ON THE RATE OF CONVERGENCE OF CERTAIN METHODS OF CENTERS

by

O. Pironneau and E. Polak

Memorandum No. ERL-M296

12 February 1971

Electronics Research Laboratory

College of Engineering
University of California, Berkeley
94720

ABSTRACT

It is shown in this paper that a theoretical method of centers, introduced by Huard, converges linearly. It is also shown, by counter-example, that a modified method of centers due to Huard and a method of feasible directions due to Topkis and Veinott cannot converge linearly even under convexity assumptions because of this, two new modified methods of centers are introduced, one theoretical and one implementable, both using a quadratic programming direction finding subroutine, and both of which are shown to converge linearly.

Research sponsored by the Joint Services Electronics Program, Grant AFOSR-68-1488 and the National Aeronautics and Space Administration, Grant NGL-05-003-016.

INTRODUCTION

The family of optimization algorithms known as methods of centers were introduced by Huard [4]. They differ from one another only in the distance function used to establish a center; an operation which must be repeated at each iteration. In their original form, these algorithms are not implementable, and because of this, various approximations, or modified methods of centers, have been proposed. The best known of these modified methods of centers is also due to Huard [5]. (These algorithms are also discussed in [8]). In this paper we shall show that a theoretical method of centers, presented in [4], and a new modified method of centers converge linearly on a class of problems. Our analysis will be based on a systematic utilization of duality theory.

The modified method of centers to be discussed in this paper is a variation of the algorithm described in [5]. Although we shall not establish the rate of convergence of the algorithm in [5] in this paper, we wish to mention that we were able to show that the algorithm in [5] converges at least as fast as $\frac{1}{\sqrt{i}}$ under the same assumptions under which our algorithm converges linearly. Furthermore, the example in the appendix indicates that the algorithm in [5] cannot converge linearly under the assumptions used in this paper. Thus, although the modified method of centers to be presented in this paper is more complex than the one in [5], it does have a better rate of convergence.

As we shall see, our analysis depends on an extention due to Geoffrion [3] of Wolfe's strong duality theorem [12],[17]. Therefore, in section 1, we shall begin by stating this theorem as well as a few other results we shall use repeatedly. Then, in section 2, we shall obtain the rate of convergence of a theoretical method of centers; finally, in section 3, we shall describe a new modified method of centers and we shall establish a bound on its rate of convergence. Since we shall be exclusively interested in rate of convergence, we shall assume that the reader is familiar with methods of centers and their convergence properties. In any event, the reader will find these described in considerable detail in [8].

Finally, we wish to note that the two algorithms discussed in this paper are not the only ones for which a rate of convergence can be obtained by a systematic use of duality theory. Similar results can also be obtained for a class of methods of feasible directions. However, because of space considerations, these will be presented separately.

SECTION I. PRELIMINARIES

Algorithms for solving problems of the form $\min\{f^0(z) | f^j(z) \leq 0, j = 1, \cdots, m\}$ usually generate sequences of points $\{z_i\}$ such that the corresponding cost sequences $\{f^0(z_i)\}$ are monotonically decreasing. The convergence of such algorithms can usually be established from the properties of the difference $f^0(z_{i+1}) - f^0(z_i)$ (see Polak [8]). We shall show that in some case, a study of the difference $f^0(z_{i+1}) - f^0(z_i)$ can also lead to a bound on the rate of convergence of the sequence $\{z_i\}$.

We recall that a sequence of points $\{z_i\}_{i=0}^{\infty}$ in a Banach space $\mathbb{B}$ is said to converge at least linearly if there exist a $\hat{z} \in \mathbb{B}$, $i_0 > 0$, $k \in (0,1)$, $\kappa > 0$ such that

1-1 $$\|z_i - \hat{z}\| \leq \kappa \cdot k^i \quad \text{for all } i \geq i_0 .$$

The sequence $\{z_i\}_{i=0}^{\infty}$ is said to converge superlinearly to $\hat{z}$ if for any $\kappa > 0$ and any $k \in (0,1)$ there exists i_0 such that 1-1 is satisfied. Our method for showing that 1-1 holds for the sequences under consideration, will be based on the observation that under suitable assumptions (see Lemma 1-20), if the sequence of costs $\{f^0(z_i)\}_{i=0}^{\infty}$ satisfies, for some $i_0 \geq 0$ and $k \in (0,1)$,

1-2 $$f^0(z_{i+1}) - f^0(\hat{z}) \leq k[f^0(z_i) - f^0(\hat{z})] \quad \text{for all } i \geq i_0,$$

then there exists a $\kappa > 0$ such that 1-1 is satisfied. Note that 1-2 is equivalent to

1-3 $$f^0(z_{i+1}) - f^0(z_i) \leq -(1-k)[f^0(z_i) - f^0(\hat{z})] \quad \text{for all } i \geq i_0$$

which, for us, will prove to be a more convenient form to work with. To establish 1-2, we shall make use of the fact that the methods of centers, to be studied in this paper, construct an upper bound $\delta(z_i) \geq [f^0(z_{i+1}) - f^0(z_i)]$, in the process of computing z_{i+1} from z_i. This bound $\delta(z_i)$ is computed as the optimal value of a minimization subproblem. In turn, to obtain an upper bound on $\delta(z_i)$, we shall make use of optimality conditions in saddle point form (Kuhn-Tucker [6]). Since this will also require conditions for existence of a saddle point, we shall make use of the strong duality theorem, stated below, which incorporates all the required results. The bound on $\delta(z_i)$ will be obtained in the form $\delta(z_i) \leq \kappa(z_i)[f^0(\hat{z}) - f^0(z_i)]$, where $\kappa(\cdot)$ will be shown to be an upper semicontinuous function. Since, under the assumptions to be introduced, the sequence $\{z_i\}$ converges to $\hat{z}$, the solution of 1-1, given any $\alpha \in (0,1)$, there exists an $i_0(\alpha) \geq 0$ such that $\kappa(z_i) \geq (i-\alpha)\kappa(\hat{z})$ for all $i \geq i_0(\alpha)$. Because of this we obtain the bound $f^0(z_{i+1}) - f^0(z_i) < (i-\alpha)\kappa(\hat{z})[f^0(\hat{z})-f^0(z_i)]$ for all $i \geq i_0(\alpha)$, which, in turn, enables us to establish linear convergence of $\{z_i\}$.

In this section we shall develop the tools needed to carry out the plan of attack outlined above. Thus, consider the problem, denoted by (P),

$$\text{(P)} \qquad \min\{g^0(z) \mid z \in \Omega\}$$

with $\Omega = \{z \in R^n \mid g^j(z) \leq 0 \; j = 1, \ldots, p; \; z \in C\}$, where $g^j: \mathbb{C} \to R$, $j = 0, 1, \cdots, p$ are <u>convex</u> and continuously differentiable functions, and $\mathbb{C}$ is a <u>convex</u> subset of R^n. The following problem, denoted by (D), is called the dual of (P):

$$\text{(D)} \qquad \max_{u \geqq 0}\{\inf_{z \in \mathbb{C}}\{g^0(z) + \sum_{j=1}^{p} u^j g^j(z)\}\} \qquad (u = (u^1, u^2, \cdots, u^p) \in R^p)$$

Let ϕ: $R^p \to R \cup \{-\infty\}$, be defined by,

1-4 $$\phi(u) = \inf_{z \in \mathbb{C}} \{g^0(z) + \sum_{j=1}^{p} u^j g^j(z)\}$$

<u>Definition</u>:

Any $\hat{z} \in \Omega$, satisfying $g^0(\hat{z}) = \min\{g^0(z) | z \in \Omega\}$, will be called a solution of (P). Any $\hat{u} \geqq 0$, satisfying, $\phi(\hat{u}) = \max_{u \geqq 0} \phi(u)$, will be called a solution of (D). $\square$

1-5 <u>Strong duality theorem</u>.

Let S_p be the set of all solutions of (P), i.e.,

1-6 $$S_p = \{z' \in \Omega | g^0(z') = \min_{z \in \Omega} g^0(z)\}$$

Suppose that S_p is not empty and that

1-7 $$\Omega' = \{z | g^j(z) < 0, j = 1, \cdots, p\}$$

is not empty[1]. Then

a) i) Problem (D) has at least one solution,

1-8 ii) $$\max_{u \geqq 0} \{\inf_{z \in \mathbb{C}} \{g^0(z) + \sum_{j=1}^{p} u^j g^j(z)\}\} = \min_{z \in \mathbb{C}} \{g^0(z) | g^j(z) \leq 0 \ j = 1, \cdots, p\};$$

iii) for any $\bar{u}$, solution of (D), and for any $\hat{z}$ solution of (P),

1-9 $$g^0(\hat{z}) = \min_{z \in \mathbb{C}} \{g^0(z) + \sum_{j=1}^{p} \bar{u}^j g^j(z)\} ;$$

b) a vector $\bar{u} \geqq 0$ in R^p is a solution of (D) if and only if there exists a

$\tilde{z}$ in Ω such that

1-10 $$\text{i)}\quad g^0(\tilde{z}) + \sum_{j=1}^{p} \bar{u}^j g^j(\tilde{z}) = \min_{z\in\mathbb{C}}\{g^0(z) + \sum_{j=1}^{p} \bar{u}^j g^j(z)\},$$

$$\text{ii)}\quad \bar{u}^j g^j(\tilde{z}) = 0 \quad j = 1,\cdots,p .$$

This theorem is a particular case of the strong duality theorem stated in Geoffrion [3]. Related theorems can be found in Rockafellar [10], Mangasarian [7]. Note that if there exists a solution z^* to $\min_{z\in\mathbb{C}}\{g^0(z) + \sum_{j=1}^{p} \bar{u}^j g^j(z)\}$, in the interior of the set $\mathbb{C}$ then for any $\hat{z} \in S_p$.

1-11 $$g^0(\hat{z}) = \max_{\substack{u\geq 0\\ z\in\mathbb{C}}}\{g^0(z) + \sum_{j=1}^{p} u^j g^j(z) \,\Big|\, \nabla g^0(z) + \sum_{j=1}^{p} u^j \nabla g^j(z) = 0\}$$

This important observation is a consequence of the fact, that if $z^* \in \overset{0}{\mathbb{C}}$, and $\tilde{u} \geqq 0$ satisfy $\nabla g^0(z^*) + \sum_{j=1}^{p} \tilde{u}^j \nabla g^j(z^*) = 0$, then z^* is a solution of $\min_{z\in\mathbb{C}}\{g^0(z) + \sum_{j=1}^{p} \tilde{u}^j g^j(z)\}$. Finally, note that when $\mathbb{C}$ is open in R^n, the solutions of $\min_{z\in\mathbb{C}}\{g^0(z) + \sum_{j=1}^{p} \tilde{u}^j g^j(z)\}$ always lie in the interior of $\mathbb{C}$.

Therefore equations 1-11 always holds for $\mathbb{C}$ open, $\hat{z}$ in S_p. $\square$

Before stating the lemma which relates 1-1 and 1-2, we shall prove, under a simple assumption, that an optimal multiplier for the minimization problem 1-13, stated below, cannot have zero as its first component.

1-12 Lemma

Consider the problem

1-13 $$\min\{f^0(z) \mid f^j(z) \leq 0 \ j = 1,\cdots, m\},$$

where f^j: $R^n \to R$, $j = 0, 1, \ldots, m$, are convex and continuously differentiable functions. Suppose that 1-13 has at least one solution and that

1-14 $$\mathbb{C}' = \{z \mid f^j(z) < 0 \ j = 1,\cdots, m\}$$

is not empty. Then, for any solution $\hat{z}$ of 1-13, the set $\Lambda(\hat{z})$ consisting of optimal multipliers $\lambda = (\lambda^0,\lambda^1,\ldots,\lambda^m) \in R^{m+1}$, and defined by

1-15 $$\Lambda(\hat{z}) = \{\lambda \in R^{m+1} \mid \sum_{j=1}^{m} \lambda^j f^j(\hat{z}) = 0; \sum_{j=0}^{m} \lambda^j \nabla f^i(\hat{z}) = 0; \sum_{j=0}^{m} \lambda^j = 1; \lambda \geq 0\}$$

is such that

1-16 $$\overline{\lambda}^0 = \min\{ \langle \lambda,e \rangle \mid \lambda \in \Lambda(\hat{z})\} > 0,$$

where $e = (1, 0, 0,\cdots, 0) \in R^{m+1}$

Proof: From 1-15 and the fact that $\hat{z}$ is a solution of 1-13, $\Lambda(\hat{z})$ is a non empty [9] compact subset of R^{m+1} and therefore there exists a $\overline{\lambda} \in \Lambda(\hat{z})$ such that $\overline{\lambda}^0 = \langle \overline{\lambda},e \rangle$. Since the functions f^j, $j = 0, 1,\ldots, m$, are convex and continuously differentiable, we must have

1-17 $$f^j(z) \geq f^j(\hat{z}) + \langle \nabla f^j(\hat{z}),z-\hat{z} \rangle, \quad j = 0, 1,\cdots, m, \text{ for all } z \in R^n.$$

Multiplying 1-17 by $\overline{\lambda}^j$, for $j = 0, 1,\cdots, m$, and adding the results, we obtain

1-18 $$\sum_{j=0}^{m} \overline{\lambda}^j f^j(z) \geq \sum_{j=0}^{m} \overline{\lambda}^j f^j(\hat{z}) + \langle \sum_{j=0}^{m} \overline{\lambda}^j \nabla f^j(\hat{z}), z-\hat{z} \rangle$$

Thus, it follows from 1-16, 1-15 that

1-19 $$\sum_{j=0}^{m} \bar{\lambda}^j f^j(z) \geq \bar{\lambda}^0 f^0(\hat{z}). \quad \text{for all } z \in R^n.$$

Hence, if $\bar{\lambda}^0 = 0$, 1-19 contradicts assumption 1-14. Consequently the lemma must be true. □

Let us now state the result which relates 1-1 and 1-2.

