

STRUCTURED UNCERTAINTY IN CONTROL SYSTEM DESIGN

John C. Doyle

Honeywell SRC and Caltech

**Abstract-** This paper reviews control system analysis and synthesis techniques for robust performance with structured uncertainty in the form of multiple unstructured perturbations and parameter variations. The structured singular value,  $\mu$ , plays a central role. The case where parameter variations are known to be real is considered.

1. Introduction

This paper will review some basic methods for analyzing the performance and robustness properties of feedback systems. The particular approach taken here is from [1]-[6] which builds on results by many other researchers.

The general framework to be used in this paper is illustrated in the diagram in Figure 1. Any linear interconnection of inputs, outputs, commands, perturbations, and a controller can be rearranged to match this diagram. For the purpose of analysis the controller may be thought of as just another system component and the diagram reduces to that in Figure 2. The analysis problem involves determining whether the error  $e$  remains in a desired set for sets of inputs  $v$  and perturbations  $\Delta$ . The interconnection structure  $G$  can be partitioned so that the transfer function from  $v$  to  $e$  can be expressed as the linear fractional transformation

$$e = F_u(G, \Delta) v = [G_{22} + G_{21}\Delta(I - G_{11}\Delta)^{-1} G_{12}] v.$$

In Section 2 only uncertainty in  $v$  will be considered so this will be referred to as *nominal performance* to indicate that  $\Delta = 0$ . Nominal performance will be seen to be equivalent to a norm test on  $G_{22}$ . The main focus of this paper will be on  $\|\bullet\|_\infty$ . Section 3 considers stability in the presence of perturbations. This will be referred to as *robust stability* with robust used here to indicate that the property of stability is maintained under perturbations. For simple unstructured perturbations, this also leads naturally to a  $\|\bullet\|_\infty$  norm test, but now on  $G_{11}$ . The  $\|\bullet\|_\infty$  norm thus provides a single norm which handles both nominal performance and robust stability. Unfortunately, norm bounds are inadequate in dealing with more realistic models of plant uncertainty involving structure and more complicated mathematical objects involving the structured singular value,  $\mu$ , are required.

The methods outlined in Sections 2 and 3 allow for assessing either nominal performance or robust stability. Obviously, it would be desirable to treat performance with both noise and perturbations occurring simultaneously. Section 4 considers this problem and shows that this also leads to tests using  $\mu$ , but now involving the entire transfer function  $G$ . Thus  $\mu$  emerges as an essential analysis tool in dealing with *robust performance* as well as with structured perturbations.

The mathematical properties and computation of  $\mu$  are briefly taken up in Sections 5 for the case of complex perturbations and 6 for the real case. Here  $\mu$  is viewed as a natural generalization of both spectral radius and spectral norm, and this viewpoint leads to useful characterizations of  $\mu$  in terms of these more familiar quantities. One consequence is that estimates for  $\mu$  can be obtained by scaling of ordinary singular values. The implications of this approach for synthesis are also briefly considered in Section 7.

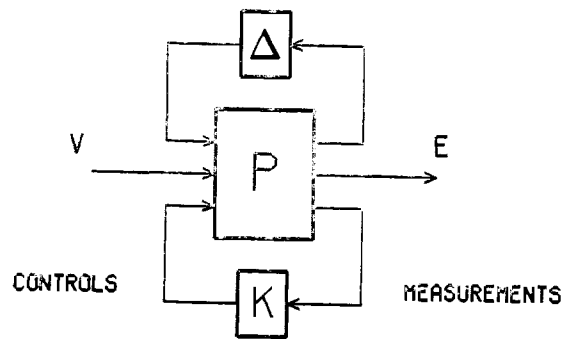


FIGURE 1

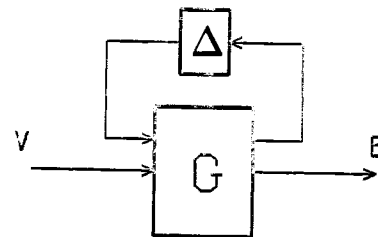


FIGURE 2

2. Nominal Performance

This section considers performance in terms of bounds on  $e$  in the presence of uncertain bounded inputs  $v$ . Bounds for both  $v$  and  $e$  are expressed in terms of signal power or energy. The terms power and energy are used here in a generalized sense to indicate that integrals of the square of the signals are involved. Suppose that  $v$  is a function of time such that on any finite interval it is square integrable. Then we may obtain bounds on  $v$  in terms of:

$$(1) \text{ Power: } BP = \left\{ v \mid \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T \|v(t)\|^2 dt < 1 \right\}$$

$$(2) \text{ Energy: } BL_2 = \left\{ v \mid \|v\|_2^2 = \int_{-\infty}^{\infty} \|v(t)\|^2 dt < 1 \right\}$$

The prefix  $B$  denotes the unit ball. The bounds are scaled to 1 since any other scaling can simply be absorbed into the interconnection structure  $G$ . Likewise, any weighting or coloring filter can be absorbed into  $G$  so that only unweighted signals need be considered. Note that in practice, the use of weightings on both  $v$  and  $e$  are essential to reflect the varying spatial and frequency content of both the input signals and the performance specifications.

