AN EXTREMAL PROBLEM IN BANACH SPACE WITH APPLICATION TO OPTIMAL CONTROL

by

P. P. Varaiya

Memorandum No. ERL-M180

12 August 1966

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

Manuscript submitted: 2 August 1966

The research reported herein was supported in part by the Joint Services Electronics Program (U. S. Army, U. S. Navy and U. S. Air Force) under Grant No. AF-AFOSR-139-65 and by the National Aeronautics and Space Administration under Grant No. NsG-354, Supplement 2.

AN EXTREMAL PROBLEM IN BANACH SPACE WITH APPLICATIONS TO OPTIMAL CONTROL

1. INTRODUCTION

The theory of optimal control has received a new impetus through the papers of Gamkrelidze [1] and Neustadt [2-4]. It seems clear now that the optimal control problem should be studied as an extremal problem in a Banach space or a locally convex space. The motivation for this generality is derived from the study of optimal control problems with trajectory constraints. This author has arrived at the problem formulated in Sec. 3 through the study of nonlinear programming in general spaces [5]. The results obtained are similar to those of Neustadt, but the method of proof and the motivation appear to be different. It is hoped that this paper will serve as a common framework for both optimal control and nonlinear programming problems.

2. NOTATIONS, DEFINITIONS, AND A PRELIMINARY RESULT

Throughout this paper, unless otherwise stated, X and Y will denote arbitrary real Banach spaces. All undefined terms can be found in Dunford and Schwartz [4].

Def. 2.1.a. A function $f: X \to Y$ is differentiable at a point x if there is a continuous linear function, $f'(x)$, mapping X into Y such that

$$\lim_{\substack{\varepsilon \to 0+ \\ w \to z}} \frac{f(x + \varepsilon w) - f(x)}{\varepsilon} = \langle f'(x), z \rangle = f'(x)(z) .$$

b. A function $f: X \to Y$ is Fréchet-differentiable at a point x if there is a continuous linear function, $f'(x)$, mapping X into Y such that

$$\lim_{|h| \to 0} \frac{f(x + h) - f(x) - \langle f'(x), h \rangle}{|h|} = 0 ;$$

f is continuously Fréchet-differentiable at x if it is differentiable in some neighborhood M of x and the mapping $z \to f'(z)$ is a continuous function of M into B(X, Y).†

† B(X, Y) denotes the Banach space of continuous linear functions from X into Y under the usual sup norm.

In addition to a linear approximation of a function at a point we shall need a "linear" approximation of a set at a point.

Def. 2.2. Let A be an arbitrary subset of X and let $x \in A$. For each neighborhood M of x let $C(A \cap M, x)$ denote the smallest closed cone, with vertex 0,[†] containing the set $\{A \cap M - x\} = \{z - x | z \in A \cap M\}$. Let $\mathfrak{M}$ be the neighborhood system at x. Then the set

$$LC(A, x) = \bigcap_{M \in \mathfrak{M}} C(A \cap M, x)$$

is called the local cone of A at x.

Def. 2.3.a. Let A be an arbitrary subset of X and let $x \in A$. The set

$$LP(A, x) = \left\{ x^* \in X^* \,\middle|\, \langle x^*, z \rangle \leq 0 \;\forall z \in LC(A, x) \right\}$$

in X^*[††] is called the local polar of A at x.

Def. 2.3.b. If L is a cone then $P(L) = LP(L, 0)$.

Remark 2.1a. The local cone is a nonempty (it always contains 0) closed cone and the local polar is a nonempty convex closed cone.

---

† All cones referred to in this paper have vertex 0.

†† X^* denotes the Banach space of all real-valued, continuous linear functions on X under the usual sup norm.

b. A useful alternative characterization of the local cone is given by the next fact.

Fact 2.1. The following statements are equivalent.

a. $z \in LC(A, x)$.

b. There exist sequences $\{x_n\} \subset A$, $\{\lambda_n\}$, with $\lambda_n > 0$ for all n, such that $x_n \to x$ and $\lambda_n(x_n - x) \to z$ as $n \to \infty$.

c. There exist sequences $\{z_n\} \subset X$, $\{\varepsilon_n\}$, such that $\varepsilon_n \to 0$, $z_n \to z$ as $n \to \infty$, and $\varepsilon_n > 0$, $(x + \varepsilon_n z_n) \in A$ for all n.

Proof. Trivially b. and c. are equivalent. The equivalence of a. and b. follows directly from Def. 2.2 using a standard Cantor diagonal argument.

Q.E.D.

The justification of the two linear approximations is provided by the following elementary but extremely useful result.

Theorem 2.1. Let f be a real-valued function defined on X, and A an arbitrary subset of X. Let $\underline{x} \in A$ be a solution of (2.1).

(2.1) $$\text{Maximize } f(x), \text{ subject to } x \in A .$$

Then, if f is differentiable (see Def. 2.1.a) at $\underline{x}$ we must have

(2.2) $$f'(\underline{x}) \in LP(A, \underline{x}) .$$

Proof. Let $z \in LC(A, \underline{x})$. We have to show that $\langle f'(\underline{x}), z \rangle \leq 0$. By Fact 2.1 there are sequences $\{z_n\}$ and $\{\mathcal{E}_n\}$ such that $z_n \to z$, $\mathcal{E}_n \to 0+$ as $n \to \infty$, and $\mathcal{E}_n > 0$, $x_n = (\underline{x} + \mathcal{E}_n z_n) \in A$ for all n. Since $\underline{x}$ is a solution of (2.1), $f(x_n) - f(\underline{x}) \leq 0$ for all n. Hence,

$$\frac{f(\underline{x} + \mathcal{E}_n z_n) - f(\underline{x})}{n} \leq 0 .$$

Passing to the limit as $n \to \infty$, we obtain (2.2) from Def. 2.1.a.

Q.E.D.

Remark 2.2. A. The definitions of differentiability (Def. 2.1.a), local cone, and local polar, are valid in arbitrary linear topological spaces. Fact 2.1 is valid if we replace "sequence" by "generalized sequence" or "net". Theorem 2.1 still remains true.

b. Theorem 2.1 shows that the elements of $LC(A, \underline{x})$ can be considered as "admissible" variations about $\underline{x}$. For many applications however, we have to consider a more restrictive class of variations. The next definition defines two such classes. Let A be an arbitrary subset of X and let $x \in A$.

