THE UNIVERSITY OF ASTON IN BIRMINGHAM

ELECTRICAL ENGINEERING DEPARTMENT

DISCONTINUOUS CONTROL OF MULTI-VARIABLE PROCESSES.

LOREN D. MEEKER.

M.Sc. Thesis. May, 1969.

SUMMARY

The research reported in this thesis is concerned with the control of physical systems which can be modelled by linear constant-coefficient vector differential equations of the form

$$\dot{x} = Ax + Gm,$$

where x and m are the n-dimensional state vector and the r-dimensional control vector, respectively. A is the n x n system matrix and G is the n x r control matrix. Examples of processes to which such equations apply are the control systems for small-angle-attitude motion of a satellite, the dynamical control of pilotless aircraft (see [5]), and the temperature-level control of processes.

The behaviour of such systems depends upon the interaction of the matrices A and G, with the control domain $\mathcal{C}$ which depends, in turn, upon the type of controllers used. This interaction is investigated in Chapter One to determine the large-scale behaviour of the system. In Chapter Two the problem of open-loop control is discussed, while in Chapter Three the time-optimal control of such systems in considered.

The methods used to investigate these processes are geometric in nature. They are based upon the geometric properties of convex sets and hyperplanes in the n-dimensional state-space of the system.

While the mathematical methods used to derive the results within may seem very abstract to the design engineer, it is hoped that the results themselves will provide useful insight into the behaviour of the systems as well as design criteria for the practising engineer.

ACKNOWLEDGMENT

The author wishes to thank Professor W. K. Roots, whose guidance, ability and organization made the research reported within not only possible, but enjoyable as well. He also thanks the other members of the Electrical Engineering and Mathematics Departments of the University of Aston, in particular Mr. M. Shridar, for many fruitful discussions throughout the course of this investigation.

The author, as recipient of Science Faculty Fellowship No. 68141, would like to acknowledge the financial support of the National Science Foundation, Washington, D.C., U.S.A.

CONTENTS

		Page
INTRODUCTION AND MATHEMATICAL PRELIMINARIES		(v)
CHAPTER ONE :	LARGE SCALE BEHAVIOUR	1
CHAPTER TWO :	OPEN LOOP CONTROL	24
CHAPTER THREE :	TIME-OPTIMAL CONTROL	33
APPENDIX A :	Linear constant-coefficient differential equations and matrix exponentials. Controllability.	61
APPENDIX B :	Orders of Magnitude.	64
BIBLIOGRAPHY :		65

Introduction and Mathematical Preliminaries.

The mathematical systems discussed in this paper are set in a state-space of n-dimensions and are described by differential equations of the form $\dot{x} = Ax + Gm$ with m, the control, or manipulated variable, lying in a fixed subset $\mathcal{C}$ of an r-dimensional space. The behaviour of the system is determined by the interaction of the two matrices A and G and the set of controls $\mathcal{C}$. As our discussion and description of this interaction will be chiefly in geometric terms, we include at this point a brief discussion of the topological and geometrical concepts required in the chapters to follow. In addition, certain other mathematical concepts are discussed in the appendices. The reader who is familiar with the geometry of higher-dimensional spaces is advised to proceed to Chapter I.

The Euclidean space of k dimensions will be denoted by $\mathbb{R}^k$, each point x of $\mathbb{R}^k$ is uniquely determined by its co-ordinates $x_1, x_2, x_3, \ldots, x_k$ which may be considered its components with respect to the <u>standard basis vectors</u> $e^1 = (1, 0, \ldots, 0)$, $e^2 = (0, 1, 0, \ldots, 0), \ldots$, $e^k = (0, \ldots, 0, 1,)$, in which case the <u>point</u> x is identified with the (<u>position</u>) <u>vector</u> $x = \sum_{i=1}^{k} x_i e^i$. Addition of these vectors is accomplished by addition of their components and their multiplication by scalars is defined as in two or three dimensions (in $\mathbb{R}^2$ or $\mathbb{R}^3$).

The <u>subspace spanned by $x^1, x^2, \ldots, x^m$</u> in $\mathbb{R}^k$ is the set of all vectors $y = \sum_{j=1}^{m} \alpha_j x^j$ where the α_j are arbitrary real numbers. If $x^1, x^2, \ldots, x^m$ are linearly independent, that is if $\sum_{j=1}^{m} \alpha_j x^j = o$ implies $\alpha_1 = \alpha_2 = .. = \alpha_m = o$, then the subspace has dimension m ($m \leqslant k$).

The distance between the two points x and y is defined by

$$|x - y| = \left(\sum_{i=1}^{k} (x_i - y_i)^2 \right)^{\frac{1}{2}} \qquad (I.1)$$

and the <u>norm</u> (<u>length</u>) of x is

$$|x| = |x - o| = \left(\sum_{i=1}^{k} x_i^2 \right)^{\frac{1}{2}}. \qquad (I.2)$$

A sequence of points $x^1, x^2, x^3, \ldots$ is said to <u>converge to a point x</u> if $\lim_{n \;\infty} |x - x^n| = o$, this is often symbolized by $x^n \rightarrow x$. A subset E of $\mathbb{R}^k$ is said to be <u>closed</u> if the limit of every convergent sequence of points in E is also in E. The set E is said to be bounded if there is a constant K with $|x| \leqslant K$ for all x in E.

The <u>k-sphere about x with radius r</u> is that subset of ^{k} consisting of all the points y with $|x - y| \; r$. A point x is called an <u>interior point</u> of a subset E if E contains a k-sphere about x. E is called <u>open</u> if every point of E is an interior point. A <u>boundary point</u> of E is a point in E which is not an interior point. If the subset E is contained in another subset F, a point x of E is called a <u>relative interior point</u> (relative to F) if E contains all the points of a k-sphere about x which are also contained in F. E is called <u>open relative to F</u> if every point of E is a relative interior point.

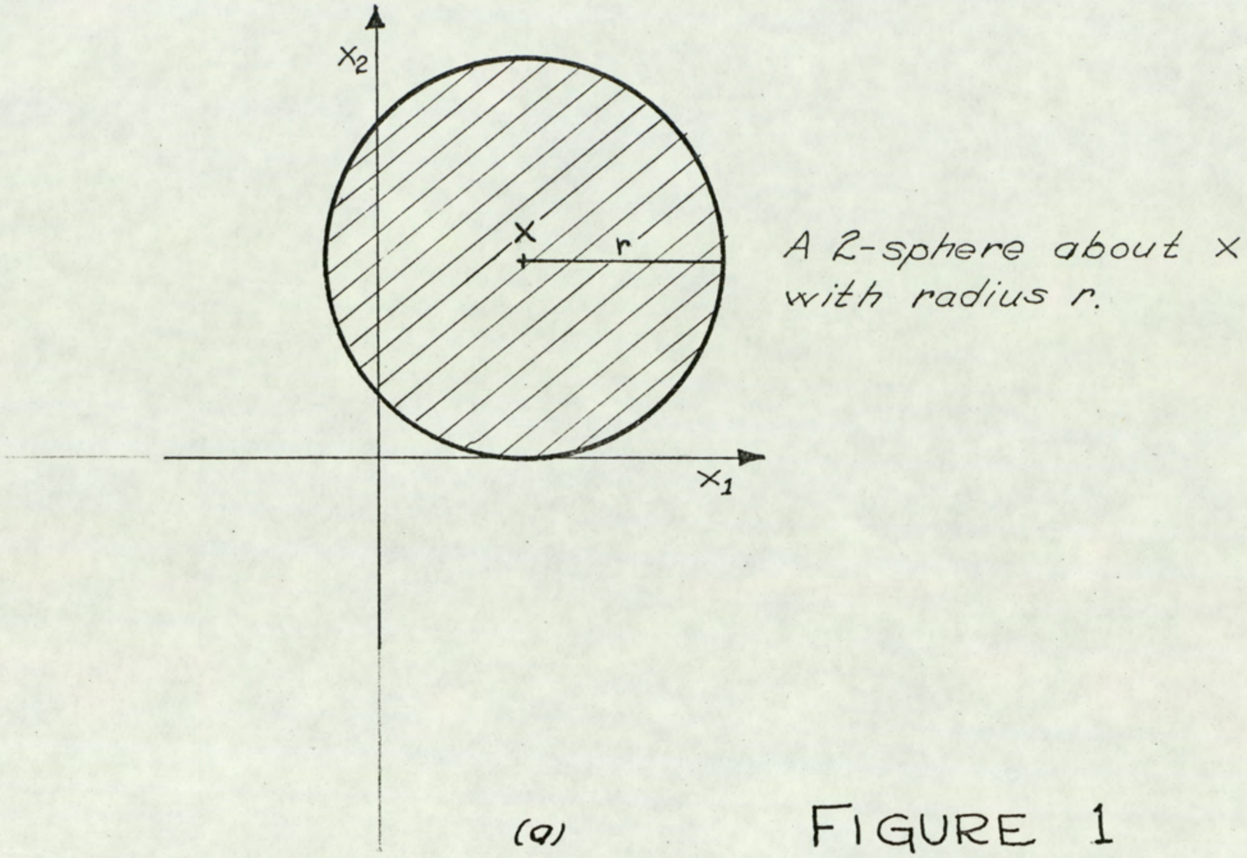
Let's consider for the moment these concepts in the familiar two-dimensional ($\mathbb{R}^2$) and the three-dimensional ($\mathbb{R}^3$) cases. The distance between points is the actual Euclidean distance, while the 2 - spheres in $\mathbb{R}^2$ are discs and the 3 - spheres in $\mathbb{R}^3$ are actual spheres. The set of points x of $\mathbb{R}^3$ with $|x| < 1$ is an open set of $\mathbb{R}^3$, any point y with $|y| = 1$ is a boundary point of this set. If F denotes the set of points x with $|x| \leq 1$ then F is closed. The subset E of F consisting of those points with $x_1 = x_2 = o$ is a closed set of $\mathbb{R}^3$ but a relatively open subset of F. These sets are exhibited in Figure 1.

Let us now develop the geometry of $\mathbb{R}^k$. This is done merely by converting the classical geometrical concepts of two or three-dimensions into analytic relations between points (vectors) and then interpreting them in the higher dimensional spaces (eg. as we did with the sphere above).

If x and y are in $\mathbb{R}^k$, <u>the line through x and y</u> is the locus of all points $z(t) = x + t\ (x - y),\ -\infty < t < \infty$. The <u>line segment between x and y</u> is that portion of the line corresponding to $o < t \leq 1$, or equivalently, the locus of points, $z = \alpha x + \beta y$ where $o \leq \alpha$, $o \leq \beta$ and $\alpha + \beta = 1$. Such a z is called a <u>convex combination</u> of x and y. This is generalized by saying z is a <u>convex combination of $x^1, x^2, \ldots, x^n$</u> if

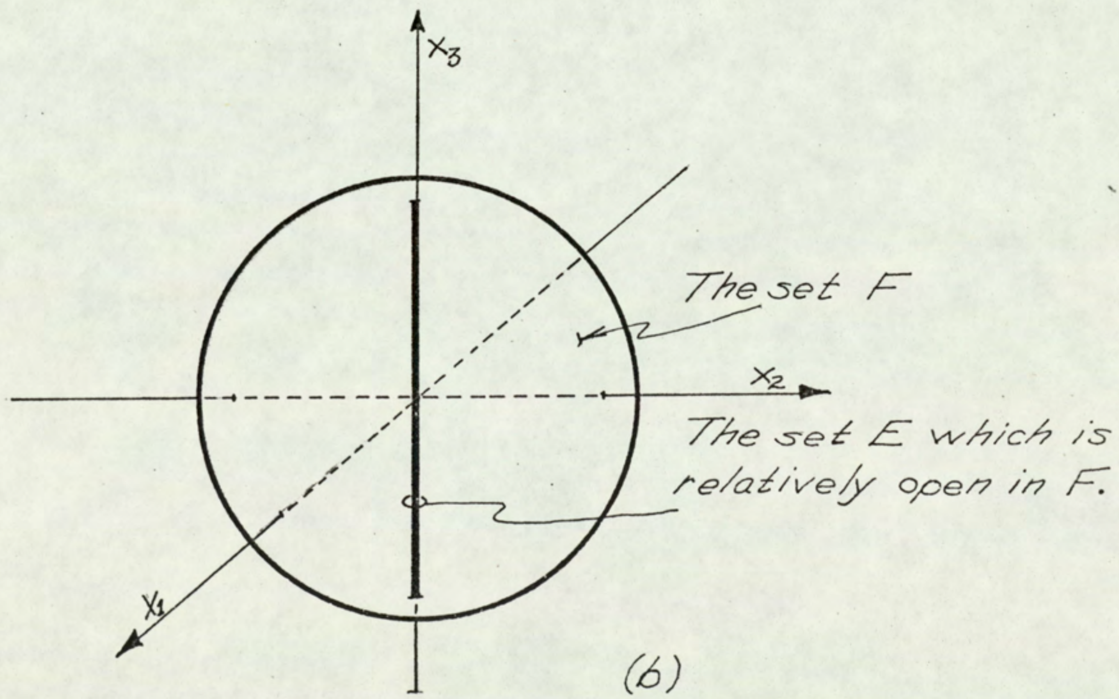
$$z = \sum_{i=1}^{n} \alpha_i x^i \quad \text{where} \quad o \leq \alpha_i,\ i = 1, 2, \ldots, n \text{ and } \sum_{i=1}^{n} \alpha_i = 1 .$$

A subset of $\mathbb{R}^k$ is called <u>convex</u> if it contains along with each pair of points the line segment between them or, equivalently, contains all the convex combinations of its points.


x_2
x
r
x_1
A 2-sphere about x
with radius r.

(a)

FIGURE 1


x_3
The set F
x_2
The set E which is
relatively open in F.
x_1

(b)

The convex hull of a set E is the smallest convex set which contains E, it consists of all convex combinations of points in E. A point x is an extreme point of the convex set C if it is not a convex combination of any two distinct points of C. A convex set is uniquely determined by its extreme points for it is the convex hull of its own extreme points; this means that every point x in a convex set can be written as $x = \sum_{j=1}^{m} \alpha_j x^j$, a convex combination of the extreme points $x^1, x^2, \ldots, x^m$ of the set.

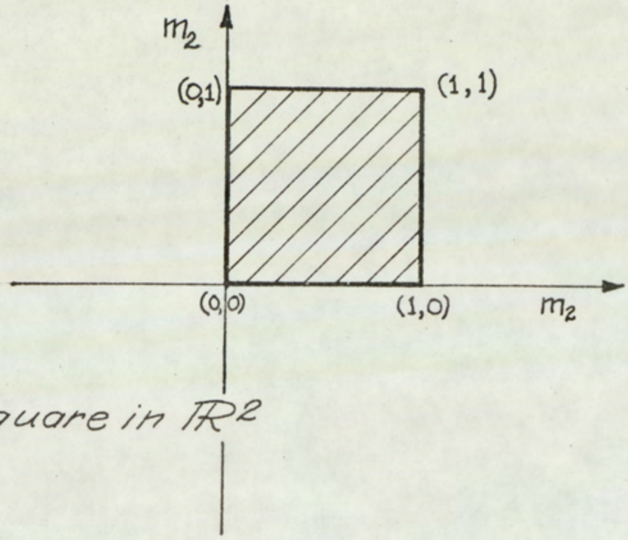
A typical convex subset of $\mathbb{R}^2$ is the unit square which consists of all points $x = (x^1, x^2)$ with $0 \leqslant x^1 \leqslant 1$, $0 \leqslant x^2 \leqslant 1$. The extreme points of the unit square are the points with co-ordinates (o,o), (1,o), (o,1), (1,1), i.e. the corners of the set. A sphere is a convex set and every boundary point is an extreme point. The convex hull of the points $x^1 = (0,0)$, $x^2 = (1,0)$, $x^3 = (3,1)$ is the triangle having these points as vertices, (see Figure 2).

It is clear from the examples immediately above, and easy to show in general, that if a convex set has an interior point x^o and a boundary point y^o then every point z on the line segment between x^o and y^o, excluding y^o itself, is an interior point of the set.

If x and y are vectors in $\mathbb{R}^k$, their inner product (scalar product) is

$$x.y = \sum_{i=1}^{k} x_i y_i = |x| \cdot |y| \cdot \cos(x,y) \qquad (I.3)$$

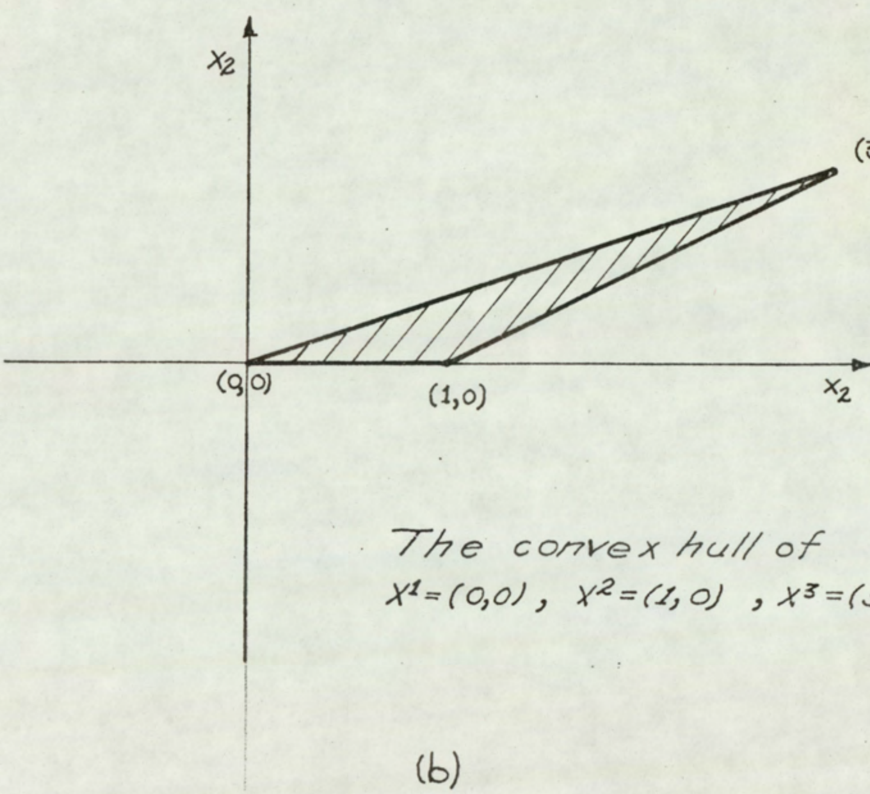
Where $\cos^{-1}(x,y)$ is the angle between x and y defined by



The unit square in $\mathbb{R}^2$

(a)

FIGURE 2



The convex hull of
$X^1 = (0,0)$, $X^2 = (1,0)$, $X^3 = (3,1)$

(b)

considering the ordinary geometric angle between x and y in the two-dimensional space spanned by x and y. x and y are called <u>(mutually)</u> <u>orthogenal</u> if x.y = o.

In three dimensions the set of points $x = (x_1, x_2, x_3)$ which lie on a plane satisfy an equation of the form $a_1x_1 + a_2x_2 + a_3x_3 = \alpha$ which may be written as $a \cdot x = \alpha$ where $a = (a_1, a_2, a_3)$ is a normal vector to the plane. Accordingly, the set of points x in $\mathbb{R}^k$ which satisfy

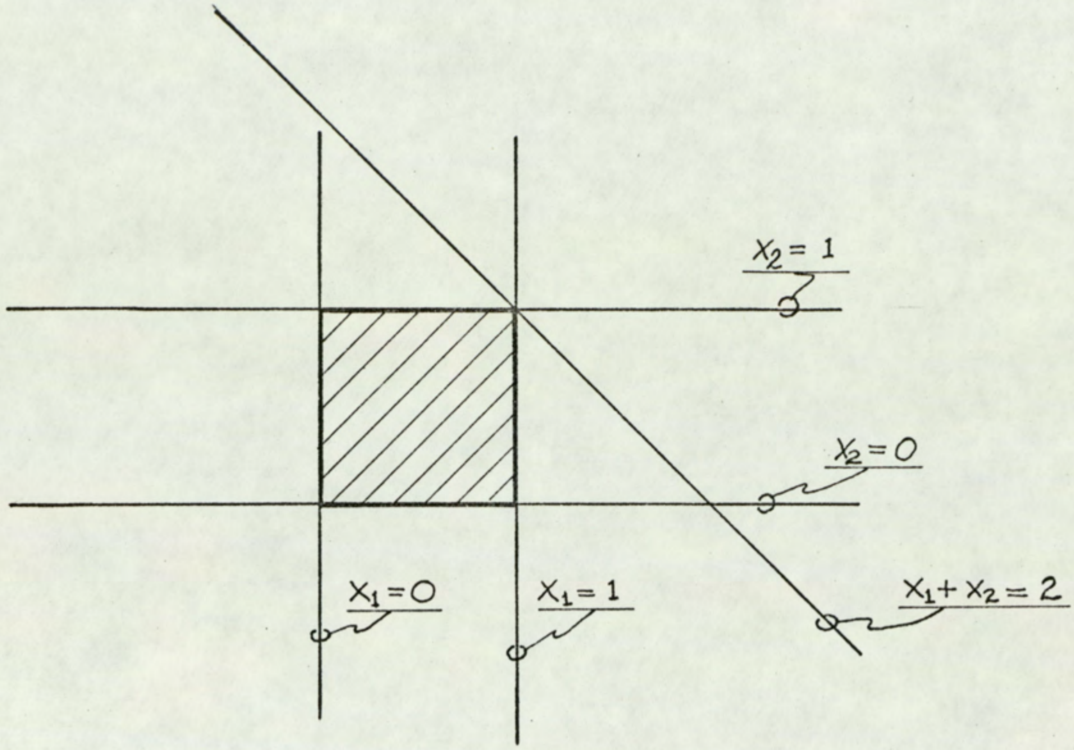
$$a \cdot x = \alpha \qquad \text{(I.4)}$$

for a fixed vector a in $\mathbb{R}^k$ and the real scalar α is called a <u>hyperplane</u>. The vector a is called a <u>normal vector to the hyperplane</u>. A hyperplane has k - 1 degrees of freedom and a hyperplane which passes through the origin is the (k - 1) - dimensional subspace spanned by the vectors orthogonal to a. The hyperplane in $\mathbb{R}^k$ defined by (I.4) divides the space into two <u>half-spaces</u>; a <u>positive half-space</u> consisting of all points x with $a \cdot x > \alpha$ and a <u>negative half-space</u> consisting of those points x with $a \cdot x < \alpha$.

There is an important connection between convex sets and hyperplanes which we shall use repeatedly in this investigation; namely, a closed convex set is the intersection of all the closed half-spaces which contain it. This implies that if C is a closed convex set and x^o is a point not in C then there is a closed half-space which contains C but does not contain x^o, analytically this means that there exists a vector a and a scalar α such that $a \cdot x \geqslant \alpha$ for all x in C but $\alpha > a \cdot x^o$. If no smaller α would suffice, i.e. there is an $\bar{x}$ in C with

$a.\bar{x} = \alpha$, then the hyperplane is called, for obvious reasons, a tangent, or supporting, hyperplane of the convex set C.

