WIS-CS-74-225
COMPUTER SCIENCES DEPARTMENT
The University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53706

Received September 4, 1974

ON THE CONVERGENCE OF ALGORITHMS WITH RESTART

by

R. R. Meyer

Technical Report #225

October 1974

Errata: TR #225

p. 5, line 9: l.c. "s" should be cap "S".

p. 14: The mapping $S^{(j)}$ of the Example may be explicitly defined as follows:

$$S^{(j)}(y_i) \equiv \{z \mid z \in \bar{S}^{(j)}(y_i),\ f(z) \leq f(x)\ \forall\ x \in \bar{S}^{(j)}(y_i)\},$$

where $\bar{S}^{(j)}(y_i) \equiv \{x \mid x = y_i - \lambda \nabla f(y_i)^T e_{j+1},\ 0 \leq \lambda \leq 1\}$,

e_{j+1} being the (j+1) st unit vector.

p. 17, line 4: "$S^{(1)}$" should be replaced by "$S^{(1)}(z)$".

ON THE CONVERGENCE OF ALGORITHMS WITH RESTART

by

R. R. Meyer

ABSTRACT

Global convergence properties are established for a class of point-to-set mathematical programming algorithms commonly termed "restart" methods. Well-known algorithms in this class include the "restart" versions of the Fletcher-Reeves conjugate gradient and Davidon-Fletcher-Powell methods. Under certain mild assumptions, it is shown that the entire sequence of iterates (as opposed to selected subsequences) generated by such algorithms converges to a "desirable" point. Some similar convergence results are also established for a related class of "inexact" algorithms, and for a class of algorithms motivated by cyclic coordinate descent methods.

1. Introduction

In Meyer [9] we introduced the concepts and general properties of restrictions and relaxations of point-to-set mappings associated with monotonic mathematical programming algorithms. (A monotonic algorithm is one for which there exists a continuous function ϕ, such that $\phi(x_{i+1}) \leq \phi(x_i)$ for successive iterates x_i and x_{i+1}. If S and S' are point-to-set mappings from a set G into its non-empty subsets, and $S(z) \subset S'(z)$ for all $z \in G$, then S is said to be a restriction of S', and S' is a relaxation of S.) All of the point-to-set mappings considered in this paper will be assumed to have a set $G \subset R^n$ as their domain, and the non-empty subsets of G as their range. Here we shall consider certain types of restrictions and relaxations of point-to-set mappings that are relevant to the study of the convergence properties of "restart" algorithms, i.e., algorithms in which, from time to time, a certain "basic" type of iteration is used and information other than the location of the last iterate is ignored. The "restart" versions of the conjugate gradient [2] and Davidon-Fletcher-Powell methods [7] are familiar examples of this type of procedure.

It will be also shown that relaxations are related to the convergence analysis of certain "inexact" algorithms. Under relatively weak assumptions we will prove global convergence of the entire sequence of iterates (rather than subsequential convergence) to a "desirable point". Finally, similar ideas are used to establish sufficient conditions for global convergence of a class of algorithms that generalizes cyclic coordinate descent methods. Related results for algorithms based on one-dimensional searches may be found in Ortega and Rheinboldt [10].

2. The Relaxation Approach

In this approach to convergence analysis, we are given a monotonic point-to-set mapping S that has certain "nice" properties, (to be described below) and we derive from it a related mapping S_α defined by

$$S_\alpha(z) \equiv \{z' | \phi(z)-\phi(z') \geq \alpha(z)(\phi(z)-\phi(\tilde{z})) \text{ for some } \tilde{z} \in S(z)\},$$

where $\alpha(z)$ is a lower semi-continuous positive function on G. In particular, if $\alpha(z) \leq 1$ for all z, then S_α will be a relaxation of S. (In most cases of interest $\alpha(z) \leq 1$. As will be seen below, the case $\alpha(z) \equiv 1$ is of interest in analyzing "restart" methods, whereas the cases $\alpha(z) \equiv \gamma < 1$ may be thought of as an "error tolerance" criterion for the amount of reduction in the function value. Rather more practical criteria for inexact line searches for variable metric and conjugate direction algorithms may be found in Lenard [4], [5] and Ritter [12].) Under some mild assumptions we will establish a global convergence result for S_α, but first we introduce some convenient terminology.

