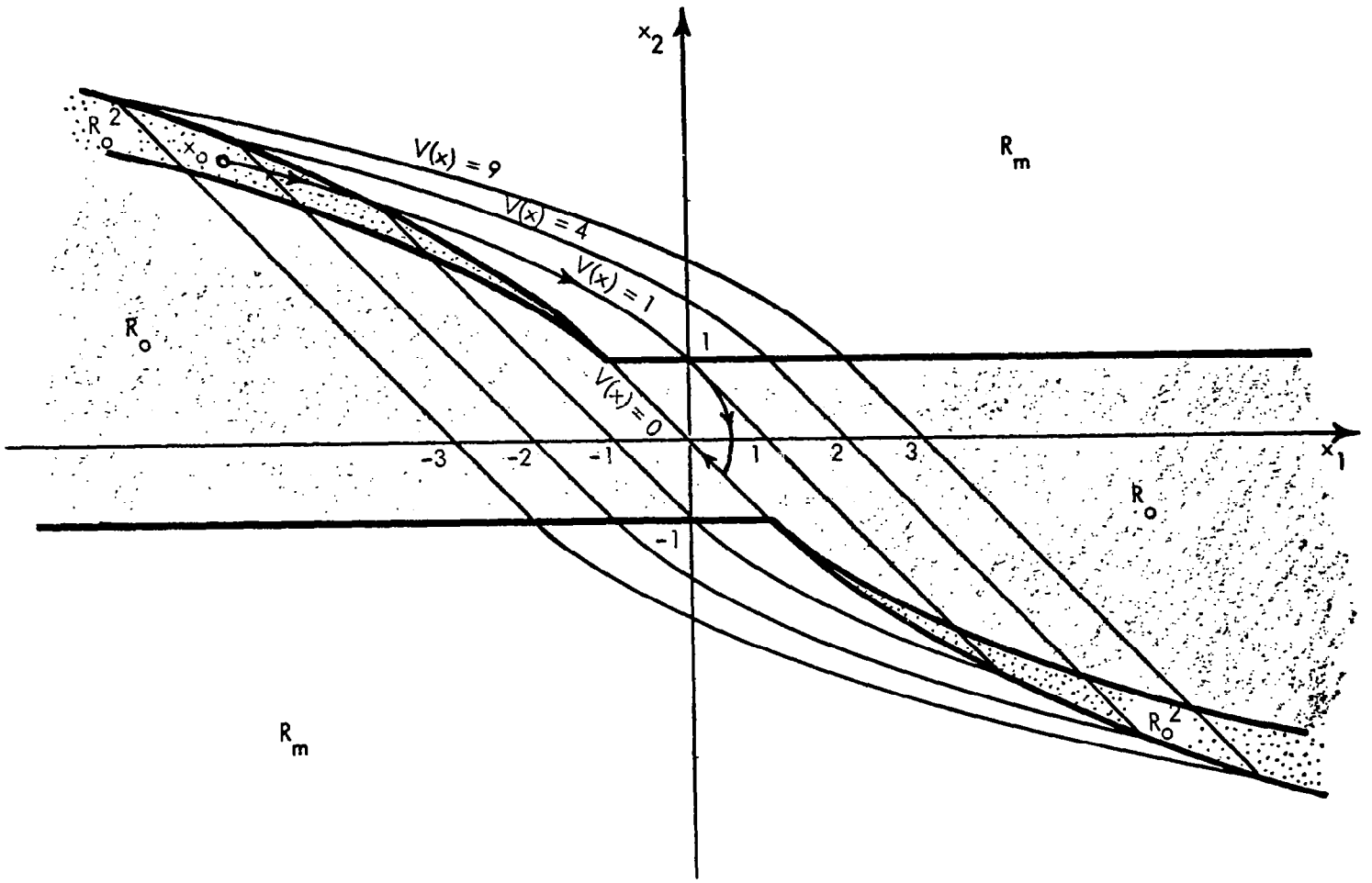

FIGURE 3

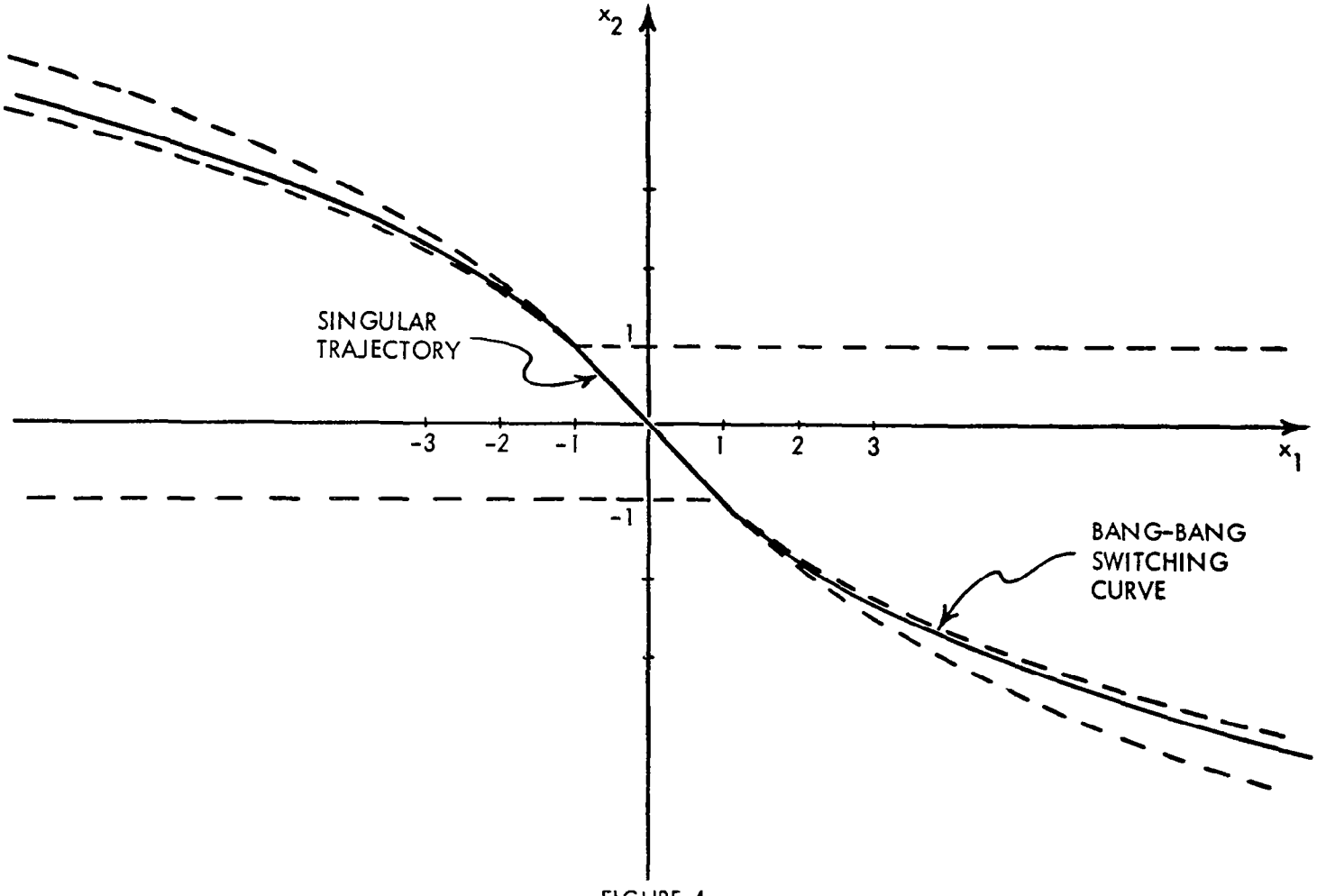
x_2
SINGULAR
TRAJECTORY
1
-3
-2
-1
1
2
3
x_1
-1
BANG-BANG
SWITCHING
CURVE

FIGURE 4

Example 4. As another special case of (1)-(5), let

$$J[u] = \max_{t_o \le t \le T} x_1^2(t) \tag{86}$$

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

$$\dot{x}_2 = -x_1 + u(t) \tag{87}$$

$$\underset{\sim}{x}(t_o) = \underset{\sim}{x}_o$$

$$\underset{\sim}{x}(T) = \underset{\sim}{0} \qquad T \text{ - unrestricted} \tag{88}$$

$$|u(t)| \le 1 \tag{89}$$

For this problem it may be verified that R_o is the connected and closed set of states $\underset{\sim}{x}$ bounded by the curves

$$\partial R_o : \begin{cases} x_2 = 0 \\ (x_1 + \operatorname{sgn} x_2)^2 + x_2^2 = 1, \quad |x_1| \le 2 \end{cases} \tag{90}$$

and lying in the second and fourth quadrants of the x_1, x_2-plane. The optimal control law for $\underset{\sim}{x} \in R_o$ can be chosen as any admissible control law which satisfies (8), (11), (88) and

$$\frac{d}{dt}(x_1^2(t)) \le 0 \qquad t_o \le t \le T \tag{91}$$

along the corresponding solution of (87). One such control law is given by

$$u^o(\underset{\sim}{x}) = \operatorname{sgn}[x_1^2 - 2|x_1| + x_2^2], \qquad \underset{\sim}{x} \in R_o \tag{92}$$

with

$$\operatorname{sgn}[0] = \begin{cases} +1 \text{ if } x_1 > 0 \\ -1 \text{ if } x_1 < 0 \end{cases}$$

which may be recognized [1] as the time-optimal control law [in the set R_o] for the dynamical system described by (87).

The set R_m is determined, as before, by solving the appropriate Mayer problem (24). In this way it is found that R_m is the two sets of points bounded by the curves

$$\partial R_m : \begin{cases} x_2 = 0 \\ x_1 + \frac{1}{4}|x_2| x_2 = 0 \\ (x_1 + \operatorname{sgn} x_2)^2 + x_2^2 = 9 \qquad |x_1| \leq 2 \end{cases} \tag{93}$$

In R_m the optimal control law can be expressed as

$$u^o(\underset{\sim}{x}) = -\operatorname{sgn} x_2, \qquad \underset{\sim}{x} \in R_m \quad \operatorname{sgn}(0) = \operatorname{sgn}(x_1) \tag{94}$$

It may be noted that the particular boundary segment of R_m defined by the last expression in (93) is an optimal trajectory which belongs to R_m.

The set $R_o^2 \subset (E^2 - R_o \cup R_m)$ is the *largest set of initial states* $\underset{\sim}{x}_o$ *with the* following property. For each $\underset{\sim}{x}_o \in R_o^2$ there exists an admissible control $u(t)$, $t_o \leq t \leq t_2$, such that, along the corresponding solution of (87), $C(\underset{\sim}{x}(t)) \leq C(\underset{\sim}{x}_o)$ and $\underset{\sim}{x}(t) \in R_o^2$ for all $t_o \leq t < t_2$ and $\underset{\sim}{x}(t_2) \in \partial R_m$. In contrast with the previous examples, the set R_o^2 for the present example is <u>not</u> the complement of $R_o \cup R_m$. Instead, it is found that R_o^2 is the two, disconnected, sets of states bounded by the curves

$$\partial R_o^2 : \begin{cases} x_2 = 0 \\ x_1 + \frac{1}{4}|x_2| x_2 = 0 \\ x_1^2 - 2|x_1| + x_2^2 = 0 \\ (x_1 + \operatorname{sgn} x_2)^2 + x_2^2 = 9 \end{cases} \tag{95}$$

and lying in the second and fourth quadrants of the x_1, x_2-plane. In R_o^2 the optimal control law can be chosen as

$$u^o(\underset{\sim}{x}) = \operatorname{sgn} x_2, \qquad \underset{\sim}{x} \in R_o^2 \tag{96}$$

The set R_m^2 is determined, as before, by solving the appropriate Mayer problem (24) using ∂R_o^2 as the "terminal manifold." In this way, it is found that R_m^2 is bounded by the curves

$$\partial R_m^2 : \begin{cases} x_2 = 0 \\ x_1 + \frac{1}{4}\,|x_2|\,x_2 = 0 \\ (x_1 + \operatorname{sgn} x_2)^2 + x_2^2 = 9 \\ (x_1 + \operatorname{sgn} x_2)^2 + x_2^2 = 25 \end{cases} \tag{97}$$

It is noted that the set R_m^2 shares a common boundary with the set R_m. In R_m^2 the optimal control law can be expressed in the same form as (94).

The remainder of E^2 is partitioned into the sets R_o^3, R_o^4, ... and $R_m^3, R_m^4, \ldots$ by repeating the process described above. It may be verified that, because the sets (R_o, R_o^2), (R_o^2, R_o^3), etc. share common boundaries, the individual control laws for the families of sets $\{R_o^i\}$ and $\{R_m^i\}$ can be replaced by the <u>one</u> control law

$$u^o(\underset{\sim}{x}) = -\operatorname{sgn}\left[x_1 + \frac{1}{4}\,|x_2|\,x_2\right], \qquad \underset{\sim}{x} \in E^2 \tag{98}$$

which is optimal throughout E^2. Alternatively, the somewhat more complex <u>time-optimal</u> control law [1] for (87)-(89) can also be used as a C-minimax optimal control law for arbitrary $\underset{\sim}{x} \in E^2$.

The particular sets R_o^i, R_m^i, $i = 1, \ldots, 3$ together with some representative V = constant contours and a typical optimal trajectory are illustrated in Figure 5.

<u>Example 5</u>. As a special case of (1)-(5), let

$$J[u] = \max_{t_o \le t \le T} \left[x_1^2(t) + x_2^2(t)\right] \tag{99}$$

with (2)-(5) the same as (39)-(42) in Example 1.

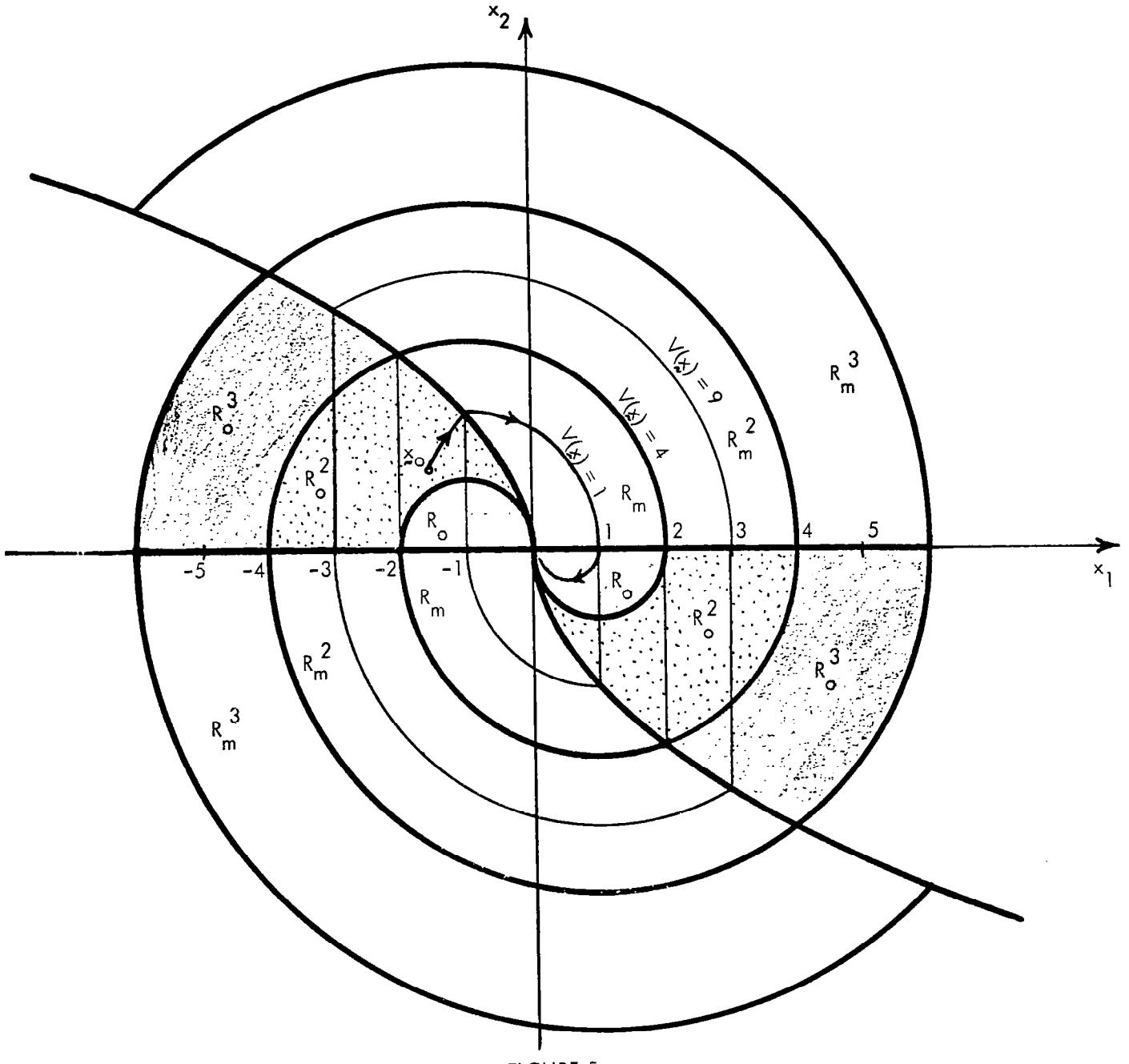

FIGURE 5

For this problem, the set R_o is the closed and connected set of states $\underset{\sim}{x}$ bounded by the curves

$$\partial R_o : \begin{cases} x_2 = 0 & |x_1| > 1 \\ x_1 + \frac{1}{2}|x_2|\, x_2 = \operatorname{sgn} x_2 \end{cases} \tag{100}$$

It may be noted that ∂R_o is not everywhere differentiable.

The optimal control for $\underset{\sim}{x}_o \in R_o$ can be chosen as any admissible control which satisfies (8), (11) and (41). An optimal control law must satisfy the additional requirement

$$\frac{d}{dt}\left[x_1^2(t) + x_2^2(t)\right] \leq 0 \qquad t_o \leq t \leq T \tag{101}$$

along the corresponding solution of (39).

The set R_m consists of the largest set of states $\underset{\sim}{x} \in (E^2 - R_o)$ for which the condition

$$\left[x_1^2(t) + x_2^2(t)\right] \leq \left[x_1^2(t_1) + x_2^2(t_1)\right] \quad t_o \leq t \leq t_1 \tag{102}$$

is satisfied naturally along solutions of the Mayer variational problem (24). Proceeding as in Example 1, it is found that the set R_m consists of two disconnected sets: (i) the set of states $\underset{\sim}{x}$ lying above the broken curve defined by

$$\partial R_m : \begin{cases} x_2 = 0 & x_1 \geq 1 \\ x_1 + \frac{1}{4} x_2^2 = +1 & x_2 \geq 0 \end{cases} \tag{103}$$

and (ii) the set of states $\underset{\sim}{x}$ lying below the broken curve defined by

$$\partial R_m : \begin{cases} x_2 = 0 & x_1 \leq -1 \\ x_1 - \frac{1}{4} x_2^2 = -1 & x_2 \leq 0 \end{cases} \tag{104}$$

In the set R_m the optimal control law can be written as

$$u^o(\underset{\sim}{x}) = -\text{sgn } x_2 \qquad \underset{\sim}{x} \in R_m \tag{105}$$

The complement of $R_o \cup R_m$ is a set of the R_o type and is therefore denoted by R_o^2. For any state $\underset{\sim}{x} \in R_o^2$ it is always possible to find an admissible control law such that along the corrrsponding solution of (39) the following conditions are satisfied

$$\underset{\sim}{x}(t) \in R_o^2 \qquad t_o \leq t < t_2 \tag{106}$$

$$\underset{\sim}{x}(t_2) \in \partial R_m \tag{107}$$

$$\frac{d}{dt}\left[x_1^2(t) + x_2^2(t)\right] \leq 0 \qquad t_o \leq t \leq t_2 \tag{108}$$

for some $t_2 > t_o$. For example, one such control law is

$$u^o(\underset{\sim}{x}) = -\text{sgn } x_2 \qquad \underset{\sim}{x} \in R_o^2 \tag{109}$$

The x_1, x_2 plane is now completely partitioned into sets of the R_o and R_m type. In the set R_o and R_o^2, (6) is given by

$$V(\underset{\sim}{x}) = (x_1^2 + x_2^2) \qquad \underset{\sim}{x} \in R_o \cup R_o^2 \tag{110}$$

and in the set R_m, (6) is given by (compare with (63))

$$V(\underset{\sim}{x}) = \left[x_1 + \frac{1}{2}\,|x_2|\,x_2\right]^2 \qquad \underset{\sim}{x} \in R_m \tag{111}$$

Since $V(\underset{\sim}{x})$ is continuous across the common boundary segments of R_m and R_o^2 the equations for those boundary segments, previously given in (103) and (104), can be obtained directly by equating (110) and (111).

The sets R_o, R_m, and R_o^2, together with some representative $V(\underset{\sim}{x})$ = constant contours and a typical optimal trajectory are illustrated in Fig. 6.

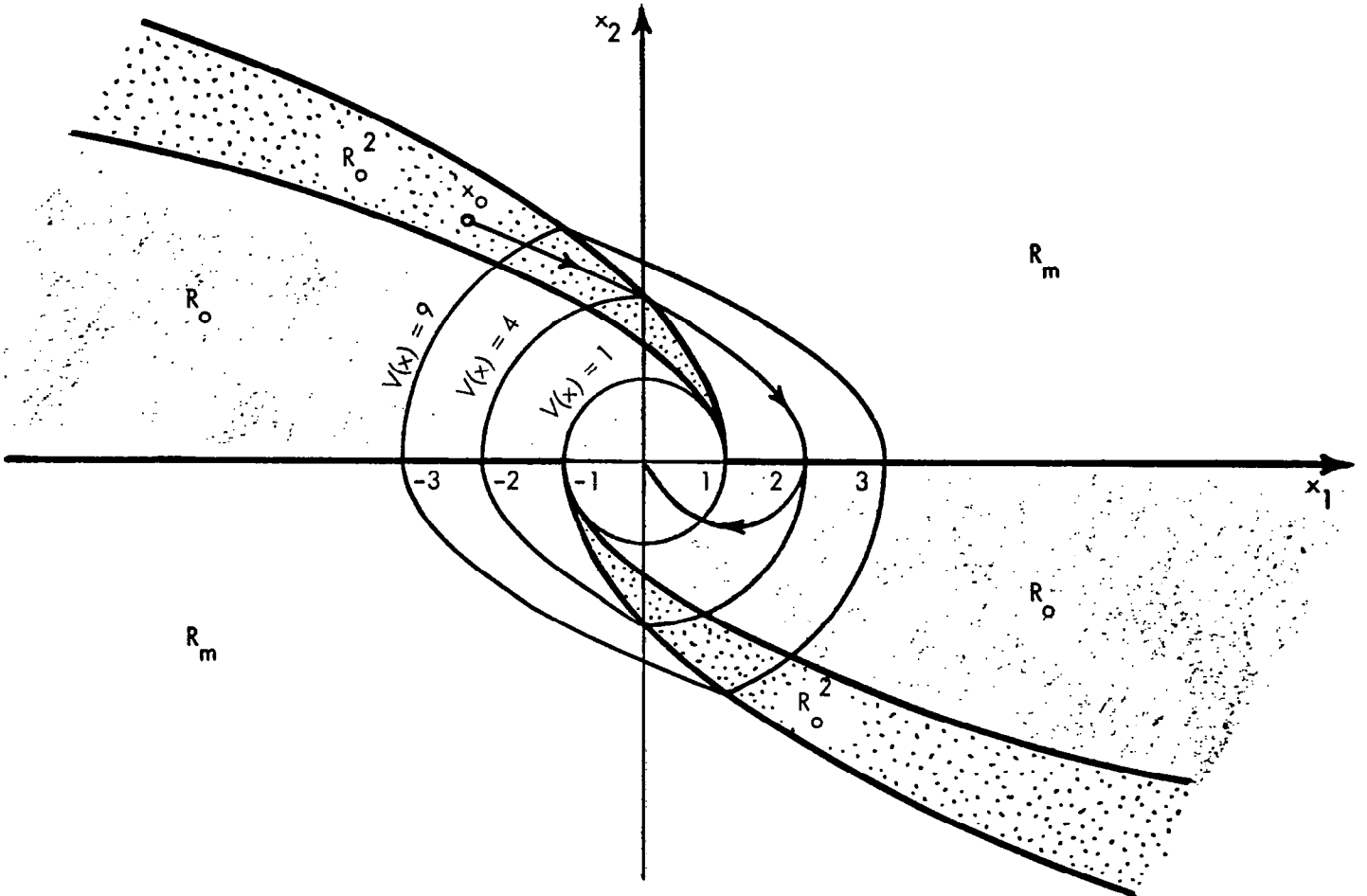

FIGURE 6

9. Performance Improvement with C-Minimax Control - A Geometric Interpretation

Suppose that the uncontrolled dynamical system described by

$$\dot{\underset{\sim}{x}} = \underset{\sim}{F}^{o}(\underset{\sim}{x}) \tag{112}$$

is asymptotically stable with respect to the terminal manifold (4) for all initial states $\underset{\sim o}{x} \in D^{o} \subset D$. Suppose further that the same dynamical system, when subjected to an external control $u(t)$, obeys the differential equation

$$\dot{\underset{\sim}{x}} = \underset{\sim}{F}(\underset{\sim}{x}, u(t)) \tag{113}$$

where $\underset{\sim}{F}(\underset{\sim}{x}, 0) = \underset{\sim}{F}^{o}(\underset{\sim}{x})$.

The maximum value of the performance index $C(\underset{\sim}{x}(t))$, $t_o \leq t \leq T$, which occurs along the solutions of the uncontrolled system (112) [with $\underset{\sim o}{x} \in D^{o}$], can be studied by the same techniques used in Section 3. By this means, the two sets $\mathcal{R}_o$ and $\mathcal{R}_m$ defined by

$$\begin{aligned} \mathcal{R}_o &= \{\underset{\sim o}{x} \in D^{o} \mid \max_{t_o \leq t \leq T} C(\underset{\sim}{x}(t)) = C(\underset{\sim o}{x})\} \\ \mathcal{R}_m &= \{\underset{\sim o}{x} \in D^{o} \mid \max_{t_o \leq t \leq T} C(\underset{\sim}{x}(t)) > C(\underset{\sim o}{x})\} \end{aligned} \tag{114}$$

can be constructed in the state space E^{n} of the system (112). Then, proceeding as in (6), one can define a function $V(\underset{\sim o}{x})$ on the subset $D^{o} \subset E^{n}$ by

$$V(\underset{\sim o}{x}) = \max_{t_o \leq t \leq T} C(\underset{\sim}{x}(t)) \qquad \underset{\sim o}{x} \in D^{o} \tag{115}$$

so that

$$V(\underset{\sim}{x}) = \begin{cases} C(\underset{\sim}{x}) \ , & \forall\ \underset{\sim}{x} \in \mathcal{R}_o \\ > C(\underset{\sim}{x}) & \forall\ \underset{\sim}{x} \in \mathcal{R}_m \end{cases} \tag{116}$$

The function $V(\underset{\sim}{x}_o)$ in (115) indicates the quality of C-minimax performance which is realized from the transient response of the uncontrolled system (112). The result (116) can be visualized geometrically in the product space $E^1 \times E^n$: $z, x_1, x_2, \cdots, x_n$ by considering the two surfaces $\mathcal{V}$ and $\mathcal{C}^o$ defined as

$$\mathcal{V} = \{(z, \underset{\sim}{x}) \in E^1 \times E^n \mid z - V(\underset{\sim}{x}) = 0\}$$
$$\mathcal{C}^o = \{(z, \underset{\sim}{x}) \in E^1 \times E^n \mid z - C(\underset{\sim}{x}) = 0\} \tag{117}$$

According to (116), the surface $\mathcal{V}$ lies "above" the surface $\mathcal{C}^o$ at each state $x \in \mathcal{R}_m$ while, at each state $x \in \mathcal{R}_o$, the surfaces $\mathcal{V}$ and $\mathcal{C}^o$ intersect (coincide). Thus, the set $\mathcal{R}_o \subset D^o$ is the projection, onto the subspace E^n, of the "points of contact" between the two surfaces $\mathcal{V}$ and $\mathcal{C}^o$.