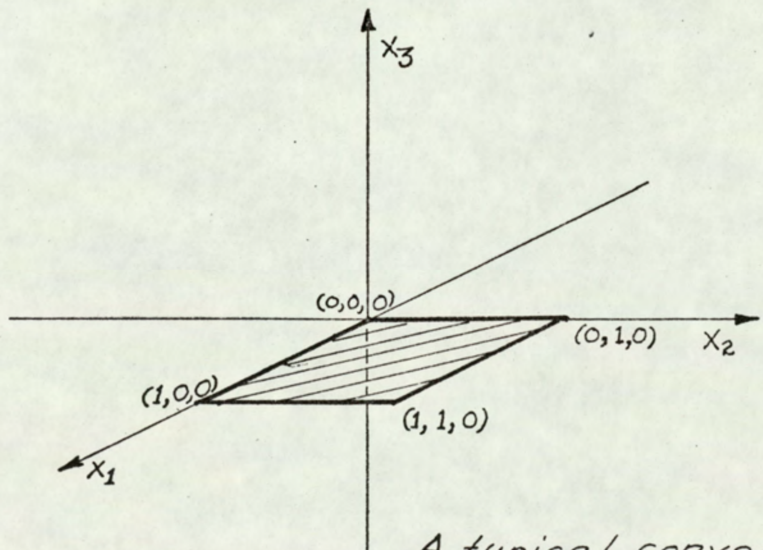
In $\mathbb{R}^2$ the hyperplanes are lines. The supporting hyperplanes for the unit square in $\mathbb{R}^2$ are the lines $x^1 = o$, $x^1 = 1$, $x^2 = o$, $x^2 = 1$, and any line passing through a vertex and containing no other point of the square. The supporting hyperplanes of the unit sphere in $\mathbb{R}^3$ (all x with $\|x\| = 1$) are those planes tangent to the sphere. If later arguments involving convex sets and hyperplanes become confusing to the reader, he is urged to consider the unit square in $\mathbb{R}^2$ as a typical convex set with interior points and the equivalent set in $\mathbb{R}^3$ consisting of the points $x = (x_1, x_2, x_3)$ with $o \leq x_1 \leq 1$, $o \leq x_2 \leq 1$, $x_3 = o$ as the typical convex set without interior points (see Figure 3).



Some supporting hyper planes of the unit square in $\mathbb{R}^2$.

FIGURE 3

(a)



A typical convex set without interior.

(b)

Chapter One. Large scale behaviour.

The systems with which we are concerned in this investigation are those which can be modelled by linear, constant-coefficient differential equations of the form

$$\dot{x} = Ax + Gm. \tag{1.1}$$

Here $x = (x_1, \ldots, x_n)$ and $m = (m_1, \ldots, m_r)$ are points of the n-dimensional state space ($\mathbb{R}^n$) and the r-dimensional control space ($\mathbb{R}^r$), $0 < r \leq n$, respectively. A is the nxn <u>system matrix</u> and G is the nxr <u>gain matrix</u>. We assume here that x and m are functions of time (with no time-lag) and that the control function $m = m(t)$ takes on only those values which lie in a closed convex, bounded subset of $\mathbb{R}^r$ which contains the origin. This subset, the <u>control domain</u> will be denoted by $\mathscr{C}$. Such a (measurable) function taking values in $\mathscr{C}$ will be called an <u>admissible control function.</u> If the control system studied is of bang-bang type, where the admissible controls take on only certain discrete values, $\mathscr{C}$ will be taken as the convex hull of these points. The gain matrix G is assumed to be one-to-one on the set $\mathscr{C}$, i.e. if $m^1 \neq m^2$ then $Gm^1 \neq Gm^2$.

As stated in the introduction, our investigations will be based primarily in geometric considerations. We begin with the local behaviour described by equation (1.1).

A curve in $\mathbb{R}^n$ whose parametric equation is given by $x(t)$, $t_0 \leq t \leq t_1$, has a tangent vector at the point $x(t)$ given by $\frac{dx}{dt}(t) \equiv \dot{x}(t)$. Thus equation (1.1) has the geometric interpretation of prescribing the tangent vector to the system trajectory at the point x. To each control m in $\mathcal{C}$ there corresponds a possible tangent vector Ax + Gm depending upon the control chosen. Such a tangent vector is an <u>admissible tangent</u> for the system at x and the collection of all such admissible tangents defines the <u>tangent cone at x</u>. Any trajectory leaving the point x must have a right-hand derivative at x which is an admissible tangent and thus lies in the tangent cone at x.

The tangent cone at x is determined by the extreme points of the control domain, $\mathcal{C}$, as any m in $\mathcal{C}$ is a convex combination of extreme points of $\mathcal{C}$. For if $m = \sum_{j=1}^{k} \alpha_j m^j$, $0 \leq \alpha_j$, $\sum_{j=1}^{k} \alpha_j = 1$ then the tangent vector,

$$Ax + Gm = Ax + G\left(\sum_{j=1}^{k} \alpha_j m^j\right) =$$

$$= \sum_{j=1}^{k} \alpha_j \left(Ax + Gm^j\right),$$

and is seen to be a convex combination of the points $Ax + Gm^j$, j =1, ..., k. Figure 4 shows the tangent cone (not to scale) for a representative system at a selection of points.

In general the designer seeks to control the system at a desired reference point which may vary with time. Success in this implies certain relationships between the reference point, the gain matrix and the control domain. It is these relationships to which we now turn our attention.

For the purpose of these large-scale investigations we shall consider only a reference point which does not vary in time.

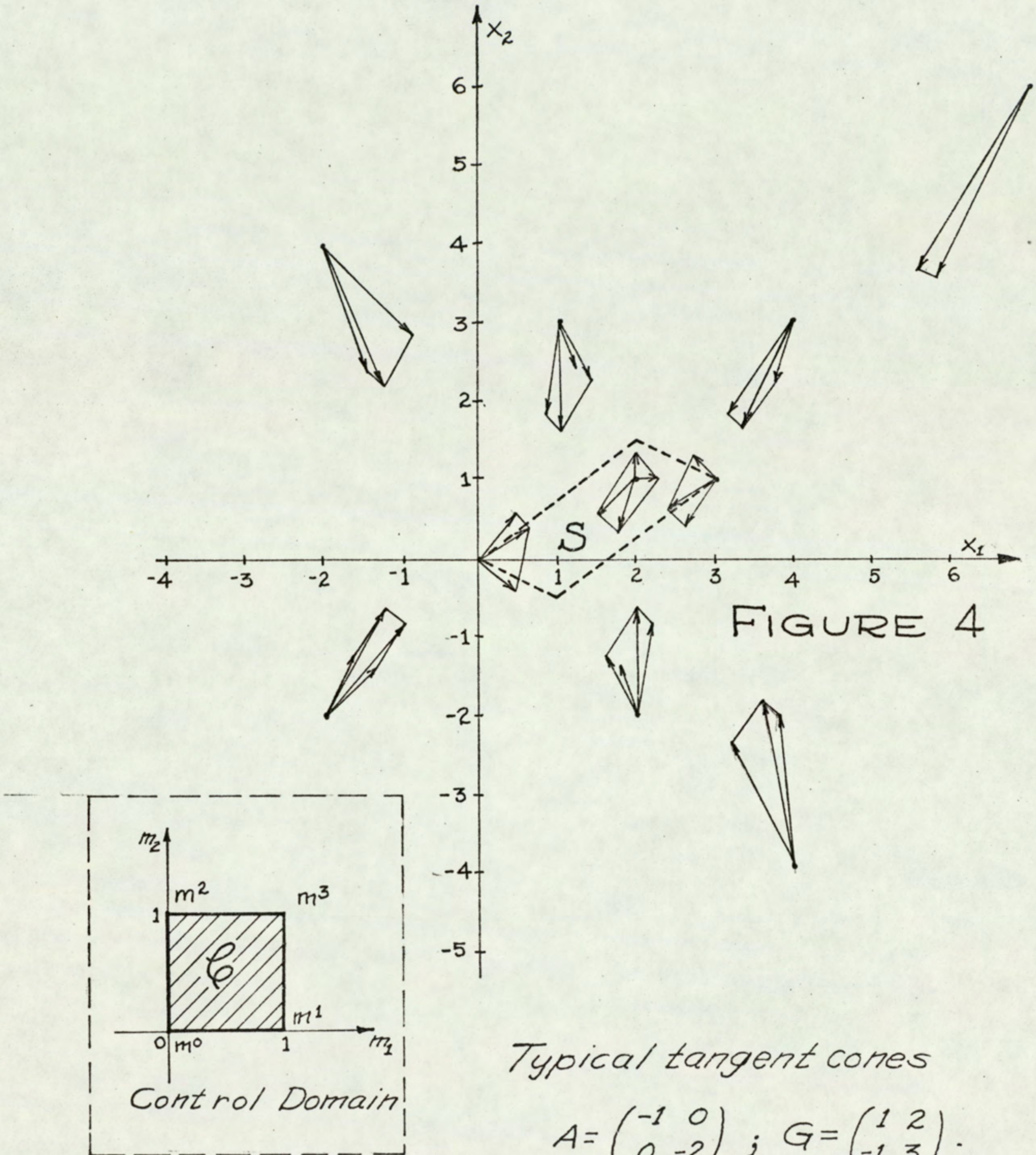


FIGURE 4

Typical tangent cones

$$A=\begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix} ; \quad G=\begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}.$$

This point will be denoted by $\bar{x}$. We shall see that the behaviour of the system depends upon the relative position of the point $-A\bar{x}$ and the set $G(\mathcal{C})$ which consists of all the points Gm with m in $\mathcal{C}$. $G(\mathcal{C})$ is a closed, bounded and convex subset of $\mathbb{R}^n$. The set consisting of those x in $\mathbb{R}^n$ with $-Ax$ in $G(\mathcal{C})$ will be denoted by S. Of course, if A is an invertible matrix ($\det A \neq 0$), $S = -A^{-1}G(\mathcal{C})$ and consists precisely of those $x = -A^{-1}Gm$ for some m in $\mathcal{C}$, S will then also be closed, bounded and convex. If A is not invertible S will be closed and convex but not bounded for, if x is in S and $Az = 0$, with $z \neq 0$, then $x+tz$ is also in S for $-\infty < t < \infty$.

Examples of the set S are given in Figures 5, 6, 7.

The equation

$$A\bar{x} + Gm = 0 \qquad (1.2)$$

will have a solution $\bar{m}$ in $\mathcal{C}$, if and only if, $-A\bar{x} = G\bar{m}$, i.e. $\bar{x}$ is in S. We now consider

<u>Case I</u> <u>$-A\bar{x}$ not in $G(\mathcal{C})$ ($\bar{x}$ not in S).</u>

Since $G(\mathcal{C})$ is closed and convex there is a supporting hyperplane $a.x = \alpha$ separating it from $-A\bar{x}$. If we choose a as the inner unit normal we have

$$a.Gm \geqslant \alpha > \alpha - \varepsilon > a.(-A\bar{x})$$

or

$$a.(A\bar{x} + Gm) > \varepsilon \qquad (1.3)$$

for all m in $\mathcal{C}$ and some $\varepsilon > 0$.

From (1.1) and (1.3) we find

$$a.\dot{x} = a.(Ax + Gm) + a.(A\bar{x} - A\bar{x}) =$$

$$= a.A(x-\bar{x}) + a.(A\bar{x} + Gm) >$$

$$> a.A(x-\bar{x}) + \varepsilon \qquad (1.4)$$

for all m in $\mathcal{C}$.

If x is sufficiently near $\bar{x}$ say $|x - \bar{x}| < \delta$ then $|a.A(x-\bar{x})| < \varepsilon/2$ and (1.4) becomes

$$a.\dot{x} > \varepsilon/2 \quad , \quad |x - \bar{x}| < \delta . \tag{1.5}$$

If the system is in this neighbourhood at time t, and remains inside until t_2 we can integrate (1.5) to find

$$a.(x(t_2) - x(t_1)) > \varepsilon/2 \cdot (t_2 - t_1). \tag{1.6}$$

Equation (1.6) clearly shows that the system cannot be maintained arbitrarily close to the desired reference point $\bar{x}$, for if it could we should find $|x(t) - \bar{x}| < \delta$ for arbitrarily long time periods. Thus the inequality (1.6) would hold with arbitrarily large right-hand side. This implies that the component of $x(t_2) - x(t_1)$ in the direction of a (and thus the length of $x(t_2) - x(t_1)$) becomes arbitrarily large which is impossible if the system remains within distance δ of $\bar{x}$.

This analysis shows that if the gain matrix and the reference point are not correctly matched so as to include $-A\bar{x}$ in $G(\mathcal{C})$ the designer must be prepared to accept certain minimal errors, or oscillations, in the system. The system must move a certain distance from $\bar{x}$ before it can possibly be brought back to the reference condition.

In certain cases much stronger conclusions can be drawn. If the unit vector a of the separating hyperplane can be chosen an eigenvector of A^T (eg. if A is diagonal the basis vectors are eigenvectors) with a real eigenvalue λ then $a.Ax = \lambda\, a.x$ and (1.4) becomes

$$a.\dot{x} > \lambda\, a.(x-\bar{x}) + \varepsilon \; . \tag{1.7}$$

If P_+ denotes the positive half-space of points x with $a.(x-\bar{x}) \geqslant -\varepsilon/2|\lambda|$, if $\lambda \neq o$, or $a.(x-\bar{x}) \geqslant -\varepsilon/2$, if $\lambda = o$, then any trajectory through a point of P_+ satisfies (from 1.7)

$$a.\dot{x} > \varepsilon/2.$$

If the system remains in P_+ throughout the time interval $t_o \leqslant t \leqslant t_1$ we integrate as before to find

$$a.(x(t_1) - x(t_o)) > \varepsilon/2\,(t_1 - t_o).$$

We therefore conclude that either the trajectory remains in P_+ and becomes unbounded or it leaves P_+. Once having left P_+ it can never re-enter as, at a boundary point, all admissible tangents have a positive component in the direction of a and hence a trajectory can cross it in at most one direction. The point $\bar{x}$ is in P_+ $(a.(x-\bar{x}) = o)$ so that a bounded trajectory (which must leave P_+) cannot be brought nearer $\bar{x}$ than the point $\bar{x} - \varepsilon/2|\lambda|\, a$ (or $\bar{x} - \varepsilon/2a$ if $\lambda = o$) which is the point on the boundary of P_+ nearest $\bar{x}$. Thus, in this instance, a trajectory either becomes unbounded or, if bounded, is at least a distance $\varepsilon/2|\lambda|$ ($\varepsilon/2$ if $\lambda = o$) from the reference point for some $\varepsilon > o$.

This analysis shows that each eigenvector of A^T corresponding to a real eigenvalue can determine a <u>region of no-return</u> in the state space of the system. This behaviour was observed in the one-dimensional case considered in [1], where in Figure 1 the regions of no-return are called the <u>on-zone</u> and the <u>off-zone</u>.

Equation (1.4) has a geometric interpretation. If we recall that the angle between any vectors x,y of $\mathbb{R}^n$ is $\cos^{-1}\left(x.y / |x|\cdot|y| \right)$ (i.e. $\cos(x,y) = \frac{x.y}{|x|\cdot|y|}$) then (1.4) furnishes a limitation on the angle between a and any admissible tangent. Indeed, if control m is chosen at $\bar{x}$, (1.4) becomes

$$a.(A\bar{x} + Gm) = |A\bar{x} + Gm| \cos(\dot{x}, a)$$

and if $d(\bar{x})$ represents the maximum of the numbers $|A\bar{x} + Gm|$ for m in $\mathcal{C}$ ($d(\bar{x})$ is the maximum distance from $-A\bar{x}$ to G ($\mathcal{C}$)) we have

$$\cos(\dot{x}, a) > \varepsilon / d(\bar{x}) > o.$$

Therefore the angle between any admissible tangent and the vector a is $< \cos^{-1}(\varepsilon / d(\bar{x})) < \pi/2$ and is inversely proportional to the maximum distance from $-A\bar{x}$ to G($\mathcal{C}$). This is clearly illustrated in Figure 4 and in the following examples.

Example I. $A = \begin{pmatrix} -1 & o \\ o & -2 \end{pmatrix}$, $G = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$, $\mathcal{C}$ is the unit square in $\mathbb{R}^2$, i.e. the set consisting of all m = (m_1, m_2) with $o \leq m_1 \leq 1$, $o \leq m_2 \leq 1$. The matrices A and G both have rank 2, A is invertible and $S = -A^{-1}G$ ($\mathcal{C}$). The basis vectors e^1 = (1,o) and e^2 = (o,1) are eigenvectors of A (=A^T) corresponding to the eigenvalues -1 and -2,respectively. The pertinent geometric properties of this system are shown in Figure 5a.

Example 2.

$$A = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 2 \\ 2 & o \end{pmatrix}, \quad \mathcal{C} \text{ as in Example I.}$$

This example differs from the preceeding in that A is not invertible and has rank 1. The eigenvalues of A are $\lambda_1 = 0$, $\lambda_2 = 2$ with corresponding eigenvectors $a_1 = (-1/\sqrt{2}, 1/\sqrt{2})$, $a_2 = (1/\sqrt{2}, 1/\sqrt{2})$ of A^T (=A). The admissible tangent cones are drawn (not to scale) at representative points, the set S is shown, and the regions of no-return are sketched in Figure 5b.

Example 3.

$$A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad G = \begin{pmatrix} 1 \\ 2 \end{pmatrix},$$ $\mathcal{C}$ is the set of one-dimensional controls $0 \leq m \leq 1$. A is invertible but has no real eigenvalues while the rank of G = 1 which is less than the dimension of the state space. The set $S = -A^{-1}G(\mathcal{C})$. This is shown in Figure 5c.

Example 4.

$$A = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 0 \\ 0 & 4 \\ -1 & 0 \end{pmatrix},$$ $\mathcal{C}$ is the unit square in $\mathbb{R}^2$ as in Examples 1 and 2. A is invertible with eigenvalues -1, -2, 1 and corresponding eigenvectors $e^1 = (1,0,0)$, $e^2 = (0,1,0)$, $e^3 = (0,0,1)$. The set S is shown in Figure 6 but, as in most higher dimensional cases, it may be easier to solve the equation $Ax + Gm = 0$ to find that

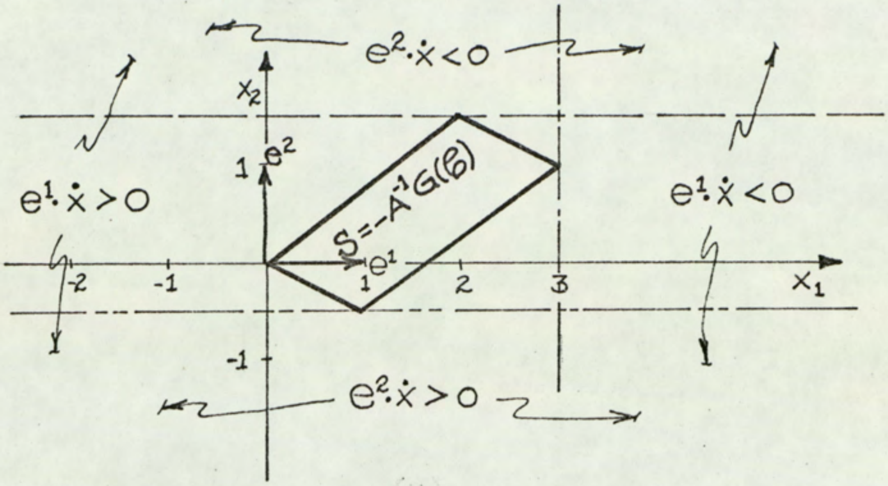
$$m_1 - x_1 = 0 \qquad 0 \leq m_1 \leq 1$$

$$4m_2 - 2x_2 = 0 \qquad 0 \leq m_2 \leq 1 \;..$$

$$x_3 - m_1 = 0$$

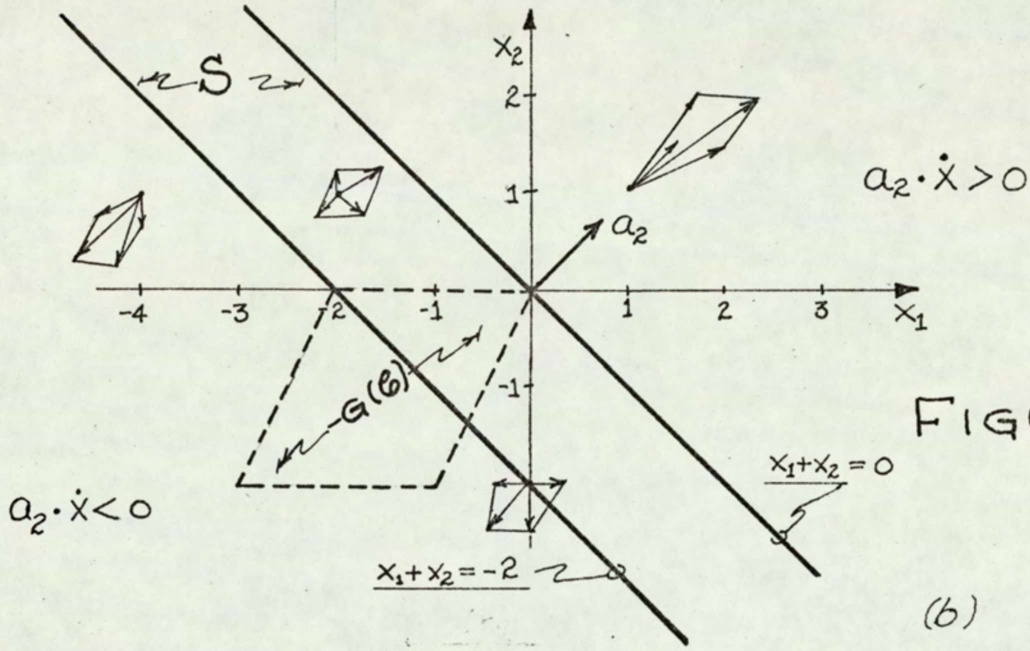
Hence $0 \leq x_1 = x_3 \leq 1 \qquad x_2 = 2m_2$, so $0 \leq x_2 \leq 2$, if (x_1, x_2, x_3) is in S.

