

Copyright © 1977, by the author(s).
All rights reserved.

Permission to make digital or hard copies of all or part of this work for personal or classroom use is granted without fee provided that copies are not made or distributed for profit or commercial advantage and that copies bear this notice and the full citation on the first page. To copy otherwise, to republish, to post on servers or to redistribute to lists, requires prior specific permission.

INTERLIBRARY LOAN DEPARTMENT
(PHOTODUPLICATION SECTION)
THE GENERAL LIBRARY
UNIVERSITY OF CALIFORNIA
BERKELEY, CALIFORNIA 94720

COMBINED PHASE I, PHASE II METHODS
OF FEASIBLE DIRECTIONS

by

E. Polak, R. Trahan and D. Q. Mayne

Memorandum No. UCB/ERL M77/39

31 May 1977

ELECTRONICS RESEARCH LABORATORY
College of Engineering
University of California, Berkeley
94720

COMBINED PHASE I, PHASE II METHODS
OF FEASIBLE DIRECTIONS

E. Polak, R. Trahan and D. Q. Mayne

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

Abstract

This paper presents several new algorithms, generalizing feasible directions algorithms, for the nonlinear programming problem, $\min\{f^0(z) \mid f^j(z) \leq 0, j=1,2,\dots,m\}$. These new algorithms do not require an initial feasible point. They automatically combine the operations of initialization (Phase I) and optimization (Phase II) in an efficient manner.

Research sponsored by the National Science Foundation (RANN) Grant ENV76-04264 and the National Science Foundation Grant ENG73-08214-A01.

I. Introduction

In solving problems in engineering design, we have found that the usual procedure for computing an initial feasible point, for later use in a method of feasible directions, was very time consuming since it was necessary to set up an auxiliary optimization problem. This was costly in programming time and required a separate computer run. Also the auxiliary problem generated a feasible point without regard to the cost function of the original problem. As a result, we attempted to construct a procedure which automatically constructs feasible points while not completely ignoring the cost function. It turned out that one can construct algorithms, derived from classical methods of feasible directions, which combine directly the operations of initialization and optimization, in a very efficient manner. The resulting algorithms are to methods of feasible directions what the combined phase I-phase II simplex algorithm is to linear programming. Our initial findings were reported in our work on computer-aided design [1,2,3,4] where we dealt with very complex problems with infinitely many inequality constraints. Since our results are scattered piecemeal and since the complexity of the engineering design problems tends to obscure our results on combined phase I-phase II methods, we present in this paper the algorithm model as well as examples of modified "optimality" functions on which our new algorithms are based, together with some previously unpublished phase I-phase II algorithms which we find to be particularly efficient computationally. From our results, it should be obvious to the reader how to construct his own phase I-phase II method should he prefer a different method of feasible directions to the ones which served us as a starting point.

II. An Algorithm Model

In this paper we consider nonlinear programming problems of the form

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

where $f^j : \mathbb{R}^n \rightarrow \mathbb{R}^1$, $j = 0, 1, 2, \dots, m$ are continuously differentiable functions. The problem of computing an initial feasible point for (1) can be seen as being related to the problem

$$\min_{z \in \mathbb{R}^n} \psi(z) \quad (2)$$

with $\psi : \mathbb{R}^n \rightarrow \mathbb{R}^1$ defined by

$$\psi(z) \triangleq \max_{j \in \underline{m}} f^j(z) \quad (3)$$

where $\underline{m} \triangleq \{1, 2, \dots, m\}$. That is, if a feasible point z_0 for (1) exists, then $\psi(z_0) \leq 0$, and such a z_0 can be computed by solving (2), since any solution z'_0 to (2) must then satisfy $\psi(z'_0) \leq \psi(z_0) \leq 0$ and hence is feasible for (1). Thus, we really have two cost functions to contend with in solving (1): $f^0(\cdot)$ which is of interest in the feasible set

$$F \triangleq \{z \mid f^j(z) \leq 0, j = 1, 2, \dots, m\} \quad (4)$$

and $\psi(\cdot)$ which is of interest in F^c , the complement of F . We shall say that a point $z \in F$ is stationary if it satisfies the F. John condition [5] for (1), viz. for some multipliers $\mu^0 \geq 0, \mu^1 \geq 0, \dots, \mu^m \geq 0$, not all zero,

$$\sum_{j=1}^m \mu^j f^j(z) = 0; \sum_{j=0}^m \mu^j \nabla f^j(z) = 0 \quad (5)$$

We shall denote by S the set of feasible stationary points ($S \subset F$). Our algorithms define an iteration map $A : \mathbb{R}^n \rightarrow 2 \mathbb{R}^n$ and have the following structure:

Algorithm Model

Data: $z_0 \in \mathbb{R}^n$.

Step 0: Set $i = 0$.

Step 1: If $z_i \in S$, stop; otherwise compute a $z_{i+1} \in A(z_i)$.

Step 2: Set $i = i+1$ and go to step 1. □

A form of the following result can be found in [1].

Theorem 1: Let f^0 , ψ be as defined earlier and suppose that

(i) $A(F) \subset F$.

(ii) For any $z \in \mathbb{R}^n$, such that $z \notin S$, there exist an $\varepsilon(z) > 0$ and a $\delta(z) < 0$ such that

$$f^0(z'') - f^0(z') \leq \delta(z) < 0, \quad \forall z' \in B(z, \varepsilon(z)) \cap F \quad (6)$$
$$\forall z'' \in A(z')$$

$$\psi(z'') - \psi(z') \leq \delta(z) < 0, \quad \forall z' \in B(z, \varepsilon(z)) \cap F^c \quad (7)$$
$$\forall z'' \in A(z')$$

where $B(z, \varepsilon) \triangleq \{z' \in \mathbb{R}^n \mid \|z' - z\| \leq \varepsilon\}$. Then every accumulation point \hat{z} , of an infinite sequence $\{z_i\}$ generated by the Algorithm Model, satisfies $\hat{z} \in S$.[†]

Proof: Suppose $\{z_i\}$ is an infinite sequence constructed by the Algorithm Model which has an accumulation point $\hat{z} \notin S$ (i.e. for some infinite subset $K \subset \{0, 1, 2, \dots\}$ $z_i \xrightarrow{K} \hat{z}$, and $\hat{z} \notin S$). We must consider two cases. (1) For some i' , $z_{i'} \in F$. Then, because of (i) $z_i \in F$ for all $i \geq i'$, and since no $z_i \in S$ (otherwise the sequence would not be infinite), it follows from (6) that $\{f^0(z_i)\}_{i=i'}^\infty$ is a monotonically decreasing sequence, i.e., $f^0(z_{i+1}) < f^0(z_i)$ for all $i \geq i'$. Since

[†] Note that the theorem does not claim that accumulation points exist. Accumulation points will obviously exist if the sequence z_i is bounded.

$f^0(\cdot)$ is continuous and $z_i \xrightarrow{K} \hat{z}$, it follows that $f^0(z_i) \xrightarrow{K} f^0(\hat{z})$. But $\{f^0(z_i)\}_{i=1}^{\infty}$ is monotonic decreasing, and hence we must have $f^0(z_i) \rightarrow f^0(\hat{z})$ as $i \rightarrow \infty$. Consequently, $\lim[f^0(z_{i+1}) - f^0(z_i)] = 0$, and hence

$$\lim_{i \in K} [f^0(z_{i+1}) - f^0(z_i)] = 0 \quad (8)$$

follows trivially. But from (6), we must have

$$\lim_{i \in K} [f^0(z_{i+1}) - f^0(z_i)] \leq \delta(\hat{z}) < 0 \quad (9)$$

and so we get a contradiction. We conclude, therefore, that $\hat{z} \in S$.