The relationship between  $v$  and  $e$  is expressed by the following

Performance Theorem (Power/Power):

$$e \in BP \text{ for all } v \in BP \\ \text{iff } \|G_{22}\|_{\infty} \leq 1$$

where

$$\|G\|_{\infty} = \sup_{\omega} \bar{\sigma}\{G(j\omega)\}$$

The same theorem holds for  $v$  and  $e$  in  $BL_2$ . This theorem is a trivial restatement of the induced norm but focuses attention on the use of  $\|\bullet\|_{\infty}$  as a test for performance.

### 3. Robust Stability

In this section, we will consider plant perturbations, a type of uncertainty entirely different from uncertain input signals. Since plant perturbations can destabilize a nominally stable system, the first issue to be addressed is robust stability. In what follows, it makes no difference whether  $\Delta$  is a constant complex, rational, or real-rational matrix so for simplicity it will be assumed constant complex. Stability will be taken to mean that the perturbed system has no closed rhp poles. Under these assumptions, we have the following simple and well-known theorem ([12],[13]):

**Theorem RSU** (Robust Stability, Unstructured):

$$\text{Stable for all } \Delta, \bar{\sigma}(\Delta) < 1 \\ \text{iff } \|G_{11}\|_{\infty} \leq 1$$

The term unstructured refers to the fact that  $\Delta$  is assumed to be bounded but otherwise unknown. Typically weights are used when modeling plant uncertainty to reflect the frequency and spatial variation of the perturbations. These weights can always be absorbed into the nominal interconnection structure so in that sense it is no loss of generality to assume a uniform norm bound on  $\Delta$ . It is in the assumption that no structural information is available for  $\Delta$  that limits the usefulness of Theorem RSU. In practical problems, it is generally the case that the uncertainty consists of parameter variations and *multiple* norm-bounded perturbations. Using only a single norm-bounded perturbation for analysis is rarely adequate. Parameter variations typically arise because of uncertain coefficients in differential equation models of physical systems and involve real scalars. Norm-bounded perturbations often arise when trying to capture the effect of unmodeled dynamics and are themselves dynamic systems. This would typically lead to norm-bounded real-rational perturbations, but for analysis, it is sufficient to instead consider constant complex matrix perturbations.

Any linear interconnection of inputs, outputs, transfer functions, parameter variations, and perturbations can be rearranged to fit the diagram in figure 2, where  $\bar{\sigma}(\Delta) < 1$  but  $\Delta$  is block-diagonal. This is an obvious consequence of the fact that composition of linear fractional transformations are linear fractional, and it holds for perturbations to transfer functions as well as elements of state-space realizations. Reducing to the uniform norm bound typically requires the absorption into the nominal interconnection structure of scalings and weights. Then  $\Delta$  will be a member of a set like

$$\mathbf{\Delta} = \{ \text{diag}(\delta_1, \delta_2, \dots, \delta_m, \Delta_1, \Delta_2, \dots, \Delta_n) \\ \mid \delta_i \in \mathbf{R}, \Delta_j \in \mathbf{C}^{k_j \times k_j} \} \quad (3.1)$$

or its bounded subset:

$$\mathbf{B}\mathbf{\Delta} = \left\{ \Delta \in \mathbf{\Delta} \mid \bar{\sigma}(\Delta) < 1 \right\}. \quad (3.2)$$

It is possible to define more general sets involving, for example, repeated perturbations, and these will be considered in Section 5. Nonsquare perturbations can easily be handled in what follows by augmenting the interconnection structure with rows or columns of zeros.

Given  $\Delta \in \mathbf{B}\mathbf{\Delta}$  Theorem RSU could be used to obtain sufficient conditions for robust stability, but the test could be arbitrarily conservative. That is, it is easy to construct examples where  $\|G_{11}\|_{\infty}$  can be made arbitrarily large but no  $\Delta \in \mathbf{B}\mathbf{\Delta}$  leads to instability. In order to obtain a precise generalization of Theorem RSU to handle structured uncertainty, we need the structured singular value,  $\mu$  [2]. The positive real-valued function  $\mu$  satisfies the property

$$\det(I - M\Delta) \neq 0 \text{ for } \forall \Delta \in \mathbf{\Delta}, \bar{\sigma}(\Delta) < \gamma \\ \text{iff } \gamma\mu(M) \leq 1. \quad (3.3)$$