If T is a point-to-set mapping, and $T(z^*) \supset \{z^*\}$, then z^* is said to be a generalized fixed-point of T. If there exists a real-valued function ϕ continuous on G such that $\bar{z} \in T(z)$ implies $\phi(\bar{z}) \leq \phi(z)$, with strict inequality holding if z is not a generalized fixed-point, then T is said to have the generalized strict monotonicity property at z (with respect to ϕ). If there exists a compact set H such that $T(z) \subset H$ for all $z \in G$, then

T is said to be uniformly compact on G. A point-to-set mapping that that is u.s.c. and uniformly compact on G, and has the generalized strict monotonicity property on G will be denoted as a CUGM mapping (uniformly compact, upper semi-continuous, generalized strict monotonicity). The algorithm that starts at an arbitrary $z_0 \in G$ and chooses $z_{i+1} \in T(z_i)$. $(i=0,1,2,\ldots)$ will be said to be the algorithm corresponding to T. Observe that if T is CUGM mapping and $\{z_i\}$ is a sequence generated by the corresponding algorithm, then every accumulation point of $\{z_i\}$ will be a generalized fixed-point of T. (This result is easily proved by contradiction since $\phi(y_i)$ must converge.)

Theorem 2.1: Let S be a CUGM mapping. If $\alpha(z) \leq 1$ for all z, and the relaxation S is uniformly compact on G, then the accumulation points of any sequence $\{y_i\}$ generated by any algorithm corresponding to S_α will be generalized fixed-points of S, and $\phi(y_i) \to \phi(y^*)$, where y^* is a generalized fixed-point of S.

Proof: Note first of all that S_α is also u.s.c. on G, since by the u.s.c. and uniform compactness of S we can show that $z_i' \in S_\alpha(z_i)$, $z_i \to z$, and $z_i' \to z'$ imply the existence of a $\tilde{z} \in S(z)$ such that $\phi(z) - \phi(z') \geq \alpha(z)(\phi(z) - \phi(\tilde{z}))$. Since it is easily seen that S_α has the generalized strict monotonicity property and that the generalized fixed-points of S_α are exactly the generalized fixed-points of S, the conclusion of the theorem thus follows from the observation preceding the theorem. ∎

A theorem analogous to Theorem 2.1 is easily derived for monotonic "restart" algorithms that employ a CUGM mapping infinitely often.

Theorem 2.2: Let $\{y_i\}$ be a sequence of points with the following properties:

(1) there exists a subsequence $\{y_{n_i}\}$ such that $y_{n_i+1} \in S(y_{n_i})$, where S is a CUGM mapping and (2) $\phi(y_{i+1}) \leq \phi(y_i)$ for all i. Then all of the accumulation points (if any exist) of $\{y_{n_i}\}$ are generalized fixed-points of s. If, in addition, $\{y_{n_i}\}$ has an accumulation point, then $\phi(y_i) \to \phi(y^*)$, where y^* is a generalized fixed-point of S.

Proof: Let y^* be the limit of a subsequence $\{y_{m_i}\}$ of $\{y_{n_i}\}$. We will show that y^* is a generalized fixed-point of S. Since S is uniformly compact, $\{y_{m_i+1}\}$ is contained in a compact set, and we may assume without loss of generality that $y_{m_i+1} \to y^{**}$. But $\phi(y^{**}) = \phi(y^*) = \lim \phi(y_i)$ and $y^{**} \in S(y^*)$, so y^* is indeed a generalized fixed-point of S. ∎

Another way of looking at Theorem 2.2 is that $y_{n_{i+1}} \in S_1(y_{n_i})$, where S_1 is the mapping derived from S_α by taking $\alpha \equiv 1$. Since, in effect, no compactness assumptions are made about S_1 in Theorem 2.2, the conclusions of the theorem are slightly weaker than those of Theorem 2.1. However, under appropriate compactness assumptions, the

convergence properties of $\{y_{n_i}\}$ would follow from the latter theorem.

