AN OUTER APPROXIMATIONS ALGORITHM FOR COMPUTER AIDED DESIGN PROBLEMS

by

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

Memorandum No. UCB/ERL M77-10

8 February 1977

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

AN OUTER APPROXIMATIONS ALGORITHM FOR COMPUTER AIDED DESIGN PROBLEMS[1]

D. Q. Mayne[2], E. Polak[3] and R. Trahan[4]

ABSTRACT

This paper presents an implementable algorithm of the outer approximations type for solving nonlinear programming problems with functional inequality constraints. The algorithm was motivated by engineering design problems in circuit tolerancing, multivariable control and shock resistant structures.

---

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

[2] Professor, Department of Computing and Control, Imperial College of Science and Technology, London, England.

[3] Professor, Department of Electrical Engineering and Computer Sciences, and the Electronics Research Laboratory, University of California, Berkeley, California 94720.

[4] Research Assistant, Department of Electrical Engineering and Computer Sciences, and the Electronics Research Laboratory, University of California, Berkeley, California 94720.

1. Introduction

Quite commonly, engineering design problems transcribe into mathematical programming problems of the form

$$\underline{P} : \min\{f(z) | g^j(z) \leq 0,\ j = 1,2,\ldots,p;\ \phi^\ell(z,\omega^\ell) \leq 0,$$

$$\omega^\ell \in \Omega^\ell,\ \ell = 1,2,\ldots,q\} \tag{1}$$

where $z \in \mathbb{R}^n$, $\omega^\ell \in \mathbb{R}^{\nu_\ell}$; f, g^j, are all continuously differentiable, on $\mathbb{R}^n$, and the ϕ^ℓ and $\nabla_z\phi^\ell$ are continuous on $\mathbb{R}^n \times \mathbb{R}^{\nu_\ell}$. The sets $\Omega^\ell \subset \mathbb{R}^{\nu_\ell}$ are assumed to be compact. In particular, circuit design, with tolerances, problems (see e.g. [Refs. 1,2,3]), multivariable control system design problems (see e.g. [Refs. 5,4,6]) and shock resistant structure design problems (see e.g. [Refs. 7,8]) have been cast in this form.

From a mathematical programming point of view, $\underline{P}$ is a particularly difficult problem, since a constraint of the form $\phi^\ell(z,\omega) \leq 0$ $\forall\omega \in \Omega^\ell$ is, in fact, an infinite number of constraints. Because of this, even to check whether a point z is feasible may require an infinite number of function evaluations, and, consequently, any conceptual algorithm for solving the general problem $\underline{P}$ is bound to be multiply infinite, i.e., it constructs an infinite sequence $\{z_i\}_{i=0}^{\infty}$, each point of which is constructed by means of an infinite number of operations. In the nonconvex case, there are basically two conceptual algorithms which can serve as prototypes for an implementable algorithm. The first is in the class of feasible directions and is due to Demyanov [Ref. 9] and the second one is in the class of outer approximations and is due to Levitin and Polyak [Ref. 10], Eaves and Zangwill [Ref. 11], and Blankenship and Falk [Ref. 12]. The set of implementable (i.e. practical) algorithms is equally scarce:

in [Ref. 6] the authors made use of Demyanov's ideas for the special case where $\nu = 1$ (i.e. all $\Omega^{\ell} \subset \mathbb{R}^1$), while in [Ref. 12] Blankenship and Falk present a slightly incomplete algorithm for the convex case. There seem to be no implementable algorithms, in the literature, for the general case.

In this paper we present a new, implementable algorithm of the outer approximations type for the general case of problem $\underline{P}$. It is an implementation of a conceptual scheme proposed by Eaves and Zangwill in [11]. We shall now explain the conceptual scheme and the obstacles which we had to overcome in constructing an implementation. For this purpose, consider the simplified form of $\underline{P}$: $\min\{f(z) \mid \phi(z,\omega) \leq 0,\ \omega \in \Omega\}$ (i.e. one constraint only) and let $M(z)$ denote the problem $\max_{\omega \in \Omega} \phi(z,\omega)$. The simplest conceptual algorithm begins with a $\omega_0 \in \Omega$ and solves $\underline{P}_0$: $\min\{f(z) \mid \phi(z,\omega_k) \leq 0,\ k = 0\}$ to obtain z_0. It then computes ω_1 by solving $M(z_0)$ and if $\phi(z_0,\omega_1) \leq 0$, it stops, otherwise it proceeds to solve $\underline{P}_1$: $\min\{f(z) \mid \phi(z,\omega_k) \leq 0,\ k = 0,1\}$ for a z_1, etc. Let $\underline{P}_i$: $\min\{f(z) \mid \phi(z,\omega_k) \leq 0, k = 1,2,\ldots,i\}$. The problems $\underline{P}_i$ have a finite number (i) of constraints only. Their feasible sets $F_i = \{z \mid \phi(z,\omega_k) \leq 0,\ k = 0,1,2,\ldots,i\}$ satisfy $F_0 \supset F_1 \supset, \ldots, \supset F \triangleq \{z \mid \phi(z,\omega) \leq 0\ \forall \omega \in \Omega\}$. Hence $f(z_0) \leq f(z_1) \leq \cdots \leq f(z_i) \leq \cdots \leq f(z^*)$, where z^* is any solution of $\underline{P}$. The theory in [Ref. 11] leads to the conclusion that any accumulation point $\hat{z}$ of $\{z_i\}_{i=0}^{\infty}$ is in F and solves $\underline{P}$. Furthermore, it is shown in [Ref. 11] that, in constructing $\underline{P}_i$, a number of the constraints on z, $\phi(z,\omega_k) \leq 0$, can be dropped from P_i if (i) $\phi(z_{i-1},\omega_k) \leq 0$ and $f(z_{i-i})$ is sufficiently larger than $f(z_k)$, and (ii) the next solution z_i satisfies $f(z_i) \geq f(z_{i-1})$. Thus, the growth of complexity of $\underline{P}_i$ can be slowed and, possibly, even arrested. To obtain an implementation we needed to invent two nontrivial modifications. (i) Since the problems $\underline{P}_i$ and $M(z)$ generally take an infinite number of iterations to solve, we had to invent an efficient approximation scheme which is compatible with the convergence of the overall algorithm. (ii) Since in the absence of convexity, $\underline{P}_i$ can only be solved in the sense that one computes a feasible <u>stationary</u> point z_i, we could not assume that the sequence $\{f(z_i)\}$ was monotonically increasing. Hence we had to invent a constraint dropping scheme which, unlike the one of Eaves and Zangwill, does not depend on monotonicity. The result is an algorithm which can be (and was) coded without further interpretations. Our computational experiments show that it works rather well and, in particular, that our constraint dropping scheme works very well in the sense that the number of constraints in $\underline{P}_i$ remains small.

2. <u>The Algorithm</u>

As already stated in the introduction, we need the following hypotheses.

<u>Assumption 2.1</u>: The functions f, g^j, $j = 1,2,\ldots,p$, from $\mathbb{R}^n$ into $\mathbb{R}^1$ are continuously differentiable. The functions $\phi^\ell : \mathbb{R}^n \times \mathbb{R}^{\nu_\ell} \to \mathbb{R}^1$ and $\nabla_z\phi^\ell : \mathbb{R}^n \times \mathbb{R}^{\nu_\ell} \to \mathbb{R}^n$, $\ell = 1,2,\ldots,q$, are continuous. $\square$

<u>Assumption 2.2</u>: The sets $\Omega^\ell \subset \mathbb{R}^{\nu_\ell}$, $\ell = 1,2,\ldots,q$, are compact. $\square$

The specific subalgorithms called by the Master Algorithm to be presented usually require additional assumptions. We shall introduce these for one subalgorithm in Appendix A.

We now develop the three optimality functions $(\theta_\Omega, \theta_{\Omega'}, \tilde{\theta}_\ell)$ which will be used in the algorithm and the analysis. These functions assume zero value at optimal points.

