Electronics Research Laboratory
University of California
Berkeley, California

THEORY OF A CLASS OF DISCRETE OPTIMAL CONTROL SYSTEMS

by

B. W. Jordan and E. Polak

The research herein reported is made possible through support received from the Departments of Army, Navy, and Air Force under grant AF-AFOSR-139-63; and Nasa grant NSG 354.

January 17, 1964

THEORY OF A CLASS OF DISCRETE OPTIMAL CONTROL SYSTEMS†

By B. W. Jordan and E. Polak

Department of Electrical Engineering
University of California
Berkeley, California

ABSTRACT

The problem considered in this paper is that of finding optimal controls for a class of fixed duration processes in systems described by nonlinear difference equations. The discrete versions of the adjoint system and the Hamiltonian are used in conjunction with the original techniques found in the proofs of the Pontryagin Maximum Principle to derive conditions necessary for a control to be optimal. These necessary conditions are shown to be related to the Pontryagin conditions for continuous systems in the following manner: the requirement of a global maximum of a Hamiltonian becomes a condition of a local maximum or of stationarity, while the transversality conditions remain identical.

1. INTRODUCTION

It has been known for some time that the Pontryagin Maximum Principle for optimal continuous time systems cannot be extended to discrete time systems, except for a few very special cases (Rozonoer 1959). This is due to the nature of the admissible control variations. In the continuous time case, in order to find necessary conditions for optimal controls, it is possible to use variations which range over all

† The research herein reported is made possible through support received from the Departments of Army, Navy, and Air Force under grant AF-AFOSR-139-63; and Nasa grant NSG 354.

admissible values of the control variable. Thus the Maximum Principle is a global result with respect to the control space. However, in the discrete time case, only amplitudinally small variations in the control may be taken and thus the most one can expect is a local condition.

Although no thoroughly exhaustive attempt has been made before to utilize the Pontryagin techniques in constructing necessary conditions for discrete time systems, some interesting results have been derived by L. I. Rozonoer (1959), S. S. L. Chang (1961), and S. Katz (1962). S. S. L. Chang considered a discretization technique for finding optimal solutions to continuous time problems by the use of a digital computer. In the process, he obtained a result somewhat analogous to condition (i) of Theorem 1 of this paper. Since his concern was with the continuous time problem, he did not consider the discrete time problem in any detail.

L. I. Rozonoer considered the extension of the Maximum Principle to systems described by linear difference equations. He obtained a modified form of the Maximum Principle which gave not only a necessary but also a sufficient condition for optimality of the control.

S. Katz showed that a further modification of the Maximum Principle did give necessary conditions for nonlinear discrete time problems with no terminal constraints on the state. However, as will be seen from the work in this paper, Katz's results were somewhat in error and only a weaker statement than his is possible.

This paper is devoted to exploring the exact nature of the general results for discrete systems which can be derived by means of techniques similar to the ones used in the derivation of the Maximum Principle. It is shown that in this case the Transversality Conditions are identical, but that the condition of a global maximum of a Hamiltonian must be changed to that of a local maximum or a stationary value.

2. FORMULATION OF THE OPTIMAL CONTROL PROBLEM

2.1 System Equations

Consider a system which satisfies the difference equations,

$$\underline{x}(k) = \underline{x}(k-1) + \underline{f}\,[\underline{x}(k-1),\ \underline{u}(k-1)] \quad k = 1, 2, \ldots \tag{1}$$

where

$$\underline{x} = \text{col}\,(x_1, \ldots, x_n) \in E^n \tag{2}$$

is the state,

$$\underline{u} = \text{col}\,(u_1, \ldots, u_r) \in U \subset E^r \tag{3}$$

is the control and

$$\underline{f} = \text{col}\,(f_1, \ldots, f_n) \in E^n \,. \tag{4}$$

U is assumed to be defined so that if $\underline{u}' \in U$ then there exists an $\epsilon > 0$ and a $\delta \underline{u}$ such that $\underline{u}' + \epsilon\, \delta \underline{u} \in U$. It is also assumed that if $\underline{u}' + \epsilon'\, \delta \underline{u}' \in U$, $\underline{u}' + \epsilon^2\, \delta \underline{u}^2 \in U$, then $\underline{u}' + \lambda\, \epsilon'\, \delta \underline{u}' + (1-\lambda)\epsilon^2\, \delta \underline{u}^2 \in U$, $0 \leq \lambda \leq 1$. It is also assumed that

$$f_i \in C' \text{ on } E^n \times U \quad i = 1, 2, \ldots, n\,. \tag{5}$$

Assume that $\underline{x}(0)$ is given. Then, if the control sequence

$$[\underline{u}(0, K-1)] \triangleq [\underline{u}(0), \ldots, \underline{u}(K-1)] \tag{6}$$

is given, the trajectory

$$[\underline{x}(0, K)] \triangleq [\underline{x}(0), \ldots, \underline{x}(K)] \tag{7}$$

can be computed.

2.2 Initial Conditions

Since the difference eqns. (1) do not depend on k, it will be assumed that the starting time is always at $k = 0$ and that the initial value of the state is always given as $\underline{x}(0)$.

2.3 Terminal Conditions

Time: It will always be assumed that the number of time steps over which the system is operated is fixed at K.

State: There are two possibilities which will be considered as terminal conditions on the state.

i) Assume that a closed, convex set $S \subset E^n$ is given. It is required that at the final time step, K, the state lie in S, [i.e., $\underline{x}(K) \in S$]. This definition allows S to be a point in E^n, a subset of E^n, or the whole of E^n. If S is of the second type (i.e., a closed, convex subset of E^n having more than one point), it will be assumed that it has no sharp edges. (i.e., at each point on its boundary surface a unique tangent plane exists).

ii) It is also possible that S might be an $(n - \ell)$-dimensional manifold described by the ℓ equations

$$S = \{\underline{x} \mid g_i[\underline{x}(K)] = 0 \qquad i = 1, 2, \ldots, \ell \geq 1\} . \tag{8}$$

In this case, S need not be convex, it will be assumed however, that the g_i have continuous partial derivatives with respect to the x_i and that $\text{grad}_{\underline{x}}\ g_i \neq 0$ for any $\underline{x} \in S$, $i = 1, \ldots, \ell$.

2.4 Cost Function

Assume that the cost of a transition from the state $\underline{x}(k - 1)$ to the state $\underline{x}(k)$ caused by the control $\underline{u}(k - 1)$ is given by

$f_0[\underline{x}(k - 1), \underline{u}(k - 1)]$. Let $x_o(k)$ be the cost of operating the system from time step zero to time step k. Then $x_o(k)$ is the solution of the difference equation

$$x_o(k) = x_o(k - 1) + f_0[\underline{x}(k - 1), \underline{u}(k - 1)] \tag{9}$$

with $x_o(0) = 0$. It is assumed that $f_0 \in C'$ on $E^n \times U$.

That this is a general cost function is demonstrated by S. Katz (1962).