(This is the only theorem in this paper in which the hypotheses do not guarantee that the iterates will be contained in a compact set. In this instance it is possible that there may be no accumulation points of $\{y_{n_i}\}$ as is shown by Example 1 in the Appendix. As the example further shows, it may then be the case that $\lim \phi(y_i)$ converges to something other than the value of ϕ at a generalized fixed-point. This behavior would be impossible if it were assumed that $\{y_{n_i}\}$ was contained in a compact set.)

The approach used in Theorem 2.2, which is similar to results of Zangwill [13] and Luenberger [16], has the disadvantage that it only deals with the behavior of a particular subsequence of the entire sequence of iterates. Nothing is said about the convergence properties of other subsequences or of the whole sequence except that $\phi(y_i) \to \phi(y^*)$. This is a very weak form of convergence, as the following example shows.

Let $G = [0,3]$,

$$S(x) \equiv \begin{cases} \{\frac{1}{2}x\} & \text{if } 0 \leq x < 1 \\ \{3\} & \text{if } x > 1 \\ \{\frac{1}{2},3\} & \text{if } x = 1, \end{cases}$$

$$\phi(x) \equiv \begin{cases} x & \text{if } x \leq 1 \\ 2 - x & \text{if } x > 1, \end{cases}$$

$\{n_i\} = \{3i \mid i=0,1,2,\ldots\}$. Then the sequence $\{y_i\}$ with $y_0 = 1$, $y_{3_i+1} = y_{3_i+3} = 2^{-i-1}$, $y_{3_i+2} = 2 - y_{3_i+1}$ $(i=0,1,2,\ldots)$, satisfies the hypotheses of Theorem 2.2, but does not converge, and, in fact, contains a subsequence converging to 2, which is not a generalized fixed-point of S. (See the Appendix for an example of non-convergence in which the accumulation points form a continuum.)

In order to establish stronger convergence properties, we must add to the hypotheses of the previous theorems. Before doing so, we will introduce some additional terminology. A point-to-set mapping S will be said to have the strict monotonicity property if $z' \in S(z)$ implies $\phi(z') < \phi(z)$ whenever $S(z) \neq \{z\}$ (those points z^* such that $S(z^*) = \{z^*\}$ will be called fixed-points of S). (For the convergence properties of CUGM mappings with the strict monotonicity property see Meyer [9].)

Theorem 2.3: Suppose that the hypotheses of Theorem 2.1 are satisfied and that S^* is a restriction of S_α such that S^* is u.s.c. at the fixed-points of S. If the fixed-points of S are fixed-points of S^* and S satisfies the strict monotonicity property, then a sequence $\{y_i\}$ generated by the algorithm corresponding to S^* will have the properties:

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

(2) the accumulation points of $\{y_i\}$ will form a continuum contained in the set of fixed-points of S if $\{y_i\}$ does not converge to a fixed-point of S.

If, in addition, the number of fixed-points of S having any given value of ϕ is finite, then $\{y_i\}$ converges to a fixed-point of S.