Note that  $\mu$  is a function of  $M$  that depends on the structure of the  $\Delta$ 's in  $\mathbf{\Delta}$ . This dependency is typically not represented explicitly. If  $\mu(M) \neq 0$ , (i.e.  $\exists \Delta \in \mathbf{\Delta}$  such that  $\det(I - M\Delta) = 0$ ) then

$$\frac{1}{\mu(M)} = \min_{\Delta \in \mathbf{\Delta}} \left\{ \bar{\sigma}(\Delta) \mid \det(I - M\Delta) = 0 \right\}. \quad (3.4)$$

Unfortunately, (3.4) is not typically very useful in computing  $\mu$  since the implied optimization problem is cumbersome and can have multiple local maxima which are not global. Computation of  $\mu$  is a complicated problem and some results will be given in Sections 5 and 6. For now, assume  $\mu$  is the function defined above.

With these definitions, the correct generalization of Theorem RSU to structured uncertainty is

**Theorem RSS** (Robust Stability, Structured)

$$\text{Stable for all } \Delta \in \mathbf{B}\mathbf{\Delta} \\ \text{iff } \|G_{11}\|_{\mu} \leq 1$$

where

$$\|G\|_{\mu} \stackrel{\text{def}}{=} \sup_{\omega} \mu\{G(j\omega)\}. \quad (3.5)$$

Note that  $\|G\|_{\mu}$  is not actually a norm, but the notation is convenient. Note also that it depends not only on  $G$  but also the assumed structure of  $\mathbf{\Delta}$ .

#### 4. Robust Performance

The methods outlined in the previous two sections allow for analyzing either nominal performance or robust stability. Obviously, it would be desirable to treat performance with both noise and perturbations occurring simultaneously [3]. The following theorem addresses exactly this problem.

**Theorem RP:**

$$F_u(G, \Delta) \text{ stable and } \|F_u(G, \Delta)\|_\infty < 1 \quad \forall \Delta \in \mathbf{B}\Delta$$

$$\text{iff } \|G\|_\mu \leq 1$$

where  $\mu$  is taken w.r.t. the structure

$$\tilde{\Delta} = \{\tilde{\Delta} = \text{diag}(\Delta, \Delta_{n+1}) \mid \Delta \in \mathbf{B}\Delta\}.$$

This theorem is the real payoff for using  $\mu$ . It's made possible by the equivalence of performance and robust stability when using  $\|\cdot\|_\infty$ . The block  $\Delta_{n+1}$  may be thought of loosely as a "performance block" used to turn the performance condition into a robust stability condition and finally into a test using  $\mu$ . Note that  $\mu$  is computed for the full  $G$  and is taken with respect to an augmented structure.

#### Analysis Summary

Performance	Stability	$e \in BP$
Perturbation		
$\Delta = 0$	No $C_+$ poles	$\ G_{22}\ _\infty \leq 1$
$\bar{\sigma}(\Delta) < 1$	$\ G_{11}\ _\infty \leq 1$	$\ G\ _\mu \leq 1$
$\Delta \in \mathbf{B}\Delta$	$\ G_{11}\ _\mu \leq 1$	$\ G\ _\mu \leq 1$

#### 5. $\mu$ For Complex Perturbations

In the previous sections, it was shown that robust performance and stability with structured uncertainty reduces to computing  $\mu$  for constant matrices  $G(j\omega)$  and then taking sup over all  $\omega$ . For this to be useful, we must have ways of computing  $\mu$  or bounds for it. This section will begin by outlining some of the mathematical properties of  $\mu$  for complex perturbations and viewing it as a natural generalization of the spectral radius  $\rho$ , and the spectral norm (maximum singular value)  $\bar{\sigma}$ . The rest of this section will focus on using scalings to characterize  $\mu$  in terms of  $\rho$  and  $\bar{\sigma}$ .

Suppose that  $\Delta$  is some subalgebra of matrices satisfying

$$\{\lambda I \mid \lambda \in \mathbf{C}\} \subset \Delta \subset \mathbf{C}^{N \times N}. \quad (5.1)$$

In this paper we will be interested in block diagonal  $\Delta$ . Define the spectrum,  $sp(M)$ , and inverse spectrum,  $isp(M)$ , of a matrix  $M \in \mathbf{C}^{N \times N}$  with respect to the subalgebra  $\Delta$  as

$$\begin{aligned} sp(M) &= \left\{ \Delta \in \Delta \mid \det(M - \Delta) = 0 \right\} \\ isp(M) &= \left\{ \Delta \in \Delta \mid \det(I - M\Delta) = 0 \right\}. \end{aligned} \quad (5.2)$$

