ON THE SOLUTION OF SINGULAR VALUE INEQUALITIES OVER A CONTINUUM OF FREQUENCIES

by

E. Polak and D. Q. Mayne

Memorandum No. UCB/ERL M80/8

9 January 1980

ELECTRONICS RESEARCH LABORATORY

College of Engineering
University of California, Berkeley
94720

ON THE SOLUTION OF SINGULAR VALUE INEQUALITIES OVER A CONTINUUM OF FREQUENCIES [†]

E. Polak[††] and D. Q. Mayne[†††]

ABSTRACT

We present an algorithm for solving singular value inequalities over a continuum of frequencies. The algorithm is in two parts: a master algorithm which constructs an infinite sequence of finite sets of inequalities and a nondifferentiable optimization subalgorithm which solves these finite sets of inequalities.

[†] Research supported by United Kingdom Science Research Council and the National Science Foundation ECS-79-13148.

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

[†††] Department of Computing and Control, Imperial College, London, SW72BZ, England

1. Introduction

We are beginning to witness the realization that important aspects of control system design can be expressed in terms of inequalities [6,7,8] and, in particular, in terms of inequalities involving the singular values of transfer function matrices, such as the compensator-plant transfer function matrix $G(x,\omega)$[1], with $x \in \mathbb{R}^k$ denoting the design parameters [1,2,3,4,5,23]. For example, the requirement that the closed loop system remain stable in the face of additive or multiplicative perturbations of $G(x,\omega)$, can be expressed in terms of singular value inequalities [2,3,5]. Another example occurs in singular perturbation problems which arise when a high order system is modeled by a low order model. The stability of the high order system follows from that of the low order system if certain singular value inequalities are satisfied [5]. Again, in [4] we find that the requirement for low sensitivity, in a linear control system to additive output noise and parameter variations can be expressed as a singular value inequality. Generally, these inequalities have the form

$$\ell(\omega) \leq \underline{\sigma}(x,\omega) \leq \bar{\sigma}(x,\omega) \leq u(\omega) \qquad (1.1)$$

$$\text{for all } \omega \in [\omega_0,\omega_f],$$

where $\underline{\sigma}$ and $\bar{\sigma}$ are the smallest and largest singular values of an appropriate transfer function matrix $G(x,\omega)$, and $\ell(\omega)$, $u(\omega)$ are continuous, real valued functions. Unfortunately, before this work, there were no known algorithms for solving such inequalities. The reason for this is that the singular values of $G(x,\omega)$ are not continuously differentiable. In fact, they are not even differentiable at multiplicity points.

---

[1] $G(x,\omega)$ is a complex valued matrix.

In general, one may wish to select a compensator vector x which not only satisfies (1.1) but a number of other specification inequalities as well (see e.g. [7,8,9]). Furthermore, it may be desirable to minimize some cost while doing this. However, the main stumbling block at present is (1.1), and once we learn how to solve (1.1), the extension of the algorithm to include additional sets of nondifferentiable or differentiable inequalities and a cost for optimization is quite straightforward. Because of this, to avoid making the exposition more complex then is absolutely necessary, we restrict ourselves to presenting an algorithm for solving the infinite set of nondifferentiable inequalities (1.1). The algorithm consists of two parts: a master outer approximation algorithm which substitutes for (1.1) an infinite sequence of finite sets of non-differentiable inequalities, defined by a few points $\omega \in [\omega_o, \omega_f]$ and a nondifferentiable optimization subalgorithm which solves these inequalities. As can be seen from [9], the extension of our algorithm to the problem of minimizing a cost subject to (1.1) and additional inequalities merely requires a fairly straighforward extension of our subalgorithm from the form of unconstrained optimization to that of constrained optimization.

2. Continuity and Differentiability of Singular Values

Let $G(x,\omega)$ be an $m \times m$ complex valued transfer function matrix, with $x \in \mathbb{R}^k$ the design, or compensator, parameter vector to be selected, and ω the frequency variable. The following assumption will hold in most design situations:

Assumption 2.1: There exists an open set $X \subset \mathbb{R}^k$ such that $G : X \times \mathbb{R}^l \to C^{m\times m}$ is analytic (componentwise). ¤

The relevant properties of the singular values $\sigma_i(x,\omega)$, $i = 1,2,..,m$ of $G(x,\omega)$ are easiest to establish by considering first the eigenvalues $y^i(x,\omega)$, $i = 1,2,..,m$, of the symmetric matrix

$$Q(x,\omega) \triangleq G^*(x,\omega)\, G(x,\omega) \tag{2.1}$$

We recall that the singular values are defined by

$$\sigma^i(x,\omega) \triangleq \sqrt{y^i(x,\omega)},\ i = 1,2,..,m. \tag{2.2}$$

Assumption 2.2: The pairs $(\bar{x}, \bar{\omega})$, $\bar{x} \in X$, $\bar{\omega} \in \mathbb{R}^l$, at which $Q(\bar{x},\bar{\omega})$ has multiple eigenvalues are isolated, i.e., for every such pair, there exists an open neighborhood $N_{\bar{x},\bar{\omega}}$ of $(\bar{x},\bar{\omega})$ in $X \times \mathbb{R}^l$ such that $(\bar{x},\bar{\omega})$ is the only pair in $X \times \mathbb{R}^l$ at which Q has multiple eigenvalues. ¤

Referring to [10] Ch. 6 Sec. 2, we see that in view of Assumption 2.1, Assumption 2.2 will be satisfied, except for some pathological cases, essentially always.

Proposition 2.1: Suppose that Assumption 2.1 holds. Then the eigenvalues $y_i(x,\omega)$, $i = 1,2,..,m$, of $Q(x,\omega)$ are locally Lipschitz continuous on $X \times \mathbb{R}^l$.

Proof: Let $x, x' \in X$, $\omega, \omega' \in \mathbb{R}^l$, and let

$$\Delta Q \triangleq Q(x',\omega') - Q(x,\omega) \tag{2.3}$$

Then ΔQ is a symmetric matrix which can be decomposed into the form

$$\Delta Q = \sum P_i \tag{2.4}$$

where P_i is a symmetric matrix with at most two nonzero elements that are equal to some element $\Delta q_{k\ell}$ of ΔQ. It now follows from Wilkinson [11], p. 41 and Parlett [12] Ch. 10 that

$$|y^i(x,\omega) - y^i(x',\omega')| \leq \sum_{k,\ell} m_{k\ell} |\Delta q_{k\ell}| \qquad \text{for } i = 1,2,..,m, \tag{2.5}$$

with $m_{k\ell} \in [0,1]$. The proposition now follows from the differentiability of $Q(\cdot,\cdot)$. ¤

Consider the matrix $Q(x,\omega)$. Let $h \in \mathbb{R}^n$, $\nu \in \mathbb{R}^1$ be arbitrary. Whenever

$$\delta y^i(x,\omega;h,\nu) \triangleq \frac{\partial}{\partial\lambda} y_i(x+\lambda h,\omega+\lambda\nu)\Big|_{\lambda = 0} \tag{2.6}$$

exists, we shall refer to it as a <u>one dimensional derivative</u>.

