CONSTRAINED MINIMIZATION PROBLEMS IN FINITE DIMENSIONAL SPACES

by

M. Canon, C. Cullum and E. Polak

ERL Technical Memorandum M-149

1 December 1965

ELECTRONICS RESEARCH LABORATORY

College of Engineering

University of California, Berkeley

94720

The research reported herein was supported by the National Science Foundation under Grant GK-569 and by the National Aeronautic and Space Administration under Grant NsG-354(supp 2).

INTRODUCTION

The entire approach to constrained minimization problems in finite dimensional spaces, as found in the field of optimal control, is substantially different from the approach to these problems found in mathematical programming. Furthermore, within each of these fields, one finds a diversity of methods and points of view. The purpose of this paper is to exhibit a unified approach to constrained minimization problems in finite dimensional spaces and to show that most of the known necessary conditions for optimality are straightforward consequences of a fairly simple, but all-encompassing theorem.

The first part of this paper is devoted to formulating the Basic Problem, i.e., the form into which most of the known, finite dimensional, constrained minimization problems can be transcribed. A necessary condition for the optimality of a solution to this Basic Problem is then derived by a geometric method, first used by McShane [1] in the Calculus of Variations and subsequently greatly popularized by Pontryagin et all [2] in their derivation of the Maximum Principle. The necessary condition for the Basic Problem is stated as an inequality which must hold for all the elements in a cone which is a suitable linearization of the constraint set. The wide range of applicability of this theorem is substantially due to the fact that one has a great deal of freedom in choosing this linearization cone.

The second part of the paper is devoted to transcribing a wide variety of minimization problems into the form of the Basic Problem, to re-deriving many classical necessary conditions, and to obtaining several new ones. In particular, it is shown that classical Lagrange multiplier theory, the results of Fritz John [3], Kuhn and Tucker [4], and Mangasarian and Fromovitz [5], in nonlinear programming theory, and the results of Jordan and Polak [6], Halkin [7], and Holtzman [8], in discrete optimal control theory, can all be obtained from the necessary condition for the Basic Problem. In addition, several new results are obtained for bounded state space, discrete optimal control problems. Presently known necessary conditions for certain bounded state space problems, such as those obtained by Rosen [9], can be seen to be special cases of the more general results presented in this paper.

It is the authors' hope that the unified approach to constrained minimization problems in E^n, presented in this paper, will facilitate the mastery of the subject and will lead to a deeper and more fruitful understanding of minimization problems in general.

I. THE BASIC PROBLEM

Statement of the Basic Problem

Let $f: E^n \to E^1$ and $r: E^n \to E^m$ be continuously differentiable functions, and let $\Omega \subset E^n$ be a subset of E^n. The Basic Problem can be stated as follows:

Find a vector $\hat{z} \in E^n$ such that

(i) $\hat{z} \in \Omega$, $r(\hat{z}) = 0$,

(ii) for all $z \in \Omega$ with $r(z) = 0$, $f(\hat{z}) \leq f(z)$.

We shall call a vector $\hat{z}$ satisfying (i) and (ii) an optimal solution to the Basic Problem.

Necessary Condition for Optimality

The necessary condition to be derived will be stated in the form of an inequality which is valid for all $\delta z = (\delta z_1, \delta z_2, \ldots, \delta z_n)$ in a convex cone "approximation" or "linearization" of the set Ω. We shall make use of two kinds of "linearizations" of the set Ω at a point z. The first one will be defined now; the second one will be defined after the proof of Theorem 1, to obtain an extension.

Definition. A convex cone* $C(z, \Omega) \subset E^n$ will be called a linearization of the first kind of the constraint set Ω at z if for any finite collection $\{\delta z^1, \delta z^2, \ldots, \delta z^k\}$ of linearly independent vectors in $C(z, \Omega)$ there exists an $\epsilon > 0$, possibly depending on z, δz^1, δz^2, $\ldots$, δz^k, such that $co\{z, z + \epsilon\, \delta z^1, \ldots, z + \epsilon\, \delta z^k\}^{\dagger} \subset \Omega$.

If the cone $C(z, \Omega)$ is a linearization of the first kind, then for every $\delta z \in C(z, \Omega)$ there exists an $\epsilon_1 > 0$ such that $z + \epsilon\, \delta z \in \Omega$ for all ϵ such that $0 \leq \epsilon \leq \epsilon_1$. The largest cone having this property is given a special name.

Definition. The radial cone to the set Ω at a point $z \in \Omega$ will be denoted by $RC(z, \Omega)$ and is defined by

$$RC(z, \Omega) = \{\delta z : z + \epsilon\, \delta z \in \Omega \text{ for all } \epsilon \text{ such that } 0 \leq \epsilon \leq \epsilon_1(z, \delta z) > 0\}$$

---

* A set C is a cone with vertex x_0 if for every $x \in C$, $x \neq x_0$, $x_0 + \lambda(x - x_0) \in C$ for all $\lambda > 0$. Since the vertex x_0 of the cone C will normally be obvious, we shall omit mentioning it.

† $co\{z, z + \epsilon\, \delta z^1, \ldots, z + \epsilon\, \delta z^k\}$ is the convex hull of z, $z + \epsilon\, \delta z^1$, $\ldots$, $z + \epsilon\, \delta z^k$, i.e., the set of all points, y, of the form $y = \mu_0 z + \mu_1(z + \epsilon\, \delta z^1) + \ldots + \mu_k(z + \epsilon\, \delta z^k)$, where $\sum_{i=0}^{k} \mu_i = 1$, μ_i 0 for all i.

Whenever the radial cone $RC(\hat{z}, \Omega)$ is a linearization of the first kind, it contains all the other linearizations of the first kind of the set Ω at $\hat{z}$. Consequently, in the various theorems to follow, the radial cone $RC(\hat{z}, \Omega)$ should always be used if possible, since this will result in stronger necessary conditions.

Next, we define the $C^{(1)}$ map $F: E^n \to E^{m+1}$

$$F(z) = (f(z), r(z)) .$$

We shall number the components of E^{m+1} from 0 to m, i.e., $y \in E^{m+1}$ is given by $y = (y^o, y^1, \ldots, y^m)$. The Jacobian matrix of the map $F(z)$, $\left(\frac{\partial F^i(z)}{\partial z^j}\right)$, will be denoted by $\frac{\partial F(z)}{\partial z}$.

For the Basic Problem stated above, the following theorem gives a necessary condition for optimality.

<u>Theorem 1.</u> If $\hat{z}$ is an optimal solution to the Basic Problem, and $C(\hat{z}, \Omega)$ is a linearization of the first kind of Ω at $\hat{z}$, then there exists a nonzero vector $\psi = (\psi^o, \psi^1, \ldots, \psi^m) \in E^{m+1}$, with $\psi^o \leq 0$, such that for all $\delta z \in \overline{C(\hat{z}, \Omega)}$ (the closure of $C(\hat{z}, \Omega)$ in E^n)

$$\left\langle \psi, \frac{\partial F(\hat{z})}{\partial z} \delta z \right\rangle \leq 0. \tag{1}$$

<u>Proof.</u> Let $K(\hat{z}) \subset E^{m+1}$ be the cone defined by

$$K(\hat{z}) = \frac{\partial F(\hat{z})}{\partial z} C(\hat{z}, \Omega) \tag{2}$$

$K(\hat{z})$ is convex because $C(\hat{z}, \Omega)$ is convex and $\frac{\partial F(\hat{z})}{\partial z}$ is a linear map. Let $\hat{y} = F(\hat{z})$. We shall now show that the cone $\{\hat{y}\} + K(\hat{z})$ must be separated from the ray

(3) $$R = \{y : y = \hat{y} + \beta(-1, 0, \ldots, 0) \ , \ \beta \geq 0\} \ ,$$

i.e. that there must exist a nonzero vector $\psi \in E^{m+1}$ such that

(4) (i) $\langle \psi, y - \hat{y} \rangle \leq 0$ for every $y \in \{\hat{y}\} + K(\hat{z})$

(ii) $\langle \psi, y - \hat{y} \rangle \geq 0$ for every $y \in R$.

