

REPORT NO. GDC DDF67-003
CONTROL NO. DCN 1-6-75-00116 SI (IF)

A study of

MINIMAX SOLUTIONS
FOR
SATURN CONTROL PROBLEMS

By C. D. Johnson

FINAL REPORT

May 1967

Prepared for
NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
GEORGE C. MARSHALL SPACE FLIGHT CENTER
HUNTSVILLE, ALABAMA
under
CONTRACT NAS8-18008

By
CONVAIR DIVISION OF GENERAL DYNAMICS
HUNTSVILLE OPERATIONS
Huntsville, Alabama

CONTENTS

List of Illustrations	iv
1. INTRODUCTION	1
2. A COMPUTATIONAL ALGORITHM FOR A CLASS OF CHEBYSHEV MINIMAX CONTROL PROBLEMS	3
2.1 Problems of Optimal Control	3
2.2 Chebyshev Minimax Optimal Control Problems	4
2.3 A Mathematical Theory for a Class of Chebyshev Minimax Optimal Control Problems	4
2.4 Generation of Chebyshev Minimax Optimal Trajectories by the Backward-Time Flooding Technique	8
2.5 An Algorithm for the Machine Solution of a Class of Chebyshev Minimax Optimal Control Problems	11
2.6 A Hybrid Analog Computer Realization of the Algorithm	15
2.7 Application of the C-Minimax Algorithm to a Load Minimizing Control Problem for a Fifth-Order Model of the Saturn Vehicle	25
2.8 Accommodation of Explicit Inequality Constraints on System State Variables	41
2.8.1 Exact Methods	42
2.8.2 Penalty Function Methods	44
2.8.3 Weakening Control Set Methods	45
References	47
3. OPTIMAL CONTROL OF THE DISTURBED LINEAR REGULATOR	49
3.1 The Optimal Regulator Problem	49
3.2 The Specific Problem	52
3.3 Form of the Solution	54
3.4 Discussion of Results	57
3.5 Extension of Results	58
3.6 Examples	59
References	65

CONTENTS

4.	A NUMERICAL ALGORITHM FOR COMPUTING THE LINEAR TRANSFORMATION $x = Ky$ WHICH TRANSFORMS AN ARBITRARY, COMPLETELY CONTROLLABLE, LINEAR DYNAMICAL SYSTEM $\dot{x} = Ax + u(t)f$ INTO THE CANONICAL (PHASE-VARIABLE) FORM $\dot{y} = A_0y + u(t)f_0$	68
4.1	Input Data	69
4.2	Algorithm for Generating K and K^{-1}	69
	References	73
5.	CONCLUSIONS	74

ILLUSTRATIONS

Chapter 2

1. General Block Diagram of the Proposed Algorithm	14
2. Essential Elements of an Analog "Track and Hold" Device	15
3. Typical Plot of Equation (33).	17
4. Analog Realization of a Free-Running Multi-Vibrator Circuit	19
5. Analog Realization of the E-Detector	21
6. Initial Condition $x_0 \in \mathcal{J}$ Generation Scheme	23
7. Hybrid Analog Circuit for the Complete Algorithm	24
8. Logic Scheme for Detecting the Inequality (99)	37
9. Analog Circuit for Generating Random Initial Conditions $x_0(\tau) \in \mathcal{J}$	39
10. Hybrid Analog Realization of the Algorithm for the Special Fifth-Order Saturn Load Minimizing Example	40

Chapter 3

1. Comparison of Responses for the System (62), (63), (71) and the System (74), (75), (76)	64
---	----

1

INTRODUCTION

This report is the final report from the Convair division of the General Dynamics Corporation on National Aeronautics and Space Administration Contract No. NAS8-18008 entitled "A Computational Algorithm for Obtaining Minimax Solutions to Saturn Control Problems".

During the study period, which began in June 1966 and continued through May 1967, the principal investigator conducted a detailed study of the problem of computing solutions for a certain class of Minimax type optimal control problems. Optimal control problems of the type considered, sometimes referred to as Chebyshev Minimax control problems, arise naturally in a variety of realistic optimization problems and have been a subject of increasing theoretical interest in recent years. In a previous study ["Study of Optimal and Adaptive Control Theory", NASA Contractor Report No. CR-715, University of Alabama Research Institute, April 1967] a mathematical theory for a certain class of Chebyshev Minimax control problems was developed and the possibility of using that theory as a basis for designing an automatic machine solution technique was suggested. In Chapter 2 of the present study, that suggestion is explored in more detail and a concrete algorithm is developed. In addition, a hybrid-analog computer realization of the algorithm is proposed and an application of the algorithm, to a certain load-minimizing control problem for the Saturn launch vehicle, is described. The problem of incorporating explicit state variable inequality constraints in the solution of Chebyshev Minimax control problems is also discussed and several alternative methods of solution are proposed.

The present study also included an investigation of a certain problem which arises in connection with practical applications of the so-called "Linear Optimal Regulator" control theory -- a theory which has been used to obtain approximate solutions to the Saturn "load-minimizing" control problem mentioned above. In particular, the problem of accommodating constant but unmeasurable (external) system disturbance inputs in the solution of the linear optimal regulator problem was considered. A mathematical theory for this class of problems was developed and the results, including several worked examples, are presented in Chapter 3.

In the study of optimal control problems associated with linear dynamical systems it is often found convenient (for both theoretical and practical reasons) to linearly transform the original state variable equations describing the physical system into a special canonical form known as the "phase-variable" form.

The algebraic theoretic properties of the transformation matrix K , required for this change of coordinates, were discussed previously in the NASA Contractor Report No. CR-715 mentioned above. In Chapter 4 of the present report an effective numerical algorithm for computing the phase-variable transformation matrix K , and its inverse K^{-1} , is presented.

This study was performed by the Convair division for, and under the direction of, the Aero-Astrodynamic Laboratory at the George C. Marshall Space Flight Center, Huntsville, Alabama. The principal investigator is especially grateful to Mr. Clyde Baker, Mr. Judson Lovingood, Mr. James Blair and Mr. Jerome Redus, of the Aero-Astrodynamic Laboratory, for their valuable suggestions and many stimulating and informative discussions during this study.

2

A COMPUTATIONAL ALGORITHM FOR A CLASS OF CHEBYSHEV MINIMAX CONTROL PROBLEMS

2.1 PROBLEMS OF OPTIMAL CONTROL

The branch of technology known as optimal control of dynamical systems is concerned with the general problem of synthesizing a forcing function $u(t)$ which will cause a given dynamical system S to respond "best" with respect to some, a priori specified, criterion of performance. The "goodness" or "quality" of the obtained response is generally measured mathematically by means of a performance index J - a functional computed along the response state trajectory of the dynamical system.

In many practical cases, the criterion of performance which is most meaningful from the physical point of view can be expressed mathematically as a time integral along the state trajectory of the system. For example, in the case of minimum-time problems the physically meaningful performance criterion is precisely expressed by the functional

$$J = \int_{t_0}^T 1 dt \quad (1)$$

where t denotes time measured along the state trajectory of S . Likewise, in the case of minimum energy, minimum fuel, and other "minimum resource consumption" type problems the physically meaningful performance criterion can be effectively expressed as a time integral of the rate of resource consumption. Optimal control problems in which the performance index is expressible as a time integral along the state trajectory have enjoyed notable popularity among both practical and theoretical control engineers. This is due, in part, to the fact that for such cases the problem can be studied within the framework of the powerful and relatively well-developed branch of mathematics known as the calculus of variations. The influence of the calculus of variations is also evident in the more recent theoretical contributions to optimal control by mathematicians such as Pontryagin, Bellman, and others. The "modern" mathematical techniques developed by these workers have generally been tailored for the same class of performance index functionals previously studied in the classical calculus of variations (eg. functionals which are, or can be reduced to, an integral along the state trajectory).

2.2 CHEBYSHEV MINIMAX OPTIMAL CONTROL PROBLEMS

Although the integral class of performance indexes has found numerous applications in practical optimization problems there are many realistic situations in which the physically meaningful performance criterion cannot be expressed mathematically as an integral along the state trajectory. For example, in certain high-speed atmospheric re-entry problems associated with space vehicles, the physical performance criterion might be stated as "keep the maximum vehicle surface temperature as small as possible" - an "optimum" condition which may be an essential factor in preventing the burning of the vehicle. This particular maximum temperature criterion cannot be expressed mathematically by an ordinary integral-type functional as used in the calculus of variations. Other practical examples can be found in the class of control problems associated with minimizing, say, maximum stress, maximum velocity, maximum force, and so on.

Optimal control problems in which the objective is to "minimize the maximum value" of a certain controlled variable are known as Chebyshev Minimax optimization problems [1] and have been the subject of increasing interest during recent years. In [1], a particular class of Chebyshev Minimax problems was studied and a relatively effective general method of solution was proposed. The results obtained in [1] included a brief mention of the possibility of devising a computational algorithm for numerically solving Chebyshev Minimax optimal control problems by completely automatic machine solution techniques. In the present study the possibility of such an automatic machine solution technique is explored in more detail and one concrete algorithm is proposed. In addition, a detailed hybrid analog computer program is developed for implementing the proposed algorithm. Before presenting the computational algorithm, the theory developed in [1] will be summarized.

2.3 A MATHEMATICAL THEORY FOR A CLASS OF CHEBYSHEV MINIMAX OPTIMAL CONTROL PROBLEMS

The particular Chebyshev Minimax optimal control problem studied in [1] can be stated as follows: In the class of piecewise continuous functions, find a control $u(t)$ which minimizes the functional

$$J[u; x_0] = \max_{t_0 \leq t \leq T} C(x(t)) \quad (2)$$

subject to the restrictions

$$\dot{x} = F(x, u(t)) \quad (3)$$

$$x(t_0) = x_0 \quad (4)$$

$$\mathcal{J}(x(T)) = 0 \quad (5)$$

$$u(t) \in U, \quad t_0 \leq t \leq T \quad (6)$$

In (2), $x = (x_1, \dots, x_n)$ is an n -vector: the system state vector, and $C(x)$ is the performance index: a real, single valued, scalar function of x defined throughout a set D of the n -dimensional euclidean state space E^n . In (3), F is a vector function continuous in u and continuously differentiable with respect to $x \in D$. Equation (5) defines the terminal manifold, $\mathcal{J} \subset D$, an m -dimensional ($m < n$) hypersurface of admissible terminal states $x(T)$. The terminal time T is specified implicitly, by (5) as the first time $t \geq t_0$ which satisfies $\mathcal{J}(x(t)) = 0$.

A piecewise continuous real valued function $u(t)$ with values belonging to the compact, convex, set U is called an admissible control. An admissible control $u = u^o(t)$ which yields an absolute minimum of the functional (2), subject to the restrictions (3)-(6), is called optimal. An optimal control of the form $u^o(t) = u^o(x(t))$ is an optimal control law. An integral curve of (3) corresponding to an optimal control, is an optimal trajectory. The set $D \subset E^n$ is taken as the set of all states x which are controllable to T . That is, for each initial state $x_0 \in D$ there exists at least one admissible control $u(t)$ such that the corresponding solution of (3) satisfies (4) and (5). It is assumed that D is non-void and $u^o(x)$ exists for all $x \in D$. It is further assumed that $C(x)$ and $\mathcal{J}(x)$ are once continuously differentiable.

The solution procedure proposed in [1] is based on the following fundamental fact. A Chebyshev Minimax optimal trajectory which starts at an arbitrary initial state $x_0 \in D$ has one or the other of the following properties: (i) the corresponding maximum value of $C(x(t))$, $t_0 \leq t \leq T$, is greater than $C(x_0)$ or (ii) the corresponding maximum value of $C(x(t))$, $t_0 \leq t \leq T$, exactly equals the value of $C(x_0)$. Thus, if $V(x)$, $x = x_0$, denotes the ordinary Caratheodory value function

$$V(x) = J[u^o; x], \quad x = x_0 \quad (7)$$

then the above mentioned fact can be expressed as the weak inequality

$$V(x) \geq C(x), \quad x = x_0 \quad (8)$$

It is assumed hereafter that $V(x)$ is continuous at each state x in the interior of D .

The result (8) suggests that the set $D \subset E^n$ can be partitioned into two subsets $\{R_0\}$ $\{R_n\}$ defined as follows

$$\{R_0\} = \{x \mid V(x) = C(x)\} \quad (9)$$

$$\{R_n\} = \{x \mid V(x) > C(x)\} \quad (10)$$

In [1] it was shown that the subsets $\{R_0\}$, $\{R_n\}$ can be identified¹ by the following procedure. Let $R_0^1 \subset D$ ($R_0^1 \supset \mathcal{J}$) be the largest set of states $x \in D$ with the property that: For each $x_0 \in D$ there exists an admissible control $u = \phi(t; x_0)$ $t_0 \leq t \leq T$, such that

(i) $dC(x(t), \phi(t; x_0))/dt \leq 0$, $t_0 \leq t \leq T$, and

(ii) $\mathcal{J}(x(T)) = 0$ for some $T \geq t_0$ along the corresponding solution of (3).

Evidently, the set R_0^1 is connected. Moreover, under the previously stated assumptions concerning existence of an optimal control law and continuity of $V(x)$, the set R_0^1 is closed, relative to D .

From (8), it is clear that when $u = \phi(t; x_0)$, $J[u]$ realizes its greatest lower bound at each $x_0 \in R_0^1$. It follows that

$$V(x) = C(x) \quad \forall x = x_0 \in R_0^1 \quad (11)$$

and $u = \phi(t; x_0)$ is an optimal control for the set R_0^1 . It is remarked that the control $\phi(t; x_0)$ is not unique, in general. Moreover the set $R_0^1 - \mathcal{J}$ might be empty.

Consider next an arbitrary initial state $x_0 \in (D - R_0^1)$ and let ∂R_0^1 denote the boundary of the set R_0^1 . In accordance with the procedure described in [1], an auxiliary, Mayer-type, variational problem (hereafter called Problem M) is defined as follows:

Problem M

Find an admissible control $u(t)$ which minimizes

¹In the identification procedure described here, and in [1], it is convenient to classify certain states x , where $V(x) = C(x)$, as belonging to the set $\{R_n\}$. Thus the partitioning described in the sequel does allow an equality sign on the right side of (10), in certain special cases.

$$\tilde{J} [u; x_0] = C(x(t_1)) \quad (12)$$

subject to the restrictions

$$\dot{x} = F(x, u(t)) \quad (13)$$

$$x(t_0) = x_0 \in (D - R_0^1) \quad (14)$$

$$x(t_1) \in \partial R_0^1, \quad t_1 - \text{unrestricted} \quad (15)$$

$$u(t) \in U \quad t_0 \leq t \leq t_1 \quad (16)$$

Let $u(t) = \gamma(t; x_0)$ be an optimal control for Problem M and let Ω denote the family of all trajectories $x(t)$, $x \in (D - R_0^1)$, generated by the solutions of $\dot{x} = F(x, \gamma(t; x_0))$, $x_0 \in (D - R_0^1)$. Further, let $R_m^1 \subset (D - R_0^1)$ be the set of states x with the (natural) property² that along each trajectory $x(t) \in \Omega$ (with $x \in R_m^1$) the value of $C(x(t))$ does not exceed the terminal value $C(x(t_1))$. It is readily verified (see [1]) that the sub-arcs of the trajectories $x(t) \in \Omega$ which belong to R_m^1 are optimal for the original C-minimax problem. In particular, if $x_0 \in R_m^1$, the C-minimax optimal control $u^0(t)$ (which takes x_0 to \mathcal{J}) can be chosen as

$$u^0(t) = \begin{cases} \gamma(t; x_0) & t_0 \leq t < t_1 & x_0 \in R_m^1 \\ \phi(t; x^*(t_1)) & t_1 \leq t \leq T & x^*(t_1) \in \partial R_0^1 \end{cases} \quad (17)$$

where $x^*(t_1)$, $t_1 = t_1(x_0)$, is the state $x \in \partial R_0^1$ at which the solution of Problem M terminates. At each state $x_0 \in R_m^1$ the functional (2) is given by

$$J[u^0; x_0] = C(x^*(t_1)) \geq C(x_0) \quad (18)$$

and therefore the value function for that set can be expressed as

$$V(x) = C(x^*(t_1(x))) \quad \forall x = x_0 \in R_m^1 \quad x^* \in \partial R_0^1 \quad (19)$$

By this means, the value function $V(x)$ and a C-minimax optimal control can be determined for an arbitrary initial state $x_0 \in R_0^1 \cup R_m^1$. Moreover, since the value function is known on the boundary of the set $R_0^1 \cup R_m^1$ that boundary can be viewed as a "new" terminal manifold, say \mathcal{J}^2 , and the identification process outlined

²It should be stressed that this property is one which occurs naturally along the family Ω --it must not be interpreted as a state variable constraint imposed upon the solutions of the original problem M.

above can be repeated for the states $x \in (D - R_0^1 \cup R_m^1)$. In this way, sets R_0^2 and R_m^2 , analogous to R_0^1 and R_m^1 , can be constructed using \mathcal{J}^2 in place of the original terminal manifold \mathcal{J} . Continuing in this way, the set D can be completely partitioned into the two families of sets $\{R_0^k\} = \{R_0^1, R_0^2, \dots\}$ and $\{R_m^k\} = \{R_m^1, R_m^2, \dots\}$. When the partitioning of D into the sets $\{R_0^k\}$, $\{R_m^k\}$ is completed, the optimal control for the original problem (1) - (5) is known. Suppose, for example, that the initial state x_0 belongs to a set $R_0^k \subset \{R_0^k\}$, $k \geq 2$. The optimal control, during the time interval, $t_0 \leq t \leq t_k$, when $x(t) \in R_0^k$, can be chosen as any admissible control for which $C(x(t)) \leq C(x_0)$ and $x(t_k) \in \partial R_m^1$ for some $R_m^1 \subset \{R_m^1\}$. The existence of at least one such control follows from the definition of the R_0 type sets. Upon entering the neighboring set R_m^1 , the continuation of the optimal control is determined by solving the appropriate, Mayer type, variational problem (12) - (16) where the "terminal manifold" is taken as the boundaries of the immediately adjoining sets of the R_0 type. In this way, the state $x(t)$ progresses alternately and optimally through the sets of the R_0 and R_m type and eventually reaches the original terminal manifold \mathcal{J} .

2.4 GENERATION OF CHEBYSHEV MINIMAX OPTIMAL TRAJECTORIES BY THE BACKWARD-TIME FLOODING TECHNIQUE

The set R_0^1 is characterized by the fact that, at each state $x \in R_0^1$, there exists at least one admissible control $\phi \in U$ such that

$$\langle \nabla C(x), F(x, \phi) \rangle \leq 0 \quad x \in R_0^1, \quad \phi \in U \quad (20)$$

Thus, in backward-time" $\tau = T - t$, ($\tau \geq 0$), one can always find, at each $x \in R_0^1$, at least one control such that

$$\frac{dC(x(\tau))}{d\tau} \geq 0 \quad (21)$$

Evidently, any state $x \in D$ which can be reached from $x(\tau = 0) \in \mathcal{J}$ by an admissible backward-time trajectory $x(\tau)$ satisfying (21) identically must belong to the set R_0^1 . Moreover, from the definition of R_0^1 it is concluded that each state $x \in R_0^1$ must be "reachable" along at least one such backward-time trajectory $x(\tau)$.

Thus, the set R_0^1 is the largest set of states x which can be reached, from $x(\tau = 0) \in \mathcal{J}$, along solution trajectories $x(\tau)$ of the system

$$\dot{x}(\tau) = -F(x, \tilde{u}(\tau)) \quad x(0) \in \mathcal{J}, \quad (\dot{} = d/d\tau) \quad (22)$$

where the control $\tilde{u}(\tau)$ is subject to the special constraint

$$\tilde{u}(\tau) \in \tilde{U}(x) \quad \forall \tau \geq 0, \quad x \in R_0^1 \quad (23)$$

and the state dependent set $\tilde{U} \subset U$ is defined by

$$\tilde{U}(x) = \{u \in U \mid \langle \nabla C(x), F(x, u) \rangle \leq 0\} \quad (24)$$

From this reachable set point of view, the total boundary ∂R_0^1 of R_0^1 is composed in general, of subsets characterized by one or the other of the following three conditions:

- (i) Some subsets of ∂R_0^1 may be built up from manifolds of trajectories of (22) which are generated by boundary controls³ $\tilde{u}(\tau) \in \partial \tilde{U}(x)$.
- (ii) Some subsets of ∂R_0^1 may coincide with the boundary of the controllable set D in which case trajectories $x(\tau)$ of (22) - (24) approach ∂R_0^1 only as $\tau \rightarrow \infty$.
- (iii) Some subsets of ∂R_0^1 may consist of states x , reached along trajectories of (22), from which the set $\tilde{U}(x)$ first becomes empty.

