A Constrained Optimization Approach to Control with Application to Flexible Structures

by

Lawrence Kent McGovern

S.B., Aeronautics and Astronautics
Massachusetts Institute of Technology, 1994

Submitted to the Department of Aeronautics and Astronautics in partial fulfillment of the requirements for the degree of Master of Science at the Massachusetts Institute of Technology

June, 1996

© 1996 Lawrence K. McGovern. All rights reserved.

The author hereby grants to MIT permission to reproduce and to distribute copies of this thesis document in whole or in part.

Signature of Author

Department of Aeronautics and Astronautics
May 10, 1996

Certified by

Dr. Brent D. Appleby
Charles Stark Draper Laboratory
Thesis Supervisor

Certified by

Professor Steven R. Hall
Department of Aeronautics and Astronautics
Thesis Advisor

Accepted by

Professor Harold Y. Wachman
Chairman, Departmental Graduate Committee
A Constrained Optimization Approach to Control
with Application to Flexible Structures

by

Lawrence Kent McGovern

Submitted to the Department of Aeronautics and
Astronautics on May 10, 1996 in partial fulfillment of the
requirements for the degree of Master of Science

Abstract

A controller design methodology which minimizes a linear or quadratic closed-loop
design metric subject to a set of linear design specifications is presented in this thesis.
Several useful convex design specifications in the time and frequency domain are given,
and posed as sets of linear constraints on the closed loop. The use of these constraints
is demonstrated in the context of a simple magnetic bearing control example.

The closed-loop optimization relies on a finite-dimensional approximation of the
achieveable space of closed-loop transfer functions. Previous work of this nature
has favored a finite impulse response (FIR) approximation. Alternative sets of or-
thonormal basis functions are explored which can lead to more efficient closed loop
approximations than the FIR filter.

The constrained optimization method was applied to an active vibration isolation
system. The complete controller designs and hardware test results are presented. Due
to an abundance of closely-spaced, lightly-damped structural modes, controller design
for the active vibration isolation system proved to be a formidable task, in which an
efficient formulation of the basis functions became essential. The controllers were
designed by directly constraining the closed loop to meet a set of performance and
robustness constraints.

Thesis Supervisor: Dr. Brent D. Appleby
Title: Senior Member of the Technical Staff

Thesis Advisor: Professor Steven R. Hall
Title: Associate Professor of Aeronautics and Astronautics
Acknowledgements

A large part of the credit for the success of this thesis belongs to my thesis supervisor, Dr. Brent Appleby. Due to his impressive “sixth sense” for control theory, I have never felt at a loss for ideas worth pursuing. I am indebted to the countless hours he has spent discussing the course of this thesis with me. He is also one of the nicest guys to have as a supervisor, which made this experience all the more rewarding.

I am also grateful to Tim Henderson, who provided the practical application for this research, and gave me the freedom to try new and non-standard methods of control design on this system. I feel very fortunate for the opportunity to ground my research in a real hardware system.

I thank Prof. Steven Hall for many useful and interesting discussions over the past year. His suggestions have been very valuable to the development of this thesis.

Nicola Elia certainly deserves a big thanks, as does Dr. Rami Mangoubi. They have each offered my invaluable help in understanding the $\ell_1$ control theory, which was the foundation of most of my research.

Finally, I would like to extend my warmest thanks to Ellen Block. Her continual love and support has always brought out the best in me. Thank you!

This thesis was prepared at The Charles Stark Draper Laboratory, Inc., under Contract 6-2872-60321. Publication of this thesis does not constitute approval by Draper or the sponsoring agency of the findings or conclusions contained herein. It is published for the exchange and stimulation of ideas. I hereby assign my copyright of this thesis to The Charles Stark Draper Laboratory, Inc., Cambridge, Massachusetts.
Contents

1 Introduction
1.1 Motivation .................................................. 13
1.2 Historical Background ........................................ 16
1.3 Contributions .................................................. 17
1.4 Organization .................................................... 18

2 Background
2.1 Signal Norms .................................................... 19
2.2 System Norms .................................................... 20
2.3 Small Gain Theorem ............................................. 22
2.4 Parameterization of All Stabilizing Controllers ............... 22
2.5 Rational Matrix Factorization ................................ 24

3 Controller Synthesis in the Time Domain
3.1 Performance Objectives ........................................ 27
  3.1.1 $H_2$ Optimization ....................................... 28
  3.1.2 $\ell_1$ Optimization ....................................... 28
3.2 Feasibility Constraints ......................................... 29
  3.2.1 Zero Interpolation Conditions ......................... 30
  3.2.2 Rank Interpolation Conditions ......................... 32
  3.2.3 Delay Augmentation Method .......................... 33
3.3 Solution to the $\ell_1$ and $H_2$ Problems .................... 34
3.4 Optimization in $Q$-space ................................... 37
4 **Linear Design Specifications**  
4.1 Time-Domain Constraints ............................................. 40  
4.1.1 Example .................................................................. 40  
4.2 Frequency-Domain Constraints ........................................ 43  
4.2.1 Example .................................................................. 45  
4.3 Stability Margin .......................................................... 46  
4.4 Robust Stability ............................................................ 48  
4.4.1 Example .................................................................. 49  
4.5 Linear Matrix Inequalities ............................................... 52

5 **Efficient Basis Functions for Closed Loop Approximations**  
5.1 Fixed Pole Model .......................................................... 56  
5.2 Laguerre Functions ....................................................... 60  
5.3 General Orthonormal Basis Functions ............................. 63  
5.4 Basis Functions in the Constrained Optimization Control Problem .. 67  
5.4.1 Performance Objective Functions ............................... 67  
5.4.2 Time-Domain Constraints ........................................... 69  
5.4.3 Frequency-Domain Constraints ................................... 69

6 **Application of Constrained Optimization Methods to an Active Vibration Isolation System**  
6.1 Structural Test Model .................................................... 72  
6.2 Design Objectives .......................................................... 74  
6.3 Phase One Design .......................................................... 76  
6.3.1 Optimization ............................................................ 77  
6.3.2 Choice of Basis .......................................................... 78  
6.3.3 Performance Specifications ......................................... 79  
6.3.4 Stability Robustness ................................................... 80  
6.3.5 Controller Reduction .................................................. 81  
6.3.6 Experimental Results .................................................. 83  
6.4 Phase Two Design .......................................................... 86
6.4.1 Performance Specifications ......................... 87
6.4.2 Stability Robustness .................................. 87
6.4.3 Controller Reduction .................................. 89
6.4.4 Experimental Results ................................. 91

7 Conclusions .............................................. 95
    7.1 Recommendations for Future Work ................. 97

Bibliography .............................................. 101
## List of Figures

2.1 Feedback connection for Small Gain Theorem. ........................... 22
2.2 Closed-loop system augmented with stable $Q$. .......................... 24

4.1 Disturbance profile. ....................................................... 41
4.2 Response to fixed input for design with time-domain constraints. ... 42
4.3 Frequency response for design with frequency-domain constraints. . 46
4.4 Response to fixed input for frequency-domain constrained design. . 47
4.5 Minimum distance $d$ from Nyquist plot to critical point. ............ 48
4.6 Gain and Phase Margins versus Sensitivity Magnitude. ............... 49
4.7 Feedback control with multiplicative uncertainty. ..................... 50
4.8 Condensed loop with uncertainty. ...................................... 50
4.9 Uncertainty for plant with structural modes. .......................... 51
4.10 Frequency response of $PK/(1 + PK)$ for design with robustness con-
     straints. ................................................................. 52
4.11 Magnitude constraint approximated with 8 linear constraints. ...... 54
4.12 Magnitude constraint posed exactly as an LMI. ....................... 54

5.1 Fixed-Pole Model approximation of $H(z)$. ............................ 58
5.2 Coefficients $\phi_n$ for a 20th order FPM. ............................ 59
5.3 Impulse response of $H(z)$. ............................................ 59
5.4 First four Laguerre functions for $a = 0.8$. ........................... 61
5.5 $\|H - \phi^T F\|_2$ as a function of $a$ for first example. ............. 63
5.6 Comparison of 50 state Laguerre model and FIR model. ............. 64
5.7 $\|H - \phi^T F\|_2$ as a function of $a$ for second example. .......... 65
6.1 Structural Test Model. ........................................... 73
6.2 Disturbance force PSD. ........................................... 75
6.3 Strut 1 open loop and closed loop using the classical controller. . . . 76
6.4 Comparison of linear model and FRF data for phase 1 design. . . . . 77
6.5 AVIS disturbance rejection problem. ................................ 78
6.6 Impulse response of the closed loop using the classical controller. . . 79
6.7 Constrained optimization and classical performance for phase 1 design. 80
6.8 Two previous AVIS controller design, and resulting frequency constraints. 82
6.9 Constrained complementary sensitivity of the phase 1 design. . . . . 82
6.10 Controller reduction for phase 1 design. ........................... 84
6.11 Wavefront error PSD performance for phase 1 design. ............... 85
6.12 Image motion performance for phase 1 design. ........................ 85
6.13 Reduced design model. ........................................... 87
6.14 Constrained optimization and classical performance. .................... 88
6.15 Complementary sensitivity of phase 2 design. ........................ 89
6.16 Sensitivity of phase 2 design. ....................................... 90
6.17 Nichols plot with linear model of design with stability margin constraints. 90
6.18 Nichols plot with FRF data of classical and constrained optimization designs. .............................. 91
6.19 Controller reduction for phase 2 design. .............................. 92
6.20 Effect of controller reduction on closed loop. .......................... 92
6.21 Wavefront error PSD performance for phase 2 design. ............... 93
6.22 Image motion performance for phase 2 design. ........................ 93
Chapter 1

Introduction

1.1 Motivation

As described by Boyd and Barratt [1], the fundamental problem of control engineering is to find a controller for a given system that meets a set of design specifications, or to determine that no such controller exists. This problem is so far unanswered by current methods of controller design. Traditional methods are often the most indirect in addressing this problem, but they are also the methods which have the widest experience base. The current state of computer technology, along with advances in optimization techniques, offers the possibility for much more direct methods to address this problem.

The techniques of classical control engineering are well established, and have been successfully used for many decades. These techniques include root-locus methods, Nyquist diagrams, Nichols charts, and varying degrees of reasoning and intuition by the engineer. However, there are many drawbacks to classical methods of control design. These methods are geared towards single-input/single-output (SISO) systems, and often have trouble with multiple-input/multiple-output (MIMO) systems with large amounts of cross coupling. Classical methods also do not address optimization. It is up to the designer to iterate until an acceptable design is found. Finally, classical methods can be very indirect at dealing with particular design specifications. Often, it is difficult to understand exactly how a design must be changed to improve the
overall performance.

Modern controller design methods provide an analytic solution to certain classes of control problems, such as the Linear Quadratic Gaussian (LQG) control problem [18]. These methods are based on a state-space representation of the system, and can easily handle MIMO problems. In the LQG framework, a cost functional is chosen to reflect the cost of error in the regulated output and of actuator use. Optimization of this cost leads to a constant state feedback matrix. A Kalman filter is then used to estimate the plant states, based on a Gaussian white noise model of plant disturbances and sensor noise. This leads to the optimal controller design as long as the model is perfect, the plant noises are correctly modeled, and the cost functional has been correctly determined. However, rarely does a control problem correctly fit into the LQG framework. More often, the control engineer uses the cost and noise weights as parameters to “tweak” a design until it has the desired properties. Robustness can be added by shaping the loop gain with frequency weights on the cost functional, by adding fictitious noise, or through the use of the Loop Transfer Recovery (LTR) method [18]. In the end, the standard method is often abandoned, and the desired controller is derived in an indirect way.

The question naturally arises whether there is a more direct method of controller design. There are a multitude of design specifications which may be placed on a closed-loop system. A few examples are maximum overshoot due to a particular input, the frequency response in a particular frequency band, or stability robustness for a set of model uncertainties. While some analytic methods have successfully integrated common design specifications into the solution (e.g., \( \mathcal{H}_\infty \) theory addresses the issue of robust control), the majority of control problems with design specifications do not have a known closed-form solution. This does not prevent these problems from being solved by means of brute force optimization. However, these problems are infinite dimensional and frequently are non-convex, meaning they may have many local minima. To be made tractable, the problem must be approximated in a finite number of dimensions (e.g., approximating a transfer function as a finite impulse response (FIR)). Even with this approximation, determining the global minimum of
a non-convex optimization problems with a reasonable amount of computations can be done only for small problems.

Convexity is therefore a very important characteristic of optimization problems. If the objective function and the set of design specifications are convex, then any local minimum is guaranteed to be the global minimum. This means that the convex optimization problem can be solved much more efficiently than the non-convex case. In [1], Boyd and Barratt show that a large number of design specifications can be posed as convex constraints on the closed loop. By only considering convex constraints, they have also shown that it is possible to efficiently determine the limits of performance achieved by any linear time-invariant (LTI) controller for an LTI system, subject to the finite dimensional approximation. This is a powerful tool for the control engineer, because it determines if a set of design specifications are overly stringent.

One of the most celebrated convex optimization routines is the Simplex Method [6], which efficiently solves the following problem:

$$\min_{x \in \mathbb{R}^n} c^T x$$
subject to
$$A_e x = b_e$$
$$A_s x \leq b_i$$

A problem in this form is known as a linear program. Closely related is the quadratic program, which includes in the objective function a quadratic term, $x^T G x$. If the constant matrix $G$ is positive definite, then the objective function is convex. A convex quadratic program also may be solved efficiently [14]. If possible, it is advantageous to formulate convex optimization problems as linear or quadratic programs. Because of their extreme efficiency, these are the optimization methods explored in this thesis. It will become clear in later chapters that many design specifications can be posed as or approximated with a set of linear design specifications.
1.2 Historical Background

In the 1960s, Fegley and colleagues published a series of papers noting the usefulness of linear and quadratic programming to the design of control systems [12]. Using these methods, they were able to find an optimal controller subject to a set of linear constraints by solving for the coefficients of the closed-loop FIR. Their examples used only simple SISO systems, and their ideas were not extended to more complicated systems, but part of this is presumably due to the state of computer technology at the time. The approach that Fegley and his colleagues outlined was an insightful way of incorporating a variety of convex constraints into the controller design process. The basis of this method, linear and quadratic programming, is of practical use today due to vastly improved computer speeds and optimization methods.

Linear programming became central to the solution of the $\ell_1$ optimal control problem by Dahleh and Pearson in 1987 [7]. The goal of the $\ell_1$ control problem is to minimize the maximum peak-to-peak gain of a closed-loop system driven by an unknown but bounded disturbance. The solution to this problem was found to be the solution of a finite-dimensional linear program. There is a striking similarity between the solution to this problem, and the types of problems being solved by Fegley. For the optimal $\ell_1$ solution, Dahleh and Pearson proved that the closed loop has an FIR structure, so the solution was formulated in terms of the closed-loop impulse response. This formulation makes the addition of linear design specifications very straightforward, but addition of these constraints was not explored until later [8].

Convex optimization in control system design was revisited by Boyd and colleagues in 1988 [2]. This paper describes a controller design method based on a software program called QDES developed by the authors. In this approach, the emphasis was placed on optimization over the free parameter $Q$ in the well known $Q$-parameterization [20]. The closed loop is affine in $Q$, so the search over all achievable closed-loop maps is replaced by a search over all stable $Q$. In QDES, $Q$ is approximated as an FIR filter, the design specifications are written as linear constraints, and the optimization carried out as a quadratic program.
New methods of convex optimization are beginning to take hold in the control community. Following recent advances in interior-point convex optimization algorithms by Nesterov and Nemirovskii [19], many more types of convex constraints may be accommodated efficiently in convex optimization programs. Linear and quadratic programming are limited to linear constraints. The interior-point methods proposed by Nesterov and Nemirovskii can accommodate linear matrix inequality (LMI) constraints. These methods are currently finding practical applications in the field of control [3, 4, 5].

