Proceedings of the 28th Conference
on Decision and Control
Tampa, Florida • December 1989

A Constrained H^∞ Smooth Optimization Technique

M.E. Fisher[1], J.B. Moore[2] and K.L. Teo[1]

1. Department of Mathematics, University of Western Australia, Nedlands, W.A. 6009, Australia.
2. Department of Systems Engineering, Australian National University, P.O. Box 4, Canberra, ACT 2601, Australia.

Abstract

In this paper, we consider a general class of constrained H^∞ optimization problems and show that these problems can be approximated by a sequence of smooth optimization problems. Thus, each of the approximate problems is readily solvable by standard optimization software packages such as those available in the NAG library or IMSL Library. The proposed approach via smooth optimization is simple in terms of mathematical content, easy to implement, and is computationally efficient.

1. Introduction and Background

A method of linear multivariable control system design which is of current interest is H^∞ optimal control [1,2]. The simplest class of such problems known as one-block problems can be formulated as follows:

$$\min_{Q \in RH^\infty} \|T_{11} + T_{12}QT_{21}\|_\infty . \qquad (1.1)$$

Here $T_{ij} \in R_p$, the class of rational proper transfer function matrices, and RH^∞ denotes the class of such which are stable. Also, for continuous time transfer functions $X(s)$,

$$\|X\|_\infty = \max_{i,\omega} \sigma_i[X(s)]_{s=j\omega} = \max_\omega \bar{\sigma}[X(\omega)]$$
$$= \{\max_\omega \bar{\lambda}[X^*(\omega)X(\omega)]\}^{1/2}, \qquad (1.2)$$

where σ_i denotes the ith singular value, $\bar{\sigma}$ the maximum singular value, $\bar{\lambda}$ the maximum eigenvalue and the superscript * denotes the conjugate transpose.

Elegant optimization techniques for (1.1) are available [1-4]. Also, these apply with modifications to the two-block and four-block problems which are formulated as in (1.1) but with Q constrained as $[Q_1 \;\; 0]$, $[Q_1^T \;\; 0]$, or the Q block diagram $(Q_1, 0)$, respectively.

In this paper, we are concerned with the optimization task where additional constraints are included in the optimization. These additional constraints may be in such a way that the controllers are of a fixed structure, perhaps conveniently parametrized in terms of a parameter vector. Also, they may be functional constraints involving singular values representing, for example, robustness requirements. Such problems cannot be solved by using the elegant techniques of [1-4]. It is well known that with these additional constraints, more complicated nonlinear programming techniques must be employed, and local minima are obtained rather than global minima.

In [5], a computational technique is developed for solving singular value inequalities over a continuum of frequencies. The case when two or more singular values are identical on an interval is excluded by assumption. In [6], an optimization problem is considered, where a differentiable cost functional is to be minimized subject to three kinds of constraints including singular valued inequalities over a continuum of frequencies. An improved algorithm which overcomes the difficulties caused by singular vector computations is then proposed in [6]. Mathematically, these two papers are highly complex. Furthermore, they cannot make use of standard optimization software packages such as are available in the NAG or IMSL library.

The key contribution of [7] is to give a simple yet efficient approach for solving functional inequality constrained optimization problems. The approach involves a new constraint transcription together with a local smoothing technique. On this basis, a sequence of approximate optimization problems can be constructed. Each of these approximate problems can be viewed as a conventional optimization problem. The approach is very simple and can make use of existing optimization software packages, yet is guaranteed to converge. From numerical studies on simple problems, the computational effort appears to be considerably less than using the nonsmooth approach of [8,9]. Our purpose in this paper is to demonstrate that under reasonable assumptions, constrained H^∞ control problems and optimizations involving singular value constraints, such as in [5, 6], can be tackled using existing smooth optimization software packages. This is achieved by noting that the objective functional (1.2) can be approximated by

$$\|\bar{\sigma}[X(\omega)]\|_{p,[0,c]} = \left\{ \int_0^c (\bar{\lambda}[X^*(\omega)X(\omega)])^p \, d\omega \right\}^{1/2p}, \qquad (1.3)$$

with an appropriate positive integer p and positive constant c, while the techniques of [7] can be trivially extended to handle the functional constraints. The extensions allow infinite intervals for the frequency range of the constraint, and certain nonsmoothnessin the constraint functions.