(2) Now suppose that $\{z_i\}_{i=0}^{\infty} \subset F^c$. It now follows from (7) that $\{\psi(z_i)\}_{i=0}^{\infty}$ is a monotonically decreasing sequence, and, since the subsequence $\{\psi(z_i)\}_{i \in K}$ must converge to $\psi(\hat{z})$ (because ψ is continuous and $z_i \xrightarrow{K} \hat{z}$), we conclude that $\psi(z_i) \rightarrow \psi(\hat{z})$ and hence that $\lim[\psi(z_{i+1}) - \psi(z_i)] = 0$, so that

$$\lim_{i \in K} [\psi(z_{i+1}) - \psi(z_i)] = 0 \quad (10)$$

But from (7) we have

$$\lim_{i \in K} [\psi(z_{i+1}) - \psi(z_i)] \leq \delta(\hat{z}) < 0 \quad (11)$$

and we have again a contradiction. We conclude again that $\hat{z} \in S$ must hold. This completes our proof. □

III. Optimality Functions for Phase I-Phase II Algorithms

The computation of a feasible direction for a classical method of feasible directions is based on the solution of a linear or quadratic program which defines an optimality function. We shall

give several examples. For any $z \in F$, and any $\epsilon \geq 0$, let

$$I_\epsilon(z) = \{j \in \underline{m} \mid f^j(z) \geq -\epsilon\} \quad (12)$$

Then we find that, among others, Zoutendijk [6] used the optimality function

$$\theta_\epsilon^1(z) \triangleq \min_{h \in C} \max_{j \in \{0\} \cup I_\epsilon(z)} \langle \nabla f^j(z), h \rangle \quad (13)$$

with $C \triangleq \{h \in \mathbb{R}^n \mid |h^i| \leq 1, i = 1, 2, \dots, n\}$. Topkis and Veinott [7] used the optimality function

$$\theta^2(z) \triangleq \min_{h \in C} \max\{\langle \nabla f^0(z), h \rangle; f^j(z) + \langle \nabla f^j(z), h \rangle, j \in \underline{m}\} \quad (14)$$

Pironneau and Polak [8] used the optimality function

$$\theta^3(z) \triangleq \max_{\mu} \left\{ \sum_{j=1}^m \mu^j f^j(z) - \frac{1}{2} \left\| \sum_{j=0}^m \mu^j \nabla f^j(z) \right\|^2 \mid \sum_{j=0}^m \mu^j = 1, \mu^j \geq 0, \right. \\ \left. j \in \{0\} \cup \underline{m} \right\} \quad (15)$$

and, more recently, Polak and Trahan [3] used the optimality function

$$\theta_\epsilon^4(z) \triangleq \max_{\mu} \left\{ -\frac{1}{2} \left\| \sum_{j \in I_\epsilon(z) \cup \{0\}} \mu^j \nabla f^j(z) \right\|^2 \mid \sum_{j \in I_\epsilon(z) \cup \{0\}} \mu^j = 1, \right. \\ \left. \mu^j \geq 0, j \in I_\epsilon(z) \cup \{0\} \right\} \quad (16)$$

Zukhovitski, Polyak, and Primak [9] used the optimality function

$$\theta_\epsilon^5(z) \triangleq \begin{cases} \min_{h \in C} \{ \langle \nabla f^0(z), h \rangle \mid \langle \nabla f^j(z), h \rangle + \epsilon \leq 0, j \in I_\epsilon(z) \} \\ \text{if a feasible } h \text{ exists} \\ 0 \text{ otherwise} \end{cases} \quad (17)$$

These optimality functions can be modified for the combined Phase I-Phase II algorithm as follows. First, we define $\psi_0 : \mathbb{R}^n \rightarrow \mathbb{R}$ by

$$\psi_0(z) \triangleq \max\{0, \psi(z)\} \quad (18)$$

Then for any $z \in \mathbb{R}^n$, and any $\epsilon \geq 0$, let

$$J_\epsilon(z) \triangleq \{j \in \underline{m} \mid f^j(z) - \psi_0(z) \geq -\epsilon\} \quad (19)$$

We now give the modified optimality function corresponding to those defined by (13)-(17).[†] Let the weighting coefficient γ satisfy

$\gamma \geq 1$; then we define

$$\tilde{\theta}_\epsilon^1(z) \triangleq \min_{h \in C} \max\{\langle \nabla f^0(z), h \rangle - \gamma \psi_0(z); \langle \nabla f^j(z), h \rangle, j \in J_\epsilon(z)\} \quad (20)$$

$$\begin{aligned} \tilde{\theta}_\epsilon^2(z) \triangleq \min_{h \in C} \max\{\langle \nabla f^0(z), h \rangle - \gamma \psi_0(z); \\ f^j(z) - \psi_0(z) + \langle \nabla f^j(z), h \rangle, j \in \underline{m}\} \end{aligned} \quad (21)$$

$$\begin{aligned} \tilde{\theta}_\epsilon^3(z) \triangleq \max_{\mu} \left\{ \sum_{j=1}^m \mu^j (f^j(z) - \psi_0(z)) - \gamma \mu^0 \psi_0(z) \right. \\ \left. - \frac{1}{2} \left\| \sum_{j=0}^m \mu^j \nabla f^j(z) \right\|^2 \mid \sum_{j=0}^m \mu^j = 1, \mu^j \geq 0, j = 0, 1, \dots, m \right\} \end{aligned} \quad (22)$$

$$\begin{aligned} \tilde{\theta}_\epsilon^4(z) \triangleq \max_{\mu} \left\{ -\gamma \mu^0 \psi_0(z) - \frac{1}{2} \left\| \sum_{j \in J_\epsilon(z) \cup \{0\}} \mu^j \nabla f^j(z) \right\|^2 \right. \\ \left. \sum_{j \in J_\epsilon(z) \cup \{0\}} \mu^j = 1; \mu^j \geq 0, j \in J_\epsilon(z) \cup \{0\} \right\} \end{aligned} \quad (23)$$

$$\tilde{\theta}_\epsilon^5(z) \triangleq \begin{cases} \min_{\substack{h^0 \leq -\epsilon, \\ h \in C}} \{ \langle \nabla f^0(z), h \rangle + \psi_0(z)(\gamma h^0 - 1) \mid \\ \langle \nabla f^j(z), h \rangle - h^0 \leq 0, j \in J_\epsilon(z) \} \\ \text{if a feasible } (h^0, h) \text{ exists} \\ 0 \text{ otherwise.} \end{cases} \quad (24)$$

[†]To obtain a uniform representation, we use the subscript ϵ on all of them whether ϵ is functional or not.

For $\pi \in \{1, 2, 5\}$, denote the solutions of the program for $\tilde{\theta}_\epsilon^\pi(z)$ by $h_\epsilon^\pi(z)$. Let $\mu_{3,\epsilon}(z)$ and $\mu_{4,\epsilon}(z)$ denote the solutions of the programs for $\tilde{\theta}_\epsilon^3(z)$ and $\tilde{\theta}_\epsilon^4(z)$ respectively. We define the "descent" direction vectors

$$h_\epsilon^3(z) \triangleq \sum_{j=0}^m \mu_{3,\epsilon}^j(z) \nabla f^j(z) \quad (25)$$

$$h_\epsilon^4(z) \triangleq \sum_{j \in J_\epsilon(z) \cup \{0\}} \mu_{4,\epsilon}^j(z) \nabla f^j(z) \quad (26)$$

Note again that in the definitions of $\tilde{\theta}_\epsilon^\pi(z)$ and $h_\epsilon^\pi(z)$ for $\pi = 2$ or 3 , ϵ is only a dummy parameter and is included for notational convenience only.

We assume the following hypotheses are satisfied by problem (1).

