ROBUST AND OPTIMAL CONTROL

ROBUST AND OPTIMAL CONTROL

KEMIN ZHOU

with

JOHN C. DOYLE and KEITH GLOVER

PRENTICE HALL, Englewood Cliffs, New Jersey 07632

TO OUR PARENTS

獻給親愛的父母

Contents

Preface

This book is inspired by the recent development in the robust and $\mathcal{H}_\infty$ control theory, particularly the state-space $\mathcal{H}_\infty$ control theory developed in the paper by Doyle, Glover, Khargonekar, and Francis [1989] (known as DGKF paper). We give a fairly comprehensive and step-by-step treatment of the state-space $\mathcal{H}_\infty$ control theory in the style of DGKF. We also treat the robust control problems with unstructured and structured uncertainties. The linear fractional transformation (LFT) and the structured singular value (known as μ) are introduced as the unified tools for robust stability and performance analysis and synthesis. Chapter 1 contains a more detailed chapter-by-chapter review of the topics and results presented in this book.

We would like to thank Professor Bruce A. Francis at University of Toronto for his helpful comments and suggestions on early versions of the manuscript. As a matter of fact, this manuscript was inspired by his lectures given at Caltech in 1987 and his masterpiece – *A Course in $\mathcal{H}_\infty$ Control Theory.* We are grateful to Professor André Tits at University of Maryland who has made numerous helpful comments and suggestions that have greatly improved the quality of the manuscript. Professor Jakob Stoustrup, Professor Hans Henrik Niemann, and their students at The Technical University of Denmark have read various versions of this manuscript and have made many helpful comments and suggestions. We are grateful to their help. Special thanks go to Professor Andrew Packard at University of California-Berkeley for his help during the preparation of the early versions of this manuscript. We are also grateful to Professor Jie Chen at University of California-Riverside for providing material used in Chapter 6. We would also like to thank Professor Kang-Zhi Liu at Chiba University (Japan) and Professor Tongwen Chen at University of Calgary for their valuable comments and suggestions. In addition, we would like to thank G. Balas, C. Beck, D. S. Bernstein, G. Gu, W. Lu, J. Morris, M. Newlin, L. Qiu, H. P. Rotstein and many other people for their comments and suggestions. The first author is especially grateful to Professor Pramod P. Khargonekar at The University of Michigan for introducing him to robust and $\mathcal{H}_\infty$ control and to Professor Tryphon Georgiou at University of Minnesota for encouraging him to complete this work.

Kemin Zhou

John C. Doyle
Keith Glover

Notation and Symbols

$\mathbb{R}$	field of real numbers
$\mathbb{C}$	field of complex numbers
$\mathbb{F}$	field, either $\mathbb{R}$ or $\mathbb{C}$
$\mathbb{R}_+$	nonnegative real numbers
$\mathbb{C}_-$ and $\overline{\mathbb{C}}_-$	open and closed left-half plane
$\mathbb{C}_+$ and $\overline{\mathbb{C}}_+$	open and closed right-half plane
$\mathbb{C}_0$, $j\mathbb{R}$	imaginary axis
$\mathbb{D}$	unit disk
$\overline{\mathbb{D}}$	closed unit disk
$\partial\mathbb{D}$	unit circle
$\in$	belong to
$\subset$	subset
$\cup$	union
$\cap$	intersection
$\square$	end of proof
$\diamondsuit$	end of example
$\heartsuit$	end of remark
$:=$	defined as
$\gtrsim$	asymptotically greater than
$\lesssim$	asymptotically less than
$\gg$	much greater than
$\ll$	much less than
$\overline{\alpha}$	complex conjugate of $\alpha \in \mathbb{C}$
$\lvert\alpha\rvert$	absolute value of $\alpha \in \mathbb{C}$
$Re(\alpha)$	real part of $\alpha \in \mathbb{C}$
$\delta(t)$	unit impulse
δ_{ij}	Kronecker delta function, $\delta_{ii} = 1$ and $\delta_{ij} = 0$ if $i \neq j$
$1_+(t)$	unit step function

I_n	$n \times n$ identity matrix
$[a_{ij}]$	a matrix with a_{ij} as its i-th row and j-th column element
$\mathrm{diag}(a_1, \ldots, a_n)$	an $n \times n$ diagonal matrix with a_i as its i-th diagonal element
A^T	transpose
A^*	adjoint operator of A or complex conjugate transpose of A
A^{-1}	inverse of A
A^+	pseudo inverse of A
A^{-*}	shorthand for $(A^{-1})^*$
$\det(A)$	determinant of A
$\mathrm{Trace}(A)$	trace of A
$\lambda(A)$	eigenvalue of A
$\rho(A)$	spectral radius of A
$\sigma(A)$	the set of spectrum of A
$\overline{\sigma}(A)$	largest singular value of A
$\underline{\sigma}(A)$	smallest singular value of A
$\sigma_i(A)$	i-th singular value of A
$\kappa(A)$	condition number of A
$\|A\|$	spectral norm of A: $\|A\| = \overline{\sigma}(A)$
Im(A), R(A)	image (or range) space of A
Ker(A), N(A)	kernel (or null) space of A
$\mathcal{X}_-(A)$	stable invariant subspace of A
$\mathcal{X}_+(A)$	antistable invariant subspace of A
$Ric(H)$	the stabilizing solution of an ARE
$g * f$	convolution of g and f
$\otimes$	Kronecker product
$\oplus$	direct sum or Kronecker sum
$\angle$	angle
$\langle , \rangle$	inner product
$x \perp y$	orthogonal, $\langle x, y \rangle = 0$
$D_\perp$	orthogonal complement of D, i.e., $\begin{bmatrix} D & D_\perp \end{bmatrix}$ or $\begin{bmatrix} D \\ D_\perp \end{bmatrix}$ is unitary
$S^\perp$	orthogonal complement of subspace S, e.g., $\mathcal{H}_2^\perp$
$\mathcal{L}_2(-\infty, \infty)$	time domain Lebesgue space
$\mathcal{L}_2[0, \infty)$	subspace of $\mathcal{L}_2(-\infty, \infty)$
$\mathcal{L}_2(-\infty, 0]$	subspace of $\mathcal{L}_2(-\infty, \infty)$
$\mathcal{L}_{2+}$	shorthand for $\mathcal{L}_2[0, \infty)$
$\mathcal{L}_{2-}$	shorthand for $\mathcal{L}_2(-\infty, 0]$
l_{2+}	shorthand for $l_2[0, \infty)$
l_{2-}	shorthand for $l_2(-\infty, 0)$
$\mathcal{L}_2(j\mathbb{R})$	square integrable functions on $\mathbb{C}_0$ including at ∞
$\mathcal{L}_2(\partial\mathbb{D})$	square integrable functions on $\partial\mathbb{D}$
$\mathcal{H}_2(j\mathbb{R})$	subspace of $\mathcal{L}_2(j\mathbb{R})$ with analytic extension to the rhp

$\mathcal{H}_2(\partial\mathbb{D})$	subspace of $\mathcal{L}_2(\partial\mathbb{D})$ with analytic extension to the inside of $\partial\mathbb{D}$
$\mathcal{H}_2^{\perp}(j\mathbb{R})$	subspace of $\mathcal{L}_2(j\mathbb{R})$ with analytic extension to the lhp
$\mathcal{H}_2^{\perp}(\partial\mathbb{D})$	subspace of $\mathcal{L}_2(\partial\mathbb{D})$ with analytic extension to the outside of $\partial\mathbb{D}$
$\mathcal{L}_\infty(j\mathbb{R})$	functions bounded on $\mathrm{Re}(s) = 0$ including at ∞
$\mathcal{L}_\infty(\partial\mathbb{D})$	functions bounded on $\partial\mathbb{D}$
$\mathcal{H}_\infty(j\mathbb{R})$	the set of $\mathcal{L}_\infty(j\mathbb{R})$ functions analytic in $\mathrm{Re}(s) > 0$
$\mathcal{H}_\infty(\partial\mathbb{D})$	the set of $\mathcal{L}_\infty(\partial\mathbb{D})$ functions analytic in $\lvert z\rvert < 1$
$\mathcal{H}_\infty^-(j\mathbb{R})$	the set of $\mathcal{L}_\infty(j\mathbb{R})$ functions analytic in $\mathrm{Re}(s) < 0$
$\mathcal{H}_\infty^-(\partial\mathbb{D})$	the set of $\mathcal{L}_\infty(\partial\mathbb{D})$ functions analytic in $\lvert z\rvert > 1$
prefix $\mathcal{B}$ or $\mathbf{B}$	*closed* unit ball, e.g. $\mathcal{BH}_\infty$ and $\mathbf{B\Delta}$
prefix $\mathbf{B}^o$	*open* unit ball
prefix $\mathcal{R}$	real rational, e.g., $\mathcal{RH}_\infty$ and $\mathcal{RH}_2$, etc
$\mathbb{R}[s]$	polynomial ring
$\mathcal{R}_p(s)$	rational proper transfer matrices
$G^\sim(s)$	shorthand for $G^T(-s)$ (continuous time)
$G^\sim(z)$	shorthand for $G^T(z^{-1})$ (discrete time)
$\left[\begin{array}{c\|c} A & B \\ \hline C & D \end{array}\right]$	shorthand for state space realization $C(sI - A)^{-1}B + D$ or $C(zI - A)^{-1}B + D$
$\mathcal{F}_\ell(M, Q)$	lower LFT
$\mathcal{F}_u(M, Q)$	upper LFT
$\mathcal{S}(M, N)$	star product

List of Acronyms

ARE	algebraic Riccati equation
BR	bounded real
CIF	complementary inner factor
DF	disturbance feedforward
FC	full control
FDLTI	finite dimensional linear time invariant
FI	full information
HF	high frequency
iff	if and only if
lcf	left coprime factorization
LF	low frequency
LFT	linear fractional transformation
lhp or LHP	left-half plane $\mathrm{Re}(s) < 0$
LQG	linear quadratic Gaussian
LQR	linear quadratic regulator
LTI	linear time invariant
LTR	loop transfer recovery
MIMO	multi-input multi-output
nlcf	normalized left coprime factorization
NP	nominal performance
nrcf	normalized right coprime factorization
NS	nominal stability
OE	output estimation
OF	output feedback
OI	output injection
rcf	right coprime factorization
rhp or RHP	right-half plane $\mathrm{Re}(s) > 0$
RP	robust performance
RS	robust stability
SF	state feedback
SISO	single-input single-output
SSV	structured singular value (μ)
SVD	singular value decomposition

1

Introduction

1.1 Historical Perspective

This book gives a comprehensive treatment of optimal $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control theory and an introduction to the more general subject of robust control. Since the central subject of this book is *state-space* $\mathcal{H}_\infty$ optimal control, in contrast to the approach adopted in the famous book by Francis [1987]: *A Course in $\mathcal{H}_\infty$ Control Theory*, it may be helpful to provide some historical perspective of the state-space $\mathcal{H}_\infty$ control theory to be presented in this book. This section is not intended as a review of the literature in $\mathcal{H}_\infty$ theory or robust control, but rather only an attempt to outline some of the work that most closely touches on our approach to state-space $\mathcal{H}_\infty$. Hopefully our lack of written historical material will be somewhat made up for by the pictorial history of control shown in Figure 1.1. Here we see how the practical but classical methods yielded to the more sophisticated modern theory. Robust control sought to blend the best of both worlds. The strange creature that resulted is the main topic of this book.

The $\mathcal{H}_\infty$ optimal control theory was originally formulated by Zames [1981] in an input-output setting. Most solution techniques available at that time involved analytic functions (Nevanlinna-Pick interpolation) or operator-theoretic methods [Sarason, 1967; Adamjan *et al.*, 1978; Ball and Helton, 1983]. Indeed, $\mathcal{H}_\infty$ theory seemed to many to signal the beginning of the end for the state-space methods which had dominated control for the previous 20 years. Unfortunately, the standard frequency-domain approaches to $\mathcal{H}_\infty$ started running into significant obstacles in dealing with multi-input multi-output (MIMO) systems, both mathematically and computationally, much as the $\mathcal{H}_2$ (or LQG) theory of the 1950's had.

Figure 1.1: A picture history of control

Not surprisingly, the first solution to a general rational MIMO $\mathcal{H}_\infty$ optimal control problem, presented in Doyle [1984], relied heavily on state-space methods, although more as a computational tool than in any essential way. The steps in this solution were as follows: parameterize all internally-stabilizing controllers via [Youla *et al.*, 1976]; obtain realizations of the closed-loop transfer matrix; convert the resulting model-matching problem into the equivalent 2×2-block general distance or best approximation problem involving mixed Hankel-Toeplitz operators; reduce to the Nehari problem (Hankel only); solve the Nehari problem by the procedure of Glover [1984]. Both [Francis, 1987] and [Francis and Doyle, 1987] give expositions of this approach, which will be referred to as the "1984" approach.

In a mathematical sense, the 1984 procedure "solved" the general rational $\mathcal{H}_\infty$ optimal control problem and much of the subsequent work in $\mathcal{H}_\infty$ control theory focused on the 2×2-block problems, either in the model-matching or general distance forms. Unfortunately, the associated complexity of computation was substantial, involving several Riccati equations of increasing dimension, and formulae for the resulting controllers tended to be very complicated and have high state dimension. Encouragement came from Limebeer and Hung [1987] and Limebeer and Halikias [1988] who showed, for problems transformable to 2×1-block problems, that a subsequent minimal realization of the controller has state dimension no greater than that of the generalized plant G. This suggested the likely existence of similarly low dimension optimal controllers in the general 2×2 case.

Additional progress on the 2×2-block problems came from Ball and Cohen [1987], who gave a state-space solution involving 3 Riccati equations. Jonckheere and Juang [1987] showed a connection between the 2×1-block problem and previous work by Jonckheere and Silverman [1978] on linear-quadratic control. Foias and Tannenbaum [1988] developed an interesting class of operators called skew Toeplitz to study the 2×2-block problem. Other approaches have been derived by Hung [1989] using an interpolation theory approach, Kwakernaak [1986] using a polynomial approach, and Kimura [1988] using a method based on conjugation.

The simple state space $\mathcal{H}_\infty$ controller formulae to be presented in this book were first derived in Glover and Doyle [1988] with the 1984 approach, but using a new 2×2-block solution, together with a cumbersome back substitution. The very simplicity of the new formulae and their similarity with the $\mathcal{H}_2$ ones suggested a more direct approach.

Independent encouragement for a simpler approach to the $\mathcal{H}_\infty$ problem came from papers by Petersen [1987], Khargonekar, Petersen, and Zhou [1990], Zhou and Khargonekar [1988], and Khargonekar, Petersen, and Rotea [1988]. They showed that for the state-feedback $\mathcal{H}_\infty$ problem one can choose a constant gain as a (sub)optimal controller. In addition, a formula for the state-feedback gain matrix was given in terms of an algebraic Riccati equation. Also, these papers established connections between $\mathcal{H}_\infty$-optimal control, quadratic stabilization, and linear-quadratic differential games.

The landmark breakthrough came in the DGKF paper (Doyle, Glover, Khargonekar, and Francis [1989]). In addition to providing controller formulae that are simple and expressed in terms of plant data as in Glover and Doyle [1988], the methods in that paper are a fundamental departure from the 1984 approach. In particular, the Youla parameterization and the resulting 2×2-block model-matching problem of the 1984 solution are avoided entirely; replaced by a more purely state-space approach involving observer-based compensators, a pair of 2×1 block problems, and a separation argument. The operator theory still plays a central role (as does Redheffer's work [Redheffer, 1960] on linear fractional transformations), but its use is more straightforward. The key to this was a return to simple and familiar state-space tools, in the style of Willems [1971], such as completing the square, and the connection between frequency domain inequalities (e.g $\|G\|_\infty < 1$), Riccati equations, and spectral factorizations. This book in some sense can be regarded as an expansion of the DGKF paper.

The state-space theory of $\mathcal{H}_\infty$ can be carried much further, by generalizing time-invariant to time-varying, infinite horizon to finite horizon, and finite dimensional to infinite dimensional. A flourish of activity has begun on these problems since the publication of the DGKF paper and numerous results have been published in the literature, not surprising, many results in DGKF paper generalize *mutatis mutandis*, to these cases, which are beyond the scope of this book.

1.2 How to Use This Book

This book is intended to be used either as a graduate textbook or as a reference for control engineers. With the second objective in mind, we have tried to balance the broadness and the depth of the material covered in the book. In particular, some chapters have been written sufficiently self-contained so that one may jump to those special topics without going through all the preceding chapters, for example, Chapter 13 on algebraic Riccati equations. Some other topics may only require some basic linear system theory, for instance, many readers may find that it is not difficult to go directly to Chapters $9 \sim 11$. In some cases, we have tried to collect some most frequently used formulas and results in one place for the convenience of reference although they may not have any direct connection with the main results presented in the book. For example, readers may find that those matrix formulas collected in Chapter 2 on linear algebra convenient in their research. On the other hand, if the book is used as a textbook, it may be advisable to skip those topics like Chapter 2 on the regular lectures and leave them for students to read. It is obvious that only some selected topics in this book can be covered in an one or two semester course. The specific choice of the topics depends on the time allotted for the course and the preference of the instructor. The diagram in Figure 1.2 shows roughly the relations among the chapters and should give the users some idea for the selection of the topics. For example, the diagram shows that the only prerequisite for Chapters 7 and 8 is Section 3.9 of Chapter 3 and, therefore, these two chapters alone may be used as a short course on model reductions. Similarly, one only needs the knowledge of Sections 13.2 and 13.6 of Chapter 13 to understand Chapter 14. Hence one may only cover those related sections of Chapter 13 if time is the factor. The book is separated roughly into the following subgroups:

- Basic Linear System Theory: Chapters $2 \sim 3$.
- Stability and Performance: Chapters $4 \sim 6$.
- Model Reduction: Chapters $7 \sim 8$.
- Robustness: Chapters $9 \sim 11$.
- $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Control: Chapters $12 \sim 19$.
- Lagrange Method: Chapter 20.
- Discrete Time Systems: Chapter 21.