It is clear that ∂R_0^1 is traversed by C-minimax optimal trajectories only in the case of condition (iii).

On the boundary of R_0^1 , the value function is given by

$$V(x) = C(x) \quad x \in \partial R_0^1 \quad (25)$$

Therefore, in the course of solving the conventional Mayer-type Problem M in the set R_n^1 , it can be shown that the values of the corresponding Lagrange-Pontryagin multipliers $(p_1(t), \dots, p_n(t)) = p(t)$, at $t = t_1$, are given by (see [1])

$$p(t_1) + \nabla C(x(t_1)) = \begin{cases} 0, & \text{if } x(t_1) \in (\partial R_0^1 - \mathcal{J}) \\ \text{normal to } \mathcal{J}, & \text{if } x(t_1) \in \mathcal{J} \end{cases} \quad (26)$$

³It should be noted that $\tilde{U}(x)$ is a closed, but not necessarily convex, set.

Suppose that $x^* \in \partial R_0^1$ is a state, reached along a C-minimax optimal trajectory $x(\tau)$ of (22) - (24), from which the set $\tilde{U}(x)$ first becomes empty [eg. condition (iii) above]. The continuation of the C-minimax optimal trajectory $x(\tau)$ across ∂R_0^1 into the region R_n^1 is determined by solving Problem M, in backward-time, starting at the state $x^* \in \partial R_0^1$. For this purpose it is necessary to solve the set of $2n$ Euler-Pntryagin canonical equations

$$\begin{aligned} \dot{x}_i(\tau) &= -F_i(x, \gamma(x, p)) \\ \dot{p}_i(\tau) &= + \sum_{j=1}^n p_j \frac{\partial F_j(x, \gamma(x, p))}{\partial x_i} \quad i = 1, \dots, n, \quad \cdot = d/d\tau \end{aligned} \quad (27)$$

where $\gamma(x, p)$ is determined by the Maximum Principle

$$\gamma(x, p) = \arg \max_{u \in U} \langle p, F(x, u) \rangle \quad (28)$$

and where the first integral

$$\langle p(\tau), F(x(\tau), \gamma(x(\tau), p(\tau))) \rangle = 0 \quad (29)$$

is satisfied identically in τ . Equations (26) and the known value of x^* effectively determine the backward-time initial conditions for (27) so that, in principle, the continuation of the C-minimax optimal trajectory $x(\tau)$, through the set R_n^1 , can be affected.⁴ According to the definition of the set R_n^1 , the continuation of a C-minimax optimal trajectory $x(\tau)$ through R_n^1 must be stopped the first time a state x is reached where any further continuation will result in the value of $C(x(\tau))$ exceeding its initial value $C(x^*) = C(x(t_1))$. Each state x determined in this manner is a boundary point for the set R_n^1 . The further backward-time continuation of the C-minimax optimal trajectory $x(\tau)$ from $x \in R_n^1$ into the set R_0^2 is carried out by the same technique used above for the set R_0^1 . In this way, the trajectory $x(\tau)$ can be continued through the sets R_0^2 , R_n^2 , R_0^3 , R_n^3 , ... and so on.

This procedure for generating C-minimax optimal trajectories has been

⁴The presence of singular solutions (to Problem M) can lead to certain technical difficulties in integrating (27) through the set R_n^1 . This subject is discussed in more detail in [2].

successfully used in a variety of more conventional optimization problems where it is commonly known as "backward-time flooding from the terminal manifold". The method is numerical in nature and is therefore primarily useful in obtaining specific "open-loop" optimal controls $u^0 = u^0(t; x_0)$. However, in some cases, for example where the optimal control is of the bang-bang type, the method can be used to monitor the switching function and thereby numerically identify points on the optimal switching surfaces in the state space. In this way, important information about the optimal control law $u^0 = u^0(x)$ can be obtained.

This procedure is particularly attractive for C-minimax optimal control problems because, as shown above, it provides a systematic method⁵ for numerically identifying points on the boundaries of the sets R_0^1, R_0^2, \dots and R_m^1, R_m^2, \dots . Moreover, the C-minimax optimal control in the sets R_m^1, R_m^2, \dots is quite often of the bang-bang type and therefore points on the optimal switching surfaces in $\{R_m^1\}$ can be identified.

A mechanization of the C-minimax backward-time flooding procedure described above requires four essential elements

- (i) A device for setting initial conditions and integrating the system differential equations and canonical equations in backward-time.
- (ii) A device for generating allowable control functions.
- (iii) A device for systematically selecting initial conditions $x_0 \in \mathcal{J}$ so that \mathcal{J} (and E^n) is flooded with a sufficiently dense covering of optimal trajectories.
- (iv) A means for monitoring and recording certain properties of the solutions $x(\tau), p(\tau)$, and re-setting the integration device.

A general block diagram for the mechanization of such a backward-time flooding procedure is described in the next section.

2.5 AN ALGORITHM FOR THE MACHINE SOLUTION OF A CLASS OF CHEBYSHEV MINIMAX OPTIMAL CONTROL PROBLEMS

In the previous section, a procedure was described for generating points on the boundaries of the sets $\{R_0^1\}$ and $\{R_m^1\}$ by systematic backward-time integration of the system and canonical equations. In the present section, a general block diagram for mechanizing this algorithm is proposed. This block diagram can be physically realized by either analog, digital, or hybrid (analog-digital) computing equipment. One possible hybrid realization will be described in the next section.

⁵ Cases in which $T = \infty$ may require special treatment (see [1]).

The general sequence of operations required for the backward-time generation of a family of C-minimax optimal trajectories from \mathcal{J} can be summarized as follows:

Algorithm

1. Select an arbitrary initial state $x_0 \in \mathcal{J}$ and set x_0 as initial condition on the system (22). The value of $x_0 \in \mathcal{J}$ might be chosen, for example, by some deterministic or random selection scheme.
2. Generate the set $\tilde{U}(x)$, defined by (24), and continuously select values \tilde{u} from $\tilde{U}(x)$ [eg. $\tilde{u} \in \tilde{U}(x)$] in some random-like manner.
3. Begin integration of the system equations (22) (starting at $x_0 \in \mathcal{J}$) using the values of $\tilde{u}(\tau) \in \tilde{U}(x(\tau))$ obtained in step (2).
4. Continue the integration in step (3) until a state $x^*(\tau)$ is reached where set $\tilde{U}(x(\tau))$ first becomes empty. When $x^*(\tau)$ is reached, the boundary of R_0^1 has been penetrated by an infinitesimal amount. At that event, place the integrating system for (22) in the hold mode and record and/or store the following data: value of x^* , value of $C(x^*)$.
5. Compute initial conditions $(x(t_1), p(t_1))$ for the canonical equations (27) by using (26) with $x(t_1) = x^*$. Note that $\nabla C(x)$ is a known function of x .
6. Generate the function $\gamma = \gamma(x, p)$ according to the rule (28).
7. Start integration of the canonical equations (27) using initial conditions from step (5) and the function $\gamma(x(\tau), p(\tau))$ from step (6). Record $x(\tau)$ and $\gamma(x(\tau), p(\tau))$ as desired.
8. Monitor the value of $C(x(\tau))$ along solutions of the canonical equations (27) and continuously compare $C(x(\tau))$ with the stored value of $C(x^*)$. Note that the sign of $dC(x(\tau))/d\tau$ should become negative when integration of the canonical equations, into the set R_n^1 , first begins. The sign may or may not change thereafter.
9. Continue integration of the canonical equations until a state $x(\tau) = \bar{x}$ is reached where $C(x(\tau))$ first exceeds $C(x^*)$. When \bar{x} is reached, the boundary of R_n^1 has been penetrated by an infinitesimal amount. At that event, place the integrating system for (27) in the hold mode and record and/or store the value of \bar{x} .
10. Set \bar{x} as initial condition on the system equations (22) and begin integration of (22) using $\tilde{u}(\tau)$ values generated by step (2). Note that the set $\tilde{U}(x)$ will not be empty at $x = \bar{x}$.

11. Repeat steps (4) - (10) of the algorithm to determine points $x^{*2}, x^{*3} \dots$ and $\bar{x}^2, \bar{x}^3, \dots$, on the boundaries of $R_0^2, R_0^3 \dots$ and $R_m^2, R_m^3 \dots$ respectively.
12. When the one continuous trajectory $x(\tau)$, generated by this procedure, has been continued sufficiently far away from \mathcal{J} , interrupt the integration in progress and return to step (1) to begin generation of a second trajectory $x(\tau)$. Repeat this procedure until a sufficiently dense set of trajectories $x(\tau)$ has been obtained.

The information generated by this algorithm consists of: (i) a finite collection of points x on the boundaries of the sets R_0^1, R_0^2, \dots and R_m^1, R_m^2, \dots , (ii) a family of C-minimax optimal trajectories, and the corresponding open loop control $\gamma(x(\tau), p(\tau))$, $\tau \geq 0$, for the sets R_m^1, R_m^2, \dots .

A general block diagram for the mechanization of this algorithm is shown in Figure 1. It should be noted that with the aid of a control gate G, the integration device F performs the integration for both equation (22) and the first of equation (27). The integrating devices F and P shown in Figure 1 can be realized by electronic analog computer integrating elements or, alternatively, by a digital computer numerical integration program. The device labeled " $\tilde{U}(x)$ computer" accepts the vector function $x(\tau)$ as an input and continuously generates a time varying "set" output consisting of all values of the real variable u which satisfy (24).

The generated set $\tilde{U}(x(\tau))$ is continuously monitored by the E-detector which places the F integrator in a momentary hold mode, operates the control gate G, and subsequently starts both F and P integrators, when $\tilde{U}(x(\tau))$ first becomes empty. The " $\tilde{u}(\tau)$ random selector" continuously selects, in some random-like manner, a sequence of values from the generated set $\tilde{U}(x(\tau))$.

The control gate G selects the input u to the integrator F to be either $u(\tau) = \tilde{u}(\tau)$ or $u(\tau) = \gamma(x(\tau), p(\tau))$ according to the gate command signals received from the E-detector and comparator K. The scalar function $\gamma(x(\tau), p(\tau))$ is generated by performing purely algebraic operations on the $2n$ scalar quantities $(x_1(\tau), \dots, x_n(\tau); p_1(\tau), \dots, p_n(\tau))$ as prescribed by equation (28).

At each successive point $x^* \in \partial R_0^i$, $i = 1, 2, \dots$, where the set $\tilde{U}(x(\tau))$ first becomes empty, the corresponding value of $C(x^*)$ is computed and stored by the sample-hold computer S. This stored (constant) value is then continuously compared (in comparator K) with the time-varying value of $C(x(\tau))$, $x(\tau) \in R_m^i$, as (27) is integrated through the adjoining set R_m^i .

The maximum interval of time allowed for the prolongation of a backward-time trajectory $x(\tau)$ can be controlled either directly (by clocks etc.) or indirectly (by limiting, say, the maximum value of the norm $\|x(\tau)\|$). In either case, when

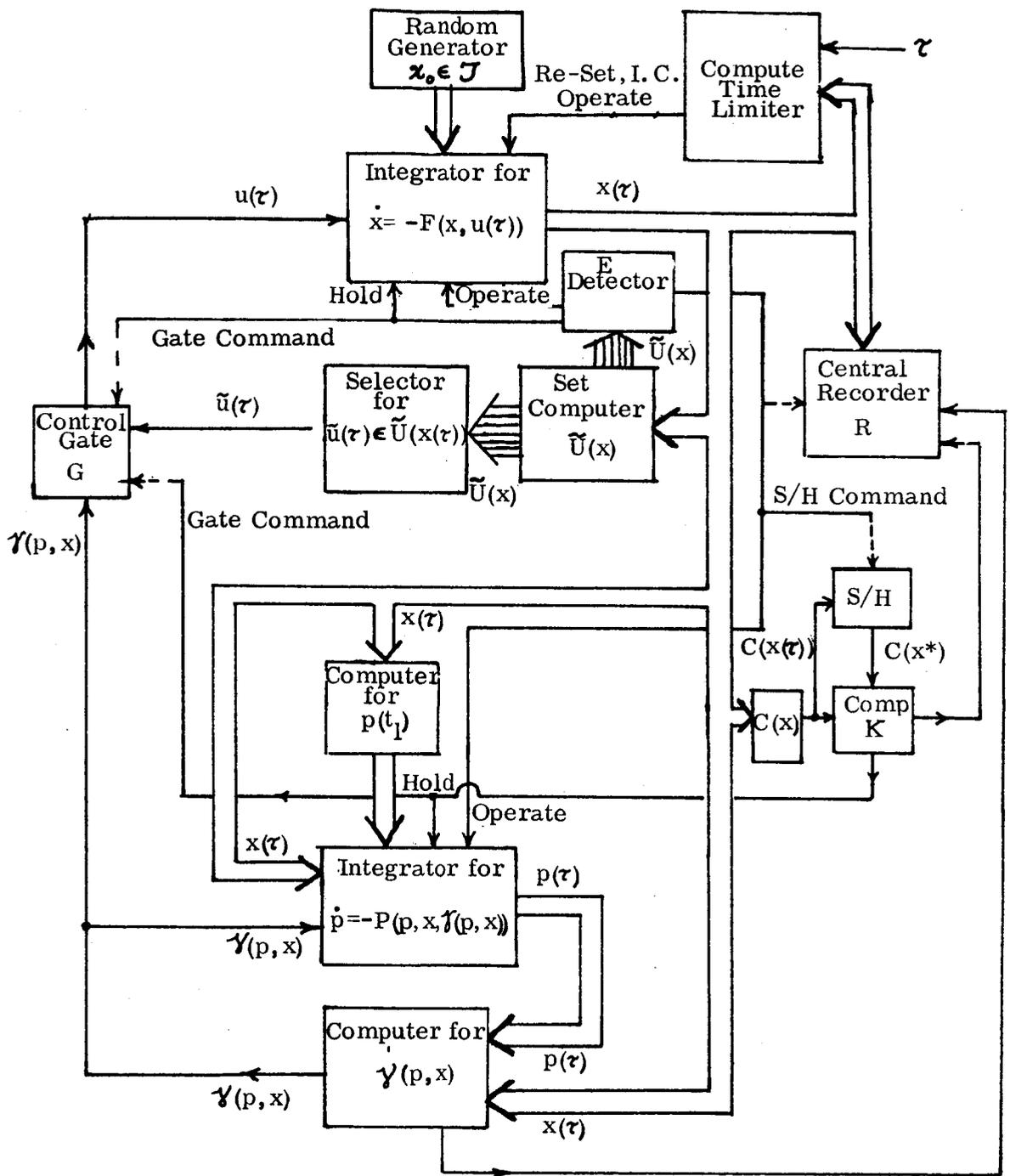


Figure 1. General Block Diagram of the Proposed Algorithm.

the limit has been reached, the integration process is interrupted and the computation procedure is returned to step #1 of the algorithm.

A hybrid analog computer realization of this algorithm is described and illustrated in the next section.

2.6 A HYBRID ANALOG COMPUTER REALIZATION OF THE ALGORITHM

The general algorithm described in Section 5 consists essentially of alternating integration between two systems of ordinary differential equations with the output of one system providing the initial condition data for the other system. Computations of this type are particularly suited for mechanization on a hybrid electronic analog computer. In this section a detailed circuit diagram is given for one such hybrid analog computer mechanization.

A hybrid electronic analog computer is characterized by the ability to logically control the operation mode of its electronic integrators and other standard analog components. By this means, a hybrid computer circuit can be so designed to automatically start and stop the analog computation, up-date and re-set integrator initial conditions, place the analog computation in "hold" mode, perform circuit switching operations, sample and store selected signal values, etc. according to logic signals generated within the computer circuit itself. [eg. in closed-loop (bootstrap) fashion].

Thus the automatically controlled integrating devices F and P shown in Figure 1 can be physically realized by standard programming of integrating and summing amplifiers on a hybrid analog computer. The functions $C(x)$, $\gamma(x, p)$ and $p(t_1)$ can likewise be realized by standard analog "algebraic function generation" schemes using ordinary diode function generators, comparitors, resolvers, etc. The sampling and storage of successive values of $C(x^*)$ is accomplished by means of a logically activated analog "track and hold" (T/H) device--a packaged component which is standard equipment on most hybrid computer systems. This component consists essentially of a single high gain integrator with a logically controlled "hold relay" as shown in Figure 2. In operation, the output of the T/H unit

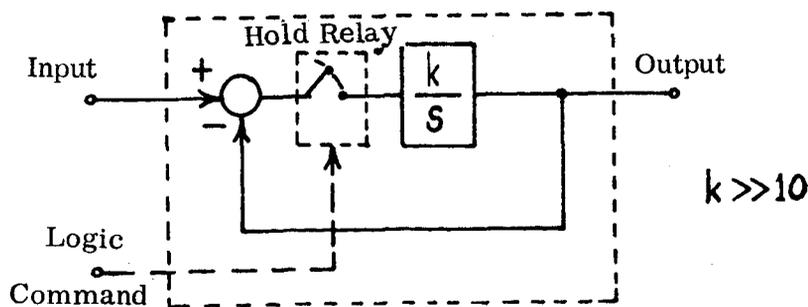


Figure 2. Essential Elements of an Analog "Track and Hold" Device

essentially equals the input as long as the logic command signal corresponds to "track" (hold relay closed).

When the logic command signal changes to "hold", the hold relay of the high gain integrator opens and the instantaneous value of the output, at that particular moment, is "frozen" (held constant)--even though the input continues to vary with time.

The continuous comparison of $C(x(\tau))$ and $C(x^*)$ is accomplished by a standard hybrid analog "logic comparator" device which accepts the signals $C(x(\tau))$ and $C(x^*)$ as inputs and generates a binary logic output signal according as $C(x(\tau)) \leq C(x^*)$ or $C(x(\tau)) > C(x^*)$. This logic output signal, in turn, serves as a command signal to logically control the P integrator circuit, the control gate G, and the central recorder R.

The control gate G can be physically realized by means of a standard hybrid analog component known as a logically controlled "latching relay". This device is essentially a switch (either mechanical or solid state) which transmits one or the other of the two input signals $\tilde{u}(t)$, or $\gamma(x(\tau), p(\tau))$ according to the two logic commands received from the E-detector and the comparator K. The position of the switch is "latched" after each transition so that the logic command which initiated the transition, say from the E-detector, cannot re-activate the switch until the logic command from the comparator K has effected a transition--and vice versa.

The non-standard computing operations in the proposed hybrid analog mechanization consist of (i) the $\tilde{U}(x)$ set computer, (ii) the E-detector, (iii) the $\tilde{u}(\tau) \in \tilde{U}(x(\tau))$ control selector and (iv) the initial condition $x(t_0) \in \mathcal{J}$ generator for the F-integrator. Components which realize these operations require special design considerations. One concrete method for constructing these components is described, in detail, below.

The device which computes the set $\tilde{U}(x)$ must accept the continuously varying n-vector x as an input and continuously generate, as an output, the instantaneous set of all values of the real variable u which simultaneously satisfy the two restrictions

$$\langle \nabla C(x), F(x, u) \rangle \leq 0 \quad (30)$$

$$u \in U \quad (31)$$

where $\nabla C(\cdot)$, and $F(\cdot, \cdot)$ are known vector functions and U is a known, compact, convex subset of the real line. The device which computes $\tilde{u}(\tau) \in \tilde{U}(x(\tau))$ must, in turn, continuously select values from the continuously varying set $\tilde{U}(x(\tau))$ in some

random-like manner. In order to realize these two operations, the following scheme was developed.

