OPTIMAL CONTROL COMPUTATIONS FOR LINEAR SAMPLED-DATA SYSTEMS

Thesis for the Degree of Ph. D.
MICHIGAN STATE UNIVERSITY
JENG NAN LIN
1968

This is to certify that the

thesis entitled

OPTIMAL CONTROL COMPUTATIONS
FOR LINEAR SAMPLED-DATA SYSTEMS

presented by

Jeng Nan Lin

has been accepted towards fulfillment
of the requirements for

Ph.D. degree in EE

Robert O. Barr
Gerald J. Park
Major professor

Date November 20, 1968

O-169

113V

ABSTRACT

OPTIMAL CONTROL COMPUTATIONS FOR LINEAR SAMPLED-DATA SYSTEMS

By

Jeng Nan Lin

This thesis presents a new algorithm for computing optimal controls for linear sampled-data systems in which easily-implemented matrix calculations are performed at each step. A comprehensive analysis is given for several classes of optimization problems and numerical data are presented for six example problems.

The thesis considers an n^{th} order, time-invariant, scalar-control, linear discrete system $\mathcal{L}$ which is governed by the following difference equation:

$$x((i+1)T) = \Phi x(iT) + bu(iT), \qquad (1)$$

where Φ is an $(n \times n)$ constant transition matrix, b is an $(n \times 1)$ non-zero vector, $u(t)$ is constant in the interval $iT \leq t < (i+1)T$ and $|u(iT)| \leq 1$ for all i.

Iteration on (1) starting with $i = 0$ and $x(0) = 0$ yields

$$x(NT) = \Phi^{N-1}bu(0)+\Phi^{N-2}bu(T)+\cdots+\Phi bu((N-2)T)+bu((N-1)T). \qquad (2)$$

Let $x_N = x(NT)$, $z_i = \Phi^{i-1}b$, $i = 1,2,\ldots,N$, and $u_j = u((N-j)T)$, $j = 1,2,\ldots,N$. Then (2) becomes

$$x_N = z_N u_N + z_{N-1}u_{N-1} + \cdots\cdots + z_2 u_2 + z_1 u_1. \qquad (3)$$

Jeng Nan Lin

The reachable set from the origin at time N, $R_N \overset{d}{=} \{x_N \in E^n : x_N = \sum_{i=1}^{N} z_i u_i,\ |u_i| \le 1,\ 1 \le i \le N\}$ can be described as a system of linear inequalities: $|c_j' x| \le 1$, where c_j is a computable $(n \times 1)$ vector, $j = 1,2,\ldots,K$, and $K \le \binom{N}{n-1}$. The properties of this set are fundamental in the computing algorithms in this thesis.

Two broad classes of problems are considered:

1. Time-optimal control problem - Given the linear system $\mathcal{L}$ starting from $x(0) = 0$, and given a desired terminal target point d, find the smallest integer number N such that $d = x_N = \sum_{i=1}^{N} z_i u_i$.

   (i) When the optimal time is $N \le n$, the solution to the time-optimal control problem always exists and is shown to be unique.

   (ii) When the optimal time is $N > n$, the solution to the time-optimal control problem is, in general, not unique. In this case, in addition to seeking a time-optimal control sequence, a choice is made to also minimize the absolute value of u_N.

An algorithm for implementing the optimal control sequence is presented, which avoids the corresponding tedious quadratic and linear programming problems. The solution exhibits the special features that x_1 is in R_1 when $u(N-1)$ is applied at the first stage, x_2 is on the boundary of R_2 when $u(N-2)$ is applied at the second stage, ..., etc., and finally x_N is on the boundary of $\overline{R}_N$, a set closely related to R_N, when $u(0)$ is applied at the N^{th} stage.

2. Terminal-error regulator problem - Given $\mathcal{L}$ with $x(0) = 0$, the terminal time N, and a desired terminal state d at time N; find a point $x \in R_N$ such that $\|d-x\|^2$ is minimum, and the corresponding optimal control sequence, where $\|\cdot\|$ denotes the Euclidean norm.

(i) In the case where $N < n$, it is shown that the control sequence always exists and is unique.

(ii) Consider the case where $N \geq n$. If $d \in \partial R_N$, the control sequence is unique; if $d \in R_N$ but $d \notin \partial R_N$, the optimal control sequence is not unique. This then leads to the time-optimal problem, and the minimum of $|u_N|$ can be required to make the time-optimal control sequence unique. If $d \notin R_N$, then $\|d-x\|^2 > 0$ and the corresponding optimal control sequence is unique. A control algorithm which terminates in a finite number of steps and guarantees the exact solution except for round-off errors is demonstrated. This algorithm does not require solving the corresponding quadratic programming problem.

These new methods include matrix calculations which are easily programmed. The above results for the time-optimal and terminal-error regulator problems can easily be extended to apply to:

(i) Control constraints given in the form of $|u_i| \le \eta_i$, $\eta_i > 0$, $i = 1,2,\ldots,N$.

(ii) Target sets given in the form of $\mathcal{S}$: $x'Sx \ge \beta$, where S is an $(n \times n)$ symmetric, positive definite matrix and $\beta > 0$.

(iii) Target sets which are time-varying, i.e., $\mathcal{S} = \mathcal{S}(i)$.

(iv) Time-varying linear discrete systems.

(v) Systems with variable sampling instants.

(vi) Linear discrete systems with the initial state $x(0) \neq 0$.

(vii) Multi-input control systems.

A brief survey of these extensions is presented.

OPTIMAL CONTROL COMPUTATIONS FOR LINEAR SAMPLED-DATA SYSTMES

By

Jeng Nan Lin

A THESIS

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

DOCTOR OF PHILOSOPHY

Department of Electrical Engineering

1968

ACKNOWLEDGEMENTS

The author wishes to express his heartfelt appreciation to Dr. R. O. Barr for his interest and valuable comments concerning this thesis. Dr. Barr's critical review of the manuscript, which resulted in a number of improvements, is gratefully acknowledged. He also wishes to extend his sincere appreciation to Dr. G. L. Park for his continued guidance and inspiration in the preparation of this work. Dr. Park's generous and kind help with the example problems and computer programming is especially extoled.

The author is indebted to Dr. H. E. Koenig who contributed in many ways to the author's graduate study, to Dr. J. S. Frame who carefully proofread the thesis and provided many fruitful comments regarding this thesis, and to Dr. R. D. Dubes who graciously served as guidance committee member.

Finally to his wife, Tzu Hsia, the author is wholeheartedly grateful for her constant encouragement during this study.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

LIST OF SYMBOLS

A	linear transformation
C	reduced echelon matrix
E^n	Euclidean n-space
H	hyperplane in E^n
H_s	support hyperplane to R_N
J	set of integers
K	total number of boundary hyperplanes of R_N
L_{ij}	elementary matrix
L	$= \begin{bmatrix} L_1 \\ \hline L_2 \end{bmatrix}$ product of elementary matrices
N	optimal time for time-optimal control problem; also terminal time for terminal-error regulator problem
P	performance index for terminal-error regulator problem
R_N	reachable set of $\mathcal{L}$ at time N from the origin
T	sampling period in seconds
$U(U_n)$	identity matrix (of dimension $(n \times n)$)
Z	$(z_1, z_2, \ldots, z_N)$, a matrix of dimension $(n \times N)$
b	$(n \times 1)$ driving non-zero vector for single-input system
b_j	$(n \times 1)$ driving non-zero vector for multi-input system

e	extreme point of R_N
$h(Y)$	the convex hull of the set of points Y in E^n
n	dimension of the state space x
$r(A)$	the rank of linear transformation A
t	time
u_i	scalar control for single-input system
u_i^j	scalar control for multi-input system
$\vec{u}_N$	control sequence of length N, $[u_1,u_2,\ldots,u_N]$ for single-input system and $[u_1^1,u_1^2,\ldots,u_1^m, u_2^1,\ldots, u_2^m,\ldots, u_N^1,\ldots, u_N^m]$ for multi-input system
u_N^*	optimal control
v	vertex of R_N
z_i	$(n \times 1)$ vector defined by $z_i = \Phi^{i-1}b$
z_i^j	$(n \times 1)$ vector defined by $z_i^j = \Phi^{i-1}b_j$
$\mathcal{C}(\Lambda)$	completement of the set Λ
Ω_i	the intersection of two closed half-spaces Γ_i^1 and Γ_i^2
Γ_i^j	closed half-space on E^n
$\mathcal{L}$	linear system to be controlled
$\mathcal{S}$	target set
Θ_N^Y	reachable subset for multi-input system
Δ_k	subspace spanned by $z_1,z_2,\ldots,z_k$
Λ	set of integers
Ψ	fundamental matrix
Φ	$(n \times n)$ real transition matrix

$\mathcal{e}$	edge
$\ell(x,y)$	the line segment between x and y
δ	integer, its value is either +1 or -1
$\forall x$	for all x
$\sim$	set difference operation
$\overline{X}$	set of points of closure of X
∂X	boundary of X
$'$	(prime) transpose of a matrix or a vector
$X \subset Y$	X is a subset of Y
$X \cap Y$	intersection of X and Y
(1)	w, x, y, z are vectors in E^n
(2)	α, β, γ, η, λ, μ, ν, ξ, ρ are scalar constants

CHAPTER 1 INTRODUCTION

During the last two decades, discrete optimal control problems have been investigated by many workers. One of the basic problems in the theory of optimal control is the time-optimal control problem. In general terms, the problem is the determination of system inputs which will take the system from a certain initial state to a prescribed terminal state in the shortest possible time, subject to various constraints. Numerous techniques for obtaining solutions to the time-optimal control problems have been proposed by many investigators [D2, D3, D4, E1, G1, H6, K6, N1, N4, T1, T2, T3, T4, W1, Z1]*. Other related problems such as minimum norm problems [C1, C2, R1], minimum energy problems [C4, R1], minimum effort problems [C3], terminal state problems [B2, N2, N3, P2], minimum time-fuel problems [T1, Z1], minimum time-energy problems [C5, H6] and problems of minimum quadratic functionals of states and/or controls [C6, C7, C8, C9, D1, G1, H4, H7, K2, P1] have also been considered. Linear programming [F1, H1, T1, Z1], non-linear programming [C5, H6, N2, N3, P1, R2], dynamic programming [B1, D1, T5], Pontryagin's maximum

*The alphanumeric numbers refer to the references at end of this dissertation.

principle [B3, C7, C8, C9, H2, H3, H7, H8, H9, R3, R4, R5] and variational methods [J1, K3, K4] are among the various techniques which have been used extensively.

One main subject of this dissertation is the discrete time-optimal control problem: determining an input signal which steers a linear discrete time-invariant n^{th} order system from the origin to a desired terminal state d in minimum time (the minimum number of sampling periods, N) subject to amplitude constraints on the input signal. When $n \geq N$, a method is proposed for finding the unique optimal solution.

When $n < N$, the optimal solution is, in general, non-unique [H6], and various procedures have been suggested for finding that time-optimal control which, among all time-optimal controls, also minimizes the energy supplied to the system [C5, H6] or which also minimizes the fuel consumption [T1, Z1]. A new algorithm is presented to find the minimum time N such that d is in R_N, the reachable set from the origin at time N. Then, a new method is proposed for finding the unique time-optimal control sequence whose first member is minimum in absolute value.

Another main subject is the terminal-error regulator problem: determining a control sequence such that the quantity $\|d - x_N\|^2$ is minimum, where N is the given terminal time. Based on the fact that the control input is constrained in magnitude, the reachable set can be described by the intersection of $2\binom{N}{n-1}$ closed half-spaces, and the boundary of the reachable set R_N is formed by subsets of hyperplanes. A new algorithm for solving the terminal-error regulator problem is developed. In contrast to a previously suggested method [N2, N3], this algorithm does not involve solving a quadratic programming problem.

The outline of the thesis is as follows. In Chapter 2 some elementary results on linear algebra, fundamental properties of the reachable set, and the computing algorithm for the optimal control u_N^* are presented. Chapter 3 contains the precise statement of the time-optimal control problem and algorithms for implementing the optimal control sequence. Chapter 4 involves the exact statement of the terminal-error regulator problem and algorithms for calculating the corresponding optimal control sequence. In Chapter 5 a brief summary of new results, and some possible extensions of these results are carefully stated. Finally, in Chapter 6 examples of time-optimal control problems and terminal-error regulator problems are given; certain pertinent conclusions are drawn by comparison with other existing techniques in the literature.

CHAPTER 2

PRELIMINARY ANALYSIS AND THEORETICAL DEVELOPMENT

In Section 2.1 some fundamental definitions and well-established theorems from vector spaces and matrix algebra needed in the sequel are presented. A description of the system to be studied is stated in Section 2.2 and the assumption of complete controllability is stated in Section 2.3. Section 2.4 contains the main body of the theory as well as properties of convex sets, polyhedra, and the reachable set. Section 2.5 involves the computation of the optimal control u_N^*.

2.1 Vector Spaces, Linear Transformations, Elementary Matrices, and Solution Properties of $Ax = y$.

A brief description of some standard definitions and well-known theorems from finite-dimensional vector spaces and linear algebra [F2, F3, H10] is given for the continuity of presentation in the sequel. Also the solution properties of the equation $Ax = y$, where A is $(n \times m)$, are discussed.

Definition If a Euclidean vector space has dimension n, then it is denoted by E^n. In E^n the scalar product is introduced by the formula:

$$(x,y) = x_1y_1 + x_2y_2 + \cdots\cdots + x_ny_n,$$

where x_i and y_i are the i^{th} components of x and y respectively. The <u>length</u> of a vector $x \in E^n$ is defined as

$$\|x\| = (x_1^2 + x_2^2 + \cdots\cdots + x_n^2)^{1/2}.$$

<u>Definition</u> Two vectors $x, y \in E^n$ are said to be <u>orthogonal</u> if their scalar product is equal to zero.

<u>Definition</u> The <u>identity matrix</u> of order n, written U or U_n, is a square matrix having ones along the main diagonal and zeros elsewhere.

$$U = \begin{bmatrix} 1 & 0 & 0 & \cdots & 0 \\ 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ \cdot & & & & \cdot \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} .$$

<u>Definition</u> Let A be an $(n \times n)$ matrix. If there exists an $(n \times n)$ matrix B which satisfies the relation

$$AB = BA = U_n,$$

then B is called the <u>inverse</u> of A.

<u>Definition</u> The <u>rank</u> of an $(m \times n)$ matrix A, written $r(A)$, is the maximum number of linearly independent columns in A.

<u>Theorem 2.1.1</u> If a square matrix A has an inverse, then so does A' and $(A')^{-1} = (A^{-1})'$, where A' denotes the

transpose of A.

Theorem 2.1.2 For any matrix A, $r(A)$, the rank of A is equal to $r(A')$.

Theorem 2.1.3 For two conformable matrices A and B, $r(AB) \leq \min\{r(A), r(B)\}$.

Theorem 2.1.4 If A is $(n \times k)$, and A has rank k, then $r(A'A) = r(A)$.

Theorem 2.1.5 Let $G = A(A'A)^{-1}A'$, where A is $(n \times k)$, $k < n$ and A has rank k. Then the image Gx under G of each vector $x \in E^n$ (also in E^n) lies in the subspace Δ_k of E^n spanned by the column vectors of A.

Proof By Theorem 2.1.4, the rank of $A'A$ is k. Since the matrix $A'A$ is $(k \times k)$, it has an inverse. Thus G is well-defined. For any vector $x \in E^n$, $Gx = A(A'A)^{-1}A'x$. Let $A = (z_1, z_2, \ldots, z_k)$ and

$$(A'A)^{-1}A'x = \begin{bmatrix} \bar{x}_1 \\ \bar{x}_2 \\ \cdot \\ \cdot \\ \cdot \\ \bar{x}_n \end{bmatrix} . \quad \text{Then} \quad Gx = \bar{x}_1 z_1 + \bar{x}_2 z_2 + \cdots + \bar{x}_k z_k .$$

Thus $Gx \in \Delta_k$.

Theorem 2.1.6 Let $G = A(A'A)^{-1}A'$. Then $G^2 = G = G'$, and Gx is orthogonal to $x - Gx$.

Proof Since $G = A(A'A)^{-1}A' = G'$, $G^2 = GG = [A(A'A)^{-1}A']$

$[A(A'A)^{-1}A'] = A(A'A)^{-1}A' = G$. $G'(U-G) = G - G^2 = 0$. $(Gx, x-Gx) = (Gx,(U-G)x) = (x,G'(U-G)x) = (x,0) = 0$. Hence Gx is orthogonal to $x - Gx$.

Definition The distinguished columns of A are the ρ nonzero columns of A, no one of which is a linear combination of its predecessors.

Definition An $(n \times k)$ matrix of rank ρ is a row echelon matrix if its last $(n-\rho)$ rows are zero, its distinguished columns are the first ρ columns of the identity matrix U_n in order, and the 1's in these columns are the first nonzero entries of their respective rows. If $\rho = n$, there are no rows of 0's and the $(n \times k)$ matrix is a reduced row echelon matrix.

The transpose of a row echelon matrix is a column echelon matrix.

Example 2.1.1 The following matrices are row echelon matrices. The fourth one is a reduced row echelon matrix.

$$\begin{bmatrix} 1 & 2 & 3 & 0 & 3 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}; \quad \begin{bmatrix} 1 & 2 & 0 & 1 & 0 \\ 0 & 0 & 1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

Definition A null matrix of dimension $(n \times k)$ is one with every entry 0.

Definition Let ϵ_{ij} denote the $(n \times k)$ matrix obtained from the $(n \times k)$ null matrix by replacing just one entry, the ij entry, by 1.

Now some solution properties of the equation $Ax = y$ are presented. Let the equation $Ax = y$ be given, where A is $(n \times k)$ known matrix of rank r and y is a known vector. When solving this equation for x, the following operations will simplify a set of n linear equations without changing the solution set (if any solutions exist):

E1. Add ν times equation j to a lower equation i (where $j < i$).

E2. Add ν times equation j to an upper equation i (where $j > i$).

E3. Multiply equation i by a scalar $\nu \neq 0$.

E4. Interchange equations i and j.

Given the equation $Ax = y$, it can be written as follows:

$$(A,y)\begin{bmatrix} x \\ -1 \end{bmatrix} = 0. \tag{1}$$

In actual computations, however, it is much simpler to work with the coefficient matrix (A,y) in (1) when y is a known constant vector. Thus the matrix (A,y) is transformed to a row echelon matrix by successive elementary operations on rows. These transformations have the same effect on the rows as the operations just described do on the equations. Each operation is effected by a left-sided multiplication by a corresponding elementary matrix, of one of the following types:

E1. Lower elementary: Add ν times row j to lower row i Matrix $L_{ij}(\nu) = U + \nu\, \epsilon_{ij}$ $j < i$.

E2. Upper elementary: Add ν times row j to upper row i Matrix $L_{ij}(\nu) = U - \nu\, \epsilon_{ij}$ $j > i$.

E3. Elementary diagonal: Multiply row i by $\nu \neq 0$ Matrix $L_i(\nu) = U + (\nu - 1)\, \epsilon_{ii}$.

E4. Transposition: Interchange rows i and j Matrix $L_{ij} = U - \epsilon_{ii} - \epsilon_{jj} + \epsilon_{ij} + \epsilon_{ji}$.

Suppose the equation $Ax = y$ is given. Let L denote the product of elementary matrices that reduce A to an echelon matrix LA, i.e.,

$$L[A, y] = \left[\begin{array}{c|c} L_1 A & L_1 y \\ \hline L_2 A & L_2 y \end{array}\right] = \left[\begin{array}{c|c} C & L_1 y \\ \hline 0 & L_2 y \end{array}\right], \quad (2)$$

where $L_2 A$ in (2) is the $(n - \pi \times k)$ null matrix, and $L_1 A = C$ is a $(\pi \times k)$ reduced echelon matrix. The following theorem is useful.

<u>Theorem 2.1.7</u> If A is an $(n \times k)$ matrix of rank π, and y is $(n \times 1)$ constant vector, the equation $Ax = y$ has a solution for x if $L_2 y = 0$ in (2).

<u>Proof</u> Since the product of elementary matrices is non-singular, the equation $Ax = y$ has the same solution set as the equation $L(Ax - y) = 0$. Equivalent to the equation

$Ax = y$ are the two equations: $L_1(Ax - y) = Cx - L_1y = 0$ and $L_2(Ax - y) = - L_2y = 0$ (see equation (2)). Thus the necessary condition for a solution is $L_2y = 0$. Sufficient conditions are $L_2y = 0$ and $Cx = L_1y$. This completes the proof.

Remark: Consider a special case of Theorem 2.1.7 where A is an $(n \times k)$ matrix with $k \le n$ and $\mathcal{r} = k$. In this case C is a reduced echelon matrix of dimension $(k \times k)$ and thus $C = U_k$. It is clear that if $L_2y = 0$, then the solution for x is given by $x = L_1y = (A'A)^{-1}A'y$.

2.2 Description of the System

Let an n^{th} order, time-invariant, scalar-control, linear discrete system $\mathcal{L}$ be governed by the following difference equation:

$$x((i+1)T) = \Phi x(iT) + bu(iT), \tag{3}$$

where i = integers 0,1,2,.....,

T = sampling period in seconds,

$x(iT) = (n \times 1)$ state vector at time iT,

$\Phi = (n \times n)$ transition matrix,

$b = (n \times 1)$ non-zero vector.

The scalar control has the following pre-determined properties:

$$u(t) = \text{constant} \qquad iT \le t < (i+1)T \tag{4}$$

$$|u(iT)| \le 1 \quad \text{for all } i. \tag{5}$$

In the subsequent development, the case where $x(0) = 0$ is considered for notational simplicity. The more general case when $x(0) \neq 0$ requires only slight modifications.

Since the system $\mathcal{L}$ is time-invariant, no loss of generality will occur if the system is started at $i = 0$ of (3). By iteration on (3), $x(NT)$ is given by

$$x(NT) = \Phi^{N-1} bu(0) + \Phi^{N-2} bu(T) + \cdots\cdots + \Phi\, bu((N-2)T) + bu((N-1)T). \qquad (6)$$

By defining $x_N = x(NT)$, $z_i = \Phi^{i-1} b$ for $i = 1,2,\ldots,N$ and $u_j = u((N-j)T)$ for $j = 1,2,\ldots,N$, (6) becomes

$$x_N = u_N z_N + u_{N-1} z_{N-1} + \cdots\cdots + u_1 z_1. \qquad (7)$$

<u>Definition</u> A <u>control sequence</u> $\vec{u}_N = [u_1,u_2,\ldots,u_N]$ of length N is called <u>admissible</u> if each u_i, $i = 1,2,\ldots,N$ satisfies (4) and (5).

2.3 Assumption of Complete Controllability

It is assumed that the system $\mathcal{L}$ is completely controllable in the sense of Kalman [K2]. This assumption is satisfied if and only if for any integer $k > 0$, the n vectors $z_k, z_{k+1},\ldots,z_{k+n-1}$ are linearly independent, where $z_i = \Phi^{i-1} b$, $i = k,k+1,\ldots,k+n-1$ [B2,K2]. Often a stronger condition than complete controllability is required; thus the following definition is useful.

<u>Definition</u> The system $\mathcal{L}$ is normal in the discrete sense if every set of n vectors $z_{i_1}, z_{i_2},\ldots,z_{i_n}$ with $i \leq i_1 < i_2 < \cdots < i_n$ is linearly independent.

2.4 Convex sets, Polyhedra, and Properties of the Reachable Set

Some standard definitions and properties concerning convex sets, polyhedra, and the reachable set which are needed in the subsequent presentation are given in this section.