In view of the above classification, one possible choice for an one-semester course on robust control would cover Chapters $4 \sim 5, 9 \sim 11$ or $4 \sim 11$ and an one-semester advanced course on $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control would cover (parts of) Chapters $12 \sim 19$. Another possible choice for an one-semester course on $\mathcal{H}_\infty$ control may include Chapter 4, parts of Chapter 5 ($5.1 \sim 5.3, 5.5, 5.7$), Chapter 10, Chapter 12 (except Section 12.6), parts of Chapter 13 ($13.2, 13.4, 13.6$), Chapter 15 and Chapter 16. Although Chapters $7 \sim 8$ are very much independent of other topics and can, in principle, be studied at any

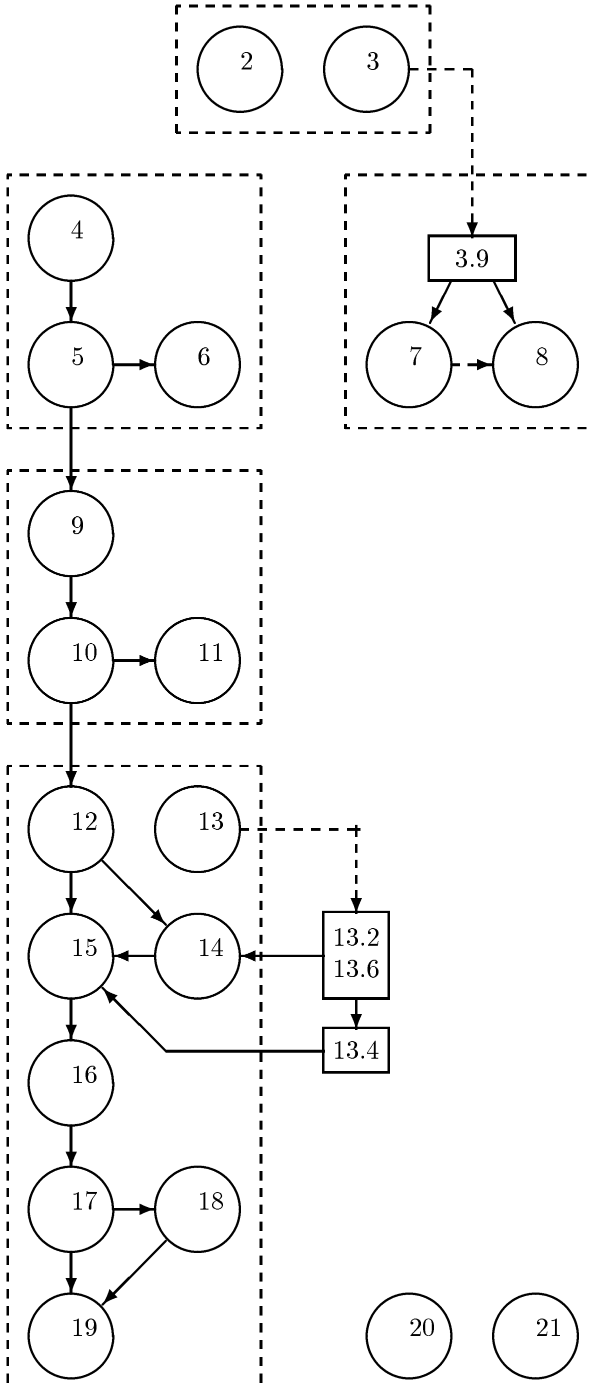


Figure 1.2: Relations among the chapters

stage with the background of Section 3.9, they may serve as an introduction to sources of model uncertainties and hence to robustness problems.

Robust Control	$\mathcal{H}_\infty$ Control	Advanced $\mathcal{H}_\infty$ Control	Model & Controller Reductions
4	4	12	3.9
5	5.1∼5.3,5.5,5.7	13.2,13.4,13.6	7
6*	10	14	8
7*	12	15	5.4,5.7
8*	13.2,13.4,13.6	16	10.1
9	15	17*	16.1,16.2
10	16	18*	17.1
11		19*	19

Table 1.1: Possible choices for an one-semester course (* chapters may be omitted)

Table 1.1 lists several possible choices of topics for an one-semester course. A course on model and controller reductions may only include the concept of $\mathcal{H}_\infty$ control and the $\mathcal{H}_\infty$ controller formulas with the detailed proofs omitted as suggested in the above table.

1.3 Highlights of The Book

The key results in each chapter are highlighted below. Note that some of the statements in this section are not precise, they are true under certain assumptions that are not explicitly stated. Readers should consult the corresponding chapters for the exact statements and conditions.

Chapter 2 reviews some basic linear algebra facts and treats a special class of matrix dilation problems. In particular, we show

$$\min_X \left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\| = \max \left\{ \left\| \begin{bmatrix} C & A \end{bmatrix} \right\|, \left\| \begin{bmatrix} B \\ A \end{bmatrix} \right\| \right\}$$

and characterize all optimal (suboptimal) X.

Chapter 3 reviews some system theoretical concepts: controllability, observability, stabilizability, detectability, pole placement, observer theory, system poles and zeros, and state space realizations. Particularly, the balanced state space realizations are studied in some detail. We show that for a given stable transfer function $G(s)$ there is a state space realization $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ such that the controllability Gramian P and the observability Gramian Q defined below are equal and diagonal: $P = Q = \Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ where

$$AP + PA^* + BB^* = 0$$

$$A^*Q + QA + C^*C = 0.$$

Chapter 4 defines several norms for signals and introduces the $\mathcal{H}_2$ spaces and the $\mathcal{H}_\infty$ spaces. The input/output gains of a stable linear system under various input signals are characterized. We show that $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms come out naturally as measures of the worst possible performance for many classes of input signals. For example, let

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \in \mathcal{RH}_\infty, \quad g(t) = Ce^{At}B$$

Then $\|G\|_\infty = \sup \frac{\|g * u\|_2}{\|u\|_2}$ and $\sigma_1 \le \|G\|_\infty \le \int_0^\infty \|g(t)\|\, dt \le 2\sum_{i=1}^n \sigma_i$. Some state space methods of computing real rational $\mathcal{H}_2$ and $\mathcal{H}_\infty$ transfer matrix norms are also presented:

$$\|G\|_2^2 = \text{trace}(B^*QB) = \text{trace}(CPC^*)$$

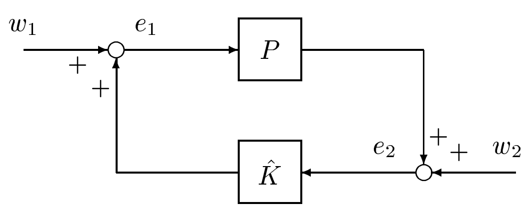
and

$$\|G\|_\infty = \max\{\gamma : \ H \text{ has an eigenvalue on the imaginary axis}\}$$

where

$$H = \begin{bmatrix} A & BB^*/\gamma^2 \\ -C^*C & -A^* \end{bmatrix}.$$

Chapter 5 introduces the feedback structure and discusses its stability and performance properties.



We show that the above closed-loop system is internally stable if and only if

$$\begin{bmatrix} I & -\hat{K} \\ -P & I \end{bmatrix}^{-1} = \begin{bmatrix} I + \hat{K}(I - P\hat{K})^{-1}P & \hat{K}(I - P\hat{K})^{-1} \\ (I - P\hat{K})^{-1}P & (I - P\hat{K})^{-1} \end{bmatrix} \in \mathcal{RH}_\infty.$$

Alternative characterizations of internal stability using coprime factorizations are also presented.

Chapter 6 introduces some multivariable versions of the Bode's sensitivity integral relations and Poisson integral formula. The sensitivity integral relations are used to study the design limitations imposed by bandwidth constraint and the open-loop unstable poles, while the Poisson integral formula is used to study the design constraints imposed by the non-minimum phase zeros. For example, let $S(s)$ be a sensitivity function, and let p_i be the right half plane poles of the open-loop system and η_i be the corresponding pole directions. Then we show that

$$\int_0^\infty \ln \overline{\sigma}(S(j\omega))d\omega = \pi\lambda_{\max}\left(\sum_i (\mathrm{Re}p_i)\eta_i\eta_i^*\right) + H_1, \quad H_1 \geq 0.$$

This equality shows that the design limitations in multivariable systems are dependent on the directionality properties of the sensitivity function as well as those of the poles (and zeros), in addition to the dependence upon pole (and zero) locations which is known in single-input single-output systems.

Chapter 7 considers the problem of reducing the order of a linear multivariable dynamical system using the balanced truncation method. Suppose

$$G(s) = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] \in \mathcal{RH}_\infty$$

is a balanced realization with controllability and observability Gramians $P = Q = \Sigma = \mathrm{diag}(\Sigma_1, \Sigma_2)$

$$\begin{aligned} \Sigma_1 &= \mathrm{diag}(\sigma_1 I_{s_1}, \sigma_2 I_{s_2}, \ldots, \sigma_r I_{s_r}) \\ \Sigma_2 &= \mathrm{diag}(\sigma_{r+1} I_{s_{r+1}}, \sigma_{r+2} I_{s_{r+2}}, \ldots, \sigma_N I_{s_N}). \end{aligned}$$

Then the truncated system $G_r(s) = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$ is stable and satisfies an additive error bound:

$$\|G(s) - G_r(s)\|_\infty \leq 2 \sum_{i=r+1}^{N} \sigma_i.$$

On the other hand, if $G^{-1} \in \mathcal{RH}_\infty$, and P and Q satisfy

$$PA^* + AP + BB^* = 0$$

$$Q(A - BD^{-1}C) + (A - BD^{-1}C)^*Q + C^*(D^{-1})^*D^{-1}C = 0$$

such that $P = Q = \mathrm{diag}(\Sigma_1, \Sigma_2)$ with G partitioned compatibly as before, then

$$G_r(s) = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$$

is stable and minimum phase, and satisfies respectively the following relative and multiplicative error bounds:

$$\left\|G^{-1}(G - G_r)\right\|_\infty \leq \prod_{i=r+1}^{N}\left(1 + 2\sigma_i(\sqrt{1+\sigma_i^2} + \sigma_i)\right) - 1$$

$$\left\|G_r^{-1}(G - G_r)\right\|_\infty \leq \prod_{i=r+1}^{N}\left(1 + 2\sigma_i(\sqrt{1+\sigma_i^2} + \sigma_i)\right) - 1.$$

Chapter 8 deals with the optimal Hankel norm approximation and its applications in $\mathcal{L}_\infty$ norm model reduction. We show that for a given $G(s)$ of McMillan degree n there is a $\hat{G}(s)$ of McMillan degree $r < n$ such that

$$\left\|G(s) - \hat{G}(s)\right\|_H = \inf \left\|G(s) - \hat{G}(s)\right\|_H = \sigma_{r+1}.$$

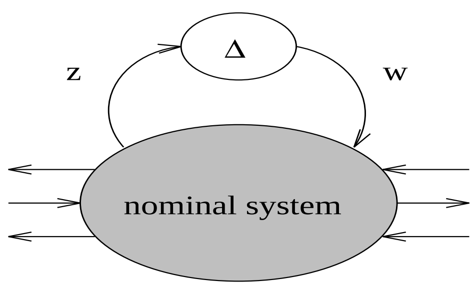
Moreover, there is a constant matrix D_0 such that

$$\left\|G(s) - \hat{G}(s) - D_0\right\|_\infty \leq \sum_{i=r+1}^{N} \sigma_i.$$

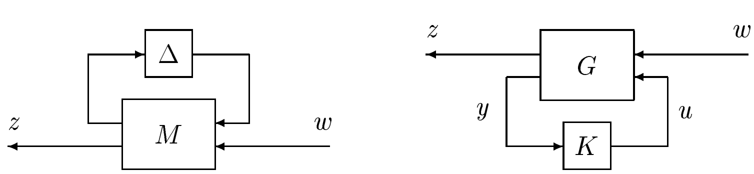
The well-known Nehari's theorem is also shown:

$$\inf_{Q \in \mathcal{RH}_\infty^-} \|G - Q\|_\infty = \|G\|_H = \sigma_1.$$

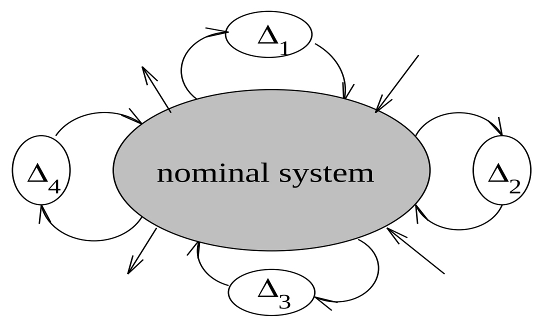
Chapter 9 derives robust stability tests for systems under various modeling assumptions through the use of a small gain theorem. In particular, we show that an uncertain system described below with an unstructured uncertainty $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty < 1$ is robustly stable if and only if the transfer function from w to z has $\mathcal{H}_\infty$ norm no greater than 1.



Chapter 10 introduces the linear fractional transformation (LFT). We show that many control problems can be formulated and treated in the LFT framework. In particular, we show that every analysis problem can be put in an LFT form with some structured $\Delta(s)$ and some interconnection matrix $M(s)$ and every synthesis problem can be put in an LFT form with a generalized plant $G(s)$ and a controller $K(s)$ to be designed.



Chapter 11 considers robust stability and performance for systems with multiple sources of uncertainties. We show that an uncertain system is robustly stable for all $\Delta_i \in \mathcal{RH}_\infty$ with $\|\Delta_i\|_\infty < 1$ if and only if the structured singular value (μ) of the corresponding interconnection model is no greater than 1.



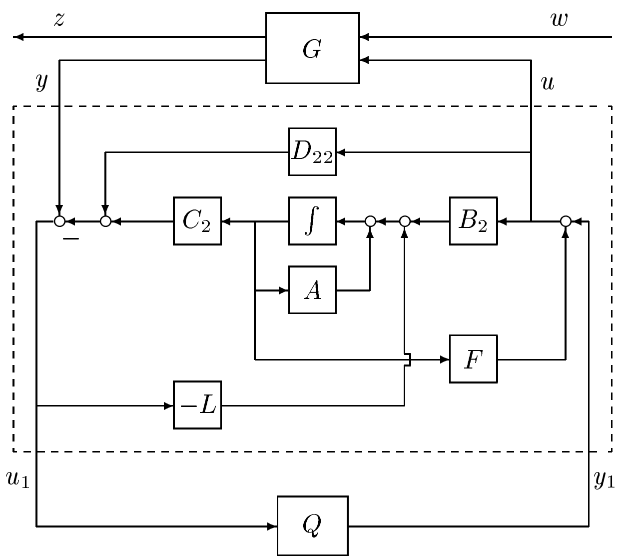
Chapter 12 characterizes all controllers that stabilize a given dynamical system $G(s)$ using the state space approach. The construction of the controller parameterization is done via separation theory and a sequence of special problems, which are so-called *full information (FI)* problems, *disturbance feedforward (DF)* problems, *full control* (FC) problems and *output estimation (OE)*. The relations among these special problems are established.

$$\begin{tikzcd}
\text{FI} \arrow[rr, leftrightarrow, "dual"] \arrow[dd, leftrightarrow, "\text{equivalent}"'] & & \text{FC} \arrow[dd, leftrightarrow, "\text{equivalent}"] \\
& & \\
\text{DF} \arrow[rr, leftrightarrow, "dual"] & & \text{OE}
\end{tikzcd}$$

For a given generalized plant

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

we show that all stabilizing controllers can be parameterized as the transfer matrix from y to u below where F and L are such that $A + LC_2$ and $A + B_2F$ are stable.



Chapter 13 studies the *Algebraic Riccati Equation* and the related problems: the properties of its solutions, the methods to obtain the solutions, and some applications. In particular, we study in detail the so-called stabilizing solution and its applications in matrix factorizations. A solution to the following ARE

$$A^*X + XA + XRX + Q = 0$$

is said to be a stabilizing solution if $A + RX$ is stable. Now let

$$H := \begin{bmatrix} A & R \\ -Q & -A^* \end{bmatrix}$$

and let $\mathcal{X}_-(H)$ be the stable H invariant subspace and

$$\mathcal{X}_-(H) = \mathrm{Im} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

where $X_1, X_2 \in \mathbb{C}^{n\times n}$. If X_1 is nonsingular, then $X := X_2X_1^{-1}$ is uniquely determined by H, denoted by $X = Ric(H)$.

A key result of this chapter is the relationship between the spectral factorization of a transfer function and the solution of a corresponding ARE. Suppose (A, B) is stabilizable and suppose either A has no eigenvalues on $j\omega$-axis or P is sign definite (i.e., $P \geq 0$ or $P \leq 0$) and (P, A) has no unobservable modes on the $j\omega$-axis. Define

$$\Phi(s) = \begin{bmatrix} B^*(-sI - A^*)^{-1} & I \end{bmatrix} \begin{bmatrix} P & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix}.$$

Then

$$\Phi(j\omega) > 0 \text{ for all } 0 \le \omega \le \infty$$

$\Longleftrightarrow$ $\exists$ a stabilizing solution X to

$$(A - BR^{-1}S^*)^*X + X(A - BR^{-1}S^*) - XBR^{-1}B^*X + P - SR^{-1}S^* = 0$$

$\Longleftrightarrow$ the Hamiltonian matrix

$$H = \begin{bmatrix} A - BR^{-1}S^* & -BR^{-1}B^* \\ -(P - SR^{-1}S^*) & -(A - BR^{-1}S^*)^* \end{bmatrix}$$

has no $j\omega$-axis eigenvalues.

Similarly,

$$\Phi(j\omega) \ge 0 \text{ for all } 0 \le \omega \le \infty$$

$\Longleftrightarrow$ $\exists$ a solution X to

$$(A - BR^{-1}S^*)^*X + X(A - BR^{-1}S^*) - XBR^{-1}B^*X + P - SR^{-1}S^* = 0$$

such that $\sigma(A - BR^{-1}S^* - BR^{-1}B^*X) \subset \overline{\mathbb{C}}_-$.