Let $\sigma(\tau)$ be an externally generated, continuous scanning signal which ranges over all values in the set $\sigma(\tau) \in U$ in some random-like (and semi-periodic) manner. For example, suppose

$$U = \{u \mid u^2 \leq 1\} \quad (32)$$

In this case $\sigma(\tau)$ might be chosen⁶ as, say,

$$\sigma(\tau) = \text{sat} [\eta(\tau) + A \sin \omega \tau] \quad (33)$$

where $\eta(\tau)$ is white noise and A is an appropriately chosen constant.⁷ A typical plot of (33) is shown in Figure 3.

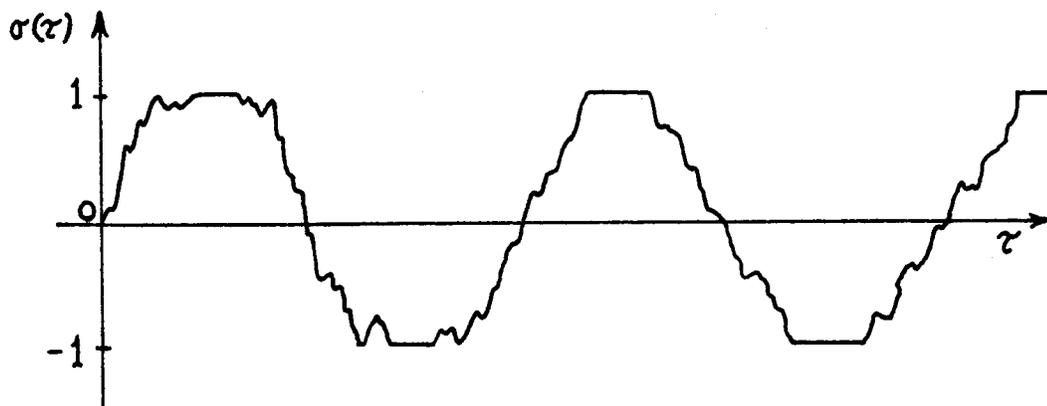


Figure 3. Typical Plot of Equation (33).

⁶The function $\text{sat}(y)$ is defined as:

$$\text{sat}(y) = \begin{cases} y: & \text{if } |y| \leq 1 \\ \text{sgn } y: & \text{if } |y| > 1 \end{cases}$$

⁷The function $A \sin \omega t$ in (33) can be replaced by any other appropriate periodic function such as a triangular wave, a saw tooth wave, etc.

Now, let the scanning signal $\sigma(\tau)$ and the n-vector $x(\tau)$ be the inputs to an ordinary algebraic function generator which computes

$$\text{sgn}[\langle \nabla C(x(\tau)), F(x(\tau), \sigma(\tau)) \rangle] \quad (34)$$

It is clear that each value $\sigma = \sigma(\tau; x(\tau))$ which makes (34) non-positive is an admissible value for $u(\tau)$. On the other hand, any value of σ which makes (34) positive is not an admissible value of $u(\tau)$. Thus, as long as the continuously varying input pair $(\sigma(\tau), x(\tau))$ generate a non-positive value for the function (34) the semi-random scanning signal $\sigma(\tau)$ can be used as the desired control $\sigma(\tau) = u(\tau)$. Whenever (34) becomes positive, at some state $x(\tau)$, the process which generates the continuous signal $\sigma(\tau)$ must somehow "skip over" the set of values of σ which makes (34) positive and "jump" to some other set of σ which does make (34) non-positive. In practice, this process can be closely approximated by massive augmentation of the (nominal) $\sigma(\tau)$ scanning frequency whenever (34) becomes positive. For example, if $\sigma(\tau)$ is generated by (33), then one could effectively "skip over" the values of σ which makes (34) positive by simply increasing the value of ω by a large factor (say, 1000) whenever (34) becomes positive. In this way, the scanning signal $\sigma(\tau)$ effectively jumps (actually experiences a very large, but finite, derivative) whenever (34) becomes positive. If the relative magnitudes of the constant A and the nominal "amplitude" N of $\eta(\tau)$ are chosen appropriately⁸ this process will always result in $\sigma(\tau)$ "jumping"⁹ to some new value which does make (34) non-positive--provided that the set $\tilde{U}(x(\tau))$ has not become exhausted (empty). If $\tilde{U}(x(\tau))$ has become empty, the scanning signal $\sigma(\tau)$ will experience a continuous, very high frequency, oscillation across the set U. The continued presence of this high frequency oscillation can be effectively used to detect the condition that $\tilde{U}(x)$ is empty--the E-detection operation. For instance, if an elapsed-time counter is activated¹⁰ each time the function (34) becomes positive, the condition that: "elapsed-time" > "one period of the high frequency scan rate" will occur if and only if $\tilde{U}(x(\tau))$ is empty. By this means, E-detector logic signals can be generated which will activate the control gate G and the F-integrator mode control whenever the set $\tilde{U}(x(\tau))$ becomes empty.

⁸The selection of appropriate values for A and N is discussed in the sequel.

⁹If the set $\tilde{U}(x(\tau))$ is not empty, (and for appropriate choices of A and N) this jump will require not more than one period of the high frequency scan rate, in general.

¹⁰It is assumed that the counter is automatically reset to zero whenever (34) becomes non-positive.

The scheme described above generates a random-like signal $\sigma(\tau)$ which can be used as the desired control $\tilde{u}(\tau)$ as long as (34) is non-positive. Moreover, whenever $\sigma(\tau^*)$ causes (34) to become positive, the scheme rapidly changes $\sigma(\tau)$ to some other admissible value $\sigma(\tau^* + \delta\tau)$ which makes (34) again non-positive. During the short transition (jump) time $\delta\tau$ the value of $u(\tau')$, $\tau^* \leq \tau' \leq \tau^* + \delta\tau$, should be chosen to satisfy (30). To accomplish this end, the integrator F can be put in momentary "hold" mode during the short interval $\delta\tau$ or, alternatively, the value of $\tilde{u}(\tau')$ can simply be held constant at the value $\tilde{u}(\tau^*)$ during the interval $\delta\tau$. This latter alternative, which will be used in the sequel, satisfies (30) only in an approximate sense, the goodness of the approximation depending on the rate of change of the left side of (30) at the time $\tau = \tau^*$. This rate can be effectively controlled by the overall time-scaling of the analog computation.

The procedure outlined above forms a practical basis for the efficient realization of the $\tilde{U}(x)$ computer, the E-detector, and the $u(\tau) \in \tilde{U}(x(\tau))$ control selector. One method for the physical implementation of these devices is described below.

The generation of the scanning signal $\sigma(\tau)$, on a hybrid analog computer, can be conveniently accomplished by implementing an analog, frequency modulated, triangular wave free-running multivibrator circuit with white noise added to the input. The multi-vibrator output signal, in this case, has the form of a pure triangular (periodic) function superposed with an integrated white noise signal. A simple analog arrangement which realizes this multi-vibrator is shown in Figure 4.

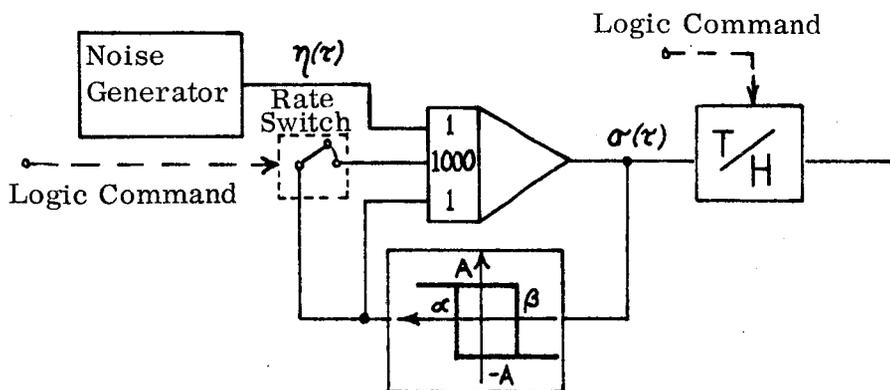


Figure 4. Analog Realization of a Free-Running Multi-Vibrator Circuit.

In Figure 4, the white noise $\eta(\tau)$ is assumed to have a "amplitude" N which satisfies the inequality

$$N < \frac{A}{2} \quad (35)$$

where $\pm A$ is the binary output of the hysteresis relay element located in the feedback path of the integrator. By this means, the sign of the derivative $d\sigma(\tau)/d\tau$ is always opposite¹¹ to the sign of the output of the relay. The "trigger-levels" α and β of the hysteresis relay are chosen to coincide with the upper and lower bounds, respectively, of the compact, convex set U . In this way, the output $\sigma(\tau)$ of the integrator is constrained to always satisfy the inequality

$$\beta \leq \sigma(\tau) \leq \alpha \quad (36)$$

Moreover, since the sign of $\dot{\sigma}(\tau)$ is always opposite the sign of the relay output ($\pm A$), the function $\sigma(\tau)$ will continually oscillate, between the values α and β , in a semi-periodic and random-like manner. The nominal "frequency" of this oscillation is controlled by the effective gain of the relay output signal through the integrator. For the circuit shown in Figure 4 this gain has been set at unity (nominal = rate switch open). The closing of the rate switch shown in Figure 4 permits the effective integrator gain of the relay output signal to be increased by a factor of 1000 when an appropriate "switch close" logic command signal is received. In this way, the derivative $\dot{\sigma}(\tau)$ is increased to a comparatively large value and the function $\sigma(\tau)$ begins to oscillate at a high "frequency". This high frequency mode is continued until the rate switch is re-opened by an appropriate logic command signal.

The output $\sigma(\tau)$ of the integrator in Figure 4 is fed into the input of an analog "track and hold" device which is logically controlled by the sign of expression (34). This T/H device transmits (tracks) $\sigma(\tau)$ as long as expression (34) is non-positive and holds $\sigma(\tau) \equiv \sigma(\tau^*)$ when (34) becomes positive.

The sign of expression (34) is determined by ordinary analog function generation of $\langle \nabla C(x), F(x, \sigma) \rangle$, where x is obtained from the output of integrator F . The output of this function generator is used to generate logic command signals for the rate switch and the T/H device shown in Figure 4. In particular, when $\langle \nabla C(x), F(x, \sigma) \rangle \leq 0$ the rate switch is open and the T/H device is in the "tracking" mode. When $\langle \nabla C(x), F(x, \sigma) \rangle > 0$ the rate switch is closed and the T/H device is put into the "hold" mode.

¹¹It is recalled that an ordinary analog integrator has a "built-in" sign change associated with the integration process.

The E-detector determines when $\tilde{U}(x)$ is empty by counting the elapsed-time while the $\sigma(\tau)$ rate switch is closed. The condition that the rate switch has remained closed for more than one "period" of the high frequency scanning signal implies that (i) $\sigma(\tau)$ has effectively ranged over all values $\beta \leq \sigma(\tau) \leq \alpha$ and (ii) no values of σ have been found which will make $\langle \nabla C(x), F(x, \sigma) \rangle \leq 0$. In this event, the set $\tilde{U}(x)$ has become empty and $x(\tau) \in \partial R_0^1$. The E-detector can be realized by logically controlling the hold and reset relays of a single (constant) input analog integrator whose output drives a logic comparator. A circuit which realizes this operation is shown in Figure 5.

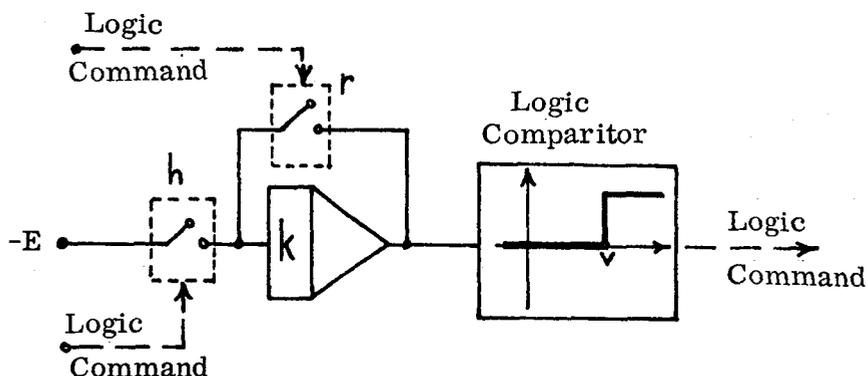


Figure 5. Analog Realization of the E-Detector

The negative constant $-E$ in Figure 5 represents a constant negative voltage which is permanently connected to the integrator input. When the $\sigma(\tau)$ rate switch is open, the logic controlled reset relay r is closed and the logic controlled hold relay h is open. Thus, the integrator output voltage in Figure 5 is constantly maintained at a zero level. When the $\sigma(\tau)$ rate switch closes, the relay r is opened and the hold relay h is closed thereby causing the integrator output to increase linearly with time τ at the rate Ek where k is the gain of the integrator. The comparator level v is chosen as: $v = Ek \Delta$ where Δ represents the established "period" of one high frequency $\sigma(\tau)$ oscillation. If the $\sigma(\tau)$ rate switch again re-opens before the integrator output exceeds the comparator level v the comparator does not generate a logic command signal and the integrator relays r and h are returned to their original positions. If the integrator output does exceed the comparator level v the comparator generates a logic command signal which: (i) places integrator F in a momentary hold, (ii) shifts the control gate G to the $u = \gamma(x, p)$ position, (iii) activates the S/H device for $C(x^*)$ and (iv) starts integrators F and P . In addition, the same logic signal can be used to command the recording of various problem variables via the central recorder R .

The synthesis of a random varying initial condition vector $x_0(\tau) \in \mathcal{J}$, for the F-integrator, can be accomplished by means of a suitably designed analog implicit function generator. For this purpose, the following (apparently original) scheme was developed. Let $\xi(\tau)$ be a uniformly bounded, random, differentiable, n-vector function generated from some external source, and let $z(\tau)$ be an n-vector function obtained by solving the system of first order integral equations

$$z(\tau) = k \int_0^\tau \mathcal{J}(\lambda(\zeta)) \nabla \mathcal{J}(\lambda(\zeta)) d\zeta \quad (37)$$

where $k > 0$ is a real, scalar constant, and

$$\lambda(\zeta) = \xi(\zeta) - z(\zeta) \quad (38)$$

It follows from (37), (38) that the vector $\lambda(\tau)$ satisfies the differential equation

$$\dot{\lambda}(\tau) = \dot{\xi}(\tau) - k \mathcal{J}(\lambda) \nabla \mathcal{J}(\lambda) \quad (39)$$

Thus, the scalar function $\mathcal{J}(\lambda)$ obeys the first order, ordinary differential equation

$$\frac{d\mathcal{J}(\lambda(\tau))}{d\tau} = \langle \nabla \mathcal{J}(\lambda), \dot{\xi}(\tau) - k \mathcal{J}(\lambda) \nabla \mathcal{J}(\lambda) \rangle \quad (40)$$

which can be written as

$$\frac{d\mathcal{J}(\lambda(\tau))}{d\tau} = -k \mathcal{J}(\lambda) \|\nabla \mathcal{J}(\lambda)\|^2 + \langle \nabla \mathcal{J}(\lambda), \dot{\xi}(\tau) \rangle \quad (41)$$

It has previously been assumed, [see Section 3], that $\nabla \mathcal{J}(x)$ exists for all $x \in D$. Now, assume further that the vector $\nabla \mathcal{J}(x)$ is non-null in some full dimensional, non-void, neighborhood of the manifold $\mathcal{J}(x) = 0$. In this case the "dynamic equilibrium solutions" of the forced first order equation (41), corresponding to $\frac{d\mathcal{J}(\lambda(\tau))}{d\tau} = 0$, are defined by

$$\mathcal{J}(\lambda) \Big|_{\text{equilibrium}} = \frac{\langle \nabla \mathcal{J}(\lambda), \dot{\xi}(\tau) \rangle}{k \|\nabla \mathcal{J}(\lambda)\|^2} \quad (42)$$

in a neighborhood of $\mathcal{J}(\lambda) = 0$. The numerator of (42) is always bounded by virtue of the previously stated assumptions. Thus, for initial conditions $\mathcal{J}(\lambda(0))$ sufficiently small, the dynamic equilibrium condition $\mathcal{J}(\lambda(\tau)) \equiv 0$ is obtained exactly for $\langle \nabla \mathcal{J}(\tau), \dot{\xi}(\tau) \rangle \equiv 0$ or $\dot{\xi}(\tau) \equiv 0$ and can otherwise be approximated arbitrarily close by choosing the positive scalar k sufficiently large. The corresponding randomly varying values of the n -vector $\lambda(\tau)$, obtained in this way, can be considered as random initial states $\lambda(\tau) = x_0(\tau)$ which satisfy (or closely approximate) the constraint $x_0(\tau) \in \mathcal{J}$.

A block diagram of this random initial condition generation scheme is shown in Figure 6.

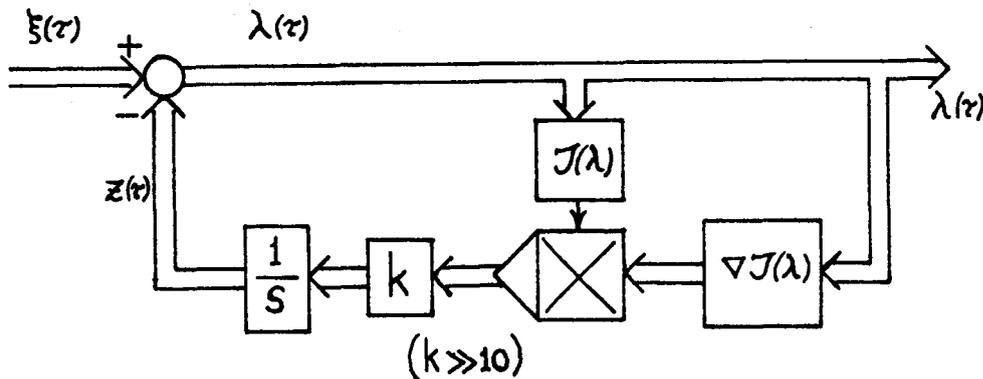


Figure 6. Initial Condition $x_0 \in \mathcal{J}$ Generation Scheme

A composite hybrid analog circuit for the complete algorithm, using the individual components described above, is illustrated in Figure 7.

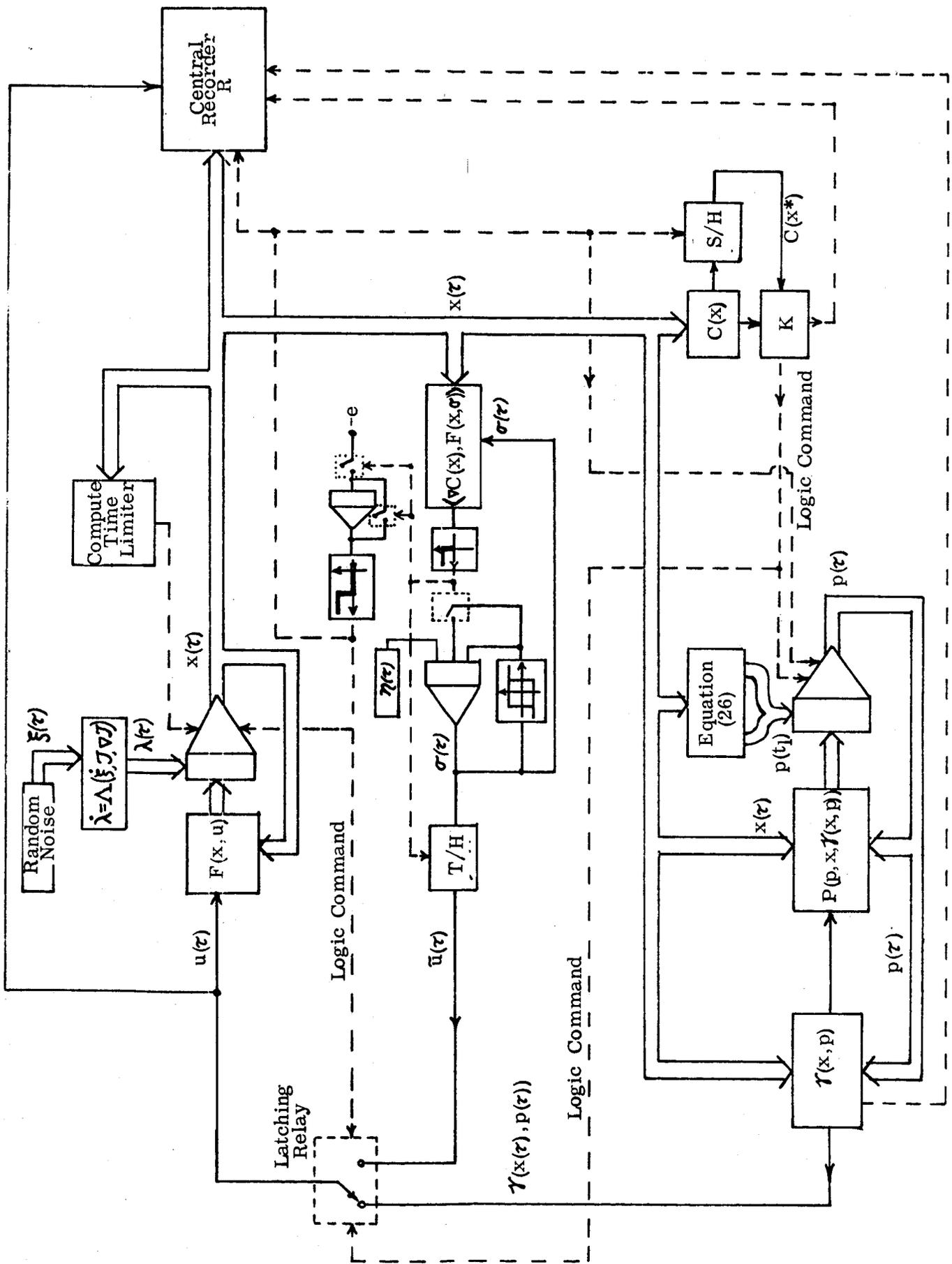


Figure 7. Hybrid Analog Circuit for the Complete Algorithm.