Proof: Since S has the strict monotonicity property, all its generalized fixed-points must be fixed-points. By Theorem 2.1, then, all accumulation points of $\{y_i\}$ must be fixed-points of S. Suppose there was a subsequence $\{y_{n_i}\}$ such that for some positive δ, $\|y_{n_i+1}-y_{n_i}\| \geq \delta$. Because of the uniform compactness of S^* we can, without loss of generality, assume that $y_{n_i} \to y^*$ and that $y_{n_i+1} \to y^{**}$. But y^* is a fixed-point of S, and therefore of S^* also, so by the u.s.c. of S^* at y^*, $y^{**} = y^*$, contradicting $\|y^*-y^{**}\| \geq \delta$. The accumulation points of $\{y_i\}$ thus must form a continuum by a theorem of Ostrowski [11] if $\{y_i\}$ does not converge. ∎

In order to show that S^* is u.s.c. at the fixed-points of S, it is sufficient to show that if z' is "near" a fixed-point z^* then the set $S^*(z')$ is also "near" z^*. For mappings S such that the fixed-points satisfy a first-order condition such as the vanishing of a gradient, this is often easy to do, as is illustrated by the example following Theorem 3.1 below.

For algorithms with periodic restart an analogous strengthening of Theorem 2.2 is possible. (An example is provided in the Appendix to illustrate the difficulties that may arise when the restarts are aperiodic.) This result and a related result for "cyclic" algorithms are described in the next section.

3. Monotonic Algorithms with "Periodic Restart"

"Restart" versions of the conjugate-gradient and quasi-Newton methods, in which the algorithm is periodically restarted by ignoring information from previous steps and simply taking a "steepest descent" step are frequently considered in the literature. In this section we will consider the analogs of Theorem 2.2 for monotonic algorithms with periodic restart and for purely cyclic algorithms (e.g., cyclic coordinate descent methods). It will be seen that, under reasonable hypotheses on the mappings involved, convergence of the entire sequence of iterates is guaranteed.

Theorem 3.1: Let $S^{(0)}$ be a CUGM mapping that is strictly monotonic on G, and assume that the number of fixed-points of $S^{(0)}$ having any given value of ϕ is finite. Let $S^{(1)}, \ldots, S^{(k-1)}$ be uniformly compact mappings such that for $j = 1, \ldots, k-1$:

(1) if $\bar{z} \in S^{(j)}(z)$, then $\phi(\bar{z}) \leq \phi(z)$,

(2) the fixed-points of $S^{(0)}$ are fixed-points of $S^{(j)}$, and

(3) $S^{(j)}$ is u.s.c. at the fixed-points of $S^{(0)}$.

Let $\{y_i\}$ be a sequence of points generated by the following algorithm:

(a) choose y_0 arbitrarily from G,

(b) given y_i, where $i = sk + r$ (s a non-negative integer and $0 \leq r \leq k-1$), choose y_{i+1} from $S^{(r)}(y_i)$.

Then $\{y_i\}$ converges to a fixed-point of $S^{(0)}$.

Proof: We will first show that every accumulation point y^* of $\{y_i\}$ is a fixed-point of $S^{(0)}$. Let $y_{n_i} \to y^*$ and denote the set $\{y_0, y_k, y_{2k}, \ldots\}$ as Y_k. Without loss of generality we may assume that there exists a convergent subsequence $\{y_{m_i}\}$ of Y_k such that $m_i \leq n_i < m_i + k$. Let y^{**} be the limit of $\{y_{m_i}\}$. We will show that $y^{**} = y^*$. Without loss of generality, we may assume that $\{y_{m_i+j}\}$ is convergent with limit $y^{(j+1)}$ for $j = 0, \ldots, k-1$. By the u.s.c. of $S^{(0)}$ at y^{**}, it follows that $y^{(1)} \in S^{(0)}(y^{**})$, but since $\phi(y^{(j)}) = \phi(y^{**})$ $(j=1,\ldots,k)$, the strict monotonicity of $S^{(0)}$ implies $y^{(1)} = y^{**}$. Thus y^{**} is a fixed-point of $S^{(0)}$ and $y^{(1)} = y^{**}$ is a fixed-point of $S^{(1)}$. By the u.s.c. of $S^{(1)}$ at $y^{(1)}$ it follows that $y^{(2)} = y^{(1)} = y^{**}$ and, by induction, we conclude that $y^{(j)} = y^{**}$ for $j = 1, \ldots, k$. Since all k subsequences $\{y_{m_i+j}\}$ $(j=0,\ldots,k-1)$ converge to y^{**}, it follows that $\{y_{n_i}\}$ converges to y^{**} also; hence $y^* = y^{**}$, a fixed-point of $S^{(0)}$.