Assumption 1. The functions $f^j : \mathbb{R}^n \rightarrow \mathbb{R}$ $j = 0, 1, 2, \dots, m$ are continuously differentiable. □

Assumption 2: For any $z \in \mathbb{R}^n$, the set of vectors $\{\nabla f^j(z), j \in J_0(z)\}$ is positive linearly independent. †
□

Assumption 2 is a sufficient condition for the Kuhn-Tucker constraint qualification to hold [10]. It guarantees that $\text{int } F = F$ and also that $\overline{\text{int } F} = F$.^{††} Furthermore, it ensures that each optimality function (with $\epsilon = 0$) is zero only at non-degenerate stationary points, i.e., $\mu^0 > 0$ in the F. John equation, (5) (which then becomes the Kuhn-Tucker condition). It also ensures that each optimality function (with $\epsilon = 0$) is strictly negative for all $z \in F^c$.

[†] We say that a set of vectors $\{\eta_j, j = 1, 2, \dots, n\}$ is positive linearly independent if the zero vector is not contained in the convex hull of $\{\eta_j, j = 1, 2, \dots, n\}$.

^{††} Int denotes interior and the overbar denotes closure.

Finally, we note that since the modified optimality functions coincide with the original ones on F , we must have $S = \{z \in F | \tilde{\theta}_0^\pi(z) = 0\}$ for any value of π . We formalize our remarks about optimality functions in the following lemma which is proved in the Appendix.

Lemma 1. Let $\pi \in \{1,2,3,4,5\}$ be given. (i) If $\hat{z} \in F$ is optimal for (1) then $\tilde{\theta}_0^\pi(\hat{z}) = 0$. (ii) For any $z \in F^c$, $\tilde{\theta}_0^\pi(z) < 0$. (iii) $\tilde{\theta}_0^\pi(z) = 0$ if and only if $z \in S$. □

It should be clear from the preceeding discussion that it is possible to modify most, if not all, optimality functions for problem (1) so as to obtain new optimality functions with the properties given by Lemma 1 for $\tilde{\theta}_0^\pi(z)$. Besides having the properties given in Lemma 1, each optimality function gives rise to a "descent" direction vector in the algorithm below. We have modified the original optimality functions so that the effect of the cost function, on this "descent" direction vector, is suppressed in a continuous manner (since $\psi(z)$ is continuous) for all $z \in F^c$. Thus, for $\pi \in \{1,2,5\}$, whenever $\psi(z) > 0$, the effect of the term $\langle \nabla f^0(z), h \rangle$ is suppressed in $\tilde{\theta}_\epsilon^\pi(z)$, in proportion to $\psi(z)$. For $\pi \in \{3,4\}$ the value of $\mu_{\pi,\epsilon}^0(z)$, (see (25), (26)) is decreased as $\psi(z)$ increases. Hence, the effect of the cost gradient $\nabla f^0(z)$ on $h_\epsilon^\pi(z)$ is also reduced for $\pi \in \{3,4\}$, when $\psi(z)$ is large. The effect of suppressing the cost in this manner is that the algorithm concentrates on decreasing $\psi(z)$ at an infeasible point z without totally ignoring the cost function. The effect of the cost function becomes progressively greater as the feasible set F is approached.

IV. The Algorithms

We now present an algorithm in which any of the above optimality functions can be used. A single data parameter, $\pi \in \{1,2,3,4,5\}$ is chosen for the desired optimality function. Each optimality function gives rise to a different direction-finding problem and hence, effectively, to a different algorithm.

The algorithm below is stated in a form which minimizes the use of "go to" statements so as to make the algorithm easier to follow.

If the algorithm were programmed in FORTRAN in this way, a number of unnecessary "if" statements would be executed. Hence, when the algorithm is coded in FORTRAN a number of "go to" statements should be inserted so as to avoid unnecessary operations.

Algorithm

Data: $\pi \in \{1,2,3,4,5\}$, $\alpha \in (0,1)$, $\beta \in (0,1)$, $\gamma \geq 1$, $\delta \in (0,1]$, $\epsilon_0 > 0$, $0 < \epsilon_1 < \epsilon_0$, $M > 0$, $z_0 \in \mathbb{R}^n$.

Step 0: Set $i=0$.

Step 1: Set $\epsilon = \epsilon_0$.

Step 2: Compute $(\tilde{\theta}_\epsilon^\pi(z_i), h_\epsilon^\pi(z_i))$. If $\pi \in \{2,3\}$, set $\epsilon = -\frac{1}{\delta} \tilde{\theta}_\epsilon^\pi(z_i)$ and $\tilde{\theta}_0^\pi(z_i) = \tilde{\theta}_\epsilon^\pi(z_i)$; else, proceed.

Step 3: If $\pi \in \{1,4,5\}$ and $\epsilon < \epsilon_1$, compute $\tilde{\theta}_0^\pi(z_i)$; else, proceed.

Step 4: If $\{\pi \in \{1,4,5\}, \epsilon < \epsilon_1, \text{ and } \tilde{\theta}_0^\pi(z_i) = 0\}$, or if $\{\pi \in \{2,3\} \text{ and } \tilde{\theta}_0^\pi(z_i) = 0\}$, stop; else, proceed.

Step 5: If $\tilde{\theta}_\epsilon^\pi(z_i) > -\delta\epsilon$, set $\epsilon = \epsilon/2$ and go to step 2; else, proceed.

Step 6: Let $\hat{M} \triangleq \max\{1, M/\|h_\epsilon^\pi(z_i)\|_\infty\}$. If $z_i \in F$, compute the largest step size $s_i = \beta^{\ell_i} \in (0, \hat{M}]$, (ℓ_i an integer) satisfying

$$f^0(z_i + s_i h_\epsilon^\pi(z_i)) - f^0(z_i) \leq -\alpha \delta \epsilon s_i \quad (27)$$

$$f^j(z_i + s_i h_\epsilon^\pi(z_i)) \leq 0 \quad j \in \underline{m} \quad (28)$$

If $z_i \in F^c$, compute the largest step size $s_i = \beta^{\ell_i} \in (0, \hat{M}]$, (ℓ_i an integer) satisfying

$$\psi(z_i + s_i h_\epsilon^\pi(z_i)) - \psi(z_i) \leq -\alpha \delta \epsilon s_i \quad (29)$$

Step 7: Set $z_{i+1} = z_i + s_i h_\epsilon^\pi(z_i)$, $i = i+1$ and go to step 1. \square

Comment: Another version of the algorithm returns from step 7 to step 2 instead of step 1. The proof for this version is somewhat harder, but quite standard and we shall omit it. \square

To establish convergence of the algorithm we show that the hypotheses of Theorem 1 are satisfied. We first state a result which shows that the algorithm is well-defined.

Lemma 2. The algorithm cannot cycle indefinitely between steps 2 and 5. \square

The lemma is obviously true for $\pi \in \{2, 3\}$. For the case, $\pi \in \{1, 4, 5\}$, the proof is identical to that of Lemma 4.3.27 in [11].

For $\pi \in \{1, 2, 3, 4, 5\}$ we can define $A^\pi : \mathbb{R}^n \rightarrow 2^{\mathbb{R}^n}$ by steps 1 through 7 of the algorithm with the program defining $\tilde{\theta}_\epsilon^\pi(z_i)$ used to calculate the direction vector $h_\epsilon^\pi(z_i)$. In order to use Theorem 1 we make use of the following result which is proved in Appendix A.