It follows from (7), (115) and (116) that, for initial states $\underset{\sim}{x}_o \in \mathcal{R}_o$, the quality of C-minimax performance of the uncontrolled system (112) cannot be improved by application of external control $u(t)$. That is, the application of external control $u(t)$ cannot improve the quality of performance if the initial state $\underset{\sim}{x}_o$ corresponds to a "point of contact" between the two surfaces $\mathcal{V}$ and $\mathcal{C}^o$. On the other hand, if $\underset{\sim}{x}_o \in \mathcal{R}_m$ the surface $\mathcal{V}$ lies "above" the surface $\mathcal{C}^o$ and therefore the quality of C-minimax performance for that initial state can be improved by the application of external control, provided that an admissible control (satisfying (4)) can be found which moves the surface $\mathcal{V}$ "closer" to the surface $\mathcal{C}^o$. The new "points of contact" between $\mathcal{V}$ and $\mathcal{C}^o$, achieved by this means, correspond to states $\underset{\sim}{x}$ which leave the set $\mathcal{R}_m$ and join the set $\mathcal{R}_o$. In addition, any initial state $\underset{\sim}{x}_o \in E^n - D^o$ which can be controlled to the terminal manifold (4) by some admissible control $u(t)$ becomes a member of one or the other of the sets $\mathcal{R}_o$ or $\mathcal{R}_m$. In this way, the set D is obtained as the union of D^o and the set of all states $\underset{\sim}{x}_o \in E^n - D^o$ which can be controlled to $\mathcal{T}$ with an admissible control.

Thus, the effect of applying a C-minimax optimal control to the system (113) can be viewed in $E^1 \times E^n$ as a "depressing" of the surface $\mathcal{V}$ down onto the surface $\mathcal{C}^o$

in such a way as to (i) increase the areas of contact between $\mathcal{V}$ and $\mathcal{C}$ wherever possible and (ii) decrease the original "distance" between $\mathcal{V}$ and $\mathcal{C}$ at those states $\underset{\sim}{x} \in D$ where contact between $\mathcal{V}$ and $\mathcal{C}$ cannot be achieved. The absolute "best" performance, in the sense of C-minimax control, is achieved when the two surfaces $\mathcal{V}$ and $\mathcal{C}$ coincide for <u>every</u> state $\underset{\sim}{x} \in D$.

10. A More General Class of C-Minimax Performance Indices

The particular class of C-minimax performance indices considered in the present study does not admit those cases in which the performance index $C(\cdot)$ is an <u>explicit</u> function of the control $u(t)$. On the other hand, there appear to be many practical situations where such a performance index is physically meaningful. It is of some interest, therefore, to consider the possibility of extending the techniques described above to the more general class of C-minimax performance indices of the form $C = C(\underset{\sim}{x}, u)$. Two methods for accomplishing this are described below.

One method which permits application of the C-minimax theory developed above to the case $C = C(\underset{\sim}{x}, u)$ consists of introducing the new state variable $x_{n+1} = u(t)$ and considering $w(t) = du(t)/dt$ as the new control variable. In this way, the additional state variable equation

$$\dot{x}_{n+1} = w(t) \tag{118}$$

can be appended to (2) and the resulting performance index $C(\underset{\sim}{x}, u) = C(\underset{\sim}{x}, x_{n+1})$ can be expressed in the form of (1). Application of this method is complicated by the necessity for selecting a suitable class of admissible control functions $w(t)$ [it may be necessary to introduce an artificial bound on admissible values of $w(t)$] and by the presence of "hard" (inequality) constraints imposed on the new state variable x_{n+1} through the original control set (5).

An alternative method for treating the case $C = C(\underset{\sim}{x}, u)$ consists of introducing, as before, the new state variable $x_{n+1} = u(t)$ and requiring that x_{n+1} satisfy the special state variable equation

$$\dot{x}_{n+1} = -k(x_{n+1} - w(t)) \tag{119}$$

where k is a positive scalar constant and w(t) is a new control function which belongs to the same class of admissible functions as the original control u(t). By this means, (119) can be appended to (2) and the performance index $C(\underset{\sim}{x}, x_{n+1})$ is reduced to the form (1). The exact solution to the original problem [i.e.: the attainment of the condition $x_{n+1}(t) = u(t) \equiv w(t)$] is obtained, through a limiting process, by letting $k \to \infty$ in (119). This method has the advantages that the new optimal control function w(t) is sought in the same class of functions as u(t) and, except for the initial condition requirement $x_{n+1}(t_o) \in U$, no inequality constraints are imposed on $x_{n+1}(t)$.

The two techniques described above can be used to study a variety of the cases $C(\underset{\sim}{x}, u)$. However, if $C(\cdot)$ has the special form $C = C(u)$, the C-minimax optimal control can be obtained by means of an essentially different method based on the theory of functional analysis. Some particular results which have been obtained by this method are described in [45]-[50].

11. Areas for Further Research

The deterministic C-minimax problem considered in the present study can be generalized to include dynamical systems with non-deterministic parameters. In particular, one might consider dynamical systems described by stochastic differential equations and replace $C(\underset{\sim}{x}(t))$ in the functional (1) by the expectation of $C(\underset{\sim}{x}(t))$. Stochastic optimal control problems of this type have apparently been studied relatively little [51], [52].

Appendix

An Integral Representation for the Maximum of a Non-Negative Continuous Function - An Elementary Proof

Theorem

Let $f(t)$, $t \in [t_o, T]$, be a real, non-negative definite, single-valued, bounded, and continuous scalar function defined on $[t_o, T]$. Suppose

$$\max_{t \in [t_o, T]} f(t) = M \tag{120}$$

Then

$$\lim_{\mu \to \infty} \left[\int_{t_o}^{T} [f(t)]^{\mu} \, dt \right]^{1/\mu} = M \tag{121}$$

Proof

Choose a constant $0 < \epsilon < M$ and define the comparison function $\hat{f}(t)$, $t \in [t_o, T]$, as follows

$$\hat{f}(t) = \begin{cases} 0, & \text{if } f(t) < (M - \epsilon) \\ M - \epsilon, & \text{if } f(t) \geq (M - \epsilon) \end{cases} \tag{122}$$

Clearly,

$$\hat{f}(t) \leq f(t) \leq M \qquad \forall\, t \in [t_o, T] \tag{123}$$

so that

$$\left[\int_{t_o}^{T} [\hat{f}(t)]^{\mu} \, dt \right]^{1/\mu} \leq \left[\int_{t_o}^{T} [f(t)]^{\mu} \, dt \right]^{1/\mu} \leq \left[\int_{t_o}^{T} M^{\mu} \, dt \right]^{1/\mu} \tag{124}$$

$$\forall\, t \in [t_o, T],\ \mu > 0$$

which can be written

$$(M - \epsilon)\left[\sum_{1}^{k} \delta_i(\epsilon)\right]^{1/\mu} \leq \left[\int_{t_o}^{T} [f(t)]^{\mu}\, dt\right]^{1/\mu} \leq M[T - t_o]^{1/\mu} \tag{125}$$

where the $\delta_i(\epsilon)$, $i = 1, \ldots, k$, denote the k positive intervals of time for which $f(t) \geq (M - \epsilon)$. Note that $0 < \sum_{1}^{k} \delta_i(\epsilon) \leq (T - t_o)$ for all $0 < \epsilon < M$. Taking the limit as $\mu \to \infty$ in (125) there obtains

$$(M - \epsilon) \leq \lim_{\mu \to \infty}\left[\int_{t_o}^{T} [f(t)]^{\mu}\, dt\right]^{1/\mu} \leq M \tag{126}$$

Since (126) is valid, in particular, for arbitrarily small positive ϵ the result (121) follows as $\epsilon \to 0$.

It can be shown [40] that the result (121), with M interpreted as the essential supremum of f(t), remains valid (almost everywhere on $[t_o, T]$) even when f(t) is only a measureable function.

Acknowledgments

The work reported herein was carried out as part of a research project initiated in 1964 and supported by the Aero-Astrodynamics Laboratory of the NASA George C. Marshall Space Flight Center, Huntsville, Alabama. The author is grateful to Mr. Clyde Baker, Mr. Judson Lovingood, and Dr. David Ford* of the Aero-Astrodynamics Laboratory for first bringing this problem to his attention and for providing many stimulating and informative discussions during this investigation. The author would also like to acknowledge the many valuable suggestions and helpful criticisms provided by Drs. W. M. Wonham, R. Isaacs, and T. F. Bridgland during the course of this study.

---

*Now with the Department of Mathematics, Emory University.

References Cited

1. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, The Mathematical Theory of Optimal Processes, John Wiley and Sons, Inc., New York, New York, 1962.

2. L. D. Berkovitz, "Variational Methods in Problems of Control and Programming," Jour. of Math. Anal. Appl., Vol. 3, No. 1, pp. 145-169, 1961.

3. R. E. Kalman, "Contributions to the Theory of Optimal Control," Boletin de la Sociedad Matematica Mexicana, pp. 111-119, 1960.

4. M. Dresher, Games of Strategy; Theory and Applications, Prentice-Hall, Inc. New Jersey, 1961.

5. D. Middleton, Introduction to Statistical Communication Theory, McGraw-Hill Book Co., New York, 1960.

6. D. Blackwell and M. A. Girshick, Theory of Games and Statistical Decisions, John Wiley and Sons, Inc., New York, 1954.

7. R. Isaacs, Differential Games, John Wiley and Sons, Inc., New York, 1965.

8. I. V. Girsanov, "Minimax Problems in the Theory of Diffusion Processes," Dokl. Akad. Nauk, Vol. 136, No. 4, 1960.

9. M. Yu. Gadzhiev, "An Application of Game Theory in Some Problems of Automatic Control," Parts 1 and 2, Avto. i Telemekhanika, Vol. 23, Nos. 8-9, 1962.

10. V. P. Grishin, "A Minimax Problem in the Theory of Analytical Design of Control Systems," Avto. i Telemakhanika, Vol. 25, No. 6. pp. 868-880, June 1964.

11. L. S. Gnoenskii. "On One Problem of Synthesizing a Control System," Dokl. Akad. Nauk, Vol. 155, No. 5, pp. 1022-1024, 1964.

12. Iu. A. Kochetkov, "Application of Pontryagin's Method to a Study of Minimax Problems of Control Processes," Tekhn. Kibernetika, No. 5, pp. 13-22, Sept.-October, 1965.

13. Y. C. Ho, A. E. Bryson and S. Baron, "Differential Games and Optimal Pursuit-Evasion Strategies," I.E.E.E. Trans. on Automatic Control, Vol. AC-10, No. 4, pp. 385-389, 1965.

14. J. Warga, "On a Class of Minimax Problems in the Calculus of Variations," Mich. Math. Jour., Vol. 12, No. 3, pp. 289-311, September 1965.

15. P. L. Tchebycheff, "Theorie des Mechanismes connus sous le Nom de Parallelogrammes," in Oeuvres de P. L. Tchebycheff, Vol. 1, Academie Imperiale des Sciences, St. Petersburg, pp. 109-143, 1899.

16. R. W. Hamming, Numerical Methods for Scientists and Engineers, McGraw-Hill Book Co., New York, Chapter 19, 1962.

17. E. Remez, "General Computational Methods for Chebyshev Approximation; Problems with Real Parameters Entering Linearly," Izdat. Akad. Nauk Ukrainsk. S.S.R., Kiev, 1957.

18. E. W. Cheney and A. A. Goldstein, "Tchebycheff Approximation and Related Extremal Problems," Jour. of Math. and Mech., Vol. 14, No. 1, pp. 87-98, 1965.

19. J. W. Young, "General Theory of Approximation by Functions Involving a Given Number of Arbitrary Parameters," Trans. Amer. Math. Soc., Vol. 8, pp. 331-344, July 1907.

20. C. Lanczos, Applied Analysis, Prentice Hall Book Co., New Jersey, 1956.

21. D. S. Carter, "A Minimum-Maximum Problem for Differential Expressions," Canadian Jour. of Math., Vol. 9, pp. 132-140, 1957.

22. R. E. Bellman, I. Glicksburg, and O. Gross, "Some Nonclassical Problems in the Calculus of Variations," Proc. Amer. Math. Soc., Vol. 7, No. 1, pp. 87-94, February 1956.

23. R. E. Bellman, "On the Application of the Theory of Dynamic Programming to the Study of Control Processes," Proc. of the Symposium on Nonlinear Circuit Analysis, April 1956, Poly. Press. Poly. Inst. of Brooklyn, Brooklyn, New York, pp. 199-213.

24. R. E. Bellman, "Notes on Control Processes -1; On the Minimum of Maximum Deviation," Quar. of Appl. Math., Vol. 14, No. 4, pp. 419-423, 1957.

25. R. E. Bellman, Dynamic Programming, Princeton Univ. Press, Princeton, New Jersey, 1957.

26. E. Sevin, "Min-Max Solutions for the Linear Mass-Spring System," A.S.M.E. Trans., Jour. of Appl. Mech., pp. 131-136, March 1957.

27. R. E. Bellman, I. Glicksburg, and O. Gross, "Some Aspects of the Mathematical Theory of Control Processes," RAND Corp. Rpt. No. R-313, January 1958.

28. A. Ya. Dubovitskii and A. A. Milyutin, "Extremum Problems with Constraints," Dokl. Akad. Nauk, No. 4, 1963.

29. _____, "Certain Optimality Problems for Linear Systems," Avto i Telemekhanika, Vol. 24, No. 12, pp. 1616-1625, December 1963.

30. J. Warga, "Minimizing Variational Curves Restricted to a Preassigned Set," Trans. of Amer. Math. Soc., Vol. 112, pp. 432-455, 1964.

31. _____, "Minimax Problems and Unilateral Curves in the Calculus of Variations," J. SIAM on Control, Ser. A, Vol. 3, No. 1, pp. 91-105, 1965.

32. L. W. Newstadt, "Optimal Control Problems as Extremal Problems in a Banach Space," Proc. of the Symp. on System Theory, April 1965, Poly. Press, Poly. Inst. of Brooklyn, Brooklyn, New York, pp. 215-224.

33. V. V. Guretskii and B. S. Fertman, "One Problem of Optimal Control," Priklad. Matemat. i Mekhan., Vol. 29, No. 5, pp. 946-949, 1965.

34. V. V. Guretskii, "On a Certain Optimal Control Problem," Izv. Akad. Nauk, Mekhanika, Vol. 1, No. 1, pp. 159-162, Jan.-Feb. 1965.

35. R. W. Bass and R. F. Webber, "Optimal Nonlinear Feedback Control Derived from Quartic and Higher-Order Performance Criteria," Proc. Third Inter. Congress I.F.A.C., 1966.

36. R. E. Bellman, Adaptive Control Processes; A Guided Tour, Princeton Univ. Press, Princeton, New Jersey, 1961.

37. J. P. LaSalle, "A Modern Verion of Liapunov's Stability Theory," to be published.

38. H. J. Kelley, R. E. Kopp, and H. G. Moyer, "Singular Extremals," Chapter 3 in Optimization - Theory and Applications; A Variational Approach, Vol. 1, ed. by G. Leitman, Academic Press, New York, 1966.

39. C. D. Johnson, "Singular Solutions in Problems of Optimal Control," Chapter 4 in Advances in Control Systems; Theory and Applications, Vol. 2, ed. by C. T. Leondes, Academic Press, New York, 1965.

40. K. Yosida, Functional Analysis, Academic Press, New York, 1965.

41. A. T. Fuller, "Relay Control Systems Optimized for Various Performance Criteria," Proc. First Inter. Congress I.F.A.C., Butterworths, London, England, pp. 510-519, 1961.

42. J. Andre and P. Seibert, "Piecewise Continuous Differential Equations," Boletin de la Sociedad Matematica Mexicana, Vol. 6, pp. 242-245, 1961.

43. W. M. Wonham and C. D. Johnson, "Optimal Bang-Bang Control with Quadratic Performance Index," A.S.M.E. Trans., Jour. of Basic Eng., Vol. 86, No. 1, pp. 107-115, March 1964.

44. C. D. Johnson and W. M. Wonham, "On a Problem of Letov in Optimal Control," A.S.M.E. Trans., Jour. of Basic Eng., Vol. 87, No. 1, pp. 81-89, March 1965.

45. N. N. Krasovskii, "On the Theory of Optimum Regulation," Avto. i Telemakhanika, Vol. 18, pp. 1005-1016, 1957.

46. L. W. Neustadt, "Minimum Effort Control," SIAM Jour. on Control, Vol. 1. pp. 16-31, 1962.

47. L. W. Neustadt, "Optimization, A Moment Problem, and Nonlinear Programming," SIAM Jour. on Control, Vol. 2, pp. 33-53, 1964.

48. V. I. Bondarenko, N. N. Krasovskii, Yu. M. Filimonov, "A Problem About Damping Linear Systems," Priklad. Matemat. i Mekhan., Vol. 29, No. 5, pp . 828-834, 1965.

49. N. N. Krasovskii, "A Problem About Damping Linear Systems with Minimum Control Intensity," Priklad. Matemat. i Mekhan., Vol. 29, No. 2, pp. 218-225, 1965.

50. A. A. Goldstein, "Minimizing Functionals on Normed-Linear Spaces," SIAM Jour. on Control, Vol. 4, No. 1, pp. 81-89, 1966.

51. M. Aoki, "On Minimum of Maximum Expected Deviation from an Unstable Equilibrium Position of a Randomly Perturbed Control System," IRE Trans. on Automatic Control, Vol. AC-7, pp. 1-12, March 1962.

52. R. Bellman, "On Minimizing the Probability of a Maximum Deviation," IRE Trans. on Automatic Control, Vol. AC-7, p. 45, July 1962.

III. Three Examples of Differential Game Problems in Optimal Control Theory

C. D. Johnson

The purpose of this study is to illustrate, by examples, the type of optimal control problems which can be solved by application of the theory of differential games. The method of solution is based on the combined application of the classical Hamilton-Jacobi-Carathéodory theory of a value function and the Principle of Optimality. The theoretical foundations for this method are omitted here since they are described, in detail, in the recently published treatise on differential games by R. Isaacs [1]. The reader is assumed to have some familiarity with this basic reference. The first example illustrates a differential game with quadratic performance index and hard inequality constraints on both controls. The second example illustrates a bounded control differential game which possesses a singular solution - a situation which is quite common in this class of problems. The last example is the differential game analogue of the classical optimal linear regulator problem.

Example 1 - A First Order System

In the class of piecewise continuous functions, find a pair of minimax controls $\{ u^\circ(t) , w^\circ(t) \}$ such that the minimax condition

$$J[u^\circ, w] \leq J[u^\circ, w^\circ] \leq J[u, w^\circ] \tag{1}$$

is satisfied for all admissible controls $u(t)$, $w(t)$

where

$$J[u, w] = \frac{1}{2} \int_0^T [x^2(t) + u^2(t)] \, dt \tag{2}$$

$$\dot{x} = -ax + u + w \tag{3}$$

and

$$x(0) = x_o$$

$$x(T) = 0$$

$$T \text{ - unrestricted}$$

$$|w(t)| \leq M$$

$$|u(t)| \leq N \qquad (N > M > 0) \tag{4}$$

In (3), a is a real scalar constant and u (t), w(t) are real scalar functions of time.

Proceeding in the spirit of differential game theory, we define the value of the game payoff V (x) as

$$V(x) = \min_{u \in U} \max_{w \in W} J[u, w] \quad , x_o = x \tag{5}$$

where U and W represent the set of admissible values for the controls u and w respectively. It is assumed here, and in all examples which follow, that the game has a proper saddle point so that

$$\min_{u \in U} \max_{w \in W} J[u,w] = \max_{w \in W} \min_{u \in U} J[u,w] \tag{6}$$

Then, the value V (x) satisfies the Hamilton-Jacobi equation

$$ax \frac{\partial V}{\partial x} - M \left|\frac{\partial V}{\partial x}\right| + N \frac{\partial V}{\partial x} \text{ sat } \left(\frac{1}{N} \frac{\partial V}{\partial x}\right) - \frac{N^2}{2} \text{ sat}^2 \left(\frac{1}{N} \frac{\partial V}{\partial x}\right) - \frac{1}{2} x^2 = 0 \tag{7}$$

and the minimax controls $u^\circ(t)$, $w^\circ(t)$ are given by

$$u^\circ(t) = - N \text{ sat}\left[\frac{1}{N} \frac{\partial V(x(t))}{\partial x}\right]$$

$$w^\circ(t) = M \text{ sgn}\left[\frac{\partial V(x(t))}{\partial x}\right] \tag{8}$$

The pair $\{ u^\circ(t), w^\circ(t) \}$ can be determined by solving (7) directly, or, alternatively, by solving for the characteristic strips of (7). Here, we solve (7) directly.

Assume that

$$0 < \frac{\partial V}{\partial x} \leq N \tag{9}$$

Then, $w^\circ = +M$ and, noting the required boundary conditions (4), it follows from (7) that

$$\frac{\partial V}{\partial x} = -(ax - M) + (\operatorname{sgn} x)\sqrt{(ax - M)^2 + x^2} \tag{10}$$

From (9), it is clear that (10) is valid in the region

$$0 < x \leq +aN + \sqrt{N^2(a^2+1) - 2NM} \tag{11}$$

In a similar manner, it is found that in the region

$$-N \leq \frac{\partial V}{\partial x} < 0 \tag{12}$$

(7) is satisfied by

$$\frac{\partial V}{\partial x} = -(ax + M) + (\operatorname{sgn} x)\sqrt{(ax + M)^2 + x^2} \tag{13}$$

which holds in the region

$$-aN - \sqrt{N^2(a^2+1) - 2NM} \leq x < 0 \tag{14}$$

It may be noted that (11) and (4) define real upper and lower boundaries, respectively, only if

$$N \geq \frac{2M}{a^2+1} \tag{15}$$

Finally, from (10) and (13) the minimax optimal controls (8) can be written in the state variable feedback form

$$u^\circ(x) = N \text{ sat } N^{-1}\left[ax - (\text{sgn } x)\,(M + \sqrt{(ax - M \text{ sgn } x)^2 + x^2})\right]$$

$$w^\circ(x) = M \text{ sgn } x \tag{16}$$

Example 2 - A Second Order System

In the class of piecewise continuous functions, find a pair of minimax controls $\{u^\circ(t), w^\circ(t)\}$ such that (1) is satisfied for all admissible controls u(t), w(t) where

$$J[u,w] = \tfrac{1}{2}\int_0^T [\, x_1^2(t) + x_2^2(t)]\, dt \tag{17}$$

$$\dot{x}_1 = x_2 + w \tag{18}$$

$$\dot{x}_2 = u \tag{19}$$

$$\underset{\sim}{x}(0) = \underset{\sim}{x}_o$$

$$x_1(T) = 0$$

$$x_2(T) = \text{unrestricted}$$

$$T \text{ - unrestricted}$$

$$|w(t)| \leq M$$

$$|u(t)| \leq N \tag{20}$$

In this case, the value $V(\underset{\sim}{x})$ of the game satisfies the Hamilton-Jacobi equation

$$x_2 \frac{\partial V}{\partial x_1} + M \left|\frac{\partial V}{\partial x_1}\right| - N \left|\frac{\partial V}{\partial x_2}\right| + \frac{1}{2} x_1^2 + \frac{1}{2} x_2^2 = 0 \tag{21}$$

and the minimax controls are given by

$$u^\circ(t) = -N \operatorname{sgn} \frac{\partial V}{\partial x_2}(\underset{\sim}{x}(t))$$

$$w^\circ(t) = -M \operatorname{sgn} \frac{\partial V}{\partial x_1}(\underset{\sim}{x}(t)) \tag{22}$$

In this example, it is not so easy to solve (21) directly. However, if there exists a singular solution [2], it may be possible to effectively solve for $\{u^\circ, w^\circ\}$ by tracing out the characteristic strips of (21) starting on the singular manifold.

Proceeding as in [2] we find that there are two possible singular solutions. The particular singular condition

$$\frac{\partial V}{\partial x_1}(\underset{\sim}{x}(t)) \equiv 0 \tag{23}$$

implies that

$$x_1(t) \equiv o$$

$$\frac{d}{dt}\left|x_2(t)\right| = \min \tag{24}$$

This yields the minimin solution; i.e., both controls "working together" to achieve an absolute minimum of (17). This solution is of no interest in the present study.

The other singular condition

$$\frac{\partial V(\underset{\sim}{x}(t))}{\partial x_2} \equiv 0 \tag{25}$$

implies that

$$u_s^o(t) = x_1(t) \quad , \qquad |x_1(t)| \leq N$$

$$w_s^o(t) = -M \operatorname{sgn} x_2(t) \tag{26}$$

From (21), (26) and the given terminal and constraint conditions (20), the singular manifold for this case is found to be described by the expression

$$x_2 = \operatorname{sgn} x_2 \left[M + \sqrt{M^2 + x_1^2} \right] \tag{27}$$

which is defined over the regions

$$\begin{aligned} -N \leq x_1 \leq 0 \quad & \text{if } x_2 > 0 \\ 0 \leq x_1 \leq +N \quad & \text{if } x_2 < 0 \end{aligned} \tag{28}$$

There is some reason to suspect that this singular solution does play a role in the present minimax problem. For instance, as $M \to 0$ (i.e. as the "player" w(t) becomes less and less potent) the singular manifold (27), (28) degenerates to

$$x_2 = -x_1 \quad , \quad |x_1| \leq N \tag{29}$$

However, the solution to the case $M = 0$ is already available [3] and it is true that, for that problem, the singular manifold (29) is optimal.[1] A rigorous proof of the minimax optimality of the singular manifold (27), (28) can be established by showing that the derived function $J[u, w]$, $\underset{\sim}{x}_o = \underset{\sim}{x}$, evaluated throughout an ϵ neighborhood of the singular manifold (27), (28), does indeed satisfy the Hamilton-Jacobi equation (21). This process usually involves somewhat lengthy calculations and will not be attempted here. Hereafter, we proceed on the assumption that this proof has been established.

On the singular manifold (27), (28) the values of $\partial V/\partial x_1$, $\partial V/\partial x_2$ are given by

---

1. In [3], the terminal condition is specified as $x_1(T) = x_2(T) = 0$.

$$\left.\frac{\partial V}{\partial x_1}\right|_s = -x_2$$

$$\left.\frac{\partial V}{\partial x_2}\right|_s = 0 \qquad (30)$$

Also, the general expressions for the characteristic strips of (21) are (setting: $p_1 = -\partial V/\partial x_1$, $p_2 = -\partial V/\partial x_2$; see [4])

$$\dot{p}_1 = x_1$$

$$\dot{p}_2 = -p_1 + x_2 \qquad (31)$$

Thus, by computing the reverse time solutions to (18), (19) and (31), using the initial conditions (30) with $u^o(T) = \pm N$, we may "flood" the $x_1 - x_2$ space with minimax "optimal" trajectories which have one end lying on the singular manifold. This process does not completely "cover" the $x_1 - x_2$ space with trajectories. The voids are filled by reverse time trajectories of (18), (19) and (31) which start on the specified terminal manifold: $x_1(T) = 0$. In this case, we replace the initial conditions (30) by corresponding values computed from the transversality condition which yields

$$\left.\frac{\partial V}{\partial x_1}\right|_{t=T} = \begin{cases} \dfrac{x_2^2}{2(M - x_2)} \; ; & x_2 > M \\ \dfrac{-x_2^2}{2(M + x_2)} \; ; & x_2 < -M \end{cases}$$

$$\left.\frac{\partial V}{\partial x_2}\right|_{t=T} = 0 \qquad (32)$$

It may be noted that there are no minimax optimal trajectories which terminate on the sector of the terminal manifold defined by: $x_1 = 0$, $|x_2| < M$. This result is characteristic of "minimax" differential games.

As the flood paths described above are traced out in backward time, the set of points $\{\underset{\sim}{x}_1(t)\}_{p_2}$, at which $p_2(t) = 0$, build up a "switching manifold" across which $u^o(t)$ switches from $\pm N$ to $\mp N$.[2] This u^o switching boundary joins with the singular manifold (27), (28). This completes the solution to this minimax problem. The results are summarized in Figure 1.