2.5 Extended System Equations

Now, for convenience, the system equations will be extended to include the cost variable by defining the $(n + 1)$-dimensional vectors

$$\tilde{\underline{x}}(k) \triangleq \text{col}\,[x_0(k), \underline{x}(k)] \tag{10}$$

$$\tilde{\underline{f}} \triangleq \text{col}\,(f_0, \underline{f}) \; . \tag{11}$$

The system equations become

$$\tilde{\underline{x}}(k) = \tilde{\underline{x}}(k - 1) + \tilde{\underline{f}}[\underline{x}(k - 1), \underline{u}(k - 1)] \; . \tag{12}$$

Then given $\tilde{\underline{x}}(0)$ and $[\underline{u}(0, K - 1)]$, $[\tilde{\underline{x}}(0, K)]$ can be determined and $x_0(K)$ will represent the cost incurred in operating the system over the time sequence $[0, K]$.

2.6 Problem Statement

The fundamental optimal control problem for the systems under consideration will be denoted by P-1 and can be stated as

P-1: Given the positive integer K and the initial state $\tilde{\underline{x}}(0)$ for the system described by (12), find the sequence of controls $[\underline{u}(0, K-1)]$ so that $\underline{u}(i) \in U$, $i = 0, \ldots, K - 1$, $\underline{x}(K) \in S$ and so that $x_0(K)$ is minimized.

Definition: The sequence $[\underline{u}^*(0, K-1)]$, which minimizes $x_0(K)$, and satisfies the boundary conditions of P-1 will be called the optimal control for P-1 and the corresponding trajectory $[\tilde{\underline{x}}^*(0, K)]$ will be called the optimal trajectory.

3. NECESSARY CONDITIONS FOR AN OPTIMAL SOLUTION

3.1 The Adjoint System

Define the $(n + 1) \times (n + 1)$ matrix

$$F(k-1) \triangleq \left.\frac{\partial \tilde{\underline{f}}}{\partial \tilde{\underline{x}}}\right|_{\substack{\underline{x}(k-1) \\ \underline{u}(k-1)}} = \left[\frac{\partial f_i}{\partial x_j}\right]_{\substack{\underline{x}(k-1) \\ \underline{u}(k-1)}} , \tag{13}$$

and the $(n + 1) \times r$ matrix

$$B(k-1) \triangleq \left.\frac{\partial \tilde{\underline{f}}}{\partial \underline{u}}\right|_{\substack{\underline{x}(k-1) \\ \underline{u}(k-1)}} = \left[\frac{\partial f_i}{\partial u_j}\right]_{\substack{\underline{x}(k-1) \\ \underline{u}(k-1)}} . \tag{14}$$

Now the adjoint variables are defined as satisfying the following system of difference equations

$$\begin{aligned} \tilde{\underline{p}}(k-1) &= \tilde{\underline{p}}(k) + F^T(k-1)\, \tilde{\underline{p}}(k) \\ \tilde{\underline{p}}(k) &= \text{col}\,[p_0(k), p_1(k), \ldots, p_n(k)] . \end{aligned} \tag{15}$$

Since this system of equations is homogeneous, all that is needed to generate the trajectory of the adjoint system is the knowledge of $\tilde{\underline{p}}(K)$.

The determination of this vector will be a major consideration of this work.

Notice that $\dfrac{\partial f_i\,[\underline{x}(k-1),\ \underline{u}(k-1)]}{\partial x_0\,(k-1)} = 0, \quad i = 0, 1, 2, \ldots, n.$

Consequently, $p_0(k-1) = p_0(k)$, $k = 1, 2, \ldots, K$.
In other words p_0 is constant for all k.

3.2 The Hamiltonian

The Hamiltonian is defined as

$$H[\tilde{\underline{p}}(k),\ \underline{x}(k-1),\ \underline{u}(k-1)] \triangleq \langle \tilde{\underline{p}}(k),\ \tilde{\underline{f}}[\underline{x}(k-1),\ \underline{u}(k-1)] \rangle . \quad (16)$$

It is seen that the system eqns. (12) and the adjoint system eqns. (15) can be written in terms of the Hamiltonian:

$$\tilde{\underline{x}}(k) = \tilde{\underline{x}}(k-1) + \frac{\partial H(k)}{\partial \tilde{\underline{p}}(k)} ;$$

$$\tilde{\underline{p}}^T(k-1) = \tilde{\underline{p}}^T(k) + \frac{\partial H(k)}{\partial \tilde{\underline{x}}(k-1)} . \quad (17)$$

Now the conditions necessary for the control $[\underline{u}^*(0, K-1)]$ to be optimal can be stated.

3.3 Theorem I

If $[\underline{u}^*(0, K-1)]$ is an optimal control for P-1 and $[\tilde{\underline{x}}^*(0, K)]$ is the corresponding optimal trajectory, then there exists a function $\tilde{\underline{p}}^*(k)$, $k = 0, 1, \ldots, K$, satisfying (15) such that

i) $H[\tilde{\underline{p}}^*(k),\ \underline{x}^*(k-1),\ \underline{u}^*(k-1)] = \langle \tilde{\underline{p}}^*(k),\ \tilde{\underline{f}}[\underline{x}^*(k-1),\ \underline{u}^*(k-1)] \rangle$ is a local maximum or stationary with respect to $\underline{u}^*(k-1) \in U$ at each time step $1 \leq k \leq K$.

ii) $p_0^*(K) \leq 0$.

Discussion: This theorem is basic and holds regardless of the terminal conditions. In other words, no matter what form S takes, the conditions of this theorem must be fulfilled. The theorem gives a test for determining if a control $[\underline{u}(0, K - 1)]$ can be optimal. This test is performed by computing $[\underline{x}(0, K)]$ and $[\underline{\tilde{p}}(0, K)]$ and inserting these values into the Hamiltonian. The Hamiltonian is then checked to see if it is a local maximum or stationary for these values of $\underline{x}(k - 1)$, $\underline{\tilde{p}}(k)$ and $\underline{u}(k - 1)$ with respect to $\underline{u}(k - 1)$, for each k, $1 \leq k \leq K$. Also, in a large number of cases, there may be only one local maximum or stationary point for $H[\underline{\tilde{p}}(k), \underline{x}(k - 1), \underline{u}(k - 1)]$. Then, using Condition i of Theorem I, $\underline{u}(k - 1)$ can be found in terms of $\underline{x}(k - 1)$ and $\underline{\tilde{p}}(k)$. The control $\underline{u}(k - 1)$ can then be eliminated from the system eqns. (12) and the adjoint system eqns. (15). There are then (2n + 2) homogeneous equations for which (2n + 2) initial conditions must be found. There are (n + 1) boundary conditions, the $\underline{\tilde{x}}(0)$. Knowledge of the (n + 1) boundary conditions $\underline{\tilde{p}}(K)$ will give the solution. The vector $\underline{\tilde{p}}(K)$ will depend upon the form of the constraint set and will be discussed.

Theorem I will be proven by examining each type of terminal constraint in turn and establishing the Transversality Conditions for each case. These Transversality Conditions are developed in Theorems II - V. It will be shown that for each type of terminal constraint, the conditions of Theorem I are necessary.