Let $\Omega \triangleq \{\Omega^1,\Omega^2,\ldots,\Omega^q\}$ and let $\Omega' \triangleq \{\Omega'^1,\Omega'^2,\ldots,\Omega'^q\}$, with $\Omega'^\ell \subset \Omega^\ell$, for $\ell = 1,2,\ldots,q$. We now define

$$\tilde{\psi}_{\Omega'}(z) \triangleq \max\{g^j(z),\ j = 1,2,\ldots,p;\ \phi^\ell(z,\omega^\ell),$$
$$\omega^\ell \in \Omega'^\ell,\ \ell = 1,2,\ldots,q\} \tag{2}$$

and

$$\psi_{\Omega'}(z) \triangleq \max\{0,\tilde{\psi}_{\Omega'}(z)\} \tag{3}$$

For reasons of computational efficiency in the implementation, we shall use the optimality function[5] $\theta_\Omega : \mathbb{R}^n \to \mathbb{R}^1$ defined by

$$\theta_\Omega(z) \triangleq \min_h\{\tfrac{1}{2}\|h\|^2 + \max\{\langle \nabla f(z),h\rangle - \psi_\Omega(z);$$
$$g^j(z) - \psi_\Omega(z) + \langle \nabla g^j(z),h\rangle,\ j = 1,2,\ldots,p;\ \phi^\ell(z,\omega^\ell) - \psi_\Omega(z)$$
$$+ \langle \nabla_z \phi^\ell(z,\omega^\ell),h\rangle,\ \omega^\ell \in \Omega^\ell,\ \ell = 1,2,\ldots,q\}\} \tag{4}$$

with $\Omega = \{\Omega^1,\Omega^2,\ldots,\Omega^q\}$. Let

$$F_\Omega \triangleq \{z \mid g^j(z) \leq 0,\ j = 1,2,\ldots,p;\ \phi^\ell(z,\omega^\ell) \leq 0,$$
$$\omega^\ell \in \Omega^\ell,\ \ell = 1,2,\ldots,q\} \tag{5}$$

i.e. F_Ω is the feasible set for P.

---

[5] It can be seen from section (4.4) in [Ref. 13] that the various optimality functions used in defining various methods of feasible directions are equivalent, in the sense that they all have the same zeros. Also, they are zero at a feasible point if and only if it is an F. John point [Ref. 14]. Many of these optimality functions are interchangeable in the Master Algorithm.

We can then write problem $\underline{P}$ in the equivalent form

$$\underline{P}_\Omega : \min\{f(z) \mid z \in F_\Omega\} \tag{6}$$

It is an easy extension of the results in [Ref. 6] to show (see Proposition 2.1 further on) that if $\hat{z} \in F_\Omega$ is optimal for $\underline{P}$, then $\theta_\Omega(z) = 0$. Given (finite) discrete sets $\Omega'^\ell = \{\omega_0^\ell, \omega_1^\ell, \ldots, \omega_{k_\ell}^\ell\} \subset \Omega^\ell$, $\ell = 1,2,\ldots,q$, we define

$$\begin{aligned} F_{\Omega'} \triangleq \{z \mid g^j(z) \le 0,\ j = 1,2,\ldots,p;\ \phi^\ell(z,\omega^\ell) \le 0, \\ \omega^\ell \in \Omega'^\ell,\ \ell = 1,2,\ldots,q\} \end{aligned} \tag{7}$$

$$\underline{P}_{\Omega'} : \min\{f(z) \mid z \in F_{\Omega'}\} \tag{8}$$

with Ω' as defined before. For the problem $\underline{P}_{\Omega'}$, we define $\theta_{\Omega'} : \mathbb{R}^n \to \mathbb{R}^1$ by

$$\begin{aligned} \theta_{\Omega'}(z) &\triangleq \min_h \{\tfrac{1}{2}\|h\|^2 + \max\{\langle \nabla f(z), h\rangle - \psi_{\Omega'}(z);\ g^j(z) - \psi_{\Omega'}(z) \\ &\quad + \langle \nabla g^j(z), h\rangle,\ j = 1,2,\ldots,p;\ \phi^\ell(z,\omega^\ell) - \psi_{\Omega'}(z) \\ &\quad + \langle \nabla_z \phi^\ell(z,\omega^\ell), h\rangle,\ \omega^\ell \in \Omega'^\ell,\ \ell = 1,2,\ldots,q\}\} \\ &= \max_{\mu \ge 0} \{-\mu_f \psi_{\Omega'}(z) + \sum_{j=1}^{p} \mu_g^j (g^j(z) - \psi_{\Omega'}(z)) \\ &\quad + \sum_{\ell=1}^{q} \sum_{j=1}^{k_\ell} \mu_\phi^{\ell,j} (\phi^\ell(z,\omega_j^\ell) - \psi_{\Omega'}(z)) - \tfrac{1}{2}\|\mu_f \nabla f(z) \\ &\quad + \sum_{j=1}^{p} \mu_g^j \nabla g^j(z) + \sum_{\ell=1}^{q} \sum_{j=1}^{k_\ell} \mu_\phi^{\ell,j} \nabla_z \phi^\ell(z,\omega_j^\ell)\|^2 \Big| \\ &\quad \mu_f + \sum_{j=1}^{p} \mu_g^j + \sum_{\ell=1}^{q} \sum_{j=1}^{k_\ell} \mu_\phi^{\ell,j} = 1\} \end{aligned} \tag{9}$$

The alternative expression follows from duality (see [Ref. 15]) and hence it is easy to see that if $\hat{z}' \in F_{\Omega'}$ is optimal for $\underline{P}_{\Omega'}$, then $\theta_{\Omega'}(\hat{z}') = 0$ and that this fact is equivalent to satisfying the F. John condition [Ref. 14]. Thus for the original problem, we see that $\theta_\Omega(z) = 0$ is a generalization of the F. John condition to the case of an infinite number of constraints. This fact will be proved in Proposition 2.1.

In our convergence analysis we shall make use of the auxiliary function $\tilde{\theta}_{\Omega'} : \mathbb{R}^n \to \mathbb{R}^1$ defined by

$$\tilde{\theta}_{\Omega'}(z) \triangleq \min\{\tfrac{1}{2}\|h\|^2 + \max\{\langle \nabla f(z),h\rangle - \psi_{\Omega'}(z);$$

$$g^j(z) - \psi_{\Omega'}(z) + \langle \nabla g^j(z),h\rangle,\ j = 1,2,\ldots,p;\ \phi^\ell(z,\omega^\ell)$$

$$- \psi_{\Omega'}(z) + \langle \nabla_z\phi^\ell(z,\omega^\ell),h\rangle,\ \omega^\ell \in \Omega^\ell,\ \ell = 1,2,\ldots,q\}\} \qquad (10)$$

with $\Omega' \subset \Omega$. Note that $\tilde{\theta}_{\Omega'}$ differs from $\theta_{\Omega'}$ because in (10) $\omega^\ell \subset \Omega^\ell$, while in (9) $\omega^\ell \in \Omega'^\ell \subset \Omega^\ell$. The important properties of θ_Ω, $\theta_{\Omega'}$ and $\tilde{\theta}_{\Omega'}$ are as follows:

<u>Proposition 2.1</u>: If Assumptions 2.1 and 2.2 are satisfied then: (i) For all $z \in \mathbb{R}^n$ and any discrete subset $\Omega' \subset \Omega$, $\theta_\Omega(z) \leq 0$, $\theta_{\Omega'}(z) \leq 0$, and furthermore, $\theta_{\Omega'}(z) \leq \tilde{\theta}_{\Omega'}(z)$.

(ii) If $\hat{z} \in F_\Omega$ is optimal for $\underline{P}_\Omega$ then $\theta_\Omega(\hat{z}) = 0$. If $\hat{z}' \in F_{\Omega'}$ is optimal for $\underline{P}_{\Omega'}$ then $\theta_{\Omega'}(\hat{z})' = 0$.

(iii) For any $\Omega' \subset \Omega$, θ_Ω, $\tilde{\theta}_{\Omega'}$ and $\theta_{\Omega'}$ are continuous on $\mathbb{R}^n$.