Example 3 - An n^{th} Order System

In the class of piecewise continuous functions, find a pair of minimax controls $\{u^o(t), w^o(t)\}$ such that (1) is satisfied for all admissible controls $u(t)$, $w(t)$, where[3]

$$J[u,w] = \frac{1}{2}\int_o^T [<\underset{\sim}{x}(t), \underset{\sim}{Q}\underset{\sim}{x}(t)> + c^2u^2(t) - r^2w^2(t)]\,dt \tag{33}$$

$$\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x} + u\underset{\sim}{f} + w\underset{\sim}{b} \tag{34}$$

and

$$\begin{aligned} \underset{\sim}{x}(0) &= \underset{\sim}{x}_o \\ \underset{\sim}{x}(T) &= \underset{\sim}{0} \\ T &\text{ - unrestricted} \end{aligned} \tag{35}$$

In (33), $\underset{\sim}{x}$ is a real n-vector, $\underset{\sim}{Q}$ is a real $n \times n$, constant, positive definite matrix, and c and r are real, non-zero scalars. In (34), $\underset{\sim}{A}$ is a real $n \times n$, constant matrix and $\underset{\sim}{f}$ and $\underset{\sim}{b}$ are real n-vectors. In this example, the values of $u(t)$, $w(t)$ are not restricted.

---

[2] It turns out that the function $p_1(t)$ has no zeros along the flood paths. Thus, $w^o(t) = M\ sgn\ x_1(t)$.

[3] Here, $<\underset{\sim}{x}, \underset{\sim}{y}>$ denotes the inner product of $\underset{\sim}{x}$ and $\underset{\sim}{y}$.

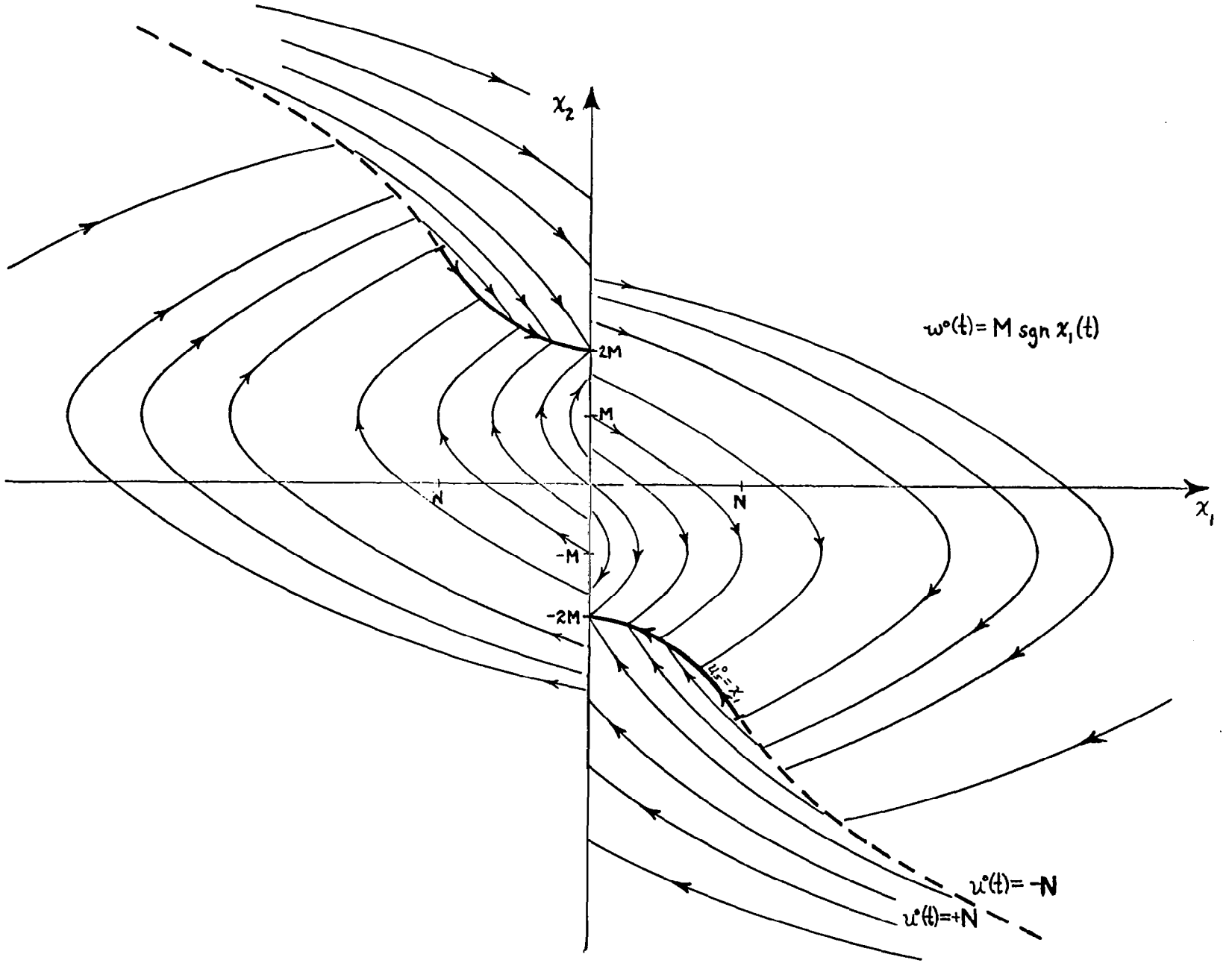

Figure 1 -- Field of Minimax Optimal Trajectories for Example 2.

We need to assume this game is proper. That is, for any initial condition $\underset{\sim}{x}_o$ and any admissible w(t) there is at least one admissible u(t) which can satisfy the terminal conditions (35). This requirement is the differential game analog of controllability. Under these assumptions, the value of the game $V(\underset{\sim}{x})$ satisfies the Hamilton-Jacobi equation

$$\langle \underset{\sim}{\nabla} V, \underset{\sim}{A}\underset{\sim}{x} \rangle - \frac{1}{2} c^{-2} \langle \underset{\sim}{\nabla} V, \underset{\sim}{f} \rangle^2 + \frac{1}{2} r^{-2} \langle \underset{\sim}{\nabla} V, \underset{\sim}{b} \rangle^2 + \frac{1}{2} \langle \underset{\sim}{x}, \underset{\sim}{Q}\underset{\sim}{x} \rangle = 0 \tag{36}$$

where $\underset{\sim}{\nabla} V = \left(\frac{\partial V}{\partial x_1}, \ldots, \frac{\partial V}{\partial x_n}\right)$, and the minimax controls are given by

$$u^o(t) = -c^{-2} \langle \underset{\sim}{\nabla} V(\underset{\sim}{x}(t)), \underset{\sim}{f} \rangle$$

$$w^o(t) = r^{-2} \langle \underset{\sim}{\nabla} V(\underset{\sim}{x}(t)), \underset{\sim}{b} \rangle \tag{37}$$

In this particular example, the direct solution of (36) can be carried out by assuming the solution $V(\underset{\sim}{x})$ is a quadratic form in $\underset{\sim}{x}$. More precisely, let

$$V(\underset{\sim}{x}) = \frac{1}{2} \langle \underset{\sim}{x}, \underset{\sim}{M}\underset{\sim}{x} \rangle, \tag{38}$$

where $\underset{\sim}{M}$ is a real, constant, positive definite,[4] n x n matrix. Then, (38) is a solution of (36) if $\underset{\sim}{M}$ satisfies the matrix equation

$$\underset{\sim}{M}\underset{\sim}{A} + \underset{\sim}{A}'\underset{\sim}{M} + \underset{\sim}{M}(r^{-2}\underset{\sim}{b}\underset{\sim}{b}' - c^{-2}\underset{\sim}{f}\underset{\sim}{f}')\underset{\sim}{M} + \underset{\sim}{Q} = 0 \tag{39}$$

Under appropriate conditions on the matrix $[r^{-2}\underset{\sim}{b}\underset{\sim}{b}' - c^{-2}\underset{\sim}{f}\underset{\sim}{f}']$, (39) has a unique *real, positive definite solution* $\underset{\sim}{M}$. *Using this solution, the minimax controls* (37) can be written in the state variable feedback form

$$u^o(\underset{\sim}{x}) = -c^{-2} \langle \underset{\sim}{M}\underset{\sim}{x}, \underset{\sim}{f} \rangle$$

$$w^o(\underset{\sim}{x}) = r^{-2} \langle \underset{\sim}{M}\underset{\sim}{x}, \underset{\sim}{b} \rangle \tag{40}$$

---

[4] It is clear from (33) that $V(\underset{\sim}{x})$ must be positive definite.

It may be noted from (40) that $u^o(\underset{\sim}{x})$ and $w^o(\underset{\sim}{x})$ are linear functions of the state variables $x_1, \ldots, x_n$.

The case of the above problem with bounded controls, $|u(t)| \leq N$, $|w(t)| \leq \bar{M}$, can be solved, in principle, by employing the techniques which were used to solve the Problem of Letov [5].

References Cited

1. R. Isaacs, "Differential Games", (book) John Wiley and Sons, Pubs. 1965.

2. C. D. Johnson and J. E. Gibson, "Singular Solutions in Problems of Optimal Control," IEEE Transactions on Automatic Control, Vol. AC-8, No. 1, January 1963.

3. W. M. Wonham and C. D. Johnson, "Optimal Bang-Bang Control with Quadratic Performance Index," Proceedings, Fourth Joint Automatic Control Conference, Minneapolis, Minn., June 1963. Also, A.S.M.E. Transactions, Journal of Basic Engineering, Vol. 86, Series D, No. 1, March 1964, pp. 107-115.

4. C. D. Johnson and J. E. Gibson, "Optimal Control with Quadratic Performance Index and Fixed Terminal Time," IEEE Transactions on Automatic Control, Vol. AC-9, No. 4, October 1964.

5. C. D. Johnson and W. M. Wonham, "On a Problem of Letov in Optimal Control," Proceedings, Fifth Joint Automatic Control Conference, Stanford California, June 1964, pp. 317-325. Also, A.S.M.E. Transactions, Journal of Basic Engineering, March 1965.

IV. A Note on the Transformation to Canonical (Phase-Variable) Form

C. D. Johnson and W. M. Wonham

Introduction

Consider the control system defined by

$$\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x} + u(t)\,\underset{\sim}{f}, \qquad \cdot = \frac{d}{dt} \tag{1}$$

where $\underset{\sim}{x} = (x_1, \ldots, x_n)$ is the state vector of the plant, $\underset{\sim}{A}$ is an (n x n) constant matrix, $\underset{\sim}{f} = (f_1, \ldots, f_n)$ is a constant n vector and u(t) is the scalar control function.

It is well known[1, 2] that, if the pair $(\underset{\sim}{A}, \underset{\sim}{f})$ is controllable, there exists a nonsingular linear transformation

$$\underset{\sim}{x} = \underset{\sim}{K}\underset{\sim}{y}$$

which reduces (1) to the canonical (phase-variable) form

$$\dot{\underset{\sim}{y}} = \underset{\sim}{A}_o\underset{\sim}{y} + u(t)\,\underset{\sim}{f}_o$$

where

$$\underset{\sim}{A}_o = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ & & & \ddots & \vdots \\ & & & & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \end{bmatrix}, \quad \underset{\sim}{f}_o = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{2}$$

This work was supported in part under Contracts No. NAS8-11231, AF 49 (638) - 1206 and AF 33 (657) - 8559.

C. D. Johnson is with the Department of Electrical Engineering, University of Alabama, Huntsville, Alabama.

W. M. Wonham is with the Center for Control Theory, Research Institute for Advanced Studies, Baltimore, Maryland.

1 R. E. Kalman, "Mathematical description of linear dynamical systems," SIAM J. on Control, ser. A, vol. 1, pp. 152-192; 1963.

2 R. E. Kalman, "When is a linear control system optimal?" ASME Trans., J. of Basic Engrg, Vol. 86, pp. 51 - 60, March, 1964.

In a previous paper,[3] a general procedure for obtaining the matrix $\underset{\sim}{K}$ was described but no explicit expressions for $\underset{\sim}{K}$ were given. In the present note, an expression for $\underset{\sim}{K}$ is derived in terms of the Vandermonde matrix and a modal matrix of $\underset{\sim}{A}$, on the assumption that the eigenvalues of $\underset{\sim}{A}$ are distinct.

Main Result

Let the eigenvalues of $\underset{\sim}{A}$ be $\lambda_1, \ldots, \lambda_n$ and let $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_n$ be a corresponding set of eigenvectors. We recall that the Vandermonde matrix of $\underset{\sim}{A}$ is the matrix $\underset{\sim}{M}_v$ with elements

$$(\underset{\sim}{M}_v)_{ij} = \lambda_j^{i-1} \quad (i, j = 1, \ldots, n). \tag{3}$$

The modal matrix $\underset{\sim}{M}$ is

$$\underset{\sim}{M} = [\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_n]$$

and has the property

$$\underset{\sim}{M}^{-1}\underset{\sim}{A}\underset{\sim}{M} = \underset{\sim}{\Lambda} \tag{4}$$

where

$$\underset{\sim}{\Lambda} = \text{diag } [\lambda_1, \ldots, \lambda_n].$$

The pair $(\underset{\sim}{A}, \underset{\sim}{f})$ is controllable if the vectors $\underset{\sim}{f}, \underset{\sim}{A}\underset{\sim}{f}, \ldots, \underset{\sim}{A}^{n-1}\underset{\sim}{f}$, are linearly independent.

It will be shown that the required transformation is

$$\underset{\sim}{K} = \underset{\sim}{M}\underset{\sim}{B}^{-1}\underset{\sim}{M}_v^{-1} \tag{5}$$

where $\underset{\sim}{B}$ is a diagonal matrix defined in the following theorem.

Theorem

Let $\underset{\sim}{A}, \underset{\sim}{f}$ be real and let $(\underset{\sim}{A}, \underset{\sim}{f})$ be controllable. Suppose the eigenvalues $\lambda_1, \ldots, \lambda_n$ of $\underset{\sim}{A}$ are distinct. Let $\underset{\sim}{M}$ be the modal matrix of $\underset{\sim}{A}$ and let $\underset{\sim}{M}_v$ be the Vandermonde matrix of $\underset{\sim}{A}$. Then a nonsingular diagonal matrix $\underset{\sim}{B}$ exists such that

$$\underset{\sim}{M}_v\underset{\sim}{B}\underset{\sim}{M}^{-1}\underset{\sim}{f} = \underset{\sim}{f}_o. \tag{6}$$

---

[3] W. M. Wonham and C. D. Johnson, "Optimal bang-bang control with quadratic performance index," ASME Trans., J. of Basic Engrg., vol. 86, pp.107-115; March, 1964.

Moreover, the matrix

$$\underset{\sim}{A}_o = \underset{\sim}{M}_v \underset{\sim}{B} \underset{\sim}{M}^{-1} \underset{\sim}{A} \underset{\sim}{M} \underset{\sim}{B}^{-1} \underset{\sim}{M}_v^{-1} \tag{7}$$

is of canonical form (2) with real elements. If

$$\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x} + u(t)\,\underset{\sim}{f}, \tag{8}$$

then the transformation

$$\underset{\sim}{x} = \underset{\sim}{M}\underset{\sim}{B}^{-1}\underset{\sim}{M}_v^{-1}\underset{\sim}{y}$$

reduces (8) to

$$\dot{\underset{\sim}{y}} = \underset{\sim}{A}_o\underset{\sim}{y} + u(t)\underset{\sim}{f}_o.$$

Proof

The transformation $\underset{\sim}{x} = \underset{\sim}{M}\underset{\sim}{z}$ reduces (8) to

$$\dot{\underset{\sim}{z}} = \underset{\sim}{\Lambda}\underset{\sim}{z} + u(t)\underset{\sim}{c}$$

where

$$\underset{\sim}{c} = \underset{\sim}{M}^{-1}\underset{\sim}{f}. \tag{9}$$

Since $(\underset{\sim}{A}, \underset{\sim}{f})$ is controllable, the vectors $\underset{\sim}{c}, \underset{\sim}{\Lambda}\underset{\sim}{c}, \ldots, \underset{\sim}{\Lambda}^{n-1}\underset{\sim}{c}$ are linearly independent; it follows, since the λ_i are distinct, that the components c_i of $\underset{\sim}{c}$ are all nonzero.

Let $\underset{\sim}{C} = \operatorname{diag}(c_1, \ldots, c_n)$; then $\underset{\sim}{M}_v\underset{\sim}{C}$ is nonsingular. Define

$$\begin{aligned}\underset{\sim}{b} &= (b_1, \ldots, b_n)' \\ &= (\underset{\sim}{M}_v\underset{\sim}{C})^{-1}\underset{\sim}{f}_o;\end{aligned} \tag{10}$$

and let

$$\underset{\sim}{B} = \operatorname{diag}(b_1, \ldots, b_n). \tag{11}$$

By (9) and (10), $\underset{\sim}{M}_v\underset{\sim}{B}\underset{\sim}{M}^{-1}\underset{\sim}{f} = \underset{\sim}{M}_v\underset{\sim}{B}\underset{\sim}{c} = \underset{\sim}{M}_v\underset{\sim}{C}\underset{\sim}{b} = \underset{\sim}{f}_o$, so that (6) is satisfied.

The b_i are all nonzero. For suppose $b_i = 0$ if $i = i_1, \ldots, i_k$. By (10), $\underset{\sim}{M}_v\underset{\sim}{C}\underset{\sim}{b} = \underset{\sim}{f}_o$. Let $\overline{\underset{\sim}{M}_v\underset{\sim}{C}}$ be the matrix obtained from $\underset{\sim}{M}_v\underset{\sim}{C}$ by deleting the last k rows and $i_1, \ldots, i_k$th columns. Then $\overline{\underset{\sim}{M}_v\underset{\sim}{C}}$ is nonsingular; the last equation implies $b_i = 0$ for the remaining b_i's; and this contradicts (6). Hence $\underset{\sim}{B}$ is nonsingular, $\underset{\sim}{A}_o$ is well defined by (7); and (4), (7) give

$$\underset{\sim}{A}_o = \underset{\sim}{M}_v\underset{\sim}{\Lambda}\,\underset{\sim}{M}_v^{-1}. \tag{12}$$

Writing $[\underset{\sim}{M}_v^{-1}]_{ij} = \mu_{ij}$ and using (3) and (12),

$$[\underset{\sim}{A}_o]_{ij} = \sum_{r=1}^{n} \sum_{s=1}^{n} \lambda_r^{i-1} \lambda_r \delta_{rs} \mu_{sj}$$

$$= \begin{cases} \delta_{i+1,j} & (i = 1, \ldots, n-1;\ j = 1, \ldots, n) \\ \sum_{r=1}^{n} \lambda_r^n \mu_{rj} & (i = n;\ j = 1, \ldots, n). \end{cases}$$

Thus $\underset{\sim}{A}_o$ is of form (2) and, since $\underset{\sim}{A}$, $\underset{\sim}{A}_o$ have the same eigenvalues,

$$\lambda^n - \sum_{i=1}^{n} a_i \lambda^{i-1} \equiv (\lambda - \lambda_1) \ldots (\lambda - \lambda_n),$$

so that the a_i are real if $\underset{\sim}{A}$ is real. This completes the proof. The transfomation to canonical form is illustrated in Fig. 1.

An Example

As an application of the theorem, consider a third-order system (1) with

$$\underset{\sim}{A} = \begin{bmatrix} -2 & -1 & 1 \\ 1 & 0 & 1 \\ -1 & 0 & 1 \end{bmatrix}; \qquad \underset{\sim}{f} = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}. \tag{13}$$

The eigenvalues of $\underset{\sim}{A}$ are

$$\lambda_1 = 1$$

$$\lambda_2 = -1 + j$$

$$\lambda_3 = -1 - j \qquad (j = \sqrt{-1})$$

and the modal matrix of A is

$$\underset{\sim}{M} = \begin{bmatrix} 0 & 5 & 5 \\ 1 & -3 - 4j & -3 + 4j \\ 1 & 2 + j & 2 - j \end{bmatrix}.$$

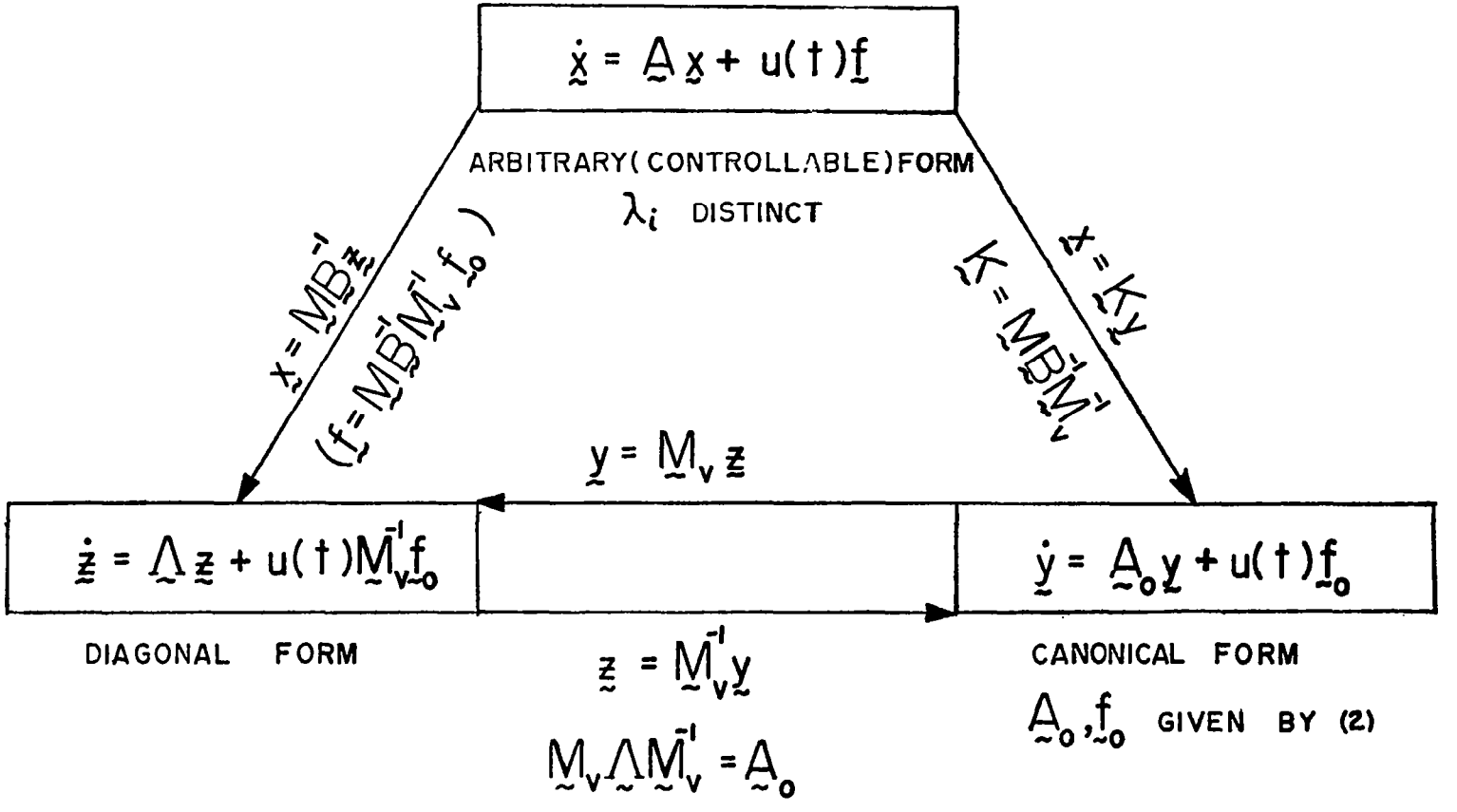
$\dot{x} = A x + u(t) f$
ARBITRARY (CONTROLLABLE) FORM
λ_i DISTINCT
$x = M B z$
$(f = M B M_v^{-1} f_o)$
$K = M B M_v^{-1}$
$x = K y$
$y = M_v z$
$\dot{z} = \Lambda z + u(t) M_v^{-1} f_o$
$\dot{y} = A_o y + u(t) f_o$
DIAGONAL FORM
$z = M_v^{-1} y$
CANONICAL FORM
A_o, f_o GIVEN BY (2)
$M_v \Lambda M_v^{-1} = A_o$

FIGURE 1

From (3), the Vandermonde matrix of $\underset{\sim}{A}$ is

$$\underset{\sim}{M}_v = \begin{bmatrix} 1 & 1 & 1 \\ 1 & -1+j & -1-j \\ 1 & -2j & 2j \end{bmatrix}.$$

With these values of $\underset{\sim}{f}$, $\underset{\sim}{M}$ and $\underset{\sim}{M}_v$, the matrix $\underset{\sim}{B}^{-1}$ is found from (6) to be

$$\underset{\sim}{B}^{-1} = \frac{1}{5}\begin{bmatrix} 20 & 0 & 0 \\ 0 & 1-3j & 0 \\ 0 & 0 & 1+3j \end{bmatrix}.$$

The matrix $\underset{\sim}{K}$ is then given by (5),

$$\underset{\sim}{K} = \begin{bmatrix} 0 & -1 & 1 \\ 0 & 3 & 1 \\ 2 & 1 & 1 \end{bmatrix},$$

and the transformation $\underset{\sim}{x} = \underset{\sim}{K}\underset{\sim}{y}$ takes (13) into the canonical (phase-variable) form

$$\underset{\sim}{A}_o = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 2 & 0 & -1 \end{bmatrix}; \qquad \underset{\sim}{f}_o = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

V. Another Note on the Transformation to Canonical (Phase-Variable) Form

C. D. Johnson* W. M. Wonham**

Introduction

The problem of determining a nonsingular linear transformation $\underset{\sim}{x} = \underset{\sim}{K}\underset{\sim}{y}$ which will take an arbitrary, completely controllable, single-input, time-invariant linear dynamical system

$$\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x} + u(t)\underset{\sim}{f} \qquad (\dot{} = d/dt) \tag{1}$$

into the canonical (phase-variable) form

$$\dot{\underset{\sim}{y}} = \underset{\sim o}{A}\,\underset{\sim}{y} + u(t)\underset{\sim o}{f} \tag{2}$$

where

$$\underset{\sim o}{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 1 & \cdots & 0 \\ \vdots & & & & \ddots & \vdots \\ & & & & 1 & 0 \\ 0 & & & & 0 & 1 \\ a_1 & a_2 & \cdots & & a_{n-1} & a_n \end{bmatrix}; \quad \underset{\sim o}{f} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{3}$$

was posed and completely solved in [1]. Since the publication of that result various features of the transformation method described therein have been rediscovered and published as "new results," [2], [3], [4]. The purpose of this *note is to point out, in more detail, some of the inherent computational features* of the transformation method originally described in [1].

This work was supported by the National Aeronautics and Space Administration under Contract No. NAS8-11231 and Grant No. NsG-381.

*Department of Electrical Engineering, University of Alabama in Huntsville, Huntsville, Alabama.

** Center for Dynamical Systems, Division of Applied Mathematics, Brown University, Providence, Rhode Island.

A General Method for Determining the Matrix K

In Appendix 1 of [1], it was shown that, given an arbitrary completely controllably pair $(\underset{\sim}{A}, \underset{\sim}{f})$, the required transformation matrix $\underset{\sim}{K}$ could be effectively computed by the following procedure.

1) Form the controllability matrix[1] $\underset{\sim}{H}$ defined by

$$\underset{\sim}{H} = [\underset{\sim}{f}, \underset{\sim}{A}\underset{\sim}{f}, \underset{\sim}{A}^2\underset{\sim}{f}, \ldots, \underset{\sim}{A}^{n-1}\underset{\sim}{f}] \tag{4}$$

2) Compute $\underset{\sim}{H}^{-1}$

3) Compute the coefficients a_i of the characteristic polynomial of $\underset{\sim}{A}$ [i.e., last row of $\underset{\sim}{A}_o$] by the following rule[2]

$$a_i = \langle \bar{h}_i, \underset{\sim}{A}^n \underset{\sim}{f} \rangle \qquad i = 1, \ldots, n \tag{5}$$

where $\bar{h}_i$ is the i^{th} row of $\underset{\sim}{H}^{-1}$.