It is now easily shown by an argument similar to that used in the proof of Theorem 2.3 that $\|y_{i+1} - y_i\| \to 0$, and since the accumulation points cannot form a continuum, $\{y_i\}$ must converge to a fixed-point of $S^{(0)}$.

Example: For an example to which the following theorem will be applicable, we will consider the solution of the problem $\min_{x \in \mathbb{R}^n} f(x)$

by the conjugate gradient method of Mc Cormick and Ritter [8] with a modified step-size. We assume that $\phi = f$, where f is continuously differentiable on $\mathbb{R}^n$. In this method the search direction d_{j+1} at iteration j $(j=0,1,2,\ldots)$ is chosen by the rule

$$d_j = \begin{cases} g_j + \gamma_{j-1} d_{j-1} & \text{if } \|\gamma_{j-1} d_{j-1}\| < \beta \|g_j\| \text{ and } j \neq mn \ (m=0,1,\ldots) \\ g_j & \text{otherwise,} \end{cases}$$

where $g_j = \nabla f(x_j)$ and

$$\gamma_{j-1} = \frac{\|g_j\|^2 - g_j' g_{j-1}}{\|g_{j-1}\|^2},$$

and β is a given positive constant. For step-size selection, we will assume that, given x_j and d_j, x_{j+1} is chosen to minimize $f(x)$ on the set $\{x \mid x = x_j - \sigma d_j, 0 < \sigma \leq K\}$ where K is a given positive constant. (More practical step-size algorithms of the Armijo [1] or Goldstein [3] variety may also be used, but this choice illustrates the underlying idea that applies equally well to all step-size procedures with a fixed upper bound.) Now let y_0 be given and assume that $L(y_0) \equiv \{x \mid f(x) \leq f(y_0)\}$ is compact, and let G be the union of the sequence $\{y_i\}$ and its accumulation points. Since G is contained in $L(y_0)$, G must be compact. Now define $S^{(r)} = \{y_{i+1}\}$ for $i = sn + r$ $(s=0,1,\ldots$ and $r=0,\ldots,n-1)$ and if y^* is an accumulation point of $\{y_i\}$ define $S^{(r)}(y^*)$ to be the set $\{y^*\}$. It is well-known that $S^{(0)}$ is a strictly monotonic CUGM mapping. We will now show that hypotheses (2) and (3) of

Theorem 3.1 are also satisfied. If y' is a fixed-point of $S^{(0)}$, then clearly $\nabla f(y') = 0$. Hence, if $y' = y_i$ for some i, then hypotheses (2) and (3) are trivially satisfied. Otherwise, y' will be an accumulation point of $\{y_i\}$, but then $y_{n_i} \to y'$ implies $\nabla f(y_{n_i}) \to 0$ and thus $d_{n_i} \to 0$. Because of the step-size procedure, it follows that if $\{y_{n_i+1}\} = S^{(r)}(y_{n_i})$ for any fixed r, then $y_{n_i+1} \to y'$. Hence hypotheses (2) and (3) are satisfied in this case also. Finally, if $L(y_0)$ or G contain at most a finite number of points at which ∇f vanishes, the preceding theorem guarantees convergence of $\{y_i\}$ to a point at which ∇f vanishes. If it is also assumed that f is pseudo-convex, then such a point will be the global minimum of $f(x)$. ∎