1-20 Lemma

Consider problem 1-13, where the functions f^j, $j = 0, 1, \cdots, m$, are now assumed to be convex and twice continuously differentiable. Suppose that an algorithm, in solving 1-13, constructs an infinite sequence $\{z_i\}_{i=0}^{\infty}$ which converges to a solution $\hat{z}$ of 1-13, and, in addition, suppose that there exist $i_0 > 0$ and $k \in (0,1)$ such that

1-21 $$f^0(z_{i+1}) - f^0(\hat{z}) \leq k[f^0(z_i) - f^0(\hat{z})],$$

1-22 $$f^j(z_i) \leq 0 \quad j = 1, \cdots, m,$$

for all $i \geq i_0$. If there exists an optimal multiplier $\lambda \in \Lambda(\hat{z})$ (see 1-15) and constants $\ell > 0$, $\varepsilon > 0$, such that

1-23 $$\langle y, \sum_{j=0}^{m} \lambda^j \frac{\partial^2 f^j}{\partial z^2}(z)\, y \rangle \geq \ell \|y\|^2 \text{ for all } y \in R^n \text{ and for all } z$$

in $B(\hat{z},\varepsilon)$,

where $B(\hat{z},\varepsilon) = \{z \mid \|z-\hat{z}\| \leq \varepsilon\}$, then there exists an integer $i_1 > 0$ such that

1-24 $\quad \|z_i - \hat{z}\|^2 \le k^{i-i_1}[\frac{2\lambda^0}{\ell}(f^0(z_{i_1}) - f^0(\hat{z}))]$ for all $i \ge i_1$.

Proof: Since $\{z_i\}$ converges to $\hat{z}$, there exists an i_2 such that $z_i \in B(\hat{z},\varepsilon)$ for all $i \ge i_2$. Let $i_1 = \max(i_0, i_2)$. Without loss of generality, we may assume that $i_1 = 0$. According to the Taylor expansion formula, for any $z_i \in \{z_i\}$ and any $\lambda \in \Lambda(\hat{z})$, there exists a $\theta_\lambda(z_i) \in (0,1)$ such that

1-25 $$\sum_{j=0}^{m} \lambda^j [f^j(z_i) - f^j(\hat{z})] = \langle z_i-\hat{z}, \sum_{j=0}^{m} \lambda^j \nabla f^j(\hat{z}) \rangle + \frac{1}{2}\langle z_i-\hat{z}, \sum_{j=0}^{m} \lambda^j \frac{\partial^2 f^j}{\partial z^2}(\xi)(z_i-\hat{z})\rangle$$

with $\xi = \theta_\lambda(z_i)z_i + [1-\theta_\lambda(z_i)]\hat{z}$.

Making use of the fact that $\lambda \in \Lambda(\hat{z})$, 1-25 becomes

1-26 $$\sum_{j=1}^{m} \lambda^j f^j(z_i) + \lambda^0[f^0(z_i) - f^0(\hat{z})] = \frac{1}{2}\langle z_i-\hat{z}, \sum_{j=0}^{m} \lambda^j \frac{\partial^2 f^j}{\partial z^2}(\xi)(z_i-\hat{z})\rangle .$$

Therefore, it follows from 1-22 and 1-23 that for any $\lambda \in \Lambda(\hat{z})$ for which 1-23 holds,

1-27 $$\lambda^0[f^0(z_i) - f^0(\hat{z})] \ge \frac{\ell}{2}\|z_i-\hat{z}\|^2.$$

Thus, by induction, 1-21 and 1-27 imply that

$$\|z_i-\hat{z}\|^2 \le \frac{2\lambda^0}{\ell} k^i (f^0(z_0) - f^0(\hat{z})),$$

which completes our proof. □

Note that hypothesis 1-23 can be replaced by the requirement that hypothesis 1-14 is satisfied and that there exists $m^0 > 0$ and $\varepsilon > 0$ such that

1-28 $$\langle y, \frac{\partial^2 f^0}{\partial z^2}(z)y \rangle \ge m^0\|y\|^2 \quad \text{for all } y \text{ in } R^n \text{ and all } z \text{ in } B(\hat{z},\varepsilon)$$

Indeed, 1-28 implies that

1-29 $$\langle y, \sum_{j=1}^{m} \lambda^j \frac{\partial^2 f}{\partial z^2}(z)y \rangle \geq m^0\lambda^0\|y\|^2 \text{ for all } y \text{ in } R^n, \text{ for all } z \text{ in } B(\hat{z},\varepsilon)$$

and for all $\lambda \in \Lambda(\hat{z})$.

Thus 1-23 is satisfied with $\ell = m^0 \bar{\lambda}^0$ where $\bar{\lambda}^0$ is defined by 1-16 and must satisfy $\bar{\lambda}^0 > 0$ because of 1-12.

Because lemma 1-20 requires 1-23 or 1-28 to be satisfied, it is convenient to introduce the following definition.

1-30 <u>Definition</u>: (linear convergence in cost)

Given a cost function f^0: $R^n \to R$, we say that the sequence $\{z_i\}_{i=0}^{\infty}$ verges to $\hat{z}$ at least linearly in cost, if there exist $k \in (0,1)$, $\kappa > 0$ and $i_0 \geq 0$ such that $|f^0(z_i) - f^0(\hat{z})| \leq k^i\kappa$ for all $i \geq i_0$.

We, now, conclude this section by giving the following result which characterizes the optimal points of problem 1-13.

1-31 <u>Lemma</u>

Consider the problem

1-32 $$\min\{f^0(z) | f^j(z) \leq 0 \quad j = 1,\cdots, m\}$$

where f^j: $R^n \to R$, $j = 0, 1,\ldots, m$, are convex and continuously differentiable functions. Let S be the set of solutions of 1-32, i.e.,

$$S = \{\hat{z} | f^0(\hat{z}) = \min\{f^0(z) | f^j(z) \leq 0 \ j = 1,\cdots, m\};\ f^j(\hat{z}) \leq 0,\ j = 1,\ldots,m\},$$

and suppose that the sets S and $\mathbb{C}' = \{z | f^j(z) < 0 \ j = 1,\ldots,m\}$ are not empty. Define k^0: $R^n \to R$ by

1-33 $$k^0(z) = \min_{(h^0,h)}\{h^0 + \frac{1}{2}\|h\|^2 \mid \langle \nabla f^0(z),h\rangle \le h^0;\ f^j(z) + \langle \nabla f^j(z),h\rangle \le h^0\ j = 1,\cdots, m\}.$$

Then

i) k^0 is well defined and is continuous on R^n,

ii) $k^0(z) < 0$ for all z in $\mathbb{C}$, z not in S; and $k^0(z) = 0$ for all z in S.

Proof: Note that from 1-33

1-34 $$k^0(z) = \min_h\{\frac{1}{2}\|h\|^2 + \max\{\langle \nabla f^0(z),h\rangle;\ f^j(z) + \langle \nabla f^j(z),h\rangle,\ j = 1,\cdots,m\}\}$$

Let $\phi(z,h) = \frac{1}{2}\|h\|^2 + \max\{\langle \nabla f^0(z),h\rangle;\ f^j(z) + \langle \nabla f^j(z),h\rangle,\ j = 1,\cdots, m\}$, then ϕ is continuous on $R^n \times R^n$ and convex in h. Furthermore $\lim_{t\to+\infty} \phi(z,th) = +\infty$ for all $(z,h) \in R^n$. Therefore, for any $z \in R^n$ there exists a vector $h(z) \in R^n$, such that $k^0(z) = \phi(z,h(z))$. Hence, $k^0(\cdot)$ is well defined.

It is straightforward to verify that the conditions of the strong duality theorem 1-5 are satisfied by problem 1-33. Therefore from 1-5(a)(ii).

1-35 $$k^0(z) = \max_{u\geqq 0}\{\inf_{(h^0,h)}\{(1 - \sum_{j=0}^{m} u^j)h^0 + \frac{1}{2}\|h\|^2 + \sum_{j=1}^{m} u^j f^j(z) + \sum_{j=0}^{m} u^j \langle \nabla f^j(z),h\rangle\}\}.$$

Next, from 1-11 and 1-5(a)(i),

1-36 $$k^0(z) = \max_{\substack{u\geqq 0\\(h^0,h)}}\{(1 - \sum_{j=0}^{m} u^j)h^0 + \frac{1}{2}\|h\|^2 + \sum_{j=0}^{m} u^j \langle \nabla f^j(z),h\rangle + \sum_{j=1}^{m} u^j f^j(z) \mid \sum_{j=0}^{1} u^j = 1;\ \sum_{j=0}^{m} u^j \nabla f^j(z) + h = 0\},$$

which is equivalent to

$$\text{1-37} \qquad k^0(z) = \max_{u \geq 0}\{\sum_{j=1}^{m} u^j f^j(z) - \frac{1}{2}\|\sum_{j=0}^{m} u^j \nabla f^j(z)\|^2 | \sum_{j=0}^{m} u^j = 1\}.$$

Thus, from 1-37 and the maximum theorem (Berge [1] pp. 116) $k^0(\cdot)$ is continuous.

It follows also from 1-37 that k^0 is negative in $\mathbb{C}$. Suppose now that $\bar{z} \in C$ and that $k^0(\bar{z}) = 0$, then from 1-37 there exists a $\bar{u} \geqq 0$ such that

$$\text{1-38} \qquad \sum_{j=1}^{m} \bar{u}^j f^j(\bar{z}) = 0; \sum_{j=0}^{m} \bar{u}^j \nabla f^j(\bar{z}) = 0; \sum_{j=0}^{m} \bar{u}^j = 1; \bar{u} \geqq 0. \; .$$

From the convexity of the functions f^j, $j = 0, 1..., m$,

$$\text{1-39} \qquad f^j(z) \geq f^j(\bar{z}) + \langle \nabla f^j(\bar{z}), z-\bar{z} \rangle \text{ for all } z \in R^n, \; j = 0, 1, \cdots, m.$$

Therefore

$$\text{1-40} \qquad \sum_{j=0}^{m} \bar{u}^j f^j(z) \geq \sum_{j=0}^{m} \bar{u}^j f^j(\bar{z}) + \langle \sum_{j=0}^{m} \bar{u}^j \nabla f^j(\bar{z}), z-\bar{z} \rangle \text{ for all } z \in R^n,$$

which becomes, because of 1-38,

$$\text{1-41} \qquad \sum_{j=1}^{m} \bar{u}^j f^j(z) \geq \bar{u}^0 (f^0(\bar{z}) - f^0(z)) \text{ for all } z \in R^n$$

If $\bar{u}^0 = 0$, it follows from 1-41 that $\mathbb{C}' = \phi$. Thus $\bar{u}^0 > 0$ and from 1-41, and the fact that $\bar{u} \geqq 0$, we conclude that $\bar{z}$ is a solution of 1-32. Conversely, if $\hat{z}$ is a solution of 1-32, there exists an optimal multiplier $\lambda \in R^{m+1}$ such that

1-42 $$\sum_{j=1}^{m} \lambda^j f^j(\hat{z}) = 0; \sum_{j=0}^{m} \lambda^j \nabla^j(\hat{z}) = 0; \sum_{j=0}^{m} \lambda^j = 1; \lambda \geqq 0$$

and 1-42 and 1-37 imply that $k^0(\hat{z}) = 0$, which completes our proof. $\square$

SECTION II. RATE OF CONVERGENCE OF A THEORETICAL METHOD OF CENTERS.

The method of centers (Huard [4]), to be studied in this section is not implementable on a computer. However it is of great theoretical interest since it leads naturaly to several implementable algorithms, known as modified methods of centers. One of these method of centers will be studied in the next section. In this section we shall present two theorems. The first theorem concludes that this method of centers under examination converges at least linearly. The second theorem shows that this method of centers converges at most linearly. We begin by recalling the algorithms and conditions for its convergence.

Consider the problem

2-1 $$\min\{f^0(z) \mid f^j(z) \leq 0,\ j = 1,\cdots, m\}$$

where $f^j\colon\ R^n \to R$, $j = 0, 1,\cdots, m$, are continuous functions.

2.2 <u>Assumptions</u>

We shall assume that there exists $z_0 \in \mathbb{C} = \{z \mid f^j(z) \leq 0\ j = 1,\cdots, m\}$ such that

2-3 i) the set $\{z \in \mathbb{C} \mid f^0(z) \leq f^0(z_0)\}$ is compact convex

ii) there exists a compact convex set $\mathbb{C}(z_0)$ containing $\{z \in \mathbb{C} \mid f^0(z) \leq f^0(z_0)\}$ in its interior such that the functions f^j_i 0, $1,\cdots,m$, are convex and continuously differentiable in $\mathbb{C}(z_0)$ and such that

iii) the function f^0 is strictly convex in $\mathbb{C}(z_0)$.

2-4 iv) $\mathbb{C}' = \{z | f^j(z) < 0 \; j = 1,\cdots, m\}$ is not empty

2-5 Algorithm (Huard [4])[2]

Step 0. Select a z_0 such that 2-2 is satisfied, and set $i = 0$.