4) Form the symmetric matrix $\underset{\sim}{L}$ defined by

$$L = \begin{bmatrix} -a_2 & -a_3 & -a_4 & \cdots & -a_{n-1} & -a_n & 1 \\ -a_3 & -a_4 & & \cdots & -a_n & 1 & 0 \\ -a_4 & & & & 1 & 0 & 0 \\ \vdots & & & \iddots & & & \vdots \\ -a_{n-1} & -a_n & & & & & \\ -a_n & 1 & & & & & \\ 1 & 0 & & \cdots & & & 0 \end{bmatrix} \tag{6}$$

5) Set

$$\underset{\sim}{K} = \underset{\sim}{H}\underset{\sim}{L} \, . \tag{7-a}$$

This gives, by direct calculation,

$$\underset{\sim}{K} = [\underset{\sim}{k}_1, \underset{\sim}{k}_2, \ldots, \underset{\sim}{k}_n] \tag{7-b}$$

---

1. It is recalled that the pair $(\underset{\sim}{A}, \underset{\sim}{f})$ is completely controllable if and only if rank $\underset{\sim}{H} = n$.
2. $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle$ denotes the scalar product of $\underset{\sim}{x}$ and $\underset{\sim}{y}$.

where

$$\underset{\sim}{k}_r = -\sum_{s=1}^{n-r} a_{r+s}\, \underset{\sim}{A}^{s-1}\underset{\sim}{f} + \underset{\sim}{A}^{n-r}\underset{\sim}{f}, \qquad r=1,2,\ldots,n-1 \qquad (7\text{-}c)$$

$$\underset{\sim}{k}_n = \underset{\sim}{f}$$

It may be noted that the vectors $\underset{\sim}{k}_r$ in the first equation of (7-c) satisfy the recursion equation

$$\underset{\sim}{k}_r = -a_{r+1}\underset{\sim}{f} + \underset{\sim}{A}\underset{\sim}{k}_{r+1} \qquad r=1,2,\ldots,n-1 \qquad (7\text{-}d)$$

The constructive procedure described above is so designed to provide a "built in" check on the validity of the assumption that the pair $(\underset{\sim}{A},\underset{\sim}{f})$ is completely controllable [eg., Step 2)]. Moreover, the apparent excess of information generated in Step 2) [to check controllability it is only necessary to compute $|\underset{\sim}{H}|$] is effectively used in Step 3) to evaluate the characteristic polynomial of $\underset{\sim}{A}$ and thereby avoid the necessity of directly expanding the determinant $|\underset{\sim}{A} - \lambda\underset{\sim}{I}|$.

Of course, if one knows a priori that the pair $(\underset{\sim}{A},\underset{\sim}{f})$ is completely controllable then Step 2) can be ignored and the elements a_i, $i = 1, \ldots, n$ in Step 3) can be determined alternatively by the more common procedure of evaluating the characteristic polynomial of $\underset{\sim}{A}$

$$|\underset{\sim}{A} - \lambda\underset{\sim}{I}| = \lambda^n - \sum_{1}^{n} a_i \lambda^{i-1} . \qquad (8)$$

A possible disadvantage of this alternative, albeit more direct, procedure is that by it the matrix $\underset{\sim}{K}$ defined in (7) can be formally constructed even when the system (1) is not completely controllable. The possibility of using the set of vectors $\underset{\sim}{k}_1, \ldots, \underset{\sim}{k}_n$ defined in (7-c), as a basis for the canonical (phase-variable) form was pointed out in [5].[3] This result has also been described in a recent textbook [6].

In practical applications of the transformation to phase-variable form one usually needs both $\underset{\sim}{K}$ and $\underset{\sim}{K}^{-1}$. Using the procedure outlined above, we have

$$\underset{\sim}{K}^{-1} = \underset{\sim}{L}^{-1}\,\underset{\sim}{H}^{-1} \qquad (9)$$

---

3. Equations (7-c) coincide with the results given in [5] if the term $a_o\underset{\sim}{g}$ in [5] is replaced by $a_1\underset{\sim}{g}$. This is apparently a typographical error.

The matrix $\underset{\sim}{H}^{-1}$ has already been computed in Step 2). The expression for $\underset{\sim}{L}^{-1}$ is easily computed[4] to be

$$\underset{\sim}{L}^{-1} = \begin{bmatrix} 0 & 0 & \cdots & 0 & 1 \\ 0 & & & 0 & 1 & b_n \\ \vdots & & & 1 & b_n & b_{n-1} \\ \vdots & & \iddots & \iddots & \iddots & \vdots \\ 0 & 0 & 1 & b_n & \iddots & \vdots \\ 0 & 1 & b_n & b_{n-1} & \cdots & b_3 \\ 1 & b_n & b_{n-1} & \cdots & b_3 & b_2 \end{bmatrix} \qquad (10)$$

where

$$b_n = a_n$$

$$b_i = a_i + \sum_{j=i+1}^{n} a_{n+i+1-j}\, b_j \quad , \quad i=n-1,\ n-2,\ \ldots,3,2.$$

Other Methods for Determining the Matrix K

The general procedure described above for computing the pair $(\underset{\sim}{K},\ \underset{\sim}{K}^{-1})$ may offer certain advantages in numerical calculation since it does not require the computation of eigenvalues or eigenvectors and does not involve an explicit evaluation of the determinant $|\underset{\sim}{A} - \lambda \underset{\sim}{I}|$. On the other hand, it is of some interest to study the algebraic structure of the matrix $\underset{\sim}{K} = \underset{\sim}{K}(\underset{\sim}{A},\ \underset{\sim}{f})$ in terms of the fundamental matrix theoretic notions of eigenvalues and eigenvectors. In such a study, the relative efficiency of numerical computing schemes is not of primary importance. The result given in [7] showed how, when the pair $(\underset{\sim}{A},\ \underset{\sim}{f})$ is completely controllable and $\underset{\sim}{A}$ has distinct eigenvalues, the matrix $\underset{\sim}{K}$ can be written as

$$\underset{\sim}{K} = \underset{\sim}{M}\ \underset{\sim}{M}_v^{-1} \qquad (11)$$

where $\underset{\sim}{M}$ is a certain modal matrix of $\underset{\sim}{A}$ (the columns of $\underset{\sim}{M}$ are n linearly independent column eigenvectors of the matrix $\underset{\sim}{A}$ which have been normalized[5] in a special way) and $\underset{\sim}{M}_v$ is the Vandermonde matrix of $\underset{\sim}{A}$. In later and independent studies of this problem, Mufti [8] and Ainsworth and Gunderson [9] generalized the result in [7] to allow for the possibility of non-distinct eigenvalues.

---

4. For example, set $\underset{\sim}{L}\underset{\sim}{L}^{-1} = \underset{\sim}{I}$.
5. In the notation of [7], $\underset{\sim}{M} = \underset{\sim}{M}\underset{\sim}{B}^{-1}$ where the columns of $\underset{\sim}{M}$ are n linearly independent column eigenvectors of $\underset{\sim}{A}$, and $\underset{\sim}{B}$ is a non-singular diagonal matrix which acts as a normalizing factor.

Relationships with Controllability and Observability

A study of controllability and the structure of $\underset{\sim}{K}$ in terms of the spectral properties of the matrix $\underset{\sim}{A}$, brings to light some facts which may have practical value. Some of these results have been described in [10]. In addition we have the following result

Proposition

Let $\underset{\sim}{A}$ be a real, constant n x n matrix with at least one repeated eigenvalue. Let $\underset{\sim}{R}$ denote a nonsingular matrix which transforms $\underset{\sim}{A}$ to the diagonal form $\underset{\sim}{\Lambda}$ = diag. $(\lambda_1, \ldots, \lambda_n)$

$$\underset{\sim}{R}^{-1} \underset{\sim}{A}\underset{\sim}{R} = \underset{\sim}{\Lambda}, \tag{12}$$

and let $\underset{\sim}{T}$ denote a nonsingular matrix which transforms $\underset{\sim}{A}$ to the companion form $\underset{\sim}{A}_o$ [defined as in (3)]

$$\underset{\sim}{T}^{-1} \underset{\sim}{A}\underset{\sim}{T} = \underset{\sim}{A}_o \ . \tag{13}$$

Then if $\underset{\sim}{R}$ exists $\underset{\sim}{T}$ does not exist.

Proof

If $\underset{\sim}{R}$ exists, $\underset{\sim}{A}$ must possess a total of n linearly independent column eigenvectors. On the other hand, if $\underset{\sim}{T}$ exists, the minimal polynomial of $\underset{\sim}{A}$ must equal the characteristic polynomial of $\underset{\sim}{A}$. The latter condition is satisfied if and only if $\underset{\sim}{A}$ has no more than one (within a constant multiplier) column eigenvector corresponding to each set of repeated eigenvalues. It follows that, when $\underset{\sim}{A}$ has repeated eigenvalues, $\underset{\sim}{T}$ cannot exist if $\underset{\sim}{R}$ exists.

In the literature on control theory, the concept of complete controllability of the linear dynamical system (1) is often illustrated by transforming the matrix $\underset{\sim}{A}$ to diagonal or Jordan canonical form and then observing the effect of this transformation on the vector $\underset{\sim}{f}$, [11], [12], [13]. As a complement to this procedure, the above result has the interesting

Corollary

Suppose the matrix $\underset{\sim}{A}$ in (1) has repeated eigenvalues. Then, if $\underset{\sim}{A}$ is similar to a diagonal matrix $\underset{\sim}{\Lambda}$, the linear dynamical

system (1) is always uncontrollable for any choice of the vector $\underset{\sim}{f}$. Moreover, if $y = \langle \underset{\sim}{c}, \underset{\sim}{x} \rangle$ is the scalar "output" of (1) then, under the conditions stated, (1) is always unobservable for any choice of the vector $\underset{\sim}{c}$.

In light of this fact, qualifying statements such as that found in the footnotes on pages 350-351 of [12] are seen to be unnecessary. The Corollary also shows that, from the viewpoint of controllability and observability, the case when $\underset{\sim}{A}$ has repeated eigenvalues actually does exhibit certain special properties.[6]

6. It is recalled that every real, symmetric matrix $\underset{\sim}{A}$ is similar to a strictly diagonal matrix $\underset{\sim}{\Lambda}$. Thus, the presence of repeated eigenvalues is a necessary and sufficient condition that (1), with a real symmetric $\underset{\sim}{A}$, be always uncontrollable and unobservable for any choice of the vectors $\underset{\sim}{f}$ and $\underset{\sim}{c}$.

References Cited

1. W. M. Wonham and C. D. Johnson, "Optimal Bang-Bang Control with Quadratic Performance Index," Proceedings, Fourth Joint Automatic Control Conference, Minneapolis, Minnesota, June 1963, pp. 101-112. See also, Trans. A.S.M.E. Jour. of Basic Engineering, Vol. 86, pp. 107-115, 1964.
2. W. G. Tuel, Jr., "On the Transformation to (Phase-Variable) Canonical Form," IEEE Transactions on Automatic Control, (to be published, July 1966).
3. L. M. Silverman, "Transformation of Time-Variable Systems to Canonical (Phase-Variable) Form, "IEEE Transactions on Automatic Control, (to be published, April 1966).
4. D. S. Rane, "A Simplified Transformation to (Phase-Variable) Canonical Form," IEEE Transactions on Automatic Control, (to be published in 1966).
5. R. E. Kalman, "Liapunov Functions for the Problem of Lure in Automatic Control," Proc. Nat'l. Acad. of Sci. U.S.A., Vol. 49, pp. 201-205, 1963.
6. S. Lefschetz, "Stability of Nonlinear Control Systems, "Academic Press, New York, 1965.
7. C. D. Johnson and W. M. Wonham, "A Note on the Transformation to Canonical (Phase-Variable) Form, " IEEE Transactions on Automatic Control, Vol. AC-9, pp. 312-313, July 1964.
8. I. H. Mufti, "On the Reduction of a System to Canonical (Phase-Variable) Form," IEEE Transactions on Automatic Control, Vol. AC-10, pp. 206-207, April 1965.
9. O. Ainsworth, Private communication to C. D. Johnson.
10. C. D. Johnson, "Invariant Hyperplanes for Linear Dynamical Systems," IEEE Transactions on Automatic Control, Vol. AC-11, pp. 113-116, January 1966.
11. L. A. Zadeh, and C. A. Desoer, "Linear System Theory, the State Space Approach," MCGraw-Hill Book Co., Inc., New York, 1963.
12. P. M. DeRusso, R. L. Roy, and C. M. Close, "State Variables for Engineers," John Wiley and Sons, New York, 1965.
13. Y. C. Ho, "What Constitutes a Controllable System?" I.R.E. Transactions on Automatic Control, p. 76, April 1962.

VI. Invariant Hyperplanes for Linear Dynamical Systems

C. D. Johnson

Abstract - In certain problems associated with the control of linear dynamical systems, the concept of invariant hyperplanes in the system state space plays an important role [1] - [8]. This paper gives conditions for the existence of invariant hyperplanes for linear dynamical systems and describes some geometric properties of these hyperplanes. In addition, some relationships between invariant hyperplanes and the concepts of controllability and observability are discussed.

Introduction

An important class of linear dynamical systems, with scalar input and output, can be described by[1]

$$\dot{\underset{\sim}{x}} = \underset{\sim\sim}{A}\underset{\sim}{x} + u(t)\underset{\sim}{f} \qquad (\dot{} = d/dt) \tag{1a}$$

$$y = \langle \underset{\sim}{h}, \underset{\sim}{x} \rangle \tag{1b}$$

where $\underset{\sim}{x} = (x_1, \ldots, x_n)$ is a real n-vector (the state vector of the system), $\underset{\sim}{A}$ is a real, constant n x n matrix, u(t) is a real, scalar function of time (the system input or control), $\underset{\sim}{f}$ and $\underset{\sim}{h}$ are real, constant, non-zero n-vectors, and y is a real scalar (the system output). Many of the mathematical properties associated with the dynamical system of (1) have convenient geometrical interpretations in the system state space, a euclidean n-space E^n whose points have coordinates $x_1, \ldots, x_n$.

This paper concerns a question about the existence of a certain property of the solutions of (1a) for the special case when $u(t) \equiv 0$. In particular, the following question is posed. What conditions are required for the existence of a linear form

$$\langle \underset{\sim}{c}, \underset{\sim}{x} \rangle = 0 \tag{2}$$

which is invariant along solutions of

$$\dot{\underset{\sim}{x}} = \underset{\sim\sim}{A}\underset{\sim}{x} \tag{3}$$

for arbitrary initial conditions satisfying (2) ? In (2), $\underset{\sim}{c}=(c_1, \ldots, c_n)$ is a real, constant, nonzero n-vector. In addition to answering this question, this paper gives a characterization of the stability of (2) along solutions of (3) and shows relationships

---

This work was supported, in part, by the National Aeronautics and Space Administration under Contract NAS8-11231 and Grant NsG-381.

The author is with the Dept. of Electrical Engineering, University of Alabama, Huntsville, Ala.

[1] The notation $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle$ denotes the inner product of $\underset{\sim}{x}$ and $\underset{\sim}{y}$.

between the form (2) and the real eigenvalues and eigenvectors of $\underset{\sim}{A}$. Finally, a solution is given for certain inverse problem associated with (2) and there is presented, for a special class of problems, a simple geometric interpretation of Kalman's concepts of controllability and observability.

The linear form (2) can be associated with an (n-1)-dimensional hyperplane in the state space E^n. This hyperplane has the property that if $\langle \underset{\sim}{c}, \underset{\sim}{x}(0) \rangle = 0$ then $\langle \underset{\sim}{c}, \underset{\sim}{x}(t) \rangle \equiv 0$ for all $t>0$ when $\underset{\sim}{x}(t)$ is a solution of (3). In the following, a linear form (2) having the above property is referred to as an invariant hyperplane $\mathcal{H}$ of (3).

If $\langle \underset{\sim}{c}, \underset{\sim}{x} \rangle = 0$ is an invariant hyperplane of (3) and if an arbitrary solution $\underset{\sim}{x}(t)$ of (3) satisfies

$$\langle \underset{\sim}{c}, \dot{\underset{\sim}{x}}(t) \rangle = \begin{cases} < 0 \ (> 0) & \text{for } \langle \underset{\sim}{c}, \underset{\sim}{x}(t) \rangle > 0 \\ > 0 \ (< 0) & \text{for } \langle \underset{\sim}{c}, \underset{\sim}{x}(t) \rangle < 0 \end{cases} \tag{4}$$

then the invariant hyperplane is called stable (unstable).

Let the eigenvalues of $\underset{\sim}{A}$ be $\lambda_1, \ldots, \lambda_n$ and let $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_n$ be a corresponding set of column eigenvectors. It is recalled that the column eigenvectors of $\underset{\sim}{A}$ are nonzero and satisfy

$$\underset{\sim}{A}\underset{\sim}{a}_i = \lambda_i \underset{\sim}{a}_i \qquad (i = 1, \ldots, n). \tag{5}$$

It is further recalled that corresponding to each eigenvalue λ_i, $i = 1, \ldots, n$, there is an associated nonzero row eigenvector β_i' which satisfies

$$\beta_i' \underset{\sim}{A} = \lambda_i \beta_i' \quad (' \text{ denotes transpose}). \tag{6}$$

Results

The conditions for the existence of invariant hyperplanes of (3) are summarized in the following.

Theorem 1

Let $\underset{\sim}{A}$ be a real, constant n x n matrix and let the real, nonzero eigenvalues of $\underset{\sim}{A}$ be denoted by $\lambda_1, \ldots, \lambda_m$; $m \leq n$. Then, corresponding to each eigenvalue λ_i $(i = 1, \ldots, m)$, there exists a real, constant, nonzero n-vector $\underset{\sim}{c}_i$ satisfying

$$\underset{\sim}{A}' \underset{\sim}{c}_i = \lambda_i \underset{\sim}{c}_i \qquad i = 1, \ldots, m \tag{7}$$

and such that

$$\langle \underset{\sim}{c}_i, \underset{\sim}{x}(t) \rangle \equiv 0. \tag{8}$$

along every solution of $\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x}$ which satisfies

$$<\underset{\sim}{c}_i, \underset{\sim}{x}(0)> = 0 \tag{9}$$

Moreover, $\underset{\sim}{c}_i'$ is a row eigenvector corresponding to λ_i and the invariant hyperplane

$$\mathcal{H}_i = \{ \underset{\sim}{x} \mid <\underset{\sim}{c}_i, \underset{\sim}{x}> = 0 \} \tag{10}$$

is stable (unstable), in the sense of (4), if $\lambda_i < 0$ (> 0).

Proof: Let $\underset{\sim}{B}$ denote the n x n matrix whose columns are

$$\underset{\sim}{c}, \underset{\sim}{A}'\underset{\sim}{c}, \underset{\sim}{A}'^2\underset{\sim}{c}, \ldots, \underset{\sim}{A}'^{n-1}\underset{\sim}{c}. \tag{11}$$

Then, from repeated differentiation of (2) it may be seen that (8) is satisfied along solutions of (3) for each $\underset{\sim}{x}(0)$ satisfying (9), if and only if rank $\underset{\sim}{B} = 1$. This latter condition is satisfied if and only if

$$\underset{\sim}{A}' \underset{\sim}{c} = \rho \underset{\sim}{c} \tag{12}$$

for some real scalar ρ. It follows from (6) that (12) has real, nonzero solutions $\rho_i = \lambda_i$, $\underset{\sim}{c}_i = \underset{\sim}{\beta}_i$ corresponding to each real nonzero eigenvalue λ_i, $i = 1, \ldots, m$. Setting $\xi_i = <\underset{\sim}{c}_i, \underset{\sim}{x}>$, $\underset{\sim}{c}_i = \underset{\sim}{\beta}_i$, the derivative of ξ_i along an arbitrary solution of (3) is found to be

$$\frac{d\xi_i(t)}{dt} = \lambda_i \xi_i(t) \qquad i = 1, \ldots, m. \tag{13}$$

It follows that the invariant hyperplane $\mathcal{H}_i = \{ x \mid <\underset{\sim}{\beta}_i, \underset{\sim}{x}> = 0 \}$ is stable (unstable) if $\lambda_i < 0 (> 0)$ $i = 1, \ldots, m$.

Corollary

Let $\underset{\sim}{a}_1, \ldots, \underset{\sim}{a}_m$ be the set of real column eigenvectors of $\underset{\sim}{A}$ and let $\underset{\sim}{c}_1, \ldots, \underset{\sim}{c}_m$ be the set of normal vectors associated with the corresponding m invariant hyperplanes of $\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x}$. Then the vectors $\underset{\sim}{c}_i$, $\underset{\sim}{a}_i$ satisfy the following orthogonality equation

$$<\underset{\sim}{c}_i, \underset{\sim}{a}_k> = 0; \ \lambda_i \neq \lambda_k \ (i, k = 1, \ldots, m). \tag{14}$$

Proof: The proof follows immediately from the well-known result that the row eigenvector corresponding to any eigenvalue is orthogonal to the column eigenvector corresponding to any different eigenvalue.

Remarks

1) If $\underset{\sim}{A}$ has repeated real eigenvalues, it is sometimes, but not always, possible to find more than one linearly independent row eigenvector corresponding to the same (nondistinct) eigenvalue. Since any linear combination of such row eigenvectors is also a row eigenvector, it follows that in such a case an infinitude of distinct invariant hyperplanes will be associated with the same (nondistinct) eigenvalue.

2) The assumption that the real λ_i are nonzero ($i = 1, \ldots, m$) assures that (13) has a unique equilibrium state $\xi(t) \equiv 0$. In this case, the corresponding invariant hyperplanes $\mathcal{H}_i$ always pass through the origin $\underset{\sim}{x} = \underset{\sim}{0}$. In the case of a distinct zero eigenvalue the corresponding vector $\underset{\sim}{c}$ still satisfies (7) but the equilibrium states of (13) are then defined by

$$\xi = \langle \underset{\sim}{c}, \underset{\sim}{x} \rangle = z \tag{15}$$

where z is an arbitrary real scalar constant. Thus, to each distinct zero eigenvalue there corresponds an infinite number of parallel invariant hyperplanes (15). The parallel hyperplanes corresponding to a zero eigenvalue are neutrally stable in the sense that along solutions of (3)

$$\langle \underset{\sim}{c}_i, \dot{\underset{\sim}{x}}(t) \rangle \equiv 0 \tag{16}$$

for all values of $\langle \underset{\sim}{c}_m, \underset{\sim}{x}(t) \rangle$.

3) The Corollary shows that the eigenvector $\underset{\sim}{a}_k$, corresponding to the real eigenvalue λ_k, lies on the intersection of the set of invariant hyperplanes $\mathcal{H}_i = \{ x \mid \langle \underset{\sim}{c}_i, \underset{\sim}{x} \rangle = 0 \}$ where $\underset{\sim}{c}_i = \underset{\sim}{\beta}_i(\lambda_i)$, $\lambda_i \neq \lambda_k$, $i, k = 1, \ldots, m$. In certain special cases, this result leads to a useful geometric interpretation of eigenvectors. Consider, for example, the case $n = 3$ and suppose that the corresponding three eigenvalues $\lambda_1, \lambda_2, \lambda_3$ are all real, distinct, and nonzero. In this case there are three distinct invariant hyperplanes which pass through the origin $x = 0$. The pairwise intersections between these three hyperplanes generate three lines that pass through the origin $x = 0$ and are collinear with the three eigenvectors a_1, a_2, and a_3. This result is illustrated in Figure 1.

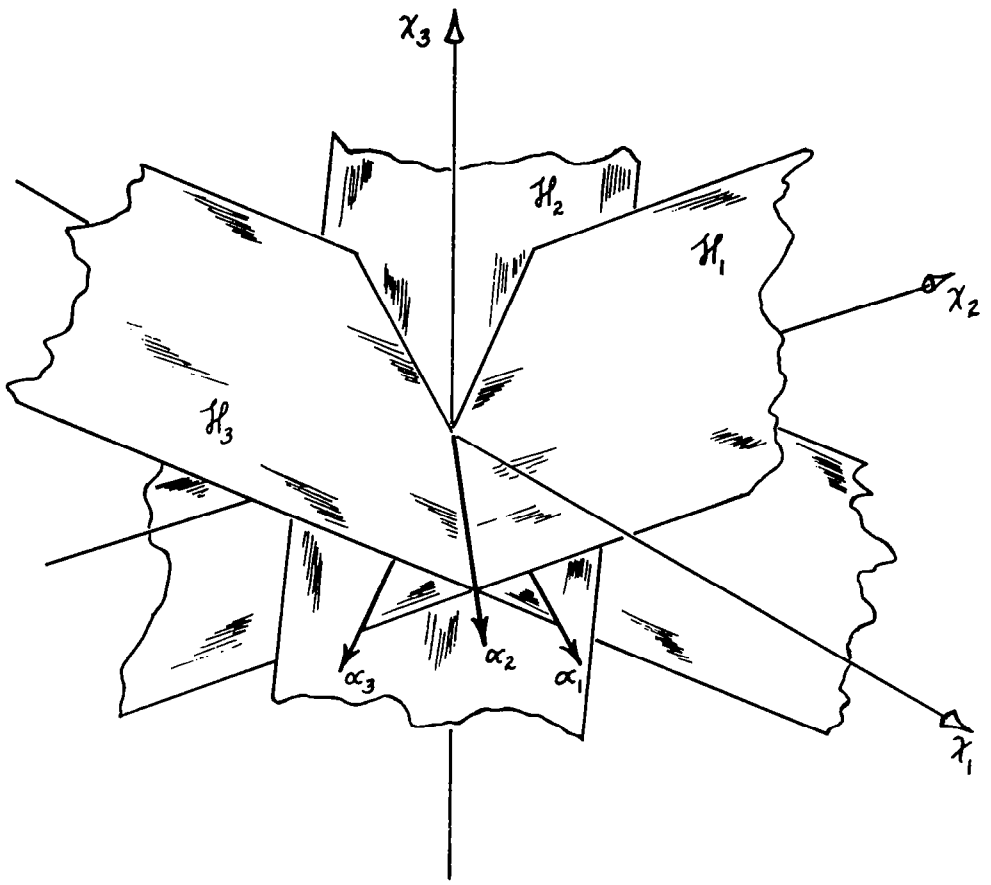

Figure 1 – Showing the eigenvectors $\underset{\sim}{a}_1$, $\underset{\sim}{a}_2$, $\underset{\sim}{a}_3$ lying on the intersections of the invariant hyperplanes $\mathcal{H}_1$, $\mathcal{H}_2$, $\mathcal{H}_3$.

4) From (13) it is clear that in the state space E^n the integral curves of the system (3) cannot cross over[2] the invariant hyperplanes of that system. Thus, the set of invariant hyperplanes of $\dot{\underset{\sim}{x}} = \underset{\sim}{\underset{\sim}{A}}\underset{\sim}{x}$ define boundaries of n-dimensional convex subsets of the state space which are invariant for the corresponding solution $\underset{\sim}{x}(t)$. This fact, together with the fact that an arbitrary solution $\underset{\sim}{x}(t)$ always monotonically approaches (recedes from)[3] each stable (unstable) $\mathcal{H}$, is useful in establishing relative bounds on components of $\underset{\sim}{x}(t)$.

An Inverse Problem

In this section the following inverse problem is considered: Find a real, constant n-vector $\underset{\sim}{\gamma} = (\gamma_1, \ldots, \gamma_n)$ such that the hyperplane

$$\mathcal{H} = \{ \underset{\sim}{x} \mid \langle \underset{\sim}{c}, \underset{\sim}{x} \rangle = 0 \} \tag{17}$$

is invariant for the linear dynamical system described by

$$\dot{\underset{\sim}{x}} = \underset{\sim}{\underset{\sim}{A}}\underset{\sim}{x} + \langle \underset{\sim}{\gamma}, \underset{\sim}{x} \rangle \underset{\sim}{f} \tag{18}$$

where $\underset{\sim}{c}$ and $\underset{\sim}{f}$ are real, constant n-vectors, and $\underset{\sim}{\underset{\sim}{A}}$ is a real, constant n x n matrix. This problem differs from the previous problem in that the hyperplane (17), (i.e., the vector $\underset{\sim}{c}$) is assumed to be specified a priori.