2. Stabilizing Controller Theory

Consider the feedback control schemes of Figure 2.1. In Figure 2.1(a), there is a nominal plant with transfer function matrix $P \in R_p$ and controller $K \in R_p$. It is known that the control loop is well posed and the controller K is stabilizing for $G = P_{22}$ if and only if

$$\begin{bmatrix} I & -K \\ -G & I \end{bmatrix}^{-1} \qquad (2.1)$$

exists and belongs to RH^∞. The theory for the class of all stabilizing controllers of [2] allows the parametrization of all stabilizing controllers K for G as a linear fractional map K(Q) in terms of arbitrary $Q \in RH^\infty$ (see Fig. 2.1(b)). The matrix $J \in R_p$ (non unique) is readily calculated from G and any stabilizing proper controller K_0 for G. From P, J of Fig. 2.1(b), a $T \in R_p$ can be constructed allowing Fig. 2.1(b) to be re-organized as in Fig. 2.1(c). A key property, namely $T_{22} = 0$, ensures that closed transfer functions are affine in Q. The stabilization theory tells us that

$$\{K \text{ stabilizes } P\} \Leftrightarrow \{Q \text{ stabilizes } T\}. \qquad (2.2)$$

Now, again referring to Fig. 2.1, let us denote the transfer function matrix from the disturbance w to the disturbance response e as W. Since this transfer function is K or Q dependent we also use the notation W(K) or W(Q) as appropriate. A standard formulation of the H^∞ optimization task is:

$$\min_{\left\{\begin{smallmatrix}\text{stabilizing } K \\ \text{for } P\end{smallmatrix}\right\}} \|W(K)\|_\infty . \qquad (2.3)$$

The equivalent task, under (2.2) and the relationships

$$T_{22} = 0, \quad W(K) = W(Q) = T_{11} + T_{12}QT_{21} \qquad (2.4)$$

is the derivative task (1.1), for which elegant solutions exist. It turns out that with P having McMillan degree n, the optimal Q leads, via the linear fractional map K(Q), to a controller of degree n-1 [10].

In practice, the controller class may be specified with more restrictions than merely that K is stabilizing for P. Let us consider the case when the structure of K is specified in terms of parameters $x \in \Xi$ with Ξ a compact subset of $\mathbb{R}^{\ell}$. It may be that the parametrization is on Q and thereby on K as $K(x) = K[Q(x)]$. Also, in practice there may be frequency domain constraints on the closed-loop system behavior other than mere stability. They may involve functional constraints on singular values. These motivate for us the following class of constrained H^{∞} optimization problems.

Constrained H^{∞} Optimization Problem

$$\min_{x \in \Xi} \|W(x)\|_{\infty}, \quad W(x) \in R_p ; \tag{2.5a}$$

subject to the constraints

$$\text{(i)} \quad h_j(x) \le 0, \quad j = 1,2,\ldots,N; \tag{2.5b}$$

$$\text{(ii)} \quad \max_{\omega \in \Omega} \phi_j(x,\omega) \le 0, \quad j = 1,\ldots,M; \tag{2.5c}$$

where $\Omega = [0,\infty)$, and $h_j:\mathbb{R}^{\ell} \to \mathbb{R}$, $\phi_j:\mathbb{R}^{\ell}\times\mathbb{R} \to \mathbb{R}$. Notice that frequency domain constraints such as stability constraints, robustness constraints, or frequency response constraints can be handled by (2.5c). These constraints can involve singular values in multivariable problems such as are studied in [4].

For convenience, let the constrained H^{∞} optimization problem (2.5) be referred to as the problem (P).

3. Reformulation as a Smooth Optimization Problem

We reformulate the constrained H^{∞} optimization problem as a smooth optimization problem, in two steps.

In Step 1, we make a smooth approximation to the objective function using the L_p norm. In Step 2, the functional constraints are approximated by conventional constraints using the technique similar to that given in [7]. To begin, let us assume that the following conditions are satisfied.

(1) $W(x,s)$ is strictly proper (zero for $s = \infty$) with no poles on the $j\omega$-axis;

(2) $W(x,s)$ for each x has distinct singular values for almost all $s = j\omega$ with no $j\omega$-axis poles;

(3) $W(x,s)$ is continuously differentiable with respect to x for almost all $s = j\omega$.

Smooth Approximation of the Objective Function

Consider the objective function given by (2.5a). By (1.2), it may be written as:

$$\min_{x \in \Xi} \|W(x)\|_{\infty} = \left\{ \min_{x \in \Xi} \max_{\omega \ge 0} \{\bar{\lambda}(x,\omega)\} \right\}^{1/2}, \tag{3.1}$$

where $\bar{\lambda}(x,\omega)$ denotes the maximum eigenvalue of $W^*(x,\omega)W(x,\omega)$.

Lemma 3.1. Under assumption (1), there exists a $x^* \in \Xi$ and an $\omega^* \in (0,\infty)$ such that

$$\min_{x \in \Xi} \max_{\omega \ge 0} \{\bar{\lambda}(x,\omega)\} = \bar{\lambda}(x^*,\omega^*) < \infty .$$

Proof. By assumption (1), it follows that for each $x \in \Xi$,

$$\max_{\omega \ge 0} \{\bar{\lambda}(x,\omega)\}$$

is attained at some finite ω. Let $\{x^i\}$ be a sequence in Ξ such that

$$f_i = \max_{\omega \ge 0} \{\bar{\lambda}(x^i,\omega)\} \longrightarrow \bar{f} = \min_{x \in \Xi} \max_{\omega \ge 0} \{\bar{\lambda}(x,\omega)\}. \tag{3.2}$$