Step 1. Compute a solution $(\delta(z_i), z_{i+1})$ of

2-6 $$\min_{(\delta,z)}\{\delta | f^0(z) - f^0(z_i) \leq \delta;\; f^j(z) \leq \delta,\; j = 1,\cdots, m;\; z \in \mathbb{C}(z_0)\}$$

Step 2. If $\delta(z_i) = 0$, set $\hat{z} = z_i$, and stop; else set $i = i+1$ and go to Step 1.

First, if the algorithm stops, then $\hat{z}$ must be a solution of 2-1 (see Polak [8] theorem 4.2.12). Next, from hypothesis 2-2 (i), (iv), and theorem 4.2.12 in Polak [8], every accumulation point of a sequence $\{z_i\}_{i=0}^{\infty}$, generated by algorithms 2-5 in solving 2-1, is a solution of 2-1. Now assumptions (i) and (iv) also imply that 2-1 has at least one solution. From 2-2 (iii) this solution is unique. From the compactness of $\mathbb{C}(z_0)$, $\{z_i\}_{i=0}^{\infty}$ has at least one accumulation point. Thus $\{z_i\}_{i=0}^{\infty}$ converges to the solution of problem 2-1.

Let $\hat{z}$ be the solution of 2-1 and let $\Lambda(\hat{z})$ be the set of optimal multipliers at $\hat{z}$, i.e.,

2-7 $$\Lambda(\hat{z}) = \{\lambda \in R^{m+1} | \sum_{j=0}^{m} \lambda^j \nabla f^j(\hat{z}) = 0;\; \sum_{j=1}^{m} \lambda^j f^j(\hat{z}) = 0;\; \sum_{j=0}^{m} \lambda^j = 1;\; \lambda \geqq 0\}$$

Referring to lemma 1-12 we see that

2-8 $$\bar{\lambda}^0 = \min\{ \langle \lambda, e \rangle | \lambda \in \Lambda(\hat{z});\; e = (1, 0, 0,\cdots, 0) \in R^{m+1}\} > 0.$$

2-9 Theorem (at least linear convergence in cost)

Let $\{z_i\}_{i=0}^{\infty}$ be an infinite sequence generated by algorithm 2-5, in solving problem 2-1, and suppose that assumptions 2-2 are satisfied. Then, given any $\alpha \in (0,1)$, there exists an integer $i_0(\alpha)$ such that

2-10 $$f^0(z_{i+1}) - f^0(\hat{z}) \leq [1 - \overline{\lambda}^0(1-\alpha)][f^0(z_i) - f^0(\hat{z})] \text{ for all } i \geq i_0(\alpha),$$

where $\hat{z}$ is the unique solution of 2-1.

Proof. To obtain a bound on $f^0(z_{i+1}) - f^0(\hat{z})$ we shall investigate the quantity $f^0(z_{i+1}) - f^0(z_i)$, which, according to 2-6, in the algorithm, is bounded by $\delta(z_i)$, i.e.,

2-11 $$f^0(z_{i+1}) - f^0(z_i) \leq \delta(z_i).$$

Then, using the strong duality theorem 1-5, we shall find a bound on $\delta(z_i)$ in terms of z_i and $f^0(z_i) - f^0(\hat{z})$. To complete the proof we shall eliminate the dependence on z_i of the bound on $\delta(z_i)$.

Thus from 2-6

2-12 $$\delta(z_i) = \min_{(\delta,z)}\{\delta \mid f^0(z) - f^0(z_i) \leq \delta;\ f^j(z) \leq \delta,\ j = 1, \ldots, m;\ z \in \mathbb{C}(z_0)\}$$

The set $\{(z,\delta) \mid f^0(z) - f^0(z_i) - \delta < 0;\ f^j(z) - \delta < 0,\ j = 1,\cdots, m\}$ is not empty for all $z_i \in \mathbb{C}(z_0)$. Therefore the strong duality theorem 1-5 can be applied to 2-12. Parts (a)(i) and (ii) of 1-5, as applied to 2-12, imply that

2-13 $$\delta(z_i) = \max_{u>0}\{\inf_{\substack{z \in \mathbb{C}(z_0) \\ \delta \in R}} \{(1 - \sum_{j=0}^{m} u^j)\delta + \sum_{j=1}^{m} u^j f^j(z) + u^0[f^0(z) - f^0(z_i)]\}\}.$$

Let $\overline{u}(z_i) = (\overline{u}^0(z_i),\cdots,\overline{u}^m(z_i)) \in R^{m+1}$ be a solution of 2-13, then from (a)(iii) of the strong duality theorem 1-5

2-14 $$\delta(z_i) = \min_{(\delta,z)}\{(1 - \sum_{j=0}^{m}\overline{u}^j(z_i))\delta + \sum_{j=1}^{m}\overline{u}^j(z_i)f^j(z) + \overline{u}^0(z_i)[f^0(z) - f^0(z_i)] \mid z \in \mathbb{C}(z_0)\}.$$

Clearly, equation 2-14 cannot hold unless

2-15 $$\sum_{j=0}^{m}\overline{u}^j(z_i) = 1\cdot$$

Consequently,

2-16 $$\delta(z_i) = \min_{z \in \mathbb{C}(z_0)} \{\sum_{j=1}^{m}\overline{u}^j(z_i)f^j(z) + u^0(z_i)[f^0(z) - f^0(z_i)]\}.$$

Upon replacing z by $\hat{z}$ in 2-16, we obtain the following bound on $\delta(z_i)$:

2-17 $$\delta(z_i) \leq \sum_{j=1}^{m}\overline{u}^j(z_i)f^j(\hat{z}) + \overline{u}^0(z_i)[f^0(\hat{z}) - f^0(z_i)].$$

Since $\hat{z} \in \mathbb{C}$, $f^j(\hat{z}) \leq 0, j = 1,\cdots, m$, which, together with the fact that $\overline{u}(z_i) \geqq 0$ implies that $\sum_{j=1}^{m}\overline{u}^j(z_i)f^j(z_i) \leq 0$. Hence,

2-18 $$\delta(z_i) \leq \overline{u}^0(z_i)[f^0(\hat{z}) - f^0(z_i)].$$

For every $z_i \in \{z_i\}_{i=0}^{\infty}$, we define $U(z_i)$ to be the set of solutions of 2-13. Let us show that every sequence $\{\overline{u}^0(z_i)\}_{i=0}^{\infty}$, consisting of the first components of vectors $\overline{u}(z_i) \in U(z_i)$, must always satisfy

2-19 $$\liminf_{i\to\infty} \bar{u}^0(z_i) \geq \bar{\lambda}^0,$$

where $\bar{\lambda}^0$ is defined by 2-8 and, according to lemma 1-12, satisfies $\bar{\lambda}^0 > 0$. Let ψ: $R^{m+1} \times R^n \to R$ be defined by

2-20 $$\psi(u,z_i) = \min_{z \in \mathbb{C}(z_0)} \{\sum_{j=1}^{m} u^j f^j(z) + u^0[f^0(z) - f^0(z_i)]\}$$

Because $\mathbb{C}(z_0)$ is compact, ψ is well defined and continuous in both arguments. Now, from 2-13, 2-14 and 2-15,

2-21 $$\delta(z_i) = \max_u\{\psi(u,z_i)|u \geqq 0; \sum_{j=0}^{m} u^j = 1\}.$$

Let Γ: $\mathbb{C} \cap \mathbb{C}(z_0) \to \mathcal{P}(R^{m+1})$(the set of all subsets of R^{m+1}) be a map defined by

2-23 $$\Gamma(z_i) = \{u \in R^{m+1}|\psi(u,z_i) = \delta(z_i);\ u \geqq 0; \sum_{j=0}^{m} u^j = 1\}$$

Let $\{\bar{u}(z_i)\}_{i=0}^{\infty}$ be any sequence such that $\bar{u}(z_i) \in U(z_i)$ for $i = 0, 1, 2,\cdots$. Because of 2-20, 2-21 and 2-23, $\bar{u}(z_i) \in U(z_i)$ implies that $\bar{u}(z_i) \in \Gamma(z_i)$. The set $\{u \in R^{m+1}|\sum_{j=0}^{m} u^j = 1,\ u \geqq 0\}$ is compact, and from 2-23 $\Gamma(z_i)$ is contained in that set for $i = 0, 1, 2,\cdots$. Therefore $\{\bar{u}(z_i)\}_{i=0}^{\infty}$ has at least one accumulation point. Let $\tilde{u} \in R^{m+1}$ be an accumulation point of $\{\bar{u}(z_i)\}_{i=0}^{\infty}$ and let $\{\bar{u}(z_{i_j})\}_{j=0}^{\infty}$ be a subsequence of $\{\bar{u}(z_i)\}_{i=0}^{\infty}$ which converges to $\tilde{u}$. From the continuity of ψ and δ, from 2-23 and from the fact that $z_{i_j} \to \hat{z}$ as $j \to \infty$, we conclude that $\tilde{u} \in \Gamma(\hat{z})$. Consequently, $\liminf_{i \to \infty} \bar{u}^0(z_i) \in \{ \langle u,e \rangle \mid u \in \Gamma(\hat{z});\ e = (1, 0, \cdots, 0) \in R^{m+1}\}$[3]

From 2-23,

$$\text{2-24} \qquad \Gamma(\hat{z}) = \{\lambda \in R^{m+1} \mid \delta(\hat{z}) = \min_{z \in \mathbb{C}(z_0)} \{\lambda^0[f^0(z) - f^0(\hat{z})] + \sum_{j=1}^{m} \lambda^j f^j(z)\};\ \sum_{j=0}^{m} \lambda^j = 1;\ \lambda \geqq 0\}.$$

We shall show that $\Gamma(\hat{z}) = \Lambda(\hat{z})$, where

$$\text{2-25} \qquad \Lambda(\hat{z}) = \{\lambda \in R^{m+1} \mid \sum_{j=1}^{m} \lambda^j f^j(\hat{z}) = 0;\ \sum_{j=0}^{m} \lambda^j \nabla^j(\hat{z}) = 0;\ \sum_{j=0}^{m} \lambda^j = 1;\ \lambda \geqq 0\}.$$

Since $\hat{z}$ is a solution of problem 2-1, $\delta(\hat{z}) = 0$. Therefore let us investigate the solutions λ of the equation

$$\text{2-26} \qquad \min_{z \in \mathbb{C}(z_0)} \{\lambda^0[f^0(z) - f^0(\hat{z})] + \sum_{j=1}^{m} \lambda^j f^j(\hat{z})\} = 0$$

which also satisfy $\lambda \geqq 0$, $\sum_{j=0}^{m} \lambda^j = 1$.

Let $\tilde{\lambda}$ be such a solution. If $\tilde{\lambda}^0 = 0$, equation 2-26 cannot be satisfied (see hypothesis 2-4). Therefore $\tilde{\lambda}^0 > 0$. From assumptions 2-2(ii) and (iii), and the fact that $\tilde{\lambda}^0 > 0$, the problem $\min_{z \in \mathbb{C}(z_0)} \{\tilde{\lambda}^0[f^0(z) - f^0(\hat{z})] + \sum_{j=1}^{m} \tilde{\lambda}^j f^j(z)\}$ has a unique solution, and the optimal value can be zero only if $\hat{z}$ is the solution (because $\sum_{j=1}^{m}\tilde{\lambda}^j f^j(\hat{z}) = \tilde{\lambda}^0[f^0(\hat{z}) - f^0(\hat{z})]$ $+\sum_{j=1}^{m}\tilde{\lambda}^j f^j(\hat{z}) \leq 0$). Hence $\sum_{j=1}^{m}\tilde{\lambda}^j f^j(\hat{z}) = 0$. Furthermore, since $\hat{z}$ belongs to the interior of $C(z_0)$, $\sum_{j=0}^{m}\tilde{\lambda}^j \nabla f^j(\hat{z}) = 0$. Conversely, $\sum_{j=1}^{m}\tilde{\lambda}^j f^j(\hat{z}) = 0$ and $\sum_{j=0}^{m}\tilde{\lambda}^j \nabla f^j(\hat{z}) = 0$ implies that 2-26 is satisfied by $\tilde{\lambda}$. Thus $\Gamma(\hat{z}) = \Lambda(\hat{z})$.

Because of the fact that $(\liminf_{i\to\infty} \bar{u}^0(z_i)) \in \{ \langle u,e \rangle \mid u \in \Gamma(\hat{z})$, $e = (1,0,\cdots,0) \in R^{m+1}\}$ implies that, given any $\alpha \in (0,1)$, there exists $i_0(\alpha)$ such that

2-27 $$\bar{u}^0(z_i) \geq \bar{\lambda}^0(1-\alpha) \quad \text{for all } i \geq i_0(\alpha).$$

Therefore, from 2-18 and 2-27 we obtain that

2-28 $$\delta(z_i) \leq \bar{\lambda}^0(1-\alpha)[f^0(\hat{z}) - f^0(z_i)] \quad \text{for all } i \geq i_0(\alpha)$$

and hence, recalling 2-11, we must have,

2-29 $$f^0(z_{i+1}) - f^0(z_i) \leq \bar{\lambda}^0(1-\alpha)[f^0(\hat{z}) - f^0(z_i)] \quad \text{for all } i \geq i_0(\alpha).$$

It now follows that

2-30 $$f^0(z_{i+1}) - f^0(\hat{z}) \leq [1 - \bar{\lambda}^0(1-\alpha)][f^0(z_i) - f^0(\hat{z})] \text{ for all } i \geq i_0(\alpha),$$

which completes our proof. $\square$

We now establish an upper bound on the rate of convergence of algorithm 2-5.

