Robustness with Parametric and Dynamic Uncertainty

Thesis by

Peter M. Young

In Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

California Institute of Technology

Pasadena, California

1993

(Submitted May 10, 1993)

To Pat, who got there first.

Acknowledgements

Whilst at Caltech I have had the pleasure of meeting, and working with, a number of very talented people. They have managed to make my stay here both productive and enjoyable. First of all I would like to thank my advisor, John Doyle, for his help on both counts. John has an insight into the "big picture," and a disdain for inelegant methods, that make him a remarkable researcher and a unique advisor. His guidance and support have been invaluable.

I would also like to thank the other members of my committee: Munther Dahleh, Joel Franklin, Manfred Morari, and Richard Murray, for their help in many ways at various times. I have enjoyed collaborating on a number of projects with visitors and fellow graduate students at Caltech, and in that regard I wish to express my gratitude to Gary Balas, Karl Åström, Matt Newlin, Richard Braatz, and Jorge Tierno for their input to the research, and for making those projects fun.

Various colleagues have influenced my work, and helped during the course of this research: I would like to thank Andy Packard, André Tits, Kemin Zhou, Roy Smith, and Steve Boyd for numerous helpful discussions. I am very grateful to Keith Glover, Michael Fan, and Ragu Balakrishnan for lending me some of their expertise, and their software. Finally I would particularly like to thank Haniph Latchman, who got me interested in robust control in the first place.

Many people have helped to put this manuscript into its final form, and of those not already mentioned I would like to thank Carolyn Beck and Raff D'Andrea, for agreeing to read the early versions, and providing many useful comments. I would also like to thank the other members of the Controls Group at Caltech for sharing their views and their friendship with me. Of course it would not have been the same without Friday night at the Athenaeum, and I would like to thank the other pool players not already mentioned, John Morris and Bobby Bodenheimer, for many a good game.

Finally I would like to offer my most heartfelt gratitude to my wife, Patricia. She has been more to me than I could have ever hoped for, and certainly more than I could try to put into words. All my thanks, and all my love, to Pat.

Abstract

In many disciplines of engineering it is often convenient, for analysis and design purposes, to approximate the real behavior of physical systems by mathematical models. For some applications however, and in particular when one wishes to design a high performance controller, the differences between the behavior of the mathematical model and the physical system can be crucial to the performance of the final design. The theory of robust control attempts to take into account these inherent inaccuracies in the model, and provide systematic analysis and design techniques in the face of this "uncertainty."

These goals can be restated as formal mathematical problems. In order to handle more realistic descriptions of physical systems, one has to allow more sophisticated models, and this leads to more difficult mathematical problems. In this thesis we will consider both the theoretical and computational aspects of such problems. In particular we will consider robustness in the presence of *both* real (e. g., parametric) and complex (e. g., dynamic) structured uncertainty.

This leads to a consideration of the general mixed μ analysis and synthesis problems. Some special cases of the analysis problem can be solved exactly, but the general problem is in fact NP hard, so that in order to develop solutions for large problems with reasonable computational requirements, we will adopt a scheme of computing and refining upper and lower bounds. By exploiting the theoretical properties of the problem, we are able to develop practical algorithms, capable of handling mixed μ analysis problems with tens of parameters, in computation times that are typically of the order of minutes. This is despite the fact that the mixed μ problem appears to have inherently combinatoric worst-case behavior.

For the synthesis problem a new "D,G-K iteration" procedure is developed to design a stabilizing controller which attempts to minimize the peak value across frequency of mixed μ. The scheme utilizes a combination of some new results from the mixed μ upper bound problem with the $\mathcal{H}_\infty$ optimal control solution. The theoretical results developed here have already been successfully applied to a number of real engineering problems, and some of these applications are briefly reviewed, to illustrate the advantages offered by the new analysis and synthesis techniques.

Table of Contents

List of Figures

List of Tables

Chapter 1

Introduction

Real physical systems cannot be exactly described by mathematical models. Nevertheless it is often convenient for analysis and design purposes to approximate the real behavior of physical systems by mathematical models, and this technique is employed in many disciplines of engineering. For many purposes these nominal models are adequate, and the engineer may design an appropriate controller on that basis.

However in some circumstances, and in particular for high performance systems, it is possible that a control design may perform well in simulation on the nominal model, and not even be stable when implemented on the real physical system. The problem arises from the fact that the model is not sufficiently accurate in some aspect, and the high performance controller is attempting to exploit knowledge of the system that is incorrect! The theory of robust control attempts to take into account these inherent inaccuracies in the model, and provide systematic analysis and design techniques in the face of this "uncertainty." Moreover, since the primary motivation is for high performance controllers, we wish to quantify in which aspects, and to what extent, the model may not be relied upon. In this way the control design may still exploit the knowledge of the system that is reliable.

In order to restate these goals as mathematical problems, one considers a nominal model subject to a perturbation. Given a property of the nominal model, we say that

it is "robust" if it holds not only for the nominal model, but also for the model subject to all allowed values of the perturbation. Two properties that we will be especially concerned with are stability and performance, where we will measure performance via the maximum gain from disturbance to error of the closed loop system. These definitions will all be made rigorously in chapter 2, but for now we can see that these concepts lead us naturally to two basic types of robustness questions. Firstly we have the analysis question: *Given a closed loop model and perturbation set, is the stated property robust?* Then there is the more difficult synthesis question: *Given an open loop model and perturbation set, can we design a controller so that the closed loop system meets the required robustness condition?* This thesis will be largely concerned with answering these two questions.

The above definitions are fairly standard in the robust control community. Note however that one may consider a variety of possible assumptions to place on the prior knowledge we have about the perturbation, and this leads to a number of different robustness paradigms. In order to handle more realistic descriptions of physical systems, one has to allow for more sophisticated models and perturbation sets, and this leads to more difficult mathematical problems.

In order to pose a meaningful problem, we assume that the perturbation is (norm) bounded. Our robustness measure will then be in terms of the *smallest* perturbation for which the stated property fails to hold. We will consider a perturbation set, or uncertainty description, which allows for structured perturbations. In particular we will consider perturbations which may be block diagonal, with any number of blocks. In fact this description is very general, and we will see in chapter 2 that a great many robustness problems can be rearranged to fit into this framework. Problems involving many different perturbations, entering different components of the model, can be rearranged into one large block diagonal perturbation to the whole system. This set-up leads us to consider the structured singular value, μ, which is a mathematical function measuring robustness with respect to such (structured) uncertainty.

We will consider briefly the history of the structured singular value, and some related concepts, in section 1.1. For now we simply note that the initial research in this field has concentrated on the complex μ problem, where it is assumed that all the blocks of the perturbation may assume complex values. This paradigm works well for many problems, and in particular it is appropriate when the perturbation is being used to represent unmodeled dynamics, which are assumed stable and bounded, but otherwise unknown.

In recent years a great deal of interest has arisen with regard to problems involving parametric uncertainty. These parameters may be coefficients in a differential equation, or may represent physically meaningful quantities, such as masses, aerodynamic coefficients, etc. In any case the uncertainty is naturally modeled with perturbations which are not only norm bounded, but also constrained to be real. Allowing these real perturbations in the uncertainty description leads us to consider mixed μ problems, where the perturbation may contain *both* real and complex blocks.

The presence of these real blocks substantially complicates the μ problem, so that the techniques developed to deal with complex μ problems are no longer adequate. In this thesis we will develop methods to address both the analysis and synthesis questions for the mixed μ problem. This will enable us to consider robustness problems for systems subject to both parametric and dynamic uncertainty.

In order to tackle the mixed μ problem we need to develop a substantial amount of mathematical and computational machinery, drawing on results from linear algebra, complex analysis, functional analysis and computational complexity theory among others. We briefly outline how this development will proceed in section 1.2, but it is important to bear in mind that, despite the level of mathematical abstraction, this theoretical research has a solid engineering motivation. In fact the techniques we will present in this thesis are currently being utilized at various industrial and academic sites, and a number of successful applications of these tools to practical engineering problems have already appeared in the literature.

1.1 Historical Review

This section is *not* intended to give a comprehensive review of the development of the structured singular value, and associated robustness concepts. The intention is merely to review some of the key ideas most relevant to this research, so as to place the present work in context. It is always possible to trace back a given piece of work to earlier influences, but an appropriate starting point for this review is the small gain theorem introduced by Zames [84] in 1965. This provided an exact robust stability test with respect to unstructured dynamic uncertainty.

A number of robust stability results were obtained for this type of problem, allowing for a single norm bounded perturbation (see, e. g., Willems [77] (1971), and Safonov [67] (1980)). This work heralded the use of singular values as an important tool in robust control. These exact results for unstructured uncertainty provide sufficient conditions for robust stability with respect to structured uncertainty. It became apparent however that for many problems these results are too crude, and it is necessary to exploit the structure of the problem to get a less conservative condition.

Several researchers looked at this problem, and the notion of rearranging the problem into block diagonal form and then using scaling matrices to reduce the conservatism gained acceptance (see [66] (1978) for an early example). The (complex) structured singular value, μ, was introduced by Doyle [23] in 1982 as a systematic means of dealing with such problems. This paper introduced the use of *both* upper and lower bounds for such problems, and stressed the need for viable methods of computation. In addition it was shown that robust performance problems could also be handled within the same framework.

The complex μ theory was subsequently extended considerably by a number of researchers. An important element of this research was the emphasis on efficient computation schemes for upper and lower bounds, rather than exact computation. Fan and Tits looked in some detail at these problems [28] (1986), and Packard developed a lower bound for the complex μ problem [59] (1988). In addition the complex μ syn-

thesis problem was also studied, and Doyle proposed the "D-K iteration" procedure [24] (1985) as a means to tackle the problem (see also Stein and Doyle [74] (1991)). This theoretical research lead to the release of the μ-Tools toolbox [3,7] by Balas et al. in 1991. This commercially available software package contains the algorithms necessary to implement (a large part of) the complex μ analysis and synthesis techniques, and these methods are now routinely applied to large engineering problems.

At the same time analysis problems involving only parametric uncertainty were also being considered, but from a somewhat different viewpoint. The idea was to consider polynomials with perturbed coefficients, where the perturbations are restricted to be real. The starting point for this research was Kharitonov's celebrated result on interval polynomials [37] (1979). Kharitonov developed an exact robust stability check for a problem with real parametric uncertainty. This result sparked enormous interest in this approach, and a number of results for this type of problem were obtained, such as the edge theorem by Bartlett et al. [11] in 1988. These results provided a variety of exact robust stability tests (see [71] (1989) and [10] (1993) for reviews). Unfortunately the problem class studied was very restrictive, so that the results were of limited value for engineering applications. Furthermore the solutions typically involved checking the vertices or edges of some polytope in the parameter space, so that they become computationally intractable for large problems.

Several researchers adopted an algorithmic approach to these analysis problems, seeking to tackle them by computing and refining upper and lower bounds. The use of Branch and Bound was suggested by de Gaston and Safonov [19] (1988) and Sideris and Sánchez Peña [70] (1989). This enabled less restrictive uncertainty descriptions to be considered, but the problem of exponential growth in computation remained.

The possibility of considering both parametric and dynamic uncertainty in the μ framework was suggested by Doyle [24] (1985). The starting point for the research in this thesis is given in [29] (1991), where Fan et al. consider the mixed μ analysis problem, and develop an upper bound.

1.2 Organization of the Thesis

Following the introductory material in chapter 1, the thesis begins with a formal definition of the mixed μ problem in chapter 2. In addition the concept of Linear Fractional Transformations is introduced, and this is used to show how many types of robustness problems can be re-cast in terms of computing mixed μ.

The properties of the mixed μ problem itself are considered in chapter 3. It is shown that many of the basic properties of complex μ generalize easily to the mixed case. However there are important fundamental differences between mixed and complex μ, and in this regard the properties of continuity and computational complexity are considered in some detail (see [58,16]). In particular we find that the mixed μ problem is NP hard. This has profound implications for the viability of certain approaches to tackling the mixed μ problem, and motivates the approach taken for the remainder of the thesis.

Since we wish to come up with tractable solutions for large problems, these results motivate us to look for approximate solutions, and in chapter 4 we present upper and lower bounds for the mixed μ problem. The mixed μ problem is first re-cast as a real eigenvalue maximization problem, and the lower bound is derived by a consideration of the conditions required for a local maximum. This leads to a power iteration for a lower bound. The upper bound is from [29], and takes the form of a linear matrix inequality minimization problem.

The properties of these bounds are examined in some detail in chapter 5, and in particular we consider under what circumstances one can guarantee that μ equals its upper bound. This material is rather technical, and probably only of interest to those already familiar with the corresponding results for the complex μ case, although in fact the machinery used here for the more general mixed case is simpler than the earlier approaches for the complex case (see [56]). This chapter also considers some particular special cases of mixed μ, where one can get stronger results than for the general problem.

Chapter 6 is concerned with the practical computation of the theoretical bounds from chapter 4. The bounds are reformulated so as to be amenable to efficient computation, and algorithms are developed to do the job. This computational software is now commercially available, and we present some results from our extensive numerical experience with the algorithms. In the course of this development we note that the upper bound has an alternative formulation as a singular value minimization problem. This gives an interesting interpretation of the upper bound in terms of a small gain type condition, and turns out to be very useful later for the synthesis problem.

Further advances in the computational aspects of the analysis problem are considered in chapter 7. Some work on adaptive power iteration schemes [75] is briefly presented. The use of Branch and Bound schemes, to iteratively refine the bounds, is also considered [51]. These techniques give us the means to compute upper and lower bounds with as small a gap as desired. Of course the computational requirements become prohibitive if we ask for very small gaps on large problems (as the NP hardness results predict).

In chapter 8 we consider an important special case in detail: the rank one mixed μ problem. In fact one can convert a number of "Kharitonov-type" robustness problems into special cases of the rank one mixed μ problem. It is then shown that for this special case exact calculation is possible, since μ is identically equal to its upper bound (which is convex). This provides an interesting link between the μ and polynomial approaches, although we note that the results are of limited use for engineering purposes, because of the restrictive assumptions inherent to the problem.

Chapters 9 and 10 are concerned with the mixed μ synthesis problem. Some technical machinery, on State Space factorization theory, is developed in chapter 9. This is then used in chapter 10 to develop a "D,G-K iteration" procedure for the mixed μ synthesis problem. The development proceeds by combining the mixed μ upper bound results from chapter 6 with the solution to the $\mathcal{H}_\infty$ optimal control

problem (see [21]). The resulting procedure is an extension of the "D-K iteration" procedure for complex μ synthesis, and finds a stabilizing controller attempting to minimize the peak value across frequency of mixed μ.

Throughout this work an important emphasis is placed on developing solutions which can be efficiently computed. All the results presented have been implemented in software, and in chapter 11 we consider some applications of these results. In addition to some example problems, results are presented from a number of applications to real engineering problems. Examples of both simulation and experimental data are presented, along with the analysis and synthesis results. Finally, in chapter 12, we conclude the thesis with a look at some directions for future research in this area.

Chapter 2

The μ Analysis Framework

In this chapter we give a brief introduction to the structured singular value, μ, and its role in the robustness analysis of linear systems. The machinery of Linear Fractional Transformations, LFT's, is also briefly presented. These two topics are intimately related, and both have received a great deal of attention over the years. We will only present a few highlights from the panoply of research results obtained in these areas, and we refer the reader to [57] and the references therein for a more comprehensive review of these ideas. Note that until recently most of the research effort has focused on the complex μ problem (see [57] for example), rather than the more general mixed case we are concerned with here. However all of the general LFT machinery for μ problems carries through easily to the mixed case, and the results in this chapter will be presented for the mixed case.

2.1 Notation and Definitions

First of all we need to establish some notation, and some basic definitions, that will be used for the remainder of this thesis. The notation we will use is fairly standard and is essentially taken from [29] and [81]. In the following suppose that M is a complex matrix, and x a complex column vector:

$\mathbb{R}$	Field of real numbers
$\mathbb{R}^{n\times m}$	Real matrix with n rows and m columns
$\mathbb{R}^n$	Real column vector with n elements
$\mathbb{C}$	Field of complex numbers
$\mathbb{C}^{n\times m}$	Complex matrix with n rows and m columns
$\mathbb{C}^n$	Complex column vector with n elements
M^T	Transpose of matrix or vector M
M^*	Complex conjugate transpose of matrix or vector M
$\overline{\sigma}$	Largest singular value of a matrix
μ	Structured singular value of a matrix
ρ	Spectral radius of a matrix
ρ_R	Real spectral radius of a matrix
$\overline{\lambda}$	Largest (real) eigenvalue of a Hermitian matrix
λ_{min}	Smallest (real) eigenvalue of a Hermitian matrix
λ_k	k^{th} largest (real) eigenvalue of a Hermitian matrix
Tr	Trace of a square matrix
det	Determinant of a square matrix
Re	Real part of a matrix
Im	Imaginary part of a matrix
Ker	Kernel of a matrix
Arg	Argument of a complex scalar
Sgn	Sign of a real scalar
Co	Convex hull of a set

$\|M\|_F$	Frobenius norm of matrix or vector M: $\|M\|_F = \sqrt{\mathrm{Tr}\,(M^*M)}$
$\|x\|$	Euclidian norm of vector x: $\|x\| = \sqrt{x^*x}$
$\|x\|_\infty$	Infinity norm of a vector $x \in \mathbb{C}^n$: $\|x\|_\infty = \max_{1\le i\le n} \|x_i\|$
I_k	$k \times k$ identity matrix
0_k	$k \times k$ zero matrix
$0_{k\times r}$	$k \times r$ zero matrix
$\mathcal{RM}$	Space of real-rational proper transfer matrices
$\mathcal{RL}_\infty$	Subset of $\mathcal{RM}$ with no $\mathbf{j}w$ axis poles
$\mathcal{RH}_\infty$	Subset of $\mathcal{RM}$ with no poles in $\mathrm{Re}\,(s) \ge 0$
$\|P\|_\infty$	Infinity norm of a (stable) transfer matrix $P(s)$
$\doteq$	Is defined to be
$\longrightarrow$	Implies
$\longleftarrow$	Is implied by
$\longleftrightarrow$	Implies and is implied by
$\square$	Q.E.D.

Occasionally we will drop the subscripts from I and 0, whence they denote identity and zero matrices respectively, of the appropriate size. For a transfer matrix $P(s)$, then we denote its State Space representation by

$$P \;=\; \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \;\doteq\; C(sI - A)^{-1}B + D. \tag{2.1}$$

For some of the above terms a little more explanation is appropriate: the real spectral radius $\rho_R(M) = \max\{|\lambda| : \lambda \text{ is a real eigenvalue of } M\}$, with $\rho_R(M) = 0$ if M has no real eigenvalues. For a transfer matrix $P \in \mathcal{RH}_\infty$, then

$$\|P\|_\infty \doteq \sup_{w\in\mathbb{R}} \overline{\sigma}(P(\mathbf{j}w)). \tag{2.2}$$

The structured singular value, μ, is a measure of the robustness of a system, and is defined formally below.

The definition of μ is dependent upon the underlying block structure of the uncertainties, which is defined as follows. Suppose we have a matrix $M \in \mathbb{C}^{n\times n}$ and three non-negative integers m_r, m_c, and m_C (with $m \doteq m_r + m_c + m_C \leq n$) which specify the number of uncertainty blocks of each type. Then the block structure $\mathcal{K}(m_r, m_c, m_C)$ is an m-tuple of positive integers

$$\mathcal{K} = (k_1, \ldots, k_{m_r}, k_{m_r+1}, \ldots, k_{m_r+m_c}, k_{m_r+m_c+1}, \ldots, k_m). \tag{2.3}$$

This m-tuple specifies the dimensions of the perturbation blocks, and we require $\sum_{i=1}^{m} k_i = n$ in order that these dimensions are compatible with M. This determines the set of allowable perturbations, namely define

$$\begin{aligned} X_{\mathcal{K}} = \{\Delta = \text{block diag}\,(&\delta_1^r I_{k_1}, \ldots, \delta_{m_r}^r I_{k_{m_r}}, \delta_1^c I_{k_{m_r+1}}, \ldots, \delta_{m_c}^c I_{k_{m_r+m_c}}, \\ &\Delta_1^C, \ldots, \Delta_{m_C}^C) : \delta_i^r \in \mathbb{R}, \delta_i^c \in \mathbb{C}, \Delta_i^C \in \mathbb{C}^{k_{m_r+m_c+i}\times k_{m_r+m_c+i}}\}. \end{aligned} \tag{2.4}$$

Note that $X_{\mathcal{K}} \subset \mathbb{C}^{n\times n}$ and that this block structure is sufficiently general to allow for (any combination of) repeated real scalars, repeated complex scalars, and full complex blocks. The purely complex case corresponds to $m_r = 0$, and the purely real case to $m_c = m_C = 0$.

Note also that all the results which follow are easily generalized to the case where the full complex blocks need not be square, and the blocks may come in any order. In fact the analysis and synthesis software we will describe in subsequent chapters has been written for the more general case. Although this generalization adds little difficulty to the problem, it does makes the notation somewhat cumbersome, and so, for ease of presentation, we will restrict our attention to the set-up in (2.4).

Definition 2.1 (Structured Singular Value [23]) *The structured singular value, $\mu_{\mathcal{K}}(M)$, of a matrix $M \in \mathbb{C}^{n\times n}$ with respect to a block structure $\mathcal{K}(m_r, m_c, m_C)$ is defined as*

$$\mu_{\mathcal{K}}(M) = \left(\min_{\Delta\in X_{\mathcal{K}}} \{\overline{\sigma}(\Delta) : \det(I - \Delta M) = 0\}\right)^{-1} \tag{2.5}$$

with $\mu_{\mathcal{K}}(M) = 0$ *if no* $\Delta \in X_{\mathcal{K}}$ *solves* $\det(I - \Delta M) = 0$.

In order to develop some of the subsequent theoretical results about μ and it's bounds, we need to define some sets of block diagonal scaling matrices (which, like μ itself, are dependent on the underlying block structure).

$$\mathcal{Q}_{\mathcal{K}} = \{\Delta \in X_{\mathcal{K}} : \delta_i^r \in [-1\ 1], \delta_i^{c*}\delta_i^c = 1, \Delta_i^{C*}\Delta_i^C = I_{k_{m_r+m_c+i}}\} \tag{2.6}$$

$$\mathcal{U}_{\mathcal{K}} = \{U \in \mathcal{Q}_{\mathcal{K}} : U^*U = I_n\} \tag{2.7}$$

$$\hat{\mathcal{D}}_{\mathcal{K}} = \{\text{ block diag }(e^{\mathbf{j}\theta_1}D_1, \ldots, e^{\mathbf{j}\theta_{m_r}}D_{m_r}, D_{m_r+1}, \ldots, D_{m_r+m_c}, d_1 I_{k_{m_r+m_c+1}}, \ldots, d_{m_C} I_{k_m}) : \theta_i \in [-\frac{\pi}{2}\ \frac{\pi}{2}], 0 < D_i = D_i^* \in \mathbb{C}^{k_i \times k_i}, 0 < d_i \in \mathbb{R}\} \tag{2.8}$$

$$\tilde{\mathcal{D}}_{\mathcal{K}} = \{\text{ block diag }(D_1, \ldots, D_{m_r+m_c}, d_1 I_{k_{m_r+m_c+1}}, \ldots, d_{m_C} I_{k_m}) : D_i = D_i^* \in \mathbb{C}^{k_i \times k_i}, d_i \in \mathbb{R}\} \tag{2.9}$$

$$\mathcal{D}_{\mathcal{K}} = \{D \in \tilde{\mathcal{D}}_{\mathcal{K}} : D > 0\} = \{D \in \hat{\mathcal{D}}_{\mathcal{K}} : \theta_i = 0, i = 1, \ldots, m_r\} \tag{2.10}$$

$$\overline{\mathcal{D}}_{\mathcal{K}} = \{\text{ block diag }(D_1, \ldots, D_{m_r+m_c}, d_1 I_{k_{m_r+m_c+1}}, \ldots, d_{m_C} I_{k_m}) : \det(D_i) \neq 0, D_i \in \mathbb{C}^{k_i \times k_i}, d_i \neq 0, d_i \in \mathbb{C}\} \tag{2.11}$$

$$\mathcal{G}_{\mathcal{K}} = \{\text{ block diag }(G_1, \ldots, G_{m_r}, O_{k_{m_r+1}}, \ldots, O_{k_m}) : G_i = G_i^* \in \mathbb{C}^{k_i \times k_i}\} \tag{2.12}$$

$$\overline{\mathcal{G}}_{\mathcal{K}} = \{\text{ block diag }(g_1, \ldots, g_{n_r}, O_{n_c}) : g_i \in \mathbb{R}\} \tag{2.13}$$

where $n_r = \sum_{i=1}^{m_r} k_i$ and $n_c = n - n_r$.

2.2 Mixed μ and Linear Fractional Transformations

Two of the main reasons why μ has proved to be such a useful tool for robustness analysis are that it answers the question of well posedness of a Linear Fractional Transformation, and it satisfies a certain property called the "Main Loop Theorem"

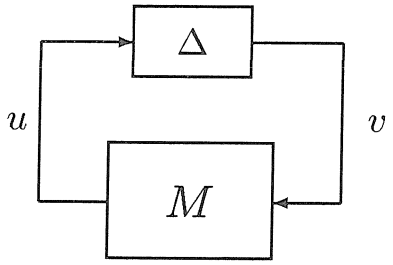


Figure 2.1: Robust stability μ analysis framework

(see [56]). These statements will be made more explicit later, but for now consider the constant matrix feedback interconnection in figure 2.1, where $\Delta \in X_{\mathcal{K}}$ represents a structured perturbation to the nominal matrix $M \in \mathbb{C}^{n\times n}$.
This feedback loop represents the loop equations

$$u = Mv$$

$$v = \Delta u.$$

Note that these equations give $(I_n - M\Delta)u = 0$ and $(I_n - \Delta M)v = 0$, so that we have a unique solution if and only if $I_n - M\Delta$ is non-singular, whence the solution is $u = v = 0$. If $I_n - M\Delta$ is singular then there are infinitely many solutions, with $|u|$ and $|v|$ arbitrarily large. In this case the feedback loop is not well posed, or is in some sense "unstable." Note that $I_n - M\Delta$ is guaranteed to be non-singular for all $\Delta \in X_{\mathcal{K}}$ with $\overline{\sigma}(\Delta) \leq 1$ if and only if $\mu_{\mathcal{K}}(M) < 1$. Thus we see that μ is defined to be the answer to the robustness question of well posedness or "stability" of the constant matrix feedback loop shown in figure 2.1.

Now we will state the feedback interpretation of μ more rigorously. Consider a

matrix $M \in \mathbb{C}^{n\times n}$ partitioned as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{2.14}$$

with $M_{11} \in \mathbb{C}^{n_1\times n_1}, M_{22} \in \mathbb{C}^{n_2\times n_2}$ and $n_1+n_2 = n$. Suppose we have block structures $\mathcal{X}_{\mathcal{K}_1}$ and $\mathcal{X}_{\mathcal{K}_2}$ compatible with M_{11} and M_{22} respectively, then the block structure $\mathcal{X}_{\hat{\mathcal{K}}}$ defined as

$$\mathcal{X}_{\hat{\mathcal{K}}} = \{\Delta = \text{block diag}\,(\Delta_1, \Delta_2) : \Delta_1 \in \mathcal{X}_{\mathcal{K}_1}, \Delta_2 \in \mathcal{X}_{\mathcal{K}_2}\} \tag{2.15}$$

is compatible with M. Now given any $\Delta_1 \in \mathcal{X}_{\mathcal{K}_1}$ the LFT $F_u(M, \Delta_1)$ is said to be well posed if and only if there exists a unique solution to the loop equations shown in figure 2.2, namely

$$\begin{aligned} u &= M_{11}v + M_{12}d \\ e &= M_{21}v + M_{22}d \\ v &= \Delta_1 u. \end{aligned}$$

It is easy to see that $F_u(M, \Delta_1)$ is well posed if and only if $(I_{n_1} - M_{11}\Delta_1)$ is invertible. When the LFT is well posed it is defined to be the unique mapping from $d \longrightarrow e$, i. e., the vectors e and d satisfy $e = F_u(M, \Delta_1)d$, where

$$F_u(M, \Delta_1) \doteq M_{22} + M_{21}\Delta_1(I_{n_1} - M_{11}\Delta_1)^{-1}M_{12}. \tag{2.16}$$

Note that in the above derivation we always assume that the feedback loop is closed around the top inputs and outputs, and hence we obtain an upper LFT (denoted F_u). This is without loss of generality, simply by reordering the inputs and outputs, but in any case one can analogously define lower LFT's and prove similar results for them. In fact one can generalize this definition of LFT to incorporate Redheffer's "star product" [63], and we refer the reader to [57,22] for a more in-depth treatment of the general properties of LFT's.

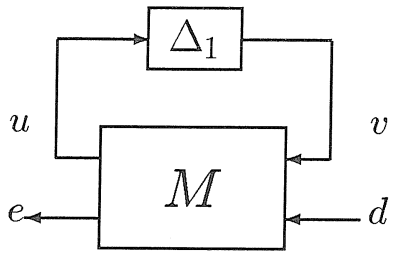


Figure 2.2: Robust performance μ analysis framework

Transfer functions represent an important example of LFT's. Consider the State Space realization of a discrete time system

$$\begin{pmatrix} x_{k+1} \\ y_k \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} x_k \\ u_k \end{pmatrix} \doteq M \begin{pmatrix} x_k \\ u_k \end{pmatrix} \tag{2.17}$$

then its transfer matrix is

$$G(z) = D + C(zI - A)^{-1}B = F_u(M, \frac{1}{z}I) \doteq F_u(M, \delta I).$$

Note that this is an LFT on a constant real matrix M, with a repeated complex scalar parameter δ, which is associated with the (inverse of the) frequency variable z.

Systems with uncertainty can also be easily represented using LFT's. One natural type of uncertainty is unknown coefficients in a State Space model. The following simple example is taken from [57]. Begin with a familiar idealized mass-spring-damper system, as shown in figure 2.3.

Suppose m, c, and k are fixed but uncertain, with

$$\begin{aligned} m &= \bar{m}(1 + w_m\delta_m) \\ c &= \bar{c}(1 + w_c\delta_c) \\ k &= \bar{k}(1 + w_k\delta_k). \end{aligned}$$

$$\ddot{y} + \frac{c}{m}\dot{y} + \frac{k}{m}y = \frac{f}{m}$$

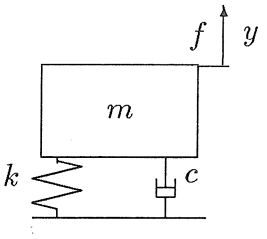


Figure 2.3: Mass-spring-damper system

where $\delta_m, \delta_c, \delta_k$ are all uncertain real scalars that are known to lie in the interval $[-1 \quad 1]$, but are otherwise unknown. The known parameters $\bar{m}, \bar{c}, \bar{k}$ represent the nominal values of m, c, k respectively, and the known weights w_m, w_c, w_k serve to normalize the uncertainty range to the unit ball.

Then defining $x_1 = y$ and $x_2 = \dot{y}$ we can write the differential equation in State Space form:

$$\begin{pmatrix} \dot{x} \\ y \end{pmatrix} = F_u(M, \Delta) \begin{pmatrix} x \\ f \end{pmatrix} \qquad \Delta = diag(\delta_m, \delta_c, \delta_k)$$

$$M = \begin{bmatrix} -w_m & \frac{-w_c}{\bar{m}} & \frac{-w_k}{\bar{m}} & \frac{-\bar{k}}{\bar{m}} & \frac{-\bar{c}}{\bar{m}} & \frac{1}{\bar{m}} \\ 0 & 0 & 0 & 0 & \bar{c} & 0 \\ 0 & 0 & 0 & \bar{k} & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ -w_m & \frac{-w_c}{\bar{m}} & \frac{-w_k}{\bar{m}} & \frac{-\bar{k}}{\bar{m}} & \frac{-\bar{c}}{\bar{m}} & \frac{1}{\bar{m}} \\ 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

Thus we obtain an LFT on a known matrix M, with the uncertainties collected together in a diagonal matrix Δ. Note that this is exactly the problem formulation assumed for the μ analysis framework. More generally, the perturbed State Space system

$$x_{k+1} = A(\delta)x_k + B(\delta)u_k$$

$$y_k = C(\delta)x_k + D(\delta)u_k$$

where δ is a vector of parameters that enter rationally can be written as an LFT on a diagonal matrix Δ made up of the (possibly repeated) elements of δ.

A fundamental property of LFT's that contributes to their importance in linear systems theory is that interconnections of LFT's are again LFT's. For example, consider a situation with three components, each with an LFT uncertainty model. The interconnection is shown in figure 2.4. By simply reorganizing the diagram, collecting all of the known systems together, and collecting all of the perturbations (the Δ_i's) together, we end up with the diagram in figure 2.5, where P depends only on G_1, G_2, G_3 and the diagram layout. Note how unstructured (or structured) uncertainty at the component level becomes *structured uncertainty at the system level.*

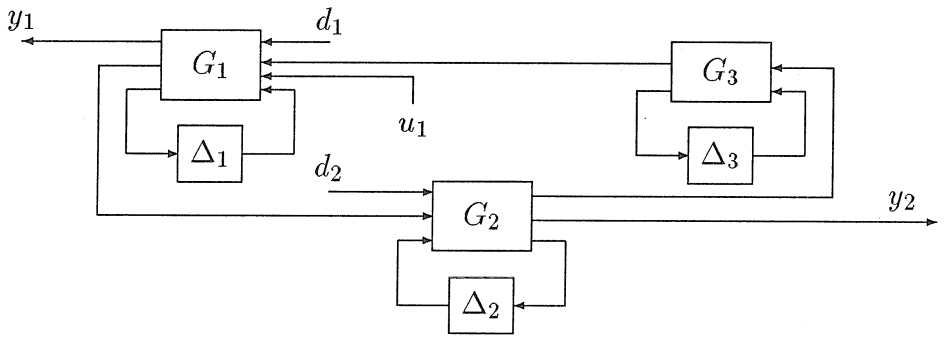


Figure 2.4: Example interconnection of LFT's

Note that problems involving additive and/or multiplicative uncertainty are special cases of linear fractional uncertainty descriptions. Furthermore by using the

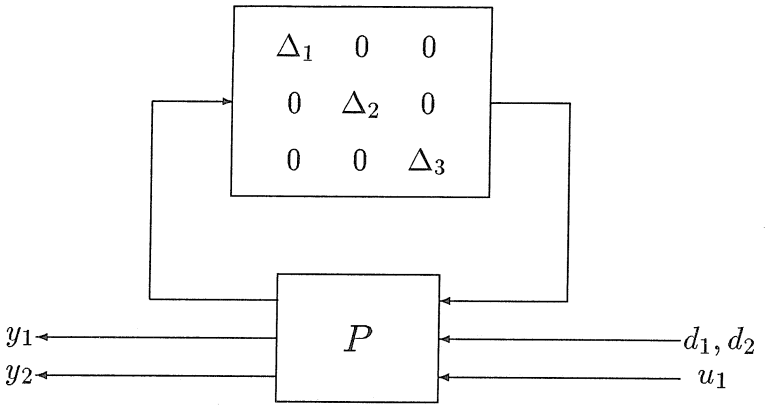


Figure 2.5: Macroscopic representation of figure 2.4

above rearrangements we can convert problems with additive and/or multiplicative uncertainty, or indeed any interconnection of systems with uncertainty entering in a linear fractional manner, into the standard μ analysis set-up of figure 2.1. Thus by allowing structured uncertainty entering the problem in a linear fractional way, we have a very general set-up that captures a great many robustness problems. Furthermore by allowing mixed real and complex uncertainties, we can capture both parametric and dynamic uncertainty in one unified framework, and we will see that we can consider both robust stability and performance problems in the same setting. Additional information on LFT's and how they arise in engineering problems is found in [22].

The following two theorems, which address robust stability and performance questions for LFT feedback interconnections of constant matrices, were proven for complex uncertainty descriptions in [56].

Theorem 2.1 (Well Posedness) *Let $M \in \mathbb{C}^{n\times n}$ and $0 < \beta \in \mathbb{R}$. The LFT $F_u(M,\Delta_1)$ is well posed for all $\Delta_1 \in \mathcal{X}_{\mathcal{K}_1}, \overline{\sigma}(\Delta_1) \leq \frac{1}{\beta}$ if and only if $\mu_{\mathcal{K}_1}(M_{11}) < \beta$.*

Theorem 2.2 (Main Loop Theorem) *Let $M \in \mathbb{C}^{n\times n}$ and $0 < \beta \in \mathbb{R}$. Then $\mu_{\hat{\mathcal{K}}}(M) < \beta$ if and only if $\mu_{\mathcal{K}_1}(M_{11}) < \beta$ and for all $\Delta_1 \in \mathcal{X}_{\mathcal{K}_1}, \overline{\sigma}(\Delta_1) \leq \frac{1}{\beta}$ we have $\mu_{\mathcal{K}_2}(F_u(M,\Delta_1)) < \beta$.*

These results extend trivially to the mixed case, the proofs being identical to those given in [56] for the complex case (there are a number of minor variations to these theorems which similarly extend to the mixed case). For a more complete feedback interpretation of these results see [56] and the references therein.

These two theorems are at the heart of the complex μ analysis methodology. The fact that they are still true in the mixed case means that mixed μ still provides an *exact* test for robust stability and robust performance problems, but now with real uncertainties allowed. The LFT machinery for rearranging various robustness problems into μ problems (such as converting robust performance problems into robust stability problems as above) is also applicable to the mixed case, so that mixed μ retains its versatility and applicability to a large class of problems. Note in particular that problems of robust stability of a polynomial with perturbed real coefficients (or a ratio of such polynomials for that matter) are yet another example of linear fractional uncertainty, and hence can be recast as mixed μ problems (see section 8.1 for example).

The structured singular value can be used to quantify robustness margins for a linear system with linear fractional uncertainty. Specifically, suppose that M in figure 2.2 is a real-rational, proper transfer matrix, denoted by $P(s)$, of size $(n_1 + n_2) \times (n_1 + n_2)$, and block structures $X_{\mathcal{K}_1} \subset \mathbb{C}^{n_1\times n_1}$ and $X_{\mathcal{K}_2} \subset \mathbb{C}^{n_2\times n_2}$ are given. Partition $P(s)$ in the obvious way.

$$P(s) = \begin{bmatrix} P_{11}(s) & P_{12}(s) \\ P_{21}(s) & P_{22}(s) \end{bmatrix}$$

For $\Delta_1 \in X_{\mathcal{K}_1}$, consider the interconnection shown in figure 2.2, with $M = P(s)$. For any $\Delta_1 \in X_{\mathcal{K}_1}$, $F_u(P(s), \Delta_1)$ is the transfer function from $d \rightarrow e$.

It is easy to show that the structured singular value is not necessarily a norm (or even a semi-norm), since it doesn't necessarily satisfy the triangle inequality. Nevertheless it does satisfy a scaling property (see section 3.1, property (a)), and so in some sense is related to the size of the matrix, and so we introduce the following notation: for a transfer matrix $P \in \mathcal{RH}_\infty$, and a block structure $\mathcal{K}$ of appropriate dimensions, define the "μ-norm" of P as

$$\|P\|_{\mathcal{K}} := \sup_{w \in \mathbb{R}} \mu_{\mathcal{K}}(P(\mathrm{j}w)).$$

Now consider $\Delta \in X_{\mathcal{K}}$ to be a block structure partitioned compatibly with P. We allow Δ_1 to be arbitrary, but assume that Δ_2 is a single complex full block. This is a "performance block" and will give us a robust performance test for the $\mathcal{H}_\infty$ norm. Since the uncertainties are typically used to cover unmodeled dynamics, then we wish to be able to consider perturbations to P which are themselves dynamical systems, with the block diagonal structure of the set $X_{\mathcal{K}}$. Associated with any block structure $X_{\mathcal{K}}$, let $\mathcal{M}(X_{\mathcal{K}})$ denote the set of all real-rational, proper, stable, block diagonal transfer matrices, with block structure like $X_{\mathcal{K}}$:

$$\mathcal{M}(X_{\mathcal{K}}) := \{\Delta \in \mathcal{RH}_\infty : \Delta(\mathrm{j}w) \in X_{\mathcal{K}} \text{ for all } w \in \mathbb{R}\}$$

The following two theorems address robust stability and performance questions for LFT's of linear systems, and give rise to the most common usage of μ: as a frequency domain robustness test.

Theorem 2.3 (Robust Stability) *Suppose that $P \in \mathcal{RH}_\infty$ and we have a real scalar $\beta > 0$. Then for all $\Delta_1 \in \mathcal{M}(X_{\mathcal{K}_1})$ with $\|\Delta_1\|_\infty \leq \beta$, the perturbed closed-loop system is well-posed and stable if and only if*

$$\|P_{11}\|_{\mathcal{K}_1} \doteq \sup_{w \in \mathbb{R}} \mu_{\mathcal{K}_1}(P_{11}(\mathrm{j}w)) < \frac{1}{\beta}.$$

Theorem 2.4 (Robust Performance) *Suppose that $P \in \mathcal{RH}_\infty$ and we have a real scalar $\beta > 0$. Then for all $\Delta_1 \in \mathcal{M}(X_{\mathcal{K}_1})$ with $\|\Delta_1\|_\infty \leq \beta$, the perturbed closed-loop system is well-posed, stable and*

$$\|F_u(P, \Delta_1)\|_\infty \doteq \sup_{w \in \mathbb{R}} \overline{\sigma}[F_u(P(\mathrm{j}w), \Delta_1(\mathrm{j}w))] < \frac{1}{\beta}$$

if and only if

$$\|P\|_{\mathcal{K}} \doteq \sup_{w \in \mathbb{R}} \mu_{\mathcal{K}}(P(\mathrm{j}w)) < \frac{1}{\beta}.$$

These results mean that we can evaluate the robustness properties of a closed loop system, by using a frequency evaluation of μ. Note that at any given frequency point we have a constant matrix μ problem, for which we can use the upper and lower bounds to be discussed in chapter 4. Note also that the block structure $\mathcal{X}_{\mathcal{K}_1}$ is inherited from the problem description, whereas we consider the augmented block structure $\mathcal{X}_{\mathcal{K}}$ because of the result in theorem 2.4. In summary then, the peak value on the μ plot of the frequency response determines the size of perturbations that the loop is robustly stable (and/or performing) against.

Note from earlier that we can write transfer functions as LFT's. This allows us to remove the frequency search from the above tests by including the frequency variable as one of the uncertain parameters (a repeated complex scalar block). The μ test then looks for the worst case parameter values, and hence the worst case frequency, at the same time. In this way we obtain a one-shot test, involving a constant matrix μ problem, for the worst case μ value across frequency. Note that since we wish to treat the frequency variable as a disk uncertainty, this State Space μ test is naturally applied to discrete time systems. However by employing a bilinear transformation from the half plane to the disk, we can obtain similar results for continuous time problems as well, and we refer the reader to [56] for the appropriate results, which generalize easily to the mixed case. This State Space μ test has several interesting connections to Lyapunov theory, and stability with respect to nonlinear uncertainties, and for a more detailed exposition of these topics we refer the interested reader to [22,56,34,57,44,69] and the references therein.

Chapter 3

The Mixed μ Problem

In this chapter we examine some of the underlying properties of the mixed μ problem, and compare them with the appropriate complex μ properties. It is seen that whilst mixed μ inherits many of the properties of complex μ, in some aspects the mixed μ problem can be fundamentally different from the complex μ problem. In particular recent results on the issues of continuity and NP completeness are reviewed [64,20,58]. These results have profound implications, particularly for the computational aspects of the mixed μ problem.

3.1 Basic Properties of Mixed μ

In this section we will present some basic properties of the mixed μ problem and contrast them with the corresponding results for the complex μ problem (see [23]). Note that these are fundamental properties of the μ problem itself, and have nothing to do with the choice of bounds or method of computation one may use for the problem.

It is not at all obvious how to compute μ from the definition in (2.5), since this definition implicitly involves an optimization problem which is not convex. In fact we will see that the computation of μ is, in general, a difficult problem, for a number of reasons, and a large part of our effort will be geared towards obtaining

good approximations to μ. Let us first start by examining some of the properties of the mixed μ problem. From the definition of mixed μ in equation (2.5), one may readily derive the following properties (the complex μ versions of these were originally presented in [23]):

(a) $\mu_{\mathcal{K}}(\gamma M) = |\gamma|\mu_{\mathcal{K}}(M)$ for all $M \in \mathbb{C}^{n\times n}$ and $\gamma \in \mathbb{R}$.

(b) $\mu_{\mathcal{K}}(I_n) = 1$ for any block structure.

(c) $\mu_{\mathcal{K}}(\Delta) = \overline{\sigma}(\Delta)$ for all $\Delta \in X_{\mathcal{K}}$.

(d) $m_r = 0, m_c = 0, m_C = 1 \quad \rightarrow \quad \mu_{\mathcal{K}}(M) = \overline{\sigma}(M)$.

(e) $m_r = 0, m_c = 1, m_C = 0 \quad \rightarrow \quad \mu_{\mathcal{K}}(M) = \rho(M)$.

(f) $m_r = 1, m_c = 0, m_C = 0 \quad \rightarrow \quad \mu_{\mathcal{K}}(M) = \rho_R(M)$.

(g) For any $M \in \mathbb{C}^{n\times n}$ the following sequence of inequalities holds

$$\rho_R(M) \leq \mu_{\mathcal{K}}(M) \leq \overline{\sigma}(M).$$

(h) For all $\Delta \in X_{\mathcal{K}}, Q \in \mathcal{Q}_{\mathcal{K}}$ then $Q\Delta \in X_{\mathcal{K}}, \Delta Q \in X_{\mathcal{K}}$ with $\overline{\sigma}(Q\Delta) \leq \overline{\sigma}(\Delta)$, $\overline{\sigma}(\Delta Q) \leq \overline{\sigma}(\Delta)$.

(i) For all $\Delta \in X_{\mathcal{K}}, U \in \mathcal{U}_{\mathcal{K}}$ then $U\Delta \in X_{\mathcal{K}}, \Delta U \in X_{\mathcal{K}}$ with $\overline{\sigma}(U\Delta) = \overline{\sigma}(\Delta)$, $\overline{\sigma}(\Delta U) = \overline{\sigma}(\Delta)$.

(j) $D\Delta D^{-1} = \Delta$ for all $\Delta \in X_{\mathcal{K}}$ and $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ (including $D \in \mathcal{D}_{\mathcal{K}}$).

(k) $\mu_{\mathcal{K}}(QM) = \mu_{\mathcal{K}}(MQ) \leq \mu_{\mathcal{K}}(M)$ for all $M \in \mathbb{C}^{n\times n}$ and $Q \in \mathcal{Q}_{\mathcal{K}}$.

(l) $\mu_{\mathcal{K}}(UM) = \mu_{\mathcal{K}}(MU) = \mu_{\mathcal{K}}(M)$ for all $M \in \mathbb{C}^{n\times n}$ and $U \in \mathcal{U}_{\mathcal{K}}$.

(m) $\mu_{\mathcal{K}}(DMD^{-1}) = \mu_{\mathcal{K}}(M)$ for all $M \in \mathbb{C}^{n\times n}$ and $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ (including $D \in \mathcal{D}_{\mathcal{K}}$).

(n) For any $M \in \mathbb{C}^{n\times n}$ the following sequence of inequalities holds

$$\max_{U\in\mathcal{U}_\mathcal{K}} \rho_R(UM) \le \max_{Q\in\mathcal{Q}_\mathcal{K}} \rho_R(QM) \le \mu_\mathcal{K}(M) \le \inf_{D\in\mathcal{D}_\mathcal{K}} \overline{\sigma}(DMD^{-1})$$

where we could equivalently use

$$\max_{U\in\mathcal{U}_\mathcal{K}} \rho_R(MU) = \max_{U\in\mathcal{U}_\mathcal{K}} \rho_R(UM), \quad \max_{Q\in\mathcal{Q}_\mathcal{K}} \rho_R(MQ) = \max_{Q\in\mathcal{Q}_\mathcal{K}} \rho_R(QM).$$

(o) For all $\Delta \in X_\mathcal{K}$ there exist $U, V \in \mathcal{U}_\mathcal{K}$ and $\Sigma = diag(\sigma_1 \dots \sigma_n)$ with $\sigma_i \in \mathbb{R}, \sigma_i \ge 0$ such that $\Delta = U\Sigma V^*$.

Note that there are important differences between some of these properties and their complex μ versions (see [23]). In particular the function ρ_R may be replaced by ρ in any of the above for complex μ problems.

The scaling property (a) follows immediately from the definition of μ in (2.5). This property is fundamental to the use of μ as a measure of the robust stability (and performance) margin of a given system, since it implies that μ is related to the size of the worst case destabilizing perturbation. Note that this property holds for $\gamma \in \mathbb{C}$ for complex μ problems.

Properties (b) - (f) are special cases for which μ is easily computed. Note from property (d) that μ for a single full complex block reduces to the maximum singular value, and in property (e) we see that μ for a single repeated complex scalar block reduces to the spectral radius. Thus μ can be thought of as a generalization of both of these quantities, and the fact that for unstructured uncertainty we have $\mu = \overline{\sigma}$, leads to the term *structured singular value.*

In order to obtain property (g) we first need the following lemma, which may be easily proven from the definition of μ.

Lemma 3.1 *Suppose we have two block structures $X_{\mathcal{K}_1}$ and $X_{\mathcal{K}_2}$, both of which are compatible with matrices in $\mathbb{C}^{n\times n}$. If we have $X_{\mathcal{K}_1} \subset X_{\mathcal{K}_2}$ then for any $M \in \mathbb{C}^{n\times n}$ it follows that*

$$\mu_{\mathcal{K}_1}(M) \le \mu_{\mathcal{K}_2}(M). \tag{3.1}$$

Note that for any uncertainty structure $X_{\mathcal{K}} \subset \mathbb{C}^{n\times n}$ we always have that

$$\{\Delta : \Delta = \delta I_n, \delta \in \mathbb{R}\} \subset \{\Delta : \Delta \in X_{\mathcal{K}}\} \subset \{\Delta : \Delta \in \mathbb{C}^{n\times n}\} \tag{3.2}$$

so that applying lemma 3.1 to the above, together with properties (d) and (f), we find that for any $M \in \mathbb{C}^{n\times n}$ and any compatible block structure $\mathcal{K}$

$$\rho_R(M) \leq \mu_{\mathcal{K}}(M) \leq \overline{\sigma}(M) \tag{3.3}$$

which is property (g). Note that for complex μ problems we have

$$\rho(M) \leq \mu_{\mathcal{K}}(M) \leq \overline{\sigma}(M). \tag{3.4}$$

These inequalities form the starting point for schemes to compute upper and lower bounds for μ. We will see that, roughly speaking, we will use a scaled (real) spectral radius condition as our lower bound for μ, and a scaled maximum singular value condition as our upper bound for μ.

The next idea exploited in the list of properties is to examine those scalings which affect either of the quantities ρ_R or $\overline{\sigma}$, but do not affect μ. Note that eigenvalues are invariant to similarity transformations, but are not invariant to unitary matrices. However for singular values the opposite is true: singular values are invariant to unitary matrices, but are not invariant to similarity transformations. It turns out that μ is invariant to an appropriate subset of unitary matrices, and to an appropriate subset of similarity transformations. This is discussed in more detail below.

Properties (h) and (i) follow from the definitions of the sets $\mathcal{U}_{\mathcal{K}}$ and $\mathcal{Q}_{\mathcal{K}}$. Note that we have $\mathcal{U}_{\mathcal{K}} \subset \mathcal{Q}_{\mathcal{K}} \subset X_{\mathcal{K}}$, so that all these sets represent allowable perturbations. Applying properties (h) and (i) to the definition of μ (2.5) we obtain properties (k) and (l). Property (l) states that μ is invariant to matrices in $\mathcal{U}_{\mathcal{K}}$ (which are block unitary). Note however that from property (k) we see that μ is not necessarily invariant to matrices in $\mathcal{Q}_{\mathcal{K}}$ (which may not be unitary, since the real parameters are not restricted to be on the boundary of the allowable set).

Applying properties (k) and (l) to the lower bound inequality in (g) we obtain the improved (scaled) lower bound inequalities in (n).

$$\max_{U\in\mathcal{U}_\mathcal{K}} \rho_R(UM) \leq \max_{Q\in\mathcal{Q}_\mathcal{K}} \rho_R(QM) \leq \mu_\mathcal{K}(M)$$

For complex μ problems this becomes

$$\max_{Q\in\mathcal{Q}_\mathcal{K}} \rho(QM) \leq \mu_\mathcal{K}(M).$$

We said earlier that matrices in $\mathcal{Q}_\mathcal{K}$ are not necessarily μ invariant, since they are not necessarily block unitary. The reason that we introduce the set $\mathcal{Q}_\mathcal{K}$ (instead of just working with $\mathcal{U}_\mathcal{K}$) is that it will be seen in section 4.1 that the second inequality in (n) is in fact an equality, whereas the first inequality is not. This contrasts with the complex μ case where $\mathcal{Q}_\mathcal{K} \equiv \mathcal{U}_\mathcal{K}$ so that one may always assume the worst case perturbation is unitary and μ-invariant (note also that one may still assume the *complex* blocks of the worst case perturbation are unitary for mixed problems). This is an important distinction between real and complex perturbations. The worst case value for a real parameter may occur at an internal point, whereas for a complex parameter the worst case value is always on the boundary. This is treated in detail in section 4.1.

Note that the set $\hat{\mathcal{D}}_\mathcal{K}$ is constructed with a block structure complementary to that of $X_\mathcal{K}$, so that it commutes with all matrices in $X_\mathcal{K}$ (and hence $\mathcal{Q}_\mathcal{K}$ and $\mathcal{U}_\mathcal{K}$ as well), which is property (j). This immediately leads to property (m) which is sufficiently important that we state it separately as a lemma.

Lemma 3.2 *For any matrix $M \in \mathbb{C}^{n\times n}$, and any compatible block structure $\mathcal{K}$ then for all $D \in \hat{\mathcal{D}}_\mathcal{K}$*

$$\mu_\mathcal{K}(M) = \mu_\mathcal{K}(DMD^{-1}). \tag{3.5}$$

Since $\mathcal{D}_\mathcal{K} \subset \hat{\mathcal{D}}_\mathcal{K}$ we have that μ is invariant to similarity transformations with matrices in $\mathcal{D}_\mathcal{K}$, and applying this to the upper bound inequality in (g) we obtain the improved

(scaled) upper bound inequality in (n)

$$\mu_{\mathcal{K}}(M) \leq \inf_{D \in \mathcal{D}_{\mathcal{K}}} \overline{\sigma}(DMD^{-1}).$$

We state these improved upper and lower bound results as a lemma.

Lemma 3.3 *For any matrix $M \in \mathbb{C}^{n \times n}$, and any compatible block structure $\mathcal{K}$*

$$\max_{Q \in \mathcal{Q}_{\mathcal{K}}} \rho_R(QM) \leq \mu_{\mathcal{K}}(M) \leq \inf_{D \in \mathcal{D}_{\mathcal{K}}} \overline{\sigma}(DMD^{-1}). \tag{3.6}$$

For complex μ problems then we have

$$\max_{Q \in \mathcal{Q}_{\mathcal{K}}} \rho(QM) \leq \mu_{\mathcal{K}}(M) \leq \inf_{D \in \mathcal{D}_{\mathcal{K}}} \overline{\sigma}(DMD^{-1}). \tag{3.7}$$

These bounds form the basis for computation techniques for complex μ problems (for additional material on upper and lower bounds for complex μ see [28] and [59]). This upper bound is exact for complex block structures with $2m_c + m_C \leq 3$ (see [55]) but is not exact for any block structures with $m_r \neq 0$. In fact it is (usually) possible to improve on this bound for mixed μ problems with $m_r \neq 0$. The mixed μ upper bound presented in [29], and examined in section 4.6, is never worse than this standard complex μ upper bound, and is frequently much better.

A great deal of research has been carried out with regard to complex μ problems, and the bounds in (3.7). They are not *guaranteed* to be "tight" for a general μ problem, but in fact they usually are for problems of engineering interest. Practical implementations of these bounds have been developed, and efficient software is available for their computation (see [3] for details of the μ-Tools toolbox). In addition an array of interesting theoretical results have been derived, and we refer the reader to [57] for a detailed review of the complex μ problem. A good deal of the work in this thesis will be geared towards extending these results and methodologies to the more general mixed μ problem.

3.2 Continuity of Mixed μ

In order to examine the continuity of the μ problem we first reformulate the definition in (2.5) as a (real) eigenvalue maximization problem. Define the unit ball in the perturbation set as

$$\mathbb{B}X_{\mathcal{K}} \doteq \{\Delta \in X_{\mathcal{K}} : \overline{\sigma}(\Delta) \leq 1\}. \tag{3.8}$$

Then the following lemma follows almost immediately from the definition of μ.

Lemma 3.4 *For any matrix $M \in \mathbb{C}^{n\times n}$, and any compatible block structure $\mathcal{K}$*

$$\mu_{\mathcal{K}}(M) = \max_{\Delta\in\mathbb{B}X_{\mathcal{K}}} \rho_R(\Delta M). \tag{3.9}$$

Now for complex μ problems this becomes

$$\mu_{\mathcal{K}}(M) = \max_{\Delta\in\mathbb{B}X_{\mathcal{K}}} \rho(\Delta M).$$

Since $\mathbb{B}X_{\mathcal{K}} \subset \mathbb{C}^{n\times n}$ is a compact set, and ρ is a continuous function, it follows that complex μ is also a continuous function of the problem data [23].

In general for mixed μ problems however we only have the expression (3.9). Although $\mathbb{B}X_{\mathcal{K}} \subset \mathbb{C}^{n\times n}$ is still a compact set, the function ρ_R is not necessarily a continuous function, but rather only upper semicontinuous. As a result we can only conclude in general that mixed μ is upper semicontinuous, as stated in the following lemma.

Lemma 3.5 (Upper Semicontinuity of Mixed μ [58]) *Mixed μ is an upper semicontinuous function of the problem data. Equivalently, suppose we have $M \in \mathbb{C}^{n\times n}$, and a compatible block structure $\mathcal{K}$. Then for any $\beta > 0$ such that $\mu_{\mathcal{K}}(M) < \beta$, there exists $\epsilon > 0$ such that $\mu_{\mathcal{K}}(\hat{M}) < \beta$ for all $\hat{M} \in \mathbb{C}^{n\times n}$ with $\overline{\sigma}(\hat{M} - M) < \epsilon$.*

See [58] for more details on the continuity properties of mixed μ, and [65] for more details on the notion of upper semicontinuous functions. Note that one can construct examples of real μ problems where μ is discontinuous in the problem data (see [9] for example).

The fact that real μ can be discontinuous clearly adds computational difficulties to the problem, since any method involving some type of search (e. g., frequency response) must address the possibility of missing a point of discontinuity. Note that for an upper semicontinuous function one may have an isolated point which is *larger* than its neighbors (see figure 3.1), so that by missing this point one would underestimate the peak value of μ, or in other words overestimate the robust stability margin for your problem (roughly speaking the size of the smallest destabilizing perturbation is $\frac{1}{\mu}$). In fact the example from [9] has a sequence of polynomials which converge to a limiting polynomial, but the robust stability margin of the limit polynomial is less than the limit of the stability margins of the convergent polynomials.

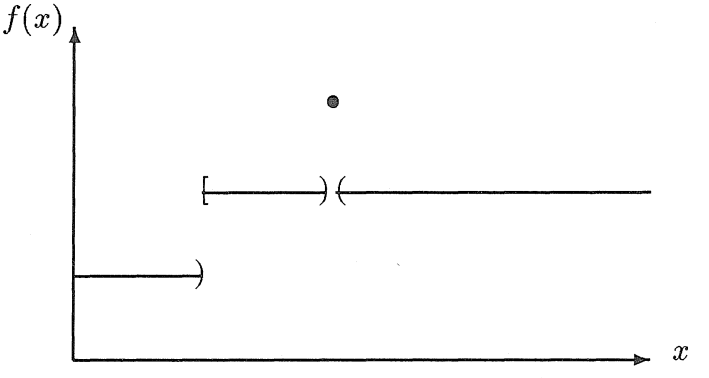


Figure 3.1: Upper semicontinuous function

More importantly however the fact that real μ can be discontinuous in the problem data sheds serious doubt on the usefulness of real μ as a *robustness* measure in such cases. This is because the system model is always a mathematical abstraction from the real world, and is only computed to finite precision, so that it would seem reasonable to require that any type of robustness measure we use be continuous in the problem data.

It is shown in [58] how to regularize these problems by essentially adding a small amount of complex uncertainty to each real uncertainty. By doing this a small amount of phase uncertainty is added to the gain uncertainty. It is then shown that the new mixed μ problem is continuous. This regularization seems reasonably well motivated from an engineering point of view, where unmodeled dynamics would always produce some phase uncertainty.

Furthermore it is shown in [58] that mixed μ problems containing some complex uncertainty are, under some mild assumptions, continuous even without the regularization procedure outlined above (whereas purely real μ problems are not). To be more explicit suppose we have a matrix M partitioned as in (2.14) and block structures $\mathcal{X}_{\mathcal{K}_1}, \mathcal{X}_{\mathcal{K}_2}, \mathcal{X}_{\hat{\mathcal{K}}}$ as in section 2.2. Further assume we have arranged the problem so that $\mathcal{X}_{\mathcal{K}_1}$ consists of purely real uncertainties and $\mathcal{X}_{\mathcal{K}_2}$ consists of purely complex uncertainties. Then the following theorem was proven in [58].

Theorem 3.1 (Continuity of Mixed μ [58]) *Suppose we have a matrix $M \in \mathbb{C}^{n\times n}$ and block structures $\mathcal{X}_{\mathcal{K}_1}, \mathcal{X}_{\mathcal{K}_2}, \mathcal{X}_{\hat{\mathcal{K}}}$ as above. Then if $\mu_{\mathcal{K}_1}(M_{11}) < \mu_{\hat{\mathcal{K}}}(M)$, μ is continuous in the problem data at M.*

Note that it is always true that $\mu_{\mathcal{K}_1}(M_{11}) \leq \mu_{\hat{\mathcal{K}}}(M)$, so that the condition $\mu_{\mathcal{K}_1}(M_{11}) < \mu_{\hat{\mathcal{K}}}(M)$ can be interpreted as meaning that the complex uncertainties enter nontrivially into the problem. Some condition of this type is clearly required since we can always trivially construct discontinuous mixed μ problems from discontinuous real μ problems by simply padding them out with zeroes. These issues are treated in greater depth in [58] where they also develop alternative conditions under which mixed μ is continuous. The continuity issue is also considered in [40], where the authors examine the continuity properties of the mixed μ upper bound.

These results, and in particular theorem 3.1, are reassuring from an engineering viewpoint since one is usually interested in robust performance problems (which therefore contain at least one complex block), or robust stability problems with some unmodeled dynamics, which are naturally covered with complex uncertainty. Thus

in problems of engineering interest, the potential discontinuity of μ should not arise, although conditioning of μ computation could be a problem and needs more study.

3.3 NP Completeness of Mixed μ

The theory of computational complexity is concerned with the tractability of computational problems. There are many important concepts in this area that can have a dramatic impact on the development of practical computation schemes for a given problem. In particular we will be concerned here with the notion of NP complete (and NP hard) problems. We refer the reader unfamiliar with these concepts to [31] for more than the very brief exposition we will present here.

The theory of NP completeness deals with decision problems, i. e., problems with a "YES/NO" answer. In particular we will be interested in the decision problem: "Is $\mu \geq k$?" for a given real positive scalar k. We will refer to this decision problem as the "μ recognition problem." Note that any lower bounds we have on the computational difficulty of a given decision problem immediately become lower bounds on the computational difficulty of the associated evaluation problem. This follows since if we can evaluate μ, then that immediately gives us an answer to the μ recognition problem. Thus when we loosely refer to a (evaluation) problem as being NP complete, we mean that the associated decision problem is NP complete, and hence the evaluation problem is NP hard (i. e., at least as hard as an NP complete problem).

The following theorem follows almost immediately from recent results in [64].

Theorem 3.2 (NP Completeness of Real μ) *The real μ recognition problem, for $M \in \mathbb{R}^{n\times n}$ and $X_{\mathcal{K}} = \{\Delta = diag(\delta_1^r, \ldots, \delta_n^r) : \delta_i^r \in \mathbb{R}\}$, is NP complete.*

The main ideas behind this result are as follows. Rohn and Poljak showed in [64] that the interval matrix problem, for a real matrix and unit intervals, is equivalent to the "max-cut" problem, which is known to be NP complete. A simple rearrangement converts that interval matrix problem into the above real μ problem, as shown below.

Suppose we have an interval matrix problem, i. e., we have a matrix $M \in \mathbb{R}^{n\times n}$, and we have an element by element uncertainty structure:

$$X_{full} = \{\Delta \in \mathbb{R}^{n\times n} : |\Delta_{ij}| \leq 1 \text{ for } i = 1,\ldots,n \text{ and } j = 1,\ldots,n\}.$$

Then the interval matrix problem asks: "Does there exist $\Delta \in X_{full}$ such that $M+\Delta$ is singular?" This problem can be transformed into the standard μ framework by simply choosing $L \in \mathbb{R}^{n\times n^2}$, $R \in \mathbb{R}^{n^2\times n}$, and $\Delta_d \in \mathbb{R}^{n^2\times n^2}$ such that

$$\Delta = L\Delta_d R$$

where the new uncertainty structure, Δ_d, is diagonal and the elements of Δ are strung out along the diagonal of Δ_d. This is a standard transformation (see [39] for more details), and L, R are known constant matrices, independent of Δ. Thus we have $\Delta_d \in X_{\mathcal{K}}$ where $m_r = n^2, m_c = 0, m_C = 0$ and

$$\mathcal{K} = (1,\ldots,1).$$

Now if M is singular then the decision problem is trivially answered as a "YES," and if not then we have:

$$\begin{aligned}\det(M+\Delta) &= \det(M + L\Delta_d R)\\ &= \det(M(I + M^{-1}L\Delta_d R))\\ &= \det(M)\det(I + (RM^{-1}L)\Delta_d)\end{aligned}$$

so that

$$\det(M+\Delta) = 0 \quad \longleftrightarrow \quad \det(I + (RM^{-1}L)\Delta_d) = 0$$

The question "Does there exist Δ_d such that $I + (RM^{-1}L)\Delta_d$ is singular?" is a real μ recognition problem. Thus we have that the given interval matrix problem polynomially reduces to a real μ problem, and so it follows that this real μ problem is NP complete as well.

While these results do not apply to the complex only case, we have the following new result from [16].

Theorem 3.3 (NP Hardness of Mixed μ [16]) *The general mixed μ recognition problem, for $M \in \mathbb{C}^{n\times n}$ and any compatible block structure $\mathcal{K}$, is NP hard.*

Note that since the special case in theorem 3.2 and the general case in theorem 3.3 are both NP hard, it follows that any problem "between" these two problems is also NP hard (e. g., the general real μ problem for $M \in \mathbb{C}^{n\times n}$).

The results in [16] are based on the fact that the indefinite (constrained) quadratic programming problem given by maximizing over x the expression

$$\max_{b_l \le x \le b_u} |x^* A x + p^* x + c| \tag{3.10}$$

for $A \in \mathbb{R}^{n\times n}$, $x, p, b_l, b_u \in \mathbb{R}^n$, and $c \in \mathbb{R}$ can be recast as a mixed μ problem. This rearrangement uses simple block diagram LFT manipulations, and the main loop theorem (see section 2.2). It can be shown easily from known results that the indefinite quadratic programming problem in (3.10) is NP hard, and it follows that the mixed μ problem is NP hard as well.

One might wonder if these computational complexity results are due to the fact that we allow pure real μ problems as a special case of mixed μ problems. Recall that the discontinuity of mixed μ was no longer a possibility if we restricted our attention to a certain class of non-trivial mixed μ problems (which are motivated by engineering considerations). Unfortunately this is not the case for the computational complexity results.

Theorem 3.4 (NP Hardness of Non Trivial Mixed μ [16]) *The mixed μ recognition problem, for $M \in \mathbb{C}^{n\times n}$ and a compatible block structure $\mathcal{K}$ as in theorem 3.1, is NP hard.*

This result means that we must deal with the issue of NP hard computation, even for non-trivial mixed μ problems. Note also that combining this result with theorem 3.1, it follows that the mixed μ recognition problem is still NP hard when we restrict our attention to the class of continuous mixed μ problems [16].

It is still a fundamental open question in the theory of computational complexity to determine the exact consequences of a problem being NP hard, and we refer the reader to [31] for an in depth treatment of the subject. However, it is generally accepted that a problem being NP hard means that it cannot be computed in polynomial time in the worst case. It is important to note that being NP hard is a property of the problem itself, not any particular algorithm. The fact that the mixed μ problem is NP hard strongly suggests that given *any* algorithm to compute μ, there will be problems for which the algorithm cannot find the answer in polynomial time. This means that for all practical purposes even moderately large examples of such problems are computationally intractable.

For the reader not familiar with these concepts, we offer the following illustration. Consider the example in table 3.1. There we have tabulated two different growth rates versus problem size. For each growth rate we have assumed that it represents two different algorithms, one which can solve a size 10 problem in 10 seconds, and one which can solve a size 10 problem in 0.01 seconds. The first growth rate is n^3 (where n is the problem size). This is a polynomial time growth rate, and is typical of algorithms for eigenvalues, singular values etc. The second growth rate is 2^n. This is an exponential (non-polynomial) time growth rate, and is typical of algorithms which require one to check all the edges or vertices of some polytope.

It is readily seen that given an algorithm with a polynomial time growth rate we can apply the algorithm to larger and larger problems with a reasonable increase in the computational requirements. In contrast, for the exponential time growth rate the increase in computational requirements is quite dramatic, and for even moderate sizes the problem rapidly becomes intractable. It is important to note that even if the exponential time algorithm is much faster on small problems it still rapidly becomes impractical as the problem size increases. The overriding implication of all this is that if we wish to be able to handle fairly large problems, we must have polynomial time algorithms, regardless of the speed on small problems. The fact that the mixed

Growth Rate	Problem Size (n)				
	10	20	30	40	50
n^3	0.01 seconds	0.08 seconds	0.27 seconds	0.64 seconds	1.25 seconds
	10 seconds	1.33 minutes	4.50 minutes	10.67 minutes	20.83 minutes
2^n	0.01 seconds	10.24 seconds	2.91 hours	124.3 days	348.7 years
	10 seconds	2.84 hours	121.4 days	340.5 years	3.49×10^5 years

Table 3.1: Comparison of polynomial and exponential time growth rates

μ problem is NP hard means that we cannot expect to find such algorithms if we attempt to solve the general problem exactly for all cases.

These results strongly suggest that it is futile to pursue exact methods for computing μ in the purely real or mixed case for even moderate (less than 100) numbers of real perturbations. One approach to overcoming this difficulty is to consider special cases of the general problem, which may be easier to solve. The difficulty with this approach is that one would like the resulting algorithm to be widely applicable to a large number of engineering problems, and it may be that the special cases that are easily solvable are too restrictive. For this reason we have concentrated on the general problem, rather than adopt this approach. Nevertheless, since special cases have been the focus of so much research, we will devote some time to considering those special cases for which computation of μ is relatively easy. In the remainder of this section, we briefly discuss some of the general issues associated with computation

of μ and special cases. Some of these special cases will be treated in more detail in chapters 5 and 8.

3.3.1 Problems with Special Structure

In light of the NP completeness results given earlier, it is natural to ask if there are special cases of the mixed μ problem that are relatively easy to compute. Essentially all such cases can be shown to involve problems where it can be verified a priori that μ is equal to its upper bound, and can therefore be computed as a convex optimization problem (see section 4.6 where the upper bound is presented in detail). Unfortunately, as we will see, these special cases are relevant to very few problems of engineering interest.

Although it is somewhat artificial, it is useful to separately consider the nominal system and the uncertainty structure (respectively P and Δ in figure 3.2), as one can get easily computable special cases from restrictions on each one. In the case of the nominal system, P, computation is easier when it is highly structured, whereas *less* structure on the uncertainty, Δ, makes computation easier. Of course, problems motivated by real engineering applications typically have general, unstructured nominal systems combined with highly structured uncertainty, exactly the opposite of what is ideal for computation.