1.3 Contributions

The most significant contribution of this thesis is a complete design of a linear controller, tested on a real hardware system. This lends credibility to the field of convex optimization as a practical tool for controller design. In demonstrating a new control method, academic examples are often used which have been developed specifically for the control theory they are demonstrating. By using a real world control example, the method must instead be tailored to the example. In the literature, the academic examples used to demonstrate constrained optimization almost universally rely on an FIR model of the closed loop. In practice, this is often not an efficient representation of a transfer function (i.e., thousands of terms may be required in the FIR). Therefore, this thesis expands an existing constrained optimization method to accommodate other basis functions. Another contribution to the method is a more general way of posing constraints on the closed-loop frequency response when the delay augmentation method is used. It will be seen in that under certain conditions, the delay augmentation method can invalidate constraints on the closed loop. A solution to this problem for frequency constraints is presented in this thesis.
1.4 Organization

This thesis focuses primarily on optimization directly in the closed-loop space using linear and quadratic programming, and neither the method used in QDES nor the use of LMIs are explored in detail. These methods are worth noting because of their close relationship with the methods used in this thesis. The methods which are used are closely based on the methods developed by Dahleh and Diaz-Bobillo [8].

The organization of this thesis falls into seven chapters. The second chapter consists of necessary background material. Chapter 3 approximates the closed loop as an FIR filter, and develops the closed-loop optimization as a linear program for $\ell_1$ optimization, and a quadratic program for $\mathcal{H}_2$ optimization. Optimization in $Q$-space rather than directly in the closed-loop space is mentioned in the final section, but does not play a significant role in the thesis. The design specifications are incorporated into the design in Chapter 4. This chapter is concerned primarily with linear constraints, because they can easily be appended to a linear or quadratic program. LMI constraints are briefly discussed at the end of this chapter, but the use of LMI constraints is left for future research. Chapter 5 explores the use of basis functions other than the FIR which may approximate the closed loop more efficiently than the FIR. Chapter 6 presents a complete controller design using constrained optimization methods for an active vibration isolation system. The controllers designed for this system were successfully implemented on a hardware testbed. Finally, the conclusions and suggestions for further research are presented in Chapter 7.
Chapter 2

Background

This chapter presents a brief overview of some well-known concepts necessary for the development of this thesis. It should serve as an aid to understanding the concepts and notation used in later chapters. Therefore, the reader may safely skip this chapter if the concepts are already familiar, and refer back to it as needed.

2.1 Signal Norms

The size of a signal can be measured by the calculation of one of its norms. The $p$-norm of an $n$-dimensional discrete signal $x[k]$ is defined as

$$
\|x\|_p = \left( \sum_{k=0}^{\infty} \sum_{i=1}^{n} |x_i[k]|^p \right)^{\frac{1}{p}}.
$$

When the $p$-norm of a signal is bounded, the signal exists in the space $\ell_p$. The norms that will be useful to this chapter are when $p = 1, 2, \infty$.

Perhaps the most widely used norm is the $\ell_2$ norm of a signal. This norm is defined by

$$
\|x\|_2 = \left( \sum_{k=0}^{\infty} x[k]^T x[k] \right)^{\frac{1}{2}}.
$$

The $\ell_2$ norm corresponds to the total amount of energy contained in a signal. Also related is the average power of a persistent signal (with infinite energy). This is also known as the root-mean-square (RMS) of the signal. Technically not a norm, the
RMS of a signal is defined as

\[ \| x \|_{\text{rms}} = \left( \lim_{N \to \infty} \frac{1}{N} \sum_{k=0}^{N-1} x[k]^T x[k] \right)^{\frac{1}{2}}. \]

For signals with large variations in amplitude, the RMS value may not be adequate as a measurement. In this case, it may be better to look at the \( \ell_\infty \) norm,

\[ \| x \|_\infty = \max \sup_{i,k} |w_i[k]|. \]

The \( \ell_\infty \) norm therefore represents the peak value the signal attains. This norm is useful to describe an unknown but bounded input signal, and to describe the maximum level achieved by an output signal.

The \( \ell_1 \) norm is important as a measurement of total resource consumption, e.g., if the signal magnitude represents the amount of fuel used per unit time. It is also an important measurement of a system impulse response, as will be discussed in the next section. The \( \ell_1 \) norm is defined as

\[ \| x \|_1 = \sum_{i=1}^{n} \sum_{k=0}^{\infty} |x_i[k]|. \]

### 2.2 System Norms

Systems are generally measured in terms of the norms of the input and output. One system measurement which will be important to this thesis is the average power output to a stochastic stationary input. For a MIMO system \( H \) with dimension \( n_z \times n_w \), and an input with power spectral density \( S_{ww}(e^{j\theta}) \), the output can be measured as

\[ \| H w \|_{\text{rms}} = \left( \text{Tr} \frac{1}{2\pi} \int_{-\pi}^{\pi} H(e^{j\theta}) S_{ww}(e^{j\theta}) H(e^{j\theta})^* d\theta \right)^{\frac{1}{2}}. \]  

(2.2)

If the input is unit variance white noise, then \( \| H w \|_{\text{rms}} \) is known as the \( \mathcal{H}_2 \) norm of the system \( H \), written \( \| H \|_2 \). In this case, \( S_{ww}(e^{j\theta}) = I \), and by Parseval’s theorem, Equation 2.2 becomes

\[ \| H \|_2 = \left( \text{Tr} \sum_{k=0}^{\infty} H[k] H[k]^T \right)^{\frac{1}{2}} = \left( \sum_{i=1}^{n_z} \sum_{j=1}^{n_w} \sum_{k=0}^{\infty} h_{ij}[k]^2 \right)^{\frac{1}{2}}. \]

(2.3)
where $H[k]$ is the matrix impulse response of $H$ with components $h_{ij}[k]$. Thus, the $\mathcal{H}_2$ norm of a system represents the RMS output given a white noise input, and can be determined by measuring the the $\ell_2$ norm of its output given a unit impulse input.

The $\mathcal{H}_\infty$ norm of a system is an induced norm, defined as

$$\|H\|_\infty = \sup_{w \neq 0} \frac{\|Hw\|_2}{\|w\|_2}.$$ 

This norm is the maximum $\ell_2/\ell_2$ system gain, and can be calculated by finding the supremum of the maximum singular values over all frequencies, i.e., $\|H\|_\infty = \sup_\omega \sigma(H(e^{i\omega T}))$.

The final system measurement which will be important to this thesis is the induced norm

$$\|H\|_1 = \sup_{w \neq 0} \frac{\|Hw\|_\infty}{\|w\|_\infty}.$$ 

This is the peak gain of the system over all possible bounded inputs, and can be found by calculating the $\ell_1$ norm of the system impulse response. To understand why the $\ell_1$ norm provides an upper bound on the system gain, consider a SISO system $h$.

Given an exogenous input with a known bound in magnitude $\|w\|_\infty \leq 1$, the system output will be bounded by

$$\|z\|_\infty = \sup_k \left| \sum_{j=0}^{\infty} h[k - j]w[j] \right| \leq \sum_{k=0}^{\infty} |h[k]| = \|h\|_1.$$ 

Therefore the system will have a bounded output if and only if the system pulse response $h \in \ell_1$, and the induced $\ell_\infty$ norm of $h$ will be bounded by its $\ell_1$ norm.

This guarantee can easily be extended to the MIMO case. Given the system $H \in \ell_1^{n_z \times n_w}$, and a bounded input disturbance $\|w\|_\infty \leq 1$, then the output $z$ is guaranteed to be bounded by

$$\|z\|_\infty \leq \|H\|_1,$$

where

$$\|H\|_1 = \max_{1 \leq i \leq n_z} \sum_{j=1}^{n_w} \sum_{k=0}^{\infty} |h_{ij}[k]|.$$
2.3 Small Gain Theorem

Figure 2.1 illustrates a feedback system between two systems, $H_1$ and $H_2$. The stability of this closed-loop system depends on the stability of the transfer function $(I - H_1H_2)^{-1}$. If both $H_1$ and $H_2$ are stable, then as long as the gain of $H_1H_2$ is less than unity at all frequencies (i.e., $||H_1H_2||_\infty < 1$), the closed loop is guaranteed to be stable. This is known as the Small Gain Theorem [9]. The Small Gain Theorem is not a necessary condition for stability, but is a good robustness condition if the phase of one of the transfer functions is unknown. An even more conservative test for stability can be applied if only the $\mathcal{H}_\infty$ norms of the systems are known. Because $||H_1||_\infty||H_2||_\infty \geq ||H_1H_2||_\infty$, the closed-loop system is known to be stable if $||H_1||_\infty||H_2||_\infty < 1$.

2.4 Parameterization of All Stabilizing Controllers

The standard control problem consists of a plant $P$ which maps exogenous input $w$ and control input $u$ to the regulated output $z$ and measurement $y$. A feedback compensator $K$ then closes the loop from the measurement to the control input. The plant can be represented by

$$
\begin{bmatrix}
z \\
y
\end{bmatrix} =
\begin{bmatrix}
P_{11} & P_{12} \\
P_{21} & P_{22}
\end{bmatrix}
\begin{bmatrix}
w \\
u
\end{bmatrix},
$$

The closed loop transfer function is

$$
P_{cl} = P_{11} + P_{12}K(I - P_{22}K)^{-1}P_{21} = \mathcal{F}_t(P, K)
$$
where $F(P, K)$ denotes the lower linear fractional transformation.

An important restriction on $K$ is that it must be internally stabilizing. To satisfy the internal stability requirement, $F(P, K)$ can be replaced by the well-known $Q$-parameterization

$$P_{cl} = T_1 + T_2 QT_3$$

(2.4)

in which $Q$ must be stable, but is otherwise arbitrary [18, 20], and $T_1$, $T_2$, and $T_3$ are described below. This parameterization has two advantages over the parameterization $F(P, K)$. First, the search space over all stabilizing controllers is replaced by a search over all stable $Q$. Secondly, the parameterization in Equation 2.4 is affine in $Q$, suggesting that it is good for optimization.

A $Q$-parameterization can be derived from any nominally stabilizing controller. If the plant $P$ has the state space description

$$x[k+1] = Ax[k] + B_1 w[k] + B_2 u[k]$$
$$z[k] = C_1 x[k] + D_{11} w[k] + D_{12} u[k]$$
$$y[k] = C_2 x[k] + D_{21} w[k] + D_{22} u[k]$$

then a stabilizing model-based controller can be constructed by choosing a controller gain matrix $F$ and an observer gain matrix $H$ such that all the eigenvalues of $A - B_2 F$ and $A - HC_2$ are inside the unit circle. Let $\hat{x}$ denote the states of the model-based controller. Adding an input $v$ and an output $e$ to the controller, where $u = -F \hat{x} + v$ and $e = C_2(x - \hat{x})$, the model-based controller can be augmented with any stable system $v = Qe$ without affecting the stability of the closed-loop system. Figure 2.2 illustrates the augmented closed loop. For this system it can be shown that the transfer function from $v$ to $e$ is zero, so $Q$ affects the closed-loop system only in the forward loop. This results in the affine representation in Equation 2.4, where

$$T_1 = \text{Transfer function from } w \text{ to } z \text{ of nominal closed-loop system}$$
$$T_2 = \text{Transfer function from } w \text{ to } e$$
$$T_3 = \text{Transfer function from } v \text{ to } z.$$
The state space representation for the augmented controller $K_s$ is

$$
\dot{x}[k + 1] = (A - B_2F - HC_2 + HD_{22}F)\dot{x}[k] + Hy[k] + (B_2 - HD_{22})v[k]
$$

$$
u[k] = -F\dot{x}[k] + v[k]
$$

$$
e[k] = -(C_2 - D_{22}F)\dot{x}[k] + y[k] - D_{22}v[k].
$$

The space of all stabilizing controllers is spanned by $K = \mathcal{F}_r(K_s, Q)$ such that $Q$ is stable. Notice that when the open-loop plant is stable, the controller and observer gain matrices in $K_s$ can go to zero. This reduces the $Q$-parameterization to $T_1 = P_{11}, T_2 = P_{12}$, and $T_3 = P_{21}$.

### 2.5 Rational Matrix Factorization

In Section 3.2, it will be necessary to use a well-known rational matrix factorization, the Smith-McMillan decomposition [18]. This factorization is defined in this section, as applied to a rational MIMO system. It is useful for characterizing the zeros in a MIMO system, which in addition to a complex number, also have a direction associated with them. The Smith-McMillan decomposition of an $n \times m$ system $G(z)$ with
rank \( r \) is found by factoring \( G(z) \) into

\[
G(z) = L(z)M(z)R(z)
\]

where \( L(z) \) and \( R(z) \) have full rank independent of \( z \), and

\[
M(z) = \begin{bmatrix}
\frac{\epsilon_1(z)}{\psi_1(z)} & 0 & \cdots & 0 \\
\vdots & \ddots & \ddots & \vdots \\
\frac{\epsilon_r(z)}{\psi_r(z)} & 0 & \cdots & 0 \\
0 & \cdots & 0 & 0 \\
\vdots & \ddots & \ddots & \ddots \\
0 & \cdots & 0 & 0 & \cdots & 0
\end{bmatrix}
\]

is an \( m \times n \) matrix in Smith-McMillan form. The polynomials \( \epsilon_i(z) \) and \( \psi_i(z) \) are coprime for each \( i \), and satisfy the following divisibility property: \( \epsilon_i \) divides \( \epsilon_{i+1} \) without remainder, and \( \psi_{i+1} \) divides \( \psi_i \) without remainder. The zeros of \( G(z) \) are found to be the roots of \( \prod_i \epsilon_i(z) \), and the poles are the roots of \( \prod_i \psi_i(z) \). Define the multiplicity of each zero \( z_0 \) in \( \epsilon_i(z) \) as \( \sigma_i(z_0) \). Then the total multiplicity of a zero in \( G(z) \) is \( \sum_{i=1}^{r} \sigma_i(z_0) \).
Chapter 3

Controller Synthesis in the Time Domain

For many types of control engineering problems, closed-loop optimization in the time domain provides a distinct advantage over the standard frequency domain techniques. Many specifications given for a closed-loop system are placed on the magnitude of the output signals based on a characteristic input signal. Common examples are peak overshoot to a step input, rise and settling time, or maximum output to an unknown but bounded input. In addition, specifications given in the frequency domain can also be placed in the time domain, such as the output amplitude to a sinusoidal input. If the closed-loop optimization is in the time domain, consideration of time domain specifications become much more straightforward. This chapter presents a control design algorithm which models the closed loop as a finite impulse response (FIR). Most of this methodology is based on the $\ell_1$ design algorithm developed by Dahleh and Diaz-Bobillo [8].

3.1 Performance Objectives

The optimality of a feedback controller requires that a particular performance metric of the closed-loop system is optimal, e.g. minimized. Section 2.2 covered several system norms which provide a way to measure the performance of a controller design.
This section shows the objectives of the $\mathcal{H}_2$ and $\ell_1$ control problems in terms of the closed-loop impulse response.