Since Ξ is compact, $\{x^i\}$ has a convergent subsequence, denoted again by the original sequence, which converges, say, to $x^* \in \Xi$. By the above, there exists some $\omega^* \in (0,\infty)$ such that

$$\{\bar{\lambda}(x^*,\omega^*)\} = \max_{\omega \ge 0} \{\bar{\lambda}(x^*,\omega)\}. \tag{3.3}$$

Since for all $i = 1,2,\ldots$ and for all $\omega \ge 0$,

$$\bar{\lambda}(x^i,\omega) \le f_i$$

it follows that for all $i = 1,2,\ldots$

$$\bar{\lambda}(x^i,\omega^*) \le f_i .$$

Therefore, by virtue of the continuity of the function f,

$$\bar{\lambda}(x^*,\omega^*) \le \lim_{i\to\infty} f_i = \bar{f}.$$

This, in turn, implies the conclusion of the Lemma.

Lemma 3.2. Under assumptions (1) and (2), there exists positive constants κ and $\hat{c}$ such that

$$\mathrm{tr}\{W^*(x,\omega)W(x,\omega)\} \le \kappa/\omega^2 \tag{3.4}$$

for all $\omega > \hat{c}$ and $x \in \Xi$.

Proof. Let $w_{ij}(x,\omega)$ be the (i,j) entry of the $q \times q$ matrix $W(x,\omega)$. Then

$$\mathrm{tr}\{W^*(x,\omega)W(x,\omega)\} = \sum_{i=1}^{q} \sum_{j=1}^{q} |w_{ij}(x,\omega)|^2 \tag{3.5}$$

remembering that for each x and ω the $w_{ij}(x,\omega)$ are complex numbers.

Choose arbitrary i and j with $1 \le i \le q$ and $1 \le j \le q$. By assumption (1) it follows that

$$w_{ij}(x,\omega) = \frac{a_n s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0}{b_m s^m + b_{m-1} s^{m-1} + \cdots + b_1 s + b_0} \tag{3.6}$$

where $s = j\omega$, for some appropriate n, m and coefficients $\{a_i\}_{i=0}^{n}$ and $\{b_j\}_{j=0}^{m}$ with $m \ge n+1$, $b_m \ne 0$ and $b_0 \ne 0$. Here, the coefficients $\{a_i\}_{i=0}^{n}$ and $\{b_j\}_{j=0}^{m}$ are functions of x. Hence

$$|w_{ij}(x,\omega)|^2 = \frac{|a_n s^n + a_{n-1} s^{n-1} + \cdots + a_1 s + a_0|^2}{|b_m s^m + b_{m-1} s^{m-1} + \cdots + b_1 s + b_0|^2} \tag{3.7}$$

Now,

$$|a_n s^n + \cdots + a_1 s + a_0|^2 \le 2(a_n^2\omega^{2n} + \cdots + a_1^2\omega^2 + a_0^2). \tag{3.8}$$

Since Ξ is a compact set, there exists positive constants α_k, $k = 0,1,\ldots,n$, such that $a_k^2 \le \alpha_k$ for all $x \in \Xi$. Let

$$\alpha = \max_{0 \le k \le n} \{\alpha_k\}.$$

Then, for $\omega \ge 1$ and all $x \in \Xi$,

$$|a_n s^n + \cdots + a_1 s + a_0|^2 \le 2\alpha(n+1)\omega^{2n}. \tag{3.9}$$

Consider m an even integer. Then

$$|b_m s^m + \cdots + b_1 s + b_0|^2 = \left((-1)^{m/2} b_m\omega^m + \cdots - b_2\omega^2 + b_0\right)^2 + \left((-1)^{(m-2)/2} b_{m-1}\omega^{m-1} + \cdots - b_3\omega^3 + b_1\omega\right)^2$$

$$\ge b_m^2\omega^{2m}\left(1 - \frac{b_{m-2}}{b_m}\omega^{-2} + \cdots - (-1)^{m/2}\frac{b_0}{b_m}\omega^{-m}\right)^2$$

$$\ge b_m^2\omega^{2m}\left(1 - 2\left|\frac{b_{m-2}}{b_m}\right|\omega^{-2} + \cdots - 2\left|\frac{b_0}{b_m}\right|\omega^{-m}\right).$$

Now, $b_m \ne 0$ and so there exists constants $\beta > 0$ and $\beta_{\ell} > 0$, $\ell = 0,1,\ldots,m-1$, such that, for all $x \in \Xi$,

$$|b_m| \ge \beta$$

and

$$\left|\frac{b_{\ell}}{b_m}\right| \le \beta_{\ell}, \quad \ell = 0,1,\ldots,m-1. \tag{3.10}$$

Let

$$\bar{\beta} = \max_{0 \le \ell \le m-1} \{\beta_\ell\} . \tag{3.11}$$

Then, from the above,

$$|b_m s^m + \cdots + b_1 s + b_0|^2$$
$$\ge \beta^2\omega^{2m}\left(1 - \frac{2\bar{\beta}}{\omega^2}(1 + \omega^{-2} + \cdots + \omega^{2-m})\right)$$
$$\ge \beta^2\omega^{2m}\left(1 - \frac{4\bar{\beta}}{\omega^2}\right), \tag{3.12}$$

for all $x \in \Xi$, provided $\omega \ge 2$. The inequality (3.12) can also be shown to hold for m an odd integer.

