Performance optimization of highly uncertain systems in H infinity

James G. Owen

BA (Oxford University) 1986.

Department of Electrical Engineering

McGill University, Montreal.

A thesis submitted to the Faculty of Graduate Studies and Research

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

January 1993

ISBN 0-315-87945-9

To my children,

Rebecca and David.

Abstract

The subject of this thesis is the attenuation of disturbances by feedback in the presence of large plant uncertainty in MIMO systems. Some relevant optimizations are formulated in an H^∞ setting, and are shown to reduce to a 'two-disc' mixed sensitivity problem. Despite its basic nature, the two-disc problem has not previously been solved, although approximations to it have been proposed. It is shown, by means of examples, that such approximations may involve large errors, highlighting the need for an accurate theory.

An accurate theory of the two-disc problem is achieved by expressing the two-disc problem as a distance minimization in a certain Banach space, and then applying Banach space duality methods to characterize the solutions. Dual and predual spaces are identified and equivalent maximizations formulated therein. Alignment conditions are obtained, which yield the conclusion that the optimal solution is 'allpass' in general and unique in the SISO case. The duality description of the problem leads naturally to a solution based on convex programming.

In an extension of the theory to time-varying discrete-time systems, optimal robust disturbance rejection is shown to reduce to an operator form of the two-disc problem in the l^2-induced norm topology. A predual description is obtained, and an existence result for optimal feedback is proved. In the case of time-invariant nominal plants subject to time-varying uncertainty, at least one optimal controller is shown to be time-invariant.

Finally, complexity based notions of uncertainty are considered. The ability of feedback to reduce metric complexity is examined when applied either prior to, or after identification.

Résumé

Dans cette thèse, on étudie l'atténuation des perturbations par feedback en présence de grande incertitude dans des systèmes à entrées et sorties multiples. Dans le contexte de la théorie H^∞, des approches appropriées d'optimisation ont été proposées. Ces approches ont démontré que le problème étudié se réduit alors au problème à deux disques à sensiblilité mixte. Malgré sa nature simple, ce problème à deux disques n'avait pas été résolu auparavant. Seules des approximations avaient été proposées. Par des exemples, on a pu montrer que ces approximations engendrent de grandes erreurs et prouver la nécèssité d'une théorie plus exacte.

Une théorie exacte du problème à deux disques est réalisée en exprimant ce problème sous forme de minimisation d'une distance dans un espace Banach et en lui appliquant les méthodes de dualité dans l'espace Banach pour caractériser les solutions. Les espaces duel et préduel sont identifiés, et des maximisations équivalentes sont dérivées. Les conditions d'alignement sont obtenues; ces conditions nous permettent de conclure que la solution est un 'passe-tout' en général et que cette solution est unique dans le cas d'un système à entrée et sortie uniques. La dualité du problème nous a permis de considérer une programmation convexe.

Dans une extension de la théorie des systèmes variables et discrets, il est démontré que le rejet optimal et robuste de perturbations se réduit à une forme operateur du problème à deux disques dans la norme L^2. Une description préduelle est obtenue, et l'existence d'un feedback optimal est prouvée. Dans le cas des systèmes constants, sujets à des perturbations variant avec le temps, on a démontré qu'il existe au moins un contrôleur optimal qui n'est pas fonction du temps.

Finalement, les notions d'incertitude basées sur la complexité sont considérées. La capacité du feedback à réduire la complexité métrique lorsqu'il est appliqué avant

ou après identification est examinée .

Acknowledgements

I would like to express my deep gratitude for the contribution of my supervisor, Prof. George Zames, to the work that led to this thesis and to my own development as a researcher. His scholarship, rigor, and depth of thought have provided me with an appreciation of the true process of scientific enquiry.

I would also like to extend my appreciation to Prof. Stephen Boyd of Stanford University for our extensive discussions during the summer of 1992, which subsequently contributed to Chapters 4 and 7.

Among past and present graduate students of the control group, with whom I have had many enlightening and lively discussions, I would particularly like to thank Lin Lin, Finn Wredenhagen, Wang Le Yi, Mark Readman, Wei Yuanjun, Yair Matas, Patrick Lie, and Hans Wehn.

Special thanks also go to Kaouthar Benameur for her expert translation of the abstract into French.

I am grateful for the financial support received from the Alma Mater and Max Bell major fellowships during the course of my studies.

Last, but by no means least, I would like to thank my wife for her love, support, encouragement and proof reading skills.

Claim of Originality

- For MIMO systems the problem of optimaly rejecting disturbances in the presence of large plant uncertainty (ORDAP) is shown to reduce to a fixed point problem expressed in terms of a MIMO extension of the two- disc problem.
- It is shown that in general the optimal robust robust disturbance attenuation may be infinitely sensitive to inaccuracies in evaluating the related two-disc problems.
- Solutions to SISO and MIMO versions of the two-disc problem are characterized using Banach space duality. Dual and predual descriptions of the problem are derived.
- Duality theory is used to prove a 'flatness' and approximate 'flatness' result for the two-disc optimization.
- Various other properties of the ORDAP and two-disc problems are derived including: A uniqueness result for the optimum, a strict monotonicity property, and well posedness w.r.t the uncertainty description.
- A numerical solution to the ORDAP and two-disc problems is derived, based on a combination of convex programming methods with duality theory.
- Explicit results are presented for the case where the weightings describing the sets of disturbances and plants are 'almost' complementary.

- The ORDAP for time-varying plants and feedback is reduced to a time-varying version of the two-disc problem. For time-invariant plants subject to time-varying perturbations, at least one of the control laws which optimaly reject l^2 disturbances is shown to be time-invariant.

- Quantifications of the effect of feedback on a measure of metric complexity (Kolmogerov ϵ-dimension) are deduced.

Contents

List of Figures

Notation

$\mathbb{R}, \mathbb{C}$ denote the fields of real and complex numbers respectively. If $x \in \mathbb{C}$ then $\overline{x}$ denotes the complex conjugate of x.

$L^p_{n\times n}$ denoted the L^p space of $n \times n$ matrix valued functions on the unit disc.

$H^p_{n\times n}$, $1 \leq p \leq \infty$ denotes the usual Hardy space of $n \times n$ matrix functions defined on the unit disc, viewed as a subspace of $L^p_{n\times n}$ of the unit disc.

$(H^1_0)_{n\times n}$ denotes the subspace of $H^1_{n\times n}$ given by $\{F \in H^1_{n\times n} : \int_0^{2\pi} F(e^{i\theta})d\theta = 0\}$

$\overline{H}^\infty_{n\times n}$ and $(\overline{H}^1_0)_{n\times n}$ are the subspaces obtained by taking complex conjugates of all functions in $H^\infty_{n\times n}$ and $(H^1_0)_{n\times n}$ respectively.

$\mathcal{R}$ denotes the field of rational functions.

$\mathcal{R}H^\infty_{n\times n}$ and $\mathcal{R}L^\infty_{n\times n}$ denote the restriction of the spaces $H^\infty_{n\times n}$ and $L^\infty_{n\times n}$ to rational functions.

$STrA$ denotes $Trace(A^*A)^{\frac{1}{2}}$.

$|\cdot|$ denotes one of the following, depending on the context. For an $n \times n$ matrix A, $|A|$ is the largest singular value. For an $n-$vector ζ, $|\zeta|$ is the Euclidean norm. For a matrix or vector valued function F on the unit circle, $|F|$ is the real-valued function on the circle satisfying $|F|(e^{i\theta}) = |F(e^{i\theta})|$.

C denotes the space of continuous functions on the unit circle.

l_2, l_2^+ denote the Hilbert spaces of infinite and semi-infinite square summable sequences.

$l^+_{2,e}$ denotes extended l_2^+.

l_p^N denotes the Banach space of N-tuples under the p norm.

$l.a.e.$ is the label used for 'Lebesgue almost everywhere'.

D and T are the unit disc and circle of the complex plane respectively.

$\sigma_{min}(\cdot)$ and $\sigma_{max}(\cdot)$ are used to denote the minimum and maximum singular values of a matrix.

$int(\cdot)$ denotes the smallest integer greater than or equal to an argument.

* denotes either the complex conjugate transpose of a vector in $\mathbb{C}^n$ or the adjoint of an operator, depending on the context.

Chapter 1

Introduction

1.1 Overview

This thesis is concerned with the effects of feedback on plant uncertainty. The first part of it is devoted to two related basic feedback problems which in a sense provide the raison d'etre for the H^∞ theory, but which have remained unsolved despite the explosive growth of the subject since the mid 1980's.

The first of these problems (Fig 1) involves the optimization of a feedback controller for a plant P subject to disturbances and sensor noise, described by the weighting functions of frequency $W(i\omega)$ and $V(i\omega)$ under a 'worst case' criterion. This problem, which is to be described in greater detail shortly, will be referred to as the <u>two-disc</u> problem. [1]. It is well known that the 'two block' problem of H^∞ theory provides an approximate solution to the two-disc problem. That approximation is off by at most $\sqrt{2}$, but an accurate solution has so far not been given.

The second, closely related but somewhat harder problem, is that of optimal disturbance rejection when there is plant uncertainty ΔP or a set of plant perturbations (Fig. 2). We will refer to this as the problem of optimal robust disturbance attenuation (ORDAP). It is well known for the SISO case [20] that the ORDAP can be reduced to a parametric version of the two-disc problem. It has therefore been widely

[1] So called because a certain level of performance is equivalent to the non-overlapping of two discs [12]

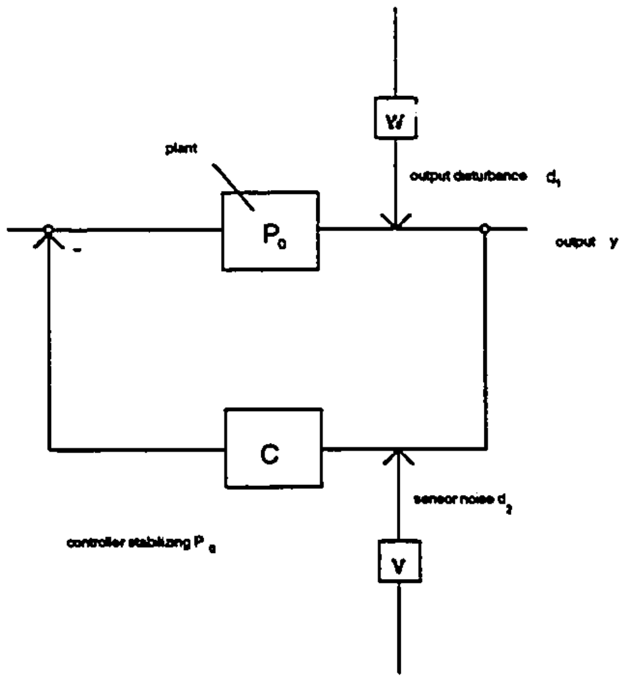

Figure 1: Control loop with known plant dynamics and two sources of disturbance

assumed that a corresponding parametric version of the two-block H^∞ problem would provide an approximate solution to the ORDAP as well.

We will show that in fact the approximations obtained in this way can be infinitely poor, even though the H^∞ approximations to the two-disc problem in the nonparametric case are never off by more than a $\sqrt{2}$ factor. It would seem therefore that the current H^∞ theory may be inadequate for dealing with even the simplest feedback problems when there is significant plant uncertainty, and that there is a corresponding need for an accurate theory of the two-disc and ORDAP problems. We will proceed to give such a theory, which will lead to a precise characterization of the solutions as well as a numerical computation.

The second part of the thesis (Chapter 7) will be devoted to the related but somewhat different question of how feedback affects metric complexity.

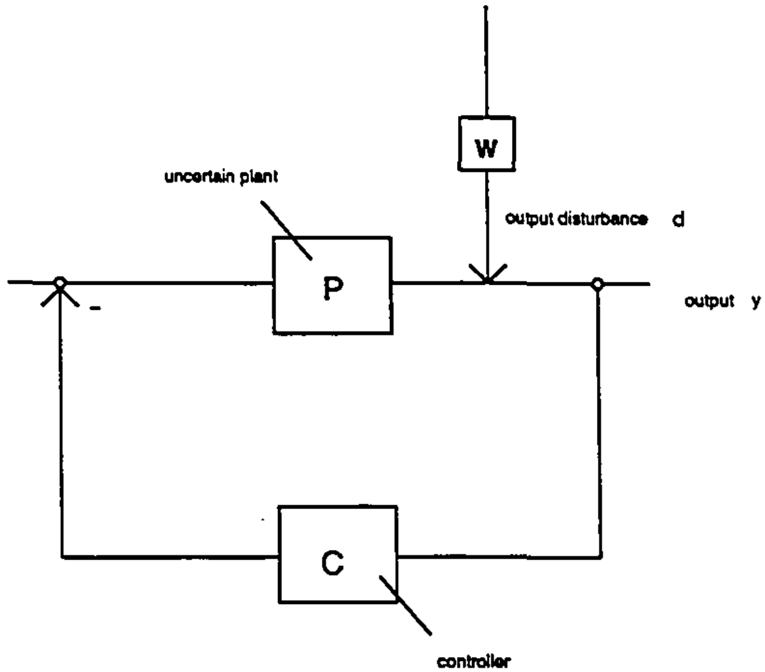

Figure 2: Control loop with uncertain plant and unknown output disturbance

1.2 The Two-Disc and ORDAP problems

The two-disc problem involves a feedback loop (Figure 1) in which there are disturbances added to both the plant output and the controller input (e.g. sensor noise), and the plant P_0 is assumed to be known exactly. The objective is to find a feedback law which optimally suppresses the effects of these disturbances on the system output. For this case, we assume that d_1, d_2 are filtered versions of signals u_1 and u_2, i.e.

$$d_1 = W u_1, \qquad d_2 = V u_2 \tag{1}$$

where u_1 and u_2 have a common upper bound u, i.e., $|u_1(i\omega)| \leq |u(i\omega)|$, $|u_2(i\omega)| \leq |u(i\omega)|$, and u has unit energy (i.e. unit L_2 norm). We seek to optimize the feedback control law C, so as to stabilize the plant P_0, and minimize the maximum energy produced at the output over all admissible disturbances d_1, d_2 satisfying (1). This reduces solving an optimization taking the form

$$\inf_{Q \in H^\infty} \| |W(1 - P_0 Q)| + |V P_0 Q| \|_\infty \tag{2}$$

which will be taken to be the defining equation for the two-disc problem in the SISO case.

In the ORDAP problem (see Fig. 2) an uncertain plant is subjected to disturbances at the output. The objective is to find a feedback control law which provides the best possible uniform attenuation of output disturbances in spite of uncertainty in the plant model. In particular, suppose that the disturbance can be any signal in the set of outputs produced by a filter with a stable transfer function W, driven by an arbitrary input of unit energy. Suppose also that the plant frequency response $P(i\omega)$ is uncertain and lies in a weighted sphere in the frequency-domain, described by the inequality

$$|P(i\omega) - P_0(i\omega)| \leq |V(i\omega) P_0(i\omega)| \qquad \forall \omega \in [0, \infty), \tag{3}$$

where P_0 is some nominal frequency response representing the 'center' of the frequency band. We seek to optimize the feedback control law C, so as to stabilize all systems whose frequency responses are described by (3), and minimize the worst case output energy produced in response to any admissible disturbance d acting at the output of any admissible [2] plant P . The infimal worst case transmitted disturbance energy μ for the set of plants described by (3) can be expressed,

$$\mu = \inf_{\substack{Q \in H^\infty \\ \|V P_0 Q\|_\infty \leq 1}} \sup_{X \in H^\infty_{n \times n}, \|X\|_\infty < 1} \left\| \frac{W(1 - P_0 Q)}{1 - X V P_0 Q} \right\|_\infty \tag{4}$$

[2]This problem is stated more formally and in more generality in Chapter 2.

What distinguishes this problem from more typical H^∞ optimizations is that it is non-convex in the parameter Q, at least when the uncertainty in the plant is not assumed to be small. (When the uncertainty is assumed to be limitingly small, (4) reduces to the well-known optimal weighted sensitivity problem of H^∞ control.) (4) can be substantially simplified, at the cost of expressing the problem in implicit form, by showing that it is equivalent to finding the smallest positive fixed-point of a function. The values of this function are defined by a family of convex, unconstrained and explicit two-disc H^∞ optimizations. For the SISO case this function is described by;

$$\chi : [0,\infty) \to [0,\infty), \qquad \chi(r) := \inf_{Q\in H^\infty} \| |W(1 - P_0Q)| + r|VP_0Q| \|_\infty \tag{5}$$

(The MIMO case is slightly different see (27)). Thus, for each value of the parameter r, the ORDAP yields the same non-standard 'sum of absolute values' minimization as the two disturbance problem for a known plant (2).

The absence of an exact theory for problems (2) and (4) has limited their analysis in the literature to approximation of (2) and (5) by a standard 'sum of squares' mixed sensitivity problem i.e., for any value of the parameter r, the optimization

$$\inf_{Q\in H^\infty} \| |W(1 - P_0Q)|^2 + r^2|VP_0Q|^2 \|_\infty, \tag{6}$$

for which exact solutions are available. However if plant uncertainty is not small, implicit dependence of the optimal disturbance attenuation on r can make accurate approximation of $\chi(\cdot)$ critical even for crude estimation of the overall performance. To justify this statement consider the case $P_0(s) = \frac{1.9-0.1s}{(1+s)(1.9+0.1s)}$, $W(s) = 0.17\left(\frac{1.8+0.2s}{1+s}\right)^3$, $V(s) = 0.22\left(\frac{0.1+2.1s}{1+s}\right)^2$ [3]. The function $\chi(\cdot)$ is shown in Fig. 3 and its fixed-point is the intersection with the line of unit slope. (The curve of Fig. 3 is

[3] This involves a lowpass nominal plant, and 'complementary' weightings W, V, whose graph is shown in Fig. 10 of page 80.

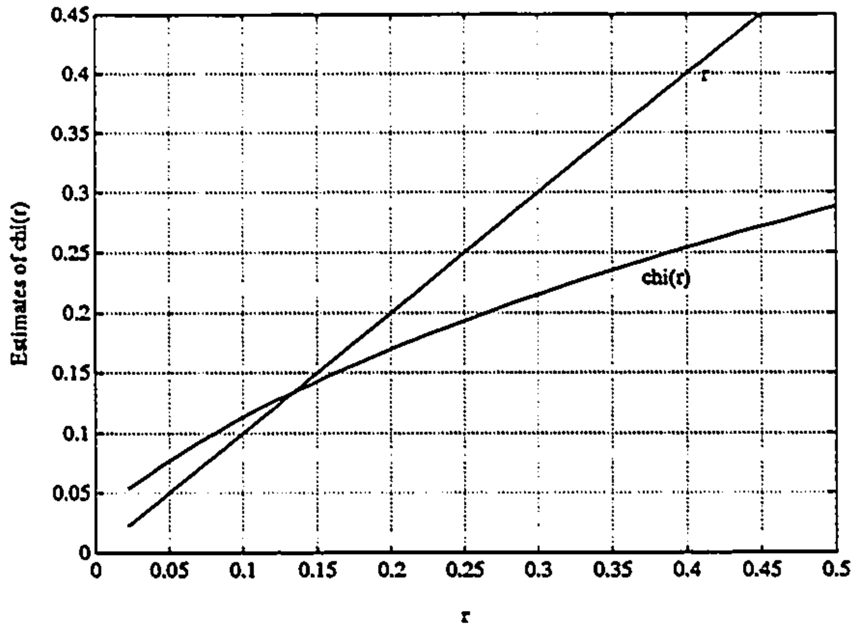

Figure 3: Plots of estimates of $\chi(r)$ vs r obtained from the algorithm described in Chapter 4

produced by the algorithm detailed in Chapter 4 and based on the theory of Chapter 3). It is apparent from Fig. 3, that for small r the location of the fixed-point would be highly sensitive to errors in the approximation of χ. Indeed, at the end of Chapter 2 it will be shown that in some cases the fixed-point may be *infinitely* sensitive to estimation errors in χ. These examples illustrate the fact that methods based on approximating χ by standard mixed sensitivity problems, with their attendant factor of $\sqrt{2}$ relative errors, will not in general yield even approximate estimates for (4) (c.f. Sect. 1.4). The conclusion, then, is that to solve disturbance attenuation problems in the presence of significant plant uncertainty, an exact theory for two-disc problems of the form (2) is needed. This motivates the first six chapters of the thesis.

1.3 Problem Statement and Outline of Results

1.3.1. Feedback and Unstructured Uncertainty

Broadly stated, the objectives of this work are to investigate the ability of feedback to reduce uncertainty, by analyzing the relevant optimizations. On a more philosophical level our objective is to understand the trade-offs that exist between open-loop uncertainty and the resulting closed-loop uncertainty sets. Since the case of limitingly small plant uncertainty is already well understood, our focus will be on situations where uncertainty may be large.

In Chapter 2 we introduce two such optimizations for MIMO systems. The first is a MIMO version of the ORDAP and the second is a similar problem which captures the potential of feedback to contract the radius of the set of plant uncertainty (c.f. Fig. 4). The ORDAP for the MIMO case is then shown to be equivalent to an implicit form of an extended two-disc optimization of (2). Chapter 2 ends with a motivating example illustrating the sensitivity of this fixed-point to inaccuracies of approximation.

A common feature in the treatments of the standard 'sum of squares' two-block problem (e.g. (6)) by [49], [18], [59] and [33] is the reduction to a minimax optimization for a maximum singular value. It is readily seen that both (5) and (2) do not fall into this category, ruling out the use of the standard H^∞ methods for a direct solution to (5). In Chapter 3 we take a completely different approach to these problems. The development there begins by recognizing that the nonstandard two-block problems obtained in Chapter 2 can be expressed geometrically, as the minimization of the distance between a vector $\begin{bmatrix} U^*W \\ 0 \end{bmatrix}$ (where U^* is the involution of the inner factor of P_0) and a subspace $S := \begin{bmatrix} W \\ V \end{bmatrix} H^\infty$, in an appropriately defined Banach space B. Duality theory is then applied to find this minimal distance. (The main steps involved in applying the duality theory are: The predual and dual spaces of B, and of the restriction of B to continuous functions are identified, as are the corresponding pre-orthogonal complement and orthogonal complement of the subspace S. Dual and predual optimizations are formulated in the respective Banach spaces. Alignment conditions are then derived relating the minimal solution of the distance minimization (equivalently the optimal Q of (2)), and the maximal solution of the dual problem.)

The results obtained in this way are applied to make various qualitative deductions about the effect of feedback on uncertainty. Some of the more important conclusions are:

- A feedback which optimally reduces uncertainty arising from either output disturbances or plant perturbations, (i.e. optimal for the ORDAP), exists and is unique in the SISO case under quite general conditions.

- The smallest achievable closed-loop plant uncertainty and plant output disturbance transmission are strict monotone increasing functions of open-loop plant uncertainty; i.e., increases in the size of the open-loop uncertainty set over *any* frequency range strictly degrade the ability of feedback to further attenuate disturbances at all other frequencies.

- The weighted sensitivity under the above optimal feedback is 'allpass'. If the feedback is 'almost' optimal then an approximate allpass condition holds.

- Feedback which produces maximal contraction of a given set of plant uncertainty in a certain sense occupies all the 'space' in the H^∞ sphere of *optimal* radius.

- Unlike the one-block weighted sensitivity minimization [46], the ORDAP is well-posed with respect to the problem data. Thus small uncertainties in the data cannot produce arbitrarily large computation errors.

The duality theory of Chapter 3 leads naturally to a numerical method of solution for the ORDAP. In Chapter 4 this theory is used to develop algorithms to solve the ORDAP by approximately reducing each of the the two-disc problems of (5) to a pair of finite variable convex optimizations. The Ellipsoid algorithm of Shor, Yudin and Nemirovsky [45] as presented in Boyd [5] is then applied to these problems to obtain polynomial-time, non-hueristic programs which find 'nearly' optimal control laws. These algorithms have been implemented numerically, and were applied to produce the curves plotted in Fig. 3 and Fig. 4 for the example of Sect. 1.2. Fig. 4 is a graph of the ability of feedback to reduce plant uncertainty versus open-loop plant uncertainty radius for the above example.

In Chapter 5 an asymptotic case of the ORDAP is examined where plant uncertainty and output disturbances occur at almost entirely different frequencies. This is motivated in part by the above observation that uncertainty over one frequency range effects the uncertainty attenuation over all other frequency ranges. This coupling between uncertainty on one frequency range and performance on another plays a role, for instance, in situations where high frequency uncertainty occurs at frequencies beyond the bandwidth of exogenous disturbances (e.g. neglected flexible mode dynamics of a robot arm). Under these conditions a more explicit approximate analysis of the ORDAP is possible. This is achieved by demonstrating that, for a limiting case,

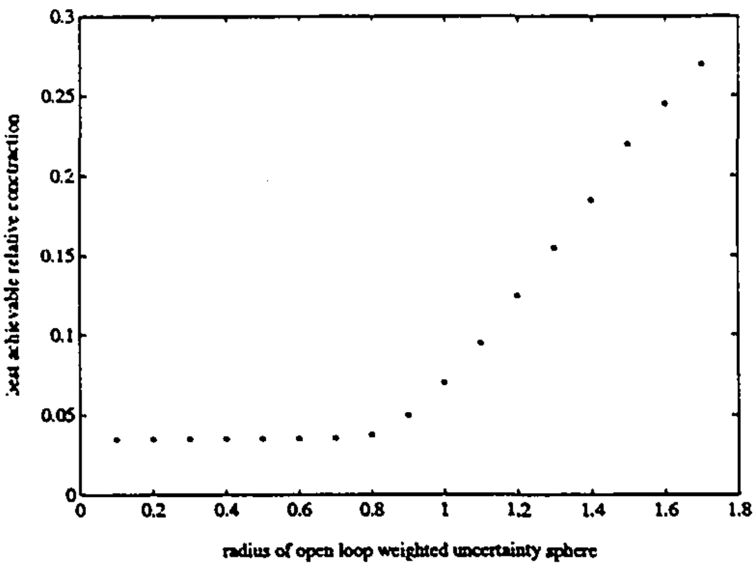

Figure 4: Optimal relative reduction of uncertainty by a single feedback control law for the example cited in Sect. 1.2

the ORDAP approximately reduces to a fixed-point problem for a function defined by a family of Hankel norms. The Hankel norms can then be found by established methods, for example as in [58], [19] and [61].

Chapters 6 and 7 are extensions of the material of Chapters 2-5. In Chapter 6 the ORDAP is formulated for time varying feedbacks and plants, and shown to reduce to a fixed-point problem based on a generalized operator version of the two-disc problem. An equivalent predual optimization is derived. Under certain conditions, time-varying discrete-time control laws which optimally reject a class of disturbances are shown to exist for sets of time-varying linear plants. For time-invariant *nominal* plants admitting possibly *time-varying* causal linear perturbations, the results of Chapter 6 reveal that at least one of the optimal feedback control laws is time-invariant. In such situations, time-varying feedback offers no advantage in reducing uncertainty over time-invariant feedback, so proving a conjecture of [29].

1.3.2 Feedback and Metric Complexity

The notion that feedback can reduce plant uncertainty leads naturally to the view that feedback and identification can be thought of as parts of the same process, namely that of reducing the amount of information that must be acquired to adequately control a system. If we ask the question, 'How much information about the input-output behavior of a system is needed to achieve some desired level of control tolerance?' then we have a basis for the quantification of the relative merits of feedback and identification. This question was posed in [56] where it was recognized that answers would depend on an information-based notion of plant uncertainty, such as Kolmogorov ϵ-dimension. Accordingly, in Chapter 7 some of the results of the preceding chapters are applied to the problem of gauging the effect of feedback on information based measures of uncertainty. Estimates are obtained for the ϵ-dimension of certain sets of plant uncertainty defined by a time-domain characteristic rather than the more usual frequency-domain specifications.

1.4 Classical Origins and Literature Review

The concept of feedback as an agent for reducing plant uncertainty goes back to the early days of classical control. In 1927, H.S. Black in his U.S patenjt application suggested that the use of high gain negative feedback could improve the accuracy of amplifier circuits in the presence of distortion. However many of the schemes that Black proposed resulted in the instability of the closed-loop system [11]. He had inadvertently encountered one of the fundamental trade-offs that lie at the heart of the ORDAP, namely that which exists between sensitivity reduction on the one hand, and the requirement of closed-loop stability on the other. The notion that there were limits to the improvement in closed-loop accuracy that could be obtained using feedback was brought into sharp focus in 1932 by Nyquist's graphical theory, which demonstrated that there were absolute constraints on the loop gain imposed by the need for stability. Horowitz [28] used the integral theorem of Bode to show that the sensitivity of a strictly proper plant could not be reduced at all frequencies; any reduction in one band would be offset by an increase elsewhere.

For the purposes of quantifying the effect of feedback on uncertainty, what was missing from the classical viewpoint was a well-founded input-output based definition of plant uncertainty. This gap was filled by the emergence of H^∞ theory [57] which captured the above trade-offs in the form of optimizations similar to those described in Sect. 1.2. In the key paper [57] Zames revitalized the subject of frequency-domain feedback design, previously based on classical rules of thumb, by proposing a coherent mathematical framework. This involved the use of the Banach algebra H^∞ to represent the space of causal time-invariant stable linear systems, and the minimization of weighted sensitivity as an objective function for feedback design. The H^∞ framework has several important properties, most notably that system interconnections can be represented by the simple algebraic operations of addition and multiplication. Of particular importance for this thesis, and for the representation of system uncertainty are two main points:

1. the H^∞ norm is the essential supremum of the frequency response.

2. there is an isometric isomorphism between H^∞ and the algebra of bounded, linear, time-invariant operators from the Hilbert space of finite energy inputs to finite energy outputs, where the norm is the induced norm. Practically, this means that two systems are close in the H^∞ metric if and only if they have similar input-output

behavior. This is not true, for example, of the state space description where very small changes in input-output behavior can radically alter both the dimensions and the entries of the state matrices.

It is the second point (2.) which makes H^∞ a natural framework in which to represent unstructured uncertainty for stable LTI systems. A sphere of uncertainty in H^∞ contains all the stable, linear, time-invariant systems whose input-output behavior is within the specified tolerance, and so is a good model of unstructured uncertainty. The first point (1.) allows this uncertainty to be represented in terms of a tolerance on the system frequency response i.e. the familiar band of uncertainty often sketched on Nyquist or Bode plots.

These observations allowed the classical idea of representing system uncertainty in terms of the frequency response, to be pinned down in an input-output framework. The representation of 'unstructured' plant uncertainty that resulted, enabled Zames to initially pose the ORDAP in [57]. For the case of limitingly small plant uncertainty he showed it to be equivalent to the problem of optimal weighted sensitivity [57], which has received a vast amount of attention in the last ten years. The initial solution appeared in [57] and [58] and has subsequently been considered in many other situations by other authors, for example [3] and [19].

Solutions to 'sum of squares' two-block problems of the form (6) were obtained in the work of Verma and Jonkheere [49], Kwakernaak [33], Foo and Postlethwaite [18], and Zames and Mitter [59] (for the infinite dimensional case).

In 1986 the ORDAP was examined in some detail by Francis and Bird in [20] and [4] for situations where the plant uncertainty was not assumed to be small. It was assumed that output disturbance signals were restricted to lie in the W-weighted sphere in H^2 described in 1.2, and the uncertain plant belonged to an additive V-weighted sphere in H^∞, whose center is the nominal system P_0 ($\in H^\infty$). Under these conditions it was shown in [20] that the optimal robust disturbance attenuation was bounded below and above by the positive fixed-points of the following two functions: [4]

$$\psi_1 : [0,\infty) \to [0,\infty) \qquad \psi_1(r) := \inf_{Q\in H^\infty} \left\| \begin{bmatrix} W(1-P_0Q) \\ rVQ \end{bmatrix} \right\|_\infty \tag{7}$$

$$\psi_2 : [0,\infty) \to [0,\infty) \qquad \psi_2(r) := \sqrt{2}\psi_1(r) \tag{8}$$

While these methods provide a means of synthesising controllers which guarantee a certain level of performance, the level of disturbance attenuation can be far from optimal as noted previously (c.f. Sect. 1.1).

It should be noted that O'Young and Francis [41] obtained more accurate estimates for the simpler but related problem of minimizing the sensitivity of an exactly known system on a frequency interval, subject to a robust stability constraint on the complementary interval.

Feedback in the presence of large plant uncertainty has also been considered by other authors in different topologies, for example the gap metric work of Georgiou and Smith [24].

There have been various attempts in the literature to generalize ideas about H^∞ uncertainty reduction to time-varying systems. Feintuch and Francis [15] studied both the optimal weighted sensitivity minimization problem and the two-block problem of [49], [33] and [18] in the time-varying case. They obtained abstract solutions to these problems and demonstrated that for time-invariant plants, time-varying control laws offered no advantage over time-invariant controllers. In [52] and [54] more concrete

[4] similar results for MIMO systems were also obtained

results were obtained for the case where the dependence of control laws on plant dynamics was taken to be causal (to allow the possibility of adaptation) for situations where the plants were 'slowly time-varying'. Khammash and Pearson [30] considered both the sensitivity minimization problem in the presence of plant uncertainty, and robust stability in the presence of diagonal structured perturbations, for time-varying systems in the l^∞ induced norm topology. However their methods could not be extended to the case of systems operating in the framework of l^2 signals.

Information-based notions of plant uncertainty were introduced into control in [56], partly motivated by the objective of a theory of identification and organization. Plant uncertainty was quantified by the metric complexity of sets of uncertainty, as measured by Kolmogorov ϵ-dimension and Kolmogorov ϵ-entropy. Such estimates of complexity provided a measure for the common objective of feedback and identification, i.e., the objective of reducing the acquisition of information needed to adequately control a system. More recent work dealing with the effects of feedback on measures of metric complexity has been reported in [37] and [51], while in [25] ideas of metric complexity were used in the context of worst-case identification. Use has been made of estimates of Kolmogorov ϵ-entropy and dimension in [50], [47], [48].

Chapter 2

The Optimal Robust Disturbance Attenuation Problem (ORDAP)

In this chapter we pose the ORDAP in the MIMO case where it takes the following form (see Fig 2). Let P be an uncertain plant, lying in a set of uncertainty described by the expression

$$\mathcal{B}(P_0, V) := \left\{ (I + VX)P_0 \ : \ X \in H^{\infty}_{n\times n}, \ \|X\|_\infty < 1 \right\} \tag{9}$$

$$\textit{where } P_0 \in H^{\infty}_{n\times n}, \ V^{\pm 1} \in H^{\infty}_{n\times n},$$

Output disturbances d lie in a set

$$\mathcal{D} \ := \ \{Wu \ : \ u \in H^2_n, \ \|u\|_2 \le 1\} \tag{10}$$

$$\textit{where } W \in H^{\infty}_{n\times n} \textit{ is outer}$$

We seek to optimize the feedback control law C so as to stabilize all plants in $\mathcal{B}(P_0, V)$ and suppress the W_1-weighted L_2 norm of the resulting output uniformly over plant and disturbance sets. This problem is expressed in the form,

$$\inf_{\substack{Q \in H^{\infty}_{n\times n} \\ \|P_0 Q V\|_\infty \le 1}} \ \sup_{X \in H^{\infty}_{n\times n}, \ \|X\|_\infty < 1} \|W_1 (I - P_0 Q)(I + VXP_0Q)^{-1} W\|_\infty \tag{11}$$

One of the purposes of this chapter will be to show that the quantity (11) is equal to the smallest fixed point of the following function of a parameter r,

$$\inf_{Q \in H^{\infty}_{n\times n}} \ ess \sup_{\theta \in [0,2\pi)} \ \sup_{\zeta \in \mathbb{C}^n, |\zeta| \le 1} \left(|W_1(I - P_0Q)W(e^{i\theta})\zeta| + r|P_0QV(e^{i\theta})\zeta| \right). \tag{12}$$

Note that (12) represents a MIMO extension of the two-disc problem. At the end of the chapter a special case of the ORDAP is analyzed, in which the optimal performance for the ORDAP is shown to be infinitely sensitive to inaccuracies in the computation of (12), illustrating the need for a precise theory for problems of this type.

Assumption. Throughout this chapter, either the control law or the plant is always assumed to be in the (radical of the algebra of causal operators, which consist of) strictly causal operators, [57] [16]. This assumption guarantees the well-posedness of all the feedback loops discussed, and ensures existence of all the operator expressions mentioned, at least in the algebra of causal time-invariant linear operators from $l^2_{+,e} \to l^2_{+,e}$ or $L^2_{+,e} \to L^2_{+,e}$. Here $l^2_{+,e}$ and $L^2_{+,e}$ denote the extended spaces of finite energy discrete and continuous-time signals respectively. Where an operator has an $H^\infty_{n\times n}$ representation no distinction is made between that representation and the associated operator, in order to simplify the presentation.

2.1 Attenuation of Output Disturbances in the Presence of Plant Uncertainty

The ORDAP is stated formally as follows:

Problem 1. We seek to minimize the largest (W_1) weighted H^2 norm of the output signal which results from any admissible disturbance $d \in \mathcal{D}$ to the output of any plant P in the uncertainty set $\mathcal{B}(P_0, V)$, over all robustly stabilizing feedbacks for $\mathcal{B}(P_0, V)$. That is, the optimal robust disturbance attenuation is defined to be

$$\mu_{opt} := \inf_{\substack{C \text{ stabilizing} \\ \text{all } P \in \mathcal{B}(P_0,V)}} \sup_{P \in \mathcal{B}(P_0,V)} \sup_{d \in \mathcal{D}} \|W_1 y\|_2 \tag{13}$$

where $W, W_1 \in H^\infty_{n\times n}$ and are outer

A first step towards simplifying optimizations of the type (13) is to express both the constraint and the quantity to be optimized, in terms of the same single unconstrained parameter. For (13) we chose the parameter to be $Q := C(I+P_0C)^{-1}$, since stabilizing control laws for the *nominal* plant P_0 are in one-to-one correspondence with $Q \in H^\infty_{n\times n}$ [57]. The constraint that a control law must stabilize a set of the type $\mathcal{B}(P_0, V)$ is shown to be equivalent to a norm bound on the parameter Q in the following lemma.