2-31 <u>Theorem</u>

Let $\{z_i\}_{i=0}^{\infty}$ be an infinite sequence generated by algorithm 2-5 in solving problem 2-1. Let $\hat{z}$ be the solution of 2-1 and let $\bar{\lambda}^0$ be defined by 2-8. If assumptions 2-2 are satisfied, then, either $\bar{\lambda}^0 = 1$ and the sequence $\{f^0(z_i)\}_{i=0}^{\infty}$ converges superlinearly to $f^0(\hat{z})$, or $\bar{\lambda}^0 < 1$ and there exists an integer i_1 such that

2-32 $$f^0(z_{i+1}) - f^0(\hat{z}) \geq (1-\bar{\lambda}^0)[f^0(z_i) - f^0(\hat{z})] \text{ for all } i \geq i_1.$$

<u>Proof</u>: Applying part (b)(ii) of the strong duality theorem 1-5 to problem 2-6, we conclude that a $\bar{u}^0(z_i)$ defined as the first component of a solution of 2-13 must satisfy

2-33 $$\bar{u}^0(z_i)[f^0(z_{i+1}) - f^0(z_i) - \delta(z_i)] = 0.$$

Now, according to 2-27, $\bar{u}^0(z_i) \geq (1-\alpha)\bar{\lambda}^0$ for all $i \geq i_0(\alpha)$. Therefore, by making use of lemma 1-12, $\bar{u}^0(z_i) > 0$ for all $i \geq i_0(\alpha)$. Hence, from 2-33, we obtain that

2-34 $$f^0(z_{i+1}) - f^0(z_i) = \delta(z_i).$$

Next, according to 2-13, for $i = 0,1,2,\ldots,$

2-35 $$\delta(z_i) = \max_{u \geqq 0}\{\inf_{\substack{z \in \mathbb{C}(z_0) \\ \delta}} \{(1 - \sum_{j=0}^{m} u^j)\delta + \sum_{j=1}^{m} u^j f^j(z) + u^0[f^0(z) - f^0(z_i]\}\}.$$

Setting $u = \lambda$, some element of $\Lambda(\hat{z})$ (defined by 2-7), we conclude that

2-36 $$\delta(z_i) \geq \inf_{z \in \mathbb{C}(z_0)} \{\sum_{j=1}^{m} \lambda^j f^j(z) + \lambda^0[f^0(z) - f^0(z_i)]\}$$

The infimum in 2-36 is achieved at $\hat{z}$ because $\sum_{j=0}^{m} \lambda^j \nabla f^j(\hat{z}) = 0$. Therefore

$$\delta(z_i) \geq \sum_{j=1}^{m} \lambda^j f^j(\hat{z}) + \lambda^0[f^0(\hat{z}) - f^0(z_i)] = \lambda^0[f^0(\hat{z}) - f^0(z_i)] \text{ for all } \lambda \in \Lambda(\hat{z}).$$

Thus

2-37 $$\delta(z_i) \geq \overline{\lambda}^0[f^0(\hat{z}) - f^0(z_i)], \text{ for } i = 0,1,2,\ldots$$

and setting $i_1 = i_0(\alpha)$, for some $\alpha \in (0,1)$,

2-38 $$f^0(z_{i+1}) - f^0(\hat{z}) \geq (1-\overline{\lambda}^0)[f^0(z_i) - f^0(\hat{z})], \text{ for } i \geq i_1,$$

which proves the second part of the theorem; the first part follows directly from theorem 2-9. This completes our proof. □

Combining 2-10 and 2-32, we see that whenever algorithm 2-5 constructs an infinite sequence $\{z_i\}_{i=0}^{\infty}$, we must have

2-39 $$\lim_{i\to\infty} \frac{f^0(z_{i+1}) - f^0(\hat{z})}{f^0(z_i) - f^0(\hat{z})} = 1 - \overline{\lambda}^0,$$

with $\overline{\lambda}^0$ given by 2-8. This situation is rather unique. In the following section we shall be able to obtain only an upper bound on $\lim_{i\to\infty} \frac{f^0(z_{i+1}) - f^0(\hat{z})}{f^0(z_i) - f^0(\hat{z})}$.

SECTION III. MODIFIED METHOD OF CENTERS.

The method of centers 2-5 requires that at each iteration we solve the problem

$$\text{3-1}\qquad \min_{(\delta,z)}\{\delta \mid f^0(z) - f^0(z_i) \le \delta;\ f^j(z) \le \delta,\ j = 1,\cdots, m\}\ .$$

In the modified method of centers [5], this problem is replaced by two subproblems: a direction finding subproblem,

$$\text{3-2}\qquad \min_{(h^0,h)}\{h^0 \mid \langle \nabla f^0(z_i),h\rangle \le h^0;\ f^j(z_i) + \langle \nabla f^j(z_i),h\rangle \le h^0,\ j = 1,\cdots, m;\ h \in S\},$$

where $S = \{h \in R^n \mid |h^i| \le 1,\ i = 1, 2,\cdots, n\}$, and a step size determination subproblem,

$$\text{3-3}\qquad \min_{(\delta_h,\mu)}\{\delta_h \mid f^0(z_i + \mu h(z_i)) - f^0(z_i) \le \delta_h;\ f^j(z_i + \mu h(z_i)) \le \delta_h,\ j = 1,\cdots, m\}.$$

Referring to the example in the appendix, we see that the algorithm using 3-2 and 3-3 does not converge linearly. Therefore, we modified 3-2 to the following form:

$$\text{3-4}\qquad \min_{(h^0,h)}\{h^0 + \tfrac{1}{2}\|h\|^2 \mid \langle \nabla f^0(z_i),h\rangle \le h^0;\ f^j(z_i) + \langle \nabla f^j(z_i),h\rangle \le h^0,\ j = 1, \cdots, m\}.$$

Problem 3-2 is a linear program with m+n+1 constraints, while 3-4 is a quadratic program with m+1 constraints. Although 3-4 is harder to solve than 3-2, it makes 3-3 easier to implement. The reason for this is (as we shall show) that an algorithm, which uses 3-4 and solves 3-3 by means of a Golden Section search of finite precision, converges, while such a result is not true when 3-2 is used.

We shall first establish linear convergence for the modified method of centers which uses the subproblems 3-3 and 3-4, with respect to problems with strictly convex inequality constraints. Then we shall show that this result is also true for the case of convex constraints. The last part of the section is devoted to establishing linear convergence for an implementable algorithm which uses a Golden Section search of finite precision to solve 3-3.

Consider the problem

3-5 $$\min\{f^0(z) \mid f^j(z) \leq 0 \;\; j = 1,\cdots, m\},$$

where f^j: $R^n \to R$, $j = 0, 1,\ldots, m$, are twice continuously differentiable functions. Let $\mathbb{C}$ be the set of feasible points, i.e.,

3-6 $$\mathbb{C} = \{z \mid f^j(z) \leq 0, \; j = 1,\cdots, m\}$$

3-7 Assumptions.

We shall suppose that

i) there exists z_0 in $\mathbb{C}$ such that $\mathbb{C}(z_0) = \{z \in \mathbb{C} \mid f^j(z) \leq f^0(z)\}$ is bounded;

ii) the set $\mathbb{C}' = \{z \mid f^j(z) < 0 \;\; j = 1,\cdots, m\}$ is not empty;

iii) the functions f^j, $j = 0, 1,\cdots, m$, are convex[4]

iv) there exist $m^j > 0$, $j = 0, 1,\cdots, m$,[5] and $\varepsilon > 0$ such that

3-8 $$m^j\|y\|^2 \leq \langle y, \frac{\partial^2 f^j}{\partial z^2}(z)y \rangle \text{ for all } y \in R^n \text{ and all } z \in B(\hat{z},\varepsilon),$$

where $\hat{z}$ is the solution of problem 3-5 and $B(\hat{z},\varepsilon) = \{z \mid \|\hat{z}-z\| \leq \varepsilon\}$.

3-9 Algorithm (modified method of centers)

Step 0. Set $i = 0$.

Step 1. Compute a solution $(h^0(z_i)\ h(z_i))$ of the problem

3-10 $$\min\{h^0 + \frac{1}{2}\|h\|^2 \mid \langle \nabla f^0(z_i),h \rangle \le h^0;\ f^j(z_i) + \langle \nabla f^j(z_i),h \rangle \le h^0,\ j = 1,\cdots, m\} .$$

Step 2. If $h^0(z_i) = 0$, set $\hat{z} = z_i$, and stop; else compute a solution $(\delta_h(z_i),\ \mu(z_i))$ of

3-11 $$\min_{(\delta_h,\mu)} \{\delta_h \mid f^0(z_i+\mu h(z_i)) - f^0(z_i) \le \delta_h;\ f^j(z_i+\mu h(z_i)) \le \delta_h\ j = 1,\cdots, m\} .$$

Step 3. Set $z_{i+1} = z_i + \mu(z_i)h(z_i)$ and go to step 1. □

Theorem. (Convergence)

Let $\{z_i\}$ be a sequence generated by algorithm 3-9, in solving problem 3-5. If assumptions 3-7 are satisfied, then, either $\{z_i\}$ is finite and its last element $\hat{z}$, is the solution of problem 3-5, or else $\{z_i\}$ converges to $\hat{z}$. □

We shall omit giving a proof of this theorem since it can be deduced easily from theorem 4.2.32 in Polak [8] and lemma 1-31.

With $\varepsilon > 0$, $\hat{z}$ and m^j, $j = 0, 1,\cdots, m$, as in 3-8, we define

3-12 $$M^j = \max\{\| \frac{\partial^2 f^j}{\partial z^2}(z)\| \mid z \in B(\hat{z},\varepsilon)\},$$

3-13 $$L = \max\{1;\ M^j,\ j = 0, 1,\cdots, m\},$$

3-14 $$\ell = \min|1;\ m^j,\ j = 0, 1,\cdots, m\}.$$

3-15 Theorem (linear convergence).

Let $\{z_i\}_{i=0}^{\infty}$ be a sequence generated by algorithm 3-9 in solving problem 3-5, whose solution is $\hat{z}$. Suppose that assumptions 3-7 are satisfied. Then $z_i \to \hat{z}$ linearly as $i \to \infty$, $f^0(z_i) \to f^0(\hat{z})$ linearly as $i \to \infty$, in accordance with the following bounds. Given any $\alpha \in (0,1)$, there exist $i_0(\alpha)$, $\kappa_0(\alpha) > 0$, such that

3-16 $$\|z_i - \hat{z}\| \leq [1 - \frac{\ell}{L} \bar{\lambda}^0 (1-\alpha)]^{1/2} \kappa_0(\alpha) \quad \text{for all } i \geq i_0(\alpha),$$

3-16' $$f^0(z_{i+1}) - f^0(\hat{z}) \leq [1 - \frac{\ell}{L} \bar{\lambda}^0 (1-\alpha)][f^0(z_i) - f^0(\hat{z})] \text{ for all } i \geq i_0(\alpha),$$

where

$$\bar{\lambda}^0 = \min\{ \langle \lambda, e \rangle \,|\, \sum_{j=0}^{m} \lambda^j \nabla f^j(\hat{z}) = 0; \sum_{j=1}^{m} \lambda^j f^j(\hat{z}) = 0; \sum_{j=0}^{m} \lambda^j = 1; \lambda \geq 0, \lambda \in R^{m+1}\},$$

$e = (1, 0, 0, \cdots 0) \in R^{m+1}$ and L, ℓ are defined by 3-13, 3-14. □

Let us outline our strategy for proving theorem 3-15. From 3-11 in the algorithm, we have

3-18 $$f^0(z_{i+1}) - f^0(z_i) \leq \delta_h(z_i).$$

We shall find an upper bound on $\delta_h(z_i)$ in two steps. First we shall show that $\delta_h(z_i) \leq \frac{1}{L} [h^0(z_i) + \frac{1}{2} \|h(z_i)\|^2]$. Then we shall prove that $h^0(z_i) + \frac{1}{2} \|h(z_i)\|^2 \leq \ell \delta(z_i)$, where $\delta(z_i)$ (see 2-12) is the optimal value of problem 3-1. Finally, making use of 2-28 (i.e. $\delta(z_i) \leq -\lambda^0(1-\alpha)[f^o(z_i) - f^0(\hat{z})]$), we shall obtain 3-16'. Inequality 3-16 will then follow from lemma 1-20. Since a part of the proof of theorem 3-15 will also be needed in proving subsequent theorems, we break up the proof of theorem 3-15 into three lemmas.

3-19 Lemma.

Suppose that assumptions 3-7 (i)-(iii) are satisfied. Let $\delta_h(z)$, $(h^0(z),h(z))$, $\hat{z}$ respectively be defined as solutions of 3-11, 3-10 and 3-5, and let L and $\mathbb{C}$ be defined by 3-13 and 3-6. Then there exists a $\gamma_1 > 0$ such that $\delta_h(z) \leq \frac{1}{L} [h^0(z) + \frac{1}{2} \|h(z)\|^2]$ for all $z \in B(\hat{z},\gamma_1) \cup \mathbb{C}$.