For simplicity, consider the standard problem of robust stability for the system in figure 3.2 where Δ is assumed to be norm bounded by 1. The least structured Δ would be a single block which would be allowed to be an arbitrary nonlinear, time-varying operator. In that case the small gain condition [84] is necessary and sufficient, and the test is simply $\|P\|_\infty < 1$. This is still true when Δ is restricted to be causal, and further restricted to either linear time-varying (LTV) or linear time-invariant (LTI).

Additional structure on Δ leads to μ tests of varying complexity, but some special cases exist when μ is equal to its upper bound. If Δ is block diagonal with any number of LTV perturbations then recent results, obtained independently by Shamma [69]

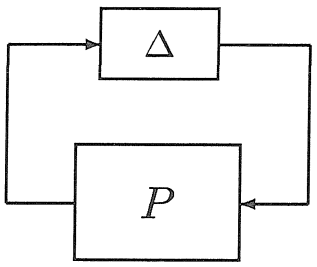


Figure 3.2: Standard robust stability problem

and Megretskii [46], show that the exact test for this case is equivalent to an upper bound for a complex μ problem. Also, if Δ consists of 3 or fewer LTI full blocks, then μ is equal to its upper bound. In general, μ is not equal to its upper bound for more complicated uncertainty structures, unless additional structure is imposed on the P. The role of structure on P will be considered briefly in the next subsection.

3.3.2 Restrictions on P and "Kharitonov-Type" Results

A popular research program over the last few years has focused on extending Kharitonov's celebrated result [37] on interval polynomials, one whose coefficients lie in intervals, to more general uncertainty structures. Kharitonov showed that one need only check 4 polynomials to determine stability of the entire family of interval polynomials. Several additional results have since been proven for other special cases, such as polynomials whose coefficients are affine in some real parameters (see [11] for example), and the solutions typically involve checking the edges or vertices of some polytope in the parameter space. It can be shown that restricting the allowed perturbation dependence to be affine (or further restricting to interval polynomials) leads to a real μ problem on a transfer matrix which is rank one.

The rank one problem is studied in detail in in chapter 8, and it will be seen that for the rank one case we can guarantee a-priori that μ equals its upper bound, so that the μ problem is computationally tractable. In fact we will see that for this case one can even develop exact analytic expressions for μ. Note however that this "rank one" assumption is very restrictive. Typically robustness problems motivated by real physical systems do not satisfy this assumption.

3.3.3 Implications for Computation Schemes

Recall that the NP completeness results strongly suggest that exact methods for computing μ in the purely real or mixed case will be computationally intractable (in the worst case) for even moderate (less than 100) numbers of real perturbations. We have seen that we can indeed find special cases of the μ problem that beat the NP hardness of the general problem, and are computationally tractable. Unfortunately the restrictions we have to put on the problem to achieve this are quite severe, and so the engineering applicability of these special cases is rather limited.

Since the general mixed μ problem is NP hard, we will not attempt to solve it exactly, but rather obtain good bounds (with reasonable computational requirements). Furthermore, recent results [20] suggest that even approximate methods are also NP hard, so we will not expect good worst case behavior from our algorithms, but rather aim for good typical behavior.

In short we will aim to develop practical algorithms for medium size problems. Here medium size means less than 100 real parameters, and "practical" means avoiding exponential (nonpolynomial) growth in computation with the number of parameters for the problems which arise in engineering applications. Practical algorithms for other NP complete problems exist and typically involve approximation, heuristics, Branch and Bound, or local search. The results we will present in chapters 4-7 are aimed at developing an intelligent combination of all these techniques, and hence a practical algorithm for the mixed μ problem.

Chapter 4

Upper and Lower Bounds

In the previous chapter we found that, in general, it is difficult to compute μ exactly, so we will focus our effort on the development and computation of upper and lower bounds. Such bounds are useful in their own right however, since an upper bound gives a (possibly conservative) limit on the size of allowable perturbations, and a lower bound yields a "problem perturbation," together with an indication of the conservatism that may be present in the upper bound. Important issues now become the efficient computation of the bounds, the degree to which they approximate μ, and techniques for refining the bounds for a better approximation. In this chapter we will concentrate on the first two of these issues, and the third one will be considered in chapter 7.

For the purely complex case a tractable upper bound was suggested in [23] involving a singular value minimization (3.7). Computation schemes for lower bounds have been developed involving a smooth optimization problem, due to Fan and Tits [28], and a power algorithm, due to Packard et al. [59]. Whilst the purely complex case is by no means completely solved, these methods are now routinely applied to large engineering problems.

The mixed case however is a fundamentally more difficult problem. An upper bound was presented by Fan et al. [29] which involves minimizing the eigenvalues of a

Hermitian matrix. This is the upper bound we will use here, and it will be presented in section 4.6. The development of a practical scheme to compute this bound will be deferred to chapter 6.

For the most part this chapter is concerned with the problem of computing a lower bound for μ in the mixed case. It is shown in section 4.1 that mixed μ can be obtained as the result of a (nonconvex) real eigenvalue maximization. Sections 4.2 through 4.5 present several important theoretical characterizations of the mixed μ problem, including the generalization of the μ decomposition to the mixed case in section 4.4. This leads to the development of a power algorithm to compute a lower bound for the mixed μ problem, which is presented in section 4.5. The algorithm not only provides a lower bound for μ, but has the property that μ is (almost) always an equilibrium point of the algorithm. This power algorithm is an extension to the mixed case of Packard's algorithm [59], which in turn is an extension of standard power iterations for eigenvalues and singular values.

4.1 Lower Bound as a Maximization

The lower bound (3.6) for the mixed case is a real eigenvalue maximization problem, namely

$$\max_{Q \in \mathcal{Q}_{\mathcal{K}}} \rho_R(QM) \leq \mu_{\mathcal{K}}(M).$$

In the purely complex case ($m_r = 0$) we can replace ρ_R by ρ and it was shown by Doyle [23] that in fact this lower bound is equal to μ. This reduces the complexity of the problem in (3.9), namely

$$\max_{\Delta \in \mathbb{B}X_{\mathcal{K}}} \rho_R(\Delta M) = \mu_{\mathcal{K}}(M)$$

since, for complex perturbations, maximization over $\mathcal{Q}_{\mathcal{K}}$ amounts to maximization over the boundary of the set $\mathbb{B}X_{\mathcal{K}}$ (i. e. , unitary perturbations). This leads to efficient computation schemes for the complex μ lower bound. In this section we show that the lower bound for the mixed case (3.6) also holds with equality, and hence it is

still sufficient to consider the complex uncertainties on their boundary. We note, however, that the definition of $\mathcal{Q}_{\mathcal{K}}$ requires us to search over the full range of the real perturbations. The following lemma is taken from [23].

Lemma 4.1 ([23]) *Let $p : \mathbb{C}^k \to \mathbb{C}$ be a (multivariable) polynomial and define $\beta = min\{|z|_\infty : p(z) = 0\}$. Then there exists a $z \in \mathbb{C}^k$ such that $p(z) = 0$ and for every i, $|z_i| = \beta$.*

This is now used to prove the main result of this section.

Theorem 4.1 *For any matrix $M \in \mathbb{C}^{n\times n}$, and any compatible block structure $\mathcal{K}$*

$$\max_{Q\in\mathcal{Q}_{\mathcal{K}}} \rho_R(QM) = \mu_{\mathcal{K}}(M). \tag{4.1}$$

Proof: Trivial from (3.6) if $\mu_{\mathcal{K}}(M) = 0$. So assume $\mu_{\mathcal{K}}(M) = \beta > 0$, and this value is achieved for some perturbation $\hat{\Delta}$, i. e. , $\det(I - \hat{\Delta}M) = 0$ and $\overline{\sigma}(\hat{\Delta}) \leq \frac{1}{\beta}$. Now *fix* the real perturbations at these "optimal" values ($\delta_i^r = \hat{\delta}_i^r, i = 1, \ldots, m_r$ with $|\hat{\delta}_i^r| \leq \frac{1}{\beta}$). Then allow the complex part of Δ to vary and consider minimizing $\overline{\sigma}(\Delta)$ subject to $\det(I - \Delta M) = 0$. Performing an SVD on Δ we obtain $\det(I - U\Sigma V M) = 0$ where U and V are (block diagonal) unitary matrices and

$$\Sigma = diag(\hat{\delta}_1^r I_{k_1}, \ldots, \hat{\delta}_{m_r}^r I_{k_{m_r}}, \delta_1^c I_{k_{m_r+1}}, \ldots, \delta_{m_c}^c I_{k_{m_r+m_c}}, \gamma_1^c, \ldots, \gamma_k^c)$$

with $k = \sum_{i=m_r+m_c+1}^{m} k_i$. This is a polynomial in $\delta_1^c, \ldots, \delta_{m_c}^c, \gamma_1^c, \ldots, \gamma_k^c$ and so applying lemma 4.1 we have a solution with $|\hat{\delta}_1^c| = \ldots = |\hat{\delta}_{m_c}^c| = |\hat{\gamma}_1^c| = \ldots = |\hat{\gamma}_k^c| = \frac{1}{\hat{\beta}}$ and $\hat{\beta} \geq \beta$. Now suppose $\hat{\beta} > \beta$, say $\hat{\beta} = \beta + \epsilon$ for some $\epsilon > 0$, then since the roots of a polynomial are continuous functions of the coefficients we can find a $\delta > 0$ so that

$$|\delta_i^r - \hat{\delta}_i^r| < \delta, i = 1, \ldots, m_r \quad \longrightarrow \quad |\delta_i^c - \hat{\delta}_i^c| < \frac{\epsilon}{2}, i = 1, \ldots, m_c$$
$$|\gamma_i^c - \hat{\gamma}_i^c| < \frac{\epsilon}{2}, i = 1, \ldots, k.$$

Then move each $|\delta_i^r|$ down by $\frac{\delta}{2}$ and we can find a Δ solving $\det(I - \Delta M) = 0$ with $\overline{\sigma}(\Delta) < \frac{1}{\beta}$ contradicting the definition of μ. Thus $\hat{\beta} = \beta$ and it is now easy to check that for this solution $\beta\hat{\Delta} = \hat{Q} \in \mathcal{Q}_{\mathcal{K}}$ with $\rho_R(\hat{Q}M) = \beta = \mu_{\mathcal{K}}(M)$. □

4.2 Facts from Matrix Theory and Linear Algebra

This section presents some basic facts from matrix (perturbation) theory and linear algebra we will need in section 4.3. The material on eigenvalue perturbation theory in section 4.2.1 is fairly standard and is presented without proof (see [36] for further details).

4.2.1 Eigenvalue Perturbation Theory

Suppose we have a complex matrix $M(t) \in \mathbb{C}^{n\times n}$ depending analytically on the real parameter t. Then denote $M_0 := M(0)$ and suppose that this matrix has a distinct (i. e., algebraic multiplicity one) eigenvalue λ_0, with right and left eigenvectors x_0 and y_0 respectively, i. e., we have (after normalizing the eigenvectors appropriately)

$$\begin{aligned} y_0^* x_0 &= 1 \\ M_0 x_0 &= \lambda_0 x_0 \\ y_0^* M_0 &= \lambda_0 y_0^*. \end{aligned}$$

Then for t in a sufficiently small neighborhood of the origin $M(t)$ has an eigenvalue $\lambda(t)$, with right and left eigenvectors $x(t)$ and $y(t)$ respectively, all of which depend analytically on t, i. e., we have

$$\begin{aligned} y(t)^* x(t) &= 1 \\ M(t)x(t) &= \lambda(t)x(t) \\ y^*(t)M(t) &= \lambda(t)y^*(t) \end{aligned}$$

with $\lambda(0) = \lambda_0$, $x(0) = x_0$, and $y(0) = y_0$. Thus we can differentiate this eigenvalue with respect to t and this yields

$$\dot{\lambda}(0) = y(0)^* \dot{M}(0)x(0) = y_0^* \dot{M}(0)x_0. \tag{4.2}$$

4.2.2 Linear Algebra Lemmas

The following two linear algebra lemmas are due to Packard [59].

Lemma 4.2 ([59]) *Let $y \in \mathbb{C}^n$ and $x \in \mathbb{C}^n$ be non-zero vectors. Then there exists a $d \in \mathbb{R}, d > 0$ such that $y = dx$ if and only if $\mathrm{Re}\,(y^*Gx) \leq 0$ for every $G \in \mathbb{C}^{n\times n}$ satisfying $G + G^* \leq 0$.*

Lemma 4.3 ([59]) *Let $y \in \mathbb{C}^n$ and $x \in \mathbb{C}^n$ be non-zero vectors. Then there exists a Hermitian, positive definite $D \in \mathbb{C}^{n\times n}$ such that $y = Dx$ if and only if $y^*x \in \mathbb{R}$ and $y^*x > 0$.*

Now define the closed half-space in the complex plane as, for some scalar $\psi \in \mathbb{R}$ (see figure 4.1 for illustration)

$$H^\psi = \{z : \mathrm{Re}\,(e^{\mathbf{j}\psi}z) \leq 0\}. \tag{4.3}$$

Then we have the following elementary linear algebra lemmas.

Lemma 4.4 *Given any set of complex scalars $\mathcal{Z} = \{z_i : i = 1, \ldots, m\}$ and any real scalar ψ, then $\mathcal{Z} \subset H^\psi$ if and only if $\sum_{i=1}^m \alpha_i z_i \in H^\psi$ for all real non-negative scalars $\alpha_i, i = 1, \ldots, m$.*

Proof: ($\longleftarrow$) For each z_k choose $\alpha_k = 1$ and $\alpha_i = 0$ for $i \neq k$.
($\longrightarrow$) For any set of real non-negative scalars $\alpha_i, i = 1, \ldots, m$, we have

$$\begin{aligned} &\mathrm{Re}\left(e^{\mathbf{j}\psi}\sum_{i=1}^m \alpha_i z_i\right) \\ =\ &\sum_{i=1}^m \alpha_i \mathrm{Re}\left(e^{\mathbf{j}\psi}z_i\right) \leq 0. \end{aligned}$$ □

Lemma 4.5 *Given any set of complex scalars $\mathcal{Z} = \{z_i : i = 1, \ldots, m\}$ define $\lambda := \sum_{i=1}^m \alpha_i z_i$ where $\alpha_i, i = 1, \ldots, m$ are real non-negative scalars. Then λ is not real and positive for any choice of the above $\alpha_i's$ if and only if $\mathcal{Z} \subset H^\psi$ for some $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$.*

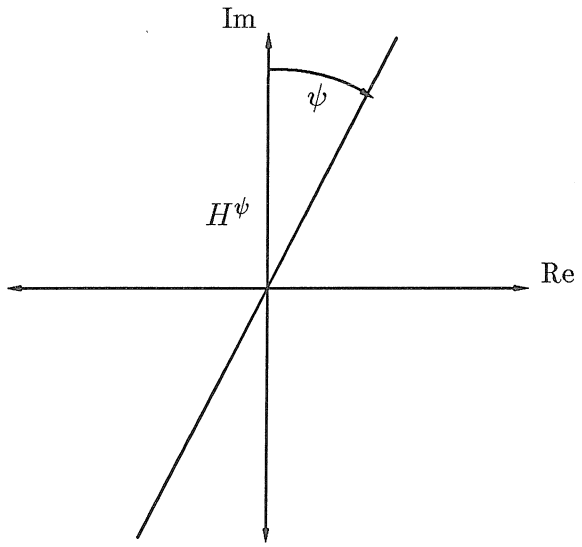


Figure 4.1: Closed half-plane H^ψ

Proof: ($\longleftarrow$) By lemma 4.4 $\mathcal{Z} \subset H^\psi$ implies $\lambda \in H^\psi$ and hence $\mathrm{Re}\,(e^{\mathbf{j}\psi}\lambda) \leq 0$. Suppose λ is real and positive. Then this implies $\mathrm{Re}\,(e^{\mathbf{j}\psi}) \leq 0$ which means $\psi \notin (-\frac{\pi}{2}\ \frac{\pi}{2})$ which is a contradiction.

($\longrightarrow$) Assume λ is never real and positive. Now suppose $\mathcal{Z} \not\subset H^\psi$ for any $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$. First choose $\psi = 0$. Then we must have at least one $z \in \mathcal{Z}$ with $\mathrm{Re}\,(z) > 0$. Now we choose $\hat{z}_1$ as the element with $\mathrm{Re}\,(z) > 0$ having minimum $|\mathrm{Arg}\,(z)|$ (which must be non-zero). Now choose $\psi = \mathrm{Arg}\,(\hat{z}_1)$, and define $\hat{\psi} = \frac{\pi}{2} - \psi$. Then since $\hat{z}_1 \in H^{\hat{\psi}}$, we must have a (non-zero) $\hat{z}_2 \in \mathcal{Z}$ with $\hat{z}_2 \notin H^{\hat{\psi}}$. Suppose

$$\hat{z}_1 = r_1(\cos\psi + \mathbf{j}\sin\psi), \qquad \hat{z}_2 = r_2(\cos\phi + \mathbf{j}\sin\phi).$$

Then by our choice of $\hat{z}_1$ and $\hat{z}_2$ straightforward trigonometry yields the following facts: $|\sin\phi| \geq |\sin\psi|$, $\mathrm{Sgn}\,(\sin\phi) = -\mathrm{Sgn}\,(\sin\psi)$, $|\cos\phi| \leq |\cos\psi|$, and if $|\cos\phi| = |\cos\psi|$ then $\cos\phi = \cos\psi$. Now choose $\hat{\alpha}_1 = \frac{1}{r_1|\sin\psi|}$ and $\hat{\alpha}_2 = \frac{1}{r_2|\sin\phi|}$. Then we

have

$$\hat{\lambda} = \hat{\alpha}_1\hat{z}_1 + \hat{\alpha}_2\hat{z}_2 = \frac{\cos\psi}{|\sin\psi|} + \frac{\cos\phi}{|\sin\phi|}.$$

Thus $\hat{\lambda}$ is real and positive which is a contradiction. □

4.3 Characterization of a Maximum Point

We are interested in computing $\mu_{\mathcal{K}}(M)$, which by (3.9) and (4.1) is given by

$$\mu_{\mathcal{K}}(M) = \max_{\Delta \in \mathbb{B}X_{\mathcal{K}}} \rho_R(\Delta M) = \max_{Q \in \mathcal{Q}_{\mathcal{K}}} \rho_R(QM).$$

For reasons of tractability we choose to consider the problem $\max_{Q\in\mathcal{Q}_{\mathcal{K}}} \rho_R(QM)$. However since this is a nonconvex problem we will in general only be able to find local maxima, and there exist examples with strictly local maxima. Thus we will obtain a lower bound for $\mu_{\mathcal{K}}(M)$ (which is the global maximum). We would like this lower bound to be "tight" (i.e., close to μ) and so we wish to rule out maxima of $\rho_R(QM)$ which we know are only local. Thus we only consider $Q \in \mathcal{Q}_{\mathcal{K}}$ which are local maxima of $\rho_R(QM)$ with respect not only to $Q \in \mathcal{Q}_{\mathcal{K}}$ but also to $Q \in \mathbb{B}X_{\mathcal{K}}$. In this section we will develop a characterization of such local maxima.

Note that for any $Q \in \mathcal{Q}_{\mathcal{K}}$ and any $\Delta \in \mathbb{B}X_{\mathcal{K}}$, then $Q\Delta \in \mathbb{B}X_{\mathcal{K}}$ and $\Delta Q \in \mathbb{B}X_{\mathcal{K}}$. Now suppose some matrix $Q \in \mathcal{Q}_{\mathcal{K}}$ achieves a local maximum of $\rho_R(QM)$ over $Q \in \mathbb{B}X_{\mathcal{K}}$. Then it is easy to show that the matrix $\hat{M} := QM$ has a local maximum of $\rho_R(\hat{Q}\hat{M})$ over $\hat{Q} \in \mathbb{B}X_{\mathcal{K}}$ at $\hat{Q} = I$. However since the real elements of Q are not restricted to be on their boundary we can say more than this. For any matrix $Q \in \mathcal{Q}_{\mathcal{K}}$ (see (2.6)) define the index sets

$$\mathcal{J}(Q) = \{i \leq m_r : |\delta_i^r| = 1\} \tag{4.4}$$

$$\hat{\mathcal{J}}(Q) = \{i \leq m_r : |\delta_i^r| < 1\} \tag{4.5}$$

and define the allowable perturbation set

$$\hat{\mathbb{B}}\Delta_\epsilon(\mathcal{J}, \hat{\mathcal{J}}) = \{\Delta \in X_{\mathcal{K}} : |\delta_i^r| \leq 1, i \in \mathcal{J}, |\delta_i^r| < 1 + \epsilon, i \in \hat{\mathcal{J}},$$
$$|\delta_i^c| \leq 1, i = 1, \ldots, m_c, \overline{\sigma}(\Delta_i^C) \leq 1, i = 1, \ldots, m_C\}. \tag{4.6}$$

We see that, for sufficiently small $\epsilon > 0$, for any $Q \in \mathcal{Q}_{\mathcal{K}}$ and any $\Delta \in \hat{\mathbb{B}}\Delta_\epsilon(\mathcal{J}(Q), \hat{\mathcal{J}}(Q))$ then $Q\Delta \in \mathbb{B}X_{\mathcal{K}}$ and $\Delta Q \in \mathbb{B}X_{\mathcal{K}}$. The point of all this is that if some matrix $Q \in \mathcal{Q}_{\mathcal{K}}$ achieves a local maximum of $\rho_R(QM)$ over $Q \in \mathbb{B}X_{\mathcal{K}}$ then the matrix $\hat{M} := QM$ has a local maximum of $\rho_R(\hat{Q}\hat{M})$ *over* $\hat{Q} \in \hat{\mathbb{B}}\Delta_\epsilon(\mathcal{J}(Q), \hat{\mathcal{J}}(Q))$ (for some $\epsilon > 0$) at $\hat{Q} = I$ (and in fact the converse is true provided we assume that for every i, $\delta_i^r \neq 0$).

The notation here is unfortunately rather cumbersome and tends to obscure what is really a rather simple concept. All that the above says is that if we are at a maximum point with some of the real perturbations at interior points (we do not need to consider this possibility for the complex perturbations) then we stay inside the allowable set, and cannot increase the function, if we move these *up* or down (in magnitude).

We introduce one further piece of notation. Suppose $M \in \mathbb{C}^{n\times n}$ has an eigenvalue λ with right and left eigenvectors x and y respectively. Then partition x and y compatibly with the block structure as

$$x = \begin{bmatrix} x_{r_1} \\ \vdots \\ x_{r_{m_r}} \\ x_{c_1} \\ \vdots \\ x_{c_{m_c}} \\ x_{C_1} \\ \vdots \\ x_{C_{m_C}} \end{bmatrix}, \qquad y = \begin{bmatrix} y_{r_1} \\ \vdots \\ y_{r_{m_r}} \\ y_{c_1} \\ \vdots \\ y_{c_{m_c}} \\ y_{C_1} \\ \vdots \\ y_{C_{m_C}} \end{bmatrix} \tag{4.7}$$

where $x_{r_i}, y_{r_i} \in \mathbb{C}^{k_i}$, $x_{c_i}, y_{c_i} \in \mathbb{C}^{k_{m_r+i}}$, $x_{C_i}, y_{C_i} \in \mathbb{C}^{k_{m_r+m_c+i}}$. These will be referred to as the "block components" of x and y, and we define the "non-degeneracy" assumption to be that for every i (in the appropriate set), $|y_{r_i}^* x_{r_i}| \neq 0, |y_{c_i}^* x_{c_i}| \neq 0, |y_{C_i}||x_{C_i}| \neq 0$.

Theorem 4.2 *Suppose the matrix $M \in \mathbb{C}^{n\times n}$ has a distinct real eigenvalue $\lambda_0 > 0$ with right and left eigenvectors x and y respectively, satisfying the non-degeneracy*

assumption. Further suppose that $\rho_R(M) = \lambda_0$. *Then if the function* $\rho_R(QM)$ *attains a local maximum over the set* $Q \in \hat{\mathbb{B}}\Delta_\epsilon(\mathcal{J}, \hat{\mathcal{J}})$ *(for some* $\epsilon > 0$*) at* $Q = I$ *then there exists a matrix* $D \in \hat{\mathcal{D}}_\mathcal{K}$, *with* $\theta_i = \pm\frac{\pi}{2}$ *for every* $i \in \hat{\mathcal{J}}$, *and a real scalar* $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$, *such that* $y = e^{\mathbf{j}\psi}Dx$.

Proof: First we parametrize the perturbation set. Consider $G \in X_\mathcal{K}$ with

$$G = \text{block diag}\,(g_1^r I_{k_1}, \ldots, g_{m_r}^r I_{k_{m_r}}, g_1^c I_{k_{m_r+1}}, \ldots, g_{m_c}^c I_{k_{m_r+m_c}}, G_1^C, \ldots, G_{m_C}^C) \quad (4.8)$$

and the added restrictions

$$\begin{aligned} g_i^r &\le 0 \quad , \quad i \in \mathcal{J} \\ \text{Re}\,(g_i^c) &\le 0 \quad , \quad i = 1, \ldots, m_c \\ G_i^C + G_i^{C*} &\le 0 \quad , \quad i = 1, \ldots, m_C. \end{aligned} \quad (4.9)$$

Now it can be shown that for some $\delta > 0$ then the set of all matrices $E(t) := (I + Gt)(I - Gt)^{-1}$ for G as above and t such that $t\overline{\sigma}(G) \in [0\ \delta)$ is an open neighborhood of $\hat{\mathbb{B}}\Delta_\epsilon(\mathcal{J}, \hat{\mathcal{J}})$ about $E(0) = I$. So now define the matrix $R(t) := E(t)M$. Then it is clear that $\rho_R(QM)$ has attained a local maximum over the set $Q \in \hat{\mathbb{B}}\Delta_\epsilon(\mathcal{J}, \hat{\mathcal{J}})$ at $Q = I$ if and only if $\rho_R(R(t))$ has attained a local maximum over $t \ge 0$ at $t = 0$ *for arbitrary* G *as above.*

Since $R(0) = M$ has a distinct real eigenvalue λ_0 we have (for some non-empty interval about the origin) an analytic function $\lambda(t)$, with $\lambda(0) = \lambda_0$, and $\lambda(t)$ an eigenvalue of $R(t)$. Thus we can differentiate to obtain

$$\dot{\lambda}(0) = y^*\dot{R}(0)x = 2y^*GMx = 2\lambda_0 y^*Gx. \quad (4.10)$$

In block notation this becomes

$$\dot{\lambda}(0) = 2\lambda_0 \left(\sum_{i=1}^{m_r} g_i^r y_{r_i}^* x_{r_i} + \sum_{i=1}^{m_c} g_i^c y_{c_i}^* x_{c_i} + \sum_{i=1}^{m_C} y_{C_i}^* G_i^C x_{C_i} \right). \quad (4.11)$$

Define the set of points

$$\begin{aligned} \mathcal{Z} = \{z_i : i = 1, \ldots, m\} = \{g_i^r y_{r_i}^* x_{r_i} : i = 1, \ldots, m_r\} \cup \\ \{g_i^c y_{c_i}^* x_{c_i} : i = 1, \ldots, m_c\} \cup \{y_{C_i}^* G_i^C x_{C_i} : i = 1, \ldots, m_C\} \end{aligned} \quad (4.12)$$

with the obvious identification for the elements z_i. Now since we are at a maximum point we have that $\dot{\lambda}(0)$ is never real and positive. Thus, noting that we may independently scale g_i^r, g_i^c, G_i^C by arbitrary non-negative scalars and still satisfy (4.9), then applying lemma 4.5 to (4.11) and (4.12) gives that this is true if and only if $\mathcal{Z} \subset H^\psi$ for some $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$ for each $G \in X_\mathcal{K}$ satisfying (4.9). Furthermore since any summation of G's satisfying (4.9) also satisfies (4.9), lemma 4.5 gives that this is true if and only if there is one H^ψ which works for *every* G, i.e., there exists $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$ such that $\mathcal{Z} \subset H^\psi$ for *all* $G \in X_\mathcal{K}$ satisfying (4.9). From the definition of H^ψ in (4.3), and G in (4.8),(4.9) this is equivalent to

$$\begin{aligned}
\mathrm{Re}\,(e^{\mathbf{j}\psi} g_i^r y_{r_i}^* x_{r_i}) \le 0 \quad &, \quad \text{for all} \ \ g_i^r \in \mathbb{R} \ \text{with} \ g_i^r \le 0, \ \ i = 1, \ldots, m_r \\
\mathrm{Re}\,(e^{\mathbf{j}\psi} g_i^r y_{r_i}^* x_{r_i}) \le 0 \quad &, \quad \text{for all} \ \ g_i^r \in \mathbb{R}, \ \ i \in \hat{\mathcal{J}} \\
\mathrm{Re}\,(e^{\mathbf{j}\psi} g_i^c y_{c_i}^* x_{c_i}) \le 0 \quad &, \quad \text{for all} \ \ g_i^c \in \mathbb{C} \ \text{with} \ \mathrm{Re}\,(g_i^c) \le 0, \ \ i = 1, \ldots, m_c \\
\mathrm{Re}\,(e^{\mathbf{j}\psi} y_{C_i}^* G_i^C x_{C_i}) \le 0 \quad &, \quad \text{for all} \ \ G_i^C \ \text{with} \ G_i^C + G_i^{C*} \le 0, \ i = 1, \ldots, m_C
\end{aligned} \tag{4.13}$$

for some $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$. It is now easy to check that the above conditions may be equivalently expressed as:

$$\begin{aligned}
\mathrm{Re}\,(e^{\mathbf{j}\psi} y_{r_i}^* x_{r_i}) \ge 0 \quad &, \quad i = 1, \ldots, m_r \\
\mathrm{Re}\,(e^{\mathbf{j}\psi} y_{r_i}^* x_{r_i}) = 0 \quad &, \quad i \in \hat{\mathcal{J}} \\
e^{\mathbf{j}\psi} y_{c_i}^* x_{c_i} \in (0\ \infty) \quad &, \quad i = 1, \ldots, m_c \\
\mathrm{Re}\,(e^{\mathbf{j}\psi} y_{C_i}^* G_i^C x_{C_i}) \le 0 \quad &, \quad \text{for all} \ \ G_i^C \ \text{with} \ G_i^C + G_i^{C*} \le 0, \ i = 1, \ldots, m_C.
\end{aligned} \tag{4.14}$$

Note that in the pure complex case the normalization condition $y^* x = 1$ implies $\psi = 0$. Since the scalar $e^{\mathbf{j}\psi}$ terms may simply be absorbed into one of the vectors we can apply lemmas 4.2 and 4.3 to each block component of x and y to obtain the equivalent conditions

$$\begin{aligned}
y_{r_i} = e^{\mathbf{j}\psi} e^{\mathbf{j}\theta_i} D_i x_{r_i} \quad &, \quad 0 < D_i = D_i^*, \ \ \theta_i \in [-\frac{\pi}{2}\ \frac{\pi}{2}], \ \ i = 1, \ldots, m_r \\
y_{r_i} = e^{\mathbf{j}\psi} e^{\mathbf{j}\theta_i} D_i x_{r_i} \quad &, \quad 0 < D_i = D_i^*, \ \ \theta_i = \pm\frac{\pi}{2}, \ \ i \in \hat{\mathcal{J}}
\end{aligned}$$

$$y_{c_i} = e^{\mathbf{j}\psi} D_i x_{c_i} \quad , \quad 0 < D_i = D_i^*, \;\; i = 1, \ldots, m_c$$
$$y_{C_i} = e^{\mathbf{j}\psi} d_i x_{C_i} \quad , \quad 0 < d_i \in \mathbb{R}, \;\; i = 1, \ldots, m_C. \tag{4.15}$$

Stacking these relations in matrix form yields $y = e^{\mathbf{j}\psi} D x$ with D of the required form. □

Remarks: We note from the proof that we immediately have a partial converse to theorem 4.2, namely that if $y = e^{\mathbf{j}\psi} D x$ under the above assumptions, then no directional derivative (in the above sense) of the eigenvalue achieving $\rho_R(QM)$ over the set $Q \in \hat{\mathbb{B}}\Delta_\epsilon(\mathcal{J}, \hat{\mathcal{J}})$ is real and positive at $Q = I$.

This alignment condition is particularly clear when interpreted from a geometric viewpoint. Consider the example illustrated in figure 4.2, which shows the block components, $y_{r_i}^* x_{r_i}$ and $y_{c_i}^* x_{c_i}$, for an example with three real blocks and any number of complex blocks. The block component $y_{r_3}^* x_{r_3}$ is associated with an internal real parameter. If we think of these components as vectors in the complex plane, then it is clear from (4.11) that with this alignment condition we can only generate $\dot{\lambda}$ in the half-space H^ψ for any allowable G. Thus we cannot make $\dot{\lambda}$ real and positive. Furthermore if we do not satisfy the alignment condition then one can easily choose G so that the summation (4.11) makes $\dot{\lambda}$ real and positive.

Note that we have made two technical assumptions in the above theorem, namely that the maximum eigenvalue was distinct, and that the block components of the eigenvectors satisfy the non-degeneracy assumption. Both of these assumptions will hold generically, and furthermore we believe that theorem 4.2 and all the results which follow can be extended to the case where both of these assumptions are removed. However this extension appears to require substantial additional technical complication. The primary motivation for the theory developed here is to lead towards a power algorithm to compute a lower bound for the mixed μ problem, and so we will not pursue these technicalities here (note that the power iteration from section 4.5 is applicable even for the cases where these assumptions do not hold).

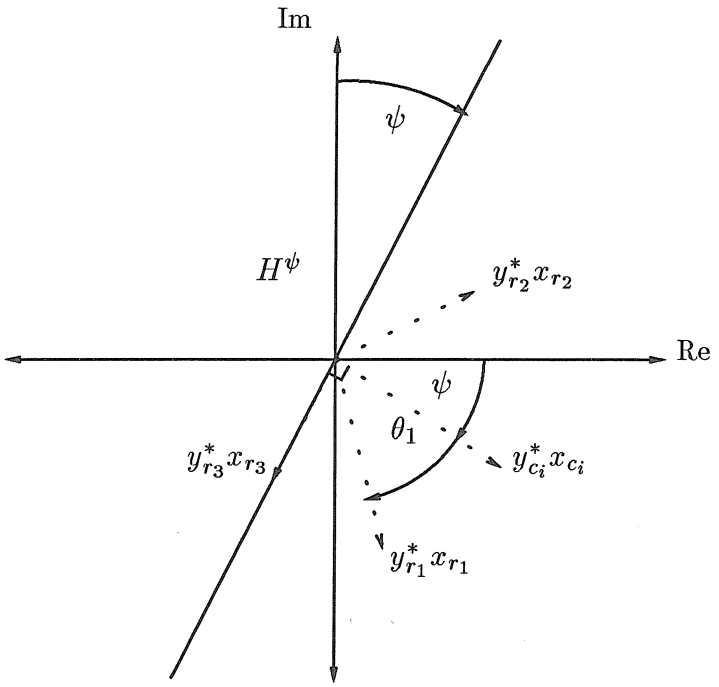


Figure 4.2: Alignment condition on eigenvector block components

4.4 A Decomposition at μ

Theorem 4.2 gives us a characterization of a maximum point of $\rho_R(QM)$ in terms of an alignment of the right and left eigenvectors of QM. This leads directly to the following decomposition.

Theorem 4.3 *Suppose $Q \in \mathcal{Q}_{\mathcal{K}}$ achieves a local maximum of $\rho_R(QM)$ over $Q \in \mathbb{B}X_{\mathcal{K}}$, and that the eigenvalue achieving $\rho_R(QM)$, denoted β, is distinct and positive. Then if the right and left eigenvectors of QM, denoted x and y respectively, satisfy the non-degeneracy assumption, there exists a matrix $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ with $D^2 \in \hat{\mathcal{D}}_{\mathcal{K}}$ and $\theta_i = \pm\frac{\pi}{4}$ for $i \in \hat{\mathcal{J}}(Q)$ such that*

$$\begin{aligned} QDMD^{-1}(Dx) &= \beta Dx \\ (x^*D^*)QD^*M(D^*)^{-1} &= \beta x^*D^* \end{aligned} \tag{4.16}$$

with $\beta \leq \mu_{\mathcal{K}}(M)$. Furthermore if the above maximum is global then $\beta = \mu_{\mathcal{K}}(M)$.

Proof: Since $Q \in \mathcal{Q}_{\mathcal{K}}$ is a local maximum of $\rho_R(QM)$ over $Q \in \mathbb{B}X_{\mathcal{K}}$, the matrix $\hat{M} := QM$ achieves a local maximum of $\rho_R(\hat{Q}\hat{M})$ over $\hat{Q} \in \hat{\mathbb{B}}\Delta_\epsilon(\mathcal{J}(Q), \hat{\mathcal{J}}(Q))$ (for some $\epsilon > 0$) at $\hat{Q} = I$. Now apply theorem 4.2 to conclude $y = e^{j\psi}\hat{D}x$ with $\hat{D} \in \hat{\mathcal{D}}_{\mathcal{K}}$ and $\hat{\theta}_i = \pm\frac{\pi}{2}$ for $i \in \hat{\mathcal{J}}(Q)$. Now define D as the unique matrix such that $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ and $D^2 = \hat{D}$. Substitution of this into the right and left eigenvalue equations of QM and simple manipulations (note that for any $Q \in \mathcal{Q}_{\mathcal{K}}$ and any $D \in \hat{\mathcal{D}}_{\mathcal{K}}$, Q and D commute) yields the results in (4.16). Finally note that from theorem 4.1 we have that $\beta \leq \mu_{\mathcal{K}}(M)$, and if the above maximum is global then $\beta = \mu_{\mathcal{K}}(M)$. □

Remarks: Employing simple manipulations of (4.16) yields a partial converse of this theorem. If we have a decomposition as in (4.16) with β real and positive and x non-zero, then we have that β is an eigenvalue of QM with right and left eigenvectors x and y respectively (thus β is a lower bound for $\mu_{\mathcal{K}}(M)$) where $y = re^{j\psi}D^2x$ with D as above, r a positive real scalar (which we could thus absorb into D), and $\psi \in [-\frac{\pi}{2}\ \frac{\pi}{2}]$. Thus defining $\hat{D} = rD^2$ we have $y = e^{j\psi}\hat{D}x$ with $\hat{D}$ as in theorem 4.2 and $\psi \in [-\frac{\pi}{2}\ \frac{\pi}{2}]$. If we add the further technical assumption that we are not in the special case of $\theta_i = \pm\frac{\pi}{4}$ for all $i = 1, \ldots, m_r$ and $m_c = 0, m_C = 0$ then we have $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$.

It is well known that for the purely complex case we have a decomposition at μ (see Packard et al. [59] and related work by Daniel et al. [18]) and (4.16) extends this result to the mixed case ($m_r \neq 0$).

Thus we (almost) always have a decomposition at μ of the form (4.16), and any such decomposition gives us a lower bound for μ. Now we reformulate this condition into a set of vector equations.

Lemma 4.6 *Suppose we have matrices $Q \in \mathcal{Q}_{\mathcal{K}}$ with $\delta_i^r \neq 0$ for $i = 1, \ldots, m_r$ and $\hat{D} \in \hat{\mathcal{D}}_{\mathcal{K}}$ with $\hat{D}^2 \in \hat{\mathcal{D}}_{\mathcal{K}}$ and $\hat{\theta}_i = \pm\frac{\pi}{4}$ for $i \in \hat{\mathcal{J}}(Q)$. Then we have a non-zero vector $\hat{x}$, and a real positive scalar β such that*

$$\begin{aligned} Q\hat{D}M\hat{D}^{-1}(\hat{D}\hat{x}) &= \beta\hat{D}\hat{x} \\ (\hat{x}^*\hat{D}^*)Q\hat{D}^*M(\hat{D}^*)^{-1} &= \beta\hat{x}^*\hat{D}^* \end{aligned} \tag{4.17}$$

if and only if there exists a matrix $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ with $\theta_i = \pm\frac{\pi}{2}$ for $i \in \hat{\mathcal{J}}(Q)$ and non-zero vectors b, a, z, w such that

$$\begin{aligned} Mb = \beta a \qquad & M^*z = \beta w \\ b = Qa \qquad & b = D^{-1}w \\ z = Q^*QDa \qquad & z = Q^*w. \end{aligned} \tag{4.18}$$

Proof: ($\longrightarrow$) Define $x = \hat{D}\hat{x}$ and b, a, z, w as

$$\begin{aligned} b = \hat{D}^{-1}x \qquad & a = \hat{D}^{-1}Q^{-1}x \\ z = \hat{D}Q^*x \qquad & w = \hat{D}x. \end{aligned}$$

Finally define $D = \hat{D}^2$ and the result follows.

($\longleftarrow$) Defining $\hat{D}$ as the unique matrix $\hat{D} \in \hat{\mathcal{D}}_{\mathcal{K}}$ such that $\hat{D}^2 = D$, and $\hat{x} = b$ the result follows directly. □

Remarks: We note that the assumption $\delta_i^r \neq 0$ for $i = 1, \ldots, m_r$ was included to ensure that Q was non-singular. This assumption was used in showing the necessity of (4.18) but was not required to show sufficiency of (4.18).

4.5 A Power Algorithm for the Lower Bound

In light of lemma 4.6 the problem of computing a lower bound for $\mu_{\mathcal{K}}(M)$ is reduced to one of finding a solution to the set of equations in (4.18) which gives us a decomposition as in (4.16). We would like to develop an algorithm for computing such a

solution, and in order to do this we first note that if we partition b, a, z, w compatibly with the block structure as in (4.7) then the set of constraint equations

$$\begin{aligned} b &= Qa & b &= D^{-1}w \\ z &= Q^*QDa & z &= Q^*w \end{aligned}$$

can be broken down into a series of m similar independent constraint equations on the block components (since Q and D are block diagonal). These equations are of three types corresponding to a repeated real scalar block, a repeated complex scalar block, or a full complex block. We now consider a generic constraint of each type. The following two lemmas are due to Packard [59].

Lemma 4.7 (Repeated Complex Scalar Block [59]) *Let $b, a, z, w \in \mathbb{C}^k$ be non-zero vectors with $a^*w \neq 0$. Then there exists a complex scalar q with $|q| = 1$, and a complex matrix $D \in \mathbb{C}^{k\times k}$ with $0 < D = D^*$ such that*

$$\begin{aligned} b &= qa & b &= D^{-1}w \\ z &= q^*qDa & z &= q^*w \end{aligned}$$

if and only if

$$z = \frac{w^*a}{|w^*a|}w \qquad b = \frac{a^*w}{|a^*w|}a. \tag{4.19}$$

Lemma 4.8 (Full Complex Block [59]) *Let $b, a, z, w \in \mathbb{C}^k$ be non-zero vectors. Then there exists a complex matrix $Q \in \mathbb{C}^{k\times k}$ with $Q^*Q = I_k$, and a real positive scalar d such that*

$$\begin{aligned} b &= Qa & b &= d^{-1}w \\ z &= Q^*Qda & z &= Q^*w \end{aligned}$$

if and only if

$$z = \frac{|w|}{|a|}a \qquad b = \frac{|a|}{|w|}w. \tag{4.20}$$

Now we consider a repeated real scalar block, bearing in mind that we have additional constraints if the real perturbation is not on the boundary (i. e. , for $i \in \hat{\mathcal{J}}(Q)$).

Lemma 4.9 (Repeated Real Scalar Block) *Let $b, a, z, w \in \mathbb{C}^k$ be non-zero vectors with $a^*w \neq 0$. Then we have a real scalar q with $|q| \leq 1$, a real scalar $\theta \in [-\frac{\pi}{2}\ \frac{\pi}{2}]$, and a complex matrix $D \in \mathbb{C}^{k\times k}$ with $0 < D = D^*$ such that*

$$\begin{aligned} b &= qa & b &= e^{-\mathrm{j}\theta}D^{-1}w \\ z &= q^*qe^{\mathrm{j}\theta}Da & z &= q^*w \end{aligned}$$

with $\theta = \pm\frac{\pi}{2}$ for $|q| < 1$ if and only if

$$z = qw \qquad b = qa \tag{4.21}$$

with

$$\begin{aligned} &\mathrm{Re}\,(a^*w) \geq 0 \quad for \quad q = 1 \\ &\mathrm{Re}\,(a^*w) \leq 0 \quad for \quad q = -1 \\ &\mathrm{Re}\,(a^*w) = 0 \quad for \quad |q| < 1. \end{aligned} \tag{4.22}$$

Proof: ($\longrightarrow$) Immediately we have $z = qw$ and $b = qa$. Thus $a^*w = \frac{1}{q}b^*w = \frac{1}{q}e^{\mathrm{j}\theta}w^*(D^*)^{-1}w$. Now $q = 1$ implies $\mathrm{Arg}\,(a^*w) = \theta$ and hence $\mathrm{Re}\,(a^*w) \geq 0$. Similarly $q = -1$ implies $\mathrm{Arg}\,(a^*w) = \theta + \pi$ and hence $\mathrm{Re}\,(a^*w) \leq 0$. Finally $|q| < 1$ implies $\mathrm{Arg}\,(a^*w) = \theta$ or $\theta + \pi$ with $\theta = \pm\frac{\pi}{2}$. Thus $\mathrm{Arg}\,(a^*w) = \pm\frac{\pi}{2}$ and so $\mathrm{Re}\,(a^*w) = 0$.

($\longleftarrow$) Immediately we have $b = qa$ and $z = q^*w$, and so $b^*w = qa^*w$. Denoting $\theta = \mathrm{Arg}\,(b^*w)$ we see that for $q = 1$ $\mathrm{Re}\,(a^*w) \geq 0$ which implies $\mathrm{Re}\,(b^*w) \geq 0$ and so $\theta \in [-\frac{\pi}{2}\ \frac{\pi}{2}]$. Similarly for $q = -1$ $\mathrm{Re}\,(a^*w) \leq 0$ which implies $\mathrm{Re}\,(b^*w) \geq 0$ and so $\theta \in [-\frac{\pi}{2}\ \frac{\pi}{2}]$. Finally for $|q| < 1$ $\mathrm{Re}\,(a^*w) = 0$ which implies $\mathrm{Re}\,(b^*w) = 0$ and so $\theta = \pm\frac{\pi}{2}$. Now $b^*(e^{-\mathrm{j}\theta}w)$ is real and positive and so applying lemma 4.3 we have a matrix $\hat{D}$ with $0 < \hat{D} = \hat{D}^*$ such that $b = e^{-\mathrm{j}\theta}\hat{D}w$. Define $D = \hat{D}^{-1}$ and we have $b = e^{-\mathrm{j}\theta}D^{-1}w$ and $z = q^*w = q^*e^{\mathrm{j}\theta}Db = q^*qe^{\mathrm{j}\theta}Da$. □

These lemmas now allow us (with a few technical assumptions) to eliminate the matrices Q and D from (4.18). In order to avoid the notation becoming excessive we

consider a simple block structure with $m_r = m_c = m_C = 1$ for the remainder of this section. *We stress that this is purely for notational convenience, and that the general formulae for an arbitrary block structure, as defined in section 2.1, are simply obtained by duplicating the appropriate formulae for each block.* So given $\mathcal{K} = (k_1, k_2, k_3)$ the appropriate scaling sets become

$$\mathcal{Q}_{sub} = \{ \text{block diag}\,(q^r I_{k_1}, q^c I_{k_2}, Q^C) : q^r \in [-1\ 1], q^{c*} q^c = 1, Q^{C*} Q^C = I_{k_3} \} \quad (4.23)$$

$$\begin{aligned} \mathcal{D}_{sub} \quad = \quad & \{ \text{block diag}\,(e^{\mathrm{j}\theta} D_1, D_2, dI_{k_3}) : \theta \in [-\frac{\pi}{2}\ \frac{\pi}{2}], \\ & 0 < D_i = D_i^* \in \mathbb{C}^{k_i \times k_i}, 0 < d \in \mathbb{R} \}. \end{aligned} \quad (4.24)$$

We partition b, a, z, w compatibly with this block structure as

$$b = \begin{bmatrix} b_1 \\ b_2 \\ b_3 \end{bmatrix}, \qquad a = \begin{bmatrix} a_1 \\ a_2 \\ a_3 \end{bmatrix}, \qquad z = \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}, \qquad w = \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix} \quad (4.25)$$

where $b_i, a_i, z_i, w_i \in \mathbb{C}^{k_i}$. Then we obtain our final form of (4.18) as in the following theorem (which will form the basis of a power iteration to compute a lower bound for $\mu_{\mathcal{K}}(M)$).

Theorem 4.4 *Suppose we have vectors $b, a, z, w \in \mathbb{C}^n$ partitioned as in (4.25) with $b_i, a_i, z_i, w_i \neq 0$ and $a_1^* w_1, a_2^* w_2 \neq 0$. Then there exist matrices $Q \in \mathcal{Q}_{sub}$ and $D \in \mathcal{D}_{sub}$, and a positive real scalar β such that*

$$\begin{aligned} Mb = \beta a \qquad & M^* z = \beta w \\ b = Qa \qquad & b = D^{-1} w \\ z = Q^* Q D a \qquad & z = Q^* w \end{aligned}$$

with $\theta \in [-\frac{\pi}{2}\ \frac{\pi}{2}]$ and $\theta = \pm\frac{\pi}{2}$ for $|q^r| < 1$ if and only if

$$\begin{aligned} & Mb = \beta a \\ & z_1 = q w_1 \qquad z_2 = \frac{w_2^* a_2}{|w_2^* a_2|} w_2 \qquad z_3 = \frac{|w_3|}{|a_3|} a_3 \\ & M^* z = \beta w \\ & b_1 = q a_1 \qquad b_2 = \frac{a_2^* w_2}{|a_2^* w_2|} a_2 \qquad b_3 = \frac{|a_3|}{|w_3|} w_3 \end{aligned} \quad (4.26)$$

for some real scalar $q \in [-1\ 1]$ *with*

$$\begin{aligned} &\mathrm{Re}\,(a_1^* w_1) \geq 0 \quad for \quad q = 1 \\ &\mathrm{Re}\,(a_1^* w_1) \leq 0 \quad for \quad q = -1 \\ &\mathrm{Re}\,(a_1^* w_1) = 0 \quad for \quad |q| < 1. \end{aligned} \tag{4.27}$$

Proof: Apply lemmas 4.7, 4.8 and 4.9 to the appropriate block components. □

Remarks: Since the relationships (4.26) and (4.27) are unaffected if we multiply b and a by an arbitrary positive real scalar α, and z and w by an arbitrary positive real scalar γ, then in searching for solutions to these equations we may impose the additional restriction $|a| = |w| = 1$.

Any solution to (4.26) and (4.27) immediately gives us a decomposition as in (4.16) and hence β is a lower bound for $\mu_{\mathcal{K}}(M)$. We also note that, under certain technical assumptions (as given), there always exists a solution to these equations with $\beta = \mu_{\mathcal{K}}(M)$. Since we would like to find the largest β we can that solves (4.26) and (4.27), we now propose finding a solution to this system of equations via the following power iteration:

$$\begin{aligned} \tilde{\beta}_{k+1} a_{k+1} &= M b_k \\ z_{1_{k+1}} &= \tilde{q}_{k+1} w_{1_k} \qquad z_{2_{k+1}} = \frac{w_{2_k}^* a_{2_{k+1}}}{|w_{2_k}^* a_{2_{k+1}}|} w_{2_k} \qquad z_{3_{k+1}} = \frac{|w_{3_k}|}{|a_{3_{k+1}}|} a_{3_{k+1}} \\ \hat{\beta}_{k+1} w_{k+1} &= M^* z_{k+1} \\ b_{1_{k+1}} &= \hat{q}_{k+1} a_{1_{k+1}} \qquad b_{2_{k+1}} = \frac{a_{2_{k+1}}^* w_{2_{k+1}}}{|a_{2_{k+1}}^* w_{2_{k+1}}|} a_{2_{k+1}} \qquad b_{3_{k+1}} = \frac{|a_{3_{k+1}}|}{|w_{3_{k+1}}|} w_{3_{k+1}} \end{aligned} \tag{4.28}$$

where $\tilde{q}_{k+1}$ and $\hat{q}_{k+1}$ evolve as

$$\tilde{\alpha}_{k+1} = \mathrm{Sgn}\,(\hat{q}_k) \frac{|b_{1_k}|}{|a_{1_{k+1}}|} + \mathrm{Re}\,(a_{1_{k+1}}^* w_{1_k})$$

$$\text{If } |\tilde{\alpha}_{k+1}| \geq 1 \qquad \text{Then } \tilde{q}_{k+1} = \frac{\tilde{\alpha}_{k+1}}{|\tilde{\alpha}_{k+1}|} \qquad \text{Else } \tilde{q}_{k+1} = \tilde{\alpha}_{k+1} \tag{4.29}$$

$$\hat{\alpha}_{k+1} = \mathrm{Sgn}\,(\tilde{q}_{k+1})\frac{|b_{1_k}|}{|a_{1_{k+1}}|} + \mathrm{Re}\,(a^*_{1_{k+1}} w_{1_{k+1}})$$

$$\text{If } |\hat{\alpha}_{k+1}| \geq 1 \qquad \text{Then } \hat{q}_{k+1} = \frac{\hat{\alpha}_{k+1}}{|\hat{\alpha}_{k+1}|} \qquad \text{Else } \hat{q}_{k+1} = \hat{\alpha}_{k+1}$$

and $\tilde{\beta}_{k+1}, \hat{\beta}_{k+1}$ are chosen positive real so that $|a_{k+1}| = |w_{k+1}| = 1$.

It is now straightforward to verify that if the algorithm converges to some equilibrium point then we satisfy the appropriate constraints on each block component and hence by lemmas 4.7, 4.8, and 4.9 we have non-zero vectors $b, a, z, w \in \mathbb{C}^n$, matrices $Q \in \mathcal{Q}_{sub}, D \in \mathcal{D}_{sub}$, and positive real scalars $\tilde{\beta}, \hat{\beta}$ such that

$$\begin{aligned} Mb &= \tilde{\beta} a & M^* z &= \hat{\beta} w \\ b &= Qa & b &= D^{-1} w \\ z &= Q^* Q D a & z &= Q^* w. \end{aligned} \tag{4.30}$$

Thus if $\tilde{\beta} = \hat{\beta}$ then we satisfy (4.18) and so have a decomposition as in (4.16), and hence $\tilde{\beta}$ is a lower bound for $\mu_{\mathcal{K}}(M)$ (associated with a local maximum of $\rho_R(QM)$). We note that if $\tilde{\beta} \neq \hat{\beta}$ then we have not found a decomposition as in (4.16). However from (4.30) we find that $QMb = \tilde{\beta} b$ and $w^* QM = \hat{\beta} w^*$. Thus we have that both $\tilde{\beta}$ and $\hat{\beta}$ are real positive eigenvalues of QM, and so by lemma 3.3, $\max(\tilde{\beta}, \hat{\beta})$ still gives us a lower bound for $\mu_{\mathcal{K}}(M)$.

Note that the equilibrium points of the algorithm are unaffected if we multiply the terms $\mathrm{Re}\,(a^*_{1_{k+1}} w_{1_k}), \mathrm{Re}\,(a^*_{1_{k+1}} w_{1_{k+1}})$ by arbitrary real positive scalars, and hence we may employ this degree of freedom to select scaling parameters so as to aid convergence. The various issues associated with developing a practical implementation of the power iteration, and the performance of the resulting algorithm, are discussed in chapter 6.

4.6 The Upper Bound

Now let us consider an upper bound for μ. As we noted earlier, one could, for the purposes of the upper bound, cover the real perturbations with complex ones

(and then use the complex μ upper bound) since this would cover the admissible perturbation set $X_{\mathcal{K}}$. Thus we obtain the familiar upper bound from complex μ theory

$$\mu_{\mathcal{K}}(M) \leq \inf_{D \in \mathcal{D}_{\mathcal{K}}} \overline{\sigma}(DMD^{-1}). \tag{4.31}$$

In order to facilitate comparison with the mixed μ upper bound from [29], note that we can reformulate this bound via the following equivalences (for $M \in \mathbb{C}^{n\times n}$ and $D \in \mathcal{D}_{\mathcal{K}}$):

$$\begin{aligned} \overline{\sigma}(DMD^{-1}) \leq \beta &\longleftrightarrow \overline{\lambda}((DMD^{-1})^*(DMD^{-1})) \leq \beta^2 \\ &\longleftrightarrow D^{-1}M^*D^2MD^{-1} - \beta^2 I_n \leq 0 \\ &\longleftrightarrow M^*D^2M - \beta^2 D^2 \leq 0. \end{aligned}$$

Since $D^2 \in \mathcal{D}_{\mathcal{K}}$ if and only if $D \in \mathcal{D}_{\mathcal{K}}$ this leads to the equivalent upper bound

$$\mu_{\mathcal{K}}(M) \leq \inf_{D \in \mathcal{D}_{\mathcal{K}}} \min_{0 \leq \beta \in \mathbb{R}} \left\{ \beta : M^*DM - \beta^2 D \leq 0 \right\}. \tag{4.32}$$

Note however that the above approach does not exploit the phase information that is present in the real perturbations, and hence this bound is frequently poor for mixed problems. The upper bound presented in [29] does exploit this phase information and gives a bound which is never worse than the standard upper bound from complex μ theory and is frequently much better. In order to present their result we first define the function $\Phi_\alpha(M, D, G)$ for matrices $M \in \mathbb{C}^{n\times n}$, $D \in \mathcal{D}_{\mathcal{K}}$, and $G \in \mathcal{G}_{\mathcal{K}}$ as

$$\Phi_\alpha(M, D, G) = \overline{\lambda}(M^*DM + \mathbf{j}(GM - M^*G) - \alpha D). \tag{4.33}$$

The following theorem, which gives an upper bound for the mixed μ problem is taken from [29].

Theorem 4.5 (Mixed μ Upper Bound [29]) *For any matrix $M \in \mathbb{C}^{n\times n}$, and any compatible block structure $\mathcal{K}$ suppose α_* is the result of the minimization problem*

$$\alpha_* = \inf_{D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}} \left[\max_{\alpha \in \mathbb{R}} \{ \alpha : \Phi_\alpha(M, D, G) \geq 0 \} \right] \tag{4.34}$$

then if $\alpha_* \leq 0$ *we have* $\mu_{\mathcal{K}}(M) = 0$, *otherwise*

$$\mu_{\mathcal{K}}(M) \leq \sqrt{\alpha_*}. \tag{4.35}$$

We can reformulate this upper bound via the following lemma.

Lemma 4.10 *For any matrices* $M \in \mathbb{C}^{n\times n}$ *and* $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ *then*

$$\max_{\alpha\in\mathbb{R}}\{\alpha : \Phi_\alpha(M, D, G) \geq 0\} = \min_{\alpha\in\mathbb{R}}\{\alpha : \Phi_\alpha(M, D, G) \leq 0\}. \tag{4.36}$$

Proof: Follows straight from the definition of $\Phi_\alpha(M, D, G)$ and continuity of eigenvalues. □

Applying this lemma to theorem 4.5 we find that the mixed μ upper bound can be alternately expressed as

$$\mu_{\mathcal{K}}(M) \leq \sqrt{\max(0, \alpha_*)} \tag{4.37}$$

where α_* is now given by

$$\alpha_* = \inf_{D\in\mathcal{D}_{\mathcal{K}}, G\in\mathcal{G}_{\mathcal{K}}} \min_{\alpha\in\mathbb{R}} \{\alpha : M^*DM + \mathbf{j}(GM - M^*G) - \alpha D \leq 0\}. \tag{4.38}$$

This expression for the upper bound can in turn be rewritten as

$$\mu_{\mathcal{K}}(M) \leq \inf_{D\in\mathcal{D}_{\mathcal{K}}, G\in\mathcal{G}_{\mathcal{K}}} \min_{0\leq\beta\in\mathbb{R}} \left\{\beta : M^*DM + \mathbf{j}(GM - M^*G) - \beta^2 D \leq 0\right\}. \tag{4.39}$$

Comparing (4.39) and (4.32) it is clear that the mixed μ upper bound in (4.39) is always at least as good as the standard upper bound from complex μ theory in (4.32), since we recover the complex μ upper bound by enforcing the choice $G = 0_n$. The G scaling matrix, which is allowed to be non-zero only for the blocks corresponding to real parameters (see the definition of $\mathcal{G}_{\mathcal{K}}$ in section 2.1), exploits the phase information that we have about the real parameters to obtain a better upper bound.

The derivation of this mixed μ upper bound is as follows. Suppose we have a feasible $Q \in \mathcal{Q}_{\mathcal{K}}$, i. e., QM has a real positive eigenvalue, or $\rho_R(QM) > 0$. Then we have a real positive scalar $0 < \gamma \in \mathbb{R}$, and a vector $x \in \mathbb{C}^{n\times n}$ with $x \neq 0$ such that

$$QMx = \gamma x.$$

Since $\overline{\sigma}(Q) \leq 1$ it follows that for any feasible x

$$\begin{aligned} \gamma^2|x|^2 &= |QMx|^2 \\ &\leq \overline{\sigma}^2(Q)|Mx|^2 \\ &\leq |Mx|^2 \end{aligned}$$

so that for any feasible x and γ we have

$$x^*(M^*M - \gamma^2 I_n)x \geq 0.$$

Now note that for this x, and *any* $G \in \mathcal{G}_{\mathcal{K}}$ we have

$$\begin{aligned} x^*GMx &= \frac{1}{\gamma}x^*M^*Q^*GMx \\ &= \frac{1}{\gamma}x^*M^*GQMx \\ &= x^*M^*Gx \end{aligned}$$

so that for *any* $G \in \mathcal{G}_{\mathcal{K}}$ and any *feasible* x we have

$$x^*(GM - M^*G)x = 0 \tag{4.40}$$

and so it follows that for such x

$$x^*(M^*M + \mathbf{j}(GM - M^*G) - \gamma^2 I_n)x \geq 0. \tag{4.41}$$

It is now easy to check, using proof by contradiction, that if we have $G \in \mathcal{G}_{\mathcal{K}}$ and $0 \leq \beta \in \mathbb{R}$ such that

$$(M^*M + \mathbf{j}(GM - M^*G) - \beta^2 I_n) \leq 0 \tag{4.42}$$

then $\gamma \leq \beta$ for all feasible γ, and hence $\mu_{\mathcal{K}}(M) \leq \beta$. Now since μ is invariant to similarity transformations with $D \in \mathcal{D}_{\mathcal{K}}$ (see lemma 3.2), we may apply this argument to $M_D \doteq DMD^{-1}$ instead of M, to reach the conclusion that if we have

$$(M_D^*M_D + \mathbf{j}(GM_D - M_D^*G) - \beta^2 I_n) \leq 0 \tag{4.43}$$

then $\mu_{\mathcal{K}}(M) \leq \beta$. Multiplying on the left and right by D does not affect the definiteness of this expression and we obtain

$$(M^*\hat{D}M + \mathbf{j}(\hat{G}M - M^*\hat{G}) - \beta^2\hat{D}) \leq 0 \tag{4.44}$$

where $\hat{D} \doteq D^2 \in \mathcal{D}_{\mathcal{K}}$ and $\hat{G} \doteq DGD \in \mathcal{G}_{\mathcal{K}}$. Since these transformations may be inverted as $D = \hat{D}^{\frac{1}{2}}, G = \hat{D}^{-\frac{1}{2}}\hat{G}\hat{D}^{-\frac{1}{2}}$ then the existence of $\hat{D} \in \mathcal{D}_{\mathcal{K}}, \hat{G} \in \mathcal{G}_{\mathcal{K}}$ satisfying (4.44) also implies $\mu_{\mathcal{K}}(M) \leq \beta$. Minimizing this expression over β and the $\hat{D}, \hat{G}$ scaling matrices gives us the upper bound in (4.39), which is equivalent to the expression in theorem 4.5.

The basic principle behind the complex μ upper bound (see (4.31) or (4.32)), which we derived in section 3.1, was to improve the crude $\overline{\sigma}$ bound via transformations that are μ-invariant (and exploit the structure of the problem). Note that here we also employ that technique to introduce the D scaling matrices to the problem. However a quite different technique is employed to introduce the G scaling matrices. There we exploit the fact that the G scaling matrices do not alter the expression

$$x^*(M^*DM + \mathbf{j}(GM - M^*G) - \beta^2 D)x$$

for any *feasible* x, though they can affect the expression for general $x \in \mathbb{C}^{n\times n}$, and hence they can affect the definiteness of the expression

$$(M^*DM + \mathbf{j}(GM - M^*G) - \beta^2 D).$$

In fact one can also derive an expression for this mixed μ upper bound in terms of scaled maximum singular values, and we will do so in chapter 6. With this expression for the upper bound there is another interesting interpretation of the G scaling matrices: we can still think of the mixed μ upper bound in terms of covering the real parameter uncertainty with complex or disk uncertainty, but now the G scales allow us to use off-axis disks (see figure 4.3). One can also think of this as a scaled small gain condition (see [84]), and for additional interpretation of the mixed μ upper bound we refer the interested reader to [34].

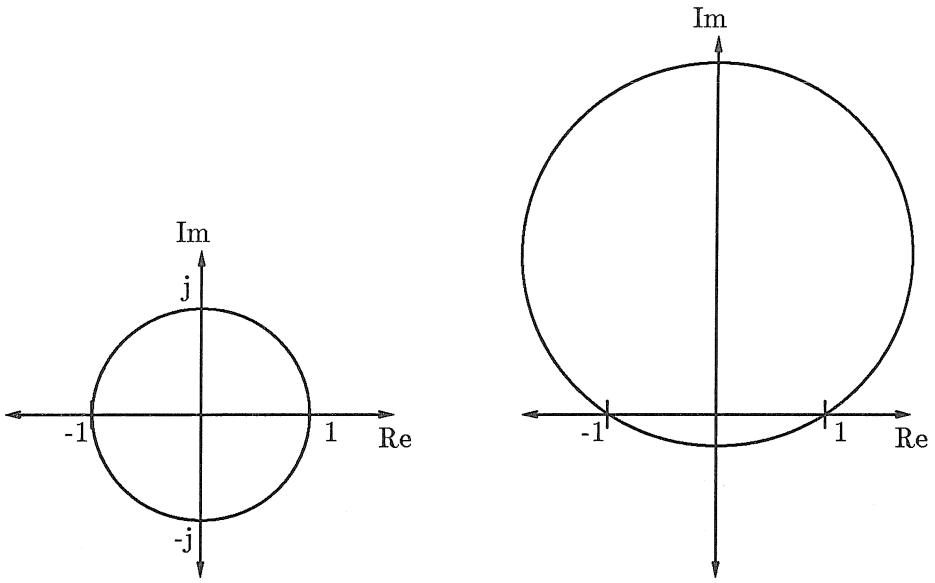


Figure 4.3: Covering real parameters with disks

The above minimization (4.39) involves a Linear Matrix Inequality, LMI, where the matrices D and G enter the problem in a linear fashion. This is very important since it implies that for fixed $\alpha \in \mathbb{R}$ the function

$$\overline{\lambda}(M^*DM + \mathbf{j}(GM - M^*G) - \alpha D) \tag{4.45}$$

is convex in both D and G (this is easily shown using the Rayleigh Quotient expression for Hermitian matrices). It is also easy to check (via the Rayleigh Quotient) that for fixed $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ the above expression (4.45) is a strictly monotonically decreasing function of α. In fact we will see in chapter 5 that the minimization over α may be simply computed as an eigenvalue problem. Thus at any point (D, G) we can easily compute the α level attained, and for any α level we have that the LMI minimization over D and G is convex. In other words we have a quasi-convex optimization problem so that all local minima are global, and hence this bound is

computationally tractable. The practical computation of the mixed μ upper (and lower) bounds will be treated in chapter 6 (see also [83]), and in the next chapter we consider some of the properties of these bounds.

Chapter 5

Properties of the Bounds

This chapter studies the relationship between μ and its bounds for the mixed case. Sections 5.1-5.3 are devoted to developing a theoretical framework for examining the equivalence between μ and its upper bound. It is hoped that this framework can then be used to identify for which block structures and/or classes of matrices the upper bound is identically equal to μ, and hence can be computed exactly (the upper bound being a convex problem). Some results in this direction are presented, and in section 5.4 we examine several special cases of the general mixed μ problem. In particular we consider the rank one case and its relation to "Kharitonov-type" analysis methods, using the above framework. This will be treated in more detail in chapter 8. The extension of the concept of 'μ values' to the mixed case is presented in section 5.5. It is shown that μ is the largest of a number of μ values, which are associated with local maxima of the lower bound function and stationary points of eigenvalues of the upper bound function (the largest such eigenvalue being associated with the upper bound). This provides a theoretical link between the upper and lower bounds, and can be used to generate guesses for the optimal value of one from the other.

5.1 Characterizing the Minimum of the Upper Bound Function

We will first consider the computation of a descent direction for the upper bound function in (4.34). This can in principle be used in a steepest descent algorithm to compute the upper bound. Our main purpose in computing it here, however, is to examine the conditions for being at a stationary point, and from these to develop some properties of the upper bound. In particular we will be concerned with characterizing the conditions under which a given pair of scaling matrices, $D_0 \in \mathcal{D}_\mathcal{K}, G_0 \in \mathcal{G}_\mathcal{K}$, represent the optimal scaling matrices for the mixed μ upper bound from theorem 4.5. It is convenient here to work with the LMI form of the mixed μ upper bound given in (4.37, 4.38). Note that this can be easily reformulated via the following lemma.

Lemma 5.1 *For any matrix $M \in \mathbb{C}^{n\times n}$ and any $D \in \mathcal{D}_\mathcal{K}, G \in \mathcal{G}_\mathcal{K}$*

$$\min_{\alpha\in\mathbb{R}}\{\alpha : (M^*DM + \mathbf{j}(GM - M^*G) - \alpha D) \leq 0\}$$

is the unique value $\overline{\alpha}$ satisfying

$$\overline{\lambda}(M^*DM + \mathbf{j}(GM - M^*G) - \overline{\alpha}D) = 0.$$

Proof: Follows straight from the Rayleigh Quotient and a simple continuity argument, since $D > 0$. □

We begin our investigation of this upper bound function with a simple lemma regarding the perturbation of negative semidefinite matrices.

Lemma 5.2 *Suppose we have matrices $A \in \mathbb{C}^{n\times n}$ and $B \in \mathbb{C}^{n\times n}$, with $A \leq 0$. Define $\mathcal{S} = \{x \in \mathbb{C}^n : x^*Ax = 0, |x| \neq 0\}$. Then we have that $A + tB < 0$ for sufficiently small $0 < t \in \mathbb{R}$ if and only if either $\mathcal{S} = \emptyset$, or $x^*Bx < 0$ for all $x \in \mathcal{S}$.*

Proof: ($\rightarrow$) Since we have that for all $x \neq 0$

$$x^*(A + tB)x = x^*Ax + tx^*Bx < 0$$

with $t > 0$, then if for any $x \neq 0$ we have $x^*Ax = 0$, for that x we must also have $x^*Bx < 0$.