The regions of no return are found from the conditions:



$e^2 \cdot \dot{x} < 0$
x_2
e^2
$e^1 \cdot \dot{x} > 0$
$S = -A^{-1}G(\mathcal{B})$
$e^1 \cdot \dot{x} < 0$
e^1
x_1
$e^2 \cdot \dot{x} > 0$

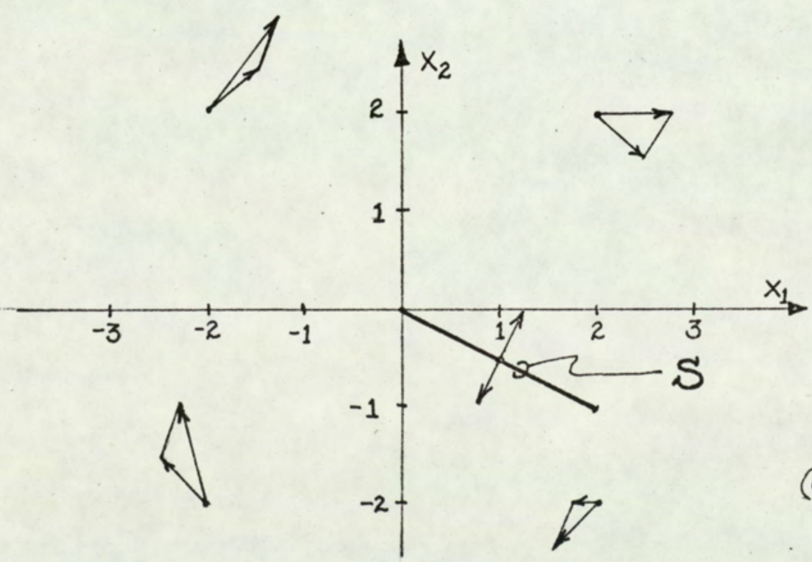
(a)



x_2
S
$a_2 \cdot \dot{x} > 0$
a_2
x_1
$-G(\mathcal{B})$
$x_1 + x_2 = 0$
$a_2 \cdot \dot{x} < 0$
$x_1 + x_2 = -2$

FIGURE 5

(b)



x_2
x_1
S

(c)

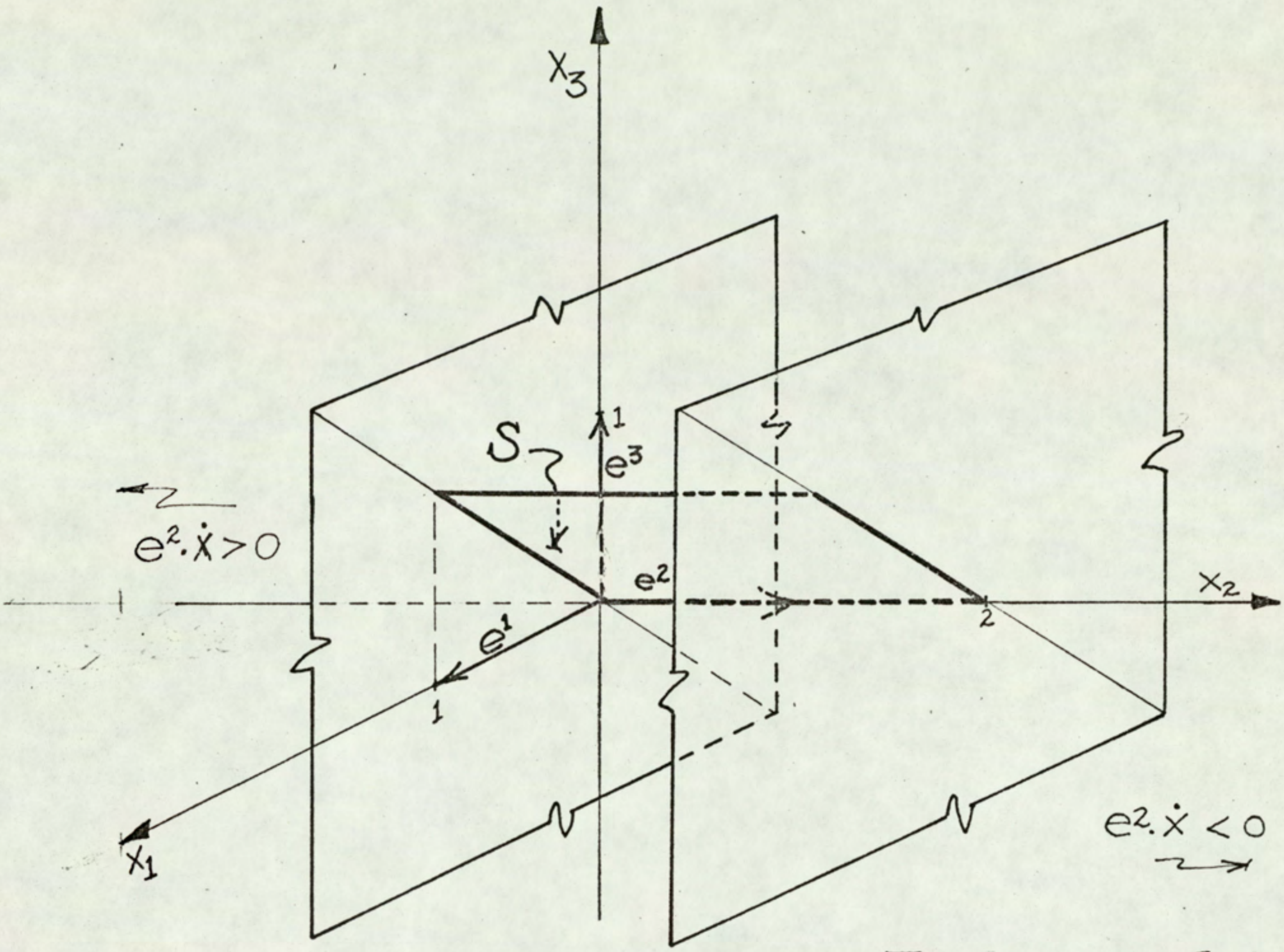


FIGURE 6

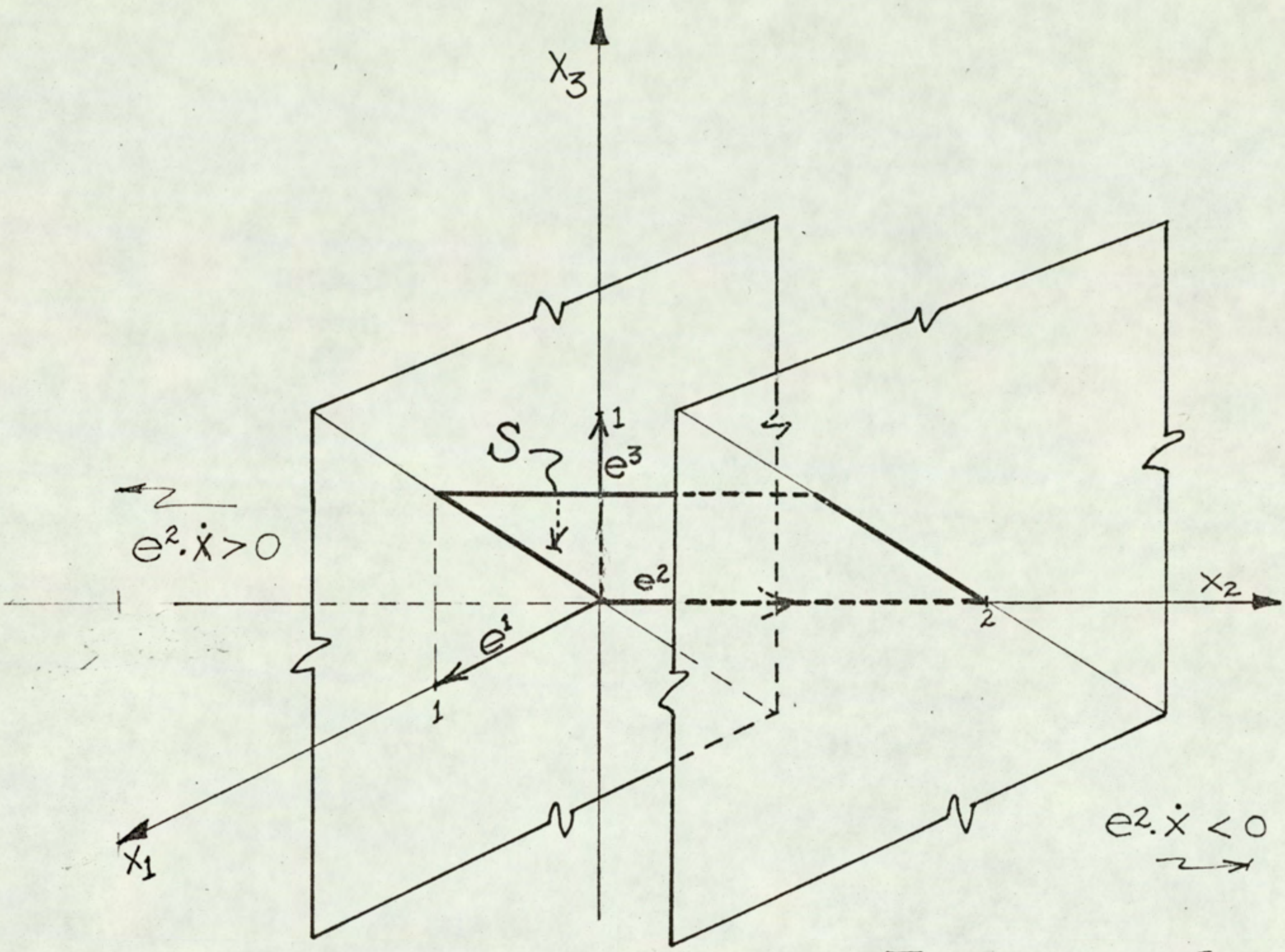


FIGURE 6

(a) $e^1 . \dot{x} > o$ if $x_1 < o$ (b) $e^2 . \dot{x} > o$ if $x_2 < o$

$< o$ if $x_1 > 1$ $< o$ if $x_2 > 2$

(c) $e^3 . \dot{x} > o$ if $x_3 > 1$

$< o$ if $x_3 < o$.

The second pair of these half-spaces is shown in Figure 6.

Keeping these examples in mind let us return to

Case 2. $-A\bar{x}$ in $G(\mathcal{C})$ ($\bar{x}$ in S).

First we suppose that $\bar{x}$ is on the boundary of S, which implies $-A\bar{x}$ is on the boundary of $G(\mathcal{C})$. Through the boundary point $-A\bar{x}$ there passes a supporting hyperplane to $G(\mathcal{C})$ with inner unit normal a such that

$$a.(-A\bar{x}) = \text{minimum } a.Gm \quad , \quad m \text{ in } \mathcal{C} . \tag{1.8}$$

This certainly implies

$$a.(A\bar{x} + Gm) \geqslant o \qquad \text{all } m \text{ in } \mathcal{C} \tag{1.9}$$

and, because $\bar{x}$ is in S, $-A\bar{x} = G\bar{m}$ for some $\bar{m}$ in $\mathcal{C}$, and hence equality can hold in (1.9).

Consider for the moment the problem of maintaining the system at $\bar{x}$. Since $A\bar{x} + G\bar{m} = o$ constant application of the control $\bar{m}$ keeps the system at the desired state ($\dot{x} = o$). However, in the case of multiple on-off controllers, the control $\bar{m}$ may not be a realizable control, it may not be one of the extreme points of $\mathcal{C}$ which represent the actual available controls. If this is the case $\bar{m}$ will still be a convex combination of extreme points, i.e. $\bar{m} = \sum_{j=1}^{k} \alpha_j m^j$, where $o < \alpha_j$ and $\sum_{j=1}^{k} \alpha_j = 1$.

Further, from (1.8),

$$a.\,(-A\bar{x}) = a.G\bar{m} = a.G\left(\sum_{j=1}^{k} \alpha_j m^j\right) =$$

$$= \sum_{j=1}^{k} \alpha_j a.Gm^j \geqslant \sum_{j=1}^{k} \alpha_j a.\,(-A\bar{x}) =$$

$$= a.\,(-A\bar{x}).$$

Now if we have $a.(-A\bar{x}) < a.Gm^j$ for even one j the inequality holds above and we conclude the impossible, $a.(-A\bar{x}) > a.(-A\bar{x})$. Therefore we must have $a.(-A\bar{x}) = a.Gm^j$, or

$$a.(A\bar{x} + Gm^j) = o \ , \quad j = 1, \ldots, k. \qquad (1.10)$$

Thus we see that the admissible tangents $A\bar{x} + Gm^j$ are all perpendicular to the vector a and thus the tangent cone at $\bar{x}$ is not restricted by an angle less than $\pi/2$ as occurred in Case 1.

We observe as well that

$$o = A\bar{x} + G\bar{m} = A\bar{x} + \sum_{j=1}^{k} \alpha_j Gm^j =$$

$$= \sum_{j=1}^{k} \alpha_j \,(A\bar{x} + Gm^j) \qquad (1.15)$$

which will, in Chapter 2, provide a discontinuous open loop control maintaining the system arbitrarily close to $\bar{x}$ for finite time intervals. This last statement is intuitively clear when the sum (1.15) contains only two terms on the right for then

$$o = \alpha_1 \,(A\bar{x} + Gm^1) + \alpha_2 \,(A\bar{x} + Gm^2)$$

and we see the two admissible tangents $A\bar{x} + Gm^1$ and $A\bar{x} + Gm^2$ are oppositely directed and therefore a "chatter" control, alternating between m^1 and m^2 over proper time intervals, should maintain the system as near $\bar{x}$ as desired.

This discussion has shown that for $\bar{x}$ on the boundary of S there exists a possibility of maintaining the system at this desired point. We have given no indication of the start-up point from which the system can be brought to $\bar{x}$. The examples we have shown indicate that there is little which can be said, in general. Indeed, Example 2 shows that points arbitrarily close to S, but not in it, are always driven to infinity, while in Example 1, with the stable matrix A, all points can be brought into a neighbourhood of the origin.

The natural continuation of our investigation would appear to be the investigation of system behaviour near a point $\bar{x}$ in the interior of S. However, the results show that the system is more complex than this. While the behaviour of the system depends upon A and G the dependence is too subtle to be distinguished by considering only A and G. The relationship between G and the powers of A, A^2, A^3, ..., A^{n-1} must also be considered. Such an investigation will provide considerable information even when S has no interior points.

The concept involved here is known as <u>controllability</u> in the literature. The system (1.1) is said to be <u>controllable</u> if it is possible to transfer the system between any given pair of points by a bounded measurable control function. This is a global property and no restraints are imposed upon the controls other than those mentioned. The system will be controllable if, and only if, the rank of the compound matrix $(G, AG, A^2G, \ldots, A^{n-1}G)$ is equal to n (see (2),page 81).

Global properties are not of primary concern in this investigation. However, the <u>controllability matrix</u>

$(G, AG, \ldots, A^{n-1}G)$ does provide considerable information of the local nature we do require.

Our chief interest at this point is in describing those start-up points which allow the system to be directed to a reference point $\bar{x}$ in S. It is easier (and sufficient) to study the points which are attainable from a given point x^o, i.e. a point y is attainable from x^o if the system can be transferred from x^o to y by an admissible control. The set of attainable points of x^o will be denoted by $K(x^o)$ and those points attainable in time t will be denoted by $K(x^o, t)$. It is sufficient to study this set of attainability of x^o because we have restricted our investigation to constant coefficient matrices.

For such systems the set of attainability of a point x has the same general properties (eg. the same dimension, being convex being open, or closed) as the set of points which can be transferred to x. Indeed, if the system (1.1) can be transferred from x^o to x^1 via the admissible control $m(t)$, $o \leq t \leq t_1$, then the system

$$\dot{x} = - Ax - Gm \qquad (1.16)$$

can be transferred from x^1 to x^o by the admissible control $m^*(t) = m(t_1 - t)$, $o \leq t \leq t_1$. The systems (1.1) and (1.16) are structurally the same ((1.16) is (1.1) with time reversed), in particular they are either both controllable or both uncontrollable, which will be the crucial issue below.

The set of attainability of x^o, $K(x^o)$, consists of all points (from Appendix A)

$$x(t,m) = e^{At}x^o + \int_o^t e^{A(t-s)}\, Gm(s)ds \qquad (1.17)$$

where $o \leq t < \infty$ and m runs through all admissible controls. In particular, $K^o(x^o, t_1)$ consists of all points $x(t_1, m) = e^{At_1}x^o + y(m)$ where y(m) is a point in the set Y consisting of the points

$$y(m) = \int_o^{t_1} e^{A(t_1-s)}\, Gm(s)ds$$

for all admissible controls m.

The set $K(x^o, t_1) = e^{At_1}x^o + Y$ is a translate of Y (see Figure 7) and therefore its overall properties are determined by those of Y.

The general theory ([2], p.69) states that Y is convex, closed and bounded. Since o is in $\mathcal{C}$, Y contains the origin in $\mathbb{R}^n$. We want to determine conditions sufficient to guarantee that Y has interior points. To do this it is clearly enough to ensure Y is an n-dimensional subset of $\mathbb{R}^n$.

If Y is contained in some n-1 dimensional subspace there exists a non-zero vector v (a normal to the subspace) such that

$$v.y(m) = \int_o^{t_1} v.e^{A(t_1-s)}Gm(s)ds = o$$

for all admissible controls. Choosing an arbitrary m^* in $\mathcal{C}$ we define the admissible control

$$m(s) = m^* \qquad o \leqslant \tau_1 < s < \tau_2 \leq t_1$$
$$= o \qquad \text{otherwise}$$

and conclude

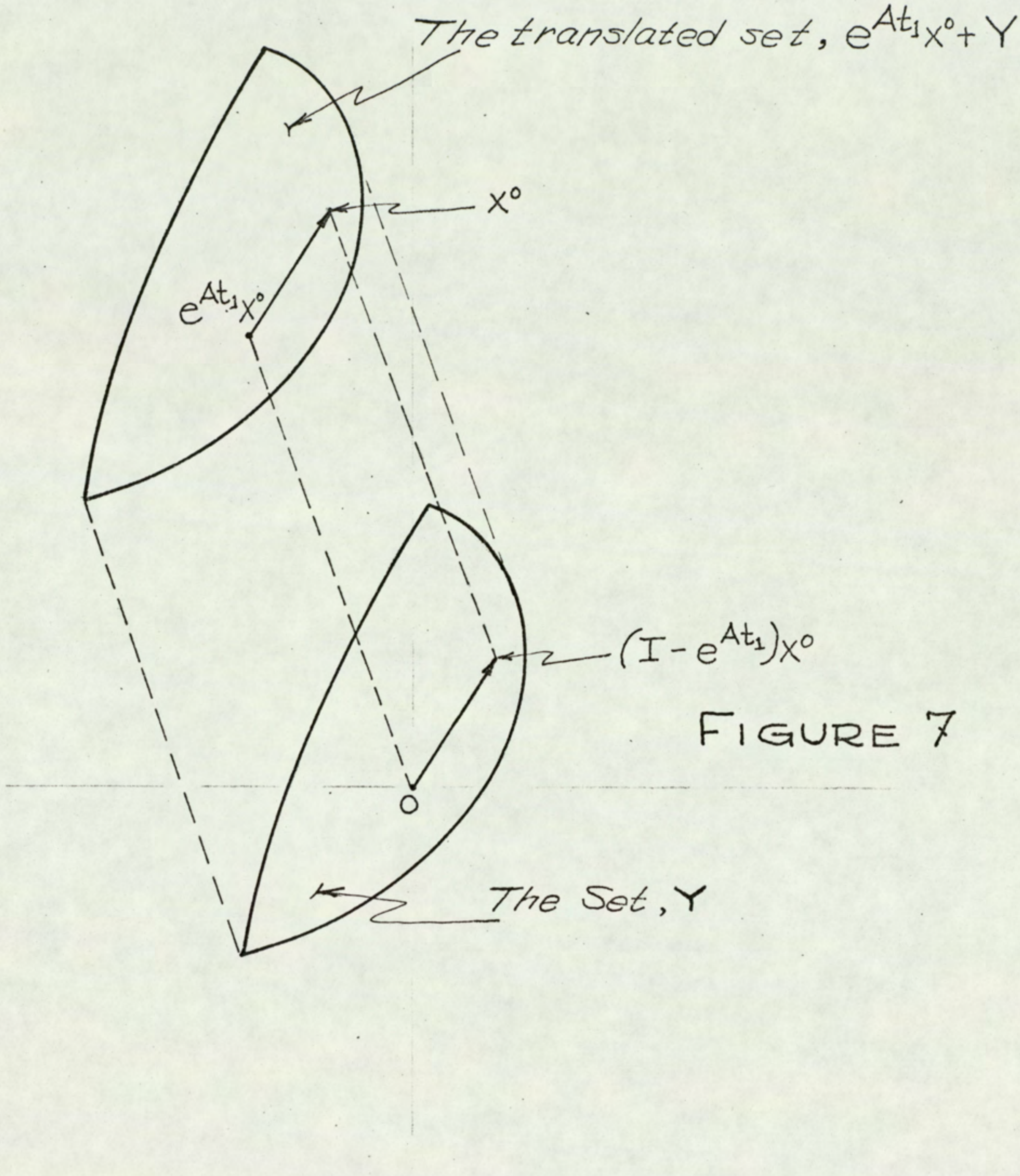


FIGURE 7

$$\int_{\tau_1}^{\tau_2} v.e^{A(t_1-s)}\ Gm^* ds = o. \tag{1.19}$$

Equation (1.19) can only hold for arbitrary τ_1, τ_2 and m^* if $v.e^{A(t_1-s)}Gm^* \equiv 0$ for $o \leq s \leq t_1$, and m^* in $\mathcal{C}$. If $v.e^{A(t_1-s)}G$ vanishes identically for $o \leq s \leq t_1$ we can differentiate n times and set $s = t_1$ to obtain

$$v.G = v.AG = \ldots = v.A^{n-1}G = o \tag{1.20}$$

which shows that the matrix $(G, AG, \ldots, A^{n-1}G)$ has rank less than n, i.e. the system (1.1) is not controllable. On the other hand, if $v.e^{A(t_1-s)}G$ does not vanish identically we conclude there exists a non-zero vector u in $\mathbb{R}^r$ ($u = v.e.^{A(t_1-s_o)}G$, for some s_o) such that $u.m^*=o$ for all m^* in $\mathcal{C}$. Thus $\mathcal{C}$ must lie in some subspace of dimension $\leq r-1$ in $\mathbb{R}^r$ and hence can have no interior points.

We have just shown:

<u>If (1) the system is controllable (rank $(G,AG,\ldots,A^{n-1}G)=n$)</u>

<u>and</u>

<u>(2) $\mathcal{C}$ has interior points (has dimension r in $\mathbb{R}^r$),</u>

then

<u>Y has dimension n in $\mathbb{R}^n$ and thus contains interior points.</u>

If the rank $(G, AG, \ldots, A^{n-1}G) < n$, Y will always have dimension $< n$, but if (1) holds there may be various control domains of different dimensions $< r$ for which Y still has dimension n. In fact, for most controllable systems, the designer can change his control methods by selecting a proper vector g and replacing (1.1) with the one-dimensional input system

$$\dot{x} = Ax + g\mu \qquad (1.21)$$

where μ now ranges over a convex subset of $\mathbb{R}$ ($\alpha \leq 0 \leq \mu \leq \beta$). The system (1.21) will be controllable provided the vectors $g, Ag, \ldots, A^{n-1}g$ are linearly independent. The existence of such a vector g is discussed in [2], page 86, and should it exist, it may be possible to modify the gain matrix G and/or the control domain $\mathcal{C}$ of (1.1) in such a way that $g = Gm^*$ for some m^* in $\mathcal{C}$. In this instance the one-dimensional subset μm^*, $\alpha \leq \mu \leq \beta$, of $\mathbb{R}^r$ is sufficient to insure that Y has dimension n.

If Y has dimension n it must, as it is convex, contain interior points. It is important for both our present investigation and the later discussion of optimization to describe those points which are interior points of Y.

A point y^o is a boundary point of the convex set Y if, and only if, there is a supporting hyperplane of Y through y^o. That is, there exists a (non-zero) vector a with

$$a.y^o = \text{maximum } a.y, \ y \text{ in } Y.$$

Hence

$$a.y^o = \int_0^{t_1} a.e^{A(t_1-s)}Gm^o(s)ds \geq \int_0^{t_1} a.e^{A(t_1-s)}Gm(s)ds \qquad (1.22)$$

where m^o is the control function corresponding to y^o and m is an arbitrary admissible control. Suppose that on the interval $\tau_1 < s < \tau_2$, the points $m^o(s) = m^I$ where m^I is an interior point of $\mathcal{C}$. Since m^I is in the interior of $\mathcal{C}$ there is a control m^* such that

$$a.e^{A(t_1-s)}G.m^I < a.e^{A(t_1-s)}Gm^*$$

on a (perhaps) smaller interval $\tau_1 \leq \tau_1' \leq s \leq \tau_2' \leq \tau_2$. If we define the admissible control

$$\underline{m}(s) = m^* \qquad \tau_1' \leqslant s < \tau_2'$$
$$= m^o(s) \qquad \text{otherwise}$$

we find

$$a.y^o = \int_o^{t_1} a.e.^{A(t_1-s)} G m^o(s) ds < \int_o^{t_1} a.e.^{A(t_1-s)} G \underline{m}(s) ds$$

which is a contradiction of (1.22). Hence y^o cannot be a boundary point of Y if its corresponding control function takes on an interior value of $\mathcal{C}$ over an interval. A more precise analysis of this situation using the theory of Lebesgue measure yields:

A point $y(\underline{m})$ is on the boundary of Y if, and only if, there is a vector a in $\mathbb{R}^n$ such that

$$a.e^{A(t_1-s)} G\underline{m}(s) = \text{maximum } a.e^{A(t_1-s)} Gm, \ m \text{ in } \mathcal{C}$$

for $o \leqslant s \leqslant t_1$, except possibly on subsets of (o, t_1) having Lebesgue measure zero.

This is perhaps the simplest version of the maximum principle of Pontryagin.

This analysis has shown that for a control, which is identically equal to the fixed m^* in the interior of $\mathcal{C}$, the corresponding $y(m^*)$ lies in the interior of Y.