Since both sets depend on  $\Delta$  it would be appropriate to subscript the symbols, but to keep notation simple this will be avoided throughout. The set  $sp(M)$  is a natural generalization of the usual notion of spectrum and is always nonempty. In this context,  $\mu$  can be viewed as a natural generalization of spectral radius since it is easily verified that

$$\mu(M) = \sup_{\Delta \in sp(M)} \underline{\sigma}(\Delta). \quad (5.3)$$

If  $\mu(M) \neq 0$  (which is equivalent to  $isp(M) \neq \emptyset$ ) then

$$\mu(M) = \sup_{\Delta \in isp(M)} \frac{1}{\bar{\sigma}(\Delta)}. \quad (5.4)$$

This characterization emphasizes the view of  $\mu$  as a generalization of  $\bar{\sigma}$  and is simply a restatement of (3.4). Indeed, in the special cases where  $\Delta$  is equal to one of its possible extreme sets in (5.1),  $\mu$  is exactly either the usual spectral radius or maximum singular value:

$$\begin{aligned} \Delta = \{\lambda I \mid \lambda \in \mathbf{C}\} &\Rightarrow \mu(M) = \rho(M) \\ \Delta = \mathbf{C}^{N \times N} &\Rightarrow \mu(M) = \bar{\sigma}(M) \end{aligned} \quad (5.5)$$

It is possible to use these two special cases to obtain bounds for  $\mu$ . For any set  $\Delta$  it easy to see that

$$\rho(M) \leq \mu(M) \leq \bar{\sigma}(M) \quad (5.6)$$

but these bounds are not directly useful for computation as matrices may be found that make the differences between the bounds and  $\mu$  as large as desired.

It is possible to improve the bounds in (5.6) by using simple properties of  $\Delta$ . Suppose that  $\mathcal{U}$  and  $\mathcal{D}$  are sets such that for any  $\Delta \in \Delta$

$$\begin{aligned} U \in \mathcal{U} &\Rightarrow \bar{\sigma}(U\Delta) = \bar{\sigma}(\Delta) \\ D \in \mathcal{D} &\Rightarrow D^{-1}\Delta D = \Delta. \end{aligned} \quad (5.7)$$

Then it is easy to see from the definition of  $\mu$  that

$$\begin{aligned} U \in \mathcal{U} &\Rightarrow \mu(MU) = \mu(M) \\ D \in \mathcal{D} &\Rightarrow \mu(DMD^{-1}) = \mu(M) \end{aligned} \quad (5.8)$$

so the bounds in (5.6) can be improved to

$$\sup_{U \in \mathcal{U}} \rho(MU) \leq \mu(M) \leq \inf_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1}). \quad (5.9)$$

The key theorems about  $\mu$  show when these inequalities are actually equalities.

Let us first consider the case where all the blocks are complex and none are repeated. Then we have the sets

$$\begin{aligned} \Delta &= \left\{ \text{diag}(\Delta_1, \Delta_2, \dots, \Delta_n) \mid \Delta_j \in \mathbf{C}^{m_j \times m_j} \right\} \\ \mathcal{U} &= \left\{ \text{diag}(U_1, U_2, \dots, U_n) \mid U_j^* U_j = I \right\} \\ \mathcal{D} &= \left\{ \text{diag}(d_1 I, d_2 I, \dots, d_n I) \mid d_i \in \mathbf{R}_+ \right\} \end{aligned} \quad (5.10)$$

It is easy to verify that (5.7) holds so that the inequalities in (5.9) apply. What is more important is that

$$\sup_{U \in \mathcal{U}} \rho(MU) = \mu(M) \quad (5.11)$$

holds for all  $M$  and  $\mathbf{\Delta}$  and

$$\mu(M) = \inf_{D \in \mathcal{D}} \bar{\sigma}(DMD^{-1}) \quad (5.12)$$

if  $n \leq 3$  (three or fewer blocks) or if  $M$  is real [2]. There are other conditions under which this upper bound is an equality but they are more cumbersome to state and generally of less interest. An example of strict inequality for the upper bound has been found for  $n = 4$ . Extensive computational experimentation has yet to find a matrix for which the upper bound exceeds  $\mu$  (actually some lower bound for  $\mu$ ) by more than 15%, and the upper bound is nearly equal for most matrices. This seems to be independent of matrix size and number of blocks. This is encouraging but additional theoretical work is needed to guarantee the quality of the upper bound in general.

The case of repeated blocks is less well understood. To see what  $\mathcal{U}$  and  $\mathcal{D}$  arise when there are repeated blocks consider the simple case where each block is a repeated scalar

$$\begin{aligned} \mathbf{\Delta} &= \left\{ \text{diag}(\delta_1 I, \delta_2 I, \dots, \delta_n I) \mid \delta_j \in \mathbf{C} \right\} \\ \mathcal{U} &= \left\{ \text{diag}(u_1 I, u_2 I, \dots, u_n I) \mid u_j \in \mathbf{C}, |u_j| = 1 \right\} \\ \mathcal{D} &= \left\{ \text{diag}(D_1, D_2, \dots, D_n) \mid D_j \text{ invertible} \right\} \end{aligned} \quad (5.13)$$

It is possible to restrict the  $D \in \mathcal{D}$  to positive definite Hermitian matrices ( $D_j = D_j^* > 0$ ) without loss of generality. As above, the inequalities in (5.9) hold and (5.11) also holds for all  $M$  and  $\mathbf{\Delta}$ . Unfortunately, it is not known under what conditions the upper bound is also an equality. The computational experience with the case of repeated blocks is much more limited than with nonrepeated blocks, but the evidence so far suggests that the upper bound is also nearly an equality. The case of repeated nonscalar blocks is just the obvious combination of the above two cases.