Def. 2.4.a. A convex cone K is <u>in the local approximation of A at x</u> if for every finite set $\{k_0, \ldots, k_m\} \subset K$ and every $\delta > 0$, there is $\mathcal{E}_0 = \mathcal{E}_0(\delta, K) > 0$ such that for every $\mathcal{E} \in (0, \mathcal{E}_0]$ there is a continuous map $\eta_{\mathcal{E},\delta} : \overline{K} \to A$ of the form

$$\eta_{\mathcal{E},\delta}(\overline{k}) = x + \mathcal{E}\left(\overline{k} + \gamma_{\mathcal{E},\delta}(\overline{k})\right)$$

with $|\gamma_{\mathcal{E},\delta}(\overline{k})| \leq \delta$. Here $\overline{K}$ denotes the convex hull of $\{k_0, \ldots, k_m\}$.

Def. 2.4.b. A convex cone K is locally in A at x if there exists $\delta > 0$ such that

$$(x + k) \in A$$

whenever $k \in K$ and $|k| \leq \delta$.

Remark 2.3. Using Fact 2.1 we see that in both instances of Def. 2.4 the cone K is contained in LC(A, x).

3. STATEMENT OF THE MAIN THEOREM AND COMMENTS

Theorem 3.1. Let X and Y be real Banach spaces. Let f be a real-valued function defined on X, and g a continuous mapping from X into Y. Le A be a subset of X and let A_Y be a closed convex cone in Y. Let $\underline{x}$ be a solution of (3.1)

(3.1) Maximize f(x), subject to $g(x) \in A_Y$ and $x \in A$.

Suppose that there is a closed convex cone $K_1 \subset X$ such that f, g, K_1, A_Y, and Y satisfy either I or II.

I. a. f and g are differentiable at $\underline{x}$ (Def. 2.1.a). Let $G = g'(\underline{x})$.

b. Y is finite dimensional, i.e., $Y = R^m$ for some $m < \infty$.

c. Either $\overline{G(K_1) - A_Y + \{g(\underline{x})\}}$† $\neq Y$, or for every $z \in K_1$ such that $z \neq 0$ and $G(x) + g(\underline{x}) \in A_Y$, there exists a convex cone K (depending on z), satisfying the following conditions:

(i) $z \in K \subset X$,

(ii) $G(K) - A_Y + \{g(\underline{x})\} = Y$,

and (iii) K is in the local approximation of A at $\underline{x}$ (See Def. 2.4a).

d. The set $A_Y + \{g(\underline{x})\}$ is a closed subset of Y.

e. The set $\{y^* \cdot G \mid y^* \in LP(A_Y, g(\underline{x}))\}$ is a closed subset of X^*.

f. The set $P(K_1) + P(K_2)$ is a closed subset of X^*, where

$$K_2 = \left\{\delta x \mid G(\delta x) \in A_Y + \{g(\underline{x})\}\right\}.$$

II. a. f is differentiable at $\underline{x}$ and g is continuously Frechet-differentiable at $\underline{x}$(Def. 2.1.b). Let $G = g'(\underline{x})$.

b. Either $\overline{G(K_1) - A_Y + \{g(\underline{x})\}} \neq Y$, or for every $z \in K_1$, such that $z \neq 0$ and $G(z) + g(\underline{x}) \in A_Y$, there exists a closed convex cone K (depending on z), satisfying the following conditions:

(i) $z \in K \subset X$,

(ii) $G(K) - A_Y + \{g(\underline{x})\} = Y$,

and (iii) K is locally in A at $\underline{x}$ (see Def. 2.4.b).

---

† $\{g(\underline{x})\}$ denotes the one-dimensional subspace of Y spanned by $g(\underline{x})$.

c. The set $A_Y + \{g(\underline{x})\}$ is a closed subset of Y.

d. The set $\left\{ y^* \cdot G \,|\, y^* \in LP(A_Y, g(\underline{x})) \right\}$ is a closed subset of X^*.

e. The set $P(K_1) + P(K_2)$ is a closed subset of X^*, where $K_2 = \left\{ \delta x \,|\, G(\delta x) \in A_Y + \{g(\underline{x})\} \right\}$.

Then there exists a number $\mu \geq 0$ and a $y^* \in Y^*$, not both zero, such that

(3.2) $$\langle \mu f'(\underline{x}), \delta x \rangle + \langle y^*, G(\delta x) \rangle \leq 0 \quad \forall \delta x \in K_1 ,$$

(3.3) $$\langle y^*, g(\underline{x}) \rangle = 0 ,$$

(3.4) $$\langle y^*, y \rangle \geq 0 \quad \forall y \in A_Y .$$

Comments: The main deference between conditions I and II lies in the fact that the requirement of finite-dimensionality of Y is dropped in II, whereas it is critical in I. But if Y is infinite-dimensional, then we require that G and K_1 have to be a "better" approximation to g and A, respectively. (Compare I.a with II.a and I.c iii) with I.b(iii).) Conditions I.c, I.d, and I.e are of a technical nature and can be shown to be redundant if A_Y is a polyhedral cone. Similarly, II.c and II.e can be shown to be redundant if A_Y is a polyhedral cone; although II.d may not be satisfied even in this case. Thus, if A_Y is a polyhedral cone, Y is finite dimensional, and K_1 is in the local approximation of A at $\underline{x}$, then I is satisfied. Similarly, if A_Y is a polyhedral cone, K_1 is locally in A at $\underline{x}$, and II.d is satisfied, then II is satisfied. Finally, it is worth noting that the conjunction of (3.3) and (3.4) is equivalent to (3.5).

$$(3.5) \qquad -y^* \in LP(A_Y, g(\underline{x})) .$$

4. PROOF OF THE MAIN THEOREM

The proof is divided into two parts; the first takes care of the degeneracies that may arise, the second case is the important one.

$$\text{Let} \quad Q = \overline{G(K_1) - A_Y + \{g(\underline{x})\}}$$

Case 1. Suppose $Q \neq Y$.

Then Q is a closed, convex cone in Y and Q is a proper subset of Y, so that there exists [6, p. 452, Theorem 10] a $y^* \in Y^*$, $y^* \neq 0$ such that

$$(4.1) \qquad \langle y^*, y \rangle \leq 0 \quad \forall y \in Q .$$

In particular, (4.1) implies that

$$\langle y^*, G(\delta x) \rangle \leq 0 \quad \forall \delta x \in K_1 ,$$

$$\langle y^*, y \rangle \geq 0 \quad \forall y \in A_Y ,$$

and $$\langle y^*, g(\underline{x}) \rangle = 0 ,$$

so that (3.2), (3.3), and (3.4) are satisfied with $y^* \neq 0$ and $\mu = 0$.

Case 2. $Q = Y$.