Proof: according to 3-11,

$$\text{3-20} \qquad \delta_h(z) = \min_{(\delta_h,\mu)}\{\delta_h | f^0(z+\mu h(z)) - f^0(z) \leq \delta_h;\ f^j(z+\mu h(z)) \leq \delta_h\ j = 1,\cdots, m\},$$

It is straightforward to show that the strong duality theorem 1-5 applies to 3-11. Therefore, if $z \in \mathbb{C}$ then, from 1-5 (a)(ii)

$$\text{3-21} \qquad \delta_h(z) = \max_{\omega\geqq 0}\{\inf_{(\delta_h,\mu)} \{(1 - \sum_{j=0}^{m}\omega^j)\delta_h + \sum_{j=1}^{m}\omega^j f^j(z+\mu h(z)) + \omega^0[f^0(z+\mu h(z)) - f^0(z)]\}\}$$

$$\text{3-22} \qquad = \max_{\omega\geqq 0}\{\inf_{\mu}\{\sum_{j=1}^{m}\omega^j f^j(z+\mu h(z)) + \omega^0[f^0(z+\mu h(z)) - f^0(z)]\} \Big| \sum_{j=0}^{m}\omega^j = 1\}$$

Applying the Taylor expansion formula and making use of equation 3-13, we obtain,

$$\text{3-23} \qquad \sum_{j=1}^{m}\omega^j f^j(z+\mu h(z)) + \omega^0[f^0(z+\mu h(z) - f^0(z)] \leq \sum_{j=1}^{m}\omega^j f^j(z) + \mu \sum_{j=0}^{m}\omega^j \langle \nabla f^j(z),h(z) \rangle + \frac{\mu^2}{2} L \|h(z)\|^2,$$

for all $\omega \in R^{m+1}$ such that $\omega \geqq 0$, $\sum_{j=0}^{m} \omega^j = 1$, and all (z,μ) such that $z + \mu h(z) \in B(\hat{z},\varepsilon)$. It follows from the strong duality theorem 1-5, as applied to problem 3-10, that

3-24 $$h^0(z) + \frac{1}{2}\|h(z)\|^2 = \max_{u \geqq 0}\{\inf_{(h^0,h)} \{(1 - \sum_{j=0}^{m} u^j)h^0 + \sum_{j=1}^{m} u^j f^j(z) + \sum_{j=0}^{m} u^j \langle \nabla f^j(z), h \rangle + \frac{1}{2}\|h\|^2\}\}$$

3-25 $$= \max_{u \geqq 0}\{\inf_h\{\sum_{j=1}^{m} u^j f^j(z) + \sum_{j=0}^{m} u^j \langle \nabla f^j(z), h \rangle + \frac{1}{2}\|h\|^2\} \,|\, \sum_{j=0}^{m} u^j = 1\}$$

and hence that

3-26 $$h^0(z) + \frac{1}{2}\|h(z)\|^2 = \max_{u \geqq 0}\{\sum_{j=1}^{m} u^j f^j(z) + \sum_{j=0}^{m} u^j \langle \nabla f^j(z), h \rangle + \frac{1}{2}\|h\|^2 \,|\, \sum_{j=0}^{m} u^j = 1;\ \sum_{j=0}^{m} u^j \nabla f^j(z) + h = 0\}$$

Therefore

3-27 $$h^0(z) + \frac{1}{2}\|h(z)\|^2 = \max_{u \geqq 0}\{\sum_{j=1}^{m} u^j f^j(z) - \frac{1}{2}\|\sum_{j=0}^{m} u^j \nabla f^j(z)\|^2 \,|\, \sum_{j=0}^{m} u^j = 1\}.$$

Let $\bar{u} \in R^{m+1}$ be a solution of 3-27, then

3-28 $$h^0(z) + \frac{1}{2}\|h(z)\|^2 = \sum_{j=1}^{m} \bar{u}^j f^j(z) - \frac{1}{2}\|\sum_{j=0}^{m} \bar{u}^j \nabla f^j(z)\|^2,$$

and from 3-26 it follows that

3-29 $$h(z) = -\sum_{j=0}^{m} u^j \nabla f^j(z)^6$$

Hence, if $z \in \mathbb{C}$,

3-30 $$h^0(z) + \|h(z)\|^2 = \sum_{j=1}^{m} u^j f^j(z) \leq 0,$$

and therefore $\frac{1}{2}\|h(z)\|^2 \leq -[h^0(z) + \frac{1}{2}\|h(z)\|^2]$. From 3-27, it is clear that $h^0(\cdot) + \frac{1}{2}\|h(\cdot)\|^2$ is continuous, negative in $\mathbb{C}$ and that it takes the value of zero at $\hat{z}$. Therefore there exists $\gamma_1 \in (0, \frac{\varepsilon}{2}]$ such that $\max\{-[h^0(z) + \frac{1}{2}\|h(z)\|^2] \mid z \in B(\hat{z},\gamma_1) \cap \mathbb{C}\} \leq \frac{\varepsilon^2}{4}$. Hence, it follows that

3-31 $$\|h(z)\| \leq \frac{\varepsilon}{2} \text{ for all } z \in B(\hat{z}_1\gamma_1) \cap \mathbb{C}.$$

which implies that $z + \mu h(z) \in B(\hat{z},\varepsilon)$ for all $\mu \in [0,1]$ and all $z \in B(\hat{z},\gamma_1) \cap \mathbb{C}$. Thus 3-23 is valid for all $\omega \geqq 0$, such that $\sum_{j=0}^{m} \omega^j = 1$, all $\mu \in [0,1]$ and all $z \in B(\hat{z},\gamma_1) \cap \mathbb{C}$. Therefore, for any $z \in B(\hat{z},\gamma_1) \cap \mathbb{C}$, from 3-22, $\delta_h(z)$

$$\leq \max_{\omega \geqq 0}\{\inf_{\mu \in [0,1]} \{\sum_{j=1}^{m} \omega^j f^j(z+\mu h(z)) + \omega^0[f^0(z+\mu h(z) - f^0(z)]\} \mid \sum_{j=0}^{m} \omega^j = 1\}$$

and from 3-23

3-32 $$\delta_h(z) \leq \max_{\omega \geqq 0}\{\inf_{\mu \in [0,1]} \{\sum_{j=1}^{m} \omega^j f^j(z) + \mu \sum_{j=0}^{m} \omega^j \langle \nabla f^j(z), h(z) \rangle + \frac{\mu^2}{2} L\|h(z)\|^2\} \mid \sum_{j=0}^{m} \omega^j = 1\}$$

By setting $\mu = \frac{1}{L}$ in 3-32 (and deleting the inf operation) we obtain

3-33 $$\delta_h(z) \le \max_{\omega \geqq 0}\{\sum_{j=1}^{m}\omega^j f^j(z) + \frac{1}{L}\sum_{j=0}^{m}\omega^j \langle \nabla f^j(z),h(z) \rangle + \frac{L}{2}\|h(z)\|^2 | \sum_{j=0}^{m}\omega^j = 1\}.$$

Since $z \in \mathbb{C}$, $f^j(z) \le 0$, $j = 1,\cdots, m$; and since $L \ge 1$, it follows that $f^j(z) \le \frac{1}{L} f^j(z)$, $j = 1,\cdots, m$. Therefore

3-34 $$\delta_h(z) \le \frac{1}{L}\max_{\omega \geqq 0}\{\sum_{j=1}^{m}\omega^j f^j(z) + \sum_{j=0}^{m}\omega^j \langle \nabla f^j(z),h(z) \rangle + \frac{1}{2}\|h(z)\|^2 | \sum_{j=0}^{m}\omega^j = 1\}.$$

By definition, $(h^0(z), h(z))$ is a solution of 3-10 with $z_i = z$. Therefore

$\langle \nabla f^0(z),h(z) \rangle \le h^0(z)$; $f^j(z) + \langle \nabla f^j(z),h(z) \rangle \le h^0(z)$, $j = 1,\ldots, m$.

Hence, from 3-34, $\delta_h(z) \le \frac{1}{L}\max_{\omega \geqq 0}\{\sum_{j=0}^{m}\omega^j h^0(z) + \frac{1}{2}\|h(z)\|^2 | \sum_{j=0}^{m}\omega^j = 1\}$, i.e.,

$\delta_h(z) \le \frac{1}{L}[h^0(z) + \frac{1}{2}\|h(z)\|^2]$, which proves the lemma. $\square$

3-35 <u>Lemma</u>

Suppose that assumptions 3-7 (i)-(iv) are satisfied. Then there exists a $\gamma_2 > 0$ such that $h^0(z) + \frac{1}{2}\|h(z)\|^2 \le \ell\delta(z)$ for all $z \in B(\hat{z},\gamma_2) \cap \mathbb{C}$, where $\hat{z}$ is the solution of problem 3-5, $(h^0(z),h(z))$, $\delta(z)$, ℓ, $\mathbb{C}$ are respectively defined by 3-10, 3-1, 3-14, 3-6.

Proof: From the continuity of $\delta(\cdot)$: $\mathbb{C} \to R$, defined by $\delta(z) = \min_{(\delta,z')}\{\delta | f^0(z') - f^0(z) \leq \delta;\ f^j(z') \leq \delta\ j = 1,\cdots, m\}$, and the fact that $(\hat{z}) = 0$, there exists $\gamma_2 \in (0,\frac{\varepsilon}{2}]$ such that[7]

3-36 $$\delta(z) = \min_{(\delta,z')}\{\delta|f^0(z') - f^0(z) \leq \delta;\ f^j(z') \leq \delta\ j = 1,\cdots m;\ z' \in B(\hat{z},\tfrac{\varepsilon}{2})\}$$

for all $z \in B(\hat{z},\gamma_2) \cap \mathbb{C}$. Upon applying the strong duality theorem 1-5 to 3-36, we obtain

3-37 $$\delta(z) = \max_{u \geqq 0}\ \{\inf_{\substack{z'\in B(\hat{z},\frac{\varepsilon}{2})\\ \delta\in R}} \{\sum_{j=1}^{m} v^j f^j(z') + v^0[f^0(z') - f^0(z)] + \{1 - \sum_{j=0}^{m} v^j]\delta\}\}$$

3-38 $$= \max_{v \geqq 0}\ \{\inf_{z'\in B(\hat{z},\frac{\varepsilon}{2})} \{\sum_{j=1}^{m} v^j f^j(z') + v^0[f^0(z')]\}|\sum_{j=0}^{m} v^j = 1\}.$$

Expanding the function inside 3-38, according to the second order Taylor expansion formula, and making use of hypothesis 3-5 (iv) we conclude that

3-39 $$\sum_{j=1}^{m} v^j f^j(z') + v^0[f^0(z') - f^0(z)] \geq \sum_{j=1}^{m} v^j f^j(z) + \sum_{j=0}^{m} v^j \langle \nabla f^j(z), z'-z \rangle + \frac{1}{2}\ell \|z'-z\|^2$$

for all $z \in B(\hat{z},\gamma_2) \cap \mathbb{C}$, for all $z' \in B(\hat{z}, \frac{\varepsilon}{2})$, and for all $v \geqq 0$ such that $\sum_{j=0}^{m} v^j = 1$. As a function of z', the right hand side of 3-39 is minimized by $z' = z - \frac{1}{\ell}\sum_{j=0}^{m} v^j \nabla f^j(z)$. Hence, 3-40 $$\sum_{j=1}^{m} v^j f^j(z') + v^0[f^0(z') - f^0(z)] \geq \sum_{j=1}^{m} v^j f^j(z) - \frac{1}{2\ell} \| \sum_{j=0}^{m} v^j \nabla f^j(z) \|^2$$

Making use of the fact that $z \in \mathbb{C}$ and that $\ell \leq 1$, we conclude from 3-40 that

3-41 $$\sum_{j=1}^{m} v^j f^j(z') + v^0[f^0(z') - f^0(z)] \geq \frac{1}{\ell}\left[\sum_{j=1}^{m} v^j f^j(z) - \frac{1}{2} \| \sum_{j=0}^{m} v^j \nabla f^j(z) \|^2\right],$$

for all $z \in B(\hat{z},\gamma_2) \cap \mathbb{C}$, for all $z' \in B(\hat{z},\frac{\varepsilon}{2})$, and for all $v \geqq 0$ such that $\sum_{j=0}^{m} v^j = 1$. Therefore, 3-38 and 3-41 imply that

3-42 $$\delta(z) \geq \frac{1}{\ell} \max_{v \geqq 0}\{\sum_{j=1}^{m} v^j f^j(z) - \frac{1}{2} \| \sum_{j=1}^{m} v^j \nabla f^j(z) \|^2 | \sum_{j=0}^{m} v^j = 1\}$$

which, according to 3-27, implies that $\delta(z_i) \geq \frac{1}{\ell} [h^0(z) + \frac{1}{2} \|h(z)\|^2]$ for all $z \in B(\hat{z},\gamma_2) \cap \mathbb{C}$. This completes our proof. $\square$

3-43 Lemma.

