WIS-CS-74-220

COMPUTER SCIENCES DEPARTMENT
University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53706

Received August 5, 1974

SUFFICIENT CONDITIONS FOR THE CONVERGENCE OF MONOTONIC MATHEMATICAL PROGRAMMING ALGORITHMS

by

R. R. Meyer

Technical Report # 220
August 1974

Errata — TR #220

p. 5 — Change (3.7) to read

"the accumulation points of $\{y_i\}$ form a continuum if $\{y_i\}$ does not converge."

— Change the last line to read

"conclusion (3.7) follows from a result of Ostrowski [8].)

p. 6 — Line -4, change "$\inf_{z \in H} \phi(z)$"

to "$\inf_{z \in S(G)} \phi(z)$"

R. R. Meyer
October 23, 1974

jp

Abstract

A global convergence theory for a broad class of "monotonic" nonlinear programming algorithms is given. The key difference between the approach presented here and previous work in this area by Zangwill, Meyer, and others, lies in the use of an appropriate definition of a fixed-point of a point-to-set mapping. The use of this fixed-point concept allows both a simplification and a strengthening and extension of previous results. In particular, actual convergence of the entire sequence of iterates (as opposed to subsequential convergence) and point-of-attraction theorems are established under weak hypotheses. Examples of the application of this theory to feasible direction algorithms are given.

1. Introduction

The purpose of this paper is to establish global convergence theorems for a broad class of "monotonic" nonlinear programming algorithms. The key difference between the approach presented here and the results of Zangwill [11] and Meyer [6] lies in the use an appropriate definition of a fixed-point of a point-to-set mapping. For those algorithms to which it is applicable, the use of this fixed-point concept allows both a simplification and a strengthening of those earlier results. In particular, it is possible under relatively weak hypotheses to prove that the entire sequence of iterates converges. Additionally, point-of-attraction theorems are proved, and an interesting characterization is obtained for a class of convergence algorithms that lack the continuity properties usually assumed. Examples of the application of this theory to the classes of feasible direction methods proposed by Topkis and Veinott [10] and Mangasarian [5] are given to illustrate the concepts.

2. A Class of Algorithms

The algorithms to be considered will be methods that address the problem:

$$\text{minimize } f(z) \qquad \text{subject to } z \in F \; , \tag{MP}$$

where f is a continuous function defined on a closed subset G of E^n, $G \supset F$. The basic characteristics common to the algorithms to be considered are (a) that they start at an arbitrary $y_0 \in G$, (b) that an iteration starting at a point y_i $(i = 0,1,2,\ldots)$ yields a point $y_{i+1} \in S(y_i)$, where S is a point-to-set mapping from G into the non-empty subsets of G, and (c) the existence of a continuous function $\phi: G \to R$ such that $\phi(y') \leq \phi(y)$ where $y' \in S(y)$. (The procedures (a) and (b) will be referred to as the algorithm corresponding to S, and when property (c) holds, the algorithm will be monotonic. Note, however, that these characteristics do not restrict the algorithms to be considered to be "primal" algorithms for which $f = \phi$, $G = F$. It is possible, for example, to deal with a dual algorithm such as Kelley's cutting plane method by taking $\phi = -f$ and G to be a set that properly contains F.) We wish to establish conditions that ensure that $||y_{i+1} - y_i|| \to 0$ and that the accumulation points of $\{y_i\}$ are fixed-points of S. Here it should be emphasized that by a fixed-point of S we mean a point $y^* \in G$ such that $S(y^*) = \{y^*\}$ (instead of simply $S(y^*) \supset \{y^*\}$).

3. Basic Convergence Theorems

To obtain our initial convergence result we need a slight strengthening of the hypotheses made in the previous section. In particular, we will need a compactness hypothesis, a continuity assumption, and a stronger form of monotonicity that requires the addition of some convenient terminology. A point-to-set mapping S will be said to be strictly monotonic (with respect to a function ϕ) at y if $y' \in S(y)$ implies $\phi(y') < \phi(y)$ whenever y is not a fixed-point of S. (The strict monotonicity property may be thought of as being vacuously satisfied at all fixed-points.) S is said to be upper semi-continuous (u.s.c.) at y if $z_i \in S(y_i)$ $(i = 0,1,2,\ldots)$, $y_i \to y$ and $z_i \to z$ imply $z \in S(y)$. These properties will be said to hold on G if they hold at all points in G. Finally a mapping S will be said to be uniformly compact on G if there exists a compact set H independent of y such that $S(y) \subset H$ for all $y \in G$. (Note that if S is also u.s.c. at y, this means that $S(y)$ is compact also.)

Theorem 3.1 Let S be a point-to-set mapping such that

(3.1) S is uniformly compact on G,

(3.2) S is u.s.c. on G, and

(3.3) S is strictly monotonic on G.

