AN ALGORITHM FOR OPTIMIZATION PROBLEMS WITH FUNCTIONAL INEQUALITY CONSTRAINTS

by

E. Polak and D. Q. Mayne

Memorandum No. ERL-M456

1 August 1974

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

An Algorithm for Optimization Problems with Functional Inequality Constraints

E. Polak & 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 an algorithm for minimizing an objective function subject to conventional inequality constraints as well as to inequality constraints of the functional type: $\max_{\omega \in \Omega} \phi(z,\omega) \leq 0$, where Ω is a closed interval in R, and $z \in R^n$ is the parameter vector to be optimized. The algorithm is motivated by a standard earthquake engineering problem and the problem of designing linear multivariable systems. The stability condition (Nyquist criterion) and disturbance suppression condition for such systems are easily expressed as a functional inequality constraints.

Research sponsored by the U.S. Army Research Office Durham Contract DAHC04-73-C-0025 and National Science Foundation Grant GI-36387.

1. Introduction.

This paper presents an algorithm for solving problems of the form:

$$\min\{f^o(z) \mid f^j(z) \leq 0,\ j = 1, \ldots m;\ g^j(z) \leq 0,\ j = 1 \ldots r\}$$

where $z \in R^n$ is the parameter vector to be optimized, $f^o\colon\ R^n \to R$ is the criterion, $g^j\colon\ R^n \to R$, $j = 1, \ldots p$, are conventional inequality constraints, and $f^j\colon\ R^n \to R$, $j = 1, \ldots m$, are inequality constraints of the functional type, i.e.

$$f^j(z) \triangleq \max_{\omega \in \Omega} \phi^j(z,\omega),\ j = 1 \ldots m \tag{1}$$

where Ω is a closed interval in R. The algorithm is motivated by problems arising in the design of earthquake resistant buildings [10] and in designing controllers for linear multivariable systems using frequency response techniques. For the latter, a sufficient condition for asymptotic stability of the closed-loop system is that (if the open-loop system is asymptotically stable) for all $\omega \in R$, $\det(T(i\omega,z))$ lies outside of a suitably specified subset B_1 of the complex plane (see Fig. 1), where $\det(T(\cdot,z))$ is the determinant of the matrix return difference of the multivariable system corresponding to parameter z. If B_1 can be defined by:

$$B_1 \triangleq \{s \in \mathbb{C} \mid \theta(s) \leq 0\} \tag{2}$$

where $\theta\colon \mathbb{C} \to R$, then the stability functional constraint is:

$$f^1(z) \leq 0$$

where $f^1(z) \triangleq \max_{\omega \in R} \phi^1(z,\omega)$, and $\phi^1(z,\omega) \triangleq \theta(|T(i\omega,z)|)$

In practice it is known that there exists a $\omega_c < \infty$ such that $\phi^1(z,\omega) < 0$ for all feasible z, for all $\omega \geq \omega_c$, so that R may be replaced by $\Omega = [0,\omega_c]$ in the definition of f^1.

Similarly, a performance constraint can be specified, for some $\delta > 0$, as $\|T^{-1}(i\omega,z)\| \leq \delta$ for all $\omega \in \Omega$.

Existing procedures for designing linear multivariable systems, when the state is not completely accessible, are either of the time-domain [1,2] or the frequency-domain type [3,4,5]. The former require relatively lengthy, complex calculations, and, usually, the specification of an artificial criterion (which may, itself, have to be repeatedly adjusted to meet basic design objectives); the latter require, at least in their current stage of development, considerable skill, and they do not necessarily yield the simplest designs. Since the latter procedures easily handle design objectives (e.g. stability and performance) which are meaningful in many practical situations, it seemed useful to develop an algorithm to achieve these objectives automatically. Standard algorithms cannot be employed since at least some of the constraints are of the functional type.

Because the problem contains functional constraints which cannot be evaluated without resorting to approximations, the algorithm is presented in two versions: the conceptual version, which assumes that the functionals can be evaluated exactly, and the implementable, which incorporates efficient approximation procedures compatible with convergence. The conceptual version is given primarily because it considerably simplifies the presentation.

In constructing these new algorithms we have made use of ideas appearing in Zontendijk's feasible directions algorithms [6] and in Demyanov's algorithm [7] for min max problems.

2. Definitions and Assumptions.

The problem considered is:

P1. $\min\{f^o(z) \mid f^j(z) \leq 0,\ j \in J_m;\ g^j(z) \leq 0,\ j \in J_p\}$

where $J_k \triangleq \{1,2, \ldots k\}$, $k = m, p$, f^j, $j \in J_m$, is defined in (1), and the following conditions are assumed to be satisfied:

A1. $f^o\colon R^n \to R$, $g^j\colon R^n \to R$, $j = 1 \ldots r$ are continuously differentiable.

A2. ϕ^j: $R^n \times \Omega \to R$, $j = 1 \ldots r$ are continuously differentiable.

Let $F \triangleq \{z \in R^n | f^j(z) \leq 0,\ j \in J_m;\ g^j(z) \leq 0,\ j \in J_p\}$ denote the <u>feasible set</u>, and $\mathring{F}$ its interior. The algorithm is an extension of a method of feasible directions, and requires the determination of the "ε-active constraints". Thus, let:

$$J_\varepsilon^f(z) \triangleq \{j \in J_m | f^j(z) - \psi_o(z) \geq -\varepsilon\} \quad (3)$$

$$J_\varepsilon^g(z) \triangleq \{j \in J_p | g^j(z) - \psi_o(z) \geq -\varepsilon\} \quad (4)$$

where ψ_o: $R^n \to R$ is defined by:

$$\psi_o(z) \triangleq \max\{0, \psi(z)\} \quad (5)$$

and ψ: $R^n \to R$ is defined by:

$$\psi(z) \triangleq \max\{f^j(z),\ j \in J_m;\ g^j(z),\ j \in J_p\} \quad (6)$$

Clearly, if $z \in F$, then $\psi_o(z) = 0$. If $j \in J_\varepsilon^f(z)$, then $\phi^j(z,\omega) - \psi_o(z) \geq -\varepsilon$ for some $\omega \in \Omega$; the set of all such ω will be denoted by:

$$\Omega_\varepsilon^j(z) \triangleq \{\omega \in \Omega | \phi^j(z,\omega) - \psi_o(z) \geq -\varepsilon\} \quad (7)$$

The following further assumptions are made:

A3. $\mathring{F} \neq \phi$, $\overline{\mathring{F}} = F$.