2.7 APPLICATION OF THE C-MINIMAX ALGORITHM TO A LOAD MINIMIZING CONTROL PROBLEM FOR A FIFTH-ORDER MODEL OF THE SATURN VEHICLE

The general algorithm described above can be used to study a variety of C-minimax optimization problems. In this section, the basic equations are derived for one particular application of the algorithm--the study of a certain "load minimizing" control system for the NASA Saturn Launch Vehicle. The dynamical model chosen for the present study is the standard fourth-order Saturn model augmented with one additional degree of freedom to permit the accommodation of certain physical constraints; namely, the gimble angle constraint, finite gimble angle "slewing" rate constraint and the fact that the load to be minimized (the maximum bending moment) is an explicit function of the gimble angle.

The state equations (3) for the standard (linearized) fourth-order Saturn model can be written as

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 \\ a_{21} & 0 & 0 & a_{24} \\ 0 & 0 & 0 & 1 \\ a_{41} & 0 & 0 & a_{44} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} + \beta(t) \begin{bmatrix} 0 \\ f_2 \\ 0 \\ f_4 \end{bmatrix} + w(t) \begin{bmatrix} 0 \\ b_2 \\ 0 \\ b_4 \end{bmatrix} \quad (43)$$

where the state variables (x_1, x_2, x_3, x_4) are defined as

x_1 = attitude angle error (ϕ)

x_2 = attitude angle error rate ($\dot{\phi}$)

x_3 = lateral drift of c. g. (z)

x_4 = rate of lateral drift of c. g. (\dot{z})

and the coefficients a_{ij}, f_i, b_i are given (in usual NASA symbols) as

$$\begin{aligned} a_{21} &= -c_1 & a_{24} &= c_1/v \\ a_{41} &= k_1 + k_2 & a_{44} &= -M'_\alpha/v \\ f_2 &= -c_2 & b_2 &= -c_1 \\ f_4 &= M'_\beta & b_4 &= M'_\alpha \end{aligned} \quad (44)$$

The scalar control $\beta(t)$ is the engine gimble angle, which is subject to the explicit inequality constraint $|\beta(t)| \leq \beta_{\max}$, and the disturbance $w(t)$ is a representation of the external wind force (usual NASA symbol α_w). In practice, the gimble angle rate $d\beta/dt$ is also bounded so that β cannot be changed instantaneously. In order to accommodate these constraints on the gimble angle and the gimble slewing rate the following additional first order dynamical equation is introduced

$$\dot{\beta} = -k_o(\beta - u(t)) \quad (45)$$

In (45), the constant $k_o > 0$ is a measure of the time-constant of the gimble positioning mechanism and $u(t)$ represents the actual (low power level) command signal (usually electrical) which activates the gimble positioning mechanism. It is assumed that admissible values of u are constrained by the inequality $|u(t)| \leq \beta_{\max}$ and the function $u(t)$ can experience simple jump discontinuities. Defining the new state variable $x_5 = \beta$, the expression (45) can be written as

$$\dot{x}_5 = -k_o(x_5 - u(t)) \quad (46)$$

which can be appended to the original set of equations (43). In this way, the gimble angle constraint, $|\beta| \leq \beta_{\max}$, is always naturally satisfied without introducing an explicit inequality constraint into the optimization problem--provided, of course, that the initial condition restriction $|x_5(t_o)| \leq \beta_{\max}$ is satisfied.

The structural bending moment M , induced on the vehicle by the engine thrust and the aerodynamic (wind) loads, can be (approximately) represented by the expression¹²

$$M(x, w(t)) = \langle c, x \rangle + M'_{\alpha} w(t) \quad (47)$$

where $c = (c_1, \dots, c_5)$ is a constant 5-vector with components defined by (usual NASA symbols)

$$\begin{aligned} c_1 &= M'_{\alpha} (= b_4) & c_3 &= 0 & c_5 &= M'_{\beta} (= f_4) \\ c_2 &= 0 & c_4 &= -(M'_{\alpha/V}) (= a_{44}) \end{aligned} \quad (48)$$

¹² $\langle x, y \rangle$ denotes the scalar product of the vectors x and y .

The design of a practical and reliable gimble angle control system is a rather involved and complicated problem, in general, owing to the presence of a variety of (often conflicting) performance criteria and physical and economic constraints. For this reason, only a very simplified version of this problem will be considered here.

One performance criterion which has been of particular concern to NASA control engineers is the maximum bending moment criterion. In this case, the primary factor of concern is that the maximum bending moment induced in the vehicle structure, during the interval of control, should be as small as possible--consistent with the other flight requirements. For example, the problem may be stated as follows. Let the state $x = (x_1, \dots, x_5)$ have the initial value $x(t_0) = x_0$ (x_0 might be zero) and suppose that the external wind disturbance $w(t)$ is a known function of time. Suppose, further, that the state $x(t)$ must satisfy the "terminal condition"¹³ $\mathcal{J}(x(t)) = 0$ at some future time $t = T$. Then among the set of all gimble angle positioning functions $\beta = \beta(t)$, $t_0 \leq t \leq T$ which satisfy the given physical constraints and boundary conditions find¹⁴ the (a) function $\beta = \beta^0(t)$ which minimizes the maximum (absolute) value of the bending moment $M(x(t), w(t))$, $t_0 \leq t \leq T$.

This non-autonomous problem (t appears explicitly in $F(\cdot, \cdot)$ and T is fixed) can be studied within the framework of the autonomous C-minimax theory developed in [1] by defining¹⁵ still another additional state variable

$$x_6 = t \tag{49}$$

where

$$x_6(t_0) = t_0, \tag{50}$$

and appending the additional differential equation

$$\dot{x}_6 = 1 \tag{51}$$

¹³Recall that the components x_1, x_2, x_3, x_4 of the state vector represent attitude and lateral drift "errors" (and their rates) with respect to the desired nominal flight path. These errors must be brought within certain tolerance limits as $t \rightarrow T$, in general.

¹⁴Alternatively, one could seek the function $u(t)$ defined in equation (45).

¹⁵See footnote #4 of [1].

to the set (43). By this device, the given, explicit, time function $w(t)$ is made to appear as a function of the state variable x_6 and the constraint $t = T$ is cast into the form of a state variable terminal condition $x_6(T) = T$ where T is no longer explicitly "fixed".

An alternative, and slightly simpler, version of the above problem is obtained by assuming that the interval¹⁶ of control $[t'_0, T]$ begins after the disturbance $w(t)$ has subsided but before the errors and error rates $x_1(t)$, $x_2(t)$, $x_3(t)$, $x_4(t)$ have been brought to zero. This is equivalent to the special case where $w(t) \equiv 0$ and the initial conditions $x_i(t_0)$, $i = 1, \dots, 4$, are not all zero. A further simplification results if the original terminal time T is assumed to be not explicitly fixed. Under these two assumptions there is no need to augment the state space with the additional coordinate x_6 .

The alternative (simplified) problem described above is particularly suited for illustrating practical application of the C-minimax algorithm developed herein. The basic equations for studying this problem, via the algorithm, will now be derived.

The simplified problem described above can be precisely stated as follows:

Statement of a Simplified Saturn Minimax Bending Moment Problem

Find a scalar control $u(t)$ which minimizes the functional¹⁷

$$J[u; x_0] = \max_{t_0 \leq t \leq T} \frac{1}{2} M^2(x(t)) = \max_{t_0 \leq t \leq T} \frac{1}{2} \langle c, x(t) \rangle^2 \quad (52)$$

subject to the restrictions

$$\dot{x} = \tilde{A}x + u(t)\tilde{f} \quad (53)$$

$$x(t_0) = x_0 ; \quad |x_3(t_0)| \leq \beta_{\max} \quad (54)$$

$$\sum_1^4 x_i^2(T) - r^2 = 0 \quad (T - \text{not restricted, } r > 0) \quad (55)$$

¹⁶That is, the particular sub-interval $[t'_0, T]$ during which the problem of minimum bending moment control is considered. Of course, the vehicle is actually "controlled" at all times of powered flight.

¹⁷The expression for $|M|$ is replaced by M^2 to permit continuous differentiation of the performance index. [See footnote #5 of [1]].

$$|u(t)| \leq \beta_{\max} \quad ; \quad \beta_{\max} = \text{fixed} > 0 \quad , \quad t_0 \leq t \leq T \quad (56)$$

where

$$\tilde{A} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ a_{21} & 0 & 0 & a_{24} & a_{25} \\ 0 & 0 & 0 & 1 & 0 \\ a_{41} & 0 & 0 & a_{44} & a_{45} \\ 0 & 0 & 0 & 0 & a_{55} \end{bmatrix} \quad ; \quad \tilde{f} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ \tilde{f}_5 \end{bmatrix} \quad (57)$$

the constants a_{21} , a_{24} , a_{41} , a_{44} , c_1 , \dots , c_5 are given by (44), (48) and

$$\begin{aligned} a_{25} &= f_2 & a_{45} &= f_4 \\ a_{55} &= -k_0 & \tilde{f}_5 &= +k_0 \end{aligned} \quad (58)$$

where f_2 , f_4 , k_0 are given by (44) and (45).

The terminal manifold $\mathcal{J}(x)$ has been chosen here as the 4-dimensional surface of the hypercylinder $\|\hat{x}\| = r$, $[\hat{x} = (x_1, \dots, x_4)]$, in the 5-dimensional state space.¹⁸ It is assumed, therefore, that the initial state x_0 lies outside of this hypercylinder and T is defined, implicitly, as the first time $t \geq t_0$ which satisfies $\sum_1^4 x_i^2(t) = r^2$.

Before discussing the programming of a numerical hybrid analog algorithm for this particular fifth-order problem it is instructive to consider some of the general analytical and geometric properties of the sought solution. For this purpose, the order n of the system will be left as an indeterminate.

The set $R_0^1 \supset \mathcal{J}$ for this problem is the largest set of states x with the following property. For each state $x_0 \in R_0^1$ there exists an admissible¹⁹ control $u(t)$, $t_0 \leq t \leq T$,

¹⁸The longitudinal axis of this hypercylinder coincides with the x_5 axis of the state space.

¹⁹Here, admissible means that $u(t)$ is piecewise continuous with values belonging to the compact, convex set $U = \{u: |u| \leq \beta_{\max}\}$.

such that the corresponding solution of (53), which starts at $x_0 \in R_0^1$, satisfies the weak inequality ('denotes transpose).

$$d/dt[\frac{1}{2} \langle c, x(t) \rangle^2] = \langle x(t), cc' \tilde{A}x(t) \rangle + u(t) \langle x(t), cc' \tilde{f} \rangle \leq 0, \quad t_0 \leq t \leq T \quad (59)$$

identically, and the boundary condition

$$\sum_1^{n-1} x_i^2(T) = r^2 \quad (60)$$

for some $T \geq t_0$. Thus, the set R_0^1 is a connected subset of the associated set Θ defined as

$$\Theta = \{x \mid \langle x, cc' \tilde{A}x \rangle + u \langle x, cc' \tilde{f} \rangle \leq 0, \text{ for some } |u| \leq \beta_{\max}\} \quad (61)$$

The admissible control value u which minimizes $d \langle c, x(t) \rangle^2 / dt$ is readily computed to be

$$\arg \min_{|u| \leq \beta_{\max}} [\langle x, cc' \tilde{A}x \rangle + u \langle x, cc' \tilde{f} \rangle] = -\text{sgn} \langle c, \tilde{f} \rangle [\beta_{\max} \text{sgn} \langle c, x \rangle] \quad (62)$$

provided that

$$\langle c, \tilde{f} \rangle \neq 0 \quad (63)$$

and

$$\langle c, x(t) \rangle \neq 0 \text{ for some positive interval of time.} \quad (64)$$

Thus, if (63) and (64) are satisfied, the set $\Theta \supset R_0^1$ can be described in the alternative, and more explicit, form

$$\Theta = \{x \mid \langle x, cc' \tilde{A}x \rangle - \beta_{\max} |\langle x, cc' \tilde{f} \rangle| \leq 0\} \quad (65)$$

It is shown in [1] that some subsets of ∂R_0^1 are defined by the equality in (65).

If $\langle c, \tilde{f} \rangle = 0$, the first time derivative of $\langle c, x(t) \rangle^2$ is not an explicit function²⁰ of the control $u(t)$. In this case, Θ is defined by

$$\Theta = \{x \mid \langle x, cc' \tilde{A}x \rangle \leq 0\} \quad (66)$$

It follows that, for this particular case, the boundary $\partial\Theta$ of Θ (recall that $\partial\Theta$ contains some subsets of ∂R_0^1 , in general) consists of the two (generally distinct) $(n-1)$ -dimensional hyperplanes H_1, H_2 defined by

$$H_1 = \{x \mid \langle c, x \rangle = 0\} \quad (67)$$

$$H_2 = \{x \mid \langle c, \tilde{A}x \rangle = 0\} \quad (68)$$

If $\langle c, x(t) \rangle \equiv 0$ is satisfied for some positive interval of time, by some appropriate choice of admissible control $u(t)$ (not necessarily $u(t) \equiv 0$), then $x(t)$ is forcibly restricted to the hyperplane H_1 and (59) is satisfied in the weak sense.²¹ This condition is usually realizable, in some proper, convex, subset $L \subset H_1$ by choosing $u(t)$ to be a linear function of $x(t)$ of the form $u(t) = \langle \mu, x(t) \rangle$ where²² μ is a constant n -vector. Further discussions of this topic are given in [3] and [4]. If the identity $\langle c, x(t) \rangle \equiv 0$ is satisfied for the special choice of zero control, $u(t) \equiv 0$, then $x(t)$ lies in some k -dimensional \tilde{A} -invariant subspace ($k \leq n$) of the n -dimensional state space. For example, $x(t)$ may lie along a one-dimensional real column eigenvector of \tilde{A} , a real two-dimensional eigenplane of \tilde{A} , etc. If $\langle c, x(t) \rangle \equiv 0$ is satisfied for every $x \in H_1$ (with $u(t) \equiv 0$) then H_1 is an invariant hyperplane²³ of \tilde{A} and the set L mentioned above equals H_1 . Moreover, when H_1 is an invariant hyperplane of \tilde{A} , it can be shown, [5], that

²⁰Some other, higher order, derivatives will be explicit functions of u , in general.

²¹The required boundary condition (60) might not be satisfied, however.

²²Portions of the boundary of the set $L \subset H_1$ are defined by $|\langle \mu, x \rangle| = \beta_{\max}$ (see [3]).

²³A general discussion of this subject is given in [5]; see also [2].

$$\langle c, \tilde{A} \rangle = \lambda c' \quad (69)$$

for some real eigenvalue λ of \tilde{A} . It follows that in this special case $H_1 = H_2$, provided²⁴ that $\lambda \neq 0$.

At each state $x \in R_0^1$ the value of the C-minimax optimal control $u^0 = \tilde{u}(x)$ is chosen²⁵ from the non-void state dependent set $\tilde{U}(x)$ defined by

$$\tilde{U}(x) = \{ |u| \leq \beta_{\max} \mid \langle x, cc' \tilde{A} x \rangle + u \langle x, cc' \tilde{f} \rangle \leq 0 \} \quad (70)$$

It is remarked that in the case of the special condition $\langle c, \tilde{f} \rangle = 0$ (eg. $\tilde{f} \in H_1$) described above, the control set $\tilde{U}(x)$ is non-void only in the closed set of states Θ defined by (66). Moreover, within the set Θ , the control set $\tilde{U}(x)$ (for this special case) coincides with the original control set $U = \tilde{U}(x) = \{u \mid |u| \leq \beta_{\max}\}$. From (11) the value function $V(x)$, in R_0^1 , is given by

$$V(x) = \frac{1}{2} \langle c, x \rangle^2 \quad \forall x \in R_0^1 \quad (71)$$

The auxiliary Problem M, for this particular example, can be stated as follows: Find an admissible control $u(t)$ which minimizes the functional

$$\tilde{J}[u; x_0] = \frac{1}{2} \langle c, x(t_1) \rangle^2 \quad (72)$$

subject to the differential equation constraint (53) and the restrictions

$$x(t_0) = x_0 \in D - R_0^1 \quad ; \quad |x_n(t_0)| \leq \beta_{\max} \quad (73)$$

$$x(t_1) \in \partial R_0^1 \quad t_1 - \text{unrestricted} \quad (74)$$

$$|u(t)| \leq \beta_{\max} \quad t_0 \leq t \leq t_1 \quad (75)$$

²⁴Note that $\lambda = 0$ implies $\langle c, \tilde{A} \rangle = 0$ which implies $H_2 = E^n$ ($H_1 \neq H_2$) where E^n is the system state space. In this case, if $H_1 = \mathcal{H}_1$, then every hyperplane parallel to \mathcal{H}_1 , is also an invariant hyperplane for the solutions of $\dot{x} = \tilde{A}x$, [5].

²⁵It should be stressed that, in general, not every value $u \in \tilde{U}(x)$ is optimal for a given state $x \in R_0^1$.

according to the Euler-Pontryagin necessary conditions for this Problem M, the optimal control $u^0 = \gamma(x, p)$ is given by

$$\gamma(x, p) = \arg \max_{|u| \leq \beta_{\max}} \langle p, \tilde{A}x + u\tilde{f} \rangle \quad (76)$$

where the two n-vectors $x(t), p(t)$ obey the differential (canonical) equations

$$\dot{x} = \tilde{A}x + \gamma(x, p)\tilde{f} \quad (77)$$

$$\dot{p} = -\tilde{A}'p \quad (78)$$

and satisfy the two-point boundary conditions

$$x(t_0) = x_0 \quad (79)$$

$$p(t_1) = \begin{cases} -cc'x(t_1); & \text{if } x(t_1) \in (\partial R_0^1 - \mathcal{J}) \\ \pi \text{ col. } (x_1(t_1), x_2(t_1), x_3(t_1), x_4(t_1), 0); & \text{if } x(t_1) \in \mathcal{J} \end{cases} \quad (80)$$

where π is a real, scalar constant. It follows that the optimal control for Problem M is of the so-called "bang-bang" type

$$\gamma(x, p) = \beta_{\max} \operatorname{sgn} \langle p, \tilde{f} \rangle \quad (81)$$

provided that the singular condition

$$\langle p(t), \tilde{f} \rangle \equiv 0 \quad (82)$$

does not occur for some positive interval of time along a non-trivial solution of (78). If the singular condition (82) does occur for some positive interval of time then (81) fails to give any information about the optimal control. In this case, special (singular solution) techniques must be used to study the problem [2]. It is easily verified that the identical vanishing of $\langle p(t), \tilde{f} \rangle$ ($p(t) \neq 0$), implies that the sequence of vectors

$$\tilde{f}, \tilde{A}f, \tilde{A}^2f, \dots, \tilde{A}^{n-1}\tilde{f} \quad (83)$$

are not linearly independent. This latter condition is recognized as the well-known necessary and sufficient condition that (53) be "not completely controllable" in the sense of Kalman [6]. Thus, if (53) is assumed to be completely controllable then singular solutions to Problem M will not occur and the optimal control for Problem M will always be a bang-bang type function well defined²⁶ by (81).