Let $A_X = \{x \in X \,|\, g(x) \in A_Y\}$ and let $K_2 = \left\{ \delta x \in X \,|\, G(\delta x) \in A_Y + \{g(\underline{x})\} \right\}$

We will now prove the important fact that if either I or II is satisfied, then

(4.2) $$LC(A \cap A_X, \underline{x}) \supset K_1 \cap K_2 .$$

Let $z \in K_1 \cap K_2$. Therefore, $z \in K_1$ and there exists a number λ such that $G(z) + \lambda g(\underline{x}) \in A_Y$. Since $g(\underline{x}) \in A_Y$, we can assume that $\lambda > 0$. Also, since $LC(A \cap A_X, \underline{x})$, K_1, and K_2 are cones, we can assume that $\lambda = 1$. Thus we have

$$z \in K_1 \text{ and } G(z) + g(\underline{x}) \in A_Y .$$

Suppose that I is satisfied:

Then because of I. c, there exists a convex cone K in X such that

(i) $z \in K$,

(ii) $G(K) - A_Y + \{g(\underline{x})\} = Y$, and (iii) K is in the local approximation of A at $\underline{x}$. By Theorem A-1 of the Appendix there exist sequences $\{z_n\} \subset X$, $\{\varepsilon_n\}$, with $\varepsilon_n > 0$, $(\underline{x} + \varepsilon_n z_n) \in A$, and $g(\underline{x} + \varepsilon_n z_n) \in A_Y$, for each n, such that $z_n \to z$, and $\varepsilon_n \to 0$ as $n \to \infty$. But then by Fact 2.1, $z \in LC(A \cap A_X, \underline{x})$.

Suppose that II is satisfied:

Then because of II.b, there exists a closed convex cone K in X such that

(i) $z \in K$,

(ii) $G(K) - A_Y + \{g(\underline{x})\} = Y$,

and (iii) K is locally in A at $\underline{x}$. By Theorem A-2 of the Appendix there exist sequences $\{z_n\} \subset X$, $\{\varepsilon_n\}$, with $\varepsilon_n > 0$, $(\underline{x} + \varepsilon_n z_n) \in A$, and $g(\underline{x} + \varepsilon_n z_n) \in A_Y$, for each n, such that $z_n \to z$, and $\varepsilon_n \to 0$ as $n \to \infty$. Once again by Fact 2.1, $z \in LC(A \cap A_X, \underline{x})$. This proves (4.2).

Now by Theorem 2.1, since f is differentiable at $\underline{x}$, we must have (4.3).

(4.3) $$f'(\underline{x}) \in LP(A \cap A_X, \underline{x}) .$$

From the definition of the local polar, and from (4.2), it is evident that

(4.4) $$LP(A \cap A_X, \underline{x}) \subset P(K_1 \cap K_2) .$$

Combining (4.3) and (4.4), we obtain (4.5).

$$f'(x) \in P(K_1 \cap K_2) \ . \tag{4.5}$$

Now K_1 and K_2 are closed, convex cones. Using this fact, the definition of local polar, and the strong separation theorem [6, p. 417, Theorem 10], it is easy to show that [5, Fact 1.3]

$$P(K_1 \cap K_2) = \overline{P(K_1) + P(K_2)} \ . \tag{4.6}$$

Because of I.f or II.e, $P(K_1) + P(K_2)$ is a closed subset of X^*, so that from (4.6) and (4.5) we can conclude

$$f'(\underline{x}) \in P(K_1) + P(K_2) \ . \tag{4.7}$$

Next we show that

$$P(K_2) = \overline{B} \ , \tag{4.8}$$

where

$$B = \left\{ y^* \cdot G \,|\, y^* \in LP\left(A_Y,\ g(\underline{x})\right) \right\} \ . \tag{4.9}$$

Let $x \in K_2$ and let $y^* \in LP\left(A_Y,\ g(\underline{x})\right)$.

Therefore,

$$(4.10) \qquad G(x) \in A_Y + \{g(\underline{x})\}$$

and,

$$(4.11) \qquad \langle y^*, y \rangle \leq 0 \quad \forall y \in A_Y ,$$

$$(4.12) \qquad \langle y^*, g(\underline{x}) \rangle = 0.$$

Combining (4.10), (4.11), and (4.12), we conclude that $\langle y^*, G(x) \rangle \leq 0$. Therefore $P(K_2) \supset \overline{B}$. Now suppoes that $P(K_2 \subset \overline{B}$. Then, by the strong separation theorem [6, p. 417, Theorem 10], it follows that there exists a vector $x \in K_2$, such that,

$$(4.13) \qquad \langle y^* \cdot G, x \rangle = \langle y^*, G(x) \rangle > 0 \qquad \forall y^* \in LP\left(A_Y, g(\underline{x})\right).$$

But since $x \in K_2$, we must have (4.10). It is easy to verify that

$$(4.14) \qquad A_Y + \{g(\underline{x})\} \subset LC\left(A_Y, g(\underline{x})\right)$$

so that from (4.10), $G(x) \in LC\left(A_Y, g(\underline{x})\right)$. But then $\langle y^*, g(x) \rangle \leq 0$ whenever $y^* \in LP\left(A_Y, g(\underline{x})\right)$. This contradicts (4.13), hence (4.8) is proved. By I.e or II.d, B is a closed subset of X^*, so that from (4.7) and (4.8) we obtain

(4.15)
$$f'(\underline{x}) \in B + P(K_1) .$$

Therefore, there exists $y^* \in Y^*$ such that

(4.16)
$$f'(\underline{x}) + y^* \cdot G \in P(K_1) ,$$

and

(4.17)
$$-y^* \in LP\left(A_Y, g(\underline{x})\right) .$$

But (4.16) is equivalent to (3.2) with $\mu = 1$, and (4.17) is equivalent to the conjunction of (3.3) with (3.4). This completes the proof of the theorem.

Q.E.D.