The lower bounds in terms of  $\rho(MU)$  have the desirable property of always achieving  $\mu$  independent of the number of blocks. Unfortunately,  $\rho(MU)$  can have multiple local maxima which are not global so direct computation of (5.11) by gradient search may not find the actual maximum. At this time there is no alternative scheme guaranteed to find the global maximum that has reasonable computational properties. Fan and Tits [14] do have an alternative scheme for a lower bound which does not guarantee that  $\mu$  will be found but appears to be very fast and has many advantages over using (5.11).

The upper bound in (5.9) is more easily found since the expression  $\bar{\sigma}(DMD^{-1})$  has only global minima. This is a direct consequence of the fact that  $\bar{\sigma}(e^D M e^{-D})$  is convex in  $D$ . This fact was used in [2] to argue that the upper bound in (5.9) probably offered a reasonable alternative to (5.11) for computation of  $\mu$ . The original proof of convexity was rather cumbersome and appeared later in [15].

Computational experience to date has indicated that it is desirable in practice to use both upper and lower bounds for  $\mu$ , since the existing bounds nicely complement each other. The upper bound is easily computed but may not give  $\mu$  except in special cases. On the other hand, it appears to be nearly equal to  $\mu$  in all cases. The existing lower bounds (including both (5.11) and those in [14]) are, in principle, equal to  $\mu$  in all cases but may fail because of local maxima. By having an upper bound it is much easier to recognize when a local maxima is not global and restart the algorithm with another initial guess. Extensive computational experience has yet to reveal a (complex)  $\mu$  problem where the bounds obtained in this way differed by more than about 15%. More research is needed to show whether this is always true. It could simply turn out that counterexamples exist but are difficult to find.

## 6. Computation of $\mu$ for Real Perturbations

The properties of  $\mu$  when  $\mathbf{\Delta}$  has some elements restricted to be real are quite different from the purely complex case. Suppose that

$$\begin{aligned} \mathbf{\Delta} &= \{ \text{diag}(\delta_1, \delta_2, \dots, \delta_m, \Delta_1, \Delta_2, \dots, \Delta_n) \\ &\quad \mid \delta_i \in \mathbf{R}, \Delta_j \in \mathbf{C}^{k_j \times k_j} \} \end{aligned} \quad (6.1)$$

and  $sp(M)$  and  $isp(M)$  are defined for  $\mathbf{\Delta}$  exactly as in (5.2). In this case, it is possible for either  $sp(M)$  or  $isp(M)$  to be the empty set. Furthermore, (5.3) is no longer a correct characterization of  $\mu$  in general and there is no natural way to view  $\mu$  as a simple generalization of the usual notion of spectral radius. Of course, (5.4) still applies provided  $\mu(M) \neq 0$  (i.e.  $isp(M) \neq \emptyset$ ).

This section will focus on upper bounds to  $\mu$  that can be obtained by scaling  $\bar{\sigma}$ . The choice of scaling is based on the following lemma which characterizes a useful class of scalings. In the following lemma, assume that

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix}$$

and  $\det(I - T_{22}M) \neq 0$  so that

$$F_l(T, M) = T_{11} + T_{12} M (I - T_{22}M)^{-1} T_{21}$$

is well-defined.

*Lemma:*

$$\begin{aligned} \text{Suppose } \exists T \text{ such that } \mathbf{B}\mathbf{\Delta} \subset \left\{ F_u(T, \Delta) \mid \bar{\sigma}(\Delta) < 1 \right\} \\ \text{then } \bar{\sigma}(F_l(T, M)) \leq 1 \Rightarrow \mu(M) \leq 1 \end{aligned}$$