The canonical equations (77) - (78), with $\gamma(x, p)$ given by (81), are a set of $2n$ piecewise linear (constant coefficient) ordinary differential equations which (in principal) can be analytically solved by straight forward application of the fundamental matrix technique[3, #7]. By this means, the bang-bang switching surface [the set of points $\{x\}$ where $\langle p, f \rangle$ experiences (isolated) zeros] can be identified in the subset $D - R_0^1$ of the system state space. Moreover, along the (optimal) solutions of the canonical equations (77), (78), the first integral

$$[\langle p, \tilde{A}x \rangle + \beta_{\max} |\langle p, \tilde{f} \rangle|] = 0 \quad (84)$$

is naturally satisfied and can be used to obtain additional information about the solution to Problem M. In particular, this result provides an additional relationship between $(p(t_0), x(t_0))$ and $(p(t_1), x(t_1))$.

The set R_n^1 is determined as the largest subset of $D - R_0^1$ with the property that along each optimal trajectory of Problem M (in R_n^1) the value of $\langle c, x(t) \rangle^2$, $t_0 \leq t < t_1$, never exceeds the value $\langle c, x(t_1) \rangle^2$, $x(t_1) \in \partial R_0^1$. In the set R_n^1 , the function $V(x)$ can be written as

$$V(x) = \frac{1}{2} \langle c, x(t_1) \rangle^2 \quad (85)$$

where $x(t_1)$ is a function of the initial condition $x_0(t_0) \in R_n^1$.

²⁶It is interesting to note that the set of C-minimax canonical equations (77), (78), with $\gamma(x, p)$ given by (81), are precisely the same form as the set of canonical equations for the time-optimal [7] control problem for, (53) - (56), in the set $D - R_0^1$. The only essential difference between these two sets of canonical equations is the required boundary conditions for $p(t_1)$. The relationship between the bang-bang switching surfaces for the C-minimax problem and the time-optimal problem (in the set $D - R_0^1$) is an interesting area for further research.

The Mayer-type Problem M described above can be cast, alternatively, as an equivalent Lagrange (integral-type) optimization problem by defining the additional state variable

$$x_{n+1} = \frac{1}{2} \langle c, x \rangle^2 \quad (86)$$

In this way, the functional (72) can be written as the time-integral of a "quadratic state-linear control" function

$$J[u; x_0] = \int_{t_0}^{t_1} [\langle x(t), Qx(t) \rangle + u(t) \langle x(t), g \rangle] dt \quad (87)$$

where $x = (x_1, \dots, x_{n+1})$, Q is an $(n+1) \times (n+1)$ constant matrix given by

$$Q = \left[\begin{array}{c|c} cc' \tilde{A} & 0 \\ \hline 0 & 0 \end{array} \right] \quad (88)$$

and g is the constant $(n+1)$ -vector

$$g = \begin{pmatrix} cc' \tilde{f} \\ \hline 0 \end{pmatrix} \quad (89)$$

The additional scalar state variable equation for x_{n+1} is

$$\dot{x}_{n+1} = \langle x, Qx \rangle + u(t) \langle x, g \rangle \quad (90)$$

which can be appended to the original set (53). It is noted that (90) is not linear in x . Another alternative format for Problem M is obtained by recalling that $|\langle c, x \rangle|$ and $\langle c, x \rangle^2$ have their minimum at the same x . Thus, one could define the additional state variable x_{n+1} in (86) alternatively as

$$x_{n+1} = |\langle c, x \rangle| \quad (91)$$

and the functional (72) could be replaced by the variationally equivalent time integral of a "(piecewise) linear state-linear control" function

$$\hat{J}[u; x_0] = \int_{t_0}^{t_1} [(\text{sgn } \langle \hat{c}, x(t) \rangle)(\langle h, x(t) \rangle + r u(t))] dt \quad (92)$$

where

$$\hat{c} = \begin{pmatrix} c \\ 0 \end{pmatrix}, \quad h = \begin{pmatrix} \tilde{A}'c \\ 0 \end{pmatrix}, \quad r = \langle c, \tilde{f} \rangle \quad (93)$$

and x_{n+1} obeys the state variable equation

$$\dot{x}_{n+1} = [\langle h, x \rangle + r u(t)](\text{sgn } \langle \hat{c}, x \rangle) \quad (94)$$

It is observed that, unlike (90), equation (94) is a piecewise linear ordinary differential equation. The study of Lagrange variational problems, with piecewise linear discontinuous integrands of the form (92), represents another interesting area for further research.

The construction of a C-minimax hybrid analog algorithm for the particular fifth-order Saturn minimax bending moment problem described above is accomplished by straightforward application of the procedure outlined in Section 6. For this purpose, the F and P integrator devices are realized by standard (linear) analog programming of summing and integrating amplifiers where

$$F(x, u(\tau)) = \tilde{A}x + u(\tau)\tilde{f} \quad (95)$$

$$P(p, x, \gamma(p, x)) = -\tilde{A}'p \quad (96)$$

It is noted that the right side of (96) is not an explicit function of the state x or the control function $\gamma(p, x)$.

The $\tilde{U}(x)$ set computer is realized by means of the special circuit shown in Figure 7 where the function $\langle \nabla C(x), F(x, \sigma) \rangle$ has the form

$$\langle \nabla C(x), F(x, \sigma) \rangle = \langle x, cc' \tilde{A}x \rangle + \sigma \langle x, cc' \tilde{f} \rangle \quad (97)$$

Expression (97) can be generated directly by standard analog non-linear function generation schemes. However, such a scheme will require a large amount of analog multiplication. Alternatively, the right side of expression (97) can be re-written in the form

$$\langle x, cc' \tilde{A}x \rangle + \sigma \langle x, cc' \tilde{f} \rangle = \langle c, x \rangle [\langle \tilde{A}c, x \rangle + \sigma \langle c, \tilde{f} \rangle] \quad (98)$$

so that the detection of the inequality

$$\langle \nabla C(x), F(x, \sigma) \rangle > 0 \quad (99)$$

is equivalent to the detection of the condition

$$\text{sgn} \langle c, x \rangle = \text{sgn} [\langle \tilde{A}c, x \rangle + \sigma \langle c, \tilde{f} \rangle] \quad (100)$$

which can be accomplished, without the aid of analog multiplication, by the simple logic scheme shown in Figure 8.

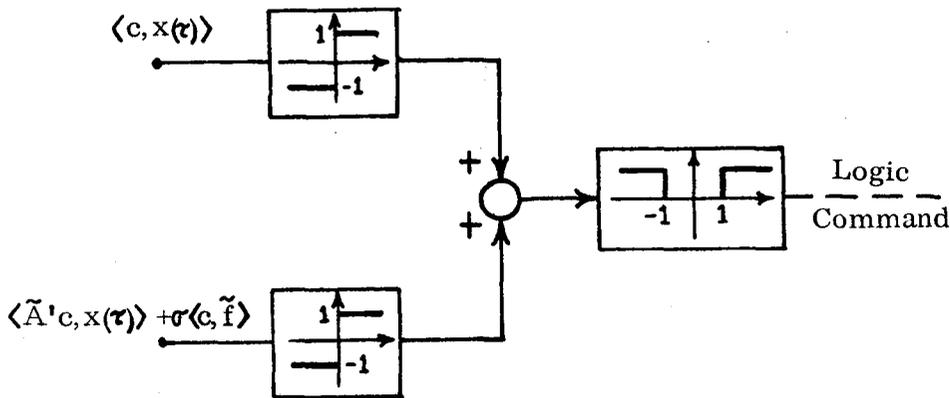


Figure 8. Logic Scheme for Detecting the Inequality (99).

It is remarked that the scheme shown in Figure 8 can be used even when $\langle c, \tilde{f} \rangle = 0$ in which case $\tilde{U}(x) = U$ (and therefore $U(x)$ is non-empty) if and only if (100) is not satisfied. In this case the separate E-detector circuit shown in Figure 7 can be eliminated. Otherwise, the E-detector and the $\tilde{u}(\tau) \in \tilde{U}(x(\tau))$ control selector for this problem can be realized by implementation of the circuits shown in Figure 7.

The $x_0 \in \mathcal{J}$ initial condition circuit for this problem can be realized in the manner illustrated in Figure 6. For this purpose, it is necessary to generate the scalar expression

$$\mathcal{J}(\lambda) = \lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 - r^2 \quad (101)$$

and the 5-vector

$$\frac{1}{2} \nabla \mathcal{J}(\lambda) = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ 0 \end{bmatrix} \quad (102)$$

and form the (scalar · vector) product

$$k \mathcal{J}(\lambda) \nabla \mathcal{J}(\lambda) = k(\lambda_1^2 + \lambda_2^2 + \lambda_3^2 + \lambda_4^2 - r^2) \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \lambda_4 \\ 0 \end{bmatrix} \quad (103)$$

where k is a "large", positive, scalar constant. The vector variable $\lambda(\tau)$ is obtained by subtracting the (vector) integration of (103) from the 5-vector (externally generated) random variable $\xi(\tau) = (\xi_1(\tau), \dots, \xi_5(\tau))$ where $\xi_5(\tau)$ satisfies the inequality constraint

$$|\xi_5(\tau)| \leq \beta_{\max} \quad \forall \tau \quad (104)$$

In this way, a 5-channel analog feedback circuit is obtained which generates the five outputs $(\lambda_1(\tau), \lambda_2(\tau), \dots, \lambda_5(\tau))$. As explained in Section 6, these five variables can be used as "initial values" for the five state variables $(x_1(\tau), \dots, x_5(\tau)) \in \mathcal{J}$ if $\xi(\tau)$ is sufficiently small and/or k is sufficiently large. An analog circuit constructed in this manner is shown in Figure 9.

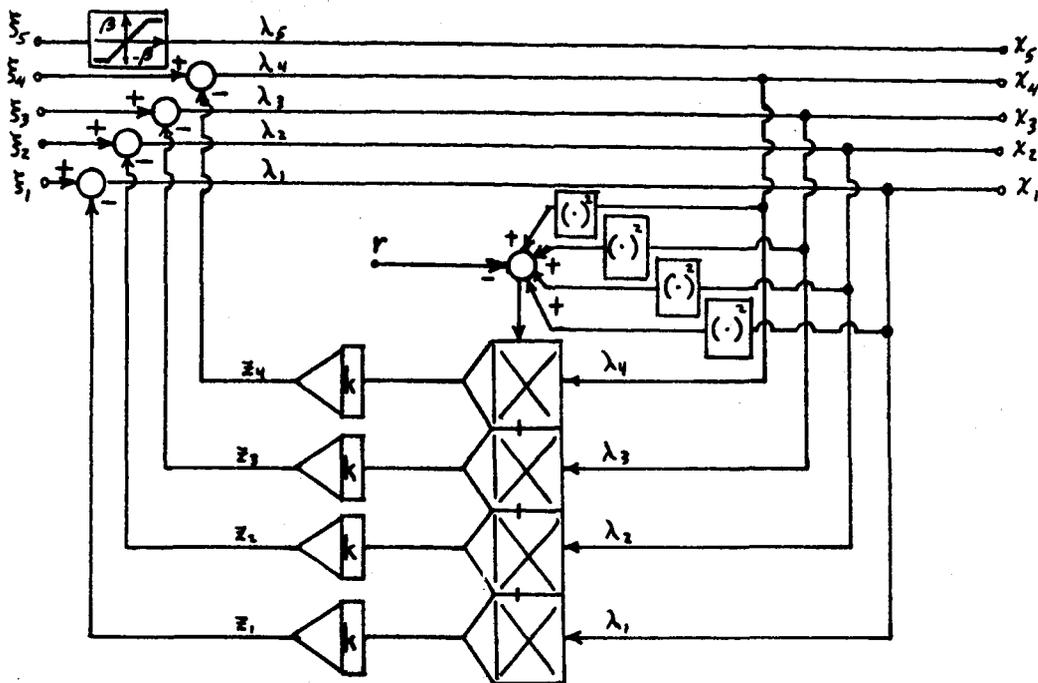


Figure 9. Analog Circuit for Generating Random Initial Conditions $x_0(\tau) \in \mathcal{J}$.

On the boundary of the $\{R_n\}$ regions the "initial condition" $p(t_1)$ for the P-integrator is generated by forming the appropriate linear function of $x(t_1)$ as prescribed in equation (80). Likewise, the control function $\gamma(x, p)$ for Problem M, prescribed by (81), can be generated by driving an analog "ideal relay" (output = $\pm \beta_{max}$) with the scalar input $\langle p(\tau), \hat{f} \rangle$ —provided that singular solutions do not exist.

The computation of the performance index $C(x(\tau)) = \frac{1}{2} \langle c, x(\tau) \rangle^2$, as required by the algorithm, can be performed by straightforward continuous analog squaring of the linear expression $\langle c, x(\tau) \rangle$. Alternatively, one can compute, instead of $C(x(\tau))$, the expression $\hat{C}(x(\tau)) = |\langle c, x(\tau) \rangle|$ since the two functions $C(\cdot)$ and $\hat{C}(\cdot)$ have their maxima and minima in common. This latter procedure has the advantage that it does not require an analog squaring device. The sampling and storage of $C(x^*)$ [or $\hat{C}(x^*)$] is accomplished by a standard analog sample-hold (S/H) device as shown in Figure 7.

The remaining components required by the hybrid-analog algorithm have no special configuration for this particular fifth-order example and can be effectively instrumented in the manner illustrated in the general circuit of Figure 7. A complete circuit of the hybrid-analog algorithm, for this particular example, is illustrated in Figure 10.

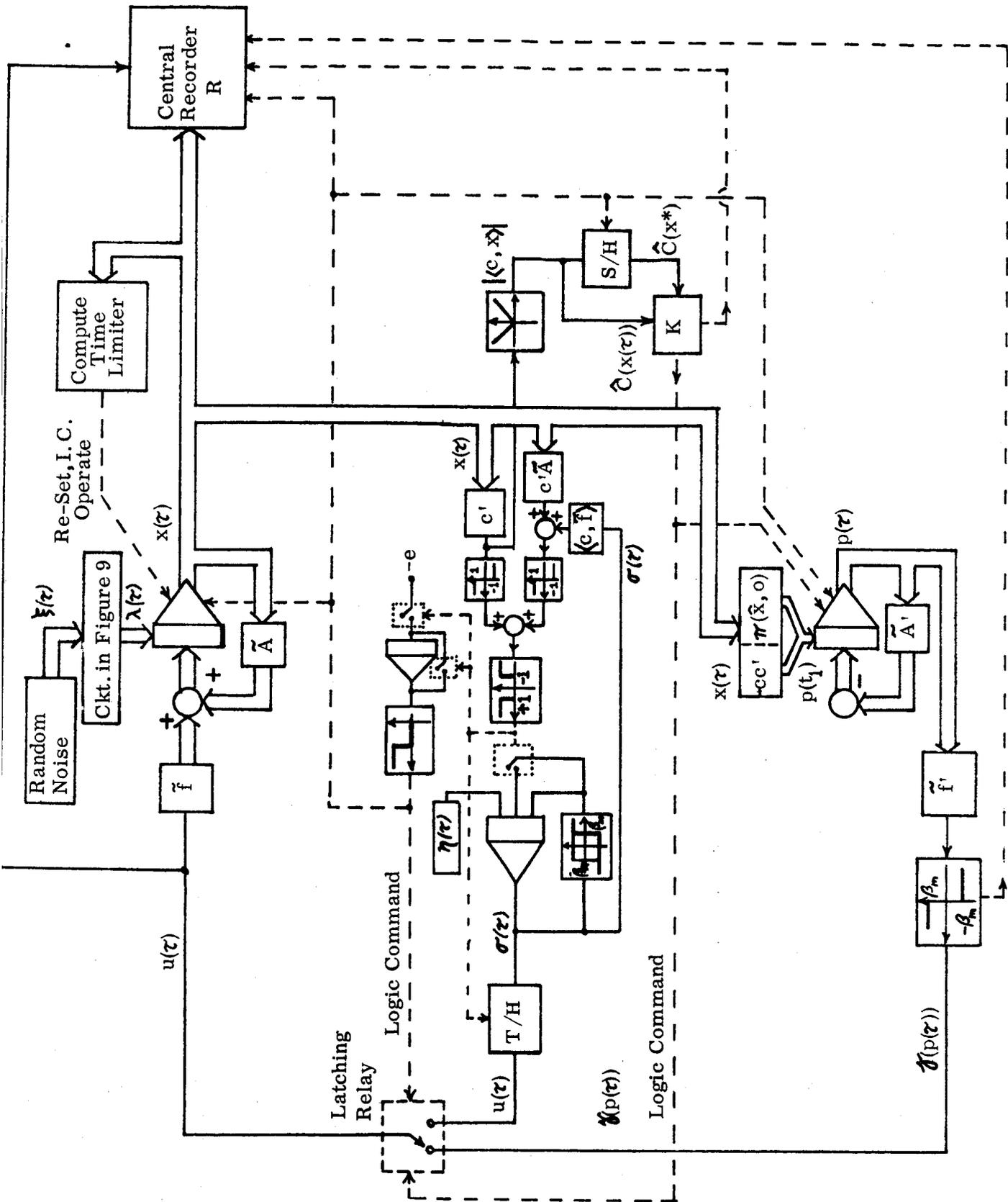


Figure 10. Hybrid Analog Realization of the Algorithm for the Special Fifth-Order Saturn Load Minimizing Example.

2.8 ACCOMMODATION OF EXPLICIT INEQUALITY CONSTRAINTS ON SYSTEM STATE VARIABLES

The principal tasks outlined in the contract for this study included an investigation of means for incorporating explicit state variable inequality constraints in the solution of Chebyshev Minimax control problems. This subject was studied from both the exact and approximate point of view and several alternative methods of solution were developed.

The explicit state variable inequality constraints considered in this study can be represented by the expressions

$$g_i(x(t)) \leq 0 \quad \begin{array}{l} i = 1, \dots, m \\ t_0 \leq t \leq T \end{array} \quad (105)$$

where the $g_i(x)$, $i = 1, \dots, m$, are continuous scalar valued functions of the state x . Thus, the particular class of C-minimax control problems with constrained state variables considered here can be stated as follows. Find a piecewise continuous control $u(t)$ which minimizes the functional (2) subject to the usual restrictions (3) - (6) and the inequalities (105).

State variable constraints of the type (105) arise naturally in a variety of optimization problems and the accommodation of such constraints, in conventional Bolza-type optimal control problems, has been a subject of considerable interest in recent years. From those studies, several alternative methods of solution have been proposed. These proposed methods can be classified into two main categories: (i) exact methods and (ii) approximate methods. The proposed exact methods are obtained by either suitable modification of the conventional theory for the Bolza problem (eg. modification of the "multiplier rule") [8], or by the introduction of certain nonlinear functional transformations [9] which effectively transform the closed set of states, defined by (105), into an equivalent open set which can then be studied by the conventional theory.

The proposed approximate methods of solution [10] - [12] are based on various forms of the "penalty function" technique introduced by Courant [13]. This technique attempts to indirectly discourage violations of the inequalities (105) by imposing severe performance penalties (degradation in system performance) whenever those inequalities are not satisfied. By this means, if the penalties are chosen sufficiently strong, the resulting optimal trajectories will tend to avoid violations of the constraints.

The methods described above for accommodating state variable inequality constraints in conventional (Bolza-type) optimization problems can be adapted to the class C-minimax optimization problems considered in the present study. In this section several such methods are discussed in detail.