<u>Proposition 2.2</u>: Suppose that Assumptions 2.1 and 2.2 are satisfied. Consider the matrix $Q(x,\omega)$. Then the bi-directional derivatives of its eigenvalues, $\delta y^i(x,\omega;h,\nu)$, $i = 1,2,\ldots,m$, exist for all $(x,\omega) \in X \times \mathbb{R}^1$, $(h,\nu) \in \mathbb{R}^k \times \mathbb{R}^1$, and are continuous in (x,ω) at all $(x,\omega) \in X \times \mathbb{R}^1$ at which $Q(x,\omega)$ has distinct eigenvalues only.

<u>Proof</u>: Let $(h,\nu) \in \mathbb{R}^k \times \mathbb{R}^1$ be arbitrary and consider the matrix $Q(x+\lambda h,\omega+\lambda\nu)$ for $\lambda \in \in[-\bar{\lambda},\bar{\lambda}]$ with $\bar{\lambda} > 0$ such that $x + \lambda h \in X$ for all $\lambda \in [-\bar{\lambda},\bar{\lambda}]$. Then the $y^i(\lambda) \triangleq y^i(x+\lambda h,\omega+\lambda\nu)$, $i = 1,2,..,m$, are the roots

of the characteristic polynomial

$$f(\lambda,y) = y^m + a_{m-1}(\lambda)y^{n-1} + \ldots + a_0(\lambda) \tag{2.7}$$

Where the $a_k(\lambda)$ are analytic functions of λ. It now follows from algebraic function theory (see e.g. [10]), that if $y^i(0)$ is a simple root, then there exists a $\lambda_i \in (0,\bar{\lambda}]$ such that in all $\lambda \in [-\lambda_i,\lambda_i]$

$$y^i(\lambda) = y_i^i(0) + \alpha_i^1\lambda + \alpha_i^2\lambda^2 + \ldots. \tag{2.7a}$$

and hence that $\partial y^i(0)/\partial\lambda = \alpha_i^1$, and consequently, the one dimensional derivative $\delta y^i(x,\omega;h,\nu) = \frac{\partial}{\partial\lambda} y^i(x+\lambda h,\omega+\lambda\nu)\big|_{\lambda=0}$ exists. Next since

$$f(\lambda,y^i(\lambda)) \equiv 0, \forall \lambda \in [-\lambda_i,\lambda_i], \tag{2.8}$$

we find that for $\lambda \in (-\lambda_i,\lambda_i)$

$$\frac{\partial f(\lambda,y^i(\lambda))}{\partial\lambda} + \frac{\partial f}{\partial y}(\lambda,y^i(\lambda))\,\frac{\partial y^i(\lambda)}{\partial\lambda} \equiv 0\,. \tag{2.9}$$

Because $y^i(0)$ is a simple eigenvalue, it follows that $\frac{\partial f}{\partial y}(0,y^i(0)) \neq 0$ and hence, from (2.9), $\partial y^i(0)/\partial\lambda$ is seen to be a continuous function of (x,ω), which also proves the continuity of $\delta y^i(x,\omega;h,\nu)$.

Now suppose that $y^i(0)$ is an eigenvalue of $Q(x,\omega)$ of multiplicity $\mu > 1$, i.e. suppose that $y^i(0) = y^{i+1}(0) = \ldots = y^{i+\mu-1}(0)$. It then follows from algebraic function theory that for some integer $\ell \in [1,\mu]$ and all $\lambda \in [-\lambda_i,\lambda_i]$, with $\lambda_i \in (0,\bar{\lambda}]$,

$$y^i(\lambda) = \sum_{k=0} \alpha_i^k \lambda^{k/\ell} \tag{2.10}$$

Now, if $\ell > 1$, then $y^i(\lambda)$ will be complex for some $\lambda \in (-\lambda_i,\lambda_i)$. But $y^i(\lambda)$ is an eigenvalue of a positive semi-definite matrix and

hence must be real. We therefore conclude that $\ell = 1$ and hence that $\frac{\partial y^i(0)}{\partial \lambda}$ exists. Consequently, $\delta y^i(x,\omega;h,\nu)$ also exists. ¤

The fact that the partial derivatives of eigenvalues of a symmetric, positive semi-definite matrix are not necessarily continuous at multiplicity points is best seen from a simple example. Consider the matrix

$$Q(x) \triangleq \begin{bmatrix} x^1 + x^2 & 1 - x^1 \\ 1 - x^1 & x^1 \end{bmatrix} \tag{2.11}$$

Its eigenvalues $y^1(x)$, $y^2(x)$ are the roots of

$$[y - (x^1+x^2)]\,[y - x^1] + (1-x^1)^2 = 0 \tag{2.12}$$

and hence

$$y^1(x) = (x^1 + \frac{x^2}{2}) + \sqrt{(1-x^1)^2 + (\frac{x^2}{2})^2} \tag{2.13a}$$

$$y^2(x) = (x^1 + \frac{x^2}{2}) - \sqrt{(1-x^1)^2 + (\frac{x^2}{2})^2} \tag{2.13b}$$

Setting $\bar{x} = (1,0)$, we get the double eigenvalue $y^1(\bar{x}) = y^2(\bar{x}) = 1$. A straightforward calculation shows that

$$(\partial y^1(\bar{x})/\partial x^1) = 0,\ (\partial y^1(\bar{x})/\partial x^2) = 1.5,$$

$$(\partial y^2(\bar{x})/\partial x^1) = 2,\ (\partial y^2(\bar{x})/\partial x^2) = -0.5.$$

Now, let $x_\varepsilon = (1,\varepsilon)$ with any $\varepsilon > 0$. Then

$$(\partial y^1(x_\varepsilon)/\partial x^1 = 1,\ (\partial y^1(x_\varepsilon)/\partial x^2 = 1.5,$$

$$(\partial y^2(x_\varepsilon)/\partial x^1 = 1,\ (\partial y_2^2(x_\varepsilon)/\partial x^2 = -0.5,$$

and we see that $\partial y^1(x)/\partial x^1$, $\partial y^1(x)/\partial x^1$ are not continuous at $\bar{x}$.

Corollary 2.1: Suppose that Assumptions 2.1 and 2.2 are satisfied that $y^i(\bar{x}, \bar{\omega})$ is a simple eigenvalue of $Q(\bar{x},\bar{\omega})$, with $\bar{x} \in X$. Then there exists an open neighborhood $N_{\bar{x},\bar{\omega}}$ in $X \times R^1$ of $(\bar{x},\bar{\omega})$ such that $\nabla_x y^i(x,\omega)$ exists and is continuous for all $(x,\omega) \in N_{\bar{x},\bar{\omega}}$.