Under an additional finiteness assumption on the fixed-points we will then show that the entire sequence of iterates $\{y_i\}$ converges to $y*$. (This behavior sharply contrasts with earlier results in which it is shown only that the accumulation points of the iterates have certain desirable properties, but for which convergence of the full sequence of iterates cannot be concluded except under strong uniqueness assumptions.) The well-known mathematical programming algorithms that belong to this class of algorithms generally have the property that their fixed-points satisfy necessary optimality conditions of (MP) . If appropriate convexity assumptions hold, then the fixed-points will also be solutions of (MP) , otherwise they may only be local optima or saddle-points. (In the basic theory to be developed in this paper, no convexity assumptions will be made.)

If $\{y_i\}$ is any sequence generated by the algorithm corresponding to S, then

(3.4) all accumulation points will be fixed-points,

(3.5) $\phi(y_i) \to \phi(y^*)$, where y^* is a fixed-point,

(3.6) $||y_{i+1} - y_i|| \to 0$, and

(3.7) the accumulation points of $\{y_i\}$ form a continuum.

Proof: Because of the compactness hypothesis, any such sequence of iterates will have at least one accumulation point, which we denote as y^*. Suppose that y^* is not a fixed-point. Then there are subsequences of iterates $\{y_{n_i}\}$ and $\{y_{n_i+1}\}$ $(i = 0,1,2,\ldots)$ such that $y_{n_i} \to y^*$ and y_{n_i+1} converges to some point y'. By upper semi-continuity, $y' \in S(y^*)$ and by the strict monotonicity property, $\phi(y') < \phi(y^*)$. However, since for all i, $\phi(y_{i+1}) \leq \phi(y_i)$, we have $\lim \phi(y_{n_i}) = \phi(y^*) = \lim \phi(y_{n_i+1}) = \phi(y')$, a contradiction. Thus y^* must be a fixed point, and monotonicity of the sequence $\{\phi(y_i)\}$ implies convergence to $\phi(y^*)$. Suppose there existed a subsequence $\{y_{k_i}\}$ such that $||y_{k_i+1} - y_{k_i}|| > \delta > 0$ $(i = 0,1,2,\ldots)$. Without loss of generality we may assume that $y_{k_i} \to \bar{y}$ and $y_{k_i+1} \to \bar{\bar{y}}$. Note that $||\bar{y} - \bar{\bar{y}}|| \geq \delta$. However, $\bar{y}$ must be a fixed point of S, and, by u.s.c., $\bar{\bar{y}} \in S(\bar{y}) = \{\bar{y}\}$, a contradiction. Since $||y_{i+1} - y_i|| \to 0$, the accumulation points form a continuum by a result of Ostrowski [8]. ∎

A number of modifications in the hypotheses can be made that leave the conclusions of the theorem unchanged. Instead of (3.3) one could assume convergence of the sequence $\{\phi(y_i)\}$ and strict monotonicity and u.s.c. only at accumulation points, but most well-known algorithms have the strict monotonicity property on the entire feasible set. It is also possible to somewhat weaken the compactness hypothesis (3.1) by dealing with level sets. Thus, defining $L(c) = \{z \mid z \in F, \ \phi(z) \leq c\}$, we could assume (instead of (3.1)) that for every c , there was a compact subset $G(c)$ of F such that $S(z) \subset G(c)$ for all $z \in L(c)$. However, whenever this type of hypothesis is satisfied, it is possible to define a new mathematical programming problem that is equivalent to the original problem and for which hypothesis (3.1) will be satisfied in its original form.

Note that the theorem implies the existence of at least one fixed-point of S in H , and also that hypotheses (3.1) and (3.3) imply that the problem of minimizing ϕ over G has an optimal solution. For, if $\{z_i\} < G$ and $\phi(z_i) \rightarrow \inf_{z \in G} \phi(z)$, then $\{S(z_i)\} < H$, which is compact and by the strict monotonicity property we conclude $\inf_{z \in G} \phi(z) = \inf_{z \in H} \phi(z)$. Thus all z^* that solve the problem $\min_{z \in G} \phi(z)$ are fixed-points of S . For primal methods in which $f = \phi$ and $F = G$, it follows that all solutions of (MP) are fixed-points of S .

The hypotheses of this theorem differ from similar results in [6] and

[11] principally in that a fixed-point $y*$ has the property that $\{y*\} = S(y*)$ rather than $\{y*\} < S(y*)$ (Points satisfying the latter condition will be called "generalized" fixed-points.) The conclusions of Theorem 3.1 are stronger than those of earlier results precisely because of this more restrictive definition, which nevertheless is suitable for most algorithms that employ local searches. Basically, this approach rules out the "oscillatory" behavior that is possible when we allow $y*$ to be a proper subset of $S(y*)$. For example, let $G = \{y_1, y_2\}$ and let $S(y_1) = S(y_2) = G$ and $\phi(y_1) = \phi(y_2) = 0$. Then the sequence $\{y_1, y_2, y_1, y_2, y_1, y_2, \ldots\}$ could be generated by the corresponding algorithm. This sequence does not have properties (3.6) or (3.7) even though its convergent subsequences converge to the "generalized fixed-points" y_1 and y_2 . Although actual convergence of the full sequence of iterates of Theorem 3.1 does not follow from the hypotheses (see the Appendix for the "limit-cycle" type of non-convergence that may occur), a finiteness assumption on the number of fixed-points will guarantee convergence of those iterates. This is not the case, when "generalized" fixed-points are allowed, as the previous example shows.