<u>Proof</u>: (i) Since $h = 0$ is allowed in the right hand sides of (4) and (9) and for any $\Omega' \subset \Omega$, $g^j(z) - \psi_{\Omega'}(z) \leq 0$, $j = 1,2,\ldots,p$, $\phi^\ell(z,\omega^\ell) - \psi_{\Omega'}(z) \leq 0$, $\omega^\ell \in \Omega'^\ell$, $\ell = 1,2,\ldots,q$, it follows that $\theta_\Omega(z) \leq 0$ and $\theta_{\Omega'}(z) \leq 0$ for all $z \in \mathbb{R}^n$. The fact that $\theta_{\Omega'}(z) \leq \tilde{\theta}_{\Omega'}(z)$ for all $z \in \mathbb{R}^n$ follows from the fact that $\Omega'^\ell \subset \Omega^\ell$, $\ell = 1,2,\ldots,q$. i.e. because the max in (10) is over a larger set than the max in the first expression in (9).

(ii) The fact that if $\hat{z}'$ is optimal for $\underline{P}_{\Omega'}$, then $\theta_{\Omega'}(\hat{z}') = 0$ follows directly from the second expression in (9), which shows that $\theta_{\Omega'}(\hat{z}') = 0$

is equivalent to the F. John condition being satisfied. To show that if $\hat{z} \in F_\Omega$ is optimal for $\underline{P}_\Omega$, then $\theta_\Omega(\hat{z}) = 0$, we proceed as follows (c.f. proof of Theorem (4.2.32) in [Ref. 13]). Since $\theta_\Omega(\hat{z}) \leq 0$ we assume, for the sake of contradiction that $\theta_\Omega(\hat{z}) < 0$. But this implies that for the $\hat{h}$ solving (4),

$$\langle \nabla f(\hat{z}), \hat{h} \rangle \leq \theta_\Omega(\hat{z}) - \frac{1}{2} \|\hat{h}\|^2 < 0 \tag{11}$$

$$g^j(\hat{z}) + \langle \nabla g^j(\hat{z}), \hat{h} \rangle \leq \theta_\Omega(\hat{z}) - \frac{1}{2} \|\hat{h}\|^2 < 0, \; j = 1,2,\dots,p \tag{12}$$

$$\phi^\ell(\hat{z}) + \langle \nabla_z \phi^\ell(\hat{z},\omega), \hat{h} \rangle \leq \theta_\Omega(\hat{z}) - \frac{1}{2} \|\hat{h}\|^2 < 0, \; \omega \in \Omega^\ell, \; \ell = 1,2,\dots,q \tag{13}$$

Since the sets Ω^ℓ are compact by assumption it can be shown (as in the proof of theorem (4.2.32) in [Ref. 13]), that there exists a $\lambda > 0$ such that $\hat{z} + \lambda\hat{h} \in F_\Omega$ and $f(\hat{z} + \lambda\hat{h}) < f(\hat{z})$, contradicting the optimality of $\hat{z}$. Thus, $\theta_\Omega(\hat{z}) = 0$, is a necessary condition of optimality for P_Ω.

(iii) To show that $\theta_\Omega(\cdot)$ is continuous on $\mathbb{R}^n$, consider any sequence $z_i \to z^* \in \mathbb{R}^n$. Let h_i be a minimizer associated with z_i in (4), and h^* with z^*.[6] Let $s : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^1$ be defined by

$$\begin{aligned} s(z,h) \triangleq \frac{1}{2} \|h\|^2 + \max\{&\langle \nabla f(z), h \rangle - \psi_\Omega(z); \\ &g^j(z) - \psi_\Omega(z) + \langle \nabla g^j(z), h \rangle, \; j = 1,2,\dots,p; \\ &\phi^\ell(z,\omega^\ell) - \psi_\Omega(z) + \langle \nabla_z \phi^\ell(z,\omega^\ell), h \rangle, \; \omega^\ell \in \Omega^\ell, \; \ell = 1,2,\dots,q\} \end{aligned} \tag{14}$$

[6]The minimizers h_i for (4) exist for all z_i, $i = 1,2,\dots$, since $s(z,h)$ is bounded from below and $s(z,h) \to \infty$ as $\|h\| \to \infty$ and $s(z,h)$ is continuous in h.

Then $\theta_\Omega(z^*) = s(z^*,h^*)$. Because of assumptions 2.1 and 2.2 and because $\psi_\Omega(\cdot)$ is continuous, $s(\cdot,\cdot)$ is continuous and therefore, for any sequence $z_i \to z^* \in \mathbb{R}^n$, as $i \to \infty$, $s(z_i,h^*) \to \theta_\Omega(z^*)$ as $i \to \infty$. But $\theta_\Omega(z_i) \leq s(z_i,h^*)$ by definition of $\theta_\Omega(z_i)$, and hence $\overline{\lim}\ \theta_\Omega(z_i) \leq \lim s(z_i,h^*) = \theta_\Omega(z^*)$, which shows that $\theta_\Omega(\cdot)$ is upper semicontinuous (u.s.c.).

To show that θ_Ω is also lower semicontinuous (l.s.c.) we first show that the sequence $\{h_i\}$ is bounded. Suppose this is not true, then there exists a subsequence $\{h_{i_k}\}$ such that $\|h_{i_k}\| \to \infty$ as $i_k \to \infty$. Now

$$\begin{aligned}\theta_\Omega(z_{i_k}) = s(z_{i_k},h_{i_k}) \geq \tfrac{1}{2}\|h_{i_k}\|^2 + \max\{&-\|\nabla f(z_{i_k})\|\ \|h_{i_k}\| - \psi_\Omega(z_{i_k});\\ &g^j(z_{i_k}) - \psi_\Omega(z_{i_k}) - \|\nabla g^j(z_{i_k})\|\ \|h_{i_k}\|,\ j = 1,2,\ldots,p;\\ &\phi^\ell(z_{i_k},\omega^\ell) - \psi_\Omega(z_{i_k}) - \|\nabla_z\phi(z_{i_k},\omega^\ell)\|\ \|h_{i_k}\|,\ \omega^\ell \in \Omega^\ell,\ \ell = 1,2,\ldots,q\}\end{aligned}$$

and hence $\theta_\Omega(z_{i_k}) \to \infty$ as $i_k \to \infty$. But, by continuity, $s(z_i,0)$ is bounded for all $i = 1,2,\ldots$, and by definition, $\theta_\Omega(z_i) \leq s(z_i,0)$ so that $\{\theta_\Omega(z_i)\}$ is a bounded sequence. This is a contradiction and therefore $\{h_i\}$ is bounded and hence $\{h_i\}$ (and any infinite subsequence of $\{h_i\}$) must have accumulation points. If we consider any accumulation point $\hat{h}$ of $\{h_i\}$ i.e., $h_{i_k} \to \hat{h}$ as $i_k \to \infty$ then $s(z_{i_k}, h_{i_k}) \to s(z^*,\hat{h}) \geq \theta_\Omega(z^*)$. Hence, for some subsequence $(z_{i_k}, h_{i_k}) \to (z^*,\hat{h})$ as $i_k \to \infty$, $\underline{\lim}\ \theta_\Omega(z_i)$ $= \lim s(z_{i_k}, h_{i_k}) = s(z^*,\hat{h}) \geq \theta_\Omega(z^*)$, and this proves that θ_Ω is l.s.c. Since $\theta_\Omega(\cdot)$ is both u.s.c. and l.s.c., it is continuous. The proof that $\tilde{\theta}_{\Omega'}$ and $\theta_{\Omega'}$ are continuous is identical to the one just given for θ_Ω. $\square$

We note at this point that, for Ω_i a discrete subset of Ω, an algorithm such as the one in Appendix A, will construct a $z_i \in \mathbb{R}^n$ satisfying $-\beta^i \leq \theta_{\Omega_i}(z_i) \leq 0$ and $0 \leq \psi_{\Omega_i}(z_i) \leq \beta^i$ in a finite number of iterations. (Note that z_i need not be in F_{Ω_i}). Because of this we shall find the following consequence of Proposition 2.1 most useful.