Definition The set of points $X \subset E^n$ is said to be convex if whenever two points x_1, x_2 belong to X, all the points of the form

$$\lambda_1 x_1 + \lambda_2 x_2,$$

where $\lambda_1, \lambda_2 \geq 0$ and $\lambda_1 + \lambda_2 = 1$, also belong to X.

Definition The line segment $\ell(x,y)$ between x and y is the set of all points of the form $\alpha x + \beta y$, where $\alpha \geq 0$, $\beta \geq 0$, and $\alpha + \beta = 1$.

Definition The convex hull $h(Y)$ of any arbitrary set of points Y in E^n is the set of points which is the intersection of all the convex sets that contain Y.

Definition A set X in E^n is said to be open in E^n if for each point x in X, there is a positive number ϵ such that every point y in E^n satisfying $\|x-y\| < \epsilon$ also belongs to the set X.

Definition A point x in E^n is called a point of closure of a set X in E^n if $x \in X$ or if for every $\epsilon > 0$ there is a point y, $y \neq x$ and $y \in X$ such that $\|x-y\| < \epsilon$.

Notation The set of points of the closure of X is denoted by $\overline{X}$.

<u>Definition</u> A set D in E^n is <u>closed</u> if $D = \overline{D}$.

<u>Definition</u> If x is in E^n, then any set which contains an open set containing x is called a <u>neighborhood</u> of x in E^n.

<u>Definition</u> A point x is said to be an interior point of a set X if X is a neighborhood of the point x.

<u>Definition</u> The difference $X \sim Y$ is the set of elements in X which are not in Y. Thus $X \sim Y = \{x : x \in X \text{ and } x \notin Y\}$.

<u>Definition</u> A point x is said to be a <u>boundary point</u> of a subset X of E^n if every neighborhood of x contains a point of X and a point of $\mathcal{C}(X) = E^n \sim X$.

<u>Definition</u> The boundary of a subset X of E^n, denoted by ∂X, is the set of all boundary points of X.

<u>Definition</u> Let R_N be the set of states of $\mathcal{L}$ which can be reached at time N, starting from $x_0 = 0$, with an admissible control sequence $\vec{u}_N$, i.e.,

$$R_N \overset{d}{=} \{x_N \in E^n : x_N = \sum_{i=1}^{N} z_i u_i \quad \text{with} \quad \vec{u}_N = [u_1, u_2, \ldots, u_n]$$

admissible}, R_N is called the <u>reachable set</u> at time N.

It follows that R_N has the following properties:

<u>Property 1</u>: R_N is convex.

<u>Proof</u> Suppose $x_1 \in R_N$, $x_2 \in R_N$, i.e.,

$$x_1 = \sum_{i=1}^{N} z_i u_i \quad \text{with} \quad |u_i| \leq 1 \quad \text{and}$$

$$x_2 = \sum_{i=1}^{N} z_i \gamma_i \quad \text{with} \quad |\gamma_i| \le 1,$$

Then for $0 \le \lambda \le 1$,

$$\lambda x_1 + (1-\lambda) x_2 = \lambda \sum_{i=1}^{N} z_i u_i + (1-\lambda) \sum_{i=1}^{N} z_i \gamma_i$$

$$= \sum_{i=1}^{N} z_i [\lambda u_i + (1-\lambda)\gamma_i].$$

By the triangle inequality,

$$|\lambda u_i + (1-\lambda)\gamma_i| \le |\lambda u_i| + |(1-\lambda)\gamma_i|$$

$$= \lambda |u_i| + (1-\lambda)|\gamma_i|$$

$$\le \lambda + (1-\lambda) = 1.$$

Hence for $0 \le \lambda \le 1$, $\lambda x_1 + (1-\lambda)\, x_2 \in R_N$ if $x_1, x_2 \in R_N$, which implies that R_N is convex.

Property 2: R_N is closed and bounded, i.e., R_N is compact.

Proof By the definition of R_N, if $x_N \in R_N$, then

$$x_N = \sum_{i=1}^{N} z_i u_i$$

$$= [z_1, z_2, \ldots\ldots, z_n] \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_N \end{bmatrix}$$

$$= \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1N} \\ z_{21} & z_{22} & \cdots & z_{2N} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ z_{n1} & z_{n2} & \cdots & z_{nN} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_N \end{bmatrix} = Z \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_N \end{bmatrix} . \quad (10)$$

Since $z_i' = (z_{1i}, z_{2i}, \ldots, z_{ni})$, $i = 1,2,\ldots,N$ are given bounded constant vectors, the linear transformation Z is continuous. And since $|u_i| \le 1$, $i = 1,2,\ldots,N$, therefore the linear transformation Z in (10) maps the closed unit cube in the control signal space onto a closed subset $R_N \subset E^n$. Since $|u_i| \le 1$, $i = 1,2,\ldots,N$, is bounded, it is clear that R_N is bounded for N finite.

<u>Property 3</u>: R_N is symmetrical with respect to the origin, and hence 0 is an interior point of R_i, $i \ge n$.

<u>Proof</u> If $x \in R_N$, then $x = \sum_{i=1}^{N} z_i u_i$ with $|u_i| \le 1$ for $i = 1,2,\ldots,N$. Clearly $-x = \sum_{i=1}^{N} z_i (-u_i) \in R_N$, because $|-u_i| = |u_i| \le 1$. Since R_N is symmetrical with respect to the origin, 0 is an interior point of R_N.

<u>Property 4</u>: $R_N \supsetneqq R_{N-1} \supsetneqq R_{N-2} \supsetneqq \cdots\cdots \supsetneqq R_1 \supsetneqq R_0$, R_0 is a set containing one single element 0, the origin.

Proof $R_k = \{x \in E^n : x = \sum_{i=1}^{k} z_i u_i,\ \vec{u}_k = [u_1, u_2, \ldots, u_k]$ admissible$\}$.

$$R_{k+1} = \{x \in E^n : x = \sum_{i=1}^{k+1} z_i u_i,\ \vec{u}_{k+1} = [u_1, u_2, \ldots, u_{k+1}] \text{ admissible}\}.$$

Suppose $x \in R_k$, then $x = \sum_{i=1}^{k} z_i u_i$ with $|u_i| \leq 1$ for $i = 1,2,\ldots,k$. Take $u_{k+1} = 0$, then $x = \sum_{i=1}^{k} z_i u_i + z_{k+1}u_{k+1}$, which implies that $x \in R_{k+1}$. Thus $R_{k+1} \supset R_k$. By finite induction on k, it follows that

$$R_N \supset R_{N-1} \supset R_{N-2} \supset \cdots\cdots \supset R_1 \supset R_0 .$$

By the assumption that the system $\mathcal{L}$ is completely controllable, $z_i \neq 0\ \forall\, i$. Since $z_{k+1} \neq 0$, $R_{k+1} \neq R_k$, for $k = 0,1,\ldots,$ therefore $R_N \supsetneq R_{N-1} \supsetneq R_{N-2} \supsetneq \cdots\cdots \supsetneq R_1 \supsetneq R_0$. Also since the system $\mathcal{L}$ is started from the origin, hence at time 0, $R_0 = 0$.

Definition A hyperplane in E^n is the set H of all points $x \in E^n$ such that $a'x = \beta$, where a is an $(n \times 1)$ non-zero constant vector called the normal to H, and β is a scalar.

Definition A support hyperplane H_s to R_N is a hyperplane such that R_N lies entirely to one side of H_s and intersects R_N in at least one point, i.e., either $a'x \leq \beta$ or $a'x \geq \beta$.

Theorem 2.4.1 Let $a'x = \beta$ be a support hyperplane to R_N, $N \geq n$. Then

$$\beta > 0 \text{ if and only if } \forall\, x_N \in R_N,\ a'x_N \leq \beta,$$

and

$$\beta < 0 \text{ if and only if } \forall\, x_N \in R_N,\ a'x_N \geq \beta.$$

Proof Only the case where $\beta > 0$ is considered; the proof for $\beta < 0$ is similar.

($\Longleftarrow$) Assume $\beta > 0$. In view of Property 4 of R_N, $0 \in R_N$. Suppose $a'x_N > \beta\ \forall\, x_N \in R_N$. Then $a'0 = 0 > \beta$, a contradiction. Hence $a'x_N \leq \beta\ \forall\, x_N \in R_N$.

($\Longrightarrow$) Assume $a'x_N \leq \beta\ \forall\, x_N \in R_N$. Since $0 \in R_N$, hence $0 \leq \beta$. By Property 4 of R_N, 0 is an interior point of R_N, hence $\beta \neq 0$; for otherwise $a'x = 0$ cannot be a support hyperplane to R_N.

Definition Let $a'x = \beta$ be a support hyperplane H_s to R_N. Then a is called the outward normal to H_s if $\beta > 0$; a is called the inward normal to H_s if $\beta < 0$.

Theorem 2.4.2 Let $a'x = \beta$ be a support hyperplane to R_N. Then $|a'x_N| \leq |\beta|\ \forall\, x_N \in R_N$.

Proof

Case 1. $\beta > 0$. By Theorem 2.4.1, $a'x_N \leq \beta\ \forall\, x_N \in R_N$. By the symmetry property of R_N, $-x_N \in R_N$, hence $-a'x_N \leq \beta\ \forall\, x_N \in R_N$. It is clear that

$$|a'x_N| \leq \beta\ \forall\, x_N \in R_N, \text{ if } \beta > 0. \qquad (11)$$

Case 2. $\beta < 0$. Similarly, it can be obtained that

$$|a'x_N| \leq -\beta\ \forall\, x_N \in R_N, \text{ if } \beta < 0. \qquad (12)$$

Combining (11) and (12), it is clear that $|a'x_N| \le |\beta| \; \forall \; x_N \in R_N$.

<u>Definition</u> A point v in R_N is called a <u>vertex</u> if (1)

$v = \sum_{i=1}^{N} z_i u_i$, where $|u_i| = 1$, $i = 1,2,\ldots,N$ and (2) there

exists at least one support hyperplane to R_N with v as the only intersection point. Examples show that points

$v = \sum_{i=1}^{N} z_i u_i$, $|u_i| = 1 \; \forall \; i$ may be interior points of R_N;

hence property (2) is essential.

<u>Theorem 2.4.3</u> Consider the reachable set from the origin at time N, R_N, and a vertex v of R_N. If a support hyperplane to R_N at v is described by $a'x = \beta$ with $\beta > 0$,

$v = \sum_{i=1}^{N} z_i u_i$ and $|u_i| = 1$ for $i = 1,2,\ldots,N$, then u_i

satisfies $u_i a' z_i \ge 0 \; \forall \; i$. For $\beta < 0$, $u_i a' z_i \le 0 \; \forall \; i$.

<u>Proof</u> Consider the case $\beta > 0$. By Theorem 2.4.1,

$$a'x \le \beta \; \forall \; x \in R_N. \tag{13}$$

Suppose for some k, $u_k a' z_k < 0$. Then

$$a'v - 2u_k a' z_k = a'(v - 2u_k z_k) = \beta - 2u_k a' z_k. \tag{14}$$

But $v_1 = v - 2u_k z_k$ of (14) is in R_N, because

$$v_1 = v - 2u_k z_k = \sum_{\substack{i=1 \\ i \ne k}}^{N} z_i u_i - u_k z_k = \sum_{\substack{i=1 \\ i \ne k}}^{N} z_i u_i + z_k(-u_k),$$

where $|u_i| = 1 \; \forall \; i$. This implies $a'v_1 > \beta$, a contradiction.

Since $u_k a'z_k < 0$ for no k, therefore $u_i a'z_i \geq 0 \ \forall\, i$. Consider $\beta < 0$. By Theorem 2.4.1, if $a'x = \beta < 0$, then $a'x \geq \beta \ \forall\, x \in R_N$, or equivalently, $a'x \leq -\beta \ \forall\, x \in R_N$. If for some k, $u_k a'z_k > 0$, then

$$-a'v + 2u_k a'z_k = -\beta + 2u_k a'z_k > -\beta.$$

Because $v_2 = -v + 2u_k z_k$ is in R_N, hence a contradiction. Therefore if $\beta < 0$, then $u_i a'z_i \leq 0 \ \forall\, i$. This completes the proof.

Definition A point e in the closed convex reachable set R_N is an extreme point if there are no two points w, and y of R_N such that

$$e = \lambda w + (1-\lambda)y, \ 0 \leq \lambda \leq 1 \quad (w \neq e, \ y \neq e).$$

Lemma 2.4.4 Vertices of R_N are extreme points of R_N and vice versa.

Proof

This part is to prove that vertices of R_N are its extreme points. Let v be a vertex of R_N and let $a'x = \beta$ describe a support hyperplane to R_N at v. Without loss of generality it can be assumed that $\beta > 0$. Two cases are considered.

Case 1. $a'x < \beta \ \forall\, x \in R_N$ (see Figure 1a).

Suppose v is not an extreme point of R_N. Then there exist two points w and $y \in R_N$, $w \neq v$, $y \neq v$ such that

$$v = \lambda w + (1-\lambda)y \quad \text{with} \quad 0 \leq \lambda \leq 1.$$

Hence

$$a'v = \lambda a'w + (1-\lambda)a'y,$$

$$\leq \lambda \max\{a'w, a'y\} + (1-\lambda) \max\{a'w, a'y\},$$

$$= \max\{a'w, a'y\},$$

$$< \beta,$$

which is a contradiction to $a'v = \beta$. Therefore v is an extreme point.

Case 2. $a'x \leq \beta \ \forall \ x \in R_N$ (see Figure 1b).

Suppose v is not an extreme point. Then $v = \lambda w + (1-\lambda)y$, where $w = \sum_{i=1}^{N} z_i \rho_i$, $y = \sum_{i=1}^{N} z_i \eta_i$ with $|\rho_i| \leq 1$, and $|\eta_i| \leq 1 \ \forall \ i$. Thus

$$a'v = \sum_{i=1}^{N} u_i a' z_i = \beta \tag{15}$$

and also

$$a'v = \lambda \sum_{i=1}^{N} \rho_i a' z_i + (1-\lambda) \sum_{i=1}^{N} \eta_i a' z_i. \tag{16}$$

Since $w \neq v$ and $y \neq v$, there exist k_1 and k_2 such that $\rho_{k_1} \neq u_{k_1}$, and $\eta_{k_2} \neq u_{k_2}$. Hence $u_{k_1} a' z_{k_1} > \rho_{k_1} a' z_{k_1}$ and $u_{k_2} a' z_{k_2} > \eta_{k_2} a' z_{k_2}$ for $a' z_{k_1} \neq 0$ and $a' z_{k_2} \neq 0$. Thus it is clear from (16) that $a'v < \beta$, contrary to (15). Therefore v must be an extreme point.

If $a' z_{k_1} = a' z_{k_2} = 0$, the proof given above is not valid. However, recall that one of the necessary conditions for $v = \sum_{i=1}^{N} u_i z_i$ to be a vertex is $|u_i| = 1$, $i = 1,2,\ldots,N$. Clearly, if there is at least one i such that $|u_i| \neq 1$, then v is not a vertex. This argument can be used in the following proof.

Again suppose v is not an extreme point, then there exist two points w and y in R_N, $w \neq v$, $y \neq v$ such that

$$v = \lambda w + (1-\lambda)y \quad \text{with} \quad 0 < \lambda < 1, \tag{17}$$

where w and y are as given above. From (17), v can be written explicitly as

$$\begin{aligned} v &= \lambda \sum_{i=1}^{N} \rho_i z_i + (1-\lambda) \sum_{i=1}^{N} \eta_i z_i \\ &= \sum_{i=1}^{N} [\, \lambda \rho_i + (1-\lambda)\, \eta_i]\, z_i. \end{aligned} \tag{18}$$

Since $w \neq y$, therefore there exists at least one k_3 such that $\rho_{k_3} \neq \eta_{k_3}$. Thus $|\lambda \rho_{k_3} + (1-\lambda)\eta_{k_3}| \leq \lambda |\rho_{k_3}| + (1-\lambda)|\eta_{k_3}|$ < 1, and clearly v is not a vertex.

We now prove that the extreme points of R_N are its vertices. Let $v = \sum_{i=1}^{N} z_i u_i$ be a vertex with $|u_i| = 1 \; \forall\, i$. It is clear that $v_1 = \sum_{\substack{i=1 \\ i \neq k}}^{N} z_i u_i + z_k \rho_k$ is not a vertex if $|\rho_k| \neq 1$. Then $-1 < \rho_k < 1$. Therefore there exist two scalars ξ_1 and ξ_2, with $\xi_1 < \rho_k < \xi_2$ such that $v_1 = \lambda w + (1-\lambda)y$ where $w = \sum_{\substack{i=1 \\ i \neq k}}^{N} z_i u_i + \xi_1 z_k$ and

$y = \sum_{\substack{i=1 \\ i \neq k}}^{N} z_i u_i + \xi_2 z_k$. Clearly $w \neq v_1$, $y \neq v_1$ and

$w, y \in R_N$. Hence v_1 is not an extreme point. Thus the proof is completed. This theorem is valid for any linear systems.

The following theorem from Eggleston [E2] is presented here without proof:

Theorem 2.4.5 (1) A support hyperplane to the closed, bounded, convex set R_N contains at least one extreme point of R_N.

(2) The closed, bounded, convex set R_N is the closure of the convex hull of its extreme points.

Definition Let X be any convex, closed and bounded set in E^n, $n \geq 2$. An edge $\mathcal{e}$ is a line segment between two vertices v_1, v_2 of X such that there exists a support hyperplane H_s of X satisfying: (1) $H_s \supset \mathcal{e}$ (2) H_s contains no vertices except v_1 and v_2. Note: $\mathcal{e} \subset X$.

Definition A face is the intersection of a support hyperplane to R_N with R_N.

Clearly, a vertex is a point and hence a zero-dimensional face, an edge is a one-dimensional face. The largest dimension for a face in E^n is $(n-1)$. An $(n-1)$-dimensional face in E^n is a hyperplane which can be described as $H = \{x \in E^n : a'x = \beta\}$, where a is an $(n \times 1)$ constant non-zero vector and β a constant. If the hyperplane passes through the origin then $\beta = 0$. Only $(n-1)$ independent vectors are required to determine the components of a. Assume $N > n$. It follows from the assumption of complete controllability that any n consecutive vectors from $\{z_1, z_2, \ldots, z_N\}$ form a linearly independent set, thus $z_1, z_2, \ldots, z_{n-1}$ determine a hyperplane in E^n.

Let $a_1' = [a_{11}, a_{12}, \ldots, a_{1n}]$ and $z_1' = [z_{11}, z_{21}, \ldots, z_{n1}]$, $z_2' = [z_{12}, z_{22}, \ldots, z_{n2}], \ldots, z_{n-1}' = [z_{1,n-1}, z_{2,n-1}, \ldots, z_{n,n-1}]$. If $a'x = 0$ passes through $z_1, z_2, \ldots, z_{n-1}$, then

$$\left\{\begin{array}{l} z_{11}a_{11} + z_{21}a_{12} + \cdots\cdots + z_{n1}a_{1n} = 0 \\ z_{12}a_{11} + z_{22}a_{12} + \cdots\cdots + z_{n2}a_{1n} = 0 \\ \cdot \\ \cdot \\ \cdot \\ z_{1,n-1}a_{11} + z_{2,n-1}a_{12} + \cdots\cdots + z_{n,n-1}a_{1n} = 0. \end{array}\right. \tag{19}$$

Since $z_1, z_2, \ldots, z_{n-1}$ are linearly independent, (19) can be solved for, say, $a_{12}, a_{13}, \ldots, a_{1n}$ in terms of a_{11}, i.e., $a_1' = [\rho_{11}, \rho_{12}, \ldots, \rho_{1n}]\, a_{11}$, where $\rho_{11} = 1$. It can further be required that this hyperplane $a_1'x = 0$ pass through a point w_1 yet to be determined. That is,

$$\beta_1 = [\rho_{11}, \rho_{12}, \ldots, \rho_{1n}]x = [\rho_{11}, \rho_{12}, \ldots, \rho_{1n}]w_1. \tag{20}$$

Let $\pi = \{1,2,\ldots,N\}$, $\Lambda_1 = \{1,2,\ldots,n-1\}$ and $\mathcal{C}(\Lambda_1) = \{n,n+1,\ldots,N\} = \pi \sim \Lambda_1$. Define $w_1 = \sum_{k\in\mathcal{C}(\Lambda_1)} \eta_{1k} z_k$, where $|\eta_{1k}| = 1$ and $\eta_{1k}[\rho_{11},\rho_{12},\ldots,\rho_{1n}]z_k > 0$ for all $k \in \mathcal{C}(\Lambda_1)$. If there is some j such that $[\rho_{11},\rho_{12},\ldots,\rho_{1n}]z_j = 0$, then z_j is contained in the hyperplane $[\rho_{11},\rho_{12},\ldots,\rho_{1n}]x = 0$. Therefore z_j can be neglected in determining w_1. Consequently from $z_1,z_2,\ldots,z_{n-1}$ an (n-1)-dimensional face can be constructed which passes through w_1 and has the following form:

$$[1, \rho_{12},\ldots,\rho_{1n}]x = \beta_1. \tag{20'}$$

Division of both sides of (20') by β_1 yields

$$\left[\frac{1}{\beta_1}, \frac{\rho_{12}}{\beta_1}, \ldots, \frac{\rho_{1n}}{\beta_1}\right] x = 1. \tag{21}$$

For brevity (21) is written as

$$c_1'x = 1, \tag{22}$$

where $c_1' = \left[\frac{1}{\beta_1}, \frac{\rho_{12}}{\beta_1}, \ldots, \frac{\rho_{1n}}{\beta_1}\right]$.

From (22), two closed half-spaces can be constructed as:

$$\Gamma_1^1 = \{x \in E^n : c_1'x \geq -1\} \text{ and}$$
$$\Gamma_1^2 = \{x \in E^n : c_1'x \leq 1\}. \tag{23}$$

In a similar manner, from $z_1,z_2,\ldots,z_{n-2},z_n$, $\Lambda_2 = \{1,2,\ldots,n-2,n\}$, $\mathcal{C}(\Lambda_2) = \{n-1,n+1,\ldots,N\}$, $w_2 = \sum_{k\in\mathcal{C}(\Lambda_2)} \eta_{2k} z_k$

with $|\eta_{2k}| = 1$ and $\eta_{2k}[1,\rho_{22},\ldots,\rho_{2n}]z_k > 0$ for $k \in \mathcal{C}(\Lambda_2)$, the hyperplane can be constructed as

$$[1,\rho_{22},\ldots,\rho_{2n}]x = \beta_2,$$

where $\beta_2 = [1,\rho_{22},\ldots,\rho_{2n}]w_2$.

Finally let $c_2' = \left[\frac{1}{\beta_2}, \frac{\rho_{22}}{\beta_2}, \ldots, \frac{\rho_{2n}}{\beta_2}\right]$, then

$$c_2' x = 1. \tag{24}$$

From (24), two closed half-spaces can be constructed as:

$$\Gamma_2^1 = \{x \in E^n : c_2' x \geq -1\} \text{ and}$$

$$\Gamma_2^2 = \{x \in E^n : c_2' x \leq 1\}.$$

This procedure can be carried out consecutively up to the final step at which Γ_p^1 and Γ_p^2 are constructed from $z_{N-n+2}, z_{N-n+3}, \ldots, z_{N-1}, z_N$, where $p = \binom{N}{n-1}$, the total number of ways of choosing $(n-1)$ vectors from N vectors. It can be observed that there may exist some $(n-1)$ vectors from $\{z_1, z_2, \ldots, z_N\}$ which are linearly dependent. In this case, these $(n-1)$ vectors do not uniquely determine a hyperplane in E^n, and they are contained in some hyperplanes which pass through these $(n-1)$ linearly dependent vectors.