A4. $\forall z \in R^n$, $\forall j \in J_0^f(z)$, $\Omega_0^j(z)$ is a finite set ($\Omega_0^j(z) = \{\omega \in \Omega | \phi^j(z,\omega) = \psi_o(z)\}$)

A5. $\forall z \in R^n$, $\forall\ \varepsilon > 0$, $\forall j \in J_\varepsilon^f(z)$, $\Omega_\varepsilon^j(z)$ is the union of a finite number of disjoint intervals $g_{\varepsilon,k}^j(z)$, $k = 1, \ldots k_\varepsilon^j(z)$, possibly of zero length, i.e.

$$\Omega_\varepsilon^j(z) = \bigcup_{k \in \mathcal{K}_\varepsilon^j(z)} I_{\varepsilon,k}^j(z) \quad (8)$$

where: $\mathcal{K}_\varepsilon^j(z) \triangleq \{1, 2, \ldots K_\varepsilon^j(z)\}$ (9)

The algorithm also uses a set of discrete frequencies $\overline{\Omega}_\varepsilon^j(z) \subset \Omega_\varepsilon^j(z)$ defined $\forall z \in R^n$, $\forall j \in J_\varepsilon^f(z)$, by:

$$\overline{\Omega}_\varepsilon^j(z) \triangleq \{\omega_{\varepsilon,k}^j(z) | k \in K_\varepsilon^j(z)\} \quad (10)$$

---

†The overbar denotes closure.

where $\omega^j_{\varepsilon k}$ is the midpoint of $I^j_{\varepsilon,k}$, $\forall k \in K^j_\varepsilon(z)$.

Most algorithms require directional derivatives of the constraint functions. This causes no difficulties for the conventional constraints g^j, $j \in J_p$ - the directional derivative corresponding to direction $h \in R^n$ is $\langle \nabla g^j(z), h \rangle$. To cope with the functional constraints, the function Df^j_ε: $R^n \times R^n \to R$, defined by:

$$
\begin{aligned}
Df^j_\varepsilon(z,h) &\triangleq \max\{ \langle \nabla_z \phi^j(z,\omega), h \rangle, \ \omega \in \Omega^j_\varepsilon(z)\}, \ j \in J^f_\varepsilon(z) \\
&\triangleq -\infty, \ j \notin J^f_\varepsilon(z) \qquad (11)
\end{aligned}
$$

$\forall \varepsilon \geq 0$, is introduced. Similarly, for convenience, Dg^j_ε: $R^n \times R^n \to R$ is defined by:

$$
\begin{aligned}
Dg^j_\varepsilon(z,h) &\triangleq \langle \nabla g_j(z), h \rangle, \ j \in J^g_\varepsilon(z) \\
&\triangleq -\infty, \ j \notin J^g_\varepsilon(z) \qquad (12)
\end{aligned}
$$

In choosing a search direction, a linear programming problem has to be solved. To make this low dimensional, the following 'approximation' $\overline{Df}^j_\varepsilon$: $R^n \times R^n \to R$ to Df^j_ε is employed:

$$
\begin{aligned}
\overline{Df}^j_\varepsilon(z) &\triangleq \max\{ \langle \nabla_z \phi^j(z,\omega), h \rangle, \ \omega \varepsilon \overline{\Omega}^j_\varepsilon(z)\}, \ j \in J^f_\varepsilon(z) \\
&\triangleq -\infty, \ j \notin J^f_\varepsilon(z) \qquad (13)
\end{aligned}
$$

Because of A4, A5, $Df^j_0 = \overline{Df}^j_0$, $\forall j \in J_m$.

For given $z \in R^n$, $\varepsilon > 0$, the search direction $h_\varepsilon(z) \in R^n$ is obtained by solving the following linear programming sub problem:

P2. Compute $h_\varepsilon(z) \in R^n$ by solving:

$$
\begin{aligned}
\overline{\theta}_\varepsilon(z) = \min_{h \in S} \max\{&-\psi_o(z) + \langle \nabla f^o(z), h \rangle; \\
&\overline{Df}^j_\varepsilon(z,h), \ j \in J_m; \ Dg^j_\varepsilon(z,h), \ j \in J_p\} \qquad (14)
\end{aligned}
$$

where:

$$
S \triangleq \{h \in R^n \mid \|h\|_\infty \leq 1\} \qquad (15)
$$

The function ψ_o makes the algorithm concentrate initially on driving z into the feasible region F. In the conceptual algorithm, once z becomes feasible it remains feasible.

To check that the computed search direction $h_\varepsilon(z)$ is satisfactory, $\tilde{\theta}_\varepsilon(z)$, defined by:

$$\tilde{\theta}_\varepsilon(z) = \max\{-\psi_o(z) + \langle \nabla f^o(z), h_\varepsilon(z) \rangle ; \; Df^j_\varepsilon(z,h_\varepsilon(z)),\ j \in J_m;\ Dg^j_\varepsilon(z,h_\varepsilon(z)),\ j \in J_p\} \tag{16}$$

is evaluated. To aid the analysis of the algorithm, the function $\theta_\varepsilon: R^n \to R$, defined, $\forall\ \varepsilon \geq 0$, by:

$$\theta_\varepsilon(z) = \min_{h \in S} \max\{-\psi_o(z) + \langle \nabla f^o(z), h \rangle ; Df^j_\varepsilon(z,h),\ j \in J_m; \; ; Dg^j_\varepsilon(z,h),\ j \in J_p\} \tag{17}$$

is introduced. $\theta_0(z) = 0$ is a necessary condition of optimality.

The final assumption can now be stated:

A6. $\forall z \in R^n$, $\{\nabla_z \phi^j(z,\omega),\ \omega \in \Omega^j_0(z),\ j \in J^f_0(z);\ \nabla g^j(z),\ j \in J^g_0(z)\}$ is a set of linearly independent vectors.[†]

This assumption is not easily verifiable. However it indicates that the algorithm may jam up, even when $z \in F$, at a point where it is not possible to reduce all the active constraints. This might indicate that the problem is not well posed, and that some constraints should be removed. The assumption ensures that there exists an $h \in R^n$ such that $Df^j_0(z,h) < 0$ $\forall j \in J_m$ and $Dg^j_0(z,h) < 0$ $\forall j \in J_p$, $\forall z \in F^c$.

3. The Conceptual Algorithm.

The conceptual algorithm can now be stated:

Algorithm I:

Data: $z_0 \in R^n$, $\alpha \in (0,1)$, $\beta \in (0,1)$, $\varepsilon_0 \in (0,\infty)$, $0 < \varepsilon_1 << 1$.

Step 0: Set $i = 0$.

Step 1: Set $\varepsilon = \varepsilon_0$.

Step 2: Compute $h_\varepsilon(z_i)$ and $\overline{\theta}_\varepsilon(z_i)$ by solving P2.

[†]This assumption is related to the Kuhn-Tucker constraint qualification.

Step 3: If $\bar{\theta}_\varepsilon(z_i) \leq -2\varepsilon$, go to Step 5.

Else, set $\varepsilon = \varepsilon/2$ and go to Step 4.

Step 4: If $\varepsilon \leq \varepsilon_1$ and $\theta_0(z_i) = 0$, stop. Else go to Step 2.[†]

Step 5: If $\tilde{\theta}_\varepsilon(z_i) \leq -\varepsilon$ go to Step 5. Else set $\varepsilon = \varepsilon/2$ and go to Step 2.

Step 6: If $z_i \in F$ compute the smallest integer ℓ_i such that:

$$f^o(z_i + \beta^{\ell_i} h_\varepsilon(z_i)) - f^o(z_i) \leq -\beta^{\ell_i}\alpha\varepsilon \tag{18}$$

$$f^j(z_i + \beta^{\ell_i} h_\varepsilon(z_i)) \leq 0, \quad \forall j \in J_m \tag{19}$$

$$g^j(z_i + \beta^{\ell_i} h_\varepsilon(z_i)) \leq 0, \quad \forall j \in J_p \tag{20}$$

If $z_i \in F^c$ (the complement of F in R^n) compute the smallest integer ℓ_i such that:

$$\psi(z_i + \beta^{\ell_i} h_\varepsilon(z_i)) - \psi(z_i) \leq -\beta^{\ell_i}\alpha\varepsilon \tag{21}$$

Step 7: Set $z_{i+1} = z_i + \beta^{\ell_i} h_\varepsilon(z_i)$, $i = i + 1$. Go to Step 1 (Step 2).[††]

#.