Furthermore, there exists a $M \in \mathcal{R}_p$ such that

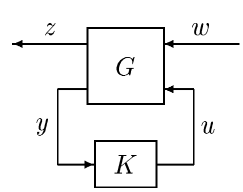
$$\Phi = M^{\sim}RM.$$

with

$$M = \left[\begin{array}{c|c} A & B \\ \hline -F & I \end{array}\right], \quad F = -R^{-1}(S^* + B^*X).$$

Chapter 14 treats the optimal control of linear time-invariant systems with quadratic performance criteria, i.e., LQR and $\mathcal{H}_2$ problems. We consider a dynamical system described by an LFT with

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right].$$



Define

$$H_2 := \begin{bmatrix} A & 0 \\ -C_1^*C_1 & -A^* \end{bmatrix} - \begin{bmatrix} B_2 \\ -C_1^*D_{12} \end{bmatrix} \begin{bmatrix} D_{12}^*C_1 & B_2^* \end{bmatrix}$$

$$J_2 := \begin{bmatrix} A^* & 0 \\ -B_1B_1^* & -A \end{bmatrix} - \begin{bmatrix} C_2^* \\ -B_1D_{21}^* \end{bmatrix} \begin{bmatrix} D_{21}B_1^* & C_2 \end{bmatrix}$$

$$X_2 := Ric(H_2) \geq 0, \quad Y_2 := Ric(J_2) \geq 0$$

$$F_2 := -(B_2^*X_2 + D_{12}^*C_1), \quad L_2 := -(Y_2C_2^* + B_1D_{21}^*).$$

Then the $\mathcal{H}_2$ optimal controller, i.e. the controller that minimizes $\|T_{zw}\|_2$, is given by

$$K_{opt}(s) := \left[\begin{array}{c|c} A + B_2F_2 + L_2C_2 & -L_2 \\ \hline F_2 & 0 \end{array}\right].$$

Chapter 15 solves a max-min problem, i.e., a full information (or state feedback) $\mathcal{H}_\infty$ control problem, which is the key to the $\mathcal{H}_\infty$ theory considered in the next chapter. Consider a dynamical system

$$\begin{aligned} \dot{x} &= Ax + B_1w + B_2u \\ z &= C_1x + D_{12}u, \quad D_{12}^*\begin{bmatrix} C_1 & D_{12} \end{bmatrix} = \begin{bmatrix} 0 & I \end{bmatrix}. \end{aligned}$$

Then we show that $\sup\limits_{w \in \mathcal{BL}_{2+}} \min\limits_{u \in \mathcal{L}_{2+}} \|z\|_2 < \gamma$ if and only if $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) \geq 0$ where

$$H_\infty := \begin{bmatrix} A & \gamma^{-2}B_1B_1^* - B_2B_2^* \\ -C_1^*C_1 & -A^* \end{bmatrix}$$

Furthermore, $u = F_\infty x$ with $F_\infty := -B_2^*X_\infty$ is an optimal control.

Chapter 16 considers a simplified $\mathcal{H}_\infty$ control problem with the generalized plant $G(s)$ as given in Chapter 14. We show that there exists an admissible controller such that $\|T_{zw}\|_\infty < \gamma$ iff the following three conditions hold:

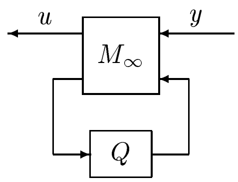
(i) $H_\infty \in$ dom(Ric) and $X_\infty := Ric(H_\infty) \geq 0$;

(ii) $J_\infty \in$ dom(Ric) and $Y_\infty := Ric(J_\infty) \geq 0$ where

$$J_\infty := \begin{bmatrix} A^* & \gamma^{-2}C_1^*C_1 - C_2^*C_2 \\ -B_1B_1^* & -A \end{bmatrix}.$$

(iii) $\rho(X_\infty Y_\infty) < \gamma^2$.

Moreover, the set of all admissible controllers such that $\|T_{zw}\|_\infty < \gamma$ equals the set of all transfer matrices from y to u in



$$M_\infty(s) = \left[\begin{array}{c|cc} \hat{A}_\infty & -Z_\infty L_\infty & Z_\infty B_2 \\ \hline F_\infty & 0 & I \\ -C_2 & I & 0 \end{array}\right]$$

where $Q \in \mathcal{RH}_\infty$, $\|Q\|_\infty < \gamma$ and

$$\hat{A}_\infty := A + \gamma^{-2} B_1 B_1^* X_\infty + B_2 F_\infty + Z_\infty L_\infty C_2$$

$$F_\infty := -B_2^* X_\infty, \quad L_\infty := -Y_\infty C_2^*, \quad Z_\infty := (I - \gamma^{-2} Y_\infty X_\infty)^{-1}.$$

Chapter 17 considers again the standard $\mathcal{H}_\infty$ control problem but with some assumptions in the last chapter relaxed. We indicate how the assumptions can be relaxed to accommodate other more complicated problems such as singular control problems. We also consider the integral control in the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ theory and show how the general $\mathcal{H}_\infty$ solution can be used to solve the $\mathcal{H}_\infty$ filtering problem. The conventional Youla parameterization approach to the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ problems is also outlined. Finally, the general state feedback $\mathcal{H}_\infty$ control problem and its relations with full information control and differential game problems are discussed.

Chapter 18 first solves a gap metric minimization problem. Let $P = \tilde{M}^{-1}\tilde{N}$ be a normalized left coprime factorization. Then we show that

$$\inf_{K \ stabilizing} \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I + PK)^{-1} \begin{bmatrix} I & P \end{bmatrix} \right\|_\infty$$

$$= \inf_{K \ stabilizing} \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I + PK)^{-1} \tilde{M}^{-1} \right\|_\infty = \left(\sqrt{1 - \left\| \begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} \right\|_H^2} \right)^{-1}.$$

This implies that there is a robustly stabilizing controller for

$$P_\Delta = (\tilde{M} + \tilde{\Delta}_M)^{-1}(\tilde{N} + \tilde{\Delta}_N)$$

with

$$\left\| \begin{bmatrix} \tilde{\Delta}_N & \tilde{\Delta}_M \end{bmatrix} \right\|_\infty < \epsilon$$

if and only if

$$\epsilon \le \sqrt{1 - \left\| \begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} \right\|_H^2}.$$

Using this stabilization result, a loop shaping design technique is proposed. The proposed technique uses only the basic concept of loop shaping methods and then a robust stabilization controller for the normalized coprime factor perturbed system is used to construct the final controller.

Chapter 19 considers the design of reduced order controllers by means of controller reduction. Special attention is paid to the controller reduction methods that preserve the closed-loop stability and performance. In particular, two $\mathcal{H}_\infty$ performance preserving reduction methods are proposed:

a) Let K_0 be a stabilizing controller such that $\|\mathcal{F}_\ell(G, K_0)\|_\infty < \gamma$. Then $\hat{K}$ is also a stabilizing controller such that $\left\|\mathcal{F}_\ell(G, \hat{K})\right\|_\infty < \gamma$ if

$$\left\|W_2^{-1}(\hat{K} - K_0)W_1^{-1}\right\|_\infty < 1$$

where W_1 and W_2 are some stable, minimum phase and invertible transfer matrices.

b) Let $K_0 = \Theta_{12}\Theta_{22}^{-1}$ be a central $\mathcal{H}_\infty$ controller such that $\|\mathcal{F}_\ell(G, K_0)\|_\infty < \gamma$ and let $\hat{U}, \hat{V} \in \mathcal{RH}_\infty$ be such that

$$\left\|\begin{bmatrix} \gamma^{-1}I & 0 \\ 0 & I \end{bmatrix} \Theta^{-1}\left(\begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix}\right)\right\|_\infty < 1/\sqrt{2}.$$

Then $\hat{K} = \hat{U}\hat{V}^{-1}$ is also a stabilizing controller such that $\|\mathcal{F}_\ell(G, \hat{K})\|_\infty < \gamma$.

Thus the controller reduction problem is converted to weighted model reduction problems for which some numerical methods are suggested.

Chapter 20 briefly introduces the Lagrange multiplier method for the design of fixed order controllers.

Chapter 21 discusses discrete time Riccati equations and some of their applications in discrete time control. Finally, the discrete time balanced model reduction is considered.

2

Linear Algebra

Some basic linear algebra facts will be reviewed in this chapter. The detailed treatment of this topic can be found in the references listed at the end of the chapter. Hence we shall omit most proofs and provide proofs only for those results that either cannot be easily found in the standard linear algebra textbooks or are insightful to the understanding of some related problems. We then treat a special class of matrix dilation problems which will be used in Chapters 8 and 17; however, most of the results presented in this book can be understood without the knowledge of the matrix dilation theory.

2.1 Linear Subspaces

Let $\mathbb{R}$ denote the real scalar field and $\mathbb{C}$ the complex scalar field. For the interest of this chapter, let $\mathbb{F}$ be either $\mathbb{R}$ or $\mathbb{C}$ and let $\mathbb{F}^n$ be the vector space over $\mathbb{F}$, i.e., $\mathbb{F}^n$ is either $\mathbb{R}^n$ or $\mathbb{C}^n$. Now let $x_1, x_2, \ldots, x_k \in \mathbb{F}^n$. Then an element of the form $\alpha_1 x_1 + \ldots + \alpha_k x_k$ with $\alpha_i \in \mathbb{F}$ is a *linear combination* over $\mathbb{F}$ of $x_1, \ldots, x_k$. The set of all linear combinations of $x_1, x_2, \ldots, x_k \in \mathbb{F}^n$ is a subspace called the *span* of $x_1, x_2, \ldots, x_k$, denoted by

$$\operatorname{span}\{x_1, x_2, \ldots, x_k\} := \{x = \alpha_1 x_1 + \ldots + \alpha_k x_k : \alpha_i \in \mathbb{F}\}.$$

A set of vectors $x_1, x_2, \ldots, x_k \in \mathbb{F}^n$ are said to be *linearly dependent* over $\mathbb{F}$ if there exists $\alpha_1, \ldots, \alpha_k \in \mathbb{F}$ not all zero such that $\alpha_1 x_2 + \ldots + \alpha_k x_k = 0$; otherwise they are said to be *linearly independent*.

Let S be a subspace of $\mathbb{F}^n$, then a set of vectors $\{x_1, x_2, \ldots, x_k\} \in S$ is called a *basis* for S if $x_1, x_2, \ldots, x_k$ are linearly independent and $S = span\{x_1, x_2, \ldots, x_k\}$. However, such a basis for a subspace S is not unique but all bases for S have the same number of elements. This number is called the *dimension* of S, denoted by $\dim(S)$.

A set of vectors $\{x_1, x_2, \ldots, x_k\}$ in $\mathbb{F}^n$ are mutually *orthogonal* if $x_i^* x_j = 0$ for all $i \neq j$ and *orthonormal* if $x_i^* x_j = \delta_{ij}$, where the superscript * denotes complex conjugate transpose and δ_{ij} is the Kronecker delta function with $\delta_{ij} = 1$ for $i = j$ and $\delta_{ij} = 0$ for $i \neq j$. More generally, a collection of subspaces $S_1, S_2, \ldots, S_k$ of $\mathbb{F}^n$ are mutually orthogonal if $x^* y = 0$ whenever $x \in S_i$ and $y \in S_j$ for $i \neq j$.

The *orthogonal complement* of a subspace $S \subset \mathbb{F}^n$ is defined by

$$S^\perp := \{y \in \mathbb{F}^n : y^* x = 0 \text{ for all } x \in S\}.$$

We call a set of vectors $\{u_1, u_2, \ldots, u_k\}$ an *orthonormal basis* for a subspace $S \in \mathbb{F}^n$ if they form a basis of S and are orthonormal. It is always possible to extend such a basis to a full orthonormal basis $\{u_1, u_2, \ldots, u_n\}$ for $\mathbb{F}^n$. Note that in this case

$$S^\perp = \operatorname{span}\{u_{k+1}, \ldots, u_n\},$$

and $\{u_{k+1}, \ldots, u_n\}$ is called an *orthonormal completion* of $\{u_1, u_2, \ldots, u_k\}$.

Let $A \in \mathbb{F}^{m \times n}$ be a linear transformation from $\mathbb{F}^n$ to $\mathbb{F}^m$, i.e.,

$$A : \mathbb{F}^n \longmapsto \mathbb{F}^m.$$

(Note that a vector $x \in \mathbb{F}^m$ can also be viewed as a linear transformation from $\mathbb{F}$ to $\mathbb{F}^m$, hence anything said for the general matrix case is also true for the vector case.) Then the *kernel* or *null space* of the linear transformation A is defined by

$$\operatorname{Ker}A = N(A) := \{x \in \mathbb{F}^n : Ax = 0\},$$

and the *image* or *range* of A is

$$\operatorname{Im}A = R(A) := \{y \in \mathbb{F}^m : y = Ax, \, x \in \mathbb{F}^n\}.$$

It is clear that $\operatorname{Ker}A$ is a subspace of $\mathbb{F}^n$ and $\operatorname{Im}A$ is a subspace of $\mathbb{F}^m$. Moreover, it can be easily seen that $\dim(\operatorname{Ker}A) + \dim(\operatorname{Im}A) = n$ and $\dim(\operatorname{Im}A) = \dim(\operatorname{Ker}A)^\perp$. Note that $(\operatorname{Ker}A)^\perp$ is a subspace of $\mathbb{F}^n$.

Let a_i, $i = 1, 2, \ldots, n$ denote the columns of a matrix $A \in \mathbb{F}^{m \times n}$, then

$$\operatorname{Im}A = \operatorname{span}\{a_1, a_2, \ldots, a_n\}.$$

The rank of a matrix A is defined by

$$\operatorname{rank}(A) = \dim(\operatorname{Im}A).$$

It is a fact that $\operatorname{rank}(A) = \operatorname{rank}(A^*)$, and thus the rank of a matrix equals the maximal number of independent rows or columns. A matrix $A \in \mathbb{F}^{m \times n}$ is said to have *full row rank* if $m \leq n$ and $\operatorname{rank}(A) = m$. Dually, it is said to have *full column rank* if $n \leq m$ and $\operatorname{rank}(A) = n$. A full rank square matrix is called a *nonsingular matrix*. It is easy to see that $\text{rank}(A) = \text{rank}(AT) = \text{rank}(PA)$ if T and P are nonsingular matrices with appropriate dimensions.

A square matrix $U \in F^{n\times n}$ whose columns form an orthonormal basis for $\mathbb{F}^n$ is called an *unitary matrix* (or *orthogonal matrix* if $\mathbb{F} = \mathbb{R}$), and it satisfies $U^*U = I = UU^*$. The following lemma is useful.

Lemma 2.1 *Let* $D = \begin{bmatrix} d_1 & \dots & d_k \end{bmatrix} \in \mathbb{F}^{n\times k}$ *$(n > k)$ be such that $D^*D = I$, so $d_i, i = 1, 2, \dots, k$ are orthonormal. Then there exists a matrix $D_\perp \in \mathbb{F}^{n\times(n-k)}$ such that $\begin{bmatrix} D & D_\perp \end{bmatrix}$ is a unitary matrix. Furthermore, the columns of $D_\perp$, $d_i, i = k+1, \dots, n$, form an orthonormal completion of $\{d_1, d_2, \dots, d_k\}$.*

The following results are standard:

Lemma 2.2 *Consider the linear equation*

$$AX = B$$

where $A \in \mathbb{F}^{n\times l}$ and $B \in \mathbb{F}^{n\times m}$ are given matrices. Then the following statements are equivalent:

(i) there exists a solution $X \in \mathbb{F}^{l\times m}$.

(ii) the columns of $B \in ImA$.

(iii) rank$\begin{bmatrix} A & B \end{bmatrix}$ =rank(A).

(iv) Ker(A^*) $\subset$Ker(B^*).

Furthermore, the solution, if it exists, is unique if and only if A has full column rank.

The following lemma concerns the rank of the product of two matrices.

Lemma 2.3 (Sylvester's inequality) *Let $A \in \mathbb{F}^{m\times n}$ and $B \in \mathbb{F}^{n\times k}$. Then*

$$\text{rank }(A) + \text{rank}(B) - n \le \text{rank}(AB) \le \min\{\text{rank }(A),\ \text{rank}(B)\}.$$

For simplicity, a matrix M with m_{ij} as its i-th row and j-th column's element will sometimes be denoted as $M = [m_{ij}]$ in this book. We will mostly use I as above to denote an identity matrix with compatible dimensions, but from time to time, we will use I_n to emphasis that it is an $n \times n$ identity matrix.

Now let $A = [a_{ij}] \in \mathbb{C}^{n\times n}$, then the *trace* of A is defined as

$$\text{Trace}(A) := \sum_{i=1}^{n} a_{ii}.$$

Trace has the following properties:

$$\text{Trace}(\alpha A) = \alpha\,\text{Trace}(A), \quad \forall \alpha \in \mathbb{C},\ A \in \mathbb{C}^{n\times n}$$

$$\text{Trace}(A + B) = \text{Trace}(A) + \text{Trace}(B), \quad \forall A,\ B \in \mathbb{C}^{n\times n}$$

$$\text{Trace}(AB) = \text{Trace}(BA), \quad \forall A \in \mathbb{C}^{n\times m},\ B \in \mathbb{C}^{m\times n}.$$

2.2 Eigenvalues and Eigenvectors

Let $A \in \mathbb{C}^{n\times n}$, then the *eigenvalues* of A are the n roots of its characteristic polynomial $p(\lambda) = \det(\lambda I - A)$. This set of roots is called the *spectrum* of A and is denoted by $\sigma(A)$ (not to be confused with singular values defined later). That is, $\sigma(A) := \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ if λ_i is a root of $p(\lambda)$. The maximal modulus of the eigenvalues is called the *spectral radius*, denoted by

$$\rho(A) := \max_{1\le i\le n} |\lambda_i|$$

where, as usual, $|\cdot|$ denotes the magnitude.

If $\lambda \in \sigma(A)$ then any nonzero vector $x \in \mathbb{C}^n$ that satisfies

$$Ax = \lambda x$$

is referred to as a *right eigenvector* of A. Dually, a nonzero vector y is called a *left eigenvector* of A if

$$y^*A = \lambda y^*.$$

It is a well known (but nontrivial) fact in linear algebra that any complex matrix admits a *Jordan Canonical* representation:

Theorem 2.4 *For any square complex matrix $A \in \mathbb{C}^{n\times n}$, there exists a nonsingular matrix T such that*

$$A = TJT^{-1}$$

where

$$J = diag\{J_1, J_2, \ldots, J_l\}$$

$$J_i = diag\{J_{i1}, J_{i2}, \ldots, J_{im_i}\}$$

$$J_{ij} = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix} \in \mathbb{C}^{n_{ij}\times n_{ij}}$$

with $\sum_{i=1}^{l}\sum_{j=1}^{m_i} n_{ij} = n$, and with $\{\lambda_i : i = 1, \ldots, l\}$ as the distinct eigenvalues of A.