Lemma 2.1 *Let* $\mathcal{B} := \left\{P_0 + AXB \;:\; X \in H^\infty_{n\times n},\ \|X\|_\infty < 1\right\}$ *where* $P_0 \in H^\infty_{n\times n}$, $B \in H^\infty_{n\times n}$, $A^{\pm 1} \in H^\infty_{n\times n}$. *Then*

$$C \textit{ robustly stabilizes } \mathcal{B} \;\Leftrightarrow \|BQA\|_\infty \leq 1 \qquad (Q := C(I+P_0C)^{-1})$$

Proof. *C robustly stabilizes* $\mathcal{B}$

$$\Leftrightarrow C(I+PC)^{-1},\ (I+PC)^{-1} \in H^\infty_{n\times n} \quad \forall P \in \mathcal{B}$$

$$\Leftrightarrow Q(I+\Delta PQ)^{-1}, (I-P_0Q)(I+\Delta PQ)^{-1} \in H^\infty_{n\times n} \quad \forall P \in \mathcal{B}$$

$$\Leftrightarrow (I+AXBQ)^{-1} \in H^\infty_{n\times n} \qquad \forall X \in H^\infty_{n\times n},\ \|X\|_\infty < 1 \;\; \textit{and} \;\; Q \in H^\infty_{n\times n}$$

$$\Leftrightarrow A(I+XBQA)^{-1}A^{-1} \in H^\infty_{n\times n} \quad \forall X \in H^\infty_{n\times n},\ \|X\|_\infty < 1 \;\; \textit{and} \;\; Q \in H^\infty_{n\times n}$$

$$\Leftrightarrow (I+XBQA)^{-1} \in H^\infty_{n\times n} \quad \forall X \in H^\infty_{n\times n},\ \|X\|_\infty < 1 \;\; \textit{and} \;\; Q \in H^\infty_{n\times n} \qquad (14)$$

Clearly $\|BQA\|_\infty \leq 1$ is a sufficient condition for (14) to hold. We will prove the necessity of this condition by showing that $\|BQA\|_\infty > 1$ leads to a contradiction. If so, there exists a $z_0 \in D$ such that $|BQA(z_0)| \geq 1+\delta$ for some $\delta > 0$. Let $BQA(z_0)$ have a singular value decomposition $U\Sigma V$. Let $X := -\frac{1}{|BQA(z_0)|}V^*U^*$ (i.e. $X \in H^\infty_{n\times n}$

is a constant unitary matrix multiplied by a scalar belonging to the interval (0, 1))

$$(I + XBQA)(z_0) = V^*(I - \frac{1}{|BQA(z_0)|}\Sigma)V \Rightarrow \sigma_{min}(I + XBQA)(z_0) = 0 \qquad (15)$$

Hence (14) cannot hold, and $\|BQA\|_\infty \leq 1$ is a necessary and sufficient condition for robust stability.

□

The quantity to be optimized in (13), $\sup_{d\in\mathcal{D}} \|W_1 y\|_2$, is also expressed in terms of Q as follows. If C is a robustly stabilizing control law for $\mathcal{B}(P_0, V)$, it can be shown [57] that the output disturbance d and plant output y are related by

$$y = (I - P_0Q)(I + \Delta PQ)^{-1}d \qquad \Delta P := P - P_0$$

for each $P \in \mathcal{B}(P_0, V)$. Lemma 2.1 can then be used to express μ_{opt} in the form

$$\mu_{opt} = \inf_{\substack{Q \in H^\infty_{n\times n} \\ \|P_0QV\|_\infty \leq 1}} \; \sup_{X\in H^\infty_{n\times n},\ \|X\|_\infty<1} \|W_1(I - P_0Q)(I + VXP_0Q)^{-1}W\|_\infty \qquad (16)$$

A further simplification for (16) will be discussed in Sect. 2.4. However we will first consider the role of feedback in modifying the radius of the set of plant uncertainty, and show that this leads to an optimization having the same form as (16).

2.2 Attenuation of Plant Uncertainty

Zames posed the following fundamental question [57] regarding the ability of feedback to reduce plant uncertainty: if a plant lies in some set of uncertainty (e.g. $\mathcal{B}(P_0, V)$), what is the smallest radius of any set of closed-loop uncertainty that can be achieved with a single feedback control law? In order to avoid the trivial answer to this question that zero closed-loop uncertainty can be achieved by disconnecting the system from the input, so recognizing that reduction of plant uncertainty is never the sole objective of feedback design, some constraint must be placed on admissible feedback structures. Such a constraint was described in [57] where it was termed a 'nominal plant invariant normalization'. It consists of limiting the two degree of freedom control law (Fig. 5), such that upon closure of the loop the action of the feedback effectively maps the nominal open-loop plant onto itself. Note that there are other normalizations which rule out disconnection of the plant input as a means of optimally reducing closed loop plant uncertainty, but nominal plant invariance has the following three advantages:

i) The normalized feedback acts only on the difference between the true system and the nominal system (i.e. the 'error' dynamics), thereby isolating contraction of plant uncertainty from other aspects of performance.

ii) The assumption of nominal plant invariance provides an immediate nominal closed-loop plant model, i.e. P_0.

iii) Plant invariant controllers retain one complete degree of freedom [57].

The formal statement of plant invariance for the feedback scheme represented in Fig. 5 is stated in terms of the elements of the equivalent flowgraph shown in Fig. 6 as follows [57],

$$(I + CP_0)^{-1}U = I \tag{17}$$

In this context, the meaning of the term 'equivalent' is that the input-output behavior between the nodes shown in Figs. 5 and 6 is identical. Zames [57], showed that for plant invariant feedbacks where $P_0 \in H^{\infty}_{n\times n}$, stability of the loop represented by Fig. 5 for $P = P_0$ is the same as the statement $C(I + P_0C)^{-1} \in H^{\infty}_{n\times n}$ for the equivalent loop of Fig.6. It also follows from [57] that if this condition holds, the stability of the loop of Fig.5 for a general $P \in H^\infty_{n\times n}$ is equivalent to $C(I+PC)^{-1} \in H^\infty_{n\times n}$. [5] If $C(I+PC)^{-1} \in H^\infty_{n\times n}$, the closed-loop map K belongs to $H^\infty_{n\times n}$ and can be expressed (Theorem 7 [57])

$$\begin{aligned} K - P_0 &= (I - P_0Q)(I + \Delta PQ)^{-1}\Delta P \;\; \textit{where} \;\; Q := C(I + P_0C)^{-1}. &(18)\\ &=: \Delta_m P_0 &(19) \end{aligned}$$

$\Delta_m \in H^\infty_{n\times n}$ *represents the closed-loop multiplicative plant perturbation*

The potential of feedback to attenuate <u>plant</u> uncertainty can now be formally quantified.

Problem 2. We seek to minimize the 'worst case' W_1-weighted closed-loop multiplicative uncertainty radius, $\sup_{P\in\mathcal{B}(P_0,V)} \|W_1\Delta_m\|_\infty$*, over all robustly stabilizing (for $\mathcal{B}(P_0,V)$) nominal plant invariant control laws.*

Using Lemma 2.1 and the expression (18), Problem 2 can be stated in the form of the following optimization,

$$\mu_{opt} = \inf_{\substack{Q\,\in\, H^\infty_{n\times n}\\ \|P_0QV\|_\infty \le 1}} \; \sup_{X\in H^\infty_{n\times n}, \|X\|_\infty<1} \|W_1(I - P_0Q)(I + VXP_0Q)^{-1}VX\|_\infty \qquad (20)$$

In Lemma 2.2 we show that the last X term in (20) can be removed without effecting the supremum in (20). This establishes that the problem of optimal attenuation of plant uncertainty (20) under a nominal plant invariant normalization is equivalent to a special case of the MIMO form of the ORDAP represented in (16).

Lemma 2.2 *If* $\|P_0QV\|_\infty \le 1$ *and* $V^{\pm 1} \in H^\infty_{n\times n}$ *then*

$$\begin{aligned} &\sup_{X\in H^\infty_{n\times n},\, \|X\|_\infty<1} \|W_1(I - P_0Q)(I + VXP_0Q)^{-1}VX\|_\infty \\ = &\sup_{X\in H^\infty_{n\times n},\, \|X\|_\infty<1} \|W_1(I - P_0Q)(I + VXP_0Q)^{-1}V\|_\infty \qquad (21)\end{aligned}$$

[5]The feedback loop of Fig. 6 is required to be stable in the sense that it is equivalent to the stable loop of Fig. 5. U and C may in general be unstable.

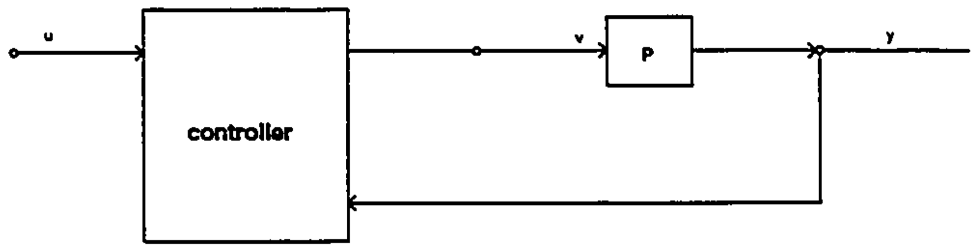

Figure 5: A two degree of freedom feedback loop.

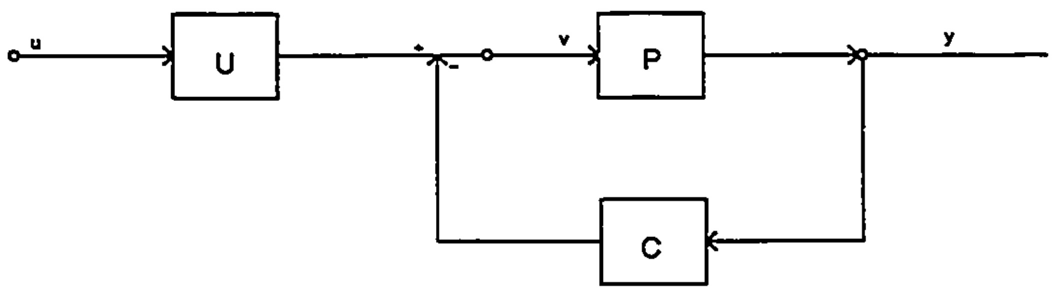

Figure 6: An equivalent representation for the loop of Fig. 5

Proof. Define $\lambda := \sup_{X \in H^{\infty}_{n \times n},\, \|X\|_{\infty} < 1} \|W_1(I - P_0Q)(I + VXP_0Q)^{-1}V\|_{\infty}$. Fix $\delta > 0$. There exists a $X \in H^{\infty}_{n \times n}, \|X\|_{\infty} < 1, z_0 \in D, \zeta \in \mathbb{C}^n, |\zeta| < 1$ such that

$$|W_1(I - P_0Q)(I + VXP_0Q)^{-1}V(z_0)\zeta| > \lambda - \delta \tag{22}$$

Let $\eta = (I + VXP_0Q)^{-1}V(z_0)\zeta \Leftrightarrow (V(z_0)^{-1} + XP_0Q(z_0))\,\eta = \zeta$
$\Leftrightarrow V(z_0)^{-1}\eta + XP_0QV(z_0)V(z_0)^{-1}\eta = \zeta$. First note that $P_0QV(z_0)V(z_0)^{-1}\eta(= P_0Q(z_0)\eta)$ is a vector of smaller Euclidean norm than $V(z_0)^{-1}\eta$ since $\sigma_{max}(P_0QV(z_0)) \leq 1$. There exists a unitary matrix Φ, such that $\Phi P_0Q(z_0)\eta$ is parallel to $-V(z_0)^{-1}\eta$, and this vector represents the closest point to $-V(z_0)^{-1}\eta$ on a sphere center zero radius $|P_0Q(z_0)\eta|$. Hence there is a scalar α in the interval $(1 - \delta, 1)$ such that

$V(z_0)^{-1}\eta + \alpha\Phi P_0 Q\eta =: \zeta'$ has Euclidean norm ≤ 1. This follows from the fact that $|V(z_0)^{-1}\eta + \Phi P_0 Q(z_0)\eta|$ must be less than or equal to $|V(z_0)^{-1}\eta + XP_0Q(z_0)\eta|$ by construction of Φ. Note that $\eta = (I + \alpha V\Phi P_0 Q)^{-1}V(z_0)\zeta'$. Since $|W_1(I - P_0Q)(z_0)\eta| > \lambda - \delta$ from (22), it follows that $|W_1(I - P_0Q)(I + \alpha V\Phi P_0Q)^{-1}V(z_0)\zeta'| > \lambda - \delta$, which in turn implies,

$$\|W(I - P_0Q)(I + \alpha V\Phi P_0Q)^{-1}V\alpha\Phi\|_\infty > \alpha(\lambda - \delta) \tag{23}$$

where in (23) Φ is taken to be the $H^\infty_{n\times n}$ function assuming the matrix Φ as its constant value. (23) has established the stated equality since δ is arbitrary.

□

Thus we have established that the optimal plant disturbance attenuation is

$$\mu_{opt} = \inf_{\substack{Q \in H^\infty_{n\times n} \\ \|P_0QV\|_\infty \leq 1}} \sup_{X \in H^\infty_{n\times n}, \|X\|_\infty < 1} \|W_1(I - P_0Q)(I + VXP_0Q)^{-1}V\|_\infty \tag{24}$$

2.3 Robust Performance: Weighted Sensitivity Minimization

Both problems 1 and 2 are in fact robust performance problems. In problem 1, the performance measure is the worst case output disturbance transmission, while in problem 2 it is the worst case deviation of the closed-loop system from the closed-loop nominal in the $H^\infty_{n\times n}$ norm. This point can be emphasized by considering an explicit form of the H^∞ robust performance problem, obtained by minimizing the weighted sensitivity norm in the presence of plant uncertainty. This problem is represented on the left in (25). The arguments of Sects. 2.1 and 2.2 can be used to show that for $\mathcal{B} = \{P \in H^\infty_{n\times n} \ : \ P = (I + XV)P_0, \ X \in H^\infty_{n\times n}, \ \|X\|_\infty < 1\}$, the optimal robust performance takes the form

$$\inf_{\substack{C \text{ stabilizing} \\ \text{all } P \in \mathcal{B}}} \|W(I+PC)^{-1}\|_\infty$$

$$= \inf_{\substack{Q \in H^\infty_{n\times n} \\ \|VP_0Q\|_\infty \le 1}} \sup_{X \in H^\infty_{n\times n},\ \|X\|_\infty < 1} \|W(I-P_0Q)(I+XVP_0Q)^{-1}\|_\infty \quad (25)$$

Thus Problem 1, Problem 2 and the explicit robust performance problem of (25), are all special cases of a generic minimization represented by (16). Therefore, the definition of the acronym ORDAP can be extended to include all three cases, in recognition of the fact that the disturbance can be taken either in the context of plants or signals.

2.4 Conversion of ORDAP to a Fixed-Point Problem

Under the assumptions of Theorem 2.1 below, the non-convex constrained optimizations of (16),(24) and (25) are equivalent to a fixed-point problem expressed in terms of a function taking values equal to the optima of the MIMO extension of the two-disc minimization of (2). Before stating this theorem, the following definition is required.

Definition 2.1 *$W, V \in H^\infty_{n\times n}$ are said to be* <u>*commensurate*</u> *if*

$$W = w_s A,\ V = v_s A \quad (26)$$

where w_s, v_s are scalar valued H^∞ functions and $A \in H^\infty_{n\times n}$.

Note that in the SISO case all pairs of H^∞ functions are commensurate.

Theorem 2.1

1.) If W and V are commensurate, $W, V^{\pm 1} \in H^\infty_{n\times n}$, and P_0 is not invertible in $H^\infty_{n\times n}$, then i. and ii. are equivalent.

i. C is a robustly stabilizing controller for the set of stable systems described by the set $\mathcal{B}(P_0, V)$, and

$\sup_{X \in H^\infty_{n\times n},\ \|X\|_\infty < 1} \|W_1(I - P_0Q)(I + VXP_0Q)^{-1}W\|_\infty \le r$ *($Q = C(I + P_0C)^{-1}$).*

ii. $\operatorname{ess\,sup}_{\theta\in[0,2\pi)} \sup_{\zeta\in\mathbb{C}^n, |\zeta|\le 1} \left(|W_1(I - P_0Q)W(e^{i\theta})\zeta| + r|P_0QV(e^{i\theta})\zeta|\right) \le r$

2.) Under the assumptions of 1.), if an optimal $Q \in H^\infty_{n\times n}$ exists in (27) for r equal to the smallest fixed-point of χ [6] and $\chi(0) > 0$, then the optimal robust disturbance attenuation μ_{opt} of (16) is equal to the smallest positive fixed-point of the function

$$\chi(r) = \inf_{Q\in H^\infty_{n\times n}} \operatorname{ess}\sup_{\theta\in[0,2\pi)} \sup_{\zeta\in\mathbb{C}^n, |\zeta|\le 1} \left(|W_1(I - P_0Q)W(e^{i\theta})\zeta| + r|P_0QV(e^{i\theta})\zeta|\right) \quad (27)$$

Remarks

1) The existence condition of Theorem 2.1 2.) is shown to be satisfied under very general conditions in Theorem 3.8 of Sect. 3.6. For example, the two conditions $V^{\pm 1} \in H^\infty_{n\times n}$ and the outer part of P_0 invertible in $H^\infty_{n\times n}$ are sufficient.

2) The weights W and V are *always* commensurate when ORDAP originates from the optimal plant disturbance attenuation problem of Sect. 2.2 (c.f (24)), or W, V and P_0 are scalar valued.

3) Theorem 2.1 can also be proven for the left weighted robust performance problem of (25) by the same argument [60].

Proof of Theorem 2.1 . 1.) The fact that W, V are commensurate, and $W, V^{\pm 1} \in H^\infty_{n\times n}$ implies the existence of a scalar H^∞ function λ such that $\lambda(z)I = V(z)^{-1}W(z)$.

[6] existence of the smallest fixed-point is proven in Appendix A

To prove i. $\Rightarrow$ ii. we note that if i. holds iff,

$$|W_1(I-P_0Q)(I+VXP_0Q)^{-1}W(z)\zeta| \le r|\zeta|$$

$$\forall X \in H^\infty_{n\times n},\ \|X\|_\infty < 1,\ \forall z \in D,\ \forall \zeta \in \mathbf{C}^n$$

$$\textit{and } \|P_0QV\|_\infty \le 1 \quad (\textit{from Lemma } 2.1)$$

$$\Leftrightarrow |W_1(I-P_0Q)V(I+XP_0QV)^{-1}V^{-1}W(z)\zeta| \le r|\zeta|$$

$$\forall X \in H^\infty_{n\times n},\ \|X\|_\infty < 1,\ \forall z \in D,\ \forall \zeta \in \mathbf{C}^n$$

$$\textit{and } \|P_0QV\|_\infty \le 1$$

$$\Leftrightarrow |\lambda(z)W_1(I-P_0Q)V(z)\eta| \le r|(I+XP_0QV(z))\eta| \tag{28}$$

$$\forall X \in H^\infty_{n\times n},\ \|X\|_\infty < 1,\ \forall z \in D,\ \forall \eta \in \mathbf{C}^n$$

$$\textit{and } \|P_0QV\|_\infty \le 1 \tag{29}$$

$$\Leftrightarrow |W_1(I-P_0Q)W(z)\eta| \le r|\eta + XP_0QV(z)\eta| \tag{30}$$

$$\forall X \in H^\infty_{n\times n},\ \|X\|_\infty < 1,\ \forall z \in D,\ \forall \eta \in \mathbf{C}^n,\ |\eta| = 1$$

$$\textit{and } \|P_0QV\|_\infty \le 1 \tag{31}$$

Now fix $z_0 \in D$, $\eta_0 \in \mathbf{C}^n$, $|\eta_0| = 1$. There exists a constant unitary matrix U such that $UP_0QV(z_0)\eta_0$ is parallel to $-\eta_0$ in $\mathbf{C}^n$. Chose the $H^\infty_{n\times n}$ function X to be constant, and equal to the matrix $(1-\delta)U$ for some arbitrary $\delta \in (0,1)$. Since $|P_0QV(z_0)\eta_0| \le |\eta_0|$, (30) implies that

$$|W_1(I-P_0Q)W(z_0)\eta_0| \le r|\eta_0| - (1-\delta)r|P_0QV(z_0)\eta_0| \tag{32}$$

Since (32) does not include X, and z_0, δ, and η_0 were chosen arbitrarily, (32) implies that

$$\sup_{z\in D} \sup_{\substack{\zeta \in \mathbf{C}^n \\ |\zeta| \le 1}} |W_1(I-P_0Q)W(z)\zeta| + r|P_0QV(z)\zeta| \le r \tag{33}$$

Invoking the lemma of Appendix E to show that no change in (33) is incurred by taking the essential supremum over the boundary of the disk instead of the interior, ii. is shown to hold.

To prove ii. $\Rightarrow$ i., first note that ii. implies the following is true (after applying the lemma of Appendix E as above),

$$|W_1(I - P_0Q)W(z)\eta| \le r\,(|\eta| - |P_0QV(z)\eta|) \;\; and \;|P_0QV(z)\eta| \le 1 \qquad (34)$$
$$\forall\eta \in \mathbf{C}^n,\; |\eta| \le 1,\; \forall z \in D$$

(Note that we have made use of the fact that the assumptions ensure that r is not zero.) Which in turn implies that (30) and (31) are satisfied for all $H^\infty_{n\times n}$ strict contractions X and for all $z \in D$, $\eta \in \mathbf{C}^n$, $|\eta| \le 1$. As shown this is equivalent to i. Hence 1. is proven.

2.) A smallest positive fixed-point of χ exists as a consequence of the lemma of Appendix A. Let this fixed-point be r_0, and let the optimal $Q \in H^\infty_{n\times n}$ for (27) when $r = r_0$ be Q_0 (which exists by assumption). From 1.) of this theorem it follows that $r_0 \ge \mu_{opt}$. To prove this inequality in the other direction, fix $\delta > 0$. From 1.) of this theorem there exists a $Q \in H^\infty_{n\times n}$ such that the statement of 1 ii. holds with $r = \mu_{opt} + \delta$. Hence $\mu_{opt} + \delta$ is an upper bound for at least one fixed-point of χ, since $\chi(\cdot)$ is continuous (lemma of Appendix A) and $\chi(0) > 0$. Thus $r_0 \le \mu_{opt}$

□

The optimization represented by χ in (27) can be simplified for commensurate weights. Since W and V are commensurate, the scalar dependence of W on V can be taken over to the left side of the first term of (27). Thus, under the assumptions of Theorem 2.1 (27) takes the following form

$$\inf_{Q\in H^\infty_{n\times n}} ess \sup_{\theta\in[0,2\pi)} \sup_{\zeta\in\mathbf{C}^n,|\zeta|\le 1} \left(|W_2(I - P_0Q)V(e^{i\theta})\zeta| + r|P_0QV(e^{i\theta})\zeta|\right) \qquad (35)$$

for some $W_2 \in H^\infty_{n\times n}$. Since by assumption $V^{\pm 1} \in H^\infty_{n\times n}$, the V term can be 'absorbed' into the free parameter Q. If P_0 has an inner-outer factorization over $H^\infty_{n\times n}$ [26] given by BH_0, and W_2B has an inner-outer factorization given by $U\tilde{W}_2$, then (35) can be written

$$\inf_{Q\in H^\infty_{n\times n}} ess \sup_{\theta\in[0,2\pi)} \sup_{\zeta\in\mathbb{C}^n, |\zeta|\le 1} \left(|(U^* W_2 V - \tilde{W}_2 H_0 Q)\zeta| + r|H_0 Q\zeta|\right) \tag{36}$$

The optimization (36) is included in the following form, where $W, \tilde{W}, \tilde{V} \in H^\infty_{n\times n}$ are general outer functions, and $U \in H^\infty_{n\times n}$ is an inner function.

$$\inf_{Q\in H^\infty_{n\times n}} ess \sup_{\theta\in[0,2\pi)} \sup_{\zeta\in\mathbb{C}^n, |\zeta|\le 1} \left(|(U^* W - \tilde{W} Q)\zeta| + r|\tilde{V} Q\zeta|\right) \tag{37}$$

Remark. The results of this chapter are substantially simpler to establish in the SISO case where the order of system cascade is not important. In this case $\chi(\cdot)$ can be expressed in the form,

$$\chi(r) = \inf_{Q\in H^\infty} \||W(1-P_0 Q)| + r|VP_0 Q|\|_\infty \tag{38}$$

2.5 A Motivating Example Based on a 'Two Arc' Result

Before beginning a general analysis, we will examine the ORDAP for an illustrative special case of SISO systems. This example is reminiscent of the 'Two Arc' theorem of complex analysis, in that it relates extremal values of the sensitivity function on one arc of the circle subject to a constraint on the complementary arc. Two purposes are served:

i. a significant limitation of the existing approach [20] is pointed out and the need for an *exact* theory of the ORDAP is highlighted.

ii. in dealing with the more general cases in later chapters, an illustration of a limiting behavior of the ORDAP is provided.

For this example we will be able to obtain precise estimates of the sensitivity of the ORDAP to errors in $\chi(r)$, and to show that a quadratic norm 'two block' approximation for $\chi(r)$ [20], with its attendant $\sqrt{2}$ ratio of upper and lower bounds, could result in arbitrarily poor estimation of ORDAP.

In Lemma 2.3 we derive estimates of $\chi(r)$ for pairs of outer H^∞ weights W, V and W_n, V_n, $n = 1, 2, \ldots$, whose magnitudes are illustrated in Figs. 7 and 8, and for nominal plants which can be approximately 'absorbed' into the free parameter Q. Note that uniform Lipschitz continuity of $log|W_n(e^{i\theta})|$ and $log|V_n(e^{i\theta})|$ implies that the boundary values of W_n, V_n are uniformly Lipschitz continuous, and so conform to assumption A2 of Chapter 3. The continuity of boundary values of W_n and V_n follows from Sect. E, Chapter V of [32].

Lemma 2.3. *If W, V, W_n, V_n are outer H^∞ weightings with magnitudes as shown in Figs. 7 and 8* then the following hold:

$$i. \quad \frac{\lambda}{\lambda+1} \le \inf_{Q \in H^\infty} \| |W(1-Q)| + |VQ| \|_\infty \le \frac{\lambda}{\lambda+1} + \delta \tag{39}$$

$$ii. \quad \lim\, inf_{n\to\infty} \inf_{Q \in H^\infty} \| |W_n(1-Q)| + |V_nQ| \|_\infty \ge \frac{\lambda}{\lambda+1} \tag{40}$$

$$\lim\, sup_{n\to\infty} \inf_{Q \in H^\infty} \| |W_n(1-Q)| + |V_nQ| \|_\infty \le \frac{\lambda}{\lambda+1} + \delta \tag{41}$$

Proof.

i. The upper bound in (39) is validated by selecting Q to be the constant $\frac{1}{\lambda+1}$. To establish the lower bound, suppose there exists an $\epsilon > 0$ such that for some

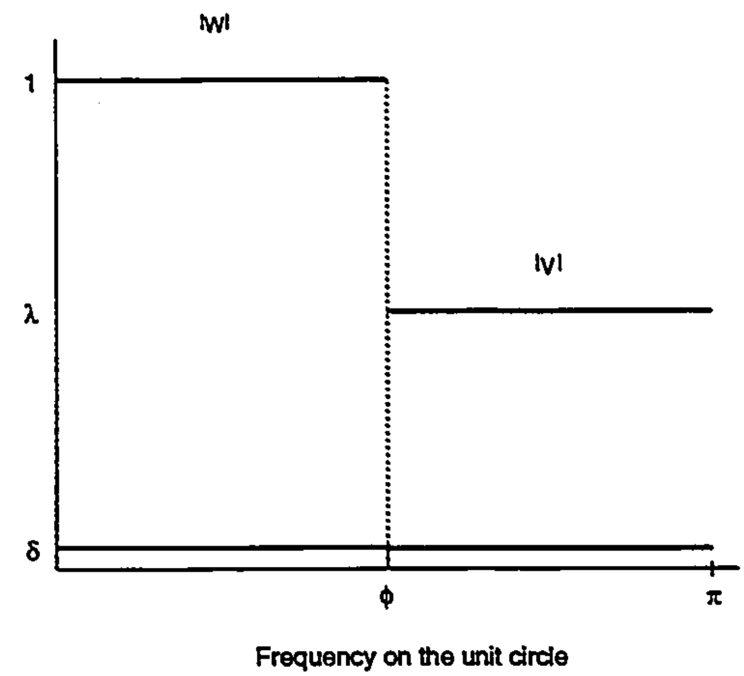

Figure 7: Magnitudes of weightings W, V in Lemma 2.1

$Q \in H^\infty$

$$|||W(1-Q)| + |VQ|||_\infty < \frac{\lambda}{\lambda+1} - \epsilon \tag{42}$$

Then it follows that

$$\begin{aligned} ReQ(e^{i\theta}) &> \frac{1}{\lambda+1} + \epsilon \quad for\ \theta \in [0,\phi] \\ ReQ(e^{i\theta}) &< \frac{1}{\lambda+1} - \frac{\epsilon}{\lambda} \quad for\ \theta \in (\phi,\pi] \end{aligned}$$

Thus $\lim_{\epsilon'\to 0}\int_{|\theta-\phi|>\epsilon'} \frac{ReQ(e^{i\theta})}{\theta-\phi}d\theta = +\infty$ which implies, in the light of the Hilbert transform for conjugate harmonic functions, that $lim_{r\uparrow 1}|ImQ(re^{i\phi})| = \infty$ (c.f. [21] Example 1.5, Chapter 4). This contradicts the assertion that $Q \in H^\infty$, with the result that (42) cannot hold and i. is proven.

ii. (41) can be established simply by chosing Q to be the constant $\frac{1}{\lambda+1}$.

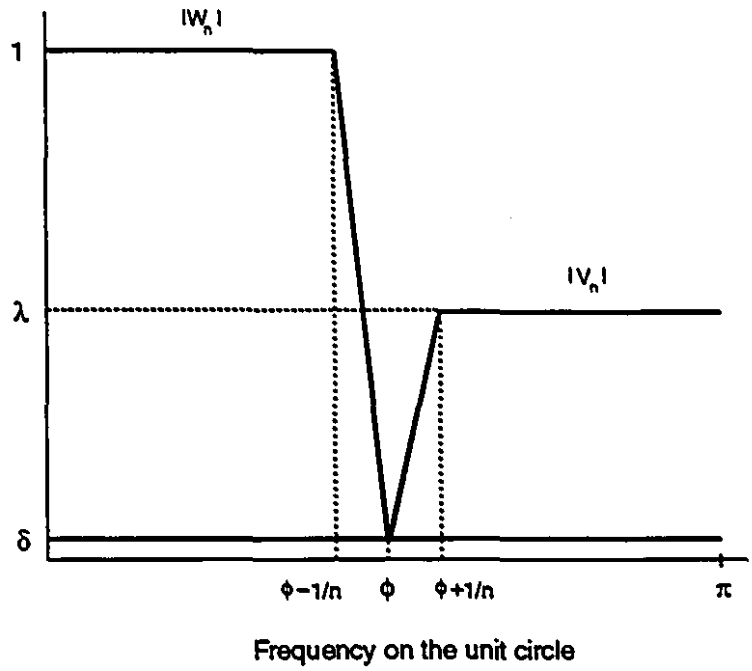

Figure 8: Magnitudes of weightings W_n, V_n in Lemma 2.1

To establish (40) suppose the contrary, i.e. there exists an $\epsilon > 0$ and a uniformly bounded sequence $Q_n \in H^\infty$ such that

$$lim\ inf_{n\to\infty} \| |W_n(1-Q_n)| + |V_nQ_n| \|_\infty < \frac{\lambda}{\lambda+1} - \epsilon \qquad where\ W, V, \tilde{Q} \in H^\infty$$

By a normal family argument we may select a subsequence of the integers $\{n_k\}_{k=1}^\infty$ such that

$$W_{n_k} \to W, \ \ V_{n_k} \to V, \ \ Q_{n_k} \to \tilde{Q} \quad (where\ W, V, \tilde{Q} \in H^\infty)$$

and the convergence is on compact subsets of the open unit disc. Thus for each $z \in D$, $|W(1-\tilde{Q})(z)| + |V\tilde{Q}(z)| \leq \frac{\lambda}{1+\lambda} - \epsilon$. This implies that $\| |W(1-\tilde{Q})| + |V\tilde{Q}| \|_\infty \leq \frac{\lambda}{\lambda+1} - \epsilon$, which violates (39), and so (40) must hold.

□

Now consider the ORDAP for the case of continuous weightings W_n, V_n as in Fig. 8, with λ set nominally equal to 1. By Lemma 2.3, for arbitrary $\epsilon > 0$, there exist nominal plants P_0 (possibly not outer) and integer n for which the optimization represented by $\chi(r)$ satisfies,

$$\frac{r}{r+1} - \epsilon \leq \chi(r) \leq \frac{r}{r+1} + \epsilon + \delta$$

The slope of $\frac{r}{r+1}$ as a function of r is equal to unity at the origin and tapers off to $\frac{1}{4}$ at $r = 1$. Thus for sufficiently small ϵ and δ, and large enough n, the smallest fixed-point of χ can be expected to be arbitrarily sensitive to inaccurate estimation of $\chi(r)$. Indeed for sufficiently small ϵ and δ, and large enough n, the fixed-point obtained by solving $\chi(r) = r$, representing the true solution to ORDAP, has an upper bound approaching $\sqrt{\epsilon + \delta}$. On the other hand, the best achievable upper bound by quadratic norm approximation is obtained from the solution of $\sqrt{2}\chi(r) = r$, which approaches $\frac{\sqrt{2}-1}{\sqrt{2}} = 0.2979$. Thus for sufficiently small $\epsilon + \delta$, the quadratic norm approximation overestimates the possible solutions to the ORDAP by a factor approaching $\frac{0.2979}{\sqrt{\epsilon+\delta}}$, which becomes arbitrarily large as $(\epsilon + \delta) \rightarrow 0$.

Remarks. The point here is that in situations where the fixed-point is sensitive to correct estimation of $\chi(r)$, a more exact optimization theory is required than that provided by the quadratic two-block problem. The duality theory developed in this thesis and the related convex optimization do provide exact estimates of $\chi(r)$. These can be used to obtain accurate solutions to ORDAP.

Chapter 3

Two Key Mixed Sensitivity Problems

Here we will consider the following pair of two-block optimizations, both of which are MIMO extensions of the two-disc problem of (2).

$$\mu_0 := \inf_{Q \in H^\infty_{n\times n}} \| \, |U^* W - \tilde{W} Q| + |\tilde{V} Q| \|_\infty \tag{43}$$

$$\mu_1 := \inf_{Q \in H^\infty_{n\times n}} ess \sup_{\theta \in [0,2\pi)} \sup_{\zeta \in \mathbb{C}^n, |\zeta| \leq 1} \left(|(U^* W - \tilde{W} Q)\zeta| + |\tilde{V} Q \zeta| \right) \tag{44}$$

As noted in the introduction these can not be handled by established methods, forcing us to find an alternative approach. This approach is the main subject of the present chapter.

In sections 3.1-3.3 Banach space duality theory will be used to characterize the solutions of (43) and (44). The former will be shown to be allpass in general and unique in the SISO case. 'Nearly' optimal control laws will be shown to satisfy an approximate allpass condition. The theory derived here will form the basis of a qualitative analysis of feedback and uncertainty discussed in Sect. 3.5, and will lead to a numerical solution method (in Chapter 4) involving a combination of duality and convex optimization.

The first optimization (43) is simpler than the second (44). In general (43) is clearly an upper bound for (44), and the two are identical in the SISO case. It will follow that by optimizing the function $\hat{\chi}(r)$ i.e. replacing $\chi(r)$ in (27) of Chapter 2 by

$$\tilde{\chi}(r) := \inf_{Q \in H^\infty_{n \times n}} \| \, |U^* W - \tilde{W} Q| + r|\tilde{V} Q| \|_\infty \tag{45}$$

we shall obtain upper bounds for the MIMO optimal robust disturbance attenuation (which are exact in the SISO case). Henceforth we shall confine our attention to (45), except in Sect. 3.6.

3.1 A Distance Problem Equivalent to (43).

The following assumption will be made initially:

(A1) *there exists* $\epsilon_0 > 0$ *such that* $(\tilde{W}^* \tilde{W} + \tilde{V}^* \tilde{V})(e^{i\theta}) \geq \epsilon_0$ *for all* $\theta \in [0, 2\pi)$. Assumption (A1) initially excludes strictly proper plants P_0 if (43) is derived from either the ORDAP or the two-disturbance problem in the manner described at the end of Sect 2.4. However, in Sect. 3.4, (A1) will be relaxed in order to allow strict propriety of the nominal system P_0.

(43) is equivalent to finding the shortest distance from a vector to a Banach subspace, defined as follows. Let B be the Banach space $L^\infty_{n \times n} \times L^\infty_{n \times n}$ consisting of pairs of matrix-valued functions on the unit circle, under the norm

$$\|K\|_B := esssup_{\theta \in [0, 2\pi)} \left(|K_1(e^{i\theta})| + |K_2(e^{i\theta})| \right), \quad K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}$$

Then (43) is equivalent to

$$\mu_0 = \inf_{Q \in H^\infty_{n \times n}} \left\| \begin{bmatrix} U^* W \\ 0 \end{bmatrix} - \begin{bmatrix} \tilde{W} \\ \tilde{V} \end{bmatrix} Q \right\|_B \tag{46}$$

(46) is the distance from $\begin{bmatrix} U^* W \\ 0 \end{bmatrix}$ to the subspace $S := \begin{bmatrix} \tilde{W} \\ \tilde{V} \end{bmatrix} (H^\infty_{n \times n})$ of B. Assumption (A1) ensures that S is a closed subspace. Note that (46) differs from the standard two-block problem of, for example, [49].

S also has an equivalent description: Let $R = \begin{bmatrix} R_1 \\ R_2 \end{bmatrix} \in H^\infty_{n\times n} \times H^\infty_{n\times n}$ be the outer isometry (see [26], Lecture vii, p66) whose range coincides with the range of $\begin{bmatrix} \tilde{W} \\ \tilde{V} \end{bmatrix}$. More explicitly R has the form $R = \begin{bmatrix} \tilde{W}\Lambda^{-1} \\ \tilde{V}\Lambda^{-1} \end{bmatrix}$ where Λ is the outer spectral factor of $\tilde{W}^*\tilde{W} + \tilde{V}^*\tilde{V}$. Then $S = R(H^\infty_{n\times n})$. *Our notation should not be confused with that for rational functions in* $H^\infty_{n\times n}$.

Let us evaluate (46) by duality.

3.2 Existence of a Predual and an Optimal Controller

A^* denotes the dual space of any Banach space A. If A_0 is a subspace of A then $(A_0)^\perp$ is the subspace of A^* which annihilates A_0, $(A_0)^\perp := \{f \in A^* : f(x) = 0 \; for \; all \; x \in A_0\}$. Isometric isomorphism between Banach spaces will be denoted by $\simeq$.

A_* is a *Predual* of A if $(A_*)^* \simeq A$, and a subspace $A_{0\perp}$ of A_* is a preannihilator of a subspace A_0 of A if, under the preceding isomorphism, $(A_{0\perp})^\perp \simeq A_0$. A standard result of Banach space duality theory asserts ([35], Ch. 5.8, Theorem 2) that when a predual and preannihilator exist as above and for $K \in B$, the identity

$$\min_{Q\in A_0} \|K - Q\|_A = \sup_{\phi\in A_{0\perp},\ \|\phi\|_{A_*}\leq 1} |\phi(K)| \tag{47}$$

holds. Let us establish the existence of a predual and determine the form of the preannihilator, for our problem, i.e. when $A := B$ and $A_0 := S$.

Introduce the notation, for any matrix A, $STr(A) = Tr(\{A^*A\}^{\frac{1}{2}}) = \sum_{i=1}^n \sigma_i(A)$, where Tr denotes the trace of A and $\sigma_i(A)$ are the singular values of A.