For future reference let $\bar{\varepsilon}(z)$ denote the value of ε generated by the algorithm (with "go to Step 1" in Step 7) given $z_i = z$, i.e. $\forall z \in R^n$, $\bar{\varepsilon}(z)$ is the largest number, of the form $\varepsilon_0/2^k$, where k is a non-negative integer, satisfying:

$$\bar{\theta}_{\bar{\varepsilon}(z)}(z) \leq -2\,\bar{\varepsilon}(z) \tag{22}$$

$$\tilde{\theta}_{\bar{\varepsilon}(z)}(z) \leq -\bar{\varepsilon}(z) \tag{23}$$

Note that Steps 1-7 of the conceptual algorithm (with "go to Step 1" in Step 7) define a Map A: $R^n \to 2^{R^n}$, so that $z_{i+1} \in A(z_i)$, with the property that $A(F) \subset F$.

[†] Note that if $\varepsilon > \varepsilon_1$, there is no need to evaluate $\theta_0(z_i)$.

[††] We shall give a proof of convergence for the case "go to Step 1" in Step 7, i.e., when ε is reset to ε_o at each iteration. The algorithm also converges when ε is forced to decrease monotonically, i.e. "go to Step 2" is used in Step 7. However, the proof of convergence is then substantially more complicated and will therefore be omitted.

The algorithm essentially has two different cost functions, one for F and one for F^c. In order to analyze its convergence properties we make use of the following general convergence theorem for algorithm models with this structure. (Theorem 1 below generalizes theorem 1.3.10 in [8]).

Theorem 1.

Suppose an algorithm A: $R^n \to 2^{R^n}$, with associated cost function c: $R^n \to R$, has the following properties (with $N_\varepsilon(z) \triangleq \{z' | \|z-z'\| \leq \varepsilon\}$):

1) $\exists\, F \subset R^n$ such that $c = c^1$ on F and $c = c^2$ on F^c, where c^1, c^2: $R^n \to R$ are continuous.

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

3) $\forall z \in R^n$, such that z is not desirable, $\exists \varepsilon(z) > 0$, $\exists \delta(z) < 0$ such that

(i) $c^1(z'') - c^1(z') \leq \delta(z) < 0, \quad \forall z' \in N_{\varepsilon(z)}(z) \cap F, \quad \forall z'' \in A(z')$

(ii) $c^2(z'') - c^2(z') \leq \delta(z) < 0, \quad \forall z' \in N_{\varepsilon(z)}(z) \cap F^c, \quad \forall z'' \in A(z')$

Then, every accumulation point of an infinite sequence $\{z_i\}$ generated according to the rule $z_{i+1} = y_i \in A(z_i)$ if $c(y_i) < c(z_i)$, $z_{i+1} = z_i$ otherwise, $i = 0, 1, 2, \ldots$, is desirable.†

Proof. Suppose $\{z_i\}$ is an infinite sequence, generated as above, which has a subsequence, indexed by K, converging to a $z \notin \Delta$, where Δ denotes the set of desirable points. From assumption 2), either there exists a finite k such that $z_i \in F\ \forall i \geq k$, or $z_i \in F^c\ \forall i$. In the first case $(j = 1)$ $c^1(z_i) \xrightarrow{K} c^1(z)$ (m in the second $(j = 2)$ $c^2(z_i) \xrightarrow{K} c^2(z)$. Hence, $\exists\ k' \in [0,\infty)$ such that, for j = 1 or 2:

$$c^j(z_{i+1}) - c^j(z_i) \leq \delta(z), \quad \forall i \in K \text{ such that } i \geq k'.$$

Hence for j = 1 or 2, (assuming, without loss of generality, that $z_0 \in F$ if $z_i \in F$ for all i sufficiently large)

$$c^j(z) - c^j(z_0) = \sum_{\substack{i \in K \\ i \geq k'}} [c^j(z_{i+1}) - c^j(z_i)]$$

---

†Note that if the sequence is finite, i.e. $z_i = z_\ell$ for all $i \geq \ell$, then z_ℓ is desirable.

$$+ \sum_{\substack{i \in K \\ i < k'}} [c^j(z_{i+1}) - c^j(z_i)]$$

$$+ \sum_{i \notin K} [c^j(z_{i+1}) - c^j(z_i)]$$

$$= - \infty.$$

since each term in the first summation is less than or equal to $- \delta(z)$ and all the other terms are negative. This contradicts the convergence of $\{c^j(z_i)\}$, j = 1 or 2. Hence z is desirable. #

To establish that the conceptual algorithm satisfies the assumptions of Theorem 1, we require the following results, which are proven in the Appendix.

Proposition 1. If $z \in F$ is optimal for P1, then $\theta_0(z) = \bar{\theta}_0(z) = \tilde{\theta}(z) = 0$. #

Proposition 2. $\forall z \in R^n$ such that $\theta_0(z) < 0$, $\exists\ \gamma > 0$, $\exists\ \varepsilon > 0$ such that:

$$\bar{\varepsilon}(z') \geq \varepsilon, \quad \forall z' \in N_\gamma(z).$$ #

Proposition 3. $\forall z \in R^n$ such that $\theta_0(z) < 0$, $\exists\ \gamma > 0$, $\exists\ \varepsilon > 0$ such that:

$$f^o(z') - f^o(z') \leq - \varepsilon, \quad \forall z' \in N_\gamma(z) \cap F, \quad \forall z'' \in A(z')$$

$$\psi(z'') - \psi(z') \leq - \varepsilon, \quad \forall z' \in N_\gamma(z) \cap F^c, \quad \forall z'' \in (A(z')),$$

where A: $R^n \to 2^{R^n}$ is defined by Steps 1 to 7 of the conceptual algorithm, with "go to Step 1" in Step 7. #

As a consequence of Proposition 1 we define the set Δ of desirable points as follows:

$$\Delta \triangleq \{z \in R^n | \theta_0(z) = 0\} \tag{24}$$

Since algorithm I stops only at $z_i \in \Delta$ we only need to consider the case where the sequence $\{z_i\}$ is infinite. Our first main result is:

Theorem 2.

Consider algorithm I with "go to Step 1" in Step 7. Then every accumulation point z^* of an infinite sequence $\{z_i\}$ generated by algorithm I is desirable.

Proof. Since f^o: $R^n \to R$ and ψ: $R^n \to R$ are continuous, the conceptual algorithm satisfies condition 1) of Theorem 1 will $c^1 \triangleq f^o$, $c^2 \triangleq \psi$.

It is obvious from Step 5 of the conceptual algorithm that if $z_i \in F$, then $z_{i+1} \in F$ and consequently, defining A: $R^n \to 2^{R^n}$ by Steps 1-7, we find that $A(z) \subset F$. Condition 2) of Theorem 1 is therefore satisfied.