5. APPLICATION OF THEOREM 3.1

A. Discrete Optimal Control

Consider a difference equation,

$$z(k+1) = z(k) + h(z(k), u(k)), \quad k = 0, 1, \ldots$$

where $z \in Z$ is the state vector, $u \in U$ is the control vector and h is a continuous mapping from $Z \times U$ into Z. Z and U are arbitrary real Banach spaces. Let N be a fixed integer representing the duration of the process. Let A_0 and A_N be subsets of Z representing the initial and target set, respectively. Let $\Omega \subset U$ be the set of available controls. The payoff is given by a real-valued function f defined on $X = Z^{N+1} \times U^N$. We are required to find a solution to the following problem:

$$(5.1) \qquad \text{Maximize } f\Big(z(0), \ldots, z(N); u(0), \ldots, u(N-1)\Big)$$

subject to

$$(5.2) \qquad z(k+1) - z(k) - h(z(k), u(k)) = 0 \quad \text{for} \quad 0 \le k \le N-1$$

and

$$(5.3) \qquad z(0) \in A_0, \; z(N) \in A_N, \; u(k) \in \Omega \quad \text{for} \quad 0 \le k \le N-1.$$

We make the following identifications:

1. $X = Z^N \times U^{N-1}$ so that $x = \Big(z(0), \ldots, z(N), u(0), \ldots, u(N-1)\Big)$
2. $Y = Z^N$
3. g is a mapping from X into Y defined by
$$g\Big(z(0), \ldots, z(N), u(0), \ldots, u(N-1)\Big) = (z_0, \ldots, z_N)$$

where $z_k = z(k+1) - z(k) - h(z(k), u(k)), \; 0 \le k \le N-1.$

4. $A_Y = \{(0, \ldots, 0)\}$ so that A_Y is a closed convex cone.
5. $A = A_0 \times Z^{N-1} \times A_N \times \Omega^N$. A is a subset of X.

Then our problem can be restated as

(5.4) Maximize $f(x)$, subject to $g(x) \in A_Y$ and $x \in A$.

Let $\underline{x} = \left(\underline{z}(0), \ldots, \underline{z}(N), \underline{u}(0), \ldots, \underline{u}(N-1)\right)$ be a solution to (5.4). Let $\hat{K}_0$ and $\hat{K}_N$ be closed convex cones in Z and let Q_i for $0 \le i \le N-1$ be closed convex cones in U, such that the closed convex cone $K_1 = \hat{K}_0 \times Z^{N-1} \times \hat{K}_N \times Q_0 \times \cdots \times Q_{N-1}$, in X is locally in A at $\underline{x}$ (Def. 2.4.b). Let f be differentiable at $\underline{x}$ and let g be continuously Fréchet-differentiable at $\underline{x}$ with $g'(\underline{x}) = G$; the latter statement is equivalent to the statement that h is continuously Fréchet-differentiable at the points $(\underline{z}(k), \underline{u}(k))$ for $0 \le k \le N-1$. We also assume that the remaining conditions in II of Theorem 3.1 are satisfied. Then there exists a number $\mu \ge 0$ and $y^* \in Y^*$, not both zero, such that

$$(5.5) \qquad \langle \mu f'(\underline{x}), \delta x \rangle + \langle y^*, G(\delta x) \rangle \le 0 \quad \forall \delta x \in K_1,$$

$$(5.6) \qquad \langle y^*, g(\underline{x}) \rangle = 0,$$

$$(5.7) \qquad \langle y^*, y \rangle \ge 0 \quad \forall y \in A_Y.$$

The statements (5.6) and (5.7) are trivial because $g(\underline{x}) = 0$ and $A_Y = \{0\}$. Since $Y = Z^N$, therefore $Y^* = (Z^*)^N$ so that

$y^* = (z^*(1), \ldots, z^*(N))$ for some $z^*(k) \in Z^*$. Also using the definitions of X, K_1 and G, we see that (5.5) is equivalent to the following:

$$(5.8) \qquad \left\langle \mu \frac{\partial f}{\partial z(0)}, \delta z \right\rangle - \left\langle z^*(1), \delta z \right\rangle - \left\langle z^*(1) \cdot \frac{\partial h}{\partial z(0)} \right\rangle \leq 0$$

$$\forall \delta z \in \hat{K}_0 ,$$

$$(5.9) \qquad \left\langle \mu \frac{\partial f}{\partial z(k)}, \delta z \right\rangle + \left\langle z^*(k) - z^*(k+1), \delta z \right\rangle$$

$$- \left\langle z^*(k+1) \cdot \frac{\partial h}{\partial z(k)}, \delta z \right\rangle \leq 0$$

$$\forall \delta z \in Z, \ 1 \leq k \leq N-1 ,$$

$$(5.10) \qquad \left\langle \mu \frac{\partial f}{\partial z(N)}, \delta z \right\rangle + \left\langle z^*(N-1), \delta z \right\rangle \leq 0 \quad \forall \delta z \in \hat{K}_N ,$$

$$(5.11) \qquad \left\langle \mu \frac{\partial f}{\partial u(k)}, \delta u \right\rangle - \left\langle z^*(k+1) \cdot \frac{\partial h}{\partial u(k)}, \delta u \right\rangle \leq 0$$

$$\forall \delta u \in Q(k), \ 0 \leq k \leq N-1 ,$$

where the derivatives are evaluated at the optimal solution $\underline{x} = \left(\underline{z}(0), \ldots, \underline{z}(N), \underline{u}(0), \ldots, \underline{u}(N-1)\right)$. The statement (5.8) and (5.10) represent the so called "transversality" conditions; (5.11) is sometimes referred to as a "local maximum principle". (5.9) can be rewritten in the familiar form (5.12).

(5.12) $$\mu \frac{\partial f}{\partial z(k)} + z^*(k) - z^*(k+1) - z^*(k+1) \cdot \frac{\partial h}{\partial z(k)} = 0$$

for $1 \le k \le N-1$.

Remarks. 1. The conditions given in [7] are a special case of the relations (5.8), (5.10), (5.12) and (5.11).

2. The fact that we allow our state variables to be infinite dimensional also enables us to consider discrete stochastic optimal control problems. See [5] for an example.

B. Continuous Optimal Control

Let $\mathcal{H}$ be the linear space whose elements $h(x, t)$ are n-dimensional real vector-valued functions defined for $x \in R^n$ and $t \in I = [t_0, t_1]$. The functions h satisfy certain smoothness conditions in x and some integrability conditions in t. Let H be a quasi-convex subset of $\mathcal{H}$. For the precise conditions and definitions the reader is referred to Gamkrelidze [1] or Neustadt [2]. The relevance of the various assumptions made in the sequel to optimal control problems is also discussed in these references.

Now for any h in H, let x(t) be an absolutely continuous function such that

(5.13) $$\dot{x}(t) = h(x(t), t), \text{ for almost all } t \in I.$$

We shall regard such a function x as an element of the Banach space X of all continuous functions from the compact interval I into R^n. We also define A as the set consisting of all those elements $x \in X$ which are absolutely continuous and satisfy (5.13) for some h in H. Now let f be a real-valued function defined on X and let $g: X \to Y = R^m$ be a continuous mapping. Let A_Y be a closed polyhedral cone in Y. We wish to solve the following problem:

(5.14) $\quad$ Maximize f(x), subject to $g(x) \in A_Y$ and $x \in A$.