Suppose that the cone $\{\hat{y}\} + K(\hat{z})$ and the ray R are not separated. Then the cone $K(\hat{z})$ must be of dimension $m+1$ and R must be an interior ray of $\{\hat{y}\} + K(\hat{z})$ (i.e., all points of R except $\hat{y}$ are interior points of $\{\hat{y}\} + K(\hat{z})$).

Let us now construct in the cone $\{\hat{y}\} + K(\hat{z})$ a simplex Σ with vertices $\hat{y}$, $\hat{y} + \delta y^1$, $y + \delta y^2, \ldots, y + \delta y^{m+1}$ such that

(i) there exists a point y on R (which we shall write as $y = \hat{y} + \delta y^o$, $\delta y^o = \gamma(-1, 0, \ldots, 0)$ with $\gamma > 0$), different from $\hat{y}$, which lies in the inteior of Σ,

(ii) there exists a set of vectors $\delta z^i \in C(\hat{z}, \Omega)$ satisfying

$$G_\alpha(x) = \hat{y} + \frac{\partial F(\hat{z})}{\partial z} ZY^{-1}(\alpha\,\delta y^o + x) - (\hat{y} + \alpha\,\delta y^o)$$

(8)
$$+\ o(ZY^{-1}(\alpha\,\delta y^o + x))$$

where $o(\cdot)$ is a continuous function such that $\lim_{\|y\| \to 0} \frac{\|o(y)\|}{\|y\|} = 0$. By definition, $\frac{\partial F(\hat{z})}{\partial z} Z = Y$, and hence (8) simplifies to

(9)
$$G_\alpha(x) = x + o(ZY^{-1}(\alpha\,\delta y^o + x)).$$

Now, for $x \in \partial(S_\alpha - \{\hat{y} + \alpha\ y^o\})$ (the boundary of the sphere) $\|x\| = \alpha r$ and we may write $x = \alpha\rho_1$, where $\|\rho_1\| = r$. Hence, for $x \in \partial(S_\alpha - \{\hat{y} + \alpha\,\delta y^o\})$

(10)
$$G_\alpha(\alpha\,\rho_1) = \alpha\,\rho_1 + o(\alpha\,ZY^{-1}(\delta y^o + \rho_1))$$

Consequently, there exists an α^*, $0 < \alpha^* \leq 1$, such that for all $\rho_1 \in E^{m+1}$, with $\|\rho_1\| = r$,

(11)
$$\|\ o(\alpha^*\ SY^{-1}(\delta y^o + \rho_1))\ \| < \alpha^*\ r$$

We now conclude from Brouwner's Fixed Point Theorem (see Appendix 1) that there exists a $\tilde{x} \in S_{\alpha^*} - \{\hat{y} + \alpha^*\ \delta y^o\}$ such that

(12)
$$G_{\alpha^*}(\tilde{x}) = 0,$$

i.e.,

$$F(\hat{z} + ZY^{-1}(\alpha^* \ \delta y^o + \tilde{x})) = \hat{y} + \alpha^* \ \delta y^o \tag{13}$$

Now $\hat{y} + \alpha^* \ \delta y^o = \text{col}(f(\hat{z}) - \alpha^* \gamma, 0, 0, \ldots, 0)$, where $\gamma > 0$. Thus, expanding (13),

$$r(\hat{z} + ZY^{-1}(\alpha^* \ \delta y^o + \tilde{x})) = 0 \tag{14}$$

and

$$f(\hat{z} + ZY^{-1}(\alpha^* \ \delta y^o + \tilde{x})) = f(\hat{z}) - \alpha^* \ \gamma < f(\hat{z}) \tag{15}$$

Furthermore, because of (6) and the fact that for any δy in the simplex $\Sigma - \{\hat{y}\}$, the vector $z = \hat{z} + ZY^{-1} \ \delta y$ belongs to $\text{co}\{\hat{z}, \hat{z} + \delta z^1, \ldots, \hat{z} + \delta z^{m+1}\}$,

$$\hat{z} + ZY^{-1}(\alpha^* \ \delta y^o + \tilde{x}) \in \Omega \tag{16}$$

Hence $\hat{z}$ is not optimal, which is a contradiction. We therefore conclude that the cone $\{\hat{y}\} + K(\hat{z})$ and the ray R must be separated, i.e., there must exist a nonzero vector $\psi \in E^{m+1}$ such that

$$\text{(i)} \quad \langle \psi, (y - \hat{y}) \rangle \leq 0 \quad \text{for every } y \in \{\hat{y}\} + K(\hat{z}) \tag{17}$$

and

(18) (ii) $\langle \psi, (y - \hat{y}) \rangle \geq 0$ for every $y \in R$

Substituting (2) in (17), we have

(19) $$\left\langle \psi, \frac{\partial F(\hat{z})}{\partial z} \delta z \right\rangle \leq 0 \text{ for every } \delta z \in C(\hat{z}, \Omega)$$

Clearly, (19) must also hold for every $\delta z \in \overline{C(\hat{z}, \Omega)}$.

Substituting for y from (3) in (18), we have

(20) $$\langle \psi, (-1, 0, \ldots, 0) \rangle = -\psi^o \geq 0 .$$

This completes the proof.

It has been pointed out by Neustadt [10] that Theorem 4 remains valid under the relaxed assumption that $C(\hat{z}, \Omega)$ is a linearization of the second kind of Ω at $\hat{z}$, defined as follows.

<u>Definition.</u> A convex cone $C(z, \Omega) \subset E^n$ will be called a <u>linearization of the second kind</u> of the constraint set Ω at z, if, for any finite collection $\{\delta z^1, \delta z^2, \ldots, \delta z^k\}$ of linearly independent vectors in $C(z, \Omega)$, there exists an $\epsilon > 0$, possibly depending on z, $\delta z^1, \ldots, \delta z^k$, and a continuous map ζ from $\text{co}\{z, z + \epsilon\, \delta z^1, \ldots, z + \epsilon\, \delta z^k\}$ into Ω, such that $\zeta(z + \delta z) = z + \delta z + o(\delta z)$, where $\lim_{\|\delta z\| \to 0} \frac{\|o(\delta z)\|}{\|\delta z\|} = 0$.

Remark. We observe that if $C(z, \Omega)$ is a linearization of the first kind of Ω at z, then it is also a linearization of the second kind of Ω at z, with the map ζ being the identity. Thus, unless we have specific cause to indicate whether a cone $C(z, \Omega)$ is a linearization of the first or second kind, we shall refer to it simply as a linearization of Ω at z. We now restate Theorem 1 in this form.

Theorem 1'. If $\hat{z}$ is an optimal solution to the basic problem and $C(\hat{z}, \Omega)$ is a linearization of Ω at $\hat{z}$, then there exists a nonzero vector $\psi = (\psi^0, \psi^1, \ldots, \psi^m) \in E^{m+1}$ with $\psi^0 \leq 0$, such that for all $\delta z \in \overline{C(\hat{z}, \Omega)}$, (the closure of $C(\hat{z}, \Omega)$ in E^n), $\left\langle \psi, \frac{\partial F(\hat{z})}{\partial z} \delta z \right\rangle \leq 0$.

The reader may easily modify the proof of Theorem 1 so as to apply to Theorem 1'. Finally, it should be pointed out that all conditions such as continuity differentiability, etc., imposed on the various functions need only hold in a neighborhood of the optimal point.

II. APPLICATIONS

We shall now show how a number of classical optimization problems can be cast in the form of the Basic Problem, and we shall then apply Theorem 1 or Theorem 1' to rederive severa classical conditions for optimality, as well as to obtain some new ones.