Finally, Proposition 3 shows that the conceptual algorithm satisfies condition 3) of Theorem 1. #

We state without proof a similar result for the alternative version of the conceptual algorithm.

Theorem 3.

Consider algorithm I with "go to Step 2" in Step 7. Then every accumulation point z^* of an infinite sequence $\{z^i\}$ generated by algorithm I is desirable. #

The interested reader can prove Theorem 3, by modifying a convergence theorem due to Klessig [9], analogously to the way we have modified Theorem 1.3.10. in [8].

4. The Implementable Algorithm.

The algorithm below takes into account the fact that it is impossible to evaluate expressions such as $\max_{\omega \in \Omega} \phi^j(z,\omega)$, with high precision, a large number of times, without incurring high computational costs. We assume that Ω is a closed interval,

$$\Omega = [\omega_o, \omega_c] \tag{25}$$

and, for q = 1, 2, ..., $\varepsilon \geq 0$, we define

$$\Omega_q \triangleq \{\omega_\ell \mid \omega_\ell = \omega_o + \ell[\omega_c - \omega_o]/2^q,\ \ell = 0, 1, \ldots, 2^q\} \tag{26}$$

$$f_q^j(z) \triangleq \max_{\omega \in \Omega_q} \phi^j(z,\omega),\ j = 1, 2, \ldots, m \tag{27}$$

$$F_q \triangleq \{z \in R^n | f_q^j(z) \leq 0,\ j \in J_m;\ g^j(z) \leq 0,\ j \in J_p\} \tag{28}$$

$$\psi_q(z) \triangleq \max\{f_q^j(z),\ j \in J_m,\ q^j(z),\ j \in J_p\} \tag{29}$$

$$\psi_{q,o}(z) \triangleq \max\{0,\ \psi_q(z)\} \tag{30}$$

$$J_{q,\varepsilon}^f(z) \triangleq \{j \in J_m | f_q^j(z) - \psi_{q,o}(z) \geq -\varepsilon\} \tag{31}$$

$$J_{q,\varepsilon}^g(z) \triangleq \{j \in J_p | g^j(z) - \psi_{q,o}(z) \geq -\varepsilon\} \tag{32}$$

$$\Omega_{q,\varepsilon}^j(z) = \{\omega \in \Omega_q | \phi^j(z,\omega) - \psi_{q,o}(z) \geq -\varepsilon\},\ j \in J_m, \tag{33}$$

$$\overline{\Omega}_{q,\varepsilon}^j(z) = \{\omega_{q,\varepsilon,k}^j(z) | k \in \mathcal{K}_{q,\varepsilon}^j(z)\} \tag{34}$$

where the $\omega_{q,\varepsilon,k}^j(z)$ are mid points of intervals $I_{q,\varepsilon,k}^j(z) = [\omega',\omega'']$, where $\omega', \omega'' \in \Omega_{q,\varepsilon}^j(z)$, $\omega' = \omega''$ is allowed, and for any $\omega^* \in \Omega_{q,\varepsilon}^j(z)$, $\omega^* \notin I_{q,\varepsilon,k}^j(z)$ $\min_{\omega \in I_{q,\varepsilon,k}^j(z)} |\omega-\omega^*| \geq 2 \cdot [(\omega_c - \omega_o)/2^q]$.

Let $I^+ \triangleq \{0, 1, 2, \ldots\}$, then $\forall j \in J_m$, $\forall q \in I^+$, $\forall\ \varepsilon \geq 0$, $Df_{q,\varepsilon}^j$ is defined as in (11) with $\Omega_{q,\varepsilon}^j$ replacing Ω_ε^j, and $\overline{D}f_{q,\varepsilon}^j$ is defined as in (13) with $\overline{\Omega}_{q,\varepsilon}^j$ replacing $\overline{\Omega}_\varepsilon^j$. Next, $\forall q \in I^+$, $\forall\ \varepsilon \geq 0$, $\forall z \in R^n$, $h_{q,\varepsilon}(z)$ and $\overline{\theta}_{q,\varepsilon}(z)$ are obtain by solving:

P3. $$\overline{\theta}_{q,\varepsilon}(z) = \min_{h \in s} \max\{-\psi_{q,o}(z) + \langle \nabla f^o(z),h \rangle;\ \overline{D}f_{q,\varepsilon}^j(z,h),\ j \in J_m;\ Dg_\varepsilon^j(z,h),\ j \in J_p\} \tag{35}$$

Finally, $\forall q \in I^+$, $\forall\ \varepsilon \geq 0$, $\tilde{\theta}_{q,\varepsilon}$: $R^n \to R$ is defined as in (16) with $Df_{q,\varepsilon}^j$ replacing Df_ε^j, $\forall j \in J_m$.

<u>Algorithm II</u>.

<u>Data</u>: $z_0 \in R^n$, $\alpha \in (0,1)$, $\beta \in (0,1)$, $\varepsilon_0 \in (0,\infty)$ $0 < \varepsilon_1 << \varepsilon_0$, $q_0 > 0$, an integer.

<u>Step 0</u>: Set $i = 0$, $q = q_0$.

Step 1: Set $\varepsilon = \varepsilon_0$.

Step 2: Compute $h_{q,\varepsilon}(z_i)$ and $\bar{\theta}_{q,\varepsilon}(z_i)$ by solving P3.

Step 3: If $\bar{\theta}_{q,\varepsilon}(z_i) \leq -2\varepsilon$, go to Step 4. Else set $\varepsilon = \varepsilon/2$ and go to Step 2.

Step 4: If $\varepsilon \leq \varepsilon_1$ and $\bar{\theta}_{q,0}(z_i) = 0$, set $y_q = z_i$, set $q = q + 1$ and go to Step 1. Else go to Step 5.

Step 5: If $\tilde{\theta}_{q,\varepsilon}(z_i) \leq -\varepsilon$, go to Step 6. Else set $\varepsilon = \varepsilon/2$ and go to Step 2.

Step 6: If $z_i \in F_q$ compute the smallest integer ℓ_i such that

$$f^o(z_i + \beta^{\ell_i} h_{q,\varepsilon}(z_i)) - f^o(z_i) \leq -\beta^{\ell_i}\alpha\varepsilon \tag{36}$$

$$f_q^j(z_i + \beta^{\ell_i} h_{q,\varepsilon}(z_i)) \leq 0 \quad \forall j \in J_m \tag{37}$$

$$g^j(z_i + \beta^{\ell_i} h_{q,\varepsilon}(z_i)) \leq 0 \quad \forall j \in J_p \tag{38}$$

If $z_i \in F_q^c$ (the complement of F_q in R^n) compute the smallest integer ℓ_i such that

$$\psi_q(z_i + \beta^{\ell_i} h_{q,\varepsilon}(z_i)) - \psi_q(z_i) \leq -\beta^{\ell_i}\alpha\varepsilon \tag{39}$$

Step 7: Set $z_{i+1} = z_i + \beta^{\ell_i} h_{q,\varepsilon}(z_i)$, $i = i+1$.

Step 8: If $z_i \in F_q$ and $\varepsilon \leq \varepsilon_1$, go to Step 9. Else go to Step 2.

