NEW CONVERGENCE THEOREMS
FOR A CLASS OF FEASIBLE DIRECTIONS ALGORITHMS

by

E. Polak and S. Tishyadhigama

Memorandum No. UCB/ERL M78/25

28 February 1980
REVISED

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

NEW CONVERGENCE THEOREMS

FOR A CLASS OF FEASIBLE DIRECTIONS ALGORITHMS

E. Polak and S. Tishyadhigama

Department of Electrical Engineering and Computer Sciences
and the Electronics Research Laboratory
University of California, Berkeley, California 94720

Abstract

This paper presents two new theorems for establishing the convergence properties of multistep constrained optimization algorithms with antijamming features. The theorems extend earlier results of Polak and Klessig and are based on a transcription of multistep methods into one-step methods in a higher dimensional space.

Research sponsored by the National Science Foundation Grant ENG73-08214-A01.

I. Introduction

Zoutendijk's introduction, in 1960 [18], of the concept of "ε-active" constraints in an antijamming device for a class of methods of feasible directions, has led to the use of similar features in a number of other algorithms. These range from algorithms for solving min max problems [1], to dual methods of feasible directions [4,8], to algorithms for solving infinitely constrained optimization problems [12], to combined phase I-phase II methods [13] as well as some others, e.g. [10,11,16]. To coin a phrase, we shall refer to all these algorithms as being ε-controlled. ε-controlled algorithms can be either one step or multistep. Zoutendijk's original methods of feasible directions [18] were multistep and the proof of their convergence was extremely involved and difficult to follow. In an effort to achieve a conceptual simplification, Polak [10] constructed a parallel family of one step methods of feasible directions and showed that their convergence followed in a simple and straightforward manner from a general convergence theorem ((1.3.10 in [10]). Nevertheless, multistep methods have their merits and many more were constructed, though a proof of convergence was generally evaded. The reason for this was that there was a lack of understanding of the essential properties of these ε-controlled, multistep methods, which ensure the convergence of these methods. The first attempts at elucidation were made by R. Klessig [3] and R. Meyer [6] who proposed closely related special purpose convergence theorems. These theorems were still fairly difficult to grasp and cumbersome to apply; nevertheless, they were unquestionably helpful in the construction of a number of algorithms (e.g. [5], [12], [13]).

In this paper we re-examine the essential properties of both one step and multistep ε-controlled algorithms. We show, by example that the simple, one-step algorithms can be used as a stepping stone to the understanding of the more complex multistep methods, and hence to relatively straightforward proofs of convergence. We then abstract out observations in the form of a one step algorithm model in an augmented space and an appropriate convergence theorem. Both this model and the convergence theorem fit in very nicely with earlier results of this type (see e.g. (1.3.10) in [10]). Our results are much simpler than Klessig's [2] and Meyer's [6] and are easier to apply.

We hope that our success in simplifying the analysis of multistep feasible directions type algorithms, by transcribing them into one-step methods in a higher dimensional space, will lead to similar successes for other classes of multistep algorithms, such as quasi-Newton and conjugate directions methods.

2. The Essential Properties of ε-controlled Algorithms

The simplest examples of ε-controlled algorithms which have all the essential properties of ε-controlled algorithms are the Polak [10] and Zoutendijk [18] methods of feasible directions. Polak's algorithm is the simpler of the two: it is one step. Zoutendijk's algorithm is multistep. Both of these algorithms solve the problem:

(1) $$\min\{f^0(z)\,|\,f^j(z) \le 0, j = 1,2,\ldots,m\},$$

where $f^j : \mathbb{R}^n \to \mathbb{R}$, $j = 0,1,\ldots,m$ are continuously differentiable, under the following assumption. Let

(1a) $$\Omega \triangleq \{z\,|\,f^j(z) \le 0,\ j \in J_m\},$$

(1b) $$J_m \triangleq \{1,2,\ldots,m\}.$$

(2) Assumption. The set Ω has an interior and it is equal to the closure of its interior, (see (4.2.5) and (4.3.2) in [10]). ###