This lemma says that if  $F_u(T, \Delta)$  "covers"  $\mathbf{B}\mathbf{\Delta}$  then  $T$  can be used to obtain an upper bound for  $\mu$ . The next step is to identify a set of  $T$ 's that satisfy the lemma. To this end define

$$\begin{aligned} \mathcal{D} &= \{ \text{diag}(d_1, d_2, \dots, d_n, d_{m+1}I, d_{m+2}I, \dots, d_{m+n}I) \\ &\quad \mid d_i \in \mathbf{R}_+ \} \\ \mathcal{C} &= \{ \text{diag}(c_1, c_2, \dots, c_m, 0, 0, \dots, 0) \mid c_i \in [-1, 1] \} \\ \mathcal{T} &= \left\{ \left[ \begin{array}{cc} jC & (I - C^2)^{1/2} D \\ D^{-1} & 0 \end{array} \right] \mid D \in \mathcal{D}, C \in \mathcal{C} \right\} \end{aligned} \quad (6.2)$$

where  $\mathcal{D}$  and  $\mathcal{C}$  are partitioned conformally with  $\mathbf{\Delta}$ . With these definitions, for  $T \in \mathcal{T}$

$$F_l(T, M) = jC + (I - C^2)^{1/2} DMD^{-1}. \quad (6.3)$$

It is a matter of some simple algebra to show that all  $T \in \mathcal{T}$  satisfy the lemma. Note that if there are no real parameters ( $m=0$ ), then  $F_l(T, M) = DMD^{-1}$  and this scaling reduces to that considered in (5.10) and (5.12).

Obtaining an upper bound based on (6.3) is somewhat more complicated than is possible in (5.12). The difficulty is that the above lemma only implies that  $\mu(M) \leq 1$  and does not scale. Using the sets in (6.2), we can define

$$\bar{\mu}(M) = \inf_{\alpha \in \mathbb{R}_+} \left\{ \alpha \mid \inf_{T \in \mathcal{T}} \bar{\sigma} \left( F_l(T, \frac{1}{\alpha} M) \right) < 1 \right\}. \quad (6.4)$$

It follows immediately from the lemma that

$$\mu(M) \leq \bar{\mu}(M)$$

and thus  $\bar{\mu}$  provides an upper bound for  $\mu$ . Again note that for no real parameters,  $\bar{\mu}(M)$  simplifies to

$$\bar{\mu}(M) = \inf_{D \in \mathcal{D}} \bar{\sigma} (DMD^{-1})$$

The natural question is how good a bound is  $\bar{\mu}$  for  $\mu$ . Recall that for  $m = 0$ ,  $n \leq 3$  that  $\bar{\mu} = \mu$  for all matrices independent of block size ( $m$  is the number of real parameters and  $n$  is the number of complex blocks). A simple extension of this result yields  $\bar{\mu} = \mu$  when  $m = 1$ ,  $n \leq 2$ . Although counterexamples exist for problems with more than these number of blocks, experience has shown that  $\bar{\mu}$  is often a good approximation to  $\mu$  even in these cases. While this experience is encouraging it is not conclusive and additional research is needed to establish the value of  $\bar{\mu}$ . Unfortunately, when there is more than one real parameter it is possible for  $\mu(M) \ll \bar{\mu}(M)$ .

## 7. $\mu$ -Synthesis

The previous sections on analysis showed that the synthesis problem reduces to finding a stabilizing controller  $K$  so that

$$\|F_l(P, K)\|_{\alpha} \leq 1 \quad \alpha = \infty \text{ or } \mu \quad (7.1)$$

where  $F_l(P, K) = P_{11} + P_{12} K(I - P_{22}K)^{-1} P_{21}$ . The  $P$  here is a real-rational, proper transfer function matrix and is not necessarily stable. It will be assumed throughout that  $P_{12}$  has at least as many rows as columns, and vice versa for  $P_{21}$ .

A complete solution to the synthesis problem for the  $\infty$ -norm was recently obtained ([1],[3]-[6]), which removed the previous restrictions that  $P_{12}$  and  $P_{21}$  be square. Another feature of this solution is that it provided an efficient computational scheme using standard real matrix operations on state-space representations. This  $H_{\infty}$ -synthesis solution can be used to provide an approach to solving the  $\mu$ -norm synthesis problem, referred to as  $\mu$ -synthesis. These results will be briefly reviewed before considering the achievable performance results.

The first step in the  $H_{\infty}$  synthesis solution involves finding  $J$  so that the substitution  $K = F_l(J, Q)$  yields

$$F_l(P, K) = F_l(P, F_l(J, Q)) = R + UQV \quad (7.2)$$

with  $F_l(P, K)$  internally stable iff  $Q \in H_{\infty}$ . This is a version of the so-called Youla parametrization [16]. Further,  $U$  is inner and  $V$  co-inner ( $U^*U = I$  and  $VV^* = I$ ), and there exist complementary inner factors  $U_{\perp}$  and  $V_{\perp}$  such that  $[UU_{\perp}]$  and  $[\begin{smallmatrix} V \\ V_{\perp} \end{smallmatrix}]$  are both square and inner. The  $U$  and  $V$  are obtained from coprime factorizations  $P_{12} = UM_1^{-1}$  and  $P_{21} = M_2^{-1}V$ .