From the results of Theorem 1, it is clear that $\underset{\sim}{\gamma}$ must be chosen so that

1) the matrix

$$\underset{\sim}{\underset{\sim}{A}} + \underset{\sim}{f}\underset{\sim}{\gamma}' \tag{19}$$

has at least one real eigenvalue, and

2) one of the invariant hyperplanes of (18), corresponding to one of the real eigenvalues of (19), is defined by (17).

A vector $\underset{\sim}{\gamma}$ which has the required properties is given in Theorem 2.

Theorem 2

Let $\underset{\sim}{\underset{\sim}{A}}$ be a real, constant n x n matrix, and let $\underset{\sim}{\gamma}$, $\underset{\sim}{f}$, and $\underset{\sim}{c}$ be real, constant n-vectors. Then, the hyperplane

$$\mathcal{H} = \{ \underset{\sim}{x} \mid \langle \underset{\sim}{c}, \underset{\sim}{x} \rangle = 0 \} \tag{20}$$

is invariant for the linear dynamical system

---

2. More precisely, if $\underset{\sim}{x}(0) \notin \mathcal{H}$, then $\underset{\sim}{x}(t)$ cannot enter $\mathcal{H}$ in finite time.

3. Here, the "distance" from $\underset{\sim}{x}(t)$ to $\mathcal{H}$ is taken as $\|\underset{\sim}{c}\|^{-1} \langle \underset{\sim}{c}, \underset{\sim}{x}(t) \rangle$.

$$\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x} + \langle \underset{\sim}{\gamma}, \underset{\sim}{x} \rangle \underset{\sim}{f} \tag{21}$$

if $\underset{\sim}{\gamma}$ satisfies

$$\underset{\sim}{\gamma} = \langle \underset{\sim}{c}, \underset{\sim}{f} \rangle^{-1} (-\underset{\sim}{A}' + k \underset{\sim}{I})\underset{\sim}{c} \tag{22}$$

where k is an arbitrary, real scalar constant.

Moreover, for the system of (21) and (22), the invariant hyperplane (20) is stable (unstable), in the sense of (4), if $k < 0$ (> 0).

Proof: Using the notation $\xi = \langle \underset{\sim}{c}, \underset{\sim}{x} \rangle$, the derivative $d\xi/dt$ along an arbitrary solution of (21) is computed to be

$$\begin{aligned} \frac{d\xi}{dt} &= \langle \underset{\sim}{c}, (\underset{\sim}{A} + \underset{\sim}{f}\,\underset{\sim}{\gamma}')\,\underset{\sim}{x}(t) \rangle \\ &= k\xi + \langle \underset{\sim}{c}, \underset{\sim}{f} \rangle \langle [\underset{\sim}{\gamma} - \langle \underset{\sim}{c}, \underset{\sim}{f} \rangle^{-1} (-\underset{\sim}{A}' + k\underset{\sim}{I})\,\underset{\sim}{c}\,], \underset{\sim}{x}(t) \rangle. \end{aligned} \tag{23}$$

The last term on the right of (23) vanishes when

$$\underset{\sim}{\gamma} = \langle \underset{\sim}{c}, \underset{\sim}{f} \rangle^{-1} (-\underset{\sim}{A}' + k\underset{\sim}{I})\,\underset{\sim}{c} \tag{24}$$

and in this case

$$\frac{d\xi}{dt} = k\xi \tag{25}$$

along an arbitrary solution of (21). This completes the proof of Theorem 2.

Remarks

Equation (18) defines the vector $\dot{\underset{\sim}{x}}$ in terms of the two components $\underset{\sim}{A}\underset{\sim}{x}$ and $\langle \underset{\sim}{\gamma}, \underset{\sim}{x} \rangle \underset{\sim}{f}$. For the special case $\langle \underset{\sim}{c}, \underset{\sim}{f} \rangle = 0$, the component $\langle \underset{\sim}{\gamma}, \underset{\sim}{x} \rangle \underset{\sim}{f}$ is always in the hyperplane $\langle \underset{\sim}{c}, \underset{\sim}{x} \rangle = 0$. It follows that, for this special case, $\langle \underset{\sim}{c}, \underset{\sim}{x} \rangle = 0$ is an invariant hyperplane of (18) if and only if $\langle \underset{\sim}{c}, \underset{\sim}{x} \rangle = 0$ is an invariant hyperplane of $\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x}$. Under this condition, the choice of the vector $\underset{\sim}{\gamma}$ is immaterial.

Relation with Controllability and Observability

The linear dynamical system (1) is said to be completely controllable [9] - [12] in the state space E^n if and only if, for each finite pair of states $(\underset{\sim}{x}_o, \underset{\sim}{x}_T) \in E^n$, there exists a finite interval $[0, T]$ and a control $u = \phi(t; \underset{\sim}{x}_o, \underset{\sim}{x}_T)$, $0 \leq t \leq T$, such that if $\underset{\sim}{x}(0) = \underset{\sim}{x}_o$ then $\underset{\sim}{x}(T) = \underset{\sim}{x}_T$ along the corresponding solution of (1). In a like manner, the linear dynamical system (1) is said to be completely observable [9]-[12] in the state space E^n if and only if, for each finite output $y(t)$, $0 \leq t \leq T > 0$, which satisfies (1) [with $u(t) \equiv 0$], there corresponds a unique initial state $\underset{\sim}{x}(0)$.

Let $\underset{\sim}{P}$ be the n x n matrix whose columns are $\underset{\sim}{f}$, $\underset{\sim}{Af}$, $\underset{\sim}{A}^2\underset{\sim}{f}$, ... , and $\underset{\sim}{A}^{n-1}\underset{\sim}{f}$, and let $\underset{\sim}{R}$ be the n x n matrix whose columns are $\underset{\sim}{h}$, $\underset{\sim}{A}'\underset{\sim}{h}$, $\underset{\sim}{A}'^2\underset{\sim}{h}$, ..., and $\underset{\sim}{A}'^{n-1}\underset{\sim}{h}$. Then a necessary and sufficient condition for the linear dynamical system (1) to be completely controllable (observable) is that rank $\underset{\sim}{P} = n$ (rank $\underset{\sim}{R} = n$) [10]. In other words, the system (1) becomes uncontrollable (unobservable) if and only if the vector $\underset{\sim}{f}(\underset{\sim}{h})$ lies in a proper $\underset{\sim}{A}$-invariant ($\underset{\sim}{A}'$-invariant) subspace[4] of E^n with dimension less than n.

The invariant hyperplanes discussed in Section II are (n-1)- dimensional $\underset{\sim}{A}$-invariant subspaces of E^n. It follows from the previous remarks that the linear dynamical system (1) becomes uncontrollable, in particular, whenever $\underset{\sim}{f}$ lies in one or more of the invariant hyperplanes of (3). Suppose, for example, that $\langle \underset{\sim}{c}_i, \underset{\sim}{f} \rangle = 0$ for some $\underset{\sim}{c}_i$ which satisfies (7). Then, the derivative of $\xi = \langle \underset{\sim}{c}_i, \underset{\sim}{x} \rangle$ along an arbitrary solution of (1a) is

$$\frac{d\xi}{dt} = \lambda_i \xi \tag{26}$$

which shows that, irrespective of the choice of u(t), the integral curve $\underset{\sim}{x}(t)$ cannot cross over the hyperplane $\langle \underset{\sim}{c}_i, \underset{\sim}{x} \rangle = 0$.

An important connection between the $\underset{\sim}{A}$-invariant and $\underset{\sim}{A}'$-invariant subspaces of E^n is summarized in the following well-known result of matrix theory: If S is an $\underset{\sim}{A}$-invariant subspace of E^n then the orthogonal complement of S is an $\underset{\sim}{A}'$-invariant subspace of E^n. Since the real column (row) eigenvectors of $\underset{\sim}{A}$ are one-dimensional $\underset{\sim}{A}$-invariant ($\underset{\sim}{A}'$-invariant) subspaces, this result shows that

1) each real row eigenvector of $\underset{\sim}{A}$ is orthogonal to an (n-1)-dimensional $\underset{\sim}{A}$-invariant subspace, and

2) each real column eigenvector of $\underset{\sim}{A}$ is orthogonal to an (n-1)-dimensional $\underset{\sim}{A}'$-invariant subspace. The first of these facts is recognized as an alternative proof of the existence of the invariant hyperplanes described in Section II. The second fact shows that the linear dynamical system (1) becomes unobservable, in particular, whenever the vector $\underset{\sim}{h}$ lies in one or more of the (n-1)-dimensional

---

4. A subspace $S \subset E^n$ is said to be $\underset{\sim}{A}$-invariant if $\underset{\sim}{x} \in S$ implies $\underset{\sim}{Ax} \in S$ for all $\underset{\sim}{x} \in S$.

hyperplanes orthogonal to the real column eigenvectors of $\underset{\sim}{A}$.

In the state space E^n, the union of all $\underset{\sim}{A}$-invariant subspaces with dimension less than n is the set $F(\underset{\sim}{A})$ of all vectors $\underset{\sim}{f}$ for which the system (1) is uncontrollable. Likewise, the union of all $\underset{\sim}{A}'$-invariant subspaces with dimension less than n is the set $H(\underset{\sim}{A})$ of all vectors $\underset{\sim}{h}$ for which the system (1) is unobservable. The case when either of the sets $F(\underset{\sim}{A})$ or $H(\underset{\sim}{A})$ is n-dimensional is particularly important since, in that event, both F(A),H(A) are n-dimensional and the system (1) is always uncontrollable and unobservable irrespective of the choice of the vectors $\underset{\sim}{f}$ and $\underset{\sim}{h}$! This degenerate condition occurs if and only if rank $\underset{\sim}{P}<n$ for arbitrary $\underset{\sim}{f}$. In other words, if and only if there exist real scalars $r_0, r_1, \ldots, r_k$, $(k \leq n-1)$, not all zero, such that

$$r_0\underset{\sim}{I} + r_1\underset{\sim}{A} + r_2\underset{\sim}{A}^2 + \ldots + r_k\underset{\sim}{A}^k = 0. \qquad (27)$$

An n x n matrix $\underset{\sim}{A}$ which satisfies (27), with $k \leq n-1$, must necessarily[5] possess repeated eigenvalues and is said to be derogatory. That is, the minimal polynomial of $\underset{\sim}{A}$ is of lower degree than the characteristic polynomial. The nature of the geometric structure that causes this degenerate condition can be illustrated by considering the special case n=2. For that case, the only proper $\underset{\sim}{A}$-invariant ($\underset{\sim}{A}'$-invariant) subspaces of dimension less than n are the real column (row) eigenvectors of $\underset{\sim}{A}$. Thus, the second-order system (1) is uncontrollable (unobservable) if and only if the vector $\underset{\sim}{f}(\underset{\sim}{h})$ is collinear with one of the real column (row) eigenvectors of $\underset{\sim}{A}$. If the 2 x 2 matrix $\underset{\sim}{A} \neq \underset{\sim}{0}$ is derogatory, it follows from (27) that the two eigenvalues of $\underset{\sim}{A}$ must be real and repeated, $\lambda_1 = \lambda_2 = \lambda \neq 0$, and $\underset{\sim}{A}$ must have the diagonal form

$$\underset{\sim}{A} = \lambda\underset{\sim}{I} \qquad \lambda \neq 0. \qquad (28)$$

It is readily verified from (5) and (6) that the column and row eigenvectors of (28) are nondistinct and can be chosen as any vector in E^2. That is, the state space portrait of (3), with $\underset{\sim}{A}$ given by (28) consists of the family of all straight lines that pass through the origin $\underset{\sim}{x} = \underset{\sim}{0}$. It follows that every vector $\underset{\sim}{f} \in E^2$ ($\underset{\sim}{h} \in E^2$)

5. The presence of repeated eigenvalues is a necessary but not sufficient condition for $\underset{\sim}{A}$ to be derogatory. A necessary and sufficient condition for $\underset{\sim}{A}$ to be derogatory is that there exists more than one linearly independent column eigenvector corresponding to the same (repeated) eigenvalue.

is collinear with one of the column (row) eigenvectors of $\underset{\sim}{A}$.

In general, the proper $\underset{\sim}{A}$-invariant and $\underset{\sim}{A}'$-invariant subspaces of E^n appear with a variety of dimensions.[6] For this reason, the loss of controllability and/or observability of the general nth-order system (1) cannot always be characterized solely in terms of the particular one-dimensional and (n - 1)-dimensional invariant subspaces previously discussed. It is possible, however, to give such characterizations for the special cases n = 2 and n = 3 since, for those two cases, the real eigenvectors and their respective orthogonal complements are the only invariant subspaces of interest. These characterizations, which follow immediately from the results previously described, may be summarized as follows.

Theorem 3

The second order, n = 2, linear dynamical system (1) is:

1) always completely controllable and completely observable, irrespective of the choices of the vectors $\underset{\sim}{f}$ and $\underset{\sim}{h}$, if and only if $\underset{\sim}{A}$ has no real eigenvalues (i.e., if and only if the system is "underdamped").

2) uncontrollable (unobservable) if and only if the vector $\underset{\sim}{f}(\underset{\sim}{h})$ is collinear with a real column (row) eigenvector of $\underset{\sim}{A}$.

3) always uncontrollable and unobservable irrespective of the choices of the vectors $\underset{\sim}{f}$ and $\underset{\sim}{h}$, if and only if $\underset{\sim}{A} = r_0\underset{\sim}{I}$ for some real scalar constant r_0.

Theorem 4

The third order, n = 3, linear dynamical system (1) is:

1) always uncontrollable (unobservable) for some choices of the vector $\underset{\sim}{f}(\underset{\sim}{h})$,

2) uncontrollable (unobservable) if and only if the vector $\underset{\sim}{f}(\underset{\sim}{h})$ is either collinear with a real column (row) eigenvector of $\underset{\sim}{A}$ or lies on a 2-dimensional plane that is orthogonal to one of the real row (column) eigenvectors of $\underset{\sim}{A}$.

3) always uncontrollable and unobservable, irrespective of the choices of the vectors $\underset{\sim}{f}$ and $\underset{\sim}{h}$, if and only if either

$$\underset{\sim}{A}^2 = r_1\underset{\sim}{A} + r_0\underset{\sim}{I}$$

or

$$\underset{\sim}{A} = r_0\underset{\sim}{I}$$

---

6. For example, there is a real 2-dimensional $\underset{\sim}{A}$-invariant subspace and a real 2-dimensional $\underset{\sim}{A}'$-invariant subspace associated with each distinct pair of complex-conjugate eigenvalues of $\underset{\sim}{A}$.

for some real scalar constants r_0, r_1.

In both Theorem 3 and Theorem 4, the absence of repeated eigenvalues is sufficient to guarantee the nonexistence of the degenerate condition 3).

Conclusion

In this paper, the set of (n-1)-dimensional hyperplanes (2) that are invariant along solutions of (3) have been identified as the orthogonal complements of the real row eigenvectors of $\underset{\sim}{A}$. A stability property of these hyperplanes, along solutions of (3), has been defined and characterized in terms of the associated real eigenvalues of $\underset{\sim}{A}$. In addition, some geometric relationships between the concepts of controllability and observability and the real column and row eigenvectors of $\underset{\sim}{A}$ have been described. By means of these relationships, the notions of controllability and observability for second and third-order linear dynamical systems (1) can be completely explained in terms of simple geometric characterizations.

References

[1] S. V. Yemel'yanov and V. A. Taran, "A class of systems of automatic control with variable structure," Izv. AN SSSR OTN, Energetika i Avtomatika, vol. 3, 1962.

[2] Ye. I. Gerashchenko, "Stability of motion in a sliding hyperplane, " Izv. AN SSSR, Tekhn. Kibernetika, vol. 4, 1963.

[3] Ye. I. Geraschenko, "Stability of a class of nonlinear systems," Tekhn. Kibernetika, vol. 2, 1964.

[4] Ye. A. Barbashin, V. A. Tabuyeva, and R. M. Eydinov, "Stability of a system of control with variable structure and disturbance of sliding conditions," Avtomat. i Telemekh. vol. 7, p. 24, 1963.

[5] V. M. Badkov and Ye. A. Barbashin, "Stabilization of a control system when controller parameters are subject to limitations," Tekhn. Kibernetika, vol. 2, 1964.

[6] Ye. B. Dudin, "Switching hyperplane in variable structure tracking system," Tekhn. Kibernetika, vol. 2, 1964.

[7] W. M. Wonham and C. D. Johnson, "Optimal bang-bang control with quadratic performance index," ASME Trans. J. Basic Eng., vol. 86, ser. D. pp. 107-115, March 1964.

[8] C. D. Johnson and W. M. Wonham, "On a problem of Letov in optimal control," J. Basic Eng., vol. 87, ser. D, pp. 81-89, March 1965.

[9] R. E. Kalman, "Contributions to the theory of optimal control," Bol. Soc. Mat. Mex., pp. 102-119, 1960.

[10] R. E. Kalman, "On the general theory of control systems," Proc. First Congress, Internat'l. Feder. of Automatic Control, T. Butterworths, London, England, vol. 1, pp. 481-492, 1961.

[11] R. E. Kalman, Y. C. Ho, and K. S. Narendra, "Controllability of linear dynamical systems," Contribs. to Differential Equations, vol. 1, no. 2, pp. 189-213, 1962.

[12] E. Kreindler and P. E. Sarachik, "On the concepts of controllability and observability in linear systems," IEEE Trans. on Automatic Control, vol. AC-9, pp. 129-136, April 1964.

VII. Optimal Control With Quadratic Performance Index And Fixed Terminal Time

C. D. Johnson[+] J. E. Gibson[++]

Summary

The conventional solution for the optimal control of a linear stationary regulator with quadratic performance index and fixed terminal time leads to a linear feedback law with time varying gain coefficients [1].[1] In addition to the usual disadvantages of time variable controllers, these time varying gain coefficients approach infinity as the specified terminal time is approached.

In the present paper, it is shown that the optimal control for the above problem can be expressed as a time invariant nonlinear feedback law. Certain parameters in the nonlinear feedback law are functions of the initial time and initial state of the system. The conventional time varying linear feedback law can be obtained directly from the time invariant nonlinear feedback law.

The results of the present paper are applicable to a more general class of optimal control problems involving linear and nonlinear systems. Two examples are given to illustrate the method.

1. Statement of the Problem

The problem is to find a control $u(t)$ which minimizes the functional[2]

$$J[u] = \frac{1}{2} \int_{t_o}^{T} [\, \langle \underset{\sim}{x}(t), \underset{\sim}{\underset{\sim}{Q}} \underset{\sim}{x}(t) \rangle + c^2 u^2(t) \,] \, dt \qquad (1)$$

[+] Electrical Engineering Department, University of Alabama in Huntsville, Huntsville, Alabama. This work was supported in part by the National Aeronautics and Space Administration under Grant No. NsG-381 and Contract No. NAS8-11231.

[++] Control and Information Systems Laboratory, School of Electrical Engineering, Purdue University, Lafayette, Indiana.

[1] Numbers in brackets designate references at end of paper.

[2] $\langle \underset{\sim}{x}, \underset{\sim}{y} \rangle$ is the scalar product of $\underset{\sim}{x}$ and $\underset{\sim}{y}$.

subject to the following conditions:

$$\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x} + u(t)\underset{\sim}{f} \qquad (\cdot = d/dt) \tag{2}$$

$$\underset{\sim}{x}(t_o) = \underset{\sim}{x}_o \tag{3}$$

$$\underset{\sim}{x}(T) = \underset{\sim}{0} \qquad (T \text{ is fixed}) \tag{4}$$

In (1), $\underset{\sim}{Q}$ is a symmetric, positive semi-definite constant matrix and c is a non-zero scalar constant. In (2), $\underset{\sim}{x} = (x_1, \ldots, x_n)$ is the state vector of the plant, $\underset{\sim}{A}$ is an (n x n) constant matrix, $\underset{\sim}{f} = (f_1, \ldots, f_n)$ is a constant n-vector and u(t) is the scalar control function. It is assumed that u(t) is piecewise continuous but otherwise unrestricted. It is further assumed that the pair $(\underset{\sim}{A}, \underset{\sim}{f})$ is controllable. Then as shown in [2], there is no loss of generality in assuming that $\underset{\sim}{A}, \underset{\sim}{f}$ have the canonical form

$$\underset{\sim}{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & & 0 \\ \vdots & & & & \vdots \\ 0 & \cdots & 0 & & 1 \\ a_1 & a_2 & a_3 & \cdots & a_n \end{bmatrix}, \qquad \underset{\sim}{f} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}. \tag{5}$$

The optimal control u for the above problem can, in principle, be found by straightforward application of the Hamilton-Jacobi theory. Since the details of the Hamilton-Jacobi formulation of this problem have already been given in [1] we will only summarize the results.

2. Form of the Optimal Control Law

Let $u^o = \phi^o(\underset{\sim}{x}, t, T)$ be an optimal control law, and let

$$V(\underset{\sim}{x}, t, T) = J[u^o]; \quad \underset{\sim}{x}(t_o) = \underset{\sim}{x}, \quad t_o = t. \tag{6}$$

Then it can be shown [1] that V satisfies the Hamilton-Jacobi equation

$$\frac{\partial V}{\partial t}(\underset{\sim}{x}, t, T) + \langle \underset{\sim}{\nabla} V(\underset{\sim}{x}, t, T), \underset{\sim}{A}\underset{\sim}{x} \rangle - \frac{c^{-2}}{2} \langle \underset{\sim}{\nabla} V(\underset{\sim}{x}, t, T), \underset{\sim}{f} \rangle^2 + \frac{1}{2} \langle \underset{\sim}{x}, \underset{\sim}{Q}\underset{\sim}{x} \rangle = 0 \tag{7}$$

where

$$\underset{\sim}{\nabla} V = \left(\frac{\partial V}{\partial x_1}, \ldots, \frac{\partial V}{\partial x_n}\right). \tag{8}$$

Further, the optimal control law is given by

$$\phi^o(\underset{\sim}{x}, t, T) = c^{-2} \langle -\underset{\sim}{\nabla} V(\underset{\sim}{x}, t, T), \underset{\sim}{f} \rangle. \tag{9}$$

The control law (9) may be determined by solving (7) directly or, alternately, by solving for the characteristic strips of (7). The direct solution of (7) is discussed below. The method of characteristic strips, which in this particular case leads to Pontryagin's canonical equations, is discussed in Appendix 1.

3. Solution of the Hamilton-Jacobi Equation

The term $\partial V/\partial t$ in (7) is related to Pontryagin's Hamiltonian function H by the relation

$$\frac{\partial V}{\partial t}(\underset{\sim}{x}(t), t, T) = H(\underset{\sim}{x}(t), -\underset{\sim}{\nabla} V(\underset{\sim}{x}(t), t, T), t) \tag{10}$$

which holds along optimal trajectories $\underset{\sim}{x}(t)$. For the problem (1) - (4) it is well known that the Hamiltonian (10) is constant along optimal trajectories.

$$H(\underset{\sim}{x}(t), -\underset{\sim}{\nabla}V(\underset{\sim}{x}(t), t, T), t) \equiv \beta \qquad (t_o \leq t \leq T) \qquad (\beta = \text{constant}) \qquad (11)$$

From (10) and (11), it is clear that a complete integral[3] of (7) must be linear in t and of the form

$$V = \mathcal{V}(\underset{\sim}{x}, \alpha_1, \ldots, \alpha_{n-1}, \beta) + \beta t + \alpha_0 \qquad (12)$$

where β is given by (11). The α_i and β in (12) are n+1 integration constants (constants of motion) which can be evaluated from the specified initial and terminal states (3), (4).[4] Thus

$$\alpha_i = \alpha_i(\underset{\sim}{x}_o, t_o, T) \qquad (i = 1, \ldots, n-1) \qquad (13)$$

$$\beta = \beta(\underset{\sim}{x}_o, t_o, T) \quad . \qquad (14)$$

Moreover, it can be shown [see Appendix 2] that along optimal trajectories there are an additional n constants of motion given by

$$\frac{\partial V}{\partial \alpha_i} = k_i \qquad (i = 1, \ldots, n-1)$$

$$\frac{\partial V}{\partial \beta} = k_n \qquad (k_j = \text{constant};\ j = 1, \ldots, n) \qquad (15)$$

It is at this point that the present method of solution differs from conventional methods. In the conventional methods of solving the problem (1) - (4); [1], [3], [4], [5], [6], [7], [8], [9], [10], it is assumed that the solution to (7) is a positive definite (or semi-definite) quadratic form[5]

---

[3] See Appendix 2.

[4] It should be noted from (6) that two boundary conditions for (12) are $V(\underset{\sim}{0}, t, T) = 0$, $\forall t_o \leq t \leq T$ and $V(\underset{\sim}{x}, T, T) = +\infty$, $\forall \underset{\sim}{x} \neq 0$.

[5] In [7], [8], [9], a solution is assumed in the form of a finite series with time variable coefficients.

$$V = \frac{1}{2} \langle \underset{\sim}{x}, \underset{\sim}{M}(t, T)\underset{\sim}{x} \rangle \tag{16}$$

where the elements of the (n x n) matrix $\underset{\sim}{M}$ are functions of time $m_{ij} = m_{ij}(t, T)$ $i,j = 1, \ldots, n$. Upon substituting (16) into (7) the $m_{ij}(t, T)$ are determined as the solution to an ordinary nonlinear matrix differential equation of the Riccati type. By this means, the assumed solution (16) leads to a time varying linear feedback control law of the form

$$\phi^o(\underset{\sim}{x}, t, T) = \langle \underset{\sim}{\gamma}(t, T), \underset{\sim}{x} \rangle \tag{17}$$

where $\underset{\sim}{\gamma}(t, T) = (\gamma_1(t, T), \ldots, \gamma_n(t, T))$ is a time varying gain vector. A well known practical disadvantage of the solution (17) is that $|\underset{\sim}{\gamma}(t, T)| \to \infty$ as $t \to T$.[6]

In the present method of solving the problem (1) - (4), the solution to (7) is sought in the form of a complete integral of the type (12).[7] By this means, the optimal control (9) is obtained as a time invariant nonlinear feedback law of the form

$$u^o = \phi^o(\underset{\sim}{x}, \underset{\sim}{x}_o, t_o, T) \tag{18}$$

where

$$\phi^o = c^{-2} \langle -\nabla v(\underset{\sim}{x}, \underset{\sim}{x}_o, t_o, T), \underset{\sim}{f} \rangle. \tag{19}$$

It is remarked that in some special cases it may be possible to obtain the expression (19) without solving for (12) explicitly.

In (19) the initial conditions $\underset{\sim}{x}_o$, t_o are arbitrary for $t_o \leq T$. Thus, the nonlinear control law (19) can easily be transformed to the conventional time varying linear control law (17) by setting $\underset{\sim}{x}_o = \underset{\sim}{x}(t)$ and $t_o = t$ in (19).[8] This

---

[6] Rekasius [10] has proposed a method for avoiding the infinite gain associated with (17) by expressing (17) in the alternate form $\phi^o(\underset{\sim}{x}, t, T) = \langle \underset{\sim}{\eta}, \underset{\sim}{x} \rangle + \psi(t, \underset{\sim}{x}_o, T)$ where $\underset{\sim}{\eta}$ = constant.

[7] Complete integrals of the form (12) may differ considerably from the quadratic form (16) [see Example 1 below].

[8] It is for this reason that (19) is termed a control "law". By definition, a control law should depend only on the instantaneous values of $\underset{\sim}{x}(t)$, t.

latter step illustrates the relationship between the two alternate forms of solutions (17) and (19). That is, in the solution (19) the time varying portion of (17) is replaced by the constants of motion (13), (14).[9] The control functions (17) and (19) are mathematically equivalent solutions which differ only in functional form.