Lemma 3. Given $\pi \in \{1,2,3,4,5\}$ and any $z \in \mathbb{R}^n$ such that $\tilde{\theta}_0^\pi(z) < 0$, there exist a $\gamma(z) > 0$ and a $\delta(z) < 0$ such that

$$\begin{aligned} f^0(z'') - f^0(z') &\leq \delta(z) < 0 \quad \forall z' \in B(z, \gamma(z)) \cap F \\ &\quad \forall z'' \in A^\pi(z') \end{aligned} \quad (30)$$

$$\begin{aligned} \psi(z'') - \psi(z') &\leq \delta(z) < 0 \quad \forall z' \in B(z, \gamma(z)) \cap F^c \\ &\quad \forall z'' \in A^\pi(z'). \end{aligned} \quad (31)$$

□

Since the algorithm stops only at points which are stationary, i.e., satisfying $\theta_0^\pi(z) = 0$, we need to consider only the case when $\{z_i\}$ is infinite. We now state our main convergence result.

Theorem 2. Given $\pi \in \{1,2,3,4,5\}$ and any infinite sequence $\{z_i\}$ constructed by the algorithm, every accumulation point \hat{z} of $\{z_i\}$ is stationary, i.e., $\theta_0^\pi(\hat{z}) = 0$.

Proof: It is obvious that $A^\pi(F) \subset F$ since step 6 maintains feasibility for any $z_i \in F$. Thus, hypothesis (i) of Theorem 1 is satisfied.

From Lemma 1 we have that $z \notin S$ if and only if

$\tilde{\theta}_0^\pi(z) < 0$. Hence, we have immediately from Lemma 3 that hypothesis (ii) of Theorem 1 is satisfied. □

V. Conclusions

In summary, we have shown that by modifying the optimality functions used in several conventional methods of feasible directions, it is possible to construct algorithms which combine directly the phase I - phase II operations. Specifically, each optimality function is modified (i) by adding a term involving $\psi_0(z) = \max\{0, f^1(z), \dots, f^m(z)\}$ to the term involving the cost function, $f^0(\cdot)$, and (ii) by extending the

definition of the ϵ -active constraint set of infeasible points as
, being the set of constraints which are ϵ -active with respect to the maximum
constraint whenever $\psi(z) > 0$. As a result of these modifications, the
new algorithms can be initialized at any point in \mathbb{R}^n . In the initial
iterations, the algorithms concentrate on decreasing $\psi(z)$
 $= \max\{f^1(z), \dots, f^m(z)\}$ (if $\psi(z) > 0$) while not completely ignoring the
cost function, whose effect becomes progressively more pronounced as
the feasible set is approached. The algorithms are demonstrably convergent
in the sense that if an infinite sequence is constructed, then any
accumulation point of the sequence (if one exists) is feasible and
satisfies a first order necessary condition of optimality.

It should be noted that the algorithms presented here can be
considered as special cases of more general combined phase I - phase II
methods of feasible directions. Each algorithm in this paper can be
modified to handle more general problems, such as those which contain
equality constraints [14], or functional constraints of the form
 $\max_{\omega \in \Omega} \phi(z, \omega) \leq 0, \Omega \subset \mathbb{R}$ (Ω a compact interval) [1].

Appendix A

We now establish the Lemmas 1 and 3 used in the body of the paper.

Lemma 1. Let $\pi \in \{1,2,3,4,5\}$ be given. (i) If $\hat{z} \in F$ is optimal for (1) then $\tilde{\theta}_0^\pi(\hat{z}) = 0$. (ii) For any $z \in F^C$, $\tilde{\theta}_0^\pi(z) < 0$. (iii) $\tilde{\theta}_0^\pi(z) = 0$ if and only if $z \in S$.

Proof:

(i) Since \hat{z} is feasible we have $\tilde{\theta}_0^\pi(\hat{z}) = \theta_0^\pi(\hat{z})$ for each $\pi \in \{1,2,3,4,5\}$. But $\theta_0^\pi(\hat{z}) = 0$, $\pi \in \{1,2,3,4,5\}$ is a well-known necessary condition of optimality for (1) [8,11].

(ii) Suppose $z \in F^C$, i.e., $\psi_0(z) > 0$. As a consequence of Assumption 3, there exists $h \in C$ such that $\langle \nabla f^j(z), h \rangle < 0$, for all $j \in J_0(z)$. Because $\psi_0(z) > 0$ and $f^j(z) - \psi_0(z) < 0$, for all $j \notin J_0(z)$, $j \neq 0$, there exists $\lambda \in (0,1)$ such that $\langle \nabla f^0(z), \lambda h \rangle - \psi_0(z) < 0$ and $f^j(z) - \psi_0(z) + \lambda \langle \nabla f^j(z), h \rangle < 0$ for all $j \notin J_0(z)$, $j \neq 0$. Since $\lambda h \in C$, this implies that $\tilde{\theta}_0^\pi(z) < 0$ for $\pi \in \{1,4,5\}$.

For $\pi \in \{3,4\}$, we first note that by duality [12] we can write $\tilde{\theta}_\epsilon^3(z)$ and $\tilde{\theta}_\epsilon^4(z)$ as

$$\begin{aligned} \tilde{\theta}_\epsilon^3(z) = \min_{h \in \mathbb{R}^n} \{ \frac{1}{2} \|h\|^2 + \max\{ \langle \nabla f^0(z), h \rangle - \gamma \psi_0(z); \\ f^j(z) - \psi_0(z) + \langle \nabla f^j(z), h \rangle, j \in \underline{m} \} \} \end{aligned} \quad (A1)$$

$$\begin{aligned} \tilde{\theta}_\epsilon^4(z) = \min_{h \in \mathbb{R}^n} \{ \frac{1}{2} \|h\|^2 + \max\{ \langle \nabla f^0(z), h \rangle - \gamma \psi_0(z); \\ \langle \nabla f^j(z), h \rangle, j \in J_\epsilon(z) \} \} \end{aligned} \quad (A2)$$

Note that the solution vectors of (A1) and (A2) are the same as those given by (25) and (26) respectively, i.e. $h_\epsilon^3(z)$ ($h_\epsilon^4(z)$) can be found

by solving either (22) or (A1) ((23) or (A2)).

Using the same arguments as above, we have that there exists a vector $\lambda h \in \mathbb{R}^n$ with $\lambda > 0$ and $h \in C$ such that

$$\begin{aligned} & \frac{1}{2} \lambda^2 \|h\|^2 + \max\{\langle \nabla f^0(z), \lambda h \rangle - \gamma \psi_0(z); \\ & f^j(z) - \psi_0(z) + \langle \nabla f^j(z), \lambda h \rangle, j \in \underline{m}\} < 0 \end{aligned} \quad (A3)$$

and

$$\begin{aligned} & \frac{1}{2} \lambda^2 \|h\|^2 + \max\{\langle \nabla f^0(z), \lambda h \rangle - \gamma \psi_0(z); \\ & \langle \nabla f^j(z), \lambda h \rangle, j \in J_0(z)\} < 0 \end{aligned} \quad (A4)$$

Hence, $\tilde{\theta}_0^\pi(z) < 0$ for $\pi \in \{3, 4\}$.

(iii) The fact that $\tilde{\theta}_0^\pi(z) = 0$ if and only if $z \in S$ follows from the fact that $\tilde{\theta}_0^\pi(z) < 0$ for all $z \in F^c$ and from the fact that for $z \in F$, $\tilde{\theta}_0^\pi(z) = 0$ if and only if $z \in S$, $\pi \in \{1, 2, 3, 4, 5\}$. \square

We now state some elementary results which are very easy to establish and, therefore, we omit the proofs.

Proposition 1. For any $z \in \mathbb{R}^n$, if $\varepsilon > \varepsilon'$, then (i) $J_\varepsilon(z) \supset J_{\varepsilon'}(z)$ and (ii) $\theta_{\varepsilon'}^\pi(z) \leq \theta_\varepsilon^\pi(z)$ for $\pi \in \{1, 2, 3, 4, 5\}$. \square

Proposition 2. For any $z \in \mathbb{R}^n$, $\varepsilon \geq 0$, there exists a $\rho > 0$ such that $J_\varepsilon(z') \subset J_\varepsilon(z)$, for all $z' \in B(z, \rho)$. \square

We now state and prove the following result which will be used in the proof of Lemma 3.