($\leftarrow$) If $\mathcal{S} = \emptyset$ then $A < 0$, and so $A + tB < 0$ for sufficiently small $t > 0$ by a simple continuity argument. Suppose instead that $\mathcal{S} \neq \emptyset$, but $x^*Bx < 0$ for all $x \in \mathcal{S}$. Define $\hat{\mathcal{S}} = \mathcal{S} \cap \mathcal{B}$ where $\mathcal{B} = \{x : |x| = 1\}$. Then $\hat{\mathcal{S}} \subset \mathcal{S}$ is compact and so by continuity there exists a set $\mathcal{V} \supset \hat{\mathcal{S}}$, which is open in $\mathcal{B}$, with $x^*Bx < 0$ for all $x \in \mathcal{V}$. Thus we have

$$x^*(A + tB)x = x^*Ax + tx^*Bx < 0$$

for all $x \in \mathcal{V}$, for *any* $t > 0$. If $\mathcal{B}\backslash\mathcal{V} = \emptyset$ we are done immediately, so assume $\mathcal{B}\backslash\mathcal{V} \neq \emptyset$. Now $\mathcal{B}\backslash\mathcal{V}$ is compact and so both x^*Ax and x^*Bx achieve maxima on $\mathcal{B}\backslash\mathcal{V}$. Suppose we have

$$\max_{x \in \mathcal{B}\backslash\mathcal{V}} x^*Ax = -\alpha$$
$$\max_{x \in \mathcal{B}\backslash\mathcal{V}} x^*Bx = \beta$$

with $\alpha > 0$ since $\mathcal{B}\backslash\mathcal{V} \cap \hat{\mathcal{S}} = \emptyset$. If $\beta \leq 0$ then $x^*(A + tB)x < 0$ for all $x \in \mathcal{B}\backslash\mathcal{V}$ for *any* $t > 0$ and we are done, so assume $\beta > 0$. Then we have

$$x^*(A + tB)x = x^*Ax + tx^*Bx \leq -\alpha + t\beta < 0 \quad \texttt{for} \quad t < \frac{\alpha}{\beta}.$$

Thus we have $x^*(A + tB)x < 0$ on $\mathcal{B}\backslash\mathcal{V}$ for sufficiently small $t > 0$. Combining this with our earlier result we have that for sufficiently small $t > 0$, $x^*(A + tB)x < 0$ for all $x \in \mathcal{B}$, and hence $A + tB < 0$. □

Note that the proof of this lemma uses simple linear algebra arguments, and does not require any results about analyticity (or even continuity for that matter) of eigenvalues. Note also that clearly one can prove an analogous result for positive semidefinite matrices. This lemma immediately provides us with a characterization of when we are at the minimum of the upper bound function, or alternately a means of checking if a given D, G pair represents a descent direction, as stated in the following theorem.

Theorem 5.1 *Suppose we have matrices $M \in \mathbb{C}^{n\times n}$, $D_0 \in \mathcal{D}_{\mathcal{K}}$ and $G_0 \in \mathcal{G}_{\mathcal{K}}$ and a real scalar $\alpha \in \mathbb{R}$ such that*

$$\overline{\lambda}(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \alpha D_0) = 0.$$

Then D_0, G_0 are the minimizing arguments of the problem

$$\inf_{D\in\mathcal{D}_{\mathcal{K}}, G\in\mathcal{G}_{\mathcal{K}}} \left[\min_{\alpha\in\mathbb{R}}\{\alpha : (M^*DM + \mathbf{j}(GM - M^*G) - \alpha D) \leq 0\}\right] \tag{5.1}$$

if and only if there does not exist a pair $D \in \tilde{\mathcal{D}}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ such that

$$x^*(M^*DM + \mathbf{j}(GM - M^*G) - \alpha D)x < 0 \tag{5.2}$$

for all $x \neq 0$ with $x^*(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \alpha D_0)x = 0$.

Proof: Since the minimization is convex we always have a descent direction unless we are at a (global) minimum. Note that in order to have a descent direction for the upper bound function we need to be able to find a pair $D \in \tilde{\mathcal{D}}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ such that for sufficiently small $t > 0$ (so that $D_0 + tD > 0$) we have $\epsilon > 0$ such that

$$(M^*(D_0 + tD)M + \mathbf{j}((G_0 + tG)M - M^*(G_0 + tG)) - (\alpha - \epsilon)(D_0 + tD)) \leq 0.$$

Clearly this is equivalent to the existence of $D \in \tilde{\mathcal{D}}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ such that for sufficiently small $t > 0$ we have

$$(M^*(D_0 + tD)M + \mathbf{j}((G_0 + tG)M - M^*(G_0 + tG)) - \alpha(D_0 + tD)) < 0$$

(since $D_0 + tD > 0$). But now define

$$\begin{aligned} A &\doteq (M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \alpha D_0) \\ B &\doteq (M^*DM + \mathbf{j}(GM - M^*G) - \alpha D) \end{aligned}$$

and apply lemma 5.2. □

Now we consider the problem of choosing a D, G pair which is the steepest descent direction, or alternately verifying that no such D, G pair (i. e. , no descent direction)

exists. Note from theorem 4.5 that we are only concerned with reducing the value of the upper bound function if $\alpha > 0$ (otherwise we already have that the upper bound equals μ) and so we make the substitution $\beta = \sqrt{\alpha}$. Now suppose we have matrices $D_0 \in \mathcal{D}_\mathcal{K}$ and $G_0 \in \mathcal{G}_\mathcal{K}$ and a real scalar $\beta > 0$ such that

$$\overline{\lambda}(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) = 0$$

with r eigenvalues coalesced at zero. Further suppose that the corresponding eigenvectors are given by $U_0 \in \mathbb{C}^{n \times r}$ where

$$(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) U_0 = 0$$

and $U_0^* U_0 = I_r$. Then we know from theorem 5.1 that $-D, -G$ (for $D \in \tilde{\mathcal{D}}_\mathcal{K}, G \in \mathcal{G}_\mathcal{K}$) is a descent direction if and only if

$$x^*(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D)x > 0 \tag{5.3}$$

for all $x \neq 0$ with $x^*(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \alpha D_0)x = 0$. Defining

$$\lambda_{min} = \min_{\eta \in \mathbb{C}^r, |\eta|=1} \eta^* \left(U_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) U_0 \right) \eta \tag{5.4}$$

we see that this is equivalent to requiring that $\lambda_{min} > 0$, which states that the matrix

$$U_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) U_0 \tag{5.5}$$

is positive definite. Now define $V_0 = M U_0$ and partition U_0, V_0 compatibly with the block structure as

$$U_0 = \begin{bmatrix} A_1 \\ \vdots \\ A_{m_r} \\ B_1 \\ \vdots \\ B_{m_c} \\ C_1 \\ \vdots \\ C_{m_C} \end{bmatrix}, \qquad V_0 = \begin{bmatrix} J_1 \\ \vdots \\ J_{m_r} \\ K_1 \\ \vdots \\ K_{m_c} \\ L_1 \\ \vdots \\ L_{m_C} \end{bmatrix} \tag{5.6}$$

where $A_i, J_i \in \mathbb{C}^{k_i \times r}$, $B_i, K_i \in \mathbb{C}^{k_{m_r+i} \times r}$, $C_i, L_i \in \mathbb{C}^{k_{m_r+m_c+i} \times r}$. Then with this notation the matrix $U_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) U_0$ may be rewritten as

$$\begin{array}{c} \sum_{i=1}^{m_r} \left(J_i^* D_i J_i - \beta^2 A_i^* D_i A_i \right) + \sum_{i=1}^{m_r} \mathbf{j} \left(A_i^* G_i J_i - J_i^* G_i A_i \right) + \\ \sum_{i=1}^{m_c} \left(K_i^* D_{m_r+i} K_i - \beta^2 B_i^* D_{m_r+i} B_i \right) + \sum_{i=1}^{m_C} d_i \left(L_i^* L_i - \beta^2 C_i^* C_i \right). \end{array} \tag{5.7}$$

Now substituting this expression into (5.4), then taking traces (since all the quantities are scalar) and exchanging the order of multiplication yields upon rearrangement

$$\lambda_{min} = \min_{\eta \in \mathbb{C}^r, |\eta|=1} \left[\begin{array}{c} \sum_{i=1}^{m_r} \mathrm{Tr} \left(D_i (J_i \eta \eta^* J_i^* - \beta^2 A_i \eta \eta^* A_i^*) \right) + \\ \sum_{i=1}^{m_r} \mathrm{Tr} \left(\mathbf{j} G_i (J_i \eta \eta^* A_i^* - A_i \eta \eta^* J_i^*) \right) + \\ \sum_{i=1}^{m_c} \mathrm{Tr} \left(D_{m_r+i} (K_i \eta \eta^* K_i^* - \beta^2 B_i \eta \eta^* B_i^*) \right) + \\ \sum_{i=1}^{m_C} d_i \left(\eta^* (L_i^* L_i - \beta^2 C_i^* C_i) \eta \right) \end{array} \right]. \tag{5.8}$$

We would like to rewrite this expression as an inner product, so first of all we define the following set of block diagonal Hermitian matrices

$$\begin{array}{rl} \mathcal{Z}_{\mathcal{K}} = \{ \text{block diag} \, (Z_1, \ldots, Z_{m_r+m_c}, z_1 I_{k_{m_r+m_c+1}}, \ldots, z_{m_C} I_{k_m}, \hat{Z}_1, \ldots, \hat{Z}_{m_r}) : & \\ Z_i = Z_i^* \in \mathbb{C}^{k_i \times k_i}, z_i \in \mathbb{R}, \hat{Z}_i = \hat{Z}_i^* \in \mathbb{C}^{k_i \times k_i} \} & (5.9) \end{array}$$

which together with the inner product

$$P, T \in \mathcal{Z}_{\mathcal{K}} \qquad \langle P, T \rangle \doteq \mathrm{Tr} \, (PT) \tag{5.10}$$

forms a real inner product space. With these definitions (5.8) takes the form

$$\lambda_{min} = \min_{\eta \in \mathbb{C}^r, |\eta|=1} \langle \hat{D}, P^\eta \rangle \tag{5.11}$$

where $P^\eta \in \mathcal{Z}_{\mathcal{K}}$ is defined by

$$\begin{array}{rcl} P_i^\eta & \doteq & J_i \eta \eta^* J_i^* - \beta^2 A_i \eta \eta^* A_i^* \, , \quad i = 1, \ldots, m_r \\ P_{m_r+i}^\eta & \doteq & K_i \eta \eta^* K_i^* - \beta^2 B_i \eta \eta^* B_i^* \, , \quad i = 1, \ldots, m_c \\ p_i^\eta & \doteq & \eta^* (L_i^* L_i - \beta^2 C_i^* C_i) \eta \, , \quad i = 1, \ldots, m_C \\ \hat{P}_i^\eta & \doteq & \mathbf{j} (J_i \eta \eta^* A_i^* - A_i \eta \eta^* J_i^*) \, , \quad i = 1, \ldots, m_r \end{array} \tag{5.12}$$

and $\hat{D} \in \mathcal{Z}_{\mathcal{K}}$ is defined as $\hat{D} \doteq \text{block diag}\,(D, \hat{G})$ with $\hat{G} \doteq \text{block diag}\,(G_1, \ldots, G_{m_r})$. Now we define the generalized gradient set $\nabla_{\mathcal{Y}}$ as the set of all such P^η, i.e.,

$$\nabla_{\mathcal{Y}} \doteq \{P^\eta \in \mathcal{Z}_{\mathcal{K}} : P_i^\eta, p_i^\eta, \hat{P}_i^\eta \text{ as in } (5.12), \eta \in \mathbb{C}^r, |\eta| = 1\} \tag{5.13}$$

where $\mathcal{Y} \doteq (M, D_0, G_0, \beta)$. Thus we obtain the (generalized gradient) expression (5.8) for a fixed D, G pair as a minimization of an inner product over a compact subset of the inner product space

$$\lambda_{min} = \min_{P \in \nabla_{\mathcal{Y}}} \langle \hat{D}, P \rangle. \tag{5.14}$$

What we've shown with this development is that the question of whether or not there exists a D, G pair for which $-D, -G$ is a descent direction for the upper bound function is equivalent to the existence of a $\hat{D} \in \mathcal{Z}_{\mathcal{K}}$ for which λ_{min} given by (5.14) is strictly positive. This question can now be answered using some results from convex analysis. Denoting the convex hull of a set X by $\text{Co}\,(X)$ we have the following.

Theorem 5.2 ([56]) *Suppose X is a finite dimensional real inner product space, and Γ is a compact subset of X. Then there exists an $\hat{x} \in X$ such that $min_{y \in \Gamma} \langle \hat{x}, y \rangle > 0$ if and only if $0 \notin \text{Co}\,(\Gamma)$.*

This result immediately provides us with an answer to the question of the existence of a descent direction.

Theorem 5.3 *Suppose we have matrices $M \in \mathbb{C}^{n \times n}$, $U_0 \in \mathbb{C}^{n \times r}$ $(r \leq n)$ and a real scalar $\beta > 0$. Then there exists a $D \in \tilde{\mathcal{D}}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ such that the matrix*

$$U_0^*(M^*DM + \mathbf{j}(GM - M^*G) - \beta^2 D)U_0$$

is strictly positive definite if and only if $0 \notin \text{Co}\,(\nabla_{\mathcal{Y}})$.

Proof: Following the development in this section it suffices to simply apply theorem 5.2 to (5.14). □

Furthermore supposing in theorem 5.2 that $0 \notin \text{Co}\,(\Gamma)$ then a method to compute such an $\hat{x}$ is given in [56]. This allows us to compute a $\hat{D} \in \mathcal{Z}_{\mathcal{K}}$ for which $\lambda_{min} > 0$

and hence a D, G pair which is a descent direction. In fact the point $\hat{D}$ turns out to be the minimum point in $\mathrm{Co}\,(\nabla_{\mathcal{Y}})$ and hence corresponds to a D, G pair representing the steepest descent direction. Putting all this together gives our final characterization for the minimum of the upper bound function.

Theorem 5.4 *Suppose we have matrices $M \in \mathbb{C}^{n\times n}$, $D_0 \in \mathcal{D}_{\mathcal{K}}$ and $G_0 \in \mathcal{G}_{\mathcal{K}}$ and a real scalar $\beta > 0$ such that*

$$\overline{\lambda}(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \beta^2 D_0) = 0$$

with r eigenvalues coalesced at zero. Further suppose that the corresponding eigenvectors are given by $U_0 \in \mathbb{C}^{n\times r}$ where

$$(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \beta^2 D_0)U_0 = 0$$

and $U_0^*U_0 = I_r$. Then D_0, G_0 are minimizing arguments of the problem (with minimum value $\alpha = \beta^2$)

$$\inf_{D\in\mathcal{D}_{\mathcal{K}},G\in\mathcal{G}_{\mathcal{K}}} \left[\min_{\alpha\in\mathbb{R}}\{\alpha : (M^*DM + \mathbf{j}(GM - M^*G) - \alpha D) \leq 0\}\right] \tag{5.15}$$

if and only if $0 \in \mathrm{Co}\,(\nabla_{\mathcal{Y}})$.

Proof: Apply theorem 5.1 and theorem 5.3. □

This theorem provides us with a means to check whether or not we have achieved the minimum of the upper bound function, in terms of the properties of the set $\nabla_{\mathcal{Y}}$. We will use this expression in section 5.3 to examine the equivalence between mixed μ and its upper bound.

5.2 Connections with the Generalized Gradient

Note that the approach taken in the previous section to computing a descent direction for the upper bound function did not involve any differentiation. It is interesting then to consider how this approach relates to computing the generalized gradient,

and this is the subject of this section. In fact we will see that one obtains exactly the same characterization from either approach, although substantially more machinery is required for the generalized gradient approach followed in this section. The reader uninterested in this comparison could skip to section 5.3 without loss of continuity.

We note from (4.34) that we are interested in computing derivatives of eigenvalues of a Hermitian matrix, and so we are guaranteed the analyticity of properly chosen branches (for both eigenvalues and eigenvectors). However since we are minimizing the maximum eigenvalue, the eigenvalues may well be coalesced, and so we must compute a generalized gradient for the (possibly) repeated eigenvalues [23]. For the sake of computation we rewrite the upper bound function in (4.34) via the following lemma.

Lemma 5.3 *For any matrices $M \in \mathbb{C}^{n\times n}$, $D \in \mathcal{D}_{\mathcal{K}}$ and $G \in \mathcal{G}_{\mathcal{K}}$ the unique value α satisfying*

$$\overline{\lambda}(M^*DM + \mathbf{j}(GM - M^*G) - \alpha D) = 0$$

is given as the largest generalized eigenvalue of the generalized eigenvalue problem

$$(M^*DM + \mathbf{j}(GM - M^*G))u = \alpha Du \tag{5.16}$$

(with generalized eigenvalue and eigenvector as α and u respectively).

Proof: Follows straight from the Rayleigh Quotient and a simple continuity argument, since $D > 0$. □

Now note that the generalized eigenvalue problem (5.16) can be rewritten as a standard eigenvalue problem

$$(M_D^*M_D + \mathbf{j}(\hat{G}M_D - M_D^*\hat{G}))\hat{u} = \alpha\hat{u} \tag{5.17}$$

where we have made the substitutions

$$\begin{aligned} \hat{D} &= D^{\frac{1}{2}} \\ M_D &= \hat{D}M\hat{D}^{-1} \end{aligned}$$

$$\begin{aligned} \hat{G} &= \hat{D}^{-1} G \hat{D}^{-1} \\ \hat{u} &= \hat{D} u. \end{aligned} \tag{5.18}$$

This is now used to compute the generalized gradient of the upper bound function in (4.34). Once again we are only concerned with reducing the value of this function if $\alpha > 0$ (otherwise we already have μ) and so we make the substitution $\beta = \sqrt{\alpha}$. Now since the left-hand side of the eigenvalue equation (5.17) is for a Hermitian matrix we can perform an eigenvalue decomposition on this matrix

$$(M_D^* M_D + \mathbf{j}(\hat{G} M_D - M_D^* \hat{G})) = \hat{U} \Sigma_\beta^2 \hat{U}^* + \hat{W} \Lambda \hat{W}^* \tag{5.19}$$

where $\hat{U}$,$\hat{W}$,Σ_β,Λ satisfy

$$\begin{aligned} \Sigma_\beta &= \mathrm{diag}\,(\beta, \ldots, \beta) \\ \Lambda &= \mathrm{diag}\,(\lambda_1, \ldots, \lambda_{n-r}) \quad \lambda_i < \beta^2 \text{ for all } i \\ \begin{pmatrix} \hat{U}^* \\ \hat{W}^* \end{pmatrix} \begin{pmatrix} \hat{U} & \hat{W} \end{pmatrix} &= I_n \end{aligned} \tag{5.20}$$

(with $\Sigma_\beta \in \mathbb{R}^{r \times r}$, $\hat{U} \in \mathbb{C}^{n \times r}$, $\Lambda \in \mathbb{R}^{n-r \times n-r}$, $\hat{W} \in \mathbb{C}^{n \times n-r}$ where r is the number of generalized eigenvalues coalesced at β^2). Thus making use of the equivalence between (5.16) and (5.17) we obtain the generalized eigenvalue decomposition

$$(M^* D M + \mathbf{j}(G M - M^* G)) = D U \Sigma_\beta^2 U^* D + D W \Lambda W^* D \tag{5.21}$$

where $U = \hat{D}^{-1} \hat{U}$ and $W = \hat{D}^{-1} \hat{W}$ satisfy

$$\begin{pmatrix} U^* D U & U^* D W \\ W^* D U & W^* D W \end{pmatrix} = \begin{pmatrix} I_r & O_{r \times n-r} \\ O_{n-r \times r} & I_{n-r} \end{pmatrix}. \tag{5.22}$$

So that post multiplying (5.21) by U we finally obtain

$$(M^* D M + \mathbf{j}(G M - M^* G)) U = D U \Sigma_\beta^2. \tag{5.23}$$

These expressions (5.21,5.22,5.23) are valid for arbitrary $D \in \mathcal{D}_\mathcal{K}$ and $G \in \mathcal{G}_\mathcal{K}$, so now suppose we are at a point $D_0 \in \mathcal{D}_\mathcal{K}, G_0 \in \mathcal{G}_\mathcal{K}$, with r generalized eigenvalues

of the upper bound function coalesced at $\beta^2 > 0$, and wish to move in a direction $D \in \tilde{\mathcal{D}}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$, i.e., consider $D(t) = D_0 + Dt$ and $G(t) = G_0 + Gt$ for some real scalar t (note that $D(t) \in \mathcal{D}_{\mathcal{K}}, G(t) \in \mathcal{G}_{\mathcal{K}}$ for sufficiently small t). Then from the earlier discussion we can choose analytic matrices $U(t)$ and $\Sigma_\beta(t) = \mathrm{diag}\,(\beta_i(t))$ with $\Sigma_\beta(0) = \beta(0)I_r = \beta I_r$ such that $U(t), \Sigma_\beta(t), D(t), G(t)$ satisfy (5.21,5.22,5.23). Thus we can substitute for $U(t), \Sigma_\beta(t), D(t), G(t)$ in (5.23) and differentiate with respect to t to obtain

$$\begin{array}{c} (M^*DM + \mathbf{j}(GM - M^*G))U(t) + \\ (M^*D(t)M + \mathbf{j}(G(t)M - M^*G(t)))\dot{U}(t) \end{array} = \begin{array}{c} D(t)U(t)\Sigma_\beta(t)\dot{\Sigma}_\beta(t) + \\ D(t)U(t)\dot{\Sigma}_\beta(t)\Sigma_\beta(t) \\ + D(t)\dot{U}(t)\Sigma_\beta^2(t) + DU(t)\Sigma_\beta^2(t). \end{array} \tag{5.24}$$

Now substitute for $U(t), \Sigma_\beta(t), D(t), G(t)$ in (5.21,5.22) to derive

$$U^*(t)(M^*D(t)M + \mathbf{j}(G(t)M - M^*G(t))) = \Sigma_\beta^2(t)U^*(t)D(t). \tag{5.25}$$

Finally premultiplying (5.24) by $U^*(t)$, then substituting for (5.25) and evaluating the expression at $t = 0$ eliminates $\dot{U}(t)$ and yields upon simplification

$$\dot{\Sigma}_\beta(0) = \frac{1}{2\beta}U^*\left(M^*DM + \mathbf{j}(GM - M^*G) - \beta^2 D\right)U \tag{5.26}$$

where $U = U(0)$. In practice we would not necessarily be able to compute U (which corresponds to the correct choice of branch for the analytic matrix of generalized eigenvectors), but it is easy to show that for any matrix $U_0 \in \mathbb{C}^{n\times r}$ satisfying

$$(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \beta^2 D_0)U_0 = 0 \tag{5.27}$$

and normalized such that $U_0^*D_0U_0 = I_r$, then $U = U_0K$ for some unitary matrix $K \in \mathbb{C}^{r\times r}$ and hence on substituting this into (5.26) we obtain

$$K\dot{\Sigma}_\beta(0)K^* = \frac{1}{2\beta}U_0^*\left(M^*DM + \mathbf{j}(GM - M^*G) - \beta^2 D\right)U_0. \tag{5.28}$$

Since K is unitary and $\dot{\Sigma}_\beta(0)$ is diagonal this represents an eigenvalue decomposition and so *the derivatives of the r generalized eigenvalues coalesced at β^2 are given as the*

eigenvalues of the matrix

$$U_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) U_0. \tag{5.29}$$

We state this as a theorem.

Theorem 5.5 *Suppose we have matrices $M \in \mathbb{C}^{n \times n}$, $D_0 \in \mathcal{D}_{\mathcal{K}}$ and $G_0 \in \mathcal{G}_{\mathcal{K}}$ and a real scalar $\beta > 0$ such that*

$$\overline{\lambda}(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) = 0$$

with r generalized eigenvalues coalesced at the maximum β^2. Further suppose that the generalized eigenvectors are given by $U_0 \in \mathbb{C}^{n \times r}$ where

$$(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) U_0 = 0$$

and $U_0^* D_0 U_0 = I_r$. Then the derivatives of the r generalized eigenvalues coalesced at β^2 are given as the eigenvalues of the matrix

$$U_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) U_0. \tag{5.30}$$

Collecting all this together immediately gives us the following theorem.

Theorem 5.6 *Suppose we have matrices $M \in \mathbb{C}^{n \times n}$, $D_0 \in \mathcal{D}_{\mathcal{K}}$ and $G_0 \in \mathcal{G}_{\mathcal{K}}$ and a real scalar $\beta > 0$ such that*

$$\overline{\lambda}(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) = 0$$

with r generalized eigenvalues coalesced at the maximum β^2. Further suppose that the generalized eigenvectors are given by $U_0 \in \mathbb{C}^{n \times r}$ where

$$(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) U_0 = 0$$

and $U_0^* U_0 = I_r$. Then D_0, G_0 are minimizing arguments of the problem (with minimum value $\alpha = \beta^2$)

$$\inf_{D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}} \left[\min_{\alpha \in \mathbb{R}} \{ \alpha : (M^* D M + \mathbf{j}(GM - M^* G) - \alpha D) \leq 0 \} \right] \tag{5.31}$$

if and only if for all $D \in \tilde{\mathcal{D}}_{\mathcal{K}}$, $G \in \mathcal{G}_{\mathcal{K}}$ *the matrix*

$$U_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) U_0$$

is not positive definite.

Proof: First we note that we can find $\hat{U}_0 \in \mathbb{C}^{n \times r}$ such that

$$\hat{U}_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) \hat{U}_0$$

is positive definite, with

$$(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0)\hat{U}_0 = 0$$

and $\hat{U}_0^* D_0 \hat{U}_0 = I_r$, if and only if we can find $U_0 \in \mathbb{C}^{n \times r}$ such that

$$U_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) U_0$$

is positive definite, with

$$(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) U_0 = 0$$

and $U_0^* U_0 = I_r$. Given one simply choose the other via

$$\hat{U}_0 = U_0 (U_0^* D_0 U_0)^{-\frac{1}{2}} \ \ and \ \ U_0 = \hat{U}_0 (\hat{U}_0^* \hat{U}_0)^{-\frac{1}{2}}$$

and it is easy to verify that the stated properties hold.

Now from theorem 5.5 we can find a descent direction D, G for the upper bound function $\Phi_\alpha(M, D_0, G_0)$ at D_0, G_0 if and only if we can find $D \in \tilde{\mathcal{D}}_{\mathcal{K}}$, $G \in \mathcal{G}_{\mathcal{K}}$ such that *all* the eigenvalues of the Hermitian matrix

$$\hat{U}_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) \hat{U}_0$$

are strictly negative, or alternately strictly positive (just choose $\pm D, \pm G$), where

$$(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0)\hat{U}_0 = 0$$

and $\hat{U}_0^* D_0 \hat{U}_0 = I_r$. Thus we have a descent direction if and only if we can make the matrix

$$U_0^* \left(M^* D M + \mathbf{j}(GM - M^* G) - \beta^2 D \right) U_0$$

positive definite. Since the minimization is convex we always have a descent direction unless we are at a (global) minimum. □

The characterization in theorem 5.6 is exactly the one we arrived at in (5.5), and the rest of the analysis proceeds from that point exactly as in section 5.1. Thus we find that the simple perturbation analysis we performed in section 5.1 did in fact give us the generalized gradient descent directions, although the machinery of analytic functions was not required.

5.3 When μ Equals the Upper Bound

We are interested in examining the conditions under which the upper bound described in the preceding sections has actually achieved μ, and when this can be guaranteed. It is apparent from theorem 5.4 that it is the set $\nabla_{\mathcal{Y}}$ which determines when we are at the minimum of the upper bound function, and in fact this set is also closely related to the question of equality with μ. First of all we note that it is possible that the "inf" in (4.34) may not be achieved, and so we cannot directly apply theorem 5.4 since D_0, G_0 are not defined. In order to address these difficulties we will now introduce some new definitions.

Suppose that we have $\beta > 0$ as a candidate solution of the minimization problem (4.34) (note that we need only concern ourselves with the case that $\beta > 0$ since otherwise we have $\mu = \textit{upper bound} = 0$ immediately). Then we must be able to find sequences D^k, G^k, β_k with $D^k \in \mathcal{D}_{\mathcal{K}}, G^k \in \mathcal{G}_{\mathcal{K}}$ such that

$$\overline{\lambda}(M^* D^k M + \mathbf{j}(G^k M - M^* G^k) - \beta_k^2 D^k) = 0$$

and $\beta_k \downarrow \beta$. Then noting that we can always normalize each element of the sequence such that $\overline{\sigma}(\text{block diag}\,(D^k, G^k)) = 1$ we can always choose D^k, G^k bounded so that

by passing to a subsequence we have $D^k \to D_0$ and $G^k \to G_0$ with $D_0 \in \tilde{\mathcal{D}}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ and $D_0 \geq 0$. Furthermore we have

$$\overline{\lambda}(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) = 0$$

with (say) r eigenvalues coalesced at zero. So now define $U_0 \in \mathbb{C}^{n \times r}$ as any matrix satisfying

$$(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) U_0 = 0$$

and normalized such that $U_0^* U_0 = I_r$, and as before define $V_0 = M U_0$. Then we define the block components of U_0, V_0 by (5.6) as before. Note that if the minimization problem (4.34) is achieved then these definitions coincide with those previously given.

Now suppose we have $\eta \in \mathbb{C}^r, |\eta| = 1$ and $q = (q_1, \ldots, q_{m_r})$ with $q_i \in [-1\ 1]$. Then we define the block diagonal matrix $P^{\eta,q} \in \mathcal{Z}_{\mathcal{K}}$ as

$$\begin{aligned}
P_i^{\eta,q} &\doteq q_i^2 J_i \eta\eta^* J_i^* - \beta^2 A_i \eta\eta^* A_i^* \ , \quad i = 1, \ldots, m_r \\
P_{m_r+i}^{\eta,q} &\doteq K_i \eta\eta^* K_i^* - \beta^2 B_i \eta\eta^* B_i^* \ , \quad i = 1, \ldots, m_c \\
p_i^{\eta,q} &\doteq \eta^* (L_i^* L_i - \beta^2 C_i^* C_i) \eta \ , \quad i = 1, \ldots, m_C \\
\hat{P}_i^{\eta,q} &\doteq \mathbf{j} q_i (J_i \eta\eta^* A_i^* - A_i \eta\eta^* J_i^*) \ , \quad i = 1, \ldots, m_r
\end{aligned} \tag{5.32}$$

and we define $\hat{\nabla}_{\mathcal{Y}}$ as the set of all such $P^{\eta,q}$, i. e. ,

$$\begin{aligned}
\hat{\nabla}_{\mathcal{Y}} \doteq \{P^{\eta,q} \in \mathcal{Z}_{\mathcal{K}} : P_i^{\eta,q}, p_i^{\eta,q}, \hat{P}_i^{\eta,q} \ &as\ in\ (5.32), \eta \in \mathbb{C}^r, |\eta| = 1, \\
&q = (q_1, \ldots, q_{m_r}), q_i \in [-1\ 1]\}.
\end{aligned} \tag{5.33}$$

Note that the set $\hat{\nabla}_{\mathcal{Y}}$ is closely related to $\nabla_{\mathcal{Y}}$, and in particular $\nabla_{\mathcal{Y}} \subset \hat{\nabla}_{\mathcal{Y}}$ for any $\mathcal{Y}$. The reason we have introduced $\hat{\nabla}_{\mathcal{Y}}$ is that this set determines whether or not μ equals its upper bound.

Theorem 5.7 *Suppose we have $M \in \mathbb{C}^{n \times n}$ together with D_0, G_0 and $\beta > 0$ as defined in the preceding discussion. Then $\beta = \mu_{\mathcal{K}}(M)$ if and only if $0 \in \hat{\nabla}_{\mathcal{Y}}$.*

Proof: The style of proof follows that for the purely complex case (see [56]), by proving the equivalence of the following four statements:

1. $0 \in \hat{\nabla}_{\mathcal{Y}}$.

2. There exists $\eta \in \mathbb{C}^r, |\eta| = 1$ and $Q \in \mathcal{Q}_{\mathcal{K}}$ such that $QV_0\eta = \beta U_0\eta$.

3. There exists $x \in \mathbb{C}^n, |x| = 1$ and $Q \in \mathcal{Q}_{\mathcal{K}}$ such that $QMx = \beta x$.

4. $\beta = \mu_{\mathcal{K}}(M)$.

1→2 By definition $0 \in \hat{\nabla}_{\mathcal{Y}}$ implies that for some $\eta \in \mathbb{C}^r, |\eta| = 1$ and some $q = (q_1, \ldots, q_{m_r}), q_i \in [-1\ 1]$ we have

$$\begin{aligned} q_i^2 J_i\eta\eta^* J_i^* - \beta^2 A_i\eta\eta^* A_i^* &= 0\ , \quad i = 1, \ldots, m_r \\ K_i\eta\eta^* K_i^* - \beta^2 B_i\eta\eta^* B_i^* &= 0\ , \quad i = 1, \ldots, m_c \\ \eta^*(L_i^* L_i - \beta^2 C_i^* C_i)\eta &= 0\ , \quad i = 1, \ldots, m_C \\ \mathbf{j}q_i(J_i\eta\eta^* A_i^* - A_i\eta\eta^* J_i^*) &= 0\ , \quad i = 1, \ldots, m_r. \end{aligned} \tag{5.34}$$

The first equation implies that we have phases $e^{\mathbf{j}\theta_i}$ such that $q_i e^{\mathbf{j}\theta_i} J_i\eta = \beta A_i\eta$, and substituting these into the fourth equation then yields $\pm q_i J_i\eta = \beta A_i\eta$. The second equation gives us phases $e^{\mathbf{j}\theta_i}$ such that $e^{\mathbf{j}\theta_i} K_i\eta = \beta B_i\eta$. The third equation gives $|L_i\eta| = \beta|C_i\eta|$ which implies that there exist unitary matrices Q_i such that $Q_i L_i\eta = \beta C_i\eta$. Stacking these relationships up in block diagonal form gives statement 2.

2→1 The block components of the relationship $QV_0\eta = \beta U_0\eta$ immediately give that the equations in (5.34) hold and hence $0 \in \hat{\nabla}_{\mathcal{Y}}$.

2→3 Substituting for $V_0 = MU_0$ in statement 2, and then defining $x = U_0\eta$ gives statement 3.

3→2 Since $\overline{\sigma}(Q) \leq 1$, $QMx = \beta x$ implies

$$|D_0^{\frac{1}{2}} Mx| \geq |QD_0^{\frac{1}{2}} Mx| = |D_0^{\frac{1}{2}} QMx| = \beta|D_0^{\frac{1}{2}} x|. \tag{5.35}$$

Also it can be shown (see section 4.6 or [29]) that since x satisfies $QMx = \beta x$, then

$$x^*(G_0M - M^*G_0)x = 0$$

for any $G_0 \in \mathcal{G}_\mathcal{K}$. Thus we have

$$x^*(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \beta^2 D_0)x = |D_0^{\frac{1}{2}}Mx|^2 - \beta^2|D_0^{\frac{1}{2}}x|^2. \quad (5.36)$$

So that applying (5.35) we obtain

$$x^*(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \beta^2 D_0)x \geq 0.$$

Since the matrix $(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \beta^2 D_0)$ is negative semidefinite this implies

$$x^*(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \beta^2 D_0)x = 0$$

and hence by Rayleigh Quotient theory $x = U_0\eta$ for some $\eta \in \mathbb{C}^r, |\eta| = 1$. Substituting for $x = U_0\eta$ and $MU_0 = V_0$ gives statement 2.

3→4 Statement 3 implies $\beta \leq \mu_\mathcal{K}(M)$ but we already have $\beta \geq \mu_\mathcal{K}(M)$ and hence $\beta = \mu_\mathcal{K}(M)$.

4→3 This follows immediately from theorem 4.1. □

Note that when then upper bound minimization problem (4.34) is achieved then by theorem 5.4 we have $0 \in \mathrm{Co}\,(\nabla_\mathcal{Y})$ and hence, since $\nabla_\mathcal{Y} \subset \hat{\nabla}_\mathcal{Y}$, we have $0 \in \mathrm{Co}\,(\hat{\nabla}_\mathcal{Y})$. We conjecture that in the general case the following holds, whether or not the "inf" is achieved in (4.34).

Conjecture 5.1 *Suppose we have $M \in \mathbb{C}^{n\times n}$ together with D_0, G_0 and $\beta > 0$ as defined in the preceding discussion. Then $\alpha_* = \beta^2$ solves the upper bound minimization problem*

$$\inf_{D\in\mathcal{D}_\mathcal{K}, G\in\mathcal{G}_\mathcal{K}} \left[\min_{\alpha\in\mathbb{R}}\{\alpha : (M^*DM + \mathbf{j}(GM - M^*G) - \alpha D) \leq 0\}\right] \quad (5.37)$$

if and only if $0 \in \mathrm{Co}\,(\hat{\nabla}_\mathcal{Y})$.

Note that whenever we can show (for some class of block structure $\mathcal{K}$ and/or some class of matrices M) that $0 \in \mathrm{Co}\,(\hat{\nabla}\mathcal{Y})$ implies $0 \in \hat{\nabla}\mathcal{Y}$ then, if true, this conjecture (with theorem 5.7) implies that μ (for these classes) is identically equal to its upper bound from theorem 4.5. It is hoped that this framework can be used to establish when this equality can be guaranteed for the mixed case. This will be a subject of further research.

The equivalence between μ and its upper bound is of particular interest since the upper bound is a convex optimization problem (see section 4.6), and hence can be computed exactly. Note that although the lower bound from (4.1) is *always* equal to μ *if* one finds the global maximum, it is a nonconvex problem (and hence one cannot guarantee to find the global maximum).

As an illustration of the use of this machinery, consider the following theorem, which was originally presented in [29], and is proven here using the methods developed above.

Theorem 5.8 ([29]) *Suppose we have $M \in \mathbb{C}^{n\times n}$, then provided the infimum in (4.34) is achieved and the corresponding largest eigenvalue of*

$$(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \alpha D_0)$$

is distinct, then $\mu_{\mathcal{K}}(M)$ equals its upper bound from theorem 4.5.

Proof: Suppose $\beta \geq 0$ is the upper bound from theorem 4.5. If $\beta = 0$ we are done immediately, so assume $\beta > 0$. Then since the upper bound is achieved theorem 5.4 implies $0 \in \mathrm{Co}\,(\nabla\mathcal{Y})$. Since the corresponding largest eigenvalue of

$$(M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \alpha D_0)$$

is distinct the set $\nabla\mathcal{Y}$ is a single point (see (5.12) and (5.13)) and so $0 \in \nabla\mathcal{Y}$. But now $\nabla\mathcal{Y} \subset \hat{\nabla}\mathcal{Y}$, so $0 \in \hat{\nabla}\mathcal{Y}$ and hence by theorem 5.7, $\beta = \mu_{\mathcal{K}}(M)$. □

Note that for this case we have that $0 \in \nabla\mathcal{Y}$. It can be shown, using methods similar to the proof of theorem 5.7 (see theorem 6.3 later), that this means we can

choose the worst case perturbation $Q \in \mathcal{Q}_{\mathcal{K}}$ so that in fact $Q \in \mathcal{U}_{\mathcal{K}}$. In other words the worst case perturbation is on a vertex. This limits the applicability of the above result, since we know that this is often not the case. The reason for this restriction is the assumption that the infimum is achieved, and in fact we can extend this result to remove this limitation.

Theorem 5.9 *Suppose we have $M \in \mathbb{C}^{n\times n}$ together with D_0, G_0 and $\beta \geq 0$ as in theorem 5.7. Then if the maximum eigenvalue of*

$$M^*D_0M + \mathbf{j}(G_0M - M^*G_0) - \beta^2 D_0$$

is distinct, $\beta = \mu_{\mathcal{K}}(M)$.

Proof: If $\beta = 0$ the result is trivial (since β is an upper bound for $\mu_{\mathcal{K}}(M)$), so assume $\beta > 0$. Choose x as the unit norm eigenvector, corresponding to the maximum eigenvalue. Then it is easy to check, via proof by contradiction and lemma 5.2, that we must have

$$\begin{aligned} x^*(M^*DM - \beta^2 D)x &\geq 0 \quad \text{for all} \quad D \in \mathcal{D}_{\mathcal{K}} \\ x^*(GM - M^*G)x &\geq 0 \quad \text{for all} \quad G \in \mathcal{G}_{\mathcal{K}}. \end{aligned}$$

But now by continuity, and the definition of $\mathcal{D}_{\mathcal{K}}, \mathcal{G}_{\mathcal{K}}$, this implies that in fact we have

$$\begin{aligned} x^*(M^*DM - \beta^2 D)x &\geq 0 \quad \text{for all} \quad D \in \tilde{\mathcal{D}}_{\mathcal{K}}, D \geq 0 \\ x^*(GM - M^*G)x &= 0 \quad \text{for all} \quad G \in \mathcal{G}_{\mathcal{K}}. \end{aligned} \tag{5.38}$$

Now suppose that $(Mx)_i$ and x_i represent one of the block components of (Mx) and x. Further suppose that D_i and G_i represent the corresponding block for D and G. We will consider separately the three types of block components. In each case we will choose every other block of D and G to be identically zero.

Consider first a full complex block. Choose $D_i = I$ and (5.38) implies that $|(Mx)_i| \geq \beta |x_i|$. Thus there exists a matrix Δ_i, with $\overline{\sigma}(\Delta_i) \leq 1$, such that $\Delta_i(Mx)_i = \beta x_i$.

For a repeated complex scalar block we immediately have the above. Then choose D_i as a positive semidefinite matrix with a kernel spanned by $(Mx)_i$. Thus we have $D_i(Mx)_i = 0$, so that applying (5.38) we find that $D_i x_i = 0$. By construction of D_i this implies there exists a complex scalar δ_i, such that $\delta_i(Mx)_i = \beta x_i$, and from earlier we may take $|\delta_i| \leq 1$.

For a repeated real scalar block we immediately have the above. Now choose $G_i = I$ and (5.38) implies that $x_i^*(Mx)_i = (Mx)_i^* x_i$. Substituting for the above this implies $\delta_i^*|(Mx)_i|^2 = \delta_i|(Mx)_i|^2$ and so we may take $\delta_i \in \mathbb{R}$.

Applying the above relationships to each block component, and stacking them up, we obtain $\Delta \in \mathbb{B}X_{\mathcal{K}}$ such that

$$\Delta Mx = \beta x.$$

Thus $\beta \leq \mu_{\mathcal{K}}(M)$ and hence $\beta = \mu_{\mathcal{K}}(M)$. □

Note that here we do not assume the infimum in the upper bound is achieved, and we find that the worst case perturbation is not necessarily on a vertex. Of course for the complex blocks we can always restrict our attention to the boundary of the uncertainty set, but for the real uncertainties this is not the case. In fact there is a strong association between the presence of internal reals, and the infimum not being achieved in the upper bound, and this will surface again in chapters 6 and 8.

5.4 Special Cases

In this section we consider the application of the results developed above to several special cases of interest. The rank one matrix case will receive a preliminary treatment here, and will be treated in detail in chapter 8.

5.4.1 Some Simple Special Cases

Here we consider some elementary special cases for which computation of μ is easy. These results are simple extensions of results for the complex μ case and we include

them here for completeness. First note that for Hermitian matrices we can trivially obtain the following result.

Lemma 5.4 *For any Hermitian matrix $M \in \mathbb{C}^{n\times n}$ and any compatible block structure $\mathcal{K}$, then $\mu_{\mathcal{K}}(M) = \overline{\sigma}(M)$.*

Proof: Noting that for a Hermitian matrix, $\rho_R(M) = \rho(M) = \overline{\sigma}(M)$, the result follows from (3.3) . □

Next we consider positive matrices, i. e., matrices whose elements are positive real numbers. For these matrices there is a wealth of results from Perron-Frobenius theory (see [33] for example), regarding eigenvalues and singular values, and these lead to the following result for μ.

Lemma 5.5 *For any positive matrix $M \in \mathbb{R}^{n\times n}$ and any compatible block structure $\mathcal{K}$ with $m_C = 0$ (i. e., only scalar uncertainties), then $\mu_{\mathcal{K}}(M) = \rho(M)$.*

Proof: Since M is positive then from standard properties of positive matrices we have that there is a real positive eigenvalue equal to $\rho(M)$ (so that $\rho_R(M) = \rho(M)$), and furthermore the corresponding right eigenvector may be taken to have all its elements real and positive (see theorem 8.2.11 in [33]). By duality it follows that we may take the left eigenvector to be positive as well. Thus denoting $\lambda \doteq \rho(M)$, we have positive (hence non-zero) vectors x, y such that

$$\begin{aligned} Mx &= \lambda x \\ y^T M &= \lambda y^T. \end{aligned}$$

Since the vectors x, y are positive we may simply define

$$d_i \doteq \sqrt{\frac{y_i}{x_i}} \quad \text{for} \quad i = 1, \ldots, n$$

and $D \doteq \text{diag}\,(d_1, \ldots, d_n)$ satisfies $D \in \mathcal{D}_{\mathcal{K}}$ and $y = D^2 x$. It is easy to check that the vector $w \doteq Dx$ is positive (and hence non-zero) and satisfies

$$\begin{aligned} DMD^{-1} w &= \lambda w \\ w^T DMD^{-1} &= \lambda w^T. \end{aligned}$$

Thus we obtain

$$(DMD^{-1})^T(DMD^{-1})w = \lambda^2 w.$$

Since the matrix $(DMD^{-1})^T(DMD^{-1})$ is positive, the fact that w is positive implies that it corresponds to the maximum eigenvalue (see corollary 8.1.30 in [33]), so that $\rho((DMD^{-1})^T(DMD^{-1})) = \lambda^2$ and hence $\overline{\sigma}(DMD^{-1}) = \lambda = \rho(M) = \rho_R(M)$. By lemma 3.3 this implies that $\rho(M) = \mu_{\mathcal{K}}(M)$. □

These two cases are not of too much interest in themselves since they are rarely encountered in practice. However they may be of some interest in providing crude bounds for more general problems. One example of such an application for the complex μ problem is provided in [68] where the optimal scalings for the positive matrix case are used to approximate the optimal scalings for a more general μ problem.

Now consider the class of matrices that have the structure

$$M = \begin{bmatrix} 0 & M_{12} \\ M_{21} & 0 \end{bmatrix} \tag{5.39}$$

with $M_{12} \in \mathbb{C}^{n_1 \times n_2}, M_{21} \in \mathbb{C}^{n_2 \times n_1}$ and $n_1 + n_2 = n$. Suppose we have block structures $\mathcal{X}_{\mathcal{K}_1} \subset \mathbb{C}^{n_1 \times n_1}$ and $\mathcal{X}_{\mathcal{K}_2} \subset \mathbb{C}^{n_2 \times n_2}$, then the block structure $\mathcal{X}_{\hat{\mathcal{K}}}$ defined as

$$\mathcal{X}_{\hat{\mathcal{K}}} = \{\Delta = \text{block diag}\,(\Delta_1, \Delta_2) : \Delta_1 \in \mathcal{X}_{\mathcal{K}_1}, \Delta_2 \in \mathcal{X}_{\mathcal{K}_2}\} \tag{5.40}$$

is compatible with M.

Lemma 5.6 *For μ problems as above and any $0 < k \in \mathbb{R}$,*

$$\mu^2_{\hat{\mathcal{K}}} \begin{pmatrix} 0 & M_{12} \\ M_{21} & 0 \end{pmatrix} < \frac{1}{k}$$

if and only if

$$\rho_R(\Delta_2 M_{21} \Delta_1 M_{12}) < \frac{1}{k} \quad \text{for all} \quad \Delta_1 \in \mathbb{B}\mathcal{X}_{\mathcal{K}_1}, \Delta_2 \in \mathbb{B}\mathcal{X}_{\mathcal{K}_2}.$$

Proof: The proof of this lemma is almost identical to the proof for the complex case given in table 11.4-1 of [47], namely

$$\mu^2_{\hat{\mathcal{K}}}\begin{pmatrix} 0 & M_{12} \\ M_{21} & 0 \end{pmatrix} < \tfrac{1}{k}$$

$$\longleftrightarrow \quad \mu_{\hat{\mathcal{K}}}\begin{pmatrix} 0 & \frac{M_{12}}{k} \\ M_{21} & 0 \end{pmatrix} < 1$$

$$\longleftrightarrow \quad \det\begin{pmatrix} I & -\frac{\Delta_1 M_{12}}{k} \\ -\Delta_2 M_{21} & I \end{pmatrix} \neq 0 \qquad \text{for all} \quad \Delta_1 \in \mathbb{B}\mathcal{X}_{\mathcal{K}_1}, \Delta_2 \in \mathbb{B}\mathcal{X}_{\mathcal{K}_2}$$

$$\longleftrightarrow \quad \det\left(I - \tfrac{\Delta_2 M_{21}\Delta_1 M_{12}}{k}\right) \neq 0 \qquad \text{for all} \quad \Delta_1 \in \mathbb{B}\mathcal{X}_{\mathcal{K}_1}, \Delta_2 \in \mathbb{B}\mathcal{X}_{\mathcal{K}_2}$$

$$\longleftrightarrow \quad \rho_R(\Delta_2 M_{21}\Delta_1 M_{12}) < \tfrac{1}{k} \qquad \text{for all} \quad \Delta_1 \in \mathbb{B}\mathcal{X}_{\mathcal{K}_1}, \Delta_2 \in \mathbb{B}\mathcal{X}_{\mathcal{K}_2}. \qquad \square$$

Combining this lemma with the special cases listed in properties (d)-(f) in section 3.1, we obtain the results listed in table 5.1. For comparison the complex μ version of these results is given in table 11.4-1 of [47].

Case	Block Structure Constraints	$\mu^2_{\hat{\mathcal{K}}}\begin{pmatrix} 0 & M_{12} \\ M_{21} & 0 \end{pmatrix}$
1	$\mathcal{X}_{\mathcal{K}_1} = \mathbb{C}^{n_1 \times n_1}, \mathcal{X}_{\mathcal{K}_2} = \mathbb{C}^{n_2 \times n_2}$	$\overline{\sigma}(M_{12})\overline{\sigma}(M_{21})$
2	$\mathcal{X}_{\mathcal{K}_1} = \{\delta^c I_{n_1} : \delta^c \in \mathbb{C}\}, m_{r_2} = 0$	$\mu_{\mathcal{K}_2}(M_{21}M_{12})$
3	$m_{r_1} = 0, \mathcal{X}_{\mathcal{K}_2} = \{\delta^c I_{n_2} : \delta^c \in \mathbb{C}\}$	$\mu_{\mathcal{K}_1}(M_{12}M_{21})$
4	$\mathcal{X}_{\mathcal{K}_1} = \{\delta^c I_{n_1} : \delta^c \in \mathbb{C}\}, \mathcal{X}_{\mathcal{K}_2} = \{\delta^c I_{n_2} : \delta^c \in \mathbb{C}\}$	$\rho(M_{12}M_{21}) = \rho(M_{21}M_{12})$
5	$\mathcal{X}_{\mathcal{K}_1} = \{\delta^r I_{n_1} : \delta^r \in \mathbb{R}\}$	$\mu_{\mathcal{K}_2}(M_{21}M_{12})$
6	$\mathcal{X}_{\mathcal{K}_2} = \{\delta^r I_{n_2} : \delta^r \in \mathbb{R}\}$	$\mu_{\mathcal{K}_1}(M_{12}M_{21})$
7	$\mathcal{X}_{\mathcal{K}_1} = \{\delta^r I_{n_1} : \delta^r \in \mathbb{R}\}, \mathcal{X}_{\mathcal{K}_2} = \{\delta^r I_{n_2} : \delta^r \in \mathbb{R}\}$	$\rho_R(M_{12}M_{21}) = \rho_R(M_{21}M_{12})$

Table 5.1: Simple special cases of mixed μ

5.4.2 The Rank One Case and "Kharitonov-Type" Results

Here we present a preliminary examination of the the rank one mixed μ problem. This will be studied in detail in chapter 8, where it will be seen that this rank one case corresponds to the so called "affine parameter variation" problem for a polynomial with perturbed coefficients which has also been examined in detail in the literature, and for which several celebrated "Kharitonov-type" results have been proven (see [11] for example).

The mixed μ problem when M is rank one is studied in [17]. The authors develop an analytic expression for the solution to this problem, which is easily computable, having sublinear growth in required computation with the problem size. The authors then examine several "Kharitonov-type" results from the literature, and they note that all these problems can be treated as special cases of "rank one μ problems" and are thus "relatively easy to solve." Even the need to check (a combinatoric number of) edges is shown to be unnecessary.

This rank one case can also be addressed within the framework developed here for examining the equivalence between μ and its upper bound. First we need a preliminary lemma.

Lemma 5.7 *Suppose we have vectors $u, v \in \mathbb{C}^n$. Let $M \in \mathbb{C}^{n\times n}$ be given by*

$$M = uv^* + vu^*. \tag{5.41}$$

Then M cannot have a non-zero repeated eigenvalue.

Proof: In fact we will prove something stronger than the stated result. Note that M is a Hermitian matrix of rank at most two. Thus M has at most two non-zero eigenvalues, and all eigenvalues are real. If M has less than two non-zero eigenvalues then we are done immediately so assume that M has two non-zero eigenvalues λ_1 and λ_2. We will show that these eigenvalues are of opposite sign. First note that by projecting into the eigenspace of λ_1, λ_2 we may without loss of generality assume

$u, v \in \mathbb{C}^2$. Then we have

$$\det(uv^* + vu^*) = \det\left([u\ v]\begin{bmatrix} v^* \\ u^* \end{bmatrix}\right) = \det([u\ v])\det\left(\begin{bmatrix} v^* \\ u^* \end{bmatrix}\right)$$
$$= \det([u\ v])(\det([v\ u]))^* = -\det([u\ v])(\det([u\ v]))^* = -|\det([u\ v])|^2 \leq 0.$$

Thus $\det(uv^* + vu^*) = \lambda_1\lambda_2 \leq 0$. □

The following theorem gives a partial answer to the rank one mixed μ problem.

Theorem 5.10 *Suppose we have a rank one matrix $M \in \mathbb{C}^{n\times n}$, then provided the infimum in (4.34) is achieved, $\mu_{\mathcal{K}}(M)$ equals its upper bound from theorem 4.5.*

Proof: Suppose $\beta \geq 0$ is the upper bound from theorem 4.5. If $\beta = 0$ we are done immediately, so assume $\beta > 0$. Suppose D, G are the minimizing arguments of the upper bound problem (4.34). Then we claim that the maximum eigenvalue (which is zero) of the matrix

$$M^*DM + \mathbf{j}(GM - M^*G) - \beta^2 D \tag{5.42}$$

is distinct. To see this first note that there is a one-to-one correspondence between the zero eigenvalues of (5.42) and the zero eigenvalues of

$$M_D^*M_D + \mathbf{j}(\hat{G}M_D - M_D^*\hat{G}) - \beta^2 I_n \tag{5.43}$$

where $M_D = D^{\frac{1}{2}}MD^{-\frac{1}{2}}$ and $\hat{G} = D^{-\frac{1}{2}}GD^{-\frac{1}{2}}$ (see the substitutions in (5.16) and (5.17)). Since $\beta > 0$ we are done by the eigenvalue shift property if the matrix

$$M_D^*M_D + \mathbf{j}(\hat{G}M_D - M_D^*\hat{G})$$

cannot have a non-zero repeated eigenvalue. But now M rank one implies that M_D is rank one and so we have vectors $x, y \in \mathbb{C}^{n\times n}$ such that $M_D = xy^*$. Substituting for this we obtain

$$M_D^*M_D + \mathbf{j}(\hat{G}M_D - M_D^*\hat{G}) = yx^*xy^* + \mathbf{j}(Gxy^* - yx^*G).$$

Let $\gamma = x^*x$ and $v = Gx$ and we obtain

$$\begin{aligned} yx^*xy^* + \mathbf{j}(Gxy^* - yx^*G) &= \gamma yy^* + \mathbf{j}vy^* - \mathbf{j}yv^* = \\ y\left(\frac{\gamma}{2}y^* - \mathbf{j}v^*\right) + \left(\frac{\gamma}{2}y + \mathbf{j}v\right)y^* &= yz^* + zy^* \end{aligned}$$

where $z = \frac{\gamma}{2}y + \mathbf{j}v$. Hence by lemma 5.7 the claim is proven. The result now follows from theorem 5.8. □

In fact theorem 5.10 extends to the general case (where the infimum may not be achieved). However the proof is fairly involved and is deferred to chapter 8 (see also [78]). This theorem (with its extension) says that for such problems μ equals its upper bound and is hence equivalent to a convex problem. This reinforces the results of [17] and offers some insight into why the problem becomes so much more difficult when we move away from the "affine parameter variation" case to the "multilinear" or "polynomial" cases (see [70]). These correspond to μ problems where M is not necessarily rank one, and hence may no longer be equal to the upper bound and so may no longer be equivalent to a convex problem (note that there exist rank two matrices for which μ does not equal its upper bound).

These results also underline why there are no practical algorithms based on "edge-type" theorems, as the results appear to be relevant only to a very special problem. Furthermore, even in the very special "affine parameter case" there are a combinatoric number of edges to check.

5.4.3 Real Matrices

As we noted earlier it is always possible to obtain an upper bound for a mixed μ problem simply by treating the real parameters as complex, and using the standard complex μ upper bound (see [55] for example). However the upper bound from theorem 4.5 is frequently much better than the complex μ upper bound because of the extra degrees of freedom we have in choosing the G scaling matrix (note that if we restrict ourselves to $G = 0_n$ we recover the complex μ upper bound). The G scaling

matrix exploits the phase information we have about the real parameters in order to reduce the bound. However it is not always possible to improve upon the complex μ upper bound via the G scaling matrix as is illustrated by the following results.

Theorem 5.11 *Given a matrix $M \in \mathbb{R}^{n\times n}$ and any block structure $\mathcal{K}$ define the following subsets of $\mathcal{D}_\mathcal{K}$ and $\mathcal{G}_\mathcal{K}$*

$$\mathcal{D}_{R\mathcal{K}} = \{D \in \mathcal{D}_\mathcal{K} : D \in \mathbb{R}^{n\times n}\} \tag{5.44}$$

$$\mathcal{G}_{R\mathcal{K}} = \{G \in \mathcal{G}_\mathcal{K} : \mathbf{j}G \in \mathbb{R}^{n\times n}\}. \tag{5.45}$$

Then we have that

$$\inf_{D\in\mathcal{D}_\mathcal{K},G\in\mathcal{G}_\mathcal{K}} \left[\min_{\alpha\in\mathbb{R}}\{\alpha : \Phi_\alpha(M,D,G) \le 0\}\right] = \inf_{D\in\mathcal{D}_{R\mathcal{K}},G\in\mathcal{G}_{R\mathcal{K}}} \left[\min_{\alpha\in\mathbb{R}}\{\alpha : \Phi_\alpha(M,D,G) \le 0\}\right]. \tag{5.46}$$

Proof: First define the quantities

$$\alpha_* = \inf_{D\in\mathcal{D}_\mathcal{K},G\in\mathcal{G}_\mathcal{K}} \left[\min_{\alpha\in\mathbb{R}}\{\alpha : \Phi_\alpha(M,D,G) \le 0\}\right] \tag{5.47}$$

$$\hat{\alpha} = \inf_{D\in\mathcal{D}_{R\mathcal{K}},G\in\mathcal{G}_{R\mathcal{K}}} \left[\min_{\alpha\in\mathbb{R}}\{\alpha : \Phi_\alpha(M,D,G) \le 0\}\right]. \tag{5.48}$$

Then clearly we have that $\alpha_* \le \hat{\alpha}$. Now suppose we have $D \in \mathcal{D}_\mathcal{K}, G \in \mathcal{G}_\mathcal{K}$ and $\alpha \in \mathbb{R}$ such that

$$M^*DM + \mathbf{j}(GM - M^*G) - \alpha D \le 0.$$

Split D and G into their real and imaginary parts as $D = D_R + \mathbf{j}D_I$, $G = G_R + \mathbf{j}G_I$ with $D_R, D_I, G_R, G_I \in \mathbb{R}^{n\times n}$. Then it is easy to show that D_R, G_R are real symmetric, and D_I, G_I are real skew symmetric. Now we have that

$$\begin{aligned} & M^*DM + \mathbf{j}(GM - M^*G) - \alpha D \le 0 \\ \longrightarrow \quad & x^*(M^*DM + \mathbf{j}(GM - M^*G) - \alpha D)x \le 0 \quad \forall x \in \mathbb{C}^n \\ \longrightarrow \quad & x^T(M^TDM + \mathbf{j}(GM - M^TG) - \alpha D)x \le 0 \quad \forall x \in \mathbb{R}^n. \end{aligned}$$

Now we note that $(M^TDM + \mathbf{j}(GM - M^TG) - \alpha D) = S + \mathbf{j}W$ where

$$\begin{aligned} S &= M^TD_RM + \mathbf{j}((\mathbf{j}G_I)M - M^T(\mathbf{j}G_I)) - \alpha D_R \\ W &= M^TD_IM + G_RM - M^TG_R - \alpha D_I. \end{aligned}$$

It is easy to check that S is real symmetric, and W is real skew symmetric, so that $x^T(S+\mathbf{j}W)x = x^TSx \quad \forall x \in \mathbb{R}^n$. Thus we have

$$\begin{aligned} & x^T(M^TDM + \mathbf{j}(GM - M^TG) - \alpha D)x \leq 0 \qquad \forall x \in \mathbb{R}^n \\ \longrightarrow \quad & x^T(M^TD_RM + \mathbf{j}((\mathbf{j}G_I)M - M^T(\mathbf{j}G_I)) - \alpha D_R)x \leq 0 \quad \forall x \in \mathbb{R}^n \\ \longrightarrow \quad & (M^*D_RM + \mathbf{j}((\mathbf{j}G_I)M - M^*(\mathbf{j}G_I)) - \alpha D_R) \leq 0. \end{aligned}$$

Similarly we can show that $D > 0 \longrightarrow D_R > 0$ and so $D_R \in \mathcal{D}_{R\mathcal{K}}, (\mathbf{j}G_I) \in \mathcal{G}_{R\mathcal{K}}$ which gives $\hat{\alpha} \leq \alpha_*$ and hence $\alpha_* = \hat{\alpha}$. □

Basically theorem 5.11 says that when computing the upper bound for real matrices we may restrict our attention to purely real $D \in \mathcal{D}_{\mathcal{K}}$ (i.e., $D \in \mathcal{D}_{\mathcal{K}}$ is real symmetric) and purely imaginary $G \in \mathcal{G}_{\mathcal{K}}$ (i.e., $G \in \mathcal{G}_{\mathcal{K}}$ is of the form $G = \mathbf{j}\hat{G}$ where $\hat{G}$ is real skew symmetric). As a consequence of this we immediately obtain the following theorem.

Theorem 5.12 *Suppose we have a real matrix $M \in \mathbb{R}^{n\times n}$ and a block structure $\mathcal{K}$ with $k_i = 1$ for $i = 1, \ldots, m_r$ (i.e., none of the real scalars are repeated), then*

$$\inf_{D\in\mathcal{D}_{\mathcal{K}}, G\in\mathcal{G}_{\mathcal{K}}} \left[\min_{\alpha\in\mathbb{R}}\{\alpha : \Phi_\alpha(M, D, G) \leq 0\}\right] = \inf_{D\in\mathcal{D}_{R\mathcal{K}}} \left[\min_{\alpha\in\mathbb{R}}\{\alpha : \Phi_\alpha(M, D, 0_n) \leq 0\}\right] \tag{5.49}$$

where $\mathcal{D}_{R\mathcal{K}}$ is defined as in theorem 5.11.

Proof: Apply theorem 5.11 to conclude we may restrict our attention in the left-hand side of (5.49) to $D \in \mathcal{D}_{R\mathcal{K}}, G \in \mathcal{G}_{R\mathcal{K}}$. Now note that for this block structure (none of the real scalars are repeated) then G is diagonal (and Hermitian) and hence pure real. Thus we have $\mathcal{G}_{R\mathcal{K}} = \{0_n\}$. □

Note that theorem 5.12 says that for μ problems involving real matrices where none of the real scalars are repeated then the choice $G = 0_n$ in the upper bound is optimal, or in other words the mixed μ upper bound equals the complex μ upper bound. This is an important class of problems. For instance one encounters μ problems where M is real when it is constructed from State Space 'A, B, C, D' matrices. Note that

theorem 5.12 does not apply if any of the real parameters are repeated, and in fact it is easy to construct examples involving real matrices with repeated real parameters where $G = 0_n$ is not optimal.

It is interesting to note that by further restricting this class to purely real μ problems we can obtain a "vertex result."

Lemma 5.8 ([25]) *Suppose we have a real matrix $M \in \mathbb{R}^{n\times n}$ and a block structure $\mathcal{K}$ with $m_r = n$ and $m_c = m_C = 0$. Further suppose that $k_i = 1$ for $i = 1, \ldots, m_r$ (i. e., none of the real scalars are repeated), then it suffices to consider perturbations at the vertices of the allowed perturbation set.*

Proof: Follows immediately from the fact that $\det(I_n - \Delta M)$ for $\Delta \in X_{\mathcal{K}}$ is a real-valued multilinear function of the δ_i^r's. □

The vertices of a problem set are those points where every parameter is at an extremal value. For these problems then we can compute μ exactly by checking a finite number of points. Note however that the required computation grows exponentially with problem size, so that this result is only applicable to small problems. This should not surprise us of course, since we know from section 3.3 that even this restricted class of the mixed μ problem is NP hard.

5.5 A Characterization of μ Values

We conclude this chapter by considering an interesting connection between the mixed μ upper and lower bounds, namely the μ values of a matrix M (with respect to some block structure $\mathcal{K}$). Roughly speaking the μ values are those values of β in $QMx = \beta x$ (for $Q \in \mathcal{Q}_{\mathcal{K}}$) which correspond to local maxima of the real eigenvalues of QM. Hence they are lower bounds for $\mu_{\mathcal{K}}(M)$, with the largest of them in fact being equal to $\mu_{\mathcal{K}}(M)$. Note that these values are associated with the existence of a decomposition as in (4.16). In this section we show that these values are also associated with stationary points of eigenvalues of the upper bound function. In order

to state this more precisely we need some additional definitions. Suppose we have $\beta > 0$ and matrices $D_0 \in \tilde{\mathcal{D}}_{\mathcal{K}}, G_0 \in \mathcal{G}_{\mathcal{K}}$ with $D_0 \geq 0$ such that for some k

$$\lambda_k(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) = 0 \tag{5.50}$$

with (say) r eigenvalues coalesced at zero. So now define $U_0 \in \mathbb{C}^{n \times r}$ as any matrix satisfying

$$(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0) U_0 = 0 \tag{5.51}$$

and normalized such that $U_0^* U_0 = I_r$, and as before define $V_0 = M U_0$. Then we define the block components of U_0, V_0 by (5.6) as before, and for any $\eta \in \mathbb{C}^r, |\eta| = 1$ and $q = (q_1, \ldots, q_{m_r})$ with $q_i \in [-1\ 1]$ we define the block diagonal matrix $P^{\eta,q} \in \mathcal{Z}_{\mathcal{K}}$ by (5.32) as before. Now we define the set $\hat{\nabla}_{\mathcal{Y},k}$ as the set of all such matrices, i.e.,

$$\begin{aligned} \hat{\nabla}_{\mathcal{Y},k} \doteq \{P^{\eta,q} \in \mathcal{Z}_{\mathcal{K}} : P_i^{\eta,q}, p_i^{\eta,q}, \hat{P}_i^{\eta,q} \text{ as in } (5.32), \eta \in \mathbb{C}^r, |\eta| = 1, \\ q = (q_1, \ldots, q_{m_r}), q_i \in [-1\ 1]\}. \end{aligned} \tag{5.52}$$

The set $\hat{\nabla}_{\mathcal{Y},k}$ plays the same role as $\hat{\nabla}_{\mathcal{Y}}$ defined earlier except that now we are requiring that the k^{th} eigenvalue of $(M^* D_0 M + \mathbf{j}(G_0 M - M^* G_0) - \beta^2 D_0)$ is zero, and not necessarily the largest eigenvalue. The precise characterization of the μ values of a matrix M is stated in the following two theorems:

Theorem 5.13 *Suppose we have matrices $M \in \mathbb{C}^{n \times n}$, $Q \in \mathcal{Q}_{\mathcal{K}}$, $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ with $D^2 \in \hat{\mathcal{D}}_{\mathcal{K}}$ and $\theta_i = \pm\frac{\pi}{4}$ for $i \in \hat{\mathcal{J}}(Q)$ satisfying a (lower bound) decomposition as in theorem 4.3, i.e., we have a non-zero vector $x \in \mathbb{C}^n$ and a positive real scalar β such that*

$$\begin{aligned} QDMD^{-1}(Dx) &= \beta Dx \\ (x^* D^*) Q D^* M (D^*)^{-1} &= \beta x^* D^*. \end{aligned}$$

Then there exist matrices $\hat{D} \in \tilde{\mathcal{D}}_{\mathcal{K}}, \hat{G} \in \mathcal{G}_{\mathcal{K}}$ with $\hat{D} \geq 0$ such that for some k

$$\lambda_k(M^* \hat{D} M + \mathbf{j}(\hat{G} M - M^* \hat{G}) - \beta^2 \hat{D}) = 0$$

and $0 \in \hat{\nabla}_{\mathcal{Y},k}$ *(with* $\mathcal{Y} = (M, \hat{D}, \hat{G}, \beta)$*). Furthermore if* $|\theta_i| < \frac{\pi}{4}$ *for all* $i = 1, \ldots, m_r$ *then* $\hat{D} > 0$.