For each $z \in \Omega$, $\varepsilon \geq 0$, we define

(3) $$J(\varepsilon,z) \triangleq \{j \in J_m | f^j(z) \geq -\varepsilon\},$$

(4) $$\theta(\varepsilon,z) \triangleq \min_{h \in S} \max\{\langle \nabla f^j(z),h \rangle | j \in J(\varepsilon,z) \cup \{0\}\},$$

where S is the L_∞-unit sphere about the origin, i.e.,

(4a) $$S \triangleq \{h \in R^n | |h^i| \leq 1, i = 1,2,\ldots,n\}.$$

Note that $\theta(\varepsilon,z) \leq 0$ for all $\varepsilon \geq 0$ and $z \in \Omega$.

We shall say that a feasible point $\hat{z}$ is a desirable point if it satisfies the first order F. John necessary condition for optimality for problem (1) [21], i.e., $\hat{z} \in \Omega$ and $0 \in co\{\nabla f^j(\hat{z}) | j \in \{0\} \cup J(0,\hat{z})\}$. Let Δ be the set of all desirable points. Then, by Theorem (1.2.8) in [10], we have

(5) $$\Delta = \{z \in \Omega | \theta(0,z) = 0\},$$

where $\theta(\cdot,\cdot)$ is defined by (4).

For the sake of compactness, we state the Polak and Zoutendijk algorithms as one, with a selector p. When p = 0, the algorithm becomes Polak's while when p = 1 the algorithm becomes Zoutendijk's. Both algorithms make use of a function $e : \mathbb{R}_+ \times \Omega \to \mathbb{R}_+$, defined, in terms of a parameter $\alpha \in (0,1)$, by

(6) $$e(\varepsilon,x) \triangleq \max\{\varepsilon' | \varepsilon' = 0 \text{ or } \varepsilon' = \alpha^k \varepsilon, k \in \mathbb{N};$$
$$\theta_{\varepsilon'}(x) \leq -\varepsilon'\}, \text{ where } \mathbb{N} \triangleq \{0,1,2,\ldots\}.$$

(7) Algorithm.

Parameters: $\alpha,\beta,\gamma \in (0,1)$.

Step 0: Compute $z_0 \in \Omega$, choose $\varepsilon_0 > 0$, set $i = 0$.

Step 1: Choose $p = 0$ or $p = 1$.

Comment: $p = 0$ indicates that Polak's antijamming scheme is chosen; $p = 1$ indicates that Zoutendijk's antijamming scheme is chosen (see Step 4).

Step 2: If $z_i \in \Delta$, stop; else, compute a vector h_i as a solution of (4) for $z = z_i$ and $\varepsilon = e(\varepsilon_i, z_i)$.

Step 3: Compute the smallest nonnegative integer ℓ_i such that $\lambda_i \triangleq \beta^{\ell_i}$ satisfies

(8) $$f^0(z_i+\lambda_i h_i) - f^0(z_i) \leq \gamma\lambda_i\theta(\varepsilon,z_i),$$

(9) $$f^j(z_i+\lambda_i h_i) \leq 0,\; j \in J_m.$$

Step 4: Set $z_{i+1} = z_i+\lambda_i h_i$. If $p = 0$, set $\varepsilon_{i+1} = \varepsilon_0$; else set $\varepsilon_{i+1} = e(\varepsilon_i,z_i)$.

Step 5: Set $i = i+1$ and go to Step 1. ###

(10) Remark. The step length selection in Step 6 can be modified by using a step-size parameter $\lambda_s > 0$, so that $\lambda = \beta^\ell\lambda_s$. ###

Referring to sec. 4.3 in [10], we find that Polak's method has the following:

(11) Essential Properties (P)

(i) Since $\varepsilon_i = \varepsilon_0$ for $i = 0,1,2,\ldots$, $\varepsilon = e(\varepsilon_i,z_i) = e(\varepsilon_0,z_i)$, as used in step 2 to compute a descent direction, is a function of z_i only and hence h_i is a function of z_i only, i.e., the method is a one step method.