Proposition 3. Let $\pi \in \{1,2,3,4,5\}$ and $\varepsilon_0 > 0$, $\delta \in (0,1]$ be given.

For any $z \in \mathbb{R}^n$ such that $\tilde{\theta}_0^\pi(z) < 0$, there exists a $\rho(z) > 0$ such that $\tilde{\theta}_{\varepsilon(z)}^\pi(z) \leq -\delta\varepsilon(z)$, and

$$\tilde{\theta}_{\varepsilon(z')}^\pi(z') \leq -\delta\varepsilon(z') \leq -\frac{1}{2}\delta\varepsilon(z) \quad \forall z' \in B(z, \rho(z)) \quad (\text{A5})$$

where $\varepsilon(z)$ ($\varepsilon(z')$) is the value of ε constructed by steps 2 through 5 of the algorithm with $z_1 = z$ ($z_1 = z'$).

Proof: (i) For $\pi=2$ or 3, and $z_1 = z$, the algorithm sets $\varepsilon(z) = -\frac{1}{\delta} \tilde{\theta}_{\varepsilon(z)}^\pi(z) = -\frac{1}{\delta} \tilde{\theta}_0^\pi(z)$. Because $\tilde{\theta}_0^2(\cdot)$ and $\tilde{\theta}_0^3(\cdot)$ are continuous, there exists a $\rho(z) > 0$ such that for $\pi=2$ or 3,

$$\tilde{\theta}_0^\pi(z') \leq \frac{1}{2} \tilde{\theta}_0^\pi(z) = -\frac{1}{2} \delta\varepsilon(z) \quad \forall z' \in B(z, \rho(z)) \quad (\text{A6})$$

Also, for $z_1 = z'$ the algorithm constructs an $\varepsilon(z')$ such that

$$\tilde{\theta}_{\varepsilon(z')}^\pi(z') = -\delta\varepsilon(z') = \tilde{\theta}_0^\pi(z') \quad (\text{A7})$$

Combining (A6) and (A7) we have the desired result for $\pi=2$ or 3.

(ii) For $\pi \in \{1,4,5\}$ and for any $z \in \mathbb{R}^n$ such that $\tilde{\theta}_0^\pi(z) < 0$, it is an immediate consequence of Lemma 2 that the algorithm constructs an $\varepsilon(z) = \varepsilon_0 2^{-j(z)} > 0$ with $j(z)$ the smallest nonnegative integer such that

$$\tilde{\theta}_{\varepsilon(z)}^\pi(z) \leq -\delta\varepsilon(z) < 0 \quad (\text{A8})$$

From Proposition 2, there exists a $\hat{\rho}(z) > 0$ such that $J_{\varepsilon(z)}^\pi(z') \subset J_{\varepsilon(z)}^\pi(z)$, for all $z' \in B(z, \hat{\rho}(z))$. Let $\tilde{\theta}^\pi : \mathbb{R}^n \rightarrow \mathbb{R}$ for $\pi \in \{1,4,5\}$ be defined by

$$\begin{aligned} \tilde{\theta}^1(z') &\triangleq \min_{h \in C} \max\{\langle \nabla f^0(z'), h \rangle - \gamma \psi_0(z'); \\ &\quad \langle \nabla f^j(z'), h \rangle, j \in J_{\varepsilon(z)}^\pi(z)\} \end{aligned} \quad (\text{A9})$$

$$\begin{aligned}
\bar{\theta}^4(z') &\triangleq \max_{\mu} \left\{ -\frac{1}{2} \left\| \sum_{j \in J_{\varepsilon(z)}(z) \cup \{0\}} \mu^j \nabla f^j(z') \right\|^2 \right. \\
&\quad \left. - \gamma \mu^0 \psi_0(z') \right\} \Big| \sum_{j \in J_{\varepsilon(z)}(z) \cup \{0\}} \mu^j = 1, \\
\mu^j &\geq 0, \quad j \in J_{\varepsilon(z)}(z) \cup \{0\}
\end{aligned} \tag{A10}$$

$$\begin{aligned}
\bar{\theta}^5(z') &\triangleq \min_{\substack{h \in C \\ h^0 \leq -\varepsilon(z)}} \left\{ \langle \nabla f^0(z'), h \rangle + \psi_0(z') (\gamma h^0 - 1) \right\} \\
&\quad \langle \nabla f^j(z'), h \rangle - h^0 \leq 0, \quad j \in J_{\varepsilon(z)}(z)
\end{aligned} \tag{A11}$$

Given $\pi \in \{1, 4, 5\}$, $\bar{\theta}^\pi(\cdot)$ is a continuous function, and therefore there exists a $\rho(z) \in (0, \hat{\rho}(z)]$ such that

$$|\bar{\theta}^\pi(z') - \bar{\theta}^\pi(z)| \leq \delta \frac{\varepsilon(z)}{2} \quad \forall z' \in B(z, \rho(z)) \tag{A12}$$

But $\bar{\theta}^\pi(z) = \tilde{\theta}_{\varepsilon(z)}^\pi(z)$ so that

$$\bar{\theta}^\pi(z') \leq \delta \frac{\varepsilon(z)}{2} + \tilde{\theta}_{\varepsilon(z)}^\pi(z) \leq -\delta \frac{\varepsilon(z)}{2} \quad \forall z' \in B(z, \rho(z)) \tag{A13}$$

Comparing $\bar{\theta}^\pi(z')$ and $\tilde{\theta}_{\varepsilon(z)}^\pi(z')$ and noting that $J_{\varepsilon(z)}(z') \subset J_{\varepsilon(z)}(z)$, for all $z' \in B(z, \rho(z))$, we obtain

$$\tilde{\theta}_{\frac{\varepsilon(z)}{2}}^\pi(z') \leq \tilde{\theta}_{\varepsilon(z)}^\pi(z') \leq -\delta \frac{\varepsilon(z)}{2}, \quad \forall z' \in B(z, \rho(z)) \tag{A14}$$

where we have made use of (A13) and Proposition 1. Let $\varepsilon(z') \triangleq \varepsilon_0 2^{-j(z')}$

where $j(z')$ is the smallest non-negative integer such that

$$\tilde{\theta}_{\varepsilon_0 2^{-j(z')}}^\pi(z') \leq -\delta \varepsilon_0 2^{-j(z')}. \quad \text{Then from (A14) we have } \varepsilon(z') \geq \frac{\varepsilon(z)}{2},$$

and hence, $\tilde{\theta}_{\varepsilon(z')}^\pi(z') \leq -\delta \varepsilon(z') \leq -\delta \frac{\varepsilon(z)}{2}$, for all $z' \in B(z, \rho(z))$. \square

Lemma 3. Let $\pi \in \{1, 2, 3, 4, 5\}$. For any $z \in \mathbb{R}^n$ such that

$\tilde{\theta}_0^\pi(z) < 0$, there exist a $\gamma(z) > 0$ and a $\delta(z) < 0$ such that

$$\begin{aligned} f^0(z'') - f^0(z') &\leq \delta(z) < 0 & \forall z' \in B(z, \gamma(z)) \cap F \\ & & \forall z'' \in A^\pi(z') \end{aligned} \quad (A15)$$

$$\begin{aligned} \psi(z'') - \psi(z') &\leq \delta(z) < 0 & \forall z' \in B(z, \gamma(z)) \cap F^c \\ & & \forall z'' \in A^\pi(z') \end{aligned} \quad (A16)$$

Proof: Given $\pi \in \{1, 2, 3, 4, 5\}$ and $z \in \mathbb{R}^n$ such that $\tilde{\theta}_0^\pi(z) < 0$, it follows from Proposition 3 that there exists a $\rho_1(z) > 0$ such that

$$\tilde{\theta}_{\varepsilon(z')}^\pi(z') \leq -\delta\varepsilon(z') \leq -\frac{1}{2} \delta\varepsilon(z) \quad \forall z' \in B(z, \rho_1(z)) \quad (A17)$$

where $\varepsilon(z')$ is the value of ε constructed by the algorithm with $z_i = z'$, and $\varepsilon(z)$ is the value corresponding to $z_i = z$.