1. Classical Theory of Lagrange Multipliers

The classical constrained minimization problem admits equality constraints only. Thus, it is the Basic Problem with $\Omega = E^n$, the entire space. Clearly, E^n is a linearization of the first kind for E^n at any point $z \in E^n$.

Thus, we conclude from Theorem 1 that if $\hat{z}$ is an optimal solution of the Basic Problem, with $\Omega = E^n$, then there exists a nonzero vector $\psi \in E^{m+1}$ such that

(21) $$\left\langle \psi, \frac{\partial F(\hat{z})}{\partial z} \delta z \right\rangle \leq 0 \text{ for all } \delta z \in E^n$$

This may be rewritten as

(22) $$\left\langle \frac{\partial F(\hat{z})^T}{\partial z} \psi, \delta z \right\rangle \leq 0 \text{ for all } \delta z \in E^n$$

Since for any $\delta z \in E^n$, $-\delta z$ is also in E^n, we conclude from (22) that

(23) $$\frac{\partial F(\hat{z})^T}{\partial z} \psi = 0$$

Now, $\frac{\partial F(\hat{z})^T}{\partial z}$ is a $n \times (m+1)$ matrix with columns $\nabla f(\hat{z})$, $\nabla r^1(\hat{z})$, . . . $\nabla r^m(\hat{z})$, where $\nabla f(\hat{z}) = \left(\frac{\partial f(\hat{z})}{\partial z_1}, \ldots, \frac{\partial f(\hat{z})}{\partial z_n} \right)$, $\nabla r^i(\hat{z}) = \left(\frac{\partial r^i(\hat{z})}{\partial z_1}, \ldots, \frac{\partial r^i(\hat{z})}{\partial z_n} \right)$. We may therefore expand (23) into the form

(24) $$\psi^o \nabla f(\hat{z}) + \sum_{i=1}^{m} \psi^i \nabla r^i(\hat{z}) = 0$$

We have thus reproved the following classical result.

Theorem 2. Let $f, r^1, r^2, \ldots, r^m$ be real valued, continuously differentiable functions on E^n. If $\hat{z} \in E^n$ minimizes $f(z)$ subject to the constraints $r^i(z) = 0$, $1, 2, \ldots, m$, then there exist scalar multipliers, $\psi^o, \psi^1, \ldots, \psi^m$, not all zero, such that the function H on E^n which they define by

(25) $$H(z) = \psi^o f(z) + \sum_{i=1}^{m} \psi^i r^i(z)$$

has a stationary point at $z = \hat{z}$, i.e., (24) is satisfied.

It is usual to assume that the gradient vectors $\nabla r^i(z)$, $i = 1, 2, \ldots, m$, are linearly independent for all z such that $r(z) = 0$. This precludes $\sum_{i=1}^{m} \psi^i \nabla r^i(\hat{z}) = 0$ and hence in (24) $\psi^o \neq 0$. Multiplying (24) by $1/\psi^o$ and letting $\hat{\lambda}_i = \psi^i/\psi^o$, $i = 1, 2, \ldots, m$, we now deduce the more commonly seen condition.

Theorem 2'. If $\hat{z}$ minimizes $f(z)$ subject to $r(z) = 0$, and the gradients $\nabla r_i(\hat{z})$, $i = 1, 2, \ldots, m$, are linearly independent, then there exists a vector $\hat{\lambda} \in E^m$ such that the Lagrangian L on $E^n \times E^m$, defined by

$$(26) \qquad L(z, \lambda) = f(z) + \sum_{i=1}^{m} \lambda^i r^i(z)$$

has a stationary point at $(\hat{z}, \hat{\lambda})$.

We note that by (24) $\frac{\partial L(\hat{z}, \hat{\lambda})}{\partial z} = 0$ and that $\frac{\partial L(\hat{z}, \hat{\lambda})}{\partial \lambda} = r(\hat{z}) = 0$, by assumption.

2. Nonlinear Programming

Let $f : E^n \to E^1$, $r : E^n \to E^m$, and $q : E^n \to E^k$ be continuously differentiable functions. The standard Nonlinear Programming Problem is that of minimizing $f(z)$ subject to the constrains that $r(z) = 0$ and $q(z) \leq 0$.

This corresponds to the special case of the Basic Problem, with $\Omega = \{z : q(z) \leq 0\}$. We shall now show how Theorem 1 can be used to obtain various commonly known necessary conditions for $\hat{z}$ to be optimal. The presentation is divided into two parts. It should be noted that the necessary conditions obtained in Part I are stronger than those obtained in Part II.

Given a particular point $z \in \Omega$, we shall often have occasion to divide the components of the inequality constraints functions, q^i, $i = 1, \ldots, k$, into two sets; those for which $q^i(z) = 0$ and those for which $q^i(z) < 0$. To simplify notation we introduce the following definition.

<u>Definition.</u> For $z \in \Omega$, let the index set $I(z)$ be defined by

(27) $$I(z) = \{i : q^i(z) = 0\}$$

The constraints q^i, $i \in I(z)$ will be called the <u>active constraints</u> at z. We shall denote by $\overline{I(z)}$ the complement of $I(z)$ in $\{1, \ldots, k\}$.

<u>Part I.</u> The set $\Omega = \{\hat{z} : q(z) \leq 0\}$ introduced above is assumed to satisfy the following condition:

<u>Assumption (A1).</u>† Let $z \in \Omega$ be an optimal solution of the nonlinear Programming Problem. Then, there exists a vector $h \in E^n$ such that

$$\langle \nabla q^i(\hat{z}), h \rangle < 0 \text{ for all } i \in I(\hat{z})$$

A sufficient condition for (A1) to be satisfied is that the vectors $\nabla q^i(\hat{z})$, $i \in I(\hat{z})$ be linearly independent (see Corollary to Lemma 3).

<u>Definition:</u> For any $z \in \Omega$, the internal cone of Ω at z, denoted by $IC(z, \Omega)$, is defined by

$$IC(z, \Omega) = \{\delta z : \langle \nabla z^i(z), \delta z \rangle < 0 \text{ for all } i \in I(z)\}$$

---

† When some of the functions q^i, $i \in I(z)$, are linear, it suffices to require that there exist a vector $h \in E^n$ such that $\langle \nabla q^i(z), h \rangle \leq 0$ for these functions and $\langle \nabla q^i(z), h \rangle < 0$ for the remaining functions q^i, $i \in I(z)$.

By assumption (Al), the convex cone $IC(\hat{z}, \Omega)$ is nonempty. It is a simple exercise in the use of Taylor's Theorem to prove the following lemma.

Lemma 1. If $IC(z, \Omega) \neq \emptyset$, the empty set, then

(i) $IC(z, \Omega)$ is a linearization of the first kind of Ω at z,

(ii) $\overline{IC(z, \Omega)} = \{\delta z : \langle \nabla q^i(z), \delta z \rangle \leq 0 \text{ for all } i \in I(z)\}$

When specialized to the Nonlinear Programming Problem, Theorem 1 assumes the following form.

Theorem 3. If $\hat{z}$ is an optimal solution to the Nonlinear Programming Problem, with (Al) satisfied, then there exists a nonzero vector $\psi \in E^{m+1}$, with $\psi^o \leq 0$, such that for all $\delta z \in \overline{IC(\hat{z}, \Omega)} = \{\delta z: \langle \nabla q^i(\hat{z}), \delta z \rangle \leq 0$ for all $i \in I(\hat{z})\}$,

$$\left\langle \frac{\partial H(\hat{z})}{\partial z}, \delta z \right\rangle \leq 0$$

where $H(z) = \psi^o f(z) + \sum_{i=1}^{m} \psi^i r^i(z)$.

Using Theorem 3 and Farkas Lemma (see Appendix 2) we obtain the following necessary condition for optimality, whihch is in a form more familiar to specialists in mathematical programming.