(ii) For every $z \in \Omega \sim \Delta$, there exist $\rho_z > 0$ and $\varepsilon_z > 0$ such that $e(\varepsilon_0, z_i) \geq \varepsilon_z$ for all $z_i \in B(z,\rho_z) \triangleq \{z' \mid \|z-z'\| \leq \rho_z\}$.

(iii) For every $\varepsilon > 0$ and $z \in \Omega \sim \Delta$, there exists a $\rho_{\varepsilon,z} > 0$ and an integer $\ell_{\varepsilon,z} \geq 0$ such that for all $z_i \in B(z,\rho_{\varepsilon,z})$ satisfying $e(\varepsilon_0,z_i) \geq \varepsilon$, ℓ_i, as computed in Step 3, satisfies $\ell_i \leq \ell_{\varepsilon,z}$; i.e. $\lambda_i \geq \beta^{\ell_{\varepsilon,z}}$ for all $z_i \in B(z,\rho_{\varepsilon,z})$. ###

It is very easy to deduce the convergence properties of Polak's algorithm from its essential properties, as follows.

(12) <u>Theorem</u> [10]. Suppose $\{z_i\}$ is an infinite sequence constructed by Polak's algorithm. Then every accumulation point of $\{z_i\}$ is in Δ.

<u>Proof</u>: Suppose that $\{z_i\}$ has an accumulation point $\hat{z}$. Then, because $\theta(\varepsilon,z_i) \leq 0$ always, it follows from (8) and the continuity of $f^0(\cdot)$ that $f^0(z_i) \searrow f^0(\hat{z})$ as $i \to \infty$. Now, for the sake of contradiction, suppose that $\hat{z} \notin \Delta$. Then according to Properties (11) (ii) and (11) (iii), there exist $\hat{\rho} > 0$, $\hat{\varepsilon} > 0$ and $\hat{\ell} \in \mathbb{N}$ such that $e(\varepsilon_0,z_i) \geq \hat{\varepsilon}$, $\ell_i \leq \hat{\ell}$ and by (8), $f^0(z_{i+1}) - f^0(z_i) \leq \gamma\lambda_i\theta(e(\varepsilon_0,z_i),z_i) \leq -\gamma\beta^{\hat{\ell}}\hat{\varepsilon} < 0$, for all $z_i \in B(\hat{z},\hat{\rho})$. But this shows that $\{f^0(z_i)\}$ is not Cauchy and hence we have a contradiction.

###

Our reexamination of Zoutendijk's method, leads to our isolation of the following:

(13) Essential Properties (Z)

(i) Since $\varepsilon_{i+1} = e(\varepsilon_i, z_i)$ depends both on ε_i and z_i, the same holds for h_i and hence the method is multistep if the sequence $\{z_i\}$ only is considered.

(ii) From Essential Property (11) (ii) and (6), it follows for Zoutendijk's method that for every $z \in \Omega \sim \Delta$ there exist $\rho_z > 0$ and $\varepsilon_z > 0$ such that

(14) $$e(\varepsilon_i, z_i) \geq \min\{\varepsilon_i, \varepsilon_z\} \text{ for all } z_i \in B(z, \rho_z)$$

(iii) From Essential Property (11) (iii), it follows for Zoutendijk's method that for every $\varepsilon > 0$ and $z \in \Omega \sim \Delta$, there exist a $\rho_{\varepsilon,z} > 0$ and an integer $\ell_{\varepsilon,z} \geq 0$ such that for all $\varepsilon_i \geq \varepsilon$ and $z_i \in B(z, \rho_{\varepsilon,z})$, satisfying $e(\varepsilon_i, z_i) \geq \varepsilon$, ℓ_i, as computed in Step 3, satisfies $\ell_i \leq \ell_{\varepsilon,z}$, i.e., the step size $\lambda_i \geq \beta^{\ell_{\varepsilon,z}}$.