The transformation T has the following form:

$$T = \begin{bmatrix} T_1 & T_2 & \ldots & T_l \end{bmatrix}$$

$$T_i = \begin{bmatrix} T_{i1} & T_{i2} & \ldots & T_{im_i} \end{bmatrix}$$

$$T_{ij} = \begin{bmatrix} t_{ij1} & t_{ij2} & \ldots & t_{ijn_{ij}} \end{bmatrix}$$

where t_{ij1} are the eigenvectors of A,

$$At_{ij1} = \lambda_i t_{ij1},$$

and $t_{ijk} \neq 0$ defined by the following linear equations for $k \geq 2$

$$(A - \lambda_i I)t_{ijk} = t_{ij(k-1)}$$

are called the *generalized eigenvectors* of A. For a given integer $q \leq n_{ij}$, the generalized eigenvectors $t_{ijl}, \forall l < q$, are called the *lower rank generalized eigenvectors of* t_{ijq}.

Definition 2.1 A square matrix $A \in \mathbb{R}^{n \times n}$ is called *cyclic* if the Jordan canonical form of A has one and only one Jordan block associated with each distinct eigenvalue.

More specifically, a matrix A is cyclic if its Jordan form has $m_i = 1, i = 1, \ldots, l$. Clearly, a square matrix A with all distinct eigenvalues is cyclic and can be diagonalized:

$$A \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix}.$$

In this case, A has the following *spectral decomposition*:

$$A = \sum_{i=1}^{n} \lambda_i x_i y_i^*$$

where $y_i \in \mathbb{C}^n$ is given by

$$\begin{bmatrix} y_1^* \\ y_2^* \\ \vdots \\ y_n^* \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix}^{-1}.$$

In general, eigenvalues need not be real, and neither do their corresponding eigenvectors. However, if A is real and λ is a real eigenvalue of A, then there is a real eigenvector corresponding to λ. In the case that all eigenvalues of a matrix A are real[1], we will denote $\lambda_{max}(A)$ for the largest eigenvalue of A and $\lambda_{min}(A)$ for the smallest eigenvalue. In particular, if A is a Hermitian matrix, then there exist a unitary matrix U and a real diagonal matrix Λ such that $A = U\Lambda U^*$, where the diagonal elements of Λ are the eigenvalues of A and the columns of U are the eigenvectors of A.

The following theorem is useful in linear system theory.

Theorem 2.5 (Cayley-Hamilton) *Let $A \in \mathbb{C}^{n \times n}$ and denote*

$$\det(\lambda I - A) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n.$$

Then

$$A^n + a_1 A^{n-1} + \cdots + a_n I = 0.$$

[1] For example, this is the case if A is *Hermitian*, i.e., $A = A^*$.

This is obvious if A has distinct eigenvalues. Since

$$A^n + a_1 A^{n-1} + \cdots + a_n I = T^{-1} \operatorname{diag}\{\ldots, \lambda_i^n + a_1 \lambda_i^{n-1} + \cdots + a_n, \ldots\} T = 0,$$

and λ_i is an eigenvalue of A. The proof for the general case follows from the following lemma.

Lemma 2.6 *Let $A \in \mathbb{C}^{n\times n}$. Then*

$$(\lambda I - A)^{-1} = \frac{1}{\det(\lambda I - A)}(R_1 \lambda^{n-1} + R_2 \lambda^{n-2} + \cdots + R_n)$$

and

$$\det(\lambda I - A) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n$$

where a_i and R_i can be computed from the following recursive formulas:

$$\begin{array}{lcllcl}
a_1 & = & -\operatorname{Trace} A & R_1 & = & I \\
a_2 & = & -\frac{1}{2}\operatorname{Trace}(R_2 A) & R_2 & = & R_1 A + a_1 I \\
 & \vdots & & & \vdots & \\
a_{n-1} & = & -\frac{1}{n-1}\operatorname{Trace}(R_{n-1} A) & R_n & = & R_{n-1} A + a_{n-1} I \\
a_n & = & -\frac{1}{n}\operatorname{Trace}(R_n A) & 0 & = & R_n A + a_n I.
\end{array}$$

The proof is left to the reader as an exercise. Note that the Cayley-Hamilton Theorem follows from the fact that

$$0 = R_n A + a_n I = A^n + a_1 A^{n-1} + \cdots + a_n I.$$

2.3 Matrix Inversion Formulas

Let A be a square matrix partitioned as follows

$$A := \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

where A_{11} and A_{22} are also square matrices. Now suppose A_{11} is nonsingular, then A has the following decomposition:

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & I \end{bmatrix} \begin{bmatrix} A_{11} & 0 \\ 0 & \Delta \end{bmatrix} \begin{bmatrix} I & A_{11}^{-1}A_{12} \\ 0 & I \end{bmatrix}$$

with $\Delta := A_{22} - A_{21}A_{11}^{-1}A_{12}$, and A is nonsingular iff Δ is nonsingular.

Dually, if A_{22} is nonsingular, then

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & A_{12}A_{22}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \hat{\Delta} & 0 \\ 0 & A_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ A_{22}^{-1}A_{21} & I \end{bmatrix}$$

with $\overline{\Delta} := A_{11} - A_{12}A_{22}^{-1}A_{21}$, and A is nonsingular iff $\overline{\Delta}$ is nonsingular. The matrix Δ ($\overline{\Delta}$) is called the *Schur complement* of A_{11} (A_{22}) in A.

Moreover, if A is nonsingular, then

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}\Delta^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}\Delta^{-1} \\ -\Delta^{-1}A_{21}A_{11}^{-1} & \Delta^{-1} \end{bmatrix}$$

and

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \overline{\Delta}^{-1} & -\overline{\Delta}^{-1}A_{12}A_{22}^{-1} \\ -A_{22}^{-1}A_{21}\overline{\Delta}^{-1} & A_{22}^{-1} + A_{22}^{-1}A_{21}\overline{\Delta}^{-1}A_{12}A_{22}^{-1} \end{bmatrix}.$$

The above matrix inversion formulas are particularly simple if A is block triangular:

$$\begin{bmatrix} A_{11} & 0 \\ A_{21} & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & 0 \\ -A_{22}^{-1}A_{21}A_{11}^{-1} & A_{22}^{-1} \end{bmatrix}$$

$$\begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}^{-1} = \begin{bmatrix} A_{11}^{-1} & -A_{11}^{-1}A_{12}A_{22}^{-1} \\ 0 & A_{22}^{-1} \end{bmatrix}.$$

The following identity is also very useful. Suppose A_{11} and A_{22} are both nonsingular matrices, then

$$(A_{11} - A_{12}A_{22}^{-1}A_{21})^{-1} = A_{11}^{-1} + A_{11}^{-1}A_{12}(A_{22} - A_{21}A_{11}^{-1}A_{12})^{-1}A_{21}A_{11}^{-1}.$$

As a consequence of the matrix decomposition formulas mentioned above, we can calculate the determinant of a matrix by using its sub-matrices. Suppose A_{11} is nonsingular, then

$$\det A = \det A_{11} \det(A_{22} - A_{21}A_{11}^{-1}A_{12}).$$

On the other hand, if A_{22} is nonsingular, then

$$\det A = \det A_{22} \det(A_{11} - A_{12}A_{22}^{-1}A_{21}).$$

In particular, for any $B \in \mathbb{C}^{m\times n}$ and $C \in \mathbb{C}^{n\times m}$, we have

$$\det \begin{bmatrix} I_m & B \\ -C & I_n \end{bmatrix} = \det(I_n + CB) = \det(I_m + BC)$$

and for $x, y \in \mathbb{C}^n$

$$\det(I_n + xy^*) = 1 + y^*x.$$

2.4 Matrix Calculus

Let $X = [x_{ij}] \in \mathbb{C}^{m\times n}$ be a real or complex matrix and $F(X) \in \mathbb{C}$ be a scalar real or complex function of X; then the derivative of $F(X)$ with respect to X is defined as

$$\frac{\partial}{\partial X} F(X) := \left[\frac{\partial}{\partial x_{ij}} F(X)\right].$$

Let A and B be constant complex matrices with compatible dimensions. Then the following is a list of formulas for the derivatives[2]:

$$\begin{aligned}
\tfrac{\partial}{\partial X}\operatorname{Trace}\{AXB\} &= A^T B^T \\
\tfrac{\partial}{\partial X}\operatorname{Trace}\{AX^T B\} &= BA \\
\tfrac{\partial}{\partial X}\operatorname{Trace}\{AXBX\} &= A^T X^T B^T + B^T X^T A^T \\
\tfrac{\partial}{\partial X}\operatorname{Trace}\{AXBX^T\} &= A^T X B^T + AXB \\
\tfrac{\partial}{\partial X}\operatorname{Trace}\{X^k\} &= k(X^{k-1})^T \\
\tfrac{\partial}{\partial X}\operatorname{Trace}\{AX^k\} &= \left(\textstyle\sum_{i=0}^{k-1} X^i A X^{k-i-1}\right)^T \\
\tfrac{\partial}{\partial X}\operatorname{Trace}\{AX^{-1}B\} &= -(X^{-1}BAX^{-1})^T \\
\tfrac{\partial}{\partial X}\log\det X &= (X^T)^{-1} \\
\tfrac{\partial}{\partial X}\det X^T &= \tfrac{\partial}{\partial X}\det X = (det X)(X^T)^{-1} \\
\tfrac{\partial}{\partial X}\det\{X^k\} &= k(det X^k)(X^T)^{-1}.
\end{aligned}$$

And finally, the derivative of a matrix $A(\alpha) \in \mathbb{C}^{m\times n}$ with respect to a scalar $\alpha \in \mathbb{C}$ is defined as

$$\frac{dA}{d\alpha} := \left[\frac{da_{ij}}{d\alpha}\right]$$

so that all the rules applicable to a scalar function also apply here. In particular, we have

$$\begin{aligned}
\frac{d(AB)}{d\alpha} &= \frac{dA}{d\alpha}B + A\frac{dB}{d\alpha} \\
\frac{dA^{-1}}{d\alpha} &= -A^{-1}\frac{dA}{d\alpha}A^{-1}.
\end{aligned}$$

[2]Note that transpose rather than complex conjugate transpose should be used in the list even if the involved matrices are complex matrices.

2.5 Kronecker Product and Kronecker Sum

Let $A \in \mathbb{C}^{m \times n}$ and $B \in \mathbb{C}^{p \times q}$, then the *Kronecker product* of A and B is defined as

$$A \otimes B := \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \in \mathbb{C}^{mp \times nq}.$$

Furthermore, if the matrices A and B are square and $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{m \times m}$ then the *Kronecker sum* of A and B is defined as

$$A \oplus B := (A \otimes I_m) + (I_n \otimes B) \in \mathbb{C}^{nm \times nm}.$$

Let $X \in \mathbb{C}^{m \times n}$ and let $\mathrm{vec}(X)$ denote the vector formed by stacking the columns of X into one long vector:

$$\mathrm{vec}(X) := \begin{bmatrix} x_{11} \\ x_{21} \\ \vdots \\ x_{m1} \\ x_{12} \\ x_{22} \\ \vdots \\ x_{1n} \\ x_{2n} \\ \vdots \\ x_{mn} \end{bmatrix}.$$

Then for any matrices $A \in \mathbb{C}^{k \times m}$, $B \in \mathbb{C}^{n \times l}$, and $X \in \mathbb{C}^{m \times n}$, we have

$$\mathrm{vec}(AXB) = (B^T \otimes A)\mathrm{vec}(X).$$

Consequently, if $k = m$ and $l = n$, then

$$\mathrm{vec}(AX + XB) = (B^T \oplus A)\mathrm{vec}(X).$$

Let $A \in \mathbb{C}^{n \times n}$ and $B \in \mathbb{C}^{m \times m}$, and let $\{\lambda_i,\ i = 1, \ldots, n\}$ be the eigenvalues of A and $\{\mu_j,\ j = 1, \ldots, m\}$ be the eigenvalues of B. Then we have the following properties:

- The eigenvalues of $A \otimes B$ are the mn numbers $\lambda_i \mu_j$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.
- The eigenvalues of $A \oplus B = (A \otimes I_m) + (I_n \otimes B)$ are the mn numbers $\lambda_i + \mu_j$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$.

- Let $\{x_i,\ i = 1, \ldots, n\}$ be the eigenvectors of A and let $\{y_j,\ j = 1, \ldots, m\}$ be the eigenvectors of B. Then the eigenvectors of $A \otimes B$ and $A \oplus B$ correspond to the eigenvalues $\lambda_i \mu_j$ and $\lambda_i + \mu_j$ are $x_i \otimes y_j$.

Using these properties, we can show the following Lemma.

Lemma 2.7 *Consider the Sylvester equation*

$$AX + XB = C \tag{2.1}$$

where $A \in \mathbb{F}^{n \times n}$, $B \in \mathbb{F}^{m \times m}$, and $C \in \mathbb{F}^{n \times m}$ are given matrices. There exists a unique solution $X \in \mathbb{F}^{n \times m}$ if and only if $\lambda_i(A) + \lambda_j(B) \neq 0,\ \forall i = 1, 2, \ldots, n$ and $j = 1, 2, \ldots, m$.

In particular, if $B = A^*$, (2.1) is called the "Lyapunov Equation"; and the necessary and sufficient condition for the existence of a unique solution is that $\lambda_i(A) + \bar{\lambda}_j(A) \neq 0,\ \forall i, j = 1, 2, \ldots, n$

Proof. Equation (2.1) can be written as a linear matrix equation by using the Kronecker product:

$$(B^T \oplus A)\text{vec}(X) = \text{vec}(C).$$

Now this equation has a unique solution iff $B^T \oplus A$ is nonsingular. Since the eigenvalues of $B^T \oplus A$ have the form of $\lambda_i(A) + \lambda_j(B^T) = \lambda_i(A) + \lambda_j(B)$, the conclusion follows. □

The properties of the Lyapunov equations will be studied in more detail in the next chapter.

2.6 Invariant Subspaces

Let $A : \mathbb{C}^n \longmapsto \mathbb{C}^n$ be a linear transformation, λ be an eigenvalue of A, and x be a corresponding eigenvector, respectively. Then $Ax = \lambda x$ and $A(\alpha x) = \lambda(\alpha x)$ for any $\alpha \in \mathbb{C}$. Clearly, the eigenvector x defines an one-dimensional subspace that is invariant with respect to pre-multiplication by A since $A^k x = \lambda^k x, \forall k$. In general, a subspace $S \subset \mathbb{C}^n$ is called *invariant* for the transformation A, or A-*invariant*, if $Ax \in S$ for every $x \in S$. In other words, that S is invariant for A means that the image of S under A is contained in S: $AS \subset S$. For example, $\{0\}$, $\mathbb{C}^n$, Ker$A$, and ImA are all A-invariant subspaces.

As a generalization of the one dimensional invariant subspace induced by an eigenvector, let $\lambda_1, \ldots, \lambda_k$ be eigenvalues of A (not necessarily distinct), and let x_i be the corresponding eigenvectors and the generalized eigenvectors. Then $S = \text{span}\{x_1, \ldots, x_k\}$ is an A-invariant subspace provided that all the lower rank generalized eigenvectors are included. More specifically, let $\lambda_1 = \lambda_2 = \cdots = \lambda_l$ be eigenvalues of A, and let $x_1, x_2, \ldots, x_l$ be the corresponding eigenvector and the generalized eigenvectors obtained through the following equations:

$$\begin{aligned}(A - \lambda_1 I)x_1 &= 0 \\ (A - \lambda_1 I)x_2 &= x_1 \\ &\vdots \\ (A - \lambda_1 I)x_l &= x_{l-1}.\end{aligned}$$

Then a subspace S with $x_t \in S$ for some $t \leq l$ is an A-invariant subspace only if all lower rank eigenvectors and generalized eigenvectors of x_t are in S, i.e., $x_i \in S,\ \forall 1 \leq i \leq t$. This will be further illustrated in Example 2.1.

On the other hand, if S is a nontrivial subspace[3] and is A-invariant, then there is $x \in S$ and λ such that $Ax = \lambda x$.

An A-invariant subspace $S \subset \mathbb{C}^n$ is called a *stable invariant subspace* if all the eigenvalues of A constrained to S have negative real parts. Stable invariant subspaces will play an important role in computing the stabilizing solutions to the algebraic Riccati equations in Chapter 13.

Example 2.1 Suppose a matrix A has the following Jordan canonical form

$$A \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & x_3 & x_4 \end{bmatrix} \begin{bmatrix} \lambda_1 & 1 & & \\ & \lambda_1 & & \\ & & \lambda_3 & \\ & & & \lambda_4 \end{bmatrix}$$

with $\mathrm{Re}\lambda_1 < 0$, $\lambda_3 < 0$, and $\lambda_4 > 0$. Then it is easy to verify that

$$\begin{array}{lllllllll} S_1 &=& \mathrm{span}\{x_1\} & S_{12} &=& \mathrm{span}\{x_1, x_2\} & S_{123} &=& \mathrm{span}\{x_1, x_2, x_3\} \\ S_3 &=& \mathrm{span}\{x_3\} & S_{13} &=& \mathrm{span}\{x_1, x_3\} & S_{124} &=& \mathrm{span}\{x_1, x_2, x_4\} \\ S_4 &=& \mathrm{span}\{x_4\} & S_{14} &=& \mathrm{span}\{x_1, x_4\} & S_{34} &=& \mathrm{span}\{x_3, x_4\} \end{array}$$

are all A-invariant subspaces. Moreover, S_1, S_3, S_{12}, S_{13}, and S_{123} are stable A-invariant subspaces. However, the subspaces $S_2 = \mathrm{span}\{x_2\}$, $S_{23} = \mathrm{span}\{x_2, x_3\}$, $S_{24} = \mathrm{span}\{x_2, x_4\}$, and $S_{234} = \mathrm{span}\{x_2, x_3, x_4\}$ are not A-invariant subspaces since the lower rank generalized eigenvector x_1 of x_2 is not in these subspaces. To illustrate, consider the subspace S_{23}. Then by definition, $Ax_2 \in S_{23}$ if it is an A-invariant subspace. Since

$$Ax_2 = \lambda x_2 + x_1,$$

$Ax_2 \in S_{23}$ would require that x_1 be a linear combination of x_2 and x_3, but this is impossible since x_1 is independent of x_2 and x_3. $\diamond$

[3]We will say subspace S is trivial if $S = \{0\}$.