Now choose $\bar{\omega} = \max\{2, 2(2\bar{\beta})^{1/2}\}$. Then, for all $x \in \Xi$ and all $\omega \ge \bar{\omega}$, we have

$$|b_m s^m + \cdots + b_1 s + b_0|^2 \ge \frac{\beta^2\omega^{2m}}{2} . \tag{3.13}$$

Combining (3.7), (3.9) and (3.13) gives

$$|w_{ij}(x,\omega)|^2 \le \frac{4\alpha(n+1)}{\beta^2\omega^{2(m-n)}} \le \frac{4\alpha(n+1)}{\beta^2\omega^2} \tag{3.14}$$

for all $x \in \Xi$ and all $\omega \ge \bar{\omega}$.

The constants α, β and $\bar{\omega}$ all depend on the values of i and j. However, if we define

$$\alpha^* = \max_{i,j} \{\alpha\}, \tag{3.15a}$$

$$\beta^* = \min_{i,j} \{\beta\} \tag{3.15b}$$

and

$$\bar{c} = \max_{i,j} \{\omega\} \tag{3.15c}$$

then, from (3.5) and (3.14),

$$\mathrm{tr}\{W^*(x,\omega)W(x,\omega)\} \le \frac{4q^2\alpha^*(n+1)}{(\beta^*)^2\omega^2} \tag{3.16}$$

for all $x \in \Xi$ and all $\omega \ge \bar{\omega}$. This proves the Lemma.

<u>Theorem 3.1.</u> Under assumptions (1) and (2), there exists a c $\in (0,\infty)$, independent of $x \in \Xi$, such that

$$\max_{\omega \ge 0} \{\bar{\lambda}(x,\omega)\} = \max_{0 \le \omega \le c} \{\bar{\lambda}(x,\omega)\} \tag{3.17}$$

for each $x \in \Xi$.

<u>Proof.</u> Clearly,

$$\bar{\lambda}(x,\omega) \le \mathrm{tr}\{W^*(x,\omega)W(x,\omega)\}.$$

By Lemma 3.1 and assumption (2), we note that

$$\min_{x \in \Xi} \max_{\omega \ge 0} \{\bar{\lambda}(x,\omega)\} = \bar{\lambda}(x,\omega^*) > 0.$$

Let

$$\bar{\lambda}(x^*,\omega^*) = a. \tag{3.18}$$

Hence, by Lemma 3.2, there exist positive constants a and $\bar{c}$ such that

$$\bar{\lambda}(x,\omega) \le \kappa/\omega^2$$

for all $\omega > \bar{c}$ and $x \in \Xi$. With c defined by

$$c = \max\{\bar{c}, \sqrt{K/a}\} \tag{3.19}$$

we then have

$$\bar{\lambda}(x,\omega) \le a \tag{3.20}$$

for all $\omega > c$ and $x \in \Xi$. Combining (3.18) and (3.20) we obtain the conclusion of the theorem.

Following from Theorem 3.1 we now consider the problem:

$$\min_x \|\bar{\sigma}(x,\omega)\|_{p,[0,c]} = \min_x \left\{\int_0^c \{\bar{\lambda}(x,\omega)\}^p \, d\omega\right\}^{1/2p} \tag{3.21a}$$

subject to the constraints

(i) $x \in \Xi$; (3.21b)

(ii) $h_j(x) \le 0$, $j = 1,2,\ldots,N$; (3.21c)

(iii) $\max_{\omega \in \Omega} \phi_j(x,\omega) \le 0$, $j = 1,\ldots,M$; (3.21d)

where p is a positive integer and c is a positive constant, both to be defined later.

For brevity, let $\mathcal{F}$ be the subset of Ξ such that the constraints (3.21b), (3.21c) and (3.21d) are satisfied. With this abbreviation, the problem (3.21) may be re-stated as:

Find a parameter vector $x \in \mathcal{F}$ such that the objective function (3.21a) is minimized over $\mathcal{F}$.

This re-stated problem will be referred to as the problem (P_p).