(iv) (Common to Polak and Zoutendijk methods). Since $\|h_i\| \leq \sqrt{n}$ for all $h_i \in S$ and $z_{i+1} = z_i + \lambda_i h_i$,

(14a) $$\lambda_i \geq \|z_{i+1} - z_i\| / \sqrt{n}$$

(v) (Common to Polak and Zoutendijk methods). From theorem (4.3.35) in [10], for every $z \in \Omega \sim \Delta$, there exist $\rho'_z > 0$ and $\varepsilon'_z > 0$ such that for all $\varepsilon_i \in [0, \varepsilon'_z]$ and $z_i \in B(z, \rho'_z)$

(14b) $$\theta(\varepsilon_i, z_i) \leq -\varepsilon'_z$$ ###

We get two important immediate consequences of properties (13)(iii)-(v). First, from (13)(iii), (13)(v) and (8), given $\varepsilon > 0$ and $z \in \Omega \sim \Delta$, there exist $\rho_{\varepsilon,z} > 0$, $\ell_{\varepsilon,z} \geq 0$ such that for all $\varepsilon_i \geq \varepsilon$ and $z_i \in B(z, \rho_{\varepsilon,z})$ such that $e(\varepsilon_i, z_i) \geq \varepsilon$,

(15) $$f^0(z_{i+1}) - f^0(z_i) \leq \gamma\lambda_i \theta(e(\varepsilon_i, z_i), z_i) \leq -\gamma\beta^{\ell_{\varepsilon,z}}\varepsilon \triangleq -\delta_{\varepsilon,z} < 0.$$

Next, from (8) and (13) (iv) and (v), we conclude that for every

$z \in \Omega \sim \Delta$ there exist $\rho'_z \in (0,\rho_z]$ and $\varepsilon'_z \in (0,\varepsilon_z]$ such that for all $\varepsilon_i > 0$ and $z_i \in B(z,\rho'_z)$

(15a) $$f^0(z_{i+1}) - f^0(z_i) \leq \gamma\lambda_i\theta(e(\varepsilon_i,z_i),z_i)$$
$$\leq \frac{-\gamma}{\sqrt{n}} \|z_{i+1}-z_i\| \, \theta(e(\varepsilon_i,z_i),z_i)$$
$$\leq \frac{-\gamma}{\sqrt{n}} \|z_{i+1}-z_i\| \, \varepsilon'_z$$
$$\triangleq -\delta'_z \|z_{i+1}-z_i\|$$

(16) <u>Theorem</u>: Suppose $\{z_i\}$ is an infinite sequence constructed by Zoutendijk's algorithm. Then every accumulation point $\{z_i\}$ is in Δ.

<u>Proof</u>: Suppose $\{z_i\}$ has an accumulation point $\hat{z}$, i.e., $z_i \xrightarrow{K} \hat{z}$, $K \subset \mathbb{N}$. Then, because $\theta(\varepsilon_i,z_i) \leq 0$ always holds, it follows from (8) and the continuity of $f^0(\cdot)$ that $f^0(z_i) \searrow f^0(\hat{z})$ as $i \to \infty$. For the sake of contradiction, suppose that $\hat{z} \notin \Delta$.

Let $\hat{\varepsilon} \triangleq \varepsilon_{\hat{z}} > 0$ and $\hat{\rho} \triangleq \min\{\rho_{\hat{z}},\rho'_{\hat{z}}\}$ be as specified in Essential Property (13)(ii) and (v) for $\hat{z}$. If $z_i \in B(\hat{z},\hat{\rho})$ for all $i \geq i_0$, for some i_0, then from (14), $\varepsilon_i \geq \min\{\varepsilon_{i_0},\hat{\varepsilon}\} \triangleq \bar{\varepsilon}$ for all $i \geq i_0$ and also $e(\varepsilon_i,z_i) \geq \bar{\varepsilon}$ for all $i \geq i_0$. It now follows from (13)(iii) and (8), via (15) that there exists an $i_1 \geq i_0$ such that