Proof: Since $y^i(\bar{x},\bar{\omega})$ is a simple eigenvalue of $Q(\bar{x},\bar{\omega})$, all the partial derivatives $\partial y^i(x,\omega)/\partial x^j$ exist and are continuous for all $(x,\omega) \in N_{\bar{x},\bar{\omega}}$. Hence, by Theorem 9.16 in [13], $\nabla_x y^i(x,\omega)$ exists and is continuous for all $(x,\omega) \in N_{\bar{x},\bar{\omega}}$. ¤

Referring to Bellman [14] p. 60, we see that when $y^i(x,\omega)$ is a distinct eigenvalue, then $\nabla_x y^i(x,\omega)$ must be given by the formula

$$\nabla_x y^i(x,\omega) = -\sum_{j=1}^{m} v_i^j(x,\omega) \frac{\partial Q_{\cdot j}(x,\omega)}{\partial x} v_i(x,\omega) \tag{2.14}$$

where $Q_{\cdot j}$ is the jth column of Q and $v_i(x,\omega)$ is the unit norm eigenvector of $Q(x,\omega)$, corresponding to $y^i(x,\omega)$. Formula (2.14) is very easy to obtain formally, as follows:

$$[y^i(x,\omega)I - Q(x,\omega)]\, v_i(x,\omega) \equiv 0. \tag{2.15}$$

Hence,

$$[y^i(x,\omega)I - Q(x,\omega)] \frac{\partial v_i(x,\omega)}{\partial x}$$

$$+ v_i(x,\omega)\, \nabla_x y_i(x,\omega)^T$$

$$+ \sum_{j=1}^{m} v_i^j(x,\omega) \frac{\partial Q_{\cdot j}(x,\omega)}{\partial x} = 0. \tag{2.16}$$

Since $\|v_i(x,\omega)\| = 1$ and $v_i(x,\omega)^*$ is also a left eigenvector of $Q(x,\omega)$, (2.14) follows directly from (2.16) after premultiplication of (2.16) by $v_i(x,\omega)^*$. The v_i and σ_i to which they correspond can be computed simultaneously by means of a singular value decomposition algorithm, such as [15].

Since the gradients $\nabla_x y_i(x,\omega)$ do not necessarily exist nor are continuous at pairs (x,ω) at which $y^i(x,\omega)$ is a multiple eigenvalue of $Q(x,\omega)$, an algorithm for solving eigenvalue inequalities has to make use of generalized gradients and their relevant properties [16] which we now summarize.

Definition 2.1. The generalized gradient $\partial_x y^i(x,\omega)$ is defined by

$$\partial_x y^i(x,\omega) \triangleq \text{co} \lim_{i\to\infty} \nabla_x y^i(x+v_i,\omega) \tag{2.17}$$

where $v_i \to 0$ is such that $\nabla_x y^i(x+v_i)$ exists, and so denotes the convex hull. ¤

Proposition 2.3: Suppose that Assumptions 2.1 and 2.2 hold. Then the generalized gradients $\partial_x y^i(x,\omega)$ $i = 1,2,..,m$, of the eigenvalues $y^i(x,\omega)$, $i = 1,2,..,m$, of $Q(x,\omega)$, exist for all $(x,\omega) \in X \times \mathbb{R}^1$. Furthermore, they are compact and upper semicontinuous (i.e. $\{x_j \to \bar{x}$, $g_j \in \partial_x y^i(x_j,\omega)$, $g_i \to \bar{g}\} \Rightarrow \{\bar{g} \in \partial_x y^i(\bar{x},\omega)\}$).

Proof: Since by Proposition 2.1, the y_i are locally Lipschitz continuous, the desired result follows directly from [16]. ¤

Generalized Mean Value Theorem 2.1 [17]. Suppose that Assumptions 2.1 and 2.2 hold. Let $y^i(x,\omega)$ be defined as before. Then for any x', $x'' \in X$, $\omega \in \mathbb{R}^1$, there exists a $\xi = x'+s(x''-x')$, with $s \in (0,1)$, such that

$$y^i(x'',\omega) - y^i(x',\omega) = \langle g_\xi, x'' - x' \rangle \tag{2.18}$$

with $g_\xi \in \partial_x y^i(\xi,\omega)$. ¤

This completes our collection of relevant results about the $y^i(x,\omega)$ and we are now ready to proceed with the construction of an algorithm for solving singular value inequalities.

3. An algorithm for singular value inequalities over a discrete set of frequencies

We begin by developing an algorithm for solving singular value inequalities at a set of frequencies. Thus, in this section, we consider the problem of finding a vector $x \in \mathbb{R}^k$ such that the singular values $\sigma^1(x,\omega)$, $\sigma^2(x,\omega)$, ..., $\sigma^n(x,\omega)$ of a complex valued m x m transfer function matrix $G(x,\omega)$, satisfy inequalities of the form

$$\ell(\omega_j) \le \sigma^i(x,\omega_j) \le u(\omega_j) \quad \text{for } i = 1,2,..,m \qquad j = 1,2,..,J \tag{3.1}$$

Since there is no essential loss in generality and considerable saving in detail, we shall develop first a version of the algorithm for the simplest case of (3.1) viz.

$$\ell(\hat{\omega}) \le \sigma^i(x,\hat{\omega}) \le u(\hat{\omega}), \qquad i = 1,2,\ldots,m, \tag{3.3}$$

where $u(\hat{\omega}) = \infty$ is allowed. Since $\hat{\omega}$ is fixed in (3.3), we may obviously drop it in our notation. Since $y^i(x) = \sigma^i(x)^2$, and $\sigma^i(x) \ge 0$ always, we can substitute for (3.3), the equivalent set

$$\ell^2 \le y^i(x) \le u^2, \qquad i = 1,2,.., m. \tag{3.4}$$

Now[2], for $i = 1,2,\ldots,m$ let

$$f^i(x) = y^i(x) - u^2 \tag{3.5a}$$

and for $i = m+1, \ldots, 2m$, let

$$f^i(x) = \ell^2 - y^{i-m}(x) \tag{3.5b}$$

Since, the $y^i(x)$ are not continuously differentiable, there is no advantage in treating them individually, as in Newton's method or in a method of feasible directions, hence we define

$$\psi(x) \triangleq \max_{i \in \underline{2m}} f^i(x), \tag{3.6}$$

with $\underline{2m} = \{1,2,..,2m\}$, which permits us to replace (3.4) by the statement: find an $x \in X$ such that

$$\psi(x) \leq 0. \tag{3.7}$$

Since the function $\psi(.)$ is obviously not continuously differentiable, it is necessary to introduce a smoothed out, "steepest descent equivalent" rule for computing descent directions. To this end, for any $\varepsilon > 0$, we define[3]

$$\partial_\varepsilon \psi(x) \triangleq \text{co} \bigcup_{\substack{I \in I_\varepsilon(x') \\ x' \in B(x,\varepsilon)}} \partial_\varepsilon f^i(x') \tag{3.8}$$

here, for a given $\alpha > 0$,

$$I_\varepsilon(x') = \{i \in \underline{2m} \mid f^i(x') - \psi(x') \geq -\alpha\varepsilon\} \tag{3.9}$$

---

[2] To extend our treatment to more than one frequency $\hat{\omega}$, all that needs to be done is that more $f^i(\cdot)$ need to be defined.

[3] It can be seen from the proofs to follow that we may have used $I_0(x')$ instead of $I_\varepsilon(x')$ in the definition of $\partial_\varepsilon\psi(x)$. However, the use of $I_\varepsilon(x')$ is bound to result in better computational behavior.

compute h_ε, we could then easily define a convergent algorithm, patterned on a phase I method of feasible directions (see [19]) which would minimize $\psi(x)$ and hence, under mild assumptions, solve (3.7) in a finite number of iterations. Unfortunately, it is impossible to compute h_ε, and therefore we can only hope to construct an algorithm based on the ideas presented so far, if we can construct, efficiently, suitable approximations to h_ε. We shall now describe such a process.