Let B_* be the Banach space $L^1_{n\times n} \times L^1_{n\times n}$ under the norm

$$\|G\|_{B_*} := \int_0^{2\pi} \Big\{Max(STrG_1, STrG_2)\Big\}(e^{i\theta})d\theta \tag{48}$$

Let $S_\perp$ be the subspace of B_*,

$$S_\perp := (I - RR^*)(L^1_{n\times n} \times L^1_{n\times n}) \oplus R(\overline{H}^1_0)_{n\times n} \tag{49}$$

where $(\overline{H}^1_0)_{n\times n}$ denotes the subspace of $L^1_{n\times n}$ consisting of integrable functions whose positive Fourier coefficients are equal to zero. Note that $\oplus$ denotes a direct sum of two subspaces.

Before identifying the dual of $B_\perp$ and the annihilator of $S_\perp$, let us establish some facts concerning bounded linear functionals on B_*. Every such functional ϕ must have the representation

$$\phi(G) = \sum_{i,j=1}^{n} \int_0^{2\pi} (G_{1,i,j}\overline{K}_{1,i,j} + G_{2,i,j}\overline{K}_{2,i,j})(e^{i\theta})d\theta \tag{50}$$

for some $K_{r,i,j} \in L^\infty$, where $r = 1,2$ and $i, j = 1, 2, ...n$. Indeed, by the Riesz theorem the representation (50) is certainly valid when ϕ is restricted to the 1 dimensional subspace of B_* spanned by $G_{r,i,j}$ for a fixed index r, i, j, and the general representation follows by linearity. (50) can be expressed more compactly in terms of the bilinear forms $[\cdot|\cdot]$ on $B_* \times B$,

$$[G|K] = \int_0^{2\pi} \{TrK_1^*G_1 + TrK_2^*G_2\}(e^{i\theta})d\theta \tag{51}$$

where $K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} = \begin{bmatrix} [K_{1,i,j}] \\ [K_{2,i,j}] \end{bmatrix} \in B$, $G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} = \begin{bmatrix} [G_{1,i,j}] \\ [G_{2,i,j}] \end{bmatrix} \in B_*$. Each functional ϕ on B_* can be expressed as $\phi(G) = [G|K]$ for some $K \in L^\infty_{n\times n} \times L^\infty_{n\times n}$. The inequality

$$\begin{aligned} \|[G|K]\| &\leq esssup_{\theta\in[0,2\pi)}\Big(|K_1(e^{i\theta})| + |K_2(e^{i\theta})|\Big) \cdot \\ &\qquad \int_0^{2\pi} Max\Big(STr(G_1(e^{i\theta})), STr(G_2(e^{i\theta}))\Big)d\theta \\ &= \|G\|_{B_*} \cdot \|K\|_B \end{aligned} \tag{52}$$

will be needed, and can be deduced from lemma of Appendix C.

Theorem 3.1. *a)* $(B_*)^* \simeq B$

b) $(S_\perp)^\perp \simeq S$ *under the isomorphism of part a).*

Proof. a) It has been stated that every bounded linear functional $\phi(.)$ on B_* has the form (51). Conversely, for any $K \in B$, the functional $\phi(G) = [G|K]$ is bounded on B_* and satisfies $\|\phi\| \le \|K\|_B$ by (52). To prove that $\|\phi\| = \|K\|_B$, observe that by definition of the essential supremum, there must exist a sequence of Lebesgue sets $\{\Omega_k\}_{k=1}^{\infty}$ of strictly positive Lebesgue measure $m(\Omega_k)$, such that for $\theta \in \Omega_k$, $|K_1(e^{i\theta})| + |K_2(e^{i\theta})| \ge \|K\|_B - \frac{1}{k}$. By the lemma of Appendix C, it is possible to define $G^{(k)}$ so that for each $\theta \in \Omega_k$, $TrG_r^{(k)}(K_r)^* = |K_r| \sum_{i=1}^{n} \sigma_i(G_r^{(k)})$ for $r = 1, 2$, and $\sum_{i=1}^{n} \sigma_i(G_r^{(k)}) = \frac{1}{m(\Omega_k)}$; and for each $\theta \notin \Omega_k$, $G_r^{(k)} = 0$. Hence $\|G\|_{B_*} = 1$ and $|\phi(G)| \ge \|K\|_B - \frac{1}{k}$, which implies that $\|\phi\| = \|K\|_B$. If $\simeq$ is taken to be the isomorphism between $K \in B$ and functionals $[\cdot|K]$ on B_*, Theorem 3.1 a) is proven.

b) If $\phi \in (B_*)^*$, then ϕ has the form $[\cdot|K]$ for some $K \in B$, and we have the following equivalences.

$$\begin{aligned}
\phi \in (S_\perp)^\perp \;\Leftrightarrow\;& [G|K] = 0 \;\forall G \in S_\perp \\
\Leftrightarrow\;& \int_0^{2\pi} Tr[K_1^* \; K_2^*]\left(RX_0 + (I - RR^*)\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right)(e^{i\theta})d\theta = 0 \\
& \qquad \forall X_0 \in (\overline{H}_0^1)_{n\times n},\; \forall X_1, X_2 \in L^1_{n\times n} \\
\Leftrightarrow\;& \int_0^{2\pi} \Big(Tr[K_1^* \; K_2^*]RX_0\Big)(e^{i\theta})d\theta = 0 \;\forall X_0 \in (\overline{H}_0^1)_{n\times n}, \\
& \qquad \textit{and} \int_0^{2\pi}\left(Tr[K_1^* \; K_2^*](I - RR^*)\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}\right)(e^{i\theta})d\theta = 0 \\
& \qquad \forall X_1, X_2 \in L^1_{n\times n} \\
\Leftrightarrow\;& [K_1^* \; K_2^*]R \in \overline{H}^{\infty}_{n\times n} \;\textit{and}\; [K_1^* \; K_2^*](I - RR^*)(e^{i\theta}) = 0 \quad a.e.(\theta) \\
\Leftrightarrow\;& K \in RH^{\infty}_{n\times n} \Leftrightarrow \phi \in S \;(S \simeq (S_\perp)^\perp)
\end{aligned}$$

$\square$

The following theorem is a corollary to Theorem 3.1, by the duality theorem cited at the beginning of Sect 3.2.

Theorem 3.2. *Under assumption A1 of Sect. 3.1 there exists at least one $Q_{opt} \in H^{\infty}_{n\times n}$ which satisfies*

$$\| |U^* W - \tilde{W} Q_{opt}| + |\tilde{V} Q_{opt}| \|_{\infty} = \inf_{Q \in H^{\infty}_{n\times n}} \| |U^* W - \tilde{W} Q| + |\tilde{V} Q| \|_{\infty} \tag{53}$$

Remark. If (43) is derived from the MIMO two-disc problem described by

$$\mu_0 = \inf_{Q \in H^{\infty}_{n\times n}} \| \, |W_1(I - P_0 Q)W| + |P_0 Q V| \, \|_{\infty} \tag{54}$$

in the manner related at the end of Sect. 2.4, then the conclusion of Theorem 3.2 can be expressed

$$\| \, |W_1(I - P_0 \tilde{Q}_{opt})W| + |P_0 \tilde{Q}_{opt} V| \, \|_{\infty} = \inf_{Q \in H^{\infty}_{n\times n}} \| \, |W_1(I - P_0 Q)W| + |P_0 Q V| \, \|_{\infty} \tag{55}$$

The $\tilde{Q}_{opt}$ of (55) and Q_{opt} of (53) are related by the equality $\tilde{Q}_{opt} = Q_{opt} V^{-1}$.

3.3 Allpass Property of the Optimum: Alignment in the Dual

In this section we assume that

(A2) $U^* W$, $\tilde{W}$, $\tilde{V}$, $U\tilde{W}$ *are continuous, as is the outer spectral factor of* $\tilde{W}^* \tilde{W} + \tilde{V}^* \tilde{V}$.

Continuity of $U^* W$ and $U\tilde{W}$ means that W and $\tilde{W}$ zero out the essential singularities of U (taking into consideration that these singularities may effect only some components of U).

Let C denote the set of continuous functions on the circle, and for any space A let A_c denote the subspace $A \cap C$. We seek the dual space $(B_c)^*$ of B_c. It will be shown that $(B_c)^*$ is isometrically isomorphic to a space $B^\sim$, consisting of $BV_{n\times n} \times BV_{n\times n}$ under a special norm, where $BV_{n\times n}$ denotes the space of functions of bounded variation on the unit circle which are assumed continuous from the right.

Accordingly, let $B^\sim$ be the space of pairs $\nu = (\nu_1,\ \nu_2) \in BV_{n\times n} \times BV_{n\times n}$, and introduce the bilinear form on $B^\sim \times B_c$,

$$< \nu|K >= \int_{[0,2\pi)} Tr\Big(K_1^* d_\theta \nu_1(\theta)\Big) + \int_{[0,2\pi)} Tr\Big(K_2^* d_\theta \nu_2(\theta)\Big), \quad K_1, K_2 \in C_{n\times n} \tag{56}$$

This form has the following equivalent representation: Let w_ν be the sum of the total variations on $[0,\theta)$ of all entries of ν_1 and ν_2. If we denote these by $w_{\nu_{1,i,j}}(\theta)$, $w_{\nu_{2,i,j}}(\theta)$, then

$$w_\nu(\theta) := \sum_{i,j=1,2,\ldots,n} w_{\nu_{1,i,j}}(\theta) + w_{\nu_{2,i,j}}(\theta), \quad \theta \in [0, 2\pi). \tag{57}$$

By the Radon Nykodym theorem, there exist $G_{\nu,r} \in L^1_{n\times n}(w_\nu)$, $r = 1,2$, in terms of which the pair of integrals (56) can be reduced to a single integral

$$< \nu|K >= \int_{[0,2\pi)} \Big\{Tr(K_1^* G_{\nu,1}) + Tr(K_2^* G_{\nu,2})\Big\} d_\theta w_\nu(e^{i\theta}) \tag{58}$$

The norm on $B^\sim$ is now defined to be

$$\|\nu\|_{B^\sim} := \int_{[0,2\pi)} Max\Big(STrG_{\nu,1}(e^{i\theta}), STrG_{\nu,2}(e^{i\theta})\Big) d_\theta w_\nu(e^{i\theta})$$

Note that the fact that (59) defines a norm on $B^\sim$, and that the metric space $B^\sim$ is complete, will follow from the isometric isomorphism between $B^\sim$ and the dual space of B_c that will be established in Theorem 3.3.

By reasoning analogous to that used in (52), the inequality

$$| < \nu | K > | \leq \|K\|_B \; \|\nu\|_{B^\sim} \tag{59}$$

is obtained.

$(S_c)^\perp$ will be identified with a subspace $S^\sim$ of $B^\sim$, defined as follows:

$$S^\sim := \left\{ \nu \in B^\sim : \; \nu(\theta) = \int_{[0,\theta]} (I - RR^*) d\nu'(\theta_1) + RG d\theta_1, \; \nu' \in B^\sim, \; G \in (\overline{H}_0^1)_{n \times n} \right\} \tag{60}$$

The following lemma establishes that the distance from $\begin{bmatrix} U^*W \\ 0 \end{bmatrix} \in B_c$ to S_c is the same as to S. Note that assumption A2 implies that R is continuous on the closed unit disc.

Lemma 3.1. *Under assumptions A1 and A2,*

$$\left\| \begin{bmatrix} U^*W \\ 0 \end{bmatrix} - RQ_{opt} \right\|_B = \inf_{X \in S_c} \left\| \begin{bmatrix} U^*W \\ 0 \end{bmatrix} - X \right\|_B = \mu_0 \tag{61}$$

Proof. This is a generalization of the proof of Lemma 1.6 of [21]. Let the subscript $r > 0$ denote the scaling of the complex disc, $F_r(z) = F(rz)$. Given any $\epsilon > 0$, it will be shown that there exists $r < 1$ for which

$$\left\| \begin{bmatrix} U^*W \\ 0 \end{bmatrix} - R(Q_{opt})_r \right\|_B \leq \mu_0 + \epsilon \tag{62}$$

Since this is true for all ϵ, and as $R(Q_{opt})_r$ is in $S \cap C$ because R and $(Q_{opt})_r$ are continuous, the lemma is true.

Write $X := \begin{bmatrix} W \\ 0 \end{bmatrix}$, $Y := \begin{bmatrix} UR_1 \\ R_2 \end{bmatrix}$. Then $X, Y \in H^\infty_{n \times n} \times H^\infty_{n \times n}$ are continuous in the closed unit disc by hypothesis (A2) and

$$\begin{aligned} \left\| \begin{bmatrix} U^*W \\ 0 \end{bmatrix} - R(Q_{opt})_r \right\|_B &= \|X - Y(Q_{opt})_r\|_B \\ &\leq \|(X - YQ_{opt})_r\|_B + \|X - X_r\|_B + \|Y - Y_r\|_B \|Q_{opt}\|_\infty \end{aligned} \tag{63}$$

$\|(X - YQ_{opt})_r\|_B$ is bounded above by $\|X - YQ_{opt}\|_B$ because for $K \in H^\infty_{n\times n} \times H^\infty_{n\times n}$, $\|K_r\|_B$ is an increasing function of $r \in [0,1]$. To show this suppose on the contrary that there exists $K = [K_1\ K_2] \in H^\infty_{n\times n} \times H^\infty_{n\times n}$ and z_0 in the open unit disc and an $\epsilon' > 0$ such that

$$\epsilon' + |K_1(e^{i\theta})| + |K_2(e^{i\theta})| < |K_1(z_0)| + |K_2(z_0)| \quad \forall \theta \in [0, 2\pi) \tag{64}$$

There exist constant unit vectors in $\mathbf{C}^n$, u_1, u_2, v_1, v_2 such that

$$\epsilon' + |u_1 K_1(e^{i\theta})v_1| + |u_2 K_2(e^{i\theta})v_2| < |u_1 K_1(z_0)v_1| + |u_2 K_2(z_0)v_2| \tag{65}$$

Denote the outer parts of the scalar H^∞ functions $u_1K_1(e^{i\theta})v_1$, $u_2K_2(e^{i\theta})v_2$ by h_1, h_2. Then (65) implies

$$\epsilon' + |h_1(e^{i\theta})| + |h_2(e^{i\theta})| < |h_1(z_0)| + |h_2(z_0)| \tag{66}$$

The magnitude of an outer H^∞ function is a positive bounded subharmonic function on the unit disc, therefore $|h_1(e^{i\theta})| + |h_2(e^{i\theta})|$ also falls into this category. Such functions satisfy a maximum modulus principle which contradicts (66). Therefore (64) cannot hold and the assertion is proven.

Given $\epsilon > 0$, there exists r for which the remaining two terms on the RHS of (63) are $\leq \frac{\epsilon}{2}$ each, because the continuity of X and Y implies that $\|X - X_r\|_B \to 0$ and $\|Y - Y_r\|_B \to 0$ as $r \to 1$. Therefore (61) is true.

□

Theorem 3.3. *(a)* $(B_c)^* \simeq B^\sim$

(b) $(S_c)^\perp \simeq S^\sim$

where $\simeq$ denotes the isometric isomorphism between $\nu \in B^\sim$ and functionals $< \cdot|\nu >$ (which equals $\overline{< \nu|\cdot >}$).

Proof. (a) Clearly $\overline{<\nu|\cdot>}$ defines a linear functional on B_c, which is bounded by (59). Conversely, if ϕ is a bounded linear functional on B_c, then we can use the Riesz representation theorem for ϕ acting on each component of B_c to get $\phi(.) = \overline{<\nu|\cdot>}$ for some $\nu \in BV_{n\times n} \times BV_{n\times n}$, or using the Radon-Nykodym theorem, the equivalent representation (58). Thus the set B^* is isomorphic to $B^\sim$, and $\|\phi\| \le \|\nu\|_{B^\sim}$ by (59). That upper bound will be shown to be a supremum, which will mean that $\|\phi\| = \|\nu\|_{B^\sim}$, i.e, the isomorphism is isometric and (a) is true.

The integral form (58) for ϕ gives a linear extension $\tilde{\phi}$ of ϕ whose domain is $L^\infty_{n\times n}(w_\nu) \times L^\infty_{n\times n}(w_\nu)$. ($L^\infty_{n\times n}(w_\nu)$ denotes functions essentially bounded w.r.t the w_ν measure.)

First, the existence of $\tilde{K} = (\tilde{K}_1, \tilde{K}_2) \in L^\infty_{n\times n}(w_\nu) \times L^\infty_{n\times n}(w_\nu)$ will be demonstrated with the property that $|\tilde{\phi}(\tilde{K})| = \|\nu\|_{B^\sim}\|\tilde{K}\|_B$. In the representation (58), let $G_{\nu,r}$, $r = 1,2$ have the singular value decomposition $G_{\nu,r}(e^{i\theta}) = U_r(e^{i\theta})D_r(e^{i\theta})V_r(e^{i\theta})$, where U_r, V_r are unitary and D_r is diagonal. On the set
$\Omega := \{\theta \in [0, 2\pi) : \ |STrG_{\nu,1}(e^{i\theta})| \ge |STrG_{\nu,2}(e^{i\theta})|\}$ define $\tilde{K}_1^*(e^{i\theta}) = V_1^*(e^{i\theta})U_1^*(e^{i\theta})$ and $\tilde{K}_2^*(e^{i\theta}) = 0$. On the complement of Ω define $\tilde{K}_2^*(e^{i\theta})$ by interchanging subscripts 1,2 in the previous definition, and set $\tilde{K}_1^*(e^{i\theta}) = 0$. Then $Tr\tilde{K}_1^*G_{\nu,1} = STrG_{\nu,1}$ on Ω and $Tr\tilde{K}_2^*G_{\nu,2} = STrG_{\nu,2}$ on Ω^c, and therefore $Tr\tilde{K}_1^*G_{\nu,1}(e^{i\theta}) + Tr\tilde{K}_2^*G_{\nu,2}(e^{i\theta}) = Max(STrG_{\nu,1}(e^{i\theta}), STrG_{\nu,2}(e^{i\theta}))$ and

$$\begin{aligned} |\tilde{\phi}(\tilde{K})| &= \int_{[0,2\pi)} Max(STrG_{\nu,1}, STrG_{\nu,2})dw_\nu(\theta) \\ &= \|\nu\|_{B^\sim} = \|\nu\|_{B^\sim}\|\tilde{K}\|_B \end{aligned} \tag{67}$$

since $|\tilde{K}_1(e^{i\theta})| + |\tilde{K}_2(e^{i\theta})| = 1 = \|\tilde{K}\|_B$ by construction.

Next, given any $\epsilon > 0$, $[\tilde{K}_1, \tilde{K}_2]$ will be approximated by a pair of continuous functions $[K_1, K_2]$ with the property that $\|[K_1,\ K_2]\|_B \leq 1$ and

$$|\tilde{\phi}(\tilde{K}_1, \tilde{K}_2) - \phi(K_1, K_2)| < \epsilon \tag{68}$$

from which, together with (67), it can be concluded that $\|\phi\| \geq \|\nu\|_{B^\sim} - \epsilon$. As ϵ was arbitrary $\|\phi\| = \|\nu\|_{B^\sim}$ as claimed.

The approximation is based on Lusin's theorem, which implies that there exist *continuous* functions $K_1', K_2' \in L^\infty_{n\times n} \cap C$ and a Borel subset Θ of the unit circle such that $[K_1'(e^{i\theta}),\ K_2'(e^{i\theta})] = [\tilde{K}_1(e^{i\theta}),\ \tilde{K}_2(e^{i\theta})]\ \forall\theta \in \Theta$, and on the complement Θ^c of Θ,

$$\int_{\Theta^c} Max(STrG_{\nu,1}, STrG_{\nu,2}) dw_\nu(\theta) \leq \frac{\epsilon}{2} \tag{69}$$

where the last integral represents a weighted Borel measure of that complement. Now let $K_r := \frac{K_r'}{Max(|K_1'|+|K_2'|,1)}$, $r = 1, 2$. Then K_r is continuous, $|K_1(e^{i\theta})| + |K_2(e^{i\theta})| \leq 1$, and (68) follows from (69).

(b). By (a), $\phi \in B^*$ can be represented by $\overline{<\nu|\cdot>}$. Then

$$\begin{aligned}
\phi \in (S_c)^\perp \ &\Leftrightarrow\ \overline{\int_{[0,2\pi)} Tr(Q^* R_1^* d\nu_1(\theta)) + Tr(Q^* R_2^* d\nu_2(\theta))} = 0 \\
&\qquad \forall Q \in H^\infty_{n\times n} \cap C \\
&\Leftrightarrow\ R^* d\nu(\theta) = G d\theta \ \ for\ some\ G \in (\overline{H}^1_0)_{n\times n} \\
&\Leftrightarrow\ d\nu(\theta) = (I - RR^*) d\nu'(\theta) + RG d\theta \qquad (70) \\
&\qquad (since\ d\nu = (I - RR^*) d\nu + RR^* d\nu\ and\ RR^* R = R)
\end{aligned}$$

and (b) is proven.

□

By the Banach space duality theorem asserting that $\inf_{Q\in A_0} \|K - Q\| = \max_{\phi\in A_0^\perp,\ \|\phi\|_{A^*}\leq 1} |\phi(K)|$ for any $K \in A$ and subspace $A_0 \subset A$, it follows from Lemmas 3.1 and Theorem 3.3 that μ_0 is attained by some extremal functional, $\mu_0 = \max_{\phi\in(S_c)^\perp,\|\phi\|_{B^*}\le 1} \left|\phi\left(\begin{bmatrix} U^*W \\ 0 \end{bmatrix}\right)\right|$ and on combining this with the results of Theorem 3.1 we get

Corollary. *Under the assumptions A1 and A2,*

$$\min_{X\in S}\left\|\begin{bmatrix} U^*W \\ 0 \end{bmatrix} - X\right\|_B = \mu_0 = \max_{\phi\in(S_c)^\perp,\ \|\phi\|_{B_c^*}\le 1}\left|\phi\left(\begin{bmatrix} U^*W \\ 0 \end{bmatrix}\right)\right| \tag{71}$$

where $B_c^* \simeq B^\sim$ and $(S_c)^\perp \simeq S^\sim$.

Let $X_{opt} = RQ_{opt}$ and ϕ_{opt} be the extremal elements which exist by the Corollary, and let

$$\mu_{00} = \inf_{Q\in C_{n\times n}} \|\ |U^*W - \tilde{W}Q| + |\tilde{V}Q|\ \|_\infty \tag{72}$$

i.e., when the open unit disc analyticity constraint is removed. In the scalar case $(n=1)$, $\mu_{00} = \|\ min(|W(e^{i\theta})|,\ |V(e^{i\theta})|)\ \|_\infty$.

Theorem 3.4. *Under assumptions A1 and A2, if $\mu_0 > \mu_{00}$ then,*

i. Any optimal Q_{opt} in (53) satisfies the flatness (' allpass') condition

$$\||(U^*W - \tilde{W}Q_{opt})(e^{i\theta})| + |\tilde{V}Q_{opt}(e^{i\theta})| = \mu_0, \ \textit{Lebesgue a.e} \tag{73}$$

ii. If Q_n is any sequence in $H^\infty_{n\times n}$ such that $\lim_{n\to\infty}\||U^*W - \tilde{W}Q_n| + |\tilde{V}Q_n|\|_\infty = \mu_0$, then

$$l.i.m_{n\to\infty}\left(|(U^*W - \tilde{W}Q_n)(e^{i\theta})| + |\tilde{V}Q_n(e^{i\theta})|\right) = \mu_0. \tag{74}$$

The condition $\mu_0 > \mu_{00}$ is sharp for both conclusions in the sense that if $\mu_0 = \mu_{00}$ then there exist W, V, P for which (73) and (74) are false.

Remark. If (43) is derived from the MIMO two-disc problem

$$\mu_0 = \inf_{Q\in H^\infty_{n\times n}} \|\ |W_1(I - P_0Q)W| + |P_0QV|\ \|_\infty \tag{75}$$

in the manner elaborated at the end of Sect. 2.4, then the statements (73) and (74) can be replaced by

$$|W_1(I-P_0\tilde{Q}_{opt})W(e^{i\theta})|+|P_0\tilde{Q}_{opt}V(e^{i\theta})| \;=\; \mu_0 \quad l.a.e \qquad (76)$$

$$and\ l.i.m_{n\to\infty}\Big(|W_1(I-P_0\tilde{Q}_n)W(e^{i\theta})|+|P_0\tilde{Q}_nV(e^{i\theta})|\Big) \;=\; \mu_0 \qquad (77)$$

where $\tilde{Q}_{opt} = Q_{opt}V^{-1}$ and $\tilde{Q}_n = Q_nV^{-1}$.

Proof of Theorem 3.4. i. Let ϕ_{opt} have the integral representation $\phi_{opt} = \overline{<\underline{\nu}|\cdot>}$ where $d\underline{\nu}(\theta) = \underline{G}d\underline{w}(\theta)$ and write $X_{opt} = \underline{X}$. Let $\tilde{\phi}_{opt}$ be the extension of ϕ_{opt} to $L^\infty_{n\times n}(\underline{w}) \times L^\infty_{n\times n}(\underline{w})$ (defined by the integral (58)), which satisfies $\tilde{\phi}_{opt}(S) = 0$. Then

$$\begin{aligned}
\mu_0 &= \left\|\left(\begin{bmatrix} U^*W \\ 0 \end{bmatrix} - RQ_{opt}\right)\right\|_B = \left|\tilde{\phi}_{opt}\left(\begin{bmatrix} U^*W \\ 0 \end{bmatrix} - \underline{X}\right)\right| \\
&\le \left|\int_{[0,2\pi)}\{Tr(U^*W-\underline{X}_1)^*\underline{G}_1 + Tr\underline{X}_2^*\underline{G}_2\}d\underline{w}(\theta)\right| \\
&\le \int_{[0,2\pi)}\{|U^*W-\underline{X}_1)|STr(\underline{G}_1)+|\underline{X}_2|STr(\underline{G}_2)\}d\underline{w}(\theta) && (78)\\
&\le \int_{[0,2\pi)}\{|U^*W-\underline{X}_1|+|\underline{X}_2|\}(e^{i\theta})d\psi(\theta) && (79)\\
&\le \|\, |U^*W-\underline{X}_1|+|\underline{X}_2|\,\|_\infty \int_{[0,2\pi)} d\psi(\theta) = \mu_0 && (80)
\end{aligned}$$

where $d\psi(\theta) := Max\big(STr(\underline{G}_1), STr(\underline{G}_2)\big)d\underline{w}(\theta)$ satisfies $\int_{[0,2\pi)} d\psi(\theta) = \|\phi_{opt}\|_{B^*} = 1$. We will show that the Borel measure induced by $d\psi(\theta)$ on the unit circle is such that Borel sets of ψ measure zero have Lebesgue measure zero. The flatness condition then follows from (79-80).

The extremal functional has the equivalent representations

$$\begin{aligned}
\overline{\phi}_{opt}(K) &= \int_{[0,2\pi)} Tr(K^*\underline{G})d\underline{w}(\theta) \\
&= \int_{[0,2\pi)} Tr\big\{K^*(I-RR^*)d\nu'(\theta)\big\} \\
&\qquad + \int_{[0,2\pi)} TrK^*RG_3d\theta && (81)
\end{aligned}$$

for some $G_3 \in (\overline{H}^1_0)_{nxn}$. Suppose that $\psi(\theta)$ (defined as in (80)) has zero increase on some Borel subset Ω of the circle of nonzero Lebesgue measure. Let M be an arbitrary constant matrix. For an arbitrary Borel set E of Ω, consider the $L^\infty_{nxn}(\underline{w}) \times L^\infty_{nxn}(\underline{w})$ function, $\hat{K}^*(e^{i\theta}) := MR^*$ for $\theta \in E$, $\hat{K}^*(e^{i\theta}) := 0$ for θ in the complement of E w.r.t. the circle. Then $\tilde{\phi}_{opt}(\hat{K}) = 0 = \int_E Tr(MG_3(e^{i\theta}))d\theta$ which implies that $G_3(e^{i\theta}) = 0$ Lebesgue a.e for $\theta \in \Omega$ since M is chosen arbitrarily. But as Ω has strictly positive Lebesgue measure and $G_3 \in (\overline{H}^1_0)_{nxn}$, G_3 must be identically zero. Therefore the maximum in (71) remains μ_0 if $(S_c)^\perp((S_c)^\perp \simeq S^\sim)$ is restricted to $(\simeq S_0^\sim)$ where,

$$S_0^\sim := \{\nu \in B^\sim :\ \nu(\theta) = \int_{[0,\theta]} (I - RR^*)d\nu'(\theta'),\ \nu' \in B^\sim\} \tag{82}$$

But under the isomorphism of Theorem 3.3, $S_0^\sim \simeq (RC_{nxn})^\perp$, by the reasoning used in Theorem 3.3b. Therefore, again using the Banach space duality theorem following (70),

$$\mu_0 = \max_{\phi \in (RC_{nxn})^\perp,\ \|\phi\|_{B^*} \le 1} \left| \phi\left(\begin{bmatrix} U^*W \\ 0 \end{bmatrix} \right) \right| = \mu_{00}$$

which contradicts the hypothesis and i. is proven.

ii. Define $X^{(n)} := RQ_n := \begin{bmatrix} X_1^{(n)} \\ X_2^{(n)} \end{bmatrix}$. Then,

$$\begin{aligned} \mu_0 &= \left|\phi_{opt}\left(\begin{bmatrix} U^*W \\ 0 \end{bmatrix}\right)\right| = \left|\tilde{\phi}_{opt}\left(\begin{bmatrix} U^*W \\ 0 \end{bmatrix} - X^{(n)}\right)\right| \\ &\le \int_{[0,2\pi)} \left(|U^*W - X_1^{(n)}| \cdot STr(\underline{G}_1) + |X_2^{(n)}| \cdot STr(\underline{G}_2)\right) d\underline{w} \\ &\le \int_{[0,2\pi)} \left(|U^*W - X_1^{(n)}| + |X_2^{(n)}|\right) d\psi \\ &\le \||U^*W - X_1^{(n)}| + |X_2^{(n)}|\|_\infty (\to \mu_0\ as\ n \to \infty) \end{aligned} \tag{83}$$

Thus,

$$\lim_{n\to\infty} \int_{[0,2\pi)} \left(\mu_0 - |(U^*W - X_1^{(n)})(e^{i\theta})| - |X_2^{(n)}(e^{i\theta})|\right) d\psi = 0 \tag{84}$$

Since $\||U^*W - X_1^{(n)}| + |X_2^{(n)}|\|_\infty \to \mu_0$ as $n \to \infty$, (84) holds with the integrand replaced by its absolute value. Therefore, $R_n(e^{i\theta}) := |(U^*W - \tilde{W}Q_n)(e^{i\theta})| + |\tilde{V}Q_n(e^{i\theta})| \to \mu_0$ as $n \to \infty$, *in* $L_1(d\psi)$. It follows that R_n converges *in* ψ *measure* to the constant μ_0. In the proof of i. it was established that Lebesgue measure is absolutely continuous w.r.t. ψ when $\mu_0 > \mu_{00}$. Hence for arbitrary $\epsilon > 0$

$$\begin{aligned} &\lim_{n\to\infty} \psi\{\theta \in [0, 2\pi) \; : \; |R_n(e^{i\theta}) - \mu_0| > \epsilon\} = 0 \\ \Rightarrow &\lim_{n\to\infty} m\{\theta \in [0, 2\pi) \; : \; |R_n(e^{i\theta}) - \mu_0| > \epsilon\} = 0 \end{aligned} \tag{85}$$

Therefore R_n converges to μ_0 *in Lebesgue measure*. Since $R_n \in L^\infty[0, 2\pi)$ we have $R_n \to \mu_0$ in $L^1[0, 2\pi)$.

If the hypothesis $\mu_0 > \mu_{00}$ is violated then the optimum is not necessarily flat, e.x., chose $W = \frac{1}{2}(1 - z)$, $V = U = 1$, in which case possibilities for Q_{opt} include $Q_{opt} = 1$ and $Q_{opt} = 0$. In the case of $Q_{opt} = 0$ the conclusions of Theorem 3.4 fail.

□

Remark. (On the uniqueness of the optimal Q.) For the SISO case, if $\mu_0 > \mu_{00}$ and W is not constant, then it can be deduced from the proof of Theorem 3.4 that the optimal Q for (73) is unique Lebesgue a.e. To show this, notice that (78) implies that

$$\begin{aligned} &arg(U^* - R_1 Q_{opt}) = arg(\underline{G}_1) \quad and \quad arg(R_2 Q_{opt}) = arg(\underline{G}_2) \\ &a.e.(\underline{w}) \;\; on\ the\ set \;\; \{\theta : \; |\underline{G}_1(e^{i\theta})| \neq 0 \;\; and \;\; |\underline{G}_2(e^{i\theta})| \neq 0\}. \end{aligned} \tag{86}$$

Define the Borel set $F := \{\theta \; : \underline{G}_1(e^{i\theta}) = \underline{G}_2(e^{i\theta}) = 0\}$. From the construction of ψ we have $\psi(F) = 0$ which implies $m(F) = 0$. Borel sets of $\underline{w}$ measure zero have ψ measure zero and so have Lebesgue measure zero. Hence (79) implies that $|\underline{G}_1| = |\underline{G}_2|$ Lebesgue a.e. if Q_{opt} is non-zero Lebesgue a.e. The last statement is true because if Q_{opt} were zero on a set of non-zero Lebesgue measure, it would be zero Lebesgue a.e., which would give a non-flat solution for non-constant W. Since $|\underline{G}_1|$ and $|\underline{G}_2|$ are equal Lebesgue a.e. and both zero only on a set of Lebesgue measure zero, (86) must hold Lebesgue a.e. This demonstrates uniqueness of Q_{opt} Lebesgue a.e. because (86) determines Q_{opt} uniquely at each $e^{i\theta}$ where it holds: the second equality in (86) determines $arg(Q_{opt})$, and then the first equality determines $|Q_{opt}|$. $\mu_0 > \mu_{00}$ is also a sharp condition for uniqueness of the optimal Q in the sense of Theorem 3.4, by the example in the proof at the end that theorem.

The statement of Theorem 3.4i could possibly be derived using the very different approach of Helton's general flatness theory [27]. However the duality-based approach taken here is quite different.

3.4 Strictly Proper Nominal Plants.

Here we shall restrict our attention to the SISO case, partly in order to simplify the presentation and partly because there there is no conceptual difference between the treatment of the SISO and MIMO cases. Recall that in the SISO case, the two-disc problems arising from the ORDAP (5), and from the two-disturbance problem (2), assumed the form

$$\inf_{Q\in H^\infty} \| |W(1-P_0Q)| + |VP_0Q| \|_\infty = \mu_0 \tag{87}$$

Assumption (A1) excludes the case of strictly proper nominal plants P_0 from the analysis of Sects. 3.1-3.3. In this section, we consider the following modification of assumption (A1) to allow strict propriety.

(A1′) *The outer factor of P_0 takes the form HY, where H is an invertible function in H^∞ and Y is a strictly proper rational outer function whose zeros on the unit circle* *are constrained to lie at the point $\{-1\}$. In addition, there exists an $\epsilon_0 > 0$ such that the weights W and V of (87) satisfy $|W(e^{i\theta})|^2 + |V(e^{i\theta})|^2 > \epsilon_0 \quad \forall \theta \in [0, 2\pi)$.*

Assumption (A1′) can be further extended to allow any finite number of zeros on the unit circle and all the following theorems and proofs can be modified accordingly. However, the simpler case is considered here for clarity.

In the following lemma we note that it is possible, under a mild condition, to 'absorb' the outer factor HY into the free parameter Q in (87)

Lemma 3.2 *If $\tilde{\mu}_0 = \inf_{Q\in H^\infty} \||W(1-P_0Q)| + |VP_0Q|\|_\infty$, $\mu_0 = \inf_{Q\in H^\infty} \||W(1-BQ)| + |VQ|\|_\infty$ then under assumptions (A1′) and (A2),*

$$\mu_0 \geq |W(e^{i\pi})| \Leftrightarrow \mu_0 = \tilde{\mu}_0$$

Proof of Lemma 3.2 ($\Leftarrow$). Suppose $\{Q_n\}$ is a sequence in H^∞ such that

$$\||W(1-P_0Q_n)| + |VP_0Q_n|\|_\infty \to \tilde{\mu}_0 \tag{88}$$

Given $\epsilon > 0$ there exists an integer M such that for any $n > M$ there is a sequence of neighborhoods $N_n \subset [0, 2\pi)$ containing π for which

$$|W(1-P_0Q_n)(e^{i\theta})| + |VP_0Q_n(e^{i\theta})| > |W(e^{i\pi})| - \epsilon \quad \forall \theta \in N_n \tag{89}$$

Thus $\tilde{\mu}_0 \geq |W(e^{i\pi})|$, since ϵ is arbitrary. Hence $\tilde{\mu}_0 = \mu_0 \Rightarrow \mu_0 \geq |W(e^{i\pi})|$.

($\Rightarrow$). Fix $\epsilon > 0$. First we establish that there exists an integer n such that $\tilde{Q} := (\frac{1+z}{2})^{\frac{1}{n}} Q_0$ where Q_0 is the optimal solution to $\inf_{Q\in H^\infty} \||W(B^* - Q)| + |VQ|\|_\infty$, satisfies

$$\||W(1-B\tilde{Q})| + |VB\tilde{Q}|\|_\infty \leq \mu_0 + \epsilon. \tag{90}$$

Since $\left(\frac{1+e^{i\theta}}{2}\right)^{\frac{1}{n}} Q_0(e^{i\theta}) \to Q_0(e^{i\theta})$ on compact subsets of $[0,2\pi)-\{\pi\}$ we have

$$\begin{aligned} X_n(e^{i\theta}) &:= \left|W\left(1-BQ_0(e^{i\theta})\left(\frac{1+e^{i\theta}}{2}\right)^{\frac{1}{n}}\right)\right| + \left|VQ_0(e^{i\theta})\left(\frac{1+e^{i\theta}}{2}\right)^{\frac{1}{n}}\right| \\ &\to |W(1-BQ_0)(e^{i\theta})| + |VQ_0(e^{i\theta})| \textit{ on compact subsets of } [0,2\pi)-\{\pi\} \end{aligned}$$

Thus to prove (90) it is sufficient to show that $X_n(e^{i\theta})$ is uniformly bounded above by $\mu_0+\epsilon$ for sufficiently large n in some neighborhood containing π. Since WB, W, V are continuous at -1 there exists a neighborhood N containing π such that $\theta \in N$ implies

$$|W(e^{i\theta})-W(-1)| + \left(|V(e^{i\theta})-V(-1)| + |WB(e^{i\theta})-WB(-1)|\right)\|Q_0\|_\infty \le \frac{1}{3}\epsilon \quad (91)$$

Thus $|W(-1)-WB(-1)Q_0(e^{i\theta})| + |V(-1)Q_0(e^{i\theta})| < \frac{\epsilon}{3}+\mu_0$ for $\theta \in N$.
Define the following convex function

$$\phi : \mathbb{C} \to \mathbb{C}, \quad \phi(z) = |W(-1)-WB(-1)z| + |V(-1)z|$$

If $\omega \in \{z \in \mathbb{C} : \phi(z) \le |W(-1)|\}$ then $r \in [0,1] \Rightarrow r\omega \in \{z \in \mathbb{C} : \phi(z) \le |W(-1)|\}$ since $r\omega$ is a convex combination of 0 and ω. Otherwise $\phi(\omega) > |W(-1)|$, implying both 0 and ω lie in the set $\{z \in \mathbb{C} : \phi(z) \le \phi(\omega)\}$, from which we conclude that convex combinations of 0 and ω must lie in the set i.e., $\phi(r\omega) \le \phi(\omega)$ for $r \in [0,1]$. Thus, multiplication of $Q_0(e^{i\theta})$ by $r \in [0,1]$ retains $\mu_0 + \frac{\epsilon}{3}$ as an upper bound for $|W(-1)-WB(-1)Q_0(e^{i\theta})r| + |V(-1)Q_0(e^{i\theta})r|$ for $\theta \in N$. Since $arg\left(\frac{1+e^{i\theta}}{2}\right)^{\frac{1}{n}} \to 0$ uniformly for all $\theta \in [0,2\pi)$, there exists an $r_\theta \in [0,1]$ for each θ such that $\left|\left(\frac{1+e^{i\theta}}{2}\right)^{\frac{1}{n}} - r_\theta\right| < \frac{\epsilon}{6Max(\|WQ_0\|_\infty, \|VQ_0\|_\infty)}$ uniformly in θ for n sufficiently large. It follows that for $\theta \in N$, $\mu_0 + \frac{2}{3}\epsilon$ is a uniform upper bound for $|W(-1) - WB(-1)Q_0(e^{i\theta})\left(\frac{1+e^{i\theta}}{2}\right)^{\frac{1}{n}}| + |V(-1)Q_0(e^{i\theta})\left(\frac{1+e^{i\theta}}{2}\right)^{\frac{1}{n}}|$ for large enough n. From (91) we arrive at the desired uniform bound for $X_n(e^{i\theta})$, and (90) must hold for $\hat{Q}$ of the stated form.