Let $\underline{x}$ be a solution of (5.14) so that

(5.15) $\quad \dot{\underline{x}}(t) = \underline{h}(\underline{x}(t), t)$ for almost all $t \in I$

for some $\underline{h} \in H$. Let [H] denote the convex hull of H, and consider the linear variational equation of (5.15),

(5.16) $$\delta x(t) = \frac{\partial \underline{h}}{\partial x}(\underline{x}(t), t)\, \delta x(t) + \Delta h(\underline{x}(t), t), \quad t \in I.$$

Here Δh is any arbitrary element of the set $\{[H] - \underline{h}\}$ and $\delta x(t_0) = \xi$ is any arbitrary n-vector. Let $\varphi(t)$ be the nonsingular matrix solution of the homogeneous matrix differential equation

$$\varphi(t) = \frac{\partial \underline{h}}{\partial x}(\underline{x}(t), t)\, \varphi(t)$$

with $\dot{\varphi}(t_0) = 1$, the identity matrix. Then the solution of (5.16) is (5.17).

$$(5.17) \qquad \delta x(t) = \varphi(t) \left\{ \xi + \int_{t_0}^{t} \varphi^{-1}(\tau)\, \Delta h(\underline{x}(\tau), \tau)\, d\tau \right\}, \quad t \in I.$$

Let $\hat{K} \subset X$ be the set of all δx which satisfy (5.17) for some $\xi \in R^n$ and some function $\Delta h \in \{[H] - \underline{h}\}$. Clearly $\hat{K}$ is convex. Let K be the convex cone generated by $\hat{K}$ and let $K_1 = \overline{K}$. Let us suppose that f and g are differentiable at $\underline{x}$ and let $G = g'(\underline{x})$. It has been shown by Neustadt [4, Theorem 3.1] that K is in the local approximation of A at $\underline{x}$. It is easy to see then, that $\overline{K}$ is also in the local approximation of A at $\underline{x}$. Since A_Y is a polyhedral cone, it follows from Theorem 3.1 that there exist numbers $\mu \leq 0, \lambda_1, \ldots, \lambda_m$ not all zero such that

$$(5.18) \qquad \mu \langle f'(\underline{x}), \delta x \rangle + \langle \lambda, G(\delta x) \rangle \leq 0 \quad \forall \delta x \in K_1,$$

$$(5.19) \qquad \langle \lambda, g(\underline{x}) \rangle = 0,$$

$$(5.20) \qquad \langle \lambda, y \rangle \geq 0 \quad \forall y \in A_Y,$$

where $\lambda = (\lambda_1, \ldots, \lambda_m)$.

Remarks. Following Neustadt [3], we can show that the relations (5.18), (5.19) and (5.20) imply the Pontryagin Maximum Principle. The assumption of quasi-convexity implies that K_1 is in the local approximation of A at $\underline{x}$. It can be shown that, in general, K_1 is not locally in A at $\underline{x}$. Hence, to apply Theorem 3.1 the finite-dimensionality of Y is essential. If, however, we are dealing with a linear control system with an admissible set of controls which is a convex polygon, then it can be shown that K_1 is locally in A at $\underline{x}$ so that we can allow Y to be an infinite dimensional Banach space.

6. ACKNOWLEDGMENTS

The author thanks Professor Neustadt for the definitions 2.1a and 2.4a and for his suggestions regarding the formulation of this problem. The author is also grateful to Mr. Nelson DaCunha for many helpful discussions. Last but not the least, sincere thanks are extended to the referee for his anonymous and painstaking help.

REFERENCES

[1] R. V. Gamkrelidze, "On some extremal problems in the theory of differential equations with applications to the theory of optimal control," J. SIAM, Ser A: Control, 3, 1965.

[2] L. W. Neustadt, "Optimal control problems as extremal problems in a Banach space," Proceedings of the Symposium on System Theory, Polytechnic Press of the Poly. Inst. of Brooklyn, Brooklyn, N. Y., 1965, pp. 215-224.

[3] L. W. Neustadt, "An abstract variational tyeory with applications to a broad class of optimization problems I: general theory," SIAM Control Journal, to be published.

[4] L. W. Neustadt, "An abstract variational theory with applications to a broad class of optimization problems II: applications," Technical Report USCEE Report 169, Electroncs Science Laboratory, University of Southern California, May 1966.

[5] P. P. Varaiya, "Nonlinear programming and optimal control," ERL Tech. Memo M-129, Electronics Research Laboratory, University of California, Berkeley, California, September 1965.

[6] N. Dunford and J. T. Schwartz, Linear Operators Part I, Interscience Publishers, Inc., New York, 1964.

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

APPENDIX

Theorem A-1. Let X be a real Banach space and g a continuous mapping from X into $Y = R^m$. Let A be a subset of X, and A_Y a closed convex cone in Y. Let $\underline{x} \in X$ be such that $\underline{x} \in A$ and $g(\underline{x}) \in A_Y$. Let g be differentiable at $\underline{x}$ and let $G = g'(\underline{x})$. Let $z \in X$ be such that $G(z) + g(\underline{x}) \in A_Y$. Suppose that there exists a convex cone $K \subset X$ such that

(i) $z \in K$,

(ii) $G(K) - A_Y + \{g(\underline{x})\} = Y$,

and (iii) K is in the local approximation of A at $\underline{x}$. Then there exist sequences $\{z_n\} \subset X$, $\{\mathcal{E}_n\}$, with $\mathcal{E}_n > 0$, $(\underline{x} + \mathcal{E}_n z_n) \in A$ and $g(\underline{x} + \mathcal{E}_n z_n) \in A_Y$ for each n, such that $z_n \to z$ and $\mathcal{E}_n \to 0$ as $n \to \infty$.

Proof. 1. First of all from (ii) and the fact that $g(\underline{x}) \in A_Y$, it is clear that 0 is in the interior of $G(K) - A_Y + g(\underline{x})$. Let Σ be a simplex in Y generated by the points $y_0, \ldots, y_m$ and containing 0 in its interior, such that $\Sigma \subset \{G(K) - A_Y + g(\underline{x})\}$. Therefore there exist vectors $k_0, \ldots, k_m$ in K and vectors $a_0, \ldots, a_m$ in A_Y such that

(A.1) $$y_i = G(k_i) - a_i + g(\underline{x}), \quad 0 \le i \le m.$$

Let $\theta > 0$ be such that if $y \in Y$ and $|y| \le \theta$, then $y \in \Sigma$.