Proof: First we split the matrix D as $\tilde{D} \doteq D^2 = D_R + \mathbf{j}D_I$ where

$$\begin{aligned} D_R \doteq \text{ block diag } (&cos\tilde{\theta}_1\tilde{D}_1, \ldots, cos\tilde{\theta}_{m_r}\tilde{D}_{m_r}, \tilde{D}_{m_r+1}, \ldots, \tilde{D}_{m_r+m_c}, \\ &\tilde{d}_1 I_{k_{m_r+m_c+1}}, \ldots, \tilde{d}_{m_C} I_{k_m}) \end{aligned}$$

$$D_I \doteq \text{block diag } (sin\tilde{\theta}_1\tilde{D}_1, \ldots, sin\tilde{\theta}_{m_r}\tilde{D}_{m_r}, 0_{k_{m_r+1}}, \ldots, 0_{k_m}).$$

Note that $D_I = D_I^*$, and $D_R = D_R^*$ with $D_R \geq 0$. Now define $\hat{D} \doteq D_R$ and $\hat{G} \doteq -\beta D_I Q$, so that $\hat{D} \in \tilde{\mathcal{D}}_{\mathcal{K}}$ with $\hat{D} \geq 0$ and $\hat{G} \in \mathcal{G}_{\mathcal{K}}$. Note also that if $|\theta_i| < \frac{\pi}{4}$ for all $i = 1, \ldots, m_r$ then $\hat{D} > 0$. From the lower bound decomposition we immediately obtain

$$M^*Q^*D^2x = \beta D^2 x.$$

Thus substituting for $x = \frac{1}{\beta}QMx$ and $D^2 = D_R + \mathbf{j}D_I$ we obtain

$$(M^*Q^*QD_RM + \mathbf{j}(M^*Q^*QD_IM - \beta^2 D_I) - \beta^2 D_R)x = 0.$$

Note that $Q^*QD_R = D_R = \hat{D}$ and furthermore

$$\begin{aligned} (M^*Q^*QD_IM - \beta^2 D_I)x &= (\beta M^*Q^*D_I - \beta D_I QM)x \\ &= (\hat{G}M - M^*\hat{G})x \end{aligned}$$

so that we have

$$(M^*\hat{D}M + \mathbf{j}(\hat{G}M - M^*\hat{G}) - \beta^2\hat{D})x = 0 \tag{5.53}$$

and hence $\lambda_k(M^*\hat{D}M + \mathbf{j}(\hat{G}M - M^*\hat{G}) - \beta^2\hat{D}) = 0$ for some k. Now equation (5.53), together with $QMx = \beta x$, implies that $0 \in \hat{\nabla}_{\mathcal{Y},k}$ (see proof of theorem 5.7). □

Theorem 5.14 *Suppose we have matrices* $M \in \mathbb{C}^{n\times n}$, $\hat{D} \in \tilde{\mathcal{D}}_{\mathcal{K}}, \hat{G} \in \mathcal{G}_{\mathcal{K}}$ *with* $\hat{D} \geq 0$ *and a positive real scalar* β *such that for some* k

$$\lambda_k(M^*\hat{D}M + \mathbf{j}(\hat{G}M - M^*\hat{G}) - \beta^2\hat{D}) = 0$$

and $0 \in \hat{\nabla}_{\mathcal{Y},k}$ *(with* $\mathcal{Y} = (M, \hat{D}, \hat{G}, \beta)$*). Then under some non-degeneracy assumptions (stated in the proof) we have matrices* $Q \in \mathcal{Q}_{\mathcal{K}}$*,* $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ *with* $D^2 \in \hat{\mathcal{D}}_{\mathcal{K}}$ *and* $\theta_i = \pm\frac{\pi}{4}$ *for* $i \in \hat{\mathcal{J}}(Q)$ *satisfying a (lower bound) decomposition as in theorem 4.3, i. e., we have a non-zero vector* $x \in \mathbb{C}^n$ *such that*

$$\begin{aligned} QDMD^{-1}(Dx) &= \beta Dx \\ (x^*D^*)QD^*M(D^*)^{-1} &= \beta x^*D^*. \end{aligned}$$

Proof: Since $0 \in \hat{\nabla}_{\mathcal{Y},k}$ we have a matrix $Q \in \mathcal{Q}_{\mathcal{K}}$ and a non zero vector $x \in \mathbb{C}^n$ such that

$$\begin{aligned} QMx &= \beta x \\ (M^*\hat{D}M + \mathbf{j}(\hat{G}M - M^*\hat{G}) - \beta^2\hat{D})x &= 0 \end{aligned} \tag{5.54}$$

(see proof of theorem 5.7). Now if we partition the vectors x and Mx compatibly with the block structure as in (4.7) then it is easy to show from the above that for $i = 1, \ldots, m_r$

$$|q_i| < 1 \quad \longrightarrow \quad |\hat{D}_i(Mx)_i| = |\hat{D}_i x_i| = 0 \tag{5.55}$$

(just note that $\overline{\sigma}(Q) \leq 1$ and use similar arguments to theorem 5.7). This implies that

$$Q^*Q\hat{D}Mx = \hat{D}Mx. \tag{5.56}$$

Now we make the first non-degeneracy assumption that $|q_i| \neq 0$ for $i = 1, \ldots, m_r$ and we define $D_R \doteq \hat{D}$ and $D_I \doteq -\frac{1}{\beta}\hat{G}Q^{-1}$. Then from (5.54) and (5.56) we obtain

$$(M^*Q^*\hat{D}QM + \mathbf{j}(\hat{G}M - M^*\hat{G}) - \beta^2\hat{D})x = 0.$$

Substituting for $QMx = \beta x$ yields upon rearrangement

$$M^*Q^*(D_R + \mathbf{j}D_I)x = \beta(D_R + \mathbf{j}D_I)x. \tag{5.57}$$

Now we consider the block components (see (4.7)) of the vector $(D_R + \mathbf{j}D_I)x$. For $i = 1, \ldots, m_r$ define the vectors

$$\gamma_i \doteq (D_{Ri} + \mathbf{j}D_{Ii})x_i.$$

Then we have that

$$\gamma_i^* x_i = x_i^* D_{Ri} x_i + \mathbf{j} x_i^* D_{Ii} x_i.$$

Since by construction $D_R = D_R^*$, $D_I = D_I^*$ and $D_R \geq 0$ this implies

$$\gamma_i^* x_i = r_i e^{\mathbf{j}\psi_i} \tag{5.58}$$

for real scalars $r_i \geq 0$ and $\psi_i \in [-\frac{\pi}{2}\ \frac{\pi}{2}]$. Now we make the second non-degeneracy assumption that for $i = 1, \ldots, m_r$ we have $r_i \neq 0$. Then the relations (5.58) imply the existence of matrices $\tilde{D}_i \in \mathbb{C}^{k_i \times k_i}$ with $\tilde{D}_i = \tilde{D}_i^*$ and $\tilde{D}_i > 0$ such that (see lemma 4.3)

$$\gamma_i = (D_{Ri} + \mathbf{j} D_{Ii}) x_i = e^{-\mathbf{j}\psi_i} \tilde{D}_i x_i \tag{5.59}$$

and furthermore from (5.55) it can be shown that $\psi_i = \pm\frac{\pi}{2}$ for $i \in \hat{\mathcal{J}}(Q)$. Now define the matrix $\tilde{D}$ as

$$\begin{aligned} \tilde{D} \ \doteq\ \text{block diag}\,(e^{-\mathbf{j}\psi_1}\tilde{D}_1, \ldots, e^{-\mathbf{j}\psi_{m_r}}\tilde{D}_{m_r}, \hat{D}_{m_r+1}, \ldots, \hat{D}_{m_r+m_c}, \\ \hat{d}_1 I_{k_{m_r+m_c+1}}, \ldots, \hat{d}_{m_C} I_{k_m}). \end{aligned}$$

Now we make the final non-degeneracy assumption that $\tilde{D}$ is non-singular (which amounts to assuming that the appropriate block components of $\hat{D}$ are positive definite rather than just positive semidefinite). Then we have that $\tilde{D} \in \hat{\mathcal{D}}_{\mathcal{K}}$ and by construction

$$(D_R + \mathbf{j} D_I) x = \tilde{D} x. \tag{5.60}$$

So that substituting (5.60) into (5.57) we obtain

$$M^* Q^* \tilde{D} x = \beta \tilde{D} x. \tag{5.61}$$

Finally we define D as the unique matrix $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ such that $D^2 = \tilde{D} \in \hat{\mathcal{D}}_{\mathcal{K}}$. (Note that we have $\theta_i = \pm\frac{\pi}{4}$ for $i \in \hat{\mathcal{J}}(Q)$.) Then by construction this satisfies

$$\begin{aligned} QMx &= \beta x \\ M^* Q^* D^2 x &= \beta D^2 x. \end{aligned}$$

Simple manipulations of these equations yield the required (lower bound) decomposition. □

These theorems provide us with a direct theoretical link between the upper and lower bounds for μ. The proofs provide the formulae to construct $\hat{D}, \hat{G}$ for the upper bound function from Q, D for the lower bound decomposition, and vice versa. Thus given a stationary point for one function we can find a stationary point for the other, and hence we can generate guesses for the optimal scaling matrices for one bound from the optimal scaling matrices for the other. This connection was established for the complex case in [59], and is closely related to the "(Major) Principal Direction Alignment" ideas (also for the complex case) in [18,38]. It is important to note here that whilst the matrix $(M^*\hat{D}M + \mathbf{j}(\hat{G}M - M^*\hat{G}) - \beta^2\hat{D})$ has a zero eigenvalue, it is not necessarily the largest eigenvalue, and hence β is not necessarily an upper bound for $\mu_{\mathcal{K}}(M)$. In fact each such β is a lower bound for μ, with the largest of them equal to μ.

Chapter 6

Practical Computation of the Bounds

In this chapter we examine the computational aspects of the upper and lower bounds for the mixed μ problem from chapter 4 (see also [29,81]). Important issues to be considered here are the efficient computation of the bounds and the degree to which they approximate μ. In chapter 7 we will also consider techniques for refining the bounds for a better approximation (at an additional computational cost).

Here we develop a practical algorithm to compute the upper and lower bounds. This has been implemented as a Matlab function (m-file) "rmu," and is currently available in a test version in conjunction with the μ-Tools toolbox [7]. The theoretical bounds described in chapter 4 require some reformulation before they can be implemented in an efficient manner, and this is described in sections 6.1 and 6.2, together with details of the algorithm construction. The bounds involve solving certain optimization problems, and it is shown that the specific structure of these problems can be exploited so as to speed up the computation considerably. Some results from our extensive numerical experience with the algorithm, regarding both the quality of the bounds and the computation time, are presented in section 6.4. These results are very encouraging, and in particular it appears that one can handle medium

size problems (less than 100 perturbations) with reasonable computational requirements. Note that this could involve optimizing several thousand parameters, so that a straightforward application of brute force optimization techniques would be unwieldy.

6.1 The Lower Bound Algorithm

First we note that the lower bound from section 4.5 takes the form of a power iteration (4.28,4.29), and each iteration of the scheme is very cheap, requiring only such operations as matrix-vector multiplications and vector inner products. This gives rise to a lower bound algorithm which is much faster than would be obtained by directly solving (4.1) via standard optimization techniques (although this maximization is carried out implicitly by the power iteration). However there is another motivation, besides speed, for this power iteration approach. As we noted earlier the maximization defining μ in lemma 3.4 is nonconvex. Of course this means that we will not be able to guarantee to find the global maximum, but in order to find as good a bound as possible we would like to avoid local methods, such as gradient search. The power iteration approach developed in chapter 4 is not based on moving towards the nearest local maximum, but rather attempting to find the *largest* point satisfying the stationarity conditions of a local maximum. In this way we attempt to find a global maximum, rather than just a local one.

In this section we will discuss some of the issues that are involved in developing a practical implementation of the power algorithm. First we note that for the purely complex case, $m_r = 0$, this algorithm reduces to that of Packard [59], and hence many of the comments made there also apply here. In particular we note that there is a potential problem with the algorithm if any of the following occur:

- $Mb_k = 0$ (or $M^* z_{k+1} = 0$) $\longrightarrow$ a_{k+1} (or w_{k+1}) is not well defined.
- $|a_{1_{k+1}}| = 0$ $\longrightarrow$ $\tilde{\alpha}_{k+1}$ and/or $\hat{\alpha}_{k+1}$ is not well defined.
- $|a^*_{2_k} w_{2_k}| = 0$ $\longrightarrow$ $z_{2_{k+1}}$ and/or $b_{2_{k+1}}$ is not well defined.

- $|a_{3_k}| = 0$ or $|w_{3_k}| = 0 \quad \longrightarrow \quad z_{3_{k+1}}$ and/or $b_{3_{k+1}}$ is not well defined.

If any of these occur then one possibility is to simply restart the algorithm from a new point (i.e., a new b_1, w_1 and $\tilde{\alpha}_2$). Of course it is possible (though not generic) that one of the above conditions will recur. For this reason the above strategy is not adopted in our code. In fact if one examines the above conditions then it can be seen that is still possible to define a sensible iteration even if some of the above occur (some terms may be arbitrary and they can simply be assigned some nominal values or left at their current values). In this way it is possible to protect the code from "divide by zero" errors when any of the above occur, and this has been implemented in the "rmu" code.

Consider the following two special cases of the complex μ problem:

- $m_r = 0, m_c = 0, m_C = 1 \quad \longrightarrow \quad \mu_{\mathcal{K}}(M) = \overline{\sigma}(M).$
- $m_r = 0, m_c = 1, m_C = 0 \quad \longrightarrow \quad \mu_{\mathcal{K}}(M) = \rho(M).$

Then, as was pointed out in [59], the power iteration for these cases reduces to iteration schemes to find $\overline{\sigma}(M)$ and $\rho(M)$ respectively (i.e., $\mu_{\mathcal{K}}(M)$ in both cases). For these cases then we have that the iteration is guaranteed to converge and find $\mu_{\mathcal{K}}(M)$. Unfortunately this is not the case in general, and in fact we do not even have guaranteed convergence. The convergence properties and performance of the algorithm will be discussed in section 6.4, and in chapter 7 we will consider some more sophisticated schemes which can in principle guarantee convergence, but for now we consider what can be done if the algorithm does indeed fail to converge. Suppose we have a matrix $M \in \mathbb{C}^{n \times n}$ partitioned as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{6.1}$$

with $M_{11} \in \mathbb{C}^{n_1 \times n_1}$, $M_{22} \in \mathbb{C}^{n_2 \times n_2}$ and $n_1 + n_2 = n$. Suppose we have block structures $\mathcal{X}_{\mathcal{K}_1}$ and $\mathcal{X}_{\mathcal{K}_2}$ compatible with M_{11} and M_{22} respectively, then the block structure

$\mathcal{X}_{\hat{\mathcal{K}}}$ defined as

$$\mathcal{X}_{\hat{\mathcal{K}}} = \{\Delta = \text{block diag}\,(\Delta_1, \Delta_2) : \Delta_1 \in \mathcal{X}_{\mathcal{K}_1}, \Delta_2 \in \mathcal{X}_{\mathcal{K}_2}\} \tag{6.2}$$

is compatible with M. Further assume we have arranged the problem so that $\mathcal{X}_{\mathcal{K}_1}$ consists of *purely* real uncertainties ($\mathcal{X}_{\mathcal{K}_1} \subset \mathbb{R}^{n_1 \times n_1}$) and $\mathcal{X}_{\mathcal{K}_2}$ consists of *purely* complex uncertainties ($\mathcal{X}_{\mathcal{K}_2} \subset \mathbb{C}^{n_2 \times n_2}$). Then we have the following lemma.

Lemma 6.1 *Suppose we have a matrix $M \in \mathbb{C}^{n \times n}$ as above, a perturbation $\Delta \in \mathcal{X}_{\hat{\mathcal{K}}}, \overline{\sigma}(\Delta) \leq 1$ as above, and a real scalar $\alpha > 0$. If $\det\,(I_{n_1} - \frac{M_{11}\Delta_1}{\alpha}) = 0$ define $\beta = \alpha$, else define β as*

$$\beta = \rho\left(\Delta_2\left(M_{22} + M_{21}\frac{\Delta_1}{\alpha}(I_{n_1} - \frac{M_{11}\Delta_1}{\alpha})^{-1}M_{12}\right)\right). \tag{6.3}$$

Then $\min(\alpha, \beta) \leq \mu_{\mathcal{K}}(M)$.

Proof: If we had $\det\,(I_{n_1} - \frac{M_{11}\Delta_1}{\alpha}) = 0$ then $\hat{\Delta} = \text{block diag}\,(\frac{\Delta_1}{\alpha}, 0_{n_2})$ sets $\det\,(I_n - \hat{\Delta}M) = 0$ so that $\min(\alpha, \beta) = \alpha \leq \mu_{\mathcal{K}}(M)$. Otherwise define λ as the maximum eigenvalue from (6.3) so that $\lambda = \beta e^{\mathrm{j}\theta}$ for some $\theta \in [0\ 2\pi]$. Then if $\beta = 0$ clearly $\min(\alpha, \beta) \leq \mu_{\mathcal{K}}(M)$, and otherwise $\hat{\Delta} = \text{block diag}\,(\frac{\Delta_1}{\alpha}, e^{-\mathrm{j}\theta}\frac{\Delta_2}{\beta})$ sets $\det\,(I_n - \hat{\Delta}M) = 0$ so that $\min(\alpha, \beta) \leq \mu_{\mathcal{K}}(M)$. □

This lemma gives us a means to compute a lower bound for μ given candidate guesses for the perturbation and the lower bound, provided we are not in the pure real case, $m_c = m_C = 0$ (whence our bound is α if $\det\,(I_{n_1} - \frac{M_{11}\Delta_1}{\alpha}) = 0$ and zero otherwise). This case is discussed in more detail in section 6.4. It turns out that one can *always* obtain candidate guesses for the perturbation and the lower bound from the current values of the b, a, z, w vectors even if the scheme has not converged. Thus one can implement the power algorithm so that it *always* returns a lower bound for μ, regardless of convergence, and the "rmu" code is implemented in this fashion. Of course it is still desirable that the power iteration converges since in that case one has more faith that the lower bound obtained is a good one (i. e., close to μ).

Finally we consider the problem of computing initial guesses to start the iteration. One possibility would be to use the results from chapter 5, where the concept of "μ values" is proven for the mixed case (see also [80]). One could compute the mixed μ upper bound and then use one of the maximizing eigenvectors to construct the initial guess for the power iteration (since this procedure produces exactly the right initial guess when the maximizing eigenvector is distinct). We will see however that the mixed μ upper bound is usually more expensive to compute than the lower bound, and so it is certainly too expensive to use as a scheme for computing initial guesses for the lower bound, unless one wishes to compute both bounds anyway. However it turns out that if one wishes to compute both bounds it is advantageous to compute the lower bound first (see [83] and section 6.2), and so we do not use this approach to compute the initial guesses for the lower bound. Note however that for numerical reasons it is desirable to "balance" the matrix before starting the lower bound power iteration, and this is the first step in Packard's scheme for computing the initial guesses b_1 and w_1 for the complex μ power algorithm [59]. It would also seem to be a cheap way to generate reasonable first guesses here. The basic outline of the scheme is:

Algorithm 6.1 (Lower Bound Initial Guess)

1. *Compute a* $D \in \mathcal{D}_{\mathcal{K}}$ *which approximately solves* $\inf_{D \in \mathcal{D}_{\mathcal{K}}} \overline{\sigma}(DMD^{-1})$. *This can be done using a generalization of Osborne's method (see [83]).*
2. *Compute* $\tilde{M} = DMD^{-1}$. *The matrix* $\tilde{M}$ *is now "balanced."*
3. *Compute* b_1 *as a right singular vector of* $\tilde{M}$ *associated with* $\overline{\sigma}(\tilde{M})$, *and set* $w_1 = b_1$.

Note that mixed μ is invariant to the transformation in step 2, (i.e., $\mu_{\mathcal{K}}(M) = \mu_{\mathcal{K}}(\tilde{M})$) so that this step is valid. Furthermore the transformation in step 2 is highly recommended since it not only aids in the initial guess computation, but numerically preconditions the matrix (see [53]), and can greatly improve the power iteration

performance. Having obtained first guesses for these vectors the power iteration is started on $\tilde{M}$ using b_1 and w_1, with the first guess for $\tilde{\alpha}_2$ simply chosen as $|\tilde{\alpha}_2| = \frac{|b_{1_1}|}{|a_{1_2}|}$ with $sgn(\tilde{\alpha}_2)$ chosen so as to minimize $|b_{1_1} - \tilde{\alpha}_2 a_{1_2}|$. Finally we note that the transformation to $\tilde{M}$ is a precursor to the upper bound computation as well (see section 6.2, and also [83]), so that this initial guess computation is particularly cheap when one wants to compute both upper and lower bounds (as is usually the case).

6.2 The Upper Bound Algorithm

Since the upper bound is a convex problem there are a whole array of numerical techniques one could use to tackle this minimization. Note however that for even medium size problems ($n < 100$) then depending on the block structure $\mathcal{K}$, the optimization over the D and G scaling matrices could involve optimizing several thousand parameters. Therefore, in order to tackle such problems with reasonable computation times, a straightforward application of brute force optimization techniques will not suffice. Instead we will exploit the specific structure of this problem, so as to develop an efficient algorithm, which can handle problems of this size.

The algorithm implementation relies heavily on the fact that the upper bound may be reformulated several different ways, as stated in the following theorem.

Theorem 6.1 *Suppose we have a matrix $M \in \mathbb{C}^{n\times n}$ and a real scalar $\beta > 0$. Then the following statements are equivalent:*

I. There exist matrices $D_I \in \mathcal{D}_\mathcal{K}, G_I \in \mathcal{G}_\mathcal{K}$ such that:

$$\overline{\lambda}\left(M^* D_I M + \mathbf{j}(G_I M - M^* G_I) - \beta^2 D_I\right) \leq 0. \tag{6.4}$$

II. There exist matrices $D_{II} \in \overline{\mathcal{D}}_\mathcal{K}, G_{II} \in \overline{\mathcal{G}}_\mathcal{K}$ (or $D_{II} \in \mathcal{D}_\mathcal{K}, G_{II} \in \mathcal{G}_\mathcal{K}$) such that:

$$\overline{\lambda}\left(M_{D_{II}}^* M_{D_{II}} + \mathbf{j}(G_{II} M_{D_{II}} - M_{D_{II}}^* G_{II})\right) \leq \beta^2 \tag{6.5}$$

where we denote $M_D \doteq DMD^{-1}$.

III. There exist matrices $D_{III} \in \overline{\mathcal{D}}_{\mathcal{K}}, G_{III} \in \overline{\mathcal{G}}_{\mathcal{K}}$ *(or* $D_{III} \in \mathcal{D}_{\mathcal{K}}, G_{III} \in \mathcal{G}_{\mathcal{K}}$*) such that:*

$$\overline{\sigma}\left(\left(\frac{M_{D_{III}}}{\beta} - \mathbf{j}G_{III}\right)(I_n + G_{III}^2)^{-\frac{1}{2}}\right) \leq 1. \tag{6.6}$$

IV. There exist matrices $D_{IV} \in \overline{\mathcal{D}}_{\mathcal{K}}, G_{IV} \in \overline{\mathcal{G}}_{\mathcal{K}}$ *(or* $D_{IV} \in \mathcal{D}_{\mathcal{K}}, G_{IV} \in \mathcal{G}_{\mathcal{K}}$*) such that:*

$$\overline{\sigma}\left((I_n + G_{IV}^2)^{-\frac{1}{4}}\left(\frac{M_{D_{IV}}}{\beta} - \mathbf{j}G_{IV}\right)(I_n + G_{IV}^2)^{-\frac{1}{4}}\right) \leq 1. \tag{6.7}$$

Proof: First note that for any of the forms in (6.5), (6.6), (6.7), we may convert between $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ and $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}, \hat{G} \in \overline{\mathcal{G}}_{\mathcal{K}}$ in the following way. Given $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ perform an eigenvalue decomposition on $G \in \mathcal{G}_{\mathcal{K}}$ to give $G = U\Lambda U^*$ (with U unitary, Λ diagonal and real). It is then easy to check that $\hat{D} = U^*D \in \overline{\mathcal{D}}_{\mathcal{K}}$ and $\hat{G} = \Lambda \in \overline{\mathcal{G}}_{\mathcal{K}}$ work. To go the other way, take $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}, \hat{G} \in \overline{\mathcal{G}}_{\mathcal{K}}$ and perform a polar decomposition on $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}$ to give $\hat{D} = VP$ (with V unitary, P Hermitian positive definite). It is then easy to check that $D = P \in \mathcal{D}_{\mathcal{K}}$ and $G = V^*\hat{G}V \in \mathcal{G}_{\mathcal{K}}$ work. Thus it suffices to prove the equivalence between I, II, III, IV for $D_I, D_{II}, D_{III}, D_{IV} \in \mathcal{D}_{\mathcal{K}}$ and $G_I, G_{II}, G_{III}, G_{IV} \in \mathcal{G}_{\mathcal{K}}$.

We will show that I $\rightarrow$ II $\rightarrow$ III $\rightarrow$ IV $\rightarrow$ I. Suppose I is true, then we have $D_I \in \mathcal{D}_{\mathcal{K}}, G_I \in \mathcal{G}_{\mathcal{K}}$ satisfying

$$M^*D_IM + \mathbf{j}(G_IM - M^*G_I) - \beta^2 D_I \leq 0.$$

Multiply by $D_I^{-\frac{1}{2}}$ on both sides (which will not affect the definiteness of the expression) to get

$$D_I^{-\frac{1}{2}}M^*D_IMD_I^{-\frac{1}{2}} + \mathbf{j}(D_I^{-\frac{1}{2}}G_IMD_I^{-\frac{1}{2}} - D_I^{-\frac{1}{2}}M^*G_ID_I^{-\frac{1}{2}}) - \beta^2 I_n \leq 0.$$

Thus defining $D_{II} = D_I^{\frac{1}{2}} \in \mathcal{D}_{\mathcal{K}}$, $G_{II} = D_I^{-\frac{1}{2}}G_ID_I^{-\frac{1}{2}} \in \mathcal{G}_{\mathcal{K}}$ and rearranging we obtain

expression II. Now note that we have

$$\begin{array}{lrcl}
 & M_{D_{II}}^* M_{D_{II}} + \mathbf{j}(G_{II} M_{D_{II}} - M_{D_{II}}^* G_{II}) & \leq & \beta^2 I_n \\
\rightarrow & \frac{1}{\beta^2}\left(M_{D_{II}}^* M_{D_{II}} + \mathbf{j}(G_{II} M_{D_{II}} - M_{D_{II}}^* G_{II})\right) & \leq & I_n \\
\rightarrow & \left(\frac{M_{D_{II}}}{\beta} - \mathbf{j}\frac{G_{II}}{\beta}\right)^* \left(\frac{M_{D_{II}}}{\beta} - \mathbf{j}\frac{G_{II}}{\beta}\right) - \frac{G_{II}^2}{\beta^2} & \leq & I_n \\
\rightarrow & \left(\frac{M_{D_{II}}}{\beta} - \mathbf{j}\frac{G_{II}}{\beta}\right)^* \left(\frac{M_{D_{II}}}{\beta} - \mathbf{j}\frac{G_{II}}{\beta}\right) & \leq & I_n + \frac{G_{II}^2}{\beta^2} \\
\rightarrow & \left(\left(\frac{M_{D_{II}}}{\beta} - \mathbf{j}\frac{G_{II}}{\beta}\right)(I_n + \frac{G_{II}^2}{\beta^2})^{-\frac{1}{2}}\right)^* \left(\left(\frac{M_{D_{II}}}{\beta} - \mathbf{j}\frac{G_{II}}{\beta}\right)(I_n + \frac{G_{II}^2}{\beta^2})^{-\frac{1}{2}}\right) & \leq & I_n \\
\rightarrow & \overline{\sigma}\left(\left(\frac{M_{D_{II}}}{\beta} - \mathbf{j}\frac{G_{II}}{\beta}\right)(I_n + \frac{G_{II}^2}{\beta^2})^{-\frac{1}{2}}\right) \leq 1 & &
\end{array}$$

so that defining $D_{III} = D_{II} \in \mathcal{D}_{\mathcal{K}}$ and $G_{III} = \frac{G_{II}}{\beta} \in \mathcal{G}_{\mathcal{K}}$ we have expression III. Now define $\hat{D} = (I_n + G_{III}^2)^{\frac{1}{4}} D_{III} \in \overline{\mathcal{D}}_{\mathcal{K}}$ and we have

$$\overline{\sigma}\left((I_n + G_{III}^2)^{-\frac{1}{4}}\left(\frac{M_{\hat{D}}}{\beta} - \mathbf{j}G_{III}\right)(I_n + G_{III}^2)^{-\frac{1}{4}}\right) \leq 1.$$

Perform an polar decomposition on $\hat{D}$ to give $\hat{D} = VP$ (with V unitary and P Hermitian positive definite). Then it is easy to see that defining $D_{IV} = P \in \mathcal{D}_{\mathcal{K}}$ and $G_{IV} = V^* G_{III} V \in \mathcal{G}_{\mathcal{K}}$ we have expression IV. Finally we note that by essentially reversing the above steps it is straightforward to verify that given $D_{IV} \in \mathcal{D}_{\mathcal{K}}, G_{IV} \in \mathcal{G}_{\mathcal{K}}$ satisfying expression IV, the matrices $D_I \in \mathcal{D}_{\mathcal{K}}, G_I \in \mathcal{G}_{\mathcal{K}}$ given by $D_I = D_{IV}(I_n + G_{IV}^2)^{-\frac{1}{2}} D_{IV}$ and $G_I = \beta D_{IV} G_{IV} (I_n + G_{IV}^2)^{-\frac{1}{2}} D_{IV}$ satisfy expression I. □

Remarks: The equivalence between I, II, III for $D_I, D_{II}, D_{III} \in \mathcal{D}_{\mathcal{K}}$ and $G_I, G_{II}, G_{III} \in \mathcal{G}_{\mathcal{K}}$ was shown in [29]. Note from the proof that we can easily obtain the formulae to convert between the various forms (there are several more equivalent forms, slight variations on the above, which can also easily be obtained).

These different formulations, whilst mathematically equivalent, have quite different numerical properties. For the purposes of developing an upper bound algorithm, we will be concerned mostly with the formulations in (6.4) and (6.7). It follows from (6.7) that one may develop an alternative form of the mixed μ upper bound.

Theorem 6.2 *Suppose we have a matrix $M \in \mathbb{C}^{n\times n}$ and a compatible block structure $\mathcal{K}$. Then we have that*

$$\mu_{\mathcal{K}}(M) \leq \inf_{\substack{\hat{D}\in\overline{\mathcal{D}}_{\mathcal{K}} \\ \hat{G}\in\overline{\mathcal{G}}_{\mathcal{K}}}} \left[\inf_{\substack{\beta\in\mathbb{R} \\ \beta>0}} \left\{ \beta : \overline{\sigma}\left((I+\hat{G}^2)^{-\frac{1}{4}} \left(\frac{\hat{D}M\hat{D}^{-1}}{\beta} - \mathbf{j}\hat{G} \right) (I+\hat{G}^2)^{-\frac{1}{4}} \right) \leq 1 \right\} \right]. \tag{6.8}$$

Furthermore this upper bound is equivalent to the one in (4.39).

Proof: Compute β_* as the solution to the minimization problem in (6.8). Then we have sequences $\hat{D}^k \in \overline{\mathcal{D}}_{\mathcal{K}}, \hat{G}^k \in \overline{\mathcal{G}}_{\mathcal{K}}, \beta_k > 0$ such that

$$\overline{\sigma}\left((I+(\hat{G}^k)^2)^{-\frac{1}{4}} \left(\frac{\hat{D}^k M(\hat{D}^k)^{-1}}{\beta_k} - \mathbf{j}\hat{G}^k \right) (I+(\hat{G}^k)^2)^{-\frac{1}{4}} \right) \leq 1$$

and $\beta_k \downarrow \beta_*$ with $\beta_* \geq 0$. Thus by theorem 6.1 we can find sequences $D^k \in \mathcal{D}_{\mathcal{K}}, G^k \in \mathcal{G}_{\mathcal{K}}$ such that

$$M^* D^k M + \mathbf{j}(G^k M - M^* G^k) - \beta_k^2 D^k \leq 0.$$

Since $\beta_k \downarrow \beta_* \geq 0$ we have that the upper bound in (4.39) is at most β_*, and hence $\mu_{\mathcal{K}}(M) \leq \beta_*$.

Now suppose we compute $\hat{\beta}$ as the solution to the mixed μ upper bound in (4.39). Then we can find sequences $D^j \in \mathcal{D}_{\mathcal{K}}, G^j \in \mathcal{G}_{\mathcal{K}}, \beta_j \geq 0$ such that

$$M^* D^j M + \mathbf{j}(G^j M - M^* G^j) - \beta_j^2 D^j \leq 0$$

with $\beta_j \downarrow \hat{\beta}$ and $\hat{\beta} \geq 0$. Defining the sequence $\hat{\beta}_j \doteq \beta_j + \frac{1}{j}$ then we have that

$$M^* D^j M + \mathbf{j}(G^j M - M^* G^j) - \hat{\beta}_j^2 D^j \leq 0$$

with $\hat{\beta}_j \downarrow \hat{\beta} \geq 0$ and $\hat{\beta}_j > 0$. Thus by theorem 6.1 we can find sequences $\hat{D}^j \in \overline{\mathcal{D}}_{\mathcal{K}}, \hat{G}^j \in \overline{\mathcal{G}}_{\mathcal{K}}$ such that

$$\overline{\sigma}\left((I+(\hat{G}^j)^2)^{-\frac{1}{4}} \left(\frac{\hat{D}^j M(\hat{D}^j)^{-1}}{\hat{\beta}_j} - \mathbf{j}\hat{G}^j \right) (I+(\hat{G}^j)^2)^{-\frac{1}{4}} \right) \leq 1.$$

Since $\hat{\beta}_j \downarrow \hat{\beta}$ with $\hat{\beta}_j > 0$ we have that the bound from (6.8) is at most $\hat{\beta}$. □

Each of these two different formulations of the upper bound problem has its own advantages. The problem statement from (4.39) has the advantages that it is linear in the matrices D and G, and is convex (and hence one will not have problems associated with local minima). The problem statement from (6.8) has the advantages that one is trying to minimize the norm of a given matrix (which offers some numerical advantages), that $\hat{D}$ enters the problem exactly as in the standard complex μ upper bound, that $\hat{G}$ enters the problem in a balanced symmetric fashion, and that $\hat{G}$ is now a real diagonal matrix.

The upper bound algorithm implemented here uses a mixture of the formulations in theorem 6.1. Initially we tackle the problem in the form of (6.8). Here we can use some methods from the complex μ bounds, together with various other techniques, to obtain a fairly good estimates of $\hat{D}$,$\hat{G}$ and β. These are then converted into an initial guess for the problem in the form of (4.39) and the algorithm then proceeds to improve on these. More specifically the algorithm proceeds as follows:

Algorithm 6.2 (Mixed μ Upper and Lower Bounds)

1. *First we balance the matrix. This proceeds by computing the scaling matrix $\hat{D}$ to solve $\inf_{\hat{D}\in\overline{\mathcal{D}}_{\mathcal{K}}} |\hat{D}M\hat{D}^{-1}|_F$ using a generalization of Osborne's method [53] (as in the standard complex μ upper bound). The matrix $\hat{M} \doteq \hat{D}M\hat{D}^{-1}$ is then balanced, and this procedure generates our initial guess for $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}$.*

2. *The lower bound is now computed using the algorithm from section 4.5, applied to the balanced matrix $\hat{M}$.*

3. *Now we have a lower bound, and $\overline{\sigma}(\hat{M})$ serves as a first guess for the upper bound. This is then improved upon in the following way. For any fixed level of β compute each block of $\hat{G}$ as $\hat{G}_i = \frac{1}{2\mathbf{j}\beta}(\hat{M}_i - \hat{M}_i^*)$ where $\hat{M}_i$ is the corresponding sub-matrix of $\hat{M}$ (i. e., $\mathbf{j}\hat{G}_i$ cancels the Skew-Hermitian part of $\frac{\hat{M}_i}{\beta}$). Then bisect on β between the lower and current upper bound to find the smallest β such that*

$$\overline{\sigma}\left((I+\hat{G}^2)^{-\frac{1}{4}}\left(\frac{\hat{M}}{\beta}-\mathbf{j}\hat{G}\right)(I+\hat{G}^2)^{-\frac{1}{4}}\right) \leq 1.$$

Finally perform an eigenvalue decomposition on $\hat{G}$ *as* $\hat{G} = U\Lambda U^*$ *(with* U *unitary,* Λ *diagonal and real), and convert to* $\hat{G} \in \overline{\mathcal{G}}_{\mathcal{K}}$ *by redefining* $\hat{G}$ *as* Λ *and absorbing the* U *matrix into* $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}$ *and* $\hat{M}$.

4. *We now have initial guesses for* $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}$ *and* $\hat{G} \in \overline{\mathcal{G}}_{\mathcal{K}}$. *The next step is to compute a descent direction for* $\hat{G} \in \overline{\mathcal{G}}_{\mathcal{K}}$ *together with an appropriate step length. A new* $\hat{G}$ *then is computed by taking this descent step. This procedure is then repeated once more.*

5. *The matrix* $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}$ *is updated by computing a diagonal matrix* $\hat{D}_d \in \overline{\mathcal{D}}_{\mathcal{K}}$ *(so that it commutes with* $\hat{G} \in \overline{\mathcal{G}}_{\mathcal{K}}$*) which minimizes*

$$\inf_{\substack{\hat{D}_d \in \overline{\mathcal{D}}_{\mathcal{K}} \\ \hat{D}_d \text{ diagonal}}} \left| \hat{D}_d (I + G^2)^{-\frac{1}{4}} \left(\frac{\hat{M}}{\beta} - \mathbf{j} G \right) (I + G^2)^{-\frac{1}{4}} \hat{D}_d^{-1} \right|_F$$

again using a generalized Osborne's method. We then absorb $\hat{D}_d$ *into* $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}$.

6. *Step 4 is repeated.*

7. *We now have guesses for* $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}, \hat{G} \in \overline{\mathcal{G}}_{\mathcal{K}}$ *and* β *for the upper bound problem in (6.8). These are converted into* $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ *which form guesses for the upper bound problem in (4.39). We now improve these guesses using a descent algorithm, which iteratively computes a descent direction, and an appropriate step length, for both* $D \in \mathcal{D}_{\mathcal{K}}$ *and* $G \in \mathcal{G}_{\mathcal{K}}$ *simultaneously. At each step we compute a new upper bound by solving the associated eigenvalue problem, and quit when the bound stops decreasing (within tolerance).*

This algorithm has been implemented as as Matlab function (m-file) "rmu," and is currently available in a test version in conjunction with the μ-Tools toolbox [7]. The algorithm returns upper and lower bounds for $\mu_{\mathcal{K}}(M)$, together with appropriate scaling matrices $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}$, $\hat{G} \in \overline{\mathcal{G}}_{\mathcal{K}}$ for the upper bound problem in (6.8), and $Q \in \mathcal{Q}_{\mathcal{K}}$ for the lower bound problem (4.1).

The balancing in step 1 of the algorithm serves several purposes. Firstly we obtain a $\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}$ which approximately solves $\inf_{\hat{D} \in \overline{\mathcal{D}}_{\mathcal{K}}} \overline{\sigma}(\hat{D} M \hat{D}^{-1})$, or in other words

the standard upper bound for the associated complex μ problem. Since we have reformulated the problem so that the $\hat{D}$ matrix enters exactly as in the complex μ upper bound, and the $\hat{G}$ matrix enters in a balanced symmetric fashion, this $\hat{D}$ matrix also serves as a good first guess for the mixed μ upper bound. A good deal of numerical experience with the generalized Osborne's method for computing complex μ upper bounds has shown that is very fast and usually works well, and so by reformulating the problem in this fashion we can exploit these properties in the mixed problem as well. Note that the conversion from $\hat{D}$ in problem (6.8) to D in problem (4.39) involves $\hat{G}$, so that we could not *directly* use the generalized Osborne's method to provide a good guess for $D \in \mathcal{D}_{\mathcal{K}}$ in problem (4.39). This balancing also numerically preconditions the problem, and can greatly improve the performance of the subsequent steps, including the lower bound computation via a power iteration in step 2 (see the discussion in section 6.1).

Step 3 of the algorithm generates our initial guess for $\hat{G}$. The approach is somewhat intuitive, but although there are no firm guarantees, it appears in general to work quite well. Thus our $\hat{D}, \hat{G}$ estimates, which require very little computation time, are usually fairly good *before* we enter the descent portion of the algorithm, and hence we can restrict ourselves to a small number of descent steps. This is crucial in obtaining a fast implementation, since the descent steps are quite computationally expensive.

Note that in step 7 we are required to compute a descent direction for $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$, together with an appropriate step length. We compute *matrix* descent directions for D, G in one shot by computing a generalized gradient of the upper bound function (see section 5.2). In this way we avoid separate computation for the individual elements of the D, G matrices. This is important not only for speed of computation, but also because in the case of repeated eigenvalues the upper bound function may not be differentiable. In that case there may not be a descent direction with respect to any individual elements of D, G, when there is a descent direction if

all the elements are allowed to move *simultaneously.* In the case that the maximum eigenvalue is distinct, then this descent direction coincides with the usual gradient direction. The step length computation is somewhat ad-hoc, but ensures that the maximum eigenvalue of the upper bound function decreases, and that we satisfy the constraint $D > 0$. Similar comments with regard to the computation of descent directions and step lengths apply to steps 4 and 6.

This implementation of the upper bound results in an algorithm which is quite efficient, and can handle medium size problems ($n < 100$) with reasonable computational requirements. Results regarding both the quality of the bounds and their computational requirements (as a function of problem size) are presented in section 6.4.

6.3 Generating Test Matrices

It was stated in chapter 3 that the mixed μ problem is NP hard, which implies that the worst case performance of our (or any other) algorithm will be poor, either in terms of the accuracy of the bounds, or the growth rate in computation. In fact we can construct examples for which the bounds in theorems 4.1 and 4.5 are arbitrarily far apart. For engineering purposes then the real issue becomes whether or not we can develop a "practical" algorithm, whose *typical* performance is acceptable. In order to examine the typical performance in section 6.4, we will run the algorithm repeatedly on a large number of test matrices, randomly generated from within certain classes, and collect statistical data. In this section we describe three specific types of random matrices that will be used.

The most straightforward way to generate random complex matrices in Matlab is with the μ-Tools "crand" command. This generates matrices whose elements are random variables, and by setting "rand('normal')" in Matlab we can choose these elements to be normally distributed with zero mean. We will refer to this type of random matrix as a *crand* matrix.

Unfortunately it is doubtful that *crand* matrices are at all representative of those of practical interest. Since the matrices that the μ software will be run on are typically obtained from control problems, a fairly natural class of random complex matrices is to randomly generate State Space 'A,B,C,D' matrices (with A stable) using the μ-Tools "sysrand" command, and then evaluate the transfer matrix at some frequency (usually placed roughly in the middle of the modes). We will refer to this type of random matrix as a *sysrand* matrix.

For the purposes of testing algorithms it is desirable to be able to generate problems for which we know the answer a-priori. The following algorithm provides us with the means to generate such problems:

Algorithm 6.3 (Generate Matrices with $\mu = 1$)

1. *Randomly generate matrices $D \in \overline{\mathcal{D}}_{\mathcal{K}}, G \in \overline{\mathcal{G}}_{\mathcal{K}}$ and $Q \in \mathcal{U}_{\mathcal{K}}$. In addition randomly generate a unitary matrix $Y \in \mathbb{C}^{n\times n}$, and a real nonnegative diagonal matrix $\Sigma = diag(\sigma_1 \ldots \sigma_n)$ with*

$$\begin{aligned} \sigma_i &= 1 \quad for \quad i = 1, \ldots, r \\ \sigma_i &< 1 \quad for \quad i = r+1, \ldots, n \end{aligned} \tag{6.9}$$

   where r is some integer satisfying $1 \leq r \leq n$. Finally generate a random unit norm vector $\eta \in \mathbb{C}^n$ with the restriction that:

$$\eta_i = 0 \quad for \quad i = r+1, \ldots, n. \tag{6.10}$$

2. *Compute $X \in \mathbb{C}^{n\times n}$ as any unitary matrix which satisfies the equation*

$$X\eta = (Q^{-1} - \mathbf{j}G)(I_n + G^2)^{-\frac{1}{2}} Y\eta. \tag{6.11}$$

   It is easy to check that the matrix $(Q^{-1} - \mathbf{j}G)(I_n + G^2)^{-\frac{1}{2}}Y$ is unitary, so that this is always possible.

3. *Compute $M \in \mathbb{C}^{n\times n}$ as*

$$M = D^{-1}\left((I_n + G^2)^{\frac{1}{4}} X\Sigma Y^* (I_n + G^2)^{\frac{1}{4}} + \mathbf{j}G\right) D. \tag{6.12}$$

Before proving that this algorithm does indeed generate problems for which we know the answer a-priori, we need a preliminary result.

Theorem 6.3 *Suppose we have $M \in \mathbb{C}^{n\times n}$ and a compatible block structure $\mathcal{K}$. Then if the infimization in the μ upper bound (4.39) is achieved, and equals $\mu_{\mathcal{K}}(M)$, we have*

$$\max_{U\in\mathcal{U}_{\mathcal{K}}} \rho_R(UM) = \mu_{\mathcal{K}}(M). \tag{6.13}$$

Proof: Recall from earlier (4.1), which says that

$$\max_{Q\in\mathcal{Q}_{\mathcal{K}}} \rho_R(QM) = \mu_{\mathcal{K}}(M)$$

so that immediately we have

$$\max_{U\in\mathcal{U}_{\mathcal{K}}} \rho_R(UM) \leq \mu_{\mathcal{K}}(M).$$

Hence the result is trivial for $\mu_{\mathcal{K}}(M) = 0$ so assume $\mu_{\mathcal{K}}(M) > 0$. Then by a simple scaling argument we may without loss of generality assume $\mu_{\mathcal{K}}(M) = 1$. Suppose we have the perturbation $Q \in \mathcal{Q}_{\mathcal{K}}$ achieving (4.1), or in other words $Q \in \mathcal{Q}_{\mathcal{K}}$ and $x \in \mathbb{C}^n$ such that

$$QMx = x$$

with $x \neq 0$. This implies that the block components of the vectors x and Mx satisfy

$$\begin{aligned} q_i^r(Mx)_{r_i} &= x_{r_i} \quad \text{for} \quad i = 1,\ldots,m_r \\ q_i^c(Mx)_{c_i} &= x_{c_i} \quad \text{for} \quad i = 1,\ldots,m_c \\ Q_i^C(Mx)_{C_i} &= x_{C_i} \quad \text{for} \quad i = 1,\ldots,m_C. \end{aligned} \tag{6.14}$$

Now by assumption we have $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ such that

$$(M^*DM + \mathbf{j}(GM - M^*G) - D) \leq 0$$

so that in particular

$$x^*(M^*DM + \mathbf{j}(GM - M^*G) - D)x \leq 0.$$

Expanding this expression out, and substituting for (6.14), one can derive that

$$\sum_{i=1}^{m_r} |D_i^{\frac{1}{2}}(Mx)_{r_i}|^2 \leq \sum_{i=1}^{m_r} |q_i^r|^2 |D_i^{\frac{1}{2}}(Mx)_{r_i}|^2 \tag{6.15}$$

(note that all the complex blocks q_i^c, Q_i^C are unitary). Since we have $|q_i^r| \leq 1$ for all $i = 1, \ldots, m_r$ this implies that

$$|q_i^r| < 1 \quad \longrightarrow \quad |D_i^{\frac{1}{2}}(Mx)_{r_i}| = 0.$$

But in this case, since $D_i > 0$, this implies $(Mx)_{r_i} = 0$, and hence by (6.14), $x_{r_i} = 0$. Thus we may just as well take $q_i^r = 1$ for all such blocks and still satisfy $QMx = x$ but now with $Q \in \mathcal{U}_{\mathcal{K}}$. $\square$

This theorem says that for problems where μ equals its upper bound, then if the infimum in the upper bound (4.39) is achieved, the worst case perturbation may be taken to be on a vertex. In general this is not the case. Note that by taking the contrapositive of this statement we find that: *for problems where μ equals its upper bound, then if the worst case perturbation must include internal reals, the infimum in the upper bound is not achieved.* Thus we see that the issue of whether or not the upper bound is achieved, is strongly associated with the presence of internal real parameters in the worst case perturbation. In the complex case this issue does not arise, and the only thing that matters in this regard is whether or not the matrix is reducible (see [28,1]). We will return to this issue in chapter 8, but for now we use this result to prove that algorithm 6.3 does as we claimed.

Theorem 6.4 *Suppose we have a matrix $M \in \mathbb{C}^{n \times n}$ and a block structure $\mathcal{K}$. Then denoting the upper bound from theorem 6.2 by $\hat{\mu}$, we have that the following two conditions are equivalent:*

1. *The infimization in theorem 6.2 is achieved and $\hat{\mu}_{\mathcal{K}}(M) = \mu_{\mathcal{K}}(M) = 1$.*
2. *M can be generated by algorithm 6.3.*

Proof: (**2**$\longrightarrow$**1**) From (6.12) we immediately have that

$$(I_n + G^2)^{-\frac{1}{4}}(DMD^{-1} - \mathbf{j}G)(I_n + G^2)^{-\frac{1}{4}} = X\Sigma Y^*$$

and hence

$$\overline{\sigma}\left((I_n+G^2)^{-\frac{1}{4}}(DMD^{-1}-\mathbf{j}G)(I_n+G^2)^{-\frac{1}{4}}\right)=1$$

so that theorem 6.2 implies that $\hat{\mu}_{\mathcal{K}}(M)\leq 1$, and we have $D\in\overline{\mathcal{D}}_{\mathcal{K}}, G\in\overline{\mathcal{G}}_{\mathcal{K}}$ achieving this bound. Now by construction $\Sigma Y^*Y\eta=\eta$, so that substituting into (6.11) we obtain

$$X\Sigma Y^*Y\eta=(Q^{-1}-\mathbf{j}G)(I_n+G^2)^{-\frac{1}{2}}Y\eta.$$

Rearranging this equation we can obtain (note that any $D\in\overline{\mathcal{D}}_{\mathcal{K}}$ and $Q\in\mathcal{U}_{\mathcal{K}}$ commute)

$$QD^{-1}\left((I_n+G^2)^{\frac{1}{4}}X\Sigma Y^*(I_n+G^2)^{\frac{1}{4}}+\mathbf{j}G\right)DD^{-1}(I_n+G^2)^{-\frac{1}{4}}Y\eta=D^{-1}(I_n+G^2)^{-\frac{1}{4}}Y\eta.$$

Recognizing M from (6.12) and defining $x=D^{-1}(I_n+G^2)^{-\frac{1}{4}}Y\eta$ we have

$$QMx=x$$

with $x\neq 0$, and hence $\mu_{\mathcal{K}}(M)\geq 1$. Together with our earlier result this implies statement 1.

(1⟶2) Statement 1, together with theorems 4.1 and 6.2, implies that we have $D\in\overline{\mathcal{D}}_{\mathcal{K}}, G\in\overline{\mathcal{G}}_{\mathcal{K}}$ and $Q\in\mathcal{Q}_{\mathcal{K}}$ such that

$$\overline{\sigma}\left((I_n+G^2)^{-\frac{1}{4}}(DMD^{-1}-\mathbf{j}G)(I_n+G^2)^{-\frac{1}{4}}\right)\leq 1 \tag{6.16}$$

$$QMx=x \tag{6.17}$$

with $x\neq 0$. By theorems 6.1 and 6.3 we may assume $Q\in\mathcal{U}_{\mathcal{K}}$ without loss of generality. Furthermore by continuity of singular values, and theorem 6.2, we must in fact have equality in (6.16). Perform a singular value decomposition on (6.16) to obtain

$$(I_n+G^2)^{-\frac{1}{4}}(DMD^{-1}-\mathbf{j}G)(I_n+G^2)^{-\frac{1}{4}}=X\Sigma Y^*. \tag{6.18}$$

Now note that our algorithm could (randomly) choose these $D\in\overline{\mathcal{D}}_{\mathcal{K}}$, $G\in\overline{\mathcal{G}}_{\mathcal{K}}$, $Q\in\mathcal{U}_{\mathcal{K}}$, Y, and Σ in step 1. Then from (6.18) we see that M satisfies (6.12) and

it remains to show that our algorithm could choose this X. Substituting for M in (6.17) we can obtain

$$X\Sigma Y^*(I_n + G^2)^{\frac{1}{4}}Dx = (Q^{-1} - \mathbf{j}G)(I_n + G^2)^{-\frac{1}{2}}(I_n + G^2)^{\frac{1}{4}}Dx.$$

Now define $y = (I_n + G^2)^{\frac{1}{4}}Dx$, so that $y \neq 0$ and we may define the unit norm vector $\hat{y} = \frac{y}{|y|}$ satisfying

$$X\Sigma Y^*\hat{y} = (Q^{-1} - \mathbf{j}G)(I_n + G^2)^{-\frac{1}{2}}\hat{y}. \tag{6.19}$$

Since $(Q^{-1} - \mathbf{j}G)(I_n + G^2)^{-\frac{1}{2}}$ and X are unitary we have that $|\Sigma Y^*\hat{y}| = 1$. But now Y is unitary, and so the structure of Σ implies that $\hat{y} = Y\eta$, where $|\eta| = 1$, and η is of the required form. Now suppose our algorithm (randomly) chose this η in step 1. Then we have that $\Sigma Y^*\hat{y} = \Sigma Y^*Y\eta = \eta$ and so

$$X\eta = (Q^{-1} - \mathbf{j}G)(I_n + G^2)^{-\frac{1}{2}}Y\eta$$

so that our algorithm could choose this X in step 2 and hence generate M. □

Remarks: The above algorithm was first developed for the purely complex case in [28, 59]. Note that we can control the number of singular values coalesced at the minimum of the upper bound function in theorem 6.2, and a simple extension to the algorithm allows us to also control the number of eigenvalues coalesced at the maximum of the lower bound function in theorem 4.1. Furthermore by scaling the level of $\overline{\sigma}(G)$ in the algorithm, it appears we can also control the typical level of complex μ for M, and hence generate problems with a large (or small) gap between mixed and complex μ (since we always have mixed μ equal to 1). Roughly speaking we find that the larger the level of $\overline{\sigma}(G)$, the larger the gap between mixed and complex μ. There is some theoretical justification for this phenomenon, but we will not go into the details here.

This algorithm allows us to randomly generate all problems with the upper bound achieved, and equal to μ, which is equal to one (together with optimal scaling matrices achieving the upper and lower bounds). Note that for these problems there is no gap

between the bounds from theorems 4.1 and 4.5, although the optimal lower bound requires the solution of a nonconvex maximization problem. We will refer to a random matrix generated by the above algorithm as a *nogap* matrix.

6.4 Algorithm Performance

In this section we consider the numerical performance of the algorithm outlined in the preceding sections. There are many questions one could ask with regard to the algorithm performance, both in terms of computation time and accuracy of the resulting bounds. We decided to focus on the algorithm performance versus matrix size for a fixed set of uncertainty descriptions.

Prior to examining the above quantities however, we first wished to examine the convergence properties of the lower bound power iteration, and for this we carried out a series of tests using *crand* matrices. Experimentally it appears that, whilst the analysis in section 4.5 offered no convergence guarantees, the power iteration converges most of the time. For a typical test run of 500 *crand* 5×5 matrices with 2 real scalar uncertainties and three complex scalar uncertainties the power algorithm converged 96% of the time (in an average of 22 iterations). The power algorithm has also been tested on a variety of other block structures, and on much larger matrices (e. g. , 100×100) and the convergence properties appear similar to those described above. An exception to this is the pure real case ($m_c = m_C = 0$). This appears to have significantly poorer properties than any other. There are important reasons for this that seem inherent to the problem, not the computation scheme. Fortunately the real-only case is of less engineering interest than the mixed case (see the discussion in chapter 3, and also [82] and the references therein).

Additionally we note that one can always terminate the power iteration after a fixed number of steps (even if it has not yet converged), and then use the procedure from section 6.1 to give a lower bound. This is the procedure currently implemented in the "rmu" code. Thus the algorithm is *guaranteed* to return a lower bound on any

problem. In fact the performance data presented in the remainder of this section was collected regardless of whether or not the power algorithm converged on the problem (i.e., no data points were excluded). We will discuss some more advanced power iterations, which further address this convergence problem, in chapter 7.

From the above tests we also found that the potential problems of certain terms in the power iteration becoming undefined do not seem to occur in practice (although it is possible to construct matrices for which this occurs). Also, as was outlined in section 6.1, we note that the algorithm can be "protected" from these kinds of problems. It does appears that we obtain $\tilde{\beta} = \hat{\beta}$ in practice, and hence the algorithm gives us not only a lower bound for $\mu_{\mathcal{K}}(M)$ but also a decomposition as in (4.16). In fact we have not seen an example where the algorithm converged with $\tilde{\beta} \neq \hat{\beta}$ and this is a subject of current research. Note that in the purely complex case $\tilde{\beta} = \hat{\beta}$ is guaranteed (see [59]).

We then wished to examine the average computational requirements of the (upper *and* lower bound) algorithm. For this purpose we used *crand* matrices (although the results are not too different for the different classes). The computational requirements versus matrix size are shown in figure 6.1 for block structures consisting of all scalar uncertainties, with 90% of them chosen as real and the rest complex. The same data for the appropriate complex μ problem is shown for comparison. The results were obtained running Matlab on a Sparc 1 workstation, and it can be seen that we can reasonably expect to handle problems of size 10 in about 10 seconds, up to problems of size 50 in about 2-3 minutes.

It can also be seen that the (experimental) growth rate in computation time for the existing implementation is approximately n^2. This is probably an artifice of the implementation in Matlab, which is an interpretive language. A more realistic measure of the computational growth rate is in terms of total floating point operations (flops). If this measure is adopted then it is seen that the (experimental) growth rate in flops is approximately n^3. In any case the algorithm growth rate appears reasonable

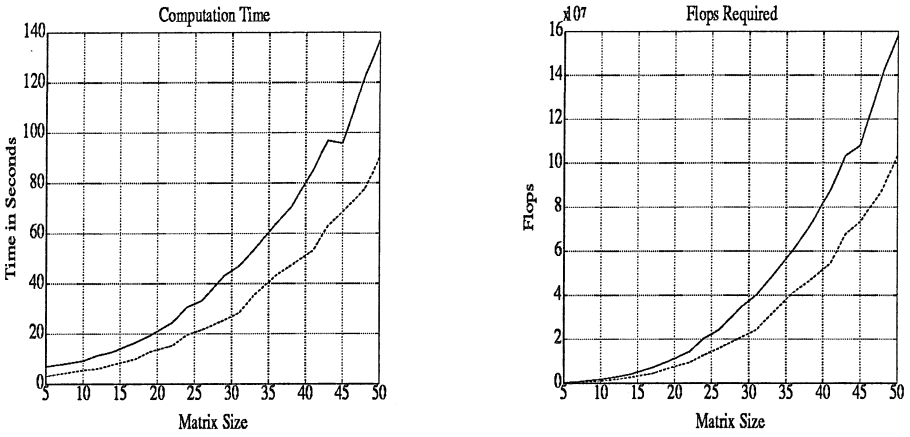


Figure 6.1: Typical computation requirements versus matrix size for mixed μ problem (solid) and complex μ problem (dashed)

whether measured in terms of time or flops required. Note that although this is only experimental data, the results are not too surprising, since it is easy to show that the cost of one iteration of the lower bound power algorithm is of order n^2, and the cost of one gradient step in the upper bound computation is of order n^3.

The next set of tests performed was aimed at evaluating the accuracy of the bounds. Again we used *crand* matrices, and the same class of block structures, except with 80% of the uncertainties chosen to be real. This time we compared the upper and lower mixed μ bounds, and also the mixed μ and complex μ upper bounds. The complex μ bounds were obtained by simply replacing all the real perturbations with complex ones, but without changing the matrix. Thus the complex upper bound is strictly larger than the mixed upper bound. The results are shown in figure 6.2, and indicate that for these problems we are obtaining fairly tight bounds, even for large problems.

It is also apparent that for these problems there is typically not much of a gap between mixed μ and complex μ. This class of matrices is interesting from the point

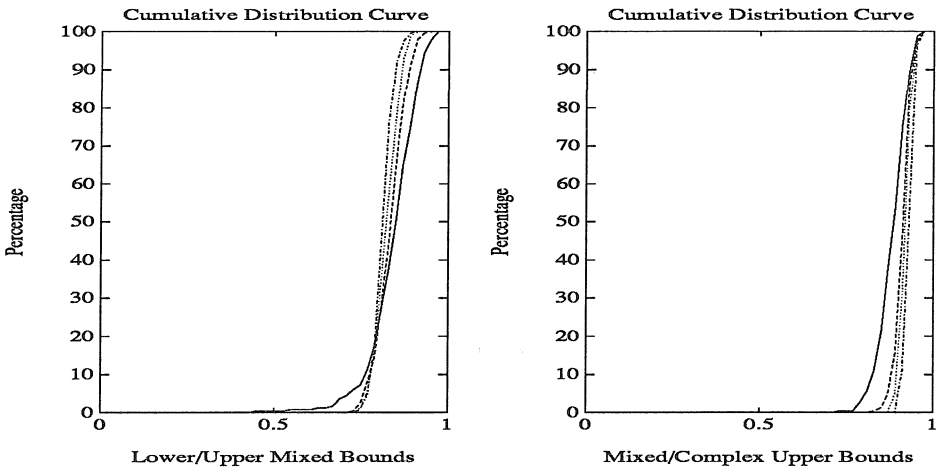


Figure 6.2: Ratios of mixed μ lower/upper bounds, and mixed μ/complex μ (upper) bounds, for a sample of *crand* matrices of sizes 10 (solid), 20 (dashed), 30 (dotted), and 50 (dashdot)

of view of the lower bound performance, since one is obtaining a mixed perturbation achieving a lower bound close to that one could obtain with a complex perturbation. However it is not too interesting from the point of view of the upper bound performance, since the G scaling matrix cannot greatly reduce the upper bound.

It is also doubtful that *crand* matrices are representative of those of practical interest. For these reasons we would like to find a class of matrices where we often encounter problems with a reasonably large gap between mixed μ and complex μ, as well as matrices which are of more practical relevance, and so the same set of tests as outlined above was performed on *sysrand* matrices. The block structure this time had 90% of the uncertainties chosen to be real. The results are shown in figure 6.3, and it can be seen that the bounds are once again reasonably tight, but now we are obtaining a wider spread of values for the gap between complex μ and mixed μ, providing a better test of the upper bound performance.

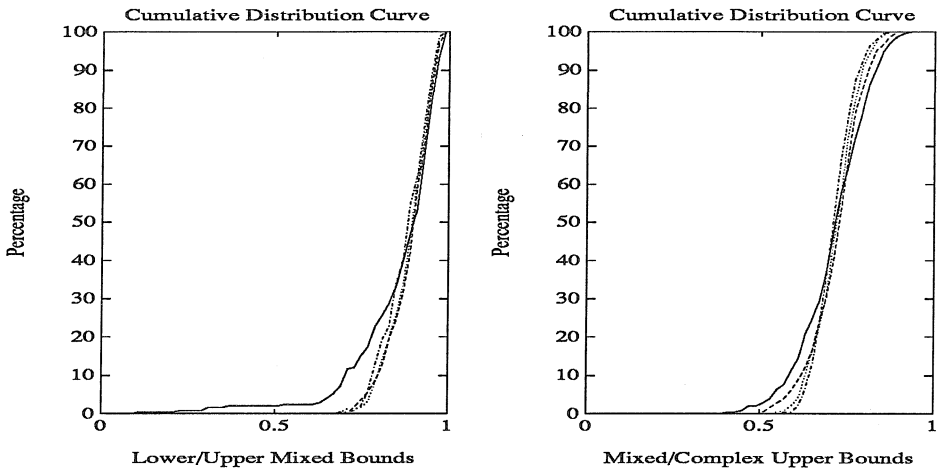


Figure 6.3: Ratios of mixed μ lower/upper bounds, and mixed-μ/complex-μ (upper) bounds, for a sample of *sysrand* matrices of sizes 10 (solid), 20 (dashed), 30 (dotted), and 50 (dashdot)

As a further test the bounds for the mixed μ problem were evaluated across a frequency range for some random stable systems (generated with "sysrand"), and compared to the bounds for the appropriate complex μ problem. Again the bounds seemed reasonably tight, and a typical example plot is shown in figure 6.4.

A number of tests were performed using the *nogap* matrices, and it was found that the upper bound computation was typically with 1-2% of the optimum for these matrices. The lower bound performance was not as good, and in fact the lower bound power iteration can fail to converge on this type of matrix, and yield a poor bound. Of course we cannot expect that our lower bound routine is *guaranteed* to find the correct answer, since it is attempting to maximize a nonconvex problem.

In addition to the above tests a number of other block structures were tested, with similar results (except for the pure real case). Note that all these tests were aimed at evaluating the typical performance of the algorithm on an essentially random selection of problems, and it appears that the algorithm is performing well for most problems.

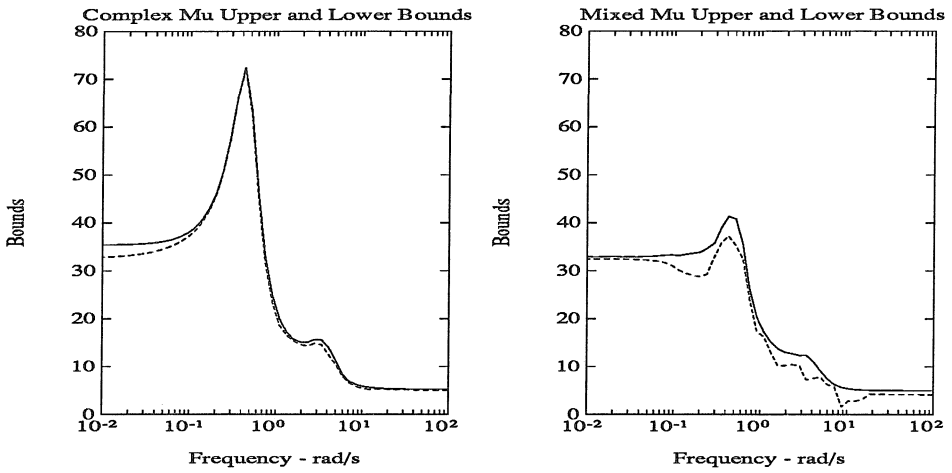


Figure 6.4: Complex μ and mixed μ upper & lower bounds versus frequency for a random system

This does not mean however that one will never encounter mixed μ problems where the gap between the upper and lower bounds is large, and it can be seen from figure 6.3 that a few such cases were found. We will discuss what can be done for these cases in the next chapter.

In addition to the numerical tests described here the "rmu" software has been applied to a number of practical problems, arising from real physical systems. These include analysis of natural frequency variations for flexible structures [42], and variation of missile autopilot dynamics with angle of attack and Mach number [6]. The software worked well for these problems, providing tight bounds for the associated mixed μ problems. We will discuss some of these in more detail in chapter 11. The software is currently being utilized at sites including Honeywell, Phillips, NASA Dryden and several universities. This will provide additional experience regarding the algorithm's performance on real engineering problems.

Chapter 7

Improved Computational Schemes

Although the algorithm presented in chapter 6 will usually provide bounds that are accurate enough for engineering purposes, in a significant number of cases of interest, it will not. In this chapter we consider what can be done for such problems. One possibility for these problems is to improve the algorithm for computing the bounds in theorems 4.1 and 4.5. We will further consider the computation of the lower and upper bounds in sections 7.1 and 7.2, and briefly present some of the latest approaches that are under development. Note however that for some problems the bounds from theorems 4.1 and 4.5 may be far apart (regardless of the computation method). For these cases we must consider improving the bounds themselves (at an additional computational cost), and in section 7.3 we consider the use of Branch and Bound techniques to achieve this goal.

7.1 Lower Bound Improvements

We would like an algorithm to compute a lower bound for μ that is fast, accurate, and reliable. For all iterative algorithms, however, there is always a tradeoff between speed, accuracy, and reliability. The power iteration algorithm for the lower bound described in chapter 6 is no exception to this rule, and there the penalty for having an algorithm that is in general fast and accurate, is that there exist cases where it

fails to converge to a solution. Although one may still obtain a lower bound for these cases, the bound may quite often be poor. In this section we will discuss how an adaptive algorithm can be used to keep the performance and speed of the standard power algorithm (4.28,4.29) for most cases, and enhance the performance for those problems where (4.28,4.29) fails to converge. The material in this section is largely taken from [75].

7.1.1 Connections with the Rank One Problem

Our first alteration to the power iteration is to note that it is attempting to force the right and left eigenvectors of QM to satisfy the alignment condition of theorem 4.2, which is associated with a local maximum of $\rho_R(QM)$. We will see in theorem 8.4 that this alignment condition is also associated with the solution to a certain rank one μ problem formed from these vectors, which leads directly to the following characterization.

Theorem 7.1 *Suppose we have $M \in \mathbb{C}^{n\times n}$ and $Q \in \mathcal{Q}_{\mathcal{K}}$ such that QM has a real positive eigenvalue (so that $\rho_R(QM) > 0$). Further suppose that $q_i^r \neq 0$ for $i = 1,\ldots,m_r$, and that the corresponding right and left eigenvectors of QM, denoted x and y respectively, satisfy the non-degeneracy assumption. Then we have $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ with $\theta_i = \pm\frac{\pi}{2}$ for $|q_i^r| < 1$ and $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$ such that*

$$y = e^{\mathbf{j}\psi} Dx$$

if and only if the matrix $Q \in \mathcal{Q}_{\mathcal{K}}$ solves the rank one μ problem

$$\max_{\hat{Q}\in\mathcal{Q}_{\mathcal{K}}} \rho_R(\hat{Q}M_{r1})$$

where $M_{r1} = \hat{x}y^*$ and $\hat{x} = Q^{-1}x$.

Proof: Note that by assumption we have $y^*x > 0$ so that $y^*Q\hat{x} > 0$ and hence $\rho_R(QM_{r1}) > 0$. The result now follows from theorem 8.4. □

This rank one problem may be easily solved (see chapter 8), giving us the means to

choose $Q \in \mathcal{Q}_{\mathcal{K}}$ which forces the alignment condition. This can be rearranged so as to fit into the power iteration framework via the following lemma.

Lemma 7.1 *Suppose we have* $M \in \mathbb{C}^{n\times n}$, $Q \in \mathcal{Q}_{\mathcal{K}}$ *with* $q_i^r \neq 0$ *for* $i = 1, \ldots, m_r$, *and* $D \in \hat{\mathcal{D}}_{\mathcal{K}}$ *with* $\theta_i = \pm\frac{\pi}{2}$ *for* $|q_i^r| < 1$ *and* $\psi \in (-\frac{\pi}{2} \ \ \frac{\pi}{2})$. *Then we have non-zero vectors* $x, y \in \mathbb{C}^n$ *satisfying the non-degeneracy assumption, and a positive real scalar* $\beta > 0$ *such that*

$$\begin{aligned} QMx &= \beta x \\ y^*QM &= \beta y^* \\ y &= e^{\mathbf{j}\psi} Dx \\ y^*x &> 0 \end{aligned}$$

if and only if there exist non-zero vectors $b, a, z, w \in \mathbb{C}^n$, *with* b, w *satisfying the nondegeneracy assumptions, such that*

$$\begin{aligned} Mb = \beta a \qquad & M^*z = \beta w \\ b = Qa \qquad & z = Q^*w \end{aligned}$$

and $Q \in \mathcal{Q}_{\mathcal{K}}$ *solves the rank one* μ *problem*

$$\max_{\hat{Q}\in\mathcal{Q}_{\mathcal{K}}} \rho_R(\hat{Q}aw^*)$$

with $\mu_{\mathcal{K}}(aw^*) > 0$.

Proof: ($\longrightarrow$) Apply theorem 7.1 and then define $b \doteq x, a \doteq \hat{x}, w \doteq y, z \doteq Q^*x$.

($\longleftarrow$) Define $x \doteq b, y \doteq w$ and apply theorem 7.1. $\square$

On the basis of this lemma, one may define a new power iteration, which is given conceptually below.

Algorithm 7.1 (Lower Bound Using Rank One Solution [75])

1. *Start with initial guesses for* $b, w \in \mathbb{C}^n$.
2. *Update* a *with a power step of the form* $\hat{\beta}a = Mb$.
3. *Compute the* $Q \in \mathcal{Q}_{\mathcal{K}}$ *that maximizes* $\rho_R(Qaw^*)$, *and update* z *as* $z = Q^*w$.
4. *Update* w *with a power step of the form* $\tilde{\beta}w = M^*z$.
5. *Compute the* $Q \in \mathcal{Q}_{\mathcal{K}}$ *that maximizes* $\rho_R(Qaw^*)$, *and update* b *as* $b = Qa$.
6. *If converged stop, else go to step 2.*

Note that for the purely complex case the solution to the rank one μ problem is particularly easy, and it is easy to check that this algorithm is identical to (4.28,4.29). For the mixed case however this algorithm provides us with a different way of choosing the real parameters, and for some problems it appears to perform better than (4.28,4.29). There is a potential problem however in that the real parameters may not be uniquely determined by the rank one solution whenever two block components, $w_{r_i}^* a_{r_i}$, have the same phase modulo $\pm$. In that case one needs to determine which solution is the required one. It seems that a combination of the routine in (4.28,4.29) and the above method should provide a good approach, and this is a subject of current research. We refer the reader to [75] for more details.

7.1.2 Adaptive Power Iteration

The problem with a standard power iteration approach is that it can only ever be stable about the largest *magnitude* eigenvalue. Thus we can never converge to a local maximum of $\rho_R(QM)$ which has $\rho_R(QM) < \rho(QM)$. For complex μ problems this is not an issue, since we are always interested in the largest magnitude eigenvalue. For mixed problems however we are interested in the largest *real* eigenvalue, and so this instability can be a serious drawback of the standard power iteration. Consider the

following matrix:

$$M = \begin{bmatrix} 5+10j & 1-2j \\ 20-10j & 0 \end{bmatrix}$$

with the block structure defined by $m_r = m_c = 1, m_C = 0$ and $\mathcal{K} = (1,1)$ (i.e., one real scalar and one complex scalar). Then one can show that for this matrix $\mu_{\mathcal{K}}(M) = 5$, and furthermore the only matrices $Q \in \mathcal{Q}_{\mathcal{K}}$ achieving the global maximum in (4.1) are $Q = \pm I_2$. For these matrices we have that $\rho_R(QM) = 5$ whereas $\rho(QM) = 10$. This means that the standard power iteration scheme described in (4.28,4.29) could not converge to the global maximum of this problem, since it is not a stable equilibrium point.

Now note that the mixed μ power algorithm represents a generalization of the complex μ power algorithm, which itself can be thought of as generalization of power algorithms for eigenvalues and singular values (see the special case results in chapter 3). It is well known that the convergence properties of standard eigenvalue and singular value power algorithms can be improved by inverse iteration, and these techniques can similarly be applied to (4.28,4.29). The idea behind inverse iteration is simply that for a square matrix $M \in \mathbb{C}^{n\times n}$, and a scalar $\lambda \in \mathbb{C}$, then the largest eigenvalue of $(M - \lambda I_n)^{-1}$ is given as $\frac{1}{\gamma-\lambda}$ where γ is the eigenvalue of M which is *closest* to λ. Note that $\frac{1}{\gamma-\lambda}$ grows in size as γ approaches λ.

In this way an inverse power iteration can be made stable about *any* of the eigenvalues of a given matrix M, by choice of λ. Furthermore since the rate of convergence of a power iteration depends on the (relative) size of the largest eigenvalue, this rate can be greatly enhanced (see [32,76] for more details).