Let us now return to $K(x^o, t_1)$ the set attainable from x^o in time t_1. x^o is in $K(x^o, t_1) = e^{At_1}x^o + Y$ if $(I - e^{At_1})x^o$ is in Y, (see Figure 7). This will occur when $-Ax^o$ is in $G(\mathcal{C})$. Further x^o will be an interior point of $K(x^o, t_1)$ (and thus of $K(x^o)$) when $(I-e^{At_1})x^o$ is an interior point of Y.

We have just seen that $(I-e^{At_1})x^o$ will be an interior point of Y if $-Ax^o = Gm^o$ for some m^o in the interior of $\mathcal{C}$, (since constant m^o control leads to $(I-e^{At_1})x^o$), thus: (see [2] page 84):

<u>If the system $\dot{x} = Ax + Gm$ is controllable (rank $(G, AG, \ldots, A^{n-1}G) = n$) and $-Ax^o = Gm^o$, where m^o is a fixed interior point of $\mathcal{C}$, then x^o is an interior point of $K(x^o)$, the set of points attainable from x^o.</u>

By considering this last result for the system (1.16) where time has been reversed we can conclude:

<u>If the system $\dot{x} = Ax + Gm$ is controllable and $-Ax^o = Gm^o$, where m^o is a fixed interior point of $\mathcal{C}$, then x^o is an interior point of the set consisting of all points controllable to x^o.</u>

We have let S denote the set of x with -Ax in $G(\mathcal{C})$, let us denote by S^o those points of S which correspond to the interior points of $\mathcal{C}$. We seek to show that the system is completely controllable within the set S^o, i.e. it can be transferred between any given pair of points by an admissible control function. To demonstrate this we take an arbitrary point x^o in S^o and let $S^o(x^o)$ be the set of points in S^o which can be reached via an admissible control in finite time (we will show $S^o(x^o) = S^o$).

$S^o(x^o)$ is a relatively open subset of S^o (see Introduction) for if x^1 is in $S^o(x^o)$ there is a neighbourhood of x^1 consisting of points attainable from x^1. If x^2 is in S^o and attainable from x^1 the control sequence $x^o \rightarrow x^1 \rightarrow x^2$ shows that x^2 is attainable from x^o. Thus x^1 has neighbourhood (relative to S^o) contained in $S^o(x^o)$ and so $S^o(x^o)$ is relatively open.

Assume now that there is y^o in S^o which is not attainable from x^o in finite time. Now, not only is y^o not in $S^o(x^o)$ but any z from which y^o is attainable is not in this set. If z were in $S^o(x^o)$ the control sequence $x^o \rightarrow z \rightarrow y^o$ would transfer the system from x^o to y^o in finite time which would mean y^o was in $S^o(x^o)$.

By considering the system (1.16) (or reversing time) we can show (by those arguments used for $S^o(x^o)$) that the set of points which are in S^o but are not attainable from x^o is also a relatively open subset of S^o. Hence S^o is divided by $S^o(x^o)$ into two relatively open sets having no points in common, x^o is in one, y^o is in the other.

Since S^o is easily seen to be convex it contains the line segment between x^o and y^o which may be parametrized by $z(t) = (1-t)x^o + ty^o$, $o \leq t \leq 1$. Now $z(o) = x^o$ is in $S^o(x^o)$ while $z(1) = y^o$ is not.

Let t_1 be the largest t with the property that for all $t < t_1$, $z(t)$ is in $S^o(x^o)$. [a] Any neighbourhood of the point $z(t_1)$ must contain points $z(t)$ for t both greater and smaller than t_1. Since $z(t_1)$ is in S^o it is either in $S^o(x^o)$ or it is not. If it is in $S^o(x^o)$ it has a relative neighbourhood contained in $S^o(x^o)$. This implies that so long as t is near t_1 , $z(t)$ is in $S^o(x^o)$ which contradicts the definition of t_1, for some of these t are greater than t_1. Similarly if $z(t_1)$ is not in $S^o(x^o)$ it has a neighbourhood containing no points of $S^o(x^o)$ which again contradicts the definition of t_1 (see Figure 8).

(a) Mathematically $t_1 = \inf\{ t: z(t) \text{ not in } S^o(x^o)\}$.

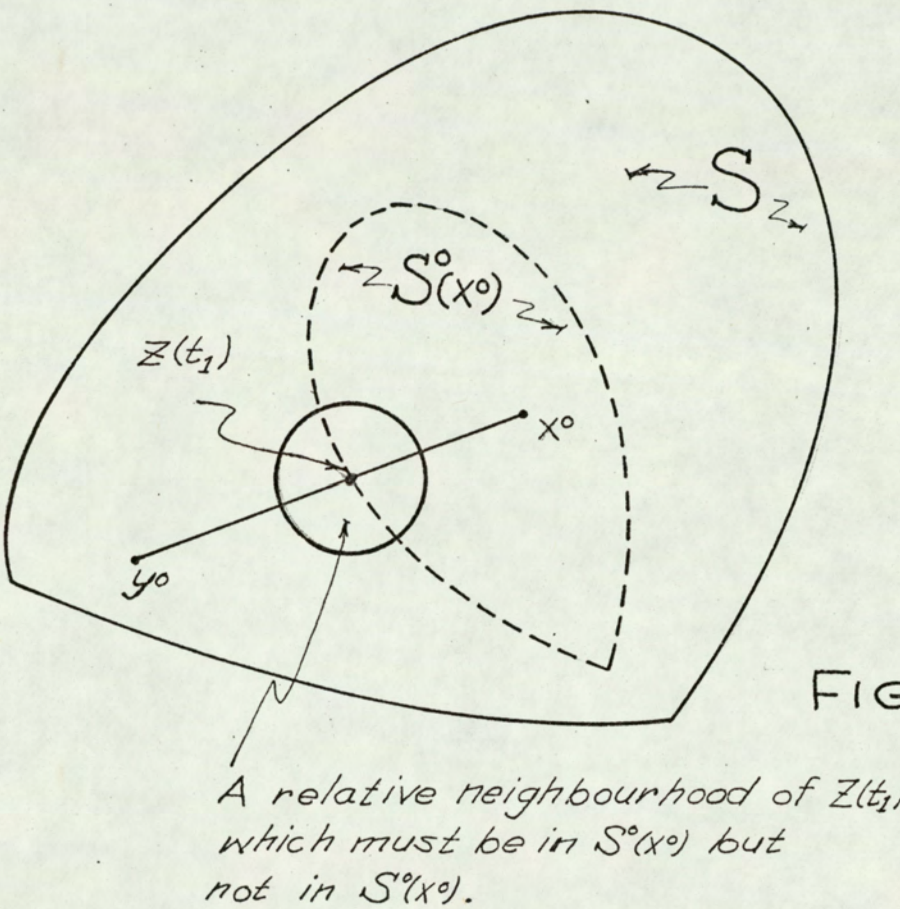

S
$S^\circ(x^\circ)$
$Z(t_1)$
x°
y°
A relative neighbourhood of $Z(t_1)$ which must be in $S^\circ(x^\circ)$ but not in $S^\circ(x^\circ)$.

FIGURE 8

This logical contradiction shows that our assumption of a point y^o in S^o but not in $S^o(x^o)$ was incorrect and that $S^o = S^o(x^o)$. Hence the system can be transferred via an admissible control between any two points of S^o. Further, since the set of attainable points from a given x^o in S^o is closed, a point which is a boundary point of S and therefore a limit point of points which are attainable from x^o is attainable, as well, from x^o. This would also mean it was attainable for the system (1.16), i.e. is controllable to x^o and hence to any other point of S. There is a further extension of this result. The <u>general bang-bang principle</u> ([2], page 79) states that if the system can be transferred from x to y by an admissible control taking values in $\mathcal{C}$ then it is also transferable from x to y by an admissible control which takes only values which are extreme points of $\mathcal{C}$, i.e. bang-bang controls.

We can summarize:

<u>If (1) rank $(G, AG, A^2G, \ldots, A^{n-1}G) = n$</u>

<u>(2) $\mathcal{C}$ has interior points</u>

<u>(3) S is the set of points with $-Ax$ in $G(\mathcal{C})$</u>

<u>then the system may be transferred between any given pair of points of S by a bang-bang control, i.e. an admissible control whose values are all extreme points of $\mathcal{C}$.</u>

In [1], page 28, Figure 1, the set S is called the cycling zone.

Let us now consider the situation which arises when (1) or (2) fail to hold. We can, without loss of generality, assume that (2) holds and (1) fails for if $\mathcal{C}$ has dimension $k < r$ we can choose an orthonormal basis $e^1, e^2, \ldots, e^k$ for $\mathcal{C}$ and let $\mathcal{C}^1$ be the set of $\mu = (\mu_1, \ldots, \mu_k)$ with $m = \sum_{i=1}^{k} \mu_i e^i$ in $\mathcal{C}$.

$\mathcal{C}^1$ is a convex and bounded, closed set of $\mathbb{R}^k$ with interior.

Further

$$Gm = \sum_{i=1}^{k} \mu_i G e^i = \sum_{i=1}^{k} g^i \mu_i = G \mu$$

where G^1 is the n x k matrix with column vectors $g^1 = Ge^1$, $g^2 = Ge^2, \ldots, g^k = Ge^k$. If $v.e^{A(t_1-s)}Gm \equiv o$ for all m in $\mathcal{C}$ we have $v.e^{A(t_1-s)}G^1\mu \equiv o$ all μ in $\mathcal{C}^1$, but since $\mathcal{C}^1$ has an interior, this implies $v.e^{(A(t_1-s)}\ G^1 \equiv o$ and therefore rank $(G^1, AG^1, \ldots, A^{n-1}G^1) < n$. We have thus transformed the system where (2) failed to another system where (2) holds and (1) fails. Accordingly we shall now consider a system for which rank $(G, AG, \ldots, A^{n-1}G) = k < n$.

For such a system there exist n-k distinct unit vectors $v^1, v^2, \ldots, v^{n-k}$ which satisfy

$$v^i.G = v^i.AG = \ldots = v^i.A^{n-1}G = o \quad i = 1,2,\ldots,n-k$$

consequently

$$v^i. \int_o^t e^{A(t-s)}Gm(s)ds = o \quad i = 1,2,\ldots,n-k \tag{1.23}$$

for all admissible controls (see Appendix A).

Consider a trajectory leaving a point x^o in S (suppose $-Ax^o = Gm^o$) given by

$$x(t,m) = e^{At}x^o + \int_o^t e^{A(t-s)}Gm(s)ds.$$

Then

$$v^i.x(t,m) = v^i.e^{At}x^o + v^i.\int_o^t e^{A(t-s)}Gm(s)ds =$$

$$= v^i.e^{At}x^o \qquad (1.24)$$

because of (1.23). If we differentiate (1.24) we find

$$\frac{d}{dt}\ v^i.x(t,m) = v^i.e^{At}Ax^o = -\ v^i.e^{At}Gm^o = o$$

since x^o is in S and hence, $v^i.x(t,m)$ is a constant equal to $v^i.x(o,m) = v^i.x^o$.

This analysis shows that any trajectory leaving a point x^o in S is constrained to lie in the hyperplanes

$$v^i.x = v^i.x^o \qquad i = 1,2,\ldots,n-k$$

therefore if y^o is also in S with $v^i.x^o \neq v^i.y^o$ it is impossible to transfer the system from x^o to y^o by an admissible control. On the other hand, if $v^i.x^o = v^i.y^o$, $i = 1,2,\ldots,n-k$ then x^o and y^o lie in a k-dimensional hyperplane. The set corresponding to Y above, i.e. the set of

$$y(m) = \int_o^{t_1} e^{A(t_1-s)}Gm(s)ds,$$

for all admissible controls m, is a k-dimensional set. If we restrict ourselves to this k-dimensional hyperplane and consider only k-dimensional neighbourhoods we can repeat our previous argument to show that the system can be transferred from x^o to y^o by an admissible control.

We can now generalize our previous result:

Suppose (1) $\mathcal{C}$ has an interior

(2) rank $(G, AG, \ldots, A^{n-1}G) = k$ and

(3) $v^1, \ldots, v^{n-k}$ are non-zero vectors with

$$v^i.G = v^i.AG = \ldots = v^i.A^{n-1}G = o, \quad i = 1, \ldots, n-k$$

(4) S is the set of x with $-Ax$ in $G(\mathcal{C})$.

Then the system $\dot{x} = Ax + Gm$ can be transferred between points x^o and y^o in S by admissible controls if, and only if,

$$v^i.x^o = v^i.y^o \quad i = 1, 2, \ldots, n-k.$$

By the general bang-bang principle if an admissible control exists transferring the system from x^o to y^o there also exists an admissible control taking on only those values which are extreme points of $\mathcal{C}$ which also transfers the system from x^o to y^o.

Summary.

In this first section we have investigated the large-scale behaviour of linear, constant coefficient control systems. The investigation has been primarily geometrical and has been based upon the set $G(\mathcal{C})$, the image of the control domain $\mathcal{C}$ under the mapping by the gain matrix G. In general $G(\mathcal{C})$ is under some control of the design engineer. For, in attempting the control of the system by a linear model with r-dimensional controls

$$m(t) = \begin{pmatrix} m_1(t) \\ m_2(t) \\ \cdot \\ \cdot \\ \cdot \\ m_r(t) \end{pmatrix}$$

(where each $m_j(t)$ may be of a different type, eg. on-off, quasi-continuous, linear with saturation, etc.), the designer has, in fact, chosen the set $\mathcal{C}$, which is the set of all points $m = m(t)$, for some t (or the convex hull of this set if bang-bang controllers are involved). The general form of the matrix G is determined by the couplings between the various components of the system but the magnitudes of the entries are determined by the gains of the various controls and so are, to a certain extent, at the designers command.

Once the form of the control system ($\mathcal{C}$) and the gain matrix (G) have been chosen the large-scale behaviour of the system near a point $\bar{x}$ depends upon the geometric location of $-A\bar{x}$ with respect to the set $G(\mathcal{C})$. In two-dimensions $G(\mathcal{C})$ may be easily plotted and the relationship between $-A\bar{x}$ and the set determined geometrically. In higher dimensions these relationships are determined by a system of equations and/or inequalities.

If $-A\bar{x}$ lies outside $G(\mathcal{C})$ the system cannot be maintained at, or even arbitrarily near, $\bar{x}$ and certain errors are inescapable. Furthermore, if $-A\bar{x}$ is too far away it may not be possible to return the system to $\bar{x}$ once it has passed near it.

When $-A\bar{x}$ is on the boundary of $G(\mathcal{C})$ the system can be maintained at $\bar{x}$ by choice of control $\bar{m}$ where $-A\bar{x} = G\bar{m}$, or arbitrarily near to $\bar{x}$ by a bang-bang controller. However, the system may be unstable, points arbitrarily near to $\bar{x}$ may not be transferable to $\bar{x}$ by any admissible control.

If $-A\bar{x} = G\bar{m}$ where $\bar{m}$ is in the interior of $\mathcal{C}$ there is a k-dimensional neighbourhood of $\bar{x}$ consisting of points transferable to $\bar{x}$ where $k = \text{rank}\ (G, AG, \ldots, A^{n-1}G)$. If $k=n$ the system can be transferred between any two points x^o, y^o which lie in S (the set of x with $-Ax$ in $G(\mathcal{C})$). Whereas when $k < n$, the set S is partitioned by a collection of k-dimensional hyperplanes and the points x^o and y^o, in addition to being in S, must lie in the same k-dimensional partition if they are to be mutually attainable.

For the practical purposes of the engineer, these results suggest:

(1) the reference point $\bar{x}$ should lie in S, with $-A\bar{x} = G\bar{m}$ for some $\bar{m}$ in the interior of $\mathcal{C}$.

(2) if possible the system should be controllable, i.e. rank $(G, AG, \ldots, A^{n-1}G) = n$.

(3) if (2) holds the system, started at any point of S, can brought to the reference point $\bar{x}$ in finite time.

(4) if (2) is not possible, only when the system is started at certain points of S can it be brought to the reference point in finite time.

(5) if the system is to operate between various reference points $\bar{x}^1$, $\bar{x}^2$, ..., $\bar{x}^j$, then these points must be chosen to be mutually attainable and hence lie in S if (2) is possible, or in the same partition of S if (2) is impossible, otherwise smooth operation of the system is impossible.

This analysis has been made without imposing any particular property on A (i.e. stability or non-degeneracy); the conclusions are valid whether A is stable or not.

Chapter Two. Open Loop Control.

In this section we discuss the problems involved in controlling a system of the type (1.1) near a chosen reference point $\bar{x}$ using bang-bang controllers. For this purpose we shall require that the system be controllable, that is, rank $(G, AG, A^2G, \ldots, A^{n-1}G) = n$ and that the control domain $\mathcal{C}$ has an interior in $\mathbb{R}^r$. The reference point $\bar{x}$ will be chosen in S and will be assumed to correspond to the control $\bar{m}$ in $\mathcal{C}$ (i.e. $-A\bar{x} = G\bar{m}$).

We have shown in Chapter One that under these conditions the system, when started at certain points (at least at another point of S), can be brought to the reference point by an admissible control. The problem of optimally choosing this control will be discussed in Chapter Three; here we shall assume the system is at the point $\bar{x}$ when t = o and shall endeavour to maintain this position via open-loop control of bang-bang controllers (controls, m, corresponding to extreme points of $\mathcal{C}$).

The control $\bar{m}$ is in $\mathcal{C}$ therefore there are extreme points $m^1, m^2, \ldots, m^k$ of $\mathcal{C}$ and constants $\alpha_j \geqslant o$, $j = 1,2,\ldots, k$ such that $\sum_{j=1}^{k} \alpha_j = 1$ and $\bar{m} = \sum_{j=1}^{k} \alpha_j m^j$. Subtracting

$$\bar{x} = e^{At}\bar{x} + \int_o^t e^{A(t-s)}G\bar{m}\,ds$$

from the solution

$$x(t,m) = e^{At}x^o + \int_o^t e^{A(t-s)}Gm(s)\,ds \qquad (2.1)$$

of $\dot{x} = Ax + Gm$, we find

$$x(t,m) - \bar{x} = e^{At}(x^o - \bar{x}) + \int_o^t e^{A(t-s)} G(m(s) - \bar{m}) ds. \quad (2.2)$$

If the system is at $\bar{x}$ when $t = o$, $x(o.m) \equiv x^o = \bar{x}$, and

$$x(t,m) - \bar{x} = \int_o^t e^{A(t-s)} G\left[m(s) - \sum_{i=1}^{k} \alpha_i m^i\right] ds \quad (2.3)$$

when $\bar{m}$ is repaced by the convex combination of extreme points. If a time $t = \bar{t}$ is fixed and if E_i denotes the set of points s in the interval $(o, \bar{t})$ where $m(s) = m^i$, (2.3) becomes

$$x(\bar{t},m) - \bar{x} = \sum_{i=1}^{k} \left[\int_{E_i} e^{A(\bar{t}-s)} - \alpha_i \int_o^{\bar{t}} e^{A(\bar{t}-s)} ds \right] Gm^i. \quad (2.4)$$

It is clear from (2.4) that if the control m is chosen so that

$$\int_{E_i} e^{A(\bar{t}-s)} ds = \alpha_i \int_o^{\bar{t}} e^{A(\bar{t}-s)} ds \quad , i = 1,2,\ldots,k \quad (2.5)$$

then $x(\bar{t},m) = \bar{x}$ and the error vanishes at $t = \bar{t}$. It is difficult, if not impossible, to satisfy all k of the equations (2.5) simultaneously. However, as our primary concern is minimizing the oscillations of the system around $\bar{x}$ we can assume $\bar{t}$ to be small (the affects of system time-lag on this assumption will be considered below). When t is small we have $e^{At} = I + O(t)$ (b) which, substituted into (2.5) yields

$$\ell(E_i)\, I = \alpha_i \bar{t} I + O(\bar{t}^2) , \quad (2.6)$$

where $\ell(E_i)$ is the length of E_i (E_i is assumed to be a union of

---

(b) see Appendix E.

intervals). Thus it is clear that the control m^i should be activated a proportion

$$\ell(E_i) / \bar{t} = \alpha_i \tag{2.7}$$

of the total time period, $\bar{t}$. Let us denote by $\hat{m}$ a control which is chosen to fulfil these conditions. For this control the error at time $\bar{t}$ is

$$x(\bar{t},\hat{m}) - \bar{x} = O(\bar{t}^2).$$

If we recycle the system, again using $\hat{m}$, we have (from (2.2))

$$x(2\bar{t},\hat{m}) - \bar{x} = e^{A\bar{t}} (x(\bar{t},\hat{m}) - \bar{x}) + O(\bar{t}^2) = O(\bar{t}^2).$$

Again recycling over a period $\bar{t}$ we find

$$x(3\bar{t},\hat{m}) - \bar{x} = e^{2A\bar{t}}(\bar{x}(t,\hat{m}) - \bar{x}) + O(\bar{t}^2)$$

and, after n-cycles

$$x(N\bar{t},\hat{m}) - \bar{x} = e^{(N-1)A\bar{t}}(x(\bar{t},\hat{m}) - \bar{x}) + O(\bar{t}^2). \tag{2.8}$$

If A is <u>not unstable</u> (i.e. has no eigenvalues with positive real part) the various powers of the matrix $e^{A\bar{t}}$ remain bounded and the error continues to be $O(\bar{t}^2)$. On the other hand, as the number of cycles gets large an unstable matrix A will lead to unbounded errors unless a correction is introduced into the system by some type of feed-back mechanism.

This analysis shows that for a system whose matrix is not unstable the constants α_i, which occur as the weightings of the extreme points m^i in the sum $\bar{m} = \sum_{i=1}^{k} \alpha_i m^i$, determine a natural oscillation of the system about $\bar{x}$ with period $\bar{t}$ whose error is no more than $O(\bar{t}^2)$. In the absence of meaningful switching lags, $\bar{t}$ can be chosen as small as desired and the oscillations will be correspondingly small.

Suppose that A is stable (has all eigenvalues with negative real part) and there are no appreciable time-lags, then equation (2.2) describes the trajectory from an arbitrary start-up point x^o. If the control is a periodic cycling of $\hat{m}$ with period $\bar{t}$ the system has, at multiples of $\bar{t}$, an error given by

$$x(N\bar{t},\hat{m}) - \bar{x} = e^{NA\bar{t}}(x^o - \bar{x}) + O(\bar{t}^2)$$

which clearly becomes $O(\bar{t}^2)$ as the number of cycles becomes large. We have shown that if $\bar{x}$ corresponds to an interior point of $\mathcal{C}$ there is a neighbourhood of $\bar{x}$ consisting of points which can be driven to $\bar{x}$ in finite time. Since, in this instance, $\bar{t}$ can be chosen arbitrarily small we are assured that after a sufficient number of cycles the system lies within this "controlled neighbourhood" of $\bar{x}$ and thus can be transferred on to $\bar{x}$ in finite time. Thus we have

If A is stable and no appreciable time-lags exist, the system can be controlled from an arbitrary start-up point to a reference point $\bar{x}$ which corresponds to an interior point of $\mathcal{C}$, i.e. $-A\bar{x} = G\bar{m}$ where $\bar{m}$ is in the interior of $\mathcal{C}$.

In general there are many convex combinations of extreme points expressing $\bar{m}$, in which case the design engineer is free to introduce other criteria in choosing the most beneficial choice of control sequence. Certain extreme points may correspond to controls which are expensive to actuate, in which case they should be excluded if possible, or, at least used in the particular combination in which they are activated over the shortest possible time. Further, if there is no reason for the designer to have a preference for certain controls, perhaps due to symmetries in the system, this symmetry can be preserved by adjusting the gain matrix G so that the control $\bar{m}$ corresponding to the reference point $\bar{x}$ is the "center of mass" of $\mathcal{C}$, i.e. $\bar{m} = \frac{1}{N}\sum_{k=1}^{N} m^k$ where $m^1, m^2, \ldots, m^N$ are <u>all</u> the extreme points of $\mathcal{C}$. This is illustrated by the following example.