2. Since g is differentiable at the point $\underline{x}$, it follows from Def. 2.1.a that

$$h(\mathcal{E}, \delta, k, \Delta x) = \frac{1}{\mathcal{E}}\left\{g\left(\underline{x} + \mathcal{E}(z + \delta\hat{k} + \Delta x)\right) - g(\underline{x})\right\} - G(z + \delta\hat{k}) \to 0$$

as $\mathcal{E} \to 0$, $\Delta x \to 0$, uniformly for $\hat{k} \in \hat{K}$ = convex hull of $\{k_0, \ldots, k_m\}$ and $\delta \in [0, 1]$. In other words, for each $\alpha > 0$ there exists $\beta = \beta(\alpha) > 0$ such that

$$|h(\mathcal{E}, \delta, \hat{k}, \Delta x)| \le \alpha$$

whenever $0 < \mathcal{E} \le \beta$, $|\Delta x| \le \beta$, $\delta \in [0, 1]$ and $\hat{k} \in \hat{K}$.

3. Since K is in the local approximation of A at $\underline{x}$ (by (iii)), it follows from Def. 2.4.a that, for each $\delta > 0$, there is an $\mathcal{E}_0 = \mathcal{E}_0(\delta) > 0$ such that for every $\mathcal{E} \in (0, \mathcal{E}_0]$, there is a continuous map $\eta_{\mathcal{E}, \delta}$ from $\{z + \overline{K}\}$ into A of the form

$$\eta_{\mathcal{E}, \delta}(z + \overline{k}) = \underline{x} + \mathcal{E}\left(z + \overline{k} + \gamma_{\mathcal{E}, \delta}(z + \overline{k})\right)$$

with $|\gamma_{\mathcal{E}, \delta}(z + \overline{k})| \le \delta$. Here $\overline{K}$ denotes the convex hull of $\hat{K} \cup \{0\}$.

4. a. Let α_n be a sequence of numbers such that $0 < \alpha_n \leq \theta$ for each n, and $\alpha_n \to 0$ as $n \to \infty$.

b. Let $\beta_n = \beta(\alpha_n)$ (see 2). We assume that $\beta_n \to 0$ as $n \to \infty$.

c. Let $\delta_n = \beta_n$, and $\mathcal{E}_0(n) = \mathcal{E}_0(\delta_n)$ used in the definition of the maps $\eta_{\mathcal{E},\delta}$ (see 3).

d. Let $\mathcal{E}_n = \min(\beta_n, \mathcal{E}_0(n), \frac{1}{2})$.

e. Let $\mu_n = \alpha_n/\theta$.

5. For each n we define a mapping of $\mu_n \Sigma$ into itself as follows:

For each $\mu_n y \in \mu_n \Sigma$ there is a unique vector $s_n(\mu_n y) = (\sigma_0, \ldots, \sigma_m)$ with $\sigma_i \geq 0$ for each i, $\sigma_0 + \cdots + \sigma_m = 1$, such that $\mu_n y = \mu_n(\sigma_0 y_0 + \cdots + \sigma_m y_m)$.

For each vector $\sigma = (\sigma_0, \ldots, \sigma_m)$ with $\sigma_i \geq 0$ for each i, and $\sigma_0 + \cdots + \sigma_m = 1$, let

$$h_n(\sigma) = \mu_n(\sigma_0 k_0 + \cdots + \sigma_m k_m)$$

$$\ell_n(\sigma) = \mu_n(\sigma_0 a_0 + \cdots + \sigma_m a_m)$$

and

$$q_n(\sigma) = -\frac{1}{\mathcal{E}_n}\left\{g\left(\eta_{\mathcal{E}_n, \delta_n}\left((z + h_n(\sigma)\right)\right) - g(\underline{x})\right\} + G\left(z + h_n(\sigma)\right).$$

Clearly the maps h_n, ℓ_n and g_n are continuous.

Let $\hat{A}$ denote the convex hull of $\{a_0, \ldots, a_m\}$. For each $\mu_n y \in \mu_n \Sigma$ let $\hat{k} \in \hat{K}$ and $\hat{a} \in \hat{A}$ be such that $h_n(s_n(\mu_n y)) = \mu_n \hat{k}$ and $\ell_n(s_n(\mu_n y)) = \mu_n \hat{a}$. Then, because of (A.1)

$$G\left(h_n\left(s_n(\mu_n y)\right)\right) - \ell_n\left(s_n(\mu_n y) + g(\underline{x})\right) \tag{A.2}$$

$$= G(\mu_n \hat{k}) - \mu_n \hat{a} + g(\underline{x}) = \mu_n y .$$

Also, for each $\mu_n y \in \mu_n \Sigma$, since $\mathcal{E}_n \le \beta_n$ and $\delta_n = \beta_n$ it can be verified that

$$\left|q_n\left(s_n(\mu_n y)\right)\right| \le \alpha_n = \mu_n \theta . \tag{A.3}$$

But this implies that $q_n\left(s_n(\mu_n \Sigma)\right) \subset \mu_n \Sigma$, so that by the Brouwer fixed point theorem, the continuous mapping $q_n \cdot s_n$ of $\mu_n \Sigma$ into itself has a fixed point, say $\mu_n y_n$. Then,

(A.4 $$\mu_n y_n = q_n\left(s_n(\mu_n y_n)\right)$$

$$= -\frac{1}{\mathcal{E}_n}\left\{g\left(\eta_{\mathcal{E}_n, \delta_n}(z + \mu_n \hat{k}_n)\right) - g(\underline{x})\right\} + G(z + \mu_n \hat{k}_n) ,$$

where $\mu_n \hat{k}_n = h_n s_n(\mu_n y_n)$. From 3, (A.5) $$\eta_{\mathcal{E}_n, \delta_n}(z + \mu_n \hat{k}_n) = \underline{x} + \mathcal{E}_n\left(z + \mu_n \hat{k}_n + \gamma_{\mathcal{E}_n, \delta_n}(z + \mu_n \hat{k}_n)\right).$$

Let

(A.6) $$z_n = z + \mu_n \hat{k}_n + \gamma_{\mathcal{E}_n, \delta_n}(z + \mu_n \hat{k}_n) .$$

Then clearly $\underline{x} + \mathcal{E}_n z_n \in A$ for each n, and $z_n \to z$ as $n \to \infty$. Also from (A.4), (A.5), and (A.6), we obtain