<u>Theorem 3.2.</u> Let assumptions (1) and (2) be satisfied and let the constant c in the problem (P_p) be defined as in (3.19). Define

$$m_p = \min_{x \in \mathcal{F}} \phi_p(x),$$

where

$$\phi_p(x) = \left\{\int_0^c \{\bar{\lambda}(x,\omega)\}^p \, d\omega\right\}^{1/2p}. \tag{3.22}$$

Furthermore, for each positive integer p, let $x^p \in \mathcal{F}$ be such that $m_p = \phi_p(x^p)$. Then, there exists a subsequence $\{p(i) : i = 1,2\ldots\}$ of the sequence $\{p : p = 1,2\ldots\}$ such that

(i) $x^{p(i)} \to x^*$, as $i \to \infty$;

(ii) x^* is an optimal parameter of the problem (P); and

(iii) $m_{p(i)} \downarrow m$, as $i \to \infty$, where $m = \min_{x \in \mathcal{F}} \phi(x)$

with

$$\phi(x) = \max_{\omega \ge 0} \bar{\lambda}(x,\omega) = \max_{0 \le \omega \le c} \bar{\lambda}(x,\omega) .$$

<u>Proof.</u> By Theorem 3.1, we have

$$\max_{\omega \ge 0} \lambda(x,\omega) = \max_{0 \le \omega \le c} \lambda(x,\omega)$$

for any $x \in \mathcal{F}$. Now, we note that $\phi_p(x) \uparrow \phi(x)$, as $p \uparrow \infty$ for each $x \in \mathcal{F}$, where $\mathcal{F}$ is a compact subset of $\mathbb{R}^\ell$. Thus, by Dini's theorem, we have $\phi_p(x) \uparrow \phi(x)$, as $p \uparrow \infty$ uniformly with respect to $x \in \mathcal{F}$. Furthermore, for all $x \in \mathcal{F}$ and for all positive integers p, $\phi_p(x) \le \phi(x)$.

For each p, let $x^p \in \mathcal{F}$ be such that

$$\min_{x \in \mathcal{F}} \phi_p(x) = \phi_p(x^p).$$

Then, $\phi_p(x^p) \le \phi(x)$ for all $x \in \mathcal{F}$ and for all positive integers p. Therefore, $\phi_p(x^p) \le m$.

Next, we recall that $\mathcal{F}$ is compact. Thus, there exists a subsequence $\{p(i) : i = 1,2,\ldots\}$ of the sequence $\{p : p = 1,2,\ldots\}$, such that

$$\lim_{i \to \infty} x^{p(i)} = x^*$$

with $x^* \in \mathcal{F}$. This implies conclusion (i) of the Theorem.

We shall establish conclusions (ii) and (iii) together. Recall that $x^* \in \mathcal{F}$. Thus,

$$\phi(x^*) \ge m.$$

Furthermore, it is clear that

$$\phi_{p(i)}(x^{p(i)}) \to \phi(x^*), \text{ as } i \to \infty.$$

Hence,

$$\lim_{i\to\infty} \phi_{p(i)}(x^{p(i)}) \ge m.$$

But,

$$\lim_{i\to\infty} \phi_{p(i)}(x^{p(i)}) \le m.$$

Thus we obtain

$$\lim_{i\to\infty} \phi_{p(i)}(x^{p(i)}) = \phi(x^*) = m.$$

This, in turn, implies conclusions (ii) and (iii). Hence, the proof is complete.

Let

$$f_p(x) = \int_0^c \{\bar{\lambda}(x,\omega)\}^p \, d\omega \,. \tag{3.23}$$

$f_p(x)$ can then be considered as the objective function for the optimization problem (P_p). The function $\bar{\lambda}(x,\omega)$ is a non-differentiable function of x with the points of non-differentiability corresponding to two or more eigenvalues coming together. This has often lead to the development of specialised non-smooth optimization algorithms for mimimizing functions of this type. By assumption (2), it follows that the eigenvalues of $W^*(x,\omega)W(x,\omega)$ are distinct except possibly at discrete values of ω. Therefore, the functions $f_p(x)$ are differentiable functions for all $x \in \Xi$. This means that standard optimization software can be used to minimize $f_p(x)$. The gradient of $f_p(x)$ is given by

$$\nabla_x f_p = p\int_0^c \{\bar{\lambda}(x,\omega)\}^{p-1} \nabla_x \bar{\lambda}(x,\omega) \, d\omega \tag{3.24}$$

where

$$\nabla_x \bar{\lambda} = 2 \text{ Real} \left(\bar{v}^* W^* \nabla_x W \bar{v} \right) \tag{3.25}$$

with $\bar{v}$ being the right singular vector corresponding to the maximum singular value $\bar{\sigma}$ of $W(x,\omega)$.

For the non-generic case when assumption (2) fails to hold, it appears that subgradient optimization techniques must be employed as in generalizing [5] to [6]. Nongeneric cases are of mathematical interest and for coping with possible ill-conditioning when singular values are close over a frequency range, and will be the subject of another paper.