(i) For $\pi \in \{3, 4\}$, let $h^\pi : \mathbb{R}^n \times \mathbb{R}^{m+1} \rightarrow \mathbb{R}^n$ be defined by

$$h^\pi(z', \hat{\mu}) = -\sum_{j=0}^m \hat{\mu}^j \nabla f^j(z')$$

Let $C^\pi \subset \mathbb{R}^n$ be the image of $B(z, \rho_1(z)) \times \{\mu \in \mathbb{R}^{m+1} \mid \sum_{j=0}^m \mu^j = 1, \mu^j \geq 0, j = 0, 1, \dots, m\}$ under h^π . Because $h^\pi(\cdot, \cdot)$ is continuous and C^π is the image of a compact set, C^π must also be compact.

(ii) For $\pi \in \{1, 2, 5\}$, let $C^\pi = C$.

In view of (i) and (ii) we have that C^π is compact for $\pi \in \{1, 2, 3, 4, 5\}$. Hence, for any $\pi \in \{1, 2, 3, 4, 5\}$, there exists a $\rho_2(z) \in (0, \rho_1(z)]$ and a $t_1(z) \in (0, 1)$ such that

$$\begin{aligned} |f^j(z' + th^\pi) - f^j(z')| &\leq (1-\alpha)\delta \frac{\varepsilon(z)}{2} & \forall z' \in B(z, \rho_2(z)) \\ & & \forall h^\pi \in C^\pi \\ & & \forall t \in [0, t_1(z)] \\ & & \forall j = 0, 1, 2, \dots, m \end{aligned} \quad (A18)$$

Also, there exists a $\rho_3(z) \in (0, \rho_2(z)]$ and $t_2(z) \in (0, t_1(z)]$ such that for $j = 0, 1, 2, \dots, m$

$$\begin{aligned}
|\langle \nabla f^j(z' + th^\pi), h^\pi \rangle - \langle \nabla f^j(z'), h^\pi \rangle| &\leq (1-\alpha)\delta \frac{\varepsilon(z)}{2} \\
\forall z' &\in B(z, \rho_3(z)) \\
\forall h^\pi &\in C^\pi \\
\forall t &\in [0, t_2(z)] \quad (A19)
\end{aligned}$$

Let $\tilde{C}^\pi(z') \subset C^\pi$ be the set of all direction vectors obtained by solving the program for $\tilde{\theta}_{\varepsilon(z')}^\pi(z')$.[†] By the mean-value theorem, for any $z' \in B(z, \rho_3(z))$ and $h_{\varepsilon(z')}^\pi(z') \in \tilde{C}^\pi(z')$, and for any $t \in [0, t_2(z)]$

$$\begin{aligned}
f^j(z' + th_{\varepsilon(z')}^\pi(z')) &= f^j(z') + t \langle \nabla f^j(z' + \xi^j h_{\varepsilon(z')}^\pi(z')), h_{\varepsilon(z')}^\pi(z') \rangle, \\
h_{\varepsilon(z')}^\pi(z') &\quad j = 0, 1, \dots, m \quad (A20)
\end{aligned}$$

where $\xi^j \in [0, t]$. From (A19) and (A20) we get

$$\begin{aligned}
f^j(z' + th_{\varepsilon(z')}^\pi(z')) - f^j(z') &\leq (1-\alpha)t\delta \frac{\varepsilon(z)}{2} \\
&+ t \langle \nabla f^j(z'), h_{\varepsilon(z')}^\pi(z') \rangle \quad \forall z' \in B(z, \rho_3(z)) \\
&\quad \forall t \in [0, t_2(z)] \\
&\quad \forall h_{\varepsilon(z')}^\pi(z') \in \tilde{C}^\pi(z') \\
&\quad j = 0, 1, 2, \dots, m \quad (A21)
\end{aligned}$$

For $\pi \in \{2, 3\}$ it is obvious from the definition of $\tilde{\theta}_{\varepsilon(z')}^2(z')$, and the dual form of $\tilde{\theta}_{\varepsilon(z')}^3(z')$ given by (A1), that for all $j \in J_{\varepsilon(z')}^\pi(z')$

$$f^j(z') - \psi_0(z') + \langle \nabla f^j(z'), h_{\varepsilon(z')}^\pi(z') \rangle \leq -\delta\varepsilon(z') \quad (A22)$$

Because $f^j(z') - \psi_0(z') \leq 0$ for all $j = 1, 2, \dots, m$, (A22) must also hold for $\pi \in \{1, 4, 5\}$ for all $j \in J_{\varepsilon(z')}^\pi(z')$. Since $t_2(z) \in (0, 1)$, multiplying (A22) by $t \in [0, 1)$ and adding to (A21) yields

[†]For $\pi \in \{3, 4\}$, the program for $\tilde{\theta}_{\varepsilon(z')}^\pi(z')$ is solved for $\mu_{\pi, \varepsilon(z')}(z')$, and $h_{\varepsilon(z')}^\pi(z')$ is then obtained by computing as in (25) or (26).

$$\begin{aligned}
f^j(z' + th_{\varepsilon(z')}^\pi(z')) - \psi_0(z') &\leq f^j(z' + th_{\varepsilon(z')}^\pi(z')) + \\
(1-t)(\psi_0(z') - f^j(z')) - \psi_0(z') & \\
&\leq (1-\alpha)t\delta \frac{\varepsilon(z)}{2} - t\delta\varepsilon(z') \\
&\leq (1-\alpha)t\delta \varepsilon(z') - t\delta \varepsilon(z') \\
&= -\alpha t\delta \varepsilon(z')
\end{aligned}$$

$$\forall z' \in B(z, \rho_3(z))$$

$$\forall t \in [0, t_2(z)]$$

$$\forall j \in J_{\varepsilon(z')}(z')$$

$$\forall h_{\varepsilon(z')}^\pi(z') \in \tilde{C}^\pi(z')$$

(A23)

For $j \notin J_{\varepsilon(z')}(z')$, $j \neq 0$, it follows from the definition of $J_{\varepsilon(z')}(z')$ that $f^j(z') - \psi_0(z') < -\varepsilon(z') \leq -\delta\varepsilon(z')$. From (A18) we obtain

$$\begin{aligned}
f^j(z' + th_{\varepsilon(z')}^\pi(z')) - f^j(z') &\leq (1-\alpha)\delta \frac{\varepsilon(z)}{2} \quad \forall z' \in B(z, \rho_3(z)) \\
\forall t &\in [0, t_2(z)] \\
\forall h_{\varepsilon(z')}^\pi(z') &\in \tilde{C}^\pi(z') \\
\forall j &\notin J_{\varepsilon(z')}(z'), j \neq 0
\end{aligned}$$

(A24)

Thus,

$$\begin{aligned}
f^j(z' + th_{\varepsilon(z')}^\pi(z')) - \psi_0(z') &\leq (1-\alpha)\delta \frac{\varepsilon(z)}{2} - \delta\varepsilon(z') \\
&\leq (1-\alpha)\delta \varepsilon(z') - \delta \varepsilon(z') \\
&= -\alpha\delta \varepsilon(z') < -\alpha t\delta \varepsilon(z')
\end{aligned}$$

$$\forall z' \in B(z, \rho_3(z))$$

$$\forall t \in [0, t_2(z)]$$

$$\forall h_{\varepsilon(z')}^\pi(z') \in \tilde{C}^\pi(z')$$

(A25)