2.7 Vector Norms and Matrix Norms

In this section, we will define vector and matrix norms. Let X be a vector space, a real-valued function $\|\cdot\|$ defined on X is said to be a *norm* on X if it satisfies the following properties:

(i) $\|x\| \geq 0$ (positivity);

(ii) $\|x\| = 0$ if and only if $x = 0$ (positive definiteness);

(iii) $\|\alpha x\| = |\alpha| \, \|x\|$, for any scalar α (homogeneity);

(iv) $\|x + y\| \leq \|x\| + \|y\|$ (triangle inequality)

for any $x \in X$ and $y \in X$. A function is said to be a *semi-norm* if it satisfies (i), (iii), and (iv) but not necessarily (ii).

Let $x \in \mathbb{C}^n$. Then we define the vector p-norm of x as

$$\|x\|_p := \left(\sum_{i=1}^n |x_i|^p \right)^{1/p}, \text{ for } 1 \leq p \leq \infty.$$

In particular, when $p = 1, 2, \infty$ we have

$$\|x\|_1 := \sum_{i=1}^n |x_i|;$$

$$\|x\|_2 := \sqrt{\sum_{i=1}^n |x_i|^2};$$

$$\|x\|_\infty := \max_{1 \leq i \leq n} |x_i|.$$

Clearly, norm is an abstraction and extension of our usual concept of length in 3-dimensional Euclidean space. So a norm of a vector is a measure of the vector "length", for example $\|x\|_2$ is the Euclidean distance of the vector x from the origin. Similarly, we can introduce some kind of measure for a matrix.

Let $A = [a_{ij}] \in \mathbb{C}^{m \times n}$, then the matrix norm *induced* by a vector p-norm is defined as

$$\|A\|_p := \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}.$$

In particular, for $p = 1, 2, \infty$, the corresponding induced matrix norm can be computed as

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^m |a_{ij}| \quad \text{(column sum)} \; ;$$

$$\|A\|_2 = \sqrt{\lambda_{\max}(A^*A)} \ ;$$

$$\|A\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}| \quad \text{(row sum)} \ .$$

The matrix norms induced by vector p-norms are sometimes called *induced p-norms.* This is because $\|A\|_p$ is defined by or induced from a vector p-norm. In fact, A can be viewed as a mapping from a vector space $\mathbb{C}^n$ equipped with a vector norm $\|\cdot\|_p$ to another vector space $\mathbb{C}^m$ equipped with a vector norm $\|\cdot\|_p$. So from a system theoretical point of view, the induced norms have the interpretation of input/output amplification gains.

We shall adopt the following convention throughout the book for the vector and matrix norms unless specified otherwise: let $x \in \mathbb{C}^n$ and $A \in \mathbb{C}^{m\times n}$, then we shall denote the Euclidean 2-norm of x simply by

$$\|x\| := \|x\|_2$$

and the induced 2-norm of A by

$$\|A\| := \|A\|_2 \ .$$

The Euclidean 2-norm has some very nice properties:

Lemma 2.8 *Let $x \in \mathbb{F}^n$ and $y \in \mathbb{F}^m$.*

1. *Suppose $n \ge m$. Then $\|x\| = \|y\|$ iff there is a matrix $U \in \mathbb{F}^{n\times m}$ such that $x = Uy$ and $U^*U = I$.*
2. *Suppose $n = m$. Then $|x^*y| \le \|x\|\,\|y\|$. Moreover, the equality holds iff $x = \alpha y$ for some $\alpha \in \mathbb{F}$ or $y = 0$.*
3. *$\|x\| \le \|y\|$ iff there is a matrix $\Delta \in \mathbb{F}^{n\times m}$ with $\|\Delta\| \le 1$ such that $x = \Delta y$. Furthermore, $\|x\| < \|y\|$ iff $\|\Delta\| < 1$.*
4. *$\|Ux\| = \|x\|$ for any appropriately dimensioned unitary matrices U.*

Another often used matrix norm is the so called *Frobenius norm.* It is defined as

$$\|A\|_F := \sqrt{\operatorname{Trace}(A^*A)} = \sqrt{\sum_{i=1}^{m}\sum_{j=1}^{n} |a_{ij}|^2} \ .$$

However, the Frobenius norm is not an induced norm.

The following properties of matrix norms are easy to show:

Lemma 2.9 *Let A and B be any matrices with appropriate dimensions. Then*

1. *$\rho(A) \leq \|A\|$ (This is also true for F norm and any induced matrix norm).*

2. *$\|AB\| \leq \|A\| \|B\|$. In particular, this gives $\left\|A^{-1}\right\| \geq \|A\|^{-1}$ if A is invertible. (This is also true for any induced matrix norm.)*

3. *$\|UAV\| = \|A\|$, and $\|UAV\|_F = \|A\|_F$, for any appropriately dimensioned unitary matrices U and V.*

4. *$\|AB\|_F \leq \|A\| \|B\|_F$ and $\|AB\|_F \leq \|B\| \|A\|_F$.*

Note that although pre-multiplication or post-multiplication of a unitary matrix on a matrix does not change its induced 2-norm and F-norm, it does change its eigenvalues. For example, let

$$A = \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}.$$

Then $\lambda_1(A) = 1, \lambda_2(A) = 0$. Now let

$$U = \begin{bmatrix} \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix};$$

then U is a unitary matrix and

$$UA = \begin{bmatrix} \sqrt{2} & 0 \\ 0 & 0 \end{bmatrix}$$

with $\lambda_1(UA) = \sqrt{2}$, $\lambda_2(UA) = 0$. This property is useful in some matrix perturbation problems, particularly, in the computation of bounds for structured singular values which will be studied in Chapter 10.

Lemma 2.10 *Let A be a block partitioned matrix with*

$$A = \begin{bmatrix} A_{11} & A_{12} & \cdots & A_{1q} \\ A_{21} & A_{22} & \cdots & A_{2q} \\ \vdots & \vdots & & \vdots \\ A_{m1} & A_{m2} & \cdots & A_{mq} \end{bmatrix} =: [A_{ij}],$$

and let each A_{ij} be an appropriately dimensioned matrix. Then for any induced matrix p-norm

$$\|A\|_p \leq \left\| \begin{bmatrix} \|A_{11}\|_p & \|A_{12}\|_p & \cdots & \|A_{1q}\|_p \\ \|A_{21}\|_p & \|A_{22}\|_p & \cdots & \|A_{2q}\|_p \\ \vdots & \vdots & & \vdots \\ \|A_{m1}\|_p & \|A_{m2}\|_p & \cdots & \|A_{mq}\|_p \end{bmatrix} \right\|_p . \tag{2.2}$$

Further, the inequality becomes an equality if the F-norm is used.

Proof. It is obvious that if the F-norm is used, then the right hand side of inequality (2.2) equals the left hand side. Hence only the induced p-norm cases, $1 \leq p \leq \infty$, will be shown. Let a vector x be partitioned consistently with A as

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_q \end{bmatrix};$$

and note that

$$\|x\|_p = \left\| \begin{bmatrix} \|x_1\|_p \\ \|x_2\|_p \\ \vdots \\ \|x_q\|_p \end{bmatrix} \right\|_p .$$

Then

$$\begin{aligned}
\|[A_{ij}]\|_p &:= \sup_{\|x\|_p=1} \|[A_{ij}]\, x\|_p = \sup_{\|x\|_p=1} \left\| \begin{bmatrix} \sum_{j=1}^q A_{1j}x_j \\ \sum_{j=1}^q A_{2j}x_j \\ \vdots \\ \sum_{j=1}^q A_{mj}x_j \end{bmatrix} \right\|_p \\
&= \sup_{\|x\|_p=1} \left\| \begin{bmatrix} \left\|\sum_{j=1}^q A_{1j}x_j\right\|_p \\ \left\|\sum_{j=1}^q A_{2j}x_j\right\|_p \\ \vdots \\ \left\|\sum_{j=1}^q A_{mj}x_j\right\|_p \end{bmatrix} \right\|_p \leq \sup_{\|x\|_p=1} \left\| \begin{bmatrix} \sum_{j=1}^q \|A_{1j}\|_p \|x_j\|_p \\ \sum_{j=1}^q \|A_{2j}\|_p \|x_j\|_p \\ \vdots \\ \sum_{j=1}^q \|A_{mj}\|_p \|x_j\|_p \end{bmatrix} \right\|_p \\
&= \sup_{\|x\|_p=1} \left\| \begin{bmatrix} \|A_{11}\|_p & \|A_{12}\|_p & \cdots & \|A_{1q}\|_p \\ \|A_{21}\|_p & \|A_{22}\|_p & \cdots & \|A_{2q}\|_p \\ \vdots & \vdots & & \vdots \\ \|A_{m1}\|_p & \|A_{m2}\|_p & \cdots & \|A_{1q}\|_p \end{bmatrix} \begin{bmatrix} \|x_1\|_p \\ \|x_2\|_p \\ \vdots \\ \|x_q\|_p \end{bmatrix} \right\|_p \\
&\leq \sup_{\|x\|_p=1} \left\| \left[\|A_{ij}\|_p \right] \right\|_p \|x\|_p \\
&= \left\| \left[\|A_{ij}\|_p \right] \right\|_p .
\end{aligned}$$

□

2.8 Singular Value Decomposition

A very useful tool in matrix analysis is *Singular Value Decomposition (SVD)*. It will be seen that singular values of a matrix are good measures of the "size" of the matrix and that the corresponding singular vectors are good indications of strong/weak input or output directions.

Theorem 2.11 *Let $A \in \mathbb{F}^{m\times n}$. There exist unitary matrices*

$$\begin{aligned} U &= [u_1, u_2, \ldots, u_m] \in \mathbb{F}^{m\times m} \\ V &= [v_1, v_2, \ldots, v_n] \in \mathbb{F}^{n\times n} \end{aligned}$$

such that

$$A = U\Sigma V^*, \quad \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix}$$

where

$$\Sigma_1 = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_p \end{bmatrix}$$

and

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p \geq 0, \;\; p = min\{m, n\}.$$

Proof. Let $\sigma = \|A\|$ and without loss of generality assume $m \geq n$. Then from the definition of $\|A\|$, there exists a $z \in \mathbb{F}^n$ such that

$$\|Az\| = \sigma \|z\|.$$

By Lemma 2.8, there is a matrix $\tilde{U} \in F^{m\times n}$ such that $\tilde{U}^*\tilde{U} = I$ and

$$Az = \sigma\tilde{U}z.$$

Now let

$$x = \frac{z}{\|z\|} \in \mathbb{F}^n, \quad y = \frac{\tilde{U}z}{\left\|\tilde{U}z\right\|} \in \mathbb{F}^m.$$

We have $Ax = \sigma y$. Let

$$V = \begin{bmatrix} x & V_1 \end{bmatrix} \in \mathbb{F}^{n\times n}$$

and

$$U = \begin{bmatrix} y & U_1 \end{bmatrix} \in \mathbb{F}^{m\times m}$$

be unitary.[4] Consequently, U^*AV has the following structure:

$$A_1 := U^*AV = \begin{bmatrix} \sigma & w^* \\ 0 & B \end{bmatrix}$$

[4]Recall that it is always possible to extend an orthonormal set of vectors to an orthonormal basis for the whole space.

where $w \in \mathbb{F}^{n-1}$ and $B \in \mathbb{F}^{(m-1)\times(n-1)}$.

Since

$$\left\| A_1^* \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \right\|_2^2 = (\sigma^2 + w^*w),$$

it follows that $\|A_1\|^2 \geq \sigma^2 + w^*w$. But since $\sigma = \|A\| = \|A_1\|$, we must have $w = 0$. An obvious induction argument gives

$$U^*AV = \Sigma.$$

This completes the proof. □

The σ_i is the i-th *singular value* of A, and the vectors u_i and v_j are, respectively, the i-th *left singular vector* and the j-th *right singular vector*. It is easy to verify that

$$\begin{aligned} Av_i &= \sigma_i u_i \\ A^*u_i &= \sigma_i v_i. \end{aligned}$$

The above equations can also be written as

$$\begin{aligned} A^*Av_i &= \sigma_i^2 v_i \\ AA^*u_i &= \sigma_i^2 u_i. \end{aligned}$$

Hence σ_i^2 is an eigenvalue of AA^* or A^*A, u_i is an eigenvector of AA^*, and v_i is an eigenvector of A^*A.

The following notations for singular values are often adopted:

$$\overline{\sigma}(A) = \sigma_{max}(A) = \sigma_1 = \text{the largest singular value of } A;$$

and

$$\underline{\sigma}(A) = \sigma_{min}(A) = \sigma_p = \text{the smallest singular value of } A\ .$$

Geometrically, the singular values of a matrix A are precisely the lengths of the semi-axes of the hyperellipsoid E defined by

$$E = \{y : y = Ax,\ x \in \mathbb{C}^n,\ \|x\| = 1\}.$$

Thus v_1 is the direction in which $\|y\|$ is largest for all $\|x\| = 1$; while v_n is the direction in which $\|y\|$ is smallest for all $\|x\| = 1$. From the input/output point of view, v_1 (v_n) is the *highest (lowest) gain input direction*, while u_1 (u_m) is the *highest (lowest) gain observing direction*. This can be illustrated by the following 2×2 matrix:

$$A = \begin{bmatrix} \cos\theta_1 & -\sin\theta_1 \\ \sin\theta_1 & \cos\theta_1 \end{bmatrix} \begin{bmatrix} \sigma_1 & \\ & \sigma_2 \end{bmatrix} \begin{bmatrix} \cos\theta_2 & -\sin\theta_2 \\ \sin\theta_2 & \cos\theta_2 \end{bmatrix}.$$

It is easy to see that A maps a unit disk to an ellipsoid with semi-axes of σ_1 and σ_2.

Hence it is often convenient to introduce the following alternative definitions for the largest singular value $\overline{\sigma}$:

$$\overline{\sigma}(A) := \max_{\|x\|=1} \|Ax\|$$

and for the smallest singular value $\underline{\sigma}$ of a *tall matrix*:

$$\underline{\sigma}(A) := \min_{\|x\|=1} \|Ax\| .$$

Lemma 2.12 *Suppose A and Δ are square matrices. Then*

(i) $|\underline{\sigma}(A+\Delta) - \underline{\sigma}(A)| \leq \overline{\sigma}(\Delta)$;

(ii) $\underline{\sigma}(A\Delta) \geq \underline{\sigma}(A)\underline{\sigma}(\Delta)$;

(iii) $\overline{\sigma}(A^{-1}) = \dfrac{1}{\underline{\sigma}(A)}$ *if A is invertible.*

Proof.

(i) By definition

$$\begin{aligned} \underline{\sigma}(A+\Delta) &:= \min_{\|x\|=1} \|(A+\Delta)x\| \\ &\geq \min_{\|x\|=1} \{\|Ax\| - \|\Delta x\|\} \\ &\geq \min_{\|x\|=1} \|Ax\| - \max_{\|x\|=1} \|\Delta x\| \\ &= \underline{\sigma}(A) - \overline{\sigma}(\Delta). \end{aligned}$$

Hence $-\overline{\sigma}(\Delta) \leq \underline{\sigma}(A+\Delta) - \underline{\sigma}(A)$. The other inequality $\underline{\sigma}(A+\Delta) - \underline{\sigma}(A) \leq \overline{\sigma}(\Delta)$ follows by replacing A by $A+\Delta$ and Δ by $-\Delta$ in the above proof.

(ii) This follows by noting that

$$\begin{aligned} \underline{\sigma}(A\Delta) &:= \min_{\|x\|=1} \|A\Delta x\| \\ &= \sqrt{\min_{\|x\|=1} x^*\Delta^* A^* A \Delta x} \\ &\geq \underline{\sigma}(A) \min_{\|x\|=1} \|\Delta x\| \\ &= \underline{\sigma}(A)\underline{\sigma}(\Delta). \end{aligned}$$

(iii) Let the singular value decomposition of A be $A = U\Sigma V^*$, then $A^{-1} = V\Sigma^{-1}U^*$. Hence $\overline{\sigma}(A^{-1}) = \overline{\sigma}(\Sigma^{-1}) = 1/\underline{\sigma}(\Sigma) = 1/\underline{\sigma}(A)$.

□

Some useful properties of SVD are collected in the following lemma.

Lemma 2.13 *Let* $A \in \mathbb{F}^{m\times n}$ *and*

$$\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > \sigma_{r+1} = \cdots = 0, \;\; r \leq min\{m, n\}.$$

Then

1. $rank(A) = r$;
2. $KerA = span\{v_{r+1}, \ldots, v_n\}$ *and* $(KerA)^\perp = span\{v_1, \ldots, v_r\}$;
3. $ImA = span\{u_1, \ldots, u_r\}$ *and* $(ImA)^\perp = span\{u_{r+1}, \ldots, u_m\}$;
4. $A \in \mathbb{F}^{m\times n}$ *has a dyadic expansion:*
$$A = \sum_{i=1}^{r} \sigma_i u_i v_i^* = U_r \Sigma_r V_r^*$$
where $U_r = [u_1, \ldots, u_r]$, $V_r = [v_1, \ldots, v_r]$, *and* $\Sigma_r = diag(\sigma_1, \ldots, \sigma_r)$;
5. $\|A\|_F^2 = \sigma_1^2 + \sigma_2^2 + \cdots + \sigma_r^2$;
6. $\|A\| = \sigma_1$;
7. $\sigma_i(U_0 A V_0) = \sigma_i(A)$, $i = 1, \ldots, p$ *for any appropriately dimensioned unitary matrices* U_0 *and* V_0;
8. *Let* $k < r = rank(A)$ *and* $A_k := \sum_{i=1}^{k} \sigma_i u_i v_i^*$, *then*
$$\min_{rank(B)\leq k} \|A - B\| = \|A - A_k\| = \sigma_{k+1}.$$

Proof. We shall only give a proof for part 8. It is easy to see that $\text{rank}(A_k) \leq k$ and $\|A - A_k\| = \sigma_{k+1}$. Hence, we only need to show that $\min\limits_{\text{rank}(B)\leq k} \|A - B\| \geq \sigma_{k+1}$. Let B be any matrix such that $\text{rank}(B) \leq k$. Then

$$\begin{aligned} \|A - B\| &= \|U\Sigma V^* - B\| = \|\Sigma - U^*BV\| \\ &\geq \left\| \begin{bmatrix} I_{k+1} & 0 \end{bmatrix} (\Sigma - U^*BV) \begin{bmatrix} I_{k+1} \\ 0 \end{bmatrix} \right\| = \left\| \Sigma_{k+1} - \hat{B} \right\| \end{aligned}$$

where $\hat{B} = \begin{bmatrix} I_{k+1} & 0 \end{bmatrix} U^*BV \begin{bmatrix} I_{k+1} \\ 0 \end{bmatrix} \in \mathbb{F}^{(k+1)\times(k+1)}$ and $\text{rank}(\hat{B}) \leq k$. Let $x \in \mathbb{F}^{k+1}$ be such that $\hat{B}x = 0$ and $\|x\| = 1$. Then

$$\|A - B\| \geq \left\| \Sigma_{k+1} - \hat{B} \right\| \geq \left\| (\Sigma_{k+1} - \hat{B})x \right\| = \|\Sigma_{k+1}x\| \geq \sigma_{k+1}.$$

Since B is arbitrary, the conclusion follows. □

2.9 Generalized Inverses

Let $A \in \mathbb{C}^{m\times n}$. A matrix $X \in \mathbb{C}^{n\times m}$ is said to be a *right inverse* of A if $AX = I$. Obviously, A has a right inverse iff A has full row rank, and, in that case, one of the right inverses is given by $X = A^*(AA^*)^{-1}$. Similarly, if $YA = I$ then Y is called a *left inverse* of A. By duality, A has a left inverse iff A has full column rank, and, furthermore, one of the left inverses is $Y = (A^*A)^{-1}A^*$. Note that right (or left) inverses are not necessarily unique. For example, any matrix in the form $\begin{bmatrix} I \\ \star \end{bmatrix}$ is a right inverse of $\begin{bmatrix} I & 0 \end{bmatrix}$.