### 3.1.1 $\mathcal{H}_2$ Optimization

If the disturbance entering a system is unknown but has a known Gaussian power spectral density, then a shaping filter can be appended to the input to create a new system with a white noise input. When the input disturbance is a unit variance white noise, it is often desirable to minimize the amount of power seen at the output due to the disturbance. This is equivalent to minimizing the expected value of the output 2-norm, $E(\|z\|_2)$, which in turn is the $\mathcal{H}_2$ norm of the closed-loop system. From Equation 2.3, the $\mathcal{H}_2$ minimization problem can be stated as

$$J_{\mathcal{H}_2 \text{ opt}} = \inf \|H\|_2^2 = \inf \left( \sum_{i=1}^{n_z} \sum_{j=1}^{n_w} \sum_{k=0}^{\infty} |h_{ij}[k]|^2 \right). \quad (3.1)$$

Frequency-domain weights are an effective tool for forcing the optimization to concentrate on a particular frequency band while placing less emphasis on other bands. This is done by minimizing the $\mathcal{H}_2$ norm of an appended system $\|WH\|_2$, where $W$ is a stable filter. The weighted $\mathcal{H}_2$ optimization problem is then

$$J_{\mathcal{H}_2 \text{ opt}} = \inf \left( \sum_{i=1}^{n_z} \sum_{j=1}^{n_w} \sum_{k=0}^{\infty} (w_i * h_{ij})[k]^2 \right). \quad (3.2)$$

where $w_i[k]$ is the impulse response of the frequency weight on output $i$.

### 3.1.2 $\ell_1$ Optimization

When the probabilistic nature of a disturbance is unknown, but the signal is known to be bounded in magnitude, i.e. $\|w\|_\infty \leq 1$, then the performance of the closed-loop system can be measured by considering the maximum possible magnitude of the regulated output, $\|z\|_\infty$. It was demonstrated in Section 2.2 that this is bounded by the $\ell_1$ norm of the closed-loop system. The $\ell_1$ control problem can therefore be stated as

$$J_{\ell_1 \text{ opt}} = \inf \|H\|_1 = \inf \left( \max_i \sum_{j=1}^{n_w} \sum_{k=0}^{\infty} |h_{ij}[k]| \right). \quad (3.3)$$
3.2 Feasibility Constraints

From Section 2.4, all achievable closed-loop transfer functions are represented by the
Q-parameterization

\[ H = \{T_1 + T_2 QT_3 \mid Q \text{ stable}\}. \]  

(3.4)

For this problem formulation, it is desirable to find the constraints on the impulse
response \( h \) such that \( H \) can be written in the form of Equation 3.4. There are
two important sets of conditions placed on the transfer function \( R = T_2 QT_3 \) which
guarantee that \( H \) is feasible. The first set of conditions is that the non-minimum phase
zeros of \( T_2 \) and \( T_3 \) must be preserved in \( R \) to guarantee the stability of \( Q \). These
constraints are known as the the zero interpolation conditions. If the problem is one-
block, meaning there are the same number of measurements as disturbances and the
same number of controls as regulated outputs, then the zero interpolation conditions
are the only feasibility constraints. If the problem is not one-block, then \( Q \) has limited
degrees of freedom. There may be fewer measurements than disturbances \( (n_y < n_w) \),
or there may be fewer controls than regulated outputs \( (n_u < n_z) \). Individually both
of these cases are called two-block problems. If both \( n_y < n_w \) and \( n_u < n_z \), then
the problem is considered four-block. This places a limitation on the structure of \( R \),
which has dimension \( n_z \times n_w \), while \( Q \) only has dimension \( n_u \times n_y \). The conditions
to ensure this are the rank interpolation conditions.

This section shows that it is possible to pose these conditions as linear constraints
on the impulse response, but does not explicitly show how these constraints should be
calculated in practice. The methods used in this section rely on the Smith-McMillan
decomposition of certain transfer matrices, which can present numerical difficulties
when implemented on a digital computer. A numerically stable method for construct-
ing the interpolation conditions which circumvents the Smith-McMillan decomposi-
tion can be found in [10].
3.2.1 Zero Interpolation Conditions

It is easiest to understand the motivation behind the zero interpolation conditions by considering a SISO example, where \( h(z) = t_1(z) + t_2(z)q(z) \). If \( r(z) = t_2(z)q(z) \), then the condition that \( r(z) \) must satisfy to insure that \( h(z) \) is feasible is that \( q(z) = r(z)/t_2(z) \) must be stable. This means that any non-minimum phase zero \( z_0 \) in \( t_2(z) \) must also be present in \( r(z) \), i.e., \( r(z_0) = 0 \). This imposes the constraint on \( h(z) \), that \( h(z_0) = t_1(z_0) \). If the non-minimum phase zero has multiplicity \( \sigma \), then this condition must also be placed on the first \( \sigma - 1 \) derivatives. Therefore, \( \left( \frac{d^n h}{dz^n} \right)(z_0) = \left( \frac{d^n t_1}{dz^n} \right)(z_0) \) for \( n = 0, \ldots, (\sigma - 1) \).

The MIMO case is not as simple, and it is helpful to use the Smith-McMillan decomposition of the transfer matrices \( T_2 \) and \( T_3 \), as developed in Section 2.5:

\[
T_2 = L T_2 M T_2 R T_2 \\
T_3 = L T_3 M T_3 R T_3
\]

where

\[
M_{T_2} = \begin{bmatrix}
\xi_1 \\
\psi_1 \\
\vdots \\
\vdots \\
\psi_{nu}
\end{bmatrix}
\]

\[
M_{T_3} = \begin{bmatrix}
\xi_1 \\
\psi_1 \\
\vdots \\
\vdots \\
\psi_{ny}
\end{bmatrix}
\]

The zero interpolation conditions will now be stated without proof. (A proof and more complete treatment of the following result can be found in [8].) Define \( R = T_2 QT_3 \). If \( \epsilon_i \) and \( \epsilon'_j \) have a non-minimum phase zero \( z_0 \) of multiplicity \( \sigma_i(z_0) \) and \( \sigma'_j(z_0) \), then the \((i, j)\) entry of the matrix \( \left( \frac{d^n h}{dz^n} L_{T_2}^{-1} R R_T^{-1} \right)(z_0) \) must also be zero for \( n = 0, \ldots, (\sigma_i(z_0) + \sigma'_j(z_0) - 1) \). To create a more compact statement of this
condition, define the following row and column vectors:

\[ a_i(z) = (L_{T_2}^{-1})(z) \quad i = 1, 2, \ldots, n_z \]
\[ \beta_j(z) = (R_{T_3}^{-1})(z) \quad j = 1, 2, \ldots, n_w \]  

(3.5)

where the indicial notation \( T_{(i,:)} \) denotes the \( i \)th row of matrix \( T \) and \( T_{(:,j)} \) denotes the \( j \)th column of matrix \( T \). Now the zero interpolation condition for each non-minimum phase zero can be written as

\[
\left( \frac{d^n}{dz^n} \alpha_i R \beta_j \right)(z_0) = 0 \text{ for } \begin{cases} i = 1, \ldots, n_u \\
j = 1, \ldots, n_y \\
n = 0, \ldots, \sigma_i(z_0) + \sigma_j'(z_0) - 1 \end{cases}
\]  

(3.6)

It is now important to show that these condition can be written as set of linear constraints on the closed-loop impulse response \( h_{ij}[k] \). Because \( R = H - T_1 \), Equation 3.6 can be rewritten as

\[
\frac{d^n}{dz^n}(\alpha_i H \beta_j)(z_0) = \frac{d^n}{dz^n}(\alpha_i T_1 \beta_j)(z_0)
\]  

(3.7)

The product \( \alpha_i H \beta_j \) is equivalent to

\[
(\alpha_i H \beta_j)(z_0) = \text{Tr}[H \beta_j \alpha_i](z_0) = \sum_{p=1}^{n_x} \sum_{q=1}^{n_w} (H_{pq} \beta_{qj} \alpha_{ip})(z_0) = \sum_{p=1}^{n_x} \sum_{q=1}^{n_w} \sum_{k=0}^{\infty} (\alpha_{ip} \beta_{qj})(z_0) h_{pq}[k] z_0^{-k}
\]

This can be substituted back into Equation 3.7 to obtain

\[
\frac{d^n}{dz^n} \left( \sum_{p=1}^{n_x} \sum_{q=1}^{n_w} \sum_{k=0}^{\infty} \alpha_{ip} \beta_{qj} h_{pq}[k] z_0^{-k} \right) \bigg|_{z=z_0} = \frac{d^n}{dz^n}(\alpha_i T_1 \beta_j)(z_0)
\]  

(3.8)

The zero interpolation constraints are now written as a set of linear constraints on the closed-loop impulse response \( h_{ij}[k] \). It is useful to vectorize this impulse response as

\[ h \equiv [h_{11}[0] \cdots h_{1n_w}[0] \quad h_{21}[0] \cdots h_{2n_w}[0] \cdots h_{nzn_w}[0] \quad h_{11}[1] \cdots]^T. \]  

(3.9)

With this definition, it is possible to rearrange the zero interpolation conditions in Equation 3.8 into a single linear operation:

\[ A_{\text{zero}} h = b_{\text{zero}}. \]
3.2.2 Rank Interpolation Conditions

The rank interpolation conditions preserve the structure of $R = T_2QT_3$ if the problem is multiblock (two or four-block). From the Smith-McMillan decomposition of $T_2$ and $T_3$ introduced in the previous section, the $i$th row of $M_{T_2}$ is zero for $n_u < i \leq n_z$, and the $j$th column of $M_{T_3}$ is zero for $n_y < j \leq n_w$. This means that the $i$th row of $(L_{T_2}^{-1}R)(z)$ should equal zero for $n_u < i \leq n_z$, and the $j$th column of $(RR_{T_3}^{-1})(z)$ should equal zero for $n_y < j \leq n_w$. Using the definitions of $\alpha$ and $\beta$ in Equation 3.5, these conditions can be stated as

$$
\begin{align*}
(\alpha_i R)(z) &= 0 \text{ for } i = n_u + 1, \ldots, n_z \\
(R\beta_j)(z) &= 0 \text{ for } j = n_y + 1, \ldots, n_w
\end{align*}
$$

(3.10)

By transforming from the $z$ to the time domain, the rank interpolation conditions appear as an infinite number of constraints on the impulse response of $R$, $r_{ij}[k]$:

$$
\sum_{p=1}^{n_z} \sum_{l=0}^{k} \alpha_{iq}[k - l]r_{pq}[l] = 0 \text{ for } \begin{cases} i = n_u + 1, \ldots, n_z \\ q = 1, \ldots, n_w \\ k \geq 0 \end{cases}
$$

(3.11)

$$
\sum_{q=1}^{n_w} \sum_{l=0}^{k} \beta_{pq}[k - l]r_{pq}[l] = 0 \text{ for } \begin{cases} j = n_y + 1, \ldots, n_w \\ p = 1, \ldots, n_z \\ k \geq 0 \end{cases}
$$

Then, by replacing $r_{ij}[k]$ with $h_{ij}[k] - (T_1)_{ij}[k]$, the zero interpolation conditions become a set of linear constraints on $h_{ij}[k]$ and can be written in the following form:

$$
A_{\text{rank}} h = b_{\text{rank}}.
$$

Notice that the constraints in Equation 3.11 are over all $k \geq 0$, which leads to an infinite number of constraints. Therefore, the range of $A_{\text{rank}}$ is infinite. This leads to practical difficulties when trying to impose these constraints using a computer. This problem is resolved in the next section, by means of the delay augmentation method. This method approximates the multiblock problem with a one block problem, therefore removing the need for rank interpolation conditions.
3.2.3 Delay Augmentation Method

The delay augmentation method was developed by Diaz-Bobillo in [10] as a method for approximating a multiblock problem as a one block problem. In doing this, the need for rank interpolation conditions is avoided, and only a finite number of interpolation constraints are needed to solve the problem. In preparation for the delay augmentation method, the general multiblock closed-loop system must be partitioned as follows:

\[
\begin{bmatrix}
H_{11} & H_{12} \\
H_{21} & H_{22}
\end{bmatrix} =
\begin{bmatrix}
T_{1,11} & T_{1,12} \\
T_{1,21} & T_{1,22}
\end{bmatrix} +
\begin{bmatrix}
T_{2,1} \\
T_{2,2}
\end{bmatrix} Q[T_{3,1} T_{3,2}],
\quad (3.12)
\]

where the dimensions of \(T_{2,1}\) and \(T_{3,1}\) are \(n_u \times n_u\) and \(n_y \times n_y\). To transform this into a one-block problem, the parameter \(Q\) must have dimensions \(n_z \times n_w\). To achieve this, \(Q\) is augmented with extra parameters, and \(T_2\) and \(T_3\) are augmented with \(N\) pure delays:

\[
\begin{bmatrix}
H_{11,N} & H_{12,N} \\
H_{21,N} & H_{22,N}
\end{bmatrix} =
\begin{bmatrix}
T_{1,11} & T_{1,12} \\
T_{1,21} & T_{1,22}
\end{bmatrix} +
\begin{bmatrix}
T_{2,1} & 0 \\
T_{2,2} & S_N
\end{bmatrix} \begin{bmatrix}
Q_{11} & Q_{12} \\
Q_{21} & Q_{22}
\end{bmatrix} \begin{bmatrix}
T_{3,1} \\
T_{3,2}
\end{bmatrix},
\quad (3.13)
\]

where \(S_N = I_{z^{-N}}\) is the system of appended delays.

The system \(H_N\) is now one-block, and \(\inf \|H_N\|\) can be solved without any rank interpolation conditions. Of course, \(Q_{12}, Q_{21},\) and \(Q_{22}\) are not present in the real system, and can be used as extra degrees of freedom in the optimal solution to Equation 3.13. Therefore, \(\inf \|H_N\|\) provides a lower bound to the optimal solution of Equation 3.12. That is,

\[
\inf_{Q \text{ stable}} \|H_N\| \leq \inf_{Q_{11} \text{ stable}, Q_{12}=Q_{21}=Q_{22}=0} \|H_N\| = \inf_{Q_{11} \text{ stable}} \|H\|. \quad (3.14)
\]

Also, an upper bound on the optimal solution is obtained by using the \(Q_{11}\) from the optimal solution to Equation 3.13, and setting the rest of the parameters to zero so the solution is feasible. For the \(\ell_1\) control problem, it was proven in [10] that as \(N \to \infty\), the lower and upper bounds converge on the optimal solution. It is assumed here that the \(H_2\) control problem should show similar behavior. This convergence can be seen
intuitively by considering the expansion of Equation 3.13:

\[ H_N = T_1 + T_2 Q_{11} T_3 + S_N R_N \]  
\[ (3.15) \]

where

\[ R_N = \begin{bmatrix} 0 & T_{21} Q_{12} \\ Q_{21} T_{31} & Q_{21} T_{32} + T_{22} Q_{12} + S_N Q_{22} \end{bmatrix}. \]

Notice that the \( H_{11,N} \) partition is unaffected by the parameters \( Q_{12}, Q_{21}, \) and \( Q_{22}. \) Also, these extra parameters will not affect any partition of \( H_N[k] \) for \( k < N \) due to the time delay in \( S_N R_N \) appearing in Equation 3.15. Therefore, the augmented parameters in \( Q \) have a decreasing effect on the closed loop as the augmented time delay is increased.

As a final point on the delay augmentation method, it should be noted that while there are no longer an infinite number of rank interpolation conditions, there are extra zero interpolation conditions due to the introduction of delays in \( T_2 \) and \( T_3. \) Because an \( N \)th order delay has the transfer function \( z^{-N} \), there are also \( N \) non-minimum phase zeros at infinity. These zeros must be accounted for in the interpolation conditions to guarantee that \( Q \) will be causal.