<u>Example 2.1</u> Let $A = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}$ and $G = \begin{pmatrix} 1 & 2 \\ -1 & 3 \end{pmatrix}$ while $\mathcal{C}$ is again the unit square in $\mathbb{R}^2$, i.e. $\mathcal{C}$ consists of all (m_1, m_2) with $0 \leq m_1, m_2 \leq 1$. Choosing $\bar{x} = (3/2, 1/2)$. the equation $-A\bar{x} = G\bar{m}$ leads to

$$\bar{m}_1 + 2\bar{m}_2 = 3/2$$

$$-\bar{m}_1 + 3\bar{m}_2 = 1$$

which has the unique solution $(\bar{m}_1, \bar{m}_2) = (\frac{1}{2}, \frac{1}{2})$. The extreme points of $\mathcal{C}$ are $m^0 = (0,0)$, $m^1 = (1,0)$, $m^2 = (0,1)$ and $m^3 = (1,1)$. For $\bar{m}$ we have the following obvious convex combinations

$$\begin{aligned} \bar{m} &= \tfrac{1}{2}m^0 + \tfrac{1}{2}m^3 \\ &= \tfrac{1}{2}m^1 + \tfrac{1}{2}m^2 \\ &= \tfrac{1}{4}m^0 + \tfrac{1}{4}m^1 + \tfrac{1}{4}m^2 + \tfrac{1}{4}m^3 \end{aligned}$$

which are all special cases of

$$\bar{m} = \delta m^0 + (\tfrac{1}{2} - \delta)\, m^1 + (\tfrac{1}{2} - \delta) m^2 + \delta m^3, \quad 0 \leq \delta \leq \tfrac{1}{2}.$$

The trajectories near $\bar{x}$ corresponding to the first three policies ($\delta = \frac{1}{2}, 0, \frac{1}{4}$) are shown in Figure 9(a).

We now turn our attentions to the difficulties occurring when the system controllers are subject to significant time-lags.

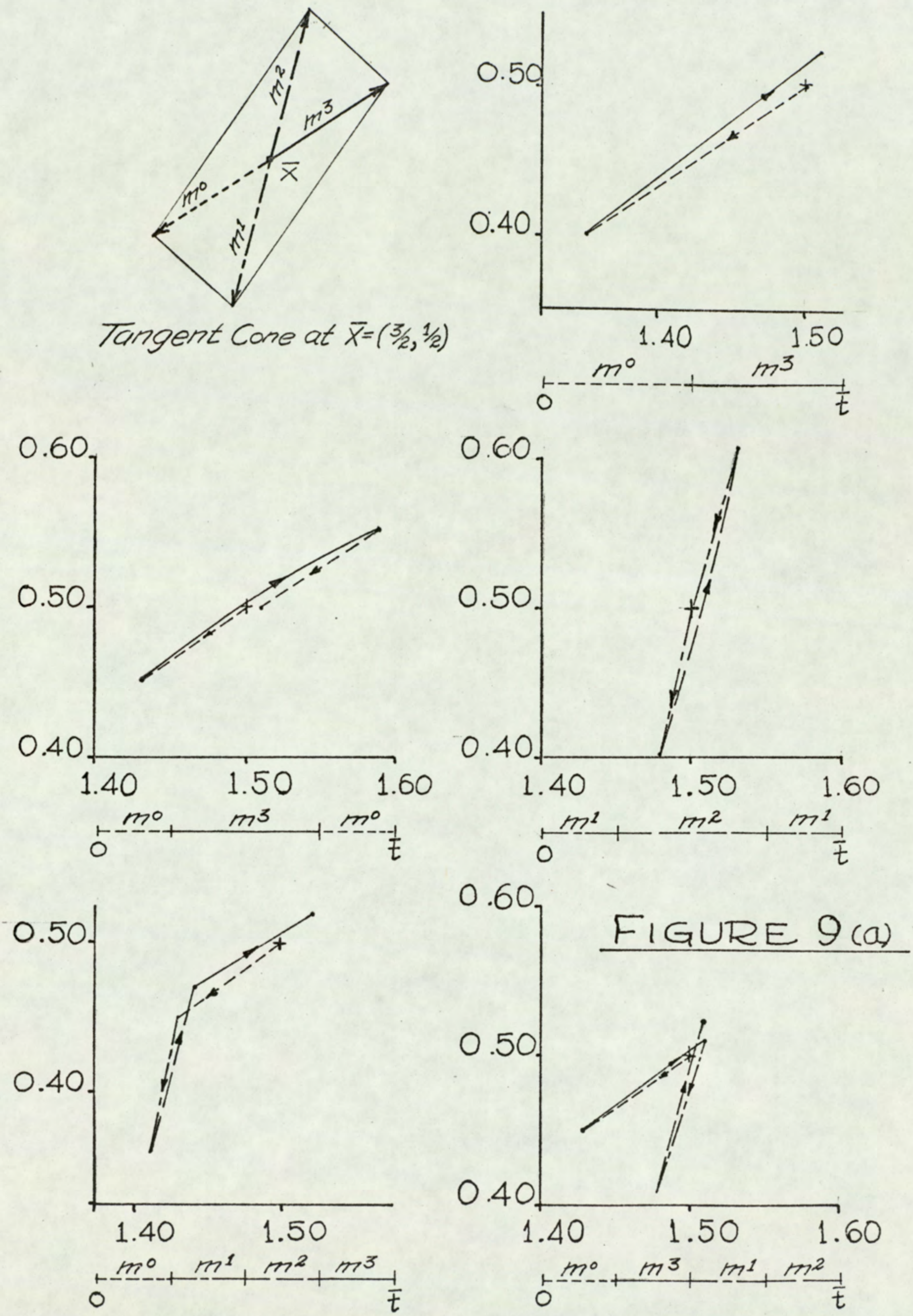


FIGURE 9(a)

SOME OPEN-LOOP CONTROL POLICIES ($\bar{t}$=0.20)

The most general case can become exceedingly complicated as one must expect, in a switch from m^i to m^j, a different time-lag $L^{\nu}_{i,j}$, $\nu = 1,2,\ldots,r$, corresponding to each component of the control vector. The motion of the system during the switching period would then be due, at each instant, to various combinations of the components of m^i and m^j. As a simplification we shall assume that during a switch from m^i to m^j all switching components of the control vector are subject to time-lags of the same order of magnitude symbolized by the quantity $L_{i,j}$. We do not assume $L_{i,j} = L_{j,i}$, however.

Above we have defined a control policy which consists in a periodic application of the control sequence m defined by $\bar{m} = \sum_{i=1}^{k} \alpha_i m^i$ in the following manner.

The sequence begins with

$$
\begin{aligned}
m(t) &= m^1 & t_1 &\leq t < t_2 \\
&= m^2 & t_2 &\leq t < t_3 \\
&\;\;\vdots & &\;\;\vdots \\
&= m^i & t_i &\leq t < t_{i+1} \\
&\;\;\vdots & &\;\;\vdots \\
&= m^k & t_k &\leq t < t_1' \\
&= m^1 & t_1' &\leq t < t_2' \\
&\;\;\vdots & &\;\;\vdots
\end{aligned}
$$

Thus the switch from m^i to m^{i+1} occurs at t_i, t_i', t_i'', etc., $t_1' - t_1 = \bar{t}$, and $t_{i+1} - t_i = \ell(E_i) = \alpha_i \bar{t}$, where $\bar{t}$ is the period allotted to each cycle. Here to minimize the number of switches we have assumed each set E_i is an interval and hence no controls are repeated within the cycle. Let us now denote by τ_i the actual time we command the switch from m^{i-1} to m^i. Clearly if switching is to be completed at time t_i we must have

$$\tau_i = t_i - L_{i-1,i} \,. \tag{2.9}$$

(see Figure 9(b)).

The time between commands must be positive, hence

$$\tau_{i+1} - \tau_i = t_{i+1} - L_{i,i+1} - (t_i - L_{i-1,i}) =$$

$$= \alpha_i \bar{t} - L_{i,i+1} + L_{i-1,i} > o$$

or

$$\bar{t} > \max \left(\frac{L_{i-1,i} - L_{i,i+1}}{\alpha_i}\right) \quad i = 1, 2, \ldots, k \tag{2.10}$$

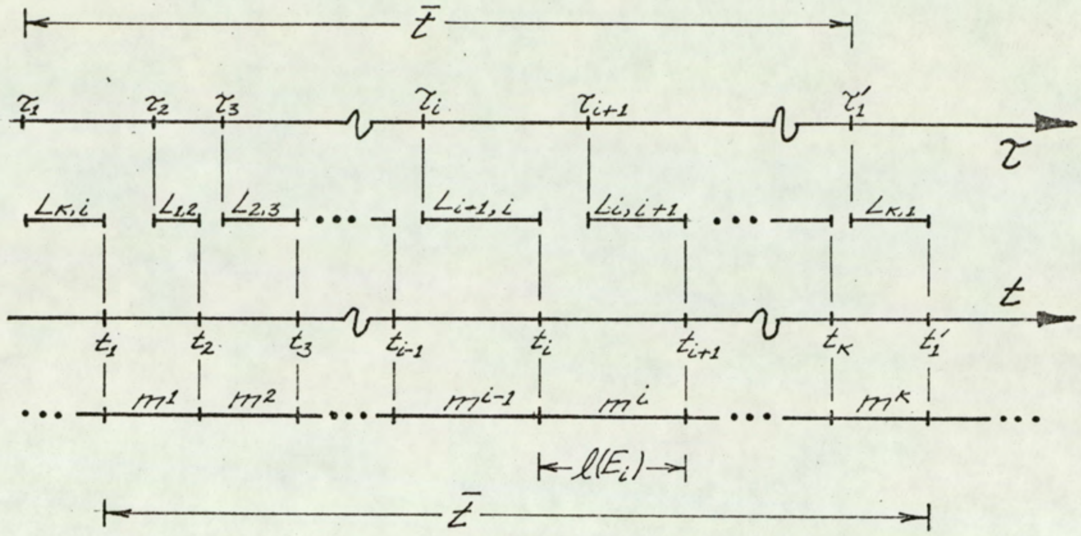
(note: here $L_{o,1} = L_{k,1}$ and $L_{k,k+1} = L_{k,1}$).

If $\bar{t}$ satisfies (2.10) then $\tau_1' - \tau_1 = \bar{t}$ and the cycling policy is feasible provided an error of the order $O(\bar{t}^2)$ is acceptable. We should observe that the condition $\tau_{i+1} - \tau_i = \alpha_i \bar{t} - L_{i,i+1} + L_{i,i-1} > o$ does not exclude the possibility $\alpha_i \bar{t} = \ell(E_i) < L_{i,i+1}$ which means that the command to switch from m^i to m^{i+1} must be given before the switch from m^{i-1} to m^i is complete, see Figure 9(b). In particular instances this might not be feasible, in which case the designer must determine the minimum cycling period from

$$\alpha_i \bar{t} > L_{i,i+1}$$

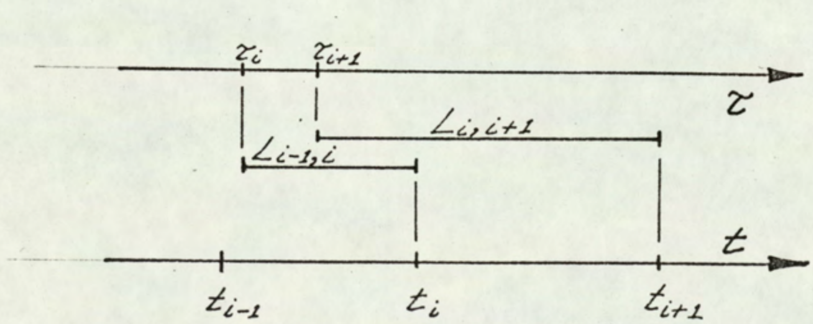
which implies

$$\bar{t} > \max \frac{L_{i,i+1}}{i} \quad i = 1, 2, \ldots, k \tag{2.11}$$



SWITCHING POLICY WITH TIME-LAGS.

FIGURE 9(b)



Situation when $L_{i,i+1} > l(E_i)$.

(as usual $L_{k+1} = L_{k,1}$).

It should be observed that in the special case of equal time-lags for all switchings (all the L's equal),(2.10) shows that the cycling period can be arbitrarily small and the results derived for the system with no time-lags are applicable. If $\bar{t}$ is determined from (2.11) instead, the condition $\bar{t} \geqslant kL$ is evident.

We have, until this point, considered the natural oscillations of the system, as derived above, only as a basis for an open-loop control system. However, these same natural oscillations can appear in systems with certain feed-back control devices. The resultant oscillations have been called <u>dynamic equilibrium cycling</u> in [1], [3] and [4] where one-dimensional temperature control is discussed. In these applications the switching times are determined by comparison of the output variable with its desired value. The switching times and the system lag, L, then determine the oscillation with period $\bar{t} \geqslant 2L$. The proportion of on-time, t_o, and off-time, t_p, to the total period is determined in accordance with the earlier analysis,

$$t_o = \alpha_o \bar{t} \ , \quad t_p = \alpha_p \bar{t} \qquad (2.12)$$

where α_o and α_p are the coefficients of the convex combination corresponding to the selected reference temperature. Equation (2.12) is valid only when the time-lag of the system is small compared to the response time of the system (i.e. major time constant) otherwise the linear approximations we have made are not valid through the entire cycle necessitated by the feed-back device.

Summary

In this chapter we have shown that a system such as 1.1 has a natural oscillation about a chosen reference point which may be used as the basis for an open-loop control system which allows a precise description of the errors involved. The oscillations are influenced by the interaction of the control domain and the reference point $\bar{x}$ in the following way. If $A\bar{x} + G\bar{m} = 0$ and $\bar{m}$ satisfies $\bar{m} = \Sigma\alpha\ m$ where the m , ..,m are extreme points of , then the control m should be activated, approximately, a proportion α of the total period of oscillation. When no time lags occur this leads to arbitrarilly small oscillations of the system about the reference points. The presence of time lags introduces a lower brand on the resultant errors and can create unacceptable performance.

Chapter Three: Optimal Control.

The final chapter is primarily concerned with the time-optimal control of systems such as (1.1) although certain of the results to be derived may be applied to systems which are to be optimized with respect to cost functions other than response time.

As is well known to all control engineers, the primary obstacle in realizing an optimum system is the determination of the switching surfaces in the state-space which are crucial to development of a feed-back control device. There have been many papers written on this problem and many computational techniques developed to approximate the optimal control and its switching surfaces, although there does not seem to be a "best possible" method at the present time. In this section we shall discuss this problem and deduce certain properties of the optimal control which in low-dimensional systems aid in contructing the switching surfaces. The study to be made will be, as in the preceeding, from a geometrical point of view, so we begin by outlining that portion of the general theory of optimal control which relates to the geometric concepts we wish to study.

We suppose we have a system of the form (1.1) with a bounded convex control domain $\mathcal{C}$ which has an interior. The reference point $\bar{x}$ is assumed to satisfy $-A\bar{x} = G\bar{m}$ for some $\bar{m}$ in the interior of $\mathcal{C}$, as we have shown in Chapter One, this condition implies that, for controllable systems, there is a neighbourhood of $\bar{x}$ all of whose points can be transferred to $\bar{x}$ in finite time. If there is an admissible control transferring a point x^0 to $\bar{x}$ then there exists a time-optimal control also transferring x^o to $\bar{x}$ (i.e. a control which accomplishes the transfer in as short a time interval possible by admissible controls, see [2], p.127, Theorem 17.) Let us denote by $T(x)$ the optimal time required to transfer x to $\bar{x}$ with the convention $T(x) = \infty$ if there is no admissible control transferring x to $\bar{x}$. We are interested in the geometric properties of the <u>isocronal</u> <u>surfaces</u>, $T(x)$ = constant. If $T(x^1) = \bar{t}$, the minimal time required to transfer x^1 to $\bar{x}$ is $\bar{t}$ and hence $\bar{x}$ is attainable from x^1 in time $\bar{t}$. We have seen that this implies that $\bar{x}$ is controllable to x^1 in time $\bar{t}$ by the control system

$$\dot{x} = -Ax - Gm \tag{3.1}$$

i.e. (1.1) with time reversed. Let us denote by $K(t)$ the set of points which are attainable from $\bar{x}$ in time t via (3.1) and an admissible control. $K(t)$ is a closed convex set whose boundary is the isocronal surface $T(x) = t$. The supporting hyperplanes to $K(t)$ are related to the optimal control problem via the Pontragin minimum principle. Indeed, if a is the outer normal to a supporting hyperplane at a boundary point x^o of $K(t)$ (i.e. $T(x^o) = \bar{t}$), then the optimal control m^* defined on $o \leq t \leq \bar{t}$ and transferring the system from x^o to $\bar{x}$ satisfies

$$\eta(t)Gm^*(t) = \underset{m \in \mathcal{C}}{\text{minimum}}\ \eta(t)Gm \tag{3.2}$$

at (almost)[c] every point of the interval $o \leq t \leq \bar{t}$ where

$$\eta(t) = a.e^{-At} \tag{3.3}$$

As this relationship between the optimal control m^* and the outer normal a will be an important part of our analysis we will

---

(c) "almost every point" refers to the theory of Lebesgue measure.

indicate the reasoning behind (3.2).

If the control $m^*(t)$ transfers the system $\dot{x} = Ax + Gm$ optimally from x^o to $\bar{x}$ in time $\bar{t}$, the control $\hat{m}(t) = m^*(\bar{t}-t)$ transfers the reversed time system $\dot{x} = -Ax - Gm$ from $\bar{x}$ to x^o optimally in time $\bar{t}$. This latter trajectory is given by

$$x^-(t) = e^{-At}\bar{x} - \int_o^t e^{A(s-t)}G\hat{m}(s)ds. \tag{3.4}$$

Because a is an outer normal to $K(\bar{t})$ at x^o

$$a.x^o = a.x^-(\bar{t}) \geqslant a.y \text{ for all } y \text{ in } K(\bar{t}). \tag{3.5}$$

The extreme property of m^* enters the discussion when we calculate

$$a.x^-(\bar{t}) = a.e^{-A\bar{t}}\bar{x} - \int_o^{\bar{t}} a.e^{A(s-\bar{t})}G\hat{m}(s)ds. \tag{3.6}$$

Suppose

$$a.e^{A(s-\bar{t})}G\hat{m}(s) = a.e^{A(s-\bar{t})}Gm^*(\bar{t}-s) >$$
$$> \min_{m \text{ in } \mathcal{C}} \left(a.e^{A(s-\bar{t})}Gm\right)$$

over some subset E of $(o,\bar{t})$ having positive Lebesgue measure. It can be shown that the control $\tilde{m}$ defined by

$$\tilde{m}(s) = \min_{m \text{ in } \mathcal{C}} a.e^{A(s-\bar{t})}Gm, \text{ for } s \text{ in } E$$
$$= \hat{m}(s) \quad \text{otherwise}$$

is admissible and therefore the response

$$\tilde{y}(\bar{t}) = e^{-A\bar{t}}\bar{x} - \int_o^{\bar{t}} e^{A(s-\bar{t})}G\tilde{m}(s)ds$$

is in $K(\bar{t})$ and satisfies

$$a.\tilde{y}(\bar{t}) - a.x^o = \int_o^{\bar{t}} \left[a.e^{A(s-\bar{t})}\, G\hat{m}(s) - a.e^{A(s-\bar{t})}\, G\tilde{m}(s)\right] ds > o$$

which contradicts (3.5). This contradiction shows that,

$$a.e^{A(s-\bar{t})}Gm^*(\bar{t}-s) = \min_{m \text{ in } \mathcal{C}} (a.e^{A(s-\bar{t})}Gm) \tag{3.7}$$

or, with $t = \bar{t} - s$,

$$a.e^{-At}Gm^*(t) = \min a.e^{-At}Gm, \quad m \text{ in } \mathcal{C}\ , \text{ (d)}$$

thus, if η is given by (3.3) we have the result (3.2).

The vector function $\eta(t) = a.e^{-At}$ is the unique solution to the adjoint system

$$\dot{\eta} = -\eta A \ , \qquad \eta(o) = a.$$

Thus, the outer normal, a, determines the optimal control corresponding to x^o by means of equation (3.2) and (3.3). The form of the optimal control is evident also from (3.2) for the minimum on the right hand side is assumed at extreme points of $\mathcal{C}$ and therefore all time-optimal controls are of bang-bang type.

We gain a geometric meaning from the equation (3.7) if we let $\xi(t) = \eta(\bar{t}-t) = a.e^{A(t-\bar{t})}$. Then

$$\dot{\xi}(t) = \xi(t)A \ , \quad \xi(\bar{t}) = a \tag{3.8}$$

and

(d) $\eta(t)$ is an outer normal to the sets, K(t), of attainability from $\bar{x}$; if we had considered the sets of attainability from x^o with system (1.1) we should be considering $-\eta(t)$ which converts (3.2) into the usual form of the maximal, rather than minimal, principle.

$$\frac{d}{dt}\left(\xi(t).x^-(t) \right) = \dot{\xi}(t).x^-(t) + \xi(t).\dot{x}^-(t) =$$
$$= \xi(t)Ax^-(t) - \xi(t)Ax^-(t) - \xi(t)G\hat{m}(t) =$$
$$= -\xi(t)G\hat{m}(t) = \max_{m \text{ in } \mathcal{G}} \left[\xi(t)(-G)m \right].^{(e)} \qquad (3.9)$$

Equations (3.8) and (3.9) may be interpreted geometrically to mean that at each point the optimal control, $\hat{m}$, is selected so as to maximize the rate of change of the projection of $x^-(t)$ on a prescribed direction $\xi(t)$ which has been determined by the adjoint system and the normal vector a.

In essence the adjoint system "pulls back" the vector a along the trajectory using equation (3.8) in order to determine the optimal control by equation (3.9). This "pull-back" is accomplished in such a manner that $\xi(t)$ is the outer normal to $K(t)$ at $x^-(t)$ for, from (3.7)

$$\xi(t).x^-(t) = \xi(t).e^{-At}\bar{x} - \int_o^t a.e^{A(t-\bar{t})}.e^{A(s-t)}G\hat{m}(s)ds =$$
$$= \xi(t).e^{-At}\bar{x} - \int_o^t a.e^{A(s-\bar{t})}G\hat{m}(s)ds \geqslant$$
$$\geqslant \xi(t).e^{-At}\bar{x} - \int_o^t a.e^{A(s-\bar{t})}Gm(s)ds =$$
$$= \xi(t).x^-(t,m),$$

where $x^-(t,m)$ is the response from $\bar{x}$ using an arbitrary admissible control m. This shows that the hyperplane $\xi(t).x = \xi(t).x^-(t)$ is a support hyperplane to $K(t)$ at the point $x^-(t)$ for each fixed t in the interval $(o,\bar{t})$(see Figure 10).

(e) The term "-G" occurs here only because we are considering system (3.1).