Step 9: If $\bar{\theta}_{q,0}(z_i) \geq -1/2^q$, set $y_q = z_i$, set $q = q + 1$ and go to Step 1. Else go to Step 2. #

The following result follows readily from the results in the appendix and the assumptions made in the beginning of the paper.

Theorem 4.

Every accumulation point y^* of the sequence $\{y_q\}$ constructed by algorithm II satisfies $y^* \in F$ and $\theta_0(y^*) = 0$. #

5. Conclusion.

It is hoped that the algorithms presented in this paper will be of use for problems with functional inequality constraints. One example of this class of problems was mentioned in the introduction: the design of linear multivariable systems when stability and performance constraints are expressed in terms of frequency response functionals. Another example is the design of buildings is withstand earthquakes [10]; here the cost would be the weight of the building, and a typical constraint the maximum deviation of a floor during an earthquake.

For the problem of designing linear multivariable systems, a set C_1 has to be chosen so that $\det(T(i\omega)) \notin B_1 \quad \forall\, \omega \in \Omega$ is sufficient to ensure "relative", as well as asymptotic, stability i.e. the locus $\det(T(i\Omega))$ $(\det(T(i\Omega)) \triangleq \{\det(T(s))s = i\omega,\ \omega \in \Omega\})$ should maintain a reasonable distance from the origin of the complex plane . It is probably worth mentioning that a prescribed degree of "relative stability" may be achieved by considering rather the locus $\det(T((\alpha + i)\Omega))$, where $\alpha < 0$ is suitably chosen, in place of $\det(T(i\Omega))$; if this is done C_1 may be chosen merely to prevent encirclement by $\det(T((\alpha + i)\Omega))$ of the origin.

The algorithm has a feature which should be of interest also for conventional optimization problems ($m = 0$), i.e. without functional constraints, namely that the initial points z_0 need not be feasible. Algorithm 1 automatically drives z into the feasible region and then maintains feasibility; i.e. unlike conventional feasible directions algorithms, it solves the "phase I" and "phase II" problems simultaneously.

REFERENCES

1. D.L. Kleinman and M. Athans, "The design of suboptimal linear time-varying systems," IEEE Trans Automatic Control, Vol. AC-13, pp 150-159, 1968.
2. E.J. Davison, "The systematic design of control systems for large multivariable linear time-invariant systems". Automatica, Vol. 9, pp 441-452, 1973.
3. H.H. Rosenbrock, "Progress in the design of multivariable control systems," Trans. Inst. Measurement Control, Vol. 4, pp 9-11, 1971.
4. A.G.J. MacFarlane, "A survey of some recent results in linear multivariable feedback theory," Automatica, Vol. 8, pp 455-492, 1972.
5. D.Q. Mayne, "The design of linear multivariable systems," to appear, Automatica, 1974.
6. G. Zoutendijk, "Method of feasible directions," Elsevier, Amsterdam, 1970.
7. V.F. Demyanov, "On the solution of certain min max problems," I, Kybernetika, Vol. 2, 1966.
8. E. Polak, "Computational methods in optimization: a unified approach," Academic Press, New York, 1971.
9. R. Klessig, "A general theory of convergence for constrained optimization algorithms that use anti-zigzagging provisions," Bell Telephone Laboratories Report, Holmdel N.J. 07733, August 1972.
10. D. Ray, K.S. Pister, A.K. Chopra, "Optimum elastic design of earthquake resistant shear buildings," Report No. EERC 74-3, Earthquake Engineering Research Center, University of California, Berkeley, 1974.

Appendix

Proposition 1. If $z \in F$ is optimal for P1, then $\theta_0(z) = \bar{\theta}_0(z) = \tilde{\theta}(z) = 0$.

Proof. (i) It follows from assumption A4, (7), (10) that, $\forall z \in R^m$, $\forall j \in J_m$, $\Omega_0^j(z) = \bar{\Omega}_0^j(z)$, and hence from (11), (13), that $\forall j \in J_m$, $Df_0^j = \overline{Df}_0^j$. It now follows from (14), (16), (17) that $\forall z \in R^n$, $\bar{\theta}_0(z) = \tilde{\theta}_0(z) = \theta_0(z)$.

(ii) If $z \in F$ is optimal, and $\theta_0(z) \leq \delta < 0$, then $\psi_o(z) = 0$, and $\exists\, h \in S$ such that:

$$Df_0^j(z,h) \leq \delta \quad \forall j \in J_m$$

$$Dg_0^j(z,h) \leq \delta \quad \forall j \in J_p.$$

It follows from the continuity of ∇f^o, $\nabla_z \phi^j$, $\forall j \in J_m$ and ∇g^j, $\forall j \in J_p$, that $\exists\, \lambda > 0$ such that $z + \lambda h \in F$ and $f^o(z + \lambda h) - f^o(z) < 0$, contradicting the optimality of z. #.

Fact 1. (i) $\forall z \in R^n$, $\bar{\theta}_0(z) = \tilde{\theta}_0(z) = \theta_0(z)$

(ii) $\forall z \in R^n$, $\forall \varepsilon \geq 0$, $\bar{\theta}_\varepsilon(z) \leq \theta_\varepsilon(z) \leq \tilde{\theta}_\varepsilon(z)$

(iii) $\forall z \in R^n$, if $\varepsilon_1 \leq \varepsilon_2$ then $\theta_{\varepsilon_1}(z) \leq \theta_{\varepsilon_2}(z)$.

Proof. (i) is proven in Proposition 1.

(ii) follows from the fact that $\forall z \in R^n$, $\forall \varepsilon \geq 0$, $\forall j \in J_m$, $\bar{\Omega}_\varepsilon^j(z) \subset \Omega_\varepsilon^j(z)$.

(iii) follows from the fact that $\varepsilon_1 \leq \varepsilon_2$ implies that

$$\Omega_{\varepsilon_1}^j(z) \subset \Omega_{\varepsilon_2}^j(z) \qquad \#$$

Fact 2. $\forall z \in R^n$, $\forall j \notin J_\varepsilon^f(z)$ $(\forall j \notin J_\varepsilon^g(z))$ $\exists\, \gamma > 0$ such that $j \notin J_\varepsilon^f(z')$ $(j \notin J_\varepsilon^f(z'))$ $\forall z' \in N_\gamma(z)$.

Proof. If $j \notin J_\varepsilon^f(z)$, then $\exists\, \delta > 0$ such that $\delta^j(z) - \psi_o(z) \leq -\varepsilon - \delta$.

The result follows directly from the continuity of f^j, ψ_o. The proof for the second case is similar. #.

Fact 3. $\forall z \in R^n$, $\forall\ \varepsilon \geq 0$, $\forall j \in J^f_\varepsilon(z)$, $\forall\ \delta > 0\ \exists\ \gamma > 0$ such that $\Omega^j_\varepsilon(z') \subset N_\delta(\Omega^j_\varepsilon(z))\ \forall z' \in N_\gamma(z)$, where N_δ, N_γ, denote, respectively, δ, γ neighborhoods of $\Omega^j_\varepsilon(z)$, and z.