Suppose that $0 \notin \partial_\varepsilon \psi(x)$ and that we have a discrete approximation $Y_k \subset \partial_\varepsilon \psi(x)$ to $\partial_\varepsilon \psi(x)$ (i.e. Y_k contains a finite number of vectors). By applying a standard quadratic programming algorithm, we can compute

$$h_k = \mathrm{Nr}(Y_k) \triangleq \arg\min \{\|h\| \,|\, h \in Y_k\} \ . \tag{3.15}$$

Since $Y_k \subset \partial_\varepsilon \psi(x)$, we must have $\|h_k\| \geq \|h_\varepsilon\|$. Now, referring to Fig. 1a, suppose we find a $g_k \in \partial_\varepsilon \psi(x)$ such that

$$\langle g_k - \tfrac{3}{4} h_k, \, h_k \rangle \leq 0 \tag{3.16}$$

and then form $Y_{k+1} = \{g_k\} \cup Y_k$. We can now compute $h_{k+1} = \mathrm{Nr}(Y_{k+1})$. Clearly, $\|h_\varepsilon\| \leq \|h_{k+1}\| < \|h_k\|$. Furthermore, as is proved in [20], this process cannot continue for an infinite number of k's, unless $h_\varepsilon = 0$. We now show that either h_k is an adequate approximation to h_ε or else a g_k as above can be constructed in a finite number of iterations. Thus, if

$$\psi(x - \frac{\varepsilon}{\|h_k\|} h_k) - \psi(x) \leq - \frac{\varepsilon}{2\|h_k\|} \|h_k\|^2 \leq - \frac{\varepsilon \|h_\varepsilon\|}{2} \tag{3.17}$$

we accept h_k, on the basis (cf.(3.14)) that it is at least "half as good" as h_ε as a descent direction. If (3.17) does not hold, then referring to Fig. 1b (since $\psi(\cdot)$ is directionally differentiable because

and

$$\partial_\varepsilon f^i(x') = \bigcup_{x'' \in B(x',\varepsilon)} \partial f^i(x'') \tag{3.10}$$

where $\partial f^i(x) = \partial y^i(x)$ for $i = 1,2,..,m$ and $\partial f^i(x) = \{g \mid -g \in \partial y^{i-m}(x)\}$ for $i = m+1,m+2,..,2m$. Now, suppose that for some $\varepsilon > 0$, $0 \notin \partial_\varepsilon \psi(x)$. Let

$$h_\varepsilon = Nr(\partial_\varepsilon \psi(x)) \triangleq \arg\min \{\|h'\| \mid h' \in \partial_\varepsilon \psi(x)\} \tag{3.11}$$

Then, for any $\lambda \in (0, \varepsilon/\|h_\varepsilon\|)$, we have $(x-\lambda h) \in B(x,\varepsilon)$ and, by the generalized mean value theorem $\psi(x-\lambda h) - \psi(x)$

$$= \max_i \{f^i(x-\lambda h_\varepsilon) - \psi(x)\}$$

$$= \max_{i \in I_0(x-\lambda h_\varepsilon)} \{f^i(x-\lambda h_\varepsilon) - \psi(x)\}$$

$$= \max_{i \in I_0(x-\lambda h_\varepsilon)} \{f^i(x) - \psi(x) - \lambda \langle g^i_\xi, h_\varepsilon \rangle\} \tag{3.12}$$

where $g^i_\xi \in \partial f^i(x-s^i\lambda h_\varepsilon)$, $i = 1,2,..,2m$ with $s^i \in (0,1)$. Now for $i = 1,2,...,2m, (x-s^i\lambda h_\varepsilon) \in B(x-\lambda h_\varepsilon, \varepsilon)$ by construction and hence all the $g^i_\xi \in \partial_\varepsilon \psi(x)$. Therefore, because of (3.11), we must have

$$\langle g^i_\xi - h_\varepsilon, h_\varepsilon \rangle \geq 0, \qquad i = 1,2,..,2m, \tag{3.13}$$

i.e. $\langle g^i_\xi, h_\varepsilon \rangle \geq \|h_\varepsilon\|^2$. In addition, $f^i(x) - \psi(x) \leq 0$ for all i. Consequently, (3.12) yields

$$\psi(x-\lambda h) - \psi(x) \leq -\lambda \|h_\varepsilon\|^2 \tag{3.14}$$

for all $\lambda \in [0, \varepsilon/\|h_\varepsilon\|]$.

Because $\partial_\varepsilon \psi(x)$ is upper semicontinuous, it follows Theorem 2, p. 116 in [18], that $\|h_\varepsilon\|$ is lower semicontinuous. If we could only

Step 2: If for $j \geq 0$, an integer, such that $\beta\varepsilon/\|h\| \leq \beta^j \leq \varepsilon/\|h\|$, we have

$$\psi(x-\beta^j h) - \psi(x) \leq -\frac{\beta^j}{2}\|h\|^2 , \tag{3.20}$$

set $h_k = h$ and go to End.

Step 3: Set $\ell = 0$, $r = \mu = \varepsilon/\|h\|$.

Step 4: Evaluate $\nabla f^i(x-\mu h)$ for $i \in I_0(x-\mu h)$ (perturbing μ slightly, if necessary to create distinct eigenvalues in this set) and set

$$d\psi(x-\mu h;h) = \max_{i \in I_0(x-\mu h)} \langle \nabla f^i(x-\mu h), h \rangle \tag{3.21}$$

Step 5: If $d\psi(x-\mu h;h) \geq -\frac{3}{4}\|h\|^2$ go to Step 8. Else continue.

Step 6: If

$$\psi(x-\mu h) - \psi(x) \leq -\frac{\mu}{2}\|h\|^2 \tag{3.22}$$

set $\ell = \mu$. Else set $r = \mu$.

Step 7: Set $\mu = (\ell+r)/2$ and go to Step 4.

Step 8: Set

$$Y = Y \cup \{\nabla f^i(x-\mu h) \mid$$

$$i \in I_0(x-\mu h),\ d\psi(x-\mu h;h) = \langle \nabla f^i(x-\mu h), h \rangle \} \tag{3.23}$$

and go to Step 1.

End. ∎

In view of the preceeding discussion, the following result is fairly obvious. For a proof, see [9,20,21,22].

the $y^i(\cdot)$ are one dimensionally differentiable) at the value of $\bar{\lambda} \in [0, \varepsilon/\|h_\varepsilon\|]$ at which $\psi(x-\bar{\lambda}h_k)-\psi(x) = -\frac{\bar{\lambda}}{2}\|h_k\|^2$, we must have a directional derivative $d\psi(x-\bar{\lambda}h_k;h_k) \geq -\frac{1}{2}\|h_k\|^2$. Now,

$$d\psi(x-\bar{\lambda}h_k;h_k) = \max_{i \in I_0(x-\bar{\lambda}h_k)} df^i(x-\bar{\lambda}h_k;h_k) \tag{3.18}$$

Since the values of λ at which the $f^i(\cdot)$ are not differentiable are isolated, we may assume that the $\nabla f^i(x-\bar{\lambda}h_k)$ exist for all $i \in I_0(x-\bar{\lambda}h_k)$ (i.e. that these are all simple eigenvalues). Let

$$g_k \in \{\nabla f^i(x-\bar{\lambda}h_k) \mid i \in I_0(x-\bar{\lambda}h_k)\}$$

be such that $d\psi(x-\bar{\lambda}h_k;h_k) = -\langle g_k, h_k\rangle$. Then we have that a) $g_k \in \partial_\varepsilon\psi(x)$ and b) $d\psi(x-\bar{\lambda}h_k;h_k) = -\langle g_k, h_k\rangle \geq -\frac{1}{2}\|h_k\|^2$. But this implies that

$$\langle g_k, h_k\rangle \leq \frac{1}{2}\|h_k\|^2 < \frac{3}{4}\|h_k\|^2 \tag{3.19}$$

and hence that g_k satisfies (3.16). Consequently, if $0 \notin \partial_\varepsilon\psi(x)$, after a finite number of augmentations of Y_k, we obtain an h_k such that (3.17) holds. Because of the continuity of the $df^i(x-\lambda h;h)$ in λ (which follows from the series (2.7a), (2.10), we can find a $\lambda_k \approx \bar{\lambda}$ and corresponding g_k which satisfies (3.16) in a finite number of operations. We now present such a process in the form of

Subroutine 3.1:

Data: $x \in X$, $\beta \in (0,1)$, $\varepsilon > 0$, $\gamma > 0$; $Y \subset \partial_\varepsilon\psi(x)$, a discrete subset; k, an integer.