More generally, if a matrix A has neither full row rank nor full column rank, then all the ordinary matrix inverses do not exist; however, the so called *pseudo-inverse*, known also as the *Moore-Penrose inverse*, is useful. This pseudo-inverse is denoted by A^+, which satisfies the following conditions:

(i) $AA^+A = A$;

(ii) $A^+AA^+ = A^+$;

(iii) $(AA^+)^* = AA^+$;

(iv) $(A^+A)^* = A^+A$.

It can be shown that pseudo-inverse is unique. One way of computing A^+ is by writing

$$A = BC$$

so that B has full column rank and C has full row rank. Then

$$A^+ = C^*(CC^*)^{-1}(B^*B)^{-1}B^*.$$

Another way to compute A^+ is by using SVD. Suppose A has a singular value decomposition

$$A = U\Sigma V^*$$

with

$$\Sigma = \begin{bmatrix} \Sigma_r & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma_r > 0.$$

Then $A^+ = V\Sigma^+U^*$ with

$$\Sigma^+ = \begin{bmatrix} \Sigma_r^{-1} & 0 \\ 0 & 0 \end{bmatrix}.$$

2.10 Semidefinite Matrices

A square hermitian matrix $A = A^*$ is said to be *positive definite (semi-definite)* , denoted by $A > 0$ (≥ 0), if $x^*Ax > 0$ (≥ 0) for all $x \neq 0$. Suppose $A \in \mathbb{F}^{n\times n}$ and $A = A^* \geq 0$, then there exists a $B \in \mathbb{F}^{n\times r}$ with $r \geq \text{rank}(A)$ such that $A = BB^*$.

Lemma 2.14 *Let $B \in \mathbb{F}^{m\times n}$ and $C \in \mathbb{F}^{k\times n}$. Suppose $m \geq k$ and $B^*B = C^*C$. Then there exists a matrix $U \in \mathbb{F}^{m\times k}$ such that $U^*U = I$ and $B = UC$.*

Proof. Let V_1 and V_2 be unitary matrices such that

$$B_1 = V_1 \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \quad C_1 = V_2 \begin{bmatrix} C_1 \\ 0 \end{bmatrix}$$

where B_1 and C_1 are full row rank. Then B_1 and C_1 have the same number of rows and $V_3 := B_1 C_1^*(C_1 C_1^*)^{-1}$ satisfies $V_3^* V_3 = I$ since $B^*B = C^*C$. Hence V_3 is a unitary matrix and $V_3^* B_1 = C_1$. Finally let

$$U = V_1 \begin{bmatrix} V_3 & 0 \\ 0 & V_4 \end{bmatrix} V_2^*$$

for any suitably dimensioned V_4 such that $V_4^* V_4 = I$. □

We can define square root for a positive semi-definite matrix A, $A^{1/2} = (A^{1/2})^* \geq 0$, by

$$A = A^{1/2} A^{1/2}.$$

Clearly, $A^{1/2}$ can be computed by using spectral decomposition or SVD: let $A = U\Lambda U^*$, then

$$A^{1/2} = U \Lambda^{1/2} U^*$$

where

$$\Lambda = \operatorname{diag}\{\lambda_1, \ldots, \lambda_n\}, \quad \Lambda^{1/2} = \operatorname{diag}\{\sqrt{\lambda_1}, \ldots, \sqrt{\lambda_n}\}.$$

Lemma 2.15 *Suppose $A = A^* > 0$ and $B = B^* \geq 0$. Then $A > B$ iff $\rho(BA^{-1}) < 1$.*

Proof. Since $A > 0$, we have $A > B$ iff

$$0 < I - A^{-1/2} B A^{-1/2} = I - A^{-1/2}(BA^{-1})A^{1/2}.$$

However, $A^{-1/2} B A^{-1/2}$ and BA^{-1} are similar, and hence $\lambda_i(BA^{-1}) = \lambda_i(A^{-1/2} B A^{-1/2})$. Therefore, the conclusion follows by the fact that

$$0 < I - A^{-1/2} B A^{-1/2}$$

iff $\rho(A^{-1/2} B A^{-1/2}) < 1$ iff $\rho(BA^{-1}) < 1$. □

Lemma 2.16 *Let $X = X^* \geq 0$ be partitioned as*

$$X = \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^* & X_{22} \end{bmatrix}.$$

Then $\operatorname{Ker} X_{22} \subset \operatorname{Ker} X_{12}$. *Consequently, if X_{22}^+ is the pseudo-inverse of X_{22}, then $Y = X_{12} X_{22}^+$ solves*

$$Y X_{22} = X_{12}$$

and

$$\begin{bmatrix} X_{11} & X_{12} \\ X_{12}^* & X_{22} \end{bmatrix} = \begin{bmatrix} I & X_{12} X_{22}^+ \\ 0 & I \end{bmatrix} \begin{bmatrix} X_{11} - X_{12} X_{22}^+ X_{12}^* & 0 \\ 0 & X_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ X_{22}^+ X_{12}^* & I \end{bmatrix}.$$

Proof. Without loss of generality, assume

$$X_{22} = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix}$$

with $\Sigma_1 = \Sigma_1^* > 0$ and $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix}$ unitary. Then

$$\text{Ker } X_{22} = \text{span}\{\text{columns of } U_2\}$$

and

$$X_{22}^+ = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Sigma_1^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_1^* \\ U_2^* \end{bmatrix}.$$

Moreover

$$\begin{bmatrix} I & 0 \\ 0 & U^* \end{bmatrix} \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^* & X_{22} \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & U \end{bmatrix} \geq 0$$

gives $X_{12}U_2 = 0$. Hence, Ker $X_{22} \subset$ Ker X_{12} and now

$$X_{12}X_{22}^+X_{22} = X_{12}U_1U_1^* = X_{12}U_1U_1^* + X_{12}U_2U_2^* = X_{12}.$$

The factorization follows easily. □

2.11 Matrix Dilation Problems*

In this section, we consider the following induced 2-norm optimization problem:

$$\min_X \left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\| \tag{2.3}$$

where X, B, C, and A are constant matrices of compatible dimensions.

The matrix $\begin{bmatrix} X & B \\ C & A \end{bmatrix}$ is a *dilation* of its sub-matrices as indicated in the following diagram:

$$\begin{tikzcd}
\begin{bmatrix} X & B \\ C & A \end{bmatrix} \arrow[r, shift right, "c"'] \arrow[d, shift left, "c"] & \begin{bmatrix} B \\ A \end{bmatrix} \arrow[l, shift right, "d"'] \arrow[d, shift left, "c"] \\
\begin{bmatrix} C & A \end{bmatrix} \arrow[u, shift left, "d"] \arrow[r, shift right, "c"'] & \begin{bmatrix} A \end{bmatrix} \arrow[l, shift right, "d"'] \arrow[u, shift left, "d"]
\end{tikzcd}$$

In this diagram, "c" stands for the operation of *compression* and "d" stands for *dilation*. Compression is always norm non-increasing and dilation is always norm non-decreasing. Sometimes dilation can be made to be norm preserving. Norm preserving dilations are the focus of this section.

The simplest matrix dilation problem occurs when solving

$$\min_X \left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\|. \tag{2.4}$$

Although (2.4) is a much simplified version of (2.3), we will see that it contains all the essential features of the general problem. Letting γ_0 denote the minimum norm in (2.4), it is immediate that

$$\gamma_0 = \|A\|.$$

The following theorem characterizes all solutions to (2.4).

Theorem 2.17 $\forall \gamma \geq \gamma_0$,

$$\left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\| \leq \gamma$$

iff there is a Y *with* $\|Y\| \leq 1$ *such that*

$$X = Y(\gamma^2 I - A^*A)^{1/2}.$$

Proof.

$$\left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\| \leq \gamma$$

iff

$$X^*X + A^*A \leq \gamma^2 I$$

iff

$$X^*X \leq (\gamma^2 I - A^*A).$$

Now suppose $X^*X \leq (\gamma^2 I - A^*A)$ and let

$$Y := X\left[(\gamma^2 I - A^*A)^{1/2}\right]^+$$

then $X = Y(\gamma^2 I - A^*A)^{1/2}$ and $Y^*Y \leq I$. Similarly if $X = Y(\gamma^2 I - A^*A)^{1/2}$ and $Y^*Y \leq I$ then $X^*X \leq (\gamma^2 I - A^*A)$. □

This theorem implies that, in general, (2.4) has more than one solution, which is in contrast to the minimization in the Frobenius norm in which $X = 0$ is the unique solution. The solution $X = 0$ is the central solution but others are possible unless $A^*A = \gamma_0^2 I$.

Remark 2.1 The theorem still holds if $(\gamma^2 I - A^*A)^{1/2}$ is replaced by any matrix R such that $\gamma^2 I - A^*A = R^*R$. ♡

A more restricted version of the above theorem is shown in the following corollary.

Corollary 2.18 $\forall \gamma > \gamma_0$,

$$\left\| \begin{bmatrix} X \\ A \end{bmatrix} \right\| \leq \gamma \ (< \gamma)$$

iff

$$\left\| X(\gamma^2 I - A^*A)^{-1/2} \right\| \leq 1 (< 1).$$

The corresponding dual results are

Theorem 2.19 $\forall \gamma \geq \gamma_0$

$$\left\| \begin{bmatrix} X & A \end{bmatrix} \right\| \leq \gamma$$

iff there is a Y, $\|Y\| \leq 1$, *such that*

$$X = (\gamma^2 I - AA^*)^{1/2} Y.$$

Corollary 2.20 $\forall \gamma > \gamma_0$

$$\left\| \begin{bmatrix} X & A \end{bmatrix} \right\| \leq \gamma \ (< \gamma)$$

iff

$$\left\| (\gamma^2 I - AA^*)^{-1/2} X \right\| \leq 1 \ (< 1).$$

Now, returning to the problem in (2.3), let

$$\gamma_0 := \min_X \left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\|. \tag{2.5}$$

The following so called Parrott's theorem will play an important role in many control related optimization problems. The proof is the straightforward application of Theorem 2.17 and its dual, Theorem 2.19.

Theorem 2.21 (Parrott's Theorem) *The minimum in (2.5) is given by*

$$\gamma_0 = \max \left\{ \left\| \begin{bmatrix} C & A \end{bmatrix} \right\|, \left\| \begin{bmatrix} B \\ A \end{bmatrix} \right\| \right\}. \tag{2.6}$$

Proof. Denote by $\hat{\gamma}$ the right hand side of the equation (2.6). Clearly, $\gamma_0 \geq \hat{\gamma}$ since compressions are norm non-increasing, and that $\gamma_0 \leq \hat{\gamma}$ will be shown by using Theorem 2.17 and Theorem 2.19.

Suppose $A \in \mathbb{C}^{n\times m}$ and $n \geq m$ (the case for $m > n$ can be shown in the same fashion). Then A has the following singular value decomposition:

$$A = U \begin{bmatrix} \Sigma_m \\ 0_{n-m,m} \end{bmatrix} V^*, \; U \in \mathbb{C}^{n\times n}, V \in \mathbb{C}^{m\times m}.$$

Hence

$$\hat{\gamma}^2 I - A^*A = V(\hat{\gamma}^2 I - \Sigma_m^2)V^*$$

and

$$\hat{\gamma}^2 I - AA^* = U \begin{bmatrix} \hat{\gamma}^2 I - \Sigma_m^2 & 0 \\ 0 & \hat{\gamma}^2 I_{n-m} \end{bmatrix} U^*.$$

Now let

$$(\hat{\gamma}^2 I - A^*A)^{1/2} := V(\hat{\gamma}^2 I - \Sigma_m^2)^{1/2}V^*$$

and

$$(\hat{\gamma}^2 I - AA^*)^{1/2} := U \begin{bmatrix} (\hat{\gamma}^2 I - \Sigma_m^2)^{1/2} & 0 \\ 0 & \hat{\gamma} I_{n-m} \end{bmatrix} U^*.$$

Then it is easy to verify that

$$(\hat{\gamma}^2 I - A^*A)^{1/2} A^* = A^*(\hat{\gamma}^2 I - AA^*)^{1/2}.$$

Using this equality, we can show that

$$\begin{bmatrix} -A^* & (\hat{\gamma}^2 I - A^*A)^{1/2} \\ (\hat{\gamma}^2 I - AA^*)^{1/2} & A \end{bmatrix} \begin{bmatrix} -A^* & (\hat{\gamma}^2 I - A^*A)^{1/2} \\ (\hat{\gamma}^2 I - AA^*)^{1/2} & A \end{bmatrix}^*$$
$$= \begin{bmatrix} \hat{\gamma}^2 I & 0 \\ 0 & \hat{\gamma}^2 I \end{bmatrix}.$$

Now we are ready to show that $\gamma_0 \leq \hat{\gamma}$.

From Theorem 2.17 we have that $B = Y(\hat{\gamma}^2 I - A^*A)^{1/2}$ for some Y such that $\|Y\| \leq 1$. Similarly, Theorem 2.19 yields $C = (\hat{\gamma}^2 I - AA^*)^{1/2} Z$ for some Z with $\|Z\| \leq 1$. Now let $\hat{X} = -YA^*Z$. Then

$$\begin{aligned}
\left\| \begin{bmatrix} \hat{X} & B \\ C & A \end{bmatrix} \right\| &= \left\| \begin{bmatrix} -YA^*Z & Y(\hat{\gamma}^2 I - A^*A)^{1/2} \\ (\hat{\gamma}^2 I - AA^*)^{1/2} Z & A \end{bmatrix} \right\| \\
&= \left\| \begin{bmatrix} Y & \\ & I \end{bmatrix} \begin{bmatrix} -A^* & (\hat{\gamma}^2 I - A^*A)^{1/2} \\ (\hat{\gamma}^2 I - AA^*)^{1/2} & A \end{bmatrix} \begin{bmatrix} Z & \\ & I \end{bmatrix} \right\| \\
&\leq \left\| \begin{bmatrix} -A^* & (\hat{\gamma}^2 I - A^*A)^{1/2} \\ (\hat{\gamma}^2 I - AA^*)^{1/2} & A \end{bmatrix} \right\| \\
&= \hat{\gamma}.
\end{aligned}$$

Thus $\hat{\gamma} \geq \gamma_0$, so $\hat{\gamma} = \gamma_0$. □

This theorem gives one solution to (2.3) and an expression for γ_0. As in (2.4), there may be more than one solution to (2.3), although the proof of theorem 2.21 exhibits only one. Theorem 2.22 considers the problem of parameterizing all solutions. The solution $\hat{X} = -YA^*Z$ is the "central" solution analogous to $X = 0$ in (2.4).

Theorem 2.22 *Suppose $\gamma \geq \gamma_0$. The solutions X such that*

$$\left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\| \leq \gamma \tag{2.7}$$

are exactly those of the form

$$X = -YA^*Z + \gamma(I - YY^*)^{1/2} W (I - Z^*Z)^{1/2} \tag{2.8}$$

where W is an arbitrary contraction ($\|W\| \leq 1$), and Y with $\|Y\| \leq 1$ and Z with $\|Z\| \leq 1$ solve the linear equations

$$\begin{aligned} B &= Y(\gamma^2 I - A^*A)^{1/2}, & (2.9)\\ C &= (\gamma^2 I - AA^*)^{1/2} Z. & (2.10) \end{aligned}$$

Proof. Since $\gamma \geq \gamma_0$, again from Theorem 2.19 there exists a Z with $\|Z\| \leq 1$ such that

$$C = (\gamma^2 I - AA^*)^{1/2} Z.$$

Note that using the above expression for C we have

$$\gamma^2 I - \begin{bmatrix} C & A \end{bmatrix}^* \begin{bmatrix} C & A \end{bmatrix}$$

$$= \begin{bmatrix} \gamma(I - Z^*Z)^{1/2} & 0 \\ -A^*Z & (\gamma^2 I - A^*A)^{1/2} \end{bmatrix}^* \begin{bmatrix} \gamma(I - Z^*Z)^{1/2} & 0 \\ -A^*Z & (\gamma^2 I - A^*A)^{1/2} \end{bmatrix}.$$

Now apply Theorem 2.17 (Remark 2.1) to inequality (2.7) with respect to the partitioned matrix $\left[\begin{array}{cc} X & B \\ \hdashline C & A \end{array}\right]$ to get

$$\begin{bmatrix} X & B \end{bmatrix} = \hat{W} \begin{bmatrix} \gamma(I - Z^*Z)^{1/2} & 0 \\ -A^*Z & (\gamma^2 I - A^*A)^{1/2} \end{bmatrix}$$

for some contraction $\hat{W}$, $\left\|\hat{W}\right\| \leq 1$. Partition $\hat{W}$ as $\hat{W} = \begin{bmatrix} W_1 & Y \end{bmatrix}$ to obtain the expression for X and B:

$$\begin{aligned} X &= -YA^*Z + \gamma W_1 (I - Z^*Z)^{1/2}, \\ B &= Y(\gamma^2 I - A^*A)^{1/2}. \end{aligned}$$

Then $\|Y\| \leq 1$ and the theorem follows by noting that $\left\| \begin{bmatrix} W_1 & Y \end{bmatrix} \right\| \leq 1$ iff there is a W, $\|W\| \leq 1$, such that $W_1 = (I - YY^*)^{1/2} W$. □

The following corollary gives an alternative version of Theorem 2.22 when $\gamma > \gamma_0$.