### 3.3 Solution to the \( \ell_1 \) and \( \mathcal{H}_2 \) Problems

The \( \ell_1 \) design problem can be easily posed as a linear program. Because of the nonlinearity (the absolute value) in Equation 3.3, a change of variables is necessary. Let \( h_{ij}[k] = h_{ij}^+[k] - h_{ij}^-[k], \) where \( h_{ij}^+[k] \geq 0 \) and \( h_{ij}^-[k] \geq 0. \) While this definition does not restrict \( h_{ij}^+[k] \) or \( h_{ij}^-[k] \) to be zero depending on the sign of \( h_{ij}[k], \) the magnitude of \( h \) can be found by

\[ |h_{ij}[k]| = \min_{h_{ij}^+[k], h_{ij}^-[k]} (h_{ij}^+[k] + h_{ij}^-[k]). \]

Using this notation, define the objective function as

\[ \mu = \max_i \sum_{j=1}^{n_w} \sum_{k=0}^{\infty} (h_{ij}^+[k] + h_{ij}^-[k]). \]
It should be understood that $\mu$ is not necessarily the $\ell_1$ norm of $H$. However, when $\mu$ is minimized, either $h^+_ij[k]$ or $h^-_ij[k]$ will be zero for every $(i, j, k)$. In this case, it is evident that $\mu$ will be the $\ell_1$-norm of $H$. Therefore, the optimal (minimum) $\ell_1$-norm of the closed loop has the property $\|H\|_1 \text{optimal} = \inf \mu$.

Next, using the vectorized closed-loop impulse response $h$ defined in Equation 3.9, define the linear operator $A_{\ell_1}$ such that

$$(A_{\ell_1}h)_i = \sum_{j=1}^{n_w} \sum_{k=0}^{\infty} h_{ij}[k] \text{ for } i = 1, \ldots, n_z.$$ 

Then the objective function is written as

$$\mu = \max_i (A_{\ell_1}(h^+ + h^-))_i.$$ 

The problem can now be posed as a linear program:

$$\|H\|_1 \text{optimal} = \inf_{\mu, h^+, h^-} \mu$$

subject to

$$[A_{\ell_1} A_{\ell_1}] \begin{bmatrix} h^+ \\ h^- \end{bmatrix} \leq 1\mu$$

$$\begin{bmatrix} A_{\text{zero}} & -A_{\text{zero}} \\ A_{\text{rank}} & -A_{\text{rank}} \end{bmatrix} \begin{bmatrix} h^+ \\ h^- \end{bmatrix} = \begin{bmatrix} b_{\text{zero}} \\ b_{\text{rank}} \end{bmatrix}$$

$$[h^+ h^-]^T \geq 0$$

where $1$ is a vector of ones with dimension $n_w$.

Similarly, the $H_2$ design problem can be posed as a convex quadratic program. Using the vector representation of the MIMO impulse response in Equation 3.9, the $H_2$ norm becomes $h^T h$. If there are frequency weights, then the $H_2$ norm is instead $h^T \tilde{h} = (w \ast h)^T (w \ast h)$. The weighted impulse response $\tilde{h}$ can be constructed as follows:

$$\begin{bmatrix} \tilde{h}[0] \\ \tilde{h}[1] \\ \tilde{h}[2] \\ \vdots \end{bmatrix} = \begin{bmatrix} w[0] & 0 & 0 & \cdots \\ w[1] & w[0] & 0 & \cdots \\ w[2] & w[1] & w[0] & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{bmatrix} \begin{bmatrix} h[0] \\ h[1] \\ h[2] \\ \vdots \end{bmatrix}.$$
Therefore $\tilde{h} = Wh$ where $W$ is a matrix operator, and the objective function is simply the quadratic $h^T W^T W h$. The quadratic program takes the form

$$\|H\|_2^{\text{optimal}} = \inf_h h^T W^T W h$$

subject to

$$A_{\text{zero}} h = b_{\text{zero}}$$
$$A_{\text{rank}} h = b_{\text{rank}}.$$  \hfill (3.17)

Immediately it should be recognized that neither Equation 3.16 nor Equation 3.17 are practical to solve. There are infinitely many variables, due to the infinite dimension of $h$. Additionally, if the problem is multiblock, there are also infinitely many constraints from the infinite range of $A_{\text{rank}}$. However, an approximate solution is within the reach of a computer. As was seen in Section 3.2.3, the delay augmentation method can be used to reduce the infinite number of constraints with a finite number of constraints. Alternative methods to delay augmentation also exist, and can be found in [8].

The infinite dimension of $h$ can be reduced to a finite dimension by assuming that $h_{ij}[k] = 0$ for $k > N$, where $N$ is a finite number of time steps. In fact, it can be proven using Duality Theory that this is in fact the case for the one-block $\ell_1$ control problem. This proof appears in [8]. Unfortunately, there is no equivalent guarantee for the $H_2$ optimal solution. However, it is not unreasonable to approximate the terms in the impulse response after a certain time as zero, because the impulse response of an $H_2$ optimal solution will be stable, and therefore will have an exponential decay towards zero.

Once the $\ell_1$ and $H_2$ problems are approximated as finite dimensional linear and quadratic programs, they can be readily solved using standard techniques. The Simplex Method still proves to be the most effective method for solving linear programs, by searching for the optimal solution along the constraint boundaries [6]. The quadratic program can also be efficiently solved if the quadratic objective function is convex, which will be the case as long as the matrix $W_T W$ is positive definite. This problem can be solved using the methods in [14].
### 3.4 Optimization in $Q$-space

In Section 2.4, it was noted that the $Q$-parameterization is affine in $Q$, meaning that if $Q_1$ and $Q_2$ each yield a feasible closed loop, then so does $\lambda Q_1 + (1 - \lambda)Q_2$ for every $\lambda \in \mathbb{R}$. This useful result does not hold for a parameterization that is made in the closed loop space. Just because $H_1$ and $H_2$ are achievable does not imply that $\lambda H_1 + (1 - \lambda)H_2$ will be achievable. However, the space of all achievable closed loops is convex, meaning that feasibility is guaranteed if $\lambda \in [0, 1]$. Thus the closed-loop representation used in this chapter (its impulse response) must be used with a set of convex specifications to insure feasibility. These are the feasibility constraints introduced in Section 3.2.

It follows that if the closed loop is represented as a linear combination of stable $Q$ transfer matrices, that there will no longer be a need for feasibility constraints. That is, represent $H$ with a set of variables $x \in \mathbb{R}^N$ with

$$H = T_1 + T_2 \sum_{n=1}^{N} x_nQ_nT_3,$$  \hspace{1cm} (3.18)

where \{Q_n\} is a sequence of stable transfer matrices. This is exactly the approximation used by Boyd and his colleagues in [2]. This paper describes a software package they have developed, called QDES, which approximates $Q$ as an FIR filter as follows:

$$Q_{ijk}(z) = E_{ij}z^{-k}, \hspace{0.5cm} 1 \leq i \leq n_u, \hspace{0.2cm} 1 \leq j \leq n_y, \hspace{0.2cm} 0 \leq k \leq N,$$

where $E_{ij}$ is a matrix with the $(i, j)$ entry as one and all the others zero.

The use of the approximation in Equation 3.18 is not explored further in this thesis, but would be worthwhile to compare with the other techniques presented in this chapter. Optimization in $Q$-space does offer several advantages over optimization in the closed-loop space. Removing the need for feasibility conditions reduces the complexity of the problem, especially for multiblock problems. Also, the optimization is only over the number of transfer functions in $Q$ rather than the number of transfer functions in $H$, which may have many more inputs and outputs. However, the optimization problem is still inherently infinite dimensional. Because this problem must be reduced to a finite dimensional subspace, the subspace cannot be chosen
arbitrarily. It was not clear what the best subspace in $Q$ would be, although an FIR basis would certainly not be a bad start. It is left as an open question whether it is more effective to solve for the closed loop in terms of $Q$ or to solve directly in the closed-loop space.
Chapter 4

Linear Design Specifications

The amount of computation required to find the solution to a constrained optimization problem can be greatly improved if the constraints are known to be convex. In a convex optimization problem, once a local minimum is found, it is guaranteed to be the global minimum. This is one reason why the $Q$-parameterization introduced in Section 2.4 is significant. This parameterization is affine in $Q$, and the space of all stable $Q$ yields the space of all feasible closed-loop transfer functions. Thus the most important design specification, feasibility, is also convex. An extensive treatment of the various kinds of closed-loop convex constraints is given in [1].

This chapter will focus on a particular subset of convex design constraints, namely linear constraints. Often, a set of convex constraints can be reasonably approximated with a finite set of linear constraints. For example, it was seen in Section 3.2 that feasibility can be approximated with a set of linear constraints. Linearity allowed the feasibility constraints to be written in the framework of a linear program for the $\ell_1$ control problem, and a quadratic program for the $H_2$ control problem. Both of these techniques converge on an answer much more rapidly than other techniques which can handle non-convex constraints. Most controller design problems will have additional design specifications beyond feasibility, many of which are convex. This chapter will present several typical design specifications that are often used, and show in detail how they can be posed as linear constraints.
4.1 Time-Domain Constraints

In Chapter 3, the closed loop was optimized in terms of its impulse response. This formulation makes it very convenient to introduce constraints on the closed loop in the time domain. A constraint on the impulse response at a particular time has the form \( g_l[k] \leq h_{ij}[k] \leq g_u[k] \) where \( g_l[k] \) and \( g_u[k] \) are the lower and upper bounds. In engineering practice, it is often more useful to look at the step response of a closed-loop system. Typical examples of step response specifications are overshoot and undershoot, asymptotic tracking, rise time and settling time. The step response can easily be derived by integrating the impulse response, and is therefore available to constrain [8]. In fact, the response to any particular fixed input \( w_j \) can be constrained in the same way, and is found by taking the convolution \( z_i = h_{ij} * w_j \). In order to put the time domain constraints into a useful form, they must be written as

\[
A_{t\text{ime}} h \leq b_{t\text{ime}},
\]

where \( h \) is defined in Equation 3.9. Equation 4.1 can then be appended to the linear and quadratic programs developed in Section 3.3.

It should be noted that if a problem is multiblock and the delay augmentation method is implemented, once the extra parameters in \( Q \) are thrown away the time domain constraints may no longer be satisfied in the solution. If the constraints were placed on \( H_{11} \), then the constraints will still be satisfied. This is because the extra parameters in \( Q \) do not affect the first block. However, if the constraints were placed on the \( H_{12} \), \( H_{21} \), or \( H_{22} \) blocks, then the constraints are guaranteed to be satisfied only for times less than the number of delays used. The reader should refer back to Section 3.2.3 for a more detailed explanation.

4.1.1 Example

Magnetic bearing control in aircraft engines provides an excellent example of the use of time-domain constraints. Magnetic bearings offer potential advantages to aircraft engine technology by allowing the rotor shaft to spin in a virtually frictionless environment. An important design requirement on the magnetic bearing controller is that
it should not allow the shaft to come into contact with the bearing walls under normal operating conditions. This event is most likely to happen as a transient response to an aircraft maneuver. This problem therefore has hard constraints on the amplitude of the output due to a specific disturbance input.

The shaft position in the magnetic bearing has two degrees of freedom which are assumed to be uncoupled for this example. The system inputs are the control and disturbance forces, and the output is the position. The rigid-body dynamics for one axis can therefore be modeled as a double integrator. With a 0.001 sec time step, the plant dynamics are represented in discrete time as

\[
y(z) = 3.667 \times 10^{-3} \frac{z^2}{(z - 1)^2} (w(z) + u(z))
\]

where the input is in units of pounds and the output is in mils.

The most extreme disturbance that may occur during a maneuver is shown in Figure 4.1. This disturbance profile is based on an aircraft maneuver which produces an 11g acceleration at the bearing. Under these conditions, the maximum allowable deflection of the shaft is 6 mils.
Figure 4.2: Response to fixed input for design with time-domain constraints.

This problem was solved as an $\mathcal{H}_2$ minimization problem. Although Chapter 3 developed the solution to this problem using an FIR basis, hundreds of terms in the FIR would have been required to reach a feasible solution, and thousands required to reach the desirable one. As will be seen in Chapter 5, the $\mathcal{H}_2$ problem is not restricted to using the FIR basis, and another more efficient basis set was chosen for the solution. A Laguerre basis (see Section 5.2) with a time scale of 0.9 and 120 free coefficients was chosen. Both the position output and control effort were weighted in the objective function. Unity scalar weights were placed on both position and control. Time-domain constraints were then placed on the output due to the disturbance in Figure 4.1, restricting the output to be between +6 and −6 mils.

The resulting transient is shown in Figure 4.2. Note that at 2 seconds, the signal actually violates the constraints. Closer inspection of the plot will show that the signal actually crept between two discrete time domain constraints (the x’s), so technically no constraint was violated. This shows a limitations of this method, which can only handle constraints at a finite number of points.
4.2 Frequency-Domain Constraints

Design specifications are frequently placed on the magnitude of an output due to a persistent sinusoidal input at a particular frequency. Consider the bound in magnitude on the SISO system $H$:

$$|H(e^{j\omega T})| \leq \gamma$$

This bound in magnitude is equivalent to the constraint

$$\Re[H(e^{j\omega T})]\cos \theta + \Im[H(e^{j\omega T})]\sin \theta \leq \gamma \quad \forall \theta \in [0, 2\pi).$$

In [8], this constraint is approximated by a finite number of linear constraints on the real and imaginary parts of $H(e^{j\omega T})$ by only considering a discrete number of angles $\theta$ evenly spaced between 0 and $2\pi$. In turn, the real and imaginary parts of $H(e^{j\omega T})$ are linear functions of $h[k]$:

$$\Re[H(e^{j\omega T})] = \sum_{k=0}^{\infty} h[k]\cos(k\omega T)$$
$$\Im[H(e^{j\omega T})] = -\sum_{k=0}^{\infty} h[k]\sin(k\omega T)$$

Finally, the frequency constraint can be written as

$$\sum_{k=0}^{\infty} h[k]\cos(k\omega T + \theta_n) \leq \gamma \text{ where } \theta_n = \{2\pi n/N \mid n = 0, 1, \ldots, N - 1\}. \quad (4.2)$$

The frequency domain constraints are guaranteed to be satisfied if the problem is one-block. If the problem is multiblock, and the delay augmentation method is used, then only the constraints placed in the first block, $H_{11}$, are guaranteed to be met. This is due to the introduction of the extra free parameters, $Q_{12}, Q_{21},$ and $Q_{22}$, which are not present in the actual system. If frequency constraints are required on an input-output pair that is not in the $H_{11}$ block, then these constraints should be reflected onto the $H_{11}$ block. Consider the constraint $|H_{(i,j)}(e^{j\omega T})| \leq \gamma$, where $H_{(i,j)}$ represents the SISO transfer function from input $j$ to output $i$, and the input-output pair $(i, j)$ does not exist in the $H_{11}$ block. This new notation is introduced to prevent confusion with the notation $H_{ij}$, which represents the $ij$ partition and is not necessarily SISO. As before, an equivalent constraint is $\Re[H_{(i,j)}(e^{j\omega T})]\cos \theta + \Im[H_{(i,j)}(e^{j\omega T})]\sin \theta \leq \gamma$. 