(17) $$f^0(z_{i+1}) - f^0(z_i) \leq -\delta_{\bar{\varepsilon},\hat{z}} < 0 \text{ for all } i \geq i_1,\ i \in K$$

and hence $\{f^0(z_i)\}$ is not Cauchy. Since this contradicts the fact that $f^0(z_i) \searrow f^0(\hat{z})$, we must have an infinite subsequence of $\{z_i\}$ outside of $B(\hat{z},\hat{\rho})$.† Let $\mu \in (0,1)$. Since $z_i \xrightarrow{K} \hat{z}$, there exists an i_2 such that $z_i \in B(\hat{z},\mu\hat{\rho})$ for all $i \geq i_2$ and $i \in K$. For any $i \in K$, $i \geq i_2$ let $j > 0$ be

†In this case, as z_i keeps reentering $B(\hat{z},\hat{\rho})$, ε_i, could, conceivably, be decreasing to zero and hence the arguments used in the proof of Theorem (12) cannot be used here.

the smallest integer such that $z_{i+j} \notin B(\hat{z},\hat{\rho})$. Then, since $\varepsilon_i > 0$ for all i, from a repeated application of (15a) and the fact $\|z_{i+j}-z_i\| \geq (1-\mu)\hat{\rho}$, we obtain, with $\hat{\delta}' \triangleq \delta'_{\hat{z}}$, that

$$f^0(z_{i+j}) - f^0(z_i) \leq -\hat{\delta}'(1-\mu)\hat{\rho} \tag{18}$$

which shows that $\{f^0(z_i)\}$ is not Cauchy. Hence we have a contradiction and the proof is complete. ###

The above proof of the convergence of Zoutendijk's method seems about as simple as it can be made. The simplification was achieved by using Polaks's method as a stepping stone and by observing that the proof of convergence for Zoutendijk's method has to differ from the one given for Polak's method by an appropriate use of properties (15) and (15a).

In the next section, we restate our observations on Zoutendijk's method in the form of an algorithm model and convergence theorem applicable to the class of algorithms discussed in the introduction.

3. A Generalization of Polak's Algorithm Model and Convergence Theorem

Let X be a closed subset of a normed linear space and let $\Sigma \subset X$ be a set of solution points. In [10], Polak considers one step algorithms defined by a map $\tilde{A} : X \to 2^X$, as follows:

(19) Algorithm Model (1.3.9) in [10]:

Step 0: Choose $x_0 \in X$; set $i = 0$.

Step 1: If $x_i \in \Sigma$, stop.

Step 2: Compute an $x_{i+1} \in \tilde{A}(x_i)$.

Step 3: Set $i = i+1$ and go to Step 1. ###

The corresponding convergence theorem reads as follows:

(20) Theorem (1.3.10) in [10]: Suppose that (i) there exists a $c : X \to \mathbb{R}$ which is either continuous or bounded from below on X, and (ii) for every $x \in X \sim \Sigma$, there exist $\rho_x > 0$, $\delta_x > 0$ such that $c(x'') - c(x') \leq -\delta_x$ for all $x' \in B(x,\rho_x) \cap X$, for all $x'' \in \tilde{A}(x')$. Then all accumulation points of an infinite sequence $\{x_i\}$ constructed by (19) are in Σ. ###

To cast both one step and multistep ε-controlled algorithms in the one step form (19), we only need to augment the set in which the iterations take place to $X \times \mathbb{R}_+$ to introduce two maps: $A : X \times \mathbb{R}_+ \to 2^X$ and $E : X \times \mathbb{R}_+ \to 2^{\mathbb{R}_+}$. Referring to Algorithm (7), we find that the appropriate generalization reads as follows:

(21) Algorithm Model.

Step 0: Choose $\varepsilon_0 > 0$, $x_0 \in X$, set $i = 0$.

Step 1: If $x_i \in \Sigma$, stop.

Step 2: Compute $\varepsilon_{i+1} \in E(\varepsilon_i,x_i)$, $x_{i+1} \in A(\varepsilon_i,x_i)$.