Approximation of Functional Constraints

Consider the functional constraints (3.21d). Clearly, they are equivalent to

$$G_j(x) = \int_\Omega g_j(x,\omega)d\omega = 0, \quad j = 1,\ldots,M,$$

where

$$g_j(x,\omega) = \max(\phi_j(x,\omega),0),$$

and $\Omega = (0,\infty)$. Since $G_j(x)$ is non-smooth in x, standard routines would have difficulty coping with these constraints. However, by using the smoothing technique suggested in [7], we define

$$g_{j,\epsilon}(x,\omega) = \begin{cases} 0 & \text{if } \phi_j(x,\omega) \le -\epsilon \,, \\ (\phi_j(x,\omega) + \epsilon)^2/4\epsilon & \text{if } |\phi_j(x,\omega)| < \epsilon \,, \\ \phi_j(x,\omega) & \text{if } \phi_j(x,\omega) \ge \epsilon \,, \end{cases}$$

and

$$G_{j,\epsilon}(x) = \int_\Omega g_{j,\epsilon}(x,\omega) \, d\omega \,. \tag{3.26}$$

Notice that

$$\lim_{\epsilon\to 0} g_{j,\epsilon}(x,\omega) = g_j(x,\omega), \quad \lim_{\epsilon\to 0} G_{j,\epsilon}(x) = G_j(x).$$

Let $\mathcal{F}_{\epsilon,\tau}$ be the subset of Ξ such that the constraints (3.21b) and (3.21c) together with the constraint

$$G_{j,\epsilon}(x) \le \tau \,. \tag{3.27}$$

With these definitions, a further approximation to the approximate problem (P_p) can be defined as:

Find a parameter vector $x \in \mathcal{F}$ such that the objective function (3.21a) is minimized over $\mathcal{F}_{\epsilon,\tau}$.

For each $\epsilon > 0$ and $\tau > 0$, let the corresponding approximate problem be denoted by $(P_{p,\epsilon,\tau})$.

In [7], it is assumed that the following conditions are satisfied:

(A) Ω is a compact interval in R;

(B) $\phi_j(x,\omega)$ is continuously differentiable in x and ω for all j. In applying the techniques of [7] to achieve a solution to the constrained H^∞ optimization task, it is important to note that the proofs of [7] can be trivially generalized to allow a relaxation of conditions (A) and (B) as:

(C) $\int_\Omega g_{j,\epsilon}(x,\omega) \, d\omega$ exists for all x for each j;

(D) $\dfrac{\partial\phi_j(x,\omega)}{\partial\omega}$ is piecewise continuous in $\omega \in \Omega$ for each x and j;

(E) $\phi_j(x,\omega)$ is continuously differentiable with respectto x for almost all ω and all j.

Under the constraint conditions (A) and (B), it is proposed in [7] that approximate problems parametrized in terms of ϵ, τ, be solved for decreasing ϵ, τ until a suitable approximation to the optimal x is found. The following theorem shows that the key theoretical result in [7] remains valid under the relaxed constraint conditions (C) to (E).

Theorem 3.2. Consider the problem (P_p) for a particular positive integer p. Then, under the relaxed constraint conditions (C) to (E), an arbitrarily close approximation to the optimal x^p for the problem (P_p) is achieved by solving the approximate problem $(P_{p,\epsilon,\tau})$ for ϵ, τ suitably small. Moreover, for each $\epsilon > 0$, the approximate solution satisfies the constraints $G_{j,\epsilon}(x) \le \tau$ for $\tau > 0$ suitably small. [There exists a $\tau(\epsilon) > 0$ such that for all $0 < \tau < \tau(\epsilon)$ the approximate solution is feasible.]

Proof. See relevant parts of [7].

Remarks.

(i) For each p, typically 3 or 4 approximate problems $(P_{p,\epsilon,\tau})$ have to be solved to achieve an "optimal" x^p for the corresponding problem (P_p).

(ii) If $\phi_j(x,\omega)$ represents singular values, then it is known that $\phi_j(x,\omega)$ is not continuously differentiable in x and ω when multiple singular values occur, so that the constraint condition (B) can fail.

Lemma 3.3. Let $\phi_j(x,\omega)$ represent singular values. Then, The relaxed constraint conditions (C) to (E) hold if and only if assumptions (1), (2) and (3) are satisfied.

Proof. It is straightforward to show that (C) $\Leftrightarrow$ (1), (D) $\Leftrightarrow$ (2), and {(D) and (E)} $\Leftrightarrow$ {(2) and (3)}.

For the nongeneric case when (E) fails, as for the function $f_p(x)$ defined in (3.23), it appears that subgradient based optimization techniques must be employed as in generalizing [5] to [6].

Derivatives of the integrals $G_{j,\epsilon}(x)$ with respect to x are obtained by making use of (3.25).

4. Computational Aspects and an Example

Consider initially problem (P) in which there are no functional constraints. As described in the previous section, we choose a sequence of positive integers p and solve approximate problems (P_p) which are:

$$\min_x \|\bar{\sigma}(x,\omega)\|_{p,[0,c]} = \min_x \phi_p(x) = \left\{\min_x f_p(x)\right\}^{1/2p} \tag{4.1}$$

with $f_p(x)$ defined by (3.23) as

$$f_p(x) = \int_0^c \{\bar{\lambda}(x,\omega)\}^p \, d\omega . \tag{4.2}$$

The point at which each minimum is attained will be denoted by x^p and so the minimum value of the corresponding objective function will be $\phi_p(x^p)$. The solution procedure for problem (P) consists of the following steps.

(I) Choose a positive integer p(1) (a suggested value is p(1) = 5) and a value of c. Set i = 1.