Theorem 4. If $\hat{z}$ is an optimal solution to the Nonlinear Programming Problem, with (A1) satisfied, then there exist a nonzero vector $\psi \in E^{m+1}$, with $\psi^o \le 0$, and a vector $\mu \in E^k$, with $\mu \le 0$, such that

$$\text{(i)} \quad \psi^o \nabla f(\hat{z}) + \sum_{i=1}^{m} \psi^i \nabla r^i(\hat{z}) + \sum_{i=1}^{k} \mu^i \nabla q^i(\hat{z}) = 0$$

and

$$\text{(ii)} \quad \sum_{i=1}^{k} \mu^i q^i(\hat{z}) = 0$$

Proof. From Theorem 3,

$$\left\langle \frac{\partial H(\hat{z})}{\partial z}, \delta z \right\rangle \le 0$$

for all δz such that $\langle \nabla q^i(\hat{z}), \delta z \rangle \le 0$, $i \in I(\hat{z})$.
By Farkas Lemma, there exist scalars $\mu^i \le 0$, $i \in I(\hat{z})$ such that

$$\frac{\partial H(\hat{z})}{\partial z} + \sum_{i \in I(\hat{z})} \mu^i \nabla q^i(\hat{z}) = 0$$

Let $\mu^i = 0$ for $i \in \overline{I(\hat{z})}$. This completes the proof.

Most of the other well-known necessary conditions for Nonlinear Programming Problems can be obtained from Theorem 4 by making additional assumptions on the functions r and q. For example, the following corollaries to Theorem 4 are immediate consequences of that theorem.

Corollary 1. If assumption (A1) is satisfied and the vectors $\nabla r^i(\hat{z})$, $i = 1, \ldots, m$, are linearly independent, then there exist vectors $\psi \in E^{m+1}$, $\mu \in E^k$ which satisfy the conditions of Theorem 4 and such that $(\psi^o, \mu) \neq 0$.

Corollary 2. If $\nabla r^i(\hat{z})$, $i = 1, \ldots, m$, together with $\nabla q^i(\hat{z})$, $i \in I(\hat{z})$, are linearly independent vectors, there exists a vector $\psi \in E^{m+1}$ satisfying the conditions of Theorem 4 with $\psi^o < 0$.

The assumption in Corollary 2 is a well-known [11] sufficient condition for the Kuhn-Tucker constraint qualification to be satisfied. When it is added to Theorem 4 we obtain a slightly restricted form† of the Kuhn-Tucker Theorem [4].

Corollary 3. If there exists a vector $h \in E^n$ such that $\langle \nabla q^i(z), h \rangle < 0$ for all $i \in I(\hat{z})$, $\langle \nabla r^i(\hat{z}), h \rangle = 0$ for $i = 1, \ldots, m$, and the vectors $\nabla r^i(\hat{z})$, $i = 1, \ldots, m$, are linearly independent, then there exists a vector $\psi \in E^{m+1}$ satisfying the conditions of Theorem 4 with $\psi^o < 0$.

The assumption in this corollary is a sufficient condition for the weakened constraint qualification [13] to be satisfied. Augmented by this assumption, Theorem 4 becomes a slightly restricted form† of the Kuhn-Tucker Theorem with the weakened constraint qualification.

† In practice, the Kuhn-Tucker constraint conditions can rarely be shown to be satisfied unless the restrictions imposed in Corollaries 2 and 3 hold.

Part II. We shall now derive a necessary condition for the Nonlinear Programming Problem which is not based on assumption (A1) and hence is weaker than the necessary condition stated in Theorem 4. This condition was first proved by Mangasarian and Fromovitz [5] using the implicit function theorem and a lemma by Motzkin [12].

Whenever the assumption (A1) is not satisfied, it is possible to show that the vectors $\nabla q^i(\hat{z})$, $i \in I(\hat{z})$ can be summed to zero with non-positive scalars. This is established in the following two lemmas.

Lemma 2. Let $K \subset E^n$ be a nonempty closed convex cone such that for every nonzero vector $d \in k$, $-d \notin K$. Then there exists a vector $h \in E^n$ such that

$$\langle h, k \rangle < 0 \text{ for all nonzero } k \in K$$

The proof of this lemma is a straightforward but somewhat tedious exercise, and is therefore omitted.

Lemma 3. Suppose that assumption (A1) is not satisfied for the set $\Omega = \{z : q(z) \le 0\}$. Then there exists a nonzero vector $\mu \in E^k$, with $\mu \le 0$, such that

$$\text{(i)} \quad \sum_{i=1}^{k} \mu^i \nabla q^i(\hat{z}) = 0$$

(ii) $\sum_{i=1}^{k} \mu^i q^i(\hat{z}) = 0.$

Proof. Consider the closed convex cone

$$K = \{v \mid v = \sum_{i \in I(\hat{z})} \alpha_i \nabla q^i(\hat{z}), \ \alpha_i \geq 0\}$$

Since there does not exist an $h \in E^n$ such that $\langle h, \nabla q^i(\hat{z}) \rangle < 0$ for every $i \in I(\hat{z})$, we conclude from Lemma 2 that there exists a nonzero vector $d \in K$ such that $-d \in K$. Thus

$$d = \sum_{i \in I(\hat{z})} \beta^i \nabla q^i(\hat{z}), \quad \beta^i \geq 0, \text{ not all } \beta^i = 0,$$

and

$$-d = \sum_{i \in I(\hat{z})} \gamma^i \nabla q^i(\hat{z}), \quad \gamma^i \geq 0, \text{ not all } \gamma^i = 0.$$

Let $\mu^i = -(\beta^i + \gamma^i)$ for $i \in I(\hat{z})$, and let $\mu^i = 0$ for $i \in \overline{I(\hat{z})}$. Then $\mu = (\mu^1, \ldots, \mu^k)$ is the desired vector.

Corollary. A sufficient condition for the assumption (A1) to be satisfied is that the vectors $\nabla q^i(\hat{z})$, $i \in I(\hat{z})$, be linearly independent.

Lemma 3 may be combined with Theorem 4 to give a necessary condition for optimality which does not require that (A1) be satisfied for this, the most general case of the Nonlinear Programming Problem, we obtain the following necessary condition for optimality.

Theorem 5. If $\hat{z}$ is an optimal solution to the nonlinear programming problem, then there exists a vector $\psi \in E^{m+1}$ and a vector $\mu \in E^k$, with $\psi^o \leq 0$ and $\mu \leq 0$, ψ and μ not both zero, such that

$$\text{(i)} \quad \psi^o \nabla f(\hat{z}) + \sum_{i=1}^{m} \psi^i \nabla r^i(\hat{z}) + \sum_{i=1}^{k} \mu^i \nabla q^i(\hat{z}) = 0$$

and

$$\text{(ii)} \quad \sum_{i=1}^{k} \mu^i q^i(\hat{z}) = 0.$$

Proof. If (A1) is satisfied, the theorem is a slightly weaker statement of Theorem 4. If (A1) is not satisfied, let μ be the vector specified in Lemma 3, and let $\psi = 0$.

Finally, we note that if we let $r \equiv 0$, Theorem 5 becomes the well-known Fritz John necessary condition for optimality [3].

We have thus shown that most of the known necessaryconditions for nonlinear programming problems, previously derived by diverse and often unrelated techniques, can now be obtained simply by applying Thereom 1 and Farkas' Lemma.

3. Optimal Control

In the field of optimal control of discrete time systems, necessary conditions for optimality have been developed by Jordan and Polak [6], Halkin [7], Holtzman [8], and Rosen [9]. By recasting the optimal control problem in the form of the Basic Problem, it is possible to obtain from Theorems 1 and 1' essentially all of the above mentioned results in a unified manner. Furthermore, the derivation given in this paper is significantly simpler in most cases. In addition, Theorem 1 and 1' together with Farkas' Lemma yield necessary conditions for optimality for a class of bounded state space problems; a result which is new with this paper.