Corollary 2.1: Let $\beta \in (0,1)$, let $\Omega_i = \{\Omega_1^1, \Omega_i^2, \ldots, \Omega_i^q\}$ $i = 1,2,\ldots,$ with $\Omega_i^\ell \subset \Omega^\ell$, $\ell = 1,2,\ldots,q$, be an infinite sequence of discrete sets. If $z_i \in \mathbb{R}^n$, $i = 1,2,\ldots$ satisfies $-\beta^i \leq \theta_{\Omega_i}(z_i) \leq 0$, $0 \leq \psi_{\Omega_i}(z_i) \leq \beta^i$ and $z_i \to z^* \in F_\Omega$ then $\theta_\Omega(z^*) = 0$.

Proof: We begin by defining the function $\sigma : \mathbb{R}^n \times \mathbb{R}^1 \to \mathbb{R}^1$ by

$$\begin{aligned}\sigma(z,\psi) \triangleq \min_h \{\tfrac{1}{2} \|h\|^2 &+ \max\{\langle \nabla f(z),h\rangle - \psi; \\ &g^j(z) - \psi + \langle \nabla g^j(z),h\rangle,\ j = 1,2,\ldots,p;\ \phi^\ell(z,\omega^\ell) \\ &- \psi + \langle \nabla_z \phi^\ell(z,\omega^\ell),h\rangle,\ \omega^\ell \in \Omega^\ell,\ \ell = 1,2,\ldots,q\}\} \qquad (15)\end{aligned}$$

$\sigma(\cdot,\cdot)$ is continuous by the same arguments used in the proof of Proposition 2.1. Letting $\psi_i \triangleq \psi_{\Omega_i}(z_i)$, $i = 1,2,\ldots,$ we have $\tilde{\theta}_{\Omega_i}(z_i) = \sigma(z_i,\psi_i)$. Because $z_i \to z^* \in F_\Omega$, and $\psi_i \to 0$ as $i \to \infty$, it follows by continuity that $\sigma(z_i,\psi_i) \to \sigma(z^*,0) = \theta_\Omega(z^*)$, where the equality follows from the definitions of σ, θ_Ω, and the fact that $\psi_\Omega(z^*) = 0$. By Proposition 2.1. $\tilde{\theta}_{\Omega_i}(z_i) \geq \theta_{\Omega_i}(z_i)$ for all $i = 1,2,\ldots,$ and $\theta_\Omega(z^*) \leq 0$. Thus, by continuity of $\tilde{\theta}_{\Omega_i}$ and θ_{Ω_i} we get $0 \geq \theta_\Omega(z^*) = \lim \sigma(z_i,\psi_i) = \lim \tilde{\theta}_{\Omega_i}(z_i)$ $\geq \lim \theta_{\Omega_i}(z_i) = 0.$ ¤

The algorithm which we shall shortly describe constructs sets Ω_i and points z_i which satisfy the conditions of the above corollary. However, before we state it, we must postulate a process for solving

$$\max_{\underline{\omega}}\{\underline{\phi}(z,\underline{\omega})\,|\,\underline{\omega} = (\omega^1,\omega^2,\ldots,\omega^q),\ \omega^\ell \in \Omega^\ell,\ \ell = 1,2,\ldots,q\} \tag{16}$$

where

$$\underline{\phi}(z,\underline{\omega}) \triangleq \max_{\ell \in \underline{q}} \phi^\ell(z,\omega^\ell) \tag{17}$$

with $\underline{q} \triangleq \{1,2,\ldots,q\}$. Since the sets Ω^ℓ are compact, we may postulate that we have an algorithm for solving (16) with the following properties which we state as an assumption.

Assumption 2.3: The algorithm for solving (16) is such that: (i) it accepts an initial point $\underline{\omega}'$ and after k iterations it constructs a point $\underline{s}(k,z,\underline{\omega}') \in \Omega = \{\Omega^1,\Omega^2,\ldots,\Omega^q\}$.

(ii) Given any compact set $Z \in \mathbb{R}^n$, the sequence $\underline{s}(k,z,\omega') \to \hat{\underline{s}}(z)$ as $k \to \infty$, a solution of (16), uniformly in z, for $z \in Z$. ¤

Generally speaking, we can always find such an algorithm for solving (16): in the worst situation we would have to use a Monte Carlo method. Next, in order to give ourselves some flexibility in the specification of the number of iterations to be performed on (16), we shall use positive integer valued, monotonically increasing "truncation" functions $t : \mathbb{Z}^+ \to \mathbb{Z}^+$, where $\mathbb{Z}^+ \triangleq \{0,1,2,\ldots\}$, with the property that $t(i) \to \infty$ as $i \to \infty$. We now have all the elements we need to state our Master Algorithm. The operations in Steps 1 and 2 call for the use of subroutines which will be discussed in the Appendices.

Master Algorithm

Data: $\tau > 0$, $\beta \in (0,1)$, $\gamma > 0$, $\mu_1 > 0$, $\mu_2 > 0$, a truncation function t.

Step 0: set $i = 0$, $k = 0$, $f_0 = -\infty$, $\delta_0 = 0$, $\Omega_0 = \{\phi,\phi,\ldots,\phi\}$. (i.e., $\Omega_0^\ell = \phi$, $\ell = 1,2,\ldots,q$. Here ϕ denotes the empty set.)

Step 1: Compute a z_i, by subroutine which solves P_{Ω_i}, (using z_{i-1} as an initial point)[7] such that $\theta_{\Omega_i}(z_i) \geq -\mu_1\beta^i$ and $\psi_{\Omega_i}(z_i) \leq \mu_2\beta^i$.

Step 2: Compute $\underline{\omega}_i = \underline{s}(t(i), z_i, \underline{\omega}_{i-1})$ by $t(i)$ iterations of subroutine for solving max $\{\underline{\phi}(z_i, \underline{\omega}) | \underline{\omega} \in \Omega\}$.

Step 3: If $\underline{\phi}(z_i, \underline{\omega}_i) \leq 0$, set $\Omega_{i+1} = \Omega_i$, set $i = i+1$ and go to step 1. Else, set $\hat{\Omega}_i^\ell = \hat{\Omega}_i^\ell$ for all ℓ such that $\phi^\ell(z_i, \omega_i^\ell) \leq 0$ and $\hat{\Omega}_i^\ell = \Omega_i^\ell \cup \{\omega_i^\ell\}$ for all ℓ such that $\phi^\ell(z_i, \omega_i^\ell) > 0$, and go to step 4.

Step 4: If

$$f(z_i) \geq f_k + \tau(1 - \beta^k)\delta_k - \gamma\beta^k \tag{18}$$

set $x_{k+1} = z_i$, set $f_{k+1} = f(z_i)$, set $\delta_{k+1} = \underline{\phi}(z_i, \underline{\omega}_i)$, set $k = k+1$, set $\Omega_{i+1}^\ell = \hat{\Omega}_i^\ell - \tilde{\Omega}_i^\ell$, where $\tilde{\Omega}_i^\ell = \{\omega \in \hat{\Omega}_i^\ell \,|\, \phi^\ell(z_i, \omega) < 0\}$, $\ell = 1,2,\ldots,q$, set $i = i+1$ and go to step 1.

If (18) is not satisfied, set $\Omega_{i+1}^\ell = \hat{\Omega}_i^\ell$, $\ell = 1,2,\ldots,q$ set $i = i+1$ and go to step 1.

□

We now turn to the analysis of the Master Algorithm.

Proposition 2.2: Suppose the sequence $\{x_k\}$ constructed by the Master Algorithm is infinite. Then, either the sequence $\{f(x_k)\}$ has no accumulation points (and hence is unbounded) or else, it converges.

Proof: Suppose $\{x_k\}$ is infinite and $\{f(x_k)\}$ has an accumulation point, f^*. Since $f_k = f(x_k)$ for $k = 1,2,\ldots$, we get from (18) that

$$f(x_{k+1}) \geq f(x_k) - \gamma\beta^k, \quad k = 1,2,\ldots, \tag{19}$$

For the purposes of contradiction, suppose now that $\{f(x_k)\}$ is unbounded.