The basic technique to be used will be to assume that the optimal control and trajectory are known. The control will then be perturbed so as to affect the trajectory only slightly. The necessary conditions which the optimal control must satisfy will then arise from the requirement that any admissible perturbed control which satisfies the terminal constraints must not result in a lower cost.

The first item, then, to be considered is the effect upon the trajectory of small perturbations in the control. Since the value of $\tilde{\underline{x}}(0)$ is given, no perturbations of its value need be considered. Only perturbations in each control vector $\underline{u}(i)$, $i = 0, \ldots, K - 1$, must be considered. It is at this point that the basic difference between this discrete time problem and the similar one for continuous time problems occurs. It is required that any perturbation must i) be such that the perturbed control is admissible, and ii) affect the trajectory only slightly. In the continuous time problem, the control is assumed to be measurable. Consequently, the perturbed control can vary from the original control by large amounts, provided the length of time, over which the perturbations are large, is small. This allows one to search out all of the control space at each time and to therefore require that the Hamiltonian be an absolute maximum at each instant of time.

In the discrete time problem, however, the only perturbations which have a small effect on the trajectory are small perturbations. Consequently, only local conditions can be obtained.

3.4 The Variational Equations

Consider, then, the effect of a perturbation on the control. Assume that the optimal control $[\underline{u}^*(0, K - 1)]$ and the optimal trajectory $[\tilde{\underline{x}}^*(0, K)]$ are known. Then perturb the control and require that the perturbed control be admissible.

Let $[\underline{u}(0, K - 1)] = [\underline{u}^*(0) + \epsilon\, \delta\, \underline{u}(0), \ldots, \underline{u}^*(K - 1) + \epsilon\, \delta\, \underline{u}(K - 1)]$ be the perturbed control where $\epsilon > 0$ is a small number independent of k. Then $[\tilde{\underline{x}}(0, K)] = [\tilde{\underline{x}}^*(0),\ \tilde{\underline{x}}^*(1) + \delta\, \tilde{\underline{x}}(1), \ldots,\ \tilde{\underline{x}}^*(K) + \delta\, \tilde{\underline{x}}(K)]$ will be the perturbed trajectory.

Then

$$\delta\tilde{\underline{x}}(k) = \tilde{\underline{x}}(k) - \tilde{\underline{x}}^*(k), \quad k = 0, 1, \ldots, K . \tag{18}$$

Since $f_i \in C'$ on $E^{n+1} \times U$, $i = 0, 1, \ldots, n+1$ and since $\delta\tilde{\underline{x}}(0) = 0$,

$$\delta\, \tilde{\underline{x}}(K) = \epsilon \sum_{i=0}^{K-1} D(i)\ \delta\underline{u}(i) + \tilde{\underline{o}}(\epsilon) \tag{19}$$

where

$$D(i) = (1 + F(K-1)) \ldots (1 + F(i+1))\, B(i)\,. \tag{20}$$

Let

$$\tilde{\underline{y}}(K) \triangleq \sum_{i=0}^{K-1} D(i)\ \delta\, \underline{u}(i) \in E_K^{n+1}\,, \tag{21}$$

where E_K^{n+1} is obtained from E^{n+1} by translating the origin of E^{n+1} to $\tilde{\underline{x}}^*(K)$. Define the set

$$\mathcal{K}_K \triangleq \left[\tilde{\underline{y}}(K) \;\middle|\; \tilde{\underline{y}}(K) = \sum_{i=0}^{K-1} D(i)\ \delta\, \underline{u}(i),\quad \begin{array}{l} \underline{u}(i) + \delta\underline{u}(i) \in U \\ i = 0, \ldots, K-1 \end{array} \right]. \tag{22}$$

This is a convex cone and will be called the "cone of attainability" due to a similar definition by Pontryagin.

It is obvious that $\mathcal{K}_K$ is convex since if

$$\tilde{\underline{y}}^1(K) = \sum_{i=0}^{K-1} D(i)\ \delta\, \underline{u}^1(i) \quad \in \mathcal{K}_K$$

and

$$\tilde{\underline{y}}^2(K) = \sum_{i=0}^{K-1} D(i)\ \delta\, \underline{u}^2(i) \quad \in \mathcal{K}_K$$

then

$$\lambda\tilde{\underline{y}}^1(K) + (1-\lambda)\, \tilde{\underline{y}}^2(K) = \sum_{i=0}^{K-1} D(i)[\,\lambda\delta\underline{u}^1(i) + (1-\lambda)\ \delta\,\underline{u}^2(i)] \in \mathcal{K}_K \quad 0 \le \lambda \le 1.$$

Notice also that $\tilde{\underline{x}}^*(K) \in \mathcal{K}_K$ and is the vertex of the cone since for $\delta \underline{u}(k) = 0, k = 0, 1, \ldots, K-1$, $\delta \tilde{\underline{x}}(K) \equiv 0$ which corresponds to $\tilde{\underline{x}}(K) = \tilde{\underline{x}}^*(K)$. Finally notice that $\tilde{\underline{y}}(k)$ satisfies the difference equation

$$\tilde{\underline{y}}(k) = \tilde{\underline{y}}(k-1) + F(k-1)\,\tilde{\underline{y}}(k-1) + B(k-1)\,\delta\,\underline{u}(k-1)\,. \tag{23}$$

3.5 Case 1: Right End Constrained to Lie on a Smooth Surface

Let the constraint set S be an $(n - \ell)$- dimensional manifold described by

$$S = [\underline{x} \mid g_j(\underline{x}) = 0, \; j = 1, 2, \ldots, \ell < n]. \tag{24}$$

Since each g_j has continuous first partial derivatives and since $\partial g_j / \partial \underline{x} \neq 0$, $i = 1, \ldots, \ell$ for $\underline{x} \in S$, there is an $(n - \ell)$-dimensional plane, T, tangent to S at each $\underline{x} \in S$ described by

$$T = [\underline{x}' \mid < \underline{x}' - \underline{x}, \; \frac{\partial g_j(\underline{x})}{\partial \underline{x}} > = 0, \; j = 1, 2, \ldots, \ell]. \tag{25}$$

Now construct the $(n + 1 - \ell)$-dimensional cylinder defined by

$$S^1 = [\tilde{\underline{x}} \quad \tilde{\underline{x}} = \text{col}\,(\underline{x}_0, x), \underline{x} \in S], \tag{26}$$

(see fig. 1). It has been assumed that $\tilde{\underline{x}}^*(K) \in S^1$. The cylinder S^1 will have an $(n + 1 - \ell)$-dimensional tangent plane at $\tilde{\underline{x}}^*(K)$ described by

$$T^1 = [\tilde{\underline{x}} \mid \tilde{\underline{x}} = \text{col}\,(x_0, x), \; \underline{x} \in T]\,. \tag{27}$$

Clearly, the projection of T^1 onto E^n is T, where T is assumed to be the tangent plane to S at $\underline{x}^*(K)$.