(II) Use a standard optimization algorithm to solve the problem ($P_{p(i)}$) and obtain $x^{p(i)}$. For each value of x the objective function $f_p(x)$ and its derivative with respect to x can be calculated as follows:

(a) For each $\omega \in [0,c]$ the value of $\bar{\lambda}(x,\omega)$ and $\nabla_x \bar{\lambda}(x,\omega)$ can be obtained as follows:

(i) Use a complex singular value decomposition to obtain the singular values and right singular vectors of $W(x,\omega)$.

(ii) Calculate $\bar{\lambda}(x,\omega) = \{\bar{\sigma}(x,\omega)\}^2$.

(iii) Calculate $\nabla_x \bar{\lambda}(x,\omega)$ from equation (3.25).

(b) Use numerical integration to calculate $f_p(x)$ and $\nabla_x f_p(x)$ from (3.23) and (3.24).

(III) Choose a value of p(i+1) > p(i) and use $x^{p(i)}$ as the initial point in the next optimization. Set i = i+1 and go to step (II).

Remarks

(i) If a complex singular valued decomposition is not available then form $W^*(x,\omega)W(x,\omega)$ and compute its eigenvalues to obtain $\bar{\lambda}(x,\omega)$. The eigenvectors of $W^*(x,\omega)W(x,\omega)$ are then the right singular vectors of $W(x,\omega)$.

(ii) In practice c does not have to be very large. For example, c = 10 will usually suffice, although this does depend on the particular problem.

(iii) Typically, only 3 or 4 different values of p will be required. For example, {p(i)} = {5,20,100} will usually give a good result. It is a relatively simple task to compute the ∞-norm for a particular value of x^p, that is, $\phi(x^p)$, and compare it with the value of the corresponding p-norm, $\phi_p(x^p)$. By Theorem 3.2, the value of $\phi_p(x^p)$, and hence $\phi(x^p)$, converges to the solution of problem (P) as $p \to \infty$.

(iv) For larger values of p, the integrand in (4.2) will need to be scaled so as to avoid numerical overflow. One means of accomplishing this is to scale the integrand by the square of $\phi(x^p)$, where x^p is the previous solution point, so then

$$\phi_{p(i)}(x) = \phi(x^{p(i-1)})\left\{\int_0^c \left[\frac{\bar{\lambda}(x,\omega)}{\{\phi(x^{p(i-1)})\}^2}\right]^p d\omega\right\}^{1/2} \tag{4.3}$$

where $\phi(x^{p(0)})$ is set to 1. The integrand in (4.3) now has a maximum of approximately 1 and so overflow is avoided. A check on the size of the integrand must be kept, however, to avoid underflow.

In the several examples we have tried the previously described computational procedure works extremely well. Admittedly we have yet to solve any particularly complicated problems, but we do not anticipate any severe difficulties.

Example. As a simple application of the procedure we consider the 2x2 transfer function matrix

$$W(x,s) = \begin{bmatrix} \frac{1}{(s+2)^2} & \frac{1}{2s^2-s+1} \\ \frac{1}{s^2-s+1} & \frac{1}{(s+1)^2} \end{bmatrix} + \frac{1}{(s+1)^2}\begin{bmatrix} x_1 & x_2 \\ x_3 & x_4 \end{bmatrix} \tag{4.4}$$

where $s = j\omega$ and $x \in \Xi \subset R^4$. The singular values of W are shown in Fig. 4.1(a) for the case x = 0. The ∞-norm of W was then minimized without any functional constraints using the sequential quadratic programming software NLPQL [12]. The starting value of x was chosen to be the zero vector. The results are summarised in Table 4.1 for p taking on the values 5, 20, 100, and 200. Fig. 4.1(b) shows the resulting singular values of W for the optimal value of x corresponding to p = 200.

The functional constraint

$$\hat{\sigma}(x,\omega) \le 0.8 \quad \text{for all } \omega \in \Omega, \tag{4.5}$$

where $\hat{\sigma}(x,\omega)$ is the minimum singular value of W, was then introduced and the problem re-solved using the same software. The resulting optimal value of the ∞-norm of W, corresponding to p = 200, was 1.86 compared to 1.41 for the unconstrained case. The corresponding singular values of W are shown in Fig. 4.2.

5. Acknowledgements

The authors wish to thank Dr. B.D. Craven for his help in the formulation and proof of Theorem 3.2 and Dr. D.J.Clements for his many helpful comments.