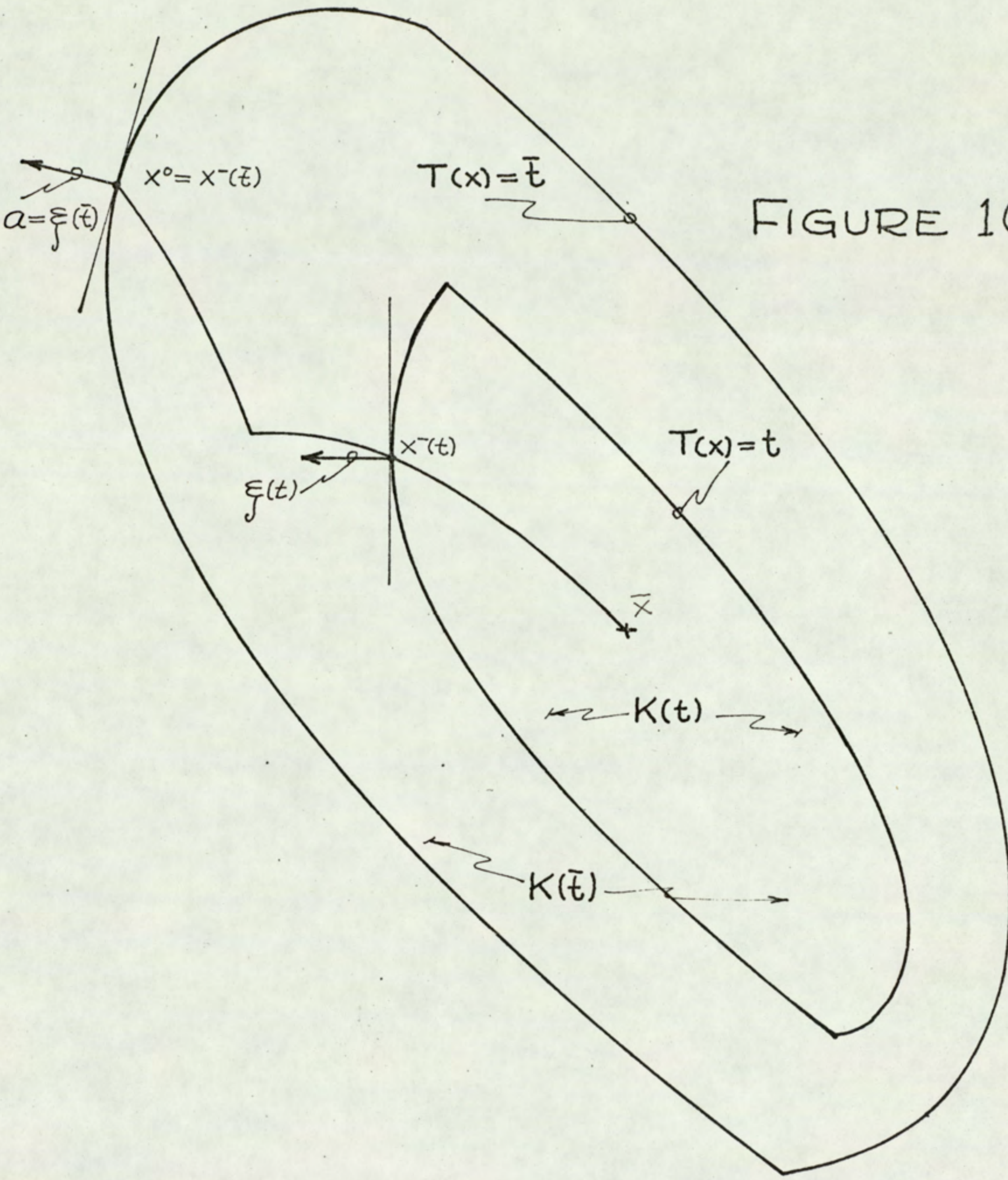


FIGURE 10

GEOMETRIC RELATIONSHIPS BETWEEN ISOCRONAL SURFACES AND ADJOINT RESPONSE, $\xi(t) = ae^{A(t-\bar{t})}$.

Now, as remarked earlier, the boundary of $K(t)$ is the isocronal surface $T(x) = t$ which has, at a point away from a switching surface, a gradient $\nabla T(x)$ which is normal to the isocronal surface.

Thus, since

$$T\,(x^-(t) = t \tag{3.10}$$

$\xi(t)$ and $\nabla T(x^-(t)\ )$ are both outer normals to $K(t)$ at $x^-(t)$ and are therefore positive multiples of one another. Indeed, upon differentiating (3.10) by t we find

$$\nabla T[x^-(t)\ ].\ \dot{x}^-(t) = \nabla T[x^-(t)\ ].(-Ax^-(t)-G\hat{m}(t)\ ) = 1$$

which shows that the optimal control at any point x away from a switching surface must be chosen so that

$$-\ \nabla T(x).(Ax + G\hat{m}) = 1.$$

Again from (3.9) and using the fact that $\xi(t) = k\,\nabla T(x^-(t)\ )$ for $k > o$, we deduce that the optimal control $\hat{m}$ at the point x satisfies

$$-\ \nabla T(x).G\hat{m} = \max_{m \text{ in } \mathcal{C}} (-\ \nabla T(x).Gm) \tag{3.11}$$

which characterizes an optimal controller in terms of the isocronal surfaces.

If the minimization process of (3.2) uniquely determines m^* the system (1.1) is said to be normal (system (3.1) will be, of course, also normal). For normal systems, with reference points $\bar{x}$ as chosen here, there exists a unique optimal control for each point x with $T(x) < \infty$. Further, the sets $K(t)$ are <u>strictly convex</u> meaning that every support hyperplane of $K(t)$ intersects $K(t)$ at a unique point.

In particular:

<u>If the system matrix A is stable every point x of $\mathbb{R}^n$ can be transferred to $\bar{x}$ in time $T(x)$ by a unique optimal control.</u>

In the examples to be discussed here, the control domain will be a <u>convex polyhedron</u>, that is a convex set whose boundary consists of portions of finitely many hyperplanes and their intersections (edges and vertices). For such control domains the system will be normal if, and only if, the vectors

$$Gu,\ AGu,\ A^2Gu,\ \ldots,\ A^{n-1}Gu \tag{3.12}$$

are linearly independent for every unit vector u which is parallel to an edge of $\mathcal{C}$.

Let us now use the properties of K(t) and T(x) developed above in an attempt to determine these sets and, through them, the time-optimal control.

For definiteness let us denote the extreme points of $\mathcal{C}$ by $m^1, m^2, \ldots, m^k$. As we have seen, an optimal control will take on only these extreme values. Further, we assume that to each m^j there corresponds a unique point x^j in S with

$$-Ax^j = Gm^j. \tag{3.13}$$

The points $s^j(\bar{t})$ which can be controlled to $\bar{x}$ in time $\bar{t}$ by constant application of the control m^j can be found by integrating (3.1) with $s(o) = \bar{x}$. Thus:

$$\dot{s}^j = -As^j - Gm^j = -As^j + Ax^j$$

which has the solution

$$s^j(\bar{t}) = e^{-A\bar{t}}(\bar{x}-x^j) + x^j. \tag{3.14}$$

Continuing, we denote by $s^{i,j}(\tau;\ \bar{t})$ the point which can be controlled to $\bar{x}$ in time $\bar{t}$ with a switch from i to j at $t = \tau$.

$s^{i,j}(\tau;\bar{t})$ is found by integrating the equation

$$\dot{s} = -As - Gm^i = -As + Ax^i$$

from the point $s^j(\bar{t} - \tau)$ over the interval (o, τ). This yields

$$s^{i,j}(\tau;\bar{t}) = e^{-A\tau}(s^j(\bar{t}-\tau) - x^i) + x^i =$$

$$= x^i + e^{-A\bar{t}}(\bar{x} - x^j) - e^{-A\tau}(x^i - x^j). \qquad (3.15)$$

Likewise, a point $s^{i,j,k}(\tau_1, \tau_2; \bar{t})$ controllable to $\bar{x}$ in time $\bar{t}$ by the control m given by

$$\begin{aligned} m(t) &= m^i & o \leq t < \tau_1 \\ &= m^j & \tau_1 \leq t < \tau_2 \\ &= m^k & \tau_2 \leq t \leq \bar{t} \end{aligned}$$

is given by

$$s^{i,j,k}(\tau_1, \tau_2; \bar{t}) = e^{-A\tau_1}\left[s^{j,k}(\tau_2 - \tau_1; \bar{t} - \tau_1) - x^i\right] + x^i$$

$$= x^i + e^{-A\bar{t}}(\bar{x} - x^k) - e^{-A\tau_1}(x^i - x^j) - e^{-A\tau_2}(x^j - x^k). \qquad (3.16)$$

The form now becomes evident. A point which can be controlled to $\bar{x}$ in time $\bar{t}$ with control

$$\begin{aligned} m(t) &= m^{j_0} & o \leq t < \tau_1 \\ &= m^{j_1} & \tau_1 \leq t < \tau_2 \\ &\vdots & \vdots \\ &= m^{j_i} & \tau_i \leq t < \tau_{i+1} \\ &\vdots & \vdots \\ &= m^{j_N} & \tau_N \leq t \leq \bar{t} \end{aligned}$$

can be shown to have the form

$$s^{j_0,\dots,j_N}(\tau_1,\dots,\tau_N;\ \bar{t}) = x^{j_0} + e^{-A\bar{t}}(\bar{x}-x^{j_N}) - \sum_{\nu=1}^{N} e^{-A\tau_\nu}(x^{j_{\nu-1}} - x^{j_\nu}) \qquad (3.17)$$

To simplify the notation we shall denote by J a <u>multi-index</u> $J = (j_o, j_1, \dots, j_N)$ where each integer j_i satisfies

$$1 \leq j_i \leq k \qquad (3.18)$$

and define

$$|J| = N. \qquad (3.19)$$

In addition to (3.18) we require that no pair of adjacent integers j_i, j_{i+1} in J be the same. This will insure that a non-trivial switch takes place at τ_{i+1}. We also let $\tau = (\tau_1, \tau_2, \dots, \tau_N)$ denote a vector in $\mathbb{R}^N$ with components

$$o \leq \tau_1 \leq \tau_2 \leq \tau_3 \dots \leq \tau_N \leq \bar{t} \qquad (3.20)$$

The set of all such points in $\mathbb{R}^N$ will be denoted by $\Delta(N;\bar{t})$.

With these conventions we have

$$s^J(\tau;\bar{t}) \equiv s^{j_o,\dots,j_N}(\tau_1,\dots,\tau_N;\bar{t})$$

and, if no confusion can result, we shall often suppress the $\bar{t}$ as well and then denote the point (3.17) by $s^J(\tau)$, τ in $\Delta(N,\bar{t})$ and the corresponding control by m^J_τ.

Now for fixed $\bar{t}$ the point $s^J(\tau)$, as τ varies in $\Delta(N,\bar{t})$, runs through a hyper-surface of dimension $\leq N$ which is contained in $K(\bar{t})$ by construction. The boundary of $K(\bar{t})$, the isocronal surface $T(x) = \bar{t}$, must also consist of portions of these hyper-surfaces and if, by some means, the particular boundary surfaces can be selected, the optimal control problem will be solved for all points on the boundary of $K(\bar{t})$.

We now turn our attention to means of discriminating between the various hyper-surfaces so that those forming the boundary of $K(\bar{t})$ may be determined.

Equation (3.2) shows that the switching locus depends upon the control domain $\mathcal{C}$. For this reason we shall restrict our analysis to two common types of control domains. The first, which corresponds to on-off controls, is the unit r-cube consisting of all $m = (m_1, \ldots, m_r)$ with $0 \leq m_i \leq 1$. The second, the symmetric r-cube, consists of those m with $-1 \leq m_i \leq 1$ and corresponds to two-way controllers. Theoretically there is no difference between these two types of controllers, they each have 2^r extreme points with all co-ordinates either 0 or 1 for on-off controls or, -1 or 1 for two-way controls. The origin always corresponds to an interior point of $\mathcal{C}$ when two-way controllers are used. The concepts to be discussed below may be easily extended to convex polyhedra of more general type than the two we are considering here.

If we let $p(t) = \eta(t).G = a.e^{-At}G$, equation (3.2) becomes

$$p(t).m^*(t) = \min_{m \text{ in } \mathcal{C}} p(t).m. \tag{3.21}$$

It is clear that if m is to be chosen to minimize the inner product $p(t).m = \sum_{i=1}^{r} p_i(t) m_i$ and to satisfy the constraints (i.e. m in $\mathcal{C}$) then each component m_i^* of the optimal control m^* must be

$$\left.\begin{aligned} m_i^* &= 0 \quad \text{if } p_i(t) > 0 \\ &= 1 \quad \text{if } p_i(t) < 0 \end{aligned}\right\} \text{ on-off controls}$$

$$\left.\begin{aligned} m_i^* &= -1 \quad \text{if } p_i(t) > 0 \\ &= 1 \quad \text{if } p_i(t) < 0 \end{aligned}\right\} \text{ two-way controls} \tag{3.22}$$

The switching instants are determined by the times the components of p(t) change sign. When $p_i(t) = o$ the component m_i* is undetermined. This is of no consequence if the zeros are isolated, that is to say $p_i(t)$ doesn't vanish on an interval, for in this case values of the control at a single point are unimportant. The vanishing of $p_i(t)$ over an interval obviously allows a choice of controls which satisfy (3.21) and so the optimal control would not be unique. This condition cannot occur when the system is normal, i.e. when (3.11) holds.

The vector p(t) depends upon the matrix A for its precise form, for example, if A is diagonal with distinct diagonal elements $\lambda_1, \lambda_2, \ldots, \lambda_n$ then the i-th component $p_i(t)$ of p(t) is the form

$$p_i(t) = \sum_{=1}^{n} q_{i,\nu} e^{-\lambda_\nu t} \qquad (3.23)$$

If the eigen values of A are not distinct, but with λ_ν having multiplicity n_ν then

$$p_i(t) = \sum_{\nu=1}^{m} q_{i,\nu}(t) e^{-\lambda_\nu t} \qquad (3.24)$$

where $q_{i,\nu}(t)$ is a polynomial with degree no larger than $n_\nu - 1$.

It will be important to know how many possible switch points p_i introduces, i.e. how many times it changes sign. If some of the eigenvalues are complex, it is easy to show that the number of zeros of p_i can be unbounded. However, if all eigenvalues are real we can determine a bound on the number of zeros of an <u>exponential</u> <u>polynomial</u> such as (3.24). Consider such an exponential polynomial

$$p(t) = \sum_{\nu=1}^{k} q_\nu(t)e^{\lambda_\nu t} \tag{3.25}$$

with each λ_ν real and q_ν a polynomial of degree no larger than $m_\nu - 1$.

If $k = 1$, $p(t) = o$ only if $q_1(t) = o$, therefore p would have no more than $m_1 - 1$ roots. If $k = 2$ we have

$$p(t)e^{-\lambda_1 t} = q_1(t) + q_2 e^{(\lambda_2 - \lambda_1)t}.$$

Differentiating m_1 times we find

$$\frac{d^{m_1}}{dt^{m_1}}\left(p(t)e^{-\lambda_1 t}\right) = Q(t)e^{(\lambda_2 - \lambda_1)t}$$

where Q is a polynomial whose degree is no larger than $m_2 - 1$. Thus the m_1-derivative of $p(t)e^{\lambda_1 t}$ can have no more than $m_2 - 1$ zeros and hence $p(t)e^{-\lambda_1 t}$ can have no more than $m_1 + m_2 - 1$ zeros. Continuing in this way (using mathematical induction) we can show that (3.25) has no more than $m_1 + m_2 + \dots + m_k - 1$ zeros and hence $p_i(t)$ has no more than $n_1 + n_2 + \dots + n_m - 1 = n - 1$ zeros. This shows that each component of the optimal control m^* will switch at most n-1 times and hence the total number of switching times is no larger than r(n-1).

If the eigenvalues of A are not all real there will be no upper bound on the number of switch points; however, if the system is normal so that no component of p can vanish over an interval, the zeros of each component will be isolated and thus for a fixed time $\bar{t}$ the number of switches for $o \leqslant t \leqslant \bar{t}$ will be bounded. Thus, in any practical situation there is an effective bound on the number of switches possible.

We now seek to determine the switching locus by constructing the isocronal surfaces. To facilitate this construction we shall summarize below the properties of the isocronal surfaces and optimal controls which have been derived.

<u>Properties of isocronal surface $T(x) = \bar{t}$.</u>

<u>P.I.</u> An isocronal surface is the boundary of a convex set.

<u>P.2.</u> A control m^* directs a point x^o to $\bar{x}$ in minimum time $\bar{t}$ if, and only if, there is a vector a^o such that

$$a^o \cdot e^{-At} G m^* = \min_{m \text{ in } \mathcal{C}} a^o \cdot e^{-At} G m \qquad (3.26)$$

for $o \leqslant t \leqslant \bar{t}$.

<u>P.3.</u> If the system is controllable, the surface $T(x) = \bar{t}$ bounds an n-dimensional body.

<u>P.4.</u> If all eigenvalues of A are real, an optimal control will have no more than $r(n-1)$ switching instants.

<u>P.5.</u> If the system is normal, the optimal controls are unique.

In utilizing these properties of the isocronal surfaces to determine the switching surfaces of the process by geometric means we are limited to state-spaces of two or three dimensions. For higher dimensional processess the geometry must be replaced by analytic methods. Nevertheless, the ease with which two dimensional problems can be solved furnishes the design engineer, or teacher, with easily constructed examples which can illustrate all pathological behaviour of such control systems which are analytic in nature. The problems arising through purely dimensional considerations must be overcome through other methods.

However, if the system can be designed in such a manner that it consists of several low-dimensional systems loosely coupled (little interference between subsystems) then the methods developed below should provide considerable information toward the development of the optimal control system.

We begin by determining when the points $s^j(\bar{t})$, obtained by constant m^j-control, are on the surface. By P.2. this occurs when there is a vector a satisfying (3.26) with $m^* = m^j$. This will be possible for sufficiently small $\bar{t}$ provided the columns of G are linearly independent for then $p(o) = a.G$ can be adjusted, by proper choice of the vector a, to have positive components in any desired combination, in particular, that combination of components which requires that the minimum of (3.26) be assumed for $m^* = m^j$. Having accomplished this for t = o, i.e.

$$p(o).m^j = \min_{m \text{ in } \mathcal{C}} p(o).m,$$

continuity of p(t) will ensure that

$$p(t).m^j = \min_{m \text{ in } \mathcal{C}} p(t).m$$

over some time interval $(o,\bar{t})$.

Consider now the locus $s^{i,j}(\tau;\bar{t})$ consisting of points controllable to $\bar{x}$ in time $\bar{t}$ by a single switch, at $t = \tau$, from m^i to m^j. We have

$$s^{i,j}(\tau;\bar{t}) = x^i + e^{-A\bar{t}}(\bar{x} - x^j) - e^{-A\tau}(x^i - x^j)$$

$$= s^j(\bar{t}) + (I - e^{-A\tau})\cdot(x^i - x^j), \qquad (3.27)$$

since $s^j(\bar{t}) = x^j + e^{-A\bar{t}}(\bar{x}-x^j)$. This shows that $s^{i,j}(o;\bar{t}) = s^j(\bar{t})$ and $s^{i,j}(\bar{t};\bar{t}) = s^i(\bar{t})$; thus as τ varies between o and $\bar{t}$ (3.27) traces a curve from $s^j(=s^j(\bar{t}))$ to $s^i(=s^i(\bar{t}))$. If the control sequence m^i to m^j is optimal this curve must lie on the isocronal surface, otherwise it lies inside the surface.

The two-switch locus (3.16) can be written (supressing the $\bar{t}$ which is understood) as $s^{i,j,k}(\tau_1,\tau_2) = s^{i,j}(\tau_1) + s^{j,k}(\tau_2) - s^j$ which satisfies

$$
\begin{aligned}
&(a) \quad s^{i,j,k}(o,o) = s^k \\
&(b) \quad s^{i,j,k}(o,\bar{t}) = s^j \\
&(c) \quad s^{i,j,k}(\bar{t},\bar{t}) = s^i \\
&(d) \quad s^{i,j,k}(o,\tau_2) = s^{j,k}(\tau_2) \\
&(e) \quad s^{i,j,k}(\tau_1,\bar{t}) = z^{i,j}(\tau_1) \\
&(f) \quad s^{i,j,k}(\tau,\tau) = z^{i,k}(\tau).
\end{aligned}
\qquad (3.28)
$$

The locus described by $s^{i,j,k}(\tau_1,\tau_2)$ as τ_1,τ_2 vary over the range $o \leq \tau_1 \leq \tau_2 \leq \bar{t}$ is a surface between the three points s^i, s^j, s^k which contains the three curves $s^{i,j}(\tau)$, $s^{j,k}(\tau)$ and $s^{i,k}(\tau)$, (see Figure 11). Similarly the N-switch locus of (3.17) will, when τ varies in $\Delta(N;\bar{t})$, consist of a set of points containing s^{j_o}, s^{j_1}, ..., s^{j_N} and the (N-1)-switch loci between each collection of N of these points. Of course in a state-space of dimension n, the N-switch loci can have dimension no larger than n. The optimal control, however, will generally consist of at least r(n-1) switching instants and hence all such multiple switching loci s^J for $|J| \leq r(n-1)$ must be considered even in the simplest case when all eigenvalues of A are real.

Let us now consider the case, $n = r = 2$, which means that if A has real eigenvalues we need consider only $|J| \leq 2$. There are $2^2 = 4$ extreme points, m^1, m^2, m^3, m^4, in the control domain

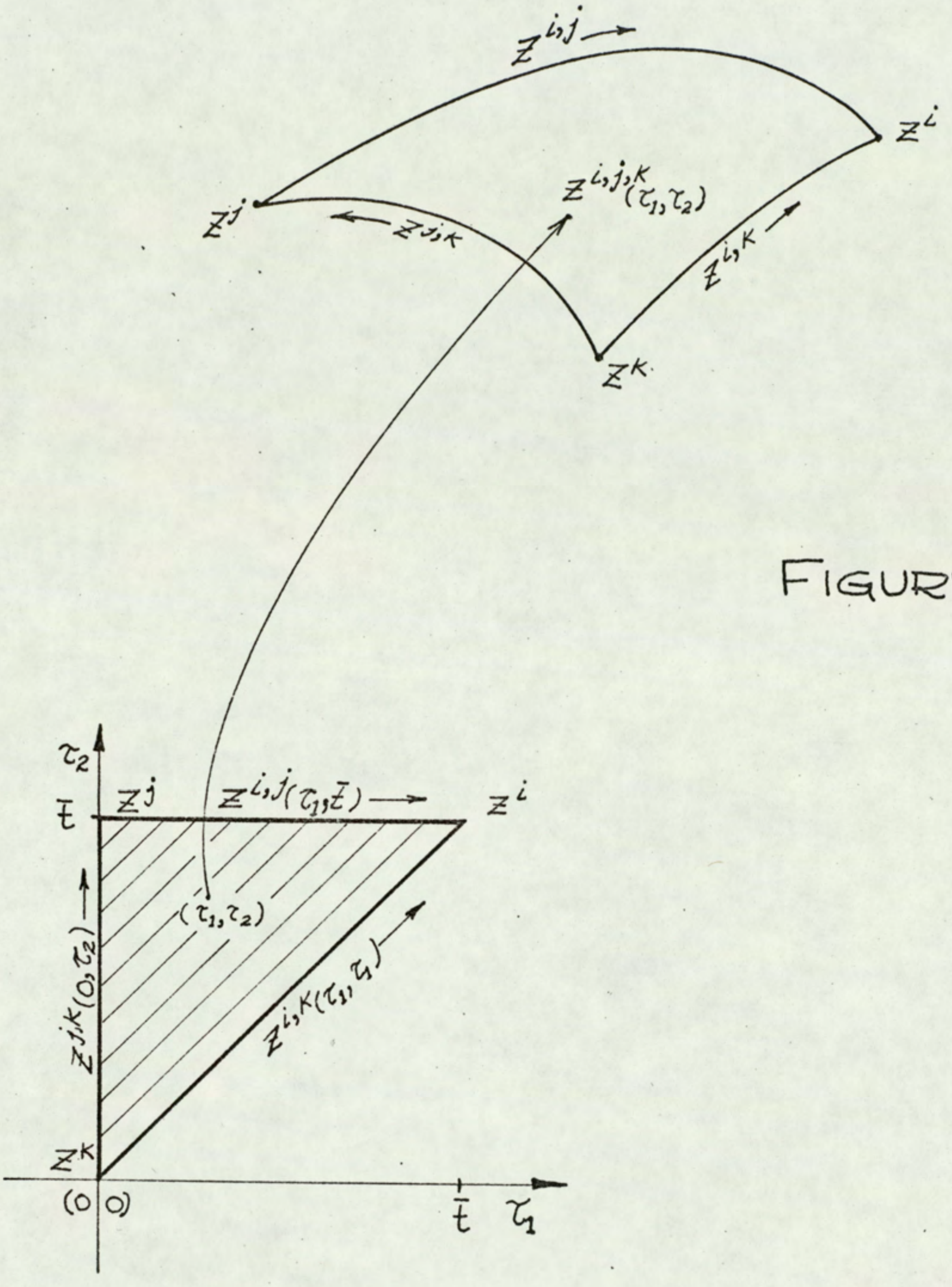


FIGURE 11

The two-switch loci

$$Z^{i,j,k}_{(\tau_1,\tau_2)} = x^i + e^{-A\bar{t}}(\bar{x} - x^k) - e^{-A\tau_1}(x^i - x^j) - e^{-A\tau_2}(x^j - x^k).$$

and so we must select, for each point x, that control sequence m^i, m^j, m^k (i,j,k = 1,2,3,4) which transfers x to $\bar{x}$ in the minimum possible time.