Step 3: Set $i = i+1$ and go to Step 1. ###

Referring to Zoutendijk's method, we see that in that case $E(\varepsilon_i,z_i) = \{e(\varepsilon_i,z_i)\}$ and $A(\cdot,\cdot)$ has the form $A(\varepsilon_i,z_i) = A'(e(\varepsilon_i,z_i),z_i)$, with the definition of $A'(\cdot,\cdot)$ obvious from the algorithm.

In Polak's algorithm, $E(\varepsilon_i,z_i) = \{\varepsilon_0\}$ and $A(\cdot,\cdot)$ is defined in the same manner as for Zoutendijk's algorithm.

As in the case of Theorem (20), we shall assume that there exists a cost function $c : X \to \mathbb{R}$ associated with Algorithm Model (21). To complete our notation, for $r > 0$, $x \in X$, we define

(22) $$B(x,r) \triangleq \{x' \in X \mid \|x'-x\| \leq r\},$$

where $\|\cdot\|$ is a particular norm associated with the space containing X.

(23) <u>Definition</u>: Let $\{(\varepsilon_i, x_i)\}$ be any sequence constructed by (21). <u>We say that</u> x_j <u>is a successor of</u> x_i <u>if</u> (ε_i, x_i) <u>and</u> (ε_j, x_j) <u>are points in the sequence and</u> $j > i$. ###

The Essential Properties (P) and (Z) now lead to the set of assumptions in the theorem below.

(24) <u>Theorem</u>: Suppose that

(i) For all $\varepsilon \geq 0$, $x \in \Sigma$, $A(\varepsilon, x) = \{x\}$ and $E(\varepsilon, x) = \{0\}$.

(ii) $c(\cdot)$ is locally bounded from below on $X \sim \Sigma$.

(iii) For each $x \in X \sim \Sigma$ there exist $\rho_x > 0$ and $\varepsilon_x > 0$ such that for any $x' \in B(x, \rho_x)$, $\varepsilon' > 0$, $\varepsilon'' \in E(\varepsilon', x')$,

(25) $$\varepsilon'' \geq \min\{\varepsilon', \varepsilon_x\}.$$

(iv) For each $x \in X \sim \Sigma$ and $\varepsilon > 0$ there exist $\rho_{\varepsilon,x} > 0$ and $\delta_{\varepsilon,x} > 0$ such that

(26) $$c(x'') - c(x') \leq -\delta_{\varepsilon,x} \quad \text{for all } \varepsilon' \geq \varepsilon,\ x' \in B(x, \rho_{\varepsilon,x}),\ x'' \in A(\varepsilon', x')$$

(v) For each $x \in X \sim \Sigma$, there exist $\rho'_x > 0$ and $\delta'_x > 0$ such that for any $r \in (0, \rho'_x]$, and $\sigma \in (0,1)$, there exists $\mu_{\sigma,r} > 0$ with the property that if $x_j \in X \sim B(x,r)$ is a successor of $x_i \in B(z, \sigma r)$ then

(27) $$c(x_j) - c(x_i) \leq -\delta'_x \, \mu_{\sigma,r}{}^{\dagger}.$$

Let $\{(\varepsilon_i, x_i)\}$ be any infinite sequence constructed by Algorithm Model (18). Then every accumulation point of $\{x_i\}$ is in Σ.

<u>Proof</u>. Because the sequence $\{\varepsilon_i, z_i\}$ to be considered is infinite, $x_i \in X \sim \Sigma$ for all i. For the sake of contradiction, assume that there exists a subsequence $\{x_i\}_{i \in K}$ such that $x_i \overset{K}{\to} \hat{x}$, with $\hat{x} \in X \sim \Sigma$. Hence by (24) (ii) $\liminf_{i \in K} c(x_i) = \hat{c} > -\infty$. Since $x_i \in X \sim \Sigma$ for all i and $\varepsilon_0 > 0$,

†This assumption is an obvious generalization of (15a).

conditions (24)(iii), (iv) imply that $\varepsilon_i > 0$ and $c(x_{i+1}) < c(x_i)$, for all i. Since the sequence $\{c(x_i)\}$ is monotonically decreasing, $c(x_i) \to \hat{c}$ as $i \to \infty$.