Corollary 3.2 Let S satisfy (3.1), (3.2) and (3.3). If the number of fixed-points having any given value of ϕ is finite, then the algorithm corresponding to S will converge to a fixed-point of S regardless of where it is started in G .

Proof: (This result follows directly from the fact that the accumulation points are fixed-points and form a continuum. However, for the sake of completeness we give a short direct proof.)

Assume that the sequence $\{y_i\}$ does not converge, so that there are at least two accumulation points. By theorem 1, all accumulation points are fixed-points, so that given an accumulation point $y*$, there exists a $d > 0$ such that $y*$ is the unique accumulation point contained in an open ball of radius d centered at $y*$. Now let M be chosen such that $||y_{i+1} - y_i|| < d/3$ for $i \geq M$. Since $y*$ is an accumulation point, $||y_s - y*|| < d/3$ for infinitely many indices s. Since there is at least one accumulation point distinct from $y*$, for infinitely indices r it is true that $||y_r - y*|| > 2d/3$. From the way M was chosen, then, there are infinitely many indices $t \geq M$ such that $d/3 < ||y_t - y*|| < 2d/3$. This leads to a contradiction, since it implies the existence of a second accumulation point in the annulus $d/3 \leq ||y - y*|| \leq 2d/3$. (See Figure 1)

Example

As an easily-developed example of an algorithm to which this type of analysis can be applied, we shall consider a feasible direction algorithm proposed by Topkis and Veinott [10]. Suppose that the problem to be solved is of the form

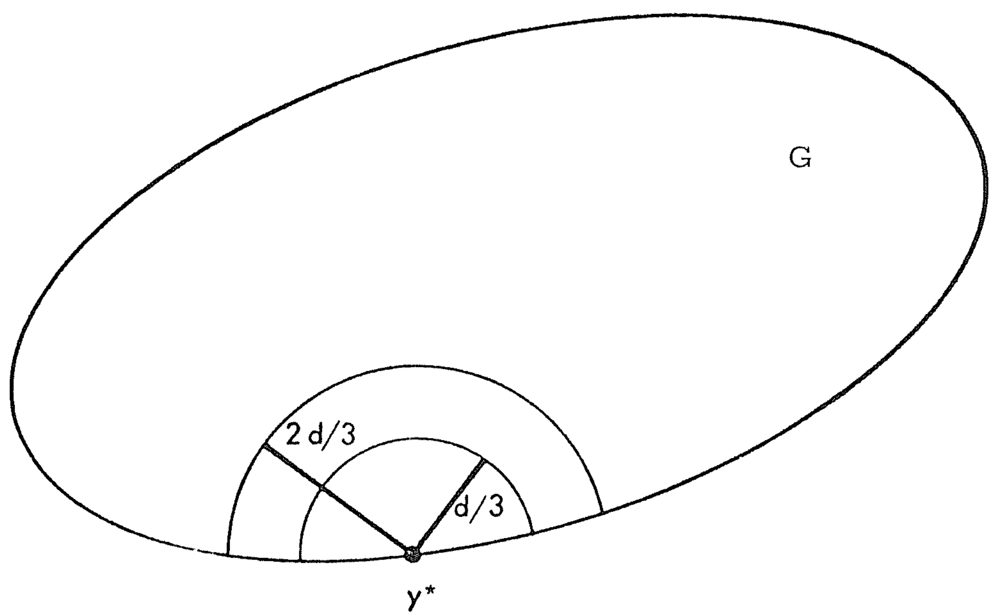

Figure 1.

$$\text{(NLP)} \qquad \begin{array}{l} \min f(x) \\ \text{s.t. } g(x) \le 0 \ , \end{array}$$

where f is a C^1 mapping from R^n to R^1 and g is a C^1 mapping from R^n to R^m. We let $F = G = H = \{x \mid g(x) \le 0\}$, and assume that this set is compact. An iteration starting at a point $x_i \in F$ consists of first finding an optimal solution (δ^*, d^*) of the approximating problem $AP(x_i)$ defined by

$$\text{(3.8)} \qquad \begin{array}{l} \min \delta \\ \text{s.t. } \nabla f(x_i)d + \frac{1}{2} dH_i d \le \delta \\ \qquad g(x_i) + \nabla g(x_i)d \le \delta e \ , \end{array}$$