As usual, if we do not assume that the set of fixed-points of $S^{(0)}$ is finite, we may still conclude that $\|y_{i+1} - y_i\| \to 0$ and that all the accumulation points of $\{y_i\}$ are fixed-points of $S^{(0)}$. It should be pointed out also that the conclusion of Theorem 3.1 does not hold if the periodicity properties of the algorithm are not assumed. This is illustrated by Example 2 in the Appendix, which exhibits the possible cumulative effect of "bad behavior" at intermediate iterations when the gaps between applications of the "good" mapping $S^{(0)}$ are allowed to become arbitrarily large. Of course, Theorem 2.2 continues to apply in Example 2, but again the behavior of the entire sequence $\{y_i\}$ is quite distinct from the behavior of $\{y_{n_i}\}$.

It should be noted that Theorem 3.1 may be generalized by assuming that, instead of being a strictly monotonic CUGM mapping, $S^{(0)}$ is uniformly compact, u.s.c at its fixed-points, and has the <u>sequential strict monotonicity property</u> (Meyer, [9]) at the remaining points of G (i.e., if $x_i' \in S^{(0)}(x_i)$ with $x_i \to x^*$, $x_i' \to x'$, then $\phi(x') < \phi(x^*)$). These weaker assumptions are easily seen to be precisely those needed for the proof of the theorem. A similar observation applies to the next theorem, which is motivated by <u>cyclic coordinate descent methods</u>. (A related result for cyclic algorithms based on <u>one-dimensional searches</u> may be found in Ortega and Rheinboldt [10].)

<u>Theorem 3.2</u>: Let $S, S^{(0)}, \ldots, S^{(k-1)}$ be CUGM mappings that are strictly monotonic on G (with respect to the same ϕ) and assume that the number of fixed-points of S having any given value of ϕ is finite. Assume in addition that if a point $z' \in G$ is <u>not</u> a fixed-point of S, then for at least one j, z' is <u>not</u> a fixed-point of $S^{(j)}$. If $\{y_i\}$ is generated as in Theorem 3.1, then $\{y_i\}$ converges to a fixed-point of S.

<u>Proof</u>: Let $\{y_{n_i}\}$ be a convergent subsequence of $\{y_i\}$ with limit y^*. Without loss of generality we may assume that for some fixed j $y_{n_i+1} \in S^{(j)}(y_{n_i})$, and that $y_{n_i+1} \to y^{**}$. Since $\phi(y^*) = \phi(y^{**})$ and $y^{**} \in S^{(j)}(y^*)$, we conclude that $y^* = y^{**}$ and $y_{n_i+1} - y_{n_i} \to 0$. By repeating the argument for $\{y_{n_i+1}\}, \ldots, \{y_{n_i+k-1}\}$ we can show

that y^* is a fixed-point of <u>all</u> the $S^{(j)}$. Hence every accumulation point of $\{y_i\}$ is a fixed-point of S, and since, by the usual argument, $y_{i+1} - y_i \to 0$, we conclude that $\{y_i\}$ must converge to a fixed-point of S. ∎

<u>Example</u>: Let f be a mapping from $\mathbb{R}^k$ into $\mathbb{R}^1$, and let S be the variant of steepest descent in which, given y_i, y_{i+1} is a point minimizing f on the set $\{z | z=y_i+\lambda\nabla f(y_i),\ 0\leq\lambda\leq 1\}$. If $S^{(j)}$ is the mapping associated with finding the minimum (within a unit step) in the direction of the $(j+1)$st unit vector vector; and if the compactness conditions are satisfied, then the preceding theorem establishes the global convergence of the corresponding cyclic coordinate descent algorithm.