2.8.1 EXACT METHODS - The most natural "exact method" for accommodating inequality constraints of the form (105) appears to be the method which defines and identifies the sets $\{R_0\}$ and $\{R_m\}$ on the particular subset $\hat{D} \subset D \subset E^n$ where \hat{D} is the largest set of states $x \in D$ with the following property. For each initial state $x_0 \in \hat{D}$ there exists at least one admissible control $u(t)$ such that the corresponding solution of (3) satisfies (4), (5) and (105). Thus the set \hat{D} is the largest set of states $x \in D$ which remain controllable to \mathcal{J} in the presence of the constraints (105). The set R_0^1 , in this case, is defined to be the largest set of states $x \in \hat{D}$ which can be controlled to the terminal manifold \mathcal{J} (by an admissible control) along a trajectory which satisfies (105) and

$$\frac{dC(x(t))}{dt} \leq 0 \quad t_0 \leq t \leq T \quad (106)$$

It is remarked that such trajectories may contain one or more subarcs which lie on the constraint surfaces defined by the equalities in (105). Moreover some subsets of the boundary ∂R_0^1 may also lie on one or more of these constraint surfaces. With $R_0^1 \in D$ identified, the auxiliary Problem \hat{M} is defined on the set $D - R_0^1$ as follows. Find an admissible control $u(t)$, $t_0 \leq t \leq t_1$, which minimizes the functional (12) subject to the restrictions

$$\dot{x} = F(x, u(t)) \quad (107)$$

$$x(t_0) = x_0 \in (\hat{D} - R_0^1) \quad (108)$$

$$x(t_1) \in \partial R_0^1 \quad ; \quad t_1 - \text{unrestricted} \quad (109)$$

and the state variable inequality constraints

$$g_i(x(t)) \leq 0 \quad i = 1, \dots, m \quad t_0 \leq t \leq t_1 \quad (110)$$

The set R_m^1 is then identified as the largest set of initial states $x_0 \in (\hat{D} - R_0^1)$ with the property that along the corresponding solution trajectory of Problem \hat{M} , defined above, the inequality

$$C(x(t)) \leq C(x(t_1)) \quad t_0 \leq t < t_1 \quad (111)$$

is satisfied identically. The auxiliary Problem \hat{M} is recognized as a conventional Mayer-type optimization problem with bounded state variables. The exact

analytical solution to this problem can be obtained, in principle, by straightforward application of the modified "multiplier rule" for the Pontryagin Maximum Principle as described in Chapter VI of [8] (see also [14] and [15])

The sets R_0^2 , R_m^2 , in \hat{D} are identified, as before, by repeating the process described above using the boundary of $R_0^1 \cup R_m^1$ as the new terminal manifold. Continuing in the manner, the set \hat{D} can be completely partitioned into the two families of sets $\{R_0\}$, $\{R_m\}$ and the "exact" C-minimax optimal control can be determined for each initial state $x_0 \in \hat{D}$.

The computational algorithm for identifying the set R_0^1 , as developed in this report, can be easily modified to accommodate the constraints (105). For this purpose it is only necessary to re-define the state dependent set (24) as²⁷

$$\tilde{U}(x) = \left\{ \begin{array}{l} \{u \in U \mid \langle \nabla C(x), F(x, u) \rangle \leq 0\}; \text{ if } g_i(x) < 0 \quad i = 1, \dots, m \\ \left\{ \begin{array}{l} \langle \nabla C(x), F(x, u) \rangle \leq 0 \text{ and} \\ \langle \nabla g_i(x), F(x, u) \rangle \leq 0, \quad \forall g_i(x) = 0 \end{array} \right\}; \text{ if } g_i(x) \neq 0 \quad i = 1, \dots, m \end{array} \right. \quad (112)$$

It is remarked that a state $x \in D$, which satisfies $g_i(x) = 0$ for some $i = 1, \dots, m$ cannot belong to $R_0^1 \subset \hat{D}$ unless the corresponding set $\tilde{U}(x)$, defined by (112), is non-void. On the other hand, the computational algorithm for identifying the set R_m^1 will require certain, non-minor modifications in order to accommodate the constraints (105) owing to the more complex "modified multiplier rule" which must be instrumented for Problem \hat{M} .

An alternative "exact" method for solving C-minimax problems with state variable constraints of the form (105) consists of introducing the m additional state variables x_{n+1} , x_{n+2} , ..., x_{n+m} and appending the additional state equations

$$\dot{x}_j = k_j [g_j^2(x) + g_j(x) |g_j(x)|] \quad j = n+1, \dots, n+m \quad (113)$$

where k_j is a real, positive constant, to the original set (3). This augmented state problem is then solved by regular C-minimax techniques, ignoring the explicit constraints (105), where the boundary conditions for the "states" x_{n+1} , ..., x_{n+m} are specified to be

²⁷ If $|g_i(x)|$ is not continuous, the second inequality in the last of (112) should be replaced by the condition that $u \in U$ does not "point" the local velocity vector $F(x, u)$ into the region in which $g_i(x) > 0$.

$$x_j(t_0) = x_j(T) \quad j = n+1, \dots, n+m \quad (114)$$

It is clear from (113) that the conditions (114) can be satisfied if and only if the constraints (105) are satisfied identically. This alternative method can be effectively applied through the computational algorithm developed in Sections 5 and 6 of the present report. It should be noted, however, that in this case not every backward time trajectory generated by the algorithm is optimal, in general, owing to the presence of the special boundary conditions (114).

2.8.2 PENALTY FUNCTION METHODS - The so-called penalty function methods seem to be the most natural "approximate" method for solving C-minimax control problems with state variable constraints of the form (105). For this purpose, the original performance $C(x)$ in (2) is augmented with an additional (additive) set of scalar terms $b_i(x)$, $i = 1, \dots, m$ which have a zero (or approximately zero) value whenever the corresponding constraints (105) are satisfied and which have a relatively large positive value whenever the corresponding constraints are violated. Thus, in the ideal case the modified C-minimax performance index

$$\mathcal{C}(x) = C(x) + \sum_1^m b_i(x) \quad (115)$$

is effectively equal to the original performance index $C(x)$ as long as the constraints (105) are not violated. When a violation does occur, the modified performance index is dominated by the term $\sum_1^m b_i(x)$ so that, effectively, $\mathcal{C}(x) \approx \sum_1^m b_i(x)$.

In practice it is desirable to choose the functions $b_i(x)$ such that, in the neighborhood of the constraint surfaces $g_i(x) = 0$, the value of $b_i(x)$ gradually increases (from ≈ 0) as $g_i(x) \rightarrow 0$ ($g_i(x) < 0$) and rapidly increases as $g_i(x)$ exceeds the value of zero. Appropriate choices for the "rate of increase" of the terms $b_i(x)$ depend upon the amount of penetration of the constraint surfaces which can be tolerated and, in general, must be determined by experimental (trial and error) techniques.

There are many admissible functions $b_i(x)$ which can be used for this purpose. For example one can choose the $b_i(x)$ as the two-parameter expression

$$b_i(x) = \epsilon_1 (g_i(x) - \epsilon_1)^{-2\nu_i} \quad \begin{array}{l} \nu_i = 1, 2, \dots, \\ i = 1, \dots, m \end{array} \quad (116)$$

where ϵ_1 is a "small" positive constant. This positive valued function, which is less than ϵ_1 in value for $g_i(x) < (\epsilon_1 - 1)$, will limit penetration of the constraint surface $g_i(x) = 0$ to be less than ϵ_1 , provided that $C(x)$ is bounded.

An alternative choice for $b_i(x)$, which has the desirable feature of being identically zero in the region $g_i(x) < 0$, outside of a small neighborhood of the surface $g_i(x) = 0$, is given by the three-parameter expression

$$b_i(x) = k_i \left[(g_i(x) + \epsilon_i)^{2\nu_i} + (g_i(x) + \epsilon_i)^{2\nu_i - 1} |g_i(x) + \epsilon_i| \right]; \quad \nu_i = \frac{1}{2}, 1, 2, 3, \dots \quad (117)$$

$i = 1, \dots, m$

where k_i and ϵ_i are, respectively, "large" and "small" positive constants. The non-negative function (117) is identically zero in value for $g_i(x) \leq -\epsilon_i$, has the positive value $2k_i \epsilon_i^{2\nu_i}$ when $g_i(x) = 0$, and has the form $2k_i (g_i(x) + \epsilon_i)^{2\nu_i}$ when $g_i(x) \geq 0$. Thus, by appropriate choice of the parameters (k_i, ϵ_i, ν_i) , the properties of the function (117) can approximate, arbitrarily close, the ideal characteristics described above.

2.8.3 WEAKENING CONTROL SET METHODS - Another method for accommodating the state variable constraints (105) consists of re-defining the original admissible control set U in such a way that, whenever $g_i(x) = 0$ for one or more $i = 1, \dots, m$, the original set U is reduced (weakened) to include only those $u \in U$ which do not cause penetration of the constraint surface(s) $g_i(x) = 0$. For this purpose, one can define the new admissible set U^* as follows:

$$U^*(x) = \begin{cases} U & \text{if } g_i(x) < 0, \forall i = 1, \dots, m \\ U_w(x) & \text{if } g_i(x) \leq 0, \forall i = 1, \dots, m \end{cases} \quad (118)$$

where

$$U_w(x) = \{u \in U \mid \langle \nabla g_i(x), F(x, u) \rangle \geq 0, \forall g_i(x) = 0, i = 1, \dots, m\} \quad (119)$$

With $U^*(x)$ defined by (118) - (119), the explicit constraints (105) can be disregarded and the C-minimax problem can be solved by ordinary means.²⁸ This method of solution is complicated by the presence of the, possibly discontinuous, state dependent control set $U^*(x)$ which must be incorporated in the solution of the Mayer Problem M. A mathematical theory for a relatively general class of Mayer

²⁸It should be noted from (119) that a C-minimax optimal control for an initial state $x_0 \in D$, where $g_i(x_0) = 0$ (for some $i = 1, \dots, m$), will not exist if the set $U_w(x_0)$ is empty.

variational problems with state dependent control sets is described in [16]. In practical applications, it would probably be advantageous (if not essential) to replace the set $U^*(x)$ with an approximating set $\tilde{U}^*(x)$ which possesses certain continuity, differentiability, and/or convexity properties. This may be necessary, for example, to insure the existence [17] of a piecewise continuous optimal control--since (unlike U), the set $U^*(x)$ will not be convex, in general. The problem of choosing mathematically appropriate and physically realistic approximating sets $\tilde{U}^*(x)$ is an interesting area for further research.

REFERENCES

1. "Optimal Control with Chebyshev Minimax Performance Index," Johnson, C. D., ASME Transactions, Jour. of Basic Engineering, Vol. 89, Series D, No. 2, June 1967. Also, Chapt. 2 of NASA Contractor Report No. CR-715 entitled "Study of Optimal and Adaptive Control Theory," April 1967.
2. "Singular Solutions in Problems of Optimal Control," Johnson, C. D., Chapt. 4 in Advances in Control Systems; Theory and Applications, Vol. 2 (book), ed. by C. T. Leondes, Academic Press, 1965.
3. "Optimal Bang-Bang Control with Quadratic Performance Index," Wonham, W. M. and Johnson, C. D., ASME Trans., Jour. of Basic Engineering, Vol. 86, Series D, No. 1, pp. 107-115, March 1964.
4. "On a Problem of Letov in Optimal Control," Johnson, C. D. and Wonham, W. M., ASME Trans., Jour. of Basic Engineering, Vol. 87, Series D, No. 1, pp. 81-89, March 1965.
5. "Invariant Hyperplanes for Linear Dynamical Systems," Johnson, C. D., IEEE Trans. on Automatic Control, Vol. AC-11, pp. 113-116, January 1966.
6. "Controllability of Linear Dynamical Systems," Kalman, R. E., Ho, Y. C., and Narendra, K. S., Contribs. to Differential Equations, Vol. 1, No. 2, pp. 189-213, 1962.
7. "Optimal Control; An Introduction to the Theory and Its Applications," Athans, M. and Falb, P. L., McGraw-Hill Book Co., New York, 1966.
8. "The Mathematical Theory of Optimal Processes," Pontryagin, L. S., Boltyanskii, V. G., Gramkrelidze, R. V., and Mishchenko, E. F., John Wiley and Sons, New York, 1962.
9. "Non-Linear Substitutions in Variational Problems," Vorobev, L. M. and Vorobeva, T. M., Automation and Remote Control, Vol. 27, No. 5, pp. 157-166, May 1966.

10. "Method of Gradients," Kelley, H. J., Chapt. 6 in Optimization Techniques (book) ed. by G. Leitmann, Academic Press, New York, 1962.
11. "Some Mathematical Theory of the Penalty Method for Solving Optimum Control Problems," Okamura Kiyohisa, SIAM Jour. on Control, ed. 2, No. 3, pp. 317-331, 1964.
12. "Penalty Functions and Bounded Phase Coordinate Control," Russell, D. L., SIAM Jour. on Control, Vol. 2, No. 3, pp. 409-422, 1964.
13. "Calculus of Variations and Supplementary Notes and Exercises," Courant R., New York University Institute of Mathematical Sciences, New York, 1956-1957.
14. "On Control Problems with Bounded State Variables," Berkovitz, L. D., Jour. of Math. Analysis and Applications, Vol. 5, pp. 488-498, 1962.
15. "The Equivalence of Some Necessary Conditions for Optimal Control Problems with Bounded State Variables," Berkovitz, L. D., and Dreyfus, S. E., Jour. of Math. Analysis and Applications, Vol. 10, pp. 275-283, 1965.
16. "Variational Methods in Problems of Control and Programming," Berkovitz, L. D., Jour. of Math. Anal. and Appl., Vol. 3, No. 1, pp. 145-169, 1961.
17. "On Certain Questions in the Theory of Optimal Control," Filippov, A. F., SIAM Jour. on Control, Vol. 1, pp. 76-84, 1963.

3

OPTIMAL CONTROL OF THE DISTURBED LINEAR REGULATOR

An optimal control problem for a linear regulator with constant, external disturbance is formulated. It is shown that, for a suitably selected quadratic-type performance index, the optimal control is not an explicit function of the external disturbance. Moreover, the optimal control can be synthesized as a time-invariant linear function of the state plus the first time-integral of a certain other time-invariant linear function of the state.

3.1 THE OPTIMAL REGULATOR PROBLEM

The optimal regulator problem for linear dynamical systems can be roughly stated as follows. In the class of piecewise continuous functions, find a vector control $u(t)$ which minimizes the functional¹

$$J [u] = \frac{1}{2} \int_0^T [\langle x(t), Qx(t) \rangle + \langle u(t), Ru(t) \rangle] dt \quad (1)$$

subject to the restrictions

$$\dot{x} = Ax + Fu(t) \quad (\dot{} = d/dt) \quad (2)$$

$$x(0) = x_0, \quad \|x_0\| < \infty \quad (3)$$

$$x(T) = 0 \quad (4)$$

$$u(t) \in U, \quad 0 \leq t \leq T \quad (5)$$

where x is an n -vector, the system state vector; u is an r -vector; Q and R are, respectively, $n \times n$ and $r \times r$ non-negative definite symmetric matrices; A and F are $n \times n$ and $n \times r$ matrices, respectively; and U is a convex subset of the r -dimensional euclidean space.

¹ $\langle x, y \rangle$ denotes the inner product of x and y .

From the point of view of design rational, the term $\langle x, Qx \rangle$ in (1) is chosen to penalize deviations of the regulated state $x(t)$ from the desired equilibrium condition $x(t) \equiv 0$ whereas the term $\langle u, Ru \rangle$ discourages the use of excessively large control effort. One version of this problem, the special case $U = E^r$, was rigorously solved in the well-known 1960 paper by Kalman [1]. Since that time a variety of other special cases of the optimal regulator problem (1) - (5) have been studied by other investigators [2] - [12].

The optimal regulator problem (1) - (5) has enjoyed notable popularity among practical control engineers primarily because, when $U = E^r$, the state variable feedback solution [ie., the control law $u^\circ(x(t), t)$] turns out to be a linear function of the state of the form

$$u^\circ(x(t), t) = K(t) x(t) \quad (6)$$

where $K(t)$ is an rxn matrix which can be effectively computed. Moreover, the reported results of practical experience seem to suggest that, if the matrices Q and R are chosen properly, the resulting "optimal" system does possess many of the same qualitative and quantitative features that are considered "good" by more conventional (classical) control system design procedures [13] - [17].

On the other hand, the ordinary optimal regulator problem, as posed above, suffers one shortcoming which makes it inapplicable in a number of practical applications--it can only accommodate "initial-condition" (or equivalently, impulse-type) disturbances. In particular, if the linear plant (2) is actually subject to both initial condition and finite input disturbances, the optimal control law for the problem (1) - (5) cannot attain and maintain the equilibrium condition $x(t) \equiv 0$, in general. For example, suppose the plant equations (2) actually have the form

$$\dot{x} = Ax + Fu(t) + Bw(t) \quad (7)$$

where B is an $n \times p$ matrix and $w(t)$ is a p -dimensional disturbance vector. Suppose further that the disturbance $w(t)$ eventually becomes, or approaches, a steady state (constant) vector, say $w(t) \equiv c \neq 0$. Then, in the presence of such a disturbance, it is clear from (7) that the linear control law (6) [computed by ignoring input disturbances] cannot satisfy the condition $x(t) \equiv 0$ for positive intervals of time, in general. That is, with the control law (6), it is only possible to "hit" the point $x = 0$ at one or more isolated moments of time [18]. This behavior may be entirely unacceptable in those regulator applications where the state $x(t)$ must be constantly maintained close to zero even in the presence of a piecewise constant disturbance $w(t)$.

It is of some interest, therefore, to consider the possibility of re-formulating the optimal regulator problem (1) - (5) in such a way that, as $t \rightarrow T$, the resulting optimal feedback control always brings the state $x(t)$ and the velocity $\dot{x}(t)$, to zero (equilibrium) in the presence of any finite, constant disturbance $w(t) = c$. Several schemes have previously been proposed for accomplishing this goal. One scheme, which in principle is applicable even for non-constant disturbances, is based on the assumption that the disturbance function $w(t)$ is completely known a priori [19], [20].² In such a case, a time-varying "bias" control vector can be computed in advance and added to the linear control law (6) to effectively cancel out steady state errors due to input disturbances. This scheme is usually impractical because the future behavior of the disturbance $w(t)$ is ordinarily not known a priori.

Another scheme which has been proposed consists of treating $w(t)$ as a non-deterministic input disturbance, with a known probability distribution, and using the theory of optimal stochastic control [21]. This method is also impractical because reliable a priori information about disturbance probability distributions is usually not available.

The practical facts of the matter are that in most regulator applications: (i) the future behavior of input disturbances is ordinarily completely unknown a priori and (ii) the instantaneous properties of input disturbances [eg., magnitude, direction, rate of change, etc.] are usually not directly accessible for measurement. Thus, a realistic scheme for optimal control of regulators with disturbances should yield a feedback control policy which requires neither instantaneous nor future information about the disturbance. This is not asking for too much. Consider, for example, the problem of driving an automobile on a highway when the wind is blowing from the side in a hard and gusty manner. In the presence of such disturbances a good driver can manipulate the steering wheel in such a way as to keep his automobile moving steadily and in close proximity to the desired direction without ever knowing the instantaneous or future values of the wind magnitude and direction. If human controllers can learn to perform with that degree of effectiveness under such conditions of uncertainty (with virtually no quantitative knowledge of the system's dynamical equations of motion), it seems plausible that the powerful analytical techniques of optimal control theory, (using relatively accurate equations of motion), should be capable of yielding mathematically optimal and physically realizable control policies which perform at least as well.

In this paper, a particular class of optimal regulator problems for linear dynamical systems with constant input disturbance is formulated and it is shown that, for the proposed performance index, the optimal control is explicitly independent of the disturbance. It is further shown that the optimal control can be expressed as the sum of a linear function of the state and the first time integral of a certain other linear function of the state. Two examples are worked in detail to illustrate application of the proposed method.

²In [20], the disturbance $w(t)$ is assumed to be added to the system output rather than applied to the input.

3.2 THE SPECIFIC PROBLEM

In the present work, only the case of time-invariant linear dynamical systems with scalar control and scalar disturbance is considered. However, the techniques used can also be applied, in principle, to time-varying deterministic systems and systems with vector control and disturbance.