Let us consider a point $s^j (= s^j(\bar{t}))$ which we assume lies on the boundary $T(x) = \bar{t}$. The single-switch control sequences associated with s^j are $s^{i,j}$ and $s^{j,k}$ for i,k = 1,2,3,4 $i \neq j$, $j \neq k$. The loci corresponding to optimal control sequences must be convex curves and bound the largest possible area. The choice between $s^{i,j}$ and $s^{j,i}$ is to be made on this basis. For normal systems these curves must be distinct, otherwise the optimal control would not be unique. It is not necessary to even sketch these loci, for it is sufficient to plot the tangent vectors to the curve at s^j. The tangent vector to $s^{i,j}(\tau)$ at $s^j = s^{i,j}(o)$ is

$$\dot{s}^{i,j}(o) = A(x^i - x^j) \equiv T^{i,j} \tag{3.29}$$

which is independent of $\bar{t}$. The tangent to $s^{j,i}$ at $s^j = s^{j,i}(\bar{t})$ is found by describing $s^{j,i}$ in the reverse order, i.e.

$$\bar{s}^{i,j}(\tau) = s^{j,i}(\bar{t} - \tau). \tag{3.30}$$

Then the tangent to (3.30) at s^j is

$$\dot{\bar{s}}^{i,j}(o) = -\dot{s}^{j,i}(\bar{t}) = Ae^{-A\bar{t}}(x^i - x^j) \equiv \bar{T}^{i,j}. \tag{3.31}$$

Simply plotting the vectors $T^{i,j}$, $\bar{T}^{i,j}$ at s^j for $i \neq j$ and selecting the two which are extreme will be sufficient to solve many two-dimensional problems. We illustrate this by an example.

Example 3.I.

Let $A = \begin{pmatrix} -1 & o \\ o & -2 \end{pmatrix}$ and $G = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$.

We shall consider the problem of optimally controlling this system to the reference point $\bar{x} = o$ by two-way controllers;

therefore, $\mathcal{C}$ is the symmetric 2-cube which has the extreme points $m^1 = (-1,-1)$, $m^2 = (1,-1)$, $m^3 = (1,1)$, $m^4 = (-1,1)$. The system is normal and controllable.

The exponential polynomials which determine the switching instants are given by

$$p(t) = (\ p_1(t),\ p_2(t)\ ) = (\alpha,\beta)e^{-At}G$$

so that

$$\begin{aligned} p_1(t) &= 2\alpha e^t - \beta e^{2t} \\ p_2(t) &= -\alpha e^t + \beta e^{2t} \end{aligned} \qquad (3.32)$$

If $s^1(\bar{t})$ is to lie on the surface $T(x) = \bar{t}$ the control m^1 must be optimal over that interval which means that there exist α and β such that

$$\begin{aligned} p_1(t) &= 2\alpha e^t - \beta e^{2t} \geqslant 0 \\ p_2(t) &= - \ e^t + \ e^{2t} \geqslant 0 \ . \end{aligned} \qquad (0 \leqslant t \leqslant \bar{t})$$

It is simple to see that this requires $\bar{t} \leqslant \log 2$. A similar result holds for $s^3(\bar{t})$ (= $-s^1(\bar{t})$). For $s^2(\bar{t})$ to lie on the boundary, we must be able to find α, β so that

$$\begin{aligned} 2\alpha e^t - \beta e^{2t} &\leqslant 0 \\ -\alpha e^t + \beta e^{2t} &\geqslant 0 \ . \end{aligned} \qquad (0 \leqslant t \leq \bar{t})$$

The first of these requires $\beta e^t \geqslant 2\alpha$ while the second gives $\beta e^t \geqslant \alpha$ so if β is chosen to be greater than the maximum of α and 2α, and positive, the inequalities hold for all $\bar{t}$. Thus $s^2(\bar{t})$ and $s^4(\bar{t})$ (= $-s^2(\bar{t})$ lie on the boundary for all $\bar{t} < \infty$.

These results indicate that the nature of the isocronal surfaces change when $\bar{t}$ passes the point log 2 so we shall consider each case separately.

For $\bar{t} = \log 2$ we plot the curves $s^j(t)$, $o \leq t \leq \log 2$. The point $s^j(\log 2)$ will again be denoted by s^j. Calculating the tangent vectors $T^{i,j} = \dot{s}^{i,j}(o)$ and $\bar{T}^{i,j} = -\dot{s}^{j,i}(\log 2)$ and plotting them at the points s^j, we see that

at s^1, $T^{4,1}$ and $\bar{T}^{2,1}$ are extreme,

at s^2, $T^{1,2}$ and $\bar{T}^{3,2}$ are extreme,

at s^3, $T^{2,3}$ and $\bar{T}^{4,3}$ are extreme,

at s^4, $\bar{T}^{1,4}$ and $T^{3,4}$ are extreme,

the other tangent vectors lying between the given vectors (see Figure 12). Of course, if $T^{4,1}$ is extreme at s^1, $\bar{T}^{1,4}$ must be extreme at s^4 as they are tangents to the same curve. If, instead, $T^{1,4}$ were extreme at s^4 it would indicate that the curve $s^{1,4}(\tau)$ was inside the curve $s^{4,1}(\tau)$ near s^4 but outside it at s^1 which, as we are working in two dimensions, would require that they cross. The crossing point would have two different optimum control sequences which is impossible. From this, we see it is only necessary to plot the tangent vectors at two of the four points s^i. We will show below that if s^i and s^j lie on the boundary and the tangent vector $T^{i,j}$ is extremal at s^j then the entire curve $s^{i,j}(\tau)$, $o \leq \tau \leq \bar{t}$, lies on the boundary. Thus the isocronal surface $T(x) = \log 2$ is formed by the curves $s^{4,1}(\tau)$, $s^{1,2}(\tau)$, $s^{2,3}(\tau)$ and $s^{3,4}(\tau)$, $o \leq \tau \leq \log 2$. If all possible system start-up locations are within this region, the problem is solved by the strategy shown in Figure 12. The switching curves are given parametrically by the curves $s^i(t) = x^i - e^{-At}x^i$, $i = 1,2,3,4$ $o \leq t \leq \log 2$.

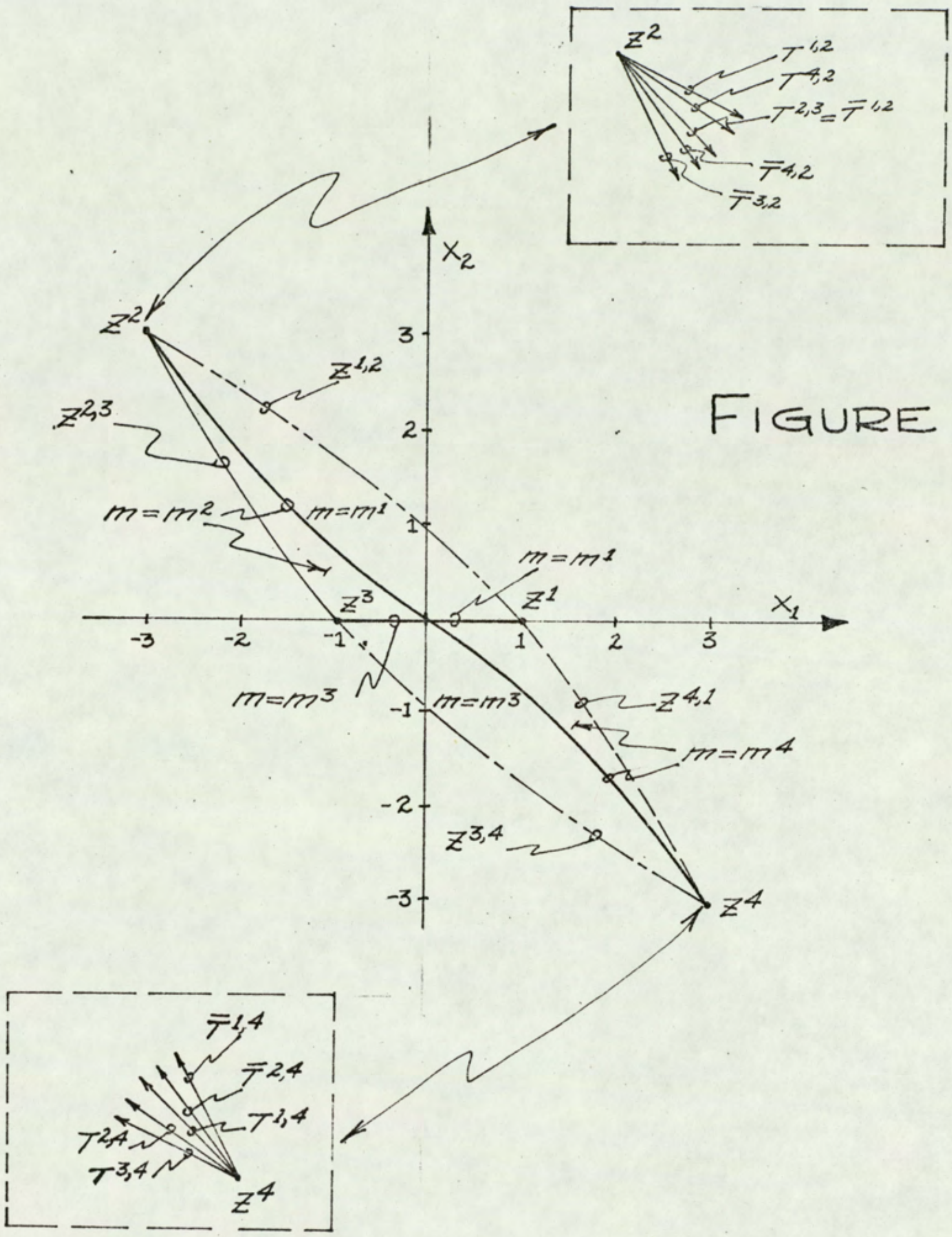


FIGURE 12

ISOCRONAL SURFACE $T(x) = \log 2$

AND

OPTIMAL SWITCHING POLICY

To illustrate the switching behaviour when $t > \log 2$ we choose $\bar{t} = \log 5$. We again plot the locus of $s^i(t)$, $o \leq t \leq \log 5$ for $i = 1,2,3,4$ and calculate the tangent vectors $T^{i,j}$ and $\bar{T}^{i,j} = -\dot{s}^{j,i}(\log 5)$. The results found above still apply for s^2 and s^4 but s^1 and s^3 are no longer on the boundary so they provide no information. We will show that in this case only a portion of the curves $s^{4,1}(\tau)$ and $s^{1,2}(\tau)$ from the boundary of $T(x) = \log 5$. To do this we suppose for the moment that s^i and s^j are on the boundary and $T^{i,j}$ and $\bar{T}^{j,i}$ (both referring to $s^{i,j}(\tau)$) are extremal at s^j and s^i respectively. If $N^{i,j}$ is an outer normal to the curve $s^{i,j}$ at s^j then $N^{i,j}.T^{i,j} = o$ and $N(\tau) = N^{i,j}e^{A\tau}$ is an outer normal to this same curve at $s^{i,j}(\tau)$ for $\dot{s}^{i,j}(\tau) = e^{-A\tau}T^{i,j}$. Now since m^j was optimal we have

$$\min_{m \text{ in } \mathscr{C}} N^{i,j}e^{-At}Gm = N^{i,j}e^{-At}Gm^j, \quad o \leq t \leq \bar{t},$$

and similarly as $N^{i,j}e^{A\bar{t}}$ is a normal at s^i

$$\min_{m \text{ in } \mathscr{C}} N^{i,j}e^{A(\bar{t}-t)}Gm = N^{i,j}e^{A(\bar{t}-t)}Gm^i, \quad o \leq t \leq \bar{t}.$$

These two results show that

$$\begin{aligned}\min_{m \text{ in } \mathscr{C}} N^{i,j}e^{At}Gm &= N^{i,j}e^{At}m^i, \qquad o < t \leq \bar{t} \\ &= N^{i,j}e^{At}m^j \qquad -\bar{t} \leq t \leq o.\end{aligned} \tag{3.33}$$

Thus when we consider the normal $N(\tau) = N^{i,j}e^{A\tau}$ at $s^{i,j}(\tau)$ we see that

$$\min_{m \text{ in } \mathscr{C}} N(\tau)e^{-At}Gm = \min_{m \text{ in } \mathscr{C}} N^{i,j}e^{A(\tau-t)}Gm$$

is assumed for $m = m^i$ when $t < \tau$ and for $m = m^j$ when $t > \tau$. Thus we have shown that the m^i, m^j control policy is optimal for each point of $s^{i,j}(\tau)$ and thus they lie on the surface.

On the other hand, if $s^i(\bar{t})$ lies on the surface only for $\bar{t} \leqslant t^1$ then the equation (3.33) can hold only for such $\bar{t}$ and the conclusion we have drawn only holds for $\tau \leqslant t_1$.

Thus, in the example we are concerned with here, $s^1(\bar{t})$ is not on the boundary if $\bar{t} > \log 2$, thus we can conclude that, when $\bar{t} > \log 2$, the points $s^{1,2}(\tau)$ lie on the boundary only for $0 \leqslant \tau \leqslant \log 2$. For $\tau > \log 2$ the policy "m^1 into m^2 at τ" is not optimal. The locus of these points which bound the region of maximum application of m^1 is given by

$$y(t) = s^{i,j}(\log 2; t) = s^j(t) + (I - e^{-A \log 2})(x^i - x^j)$$

where $\log 2 \leqslant t < \infty$ (see Figure 13). The remainder of the isocronal surface must be constructed of points whose optimal control policy involves two switching instants.

Even for this low dimensional problem there are 24 distinct switching combinations involving two switching instants; however, as the single switch policies are optimal around the points s^2 and s^4, the optimal two-switch locus must lie inside the single-switch locus near these points and only emerge when the single-switch policy is no longer adequate. The only two-switch policies which agree with the single-switch policies near s^2 and s^4 are $s^{4,1,2}_{(\tau_1,\tau_2)}$ and $s^{2,3,4}_{(\tau_1,\tau_2)}$ as is shown in Figure 14. Once having determined the isocronal surface the switching surfaces are easily determined, as indicated in Figure 13.

This example illustrates a typical two-dimensional problem. If the problem is not normal, two or more of the curves $z^{i,j}(\tau)$ will coincide. If the eigenvalues of A are complex the number of switches can be large but, practically, one can restrict

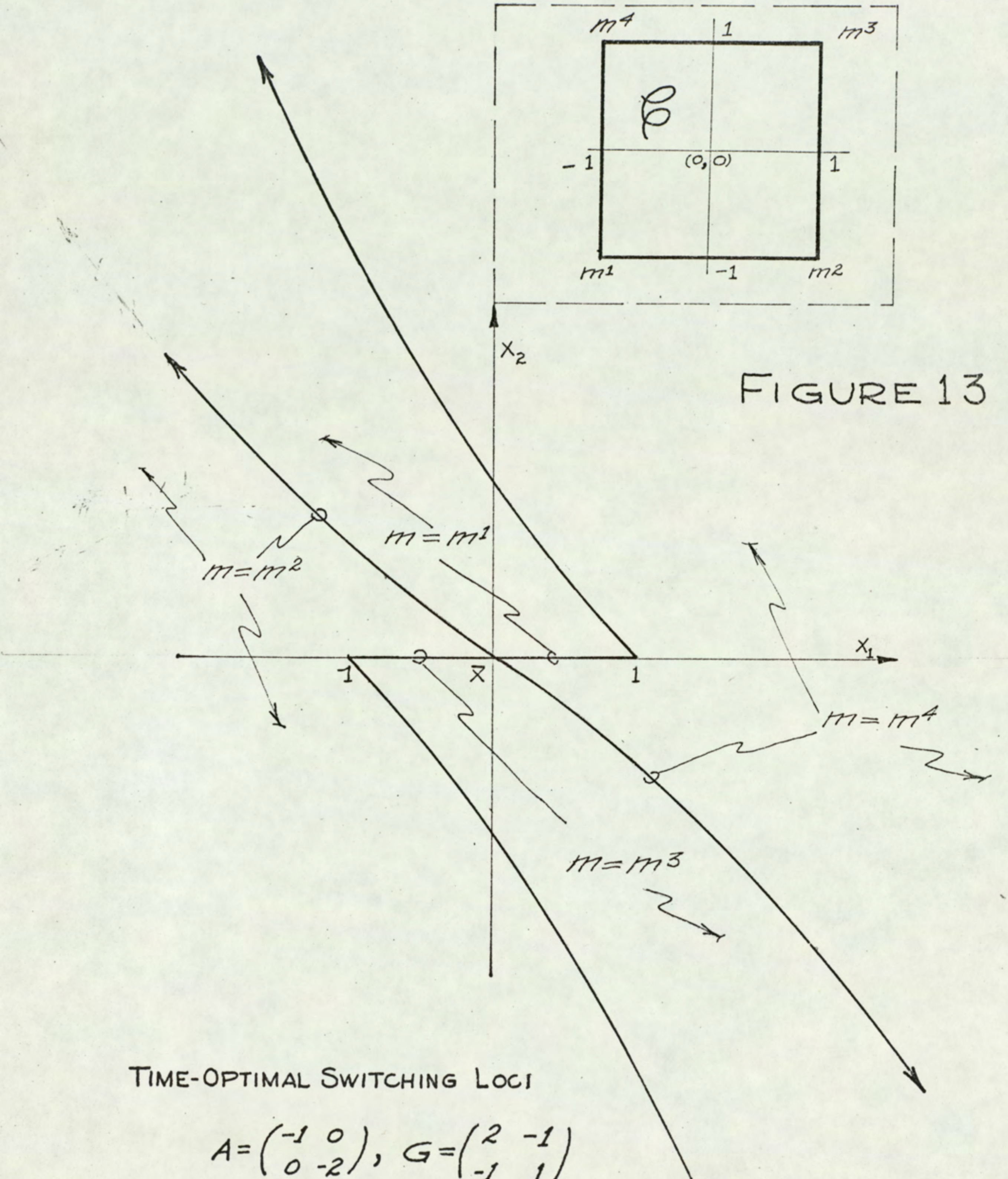


FIGURE 13

TIME-OPTIMAL SWITCHING LOCI

$$A = \begin{pmatrix} -1 & 0 \\ 0 & -2 \end{pmatrix}, \quad G = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}$$

$$\bar{x} = 0$$

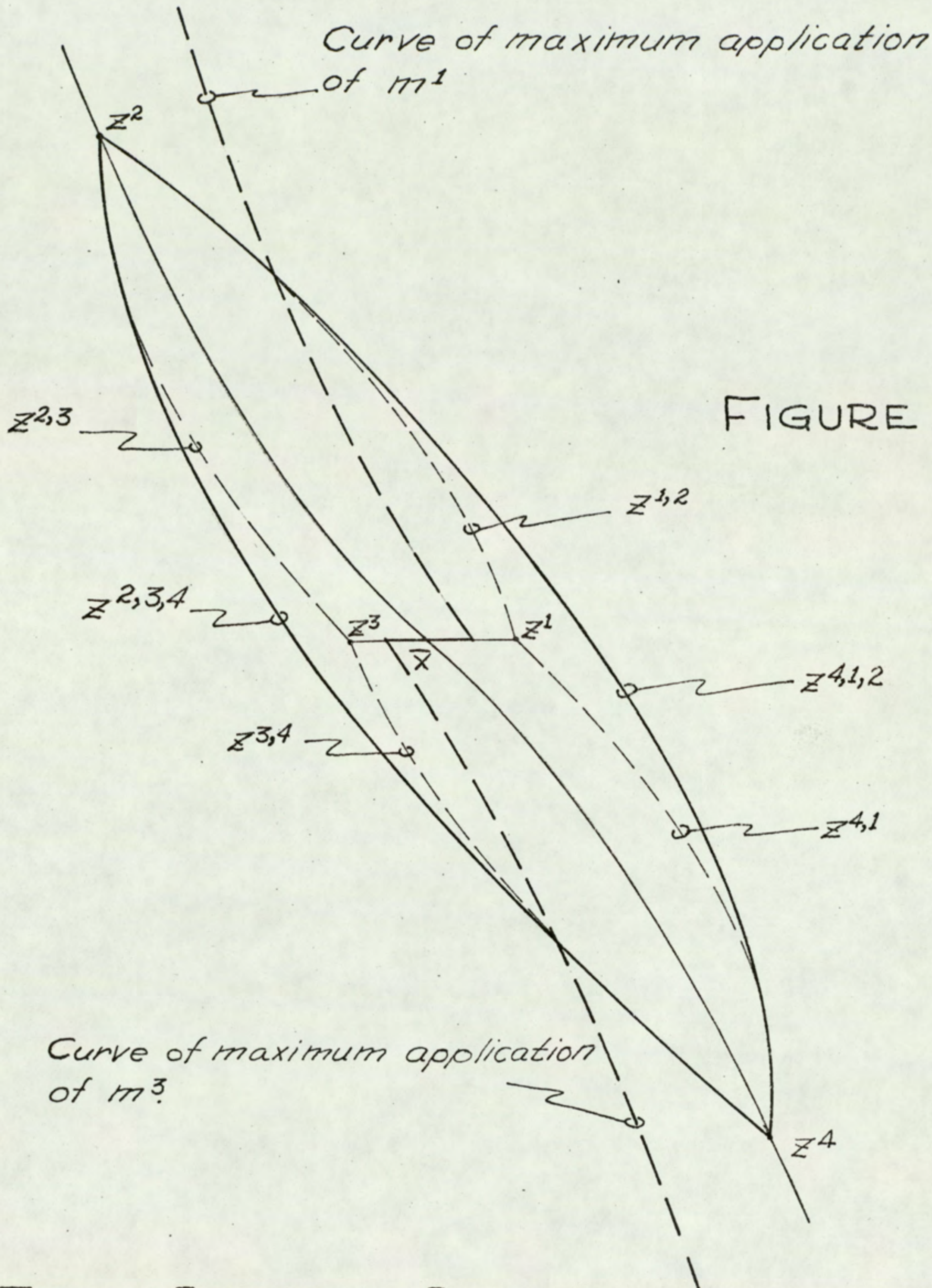


FIGURE 14

TYPICAL ISOCRONAL SURFACE

$\bar{t} > \log 2$

consideration to a finite $\bar{t}$, in which case only a finite number of switches occur, the number diminishing as $\bar{t}$ decreases. We will illustrate this with a simple example.