[7] For $i = 0$, set z_{-1} and $\underline{\omega}_{-1}$ arbitrary.

Since because of (19) $f(x_k) \geq f(x_1) - \gamma\beta/1-\beta$ for $k = 1,2,\ldots,$ the fact that $\{f(x_k)\}$ is unbounded implies that there is an infinite subsequence, with indices in $K \subset \{1,2,\ldots\}$ such that $f(x_k) \overset{K}{\to} \infty$ as $k \to \infty$. Therefore, there is a $k_0 \in K$ such that $f(x_{k_0}) \geq f^* + 2\gamma\beta/(1-\beta)$. But from (19),

$$f(x_k) \geq f(x_{k_0}) - \gamma\beta/(1-\beta) \geq f^* + \gamma\beta/1-\beta \tag{20}$$

for all $k \geq k_0$, and hence f^* cannot be an accumulation point of $\{f(x_k)\}$. Thus, if $\{f(x_k)\}$ has an accumulation point, then $\{f(x_k)\}$ must be bounded. A similar argument shows that f^* must then be the unique accumulation point. Thus, either $\{f(x_k)\}$ converges, or it has no accumulation points.

¤

<u>Lemma 2.1</u>: Suppose the Master Algorithm constructs an infinite sequence $\{x_k\}_{k=1}^{\infty}$. Then every accumulation point x^* is feasible, i.e. $x^* \in F_\Omega$, and satisfies $\theta_\Omega(x^*) = 0$.

<u>Proof</u>: Suppose that $\{x_k\}$ has an accumulation point x^*. Then, by continuity, $\{f(x_k)\}$ must have an accumulation point and therefore, by Proposition 2, $f(x_k) \to f(x^*) \triangleq f^*$ as $k \to \infty$. Since by construction $f_k = f(x_k)$ for $k = 1,2,\ldots,$ it follows from (18) that

$$f(x_{k+1}) \geq f(x_k) + \tau(1-\beta^k)\delta_k - \gamma\beta^k \tag{21}$$

and hence, since $\beta^k \to 0$ as $k \to 0$ and $f(x_k) \to f^*$,

$$\overline{\lim}\ \delta_k \leq 0 \tag{22}$$

But, by Assumption 2.3, and the construction in the Master Algorithm

$$[\max_{\underline{\omega} \in \Omega} \underline{\phi}(x_k,\underline{\omega}) - \delta_k] \to 0 \text{ as } k \to \infty \tag{23}$$

and hence, since $\max_{\omega \in \Omega} \underline{\phi}(x,\underline{\omega})$ is continuous in x, we have, because of (22) and (23),

$$\max_{\underline{\omega} \in \Omega} \underline{\phi}(x^*,\underline{\omega}) \leq \overline{\lim} \max_{\underline{\omega} \in \Omega} \underline{\phi}(x_k,\underline{\omega}) \leq 0 \tag{24}$$

Next, by construction, each x_k satisfies $g^j(x_k) \leq \mu_2\beta^k$ for $j = 1,2,\ldots,p$, and therefore, by continuity, $g^j(x^*) \leq 0$ for $j = 1,2,\ldots,p$. Thus we have, $g^j(x^*) \leq 0$ for $j = 1,2,\ldots,p$, and $\max_{\underline{\omega} \in \Omega} \underline{\phi}(x^*,\underline{\omega}) = \max_{\ell \in q} \max_{\omega \in \Omega^\ell} \phi^\ell(x^*,\omega^\ell) \leq 0$, i.e., $x^* \in F_\Omega$.

Next, we define i(k) as the index i at which $x_k = z_i$ was constructed. Since $i(k) \geq k$ for all $k = 1,2,\ldots$, the test in step 1 of the Master Algorithm gives $-\mu_1\beta^k \leq -\mu_1\beta^{i(k)} \leq \theta_{\Omega_{i(k)}}(x_k) \leq 0$ and $0 \leq \psi_{\Omega_{i(k)}}(x_k) \leq \mu_2\beta^{i(k)} \leq \mu_2\beta^k$. It now follows from the Corollary to Proposition 2.1 that $\theta_\Omega(x^*) = 0$, which completes our proof. ◻

<u>Lemma 2.2</u>: Suppose that the sequence $\{x_k\}$ constructed by the Master Algorithm is finite and that the sequence $\{z_i\}$ is infinite, then any accumulation point z^* of $\{z_i\}$ satisfies $z^* \in F_\Omega$ and $\theta_\Omega(z^*) = 0$.

<u>Proof</u>:

<u>Case 1</u>: Suppose $z_i \xrightarrow{K} z^*$ as $i \to \infty$, $K \subset \{0,1,2,\ldots\}$ and $\underline{\phi}(z_i,\underline{\omega}_i)] \leq 0$ for all $i \in K$. Then we must have $\overline{\lim}\ \underline{\phi}(z_i,\underline{\omega}_i) \leq 0$. But, by Assumption 2.1, $[\max_{\underline{\omega} \in \Omega} \underline{\phi}(z_i,\underline{\omega}) - \phi(z_i,\underline{\omega}_i)] \to 0$ as $i \to \infty$ and hence, since $z_i \xrightarrow{K} z^*$ as $i \to \infty$, $K \subset \{0,1,2,\ldots\}$ results in $\max_{\underline{\omega} \in \Omega} \underline{\phi}(z_i,\underline{\omega}) \xrightarrow{K} \max_{\underline{\omega} \in \Omega} \phi(z^*,\underline{\omega})$, we must have $\max_{\underline{\omega} \in \Omega} \underline{\phi}(z^*,\underline{\omega}) \leq 0$. Also $g^j(z_i) \leq \mu_2\beta^i$ for all $j = 1,2,\ldots,p$ and all $i = 1,2,\ldots$, by construction in step 1. It now follows by continuity that $g^j(z^*) \leq 0$ for all $j = 1,2,\ldots,p$. Thus, we have shown that $z^* \in F_\Omega$. It now follows from Proposition 2.1 and its Corollary that $\theta_\Omega(z^*) = 0$.

Case 2: Suppose $z_i \xrightarrow{K} z^*$ as $i \to \infty$, and there exists an infinite subsequence, indexed by $K' \subset K$, such that $\underline{\phi}(z_i, \underline{\omega}_i) > 0$ for all $i \in K'$. Since the sequence $\{x_k\}$ is finite, there exists an i_0 such that $\Omega_i \subset \Omega_{i+1}$ for all $i \geq i_0$ and hence $\underline{\omega}_i \in \Omega_{i+j}$ for all $i \geq i_0$, $j = 1,2,\ldots$ Suppose, for the sake of contradiction, that

$$\max_{\underline{\omega} \in \Omega} \underline{\phi}(z^*, \underline{\omega}) = 2\gamma > 0 \tag{25}$$

Now we have, since $z_i \xrightarrow{K'} z^*$,

$$|\underline{\phi}(z_{i+j}, \underline{\omega}_i) - \underline{\phi}(z_i, \underline{\omega}_i)| \to 0 \tag{26}$$

as $i \to \infty$, $i \in K'$, $i+j \in K'$, and

$$\underline{\phi}(z_{i+j}, \underline{\omega}_i) \leq \mu_2 \beta^{i+j} \tag{27}$$

for all $i \in K'$, $i \geq i_0$ and all $j = 1,2,\ldots$ by construction in step 1 of the Master Algorithm. Next, because of the properties of the map s (Assumption 2.3) we have

$$|\underline{\phi}(z_i, \underline{\omega}_i) - \max_{\underline{\omega} \in \Omega} \underline{\phi}(z_i, \underline{\omega})| \to 0 \text{ as } i \to \infty \tag{28}$$

Combining (28), (25) and (26), we conclude that there is an index $i_1 \geq i_0$ in K' such that

$$\underline{\phi}(z_{i+j}, \underline{\omega}_i) > \gamma > \mu_2 \beta^{i+j} \tag{29}$$

for all $i \geq i_1$, $i \in K'$, $i+j \in K'$, which contradicts (27). Hence we must have $\max_{\underline{\omega} \in \Omega} \underline{\phi}(z^*, \underline{\omega}) \leq 0$. Since $g^j(z_i) \leq \mu_2 \beta^i$ for all $i = 0,1,2,\ldots$, for all $j = 1,2,\ldots,p$, it follows, by continuity, that $g^j(z^*) \leq 0$, for all $j = 1,2,\ldots,p$, and hence that $z^* \in F_\Omega$. It now follows from the Corollary to Proposition 2.1 that $\theta_\Omega(z^*) = 0$. This completes the proof. ◻

We can summarize the conclusions of Lemmas 2.1 and 2.2 as follows.