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

$$J [u] = \frac{1}{2} \int_0^T [\langle x(t), Qx(t) \rangle + r^2 \dot{u}^2(t)] dt \quad (8)$$

subject to the restrictions

$$\dot{x} = Ax + u(t)f + w(t)b \quad (\dot{} = d/dt) \quad (9)$$

$$x(0) = x_0, \quad \| x_0 \| < \infty \quad (10)$$

$$\lim_{t \rightarrow T} x(t) = \lim_{t \rightarrow T} \dot{x}(t) = 0, \quad T\text{-unrestricted} \quad (11)$$

$$u(0) = u_0 \quad (12)$$

$$w(t) \equiv c = \text{scalar constant}, \quad t \geq 0, \quad |c| < \infty \quad (13)$$

where x is an n -vector, Q is an $n \times n$ non-negative definite constant matrix, r is a positive scalar constant, A is an $n \times n$ constant matrix and f and b are constant non-zero n -vectors.

From the design criteria point of view, the performance index (8) differs from the ordinary quadratic functional (1) in that large values of control are discouraged indirectly by penalizing the rate of change of control rather than the control itself. In addition, the initial state $x(0)$ of the system (just before application of the disturbance) is allowed to be zero (the desired operating condition). This provision seems to reflect more accurately³ the actual situation in practical applications. The initial condition u_0 on the control $u(t)$ is assumed to be specified

³It is recalled that in the usual formulation of the optimal regulator problem (1) - (5), the interval of control $[0, T]$ is assumed to start after the disturbance has subsided but before the perturbed state $x(t)$ has returned to the desired operating point.

(chosen) a priori. One rationale for choosing this initial value is discussed in a later section.

It is clear from (9) that the equilibrium condition (11) can be satisfied only if there exists an admissible control $u = u(c)$ such that

$$u(c) f = -cb \quad (14)$$

for every admissible value of c . Thus, for $c \neq 0$, it is necessary to assume that the vectors f and b are collinear. That is

$$b = \alpha f \quad (15)$$

for some non-zero scalar constant α .

It is further necessary to assume (see [1], [20]) that the pair (A, f) is completely controllable. That is, the vectors

$$f, Af, A^2f, \dots, A^{n-1}f \quad (16)$$

are linearly independent. Under these two assumptions, it can be shown, [3], that there is no loss of generality⁴ in assuming that the triple (A, f, b) has the canonical (phase-variable) form

$$A = A_0 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 & 0 \\ 0 & 0 & 1 & \dots & 0 & 0 \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ \cdot & & & & & \cdot \\ 0 & 0 & 0 & \dots & 0 & 1 \\ a_1 & a_2 & a_3 & \dots & a_{n-1} & a_n \end{bmatrix}; f = f_0 = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix}; b = \alpha f_0 \quad (17)$$

The identification of generally applicable necessary and sufficient conditions

⁴That is, if the pair (A, f) is completely controllable, it is always possible to find a nonsingular linear transformation $x = Ky$ such that $K^{-1}AK = A_0$ and $K^{-1}f = f_0$. Some algorithms for constructing the matrix K are described in [3], [22], [23]. A numerical program for implementing one of those algorithms is described in Chapter 4 of this report.

for the existence of optimal controls is still an unsolved⁵ problem--even for the ordinary regulator problem (1) - (5).⁶ We will, therefore, elude this question by simply assuming that, for each pair (x_0, c) an optimal control exists and is unique.

3.3 FORM OF THE SOLUTION

Using the scalar constants defined by (13), (15), an auxiliary state variable is introduced as

$$x_{n+1}(t) = u(t) + c\alpha \quad (18)$$

$$x_{n+1}(0) = u_0 + c\alpha \quad (19)$$

and the additional differential equation

$$\dot{x}_{n+1} = v(t) \quad (20)$$

is appended to (9). Incorporating (13), (18), and (20) into (9) and setting $\tilde{x} = (x_1, \dots, x_n, x_{n+1})$ it is found that $\tilde{x}(t)$ obeys the equation

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} + v(t)\tilde{f} + \tilde{0}b \quad (21)$$

where

$$\tilde{A} = \begin{bmatrix} A_0 & | & f_0 \\ \hline 0 & | & 0 \end{bmatrix}; \quad \tilde{f} = \begin{bmatrix} 0 \\ \hline f_0 \end{bmatrix}; \quad \tilde{b} = \begin{bmatrix} \alpha f_0 \\ \hline 0 \end{bmatrix} \quad (22)$$

⁵The problem of existence of optimal controls, for a very general class of problems has been studied by Markus and Lee [24] and by Bridgland [25]. Sufficient conditions for the existence of solutions to the optimal regulator problem (1) - (5) have been given in [1], [26], [27], and [28].

⁶One necessary condition for the existence of a solution to the problem (1) - (5) is that the non-negative definite quadratic form $\langle x, Qx \rangle$ does not vanish identically along a periodic solution of $\dot{x} = Ax$. This vanishing condition, for the special case $Q = qq^T$, (q n -vector), is discussed in [29] (see also [6]).

The original problem (8) - (13) can now be stated in the following alternative form: Find a piecewise continuous scalar function $v(t)$ which minimizes

$$J[v] = \frac{1}{2} \int_0^T [\langle \tilde{x}(t), \tilde{Q}\tilde{x}(t) \rangle + r^2 v^2(t)] dt \quad (23)$$

subject to the restrictions

$$\dot{\tilde{x}} = \tilde{A}\tilde{x} + v(t)\tilde{f} \quad (24)$$

$$\tilde{x}(0) = (x_0, u_0 + c\alpha)' \quad (' \text{ denotes transpose}) \quad (25)$$

$$\tilde{x}(T) = 0 \quad T\text{-unrestricted} \quad (26)$$

where \tilde{A} , \tilde{f} are given by (22) and \tilde{Q} is an $(n+1) \times (n+1)$ non-negative definite matrix obtained by adding an additional row and column of zero elements to the matrix Q in (8)

$$\tilde{Q} = \begin{bmatrix} Q & | & 0 \\ \hline & & \\ 0 & | & 0 \end{bmatrix} \quad (27)$$

The alternative problem (23) - (26) is recognized as the scalar control case of the ordinary undisturbed, unbounded control, optimal regulator problem (1) - (5). The solution to this latter problem is given, in the control law form $v^0(\tilde{x})$, by the well-known [20] expression

$$v^0(\tilde{x}) = \langle \gamma, \tilde{x} \rangle \quad (28)$$

where

$$\gamma = -r^{-2} M\tilde{f} \quad (29)$$

and M is the unique, constant, $(n+1) \times (n+1)$ positive definite,⁷ symmetric matrix satisfying

$$\tilde{A}'M + M\tilde{A} - r^{-2} M\tilde{f}\tilde{f}'M + \tilde{Q} = 0 \quad (30)$$

⁷The matrix M is positive definite if $\langle \tilde{x}(t), \tilde{Q}\tilde{x}(t) \rangle$ does not vanish identically on a non-trivial solution of $\dot{\tilde{x}} = \tilde{A}\tilde{x}$. [See footnote 6].

From (18), (20), the optimal control $u^\circ(t)$ for the disturbed regulator problem (8) - (13) is found to be

$$u^\circ(t) = x_{n+1}(0) + \int_0^t v^\circ(\tilde{x}(\tau)) d\tau - c\alpha \quad (31)$$

or

$$u^\circ(t) = \sum_1^{n+1} \gamma_i \int_0^t x_i(\tau) d\tau \quad (32)$$

However, from (24)

$$x_j = \dot{x}_{j-1} \quad j = 2, 3, \dots, n \quad (33)$$

and

$$x_{n+1} = \dot{x}_n - \sum_1^n a_i x_i$$

so that (32) can be expressed solely in terms of the state variables $x_1(t), \dots, x_n(t)$ as follows

$$u^\circ(t) = \sum_1^n \beta_i x_i(t) + \beta_0 \int_0^t x_1(\tau) d\tau ; u^\circ(0) = u_0 \quad (34)$$

where the scalar coefficients $\beta_0, \beta_1, \dots, \beta_n$ are independent of c and are defined by

$$\beta_i = \gamma_{i+1} - \gamma_{n+1} a_{i+1} \quad i = 0, 1, 2, \dots, n-1 \quad (35)$$

$$\beta_n = \gamma_{n+1}$$

and the initial condition on the integral term in (34) is chosen to satisfy $u^\circ(0) = u_0$.

Thus, the optimal control for the system (9), with performance index (8), results in a constant coefficient $(n+1)$ - order, linear dynamical system

$$\dot{x} = Ax + \left(\sum_1^n \beta_i x_i + \beta_0 \int_0^t x_1(\tau) d\tau \right) f + cb \quad (36)$$

which is globally asymptotically stable with respect to the equilibrium state $x = 0$ for every constant $|c| < \infty$. It should be noted that the terminal condition (11) is satisfied only as $t \rightarrow \infty$.

It is recalled [see (17) and footnote 4] that for mathematical convenience the state vector x in (9) was assumed to be in (or have previously been linearly transformed to) phase-variable form: $\dot{x}_i = x_{i+1}$, $i=1 \dots, n-1$. For this reason, each phase-variable coordinate x_1, \dots, x_n appearing on the right of (34) actually represents a linear combination of the original (physical) state variables for the problem, in general. Therefore, in terms of more general (non phase-variable) state variables y_1, \dots, y_n the optimal control (34) can be expressed as the sum of: (i) a linear, constant coefficient, combination of the state variables and (ii) the first time integral of a certain other linear, constant coefficient, combination⁸ of the state variables.

$$u^{\circ}(t) = \sum_1^n \rho_i y_i(t) + \beta_0 \int_0^t \left(\sum_1^n \eta_i y_i(\tau) \right) d\tau, \quad u^{\circ}(0) = u_0 \quad (37)$$

$$\rho_i, \eta_i = \text{constant } i = 1, \dots, n$$

3.4 DISCUSSION OF RESULTS

It was assumed in (12) that the initial value $u^{\circ}(0) = u_0$ was specified (fixed) a priori. In this case the control (34) is optimal, with respect to the functional (8), for every pair (x_0, c) . Although it is the designer's privilege to arbitrarily choose the value u_0 , there is one rather natural procedure for selecting this

⁸This special linear combination of state variables $\sum_1^n \eta_i y_i$, corresponding to the particular coordinate x_1 in the canonical phase-variable coordinate system, plays an important role in many optimal control problems. If H is the matrix whose columns are $f, Af, \dots, A^{n-1}f$ [the so-called controllability matrix] then, it can be shown [23] that, in general

$$x_1 = \sum_1^n \eta_i y_i = \langle \bar{h}_n, y \rangle \quad (38)$$

where \bar{h}_n is the n th row of H^{-1} .

parameter. If one attempts to solve for the theoretically optimum choice of u_0 it is found⁹ that u_0 depends explicitly on (x_0, c) . However, since $w(t)$ is assumed to be not accessible for measurement, this choice for u_0 is not physically realizable. On the other hand, suppose that, for $t < 0$, the disturbance $w(t)$ is constant, say $w(t) \equiv c_0 \neq c$. In this case, if one agrees to choose u_0 as the particular value which maintains equilibrium $x(t) \equiv 0$, for all $t < 0$, the control (34) always naturally approaches (asymptotically) this desired "initial value" [$u_0 = -c_0 \alpha$] because

$$\lim_{(t \rightarrow \infty)} u^\circ(t) = -c\alpha \quad (39)$$

along every solution of the optimally controlled system (9) for arbitrary initial values of $u^\circ(t)$. Thus, in practical applications where the interval of control is actually finite (and the disturbance is more nearly "piecewise constant") one can effectively disregard the explicit "setting" of u_0 and simply let $u_0(t)$, $t < 0$, seek its natural equilibrium value (39).

The linear control law (6) has often been described as the modern optimal control version of the classical "proportional feedback" control principle [20]. In a like manner, the linear functional optimal control (34) can be viewed as a modern version of the classical "proportional plus integral" feedback control principle. This latter principle, sometimes called the "follow-up" control principle, is an old and well-known technique for reducing or eliminating "offset-errors" in regulator-type controllers.

3.5 EXTENSION OF RESULTS

The technique described above can be extended to a more general class of input disturbances. In particular, if the disturbance $w(t)$ is an m th degree polynomial in t , $m \geq 0$, and if the performance index (8) has the form

$$J[u] = \frac{1}{2} \int_0^T [\langle x(t), Qx(t) \rangle + r^2 (d^{m+1} u(t) / dt^{m+1})] dt ; m \geq 0 \quad (40)$$

⁹The value of u_0 which minimizes (8), for fixed (x_0, c) is given by

$$u_0 = x_{n+1}^* - c\alpha$$

where

$$x_{n+1}^* = \arg \min_{x_{n+1}} \langle \tilde{x}_0, M \tilde{x}_0 \rangle, \tilde{x}_0 = (x_0, x_{n+1})$$

and M is the positive definite, symmetric, solution of (30).

then, following the method of solution outlined above, the optimal control $u^{\circ}(t)$ is obtained as the sum of (i) a linear combination of the state variables $x_i(t)$, $i = 1, \dots, n$, and (ii) the $(m+1)$ th. time integral of a certain other linear combination of the state variables. Moreover, for fixed $u^{\circ}(0)$, this latter control is invariantly optimal in the sense that it is simultaneously optimal for all polynomial input disturbances $w(t)$ of degree $0 \leq l \leq m$.

3.6 EXAMPLES

Example 1 - A First Order System

As a special case of (8) - (13), let

$$J[u] = \frac{1}{2} \int_0^T [qx_1^2(t) + r^2 \dot{u}^2(t)] dt, \quad q > 0, r > 0 \quad (41)$$

and

$$\dot{x}_1 = u(t) + \alpha w(t), \quad w(t) \equiv c \quad (42)$$

with

$$x_1(0) = x_0$$

$$x_1(T) = 0 \quad T\text{-unrestricted} \quad (43)$$

$$u(0) = u_0 \quad (44)$$

Preceding as in (18), define the auxiliary state variable

$$x_2(t) = u(t) + \alpha c$$

$$x_2(0) = u_0 + \alpha c \quad (45)$$

and set

$$\dot{x}_2 = \dot{u}(t) = v(t) \quad (46)$$

Equations (41), (42) can now be written as

$$J[v] = \frac{1}{2} \int_0^T [qx_1^2(t) + r^2 v^2(t)] dt \quad (47)$$

$$\dot{x}_1 = x_2 \quad (48)$$

$$\dot{x}_2 = v(t) \quad (49)$$

The unique, positive definite, symmetric solution M of (30), with

$$\tilde{A} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}; \quad \tilde{f} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}; \quad \tilde{Q} = \begin{bmatrix} q & 0 \\ 0 & 0 \end{bmatrix} \quad (50)$$

is readily found to be

$$M = \begin{bmatrix} \left(2rq\right)^{\frac{1}{2}} & r(q)^{\frac{1}{2}} \\ r(q)^{\frac{1}{2}} & \left(2r^3q\right)^{\frac{1}{2}} \end{bmatrix} \quad (51)$$

where $(\cdot)^k$, $k > 0$, denotes the positive k^{th} root of (\cdot) . The expression for $v^\circ(x)$ is then given by

$$v^\circ(\tilde{x}) = -r^{-2} \langle M\tilde{f}, \tilde{x} \rangle, \quad \tilde{x} = (x_1, x_2) \quad (52)$$

or

$$v^\circ(\tilde{x}) = -r^{-1}(q)^{\frac{1}{2}} x_1 - \left(2r^{-1}\right)^{\frac{1}{2}} (q)^{\frac{1}{4}} x_2 \quad (53)$$

Finally, the optimal control $u^\circ(t)$ is obtained from (34) as

$$u^\circ(t) = \gamma_1 \int_0^t x_1(\tau) d\tau + \gamma_2 x_1(t) + u_0 \quad (54)$$

where

$$\begin{aligned} \gamma_1 &= -r^{-1}(q)^{\frac{1}{2}} < 0 \\ \gamma_2 &= -\left(2r^{-1}\right)^{\frac{1}{2}} (q)^{\frac{1}{4}} < 0 \end{aligned} \quad (55)$$

The optimally controlled plant (42) is therefore given by

$$\dot{x}_1 = \gamma_1 \int_0^t x_1(\tau) d\tau + \gamma_2 x_1 + \alpha w(t) + u_0, \quad w(t) \equiv c \quad (56)$$

which is asymptotically stable, with respect to $x_1 = 0$, for all finite c .

The control (54) is optimal for every fixed value of u_0 . A practical scheme for choosing u_0 was described above [see eg. (39)]. For comparison purposes the theoretically optimum choice for u_0 , with (x_0, c) fixed, is found to be [see footnote #9]

$$u_0 = - q^{\frac{1}{4}} (2r)^{-\frac{1}{2}} x_0 - c\alpha$$

which is not physically realizable.

It is interesting to compare the optimal control (54) of the present example with the solution obtained for the following conventional, undisturbed linear regulator problem. Minimize

$$J[u] = \frac{1}{2} \int_0^T [qx_1^2(t) + \rho^2 u^2(t)] dt \quad q > 0, \rho > 0 \quad (57)$$

subject to

$$\dot{x}_1 = u(t) + \alpha w(t), \quad w(t) \equiv 0 \quad (58)$$

$$x_1(0) = x_0$$

$$x_1(T) = 0, \quad T\text{-unrestricted} \quad (59)$$

The optimal control law for this problem is well-known [20] and is given by

$$u^0(x) = -\rho^{-1} (q)^{\frac{1}{2}} x_1 \quad (60)$$

Example 2 - A Second Order Example

As another special case of (8) - (13), let

$$J[u] = \frac{1}{2} \int_0^T [x_1^2(t) + x_2^2(t) + \dot{u}^2(t)] dt \quad (61)$$

and

$$\dot{x}_1 = x_2 \quad (62)$$

$$\dot{x}_2 = u(t) + w(t) , \quad w(t) \equiv c \quad (63)$$

$$x(0) = x_0$$

$$x(T) = 0 , \quad T\text{-unrestricted} \quad (64)$$

$$u(0) = u_0 \quad (65)$$

Proceeding as in the previous example, the auxiliary state variable $x_3(t)$ is defined as

$$x_3(t) = u(t) + c \quad (66)$$

$$x_3(0) = u_0 + c$$

so that, setting $v(t) = \dot{u}(t)$, (61) - (63) can be written as the equivalent third order system

$$J[v] = \frac{1}{2} \int_0^T [x_1^2(t) + x_2^2(t) + v^2(t)] dt \quad (67)$$

$$\dot{x}_1 = x_2 \quad (68)$$

$$\dot{x}_2 = x_3 \quad (69)$$

$$\dot{x}_3 = v(t) \quad (70)$$

From (28) - (35) the optimal control is found to be

$$u^o(t) = \gamma_1 \int_0^t x_1(\tau) d\tau + \gamma_2 x_1 + \gamma_3 x_2 + u_0 \quad (71)$$

where the γ_i , $i = 1, 2, 3$, are given (approximately) by

$$\begin{aligned}\gamma_1 &= -1.00 \\ \gamma_2 &= -2.31 \\ \gamma_3 &= -2.15\end{aligned}\tag{72}$$