Proof. First, it follows from (5) and (6) that ψ_o: $R^n \to R$ is continuous. If $N \triangleq N_\delta(\Omega^j_\varepsilon(z)) = \Omega$, the result is trivially true. If $N \neq \Omega$, then $N^c \subset (\Omega^j_\varepsilon(z))^c$, and $\exists\ \eta > 0$ such that $\forall j \in J^f_\varepsilon(z)$ $\sup_{\omega \in N^c} [\phi^j(z,\omega) - \psi_o(z)] \leq -\varepsilon - \eta$. (for, if not, $\sup_{\omega \in N^c} [\phi^j(z,\omega) - \psi_o(z)] \geq -\varepsilon$, and $\exists$ a sequence $\{\omega_\ell\}$ such that $\omega_\ell \in N^c$ and $[\phi^j(z,\omega_\ell) - \psi_o(z)] < -\varepsilon$ $\forall\ \ell \geq 0$ and $\omega_\ell \to \omega^* \in \overline{N}^c$ where ω^* satisfies $[\phi^j(z,\omega^*) - \psi_o(z)] \geq -\varepsilon$; however this implies that $\omega^* \in \Omega^j_\varepsilon(z)$, i.e. $\Omega^j_\varepsilon(z) \cap \overline{N}^c \neq \phi$, contradicting the fact that $\Omega^j_\varepsilon(z) \subset \text{int } N$). Hence $\exists\ \gamma > 0$ such that:

$$\sup_{\omega \in N^c} [\phi^j(z',\omega) - \psi_o(z')] = \max_{\omega \in \overline{N}^c} [\phi^j(z',\omega) - \psi_o(z')] \leq -\varepsilon - \eta/2 < -\varepsilon$$

$\forall\ z' \in N_\gamma(z)$. Thus $\Omega^j_\varepsilon(z') \subset N$, $\forall\ z' \in N_\gamma(z)$. #

Fact 4. $\forall\ z \in R^n$, $\forall\ j \in J^f_0(z)$, $\Omega^j_\varepsilon(z) \to \Omega^j_0(z)$, with respect to inclusion, as $\varepsilon \to 0$.

Proof. The result follows from assumption A4, and the continuity of $\phi^j - \psi_o$, $\forall\ j \in J_m$. #

Corrollary. $\forall\ z \in R^n$ such that $\theta_0(z) < 0$, $\exists\ \varepsilon > 0$ such that $\theta_{\varepsilon'}(z) < 0$, $\forall\ \varepsilon' \in [0,\varepsilon]$. #

Fact 5. $\forall\ z \in R^n$, $\forall\ \delta > 0$, $\forall\ \gamma > 0$, $\exists\ \varepsilon > 0$ such that $\mu(\Omega^j_{\varepsilon'}(z')) \leq \delta$, $\forall\ z' \in N_\gamma(z)$, $\forall\ \varepsilon' \in [0,\varepsilon]$, $\forall\ j \in J_m$, where $\mu(\cdot)$ denotes measure (total length).

$\underline{\text{Proof}}$. If $j \notin J^f_{\varepsilon'}(z')$, $\mu(\Omega^j_{\varepsilon'}(z')) = 0$. Let $j \in J^f_{\varepsilon'}(z')$ and assume that Fact 5 is false, i.e. $\sup_{z' \in N_\gamma(z)} \mu(\Omega^j_{\varepsilon'}(z')) > \delta$, $\forall\, \varepsilon' > 0$. Hence $\exists$ sequences $\{z_i\}$, with $z_i \to z^*$, and $\{\varepsilon_i\}$, with $\varepsilon_i \to 0$, such that $z_i \in N_\gamma(z)$ and $\mu(\Omega'_{\varepsilon_i}(z_i)) \geq \delta/2$, $\forall\, i \geq 0$. But, by Fact 4, $\exists\, \varepsilon^* > 0$ such that $\mu(\Omega^j_{\varepsilon^*}(z^*)) \leq \delta/8$, and $\exists\, i^*$ such that $\varepsilon_i < \varepsilon^*$ $\forall i \geq i^*$, and (by Fact 3):

$$\Omega^j_{\varepsilon_i}(z_i) \subset \Omega^j_{\varepsilon^*}(z_i) \subset N_{\delta/8}(\Omega^j_{\varepsilon^*}(z^*))$$

$\forall\, i \geq i^*$. Since $\mu(\Omega^j_{\varepsilon^*}(z^*)) \leq \delta/8$ we have $\mu(\Omega^j_{\varepsilon_i}(z_i)) \leq \delta/4$, $\forall\, i \geq i^*$. But this contradicts $\mu(\Omega_{\varepsilon_i}(z_i)) \geq \delta/2$, $\forall\, i \geq 0$. #.

$\underline{\text{Fact 6}}$. $\forall\, z \in R^n$, $\forall\, h \in S$, $\forall\, \delta > 0$, $\exists\, \gamma > 0$, $\exists\, \varepsilon > 0$ such that:

$$Df^j_{\varepsilon'}(z',h) \leq \overline{D}f^j_{\varepsilon'}(z',h) + \delta, \ \forall\, z' \in N_\gamma(z), \quad \forall\, \varepsilon' \in [0,\varepsilon], \quad \forall\, j \in J_m \quad \text{(A1)}.$$

$\underline{\text{Proof}}$. If $j \notin J^f_\varepsilon(z')$, then $j \notin J^f_{\varepsilon'}(z')$, for $\varepsilon' \in [0,\varepsilon]$, and $Df^j_{\varepsilon'}(z',h) = -\infty$ and (A1) is trivially satisfied. From the continuity of $\{\nabla_z \phi^j,\ j \in J_m\}$, $\forall\, z \in R^n$, $\forall\, \delta > 0$, $\exists\, \delta_1 > 0$, $\exists\, \gamma_1 > 0$ such that:

$$\|\nabla_z \phi^j(z',\omega') - \nabla_z \phi^j(z,\omega)\| \leq \delta/\sqrt{n}$$

$\forall\, z' \in N_{\gamma_1}(z)$, $\forall\, \omega, \omega' \in \Omega$ such that $|\omega'-\omega| \leq \delta_1$, $\forall\, j \in J^f_\varepsilon(z)$. But from Fact 5, $\exists\, \varepsilon > 0$, $\exists\, \gamma \in (0,\gamma_1]$ such that $\mu(\Omega^j_{\varepsilon'}(z')) \leq \delta_1$, $\forall\, z' \in N_\gamma(z)$, $\forall\, \varepsilon' \in [0,\varepsilon]$, $\forall\, j \in J^f_\varepsilon(z)$. Hence $\|\nabla_z \phi^j(z',\omega') - \nabla_2 \phi^j(z,\omega)\| \leq \delta/\sqrt{n}$, $\forall\, z' \in N_\gamma(z)$, $\forall\, \omega, \omega'$ in the $\underline{\text{same}}$ subinterval of $\Omega^j_{\varepsilon'}(z')$, $\forall\, \varepsilon' \in [0,\varepsilon]$, $\forall\, j \in J^f_\varepsilon(z)$. The desired result follows from the definitions of Df^j_ε, $\overline{D}f^j_\varepsilon$, and the fact that $\|h\| \leq \sqrt{n}$ $\forall\, h \in S$. #