Therefore, from (A23) and (A25) we obtain

$$\psi(z' + th_{\varepsilon(z')}^{\pi}(z')) - \psi_0(z') \leq -\alpha t \delta \varepsilon(z')$$

$$\forall z' \in B(z, \rho_3(z))$$

$$\forall t \in [0, t_2(z)]$$

$$\forall h_{\varepsilon(z')}^{\pi}(z') \in \tilde{C}^{\pi}(z') \quad (A26)$$

Let $\hat{k}(z) \geq 0$ be an integer such that

$$\beta^{\hat{k}(z)} \leq t_2(z) \leq \beta^{\hat{k}(z)-1} \quad (A27)$$

In step 6 of the algorithm, the smallest integer $k(z')$ is calculated such that $s(z') = \beta^{k(z')} \in (0, \hat{M}]$ satisfies

$$\psi(z' + s(z')h_{\varepsilon(z')}^{\pi}(z')) - \psi(z') \leq -s(z')\alpha\delta\varepsilon(z') \quad (A28)$$

whenever $z' \in F^c$. Therefore, $k(z') \leq \hat{k}(z)$ and $-\beta^{k(z')} \leq -\beta^{\hat{k}(z)}$ which gives

$$\psi(z' + s(z')h_{\varepsilon(z')}^{\pi}(z')) - \psi(z') \leq -\alpha\delta\beta^{\hat{k}(z)} \frac{\varepsilon(z)}{2}$$

$$\forall z' \in B(z, \rho_3(z)) \cap F^c$$

$$\forall h_{\varepsilon(z')}^{\pi}(z') \in \tilde{C}^{\pi}(z') \quad (A29)$$

We now consider $z' \in F$, i.e. $\psi_0(z') = 0$. From the definition of $\tilde{\theta}_{\varepsilon(z')}^{\pi}(z')$ (or from the dual forms of $\tilde{\theta}_{\varepsilon}^3(\cdot)$ and $\tilde{\theta}_{\varepsilon}^4(\cdot)$ in (A1) and (A2)) it is easily seen that

$$\langle \nabla f^0(z'), h_{\varepsilon(z')}^{\pi}(z') \rangle \leq -\delta\varepsilon(z') \quad \forall z' \in B(z, \rho_3(z)) \cap F$$

$$\forall h_{\varepsilon(z')}^{\pi}(z') \in \tilde{C}^{\pi}(z') \quad (A30)$$

By combining (A21) and (A30) we get

$$\begin{aligned}
 f^0(z' + h_{\varepsilon(z')}^{\pi}(z')) - f^0(z') &\leq (1-\alpha)t\delta \frac{\varepsilon(z)}{2} - t\delta\varepsilon(z') \\
 &\leq (1-\alpha)t\delta\varepsilon(z') - t\delta\varepsilon(z') \\
 &= -\alpha t\delta\varepsilon(z')
 \end{aligned}$$

$$\begin{aligned}
 \forall z' &\in B(z, \rho_3(z)) \cap F \\
 \forall t &\in [0, t_2(z)] \\
 \forall h_{\varepsilon(z')}^{\pi}(z') &\in \tilde{C}^{\pi}(z') \quad (A31)
 \end{aligned}$$

From (A26) we have

$$\begin{aligned}
 \psi(z' + th_{\varepsilon(z')}^{\pi}(z')) &\leq 0 \quad \forall z' \in B(z, \rho_3(z)) \cap F \\
 \forall t &\in [0, t_2(z)] \\
 \forall h_{\varepsilon(z')}^{\pi}(z') &\in \tilde{C}^{\pi}(z') \quad (A32)
 \end{aligned}$$

In Step 6 of the algorithm, the smallest integer $k(z')$ is calculated such that $s(z') = \beta^{k(z')} \in (0, \hat{M}]$ satisfies

$$\begin{aligned}
 f^0(z' + s(z')h_{\varepsilon(z')}^{\pi}(z')) - f^0(z') &\leq -s(z')\alpha\delta\varepsilon(z') \\
 f^j(z' + s(z')h_{\varepsilon(z')}^{\pi}(z')) &\leq 0 \quad j = 1, 2, \dots, m \quad (A33)
 \end{aligned}$$

whenever $z' \in F$. Again, we have that $k(z') \leq \hat{k}(z)$ and $-\beta^{k(z')} \leq -\beta^{\hat{k}(z)}$.

Hence,

$$\begin{aligned}
 f^0(z' + s(z')h_{\varepsilon(z')}^{\pi}(z')) - f^0(z') &\leq -\beta^{\hat{k}(z)}\alpha\delta \frac{\varepsilon(z)}{2} \triangleq \delta(z) \\
 \forall z' &\in B(z, \rho_3(z)) \cap F \\
 \forall h_{\varepsilon(z')}^{\pi}(z') &\in \tilde{C}^{\pi}(z') \quad (A34)
 \end{aligned}$$

Let $\gamma(z) \triangleq \rho_3(z)$, then we are done.

□

Appendix B.

In order to illustrate the behavior of the combined phase I-phase II algorithms, consider the following optimization problem which is a modification of Problem No. 16 in Himmelblau [13].

$$\min f^0(z) \triangleq -0.5(z^1 z^4 - z^2 z^3 + z^3 z^5 - z^5 z^8 - z^6 z^7) \quad (B1)$$

$$\text{subject to: } f^1(z) \triangleq -1 + (z^3)^2 + (z^4)^2 \leq 0 \quad (B2)$$

$$f^2(z) \triangleq -1 + (z^5)^2 + (z^6)^2 \leq 0 \quad (B3)$$

$$f^3(z) \triangleq -1 + (z^1)^2 + (z^2 - 1)^2 \leq 0 \quad (B4)$$

$$f^4(z) \triangleq -1 + (z^1 - z^5)^2 + (z^2 - z^6)^2 \leq 0 \quad (B5)$$

$$f^5(z) \triangleq -1 + (z^1 - z^7)^2 + (z^2 - z^8)^2 \leq 0 \quad (B6)$$

$$f^6(z) \triangleq -1 + (z^3 - z^5)^2 + (z^4 - z^6)^2 \leq 0 \quad (B7)$$

$$f^7(z) \triangleq -1 + (z^3 - z^7)^2 + (z^4 - z^8)^2 \leq 0 \quad (B8)$$

$$f^8(z) \triangleq -1 + (z^7)^2 + (z^8 - 1)^2 \leq 0 \quad (B9)$$

$$f^9(z) \triangleq -z^1 z^4 + z^2 z^3 \leq 0 \quad (B10)$$

$$f^{10}(z) \triangleq -z^3 \leq 0 \quad (B11)$$

$$f^{11}(z) \triangleq z^5 \leq 0 \quad (B12)$$

$$f^{12}(z) \triangleq -z^5 z^8 + z^6 z^7 \leq 0 \quad (B13)$$

where $z \triangleq (z^1, z^2, \dots, z^8)^T \in \mathbb{R}^8$. The original problem in [13], to maximize the area of a hexagon in which the maximum diameter is unity, was modified slightly so that Assumption 3 could be satisfied.

The parameters used in the algorithm were

$$\pi=3; \delta=10^{-3}; \epsilon_0=10^{-1}; \epsilon_1=10^{-6}; \alpha=0.3; \beta=0.8; M=1.0$$

Note that with $\pi=3$, the modified Polak-Trahan optimality function was used. In order to compare the separate phase I-phase II method to the new combined method, the following phase I problem was formulated

$$\min\{z^9 | f^j(z) - z^9 \leq 0, j = 1, 2, \dots, 12\} \quad (B14)$$

The value of z_0^9 was chosen so that the initial point was feasible for (B14); i.e. $z_0^9 = \max\{f^j(z_0), j=1, 2, \dots, 12\}$. The algorithm was then applied to (B14) until z^9 became negative. The results are tabulated in Table 1a.