where e is m-vector of 1's and H_i is a positive definite matrix with eigenvalues in a fixed interval $[M_1, M_2]$. If (d^*, δ^*) is an optimal solution of (3.8), then let K be a positive constant independent of i, let $\Lambda(x_i, d^*, \delta^*) \equiv \{\lambda \mid x_i + \lambda d^* \epsilon F \ , \ 0 \le \lambda \le K\delta^*\}$ and let x_{i+1} be any point of the form $x_{i+1} = x_i + \lambda_i \, d^*$, where $\lambda_i \in \{\lambda \mid \lambda \in \Lambda(x_i, d^*, \delta^*) \ , \ \rho \le \gamma_i(\lambda, d^*) \le 1-\rho\} \cup \Lambda^*(x_i, d^*, \delta^*)$, where $\gamma_i(\lambda, d^*) = (f(x_i) - f(x_i + \lambda d^*))/(-\nabla f(x_i)\lambda d^*)$, $\rho \in (0, \frac{1}{2}]$ is a positive constant independent of i, and $\Lambda^*(x_i, d^*, \delta^*)$ is the largest λ in $\Lambda(x_i, d^*, \delta^*)$ satisfying $\gamma_i(\lambda, d^*) \ge \rho$. Letting $S(x_i)$ be the set of all possible successors that may be obtained by this procedure, it may be shown that all the hypotheses of Theorem 3.1 are satisfied. In this case the fixed-points of S will be the points that satisfy the Fritz-John necessary optimality conditions conditions [4] for the original problem (NLP). If there are at most a finite number of Fritz-John points for (NLP), then we may conclude by Corollary 3.2 that the feasible direction algorithm above is globally convergent to a Fritz-John point. In the traditional approach used by Topkis and Veinott only the Fritz-John property of the accumulation points is established.

In order to condense the presentation, henceforth any mapping S that satisfies (3.1), (3.2), and (3.3) will be termed a <u>CUM mapping</u> (an acronym for uniformly <u>c</u>ompact, <u>u</u>pper semi-continuous, and strictly <u>m</u>onotonic). In this terminology, Corollary 3.2 says that any CUM mapping with a finite number of fixed-points is <u>globally convergent</u> to a fixed-point. Polak [9] states a result similar to Corollary 3.2 for algorithms allowing generalized fixed-points, but requires as an <u>assumption</u> that $||y_{i+1} - y_i|| \to 0$. As indicated by the example following Theorem 3.1, this assumption is readily violated when generalized fixed-points are allowed.

We state now for future reference an additional Corollary that may be proved in much the same manner:

<u>Corollary 3.3</u>: Let S be a CUM mapping, and let $\{y_i\}$ be a sequence obtained by the corresponding algorithm. If $\{y_i\}$ has an accumulation point $y*$ that is an <u>isolated</u> fixed-point of S (i.e., not the limit of a sequence of distinct fixed-points), then $\{y_i\}$ converges to $y*$.

The preceding theorem and its corollaries are <u>global</u> convergence results that require no hypotheses on the starting point y_0. In order to develop additional insight into the behavior of the algorithms considered here, we will establish a "point-of-attraction" result for "locally" CUM mappings after developing a useful preliminary lemma.

<u>Lemma 3.4</u>: Let $z*$ be a fixed-point that is also a strong local minimum of ϕ on G. Let S be strictly monotonic relative to ϕ on G. Assume that there exists an open set B containing $z*$ such that S is uniformly compact and strictly monotonic on $B \cap G$ ($\exists$ compact set H such that $S(z) \subset H$ for all $z \in B \cap G$). If S is u.s.c. at $z*$, then given any open set B_1 containing $z*$, there exists an open set B_2 containing $z*$ such that $y_0 \in B_2 \cap G$ implies $\{y_i\} \subset B_1 \cap G$.

<u>Proof:</u> Choose $\epsilon > 0$ such that $B_3 \equiv \{z| \ ||z-z*|| \leq \epsilon\}$ is a subset of B_1 and of B and has the property that $z \in B_3 \cap G$ implies $\phi(z) > \phi(z*)$. Choose $\epsilon' \leq \epsilon$ such that $B_4 \equiv \{z| \ ||z-z*|| < \epsilon'\}$ has the property that $z \in B_4 \cap G$ implies $S(z) \subset B_3$.