Definition All the hyperplanes $c_i'x = 1$ and $c_i'x = -1$, or in short form $|c_i' x| = 1$, $i = 1,2,\ldots,K$, $K \leq \binom{N}{n-1}$, are called boundary hyperplanes of R_N. For notational simplicity, let the total number of the boundary hyperplanes of

R_N be denoted by K.

Thus, given z_i's, it is possible to compute boundary hyperplanes of R_N, and hence Γ_i^j's.

<u>Example 2.4.1</u> Given $z_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$, $z_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$, $z_3 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ as in Figure 2, then $N = 3$, $n = 2$, and hence $p = \binom{N}{n-1} = \binom{3}{1} = 3$. Find Γ_1^1, Γ_1^2, Γ_2^1, Γ_2^2, Γ_3^1, Γ_3^2.

<u>Solution</u>: Let $a_1' = [a_{11}, a_{12}]$, $z_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}$. Then from (19) it follows that $2a_{11} + a_{12} = 0$ i.e., $a_{12} = -2a_{11}$.

Hence $a_1' = [1,-2]a_{11}$.

Now $\pi = \{1,2,3\}$, $\Lambda_1 = \{1\}$, $\mathcal{C}(\Lambda_1) = \{2,3\}$.

Define $w_1 = \eta_{12}z_2 + \eta_{13}z_3$, with $|\eta_{12}| = |\eta_{13}| = 1$.

Since $\eta_{12}[1,-2]\begin{bmatrix} 0 \\ 2 \end{bmatrix} = -4\eta_{12} > 0$, hence $\eta_{12} = -1$.

Since $\eta_{13}[1,-2]\begin{bmatrix} 1 \\ -1 \end{bmatrix} = 3\eta_{13} > 0$, hence $\eta_{13} = 1$.

Therefore $w_1 = -z_2 + z_3 = -\begin{bmatrix} 0 \\ 2 \end{bmatrix} + \begin{bmatrix} 1 \\ -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -3 \end{bmatrix}$.

$\beta_1 = [1,-2]\begin{bmatrix} 1 \\ -3 \end{bmatrix} = 7$.

Let $c_1' = \left[\frac{1}{7}, \frac{-2}{7}\right]$. Then $\Gamma_1^1 = \{x \in E^2 : \left[\frac{1}{7}, \frac{-2}{7}\right] x \geq -1\}$

and $\Gamma_1^2 = \{x \in E^2 : \left[\frac{1}{7}, \frac{-2}{7}\right] x \leq 1\}$.

Let $a_2' = [a_{21}, a_{22}]$, $z_2 = \begin{bmatrix} 0 \\ 2 \end{bmatrix}$. Then from (19) it follows that $0a_{21} + 2a_{22} = 0$ i.e., $a_{22} = 0$, and a_{21} any number.

$\therefore a_2' = [1,0]a_{21}$.

Now $\Lambda_2 = \{2\}$, and $\mathcal{C}(\Lambda_2) = \{1,3\}$.

Define $w_2 = \eta_{21}z_1 + \eta_{23}z_3$, with $|\eta_{21}| = |\eta_{23}| = 1$.

Since $\eta_{21}[1,0]\begin{bmatrix}2\\1\end{bmatrix} = 2\eta_{21} > 0$, $\therefore\ \eta_{21} = 1$.

Since $\eta_{23}[1,0]\begin{bmatrix}1\\-1\end{bmatrix} = \eta_{23} > 0$, $\therefore\ \eta_{23} = 1$.

Therefore $w_2 = z_1 + z_3 = \begin{bmatrix}2\\1\end{bmatrix} + \begin{bmatrix}1\\-1\end{bmatrix} = \begin{bmatrix}3\\0\end{bmatrix}$.

$$\beta_2 = [1,0]\begin{bmatrix}3\\0\end{bmatrix} = 3.$$

Let $c'_2 = [\frac{1}{3}, 0]$. Then $\Gamma^1_2 = \{x \in E^2 : [\frac{1}{3},0]x \geq -1\}$

and $\Gamma^2_2 = \{x \in E^2 : [\frac{1}{3},0]x \leq 1\}$.

Finally let $a'_3 = [a_{31}, a_{32}]$, and $z_3 = \begin{bmatrix}1\\-1\end{bmatrix}$. Then from (19), it follows that

$$a_{31} - a_{32} = 0 \qquad \text{i.e.,} \quad a_{32} = a_{31}.$$

$$\therefore\ a'_3 = [1,1]a_{31}.$$

Now $\Lambda_3 = \{3\}$, and $\mathcal{C}(\Lambda_3) = \{1,2\}$.

Define $w_3 = \eta_{31}z_1 + \eta_{32}z_2$, with $|\eta_{31}| = |\eta_{32}| = 1$.

Since $\eta_{31}[1,1]\begin{bmatrix}2\\1\end{bmatrix} = 3\eta_{31} > 0$, $\therefore\ \eta_{31} = 1$.

Since $\eta_{32}[1,1]\begin{bmatrix}0\\2\end{bmatrix} = 2\eta_{32} > 0$, $\therefore\ \eta_{32} = 1$.

Therefore $w_3 = z_1 + z_2 = \begin{bmatrix}2\\1\end{bmatrix} + \begin{bmatrix}0\\2\end{bmatrix} = \begin{bmatrix}2\\3\end{bmatrix}$.

$\beta_3 = [1\ 1]\begin{bmatrix}2\\3\end{bmatrix} = 5$.

Let $c'_3 = \begin{bmatrix}\frac{1}{5} & \frac{1}{5}\end{bmatrix}$. Then $\Gamma^1_3 = \{x \in E^2 : \begin{bmatrix}\frac{1}{5} & \frac{1}{5}\end{bmatrix} x \geq -1\}$

and $\Gamma^2_3 = \{x \in E^2 : \begin{bmatrix}\frac{1}{5} & \frac{1}{5}\end{bmatrix} x \leq 1\}$.

It is shown now that the reachable set R_N is directly related to the sets Γ^j_i's.

<u>Definition</u> Let $\Omega_1 = \Gamma_1^1 \cap \Gamma_1^2$, $\Omega_2 = \Gamma_2^1 \cap \Gamma_2^2$, $\ldots$, $\Omega_K = \Gamma_K^1 \cap \Gamma_K^2$.

<u>Theorem 2.4.6</u> For the reachable set at time N, R_N, $R_N = \Omega_1 \cap \Omega_2 \cap \ldots \cap \Omega_K$, where Ω_i, $i = 1,2,\ldots,K$ are defined above.

<u>Proof</u> First prove $R_N \subset \Omega_1$. Any point $x_N \in R_N$ can be written as $x_N = \sum_{i=1}^{N} z_i u_i$, $|u_i| \le 1$, $i = 1,2,\ldots,N$. Hence $c_1' x_N = \sum_{i=1}^{N} c_1' z_i u_i$. Since $\Omega_1 = \Gamma_1^1 \cap \Gamma_1^2$, $c_1' z_i = 0$ for $i = 1,2,\ldots,n-1$ and $\sum_{k\in\mathcal{C}(\Lambda_1)} \eta_{1k} c_1' z_k = 1$ where $|\eta_{1k}| = 1$ for $k \in \mathcal{C}(\Lambda_1) = \{n,n+1,\ldots,N\}$, hence

$$|c_1' x_N| = \left|\sum_{i=1}^{N} c_1' z_i u_i\right| \le \sum_{k\in\mathcal{C}(\Lambda_1)} |u_i||c_1' z_k| \le |c_1' z_k|. \tag{25}$$

The last inequality of (25) follows because of $|u_i| \le 1$, $i = 1,2,\ldots,N$. But

$$|c_1' z_k| = \left|\left[\frac{1}{\beta_1}, \frac{\rho_{12}}{\beta_1}, \ldots, \frac{\rho_{1n}}{\beta_1}\right] z_k\right| = \eta_{1k}\left[\frac{1}{\beta_1}, \frac{\rho_{12}}{\beta_1}, \ldots, \frac{\rho_{1n}}{\beta_1}\right] z_k > 0, \tag{26}$$

where $|\eta_{1k}| = 1$ for $k \in \mathcal{C}(\Lambda_1)$.

Then
$$\sum_{k\in\mathcal{C}(\Lambda_1)} |c_1' z_k| = \sum_{k\in\mathcal{C}(\Lambda_1)} \eta_{1k}\left[\frac{1}{\beta_1}, \frac{\rho_{12}}{\beta_1}, \ldots, \frac{\rho_{1n}}{\beta_1}\right] z_k$$
$$= \frac{1}{\beta_1} \sum_{k\in\mathcal{C}(\Lambda_1)} [1,\rho_{12},\ldots,\rho_{1n}] z_k = \frac{1}{\beta_1}\beta_1 = 1. \tag{27}$$

From (25), (26) and (27), it follows that

$$|c_1' x_N| \leq 1. \tag{28}$$

From (28), it follows that $x_N \in \Omega_1 = \Gamma_1^1 \cap \Gamma_1^2$. By similar arguments it can be shown that $x_N \in \Omega_2, \ldots, \Omega_K$. Hence $x_N \in \Omega_1 \cap \Omega_2 \cap \cdots \cap \Omega_K$.

Next it remains to be shown that $\Omega_1 \cap \Omega_2 \cap \ldots \cap \Omega_K \subset R_N$. This is equivalent to show that $x_N = \sum_{i=1}^{N} z_i u_i \notin R_N$ implies that $x_N \notin \Omega_1 \cap \Omega_2 \cap \ldots \cap \Omega_K$. By definition of a vertex v, there exists at least one support hyperplane $a'x = \beta$ to R_N with v as an intersection point, and v satisfies $v = \sum_{i=1}^{N} z_i u_i$ with $|u_i| = 1$, $i = 1,2,\ldots,N$. Since R_N is closed, convex, and symmetrical with respect to the origin, therefore c_i is an outward normal to R_N and perpendicular to the hyperplane $c_1' x = 1$. By definition, Ω_i is constructed from $c_i' x = 1$, with $x = \sum_{k=1}^{N} z_k u_k$, $|u_k| = 1 \; \forall \; k \in \mathcal{C}(\Lambda_i)$. For those $j \in \Lambda_i$, $c_i' z_j = 0$, hence u_j can be taken as either 1 or -1. Since there are $(n-1)$ elements in Λ_i, hence $c_i' x = 1$ actually passes through 2^{n-1} vertices. All of these 2^{n-1} v's satisfy $c_i' v = 1$, while $c_i' x_N \leq 1 \; \forall \; x_N \in R_N$. The corresponding symmetrical hyperplane is defined by $c_i' v = -1$, while $c_i' x_N \geq -1 \; \forall \; x_N \in R_N$.

By Theorem 2.4.5, the reachable set R_N is the convex hull of its vertices, thus if $x_N \notin R_N$, then $|c_k' x_N| > 1$ for at least one k. This implies that if $x_N \notin R_N$, then $x_N \notin \Omega_k$ for at least one k. Consequently $x_N \notin \Omega_1 \cap \Omega_2 \cap \ldots \cap \Omega_K$. This establishes the proof.

By Theorem 2.4.6, the reachable set from the origin at time N, R_N, is described as $\Omega_1 \cap \Omega_2 \cap \ldots \cap \Omega_K$, where $K \leq \binom{N}{n-1}$ and $\Omega_i = \{x \in E^n : |c_i' x| \leq 1\}$ for $i = 1,2,\ldots,K$. Here c_i's are constructed from a set of $(n-1)$ vectors in $\{z_1, z_2, \ldots, z_N\}$. R_{N+1} is described in a similar manner, the only difference is that for R_{N+1} the set is $\{z_1, z_2, \ldots, z_N, z_{N+1}\}$. All the combinations of $(n-1)$ vectors from $\{z_1, z_2, \ldots, z_{N+1}\}$ consist of two parts: one part contains z_{N+1}, the other does not. That part which does not contain z_{N+1} is simply all the combinations of $(n-1)$ vectors from the set $\{z_1, z_2, \ldots, z_N\}$. The other part which contains z_{N+1} corresponds to all the combinations of $(n-2)$ vectors from the set $\{z_1, z_2, \ldots, z_N\}$ plus z_{N+1} with a total number of $(n-1)$ vectors. Thus it is obvious that there is a certain relation between the c_i's of R_N and those $\tilde{c}_i$'s of R_{N+1}. The following theorem states this relation.

<u>Theorem 2.4.7</u> If R_N has $|c_i' x| = 1$ as a boundary hyperplane, then R_{N+1} has $|\tilde{c}_i' x| = |\frac{c_i'}{1+|\xi|} x| = 1$ as its

corresponding boundary hyperplane, where $\xi = c_1' z_{N+1}$.

<u>Proof</u> Since the c_i, for $i = 1,2,\ldots,K$ with $K \le \binom{N}{n-1}$, are constructed in exactly similar manner, the theorem is proved without loss of generality for the first case $i = 1$. Thus consider c_1 and $\tilde{c}_1$ be constructed from $z_1, z_2, \ldots, z_{n-1}$. For R_N and R_{N+1}, $[\rho_{11}, \rho_{12}, \ldots, \rho_{1n}]$ are the same, as demonstrated in equation (18). For R_N, w_1 is defined as $w_1 = \sum_{k \in \mathcal{C}(\Lambda_1)} \eta_{1k} z_k$, where $|\eta_{1k}| = 1$ and $\eta_{1k}[\rho_{11}, \rho_{12}, \ldots, \rho_{1n}] z_k > 0$. Then define $\beta_1 = [\rho_{11}, \rho_{12}, \ldots, \rho_{1n}] w_1$ and $c_1' = \left[\frac{\rho_{11}}{\beta_1}, \frac{\rho_{12}}{\beta_1}, \ldots, \frac{\rho_{1n}}{\beta_1}\right]$. Now for R_{N+1}, $\tilde{w}_1 = \sum_{k \in \mathcal{C}(\Lambda_1)} \eta_{1k} z_k + \eta_{1,N+1} z_{N+1} = w_1 + \eta_{1,N+1} z_{N+1}$, where $|\eta_{1,N+1}| = 1$ and $\eta_{1,N+1}[\rho_{11}, \rho_{12}, \ldots, \rho_{1n}]\, z_{N+1} > 0$. Define $\tilde{\beta}_1 = [\rho_{11}, \rho_{12}, \ldots, \rho_{1n}] \tilde{w}_1$, which is the same as $\frac{\tilde{\beta}_1}{\beta_1} = \left[\frac{\rho_{11}}{\beta_1}, \frac{\rho_{12}}{\beta_1}, \ldots, \frac{\rho_{1n}}{\beta_1}\right] \tilde{w}_1$. Clearly $\frac{\tilde{\beta}_1}{\beta_1} = 1 + \left| \left[\frac{\rho_{11}}{\beta_1}, \frac{\rho_{12}}{\beta_1}, \ldots, \frac{\rho_{1n}}{\beta_1}\right] z_{N+1} \right| = 1 + |\xi|$, where $\xi = c_1' z_{N+1}$. Therefore $\tilde{c}_1' = \left[\frac{\rho_{11}}{\beta_1(1+|\xi|)}, \frac{\rho_{12}}{\beta_1(1+|\xi|)}, \ldots, \frac{\rho_{1n}}{\beta_1(1+|\xi|)}\right] = \frac{1}{1+|\xi|} c_1'$. This completes the proof.

Thus in constructing the boundary hyperplanes $|\tilde{c}_i' x| = 1$ of R_{N+1}, it is only required to solve $\binom{N}{n-2}$ additional systems of $(n-1)$ linear equations in $(n-1)$

unknowns. All the others are constructed from those $|c_i' x| = 1$ of R_N by merely a simple translation with an adjustment factor of $|\xi| = |c_i' z_{N+1}|$. Considerable time is reduced in using this property to construct the boundary hyperplanes $|\tilde{c}_i' x| = 1$ of R_{N+1} from those $|c_i' x| = 1$ of R_N.

2.5 Computation of the Optimal Control u_N^*

The method for constructing R_N was presented in Section 2.4, for given z_i's. For a certain point $f \in E^n$, if $f \in R_N$ and $f \notin R_{N-1}$, then the optimal control u_N^* can be found. Theorem 2.5.1 establishes the optimal condition that u_N^* satisfies. An algorithm for finding the u_N^* is proposed.

Theorem 2.5.1 Given $N > n$, $z_1, z_2, \ldots, z_N$ and $f \in R_N - R_{N-1}$. If u_N^* is the value u_N such that $f - u_N z_N \in R_{N-1}$ and $|u_N|$ = minimum, define $f_0 = f - u_N^* z_N \in R_{N-1}$. Then f_0 is unique and lies on the boundary of R_{N-1}.

Proof (1) Because $|u_N| > 0$ for $u_N \neq 0$ and 0 for $u_N = 0$, its minimum is unique and at $u_N = 0$ if no restriction is imposed on u_N (see Figure 3). But since it is required that $f = f_0 + u_N z_N$, where $f \in R_N - R_{N-1}$ and $f_0 \in R_{N-1}$, therefore u_N cannot be zero. It is shown in Figure 4 that u_N^* is the minimum of $|u_N|$ such that $f = f_0 + u_N z_N$. It is clear that if u_N^* is the minimum then

$-u_N^*$ cannot be the minimum. For if $f = f_0 + u_N^* z_N$ and $f = f_0 - u_N^* z_N$, then $f = f_0$, which is a contradiction. Therefore u_N^* is unique, and hence f_0 is unique.

(2) By the result of (1), $|u_N^*|$ is the minimum distance from f to $f - u_N z_N$ (in the direction of z_N), where $f - u_N z_N \in R_{N-1}$. Since R_{N-1} is a closed convex subset of E^n and $f \notin R_{N-1}$, hence $f_0 = f - u_N^* z_N$ is on the boundary of R_{N-1}.

Let $f = f_0 + u_N^* z_N$, where u_N^* is as defined in Theorem 2.5.1. Also let $\overline{z}_N = u_N^* z_N$, then $f = f_0 + \frac{u_N^*}{|u_N^*|}\overline{z}_N$ $= f_0 + \gamma_N^* \overline{z}_N$. Define $\overline{R}_N = \{x_N \in E^n : x_N = \sum_{i=1}^{N-1} z_i u_i + \overline{z}_N u_N,$ $|u_i| \leq i = 1,2,\ldots,N\}$, then f is on the boundary of $\overline{R}_N$. Let the boundary hyperplanes of $\overline{R}_N$ be specified by $|\overline{c}_i' x| = 1$ and those of R_N by $|c_i' x| = 1$.

<u>Theorem 2.5.2</u> Let u_N^* satisfy Theorem 2.5.1, i.e., u_N^* is the value u_N such that $f - u_N z_N \in R_{N-1}$ and $|u_N|$ = minimum, and f_0 is defined as $f_0 = f - u_N^* z_N$. Also let $\gamma_N^* = \frac{u_N^*}{|u_N^*|}$, $\overline{z}_N = u_N^* z_N$ and let f satisfy $\overline{c}_i' f = \delta$, where δ is either 1 or -1. Then $\overline{c}_i' f > 0$ if and only if $\overline{c}_i' f_0 > 0$; and similarly $\overline{c}_i' f < 0$ if and only if $\overline{c}_i' f_0 < 0$.

<u>Proof</u> Consider $\delta = 1$. Then $\overline{c}_i' f = 1$ and $\overline{c}_i$ is the outward normal to $\overline{R}_N$. Clearly $\overline{c}_i$ is perpendicular to

the hyperplane $\overline{c}_i' x = 1$. Since $\overline{R}_N \supsetneq R_{N-1}$ and $\overline{R}_N$ contains the origin, hence $\overline{c}_i' f > \overline{c}_i' f_0$.

(i) Suppose $\overline{c}_i' f_0 > 0$. It is obvious that $\overline{c}_i' f > c_i' f_0 > 0$.

(ii) Assume $\overline{c}_i' f = 1$. Since $f = f_0 + \gamma_N^* \overline{z}_N$, thus

$$\overline{c}_i' (f_0 + \gamma_N^* z_N) = 1. \tag{29}$$

Because $\overline{z}_N \in R_N$, thus $\overline{c}_i' \overline{z}_N < 1$. From (29), it follows that

$$\overline{c}_i' f_0 = 1 - \gamma_N^* \overline{c}_i' \overline{z}_N > 0. \tag{30}$$

The last inequality of (30) follows because $|\gamma_N^*| = 1$.
The case $\delta = -1$ can be proved similarly by remembering $\overline{c}_i' f < \overline{c}_i' f_0 < 0$. Thus the proof is completed.

Remark For a special case of the theorem when $f \in \partial R_N$, then $|u_N^*| = 1$ and this theorem states: $c_i' f > 0$ if and only if $c_i' f_0 > 0$; and similarly $c_i' f < 0$ if and only if $c_i' f_0 < 0$.

Corollary 2.5.3 Let the boundary hyperplanes of R_N be specified by $|c_i' x| = 1$, $i = 1,2,\ldots,K$, where $K \le \binom{N}{n-1}$.
If $f \in R_N$, $f \in R_{N-1}$, and $f_0 \in \partial R_{N-1}$, then $c_i' f > 0$ if and only if $c_i' f_0 > 0$; and similarly $c_i' f < 0$ if and only if $c_i' f_0 < 0$.

Proof The necessary part is obvious, since $f_0 \in R_{N-1} \subsetneq R_N$, and $f \in R_N$. Now suppose that $c_i' f > 0$, then

$$c_i' f = c_i'(f_0 + u_N z_N) > 0. \tag{31}$$

But c_i is found such that

$u_N c_i' z_N \geq 0.$ (32)

Equality of (32) holds if $c_i' z_N = 0$. Thus it is clear from (32) that $c_i' f_0 > 0$. An alternative proof is easily seen by applying Theorems 2.4.7 and 2.5.2. The proof that $c_i' f < 0$ if and only if $c_i' f_0 < 0$ can be carried out similarly.

<u>Corollary 2.5.4</u> Let the boundary hyperplanes of R_{N-1} be specified by $|g_i' x| = 1$, $i = 1,2,\ldots,K'$, where $K' \leq \binom{N-1}{n-1}$. If $f \in R_N$, $f \notin R_{N-1}$, and $f_0 \in \partial R_{N-1}$, then $g_i' f > 0$ if and only if $g_i' f_0 > 0$; similarly $g_i' f < 0$ if and only if $g_i' f_0 < 0$.

<u>Proof</u> Only the first half and the sufficient part is proved. Suppose $g_i' f > 1 > 0$, then there exists a corresponding boundary hyperplane $c_i' x = 1$ of R_N such that $c_i' f > 0$, where $c_i = \dfrac{g_i}{1+|g_i' z_N|}$ by Theorem 2.4.7. Since $c_i' f > 0$, it is clear that $c_i' f_0 > 0$ by Corollary 2.5.3. Consequently $g_i' f_0 > 0$.

By applying Theorem 2.5.2 and Corollaries 2.5.3 and 2.5.4, the unique u_N^* satisfying Theorem 2.5.1 can be found by the following

<u>Algorithm for Computing u_N^*</u> Let $f \in R_N$ but $f \notin R_{N-1}$. Assume that the boundary hyperplanes of R_{N-1} are specified by $|g_i' x| = 1$, $i = 1,2,\ldots,K'$, where $K' \leq \binom{N-1}{n-1}$. Since by the assumption that $f \notin R_{N-1}$, then $|g_i' f| > 1$ for at

least one i. It is desired to find that g_i for which $g_i'(f - u_N z_N) = \delta = \pm 1$ and further $(f - u_N z_N) \in \partial R_{N-1}$, or equivalently

$$g_i' f = \delta + (g_i' z_N) u_N. \tag{33}$$

From (33), u_N is given as

$$u_N = \frac{g_i' f - \delta}{(g_i' z_N)} . \tag{34}$$

If this u_N satisfies $(f - u_N z_N) \in \partial R_{N-1}$, then it is u_N^*.