Step 1: Compute $h = Nr(coY)$. If $\|h\| \leq \gamma\varepsilon$, set $h_k = h$ and go to End.

Step 6: Set $x_{k+1} = x_k - \beta^{\ell_k} h_k$.

Step 7: If $\psi(x_{k+1}) \leq 0$, stop. Else, set $k = k + 1$ and go to Step 1. □

Theorem 3.1: Suppose that $0 \notin \partial_0\psi(x)$ for all $x \in X$ such that $\psi(x) > 0$, and that Algorithm 3.1 constructs a bounded sequence $\{x_k\}_{k=0}^{\infty}$ in X. If $\{x_k\}$ is infinite, then any accumulation point x^* of $\{x_k\}$ satisfies $\psi(x^*) = 0$.

Proof: First, suppose that x^* is such that $\psi(x^*) > 0$, so that $0 \notin \partial_0\psi(x^*)$. Then, because of the continuity of the $f^i(\cdot)$ and the upper semicontinuity of the $\partial f^i(\cdot)$, there exists an $\varepsilon_1 > 0$ such that $0 \notin \partial_{\varepsilon_1}\psi(x^*)$. Since for all $\varepsilon \in [0,\varepsilon_1]$, $\partial_\varepsilon\psi(x^*) \subset \partial_{\varepsilon_1}\psi(x^*)$, there must exist an $\varepsilon^* = \varepsilon_o/2^i \in [0, \varepsilon_1]$, such that if $h_{\varepsilon^*}(x^*) = Nr\,(\partial_{\varepsilon^*}\psi(x^*))$, then $\|h_{\varepsilon^*}\| \geq 2\delta\varepsilon^* > \delta\varepsilon^*$. Since $\|h_{\varepsilon^*}(x)\|$ is lower semicontinuous ($h_{\varepsilon^*}(x) \triangleq Nr(\partial_{\varepsilon^*}\psi(x))$) there must exist a $\rho^* > 0$ such that

$$\|h_{\varepsilon^*}(x)\| \geq \delta\varepsilon^* \quad \forall\, x \in B(x^*,\rho^*)\,. \tag{3.27}$$

Now, suppose that $\{x_k\}$, as constructed by Algorithm 3.1 is infinite, that x^* is one of its limit points and that $\psi(x^*) > 0$. Then we have: a) $\psi(x_k)$ is a monotonically decreasing sequence. b) There exists a subsequence $\{x_k\}_{k \in K}$ such that $x_k \overset{k}{\to} x^*$ and hence there exists a k^* such that $x_k \in B(x^*,\rho^*)$ for all $k \in K$, $k \geq k^*$, with ρ^* as in (3.27). Hence for $k \in K$, $k \geq k^*$, Step 5 will be reached with $\varepsilon \geq \varepsilon^*$, with ε^* as in (3.27), and hence, from (3.20), (3.26) and (3.27), we must have, with ℓ^* such that $\beta^{\ell^*}\delta \leq 1$, that $\ell_k \leq \ell^*$ for all $k \in K$, $k \geq k^*$, so that

$$\begin{aligned}(x_{k+1}) - \psi(x_k) &\leq -\frac{\beta^{\ell_k}}{2}\|h_k\|^2 \\ &\leq -\frac{\beta^{\ell_k}}{2}\delta\varepsilon^* \quad \forall k \in K,\ k \geq k^*\,.\end{aligned} \tag{3.28}$$

Proposition 3.1: If $0 \notin \partial_\varepsilon \psi(x)$, then Subroutine 3.1 will construct an h_k satisfying (3.20) in a finite number of iterations. ¤

We are finally ready to state an algorithm for solving the problem of finding an $x \in X$ such that $\psi(x) \leq 0$.

Algorithm 3.1:

Data: x_0, $\varepsilon > 0$, $\alpha > 0$, $\beta \in (0,1)$, $\gamma > 0$, $\delta > 0$.

Step 0: set $k = 0$.

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

Step 2: For $i \in I_\varepsilon(x_k)$,

If $f^i(x_k)$ corresponds to a simple eigenvalue of $Q(x_k)$, compute

$$g_i = \Delta f^i(x_k) \tag{3.24}$$

making use of the formula (2.14).

If $f^i(x_k)$ corresponds to a multiple eigenvalue of $Q(x_k)$, compute

$$g_i = \nabla f^i(x_k+d) \tag{3.25}$$

Where d is a randomly selected vector, with $\|d\| \leq 0.9\varepsilon$.

Set $Y = \{g_i | i \in I_\varepsilon(x_k)\}$.

Step 3: Transfer x_k, Y, ε, γ, k to subroutine 3.1 to compute h_k.

Step 4: If $\|h_k\| < \delta\varepsilon$, set $\varepsilon = \varepsilon/2$ and go to Step 2. Else continue.

Step 5: Compute the smallest integer $\ell_k \geq 0$ such that

$$\psi(x_k - \beta^{\ell_k} h_k) - \psi(x_k) \leq -\frac{\beta}{2}^{\ell_k} \|h_k\|^2 \tag{3.26}$$

$$\ell(\omega) \leq \sigma^i(x,\omega) \leq u(\omega), \qquad \text{for } i = 1,2,..,n$$
$$\omega \in [\omega_o, \omega_f] \tag{4.1a}$$

Where $\ell, u : \mathbb{R}^1 \to \mathbb{R}^1_+$ are continuous functions and the $\sigma^i(x,\omega)$ are the singular values of a m x m transfer function matrix $G(x,\omega)$. As in the preceeding section, we rewrite (4.1a) into the equivalent problem of finding an $x \in X$ satisfying

$$\ell(\omega)^2 \leq y^i(x,\omega) \leq u(\omega)^2, \qquad \text{for } i = 1,2,..,m$$
$$\omega \in [\omega_o, \omega_f] , \tag{4.1b}$$

where the $y^i(x,\omega) = \sigma^i(x,\omega)^2$ are the eigenvalues of the positive semidefinite m x m matrix

$$Q(x,\omega) \triangleq G(x,\omega) = G(x,\omega)^* G(x,\omega). \tag{4.2}$$

As in section 3, we define again

$$f^i(x,\omega) \triangleq y^i(x,\omega) - u(\omega)^2, \quad \text{for } i = 1,2,..,m \tag{4.3a}$$

and

$$f^i(x,\omega) \triangleq \ell(\omega)^2 - y^{i-n}(x,\omega), \text{ for } i = m+1, m+2, .., 2m . \tag{4.3}$$

Next, we define

$$\psi(x,\omega) \triangleq \max_{i \in \underline{2m}} f^i(x,\omega) \tag{4.4}$$

Where $\underline{2m} = 1,2,..,2m$.