(A.7) $$\mu_n y_n = -\frac{1}{\mathcal{E}_n}\left\{g(\underline{x} + \mathcal{E}_n z_n) - g(\underline{x})\right\} + G(z + \mu_n \hat{k}_n) .$$

Defining $\hat{a}_n \in \hat{A}$ by $h_n s_n(\mu_n y_n) = \mu_n \hat{a}_n$, from (A.2) we obtain

(A.8) $$\mu_n y_n = G(\mu_n \hat{k}_n) - \mu_n \hat{a}_n + g(\underline{x}) ,$$

so that on substituting in (A.7) we have

(A.9) $$-\mu_n \hat{a}_n + g(\underline{x}) = -\frac{1}{\mathcal{E}_n}\left\{g(\underline{x} + \mathcal{E}_n z_n) - g(\underline{x})\right\} + G(z) .$$

Therefore,

(A.10) $$g(\underline{x} + \mathcal{E}_n z_n) = \mathcal{E}_n \mu_n \hat{a}_n + \mathcal{E}_n\left(G(z) + g(\underline{x})\right) + (1 - 2\mathcal{E}_n) g(\underline{x}).$$

Since $\hat{a}_n \in \hat{A} \subset A_Y$ and $\mathcal{E}_n \mu_n > 0$, therefore $\mathcal{E}_n \mu_n \hat{a}_n \in A_Y$. Also $G(z) + g(\underline{x}) \in A_Y$ by hypothesis; $(1 - 2\mathcal{E}_n) \geq 0$ and $g(\underline{x}) \in A_Y$, so that $(1 - 2\mathcal{E}_n)\, g(\underline{x}) \in A_Y$. Since A_Y is a convex cone, we conclude that

$$g(\underline{x} + \mathcal{E}_n z_n) \in A_Y \text{ for each } n \tag{A.11}$$

and the theorem is proved.

Q.E.D.

Theorem A-2. Let X and Y be real Banach spaces and let be be a continuous mapping of X into Y. Let A be a subset of X, and let A_Y be a closed, convex cone in Y. Let $\underline{x} \in X$ be such that $\underline{x} \in A$ and $g(\underline{x}) \in A_Y$. Let g be continuously Fréchet-differentiable at $\underline{x}$ and let $G = g'(\underline{x})$. Let $z \in X$ be such that $G(z) + g(\underline{x}) \in A_Y$. Suppose that there exists a closed convex cone $K \subset X$ such that

(i) $z \in K$,

(ii) $G(K) - A_Y + \{g(\underline{x})\} = Y$,

and (iii) K is locally in A at $\underline{x}$.

Then there exist sequences $\{z_n\} \subset X$, $\{\mathcal{E}_n\}$, with $\mathcal{E}_n > 0$, $(\underline{x} + \mathcal{E}_n z_n) \in A$, and $g(\underline{x} + \mathcal{E}_n z_n) \in A_Y$ for each n, such that $z_n \to z$ and $\mathcal{E}_n \to 0$ as $n \to \infty$.

Proof. The theorem is proved with the aid of Lemmas A-1 and A-2.

Lemma A.1 Let W and Y be real Banach spaces and let $\tilde{G}$ be a continuous linear mapping from W into Y. Let $\tilde{K}$ be a closed convex cone in W such that $\tilde{G}(\tilde{K}) = Y$. For each $\rho > 0$ let $\tilde{K}_\rho = \{w \in W \mid |w| \leq \rho, \; w \in \tilde{K}\}$. Then there is a number $m > 0$, independent of ρ, such that

$$\tilde{G}(\tilde{K}_\rho) \supset S_{m\rho}, \tag{A.12}$$

where $S_{m\rho}$ is the closed sphere in Y of center 0 and radius ρ.

Proof. This result is a generalization of the Interior Mapping Principle. Although the proof is long, it is a straightforward modification of that given by Dunford and Schwartz [6, pp. 55-56].

Q.E.D.

Lemma A.2. Let W, Y be real Banach spaces, and let $\tilde{g}$ be a mapping from W into Y such that $\tilde{g}$ is continuously Frechet-differentiable at a point $\underline{w} \in W$. Let $\tilde{G} = \tilde{g}'(\underline{w})$. Let $\tilde{K}$ be a closed, convex cone in W such that $\tilde{G}(\tilde{K}) = Y$. Let $\hat{w} \in \tilde{K}$ be any fixed vector. Then there exist sequences $\{\hat{w}_n\} \subset \tilde{K}$, $\{\mathcal{E}_n\}$ with $\mathcal{E}_n > 0$ and $\tilde{g}(\underline{w} + \mathcal{E}_n \hat{w}_n) = \tilde{g}(\underline{w}) + \tilde{G}(\mathcal{E}_n \hat{w})$ for each n, such that $\mathcal{E}_n \to 0$ and $\hat{w}_n \to \hat{w}$ as $n \to \infty$.

Proof. Let $m > 0$ be any number which satisfies (A.12). Let $v : W \to Y$ be the function defined by $v(w) = \tilde{g}(\underline{w} + w) - \tilde{g}(\underline{w}) - \tilde{G}(w)$. Then,

$$|v(\xi\hat{w} + w_1) - v(\xi\hat{w} + w_2)|$$

$$= |\tilde{g}(\underline{w} + \xi\hat{w} + w_1) - \tilde{g}(\underline{w} + \xi\hat{w} + w_2) - \tilde{G}(w_1 - w_2)|$$

$$= |\langle \tilde{g}'(\underline{w} + \xi\hat{w} + w_1),\ w_1 - w_2 \rangle + o_1(|w_1 - w_2|) - \tilde{G}(w_1 - w_2)| .$$

Therefore,

$$\frac{|v(\xi\hat{w} + w_1) - v(\xi\hat{w} + w_2)|}{|w_1 - w_2|} \leq \|\tilde{g}'(\underline{w} + \xi\hat{w} + w_1) - \tilde{G}\| + \frac{o_1(|w_1 - w_2|)}{|w_1 - w_2|} .$$

Also,

$$|v(\xi\hat{w} + w_1)| = |\tilde{g}(\underline{w} + \xi\hat{w} + w_1) - \tilde{g}(\underline{w}) - \tilde{G}(\xi w_1)| = o_2(|\xi\hat{w} + w_1|) .$$

Let $\xi_0 > 0$ be such that for $0 < \xi < \xi_0$,

(A.13) $$\frac{|v(\xi\hat{w} + w_1) - v(\xi\hat{w} + w_2)|}{|w_1 - w_2|} < \frac{m}{4} \text{ when } |w_i| < \xi_0$$

for $i = 1, 2$

and

(A.14) $$0_2(|\xi\hat{w} + w_1|) = o(\xi) < \frac{m}{4} \quad \text{for } |w_1| < \xi .$$

Let $\xi \in (0, \xi_0)$ be fixed.