Proof of the Algorithm Since $f = f_0 + u_N z_N$, thus

$$g_i' f = g_i'(f_0 + u_N z_N) = g_i' f_0 + u_N(g_i' z_N). \tag{35}$$

If $g_i' f > 1 > 0$, then $g_i' f_0 > 0$ by Corollary 2.5.4. Take $g_i' f_0 = 1$. Since f_0 is required to be on the boundary hyperplane $g_i' x = 1$ of R_{N-1}, thus

$$g_i' f = 1 + u_N(g_i' z_N) = 1 + \sigma. \tag{35'}$$

From (35'), it yields

$$u_N = \frac{\sigma}{(g_i' z_N)} . \tag{36}$$

Similarly, if $g_i' f < -1 < 0$, then $g_i' f_0 < 0$. Thus

$$g_i' f = -1 + u_N(g_i' z_N) = -1 + \sigma'. \tag{37}$$

From (37) then

$$u_N = \frac{\sigma'}{(g_i' z_N)} . \tag{38}$$

After having found u_N by either (36) or (38), it can be checked whether $(f - u_N z_N) \in \partial R_{N-1}$. If indeed $(f - u_N z_N) \in \partial R_{N-1}$, then this u_N is u_N^*, the optimal solution. On the other hand, if $(f - u_N z_N) \notin \partial R_{N-1}$, then other g_i's have to be considered for which $|g_i' f| > 1$, until this unique u_N^* is found such that $(f - u_N^* z_N) \in \partial R_{N-1}$.

Theorem 2.5.5 If u_N^* satisfies Theorem 2.5.1, then $x_i = \sum_{k=1}^{i} u_K z_K \in \partial R_i$, $i = 2,3,\ldots,N-1$, with $x_1 \in R_1$ and $x_N \in \partial \bar{R}_N$.

Proof By Theorem 2.5.1, $f_0 \in \partial R_{N-1}$. Clearly $f_0 = x_{N-1} \in \partial R_{N-1}$. Suppose for some $i \neq 1$ such that $x_i \notin \partial R_i$. Then either $x_i \notin R_i$ or x_i is an interior point of R_i. $x_i \notin R_i$ is not possible. If x_i is an interior point of R_i, then x_{i+j} is an interior point of R_{i+j}, $j = 1,2,\ldots,N-i-1$. In particular, x_{N-1} is an interior point of R_{N-1}. This is a contradiction to $x_{N-1} \in \partial R_{N-1}$. Therefore for $i = 2,3,\ldots,N-1$, $x_i \in \partial R_i$, and $x_N \in \partial \bar{R}_N$. For all $i = 2,3,\ldots,N-1$, $x_i \in \partial R_i$, it is not required that $x_1 \in \partial R_1$. Actually, ∂R_1 is simply the set $\{z_1, -z_1\}$. As Figure 5 shows, $x_2 \in \partial R_2$ without x_1 being on the boundary of R_1.

CHAPTER 3 TIME-OPTIMAL CONTROL PROBLEM

3.1 Statement of Time-Optimal Control Problem

Given the linear system $\mathcal{L}$ as described in Section 2.2, and a desired target point $d \in E^n$, $d' = (d_1, d_2, \ldots, d_n)$, find the smallest integer N and an admissible control sequence $\vec{u}_N = [u_1, u_2, \ldots, u_N]$ such that $d \in R_N$, i.e.,

$$u_1 z_{i1} + u_2 z_{i2} + \cdots + u_N z_{iN} = d_i, \; i = 1,2,\ldots,n \quad (1)$$

$$|u_j| \le 1, \; j = 1,2,\ldots,N \quad (2)$$

where d_i is the i^{th} element of d, and z_{ij} is the i^{th} element of z_j.

Observe that equation (1) can be rewritten in several equivalent forms:

$$z_1 u_1 + z_2 u_2 + \cdots\cdots + z_N u_N = d, \quad (3)$$

$$\begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1N} \\ z_{21} & z_{22} & \cdots & z_{2N} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ z_{n1} & z_{n2} & \cdots & z_{nN} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_N \end{bmatrix} = \begin{bmatrix} d_1 \\ d_2 \\ \cdot \\ \cdot \\ \cdot \\ d_n \end{bmatrix}, \quad (4)$$

$$Z\,u = d, \tag{5}$$

where $Z = (z_1, z_2, \ldots, z_N)$ is an $(n \times N)$ constant matrix.

It can be shown [D4, H6, K1] that for unstable systems, a solution exists for (1) and (2) while for stable systems, a solution exists for very limited cases.

Consider the last n columns of the matrix Z. Then it can be asserted that : 1. The last n columns of Z are linearly independent, and thus form a basis for the n-dimensional state space. 2. The matrix Φ^N becomes unbounded when $N \to \infty$ for an unstable system.

From the preceding two statements, it is a simple matter to conclude that an unstable system with arbitrary constraints can always be steered to any desired terminal state d, $\|d\| < \infty$, by choosing as control policy zero controls for the first $N-n$ members and some appropriate control signals for the last n members for N sufficiently large. On the other hand, only a sufficiently small region around the origin can be reached by a stable system starting from the origin.

3.2 Solution Properties and Computing Algorithm for Time-Optimal Control Problem

Because a direct method to find the minimum control time required is not available, a systematic iterative process is employed. To ensure that the N found is the smallest integer such that $d \in R_N$, N is set initially to 1. The computer algorithm for the time-optimal control problem proceeds in four phases as is illustrated by Figure 6.

Phase 1. For $N \le n$, does $Zu = d$ for some $(N \times 1)$ vector u?

Let $N = 1$, does $Zu = d$ for some $(N \times 1)$ vector u? If the answer is in the affirmative, Phase 2 is started. If the answer is in the negative, set $N = 2$ and Phase 1 is repeated. This process is carried out iteratively for $N \le n$. If $N = n$ and $d \notin R_N$, then set $N = n + 1$, and Phase 3 is initiated. In Phase 1, the minimal integer N has to be found for which (5) is satisfied. To see whether there is a solution for $Zu = d$, Theorem 2.1.7 can be employed. Let L denote the product of the elementary matrices that convert Z to an echelon matrix

$$L\,(Z,d) = \left[\begin{array}{c|c} L_1 Z & L_1 d \\ \hline L_2 Z & L_2 d \end{array}\right] = \left[\begin{array}{c|c} C & L_1 d \\ \hline 0 & L_2 d \end{array}\right] . \qquad (6)$$

If $L_2 d \neq 0$, then there is no solution for u_i, $i = 1,2,\ldots,N$ such that (5) is satisfied. In this case, N is increased by 1 and Phase 1 computations are carried out repeatedly. If $L_2 d \neq 0$ for $N = 1,2,\ldots,n$, then N is set equal to $n + 1$ and the calculations are done in

Phase 3. If for some $N \leq n$, $L_2 d = 0$, then Phase 2 calculations are performed to find the unique u_i's for which (4) is satisfied.

Phase 2. Compute $u = L_1 d$, is u admissible?

If the control sequence $u = L_1 d$ found is admissible, then the time-optimal control problem is solved and the unique optimal control is $\vec{u}_N = u$. However, if u is not admissible, then Phase 3 is started.

Following the assumption that the system $\mathcal{L}$ is completely controllable, the N vectors $z_i = \Phi^{i-1} b$, $i = 1, 2, \ldots, N \leq n$ are linearly independent. Therefore it is obvious that the control sequence is unique and is given by $C^{-1} L_1 d = L_1 d$, where the i^{th} component of $L_1 d$ corresponds to u_i. Now it is clear whether this control sequence is admissible. Suppose the control sequence is not admissible, then the desired state d cannot be reached in less than or equal to n sampling periods, as the following Lemma 3.2.1 shows. Thus Phase 3 has to be considered.

Lemma 3.2.1 Let $d \in E^n$, and $u_1'' z_1 + u_2'' z_2 + \cdots + u_k'' z_k = d$ with $k \leq n$. Then there exist no u_i's satisfying

$$u_1 z_1 + u_2 z_2 + \cdots + u_k z_k + u_{k+1} z_{k+1} + \cdots + u_n z_n = d,$$

where $u_i'' \neq u_i$ for $i = 1, 2, \ldots, k$.

Proof Assume the converse is true. Then

$$u_1'' z_1 + u_2'' z_2 + \cdots + u_k'' z_k = d \tag{7}$$

and $$u_1z_1 + u_2z_2 + \cdots + u_kz_k + u_{k+1}z_{k+1} + \cdots + u_nz_n = d. \tag{8}$$

Subtract (7) from (8), then

$$(u_1-u_1'')z_1 + \cdots + (u_k-u_k'') + u_{k+1}z_{k+1} + \cdots + u_nz_n = 0. \tag{9}$$

Since $z_1, z_2, \ldots, z_n$ are linearly independent, hence $u_1 = u_1''$, $u_2 = u_2''$, ..., $u_k = u_k''$, and $u_{k+1} = 0$, ..., $u_n = 0$. This completes the proof.

Remark: For $N = n$, Z has an inverse. Thus $u = Z^{-1} d$ can be obtained directly and checked whether it is admissible.

It has been established in either Phase 1 or Phase 2 that $N \le n$ is not the minimal time. The calculations to find the minimal time N and the corresponding optimal control for $N > n$ are considered in Phases 3 and 4. For the case when $N > n$, it has been shown [H6] that there exists either a unique or an infinity of control sequences to the time-optimal control problem. The calculations in Phase 3 will indicate whether the control sequence is unique or not.

Phase 3. For $N > n$, is $d \in R_N$?

By Theorem 2.4.6, $R_N = \Omega_1 \cap \Omega_2 \cap \ldots \cap \Omega_K$. Thus $d \notin R_N$ if and only if $d \notin \Omega_i$, for at least one i, where $1 \le i \le K$. Since $\Omega_i = \{x \in E^n : |c_i' x| \le 1\}$, to find whether $d \notin \Omega_i$, or equivalently $d \notin R_N$, it is only necessary to check whether $|c_i' d| > 1$ for some i. If $d \notin R_N$, then N is increased by 1, and check whether $d \in R_N$, etc. The procedure can be continued up to a certain N, such that $d \in R_N$. If $d \in R_N$, this algorithm will indicate whether the control sequence is unique. As is shown in Theorem 3.2.2, any point $d \in \partial R_N$ with $N > n$ corresponds to a unique optimal control sequence. Thus if $d \in R_N$, and if $|c_j' d| = 1$ for at least one j, then $d \in \partial R_N$, and the optimal control sequence is unique. If $d \in R_N$, and the solution for control is not unique, then Phase 4 has to be started.

Phase 4. Find $\vec{u}_N$ for which $|u_N|$ = minimum, and the algorithm terminates.

When an infinity of control sequences exists, it is desirable to choose, among those which satisfy the minimum-time, a unique control sequence which also minimizes $|u_N|$, where $u_N = u(0)$ is the first member of the control sequence. In a more compact form the problem can be stated as :

$$\text{to find} \quad \min_{|u_N| \le 1} \{|u_N|\}, \tag{10}$$

where $d = d_0 + u_N z_N$, $d_0 \in R_{N-1}$

Minimization of (10) can be obtained by directly applying Theorem 2.5.1; the solution is unique as is shown in

Theorem 3.2.4. Let $f = d$, and $f_0 = d_0$, then $f_0 \in \partial R_{N-1}$. As proved below in Theorem 3.2.2, $f_0 \in \partial R_{N-1}$ has a unique control sequence, and thus the time-optimal control problem is completely solved.

Theorem 3.2.2 For a linear normal system a point $f \in \partial R_N$ corresponds to a unique control sequence $\vec{u}_N = [u_1, u_2, \ldots, u_N]$.

Proof Suppose f is on the hyperplane $c_i' x = 1$, i.e., $c_i' f = 1$. From the way c_i is constructed, there are n-1 vectors z_j, $j \in \mathcal{C}(\Lambda_i)$, such that $c_i' z_j = 0$; and there are N-n vectors z_m, $m \in \Lambda_i$, such that $c_i' z_m \neq 0$. For those z_m with $c_i' z_m \neq 0$, the corresponding control is such that $|u_m| = 1$ and $u_m c_i' z_m > 0$. Therefore these N-n controls with magnitude unity are determined and unique. On the other hand, when $c_i' z_j = 0$, $j \in \mathcal{C}(\Lambda_i)$, these controls u_j are not specified. However, any set of n vectors from $\{z_1, z_2, \ldots, z_N\}$ forms a linearly independent set in E^n. In particular, z_j, $j \in \mathcal{C}(\Lambda_i)$, is a linearly independent set. Since $\sum_{j \in \mathcal{C}(\Lambda_i)} z_j u_j + \sum_{m \in \Lambda_i} z_m u_m = d$, thus

$$\sum_{j \in \mathcal{C}(\Lambda_i)} z_j u_j = d - \sum_{m \in \Lambda_i} z_m u_m. \qquad (11)$$

Solution of (11), i.e., determination of u_j, $j \in \mathcal{C}(\Lambda_i)$ is guaranteed by the linear independence of z_j, $j \in \mathcal{C}(\Lambda_i)$, and the solution is unique. Therefore it is concluded that corresponding to a point on the boundary of R_N, the control sequence is unique and has at most (n-1) members with magnitude less than unity.

Corollary 3.2.3 A point $\tilde{f} \in \partial R_k$, $k = N-1, N-2, \ldots, n+1$, corresponds to a unique control seuqence $\vec{u}_k = [u_1, u_2, \ldots, u_k]$.

Theorem 3.2.4 The solution to (10) is unique.

Proof Two different situations will be considered.

(i) When $d = f \in \partial R_N$.

This is proved in Theorem 3.2.2.

(ii) When $d = f \notin \partial R_N$, but $d = f \in R_N$ and $d = f \notin R_{N-1}$.

Minimization of (10) is obtained by applying Theorem 2.5.1. Since the minimal u_N, u_N^* is unique for (10), then $d = f = f_0 + u_N^* z_N = d_0 + u_N^* z_N$, where $d_0 = f_0 \in \partial R_{N-1}$. By Corollary 3.2.3, the optimal control sequence is unique. This completes the proof.

CHAPTER 4 TERMINAL-ERROR REGULATOR PROBLEM

4.1 Statement of Terminal-Error Regulator Problem

The system considered is the one described in Section 2.2. It is desired to determine a control sequence $\vec{u}_N = [u_1, u_2, \ldots, u_N]$, $|u_i| \le 1$, $1 \le i \le N$, such that the scalar quantity

$$P = (d-x_N)'Q\,(d-x_N) \tag{1}$$

is minimum, where d is the desired terminal state of the system, Q is an $(n \times n)$ positive definite metrix, and the terminal time N is given in advance. It is well known [01] that any positive definite matrix can be decomposed into the product of the transpose of some $(n \times n)$ matrix W and itself, i.e., $Q = W'W$. By a change of variable, $w = Wx$, P becomes

$$P = (w_d-w_N)'(w_d-w_N), \tag{2}$$

where $w_d = Wd$ and $w_N = Wx_N$. Consequently, it is clear from (2) that no loss of generality will occur if the original matrix Q in (1) is assumed to be the identity matrix U_n. Since the reachable set R_N is compact, it is clear that there exists a unique point x_N such that (1) is minimum. However, there may be several control sequences which generate x_N. The solution properties and computing algorithms for generating x_N and some $\vec{u}_N$ are considered separately depending upon whether $n > N$ or $n \leq N$.

4.2 Solution Properties and Computing Algorithm for the Terminal-Error Regulator Problem; $n > N$

It is shown that the terminal-error regulator optimal control sequence must exist and is unique. Moreover, the procedures for finding the optimal control sequence are described thoroughly.

Given the terminal time N, let Z be defined as $Z = (z_1, z_2, \ldots, z_N)$, an $(n \times N)$ matrix whose columns consist of $z_1, z_2, \ldots, z_N$. By the controllability assumption of the system $\mathcal{L}$, the vectors $z_1, z_2, \ldots, z_N$ with $N \leq n$ are linearly independent, and therefore $r(Z) = N$. From Theorem 2.1.4, $Z'Z$ is non-singular and its inverse exists. Let $G = Z(Z'Z)^{-1}Z'$. For any vector $d \in E^n$, the vector Gd is a projection of d on Δ_N, which is the subspace spanned by $z_1, z_2, \ldots, z_N$. These results follow from Theorem 2.15 and Theorem 2.1.6. Cases $Gd \in R_N$ and $Gd \notin R_N$ are considered separately.

Case 1. If $Gd \in R_N$, then the optimal control sequence is unique. This is a consequence of the remark following Theorem 2.1.7.

Case 2. If $Gd \notin R_N$, the optimal control sequence is also unique, and a point $x_N \in R_N$ can be found such that $\|d-x_N\|^2$ is minimum. Since $\|d-x_N\|^2 = \|d-Gd\|^2 + \|Gd-x_N\|^2$, where $\|d-Gd\|^2$ is a fixed number for given G and d, thus $\|Gd-x_N\|^2$ is minimum implies that $\|d-x_N\|^2$ is minimum. The problem then becomes finding a point $x_N \in R_N$ such that $\|Gd-x_N\|^2$ is minimum.

As has been described, for both cases $Gd \in R_N$ and $Gd \notin R_N$ the optimal control sequence is unique. The procedures for finding the optimal control sequence are presented now.

Given the terminal time $N \le n$, and the desired terminal state d, then $Z = (z_1, z_2, \ldots, z_N)$ and $G = Z(Z'Z)^{-1}Z'$ are well-defined. To check whether $Gd \in R_N$ or not, Theorem 2.1.7 is employed. Apply elementary transformations to convert (Z,Gd) to an echelon matrix, then

$$L(Z,Gd) = \left[\begin{array}{c|c} L_1 Z & L_1 Gd \\ \hline L_2 Z & L_2 Gd \end{array}\right] = \left[\begin{array}{c|c} C & L_1 Gd \\ \hline 0 & L_2 Gd \end{array}\right] . \tag{3}$$

If $L_2 Gd = 0$ in (3), then $Gd \in R_N$. Also $C = L_1 Z = U$. The unique optimal control sequence is given by $(L_1 Z)^{-1} L_1 Gd = C^{-1} L_1 Gd = L_1 Gd$, where the i^{th} component of $L_1 Gd$ is u_i, $i = 1,2,\ldots,N$. On the other hand, if

$L_2Gd \neq 0$, then $Gd \notin R_N$. In this case, a different approach is taken to implement the optimal control sequence by using Gram-Schmidt orthogonalization process. As a consequence, the problem becomes an equivalent one where the dimension n is reduced to N.

Since $z_1, z_2, \ldots, z_N$ are linearly independent, they can be converted into an orthogonal set in E^n as is illustrated in the following:

$$\begin{cases} z_1 \to \tilde{z}_1 = z_1 \\ z_2 \to \tilde{z}_2 = a_{22}z_2 + a_{21}\tilde{z}_1 \\ z_3 \to \tilde{z}_3 = a_{33}z_3 + a_{32}\tilde{z}_2 + a_{31}\tilde{z}_1 \\ \cdot \\ \cdot \\ \cdot \\ z_N \to \tilde{z}_N = a_{N,N}z_N + a_{N,N-1}\tilde{z}_{N-1} + \cdots + a_{N,2}\tilde{z}_2 + \\ \qquad a_{N,1}\tilde{z}_1, \end{cases} \tag{4}$$

where $\tilde{z}_1, \tilde{z}_2, \ldots, \tilde{z}_N$ are mutually orthogonal. Clearly, other orthogonal sets are also available but will not affect later computations.

From (4) it follows directly that

$$\begin{cases} z_1 = \tilde{z}_1 \\ z_2 = -\dfrac{a_{21}}{a_{22}}\tilde{z}_1 + \dfrac{1}{a_{22}}\tilde{z}_2 \\ z_3 = -\dfrac{a_{31}}{a_{33}}\tilde{z}_1 - \dfrac{a_{32}}{a_{33}}\tilde{z}_2 + \dfrac{1}{a_{33}}\tilde{z}_3 \\ \cdot \\ \cdot \\ \cdot \\ z_N = -\dfrac{a_{N,1}}{a_{N,N}}\tilde{z}_1 - \dfrac{a_{N,2}}{a_{N,N}}\tilde{z}_2 - \cdots - \dfrac{a_{N,N-1}}{aN,N}\tilde{z}_{N-1} + \dfrac{1}{a_{N,N}}\tilde{z}_n . \end{cases} \tag{5}$$

Observe that $\tilde{z}_1,\tilde{z}_2,\ldots,\tilde{z}_N$ is a basis for the subspace spanned by $z_1,z_2,\ldots,z_N$. If $z_1,z_2,\ldots,z_N$ are expressed as linear combinations of $\tilde{z}_1,\tilde{z}_2,\ldots,\tilde{z}_N$ as in (5), then the coefficient matrix is $[y_1,y_2,\ldots,y_N]'$, where

$$\begin{cases} y_1' = (1,0,\ldots,0) \\ y_2' = (-\dfrac{a_{21}}{a_{22}}, \dfrac{1}{a_{22}}, 0, \ldots, 0) \\ \cdot \\ \cdot \\ \cdot \\ y_N' = (-\dfrac{a_{N,1}}{a_{N,N}}, -\dfrac{a_{N,2}}{a_{N,N}}, \ldots, -\dfrac{a_{N,N-1}}{a_{N,N}}, \dfrac{1}{a_{N,N}}) . \end{cases} \tag{6}$$

Using the correspondence between z_i and y_i, $i = 1,2,\ldots,N$, then the reachable set at time N from the origin, namely,

$$R_N = \{ x_N \in E^n : x_N = \sum_{i=1}^{N} z_i u_i, \ |u_i| \le 1, \ i=1,2,\ldots,N\}, \tag{7}$$

can be converted to

$$\tilde{R}_N = \{ \tilde{x}_N \in E^N : \tilde{x}_N = \sum_{i=1}^{N} y_i u_i, \ |u_i| \le 1, \ i = 1,2,\ldots,N\}. \qquad (8)$$

Now a similar expression can be found for Gd. Applying Theorem 2.1.5,

$$Gd = Z(Z'Z)^{-1}Z'd = (z_1, z_2, \ldots, z_N)(Z'Z)^{-1}Z'd$$

$$= \tilde{d}_1 z_1 + \tilde{d}_2 z_2 + \cdots + \tilde{d}_N z_N, \qquad (9)$$

where $\tilde{d}_i$ = the i^{th} component of $(Z'Z)^{-1}Z'd$. After some simple algebraic manipulations, Gd can be shown to correspond to

$$y_d = d_1'' y_1 + d_2'' y_2 + \cdots + d_N'' y_N. \qquad (10)$$

Note that the dimension of y_i's in (6) is N. Then the problem becomes: given $y_1, y_2, \ldots, y_N$, and the terminal state y_d, find a point $\tilde{x}_N \in \tilde{R}_N$ and the corresponding control sequence such that $\|y_d - \tilde{x}_N\|^2$ = minimum. This equivalent problem has $n = N$ and thus can be solved by the procedure of the next section.

4.3 Solution Properties and Computing Algorithm for the Terminal-Error Regulator Problem; $n \le N$

Now the case when $n \le N$ is considered. In the case where the desired terminal state d lies outside the reachable set R_N at time N, then the optimal control sequence is unique and $P > 0$. When the desired terminal state d is an element of R_N, it is clear that $x_N = d$,

$P = 0$, and the optimal control sequence is, in general, not unique.

Case 1. $P = 0$.