$\underline{\text{Fact 7}}$. $\forall\, z \in R^n$, $\forall\, \delta > 0$, $\exists\, \gamma > 0$, $\exists\, \varepsilon > 0$ such that:

$$\tilde{\theta}_{\varepsilon'}(z') \leq \overline{\theta}_{\varepsilon'}(z') + \delta, \quad \forall\, \varepsilon' \in [0,\varepsilon], \quad \forall\, z' \in N_\gamma(z).$$

Proof. The result follows from the definitions of $\bar{\theta}_{\varepsilon'}$, $\tilde{\theta}_{\varepsilon'}$, the continuity of ψ_o and ∇f^o, and Fact 6. #

Fact 8. $\forall\, z \in R^n$ such that $\theta_0(z) < 0$, $\exists\, \gamma > 0$, $\exists\, \varepsilon > 0$ such that:

$$\theta_{\varepsilon'}(z') \le \frac{1}{2}\,\theta_\varepsilon(z) < 0 \quad \forall\, z' \in N_\gamma(z), \quad \forall\, \varepsilon' \in [0,\varepsilon]$$

Proof. By Fact 4, $\forall\, z \in R^n$ such that $\theta_0(z) < 0$, $\exists\, \varepsilon > 0$ such that $\theta_{\varepsilon'}(z) < 0 \;\; \forall\, \varepsilon' \in [0,\varepsilon]$. Now:

$$\theta_\varepsilon(z) \triangleq \min_{h \in S} \max\{-\psi_o(z) + \langle \nabla f^o(z), h\rangle;\ \langle \nabla_z \phi^j(z,\omega), h\rangle,$$
$$\omega \in \Omega^j_\varepsilon(z),\ j \in J^f_\varepsilon(z);\ Dg^j_\varepsilon(z),\ j \in J_p\}$$

Let $\delta > 0$ be such that $\hat{\theta}_\varepsilon(z) \le 3/4\, \theta_\varepsilon(z)$, where $\hat{\theta}_\varepsilon : R^n \to R$ is defined by:

$$\hat{\theta}_\varepsilon(z') \triangleq \min_{h \in S} \max\{-\psi_o(z') + \langle \nabla f^o(z'), h\rangle;\ \langle \nabla_z \phi^j(z',\omega), h\rangle,$$
$$\omega \in N_\delta(\Omega^j_\varepsilon(z)),\ j \in J^f_\varepsilon(z');\ Dg^j_\varepsilon(z'),\ j \in J_p\}$$

By Fact 3, $\forall\, z \in R^n$, $\exists\, \gamma_1 > 0$ such that $\Omega^j_\varepsilon(z') \subset N_\delta(\Omega^j_\varepsilon(z))$, $\forall\, j \in J^f_\varepsilon(z')$,

So that $\theta_\varepsilon(z') \le \hat{\theta}_\varepsilon(z') \;\; \forall\, z' \in N_{\gamma_1}(z)$.

Let $\gamma \in (0,\gamma_1]$ be such that $\hat{\theta}_\varepsilon(z') \le \frac{2}{3}\,\hat{\theta}_\varepsilon(z) \;\; \forall\, z' \in N_\gamma(z)$. Such a γ exists because of the continuity of ψ_o, ∇f^o, $\nabla_z \phi^j \;\; \forall\, j \in J_m$, $\nabla g^j \;\; \forall\, j \in J_p$.

Hence, $\theta_{\varepsilon'}(z') \le \hat{\theta}_{\varepsilon'}(z') \le \hat{\theta}_\varepsilon(z') \le \frac{2}{3}\,\hat{\theta}_\varepsilon(z) \le \frac{1}{2}\,\theta_\varepsilon(z) < 0 \quad \forall\, z' \in N_\gamma(z)$, $\forall\, \varepsilon' \in [0,\varepsilon]$. #.

Proposition 2. $\forall\, z \in R^n$ such that $\theta_0(z) < 0$, $\exists\, \gamma > 0\ \exists\, \varepsilon > 0$ such that $\bar{\varepsilon}(z') \ge \varepsilon \;\; \forall\, z' \in N_\gamma(z)$, where $\bar{\varepsilon}(z)$ is the ε generated by algorithm I in Step 5 (with $z_1 = z$).

Proof. By Fact 8, $\forall\, z \in R^n$ such that $\theta_0(z) < 0$, $\exists\, \gamma_1 > 0$, $\exists\, \varepsilon_1 > 0$ such that: $\theta_{\varepsilon'}(z') \le \frac{1}{2}\,\theta_\varepsilon(z) \triangleq \hat{\theta} < 0 \;\; \forall\, z' \in N_{\gamma_1}(z)$, $\forall\, \varepsilon' \in [0,\varepsilon_1]$.

Hence, by Fact 1, $\bar{\theta}_{\varepsilon'}(z') \leq \hat{\theta}$ $\forall z' \in N_{\gamma_1}(z)$, $\forall \varepsilon' \in [0,\varepsilon_1]$.

Let $\delta = \min\{\varepsilon_1, -\hat{\theta}/2\}$. By Fact 7, $\exists\ \gamma \in (0,\gamma_1]$, $\exists\ \varepsilon \in (0,\delta]$ such that

$\tilde{\theta}_{\varepsilon'}(z') \leq \bar{\theta}_{\varepsilon'}(z') + \delta$, i.e. $\tilde{\theta}_{\varepsilon'}(z') \leq \hat{\theta} + \delta$ $\forall z' \in N_\gamma(z)$, $\forall \varepsilon' \in [0,\varepsilon]$.

Now $\hat{\theta} \leq 2\delta$ and $\varepsilon \leq \delta \leq \varepsilon_1$ so that:

$$\bar{\theta}_\varepsilon(z') \leq -2\delta \leq -2\varepsilon \tag{A2}$$

$$\tilde{\theta}_\varepsilon(z') \leq -\delta \leq -\varepsilon \tag{A3}$$

$\forall z' \in N_\gamma(z)$. Since $\bar{\varepsilon}(z')$ is the largest number of the form $\varepsilon_0/2^\ell$, ℓ a positive integer, satisfying (A2) and (A3), the desired result is proven. #

<u>Proposition 3</u>. $\forall z \in R^n$ such that $\theta_0(z) < 0$, $\exists\ \gamma > 0$, $\exists\ \varepsilon > 0$ such that:

$f^o(z'') - f^o(z') \leq -\varepsilon$, $\forall z' \in N_\gamma(z) \cap F$, $\forall z'' \in A(z')$, and $\psi(z'')-\psi(z') \leq -\varepsilon$,

$\psi(z'') - \psi(z') \leq -\varepsilon$, $\forall z' \in N_\gamma(z) \cap F^c$, $\forall z'' \in A(z')$ (where A: $R^n \to 2^{R^n}$

is defined by Steps 1 to 6 of the conceptual algorithm with "go to Step 1" in Step 6) .

<u>Proof</u>.

By Proposition 2, $\forall z \in R^n$ such that $\theta_0(z) < 0$, $\exists\ \gamma_1 > 0$, $\exists\ \varepsilon_1 > 0$ such that $\bar{\varepsilon}(z') \geq \varepsilon_1$, and $\tilde{\theta}_{\bar{\varepsilon}(z')} \leq -\varepsilon_1$ $\forall z' \in N_{\gamma_1}(z)$.

(a) (upper bound on $\psi(z'')$, $z'' \in A(z')$).