Finally, we define $\phi : X \to \mathbb{R}^1$ by

$$\phi(x) \triangleq \max_{\omega \in \Omega} \psi(x,\omega) \tag{4.5}$$

with $\Omega = [\omega_o, \omega_f]$. We note that $\psi(\cdot,\cdot)$ is continuous and hence, since

But (3.28) implies that $\psi(x_k) \to -\infty$ which, by continuity of $\psi(\cdot)$, contradicts our assumption that $\psi(x^*) > 0$. Hence we are done. ¤

An important assumption in Theorem 3.1 is that the sequence $\{x_k\}$ constructed by Algorithm 3.1 is in X, the subset of $\mathbb{R}^k$ in which the transfer function $G(x,\hat{\omega})$ matrix is componentwise analytic. If we assume that for $\hat{x} \notin X$, $G(x,\hat{\omega})$ has a pole at $\hat{x}$, then $\psi(\hat{x}) = \infty$ and, provided an upper bound on the singular values is included or, we see that the Algorithm 3.1 will automatically avoid such points.

Finally, an important and rather obvious extension of Theorem 3.1 is as follows.

Corollary 3.1: Suppose that $0 \notin \partial_0 \psi(x)$ for all $x \in X$ such that $\psi(x) \geq 0$. If Algorithm 3.1 constructs a bounded sequence $\{x_k\}$, then this sequence must be finite and its last element, say $x_{\hat{k}}$, satisfies $\psi(x_{\hat{k}}) \leq 0$. ¤

The assumptions in Theorem 3.1 and Corollary 3.1 that $0 \notin \partial_0 \psi(x)$ for all $x \in X$ such that $\psi(x) > 0$ (or $\psi(x) \geq 0$) is totally analogous to a similar assumption made in phase I methods of feasible directions (see [19]) and is weaker than appropriate assumptions for Newton's method to work in solving differentiable inequalities.

We are now ready to proceed to our final stage, namely the construction of an algorithm for solving singular value inequalities over a continuum of frequencies.

4. A Master Algorithm for Singular Value Inequalities Over a Continuum of Frequencies.

We now turn to our problem in its most general form, namely, find an $x \in X$ such that

which is why the term "outer approximation" is used for the set on the left. The Master Algorithm below, makes use of a double subscripted sequence $\{\varepsilon_{jk}\}_{k \geq j}$ which must satisfy the following hypotheses:

$$\text{H1:} \quad \varepsilon_{kk} = 0 \text{ for all } k \text{ and } \varepsilon_{jk} > 0 \text{ for all } k > j, \tag{4.11a}$$

$$\text{H2:} \quad \varepsilon_{jk} \nearrow \hat{\varepsilon}_j \text{ as } k \to \infty; \tag{4.11b}$$

$$\text{H3:} \quad \hat{\varepsilon}_j \searrow 0 \text{ as } j \to \infty. \tag{4.11c}$$

There is some skill involved in selecting a sequence $\{\varepsilon_{jk}\}$ which is particularly well matched to one's problem. However, a good rule of thumb is to keep the ε_{jk} as large as possible for as long as possible since this results in the smallest sets of finite inequalities that must be solved at each iteration of the Master Algorithm. A particular example of such a sequence $\{\varepsilon_{jk}\}$ is given by $\varepsilon_{jk} = 100\ [(\frac{1}{j+1})^{1/10} - (\frac{1}{k+1})^{1/10}]$, $k = 0,1,2,\ldots,$ and $j \leq k$.

Finally, if we wish to be sure that Algorithm 3.1 will solve the finite sets of inequalities to be constructed by the Master Algorithm, we must make the following assumption.

Assumption 4.1: For any finite subset $\Omega_k \subset \Omega$, for every $x \in X$ such that $\max_{\omega \in \Omega_k} \psi(x,\omega) \geq 0$,

$$0 \notin \text{co} \bigcup_{\omega \in \Omega_k(x)} \partial\psi_0(x,\omega), \tag{4.12}$$

where

$$\Omega_k(x) \triangleq \{\omega' \in \Omega_k \mid \psi(x,\omega') = \max_{\omega \in \Omega_k} \psi(x,\omega')\}. \tag{4.13}$$

□

Ω is a fixed compact set, by the Maximum Theorem p. 116 in [18], $\phi(\cdot)$ is also continuous. Thus, the problem we wish to solve, can be stated as that of finding an $\hat{x} \in X$ such that

$$\phi(\hat{x}) \leq 0. \tag{4.6}$$

We propose to solve (4.6) by decomposing this problem into an infinite sequence of finite sets of inequalities which need to be solved only approximately by Algorithm 3.1. Our method is extracted from the outer approximations methods proposed for infinitely constrained optimization problem in [23].

First, we must introduce an acceptable scheme for approximately evaluating $\phi(x)$. Let $\ell : \mathbb{N}_+ \to \mathbb{N}_+$ (with $\mathbb{N}_+ = \{0,1,2,3,\ldots\}$) be such that $\ell(k) \to \infty$ monotonically as $k \to \infty$ and, for $k \in \mathbb{N}_+$, let

$$\tilde{\Omega}_k \triangleq \{\omega | \omega = \omega_o + j\,\frac{\omega_f - \omega_o}{\ell(k)}, \quad j = 1,2,..,\ell(k)\} \tag{4.7}$$

and let

$$\phi_k(x) \triangleq \max_{\omega \in \tilde{\Omega}_k} \psi(x,\omega) \tag{4.8}$$

Obviously, $\phi_k(x)$ is always computable in a finite number of operations (assuming that the σ^i computations are finite). Also, given any compact set $C \subset X$,

$$|\phi_k(x) - \phi(x)| \to 0 \text{ as } k \to \infty\,, \tag{4.9}$$

uniformly in C. Note also that

$$\{x | \phi_k(x) \leq 0\} \supset \{x | \phi(x) \leq 0\}, \tag{4.10}$$

Proof: We distinguish between two cases. First suppose that the Master Algorithm jams up, cycling between Steps 2 and 3 and generates an infinite sequence $x_{k_o} = x_{k_o+1} = - - - .$, i.e., $x_k = x_{k_o}$ for all $k \geq k_o$. Then we must have $\phi_k(x_{k_o}) \leq 0$ for all $k \geq k_o$ and hence, $\phi(x_{k_o}) \leq 0$.