The determination of the smallest m such that $d \in R_m$ and $d \notin R_{m-1}$ is the time-optimal control problem which can be solved by the algorithm in Section 3.2. Therefore an effective procedure for the terminal-error regulator problem with $n \leq N$ can begin by using the algorithm of Section 3.2 subject to $m \leq N$. If $d \notin R_N$, then $P > 0$ and Case 2 is used. If $d \in R_N$, then $P = 0$ and the time-optimal solution m satisfies $m \leq N$. If $m = N$, then the time-optimal control sequence $\vec{u}_m$ equals $\vec{u}_N$, the terminal-error regulator optimal control sequence. If $m < N$, then $\vec{u}_N = [u_1, u_2, \ldots, u_m, 0, 0, \ldots, 0]$, i.e., the last $(N-m)$ members of $\vec{u}_N$ are zero.

Case 2. $P > 0$.

Clearly $P > 0$ if and only if $d \notin R_N$. Since $d \notin R_N$, a unique point $x_N \in \partial R_N$ has to be found such that $\|d-x_N\|^2$ is minimum. From Theorem 2.4.6, $R_N = \Omega_1 \cap \Omega_2 \cap \cdots \cap \Omega_K$, and thus the boundary of R_N is contained in the hyperplane specified by $c_i' x = \delta$, $i = 1,2,\ldots,K$ where $\delta = \pm 1$ and $K \leq \binom{N}{n-1}$. Obviously, x_N is an extreme point (i.e., vertex) of R_N or is contained in the intersection of at most $(n-1)$ hyperplanes: $c_i' x = \delta$. The problem of finding

$x_N \in R_N$ such that $\|d-x_N\|^2$ is minimum can be solved in three parts.

1. Verification of vertex of R_N

Recall that v is a vertex of R_N if i) $v = \sum_{i=1}^{N} \alpha_i z_i$, $|\alpha_i| = 1$, $i = 1,2,\ldots,N$ and ii) there exists at least one support hyperplane to R_N with v as an intersection point. By Theorem 2.4.6 $\forall\ x \in R_N$ it is true that $|c_i' x| \leq 1$, $i = 1,2,\ldots,K$, where $K \leq \binom{N}{n-1}$ and $c_i' x = \delta$ are the boundary hyperplanes as defined previously in Section 2.4. It is obvious that if $v = \sum_{i=1}^{N} \alpha_i z_i$ satisfies i) and $|c_j' v| = 1$ for some j, then v is a vertex of R_N.

2. Determination of a closest vertex of R_N to d

The procedure described below permits finding one of the closest vertices of R_N to d. It can be observed that there may be several different vertices of R_N with the smallest equal distance to d. Let $v^{(1)} = \sum_{i=1}^{N} \alpha_i^{(1)} z_i$ be any vertex of R_N. By changing signs of some of the controls $\alpha_i^{(1)}$, a new vertex $\tilde{v}^{(1)} = \sum_{i=1}^{N} \tilde{\alpha}_i^{(1)} z_i$ can be found. Compute $\|d-v^{(1)}\|^2$ and $\|d-\tilde{v}^{(1)}\|^2$. If $\|d-v^{(1)}\|^2 \leq \|d-\tilde{v}^{(1)}\|^2$, then set $v^{(2)} = v^{(1)}$; if $\|d-v^{(1)}\|^2 > \|d-\tilde{v}^{(1)}\|^2$, then set $v^{(2)} = \tilde{v}^{(1)}$. Thus it is clear that $\|d-v^{(1)}\|^2 \geq \|d-v^{(2)}\|^2$. In a similar manner

$\tilde{v}^{(2)}, v^{(3)}, \tilde{v}^{(3)}, v^{(4)}, \ldots,$ etc. can be found and they satisfy $\|d-v^{(2)}\|^2 \geq \|d-v^{(3)}\|^2 \geq \|d-v^{(4)}\|^2 \geq \ldots,$ etc. Since vertices of R_N are finite in number, it is evident that the process will terminate in a finite number of steps and determine one of the closest vertices of R_N to d.

3. Determination of the optimal u_i's satisfying $|u_i| \leq 1$

Observe that the point $x_N \in R_N$ satisfying $\|d-x_N\|^2 =$ minimum is contained in a zero-dimensional face (i.e., vertex), one-dimensional face (i.e., edge), two-dimensional face, ..., or (n-1)-dimensional face of R_N. Consider $F_d = d-v$, where v is determined in part 2). Let $J = \{j : (F_d, \alpha_j z_j) < 0\} = \{j : (v, \alpha_j z_j) > (d, \alpha_j z_j)\}$. The following theorem is useful in identifying which members of the control sequence obtained in part 2) can be modified to yield smaller length of $d - x$ for some $x \in R_N$.

<u>Theorem 4.3.1</u> Let $v = \sum_{i=1}^{N} \alpha_i z_i$, $|\alpha_i| = 1$ be a vertex of R_N which is one of the closest to d. Let d be the desired terminal target point and $d \notin R_N$. If there is a k such that $(d-v, \alpha_k z_k) < 0$, then there exists a point $x \in R_N$ such that $\|d-x\| < \|d-v\|$.

<u>Proof</u> Let $x = \sum_{\substack{i=1 \\ i \neq k}}^{N} \alpha_i z_i + \tilde{\alpha}_k z_k$, where $\tilde{\alpha}_k \neq \alpha_k$. Then the squared lengths of $d-v$ and $d-x$ can be computed:

$$\|d-v\|^2 = \|d\|^2 + \|\sum_{\substack{i=1\\i\neq k}}^{N} \alpha_i z_i\|^2 - 2(d,\alpha_k z_k) - 2(d, \sum_{\substack{i=1\\i\neq k}}^{N} \alpha_i z_i)$$

$$+ \|\alpha_k z_k\|^2 + 2(\alpha_k z_k, \sum_{\substack{i=1\\i\neq k}}^{N} \alpha_i z_i) \tag{11}$$

$$\|d-x\|^2 = \|d\|^2 + \|\sum_{\substack{i=1\\i\neq k}}^{N} \alpha_i z_i\|^2 - 2(d,\tilde{\alpha}_k z_k) - 2(d, \sum_{\substack{i=1\\i\neq k}}^{N} \tilde{\alpha}_i z_i)$$

$$+ \|\tilde{\alpha}_k z_k\|^2 + 2(\tilde{\alpha}_k z_k, \sum_{\substack{i=1\\i\neq k}}^{N} \alpha_i z_i). \tag{12}$$

To compare $\|d-v\|^2$ and $\|d-x\|^2$, their difference is taken:

$$\|d-v\|^2 - \|d-x\|^2 = \|\alpha_k z_k\|^2 - \|\tilde{\alpha}_k z_k\|^2$$

$$+ 2(\alpha_k z_k - \tilde{\alpha}_k z_k, \sum_{\substack{i=1\\i\neq k}}^{N} \alpha_i z_i - d) \tag{13}$$

$$= \|\alpha_k z_k\|^2 - \|\tilde{\alpha}_k z_k\|^2$$

$$+ 2(\alpha_k z_k - \tilde{\alpha}_k z_k, v-d - \alpha_k z_k). \tag{14}$$

Equation (14) can be examined more closely by assigning either $\alpha_k = +1$ or $\alpha_k = -1$. Consider $\alpha_k = +1$. It is clear that $x \in R_N$ if $\tilde{\alpha}_k = 1 - \epsilon$, $0 < \epsilon \leq 2$. If this value is substituted for $\tilde{\alpha}_k$ in (14), then

$$\|d-v\|^2 - \|d-x\|^2 = 2(z_k, v-d) - \epsilon \|z_k\|^2. \tag{15}$$

Since $(z_k, v-d) > 0$ by hypothesis, there always exists an

ϵ such that $\|d-x\|^2 < \|d-v\|^2$. Clearly $\tilde{\alpha}_k \in [-1,1)$. For $\alpha_k = -1$, it is also true that there exists at least one $\tilde{\alpha}_k \in (-1,1]$ such that $\|d-x\|^2 < \|d-v\|^2$. This completes the proof.

As demonstrated in Figure 7, $\tilde{\alpha}_k$ can be taken to be $\lambda\alpha_k$, $0 < \lambda < 1$, to improve the minimum squared length of F_d.

It follows from the previous theorem that if J is empty, then the v found in part 1) is x_N, which is the intersection of some n boundary hyperplanes $c_i' x = \delta$ of R_N, and clearly the optimal control sequence is given by $[\alpha_1, \alpha_2, \ldots, \alpha_N]$. If J has one element, then x_N is the intersection of some $(n-1)$ boundary hyperplanes of R_N. In general, if J has k elements, then x_N is the intersection of $(n-k)$ boundary hyperplanes of R_N.

Now it is shown that J can have at most $(n-1)$ elements. Obviously the closest vertex v to d is contained in some $(n-1)$-dimensional support hyperplane, say, $c_i' x = \delta$. Since $c_i' x = \delta$ is a support hyperplane, then from Theorem 2.4.3 u_j satisfies $u_j c_i' z_j \geq 0$ for $\delta = 1$ and $u_j c_i' z_j \leq 0$ for $\delta = -1$, $j = 1,2,\ldots,N$. But $c_i' z_j = 0$ for exactly $(n-1)$ z_j's (see Section 2.4) and $c_i' z_j \neq 0$ for $(N-n+1)$ z_j's. Thus those controls u_j corresponding to $u_j c_i' z_j \neq 0$ are fixed at ± 1 and the $(n-1)$ controls u_j corresponding to $c_i' z_j = 0$ are not specified. If there are more than n elements in J then this implies that the v found in part 1) is not on the hyperplane $c_i' x = \delta$. This is a contradiction. Hence it can be concluded that J has at most (n-1) elements.

Let $j_1, j_2, \ldots, j_k, k \leq n-1$, denote the elements of J. Then the original problem becomes determination of $u_{j_1}, u_{j_2}, \ldots, u_{j_k}$ such that $\| d - \sum_{\substack{i=1 \\ i \notin J}}^{N} \alpha_i z_i - \sum_{j_i \in J} u_{j_i} z_{j_i} \|^2 = \text{minimum}$, subject to the admissibility conditions on u_{j_i}, i.e., $|u_{j_i}| \leq 1$ for all j_i. This problem can be solved by using the algorithm in Section 4.2, with the new desired terminal state replaced by $d - \sum_{\substack{i=1 \\ i \notin J}}^{N} \alpha_i z_i$.

To summarize, the following is sketched:

Given the terminal time N, with $N \geq n$, and the desired terminal state d, if $d \in R_N$, then the control is, in general, not unique. In order to have a unique optimal control sequence, an additional requirement that u_N be minimum in absolute value is associated with the control sequence, as has been solved in Section 3.2. If $d \notin R_N$, then the scalar quantity

$$P = (d-x_N)'(d-x_N) > 0$$

and the optimal control sequence is unique. In this case

x_N is on the boundary of R_N. It has been shown that x_N may be on one of the boundary hyperplanes or at the intersection of more than one boundary hyperplanes of R_N. A method is presented to find the unique point $x_N \in R_N$ and the corresponding optimal control, which avoids solving the quadratic programming problem. The computer algorithm for the terminal-error regulator problem is illustrated in Figure 8.

CHAPTER 5 SUMMARY AND EXTENSIONS

Some important features of the computing procedures described in Chapters 2, 3, and 4 are summarized in Section 5.1. In Section 5.2 certain possible extensions of these results are depicted.

5.1 Summary

This section contains brief summaries of the computing procedures developed in Chapters 2, 3, and 4 and some of the more prominent properties of the algorithms. In adiition, a refinement which is useful for calculating the vectors c_i, which describe the boundary hyperplanes of R_N, is included in part (iv). This often saves computational time for both the time-optimal and terminal-error regulator problems.

(i) Determining whether $d \in R_N$ or $d \notin R_N$.

In this dissertation a simple algorithm is proposed for determining whether a given point $d \in E^n$ is in R_N. Depending upon whether $n \geq N$ or $n < N$, two different algorithms are considered.

<u>Case 1</u>. $n \geq N$.

Let $Z = (z_1, z_2, \ldots, z_N)$ be an $(n \times N)$ matrix. By performing row operations (see Section 3.2, Phases 1 and 2),

(Z,d) is transformed to an echelon matrix:

$$L(Z,d) = \left[\begin{array}{c|c} L_1 Z & L_1 d \\ \hline L_2 Z & L_2 d \end{array}\right] = \left[\begin{array}{c|c} C & L_1 d \\ \hline 0 & L_2 d \end{array}\right] ,$$

where $L = \left[\begin{array}{c} L_1 \\ \hline L_2 \end{array}\right]$ is the product of elementary matrices. If $L_2 d = 0$, and furthermore $u = L_1 d$ is admissible, then $d \in R_N$. If $L_2 d \neq 0$ or if $L_2 d = 0$ and $u = L_1 d$ is not admissible, then $d \notin R_N$. When $L_2 d = 0$ and $u = L_1 d$ is not admissible, it is also true that $d \notin R_n$.

Case 2. $n < N$.

From Theorem 2.4.6, the boundary of R_N consists of subsets of $2K$ hyperplanes of the form $c_i' x = \delta$, $\delta = \pm 1$. The $(n \times 1)$ vector c_i can be determined by solving $K \leq \binom{N}{n-1}$ systems of $(n-1)$ simultaneous equations in $(n-1)$ unknowns (see Section 2.4). Then $d \in R_N$ if and only if $|c_i' d| \leq 1$, $i = 1,2,\ldots,K$.

(ii) Properties of the Algorithm for Time-Optimal Control Problem

The calculations in (i) can be performed sequentially for $N = 1,2,\ldots,$ etc. to determine the smallest integer N such that $d \in R_N$. This N is the optimal time. If $n \geq N$, the time-optimal control sequence exists and is unique.

Moreover, it is given by L_1d, where the i^{th} component of L_1d is u_i, $i = 1,2,\ldots,N$. If $n < N$, the time-optimal control sequence either is unique or has an infinite number of solutions. If the control sequence is unique, then there are at most (n-1) controls with magnitude less than unity. If the control sequence is not unique, an additional requirement that u_N be minimum in absolute value is imposed. In this case, the resulting optimal control sequence is unique and has at most n components with magnitude less than unity.

(iii) Properties of the Algorithm for Terminal-Error Regulator Control Problem

In the terminal-error regulator problem the terminal time N and the desired terminal state d are given. The optimal control sequence is the one which minimizes the scalar quantity: $P = \|d-x_N\|^2$, where $x_N \in R_N$. Two cases are considered for $n > N$ and $n \leq N$.

Case 1. $n > N$.

If $n > N$, the optimal control sequence for the terminal-error regulator problem is shown to be unique. Furthermore, it is given by L_1Gd if $Gd \in R_N$, where $G = Z(Z'Z)^{-1}Z'$, and $Z = (z_1,z_2,\ldots,z_N)$. If $Gd \notin R_N$, then the optimal control sequence can be found by employing the Gram-Schmidt orthogonalization process to reduce the problem to an equivalent one with lower dimension. Then the method of Case 2 is used.

Case 2. $n \leq N$.

For $n \leq N$, if $d \in \partial R_N$ then the optimal control sequence is unique. If $d \in R_N$ but $d \notin \partial R_N$, then $P = 0$ and the problem becomes the time-optimal control problem considered in Section 3.2, Phase 3 and Phase 4. If $d \notin R_N$, then $P > 0$ and the optimal control sequence is unique. The algorithm which is proposed requires computation of the c_i which describe the boundary hyperplanes of R_N, but avoids solving the corresponding quadratic programming problem.

(iv) Construction of Boundary Hyperplanes for R_{N+1} from Those of R_N

In part (i) case 2 and part (iii) case 2 it is required to calculate the c_i which describe the boundary hyperplanes of R_N. The boundary hyperplanes of R_{N+1} are also needed frequently and computational time can be saved by computing them indirectly from those of R_N. It may be recalled that the direct computation required solving $\binom{N+1}{n-1}$ systems of (n-1) simultaneous equations in (n-1) unknowns. All the subsets of the hyperplanes which constitute the boundary of R_N will be subsets of the boundary hyperplanes of R_{N+1} by a simple translation:

If subset of $|c_i' x| = 1$ is one of the subsets of the boundary hyperplanes of R_N, then

$$|\tilde{c}_i' x| = |\frac{c_i'}{1+|\xi|} x| = 1$$

is the corresponding subset of the boundary hyperplanes of R_{N+1}, where $\xi = c_i' z_{N+1}$.

The new subsets of the boundary hyperplanes of R_{N+1}, i.e., those which cannot be obtained by a translation from subsets of the boundary hyperplanes of R_N, can be calculated by (n-1) vectors, which consist of z_{N+1} and (n-2) vectors from $\{z_1, z_2, \ldots, z_N\}$. Thus the total number of subsets of the boundary hyperplanes of R_{N+1} is $2[\binom{N}{n-2} + \binom{N}{n-1}] = 2\binom{N+1}{n-1}$. It is clear that at each step from R_N to R_{N+1}, subsets of the boundary hyperplanes of R_{N+1} can be obtained by solving only $\binom{N}{n-2}$ systems of (n-1) simultaneous equations in (n-1) unknowns and $\binom{N}{n-1}$ simple translations.

5.2 Extensions

In chapters 2, 3, and 4 two basic control problems are examined. There are a number of extensions which can be considered. The extension to a higher-dimensional control signal is of primary importance. Other various extensions are briefly investigated.

(i) Alternate Control Constraints Given by $|u_i| \leq \eta_i$

The actual constraints on the control signal, $|u_i| \leq 1$,

$i = 1,2,\ldots,N$, have been fully utilized to construct the boundary of R_N. When the constraints are given in the form, $|u_i| \le \eta_i$, $i = 1,2,\ldots,N$, the construction of the boundary of R_N presents no difficulty at all. In fact, if z_i is replaced by $\eta_i z_i$ for all i, then the method presented applies without any modifications.

(ii) Target Sets Given in the Form of $x'Sx \ge \beta$

Let the target set $x'Sx \ge \beta$ be given, where S is an $(n \times n)$ symmetric positive definite matrix and $\beta > 0$. By the well-known result of matrix theory [01], S can be written as $S = W'W$. By a change of variable, $w = Wx$, the form of the target set can be rewritten as: $w'w = \|w\| \ge \beta$. Hence without loss of generality S can be assumed to be an identity matrix. If $x(0) = x_0 = 0$, the time-optimal control problem is to find a point $x_N \in R_N$ such that $\|x_N\|^2 \ge \beta$ and to find the corresponding control sequence. Because of the symmetry of R_N and the convexity of the set $x'Sx \le \beta$, it is clear that when N is the minimum time the sets R_N and $x'Sx \ge \beta$ will intersect at a finite number of vertices or at an infinite number of points (see Figure 9). In any case controllability implies that there exist N and a vertex $v \in R_N$ such that v and $-v$ are solutions. Thus to get a solution, the vertices of R_N can be computed using part 1) of Section 4.3 for each choice of $N = 1,2,\ldots,$

etc. The control sequence is easily implemented (see Theorem 3.2.2).

(iii) Target Sets Which Are Time-Varying

Consider target sets $\mathcal{S}(i) : x'S(i)x \geq \beta_i$, $i = 1, 2, \ldots,$ etc. In this case the analysis in part (ii) is modified to consider for each N the set $\mathcal{S}(N)$, etc. Figure 10 illustrates an example where the optimal time N is 4.

(iv) Time-Varying Linear Discrete Systems

It is assumed that the system to be considered satisfies the following difference equation:

$$x(i + 1) = \Phi(i)x(i) + b(i)u(i).$$

Note that $x(i)$ can be written as

$$x(i) = \Psi(i)x(0) + \Psi(i) \sum_{j=1}^{i} \Psi(j)b(j-1)u(j-1),$$

where $\Psi(i)$ is an $(n \times n)$ matrix with $\Psi(i+1) = \Phi(i)\Psi(i)$, $\Psi(0) = U$. Also note that $\Psi(i) = \Phi(i-1)\cdot\Phi(i-2)\cdot\cdots\cdot\Phi(0)$ and

$$\Psi(i)\Psi^{-1}(j) = \Phi(i-1)\cdot\Phi(i-2)\cdots\Phi(0)\cdot\Phi^{-1}(0)\cdot\Phi^{-1}(1)\cdots\Phi^{-1}(j-1)$$
$$= \Phi(i-1)\cdot\Phi(i-2)\cdot\cdots\cdot\Phi(j) \quad \text{for} \quad i > j.$$

By defining $z_i = \Psi(N)\Psi^{-1}(i)b(i-1)$, $i = 1,2,\ldots,N$, $u_i = u(i-1)$, $i = 1,2,\ldots,N$, $x_i = x(i)$, $i = 0,1,\ldots,N$, then $x_N = \Psi(N)x_0 + z_1u_1 + z_2u_2 + \cdots\cdots + z_Nu_N$. The time-optimal control problem is to find a point $x_N \in R_N$ and a corresponding control sequence such that $x_N = d$, the desired terminal state. This problem is essentially the same as the one considered except now z_i's, $i = 1,2,\ldots,N$ are time-varying and for each N, z_i's have to be recalculated. The same results apply to this time-varying system.

(v) Systems with Variable Sampling Instants

Consider the linear system governed by the differential equation: $x(t) = \Phi x(t) + bu(t)$. The solution is given by

$$x(t_N) = \Psi(t_N)\,[\,x(0) + \int_0^{t_N} \Psi(-s)bu(s)ds\,]$$

$$= \Psi(t_N)\,[\,x(0) + \sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \Psi(-s)dsbu(i)\,] \quad ,$$

where $\Psi(t_N) = e^{\Phi t_N}$, $u(t) = u(i)$, $t_i \leq t < t_{i+1}$.

Let $z_i = \Psi(t_N) \int_{t_{i-1}}^{t_i} \Psi(-s)dsb$ and $u_i = u(i-1)$, $i = 1,2,\ldots,N$.

Then

$$x(t_N) = \Psi(t_N)x(0) + \sum_{i=1}^{N} z_i u_i . \tag{1}$$

Thus (1) is the same equation as (11) of Section 2.2 with Φ^N replaced by $\Psi(t_N)$. It is clear that the algorithms can be applied to systems with variable sampling instants. All the z_i and $\Psi(t_N)$ have to be recalculated for each N.

(vi) Linear Discrete Systems with the Initial State $x_0 \neq 0$

If $x_0 \neq 0$ and d is given, translation of coordinates such that x_0 becomes the new origin yields the formulation of Chapter 2, and all the algorithms apply (see Figure 11).

(vii) Multi-Input Control Systems

For multi-input control system, the construction of R_N presents no real difficulty, although it is more complicated than in the single-input system. However, further constraints are necessary to insure that the control sequence is unique. The following considerations are important.