For the phase II mode, the algorithm was applied to the original problem, (B1)-(B13), using, as an initial point, the resulting feasible point from the phase I mode. The phase II results are tabulated in Table 1b. The last iteration shown, $i = 47$, is the number of iterations needed for each z^i , $i = 1, 2, \dots, 8$, to be a solution accurate to at least four decimal places. The algorithm was actually run for more iterations, but this resulted in no change in the first four decimal places.

The combined phase I-phase II algorithm was run with the same initial point as that for the separate phase I. The results are shown in Tables 2 and 3 for different values of γ . For $\gamma = 1.0$, the computer run was terminated after the maximum limit of 100 iterations was reached without obtaining a four decimal place accuracy solution. For $\gamma = 2.0$, an accuracy of four decimal places was reached after 43 iterations. These results clearly show that the new method works very well when γ is chosen properly. Our computational experience with this example, as well as with others, indicates that γ should be chosen strictly greater than 1 so that the algorithm concentrates more heavily on becoming feasible in the initial iterations.

Upon comparing Tables 1 and 3, it can be seen that the new combined phase I-phase II method is somewhat faster than the separate phase I-phase II method. Programming time is also saved by using the new methods since it is necessary to set up only one problem instead of two separate problems.

REFERENCES

- [1] E. Polak and D. Q. Mayne, "An Algorithm for Optimization Problems with Functional Inequality Constraints," IEEE Transactions on Automatic Control, Vol. AC-21, No. 2, April 1976, pp. 184-193.
- [2] E. Polak, K. S. Pister, and D. Ray, "Optimal Design of Framed Structures Subjected to Earthquakes," Engineering Optimization, Vol. 2, pp. 65-71, 1976.
- [3] E. Polak and R. Trahan, "An Algorithm for Computer Aided Design Problems," Proceedings of the 1976 IEEE Conference on Decision and Control, Dec. 1976, pp. 537-542.
- [4] D. Q. Mayne, E. Polak, and R. Trahan, "An Outer Approximations Algorithm for Computer Aided Design Problems," Submitted to JOTA.
- [5] F. John, "Extremum Problems with Inequalities as Side Conditions," in K. O. Friedrichs, O. W. Neugebauer, and J. J. Stoker (eds.), Studies and Essays: Courant Anniversary Volume, Interscience Publishers, Inc., New York, New York, 1948.
- [6] G. Zoutendijk, Methods of Feasible Directions, Elsevier, Amsterdam, 1970.
- [7] Donald M. Topkis and Arthur F. Veinott, Jr., "On the Convergence of Some Feasible Direction Algorithms for Nonlinear Programming," J. SIAM Control, Vol. 5, No. 2, 1967, pp. 268-279.
- [8] O. Pironneau and E. Polak, "Rate of Convergence of a Class of Methods of Feasible Directions," SIAM J. Numer. Anal. Vol. 10, No. 1, March 1973, pp. 161-173.

- [9] S. I. Zukhovitskii, R. A. Polyak, and M. E. Primak, "An Algorithm for the Solution of Convex Programming Problems," DAN USSR 153 (5), 1963, pp. 991-1000.
- [10] M. D. Canon, C. D. Cullum, Jr., and E. Polak, Theory of Optimal Control and Mathematical Programming, McGraw-Hill, New York, 1970.
- [11] E. Polak, Computational Methods in Optimization: A Unified Approach, Academic Press, New York, 1971.
- [12] O. Pironneau and E. Polak, "On the Rate of Convergence of Certain Methods of Centers," Mathematical Programming, Vol. 2, No. 2, April 1972, pp. 230-257.
- [13] D. M. Himmelblau, Applied Nonlinear Programming, McGraw Hill, New York, 1972.
- [14] D. Q. Mayne, E. Polak, "Feasible Directions Algorithms for Optimization Problems with Equality and Inequality Constraints," Mathematical Programming 11, 1976, pp. 67-80.

Table 1a

Phase 1

i	z^1	z^2	z^3	z^4	z^5	z^6	z^7	z^8	z^9
0	1.000	0.000	1.000	1.000	-1.000	1.000	-1.000	0.000	4.000
6	0.333	0.337	0.333	0.666	-0.333	0.666	-0.333	0.337	-0.110

Table 1b

Phase 2

i	z^1	z^2	z^3	z^4	z^5	z^6	z^7	z^8	f(z)
0	0.3330	0.3370	0.3330	0.6660	-0.3330	0.6660	-0.3330	0.3370	-0.4426
10	0.4934	0.3891	0.3404	0.9253	-0.3404	0.9253	-0.4934	0.3891	-0.6646
20	0.5000	0.4032	0.3443	0.9386	-0.3443	0.9386	-0.5000	0.4032	-0.6747
30	0.5000	0.4022	0.3436	0.9391	-0.3436	0.9391	-0.5000	0.4032	-0.6750
40	0.5000	0.4024	0.3438	0.9390	-0.3438	0.9390	-0.5000	0.4024	-0.6750
43	0.5000	0.4023	0.3438	0.9391	-0.3438	0.9391	-0.5000	0.4023	-0.6750
47	0.5000	0.4024	0.3438	0.9391	-0.3438	0.9391	-0.5000	0.4024	-0.6750

Table 2

Combined Phase I-Phase II; $\gamma = 1.0$

i	z^1	z^2	z^3	z^4	z^5	z^6	z^7	z^8	f(z)	$\psi_0(z)$
0	1.0000	0.0000	1.0000	1.0000	-1.0000	1.0000	-1.0000	0.0000	-2.0000	4.0000
20	0.4426	0.2663	0.4426	0.7337	-0.4426	0.7337	-0.4426	0.2663	-0.6497	0.0022
40	0.4331	0.2501	0.4331	0.7499	-0.4331	0.7499	-0.4331	0.2501	-0.6495	8.5×10^{-8}
47	0.4329	0.2499	0.4331	0.7500	-0.4331	0.7500	-0.4329	0.2499	-0.6495	0.0000
60	0.3936	0.0989	0.5170	0.6335	-0.4685	0.5916	-0.3248	0.1069	-0.6629	0.0000
80	0.3565	0.0669	0.5024	0.5977	-0.4958	0.5875	-0.3390	0.0609	-0.6733	0.0000
100	0.3448	0.0615	0.5003	0.5978	-0.4996	0.5969	-0.3432	0.0608	-0.6749	0.0000

Table 3

Combined Phase I-Phase II; $\gamma = 2.0$

i	z^1	z^2	z^3	z^4	z^5	z^6	z^7	z^8	$f(z)$	$\psi_0(z)$
0	1.0000	0.0000	1.0000	1.0000	-1.0000	1.0000	-1.0000	0.0000	-2.0000	4.0000
1	0.0385	0.1202	0.0385	0.8799	-0.0385	0.8799	-0.0385	0.1202	-0.6768	0.0000
10	0.4759	0.3943	0.3855	0.9021	-0.3855	0.9021	-0.4759	0.3943	-0.6628	0.0000
20	0.4998	0.4028	0.3443	0.9388	-0.3443	0.9388	-0.4998	0.4028	-0.6748	0.0000
30	0.5000	0.4023	0.3437	0.9391	-0.3437	0.9391	-0.5000	0.4023	-0.6750	0.0000
40	0.5000	0.4023	0.3438	0.9391	-0.3438	0.9391	-0.5000	0.4023	-0.6750	0.0000
43	0.5000	0.4024	0.3438	0.9391	-0.3438	0.9391	-0.5000	0.4024	-0.6750	0.0000