The next step involves using a rational matrix version of the Davis-Kahan-Weinberger matrix dilation results [17] to further reduce the problem to one of finding  $\hat{Q} \in RH_{\infty}$  such that

$$\|G + \hat{Q}\|_{\infty} \leq 1 \quad (7.3)$$

where  $G \in RL_{\infty}$ . This problem can then be solved using the Hankel norm approximation methods developed by Glover [18]. The resulting optimal  $\hat{Q}$  can then be used to find first the optimal  $Q$  and then the optimal  $K$ .

The  $\mu$ -synthesis problem does not yet have as complete a solution as does the  $H_{\infty}$  synthesis problem. A reasonable approach

would be to try to find a stabilizing controller  $K$  and scaling  $D$  so that

$$\|DF_l(P, K)D^{-1}\|_{\infty} \leq 1. \quad (7.4)$$

One method to do this is to alternately minimize the above expression for either  $K$  and  $D$  while holding the other constant. For fixed  $D$  the left-hand side of (7.4) is just an  $H_{\infty}$  control problem and can be solved using the methods reviewed above. For fixed  $K$ , the left-hand side of (7.4) can be minimized at each frequency as a convex optimization problem in  $D$ . The resulting  $D$  can be fit with a stable, rational transfer function with stable inverse (the phase of  $D$  does not affect the norm).

This approach to  $\mu$ -synthesis has been successfully applied to several example problems. In principle, it could be used to obtain controllers that are arbitrarily close to  $\mu$ -optimal in the case of 3 or fewer blocks and provide nearly optimal controllers for the general case. This would depend on the suggested iterative scheme converging to the global optimal  $K$  and  $D$ . Unfortunately, individual convexity in the two parameters of an optimization problem does not imply joint convexity, and this scheme is not always guaranteed to converge globally to the best  $K$  and  $D$ .

To better understand the properties of the problem in (7.4) it is useful to consider the constant matrix problem. Using (7.2), we can reduce (7.2) to

$$\|D(R + UQV)D^{-1}\|_{\infty} \leq 1. \quad (7.5)$$

for constant  $R, U, V$  with  $U^*U = 1$  and  $VV^* = 1$ . For  $D = I$ , it follows from [18] that

$$\min_Q \bar{\sigma}(R + UQV) = \max(\bar{\sigma}(U_{\perp}^*R), \bar{\sigma}(RV_{\perp}^*)) \quad (7.6)$$

where  $U_{\perp}^*$  and  $V_{\perp}^*$  are chosen so that  $[U \ U_{\perp}]$  and  $[\begin{smallmatrix} V \\ V_{\perp} \end{smallmatrix}]$  are both square and unitary. All of these quantities are easily computed using standard SVD routines.

Posing (7.5) as an optimization problems gives

$$\min_{D, Q} \bar{\sigma} (D(R + UQV)D^{-1}). \quad (7.7)$$

It is known that this problem is convex in either  $D$  (actually  $\ln(D)$ ) or  $Q$  individually when the other is held fixed, but is not convex in both variables jointly. This means that the iterative scheme suggested as a possible approach to  $\mu$ -synthesis is not guaranteed to converge even in the constant matrix case. It is possible, however, to compute the desired  $D$  in (7.7) directly.

The result in (7.6) may be applied to (7.8) to obtain

$$\begin{aligned} \min_Q \bar{\sigma}(D(R + UQV)D^{-1}) = \\ \max \{ \bar{\sigma}((DU)_\perp^* DR), \bar{\sigma}(RD^{-1}(VD^{-1})_\perp^*) \} \end{aligned} \quad (7.8)$$

where

$$(DU)_\perp \triangleq D^{-1}U_\perp(U_\perp^*D^{-2}U_\perp)^{-1/2} \quad (7.9)$$

and  $(VD^{-1})_\perp$  is defined similarly. Note that  $[DU(U^*D^2U)^{-1/2} (DU)_\perp]$  is unitary. It can be shown that the right hand side of (7.8) is convex in  $\ln(D)$  so that the "optimal" scaling for (7.7) may be computed by search in advance. This gives a tight lower bound for (7.7) and the resulting  $D$  scaling may be used to compute the optimal  $Q$ .

A simple example will illustrate all the essential features of this possibly confusing sequence of ideas. Consider the problem

$$\min_q \mu \left( \begin{bmatrix} -1 & 1 \\ q & 1 \end{bmatrix} \right) = \min_{q,d} \bar{\sigma} \left( \begin{bmatrix} -1 & d \\ q/d & 1 \end{bmatrix} \right). \quad (7.10)$$

The  $\mu$ -optimal  $q$  is  $q = 0$  which gives  $\mu = 1$ . For fixed  $d$  the  $\bar{\sigma}$ -optimal  $q = d^2$  and for fixed  $q > 0$  the  $\bar{\sigma}$ -optimal  $d$  is  $d = \sqrt{q}$ . Thus, iteratively solving for either  $q$  or  $d$  will immediately converge to the curve  $q = d^2$ . For example, with the initial guess of  $q = d = 1$ , the iterative scheme will not change either  $q$  or  $d$  and will thus fail to find the global optimum.