APPENDIX

Example 1: Let $G = \{x \mid x \geq -1\}$,

$$\phi(x) = \begin{cases} e^{-x} & \text{for } x \geq 1 \\ e^{-1}x & \text{for } -1 \leq x \leq 1, \end{cases}$$

$$S(x) = \begin{cases} \{e^{-x}\} & \text{for } x \geq e \\ \{e^{-x} - (e-x)\} & \text{for } x < e, \end{cases}$$

$y_0 = y_1 = y_2 = y_3 = 3$, and $\{n_i\} = \{3, 5, 7, \ldots\}$. Taking $y_{n_i} = n_i$, all of the hypotheses of Theorem 2.2 are satisfied with $\phi(y_{n_i}) = e^{-n_i}$, $y_{n_i+1} = e^{-n_i}$, and $\phi(y_{n_i+1}) = e^{-n_i+1}$. The only fixed-point of S is -1, and $\phi(-1) = -e^{-1}$. However, $\{y_{n_i}\}$ has no accumulation point, and $\phi(y_i) \to 0$.

Example 2: Let $\{(x,y) \mid x^2+2y^2=1\} \equiv E$ and for $z = (x,y)$ in the set $E' \equiv \{(x,y) \mid x^2+2y^2 \geq 1\}$, let $I(z)$ be the intersection of E with the line between z and $(0,0)$, and let $S^{(0)}(z) \equiv \{\frac{1}{2}z + \frac{1}{2}I(z)\}$. (Thus z^* is a fixed-point of $S^{(0)}$ if and only if $z^* \in E$.) Now let S be the point-to-set mapping defined as follows: let $C^{(j)}$ $(j=0,1,2,\ldots)$ be the family of circles defined by $C^{(j)} = \{(x,y) \mid x^2+y^2 \leq 1+1/(j+1)\}$; let $C = \{(x,y) \mid x^2+y^2=1\}$ and $C' = \{(x,y) \mid x^2+y^2 \geq 1\}$; given a $z \in C'/C$, let $C(z)$ be the $C^{(j)}$ with the smallest index such that z does not lie in $C^{(j)}$; and, finally, let $S(z)$ be the set consisting of the point on $C(z)$ such that the line connecting z and $S(z)$ is tangent to $C(z)$ and the movement in going from z to $S(z)$ has a clockwise orientation. We now derive a mapping $\overline{S}(z)$ from the mapping

$S(z)$ by the rule that $\overline{S}(z)$ is the set consisting of the point of intersection with the x-axis of the relative interior of the line segment connecting z and $S(z)$ if this point exists, and $\overline{S}(z) = S(z)$ otherwise. Finally, we define $S^{(1)} = \{z\}$ if $z \in E' \cap \{(x,y) | x^2+y^2<1\}$ and $S^{(1)}(z) = \overline{S}(z)$ otherwise.

Consider the following algorithm, which we will denote as algorithm A:

(a) Choose y_0 arbitrarily from C'.

(b) Given y_i, determine y_{i+1} by the following rules:

(i) if $y_i \in C$, let $y_{i+1} = y_i$,

(ii) if y_i is on the x-axis and y_{i-1} was not on the x-axis, choose $y_{i+1} \in S^{(0)}(y_i)$;

(iii) if neither of the assumptions in (i) and (ii) holds, then choose $y_{i+1} \in S^{(1)}(y_i)$.

If the starting point $y_0 \in C'/C$, then we may show that the sequence $\{y_i\}$ generated by the algorithm will use both of the mappings $S^{(0)}$ and $S^{(1)}$ an infinite number of times and have C as its set of accumulation points. The mappings $S^{(0)}$ and $S^{(1)}$ are both strictly monotonic CUGM mappings when we take $G = E' \cap \{(x,y) | x^2+y^2 \leq 2\}$ and $\phi(x,y) = x^2 + y^2$, however, with the exception of $(-1,0)$ and $(1,0)$, the accumulation points of the sequence $\{y_i\}$ are not fixed-points of $S^{(0)}$. Theorem 3.1 does not apply to Algorithm A solely because the "gaps" between uses of $S^{(0)}$ are arbitrarily large.