Now define $\hat{Q}_n := \tilde{Q}\frac{\psi_n}{YH}$ where $\psi_n(z) := [\frac{n(1+z)}{(1+n)+(n-1)z}]^k$, k is the order of the zero of Y at $e^{i\pi}$. Using the same argument as that used in [58] we can show that

$$limsup_{n\rightarrow\infty}|||W(1-P_0\hat{Q}_n)|+|VP_0\hat{Q}_n|||_\infty \leq \mu_0+\epsilon \tag{92}$$

Since ϵ is arbitrary we have established that $\tilde{\mu}_0 \leq \mu_0$. Inequality in the other direction follows from the fact that $HYBH^\infty \subset BH^\infty$.

□

Remark. The assumption that $\mu_0 \geq |W(e^{i\pi})|$ will certainly be satisfied when high frequency plant uncertainty is more pronounced than high frequency disturbances (i.e. $|W(e^{i\pi})| \leq |V(e^{i\pi})|$). This follows because if $|V(e^{i\pi})| \geq |W(e^{i\pi})|$, then $\mu_0 \geq \|Min(|W|,|V|)\|_\infty \geq |W(e^{i\pi})|$.

For an important class of circumstances, strict propriety of the nominal plant P_0 rules out the existence of an optimal control law for both the ORDAP and the two-disc problem (87), as shown in Theorem 3.5 below. This contrasts with the case of nonstrictly proper plants (A1).

Theorem 3.5 *Under the assumptions of Theorem 3.4 with (A1′) replacing (A1), if $\mu_0 > |W(e^{i\pi})|$ then there does not exist an optimal Q for the optimization*

$$\inf_{Q\in H^\infty}|||W(1-P_0Q)|+|VP_0Q|||_\infty := \tilde{\mu}_0$$

Proof. Suppose on the contrary that there exists a $Q_{opt} \in H^\infty$ for which

$$|||W(1-BHYQ_{opt})|+|VBHYQ_{opt}|||_\infty = \tilde{\mu}_0 = \mu_0 \tag{93}$$

and $\mu_0 > |W(e^{i\pi})|$. Then for arbitrary $\epsilon > 0$ there exists a neighborhood N of π such that for $\theta \in N$

$$|W(1 - BHYQ_{opt})(e^{i\theta})| + |VBHYQ_{opt}(e^{i\theta})| \leq |W(e^{i\pi})| + \epsilon \tag{94}$$

But Theorem 3.4 part (b) i.e., the flatness property of the optimum, and the continuity of WB and V, $\mu_0 \leq |W(e^{i\pi})| + \epsilon$, and since ϵ is arbitrary, $\mu_0 \leq |W(e^{i\pi})|$ contradicting the assumption.

□

When an optimal control law does not exist, obviously one cannot draw any conclusions about *the* optimal behavior along the lines of Theorems 3.2 and 3.4. However, in these situations it is meaningful to analyze the properties of 'nearly' optimal solutions, where the performance is very close to the optimum. Accordingly, in Theorem 3.4ii it was shown that the optimal flatness result of Theorem 3.4i. holds in a limiting sense for 'nearly' optimal control laws, enabling us to deduce an 'approximate flatness' result for strictly proper nominal plants satisfying $\mu_0 \geq |W(e^{i\pi})|$.

In summary then, if the high frequency uncertainty is more pronounced than the high frequency disturbances (see above Remark), then the value of the optimal robust disturbance attenuation for a strictly proper nominal plant is the same as that of a non-strictly proper counterpart with the same inner factor. However, strict propriety rules out the existence of an optimal control law and its attendant properties.

3.5 Qualitative Implications for Feedback

In the following subsections the duality theory as discussed in Sects. 3.1-3.3 is used to examine the ORDAP for SISO plants, with particular emphasis on the qualitative aspects. The application of this theory to the synthesis of 'nearly' optimal control laws is deferred to Chapter 4.

3.5.1 Properties of the Optimal Behavior of the ORDAP

As shown in Chapter 2, the ORDAP for SISO systems reduces to a fixed-point problem involving a two-disc optimization of the form (38) (also see [20]). Here, the theory of Sects. 3.1-3.4 for these two-disc optimizations (43) is applied to gain insight into the ORDAP for the SISO case. It should be noted that in order to apply the ideas of Sect 3.1-3.3 to situations where V is scaled by a parameter r, assumption A2 must be strengthened to:

(A3) *B^*W, W, V, BW are continuous on T, as is the outer spectral factor of $|W|^2 + r^2|V|^2$ for any real r, and the outer factor of the plant P_0 is invertible in H^∞.*

Recalling the representation of the ORDAP in the MIMO case (16), the representation of the ORDAP in the SISO case assumes the form

$$\mu_{opt} := \inf_{\substack{Q \in H^\infty \\ \|VP_0Q\|_\infty \leq 1}} \sup_{X \in H^\infty, \|X\|_\infty < 1} \left\| \frac{W(1 - P_0Q)}{1 + XVP_0Q} \right\|_\infty \tag{95}$$

In the following theorem, optimal feedback laws for (95) are shown to exist, uniqueness of the optimal feedback is proven, and an expression is obtained for the magnitude of the sensitivity function under optimality.

Theorem 3.6 *a) If the outer factor of P_0 is invertible in H^∞, $|W(e^{i\theta})|^2 + |V(e^{i\theta})|^2$ is uniformly bounded below, and $\mu_{opt} > 0$, there exists at least one control law stabilizing every system in $B(P_0, V)$ such that*

$$\sup_{X \in H^\infty, \|X\|_\infty < 1} \left\| \frac{W(1 - P_0Q)}{1 + XVP_0Q} \right\|_\infty = \mu_{opt} \tag{96}$$

b) If, in addition, assumption A3 holds and $\mu_{opt} > \|Min(|W(e^{i\theta})|, \mu_{opt}|V(e^{i\theta})|)\|_\infty$, then there exists one, and only one controller C_0 which stabilizes every system in $B(P_0, V)$ and achieves equality in (96). For this optimal case, the magnitude of the *sensitivity function for the nominal plant P_0 satisfies*

$$\left|\frac{W}{1+P_0C_0}(e^{i\theta})\right| \;=\; \mu_{opt} - \mu_{opt}\left|\frac{VP_0C_0}{1+P_0C_0}(e^{i\theta})\right| \quad \textit{Lebesgue a.e} \tag{97}$$

Remark. The last equality (97) points to a trade-off in the optimal case, between the performance of the nominal system as represented by the sensitivity, and the stability margin as represented by the reciprocal of $|VP_0Q(e^{i\theta})|$ at each frequency θ. Note that the reciprocal of $|VP_0Q(e^{i\theta})|$ at each θ represents the largest allowable weighted multiplicative plant perturbation at frequency θ for which closed-loop stability can be guaranteed. For the nominal plant under optimal feedback, (97) suggests that at those frequencies where nominal sensitivity is relatively small, $|VP_0Q(e^{i\theta})|$ is close to unity, and the stability margin is close to the minimum specified in the problem formulation.

Proof of Theorem 3.6 a) Under assumptions stated in a), it follows from Theorem 3.2 that for each r there exists an optimal $Q \in H^\infty$ for optimizations represented by $\chi(r)$ in (38) for the SISO case. From Theorem 2.1 μ_{opt} is the smallest fixed-point of $\chi(\cdot)$. Hence if Q_{opt} is the minimal $Q \in H^\infty$ for the optimization $\chi(\mu_{opt})$ then,

$$\|\; |W(1-P_0Q_{opt})| + \mu_{opt}|VP_0Q_{opt}| \;\|_\infty = \chi(\mu_{opt}) = \mu_{opt} \tag{98}$$

$$\Rightarrow \|VP_0Q_{opt}\|_\infty \leq 1 \;\; (\textit{since } \mu_{opt} > 0) \quad \textit{and} \tag{99}$$

$$|W(1-P_0Q_{opt})(e^{i\theta})| \leq \mu_{opt}\left|1 + (P-P_0)Q_{opt}(e^{i\theta})\right| \quad \forall P \in B(P_0,V)\; \forall\theta \in [0,2\pi)$$

$$\Rightarrow \left\|\frac{W(1-P_0Q_{opt})}{1+(P-P_0)Q_{opt}}\right\|_\infty \leq \mu_{opt} \qquad \forall P \in B(P_0,V)$$

The above fractional term is in H^∞ since $\|(P-P_0)Q_{opt}\|_\infty \leq \|XVP_0Q_{opt}\|_\infty < 1$ (since $\|X\|_\infty < 1$ for all admissible P). An optimal controller is given by $C_{opt} = \frac{Q_{opt}}{1-Q_{opt}P_0}$.

b) Rewriting (98)

$$\chi(\mu_{opt}) = \mu_{opt} = \min_{Q\in H^\infty} \| |W(1-P_0Q)| + \mu_{opt}|VP_0Q| \|_\infty \tag{100}$$

$$= \| |W(1-P_0Q_{opt})| + \mu_{opt}|VP_0Q_{opt}| \|_\infty \tag{101}$$

Theorem 3.4 can now be applied to the optimization (100). Assumption A3 guarantees that A2 holds for the weights W and rV, with the result that all the conditions for the application Theorem 3.4 to this optimization are met. The hypothesis $\mu_0 > \mu_{00}$ in the statement of this theorem, for the case (100) becomes $\mu_{opt} > \|Min\left(|W(e^{i\theta})|, \mu_{opt}|V(e^{i\theta})|\right)\|_\infty$. Thus, the Q_{opt} which satisfies (101) is unique, and

$$|W(1-P_0Q_{opt})(e^{i\theta})| + \mu_{opt}|VP_0Q_{opt}(e^{i\theta})| = \mu_{opt} \qquad l.a.e. \tag{102}$$

proving both statements of b).

□

3.5.2 Growth of Optimal Robust Disturbance Attenuation with Plant Uncertainty

The flatness condition (73) of Theorem 3.4 has significant implications for the growth of the disturbance transmission with plant uncertainty. If the size of the set of open-loop plant uncertainty is strictly increased, i.e. V is replaced by V' where $|V'(e^{i\theta})| \geq |V(e^{i\theta})|$ $l.a.e$ and $|V'(e^{i\theta})| > |V(e^{i\theta})|$ on a set of nonzero measure, then robust performance for the ORDAP is strictly increased from μ to μ', when the assumptions of Theorem 3.4 apply for W, V and P_0, and $|W(e^{i\theta})|$ is not identically constant. To prove this, suppose on the contrary that $\mu = \mu'$ for the situation described in the previous sentence. By assumption, we have $\mu > \|Min\,(|W|, \mu|V|)\|_\infty$. Thus there exists $Q_0 \in H^\infty$ (see Theorem 3.6) such that

$$|W(1-BQ_0)| + \mu|VQ_0| = \mu \qquad l.a.e. \tag{103}$$

Since $\mu = \mu'$ we also have for some $Q_1 \in H^\infty$ (from Theorem 2.1 and Theorem 3.6a)

$$|W(1 - BQ_1)| + \mu|V'Q_1| \leq \mu \qquad l.a.e. \tag{104}$$

The uniqueness of optimal solutions of Theorem 3.4i. implies that (103) and (104) can only both be true if $Q_0 = Q_1$ $l.a.e.$. Subtracting (103) from (104) then gives

$$\mu\left(|V'| - |V|\right)|Q_0| = 0 \quad l.a.e. \tag{105}$$

which in turn implies that Q_0 must be zero on a set of positive Lebesgue measure, and so must be identically zero since it is an element of H^∞. In the light of (103) this would imply that $|W| = \mu_{opt}$ $l.a.e.$, which is ruled out by the assumption that $|W|$ is not identically constant. QED. A conceptually identical argument yields the same conclusion for the case where the weighting W is replaced by a larger weighting which is strictly larger on a set of strictly positive Lebesgue measure, under the assumption that $|V|$ is not identically constant.

When the conclusion of Lemma 3.2 holds, the same strict monotonicity property applies for strictly proper nominal plants (i.e. when (A1) is replaced by (A1')).

The above conclusions on the ORDAP will now be applied specifically to the SISO case of the plant disturbance attenuation problem of Section 2.2, in which W of (95) takes the form $W = W_1 V$ for W_1 and V as in (24). The result is that under the assumptions of Theorem 3.6b, the closed-loop uncertainty set completely 'fills out' a W_1^{-1} weighted sphere in H^∞ of radius μ_{opt}. In this context, the term 'fills out' refers to the situation where the a posteriori plant uncertainty set cannot be contained in a smaller weighted sphere in H^∞. This result is proven by showing that its negation would contradict the above strict monotonicity property. Accordingly, suppose there exists a smaller sphere of containment. That is, there exists a weighting W_1' such that

$|W_1'| \geq |W_1|$ *l.a.e.*, which is *strictly* larger on a set of positive Lebesgue measure, and for which the optimal robust disturbance attenuation (95) remains unchanged when $W = W_1 V$ is replaced by $W_1' V$. However, this possibility is ruled out by the strict monotonicity of the optimal robust disturbance w.r.t W mentioned above. The conclusion is, therefore, that a feedback which optimally contracts H^∞ unstructured plant uncertainty uses all the 'space' in the sphere of optimal closed-loop radius. In practical terms, this means that if certain frequencies are deemphasized, i.e. lightly weighted in the ORDAP, the resulting optimal closed-loop uncertainty set will have proportionately greater radius in those frequency ranges.

3.5.3 Well-Posedness of the ORDAP and Uniqueness of the Fixed-Point of χ

In this subsection we investigate the well-posedness of the ORDAP w.r.t problem data. In this context, well-posedness refers to the property that optimal performance depends continuously on the problem data [46]. If the ORDAP is to provide a basis for robust control synthesis, at the very least the optimal performance must depend continuously on the problem data, otherwise the slightest inaccuracies in a priori information could give rise to large errors. Moreover, from a more abstract point of view, any ill-posedness would suggest that some latent physical constraint in the problem had been neglected [46].

Smith's paper [46] motivated the analysis of this subsection with the observation that well-posedness and robustness were distinct concepts, in the sense that robust synthesis problems such as ORDAP could still be ill-posed in the problem data. This point was illustrated in [46] with an example (Example 4) in which the optimal performance in the ORDAP was a discontinuous function of the open-loop uncertainty radius. This suggested that the ORDAP could be ill-posed w.r.t. this quantity, at least for some cases. Accordingly, in this subsection we use the theory of Sects. 3.1-3.4 to examine the dependence of optimal robust disturbance attenuation on the radius of the a priori uncertainty. We show that Example 4 of [46] is really a special case, and that for quite general situations the ORDAP is well-posed for perturbations to the radius of a priori uncertainty.

In contrast to the behavior exhibited in Example 4 of [46], Smith showed that the ORDAP was well-posed w.r.t. perturbations to the nominal plant P_0 for additive descriptions of plant uncertainty. Hence, in those situations where the ORDAP is well-posed w.r.t uncertainty radius, we can conclude well-posedness w.r.t. the complete additive uncertainty description.

Note that well-posedness of the ORDAP with respect to the weighting function $W \in H^\infty$ follows directly from the representation (95).

Henceforth, we define the function $\chi_0 : [0,\infty) \to [0,\infty)$

$$\chi_0(r) := \|Min\,(|W|, r|V|)\,\|_\infty \tag{106}$$

The following theorem identifies the conditions under which the optimal robust disturbance attenuation depends continuously on the uncertainty radius. We normalize W such that $\|W\|_\infty = 1$.

Theorem 3.7 *For SISO systems where the nominal plant is not purely outer, where its outer factor is invertible in H^∞, and where $|W(e^{i\theta})|^2 + |V(e^{i\theta})|^2$ is assumed to be uniformly bounded below by a strictly positive quantity, the following is true:*

1) a) and b) are equivalent conditions,

a) There is no subinterval $\mathcal{I}$ of [0,1] with non-zero length such that

$$\chi(r) = \chi_0(r) = r \qquad \textit{for all } r \in \mathcal{I} \tag{107}$$

b) There exists one and one only positive fixed-point of $\chi(\cdot)$.

2) The optimal robust disturbance attenuation

$$\inf_{\substack{Q \in H^\infty \\ \|rVP_0Q\|_\infty \leq 1}} \quad \sup_{X \in H^\infty_{n\times n}} \left\| \frac{W(1-P_0Q)}{1-rXVP_0Q} \right\|_\infty \tag{108}$$

is a continuous function of r *at* $r = x$ *if and only if the conditions of 1) are satisfied for the nominal plant* P_0 *and the weights* W, xV.

Proof.

1) Suppose $\chi(\cdot)$ has more than one fixed-point. Then there exists $r_1, r_2 \in (0,1]$ such that $r_1 < r_2$ and

$$\chi(r_1) = r_1 \qquad\qquad \chi(r_2) = r_2 \tag{109}$$

Under this condition we shall establish the following two claims, which will prove the implication a) $\Rightarrow$ b).

Claim 1: $\chi(\rho) = \rho$ for all $\rho \in [r_1, r_2]$

Claim 2: $\chi(\rho) = \chi_0(\rho)$ for all $\rho \in (r_1, r_2]$

To prove claim 1 we note that $\chi(r_1) = r_1$, which implies (Theorem 3.2) that there exists a $Q_1 \in H^\infty$ such that

$$|W(1-BQ_1)(e^{i\theta})| + r_1|VQ_1(e^{i\theta})| \leq r_1 \qquad \textit{l.a.e.} \tag{110}$$

From (110) we have that $\|VQ_1\|_\infty \leq 1$ (since $r_1 > 0$ because the nominal plant is not purely outer). Hence it follows that

$$|W(1-BQ_1)(e^{i\theta})| + \rho|VQ_1(e^{i\theta})| \leq \rho \qquad \textit{for l.a.e } \theta \textit{ and for all } \rho \in [r_1, r_2] \tag{111}$$

Thus $\chi(\rho) \leq \rho$ for $\rho \in [r_1, r_2]$. To prove that $\chi(\rho) = \rho \;\; \forall \rho \in [r_1, r_2]$ suppose on the contrary $\chi(\rho) < \rho$ for some $\rho \in (r_1, r_2)$. Then there exists a $\tilde{Q} \in H^\infty$ such that

$$|W(1-B\tilde{Q})(e^{i\theta})| + \rho|V\tilde{Q}(e^{i\theta})| \;<\; \rho - \delta \;\; \textit{l.a.e} \quad \textit{for some } \delta > 0 \tag{112}$$

$$\Rightarrow |W(1-B\tilde{Q})(e^{i\theta})| + r_2|V\tilde{Q}(e^{i\theta})| \;<\; r_2 - \delta \;\; l.a.e. \tag{113}$$

However (113) is impossible since $\chi(r_2) = r_2$. Hence the assertion $\chi(\rho) < \rho$ is false, and we must have $\chi(\rho) = \rho \;\; \forall \rho \in [r_1, r_2]$, proving claim 1.

To prove claim 2, we suppose that $\chi(\rho) > \chi_0(\rho)$ for some $\rho \in (r_1, r_2)$ and show that is leads to a contradiction. This assertion implies, invoking Theorem 3.6 and using the fact that $\chi(\rho) = \rho$ from claim 1, that there exists a $Q_3 \in H^\infty$ such that

$$|W(1-BQ_3)(e^{i\theta})| + \rho|VQ_3(e^{i\theta})| = \rho \qquad l.a.e. \tag{114}$$

and that the set

$$\mathcal{A}_\rho := \left\{ Q \in H^\infty \;:\; |W(1-BQ)(e^{i\theta})| + \rho|VQ(e^{i\theta})| \le \rho \qquad l.a.e. \right\} \tag{115}$$

is the singleton $\{Q_3\}$. From (111), Q_1 must also lie in $\mathcal{A}_\rho$, thus we conclude that

$$Q_1(e^{i\theta}) = Q_3(e^{i\theta}) \;\; l.a.e. \tag{116}$$

Subtracting (110) from (114) in the light of (116) implies that $(\rho - r_1)|VQ_3(e^{i\theta})| \ge \rho - r_1 \;\; l.a.e.$ which in turn implies that $|VQ_3(e^{i\theta})| \ge 1 \;\; l.a.e.$. From (114) this gives, $|W(1-BQ)(e^{i\theta})| = 0 \;\; l.a.e.$ This last conclusion is ruled out by the assumption that the nominal plant is not purely outer. Thus the initial assertion that $\chi(\rho) > \chi_0(\rho)$ is false, claim 2 must be valid, and a) $\Rightarrow$ b).

The implication b) $\Rightarrow$ a) is established by showing the that negation of a) $\Rightarrow$ negation of b), which follows from the definition of fixed-point.

2) From Theorem 2.1 and 3.2 the quantity in (108) is the smallest fixed-point of

$$\chi_x(r) := \inf_{Q \in H^\infty} \| |W(1-P_0Q)| + r|xVP_0Q| \|_\infty \quad \textit{as a function of } r. \tag{117}$$

If this fixed-point is unique for a particular x, then its position must be continuous w.r.t perturbations to x, since $\chi_x(\cdot)$ is a continuous, non-decreasing function. This follows from the fact that if a real continuous function taking positive and negative values on a compact interval of $\mathbb{R}$ possesses a unique root, then the location of the smallest root is continuous w.r.t. perturbations to the function in the uniform metric. If, on the other hand, the positive fixed-point of χ_x is not unique then, following the proof of 1), positive fixed-points must constitute an interval $[r_1, r_2]$. If $\delta > 0$ then

$$\chi_{x+\delta}(r) = \inf_{Q \in H^\infty} \| |W(1 - BQ)| + x\frac{x+\delta}{x} r |VQ| \|_\infty \tag{118}$$

$$= \frac{x+\delta}{x} r \quad for \ \frac{x+\delta}{x} r \in [r_1, r_2] \tag{119}$$

It follows that fixed-points of $\chi_{x+\delta}$ are excluded from $[\frac{x}{x+\delta} r_1, \frac{x}{x+\delta} r_2]$. In addition, fixed-points of $\chi_{x+\delta}$ must be bounded below by r_1, or r_1 could not be the smallest fixed-point of χ_x. Since δ is arbitrary and $r_2 > r_1$, the smallest fixed-point of $\chi(\cdot)$ is not continuous at x.

$\square$

Statement 1a) of Theorem 3.7 is actually very weak, and almost always holds in non-pathological situations. For example, in *any* of the following cases 1a) must be a true statement:

i. $W, V \in C^1[0, 2\pi)$, and $|V(e^{i\theta})| = 1$ for at most one single $\theta \in [0, \pi)$ at which $|V(e^{i\theta})|$ does not achieve a maximum.

ii. $\|V\|_\infty < 1$

iii. $\|WV\|_\infty$ is bounded above by the optimal robust disturbance attenuation.

Note that none of these conditions apply in Example 4 of [46].

3.6 ORDAP for MIMO Systems: An Exact Characterization.

In this section, we will deal directly with χ, in order to obtain an exact description of the ORDAP for MIMO systems in a predual space. The development will be similar to that of Sect. 3.1 with a modified matrix norm, and will enable the implementation of the same convex optimization based numerical techniques as in Chapter 4, for MIMO systems.

Define $\hat{B}$ to be the Banach space $L^{\infty}_{n\times n} \times L^{\infty}_{n\times n}$ consisting of pairs of $n \times n$ matrix-valued functions defined on the unit circle under the norm,

$$\|K\|_{\hat{B}} := ess \sup_{\theta\in[0,2\pi)} \sup_{\zeta\in\mathbb{C}^n, |\zeta|\leq 1} \left(|K_1(e^{i\theta})\zeta| + |K_2(e^{i\theta})\zeta|\right) \tag{120}$$
$$K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix}$$

From Theorem 2.1, the optimal robust disturbance attenuation problem for the MIMO case reduces to an optimization of the form

$$\mu_0 = \inf_{Q\in H^{\infty}_{n\times n}} \left\| \begin{bmatrix} \bar{U}^* W \\ 0 \end{bmatrix} - RQ \right\|_{\hat{B}} \tag{121}$$

The rest of this section is devoted to obtaining a description of the optimization (121) in the predual space.

We define X_n to be the Euclidean space $\mathbb{C}^n$ with the Euclidean inner product, and Y_n to be the Banach space $\mathbb{C}^n \times \mathbb{C}^n$ under the norm defined by

$$\left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_{Y_n} := Max\,(|x|, |y|) \tag{122}$$

From Lemma 6.1 we have that

$$X_n^* \simeq X_n, \quad X_n^{**} \simeq X_n \tag{123}$$

$$Y_n^* \simeq Z_n, \quad Y_n^{**} \simeq Y_n \tag{124}$$

where Z_n is the Banach space $\mathbb{C}^n \times \mathbb{C}^n$ under a norm defined by

$$\left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_Z := |x| + |y| \tag{125}$$

The Banach space of bounded linear operators from $X_n^* \to Y_n^*$ under the induced operator norm (denoted by $B(X_n^*, Y_n^*)$) is isometrically isomorphic to the Banach space of matrix pairs in $\mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$ under the norm defined by

$$\left\| \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \right\|_\sim := \sup_{\zeta \in \mathbb{C}^n, |\zeta| \le 1} (|A_1\zeta| + |A_2\zeta|) \tag{126}$$

The following definition of the *nuclear* norm for pairs of matrices is a special case of the more general definition given in Chapter 6, Definition 6.1.

Definition 3.1 *If $A_1, A_2 \in \mathbb{C}^{n\times n}$ then the nuclear norm $\left\| \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} \right\|_{nuc}$ is defined to be the infimum of all sums $\sum_k \|x_k\|_{X_n} \cdot \|y_k\|_{Y_n}$, $x_k \in \mathbb{C}^n$, $y_k \in Y_n$ such that*

$$\begin{bmatrix} A_1 \\ A_2 \end{bmatrix} u = \sum_k (x_k^T u) y_k \quad \forall u \in \mathbb{C}^n \tag{127}$$

Lemma 3.3 *If Υ, Φ, A, $B \in \mathbb{C}^{n\times n}$ then*

$$|Tr\Upsilon^* A + Tr\Phi^* B| \le \sup_{\zeta \in \mathbb{C}^n, |\zeta| \le 1} (|A\zeta| + |B\zeta|) \cdot \left\| \begin{bmatrix} \Upsilon \\ \Phi \end{bmatrix} \right\|_{nuc} \tag{128}$$

In addition for every A, B there is a choice of Υ, Φ which makes (128) an equality.

Proof. Consider the space of linear operators from $X_n \to Y_n$ equipped with the nuclear norm of Definition 6.1. This is a Banach space after [22] and [8], and following the notation of the latter reference it is denoted by $N(X_n, Y_n)$. As in the case of $B(X_n^*, Y_n^*)$ above, there is an isometric isomorphism between $N(X_n, Y_n)$ and a Banach space of matrices $\mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$ under the nuclear norm of definition 3.1. It follows from Theorem 2.10 of [8] that the dual of $N(X_n, Y_n)$ is identified with $B(X_n^*, Y_n^*)$.

Evaluation of linear functionals can be expressed in terms of the respective matrix representations of the argument and functional in the following manner. If ϕ is a bounded linear functional on $N(X_n, Y_n)$ then the action of ϕ on an element of $N(X_n, Y_n)$ possessing a matrix representation $\begin{bmatrix} \Upsilon \\ \Phi \end{bmatrix} \in \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$ can be expressed

$$Tr\Upsilon A^* + Tr\Phi B^*. \tag{129}$$

$\begin{bmatrix} A \\ B \end{bmatrix} \in \mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$ is the matrix representation (under the isomorphism defined prior to (126)) of the unique element in the dual space $B(X_n^*, Y_n^*)$ corresponding to the functional ϕ. The statement of the isometric isometry between the dual of $N(X_n, Y_n)$ and $B(X_n^*, Y_n^*)$, when expressed in terms of the respective matrix representations, establishes both (128) and the fact that (128) can be made arbitrarily close to equality by appropriate choice of $\Upsilon, \Phi \in \mathbb{C}^{n\times n}$. Exact equality for some choice of Υ and Φ, then follows from the compactness of closed bounded subsets of $\mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$ in the nuclear norm.

□

Now define the Banach space $\hat{B}_*$ to be $L^1_{n\times n} \times L^1_{n\times n}$ under the norm

$$\|G\|_{\hat{B}_*} := \int_0^{2\pi} \left\| \begin{bmatrix} G_1(e^{i\theta}) \\ G_2(e^{i\theta}) \end{bmatrix} \right\|_{nuc} d\theta \tag{130}$$

where the subscript $_{nuc}$ denotes the matrix nuclear norm of definition 3.1.

Remark. To establish that $\hat{B}_*$ is indeed a Banach space, the only non-trivial step is to demonstrate that it is complete as a metric space (other properties follow from the fact that $\|\cdot\|_{nuc}$ is a matrix norm). Both the norm of B_* and the norm of $\hat{B}_*$ are defined in terms of an integral over $[0, 2\pi)$ of a matrix norm on $\mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$. In the former case, that matrix norm takes the form $Max\,(Str(G_1), STr(G_2))$ while in the latter case it takes the form $\left\| \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \right\|_{nuc}$ for pairs of $\mathbb{C}^{n\times n}$ matrices G_1 and

G_2. However, since all norms are equivalent on $\mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$, we conclude that the norms of B^* and $\hat{B}_*$ must also be equivalent. Since the Banach space B_* is complete (Appendix A), we conclude that $\hat{B}_*$ must also be complete. Thus $\hat{B}_*$ is a Banach space.

Define the subspace of $\hat{S}_\perp$ of $\hat{B}_*$ as in (49). The following theorem identifies the predual of $\hat{B}_*$ and the preorthogonal complement of $\hat{S}_\perp$.

Theorem 3.8

$$i. \quad (\hat{B}_*)^* \simeq \hat{B} \tag{131}$$

$$ii. \quad (\hat{S}_*)^\perp \simeq S \ \textit{ under the isomorphism of } i. \tag{132}$$

Remark. A corollary to Theorem 3.8 establishes the existence of an optimal $Q \in H^\infty_{n\times n}$ for the optimization described by

$$\inf_{Q\in H^\infty_{n\times n}} ess \sup_{\theta\in[0,2\pi)} \sup_{\zeta\in\mathbb{C}^n, |\zeta|\leq 1} \left(|W_1(I-P_0Q)W(e^{i\theta})\zeta| + |P_0QV(e^{i\theta})\zeta|\right) \tag{133}$$

when assumption (A1) applies to the weights $\tilde{W}$ and $\tilde{V}$ derived from W, W_1, V, P_0 in (133) in the manner described in Sect. 2.4 (c.f. remark following Theorem 3.2).

Proof of Theorem 3.8. i. As in the proof of Theorem 3.1, ϕ is a bounded linear functional on $\hat{B}_*$ iff it has the representation,

$$\phi(G) = [G|K] \tag{134}$$

where $[\cdot|\cdot]$ is a bilinear form on $\hat{B}_* \times \hat{B}$ taking the form of (51), $K = \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \in \hat{B}$, $G = \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \in \hat{B}_*$. Exactly as in the proof of Theorem 3.1, we can establish that the K of (134) is uniquely specified by ϕ, and so the required isomorphism is established. Next we prove that the isomorphism is isometric. From Lemma 3.3,

$$|[G|K]| \leq \int_0^{2\pi} \sup_{\zeta\in\mathbb{C}^n, |\zeta|\leq 1} \left(|K_1(e^{i\theta})\zeta| + |K_2(e^{i\theta})\zeta|\right) \cdot \left\| \begin{bmatrix} G_1(e^{i\theta}) \\ G_2(e^{i\theta} \end{bmatrix} \right\|_{nuc} d\theta \tag{135}$$

$$\leq \ \|K\|_{\hat{B}} \cdot \|G\|_{\hat{B}_*} \tag{136}$$

Thus the bilinear form $[\cdot|K]$ defines a bounded linear functional on $\hat{B}_*$ with induced norm bounded above by $\|K\|_{\hat{B}}$. To prove that this upper bound is in fact equal to the induced operator norm, we mimic the argument used in the proof of Theorem 3.1. From the definition of the essential supremum there exists a sequence of Lebesgue sets $\{\Omega_k\}_{k=1}^{\infty}$ of strictly positive Lebesgue measure $m(\Omega_k)$ such that for $\theta \in \Omega_k$

$$\sup_{\zeta \in \mathbb{C}^n, |\zeta| \leq 1} \left(|K_1(e^{i\theta})\zeta| + |K_2(e^{i\theta})\zeta|\right) \geq \|K\|_{\hat{B}} - \frac{1}{k} \tag{137}$$

From Lemma 3.3, it is possible to define $G^{(k)} \in \hat{B}_*$ for $k = 1, 2$ such that,

$$\begin{aligned} for\ \theta \in \Omega_k \ \ & TrG_1^{(k)}K_1^*(e^{i\theta}) + TrG_2^{(k)}K_2^*(e^{i\theta}) \\ & = \sup_{\zeta \in \mathbb{C}^n, |\zeta| \leq 1} \left(|K_1(e^{i\theta})\zeta| + |K_2(e^{i\theta})\zeta|\right) \cdot \left\| \begin{bmatrix} G_1^{(k)}(e^{i\theta}) \\ G_2^{(k)}(e^{i\theta}) \end{bmatrix} \right\|_{nuc} \qquad (138) \\ & \qquad and \left\| \begin{bmatrix} G_k^{(1)}(e^{i\theta}) \\ G_k^{(2)}(e^{i\theta}) \end{bmatrix} \right\|_{nuc} = \frac{1}{m(\Omega_k)} \\ for\ \theta \notin \Omega_k \ \ & G_k^{(1)}(e^{i\theta}) = G_k^{(2)}(e^{i\theta}) = 0 \qquad (139) \end{aligned}$$

Hence $\|G\|_{\hat{B}_*} = 1$ and $|[G|K]| \geq \|K\|_{\hat{B}} - \frac{1}{k}$. Since k is arbitrary, we have deduced that the induced norm of ϕ is indeed equal to $\|K\|_{\hat{B}}$.

ii. (132) follows directly from the proof of Theorem 3.1b, since this argument does not exploit any topological properties of the space B_{**}.

□

Chapter 4

Summary of the Numerical Solution to ORDAP based on Convex Programming

In this chapter we will briefly summarize the arguments that allow convex programming techniques to be used in conjunction with duality theory to produce a numerical solution to the two-disc optimization of (2).

4.1 General Remarks

The results of Chapter 3 included predual and dual representations of the two-disc problem of (43). These comprise maximizations or suprema, in contrast to the infimum of the original form (43). Thus, the results of Chapter 3 lead naturally to a dual pair of numerical solutions, which approach the optimal μ_0 from opposite directions, and have the virtue of producing estimates of μ_0 together with tolerances on these estimates. The numerical solutions, which are based on convex programming methods (see Boyd [5]), will be briefly summarized here for the SISO case and used to compute the example cited in Chapter 1. For convenience, W is normalized such that $\|W\|_\infty \leq 1$.

The first of these solutions, which will be referred to as the 'primary', exploits the fact that $\||W(I - P_0Q)| + |VP_0Q|\|_\infty$ is a convex function of Q. The problem (2) is infinite-dimensional in the sense that there is no finite limit to the number of parameters generally required to specify Q. However, (2) can be approximated by a finite variable convex optimization in the following manner: restrict Q to lie in the space $\mathcal{P}_m$ consisting of degree m analytic trigonometric polynomials of the form $a_0 + a_1 z + ... + a_m z^m$ with real coefficients, and then discretize the unit circle sufficiently finely w.r.t. m. [7] This yields a convex problem in the variables $a_0, ..., a_m$. For any fixed m, these convex problems generate upper bounds for μ_0 and suboptimal control laws, since Q is restricted to a proper subspace of H^∞. Such problems are then standard applications of convex programming techniques. The technique that will be employed here is the Ellipsoid algorithm of Shor, Yudin and Nemirovski [45]. This algorithm is chosen primarily for its non-hueristic stopping criterion [5]. That is, for a prespecified tolerance ϵ the Ellipsoid algorithm will terminate only when the estimate of the optimum is guaranteed to be within ϵ of the true minimum. This algorithm has the advantages of simplicity, robustness, and polynomial execution time. However it can be slower than other methods such as the method of analytic centers.

The second or 'dual' solution exploits the representation of (2) in the predual, as follows. μ_0 can be expressed in terms of the following minimization,

$$\mu_0 = \sup_{\phi \in S_\perp, \|\phi\|_{B_*} \leq 1} \left| \phi\left(\begin{bmatrix} U^* W \\ 0 \end{bmatrix} \right) \right| = - \inf_{\phi \in S_\perp, \|\phi\|_{B_*} \leq 1} Re\left(\phi\left(\begin{bmatrix} U^* W \\ 0 \end{bmatrix} \right) \right). \quad (140)$$

(140) holds, since for fixed U^*W, $\phi\left(\begin{bmatrix} U^* W \\ 0 \end{bmatrix} \right)$ is a linear function of ϕ and $S_\perp$ is a subspace. The dual solution is therefore -1 times the convex minimization shown on the right of (140). For the purposes of simplifying this minimization, the preorthogonal complement $S_\perp$ described in (49) can be shown to take the form

$$S_\perp = \left\{ \frac{1}{|W|^2 + |V|^2} \left(\begin{bmatrix} V^* \\ -W^* \end{bmatrix} X + \begin{bmatrix} W \\ V \end{bmatrix} Y \right) \ : \ X \in L^1, Y \in \overline{H}_0^1 \right\}. \quad (141)$$

[7] the meaning of the term 'sufficiently finely' is discussed in 4.2

X is then restricted to lie in the subspace of L^1 consisting of trigonometric polynomials of the form $a_{-\frac{m}{2}}z^{-\frac{m}{2}} + \ldots + a_0 + a_1 z + \ldots + a_{\frac{m}{2}} z^{\frac{m}{2}}$ denoted by $\mathcal{G}_m$. Y is restricted to the subspace of $\overline{H}_0^1$ comprising m-dimensional anti-analytic trigonometric polynomials of the form $a_{-1}z^{-1} + \ldots + a_{-m}z^{-m}$ denoted by $\mathcal{H}_m$ (m is assumed to be even). After sufficiently fine discretization of the unit circle, the resulting finite variable constrained convex optimization yields upper bounds for the infimum in (140) and lower bounds for the optimization represented by μ_0. It can be shown that as $m \to \infty$, the upper and lower bounds obtained by this procedure converge. This follows from the fact that any continuous H^∞ function on the unit circle can be approximated uniformly by H^∞ polynomials, and any integrable H_0^1 function on the unit circle can be approximated in the mean by H_0^1 polynomials. Thus, both finite dimensional convex minimizations approach the infinite-dimensional minimum.