To guarantee that the constraint on \( H_{(i,j)} \) is satisfied, it is necessary to reflect it as a constraint on the \( n_u \times n_y \) transfer function \( H_{11} \). \( H_{11} \) and \( H_{(i,j)} \) are written as

\[
H_{11} = T_{1,11} + T_{2,1}Q T_{3,1} \\
H_{(i,j)} = T_{1(i,j)} + T_{2(i,:)\mu}Q T_{3(:,j)}
\]

The indicial notation \( T_{(i,:)} \) denotes the \( i \)th row of transfer matrix \( T \), and \( T_{(:,j)} \) denotes the \( j \)th column of \( T \). Solving for \( Q \) in terms of \( H_{11} \), and then substituting back into the expression for \( H_{(i,j)} \) yields

\[
H_{(i,j)} = T_{1(i,j)} - T_{2(i,:)\mu}T_{21}^{-1}T_{1,11}T_{3,1}^{-1}T_{3(:,j)} + T_{2(i,:)\mu}T_{21}^{-1}H_{11}T_{3,1}^{-1}T_{3(:,j)}. \ 
\]

(4.3)

Note that this expression is to be evaluated at a particular frequency, so only complex matrices, rather than MIMO transfer functions, will have to be inverted. Equation 4.3 is exactly the expression needed to pose constraints on \( H_{(i,j)} \) in terms of \( H_{11} \). However, it is not yet in a useful form. The next step is to express Equation 4.3 at a particular frequency as a linear function of the impulse response of \( H_{11} \). To do this, first arrange \( H_{11} \) into a \( n_u n_y \times 1 \) dimensional vector transfer function as follows:

\[
\tilde{H}_{11} = \begin{bmatrix} H_{1,1} \\ \vdots \\ H_{1,n_y} \\ H_{2,1} \\ \vdots \\ H_{n_u,n_y} \end{bmatrix} 
\]

Using this new notation for the \( H_{11} \) block, \( H_{(i,j)} \) can be written in the form

\[
H_{(i,j)} = L + M \tilde{H}_{11}
\]

where \( L \) is a SISO transfer function and \( M \) is a \( 1 \times n_u n_y \) transfer function. The next step is to evaluate each transfer function at the frequency \( \omega \), and separate into real and imaginary parts. Define \( L_R(\omega) = \Re[L(e^{i\omega T})] \) and \( L_I(\omega) = \Im[L(e^{i\omega T})] \). \( M \), and \( \tilde{H}_{11} \) have similar definitions. \( H_{(i,j)} \) can now be broken into

\[
\Re[H_{(i,j)}(e^{i\omega T})] = L_R(\omega) + M_R(\omega)\tilde{H}_{11,R}(\omega) - M_I(\omega)\tilde{H}_{11,I}(\omega) \\
\Im[H_{(i,j)}(e^{i\omega T})] = L_I(\omega) + M_R(\omega)\tilde{H}_{11,I}(\omega) + M_I(\omega)\tilde{H}_{11,R}(\omega)
\]
Next, let $\tilde{h}[k]$ be the vector impulse response of $\tilde{H}_{11}$. From this definition, $	ilde{H}_{11,R}(\omega) = \sum_{k=0}^{\infty} \tilde{h}[k] \cos(k\omega T)$ and $\tilde{H}_{11,I}(\omega) = -\sum_{k=0}^{\infty} \tilde{h}[k] \sin(k\omega T)$. The frequency constraint can now be written in a convenient form:

$$\sum_{k=0}^{\infty} [M_R(\cos(k\omega T) \cos \theta_n - \sin(k\omega T) \sin \theta_n) + M_I(\cos(k\omega T) \sin \theta_n + \sin(k\omega T) \cos \theta_n)] \tilde{h}[k] \leq \gamma - L_R \cos \theta_n - L_I \sin \theta_n. \quad (4.4)$$

Equation 4.4 is now written purely in terms of the impulse response of $H_{11}$, and will therefore be unaffected by the introduction of the terms $Q_{12}, Q_{21},$ and $Q_{22}$ in the delay augmentation method. There is another useful feature of this equation. If there is an output which needs to be constrained but does not appear in the objective function, then the frequency constraints can be posed in terms of other variables already present in the problem.

Both Equations 4.2 and 4.4 are linear constraints on the vector $h$, and can be written in the form $A_{freq}h \leq b_{freq}$. This inequality is easily incorporated into the linear and quadratic programs developed in Chapter 3. Examples of the use of frequency constraints in controller design will come in subsequent sections.

### 4.2.1 Example

Although the transient in Figure 4.2 from the example in Section 4.1 does meet the time-domain constraints, there is still something unsatisfying about this solution. There appears to be a significant steady-state error to the disturbance input. This problem could be corrected by placing additional time-domain constraints which force the response to decay towards zero. This problem can also be approached in the frequency domain by forcing the closed loop to exhibit integral control behavior. By forcing the closed loop gain $|P/(1 + PK)| \rightarrow 0$ as $e^{\omega T} \rightarrow 1$ at 20 dB/decade, the steady-state error due to a step input will be zero, and the steady-state error due to a ramp input will be bounded. This can be done with frequency constraints placed on $P/(1 + PK)$ at low frequency.

Figure 4.3 shows the effect of placing these constraints on the closed-loop frequency response. For comparison, both the cases with and without the time domain
Figure 4.3: Frequency response for design with frequency-domain constraints.

constraints are shown. Figure 4.4 shows the resulting response to the disturbance in Figure 4.1.

4.3 Stability Margin

The stability margin of a closed-loop system is the amount of perturbation in the loop transfer function the system can handle before it will become unstable [11]. For a SISO controller $K$, the stability margin is the minimum distance from the Nyquist plot of the loop gain $PK$, where $P$ is the plant dynamics, to the critical point $-1$. This distance, $d$, can be calculated as

$$d = \inf_{\omega} |P(e^{i\omega T})K(e^{i\omega T}) + 1|$$

$$= \left\{ \sup_{\omega} \left[ \frac{1}{1 + P(e^{i\omega T})K(e^{i\omega T})} \right] \right\}^{-1}$$

$$= ||S||_{\infty}^{-1},$$
where $S$ is the sensitivity transfer function $1/(1 + PK)$. An example of the minimum distance $d$ is shown in Figure 4.5.

Two metrics of stability margin common in classical control are gain margin and phase margin. Gain margin represents the minimum amount of change in gain of a plant necessary to destabilize the closed loop. The upper gain margin is the minimum increase in gain which will induce instability, and the lower gain margin is the minimum decrease in gain. Gain margin is written in decibels as $20 \log_{10}(|(PK)(\omega_{G.M.})|^{-1})$, where $\omega_{G.M.}$ is the gain margin frequency. At $\omega_{G.M.}$, the phase of the loop gain is $180^\circ$, therefore the sensitivity is $1/(1 - |PK|)$. From this, the gain margin as a function of sensitivity is

$$G.M. = 20 \log_{10} \left[ \frac{S(\omega_{G.M.})}{S(\omega_{G.M.}) - 1} \right] \quad (4.5)$$

Similarly, phase margin represents the amount of change in phase necessary to induce instability. At the frequency which determines phase margin, $\omega_{P.M.}$, the loop gain is 0 dB, therefore the sensitivity is $1/(1 + e^{i\theta})$, where $\theta$ is the phase of $PK$. The distance between $PK$ and $-1$ at this point is $1/|S|$, and the angle between these
two points from the origin is the phase margin, so from simple geometry, the phase margin is calculated to be

$$P.M. = 2 \sin^{-1} \left( \frac{1}{2|S(\omega_{P.M.})|} \right).$$

(4.6)

Figure 4.6 graphically shows this relationship between sensitivity magnitude and gain and phase margin. This provides a convenient way to incorporate gain and phase margin specifications into the controller design problem. By referring to Figure 4.6, an appropriate sensitivity magnitude for a given gain and phase margin constraint can be chosen. Then by using the method presented in Section 4.2, the sensitivity can be constrained to stay below these magnitudes near the gain and phase margin frequencies.

### 4.4 Robust Stability

In the previous section, the minimum distance from the Nyquist plot to the critical point was shown to be $\|S\|_{\infty}^{-1}$. Therefore, as long any perturbation of the loop gain has magnitude less than this distance, the system will be stable. This is a conservative test, and can be improved by developing a frequency dependent model of the plant uncertainty. One such model is the multiplicative uncertainty model $P_{\text{pert}} = P(1 + \ldots$
Figure 4.6: Gain and Phase Margins versus Sensitivity Magnitude.

\[ \Delta(\omega)W(\omega) \], where \( W(\omega) \) is a frequency dependent weight, and \( \Delta(\omega) \) is an uncertain, stable transfer function such that \( \|\Delta\|_\infty \leq 1 \). This is known as an unstructured uncertainty model, and is well suited for capturing unmodeled dynamics. Figure 4.7 illustrates a feedback system with a multiplicative uncertainty model. Ignoring the input and output, this system can be condensed into the configuration in Figure 4.8, where \( C = PK/(1 + PK) \), which is the complimentary sensitivity. According to the small-gain theorem (see Section 2.3), the system in Figure 4.8 is guaranteed to be stable if and only if the transfer function \( \|WC\|_\infty < 1 \). This leads to a useful result. Stability robustness constraints can be brought into the controller design by imposing the following set of frequency constraints:

\[ |C(\omega)| \leq \left| \frac{1}{W(\omega)} \right| \quad \forall \omega. \]

4.4.1 Example

The magnetic bearing controllers designed in Sections 4.1.1 and 4.2.1 were based only on the rigid-body dynamics of the rotor. In reality, the rotor also has flexible dynamics...
Figure 4.7: Feedback control with multiplicative uncertainty.

Figure 4.8: Condensed loop with uncertainty.
which can destabilize the closed loop. These structural modes are time-varying (the modal frequencies shift with the normal operating rotational speed variations as well as when the aircraft engine is powered up). Therefore, it is desirable to be robust to a variety of different flexible mode dynamics.

Figure 4.9 shows the plant dynamics at one stage of operation. Associated with this is an uncertainty over other possible locations of the plant modes. The nominal plant is the double integrator, and this can be combined with the uncertainty to form a multiplicative uncertainty model. The inverse of the multiplicative uncertainty forms the constraints for the transfer function $PK/(1 + PK)$, which is illustrated in Figure 4.10. Imposing these constraints on the closed loop has little impact on the transient response, so the conservatism of the unstructured uncertainty model has not hurt this design. More aggressive constraints may be used to force the compensator to roll off, if that is desired. Constraints on the stability margins were not required in this design, because already the gain margins were at +10 dB and -13 dB, and the phase margin was 50°.
Figure 4.10: Frequency response of $PK/(1 + PK)$ for design with robustness constraints.

As a final point on the magnetic bearing control design, in all cases the compensator found turned out to be unstable. Although it is certainly possible to implement an unstable compensator, most control engineers are much more comfortable with using a stable compensator instead. Therefore, it would be nice to place additional constraints on the closed loop to guarantee stability of the compensator. Unfortunately, controller stability (also known as strong stability) is a non-convex constraint [1]. This means that either a non-convex optimization routine must be used, which is frequently not computationally practical, or the non-convex region of feasible solutions must be restricted to a convex sub-region.

4.5 Linear Matrix Inequalities

So far, it has been seen that control problems with linear design specifications can be efficiently solved through linear and quadratic programming. However, most convex
specifications are not linear. One solution may be to approximate the specifications with a set of linear constraints, as was done in Section 4.2 with the frequency constraints. This can be very inefficient, especially if the constraint needs to be accurate. Fortunately, there are efficient algorithms that have recently emerged which can handle a wider range of convex constraints. Nesterov and Nemirovskii have developed interior-point methods for convex optimization problems which involve linear matrix inequality (LMI) constraints [19]. Using these methods, one common LMI problem which may be solved is

\[
\begin{align*}
\min_{x \in \mathbb{R}^m} &\quad c^T x \\
\text{subject to} &\quad F(x) = F_0 + \sum_{i=1}^{m} x_i F_i > 0
\end{align*}
\]

where \( F_i = F_i^T \in \mathbb{R}^{n \times n} \). The inequality constraint means that \( F(x) \) must be positive definite, which is a convex constraint on \( x \). When \( F_i \) is diagonal, then the constraints are reduced to a set of linear inequalities, therefore all the constraints introduced in this chapter can be put into the LMI framework.

A constraint on the maximum singular value of a matrix \( Z(x) \in \mathbb{R}^{p \times q} \) may easily be posed as an LMI:

\[
\begin{bmatrix}
I & Z(x) \\
Z(x)^T & I
\end{bmatrix} > 0 \iff I - Z(x)^T Z(x) > 0 \iff \sigma(Z(x)) < 1
\]

This would have been difficult to approximate with a set of linear constraints, especially if the dimension of \( Z(x) \) is large. This leads to a very compact representation of the frequency constraint \(|H(e^{j\omega T})| < \gamma\). Recall that this was approximated with a set of linear constraints as

\[
\Re[H(e^{j\omega T})] \cos \theta + \Im[H(e^{j\omega T})] \sin \theta < \gamma \text{ where } \theta_n = \{2\pi n/N \mid n = 0, 1, \ldots, N - 1\}.
\]

Figure 4.11 shows this approximation for \( N = 8 \). If the constraint is instead written
as the LMI
\[
\begin{bmatrix}
\gamma & 0 & \Re[H(e^{j\omega T})] \\
0 & \gamma & \Im[H(e^{j\omega T})] \\
\Re[H(e^{j\omega T})] & \Im[H(e^{j\omega T})] & \gamma
\end{bmatrix} > 0
\]
then it is an exact constraint on the magnitude of $H(e^{j\omega T})$, as shown in Figure 4.12.

In the LMI framework, frequency constraints are not limited to a finite number of frequency points, but a constraint may be placed on an entire transfer function at once. For a transfer function $H(z) = C(zI - A)^{-1}B + D$, it can be proved that $\|H\|_\infty < \gamma$ if and only if there exists a matrix $X = X^T$ that satisfies the LMI
\[
\begin{bmatrix}
A^TXA - X & A^TXB & C^T \\
B^TXA & B^TXB - \gamma I & D^T \\
C & D & -\gamma I
\end{bmatrix} < 0.
\]
This is useful in developing a solution to the mixed $\mathcal{H}_2/\mathcal{H}_\infty$ control problem [4], and the mixed $\ell_1/\mathcal{H}_\infty$ control problem [5]. The application of LMI constraints to a variety of problems in system and control theory is explored in detail in [3].
Chapter 5

Efficient Basis Functions for Closed Loop Approximations

The optimization techniques used in this thesis (linear and quadratic programming) rely on the representation of the closed loop as a finite number of linear coefficients. Any transfer function can be represented with an infinite set of coefficients. Consider a SISO transfer function $H(z)$. Given a sequence of orthonormal transfer functions $\{F_n(z)\}$, $H(z)$ can be written as

$$H(z) = \sum_{n=0}^{\infty} \phi_n F_n(z), \quad (5.1)$$

where $\{\phi_n\}$ is the corresponding sequence of independent coefficients. Until now, the only type of model used has been the finite impulse response (FIR), where $F_n(z) = z^{-n}$, and $\phi_n = 0$ for $n > N$, where $N$ is some finite integer. This model can be used to approximate stable transfer functions. Because the closed loop must be stable, the terms in the impulse response will become negligibly small after a finite time. Additionally, it has been proven that the optimal $\ell_1$ closed loop for one-block problems is FIR [7], so after a finite time all the terms will be exactly zero. The downside to this model is that for some problems, a large number of terms may still be required in the optimal solution. This is especially true for structural systems with lightly damped modes. In this case, thousands of terms in the impulse response may be needed to approximate the optimal closed-loop impulse response. The amount of
computational effort required to find the solution is determined in part by the number of coefficients to solve for, so it is desirable to keep this number as small as possible.

Because Equation 5.1 uses an infinite number of coefficients, some knowledge of the closed loop must be brought along to approximate this space with a finite subspace. In the case of the FIR model, the only knowledge used was the fact that the closed loop is stable, which still may require a large subspace. At the other extreme, knowledge of the closed loop may include knowledge of the exact location of all the poles. In this case the number of coefficients required is the number of poles, and the coefficients become the residues of the poles. The point is that a set of basis functions, such as the FIR, may be very inefficient in its representation of the optimal closed-loop transfer function and may require a large number of coefficients for a good approximation. Some knowledge of the closed loop can lead to a much more efficient set of basis functions which may offer a better alternative to the FIR basis. This chapter explores several alternative sets of basis functions, and shows under what conditions they will be more efficient than the FIR basis.