Next choose a $\lambda > \phi(z^*)$ such that $\phi(z) < \lambda$ and $z \in B_3$ imply $z \in B_4 \cap G$. Finally, let $B_2 \equiv \{z | \; ||z - z^*|| < \epsilon''\}$, where ϵ'' is chosen so that $B_2 \subset B_4$ and so that $z \in B_2 \cap G$ implies $\phi(z) < \lambda$. Clearly $y_0 \in B_2 \cap G$ implies $y_0 \in B_1 \cap G$ since $B_2 \subset B_4 \subset B_3 \subset B_1$. The proof will be completed by showing that if $y_i \in B_4 \cap G$, then $y_{i+1} \in B_4 \cap G \subset B_1 \cap G$. By construction $y_i \in B_4 \cap G$ implies $y_{i+1} \in B_3 \cap G$. By the monotonicity property $\phi(y_{i+1}) \leq \phi(y_0) < \lambda$, hence $y_{i+1} \in B_4 \cap G$. $\blacksquare$

A point-of-attraction theorem is now easily obtained for strong local minima of ϕ by a slight strengthening of the hypotheses.

<u>Theorem 3.5</u>: If the hypotheses of Lemma 3.4 are satisfied, and, if, in addition S is u.s.c. on B and z^* is an <u>isolated</u> fixed-point of S, then there exists an open set B^* such that $y_0 \in B^* \cap G$ implies $\{y_i\}$ converges to z^*.

<u>Proof</u>: Since z^* is isolated, choose B_1 to be an open set such that $z^* \in B_1$ and z^* is the only fixed-point in that the closure of $B_1 \cap G$, which in turn is contained in $B \cap G$. Now apply Lemma 3.4 and let $B^* = B_2$. By Theorem 3.1 and Lemma 3.4, all accumulation points of $\{y_i\}$ are fixed-points of S and lie in the closure of $B_1 \cap G$, hence

z^* is the unique accumulation point and thus the limit of the bounded sequence $\{y_i\}$. $\blacksquare$

Thus we have shown that any strong local minimum of ϕ on G is an "attractive" fixed-point of a CUM mapping S , provided that it is an isolated fixed-point of S . Note that Corollary 3.3 may be interpreted as saying that an *arbitrary* isolated fixed-point of S is "quasi-attractive" in the sense that if it is approached *arbitrarily closely* by the iterates, then the iterates will converge to it.

We will now show that both Corollary 3.3 and Theorem 3.5 are readily extended to mappings that allow *"generalized" fixed-points*, i.e., those $z^* \in G$ such that $S(z^*) \supset \{z^*\}$ (see Zangwill [11], Meyer [6], or Luenberger [3] for a discussion of the "subsequential" convergence properties of these algorithms). In thise case, instead of the strict monotonicity property, we assume the *"generalized" strict monotonicity property* at y: if $y' \in S(y)$, then $\phi(y') < \phi(y)$ if y is *not* a generalized fixed-point, and $\phi(y') \leq \phi(y)$ otherwise.

Theorem 3.6 Let S be a point-to-set mapping such that S is uniformly compact and u.s.c. on G , and has the *generalized* strict monotonicity property on G . Let $\{y_i\}$ be a sequence obtained by the corresponding algorithm. If $\{y_i\}$ has an accumulation point y^* that is an *isolated* generalized fixed-point of S satisfying $\{y^*\} = S(y^*)$, then $y_i \to y^*$. If, in addition, y^* is a strong local minimum of ϕ on G , then there exists an open neighborhood B of y^* such that if the algorithm is started in $B \cap G$, then the iterates will converge to y^* .

<u>Proof</u>: As shown in [6], the accumulation points of $\{y_i\}$ must be generalized fixed-points. Since $S(y^*) = \{y^*\}$ we can choose d^* such that the distance d from y^* to the nearest generalized fixed-point distinct from y^* satisfies $d \geq d^*$ and such that $||y_j - y^*|| < d^*/3$ implies $||y_{j+1} - y_j|| < d/3$. If we now assume that $y_i \not\to y^*$ we can establish a contradiction by demonstrating in a manner similar to the proof of Corollary 3.2 that there is a generalized fixed-point in the annulus $d/3 \leq ||y - y^*|| \leq 2d/3$.

If, in addition, y^* is a strong local minimum of ϕ , the point-of-attraction result is proved in a manner precisely analogous to the proof of Theorem 3.5. ∎

The preceding theorems and its analogs for CUM mappings may be thought as the mathematical programming analogs of Liapunov stability theory. This analogy for point-to-point mappings is described by Zangwill [11].

4. Restrictions and Relaxations of Point-to-Set Mappings

In many cases an algorithm is "derived" from a CUM mapping by introducing computationally expedient modifications. These modifications might consist of substituting "inexact" line searches such as those of Armijo [1] or Golstein [2] for "exact" line searches, the introduction of quasi-Newton steps to accelerate the convergence of steepest descent as in the Davidon-Fletcher-Powell method with restart [3], or the deletion of "inactive" constraints [5]. Such modifications may result in an algorithm that does not have the strict monotonicity or upper semi-continuity properties of the original algorithm, or for which upper semi-continuity may be very difficult to establish. Nevertheless, the resulting algorithms generally still have nice convergence properties, and in this section we will extend the theory of CUM mappings in order to establish these convergence properties. It will be seen that an appropriate extension of the theory is obtained by considering certain classes of restrictions and relaxations of CUM mappings: if T_1 and T_2 are point-to-set mappings from G to into the non-empty subsets of G such that $T_1(z) \supset T_2(z)$ for all $z \in G$ then T_1 will be said to be a relaxation of T_2 and, conversely, T_2 will be said to be a restriction of T_1.