Let $w_0 = 0$. Therefore, $\tilde{G}(w_0) = 0$.

Let $w_1 \in \tilde{K}$ such that $\tilde{G}(w_1 - w_0) = -v(\xi \hat{w} + w_0)$, and $|w_1 - w_0| < \frac{1}{m} |v(\xi\hat{w} + w_0)| < \frac{1}{m} o(\xi)$.

For $n \leq 1$, let $w_{n+1} \in \tilde{K}$ such that $\tilde{G}(w_{n+1} - w_n) = -v(\xi\hat{w} + w_n) + v(\xi\hat{w} + w_{n-1})$ and $|w_{n+1} - w_n| < \frac{1}{m} o(\xi)$. We first show that for $n \geq 0$, $|w_n| < \xi$ so that the above inequalities are valid. Firstly,

$$|w_0| = |0| < \frac{1}{m} o(\xi) < \frac{1}{4}\xi \text{ and } |w_1 - w_0| < \frac{1}{m} o(\xi) < \frac{1}{4} \quad .$$

Therefore, $|w_1| < \frac{1}{4}\xi$.

By induction on n,

$$|w_{n+1} - w_n| < \left(\frac{1}{m} o(\xi)\right)^n |w_1 - w_0| < \left(\frac{1}{4}\xi\right)^n |w_1 - w_0| .$$

Hence, for any integer $p \geq 0$,

$$|w_{n+p} - w_n| < \left(\frac{1}{4}\xi\right)^p \frac{1}{1 - (1/4)\xi} |w_1 - w_0| < \left(\frac{1}{4}\xi\right)^p \frac{2}{m} o(\xi) .$$

In particular, $|w_{n+1} - w_1| < \xi/2$ so that $|w_{n+1}| < \xi$. Also w_n converges. Let

$$\lim_{n\to\infty} w_n = \hat{w}(\xi).$$

Then $|\hat{w}(\xi)| < \frac{4}{m} o(\xi)$ and $\hat{w}(\xi) \in \tilde{K}$.

Now,

$$\tilde{G}(w_0) = 0,$$

$$\tilde{G}(w_1) - \tilde{G}(w_0) = -v(\xi\hat{w} + w_0),$$

$$\tilde{G}(w_2) - \tilde{G}(w_1) = -v(\xi\hat{w} + w_1) + v(\xi\hat{w} + w_0),$$

$$\vdots$$

$$\hat{G}(w_{n+1}) - \hat{G}(w_n) = -v(\xi\hat{w} + w_n) + v(\xi\hat{w} + w_{n-1}).$$

Adding both sides, we obtain

$$\tilde{G}(w_{n+1}) = -v(\xi\hat{w} + w_n) = -\tilde{g}(\underline{w} + \xi\hat{w} + w_n) + \tilde{g}(\underline{w}) + \tilde{G}(\xi\hat{w} + w_n).$$

Hence,

$$\tilde{g}(\underline{w} + \xi\hat{w} + w_n) = \tilde{g}(\underline{w}) + \tilde{G}(\xi\hat{w}) + \tilde{G}(w_n - w_{n+1}).$$

Passing to the limit, as $n \to \infty$, we obtain

$$\tilde{g}(\underline{w} + \quad \hat{w} + \hat{w}(\xi)) = \tilde{g}(\underline{w}) + \tilde{G}(\xi\hat{w}).$$

Let $\{\xi_n\}$ be any sequence of positive numbers such that $\xi_n \to 0$ as $n \to \infty$, and define $\hat{w}_n = \hat{w} + \frac{1}{\xi_n}\hat{w}(\xi_n)$. Since $\left|\frac{1}{\xi_n}\hat{w}(\xi_n)\right| < \frac{1}{\xi_n} \cdot \frac{1}{m} \cdot o(\xi_n)$, $\hat{w}_n \to \hat{w}$ as $n \to \infty$ and the assertion is proved.

Q. E. D.

Proof of Theorem A. 2.

We make the following appropriate identifications:

1. $W = X \times Y \times Y$,

2. $\tilde{K} = K \times A_Y \times \{g(\underline{x})\}$,

3. $\tilde{g}(x, y_1, y_2) = g(x_1) - y_1 + y_2$,

4. $\underline{w} = (\underline{x}, 0, 0)$ and $\hat{w} = (z, 0, 0)$.

5. $\tilde{G} = \tilde{g}'(\underline{w})$ so that, $\tilde{G}(x, y_1, y_2) = G(x_1) - y_1 + y_2$.

Then by Lemma A.2, there exist sequences $\{\hat{w}_n\} \subset \tilde{K}$, $\{\mathcal{E}_n\}$, with $\mathcal{E}_n > 0$ and $\tilde{g}(\underline{w} + \mathcal{E}_n \hat{w}_n) = \tilde{g}(\underline{w}) + \tilde{G}(\mathcal{E}_n \hat{w})$ for each n, such that $\mathcal{E}_n \to 0$ and $\hat{w}_n \to \hat{w}$ as $n \to \infty$.

Let $\hat{w}_n = (z_n, y_n^1, y_n^2)$. Then $z_n \in K$, $y_n^1 \in A_Y$, and $y_n^2 \in \{g(\underline{x})\}$ for each n. Let $y_n^2 = \lambda_n g(\underline{x})$. From 3 and 4 we obtain for each n,

$$g(\underline{x} + \mathcal{E}_n z_n) - \mathcal{E}_n y_n^1 + \mathcal{E}_n \lambda_n g(\underline{x}) = g(\underline{x}) + G(\mathcal{E}_n z) .$$

or

$$g(\underline{x} + \mathcal{E}_n z_n) = \mathcal{E}_n y_n^1 + \mathcal{E}_n(g(\underline{x}) + G(z)) + (1 - \mathcal{E}_n - \lambda_n \mathcal{E}_n) g(\underline{x}) .$$

Now, $y_n^1 \in A_Y$, $g(\underline{x}) \in A_Y$, $g(\underline{x}) + G(z) \in A_Y$. Also $\lambda_n \to 0$ as $n \to \infty$, so that for n sufficiently large $g(\underline{x} + \xi_n z_n) \in A_Y$. Also since K is locally in A at $\underline{x}$, for sufficiently large n we must have $(\underline{x} + \xi_n z_n) \in A$ and the theorem is proved.

Q.E.D.