Let $r = \min\{\rho_{\hat{x}}, \rho_{\hat{x}}^!\}$, where $\rho_{\hat{x}}$ and $\rho_{\hat{x}}^!$ are positive constants as in (24)(iii) and (24)(v) respectively. We may assume that $x_i \in B(x,r)$ for all $i \in K$. Two cases may arise:

(i) There exists an integer i_0 such that $x_i \in B(\hat{x},r)$ for all $i \geq i_0$.

(ii) There exists an infinite subsequence $\{x_j\}_{j \in K'}$ such that $x_j \in X - B(\hat{x},r)$, for all $j \in K'$.

In case (i), let $\varepsilon_{\hat{x}}$ be as specified in (24) (iii), then for all $i \geq i_0$, $\varepsilon_i \geq \min\{\varepsilon_{i_0}, \varepsilon_{\hat{x}}\} \triangleq \hat{\varepsilon}$. Let $\delta_{\hat{\varepsilon},\hat{x}}$ be as specified in (24)(iii), then (26) implies that $c(x_{i+1}) - c(x_i) \leq -\delta_{\hat{\varepsilon},\hat{x}}$ for all $i \geq i_1 \geq i_0$, $i \in K$. But this implies that $\{c(x_i)\}$ is not Cauchy, which is a contradiction.

We now consider case (ii). Let $\sigma \in (0,1)$ be arbitrary and let $\mu_{\sigma,r}$ be the constant in (24)(v). Let i_2 be an integer such that $x_i \in B(\hat{x},\sigma r)$ for all $i \geq i_2$ and $i \in K$. Let $K^* \subset K \cup K'$ be an infinite subset of integers constructed as follows. We denote the elements of K^* by $i(k)$, with $i(k+1) > i(k)$, for all $k = 0,1,\ldots$. We choose $i(0) \in K$ such that $i(0) \geq i_2$. For $k = 0,1,\ldots$, given $i(k) \in K^*$, if k is even, we let $i(k+1) \in K'$; if k is odd, we let $i(k+1) \in K$. (The choice of the elements in K, K' is arbitrary except that the ordering $i(k+1) > i(k)$ must be observed.) Making use of (24)(iv), we obtain for all $i(k) \in K^*$ and k even,

$$c(x_{i(k+1)}) - c(x_{i(k)}) \leq -\delta_{\hat{x}}^! \mu_{\sigma,\mu}$$

This implies that $\{c(x_i)\}_{i \in K^*}$ is not Cauchy which contradicts the fact that every subsequence of a Cauchy sequence is a Cauchy sequence. Hence, our proof is complete. ###

Although Theorem (24) seems to be in the most general form possible, it suffers from the aesthetic drawback that assumption (v) is really a multistep assumption that is made for a one step algorithm. At a slight loss of generality this matter can be remedied as follows (c.f. (15a)).

(28) Corollary: Suppose that assumption (v) of Theorem (24) is replaced with: (iv)' For every $x \in X \sim \Sigma$, there exist $\rho_x' > 0$ and $\delta_x' > 0$ such that for all $x' \in B(x,\rho_x')$ and any $\varepsilon' \geq 0$,

(29) $$c(x'') - c(x') \leq -\delta_x' \| x''-x' \| \text{ for all } x'' \in A(x',\varepsilon')$$

Then the conclusions of theorem (24) remain valid. ###

Quite obviously the above assumption (iv)' implies assumption (iv) in Theorem (24) and hence there is no need for a proof of the corollary.

Algorithm Model (21) follows the pattern set in Algorithm Model (19) in that it suppresses all the detail of the structure of the maps $E(\cdot,\cdot)$ and $A(\cdot,\cdot)$.

Conclusion

This paper was motivated by a desire to simplify the proofs of convergence of a class of algorithms which, in structure, resemble Zoutendijk's methods of feasible directions. Our approach was to establish the essential properties of these algorithms as a consequence of the essential properties of the simpler, Polak type methods of feasible directions and to present the main steps one needs to follow in constructing a simple proof. We then abstracted our findings in the form of an algorithm model and a general convergence theorem. We hope that the results in this paper will eliminate a good part of the difficulties in understanding ε-controlled algorithms and their convergence properties.