The general Optimal Control Problem that we will consider takes the following form:

Given a system described by the difference equation

$$x_{i+1} - x_i = f_i(x_i, u_i), \quad x_i \in E^n, \; u_i \in E^m, \; i = 0, \ldots, k-1 \tag{28}$$

Find a control sequence $(u_0, u_1, \ldots, u_{k-1})$ and a corresponding trajectory $(x_o, x_1, \ldots, x_k)$ such that

(i) $u_i \in U_i \subset E^m$ for $i = 0, \ldots, k-1$ (Control Constraints)

(ii) $x_i \in \Omega_i = \{x_i : q_i(x_i) \leq 0\}$, $q_i : E^n \to E^{n_i}$, $i = 0, \ldots, k$
(State Space Constraints)

and, in addition, the initial and terminal states, x_0 and x_k, satisfy

(29) (iii) $g_0(x_0) = 0, \quad g_0 : E^n \to E^{\ell_0}$ (Initial Manifold Constraint)

(iv) $g_k(x_k) = 0, \quad g_k : E^n \to E^{\ell_k}$ (Terminal Manifold Constraint)

and such that

$$\sum_{i=0}^{k-1} f_i^o(x_i, u_i) \text{ is minimized.}$$

We make the following assumptions on the various sets and functions appearing above.

Assumptions

(a) $f_i : E^n \times E^m \to E^n$ is a $C^{(1)}$ function for $i = 0, \ldots, k-1$,

(b) For every $u_i \in U_i$, and for all $i = 0, \ldots, k-1$, the radial cone $RC(u_i, U_i)$ is a linearization of the first kind for U_i at u_i,

(30) (c) g_0 and g_k are $C^{(1)}$ functions whose Jacobian matrices have maximum rank,

(d) For all $x_i \in \Omega$, $i = 0, \ldots, k$, the gradients of the active constraints, $\nabla q_i^j(x_i)$, $j \in I(x_i)$ [see (27)], are linearly independent vectors,

(e) $f_i^o : E^n \times E^m \to E^1$ is a $C^{(1)}$ function for $i = 0, \ldots, k-1$.

This problem may be reformulated in the form of the Basic Problem, i.e., $\{\min\ f(z) : r(z) = 0,\ z \in \Omega\}$, by making the following identifications. Let $z = (x_0, x_1, \ldots, x_k, u_0, \ldots, u_{k-1}) \in E^{(k+1)n+km}$, and let f, r, and Ω be defined by

$$\text{(i)} \quad f(z) = \sum_{i=0}^{k-1} f_i^o(x_i, u_i)$$

$$\text{(31)} \qquad \text{(ii)} \quad r(z) = \begin{bmatrix} x_1 - x_0 - f_0(x_0, u_0) \\ \vdots \\ x_k - x_{k-1} - f_{k-1}(x_{k-1}, u_{k-1}) \\ g_0(x_0) \\ g_k(x_k) \end{bmatrix}$$

$$\text{(iii)} \quad \Omega = \Omega_0 \times \Omega_1 \times \ldots \times \Omega_k \times U_0 \times U_1 \times \ldots \times U_{k-1}$$

Clearly f and r have the required differentiability properties. The cone

$$\text{(32)} \qquad C(z, \Omega) = IC(x_0, \Omega) \times \ldots \times IC(x_k, \Omega_k) \times RC(u_0, U_0) \times \ldots \times RC(u_{k-1}, U_{k-1}),$$

where $IC(x_i, \Omega_i)$ and $RC(u_i, U_i)$ were defined earlier, is obviously a linearization of the first kind for Ω at z since assumption (d) and

Lemma 3 guarantee that $IC(x_i, \Omega_i)$ is nonempty for every $i = 0, \ldots, k-1$, and by Lemma 1 it is a linearization of the first kind for Ω_i at x_i, while $RC(u_i, U_i)$, for $i = 0, \ldots, k-1$, is a linearization of the first kind by assumption (b). Therefore, we may apply Theorem 1, from which we conclude that if $\hat{z}$ is an optimal solution to the Optimal Control Problem, then there exists a nonzero vector $\psi = (p^o, \pi)$, with $p^o \leq 0$, and $\pi = (-p_1, \ldots, -p_k, \mu_0, \mu_k)$, where $p_i \in E^n$, $\mu_0 \in E^{\ell_0}$, $\mu_k \in E^{\ell_k}$, such that

$$(33) \qquad p^o \frac{\partial f(\hat{z})}{\partial z} \delta z + \left\langle \pi, \frac{\partial r(\hat{z})}{\partial z} \delta z \right\rangle \leq 0$$

for all $\delta z \in \overline{C(\hat{z}, \Omega)}$. Substituting for f and r in (33) and expanding, we get

$$p^o \sum_{i=0}^{k-1} \frac{\partial f_i^o(\hat{x}_i, \hat{u}_i)}{\partial x_i} \delta x_i + \sum_{i=0}^{k-1} \frac{\partial f_i^o(\hat{x}_i, \hat{u}_i)}{\partial z_i} \delta u_i$$

$$+ \sum_{i=0}^{k-1} \left\langle -p_{i+1}, \delta x_{i+1} - \delta x_i - \frac{\partial f_i(\hat{x}_i, \hat{u}_i)}{\partial xi} \delta x_i - \frac{\partial f_i(\hat{x}_i, \hat{u}_i)}{\partial x_i} \delta u_i \right\rangle$$

$$(34) \qquad + \left\langle \mu_0, \frac{\partial g_0(\hat{x}_0)}{\partial x_0} \delta x_0 \right\rangle + \left\langle \mu_k, \frac{\partial g_k(\hat{x}_k)}{\partial x_k} \delta x_k \right\rangle \leq 0$$

for every $\delta x = (\delta x_0, \ldots, \delta x_k, \delta u_0, \ldots, \delta u_{k-1}) \in \overline{C(\hat{z}, \Omega)}$.

The usual form of the necessary conditions in terms of a Hamiltonian, adjoint equation, transversality conditions, etc. are obtained by considering special forms of δz. The conditions obtainable by this procedure are summarized in Theorem 6 below.

Theorem 6. If $\hat{z} = (\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_k, \hat{u}_0, \ldots, \hat{u}_{k-1})$ is an otpimal solution to the Optimal Control Problem, then there exist vectors p_0, $p_1, \ldots, p_k$ in E^n, $\lambda_0, \lambda_1, \ldots, \lambda_k$, $\lambda_i \in E^{n_i}$, with $\lambda_i \leq 0$, $\mu_0 \in E^{\ell_0}$ $\mu_k \in E^{\ell_k}$, and a scalar $p^o \leq 0$, such that

(i) Not all of the quantities p^o, p_0, $p_1, \ldots, p_k$, μ_0, μ_k are zero

(ii) $$p_i - p_{i+1} = \left[\frac{\partial f_i(\hat{x}_i, \hat{u}_i)}{\partial x_i}\right]^T p_{i+1} + \left[\frac{\partial f_i^o(\hat{x}_i, \hat{u}_i)}{\partial x_i}\right]^T p^0$$

$$+ \left[\frac{\partial q_i(\hat{x}_i)}{\partial x_i}\right]^T \lambda_i \text{ for } i = 0, 1, 2, \ldots, k-1$$

(III) $$p_k = \left[\frac{\partial g_k(\hat{x}_k)}{\partial x_k}\right]^T \mu_k + \left[\frac{\partial q_k(\hat{x}_k)}{\partial x_k}\right]^T \lambda_k ;$$

(iv) $$p_0 = - \left[\frac{\partial g_0(\hat{x}_0)}{\partial x_0}\right]^T \mu_0$$

(v) $\left\langle \lambda_i, q_i(x_i) \right\rangle = 0$ for $i = 0, \ldots, k$

$$\text{(vi)} \quad \left\langle \left[\frac{\partial f_i^o(\hat{x}_i, \hat{u}_i)}{\partial u}\right]^T p^o + \left[\frac{\partial f_i(\hat{x}_i, \hat{u}_i)}{\partial u}\right]^T p_{i+1}, \delta u_i \right\rangle \leq 0$$

for all $\delta u_i \in \overline{RC(\hat{u}_i, U_i)}$.