We will first consider restrictions of CUM mappings, since these behave essentially the same as CUM mappings, although they may not be u.s.c.

Lemma 4.1 If T_2 is a restriction of a CUM mapping T_1, then T_2 is uniformly compact and strictly monotonic on G, and $\{z|z$ is a fixed-point of $T_1\} = \{z|z$ is a fixed-point of $T_2\}$. Moreover (3.4) - (3.7) hold for any sequence $\{y_i\}$ generated by the algorithm corresponding to T_2. If, in addition, the number of fixed-points having any given value of ψ is finite, then $\{y_i\}$ will converge to a fixed-point of T_2.

Proof: Suppose $F^* \equiv \{z|z$ is a fixed-point of $T_1\}$ and $z^* \in F^*$. Since $T_2(z^*)$ is non-empty and contained in $T_1(z^*)$, it follows that z^* is also a fixed-point of S. Conversely, if $T_2(z^*) = \{z^*\}$, then $z^* \in T_1(z^*)$ and by the strict monotonicity property $T_1(z^*) = \{z^*\}$. The strict monotonicity and compactness of T_2 follow directly from the definitions.

Finally, if $\{y_i\}$ was generated by the algorithm corresponding to T_2, then $\{y_i\}$ could also have been generated by the algorithm corresponding to T_1, hence (3.4) - (3.7) and the convergence conclusion also hold. ∎

Example:

It should be noted that a restriction of a CUM mapping need not be u.s.c. For example, suppose that instead of the Goldstein step-size procedure in the example following Corollary 3.2 we use a restriction corresponding to an Armijo-type step-size procedure: choose λ_i to be the first power of $\frac{1}{2}$ in $\Lambda(x_i, d^*, \delta^*)$ that satisfies $\gamma_i(\lambda_i, d^*) \geq \rho$. If at x_i

the latter inequality is satisfied as an equality, then the corresponding mapping may not be u.s.c. at x_i . $\blacksquare$

Two further observations are in order with regard to the u.s.c. properties of a restriction T_2 of a CUM mapping T_1 . First of all, T_2 must be u.s.c. at its <u>fixed-points.</u> For if we assume otherwise, we can obtain a contradiction by using the uniform compactness of T_2. Secondly, even if the restriction T_2 of a CUM mapping is not u.s.c. at a point $\bar{z}$ (which cannot be a fixed-point of T_2 by the preceding observation), T_2 must satisfy the following <u>sequential strict monotonicity property</u> at $\bar{z}$: if $\bar{z}_i \to \bar{z}$ and $z_i^* \to z^*$ with $z_i^* \in T_2(\bar{z}_i)$ for all i , then $\phi(z^*) < \phi(\bar{z})$. (Note that we do <u>not</u> assume that $z^* \in T_2(\bar{z})$.) It should be observed that the sequential strict monotonicity property at $\bar{z}$ implies the strict monotonicity property at $\bar{z}$, but not vice-versa. Moreover, by the same reasoning, the sequential strict monotonicity property <u>cannot</u> hold at a fixed-point. The sequential strict monotonicity property is of interest because, as will be seen, it furnishes us with means of determining if an algorithm is a restriction of a CUM mapping. This determination might otherwise be difficult to perform if the CUM relaxation were not known beforehand.

Consider a mapping S that is uniformly compact and strictly monotonic on G , but may not be u.s.c. on all of G . Let $G' \equiv \{z | z \in G ,$ S is not u.s.c. at $z\}$ be the subset of G on which we do <u>not</u> have the u.s.c. property. (If G' is empty, then S is a CUM mapping, so the main interest is in the case when G' is non-empty.) If S is sequentially strictly monotonic on G', we will show that all the convergence properties of a CUM mapping hold for S, because S can be shown to be a restriction of a CUM mapping. In particular we define the completion of S to be the point-to-set mapping $\hat{S}(z) \equiv \{z' \mid \exists \{z_i\} \to z, \{z_i'\} \to z',$ where $z_i' \in S(z_i)\}$. Note that $S(z)$ thus is a restriction of $\hat{S}(z)$, and that $\hat{S}(z) = S(z)$ if and only if S is u.s.c. at z.