On the other hand,

$$\begin{aligned} \min_q \bar{\sigma} \begin{bmatrix} -1 & d \\ q/d & 1 \end{bmatrix} = \max \left( \bar{\sigma}([-1 \ d]), \bar{\sigma} \begin{bmatrix} d \\ 1 \end{bmatrix} \right) \\ = \sqrt{1 + d^2}. \end{aligned} \quad (7.11)$$

Thus,

$$\min_d \left( \min_q \bar{\sigma} \begin{bmatrix} -1 & d \\ q/d & 1 \end{bmatrix} \right) = \min_d \sqrt{1 + d^2} \quad (7.12)$$

which is clearly convex and achieves its minimum as  $d \rightarrow 0$ . If the expression in (7.8) were used to compute the  $d$  in advance, it would be possible to find the optimal achievable level for (7.9). This example also illustrates why, strictly speaking, inf, not min must be used for the  $D$  scalings as in (5.9). This issue will not be taken up in this paper. It turns out to be of little significance anyway.

The simplest application of these ideas to the selection of the  $D$  scalings for the  $\mu$ -synthesis problem is to compute an initial guess for  $D$  at each frequency using (7.8). This would be the optimal  $D$  for an acausal controller, and should provide a good initial guess for the optimal  $D$  for the causal controller problem. A deeper question is whether some generalization of (7.8) and its convexity properties applies to the rational case. While this seems likely, the details have not been worked out and the practical implications are uncertain. For some additional results on  $\mu$ -synthesis, see [19].

## Acknowledgements

The results in this paper were inspired by many stimulating discussions with colleagues. I would particularly like to thank my coworkers at Honeywell Systems and Research Center, and also thank E. Zafiriou, P. Khargonekar, M. Morari, S. Boyd, W. Helton, B. Francis, D. Sarason, and M. Safonov. This work was supported by Honeywell Internal Research and Development Funding, the Office of Naval Research under ONR Research Grant N00014-82-C-0157, and the U.S. Air Force of Scientific Research F49620-82-C-0090.

## References

- [1] J.C. Doyle, *Lecture Notes*, 1984 ONR/Honeywell Workshop on Advances in Multivariable Control
- [2] J.C. Doyle, "Analysis of feedback systems with structured uncertainty," *IEE Proceedings, Part D*, V129, No. 6, Nov., 1982
- [3] J.C. Doyle, J.E. Wall, and G. Stein, "Performance and robustness analysis with structured uncertainty," 1982 CDC
- [4] J.C. Doyle, "Synthesis of robust controllers and filters," 1983 CDC
- [5] C.C. Chu and J.C. Doyle, "On inner-outer and spectral factorizations," 1984 CDC
- [6] C.C. Chu and J.C. Doyle, "The general distance problem in  $H_\infty$  synthesis," 1985 CDC
- [7] I.M. Horowitz, *Synthesis of Feedback Systems*, 1963
- [8] G. Zames, "Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms and approximate inverses," *IEEE Trans. Auto. Control*, AC-26, 1981
- [9] G. Zames and B.A. Francis, "Feedback, minimax sensitivity, and optimal robustness," *IEEE Trans. Auto. Control*, AC-28, April, 1983
- [10] M. Vidyasagar, to appear
- [12] G. Zames, "On the input-output stability of nonlinear time-varying feedback systems, parts I and II," *IEEE Trans. Auto. Control*, AC-11, 1966
- [13] J.C. Doyle and G. Stein, "Multivariable feedback design: concepts for a classical/modern synthesis," *IEEE Trans. Auto. Control*, AC-26, Feb., 1981
- [14] M. Fan and A. Tits, "Characterization and efficient computation of the structured singular value," submitted to *IEEE Trans. Auto. Control*
- [15] M.G. Safonov and J.C. Doyle, "Minimizing conservativeness of robustness singular values," in S.G. Tzefestas (ed.), *Multivariable Control*, 1984, D. Reidel Publishing Company
- [16] D.C. Youla, H.A. Jabr, and J.J. Bongiorno, Jr., "Modern Wiener-Hopf design of optimal controllers: Part II," *IEEE Trans. Auto. Control*, AC-21, 1976
- [17] C. Davis, W.M. Kahan, and H.F. Weinberger, "Norm-preserving dilations and their applications to optimal error bounds," *SIAM J. Num. Anal.*, 1982
- [18] K. Glover, "All optimal Hankel-norm approximations of linear multivariable systems and their  $L^\infty$ -error bounds," *Int. J. Control*, 1984
- [19] M.G. Safonov, "Optimal diagonal scaling for infinity norm optimization," 1985 ACC