Let $\{z_i\}_{i=0}^{\infty}$ be a sequence generated by algorithm 3-9 in solving problem 3-5. Suppose that assumptions 3-7 (i)-(iii) are satisfied and suppose that the function f^0 is strictly convex in a convex neighborhood of the solution $\hat{z}$ of 3-5. Furthermore, suppose that the functions f^j, $j = 0, 1, \cdots, m$, are such that there exist $\gamma > 0$, $\kappa > 0$ such that

3-44 $$\delta_h(z) \leq \kappa\delta(z) \text{ for all } z \in B(\hat{z},\gamma) \cap \mathbb{C}.$$

Then, given any $\alpha \in (0,1)$, there exists an integer $i_0(\alpha)$ such that

3-45 $\quad f^0(z_{i+1}) - f^0(\hat{z}) \leq [1 - \kappa\bar{\lambda}^0(1-\alpha)][f^0(z_i) - f^0(\hat{z})]$ for all $i \geq i_0(\alpha)$

Proof: According to 2-28, given any $\alpha \in (0,1)$, there exists an integer $i_3(\alpha)$ such that $\delta(z_i) \leq \lambda^0(1-\alpha)[f^0(\hat{z})-f^0(z_i)]$ for all $i \geq i_0(\alpha)$. It therefore follows from 3-44 that

3-46 $\quad \delta_h(z_i) \leq \kappa(1-\alpha)\bar{\lambda}^0[f^0(\hat{z}) - f^0(z_i)]$ for all $i \geq \min\{i_0(\alpha), i_3(\alpha)\}$

Hence from 3-18,

3-47 $\quad f^0(z_{i+1}) - f^0(z_i) \leq \delta_h(z_i) \leq \kappa(1-\alpha)\bar{\lambda}^0[f^0(\hat{z}) - f^0(z_i)]$

which, rearranged, becomes

3-48 $\quad f^0(z_{i+1}) - f^0(\hat{z}) \leq [1 - \kappa\,\bar{\lambda}^0(1-\alpha)][f^0(z_i) - f^0(\hat{z}_i)]$,

and completes our proof. □

Proof of Theorem 3-15:

Inequality 3-16' follows directly from the lemmas 3-19, 3-35 and 3-43. Inequality 3-16 follows from 3-16' and lemma 1-20. This completes the proof of theorem 3-15. □

In proving theorem 3-15, assumption 3-7 (iv) was used only to establish lemma 3-35. We shall now prove that a theorem similar to 3-15 can also be derived when assumption 3-7 (iv) is weakened as follows.

3-49 Assumption: There exist $\varepsilon' > 0$ and $m^0 \in (0,1)$ such that $\langle y, \frac{\partial^2 f^0}{\partial z^2}(z)\, y \rangle \geq m^0\|y\|^2$ for all $y \in R^n$, and all $z \in B(\hat{z},\varepsilon) \cap \mathbb{C}$, where $\hat{z}$ is the unique solution of problem 3-5 and $\mathbb{C}$ is defined by 3-6. □

3-50 Theorem. (Linear convergence in the case of non strictly convex constraints).

Let $\{z_i\}_{i=0}^{\infty}$ be a sequence generated by algorithm 3-9 in solving problem 3-5. Suppose that assumptions 3-7 (i)-(iii) and 3-49 are satisfied. Then $z_i \to \hat{z}$ linearly as $i \to \infty$, $f^0(z_i) \to f^0(\hat{z})$ linearly as $i \to \infty$, in accordance with the following bounds: given any $\alpha \in (0,1)$ there exist $i_0(\alpha)$, $\kappa_0(\alpha) > 0$, such that

3-51 $$\|z_i - \hat{z}\| \leq [1 - \bar{\lambda}^{0^2} \frac{m^0}{L} (1-\alpha)^2]^{1/2} \kappa_0(\alpha) \quad \text{for all } i \geq i_0(\alpha),$$

3-52 $$f^0(z_{i+1}) - f^0(\hat{z}) \leq [1 - \bar{\lambda}^{0^2} \frac{m^0}{L} (1-\alpha)^2][f^0(z_i) - f^0(\hat{z})] \quad \text{for all } i \geq i_0(\alpha),$$

where m^0, L, $\bar{\lambda}^0$ are respectively defined by 3-49, 3-13, 3-17.

Proof: Theorem 3-40 follows directly from lemmas 3-19, 3-43, 1-20 and lemma 3-53 below.

□

3-53 Lemma.

Suppose that assumptions 3-7 (i)-(iii) and 3-49 are satisfied. Then, given any $\alpha \in (0,1)$, there exists a $\gamma_3(\alpha) > 0$ such that $h^0(z) + \frac{1}{2}\|h(z)\|^2 \leq m^0 \bar{\lambda}^0 (1-\alpha)\delta(z)$ for all z in $B(\hat{z},\gamma_3(\alpha)) \cap \mathbb{C}$, where $\hat{z}$, $(h^0(z),h(z))$, m^0, $\bar{\lambda}^0$, $\delta(z)$, $\mathbb{C}$, are respectively defined by 3-5, 3-10, 3-49, 3-17, 3-1, 3-6.

Proof: According to lemma 3-35 (see 3-38), for any $z \in B(\hat{z},\gamma_2) \cap \mathbb{C}$

3-54 $$\delta(z) = \max_{v \geq 0} \{\inf_{z' \in B(\hat{z}, \frac{\varepsilon}{2})} \{\sum_{j=1}^{m} v^j f^j(z') + v^0[f^0(z') - f^0(z)]\} \mid \sum_{j=0}^{m} v^j = 1\}.$$

Therefore

$$\text{3-55} \qquad \delta(z) \geq \inf_{z' \in B(\hat{z}, \frac{\varepsilon}{2})} \{\sum_{j=1}^{m} v^j f^j(z') + v^0[f^0(z') - f^0(z)]\}$$

for all $v \geqq 0$ such that $\sum_{j=0}^{m} v^j = 1$ and all $z \in B(\hat{z},\gamma_2) \cap \mathbb{C}$.

Applying the Taylor second order expansion formula and using hypotheses 3-49 and 3-7(iii), we obtain

$$\text{3-56} \qquad \delta(z) \geq \inf_{z' \in B(\hat{z}, \frac{\varepsilon}{2})} \{\sum_{j=1}^{m} v^j f^j(z) + \sum_{j=0}^{m} v^j \langle \nabla f^j(z), z'-z \rangle + \frac{m^0}{2} v^0 \| z'-z \|^2\}$$

for all $v \geqq 0$ such that $\sum_{j=0}^{m} v^j = 1$, and all $z \in B(\hat{z},\gamma_2) \cap \mathbb{C}$.

By deleting the constraint $z' \in B(\hat{z},\frac{\varepsilon}{2})$ in 3-56, we obtain

$$\text{3-57} \qquad \delta(z) \geq \sum_{j=1}^{m} v^j f^j(z) - \frac{1}{2m^0 v^0} \| \sum_{j=0}^{m} v^j \nabla f^j(z) \|^2$$

for all $v \geqq 0$ such that $\sum_{j=0}^{m} v^j = 1$ and all $z \in \hat{z},\gamma_2) \cap \mathbb{C}$. Let $\bar{v}(z) \in R^{m+1}$

be a solution of problem 3-27. Then, since $m^0\bar{v}^0(z) \leq 1$ and $\sum_{j=1}^{m} \bar{v}^j(z) f^j(z) \leq 0$,

we must have, $\sum_{j=1}^{m} \bar{v}^j(z) f^j(z) \geq \frac{1}{m^0\bar{v}^0(z)} \sum_{j=1}^{m} \bar{v}^j(z) f^j(z)$.

Hence, from 3-57, and 3-27,

3-58 $$\delta(z) \geq \frac{1}{m^0 \bar{v}^0(z)} \left[\sum_{j=1}^{m} \bar{v}^j(z) f^j(z) - \frac{1}{2} \| \sum_{j=0}^{m} \bar{v}^j(z) \nabla f^j(z) \|^2 \right] = \frac{1}{m^0 \bar{v}^0(z)} \left[h^0(z) + \frac{1}{2} \| h(z) \|^2 \right]$$

By following the same pattern of reasoning as in the theoretical method of centers, it is easy to prove that

3-59 $$\liminf_{\substack{z \to \hat{z} \\ z \in \mathbb{C}}} \bar{v}^0(z) \geq \bar{\lambda}^0$$

where $\bar{\lambda}^0$ is defined by 3-17 and, according to lemma 1-12 satisfies $\bar{\lambda}^0 > 0$. It now follows from 3-59 that, given any $\alpha \in (0,1)$, there exists $\gamma_3(\alpha) \in (0,\gamma_2)$ such that $\bar{v}^0(z) \geq \bar{\lambda}^0(1-\alpha)$ for all $z \in B(\hat{z},\gamma_3(\alpha)) \cap \mathbb{C}$. Therefore 3-58 implies that $\delta(z) \geq \frac{1}{m^0\bar{\lambda}^0(1-\alpha)} [h^0(z) + \frac{1}{2} \| h(z) \|^2]$, which completes the proof of the lemma.

□

Algorithm 3-9 is not implementable on a computer because of the exact minimization required in 3-11. Let Step 2 in algorithm 3-9 be replaced by Step 2'

3-60 <u>Step 2'</u> If $i = 0$ select a $\beta > 0$, else

If $h^0(z_i) = 0$ set $\hat{z} = z_i$, and stop; else apply the Golden Section Search (or any similar scheme) to the function θ: $R \to R$ defined by

3-61 $$\theta(\mu) = \max\{f^0(z_i + \mu h(z_i)) - f^0(z_i);\ f^j(z_i + \mu h(z_i))\ j = 1,\cdots, m\}$$

to find two points $\mu(z_i)$, $\mu'(z_i)$ with $\mu'(z_i) > \mu(z_i) > 0$, such that $[\mu(z_i), \mu'(z_i)]$ contains the minimizer $\bar{\mu}(z_i)$ of θ and such that

3-62 $$\theta(\mu(z_i)) \leq \beta[\mu'(z_i) - \mu(z_i)]\langle \nabla f^0(z_i), h(z_i)\rangle$$ □

3-63 <u>Theorem</u> (linear convergence) .

Let $\{z_i\}_{i=0}^{\infty}$ be a sequence generated by algorithm 3-9, modified to use Step 2' above, in solving 3-5. Suppose that assumptions 3-7 (i)-(iii) and 3-49 are satisfied. Then, either $\{z_i\}$ is finite its last element is $\hat{z}$, the unique solution of 3-5, or $\{z_i\}$ is infinite and $z_i \to \hat{z}$ linearly as $i \to \infty$; $f^0(z_i) \to f^0(\hat{z})$ linearly as $i \to \infty$, in accordance with the following bounds. Given any $\alpha \in (0,1)$, there exist $i_0(\alpha)$, $\kappa_0(\alpha)$ such that,

3-64 $$\| z_i - \hat{z} \| \leq [1 - \frac{m^0}{L}[\bar{\lambda}^0(1-\alpha)]^2 \frac{\beta}{1+\beta}]^{i/2}\kappa_0(\alpha) \quad \text{for all } i \geq i_0(\alpha)$$

3-65 $$f^0(z_{i+1}) - f^0(z_i) \leq [1 - \frac{m^0}{L}[\bar{\lambda}^0(1-\alpha)]^2 \frac{\beta}{1+\beta}][f^0(z_i) - f^0(\hat{z})] \quad \text{for all } i \geq i_0(\alpha)$$

where m^0, $\bar{\lambda}^0$, L are respectively defined by 3-49, 3-17 and 3-13.

<u>Proof</u>. The function θ: $R \to R$ defined by 3-61 is convex. Therefore

3-66 $$\theta(\mu) \geq \mu\langle \nabla f^0(z_i), h(z_i)\rangle$$

because $f^j(z_i) < 0, j = 1,\cdots, m$ and $i = 1, 2, 3,\cdots$, and hence $\frac{d\theta(0+)}{d\mu} = \langle \nabla f^0(z_i), h(z_i)\rangle$ for all $i \geq 1$. It follows from 3-62 and 3-66 that $\mu(z_i) \geq \tilde{\mu}$, where $\tilde{\mu}$ satisfies $\tilde{\mu}\langle \nabla f^0(z_i), h(z_i)\rangle = \beta(\tilde{\mu}(z_i) - \tilde{\mu})\langle \nabla f^0(z_i), h(z_i)\rangle$

(see Fig. 1). If the point z_i is not a solution of 3-5, then from lemma 1-31 and 3-10, $\langle \nabla f^0(z_i), h(z_i) \rangle < 0$. Thus

$$\text{3-67} \qquad \mu(z_i) \geq \tilde{\mu} = \frac{\beta}{1+\beta} \bar{\mu}(z_i) \cdot$$

Next, it follows from the convexity of $\theta(\cdot)$ and the fact that $\bar{\mu}(z_i)$ is a minimizer of $\theta(\cdot)$ that

$$\text{3-68} \qquad \theta(\mu(z_i)) \leq \frac{\mu(z_i)}{\bar{\mu}(z_i)} \theta(\bar{\mu}(z_i)),$$

which, together with 3-67 implies that

$$\text{2-69} \qquad \theta(\mu(z_i)) \leq \frac{\beta}{1+\beta} \theta(\bar{\mu}(z_i)).$$ [8]

Therefore, from 3-61 and 3-11

$$\text{3-70} \qquad f^0(z_{i+1}) - f^0(z_i) \leq \frac{\beta}{1+\beta} \delta_h(z_i)$$

where $\delta_h(z_i)$ is defined by 3-11.

Thus the convergence of algorithm 3-9, modified to use Step 2', follows from 3-70, theorem 1.3.10 in Polak [8] and the proof of convergence of algorithm 3-9. Inequalities 3-66 and 3-67 follow from lemmas 3-19, 3-43, 3-53, 1-20 and inequality 3-70. □

The following theorem highlights an important facet of algorithm 3-9 when modified to use Step 2' instead of Step 2.

3-71 <u>Theorem</u>.