Example:

In Mangasarian [5] it is shown that it is sufficient for convergence purposes to consider only the "ϵ-active" constraints in constructing the subproblem (3.8). That is, rather than considering all the constraints of (NLP) in constructing $AP(x_i)$ we need only consider those constraints satisfying $-\epsilon \leq g_j(x_i) \leq 0$, where ϵ is a positive constant independent of i. This device, however, may cause the corresponding mapping to fail to be u.s.c. at those x_i which are not fixed-points and for which $g_j(x_i) = -\epsilon$ for some j, since the corresponding constraint will not be included in the subproblem for x near x_i for which $g_j(x) < -\epsilon$. However, in this case it is easily shown the sequential strict monotonicity property will hold at such x_i, so that global convergence is assured by

Lemma 4.1. In this case the completion of the corresponding point-to-set mapping has the property that at x_i a finite <u>family</u> of subproblems may be considered, which would include both the subproblems in which g_j is considered and in which g_j is not considered, provided that $g_j(x_i) = -\epsilon$ and that x_i is the limit of a sequence $\{z_1, z_2, \ldots\} < G$ such that $g_j(z_k) < -\epsilon$.

<u>Theorem 4.2</u> If a point-to-set mapping S is strictly monotonic on G and sequentially strictly monotonic on G' , then its completion $\hat{S}$ is strictly monotonic and u.s.c. on G .

<u>Proof:</u> Let $\bar{z}$ be an arbitrary point in G . If $\hat{S}(\bar{z}) = S(\bar{z})$, then clearly $\hat{S}$ is strictly monotonic at $\bar{z}$; otherwise, by the <u>sequential</u> strict monotonicity property, it follows that $\hat{S}$ is strictly monotonic at $\bar{z}$. Let $\bar{z}_i \to \bar{z}$ and $z^*_i \to z^*$, where $z^*_i \in \hat{S}(\bar{z}_i)$, and let $\{\bar{z}_{i,j}\}$ and $\{z^*_{i,j}\}$ be sequences such that $\lim_{j\to\infty} \bar{z}_{i,j} = \bar{z}_i$ and $\lim_{j\to\infty} z^*_{i,j} = z^*_i$, where $z^*_{i,j} \in S(\bar{z}_{i,j})$. By a result of Meyer [6, Appendix], there exist subsequences $\{\bar{z}_{i(j),j}\}$ and $\{z^*_{i(j),j}\}$ such that $\lim_j \bar{z}_{i(j),j} = \bar{z}$ and $\lim_j z^*_{i(j),j} = z^*$. From the definition of $\hat{S}(\bar{z})$ it follows that $z^* \in \hat{S}(\bar{z})$, and we have thus shown that $\hat{S}$ is u.s.c. at an arbitrary $\bar{z} \in G$. ∎

We can summarize our results with respect to restrictions as follows: (1) restrictions of CUM mappings have the same convergence properties as CUM mappings; (2) a mapping T_2 is the restriction of a CUM mapping T_1 <u>if and only if</u> T_2 is uniformly compact and strictly monotonic on G, and sequentially strict monotonic on G'.

As a final example of restriction, we shall consider a technique for obtaining a CUM restriction of a uniformly compact, u.s.c. mapping that has the generalized strict monotonicity property.

<u>Theorem 4.3</u>: Let S be a point-to-set mapping that is uniformly compact and u.s.c. on G, and that satisfies the generalized strict monotonicity property on G. If S is also <u>lower</u> semi-continuous at its generalized fixed-points, then there exists a restriction of S that is a CUM mapping.

<u>Proof:</u> For all $z \in G$, $S(z)$ is closed because of the u.s.c. property. Let δ be a metric on G and define $S^*(z) = \{z' \mid z' \in S(z)$, $\delta(z',z) \leq \delta(y',z)$ for all $y' \in S(z)\}$. Note that if z^* is a generalized fixed-point of S, l.s.c. of S at z^* implies that if $\{z_i\} \subset G$ and $z_i \to z^*$, then there exists a sequence $\{z_i'\}$ with $z_i' \in S(z_i)$ and $z_i' \to z^*$, so that $\delta(z_i', z_i) \to 0$. From this observation it follows that if $z_i^* \in S^*(z_i)$ and $z_i^* \to \bar{z}$, then $\bar{z} = z^*$. Since $S^*(z^*) = \{z^*\}$, we conclude that S^* is u.s.c. at the generalized fixed-points of S, which, in turn, are fixed-

points of S^*. We will now take care of the remaining points of G by showing that the sequential strict monotonicity property holds for S^* at any point $\hat{z}$ that is <u>not</u> a generalized fixed-point of S. If $\hat{z}_i \to \hat{z}$ and $\tilde{z}_i \to \tilde{z}$ with $\tilde{z}_i \in S^*(\hat{z}_i)$, then by the u.s.c. of S we have $\tilde{z} \in S(\hat{z})$, and by generalized strict monotonicity $\phi(\tilde{z}) < \phi(\hat{z})$, proving the sequential strict monotonicity for S^* at $\hat{z}$. Now by applying Theorem 4.2, we conclude that the completion $\hat{S}^*$ of S^* is a CUM mapping, and, moreover, by the preceding proof of sequential strict monotonicity, it follows that $\hat{S}^*$ is also a restriction of S. ∎

For monotonic algorithms which may be thought of as employing CUM mappings <u>periodically</u> rather than at every iteration (such as, for example, the "restarted" quasi-Newton or conjugate gradient methods, in which periodically a search is made in the gradient direction), the theorems for CUM mappings may be extended appropriately by considering <u>relaxations</u> of CUM mappings. These extensions of the theory and related results may be found in Meyer [7].