<u>Theorem 2.1</u>: Suppose that the Master Algorithm constructs a sequence $\{z_i\}$, possibly dropping constraints and constructing a corresponding subsequence $\{x_k\}$. If the subsequence $\{x_k\}$ is finite (i.e. if no constraints are dropped beyond a certain point) and $\{z_i\}$ is infinite, then every accumulation point z^* of $\{z_i\}$ satisfies $z^* \in F_\Omega$, $\theta_\Omega(z^*) = 0$. If the (sub) sequence $\{x_k\}$ is infinite then every accumulation point x^* of $\{x_k\}$ satisfies $x^* \in F_\Omega$, $\theta_\Omega(x^*) = 0$. ¤

Appendix A

The subalgorithm presented here is a new dual method of feasible directions. It is included as a method for executing step 1 of the Master Algorithm. This subalgorithm has two highly desirable features which were evolved from earlier work.[Refs. 6,15]. The first feature is a dual type direction finding subproblem which (cf. [Ref. 15]) provides a better descent direction and results in a faster speed of convergence than conventional methods of feasible directions (which use box constraints). The second feature is the elimination of the requirement of starting at an initial feasible point. The method developed in [Ref. 6] for starting at arbitrary initial points has been modified to be used in conjuction with the dual direction finding subproblem (see also [Ref. 5]).

AI. Assumptions and Definitions

We begin by stating our first hypothesis for problem P. The assumptions which we use are standard for methods of feasible directions.

Assumption A1.

For any discrete set $\Omega' \subset \Omega$ the set $F_{\Omega'}$ satisfies: (a) int $F_{\Omega'} \neq \phi$, (b) $\overline{\text{Int } F_{\Omega'}} = F_{\Omega'}$,[8] and (c) $F_{\Omega'}$ is compact. ¤

For any $y \in \mathbb{R}^n$, for any discrete $\Omega' \subset \Omega$, and $\varepsilon \geq 0$, we define the "ε-active" constraint sets:

$$J_{\Omega',\varepsilon}(y) \triangleq \{j \mid g^j(y) - \psi_{\Omega'}(y) \geq -\varepsilon, \quad j = 1,2,\ldots,p\} \tag{A1}$$

$$L_{\Omega',\varepsilon}(y) \triangleq \{\ell \mid \max_{\omega^\ell \in \Omega'^\ell} \phi^\ell(y,\omega^\ell) - \psi_{\Omega'}(y) \geq -\varepsilon, \quad \ell = 1,2,\ldots,q\} \tag{A2}$$

8 Int denotes the interior of a set and the overbar indicates closure.

$$\Omega_\varepsilon^{\prime\ell}(y) \triangleq \{\omega^\ell \in \Omega^{\prime\ell} \mid \phi^\ell(y,\omega^\ell) - \psi_{\Omega'}(y) \geq -\varepsilon\} \quad \ell = 1,2,\ldots,q \qquad \text{(A3)}$$

where $\psi_{\Omega'}(\cdot)$ is defined in (2). For any discrete set $\Omega' \subset \Omega$ and $\varepsilon \geq 0$ we define the optimality function[9] $\theta_{\Omega',\varepsilon} : \mathbb{R}^n \to \mathbb{R}$ by

$$\theta_{\Omega',\varepsilon}(y) \triangleq \min_h \{\tfrac{1}{2}\|h\|_2^2 + \max\{\langle \nabla f(y),h\rangle - \psi_{\Omega'}(y);$$

$$g^j(y) - \psi_{\Omega'}(y) + \langle \nabla g^j(y),h\rangle,\ j \in J_{\Omega',\varepsilon}(y);$$

$$\phi^\ell(y,\omega^\ell) - \psi_{\Omega'}(y) + \langle \nabla\phi^\ell(y,\omega^\ell),h\rangle,\ \omega^\ell \in \Omega_\varepsilon^{\prime\ell}(y),\ \ell \in L_{\Omega',\varepsilon}(y)\}\}$$

$$= \max_{\mu \geq 0} \{-\tfrac{1}{2}\|\mu_f \nabla f(y) + \sum_{j \in J_{\Omega',\varepsilon}(y)} \mu_g^j \nabla g^j(y)$$

$$+ \sum_{\ell \in L_{\Omega',\varepsilon}(y)} \sum_{\omega^\ell \in \Omega_\varepsilon^{\prime\ell}} \mu_\phi^{\ell,\omega^\ell} \nabla_y \phi^\ell(y,\omega^\ell)\|_2^2 - \mu_f \psi_{\Omega'}(y)$$

$$+ \sum_{j \in J_{\Omega',\varepsilon}(y)} \mu_g^j (g^j(y) - \psi_{\ell'}(y))$$

$$+ \sum_{\ell \in L_{\Omega',\varepsilon}(y)} \sum_{\omega^\ell \in \Omega_\varepsilon^{\prime\ell}(y)} \mu_\phi^{\ell,\omega^\ell} (\phi^\ell(y,\omega^\ell) - \psi_{\Omega'}(y)) \mid \mu_f$$

$$+ \sum_{j \in J_{\Omega',\varepsilon}(y)} \mu_g^j + \sum_{\ell \in L_{\Omega',\varepsilon}(y)} \sum_{\omega^\ell \in \Omega_\varepsilon^{\prime\ell}(y)} \mu_\phi^{\ell,\omega^\ell} = 1\}$$

(A4)

where the second expression follows from duality [Ref. 15]. We define the descent direction vector by

$$h_{\Omega',\varepsilon}(y) \triangleq -\mu_f \nabla f(y) - \sum_{j \in J_{\Omega',\varepsilon}(y)} \mu_g^j \nabla g^j$$

$$- \sum_{\ell \in L_{\Omega',\varepsilon}(y)} \sum_{\omega^\ell \in \Omega_\varepsilon^{\prime\ell}} \mu_\phi^{\ell,\omega^\ell} \nabla_y \phi^\ell(y,\omega^\ell) \qquad \text{(A5)}$$

---

[9] It is easily shown that $\theta_{\Omega',0}$ is an u.s.c. function with the property that if $\hat{z}$ is optimal for $P_{\Omega'}$, then $\theta_{\Omega',0}(\hat{z}) = 0$. Furthermore, $\theta_{\Omega',0}(\hat{z}) = 0$ if and only if $\hat{z} \in F_{\Omega'}$ is an F. John point.

where the μ's are obtained by solving the second expression in (A4) which is a quadratic program.

We also require the following condition to hold:

Assumption A2.

For any discrete set $\Omega' \in \Omega$ and for any $y \in \mathbb{R}^n$ the set of vectors $\{\nabla g^j(y), j \in J_{\Omega',0}(y);\ \nabla_y \phi^\ell(y,\omega^\ell),\ \omega^\ell \in \Omega'^\ell(y),\ \ell \in L_{\Omega',0}(y)\}$ is positive linearly independent.[10]

$\square$

Satisfaction of this assumption ensures that for any $y \in F^c_{\Omega'}$, (the complement of $F_{\Omega'}$), $\theta_{\Omega',0}(y) < 0$, and hence ensures that the subalgorithm to be presented cannot jam up at an infeasible point. It also ensures that the subalgorithm cannot converge to trivial F. John points in $F_{\Omega'}$.

AII. The Subalgorithm

Given any $i \in \{0,1,2,\ldots\}$ and the corresponding Ω_i and z_{i-1} from the Master Algorithm, the following subalgorithm finds a point z_i such that $\theta_{\Omega_i}(z_i) \geq -\mu_1\beta^i$ and $\psi_{\Omega_i}(z_i) \leq \mu_2\beta^i$.