Let the multi-input control system be governed by the following difference equation:

$$x((i+1)T) = \Phi x(iT) + Bu(iT), \qquad (2)$$

where $B = (n \times m)$ constant control matrix with columns $b_1, b_2, \ldots, b_m$, where $m \leq n$, $u(iT) = (m \times 1)$ control vector with components $u^1(iT), u^2(iT), \ldots, u^m(iT)$, and i, T, Φ, x(iT) are the same as previously defined. The control has the following pre-determined properties:

$$u^j(t) = \text{constant for } j = 1,2,\ldots,m \quad iT \leq t < (i+1)T \qquad (3)$$

$$|u^j(iT)| \leq 1 \text{ for all } i \text{ and for } j = 1,2,\ldots,m. \qquad (4)$$

It is assumed that the system is completely controllable [K2]. The system is completely controllable if and only if $r[\Phi^{n-1}B, \Phi^{n-2}B, \ldots, \Phi B, B] = n$ [B2]. By iteration on (1), x(NT) is given by

$$x(NT) = \Phi^{N-1}Bu(0) + \Phi^{N-2}Bu(T) + \cdots + \Phi Bu((N-2)T) + Bu((N-1)T) \qquad (5)$$

or in matrix form

$$x(NT) = [\Phi^{N-1}B, \Phi^{N-2}B, \ldots, \Phi B, B] \begin{bmatrix} u(0) \\ u(T) \\ \cdot \\ \cdot \\ \cdot \\ \cdot \\ u((N-2)T) \\ u((N-1)T) \end{bmatrix} . \qquad (6)$$

By letting $x_N = x(NT)$, $z_i^j = \Phi^{i-1}b_j$, for $i = 1,2,\ldots,N$, $j = 1,2,\ldots,m$ and $u_i^j = u^j((N-i)T)$ for $i = 1,2,\ldots,N$, $j = 1,2,\ldots,m$, (4) becomes

$$x_N = u_1^1 z_1^1 + u_1^2 z_1^2 + \cdots + u_1^m z_1^m + u_2^m z_2^m + \cdots + u_N^1 z_N^1 + \cdots + u_N^m z_N^m. \qquad (7)$$

<u>Definition</u> The <u>reachable subset</u> from the origin at time N is defined as

$$\Theta_N^\gamma = \{x \in E^n : x = \sum_{i=1}^{N-1} \sum_{j=1}^{m} u_i^j z_i^j + \sum_{k=1}^{\gamma} u_N^k z_N^k, \quad |u_i^j| \le 1 \ \forall \ i \text{ and } j\},$$

where $1 \le \gamma \le m$ and $N \ge 1$. Clearly $R_N = \Theta_{N-1}^m$.

It is clear from the above definition that given N and γ, $1 \le \gamma \le m$, the construction of Θ_N^γ does not present any difficulty.

In the time-optimal control problem, for any desired terminal state d with $\|d\| < \infty$, there exist γ and N such that $d \in \Theta_N^{\gamma}$ and $d \in \Theta_N^{\gamma-1}$, where $2 \leq \gamma \leq m$. If $\gamma = m$, and $d \in \partial\Theta_N^{\gamma}$ then the time-optimal control sequence is unique. If $\gamma < m$, or $\gamma = m$ but $d \notin \partial\Theta_N^{m}$ then the optimal control is, in general, not unique. In this case, the time-optimal control sequence is defined as the one with $|u_N^{\gamma}|$ = minimum and $u_N^j = 0$, $j = \gamma + 1, \gamma + 2, \ldots, m$.

All the above extensions require only slight modifications of the algorithms in Chapters 2, 3, and 4. For non-symmetrical constraints on the control or target sets which are non-symmetrical considerable modifications are required and they are not treated here. Another problem in which it is difficult to extend the algorithms is: initial state $x_0 \neq 0$, target set $\mathcal{S}$: $x'Sx \leq \beta$. Several configurations which may arise in this problem are illustrated in Figure 12.

CHAPTER 6 NUMERICAL EXAMPLES AND CONCLUSIONS

The study of control of a system with a discrete model is of value in its own right because of the close relation of such models with various physical, biological, and socio-economic processes [K5]. In some cases, a physical system is modeled by an ordinary differential equation whose solution is assumed to approximate the actual evolution of the system. In general, the continuous system with differential equation model is solved on a digital computer, which is in fact a process of discrete approximation to the continuous problem. In the following examples, the first two are given as discrete models and the remaining four as continuous models. The continuous ones are approximated by a corresponding discrete version for computational purposes.

Consider the system governed by the vector differential equation:

$$\dot{x}(t) = A\,x(t) + D\,u(t), \qquad (1)$$

where A is an $(n \times n)$ constant transition matrix and D is an $(n \times m)$ control matrix.

The solution to the vector differential equation (1) is

$$x(t) = \psi(t)\,[\,x(0) + \int_0^t \psi(-\tau) D\,u(\tau)\,d\tau\,]\,, \qquad (2)$$

where $\psi(t) = e^{At}$ is the fundamental matrix, and satisfies $\dot{\psi}(t) = A\psi(t)$, and $\psi(0) = U$, the identity matrix.

If the control signal satisfies

$$u(t) = u(i) = \text{constant} \qquad iT \le t < (i+1)T \tag{3}$$

and k is a positive integer, then

$$\begin{aligned} x((k+1)T) &= \psi((k+1)T)[x(0) + \int_0^{(k+1)T} \psi(-\tau)Du(\tau)d\tau] \\ &= \psi((k+1)T)[x(0) + \sum_{i=1}^{k+1} \int_{(i-1)T}^{iT} \psi(-\tau)D\, d\tau u(i-1)] \, . \end{aligned} \tag{4}$$

Let $k = k-1$. Then from (4)

$$x(kT) = \psi(kT)[x(0) + \sum_{i=1}^{k} \int_{(i-1)T}^{iT} \psi(-\tau)D\, d\tau u(i-1)] \, . \tag{5}$$

It is clear that (5) can be solved for $x(0)$. Thus

$$x(0) = \psi^{-1}(kT)[x(kT) - \sum_{i=1}^{k} \int_{(i-1)T}^{iT} \psi(-\tau)D\, d\tau u(i-1)] \, . \tag{6}$$

Substitute (6) into (4), then

$$x((k+1)T) = \psi((k+1)T)[\psi^{-1}(kT)x(kT) + \int_{kT}^{(k+1)T} \psi(-\tau)Dd\tau u(k)] \tag{7}$$

$$= \psi((k+1)T)\psi^{-1}(kT)x(kT) + \psi((k+1)T)\int_{kT}^{(K+1)T} \psi(-\tau)Dd\tau u(k) . \tag{8}$$

By change of variables, $t = \tau - kT$, (8) gives

$$\begin{aligned} x((k+1)T) &= \psi(T)x(kT) + \psi((k+1)T)\int_0^T \psi(-t-kT)D\, dt u(k) \\ &= \psi(T)x(kT) + \psi(T)\int_0^T \psi(-t)D\, dt u(k) . \end{aligned} \tag{9}$$

Let $x(k+1) = x((k+1)T)$, $x(k) = x(kT)$ and $B = \psi(T)\int_0^T \psi(-t)D dt$.

Then (9) can be written as

$$x(k+1) = \psi(T)\, x(k) + B\, u(k). \tag{10}$$

Now it is clear that (10) is an appropriate system equation which has been described in Section 2.2. Several examples are presented in the following to illustrate the application of the suggested techniques.

<u>Example 6.1</u> Given $z_1' = (1,2,0)$, $z_2' = (1,2,1)$, $z_3' = (3,4,1)$, $z_4' = (1,1,3)$, $z_5' = (1,3,7)$, and $d' = (-3.5,-6.0,1.0)$, find the minimal time N such that $d \in R_N$ and the corresponding control sequence. If the control sequence is not unique, find the unique control sequence for which $|u_N|$ is a minimum.

<u>Solution</u>: Inspection of z_1, z_2, and d clearly implies $d \notin R_1$ and $d \notin R_2$. Let $N = 3$. The coefficients of the boundary hyperplanes of R_3 are calculated and listed in Table 6.1.

Table 6.1 Boundary Hyperplane Coefficients of R_3

No.i	Subscripts of z's determining the coefficients	c_i'
1	1,2	1.0, -0.5, 0.0
2	1,3	1.0, -0.5, -1.0
3	2,3	1.0, -1.0, 1.0

Now $c_1'd = -0.5$, $c_2'd = -1.5 < -1$. Since $|c_2' d| > 1$, hence $d \notin R_3$. Let $N = 4$. The coefficients of the boundary hyperplanes of R_4 are calculated and listed in Table 6.2.

Table 6.2 Boundary Hyperplane Coefficients of R_4

No. i	Subscripts of z's determining the coefficients	c_i'		
1	1, 2	0.66667,	-0.33333,	0.0
2	1, 3	0.28571,	-0.14286,	-0.28571
3	1, 4	1.0 ,	-0.5 ,	-0.16667
4	2, 3	0.25 ,	-0.25 ,	0.25
5	2, 4	0.71429,	-0.28751,	-0.14286
6	3, 4	1.0 ,	-0.72727,	-0.09091

$|c_i' d|$, $i = 1,2,\ldots,6$, can be calculated:

$|c_1' d| = 0.86667$, $|c_2' d| = 0.97143$, $|c_3' d| = 0.51667$,

$|c_4' d| = 0.5$, $|c_5' d| = 0.82571$, $|c_6' d| = 0.94546$.

Since $|c_i' d| < 1$ for $i = 1,2,\ldots,6$, thus it is clear that $d \in R_4$. The corresponding control is found to be [-1,0, -0.5833, -0.9167, 0.8333], where $|0.8333|$ is a minimum.

<u>Example 6.2</u>

Let the difference equation [K5] $x(i+1) = \begin{bmatrix} 1 & 0 & 3 \\ 1 & 2 & -1 \\ 0 & 1 & 4 \end{bmatrix} x(i) + \begin{bmatrix} 1 & 0 \\ 0 & 2 \\ 2 & 1 \end{bmatrix} \begin{bmatrix} u^1(i) \\ u^2(i) \end{bmatrix}$ be given, with $|u^1(i)| \leq 1$ and $|u^2(i)| \leq 1$ for all i, and the initial state be 0, the terminal state be $d' = (2682.7, 154.1, 6020.2)$. The problem is to find the minimal integer N and the corresponding control sequence such that $d \in \Theta^{\gamma}_{N-1}$, $\gamma = 1$ or 2, and $|u^{\gamma}_{N-1}|$ = minimum.

Solution It is clear that z^1_1, z^2_1, z^1_2, z^2_2, ..., can be calculated fairly easily by recursion relations (see part (vii) of Section 5.2). The computer output shows that $d \notin R_j$, $1 \le j \le 14$. Some of the numbers $|c_i' d|$ are shown in Table 6.3, where c_i's correspond to the boundary hyperplane coefficients of Θ^1_{14}.

Table 6.3 Some $|c_i' d|$ for Θ^1_{14}

i	$\|c_i' d\|$
1	0.000000706
2	0.000000330
3	0.000000800
4	0.000001146
5	0.000001742
6	0.000003187
7	0.017966655
	⋮
	⋮
401	0.55360002
402	0.32617638
403	0.77601498
404	0.55480775
405	0.12834193
406	0.79181064

From Table 6.3, $|c_i' d| \le 1$ for all i. Thus $d \in \Theta^1_{14}$. The optimal control sequence is found to be $[u^1_1, u^2_1, u^1_2, \ldots, u^2_{14}]$ = [1.0, -1.0, 1.0, -1.0, -1.0, -1.0, -1.0, -1.0, 1.0, -1.0, 1.0, -1.0, 1.0, -1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, -1.0, 1.0,

-1.0, 1.0, -1.0, 1.0, 0.037, 0.848, -0.2088, 0.0].

Example 6.3

Given (1) The system equation [H11] $\dot{x}(t) = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 2\omega \\ 0 & 0 & 0 & 1 \\ 0 & -2\omega & 3\omega^2 & 0 \end{bmatrix} x(t)$

$+ \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} u^1(t) \\ u^2(t) \end{bmatrix}$, where $\omega = 0.00111$ ($^{rad}/_{sec}$).

(2) Constant sampling periods of 10 seconds.

(3) The control constraints are $|u^1(t)| \le 1$ and $|u^2(t)| \le 1$.

(4) The initial state $x_0' = (0\ 0\ 0\ 0)$.

(5) Desired terminal state $d' = (1402.5, 44.5, 149.8, -8.8)$.

Find the smallest integer N and the corresponding optimal control sequence satisfying $d \in \Theta_{N-1}^{\gamma}$, $\gamma = 1$ or 2, and $|u_{N-1}^{\gamma}| =$ minimum.

Solution The continuous system equation is approximated by equation (10). Furthermore, the fundamental matrix $\psi(T) = e^{AT}$ is evaluated by an approximation: $e^{AT} = U + \sum_{i=1}^{M} \frac{A^i}{i!} T^i$. In this example and Example 6.6 M is taken to be 10. Thus a discrete model as described in Section 2.2 is realized after these approximations, and hence the techniques developed in Chapters 3 and 4 are applicable.

The smallest integer N is found to be 6 and γ to be 1, i.e., $d \in \Theta_5^1$. The corresponding optimal control sequence is $[u_1^1, u_1^2, \ldots, u_5^2] = [0.09173, -0.02240, 1.0, -1.0, 1.0, -1.0, 1.0, 0.45176, 1.0, 1.0, 0.32827, 0.0]$.

Example 6.4 Let the same system equation of Example 6.3 be given and let $\gamma = 2$, $N = 5$, $d' = (1402.5, 45.2, 139.5, -10.8)$. The problem is to find a point $x \in \Theta_{N-1}^{\gamma}$ and the correspon-ding control sequence such that $\|d-x\|^2$ is minimum.

Solution The point x is one of the vertices of Θ_4^2 and the corresponding optimal control sequence is [1.0, -1.0, 1.0, -1.0, 1.0, 1.0, 1.0, -1.0, 1.0, 1.0].

Example 6.5

Given (1) the system equation [P2]

$$\dot{x}(t) = \begin{bmatrix} -\frac{1}{T_1} & 0 & 0 & 0 & 0 \\ \frac{1}{T_3} & -\frac{1}{T_3} & 0 & 0 & 0 \\ 0 & \frac{K_2}{T_4} & -\frac{T_4+T_5}{T_4T_5} & -\frac{1}{T_4T_5} & 0 \\ 0 & \eta & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{M} & -\frac{D}{M} \end{bmatrix} x(t) + \begin{bmatrix} -1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} u(t),$$

where $\eta = \frac{1}{T_4 T_5} - \frac{(T_4+T_5)K_2}{T_4^2 T_5}$, $T_1 = 94.25$, $T_3 = 3.8$,

$T_4 = 113.1$, $T_5 = 4524$, $K_2 = 0.324$, $M = 18221.6$, and $D = 5.94$.

(2) Constant sampling periods of 5 seconds.

(3) The control constraint is $|u(t)| \leq 1$.

(4) The initial state $x_0' = (0\ 0\ 0\ 0)$.

(5) The terminal state $d' = (-43.88, -41.77, -3.09, -63.7, 0.052)$.

Find the smallest integer N and the corresponding optimal control sequence satisfying $d \in R_N$ and $|u_N| =$ minimum.

Solution The method used to approximate the continuous equation is the same as Example 6.4, and hence it is not repeated. The smallest integer is 12, and the optimal control sequence is $\vec{u}_N = [u_1, u_2, \ldots, u_{12}] = [1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 1.0, 0.985, 0.85, 1.0, 1.0]$.

Example 6.6 Let the same system equation of Example 6.5 be given and let $\gamma = 2$, $N = 6$, $d' = (-43.88, 21.77, -3.09, 26.70, -2.05)$. The problem is to find a point $x \in \Theta_{N-1}^{\gamma}$ and the corresponding control sequence such that $\|d-x\|^2$ is minimum.

Solution The point x is one of the vertices of Θ_5^2 and the corresponding optimal control sequence is [1.0, 1.0, 1.0, 1.0, 1.0, -1.0, -1.0, 1.0, 1.0, -1.0, -1.0, -1.0].

The computations for these examples were performed using the CDC 3600 computer at Michigan State University Computer Laboratory. The acutal running time for each of the six example problems were: 0.15, 11.458, 2.143, 5.936, 6.918, 29.013 seconds respectively.

The examples are intended to illustrate the application of the theory developed in Chapter 2 and the algorithms in Chapters 3, and 4 to time-optimal control problems and terminal-error regulator problems. Some general results and particular comments on the algorithms will be discussed in the following:

(1) The speed of the algorithm for checking $d \in R_N$ is dependent considerably upon the nature of a given specific problem. For instance, in the third-order system of Example 6.1, the subsets of hyperplanes which constitute the boundary of R_4 are $|c_i' x| = 1$, $i = 1,2,\ldots,6$, as listed is Table 6.2. The computer program is written to check whether $|c_i' d| > 1$ or $|c_i' d| \le 1$ for $i = 1,2,\ldots,6$, in order. If for certain desired terminal state d, $|c_1'd| > 1$, then it is clear that $d \notin R_4$. On the other hand, if $|c_i' d| \le 1$ for $i = 1, 2,\ldots,5$, but $|c_6' d| > 1$, then it is obvious that five more calculations are needed before the conclusion that $d \notin R_4$ can be made than it would in the former case. For systems of various orders and different N's, similar considerations

hold.

The time-optimal algorithm compares favorably with other methods for finding the smallest N such that $d \in R_N$, i.e.,

$$d = Z\,u = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1N} \\ z_{21} & z_{22} & \cdots & z_{2N} \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ \cdot & & & \cdot \\ z_{n1} & z_{n2} & \cdots & z_{nN} \end{bmatrix} \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_N \end{bmatrix} = [z_1, z_2, \ldots, z_N] \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_N \end{bmatrix} \tag{11}$$

$$\text{and} \quad |u_i| \le 1, \quad i = 1,2,\ldots,N \tag{12}$$

Assume $N \ge n$. From (11), $u_1, u_2, \ldots,$ and u_{n-1} can be solved in terms of $u_n, u_{n+1}, \ldots, u_N$. Geometrically this defines an (N-n)-flat in E^N. On the other hand, (12) defines a hypercube M around the origin in E^N. It is clear that $d \in R_N$ if and only if the intersection of M and this (N-n)-flat is not empty [H6]. But no further implementation of this method was suggested in [H6]. Koepcke [K8] solved the problem by storing all the boundary hyperplanes of R_N, $c_i'x = \eta_i$. In this dissertation only c_i of $|c_i'\,x| = 1$ are stored, hence no additional storage is required for storing η_i's. Koepcke did not treat the most general problems in that c_i's are constructed from the linearly independent columns of Z. A simple example can confirm this claim. For example, let

$$Z = [z_1, z_2, z_3] = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 2 & 3 \\ 0 & 2 & 2 \end{bmatrix} .$$

Hence $z_3 = z_1 + z_2$, but $\{z_1, z_2\}$, $\{z_1, z_3\}$, and $\{z_2, z_3\}$ determine six subsets of the boundary hyperplanes of R_3, which is indeed correct. Neustadt's method [N4] consists of finding

$$\alpha(N) = \min_k \frac{\sum_{i=1}^{N} \sum_{j=1}^{m} |(c_k, z_i^j)|}{|(c_k, d)|} \quad * .$$

If $\alpha(N) \leq 1$, then N is the minimum time. It is clear that for some N, to check whether $|c_k' d| = |(c_k, d)| \leq 1$ is easier and faster than to find $\alpha(N)$ and check whether $\alpha(N) \leq 1$. In order to use the linear programming technique [T1, Z1], some transformations must be made to (11) and (12). Let $u_{im} = u_i + 1$, $i = 1,2,\ldots,N$. Then (11) and (12) become

$$z_1 u_{1m} + z_2 u_{2m} + \cdots + z_N u_N = d - \sum_{i=1}^{N} z_i = \bar{d} = \begin{bmatrix} \bar{d}_1 \\ \bar{d}_2 \\ \cdot \\ \cdot \\ \cdot \\ \bar{d}_n \end{bmatrix} \quad (13)$$

and

$$0 \leq u_{jm} \leq 2, \quad j = 1,2,\ldots,N \quad (14)$$

respectively. Depending upon whether $\bar{d}_i$ is positive or negative, either λ_i or $-\lambda_i$ is added to the left

---

* This applies also when the input signal to the system is scalar, i.e., $m = 1$.

side of (13). Then (13) becomes

$$z_{i1}u_{1m} + z_{i2}u_{2m} + \cdots + z_{iN}u_{Nm} \pm \lambda_i = \bar{d}_i, \; i = 1,2,\ldots,n. \quad (15)$$

Add $q_j \geq 0$ to the left side of (14), then

$$u_{jm} + q_j = 2, \; j = 1,2,\ldots,N. \quad (16)$$

To find whether $d \in R_N$ by linear programming technique, $\Sigma\lambda_i$ has to be minimized. It can be shown that if any feasible solution exists to Equations (11) and (12), then a basic feasible solution also exists. Therefore, unless there is no solution to the original constraints, the optimal solution to this auxiliary linear programming problem will be $\Sigma\lambda_i = 0$. Since the λ_i's were required to be non-negative, their sum can be zero only when each variable is itself zero.

In general, the total number of multiplications and divisions required by a linear programming technique is

$$\sum_{i=n}^{N} \{2i+(i+n)+2i(i+n)\}2i = (6N + 8n + 24)\binom{N+1}{3}+(6n+10)\cdot \binom{N+1}{2} + (6-14n)\binom{n+1}{3}. \quad (17)$$

while that by using the method presented in this dissertation is

$$\left\{\sum_{i=1}^{n-1}[(n-i)(n-i-1) + (n-i)] + \frac{n(n-1)}{2}\right\} \cdot \binom{N}{n-1} = 2\binom{n+1}{3}\cdot\binom{N}{n-1}. \quad (18)$$

At this point it may seem groundless to try to conclude which of these two methods requires less operations and hence results in less computer time. However, while the latter provides complete information about the boundary of R_N, the former gives no insight into the problem. When the method presented in this dissertation is used, it requires little time to compute the time-optimal control sequence after the minimum time N is found. For certain specific values of N and n, (17) and (18) can be evaluated and compared. In Example 6.1, 6.2, 6.3, and 6.5 equation (17) gives 856, 1037736, 38760, 40736; equation (18) gives 48, 3480, 4400, 19800 respectively.

(2) $|c_i' x| = 1$ not only provides information for checking whether $d \in R_N$, but also furnishes the boundary of R_N. This latter feature is the basis of most of the extension work.

(3) The effect due to round-off errors is present. For example, the boundary hyperplane coefficients c_1' of R_4 in Example 6.1 are [0.66667, -0.33333, 0.00000] with five-decimal-places round-off accuracy, while the actual values are $[\frac{2}{3}, -\frac{1}{3}, 0]$. Thus in concluding whether $d \in R_N$ by testing $|c_i' d| \leq 1$ or $|c_i' d| > 1$, there is flexibility in using either $1 + \epsilon$ or $1 - \epsilon$ instead of 1, where ϵ is a small number.

(4) When some (n-1) vectors from $\{z_1, z_2, \ldots, z_N\}$ are nearly linearly dependent, the method of finding the

coefficients of the boundary hyperplanes of R_N, which, in essence, is finding solutions to systems of simultaneous linear equations, may suffer difficulties.

(5) In the multi-input system, any (n-1) vectors from

$$\{z_1^1, z_1^2, \ldots, z_1^m, z_2^1, \ldots, z_2^m, \ldots, z_N^1, z_N^2, \ldots, z_N^m\}$$

may not be linearly independent. Since the algorithm includes all the possibilities of taking (n-1) vectors from the Nm vectors, it is clear that if any (n-1) vectors are linearly dependent, then they are contained in some hyperplane determined by other (n-1) vectors from the Nm vectors. Hence the construction of R_N or Θ_N^γ does not exhibit any problem. As was pointed out by Hankley and Tou [H5], Desoer and Wing's method [D2, D3, D4, W1] in constructing the critical surfaces is not, in general, valid.

(6) The computer storage and time requirements depend mainly on the terminal time N and the order of the system n. But, in general, as described in (1), the algorithm for finding the smallest N such that $d \in R_N$ requires less computer storage and computer time.

(7) In the terminal-error regulator problem the terminal time N and the desired target point d are given in advance. It is required to find a point $x_N \in R_N$ and the corresponding control sequence such that $\|d-x_N\|^2$ is minimum. If $N < n$ and $Gd \in R_N$, then L_1Gd is the

unique optimal control sequence (see equation (3)). If $N \geq n$ or if $N < n$ and $Gd \notin R_N$, then x_N is either a vertex (then the problem is completely solved) or the problem is reduced to an equivalent one with lower dimension. In the first case where $N \geq n$, the problem is reduced to k-dimensional one ($k \leq n-1$, see pp.57). In the second case when $N < n$, and $Gd \notin R_N$, then Gram-Schmidt orthogonalization process (see part (ii) of Section 4.2) is used to reduce the problem to N-dimensional one, where $N = n$. The procedure then repeats for $N \leq n$. This method, as contrast to the quadratic programming technique suggested by Nahi and Wheeler [N2, N3], mainly involves matrix calculations and is easily programmed on a computer.