APPENDIX

Example: Let D' be the sequence $\{y_1, y_2, \ldots\}$ in E^2 defined in the following manner: $y_1 = (0,2)$; given y_n , let y_{n+1} be the point on the circle with center at the origin and radius $1 + 1/n+1$ such that (1) the line segment connecting y_n and y_{n+1} is tangent to that circle, and (2) the movement in going from y_n to y_{n+1} has a clockwise orientation (see Figure 2). It can be shown that the set of accumulation points of D' is the unit circle. (This is essentially a consequence of the geometry and the fact that the distance between y_n and y_{n+1} is $[(2n^2 + 4n + 1)/(n^4 + 2n^3 + n^2)]^{1/2}$, which is of the order of $1/n$.)

Now let D be the closure of D' , which is the union of D' and the unit circle. Define a mapping S on D as follows:

(1) $S(y_n) = \{y_{n+1}\}$ $n = 1, 2, \ldots$

(2) $S(y) = \{y\}$ for y on the unit circle

With $\phi(z)$ defined as the distance from the origin, and F taken to be D it is easily verified that S satisfies the hypotheses of theorem 3.1. However, if any point of D' is chosen as a starting point, the corresponding sequence of iterates does not converge, but rather "spirals" around the unit circle. (This example is derived from a similar example given by Topkis and Veinott [10] to illustrate a different observation.)

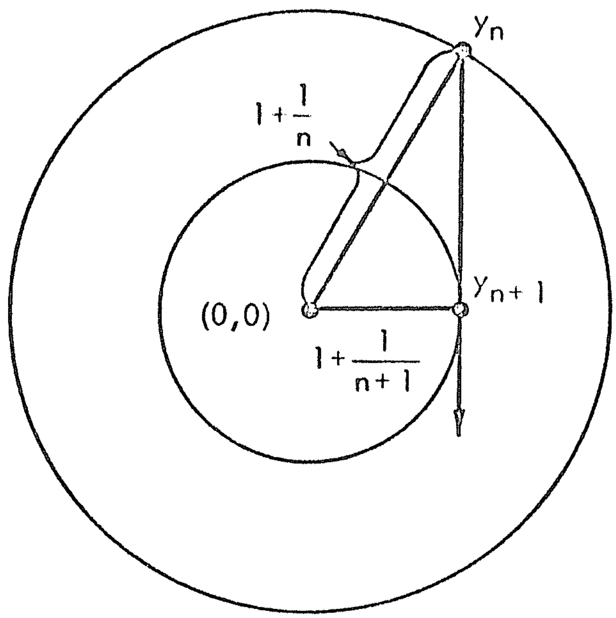

Figure 2.

ACKNOWLEDGEMENT

The author is indebted to Professor S. M. Robinson for pointing out the possibility of using a restriction procedure similar to that used in the proof of Theorem 4.3 in order to obtain a strictly monotonic algorithm from an algorithm with the generalized strict monotonicity property.

REFERENCES

1. L. Armijo, "Minimization of functions having Lipschitz continuous first partial derivatives", Pacific J. Math., 16 (1966), pp. 1-3.

2. A. A. Goldstein, Constructive Real Analysis, Harper and Row, New York, 1967.

3. D. G. Luenberger, Introduction to Linear and Nonlinear Programming, Addison-Wesley, Reading, Mass., 1973.

4. O. L. Mangasarian, Nonlinear Programming, McGraw-Hill, New York, 1969.

5. O. L. Mangasarian, "Dual, Feasible Direction Algorithms", in Techniques of Optimization, ed by A. V. Balakrishnan, Academic Press, New York, 1972.

6. R. R. Meyer, "The Validity of a Family of Optimization Methods", SIAM J. Control, 8 (1970), pp. 41-54.

7. R. R. Meyer, "On the Convergence of Algorithms with Restart", forthcoming University of Wisconsin Computer Sciences Department Technical Report.

8. A. M. Ostrowski, Solution of Equations and Systems of Equations, Academic Press, New York, 1966.

9. E. Polak, Computational Methods in Optimization, Academic Press, New York, 1971.

10. D. M. Topkis and A. F. Veinott, "On the Convergence of Some Feasible Direction Algorithms for Nonlinear Programming," SIAM J. Control, 5 (1967), pp. 268-279.

11. W. I. Zangwill, Nonlinear Programming, Prentice-Hall, Englewood Cliffs, N. J., 1969.