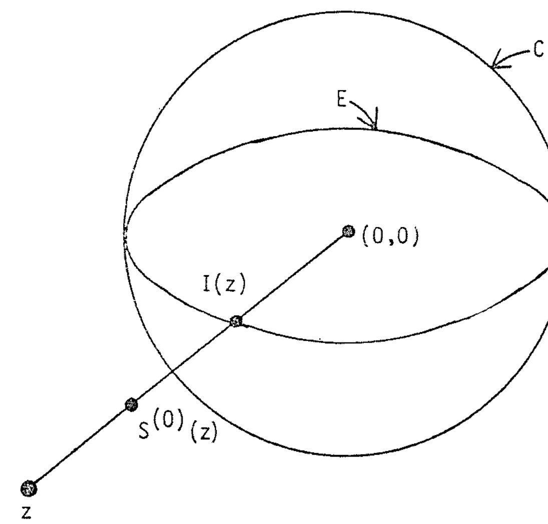

FIGURE 1 — The Mapping $S^{(0)}$

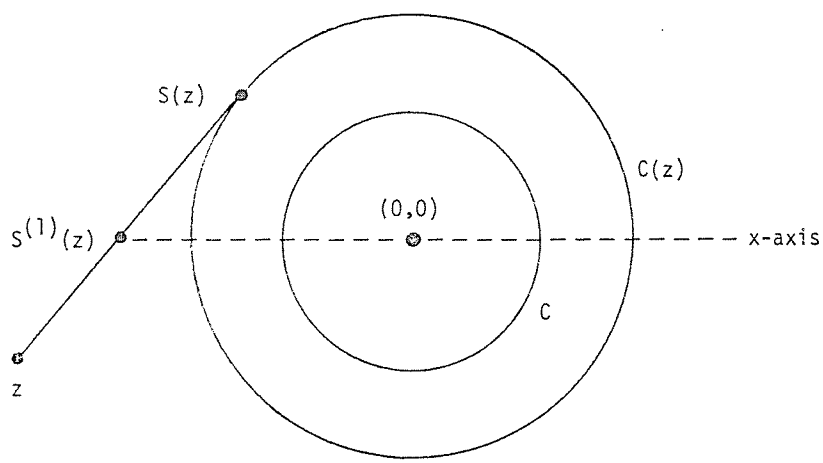

FIGURE 2 — The Mapping $S^{(1)}$

REFERENCES

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

2. R. Fletcher and C. M. Reeves, Function minimization by conjugate gradients", Computer J., 7 (1964), pp. 149-154.

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

4. Melanie L. Lenard, "Practical convergence conditions for restarted conjugate gradient methods", Mathematics Research Center TSR #1373, Madison, Wisconsin, December 1973.

5. Melanie L. Lenard, "Practical convergence conditions for the Davidon-Fletcher-Powell method", Mathematics Research Center TSR #1356, Madison, Wisconsin, April 1974.

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

7. Garth P. Mc Cormick and J. D. Pearson, "Variable metric methods and unconstrained optimization", in Optimization, ed. by R. Fletcher, Academic Press, New York, 1969.

8. Garth P. Mc Cormick and Klaus Ritter, "On the convergence and rate of convergence of the conjugate gradient method", Mathematics Research Center TSR #1118, Madison, Wisconsin, June 1971.

9. R. R. Meyer, "Sufficient conditions for the convergence of monotonic mathematical programming algorithms", University of Wisconsin Computer Sciences Department Technical Report #220, Madison, Wisconsin,

10. James M. Ortega and Werner C. Rheinboldt, "A general convergence result for unconstrained minimization methods", S.I.A.M. J. Numer. Anal., 9 (1972), pp. 40-43.

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

12. K. Ritter, "Accelerating procedures for methods of conjugate directions", Mathematics Research Center TSR #1398, Madison, Wisconsin, April 1974.

13. W. I. Zangwill, _Nonlinear Programming_, Prentice-Hall, Englewood Cliffs, N.J., 1969.