If a Golden Section Search is used, in Step 2' of algorithm 3-9, to compute $\mu(z_i)$, $\mu'(z_i)$ which satisfy 3-62, then

3-72 $$\lim_{i\to\infty}[\mu'(z_i) - \mu(z_i)] > 0.$$

<u>Proof</u>:

Since, $\theta(\cdot)$ is convex, $\theta(\mu) \leq \delta_h(z_i) - (\bar{\mu}(z_i) - \mu)\langle \nabla f^0(z_i),h(z_i)\rangle$ for all $\mu \leq \bar{\mu}(z_i)$. Therefore, from lemma 3-19, there exists an $i_1 > 0$ such that

3-73 $$\theta(\mu) \leq \frac{1}{L}\,[h^0(z_i) + \frac{1}{2}\,\|h(z_i)] - (\bar{\mu}(z_i)-\mu)\langle \nabla f^0(z_i),h(z_i)\rangle \quad \text{for all } i \geq i_1 \text{ and all } \mu \in [0,\bar{\mu}(z_i)],$$

where $(h^0(z_i),h(z_i))$ is a solution of 3-10. It follows from 3-30 and 3-73 that

3-74 $$\theta(\mu(z_i)) \leq \frac{h^0(z_i)}{2L} - (\bar{\mu}(z_i) - \mu)\langle \nabla f^0(z_i),h(z_i)\rangle \quad \text{for all } i \geq i_1 \text{ and all } \mu \in [0,\bar{\mu}(z_i)].$$

According to part (b)(ii) of the strong duality theorem 1-5, as applied to 3-24,

3-75 $$v^0(z_i)[\langle \nabla f^0(z_i),h(z_i)\rangle - h^0(z_i)] = 0$$

where $v^0(z_i)$ is the first component of a solution of 3-24. From lemma 3-53 (see 3-59) there exists $i_2 \geq i_1$ such that $v^0(z_i) > 0$ for all $i \geq i_2$. Thus from 3-75.

3-76 $$h^0(z_i) = \langle \nabla f^0(z_i),h(z_i)\rangle \quad \text{for all } i \geq i_2$$

and therefore 3-72 becomes,

$$\text{3-77} \qquad \theta(\mu) \le [\frac{1}{2L} - (\bar{\mu}(z_i)-\mu)] \langle \nabla f^0(z_i), h(z_i) \rangle, \text{ for all } \mu \in [0, \bar{\mu}(z_i)] \text{ and all } i \ge i_2$$

In Step 2', starting from an interval $[\mu^0, \mu^{0)}]$ containing $\bar{\mu}(z_i)$, the Golden Section search generates a sequence of interval $\{[\mu^k, \mu'^k]\}_{k=0}^{p}$ such that $\bar{\mu}(z_i) \in [\mu^k, \mu'^k]$ and $\dfrac{\mu'^{k+1} - \mu^{k+1}}{\mu'^k - \mu^k} = \Delta$ ($\Delta \simeq 0.68$) and $\mu^p \Rightarrow \mu(z_i)$; $\mu'^p = \mu'(z_i)$. Since the search for $\mu(z_i)$ did not stop at μ^{p-1}, μ^{p-1} must have failed the test 3-62 i.e.

$$\text{3-78} \qquad \theta(\mu^{p-1}) > \beta[\mu'^{p-1} - \mu^{p-1}] \langle \nabla f^0(z_i), h(z_i) \rangle$$

and therefore

$$\text{3-79} \qquad \theta(\mu^{p-1}) > \frac{\beta}{\Delta}[\mu'(z_i) - \mu(z_i)] \langle \nabla f^0(z_i), h(z_i) \rangle \;.$$

Hence, from 3-75,

$$\text{3-80} \qquad \frac{\beta}{\Delta}[\mu'(z_i) - \mu(z_i)] \ge \frac{1}{2L} - [\bar{\mu}(z_i) - \mu^{p-1}]$$

Making use of the fact that $\bar{\mu}(z_i) - \mu^{p-1} \le \frac{1}{\Delta}[\mu'(z_i) - \mu(z_i)]$

we obtain

$$\text{3-81} \qquad [\mu'(z_i) - \mu(z_i)] \ge \frac{1}{2L} \cdot \frac{\Delta}{1+\beta},$$

which completes our proof. □

3-82 Theorem

Let L, m^0, $\bar{\lambda}^0$ and $\{z_i\}$ be defined as in theorem 3-63 and suppose that the assumptions 3-7 (i) - (iii) and 3-49 are satisfied. Then there exists and integer i_1 such that

3-83 $$\mu^*(z_i) \leq \frac{1}{m^0} + 4(L-\bar{\lambda}^0 m^0)\left(\frac{L}{m^0\bar{\lambda}^0}\right)^2 \qquad \text{for all } i \geq i_1$$

where $\mu^*(z_i)$ is defined as the strictly positive root of $\theta(\mu)$, ($\theta(\cdot)$ as in 3-61).

Proof. From 3-61,

3-84 $$\theta(\mu) \geq f^0(z_i + \mu\, h(z_i)) - f^0(z_i) \qquad \text{for all } \mu \geq 0$$

Applying the second order Taylor expansion formula to the right hand side of 3-84, we obtain

3-85 $$\theta(\mu) \geq \mu\ \langle \nabla f^0(z_i) \rangle + \langle \nabla f^0(z_i), h(z_i) \rangle + \frac{\mu^2}{2} m^0 \|h(z_i)\|^2, \quad \text{for all } \mu \geq 0,$$

which, because of 3-76, becomes

3-86 $$\theta(\mu) \geq \mu[h^0(z_i) + \frac{m^0\mu}{2} \|h(z_i)\|^2 \quad \text{for all } i \geq i_2, \text{ for all } \mu \geq 0.$$

The right hand side of 3-86 is strictly positive for all $\mu > \frac{2}{m^0} \frac{h^0(z_i)}{\|h(z_i)\|^2}$.

Consequently, $\mu^*(z_i)$, the strictly positive root of $\theta(\cdot)$, is such that

3-87 $$\mu^*(z_i) \leq -\frac{2}{m^0}\frac{h^0(z_i)}{\|h(z_i)\|^2} \qquad \text{for all } i \geq i_2.$$

Let $(\lambda^0, \lambda^1, \ldots \lambda^m) \in \Lambda(\hat{z})$ (see 2-7), then it follows from 3-10 that

$$\text{3-88} \qquad \sum_{j=1}^{m} \lambda^j f^j(z_i) + \langle \sum_{j=0}^{m} \lambda^j \nabla f^j(z_i), h(z_i) \rangle \leq h^0(z_i).$$

Expanding the left hand side of 3-88 about $\hat{z}$ and making use of 2-7 and 3-49, we conclude that there exists an integer i_1 such that

$$\text{3-89} \qquad \frac{1}{2} m^0 \lambda^0 \|\hat{z}-z_i\|^2 - L \|\hat{z}-z_i\| \|h(z_i)\| \leq h^0(z_i) \leq 0 \quad \text{for all } i \geq i_1.$$

Therefore

$$\text{3-90} \qquad \|h(z_i\| \geq \frac{m^0\lambda^0}{2L} \|\hat{z}-z_i\|, \quad \text{for all } i \geq i_1.$$

Next, from 3-27,

$$\text{3-91} \qquad h^0(z_i) + \frac{1}{2} \|h(z_i)\|^2 \geq \sum_{j=1}^{m} \lambda^j f^j(z_i) - \frac{1}{2} \| \sum_{j=0}^{m} \lambda^j \nabla f^j(z_i)\|^2,$$

which, with the aid of the Taylor expansion formula, yields

$$\text{3-92} \qquad h^0(z_i) + \frac{1}{2} \|h(z_i)\|^2 \geq \frac{1}{2}(\lambda^0 m^0 - L)\|\hat{z}-z_i\|^2.$$

Combining 3-90 and 3-92, we obtain

$$\text{3-93} \qquad \frac{h^0(z_i)}{\|h(z_i)\|^2} \geq -\frac{1}{2} - \frac{1}{2}(L-\lambda^0 m^0) \left(\frac{2L}{m^0\lambda^0}\right)^2 .$$

Inequality 3-83 now follows from 3-93 and 3-87. $\square$

3-94 Corollary.

Let $\{z_i\}$ be as in theorem 3-63 and suppose that for $i = 0, 1, 2, \ldots$ k_i is the smallest integer such that $\theta(z_i + k_i\, h(z_i)) > 0$. Then

$$k_i \le \frac{1}{m^0} + 4(L-\bar{\lambda}^0 m^0)\, \frac{L}{m^0\bar{\lambda}^0}^{\,2} + 1, \text{ for all } i \ge i_1, \text{ where } i_1 \text{ is as m 3-82.}$$

Furthermore if $\mu(z_i)$, $\mu'(z_i)$ in 3-60 are computed by means of the Golden Section search using the initial interval $[0,k_i]$ then there exists an integer $\hat{k}$ such that $\mu(z_i)$, $\mu'(z_i)$ are computed with no more than $\hat{k}$ evaluation of $\theta(\mu)$, for $i = 1, 2, \ldots$. $\square$

Corollary 3-94 together with theorem 3-63 show that algorithm 3-9 modified to use Step 2', implemented as in Corollary 3-94 converges linearly.

CONCLUSION:

We have seen in this paper that duality theory can be used to construct bounds needed for determining the rate of convergence of a class of optimization algorithms. We have also demonstrated how the use of duality theory can influence the construction of new algorithms. The approach used in this paper has a certain amount of generality, since it can also be used to deal with a number of methods of feasible directions and optimal control algorithms. We shall present our work on these other algorithms under separate cover.

APPENDIX A: Counter example to the linear convergence of the modified method of centers in [5] and of the Topkis-Veinnott method of feasible directions [11].

1) Topkis-Veinnott method of feasible directions[11]

Consider the problem

A-1 $$\min \{x^2 + y^2 | y \geq 0\}$$

and let $z_0 = (x_0, y_0)$ be such that $0 < |x_0| \leq y_0 \leq \frac{1}{2}$. To solve a problem of the type $\min \{f^0(z) | f^j(z) \leq 0 \ j = 1, \ldots m)$ the Topkis-Veinnott method of feasible directions, computes a feasible direction $h(z_i)$ at z_i as a solution of $\min_{|h^k|<1} \{\max\{<\nabla f^0(z_i), h>; f^j(z_i) + \langle \nabla f^j(z_i), h \rangle \ j = 1, \ldots, m\}\}$ which, for $f^0(z) = x^2 + y^2$, $f^1(z) = -y$, $m = 1$, becomes, at z_0,

A-2 $$\min_{\substack{-1 \leq h^1 \leq 1 \\ -1 \leq h^2 \leq 1}} \{\max\{2x_0 h^1 + 2y_0 h^2, \ - h^2 - y_0\}\} .$$

Hence

A-3 $$h^1(z_i) = -\mathrm{sgn}\ x_0 \ ; \ h^2(z_i) = - \frac{y_0 - 2|x_0|}{1 + 2y_0}$$

The step size $\mu(z_i)$ is computed as the solution of $\min\{f^0(z_i + \mu\ h(z_i)) | f^j(z_i + \mu\ h(z_i)) \leq 0 \ j = 1, \ldots, m\}$. Therefore $\mu(z_0)$ is found as the solution of

A-4 $$\min_{\mu}\{(x_0 - \mu\ \mathrm{sgn}\ x_0)^2 + [y_0 - \mu \frac{y_0 - 2|x_0|}{1 + 2y_0}]^2 \ | \ y_0 - \mu \frac{y_0 - 2|x_0|}{1 + 2y_0} \geq 0\}$$

From A-3 and A-4, (see Fig. 2), it is clear that the constraints in A-4 is not active at $\mu(z_0)$. Therefore $\mu(z_0)$ satisfies

A-5 $$(|x_0| - \mu(z_0)) + \frac{y_0 - 2|x_0|}{1 + 2y_0} (y_0 - \mu(z_0) \frac{y_0 - 2|x_0|}{1 + 2y_0}) = 0.$$

Let $z_1 = (x_1, y_1)$ be the next point computed by the algorithm. Then from A-3

A-6 $$x_1 = x_0 - \mu(z_0) \text{ sgn } x_0 \quad ; y_1 = y_0 - \mu(z_0) \frac{y_0 - 2|x_0|}{1 + 2y_0}$$

Hence, from A-5 and the assumption on x_0, y_0,

A-7 $$\left|\frac{x_1}{y_1}\right| = \frac{(y_0 - 2|x_0|)}{1 + 2y_0} = y_0 \frac{\left|1 - 2\frac{|x_0|}{y_0}\right|}{1 + 2y_0}$$

which implies that

A-8 $$\left|\frac{x_1}{y_1}\right| \leq y_0 .$$

Let $\{z_i\}$ be the sequence generated by the Topkis-Veinnott method of feasible direction, in solving problem A-1, and starting from $z_0 = (x_0, y_0)$ such that $0 < |x_0| \leq y_0 \leq \frac{1}{2}$. At each iteration the algorithm decreases the cost therefore if y_0 is chosen such that $y_0 \leq \frac{1}{4}$, by making repeated use of A-8, we obtain $0 < |x_i| \leq y_i \leq \frac{1}{2}$ for all $i \geq 1$. From A-3 it follows that $h^2(z_i) \leq 0$ for all $i \geq 1$. Hence (see Fig. 2)

A-9 $$\frac{\|z_{i+1} - z\|}{\|z_i - \hat{z}\|} = \cos \gamma_i$$

where

A-10 $$\gamma_i = \pi - \beta_{i+1} - \beta_i$$

From A-8, $\tan \beta_{i+1} = \frac{y_{i+1}}{x_{i+1}} \geq \frac{1}{y_i}$. Hence, $\beta_i \to \frac{\pi}{2}$ as $i \to \infty$, and from A-10, $\gamma_i \to 0$ as $i \to \infty$. Therefore, from A-9, $\lim_{i\to\infty} \frac{\|z_{i+1} - \hat{z}\|}{\|z_i - \hat{z}\|} = 1$, which proves that $\{z_i\}$ cannot converge linearly.