5.1 Fixed Pole Model

The FIR model approximates a transfer function with

$$H(z) \approx \sum_{n=0}^{N} \frac{\phi_n}{z^n}.$$ 

Therefore, $H(z)$ can be approximated by an $N$th order model with all its poles at the origin. In [17], the Fixed Pole Model (FPM) was developed to reduce the number of parameters needed in the FIR by placing the poles closer to their actual location rather than at the origin. The general FPM is defined as

$$H(z) \approx \frac{\sum_{n=0}^{N-1} \phi_n z^n}{A(z)}, \quad (5.2)$$

where $A(z) = \sum_{n=0}^{N-1} a_k z^n$ is a polynomial containing the approximate location of the poles of $H(z)$. The accuracy possible using this model depends on the order $N$ and the closeness of the roots of $A(z)$ to the actual poles. A reasonable approximation of
the pole locations can substantially reduce the number of parameters needed in the sequence \( \{\phi_n\} \) over the number of parameters needed in the FIR model.

To demonstrate the increased efficiency of the FPM over the FIR model, consider the transfer function

\[
H(s) = \frac{(s + 10)(s^2 + 30s + 325)}{(s^2 + 50s + 1850)(s^2 + 10s + 4025)}.
\]

Converting to discrete time using a zero order hold with a sampling time of 0.005 sec, the poles of the transfer function become \( z = 0.5270 \pm 0.8207i, 0.8690 \pm 0.1536i \). The pole locations for the FPM are chosen to be \( z = 0.51 \pm 0.79i, 0.83 \pm 0.15i \), which is within five percent of their actual locations. The denominator of the FPM becomes

\[
A(z) = (z^4 - 2.6800z^3 + 3.2888z^2 - 2.1934z + .6290)^n,
\]

where the order of the FPM is \( 4n \). The objective is to find the best approximation of the plant \( H(z) \) over all the coefficients \( \phi_n \). This is found by solving the minimization problem

\[
\phi_{opt} = \arg \min_{\phi \in \mathbb{R}^N} \| H(z) - \sum_{n=0}^{N-1} \phi_n \left( \frac{z^n}{A(z)} \right) \|_2.
\] (5.3)

This \( \mathcal{H}_2 \) minimization problem will be useful in subsequent sections to evaluate the efficiency of different basis functions, so the solution will be developed here. Let the FPM equal \( \phi^T F \) where \( \phi \) is the vector of free coefficients and \( F \) is an \( N \times 1 \) transfer function. The state-space representation of \( F \) can be written as \((A_F, B_F, I, 0)\) with

\[
A_F = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \vdots \\
0 & 0 & 0 & \cdots & 1 \\
-a_0 & -a_1 & -a_2 & \cdots & -a_{N-1}
\end{bmatrix}, \quad B_F = \begin{bmatrix} 0 \\
0 \\
\vdots \\
0 \\
1 \end{bmatrix}
\]

The steady-state term \( D_F \) term is neglected here because it can always be set exactly to the steady state of \( H(z) \). The transfer function \( H - \phi^T F \) can then be written in state-space representation as

\[
H - \phi^T F \sim \begin{bmatrix}
A_H & 0 & B_H \\
0 & A_F & B_F \\
C_H & -\phi^T & 0
\end{bmatrix} = \begin{bmatrix} A \\ B \\ C \end{bmatrix}. \]

57
The \( H_2 \) norm \( \| H - \phi^T F \|_2 \) can now be calculated as

\[
\| H - \phi^T F \|_2^2 = \text{Tr}(C P C^T) = [C_H - \phi^T] P_1 P_{12} \begin{bmatrix} C_H^T \\ P_{21} P_{22} \end{bmatrix} \begin{bmatrix} \phi \\
-\phi \end{bmatrix}
\]  

where \( P \) is found by solving the discrete Lyapunov Equation

\[
A P A^T + B B^T = P.
\]

The minimization problem in Equation 5.3 can be easily solved by finding the \( \phi \) which minimizes Equation 5.4. Setting the derivative (with respect to \( \phi \)) of Equation 5.4 to zero leads to the result

\[
\phi_{\text{opt}} = P_{22}^{-1} P_{21} C_H^T.
\]

Figure 5.1 shows the approximation of \( H(z) \) using an increasing number of states in the FPM. The coefficients \( \phi_n \) are plotted in Figure 5.2 for the 20th order FPM. For comparison, Figure 5.3 shows the impulse response of \( H(z) \) to demonstrate the number of terms which would be necessary for an FIR approximation.
Figure 5.2: Coefficients $\phi_n$ for a 20th order FPM.

Figure 5.3: Impulse response of $H(z)$. 
The FPM is excellent for capturing the dynamics of systems with lightly damped modes if the locations of the poles are approximately known. However, the FPM basis functions are highly non-orthogonal. Each basis function \( \frac{z^n}{A(z)} \) contains the same dynamics, separated only by a time delay. This means that an FPM with \( N \) coefficients may not span \( N \) dimensions, or may span this space very poorly. Due to its non-orthogonality, a large dimensional FPM can lead to numerical problems in the optimization. For this reason, the FPM as it is presented in this section has limited application to the constrained optimization control problem.

## 5.2 Laguerre Functions

It was seen in the last section that a set of basis functions that approximates a transfer function more efficiently than the FIR can be found if the poles in the basis are moved away from the origin and closer to the actual location of the transfer function poles. However, the major problem encountered with the FPM was that once new locations for the poles of the basis were chosen, the basis functions were no longer orthogonal. Orthonormality is not a necessary characteristic of the basis functions, but it is desirable because it leads to numerically more reliable solutions. The Laguerre functions satisfy this property, and are a viable alternative to the FIR.

In the \( z \) domain, the Laguerre functions take the form

\[
F_n(z) = \sqrt{1 - a^2} \frac{(1 - az)^{n-1}}{z - a}.
\]

where \( a \) is the time scale, such that \( |a| < 1 \). A detailed discussion of the discrete Laguerre functions can be found in [16]. \( F_n(z) \) is all-pass (has unity magnitude at all frequencies), consisting of \( n \) stable poles at \( a \) and \( n - 1 \) non-minimum phase zeros at \( \frac{1}{a} \). It will be seen shortly that \( a \) can be chosen to reflect the dynamics of the transfer function to be approximated. When \( a = 0 \), the Laguerre basis functions become the FIR basis. Therefore, given the right information about the transfer function, it should be possible to choose \( a \) so the transfer function is approximated better or with less terms than the FIR.
In the time domain, the Laguerre functions can be written as

\[ f_n[k] = \sqrt{1 - a^2} \left[ \sum_{j=0}^{n} (-1)^{n+j} \binom{n}{j} \binom{k + n - j}{n} a^{n+k-2j} u[k - j] \right], \]

where \( u[k] \) is the unit step function. The first four Laguerre functions are illustrated in Figure 5.4 for \( a = 0.8 \). These functions are orthonormal, meaning they satisfy

\[ \sum_{k=0}^{\infty} f_i[k] f_j[k] = \delta_{ij}, \]

where \( \delta_{ij} \) is the Kronecker delta. The Laguerre functions can be calculated recursively using the following relations:

\[ f_0[0] = \sqrt{1 - a^2} \]
\[ f_{n+1}[0] = -a f_n[0] \]
\[ f_n[k + 1] = a f_n[k] + \eta \sum_{i=0}^{n-1} (-a)^{n-i-1} f_i[k], \]

with \( \eta = 1 - a^2 \). This leads to the state-space representation of transfer function \( F \).
as \((A_F, B_F, I, 0)\), with

\[
A_F = \begin{bmatrix}
a & 0 & 0 & \cdots & 0 \\
\eta & a & 0 & \cdots & 0 \\
-a\eta & \eta & a & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & 0 \\
(-a)^{N-2}\eta & (-a)^{N-3}\eta & \cdots & \eta & a
\end{bmatrix}
\]

\[
B_F = \begin{bmatrix}
\eta \\
-a\eta \\
\vdots \\
(-a)^{N-2}\eta \\
(-a)^{N-1}\eta
\end{bmatrix}
\]

A given plant \(H(z)\) can be approximated using the Laguerre functions as

\[
H(z) \approx \sum_{n=0}^{N-1} \phi_n F_n(z) = \phi^T F,
\]

where \(\phi\) is a vector of the coefficients \(\phi_n\). Using the method for minimizing \(\|H - \phi^T F\|_2\) outlined Section 5.1, the best possible approximation (in the \(\mathcal{H}_2\) norm sense) of \(H\) using the basis transfer function \(F\) can be found.

When constructing the Laguerre functions, it is important to choose the time scale \(a\) appropriately. There always exists a time scale that will approximate the transfer function as well or better than the FIR (\(a\) can always be chosen to be zero, which will yield the FIR basis). However, a poorly chosen time scale can actually perform worse than the FIR. For any plant, there exists an optimal time scale. This time scale was derived in [13]. Given a SISO transfer function \(H\), with impulse response \(h[k]\), define the quantities

\[
M_1 = \frac{1}{\|h\|^2} \sum_{k=0}^{\infty} kh^2[k], \quad M_2 = \frac{1}{\|h\|^2} \sum_{k=0}^{\infty} k(\Delta h[k])^2,
\]

where \(\Delta h[k] = h[k+1] - h[k]\). Then the optimum \(a\) is found to be

\[
a_{opt} = \frac{2M_1 - 1 - M_2}{2M_1 - 1 + \sqrt{4M_1M_2 - M_2^2 - 2M_2}}.
\]

From the example developed in Section 5.1, the optimal time scale for the transfer function \(H(z)\) is \(a_{opt} = 0.293\). The optimal Laguerre model shows modest improvement over the optimal FIR model. For a 50th order approximation, the optimal error \(\inf \|H - \phi^T F\|_2\) is 3.50 for Laguerre, whereas for the FIR it is 4.46 (the error for the 20th order FPM was 0.19). For this example, a set of curves which show the \(\mathcal{H}_2\) norm of the error for various approximation orders and time scales is shown in Figure 5.5.
Intuitively, the Laguerre functions offer the most improvement over the FIR model for transfer functions where the optimal time scale is close to one. As an example, consider a plant with the poles and zeros given in Table 5.1. This plant has an optimal time scale of $a = 0.928$. Figure 5.6 shows the optimal 50 state approximation using the Laguerre functions and an FIR model. The 2-norm of the error with the Laguerre approximation was 0.1413 compared to 69.07 for the FIR model. Figure 5.7 shows the error for other time scales and approximation orders.

5.3 General Orthonormal Basis Functions

The three types of basis functions described above (FIR, FPM, and Laguerre) have shortcomings which hinder their use in many applications. Of all, the set of Laguerre functions is the most flexible, because it is an orthonormal set which can efficiently deal with a wide range of dynamics. However, the Laguerre functions only allow the placement of the poles at one location. If the dynamics of the transfer function are
Table 5.1: Pole and zero locations for Laguerre basis example.

<table>
<thead>
<tr>
<th>Poles</th>
<th>Zeros</th>
</tr>
</thead>
<tbody>
<tr>
<td>.6588 ± .4021i</td>
<td>.9981 ± .0060i</td>
</tr>
<tr>
<td>.8986 ± .2669i</td>
<td>.7410 ± .2644i</td>
</tr>
<tr>
<td>.9764 ± .0607i</td>
<td>.7127 ± .3312i</td>
</tr>
<tr>
<td>.9924 ± .0170i</td>
<td>1.0</td>
</tr>
<tr>
<td>.8496</td>
<td>−1 (×2)</td>
</tr>
</tbody>
</table>

Figure 5.6: Comparison of 50 state Laguerre model and FIR model.
known to a certain degree of accuracy, then it may be desirable to incorporate these specific dynamics into the basis. Despite its non-orthonormality, the FPM showed the greatest efficiency by being able to handle specific dynamics. This section shows how an arbitrary set of basis functions, such as the FPM basis, can be made orthonormal.

Consider that a set of $N$ basis functions have been constructed which span a desirable space, but are not orthonormal. These basis functions can be put into an $N \times 1$ transfer function $F(z)$ with a state-space representation $(A, B, C, D)$. $F(z)$ can be expanded into its impulse response using the geometric series as

$$F(z) = C(Iz - A)^{-1}B + D = \sum_{k=0}^{\infty} CA^k B z^{-(k+1)} + D$$

Next, define the impulse response matrix

$$R = [D \ CB \ CAB \ CA^2B \ CA^3B \cdots CA^K B].$$

for some large $K$, such that the terms in $CA^k B$ are negligibly small for $k > K$. It is desired to find a linear transformation which will make the rows of $R$ orthonormal.
This can easily be done by finding the singular value decomposition of $R$:

$$R = U\Sigma V^T = \begin{bmatrix} \sigma_1 & 0 & \vdots & \sigma_N \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \cdots & \cdots & \cdots \end{bmatrix} \begin{bmatrix} u_1 & \cdots & u_N \end{bmatrix} \begin{bmatrix} v_1 \\ \vdots \\ v_N \end{bmatrix}.$$  

The $N$ columns in $U$ span the column space of $R$ for $\sigma_n \neq 0$ and span the left nullspace for $\sigma_n = 0$. Similarly, the rows of $V^T$ span the row space of $R$ for $\sigma_n \neq 0$ and the nullspace for $\sigma_n = 0$. Furthermore, the columns of $U$ and the rows of $V^T$ are orthonormal. Define $V_m^T$ to be a matrix of the first $m$ rows of $V^T$, where $m$ is the number of non-zero singular values of $R$ (i.e., the rank of $R$). Then $R = U_m\Sigma_m V_m^T$, where $U_m$ is a matrix of the first $m$ columns of $U$, and $\Sigma_m$ is a diagonal matrix consisting of the non-zero singular values. $V_m^T$ has orthonormal rows which span the row space of $R$, therefore the rows of $V_m^T$ would make a good set of basis functions. $V_m^T$ is found to be

$$V_m^T = S R = [SD \ SCB \ SCAB \ SCA^2B \ SCA^3B \cdots SCA^KB].$$

where

$$S = (\Sigma_m^T U_m^T U_m \Sigma_m)^{-1} \Sigma_m^T U_m.$$  

A set of basis functions with orthonormal impulse response $V_m^T$ can now be found with state-space representation $(A, B, SC, SD)$.

This orthogonalization technique is extremely useful for blending together different sets of basis functions, or constructing entirely new ones. For example, Section 5.2 showed that a correctly chosen time scale could lead to an efficient set of basis functions. It is reasonable to assume that even more efficient basis functions are possible by blending together two sets of Laguerre functions with two different time scales. Blending together different sets of low order FPMs with different dynamics also has a strong potential. By doing this, it would be possible to construct an orthonormal set of basis functions with many poles scattered over the unit circle. Even if nothing is known about the optimal closed-loop dynamics, a set of basis functions like this may be a much better starting point than the FIR basis.
5.4 Basis Functions in the Constrained Optimization Control Problem

The solution to the constrained optimization control problem using the FIR basis was presented in Chapter 3. This basis is a logical starting point for representing the closed loop, but there are also many other sets of basis functions that might also be used. As was seen in the previous section, the number of coefficients needed to closely approximate a particular transfer function depends greatly on that transfer function and on the basis selected. Because the problem must be reduced from an infinite dimensional to a finite dimensional optimization problem, it is important to choose the subspace carefully. The FIR is not necessarily an efficient representation of the closed loop, and if any additional information is known about the desired closed loop (before actually solving for it), then this information can be used to select a more efficient basis. For example, if there are a set of poles in the open-loop plant which should not change very much in the desired closed loop, then these poles can be added to the basis. This is particularly useful for lightly-damped modes at high frequencies where the controller gain should be small. The Laguerre functions are also useful if the closed loop dynamics are relatively unknown but known to be slow (or significantly slower than the discrete time step of the controller). If necessary, different types of basis functions can be blended together as long as they are all made orthonormal, which can be done using the steps outlined in Section 5.3. The following subsections describe how to transform the constrained optimization problem to accept an arbitrary basis.