References

[1] V. F. Demjanov, "Algorithms for some minimax problems," Journal of Computer and System Science, 2, (1968) 342-380.

[2] P. Huard, "Programmation matématique convexe," R. I. R. O., 7, (1968) 43-59.

[3] R. Klessig, "A general theory of convergence for constrained optimization algorithms that use antizigzagging provisions," SIAM J. Control, 12 (4), (1974) 598-608.

[4] O. L. Mangasarian, "Dual, feasible direction algorithms," in Techniques of Optimization (A. V. Balakrishnan, ed.), Academic Press, New York, (1969) 67-88.

[5] D. Q. Mayne and E. Polak, "Feasible directions algorithms for optimization problems with equality and inequality constraints," Mathematical Programming, 11, (1976) 67-80.

[6] R. R. Meyer, "A convergence theory for a class of anti-jamming strategies," Tech. Rept. 1481, Mathematical Research Center, University of Wisconsin, Madison, Wisconsin, (1974).

[7] J. M. Ortega and W. C. Rheinboldt, Iterative Solution of Nonlinear Equations in Several Variables, Academic Press, New York, (1970).

[8] O. Pironneau and E. Polak, "A dual method for optimal control problems," SIAM J. Control, 11 (3), (1973).

[9] O. Pironneau and E. Polak, "Rate of convergence of a class of methods of feasible directions," SIAM J. Numerical Analysis, 10 (1), (1973) 161-174.

[10] E. Polak, Computational Methods in Optimization: A Unified Approach, Academic Press, New York, (1971).

[11] E. Polak and D. Q. Mayne, "First-order strong variation algorithms for optimal control problems with terminal inequality constraints," J. O. T. A., 16 (3-4), (1975) 303-325.

[12] E. Polak and D. Q. Mayne, "An algorithm for optimization problems with functional inequality constraints," IEEE Trans. Automatic Control, AC-12 (2), (1976), 184-193.

[13] E. Polak, R. Trahan and D. Q. Mayne, "Combined phase I-phase II methods of feasible directions," Memorandum No. UCB/ERL M77/39, University of California, Berkeley, California, (1977).

[14] S. W. Rauch, "A convergence theory for a class of nonlinear programming problems," SIAM J. Numerical Analysis, 10 (1), (1973), 207-228.

[15] S. Tishyadhigama, E. Polak and R. Klessig, "A comparative study of several general convergence conditions for algorithms modeled by point-to-set maps," Memorandum No. UCB/ERL M77/78, University of California, Berkeley, California, (1977), (to appear in Mathematical Programming Study, 7, 1979).

[16] D. M. Topkis and A. F. Veinott, Jr., "On the convergence of some feasible direction algorithms for nonlinear programming," SIAM J. Control, 5 (2), (1967) 268-279.

[17] W. I. Zangwill, Nonlinear Programming: A Unified Approach, Prentice-Hall, Englewood Cliffs, New Jersey, (1969).

[18] G. Zoutendijk, Methods of Feasible Directions, Elsevier Publishing Company, Amsterdam, Holland, (1960).

[19] G. Zoutendijk, "Some algorithm based on the principle of feasible directions," in Nonlinear Programming (J. B. Rosen, O. L. Mangasarian and K. Ritter, eds.), Academic Press, New York, (1970).

[20] S. I. Zukhovitskii, R. A. Polyak and M. E. Primak, "An algorithm for the solution of convex Tchebycheff approximation problems," Doklady Akademii Nauk SSSR, 151, (1963) 27-30.

[21] F. John, "Extremum problems with inequalities as side conditions," in Courant Anniversary Volume (K. O. Friedrichs, O. E. Neugebauer and J. J. Stoker, eds.), Interscience, New York, (1948) 187-204.

[22] S. Tishyadhigama, "General convergence theorems: Their relationships and applications," Ph.D. Dissertation, Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, (1977).