Construct the hyperplane C passing through $\tilde{\underline{x}}^*(K)$ perpendicular to the x_0 axis.

$$C = [\underline{\tilde{x}} \mid x_0 = x_0^*(K)] \ . \tag{28}$$

The hyperplane C will cut T^1 into two semi-infinite planes

$$T^{1+} = [\underline{\tilde{x}} \mid \underline{\tilde{x}} \in T^1, \ x_0 \geq x_0^*(K)] \tag{29}$$

$$T^{1-} = [\underline{\tilde{x}} \mid \underline{\tilde{x}} \in T^1, \ x_0 \leq x_0^*(K)] \tag{30}$$

with the common boundary

$$T^{1^o} = C \cap T^1 \ . \tag{31}$$

C will also cut the cylinder S^1 into two semi-infinite cylinders

$$S^{1+} = [\underline{\tilde{x}} \mid \underline{\tilde{x}} \in S^1, \ x_0 \geq x_0^*(K)] \tag{32}$$

$$S^{1-} = [\underline{\tilde{x}} \mid \underline{\tilde{x}} \in S^1, \ x_0 \leq x_0^*(K)] \tag{33}$$

with the common boundary

$$S^{1^o} = C \cap S^1 \ .$$

Let

$$\underline{\xi} = \text{col}\ (\xi_1, \xi_2, \dots \xi_n) \tag{34}$$

be an arbitrary n-dimensional vector lying in T. Let $\underline{p}(K)$ be a vector consisting of the last n-components of $\underline{\tilde{p}}(K)$.

$$\underline{\tilde{p}}(K) \triangleq \text{col}\ [p_1(K), \dots, p_n(K)] \ . \tag{35}$$

Then for this case, the following theorem holds:

Theorem II

Consider the problem P-1 when the constraint set S is an $(n - \ell)$-dimensional smooth manifold defined by (30). Then, necessary conditions that $[\underline{u}^*(0, K - 1)]$ be an optimal control are

i) The conditions stated in Theorem I.

ii) $< \underline{p}^*(K), \underline{\xi} > = 0$

where $\underline{p}(K)$ was defined in (35) and $\underline{\xi}$ is any vector lying in T tangent to S at $\underline{x}^*(K)$.

Proof

Since $[\underline{u}^*(0, K-1)]$ is an optimal control, it is necessary that any admissible perturbed control, whose corresponding trajectory satisfies the terminal conditions, not give a lower cost. For this requirement to be fulfilled, it is necessary that there exist a hyperplane separating $\mathcal{K}_K$ and T^{1-}. This is shown by establishing Lemma I.

Lemma I

Let $[\tilde{\underline{x}}(0), K)]$ be the trajectory corresponding to $[\underline{u}(0, K - 1)]$ and starting at $\tilde{\underline{x}}(0)$. Let G be a $\ell \leq n$-dimensional smooth manifold with an edge, G_e, in E^{n+1}, and let $\tilde{\underline{x}}(K) \in G_e$. Let L be the half-plane tangent to G at $\tilde{\underline{x}}(K)$.

If the cones $\mathcal{K}_K$ and L, having a common vertex at $\tilde{\underline{x}}(K)$, are not separated, then there exists a control $[\underline{u}'(0, K - 1)]$ with a corresponding trajectory $[\tilde{\underline{x}}'(0, K)]$, starting at $\tilde{\underline{x}}(0)$, such that $\tilde{\underline{x}}'(K) \in G$ but $\tilde{\underline{x}}'(K) \notin G_e$.

Proof of Lemma I

This lemma can be proven in a manner identical to the proof of Lemma 10 in L. S. Pontryagin (1962).

Let us apply this lemma to the proof of Theorem II. It follows from Lemma I that if the cones $\mathcal{K}_K$ and T^{1-}, having the common vertex $\tilde{\underline{x}}^*(K)$, are not separated, then there exists a control $[\underline{u}'(0, K-1)]$ with a corresponding trajectory $[\tilde{\underline{x}}'(0, K)]$ such that $\tilde{\underline{x}}'(K)$ lies in S^{1-} but not on the edge of S^{1-} and consequently will satisfy the constraints and have a lower cost.

Therefore, for $[\underline{u}^*(0, K-1)]$ and $[\tilde{\underline{x}}^*(0, K)]$ to be optimal, it is necessary that there exist a hyperplane, call it A, separating $\mathcal{K}_K$ and T^{1-}. Let the (n + 1)-dimensional vector $\tilde{\underline{a}} = \mathrm{col}(a_0, a_1, \ldots, a_{n+1})$ be the normal to A. Choose the direction of $\tilde{\underline{a}}$ so that

$$< \tilde{\underline{x}} - \tilde{\underline{x}}^*(K), \tilde{\underline{a}} > \;\leq 0 \text{ if } [\tilde{\underline{x}} - \tilde{\underline{x}}^*(K)] \in \mathcal{K}_K, \tag{36}$$

then

$$< \tilde{\underline{x}} - \tilde{\underline{x}}^*(K), \tilde{\underline{a}} > \;\geq 0 \text{ if } [\tilde{\underline{x}} - \tilde{\underline{x}}^*(K)] \in T^{1-}. \tag{37}$$

Clearly, the hyperplane A contains $\tilde{\underline{x}}^*(K)$ and T^{1o}. Let $\underline{\xi} = \mathrm{col}(\xi_1, \ldots, \xi_n)$ be any vector in T. Then $\tilde{\underline{\xi}} = \mathrm{col}(0, \underline{\xi})$ will be parallel to T^{1o}. Since $T^{1o} \subset A$, $< \tilde{\underline{a}}, \tilde{\underline{\xi}} > = 0$. But $\xi_0 = 0$. Therefore

$$< \tilde{\underline{a}}, \tilde{\underline{\xi}} > = \sum_{i=1}^{n} a_i \xi_i = 0 . \tag{38}$$

Since x_0 does not appear in the constraint relation (30), any point lying on the vector emanating from $\tilde{\underline{x}}^*(K)$ and pointing in the direction $\underline{\eta} = \mathrm{col}(-1, 0, 0, \ldots, 0)$ belongs to T^{1-}. Consequently, from the way $\tilde{\underline{a}}$ was chosen, it follows that

$$< \tilde{\underline{a}}, \tilde{\underline{\eta}} > \;\geq 0 .$$

But $< \tilde{\underline{a}}, \tilde{\underline{\eta}} > = -a_0 \geq 0$ and therefore

$$a_0 \leq 0 . \tag{39}$$

Since $\tilde{\underline{y}}(K) \in \mathcal{H}_K$, it follows from (36) that

$$< \tilde{\underline{a}}, \tilde{\underline{y}}(K) > \; \leq 0. \tag{40}$$

Now choose a special perturbation in the control. Let the control be perturbed at only the νth time step, $0 \leq \nu \leq K - 1$. Then

$$\underline{u}(k) = \underline{u}^*(k) \qquad k = 0, \ldots, \nu - 1, \; \nu + 1, \ldots, K - 1 \tag{41}$$

$$\underline{u}(\nu) = \underline{u}^*(\nu) + \epsilon \, \delta \, \underline{u}(\nu) \tag{42}$$

and

$$\delta \, \tilde{\underline{x}}(k) = 0, \; k \leq \nu \tag{43}$$

$$\delta \, \tilde{\underline{x}}(\nu + 1) = \tilde{\underline{f}}[\tilde{\underline{x}}^*(\nu), \; \underline{u}^*(\nu) + \epsilon \, \delta \, \underline{u}(\nu)] - \tilde{\underline{f}}[\tilde{\underline{x}}^*(\nu), \; \underline{u}^*(\nu)]. \tag{44}$$