Example 3.2

A simple control problem involving the harmonic oscillator considers the equation $\ddot{u} + u = m$. This leads to the vector equation (when $x_1 = u$, $x_2 = \dot{u}$),

$$\dot{x} = \begin{pmatrix} o & 1 \\ -1 & o \end{pmatrix} x + \begin{pmatrix} o \\ 1 \end{pmatrix} m.$$

We shall consider the time-optimal problem of controlling this system at the point $\bar{x} = (\frac{3}{4}, o)$ by on-off controllers. The control domain is the interval $o \leq m \leq 1$. The system is normal and controllable. The eigenvalues of A are $\pm i$ which lead to

$$e^{At} = \begin{pmatrix} \cos t & \sin t \\ -\sin t & \cos t \end{pmatrix}.$$

The exponential polynomial equation which determines the switching instants is

$$(\alpha, \beta) e^{-At} g = -\alpha \cos t + \beta \sin t = \cos(t + \theta)$$

if $\alpha = -\cos\theta$, $\beta = \sin\theta$. The control $m^o = o$ will be optimal over the interval $(o, \bar{t})$ provided θ is such that $\cos(t + \theta) \geq o$, $o \leq t \leq \bar{t}$. The choice $\theta = -\pi/2$ makes $\bar{t} = \pi$ the maximum possible. Similarly $m^1 = 1$ is optimal only over a time interval $o \leq t \leq \pi$, as we see when $\theta = \pi/2$. As we have seen in Example 3.1, provided a time interval of length π is sufficient to bring all necessary start-up points to $\bar{x}$, we can complete the analysis by constructing the single-switch loci $s^{1,o}_{(\tau)}$ and $s^{o,1}_{(\tau)}$, $o \leq \tau \leq \pi$, which define the boundary of the isocronal surface $T(x) = \pi$. On the other hand it may be necessary to extend the time interval, in which case multiple-switch loci must be determined. For the purposes of this example, we take $\bar{t} = 3\pi/2$.

From the switching equation $\eta_o e^{-At} gm = \cos(t + \theta)m$ we observe that the switching instants on any optimal control occur π time-units apart so that for $\bar{t} = 3\pi/2$ there can be no more than 2 switches in any optimal control.

Consider now the single-switch loci $s^{1,o}(\tau)$ and $s^{o,1}(\tau)$ $o \leq \tau \leq 3\pi/2$. In this one-dimensional control situation both curves are required; however, as the points $s^1(3\pi/2) = s^{o,1}(o)$ and $s^o(3\pi/2) = s^{1,o}(o)$ are not on the boundary, we cannot assert that any part of these loci form the boundary. The loci of maximal m^o application is found from $s^{o,1}(\pi;t) = x^o + e^{-At}(\bar{x}-x^1) -e^{-A\pi}(x^o - x^1)$, $\pi \leq t \leq 2\pi$ which with $x^o = (o,o)$, $x^1 = 1,o)$, $\bar{x} = (\frac{3}{4},o)$ becomes

$$s^{o,1}(\pi;t) = \begin{pmatrix} -\frac{1}{4}\cos t - 1 \\ -\frac{1}{4}\sin t \end{pmatrix} \qquad \pi \leq t \leq 2\pi$$

which is a semi-circle about $(-1,o)$ with radius $\frac{1}{4}$. This curve is a switching curve "m^1 into m^o". The switching curve "m^o into m^1" is found as the loci of maximal application of m^1 which leads to the curve

$$s^{1,o}(\pi;t) = \begin{pmatrix} \frac{3}{4}\cos t + 2 \\ \frac{3}{4}\sin t \end{pmatrix} \qquad \pi \leq t \leq 2\pi$$

which is again a circle which is centred at $(2,o)$ with radius $\frac{3}{4}$. The two-switch loci $s^{1,o,1}$ and $s^{o,1,o}$ form the remainder of the boundary of $T(x) = 3\pi/2$. These curves are shown in Figure 15.

Let us now consider a three dimension problem. We shall briefly discuss only the simplest case of a one dimensional controller and a matrix A having real eigenvalues.

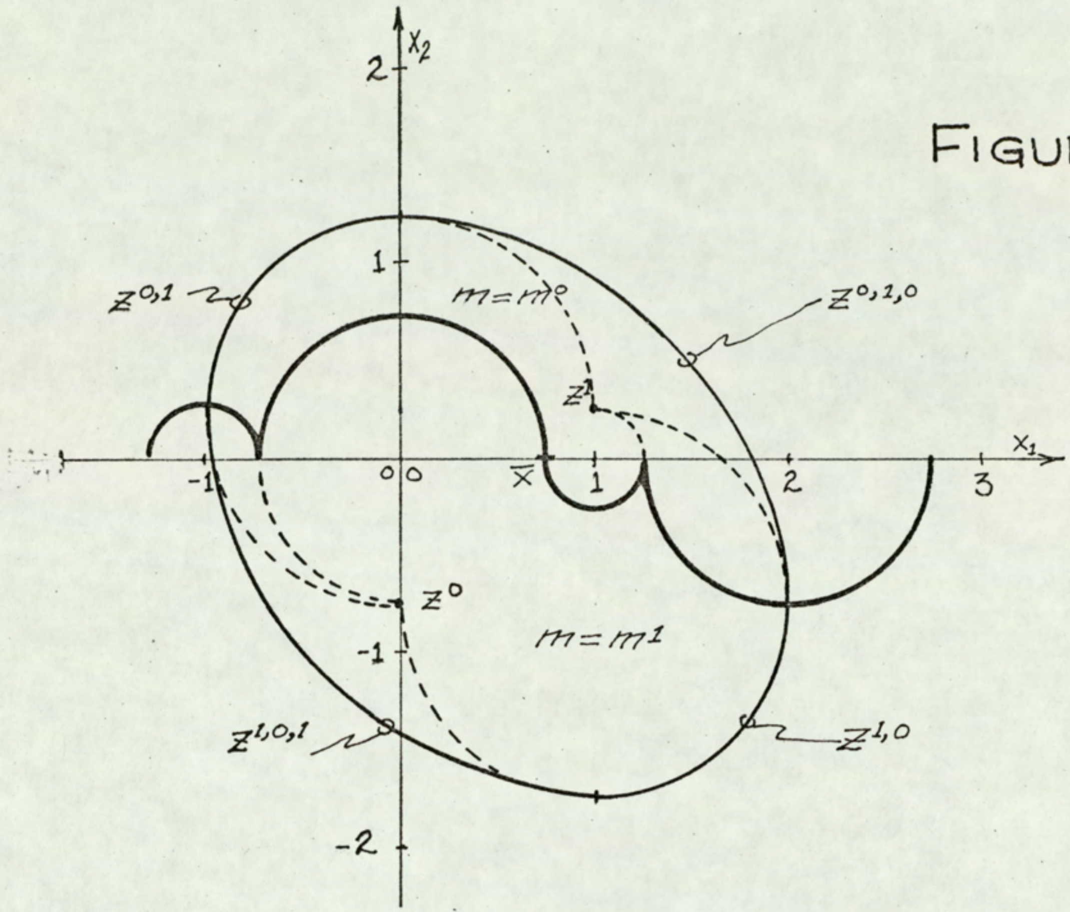


FIGURE 15

ISOCRONAL SURFACE AND SWITCHING CURVES

$$A=\begin{pmatrix}0 & 1\\ -1 & 0\end{pmatrix} \quad G=\begin{pmatrix}0\\ 1\end{pmatrix}$$

$$m^0=0 \; ; \; m^1=1 \; ; \; \bar{t}=3\pi/2 .$$

Example 3.3

We take

$$A = \begin{pmatrix} 1 & o & o \\ o & -1 & o \\ o & o & -2 \end{pmatrix} , \quad g = \begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix}$$

and seek to control the system

$$\dot{x} = Ax + gm$$

at the point $\bar{x} = (-\frac{1}{2}, 1, -\frac{1}{4})$ by on-off controllers. Accordingly $\mathcal{C}$ is the unit interval $o \leqslant m \leqslant 1$. The extreme points of $\mathcal{C}$ are $m^o = o$, $m^1 = 1$.

Examination of $p(t) = \eta_o e^{-At}.g$ shows that m^o and m^1 are optimal for all $\bar{t}$ and hence the corresponding points $s^o(\bar{t})$ and $s^1(\bar{t})$ lie on the isocronal surface $T(x) = \bar{t}$ for all $\bar{t}$. As in the previous example, the single-switch loci $s^{1,o}(\tau)$, $s^{o,1}(\tau)$, $o \leqslant \tau \leqslant \bar{t}$ are both required, because of the one-dimensional controller, but here they both lie on the boundary surface as $s^1(\bar{t})$ and $s^2(\bar{t})$ lie on it. The two-switch loci complete the surface $T(x) = \bar{t}$ as, in this case, $r.(n-1) = 2$ is the maximal number of switches in any optimal control. A sketch of the surface $T(x) = \bar{t}$ is shown in Figure 16.

Analysis of Figure 16 shows that the state-space is divided into two regions. The control m is o in one and 1 in the other. Further we have a parametric description of the switching surfaces.

A switch from o to 1 is made at a point on the one-dimensional curve $s^1(t)$, $o \leqslant t < \infty$ or on the two-dimensional surface $s^{1,o}(\tau;t)$, $o < \tau \leqslant t < \infty$.

A switch from 1 to o is made at a point on the one-dimensional

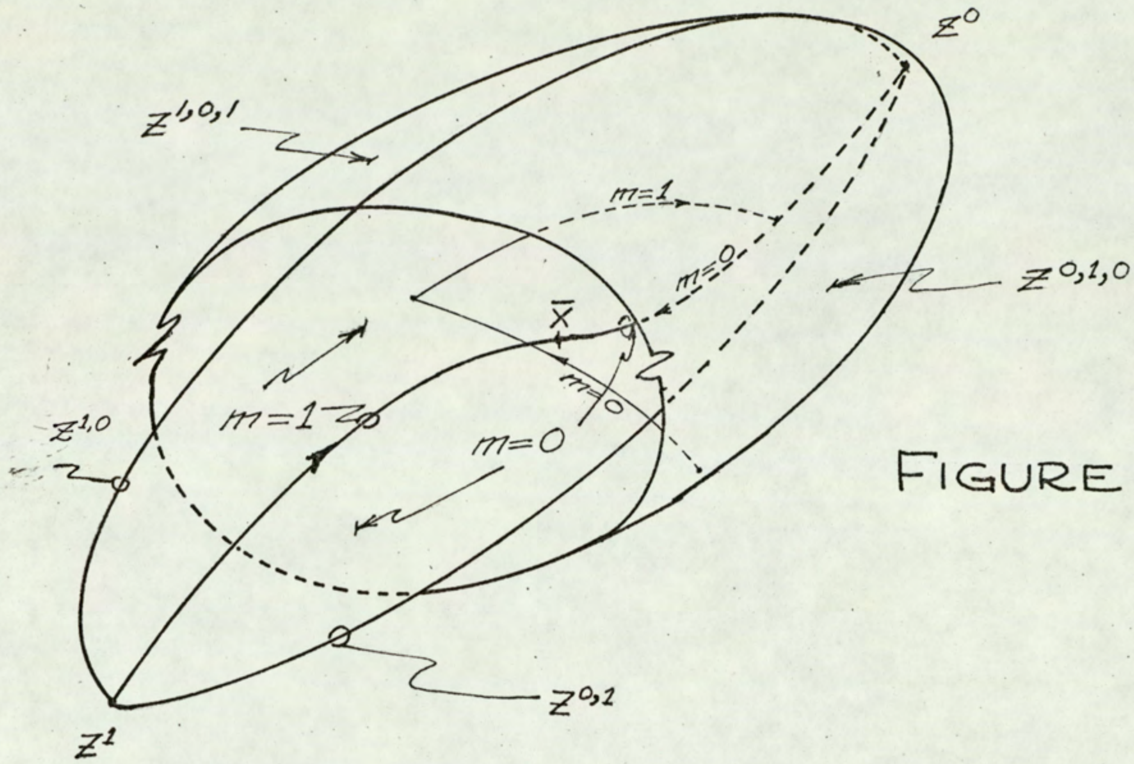


FIGURE 16

TYPICAL ISOCRONAL SURFACE
AND SWITCHING SURFACES

($n=3$, $r=1$, real eigenvalues)

curve $s^{o}(t)$, $o \leq t < \infty$ or on the two-dimensional surface $s^{o,1}(\tau;t)$, $o < \tau \leq t < \infty$.

The last example is valuable in that it indicates what can be expected in higher dimensional problems. There we should expect to have a sequence of switching surfaces having increasing dimension 1, 2, 3, ..., n-1. Adequately describing these loci is extremely difficult, however the analysis does provide a parametric description of the switching surfaces. The parametric description can be used to construct linear or quadratic (or higher degree) approximations to the switching surfaces which can be easily simulated by computers and thereby provide a feed-back control device. Approximations of this type could be used to provide sub-optimal controls of an efficient type.

Summary of Chapter Three.

In this chapter we have considered the time-optimal control of linear constant-coefficient systems. The set of points controllable to a fixed reference point in time $\bar{t}$ has been shown to be a closed convex set containing the reference point. This convex set is bounded by the isocronal surface $T(x) = \bar{t}$. The normal vectors to the isocronal surface determine the optimal control through the adjoint differential equation and the Pontryagin Minimum Principle. We have studied the geometric properties of the isocronal surfaces and their relation to the loci of points parametrized by various switching policies. It has been shown that in two dimensional problems this provides a facile method of determining the switching surfaces which are described by parametric equations. In higher dimensional problems with a one-dimensional control the methods developed here also provide insight into the behaviour of the optimal system and, further, a parametric description of the switching surfaces which may then be approximated to provide sub-optimal control.

The methods used in the lower-dimensional cases can also be used to study complex systems consisting of several lower-dimensional problems weakly coupled together. For general systems with several degrees of freedom and controllers of more than one dimension, the geometric methods are not directly applicable but must be translated into analytical form. Due to the complexity of such systems, this will require use of a computer for, in the simplest case of a three-dimensional system and two-dimensional control, the optimal controls will have four switching instants, with a choice of four controls, making a total of $4.3^4 = 324$ possible distinct four-switch loci. It is hoped that further study will greatly simplify this problem.

Even in the higher dimensional systems, the concepts developed here are useful. The usual procedure of finding the optimal control policy for such systems is to use the two-point boundary problem arising from the maximum principle, or a modified method of steepest descent, to calculate the optimal control for a fixed point. This information is stored in the computer and the process repeated. From the analysis here we see that the calculation of one optimal control furnishes a point $x^1 = s^J(\bar{\tau},.\bar{t})$ on the isocronal surface with switching points given by $\bar{\tau} = (\bar{\tau}_1, \bar{\tau}_2, \ldots, \bar{\tau}_N)$, $(0 < \bar{\tau}_1 < \bar{\tau}_2 < \ldots < \bar{\tau}_N < \bar{t})$ and with optimal policy $m^{j_0} \rightarrow m^{j_1} \rightarrow \ldots \rightarrow m^{j_N}$ where $J = (j_0, j_1, \ldots, j_N)$. Since the functions have been described (see (3.17)) and are continuous we can vary all the parameters $\tau_1, \tau_2, \ldots, \tau_N$ and t slightly to obtain the optimal strategy for points near x^1. This also furnishes the optimal control for all points on the trajectories from $s^J(\tau; t)$ to the reference point. Thus it is seen that a calculation of the optimal control at x^1 furnishes the optimal control policy for all points in a "tube" about the calculated optimal trajectory.

From a design standpoint we have shown under what conditions a two-dimensional control system can be optimally controlled by a single-switch trajectory. This will occur when the constant application of the extreme controls are optimal. If the columns of the gain matrix are linearly independent this will occur for at least a positive time period. It will occur for all times provided the system matrix A has real roots and the product $A^{-1}G$ is not "too skew". This is seen by considering the extreme cases of, first, $A^{-1}G = I$, in which case the image of $\mathcal{C}$ (the set S of Chapter 1) is again $\mathcal{C}$ and all controls are treated equally and all constant controls are optimal; or secondly, when $A^{-1}G$ maps $\mathcal{C}$ into a line segment and completely suppresses the effect of two of the controls and only two of the constant extreme controls are optimal. This behaviour furnishes another factor somewhat under the control of the designer as it again depends upon the gain matrix. Below we summarize the approach developed in this chapter for optimizing two-dimensional systems.

1. Plot the loci of the $s^i(\bar{t})$ obtained from the constant application of $m = m^i$.

2. Examine the exponential-polynomial switching functions obtained from the adjoint system to determine for what values of $\bar{t}$, $z^i(\bar{t})$ is on the boundary $T(x) = \bar{t}$.

3. If $s^j(\bar{t})$ is on the boundary, plot the tangent vectors,

$$T^{i,j} = A(x^i - x^j) = G(m^i - m^j)$$

and

$$\bar{T}^{i,j} = e^{-A\bar{t}}T^{i,j},$$

to the single-switch loci at $s^j(\bar{t})$.

4. Determine the extreme tangent vectors at each point. If $T^{i,j}$ is extreme then $s^{i,j}(\tau)$ is optimal for all τ, $o \leqslant \tau \leqslant \bar{t}$, when $s^i(\bar{t})$ is on the boundary, and for $o \leqslant \tau \leqslant t_1$, if $s^i(t)$ is only on the boundary for $o \leq t \leq t_1$.

5. If some $s^i(\bar{t})$ is not on the surface $T(x) = \bar{t}$, determine the two-switch policies which agree with the single-switch policies found in 4, where they are optimal.

6. If the system matrix has real eigenvalues the solution is complete, otherwise policies with additional switching must be considered.

7. The switching curves are given parametrically by $s^j(t) = x^j + e^{-At}(\bar{x} - x^j)$ $(j = 1,2,3,4)$ for $o \leq t \leq t_1$, where t_1 is the largest value with $T(s^j(t_1)) = t_1$ (i.e. the largest t_1 with $s^j(t_1)$ on the isocronal surface $T(x) = t_1$) and are then given by

$$s^{j,k}(t_1;t) = x^j + e^{-At}(\bar{x} - x^k) - e^{-At_1}(x^j - x^k)$$

for $t_1 \leqslant t < \infty$, where $s^{j,k}(\tau;t_1)$ is the optimal control loci for $\bar{t} = t_1$ selected by the methods of 4.

We have not considered the effects of time lags in the control operations here. This adds another complication which will be studied in further work. In this respect the above analysis represents the behaviour of systems without inertia and with ideal relays. Any actual control system will be sub-optimal but can be compared with the ideal situation as found by the methods developed here.

There is also much further work to be done in modifying the approach used here so as apply to systems with other cost functions, to systems with time varying coefficients and to non-linear systems linearized about some optimal trajectory.

APPENDIX A.

Linear constant-coefficient differential equations and matrix exponentials.

Given two n x n matrices $A = (a_{i,j})$ and $B = b_{i,j})$ it is possible to define not only their sum, $A + B = (a_{ij} + b_{ij})$, and their product, $A.B = (\sum_{k=1}^{n} a_{ik} b_{kj})$, but also a measure of "distance" between them, $|A-B| = \left[\sum_{i,j=1}^{n} (a_{ij} - b_{ij})^2\right]^{\frac{1}{2}}$.

$|A| = |A - O|$ is called the <u>norm</u> of A. It is easy to show that $|A.B| \leq |A|.|B|$.

With this notion of distance between matrices, we can talk about convergence of sequences, i.e. a sequence of matrices $A_1, A_2, \ldots, A_n, \ldots$ converges to the matrix A if, and only if, $\lim_{n\,\infty} |A-A_n| = o$. In particular we can define power series of matrices as limits of partial sums of the form $S_n = \sum_{k=o}^{n} \alpha_k A^k$. The partial sums satisfy $\left|\sum_{k=o}^{n} \alpha_k A^k\right| \leq \sum_{k=o}^{n} |\alpha_k| \; |A^k| \leq \sum_{k=o}^{n} |\alpha_k| \, (|A|)^k$ and it can be shown that if $\sum_{k=o}^{n} |\alpha_k| \, (|A|)^k$ converges (as $n \rightarrow \infty$) to a real number, the sequence of partial sums $\sum_{k=o}^{n} \alpha_k A^k$ converge to a unique matrix. Since the series $\sum_{k=o}^{\infty} \frac{1}{k!} |A|^k$ converges for all $|A|$, we find that the matrix series $\sum_{k=o}^{\infty} \frac{1}{k!} A^k$ also converges for each matrix A. This defines the <u>matrix exponential</u>

$$e^A \equiv \sum_{k=o}^{\infty} \frac{1}{k!} A^k .$$

If the matrices A and B commute (i.e. AB = BA) then

$$e^{(A+B)} = e^A e^B \;;$$

since A and -A always commute we find

$$e^A . e^{-A} = e^o = I$$

so that $e^{-A} = (e^A)^{-1}$. Further, since each matrix satisfies its own characteristic equation, we have

$$A^n = \sum_{k=o}^{n-1} \alpha_k A^k$$

for some $\alpha_o, \alpha_1, \ldots, \alpha_{n-1}$ and, as the powers of A^k are therefore seen to be dependent, we can write

$$e^A = \sum_{k=o}^{n-1} \beta_k A^k$$

for some choice of the β_k, k=o, 1, 2,..., n-1.

If A is replaced by tA where t is real we find

$$\frac{d}{dt} e^{tA} = Ae^{tA} = e^{tA}A.$$

Thus, in seeking to solve the vector differential equation

$$\dot{x} = Ax + F(t),$$

where F is an n-dimensional vector function, we multiply by e^{-tA} and observe, as in the scalar case,

$$e^{-tA}\dot{x} - e^{-tA}Ax = \frac{d}{dt}\left(e^{-tA}x\right) = e^{-tA}F(t)$$

and

$$e^{-tA}x = x(o) + \int_o^t e^{-sA}F(s)ds$$

or

$$x(t) = e^{tA}x(o) + \int_o^t e^{(t-s)A}F(s)ds$$

which is the solution to the differential equation with initial value x(o).

For the control systems considered here the function $F(t) = Gm(t)$ where G is a constant n x r matrix and m(t) is the r-dimensional control vector. In this case the solution becomes

$$x(t) = e^{At}x(o) + \int_o^t e^{(t-s)A}Gm(s)ds.$$

<u>Controllability.</u>

In Chapter One the concept of controllability is discussed. Within this discussion it is found necessary to characterize those matrices A and G for which a vector v can be found which satisfies

$$v.e^{tA}G \equiv 0$$

As we have stated, functions $\alpha_o(t), \alpha_1(t), \ldots, \alpha_{n-1}(t)$ can be found so that $e^{tA} = \sum_{k=o}^{n-1} \alpha_k(t)A^k$. Using this we see that $v.e^{tA}G = \sum_{k=o}^{n-1} \alpha_k(t)\ v.A^kG \equiv o$. Since the functions $\alpha_o, \alpha_1, \ldots, \alpha_{n-1}$ are independent this means $v.G = v.AG = \ldots = v.A^{n-1}G = 0$, which implies rank $[G, AG, \ldots, A^{n-1}G] < n$. A system is called <u>controllable</u> when this can't happen, hence, for controllable systems, rank $[G, AG, A^2G, \ldots, A^{n-1}G] = n$.

APPENDIX B.

Orders of Magnitude.

The comparison of the growth behaviour of functions is facilitated by defining orders of magnitude which represent a comparison of the functions with certain standard functions. For example, a function $f(t)$ is said to be $O(t)$ near zero if

$$|f(t)| \leq K|t|$$

for some constant K and sufficient small t, e.g. $\sin t = O(t)$ for t near zero.

Similarly, $f(t)$ is said to be $O(t^n)$ near o if a constant K exists so that

$$|f(t)| \leq K|t|^n$$

for t sufficiently small.

If $F(t)$ is a vector function, we say $F(t)$ is $O(t^n)$ near zero if its length, $|F(t)|$, is $O(t^n)$ near zero.

BIBLIOGRAPHY

[1] Roots, W.K. and Nightingale, J.M. :
Temperature control in electric space heating.
IEEE Transactions on Applications and Industry,
Jan. 1964, pp 1 - 69.

[2] Lee, E.B. and Marckus, L. :
Foundations of optimal control.
Wiley, 1967.

[3] Roots, W.K. and Walker, F. :
Closed-loop controlled electric space-heating and space-cooling processes.
IEEE Transactions on Industry and General
Applications, Sept./Oct. 1966, pp 402 - 416.

[4] Roots, W.K. and Walker, F. :
Dynamic equilibrium cycling in discontinuously controlled electric process heating.
IEEE Transactions on Industry and General
Applications, March/April 1967.

[5] Flugge-Lotz, I. :
Discontinuous and Optimal Control.
McGraw-Hill, 1968.