In the problem (1) - (4), if the fixed terminal time T is infinite (or, equivalently, if T is unrestricted) then β in (11) becomes zero and the optimal control (9) reduces to a time invariant linear feedback law of the form [11]

$$\phi^{o}(\underset{\sim}{x}) = \langle \underset{\sim}{\gamma}, \underset{\sim}{x} \rangle \tag{20}$$

where $\underset{\sim}{\gamma} = (\gamma_1, \ldots, \gamma_n)$ is a constant n-vector.

A variation of the problem (1) - (4) is obtained by fixing the elapsed time $\tau = T - t_o$. In this case, one may arbitrarily set $t_o = 0$, $T = \tau$ and (19) contains one less parameter.

4. Comparison of Alternate Solutions

The conventional time varying linear feedback law (17) is illustrated in Figure 1. This form of solution has the advantage of being independent of the initial state $\underset{\sim}{x}(t_o)$. That is, the control $u^{o}(\underset{\sim}{x}, t, T)$ is always optimal with respect to any instantaneous state $\underset{\sim}{x}(t)$, $(t \leq T)$. A practical disadvantage of this solution is the physical unrealizability and extreme sensitivity of the feedback controller as $t \to T$ and $|\underset{\sim}{\gamma}(t, T)| \to \infty$.

The time-invariant, nonlinear feedback law (19) is illustrated in Figure 2. The switches s_1, s_2 represent devices which, when activated, will sample and hold the initial conditions $\underset{\sim}{x}(t_o)$, t_o. An apparent disadvantage of this solution is the fact that the control $u^{o}(\underset{\sim}{x}, \underset{\sim}{x}_o, t_o, T)$ is optimal only for states $\underset{\sim}{x}(t)$ lying on the optimal trajectory passing through the initial state $\underset{\sim}{x}(t_o)$. Thus if, after

---

[9] Note that the reverse transformation [from (17) to (19)] requires a priori knowledge of the constants of motion.

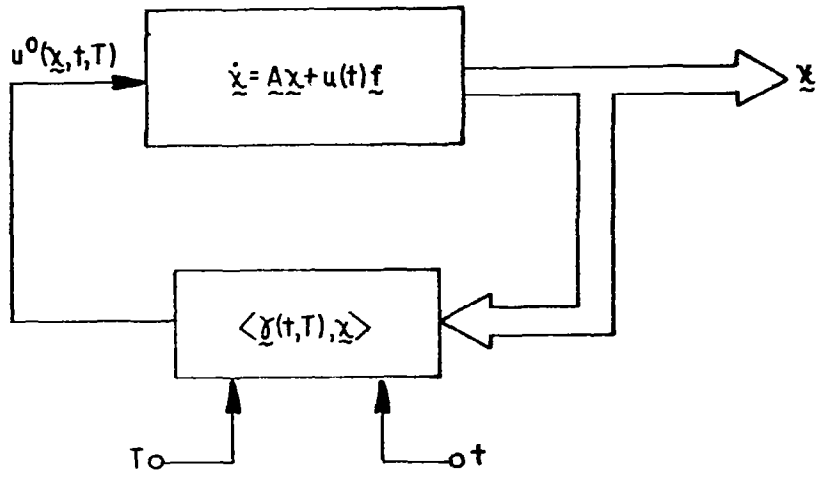

FIGURE 1

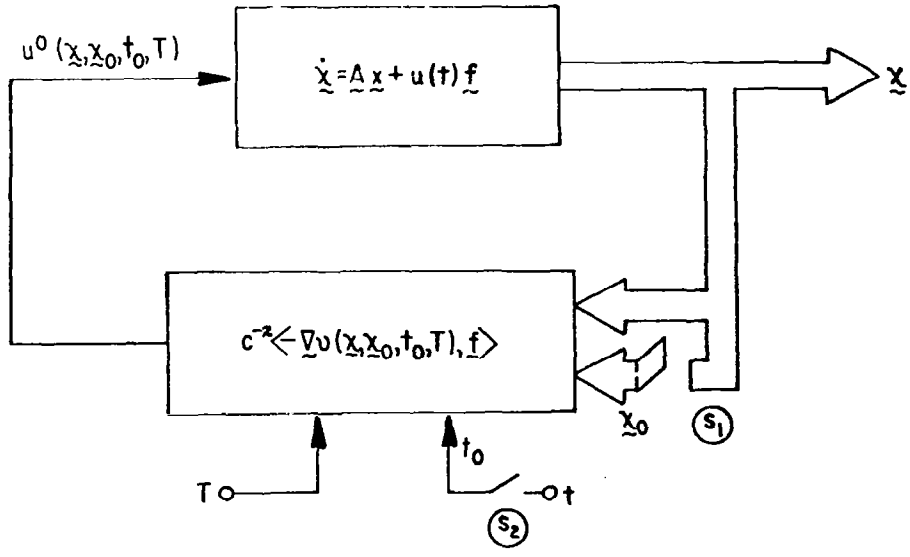

FIGURE 2

sampling the initial conditions $\underset{\sim}{x}(t_o)$, t_o, the actual state $\underset{\sim}{x}(t)$ should deviate from the theoretical optimum trajectory (for instance, due to momentary external disturbances) then the control $u^o(\underset{\sim}{x}, \underset{\sim}{x}_o, t_o, T)$ would not continue to be optimal with respect to the actual state $\underset{\sim}{x}(t)$. However, if such a disturbance should occur, the optimal control for the disturbed state $\underset{\sim}{x}(t)$ can be obtained by momentarily activating the sample and hold devices s_1 and s_2. By this means, the disturbed values $x(t), t$ become the new initial conditions $\underset{\sim}{x}(t_o)$, t_o. This technique allows the step-wise readjustment of the nonlinear feedback law to account for any deviations of $\underset{\sim}{x}(t)$ from the original optimal trajectory.[10] In the limit, as s_1 and s_2 sample $\underset{\sim}{x}(t)$ and t continuously, the control $u^o(\underset{\sim}{x}, \underset{\sim}{x}_o, t_o, T)$ becomes $u^o(\underset{\sim}{x}, t, T)$ and the alternate solutions of Figures 1 and 2 become identical. By employing this dual mode property of the solution (19) it may be possible to eliminate some of the practical disadvantages of the conventional solution (17).

5. Example 1 - A Linear Regulator of First Order

As a special case of (1), (2), let[11]

$$\dot{x} = -x + u(t) \qquad (x = \text{scalar}) \tag{21}$$

$$J[u] = \frac{1}{2} \int_{t_o}^{T} (x(t)^2 + u(t)^2)\, dt \tag{22}$$

with

$$x(t_o) = x_o$$

$$x(T) = 0 \qquad (T = \text{fixed}). \tag{23}$$

The Hamilton-Jacobi equation (7) is

$$\frac{\partial V}{\partial t} - x\frac{\partial V}{\partial x} - \frac{1}{2}\left(\frac{\partial V}{\partial x}\right)^2 + \frac{1}{2}x^2 = 0 \tag{24}$$

---

[10] The readjustment of the nonlinear control law is equivalent to a re-evaluation of the constants of motion (13), (14) for the disturbed state $\underset{\sim}{x}(t)$. This may be viewed as a mechanization of the Principle of Optimality [12].

[11] This example has been considered by Rekasius [10].

and a complete integral of (24) is of the form

$$V = v(x, \beta) + \beta t + \alpha. \tag{25}$$

It can be verified that a complete integral of (24) is

$$V = \frac{1}{2}\left(-x^2 \pm x\sqrt{2(x^2 + \beta)}\right) \pm \frac{\beta}{\sqrt{2}} \operatorname{Ln}\left(\sqrt{2}\, x + \sqrt{2(x^2 + \beta)}\right) + \beta t + \alpha \tag{26}$$

where

$$\alpha = \alpha(x_o, t_o, T)$$

$$\beta = \beta(x_o, t_o, T).$$

From (19) and (23) the optimal control is[12]

$$\phi^o = +x - (\operatorname{sgn} x)\sqrt{2(x^2 + \beta)} \qquad (\operatorname{sgn} 0 \triangleq 0). \tag{27}$$

Substituting (27) in (21), the expression for β is found to be

$$\beta = x_o^2 \operatorname{csch}^2\left[\sqrt{2}\,(T - t_o)\right]. \tag{28}$$

Therefore, the optimal control (27) can be written

$$\phi^o(\underset{\sim}{x}, \underset{\sim}{x}_o, t_o, T) = +x - (\operatorname{sgn} x)\sqrt{2x^2 + 2x_o^2 \operatorname{csch}^2\left[\sqrt{2}\,(T - t_o)\right]}. \tag{29}$$

The field of trajectories for the plant (21) with control (29) is illustrated in Figure 3. A plot of the constant of motion β given by (28) is shown in Figure 4. The contours β = constant in Figure 4 may be interpreted as optimal trajectories. If $T = \infty$ (or, if T = unrestricted) then β in (28) becomes zero and the optimal control (29) becomes

---

[12] Since (21) is only first order, this result can be obtained directly from (24) by observing that $\phi^o = -\partial V/\partial x$.

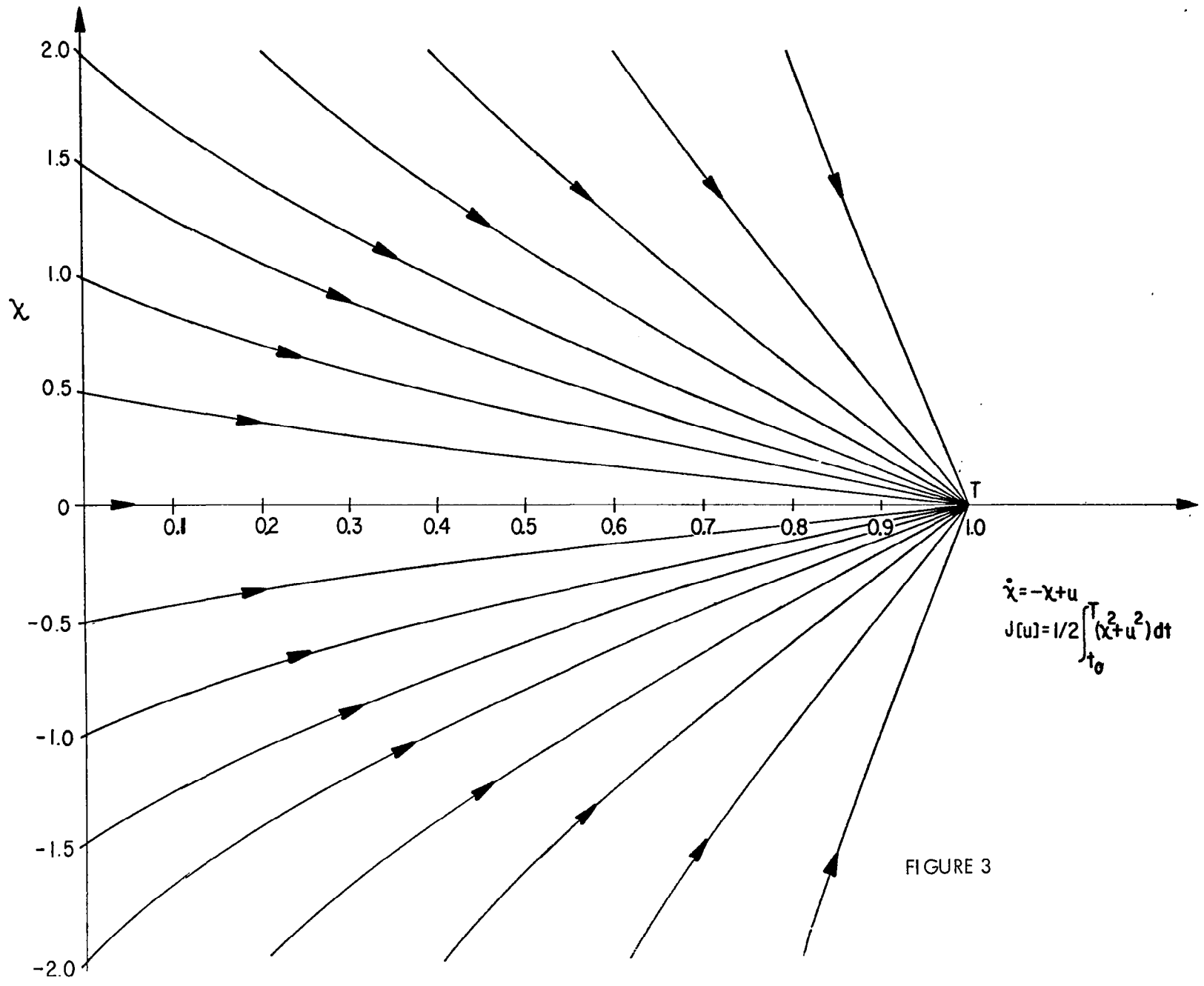
2.0
1.5
1.0
χ
0.5
0
-0.5
-1.0
-1.5
-2.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0
T
$\dot{\chi} = -\chi + u$
$J[u] = 1/2 \int_{t_0}^{T} (\chi^2 + u^2)\, dt$

FIGURE 3

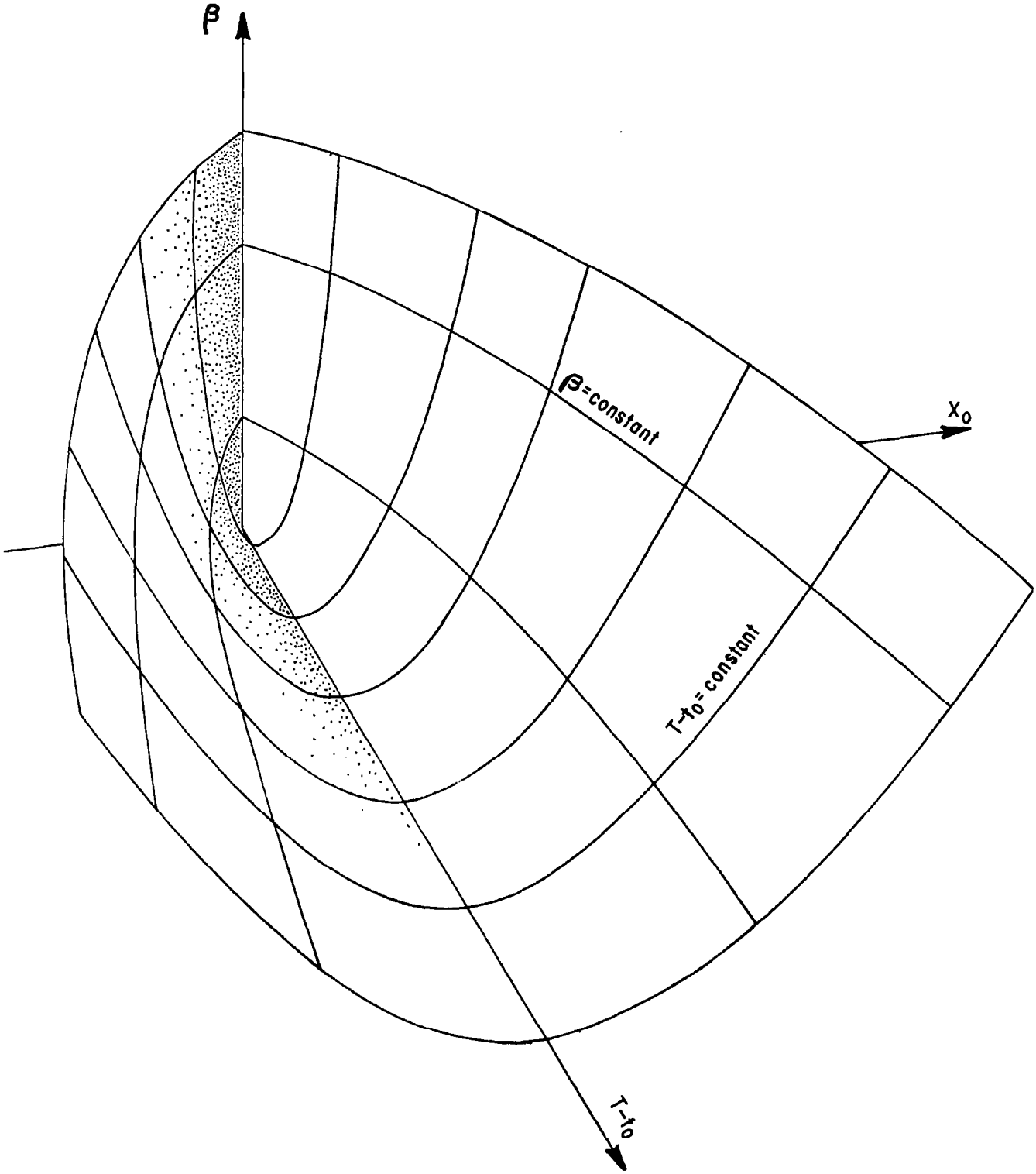

FIGURE 4

$$\phi^{o}(x)\Big|_{T=\infty} = (1 - \sqrt{2})x. \tag{30}$$

The conventional time varying linear control law (17) can be obtained directly from (29) by setting $x_o = x$, $t_o = t$ in (29). The result is

$$\phi^{o}(x, t, T) = [1 - \sqrt{2}\ \mathrm{ctnh}\,[\sqrt{2}\,(T - t)\,]\,]\,x. \tag{31}$$

It may be noted in (31) that $\mathrm{ctnh}\,[\sqrt{2}\,(T - t)] \to \infty$ as $t \to T$.

The expressions for the constants of motion provide a simple method for deriving expressions for x(t). In the present example, (28) yields directly the expression

$$x(t) = x_o \frac{\sinh\,[\sqrt{2}\,(T - t)]}{\sinh\,[\sqrt{2}\,(T - t_o)]}\,. \tag{32}$$

6. Example 2 - A More General First Order Problem

An advantage of the complete integral method of solution is that, in principle, it may be applied to a more general class of problems. Consider, for example, the problem of finding a control u(t) which minimizes the functional

$$J[u] = \int_{t_o}^{T} [g(x(t)) + c^2 u^2(t)\,]\,dt \tag{33}$$

subject to the following conditions:

$$\dot{x} = f(x) + u\,b(x), (x = \text{scalar}) \tag{34}$$

$$x(t_o) = x_o$$

$$x(T) = 0 \qquad (T = \text{fixed}) \tag{35}$$

In (33), g(x) is a non-negative definite continuous function of x and c is a non-zero scalar constant. In (34) f(x) and b(x) are continuous functions of the scalar x.

It is assumed that u(t) is piecewise continuous but otherwise unrestricted. Then, using the complete integral method described above, it is found that the optimal control u^o (if it exists) can be expressed as

$$u^o(\underset{\sim}{x}, \underset{\sim}{x}_o, t_o, T) = -\frac{f(x)}{b(x)} - \frac{(\text{sgn } x)}{b(x)} \sqrt{[f(x)]^2 + c^{-2}[b(x)]^2 [g(x) + \beta]} \quad (36)$$

The $\beta = \beta(x_o, t_o, T)$ in (36) is a constant of motion which can be evaluated from the initial and terminal conditions (35).

If the terminal time T in (35) is unrestricted, then β is zero and (36) becomes

$$u^o(\underset{\sim}{x}) = -\frac{f(x)}{b(x)} - \frac{(\text{sgn } x)}{b(x)} \sqrt{[f(x)]^2 + c^{-2}[b(x)]^2 g(x)} \quad (37)$$

This method of solution can, in principle, be extended to problems of higher order. In the case of higher order problems, it may be necessary to evaluate several of the constants of motion (13), (14), (15).

7. Conclusions

The conventional solution for the optimal control of a linear regulator with quadratic performance index and fixed terminal time leads to a time varying linear control law which is physically unrealizable. It has been shown that the optimal control for this problem can be expressed as a time invariant nonlinear feedback law. Certain parameters in the nonlinear law are functions of the initial conditions $\underset{\sim}{x}(t_o)$, t_o.

The time invariant nonlinear control law can be transformed to the conventional time variable control law by setting $\underset{\sim}{x}(t_o) = \underset{\sim}{x}(t)$ and $t_o = t$. By this means, it may be possible to design a physically realizable optimal controller which retains some of the desirable features of the conventional linear control law. The method of solution used here is applicable, in principle, to a more general class of optimal control problems.

The purpose of the present paper is to point out the possibility of employing the complete integral method in the solution of certain problems of optimal control. Although the principle is relatively simple, application of this method is complicated by the practical difficulties of finding complete integrals.

8. Appendix 1 - Characteristic Strips of the Hamilton-Jacobi Equation

The equations of the characteristic strips of the Hamilton-Jacobi equation (7) are [13]

$$\dot{\underset{\sim}{x}} = \underset{\sim}{A}\underset{\sim}{x} + c^{-2}\langle \underset{\sim}{p}, \underset{\sim}{f}\rangle \underset{\sim}{f}$$

$$\dot{\underset{\sim}{p}} = \underset{\sim}{Q}\underset{\sim}{x} - \underset{\sim}{A}'\underset{\sim}{p} \qquad (' \text{ denotes transpose}) \quad (38)$$

$$\dot{q} = 0 \qquad (39)$$

where

$$\underset{\sim}{p} = -\nabla V$$

$$q = \frac{\partial V}{\partial t}.$$

Equations (38) are equivalent to Pontryagin's canonical equations [14] and (39) is equivalent to the relation $\frac{dH}{dt} = 0$.

When $\underset{\sim}{A}, \underset{\sim}{f}$ are of the canonical form (5), the first of equations (38) can be written as the single n^{th} order differential equation

$$s^n x_1 - \langle \underset{\sim}{a}, \underset{\sim}{\pi}(s)\rangle x_1 - c^{-2} p_n = 0 \qquad (40)$$

and the set of equations (38) can be written as one $2n^{th}$ order linear differential equation of the form

$$s^{2n} x_1 - s^n \langle \underset{\sim}{a}, (\underset{\sim}{\pi}(s) + (-1)^n \underset{\sim}{\pi}(-s))\rangle x_1 + (-1)^n c^{-2} \langle \underset{\sim}{\pi}(s), (\underset{\sim}{Q} + c^2 \underset{\sim}{a}\underset{\sim}{a}')\underset{\sim}{\pi}(-s)\rangle x_1 = 0 \quad (41)$$

where

$$s^k x_1 = \frac{d^k x_1}{dt^k}$$

$$\underset{\sim}{a} = (a_1, \ldots, a_n)$$

$$\underset{\sim}{\pi}(s) = (1, s, s^2, \ldots, s^{n-1}).$$

It may be noted that the 2n roots of the characteristic equation of (41) occur in pairs $(\lambda, -\lambda)$. Equation (40) represents the differential equation of the optimally controlled plant with $u^o(t) = c^{-2} p_n(t)$. Equation (41) may be considered as a $2n^{th}$ order differential equation which is obtained by taking n successive derivatives of (40).

The order of (41) can be reduced to n by taking n successive first integrals of (41) to obtain

$$\theta(x_1, sx_1, \ldots, s^n x_1, c_1, \ldots, c_n) = 0 \qquad (42)$$

where $c_1, \ldots, c_n$ are n constants of integration. Equations (40) and (42) may now be solved jointly to obtain

$$u^o = c^{-2} p_n = \phi^o(\underset{\sim}{x}, c_1, \ldots, c_n). \qquad (43)$$

The constants of integration c_i $(i = 1, \ldots, n)$ are chosen to satisfy the specified boundary conditions $\underset{\sim}{x}(t_o) = \underset{\sim}{x}_o$, $\underset{\sim}{x}(T) = \underset{\sim}{0}$. By this means, there is obtained

$$c_i = c_i(\underset{\sim}{x}_o, t_o, T) \qquad (i = 1, \ldots, n). \quad (44)$$

However, since any state $\underset{\sim}{x}(t)$ along an optimal trajectory can be considered as the instantaneous initial state $\underset{\sim}{x}(t_o)$, the constants of integration in (44) can be written as

$$c_i = c_i(\underset{\sim}{x}, t, T) \qquad (i = 1, \ldots, n). \tag{45}$$

In this way, the control function (43) can be expressed in the form

$$u^o = \phi^o(\underset{\sim}{x}, t, T) \tag{46}$$

which leads to the conventional time varying linear control law (17).

9. Appendix 2 - Complete Integrals of the Hamilton-Jacobi Equation and Constants of Motion Along Optimal Trajectories

The Hamilton-Jacobi equation can be written as [1]

$$\frac{\partial V}{\partial t} - H(\underset{\sim}{x}, -\underset{\sim}{\nabla} V, t, u^o(\underset{\sim}{x}, t, -\underset{\sim}{\nabla} V)) = 0 \tag{47}$$

where $H(\underset{\sim}{x}, \underset{\sim}{p}, t, u^o(\underset{\sim}{x}, t, \underset{\sim}{p}))$ is Pontryagin's Hamiltonian function [14] and $\underset{\sim}{p} = -\underset{\sim}{\nabla} V$ is the so-called conjugate variable.

If $\underset{\sim}{x}$ is an n-vector then, following Lagrange, a complete integral of (47) is defined [13] as any solution of (47) which contains n essential constants of integration a_i $(i = 1, \ldots, n)$. Thus a complete integral of (47) is of the form[13]

$$V = V(x_1, \ldots, x_n, t, a_1, \ldots, a_n) \tag{48}$$

The $2n+1$ canonical or characteristic strip equations associated with (47) are [13]

$$\frac{dx_i(t)}{dt} = \frac{\partial H(\underset{\sim}{x}(t), \underset{\sim}{p}(t), t, u^o(t))}{\partial p_i} \quad ; \quad \frac{dp_i(t)}{dt} = -\frac{\partial H(\underset{\sim}{x}(t), \underset{\sim}{p}(t), t, u^o(t))}{\partial x_i} \qquad (i = 1, \ldots, n) \tag{49}$$

and

$$\frac{dq(t)}{dt} = \frac{\partial H(\underset{\sim}{x}(t), \underset{\sim}{p}(t), t, u^o(t))}{\partial t} \quad ; \quad q = \frac{\partial V}{\partial t} \, . \tag{50}$$

---

[13] Since V does not appear explicitly in (47), it is always possible to append on arbitrary additive constant to a solution (48).

It will now be shown that, if V is sufficiently continuous in its arguments, the expressions

$$\frac{\partial V}{\partial a_i} = k_i \qquad (i = 1, \ldots, n) \tag{51}$$

are constants of motion (first integrals) along solutions of the canonical equations.[14]

The time derivative of (51) is

$$\frac{d}{dt}\left(\frac{\partial V}{\partial a_i}\right) = \frac{\partial^2 V}{\partial t \partial a_i} + \sum_{k=1}^{n} \frac{\partial^2 V}{\partial x_k \partial a_i} \frac{dx_k}{dt} \; . \tag{52}$$

If

$$\frac{\partial^2 V}{\partial t \partial a_i} = \frac{\partial^2 V}{\partial a_i \partial t}$$

and

$$\frac{\partial^2 V}{\partial a_i \partial x_k} = \frac{\partial^2 V}{\partial x_k \partial a_i} \quad (i, k = 1, \ldots, n) \tag{53}$$

then (47) yields

$$\frac{\partial^2 V}{\partial t \partial a_i} = -\sum_{k=1}^{n} \frac{\partial H}{\partial p_k} \frac{\partial^2 V}{\partial x_k \partial a_i} \; . \tag{54}$$

Substituting (54) in (52) there is obtained

$$\frac{d}{dt}\left(\frac{\partial V}{\partial a_i}\right) = \sum_{k=1}^{n} \frac{\partial^2 V}{\partial x_k \partial a_i}\left(\frac{dx_k}{dt} - \frac{\partial H}{\partial p_k}\right) . \tag{55}$$

It is clear from (55) that along solutions of (49)

$$\frac{\partial V}{\partial a_i} = \text{constant} = k_i \qquad (i = 1, \ldots, n). \tag{56}$$

---

[14] This result is well known in classical mechanics [15].

From the expressions $p_i = -\frac{\partial V}{\partial x_i}(x_1, \ldots, x_n, t, a_1, \ldots, a_n)$ and $\frac{\partial V}{\partial a_i}(x_1, \ldots, x_n, t, a_1, \ldots, a_n) = k_i$, there is obtained

$$x_i = x_i(t, a_1, \ldots, a_n, k_1, \ldots, k_n)$$
$$p_i = p_i(t, a_1, \ldots, a_n, k_1, \ldots, k_n) \qquad (57)$$
$$(i = 1, \ldots, n)$$

which constitute a general solution of the 2n canonical equations (49). The 2n constants a_i, k_i $(i = i, \ldots, n)$ are evaluated from the specified initial and terminal conditions of (57).