The subalgorithm can now be stated.

Subalgorithm

Data: $i \in \{0,1,2,\ldots\}$, $z_{i-1} \in \mathbb{R}^n$, $\Omega_i \subset \Omega$, $\mu_1 > 0$, $\mu_2 > 0$, $\beta \in (0,1)$, $\alpha \in (0,1)$, $\bar{\beta} \in (0,1)$, $S > 0$, $\delta > 0$, $\varepsilon_0 > 0$, $0 < \varepsilon_1 << 1$.

Step 0: Set $k = 0$, $\varepsilon = \varepsilon_0$, $y_0 = z_{i-1}$.

Step 1: Compute $\theta_{\Omega_i,\varepsilon}(y_k)$ and $h_{\Omega_i,\varepsilon}(y_k)$ by solving (A4) and (A5).

[10] This assumption is related to the Kuhn-Tucker constraint qualification [Ref. 16]. We say a set of vectors $\{\eta_j\}_{j=1}^{s}$ is positive linearly independent if the zero vector is not contained in the convex hull of $\{\eta_j\}_{j=1}^{s}$.

Step 2: If $\theta_{\Omega_i,\varepsilon}(y_k) \leq -\delta\varepsilon$, go to step 4; else set $\varepsilon = \varepsilon/2$ and go to step 3.

Step 3: If $\theta_{\Omega_i,\varepsilon}(y_k) \geq -\mu_1\beta^i$ and $\psi_{\Omega_i}(y_k) \leq \mu_2\beta^i$ set $z_i = y_k$ and stop; else go to step 1.

Step 4: If $\psi_{\Omega_i}(y_k) = 0$ compute the largest step size $s_k = \bar{\beta}^{\lambda_k} \in (0, S/\|h_{\Omega_i,\varepsilon}(y_k)\|_\infty]$ (λ_k an integer) satisfying

$$f(y_k + s_k h_{\Omega_i,\varepsilon}(y_k)) - f(y_k) \leq -\alpha\delta\varepsilon s_k \tag{A6}$$

$$g^j(y_k + s_k h_{\Omega_i,\varepsilon}(y_k)) \leq 0 \quad j = 1,2,\ldots,p \tag{A7}$$

$$\phi^\ell(y_k + s_k h_{\Omega_i,\varepsilon}(y_k),\omega^\ell) \leq 0 \quad \omega^\ell \in \Omega_i^\ell, \quad \ell = 1,2,\ldots,q \tag{A8}$$

If $\psi_{\Omega_i}(y_k) > 0$ compute the largest step size $s_k = \bar{\beta}^{\lambda_k} \in (0, S/\|h_{\Omega_i,\varepsilon}(y_k)\|_\infty]$ (λ_k an integer) satisfying

$$\tilde{\psi}_{\Omega_i}(y_k + s_k h_{\Omega_i,\varepsilon}(y_k)) - \tilde{\psi}_{\Omega_i}(y_k) \leq -\alpha\delta\varepsilon s_k \tag{A9}$$

Step 5: Set $y_{k+1} = y_k + s_k h_{\Omega_i,\varepsilon}(y_k)$, $\varepsilon_{k+1} = \varepsilon$, $k = k+1$, and go to step 1.

¤

Theorem A1: Under Assumptions 2.1 and 2.2 and A1 and A2, the subalgorithm computes a point z_i, satisfying $\psi_{\Omega_i}(z_i) \leq \mu_2\beta^i$ and $\theta_{\Omega_i}(z_i) \geq -\mu_1\beta^i$ in a finite number of iterations.

Proof: By modifying a convergence theorem due to Klessig [Ref. 17] in the same manner that Theorem (1.3.10) of [Ref. 13] was modified in [Ref. 6] we can conclude that if the sequence $\{y_k\}$ constructed by the subalgorithm is infinite, then every accumulation point y^* of $\{y_k\}$ satisfies $\psi_{\Omega_i}(y^*) = 0$.

[1] It is tacitly assumed that any sequence $\{y_k\}$ constructed by the subalgorithm is contained in a compact set.

Also, the sequence $\{\varepsilon_k\}$ must satisfy $\varepsilon_k \to 0$ as $k \to \infty$.

Now for the sake of contradiction, suppose that $\{y_k\}$ is infinite, i.e. the stop command in step 3 of the subalgorithm is not executed. Since $\varepsilon_k \to 0$ and $k \to \infty$, there exists an infinite subsequence $K \subset \{0,1,2,\ldots\}$ such that (in step 2) $\theta_{\Omega_i,\varepsilon_k}(y_k) > -\delta\varepsilon_k$ for all $k \in K$ and $\varepsilon_{k+1} < \varepsilon_k$ is produced. Hence, there exists a finite integer k' such that $-\mu_1\beta^i \leq -\delta\varepsilon_k < \theta_{\Omega_i,\varepsilon_k}(y_k)$ for all $k \geq k'$, $k \in K$. By comparing (A4) and (9) we get $\theta_{\Omega_i,\varepsilon}(y) \leq \theta_{\Omega_i}(y)$ for all $\varepsilon \geq 0$ and for all $y \in \mathbb{R}^n$. Hence, $-\mu_1\beta^i < \theta_{\Omega_i}(y_k)$ for all $k \geq k'$, $k \in K$. Let y^* be any accumulation point of $\{y_k\}_{k \in K}$, i.e., $y_k \to y^*$, $k \in K' \subset K$. Since $\psi_{\Omega_i}(y^*) = 0$, there exists a $k'' \geq k'$ such that $\psi_{\Omega_i}(y_k) \leq \mu_2\beta^i$ for all $k \geq k''$, $k \in K'$. Thus, the stop command in step 4 of the algorithm would be executed at the first $k \in K'$ satisfying $k \geq k''$. This is a contradiction to the assumption that $\{y_k\}$ is infinite. ¤

Appendix B

We now present an example of a control system design problem which is in the form of problem P. Given the system in Figure 1 with

$$G(s) = \frac{1}{(s+3)(s^2+2s+2)} \tag{B1}$$

we wish to design a PID series compensator $H(z,s) = z^1 + z^2/s + z^3 s$ to give the closed loop system a phase margin not smaller than 45°, and to minimize mean square error in the zero-state response to a step input. The cost becomes

$$f(z) = \int_0^\infty e^2(z,t)dt = \frac{z^2(122+17z^1+6z^3-5z^2+z^1z^3) + 180z^3 - 36z^1 + 1224}{z^2(408+56z^1-50z^2+60z^3+10z^1z^3-2(z^1)^2)} \tag{B2}$$

where we have used Parseval's theorem and the tables in [Ref. 18].

The phase margin requirements are formulated as an inequality constraint $\phi^1(z,\omega) \leq 0$ for $\omega \in \Omega^1 \triangleq [10^{-6}, 30.0]$ where

$$\phi^1(z,\omega) \triangleq \mathrm{Im}\, T(z,\omega) - 3.33(\mathrm{Re}\, T(z,\omega))^2 + 1.0 \qquad \text{(B3)}$$

and $T(z,\omega) = 1 + H(z,j\omega)\, G(j\omega)$. With this constraint satisfied the Nyquist plot will remain outside a parabolic region as shown in Figure 2. Conventional constraints are placed on the gains, $0 \leq z^1 \leq 100$, $0.1 \leq z^2 \leq 100$, $0 \leq z^3 \leq 100$.

The max calculation required in step 2 of the Master Algorithm is performed by discretizing Ω^1 into $t(i)$ equally spaced points and calculating the maximum of $\phi^1(z,\omega)$ over this discrete set. The iteration truncation function $t : \mathbb{Z}^+ \to \mathbb{Z}^+$ is given by $t(i) = 2^{\max\{5,i\}} + 1$. This approximation method satisfies the requirements of Assumption 3. The parameters used in the Master Algorithm are:

$$\tau = 10^{-3},\ \beta = 0.5,\ \gamma = 10^{-3},\ \mu_1 = 10^{-8},\ \mu_2 = 10^{-4}$$

The parameters used in the subalgorithm in Appendix A are:

$$\alpha = 0.2,\ \bar{\beta} = 0.3,\ S = 15.0,\ \delta = 10^{-3},\ \varepsilon_0 = 0.02$$

Our results are tabulated in Table 1. The column marked "Iterations Subalgo" gives the number of iterations of the feasible directions sub-algorithm needed to execute step 1 of the Master Algorithm. The total CPU time used for the 12 iterations shown in Table 1 was approximately 23 seconds on a CDC 6400 computer.