To prove all of the above conditions would be somewhat laborious, therefore, we will only derive condition (vi) to demonstrate how one proceeds.

Let $\delta z = (0, \ldots, 0, \delta u_i, 0, \ldots, 0)$ with $\delta u_i \in \overline{RC(\hat{u}_i, U_i)}$. Clearly $\delta z \in \overline{C(\hat{z}, \quad)}$, and the inequality (34) reduces to

$$p^o \frac{\partial f_i^o(\hat{x}_i, \hat{u}_i)}{\partial u_i} \delta u_i + \left\langle p_{i+1}, \frac{\partial f_i(\hat{x}_i, \hat{u}_i)}{\partial u_i} \delta u_i \right\rangle \leq 0.$$

Simple rearrangement yields (vi).

It should be remarked at this point that the derivation of conditions (ii), (iii), and (v) require the use of Farkas' Lemma (see Appendix 2), while condition (iv) is simply a definition.

To the authors' knowledge the above, quite general, necessary condition has not been obtained previously, although Rosen [9] did obtain a similar result under substantially more restrictive assumptions on the sets U_i.

In the special case when there are no state space constraints, i.e., $q_i \equiv 0$ for $i = 0, \ldots, k$, Theorem 6 reduces to the following.

Corollary. If the functions $q_i \equiv 0$ for $i = 0, \ldots, k$, and $\hat{z}$ is an optimal solution to the Optimal Control Problem, then there exist vectors $p_0, \ldots, p_k$ in E^n, $\mu_0 \in E^{\ell_0}$, $\mu_k \in E^{\ell_k}$, and a scalar $p^o \leq 0$, such that

(i) not all of the quantities $p^o, p_0, \ldots, p_k$ are zero,

(ii) $$p_i = p_{i+1} = \left[\frac{\partial f_i(\hat{x}_i, \hat{u}_i)}{\partial x_i}\right]^T p_{i+1} + \left[\frac{\partial f_i^o(\hat{x}_i, \hat{u}_i)}{\partial x_i}\right]^T p^o,$$

$i = 0, \ldots, k-1$

(iii) $$p_k = \left[\frac{\partial g_k(\hat{x}_k)}{\partial x_k}\right]^T \mu_k$$

(iv) $$p_0 = -\left[\frac{\partial g_0(\hat{x}_0)}{\partial x_0}\right]^T \mu_0$$

(v) $$\left[\frac{\partial f_i^o(\hat{x}_i, \hat{u}_i)}{\partial u_i}\right]^T p^o + \left\langle \left[\frac{\partial f_i(\hat{x}_i, \hat{u}_i)}{\partial u}\right]^T p_{i+1}, \delta u_i \right\rangle \leq 0$$

for all $\delta u_i \in \overline{RC(\hat{u}_i, U_i)}$.

This is the condition derived by Jordan and Polak [6].

A Maximum Principle. Halkin [7] and Holtzman [8] have shown that by making some additional assumptions, condition (v) in the above corollary may be replaced by a stronger condition, which is usually called a Maximum Principle. Both Halkin's and Holtzman's results can be obtained from Theorem 1', but, for simplicity, we shall only show how Halkin's results are obtained.

The optimal control problem considered by Halkin differs from the Optimal Control Problem stated at the beginning of this section in the following way.

(i) There are no state space constraints other than the initial and terminal manifold constraints, i.e., $q_i \equiv 0$ for $i = 0, 1, \ldots, k$.

(ii) Assumptions (a) and (e) for the Optimal Control Problem are replaced by the following:

(a') For every $u_i \in U_i$, the functions $f_i(\cdot, u_i)$, $i = 0, \ldots, k-1$, are continuously differentiable on E^n.

(35) (e') For every $u_i \in U_i$, the functions $f_i^o(\cdot, u_i)$, are continuously differentiable functions on E^n.

(iii) Assumption (b) is replaced by the following:

(b') For every $x \in E^n$ and every $i = 0, 1, \ldots, k-1$, the sets $\underline{f}_i(x, U_i)$ are convex, where $\underline{f}_i = E^n \times E^m \to E^{n+1}$ is defined by $\underline{f}_i(x, u) = [f_i^o(x, u), f_i(x, u)]$

The reformulation of the Halkin Problem as a Basic Problem differs only slightly from that used for the Optimal Control Problem.

First, we introduce new variables $\underline{v}_i = (v_i^o, v_i) \in E^{n+1}$ with $v_i = (v_i^1, v_i^2, \ldots, v_i^n) \in E^n$ where $i = 0, \ldots, k\text{-}1$. Then we let $z = (x_0, x_1, \ldots, x_k, \underline{v}_0, \ldots, \underline{v}_{k-1}) \in E^{(k+1)n+k(n+1)}$, and we define the functions f and r and the set Ω by

$$\text{(i)} \quad f(z) = \sum_{i=0}^{k-1} v_i^o ,$$

$$\text{(36)} \qquad \text{(ii)} \quad r(z) = \begin{bmatrix} x_1 - x_0 - v_0 \\ \vdots \\ x_k - x_{k-1} - v_{k-1} \\ g_0(x_0) \\ g_k(x_k) \end{bmatrix}$$

$$\text{(iii)} \quad \Omega = \{z = (x_0, \ldots, x_k, \underline{v}_0, \ldots, v_{k-1}) : \underline{v}_i \in \underline{f}_i(x_i, U_i)\} .$$

For the linearization of the set Ω at $\hat{z}$ we take the cone

$$\text{(37)} \qquad C(\hat{z}, \Omega) = \Big\{ \delta z : \delta z = (\delta x_0, \ldots, \delta x_k, \delta \underline{v}_0, \ldots, \delta \underline{v}_k) \text{ and } \\ \delta \underline{v}_i \in \left\{ \frac{\partial f_i(\hat{x}_i, \hat{u}_i)}{\partial x_i} \delta x_i \right\} + RC(\hat{\underline{v}}_i, \underline{f}_i(\hat{x}_i, U_i)) \Big\} .$$

Clearly $C(\hat{z},\Omega)$ is a convex cone. We shall now show that $C(\hat{z},\Omega)$ is a linearization of the second of Ω at $\hat{z}$.

Let $\delta z^1, \ldots, \delta z^r$ be any finite collection of linearly independent vectors in $C(\hat{z}, \Omega)$, with $\delta z^i = (\delta x_0^i, \ldots, \delta x_k^i, \delta \underline{v}_0^i, \ldots, \delta \underline{v}_{k-1}^i)$. For each $i=1, \ldots, r$, and for each $j=0, \ldots, k-1$, there exists an $\epsilon^i j > 0$ such that $\underline{v}_j + \epsilon^i j(\delta \underline{v}_j^i - \frac{\partial f_j(\hat{x}_j, \hat{u}_j)}{\partial x_j} \delta x_j^i) \in \underline{f}_j(\hat{x}_j, U_j)$. As a consequence, there exists an $\epsilon > 0$ and vectors $u_j^i \in U_j$ such that

$$(38) \qquad \underline{v}_j + \epsilon \, \delta \underline{v}_j^i = \epsilon \frac{\partial f_j(\hat{x}_j, \hat{u}_j)}{\partial x_j} \delta x_j^i + \underline{f}_j(\hat{x}_j, u_j^i).$$

Let $C_0 = \text{co}\{\hat{z}, \hat{z} + \epsilon \, \delta z^1, \ldots, \hat{z} + \epsilon \, \delta z^r\}$

Let $z \in C_0$ be arbitrary, and let $\delta z = z - \hat{z}$. Then we may write

$$\delta z = \sum_{i=1}^{r} \mu^i \epsilon \, \delta z^i \quad \text{where} \quad \mu^i \geq 0, \; \sum_{i=1}^{r} \mu^i \leq 1,$$

or

$$\delta z = Z_\mu$$

where $Z = (\epsilon \delta z^1, \epsilon \delta z^2, \ldots, \epsilon \delta z^r)$ is a matrix with columns $\epsilon \delta z^i$, and $\mu = (\mu^1, \ldots, \mu^r)$ is an r-vector. For every $z \in C_0$, the vector μ is uniquely determined by the expression

$$\mu = Y \delta z$$

where Y is a matrix whose rows, y_i, $i=1, \ldots, r$, satisfy $\langle y_i, \epsilon \delta z^j \rangle = \delta_{ij}$, the Kronecker delta, for $i, j=1, \ldots, r$.