Let

$$\tilde{\underline{y}}(\nu + 1) = \frac{\partial \, \tilde{\underline{f}}[\underline{x}^*(\nu), \; \underline{u}^*(\nu)]}{\partial \, \underline{u}} \; \delta \, \underline{u}(\nu). \tag{45}$$

Now consider the adjoint system (15)

$$\tilde{p}(k - 1) = \tilde{\underline{p}}^T(k) + F^T(k - 1) \, \tilde{\underline{p}}(k) \quad . \tag{46}$$

Since $\delta \, \underline{u}(k) = 0 \;\; k = \nu + 1, \ldots, K - 1$.

$$\begin{aligned} < \tilde{\underline{p}}(k), \tilde{\underline{y}}(k) > &= < \tilde{\underline{p}}(k), \; F(k - 1) \, \tilde{\underline{y}}(k - 1) + \tilde{\underline{y}}(k - 1) > \\ &= < F^T(k - 1) \, \tilde{\underline{p}}(k), \; \tilde{\underline{y}}(k - 1) > + < \tilde{\underline{p}}(k), \; \tilde{\underline{y}}(k - 1) > \\ &= < \tilde{\underline{p}}(k - 1) - \tilde{\underline{p}}(k), \; \tilde{\underline{y}}(k - 1) > + < \tilde{\underline{p}}(k), \; \tilde{\underline{y}}(k - 1) > \\ &= < \tilde{\underline{p}}(k - 1), \; \tilde{\underline{y}}(k - 1) > . \end{aligned} \tag{47}$$

for $k = \nu + 2, \ldots, K$.

Therefore

$$< \underline{\tilde{p}}(K), \ \underline{\tilde{y}}(K) > = < \underline{\tilde{p}}(\nu + 1), \ \underline{\tilde{y}} (\nu + 1) > . \tag{48}$$

Now let $\underline{\tilde{p}}(K) = \underline{\tilde{a}}$. Then since $\underline{\tilde{y}}(K) \in \mathcal{K}_K$, the necessary condition (40) becomes

$$< \underline{\tilde{p}} (K), \ \underline{\tilde{y}} (K) > \ \leq \ 0. \tag{49}$$

From (48), (49) becomes,

$$< \underline{\tilde{p}}(\nu + 1), \ \underline{\tilde{y}} (\nu + 1) > \ \leq \ 0 \tag{50}$$

or from (45)

$$< \underline{\tilde{p}}(\nu + 1), \ \frac{\partial \ \underline{\tilde{f}}[\underline{x}^*(\nu), \ \underline{u}^*(\nu)]}{\partial \ \underline{u}} \ \delta \, \underline{u}(\nu) > \ \leq 0 \ . \tag{51}$$

Equivalently, it is necessary that $H(\nu + 1)$ be a local maximum or stationary with respect to $\underline{u}^*(\nu)$. Since the choice of ν was arbitrary, a necessary condition that $[\underline{u}^*(0, K - 1)]$ be an optimal control is that $H(\nu)$ be a local maximum or stationary for $1 \leq \nu \leq K$.

This shows that for these terminal conditions, condition (i) of Theorem I is necessary. Since $p_0(K) = a_0 \leq 0$ condition (ii) of Theorem I is necessary. Finally, from (50)

$$< p(K), \ \underline{\xi} > \ = \ \sum_{i=1}^{n} a_i \, \xi_i = 0 \ . \tag{52}$$

This completes the proof of Theorem II.

3.6 Case 2: Right End Constrained to Lie at a Point

Next consider the problem when S is a point in E^n. For this problem, the following theorem holds.

Theorem III

Consider the problem P-1 when the constraint set S is a point in E^n. Then, necessary conditions for $[\underline{u}^*(0, K-1)]$ to be an optimal control are

i) The conditions stated in Theorem I.

Proof

Since S is a point, S^1 is a line perpendicular to E^n (i.e., parallel to the x_0 axis) and passing through S. S^{1-} is the semi-infinite line consisting of those points in S^1 below or in C. Clearly, $T^{1-} = S^{1-}$. It follow from Lemma I that $\mathcal{K}_K$ and S^{1-} must be separated by the hyperplane A, with its normal $\tilde{\underline{a}}$. As in Theorem II, one may choose $a_0 \leq 0$. However, since T^{1o} is a point, no transversality conditions need be imposed on $\tilde{\underline{a}}$. Proceeding as in Theorem II, it is found by letting $\tilde{\underline{p}}(K) = \tilde{\underline{a}}$ that conditions i) and ii) of Theorem I are necessary for $[\underline{u}^*(0, K-1)]$ to be an optimal control. Q.E.D.

Remark

Note that no conditions are imposed on the values of the last n-components of $\tilde{\underline{p}}(K)$.

3.7 Case 3: Free Right End

Consider the problem where S is the whole space, E^n. In other words, S can lie anywhere in E^n. For this set of terminal conditions the following theorem holds.

Theorem IV

Consider the problem P-1 when the constraint set $S = E^n$. Then, necessary conditions for $[\underline{u}^*(0, K-1)]$ to be an optimal control are,

i) the conditions stated in Theorem I

ii) $p_i^*(K) = 0 \quad i = 1, 2, \ldots, n.$

Proof

Clearly, S^{1-} will be the closed half-space in E^{n+1} consisting of all points lying below or in C. Let T^{1o} be an arbitrary line in C passing through $\tilde{\underline{x}}^*(K)$ and let T^{1-} be the semi-infinite hyperplane consisting of those points lying in T^{1o} or directly below it. It follows from Lemma I that T^{1-} and $\mathcal{K}_K$ must be separated by some hyperplane A with its normal $\tilde{\underline{a}}$. It follows from Theorem II that $\tilde{\underline{a}}$ must be perpendicular to T^{1o}. But T^{1o} has arbitrary direction in C, consequently, $\tilde{\underline{a}}$ must be orthogonal to any vector in C, i. e., to C itself. It can only have the value

$$\tilde{\underline{a}} = \mathrm{col}\,(-1, 0, 0, \dots, 0) \tag{53}$$

and therefore A coincides with C. Proceeding as in Theorem II and letting $\tilde{\underline{p}}(K) = \tilde{\underline{a}}$, the conditions of Theorem I are shown to be necessary and in addition the Transversality Conditions ii) of Theorem IV are shown to be necessary.

3.8 Case 4: Right End Constrained to Lie in an n-Dimensional Subset of E^n

The problems where S is a point in E^n, a manifold of dimension $n - \ell < n$ and the whole of E^n have been considered. The only problem left is that where S is an n-dimensional proper subset of E^n. Assume that S is closed and convex.

S^{1-} will be a semi-infinite cylinder consisting of all those points in S^1 which lie in or below C. Two possibilities can occur. (see fig. 5).