References

1. R. E. Kalman, Contributions to the Theory of Optimal Control, Boletin de la Sociedad Matematica Mexicana, pp. 111-119, 1960.

2. W. M. Wonham and C. D. Johnson, Optimal Bang-Bang Control with Quadratic Performance Index, Fourth Joint Automatic Control Conference, Minneapolis, Minnesota, 1963; also, ASME Trans., Journal of Basic Engineering, March 1964.

3. R. E. Kalman, The Theory of Optimal Control and the Calculus of Variations, RIAS Technical Report 61-3, pp. 25-27, Baltimore, Maryland.

4. W. M. Wonham, Variational Formulation of Control Problems, Lecture Notes, Control and Information Systems Laboratory, School of Electrical Engineering, Purdue University, Lafayette, Indiana, pp. 86-92, 1962.

5. R. E. Kalman and R. W. Koepcke, Optimal Synthesis of Linear Sampling Control Systems Using Generalized Performance Indices. Trans. ASME, Vol. 80, November 1958, pp. 1820-1826.

6. J. D. Pearson, Approximation Methods in Optimal Control, Journal of Electronics and Control, Vol. 13, No. 5, November 1962, pp. 453-469.

7. C. W. Merriam III, A Class of Optimum Control Systems, Journal of the Franklin Institute, Vol. 267, April 1959, pp. 267-281.

8. C. W. Merriam III, Use of a Mathematical Error Criterion in the Design of Adaptive Control Systems, Trans. AIEE, Vol. 78, pt. II, Jan. 1959, pp. 506-512.

9. E. B. Lee, Design of Optimum Multivariable Control Systems, Trans. ASME, Journal of Basic Engineering, Vol. 83, 1961, pp. 85-90.

10. Z. V. Rekasius, An Alternate Approach to the Fixed Terminal Point Regulator Problem, 9th Annual East Coast Conference of IRE and PGANE, October 24, 1962, Baltimore, Maryland; also, IEEE Transactions PTGAC, Vol. AC-9, July, 1964.

11. C. D. Johnson and W. M. Wonham, On a Problem of Letov in Optimal Control, Fifth Joint Automatic Control Conference, Stanford, California, June 1964. To appear, ASME Trans., Journal of Basic Engineering. March, 1965.

12. R. E. Bellman, Dynamic Programming, Princeton University, Princeton, New Jersey, 1957.

13. E. Goursat, A Course in Mathematical Analysis, Vol. II, Part 2, Dover Publications, New York, 1959.

14. L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze and E. F. Mishchenko, The Mathematical Theory of Optimal Processes, John Wiley and Sons, New York, 1962.

15. H. Goldstein, Classical Mechanics, Addison Wesley, 1950, pp. 282-283.

VIII. On a Problem of Letov in Optimal Control[1]

C. D. JOHNSON
Electrical Engineering Department,
University of Alabama,
Huntsville Center, Huntsville, Ala.

W. M. WONHAM
Center for Control Theory,
RIAS, Baltimore, Md.

In a series of papers [1, 2],[2] A. M. Letov discussed an optimal regulator problem for a linear plant with bounded control variable and quadratic performance index. This problem was also discussed by Chang [3]. Krasovskii and Letov observed later [4] that the solution proposed in [1, 2, and 3] may be correct only for special choices of the initial value of the state vector. In the present note, further aspects of the solution in the general case are described and three examples are given. The possible existence of a regime of unsaturated-nonlinear optimal control is demonstrated. The presence of this regime in the optimal control law was apparently overlooked in [1–4].

Statement of the Problem

THE problem is to find a continuous control function $u = u(t)$ which minimizes the functional[3]

$$J[u] = \frac{1}{2}\int_0^T [\langle \mathbf{x}(t), \mathbf{Q}\mathbf{x}(t)\rangle + c^2u^2(t)]dt \tag{1}$$

subject to the following conditions:

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + u\mathbf{f} \qquad (\cdot = d/dt) \tag{2}$$

$$\mathbf{x}(0) = \mathbf{x}_0 \tag{3}$$

$$\mathbf{x}(T) = \mathbf{0} \qquad (T \text{ is unrestricted}) \tag{4}$$

$$|u(t)| \leqq 1 \qquad (0 \leqq t \leqq T) \tag{5}$$

In (1), $\mathbf{Q}$ is a positive semidefinite symmetric constant matrix and c is a scalar constant. In (2), $\mathbf{x} = (x_1, \ldots, x_n)$ is the state vector of the plant, $\mathbf{A}$ is an $(n \times n)$ constant matrix, $\mathbf{f} = (f_1, \ldots, f_n)$ is a constant n-vector and u is the scalar control function. It is assumed that the pair $(\mathbf{A}, \mathbf{f})$ is controllable; that is, the vectors

$$\mathbf{f}, \mathbf{A}\mathbf{f}, \ldots, \mathbf{A}^{n-1}\mathbf{f} \tag{6}$$

are linearly independent. Then, as shown in [5], there is no loss of generality in assuming that $\mathbf{A}$, $\mathbf{f}$ have the form

$$\mathbf{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ 0 & \cdots & 0 & & 1 \\ a_1 & \cdots & & & a_n \end{bmatrix}, \quad \mathbf{f} = \begin{bmatrix} 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix} \tag{7}$$

It is further assumed that the state $\mathbf{x} = \mathbf{0}$ is reachable from $\mathbf{x}_0$ using an admissible control (5). Then, if $\mathbf{Q}$ is appropriately restricted, an optimal control u exists [9] and can in principle be found by straightforward application of Pontryagin's principle or dynamic programming. Since the details have already been given in [2 and 4] we first summarize the results.

Form of the Optimal Control Law

1 Let $u^0 = \phi^0(\mathbf{x})$ be an optimal control law, and let

$$V(\mathbf{x}) = J[u^0], \qquad \mathbf{x}(0) = \mathbf{x} \tag{8}$$

We shall assume that V is continuously differentiable in $\mathbf{x}$. Then V satisfies the Hamilton-Jacobi equation

$$\langle \nabla V(\mathbf{x}), \mathbf{A}\mathbf{x}\rangle - \frac{c^{-2}}{2}\langle \nabla V(\mathbf{x}), \mathbf{f}\rangle^2 + \frac{1}{2}\langle \mathbf{x}, \mathbf{Q}\mathbf{x}\rangle = 0,$$
$$\text{if} \quad |c^{-2}\langle \nabla V(\mathbf{x}), \mathbf{f}\rangle| \leqq 1 \tag{9a}$$

and

$$\langle \nabla V(\mathbf{x}), \mathbf{A}\mathbf{x}\rangle - |\langle \nabla V(\mathbf{x}), \mathbf{f}\rangle| + \frac{c^2}{2} + \frac{1}{2}\langle \mathbf{x}, \mathbf{Q}\mathbf{x}\rangle = 0,$$
$$\text{if} \quad |c^{-2}\langle \nabla V(\mathbf{x}), \mathbf{f}\rangle| \geqq 1 \tag{9b}$$

In (9)

$$\nabla V = \left(\frac{\partial V}{\partial x_1}, \ldots, \frac{\partial V}{\partial x_n}\right) \tag{10}$$

Further, the optimal control is given by

$$\phi^0(\mathbf{x}) = \text{sat}[c^{-2}\langle -\nabla V(\mathbf{x}), \mathbf{f}\rangle] \tag{11}$$

where

$$\text{sat } y = \begin{cases} y, & |y| \leqq 1 \\ \text{sgn } y, & |y| \geqq 1 \end{cases} \tag{12}$$

Equation (9a) holds in the set of states $\mathbf{x}$ where the control is unsaturated (i.e., $|\phi^0(\mathbf{x})| < 1$) and (9b) holds in the set of states $\mathbf{x}$ where the control is saturated (i.e., $|\phi^0(\mathbf{x})| = 1$). Consider first the set where (9a) is satisfied. In the absence of the constraint (5), the restriction $|c^{-2}\langle \nabla V(\mathbf{x}), \mathbf{f}\rangle| \leqq 1$ disappears and (9a) holds at all states $\mathbf{x}$. For this case it is well known [6] that the solution of (9a) is

$$V(\mathbf{x}) = \frac{1}{2}\langle \mathbf{x}, \mathbf{M}\mathbf{x}\rangle \tag{13}$$

where the matrix $\mathbf{M}$ is symmetric, positive definite, and uniquely defined by

$$\mathbf{A}'\mathbf{M} + \mathbf{M}\mathbf{A} - c^{-2}\mathbf{M}\mathbf{f}\mathbf{f}'\mathbf{M} + \mathbf{Q} = \mathbf{0}$$
($'$ denotes transpose) (14)

Further, the optimal control law is linear and is given by

$$\phi_L(\mathbf{x}) = c^{-2}\langle -\nabla V(\mathbf{x}), \mathbf{f}\rangle = \langle \boldsymbol{\gamma}, \mathbf{x}\rangle \tag{15}$$

where

$$\boldsymbol{\gamma} = -c^{-2}\mathbf{M}\mathbf{f} \tag{16}$$

2 To introduce the constraint $|u| \leqq 1$ we proceed as in [5]. Let L be the set of states $\mathbf{x}_0$ such that, if $\mathbf{x}(0) = \mathbf{x}_0$ and if $u = \phi_L(\mathbf{x})$ in (2), then $|\phi_L[\mathbf{x}(t)]| \leqq 1$ for $t \geqq 0$. In other words, L is the set of initial states for which the constraint $|u| \leqq 1$ is satisfied along the corresponding trajectories when the control law is $\phi_L(\mathbf{x})$. It is clear that if $\mathbf{x}(0) \in L$ then $\mathbf{x}(t) \in L$ for $t \geqq 0$; and it can be verified that L is an n-dimensional, convex, and in general *proper* subset of the set of states $\mathbf{x}$ defined by the inequality

[1] This research was supported in part at the respective institutions by the National Aeronautics and Space Administration under Grant No. NsG-381 (and Contract No. NAS 8-11231) and Grant No. NASw-845.

[2] Numbers in brackets designate References at end of paper.

[3] $\langle \mathbf{x}, \mathbf{y}\rangle$ is the scalar product of $\mathbf{x}$ and $\mathbf{y}$.

$$|\phi_L(\mathbf{x})| \leqq 1 \tag{17}$$

Obviously L contains the origin $\mathbf{x} = \mathbf{0}$. It follows that

$$\phi^0(\mathbf{x}) = \phi_L(\mathbf{x}), \quad \text{when} \quad \mathbf{x} \in L \tag{18}$$

In general $\phi^0(\mathbf{x})$ is not given by (18) in the entire strip (17); the set L coincides with the strip (17) only for special choices of the matrices $\mathbf{A}$ and $\mathbf{Q}$ (see section, "Example 3"). Moreover $\phi^0(\mathbf{x})$ is not given, in general, by the simple rule

$$\phi^0(\mathbf{x}) = \text{sat } \phi_L(\mathbf{x}), \quad \text{for all } \mathbf{x} \tag{19}$$

The control law (19) is the solution (in general, incorrect) proposed in [1-3].

The set L is the largest set of states $\mathbf{x}$ for which one can state *a priori* that (13) is a valid solution of (9*a*).[4] In general (9*a*) has solutions different from (13) which are valid in a certain region N. For $\mathbf{x}$ in N, the optimal control is again unsaturated, i.e., $|c^{-2}\langle \nabla V(\mathbf{x}), \mathbf{f}\rangle| < 1$, but $\phi^0(\mathbf{x})$ is a nonlinear function $\phi_N(\mathbf{x})$ of the state $\mathbf{x}$. The possible existence of N was apparently overlooked in [1-4]. Failure to include the nonlinear regime, when it exists, will lead to the apparent discontinuities in $V(\mathbf{x})$ which were described in the last paragraph of [4].

The interior of $L \cup N$ is the set of states $\mathbf{x}$ where the optimal control $\phi^0(\mathbf{x})$ is unsaturated ($|\phi^0(\mathbf{x})| < 1$). In the complement of $L \cup N$, which we shall denote by S, the control is saturated; i.e., $|c^{-2}\langle \nabla V(\mathbf{x}), \mathbf{f}\rangle| > 1$ and $\phi^0(\mathbf{x}) = \phi_S(\mathbf{x}) = \pm 1$. In principle, analytic expressions for $\phi_N(\mathbf{x})$ and the boundaries of N and S are determined by solving (9) in N and S and then applying (11). In practice this procedure is complicated by the fact that the solution of (9) does not have the same analytic form throughout N and S. Some results of applying this procedure to a concrete example are given later in Example 1. Further research is needed to determine more practical methods for obtaining, or approximating, $\phi_N(\mathbf{x})$.

3 The theory of characteristic strips [7] suggests an alternative and practical technique for determining the boundaries of N and S. In this technique the equations of the characteristic strips of (9) are integrated in reversed time, starting from states on the boundary of L. For this problem the equations of the characteristic strips are equivalent to Pontryagin's canonical equations [8] and are

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + u^0\mathbf{f} \\ u^0(t) &= \text{sat } [c^{-2}\langle \mathbf{p}(t), \mathbf{f}\rangle] \\ \dot{\mathbf{p}} &= \mathbf{Q}\mathbf{x} - \mathbf{A}'\mathbf{p} \end{aligned} \tag{20}$$

where $\mathbf{p} = -\nabla V(\mathbf{x})$. At states $\mathbf{x} \in L$ (13) yields

$$\mathbf{p} = -\mathbf{M}\mathbf{x} \tag{21}$$

where $\mathbf{M}$ is given by (14). Integration of (20) for $t \leqq 0$ (with initial conditions $\mathbf{x}(0) = \mathbf{x}$, $\mathbf{p}(0) = -\mathbf{M}\mathbf{x}$ and $\mathbf{x}$ on the boundary of L) presents no problem in principle, since the "sat" function is Lipschitz-continuous. In this way states on the common boundary of N and S are determined as the values of $\mathbf{x}(t)$ when $c^{-2}\langle \mathbf{p}(t), \mathbf{f}\rangle = \pm 1$.[5]

Example 1

As a special case of (1), (2), let

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u, \qquad |u(t)| \leqq 1 \end{aligned} \tag{22}$$

$$J[u] = \frac{1}{2}\int_0^T (x_1^2 + x_2^2 + c^2u^2)dt \tag{23}$$

The linear control law is found from (14), (15) to be

$$\phi_L(\mathbf{x}) = -c^{-1}x_1 - c^{-1}(1 + 2c)^{1/2}x_2 \tag{24}$$

The set L is the largest subset of the strip $|\phi_L(\mathbf{x})| \leqq 1$ which is invariant for the system

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= \phi_L(\mathbf{x}) \qquad (t \geqq 0) \end{aligned} \tag{25}$$

Thus L is bounded in part by the straight lines $\phi_L(\mathbf{x}) = \pm 1$, and in part by the two trajectories of the system (25) which are tangent to these lines, Fig. 1. A trajectory of (25) is tangent to the line $\phi_L(\mathbf{x}) = \pm 1$ at the state $\pm\mathbf{x}_\theta$ where

$$\mathbf{x}_\theta = [1 + c, \quad -(1 + 2c)^{1/2}] \tag{26}$$

The state $\pm\mathbf{x}_t$, at which a trajectory tangent to $\phi_L(\mathbf{x}) = \pm 1$ intersects the opposite boundary $\phi_L(\mathbf{x}) = \mp 1$ is determined by integration of (25) for $t \leqq 0$, with $\mathbf{x}(0) = \pm\mathbf{x}_\theta$.

The boundaries of N and S are now established by integrating the canonical equations for $t \leqq 0$. The matrix $\mathbf{M}$ is found from (14) to be

$$\mathbf{M} = \begin{bmatrix} (1 + 2c)^{1/2} & c \\ c & c(1 + 2c)^{1/2} \end{bmatrix} \tag{27}$$

and the canonical equations (20) become

$$\begin{aligned} \dot{x}_1 &= x_2 & \dot{p}_1 &= x_1 \\ \dot{x}_2 &= \text{sat } (c^{-2}p_2) & \dot{p}_2 &= -p_1 + x_2 \end{aligned} \tag{28}$$

The integration of (28) for $t \leqq 0$ need be started only from states $\mathbf{x}$ on the linear boundary segments $[-\mathbf{x}_t, \mathbf{x}_\theta]$ and $[\mathbf{x}_t, -\mathbf{x}_\theta]$.[6] The corresponding initial values of p_1, p_2 are given by (21) and (27). For $c = 1$, the results, obtained with an analog computer, are shown in Fig. 2. As an optimal trajectory is traced backward from L, the state trajectory first enters S, where the control $u^0(t) = \text{sat } [c^{-2}p_2(t)]$ remains constant at the saturation level ± 1. The trajectory then passes through N (the curved strip in Fig. 2), where $u^0(t)$ varies continuously from ± 1 to ∓ 1; and so on. As shown in Fig. 2 the set S can be divided into two subsets $S_\pm$ where $\phi^0(\mathbf{x}) = \pm 1$. The behavior of $p_2(t)$ and of $u^0(t)$, $t \leqq 0$, is illustrated in Fig. 3.

As $c \to 0$ the set L reduces to the linear segment $x_1 + x_2 = 0$, $|x_1| \leqq 1$, and the boundaries of N approach a common switching curve. The optimal control then has a bang-bang and a singular mode, Fig. 4. The general problem (1)-(5) with $c = 0$ has been discussed in [5].

The following results (for $c = 1$) were obtained by analytic solution of the Hamilton-Jacobi equation. From (9),

$$x_2\frac{\partial V}{\partial x_1} - \frac{1}{2}\left(\frac{\partial V}{\partial x_2}\right)^2 + \frac{1}{2}(x_1^2 + x_2^2) = 0, \qquad \left|\frac{\partial V}{\partial x_2}\right| \leqq 1 \tag{29a}$$

$$x_2\frac{\partial V}{\partial x_1} - \left|\frac{\partial V}{\partial x_2}\right| + \frac{1}{2} + \frac{1}{2}(x_1^2 + x_2^2) = 0, \qquad \left|\frac{\partial V}{\partial x_2}\right| \geqq 1 \tag{29b}$$

In the set L, the solution of (29*a*) is given by (13) and (27):

$$V_L(\mathbf{x}) = \frac{1}{2}(\sqrt{3}x_1^2 + 2x_1x_2 + \sqrt{3}x_2^2) \tag{30}$$

In the subset S_1 of S, shown in Fig. 5, (29*b*) has the solution

$$V_{S_1}(\mathbf{x}) = (1/30)[-3\sqrt{3} + 15x_1^2x_2 + 10x_1x_2^3 + 15x_2 + 5x_2^3 + 2x_2^5 + 2|(1 + 2x_1 + x_2^2)^{5/2}|]. \tag{31}$$

[4] Here the term "valid" means that $V(\mathbf{x})$ satisfies the definition (8).

[5] The computation of $u^0(t)$, $t \leqq 0$, also allows one to determine numerical values of $V(\mathbf{x})$. This information may be useful in comparing various suboptimal control laws.

[6] Integration need not be started from the segments $(\pm\mathbf{x}_t, \pm\mathbf{x}_\theta)$ since these are characteristic curves of (9).

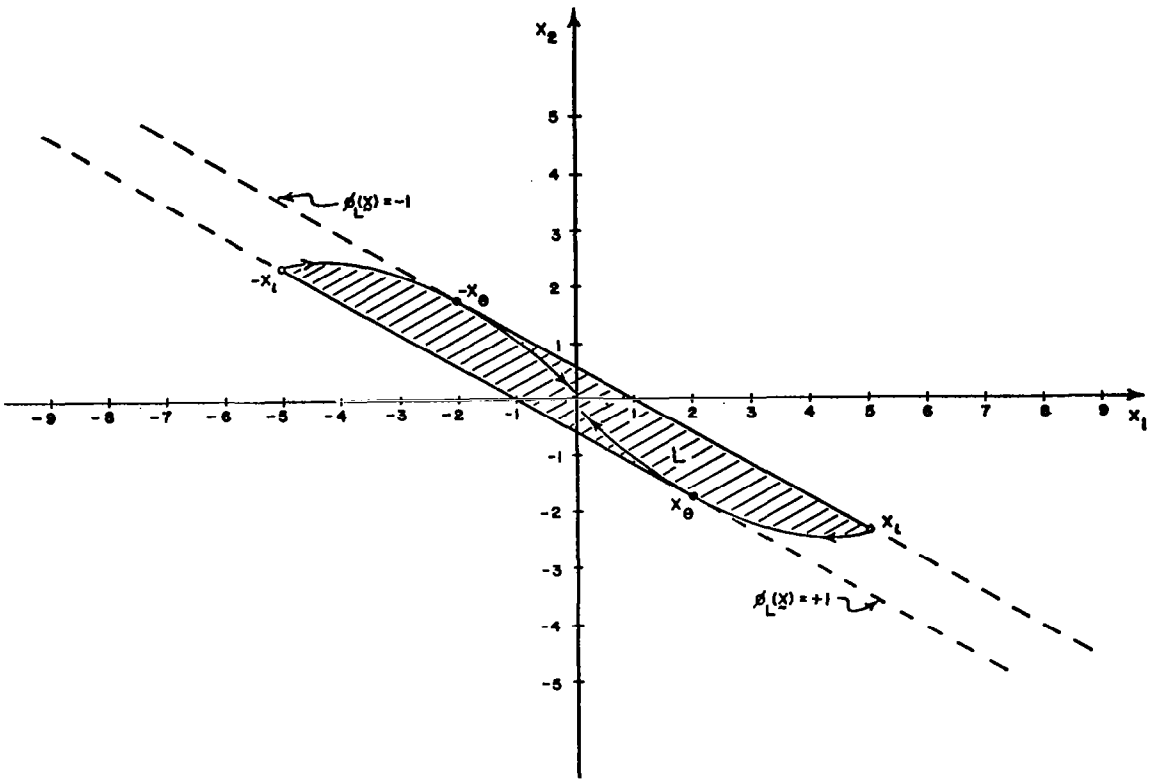

Fig. 1 Construction of set L for Example 1

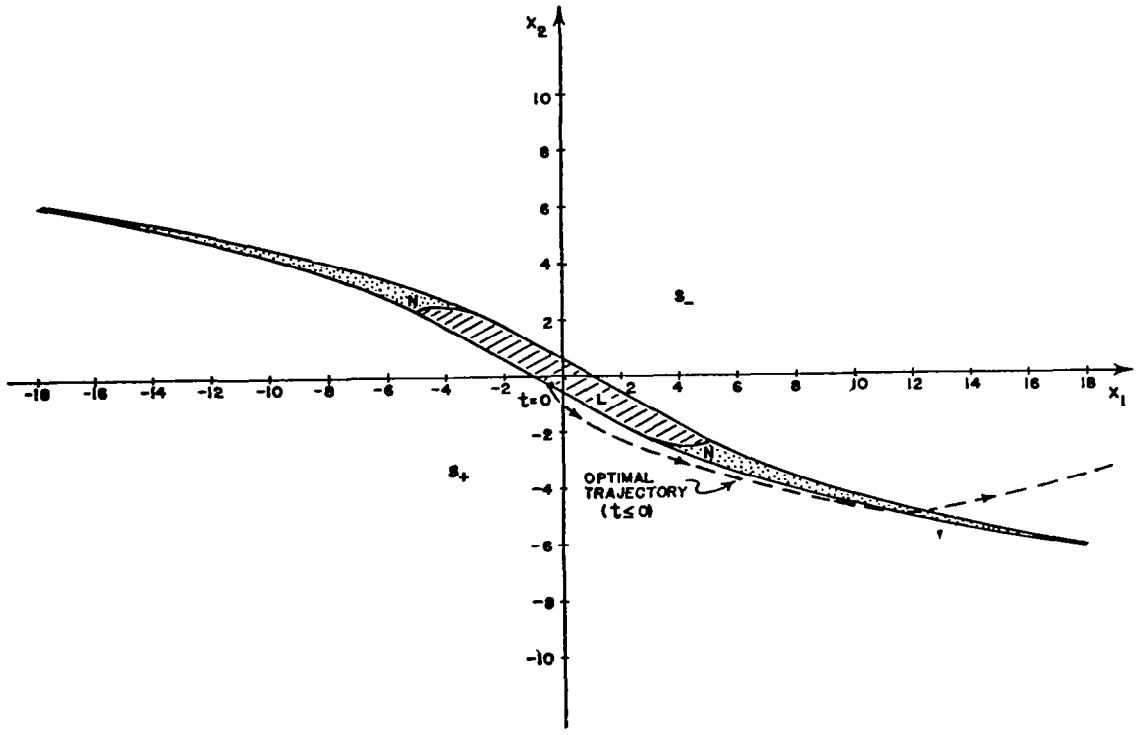

Fig. 2 L, N, and S-regions for Example 1

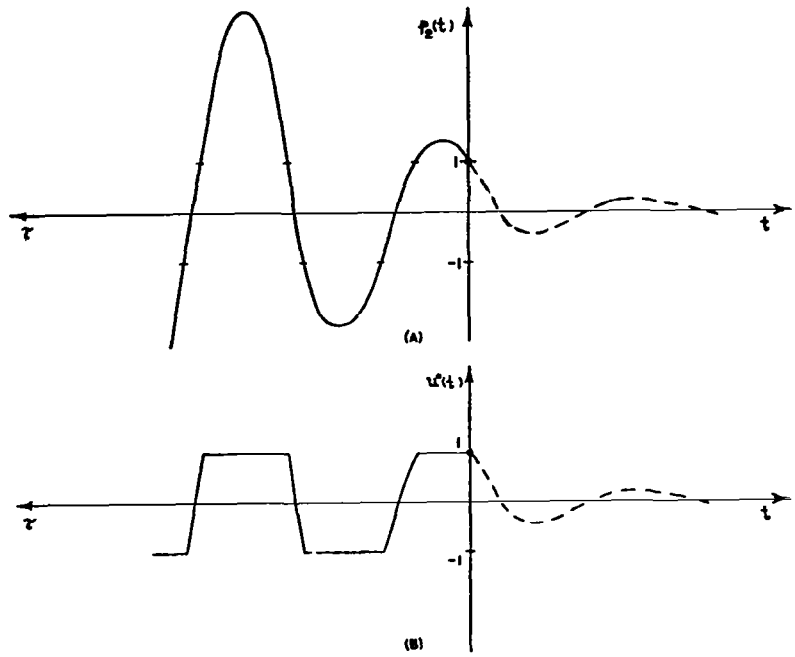

Fig. 3 Typical plots of $p_2(t)$ and $u^0(t)$ for $t \leq 0$. (a) $p_2(t)$; (b) $u^0(t)$

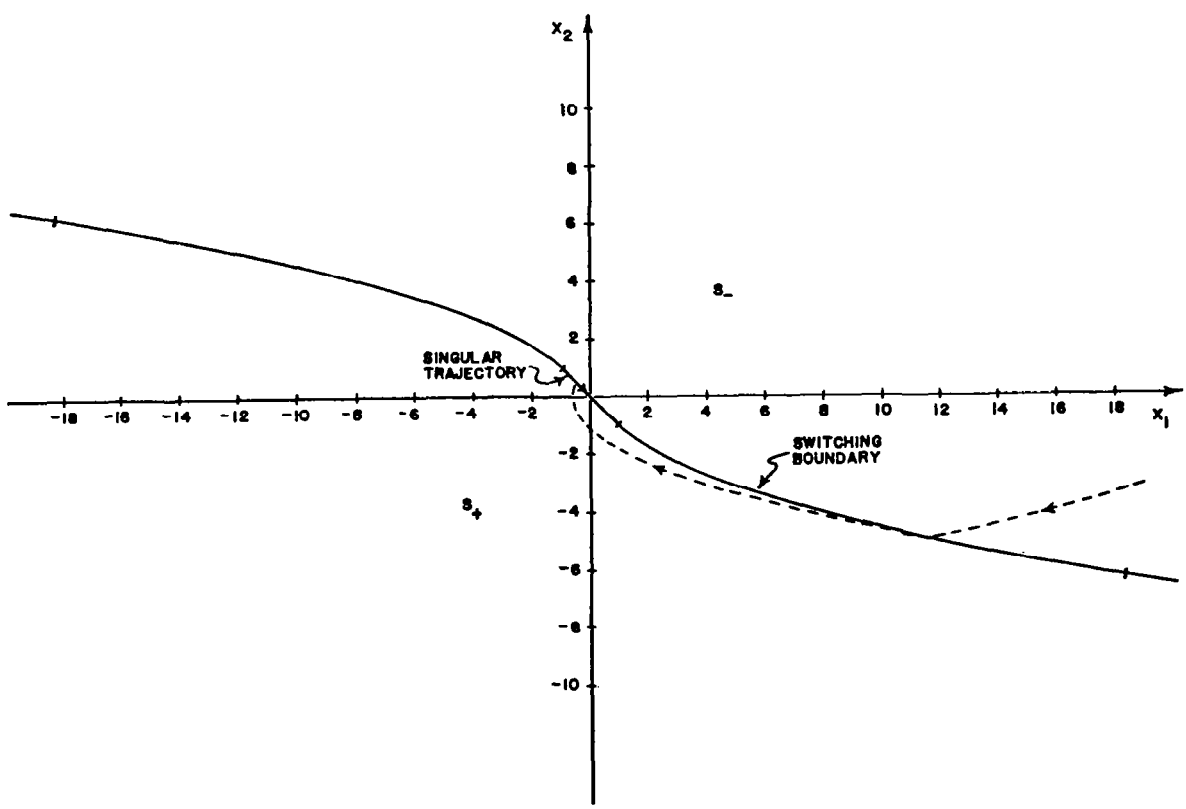

Fig. 4 Results for $c = 0$ in Example 1

It can be verified that the solution (31) satisfies the requirement that V be continuous on the (L, S_1) boundary, defined by $x_1 + \sqrt{3}x_2 = 1$, $|x_2| \leqq \sqrt{3}$.