The map $\zeta : C_0 \to \Omega$ is defined as follows. For every $z = (x_o, \ldots, x_k, \underline{v}_0, \ldots, \underline{v}_{k-1}) \in C_0$, and corresponding

$\delta z = (\delta x_0, \ldots, \delta x_k, \underline{\delta v}_0, \ldots, \underline{\delta v}_{k-1}) = z - \hat{z}$, let $\zeta(z) = (y_0, \ldots, y_k, w_0, \ldots, w_{k-1})$ with

$$\text{(39)} \quad \text{(i)} \quad y_j = z_j, \quad j = 0, \ldots, k$$

$$\text{(ii)} \quad w_j = \underline{f}_j(x_j, \hat{u}_j) + \sum_{i=1}^{r} \mu^i(\delta z) \left[\underline{f}_j(x_j, u_j^i) - \underline{f}_j(x_j, \hat{u}_j) \right]$$

for $j = 0, \ldots, k-1$

where

$$\mu(\delta z) = (\mu^1(\delta z), \ldots, \mu^r(\delta z)) = Y \, \delta z, \quad \text{and}$$

u_j^i, $i = 1, \ldots, r$, $j = 0, \ldots, k-1$ were defined in (38).

The range of $\zeta(z)$ is contained in Ω because of the convexity of Ω. Since it is clear that $\zeta(z)$ is continuously differentiable, the reader may verify that $\zeta(z)$ is the identity map plus a small term, as required in the definition of a linearization of the second kind, by expanding $\zeta(z)$ about z.

Theorem 1' may now be applied to this problem to obtain the usual separation results, i.e., if $\hat{z} = (\hat{x}_0, \hat{x}_1, \ldots, \hat{x}_k, \hat{u}_0, \hat{u}_1, \ldots, \hat{u}_{k-1})$ is an optimal solution to the Halkin Problem, then there exists a nonzero vector $\psi = (p^o, \pi)$ in the $p^o \in 0$ and $\pi = (-p \qquad , \ldots, -p_k, \mu_0, \mu_k)$, where $p_i \in E^n$, $\mu_0 \in E^{\ell o}$, $\mu_k \in E^{\ell k}$, such that

$$p^o \sum_{i=0}^{k-1} \delta v_i^o + \sum_{i=0}^{k-1} \langle -p_{i+1}, \delta x_{i+1} - \delta v_i \rangle + \langle \mu_0, \frac{\partial g_0(\hat{x}_o)}{\partial x_0} \delta x_0 \rangle$$

$$\text{(40)} \qquad + \langle \mu_k, \frac{\partial g_k(x_k)}{\partial x_k} \delta x_k \rangle \leq 0$$

for all $\delta z = (\delta x_0, \dots, \delta x_k, \delta \underline{v}_0, \dots, \delta \underline{v}_k) \in C(z,\Omega)$, where $p^o \leq 0$, and the vector $(p^o, -p_1, \dots, -p_k, \mu_0, \mu_k) \neq 0$. By taking appropriate perturbations we can obtain Halkin's necessary condition [7].

<u>Theorem 7.</u> If $(\hat{u}_0, \hat{u}_1, \dots, \hat{u}_{k-1})$ is an optimal control sequence and $(\hat{x}_0, \hat{x}_1, \dots, \hat{x}_k)$ is a corresponding optimal trajectory for the Halkin Problem, then there exist vectors $p_0, \dots, p_k \in E^n$, $\mu_0 \in E^{\ell o}$, $\mu_k \in E^{\ell k}$, and a scalar $p^o \leq 0$, such that

(i) not all of the quantities $p^o, p_0, \dots, p_k, \mu_0, \mu_k$ are zero,

(ii) $$p_i - p_{i+1} = \left[\frac{\partial f_i(\hat{x}_i, \hat{u}_i)}{\partial x_i}\right]^T p_{i+1} + \left[\frac{\partial f_i^0(\hat{x}_i, \hat{u}_i)}{\partial x_i}\right]^T p^o, \quad i = 0, \dots, k-1$$

(iii) $$p_k = \left[\frac{\partial g_k(\hat{x}_k)}{\partial x_k}\right]^T \mu_k$$

(iv) $$p_o = -\left[\frac{\partial g_0(\hat{x}_0)}{\partial x_0}\right]^T \mu_0$$

(v) $$p^0 f^0(\hat{x}_i, \hat{u}_i) + \langle p_{i+1}, f_i(\hat{x}_i, \hat{u}_i)\rangle \leq p^o f^o(\hat{x}_i, u_i) + \langle p_{i+1}, f_i(\hat{x}_i, u_1)\rangle$$

for all $u_i \in U_i$.

The reader can obtain most of the results by straightforward substitution of appropriate perturbations, δz, in (40). We shall prove only condition (v).

Let $\delta z = (0, \ldots, \delta \underline{v}_i, \ldots, 0)$, with $\delta \underline{v}_i \in RC(\hat{\underline{v}}_i, \underline{f}_i(\hat{x}_i, U_i)$. This is certainly an admissible perturbation, and, for this δz, (40) reduces to

$$(41) \qquad p^o \delta v_i^o + \quad p_{i+1}, \delta v_i \quad \leq 0 \quad \text{for all} \quad \underline{\delta v}_i \in \overline{RC(\hat{\underline{v}}_i, \underline{f}_i(\hat{x}_i, U_i)}$$

Since $\underline{f}_i(\hat{x}_i, U_i)$ is a convex set, the vectors $\underline{f}_i(\hat{x}_i, u_i) - \underline{f}_i(\hat{x}_i, \hat{u}_i)$ belong to $RC(\hat{\underline{v}}_i, \underline{f}_i(\hat{x}_i, U_i))$ for all $u_i \in U_i$. Therefore, from (41), we get

$$p^o \left[f_i^o(\hat{x}_i, u_i) - f_i^o(\hat{x}_i, \hat{u}_i) \right] + \left\langle p_{i+1}, \; f_i(\hat{x}_i, u_i) - f_i(\hat{x}_i, \hat{u}_i) \right\rangle \leq 0$$

for all $u_i \in U_i$. Condition (v) follows immediately.

Holtzmann obtains exactly the same result as Halkin, (i.e., Theorem 7), under the less restrictive assumption that the sets, $f_i(x_i, U_i)$, are only directionally convex (see Holtzmann [8]). The derivation of this result from Theorem 1' proceeds in essentially the same manner as the derivation of Theorem 7 above.

Remark. It has already been pointed out that Theorem 7 differs from the corollary to Theorem 6 only in the condition (v). In fact, using the method outlined above, a Maximum Principle can be derived in the presence of state space constraints of the type considered in (29, (iii)), provided all the other assumptions of Halkin or Holtzmann are satisfied. One then gets a theorem identical to Theorem 6 except that condition (vi) is replaced by the Maximum Principle, i.e., condition (v) of Theorem 7. Theorem 7 then becomes a corollary to this more general result.