Case A: $\tilde{\underline{x}}^*(K)$ may lie on the surface of S^1. It will also lie in C. Let T^1 be the tangent plane to S^1 at $\tilde{\underline{x}}^*(K)$. T^1 will be cut in half by C. Let T^{1+} and T^{1-} be the upper and lower halves as before with the common boundary $T^{1o} = C \cap T^1$. Let $\tilde{\underline{h}}$ be the normal to T^1 at $\tilde{\underline{x}}^*(K)$ which points away from S^1

$$\tilde{\underline{h}} = \mathrm{col}\,(0, h_1, \dots, h_n)\,. \tag{54}$$

Let $\tilde{\underline{c}}$ be the upward pointing normal to C,

$$\tilde{\underline{c}} = \mathrm{col}\,(1, 0, \ldots, 0). \tag{55}$$

Case B: $\tilde{\underline{x}}^*(K)$ may lie in the interior of S^1. It will still lie in C. In this case there will be no concept of a tangent plane and $\tilde{\underline{h}}$ will be defined as the zero vector

$$\tilde{\underline{h}} = \mathrm{col}\,(0, 0, \ldots, 0)\,. \tag{56}$$

The vector $\tilde{\underline{c}}$ will be as in case A .

Then in either case the following theorem holds.

Theorem 5

Consider the problem P-1 when the constraint set S is a closed, convex, n-dimensional subset of E^n. Then, necessary conditions for $[\underline{u}^*(0, K - 1)]$ to be an optimal control are,

i) the conditions stated in Theorem I

ii) $\tilde{\underline{p}}^*(K) = \lambda\tilde{\underline{h}} + \mu\tilde{\underline{c}}$

where λ, μ are nonpositive constants.

Proof

Consider Case B. This situation is identical with that of the free right end and Theorem IV holds. Since $\tilde{\underline{h}} = 0$ in this case

$$\tilde{\underline{p}}^*(K) = \mu\tilde{\underline{c}} = \mathrm{col}\,(\mu, 0, 0, \ldots, 0)\,. \tag{57}$$

But from Theorem IV it is seen that $\mu \leq 0$. Therefore Theorem V is true for Case B.

Consider Case A. Let L be a half-hyperplane having T^{1o} as an edge and intersecting S^{1-}. Then it follows from Lemma 1 that $\mathcal{K}_K$ must be separated from L. Therefore, $\mathcal{K}_K$ must be separated from the closed quarter-space, Q, bounded from above by C and from the side by T^{1-}. Let A be the hyperplane which separates $\mathcal{K}_K$ from Q and let $\underset{\sim}{a}$ be the normal to A which points into Q. Then $-\underset{\sim}{a}$ will lie in the space spanned by $\underset{\sim}{h}$ and $\underset{\sim}{c}$ and will lie between them. Therefore $\underset{\sim}{a}$ can be written as a negative linear combination of $\underset{\sim}{h}$ and $\underset{\sim}{c}$.

$$\underset{\sim}{a} = \lambda \underset{\sim}{h} + \mu \underset{\sim}{c} \qquad \lambda, \mu \leq 0 . \tag{58}$$

Then, proceeding as in Theorem II and letting $\underset{\sim}{p}^*(K) = \underset{\sim}{a}$, the conditions of Theorem I are shown to be necessary and in addition the Transversality Conditions (ii) of Theorem V are shown to be necessary.

The conditions of Theorem I have been shown to be necessary for each terminal constraint under consideration. Therefore, the proof of Theorem I is completed.

4. CONCLUSIONS

This paper demonstrates the extent to which the techniques used in the construction of the Maximum Principle can be used to obtain a related necessary optimality condition for discrete time problems but which is not necessarily a maximum condition.

It is interesting to see that Rozonoer's assertion, that the "extension of the Maximum Principle to discrete systems is possible, generally speaking, only in the linear case," is correct and that the corresponding necessary conditions for the nonlinear case are, in fact, weaker than those given by Pontryagin, (i.e., the Hamiltonian is required to be only a local maximum or stationary rather than an absolute maximum). In many systems, however, the Hamiltonian will have only one local maximum or stationary point and for these problems, the results derived here are as useful as those derived by Pontryagin.

Also noteworthy is the fact that Katz's conclusion that the Hamiltonian must be a local maximum, is not quite complete due to his neglect of second order terms. Rather, as shown in this paper, it is only necessary that the Hamiltonian be a local maximum or stationary.

There is one specific problem in which all the assumptions of this paper need not be met. In the free right-end case (see Section 3.7 of this paper) which was considered by Katz, the control constraint set U need not be restricted as in the other cases. For this case it is only necessary that for any $\underline{u}' \in U$, there exist some ϵ and some $\delta \underline{u}$ such that $\underline{u}' + \epsilon \, \delta \underline{u} \in U$; i.e., it is not necessary that the set $[\delta \underline{u} \mid \underline{u}' + \epsilon \, \delta \underline{u} \in U \text{ for some } \epsilon]$ be convex. The reason for this is that S^{1-} for this case is a whole half space in E^{n+1} and the separating hyperplane is uniquely defined. Consequently, the proofs no longer depend on the convexity of $\mathcal{K}_K$.

For all cases considered here, it is possible for the control constraint set to change with the time step k [i.e., U = U(k)], provided that each of the U(k), k = 0, 1, ..., K - 1, is an admissible constraint set.

It is hoped that the results presented in this paper will help to complete the theory of optimal control systems.