Next, suppose the sequence $\{x_k\}$ was constructed without jamming up in the loop defined by Steps 2 and 3, and suppose that $x_k \overset{K}{\to} \hat{x}$. For the sake of contradiction, suppose that $\phi(\hat{x}) > 0$. Hence, since by continuity $\phi(x_k) \overset{K}{\to} \phi(\hat{x})$ and $|\phi_k(x_k) - \phi(x_k)| \to 0$ as $k \to \infty$, and $\hat{\varepsilon}_j \searrow 0$, by construction, there must exist a j_o such that

$$\phi_j(x_j) \geq \phi(\hat{x})/2 \geq \hat{\varepsilon}_j \geq \varepsilon_{jk} \tag{4.16}$$

for all $j \in K$ such that $j \geq j_o$ and $k \geq j$. Since by construction of x_k and Ω_k, $\psi(x_k,\omega) \leq 0$ for all $\omega \in \Omega_k$, and $\omega_j \in \Omega_k$ for all $j \in K$, $j \geq j_o$ and $k > j$, we must have

$$\psi(x_k,\omega_j) \leq 0 \text{ for all } j \in K,\ j \geq j_o \text{ and } k \geq j+1 . \tag{4.17}$$

Now, Ω is compact and hence $\psi(\cdot,\omega)$ is continuous, uniformly in $\omega \in \Omega$. Therefore we must have

$$|\psi(x_k,\omega_j) - \psi(x_j,\omega_j)| \to 0 , \tag{4.18}$$

as $j \to \infty$, with $j, k \in K$ and $k \geq j+1$. But (4.17) and (4.18) imply that

$$\overline{\lim_{j \in K}}\ \phi_j(x_j) \leq 0 , \tag{4.19}$$

which, in turn, implies that $\overline{\lim_{j \in K}}\ \phi(x_j) \leq 0$. Hence since $x_j \overset{K}{\to} \hat{x}$ and $\phi(\cdot)$

We are now ready to introduce the

Master Algorithm 4.1

Data: A set of discrete frequencies $\Omega_0 \subset \Omega$ and a double subscripted sequence $\{\varepsilon_{jk}\}_{k \geq j}$ satisfying (4.11a) - (4.11c). A function $\ell : \mathbb{N}_+ \to \mathbb{N}_+$ such that $\ell(k) \nearrow \infty$ as $k \to \infty$.

Step 0: Set $k = 0$.

Step 1: Compute by means of Algorithm 3.1 an x_k such that

$$\psi(x_k, \omega) \leq 0 \text{ for all } \omega \in \Omega_k . \tag{4.14}$$

Step 2: Compute $\phi_k(x_k)$ and an $\omega_k \in \tilde{\Omega}_k$ such that $\psi(x_k, \omega_k) = \phi_k(x_k)$.

Step 3: If $\psi_k(x_k) \leq 0$, set $x_{k+1} = x_k$, set $k = k+1$ and go to Step 2.

Step 4: Construct

$$\Omega_{k+1} = \{\omega_k\} \cup \{\omega_j \in \Omega_k \,|\, \psi(x_j, \omega_j) \geq \varepsilon_{jk}\} . \tag{4.15}$$

Set $k = k+1$ and to to Step 1. ¤

Examining (4.15) we see the advantage in keeping the ε_{jk} as large as possible for as long as possible: this results in the largest removal of points ω_j from Ω_k in the construction of Ω_{k+1}. It should also be noted that a good thing can be overdone and it is best to adjust the ε_{jk} sequence interactively, as the computation progresses.

Theorem 4.1: Suppose that Assumptions 2.1, 2.2 and 4.1 are satisfied. Let $\{x_k\}_{k=0}^{\infty}$ be any sequence constructed by the Master Algorithm 4.1. Then any accumulation point $\hat{x}$ of $\{x_k\}_{k=0}^{\infty}$ satisfies $\psi(\hat{x}) \leq 0$.

REFERENCES

1. A. J. Laub, "Linear Multivariable Control: numerical considerations," Report No. ESL-P-833, Electronics Systems Lab, MIT 197.
2. A. J. Laub "An inequality and some computations related to the robust stability of linear dynamic systems," IEEE Trans. Automat. Contr., Vol. AC-24, No. 2, pp. 318-320, 1979.
3. J. C. Doyle, "Robustness of mutliloop linear feedback systems," in Proc. 17th IEEE Conf. on Decision and Control, 1979.
4. J. C. Doyle, "Achievable performance in multivariable feedback systems," in Proc. 18th IEEE Conf. on Decision and Control, 1970.
5. N. R. Sandel, "Robust stability of systems with applications to singular perturbations," Automatica,Vol. 15, p. 467, 1979.
6. V. Zakian and V. Al-Naib, Design of dynamical and control systems by the method of inequalities," Proc. IEEE, 120, No. 11, pp. 1421-1427, 1973.
7. E. Polak and R. Trahan, "An algorithm for computer aided design problems of control systems," Proc. IEEE Conf. on Decision and Control, 1976.
8. D. Q. Mayne, E. Polak and A. Sangiovanni-Vincentelli, "Computer aided design via optimization," Proc. IFAC Workshop on Control Applications of Nonlinear Programming, Denver, Co., June 1979 (in press).
9. E. Polak and A. Sangiovanni-Vincentelli, "Theoretical and computational aspects of the optimal design centering, tolerances and tuning problems," IEEE Trans. Automat. Contr., Vol. CAS-26, No. 9, p. 295-813, 1979.
10. L. V. Ahlfors, Complex Analysis, McGraw Hill, N.Y., 1953.
11. J. H. Wilkinson, The Algebraic Eigenvalue Problem, Clarendon Press, Oxford, 1961.

is continuous, we must have $\phi(\hat{x}) \leq 0$. But this contradicts our hypothesis and hence we are done. ¤

Although we are not able to establish that the cardinality of the sets Ω_k remains bounded, computational experience leads one to expect the set Ω_k to be small and to stay small, and hence the computation of the x_k is not too difficult. The computation of x_{k+1} is further facilitated by using x_k as an initial point for Algorithm 3.1.

5. Conclusion:

We have shown that distributed singular value inequalities can be solved by combining concepts of outer approximations with those of nondifferentiable optimization. The algorithm is usable as presented. However, since it is the first algorithm for this class of problems, we expect with time to evolve more efficient versions which make greater use of the properties of singular values.

FOOTNOTES

[1] $G(x,\omega)$ is a complex valued matrix.

[2] To extend our treatment to more than one frequency $\hat{\omega}$, all that needs to be done is that more $f^i(\cdot)$ will need to be defined.

[3] It can be seen from the proofs to follow that we may have used $I_0(x')$ instead of $I_\varepsilon(x')$ in the definition of $\partial_\varepsilon\psi(x)$. However, the use of $I_\varepsilon(x')$ is bound to result in better computational behavior.

12. B. N. Parlett, The symmetric eigenvalue problem, Prentice Hall, N.Y., 1980.

13. W. Rudin, Principles of mathematical analysis, McGraw Hill, N.Y., 1964.

14. R. Bellman, Introduction to matrix analysis, McGraw-Hill, N.Y., 1960.

15. G. Golub and C. Reinsh, "Singular value decomposition and least squares solutions," Num. Math. Vol. 14, pp. 403-420, 1970.

16. F. H. Clarke, "Generalized gradients and applications," Trans. Amer. Math. Soc., Vol. 205, pp. 247-262, 1975.

17. G. Lebourg, "Valeur moyenne pour gradient généralisé," C.R. Acad. Sci., Paris, Vol. 281, pp. 795-797, 1975.

18. C. Berge, Topological Spaces, Macmilan Co., N.Y., 1963.

19. E. Polak, R. Trahan and D.Q. Mayne, "Combined phase I-phase II methods of feasible directions," Math Programming, Vol. 17, p. 61-73, 1979

20. R. Mifflin, "An algorithm for constrained optimization with semi-smooth functions," Math. of Oper. Research, Vol. 2, No. 2, pp. 191-207, 1977.

21. C. Lemarech "An extension of Davidon methods to nondifferentiable problems," in Mathematical Programming Studies 3, M.L. Balinski and R. Wolfe Eds., North Holland, Amsterdam, pp. 95-109, 1975.