Recall that the lower bound power algorithm for μ is derived by writing a set of equations characterizing the local maxima of $\rho_R(QM)$ and deriving from them a set of recursive formulae whose equilibrium points verify those equations. There is no unique way of deriving such recursive formulae, and different formulations will lead to algorithms with different convergence properties. In the remainder of this section we present a new formulation, which was originally derived in [75], that uses mixed

power and inverse-power iteration. In this way the algorithm is always powering to find the *largest* eigenvalue of some matrix (even for mixed problems), so that the instability problem described earlier is alleviated.

In order to do this we introduce the following notation. Separate the perturbation $Q \in \mathcal{Q}_{\mathcal{K}}$ into its real, Q_1, and complex (both scalars and full blocks), Q_2, components.

$$Q = \begin{pmatrix} Q_1 & 0 \\ 0 & Q_2 \end{pmatrix}.$$

Analogously partition the vectors $b, a, z, w \in \mathbb{C}^n$:

$$b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} \qquad a = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} \qquad z = \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \qquad w = \begin{pmatrix} w_1 \\ w_2 \end{pmatrix}$$

and the matrices $M \in \mathbb{C}^{n\times n}$, $D \in \hat{\mathcal{D}}_{\mathcal{K}}$:

$$M = \begin{pmatrix} M_{11} & M_{21} \\ M_{12} & M_{22} \end{pmatrix} \qquad D = \begin{pmatrix} D_1 & 0 \\ 0 & D_2 \end{pmatrix}.$$

Theorem 7.2 ([75]) *Suppose we have $M \in \mathbb{C}^{n\times n}$, together with a compatible block structure $\mathcal{K}$, and $D \in \hat{\mathcal{D}}_{\mathcal{K}}, Q \in \mathcal{Q}_{\mathcal{K}}$ with $q_i^r \neq 0$ for $i = 1, \ldots, m_r$. Further suppose we have a real positive scalar $\beta > 0$ such that $\beta I_{n_r} - M_{11}Q_1$ is non-singular. Then there exist non-zero vectors $b, z, a, w \in \mathbb{C}^n$ such that*

$$Mb = \beta a \qquad M^*z = \beta w \tag{7.1}$$

$$b = Qa \qquad b = D^{-1}w \tag{7.2}$$

$$z = Q^*QDa \qquad z = Q^*w$$

if and only if there exist non-zero vectors $b_2, z_2, a_2, w_2 \in \mathbb{C}^{n_c}$ such that

$$M_C b_2 = \beta a_2 \qquad M_C^* z_2 = \beta w_2 \tag{7.3}$$

$$b_2 = Q_2 a_2 \qquad b_2 = D_2^{-1} w_2 \tag{7.4}$$

$$z_2 = Q_2^* Q_2 D_2 a_2 \qquad z_2 = Q_2^* w_2$$

$$Q_1(\beta I_{n_r} - M_{11}Q_1)^{-1}M_{12}b_2 = D_1^{-1}(\beta I_{n_r} - M_{11}^*Q_1)^{-1}M_{21}^* z_2 \tag{7.5}$$

where $M_C = (M_{22} + M_{21}Q_1(\beta I_{n_r} - M_{11}Q_1)^{-1}M_{12})$.

Proof: ($\longrightarrow$) Partition b, z, a, w as above and we may rewrite (7.1) as

$$\beta a_1 = M_{11}Q_1 a_1 + M_{12}Q_2 a_2 \tag{7.6}$$

$$\beta a_2 = M_{21}Q_1 a_1 + M_{22}Q_2 a_2 \tag{7.7}$$

$$\beta z_1 = Q_1 M_{11}^* z_1 + Q_1 M_{21}^* z_2 \tag{7.8}$$

$$\beta z_2 = Q_2^* M_{12}^* z_1 + Q_2^* M_{22}^* z_2. \tag{7.9}$$

Note that it is easy to argue by contradiction that $b_2, z_2, a_2, w_2 \neq 0$. Solving for a_1 and z_1 in (7.6) and (7.8), and substituting in (7.7) and (7.9), we obtain (7.3). Equations (7.4) and (7.5) derive directly from (7.2).

($\longleftarrow$) Define the following vectors

$$\begin{aligned} a_1 &= (\beta I_{n_r} - M_{11}Q_1)^{-1} M_{12} b_2 \\ w_1 &= (\beta I_{n_r} - M_{11}^* Q_1)^{-1} M_{21}^* z_2 \\ b_1 &= Q_1 a_1 \\ z_1 &= Q_1 w_1. \end{aligned}$$

Then (7.3) implies (7.6-7.9). Making the obvious definitions for b, z, a, w these equations imply $b, z, a, w \neq 0$ and (7.1). Finally (7.4) and (7.5) imply (7.2). □

These conditions are easily rewritten in terms of the block components as before. Assuming the usual non-degeneracy conditions we may once more apply lemmas 4.7,4.8 and 4.9 to obtain the equivalent conditions:

$$\begin{aligned} \beta a_2 &= (M_{22} + M_{21}Q_1(\beta I_{n_r} - M_{11}Q_1)^{-1} M_{12}) b_2 \\ z_{c_i} &= \frac{w_{c_i}^* a_{c_i}}{|w_{c_i}^* a_{c_i}|} w_{c_i} \qquad \text{for } i = 1, \ldots, m_c \\ z_{C_i} &= \frac{|w_{C_i}|}{|a_{C_i}|} a_{C_i} \qquad \text{for } i = 1, \ldots, m_C \\ \beta w_2 &= (M_{22}^* + M_{12}^*(\beta I_{n_r} - Q_1 M_{11}^*)^{-1} Q_1 M_{21}^*) z_2 \\ b_{c_i} &= \frac{a_{c_i}^* w_{c_i}}{|a_{c_i}^* w_{c_i}|} a_{c_i} \qquad \text{for } i = 1, \ldots, m_c \\ b_{C_i} &= \frac{|a_{C_i}|}{|w_{C_i}|} w_{C_i} \qquad \text{for } i = 1, \ldots, m_C \end{aligned} \tag{7.10}$$

where the real parameters, q_i^r, satisfy

$$\begin{aligned} \mathrm{Re}\,(a_{r_i}^* w_{r_i}) &\geq 0 \quad \text{if} \quad q_i^r = 1 \\ \mathrm{Re}\,(a_{r_i}^* w_{r_i}) &\leq 0 \quad \text{if} \quad q_i^r = -1 \qquad \text{for } i = 1, \ldots, m_r \\ \mathrm{Re}\,(a_{r_i}^* w_{r_i}) &= 0 \quad \text{if} \quad |q_i^r| < 1 \end{aligned} \tag{7.11}$$

and a_{r_i}, w_{r_i} are formed by partitioning the vectors a_1, w_1, where a_1, w_1 are given by

$$\begin{aligned} a_1 &= (\beta I_{n_r} - M_{11} Q_1)^{-1} M_{12} b_2 \\ w_1 &= (\beta I_{n_r} - M_{11}^* Q_1)^{-1} M_{21}^* z_2. \end{aligned} \tag{7.12}$$

These equations suggest the following iterative algorithm, which has (7.10,7.11,7.12) as equilibrium conditions.

Algorithm 7.2 (Lower Bound: Mixed Power/Inverse Power Iteration [75])

1. *Start with initial guesses for* $b_2, w_2 \in \mathbb{C}^{n_c}$, Q_1 *and* $\beta > 0$.
2. *Update* $M_C = (M_{22} + M_{21} Q_1 (\beta I_{n_r} - M_{11} Q_1)^{-1} M_{12})$.
3. *Update* a_2 *with a power step of the form* $\hat{\beta} a_2 = M_C b_2$.
4. *Compute the* Q_2 *that maximizes* $\rho_R(Q_2 a_2 w_2^*)$, *and update* z_2 *as* $z_2 = Q_2^* w_2$.
5. *Update* w_2 *with a power step of the form* $\tilde{\beta} w_2 = M_C^* z_2$.
6. *Compute the* Q_2 *that maximizes* $\rho_R(Q_2 a_2 w_2^*)$, *and update* b_2 *as* $b_2 = Q_2 a_2$.
7. *Compute* a_1, w_1 *using (7.12).*
8. *Compute the* $Q \in \mathcal{Q}_{\mathcal{K}}$ *that maximizes* $\rho_R(Q a w^*)$, *and update* Q_1.
9. *Update* β *as* $\frac{1}{2}(\hat{\beta} + \tilde{\beta})$.
10. *If converged stop, else go to step 2.*

Although theorem 7.2 represents a simple algebraic rearrangement of the conditions in (7.1,7.2), it has a great impact on the behavior of the power algorithm derived from them. This is because the power steps on M_C in the above algorithm are *always* attempting to find the *largest* eigenvalue. Essentially this may be thought of

as implementing inverse power iteration on the real parts of the problem (where we need the largest real eigenvalue), and power iteration on the complex parts of the problem (where we simply need the largest eigenvalue). In this way the new algorithm removes the aforementioned instability problem for equilibrium points where $\rho_R(QM) < \rho(QM)$, and so we would expect the (convergence) properties to be significantly enhanced.

Note that in the above algorithm the matrices M_C and Q_2 don't actually have to be formed as we are only interested in the products of these matrices with vectors. By optimizing the way we actually implement these operations, we can substantially reduce the computation time. Nevertheless it is still true that the above algorithm involves more computation per iteration than (4.28,4.29). Recall that the motivation for investigating more advanced power iteration schemes was to develop an algorithm that would preserve the numerical properties (speed, accuracy, growth rate) of the standard power iteration for those cases where it works well, and improve its convergence properties (and hence accuracy) in the other cases, without excessively degrading its speed. A reasonable way to achieve this purpose is to have an adaptive power algorithm, which has available a number of different power iterations of increasing accuracy and complexity, together with a scheduling rule that starts with the fastest one and progressively shifts to the others, if they fail to converge. The results obtained using a simple scheduling rule are presented in the next section.

7.1.3 Algorithm Performance

Recall that the standard power iteration (4.28,4.29) appeared to perform well when tested on *crand* matrices (see section 6.4). For instance, in a typical test run of 500 *crand* 5×5 matrices with 2 real scalar uncertainties and three complex scalar uncertainties the standard power iteration converged 96% of the time (in an average of 22 iterations). In fact the new iteration was found to further improve on this performance, increasing the convergence rate to 99% for this test.

However these results are not so interesting, since the performance of the original power iteration is already satisfactory for these problems. What is required is a class of problems where the standard power iteration often performs poorly. It turns out that the *nogap* matrices can be generated so that they often satisfy $\rho_R(QM) \ll \rho(QM)$ at the global maximum, and this causes problems for the standard power iteration.

The results obtained for a random sampling of 100 *nogap* matrices for two different uncertainty structures are shown in figures 7.1 and 7.2 (these figures are taken from [75]). Note from figure 7.1 that for this case over 50% of the answers given by the new algorithm are larger than 0.9 as compared to less than 15% for the standard power iteration (recall that by construction the *nogap* matrices satisfy $\mu_{\mathcal{K}}(M) = 1$).

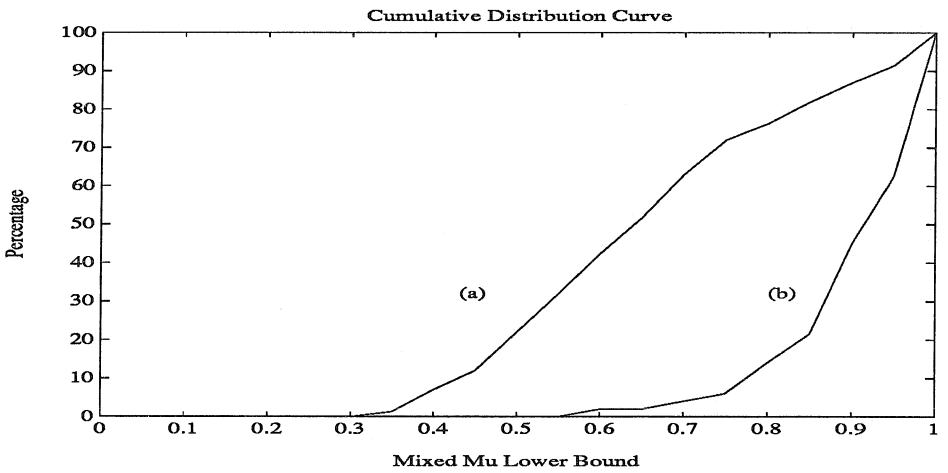


Figure 7.1: Comparison of (a) standard power iteration, and (b) adaptive power iteration on *nogap* matrices, with the uncertainty structure consisting of 4 real scalars, 2 complex scalars, and a 2×2 full complex block

The performance degrades somewhat when the number of real parameters increases, but it can be seen from figure 7.2 that the degradation is much more graceful

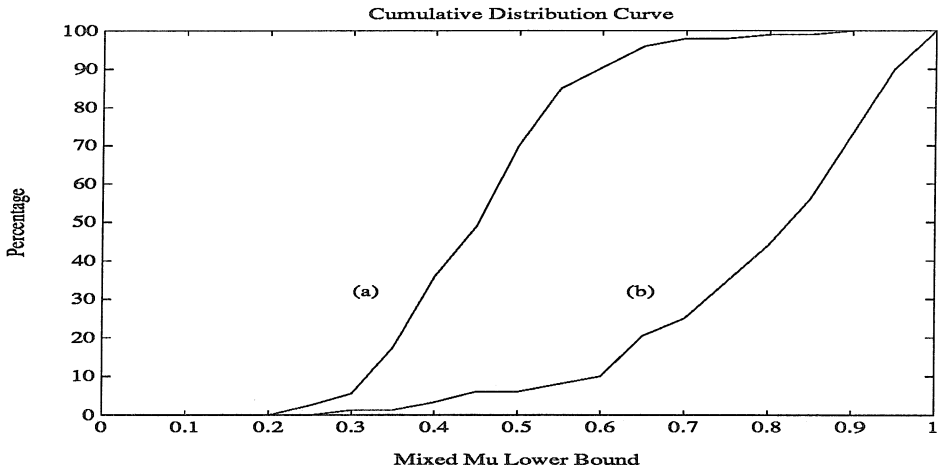


Figure 7.2: Comparison of (a) standard power iteration, and (b) adaptive power iteration on *nogap* matrices, with the uncertainty structure consisting of 8 real scalars, 2 complex scalars, and a 2×2 full complex block

for the new algorithm. Note also that in both cases the new algorithm converged to an equilibrium point 95% of the time (so that the bound obtained is associated with a local maximum of $\rho_R(QM)$). This is for a class of problems where the standard power iteration typically fails to converge.

Note that the new algorithm is still not *guaranteed* to converge. However there are further possible modifications to the power iteration. First note from (7.1,7.2) that we have

$$\begin{aligned} Mb &= \beta a \\ \longrightarrow \quad QMb &= \beta b \\ \longrightarrow \quad (QM)^k b &= \beta^k b \end{aligned}$$

so that one could increase the number of power steps taken during each iteration, and consequently improve the current estimates of the eigenvalue and associated eigenvectors. Then note that if instead of updating Q as currently proposed, we

instead take a relaxed step of the form

$$Q_{new} = (1 - t)Q_{old} + tQ_{rank\ 1}$$

(where $t \in (0\ 1]$ and $Q_{rank\ 1}$ maximizes $\rho_R(Qaw^*)$ as before), one can show that if we are sufficiently close to the equilibrium conditions, then for t small enough, Q is updated in the direction of the gradient (and for $t = 1$ we have the current scheme [75]). Since gradient search techniques are reliable (but slow), these two facts can in principle be used to develop an adaptive power iteration with *guaranteed* convergence, and a first attempt at such a scheme may be found in [75]. The challenge is to develop an adaptive algorithm which has guaranteed convergence, without an excessive degradation in speed, and this is a subject of current research.

7.2 Upper Bound Improvements

The mixed μ upper bound, in the form of (4.39), can be viewed as a special case of a class of LMI problems. The solution of LMI's is a subject of much research interest right now, since they appear in many control problems (see [22]). Of course one could tackle these problems with standard optimization techniques, such as gradient search and related descent methods, although these methods may not be very efficient (see [45] for a review of some standard optimization techniques). In particular, difficulties can occur because of the fact that the function to be minimized may not be differentiable at a point where the eigenvalues are coalesced (which may well be the case near the optimum). Since LMI optimization problems are convex in the free variables, there are a variety of numerical techniques from convex programming one can employ for these problems, such as ellipsoid and cutting plane methods (see [14] for a review of these methods), although once again the difficulty is not so much with developing an algorithm which is guaranteed to converge to the solution, but finding one that does it efficiently.

A number of researchers are investigating these types of problems. Fan has pro-

posed an algorithm which has quadratic convergence [27], under certain assumptions about the minimizing solution. Overton has developed descent techniques which specifically address the problem of repeated eigenvalues and non-differentiability [54]. The use of interior point methods has been investigated by Nesterov and Nemirovsky [48], and appears to have great promise. Several methodologies for LMI solution have been based on this approach (see [15,35] for example). Recent work by Beck and Doyle [13] examines the use of hybrid algorithms for LMI problems. This research area is quite active, and it is expected that improved algorithms for LMI solution will soon appear.

The algorithm developed in chapter 6 represents a first attempt at solving one particular LMI. As more refined algorithms for the solution of LMI's appear, then they can be used to improve the μ upper bound computation. We refer the interested reader to [12] for a review of LMI optimization techniques, and [13] for the application of LMI techniques to the mixed μ upper bound problem.

7.3 Branch and Bound

As we mentioned earlier there are problems for which the bounds from theorems 4.1 and 4.5 may be far apart (regardless of the computation method). For these cases we must consider improving the bounds themselves. The use of Branch and Bound schemes, to improve upon existing bounds, has been suggested by several authors (see [70,19,2] and the references therein), and in this section we consider the application of Branch and Bound techniques to the mixed μ problem.

7.3.1 Mixed μ & Branch and Bound

Branch and Bound is a general technique for computing improved bounds for a given optimization problem. The fundamental idea is to note that the gaps between the upper and lower bounds for the optimization problem depend on the domain of optimization. Thus if we partition the original domain of optimization, we obtain easier

problems (i. e., problems with smaller gaps) and so combining these results we get better bounds on the original domain. The finer we partition the domain of optimization, the smaller the gap between the upper and lower bounds.

In the context of the mixed μ problem then, one may 'chop' the subspace of real parameters into two subdomains and then evaluate the bounds on each subdomain (branch). One thus obtains upper and lower bounds for each subdomain of the partitioned space, and by choosing the *largest* of each of these we obtain new upper and lower bounds for the original problem. This branching process is then repeated as often as necessary, to refine the bounds as accurately as desired. This is stated more formally below.

Define the (maximum positive real eigenvalue) function $\overline{\lambda}_R : \mathbb{C}^{n\times n} \to \mathbb{R}$ as

$$\overline{\lambda}_R(M) \doteq \max\{\lambda : \lambda \text{ is a positive real eigenvalue of } M\} \tag{7.13}$$

with $\overline{\lambda}_R(M) = 0$ if M has no positive real eigenvalues. Then it is easy to show that an equivalent definition of μ is given by

$$\mu_{\mathcal{K}}(M) = \max_{\Delta\in\mathbb{B}X_{\mathcal{K}}} \overline{\lambda}_R(\Delta M). \tag{7.14}$$

Clearly if we have sets $\mathbb{B}X_j \subset X_{\mathcal{K}}$ with $\bigcup \mathbb{B}X_j = \mathbb{B}X_{\mathcal{K}}$ and we define

$$\mu_j(M) = \max_{\Delta\in\mathbb{B}X_j} \overline{\lambda}_R(\Delta M) \tag{7.15}$$

then it follows that

$$\mu_{\mathcal{K}}(M) = \max_j \mu_j(M). \tag{7.16}$$

Now denote upper and lower bounds for μ_j as ub_j and lb_j respectively, so that

$$lb_j \leq \mu_j \leq ub_j$$

and define the quantities

$$\begin{aligned} L &\doteq \max_j lb_j \\ U &\doteq \max_j ub_j. \end{aligned} \tag{7.17}$$

Then immediately we have that

$$\max_j lb_j \leq \max_j \mu_j \leq \max_j ub_j$$

so that

$$L \leq \mu \leq U. \tag{7.18}$$

Thus L and U are upper and lower bounds for $\mu_{\mathcal{K}}(M)$ which depend on the "local" bounds lb_j and ub_j, and on the "partitioning" $\bigcup \mathbb{B}X_j = \mathbb{B}X_{\mathcal{K}}$.

One simple procedure for partitioning $\mathbb{B}X_{\mathcal{K}}$ into the $\mathbb{B}X_j$'s is to chop one of the 'longest sides' [2]. This leads to the following Branch and Bound scheme for the mixed μ problem.

Algorithm 7.3 (Branch and Bound for Mixed μ [51])

Initialize $\{\mathbb{B}X_j\} = \mathbb{B}X_{\mathcal{K}}$

Let $U = \max_j ub_j$

$L = \max_j lb_j$

while $U - L > \epsilon$

Let $\mathbb{B}X_a$ be any element of $\{\mathbb{B}X_j\}$ with $ub_a = U$.

Partition $\mathbb{B}X_a$ into $\mathbb{B}X_b$ and $\mathbb{B}X_c$ by bisecting $\mathbb{B}X_a$ along one of its longest edges.

Add $\mathbb{B}X_b$ and $\mathbb{B}X_c$ to $\{\mathbb{B}X_j\}$.

Remove $\mathbb{B}X_a$ from $\{\mathbb{B}X_j\}$.

endwhile

With mild assumptions on the bounds lb_j and ub_j, it is easy to prove that this algorithm has *guaranteed* convergence [51]. There are a number of enhancements one can make to this scheme, with more sophisticated partitioning procedures, and they similarly can be shown to have guaranteed convergence (see [51] for details). However it is immediately apparent that one has the potential to encounter problems with exponential growth rates using this approach. In fact it can be shown that if one can construct a matrix where the bounds are not within the required tolerance at

the first step (which we can for even our best bounds), then one can build from this examples which will require exponential time growth rate for the Branch and Bound scheme. Essentially what this result says is that if one has to Branch and Bound at all, then it will be exponential time in the worst case. This is not at all surprising since this scheme provides us with an algorithm to compute *guaranteed* bounds for mixed μ, which we know to be an NP hard problem. The real issue then is whether or not we can produce a "practical" scheme, whose *typical* computation time is polynomial (despite the fact that the worst case computation time is exponential). There are some important issues and tradeoffs to be considered in implementing such a scheme, which can greatly impact the performance. These issues will be examined in the next subsection, but prior to that we show how the mixed μ problem can be put in a Branch and Bound Framework. This material is largely taken from [51].

Now from the above it follows that in order to implement a Branch and Bound algorithm for the mixed μ problem we need to be able to compute upper and lower bounds for the quantity $\mu_j(M)$, where the parameter space has been partitioned into subdomains of the form:

$$\mathbb{B}X_j = \{\Delta \in \mathbb{B}X_{\mathcal{K}} : \delta_i^r \in [a_i\ b_i] \text{ with } a_i, b_i \in [-1\ 1], a_i \leq b_i, \text{ for } i = 1, \ldots, m_r\}. \tag{7.19}$$

Consider first the lower bound. It is a simple extension of theorem 4.1 to show that

$$\max_{\Delta \in \mathbb{B}X_j} \overline{\lambda}_R(\Delta M) = \max_{Q \in \mathcal{Q}_j} \overline{\lambda}_R(QM) \tag{7.20}$$

where the set $\mathcal{Q}_j$ is defined as

$$\mathcal{Q}_j = \{Q \in \mathcal{Q}_{\mathcal{K}} : q_i^r \in [a_i\ b_i] \text{ with } a_i, b_i \in [-1\ 1], a_i \leq b_i, \text{ for } i = 1, \ldots, m_r\}. \tag{7.21}$$

From this point on we may essentially repeat the machinery of chapter 4 to reduce our problem to one of finding vectors $b, z, a, w \in \mathbb{C}^n$, and a real scalar $\beta > 0$ satisfying the equations

$$
\begin{aligned}
\beta a &= Mb \\
z_{r_i} &= q_i^r w_{r_i} && \text{for } i = 1, \ldots, m_r \\
z_{c_i} &= \frac{w_{c_i}^* a_{c_i}}{|w_{c_i}^* a_{c_i}|} w_{c_i} && \text{for } i = 1, \ldots, m_c \\
z_{C_i} &= \frac{|w_{C_i}|}{|a_{C_i}|} a_{C_i} && \text{for } i = 1, \ldots, m_C \\
\beta w &= M^* z && \qquad (7.22) \\
b_{r_i} &= q_i^r a_{r_i} && \text{for } i = 1, \ldots, m_r \\
b_{c_i} &= \frac{a_{c_i}^* w_{c_i}}{|a_{c_i}^* w_{c_i}|} a_{c_i} && \text{for } i = 1, \ldots, m_c \\
b_{C_i} &= \frac{|a_{C_i}|}{|w_{C_i}|} w_{C_i} && \text{for } i = 1, \ldots, m_C
\end{aligned}
$$

where the real parameters, q_i^r, satisfy

$$
\begin{aligned}
&\mathrm{Re}\,(a_{r_i}^* w_{r_i}) \geq 0 \quad \text{if} \quad q_i^r = b_i \\
&\mathrm{Re}\,(a_{r_i}^* w_{r_i}) \leq 0 \quad \text{if} \quad q_i^r = a_i \qquad \text{for } i = 1, \ldots, m_r \qquad (7.23) \\
&\mathrm{Re}\,(a_{r_i}^* w_{r_i}) = 0 \quad \text{if} \quad q_i^r \in (a_i\ b_i).
\end{aligned}
$$

As before this set of equations immediately leads to a power iteration to compute a lower bound for $\mu_j(M)$. This involves only a slight modification to (4.28,4.29), and furthermore we may also apply the more advanced power iteration techniques from section 7.1 to further enhance the performance.

The basic idea behind computing an upper bound for $\mu_j(M)$ is, as before, to cover the real uncertainty with some sort of shifted disk (see figure 4.3) and then apply standard results for disk uncertainty. Since we have greater restrictions on the range of the real parameters we may use smaller shifted disks to get a better bound. To this end define the center and radius matrices, C_j and R_j respectively, as diagonal real matrices

$$
\begin{aligned}
C_j &= \text{block diag}(c_1, \ldots, c_{n_r}, 0_{n_c}) \\
R_j &= \text{block diag}(r_1, \ldots, r_{n_r}, I_{n_c}) \qquad (7.24)
\end{aligned}
$$

(with $c_i, r_i \in \mathbb{R}$ and $r_i \geq 0$ for $i = 1, \ldots, n_r$) which define the perturbation set

$$\mathbb{B}X_j = \{\Delta : \Delta = C_j + R_j\hat{\Delta} \text{ for } \hat{\Delta} \in \mathbb{B}X_{\mathcal{K}}\}. \tag{7.25}$$

Then we have the following upper bound for $\mu_j(M)$.

Theorem 7.3 ([51]) *Suppose we have a matrix $M \in \mathbb{C}^{n\times n}$, a compatible block structure $\mathcal{K}$, and a partitioned perturbation set $\mathbb{B}X_j$. Further suppose that* $\det(I_n - \frac{M}{\alpha}C_j) \neq 0$, *and define*

$$\hat{M} \doteq R_j^{\frac{1}{2}}(I_n - \frac{M}{\alpha}C_j)^{-1}\frac{M}{\alpha}R_j^{\frac{1}{2}}. \tag{7.26}$$

Then we have that

$$\mu_j(M) \leq \inf_{\substack{D\in\mathcal{D}_{\mathcal{K}} \\ G\in\mathcal{G}_{\mathcal{K}}}} \left[\min_{\substack{\hat{\alpha}\in\mathbb{R} \\ \hat{\alpha}\geq 0}} \{\hat{\alpha} : (\hat{M}^*D\hat{M} + \mathbf{j}(G\hat{M} - \hat{M}^*G) - D) \leq 0 \text{ for all } \alpha \geq \hat{\alpha}\}\right]. \tag{7.27}$$

Note that if $\det(I_n - \frac{M}{\alpha}C_j) = 0$ then α is a lower bound for $\mu_j(M)$. In order to compute this quantity then define the (maximum real eigenvalue) function $\overline{\lambda}_r : \mathbb{C}^{n\times n} \to \mathbb{R}$ as

$$\overline{\lambda}_r(M) \doteq \max\{\lambda : \lambda \text{ is a real eigenvalue of } M\} \tag{7.28}$$

with $\overline{\lambda}_r(M) = -\infty$ if M has no real eigenvalues, and we have the following result.

Theorem 7.4 ([51]) *Suppose we have the set-up of theorem 7.3. Then for any fixed $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ we have that*

$$\begin{aligned} &\min_{\hat{\alpha}\in\mathbb{R}}\{\hat{\alpha} : (\hat{M}^*D\hat{M} + \mathbf{j}(G\hat{M} - \hat{M}^*G) - D) \leq 0 \text{ for all } \alpha \geq \hat{\alpha}\} \\ = \overline{\lambda}_r &\begin{bmatrix} C_jM^* - \mathbf{j}R_j^{\frac{1}{2}}GD^{-1}R_j^{\frac{1}{2}}M^* & R_j^{\frac{1}{2}}DR_j^{\frac{1}{2}}M + R_j^{\frac{1}{2}}GD^{-1}GR_j^{\frac{1}{2}}M \\ R_j^{\frac{1}{2}}D^{-1}R_j^{\frac{1}{2}}M^* & C_jM + \mathbf{j}R_j^{\frac{1}{2}}D^{-1}GR_j^{\frac{1}{2}}M \end{bmatrix}. \end{aligned} \tag{7.29}$$

These two theorems give us the means to compute an upper bound for $\mu_j(M)$. Note that we can apply the machinery of chapter 6 to also develop a $\overline{\sigma}$ formulation of this upper bound, and a practical algorithm for its computation (see [51]). Furthermore since this bound also has an LMI formulation we can expect that improved LMI

algorithms can be used here as well. In fact it is easy to check that the standard upper and lower bounds for the mixed μ problem, presented in chapter 4, are special cases of the bounds presented here for the partitioned μ_j problem. This is also related to work in [41] on one-sided μ problems, though here we allow an arbitrary partition of the real parameter space, rather than merely a sign restriction on the real parameters. These results now enable us to implement a Branch and Bound algorithm for the mixed μ problem.

7.3.2 A Practical Branch and Bound Scheme

Important issues one must consider in developing a practical Branch and Bound scheme include the tradeoff between the computational cost versus accuracy of the bounds themselves, for any given sub-problem, and also the amount of computational cost one is prepared to pay in order to evaluate a good direction to chop the remaining subspace. In order to carry out a preliminary examination of these issues we first implemented two different Branch and Bound schemes, which we will refer to as scheme A and scheme B. These Branch and Bound schemes computed a certain robust stability margin for a problem closely related to the mixed μ problem (see [82] for details). Scheme A used the best bounds we had available for the problem, which involved a significant computational cost. In Scheme B the upper and lower bounds were computed using straightforward norm inequalities, with no optimization or scaling matrices, and as such represent cheap but crude bounds. Additionally Scheme A used a more intelligent (but more expensive) chopping criterion than Scheme B. A series of tests were carried out running these two schemes on problems involving *sysrand* matrices (see section 6.3). The block structures consisted of 2^p real (unrepeated) scalar uncertainties and one $p \times p$ full complex block, for various choices of p (with the aforementioned quantities rounded to the nearest integer). The results are shown in figure 7.3. For each choice of problem size the two schemes were run on the same 100 test matrices, and the plots indicate the number of 'steps' required

by each scheme as a function of problem size (where the initial bounds were counted as the first step, and each chop was counted as an additional step). The labeling is best explained by example. The label "B50" means that this is the worst problem encountered by Scheme B, from among the easiest 50% of the problems (for Scheme B). Note that each scheme was allowed a maximum of 1000 steps (hence the curves for scheme B appear to terminate prematurely, because the next points were larger than 1000).

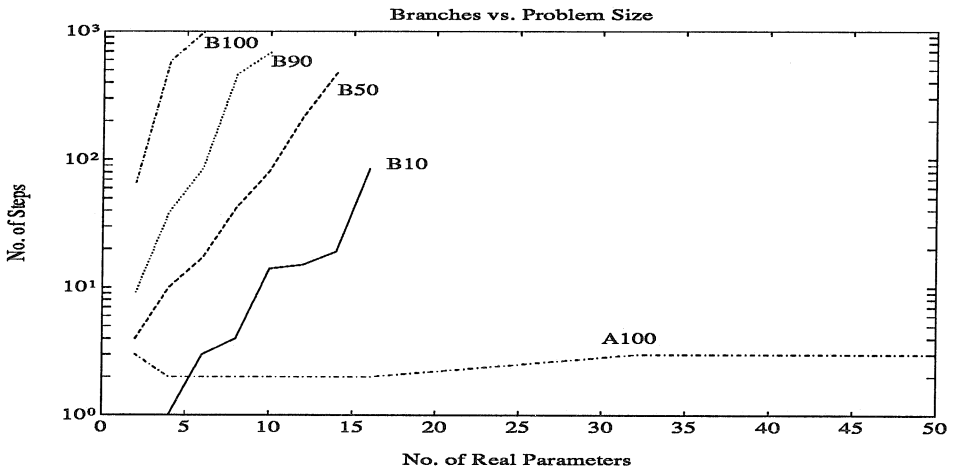


Figure 7.3: Growth rate of Branch and Bound computation steps, for schemes with expensive versus crude bounds

It can be readily seen from Figure 7.3 that the (experimental) growth rate for Scheme B is exponential on the problem data for *any* of the levels "B100"-"B10" (note that the results are plotted on a log-linear scale). Scheme A however appears to have a quite reasonable growth rate on these problems, even for the "A100" level. Note also that Scheme B failed to converge (in 1000 steps) on many of the problems, whereas Scheme A converged fairly rapidly on all the problems.

These results clearly indicate that for a practical Branch and Bound scheme the methods for computing the bounds and the chopping criterion are absolutely critical to the performance on even medium sized problems (see [82] for more details). One is prepared to spend a high computational cost on both of these, provided it is still polynomial time, since one is potentially avoiding exponential time growth in the behavior of the Branch and Bound scheme (note that if any branch yields no improvement in the bounds then the subsequent computation can be doubled, since the same computation may have to be performed for each branch).

We would like to know what kind of performance level we can expect to achieve from a Branch and Bound scheme for the mixed μ problem. It is clear from the above discussion that we need to use sophisticated bounds (despite their computational expense) if we expect to get any kind of high performance scheme with reasonable computational requirements for fairly large problems. In order to examine the properties of such a scheme we implemented a Branch and Bound scheme for the mixed μ problem using the best currently available bounds (including a preliminary version of the lower bound from section 7.1 [75]). This was then used to collect statistical data on the performance, by running the scheme repeatedly on random problems (again, we are interested in typical, rather than worst case, performance for reasons discussed earlier).

This Branch and Bound scheme was used to compute upper and lower bounds for mixed μ problems on *sysrand* matrices (see section 6.3). The uncertainties consisted of m_r real scalars, and (approximately) $\frac{m_r}{4}$ complex scalars, where m_r ranged from 2 to 64. The results from one such batch of tests are shown in figure 7.4. There we have plotted the required number of branches versus number of real parameters for a series of Branch and Bound tests. Thus the curves represent required computational effort versus problem size. For each curve we have plotted the worst problem encountered from a pre-set number of runs, where for each problem the requirement for convergence was to reach a pre-specified tolerance between the upper and lower bounds, as labeled

on the curve. Tolerances of $1\%, 5\%, 10\%$ and 20% were considered, and for any problem the run was terminated if it failed to converge to the required tolerance within 100 branches (hence some of the curves terminate prematurely if the next problem size did not converge in time). Note that the plots are on a log-linear scale, so that any straight line with non-zero slope represents an exponential growth rate.

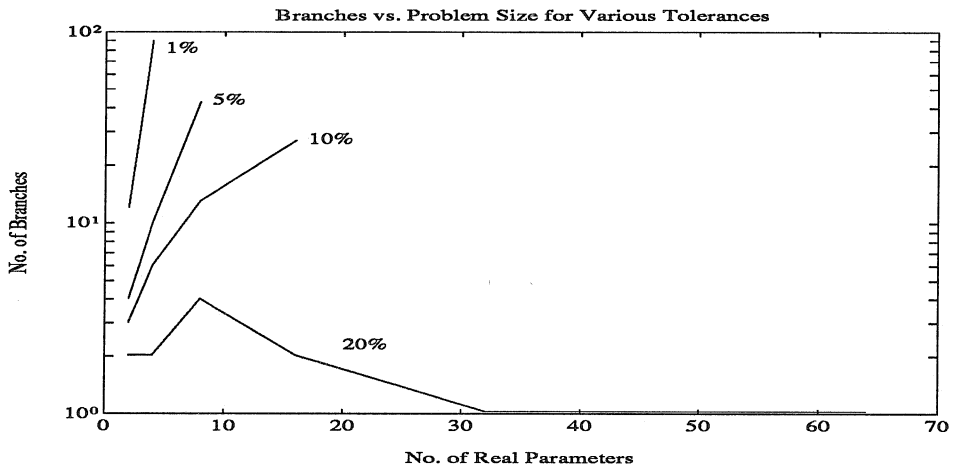


Figure 7.4: Branch and Bound computational requirements for varying degrees of required accuracy

It is clear from figure 7.4 that if the tolerance is set tight enough then the typical growth rate is unacceptable (see the 1% curve for example). Thus as the problem size increases the required computation quickly becomes impractical, and so we cannot expect to be able to achieve these tolerances. Note however that for the 20% curve the computational requirements remain modest even for the largest problems tested. Thus we can reasonably expect to be able to achieve this level of accuracy. Fortunately this degree of accuracy is quite sufficient for engineering purposes. It is important to keep in mind that our mathematical models are only approximations to real physical

systems, and the uncertainties are intended to cover the deficiencies in our knowledge of that system. Thus it is somewhat naive to think that we can have precise knowledge of the uncertainty levels in real engineering problems.

It is interesting to note that for the 20% level the bounds were usually within tolerance at the first try, so that it was usually not necessary to branch at all. This suggests that if one is interested in solving fairly large problems, then one can only expect the Branch and Bound scheme to achieve a degree of accuracy that the bounds usually get anyway! Thus the Branch and Bound scheme is not being used as a general computation scheme per se, but only to fix the occasional problems for which the bounds are poor, and for these problems to achieve the degree of accuracy which the bounds typically get. This reinforces the results in [82] and emphasizes the necessity for good bounds.

To further illustrate this point consider the plot in figure 7.5. This plot shows a mixed μ computation for a problem with 4 real and 1 complex scalar uncertainties, where the initial bounds were quite poor (85% relative gap as opposed to a typical level of less than 20%). We have plotted the current upper and lower bounds for the problem versus the number of branches, so that the progress of the Branch and Bound scheme on the problem can be seen. It is readily apparent that initially quite rapid progress is made so that in only 29 steps the new bounds are within 20%. However it is also apparent that the progress of the scheme slows quite dramatically after this point, so that achieving greater levels of accuracy requires substantially more computational effort, and rapidly becomes impractical.

The preliminary study of the use of cheap versus sophisticated bounds, which we discussed earlier, employed Branch and Bound schemes using methods from the extreme ends of the spectrum (see [82]). In other words the best currently available bounds (which are quite computationally intensive) were compared to some very crude bounds (which are cheap to compute), when employed in a Branch and Bounds scheme. The overwhelming conclusion was in favor of the more sophisticated bounds.

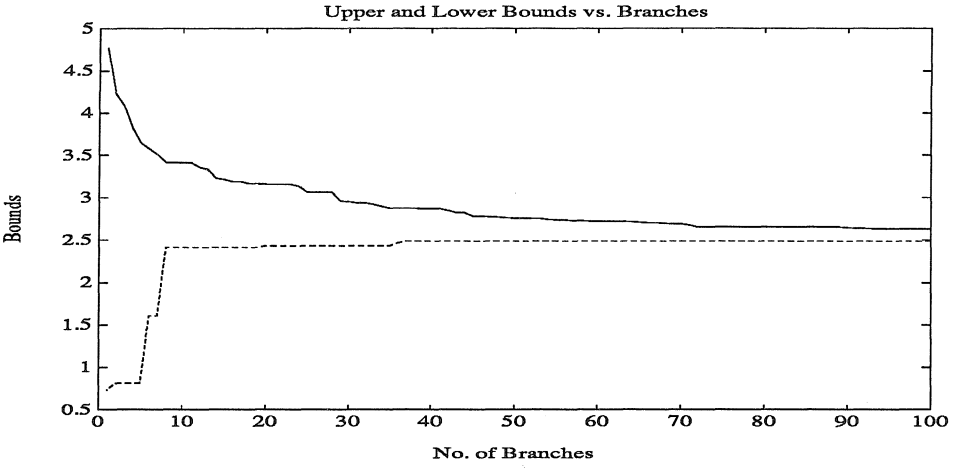


Figure 7.5: Progress of Branch and Bound for a hard problem

In order to examine this question more deeply we compared the use of the best bounds we had available to the next best we could use in a Branch and Bound scheme. The results are plotted in figure 7.6. The left-hand plot was generated using a Branch and Bound scheme employing the bounds previously discussed. We will refer to this as scheme C. The results in the right-hand plot came from a scheme employing the same lower bound, and an upper bound obtained by covering the real parameters with complex ones, and then evaluating the complex μ upper bound. Essentially this amounts to enforcing the choice $G = 0_n$ in (4.34), and so this bound is a little cheaper to compute, but not quite as good, as (4.34). We will refer to this scheme as scheme D. The results are shown for a series of mixed μ problems with 4 real and 1 complex scalar uncertainties. We have plotted the relative gap between the bounds versus the number of branches on a log-log scale. Thus we see the progress of the Branch and Bound schemes with time, and for clarity a number of tolerance levels between the bounds are labeled. Note that for scheme C all the problems reached tolerances of 10% within 6 branches whereas for scheme D several problems failed to reach 10% within the allowed 100 branches. Furthermore the typical performance for scheme

D can be clearly seen to be inferior to scheme C. It is clear that even this level of reduction in the quality of the bounds markedly affects the performance of the overall scheme. Thus we are lead to conclude once more that the performance of the bounds is crucial to the performance of the overall scheme, and that for a high performance Branch and Bound scheme it is important to use the best bounds available.

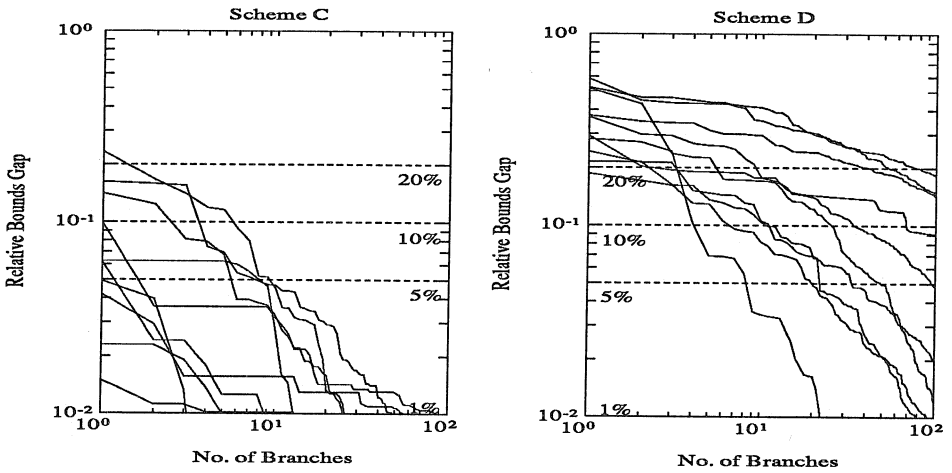


Figure 7.6: Comparison of Branch and Bound schemes

Note that this represents a somewhat different philosophy of Branch and Bound than the one suggested in [2], for example. Here we are performing a relatively small number of quite expensive branches, as opposed to a large number of cheap ones. It should be noted that the approach in [2] might actually work better on small (< 10 parameters) problems. We believe that the scheme outlined here constitutes a practical algorithm for the mixed μ problem, with reasonable accuracy and modest computational requirements for problems of medium size (< 100 parameters) that are of engineering interest. This is despite the fact that the mixed μ problem appears to have inherently combinatoric worst-case behavior.

Chapter 8

The Rank One Mixed μ Problem

We have already discussed in some detail the computational complexity of the mixed μ problem. In particular we have seen that the general mixed μ problem is NP hard, and this strongly suggests that the exact solution of the general problem is computationally intractable, except for small problems. The subsequent work undertaken in chapters 6 and 7 was aimed at developing practical computation schemes for this problem, which give approximate answers with reasonable amounts of computation.

In this chapter we want to consider not the general problem, but a particular special case of this problem: the rank one mixed μ problem. The rank one mixed μ problem is that of computing $\mu_{\mathcal{K}}(M)$ for $M = uv^*$ with $u, v \in \mathbb{C}^n$, and no additional assumptions about the block structure, $\mathcal{K}$. Recalling the discussion from section 3.3.1, this is a case of imposing "structure on M" to simplify the problem, whilst retaining the general uncertainty description $\mathcal{K}$. Note that imposing the condition that M be a dyad is a severe restriction, and this will greatly limit the applicability of this analysis to real engineering applications.

The reason for our interest in this particular problem is that it turns out that a special case of this problem is equivalent to the so called "affine parameter variation" problem (for a polynomial with perturbed coefficients) which has been examined in detail in the literature, and for which several celebrated "Kharitonov-type" results

have been proven. These results provide *exact* robust stability tests for such problems, with respect to real parametric uncertainty (see [71] and the references therein). It will be seen that this special problem does indeed avoid the NP hardness issues of the general problem, and by examining the rank one μ problem directly we are able to obtain a complete (and exact) solution, in terms of quantities which are easily computed.

Moreover we will show that the upper bound is guaranteed to achieve μ for rank one problems. This means that in fact the standard analysis techniques for the general problem, from chapters 4 and 6, exploit all the information that we have about this special problem, and also yield exact results.

8.1 "Kharitonov-Type" Analysis

Before embarking on a study of the rank one mixed μ problem, we first place this problem in context, by considering the "affine parameter variation" problem, for a polynomial with perturbed coefficients. This formulation of the "affine parameter variation" problem is fairly standard, and is taken from [61].

Consider a real monic polynomial in the complex variable s, whose coefficients are affine functions of a real vector of uncertainties, $k \in \mathbb{R}^m$

$$p(s,k) = s^n + a_1(k)s^{n-1} + a_2(k)s^{n-2} + \quad \ldots \quad + a_n(k) \tag{8.1}$$

where $a_i(k)$ for $i = 1, \ldots, n$ are affine functions of k, i. e., there exists $F \in \mathbb{R}^{n \times m}$ and $g \in \mathbb{R}^m$ such that

$$a(k) \doteq [a_1(k) \quad a_2(k) \quad \ldots \quad a_n(k)]^T = Fk + g. \tag{8.2}$$

Thus we can rewrite this set of polynomials as

$$p(s,k) = s^n + [s^{n-1} \quad s^{n-2} \quad \ldots \quad 1](Fk + g). \tag{8.3}$$

Since this polynomial will typically be the closed loop characteristic polynomial of some uncertain system, we will say that it is stable if it has all its roots in the open

left half-plane. We assume nominal stability, i. e., $p(s,0)$ is stable. Thus in order to check robust stability we can show by simple continuity arguments that it suffices to check that the polynomial has no roots on the imaginary axis for any k. Assume in the following analysis that we are considering points on the imaginary axis, i. e., $s = \mathbf{j}w$ where $w \in \mathbb{R}$. Note that since we have nominal stability $p(s,0) \neq 0$ for all $s = \mathbf{j}w$. Thus we have

$$p(s,0) = s^n + [s^{n-1} \quad s^{n-2} \quad \dots \quad 1]g \quad \neq \quad 0. \tag{8.4}$$

Now note that we have a root of the uncertain polynomial on the imaginary axis if and only if for some $s = \mathbf{j}w$ and some $k \in \mathbb{R}^m$, $p(s,k) = 0$. This can be restated as

$$\begin{aligned} p(s,k) = 0 \quad &\longleftrightarrow \quad [s^{n-1} \quad s^{n-2} \quad \dots \quad 1]Fk = -s^n - [s^{n-1} \quad s^{n-2} \quad \dots \quad 1]g \\ &\longleftrightarrow \quad v^*k = 1 \\ &\longleftrightarrow \quad 1 - v^*k = 0 \end{aligned} \tag{8.5}$$

where $v \in \mathbb{C}^m$ is given by

$$\begin{aligned} v \quad &= \quad \left(\frac{1}{-s^n - [s^{n-1} \quad s^{n-2} \quad \dots \quad 1]g}\right)^* F^T[s^{n-1} \quad s^{n-2} \quad \dots \quad 1]^* \\ &= \quad \left(\frac{-1}{p(s,0)}\right)^* F^T[s^{n-1} \quad s^{n-2} \quad \dots \quad 1]^*. \end{aligned} \tag{8.6}$$

But now define the quantities $\Delta \in \mathbb{R}^{m \times m}$ and $u \in \mathbb{R}^m$ by

$$\Delta \quad \doteq \quad \mathrm{diag}\,(k_1, \dots, k_m) \tag{8.7}$$

$$u \quad \doteq \quad (1 \quad \dots \quad 1)^T \tag{8.8}$$

and we have that $\Delta u = k$. Thus an equivalent condition for the existence of an imaginary axis root of the uncertain polynomial is given by

$$\begin{aligned} 1 - v^*k = 0 \quad &\longleftrightarrow \quad 1 - v^*\Delta u = 0 \\ &\longleftrightarrow \quad \det\,(I_m - \Delta u v^*) = 0 \\ &\longleftrightarrow \quad \det\,(I_m - \Delta M) = 0 \end{aligned} \tag{8.9}$$

where $M \in \mathbb{C}^{m\times m}$ is the dyad $M = uv^*$. Checking this condition is exactly a rank one μ problem and thus we see that the "affine parameter variation" problem, for a polynomial with perturbed coefficients, is a special case of the rank one μ problem (with only real perturbations).

It is possible to consider a number of different stability problems arising from this set-up, by allowing for different stability regions, and different norms to measure the size of k. We will only be interested in the case where stability is associated with all the roots in the open left half-plane, and we use $|k|_\infty \doteq \max_{i\leq m} |k_i|$ to measure the size of k. In this case we find that we can use the standard definition of μ, and we are required to compute the peak value across frequency of μ, for a transfer matrix which is rank one. The treatment of other norms/regions is discussed in [17].

A number of different stability results have been presented for this type of problem. One of the strongest motivations for pursuing these problems was provided by Kharitonov's celebrated result for "interval polynomials" [37]. This is a special case of the above setting where one further restricts the uncertainty description to be of the form

$$a_i(k) = a_i + b_i k_i \quad \text{for} \quad i = 1, \ldots, m \tag{8.10}$$

where $a_i \in \mathbb{R}$, $b_i \in \mathbb{R}$ and $m = n$. Thus the coefficients of the polynomial are independent of each other, and known only to lie within certain intervals. For this problem it was shown in [37] that one need only check four specific polynomials to establish stability of the whole family. This is clearly a polynomial time computation, and we have restricted the problem sufficiently to beat the NP-hardness of the general problem. In doing so however we have placed quite severe restrictions on the allowable problem class, and so the applicability of the result is rather limited.

If we consider the "affine parameter" case, then it was shown in [11] that it suffices to check stability of the edges of the parameter space, i. e., we may take every element of k except one to be at an extremal value of it's allowed range. Note that this requires checking a combinatoric number of edges, so that even this increase in the generality

of the problem produces dramatic increases in computation. Exact results for this type of problem typically involve checking the vertices or edges of some polytope in the parameter space, and hence involve exponential growth in computation (see [71, 10]). If one is prepared to allow a frequency search then this exponential growth can be avoided [61]. This can also be seen from the framework we will develop here, since these problems can all be tackled as rank one μ problems, which we will see can be easily solved.

Thus we will find that it is possible to develop exact robust stability tests in the μ framework as well for this type of problem. Of course we must note once more that the applicability of this rank one μ analysis is rather limited, and the fact that the general problem is NP hard strongly suggests that results for this case cannot be usefully extended to the general case. This is the reason why the "Kharitonov-type" analysis methods do not extend to the more general "multilinear" or "polynomial" cases (which correspond to more general μ problems), and one is forced to use approximate and/or iterative methods (see [70] for example).

8.2 Equivalence with the Upper Bound

This section is devoted to proving that for a rank one μ problem, the upper bound always achieves μ, regardless of the block structure. A preliminary version of this result was proved in section 5.4.2, where additional assumptions were imposed to make the proof fairly simple. Here we will not make any such assumptions on the problem, and it turns out that this makes the proof substantially more difficult.

Theorem 8.1 *Suppose we have a rank one matrix $M \in \mathbb{C}^{n\times n}$, then for any block structure, $\mathcal{K}$, $\mu_{\mathcal{K}}(M)$ equals its upper bound from theorem 4.5.*

Before tackling the proof of this theorem, we need a few preliminary results.

Lemma 8.1 *Suppose we have a rank one matrix $M = uv^*$, with $u, v \in \mathbb{C}^n$, and a block structure $\mathcal{K}$. Then there exists a sequence of matrices $D^j \in \mathcal{D}_{\mathcal{K}}$ such that the*

following is true:

1. *Define* $u^j \doteq D^j u$, $v^j \doteq (D^j)^{-1} v$ *and the following limits exist:*

$$\begin{aligned} \lim_{j\to\infty} u^j &= \overline{u} \\ \lim_{j\to\infty} v^j &= \overline{v}. \end{aligned} \tag{8.11}$$

2. *Partition* $u, v, \overline{u}, \overline{v}$ *compatibly with the block structure as in (4.7). Then we have that the following is true:*

$$\begin{aligned} \overline{v}_{r_i}^* \overline{u}_{r_i} &= v_{r_i}^* u_{r_i} \quad \text{for} \quad i = 1, \ldots, m_r \\ \overline{v}_{c_i}^* \overline{u}_{c_i} &= v_{c_i}^* u_{c_i} \quad \text{for} \quad i = 1, \ldots, m_c \\ \overline{v}_{C_i}^* \Delta \overline{u}_{C_i} &= v_{C_i}^* \Delta u_{C_i} \quad \text{for} \quad i = 1, \ldots, m_C \quad \text{for any } \Delta \in \mathbb{C}^{k_{m_r+m_c+i} \times k_{m_r+m_c+i}}. \end{aligned} \tag{8.12}$$

3. *With this notation we also have that*

$$\begin{aligned} |\overline{v}_{r_i}^* \overline{u}_{r_i}| = |\overline{v}_{r_i}|^2 = |\overline{u}_{r_i}|^2 \quad &\text{for} \quad i = 1, \ldots, m_r \\ |\overline{v}_{c_i}^* \overline{u}_{c_i}| = |\overline{v}_{c_i}|^2 = |\overline{u}_{c_i}|^2 \quad &\text{for} \quad i = 1, \ldots, m_c \\ |v_{C_i}||u_{C_i}| = |\overline{v}_{C_i}||\overline{u}_{C_i}| = |\overline{v}_{C_i}|^2 = |\overline{u}_{C_i}|^2 \quad &\text{for} \quad i = 1, \ldots, m_C. \end{aligned} \tag{8.13}$$

4. *Define* $M^j \doteq u^j v^{j*}$ *and the following limit exists*

$$\lim_{j\to\infty} M^j = \lim_{j\to\infty} D^j M (D^j)^{-1} = \overline{M}. \tag{8.14}$$

5. *We have* $\overline{M} = \overline{u}\,\overline{v}^*$, *with* $\mu_{\mathcal{K}}(\overline{M}) = \mu_{\mathcal{K}}(M)$.

Proof: Consider first the repeated real scalar blocks, i.e. , u_{r_i}, v_{r_i} for $i = 1, \ldots, m_r$. Suppose first that for some i in this range we have that $|v_{r_i}^* u_{r_i}| \neq 0$. Then we have

$$v_{r_i}^* u_{r_i} = \gamma e^{\mathbf{j}\theta}$$

for some $\gamma, \theta \in \mathbb{R}$ with $\gamma > 0$. Thus we have that

$$v_{r_i}^* (e^{-\mathbf{j}\theta} u_{r_i}) = \gamma > 0$$

and so by lemma 4.3 we can conclude that there exists $D_i = D_i^* > 0$ such that

$$v_{r_i} = e^{-\mathrm{j}\theta}(D_i)^2 u_{r_i}.$$

Choose $D_i^j = D_i$ for all j and so for this block we have for all j

$$|v_{r_i}^j| = |(D_i)^{-1} v_{r_i}| = |e^{-\mathrm{j}\theta} D_i u_{r_i}| = |D_i u_{r_i}| = |u_{r_i}^j|$$

and furthermore

$$|v_{r_i}^* u_{r_i}| = |v_{r_i}^* e^{-\mathrm{j}\theta} u_{r_i}| = |v_{r_i}^* (D_i)^{-2} v_{r_i}| = |(D_i)^{-1} v_{r_i}|^2 = |v_{r_i}^j|^2.$$

Finally note that

$$v_{r_i}^{j*} u_{r_i}^j = v_{r_i}^* (D_i)^{-1} D_i u_{r_i} = v_{r_i}^* u_{r_i}.$$

For this block then properties $1, 2, 3$ all hold.

Now suppose that for some $i \leq m_r$ we have $|v_{r_i}^* u_{r_i}| = 0$, or in other words v_{r_i} orthogonal to u_{r_i} (if we have $|v_{r_i}| = 0$ and/or $|u_{r_i}| = 0$ then see the treatment for the full complex blocks later). Thus we can choose a Hermitian positive definite matrix D_i^j with u_{r_i} as an eigenvector corresponding to the smallest eigenvalue of D_i^j, and v_{r_i} as an eigenvector corresponding to the largest eigenvalue. Now simply choose such D_i^j with

$$\begin{aligned} \overline{\lambda}(D_i^j) \uparrow \infty \quad & \text{as} \quad j \uparrow \infty \\ \lambda_{min}(D_i^j) \downarrow 0 \quad & \text{as} \quad j \uparrow \infty \end{aligned} \tag{8.15}$$

and we have that

$$\begin{aligned} \lim_{j\to\infty} |v_{r_i}^j| \quad &= \quad \lim_{j\to\infty} |(D_i^j)^{-1} v_{r_i}| \quad = \quad 0 \\ \lim_{j\to\infty} |u_{r_i}^j| \quad &= \quad \lim_{j\to\infty} |D_i^j u_{r_i}| \quad = \quad 0 \end{aligned} \tag{8.16}$$

so that properties $1, 2, 3$ hold once more. The treatment for the repeated complex scalar blocks (i.e., u_{c_i}, v_{c_i} for $i = 1, \ldots, m_c$) is identical.

Now consider the full complex blocks, i.e., u_{C_i}, v_{C_i} for $i = 1, \ldots, m_C$. First consider a block where (for some i in the appropriate range) $|v_{C_i}||u_{C_i}| \neq 0$. Choose the scalar d_i as

$$d_i = \sqrt{\frac{|v_{C_i}|}{|u_{C_i}|}} \tag{8.17}$$

and choose $D_i^j = d_i I_{k_{m_r+m_c+i}}$ for all j. Then for this block we have that

$$\begin{aligned} |u_{C_i}^j|^2 &= |d_i u_{C_i}|^2 = (d_i)^2 |u_{C_i}|^2 = \frac{|v_{C_i}|}{|u_{C_i}|} |u_{C_i}|^2 = |u_{C_i}||v_{C_i}| \\ |v_{C_i}^j|^2 &= |\frac{1}{d_i} v_{C_i}|^2 = \frac{1}{(d_i)^2} |v_{C_i}|^2 = \frac{|u_{C_i}|}{|v_{C_i}|} |v_{C_i}|^2 = |u_{C_i}||v_{C_i}| \end{aligned} \tag{8.18}$$

and furthermore for any $\Delta \in \mathbb{C}^{k_{m_r+m_c+i} \times k_{m_r+m_c+i}}$ we have

$$v_{C_i}^* \Delta u_{C_i} = \left(\frac{1}{d_i} v_{C_i}\right)^* \Delta (d_i u_{C_i}) = v_{C_i}^{j*} \Delta u_{C_i}^j.$$

Note that this implies

$$|v_{C_i}^j||u_{C_i}^j| = |v_{C_i}||u_{C_i}|$$

so that properties $1, 2, 3$ all hold for this block.

Suppose now that for some i we have $|v_{C_i}||u_{C_i}| = 0$, and hence $|u_{C_i}| = 0$ and/or $|v_{C_i}| = 0$. If both are zero simply choose $D_i^j = I_{k_{m_r+m_c+i}}$ for all j, and it is easy to check that properties $1, 2, 3$ all (trivially) hold. If not then choose $D_i^j = d_i^j I_{k_{m_r+m_c+i}}$ where d_i^j is chosen to satisfy

$$\begin{aligned} d_i^j &\uparrow \infty \quad \text{as} \quad j \uparrow \infty \quad \text{if} \quad |u_{C_i}| = 0 \\ d_i^j &\downarrow 0 \quad \text{as} \quad j \uparrow \infty \quad \text{if} \quad |v_{C_i}| = 0. \end{aligned} \tag{8.19}$$

In either case we have that

$$\begin{aligned} \lim_{j\to\infty} |v_{C_i}^j| &= \lim_{j\to\infty} |\frac{1}{d_i^j} v_{C_i}| = 0 \\ \lim_{j\to\infty} |u_{C_i}^j| &= \lim_{j\to\infty} |d_i^j u_{C_i}| = 0 \end{aligned} \tag{8.20}$$

so that properties $1, 2, 3$ hold once more. Thus stacking up the blocks we have constructed we obtain our $D^j \in \mathcal{D}_{\mathcal{K}}$ satisfying properties $1, 2, 3$.

Now note that

$$M^j \doteq u^j v^{j*} = D^j u v^* (D^j)^{-1} = D^j M (D^j)^{-1}$$

is a product of convergent sequences, u^j and v^j, and hence converges, so we have property 4. Furthermore by standard properties of limits we have that $\overline{M} = \overline{u}\ \overline{v}^*$. Note finally that employing properties $1, 2, 3, 4$ we obtain that for any $\Delta \in X_{\mathcal{K}}$

$$\begin{aligned}
\det(I_n - \Delta\overline{M}) &= \det(I_n - \Delta\overline{u}\,\overline{v}^*) \\
&= 1 - \overline{v}^* \Delta \overline{u} \\
&= 1 - \sum_{i=1}^{m_r} \delta_i^r \overline{v}_{r_i}^*\ \overline{u}_{r_i} - \sum_{i=1}^{m_c} \delta_i^c \overline{v}_{c_i}^*\ \overline{u}_{c_i} - \sum_{i=1}^{m_C} \overline{v}_{C_i}^* \Delta_i^C \overline{u}_{C_i} \\
&= 1 - \sum_{i=1}^{m_r} \delta_i^r v_{r_i}^*\ u_{r_i} - \sum_{i=1}^{m_c} \delta_i^c v_{c_i}^*\ u_{c_i} - \sum_{i=1}^{m_C} v_{C_i}^* \Delta_i^C u_{C_i} \\
&= \det(I_n - \Delta M) \qquad (8.21)
\end{aligned}$$

and hence $\mu_{\mathcal{K}}(\overline{M}) = \mu_{\mathcal{K}}(M)$, which is property 5. □

Remarks: Note the the sequence of matrices $D^j \in \mathcal{D}_{\mathcal{K}}$ satisfies the following:

$$\begin{aligned}
\lim_{j\to\infty} \overline{\sigma}(D^j M (D^j)^{-1}) &= \overline{\sigma}(\overline{M}) = |\overline{u}||\overline{v}| = |\overline{u}|^2 \\
&= \sum_{i=1}^{m_r} |\overline{u}_{r_i}|^2 + \sum_{i=1}^{m_c} |\overline{u}_{c_i}|^2 + \sum_{i=1}^{m_C} |\overline{u}_{C_i}|^2 \\
&= \sum_{i=1}^{m_r} |\overline{v}_{r_i}^*\ \overline{u}_{r_i}| + \sum_{i=1}^{m_c} |\overline{v}_{c_i}^*\ \overline{u}_{c_i}| + \sum_{i=1}^{m_C} |\overline{v}_{C_i}||\overline{u}_{C_i}| \\
&= \sum_{i=1}^{m_r} |v_{r_i}^*\ u_{r_i}| + \sum_{i=1}^{m_c} |v_{c_i}^*\ u_{c_i}| + \sum_{i=1}^{m_C} |v_{C_i}||u_{C_i}|. \qquad (8.22)
\end{aligned}$$

This final expression is exactly μ for the associated complex μ problem. Therefore this lemma proves that complex μ equals its upper bound for rank one matrices (which we already knew), and in the process we explicitly constructed the sequence of scaling matrices that does the job. Thus the sequence $D^j \in \mathcal{D}_{\mathcal{K}}$ is exactly the optimal scalings from the upper bound of the associated complex μ problem.

Note that this lemma provides us with a μ invariant transformation from M to $\overline{M}$. The point of carrying out this transformation is that property 3 implies that the

vectors $\overline{u}, \overline{v}$ of the dyad $\overline{M}$ are perfectly balanced in the sense that each sub block of $\overline{u}$ has the same norm as the corresponding sub block of $\overline{v}$. Consequently we have that for each sub block of $\overline{u}$ and $\overline{v}$ we either satisfy the non degeneracy assumptions (see section 4.3), or the corresponding sub blocks of $\overline{u}$ and $\overline{v}$ are *both* identically zero.

Theorem 8.2 *Suppose we have $M = uv^*$ with $u, v \in \mathbb{C}^n$, and a compatible block structure $\mathcal{K}$, with $\mu_{\mathcal{K}}(M) > 0$. Partition u, v with respect to this block structure as in (4.7), and assume that for each sub block of u and v we either satisfy the non-degeneracy assumptions, or the corresponding sub blocks of u and v are both identically zero. Then there exist matrices $Q \in \mathcal{Q}_{\mathcal{K}}$ and $\hat{Q} \in \mathcal{Q}_{\mathcal{K}}$ with*

$$\hat{Q}_{ij} = \begin{cases} 1 & \text{if } i = j \leq m_r \text{ and } Q_{ij} = 0 \\ Q_{ij} & \text{otherwise} \end{cases} \tag{8.23}$$

together with matrices $D_R \in \tilde{\mathcal{D}}_{\mathcal{K}}, D_R \geq 0$ and $D_I \in \mathcal{G}_{\mathcal{K}}$, and a real scalar $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$ such that

$$v = e^{\mathbf{j}\psi} \left(D_R Q + \mathbf{j} D_I \hat{Q} \right) u \tag{8.24}$$

with $0 < v^* Q u = \mu_{\mathcal{K}}(M)$.

Proof: The proof of this theorem is essentially an application of the same machinery used to prove theorem 4.2. First suppose that $\mu_{\mathcal{K}}(M) = 1$. Then by theorem 4.1 there exists $Q \in \mathcal{Q}_{\mathcal{K}}$ with $\rho_R(QM) = 1$. Since we may absorb a factor of $\pm$ into Q we may assume that the eigenvalue achieving $\rho_R(QM)$ is positive. But note that this implies

$$0 = \det\left(I_n - QM\right) = \det\left(I_n - Quv^*\right) = 1 - v^* Q u \tag{8.25}$$

and hence $0 < v^* Q u = 1 = \mu_{\mathcal{K}}(M)$.

Since M is rank one, so is QM, and hence QM has at most one non-zero eigenvalue (not repeated). Thus the eigenvalue at one is distinct and furthermore we have that

$$\begin{aligned} QM(Qu) &= Quv^*Qu &&= (Qu) \\ v^*QM &= v^*Quv^* &&= v^*. \end{aligned} \tag{8.26}$$

Since Qu and v are both non-zero this implies that they are the right and left eigenvectors of QM corresponding to the unity eigenvalue, and furthermore they are normalized with $v^*(Qu) = 1$. Since this eigenvalue is distinct we can differentiate it, and applying the machinery of theorem 4.2 to $\hat{M} \doteq QM$ we can derive that the following relations hold

$$\begin{aligned}
\mathrm{Re}\,(e^{\mathbf{j}\psi} q_i^r v_{r_i}^* u_{r_i}) \geq 0 \quad &, \quad i = 1, \ldots, m_r \\
\mathrm{Re}\,(e^{\mathbf{j}\psi} q_i^r v_{r_i}^* u_{r_i}) = 0 \quad &, \quad \text{if } i \leq m_r \text{ and } |q_i^r| < 1 \\
e^{\mathbf{j}\psi} q_i^c v_{c_i}^* u_{c_i} \in [0\ \infty) \quad &, \quad i = 1, \ldots, m_c \\
\mathrm{Re}\,(e^{\mathbf{j}\psi} v_{C_i}^* G_i^C Q_i^C u_{C_i}) \leq 0 \quad &, \quad \text{for all } G_i^C \in \mathbb{C}^{k_{m_r+m_c+i} \times k_{m_r+m_c+i}} \\
&\qquad \text{with } G_i^C + G_i^{C*} \leq 0, \ i = 1, \ldots, m_C
\end{aligned} \tag{8.27}$$

for some $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$. Now we may apply lemmas 4.2 and 4.3 to each sub block to obtain

$$\begin{aligned}
v_{r_i} = e^{\mathbf{j}\psi} e^{\mathbf{j}\theta_i} q_i^r D_i u_{r_i} \quad &, \quad 0 < D_i = D_i^*, \ \theta_i \in [-\frac{\pi}{2}\ \frac{\pi}{2}], \ i \leq m_r \text{ and } |q_i^r| = 1 \\
v_{r_i} = e^{\mathbf{j}\psi} e^{\mathbf{j}\theta_i} q_i^r D_i u_{r_i} \quad &, \quad 0 < D_i = D_i^*, \ \theta_i = \pm\frac{\pi}{2}, \ i \leq m_r \text{ and } 0 < |q_i^r| < 1 \\
v_{r_i} = e^{\mathbf{j}\psi} e^{\mathbf{j}\theta_i} D_i u_{r_i} \quad &, \quad 0 < D_i = D_i^*, \ \theta_i = \pm\frac{\pi}{2}, \ i \leq m_r \text{ and } |q_i^r| = 0 \\
v_{c_i} = e^{\mathbf{j}\psi} q_i^c D_{m_r+i} u_{c_i} \quad &, \quad 0 < D_{m_r+i} = D_{m_r+i}^*, \ i = 1, \ldots, m_c \\
v_{C_i} = e^{\mathbf{j}\psi} d_i Q_i^C u_{C_i} \quad &, \quad 0 < d_i \in \mathbb{R}, \ i = 1, \ldots, m_C.
\end{aligned} \tag{8.28}$$

Note that in order to apply lemmas 4.2 and 4.3 we need to assume that the non-degeneracy assumptions are satisfied for that sub block. However we have assumed at the outset that u, v either satisfy the non-degeneracy assumption for a given sub block, or have both sub blocks of u and v identically zero, in which case the above relationships hold trivially. Applying this argument to v and Qu we get that the above relationships hold for *every* block. The only case where this argument breaks down is for the repeated real scalar blocks with $q_i^r = 0$. For these blocks however we can show that the above relationships hold by a simple geometric argument.

All that remains now is to define the appropriate quantities. Define $\hat{Q} \in \mathcal{Q}_{\mathcal{K}}$ directly from $Q \in \mathcal{Q}_{\mathcal{K}}$ via (8.23). For $i \leq m_r$ note that

$$e^{\mathbf{j}\theta_i} = \cos(\theta_i) + \mathbf{j}\sin(\theta_i)$$

so that we may split each scaling matrix as

$$e^{\mathbf{j}\theta_i} D_i = \cos(\theta_i) D_i + \mathbf{j}\sin(\theta_i) D_i \doteq D_{R_i} + \mathbf{j} D_{I_i} \quad \text{for} \quad i = 1, \ldots, m_r. \tag{8.29}$$

For the complex blocks we simply define

$$\begin{aligned} D_{R_{m_r+i}} &= D_{m_r+i} \quad \text{for} \quad i = 1, \ldots, m_c \\ d_{R_i} &= d_i \quad \text{for} \quad i = 1, \ldots, m_C. \end{aligned} \tag{8.30}$$

Now stack these definitions up to define $D_R \in \tilde{\mathcal{D}}_{\mathcal{K}}$ and $D_I \in \mathcal{G}_{\mathcal{K}}$. Since for each θ_i we have that $\theta_i \in [-\frac{\pi}{2}\ \frac{\pi}{2}]$, it follows that $\cos(\theta_i) \geq 0$ and hence $D_R \geq 0$. It is now easy to verify that with these definitions the relationships in (8.28) may be written in matrix form as $v = e^{\mathbf{j}\psi}\left(D_R Q + \mathbf{j} D_I \hat{Q}\right) u$. This proves the result for $\mu_{\mathcal{K}}(M) = 1$. The result for $\mu_{\mathcal{K}}(M) > 0$ follows immediately from this by simply scaling u so that $\mu_{\mathcal{K}}(M) = 1$, applying the result for $\mu_{\mathcal{K}}(M) = 1$, and then reabsorbing the scaling back into D_R and D_I. $\square$

Remarks: Although it appears at first sight as a rather unmotivated mathematical abstraction, we will see that in fact this alignment condition between v and u is the key to the equivalence between μ and it's upper bound. Note also that, as in theorem 4.2, we have such an alignment for any $Q \in \mathcal{Q}_{\mathcal{K}}$ achieving a *local* maximum of $\rho_R(QM)$ over $Q \in \mathbb{B}X_{\mathcal{K}}$ with $\rho_R(QM) > 0$. This follows since we derived the alignment condition simply from stationarity conditions, and did not use the fact that $\mu_{\mathcal{K}}(M)$ is the *global* maximum at all.