It should be noted that the dual formulations of (47) are critical to the implementation of this method. While convex programs based on the finite variable approximation to the primary problem produce upper bounds for μ_0, they give no indication of how far these estimates are from the optimum, and so by themselves are of limited utility. The lower bounds obtained from the dual problem, though, enable such estimates of μ_0 to be expressed to a known tolerance.

4.2 Approximate Representation by Euclidean Convex Optimizations

4.2.1 The Primary Problem

To establish a Euclidean vector problem which approximates the primary optimization (2) we use the following representation

$$\mu_0 = \inf_{Q \in H^\infty} \left\| \begin{bmatrix} U^* W \\ 0 \end{bmatrix} - \begin{bmatrix} W\Lambda^{-1} \\ V\Lambda^{-1} \end{bmatrix} Q \right\|_B \tag{142}$$

If Q is then restricted to lie in the subspace $\mathcal{P}_m$, the resulting optimum is defined to be

$$\mu_0^{(m)} := \inf_{Q \in \mathcal{P}_m} \left\| \begin{bmatrix} U^* W \\ 0 \end{bmatrix} - \begin{bmatrix} W\Lambda^{-1} \\ V\Lambda^{-1} \end{bmatrix} Q \right\|_B \tag{143}$$

If $\|W\|_\infty \leq 1$, then $\mu_0^{(m)} \leq 1$. Thus Q can further be restricted to those elements of $\mathcal{P}_m$ which satisfy the inequality (144) below, without effecting the infimum in (143).

$$\| \, |U^*W - W\Lambda^{-1}Q| + |V\Lambda^{-1}Q| \, \|_\infty \leq 1 \tag{144}$$

If (144) holds then we have that

$$\begin{aligned} \frac{|W|^2}{|W|^2+|V|^2}(e^{i\theta}) \cdot \left|(\Lambda U^* - Q)(e^{i\theta})\right|^2 & \\ + \frac{|V|^2}{|W|^2+|V|^2}(e^{i\theta}) \cdot |Q(e^{i\theta})|^2 &\leq 1 \;\; l.a.e \qquad (145) \\ \Rightarrow Min\left(\left|(\Lambda U^* - Q)(e^{i\theta})\right|, \left|Q(e^{i\theta})\right|\right) &\leq 1 \;\; l.a.e. \qquad (146) \\ \Rightarrow \|Q\|_\infty &\leq 1 + \|\Lambda\|_\infty \qquad (147) \end{aligned}$$

Since Q is a degree m trigonometric polynomial in $\mathcal{P}_m$ which satisfies (147), then by Bernstein's theorem (see for example [42]) we conclude that $\|Q'\|_\infty \leq m(1 + \|\Lambda\|)$. Thus, we have established that $\mu_0^{(m)}$ can be found by searching over the *subset* of $\mathcal{P}_m$ defined by (147). Moreover the derivatives of all such Q are uniformly bounded above by $m(1+\|\Lambda\|_\infty)$. This second point is a key issue because the uniform boundedness of this derivative, coupled with the assumption that W, V, U^*W are uniformly Lipschitz, allows us to compute (143) based on inspection of a sufficiently fine partition of discrete frequencies. In particular, for each $Q \in \mathcal{P}_m$ which satisfies (147), the function of θ described by

$$|(U^*W - W\Lambda^{-1}Q)(e^{i\theta})| + |\Lambda^{-1}VQ(e^{i\theta})| \tag{148}$$

is uniformly Lipschitz in θ, with Lipschitz constant bounded above by

$$Lip(U^*W) + (1+\|\Lambda\|_\infty)\cdot\left\{Lip(\frac{W}{\Lambda}) + Lip(\frac{V}{\Lambda}) + 2m\right\} := L_\epsilon \qquad (149)$$

The function $Lip(\cdot)$ denotes the l.u.b of all uniform Lipschitz constants associated with its argument. Define, for a fixed ϵ, $N_s := int\left(\frac{\pi L_\epsilon}{2\epsilon}\right)$, and $\theta_k := \frac{k\pi}{N_s}$ for integers $k \in [1, N_s]$. From the uniform Lipschitz continuity of (148), all $Q \in \mathcal{P}_m$ which satisfy (147), also satisfy (150).

$$\begin{aligned}&\left\|\, |U^*W - W\Lambda^{-1}Q| + |V\Lambda^{-1}Q| \,\right\|_\infty \\ &-Max_{k\in\mathbb{Z}\cap[1,N_s]}\left(|(U^*W - W\Lambda^{-1}Q)(e^{i\theta_k})| + |V\Lambda^{-1}Q(e^{i\theta_k})|\right) \in [-\epsilon,\epsilon]\end{aligned} \qquad (150)$$

It should be noted that the uniform partition of the unit circle may be conservative, in the sense that non-uniform partitions with fewer elements may be found, for which (150) is also satisfied. From the point of view of actual computation this would mean great savings in memory and speed of the resulting programs. The uniform case is considered here for the clarity of presentation.

The optimization represented by the quantity $\mu_0^{(m)}$ can now be reduced to a Euclidean vector convex problem, incurring an error smaller than ϵ in the following manner. Define $x \in \mathbb{R}^{m+1}$ to be the vector of coefficients of $Q \in \mathcal{P}_m$ in ascending powers. Let $S_m \subset \mathbb{R}^{m+1}$ denote the convex set consisting of all coefficients x of those $Q \in \mathcal{P}_m$ which satisfy (147). Throughout this chapter, for cases where P_0, W and V have real Fourier coefficients, the polynomials of $\mathcal{P}_m, \mathcal{G}_m$ and $\mathcal{H}_m$ can be taken to have real coefficients without effecting any of the optima cited. The computation of $\mu_0^{(m)}$ is, to accuracy ϵ, approximately equivalent to

$$\inf_{x\in S_m} Max_{k\in\mathbb{Z}\cap[1,N_s]}\Big(\Big|U^*W(e^{i\theta_k}) - W\Lambda^{-1}(e^{i\theta_k})\sum_{r=0}^{m} x_r e^{ir\theta_k}\Big|$$

$$+\left|V\Lambda^{-1}(e^{i\theta_k})\sum_{r=0}^{m}x_r e^{ir\theta_k}\right|\Big) =: \tilde{\mu}_0^{(m)} \tag{151}$$

(151) can then be expressed in the following Euclidean vector form.

$$\tilde{\mu}_0^{(m)} = \inf_{x\in S_m} \|abs(b-Ax)+abs(Cx)\|_{l_\infty^{N_s}} \tag{152}$$

where $abs(\cdot)$ denotes the Euclidean *vector* obtained by taking the magnitudes of each element of the argument vector, b is the N_s dimensional column vector defined by $[b]_k = U^*W(e^{i\theta_k})$, A is the $N_s\times(m+1)$ matrix defined by $[A]_{r,s} = W\Lambda^{-1}(e^{i\theta_r})e^{i(s-1)\theta_r}$, and C is the $N_s\times(m+1)$ matrix defined by $[C]_{r,s} = V\Lambda^{-1}(e^{i\theta_r})e^{i(s-1)\theta_r}$.

4.2.2 The Dual case

The first step in obtaining an approximately equivalent Euclidean vector optimization for the dual problem (140), is to obtain representations for $S_\perp$ (49) and the bilinear forms of (51) in terms of W and V. The two terms in (49) that must be expanded are $(I-RR^*)(L^1\times L^1)$ and $R\overline{H}_0^1$. Since $R = \Lambda^{-1}\begin{bmatrix} W \\ V \end{bmatrix}$, $(I-RR^*)$ can be manipulated into the form

$$(I-RR^*)(e^{i\theta}) = \frac{1}{|W|^2+|V|^2}\begin{bmatrix} V^* \\ -W^* \end{bmatrix}\begin{bmatrix} V & -W \end{bmatrix} \tag{153}$$

Thus $(I-RR^*)(L^1\times L^1) = \frac{1}{|W|^2+|V|^2}\begin{bmatrix} V^* \\ -W^* \end{bmatrix}L^1$ from assumption A1 $R\overline{H}_0^1 = \Lambda^{-1}\begin{bmatrix} W \\ V \end{bmatrix}\overline{H}_0^1 = \frac{1}{|W|^2+|V|^2}\begin{bmatrix} W \\ V \end{bmatrix}\Lambda_-\overline{H}_0^1$. Since Λ_- is an invertible element of $\overline{H}^\infty$ (from assumption A1), we obtain $R\overline{H}_0^1 = \frac{1}{|W|^2+|V|^2}\begin{bmatrix} W \\ V \end{bmatrix}\overline{H}_0^1$. Thus,

$$S_\perp = \left\{\frac{1}{|W|^2+|V|^2}\left(\begin{bmatrix} V^* \\ -W^* \end{bmatrix}X + \begin{bmatrix} W \\ V \end{bmatrix}Y\right) \;:\; X\in L^1, Y\in\overline{H}_0^1\right\} \tag{154}$$

Using this representation, the constraint that $G \in S_\perp$ must satisfy $\|G\|_{B_*} \leq 1$, can be expressed

$$\left\| \frac{1}{|W|^2+|V|^2}\left(\begin{bmatrix} V^* \\ -W^* \end{bmatrix} X + \begin{bmatrix} W \\ V \end{bmatrix} Y \right) \right\|_{B_*} \leq 1 \tag{155}$$

and the bilinear forms of (51), for $K = \begin{bmatrix} U^*W \\ 0 \end{bmatrix}$ and $G \in S_\perp$ can be expressed,

$$[G|K] = \frac{1}{2\pi}\int_0^{2\pi} \frac{1}{|W|^2+|V|^2} \begin{bmatrix} UW^* & 0 \end{bmatrix} \left(\begin{bmatrix} V^* \\ -W^* \end{bmatrix} X + \begin{bmatrix} W \\ V \end{bmatrix} Y \right) dt \tag{156}$$

The second step towards attaining the desired representation mirrors the argument of 4.2.1. X and Y in (154) are respectively restricted to $\mathcal{G}_m$ and $\mathcal{H}_m$, with the result that lower bounds for the optimization of (140) are obtained. For $X \in \mathcal{G}_m$ and $Y \in \mathcal{H}_m$, we will show that the constraint that the B_* norm of each element of $S_\perp$ in (154) be less than or equal to unity (155), implies that the derivatives of X and Y on the unit circle are uniformly bounded. The assumption of Lipschitz continuity of W and V, enables the integral of (156) and the integral implicit in (155), to be approximated by certain weighted sums. Each integral represents the evaluation of the functional represented by K, and the norm restriction in (155) respectively.

If an element of $S_\perp$ has representation (154) and has a norm bounded by unity i.e. (155) holds, then

$$\left\| \frac{1}{|W|^2+|V|^2}(V^*X + WY) \right\|_1 \leq 1 \;\; and \tag{157}$$

$$\left\| \frac{1}{|W|^2+|V|^2}(-W^*X + VY) \right\|_1 \leq 1 \tag{158}$$

Multiplying the term inside the norm symbols in (157) by V, then multiplying the corresponding term in (158) by W and subtracting, followed by the application of the triangle inequality gives $\|X\|_1 \leq 1 + \|V\|_\infty$. A similar argument yields $\|X\|_1 \leq$

$1+\|V\|_\infty$. It follows that each of the Fourier coefficients of X and Y are bounded by $1+\|V\|_\infty$. Thus if $X \in \mathcal{G}_m, Y \in \mathcal{H}_m$ then $\|X\|_\infty \leq (m+1)(1+\|V\|_\infty)$, $\|Y\|_\infty \leq m(1+\|V\|_\infty)$ and the derivatives w.r.t θ of $X(e^{i\theta})$ and $Y(e^{i\theta})$ are such that $\|X'(e^{i\theta})\|_\infty \leq \frac{m(m+2)}{4}(1+\|V\|_\infty)$, $\|Y'(e^{i\theta})\|_\infty \leq \frac{m(m+1)}{2}(1+\|V\|_\infty)$. Thus for $X \in \mathcal{G}_m, Y \in \mathcal{H}_m$ each element of $S_\perp$ in (154) has uniform Lipschitz constant bounded above by

$$(1+\|V\|_\infty)\left\{(m+1)\left[Lip\left(\frac{W}{|W|^2+|V|^2}\right)+Lip\left(\frac{V}{|W|^2+|V|^2}\right)\right]+\frac{m^2+1}{\sqrt{2}}\right\} =: L_1$$

The integrand of (156) has uniform Lipschitz constant bounded above by

$$(1+\|V\|_\infty)\left\{(m+1)Lip\left(\frac{U^*WV}{|W|^2+|V|^2}\right)+mLip\left(\frac{U|W|^2}{|W|^2+|V|^2}\right)+\frac{5}{8}m^2+\frac{3}{4}m\right\} := L_2$$

If θ_k are chosen as in 4 .1 with $N_s = int\frac{\pi Mcx(L_1,L_2)}{2\epsilon}$ then the norm in (155) is within ϵ of

$$\frac{1}{N_s}\sum_{k=1}^{N_s}\frac{1}{|W|^2+|V|^2}(e^{i\theta_k})\cdot Max\left(|V^*X(e^{i\theta_k})+WY(e^{i\theta_k})|, |-W^*X(e^{i\theta_k})+VY(e^{i\theta_k})|\right) \tag{159}$$

and the integral of (156) is within ϵ of

$$\frac{1}{N_s}\sum_{k=1}^{N_s}\frac{UW^*}{|W|^2+|V|^2}(V^*X+WY)(e^{i\theta_k}) \tag{160}$$

If $x \in \mathbb{R}^{m+1}$, $y \in \mathbb{R}^m$ are respectively chosen to be the coefficients of X and Y in ascending order, then (159) can be written,

$$\left\|Max\left\{abs\begin{bmatrix} B^{(1)} & B^{(2)}\end{bmatrix}\begin{bmatrix} x \\ y\end{bmatrix}, abs\begin{bmatrix} C^{(1)} & C^{(2)}\end{bmatrix}\begin{bmatrix} x \\ y\end{bmatrix}\right\}\right\|_{l^1_{N_s}} \tag{161}$$

and (160) can be written,

$$w_{N_s}^T \left(\left[\begin{array}{cc} A^{(1)} & A^{(2)} \end{array} \right] \left[\begin{array}{c} x \\ y \end{array} \right] \right) \tag{162}$$

where $w_{N_s}^T$ is an N_s-dimensional row vector, whose entries are all 1, Max denotes the vector formed by taking the element-by-element maximum of its two arguments, and

$$\begin{aligned}
[A^{(1)}]_{k,r} &= \frac{UW^*V^*}{N_s(|W|^2+|V|^2)}(e^{i\theta_k})e^{i\theta_k(r-\frac{m}{2}-1)}, \ [A^{(2)}]_{k,s} = \frac{U|W|^2}{N_s(|W|^2+|V|^2)}(e^{i\theta_k})e^{-i\theta_k s} \\
[B^{(1)}]_{k,r} &= \frac{V^*}{N_s(|W|^2+|V|^2)}(e^{i\theta_k})e^{i\theta_k(r-\frac{m}{2}-1)}, \ [B^{(2)}]_{k,s} = \frac{W}{N_s(|W|^2+|V|^2)}(e^{i\theta_k})e^{-is\theta_k} \\
[C^{(1)}]_{k,r} &= \frac{-W^*}{N_s(|W|^2+|V|^2)}(e^{i\theta_k})e^{i\theta_k(r-\frac{m}{2}-1)}, \ [C^{(2)}]_{k,s} = \frac{V}{N_s(|W|^2+|V|^2)}(e^{i\theta_k})e^{-is\theta_k}
\end{aligned}$$

where k is an integer between 1 and N_s, r is an integer between 1 and $m+1$, and s is an integer between 1 and m. Thus the Euclidean vector convex optimization,

$$\begin{aligned}
-\inf \Bigg\{ & Re\left(w_{N_s}^T \left[\begin{array}{cc} A^{(1)} & A^{(2)} \end{array} \right] \left[\begin{array}{c} x \\ y \end{array} \right] \right) + 2\epsilon \ : \ \left[\begin{array}{c} x \\ y \end{array} \right] \in \mathbb{R}^{2m+1}, \\
& \left\| Max\left\{ abs \left[\begin{array}{cc} B^{(1)} & B^{(2)} \end{array} \right] \left[\begin{array}{c} x \\ y \end{array} \right], abs \left[\begin{array}{cc} C^{(1)} & C^{(2)} \end{array} \right] \left[\begin{array}{c} x \\ y \end{array} \right] \right\} \right\|_{l^1_{N_s}} \leq 1 \Bigg\}
\end{aligned} \tag{163}$$

is a lower bound for μ_0 (43).

4.3 The Ellipsoid Algorithm of Shor, Yudin and Nemirovsky

Both optimizations of the type (152) and (163) are fairly standard applications of numerical convex programming techniques [5]. Outlined here is the Ellipsoid algorithm of Shor, Yudin and Nemirovski [45], whose implementation is discussed in detail in [5]. The only non-standard feature of (152) is that the convex set S_m has no simple description in coefficient space $\mathbb{R}^{m+1}$. This problem is dealt with by taking the starting ellipsoid in the Ellipsoid algorithm to contain S_m and then treating (152) as an unconstrained optimization. This is justifiable so long as the final x obtained for (152) does indeed lie in the set S_m. Although this is not guaranteed a priori, the condition that x correspond to a $Q \in S_m$ can be tested for upon termination of the algorithm. We have found it to hold in the numerical implementations that have been tried. If the final x does not correspond to a $Q \in S_m$ then an alternative to the development of 4.2.1 must be used in which the set S_m is replaced by a Euclidean sphere in coefficient space, and new derivative bounds established for Q as in the argument of 4.2.2. This would result in larger derivative upper bounds for Q, and necessitate a finer partition $\{\theta_k\}$. This could have been done at the outset, thereby eliminating any potential problem, at the cost of slower algorithm execution and more memory usage.

Coding the Ellipsoid Algorithm is merely a matter of substituting the above descriptions of the matrices $A, b, \left[\begin{array}{cc} A^{(1)} & A^{(2)} \end{array}\right], \left[\begin{array}{cc} B^{(1)} & B^{(2)} \end{array}\right], \left[\begin{array}{cc} C^{(1)} & C^{(2)} \end{array}\right]$ into the standard forms given in pages 326-328 of [5]. Implementation of the Ellipsoid Algorithm requires the computation of subgradients of the objective convex functionals (152) and (162), and of the constraint functional for the dual problem (161).

Let us briefly summarize the derivation of the appropriate subgradients. In order to find the subgradient associated with the functional represented by the unconstrained form of (152), let ψ_k be the convex functional on $\mathbb{R}^{m+1}$ defined by

$$\psi_k(x) := |e_k^T(Ax - b)| = Max_{\theta \in [0,2\pi)} Re\left(e^{i\theta} e_k^T(Ax - b)\right) \tag{164}$$

The subgradient set for ψ_k at $\bar{x}$ (where $\psi_k(\bar{x}) \neq 0$) contains the subgradient set of the convex functional defined by

$$\begin{gathered} x \to Re\left(e^{i\theta_{active}} e_k^T(Ax - b)\right) \\ where\ e^{i\theta_{active}} = \frac{\overline{e_k^T(A\bar{x} - b)}}{|e_k^T(A\bar{x} - b)|} \end{gathered} \tag{165}$$

Thus, a subgradient of the functional ψ_k at $x = \tilde{x}$, $\psi_k(\tilde{x}) \neq 0$ is

$$Re\left(\frac{\overline{e_k^T(A\tilde{x}-b)}}{|e_k^T(A\tilde{x}-b)|}A^T e_k\right) \tag{166}$$

On the other hand, if $\psi_k(\tilde{x}) = 0$ then $\tilde{x}$ is a global minimum for ψ_k, and 0 is a subgradient. A similar argument shows that the subgradient set of the convex functional represented by

$$\hat{\psi}_k(x) = |e_k^T Cx| \tag{167}$$

includes the element

$$Re\left(\frac{\overline{e_k^T C\tilde{x}}}{|e_k^T C\tilde{x}|}C^T e_k\right) \quad if\ \hat{\psi}_k(\tilde{x}) \neq 0. \tag{168}$$

If $\hat{\psi}_k(\tilde{x}) = 0$ then $\tilde{x}$ is a global minimum of ψ_k and 0 is a subgradient. Thus, the subgradient set at $\tilde{x}$ for the functional obtained by taking the maximum of $\psi_k + \hat{\psi}_k$ over all integers k between 1 and N_s (i.e. the expression (152)), contains the element

$$Re\left(\frac{\overline{e_{k_0}^T(A\tilde{x}-b)}}{|e_{k_0}^T(A\tilde{x}-b)|}A^T e_{k_0} + \frac{\overline{e_{k_0}^T C\tilde{x}}}{|e_{k_0}^T C\tilde{x}|}C^T e_{k_0}\right) \tag{169}$$

where $k_0 := argmax_{1\leq k\leq N_s} e_k^T\left[abs(b - Ax) + abs(Cx)\right]$. Note that the first term is taken as zero if $\psi_{k_0}(\tilde{x}) = 0$, and the second term is taken as zero if $\hat{\psi}_{k_0}(\tilde{x}) = 0$. Thus, (169) represents a subgradient of the convex functional defined in (152).

Define a convex functional ϕ on $\mathbb{R}^{2m+1}$ by

$$\phi(\omega) := Re\left(w_{N_s}^T \begin{bmatrix} A^{(1)} & A^{(2)} \end{bmatrix} \omega\right) \tag{170}$$

The subgradient set for ϕ comprises the element

$$Re\left(\begin{bmatrix} A^{(1)T} \\ A^{(2)T} \end{bmatrix} w_{N_s}\right) \tag{171}$$

Thus (171) gives a subgradient for the functional defined by (162).

In order to obtain a subgradient for the constraint functional of (161), define the convex functionals $\phi_k^{(1)}, \phi_k^{(2)}$ on $\mathbb{R}^{2m+1}$,

$$\phi_k^{(1)}(\omega) := \left|e_k^T \begin{bmatrix} B^{(1)} & B^{(2)} \end{bmatrix} \omega\right| = Max_{\theta\in[0,2\pi)} Re\left(e^{i\theta} e_k^T \begin{bmatrix} B^{(1)} & B^{(2)} \end{bmatrix} \omega\right)$$
$$\phi_k^{(2)}(\omega) := \left|e_k^T \begin{bmatrix} C^{(1)} & C^{(2)} \end{bmatrix} \omega\right| = Max_{\theta\in[0,2\pi)} Re\left(e^{i\theta} e_k^T \begin{bmatrix} C^{(1)} & C^{(2)} \end{bmatrix} \omega\right)$$

The subgradient set for $\phi_k^{(1)}$ at $\tilde{\omega} \in \mathbb{R}^{2m+1}$ which satisfies $\phi_k^{(1)}(\tilde{\omega}) \neq 0$ contains the vector

$$Re\left(\frac{\overline{e_k^T \begin{bmatrix} B^{(1)} & B^{(2)} \end{bmatrix} \tilde{\omega}}}{|e_k^T \begin{bmatrix} B^{(1)} & B^{(2)} \end{bmatrix} \tilde{\omega}|} \begin{bmatrix} B^{(1)*} \\ B^{(2)*} \end{bmatrix} e_k\right) =: g_k(1) \tag{172}$$

The subgradient set for $\phi_k^{(2)}$ at $\tilde{\omega}$ such that $\phi_k^{(2)}(\tilde{\omega}) \neq 0$ contains $g_k(2)$, which is defined as for $g_k(1)$ but with B and C interchanged. If $\phi_k^{(2)}(\tilde{\omega}) = 0$ or $\phi_k^{(2)}(\tilde{\omega}) = 0$ then the respective subgradient sets contain 0, and the respective g_k term is defined to be zero.

Thus, a subgradient of the constraint functional defined by (161) is

$$\sum_{k=1}^{N_s} g_k\left(argmax_{r=1,2}[\phi_k^{(r)}(\omega)]\right) \tag{173}$$

4.4 An Example.

In this subsection the methods of this chapter are used to estimate $\chi(r)$ for a case where the location of the fixed point is highly sensitive to inaccuracies in the estimates of $\chi(r)$. The results of this computation were used to plot the curve of Fig. 3 in the introduction.

We plot estimates of $\chi(r)$ (c.f. (152)) and its related predual problem as a function of $r \in [0,1]$, in order to estimate the optimal robust disturbance attenuation. The weights W, V are $0.17\left(\frac{1.8+0.2s}{1+s}\right)^3$ and $0.22\left(\frac{0.1+2.1s}{1+s}\right)^2$ respectively. The nominal plant P_0 is $\frac{1.9-0.1s}{(1+s)(1.9+0.1s)}$, as stated in the introduction.

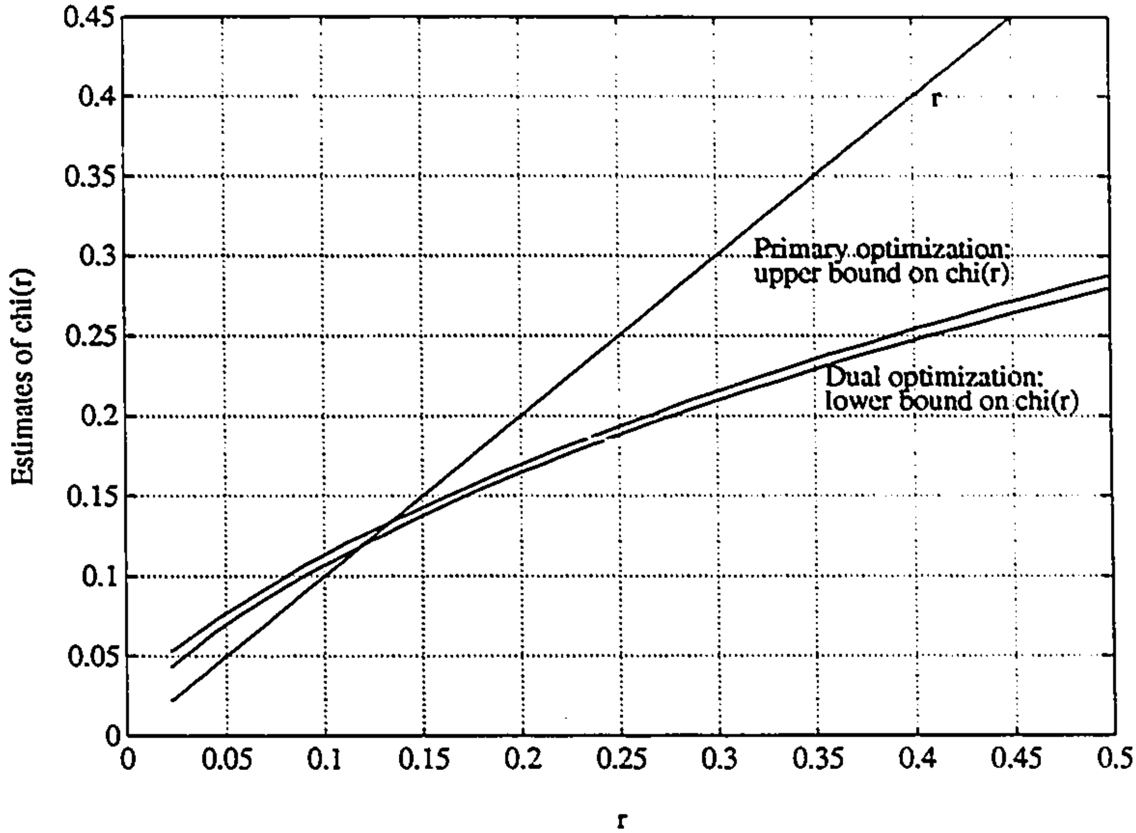

Figure 9: Plots of estimates of $\chi(r)$ vs r based on convex optimization

The lower curve in fig. 9 is generated by a convex program for the finite variable problem ($m = 6$) resulting from the infimum in (163). The upper curve is produced by a convex program for the finite variable problem ($m = 6$), resulting from (152). The curves are generated by an ellipsoid algorithm for the finite variable approximations to the primary and dual problems. The parameter r takes on the values of successive multiples of 0.01, from 0.02 to 0.43 i.e., a total of 42 numerical optimizations, each of which is computed to an accuracy better than ± 0.002.

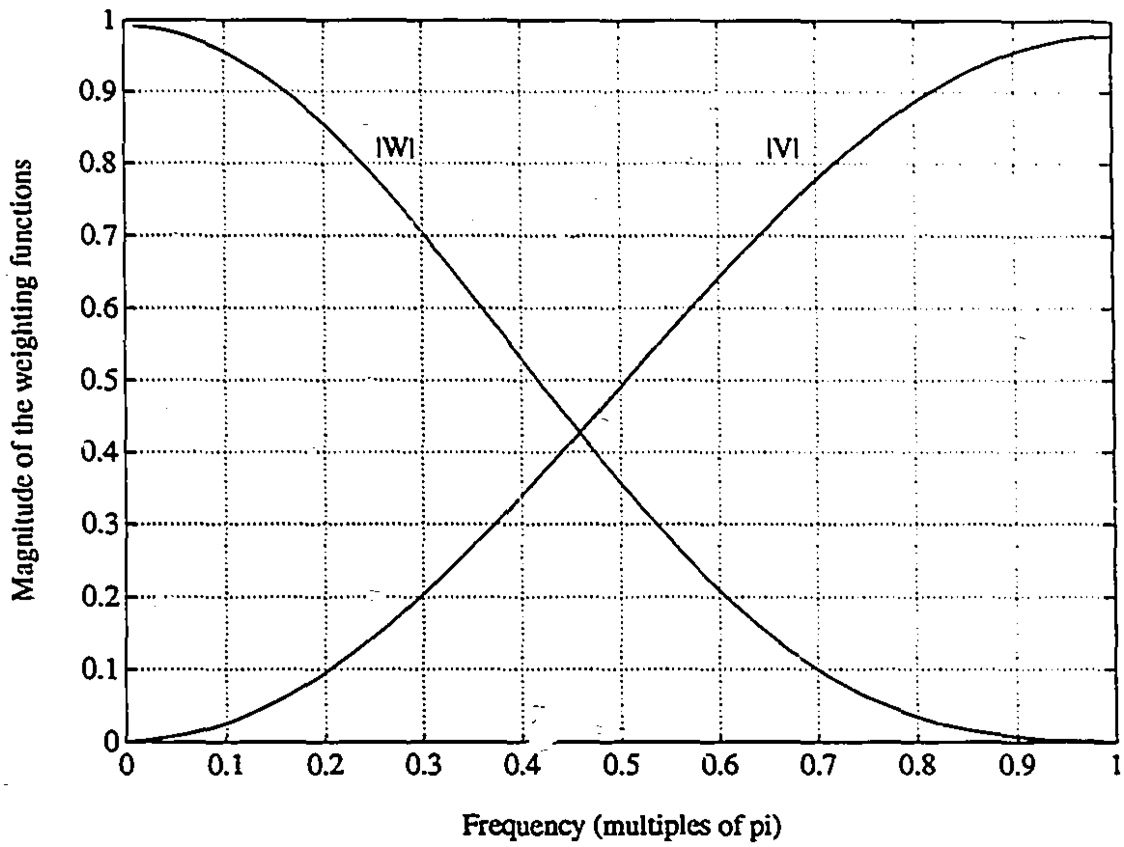

Figure 10: Magnitudes of the weights W and V vs frequency

Chapter 5

The Asymptotic Case of Almost Complementary Weightings

In this chapter we investigate the ORDAP for the case where the infinity norm of the product of the weightings W and V is small compared with the optimal robust performance. Under this condition we label the weightings W and V 'almost' complementary. For almost-complementary weightings, approximate solutions of the two-disc problems are obtained, which are expressed in terms of Hankel norms that can be computed by standard means. When combined with a bound on the slope of the function $\chi(\cdot)$ in the vicinity of the fixed-point, these approximations yield approximates solutions to the ORDAP which are accurate in the limit of small $\|W^*V\|_\infty$. Also, explicit tolerance limits are derived and shown to be proportional to $\|W^*V\|_\infty$. Finally, bounds are obtained on the growth of optimal robust performance vs the radius of the sphere of uncertainty $\mathcal{B}(P_0, V)$. The bounds are independent of any plant or weighting characteristics and strengthen the continuity result of Theorem 3.7.

One of the qualitative deductions of Chapter 3 was that increases in the plant uncertainty on one frequency range produce a strict reduction in the potential for disturbance rejection over all other frequency ranges (see Sect. 3.5.2). Part of the objective of this chapter is to gain greater insight into this phenomenon by explicitly asking; what is the coupling between sensitivity on one frequency interval and uncertainty on another? In particular, if frequency response information is available over a limited frequency range, what are the characteristics which limit or enhance the ability of feedback to reject disturbances on that range? Implicit in this question is form of the ORDAP, since there is uncertainty outside the range where frequency information is available. Although the material of the last four chapters has illuminated the ORDAP in both qualitative and quantitative settings, it lacks the explicit nature of a closed form solution and so cannot be used directly to gain an intuition for these questions. By contrast the approximate results of this chapter, which apply where disturbances and plant uncertainty occur over predominantly different frequency ranges, are more explicit. They enable us to isolate the factors affecting the coupling between uncertainty and performance.

The following assumption, which includes the notion that W and V are almost complementary will be needed throughout this chapter.

A4 W and V are commensurate (see definition 2.1) and $\|W^*V\|_\infty = \epsilon << 1$.

5.1 Mixed Norm Problems and the Hankel/Toeplitz Approximates

For this section, we shall assume that the ORDAP takes the form of (16) of Chapter 2, with $W_1 = I$ and W and V commensurate. The resulting mixed norm two-disc problem takes the form

$$\inf_{Q \in H^\infty_{n\times n}} \left\| \begin{bmatrix} (I - P_0Q)W \\ P_0QV \end{bmatrix} \right\|_B =: \mu \tag{174}$$

The initial objective is to approximate (174) by a standard H^∞ 'two-block' optimization. This will then be minimized by Hankel methods to give upper and lower bounds for μ, which should should be accurate in the limit as $\|W^*V\|_\infty \to 0$. For this purpose introduce the cost functions,

$$J(Q, \theta, \zeta) := \{|(I - X)W\zeta| + |XV\zeta|\}(e^{i\theta})$$

$$J_{(2)}(Q,\theta,\zeta) := \left\{|(I-X)W\zeta|^2 + |XV\zeta|^2\right\}^{\frac{1}{2}}(e^{i\theta})$$

where P_0Q is denoted by X. Then,

$$\mu = \inf_{Q\in H^\infty_{nxn}} esssup\{J(Q,\theta,\zeta) : \theta\in[0,2\pi), \zeta\in\mathbf{C}^n, |\zeta|=1\}$$

We start with the identity

$$J^2(Q,\theta,\zeta) - J^2_{(2)}(Q,\theta,\zeta) = 2\{|(I-X)W\zeta|.|XV\zeta|\}(e^{i\theta}) \tag{175}$$

from which we will get an approximation of the form

$$\begin{aligned} J^2(Q,\theta,\zeta) &= J^2_{(2)}(Q,\theta,\zeta) + 2\{|(I-X)W\zeta|.|XV\zeta|\}(e^{i\theta}) \\ &= J^2_{(*)}(Q,\theta,\zeta) + \Delta(\theta,\zeta) \end{aligned} \tag{176}$$

Where $J^2_{(*)}(Q,\theta,\zeta)$ is a nominal cost consisting of terms which are quadratic in X or do not involve X, and Δ is a residual term.

Since W and V are assumed commensurate they can be expressed in the form $W = w_sA$, $V = v_sA$ where $w_s, v_s \in H^\infty$ and $A \in H^\infty_{nxn}$, $|A(e^{i\theta})| = 1$ $l.a.e.$.

Since $W, V \in H^\infty_{nxn}$, $\int log|W|d\theta > -\infty$, $\int log(|W|+|V|)d\theta > -\infty$, which implies that $|W|$ and $(|W|+|V|)$ admit spectral factorizations

$$|W| = (|W|)_-(|W|)_+ \qquad (|W|+|V|) = (|W|+|V|)_-(|W|+|V|)_+$$

where $(.)_+$,$(.)_-$ denote functions which are outer, or have outer adjoint, respectively. The functions $(1+\frac{|V|}{|W|})_+$ and $(1+\frac{|V|}{|W|})_-$ are defined by the identity, $\left(1+\frac{|V|}{|W|}\right)_\pm := \frac{(|W|+|V|)_\pm}{(|W|)_\pm}$. Define M to be the unimodular scalar function which satisfies $M^* = \left(1+\frac{|V|}{|W|}\right)_-^{-1}\left(1+\frac{|V|}{|W|}\right)_+$.

Let $\mathbf{\Gamma}_A$ denote the Hankel operator with symbol A. The first result is an estimate for μ^2.

Theorem 5.1 *For commensurate weights W and V,*

$$\mu^2 := \inf_{Q \in H^\infty_{n \times n}} ess\sup\{J^2(Q,\theta,\zeta)\} : \theta \in [0,2\pi), \zeta \in \mathbf{C}^n, |\zeta| = 1\}$$

satisfies,

$$\mu^2 = \inf_{Q \in H^\infty_{n \times n}} \left\| \begin{bmatrix} B^* M^* W - H_0 Q \\ \frac{1}{\sqrt{2}}(|W|.|V|)_+ A \end{bmatrix} \right\|_\infty^2 + \delta \text{ where } -\frac{1}{2}\|W^* V\|_\infty \leq \delta \leq \frac{1}{2}\|W^* V\|_\infty \tag{177}$$

$$\text{which implies that } \mu^2 = \|\Gamma_{M^* B^* W}\|^2 + \delta' \text{ where } 0 \leq \delta' \leq \|W^* V\|_\infty \tag{178}$$

A function $Q \in H^\infty_{n \times n}$ whose cost $J(Q,\theta,\zeta)$ satisfies (177) can be found as a solution to the Hankel Toeplitz problem determined in (182) of the following.