6. References

1. Francis, B.A., A Course in H^∞ Control Theory, Springer Verlag, 1987
2. Glover, K., All Optimal Hankel-norm Approximations of Linear Multi- variable Systems and Their L^∞-error Bounds, International Journal of Control, Vol. 39, pp.1115-1193, 1984.
3. Doyle, J.C., Glover, K., Khargonekar, P., and Francis, B., State-Space Solutuions to Standard H^2 and H^∞ Control Problems, IEEE American Control Conference, Atlanta, June 1988.
4. Limebeer, D.J.N., Kasenally, L.M., and Safonov, M.G., A Characterization of All Solutions to the Four Block General Distance Problem, submitted for publication.
5. Polak, E., and Wardi, Y., Nondifferentiable Optimization Algorithm for Designing Control Systems Having Singular Value Inequalities, Automatica, Vol. 18, pp.267-283, 1982.
6. Polak, E., and Mayne, D. Q., On the Solution of Singular Value Inequalities Over a Continuum of Frequencies, IEEE Trans. Automat. Control, Vol. AC-26, pp.690-695, 1981.
7. Jennings, L.S. and Teo, K.L., A Computational Algorithm for Functional Inequality Constrained Optimization Problems, to appear in Automatica, 1989.
8. Polak, E., and Mayne, D.Q., An Algorithm for Optimization Problems with Functional Inequality Constraints, IEEE Trans. Automat. Contr., Vol. AC-21, pp.184-193, 1976.
9. Schittkowski, K, NLPQL: A FORTRAN Subroutine for Solving Constrained Nonlinear Programming Problems, Operations Research Annals, Vol. 5, pp.485-500, 1985.

p	x^p				$\phi_p(x^p)$	$\phi(x^p)$
5	-0.257	0.185	0.165	-2.00	1.299	1.446
20	-0.253	0.292	-0.05	-1.97	1.359	1.418
100	-0.259	0.349	-0.05	-1.99	1.396	1.413
200	-0.259	0.355	-0.05	-1.99	1.403	1.413

Table 4.1. Values of the p-norm and the ∞-norm for the example.

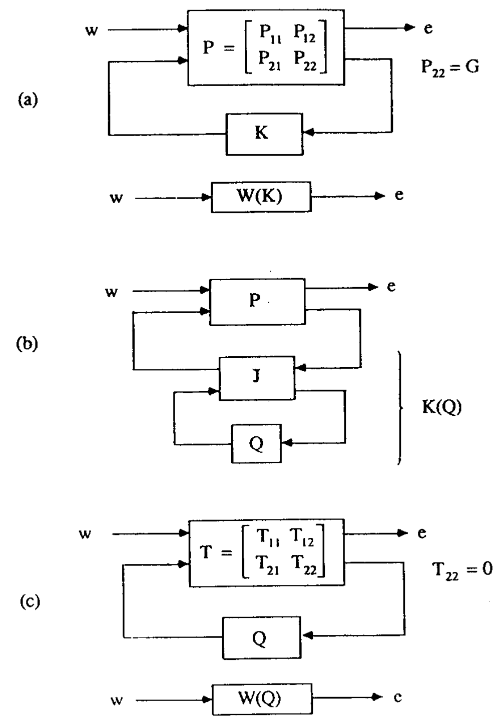


Figure 2.1 Stabilizing Controller Schemes.

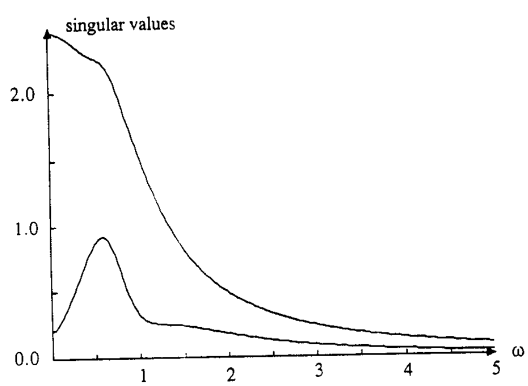
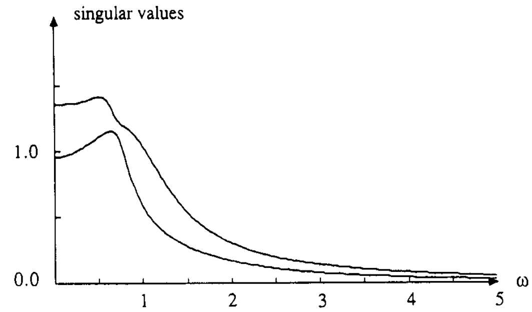


Figure 4.1 Singular values of $W(x,s)$ for

(a) $x = [0,\ 0,\ 0,\ 0]^T$;



(b) $x = [-0.259,\ 0.349,\ 0.0525,\ 1.99]^T$.

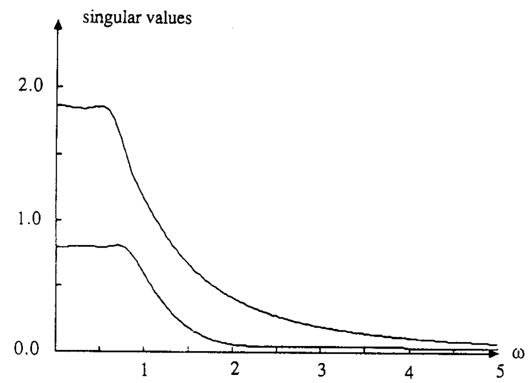


Figure 4.2 Singular values of $W(x,s)$ for the constrained case:

$x = [-0.130,\ 0.309,\ 0.0135,\ -3.15]^T$.