2) Modified method of centers

Huard's modified method of centers [5] has the same direction determination subproblem as the Topkis-Veinnott algorithms just described, but the step size $\mu(z_i)$ is given by $\min_{\mu}\{\max\{f^0(z_i + \mu h(z_i)) - f^0(z_i); f^j(z_i + \mu h(z_i))$ $, j = 1, \ldots, m\}\}$ which for problem A-1 becomes, at z_0,

A-11 $$\min_{\mu}\{\max\{(x_0 - \mu \operatorname{sgn} x_0)^2 + \left(y_0 - \mu \frac{y_0 - 2|x_0|}{1 + 2y_0}\right)^2 ; y_0 - \mu \frac{y_0 - 2|x_0|}{1 + 2y_0}\}\}.$$

Consequently, $\mu(z_0)$ satisfies either A-5 or

A-12 $$(|x_0| - \mu)^2 + \left(y_0 - \mu \frac{y_0 - 2|x_0|}{1 + 2y_0}\right)^2 = y_0 - \mu \frac{y_0 - 2|x_0|}{1 + 2y_0}$$

Therefore, the point $z_1 = (x_1,y_1)$ computed by the Huard modified method of centers, after z_0 is given by A-6 with $\mu(z_1) > 0$ being the smallest strictly positive number which satisfies either A-5 or A-12. Note that if $\mu(z_1)$ is given by A-12 then $(x_1 y_1)$ satisfies $x_1^2 + y_1^2 = y_1$. Hence it is clear (see Fig. 2) that if (x_0, y_0) is close enough to $(0,0)$, $\bar{\lambda}$ will never be given by A-12, which implies that, for problem A-1, and starting from z_0 such that $0 < |x_0| \leq Y_0 \leq \varepsilon$, the Topkis-Veinnott and the Huard algorithms compute the same sequence of points and hence neither of those two algorithms can converge linearly.

<u>Footnotes</u>

1. This condition is stronger than the one used by Geoffrion [3].
2. This algorithm is not implementable because of the exact minimization required in 2-6.
3. Note that the map Γ^0: $R^n \to \mathcal{P}(R)$ defined by $\Gamma^0(z) = \{ \langle u,e \rangle \mid u \in \Gamma(z)\}$ is closed (see Berge [1]).
4. Actually, the convexity of f^j, $j = 0, 1, \ldots, m$, is needed only in a compact convex subset of R^n containing $\mathbb{C}(z_0)$ in its interior. It is only for the sake of simplicity that we assumed global convexity.
5. As we shall see later, it is sufficient for 3-8 to hold for $j = 0$ only, to insure linear convergence, but in that case the bound on the rate is larger.
6. Note that in some cases it might be easier to solve 3-27 and use 3-29, than to solve 3-10 in algorithm 3-9.
7. The mapping Γ: $R^n \to \mathcal{P}(R^n)$ define by $\Gamma(z) = \{z' | f^0(z') - f^0(z) \leq \delta(z)\ ;\ f^j(z') \leq \delta(z)\}$ is upper semi continuous (see Berge [1]).
8. The process of replacing the computation of a minimizer $\bar{\mu}$ of a convex function θ: $[0,+\infty) \to (-\infty,0]$ by the computation of two points μ,μ' such that $0 < \mu \leq \bar{\mu} \leq \mu'$ and such that $\theta(\mu) \leq \beta(\mu'-\mu)\,\frac{d\theta}{d\mu}\,(0^+)$ constitutes, because of 3-69, a general procedure for implementing algorithms of the type of 3-9.

References

[1] Berge, C., Topological Spaces, Mac Millan, New York, 1963.

[2] Frank, M., and P. Wolfe, "An Algorithm for Quadratic Programming," Naval Log. Quart., Vol. 3 (1956), pp. 95-110.

[3] Geoffrion, A., "Duality in Nonlinear Programming," Working Paper No. 150, University of California, Los Angeles, Western Management Science Institute, August 1969.

[4] Huard, P., "The Method of Centers" in Nonlinear Programming A Course, North-Holland Publication Company, Amsterdam, 1965.

[5] Huard, P., "Programmation Mathematique Convex," RIRO, R7, (1968), pp. 43-59.

[6] Kuhn, H. W., and A. W. Tucker, "Nonlinear Programming," Proc. Second Berkeley Symp. Math. Stat. Probability, University of California Press, Berkeley, Calif. 1951, pp. 481-492.

[7] Mangasarian, O., Nonlinear Programming, McGraw-Hill Co., New York, 1969.

[8] Polak, E., Computational Methods in Optimization, Academic Press, New York, 1971.

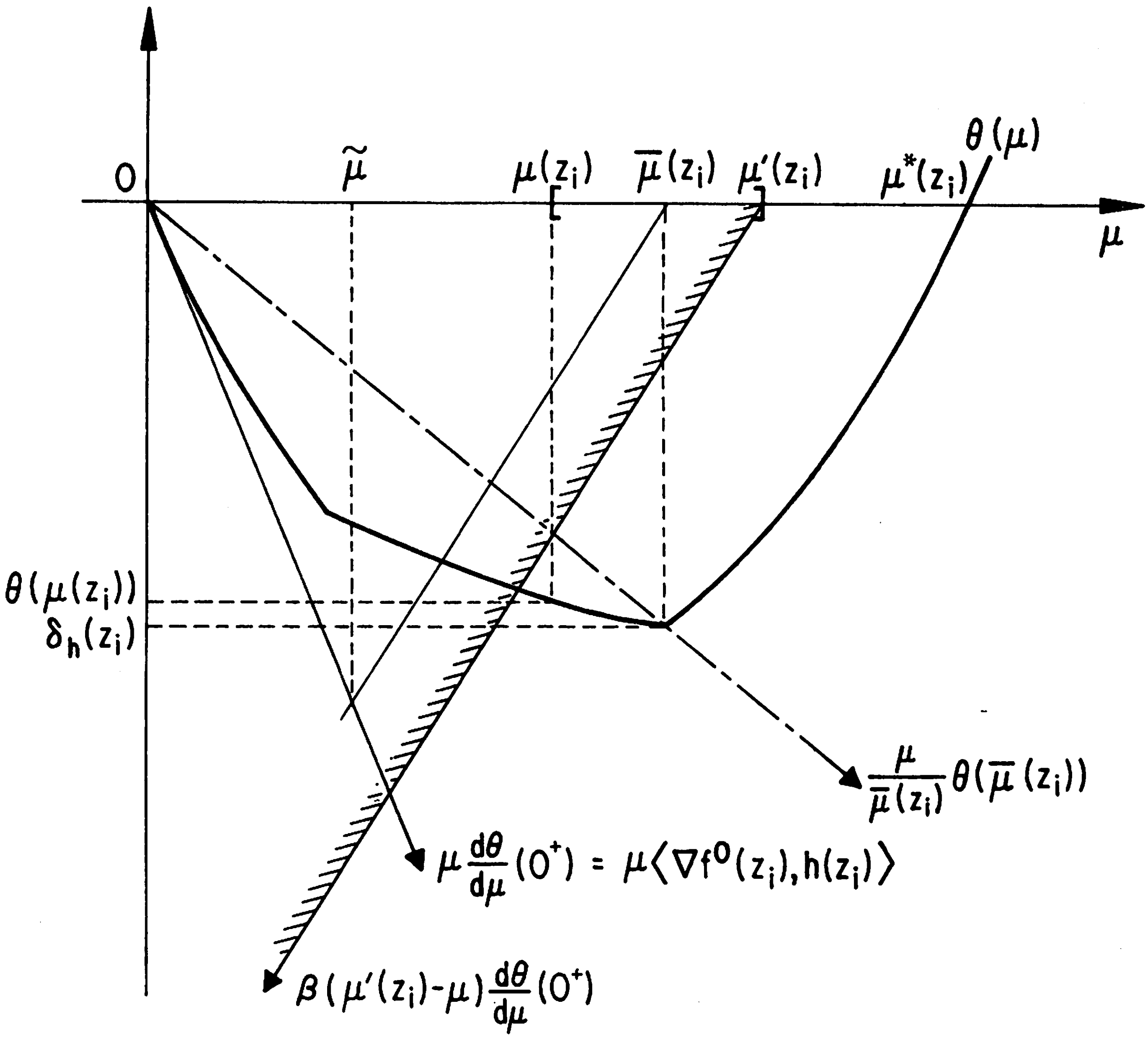
[9] Polak, E., Theory of Optimal Control and Mathematical Programming, McGraw-Hill Co., New York, 1970, theorem 19, p. 64 and lemma 1, p. 68.

[10] Rockafellar, R. T., Convex Analysis, Princeton University Press, Princeton, New Jersey, 1970.

[11] Topkis, D., and A. Veinnott, "On the Convergence of Some Feasible Direction Algorithms for Nonlinear Programming," SIAM J. on Control,

[12] Wolfe, P., "A Duality Theorem for Nonlinear Programming, Quarterly of Applied Mathematics, 1961, pp. 239-244.

[13] Zoutendijk, G., "Method of Feasible Direction," Elsevier Publishing Co., Amsterdam, 1960.



0
$\tilde{\mu}$
$\mu(z_i)$
$\bar{\mu}(z_i)$
$\mu'(z_i)$
$\mu^*(z_i)$
$\theta(\mu)$
μ
$\theta(\mu(z_i))$
$\delta_h(z_i)$
$\frac{\mu}{\bar{\mu}(z_i)}\theta(\bar{\mu}(z_i))$
$\mu\frac{d\theta}{d\mu}(0^+) = \mu\langle\nabla f^0(z_i), h(z_i)\rangle$
$\beta(\mu'(z_i)-\mu)\frac{d\theta}{d\mu}(0^+)$

Fig. 1.

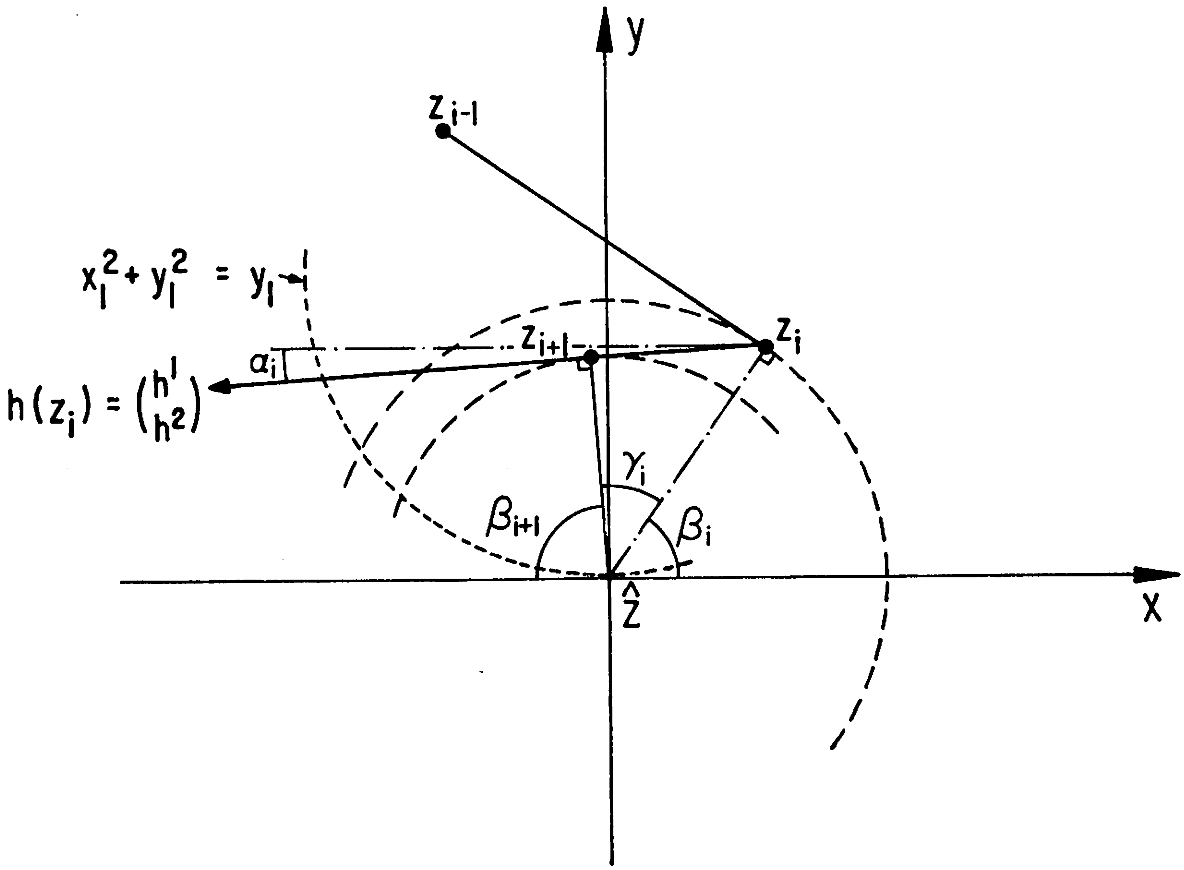

y
z_{i-1}
$x_1^2 + y_1^2 = y_1$
z_i
z_{i+1}
α_i
$h(z_i) = \binom{h^1}{h^2}$
γ_i
β_{i+1}
β_i
$\hat{z}$
x

Fig. 2.