For comparison purposes, the optimal control law for the conventional, undisturbed linear regulator problem with

$$J[u] = \frac{1}{2} \int_0^T [x_1^2(t) + x_2^2(t) + u^2(t)] dt\tag{73}$$

and

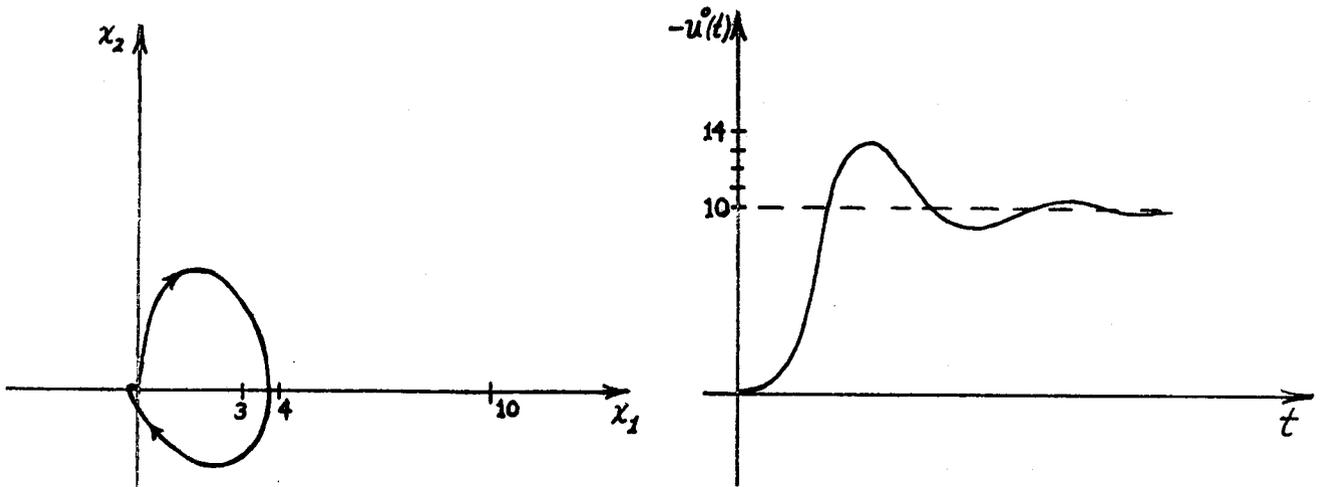
$$\dot{x}_1 = x_2\tag{74}$$

$$\dot{x}_2 = u(t) + w(t), \quad w(t) \equiv 0\tag{75}$$

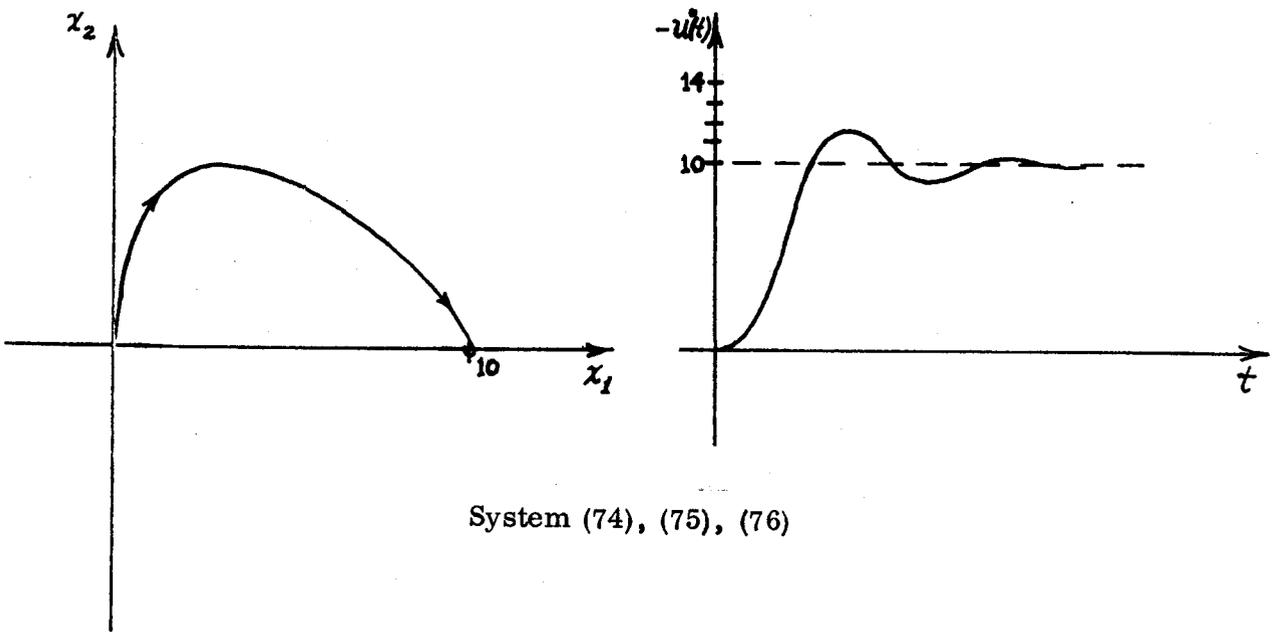
is found to be [6]

$$u^o(x) = -x_1 - 1.73 x_2\tag{76}$$

The responses of the two optimally controlled systems (62), (63), (71) and (74) - (76), for a value of $w(t) \equiv 10$, are shown in Figure 1. It is observed that the two control functions $u^o(t)$, given by (71) and (76), are quite similar although the latter, of course, does not satisfy the specified boundary condition $x(T) = 0$.



System (62), (63), (71)



System (74), (75), (76)

Figure 1. Comparison of Responses for the System (62), (63), (71) and the System (74), (75), (76).

REFERENCES

1. "Contributions to the Theory of Optimal Control," Kalman, R. E., Bol. Soc. Mat. Mexicana, Second Ser. Vol. 5, pp. 102-119, 1960.
2. "Analytical Controller Design: Parts 1, 2, 3," Letov, A. M., Avtomatika i Telemekhanika, Vol. 21, nos. 4, 5, 6, pp. 436-441, 561-568, 661-665, April, May, June 1960.
3. "Optimal Bang-Bang Control with Quadratic Performance Index," Wonham, W. M. and Johnson, C. D., Proceedings, Fourth Joint Automatic Control Conference, Minneapolis, Minn., 1963. Also; ASME Trans., Jour. of Basic Engineering, Series D, Vol. 86, pp. 107-115, March 1964.
4. "Optimal Control with Quadratic Performance Index and Fixed Terminal Time," Johnson, C. D. and Gibson, J. E., IEEE Trans. on Automatic Control, Vol. AC-9, no. 4, October 1964.
5. "An Alternate Approach to the Fixed Terminal Point Regulator Problem," Rekasius, Z. V., 9th Annual East Coast Conference of IRE and PGANE, Baltimore, Md., October 24, 1962. Also; IEEE Trans. on Automatic Control, Vol. AC-9, no. 3, July 1964.
6. "On a Problem of Letov in Optimal Control," Johnson, C. D. and Wonham, W. M., Proc. Fifth Joint Automatic Control Conference, Stanford, California, June 1964. Also; ASME Trans., Jour. of Basic Engineering, Vol. 87, Series D, no. 1, pp. 81-89, March 1965.
7. "Optimal Control of Variable Structure Systems in the Presence of Constant Disturbance," Buyakas, V. I., Avtomatika i Telemekhanika, Vol. 27, no. 7, pp. 92-101, July 1966.
8. "Optimal Control of Variable Structure Systems," Buyakas, V. I., Avtomatika i Telemekhanika, Vol. 27, no. 4, pp. 57-68, April 1966.
9. "Synthesis of Optimal Control for a System of Second Order," Kozirov, L. M. and Kupervasser, Yu. I., Tekhnicheskaya Kibernetika, Vol. , no. 4, pp. 154-162, 1965.

10. "Control Algorithm Design of Variable Optimality Criterion," Kuznetsov, N. A., *Avtomatika i Telemekhanika*, Vol. 27, no. 5, pp. 5-14, May 1966.
11. "High Order System Design Via State-Space Considerations," Bass, R. W. and Gura, I., *Proceedings, Sixth Joint Automatic Control Conference*, Troy, New York, June 1965.
12. "On Synthesis of Optimal Bang-Bang Feedback Control Systems with Quadratic Performance Index," Bass, R. W. and Webber, R. F., *Proceedings, Sixth Joint Control Conference*, Troy, New York, June 1965.
13. "An Application of the Quadratic Penalty Function Criterion to the Determination of a Linear Control for a Flexible Vehicle," E. E. Fisher, *Jour. AIAA*, Vol. 3, no. 7, pp. 1262-1267, July 1965.
14. "The Multivariable Root Square Locus - An Optimal System Design Aid," Rynaski, E. G., *AIAA Conf. on Guidance and Control*, Minneapolis, Minn., August 1965.
15. "Optimal Control: A Review of Theory and Practice," Paiewonsky, B., *Jour. AIAA*, Vol. 3, no. 11, pp. 1985-2006, November 1965.
16. "Optimal Helicopter Stationkeeping," Rynaski, E. G., *IEEE Trans. on Automatic Control*, Vol. AC-11, no. 3, pp. 346-355, July 1966.
17. "On the Optimal Error Regulation of a String of Moving Vehicles," Levine, W. S. and Athans, M., *IEEE Trans. on Automatic Control*, Vol. AC-11, no. 3, pp. 355-361, July 1966.
18. "The Optimal Problem of a 'Hit' in Transient Linear Systems," Tkhan'bank, Nguyen, *Tekhnicheskaya Kibernetika*, Vol. , no. 1, Jan.-Feb. 1965.
19. "The Theory of Optimal Control and the Calculus of Variations," Kalman, R. E., Chapter 16 of *Mathematical Optimazation Techniques* (book), Univ. of California Press, 1963.
20. "Optimal Control; An Introduction to the Theory and Its Applications," Athans, M. and Falb, P., (book) McGraw-Hill, New York, 1966.
21. "On the Status of Optimal Control and Stability for Stochastic Systems," Kushner, H. J., *IEEE International Conv. Record*, Part 6, pp. 143-151, March 1966.
22. "A Note on the Transformation to Canonical (Phase-Variable) Form," Johnson, C. D. and Wonham, W. M., *IEEE Trans. on Automatic Control*, Vol. AC-9, no. 3, pp. 312-313, July 1964.

23. "Another Note on the Transformation to Canonical (Phase-Variable) Form," Johnson, C. D. and Wonham, W. M., IEEE Trans. on Automatic Control, Vol. AC-11, no. 3, July 1966.
24. "On the Existence of Optimal Controls," Markus, L. and Lee, E. B., ASME Trans., Jour. of Basic Engineering, Vol. 84, pp. 13-22, March 1962.
25. "On the Existence of Optimal Feedback Controls; Parts 1 and 2," Bridgland, T. F., Jr., SIAM, Journal on Control, Ser. A, Vol. 1, no. 3, pp. 261-274, 1963 and Vol. 2, no. 2, pp. 127-150, 1965.
26. "An Optimal Regulator Problem," Chang, A., SIAM Jour. on Control, Vol. 2, pp. 220-233, 1964.
27. "The Existence of Optimal Controls for a Performance Index with Positive Integrand," Anvari, M. and Datko, R. F., SIAM Jour. on Control, Vol. 4, no. 2, pp. 372-381, 1966.
28. "On a Solution of an Optimization Problem in Linear Systems with Quadratic Performance Index," Sakawa, Y., SIAM Jour. on Control, Vol. 4, no. 2, pp. 382-395, 1966.
29. "Invariant Hyperplanes for Linear Dynamical Systems," Johnson, C. D., I.E.E.E. Trans. on Automatic Control, Vol. AC-11, pp. 113-116, January 1966.

4

A NUMERICAL ALGORITHM FOR COMPUTING THE LINEAR TRANSFORMATION $x = Ky$ WHICH TRANSFORMS AN ARBITRARY, COMPLETELY CONTROLLABLE, LINEAR DYNAMICAL SYSTEM $\dot{x} = Ax + u(t)f$ INTO THE CANONICAL (PHASE-VARIABLE) FORM $\dot{y} = A_0y + u(t)f_0$

The algorithm described below generates a nonsingular matrix K (and its inverse K^{-1}) with the following property: Given any completely controllable¹ pair (A, f) where A is a real $n \times n$ matrix and f is a real n -vector, the associated pair (A_0, f_0) defined by

$$A_0 = K^{-1}AK$$

$$f_0 = K^{-1}f$$

has the canonical (phase-variable) form

$$A_0 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & & & & 1 \\ \alpha_1 & \alpha_2 & \alpha_3 & & \alpha_n \end{bmatrix}; f_0 = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \\ 1 \end{bmatrix}$$

This algorithm, which is based on a result originally described in [1] (see also [2]-[5]), accepts as inputs the elements of the pair (A, f) and generates, as outputs, the elements of: K, K^{-1}, A_0, f_0 .

¹A pair (A, f) is said to be completely controllable [and the nonsingular matrix K exists] if and only if the sequence of vectors $f, Af, A^2f, A^3f, \dots, A^{n-1}f$, are linearly independent.

This routine calls for one matrix inversion operation (4.2, Step 2) which must be carried out by means of an external "matrix inversion sub-routine". All other operations are straight forward multiplications of scalars and matrices.

4.1 INPUT DATA

- a. Then n^2 elements a_{ij} ($i, j = 1, 2, \dots, n$) of the $n \times n$ matrix A are input and stored.
- b. The n elements f_i ($i = 1, 2, \dots, n$) of the n -dimensional vector f are input and stored.

4.2 ALGORITHM FOR GENERATING K AND K^{-1}

- Step 1. Compute and store the n^2 elements h_{ij} ($i, j = 1, 2, \dots, n$) of the controllability matrix H by using the recursive relation:

$$\begin{aligned}
 h_{i1} &= f_i & i &= 1, 2, \dots, n \\
 h_{ij} &= \sum_{s=1}^n a_{is} h_{s, j-1} & i &= 1, 2, \dots, n \\
 & & j &= 2, 3, \dots, n
 \end{aligned}$$

Also, compute and store the n elements b_i ($i = 1, 2, \dots, n$) where:

$$b_i = \sum_{s=1}^n a_{is} h_{sn}$$

- Step 2. Compute and store the n^2 elements of H^{-1} (the inverse of H). Let the elements of H^{-1} be denoted by \bar{h}_{ij} ($i, j = 1, 2, \dots, n$)

(NOTE: This matrix inversion operation requires an externally supplied "matrix inversion sub-routine".)

- Step 3. If H^{-1} does not exist, stop computation and print: THE PAIR (A, f) IS NOT COMPLETELY CONTROLLABLE. Otherwise, go on to step 4.

Step 4. Compute and store the n scalars (numbers) α_i ($i = 1, 2, \dots, n$) where

$$\alpha_i = \sum_{j=1}^n \bar{h}_{ij} b_j \quad i = 1, 2, \dots, n$$

(NOTE: The elements \bar{h}_{ij} and b_j are called from storage.)

Step 5. Compute and store the n^2 elements k_{ij} ($i, j = 1, 2, \dots, n$) of K by the following (backward) recursive rule:

First compute

$$k_{in} = f_i \quad , \quad i = 1, 2, \dots, n$$

then compute (for descending values of j)

$$k_{ij} = -\alpha_{j+1} f_i + \sum_{s=1}^n a_{is} k_{s,j+1} \quad \begin{array}{l} i = 1, 2, \dots, n \\ j = (n-1), (n-2), \dots, 2, 1 \end{array}$$

(NOTE: The elements α_i and a_{ij} are called storage.)

Step 6. Compute and store the n elements β_i ($i = 1, 2, \dots, n$) by the following (backward) recursive rule:

First compute

$$\beta_n = \alpha_n$$

then compute (for descending values of i)

$$\beta_i = \alpha_i + \sum_{s=i+1}^n \alpha_{n+i+1-s} \beta_s \quad , \quad i = (n-1), (n-2), \dots, 3, 2.$$

Step 7. Compute and store the elements m_{ij} ($i, j = 1, 2, \dots, n$) of the matrix M by the following rule:

$$m_{ij} = \left\{ \begin{array}{ll} 0 & \text{if } (i+j) \leq n \\ +1 & \text{if } (i+j) = n+1 \\ \beta_{2n+2-i-j} & \text{if } (i+j) \geq n+2 \end{array} \right.$$

(NOTE: The elements β_i are called from storage.)

Step 8. Compute and store the n^2 elements of the $n \times n$ matrix K^{-1} where K^{-1} is computed by forming the following matrix product:

$$K^{-1} = MH^{-1}$$

(NOTE: The elements of M and H^{-1} are called from storage.)

Step 9. Compute and store the n^2 elements of the $n \times n$ matrix A_0 where A_0 is computed by forming the following triple matrix product:

$$A_0 = K^{-1}AK$$

(NOTE: The elements of K^{-1} , A , and K are called from storage.)

Step 10. Compute and store the n elements of the n -vector f_0 where

$$f_0 = K^{-1}f$$

(NOTE: The elements of K^{-1} and f are called from storage.)

Step 11. PRINT OUT THE FOLLOWING MATRICES (AND VECTORS) FROM THE STORED DATA:

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$$

$$f = \begin{bmatrix} f_1 \\ f_2 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{bmatrix}$$

$$K = \begin{bmatrix} k_{11} & k_{12} & \dots & k_{1n} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ k_{n1} & k_{n2} & \dots & k_{nn} \end{bmatrix}$$

$$K^{-1} = \begin{bmatrix} \bar{k}_{11} & \bar{k}_{12} & \dots & \bar{k}_{1n} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \bar{k}_{n1} & \bar{k}_{n2} & \dots & \bar{k}_{nn} \end{bmatrix}$$

$$A_0 = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & 0 & \dots & & 1 \\ \alpha_1 & \alpha_2 & \dots & & \alpha_n \end{bmatrix}$$

$$f_0 = \begin{bmatrix} 0 \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 1 \end{bmatrix}$$

ALSO, PRINT OUT THE ELEMENTS:

$\alpha_1, \alpha_2, \alpha_3, \dots, \alpha_n$

$b_1, b_2, b_3, \dots, b_n$

REFERENCES

1. "Optimal Bang-Bang Control with Quadratic Performance Index," Wonham, W. M. and Johnson, C. D., Proc. 4th JACC, pp. 101-112, 1963. See also, J. Basic Engrg., Trans. ASME, vol. 86, pp. 107-115, March 1964.
2. "A Note on the Transformation to Canonical (Phase-Variable) Form," Johnson, C. D. and Wonham, W. M., IEEE Trans. on Automatic Control (Short Papers), vol. AC-9, pp. 312-313, July 1964.
3. "On the Reduction of a System to Canonical (Phase-Variable) Form," Mufti, I. H., IEEE Trans. on Automatic Control (Correspondence), vol. AC-10, pp. 206-207, April 1965.
4. "Transformation of Time-Variable Systems to Canonical (Phase-Variable) Form," Silverman, L. M., IEEE Trans. on Automatic Control, (Short Paper), vol. AC-11, pp. 300-303, April 1966.
5. "Another Note on the Transformation to Canonical (Phase-Variable) Form," Johnson, C. D. and Wonham, W. M., IEEE Trans. on Automatic Control, vol. AC-11, no. 3, pp. 609-610, July 1966.

5

CONCLUSIONS

The algorithm described in Chapter 2 for computing C-minimax optimal controls should prove useful in obtaining numerical-type descriptions for the boundaries of the R_0^i and R_m^i sets, $i = 1, 2, \dots$. Moreover, in those cases where the optimal control in $\{R_m\}$ is of the bang-bang type, (a situation very likely to occur when the state variable equations are linear in the control u , [2, Ch. 2], the algorithm will permit the numerical determination of points on the bang-bang switching surface. The question of how to effectively use numerical data, obtained from algorithms such as presented here, in the practical synthesis of higher order feedback control systems is a common, and still unresolved, problem in the application of optimal control theory. This important subject should be given further study in the future with particular emphasis on (i) the "fitting" of higher-dimensional "surfaces" to sets of numerical data and (ii) the possibility of reducing the dimensionality of the state space by selectively ignoring certain state variables. This latter topic is closely associated with the more general problem of choosing the most appropriate set of state variables to describe a dynamical system.

The algorithm will also have practical application in determining the relative effectiveness of various sub-optimal C-minimax controls which may be proposed. For this purpose, the exact C-minimax control can be computed, via the algorithm, in backward time, starting at a selected "terminal state" $x(T) \in \mathcal{J}$. The backward time optimal trajectory generated by this means is then stopped at some selected time τ' and $x(\tau')$ is noted. Then, $x(\tau')$ is used as the forward-time initial condition for the same system with the proposed sub-optimal control, and the corresponding forward time sub-optimal trajectory is determined. Comparison of the maximum value of $C(x)$ along each trajectory will then reveal the degree of sub-optimality of the proposed non-optimal control.

The theory developed in Chapter 3 for linear regulators with constant disturbance inputs provides one solution for the problem of eliminating steady-state "offset" error in regulator control systems. Although the theory is developed for strictly constant disturbance inputs it would be interesting to study, experimentally, the performance characteristics of such a system with various non-constant disturbances. In addition, an investigation of the relative effectiveness of such systems, when used as sub-optimal C-minimax controllers, would provide useful information.