REFERENCES

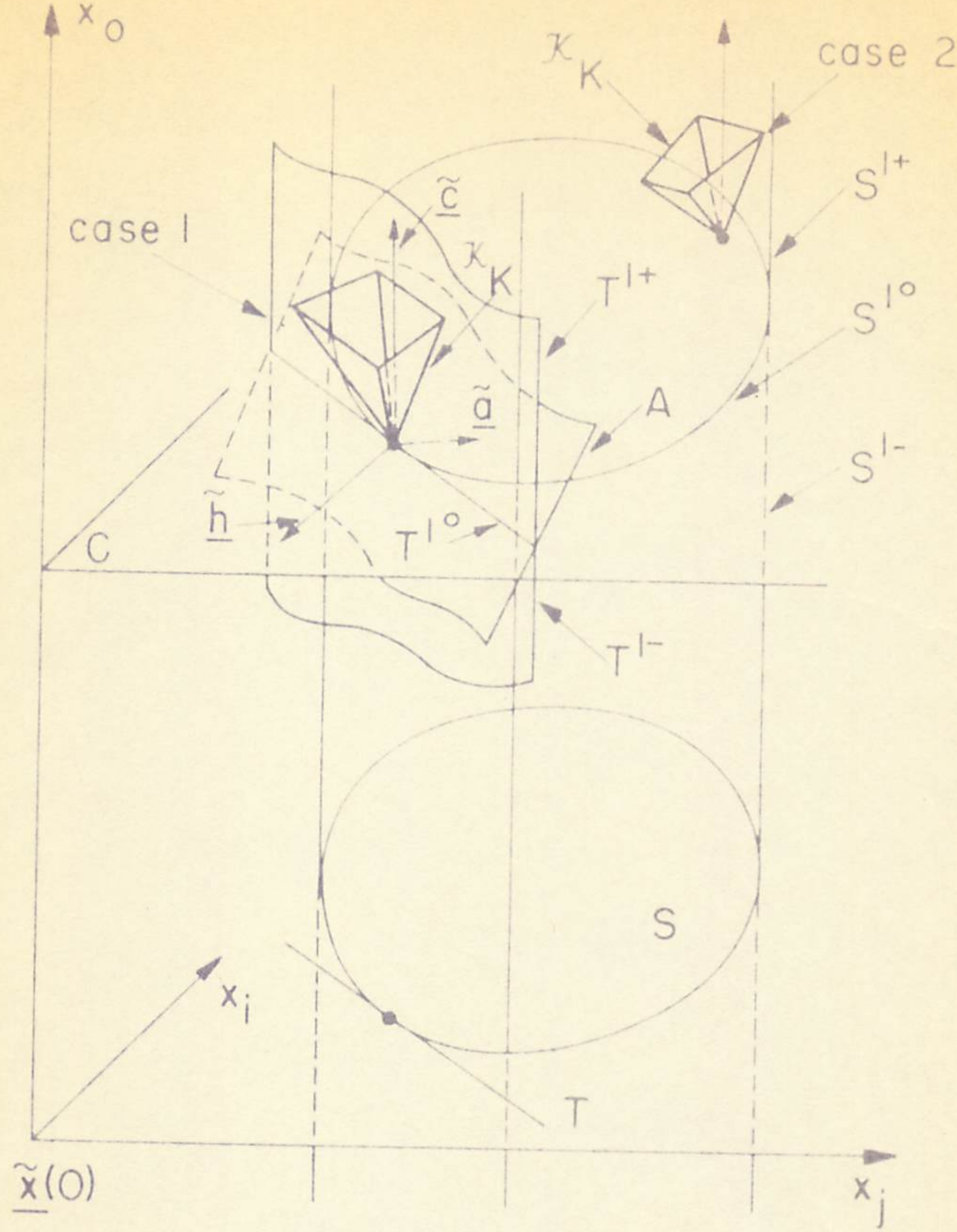
1. PONTRYAGIN, L. S., et al., 1962, The Mathematical Theory of Optimal Processes (New York: John Wiley & Sons).

2. ROZONOER, L. I., 1959, Automatika i Telemekh, **20**, I, II, & III, 1320-1334, 1441-1458, 1561-1578. [English translation in Automation and Remote Control, **20**, I, II, & III, 1288-1302, 1405-1421, 1517-1532 (1959).]

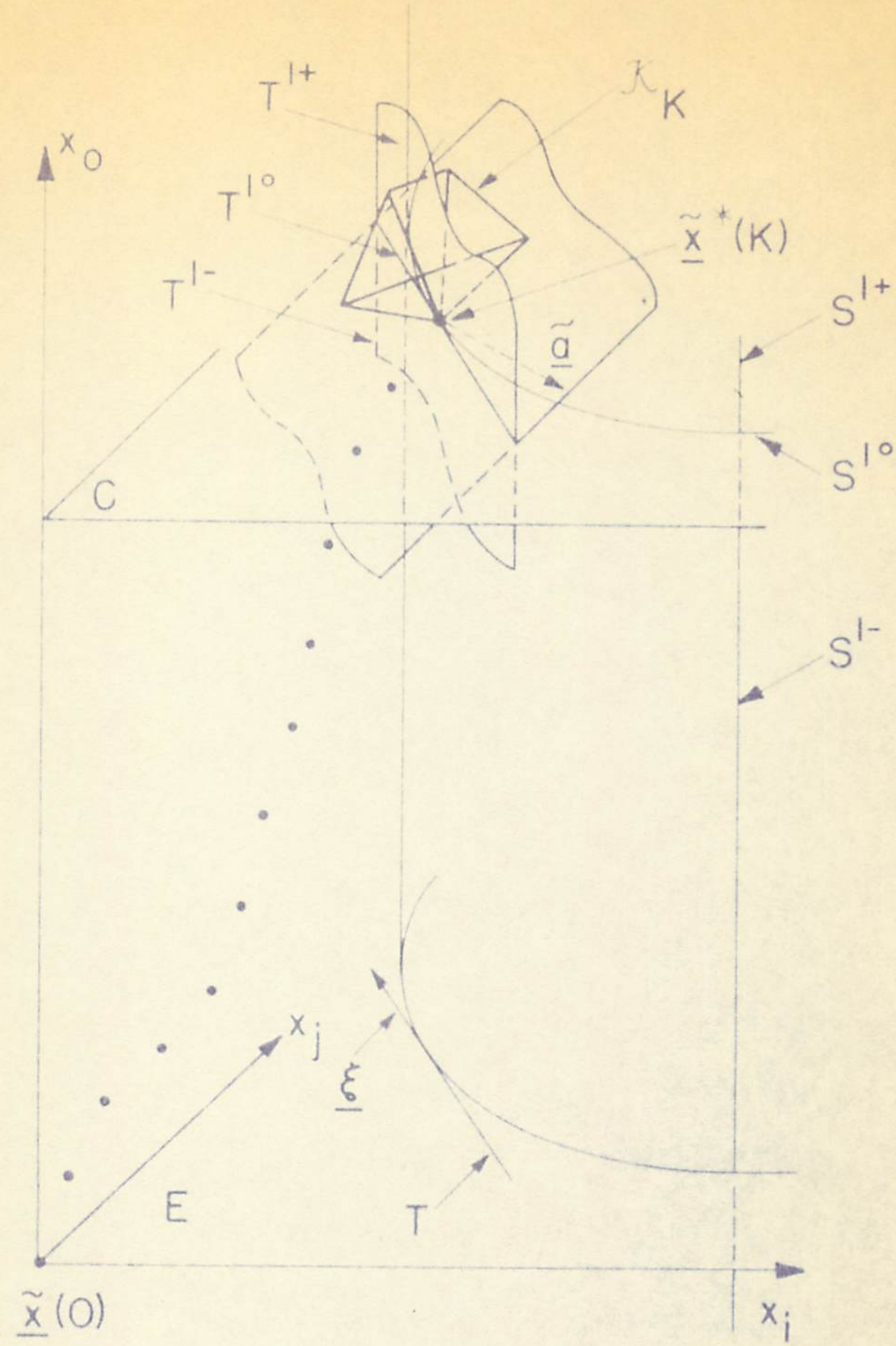
3. KATZ, S., 1962, J. Electron. Contr., **3**, No. 2, 179.