Proof. (175) assumes the form

$$\begin{aligned} &2\Big\{|(I-X)W\zeta|.|XV\zeta|\Big\}(e^{i\theta}) \\ &= 2\Big\{|w_s|.|v_s|.|(I-X)A\zeta|.|XA\zeta|\Big\}(e^{i\theta}) \qquad (\text{ as } |w_s| = |W|,\ |v_s| = |V|) \\ &= \Big\{|W|.|V|.[|(I-X)A\zeta|^2 + |XA\zeta|^2 - (|(I-X)A\zeta| - |XA\zeta|)^2]\Big\}(e^{i\theta}) \end{aligned}$$

Now, $(|(I-X)A\zeta| - |XA\zeta|)^2(e^{i\theta}) \leq |A\zeta(e^{i\theta})|^2 \leq 1$. Therefore,

$$\begin{aligned} &2\Big\{|(I-X)W\zeta|.|XV\zeta|\Big\}(e^{i\theta}) \qquad (179) \\ &= \zeta^*\Big\{|W|.|V|.[A^*(I-X)^*(I-X)A + A^*X^*XA - \frac{1}{2}A^*A]\Big\}(e^{i\theta})\zeta + \Delta(\theta,\zeta) \end{aligned}$$

where $\Delta(\theta,\zeta) := \{|W|.|V|.[\frac{1}{2}|A\zeta|^2 - (|(I-X)A\zeta| - |XA\zeta|)^2]\}$ which satisfies

$$\begin{aligned} -\frac{1}{2}(|W|.|V|.|A\zeta|^2)(e^{i\theta}) \ &\leq\ \Delta(\theta,\zeta) \leq \frac{1}{2}(|W|.|V|.|A\zeta|^2)(e^{i\theta}) \\ so\ -\frac{1}{2}\|W^*V\|_\infty \ &\leq\ \Delta(\theta,\zeta) \leq \frac{1}{2}\|W^*V\|_\infty \end{aligned}$$

It follows from (175) and (179) that J can be approximated by a quadratic function $J_{(-)}$,

$$J^2(Q,\theta,\zeta) = J^2_{(-)}(Q,\theta,\zeta) + \Delta(\theta,\zeta) \tag{180}$$

where,

$$\begin{aligned} J^2_{(-)}(Q,\theta,\zeta) &= \zeta^*\Big\{A^*(I-X)^*(|W|^2+|W|.|V|)(I-X)A \\ &\quad +A^*X^*(|V|^2+|W|.|V|)XA - \frac{1}{2}|W|.|V|A^*A\}(e^{i\theta})\zeta \\ &= \zeta^*\{A^*X^*(|W|+|V|)^2XA - A^*X^*(|W|^2+|W|.|V|)A \\ &\quad -(|W|^2+|W|.|V|)A^*XA + (|W|^2+\frac{1}{2}|W|.|V|)A^*A\}(e^{i\theta})\zeta \end{aligned}$$

Upon completing the square, get

$$\begin{aligned} J^2_{(-)}(Q,\theta,\zeta) &= \Big|\, [(|W|+|V|)^2_+XA - (|W|+|V|)^{-2}_-(|W|+|V|)|W|A]\zeta\, \Big|^2 \\ &\quad +\frac{1}{2}\Big\{|W|.|V|\zeta^*A^*A\zeta\Big\}(e^{i\theta}) \end{aligned} \tag{181}$$

Recall that $X = P_0Q$, $P = BH_0$. Define the modified weighting function $W_0 = (|W|+|V|)^2_+A$. Observe that in (181) we can equate the factor $(|W|+|V|)^{-2}_-(|W|+|V|)|W|A$ to the ratio of Nevanlinna class functions introduced above, $\left(1+\frac{|V|}{|W|}\right)^{-1}_-\left(1+\frac{|V|}{|W|}\right)_+W = M^*W$ (since $(|W|)^2_+A = W$). Then (181) takes the form,

$$J^2_{(-)}(Q,\theta,\zeta) = \Big\{|\,(H_0QW_0 - B^*M^*W)\zeta\,|^2 + \frac{1}{2}|W|.|V|.|A\zeta|^2\Big\}(e^{i\theta}) \tag{182}$$

The minimization of $ess\sup\{J^2_{(-)}(Q,\theta,\zeta) : \theta \in [0,2\pi), \zeta \in \mathbb{C}^n, |\zeta| = 1\}$ is now a standard Hankel-Toeplitz ('two block') problem, whose solution gives (177).

□

Remark. In the limit as $\|W^*V\|_\infty \to 0$, μ approaches $\|\Gamma_{B^*M^*W}\|$ which is a solution to a 'one block' problem (i.e with $V = 0$) but with symbol changed from B^*W to

B^*M^*W. The effect of the non-zero V is captured by the unimodular multiplicand M^*. In general M^* does not approach the identity as $\|W^*V\|_\infty \to 0$. The motivating example of Sect. 2.5 of Chapter 2 is an instance of this phenomenon.

5.2 Computability of the Hankel/Hankel Toeplitz Approximation to γ

If we are to exploit these 'sum of squares' two block estimates in the approximate solution of the ORDAP, either they must be explicitly computable or there must be some method of approximating them by numerically accessible quantities. The question arises because, even for rational W and V, M^* may be irrational (in general it is), and so the estimates cannot always be found directly. Note that [61] provided a means of finding $\|\Gamma_{M^*W}\|$ for non-rational inner functions M and rational $W \in H^\infty$, however in this case the M is not necessarily in H^∞, so these methods cannot be applied.

Lemma 5.1 gives a necessary and sufficient condition for M^* to be a rational function, which however can be quite restrictive. Lemma 5.2 exhibits a sufficient condition for the weaker conclusion that WM^* is continuous on $[0, 2\pi)$, allowing it to be approximated by rational functions. B^*W continuous on the unit circle is a sufficient condition for both $\|\Gamma_{M^*B^*W}\|$ and $\inf_{Q \in H^\infty_{n\times n}} \left\| \begin{bmatrix} M^*B^*W - Q \\ \frac{1}{\sqrt{2}}(|W|.|V|)_+ A \end{bmatrix} \right\|_\infty$ to be approximated to arbitrary accuracy by standard H^∞ optimization problems for finite dimensional systems.

Lemma 5.1 *If $W, V^{\pm 1} \in H^\infty_{n\times n}$ and W and V are commensurate then M^* is rational if and only if $W = v^2 V$ for some rational H^∞ function v.*

Lemma 5.2 *If in addition to the assumptions of Lemma 5.1 , $W(e^{i\theta})$ and $V(e^{i\theta})$ are uniformly Lipschitz on $[0, 2\pi)$ then there exists an explicitly computable sequence of* *rational functions R_k such that $R_k \to WM^*$ uniformly on the unit circle.*

Proof of Lemma 5.1 Let λ be the H^∞ function defined by the identity $\lambda I = V^{-1}W$.

$$M^* \in \mathcal{R}L^\infty \Leftrightarrow \left(1 + \frac{|V|}{|W|}\right)_+ \in \mathcal{R} \tag{183}$$

$$\Leftrightarrow 1 + |\lambda^{-1}| \in \mathcal{R} \Leftrightarrow |\lambda^{-1}| \in \mathcal{R} \tag{184}$$

Hence, since $\int_0^{2\pi} log|\lambda(e^{i\theta})|d\theta > -\infty$ (because $\lambda \in H^\infty$), there exists a unique outer rational H^∞ function v which satisfies the equation ([7]),

$$|v(e^{i\theta})|^2 = |\lambda(e^{i\theta})| \quad for\ almost\ all\ \theta \in [0, 2\pi) \tag{185}$$

Thus $|v^2| = |\lambda|$ *a.e.* Because both h^2 and v are outer functions (185) is equivalent to $\lambda = v^2$ and the lemma is proven.

□

Proof of Lemma 5.2 If we can show that $WM^* \in C[0, 2\pi)$ then the lemma will be proven. This follows because partial Cesaro sums of Fourier series of continuous functions converge uniformly on the unit circle.

$$\begin{aligned} arg(M_r^*)(e^{i\theta}) &= arg\frac{\left(1 + \frac{1}{|\lambda|}\right)_+}{\left(1 + \frac{1}{|\lambda|}\right)_-}(e^{i\theta}) = arg\frac{(|\lambda|+1)_+}{(|\lambda|+1)_-}\frac{(|\lambda|)_-}{(|\lambda|)_+} \\ &= 2arg\,(|\lambda|+1)_+ - 2arg\,(|\lambda|)_+ = 2arg\,(|\lambda|+1)_+ - arg\lambda(e^{i\theta}), \end{aligned} \tag{186}$$

if $W(e^{i\theta}) \neq 0$. From the assumptions of uniform Lipschitz continuity, commensurate weights and $V^{\pm 1} \in H^\infty_{n \times n}$, it follows that $\lambda(e^{i\theta})$ is continuous on $[0, 2\pi)$. Define the set $\mathcal{C} := \{\theta \in [0, 2\pi) \ : \ |W(e^{i\theta})| > 0\}$. If we can show $arg\,(|\lambda|)_+(\cdot)$ and $arg\,(|\lambda|+1)_+(\cdot)$ are in $C[0, 2\pi)$ then $M^*(e^{i\theta})$ would be continuous on $\mathcal{C}$. $WM^*(e^{i\theta})$ is forced to be continuous on the set $[0, 2\pi) - \mathcal{C}$, since W is continuous and zero on this set, and M^* is allpass. Since $(|\lambda|)_+^2 = \lambda$, we need only show that $arg\,(|\lambda|+1)_+(\cdot)$ is continuous on C, and the lemma will be proven.

$$arg\,(|\lambda|+1)_{+}\,(e^{i\theta}) = \frac{1}{2}\int_{0}^{\pi} K(\theta,t)log(|\lambda|+1)(e^{it})dt \tag{187}$$

Since $V^{-1} \in H^{\infty}_{n\times n}$ and $V(e^{i\theta}) \in LIP$, (where LIP denotes the class of uniformly Lipschitz continuous functions on $[0,2\pi)$), we have $V(e^{i\theta})^{-1} \in LIP$. Hence $\lambda \in LIP \Rightarrow |\lambda|+1 \in LIP$. This implies that $log(|\lambda|+1) \in LIP$. Thus from Koosis [32] Chapter V, Sect. E pp. 140, the integral expression on the RHS of (187), as a function of θ, is uniformly continuous on $[0,2\pi)$. Hence we have continuity of WM^{*} on the unit circle and the lemma is proven.

□

Remarks.

1. $\|\Gamma_{B^{*}R_k}\|$ can be found by standard methods (e.g. [58], [19], [61]), and $\|\Gamma_{B^{*}R_k}\| \to \|\Gamma_{M_r^{*}B^{*}W}\|$.

2. The assumption of Lipschitz continuity in Lemma 5.2 is satisfied by any $H^{\infty}_{n\times n}$ rational functions W and V.

5.3 Estimates of the Slope of $\chi(\cdot)$.

In this subsection we derive an estimate for the slope of the function $\chi(\cdot)$, under assumption A5. This estimate, given in Lemma 5.3, applies for ϵ sufficiently small in the statement of A4. It bounds the slope of chords joining $(r_0, \chi(r_0))$ to $(r_0+x, \chi(r_0+x))$ for $x \geq 0$ to be strictly less than a constant which is strictly less than unity. This estimate will enable us to conclude, in section 5.4, that the approximations to χ derived in Sect 5.1 will yield approximations to the fixed-point(s) of χ.

Lemma 5.3 *Under assumptions A1 and A4, if $\epsilon < \frac{r_0^2}{16}$ where r_0 is any fixed-point of χ, then for all $x > 0$,*

$$\chi(r_0 + x) - \chi(r_0) \leq x \left(\frac{4 + \frac{r_0}{2}}{4 + r_0} \right) \tag{188}$$

Remark. The quantity $\frac{4+\frac{r_0}{2}}{4+r_0}$ is close to unity for small r_0, resulting in cruder tolerances for the estimates of the optimal robust disturbance attenuation which are based on approximations of χ. In the absence of any additional assumptions this is unavoidable, since in the motivating example of Chapter 2, Sect. 2.5, a function $\chi(\cdot)$ was exhibited, satisfying the assumptions of Lemma 5.3, for which $\chi(r_0 + x) - \chi(r_0) \simeq x(1 - 2r_0)$ for small r_0.

Proof of Lemma 5.3. Fix $\delta > 0$. If $\chi(r_0) = r_0$ there exists $\hat{Q} \in H^{\infty}_{n\times n}$ such that

$$ess \sup_{\theta\in[0,2\pi)} \sup_{\zeta\in\mathbb{C}^n,\ |\zeta|\leq 1} \left(|(I - B\hat{Q})W(e^{i\theta})\zeta| + r_0|B\hat{Q}V(e^{i\theta})\zeta| \right) \leq r_0 + \delta \tag{189}$$

Define $\tilde{Q} := \frac{r_0+\lambda x}{r_0+x}\hat{Q}$ where $x > 0$ and some $\lambda \in (0,1)$ which will be defined below. (188) will be proven by exhibiting an upper bound for

$$ess \sup_{\theta\in[0,2\pi)} \sup_{\zeta\in\mathbb{C}^n,\ |\zeta|\leq 1} \left(|(I - U\tilde{Q})W(e^{i\theta})\zeta| + (r_0 + x)|U\tilde{Q}V(e^{i\theta})| \right) \tag{190}$$

for an appropriately chosen constant λ. This is in turn an upper bound for $\chi(r_0 + x)$. Define $\phi(e^{i\theta}) := \sup_{\zeta\in\mathbb{C}^n,\ |\zeta|\leq 1} \left(|(I - U\tilde{Q})W(e^{i\theta})\zeta| + (r_0 + x)|U\tilde{Q}V(e^{i\theta})| \right)$.

The upper bound is derived by applying two sets of inequalities for ϕ to the following two disjoint sets of the unit circle.

$$A := \{\theta \in [0, 2\pi) \ : \ \frac{|W(e^{i\theta})|}{r_0} < \frac{1}{2}\} \tag{191}$$

$$B := \{\theta \in [0, 2\pi) \ : \ |V(e^{i\theta})| \leq \frac{2\epsilon}{r_0}\} \cap A^c \tag{192}$$

Since W and V are commensurate and $\| \, |W| \cdot |V| \, \|_\infty \leq \epsilon$ we have $A \cup B = [0, 2\pi)$. Since $\tilde{Q} = \frac{r_0 + \lambda x}{r_0 + x} \hat{Q}$,

$$\phi(e^{i\theta}) \leq \sup_{\zeta \in \mathbb{C}^n,\ |\zeta| \leq 1} \Big(\frac{r_0 + \lambda x}{r_0 + x} \left\{ |(I - B\hat{Q})W(e^{i\theta})\zeta| + (r_0 + x)|B\hat{Q}V(e^{i\theta})\zeta| \right\} + \frac{(1-\lambda)x}{r_0 + x} |W(e^{i\theta})\zeta| \Big) \; for \; a.e \; \theta \in [0, 2\pi) \tag{193}$$

$$\leq \frac{r_0 + \lambda x}{r_0 + x} \left(r_0 + \delta + x|B\hat{Q}V(e^{i\theta})| \right) + \frac{(1-\lambda)x}{r_0 + x} |W(e^{i\theta})| \;\; l.a.e \tag{194}$$

From (189) we have $|B\hat{Q}V(e^{i\theta})| \leq \frac{r_0 + \delta}{r_0} \; \forall \theta \in [0, 2\pi)$. If $\theta \in A$, $|W(e^{i\theta})| \leq \frac{r_0}{2}$. In the light of these two inequalities (194) gives,

$$\theta \in A \Rightarrow \phi(e^{i\theta}) \leq r_0 + x\frac{1+\lambda}{2} + \delta(1 + \frac{\lambda x}{r_0}) l.a.e \tag{195}$$

We may also write,

$$\phi(e^{i\theta}) \leq \sup_{\zeta \in \mathbb{C}^n,\ |\zeta| \leq 1} \Big(|(I - B\hat{Q})W(e^{i\theta})\zeta| + r_0|B\hat{Q}V(e^{i\theta})\zeta| + \frac{(1-\lambda)x}{r_0 + x} |\hat{Q}W(e^{i\theta})\zeta| + \lambda x |B\hat{Q}V(e^{i\theta})\zeta| \Big) \tag{196}$$

$$\leq \sup_{\zeta \in \mathbb{C}^n,\ |\zeta| \leq 1} \left(r_0 + \delta + \frac{(1-\lambda)x}{r_0 + x} |B\hat{Q}W(e^{i\theta})\zeta| + \lambda x |B\hat{Q}V(e^{i\theta})\zeta| \right) \tag{197}$$

where each inequality holds l.a.e in θ. From (189) we have

$$\sup_{\zeta \in \mathbb{C}^n,\ |\zeta| \leq 1} \left(|B\hat{Q}W(e^{i\theta})\zeta| + r_0|B\hat{Q}V(e^{i\theta})\zeta| \right) \leq r_0 + 1 + \delta \tag{198}$$

$$\Rightarrow (|W(e^{i\theta})| + r_0|V(e^{i\theta})|)|B\hat{Q}A(e^{i\theta})| \leq r_0 + 1 + \delta \tag{199}$$

$$(since \; W \; and \; V \; are \; commensurate)$$

$$\Rightarrow |B\hat{Q}V(e^{i\theta})| \leq (r_0 + 1 + \delta) \frac{|V(e^{i\theta})|}{|V(e^{i\theta})| + |W(e^{i\theta})|} \tag{200}$$

Thus for $\theta \in B$, $|B\hat{Q}V(e^{i\theta})| \leq 4\epsilon \frac{(r_0 + 1 + \delta)}{r_0^2}$.

(197) and (200) taken together give for $\theta \in B$

$$\phi(e^{i\theta}) \leq r_0 + \delta + (1-\lambda)\frac{x(r_0+1+\delta)}{r_0+x} + \frac{4\epsilon\lambda x(r_0+1+\delta)}{r_0^2} \tag{201}$$

Since $\epsilon < \frac{r_0^2}{16}$, λ can be chosen as $\frac{4}{4+r_0}$ to obtain

$$\begin{aligned} \phi(e^{i\theta}) &\leq r_0 + x\left(\frac{4+\frac{r_0}{2}}{4+r_0}\right) \quad \textit{for a.e } \theta \in B & (202) \\ \phi(e^{i\theta}) &\leq r_0 + x\left(\frac{4+\frac{r_0}{2}}{4+r_0} + \delta(1+\frac{\lambda x}{r_0})\right) \quad \textit{for a.e } \theta \in A & (203) \end{aligned}$$

Because $\chi(r_0+x) \leq \sup_{\theta\in[0,2\pi)} \phi(e^{i\theta})$ (202) and (203) hold with $\chi(r_0+x)$ in place of ϕ. Since δ is arbitrary the lemma is proven.

□

5.4 An Approximation for the Optimal Robust Disturbance Attenuation

In this section we examine the relationship between the proximity of the function χ and an approximate $\hat{\chi}$, and their corresponding smallest fixed-points. $\hat{\chi}$ is defined by $\hat{\chi} : [0,\infty) \to [0,\infty)$, $\hat{\chi}(r) := \|\Gamma_{M_r^* B^* w}\| + \|W^* V\|_\infty$, where M_r^* is defined as for M^* but with V replaced by rV. Theorem 5.1 allows us to conclude that $|\chi(r) - \hat{\chi}(r)| \leq \|W^* V\|_\infty$ for $r \in [0,1]$. Lemma 5.4 below demonstrates the existence of fixed-points for χ and $\hat{\chi}$ and establishes the continuity of χ and $\hat{\chi}$.

Lemma 5.4 *If W is normalized such that $\|W\|_\infty = 1$,*

1.) χ is a continuous, non-decreasing function on $[0,\infty)$ for which there exists a smallest fixed-point, which lies in the interval $[\chi(0), \chi(1)]$.

2.) If $W(e^{i\theta})$ and $V(e^{i\theta})$ are commensurate and uniformly Lipschitz for $\theta \in [0,2\pi)$, and A1 holds, then $\hat{\chi}(r) := \|\Gamma_{M_r^* B^* w}\| + \|W^* V\|_\infty$ is a continuous function of r on $[0,\infty)$, for which there exists a smallest fixed-point.

Proof. 1.) Continuity, the non-decreasing property of $\chi(\cdot)$, and the existence of a smallest fixed-point all follow from Appendix A.

Let r_f be any fixed-point of χ. Since $\chi(r_f) \in [0,1]$ and $\chi(r_f) = r_f$ it follows that $r_f \in [0,1]$. The non-decreasing property of χ, therefore, implies that $r_f = \chi(r_f) \leq \chi(1)$ and $r_f = \chi(r_f) \geq \chi(0)$.

2.) In order to show that $\hat{\chi}(\cdot)$ is continuous, it is sufficient to show that WM_r^* is continuous as a function of r in the metric of $L^\infty[0,2\pi)$. Fix $\frac{\pi}{4} > \delta > 0$ and define $\Phi(e^{i\theta}, e^{it}, x) := \left(1 + x\frac{|V|}{|W|+r|V|}(e^{i\theta})\right)^{-1}\left(1 + x\frac{|V|}{|W|+r|V|}(e^{it})\right)$.

$$
\begin{aligned}
&|(WM_{r+x}^* - WM_r^*)(e^{i\theta})| \\
&\leq |W(e^{i\theta})| \left| \fint_0^\pi K(\theta,t) log\left(1 + x\frac{|V|}{|W|+r|V|}(e^{it})\right) dt \right| && (204)\\
&= |W(e^{i\theta})| \left| \fint_0^\pi K(\theta,t) log\Phi(e^{i\theta}, e^{it}, x) dt \right| && (205)\\
&\leq |W(e^{i\theta})| \left| \int_{[0,\pi)-[\theta-\delta,\theta+\delta]} K(\theta,t) log\Phi(e^{i\theta}, e^{it}, x) dt \right| \\
&\quad + |W(e^{i\theta})| \left| \fint_{(\theta-\delta,\theta+\delta)} K(\theta,t) log\Phi(e^{i\theta}, e^{it}, x) dt \right| && (206)
\end{aligned}
$$

The first integral term of (206) converges to zero uniformly for $\theta \in [0,2\pi)$ as $x \to 0$, since the integrand converges to zero uniformly for t and θ in $\mathbf{W}_\delta := \{(\theta,t) \in T \times T \ : \ |\theta - t| > \delta\}$. $K(\theta,t)$ has the property that $|K(\theta,t)| \leq \frac{2\pi}{|t-\theta|}$ for $|t-\theta| < \frac{\pi}{4}$, hence the second integral term of (206) is bounded above by $Cx\delta|W(e^{i\theta})|$, for some constant C which is finite and depends on the Lipschitz constants of W and V. Thus the second integral term of (206) converges to zero uniformly for $\theta \in [0,2\pi)$ as $x \to 0$. Thus the continuity of M_r^*W as an L^∞ valued function of r is proven. Existence of a smallest fixed point for $\hat{\chi}(\cdot)$ follows from an identical argument to that used to prove the lemma of Appendix A.

□

The following theorem combines the estimate of the slope of χ of Lemma 5.3 with the approximations of 5.1, to obtain an implicit estimate for μ_{opt}, based on finding the smallest fixed-point of $\hat{\chi}$.

Theorem 5.2 *Under assumptions A1 and A4, if* $\|W^{*}V\|_{\infty} \leq \frac{\mu_{opt}^2}{16}$ *then the optimal robust disturbance attenuation for* $\mathcal{B}(P_0, V)$ *satisfies,*

$$\hat{r}_0 - \frac{10}{\hat{r}_0}\|W^{*}V\|_{\infty} \leq \mu_{opt} \leq \hat{r}_0 \tag{207}$$

where $\hat{r}_0$ *is the smallest fixed-point of* $\|\Gamma_{M_r^{*}B^{*}W}\| + \|W^{*}V\|_{\infty}$ *as a function of* r.

Proof.

Upper bound. From Theorem 5.1, $\|\Gamma_{M_r^{*}B^{*}W}\| \leq \chi(r) \leq \|\Gamma_{M_r^{*}B^{*}W}\| + \|W^{*}V\|_{\infty} = \hat{\chi}(r)$. μ_{opt} is the smallest fixed-point of χ from Theorem 2.1. Because $\hat{r}_0$ is the smallest fixed-point of $\hat{\chi}$, $r < \mu_{opt} \Rightarrow \chi(r) > r \Rightarrow \hat{\chi}(r) > r$. Thus $\hat{r}_0 \geq \mu_{opt}$.

Lower bound. From Lemma 5.3, since $\|W^{*}V\|_{\infty} \leq \frac{\mu_{opt}^2}{16}$ and $\hat{r}_0 \geq \mu_{opt}$,

$$\chi(\hat{r}_0) - \chi(\mu_{opt}) \leq (\hat{r}_0 - \mu_{opt})\left(\frac{4 + \frac{\mu_{opt}}{2}}{4 + \mu_{opt}}\right) \tag{208}$$

But $\chi(\hat{r}_0) \geq \|\Gamma_{M_{\hat{r}_0}^{*}B^{*}W}\| = \hat{r}_0 - \|W^{*}V\|_{\infty}$ (since $\hat{r}_0$ is a fixed-point of $\hat{\chi}$). Combining this observation with (208) gives,

$$\hat{r}_0 - \epsilon - \mu_{opt} \leq (\hat{r}_0 - \mu_{opt})\left(\frac{4 + \frac{\mu_{opt}}{2}}{4 + \mu_{opt}}\right) \tag{209}$$

Rearranging this gives $\mu_{opt} \geq \hat{r}_0 - \frac{10}{\hat{r}_0}\epsilon$

□

5.5 An Upper Bound for the Growth of Optimal Robust Disturbance Attenuation with A Priori Uncertainty.

In this section we investigate the effect of the radius of a priori plant uncertainty on the ability of feedback to reject uncertainty in a weighted ball. Here we can derive explicit estimates, in contrast to the general statement of Theorem 3.7 of Sect. 3.5.3, by making use of assumption A4 and Lemma 5.3. To this end, we introduce a scaling of the multiplicative ball of uncertainty $\mathcal{B}(P_0, V)$ by a parameter λ such that it becomes $\mathcal{B}(P_0, \lambda V)$. The behaviour of the optimal robust disturbance attenuation for $\mathcal{B}(P_0, \lambda V)$, represented by $\mu_{opt}(\lambda)$, as a function of the uncertainty radius λ is then examined.

Theorem 5.3 *Under the assumptions A1 and A4 where* $\lambda \|W^* V\|_\infty \leq \frac{\mu_{opt}(\lambda)^2}{16\lambda}$ *and* $x < \frac{\lambda \mu_{opt}(\lambda)}{8+\mu_{opt}(\lambda)}$,

$$\mu_{opt}(\lambda + x) - \mu_{opt}(\lambda) \leq \left(\frac{\mu_{opt}(\lambda)}{\frac{\lambda\mu_{opt}(\lambda)}{8+\mu_{opt}(\lambda)} - x} \right) x \tag{210}$$

Remark.

The condition on x ensures that the term $\left(\frac{\mu_{opt}(\lambda)}{\frac{\lambda\mu_{opt}}{8+\mu_{opt}} - x} \right)$ is always finite and positive. For small x (i.e. local estimation of the slope of $\mu_{opt}(\cdot)$ this term is bounded above by $\frac{9}{\lambda}$).

Proof of Theorem 5.3 From Theorem 2.1, $\mu_{opt}(\lambda)$ is the smallest fixed-point of

$$\begin{aligned} \inf_{Q \in H^\infty_{n\times n}} \; ess \sup_{\theta \in [0,2\pi)} \sup_{\zeta \in \mathbb{C}^n,\ |\zeta| \leq 1} & \left(|(I - P_0 Q) W(e^{i\theta})\zeta| + \lambda r |P_0 Q V(e^{i\theta})\zeta| \right) && (211) \\ := \chi_\lambda(r) & && (212) \end{aligned}$$

as a function of r. Define $S := \frac{4+\frac{\mu_{opt}(\lambda)}{2}}{4+\mu_{opt}(\lambda)}$. Let $x > 0,\ y > 0$.

$$\begin{aligned} \chi_{\lambda+x}\left(\mu_{opt}(\lambda) + y\right) &= \chi_1(\lambda\mu_{opt}(\lambda) + \lambda y + x\mu_{opt}(\lambda) + xy) && (213) \\ &= \chi_\lambda \left(\mu_{opt}(\lambda) + y + \frac{x\mu_{opt}(\lambda)}{\lambda} + \frac{xy}{\lambda} \right) && (214) \\ &\leq \mu_{opt}(\lambda) + S\left(y + \frac{x\mu_{opt}(\lambda)}{\lambda} + \frac{xy}{\lambda} \right) && (215) \end{aligned}$$

The last inequality follows since $\lambda\|W^{*}V\|_{\infty} \leq \frac{\mu_{opt}(\lambda)}{16}$ by assumption and $\mu_{opt}(\lambda)$ is the smallest fixed-point of χ_{λ}, so Lemma 5.3 is applicable. Suppose

$$S(y + \frac{x\mu_{opt}(\lambda)}{\lambda} + \frac{xy}{\lambda}) < y \tag{216}$$

for all $y \geq y_0 > 0$. (215) then would imply that $y_0 + \mu_{opt}(\lambda)$ must be an upper bound to the smallest fixed-point of $\chi_{\lambda+x}$, or equivalently $\mu_{opt}(\lambda + x)$. Simply by rearrangement we can show that if $x < \lambda\left(\frac{1}{S} - 1\right)$ then (216) is satisfied for any y greater than $\frac{Sx\mu_{opt}(\lambda)}{\lambda(1-S)-Sx}$. Hence the Theorem is proven.

□

Chapter 6

The ORDAP for Time-Varying Linear Systems

In this chapter the ORDAP is posed in the more general setting of time-varying linear systems. The statement of the ORDAP is the same as that for the linear time-invariant case, except the set of plants is a sphere of bounded linear causal operators rather than a sphere in H^∞, and moreover feedback control laws are allowed to be time-varying. Under certain conditions, the time-varying version of the ORDAP is reduced to an implicit form of an operator-based two-disc problem taking the form

$$\inf_{Q\in B_c(l^2_+,l^2_+)} \sup_{\zeta\in l^2_+,\|\zeta\|_2\le 1} \left(\|W(I-P_0Q)\zeta\|_2 + r\|VP_0Q\zeta\|_2\right) \tag{217}$$

The two-disc optimization is then shown to be expressible as a distance minimization in a Banach space of bounded linear operators, and an equivalent predual maximization is derived. A corollary of this result establishes, under certain conditions, the existence of an optimal control law for the ORDAP in the time-varying case. In Sect. 6.5, this theory is applied to a comparison of the effectiveness of linear time-varying and linear time-invariant control laws for cases where the *nominal* plant P_0 is time-invariant. As hypothesised in [29], it is shown that time-varying feedback offers no advantage over time-invariant feedback for l^2 disturbance rejection in the presence of time-varying plant uncertainty in the l^2 induced norm.

6.1 Definitions and Notation for Chapter 6

It is necessary to introduce some additional notation solely for Chapter 6 in order to handle more general spaces of linear operators (some of this notation is borrowed from [8]). The analysis in this chapter makes use of operator theory rather than the analytic function theory used previously because there is no isometric isomorphism between spaces of such systems and the space H^∞.

$B(L_1, L_2)$ denotes the Banach space of bounded linear operators from a Banach space L_1 to a Banach space L_2, where the norm is the operator norm.

l^2_+ denotes the usual Hilbert space of forward one-sided sequences

l^2 denotes the Hilbert space of two sided sequences.

S_k for some integer k denotes the shift k steps ahead. If k is negative then there may be an implicit truncation. Note that the domain and co-domain of S_k are intentionally left unspecified, and will depend on the context.

P_k for some integer k denotes the forward truncation operator which sets all outputs after time k to zero. Again the domain and co-domain depend on the context.

Causal operators $\Phi \in B(L_1, L_2)$ are defined to be those which satisfy $P_n\Phi(I - P_n) = 0$ for all positive integers n, where L_1 and L_2 are Hilbert resolution spaces [16].

Strictly causal operators $\Phi \in B(L_1, L_2)$ are those which satisfy $P_{n+1}\Phi(I-P_n) = 0$ for all positive integers n, where L_1 and L_2 are Hilbert resolution spaces.

The subscripts $_{sc}$ and $_c$ denote the restriction of a subspace of operators to its intersection with causal and strictly causal operators respectively.

$< \cdot, \cdot >$ denotes either the inner product in l^2 or the binary operation of functional evaluation, depending on the context.

6.2 Formulation of the Problem.

The statement of the ORDAP we use for time-varying systems follows. Let $P_0 \in B_c(l^2, l^2)$ be a nominal, stable plant (possibly time-varying) and denote the set of plant uncertainty by

$$C(P_0, V) := \{P \in B_c(l^2, l^2) \ : \ P = XVP_0 + P_0, X \in B_c(l^2, l^2), \|X\| < 1\}$$

where V is a causal, linear, time-invariant weighting function, and P_0 is assumed to be strictly causal in order that the feedback loop is well-posed. The optimal robust disturbance attenuation problem (ORDAP) comprises finding the smallest weighted l^2_+ induced norm of the sensitivity operator achievable by a single causal feedback control law for all plants in $C(P_0, V)$. In mathematical language, the equivalent statement is:

$$\mu := \inf_{\substack{C_0 \text{ stabilizing} \\ \forall P \in C(P_0, V)}} \ \sup_{P \in C(P_0, V)} \|W(I + PC_0)^{-1}\| \tag{218}$$

From [57] we can express (218) in the form

$$\begin{aligned}\mu := \inf\Big\{ & \sup_{P \in C(P_0, V)} \|W(I - P_0Q)(I + \Delta PQ)^{-1}\| \ : \\ & Q \in B_c(l^2_+, l^2_+) \textit{ such that } (I + (P - P_0)Q)^{-1} \in B_c(l^2_+, l^2_+) \ \ \forall P \in C(P_0, V)\Big\},\end{aligned}$$

where ΔP denotes $P - P_0$ and W is a linear time-invariant weighting function.

If a particular robustly stabilizing feedback control law C achieves a 'worst case' weighted sensitivity induced norm from $l^2_+ \to l^2_+$ which is less than γ then,

$$\|W(I - P_0Q)(I + \Delta PQ)^{-1}\| \leq \gamma \textit{ and } (I + (P - P_0)Q)^{-1} \in B_c(l^2_+, l^2_+) \tag{219}$$
$$\forall P \in C(P_0, V)$$

(219) is equivalent to,

$$\|W(I-P_0Q)(I+\Delta PQ)^{-1}\zeta\|_2 \le \gamma\|\zeta\|_2 \ and\ (I+\Delta PQ)^{-1} \in B_c(l_+^2, l_+^2)$$
$$\forall\zeta \in l_+^2,\ \ \forall P \in \mathcal{C}(P_0, V) \tag{220}$$

Since $(I+\Delta PQ)$ has a bounded inverse in $B(l_+^2, l_+^2)$, we obtain

$$\|W(I-P_0Q)\eta\|_2 \le \gamma\|(I+\Delta PQ)\eta\|_2 \ \ and\ \ (I+\Delta PQ)^{-1} \in B_c(l_+^2, l_+^2)$$
$$\forall\eta \in l_+^2,\ \ \forall P \in \mathcal{C}(P_0, V) \tag{221}$$

which is equivalent to (220). Certainly (221) is implied by (Sect.4 Chapter 5 of [16])

$$\|W(I-P_0Q)\zeta\|_2 \le \gamma\|\zeta\|_2 - \gamma\|VP_0Q\zeta\|_2 \quad \forall\zeta \in l_+^2 \tag{222}$$

(222) is equivalent to

$$\|W(I-P_0Q)\zeta\|_2 + \gamma\|VP_0Q\zeta\|_2 \le \gamma \quad \forall\zeta \in l_+^2,\ \|\zeta\|_2 \le 1 \tag{223}$$

Therefore we conclude that, under the existence assumption of Theorem 6.1, the optimal robust disturbance attenuation μ is bounded above by the smallest positive fixed-point of

$$\chi(r) := \inf_{Q\in B_c(l_+^2, l_+^2)} \sup_{\zeta\in l_+^2, \|\zeta\|_2\le 1} (\|W(I-P_0Q)\zeta\|_2 + r\|VP_0Q\zeta\|_2)) \tag{224}$$

on the interval $[0,1]$. The existence of a smallest fixed-point follows from the fact that the function $\chi(\cdot)$ in (224) is a continuous, positive, non-decreasing function on $[0,1]$. This is proven by the same argument as that used in [20]. This reasoning has proven b) of Theorem 6.1

Theorem 6.1 *If there exists an optimal $Q \in B_c(l_+^2, l_+^2)$ for each value of $r \in (0,1]$ in the optimizations (226) below, then the following hold:*

a) if P_0 is time-invariant and

$$\mu := \inf \left\{ \sup_{P \in C(P_0, V)} \|W(I - P_0 Q)(I + \Delta P Q)^{-1}\| \ : \ Q \in B_c(l_+^2, l_+^2) \right.$$
$$\left. \text{such that } (I + (P - P_0)Q)^{-1} \in B_c(l_+^2, l_+^2) \ \forall P \in C(P_0, V) \right\}, \qquad (225)$$

then μ is the smallest positive fixed-point of

$$\chi(r) := \inf_{Q \in B_c(l_+^2, l_+^2)} \sup_{\zeta \in l_+^2, \|\zeta\|_2 \le 1} (\|W(I - P_0 Q)\zeta\|_2 + r\|V P_0 Q \zeta\|_2) \qquad (226)$$

b) if P_0 is time-varying then μ is bounded above by the smallest positive fixed-point of χ.

Remarks.

1) As in Theorem 2.1, the existence of an optimal Q for the optimization of (226) will be established under quite general conditions in Theorem 6.2.

2) Theorem 6.1 implies that the ORDAP for time-invariant nominal plants with possibly time-varying perturbations and control laws, reduces to evaluation of the following type of optimization

$$\mu_0 := \inf_{Q \in B_c(l_+^2, l_+^2)} \sup_{\zeta \in l_+^2, \|\zeta\|_2 \le 1} (\|W(I - P_0 Q)\zeta\|_2 + \|V P_0 Q \zeta\|_2) \qquad (227)$$

Where W, V, P_0 are linear causal and time-invariant.

Proof of Theorem 6.1a). Let $\tilde{Q} \in B_c(l_+^2, l_+^2)$ be such that

$$\|W(I - P_0 \tilde{Q})(I + XVP_0\tilde{Q})^{-1}\| \le \mu' \text{ and } (I + XVP_0\tilde{Q})^{-1} \in B_c(l_+^2, l_+^2)$$
$$\forall X \in B_c(l_+^2, l_+^2), \ \|X\| \le 1 \qquad (228)$$

where $\mu' > \mu$.

We divide the proof of part (a) of this theorem into two claims.