(8) For $N < n$, the computational aspects of both time-optimal and terminal-error regulator problems have never been before considered in the literature. In this dissertation, all possible relationships between n and N are considered. In the time-optimal control problem, if the minimum time is $N \geq n$ and $d \in \partial R_N$, then optimal control sequence has at most (n-1) members with magnitude different from unity. If the minimum time $N \geq n$ and $d \notin \partial R_N$, then the optimal control sequence has at most n members with magnitude different from unity. In the terminal-error regulator problem, if the given time $N \geq n$ and $d \notin R_N$, then

the optimal control sequence has at most (n-1) members with magnitude different from unity. On the other hand, if $N \geq n$ and $d \in R_N$, and $d \notin R_{N-1}$, then the optimal control sequence has at most n members with magnitude different from unity.

REFERENCES

[B1] R. Bellman, Dynamic Programming, Princeton University Press, Princeton, New Jersey, 1957.

[B2] J.E. Bertram, P.E. Sarachik, "On Optimal Computer Control," Pro. First Int'l Congress of Automatic Control, Butterworth, pp. 419-422, 1960.

[B3] A.G. Butkovskii, "The Necessary and Sufficient Conditions for Optimality of Discrete Control Systems," Automation and Remote Control, Vol. 24, No. 8, pp. 963-970, January, 1963.

[C1] J.A. Cadzow, "A Study of Minimum Norm Control for Sampled-Data Systems," Pro. JACC, pp. 545-550, 1965.

[C2] J.A. Cadzow, "An Extension of the Minimum Controller for Discrete Systems," IEEE Trans. Auto. Control, Vol. AC-12, No.2, pp. 202-203, April, 1967.

[C3] J.A. Cadzow, "Minimum-Effort Control for Discrete Systems," IEEE Trans. Auto. Control, Vol. AC-11, No.2, pp. 309-310, April, 1966.

[C4] J.A. Cadzow, "Minimum Energy Regulator for a Linear Discrete System," IEEE Trans. Auto. Control, Vol. AC-12, No. 2, pp. 183-185, 1967.

[C5] M.D. Canon, J.H. Eaton, "A New Algorithm for a Class of Quadratic Programming Problems with Application to Control," J. SIAM Control, Vol. 4, No.1, pp. 34-45, 1966.

[C6] D.H. Chyung, "Discrete Linear Optimal Control Systems with Essentially Quadratic Cost Functionals," IEEE Trans. Auto. Control, Vol. AC-11, No. 3, pp. 404-413, July, 1966.

[C7] D.H. Chyung, "Discrete Linear System with Convex Cost Functional," Presented at the 4th Allerton Conf., Univ. of Illinois, Urbana, Ill., October, 1966.

[C8] D.H. Chyung, "Linear Discrete Optimal Systems with Side Constraints," Department of Electrical Engineering, University of South Carolina, Columbia, S.C., 1966.

[C9] D.H. Chyung, "An Approximation to Bounded Phase Coordinate Control Problem for Linear Discrete Systems," IEEE Trans. Auto. Control, Vol. AC-12, No. 1, pp. 37-42, February, 1967.

[D1] G.W. Deley, G.F. Franklin, "Optimal Bounded Control of Linear Sampled-Data Systems with Quadratic Loss," Trans. ASME, Vol. 87, No. 1, pp. 135-141, March, 1965.

[D2] C.A. Desoer, J. Wing, "An Optimal Strategy for a Saturating Sampled-Data System," IRE Trans. Auto. Control, Vol. AC-6, pp. 5-15, February, 1961.

[D3] C.A. Desoer, J. Wing, "A Minimum Time Discrete Systems," IRE Trans. Auto. Control, Vol. AC-6, pp. 111-125, May, 1961.

[D4] C.A. Desoer, J. Wing, "The Minimum Time Regulator Problem for Linear Sampled-Data Systems: General Theory," J. Franklin Inst., Vol. 272, No. 9, pp. 208-228, September, 1961.

[E1] J.H. Eaton, "An On Line Solution to Sampled-Data Time Optimal Control," J. Electronics and Control, Vol. 15, pp. 333-341, 1963.

[E2] H.G. Eggleston, "Convexity," Cambridge Tracts in Math and Math. Physics No. 47, Cambridge University Press, 1958.

[F1] K.A. Fegley, M.I. Hsu, "Optimal Discrete Control by Linear Programming," IEEE Trans. Auto. Control, Vol. AC-10, No. 1, pp. 114-115, Jan. 1965.

[F2] J.S. Frame, "Matrix Theory and Linear Algebra with Applications," (Class Notes) Department of Mathematics, Michigan State University.

[F3] J.S. Frame, H.E. Koenig, "Matrix Functions and Applications," IEEE Spectrum, March-July, 1964. pp. 208-220, 102-108, 100-109, 123-131.

[G1] T.L. Gunkel, G.F. Franklin, "A General Solution for Linear Sampled-Data Control," Trans. ASME, Vol. 85, No. 2, pp. 197-203, June, 1963.

[H1] G. Hadley, Linear Programming, Addison-Wesley Pub. Co., Reading, Mass., 1962.

[H2] H. Halkin, "Optimal Control for Systems Described by Difference Equations," in Advances in Control Systems: Theory and Applications, C.T. Leondes, Ed., Vol. 1, pp. 173-196. New York: Academic, 1964.

[H3] H. Halkin, "A maximum Principle of the Pontryagin Type for Systems Described by Nonlinear Difference Equations," J. SIAM Control, Vol. 4, No. 1, pp. 90-111, February, 1966.

[H4] H. Halkin, B.W. Jordan, E. Polak, J.B. Rosen, "Theory of Optimal Discrete Time Systems," Pro. 3rd Int'l Congress of Auto. Control, pp. 28B.1-28B.7, London, 1966.

[H5] W.J. Hankley, J.T. Tou, "Note on Control of Multiple-Input Discrete Systems," IEEE Trans. Auto. Control, Vol. AC-12, pp. 469-470, August, 1967.

[H6] Y.C. Ho, "Solution Space Approach to Optimal Control Problems," Trans. ASME, Vol. 83, pp. 53-58, March, 1961.

[H7] J. Holtzman, "Convexity and the Maximum Principle for Discrete Systems," IEEE Trans. Auto. Control, Vol. AC-11, No. 1, pp. 30-35, January, 1966.

[H8] J. Holtzman, "On the Maximum Principle for Nonlinear Discrete-Time Systems," IEEE Trans. Auto. Control, Vol. AC-11, pp. 273-274, April, 1966.

[H9] J. Holtzman, H. Halkin, "Directional Convexity and the Maximum Principle for Discrete Systems," J. SIAM Control, Vol. 4, No. 2, pp. 263-275, May, 1966.

[H10] P.R. Halmos, Finite-Dimensional Vector Spaces, D. Van Nostrand Co., Princeton, N. J., Second Ed., 1958.

[H11] R.D. Hutcheson, "The Computation of Optimal Controls by an Iterative Procedure," Thesis for the degree of Aeronautical and Astronautical Engineer, University of Michigan, Ann Arbor, Michigan, 1965.

[J1] B.W. Jordan, E. Polak, "Theory of a Class of Discrete Optimal Control Systems," J. Electornics and Control, Vol. 17, pp. 697-711, 1964.

[K1] R.E. Kalman, "Optimal Nonlinear Control of Saturating Systems by Intermittent Action," IRE WESCON Convention Record, Part. 4, pp. 130-135, 1957.

[K2] R.E. Kalman, "On the General Theory of Control Systems," Pro. 1st Int'l Congress of Auto. Control, pp. 481-492, Moscow, 1960.

[K3] S. Katz, "A Discrete Version of Pontryagin's Maximum Principle," J. Electronics and Control, Vol. 13, No. 2, pp. 179-184, August, 1962.

[K4] S. Katz, "A General Maximum Principle for End-point Control Problems," Vol. 16, No. 2, pp. 189-222, February, 1964.

[K5] H.E. Koenig, Y. Tokad, H.K. Kesavan, Analysis of Discrete Physical Systems, McGraw-Hill Book Co., New York, N.Y., 1967.

[K6] R.W. Koepcke, "A Solution to the Sampled, Minimum-Time Problem," Trans. ASME Vol. 81, No. 1, pp. 145-150, March, 1964.

[L1] E.B. Lee, "A Computational Scheme for Discrete Systems," Trans. IEEE Auto. Control, Vol. AC-9, No. 1, pp. 115, 1964.

[N1] A. Nagata, S. Kodama, S. Kumagai, "Time Optimal Discrete Control System with Bounded State Variable," IEEE Trans. Auto. Control, Vol. AC-10, pp. 155-164, April, 1965.

[N2] N.E. Nahi, L.A. Wheeler, "Discrete Terminal Control of Space Vehicles via Mathematical Programming," Univ. of Southern California, Electrical Engineering Department Report 141, October, 1965.

[N3] N.E. Nahi, L.A. Wheeler, "An Iterative Procedure for Solving the Discrete Terminal Control Problem," Pro. Nat'l Electronics Conf., Vol. XXII, pp. 671-677, 1966.

[N4] L.W. Neustadt, "Discrete Time Optimal Control Systems," Int'l Symposium on Nonlinear Differential Equations and Nonlinear Mechanics, Edited by J.P. LaSalle and S. Lefschetz, Academic Press, New York, N.Y., pp. 267-283, 1963.

[O1] K. Ogata, State Space Analysis of Control Systems, Prentice-Hall, Englewood Cliffs, N.J., 1967.

[P1] J.B. Pearson, R. Sridhar, "A Discrete Optimal Control Problem," Trans. IEEE Auto. Control, Vol. AC-11, No.2, pp. 171-174, April, 1966.

[P2] J. Preminger, G.L. Park, "Analysis of Dynamic Stability of a Power System under Deterministic Load Changes," to be published, Proceedings 1969 I.F.A.C., Warsaw.

[P3] A.I. Propoi, "Methods of Feasible Directions in Problems of Discrete Optimal Control," Automation and Remote Control, Vol. 28, No. 2, pp. 234-245, 1967.

[R1] A.M. Revington, J.C. Hung, "Time Weighted Energy Control of Discrete-Data Control Systems," Trans. IEEE Auto. Control, Vol. AC-11, No. 4, pp. 758-759, 1966.

[R2] J.B. Rosen,"Optimal Control and Convex Programming," Nonlinear Programming, J. Abadie, Ed., North-Holland, Amsterdam, 1967.

[R3] L.I. Rozonoer, "L.S. Pontryagin Maximum Principle in the Theory of Optimal Systems I," Automation and Remote Control, Vol. 20, No. 10, pp. 1288-1302, October, 1959.

[R4] L.I. Rozonoer, "L.S. Prontryagin's Maximum Principle in Optimal System Theory II," Automation and Remote Control, Vol. 20, No. 11, pp. 1405-1421, November, 1959.

[R5] L.I. Rozonoer, "The Maximum Principle of L.S. Pontryagin in Optimal-System Theory III,"Automation and Remote Control, Vol. 20, No. 12, pp. 1517-1561, December, 1959.

[T1] H.C. Torng, "Optimization of Discrete Control Systems Through Linear Programming," J. Franklin Inst., Vol. 277, No. 7, pp. 28-44, July, 1964.

[T2] J.T. Tou, "Synthesis of Descrete Systems Subject to Saturation," J. Franklin Inst., Vol. 277, No. 5, pp. 401-413, May, 1964.

[T3] J.T. Tou, "Optimum Control of Discrete Systems Subject to Saturation," IEEE Pro., Vol. 52, No. 1, pp. 88-89, 1964.

[T4] J.T. Tou, Optimal Design of Digital Control Systems, Academic Press, New York, 1963.

[T5] J.T. Tou, B. Vadhanaphuti, "Optimum Control of Nonlinear Discrete-Data Systems," Trans. AIEE., Vol. 80, pp. 166-171, 1961.

[W1] J. Wing, C.A. Desoer, "The Multiple-Input Minimal Time Regulator Problem: General Theory," IEEE Trans. Auto. Control, Vol. AC-8, pp. 125-136, April, 1963.

[Z1] L.A. Zadeh, B.H. Whalen, "On Optimal Control and Linear Programming," IRE Trans. Auto. Control, Vol. Ac-7, pp. 45-46, 1962.

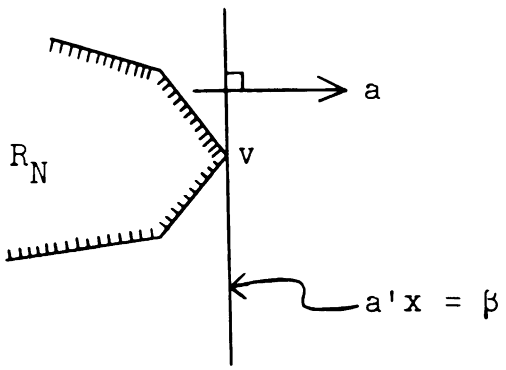

Fig. 1a Intersection of R_N with hyperplane $a'x = \beta$, where $a'x_N < \beta \; \forall \; x_N \in R_N$.

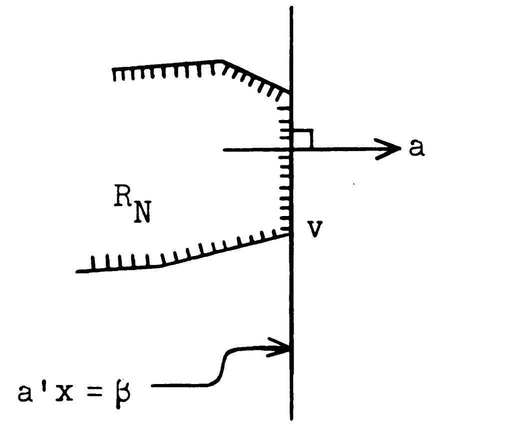

Fig. 1b Intersection of R_N with hyperplane $a'x = \beta$, where $a'x_N \leq \beta \; \forall \; x_N \in R_N$.

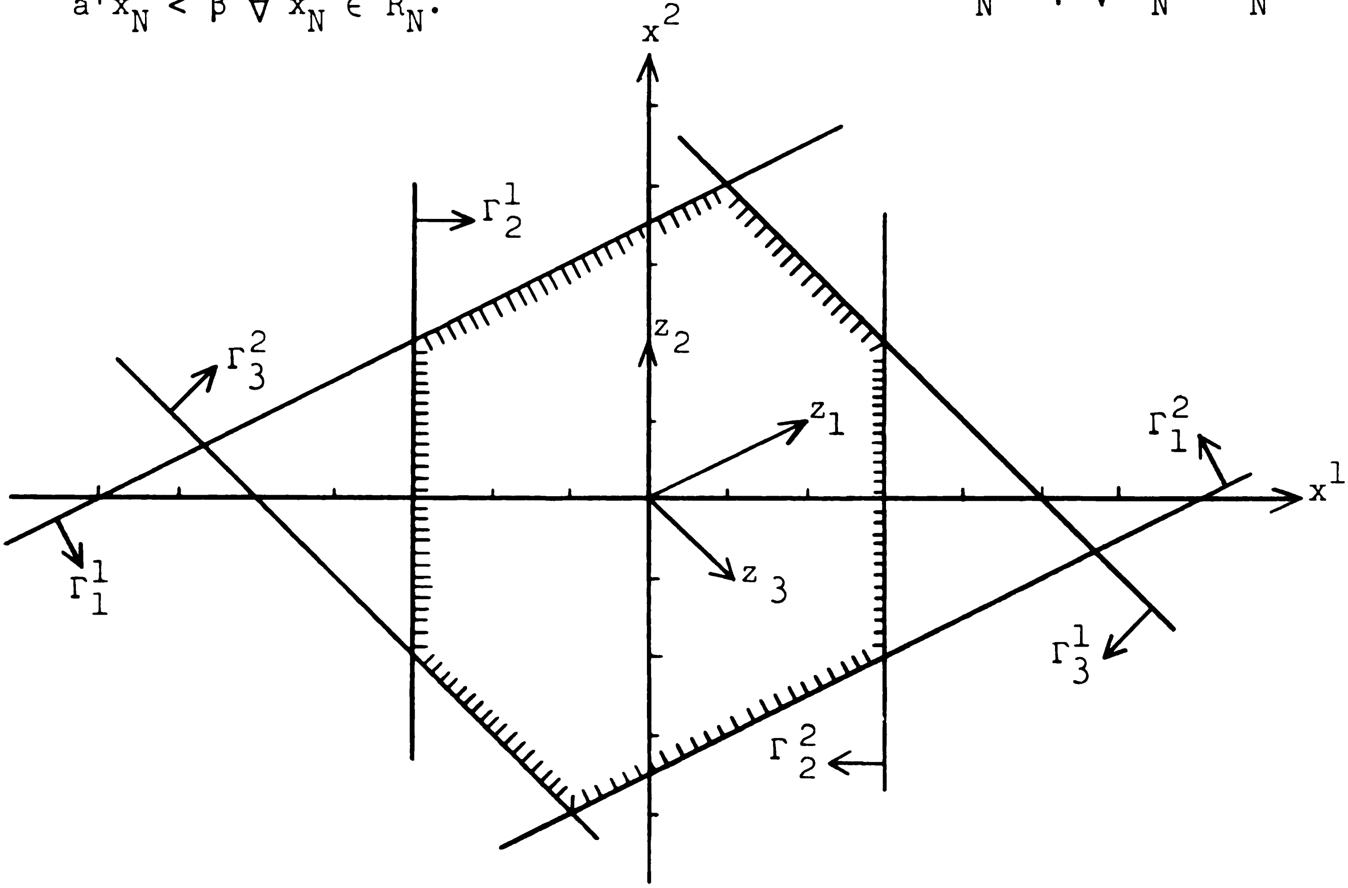

Fig. 2 Closed half-spaces for Example 2.4.1.

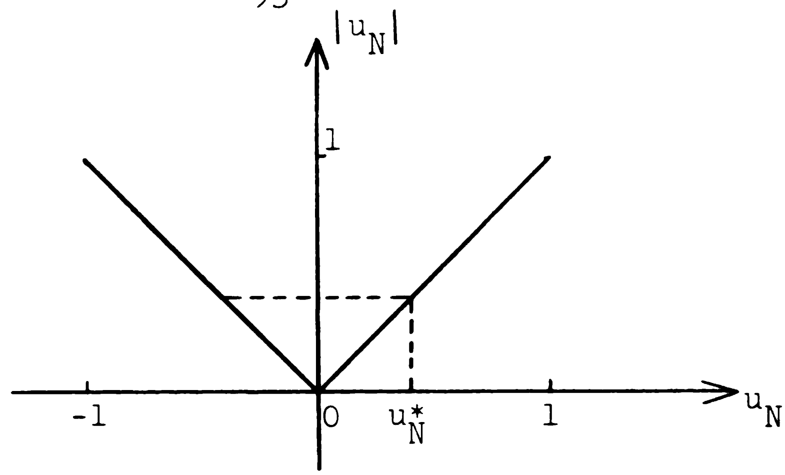

Fig. 3 $|u_N|$ vs. u_N with $|u_N| \leq 1$.

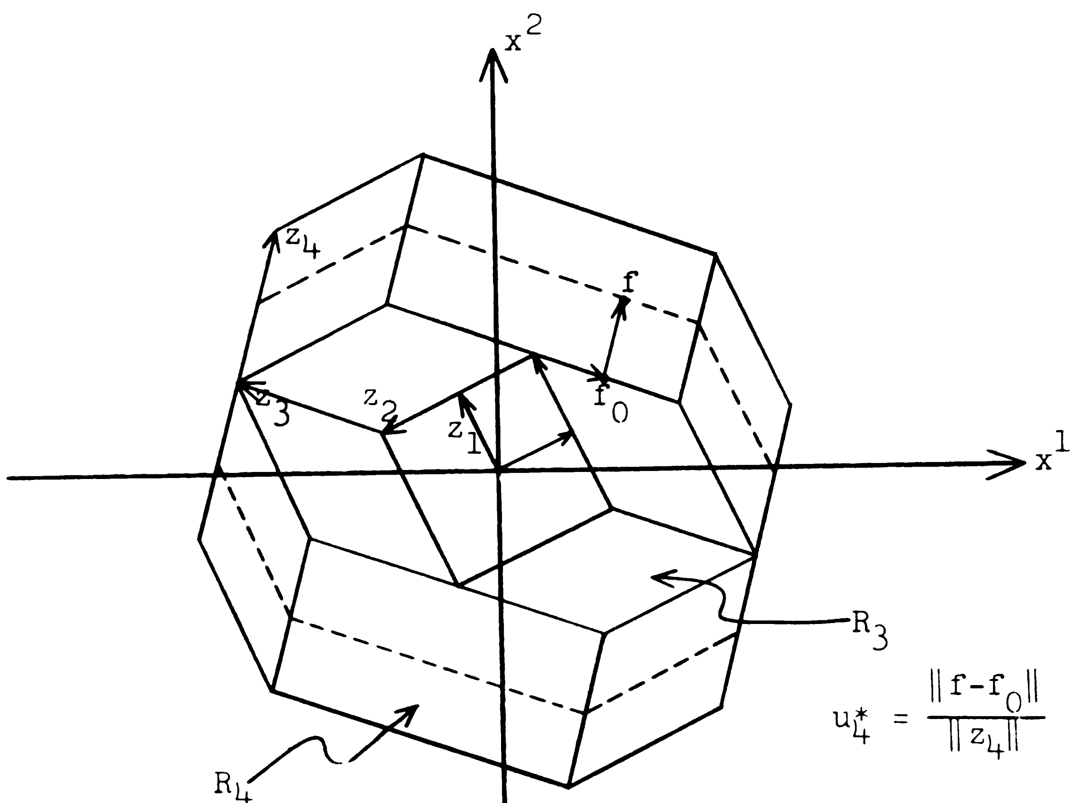

Fig. 4 u_N^* is the minimum of $|u_N|$ such that $f=f_0+u_Nz_N$ (with $f_0 \in \partial R_3, N = 4$)