Corollary 2.23 *For* $\gamma > \gamma_0$.

$$\left\| \begin{bmatrix} X & B \\ C & A \end{bmatrix} \right\| \leq \gamma \ (< \gamma) \tag{2.11}$$

iff

$$\left\| (I - YY^*)^{-1/2}(X + YA^*Z)(I - Z^*Z)^{-1/2} \right\| \leq \gamma \ (< \gamma) \tag{2.12}$$

where

$$Y = B(\gamma^2 I - A^*A)^{-1/2}, \tag{2.13}$$
$$Z = (\gamma^2 I - AA^*)^{-1/2}C. \tag{2.14}$$

Note that in the case of $\gamma > \gamma_0$, $I - YY^*$ and $I - Z^*Z$ are invertible since Corollary 2.18 and 2.20 clearly show that $\|Y\| < 1$ and $\|Z\| < 1$. There are many alternative characterizations of solutions to (2.11), although the formulas given above seem to be the simplest.

As a straightforward application of the dilation results obtained above, consider the following matrix approximation problem:

$$\gamma_0 = \min_Q \|R + UQV\| \tag{2.15}$$

where R, U, and V are constant matrices such that $U^*U = I$ and $VV^* = I$.

Corollary 2.24 *The minimum achievable norm is given by*

$$\gamma_0 = \max\left\{ \|U_\perp^* R\|,\ \|RV_\perp^*\| \right\},$$

and the parameterization for all optimal solutions Q can be obtained from Theorem 2.22 with $X = Q + U^*RV^*$, $A = U_\perp^* RV_\perp^*$, $B = U^*RV_\perp^*$, and $C = U_\perp^* RV^*$.

Proof. Let $U_\perp$ and $V_\perp$ be such that $\begin{bmatrix} U & U_\perp \end{bmatrix}$ and $\begin{bmatrix} V \\ V_\perp \end{bmatrix}$ are unitary matrices. Then

$$\begin{aligned} \gamma_0 &= \min_Q \|R + UQV\| \\ &= \min_Q \left\| \begin{bmatrix} U & U_\perp \end{bmatrix}^* (R + UQV) \begin{bmatrix} V \\ V_\perp \end{bmatrix}^* \right\| \\ &= \min_Q \left\| \begin{bmatrix} U^*RV^* + Q & U^*RV_\perp^* \\ U_\perp^*RV^* & U_\perp^*RV_\perp^* \end{bmatrix} \right\|. \end{aligned}$$

The result follows after applying Theorem 2.21 and 2.22. $\square$

A similar problem arises in $\mathcal{H}_\infty$ control theory.

2.12 Notes and References

A very extensive treatment of most topics in this chapter can be found in Brogan [1991], Horn and Johnson [1990,1991] and Lancaster and Tismenetsky [1985]. Golub and Van Loan's book [1983] contains many numerical algorithms for solving most of the problems in this chapter. The matrix dilation theory can be found in Davis, Kahan, and Weinberger [1982].

3

Linear Dynamical Systems

This chapter reviews some basic system theoretical concepts. The notions of controllability, observability, stabilizability, and detectability are defined and various algebraic and geometric characterizations of these notions are summarized. Kalman canonical decomposition, pole placement, and observer theory are then introduced. The solutions of Lyapunov equations and their connections with system stability, controllability, and so on, are discussed. System interconnections and realizations, in particular the balanced realization, are studied in some detail. Finally, the concepts of system poles and zeros are introduced.

3.1 Descriptions of Linear Dynamical Systems

Let a finite dimensional linear time invariant (FDLTI) dynamical system be described by the following linear constant coefficient differential equations:

$$\dot{x} = Ax + Bu, \quad x(t_0) = x_0 \tag{3.1}$$

$$y = Cx + Du \tag{3.2}$$

where $x(t) \in \mathbb{R}^n$ is called the system *state*, $x(t_0)$ is called the *initial condition* of the system, $u(t) \in \mathbb{R}^m$ is called the system *input*, and $y(t) \in \mathbb{R}^p$ is the system *output*. The A, B, C, and D are appropriately dimensioned real constant matrices. A dynamical system with single input ($m = 1$) and single output ($p = 1$) is called a SISO (single input and single output) system, otherwise it is called MIMO (multiple input and multiple output) system. The corresponding transfer matrix from u to y is defined as

$$Y(s) = G(s)U(s)$$

where $U(s)$ and $Y(s)$ are the Laplace transform of $u(t)$ and $y(t)$ with zero initial condition ($x(0) = 0$). Hence, we have

$$G(s) = C(sI - A)^{-1}B + D.$$

Note that the system equations (3.1) and (3.2) can be written in a more compact matrix form:

$$\begin{bmatrix} \dot{x} \\ y \end{bmatrix} = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix}.$$

To expedite calculations involving transfer matrices, the notation

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := C(sI - A)^{-1}B + D$$

will be used. Other reasons for using this notation will be discussed in Chapter 10. Note that

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}$$

is a real block matrix, not a transfer function.

Now given the initial condition $x(t_0)$ and the input $u(t)$, the dynamical system response $x(t)$ and $y(t)$ for $t \geq t_0$ can be determined from the following formulas:

$$x(t) = e^{A(t-t_0)}x(t_0) + \int_{t_0}^{t} e^{A(t-\tau)}Bu(\tau)d\tau \tag{3.3}$$

$$y(t) = Cx(t) + Du(t). \tag{3.4}$$

In the case of $u(t) = 0$, $\forall t \geq t_0$, it is easy to see from the solution that for any $t_1 \geq t_0$ and $t \geq t_0$, we have

$$x(t) = e^{A(t-t_1)}x(t_1).$$

Therefore, the matrix function $\Phi(t, t_1) = e^{A(t-t_1)}$ acts as a transformation from one state to another, and thus $\Phi(t, t_1)$ is usually called the *state transition matrix*. Since the state of a linear system at one time can be obtained from the state at another through the transition matrix, we can assume without loss of generality that $t_0 = 0$. This will be assumed in the sequel.

The impulse matrix of the dynamical system is defined as

$$g(t) = \mathcal{L}^{-1}\left\{G(s)\right\} = Ce^{At}B1_+(t) + D\delta(t)$$

where $\delta(t)$ is the unit impulse and $1_+(t)$ is the unit step defined as

$$1_+(t) := \begin{cases} 1, & t \geq 0; \\ 0, & t < 0. \end{cases}$$

The input/output relationship (i.e., with zero initial state: $x_0 = 0$) can be described by the convolution equation

$$y(t) = (g * u)(t) := \int_{-\infty}^{\infty} g(t-\tau)u(\tau)d\tau = \int_{-\infty}^{t} g(t-\tau)u(\tau)d\tau.$$

3.2 Controllability and Observability

We now turn to some very important concepts in linear system theory.

Definition 3.1 The dynamical system described by the equation (3.1) or the pair (A, B) is said to be *controllable* if, for any initial state $x(0) = x_0$, $t_1 > 0$ and final state x_1, there exists a (piecewise continuous) input $u(\cdot)$ such that the solution of (3.1) satisfies $x(t_1) = x_1$. Otherwise, the system or the pair (A, B) is said to be *uncontrollable*.

The controllability (or observability introduced next) of a system can be verified through some algebraic or geometric criteria.

Theorem 3.1 *The following are equivalent:*

(i) (A, B) *is controllable.*

(ii) The matrix

$$W_c(t) := \int_0^t e^{A\tau}BB^*e^{A^*\tau}d\tau$$

is positive definite for any $t > 0$.

(iii) The controllability matrix

$$\mathcal{C} = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix}$$

has full row rank or, in other words, $\langle A \,|ImB\rangle := \sum_{i=1}^n Im(A^{i-1}B) = \mathbb{R}^n$.

(iv) The matrix $[A - \lambda I,\ B]$ *has full row rank for all* λ *in* $\mathbb{C}$.

(v) Let λ *and* x *be any eigenvalue and any corresponding left eigenvector of* A*, i.e.,* $x^*A = x^*\lambda$*, then* $x^*B \neq 0$.

(vi) The eigenvalues of $A+BF$ *can be freely assigned (with the restriction that complex eigenvalues are in conjugate pairs) by a suitable choice of* F.

Proof.

$(i) \Leftrightarrow (ii)$: Suppose $W_c(t_1) > 0$ for some $t_1 > 0$, and let the input be defined as

$$u(\tau) = -B^* e^{A^*(t_1-\tau)} W_c(t_1)^{-1}(e^{At_1}x_0 - x_1).$$

Then it is easy to verify using the formula in (3.3) that $x(t_1) = x_1$. Since x_1 is arbitrary, the pair (A, B) is controllable.

To show that the controllability of (A, B) implies that $W_c(t) > 0$ for any $t > 0$, assume that (A, B) is controllable but $W_c(t_1)$ is singular for some $t_1 > 0$. Since $e^{At}BB^*e^{A^*t} \geq 0$ for all t, there exists a real vector $0 \neq v \in \mathbb{R}^n$ such that

$$v^* e^{At} B = 0, \quad t \in [0, t_1].$$

Now let $x(t_1) = x_1 = 0$, and then from the solution (3.3), we have

$$0 = e^{At_1}x(0) + \int_0^{t_1} e^{A(t_1-\tau)}Bu(\tau)d\tau.$$

Pre-multiply the above equation by v^* to get

$$0 = v^* e^{At_1} x(0).$$

If we chose the initial state $x(0) = e^{-At_1}v$, then $v = 0$, and this is a contradiction. Hence, $W_c(t)$ can not be singular for any $t > 0$.

$(ii) \Leftrightarrow (iii)$: Suppose $W_c(t) > 0$ for all $t > 0$ (in fact, it can be shown that $W_c(t) > 0$ for all $t > 0$ iff, for some t_1, $W_c(t_1) > 0$) but the controllability matrix $\mathcal{C}$ does not have full row rank. Then there exists a $v \in \mathbb{R}^n$ such that

$$v^* A^i B = 0$$

for all $0 \leq i \leq n-1$. In fact, this equality holds for all $i \geq 0$ by the Cayley-Hamilton Theorem. Hence,

$$v^* e^{At} B = 0$$

for all t or, equivalently, $v^* W_c(t) = 0$ for all t; this is a contradiction, and hence, the controllability matrix $\mathcal{C}$ must be full row rank. Conversely, suppose $\mathcal{C}$ has full row rank but $W_c(t)$ is singular for some t_1. Then there exists a $0 \neq v \in \mathbb{R}^n$ such that $v^* e^{At} B = 0$ for all $t \in [0, t_1]$. Therefore, set $t = 0$, and we have

$$v^* B = 0.$$

Next, evaluate the i-th derivative of $v^* e^{At} B = 0$ at $t = 0$ to get

$$v^* A^i B = 0, \quad i > 0.$$

Hence, we have

$$v^* \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix} = 0$$

or, in other words, the controllability matrix $\mathcal{C}$ does not have full row rank. This is again a contradiction.

$(iii) \Rightarrow (iv)$: Suppose, on the contrary, that the matrix

$$\begin{bmatrix} A-\lambda I & B \end{bmatrix}$$

does not have full row rank for some $\lambda \in \mathbb{C}$. Then there exists a vector $x \in \mathbb{C}^n$ such that

$$x^* \begin{bmatrix} A-\lambda I & B \end{bmatrix} = 0$$

i.e., $x^*A = \lambda x^*$ and $x^*B = 0$. However, this will result in

$$x^* \begin{bmatrix} B & AB & \dots & A^{n-1}B \end{bmatrix} = \begin{bmatrix} x^*B & \lambda x^*B & \dots & \lambda^{n-1}x^*B \end{bmatrix} = 0$$

i.e., the controllability matrix $\mathcal{C}$ does not have full row rank, and this is a contradiction.

$(iv) \Rightarrow (v)$: This is obvious from the proof of $(iii) \Rightarrow (iv)$.

$(v) \Rightarrow (iii)$: We will again prove this by contradiction. Assume that (v) holds but rank $\mathcal{C} = k < n$. Then in section 3.3, we will show that there is a transformation T such that

$$TAT^{-1} = \begin{bmatrix} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{\bar{c}} \end{bmatrix} \quad TB = \begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix}$$

with $\bar{A}_{\bar{c}} \in \mathbb{R}^{(n-k)\times(n-k)}$. Let λ_1 and $x_{\bar{c}}$ be any eigenvalue and any corresponding left eigenvector of $\bar{A}_{\bar{c}}$, i.e., $x_{\bar{c}}^*\bar{A}_{\bar{c}} = \lambda_1 x_{\bar{c}}^*$. Then $x^*(TB) = 0$ and

$$x = \begin{bmatrix} 0 \\ x_{\bar{c}} \end{bmatrix}$$

is an eigenvector of TAT^{-1} corresponding to the eigenvalue λ_1, which implies that (TAT^{-1}, TB) is not controllable. This is a contradiction since similarity transformation does not change controllability. Hence, the proof is completed.

$(vi) \Rightarrow (i)$: This follows the same arguments as in the proof of $(v) \Rightarrow (iii)$: assume that (vi) holds but (A, B) is uncontrollable. Then, there is a decomposition so that some subsystems are not affected by the control, but this contradicts the condition (vi).

$(i) \Rightarrow (vi)$: This will be clear in section 3.4. In that section, we will explicitly construct a matrix F so that the eigenvalues of $A + BF$ are in the desired locations.

□

Definition 3.2 An unforced dynamical system $\dot{x} = Ax$ is said to be *stable* if all the eigenvalues of A are in the open left half plane, i.e., $\mathrm{Re}\lambda(A) < 0$. A matrix A with such a property is said to be stable or Hurwitz.

Definition 3.3 The dynamical system (3.1), or the pair $(A,\, B)$, is said to be *stabilizable* if there exists a state feedback $u = Fx$ such that the system is stable, i.e., $A + BF$ is stable.

Therefore, it is more appropriate to call this stabilizability the *state feedback stabilizability* to differentiate it from the *output feedback stabilizability* defined later.

The following theorem is a consequence of Theorem 3.1.

Theorem 3.2 *The following are equivalent:*

(i) $(A,\, B)$ *is stabilizable.*

(ii) The matrix $[A - \lambda I,\, B]$ *has full row rank for all* $Re\lambda \geq 0$.

(iii) For all λ *and* x *such that* $x^*A = x^*\lambda$ *and* $Re\lambda \geq 0$, $x^*B \neq 0$.

(iv) There exists a matrix F *such that* $A + BF$ *is Hurwitz.*

We now consider the dual notions of observability and detectability of the system described by equations (3.1) and (3.2).

Definition 3.4 The dynamical system described by the equations (3.1) and (3.2) or by the pair (C, A) is said to be *observable* if, for any $t_1 > 0$, the initial state $x(0) = x_0$ can be determined from the time history of the input $u(t)$ and the output $y(t)$ in the interval of $[0, t_1]$. Otherwise, the system, or (C, A), is said to be *unobservable.*

Theorem 3.3 *The following are equivalent:*

(i) (C, A) *is observable.*

(ii) The matrix

$$W_o(t) := \int_0^t e^{A^*\tau} C^* C e^{A\tau} d\tau$$

is positive definite for any $t > 0$.

(iii) The observability matrix

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

has full column rank or $\bigcap_{i=1}^{n} Ker(CA^{i-1}) = 0$.

(iv) The matrix $\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}$ *has full column rank for all* λ *in* $\mathbb{C}$.

(v) *Let λ and y be any eigenvalue and any corresponding right eigenvector of A, i.e., $Ay = \lambda y$, then $Cy \neq 0$.*

(vi) *The eigenvalues of $A+LC$ can be freely assigned (with the restriction that complex eigenvalues are in conjugate pairs) by a suitable choice of L.*

(vii) *(A^*, C^*) is controllable.*

Proof. First, we will show the equivalence between conditions (i) and (iii). Once this is done, the rest will follow by the duality or condition (vii).

$(i) \Leftarrow (iii)$: Note that given the input $u(t)$ and the initial condition x_0, the output in the time interval $[0, t_1]$ is given by

$$y(t) = Ce^{At}x(0) + \int_0^t Ce^{A(t-\tau)}Bu(\tau)d\tau + Du(t).$$

Since $y(t)$ and $u(t)$ are known, there is no loss of generality in assuming $u(t) = 0, \forall t$. Hence,

$$y(t) = Ce^{At}x(0), \quad t \in [0, t_1].$$

From this equation, we have

$$\begin{bmatrix} y(0) \\ \dot{y}(0) \\ \vdots \\ y^{(n-1)}(0) \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} x(0)$$

where $y^{(i)}$ stands for the i-th derivative of y. Since the observability matrix $\mathcal{O}$ has full column rank, there is a unique solution $x(0)$ in the above equation. This completes the proof.

$(i) \Rightarrow (iii)$: This will be proven by contradiction. Assume that (C, A) is observable but that the observability matrix does not have full column rank, i.e., there is a vector x_0 such that $\mathcal{O}x_0 = 0$ or equivalently $CA^i x_0 = 0$, $\forall i \geq 0$ by the Cayley-Hamilton Theorem. Now suppose the initial state $x(0) = x_0$, then $y(t) = Ce^{At}x(0) = 0$. This implies that the system is not observable since $x(0)$ cannot be determined from $y(t) \equiv 0$.

$\square$

Definition 3.5 The system, or the pair (C, A), is *detectable* if $A + LC$ is stable for some L.

Theorem 3.4 *The following are equivalent:*

(i) (C, A) *is detectable.*

(ii) The matrix $\begin{bmatrix} A - \lambda I \\ C \end{bmatrix}$ *has full column rank for all* $Re\lambda \geq 0$.