CONCLUSION

We have shown that a wide class of constrained minimization problems can be reduced to a common canonical form, the so-called Basic Problem, for which we have derived necessary conditions of optimality. It is rather clear that the present paper does not exhaust all the possible permutations and combinations of necessary conditions or minimization problems that can be treated by reduction to the Basic Problem. To name but a few, not discussed herein explicitly, we can point out optimal control problems with nonseparable constraints, such as total energy, total fuel, or else involving products of trajectory and control variables, which can also be reduced to the Basic Problem. However, one gets for these problems a necessary condition which applies to the entire trajectory and which does not necessarily break down into a series of conditions applicable at each sampling instant. One can also consider optimal control or nonlinear programming problems in which the "trajectory" constraint sets are specified in more general form than equalities or inequalities. The necessary conditions derived in this paper can be suitably modified to cover such cases, yielding transversality conditions in terms of polar cones rather than in terms of gradient vectors.

Although nothing has been said in this paper about sufficient conditions, it is clear that under assumptions such as convexity, it is possible to show that some of the necessary conditions given here are also sufficient.

Finally, it should be pointed out that the general approach presented in this paper, is the result of hindsight, an irritation with fragmentation, and the authors' conviction that in terms of problem solving, the geometric approach taken has great conceptual and intuitive advantages.

APPENDIX I. THE BROUWER FIXED POINT THEOREM

In proving Theorem 1, the authors have used a modified version of the Brouwer Fixed Point Theorem. The conventional form of the theorem, which is stated and proved in reference [14], is worded as follows.

Brouwer Fixed Point Theorem. If $f(\cdot)$ is a continuous map from the unit sphere in E^n into the unit sphere in E^n, then $f(\cdot)$ has a fixed point.

The version used in this paper is stated without proof by Dieudonné [15]. Since the proof is very short, it is included here.

Theorem. If $f(\cdot)$ is a continuous map from the unit sphere in E^n into E^n with $f(x) = x + g(x)$, where $\|g(x)\| < 1$ for all x with $\|x\| = 1$, then the origin is contained in the range of $f(\cdot)$.

Proof. To say that the origin is contained in the range of $f(\cdot)$ is equivalent to saying that the function $h(x) = -g(x)$ has a fixed point. Let us define the function $h_1(\cdot)$ by

$$h_1(x) = \begin{cases} -g(x) & \text{if } \|g(x)\| \leq 1 \\ -g(x)/\|g(x)\| & \text{if } \|g(x)\| > 1 \end{cases}$$

Clearly, $h_1(\cdot)$ is a continuous function from the unit sphere in E^n into the unit sphere in E^n. Therefore, by the Brouwer Fixed Point Theorem, $h_1(\cdot)$ has a fixed point, say x_1. If $\|x_1\| = 1$, then $\|g(x_1)\| < 1$, by hypothesis, and hence $\|h_1(x_1)\| < 1$, a contradiction. Thus $\|x_1\| = \|h_1(x_1)\| < 1$, which implies that $h_1(x_1) = -g(x_1)$ and hence x_1 is a fixed point of $-g(\cdot)$.

APPENDIX II. FARKAS' LEMMA

Farkas' Lemma [16] is a frequently quoted result in mathematical programming, which is also well known in other fields. However, the authors are not aware of a readily accessible, simple proof for this lemma and, consequently, include one here.

Farkas Lemma. Given a set of vectors $a_1, a_2, \ldots, a_m$ belonging to E^n, a vector b in E^n satisfies

$$\langle b, x \rangle \leq 0 \text{ for all } x \in \{x : \langle a_i, x \rangle \leq 0,\ i = 1, 2, \ldots, m\}$$

if and only if there exist scalars $\lambda_i \geq 0$, $i = 1, 2, \ldots, m$, such that

$$b = \sum_{i=1}^{m} \lambda_i a_i$$

Proof. $\Leftarrow$ Obvious.

$\Rightarrow$ Let $C = \{x : x = \sum_{i=1}^{m} \lambda_i a_i, \ \lambda_i \geq 0 \text{ for } i = 1, 2, \ldots, m\}$, and let $PC = \{\bar{x} : \langle \bar{x}, x \rangle \leq 0 \text{ for all } x \in C\}$. The set PC is called the polar to the cone C. Clearly, both C and PC are closed convex cones. At this point we digress to prove a proposition.

Proposition. If C is a closed convex cone in E^n, then $P(PC) = C$.

Proof. Clearly, $P(PC) \supset C$. Suppose $y \in P(PC)$, but $y \notin C$. Since C is a closed convex cone, there exists a vector $z \in E^n$ such that $\langle z, y \rangle > 0$. Clearly, $z \in PC$. But then $y \notin P(PC)$, which is a contradiction. Thus, $P(PC) = C$.

In our case, $PC = \{\bar{x} : \langle \bar{x}, \sum_{i=1}^{m} \mu_i a_i \rangle \leq 0, \ \mu_i \geq 0\}$, i.e.,

$PC = \{\bar{x} : \sum_{i=1}^{m} \mu_i \langle x, a_i \rangle \leq 0 \text{ for all } \mu_i \geq 0\}$, and hence

$PC = \{\bar{x} : \langle \bar{x}, a_i \rangle \leq 0 \text{ for } i = 1, 2, \ldots, m\}$. Then $b \in P(PC)$ since $\langle b, \bar{x} \rangle \leq 0$ for all $\bar{x} \in PC$ by hypothesis. By the proposition we just proved, $b \in C$. Q.E.D.

REFERENCES

1. E. J. McShane, "On multipliers for Lagrange problems," Am. J. Math., 61, 1939, pp. 809-819.

2. L. S. Pontryagin et. a., The Mathematical Theory of Optimal Processes, Interscience, New York, N. Y., 1962.

3. F. John, "Extremum problems with inequalities as side conditions" in Studies and Essays, Courant Anniversary Volume, edited by K. O. Friedrichs, O. E. Neugebauer and J. J. Stoker, Interscience, New York, N. Y., 1948, pp. 187-204.

4. H. W. Kuhn and A. W. Tucker, "Nonlinear Programming," Proc. of the Second Berkeley Symposium on Mathematic Statistics and Probability," Univ. of California Press, Berkeley, California, 1951, pp. 481-492.

5. O. L. Mangasarian and S. Fromovitz, "The Fritz John necessary optimality conditions in the presence of equality and inequality constraints," Sheel Development Company Paper No. P1433, 1965.

6. B. W. Jordan and E. Polak, "Theory of a class of discrete optimal control systems," J. Electronics and Control, 17, 1964, pp. 697-713.

7. H. Halkin, "A maximum principle of the Pontryagin type for systems described by nonlinear difference equations," to appear in SIAM Journal on Control.

8. J. M. Holtzman, "On the maximum principle for nonlinear discrete-time systems," to appear in Trans. of PGAC, IEEE.

9. J. B. Rosen, "Optimal control and convex programming," MRC Technical Report No. 547, Math. Research Center, Univ. of Wisconsin, February 1965.

10. L. Neustadt, Univ. of Southern California, personal communication.

11. K. Arrow and L. Hurwicz, "Reduction of constrained maxima to saddle-point problems," Proc. of the Third Berkeley Symposium on Mathematical Statistics and Probability, Univ. of California Press, Berkeley, California, 1956.

12. T. S. Motzkin, "Two consequences of the transportation theorem of linear inequalities," Econometrica, 19, 1951, pp 184-185.

13. K. Arrow, L. Hurwicz, and H. Uzawa, "Constraint qualifications in maximization problems," Naval Research Logistics Quarterly, 8, 1961, pp 175-191.

14. W. Hurewicz and H. Wallman, Dimension Theory, Princeton Univ. Press, Princeton, N. J., 1948, pp. 40-41.

15. J. Dieudonné, Foundations of Modern Analysis, Academic Press, New York, N. Y., 1960, p. 269.

16. J. Farkas, "Über die Theorie der einfachen Ungleichungen," Journal für die reine und angewandte Mathematik, Vol. 124, 1901, pp. 1-27.