Note that the conditions on u and v assumed in theorem 8.2 are exactly those guaranteed for the transformed vectors in lemma 8.1. Thus by first transforming the dyad as in lemma 8.1 we can show that (for the transformed dyad) we *always* have an

alignment condition as in theorem 8.2, *without* requiring any type of non-degeneracy assumptions (except that $\mu_{\mathcal{K}}(M) > 0$).

We are now in a position to combine these results to prove the main result of this chapter.

Proof of Theorem 8.1: We will consider separately the cases $\mu_{\mathcal{K}}(M) > 0$ and $\mu_{\mathcal{K}}(M) = 0$. First suppose that $\mu_{\mathcal{K}}(M) > 0$ and for the moment further assume that in fact $\mu_{\mathcal{K}}(M) = 1$. Carry out the transformation of lemma 8.1 to define $\overline{u}, \overline{v}, \overline{M}$ with $\mu_{\mathcal{K}}(\overline{M}) = 1$. Now $\overline{u}, \overline{v}, \overline{M}$ satisfy the assumptions of theorem 8.2, so we may apply this theorem to conclude that we have matrices $Q \in \mathcal{Q}_{\mathcal{K}}$ and $\hat{Q} \in \mathcal{Q}_{\mathcal{K}}$ with $\hat{Q}$ as in (8.23), together with matrices $D_R \in \tilde{\mathcal{D}}_{\mathcal{K}}, D_R \geq 0$ and $D_I \in \mathcal{G}_{\mathcal{K}}$, and a real scalar $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$ such that

$$\overline{v} = e^{\mathbf{j}\psi}\left(D_R Q + \mathbf{j} D_I \hat{Q}\right)\overline{u} \tag{8.31}$$

with $\overline{v}^* Q \overline{u} = 1$. Note that this implies

$$\overline{v}^* Q \overline{u} = e^{-\mathbf{j}\psi}\overline{u}^*(Q^* D_R - \mathbf{j}\hat{Q}^* D_I)Qu = 1$$

and hence

$$\overline{u}^* Q^* D_R Q \overline{u} - \mathbf{j}\overline{u}^* \hat{Q}^* D_I Q \overline{u} = e^{\mathbf{j}\psi} = \cos(\psi) + \mathbf{j}\sin(\psi). \tag{8.32}$$

Now we have immediately that $0 \leq \overline{u}^* Q^* D_R Q \overline{u} \in \mathbb{R}$ and furthermore

$$\overline{u}^* \hat{Q}^* D_I Q \overline{u} = \overline{u}^* \hat{Q}^* Q D_I \overline{u} = \overline{u}^* Q^* Q D_I \overline{u} = \overline{u}^* Q^* D_I Q \overline{u}$$

so that $\overline{u}^* \hat{Q}^* D_I Q \overline{u} \in \mathbb{R}$ as well, and hence by comparing real and imaginary parts in (8.32)

$$\cos(\psi) = \overline{u}^* Q^* D_R Q \overline{u} = \overline{u}^* Q^* Q D_R \overline{u} = \overline{u}^* D_R \overline{u}. \tag{8.33}$$

Define $\hat{G} \in \mathcal{G}_{\mathcal{K}}$ as $\hat{G} \doteq -D_I \hat{Q}$, and substituting D_R and $\hat{G}$ into the upper bound expression we obtain

$$\begin{aligned} & \overline{M}^* D_R \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^* \hat{G}) - D_R \\ = \ & \overline{v}\,\overline{u}^* D_R \overline{u}\,\overline{v}^* + \mathbf{j}(-D_I \hat{Q}\overline{u}\,\overline{v}^* + \overline{v}\,\overline{u}^* D_I \hat{Q}) - D_R \\ = \ & -\overline{v}\,\overline{v}^*(\overline{u}^* D_R \overline{u}) - D_R + D_R Q \overline{u}\,\overline{v}^* + \overline{v}\,\overline{u}^* Q^* D_R \end{aligned} \tag{8.34}$$

where we have made use of the substitution $-\mathbf{j}D_I\hat{Q}\overline{u} = -e^{-\mathbf{j}\psi}\overline{v} + D_RQ\overline{u}$ and (8.33). Thus for any $\eta \in \mathbb{C}^n$ we have that

$$\begin{aligned} & \eta^* \left(\overline{M}^* D_R \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - D_R\right)\eta \\ = \;& 2\mathrm{Re}\,((\eta^* D_R Q\overline{u})(\overline{v}^*\eta)) - (\overline{u}^* D_R \overline{u})|\overline{v}^*\eta|^2 - \eta^* D_R \eta \\ = \;& -((\overline{v}^*\eta)Q\overline{u} - \eta)^* D_R ((\overline{v}^*\eta)Q\overline{u} - \eta) \;\leq\; 0. \end{aligned} \tag{8.35}$$

Now note that we get equality (to zero) in (8.35) if and only if

$$(\overline{v}^*\eta)Q\overline{u} - \eta = -x \tag{8.36}$$

where $x \in \mathrm{Ker}\,(D_R)$, because $D_R \geq 0$. But this implies

$$\overline{v}^*\eta = (\overline{v}^* Q\overline{u})(\overline{v}^*\eta) + \overline{v}^* x = \overline{v}^*\eta + \overline{v}^* x$$

and hence $\overline{v}^* x = 0$. From this it is easy to check that

$$\eta = (\overline{v}^*\eta)Q\overline{u} + x \quad \longleftrightarrow \quad \eta = \alpha Q\overline{u} + x \;\text{ for some }\; \alpha \in \mathbb{C}.$$

To summarize, we have shown that

$$\eta^* \left(\overline{M}^* D_R \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - D_R\right)\eta \;\leq\; 0 \tag{8.37}$$

with equality if and only if η is of the form

$$\eta = \alpha Q\overline{u} + x \tag{8.38}$$

for some $\alpha \in \mathbb{C}$ and $x \in \mathbb{C}^n$ with $D_R x = 0$ and $\overline{v}^* x = 0$. Now define $\hat{D} \doteq D_R + tI_n$, where $0 < t \in \mathbb{R}$ will be chosen later, and note that for $t > 0$, $\hat{D} \in \mathcal{D}_{\mathcal{K}}$. Now choose *any* $\epsilon > 0$ and we have that

$$\begin{aligned} \overline{M}^*\hat{D}\overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - (1+\epsilon)\hat{D} \;&=\; \left(\overline{M}^* D_R \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - D_R - \epsilon D_R\right) \\ &+\; t\left(\overline{M}^*\overline{M} - (1+\epsilon)I_n\right) \\ &\doteq\; A + tB. \end{aligned} \tag{8.39}$$

We wish to apply lemma 5.2 to (8.39). Note that since we have

$$\begin{aligned} \overline{M}^* D_R \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - D_R &\leq 0 \\ -\epsilon D_R &\leq 0 \end{aligned}$$

we immediately have that $A \leq 0$, and furthermore $\eta^* A \eta = 0$ if and only if

$$\begin{aligned} \eta^* \left(\overline{M}^* D_R \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - D_R \right) \eta &= 0 \\ \eta^* D_R \eta &= 0. \end{aligned}$$

From our earlier results this implies

$$\eta = \alpha Q \overline{u} + x$$

with $D_R x = 0$, $\overline{v}^* x = 0$, and $\eta^* D_R \eta = 0$. But since $D_R \geq 0$ this implies that $\eta \in \text{Ker}\,(D_R)$, and hence $\alpha Q\overline{u} \in \text{Ker}\,(D_R)$. Now note that this gives

$$0 = (\alpha Q\overline{u})^* D_R (\alpha Q\overline{u}) = |\alpha|^2 \overline{u}^* Q^* D_R Q \overline{u} = |\alpha|^2 \cos(\psi).$$

Since $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$ we have $\cos(\psi) > 0$, and so this expression vanishes if and only if $\alpha = 0$. Thus we obtain $\eta = x$ and so we have

$$\eta^* A \eta = 0 \quad \longleftrightarrow \quad D_R \eta = 0 \ \text{ and } \ \overline{v}^* \eta = 0.$$

Now note that for such η we have

$$\begin{aligned} \eta^* B \eta &= \eta^* \left(\overline{M}^* \overline{M} - (1+\epsilon) I_n \right) \eta \\ &= \eta^* \left(\overline{v}\ \overline{u}^* \overline{u}\ \overline{v}^* - (1+\epsilon) I_n \right) \eta \\ &= -(1+\epsilon)|\eta|^2 \ < \ 0 \ \text{ for } \ \eta \neq 0. \end{aligned} \tag{8.40}$$

Thus applying lemma 5.2 we have that for sufficiently small $t > 0$, $A + tB < 0$ and hence choosing such a t we obtain

$$\overline{M}^* \hat{D} \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - (1+\epsilon)\hat{D} \ < \ 0 \tag{8.41}$$

with $\hat{D} \in \mathcal{D}_{\mathcal{K}}$ and $\hat{G} \in \mathcal{G}_{\mathcal{K}}$. It now remains to unwrap the transformation from $\overline{M}$, to recover the result for M. Recall from lemma 8.1 that we have a sequence $D^j \in \mathcal{D}_{\mathcal{K}}$ such that

$$\lim_{j\to\infty} D^j M (D^j)^{-1} = \overline{M}.$$

Since this implies that we may choose $D^j \in \mathcal{D}_{\mathcal{K}}$ so that $D^j M (D^j)^{-1}$ is arbitrarily close to $\overline{M}$, by continuity we may choose $D^j \in \mathcal{D}_{\mathcal{K}}$ so that

$$(D^j M (D^j)^{-1})^* \hat{D} (D^j M (D^j)^{-1}) + \mathbf{j}(\hat{G}(D^j M (D^j)^{-1}) - (D^j M (D^j)^{-1})^* \hat{G}) - (1+\epsilon)\hat{D} < 0. \tag{8.42}$$

Since $D^j > 0$ we may multiply both sides of this expression by D^j without affecting the definiteness to yield

$$M^*(D^j \hat{D} D^j) M + \mathbf{j}((D^j \hat{G} D^j) M - M^*(D^j \hat{G} D^j) - (1+\epsilon)(D^j \hat{D} D^j) \;<\; 0 \tag{8.43}$$

so that defining $D \doteq D^j \hat{D} D^j$ and $G \doteq D^j \hat{G} D^j$ we have

$$M^* D M + \mathbf{j}(GM - M^* G) - (1+\epsilon) D \;<\; 0 \tag{8.44}$$

with $D \in \mathcal{D}_{\mathcal{K}}$ and $G \in \mathcal{G}_{\mathcal{K}}$. But this shows that the upper bound achieves $\sqrt{1+\epsilon}$ where $\epsilon > 0$ was arbitrary. Thus by taking the infimum we find that the upper bound gives 1, which by assumption equals $\mu_{\mathcal{K}}(M)$. This proves the result for $\mu_{\mathcal{K}}(M) = 1$. The extension to $\mu_{\mathcal{K}}(M) > 0$ is easy. Suppose $\mu_{\mathcal{K}}(M) = \beta > 0$. Then $\mu_{\mathcal{K}}(\frac{M}{\beta}) = 1$ so we may apply this result to obtain, for any $\hat{\epsilon} > 0$, $\hat{D} \in \mathcal{D}_{\mathcal{K}}$ and $\hat{G} \in \mathcal{G}_{\mathcal{K}}$ such that

$$\frac{M^*}{\beta} \hat{D} \frac{M}{\beta} + \mathbf{j}(\hat{G}\frac{M}{\beta} - \frac{M^*}{\beta}\hat{G}) - (1+\hat{\epsilon})\hat{D} \;<\; 0. \tag{8.45}$$

Now given *any* $\epsilon > 0$ choose $\hat{\epsilon} = (1 + \frac{\epsilon}{\beta})^2 - 1$, and choose $\hat{D}, \hat{G}$ as above. Then this expression gives

$$M^* \hat{D} M + \mathbf{j}((\beta\hat{G})M - M^*(\beta\hat{G})) - (\beta+\epsilon)^2 \hat{D} < 0$$

so that defining $D \doteq \hat{D}$ and $G \doteq \beta\hat{G}$ we obtain

$$M^* D M + \mathbf{j}(GM - M^* G) - (\beta+\epsilon)^2 D \;<\; 0$$

with $D \in \mathcal{D}_{\mathcal{K}}$ and $G \in \mathcal{G}_{\mathcal{K}}$, and hence by the same reasoning as before this implies that the upper bound achieves β, which by assumption equals $\mu_{\mathcal{K}}(M)$. This proves the result for $\mu_{\mathcal{K}}(M) > 0$.

The only case we have left to deal with is when $\mu_{\mathcal{K}}(M) = 0$. This case is not covered by the previous analysis since in this case there is no destabilizing perturbation ($Q \in \mathcal{Q}_{\mathcal{K}}$). Start first by employing the transformation to $\overline{M}$ in lemma 8.1. If $\overline{\sigma}(\overline{M}) = 0$ we are done, so assume not, i.e., $|\overline{u}||\overline{v}| \neq 0$. Note immediately that all the complex blocks must be zero,

$$\begin{aligned} |\overline{u}_{c_i}| = |\overline{v}_{c_i}| = 0 \quad \text{for} \quad i = 1, \ldots, m_c \\ |\overline{u}_{C_i}| = |\overline{v}_{C_i}| = 0 \quad \text{for} \quad i = 1, \ldots, m_C \end{aligned} \tag{8.46}$$

else we could find $Q \in \mathcal{Q}_{\mathcal{K}}$ with $\rho_R(QM) > 0$ simply by choosing all the real perturbations to be zero, and appropriate complex perturbations. Furthermore it is easy to check via a simple geometric argument that for every non-zero repeated real scalar sub block, the quantity $\frac{\overline{v}_{r_i}^* \overline{u}_{r_i}}{|\overline{v}_{r_i}^* \overline{u}_{r_i}|}$ must be the same modulo $\pm$, (for $i = 1, \ldots, m_r$), and this quantity cannot be purely real. If this were not so then once again we could find $Q \in \mathcal{Q}_{\mathcal{K}}$ with $\rho_R(QM) > 0$. These conditions are equivalently expressed as

$$\overline{v}_{r_i}^* \overline{u}_{r_i} = \gamma_i e^{-\mathbf{j}\psi} e^{-\mathbf{j}\theta_i} \text{ for each } i \leq m_r \text{ with } |\overline{v}_{r_i}^* \overline{u}_{r_i}| \neq 0 \tag{8.47}$$

where for each i we have $\theta_i = \pm\frac{\pi}{2}$, $0 < \gamma_i \in \mathbb{R}$, and $\psi \in (-\frac{\pi}{2}\ \frac{\pi}{2})$. Thus applying lemma 4.3 to each of these blocks we have

$$\overline{v}_{r_i} = e^{\mathbf{j}\psi} e^{\mathbf{j}\theta_i} D_i \overline{u}_{r_i} \text{ for each } i \leq m_r \text{ with } |\overline{v}_{r_i}^* \overline{u}_{r_i}| \neq 0 \tag{8.48}$$

where for each i we have $0 < D_i = D_i^*$. Stacking these relationships up we obtain $D_I \in \mathcal{G}_{\mathcal{K}}$ such that

$$\overline{v} = \mathbf{j} e^{\mathbf{j}\psi} D_I \overline{u}. \tag{8.49}$$

Note that for the zero sub blocks of $\overline{u}$ and $\overline{v}$ this relationship holds trivially, and for the non-zero sub blocks it follows from (8.48). Now choose $D_R = 0_n$ and $\hat{G} = -D_I$.

Note that $D_R \in \mathcal{D}_\mathcal{K}, D_R \geq 0$, $\hat{G} \in \mathcal{G}_\mathcal{K}$, and for *any* $0 < \epsilon \in \mathbb{R}$ we have

$$\begin{aligned} & \overline{M}^* D_R \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - \epsilon D_R \\ = \; & \mathbf{j}(-D_I \overline{u}\, \overline{v}^* + \overline{v}\, \overline{u}^* D_I) \\ = \; & -2\cos(\psi)\overline{v}\, \overline{v}^* \quad \leq \quad 0 \end{aligned} \tag{8.50}$$

where we have used the substitution $\mathbf{j} D_I \overline{u} = e^{-\mathbf{j}\psi}\overline{v}$. Now define $\hat{D} \doteq D_R + tI_n = tI_n$, where $0 < t \in \mathbb{R}$ will be chosen later, and note that for $t > 0$, $\hat{D} \in \mathcal{D}_\mathcal{K}$. Then we have that

$$\begin{aligned} \overline{M}^* \hat{D} \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - \epsilon\hat{D} \; &= \; \left(\overline{M}^* D_R \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - \epsilon D_R\right) \\ &+ \; t\left(\overline{M}^*\overline{M} - \epsilon I_n\right) \\ &\doteq \; A + tB. \end{aligned} \tag{8.51}$$

By our earlier results $A \leq 0$ and furthermore we have

$$\eta^* A \eta = -2\cos(\psi)|\overline{v}^*\eta|^2 \quad \leq \quad 0 \tag{8.52}$$

with equality if and only if $\overline{v}^*\eta = 0$ (since $\cos(\psi) > 0$). For such η we have that

$$\begin{aligned} \eta^* B \eta \; &= \; \eta^*\left(\overline{M}^*\overline{M} - \epsilon I_n\right)\eta \\ &= \; \eta^*\left(\overline{v}\, \overline{u}^* \overline{u}\, \overline{v}^* - \epsilon I_n\right)\eta \\ &= \; -\epsilon|\eta|^2 \quad < \quad 0 \text{ for } \eta \neq 0. \end{aligned} \tag{8.53}$$

Thus applying lemma 5.2 we obtain that $A + tB < 0$ for sufficiently small $t > 0$. Choosing such a t we have that

$$\overline{M}^* \hat{D} \overline{M} + \mathbf{j}(\hat{G}\overline{M} - \overline{M}^*\hat{G}) - \epsilon\hat{D} \quad < \quad 0 \tag{8.54}$$

with $\hat{D} \in \mathcal{D}_\mathcal{K}$ and $\hat{G} \in \mathcal{G}_\mathcal{K}$. Now we can unwrap the transformation in lemma 8.1 exactly as before to obtain $D \in \mathcal{D}_\mathcal{K}$ and $G \in \mathcal{G}_\mathcal{K}$ satisfying

$$M^* D M + \mathbf{j}(GM - M^*G) - \epsilon D \quad < \quad 0. \tag{8.55}$$

But this shows that the upper bound achieves $\sqrt{\epsilon}$ where $\epsilon > 0$ was arbitrary. Thus by taking the infimum we have that the upper bound achieves 0 which by assumption equals $\mu_{\mathcal{K}}(M)$. This concludes the $\mu_{\mathcal{K}}(M) = 0$ case, and combining this with our earlier result for $\mu_{\mathcal{K}}(M) > 0$ the theorem is proved. □

8.3 Additional Properties

As the reader was warned, this proof is substantially more involved than the one given in section 5.4.2, where additional simplifying assumptions were made. Note however that this is a constructive proof, and so it actually gives us quite a bit more information about the rank one problem than we obtained with the earlier result. We are able not only to say that the upper bound achieves μ, but to explicitly construct the D, G scaling matrices that do the job. This allows us to examine the properties of these scaling matrices as a function of the problem data, and to arrive at several interesting conclusions.

First note that the construction of the optimal sequence of D, G scaling matrices was based on employing the "μ-values" construction of section 5.5. Note that theorem 8.2 holds for any of the "μ-values" (see the remarks following theorem 8.2). To be more specific this means that given any $Q \in \mathcal{Q}_{\mathcal{K}}$ achieving a *local* maximum over $Q \in \mathbb{B}X_{\mathcal{K}}$ of $\rho_R(QM)$ with $\rho_R(QM) = \beta > 0$, then we can employ the machinery of lemma 8.1, and theorems 8.1 and 8.2, to construct (a sequence of) $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ verifying that $\mu_{\mathcal{K}}(M) \leq \beta$. But since we already have $\beta = \rho_R(QM) \leq \mu_{\mathcal{K}}(M)$ this gives $\mu_{\mathcal{K}}(M) = \beta$. Thus we find that for the rank one problem any non-zero local maximum of $\rho_R(QM)$ (as defined above) is global. We state this as a theorem.

Theorem 8.3 *Suppose we have $M = uv^*$, with $u, v \in \mathbb{C}^n$, and a compatible block structure $\mathcal{K}$. Further suppose we have $Q \in \mathcal{Q}_{\mathcal{K}}$ such that $\rho_R(QM) = \beta > 0$ is a local maximum of $\rho_R(QM)$ over $Q \in \mathbb{B}X_{\mathcal{K}}$. Then $\beta = \mu_{\mathcal{K}}(M)$.*

This offers further insight into why the rank one problem is easy. For the general problem we do not have any such guarantees about local maxima, and in fact one

can easily construct problems with local maxima that are not global.

In fact we can characterize the solution to the rank one μ problem in terms of this alignment condition.

Theorem 8.4 *Suppose we have $M = uv^*$, with $u, v \in \mathbb{C}^n$ satisfying the non-degeneracy assumptions, and a compatible block structure $\mathcal{K}$. Further suppose we have $Q \in \mathcal{Q}_\mathcal{K}$ with $q_i^r \neq 0$ for $i = 1, \ldots, m_r$ such that $\rho_R(QM) = \beta > 0$. Then we have $\beta = \mu_\mathcal{K}(M)$ if and only if there exists $D \in \hat{\mathcal{D}}_\mathcal{K}$ with $\theta_i = \pm\frac{\pi}{2}$ for $|q_i^r| < 1$ and $\psi \in (-\frac{\pi}{2} \ \frac{\pi}{2})$ such that*

$$v = e^{\mathbf{j}\psi} DQu. \tag{8.56}$$

Proof: ($\rightarrow$) Note that as in the proof of theorem 8.1 we can show that Qu and v are the right and left eigenvectors of QM corresponding to the distinct eigenvalue β, with $v^*Qu = \beta$. Since $\beta = \mu_\mathcal{K}(M)$ we are at a local maximum by assumption and the result follows from theorem 4.2.

($\leftarrow$) Again we have that Qu and v are the right and left eigenvectors of QM corresponding to the distinct eigenvalue β, with $v^*Qu = \beta$. But now for each $i \leq m_r$ we have

$$e^{\mathbf{j}\theta_i} = \cos(\theta_i) + \mathbf{j}\sin(\theta_i)$$

so that defining $D_R \in \mathcal{D}_\mathcal{K}, D_R \geq 0$ and $D_I \in \mathcal{G}_\mathcal{K}$ by

$$\begin{aligned} D_{R_i} = \cos(\theta_i) D_i \ \text{ and } \ D_{I_i} = \sin(\theta_i) D_i \quad &\text{for} \quad i = 1, \ldots, m_r \\ D_{R_{m_r+i}} = D_{m_r+i} \quad &\text{for} \quad i = 1, \ldots, m_c \\ d_{R_{m_r+m_c+i}} = d_{m_r+m_c+i} \quad &\text{for} \quad i = 1, \ldots, m_C \end{aligned}$$

we have $v = e^{\mathbf{j}\psi}(D_R + \mathbf{j}D_I)Qu$ as in theorem 8.2. From this alignment we can construct (a sequence of) $\hat{D} \in \mathcal{D}_\mathcal{K}, \hat{G} \in \mathcal{G}_\mathcal{K}$ showing that $\mu_\mathcal{K}(M) \leq \beta$ (as in theorem 8.1) and hence $\beta = \mu_\mathcal{K}(M)$. $\square$

Note that for this theorem we added some technical assumptions. These can once again be dealt with via the machinery of lemma 8.1, but since the main use of this

theorem is for the improved lower bound power iteration discussed in chapter 7 we do not bother with this added complication. Roughly speaking this theorem says that Q achieves μ if and only if it aligns v and Qu (as above). We will return to this fact once more in section 8.5 when we look at the rank one μ problem from a graphical perspective. From that viewpoint the intuition behind this alignment property becomes clear.

It is interesting to consider the cases when the infimum is not achieved in theorem 4.5. This occurs when the D matrix is tending towards semidefiniteness. Note that since we may arbitrarily scale D and G by any positive real scalar and not affect the definiteness of the expression

$$M^*DM + \mathrm{j}(GM - M^*G) - \alpha D$$

then we could alternately think of this phenomenon as D and/or G blowing up. It is just a matter of the choice of normalization. It is possible that we can have $\lambda_{min}(D) \downarrow 0$ because we fail to satisfy the non-degeneracy conditions on some particular block. This occurs in the complex μ case as well, and is associated with the reducibility of M (see [28] and [57] for more details). Note that in this case we can perform a μ invariant transformation to $\overline{M}$ via lemma 8.1. For this matrix then we have that $\overline{M} = \overline{u}\,\overline{v}^*$ where the blocks of $\overline{u}$ and $\overline{v}$ either satisfy the non-degeneracy assumption or are identically zero. Thus we may delete these zero blocks of $\overline{u}$ and $\overline{v}$, to form reduced vectors $\hat{u}$ and $\hat{v}$. The matrix $\hat{M} \doteq \hat{u}\hat{v}^*$ thus corresponds to a reduction of $\overline{M}$ with the appropriate zero rows and columns deleted. At the same time define a reduced block structure $\hat{\mathcal{K}}$, which is obtained as a reduction of $\mathcal{K}$ by deleting these same blocks. Now we have a reduced matrix $\hat{M}$, together with a reduced block structure $\hat{\mathcal{K}}$ such that $\hat{M} = \hat{u}\hat{v}^*$, where every block of $\hat{u}$ and $\hat{v}$ satisfies the non-degeneracy conditions. Furthermore it is easy to check that

$$\mu_{\hat{\mathcal{K}}}(\hat{M}) = \mu_{\mathcal{K}}(\overline{M}) = \mu_{\mathcal{K}}(M). \tag{8.57}$$

Thus whenever the non-degeneracy conditions are not satisfied for a dyad M, we have

a μ invariant reduction to a smaller dyad $\hat{M}$, which does satisfy the non-degeneracy conditions.

For complex μ problems this issue of non-degeneracy is the whole story as far as whether or not the infimum is achieved in the upper bound, since in the complex case the infimum is always achieved for a dyad satisfying the non-degeneracy conditions. In particular the infimum in the upper bound will be achieved for $\hat{M}$. This can be seen in more detail by examining the proof of lemma 8.1, where it can be seen that if every block satisfies the non-degeneracy assumptions, then we can solve for the $D \in \mathcal{D}_{\mathcal{K}}$ that achieves μ. The mixed case however is somewhat more interesting, since it is possible that the infimum in the upper bound is *not* achieved for a dyad that *does* satisfy the non-degeneracy conditions.

For the mixed case, since we have already dealt with the non-degeneracy conditions, suppose that we satisfy the conditions of theorem 8.4 Then consider the alignment condition

$$v = e^{j\psi} DQu$$

where $D \in \hat{\mathcal{D}}_{\mathcal{K}}$. Note from the definition of D_{R_i} in the proof theorem 8.4

$$D_{R_i} = \cos(\theta_i) D_i$$

that we have $D_{R_i} = 0_{k_i}$ whenever $\cos(\theta_i) = 0$, and $D_{R_i} > 0$ whenever $\cos(\theta_i) > 0$. But since $\theta_i \in [-\frac{\pi}{2} \quad \frac{\pi}{2}]$ we have that $\cos(\theta_i) \geq 0$ with $\cos(\theta_i) = 0$ if and only if $\theta_i = \pm\frac{\pi}{2}$. Thus we have D_{R_i} as positive definite except for those blocks where $\theta_i = \pm\frac{\pi}{2}$. In the proof of theorem 8.1 we can see from (8.37) that if D_R is positive definite then the infimum is achieved (with $D_R \in \mathcal{D}_{\mathcal{K}}$ and $\hat{G} \in \mathcal{G}_{\mathcal{K}}$ from (8.37)). Thus we are guaranteed to achieve the infimum except for cases where we have blocks with $\theta_i = \pm\frac{\pi}{2}$. Note that for internal reals (i. e., $|q_i^r| < 1$) we have $\theta_i = \pm\frac{\pi}{2}$, and for non-internal reals this is generically not the case. Thus we find that, roughly speaking, if we don't need internal reals, then the infimum is achieved.

This result is (almost) the converse to theorem 6.3, which showed that in general if the upper bound equals μ (which it does for rank one problems), and the worst

case perturbation must include internal reals, then the infimum is not achieved. This intimate association between the presence of internal reals in the worst case perturbation and the infimum being achieved in the upper bound is quite enlightening. It indicates that for mixed problems we can expect that the infimum is frequently *not* achieved, and this is a fundamental property of the bound. This is in contrast with the complex case where the infimum is achieved except for a thin set of (reducible) matrices [28,1].

8.4 An Exact Expression

Note that in the previous section we remarked that we can solve for the D, G scaling matrices in the upper bound *without* resorting to actually solving the associated convex optimization problem. In order to do this however we needed to know the value of $\mu_{\mathcal{K}}(M)$, and the associated destabilizing perturbation $Q \in \mathcal{Q}_{\mathcal{K}}$ (if there is one). In this and the following section we will see that in fact we can also obtain both of these quantities in closed form, so that we have a complete solution to the rank one μ problem in closed form. Thus we can easily compute all the relevant quantities *without* ever having to resort to numerical solution of an optimization problem.

The following result is due to Chen et al. [17].

Theorem 8.5 ([17]) *Suppose we have $M = uv^*$, with $u, v \in \mathbb{C}^n$, and a compatible block structure $\mathcal{K}$. Then we have that*

$$\mu_{\mathcal{K}}(M) = \min_{x \in \mathbb{R}} \left(\sum_{i=1}^{m_r} |\mathrm{Re}(v_{r_i}^* u_{r_i}) + x \mathrm{Im}(v_{r_i}^* u_{r_i})| + \left(\sum_{i=1}^{m_c} |v_{c_i}^* u_{c_i}| + \sum_{i=1}^{m_C} |v_{C_i}||u_{C_i}| \right) \sqrt{1+x^2} \right). \tag{8.58}$$

The proof of this theorem uses simple linear algebra techniques, together with the notion of dual spaces and the Hahn-Banach theorem [65]. It is given in full in [17] and [61], and we will not go into any of the details here.

Note that this expression is simply a convex minimization of a continuous function of one variable. Furthermore one can easily show that we may restrict our attention

to a finite region of the real axis so that the "minimum" in the above makes sense. Thus this theorem gives an exact expression for μ in the rank one case, which is simple to compute. In fact Chen et al. have shown that one need not resort to numerical methods to solve this problem, but one may simply pick from a finite number of choices for x, the number of choices growing linearly with m_r. This result will also be apparent from the analysis to follow in section 8.5. Note that in the purely complex case ($m_r = 0$) then (8.58) is trivially solved by $x = 0$ to give

$$\mu_{\mathcal{K}}(M) = \sum_{i=1}^{m_c} |v_{c_i}^* u_{c_i}| + \sum_{i=1}^{m_C} |v_{C_i}||u_{C_i}| \tag{8.59}$$

which is the standard expression for the rank one complex μ problem [56].

It is possible to take Chen's expression (8.58) for the rank one μ problem, and from this to derive the destabilizing perturbation (if there is one). Thus we obtain the worst case $Q \in \mathcal{Q}_{\mathcal{K}}$ in closed form, as we claimed earlier. In doing this construction for $Q \in \mathcal{Q}_{\mathcal{K}}$ one discovers that in fact one need only take at most one of the real perturbations to be internal. Thus we find that we may take the worst case perturbation to be on an edge for a rank one μ problem, even if both real and complex perturbations are included. In fact both of these results also follow from the graphical interpretation in the following section, and so we will not present the details of how to derive them from (8.58).

8.5 A Graphical Interpretation

It is interesting to consider a graphical interpretation of the rank one mixed μ problem, in the complex plane. Suppose that we have $M = uv^*$, with $u, v \in \mathbb{C}^n$, and some compatible block structure $\mathcal{K}$. Then note that for any $\Delta \in \mathbb{B}X_{\mathcal{K}}$ and $0 < \alpha \in \mathbb{R}$ we have that

$$\begin{aligned} \det\left(I_n - \frac{\Delta M}{\alpha}\right) = 0 \quad &\longleftrightarrow \quad \det\left(\alpha I_n - \Delta M\right) = 0 \\ &\longleftrightarrow \quad v^* \Delta u = \alpha \end{aligned} \tag{8.60}$$

$$\longleftrightarrow \quad \sum_{i=1}^{m_r} \delta_i^r v_{r_i}^* u_{r_i} + \sum_{i=1}^{m_c} \delta_i^c v_{c_i}^* u_{c_i} + \sum_{i=1}^{m_C} v_{C_i}^* \Delta_i^C u_{C_i} = \alpha.$$

This equation forms the basis of our graphical interpretation. If we think of the components, $\delta_i^r v_{r_i}^* u_{r_i}$, $\delta_i^c v_{c_i}^* u_{c_i}$, $v_{C_i}^* \Delta_i^C u_{C_i}$, as vectors in the complex plane, then the rank one mixed μ problem simply amounts to choosing δ_i^r, δ_i^c, Δ_i^C so that these vectors add up to a positive real number, which is as large as possible. Note that it is now evident that we do not affect the problem if we throw away the degenerate blocks (with $|v_{r_i}^* u_{r_i}| = 0$, or $|v_{c_i}^* u_{c_i}| = 0$, or $|v_{C_i}||u_{C_i}| = 0$).

First of all it is clear that, since the complex perturbations can have arbitrary phase, we will always wish to choose $|\delta_i^c| = 1$ and $\overline{\sigma}(\Delta_i^C) = 1$ so that these vectors have their maximum length, namely

$$\begin{aligned} \delta_i^c v_{c_i}^* u_{c_i} &= e^{\mathbf{j}\psi_{c_i}} |v_{c_i}^* u_{c_i}| \\ v_{C_i}^* \Delta_i^C u_{C_i} &= e^{\mathbf{j}\psi_{C_i}} |v_{C_i}||u_{C_i}|. \end{aligned} \tag{8.61}$$

Furthermore it is also clear from the geometry of the problem that in fact we will wish to align all these vectors, so as to make one vector of maximal length whose phase we are free to choose. Thus we may take $\psi_{c_i} = \psi$ and $\psi_{C_i} = \psi$ (with appropriate ranges for i), and so we have

$$\sum_{i=1}^{m_c} \delta_i^c v_{c_i}^* u_{c_i} + \sum_{i=1}^{m_C} v_{C_i}^* \Delta_i^C u_{C_i} = e^{\mathbf{j}\psi} \left(\sum_{i=1}^{m_c} |v_{c_i}^* u_{c_i}| + \sum_{i=1}^{m_C} |v_{C_i}||u_{C_i}| \right) \tag{8.62}$$

where we have one free parameter, ψ, left to choose for the complex blocks. Note that it is easy to choose δ_i^c and Δ_i^C so that we satisfy (8.62). With this observation denote $0 \leq L_C \in \mathbb{R}$ as

$$L_C = \sum_{i=1}^{m_c} |v_{c_i}^* u_{c_i}| + \sum_{i=1}^{m_C} |v_{C_i}||u_{C_i}| \tag{8.63}$$

and the rank one mixed μ problem reduces to choosing real numbers $\delta_i^r \in [-1 \;\; 1]$ and $\psi \in [-\pi \;\; \pi]$ so as to maximize $0 < \alpha \in \mathbb{R}$ where

$$\sum_{i=1}^{m_r} \delta_i^r v_{r_i}^* u_{r_i} + e^{\mathbf{j}\psi} L_C = \alpha. \tag{8.64}$$

The solution to this problem can be obtained geometrically, by thinking of (8.64) as a vector sum in the complex plane. See figure 8.1 for an example illustration. Note that having chosen the values of δ_i^r then it is easy to choose ψ so as to maximize α, or alternately to verify that no ψ exists to make the summation in (8.64) add up to a real number. The real question then is how to choose the values of δ_i^r. Consider the following algorithm (some of the key ideas for this approach are from Newlin [49]).

Algorithm 8.1 (Rank One μ Solution)

1. *Choose starting values for the real perturbations as* $\delta_i^r = \mathrm{Sgn}(\mathrm{Re}(v_{r_i}^* u_{r_i}))$. *Then for all* δ_i^r *we have* $\mathrm{Re}(\delta_i^r v_{r_i}^* u_{r_i}) \geq 0$. *Now compute*

$$S = \mathrm{Sgn}\left(\mathrm{Im}\left(\sum_{i=1}^{m_r} \delta_i^r v_{r_i}^* u_{r_i}\right)\right).$$

2. *Rank all the components* $\delta_i^r v_{r_i}^* u_{r_i}$ *by argument, so that the highest rank is assigned to the greatest value of* $\mathrm{Arg}(S\delta_i^r v_{r_i}^* u_{r_i})$.

3. *Consider the highest rank component which has not yet been looked at. Compute* δ_{opt}, *which is the optimal value of this* δ_i^r, *for* $\delta_{opt} \in \mathbb{R}$ *unrestricted in sign or magnitude, and all the other real perturbations fixed.*

4. *If* $\mathrm{Sgn}(\delta_{opt}) = -\mathrm{Sgn}(\delta_i^r)$ *and* $|\delta_{opt}| > 1$ *and not all the components have been looked at, then set* $F = 1$, *else set* $F = 0$. *Now reassign* $\delta_i^r = \max(-1, \min(1, \delta_{opt}))$.

5. *If* $F = 1$ *then go to step 3.*

6. *For these values of* δ_i^r *compute the optimal value of* ψ, *and hence* $\mu_{\mathcal{K}}(M)$ *and the destabilizing perturbation* $Q \in \mathcal{Q}_{\mathcal{K}}$ *(if there is one).*

This algorithm guarantees to compute μ exactly for a rank one mixed μ problem, together with an optimal destabilizing perturbation $Q \in \mathcal{Q}_{\mathcal{K}}$ (if there is one). The reasoning behind this algorithm is simple if one thinks of the problem geometrically. The objective is to make the summation

$$\sum_{i=1}^{m_r} \delta_i^r v_{r_i}^* u_{r_i} + e^{\mathbf{j}\psi} L_C$$

add up to a real number which is as large as possible. First of all one chooses the perturbations, δ_i^r, so that the real parts of each component, $\delta_i^r v_{r_i}^* u_{r_i}$, are nonnegative. Then consider the summation

$$\sum_{i=1}^{m_r} \delta_i^r v_{r_i}^* u_{r_i} \tag{8.65}$$

to which we must add the complex component, $e^{\mathbf{j}\psi} L_C$, so as to make it add up to a real number. Suppose the imaginary part of the summation in (8.65) is nonnegative (symmetric arguments apply for the other case). Then we consider the component $\delta_i^r v_{r_i}^* u_{r_i}$ with the largest argument. This component is ranked the "worst" component in the summation, in the sense that it contributes the most positive imaginary part (which we have too much of) for a given positive real part (which we want). Then we compute the optimal value, δ_{opt}, for this parameter, δ_i^r (with all the other real perturbations fixed), and reassign δ_i^r with the value of δ_{opt}, clipped to the interval $[-1 \quad 1]$. If δ_{opt} is not both of opposite sign to the original value of δ_i^r, and greater than one, then it says that you could not improve the summation by further reducing the imaginary contribution from this component, $\delta_i^r v_{r_i}^* u_{r_i}$. But this component had the "worst" ratio of imaginary to real contribution, so you could not improve by changing any other component, and hence you are done. If the above condition is not met then you could get further improvement with this component so you check the next rank component, until you meet the condition or you have checked them all. In this way you proceed with at most a linear search over the real parameters to obtain the optimal values for all the real perturbations, δ_i^r. Given these, then it is easy to compute the remainder of the solution.

It is easy to verify that the computation of δ_{opt} in step 3 and ψ in step 6 boil down to just simple trigonometry. It can also be verified that we always have $\psi \in (-\frac{\pi}{2} \quad \frac{\pi}{2})$ and denoting by θ_i the angle between $e^{\mathbf{j}\psi} L_C$ and $\delta_i^r v_{r_i}^* u_{r_i}$, then $\theta_i \in [-\frac{\pi}{2} \quad \frac{\pi}{2}]$. Furthermore if $|\delta_i^r| < 1$ then $\theta_i = \pm\frac{\pi}{2}$. But now if we redraw the vector summation in figure 8.1 then we obtain the picture shown in figure 4.2. Thus we see that this is exactly the alignment condition derived in theorems 4.2 and 8.2. From the geometric

viewpoint it is now clear why this alignment must hold at the maximum of $\rho_R(QM)$.

Note that algorithm 8.1 requires at most a search over the real parameters, which grows linearly with m_r. All the computations required can be performed via simple trigonometry, so that this *algorithm* is really a shorthand notation for the closed form solution to the rank one mixed μ problem (which would otherwise be cumbersome to write, involving a "max" over m_r possibilities). Thus we have a closed form solution, with trivial computational requirements, for both $\mu_{\mathcal{K}}(M)$ and the associated $Q \in \mathcal{Q}_{\mathcal{K}}$. As we pointed out in the preceding section, we could also obtain closed form expressions for these quantities from Chen's result (8.58).

The fact that the graphical solution to the rank one μ problem is so simple, really lets us see what is going on with the problem. As an illustration of this we immediately obtain the following theorem.

Theorem 8.6 *Suppose we have $M = uv^*$, with $u, v \in \mathbb{C}^n$, and a compatible block structure $\mathcal{K}$. Then in computing $\mu_{\mathcal{K}}(M)$ it suffices to consider perturbations $\Delta \in X_{\mathcal{K}}$ with at most one of the real variables δ_i^r having $|\delta_i^r| < 1$.*

Proof: Note that algorithm 8.1 starts out by assigning all the real variables at extremal values. It can be seen from steps 4 and 5 that the algorithm quits if ever any variable is reassigned internally. Since algorithm 8.1 guarantees to find $\mu_{\mathcal{K}}(M)$ we find that an optimal destabilizing perturbation can be found with at most one real variable internal. □

This is the mixed μ counterpart of the well known "edge theorem" [11] for the "affine parameter variation" case for a polynomial with perturbed coefficients (see section 8.1). Note that the result holds for pure real or mixed μ problems. Once again the reason for this result is clear when we look at the problem geometrically: if we have more than one variable internal then we can always increase one of them (in magnitude), and compensate with the other so that the summation (8.64) stays real and does not decrease. We simply do this until all but one (or none) of the real variables

is at its extremal value. Note that this "edge result" can also be derived from Chen's expression (8.58) for the rank one μ problem.

In fact if one considers the geometry of the problem, then it is possible to state a slightly stronger version of this "edge result": aside from cases where we have real degenerate blocks (with $|v_{r_i}^* u_{r_i}| = 0$), or real blocks with the same phase modulo $\pm$ (i. e., $\mathrm{Arg}(v_{r_i}^* u_{r_i}) = \mathrm{Arg}(v_{r_j}^* u_{r_j})$ or $\mathrm{Arg}(v_{r_j}^* u_{r_j}) + \pi$ for $i \neq j$), then the *only* optimal destabilizing perturbations are on the edges.

Thinking about the problem graphically makes it easy to construct examples with a particular value for $\mu_{\mathcal{K}}(M)$, and particular properties for the alignment condition at the maximum of $\rho_R(QM)$. To conclude this chapter we present a series of such examples which illustrate certain facts about μ and rank one problems. These facts may not be obviously true (or false) from the definition of μ, but are immediately clear when one considers the graphical interpretation for the corresponding rank one example.

Fact 1: *For a rank one mixed μ problem, it is not generic that the worst case perturbation is on a vertex.* Consider the following example:

$$M = \begin{pmatrix} \mathrm{j} \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} \qquad \Delta = \mathrm{diag}\,(\delta^r, \delta^c)$$

with $\delta^r \in \mathbb{R}$ and $\delta^c \in \mathbb{C}$. Then it is easy to see that the worst case perturbation is $\delta^r = 0, \delta^c = 1$. This has δ^r internal, and this property holds for small perturbations to the problem.

Fact 2: *For a rank one mixed μ problem, it is not generic that the worst case perturbation is not on a vertex.* Consider the following example:

$$M = \begin{pmatrix} 1 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 1 \end{pmatrix} \qquad \Delta = \mathrm{diag}\,(\delta^r, \delta^c)$$

with $\delta^r \in \mathbb{R}$ and $\delta^c \in \mathbb{C}$. Then it is easy to see that the worst case perturbation is $\delta^r = 1, \delta^c = 1$. This has δ^r at a vertex, and this property holds for small perturbations to the problem.

Fact 3: *For a rank one pure real μ problem, it is generic that the worst case perturbation is not on a vertex.* This follows by noting that given any summation as in (8.65), then we can perturb the components by an arbitrarily small amount so that the summation cannot be made purely real with every $\delta_i^r = \pm 1$.

Fact 4: *For a rank one pure real or mixed μ problem, with at least two uncertainty blocks, it is generic that $\mu_{\mathcal{K}}(M) > 0$.* This follows by noting that for any problem with at least two blocks we can perturb the components by an arbitrarily small amount so that the summation (8.64) can be made real and positive.

Fact 5: *One can have μ problems, where the worst case perturbation has all the real variables internal, or even zero.* This follows from the example given in fact 1.

Note that the examples in facts 1 and 2 also illustrate that for rank one problems, the infimum being achieved in the upper bound is not generic, and the infimum not being achieved is also not generic.

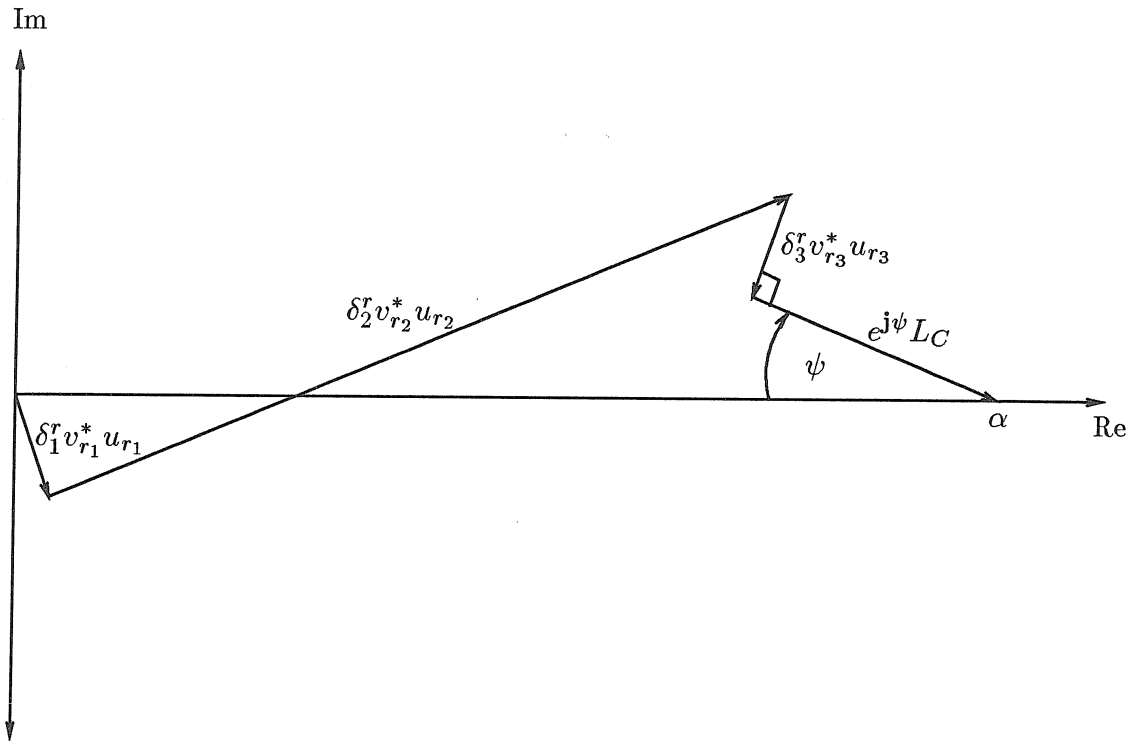


Figure 8.1: Graphical solution of the rank one mixed μ problem

Chapter 9

State Space Factorization Theory

Up to now we have been concerned with the mixed μ analysis question, i.e., the problem of assessing the level of robust stability and/or performance attained by a given (closed loop) system with respect to structured mixed uncertainty. We now wish to consider the mixed μ synthesis question, i.e., how to design a controller so that the resulting closed loop system attains a certain level of robust performance. This goal will be more formally defined and considered in chapter 10, but prior to this we need to develop some State Space factorization machinery, which is the subject of this chapter.

We will develop State Space formulae for specific coprime, inner-outer, and spectral factorizations that will be needed in chapter 10. These factorizations are extensions of standard results (see [85] for example), and as such are of some interest in their own right.

9.1 Basic Concepts

We begin with some basic concepts about transfer matrices. For a more detailed exposition on this subject see [85,30].

Definition 9.1 *Suppose we have a transfer matrix* $\mathbf{D} \in \mathcal{RM}$. *Then* $\mathbf{D}$ *is said to be inner if* $\mathbf{D} \in \mathcal{RH}_\infty$ *and* $\mathbf{D}^\sim\mathbf{D} = I$, *and co-inner if* $\mathbf{D} \in \mathcal{RH}_\infty$ *and* $\mathbf{D}\mathbf{D}^\sim = I$.

If we are interested in square transfer matrices, then these notions can be simplified somewhat.

Lemma 9.1 *Suppose we have a square transfer matrix* $\mathbf{D} \in \mathcal{RM}$. *Then the following conditions are equivalent:*

1. $\mathbf{D}^{\sim}\mathbf{D} = I$.
2. $\mathbf{D}\mathbf{D}^{\sim} = I$.
3. $\mathbf{D}$ *is invertible and* $\mathbf{D}^{-1\sim}\mathbf{D}^{-1} = I$.
4. $\mathbf{D}$ *is invertible and* $\mathbf{D}^{-1}\mathbf{D}^{-1\sim} = I$.

Proof: Follows from straightforward manipulations. □

Lemma 9.2 *Suppose we have a square transfer matrix* $\mathbf{D} \in \mathcal{RM}$. *Then* $\mathbf{D}$ *is inner if and only if* $\mathbf{D}$ *is co-inner.*

Proof: Apply lemma 9.1 to $\mathbf{D} \in \mathcal{RH}_\infty$. □

A closely related concept is that of an all pass transfer matrix.

Definition 9.2 *Given a square transfer matrix* $\mathbf{D} \in \mathcal{RM}$, *then* $\mathbf{D}$ *is said to be all pass if* $\mathbf{D} \in \mathcal{RL}_\infty$ *and* $\mathbf{D}^{\sim}\mathbf{D} = I$.

Note from lemma 9.1 that we could equivalently require $\mathbf{D} \in \mathcal{RL}_\infty$ and $\mathbf{D}\mathbf{D}^{\sim} = I$. For square matrices then an inner (or co-inner) transfer matrix is exactly a stable all pass transfer matrix.

Lemma 9.3 *Suppose we have a square transfer matrix* $\mathbf{D} \in \mathcal{RM}$. *Then* $\mathbf{D}$ *is all pass if and only if* $\mathbf{D}$ *is invertible and* $\mathbf{D}^{-1}$ *is all pass.*

Proof: ($\longrightarrow$) Suppose $\mathbf{D}$ is all pass, and hence $\mathbf{D} \in \mathcal{RL}_\infty$. From lemma 9.1 we immediately find that $\mathbf{D}$ is invertible and $\mathbf{D}^{-1\sim}\mathbf{D}^{-1} = I$. But now this implies $\mathbf{D}^{-1} = \mathbf{D}^{\sim}$, and hence $\mathbf{D}^{-1} \in \mathcal{RL}_\infty$, which together with the earlier result implies $\mathbf{D}^{-1}$ is all pass.

($\longleftarrow$) Suppose $\mathbf{D}$ is invertible and $\mathbf{D}^{-1}$ is all pass. Applying the above proof we obtain $\mathbf{D} = (\mathbf{D}^{-1})^{-1}$ is all pass. $\square$

The term all pass is motivated by the fact that at any given frequency the gain matrix of an all pass transfer matrix is unitary, or, in other words, has unity magnitude gain in all directions. To be more explicit, we have the following well known results (stated without proof).

Lemma 9.4 *Suppose N_1, N_2 are all pass transfer matrices, and $\mathbf{D} \in \mathcal{RM}$. Then for any $w \in \mathbb{R}$ we have*

$$\overline{\sigma}(N_1(\mathrm{j}w)\mathbf{D}(\mathrm{j}w)N_2(\mathrm{j}w)) = \overline{\sigma}(\mathbf{D}(\mathrm{j}w)).$$

Lemma 9.5 *Suppose N_1, N_2 are square inner transfer matrices, and $\mathbf{D} \in \mathcal{RH}_\infty$. Then we have $N_1\mathbf{D}N_2 \in \mathcal{RH}_\infty$ and*

$$\|N_1\mathbf{D}N_2\|_\infty = \|\mathbf{D}\|_\infty.$$

Thus the maximum singular value frequency response of a system is invariant to an all pass transfer matrix (and hence the $\mathcal{H}_\infty$ norm is invariant to a square inner transfer matrix). This invariance is the reason we are interested in these concepts. The μ upper bound is based on computing the maximum singular value of a scaled matrix, and it will be seen in chapter 10 that the synthesis problem can be tackled by employing a scaled $\mathcal{H}_\infty$ norm optimization procedure, which will rely on the above invariance properties.

The remainder of this chapter will make use of the following notions. The term *stable* refers to a system (matrix) with all its poles (eigenvalues) in the open left half-plane, and *marginally stable* refers to a system (matrix) with all its poles (eigenvalues) in the closed left half-plane. The term *antistable* refers to a system (matrix) with all its poles (eigenvalues) in the open right half-plane.

Additionally we will make use of some fairly standard notation and results regarding the solution and properties of Riccati equations. We refer the reader to [21] for a concise review of this area, and [85] for a more detailed exposition.

9.2 Coprime Factorization

Now we wish to develop certain transfer matrix factorizations, the first of which is a coprime factorization. Essentially this amounts to splitting a transfer matrix into stable numerator and denominator matrices, without allowing stable common factors between them. We will use this concept to factor out the antistable part of the original transfer matrix.

Definition 9.3 *Suppose we have two transfer matrices $N, M \in \mathcal{RM}$. Then N, M are said to be right coprime (over $\mathcal{RH}_\infty$) if $N, M \in \mathcal{RH}_\infty$, they have the same number of columns, and $\begin{pmatrix} M \\ N \end{pmatrix}$ is left invertible in $\mathcal{RH}_\infty$.*

Given a transfer matrix $\mathbf{D} \in \mathcal{RM}$, then a right coprime factorization of $\mathbf{D}$ (if it exists) consists of two right coprime transfer matrices N, M such that $\mathbf{D} = NM^{-1}$ (implicitly we have that M is square and invertible). The following theorem addresses the existence of such a factorization.

Theorem 9.1 *Suppose we have a transfer matrix $\mathbf{D} \in \mathcal{RM}$, and (without loss of generality) that a minimal realization for $\mathbf{D}$ is given by*

$$\mathbf{D} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] \tag{9.1}$$

where A_1 is marginally stable and A_2 is antistable. Now defining

$$H = \begin{pmatrix} A_2 & B_2 B_2^T \\ 0 & -A_2^T \end{pmatrix} \tag{9.2}$$

we have that $H \in dom(Ric)$, and so we can define

$$X = Ric(H) \leq 0 \tag{9.3}$$

$$N = \left[\begin{array}{cc|c} A_1 & B_1 B_2^T X & B_1 \\ 0 & A_2 + B_2 B_2^T X & B_2 \\ \hline C_1 & C_2 + D B_2^T X & D \end{array}\right] \tag{9.4}$$

$$M = \left[\begin{array}{c|c} A_2 + B_2 B_2^T X & B_2 \\ \hline B_2^T X & I \end{array}\right]. \tag{9.5}$$

Then we have that $\mathbf{D} = NM^{-1}$, *with* N *marginally stable,* M *inner, and (9.5) a minimal realization of* M. *Furthermore if* $\mathbf{D} \in \mathcal{RL}_\infty$ *then* $N \in \mathcal{RH}_\infty$, *and* N, M *are right coprime over* $\mathcal{RH}_\infty$.

Proof: Note first that H clearly has no imaginary axis eigenvalues since A_2 is anti-stable. Now our assumption of a minimal realization for $\mathbf{D}$ implies that (A_2, B_2) is controllable, and hence $(A_2, B_2 B_2^T)$ is controllable. Since $B_2 B_2^T$ is positive semidefinite, these facts imply that $H \in dom(Ric)$, so $X = Ric(H)$ is well defined, and is a real symmetric matrix satisfying the Riccati equation

$$A_2^T X + X A_2 + X B_2 B_2^T X = 0 \tag{9.6}$$

with $A_2 + B_2 B_2^T X$ stable (see section 7.2 in [30]). Rewriting this we obtain

$$(A_2 + B_2 B_2^T X)^T X + X(A_2 + B_2 B_2^T X) - X B_2 B_2^T X = 0.$$

This is a Lyapunov equation, and so the fact that $A_2 + B_2 B_2^T X$ is stable and we have $-X B_2 B_2^T X \leq 0$ implies that $X \leq 0$. By construction we have that N is marginally stable, and $M \in \mathcal{RH}_\infty$ is square and invertible. Now we have

$$\mathbf{D}M = \left[\begin{array}{ccc|c} A_1 & 0 & B_1 B_2^T X & B_1 \\ 0 & A_2 & B_2 B_2^T X & B_2 \\ 0 & 0 & A_2 + B_2 B_2^T X & B_2 \\ \hline C_1 & C_2 & D B_2^T X & D \end{array}\right].$$

Applying the state transformation

$$T = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & I & I \end{pmatrix}$$

we obtain

$$\mathbb{D}M = \left[\begin{array}{ccc|c} A_1 & B_1B_2^TX & B_1B_2^TX & B_1 \\ 0 & A_2+B_2B_2^TX & B_2B_2^TX & B_2 \\ 0 & 0 & A_2 & 0 \\ \hline C_1 & C_2+DB_2^TX & DB_2^TX & D \end{array}\right] = \left[\begin{array}{cc|c} A_1 & B_1B_2^TX & B_1 \\ 0 & A_2+B_2B_2^TX & B_2 \\ \hline C_1 & C_2+DB_2^TX & D \end{array}\right] = N$$

and hence $\mathbb{D} = NM^{-1}$. Furthermore note that

$$\begin{aligned} M^{\sim}M &= \left[\begin{array}{cc|c} -A_2^T-XB_2B_2^T & -XB_2B_2^TX & -XB_2 \\ 0 & A_2+B_2B_2^TX & B_2 \\ \hline B_2^T & B_2^TX & I \end{array}\right] \\ &= \left[\begin{array}{cc|c} -A_2^T-XB_2B_2^T & A_2^TX+XA_2+XB_2B_2^TX & 0 \\ 0 & A_2+B_2B_2^TX & B_2 \\ \hline B_2^T & 0 & I \end{array}\right] \\ &= \left[\begin{array}{cc|c} -A_2^T-XB_2B_2^T & 0 & 0 \\ 0 & A_2+B_2B_2^TX & B_2 \\ \hline B_2^T & 0 & I \end{array}\right] = I \end{aligned}$$

(where we made use of the state transformation $T = \begin{pmatrix} I & -X \\ 0 & I \end{pmatrix}$ and the Riccati equation (9.6)) so that M is inner. Now we know that (A_2, B_2) is controllable, A_2 is antistable, and $A_2 + B_2B_2^TX$ is stable. Applying the PBH tests to (9.5), and using these three facts, it is easy to show that $(A_2 + B_2B_2^TX, B_2)$ is controllable and $(B_2^TX, A_2 + B_2B_2^TX)$ is observable, so that (9.5) is a minimal realization of M. Finally we note that if $\mathbb{D} \in \mathcal{RL}_\infty$ then A_1 is stable and hence $N \in \mathcal{RH}_\infty$. The fact that N, M are then right coprime over $\mathcal{RH}_\infty$ may be verified by choosing L such that $A + LC$ is stable (which we can do since (C, A) is observable) and checking that

$$P = \left[\begin{array}{cc|cc} A+LC & & -B-LD & L \\ \hline 0 & B_2^TX & I & 0 \end{array}\right] \in \mathcal{RH}_\infty \tag{9.7}$$

is a left inverse of $\begin{pmatrix} M \\ N \end{pmatrix}$. □

This theorem provides us with State Space formulae to construct a particular factorization, namely one where the denominator matrix M is inner. When $\mathbf{D} \in \mathcal{RL}_\infty$ this is a right coprime factorization, and in fact it is easy to show that such a factorization exists if and only if $\mathbf{D} \in \mathcal{RL}_\infty$. Note from the proof that for $\mathbf{D} \notin \mathcal{RL}_\infty$ A_1 has imaginary axis eigenvalues and so $N \notin \mathcal{RH}_\infty$. However we still have an inner M, and a marginally stable N, such that $\mathbf{D} = NM^{-1}$, with $P \in \mathcal{RH}_\infty$ (see (9.7)) a left inverse of $\begin{pmatrix} M \\ N \end{pmatrix}$.

The existence of a right coprime factorization for $\mathbf{D} \in \mathcal{RL}_\infty$ is well known. What is new here is that the factorization in theorem 9.1 may be carried out for arbitrary $\mathbf{D} \in \mathcal{RM}$ and, more importantly, that we only solve a Riccati equation of the same size as A_2, i.e., the number of antistable poles in $\mathbf{D}$. Thus we only invert out the antistable part of $\mathbf{D}$ and hence obtain a minimal realization for M. This offers clear numerical advantages over implementing the standard formulae for coprime factorization (see [85] for example), which would require solving a Riccati equation of the same dimension as A, and would result in unobservable modes in the realization for M. The numerical advantages of the formulation presented here are particularly important when there is pole symmetry about the imaginary axis, which occurs, for example, when $\mathbf{D}$ is self-adjoint. In that case the Hamiltonian matrix from the standard formulae (see [85,30] for example)

$$H = \begin{pmatrix} A & BB^T \\ 0 & -A^T \end{pmatrix}$$

has every eigenvalue repeated. A standard approach to solving the Riccati equation relies on computing the eigenspaces of the Hamiltonian matrix H. It is well known that a non-trivial Jordan form can present severe problems for numerical computation of the eigenvalue decomposition, so that a straightforward implementation of the standard formulae for coprime factorization could run into numerical difficulties on this type of problem.

The analogous definitions and results for left coprime factorizations follow by duality.

Definition 9.4 *Suppose we have two transfer matrices $\tilde{N}, \tilde{M} \in \mathcal{RM}$. Then $\tilde{N}, \tilde{M}$ are said to be left coprime (over $\mathcal{RH}_\infty$) if $\tilde{N}, \tilde{M} \in \mathcal{RH}_\infty$, they have the same number of rows, and $\begin{pmatrix} \tilde{M} & \tilde{N} \end{pmatrix}$ is right invertible in $\mathcal{RH}_\infty$.*

Given a transfer matrix $\mathbf{D} \in \mathcal{RM}$, then a left coprime factorization of $\mathbf{D}$ (if it exists) consists of two left coprime transfer matrices $\tilde{N}, \tilde{M}$ such that $\mathbf{D} = \tilde{M}^{-1}\tilde{N}$.

Theorem 9.2 *Suppose we have a transfer matrix $\mathbf{D} \in \mathcal{RM}$, and (without loss of generality) that a minimal realization for $\mathbf{D}$ is given by*

$$\mathbf{D} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] \tag{9.8}$$

where A_1 is marginally stable and A_2 is antistable. Now defining

$$H = \begin{pmatrix} A_2^T & C_2^T C_2 \\ 0 & -A_2 \end{pmatrix} \tag{9.9}$$

we have that $H \in dom(Ric)$, and so we can define

$$X = Ric(H) \leq 0 \tag{9.10}$$

$$\tilde{N} = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ XC_2^TC_1 & A_2 + XC_2^TC_2 & B_2 + XC_2^TD \\ \hline C_1 & C_2 & D \end{array}\right] \tag{9.11}$$

$$\tilde{M} = \left[\begin{array}{c|c} A_2 + XC_2^TC_2 & XC_2^T \\ \hline C_2 & I \end{array}\right]. \tag{9.12}$$

Then we have that $\mathbf{D} = \tilde{M}^{-1}\tilde{N}$, with $\tilde{N}$ marginally stable, $\tilde{M}$ inner, and (9.12) a minimal realization of $\tilde{M}$. Furthermore if $\mathbf{D} \in \mathcal{RL}_\infty$ then $\tilde{N} \in \mathcal{RH}_\infty$, and $\tilde{N}, \tilde{M}$ are left coprime over $\mathcal{RH}_\infty$.

Proof: Apply theorem 9.1 to $\mathbf{D}^T$ and the results follow. □

9.3 Inner-Outer Factorization

In this section we will make use of the following notions. The term *minimum phase* refers to a system (matrix) with all its zeros (eigenvalues) in the open left half-plane, and *marginally minimum phase* refers to a system (matrix) with all its zeros (eigenvalues) in the closed left half-plane. The term *anti minimum phase* refers to a system (matrix) with all its zeros (eigenvalues) in the open right half-plane.

We now wish to develop a factorization which will factor out the anti minimum phase part of a transfer matrix. It is useful to consider another special type of transfer matrix.

Definition 9.5 *Suppose we have a transfer matrix* $\mathbf{D} \in \mathcal{RM}$. *Then* $\mathbf{D}$ *is outer if* $\mathbf{D} \in \mathcal{RH}_\infty$ *and* $\mathbf{D}$ *is right invertible in* $\mathcal{RH}_\infty$.

It is easy to show that a square transfer matrix $\mathbf{D} \in \mathcal{RM}$ is outer if and only if $\mathbf{D}, \mathbf{D}^{-1} \in \mathcal{RH}_\infty$. Thus an outer matrix is stable minimum phase, i. e., it has all its poles and zeros in the open left half-plane.

Note that in an analogous fashion we can define a matrix $\mathbf{D} \in \mathcal{RM}$ as co-outer if $\mathbf{D} \in \mathcal{RH}_\infty$ and $\mathbf{D}$ is left invertible in $\mathcal{RH}_\infty$. It is clear that for square transfer matrices $\mathbf{D}$ is outer if and only if $\mathbf{D}$ is co-outer.

If we can factor a matrix $\mathbf{D} \in \mathcal{RM}$ as $\mathbf{D} = \mathbf{D}_i\mathbf{D}_o$, with $\mathbf{D}_i$ inner, and $\mathbf{D}_o$ outer, then this is referred to as an inner-outer factorization of $\mathbf{D}$. Clearly we must have $\mathbf{D} \in \mathcal{RH}_\infty$ for such a factorization to exist, and in fact it can be shown that such a factorization exists for all $\mathbf{D} \in \mathcal{RH}_\infty$ (see [85,30]).

Here we need only concern ourselves with a special case of this factorization, namely when $\mathbf{D}$ is square and invertible.

Theorem 9.3 *Suppose* $\mathbf{D} \in \mathcal{RM}$ *is square, invertible, and marginally stable. Further suppose (without loss of generality) that we have a minimal realization for* $\mathbf{D}^{-1}$

$$\mathbf{D}^{-1} = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] \tag{9.13}$$

with A_1 *marginally stable, and* A_2 *anti-stable, so that defining*

$$\begin{aligned} \hat{A} &= A - BD^{-1}C \\ \hat{B} &= BD^{-1} \\ \hat{C} &= -D^{-1}C \\ \hat{D} &= D^{-1} \end{aligned} \tag{9.14}$$

we have that

$$\mathbf{D} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] = \left[\begin{array}{cc|c} \hat{A} & & \hat{B} \\ \hat{C}_1 & \hat{C}_2 & \hat{D} \end{array}\right] \tag{9.15}$$

is a minimal realization for $\mathbf{D}$. *Now defining*

$$H = \begin{pmatrix} A_2 & B_2 B_2^T \\ 0 & -A_2^T \end{pmatrix} \tag{9.16}$$

we have that $H \in dom(Ric)$, *and so we can define*

$$X = Ric(H) \leq 0 \tag{9.17}$$

$$\mathbf{D}_i = \left[\begin{array}{c|c} A_2 + B_2 B_2^T X & B_2 \\ \hline B_2^T X & I \end{array}\right] \tag{9.18}$$

$$\mathbf{D}_o = \left[\begin{array}{cc|c} \hat{A} & & \hat{B} \\ \hline \hat{C}_1 & \hat{C}_2 - B_2^T X & \hat{D} \end{array}\right]. \tag{9.19}$$

Then we have that $\mathbf{D} = \mathbf{D}_i \mathbf{D}_o$ *with* $\mathbf{D}_i$ *inner, and (9.18) a minimal realization for* $\mathbf{D}_i$, *and* $\mathbf{D}_o$ *marginally stable, marginally minimum phase. Furthermore if* $\mathbf{D} \in \mathcal{RH}_\infty$ *then* $\mathbf{D}_o \in \mathcal{RH}_\infty$ *and if* $\mathbf{D}^{-1} \in \mathcal{RL}_\infty$ *then* $\mathbf{D}_o$ *is (right) invertible in* $\mathcal{RH}_\infty$.

Proof: Apply theorem 9.1 to factor $\mathbf{D}^{-1}$ as $\mathbf{D}^{-1} = NM^{-1}$ with

$$N = \left[\begin{array}{cc|c} A_1 & B_1B_2^TX & B_1 \\ 0 & A_2 + B_2B_2^TX & B_2 \\ \hline C_1 & C_2 + DB_2^TX & D \end{array}\right] \tag{9.20}$$

$$M = \left[\begin{array}{c|c} A_2 + B_2B_2^TX & B_2 \\ \hline B_2^TX & I \end{array}\right] \tag{9.21}$$

where $X = Ric(H) \leq 0$. Note that N is marginally stable, M is inner, and (9.21) is a minimal realization of M. Note also that N is square and invertible (as is M). Now define $\mathbf{D}_i = M$ and $\mathbf{D}_o = N^{-1}$, and we have

$$\mathbf{D} = (\mathbf{D}^{-1})^{-1} = (NM^{-1})^{-1} = MN^{-1} = \mathbf{D}_i\mathbf{D}_o$$

with $\mathbf{D}_i$ inner, (9.18) a minimal realization of $\mathbf{D}_i$, and $\mathbf{D}_o$ marginally minimum phase (since $\mathbf{D}_o^{-1}$ is marginally stable). Now by construction we have

$$\mathbf{D}_o = \left[\begin{array}{cc|c} A_1 & B_1B_2^TX & B_1 \\ 0 & A_2 + B_2B_2^TX & B_2 \\ \hline C_1 & C_2 + DB_2^TX & D \end{array}\right]^{-1} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C}_1 & \hat{C}_2 - B_2^TX \mid \hat{D} \end{array}\right] \tag{9.22}$$

so that $\mathbf{D}_o$ is marginally stable. Finally note that if $\mathbf{D} \in \mathcal{RH}_\infty$ then $\hat{A}$ is stable so $\mathbf{D}_o \in \mathcal{RH}_\infty$, and if $\mathbf{D}^{-1} \in \mathcal{RL}_\infty$ then $N \in \mathcal{RH}_\infty$, so $\mathbf{D}_o$ is (right) invertible in $\mathcal{RH}_\infty$ (with $\mathbf{D}_o^{-1} = N$). □

In the case that $\mathbf{D} \in \mathcal{RH}_\infty, \mathbf{D}^{-1} \in \mathcal{RL}_\infty$ we have that $\mathbf{D}_o$ is outer and hence this is exactly an inner-outer factorization. Note once again that although the inner-outer factorization is well known, we are here only solving a Riccati equation of the same dimension as A_1, i.e., the anti minimum phase zeros of $\mathbf{D}$. Thus by only factoring out the anti minimum phase part of $\mathbf{D}$ we obtain a minimal realization for $\mathbf{D}_i$. This offers numerical advantages over the standard approach, and similar comments to those made in section 9.2 apply.