Conclusion

To a very large extent, the efficiency of an outer approimation algorithm depends on how successful it is in dropping constraints. So far, it is impossible to predict theoretically that all but a finite number of constraints must be dropped. However, our very easily satisfied constraint dropping test (18) coupled with the fact that, the subalgorithm in Appendix A constructs points z_i in the interior of the feasible set, lead us to expect that constraints would be dropped en masse. Our experimental results bear this out. The current algorithm tends to be more efficient even than the highly specialized algorithm described in [Ref. 6], because the new algorithm spends much less time computing step size. It is also more efficient than penalty function methods which tend to stall at fairly low precision on this type of problem.

REFERENCES

1. BANDLER, John W., LIU, Peter C., CHEN, James H., Worst Case Network Tolerance Optimization, IEEE Transactions on Microwave Theory and Technology, Vol. MTT-23, No. 8, pp. 630-641, Aug. 1975.

2. BANDLER, J. W., Optimization of Design Tolerances using Nonlinear Programming, Journal of Optimization Theory and Applications, Vol. 14, No. 1, pp. 99-114, July 1974.

3. BANDLER, John W., LIU, Peter C., TROMP, Herman, A Nonlinear Programming Approach to Optimal Design Centering, Tolerancing, and Tuning, IEEE Transactions on Circuits and Systems, Vol. CAS-23, No. 3, pp. 155-165, March 1976.

4. ZAKIAN, V., AL-NAIB, U., Design of Dynamical and Control Systems by the Method of Inequalties, Proceedings of the IEEE, Vol. 120, No. 11, pp. 1421-1427, Nov. 1973.

5. POLAK, E. and TRAHAN, R., An Algorithm for Computer Aided Design Problems, Proceedings of the 1976 IEEE Conference on Decision and Control, Dec. 1976.

6. POLAK, E. and MAYNE, D. Q., An Algorithm for Optimization Problems with Functional Inequality Constraints, IEEE Transactions on Automatic Control, Vol. AC-21, No. 2, pp. 184-193, April 1976.

7. POLAK, E., PISTER, K. S., RAY, D. Optimal Design of Framed Structures Subjected to Earthquakes, Engineering Optimization, Vol. 2, pp. 65-71, 1976.

8. RAY, D., PISTER, K. S., CHOPRA, A. K., Optimum Design of Earthquake Resistant Shear Buildings, University of California, Berkeley, Earthquake Engineering Research Center, Report EERC 74-3, 1974.

9. DEMYANOV, V. F., On the Solution of Certain Min Max Problems I, Kibern, Vol. 2, pp. 47-53, 1966.

10. LEVITIN, E. S., POLYAK, B. T., Constrained Minimization Methods, Zh. Vychisl. Mat. Mat. Fiz., 6, 5, pp. 787-823, 1966.

11. EAVES, B. CURTIS, ZANGWILL, W. I., Generalized Cutting Plane Algorithms, SIAM Journal of Control, Vol. 9, No. 4, pp. 529-542, Nov. 1971.

12. BLANKENSHIP, Jerry W. and FALK, James E., Infinitely Constrained Optimization Problems, The George Washington University, Institute for Management Science and Engineering, Serial T-301, June 1974.

13. POLAK, E., Computational Methods in Optimization: A Unified Approach, Academic Press, New York, New York, 1971.

14. JOHN, F., 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.

15. PIRONNEAU, O. and POLAK, E., Rate of Convergence of a Class of Methods of Feasible Directions, SIAM Journal of Numerical Analysis, Vol. 10, No. 1, pp. 161-174, March 1973.

16. CANON, M. D., CULLUM, C. D., POLAK, E., Theory of Optimal Control and Mathematical Programming, McGraw-Hill, New York, New York, 1970.

17. KLESSIG, R., A General Theory of Convergence for Constrained Optimization Algorithms that use Anti-Zigzagging Provisions, Bell Telephone Laboratories, Holmdel, New Jersey, Technical Report, Aug. 1972.

18. CHEN, C. T., Analysis and Synthesis of Linear Control Systems, Holt, Rinehart and Winston, New York, New York, 1975.

i	k	z_i^1	z_i^2	z_i^3	$f(z_i)$	ω_i	$\bar{\phi}(z_i,\omega_i)$	Iterations Subalgo
0	0	34.641	56.797	99.999	0.1274	10.312	4.949×10^{-1}	32
1	1	28.778	57.985	64.375	0.1392	7.500	3.615×10^{-1}	18
2	2	23.281	57.619	41.538	0.1597	6.562	1.387×10^{-1}	17
3	3	20.189	48.874	36.455	0.1686	5.625	6.259×10^{-1}	72
4	4	16.780	46.776	34.827	0.1746	5.625	≤ 0	55
5	4	16.631	46.413	34.905	0.1746	5.625	≤ 0	8
6	4	16.694	45.945	34.852	0.1746	5.625	≤ 0	21
7	4	16.694	45.944	34.852	0.1746	5.625	≤ 0	1
8	4	16.724	45.729	34.827	0.1746	5.625	≤ 0	24
9	4	16.724	45.729	34.827	0.1746	5.684	1.921×10^{-4}	1
10	5	17.140	45.261	34.571	0.1746	5.654	1.704×10^{-4}	97
11	6	16.966	45.362	34.665	0.1746	5.654	≤ 0	118
12	6	16.967	45.362	34.665	0.1746	5.654	≤ 0	2

Table 1

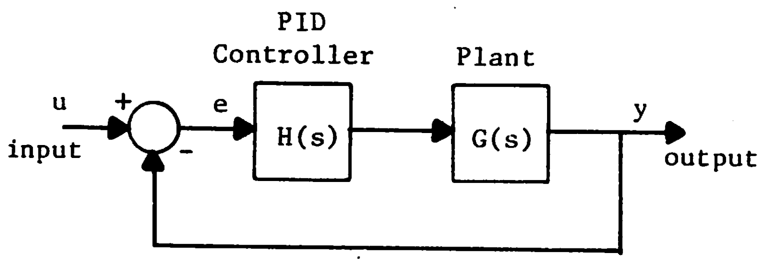

Fig. 1. System block diagram for PID controller design problem.

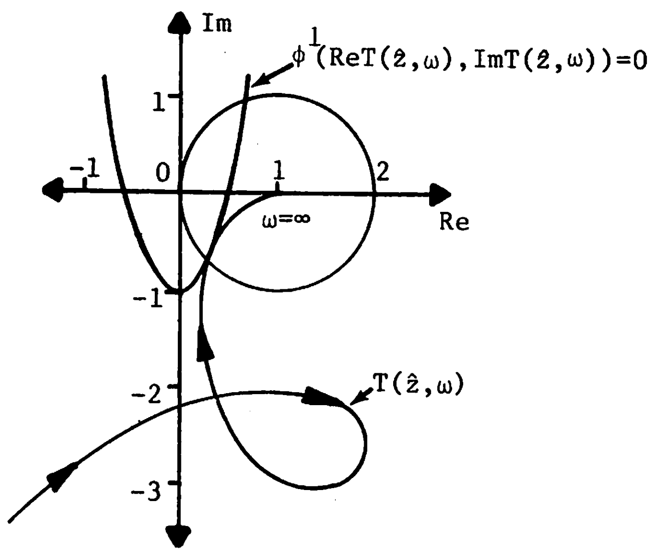

Fig. 2. Nyquist plot for system of Fig. 1 with gains $\hat{z}^1=16.97$, $\hat{z}^2=45.36$, and $\hat{z}^3=34.66$.