22. C. Gonzaga and E. Polak, "On constraint dropping schemes and optimality functions for a class of outer approximations algorithms," SIAM J. Contr. and Optimization, Vol. 17, pp. 477-493, 1979.

23. C. A. Desoer and M. J. Chen, "Design of multivariable control systems with stable plant," 1980 JACC.

and let U_ε be a matrix of orthonormal eigenvectors associated with the eigenvalues $f^i(x)$ of $Q(x)$. Finally, let $E_{\varepsilon(x-\mu h,h)}$ be a diagonal matrix whose elements are the ordered eigenvalues of $U_\varepsilon^* \frac{\partial Q}{\partial \lambda}(x-\lambda h)\big|_{\lambda=\mu} U_\varepsilon$. Then

$$\Lambda_\varepsilon(x-mh,h) \simeq E_\varepsilon(x-\mu h,h) \tag{A3}$$

Thus, rather than performing a <u>full</u> singular value decomposition at $x-\mu h$ we only need to perform of singular value decomposition for the much smaller matrix $U_\varepsilon^* \frac{\partial Q}{\partial x}(x-\lambda h)\big|_{\lambda=\partial} U_\varepsilon$. Furthermore, if V is a matrix of eigenvectors for this smaller matrix, then the required eigenvectors for the gradient formulas are the columns of the matrix $U_\varepsilon V$. We note that whereas the calculation of eigenvectors in the vicinity of multiple singular values by SVD is highly ill conditioned, the above described process can be expected to be quite well conditioned.

Appendix: Computational Aspects of Subroutine 3.1

We shall now examine some of the computational aspects of subroutine (3.1). First, since in step 5, the test $d\psi(x-mh;h) \geq -\frac{3}{4}\|h\|^2$ is used, the next h computed in step 1 (say, h_1)will satisfy, approximately, $\|h_1\| \leq \frac{3}{4}\|h\|$. Obviously this process is not likely to go on for more than a few iterations when $\|Nr(\partial_\varepsilon\psi(x))\|$ is comparable to the "diameter" of that set. As $\|Nr(\partial_\varepsilon\psi(x))\|$ goes to zero, progressively more work would be required. However, since we are not minimizing $\psi(x)$ but only looking for an x such that $\psi(x) \leq 0$, it can be expected that the work spent in Subroutine 3.1 will never be excessive. It is possible to decrease this work even further at the expense of risking to accept a somewhat worse direction by replacing the factor $\frac{1}{2}$ in (3.20), (3.22) and (3.26) by 0.1, say, and the factor $\frac{3}{4}$ in Step 5, by 0.15, say, which results in the successive vector h produced in step 1 of Subroutine 3.1 to drop in length by a factor of approximately 0.15, (i.e. $\|h_{new}\| \leq 0.15\|h_{old}\|$).

Next, it has been pointed out to us by Prof. W. Kahan, of the Department of Electrical Engineering and Computer Sciences, University of California, Berkeley, that the computation required in (3.21) can be carried out approximately without performing a full singular value decomposition, thus resulting in considerable savings. This suggestion is based on the observation that for ε small,

$$d(x-mh;h) \approx \max_{i\in I_\varepsilon(x)} df^i(x-\mu h;h) \tag{A1}$$

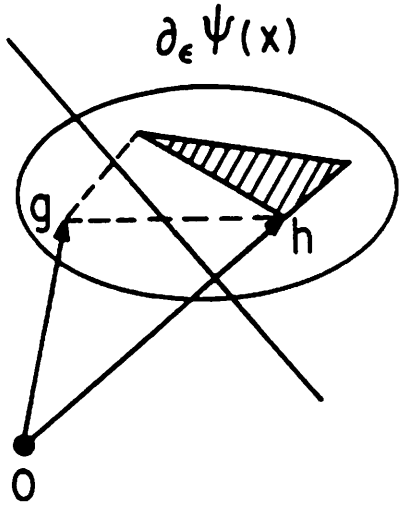
and that the directional derivatives $df^i(x-mh;h)$, $i \in I_\varepsilon(x)$ are approximately given by the following formula: Let

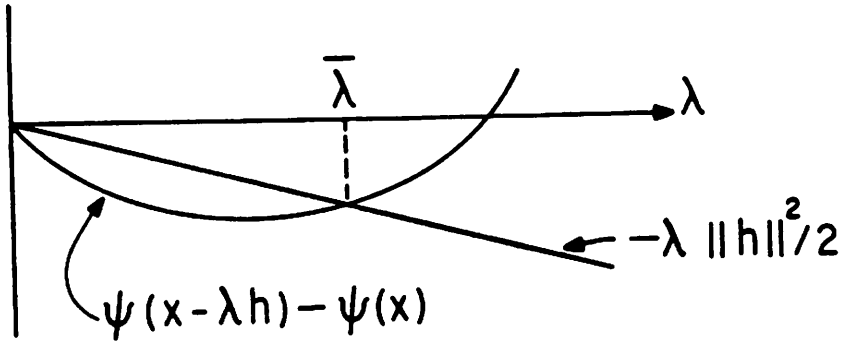
$$\Lambda_\varepsilon(x-mh;h) \triangleq \text{diag}(df^i(x-mh;h)) \tag{A2}$$

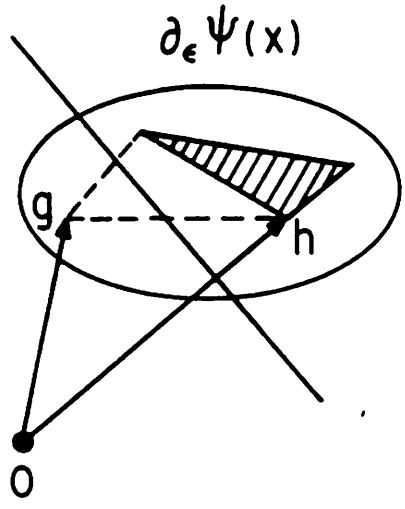
FIGURE CAPTIONS

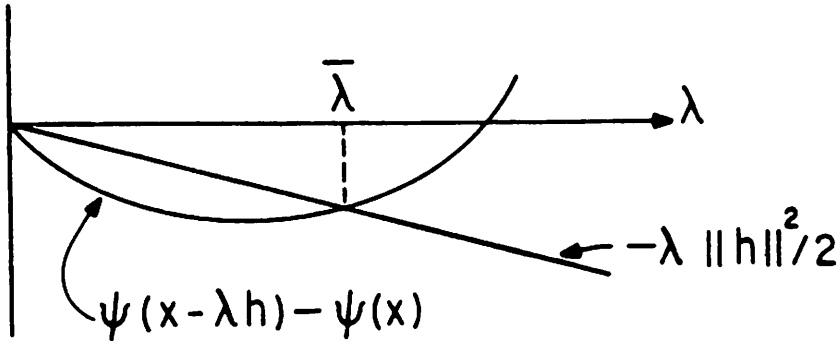
Fig. 1a. Successive improvement of h.

Fig. 1b. Definition of $\overline{\lambda}$.


$\partial_\epsilon \psi(x)$
g
h
0


$\bar{\lambda}$
λ
$-\lambda \|h\|^2/2$
$\psi(x-\lambda h)-\psi(x)$


$\partial_\epsilon \psi(x)$
g
h
0


$\bar{\lambda}$
λ
$-\lambda \|h\|^2/2$
$\psi(x-\lambda h)-\psi(x)$