Claim 1: $\lim_{n\to\infty} \sup_{\substack{\zeta \in l_2^+ \\ \|\zeta\|_2 \leq 1}} \|VP_0\tilde{Q}S_n\zeta\|_2 \leq 1$

To prove claim 1 suppose that, on the contrary, there exists an $\epsilon > 0$ such that

$$\sup_{\substack{\zeta \in l_2^+ \\ \|\zeta\|_2 \leq 1}} \|VP_0\tilde{Q}S_n\zeta\|_2 > 1 + \epsilon \quad \textit{for all positive integers } n > 0 \tag{229}$$

Fix $\delta > 0$. Consider the following construction of a vector $\eta \in l_+^2$.

step 1. Define an integer $n_0 := 0$. Select $\zeta_1 \in l_+^2$, $\|\zeta_1\|_2 \leq 1$ which has only finitely many non-zero elements and satisfies (229) for $\zeta = \zeta_1$ and $n = 0$. Since $VP_0\tilde{Q}\zeta_1 \in l_+^2$ there exists an integer n_1 such that

$$\|(I - P_{n_1})VP_0\tilde{Q}\zeta_1\|_2 < \delta \;\; \textit{and} \;\; n_1 > \textit{length}(\zeta_1). \tag{230}$$

step k. Select $\zeta_k \in l_+^2$ with only finitely many non-zero elements such that $\|\zeta_k\|_2 \leq 1$, $P_{n_{k-1}}\zeta_k = 0$, ζ_k satisfies (229) for $\zeta = \zeta_k$ and $n = 0$. Define n_k to be an integer such that

$$\|(I - P_{n_k})VP_0\tilde{Q}\zeta_k\|_2 < \delta \;\; \textit{and} \;\; n_k > \textit{length}(\zeta_k) \tag{231}$$

Note that $\textit{length}(\cdot)$ denotes the smallest integer such that all entries of the argument are zero at positions with greater index.

Define $\eta := \frac{1}{N}\sum_{k=1}^{N} \zeta_k$. If δ is chosen to be sufficiently small w.r.t. the quantities $\|VP_0\tilde{Q}\|$ and ϵ we obtain

$$\|VP_0\tilde{Q}\eta\|_2 > 1 + \frac{\epsilon}{2} \tag{232}$$

Define $X \in B_c(l^2_+, l^2_+)$ to be the following strictly contractive, causal, compact, linear operator from $l^2_+ \to l^2_+$,

$$Xu := \sum_{k=1}^{N-1} \lambda_k < \nu_k, u > \zeta_{k+1}, \tag{233}$$

where $\nu_k := \frac{(P_{n_k} - P_{n_{k-1}})VP_0\tilde{Q}\zeta_k}{\|(P_{n_k} - P_{n_{k-1}})VP_0\tilde{Q}\zeta_k\|_2}$ and $\lambda_k := \frac{1}{\|VP_0\tilde{Q}\zeta_k\|_2} \leq \frac{1}{1+\epsilon}$. Note that X maps the vector $(P_{n_k} - P_{n_{k-1}})VP_0\tilde{Q}\zeta_k$ to $\frac{\|P_{n_k} - P_{n_{k-1}})VP_0\tilde{Q}\zeta_k\|_2}{\|VP_0\tilde{Q}\zeta_k\|_2}\zeta_{k+1}$ for $1 \leq k \leq N-1$, enabling the deduction of inequality (235).

$$\begin{aligned}
\|(I - XVP_0\tilde{Q})\eta\|_2 &\leq \left\|\frac{\zeta_1}{N}\right\|_2 + \left\|\frac{XVP_0\tilde{Q}\zeta_N}{N}\right\|_2 + \sum_{k=2}^{N}\left\|\frac{\zeta_k - XVP_0\tilde{Q}\zeta_{k-1}}{N}\right\|_2 && (234)\\
&\leq \frac{1 + \|VP_0\tilde{Q}\|}{N} + 2\delta && (235)
\end{aligned}$$

Thus, if δ was chosen sufficiently small, for large enough N, (235) contradicts (232) and so claim 1 is proven.

claim 2:

$$lim_{n\to\infty} \sup_{\zeta\in l^2_+, \|\zeta\|_2\leq 1} \left(\|W(I - P_0\tilde{Q})S_n\zeta\|_2 + \mu'\|VP_0\tilde{Q}S_n\zeta\|_2\right) \leq \mu'$$

We will also prove claim 2 by contradiction. Suppose, on the contrary, that claim 2 is false, then there exists an $\epsilon > 0$ such that

$$\sup_{\zeta\in l^2_+, \|\zeta\|_2\leq 1} \left(\|W(I - P_0\tilde{Q})S_n\zeta\|_2 + \mu'\|VP_0\tilde{Q}S_n\zeta\|_2\right) > \mu' + \epsilon \tag{236}$$

for all positive integers n. If (236) holds then we follow a very similar procedure to that used to construct the vector η in the proof of claim 1. Note that all the variables are reset.

step 1. Define $n_0 = 0$. Select a $\zeta_1 \in l^2_+, \|\zeta_1\|_2 \leq 1$ with finitely many non-zero elements such that (236) is satisfied for $\zeta = \zeta_1$, $n = 0$ and $\|VP_0\tilde{Q}\zeta_1\|_2 < 1 + \delta$

(c.f. claim 1). There exists an integer n_1 such that $\|(I-P_{n_1})W(I-P_0\tilde{Q})\zeta_1\|_2 \leq \delta$, $\|(I-P_{n_1})VP_0\tilde{Q}\zeta_1\|_2 \leq \delta$, and $n_1 > length(\zeta_1)$.

step k. Select a $\zeta_k \in l_+^2$, $\|\zeta_k\|_2 \leq 1$, $P_{n_{k-1}}\zeta_k = 0$ with finitely many non-zero elements such that (236) is satisfied for $\zeta = \zeta_k$, $n = 0$, and $\|VP_0\tilde{Q}\zeta_k\|_2 < 1+\delta$. The existence of such a vector follows from (236) and claim 1. There exists an integer n_k such that $\|(I-P_{n_k})W(I-P_0\tilde{Q})\zeta_k\|_2 \leq \delta$, $\|(I-P_{n_k})VP_0\tilde{Q}\zeta_k\|_2 \leq \delta$ and $n_k > length(\zeta_k)$.

Chose $\eta = \frac{1}{N}\sum_{k=1}^{N}\zeta_k$. For sufficiently small δ we obtain

$$\|W(I-P_0\tilde{Q})\eta\|_2 + \mu'\|VP_0\tilde{Q}\eta\|_2 \geq \mu' + \frac{\epsilon}{2} \tag{237}$$

Define $X \in B_c(l^2, l^2)$ to be the following causal, compact, strictly contractive linear operator

$$Xu := \frac{1}{1+\delta}\sum_{k=1}^{N-1} < \nu_k, u > \zeta_{k+1} \quad where\ \nu_k := \frac{(P_{n_k}-P_{n_{k-1}})VP_0\tilde{Q}\zeta_k}{\|(P_{n_k}-P_{n_{k-1}})VP_0\tilde{Q}\zeta_k\|_2}$$

In the event that the denominator and numerator of the expression for ν_k are both zero, ν_k is set equal to zero. Note that X maps $(P_{n_k}-P_{n_{k-1}})VP_0\tilde{Q}\zeta_k \to \frac{1}{1+\delta}\|(P_{n_k} - P_{n_{k-1}})VP_0\tilde{Q}\zeta_k\|_2\zeta_{k+1}$ for $1 \leq k \leq N-1$. Hence

$$\begin{aligned}\|(I-XVP_0\tilde{Q})\eta\|_2 &\leq \|\frac{\zeta_1}{N}\|_2 + \|\frac{XVP_0\tilde{Q}\zeta_N}{N}\|_2 + \sum_{k=2}^{N}\|\frac{\zeta_k - XVP_0\tilde{Q}\zeta_{k-1}}{N}\|_2 \\ &\leq \delta + \frac{2}{N} + \sum_{k=2}^{N}\|\frac{\zeta_k - X(P_{n_{k-1}}-P_{n_{k-2}})VP_0\tilde{Q}\zeta_{k-1}}{N}\|_2 \end{aligned} \tag{238}$$

Each summand in (238) is bounded above by
$\frac{1}{N}\left(1 - \|(P_{n_{k-1}}-P_{n_{k-2}})VP_0\tilde{Q}\zeta_{k-1}\|_2 + 3\delta\right)$ since $\|VP_0\tilde{Q}\zeta_{k-1}\|_2 < 1+\delta$. For sufficiently small δ, sufficiently large N,

$$\|(I-XVP_0\tilde{Q})\eta\|_2 \leq \frac{\epsilon}{3} + 1 - \|VP_0\tilde{Q}\eta\|_2$$

Thus $\|W(I - P_0\tilde{\zeta})\eta\|_2 \le \mu'(1+\frac{\epsilon}{3}) - \mu'\|VP_0\tilde{Q}\eta\|_2$ from (221). This contradicts (237) and so claim 2 must be true.

Claim 2 implies that (239) holds below.

$$\Rightarrow lim_{n\to\infty} \sup_{\zeta\in l^2_+, \|\zeta\|_2\le 1} \|W(I - P_0S_{-n}\tilde{Q}S_n)\zeta\|_2 + \mu'\|VP_0S_{-n}\tilde{Q}S_n\zeta\|_2 \le \mu'$$
$$\textit{since } P_0, W, V \textit{ are time invariant and causal} \tag{239}$$

$$\Rightarrow \inf_{Q\in B_c(l^2_+, l^2_+)} \sup_{\zeta\in l^2_+, \|\zeta\|_2\le 1} (\|W(I - P_0Q)\zeta\|_2 + \mu'\|VP_0Q\zeta\|_2) \le \mu' \tag{240}$$
$$\textit{since } S_{-n}\tilde{Q}S_n \textit{ is causal.}$$

The rest of the proof is conceptually identical to the proof of Theorem 2.1 b), and so will not be repeated in this case.

□

Remark. The proof of Theorem 6.1 also shows that the conclusion of a) holds for time-varying nominal plants P_0 which satisfy the following condition.

$$\inf_{Q\in B_c(l^2_+, l^2_+)} \sup_{\substack{\zeta\in l^2_+ \\ \|\zeta\|_2\le 1}} (\|W(I - P_0Q)S_n\zeta\|_2 + r\|VP_0QS_n\zeta\|_2)$$
$$\textit{is independent of } n \textit{ for all } r\in[0,1]. \tag{241}$$

(241) is satisfied, for example, by any periodically time-varying nominal system P_0.

6.3 A Distance Problem in Banach Space

Here we show that (227) can be expressed as a distance minimization in a Banach space of linear operators. To this end let Z be the Banach space $l^2_+ \times l^2_+$ under the norm defined by

$$\left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_Z := \|x\|_2 + \|y\|_2, \quad x, y \in l^2_+$$

Since $W(I-P_0O)$ and VP_0Q are bounded operators from $l_+^2 \rightarrow l_+^2$, $\begin{bmatrix} W(I-P_0Q) \\ VP_0Q \end{bmatrix}$ represents a bounded operator from $l_+^2 \rightarrow Z$. The induced operator norm is

$$\left\| \begin{bmatrix} W(I-P_0Q) \\ VP_0Q \end{bmatrix} \right\|_L = \sup_{\zeta \in l_+^2, \|\zeta\|_2 \leq 1} \left\| \begin{bmatrix} W(I-P_0Q) \\ VP_0Q \end{bmatrix} \zeta \right\|_Z \tag{242}$$
$$= \sup_{\zeta \in l_+^2, \|\zeta\|_2 \leq 1} (\|W(I-P_0Q)\zeta\|_2 + \|VP_0Q\zeta\|_2)$$

where the subscript L denotes the induced operator norm *of* $B(l_+^2, Z)$. Thus we can express μ_0 in the form

$$\mu_0 = \inf_{Q \in B_c(l_+^2, l_+^2)} \left\| \begin{bmatrix} W \\ 0 \end{bmatrix} - \begin{bmatrix} W \\ V \end{bmatrix} P_0 Q \right\|_L$$

As in Chapter 3, we assume that

(A1) $|W(e^{i\theta})|^2 + |V(e^{i\theta})|^2 \geq \epsilon_0 > 0 \qquad \forall\theta \in [0, 2\pi)$

Under this condition there exists a function $\Lambda \in H^\infty$, invertible in H^∞, such that $|\Lambda(e^{i\theta})|^2 = |W(e^{i\theta})|^2 + |V(e^{i\theta})|^2$. Next we assume (c.f. [15])

(A5) ΛP_0 *as an operator in* $B_c(l^2, l^2)$ *has a factorization* UH *where* $H^{\pm 1} \in B_c(l^2, l^2)$ *and* U *is a causal unitary operator from* $l^2 \rightarrow l^2$. [8]

(A1) and (A5) allow us to express μ_0 as a distance problem in the space of Z-valued operators on l_+^2,

$$\mu_0 = \inf_{Q \in B_c(l_+^2, l_+^2)} \left\| \begin{bmatrix} W \\ 0 \end{bmatrix} - RUQ \right\|_L \tag{243}$$

where $R := \begin{bmatrix} W \\ V \end{bmatrix} \Lambda^{-1}$.

[8]The restriction of U to l_+^2 results in the more familiar partial isometry.

6.4 Formulation in the Predual

Throughout the rest of this chapter, we define X to be the Hilbert space l_+^2 and Y to be the Banach space $l_+^2 \times l_+^2$ under the norm

$$\left\| \begin{bmatrix} x \\ y \end{bmatrix} \right\|_Y := Max\left(\|x\|_2, \|y\|_2\right), \quad x, y \in l_+^2$$

Lemma 6.1

$$i. \quad X^* \simeq l_+^2,\ X^{**} \simeq l_+^2$$

$$ii. \quad Y^* \simeq Z,\ Y^{**} \simeq Y$$

Remark. If we take finite spaces of n-tuples rather than the spaces of infinite sequences considered here, then the same relations can be obtained with a conceptually identical proof.

Proof. i. is standard, since X is a Hilbert space.

ii. ϕ is a bounded linear functional on Y iff ϕ has the representation,

$$\phi\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) = \phi_1(x) + \phi_2(y)$$

where ϕ_1, ϕ_2 are bounded linear functionals on the Hilbert space l_+^2. Thus $\phi\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) =< \alpha, x >_{l_+^2} + < \beta, y >_{l_+^2}$ for some $\alpha, \beta \in l_+^2$.

$$\begin{aligned} \left|\phi\left(\begin{bmatrix} x \\ y \end{bmatrix}\right)\right| &\leq \|\alpha\|_2 \cdot \|x\|_2 + \|\beta\|_2 \cdot \|y\|_2 \\ &\leq \left(\|\alpha\|_2 + \|\beta\|_2\right) Max\left(\|x\|_2, \|y\|_2\right) \end{aligned} \tag{244}$$

Equality in (244) can be achieved by chosing $x = \frac{\alpha}{\|\alpha\|_2}$, $y = \frac{\beta}{\|\beta\|_2}$, if $\alpha \neq 0,\ \beta \neq 0$, $x = 0$ if $\alpha = 0$ and $y = 0$ if $\beta = 0$. Thus $\|\phi\| = \|\alpha\|_2 + \|\beta\|_2$ where the former norm is the induced norm of the functional ϕ. Thus we have established the first isometric isomorphism of ii. In a similar manner, if ϕ is a bounded linear functional on Y^*, and $\begin{bmatrix} x \\ y \end{bmatrix} \in Y^*$ with $\begin{bmatrix} \tilde{x} \\ \tilde{y} \end{bmatrix}$ as its equivalent in Z,

$$\begin{aligned} \phi\left(\begin{bmatrix} x \\ y \end{bmatrix}\right) &= \phi_1(\tilde{x}) + \phi_2(\tilde{y}) \quad \textit{where } \phi_1, \phi_2 \textit{ are bounded linear functionals on } l_+^2 \\ &= < v, \tilde{x} >_{l_+^2} + < \theta, \tilde{y} >_{l_+^2} \end{aligned}$$

where $v, \theta \in l_+^2$ represent the functionals ϕ_1, ϕ_2. Thus

$$\left|\phi\begin{pmatrix} x \\ y \end{pmatrix}\right| \le (\|\tilde{x}\|_2 + \|\tilde{y}\|_2)\, Max\,(\|v\|_2, \|\theta\|_2) \tag{245}$$

Equality is achieved in (245) by setting either $\tilde{x} = \frac{v}{\|v\|_2}, \tilde{y} = 0$, or $\tilde{x} = 0, \tilde{y} = \frac{\theta}{\|\theta\|_2}$ depending on whether $\|v\|_2 \ge \|\theta\|_2$ or $\|v\|_2 < \|\theta\|_2$. Thus the isometric isomorphism for the second relation of ii. is established.

□

Definition 6.1 (c.f. [8]) *$N(B_1, B_2)$ is defined to be the Banach space of nuclear operators mapping the Banach space B_1 to the Banach space B_2 under the nuclear norm.* [9] *An operator $\Phi : B_1 \to B_2$ is said to be nuclear if it has the representation*

$$\Phi b = \sum_n < b_n^*, b > \omega_n, \quad \textit{where } \omega_n \in B_2,\ b_n^* \in B_1^* \tag{246}$$

$$\textit{and} \quad \sum_n \|b_n^*\| \cdot \|\omega_n\| < \infty \tag{247}$$

The nuclear norm $\|\Phi\|_{nuc}$ is defined to be the infimum over all sums in (247) corresponding to representations (246).

Definition 6.2 (c.f. [8]) *Tr (the trace) is defined to be the following linear functional on $N(B_1, B_1)$*

$$Tr\Phi := \sum_n < b_n^*, \omega_n > \tag{248}$$

[9] $N(B_1, B_2)$ was shown to be a Banach space in [22]

for Φ as in (246) (N.B. $B_1 = B_2$). Note that (248) is well defined since the quantity on the right of (248) is independent of the particular representation chosen for Φ.

Lemma 6.1 ii. has established that the L norm defined in (242) is the induced operator norm of the Banach space $B(X^*, Y^*)$. Therefore $B(X^*, Y^*)$ is the appropriate Banach space in which to represent the distance problem (243).

The following key lemma applies to the Banach spaces X and Y.

Lemma 6.2 (Diestel and Uhl [10] also see [8])

$$B(X^*, Y^*) \simeq N(X, Y)^*$$

The isometric isomorphism $\simeq$ is generated by the representation

$$\phi(T) = < B, T > = TrT^*B = TrBT^*, \quad where\ T \in N(X, Y),\ B \in B(X^*, Y^*)$$

possessed by any bounded linear functional ϕ on $N(X, Y)$.

If we are to represent (243) as a maximal problem in the predual space ($N(X, Y)$) following Chapter 3, we need to identify a subspace S^0 of $N(X, Y)$ such that $(S^0)^\perp \simeq RUB_c(X^*, X^*)$, where $\simeq$ is the isomorphism of Lemma 6.2. To this end, in lemma 6.3 below, the causal operators in $B_c(l^2, l^2)$ are characterized as the annihilators of the adjoints of a space of strictly causal operators.

Lemma 6.3 *If $\Phi \in B(l^2, l^2)$ the following are equivalent,*

$$i. \quad Tr T\Phi = 0 \quad \forall T \in N_{sc}(l^2, l^2) \tag{249}$$

$$ii. \quad \Phi \in B_c(l^2, l^2) \tag{250}$$

Proof. ($i. \Rightarrow ii.$) Define, for any integer $M \geq 0$, $R_M \in N(l^2, l^2)$ to be the operator $(I - P_M)\Theta P_M$, where Θ is an arbitrary operator in $N(l^2, l^2)$. From the projection property of P_M, we conclude that R_M is strictly causal. From (249) we have,

$$\begin{aligned} Tr\Phi(I - P_M)\Theta P_M &= 0 \quad \forall \Theta \in N(l^2, l^2) \\ \Rightarrow Tr P_M \Phi(I - P_M)\Theta &= 0 \quad \forall \Theta \in N(l^2, l^2) \\ \Rightarrow P_M \Phi(I - P_M) &= 0 \end{aligned}$$

Since M is arbitrary, $\Phi \in B_c(l^2, l^2)$.

($ii. \Rightarrow i.$) It is necessary only to show that if $\Phi \in B_c(l^2, l^2)$ and $T \in N_{sc}(l^2, l^2)$ then $TrT\Phi = 0$. Under the conditions of ii $\Phi T \in N_{sc}(l^2, l^2)$ (c.f. property 2.5 [8]). Thus, since ΦT is nuclear, for $v \in l^2$, $\Phi T v = \sum_n < x_n^{**}, v > x_n^*$ for some $x_n^{**} \in l^2$ and $x_n^* \in l^2$. Hence, we can write

$$(\Phi T e_k)_k = \sum_n \overline{(x_n^{**})_k}(x_n^*)_k \tag{251}$$

where e_k is the k-th standard basis vector of l^2 and $(\cdot)_k$ denotes the k-th component of a vector in l^2. From the strict causality of ΦT, $(\Phi T e_k)_k = 0$. Thus, we can express

$$\begin{aligned} Tr\Phi T &= \sum_n < x_n^{**}, x_n^* > = \sum_n \sum_{k=1}^{\infty} \overline{(x_n^{**})_k}(x_n^*)_k \\ &= \sum_{k=1}^{\infty} \sum_n \overline{(x_n^{**})_k}(x_n^*)_k \quad \textit{(by Fubini's theorem and the nuclearity of } T\textit{)} \\ &= \sum_{k=1}^{\infty} (\Phi T e_k)_k = 0 \end{aligned}$$

□

Note that in what follows, we shall suppress the distinction between isometrically isomorphic Banach spaces in order to simplify the presentation.

Define the following subspace of $N(Y^*, X^*)$,

$$S_0 := \{\mathcal{P}\,(\Pi(I - RR^*) + KU^*R^*)\,|_{Y^*} \;:\; \Pi \in N(Z_\sim, l^2), K \in N_{sc}(l^2, l^2)\}$$

where $\mathcal{P} : l^2 \to l^2_+$ denotes the canonical projection, and $Z_\sim$ is the Banach space $l^2 \times l^2$ under the Z norm. Note that $(I - RR^*)$, R^* are multiplication operators and * denotes the involution operation. Define S^0 to be the following subspace of $N(X, Y)$

$$S^0 := \{T \in N(X, Y) \;:\; T^* \in S_0\}$$

Lemma 6.4 proves that $(S^0)^\perp \simeq RUB_c(X^*, X^*)$ which establishes S^0 as the desired pre-orthogonal complement.

Lemma 6.4 *If $\Phi \in B(X^*, Y^*)$ then*

$$< \Phi, T >= 0 \;\; for\ all\ T \in S^0 \Leftrightarrow \Phi \in RUB_c(X^*, X^*) \tag{252}$$

Proof. The reflexivity of the Banach spaces X and Y (Lemma 6.1) and the representation of linear functionals in terms of inner products on l^2_+ imply that any element of $N(Y^*, X^*)$ is the adjoint of some element of $N(X, Y)$. Thus the LHS of (252) is equivalent to

$$Tr\,(\mathcal{P}\Pi(I - RR^*)|_{Y^*}\Phi + \mathcal{P}KU^*R^*|_{Y^*}\Phi) = 0 \quad \forall \Pi \in N(Z_\sim, l^2),\;\; K \in N_{sc}(l^2, l^2)$$

$$\Leftrightarrow (I - RR^*)|_{Y^*}\Phi\mathcal{P} = 0 \;\; and\;\; U^*R^*|_{Y^*}\Phi\mathcal{P} \in B_c(l^2, l^2) \;\; (from\ Lemma\ 6.3)$$

$$\Leftrightarrow (I - RR^*)|_{Y^*}\Phi\mathcal{P} = 0 \;\; and\;\; RR^*|_{Y^*}\Phi\mathcal{P} \in RUB_c(l^2, l^2)$$

$$\Leftrightarrow \Phi \in RUB_c(X^*, X^*)$$

□

The characterization of the predual of $B(X^*, Y^*)$ and the pre-orthogonal complement of $RUB_c(X^*, X^*)$ has demonstrated the following result.

Theorem 6.2 *Under assumption (A1),*

$$Min_{Q \in B_c(l^2_+, l^2_+)} \left\| \begin{bmatrix} W \\ 0 \end{bmatrix} - RUQ \right\|_L = \sup_{T \in S^0, \|T\|_{N(X,Y)} \leq 1} \left| TrT^* \begin{bmatrix} W \\ 0 \end{bmatrix} \right| \tag{253}$$

Remarks. Implicit in the statement of Theorem 6.2 is the existence of a linear time-varying control law for the optimization (243). Theorem 6.2 also establishes that under assumption (A5) optimal feedback laws exist both for the optimizations (224) in general, and for the time-varying ORDAP, when P_0 is time-invariant or satisfies (241).

6.5 Linear Time-Varying vs Linear Time-Invariant Control Laws

In this section, we use the theory of Sect. 6.4 to compare linear time-invariant and linear time-varying control laws for the ORDAP, when the nominal plant P_0 is constrained to be time-invariant. Under this condition, the operator U in (243) becomes a multiplication operator with a symbol equal to an inner function in H^∞. If the controllers are restricted to being linear and time-invariant, then (243) becomes

$$\mu_0^{ti} = \inf_{Q \in B_c^{ti}(l^2_+, l^2_+)} \left\| \begin{bmatrix} W \\ 0 \end{bmatrix} - RUQ \right\|_L \tag{254}$$

where $B_c^{ti}(l^2, l^2)$ is the subspace of time invariant operators in $B_c(l^2, l^2)$.

Lemma 6.5 *If P_0 is linear, time-invariant and causal, and if assumption (A1) holds, then*

$$\mu_0 = \mu_0^{ti}$$

Moreover at least one of the optimal feedback laws must be time-invariant

Proof. (Applying an idea of Shamma and Dahleh [44]). Let $Q \in B_c(l^2, l^2)$ be any minimal Q in (253). Define

$$Q_n := \frac{1}{n+1} \sum_{k=0}^{n} S_{-k} Q S_k$$

From the definition of the shift and the L norm, coupled with the fact that RU is time-invariant and causal we get

$$\begin{aligned} \mu_0 \geq \left\| S_{-k} \left(\begin{bmatrix} W \\ 0 \end{bmatrix} - RUQ \right) S_k \right\|_L &= \left\| \begin{bmatrix} W \\ 0 \end{bmatrix} - RUS_{-k}QS_k \right\|_L \\ \Rightarrow \left\| \begin{bmatrix} W \\ 0 \end{bmatrix} - RUQ_n \right\|_L \leq \mu_0 & \qquad (255) \end{aligned}$$

As in [44] we have $\|Q_n\| \leq \|Q\| \quad \forall n$ where $\|\cdot\|$ is the operator norm of $B(X^*, X^*)$. Thus $\|RUQ_n\|_L$ is uniformly bounded. Since $B(X^*, Y^*)$ has a separable predual i.e. $N(X, Y)$ (see [8]), there exists a subsequence of the integers $\{n_k\}$ such that

$$RUQ_{n_k} \underset{wk^*}{\rightarrow} J \ (\in B(X^*, Y^*)). \qquad (256)$$

This follows from Alaoglu's theorem which asserts the wk^* compactness of the unit ball in such a Banach space. (256) implies $\begin{bmatrix} W \\ 0 \end{bmatrix} - RUQ_{n_k} \underset{wk^*}{\rightarrow} \begin{bmatrix} W \\ 0 \end{bmatrix} - J$ from which we conclude that

$$\begin{aligned} \mu_0 &\geq \liminf_{k \to \infty} \left\| \begin{bmatrix} W \\ 0 \end{bmatrix} - RUQ_{n_k} \right\|_L \\ &\geq \left\| \begin{bmatrix} W \\ 0 \end{bmatrix} - J \right\|_L \qquad (257) \end{aligned}$$

the first inequality follows from (255), the second from Alaoglu's Theorem. Next we claim that J has the form $J = RUG$ where $G \in B_c(l^2_+, l^2_+)$. To prove this, suppose T is an arbitrary element of S^0. From Lemma 6.4 we have $< RUQ_{n_k}, T >= 0$, and hence from (256) we obtain $< J, T >= 0$. Thus from Lemma 6.4, J has the conjectured form.

The lemma will be proven if we can show that RUG is time-invariant. Let T be an arbitrary operator in $N(X, Y)$, then

$$\begin{aligned} < SRUQ_{n_k}, T > &= TrT^*SRUQ_{n_k} = Tr(S_{-1}T)^*RUQ_{n_k} \\ &= < RUQ_{n_k}, S_{-1}T > \rightarrow < RUG, S_{-1}T > \quad as\ k \rightarrow \infty \\ but \quad < RUG, S_{-1}T > &= TrT^*SRUG = < SRUG, T > \end{aligned}$$

$$Thus, \qquad SRUQ_{n_k} \underset{wk^*}{\rightarrow} SRUG \quad since\ T \in N(X, Y)\ was\ arbitrary \tag{258}$$

S_{-1} denotes the backward shift with truncation, mapping $Y \rightarrow Y$. A similar argument shows that

$$RUQ_{n_k}S \underset{wk^*}{\rightarrow} RUGS \tag{259}$$

From [44], $\|SQ_{n_k} - Q_{n_k}S\| \rightarrow 0$ as $k \rightarrow \infty$, from which we deduce $\|SRUQ_{n_k} - RUQ_{n_k}S\|_L \rightarrow 0$ as $k \rightarrow \infty$, since RU is time-invariant. Thus, for arbitrary $T \in N(X, Y)$

$$| < T, SRUQ_{n_k} > - < T, RUQ_{n_k}S > | \rightarrow 0 \quad as\ k \rightarrow \infty$$

Since $T \in N(X, Y)$ is arbitrary, we conclude from (258) and (259) that

$$SRUG = RUGS$$

It follows that $J = RUG$ is a time-invariant linear causal element of $B(X^*, Y^*)$, thereby in light of (257) the lemma is proven.

□

Lemma 6.5 shows that for the ORDAP where the nominal plant P_0 is time-invariant, there is no loss of performance inherent in the restriction to time-invariant feedback. Theorem 6.3 is a corollary to Theorems 6.1, 6.2 and Lemma 6.5.

Theorem 6.3 *If P_0 is time-invariant and assumptions (A1) and (A5) hold then,*

$$
\begin{aligned}
\mu &= \inf\Big\{ \sup_{P\in C(P_0,V)} \|W(I-P_0Q)(I+\Delta PQ)^{-1}\| \;: \qquad (260)\\
&\qquad Q \in B_c(l^2_+, l^2_+) \text{ and } (I+(P-P_0)Q)^{-1} \in B(l^2_+, l^2_+) \;\; \forall P \in C(P_0, V)\Big\}\\
&= \inf_{\substack{Q\in B_c^{ti}(l^2_+, l^2_+)\\ \|VP_0Q\|\le 1}} \sup_{P\in C(P_0,V)} \|W(I-P_0Q)(I+\Delta PQ)^{-1}\| \qquad (261)
\end{aligned}
$$

Moreover an optimal time-invariant feedback exists for the optimizations described by (260) and (261).

Concluding Remark. Theorem 6.3 provides a justification for the restriction to *time-invariant* linear feedback in chapters 2,3,4 and 5, since it suggests that for SISO discrete time systems at least, no advantage can be gained in performance by widening the class of linear control laws to include those varying in time.

Chapter 7

Information-Based Notions of Uncertainty

In this chapter system uncertainty will be quantified using a measure of metric complexity known as Kolmogorov ϵ-dimension. Initially, estimates are obtained for the ϵ-dimension of a class of H^∞ discrete-time systems which satisfy an exponential bound on the impulse response. Subsequently, the ability of feedback to the reduce the ϵ-dimension of sets of plant uncertainty is investigated for SISO discrete-time systems. Various cases are considered, including situations where feedback is applied both before and after identification. For certain open-loop sets of multiplicative plant uncertainty, feedback is shown to reduce asymptotically (i.e. ϵ small) the complexity of identification in cases where the a posteriori objective tolerance is defined by a W-weighted H^∞ sphere, and the identification is constrained to produce an unweighted H^∞ tolerance a posteriori. Large-ϵ (i.e. non-asymptotic) results are obtained for the simpler case of quantifying the effect of feedback on the ϵ-dimension of a set of additive output disturbances. We show that for this case the ability of feedback to reduce the ϵ dimension is a decreasing function of ϵ, and that the reduction is only significant when ϵ is not small in a certain sense.

7.1 Measures of Metric Complexity

The notions of ϵ-dimension and ϵ-entropy were introduced into feedback theory in [56], [55] where they were proposed as measures of the complexity of identification and the effectiveness of feedback in reducing identification costs.

In identification experiments we start with some a priori information about the uncertain plant. This is expressed here by assuming that it belongs to a subset $\mathbf{S}$ of possible plants in a Banach space. For purposes of this chapter assume the Banach space to be H^∞, and the subset to be a sphere (although the main ideas carry over to more general Banach spaces and subsets). The objective is to identify the plant to a tolerance ϵ. The *ϵ-dimension* of $\mathbf{S}$, denoted by $\mathcal{N}_{\mathbf{S}}(\epsilon)$ is the dimension of the smallest subspace of H^∞ whose distance from $\mathbf{S}$ does not exceed ϵ (for more details see [56] or [34] Chapter 9). $\mathcal{N}_{\mathbf{S}}(\epsilon)$ can be viewed as a measure of the intrinsic complexity of identification by linear schemes. A measure which does not depend on linearity is *ϵ-entropy*, $\mathcal{I}_{\mathbf{S}}(\epsilon)$, defined to be log_2 of the smallest number of ϵ-balls in H^∞ needed to cover $\mathbf{S}$ (see [56] or [34] Chapter 10). There has recently been renewed interest in such measures, (see, e.g. [51], [25]).

The starting point for the discrete case in Sect. 3 of [56] were the well known results of Tichomirov [48] p93 and Vitushkin [50] p36, Theorem 1, giving $\mathcal{N}_{\mathbf{S}}(\epsilon)$ and $\mathcal{I}_{\mathbf{S}}(\epsilon)$ respectively for H^∞ functions analytic and bounded in an enlarged disc. Let us summerize these results. Let H^∞_a denote H^∞ of the unit disc of radius e^a, $a \geq 0$. For any $C > 0$, $a \geq 0$, let $\mathbf{b}(C, a)$ denote the ball $\{K \in H^\infty \ : \ K \in H^\infty_a \ and \ |k(z)| < C \ for \ |z| \leq e^a\}$. Then, for $\mathbf{S} = \mathbf{b}(C, a), a > 0$ and ϵ measured in the H^∞ norm, the estimates

$$\mathcal{N}_{\mathbf{S}}(\epsilon) = \left[\frac{1}{a} ln \frac{C}{\epsilon}\right]^{int} \tag{262}$$

$$\mathcal{I}_{\mathbf{S}}(\epsilon) = \left[\frac{1}{a}\left(ln\frac{C}{\epsilon}\right)^2\right]^{int} + O(ln\frac{C}{\epsilon} ln \ ln\frac{C}{\epsilon}) \tag{263}$$

hold. $O(x)$ signifies a function satisfying $|O(x)| < \ Const. \ |x|$. The formula (263) is asymptotically accurate in the sense that the second term on the RHS of (263) becomes negligible with respect to the first. With a few notable exceptions such as (262), complexity theory provides only asymptotic formulas which here take the form

$$\mathcal{N}_{\mathbf{S}}(\epsilon) \sim \frac{1}{a} ln \frac{C}{\epsilon} \tag{264}$$

$$\mathcal{I}_{\mathbf{S}}(\epsilon) \sim \frac{1}{a} \left(ln \frac{C}{\epsilon} \right)^2 \tag{265}$$

where the notation $f(\epsilon) \sim g(\epsilon)$ means that $f(\epsilon) = g(\epsilon)[1 + O(\epsilon)]$. Another exception to the generalization of (262) by Taikov [47], which concerns the set $\mathbf{b}_m(C, a) = \{K \in H^\infty \ : \ K^{(m)} \in \mathbf{b}(C, a)\}$, where $K^{(m)}$ denotes the m-th derivative, $m = 0, 1, ...$; we shall only use the case $m = 1, \ a = 0$, for which

$$\mathcal{N}_{\mathbf{b}_1(C,0)}(\epsilon) = \left[\frac{C}{\epsilon} \right]_{int}. \tag{266}$$

We note that the a priori information in (262)-(265) involves analyticity in the frequency domain, whereas in identification problems a priori information usually pertains to the time-domain, even where a posteriori tolerances are specified in H^∞. Accordingly, the idea here is to obtain formulas similar to (262) and (263) for plants whose impulse responses satisfy an exponential bound. Let

$$\mathbf{h}(C, a) := \{K \in H^\infty \ : \ |k(n)| \leq Ce^{-an}\}$$

where $k(\cdot)$ denotes the discrete Fourier inverse transform of K.

7.2 ϵ-Dimension of $\mathbf{h}(C, a)$

Theorem 7.1 If $\mathcal{N}_{\mathbf{h}(C,a)}(\epsilon)$ is the ϵ dimension of the set $\mathbf{h}(C, a)$ in the Banach space H^∞ then,

$$\left[\frac{1}{a} ln \frac{C}{\epsilon} \right]^{int} + C_1 \leq \mathcal{N}_{\mathbf{h}(C,a)}(\epsilon) \leq \left[\frac{1}{a} ln \frac{C}{\epsilon} + C_2 \right]^{int} \tag{267}$$

where C_1 is a constant which satisfies $[\frac{1}{2(e^{2a}-1)}]_{int} \leq C_1 \leq \left[\frac{1}{2(1-e^{-2a})}\right]_{int}$, $C_2 := \frac{1}{a} ln \frac{1}{1-e^{-a}}$ and $[\cdot]_{int}$ denotes rounding to the nearest lower integer.

Proof. Lower Bound. Let $Ball(H^\infty, \epsilon)$ denote the H^∞ ball of radius ϵ, and M a k-dimensional subspace of H^∞. We will employ the following Lemma: If $\mathbf{S}$ contains a set of the form $\mathbf{S}_0 := M \cap Ball(H^\infty, \epsilon')$, $\epsilon' > \epsilon$, then $\mathcal{N}_{\mathbf{S}}(\epsilon) \geq k$. The lemma is a corollary to a theorem on n-th width ([34] Theorem 2, Sect. 9.3), which states that in any Banach space , the intersection of the ϵ-ball with a k-dimensional subspace has ϵ as its $k-1$ width. By that theorem, the $k-1$ width of $\mathbf{S}_0$ is ϵ'; therefore there does not exist a $k-1$ dimensional subspace of H^∞ whose distance from $\mathbf{S}_0$ is ϵ, as $\epsilon < \epsilon'$; consequently $\mathcal{N}_{\mathbf{S}_0}(\epsilon) > k-1$, i.e., $\mathcal{N}_{\mathbf{S}_0} \geq k$, which implies the Lemma.

Here $\mathbf{S} = \mathbf{h}(C, a)$. We will now construct a subspace $M = span\{\Phi_1, ..., \Phi_k\}$ determined by basis vectors $\Phi_i \in \mathbf{S} \cap Ball(H^\infty, \epsilon')$, whose inverse z-transforms ϕ_i have nonoverlapping finite supports in the integers, and consequently are orthogonal in H^2. The Φ_i will be normalized, $\|\phi\|_{H^2} = \epsilon'$, $i = 1, ..., k$. Then any $P \in M \cap Ball(H^\infty, \epsilon')$ must have the form of a linear combination $P = \sum_{i=1}^k \alpha_i \Phi_i$, where $\sum_{i=1}^k |\alpha_i|^2 \cdot \|\Phi_i\|_{H^2}^2 \leq \epsilon'^2$ and consequently $|\alpha_i| \leq 1$, $i = 1, ..., k$: But since the $\Phi_i \in \mathbf{S}$ are chosen so that ϕ_i have nonoverlapping supports in l^2, such linear combinations must also belong to $\mathbf{S}$, and therefore $M \cap Ball(H^\infty, \epsilon')$ is a subset of $\mathbf{h}(C, a)$, establishing the lower bound k.