5.4.1 Performance Objective Functions

The two types of objective functions explored in Chapter 3 were the system $\ell_1$ and $\mathcal{H}_2$ norms. The $\ell_1$ norm is a nonlinear function of the impulse response, which presents a difficulty when placing the $\ell_1$ norm as the objective function of a linear program. For the FIR basis, this was solved by separating each term in the impulse response into
$h^+$ and $h^-$ (see Section 3.3). This is a particular feature of the FIR basis, and cannot be used with arbitrary basis functions. For this reason, the use of basis functions other than the FIR basis for the $\ell_1$ control problem will not be explored in this thesis.

On the other hand, the $\mathcal{H}_2$ norm can easily be computed as a quadratic of the basis function coefficients for an arbitrary basis. Given that a set of basis functions is orthonormal (if the functions are not, then they can be made orthonormal as in Section 5.3), then the square of the system $\mathcal{H}_2$ norm is $\phi^T\phi$, where $\phi$ is a vector containing the basis function coefficients. The vector $\phi$ can represent a MIMO transfer function, and is constructed in a similar way that the vector $h$ was constructed in Equation 3.9.

Frequency-domain weights can be added to the cost functional by appending a stable transfer function $W$ to the output, so that $z = WHw$. The impulse response of $WH$ is found by the convolution $\tilde{h} = w_f * h$, where $w_f[k]$ is the impulse response of $W$. In matrix form, the convolution is written as

$$
\begin{bmatrix}
\tilde{h}[0] \\
\tilde{h}[1] \\
\tilde{h}[2] \\
\vdots
\end{bmatrix} =
\begin{bmatrix}
w_f[0] & 0 & 0 & \cdots \\
w_f[1] & w_f[0] & 0 & \cdots \\
w_f[2] & w_f[1] & w_f[0] & \cdots \\
\vdots & \vdots & \vdots & \ddots
\end{bmatrix}
\begin{bmatrix}
h[0] \\
h[1] \\
h[2] \\
\vdots
\end{bmatrix}.
$$

In turn, the vector $h$ is a function of the basis coefficients $\phi$ as follows:

$$
\begin{bmatrix}
h[0] \\
h[1] \\
\vdots
\end{bmatrix} =
\begin{bmatrix}
f_0[0] & \cdots & f_{N-1}[0] \\
f_0[1] & \cdots & f_{N-1}[1] \\
\vdots & \vdots & \vdots
\end{bmatrix}
\begin{bmatrix}
\phi_0 \\
\phi_1 \\
\vdots
\end{bmatrix},
$$

where $f_n[k]$ is the impulse response of the basis function $F_n(z)$. Thus, the frequency-weighted impulse response $\tilde{h}$ can be written as a function of $\phi$ in matrix form as $\tilde{h} = M\phi$. The quadratic objective function $\tilde{h}^T\tilde{h}$ is then $\phi^TMTM\phi$, where $MTM$ is an $N \times N$ positive definite matrix.
5.4.2 Time-Domain Constraints

Time domain constraints can be placed on the coefficients $\phi_n$ very similar to the way frequency-domain weights were introduced to the objective function. An output $z$ due to a specific disturbance $w$ can be found by the convolution $z[k] = (h \ast w)[k]$, where the impulse response $h[k]$ is a linear function of the coefficients $\phi_n$ as described earlier. The time-domain constraints are then written as

$$g_l[k] \leq \sum_{n=0}^{N-1} (w \ast f_n)[k] \phi_n \leq g_u[k]$$

where $g_l[k]$ and $g_u[k]$ are the lower and upper bounds of $z[k]$.

5.4.3 Frequency-Domain Constraints

It was seen in Section 4.2 that constraints on the magnitude of a system $H$ can be posed as a set of linear constraints:

$$\Re[H(e^{i\omega T})] \cos \theta + \Im[H(e^{i\omega T})] \sin \theta \leq \gamma \quad \forall \theta \in [0, 2\pi).$$

When the system is represented by the sum of a finite number of basis functions $\phi_n F_n(z)$, this set of constraints is equivalent to

$$\sum_{n=0}^{N-1} (\Re[F_n(e^{i\omega T})] \cos \theta + \Im[F_n(e^{i\omega T})] \sin \theta) \phi_n \leq \gamma \quad \forall \theta \in [0, 2\pi). \quad (5.5)$$

Equation 5.5 only applies to frequency constraints placed within the $H_{11}$ block of the closed loop system. As described in Section 4.2, any frequency constraints which are placed in the $H_{12}$, $H_{21}$, or $H_{22}$ blocks of the problem must be reflected back to the $H_{11}$ block. Section 4.2 also gives a method for doing this. This method can be used with an arbitrary set of basis functions simply by substituting the real and imaginary parts of $F_n(e^{i\omega T})$ in the appropriate places.
Chapter 6

Application of Constrained Optimization Methods to an Active Vibration Isolation System

The Active Vibration Isolation System (AVIS) is an ongoing project at Draper Laboratories to demonstrate technology in active structural control. It is part of a larger study investigating the design of a spaceborne optical system. Telescope optics require a highly accurate alignment to be effective. In the space environment, this accuracy can be disrupted by the spacecraft attitude control system, thermal deformations, or anything else capable of producing a vibrational disturbance. This problem is being approached through the use of a high bandwidth structural controller and a low bandwidth optical controller. The optical control system involves the control of independent segments of the primary mirror to correct for low frequency (below \( \sim 0.1 \) Hz) wavefront distortion. The purpose of the structural control system, which includes AVIS, is to reduce high frequency errors in the wavefront and position of the telescope image on the detector due to structural vibrations.

The telescope structural control system consists of both AVIS, which is designed to isolate the telescope optics from the disturbance environment, and an active damping control system, which reduces any residual vibrations seen in the telescope. AVIS consists of six active struts which mount the telescope to the spacecraft. Linear models
for individual struts from the strut actuators to the strut sensors were provided for
control design. Therefore, all the controller designs for this project were SISO, and
coupling between the struts was neglected.

This chapter presents a complete design for AVIS using the methodologies dis-
cussed in the preceding chapters. The design was carried out over two phases. In
the first phase, controllers were designed using detailed linear models of the struts.
Although these controllers were stable with the design model as well as with the
frequency response function (FRF) data taken directly from the hardware, coupling
between struts as well as nonlinearities in the actual hardware presented difficulties in
implementing the controllers. Most of the discrepancies between the models and the
hardware were at high frequency, near the region of crossover. The second design and
test phase used simplified models which did not include any high frequency modes.
These controller designs were necessarily more conservative, but were much easier to
implement in the hardware.

6.1 Structural Test Model

The structural test model (STM) was developed by Eastman Kodak as a hardware
test bed for telescope structural control systems [15]. A diagram of this test bed
is shown in Figure 6.1. The optical control system was implemented on a different
set of test hardware, so the primary and secondary mirrors on the STM are actually
mass simulators. The major components of the STM are the Aft Metering Structure
(AMS), the K tubes, and the main mounts. The AMS houses the primary mirror,
and supports the K tubes, to which the secondary mirror assembly is mounted. The
primary mirror is 102 inches in diameter, which is roughly the size of the Hubble
Space Telescope aperture. The main mounts are the active struts, which connect the
AMS to a large granite block.

The active struts were developed under Draper subcontract by the Jet Propulsion
Laboratory. They each contain a piezoelectric actuator and two load sensors. The
actuator is contained within an overload protection mechanism, which prevents the
Figure 6.1: Structural Test Model.
actuator from experiencing excessive compressive loads. One of the two load sensors is located next to the actuator, which provides excellent collocation out to 5000 Hz. However, due to the overload protection device, this internal sensor does not measure the complete load path. The second load sensor is located outside of the overload protection system in order to measure the complete load path. This sensor is collocated with the actuator out to about 700 Hz, after which the local strut dynamics become significant.

The active damping control system is mounted to the lower part of the K tubes. This system uses a set of self-sensing piezoelectric actuators to damp the bending modes in the K tubes. The active damping control system loop was designed independent of the AVIS loop due to minimal dynamic interaction between the two loops.

For open and closed-loop testing, an input disturbance is applied to the granite block along the X and Z axes. It is applied as a stochastic force with the PSD shown in Figure 6.2.

Control system performance is evaluated from a set of 29 accelerometers mounted to the structure: 21 sensors on the primary mirror, 3 sensors on the secondary mirror, and 5 sensors at the detector location. Data from these sensors is used to determine the amount of wavefront error at the primary mirror and image motion on the detector which would be seen if the optics were present. These performance metrics are evaluated over a 10–200 Hz bandwidth.

6.2 Design Objectives

The ultimate objective of AVIS is to move the open-loop telescope-on-struts modes from the 30–65 Hz range down to the 5–10 Hz range in order to provide > 30 dB of isolation in the 10–200 Hz performance band. This can be accomplished by minimizing the closed-loop gain from disturbance input to load sensor output in this frequency range. However, this is also subject to the design constraint that these modes stay above 5 Hz, which limits the amount that the closed-loop gain can be minimized.
A controller designed by classical methods has been implemented on the AVIS hardware. It succeeds in reducing the closed-loop gain in the 10–200 Hz range, while keeping the lowest modal frequency above 5 Hz. Figure 6.3 shows the closed-loop of a linear model of strut 1 with the classical controller. Also on this Bode plot is the open-loop linear model of strut 1. The classical controller has approximately 2 dB of gain margin and 10° of phase margin with the FRF data. Although these types of margins would be too close to be used in flight hardware, they were sufficient for this study where the goal was to maximize closed-loop performance. A complete discussion of the classical controller designed for AVIS can be found in [15].

The design objectives of the constrained optimization controller are to:

1. Improve upon the closed-loop classical performance in the 10–200 Hz range.
2. Meet or exceed the gain and phase margins of the classical controller.
3. Keep the primary structural mode above 5 Hz in the closed loop.
4. Maintain closed-loop stability with all six strut loops closed.

### 6.3 Phase One Design

The first set of constrained optimization controllers for AVIS were based on detailed linear models designed by Kodak for each strut from the actuator to the external load sensor. These models were in discrete time with a time step of 0.0001 sec. The models were derived from a blending of system identifications on low, mid, and high frequency regions. The linear model for strut 1 compared to the FRF data is shown in Figure 6.4. This model distinctly shows the collocated (below 200 Hz) and non-collocated (above 700 Hz) regions. The designs presented in this section and Section 6.4 are only for the first strut. Similar designs were carried out for the other five struts.
6.3.1 Optimization

The controller design was posed as a standard disturbance rejection problem, where the disturbance impact on the external load cell and actuator output was minimized. The control and sensor outputs were appended with frequency weights in the objective function. Minimization of the closed loop disturbance to sensor transfer function was only important for frequencies below 200 Hz, so the external load cell output was appended with a frequency weight emphasizing these frequencies. In continuous time, the sensor weight was

\[ W_1(s) = \left( \frac{s + 1000 \times 2\pi}{s + 200 \times 2\pi} \right)^4. \]

Control action at high frequency was undesirable, due to unmodeled dynamics and sensor noise. Therefore the control output was appended with a second order high-pass filter. The control weight used was

\[ W_2(s) = \left( \frac{15s}{s + 2000 \times 2\pi} \right)^2. \]

Figure 6.5 illustrates the closed loop as a disturbance rejection problem. The objec-
Figure 6.5: AVIS disturbance rejection problem.

tive function was the $\mathcal{H}_2$ norm of the weighted sensor plus control outputs. Notice that this is a two-block problem (there are more regulated outputs than control inputs). Therefore, the delay augmentation problem was required in the solution of the optimization problem.

6.3.2 Choice of Basis

Due to the presence of high frequency lightly damped poles, the FIR basis set proved to be inefficient for this problem. This is made clear by examining the behavior of the closed loop using the classical controller. Because the controller rolls off at high frequencies, the high frequency dynamics in the open loop stay roughly the same in the closed loop. Figure 6.3 shows the frequency response of the open and closed loop, which shows very little difference at frequencies above 1000 Hz. Figure 6.6 shows the first 500 samples in the impulse response of the classical closed loop. To reasonably approximate this with an FIR, at least a thousand terms would be required. However, most of the modes that cause this impulse response to be so long are already approximately known. Therefore, a much more efficient basis can be selected for the closed loop by combining the high frequency modes with a much shorter FIR basis.

A blended FIR/FPM basis set was used for the closed loop optimization. For each
mode above 900 Hz, a pair of FPM basis functions was created. All of these basis functions were then combined with a 60th order FIR basis, creating the following transfer function:

\[ H(z) = \sum_{n=0}^{N-1} \frac{\phi_{2n}}{A_n(z)} + \frac{z\phi_{2n+1}}{A_n(z)} + \sum_{n=0}^{59} \frac{\phi_{2N+n}}{z^n} \]

where \( A_n(z) = (z-p_n)(z-p_n^*) \). This set of basis functions was then orthonormalized using the technique given in Section 5.3, and the basis functions with small singular values were removed.

### 6.3.3 Performance Specifications

To restate the performance objectives of AVIS, it is desired to attenuate the sensor output due to the disturbance in the 10-200 Hz frequency band. The closed loop gain of the classical design is about \(-24\) dB between 10-60 Hz, but quickly loses performance above 60 Hz. It was decided that the constrained optimization design should be more aggressive in performance above 60 Hz at the expense of performance between 10-60 Hz. Closed-loop frequency constraints were chosen at \(-20\) dB between
Figure 6.7: Constrained optimization and classical performance for phase 1 design.

10–150 Hz. Placing aggressive constraints above 150 Hz caused the $\mathcal{H}_2$ design to become infeasible, so loss in performance between 150–200 Hz was tolerated. Figure 6.7 shows the closed-loop designs of the classical and constrained optimization controllers using the FRF data.

6.3.4 Stability Robustness

There was known to exist a high degree of model uncertainty at high frequencies, but it is difficult to quantify how much. Controllers which stabilized the closed-loop system with the FRF data frequently caused the real closed-loop system to go unstable. Therefore, differences between the FRF data and the linear model did not account for the uncertainty. Often, the real system would remain stable if the loop around an individual strut was closed, but went unstable if more than one or two loops were closed. This suggested that some of the uncertainty was due to dynamic coupling between the struts. Nonlinearities also played a role. The FRF data, and indeed the
models, were derived from open-loop test data taken in response to actuator inputs. These inputs were set in order to maximize the signal-to-noise ratio of the sensors. These high input signals are not representative of the lower amplitude signals typically seen in closed-loop operation, and the nonlinearities of the system accounted for as much as a 5 dB difference in gain in the region of crossover (between 700 and 1200 Hz).

Previous AVIS tests provided examples of controllers that were successfully able to stabilize the closed loop, and those that were not. It was noticed that for some controllers, once multiple loops were closed, a mode in the 1000 Hz region would induce instability. Therefore, it was realized that in the 1000 Hz region the robustness constraints on the complementary sensitivity $PK/(1 + PK)$ would have to be below the magnitude of the complementary sensitivity for controllers which produced unstable designs. It was also desired to force the controller to roll off at high frequencies, so a first order roll off constraint was placed on the complementary sensitivity. Figure 6.8 illustrates the constraints chosen for strut 1, along with two previous loop-shaping design which were and were not able to stabilize the real system. Figure 6.9 shows these constraints applied to the actual constrained optimization design. The apparent constraint violation at 3000 Hz actually does not violate any constraints, but has crept between two separate frequency constraints.

6.3.5 Controller Reduction

The constrained optimization method relies on solving for the desired closed loop rather than solving for the controller directly. Deriving the controller that produces a particular closed loop can lead to a fairly large ordered controller. For example, the strut 1 model began with 86 states, and the closed loop in the optimization was limited to 75. Because the model was stable, it was possible to find a $Q$-parameterization where $T_1$, $T_2$, and $T_3$ each were limited to 86 states. $Q$ can then be found by

$$Q = T_2^{-1}(H - T_1)T_3^{-1}$$
Figure 6.8: Two previous AVIS controller design, and resulting frequency constraints.