Once again we may define dual version of this factorization, namely if we can factor a matrix $\mathbf{D} \in \mathcal{RM}$ as $\mathbf{D} = \mathbf{D}_{co}\mathbf{D}_{ci}$, with $\mathbf{D}_{ci}$ co-inner, and $\mathbf{D}_{co}$ co-outer, then

this is referred to as a co-outer-co-inner factorization of $\mathbf{D}$. Again the results follow easily by duality, and we do not include them here

9.4 Combined Factorizations

The machinery developed in the preceding subsections enables us to carry out a combined factorization, which splits off the anti-stable poles and anti minimum phase zeros of $\mathbf{D}$ using all pass functions.

Theorem 9.4 *Suppose we have a square invertible transfer matrix* $\mathbf{D} \in \mathcal{RM}$. *Then* $\mathbf{D}$ *may be factored as*

$$\mathbf{D} = \mathbf{D}_{apa}\mathbf{D}_{aps}\mathbf{D}_{smp} \tag{9.23}$$

where $\mathbf{D}_{apa}$ *is anti stable all pass,* $\mathbf{D}_{aps}$ *is stable all pass, and* $\mathbf{D}_{smp}$ *is marginally stable, marginally minimum phase. Furthermore if* $\mathbf{D} \in \mathcal{RL}_\infty$ *then* $\mathbf{D}_{smp}$ *is stable, and if* $\mathbf{D}^{-1} \in \mathcal{RL}_\infty$ *then* $\mathbf{D}_{smp}$ *is minimum phase.*

Proof: Apply theorem 9.2 to obtain

$$\mathbf{D} = \tilde{M}^{-1}\tilde{N}$$

with $\tilde{M}$ inner, and $\tilde{N}$ marginally stable. Define $\mathbf{D}_{apa} = \tilde{M}^{-1}$ and it is easy to check from (9.12) that $\tilde{M}^{-1}$ is anti-stable. Now apply theorem 9.3 to obtain

$$\tilde{N} = \mathbf{D}_{aps}\mathbf{D}_{smp}$$

with $\mathbf{D}_{aps}$ inner and $\mathbf{D}_{smp}$ marginally stable, marginally minimum phase. The results now follow. □

Note that for $\mathbf{D}, \mathbf{D}^{-1} \in \mathcal{RL}_\infty$ (which will be the case when this result is employed in the mixed μ synthesis problem in chapter 10) this theorem allows us to factor $\mathbf{D}$ such that

$$\mathbf{D} = \mathbf{D}_{ap}\mathbf{D}_{smp} \tag{9.24}$$

where $\mathbf{D}_{ap} = \mathbf{D}_{apa}\mathbf{D}_{aps}$ is an all pass function, and $\mathbf{D}_{smp}$ is stable minimum phase. The State Space formulae to perform this factorization are given in theorems 9.2 and 9.3. In order to compute this factorization we need to solve two Riccati equations. Note that these two Riccati equations have dimension equal to the number of anti-stable poles of $\mathbf{D}$, and the number of anti minimum phase zeros of $\mathbf{D}$ respectively. In this way at each stage we avoid some potential numerical problems that one could encounter with the standard formulae (for which each Riccati equation would have the same dimension as the number of poles (or zeros) of $\mathbf{D}$), and also we factor out a minimal realization for the all pass function.

9.5 Spectral Factorization

Finally we consider spectral factorization, which is essentially a transfer matrix equivalent of the square root. We are concerned with a particular form of spectral factorization, that arises from the mixed μ upper bound.

Theorem 9.5 *Suppose we have a square transfer matrix $G \in \mathcal{RM}$. Further suppose (without loss of generality) that a minimal realization for G is given by*

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]. \tag{9.25}$$

Now define the symmetric positive definite matrices R and Q as

$$\begin{aligned} R &= I + D^T D \\ Q &= I + DD^T \end{aligned} \tag{9.26}$$

and define the Hamiltonian matrix

$$H = \begin{pmatrix} A - BR^{-1}D^T C & -BR^{-1}B^T \\ -C^T Q^{-1} C & -(A - BR^{-1}D^T C)^T \end{pmatrix}. \tag{9.27}$$

Then $H \in dom(Ric)$ so we may define

$$X = Ric(H) > 0 \tag{9.28}$$

$$G_h = \left[\begin{array}{c|c} A - BR^{-1}D^TC - BR^{-1}B^TX & BR^{-\frac{1}{2}} \\ \hline -R^{-1}(D^TC + B^TX) & R^{-\frac{1}{2}} \end{array}\right]. \tag{9.29}$$

Then we have that $G_h \in \mathcal{RH}_\infty$ *and*

$$(I + G^\sim G)^{-1} = G_h G_h^\sim \tag{9.30}$$

with $GG_h \in \mathcal{RH}_\infty$ *given by*

$$GG_h = \left[\begin{array}{c|c} A - BR^{-1}D^TC - BR^{-1}B^TX & BR^{-\frac{1}{2}} \\ \hline Q^{-1}(C - DB^TX) & DR^{-\frac{1}{2}} \end{array}\right]. \tag{9.31}$$

Proof: First we make the definitions

$$\begin{aligned} \hat{A} &= A - BR^{-1}D^TC \\ \hat{B} &= BR^{-\frac{1}{2}} \\ \hat{C} &= Q^{-\frac{1}{2}}C. \end{aligned}$$

It is easy to show that $(\hat{A}, \hat{B})$ is controllable and $(\hat{C}, \hat{A})$ is observable. But now we note that H has the form

$$H = \begin{pmatrix} \hat{A} & -\hat{B}\hat{B}^T \\ -\hat{C}^T\hat{C} & -\hat{A}^T \end{pmatrix}$$

and hence we have that $H \in dom(Ric)$, so that $X = Ric(H)$ is well defined and $X > 0$ (see lemma 3 in [21]) satisfies the Riccati equation

$$(A - BR^{-1}D^TC)^TX + X(A - BR^{-1}D^TC) - XBR^{-1}B^TX + C^TQ^{-1}C = 0 \tag{9.32}$$

with $A - BR^{-1}D^TC - BR^{-1}B^TX$ stable. Thus by construction $G_h \in \mathcal{RH}_\infty$, and G_h is square and invertible. Now define

$$N = G_h^{-1} = \left[\begin{array}{c|c} A & B \\ \hline R^{-\frac{1}{2}}(D^TC + B^TX) & R^{\frac{1}{2}} \end{array}\right] \tag{9.33}$$

and we have

$$N^\sim N = \left[\begin{array}{cc|c} A & 0 & B \\ -(C^TD + XB)R^{-1}(D^TC + B^TX) & -A^T & -C^TD - XB \\ \hline D^TC + B^TX & B^T & R \end{array}\right]$$

$$
\begin{aligned}
&= \left[\begin{array}{cc|c} A & 0 & B \\ -(A^TX + XA + C^TC) & -A^T & -C^TD - XB \\ \hline D^TC + B^TX & B^T & R \end{array}\right] \\
&= \left[\begin{array}{cc|c} A & 0 & B \\ -C^TC & -A^T & -C^TD \\ \hline D^TC & B^T & R \end{array}\right] \\
&= I + G^{\sim}G.
\end{aligned}
$$

Note we have made use of the Riccati equation (9.32) and the similarity transformation $T = \begin{pmatrix} I & 0 \\ -X & I \end{pmatrix}$. Thus we have

$$(I + G^{\sim}G)^{-1} = (N^{\sim}N)^{-1} = N^{-1}N^{-1\sim} = G_hG_h^{\sim}.$$

Finally we have that

$$
\begin{aligned}
GG_h &= \left[\begin{array}{cc|c} A - BR^{-1}D^TC - BR^{-1}B^TX & 0 & BR^{-\frac{1}{2}} \\ -BR^{-1}D^TC - BR^{-1}B^TX & A & BR^{-\frac{1}{2}} \\ \hline -DR^{-1}(D^TC + B^TX) & C & DR^{-\frac{1}{2}} \end{array}\right] \\
&= \left[\begin{array}{cc|c} A & 0 & 0 \\ -BR^{-1}D^TC - BR^{-1}B^TX & A - BR^{-1}D^TC - BR^{-1}B^TX & BR^{-\frac{1}{2}} \\ \hline -DR^{-1}(D^TC + B^TX) & C - DR^{-1}(D^TC + B^TX) & DR^{-\frac{1}{2}} \end{array}\right] \\
&= \left[\begin{array}{c|c} A - BR^{-1}D^TC - BR^{-1}B^TX & BR^{-\frac{1}{2}} \\ \hline Q^{-1}(C - DB^TX) & DR^{-\frac{1}{2}} \end{array}\right]
\end{aligned}
$$

(where we have used the similarity transformation $T = \begin{pmatrix} I & I \\ 0 & I \end{pmatrix}$) and hence $GG_h \in \mathcal{RH}_\infty$. □

Essentially G_h is a stable realization of $(I + G^{\sim}G)^{-\frac{1}{2}}$. Basically in order to compute G_h we first form $(I + G^{\sim}G)$, which is self adjoint, and hence has pole/zero symmetry about the imaginary axis. Then we form N by selecting out half of the poles and zeros of $(I + G^{\sim}G)$, keeping the original poles of G (which are also poles of $(I + G^{\sim}G)$) and

the minimum phase zeros of $(I + G^{\sim}G)$. This requires solving one Riccati equation of the same dimension as A, and we are guaranteed the existence of a solution since $(I + G^{\sim}G)$ cannot have zeros on the imaginary axis. Then $G_h = N^{-1}$ has stable poles, and it's zeros are at the poles of G, so that GG_h has the same poles as G_h, and hence is stable as well.

Note that we do not make *any* restrictions on the pole/zero locations of G for the existence of this factorization. In particular G may have poles and/or zeros on the imaginary axis, and we still obtain stable realizations for G_h and GG_h. This is in contrast with the usual approach to spectral factorization, where some such restriction is typically imposed, and it will be seen in chapter 10 that this extra degree of generality is required for the mixed μ synthesis problem.

Normally one solves two Riccati equations in a spectral factorization, one for the poles, and one for the zeros. This yields a spectral factor which is both stable and minimum phase. In fact we can extend this factorization to obtain a stable (marginally) minimum phase realization for G_h if required, by solving one more Riccati equation, with dimension equal to the number of anti-stable poles of G. Note that the formula (9.33) for N gives

$$G_h^{-1} = \left[\begin{array}{c|c} A & B \\ \hline R^{-\frac{1}{2}}(D^TC + B^TX) & R^{\frac{1}{2}} \end{array}\right]$$

and hence G_h is (marginally) minimum phase if G is (marginally) stable. Now note from earlier (theorem 9.2) that we can first perform a left coprime factorization on G to yield

$$G = \tilde{M}^{-1}\tilde{N}$$

with $\tilde{M}$ inner and $\tilde{N}$ (marginally) stable. But now note that

$$G^{\sim}G = (\tilde{M}^{-1}\tilde{N})^{\sim}(\tilde{M}^{-1}\tilde{N}) = \tilde{N}^{\sim}\tilde{M}^{-1\sim}\tilde{M}^{-1}\tilde{N} = \tilde{N}^{\sim}\tilde{N}$$

so that we can now apply theorem 9.5 to $\tilde{N}$, and obtain the required factorization. In this way we obtain G_h as stable, marginally minimum phase, satisfying

$$(I + G^{\sim}G)^{-1} = G_h G_h^{\sim}$$

and if $G \in \mathcal{RL}_\infty$ then G_h is stable minimum phase. Finally we note that, as before, one can obtain a number of dual results by considering the dual (transpose) system, and we do not include them here.

Chapter 10

Mixed μ Synthesis

The problem of synthesising a controller which is (optimally) robust to structured mixed uncertainty is very difficult, since the associated optimization problem is not convex. Furthermore it seems intuitively clear that the synthesis problem is at least as hard as the analysis problem, which is known to be NP hard. Some exact solutions have been presented for special cases of the synthesis problem (see [62] for example, which reduces the rank one μ synthesis problem to a convex optimization problem), but these are all cases for which the analysis problem also simplifies considerably. As yet there is no globally optimal solution to the general synthesis problem (even in the purely complex case), and no indication that one will be forthcoming in the foreseeable future.

Nevertheless the (complex) μ-synthesis procedure first outlined in [24] has been successfully applied to a large number of engineering problems (see [8] for example). This procedure involves a "D-K iteration" between computing the μ upper bound, and solving for an H_∞ (sub) optimal controller (both of which are convex problems). This procedure, which was developed for μ problems involving *only* complex blocks, does not guarantee to find the globally μ-optimal controller (since the problem is not *jointly* convex in D *and* K), but has often been found to work well in practice.

The approach taken here to the mixed μ synthesis problem is to extend the above

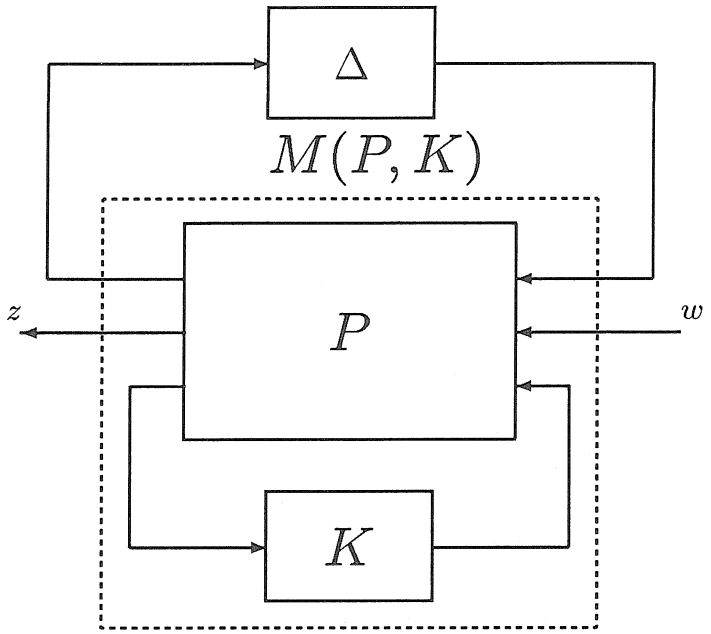


Figure 10.1: Feedback interconnection for μ synthesis

procedure to the mixed case, by exploiting the new analysis tools for the mixed μ upper bound described in the preceding chapters. In this way a "D,G-K iteration" procedure is developed, which finds a controller that stabilizes the nominal system, and attempts to minimize the peak value (across frequency) of mixed μ.

10.1 Complex μ Synthesis and "D-K Iteration"

Prior to tackling the mixed μ synthesis problem, we first include a very brief review of complex μ synthesis. For a more detailed exposition see [24,74].

In order to facilitate study of the problem of synthesising a controller to meet a certain robust performance requirement, consider placing the problem in a standard framework as shown in figure 10.1. The transfer matrix $\mathbf{P}$ is assumed here to contain all the appropriate weights for the problem, as well as the interconnection structure

pertaining to the model of the physical plant. Note that we can define the transfer matrix $\mathbf{M}(\mathbf{P},\mathbf{K})$, by absorbing the controller $\mathbf{K}$ into $\mathbf{P}$ as shown in figure 10.1, and so place the problem in the standard robust performance μ analysis framework of figure 2.2. In light of theorem 2.4 we wish to consider the problem of finding a controller $\mathbf{K}$ achieving

$$\inf_{\mathbf{K}\in\mathcal{K}_\mathcal{S}} \|\mathbf{M}(\mathbf{P},\mathbf{K})\|_\mathcal{K} = \inf_{\mathbf{K}\in\mathcal{K}_\mathcal{S}} \sup_{w\in\mathbb{R}} \mu_\mathcal{K}(\mathbf{M}(\mathbf{P},\mathbf{K})(\mathrm{j}w)) \tag{10.1}$$

where $\mathcal{K}_\mathcal{S}$ denotes all real rational proper controllers that nominally stabilize $\mathbf{P}$ (i. e., render $\mathbf{M}(\mathbf{P},\mathbf{K})$ internally stable). This problem is not tractable however and so we consider the problem given by replacing complex μ by its upper bound in (10.1), namely

$$\inf_{\mathbf{K}\in\mathcal{K}_\mathcal{S}} \sup_{w\in\mathbb{R}} \inf_{D(w)\in\mathcal{D}_\mathcal{K}} \overline{\sigma}(D(w)\mathbf{M}(\mathbf{P},\mathbf{K})(\mathrm{j}w)D^{-1}(w)). \tag{10.2}$$

Note that if we fix $\mathbf{K}$ then the problem of finding $D(w)$ is just the standard complex μ upper bound problem (across frequency) which is a convex problem and can be efficiently solved. If we choose $D(w)$ matrices at a set of frequency points (from the μ upper bound) we can fit a real rational, stable, minimum-phase transfer matrix to them. If we fix this $\mathbf{D}$ transfer matrix then the problem of finding $\mathbf{K}$ reduces to a standard $\mathcal{H}_\infty$ problem, as follows. Consider the feedback interconnection in figure 10.2. Assume once again that we have collected all the relevant quantities in the nominal generalized plant $\hat{P}$, and denote by T_{zw} the closed loop transfer function from w to z. Then the problem of choosing a real rational proper controller $\mathbf{K}$, so as to render this feedback interconnection internally stable, and minimize $\|T_{zw}\|_\infty$, is the standard $\mathcal{H}_\infty$ optimal control problem. This problem is convex and a solution was obtained in [21], for which commercial software is now available [7].

The above approach leads to the following "D-K iteration" scheme, which attempts to find a complex μ optimal controller (and hence the best achievable robust performance for problems with structured complex uncertainty):

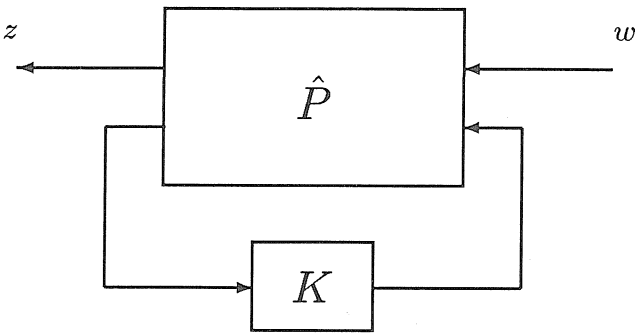


Figure 10.2: Standard framework for $\mathcal{H}_\infty$ optimal control

Procedure 10.1 (D-K Iteration)

1. *Find an initial estimate of the scaling matrices $D(w)$ pointwise across frequency. One possibility is to use the identity matrix at each point.*
2. *Find a State Space realization, $\mathbf{D}$, fitting the pointwise scaling matrices $D(w)$ with a stable minimum phase system (so that $\mathbf{D}$ and $\mathbf{D}^{-1}$ are stable). Augment this with an identity matrix of the appropriate size so that $\mathbf{D}$ is compatible with $\mathbf{P}$. Construct the State Space system $\mathbf{P_D} = \mathbf{DPD}^{-1}$.*
3. *Find the $\mathcal{H}_\infty$ optimal controller $\hat{\mathbf{K}}$ minimizing $\|\mathbf{M}(\mathbf{P_D}, \mathbf{K})\|_\infty$ over all all stabilizing, proper, real rational controllers $\mathbf{K}$.*
4. *Find $\hat{D}(w)$ solving the minimization problem*

$$\inf_{D(w)\in\mathcal{D}_\mathcal{K}} \overline{\sigma}(D(w)\mathbf{M}(\mathbf{P}, \hat{\mathbf{K}})(\mathbf{j}w)D^{-1}(w))$$

   pointwise across frequency.
5. *Compare $\hat{D}(w)$ with the previous estimate $D(w)$. Stop if they are close, else replace $D(w)$ with $\hat{D}(w)$ and return to step 2.*

This iteration (assuming perfect State Space realizations of $D(w)$) is monotonically nonincreasing, so that we are guaranteed convergence to a local minimum of the

problem. Having converged to such a point the controller $\hat{\mathbf{K}}$ from step 3 is the resulting complex μ synthesis controller. Note that although the individual problems (the μ upper bound and $\mathcal{H}_\infty$ optimal control) are convex, the joint problem (μ optimal control) is not convex (see [24] for a counterexample to convexity). Thus the "*D-K* iteration" described above is not guaranteed to converge to the global optimum of the problem. However many designs have been performed using this technique in recent years (see [4,5] for example), and it has usually been found to work well in practice. There are further subtleties to the above procedure, which we will not go into here. For a more detailed review of complex μ synthesis see [74,3].

This "*D-K* iteration" scheme is the basis of complex μ synthesis controller design. The design process consists essentially of the following stages.

Procedure 10.2 (Complex μ Synthesis)

1. *Decide on an appropriate interconnection structure to model the system, including the uncertainty structure against which robustness is desired.*
2. *Choose appropriate weights to reflect the desired performance specifications, and any information known about the uncertainties.*
3. *Implement the above "D-K iteration". Note that this involves deciding on an appropriate frequency range of interest, and selecting the order of the State Space fit in step 2.*

This procedure is by no means mechanical, and a great deal of engineering judgement is still required, particularly in steps 1 and 2. Also, as with any control design technique, the full design process may well involve an iterative application of the above process, together with an evaluation of each resulting controller.

10.2 Mixed μ Synthesis and "*D*,*G*-*K* Iteration"

Note from theorem 6.1 that the mixed μ upper bound can be reformulated in several different ways, and in particular it can be cast as a singular value minimization

problem. Although the LMI form of the upper bound has proven extremely useful for analysis and computational issues, it is the singular value form of the upper bound that we will use to tackle the synthesis problem.

As in the complex case, theorem 2.4 leads us to consider the problem of finding a controller $\mathbf{K}$ achieving

$$\inf_{\mathbf{K}\in\mathcal{K}_{\mathcal{S}}} \|\mathbf{M}(\mathbf{P},\mathbf{K})\|_{\mathcal{K}} = \inf_{\mathbf{K}\in\mathcal{K}_{\mathcal{S}}} \sup_{w\in\mathbb{R}} \mu_{\mathcal{K}}(\mathbf{M}(\mathbf{P},\mathbf{K})(\mathrm{j}w)) \tag{10.3}$$

(where $\mathcal{K}_{\mathcal{S}}$ denotes all real rational proper controllers that nominally stabilize $\mathbb{P}$). Since this problem is not tractable we once again consider the problem given by replacing μ by its upper bound in (10.3). As pointed out earlier there are several equivalent formulations of this upper bound, and we will work with formulation III from theorem 6.1 which leads us to the following synthesis problem.

$$\inf_{\mathbf{K}\in\mathcal{K}_{\mathcal{S}}} \sup_{w\in\mathbb{R}} \inf_{D(w)\in\mathcal{D}_{\mathcal{K}},G(w)\in\mathcal{G}_{\mathcal{K}}} \inf_{\beta(w)\in\mathbb{R},\beta(w)>0} \{\beta(w) : \Gamma \leq 1\} \tag{10.4}$$

where

$$\Gamma = \overline{\sigma}\left(\left(\frac{D(w)\mathbf{M}(\mathbf{P},\mathbf{K})(\mathrm{j}w)D^{-1}(w)}{\beta(w)} - \mathrm{j}G(w)\right)(I+G^2(w))^{-\frac{1}{2}}\right). \tag{10.5}$$

The basic idea behind our approach to mixed μ synthesis is evident from these equations (10.4,10.5). For fixed $\mathbf{K}$ then the problem of finding $D(w), G(w), \beta(w)$ is just the mixed μ upper bound problem (across frequency), which is a convex problem and can be efficiently solved. Having chosen $D(w), G(w), \beta(w)$ at a set of frequency points (from the μ upper bound), we will fit real rational transfer matrices to them, in such a way that the interconnection is stable. For fixed $\mathbf{D}, \mathbf{G}, \beta$ transfer matrices then the problem of finding $\mathbf{K}$ will be reduced to a standard $\mathcal{H}_\infty$ problem which again is convex and can be efficiently solved. This will lead to a "D,G-K iteration" scheme, for the mixed μ synthesis problem.

Of course we will suffer from the same limitations as the "D-K iteration" procedure for complex μ synthesis, namely that the joint problem of optimizing $\mathbf{D}, \mathbf{G}$ *and* $\mathbf{K}$ is not convex. Thus the scheme will not guarantee to find the global optimum

of the problem. Nevertheless we will obtain a procedure which exploits the phase information in the real uncertainties, via the G scaling matrix, and so has the potential of greatly improving on the present approach, where one simply covers the real uncertainties with complex ones, and applies complex μ synthesis.

10.2.1 Reformulating the Mixed μ Upper Bound

Note that the problem statement in equations (10.4,10.5) is not quite in the form of a (scaled) singular value minimization, but rather a minimization subject to a (scaled) singular value constraint. In order to proceed we need to exploit some properties of the mixed μ upper bound.

Theorem 10.1 *Suppose we have matrices $M \in \mathbb{C}^{n\times n}, D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ and a real scalar $\beta > 0$ such that*

$$\overline{\sigma}\left(\left(\frac{DMD^{-1}}{\beta} - \mathbf{j}G\right)(I+G^2)^{-\frac{1}{2}}\right) \leq 1.$$

Then for any real $\hat{\beta} > \beta$ there exist matrices $\hat{D} \in \mathcal{D}_{\mathcal{K}}$ and $\hat{G} \in \mathcal{G}_{\mathcal{K}}$ such that

$$\overline{\sigma}\left(\left(\frac{\hat{D}M\hat{D}^{-1}}{\hat{\beta}} - \mathbf{j}\hat{G}\right)(I+\hat{G}^2)^{-\frac{1}{2}}\right) < 1.$$

Proof: For convenience define $\hat{M} \doteq DMD^{-1}$, and we have

$$\begin{array}{rcccl}
 & \overline{\sigma}\left(\left(\frac{\hat{M}}{\beta} - \mathbf{j}G\right)(I+G^2)^{-\frac{1}{2}}\right) & \leq & 1 & \\
\longrightarrow & \left(\frac{\hat{M}^*}{\beta} + \mathbf{j}G\right)\left(\frac{\hat{M}}{\beta} - \mathbf{j}G\right) & \leq & I+G^2 & \\
\longrightarrow & \hat{M}^*\hat{M} + \mathbf{j}\beta(G\hat{M} - \hat{M}^*G) & \leq & \beta^2 I & < \quad \hat{\beta}^2 I \\
\longrightarrow & \frac{\hat{M}^*\hat{M}}{\hat{\beta}^2} + \mathbf{j}\frac{\beta}{\hat{\beta}^2}(G\hat{M} - \hat{M}^*G) & < & I. &
\end{array}$$

Thus defining $\hat{D} \doteq D$, $\hat{G} \doteq \frac{\beta}{\hat{\beta}}G$ we have

$$\begin{array}{rcccl}
 & \frac{\hat{M}^*\hat{M}}{\hat{\beta}^2} + \frac{\mathbf{j}}{\hat{\beta}}(\hat{G}\hat{M} - \hat{M}^*\hat{G}) & < & I & \\
\longrightarrow & \left(\frac{\hat{M}^*}{\hat{\beta}} + \mathbf{j}\hat{G}\right)\left(\frac{\hat{M}}{\hat{\beta}} - \mathbf{j}\hat{G}\right) & < & I+\hat{G}^2 & \\
\longrightarrow & \overline{\sigma}\left(\left(\frac{\hat{D}M\hat{D}^{-1}}{\hat{\beta}} - \mathbf{j}\hat{G}\right)(I+\hat{G}^2)^{-\frac{1}{2}}\right) & < & 1. & \qquad \square
\end{array}$$

Theorem 10.2 *Suppose we have matrices* $M \in \mathbb{C}^{n\times n}, D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ *and real scalars* $\beta > 0$ *and* $0 < r \leq 1$ *such that*

$$\overline{\sigma}\left(\left(\frac{DMD^{-1}}{\beta} - \mathbf{j}G\right)(I+G^2)^{-\frac{1}{2}}\right) \leq r.$$

Then there exists $\hat{D} \in \mathcal{D}_{\mathcal{K}}, \hat{G} \in \mathcal{G}_{\mathcal{K}}$ *such that*

$$\overline{\sigma}\left(\left(\frac{\hat{D}M\hat{D}^{-1}}{r\beta} - \mathbf{j}\hat{G}\right)(I+\hat{G}^2)^{-\frac{1}{2}}\right) \leq 1.$$

Proof: Now define $\hat{M} \doteq \frac{DMD^{-1}}{\beta}$ and we have

$$\begin{array}{rcccccc}
 & \overline{\sigma}\left(\left(\hat{M} - \mathbf{j}G\right)(I+G^2)^{-\frac{1}{2}}\right) & \leq & r & & \\
\longrightarrow & (\hat{M}^* + \mathbf{j}G)(\hat{M} - \mathbf{j}G) & \leq & r^2(I+G^2) & & \\
\longrightarrow & \hat{M}^*\hat{M} + \mathbf{j}(G\hat{M} - \hat{M}^*G) & \leq & r^2I + (r^2-1)G^2 & \leq & r^2I \\
\longrightarrow & \frac{\hat{M}^*\hat{M}}{r^2} + \frac{\mathbf{j}}{r^2}(G\hat{M} - \hat{M}^*G) & \leq & I & &
\end{array}$$

so that defining $\hat{D} \doteq D$, $\hat{G} \doteq \frac{G}{r}$ we have

$$\begin{array}{rcccc}
 & \frac{\hat{M}^*\hat{M}}{r^2} + \frac{\mathbf{j}}{r}(\hat{G}\hat{M} - \hat{M}^*\hat{G}) & \leq & I & \\
\longrightarrow & \left(\frac{\hat{M}^*}{r} + \mathbf{j}\hat{G}\right)\left(\frac{\hat{M}}{r} - \mathbf{j}\hat{G}\right) & \leq & I + \hat{G}^2 & \\
\longrightarrow & \overline{\sigma}\left(\left(\frac{\hat{D}M\hat{D}^{-1}}{r\beta} - \mathbf{j}\hat{G}\right)(I+\hat{G}^2)^{-\frac{1}{2}}\right) & \leq & 1. & \square
\end{array}$$

These two theorems tell us something about how the mixed μ upper bound scales. Note that both of these results are trivial in the complex case, where $G = 0_n$. For the mixed case however the results are not so obvious, and in particular it is important to note that theorem 10.2 applies *only if* $r \leq 1$. From this theorem we see that whenever we have the appropriate scaling matrices and $r \leq 1$, then $r\beta$ is an upper bound for $\mu_{\mathcal{K}}(M)$. However when $r > 1$ we *cannot* conclude that $r\beta$ is an upper bound in the mixed case, even though this conclusion is obviously true in the complex case. In fact it is easy to construct mixed examples with $r > 1$ where $r\beta$ is strictly less than $\mu_{\mathcal{K}}(M)$.

We now propose the following "D,G-K iteration" for the mixed μ synthesis problem:

Procedure 10.3 (D,G-K **Iteration**)

1. *Find initial estimates of the scaling matrices $D(w)$, $G(w)$, and the real positive scalar β_*. One possibility for the scalings $D(w)$ is to choose them as the identity matrix at each frequency point. If $G(w)$ is chosen to be the zero matrix at each frequency point, then β_* is arbitrary (so choose say $\beta_* = 1$), otherwise it must satisfy*

$$\overline{\sigma}\left(\left(\frac{D(w)\mathbf{M}(\mathbf{P},\mathbf{K})(\mathbf{j}w)D^{-1}(w)}{\beta_*} - \mathbf{j}G(w)\right)(I+G(w)^2)^{-\frac{1}{2}}\right) \leq 1$$

for all w, for some stabilizing controller $\mathbf{K}$.

2. *Fit State Space realizations $\mathbf{D}$ and $\mathbf{G}$ to the pointwise scaling matrices $D(w)$ and $\mathbf{j}G(w)$, so that $\mathbf{D}(\mathbf{j}w)$ approximates $D(w)$, and $\mathbf{G}(\mathbf{j}w)$ approximates $\mathbf{j}G(w)$. Now replace $\mathbf{D}$ and $\mathbf{G}$ with appropriate factors so that $\mathbf{D}$, $\mathbf{D}^{-1}$, $\mathbf{G}_h$, and $\mathbf{G}\mathbf{G}_h$ are all stable, where $\mathbf{G}_h$ is a spectral factor satisfying $(I+\mathbf{G}^{\sim}\mathbf{G})^{-1} = \mathbf{G}_h\mathbf{G}_h^{\sim}$. Augment $\mathbf{D}$ and $\mathbf{G}_h$ with identity matrices, and $\mathbf{G}$ with a zero matrix, of appropriate dimensions so that $\mathbf{D},\mathbf{G},\mathbf{G}_h$ are all compatible with $\mathbf{P}$. Form the State Space system $\mathbf{P}_{\mathbf{DG}} = (\mathbf{D}\mathbf{P}\mathbf{D}^{-1} - \beta_*\mathbf{G})\mathbf{G}_h$.*

3. *Find the $\mathcal{H}_\infty$ optimal controller $\hat{\mathbf{K}}$ minimizing $\|\mathbf{M}(\mathbf{P}_{\mathbf{DG}},\mathbf{K})\|_\infty$ over all all stabilizing, proper, real rational controllers $\mathbf{K}$.*

4. *Compute β_* as*

$$\beta_* = \sup_{w\in\mathbb{R}} \inf_{\tilde{D}(w)\in\mathcal{D}_\mathcal{K},\tilde{G}(w)\in\mathcal{G}_\mathcal{K}} \inf_{\beta(w)\in\mathbb{R},\beta(w)>0} \{\beta(w) : \Gamma \leq 1\}$$

where

$$\Gamma = \overline{\sigma}\left(\left(\frac{\tilde{D}(w)\mathbf{M}(\mathbf{P},\mathbf{K})(\mathbf{j}w)\tilde{D}^{-1}(w)}{\beta(w)} - \mathbf{j}\tilde{G}(w)\right)(I+\tilde{G}^2(w))^{-\frac{1}{2}}\right).$$

5. *Find $\hat{D}(w),\hat{G}(w)$ solving the minimization problem*

$$\inf_{\hat{D}(w)\in\mathcal{D}_\mathcal{K},\hat{G}(w)\in\mathcal{G}_\mathcal{K}} \overline{\sigma}\left(\left(\frac{\hat{D}(w)\mathbf{M}(\mathbf{P},\mathbf{K})(\mathbf{j}w)\hat{D}^{-1}(w)}{\beta_*} - \mathbf{j}\hat{G}(w)\right)(I+\hat{G}^2(w))^{-\frac{1}{2}}\right)$$

pointwise across frequency.

6. *Compare the new scaling matrices* $\hat{D}(w), \hat{G}(w)$ *with the previous estimates* $D(w)$, $G(w)$. *Stop if they are close, else replace* $D(w), G(w)$ *with* $\hat{D}(w), \hat{G}(w)$ *respectively and return to step 2.*

We will see that this iteration (assuming perfect State Space realizations of $D(w)$ and $G(w)$) is monotonically nonincreasing, so that we are guaranteed convergence to a local minimum of the problem. Having converged to such a point then the controller $\hat{\mathbf{K}}$ from step 3 is the resulting mixed μ synthesis controller, and it satisfies

$$\sup_{w \in \mathbb{R}} \mu_{\mathcal{K}}(\mathbf{M}(\mathbf{P}, \hat{\mathbf{K}})(\mathrm{j}w)) \leq \beta_*.$$

Note that, as in the complex μ case, we use the $\mathcal{H}_\infty$ optimal control solution to synthesize the controller. By construction the scaled system $\mathbf{P_{DG}}$ fits into the standard $\mathcal{H}_\infty$ optimal control framework, as shown in figure 10.3.

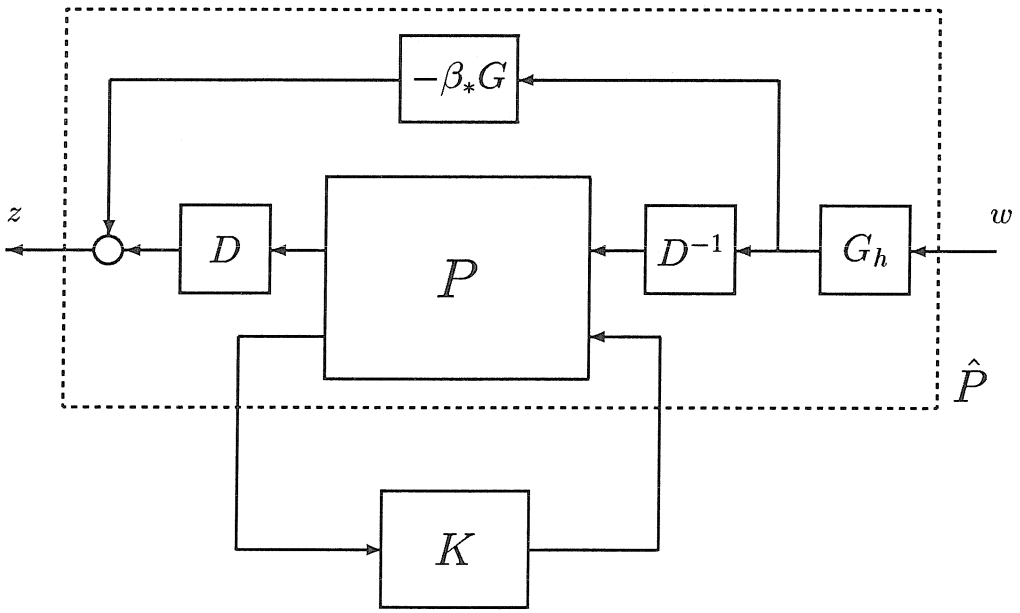


Figure 10.3: The "K" iteration of the "D,G-K iteration" procedure

The fact that this iteration is monotonically nonincreasing follows from theorems 10.1 and 10.2 in the following way. Since β_* was defined in step 4 as the maximum across frequency of the upper bound we have that $\beta_* \geq \beta(w)$ for all w, with $\beta_* = \beta(w)$ only for those frequency points whose upper bound achieves the maximum across frequency. Then theorem 10.1 guarantees us that in step 5 we can achieve

$$\overline{\sigma}\left(\left(\frac{\hat{D}(w)\mathbf{M}(\mathbf{P},\mathbf{K})(\mathrm{j}w)\hat{D}^{-1}(w)}{\beta_*} - \mathrm{j}\hat{G}(w)\right)(I+\hat{G}^2(w))^{-\frac{1}{2}}\right) < 1$$

for all frequency points with $\beta_* > \beta(w)$ and

$$\overline{\sigma}\left(\left(\frac{\hat{D}(w)\mathbf{M}(\mathbf{P},\mathbf{K})(\mathrm{j}w)\hat{D}^{-1}(w)}{\beta_*} - \mathrm{j}\hat{G}(w)\right)(I+\hat{G}^2(w))^{-\frac{1}{2}}\right) = 1$$

for points with $\beta_* = \beta(w)$. Now in step 3 we have that our previous controller, $\hat{\mathbf{K}}$, stabilized the interconnection $\mathbf{M}(\mathbf{P},\hat{\mathbf{K}})$, and hence, by our choice of factorizations for $\mathbf{D}$ and $\mathbf{G}_h$, $\hat{\mathbf{K}}$ stabilizes $\mathbf{M}(\mathbf{P_{DG}},\hat{\mathbf{K}})$ as well. Note that the above formulae imply that $\hat{\mathbf{K}}$ achieved $\|\mathbf{M}(\mathbf{P_{DG}},\hat{\mathbf{K}})\|_\infty = \beta_*$, so that we are guaranteed the existence of a stabilizing controller achieving at least this level for $\|\mathbf{M}(\mathbf{P_{DG}},\hat{\mathbf{K}}_{new})\|_\infty$. Thus our new $\mathcal{H}_\infty$ optimal controller in step 3 will satisfy $\|\mathbf{M}(\mathbf{P_{DG}},\hat{\mathbf{K}}_{new})\|_\infty = r\beta_*$ with $r \leq 1$. But theorem 10.2 implies that $r\beta_*$ is now an upper bound for μ across frequency, and so the new computation of β_{*new} in step 4 will yield $\beta_{*new} \leq r\beta_* \leq \beta_*$. Thus the iteration is monotonically nonincreasing, and furthermore if the new $\mathcal{H}_\infty$-optimal controller from step 3 achieves any improvement over the old one (i.e., $r < 1$) then the peak value of the μ upper bound across frequency is strictly decreased.

Note also that the $\mathcal{H}_\infty$ norm of the interconnection $\mathbf{M}(\mathbf{P_{DG}},\hat{\mathbf{K}})$ (formed by closing the new interconnection $\mathbf{P_{DG}}$ from step 2 with the previous controller $\hat{\mathbf{K}}$) is determined by those frequencies which have the maximum value for the mixed μ upper bound. This places the emphasis for the $\mathcal{H}_\infty$ norm minimization at the frequencies where mixed μ is largest (which is just what we want). Furthermore the way we have constructed the interconnection $\mathbf{P_{DG}}$ *guarantees* that if the new controller $\hat{\mathbf{K}}_{new}$ reduces the $\mathcal{H}_\infty$ norm of $\mathbf{M}(\mathbf{P_{DG}},\hat{\mathbf{K}}_{new})$, then the peak value of the μ upper bound across frequency is reduced. It is important to have this guarantee because of

the non-uniform way in which the scaling $\beta(w)$ appears in the mixed μ upper bound (see step 5). It ensures that we do not suffer from any "pop-up" type phenomena, where for some frequencies a small increase in the maximum singular value creates a very large increase in μ. The possibility of this type of behavior stems from the fact that for mixed problems theorem 10.2 does not hold for $r > 1$. The proposed "D,G-K iteration" explicitly forbids this "pop-up" type of behavior.

Note that in step 4 we compute scalings $\tilde{D}(w), \tilde{G}(w)$ which minimize the scalar $\beta(w)$ at each frequency subject to a (scaled) singular value constraint, and this is exactly the mixed μ upper bound computation across frequency. A closely related problem is solved in step 5, where we compute scalings $\hat{D}(w), \hat{G}(w)$ to minimize a (scaled) singular value problem with β_* *fixed* across frequency. If we consider complex problems then the separate computation of $\tilde{D}(w), \tilde{G}(w)$ in step 4, and then $\hat{D}(w), \hat{G}(w)$ in step 5 is unnecessary, since it is easy to see that the two sets of optimal scalings may be taken to be the same. Thus for complex problems we do not need the second pass, taken in step 5, to compute the scaling matrices that are to be fitted for the synthesis problem, but rather we just use the scalings from the analysis problem (solved in step 4) directly. For mixed problems however the two sets of scalings in steps 4 and 5 can be quite different, and it is important to use the scalings $\hat{D}(w), \hat{G}(w)$ from step 5 for fitting (in step 2), rather than the scalings $\tilde{D}(w), \tilde{G}(w)$ from the mixed μ analysis problem in step 4.

10.2.2 Forming The New Interconnection

In order to implement the "D,G-K iteration" described in the preceding subsection we need to be able to carry out certain factorizations, required in step 2. These factorizations are used to ensure that if $\hat{\mathbf{K}}$ stabilizes $\mathbf{M}(\mathbf{P}, \hat{\mathbf{K}})$, then $\hat{\mathbf{K}}$ also stabilizes $\mathbf{M}((\mathbf{DPD}^{-1} - \beta_* \mathbf{G})\mathbf{G}_h, \hat{\mathbf{K}})$, and vica versa (for the converse direction we need a slight technical assumption that we don't get right half-plane pole/zero cancellations when forming $\mathbf{DPD}^{-1}\mathbf{G}_h$). The fact that $\hat{\mathbf{K}}$ stabilizes *both* these interconnections was used

when showing that the "D,G-K iteration" is monotonically nonincreasing (and hence converges).

The problem of fitting pointwise frequency data $D(w)$,$G(w)$ with State Space realizations $\mathbf{D}, \mathbf{G}$ will be considered in the next subsection. For now let us assume that we can obtain such fits. We will see in section 10.3 that we may assume that $\mathbf{D}$ is square and invertible with $\mathbf{D}, \mathbf{D}^{-1} \in \mathcal{RL}_\infty$, but the only assumption we will make about $\mathbf{G}$ is that it is square.

Given State Space realizations for $\mathbf{D}$ and $\mathbf{G}$ we want to factor them as outlined in step 2 of the "D,G-K iteration," described in procedure 10.3, so that for the new $\mathbf{D}, \mathbf{G}, \mathbf{G}_h$ the interconnection $\mathbf{M}((\mathbf{D}\mathbf{P}\mathbf{D}^{-1} - \beta_* \mathbf{G})\mathbf{G}_h, \hat{\mathbf{K}})$ is stable, and we would like to do this without affecting the value of

$$\overline{\sigma}(\mathbf{M}((\mathbf{D}(\mathrm{j}w)\mathbf{P}(\mathrm{j}w)\mathbf{D}^{-1}(\mathrm{j}w) - \beta_* \mathbf{G}(\mathrm{j}w))\mathbf{G}_h(\mathrm{j}w), \hat{\mathbf{K}}(\mathrm{j}w)))$$

for any frequency. This can be accomplished by using the machinery developed in chapter 9, applied to each block of $\mathbf{D}$ and $\mathbf{G}$.

First of all suppose we have fitted $D(w)$ with a transfer matrix $\mathbf{D}$, but have not yet fitted $G(w)$. Apply theorem 9.4 to each block of $\mathbf{D}$ to factor out an all pass function so that the resulting block of $\mathbf{D}$ is stable minimum phase (hence $\mathbf{D}$ and $\mathbf{D}^{-1}$ are stable). Thus we have a new transfer matrix $\mathbf{D}_{smp}$ which is stable minimum phase, and at each frequency point we have

$$\mathbf{D}(\mathrm{j}w) = U(w)\mathbf{D}_{smp}(\mathrm{j}w) \tag{10.6}$$

where U(w) is a unitary block diagonal matrix (with the same block structure as $\mathbf{D}$). Then denote $\mathbf{D}(\mathrm{j}w), U(w)$ and $\mathbf{D}_{smp}(\mathrm{j}w)$ by D_w, U_w and D_{sw} respectively. Finally denote the pointwise frequency data $G(w)$ (to be fitted) by G_w, and we have the

following equivalences:

$$\begin{aligned}
& \overline{\sigma}\left(\left(D_w\mathbf{M}(\mathbf{P},\hat{\mathbf{K}})(\mathbf{j}w)D_w^{-1}-\mathbf{j}\beta_* G_w\right)(I+G_w^2)^{-\frac{1}{2}}\right) \\
= \;& \overline{\sigma}\left(\left(U_w D_{sw}\mathbf{M}(\mathbf{P},\hat{\mathbf{K}})(\mathbf{j}w)D_{sw}^{-1}U_w^*-\mathbf{j}\beta_* G_w\right)(I+G_w^2)^{-\frac{1}{2}}\right) \\
= \;& \overline{\sigma}\left(\left(D_{sw}\mathbf{M}(\mathbf{P},\hat{\mathbf{K}})(\mathbf{j}w)D_{sw}^{-1}-\mathbf{j}\beta_* U_w^* G_w U_w\right)U_w^*(I+G_w^2)^{-\frac{1}{2}}U_w\right) \\
= \;& \overline{\sigma}\left(\left(D_{sw}\mathbf{M}(\mathbf{P},\hat{\mathbf{K}})(\mathbf{j}w)D_{sw}^{-1}-\mathbf{j}\beta_*\hat{G}_w\right)(I+\hat{G}_w^2)^{-\frac{1}{2}}\right)
\end{aligned} \tag{10.7}$$

where $\hat{G}_w = U_w^* G_w U_w$. Note that $\hat{G}_w \in \mathcal{G}_{\mathcal{K}}$, so now choose the transfer matrix $\mathbf{G}$ so that $\mathbf{G}(\mathbf{j}w)$ is a fit to $\mathbf{j}\hat{G}_w$ (rather than $\mathbf{j}G_w$). We may now apply theorem 9.5 to each non-zero block of $\mathbf{G}$ to construct $\mathbf{G}_h$, with $\mathbf{G}_h$ and $\mathbf{G}\mathbf{G}_h$ stable, and $\mathbf{G}_h$ having the same block diagonal structure as $\mathbf{G}$. Furthermore at each frequency $\mathbf{G}_h$ satisfies

$$(I+\mathbf{G}^{\sim}(\mathbf{j}w)\mathbf{G}(\mathbf{j}w))^{-1} = (I+\hat{G}_w^2)^{-1} = \mathbf{G}_h(\mathbf{j}w)\mathbf{G}_h^{\sim}(\mathbf{j}w). \tag{10.8}$$

Now for convenience denote $D_{sw}\mathbf{M}(\mathbf{P},\hat{\mathbf{K}})(\mathbf{j}w)D_{sw}^{-1}$ by M_{D_w}, and $\mathbf{G}_h(\mathbf{j}w)$ by G_{hw}. Then note that we have

$$\begin{aligned}
& \overline{\sigma}^2\left(\left(M_{D_w}-\mathbf{j}\beta_*\hat{G}_w\right)(I+\hat{G}_w^2)^{-\frac{1}{2}}\right) \\
= \;& \overline{\lambda}\left((I+\hat{G}_w^2)^{-\frac{1}{2}}\left(M_{D_w}^*-\beta_*\mathbf{G}^*(\mathbf{j}w)\right)\left(M_{D_w}-\beta_*\mathbf{G}(\mathbf{j}w)\right)(I+\hat{G}_w^2)^{-\frac{1}{2}}\right) \\
= \;& \overline{\lambda}\left(\left(M_{D_w}^*-\beta_*\mathbf{G}^*(\mathbf{j}w)\right)\left(M_{D_w}-\beta_*\mathbf{G}(\mathbf{j}w)\right)(I+\hat{G}_w^2)^{-1}\right) \\
= \;& \overline{\lambda}\left(\left(M_{D_w}^*-\beta_*\mathbf{G}^*(\mathbf{j}w)\right)\left(M_{D_w}-\beta_*\mathbf{G}(\mathbf{j}w)\right)G_{hw}G_{hw}^{\sim}\right) \\
= \;& \overline{\lambda}\left(G_{hw}^{\sim}\left(M_{D_w}^*-\beta_*\mathbf{G}^*(\mathbf{j}w)\right)\left(M_{D_w}-\beta_*\mathbf{G}(\mathbf{j}w)\right)G_{hw}\right) \\
= \;& \overline{\sigma}^2\left(\left(M_{D_w}-\beta_*\mathbf{G}(\mathbf{j}w)\right)G_{hw}\right)
\end{aligned} \tag{10.9}$$

so that this factorization does just what we require, and the appropriate State Space formulae for $\mathbf{D}, \mathbf{G}_h$ and $\mathbf{G}\mathbf{G}_h$ may be obtained by applying the results from chapter 9 to each block of $\mathbf{D}$ and $\mathbf{G}$.

Note that the final step of this process is to augment $\mathbf{D}$ and $\mathbf{G}_h$ with an identity matrix, and $\mathbf{G}$ with a zero matrix, of appropriate dimensions so that $\mathbf{D}, \mathbf{G}, \mathbf{G}_h$ are all compatible with $\mathbf{P}$, and then form the interconnection $\mathbf{P}_{\mathbf{DG}} = (\mathbf{D}\mathbf{P}\mathbf{D}^{-1}-\beta_*\mathbf{G})\mathbf{G}_h$. Note that when forming $\mathbf{G}\mathbf{G}_h$ we were able to employ state cancellations to obtain a stable realization with the same number of states as $\mathbf{G}_h$. It turns out that we

may once again exploit the special structure of $\mathbf{G}$ and $\mathbf{G}_h$ to obtain further state cancellations when forming the interconnection $\mathbf{P_{DG}}$. This state reduction comes about by cancelling uncontrollable and/or unobservable modes in the interconnection, and is highly desirable from a numerical viewpoint since the $\mathcal{H}_\infty$ optimal controller computed in step 3 of procedure 10.3 will have the same number of states as our realization for $\mathbf{P_{DG}}$.

Note from theorem 9.5 that the State Space formulae for the blocks of $\mathbf{G}_h$ and $\mathbf{GG}_h$ share the same "A" and "B" matrices. In other words we have realizations for each block of $\mathbf{G}_h$ and $\mathbf{GG}_h$ which take the form

$$(\mathbf{G}_h)_i = \left[\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right] \qquad (\mathbf{GG}_h)_i = \left[\begin{array}{c|c} A_i & B_i \\ \hline \hat{C}_i & \hat{D}_i \end{array}\right]. \tag{10.10}$$

Now that it is easy to show that since each block of $\mathbf{G}_h$ and $\mathbf{GG}_h$ satisfies these relationships, we may choose the block diagonal system realizations of $\mathbf{G}_h$ and $\mathbf{GG}_h$ to satisfy them as well. Furthermore we may scale $\mathbf{GG}_h$ by a real scalar (namely $-\beta_*$) and still preserve these relationships, since one may easily check that for any $\beta \in \mathbb{R}$

$$\beta\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{c|c} A & B \\ \hline \beta C & \beta D \end{array}\right]. \tag{10.11}$$

Thus we may choose our realizations for $\mathbf{G}_h$ and $-\beta_*\mathbf{GG}_h$ to be of the form

$$\mathbf{G}_h = \left[\begin{array}{c|c} A_G & B_G \\ \hline C_H & D_H \end{array}\right] \qquad -\beta_*\mathbf{GG}_h = \left[\begin{array}{c|c} A_G & B_G \\ \hline C_G & D_G \end{array}\right]. \tag{10.12}$$

Now suppose that our State Space realization for $\mathbf{DPD}^{-1}$ is

$$\mathbf{DPD}^{-1} = \left[\begin{array}{c|c} A_D & B_D \\ \hline C_D & D_D \end{array}\right]. \tag{10.13}$$

Then the interconnection $\mathbf{P_{DG}}$ is given by

$$\begin{aligned}
\mathbf{P_{DG}} &= (\mathbf{DPD^{-1}} - \beta_* \mathbf{G})\mathbf{G}_h \\
&= \mathbf{DPD^{-1}G}_h - \beta_* \mathbf{GG}_h \\
&= \left[\begin{array}{c|c} A_D & B_D \\ \hline C_D & D_D \end{array}\right]\left[\begin{array}{c|c} A_G & B_G \\ \hline C_H & D_H \end{array}\right] + \left[\begin{array}{c|c} A_G & B_G \\ \hline C_G & D_G \end{array}\right] \\
&= \left[\begin{array}{ccc|c} A_D & B_D C_H & 0 & B_D D_H \\ 0 & A_G & 0 & B_G \\ 0 & 0 & A_G & B_G \\ \hline C_D & D_D C_H & C_G & D_D D_H + D_G \end{array}\right] \\
&= \left[\begin{array}{ccc|c} A_D & B_D C_H & 0 & B_D D_H \\ 0 & A_G & 0 & B_G \\ 0 & 0 & A_G & 0 \\ \hline C_D & D_D C_H + C_G & C_G & D_D D_H + D_G \end{array}\right] \\
&= \left[\begin{array}{cc|c} A_D & B_D C_H & B_D D_H \\ 0 & A_G & B_G \\ \hline C_D & D_D C_H + C_G & D_D D_H + D_G \end{array}\right]
\end{aligned} \tag{10.14}$$

where we have made use of the similarity transformation $T = \begin{pmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & I & I \end{pmatrix}$. Note that in the final formula for $\mathbf{P_{DG}}$ we only have the states from $\mathbf{DPD^{-1}}$ and $\mathbf{G}_h$ appearing once each.

10.3 Fitting D and G

Finally we consider the problem of fitting State Space system realizations to frequency response data. Of course we will not attempt to exactly interpolate the data at each frequency point since that would require, in general, too many states. Rather we will attempt to obtain realizations that approximate the data over the frequency range of interest. Note that since we will be concerned with the $\mathcal{H}_\infty$ norm of the resulting

interconnection, we will still require acceptable behavior from the State Space fits even outside the frequency range we are especially interested in. In particular we will require that the fits behave reasonably as the frequency tends to zero or infinity.

First of all consider the problem of fitting $D(w)$. This problem is encountered when implementing a "D-K iteration" for complex μ synthesis, and hence has already received a good deal of attention. Several methods have been developed to tackle this problem, and although there is certainly room for improvement, they often work quite well.

In particular the μ-Tools toolbox [3] contains two algorithms for fitting scalar magnitude data with a SISO system. One algorithm, *fitmag*, uses the complex-cepstrum (see [52]) to generate an appropriate phase, and then uses the algorithm *invfreqs* from the signal processing toolbox [43], which uses an equation error method to identify a system model from the frequency response data. Another μ-Tools algorithm, *fitmaglp*, uses a linear programming technique to find a SISO system, which attempts to fit the magnitude data. In addition to these two algorithms for fitting magnitude data, there is the algorithm *fitsys*, which attempts to fit given scalar frequency response data in both magnitude and phase (with a SISO system). This algorithm is based heavily on the algorithm *invfreqs* (see [43] for more details).

Now one may easily build a routine to fit $D(w)$ by using the above algorithms to fit the (scalar) elements of $D(w)$ with SISO systems. The diagonal elements can be fitted (in magnitude only) with *fitmag* or *fitmaglp*, and the non-zero off diagonal elements can be fitted with *fitsys*. Putting these systems together appropriately we can form $\mathbf{D} \in \mathcal{RM}$ fitting $D(w)$. This is the procedure we will use here, and we will not go into any further details, since it is the same procedure one typically uses in the complex μ synthesis case. Note that since $\mathbf{D}$ is a State Space system, $\mathbf{D}(\mathrm{j}w)$ will tend to a real matrix as w goes to zero or infinity. But now we have that $\mathbf{M}(\mathbf{P}, \hat{\mathbf{K}})(\mathrm{j}w)$ will also tend to a real matrix (for the same reason) and so theorem 5.11 implies that the optimal $D(w)$, at zero or infinite frequency, may be taken to be real as well. Thus

the low/high frequency limits of $D(w)$ are amenable to fitting with a State Space system $\mathbf{D}(\mathbf{j}w)$. Note also that all blocks of $\mathbf{D}$ are square, and furthermore each block of $\mathbf{D}(\mathbf{j}w)$ may be taken to be Hermitian positive definite. Thus we do not need to allow sign changes in the (eigenvalues of) the blocks of $\mathbf{D}(\mathbf{j}w)$ and hence we do not need poles or zeros on the imaginary axis. Thus we may restrict our fitting routines to return $\mathbf{D} \in \mathcal{RM}$ invertible, with $\mathbf{D}, \mathbf{D}^{-1} \in \mathcal{RL}_\infty$ (which is typically done in the complex μ synthesis case, for the same reasons) as we required in section 10.2.2.

Now let's turn our attention to the problem of fitting $\mathbf{j}G(w)$. Once again we will tackle this problem by fitting the (scalar) elements of $\mathbf{j}G(w)$ with SISO systems, and using these to build up $\mathbf{G}$. For the non-zero off diagonal elements we can once again fit them using *fitsys*. Note that at zero or infinite frequency theorem 5.11 implies that the optimal $G(w)$ may be taken to be pure imaginary, so that $\mathbf{j}G(w)$ may be taken to be pure real, and hence amenable to fitting with a State Space system $\mathbf{G}(\mathbf{j}w)$. For the diagonal elements of $\mathbf{j}G(w)$ we have additional constraints, which are explored in more detail below.

Note that when choosing $\mathbf{D}$ to fit $D(w)$ we could always factor out an appropriate block diagonal unitary matrix, without affecting the maximum singular value of the interconnection. For the diagonal elements of $\mathbf{D}(\mathbf{j}w)$ then, we could use this unitary matrix to generate any required phase, so we needed them to fit $D(w)$ in *magnitude only*, and we could do this using existing tools. When choosing $\mathbf{G}$ to fit $\mathbf{j}G(w)$ this is not the case. We do not have the luxury of factoring out a unitary matrix, and so we will need to fit in both magnitude *and* phase. This means that we will require that the phase of the diagonal elements of $\mathbf{G}$ is purely imaginary for *all frequency.*

In order to see what constraints this places on $\mathbf{G}$ consider first the conditions under which a scalar transfer function $\mathbf{g}(s)$ is purely real for all frequency.

Lemma 10.1 *Consider a scalar transfer function* $\mathbf{g} \in \mathcal{RM}$. *Then the following statements are equivalent*

1. $\mathbf{g}(\mathbf{j}w)$ *is purely real for all* $w \in \mathbb{R}$.

2. $\mathbf{g}^{\sim}(s) \doteq \mathbf{g}(-s) = \mathbf{g}(s)$.

3. $\mathbf{g}(s)$ *has pole and zero symmetry about the imaginary axis, with the total number of poles and zeros at the origin being even.*

4. $\mathbf{g}(s) = \frac{z(s^2)}{p(s^2)}$ *where* $z(s^2)$ *and* $p(s^2)$ *are polynomials in* s^2.

Proof: We will prove that $1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 4 \longrightarrow 1$. Suppose first $\mathbf{g}(\mathrm{j}w)$ is purely real for all $w \in \mathbb{R}$. Then

$$\mathbf{g}(\mathrm{j}w) = \mathbf{g}^*(\mathrm{j}w) = \mathbf{g}(-\mathrm{j}w)$$

and hence $\mathbf{g}(s) - \mathbf{g}(-s)$ is zero for all $s = \mathrm{j}w$. But since $(\mathbf{g}(s) - \mathbf{g}(-s)) \in \mathcal{RM}$ it has finitely many zeros and hence we must have $\mathbf{g}(s) - \mathbf{g}(-s) = 0$, or in other words $\mathbf{g}^{\sim}(s) \doteq \mathbf{g}(-s) = \mathbf{g}(s)$. Now $\mathbf{g}(s) = \mathbf{g}(-s)$ gives that whenever s_0 is a zero (pole) of $\mathbf{g}(s)$, then $-s_0$ is a zero (pole) of $\mathbf{g}(s)$. Together with the fact that for $\mathbf{g} \in \mathcal{RM}$ the poles and zeros come in conjugate symmetric pairs, this implies that we have pole and zero symmetry about the imaginary axis. Since this implies that all the poles and zeros not at the origin must come in pairs (or fours), it is easy to check that we must also have the total number of poles and zeros at the origin being even. By collecting terms in pairs it is easy to check that, if we have pole and zero symmetry about the imaginary axis, and the total number of poles and zeros at the origin is even, then we may write $\mathbf{g}(s) = \frac{z(s^2)}{p(s^2)}$. Finally it is clear that if $\mathbf{g}(s) = \frac{z(s^2)}{p(s^2)}$ then $\mathbf{g}(\mathrm{j}w)$ is purely real for all $w \in \mathbb{R}$. □

A simple extension of these arguments gives us the conditions under which $\mathbf{g}(s)$ is purely imaginary for all frequency.

Lemma 10.2 *Consider a scalar transfer function* $\mathbf{g} \in \mathcal{RM}$. *Then the following statements are equivalent*

1. $\mathbf{g}(\mathrm{j}w)$ *is purely imaginary for all* $w \in \mathbb{R}$.

2. $\mathbf{g}^{\sim}(s) \doteq \mathbf{g}(-s) = -\mathbf{g}(s)$.

3. $\mathbf{g}(s)$ *has pole and zero symmetry about the imaginary axis, with the total number of poles and zeros at the origin being odd.*

4. $\mathbf{g}(s) = s\frac{z(s^2)}{p(s^2)}$ *where* $z(s^2)$ *and* $p(s^2)$ *are polynomials in* s^2.

Proof: Again we prove that $1 \longrightarrow 2 \longrightarrow 3 \longrightarrow 4 \longrightarrow 1$. First suppose that $\mathbf{g}(\mathbf{j}w)$ is purely imaginary for all $w \in \mathbb{R}$. Then

$$\mathbf{g}(\mathbf{j}w) = -\mathbf{g}^*(\mathbf{j}w) = -\mathbf{g}(-\mathbf{j}w)$$

and hence $\mathbf{g}(s) + \mathbf{g}(-s)$ is zero for all $s = \mathbf{j}w$. But since $(\mathbf{g}(s) + \mathbf{g}(-s)) \in \mathcal{RM}$ it has finitely many zeros and hence we must have $\mathbf{g}(s) + \mathbf{g}(-s) = 0$, or in other words $\mathbf{g}^{\sim}(s) \doteq \mathbf{g}(-s) = -\mathbf{g}(s)$. Now $\mathbf{g}(s) = -\mathbf{g}(-s)$ gives that whenever s_0 is a zero (pole) of $\mathbf{g}(s)$, then $-s_0$ is a zero (pole) of $\mathbf{g}(s)$. Together with the fact that for $\mathbf{g} \in \mathcal{RM}$ the poles and zeros come in conjugate symmetric pairs, this implies that we have pole and zero symmetry about the imaginary axis. Since this implies that all the poles and zeros not at the origin must come in pairs (or fours), it is easy to check that we must also have the total number of poles and zeros at the origin being odd. By collecting terms in pairs it is easy to check that, if we have pole and zero symmetry about the imaginary axis, and the total number of poles and zeros at the origin is odd, then we may write $\mathbf{g}(s) = s\frac{z(s^2)}{p(s^2)}$. Finally it is clear that if $\mathbf{g}(s) = s\frac{z(s^2)}{p(s^2)}$ then $\mathbf{g}(\mathbf{j}w)$ is purely imaginary for all $w \in \mathbb{R}$. □

Thus we see that a purely imaginary transfer function is obtained by adding a zero at the origin to a purely real transfer function. Our algorithm to fit purely imaginary SISO transfer functions, *fitmagreal*, which we will use to fit the diagonal elements of $\mathbf{j}G(w)$, exploits this fact and is described briefly below.

Suppose that $g(w)$ is a diagonal element of $G(w)$. Then we will first fit $h(w) \doteq \frac{g(w)}{w}$ with a purely real SISO transfer function, $\hat{\mathbf{g}}$. In order to do this we will parametrize our allowable transfer functions for the fit as

$$\hat{\mathbf{g}}(s) = \frac{z(s^2)}{p(s^2)}$$

where the degree of $p(s^2)$ is chosen by the user, and the degree of $z(s^2)$ is chosen so that $\hat{\mathbf{g}}(s)$ is *strictly proper*. Thus we are left with the problem of finding the coefficients of z and p so that $\hat{\mathbf{g}}(\mathbf{j}w) = \frac{z(-w^2)}{p(-w^2)}$ best approximates $h(w)$. Note that all the quantities concerned in this problem are real numbers. We tackle this problem using an algorithm *realfit*, which is a modification of the algorithm *fitmaglp*, and is based on a linear programming approach. Note that $h(w)$ is not constrained in sign, so that we will wish to allow sign changes in $\hat{\mathbf{g}}(\mathbf{j}w)$. Hence we will not place any restrictions on the pole/zero locations of $\hat{\mathbf{g}}(s)$, and in particular it may have poles and/or zeros on the imaginary axis.

Having obtained $\hat{\mathbf{g}}(s)$ as a strictly proper transfer function, say

$$\hat{\mathbf{g}} = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \tag{10.15}$$

then we form $\mathbf{g}$ as

$$\mathbf{g}(s) \doteq s\hat{\mathbf{g}}(s) \quad \longrightarrow \quad \mathbf{g} = \left[\begin{array}{c|c} A & B \\ \hline CA & CB \end{array}\right]. \tag{10.16}$$

Thus we obtain $\mathbf{g}$ as a purely imaginary transfer function, with $\mathbf{g}(\mathbf{j}w)$ approximating $\mathbf{j}g(w)$. Note from this construction that the transfer matrix $\mathbf{G}$, built from these transfer functions, may contain poles and zeros located on the imaginary axis. For this reason the factorization of $\mathbf{G}$ used in section 10.2.2 did not make any assumptions about the pole/zero locations of $\mathbf{G}$.

Note that lemma 10.2 implies that any non-zero purely imaginary (proper) transfer function $\mathbf{g} \in \mathcal{RM}$ is in fact strictly proper, and indeed our construction for $\mathbf{g}$ will yield a strictly proper transfer function. Thus $\mathbf{g}(\mathbf{j}w)$ tends to zero as w tends to infinity. Furthermore, assuming $\hat{\mathbf{g}}(s)$ has no poles at the origin, then $\mathbf{g}(s)$ has a zero at the origin, so that $\mathbf{g}(\mathbf{j}w)$ also tends to zero as w tends to zero. Now by earlier arguments we know that theorem 5.11 implies that the optimal $G(w)$ may be taken to be purely imaginary at zero or infinite frequency. But $G(w)$ is Hermitian and so the diagonal elements are purely real, and hence the diagonal elements of the optimal

$G(w)$ may be taken to be zero at zero or infinite frequency. Thus the high and low frequency limits of $\mathbf{g}(\mathbf{j}w)$ are compatible with the data it is attempting to fit (i.e., both $\mathbf{j}g(w)$ and $\mathbf{g}(\mathbf{j}w)$ tend to zero at high/low frequency).

Finally we note that since the routine *fitsys* can be used to fit a SISO transfer function in both magnitude and phase, we could apply this algorithm directly to the problem of fitting the diagonal elements of $G(w)$, rather than the approach taken above. Note however that the above approach *enforces* the phase constraints for all frequency, and enforces the correct behavior at the high and low frequency limits. We would not obtain these guarantees by using *fitsys*. Furthermore by simply applying *fitsys* we would not be exploiting all the information that we have about the structure of $\mathbf{g}$ (as is done above), and in fact we found that the *fitsys* routine performed poorly for many problems when *fitmagreal* was able to obtain quite good fits.

10.4 Some Practical Considerations

These procedures have been implemented in software as Matlab functions (m-files) "rmufit" and "rmuflp." Together with the analysis software, "rmu," from chapter 6, they allow the user to implement the "D,G-K iteration" described in the preceding subsections, and hence to design mixed μ synthesis controllers.

It should be noted that for complex μ problems then the "D,G-K iteration" reduces exactly to the familiar "D-K iteration" for complex μ synthesis. Thus many of the techniques required to obtain a practical design from a complex μ synthesis procedure carry over into the mixed case as well. There are many issues to be considered, and we will not go into any of the details here, except to make the following comments. The procedures developed in this chapter did not make any restrictions on the block structure, $\mathcal{K}$, of the uncertainties in the underlying robustness problem. In particular repeated scalar blocks (real or complex) were allowed. It should be noted however that for real or complex μ problems the presence of repeated scalar uncertainty blocks gives rise to full blocks in the scaling matrices $(D(w), G(w))$. Since the

present approach fits element by element, then this may well require a lot of of states if the blocks are large (in dimension), so that this approach becomes impractical for problems with large (in dimension) repeated scalar uncertainty blocks.

Fortunately for many problems we do not need repeated scalar uncertainty blocks, and so we only need diagonal scaling matrices. In fact the currently available software in the μ-Tools toolbox (for complex μ synthesis) only handles this case. For this reason most of the effort in our approach has been to develop a method which will work well for that case (hence the emphasis on developing good routines for fitting the diagonal elements of $D(w)$ and $\mathbf{j}G(w)$). A practical solution to the more general case is a subject of ongoing research, and this is discussed in more detail in chapter 12.

Chapter 11

Applications and Examples

In the course of studying the mixed μ problem we have become involved in a fair amount of mathematical abstraction. Nevertheless the underlying motivation for this research is the engineering problem of trying to develop control strategies and methodologies for real physical systems. The mixed μ analysis and synthesis techniques developed here have been implemented in software, and the bounds computation routines are now commercially available. This has already lead to a number of interesting applications, and in this chapter we will briefly look at some of these.