The boundary segment between $L \cup N$ and S_1, Fig. 5, is determined by setting $\partial V(\mathbf{x})/\partial x_2 = +1$ in (31). The result is

$$-3 + 3x_1^2 + 6x_1x_2^2 + 3x_2^3 + 2x_2^4 + 2x_2|(1 + 2x_1 + x_2^2)^{3/2}| = 0 \quad (32)$$

If $|x_2| \leqq \sqrt{3}$, the left side of (32) vanishes when $x_1 + \sqrt{3}x_2 = 1$; if $x_2 \geqq \sqrt{3}$, the mode boundary (N, S_1) defined by (32) is a curved segment lying between the parabola $1 + 2x_1 + x_2^2 = 0$ and the line $x_1 + \sqrt{3}x_2 = 1$. The fact that the boundary segment (N, S_1) is not linear shows that the control law in the region N must be a nonlinear function of the state variables.

Example 2

The following example was discussed briefly in [4]. Let

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= -x_1 + u \qquad |u(t)| \leqq 1 \\ J[u] &= \frac{1}{2}\int_0^T (x_2^2 + u^2)dt \end{aligned} \quad (33)$$

The construction of the L, N, and S boundaries for this problem proceeds in the same manner as for Example 1 and the results are shown in Fig. 6. The trajectories in $S+(S-)$ are circular arcs centered at $x_1 = +1(-1)$, $x_2 = 0$.

Example 3 Optimal Dual-Mode Control

As mentioned earlier in the paper, the set L coincides with the entire strip $|\phi_L(\mathbf{x})| \leqq 1$ only for special choices of the matrices $\mathbf{A}$ and $\mathbf{Q}$. Expressed geometrically, a necessary condition is that the hyperplane $\phi_L(\mathbf{x}) = 0$ be invariant for the linear system defined by setting $u = \phi_L$ in (2). Sufficient conditions can be simply expressed algebraically when $\mathbf{A}$, $\mathbf{f}$ are of canonical form (7), and in this case we have the following result.

Theorem

Let the eigenvalues of $\mathbf{A}$ *be* $\alpha_1, \ldots, \alpha_n$. *If* (i) Re $\alpha_m < 0$ *and the* α_m *are distinct,* $m = 1, \ldots, n-1$, (ii) α_n *is real,* (iii) $\mathbf{Q} = [q_{jk}]$ *is so chosen that* $q_{11} > 0$, $\mathbf{Q}$ *is symmetric positive semidefinite, and*

$$\sum_{j=1}^{n}\sum_{k=1}^{n} q_{jk}\alpha_m^{j-1}(-\alpha_m)^{k-1} = 0, \qquad m = 1, \ldots, n-1 \quad (34)$$

then the optimal control law is

$$\phi^0(\mathbf{x}) = \text{sat } [\phi_L(\mathbf{x})] \quad (35)$$

If $\alpha_n \leqq 0$, *the origin* $\mathbf{x} = \mathbf{0}$ *is reachable from all* $\mathbf{x}_0$; *if* $\alpha_n > 0$ *the origin is reachable from* $\mathbf{x}_0$ *if, and only if,*

$$|\phi_L(\mathbf{x}_0)| < 1 + [1 + q_{11}(c\,a_1)^{-2}]^{1/2}$$

A proof is given in the Appendix.

Under the conditions of the theorem the optimal control law is of the very simple "dual-mode" form proposed in [1–3].

If $n = 1$ the conditions hold trivially.

For $n \geqq 2$ the condition Re $\alpha_m < 0$ $(m = 1, \ldots, n-1)$ is somewhat restrictive and cannot be relaxed. However, it is always possible to choose $\mathbf{Q}$ such that (iii) is satisfied; for instance, choose real numbers $e_1 = 1, e_2, \ldots, e_n$ such that

$$\sum_{j=1}^{n} e_j\alpha_m^{j-1} = 0 \qquad (m = 1, \ldots, n-1)$$

and put

$$q_{ij} = e_ie_j \qquad (i, j = 1, \ldots, n)$$

If the conditions of the theorem are satisfied, then as $c \to 0$, the strip $|\phi_L(\mathbf{x})| \leqq 1$ reduces to the $(n-1)$-dimensional hyperplane $\phi_L(\mathbf{x}) = 0$ and the optimal control law becomes

$$\phi^0(\mathbf{x}) = \begin{cases} \text{sgn } \phi_L(\mathbf{x}), & \phi_L(\mathbf{x}) \neq 0 \\ 0, & \phi_L(\mathbf{x}) = 0 \end{cases} \quad (36)$$

Here the control $\phi^0(\mathbf{x}) = 0$ is singular. This case of optimal linear switching has been discussed in [5].

As an application of the theorem let

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_1 + u, \qquad |u(t)| \leqq 1 \end{aligned} \quad (37)$$

$$J[u] = \frac{1}{2}\int_0^T (x_1^2 + x_2^2 + c^2u^2)dt \quad (38)$$

By (35) the optimal control is

$$\phi^0(\mathbf{x}) = -\text{sat } [\{1 + (1 + c^{-2})^{1/2}\}(x_1 + x_2)] \quad (39)$$

and the origin is reachable from $\mathbf{x} = (x_1, x_2)$ provided

$$|x_1 + x_2| < 1 \quad (40)$$

The results, for $c = 1$, are shown in Fig. 7.

As $c \to 0$ the strip L in Fig. 7 reduces to the line $x_1 + x_2 = 0$ and the optimal control law becomes

$$\phi^0(\mathbf{x}) = \begin{cases} -\text{ sgn } (x_1 + x_2), & 0 < |x_1 + x_2| < 1 \\ 0, & x_1 + x_2 = 0 \end{cases} \quad (41)$$

The results for this case are shown in Fig. 8.

Conclusions

The linear-saturation control law proposed in [1–3] is correct only for special choices of the problem parameters. Sufficient conditions are obtained for validity of this law. In general the optimal control law has three modes; namely, linear, nonlinear, and saturation. Some aspects of the general case have been illustrated with an example. A scheme has been proposed for computing the boundaries of the regions of linear, nonlinear and saturated optimal control. Further research is needed to determine explicit expressions or suitable approximations for the nonlinear control law.

The examples suggest some interesting theoretical problems. One is to obtain a more explicit description of the regions L and N. Another is to relate the mode boundaries with the switching surface of time-optimal control.

References

1 A. M. Letov, "Analytic Controller Design II," *Automation and Remote Control*, vol. 21, 1960, pp. 389–393.

2 A. M. Letov, "The Analytical Design of Control Systems," *Automation and Remote Control*, vol. 22, 1961, pp. 363–372.

3 Chang Jen-Wei, "A Problem in the Synthesis of Optimal Systems Using the Maximum Principle," *Automation and Remote Control*, vol. 22, 1961, pp. 1170–1176.

4 N. N. Krasovskii and A. M. Letov, "The Theory of Analytic Design of Controllers," *Automation and Remote Control*, vol. 23, 1962, pp. 649–656.

5 W. M. Wonham and C. D. Johnson, "Optimal Bang-Bang Control With Quadratic Performance Index," Proceedings, Joint Automatic Control Conference, Minneapolis, Minn., 1963; JOURNAL OF BASIC ENGINEERING, TRANS. ASME, Series D, vol. 86, 1964, pp. 107–115.

6 A. M. Letov, "Analytical Controller Design I," *Automation and Remote Control*, vol. 21, 1960, pp. 303–306.

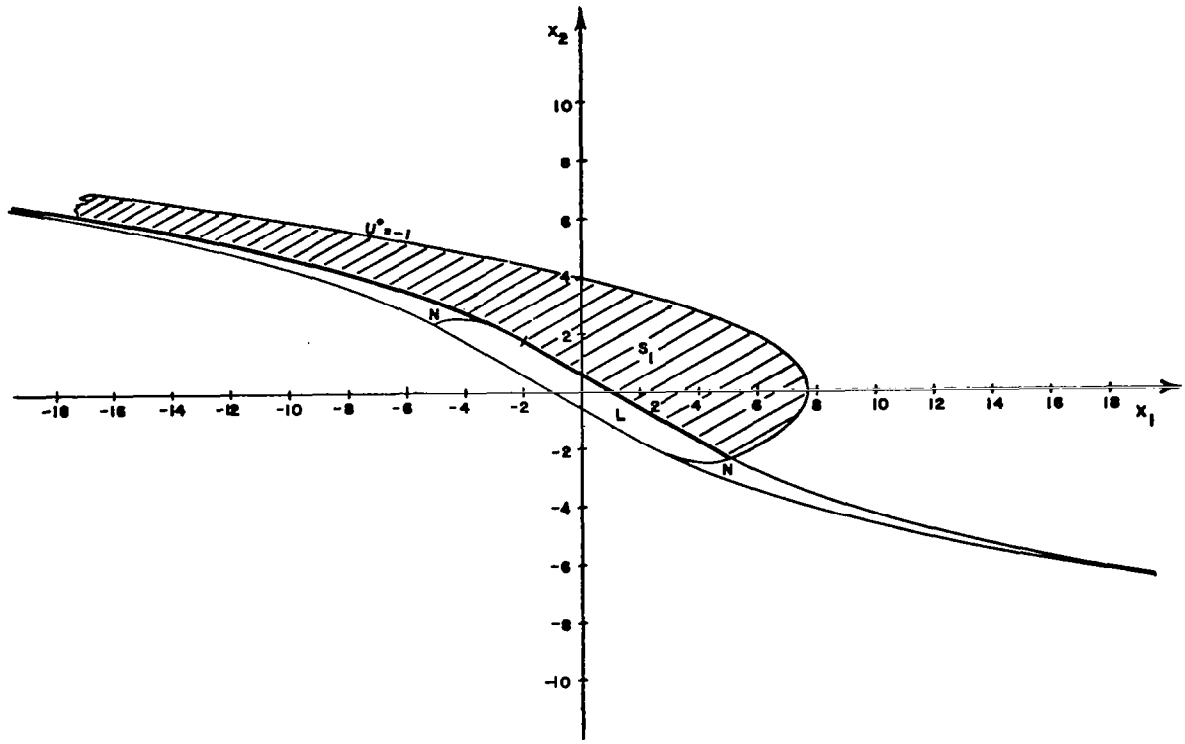

Fig. 5 Subset S_1 in which (31) is a valid solution of (29b)

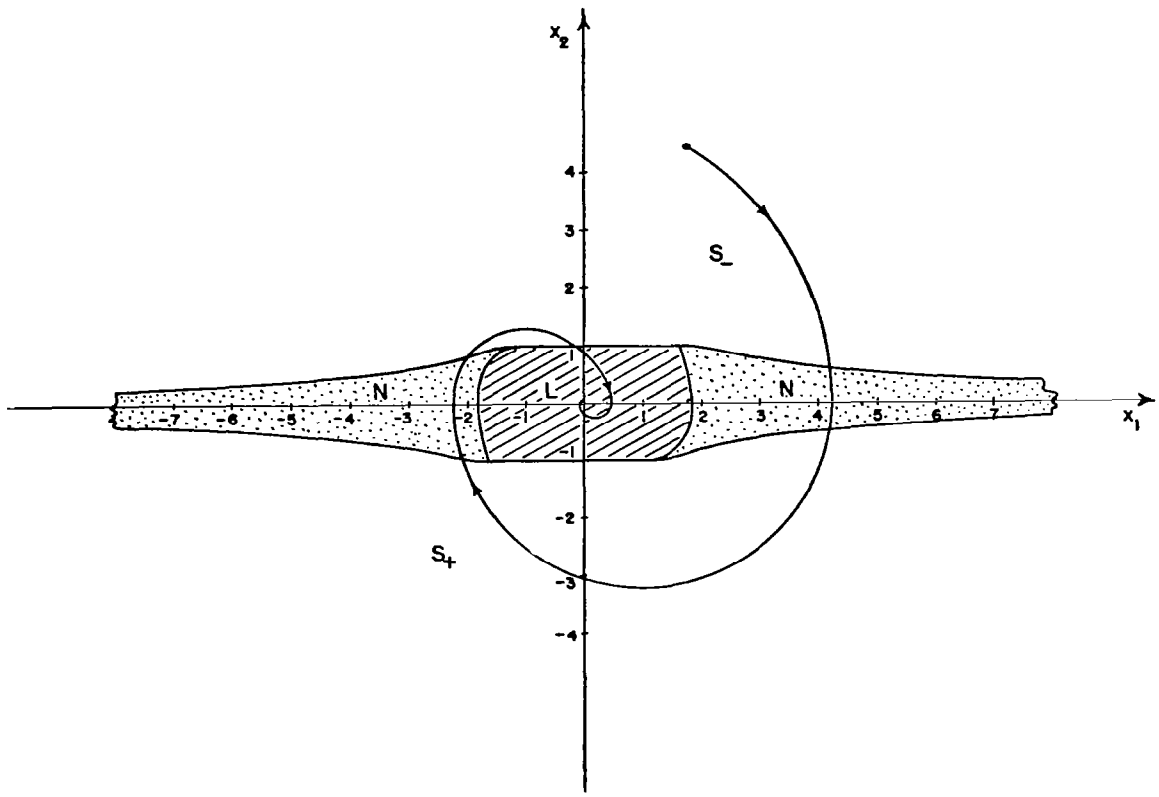

Fig. 6 *L*, *N*, and *S*-regions for Example 2

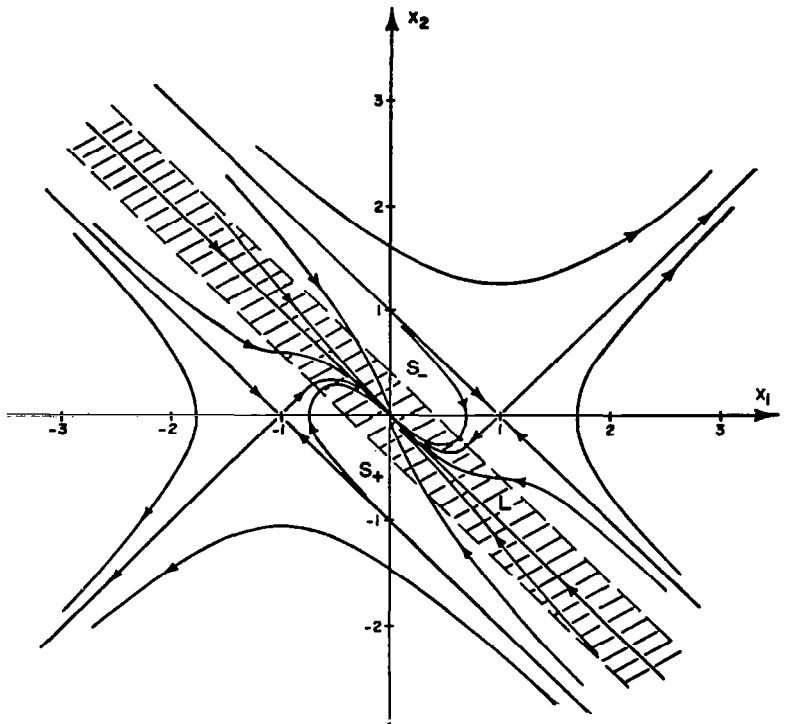

Fig. 7 L and S-regions for Example 3

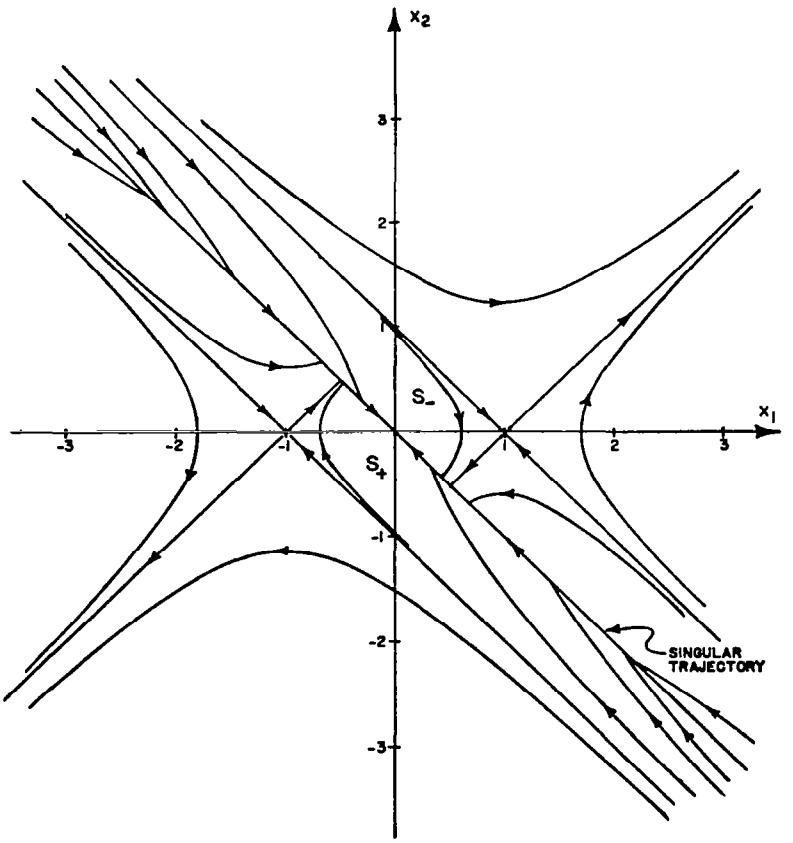

Fig. 8 Results for $c = 0$ in Example 3

7 G. F. D. Duff, *Partial Differential Equations*, University of Toronto Press, Toronto, Canada, 1956, chapter 3.
8 L. S. Pontryagin, V. G. Boltyanskii, R. V. Gamkrelidze, and E. F. Mishchenko, *The Mathematical Theory of Optimal Processes*, John Wiley & Sons, Inc., New York, N. Y., 1962.
9 R. Datko, "On Existence of Optimal Controls for a Performance Index With Positive Integrand," to be published.

APPENDIX

Proof of Theorem, Example 3

1 Let $\mathbf{x} = \mathbf{x}(t)$, $\mathbf{p} = \mathbf{p}(t) = -\nabla V_L(x(t))$ be a characteristic strip of V_L. We shall put $c = 1$. Then from (20)

$$\begin{aligned} \dot{\mathbf{x}} &= \mathbf{A}\mathbf{x} + \langle \mathbf{p}, \mathbf{f} \rangle \mathbf{f} \\ \dot{\mathbf{p}} &= \mathbf{Q}\mathbf{x} - \mathbf{A}'\mathbf{p} \end{aligned} \tag{42}$$

When $\mathbf{A}$ and $\mathbf{f}$ are given by (7) the characteristic polynomial of the system (42) is easily found to be

$$P(\lambda) = \sum_1^n \sum_1^n q_{jk}\lambda^{j-1}(-\lambda)^{k-1} + \left[\lambda^n - \sum_1^n a_k\lambda^{k-1}\right]\left[(-\lambda)^n - \sum_1^n a_k(-\lambda)^{k-1}\right] \tag{43}$$

Let the zeros of $P(\lambda)$ be $(\lambda_1, -\lambda_1), \ldots, (\lambda_n, -\lambda_n)$, where $\operatorname{Re}\lambda_m \leqq 0$, $m = 1, \ldots, n$. Actually $\operatorname{Re}\lambda_m < 0$ $(m = 1, \ldots, n)$, for if $\lambda = i\nu$ (ν real)

$$P(i\nu) = \left|\sum_1^n \sum_1^n q_{jk}(i\nu)^{j-1}(-i\nu)^{k-1}\right| + \left|(i\nu)^n - \sum_1^n a_k(i\nu)^{k-1}\right|^2 \geqq \begin{cases} \left|(i\nu)^n - \sum_1^n a_k(i\nu)^{k-1}\right|^2 > 0, & \nu \neq 0 \\ q_{11} > 0, & \nu = 0 \end{cases}$$

Hence the optimal linear control law $\phi_L(\mathbf{x}) = \langle \gamma, \mathbf{x} \rangle$ exists; and the optimal linear system is

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \langle \gamma, \mathbf{x} \rangle \mathbf{f} \tag{44}$$

with eigenvalues $\lambda_1, \ldots, \lambda_n$.

2 It will first be shown that the set L coincides with the strip $|\langle \gamma, \mathbf{x} \rangle| \leqq 1$. Equivalently,

$$\langle \gamma, \dot{\mathbf{x}} \rangle < 0 \quad (\text{or } > 0) \tag{45}$$

if

$$\langle \gamma, \mathbf{x} \rangle = +1 \quad (\text{or } -1) \tag{46}$$

where $\dot{\mathbf{x}}$ is given by (44).

From (34) and (43)

$$\begin{aligned} P(\lambda) &= q_{11}(\alpha_1 \ldots \alpha_{n-1})^{-2} \prod_1^{n-1} (\lambda - \alpha_m)(-\lambda - \alpha_m) \\ &\quad + \prod_1^n (\lambda - \alpha_m)(-\lambda - \alpha_m) \\ &= (\lambda - \lambda_n)(-\lambda - \lambda_n) \prod_1^{n-1} (\lambda - \alpha_m)(-\lambda - \alpha_m) \end{aligned} \tag{47}$$

where

$$\lambda_n = -|\alpha_n|a_1|^{-1}(a_1^2 + q_{11})^{1/2} \tag{48}$$

and we have used the fact that

$$|a_1| = |\alpha_1 \ldots \alpha_n| \tag{49}$$

Thus, from (47)

$$\lambda_m = \alpha_m, \qquad m = 1, \ldots, n-1 \tag{50}$$

Since $\lambda_1, \ldots, \lambda_n$ are the eigenvalues of the system (44), we have also

$$\lambda_m{}^n - \sum_{k=1}^n (a_k + \gamma_k)\lambda_m{}^{k-1} = 0, \qquad m = 1, \ldots, n \tag{51}$$

Hence by (50)

$$\sum_{k=1}^n \gamma_k \alpha_m{}^{k-1} = 0, \qquad m = 1, \ldots, n-1 \tag{52}$$

A simple calculation from (48), (50), and (51) shows that $q_{11} > 0$ implies $\gamma_n \neq 0$. From (52) we now have the identities

$$\sum_1^n \gamma_k \lambda^{k-1} = \gamma_n \prod_1^{n-1} (\lambda - \alpha_m) \tag{53}$$

and

$$\sum_1^n \gamma_{k-1}\lambda^{k-1} = \lambda\gamma_n \prod_1^{n-1} (\lambda - \alpha_m) - \gamma_n\lambda^n \tag{54}$$

where $\gamma_0 \equiv 0$. Also from (47) and (51)

$$\lambda^n - \sum_1^n (a_k + \gamma_k)\lambda^{k-1} = (\lambda - \lambda_n) \prod_1^{n-1} (\lambda - \alpha_m) \tag{55}$$

On combining (53)–(55) it now follows that

$$\sum_{k=1}^n [\gamma_{k-1} + \gamma_n(a_k + \gamma_k)]\lambda^{k-1} = \lambda_n \sum_1^n \gamma_k\lambda^{k-1}$$

and therefore

$$\gamma_{k-1} + \gamma_n(a_k + \gamma_k) = \lambda_n\gamma_k, \qquad k = 1, \ldots, n \tag{56}$$

From (56) we have

$$\langle \gamma, \dot{\mathbf{x}} \rangle = \lambda_n \langle \gamma, \mathbf{x} \rangle \tag{57}$$

for all $\mathbf{x}$ in L. Since $\lambda_n < 0$, (46) implies (45), as was to be shown.

3 It will now be shown that

$$\phi^0(\mathbf{x}) = \operatorname{sgn} \langle \gamma, \mathbf{x} \rangle \quad \text{if} \quad |\langle \gamma, \mathbf{x} \rangle| \geqq 1 \tag{58}$$

To this end the original optimization problem will be reduced to an equivalent problem for a system of first order. By (34), (52), and a slight extension of the results in Section 5 of [5] we can write

$$\sum_{j=1}^n \sum_{k=1}^n q_{jk}x_jx_k + u^2 \equiv q_{11}\gamma_1{}^{-2}\langle \gamma, \mathbf{x} \rangle^2 + u^2 - 2\frac{d}{dt} V_0(\mathbf{x}) \tag{59}$$

where V_0 is a homogeneous quadratic form in $x_1, \ldots, x_{n-1}$. Let

$$\xi = q_{11}{}^{1/2}\gamma_1{}^{-1}\langle \gamma, \mathbf{x} \rangle \tag{60}$$

We shall show that ξ satisfies a first-order differential equation. From (44) and (60)

$$q_{11}{}^{-1/2}\gamma_1\dot{\xi} = \langle \mathbf{A}'\gamma, \mathbf{x} \rangle + \gamma_n u \tag{61}$$

By (56)

$$\gamma_{k-1} + a_k\gamma_n = (\lambda_n - \gamma_n)\gamma_k, \qquad k = 1, \ldots, n;$$

hence

$$\mathbf{A}'\gamma = (\lambda_n - \gamma_n)\gamma$$

Thus (61) can be written

$$\dot{\xi} = (\lambda_n - \gamma_n)\xi + q_{11}{}^{1/2}\gamma_1{}^{-1}\gamma_n u \tag{62}$$

It is seen now that the original problem is equivalent to that of minimizing

$$J[u] = \frac{1}{2}\int_0^T (\xi^2 + u^2)dt \tag{63}$$

subject to (62) and the condition $\xi(T) = 0$. Since (62) is of first order this problem is easily solved. The result is

$$\phi^0(x) = \text{sat}\,[\theta\langle\gamma, x\rangle] \tag{64}$$

where

$$\theta = \frac{\gamma_n - \lambda_n - [(\gamma_n - \lambda_n)^2 + q_{11}\gamma_1^{-2}\gamma_n^2]^{1/2}}{\gamma_n} \tag{65}$$

It remains to check that $\theta = 1$. From (55)

$$a_n + \gamma_n = \lambda_n + \sum_1^{n-1} \alpha_m$$

or, since

$$a_n = \text{tr}\,\mathbf{A} = \sum_1^n \alpha_m,$$

$$\gamma_n = \lambda_n - \alpha_n \tag{66}$$

Also, from (53)

$$\gamma_1 = (-1)^{n-1}\gamma_n\alpha_1 \ldots \alpha_{n-1}$$

or

$$\gamma_1^2 = \gamma_n^2\alpha_n^{-2}a_1^2 \tag{67}$$

Substitution of (48), (66), and (67) into (65) gives the desired result.

The second statement of the theorem can be verified easily from (62) and (66). This completes the proof.