Because of the compactness of the various sets involved, $\exists\ \lambda_1 \in (0,1]$ such that:

$$|\phi^j(z' + \lambda h,\omega) - \phi^j(z',\omega)| \leq \varepsilon_1/2 \tag{A4}$$

$\forall z' \in N_{\gamma_1}(z)$, $\forall \omega \in \Omega$, $\forall h \in S$, $\forall j \in J_m$, $\forall \lambda \in [0,\lambda_1]$.

Let $Z \triangleq \{(j,\omega) | (j \notin J^f_{\bar{\varepsilon}(z')}(z'), \omega \in \Omega), (j \in J^f_{\bar{\varepsilon}(z')}(z'), \omega \notin \Omega^j_{\bar{\varepsilon}(z')}(z'))\}$;

then $\phi^j(z',\omega) - \psi(z') < -\varepsilon_1$, $\forall\ (j,\omega) \in Z$. It follows from (A4) that:

$$\phi^j(z' + \lambda\bar{h}(z'),\omega) - \psi(z') \le -\varepsilon_1/2 \tag{A5}$$

$\forall\ z' \in N_{\gamma_1}(z)$, for all $(j,\omega) \in Z$ where $\bar{h}(\cdot)$ denotes $h_{\bar{\varepsilon}(\cdot)}(\cdot)$.

Also:

$$\phi^j(z' + \lambda\bar{h}(z'),\omega) - \phi^j(z',\omega) \le \lambda[\langle \nabla_z\phi^j(z',\omega),\bar{h}(z')\rangle$$

$$+ \sqrt{n}\int_0^1 \|\nabla_z\phi^j(z' + s\lambda\bar{h}(z'),\omega) - \nabla_z\phi^j(z',\omega)\|ds]$$

since $\|h\| \le \sqrt{n}\ \ \forall\ h \in S$. Let Z^c denote the complement of Z in $J_m \times \Omega$, i.e. $Z^c = \{(j,\omega) | (j \in J^f_{\bar{\varepsilon}(z')}(z'),\ \omega \in \Omega^j_{\bar{\varepsilon}(z')}(z'))\}$. Now $\langle \nabla_z\phi^j(z',\omega),\ \bar{h}(z')\rangle \le \tilde{\theta}\bar{\varepsilon}(z') \le -\varepsilon_1\ \ \forall\ (j,\omega) \in Z^c$; it follows from the continuity of $\nabla_z\phi^j$, that $\exists\ \lambda_2 \in (0,\lambda_1]$ such that

$$\phi^j(z' + \lambda\bar{h}(z'),\omega) - \psi(z') \le -\lambda\ \alpha\varepsilon_1 \tag{A6}$$

$\forall\ (j,\omega) \in Z^c$, $\forall\ z' \in N_{\gamma_1}(z)$, $\forall\lambda \in [0,\lambda_2]$.

Combining A5, A6 we obtain:

$$f^j(z' + \lambda\bar{h}(z')) - \psi(z') \le -\lambda\alpha\varepsilon_1 \tag{A7}$$

$\forall\ z' \in N_{\gamma_1}(z)$, $\forall\ \lambda \in (0,\lambda^f]$, $\forall\ j \in J_m$, where $\lambda^f \in (0,\lambda_2]$ satisfies $\lambda^f\alpha\varepsilon_1 \le \varepsilon_1/2$.

Similarly, it can be shown that $\exists\ \lambda^g > 0$ such that:

$$g^j(z' + \lambda\bar{h}(z')) - \psi(z') \le -\lambda\alpha\varepsilon_1 \tag{A8}$$

$\forall\ z' \in N_\gamma(z)$, $\forall\ \lambda \in (0,\lambda^g]$, $\forall\ j \in J_p$.

Let $\lambda_3 = \beta^{k_1} > 0$, where k_1 is the smallest non-negative integer, satisfying $\lambda_3 \le \max\{\lambda^f,\lambda^g\}$. From (A7) and (A8) and the definition of ψ we obtain:

$$\psi(z' + \lambda_3\bar{h}(z')) - \psi(z') \leq - \lambda_3\alpha\varepsilon_1 \qquad \text{(A9)}$$

$\forall\ z' \in N_{\gamma_1}(z)$.

Now, for $z_i = z'$, Step 5 of the algorithm chooses the smallest non-negative integer $\ell(z')$ and a corresponding $\bar{\lambda}(z') = \beta^{\ell(z')}$ such that:

$$\psi(z'') - \psi(z') = \psi(z' + \bar{\lambda}(z')\bar{h}(z')) - \psi(z')$$
$$\leq - \bar{\lambda}(z')\alpha\varepsilon(z') \leq - \bar{\lambda}(z')\ \alpha\ \varepsilon_1, \qquad \text{(A10)}$$

with $z'' = z' + \bar{\lambda}(z')\bar{h}(z')$.

Comparing (A9) with (A10) we obtain

$$\bar{\lambda}(z') \geq \lambda_3$$

$\forall\ z' \in N_{\gamma_1}(z) \cap F^c$. Hence:

$$\psi(z'') - \psi(z') \leq - \beta\lambda_3\alpha\varepsilon_1 \qquad \text{(A11)}$$

$\forall\ z' \in N_{\gamma_1}(z)$, $\forall\ z'' \in A(z')$

(b) (upper bound on $f^o(z'')$ $z'' \in A(z')$). Let $z' \in F$ ($\psi_o(z') = 0$, $\psi(z') \leq 0$).

Since $\psi_o(z') = 0$, it follows that:

$$\langle \nabla f^o(z'), \bar{h}(z') \rangle \leq \theta_{\bar{\varepsilon}(z')}(z') \leq - \varepsilon_1,\ \forall z' \in N_{\gamma_1}(z).$$

Hence $\exists\ \lambda_4 > 0$ such that:

$$f^o(z' + \lambda\bar{h}(z')) - f^o(z') \leq - \lambda\alpha\varepsilon_1$$

$\forall\ \lambda \in (0,\lambda_4]$, $\forall\ z' \in N_{\gamma_1}(z) \cap F$.

It follows from the discussion in (a) above (which is not restricted to the case $z' \in F^c$) and, in particular, from (A9), since $\psi(z') \leq 0$, that $\exists\ \lambda_5 > 0$ such that:

$$\psi(z' + \lambda\bar{h}(z')) \leq 0 \qquad \text{(A12)}$$

$\forall\, z' \in N_{\gamma_1}(z)$, $\forall\, \lambda \in (0,\lambda_5]$. Let $\lambda_6 = \beta^{k_2} > 0$ be such that $\lambda_6 \leq \max\{\lambda_4,\lambda_5\}$. If $\ell(z')$ is generated by the algorithm in Step 5, then $\bar{\lambda}(z') = \beta^{\ell(z')}$ satisfies $\bar{\lambda}(z') \geq \lambda_6$ $\forall\, z' \in N_{\gamma_1}(z) \cap F$. Hence:

$$f^{o}(z'') - f_o(z') \leq - \lambda_6\alpha\varepsilon_1 \qquad \text{(A13)}$$

$\forall\, z' \in N_{\gamma_1}(z)$, $\forall\, z'' \in A(z')$.

(A11) and (A13) establish Proposition 3. #