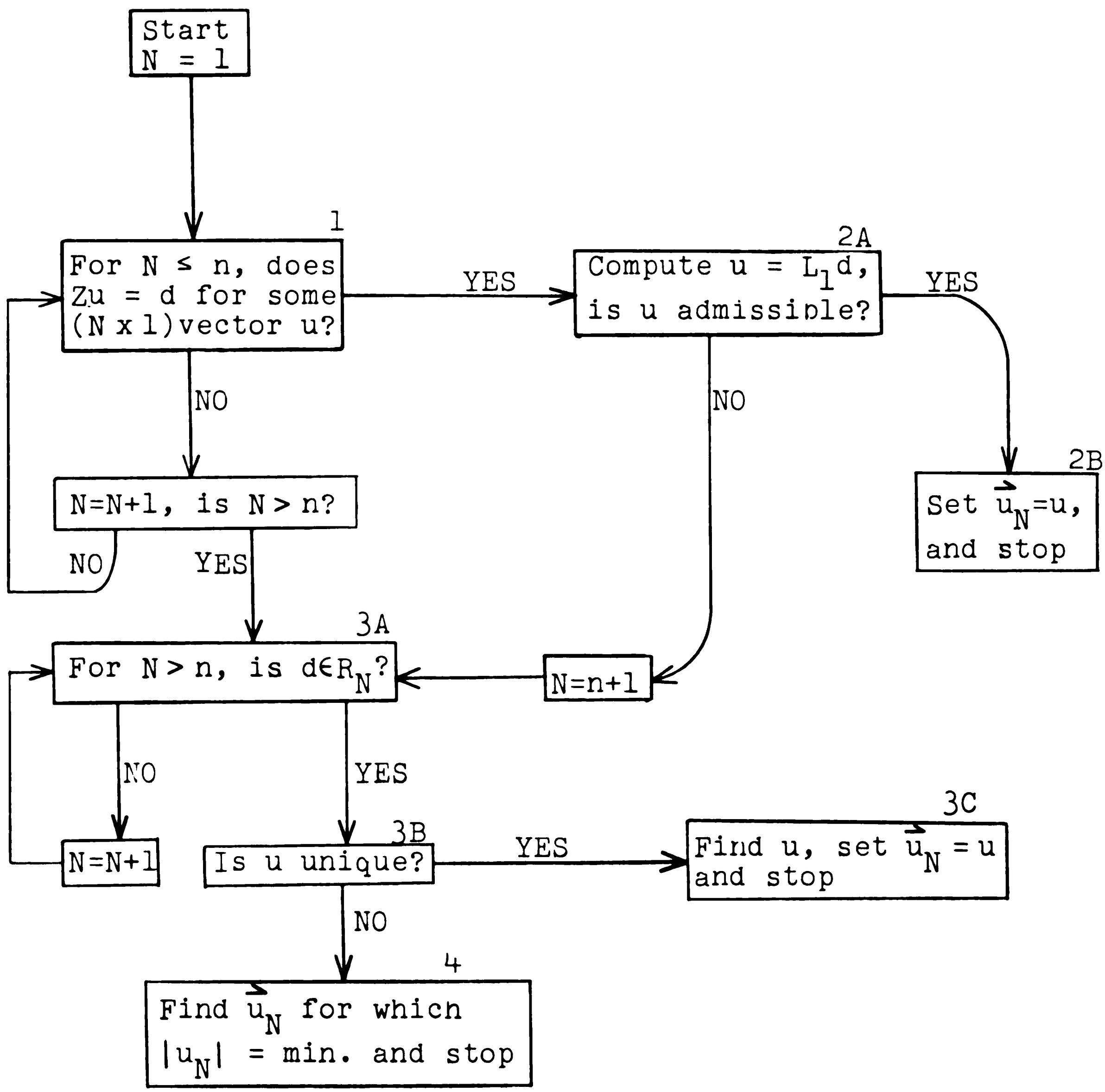

Fig. 6 Computer Flow Diagram for Time-Optimal Control Problem; Phases 1, 2, 3, 4

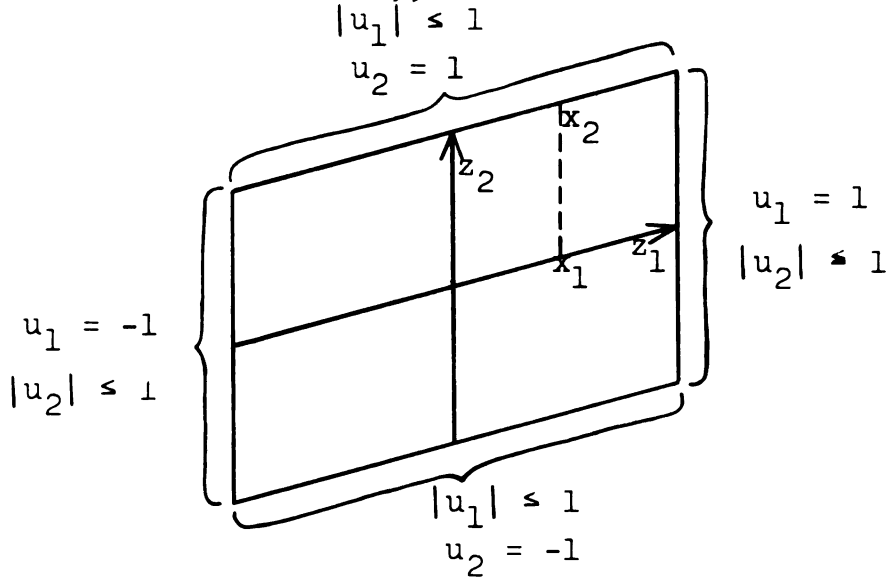

Fig. 5 $x_2 \in \partial R_2$ with $x_1 \notin \partial R_1$.

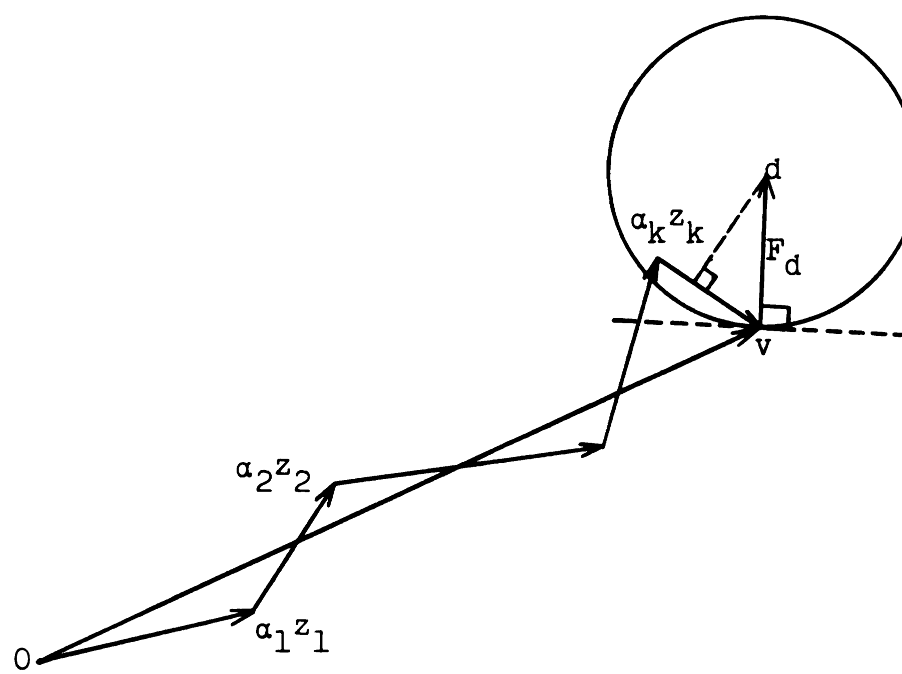

Fig. 7 Minimum squared length of F_d can be improved by finding an appropriate α_k, if $(F_d, \alpha_k z_k) < 0$.

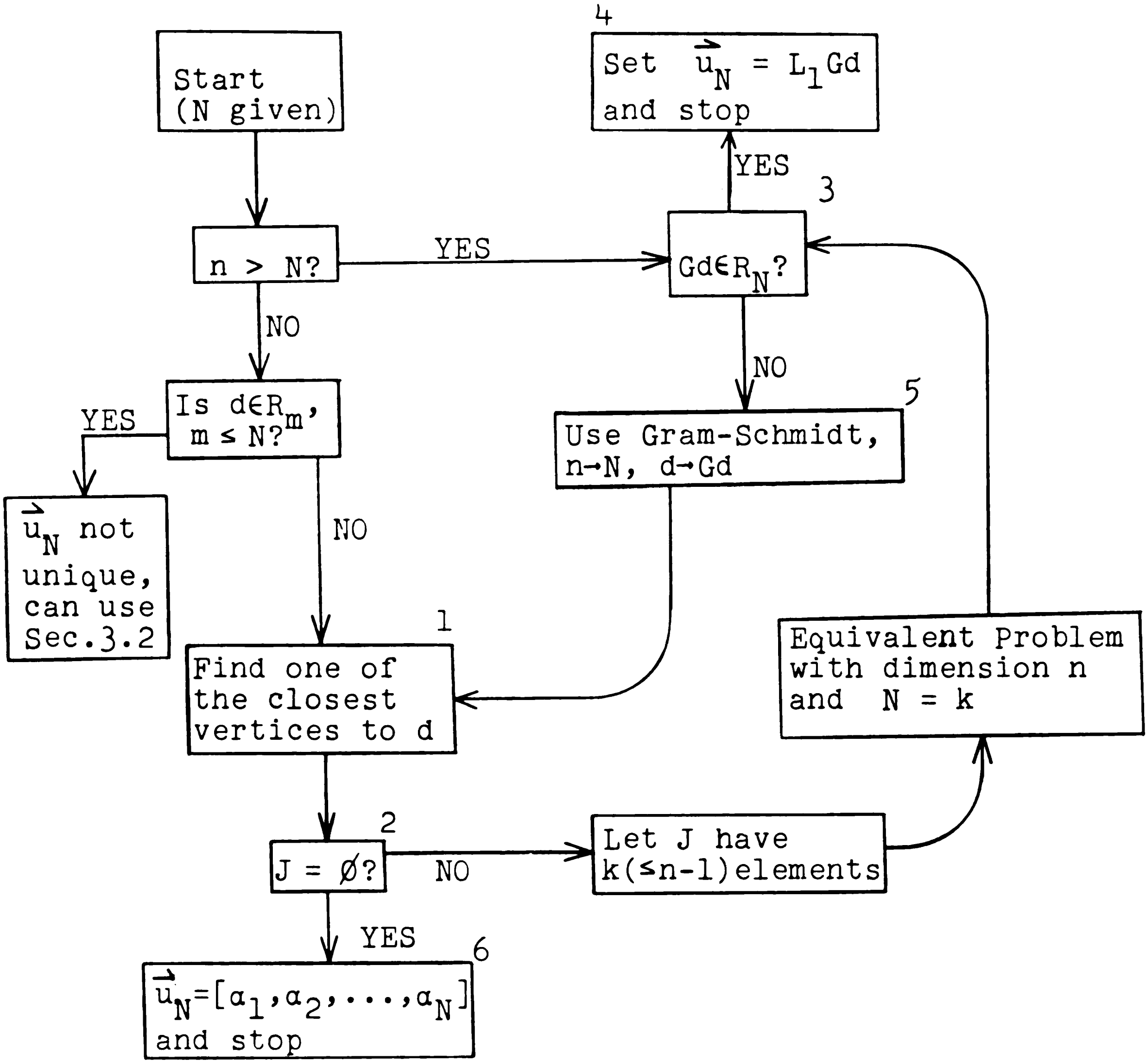

Fig. 8 Computer Flow Diagram for Terminal-Error Regulator Problem; Main Computational Steps 1, 2, 3, 4, 5 and 6

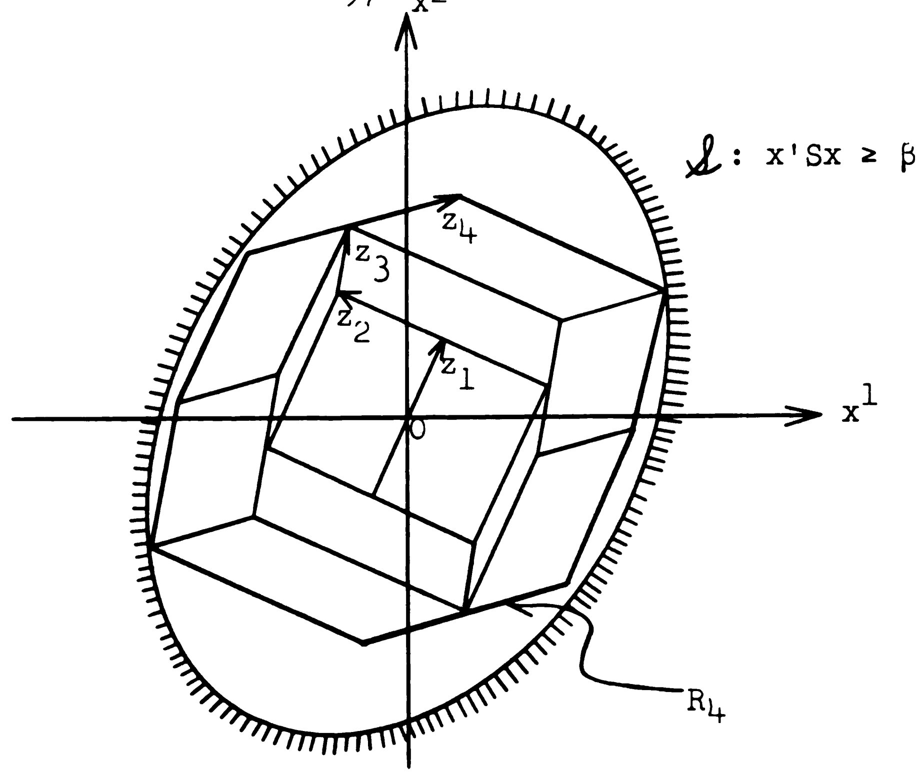

Fig. 9 Intersection of R_4 and the Target Set $\mathcal{S}$: $x'Sx \geq \beta$

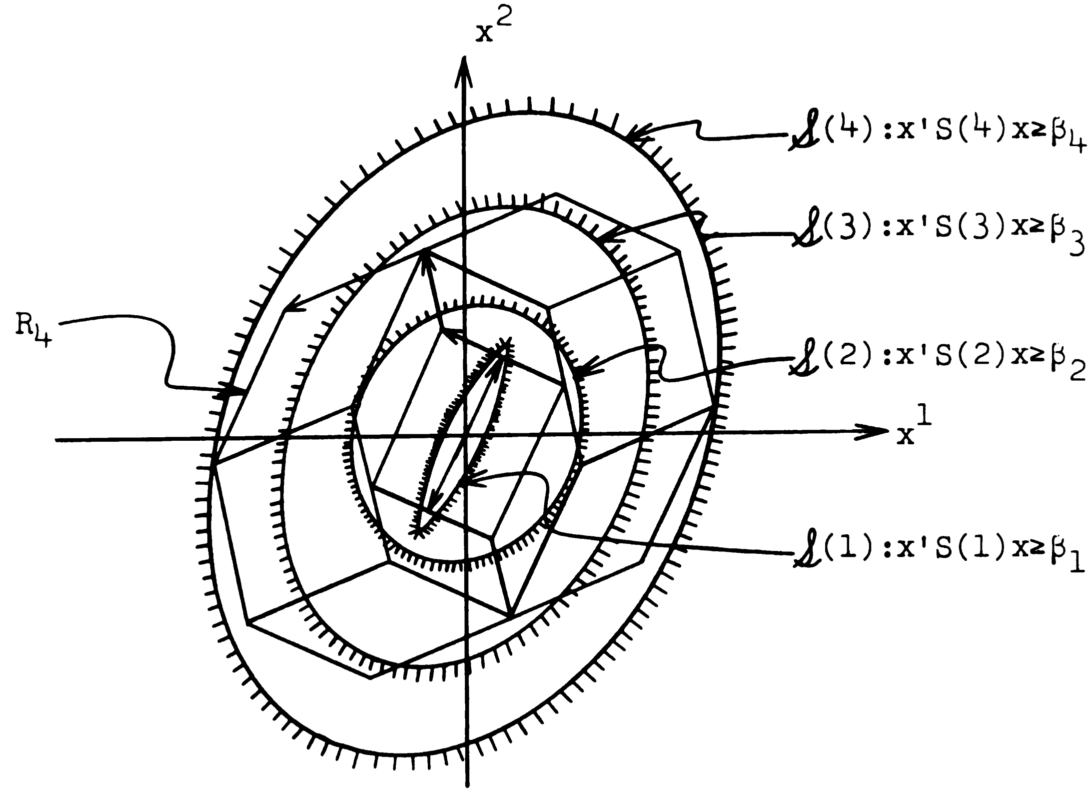

Fig. 10 Time-varying targets; intersection of $\mathcal{S}(4)$ and R_4

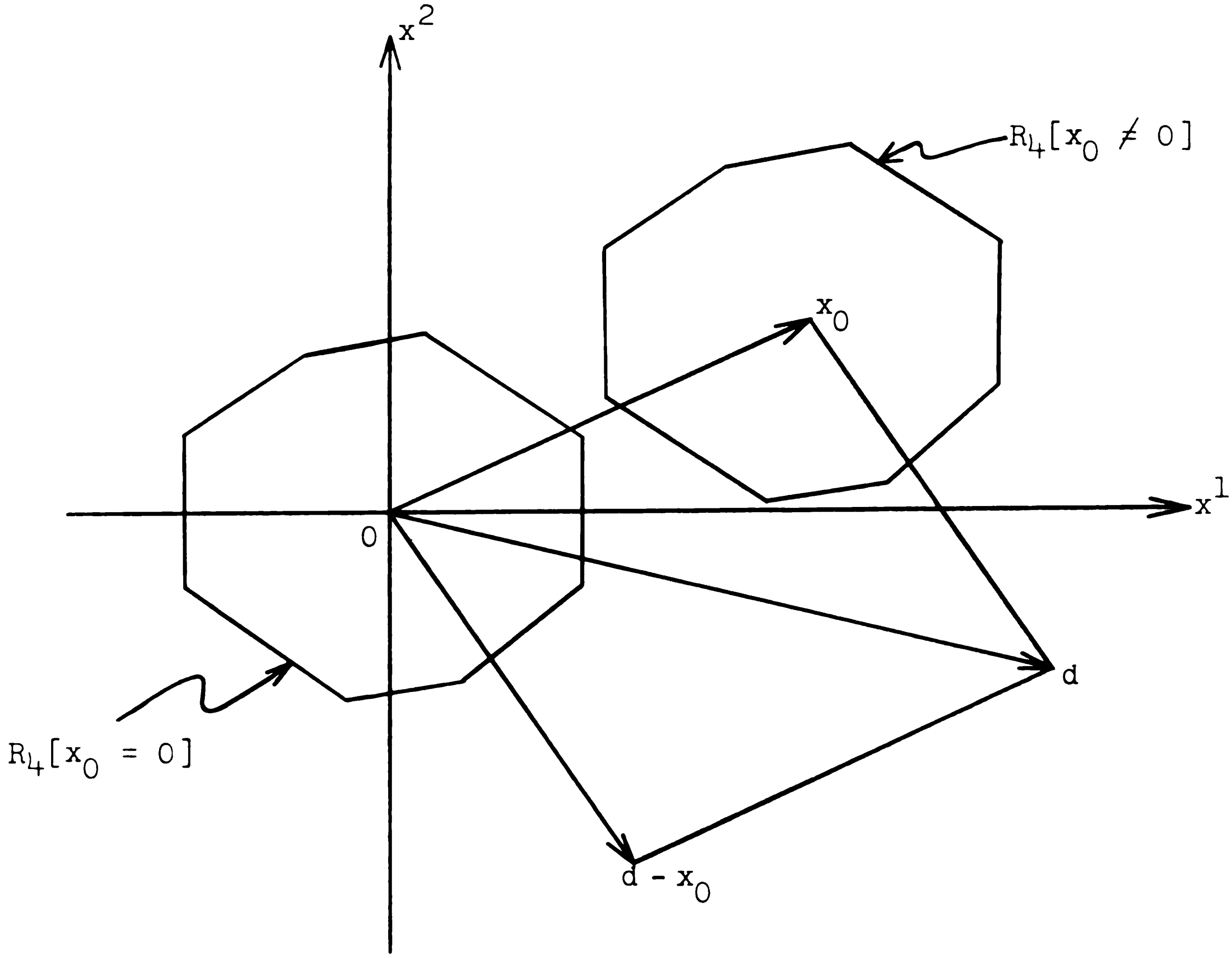

Fig. 11 Translation of Coordinates When $x_0 \neq 0$ Resulting in a New Equivalent Problem with $x_0 = 0$.

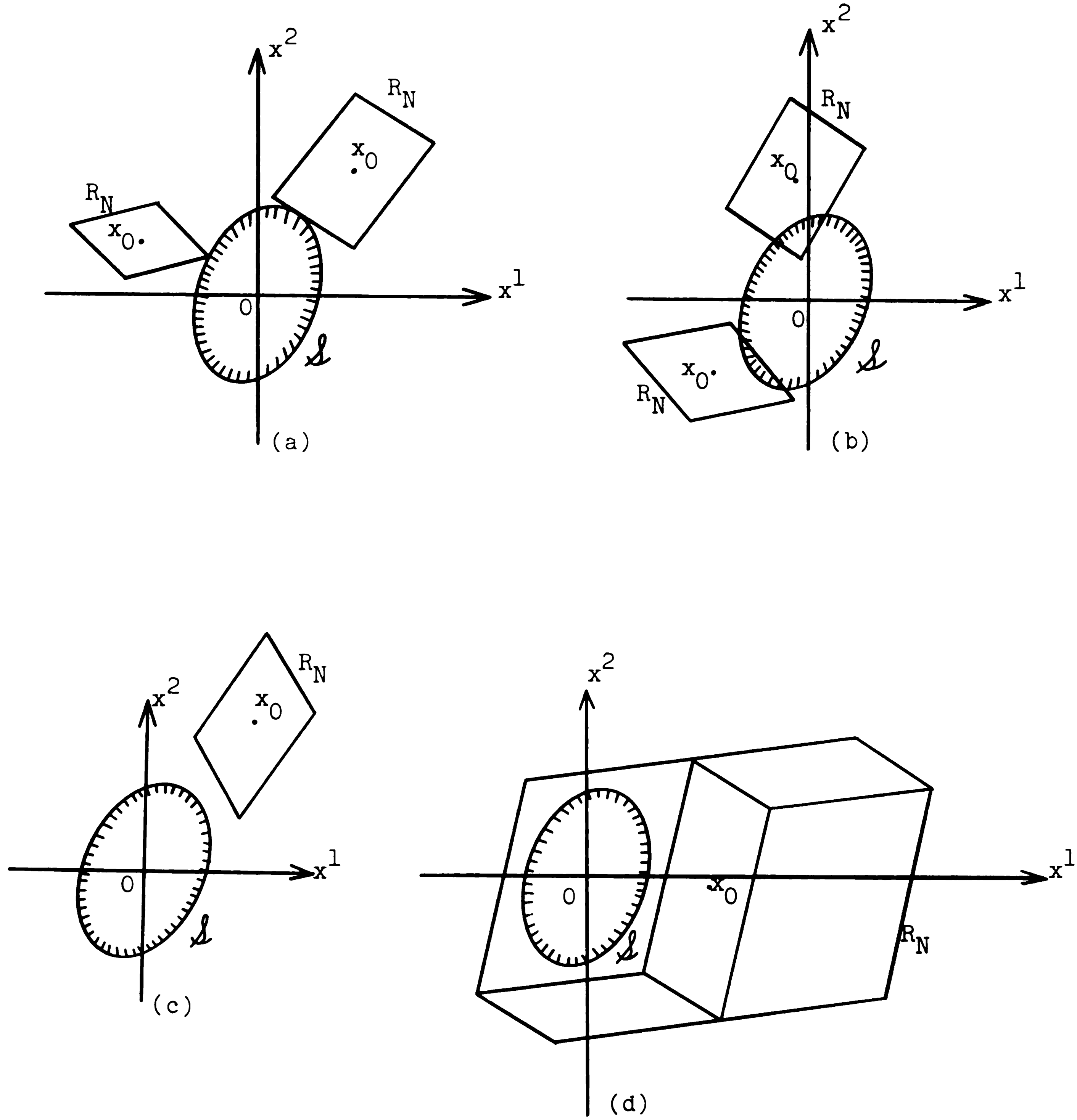

Fig. 12 Some Relations between R_N and $\mathcal{S}$: $x'Sx \le \beta$ (when $x_0 \neq 0$): (a) $R_N \cap \mathcal{S}$ has one point; (b) $R_N \cap \mathcal{S}$ has an infinite number of points; (c) $R_N \cap \mathcal{S} = \emptyset$ (empty set), (d) $R_N \supset \mathcal{S}$.