4. CHANG, S. S. L., 1961, IRE Int. Conv. Record, **9**, Pt. 4.

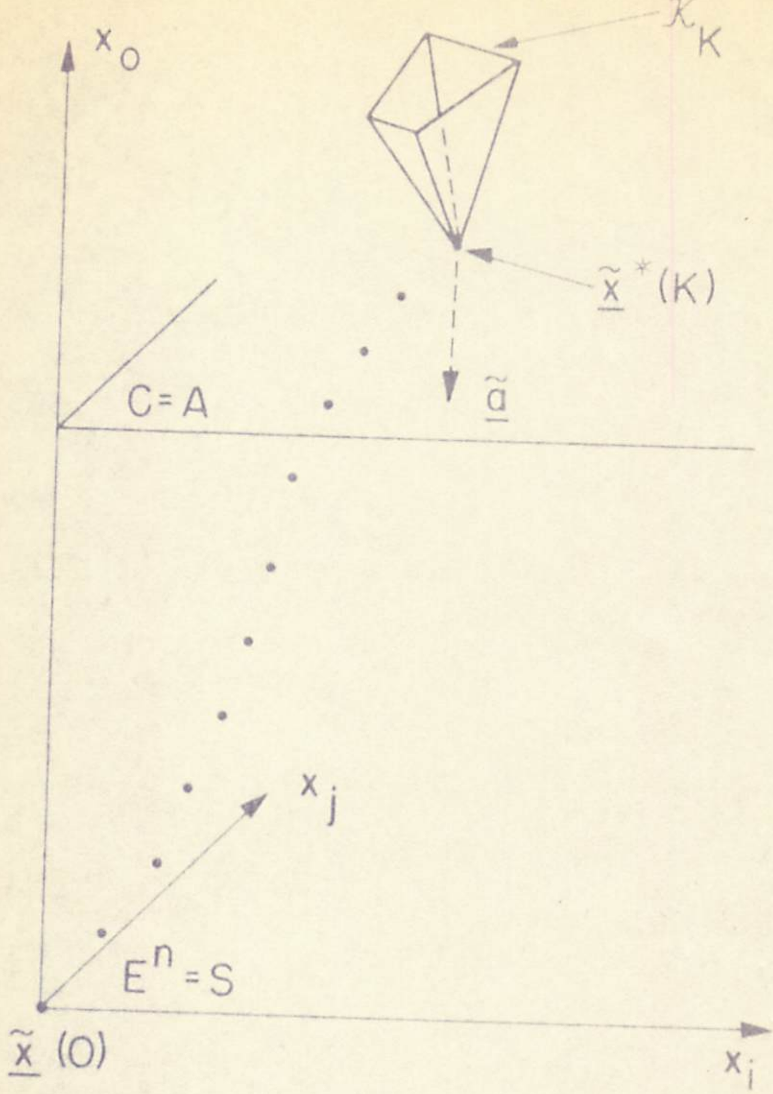
5. CHANG, S. S. L., 1960, Proc. IRE, **48**, No. 12, 2030.



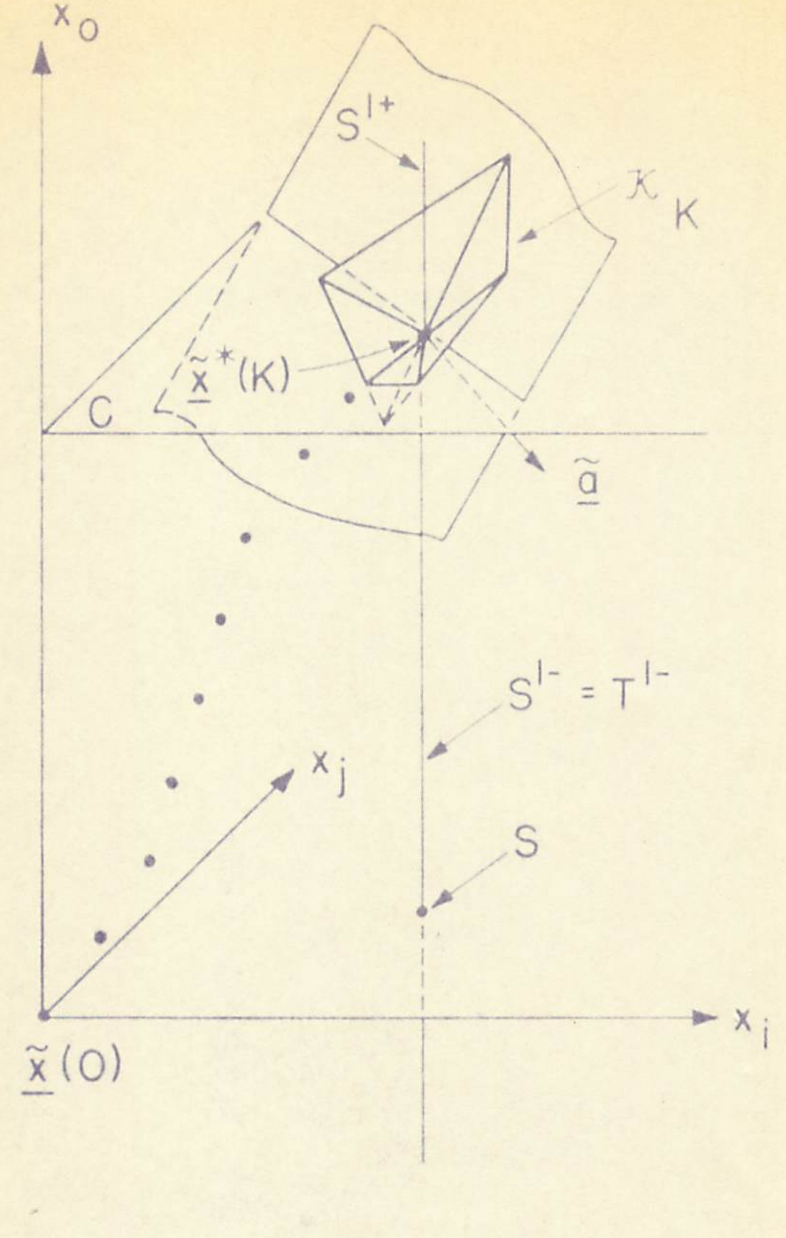
x_0
$\mathcal{K}_K$
case 2
S^{1+}
case 1
$\underline{\tilde{c}}$
$\mathcal{K}_K$
T^{1+}
S^{1o}
A
$\underline{\tilde{a}}$
S^{1-}
$\underline{\tilde{h}}$
C
T^{1o}
T^{1-}
S
x_i
T
$\underline{\tilde{x}}(0)$
x_j



T^{l+}
$\mathcal{K}_K$
x_o
$T^{l\circ}$
$\underline{\tilde{x}}^*(K)$
T^{l-}
S^{l+}
$\underline{\tilde{a}}$
$S^{l\circ}$
C
S^{l-}
x_j
$\underline{\xi}$
E
T
$\underline{\tilde{x}}(0)$
x_i



x_o
$\mathcal{K}_K$
$\underset{\sim}{\tilde{x}}^*(K)$
C=A
$\underset{\sim}{\tilde{a}}$
x_j
$E^n = S$
$\underset{\sim}{\tilde{x}}(0)$
x_i



x_0
S^{1+}
$\mathcal{K}_K$
$\underline{\tilde{x}}^*(K)$
C
$\underline{\tilde{a}}$
$S^{1-} = T^{1-}$
x_j
S
x_i
$\underline{\tilde{x}}(0)$