11.1 A Simple Numerical Example

Before tackling any real engineering problems, we first present a simple (constant matrix) example. The intention here is simply to provide the reader with a numerical example which illustrates the techniques involved in mixed μ analysis. Consider the following matrix:

$$M = \begin{pmatrix} 0.41 - 0.73\mathbf{j} & 0.32 - 0.21\mathbf{j} & -0.51 + 0.35\mathbf{j} \\ -0.98 - 0.45\mathbf{j} & 1.52 + 0.56\mathbf{j} & 0.89 + 1.13\mathbf{j} \\ -1.02 - 0.61\mathbf{j} & 0.75 - 1.06\mathbf{j} & -0.25 + 0.15\mathbf{j} \end{pmatrix}.$$

We will consider structured uncertainties with two real scalars and one complex scalar, i. e., we have $\Delta \in X_{\mathcal{K}}$ where the block structure takes the form

$$X_{\mathcal{K}} = \{ \operatorname{diag}(\delta_1^r, \delta_2^r, \delta^c) : \delta_1^r, \delta_2^r \in \mathbb{R}, \delta^c \in \mathbb{C} \}.$$

Using the "rmu" software from chapter 6 to compute upper and lower bounds we find that

$$2.1768 \quad \leq \quad \mu_{\mathcal{K}}(M) \quad \leq \quad 2.1802$$

so that for this example the bounds are tight. The upper bound is verified by the fact that we have matrices $D \in \mathcal{D}_{\mathcal{K}}, G \in \mathcal{G}_{\mathcal{K}}$ given as

$$D = \operatorname{diag}(3.7139 \times 10^{-4}, 0.8415, 1) \qquad , \qquad G = \operatorname{diag}(-6.2604, 1.4106, 0)$$

for which one may check that

$$\overline{\lambda}(M^* D M + \mathbf{j}(GM - M^* G) - \beta_{up}^2 D) = 0$$

where $\beta_{up} = 2.1802$. These D, G scaling matrices were obtained by solving the (convex) LMI minimization problem from theorem 4.5 via the machinery of algorithm 6.2. Of course, as with any practical algorithm, we do not pursue this minimization to its exact optimal solution, but rather quit at some pre-specified level of accuracy. In fact, in this case, it appears that the infimum in the upper bound is not achieved, since one can see that D_1 is tending to zero, so that the D scaling matrix is tending towards singularity. This can also be seen if we use the machinery of theorem 6.1 to convert these scaling matrices into ones suitable for the upper bound from theorem 6.2. Then we find that we have $\overline{D} \in \overline{\mathcal{D}}_{\mathcal{K}}, \overline{G} \in \overline{\mathcal{G}}_{\mathcal{K}}$ given as

$$\overline{D} = \operatorname{diag}(1.6946, 1.0303, 1) \qquad , \qquad \overline{G} = \operatorname{diag}(-7.7317 \times 10^3, 0.7689, 0)$$

for which one may verify that

$$\overline{\sigma}\left((I_3 + \overline{G}^2)^{-\frac{1}{4}} \left(\frac{\overline{D} M \overline{D}^{-1}}{\beta_{up}} - \mathbf{j}\overline{G} \right) (I_3 + \overline{G}^2)^{-\frac{1}{4}} \right) = 1.$$

Note that in these coordinates the $\overline{D}$ matrix is not becoming ill conditioned, but $\overline{G}_1$ is becoming unbounded, so that here the $\overline{G}$ scaling matrix is blowing up, and once again the infimum is not achieved.

The lower bound is obtained via the power iteration from section 4.5. The algorithm converges and we recover the vectors b, z, a, w as

$$b = \begin{pmatrix} -0.1157 + 0.0837\mathbf{j} \\ 0.5332 - 0.6535\mathbf{j} \\ -0.4529 - 0.2092\mathbf{j} \end{pmatrix} \quad , \quad z = \begin{pmatrix} -0.2168 - 0.3001\mathbf{j} \\ 0.0585 - 0.7398\mathbf{j} \\ 0.0082 - 0.4273\mathbf{j} \end{pmatrix},$$

$$a = \begin{pmatrix} 0.1614 - 0.1167\mathbf{j} \\ 0.5333 - 0.6535\mathbf{j} \\ 0.0096 - 0.4988\mathbf{j} \end{pmatrix} \quad , \quad w = \begin{pmatrix} 0.3023 + 0.4185\mathbf{j} \\ 0.0585 - 0.7399\mathbf{j} \\ -0.3880 - 0.1792\mathbf{j} \end{pmatrix}.$$

These vectors satisfy the lower bound decomposition equations in (4.18), namely

$$\begin{aligned} Mb = \beta_{low} a \qquad & M^* z = \beta_{low} w \\ b = Qa \qquad & b = \hat{D}^{-1} w \\ z = Q^* Q \hat{D} a \qquad & z = Q^* w, \end{aligned}$$

where $\beta_{low} = 2.1768$, and the matrices $Q \in \mathcal{Q}_{\mathcal{K}}, \hat{D} \in \hat{\mathcal{D}}_{\mathcal{K}}$ are given by

$$\begin{aligned} Q &= \operatorname{diag}(-0.7172, 0.9999, 0.4017 - 0.9158\mathbf{j}) \\ \hat{D} &= \operatorname{diag}(0.0025 - 3.6138\mathbf{j}, 0.7236 - 0.5009\mathbf{j}, 0.8566). \end{aligned}$$

Note that we have

$$\begin{aligned} \hat{D}_1 = 3.6138 e^{-1.5701\mathbf{j}} \quad &\longrightarrow \quad \theta_1 = -89.96^\circ \\ \hat{D}_2 = 0.88 e^{-0.6055\mathbf{j}} \quad &\longrightarrow \quad \theta_2 = -34.69^\circ \end{aligned}$$

which satisfies the alignment criterion since $|q_1^r| < 1$ whereas $|q_2^r| = 1$ (approximately). Note also that the presence of an internal real variable fits in with our earlier observation that it appears that the upper bound is tight, but the infimum is not achieved (see theorem 6.3).

For this problem we find that the maximum eigenvalue in the upper bound function is distinct. It follows from theorem 5.9 that for this problem the upper bound equals μ. In fact if we employ the "μ-values" construction from section 5.5, applied to the lower bound decomposition, we obtain $\tilde{D} \in \mathcal{D}_{\mathcal{K}}, \tilde{G} \in \mathcal{G}_{\mathcal{K}}$ as

$$\tilde{D} = \mathrm{diag}\,(2.9086 \times 10^{-3}, 0.8447, 1) \quad , \quad \tilde{G} = \mathrm{diag}\,(-6.5863, 1.2726, 0)$$

for which

$$\overline{\lambda}(M^*\tilde{D}M + \mathbf{j}(\tilde{G}M - M^*\tilde{G}) - \hat{\beta}_{up}^2\tilde{D}) = 0$$

with $\hat{\beta}_{up} = 2.1770$. Of course in general it is not guaranteed that the bounds will be tight on any given problem, so that further computation (via Branch and Bound) may be required.

11.2 Engineering Application of Analysis Techniques

In this section we briefly present two engineering problems to which mixed μ analysis techniques have been applied, one from the area of flight control, and one from flexible structures. For more details on these two examples see [6,8].

Missile Autopilot Problem

The control design of a longitudinal missile autopilot is considered in [6]. The objective is to control the missile pitch via tail fin deflection. The performance specifications, in terms of steady-state error, overshoot etc. are given in [6]. What is of interest here is that it is desired that a single controller should be able to meet the required performance specifications over a range of $\pm 20°$ for the angle of attack, and missile speeds between Mach 2 and 4. Furthermore the controller must be able to cope with unmodeled high frequency flexible body modes in the missile.

Having designed a candidate controller the closed loop system may be recast into the standard μ analysis framework (see figure 10.1). This results in an interconnection

transfer matrix, $M(s)$, with 16 states, 17 inputs, and 16 outputs. The uncertainty structure is given as

$$X_{\mathcal{K}} = \{ \text{block diag}\,(\delta_1^c, \delta_2^c, \delta_1^r I_4, \delta_2^r I_8, \Delta^C) : \delta_i^c \in \mathbb{C}, \delta_i^r \in \mathbb{R}, \Delta^C \in \mathbb{C}^{3\times 2} \}.$$

Note that here the blocks do not come in the usual order, and the full complex block is not square. However, as we mentioned earlier, all the results are easily generalized to this case (though somewhat more messy to present), and the software implementation allows this extra generality. The block structured uncertainty thus consists of 5 blocks, and includes both real and complex uncertainties. The first two are complex scalar uncertainties associated with additive and multiplicative uncertainties in the plant model, and are used to cover unmodeled dynamics, as described above. The two repeated real scalar uncertainty blocks are used to cover perturbations in angle of attack and Mach number. Finally we have a full complex block associated with an $\mathcal{H}_\infty$ norm performance criterion (see theorem 2.4).

The "rmu" software is used to compute upper and lower bounds for $\mu_{\mathcal{K}}(M(\mathrm{j}w))$ across a frequency range of interest. The robust performance mixed μ plots for this problem, and the associated complex μ problem (simply 'covering' the real uncertainties with complex ones), are shown in figure 11.1. These plots determine the robustness (in performance) of the control design to the specified perturbations (see theorem 2.4).

Note that for both the mixed and complex μ problems the bounds capture the peak value of μ across frequency quite well (though the mixed μ bounds are somewhat less accurate in certain frequency ranges). It can also be seen that mixed μ is quite different to complex μ. In particular the complex μ approximation to the problem indicates that the controller robustness properties are poor around 40 rad/s (where complex μ peaks), whereas the mixed μ analysis indicates that this frequency is not a particular problem. Simulations of the closed loop system, using a more sophisticated nonlinear model of the missile, supported the controller/performance predictions from the mixed μ bounds. A number of different designs were carried out with similar

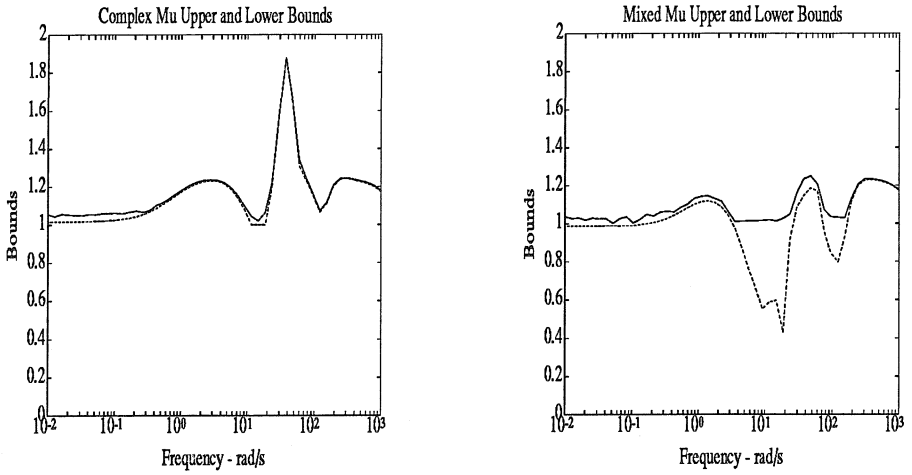


Figure 11.1: Robust performance μ plots for the missile autopilot problem

results. In fact it was found that designs which improved (the peak value of) complex μ at the expense of mixed μ had degraded performance in closed loop simulations (see [6] for details).

Flexible Structure Problem

The NASA Langley Minimast facility is a deployable/retractable beam truss, designed to represent future deployable space trusses. This flexible structure has 18 bays (total length 20.2m), and is equipped with accelerometers and torque wheel actuators, so that it provides an excellent testbed for applied control research. For more details of the facility see [8] and the references therein.

This set-up is used to investigate a variety of control problems in [8], one of which we will briefly examine here. The basic control problem is to actively damp out the vibration due to an impulse input disturbance of the structure. The open loop response due to the disturbance is shown in figure 11.2. This is experimental data, and the responses are recorded in three different directions (A,B,C) at two different bays (18 and 10) on the structure. It can be seen from these responses that, without

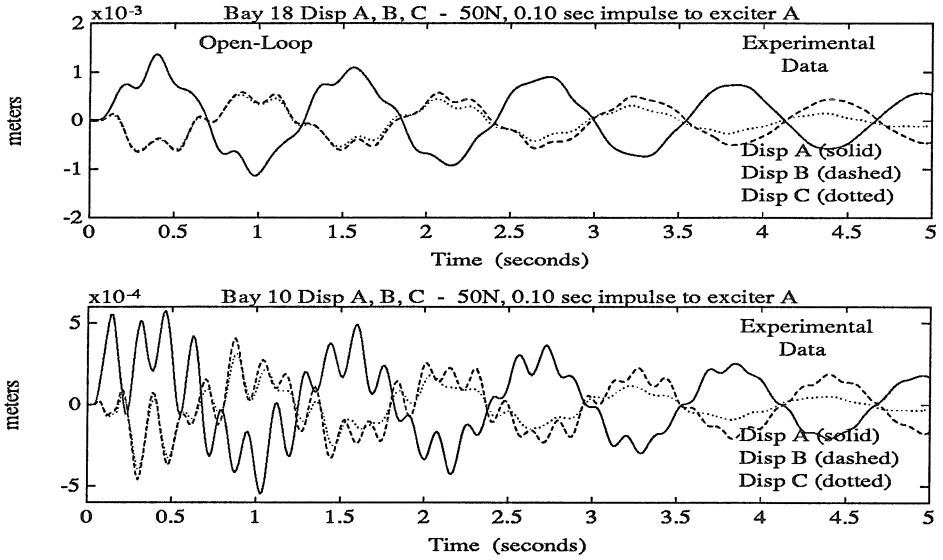
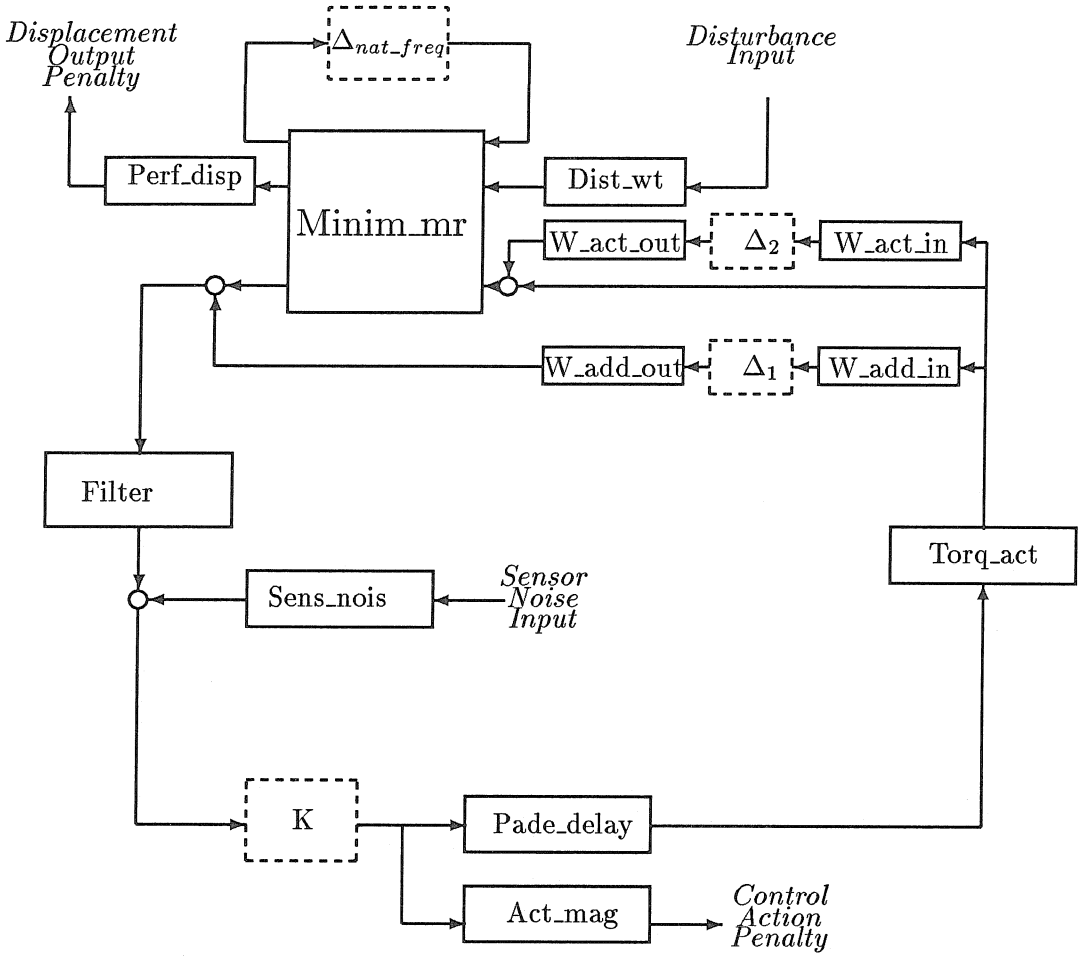


Figure 11.2: Experimental open loop response of Langley Minimast structure to impulse input disturbance

active control, the structure has a number of very lightly damped modes.

An accurate model of the structure is available. However, in order to investigate the possibility of designing fairly high performance controllers *without* accurate plant models, the following study was performed. The plant model was deliberately made "incorrect," by altering the natural frequencies of the first 5 modes in the model by 20%, and controllers were then designed using this "incorrect" model. The controllers were designed to be robust with respect to the interconnection shown in figure 11.3. For details of the actuator, sensor, and plant models, together with the selection of appropriate weights, (and the design procedure) see [8].

Note that the natural frequency variations are covered with real scalar uncertainties. In addition there is an additive uncertainty, Δ_1, to cover unmodeled high-order dynamics in the structure model, and a multiplicative uncertainty, Δ_2, to cover inaccuracies in the actuator model. Together with a performance block for the $\mathcal{H}_\infty$ norm



$$\Delta_{nat\_freq} = \begin{bmatrix} \delta_1^r & 0 & 0 & 0 & 0 \\ 0 & \delta_2^r & 0 & 0 & 0 \\ 0 & 0 & \delta_3^r & 0 & 0 \\ 0 & 0 & 0 & \delta_4^r & 0 \\ 0 & 0 & 0 & 0 & \delta_5^r \end{bmatrix}$$

Figure 11.3: Interconnection for robust performance mixed μ analysis of the Langley Minimast structure

performance criterion, this gives the uncertainty description as

$$X_{\mathcal{K}} = \{\,\text{block diag}\,(\delta_1^r, \delta_2^r, \delta_3^r, \delta_4^r, \delta_5^r, \Delta_1^C, \Delta_2^C, \Delta_3^C) : \\ \delta_i^r \in \mathbb{R}, \Delta_1^C \in \mathbb{C}^{4\times 3}, \Delta_2^C \in \mathbb{C}^{3\times 3}, \Delta_3^C \in \mathbb{C}^{7\times 9}\}.$$

This leads to a mixed μ problem with 5 real scalar uncertainties and 3 full complex blocks, which is used to evaluate the candidate controllers.

The "rmu" software is used to compute upper and lower bounds for this mixed μ problem across a frequency range of interest. The robust performance mixed μ plots for this problem, and the associated complex μ problem, are shown in figure 11.4. These plots determine the robustness (in performance) of the control design to the

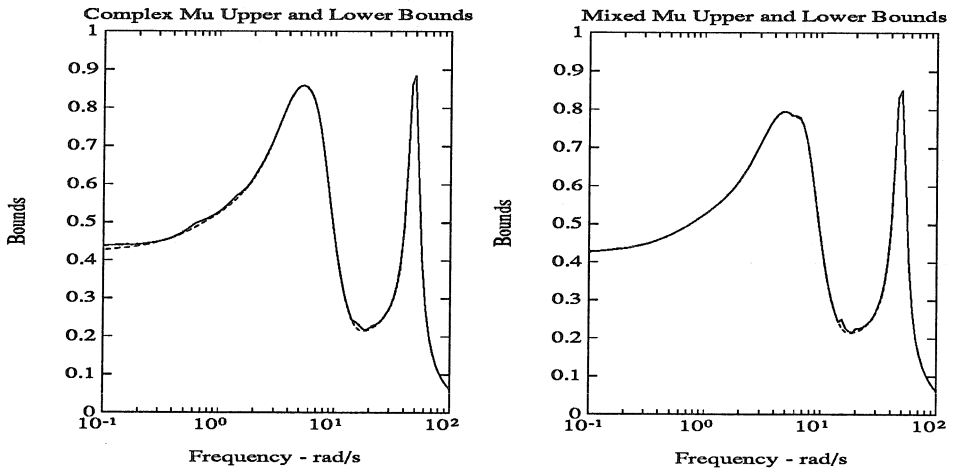


Figure 11.4: Robust performance μ plots for the flexible structure problem

specified uncertainties in the natural frequencies, and unmodeled dynamics.

Note that for this problem the upper and lower bounds for mixed or complex μ are tight across the entire frequency range. In either case we find that the μ computation predicts that this control design is indeed robust with respect to the

prescribed uncertainties. This controller was implemented on the structure, and the closed loop response to the impulse input disturbance is shown in figure 11.5. On

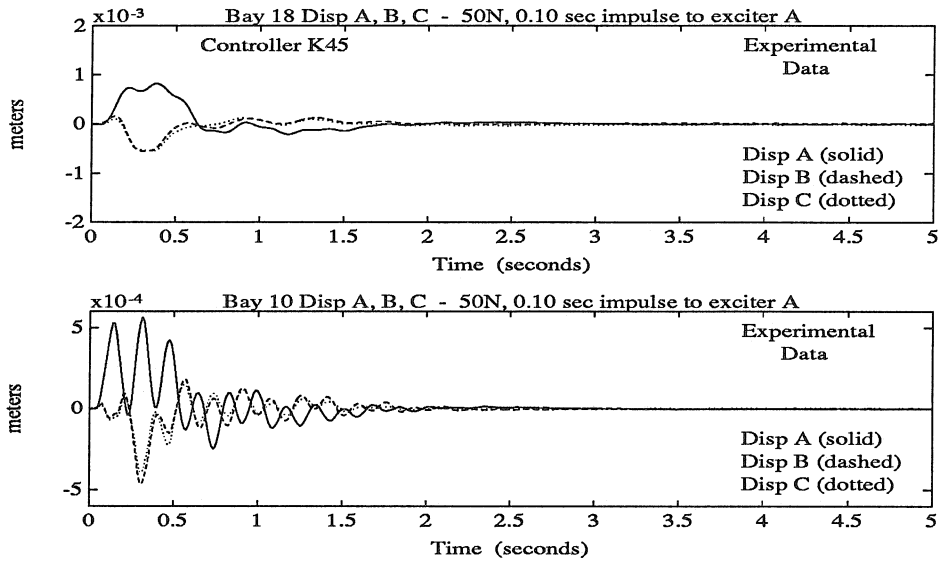


Figure 11.5: Experimental closed loop response of Langley Minimast structure to impulse input disturbance

comparison with figure 11.2 it is readily seen that the damping of the structure has been substantially improved by active control (see [8] for a detailed analysis). Note once more that this level of (experimental) controller performance was achieved using an inaccurate model of the structure, by ensuring the controller was robust to these inaccuracies.

Interestingly in this case, because of the way the uncertainties entered the system, mixed and complex μ are seen to be very close (see [8] for a physical interpretation of these results). Of course this is not the case in general (see the previous example for instance), so that it is important to have both upper and lower bounds. The upper bound provides a robustness guarantee, and the lower bound provides a measure of the

potential conservatism of the upper bound (together with a "problem perturbation").

The "rmu" software is now available commercially as part of the μ-Tools software package. This software is currently being used at several industrial and academic sites, including Honeywell, NASA Dryden, Phillips, Caltech, U. C. Berkeley, University of Minneapolis, and Cambridge University. This has already lead to a number of engineering applications in addition to the above examples, on problems ranging from C. D. players to the Space Shuttle. For these problems the software has worked well, providing reasonably tight bounds for the associated mixed μ problems.

11.3 Synthesis and Controller Design

In this section we will look at some fairly simple examples of applying the mixed μ synthesis techniques described in chapters 9 and 10 to design controllers.

Optimal Gain Margin Problem

The problem of designing a controller to maximize the gain margin of a SISO plant can be solved analytically (see section 11.3 of [26]). This question can (almost) be restated as a special case of a mixed μ synthesis problem, so that it provides us with a simple benchmark test for the "D,G-K iteration" procedure.

Consider the feedback interconnection in figure 11.6. The perturbed plant $\tilde{P}$ is given by

$$\tilde{P}(s) = \left(\frac{s-1.2}{1-1.2s}\right)(1+\delta^r)$$

where $\delta^r \in \mathbb{R}$ is a real scalar uncertainty. This set-up fits into the standard mixed μ synthesis framework of figure 10.1. If we choose the performance weights as

$$W_{S_1} = W_{S_2} = W_{T_1} = W_{T_2} = 0.01$$

then the performance requirements for this problem are negligible. Thus we find that this robust performance problem is effectively a robust stability problem, and is in fact (approximately) the optimal gain margin problem.

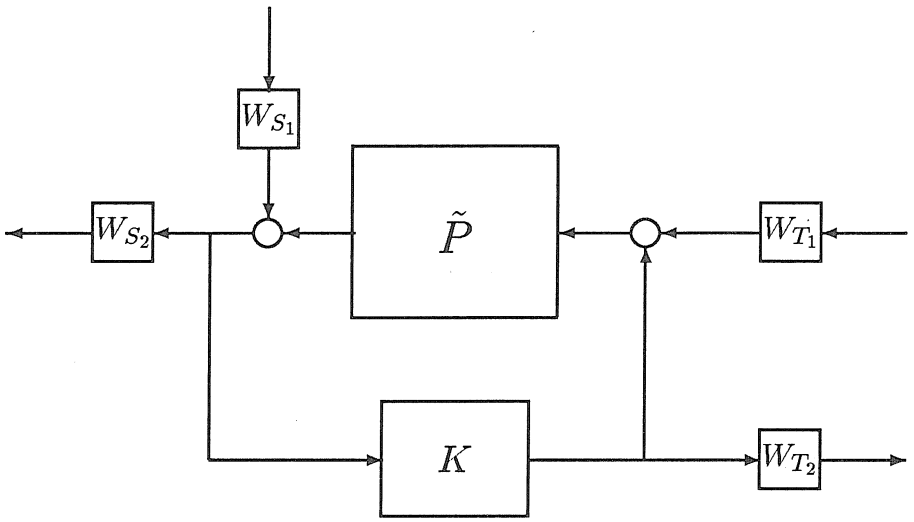


Figure 11.6: Interconnection for mixed μ synthesis examples

Note that the nominal plant is non-minimum phase and unstable. It can be stabilized simply with unity negative feedback, and the Nyquist plot is shown in figure 11.7 (for $\delta^r = 0$). In addition the mixed and complex μ plots across frequency for the closed loop system are also shown. Note that in this example the upper and lower bounds were tight, so that only the upper bounds are shown. For a plant of the form

$$P = \frac{s - \alpha}{1 - \alpha s}$$

with $\alpha > 1$, then the optimal gain margin controller is given by [25]

$$K = -\frac{1 + \beta s}{s + \beta}$$

where $\beta \downarrow \frac{1}{\alpha}$. The results for the optimal gain margin controller are shown in figure 11.8, and it can be seen that this controller flattens out the Nyquist curve so as to substantially improve the gain margin. As a result the peak value across frequency of mixed μ is substantially improved (from 6.0 to 3.35), at the expense of complex μ.

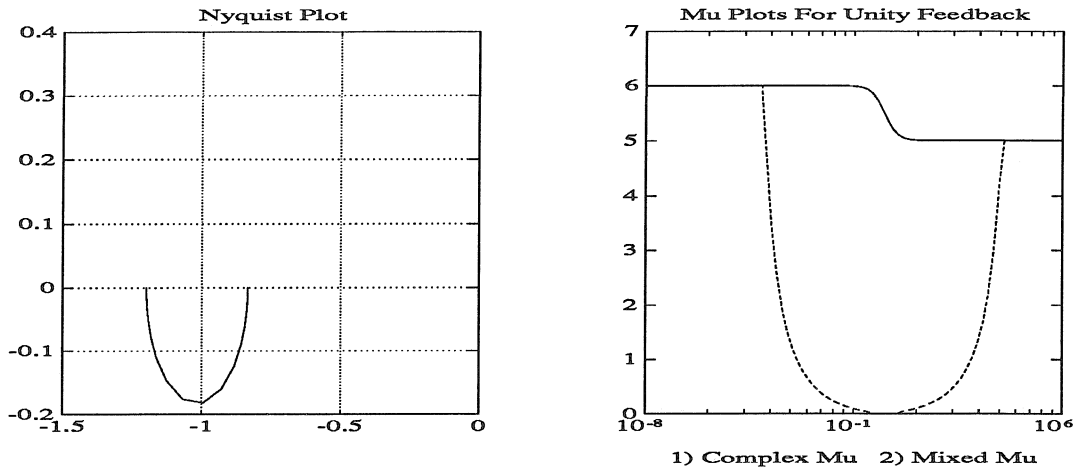


Figure 11.7: Nyquist and μ plots for unity (negative) feedback controller (gain margin problem)

The standard "D-K iteration" procedure for complex μ synthesis was applied to this problem, and the results are shown in figure 11.9. Note that this controller gives a slight improvement over unity feedback for both mixed and complex μ (from 6.0 to 5.57). However the Nyquist plot is not greatly modified (other than the roll off at high frequency), so that the gain margin is not greatly improved. Applying the mixed μ synthesis procedure to this problem results in the plots in figure 11.10. Note that this controller achieves about the same level for mixed μ as the optimal (gain margin) controller. Once again this is achieved by flattening out the Nyquist curve so as to improve the gain margin at the expense of complex μ. For this simple example we see that the mixed μ synthesis procedure does exactly the right thing, and shapes the frequency response so as to minimize the peak value across frequency of mixed μ.

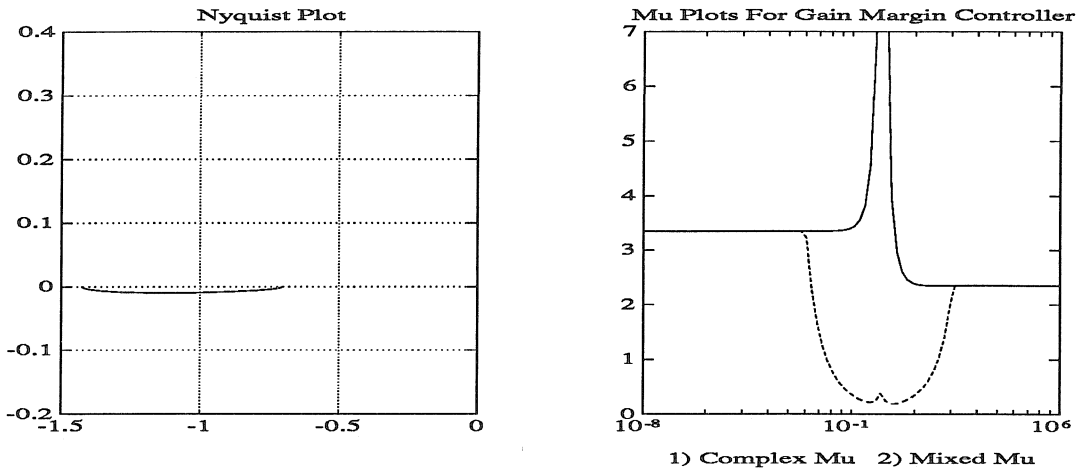


Figure 11.8: Nyquist and μ plots for optimal gain margin controller (gain margin problem)

C. D. Player Problem

The previous example was simple enough that it was possible to solve for the optimal controller. Thus we could compare the mixed μ synthesis controller to the optimal solution, and interpret the results in terms of the Nyquist plot. However this example (effectively) included no performance requirements, so that although it provides a nice test of the procedures involved, it does not result in a practical controller.

Consider now an uncertain plant of the form

$$\tilde{P}(s) = \frac{1}{s^2}(5.05 + 4.95\delta^r)$$

where $\delta^r \in \mathbb{R}$ is a real scalar uncertainty. This example is from [79], and is motivated by an idealized model of the dynamics of a C. D. player servo. Note that it consists of a double integrator with an uncertain gain, i. e., $\tilde{P}(s) = \frac{k_p}{s^2}$ where k_p is variable. The problem is scaled so that for $\delta^r \in [-1\ 1]$ we have $0.1 \leq k_p \leq 10$.

It is desired to control this plant for any value of the gain k_p in the above range. The performance requirement is to maximize the closed loop bandwidth of the system,

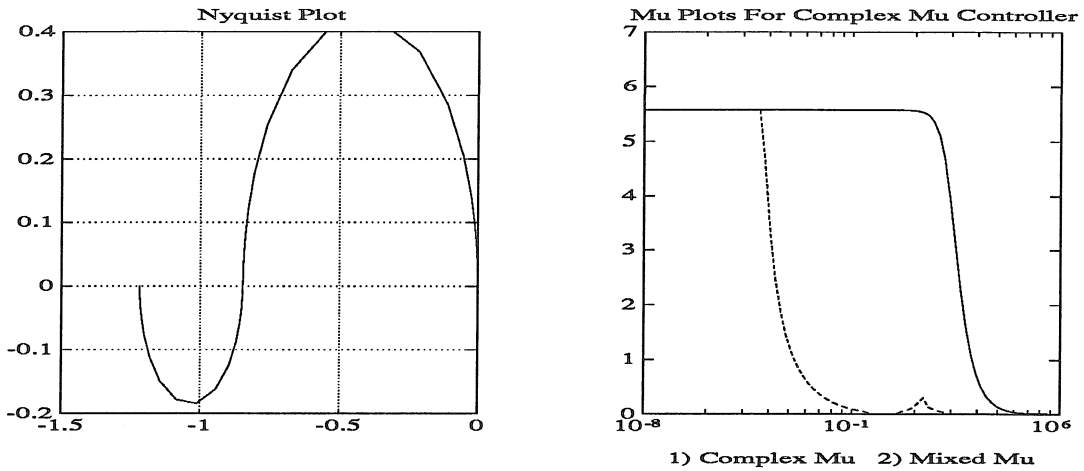


Figure 11.9: Nyquist and μ plots for complex μ synthesis controller (gain margin problem)

for a certain tolerated level of sensitivity. In effect this maximizes the speed of the step response for a certain tolerated level of overshoot. At the same time the control action must not be too large, and the design must work for all allowable values of k_p. This is discussed in more detail in [79], and the sensitivity and complementary sensitivity weights selected to reflect these requirements were

$$\begin{aligned} W_{S_1} = W_{S_2} &= \frac{0.03}{s + 0.05} \\ W_{T_1} = W_{T_2} &= \frac{5\,(s + 0.001)}{s + 5} \end{aligned}$$

for the interconnection shown in figure 11.6.

First of all a complex μ synthesis controller was designed for this system and the results are shown in figure 11.11. The Nyquist plot is shown for $k_p = 1$, together with complex and mixed μ plots across frequency. Once again the upper and lower bounds were tight so that only the upper bounds are shown. Note that, as one might expect, this controller flattens out complex μ so as to minimize it's peak value (at 1.54), with the peak value for mixed μ being about the same.

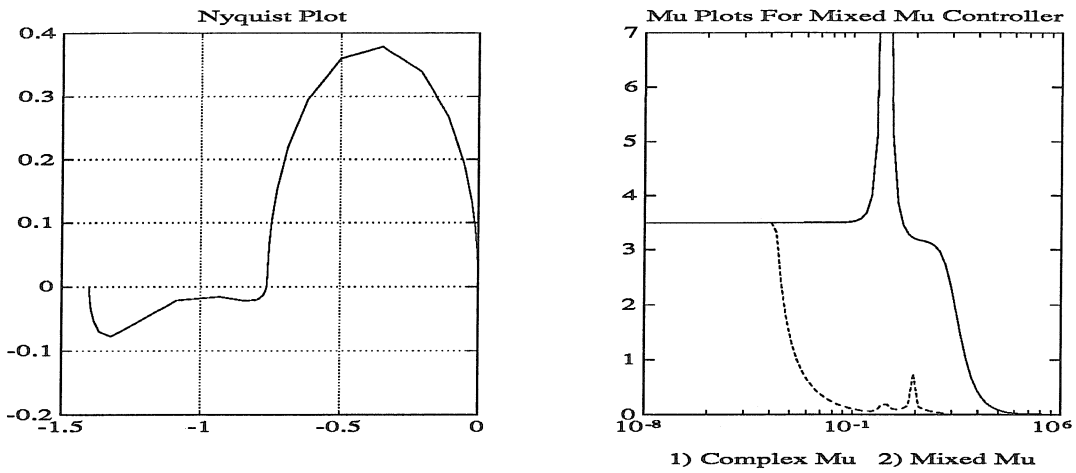


Figure 11.10: Nyquist and μ plots for mixed μ synthesis controller (gain margin problem)

Note that since we are required to stabilize a double integrator, the controller must introduce some phase lead (with negative feedback). Furthermore this phase lead must be maintained over a wide frequency range, so as to allow for the gain variation in the plant. At the same time we do not want to introduce too much gain over this frequency range, since we want to keep the control action reasonable. This implies that we do not want the *maximum* phase lead to be too large. From a classical viewpoint then it is clear that the ideal strategy is to use a controller whose phase response is approximately flat over the desired frequency range. This will correspond to a Nyquist plot for the compensated system which is approximately a straight line near the critical point. Thus the gain variation merely changes the frequency calibration of the Nyquist curve, but does not alter it's shape. In this way the controller ensures that the shape of the step response is invariant to the plant gain (within a certain range), and only the speed of response changes. It then remains to choose the slope of this line (i. e., the phase margin of the compensated system) so as to give the desired tradeoff between overshoot and speed of response. We will refer

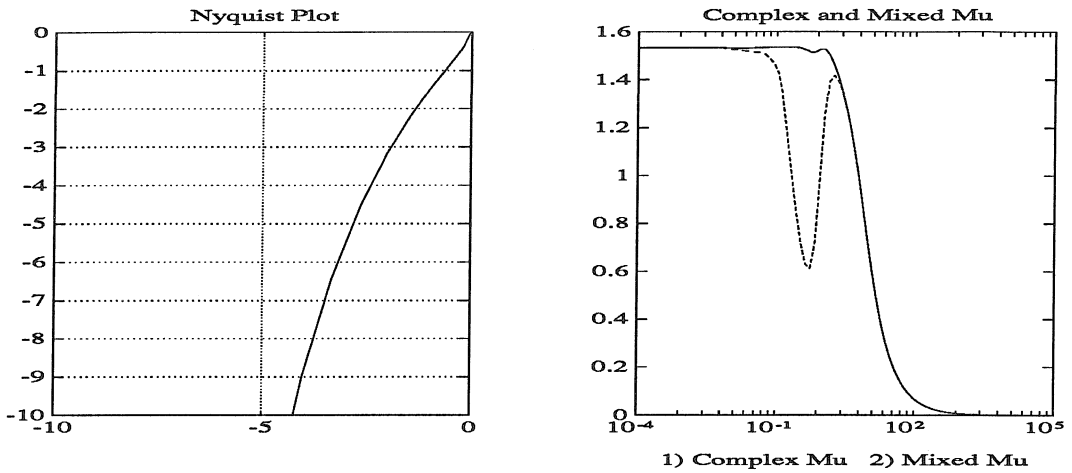


Figure 11.11: Nyquist and μ plots for complex μ synthesis controller (C.D. player problem)

to this controller as the Bode controller, and for more details of its motivation and design see [79] and the references therein.

The exact realization of the Bode controller is infinite dimensional, but it is readily approximated as

$$K = -0.1\frac{(s+\frac{1}{32})(s+\frac{1}{8})(s+\frac{1}{2})}{(s+\frac{1}{16})(s+\frac{1}{4})(s+1)}.$$

The results obtained with the Bode controller are shown in figure 11.12. Note that although this controller is poor with respect to complex μ (peaks around 2.65), it actually improves on mixed μ (peaks around 1.42). Of course this is what one would expect since the design process for the Bode controller assumes that the uncertain gain is a real parameter, rather than a disk uncertainty.

Finally a mixed μ synthesis controller was designed for this problem, and the results are shown in figure 11.13. This controller looks very similar to the Bode controller. It actually achieves the lowest level of mixed μ (peaks around 1.25), and once again this is achieved at the expense of complex μ.

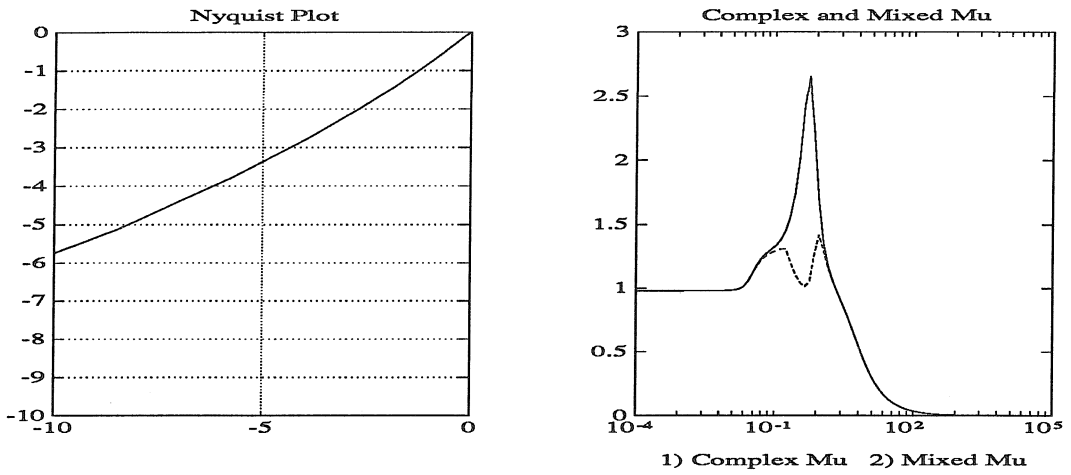


Figure 11.12: Nyquist and μ plots for Bode controller (C. D. player problem)

For this problem we find that the mixed μ synthesis routine comes up with a controller design that emulates the performance of the classical design. The step responses for both compensated systems, for a number of values of the plant gain, are shown in figure 11.14. It can be seen that the responses are very comparable, and in fact they represent slightly different choices for the tradeoff, with the mixed μ synthesis controller being a little slower, but with less overshoot. The designs could be made to look even closer by fine tuning the weights if desired.

A Modal Example

The previous two examples involved only one uncertain real parameter, so that it was possible in both cases to exploit our knowledge of the problem, and find a good controller design by other (simpler) methods for the purposes of comparison. In both cases we found that the mixed μ synthesis routine was able to find a comparable design. Of course the main point in developing the mixed μ synthesis machinery was to have a systematic design procedure, capable of dealing with more complicated problems, where the earlier simple approaches are no longer available. We conclude this

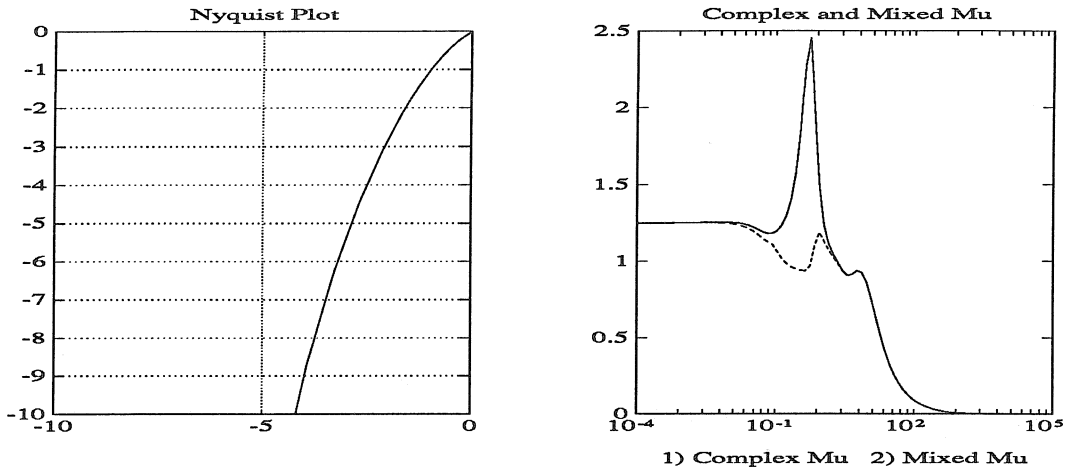


Figure 11.13: Nyquist and μ plots for mixed μ synthesis controller (C. D. player problem)

chapter with a synthesis example for a problem with two uncertain real parameters.

Consider once again the interconnection in figure 11.6. As in the gain margin problem we make this essentially a robust stability problem by choosing the weights as

$$W_{S_1} = W_{S_2} = W_{T_1} = W_{T_2} = 0.01$$

so that the performance requirements are negligible. Now consider an uncertain plant of the form

$$\tilde{P}(s) = \frac{s^2 - 2\zeta\omega s + \omega^2(1+\delta_1^r)}{s^2 + 2\zeta\omega s + \omega^2(1+\delta_2^r)}$$

where $\delta_1^r, \delta_2^r \in \mathbb{R}$ are uncertain real parameters. This is a second-order transfer function intended to represent a lightly damped mode in a flexible structure. This can be put in standard μ synthesis form by rewriting as $\tilde{P} = F_u(G, \Delta)$ where $\Delta =$

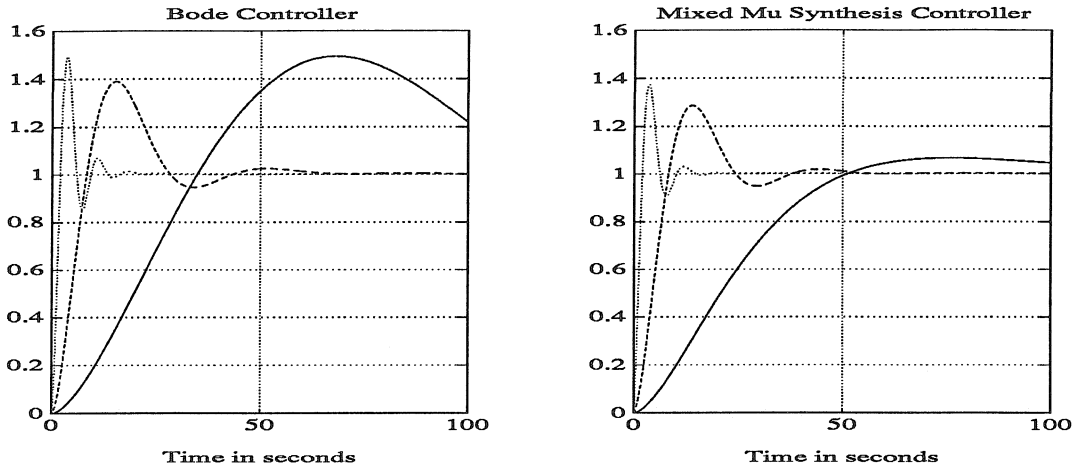


Figure 11.14: Closed loop step responses for $k_p = 0.1,1,10$ for the Bode and mixed μ synthesis controllers (C. D. player problem)

diag (δ_1^r, δ_2^r) and

$$G = \left[\begin{array}{cc|ccc} 0 & 1 & -\frac{1}{\omega^2} & -\frac{1}{\omega^2} & 0 \\ -\omega^2 & -2\zeta\omega & 0 & 0 & -4\zeta\omega \\ \hline 0 & 0 & 0 & 0 & \omega^2 \\ 0 & -\omega^2 & 0 & 0 & -\omega^2 \\ 0 & 1 & 0 & 0 & 1 \end{array}\right].$$

For this example we choose $\omega = 1$ and $\zeta = 0.2$, so that the mode is lightly damped, with a pair of complex conjugate poles at $-0.2 \pm 0.9798\mathbf{j}$ and a pair of complex conjugate zeros at $0.2 \pm 0.9798\mathbf{j}$ (for $\delta_1^r = \delta_2^r = 0$).

Consider for a moment what happens to these pole/zero locations as we start to move δ_1^r, δ_2^r away from zero. If we restrict $\delta_1^r, \delta_2^r \in \mathbb{R}$ then we get the motion indicated by the arrows in figure 11.15, so that initially the poles stay in the left half-plane and the zeros stay in the right half-plane. However if we allow $\delta_1^r, \delta_2^r \in \mathbb{C}$ then we get motion inside the indicated circles. In this case we find that the poles and zeros can cross the imaginary axis for much smaller values of the uncertainty. Thus we

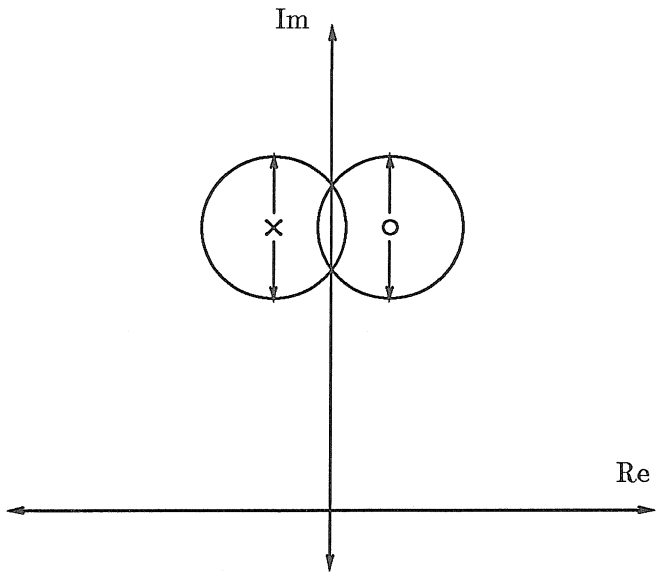


Figure 11.15: Pole/zero motion for the modal example

would expect to see a fairly large gap between mixed and complex μ for this type of problem.

A complex μ synthesis controller was designed for this problem and the complex and mixed μ plots are shown in the right-hand plot of figure 11.16. Once again only the upper bounds are shown. This controller achieves a peak value for complex μ of about 2.6, with mixed μ being about the same. The results for the mixed μ synthesis controller are shown in the left-hand plot of figure 11.16. This controller does not greatly affect the peak value of complex μ, but reduces the peak value of mixed μ to about 1.1. Thus we do indeed find that there is a large gap between mixed and complex μ for this problem, and the mixed μ synthesis procedure is able to exploit that fact to significantly improve on the complex μ synthesis controller.

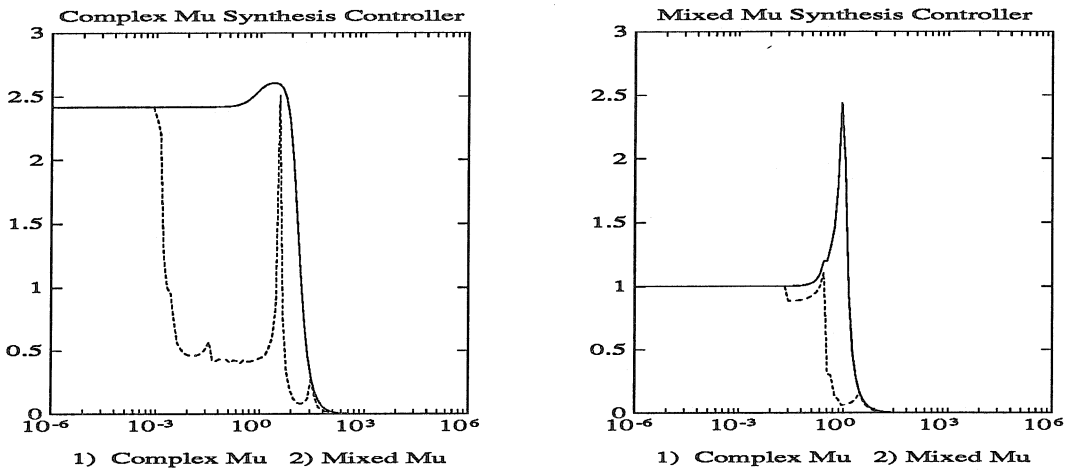


Figure 11.16: Complex and mixed μ plots for the modal example

Chapter 12

Concluding Remarks

We conclude the thesis with a brief summary of the work, and a look at some related topics, together with suggestions for future research.

12.1 Summary

In recent years a great deal of interest has arisen with regard to robustness problems involving parametric uncertainty. These problems involve uncertain parameters that are not only norm bounded, but also constrained to be real. They arise naturally in many settings when one wishes to allow for a possible *range* of values for certain parameters in the model (masses, aerodynamic coefficients etc.).

We have seen that robustness problems involving parametric uncertainty can be reformulated as mixed μ problems, where the block structured uncertainty description is now allowed to contain both real and complex blocks. The complex μ problem (where the block structured uncertainty description contains only complex blocks) has seen substantial progress in the last ten years [57]. Although it is by no means completely solved, analysis and synthesis techniques have been developed, and efficient computational software is now available [7], so that complex μ techniques are now routinely applied to large engineering problems.

Unfortunately complex μ can be a poor approximation for mixed μ, so that these

techniques are no longer adequate when the robustness problem includes parametric uncertainty. The main goal of the work described in this thesis was to develop analysis and synthesis techniques for the mixed μ problem.

It has been shown that mixed μ inherits many of the (appropriately generalized) properties of complex μ. However, in some aspects the mixed μ problem can be fundamentally different from the complex μ problem, and in particular the general mixed μ problem is NP hard. This prohibits its exact computation, except for small problems or special cases. Nevertheless practical computation schemes for the mixed μ analysis problem were developed, which yield upper and lower bounds.

These bounds take the form of generalizations of the bounds for the complex μ problem (i. e., by applying the mixed μ bounds to complex μ problems, one recovers the standard complex μ bounds). They rely heavily on exploiting the theoretical properties of the problem, so as to come up with approximations that are both easy to compute, and usually fairly accurate. In fact the upper (and lower) bounds reduce to exact expressions for μ for any of the special cases that we know how to solve.

One such example is the rank one μ problem, a special case of which is equivalent to the "affine parameter variation" problem for a polynomial with perturbed coefficients. Variations on this problem have received a good deal of attention in the literature, and a number of "Kharitonov-type" exact stability results have been obtained (see [71,10] for surveys). It was shown that for this problem μ is identically equal to its upper bound (which is convex) so that exact analysis is also possible in the μ framework for this special case. In fact it is possible to solve for the optimal scaling matrices, so that one obtains an analytic solution for the rank one mixed μ problem.

Of course the NP-hardness results preclude this type of approach (or any other exact approach) for the general problem. In that case, if the upper and lower bounds are not sufficiently accurate, one may use a Branch and Bound scheme to iteratively refine them (at extra computational cost). Using this approach a Branch and Bound algorithm was developed, which is capable of handling fairly large problems (tens

of parameters) with reasonable computational requirements, and typically produces answers with acceptable accuracy. This is despite the fact that the mixed μ problem appears to have inherently combinatoric worst-case behavior.

These analysis tools have been implemented in software which is now (partially) available commercially as part of the μ-Tools software package. This software is currently being used at several industrial and academic sites, including Honeywell, NASA Dryden, Phillips, Caltech, U. C. Berkeley, University of Minneapolis, and Cambridge University. This has already lead to a number of engineering applications, and in particular several studies in the areas of flexible structures (e. g., NASA Langley Minimast) and flight control (e. g., Space Shuttle) have utilized the above tools.

The mixed μ synthesis problem, that of designing a controller to minimize mixed μ, has also been studied. A "D,G-K iteration" procedure was developed, which finds a stabilizing controller achieving a local minimum of the problem (which is not convex). Once again this takes the form of a generalization of the appropriate procedure for complex μ. Initial applications of this procedure have been very encouraging, and produced controllers with a lower (peak) value for mixed μ than one could obtain with standard complex μ synthesis techniques.

The theoretical tools developed here, together with the associated computational software, provide the systematic means to handle a broad class of robustness problems which include both parametric and dynamic uncertainty. This should aid the engineer in understanding the fundamental tradeoffs and limitations inherent to a given physical system, and in designing an appropriate controller.

12.2 Related Work and Future Directions

The mixed μ problem has important connections to a number of other topics in systems theory. Exploring these connections more fully, as well as extending the

current theoretical and computational tools for mixed μ, provides a number of avenues for further research:

Real Full Blocks: The uncertainty description allowed for in the present theory includes repeated real scalars, repeated complex scalars, and full complex blocks. This is a very general set-up. In some cases however it would be useful to be able to consider even more general classes of uncertainty, and one such example is real full blocks, i. e., $\Delta_i^R \in \mathbb{R}^{k_i \times k_i}$ an arbitrary real matrix (of the appropriate size). Recent work has enabled exact computation with respect to a single real full block [60], and the computation involves a singular value minimization problem reminiscent of the μ upper bound (though it involves the *second* largest singular value). It may be possible to combine this result with the work presented here, to allow for general mixed μ problems which include real full blocks.

Nonlinear and/or Time Varying Uncertainties: Recent results obtained independently by Shamma [69] and Megretsky [46] have shown that problems involving structured nonlinear and/or time varying uncertainties (with an arbitrary number of blocks) can be solved exactly as a complex μ upper bound problem. Once again it would be useful to allow for more general mixed μ problems, which included these types of blocks.

Related μ Problems: A number of extensions to the μ theory are currently under investigation. Smith [72] and Newlin [50] have looked at the model validation problem, and Smith and Doyle [73] have examined its connection to the system identification problem. In both cases the problem one needs to solve is intimately associated with the mixed μ problem. A number of other avenues of current research, including time domain and nonlinear μ theories, also require extensions to the mixed μ theory.

Bounds Computation: The computation schemes for both the upper and lower bounds are still being improved, and it is expected that a next generation of algorithms is not too far away. For the lower bound it is possible to consider more sophisticated adaptive power iterations [75]. These algorithms change as the iteration proceeds, increasing the reliability of the algorithm (at additional computational expense) if it is not converging. In this way one can guarantee the convergence of the lower bound power iteration, and the goal is to do this without reducing its speed. The upper bound can be cast as an LMI minimization problem, and a number of research directions are being pursued for more efficient solution of such problems (see [12] for a review).

Branch and Bound: The present Branch and Bounds scheme is a fairly simple one. It seems to work pretty well, but there are a number of areas where one may consider possibilities for improvement. These include a better heuristic for the chopping criterion, and the use of (optimized) off-center cuts. Both of these extensions have the potential to greatly enhance the performance (see [51]).

Synthesis: The current synthesis techniques are very preliminary, and it is expected that these can be improved upon considerably in the fairly near future, by more fully exploiting the properties of the $\mathcal{H}_\infty$ optimal controller. In particular the present scheme is not really practical when there are large repeated scalar blocks, because of the difficulty of obtaining good fits to the scaling matrices, and this area deserves further study.

Many of these topics are currently being investigated (by a number of researchers), and we believe that this research area will continue to be active for some time. Indeed in some cases the results presented in this thesis are already being used as building blocks towards more powerful robustness theories and methodologies.

Bibliography

[1] V. Balakrishnan and S. P. Boyd, "Existence and uniqueness of optimal matrix scalings," in *Proceedings of the 31st Conference on Decision and Control*, pp. 2010–2011, 1992.

[2] V. Balakrishnan, S. P. Boyd, and S. Balemi, "Branch and Bound algorithm for computing the minimum stability degree of parameter-dependent linear systems," technical report, Information Systems laboratory, Department of Electrical Engineering, Stanford University, 1991.

[3] G. J. Balas, J. C. Doyle, K. Glover, A. K. Packard, and R. S. Smith, "The μ analysis and synthesis toolbox," MathWorks and MUSYN, 1991.

[4] G. J. Balas, C. C. Chu, and J. C. Doyle, "Vibration damping and robust control of the JPL/AFAL experiment using μ-synthesis," in *Proceedings of the 28th Conference on Decision and Control*, pp. 2689–2694, 1989.

[5] G. J. Balas and J. C. Doyle, "Robustness and performance tradeoffs in control design for flexible structures," in *Proceedings of the 29th Conference on Decision and Control*, pp. 2999–3010, 1990.

[6] G. J. Balas and A. K. Packard, "Development and application of time-varying μ-synthesis techniques for control design of missile autopilots," John Hopkins Applied Physics Laboratories final report, January, 1992.

[7] G. J. Balas, A. K. Packard, J. C. Doyle, K. Glover, and R. S. Smith, "Development of advanced control design software for researchers and engineers," in *Proceedings of the American Control Conference*, pp. 996–1001, 1991.

[8] G. J. Balas, P. M. Young, and J. C. Doyle, "μ based control design as applied to a large space structure: Control design for the minimast facility," NASA CSI/GI final report, June 1992.

[9] B. Barmish, P. Khargonekar, Z. Shi, and R. Tempo, "Robustness margin need not be a continuous function of the problem data," *Systems & Control Letters*, vol. 15, pp. 91–98, 1989.

[10] B. R. Barmish and H. I. Kang, "A survey of extreme point results for robustness of control systems," *Automatica*, vol. 29, pp. 13–35, 1993.

[11] A. C. Bartlett, C. V. Hollot, and H. Lin, "Root locations of an entire polytope of polynomials: It suffices to check the edges," in *Mathematics of Control, Signals and Systems*, Springer Verlag, 1988.

[12] C. Beck, "Computational issues in solving LMIs," in *Proceedings of the 30^{th} Conference on Decision and Control*, pp. 1259–1260, 1991.

[13] C. Beck and J. C. Doyle, "Mixed μ upper bound computation," in *Proceedings of the 31^{st} Conference on Decision and Control*, pp. 3187–3192, 1992.

[14] S. P. Boyd and C. H. Barratt, *Linear Controller Design: Limits of Performance*, Prentice Hall, New Jersey, 1991.

[15] S. P. Boyd and L. El Ghaoui, "Method of centers for minimizing generalized eigenvalues," to appear in special issue of *Linear Algebra and its Applications*, on Numerical Linear Algebra Methods in Control, Signals and Systems, 1992.

[16] R. D. Braatz, P. M. Young, J. C. Doyle, and M. Morari, "Computational complexity of μ calculation," to appear in *IEEE Transactions on Automatic Control.*

[17] J. Chen, M. K. H. Fan, and C. N. Nett, "The structured singular value and stability of uncertain polynomials: A missing link," *Control of Systems with Inexact Dynamic Models, ASME*, pp. 15–23, 1991.

[18] R. W. Daniel, B. Kouvaritakis, and H. A. Latchman, "Principal direction alignment: A geometric framework for the complete solution to the μ problem," *IEE Proceedings, Part D*, vol. 133, pp. 45–56, March 1986.

[19] R. R. E. de Gaston and M. G. Safonov, "Exact calculation of the multiloop stability margin," *IEEE Transactions on Automatic Control*, vol. 33, pp. 156–171, 1988.

[20] J. Demmel, "The componentwise distance to the nearest singular matrix," *SIAM Journal on Matrix Analysis and Applications*, vol. 13, pp. 10–19, 1992.

[21] J. C. Doyle, K. Glover, P. Khargonekar, and B. A. Francis, "State space solutions to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems," *IEEE Transactions on Automatic Control*, vol. 34, pp. 831–847, August 1989.

[22] J. C. Doyle, A. K. Packard, and K. Zhou, "Review of LFTs, LMIs and μ," in *Proceedings of the 30^{th} Conference on Decision and Control*, pp. 1227–1232, 1991.

[23] J. C. Doyle, "Analysis of feedback systems with structured uncertainty," *IEE Proceedings, Part D*, vol. 129, pp. 242–250, November 1982.

[24] J. C. Doyle, "Structured uncertainty in control system design," in *Proceedings of the 24^{th} Conference on Decision and Control*, pp. 260–265, 1985.

[25] J. C. Doyle, private communication, 1992.

[26] J. C. Doyle, B. A. Francis, and A. R. Tannenbaum, *Feedback Control Theory*, Macmillan, New York, 1992.

[27] M. K. H. Fan, "A quadratically convergent algorithm on minimizing the largest eigenvalue of a symmetric matrix," submitted to *Linear Algebra and its Applications.*

[28] M. K. H. Fan and A. L. Tits, "Characterization and efficient computation of the structured singular value," *IEEE Transactions on Automatic Control*, vol. 31, pp. 734–743, 1986.

[29] M. K. H. Fan, A. L. Tits, and J. C. Doyle, "Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics," *IEEE Transactions on Automatic Control*, vol. 36, pp. 25–38, 1991.

[30] B. A. Francis, *A Course in $\mathcal{H}_\infty$ Control Theory*, Springer Verlag, New York, 1987.

[31] M. R. Garey and D. S. Johnson, *Computers and Intractability: A Guide to the Theory of NP Completeness*, W. H. Freeman, New York, 1979.

[32] G. H. Golub and C. F. Van Loan, *Matrix Computations*, The Johns Hopkins University Press, Baltimore, 1987.

[33] R. A. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.

[34] J. P. How and S. R. Hall, "Connections between the Popov criterion and bounds for real parametric uncertainty," submitted to *IEEE Transactions on Automatic Control.*

[35] F. Jarre, "An interior point method for minimizing the maximum eigenvalue of a linear combination of matrices," technical report, SOL 91-8, Department of Operations Research, Stanford University, 1991.

[36] T. Kato, *A Short Introduction to Perturbation Theory for Linear Operators*, Springer Verlag, New York, 1982.

[37] V. L. Kharitonov, "Asymptotic stability of an equilibrium position of a family of systems of linear differential equations," *Differential Equations*, vol. 14, pp. 1483–1485, 1979.

[38] B. Kouvaritakis and H. A. Latchman, "Necessary and sufficient stability criterion for systems with structured uncertainties: The major principal direction alignment principle," *International Journal of Control*, vol. 42, pp. 575–598, 1985.

[39] H. A. Latchman, *Frequency Response Methods for Uncertain Multivariable Systems*, Ph.D. dissertation, Oxford University, 1986.

[40] L. Lee and A. L. Tits, "On continuity/discontinuity in robustness indicators," in *Proceedings of the 30th Conference on Decision and Control*, pp. 547–548, 1991.

[41] L. Lee and A. L. Tits, "Mixed one and two sided real uncertainty: A vertex result," in *Proceedings of the American Control Conference*, pp. 2197–2199, 1992.

[42] K. Lim and G. J. Balas, "Line-of-sight control of the CSI evolutionary model: $\mathcal{H}_\infty$ and μ control," in *Proceedings of the American Control Conference*, pp. 1996–2000, 1992.

[43] J. N. Little and L. Shure, *The Signal Processing Toolbox*, MathWorks, 1988.

[44] W. M. Lu, K. Zhou, and J. C. Doyle, "Stabilization of LFT systems," in *Proceedings of the 30th Conference on Decision and Control*, pp. 1239–1244, 1991.

[45] D. G. Luenberger, *Linear and Nonlinear Programming*, Addison Wesley, Reading, Massachusetts, 1989.

[46] A. Megretsky, "Power distribution approach in robust control," technical report, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden, 1992.

[47] M. Morari and E. Zafiriou, *Robust Process Control*, Prentice Hall, New Jersey, 1989.

[48] J. E. Nesterov and A. S. Nemirovsky, "Self-concordant functions and polynomial-time methods in convex programming," technical report, Central Economical and Mathematical Institute, USSR Academy of Science, Moscow, USSR, 1989.

[49] M. P. Newlin, private communication, 1992.

[50] M. P. Newlin and R. S. Smith, "Model validation and a generalization of μ," in *Proceedings of the 30th Conference on Decision and Control*, pp. 1257–1258, 1991.

[51] M. P. Newlin and P. M. Young, "Mixed μ problems and Branch and Bound techniques," in *Proceedings of the 31st Conference on Decision and Control*, pp. 3175–3180, 1992.

[52] A. V. Oppenheim and R. W. Schaffer, *Digital Signal Processing*, Prentice Hall, New Jersey, 1975.

[53] E. E. Osborne, "On preconditioning of matrices," *Journal of the Association for Computing Machinery*, vol. 7, pp. 338–345, 1960.

[54] M. L. Overton, "On minimizing the maximum eigenvalue of a symmetric matrix," *SIAM Journal of Matrix Analysis and its Applications*, vol. 9, pp. 256–268, 1988.

[55] A. K. Packard and J. C. Doyle, "Structured singular value with repeated scalar blocks," in *Proceedings of the American Control Conference*, pp. 1213–1218, 1988.

[56] A. K. Packard, *What's new with μ: Structured Uncertainty in Multivariable Control*, Ph.D. dissertation, University of California, Berkeley, 1988.

[57] A. K. Packard and J. C. Doyle, "The complex structured singular value," *Automatica*, vol. 29, pp. 71–109, 1993.

[58] A. K. Packard and P. Pandey, "Continuity properties of the real/complex structured singular value," *IEEE Transactions on Automatic Control*, vol. 38, pp. 415–428, March 1993.

[59] A. K. Packard, M. K. H. Fan, and J. C. Doyle, "A power method for the structured singular value," in *Proceedings of the 27^{th} Conference on Decision and Control*, pp. 2132–2137, 1988.

[60] L. Qiu, B. Bernhardsson, A. Rantzer, E. J. Davison, P. M. Young, and J. C. Doyle, "On the real stability radius," to appear in the *International Federation of Automatic Control, 12^{th} World Congress*, 1993.

[61] L. Qiu and E. J. Davison, "A simple procedure for the exact stability robustness computation of polynomials with affine coefficient perturbations," *Systems & Control Letters*, vol. 13, pp. 413–420, 1989.

[62] A. Rantzer and A. Megretsky, "A convex parameterization of robustly stabilizing controllers," technical report, Department of Mathematics, Royal Institute of Technology, Stockholm, Sweden, 1992.

[63] R. Redheffer, "Inequalities for a matrix Riccati equation," *Journal of Mathematics and Mechanics*, vol. 8, 1959.

[64] J. Rohn and S. Poljak, "Checking robust nonsingularity is NP-hard," to appear in *Mathematics of Control, Signals and Systems.*

[65] W. Rudin, *Real and Complex Analysis*, Mc Graw Hill, New York, 1987.

[66] M. G. Safonov, "Tight bounds on the response of multivariable systems with component uncertainty," in *Proceedings of the Annual Allerton Conference on Communication, Control and Computing*, pp. 451–460, 1978.

[67] M. G. Safonov, *Stability and Robustness of Multivariable Feedback Systems*, MIT Press, Boston, Massachusetts, 1980.

[68] M. G. Safonov, "Stability margins for diagonally perturbed multivariable feedback systems," *IEE Proceedings, Part D*, vol. 129, pp. 251–256, 1982.

[69] J. S. Shamma, "Robustness analysis for time-varying systems," in *Proceedings of the 31^{st} Conference on Decision and Control*, pp. 3163–3168, 1992.

[70] A. Sideris and R. S. Sánchez Peña, "Fast computation of the multivariable stability margin for real interrelated uncertain parameters," *IEEE Transactions on Automatic Control*, vol. 34, pp. 1272–1276, December 1989.

[71] D. D. Siljak, "Parameter space methods for robust control design: A guided tour," *IEEE Transactions on Automatic Control*, vol. 34, pp. 674–688, 1989.

[72] R. S. Smith, *Model Validation for Uncertain Systems*, Ph.D. dissertation, California Institute of Technology, 1990.

[73] R. S. Smith and J. C. Doyle, "Towards a methodology for robust parameter identification," in *Proceedings of the American Control Conference*, pp. 2394–2399, 1990.

[74] G. Stein and J. C. Doyle, "Beyond singular values and loop shapes," *Journal of Guidance, Control and Dynamics*, vol. 14, pp. 5–16, January 1991.

[75] J. E. Tierno and P. M. Young, "An improved μ lower bound via adaptive power iteration," in *Proceedings of the 31^{st} Conference on Decision and Control*, pp. 3181–3186, 1992.

[76] J. H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, Oxford, 1965.

[77] J. C. Willems, *The Analysis of Feedback Systems*, MIT Press, Cambridge, Massachusetts, 1971.

[78] P. M. Young, "The rank one mixed μ problem and "Kharitonov-type" analysis," submitted to *Automatica*.

[79] P. M. Young and K. J. Åström, "μ meets Bode," manuscript in preparation.

[80] P. M. Young and J. C. Doyle, "Properties of the mixed μ problem and its bounds," submitted to *IEEE Transactions on Automatic Control*.

[81] P. M. Young and J. C. Doyle, "Computation of μ with real and complex uncertainties," in *Proceedings of the 29^{th} Conference on Decision and Control*, pp. 1230–1235, 1990.

[82] P. M. Young, M. P. Newlin, and J. C. Doyle, "μ analysis with real parametric uncertainty," in *Proceedings of the 30^{th} Conference on Decision and Control*, pp. 1251–1256, 1991.

[83] P. M. Young, M. P. Newlin, and J. C. Doyle, "Practical computation of the mixed μ problem," in *Proceedings of the American Control Conference*, pp. 2190–2194, 1992.

[84] G. Zames, "On the input-output stability of nonlinear time-varying feedback systems, parts I and II," *IEEE Transactions on Automatic Control*, vol. 11, pp. 228–238 and pp. 465–476, 1966.

[85] K. Zhou, J. C. Doyle, and K. Glover, *Robust and Optimal Control*, book in preparation.