Let us construct the $\{\Phi_i\}_{i=1}^k$. Let k_1 be the unique integer satisfying $Ce^{-ak_1} > \epsilon \geq Ce^{-a(k_1+1)}$, i.e., $k_1 = \left[\frac{1}{a} ln \frac{C}{\epsilon}\right]^{int} - 1$, where $[\cdot]_{int}$ ($[\cdot]^{int}$) denotes rounding to the nearest smaller (larger) integer and let ϵ' satisfy $Ce^{-ak_1} \geq \epsilon' > \epsilon \geq Ce^{-a(k_1+1)}$. Let ψ denote the exponential function $\psi(n) = Ce^{-an}$, $n = 0, ...$; I_i, $i = 1, ...$ denote the consecutive intervals of integers, $I_i = [n_{i-1}, n_i)$; and $\psi_i \in l^2$ denote the functions with support on I_i and coinciding with ψ there, $\psi_i(n) = \psi(n)$ for $n \in I_i$. The end points of the intervals can be selected to satisfy $\|\psi_i\|_{l^2} \geq \epsilon'$ and $\|\psi_i|I_i - \{n_i\}\|_{l^2} < \epsilon'$, where the last restriction of ψ_i is obtained by eliminating the end point $\{n_i\}$ of the i-th interval for $i = 1, .., k$, k being the maximum number of such intervals. Let $\Phi_i \in L^2 \cap H^\infty$ be the function $\Phi_i := \epsilon' \frac{\Psi_i}{\|\Psi_i\|_{L^2}}$, (where Ψ_i is the z transform of ψ_i) normalized to have norm ϵ' in L^2. By construction $|\phi_i(n)| \leq |\psi_i(n)|$ for all positive integers n, and so $\Phi_i \in \mathbf{S} \cap Ball(H^\infty, \epsilon')$. The first $k_1 + 1$ of the functions $\{\phi_i\}$ will have support on an interval containing a single integer, $I_i = \{i-1\}$. The remaining k_2 functions, $\{\phi_i\}_{i=k_1+2}^{k}$ ($k_2 = k - k_1 - 1$) will have support on an interval which contains two or more consecutive integers in $[k_2 + 1, \infty)$. It follows that k_2 is bounded below by the number of times the sum $\sum_{r=k_1+1}^{\infty} C^2 e^{-2ar}$ can be subdivided into consecutive partial sums each exceeding ϵ'^2 in magnitude. This is, in turn, bounded below by

$$\left[\frac{1}{2\epsilon'^2}\sum_{r=k_1+1}^{\infty} C^2 e^{-2ar}\right]_{int} = \left[\frac{1}{2}\frac{C^2 e^{-2ak_1}}{\epsilon'^2}\frac{1}{e^{2a}-1}\right]_{int} \tag{268}$$

Since $C^2 e^{-2ak_1} \geq \epsilon'^2 > C^2 e^{-2a(k_1+1)}$ by definition of ϵ', the lower bound for k_2 lies in the stated range for C_1 in the statement of the theorem. The lower bound for $\mathcal{N}_{\mathbf{H}(C,a)}(\epsilon)$ then follows by observing that $k = k_1 + k_2 + 1$.

Upper Bound. In order to conclude that $\mathcal{N}_{\mathbf{S}}(\epsilon) \leq k$, it is enough to exhibit a k-dimensional subspace of H^∞, $N = span\{\Gamma_1, ..., \Gamma_k\}$ such that

$$\sup_{g \in \mathbf{S}} \inf_{f \in N} \|f - g\|_\infty \leq \epsilon \tag{269}$$

Define k to be the smallest integer such that $\sum_{r=k}^{\infty} Ce^{-ar} \leq \epsilon$, i.e., $k = \left[\frac{1}{a} ln\frac{C}{\epsilon} + \frac{1}{a} ln\frac{1}{1-e^{-a}}\right]^{int}$. Then the subspace $N = span\{z^0, z^1, ..., z^{k-1}\}$ can be shown to have the required property (269), by simply chosing $f(z) = \sum_{r=0}^{k-1} \alpha_r z^r$ where α_r is the rth Fourier coefficient of g.

□

(264) is now a corollary to the Theorem and the fact that the constants C_1, C_2 are independent of ϵ. (The proof of (265) is similar to the proof of (263) in [34] Chapter 10).

7.3 Reduction of Metric Complexity by Feedback.

It was pointed out in [55], [56] that feedback can reduce the metric dimension and entropy associated with identification, i.e., identification and feedback can each be used to shrink balls of plant uncertainty, and the more effective the latter is, the less complex the former. We would like to detail this in the present context. However let us first consider a simpler case which can serve as a prototype for our theory, namely how identification and feedback reduce the complexity of *additive* disturbances at the plant output [6].

Here P and W will denote a plant and weighting in $H^\infty(disc)$, and $\underline{W} : l^2(-\infty,\infty) \to l^2(-\infty,\infty)$ the convolution operator with frequency response W acting in the Hilbert space $l^2(-\infty,\infty)$. Suppose that the a priori information places disturbances in the set $S := \{\underline{W}d \ : \ d \in l^2[0,n-1], \ \|d\|_{l_2} \leq 1\}$, consisting of the image under $\underline{W}$ of sequences of length n in the unit ball of l^2. Denote by $\eta(W,\epsilon)$ the Lebesgue measure of the set $\{\theta \in [0,2\pi) \ : \ |W(e^{i\theta})| \geq \epsilon\}$. It can be deduced from Szego's theory that the metric complexity of S relative to the set $\{d \in l^2[0,n-1] \ : \ \|d\|_2 \leq 1\}$, as measured by the relative Kolmogorov ϵ-dimension, is

$$\eta(W,\epsilon) + \Delta(n)$$

where $\lim_{n\to\infty} \frac{\Delta(n)}{n} = 0$, which can be interpreted as the complexity of identification without feedback. Denote the restriction of $\underline{W}$ to $l^2[0,n-1]$ by $\underline{W}_{(n)}$. The operator

$\underline{W}^*_{(n)}\underline{W}_{(n)}$ is Toeplitz in Euclidean n-space, and has normalized eigenvectors ζ_i and eigenvalues λ_i^2, $i = 1, ..., n$. The set S is now an n-dimensional polytope in l^2 with sides $\{\lambda_i \underline{W}_{(n)}(\zeta_i)\}_{i=1}^n$. The ϵ-dimension of this set is (see [42] Chapter 4)

$$k(\underline{W}_{(n)}, \epsilon) \;:=\; \textit{Number of singular values of the matrix } W_{(n)}$$
$$\textit{which are greater than } \epsilon$$

The limit of $\frac{1}{n}k(\underline{W}_{(n)}, \epsilon)$ as $n \to 0$ can be shown to be equal to $\eta(W, \epsilon)$, by applying Szego's theorem on Toeplitz forms [53](Corollary 1, pp 205, and Theorem 6, pp 202). Feedback multiplies the disturbance frequency responses by a factor $(1 - PQ) \in H^\infty$. The relative reduction in ϵ-dimension produced by feedback in the limit as $n \to \infty$ is therefore

$$1 - \frac{\eta(W(1 - PQ), \epsilon)}{\eta(W, \epsilon)} \tag{270}$$

If the optimal weighted sensitivity is $\inf_{Q \in H^\infty} \|W(1 - PQ)\|_\infty =: \mu$ and $\epsilon > \mu$, then the relative reduction in dimension is 1, i.e., no identification is needed to shrink disturbances at the plant output to a tolerance ϵ. However, for smaller values of ϵ the reduction is smaller and, for strictly causal systems for example ($P(0) = 0$) it approaches 0 as $\epsilon \to 0$. When $W = 1$ i.e. the disturbances are unweighted, (270) is a non-increasing function of ϵ.

Let us now return to the more interesting case of uncertainty in the plant. Let $\mathbf{A}$ be a Banach subspace of H^∞, and the set $\mathbf{S}$ of uncertain plants be a ball in $\mathbf{A}$ centered at $P_0 \in \mathbf{A}$ and of radius C, denoted by $\mathbf{S}(P_0, C) := \{P \in \mathbf{A} \;:\; \|P - P_0\|_{\mathbf{A}} \leq C\}$. We will consider identification schemes which are constrained to replace $\mathbf{S}(P_0, C)$ by some smaller sphere $\mathbf{S}(P_1, \epsilon)$, with new center P_1 somewhere in $\mathbf{S}(P_0, C)$. Note that feedback is employed in the nominal plant invariant form outlined in Chapter 2 and is constrained such that $C(I+PC)^{-1} =: Q \in \mathbb{A}$. For such a feedback, the open-loop to closed-loop map $P \to \mathbb{K}_Q(P)$ has the parameterization (see [57] for details),

$$\mathbb{K}_Q(P) = P(I + Q\Delta P)^{-1}, \quad \Delta P := P - P_0,$$

and closed-loop perturbations satisfy (see Chapter 2)

$$\Delta\mathbb{K}_Q(P) = \mathbb{K}_Q(P) - \mathbb{K}_Q(P_0) = (I - P_0Q)(I + \Delta PQ)^{-1}\Delta P \tag{271}$$

For a weighting W such that $W^{\pm 1} \in \mathbf{A}$, $\|W\|_{\mathbf{A}} \leq 1$, assumed fixed throughout, if such a feedback is used to shrink weighted perturbations in a ball $\mathbf{S}(P_0, C)$, the optimal shrinkage is

$$\begin{aligned} \mu(P_0, C) &:= \inf_{Q \in \mathbb{A}} \sup_{\Delta P} \frac{1}{C} \|W\Delta\mathbb{K}_Q(P)\|_{\mathbf{A}} \qquad (272) \\ &\leq \inf_{Q \in \mathbb{A}} \|W(I - P_0Q)\|_{\mathbf{A}}(1 - C\|Q\|_{\mathbf{A}})^{-1} \end{aligned}$$

the last inequality following from (271). (For stable plants) $\mu(P_0, C) \leq 1$ always. When $\mathbf{A} = H^\infty$ then (272) is equivalent to the ORDAP. The objective now is to shrink an a priori ball of uncertainty $\mathbf{S}(P_0, C)$ to achieve a W-weighted tolerance $\epsilon > 0$, relying on feedback to achieve as much of the shrinkage as possible, and on identification for the rest. It is assumed that the identification process starts with an unweighted ball of uncertainty and shrinks its (unweighted) radius. There are two main possibilities.

Case 1. An example in which feedback is applied prior to identification. Let $\mathbf{A}$ be the Banach space H^∞_a where the norm is taken as the infinity norm on the circle radius e^a. Suppose for case 1 that the complexity of the of the closed-loop set of uncertainty is taken to be the Kolmogorov ϵ-dimension of the smallest H^∞ sphere of containment in $\mathbf{A}$ with center P_0. If the assumptions of Theorem 3.4 of Chapter

3 hold for the ORDAP (with H^∞ replaced by H^∞_a) represented by (272), then the smallest such sphere containing the set $W\mathbb{K}_{Q_{opt}}(\mathbf{S}-P_0)$ is $\mu(P_0,C)(\mathbf{S}-P_0)$. Thus, recalling that the control tolerance is measured in the W- weighted H^∞ metric, we obtain that the relative reduction of Kolmogorov ϵ-dimension due to the optimal feedback of (272) is

$$\frac{\mathcal{N}_{W(\mathbf{S}-P_0)}(\epsilon)-\mathcal{N}_{\mu(P_0,C)(\mathbf{S}-P_0)}(\epsilon)}{\mathcal{N}_{W(\mathbf{S}-P_0)}(\epsilon)} \quad where \;\; \mu=\mu(P_0,C) \tag{273}$$

(273) follows, since it can be shown that from the point of view of ϵ-dimension it is irrelevant whether the weighting W acts on the norm or on the set. If W is bounded away from zero on the unit circle the quotient (273) converges to zero as $\epsilon \to 0$.
Case 2. Feedback is applied after identification. If feedback is capable of improving tolerances after identification, say from ϵ_1 to ϵ ($\epsilon_1 > \epsilon$), then it is enough to identify to the larger tolerance, with an attendant gain in dimension of $\mathcal{N}_{\mathbf{S}}(\epsilon_1)-\mathcal{N}_{\mathbf{S}}(\epsilon)$, where the a priori ball is $\mathbf{S} := \mathbf{S}(P_0,C) \subset \mathbf{A}_0 \subset H^\infty$. The feedback action is possible whenever *each* plant P in $\mathbf{S}$ is sufficiently invertible, satisfying $\epsilon_1\mu(P_0,\epsilon_1) \leq \epsilon$.

7.4 A Case of Multiplicative Uncertainty.

A variant of Case 2 involves multiplicative plant uncertainty for which identification tolerances are expressed logarithmically. This will involve plants with factorizations $P = VU$, where $U \in H^\infty$ is a fixed known inner function, and $V \in H^\infty_a$ is an uncertain outer function. A priori lnV is assumed to be in the set (see (266)).

$$\mathbf{V}_0 = \{lnV \in H^\infty_a \;:\; lnV - lnV_0 \in \mathbf{b}_m(C,a)\}$$

for some fixed $a \geq 0$ and integer m; i.e, if $m=0$, V is a logarithmic sphere with center at lnV_0. The function to be identified here is lnV, and after identification it lies in the shrunken logarithmic sphere,

$$\mathbf{V}_1(P_1,\delta) := \{lnV \in H_a^\infty \ : \ lnV - lnV_1 \in Ball(H^\infty, \delta)\}$$

The center of such an identified sphere can be shifted back to lnV_0 without affecting the radius δ, by multiplying V_1 by $\frac{V_0}{V_1}$, thus keeping the center invariant under identification. Assume the objective of the combination of identification and feedback is to achieve a weighted fractional tolerance $\|\frac{W\Delta V}{V_0}\|_{H^\infty} \le \epsilon$, $\Delta V := V - V_0$, where $W \in H^\infty$, $\|W\|_{H^\infty} = 1$. In the absence of feedback, if identification is constrained to produce a ball $V_1(P_0,\delta)$, then the ball needs to conform to the unweighted tolerance $\|\frac{\Delta V}{V_0}\|_{H^\infty} \le \epsilon$, which establishes the relation $\delta = \phi^{-1}(\epsilon)$ where

$$\epsilon = sup\{|x| \ : \ |ln(1+x)| \le \delta, \ x \in \mathbb{C}\} =: \phi(\delta)$$

which has the property that $\phi^{-1}(\epsilon) \to \epsilon$ as $\epsilon \to 0$.

A larger identification tolerance $\delta_1 := \phi^{-1}(\epsilon_1)$ $(\delta_1 \ge \delta)$ is sufficient if feedback can shrink ϵ_1 to ϵ $(\epsilon_1 > \epsilon)$. Suppose that feedback minimizes sensitivity for the nominal plant P_0, let $\mu(P_0) := \inf_{Q\in H^\infty} \|W(1-P_0Q)\|_\infty$, and

$$\begin{aligned} \mu_*(P_0,\epsilon_1) &:= \inf_{Q\in H^\infty} \sup\left\{\frac{1}{\epsilon_1}\left\|\frac{W\Delta\mathbb{K}_{Q(P_0)}(P)}{\mathbb{K}_{Q(P_0)}(P_0)}\right\|_{H^\infty} \ : \ \|\frac{\Delta P}{P_0}\|_{H^\infty} \le \epsilon_1\right\} \\ &\le \frac{\mu(P_0)}{(1-\epsilon_1\|P_0Q_{(P_0)}\|_{H^\infty})} \end{aligned}$$

Feedback reduces the ϵ-dimension of identification from $\mathcal{N}(\epsilon) := \mathcal{N}_{\mathbf{b}_m(C,a)}(\phi^{-1}(\epsilon))$ to $\mathcal{N}(\epsilon_1) := \mathcal{N}_{\mathbf{b}_m(C,a)}(\phi^{-1}(\epsilon_1)$. Let us compute this advantage.

<u>Case 2a</u>. $a > 0$, $m = 0$. This involves an a priori bound on lnV on the enlarged disc, but no explicit constraint on its derivative. Here

$$\begin{aligned} \mathcal{N}(\epsilon) - \mathcal{N}(\epsilon_1) &= \left[\frac{1}{a}ln\phi^{-1}(\epsilon)\right]_{int} - \left[\frac{1}{a}ln\phi^{-1}(\epsilon_1)\right]_{int} \\ &\ge \left[\frac{1}{a}ln\mu\right]_{int} - 1 \quad \textit{for } \epsilon \textit{ small enough} \end{aligned}$$

Here feedback is asymptotically insignificant as $\mathcal{N}(\epsilon_1) \sim \mathcal{N}(\epsilon)$. Of course, it may be significant for large values of ϵ.

Case 2b. $a = 0$, $m = 1$, i.e. there is an a priori bound on the derivative $\frac{d}{d\theta}(lnV)(e^{i\theta})$ on the unit disc. By Taikov's theorem (266),

$$\frac{\mathcal{N}(\epsilon_1)}{\mathcal{N}(\epsilon)} = \frac{\left[\frac{C}{\epsilon_1}\right]^{int}}{\left[\frac{C}{\epsilon}\right]^{int}} \longrightarrow \frac{\epsilon}{\epsilon_1} = \mu(P_0) \;\; as \; \epsilon \to 0$$

Remark. Under the assumption of a derivative bound on the log frequency response, feedback reduces the complexity by a factor approaching $\mu(P_0)$ as $\epsilon \to 0$, which is asymptotically significant provided $\mu(P_0) < 1$. A comparison of cases 2a, 2b, suggests that (for at least the constrained identification problem defined here) feedback is asymptotically significant where the Fourier coefficients of the transfer function to be identified, here $\widehat{lnP_0}(n)$, decrease linearly as $n \to \infty$, but is not significant where they decrease exponentially.

Chapter 8

Concluding Remarks

8.1 Synopsis

The main subject of Chapters 1-6 of the thesis was an analysis of the ORDAP and two-disc problems, first in the time-invariant and then the time-varying cases. The fundamental nature of these problems, coupled with the fact that they were intractable by established methods (even approximately in the case of the ORDAP) provided the motivation for the re-examination undertaken in the first six chapters. We began by showing, for both MIMO time-invariant and SISO time-varying systems, that the ORDAP could be reduced to an implicit form of a two-disc type problem. The recognition that these two-disc problems were in fact distance minimizations in a certain non-standard Banach space of linear systems, enabled predual and dual representations to be obtained. These representations allowed the following two insights into the problem. Convex programming methods, in the form of the ellipsoid algorithm of Shor, Yudin and Nemirovsky [45], were used to derive non-hueristic algorithms for numerically solving the ORDAP. On a more abstract level, alignment conditions relating the closest element in the in the distance minimization to the maximal element in the dual optimization were obtained. This provided a geometrical framework for the ORDAP and two-disc problems which revealed various qualitative properties of the optimal solution, and shed light on how the potential of feedback to reject disturbances depends on the radius of the open-loop plant uncertainty. Some examples of the conclusions obtained in this manner included the following observations: A 'flatness' property of optimal solutions of the ORDAP; strict monotonicity of the dependence of optimal robust performance on open-loop uncertainty radius at all frequencies; existence and uniqueness (for the SISO case) of optimal control laws for the ORDAP.

The last chapter of the thesis was concerned with information-based measures of plant uncertainty i.e. metric complexity. The objective was to understand when and by how much feedback could reduce the quantity of information necessary to control an uncertain system to some desired tolerance. It was shown that for certain open-loop multiplicative spheres of uncertainty, feedback could asymptotically reduce the measure of metric complexity known as Kolmogorov ϵ-dimension, when applied after constrained (in a certain sense) identification. The action of feedback on the metric complexity of a class of additive disturbances was also considered for the non-asymptotic case. The conclusion in this case was that feedback could reduce the ϵ-dimension if and only if ϵ was not small.

8.2 Directions for Further Research

Among the more immediate goals of further research are the development of a robust control synthesis software package for both the SISO and MIMO cases of the ORDAP, based on the numerical solution articulated in the algorithms of Chapter 4. Other objectives include an analysis of the ORDAP for time-varying continuous-time systems, extending the discrete-time results of Chapter 6, and an extension of the duality methods of Chapter 3 to handle convex constraints representing other feedback objectives.

Longer-term goals of future research stemming from this work include:

1) The development of a systematic means of extracting nominal models from raw frequency response data, justifying the 'nominal plant plus weighted uncertainty' descriptions used in this thesis and elsewhere.

2) Establishing a more complete theory of the effect of feedback on complexity, unifying the objectives of feedback and identification in form of the single goal of complexity reduction.

Appendices

Appendix A: Properties of $\chi(\cdot)$

Lemma. $\chi : [0,\infty) \rightarrow [0,\infty)$ *defined in (27), is a continuous, non-decreasing function for which there exists a smallest positive fixed-point*

Proof. Continuity of χ follows form the argument used in [20] for a similar quantity. The non-decreasing nature of χ is a consequence of the fact that the term inside the brackets in (27) is an increasing function of r.

Let $s(r) := \chi(r) - r$. $s(\cdot)$ is a continuous function for which

$$s(0) = \inf_{Q \in H^{\infty}_{n\times n}} \|W_1(I - P_0 Q)W\|_{\infty} > 0 \quad and \quad s(\|W_1 W\|_{\infty} + 1) < 0 \tag{274}$$

The last inequality is a consequence of the fact that $\chi(r) \leq \|W_1 W\|_{\infty}\ \forall r \geq 0$. Because continuous functions map connected sets to connected sets there must exist at least one $\hat{r} \in [0, 1 + \|W_1 W\|_{\infty}]$ such that $s(\hat{r}) = 0$. The continuity of χ also implies that the set $\{r \in [0,\infty) \ : \ s(r) = 0\}$ is closed in the topology of the reals, and thus there must exist a minimum element (since it is non-empty). Hence χ has a minimum fixed-point in $(0,\infty)$.

Appendix B: Proof That B_* is a Banach Space

Lemma *B_* is a Banach space*

Proof. The integrand of (48) represents a norm on the linear space $\mathbb{C}^{n\times n} \times \mathbb{C}^{n\times n}$. Thus (48) defines a norm on the linear space B_*. To prove the completeness of B_* suppose that $G_k := [G_1^{(k)}\ G_2^{(k)}]$, $G_1^{(k)}, G_2^{(k)} \in L^1_{n\times n}$ is a Cauchy sequence in the metric defined by the norm (48). It follows from the definition of STr that the L^1 sequence resulting from the restriction of G_k to a single entry, is Cauchy in the Banach space L^1. Hence there is element-by-element convergence in the L^1 topology of G_k to some $G \in B_*$. Denote G by $[G_1\ G_2]$. The following inequality proves that $G_k \to G$ in the B_* metric.

$$
\begin{aligned}
&\int_0^{2\pi} Max\left(STr(G_1^{(k)}), STr(G_2^{(k)})\right) d\theta \\
&\quad \le n \int_0^{2\pi} \left(\sum_{i,j=1}^{n} \left|(G_1^{(k)} - G_1)_{i,j}\right| + \sum_{i,j=1}^{n} \left|(G_2^{(k)} - G_2)\right| \right)_{i,j} d\theta \qquad (275) \\
&\quad \to 0 \quad as \quad n \to \infty
\end{aligned}
$$

□

Appendix C: Matrix Trace Inequalities.

Lemma.

i. If X, Y are two $n \times n$ matrices then,

$$|TrXY| \leq |X| \sum_{i=1}^{n} \sigma_i(Y) \tag{276}$$

ii. If X has singular value decomposition $X = VDU$ equality in (276) is achieved for $Y = U^* D_1 V^*$ where D_1 is the diagonal matrix with diagonal $\{1,\ 0, ...,\ 0\}$.

Proof. i. Let $Y = UDV$ be the singular value decomposition of Y. Then,

$$|TrXY| \leq |TrXUDV| = |TrVXUD|$$

Since for square matrices the trace is merely the sum of the diagonal elements, $TrVXUD = \sum_{i=1}^{n} (VXU)_{i,i} d_i$ where $\{d_i\}_{i=1}^{n}$ are the positive diagonal entries of D. Thus,

$$|TrVXUD| \leq \sum_{i=1}^{n} |(VXU)_{i,i}| d_i \leq Max_{i=1,...,n} |(VXU)_{i,i}| \cdot TrD$$

Since the magnitude of each element of VXU is bounded above by $|VXU|$

$$|TrVXUD| \leq |X| \cdot |TrD| = |X| \sum_{i=1}^{n} \sigma_i(Y)$$

ii. Follows by substitution.

□

Appendix D: The Gain-Phase Relationship.

The Bode gain-phase relationships for a minimum phase system were originally derived from the Hilbert transform formula for analytic functions. In this appendix we summerize the Hilbert transform form of the gain-phase relationship, and state some of the salient properties.

For a normalized outer H^∞ scalar values function, we have the following relationship between phase and gain on the unit circle [32],

$$arg h(e^{i\theta}) = \frac{1}{2\pi} \int_0^{2\pi} cot\left(\frac{t-\theta}{2}\right) log|h(e^{it})|dt \tag{277}$$

For H^∞ functions with real Fourier coefficients (corresponding to real life systems) $|h(e^{i\theta})| = |h(e^{-i\theta})| \;\; \forall t \in [0, 2\pi)$. (277) becomes,

$$\begin{aligned} arg h(e^{i\theta}) &= \int_0^{\pi} K(\theta, t) log|h(e^{it})|dt & (278) \\ where \;\; K(\theta, t) &:= -\frac{1}{2\pi}\left(cot(\frac{\theta - t}{2}) + cot(\frac{\theta + t}{2})\right) & (279) \end{aligned}$$

Appendix E: A Maximum Modulus Principle

For H^∞ functions the maximum modulus theorem states that the supremum of the function over the interior of the unit disc is the same as the essential supremum over the unit circle. The same property can be exhibited for the norm defined by the two suprema of (44) for pairs of $H^\infty_{n\times n}$ functions i.e., the norm of the Banach space $\hat{B}$.

Lemma. *If* $X, Y \in H^\infty_{n\times n}$ *then*

$$\sup_{z\in D} \sup_{\substack{\zeta\in\mathbb{C}^n \\ |\zeta|\le 1}} |X(z)\zeta| + |Y(z)\zeta| = ess \sup_{\theta\in[0,2\pi)} \sup_{\substack{\zeta\in\mathbb{C}^n \\ |\zeta|\le 1}} |X(e^{i\theta})\zeta| + |Y(e^{i\theta})\zeta| \qquad (280)$$

Proof. The inequality,

$$\sup_{z\in D} \sup_{\substack{\zeta\in\mathbb{C}^n \\ |\zeta|\le 1}} |X(z)\zeta| + |Y(z)\zeta| \ge ess \sup_{\theta\in[0,2\pi)} \sup_{\substack{\zeta\in\mathbb{C}^n \\ |\zeta|\le 1}} |X(e^{i\theta})\zeta| + |Y(e^{i\theta})\zeta| \qquad (281)$$

follows from the fact that $X(e^{i\theta})$ and $Y(e^{i\theta})$ are the non-tangential limits of X and Y in D for almost every $\theta \in [0, 2\pi)$.

Suppose the inequality in (281) is strict, i.e.,

$$\sup_{z\in D} \sup_{\substack{\zeta\in\mathbb{C}^n \\ |\zeta|\le 1}} |X(z)\zeta| + |Y(z)\zeta| > ess \sup_{\theta\in[0,2\pi)} \sup_{\substack{\zeta\in\mathbb{C}^n \\ |\zeta|\le 1}} |X(e^{i\theta})\zeta| + |Y(e^{i\theta})\zeta| \qquad (282)$$

Then there must exist constant unit vectors $\eta_1, \eta_2, \zeta_1 \in \mathbb{C}^n$ such that,

$$\sup_{z\in D} \{|\eta_1^* X(z)\zeta_1| + |\eta_2^* Y(z)\zeta_1|\} > ess \sup_{\theta\in[0,2\pi)} \left\{|\eta_1^* X(e^{i\theta})\zeta_1| + |\eta_2^* Y(e^{i\theta})\zeta_1|\right\} \qquad (283)$$

Define the scalar valued H^∞ functions x, y by, $x(z) := \eta_1 X(x)\zeta_1$, $y(z) := \eta_2 Y(z)\zeta_1$ $z \in D$. Thus (283) becomes,

$$\sup_{z\in D}\left(|x(z)| + |y(z)|\right) > ess \sup_{\theta\in[0,2\pi)}\left(|x(e^{i\theta})| + |y(e^{i\theta})|\right) + \delta \tag{284}$$

for some $\delta > 0$. From the definition of the supremum, there exists $z_0 \in D$ such that $|x(z_0)| + |y(z_0)| \geq \sup_{z\in D}\left(|x(z)| + |y(z)|\right) - \frac{\delta}{2}$. Thus,

$$\sup_{z\in D}|x(z) + e^{i\gamma}y(z)| \geq ess \sup_{\theta\in[0,2\pi)}\left(|x(e^{i\theta})| + |y(e^{i\theta})|\right) + \frac{\delta}{2} \tag{285}$$

where $e^{i\gamma}$ is the unit magnitude constant for which $|x(z_0)|+|y(z_0)| = |x(z_0)+e^{i\gamma}y(z_0)|$. But (285) violates the maximum modulus principle for the scalar valued H^∞ function $x+e^{i\gamma}y$. Hence (282) cannot hold and (281) must be an equality, proving the lemma.

□

Bibliography

[1] V.M. Adamjan, D. Z. Arov, and M.G. Krein, Analytic Properties of Schmidt Pairs for a Hankel Operator and the Generalized Shur-Takagi Problem, *Math. USSR Sbornic* 15, 1971, pp. 31-73.

[2] Akhiezer and Glazman, *Theory of Linear Operators in Hilbert Space*, Vols. 1 and 2, Ungar, New York, 1963.

[3] J. Ball and J. W. Helton, A Beurling-Lax Theorem for the Lie Group $U(m, n)$ Which Contains Most Classical Interpolation Theory, *J. Operator Theory*, 9, 1983, pp.107-142.

[4] J. S. Bird and B. A. Francis, On the Robust Disturbance Attenuation Problem, *Proc. CDC*, Athens, Greece, 1986.

[5] S.P. Boyd and C. Barratt, *Linear Control System Design: Limits of Performance*, Prentice-Hall, New Jersey, 1991.

[6] S. P. Boyd, *personal communication*, July 1992.

[7] P.E. Caines, *Linear Stochastic Systems*, Wiley, 1988.

[8] H. Chapellat and M. Dahleh, Analysis of Time-Varying Control Strategies for Optimal Disturbance Rejection and Robustness, *University of California at S.B Report*, CCEC-91-1009, 1991.

[9] H. Chapellat and M. Dahleh, Analysis of Time-Varying Control Strategies for Optimal Disturbance Rejection and Robustness, *IEEE Trans. Automatic Contr.*, AC-37, 11, 1992, pp.1734-1746.

[10] J. Diestel and J.J. Uhl, Vector Measures, *Mathematical Surveys*, 15, American Mathematical Society, R.I., 1977.

[11] P. Dorato, *Robust Control*, IEEE, New York, 1987.

[12] J. Doyle, B. Francis, and A. Tannenbaum, *Feedback Control Theory*, Macmillan, New York, 1992.

[13] J.C. Doyle and G. Stein, Multivariable Feedback Design: Concepts for a Classical/Modern Synthesis, *IEEE Transactions on Automatic Control*, Vol. AC-26, 1, 1981, pp. 4-16.

[14] P.L. Duren, *Theory of H^p Spaces*, Academic Press, New York, 1970.

[15] A. Feintuch and B.A. Francis, Uniformly Optimal Control of Linear Time-Varying Systems, *Systems and Control Letters* , 15, 1985, pp.67-71.

[16] A. Feintuch and R. Saeks, *Systems Theory: A Hilbert Space Approach*, Academic Press, 1982.

[17] D. Flamm, Outer Factor 'Absorption' for H^∞ Control Problems, *ISS Report*, #54, Princeton University, 1990.

[18] Y.K. Foo and I. Postlethwaite, An H^∞ Minimax approach to the Design of Robust Control Systems, *Systems and Control Letters*, 5, 1984, pp. 81-88.

[19] B. Francis, *A Course in H^∞ Control Theory*, Lecture Notes in Control and Information Sciences, Vol. 88, Springer-Verlag, 1987

[20] B. Francis, On Disturbance Attenuation with Plant Uncertainty, *Workshop on New Perspectives in Industrial Control System Design*, Oxford, UK, 1986.

[21] J.B. Garnett, *Bounded Analytical Functions*, Academic Press, New York, 1981.

[22] Grothendiek, Produits Tensoriels Topologiques et Espaces Nucleaires, *Memoirs of the American Mathematical Society*, 16, 1955.

[23] G. Gu and P.P. Khargonekar, Linear and Nonlinear Algorithms for Identification in H^∞ with Error Bounds, *IEEE Trans. Automatic Contr.*, AC-37, 7, 1992, pp. 953-964.

[24] T. Georgiou and M. Smith, Robust Stabilization in the Gap Metric: Controller Design for Distributed Plants, *IEEE Trans. Automatic Contr.*, Vol AC-37, 8, 1992, pp. 1133-1144.

[25] A.J. Helmicki, C.A. Jacobson, and C.N. Nett, Control Oriented System Identification: A Worst Case/Deterministic Approach in H_∞, *IEEE Trans. Automatic Contr.*, Vol. AC-36, 1991, pp. 1163-1176.

[26] H. Helson, Lectures On Invariant Subspaces, Academic Press, New York, 1964.

[27] J.W. Helton and R.E. Howe, A Bang-Bang Theorem for Optimization over Spaces of Analytic Functions, *J. Approximation Theory*, Vol 47, No. 2, 1986, pp101-120.

[28] I. Horowitz, *Synthesis of Feedback Systems*, Academic Press, New York, 1963.

[29] M. Khammash and M. Dahleh, Time-Varying Control and the Robust Performance of Systems with Structured Norm-Bounded Perturbations, *Proc. CDC*, Brighton, England, 1991.

[30] M. Khammash and J.B. Pearson, Performance Robustness of Discrete- Time Systems with Structured Uncertainty, *IEEE Trans. Automat.Control*, AC-36, 4, 1991, pp.398-412.

[31] P. Khargonakar and K. Poolla, Robust Control of Linear Time-Invariant Plants Using Switching and Non-Linear Feedback, *Proc. CDC*, Tampa, Florida, 1989.

[32] P. Koosis, *An Introduction to H^p Spaces*, Cambridge University Press, 1980.

[33] H. Kwakernaak, Minimax Frequency-Domain Performance and Robustness Optimization of Linear Feedback Systems, *IEEE Tran. Automat. Control*, AC-30, No. 10, 1985.

[34] G.G. Lorentz, *Approximation of Functions*, Chelsea Publishing, New York, 1986.

[35] D.G. Luenberger, *Optimization by Vector Space Methods*, John Wiley, 1968.

[36] N.K. Nikol'skii, *A Treatise on the Shift Operator*, Springer Verlag, 1986.

[37] J. G. Owen, A Note on Metric Dimension and Feedback in Discrete Time, *Proc. ACC*, Chicago, Ill, 1992.

[38] J.G. Owen and G. Zames, Unstructured Uncertainty in H^∞, *Proc. Workshop on Robust control*, San Antonio 1991.

[39] J.G. Owen and G. Zames, Duality Theory of Robust Disturbance Attenuation, *to be published in Automatica, May 1993.*

[40] J.G. Owen and G. Zames, Robust H^∞ Disturbance Minimization by Duality, *Proc. CDC*, Brighton, England, 1991.

[41] S. O'Young and B. Francis, Optimal Performance and Robust Stabilization, *Automatica*, Vol. 22, 1986, pp. 171-183.

[42] A. Pinkus, *n-Widths in Approximation Theory*, Springer-Verlag, 1985.

[43] W. Rudin, *Real and Complex Analysis*, McGraw Hill, New York, 1978.

[44] J.S. Shamma and M.A. Dahleh, Time-Varying vs. Time-Invariant Compensation for Rejection of Persistent Bounded Disturbances and Robust Stabilization, *IEEE Trans. Automatic Contr.*, AC-36, 7, 1991, pp.838-847.

[45] N. Z. Shor, Minimization Techniques for Non-differentable Functions, *Springer Series in Computational Mathematics*, Springer-Verlag, 1985.

[46] M.C. Smith, Well-Posedness of H^∞ Optimal Control Problems, *SIAM J. Control and Optimization*, Vol 28, No. 2, 1988, pp. 342-358.

[47] L.V. Taikov, Diameters of Certain Classes of Analytic Functions, Mat. Zametki 22, 1977, pp. 535-542.

[48] V.M. Tichomirov, Diameters of Sets in Function Spaces, *Uspehi Mat. Nauk, Russian Math Surveys*, Vol. 15, 1960, pp. 75-111.

[49] M. Verma and E. Jonckheere L^∞ Compensation With Mixed Sensitivity as a Broadband Matching Problem, *Systems and Control Letters*, 4, 1984, pp.125-129.

[50] A.G. Vitushkin, *Theory of the Transmission and Processing of Information*, Pergamon Press, New York, 1961.

[51] L.Y. Wang and L. Lin, ϵ-Dimension of Exponentially Stable Systems, Wayne State Univ. (Detroit) Report, 1991.

[52] L. Wang and G. Zames, Local-global Double Algebras for Slow H^∞ Adaptation, Part II: Optimization for Stable Plants, *IEEE Trans. Automatic Contr.*, Vol AC-36, 1991, pp. 143-151.

[53] H. Widom, Toeplitz Matrices, in *Studies in Real and Complex Analysis* (Math. Ass. Amer., Studies in Mathematics), I.I. Hirschmann, Jr., Ed. Englewood Cliffs, N.J.: Prentice Hall, 1965.

[54] G. Zames and L. Wang, Double Algebras for Slow H^∞ Adaptation, Part I: Inversion and Stability, *IEEE Trans. Automatic Contr.*, Vol AC-36, 1991, pp. 130-142.

[55] G. Zames, Feedback and Complexity, *Proc. IEEE Conf. Decision Contr.*, 1976, addenda, pp. 1-2.

[56] G. Zames, On the Metric Complexity of Causal Linear Systems, ϵ-entropy and ϵ-dimension for Continuous Time, *IEEE Trans. Automatic Contr.*, vol. AC-12, 1979, pp.222-230.

[57] G. Zames, Feedback and Optimal Sensitivity: Model Reference Transformations, Multiplicative Seminorms, and Approximate Inverses, *IEEE Trans. Automatic Contr.*, AC-23, 1981, pp. 301-320.

[58] G. Zames and B.A. Francis, Feedback, Minimax Sensitivity and Optimal Robustness, *IEEE Trans. Automatic Contr.*, AC-28, 5, 1983, pp. 585-601.

[59] G. Zames and S.K. Mitter, A Note on the Essential Spectrum and Norms of Mixed Hankel-Toeplitz Operators, *Systems and Control Letters* , 10, 1988, pp. 159-165.

[60] G. Zames and J.G Owen, Duality Theory for MIMO Robust Disturbance Rejection, *to be published in the IEEE Trans. Automatic Contr. in May 1993.*

[61] G. Zames, A. Tannenbaum, and C. Foias, Optimal H^∞ Interpolation: A New Approach, *Proc. CDC*, Athens, Greece, 1986, pp. 350-355.