Figure 6.9: Constrained complementary sensitivity of the phase 1 design.
as long as $T_2$ and $T_3$ are invertible. This leads to a $Q$ with 333 states. Most of these disappear by removing all the near pole-zero cancellations, and with standard model reduction methods. However, another 86 states will be added when deriving the controller from $K = \mathcal{F}_c(K_s, Q)$. In sum, it is very likely that the controller dimension will become very large. Model reduction is essential to producing practical controllers.

Initial controller reduction led to controllers with 40 states and still preserved the basic shape of the original controllers. These controllers showed good stability characteristics with the FRF data, and it was believed that they would also stabilize the real system. However, control action at high frequency caused the system to go unstable when more than one loop was closed, so it was necessary to further reduce the controllers. Also, at low frequencies the controller gain became very large. This caused the open loop modes between 10–200 Hz to shift below 5 Hz, violating the original design specifications. This was easily corrected by removing the low frequency dynamics of the controller, forcing the controller gain to remain constant at low frequencies. Figure 6.10 shows the reduction process for the strut 1 controller. Out of all three controllers shown in this figure, only the fourth order controller was able to stabilize the real system.

### 6.3.6 Experimental Results

The closed-loop wavefront error PSD for the phase 1 design is shown in Figure 6.11. Figure 6.12 shows the image motion performance metric for this design, along with the frequency-dependent performance goal. The performance of the classical controller is shown for comparison. It is clear from these plots that by not being as aggressive as the classical controller at low frequencies, the constrained optimization controller has sacrificed a fair amount of performance. The performance goal for the wavefront error is 0.02 waves RMS, and although neither design meets the specification, the classical design is closer than the constrained optimization design. The image motion with the constrained optimization design was substantially higher than the motion for the classical design at low frequencies, and only showed modest improvement at
Figure 6.10: Controller reduction for phase 1 design.
Figure 6.11: Wavefront error PSD performance for phase 1 design.

Figure 6.12: Image motion performance for phase 1 design.
high frequencies.

Testing the phase 1 controllers gave insight that became important in designing the second set of constrained optimization controllers for AVIS. It was believed that some of the performance should be given up at low frequencies to allow a much greater improvement in performance at higher frequencies. The experimental results showed that the loss in performance at low frequencies outweighed the gain at high frequencies. From this, it was decided that the next set of performance constraints should be set to give at least as much performance as the classical controller at all frequencies between 10–200 Hz.

6.4 Phase Two Design

The lessons learned from testing the phase 1 controller designs led to a much simplified linear model of the plant. New models were derived which did not include the fine detail between 10–200 Hz. It was found that knowing these modes did not offer anything beneficial to the controller design process, and only made the plant models more cumbersome. It was also discovered that the modes in the model above 800 Hz to a large extent were not valid in the real system, due to coupling and nonlinearities. Model-based control design often attempts to invert the plant, which offers some explanation why the 40th order controller in Figure 6.10 shows so much activity at high frequency. To prevent the controller from inverting unknown plant dynamics, these dynamics were removed from the linear model. This reduction of the plant model can be seen in Figure 6.13. Finally, a one-step time delay was added to the plant to account for the processor time delay and zero order hold present in the actual implementation. The effect of this time delay on the design model is visible as a reduction in phase in Figure 6.13. In the end, the strut 1 design model ended up having 16 states, compared to 87 for the phase 1 design. Similar reduction was seen in the other five strut models.
6.4.1 Performance Specifications

For the second phase of controller designs, it was decided that constrained optimization performance should meet or exceed the closed-loop performance of the classical design at all frequencies between 10–200 Hz. This was slightly different than the phase 1 design goal, where the constrained optimization performance was set at −20 dB regardless of classical performance. The performance design constraints were set at −26 dB between 10–100 Hz. After 100 Hz, the constraints were linearly increased to −3 dB at 200 Hz. These constraints were found to be nearly the most aggressive performance constraints that could be placed on the closed loop and still be achievable. Figure 6.14 shows the constrained optimization closed loop with the FRF data. For comparison, the classical closed loop is also shown.

6.4.2 Stability Robustness

Experience showed that the controllers with excessive high frequency dynamics tended to destabilize the closed loop, whereas those with a smooth roll off were stabilizing
Figure 6.14: Constrained optimization and classical performance.

(e.g., see Figure 6.10). Therefore, the frequency constraints on $PK/(1 + PK)$ were chosen to encourage a smooth, first order roll off in the controller. When the loop gain magnitude is much less than one (i.e., $|PK| \ll 1$) then it approaches the gain of $PK/(1 + PK)$. This should be true at high frequencies, where the closed loop must be gain stabilized. A roll off of $\frac{200}{\omega}$ was chosen for the controller. This was approximated as a constraint on the complementary sensitivity:

$$|PK/(1 + PK)| \leq \frac{200}{\omega}|P|.$$  

Figure 6.15 illustrates the constrained complementary sensitivity.

The constrained optimization design goal was not only to improve upon performance over the classical design, but to provide better stability margins. The initial constrained optimization designs showed a poor gain margin, so sensitivity was constrained to improve this margin. Figure 6.16 shows the constraints on the sensitivity, and Figure 6.17 shows the resulting Nichols plot. The apparent constraint violation at 450 Hz is because this frequency constraint has only been approximated by a set
of 16 linear constraints (See Section 4.5). For this design, the phase margin frequency was about 450 Hz, and the gain margin frequency was about 650 Hz. Due to the frequency constraints, the sensitivity shows a reduction in gain from 13.7 dB to 9.4 dB at the gain margin frequency. Referring back to the gain margin equation in Section 4.3 (Equation 4.5), the predicted improvement in gain margin is from −2 dB to −3.6 dB, which is confirmed by the Nichols plot in Figure 6.17. The Nichols plot of the constrained optimization and classical designs in series with the FRF data is shown in Figure 6.18. Notice that the constrained optimization design has an upper gain margin, whereas the classical design does not. This is a direct consequence of the increased performance achieved by the constrained optimization design.

6.4.3 Controller Reduction

Controller implementation was much easier for the phase 2 designs than the phase 1 designs. At high frequencies, model reduction techniques were used to eliminate unnecessary modes without affecting the overall shape of the controller. As before,
Figure 6.16: Sensitivity of phase 2 design.

Figure 6.17: Nichols plot with linear model of design with stability margin constraints.
the controller gain at low frequencies was too high, and shifted the closed loop modes below 5 Hz. Removing the low frequency dynamics from the controller provided a design which met the 5 Hz specification. Figures 6.19 and 6.20 show the effect of this reduction on the controller and closed loop. The low frequency model reduction did mean that the closed loop no longer met the frequency constraints. Even with this slight loss in performance, the reduced constrained optimization controller still performed better than the classical classical controller in the closed loop between 10–200 Hz.

6.4.4 Experimental Results

The closed-loop wavefront error PSD for the phase 2 design is shown in Figure 6.21. The image motion, along with the frequency-dependent performance goal, is shown in Figure 6.22. The most obvious conclusion to draw from the experimental data is that the constrained optimization design has offered almost no improvement over
Figure 6.19: Controller reduction for phase 2 design.

Figure 6.20: Effect of controller reduction on closed loop.
Figure 6.21: Wavefront error PSD performance for phase 2 design.

Figure 6.22: Image motion performance for phase 2 design.
the classical design in meeting the performance specifications. The wavefront error RMS has improved from 0.070 to 0.069, which hardly moves AVIS towards the design goal of 0.02 RMS. The situation is similar for the image motion performance metric. The classical design does not meet the performance goal at 12 Hz, 31 Hz, 38 Hz, and 60 Hz. This goal is met at 60 Hz with the constrained optimization design, but still fails for the other three modes.

Even though these results show that the classical design has performed just as well as the constrained optimization design, this should not be misinterpreted as saying that constrained optimization offers nothing over classical methods of controller design. For this particular application, the regulated outputs of the design model and the performance measurements for the real system are fairly different. For example, the modes that appear in the closed-loop image motion at 12 Hz, 31 Hz, 38 Hz, and 60 Hz are virtually uncontrollable and unobservable to the controller, and do not appear in the regulated output of the model. These modes represent the bending modes of the struts, so failure to meet specification is more a limitation of the hardware than a poor controller design.

In judging how effective the constrained optimization method has been compared to classical methods of designing controllers for this application, it is more fair to look at the design model performance. In this case, the constrained optimization method has given a design which performs better than the classical controller (see Figure 6.14) and has a better stability margin (see Figure 6.17). The constrained optimization design has also given insight into the limits of performance of this particular application. By methodically choosing more aggressive performance constraints for each design iteration, eventually the design constraints were found to be infeasible. This information is not available with classical control methods. The constrained optimization method has provided a controller design which pushes the performance near the absolute limit. Therefore, because this controller fails to meet the performance specification, it seems likely that there does not exist a linear controller which does meet these specifications with the existing actuator/sensor configuration.
Chapter 7

Conclusions

A control design methodology for constrained optimization has been presented in this thesis. The method was applied to an active vibration isolation system, demonstrating that it is of practical use for real-world control problems. Although the experimental results showed little difference in performance between the constrained optimization controller and a controller designed with classical methods, the strength of the constrained optimization method should not be underestimated. It is frequently not obvious how to choose a classical controller for a particular system. The design process must rely on the skill and intuition of the engineer, not to mention a certain degree of luck. Constrained optimization offers a much more direct way of designing controllers. It allows the specifications to be chosen directly, and avoids much of the trial-and-error iterations that confront a designer using classical or modern methods. Also, it is an almost general method which can handle a wide variety of MIMO LTI systems.

Ideally, there would exist a method that would take an arbitrary system, set of design specifications, and performance objective function, and would output the optimal controller, if one exists. Unfortunately, even restricted to LTI systems, this is an infinite-dimensional optimization problem which is frequently non-convex. In this thesis, several simplifying assumptions were made that make this problem tractable. First, the optimization was restricted to a finite-dimensional subspace which was felt would closely approximate the optimal answer (e.g., the solution presented in
Chapter 3 restricted the closed loop to be an FIR filter). The second simplification was that only linear design specifications were considered. This allowed the problem to be solved efficiently as a linear or quadratic program, depending on the objective function. Although it was seen in Section 4.5 that some nonlinear convex constraints could be posed as LMIs, which also have an efficient solution, the nonlinear constraints required for this thesis could be adequately approximated with a finite set of linear constraints.

With these simplifications, the constrained optimization method is tractable and can still handle a large variety of control problems. The biggest limitation on this method is the available computing power. Depending on the number of variables and constraints in the optimization problem, the solution may take too long to find, or the problem may exceed the memory limitations of the computer. Obviously, small academic examples which work well with this method can be found, and some may even be solved by hand. An important goal of this thesis was to make this method work for a real engineering problem, where the computer limitations became very relevant. The application chosen was a large flexible structure, with an abundance of closely-spaced, lightly-damped modes. This type of problem presents a difficulty when optimizing over a small subspace, because the impulse response of the open-loop system has a slow decay rate. If the closed loop is approximated as an FIR filter (a common approximation in the literature), then the relevant subspace may require thousands of terms in the FIR. For this reason, Chapter 5 examined the use of alternative basis functions which may approximate the optimal closed loop more efficiently than the FIR filter. A basis which combined some of the open-loop pole locations with an FIR filter was successfully used for the structural control problem. Also, the magnetic bearing example developed in Chapter 4 used a basis of Laguerre functions. In both cases, the answers derived would have been difficult, if not impossible, to find with an FIR basis, due to memory limitations of the computer.

On the whole, the constrained optimization method proved to be an effective controller design technique for AVIS. It provided a set of controllers with a sufficient amount of robustness which were pushed close to the limits of performance. The
design was driven mostly by the constraints, which were chosen in an obvious way. Subject to the simplifying assumptions required by the method, it answered the fundamental problem posed in Chapter 1. That is, this method was able to find a controller which met the design specifications when one existed.

7.1 Recommendations for Future Work

This thesis has dealt with only one application of constrained optimization to a real system. There are presumably many other practical problems that would be well suited for constrained optimization. Practical design problems play an important role in transforming a method that works in principle into a useful design tool. These problems also help to uncover limitations in the method. More experience using the constrained optimization design method is needed to further develop and understand this method.

There are certainly many ways in which constrained optimization may be used in control, and only a few could be touched on in this thesis. One unexplored possibility is the use of “soft” constraints. The designs in this thesis used only hard constraints. Softer constraints can be constructed by placing the magnitude of the constraint in the objective function.

The availability of time and frequency data was not fully exploited, and leaves another potential use of constrained optimization open for development. The AVIS controller designs in Chapter 6 depended on a linear model of the plant, even though the FRF data was available. It is conceivable that the methods presented in this thesis could be used to design a controller with only time and frequency data, rather than a model of the plant.

Choosing an efficient set of basis functions is an essential step in reducing the dimensionality of a design problem. Although this thesis did demonstrate that there are other sets of basis functions which can be more efficient than an FIR basis, there was little discussion about how a specific basis should be chosen. Selecting this basis requires some knowledge about the optimal closed loop, which may appear to be
self defeating. However, the bases used for AVIS and the magnetic bearing example were selected with nothing more than a little intuition. It would be useful to develop a general method, or even some rules of thumb, for selecting a good set of basis functions for the closed loop given an open loop system.

Several unexpected problems developed for both the AVIS design and the magnetic bearing example. The most frustrating was a tendency of the frequency response to sneak between the discrete set of frequency constraints, especially at high frequencies. An example of this can be seen by referring back to Figure 6.15. This was dealt with by increasing the density of the frequency constraints. The drawback to this approach is that it increases the total number of constraints in the optimization problem, therefore requiring much more computational power. If there exists a method of constraining the slope of the frequency response at a particular frequency point, as well as the magnitude, then it may be possible to use many less frequency constraints. This problem may also be a byproduct of the optimization. It is possible that the optimization is using the extra degrees of freedom between the constraints to further reduce the objective function. It would be interesting to examine the solution at various stages of the optimization, and see if there is a point at which the closed loop behavior starts to become undesirable.

Another problem encountered was finding an effective way of forcing the controller to roll off at high frequencies. Using modern control methods, this is often accomplished by increasing frequency weights on the control effort and sensor noise at high frequency. For some reason, frequency weights proved to be ineffective with constrained optimization. If this problem consistently appears in future designs using this method, then there should be some research into the cause behind it, and how it might be corrected.

Optimization over $Q$ rather than the closed loop was briefly discussed in Section 3.4, but was not used in the designs. There would be numerous advantages to this approach. There would be no interpolation conditions. The dimension of $Q$ is always less than or equal to the dimension of the closed loop. Also, because the $Q$-parameterization is based on a nominal closed loop with a stabilizing controller, it
would be possible to devise methods that improve upon existing designs. This may
lead to an iterative approach that converges on the optimal controller. One potential
drawback is that it may be difficult to choose a good set of basis functions for $Q$. The
direct use of $Q$ for optimization deserves more investigation.

Finally, it would be worthwhile to compare the use of LMI constraints with linear
constraints in control design. Many convex constraints that can only be approxi-
mated with a set of linear constraints can be posed exactly as an LMI. Depending
on the constraint, this advantage may or may not be important. If LMI problems
offer every advantage over problems which only use linear constraints, then the op-
timization techniques in this thesis should be abandoned in favor of LMI methods.
However, it is more probable that LMI optimization methods also have inherent dis-
advantages, For example, LMI optimization may be slower and use more memory
than the simplex method. There should be a more thorough understanding about
when LMI optimization is an appropriate tool for control design.
Bibliography