(iii) For all λ *and* x *such that* $Ax = \lambda x$ *and* $Re\lambda \geq 0$, $Cx \neq 0$.

(iv) There exists a matrix L *such that* $A + LC$ *is Hurwitz.*

(v) (A^*, C^*) *is stabilizable.*

The conditions (iv) and (v) of Theorem 3.1 and Theorem 3.3 and the conditions (ii) and (iii) of Theorem 3.2 and Theorem 3.4 are often called Popov-Belevitch-Hautus (PBH) tests. In particular, the following definitions of modal controllability and observability are often useful.

Definition 3.6 Let λ be an eigenvalue of A or, equivalently, a mode of the system. Then the mode λ is said to be controllable (observable) if $x^*B \neq 0$ $(Cx \neq 0)$ for *all* left (right) eigenvectors of A associated with λ, i.e., $x^*A = \lambda x^*$ $(Ax = \lambda x)$ and $0 \neq x \in \mathbb{C}^n$. Otherwise, the mode is said to be uncontrollable (unobservable).

It follows that a system is controllable (observable) if and only if every mode is controllable (observable). Similarly, a system is stabilizable (detectable) if and only if every unstable mode is controllable (observable).

For example, consider the following 4th order system:

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cccc|c} \lambda_1 & 1 & 0 & 0 & 0 \\ 0 & \lambda_1 & 1 & 0 & 1 \\ 0 & 0 & \lambda_1 & 0 & \alpha \\ 0 & 0 & 0 & \lambda_2 & 1 \\ \hline 1 & 0 & 0 & \beta & 0 \end{array}\right]$$

with $\lambda_1 \neq \lambda_2$. Then, the mode λ_1 is not controllable if $\alpha = 0$, and λ_2 is not observable if $\beta = 0$. Note that if $\lambda_1 = \lambda_2$, the system is uncontrollable and unobservable for any α and β since in that case, both

$$x_1 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \end{bmatrix} \quad \text{and} \quad x_2 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}$$

are the left eigenvectors of A corresponding to λ_1. Hence any linear combination of x_1 and x_2 is still an eigenvector of A corresponding to λ_1. In particular, let $x = x_1 - \alpha x_2$, then $x^*B = 0$, and as a result, the system is not controllable. Similar arguments can be applied to check observability. However, if the B matrix is changed into a 4×2 matrix with the last two rows independent of each other, then the system is controllable even if $\lambda_1 = \lambda_2$. For example, the reader may easily verify that the system with

$$B = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ \alpha & 1 \\ 1 & 0 \end{bmatrix}$$

is controllable for any α.

In general, for a system given in the Jordan canonical form, the controllability and observability can be concluded by inspection. The interested reader may easily derive some explicit conditions for the controllability and observability by using Jordan canonical form and the tests (iv) and (v) of Theorem 3.1 and Theorem 3.3.

3.3 Kalman Canonical Decomposition

There are usually many different coordinate systems to describe a dynamical system. For example, consider a simple pendulum, the motion of the pendulum can be uniquely determined either in terms of the angle of the string attached to the pendulum or in terms of the vertical displacement of the pendulum. However, in most cases, the angular displacement is a more natural description than the vertical displacement in spite of the fact that they both describe the same dynamical system. This is true for most physical dynamical systems. On the other hand, although some coordinates may not be natural descriptions of a physical dynamical system, they may make the system analysis and synthesis much easier.

In general, let $T \in \mathbb{R}^{n\times n}$ be a nonsingular matrix and define

$$\bar{x} = Tx.$$

Then the original dynamical system equations (3.1) and (3.2) become

$$\begin{aligned} \dot{\bar{x}} &= TAT^{-1}\bar{x} + TBu \\ y &= CT^{-1}\bar{x} + Du. \end{aligned}$$

These equations represent the same dynamical system for any nonsingular matrix T, and hence, we can regard these representations as equivalent. It is easy to see that the input/output transfer matrix is not changed under the coordinate transformation, i.e.,

$$G(s) = C(sI - A)^{-1}B + D = CT^{-1}(sI - TAT^{-1})^{-1}TB + D.$$

In this section, we will consider the system structure decomposition using coordinate transformation if the system is not completely controllable and/or is not completely observable. To begin with, let us consider further the dynamical systems related by a similarity transformation:

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \longmapsto \left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline \bar{C} & \bar{D} \end{array}\right] = \left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array}\right].$$

The controllability and observability matrices are related by

$$\bar{\mathcal{C}} = T\mathcal{C} \quad \bar{\mathcal{O}} = \mathcal{O}T^{-1}.$$

Moreover, the following theorem follows easily from the PBH tests or from the above relations.

Theorem 3.5 *The controllability (or stabilizability) and observability (or detectability) are invariant under similarity transformations.*

Using this fact, we can now show the following theorem.

Theorem 3.6 *If the controllability matrix $\mathcal{C}$ has rank $k_1 < n$, then there exists a similarity transformation*

$$\bar{x} = \begin{bmatrix} \bar{x}_c \\ \bar{x}_{\bar{c}} \end{bmatrix} = Tx$$

such that

$$\begin{aligned} \begin{bmatrix} \dot{\bar{x}}_c \\ \dot{\bar{x}}_{\bar{c}} \end{bmatrix} &= \begin{bmatrix} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{\bar{c}} \end{bmatrix} \begin{bmatrix} \bar{x}_c \\ \bar{x}_{\bar{c}} \end{bmatrix} + \begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix} u \\ y &= \begin{bmatrix} \bar{C}_c & \bar{C}_{\bar{c}} \end{bmatrix} \begin{bmatrix} \bar{x}_c \\ \bar{x}_{\bar{c}} \end{bmatrix} + Du \end{aligned}$$

with $\bar{A}_c \in \mathbb{C}^{k_1 \times k_1}$ and $(\bar{A}_c, \bar{B}_c)$ controllable. Moreover,

$$G(s) = C(sI - A)^{-1}B + D = \bar{C}_c(sI - \bar{A}_c)^{-1}\bar{B}_c + D.$$

Proof. Since $\operatorname{rank} \mathcal{C} = k_1 < n$, the pair (A, B) is not controllable. Now let $q_1, q_2, \ldots, q_{k_1}$ be any linearly independent columns of $\mathcal{C}$. Let $q_i, i = k_1 + 1, \ldots, n$ be any $n - k_1$ linearly independent vectors such that the matrix

$$Q := \begin{bmatrix} q_1 & \cdots & q_{k_1} & q_{k_1+1} & \cdots & q_n \end{bmatrix}$$

is nonsingular. Define

$$T := Q^{-1}.$$

Then the transformation $\bar{x} = Tx$ will give the desired decomposition. To see that, note that for each $i = 1, 2, \ldots, k_1$, Aq_i can be written as a linear combination of $q_i, i = 1, 2, \ldots, k_1$ since Aq_i is a linear combination of the columns of $\mathcal{C}$ by the Cayley-Hamilton Theorem. Therefore, we have

$$\begin{aligned} AT^{-1} &= \begin{bmatrix} Aq_1 & \cdots & Aq_{k_1} & Aq_{k_1+1} & \cdots & Aq_n \end{bmatrix} \\ &= \begin{bmatrix} q_1 & \cdots & q_{k_1} & q_{k_1+1} & \cdots & q_n \end{bmatrix} \begin{bmatrix} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{\bar{c}} \end{bmatrix} \\ &= T^{-1} \begin{bmatrix} \bar{A}_c & \bar{A}_{12} \\ 0 & \bar{A}_{\bar{c}} \end{bmatrix} \end{aligned}$$

for some $k_1 \times k_1$ matrix $\bar{A}_c$. Similarly, each column of the matrix B is a linear combination of $q_i, i = 1, 2, \ldots, k_1$, hence

$$B = Q \begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix} = T^{-1} \begin{bmatrix} \bar{B}_c \\ 0 \end{bmatrix}$$

for some $\bar{B}_c \in \mathbb{C}^{k_1 \times m}$.

To show that $(\bar{A}_c, \bar{B}_c)$ is controllable, note that rank $\mathcal{C} = k_1$ and

$$\mathcal{C} = T^{-1} \begin{bmatrix} \bar{B}_c & \bar{A}_c\bar{B}_c & \cdots & \bar{A}_c^{k_1-1}\bar{B}_c & \cdots & \bar{A}_c^{n-1}\bar{B}_c \\ 0 & 0 & \cdots & 0 & \cdots & 0 \end{bmatrix}.$$

Since, for each $j \geq k_1$, $\bar{A}_c^j$ is a linear combination of $\bar{A}_c^i, i = 0, 1, \ldots, (k_1 - 1)$ by Cayley-Hamilton theorem, we have

$$\operatorname{rank} \begin{bmatrix} \bar{B}_c & \bar{A}_c\bar{B}_c & \cdots & \bar{A}_c^{k_1-1}\bar{B}_c \end{bmatrix} = k_1,$$

i.e., $(\bar{A}_c, \bar{B}_c)$ is controllable. □

A numerically reliable way to find a such transformation T is to use QR factorization. For example, if the controllability matrix $\mathcal{C}$ has the QR factorization $QR = \mathcal{C}$, then $T = Q^{-1}$.

Corollary 3.7 *If the system is stabilizable and the controllability matrix $\mathcal{C}$ has rank $k_1 < n$, then there exists a similarity transformation T such that*

$$\left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array}\right] = \left[\begin{array}{cc|c} \bar{A}_c & \bar{A}_{12} & \bar{B}_c \\ 0 & \bar{A}_{\bar{c}} & 0 \\ \hline \bar{C}_c & \bar{C}_{\bar{c}} & D \end{array}\right]$$

with $\bar{A}_c \in \mathbb{C}^{k_1 \times k_1}$, $(\bar{A}_c, \bar{B}_c)$ controllable and with $\bar{A}_{\bar{c}}$ stable.

Hence, the state space $\bar{x}$ is partitioned into two orthogonal subspaces

$$\left\{ \begin{bmatrix} \bar{x}_c \\ 0 \end{bmatrix} \right\} \quad \text{and} \quad \left\{ \begin{bmatrix} 0 \\ \bar{x}_{\bar{c}} \end{bmatrix} \right\}$$

with the first subspace controllable from the input and second completely uncontrollable from the input (i.e., the state $\bar{x}_{\bar{c}}$ are not affected by the control u). Write these subspaces in terms of the original coordinates x, we have

$$x = T^{-1}\bar{x} = \begin{bmatrix} q_1 & \cdots & q_{k_1} & q_{k_1+1} & \cdots & q_n \end{bmatrix} \begin{bmatrix} \bar{x}_c \\ \bar{x}_{\bar{c}} \end{bmatrix}.$$

So the controllable subspace is the span of $q_i, i = 1, \ldots, k_1$ or, equivalently, Im $\mathcal{C}$. On the other hand, the uncontrollable subspace is given by the complement of the controllable subspace.

By duality, we have the following decomposition if the system is not completely observable.

Theorem 3.8 *If the observability matrix $\mathcal{O}$ has rank $k_2 < n$, then there exists a similarity transformation T such that*

$$\left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array}\right] = \left[\begin{array}{cc|c} \bar{A}_o & 0 & \bar{B}_o \\ \bar{A}_{21} & \bar{A}_{\bar{o}} & \bar{B}_{\bar{o}} \\ \hline \bar{C}_o & 0 & D \end{array}\right]$$

with $\bar{A}_o \in \mathbb{C}^{k_2 \times k_2}$ and $(\bar{C}_o, \bar{A}_o)$ observable.

Corollary 3.9 *If the system is detectable and the observability matrix $\mathcal{C}$ has rank $k_2 < n$, then there exists a similarity transformation T such that*

$$\left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array}\right] = \left[\begin{array}{cc|c} \bar{A}_o & 0 & \bar{B}_o \\ \bar{A}_{21} & \bar{A}_{\bar{o}} & \bar{B}_{\bar{o}} \\ \hline \bar{C}_o & 0 & D \end{array}\right]$$

with $\bar{A}_o \in \mathbb{C}^{k_2 \times k_2}$, $(\bar{C}_o, \bar{A}_o)$ observable and with $\bar{A}_{\bar{o}}$ stable.

Similarly, we have

$$G(s) = C(sI - A)^{-1}B + D = \bar{C}_o(sI - \bar{A}_o)^{-1}\bar{B}_o + D.$$

Carefully combining the above two theorems, we get the following *Kalman Canonical Decomposition*. The proof is left to the reader as an exercise.

Theorem 3.10 *Let an LTI dynamical system be described by the equations (3.1) and (3.2). Then there exists a nonsingular coordinate transformation $\bar{x} = Tx$ such that*

$$\begin{aligned} \begin{bmatrix} \dot{\bar{x}}_{co} \\ \dot{\bar{x}}_{c\bar{o}} \\ \dot{\bar{x}}_{\bar{c}o} \\ \dot{\bar{x}}_{\bar{c}\bar{o}} \end{bmatrix} &= \begin{bmatrix} \bar{A}_{co} & 0 & \bar{A}_{13} & 0 \\ \bar{A}_{21} & \bar{A}_{c\bar{o}} & \bar{A}_{23} & \bar{A}_{24} \\ 0 & 0 & \bar{A}_{\bar{c}o} & 0 \\ 0 & 0 & \bar{A}_{43} & \bar{A}_{\bar{c}\bar{o}} \end{bmatrix} \begin{bmatrix} \bar{x}_{co} \\ \bar{x}_{c\bar{o}} \\ \bar{x}_{\bar{c}o} \\ \bar{x}_{\bar{c}\bar{o}} \end{bmatrix} + \begin{bmatrix} \bar{B}_{co} \\ \bar{B}_{c\bar{o}} \\ 0 \\ 0 \end{bmatrix} u \\ y &= \begin{bmatrix} \bar{C}_{co} & 0 & \bar{C}_{\bar{c}o} & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_{co} \\ \bar{x}_{c\bar{o}} \\ \bar{x}_{\bar{c}o} \\ \bar{x}_{\bar{c}\bar{o}} \end{bmatrix} + Du \end{aligned}$$

or equivalently

$$\left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array}\right] = \left[\begin{array}{cccc|c} \bar{A}_{co} & 0 & \bar{A}_{13} & 0 & \bar{B}_{co} \\ \bar{A}_{21} & \bar{A}_{c\bar{o}} & \bar{A}_{23} & \bar{A}_{24} & \bar{B}_{c\bar{o}} \\ 0 & 0 & \bar{A}_{\bar{c}o} & 0 & 0 \\ 0 & 0 & \bar{A}_{43} & \bar{A}_{\bar{c}\bar{o}} & 0 \\ \hline \bar{C}_{co} & 0 & \bar{C}_{\bar{c}o} & 0 & D \end{array}\right]$$

where the vector $\bar{x}_{co}$ *is controllable and observable,* $\bar{x}_{c\bar{o}}$ *is controllable but unobservable,* $\bar{x}_{\bar{c}o}$ *is observable but uncontrollable, and* $\bar{x}_{\bar{c}\bar{o}}$ *is uncontrollable and unobservable. Moreover, the transfer matrix from* u *to* y *is given by*

$$G(s) = \bar{C}_{co}(sI - \bar{A}_{co})^{-1}\bar{B}_{co} + D.$$

One important issue is that although the transfer matrix of a dynamical system

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

is equal to its controllable and observable part

$$\left[\begin{array}{c|c} \bar{A}_{co} & \bar{B}_{co} \\ \hline \bar{C}_{co} & D \end{array}\right]$$

their internal behaviors are very different. In other words, while their input/output behaviors are the same, their state space response with nonzero initial conditions are very different. This can be illustrated by the state space response for the simple system

$$\begin{aligned}
\begin{bmatrix} \dot{\bar{x}}_{co} \\ \dot{\bar{x}}_{c\bar{o}} \\ \dot{\bar{x}}_{\bar{c}o} \\ \dot{\bar{x}}_{\bar{c}\bar{o}} \end{bmatrix} &= \begin{bmatrix} \bar{A}_{co} & 0 & 0 & 0 \\ 0 & \bar{A}_{c\bar{o}} & 0 & 0 \\ 0 & 0 & \bar{A}_{\bar{c}o} & 0 \\ 0 & 0 & 0 & \bar{A}_{\bar{c}\bar{o}} \end{bmatrix} \begin{bmatrix} \bar{x}_{co} \\ \bar{x}_{c\bar{o}} \\ \bar{x}_{\bar{c}o} \\ \bar{x}_{\bar{c}\bar{o}} \end{bmatrix} + \begin{bmatrix} \bar{B}_{co} \\ \bar{B}_{c\bar{o}} \\ 0 \\ 0 \end{bmatrix} u \\
y &= \begin{bmatrix} \bar{C}_{co} & 0 & \bar{C}_{\bar{c}o} & 0 \end{bmatrix} \begin{bmatrix} \bar{x}_{co} \\ \bar{x}_{c\bar{o}} \\ \bar{x}_{\bar{c}o} \\ \bar{x}_{\bar{c}\bar{o}} \end{bmatrix}
\end{aligned}$$

with $\bar{x}_{co}$ controllable and observable, $\bar{x}_{c\bar{o}}$ controllable but unobservable, $\bar{x}_{\bar{c}o}$ observable but uncontrollable, and $\bar{x}_{\bar{c}\bar{o}}$ uncontrollable and unobservable.

The solution to this system is given by

$$\begin{aligned}
\begin{bmatrix} \bar{x}_{co}(t) \\ \bar{x}_{c\bar{o}}(t) \\ \bar{x}_{\bar{c}o}(t) \\ \bar{x}_{\bar{c}\bar{o}}(t) \end{bmatrix} &= \begin{bmatrix} e^{\bar{A}_{co}t}\bar{x}_{co}(0) + \int_0^t e^{\bar{A}_{co}(t-\tau)}\bar{B}_{co}u(\tau)d\tau \\ e^{\bar{A}_{c\bar{o}}t}\bar{x}_{c\bar{o}}(0) + \int_0^t e^{\bar{A}_{c\bar{o}}(t-\tau)}\bar{B}_{c\bar{o}}u(\tau)d\tau \\ e^{\bar{A}_{\bar{c}o}t}\bar{x}_{\bar{c}o}(0) \\ e^{\bar{A}_{\bar{c}\bar{o}}t}\bar{x}_{\bar{c}\bar{o}}(0) \end{bmatrix} \\
y(t) &= \bar{C}_{co}\bar{x}_{co}(t) + \bar{C}_{\bar{c}o}\bar{x}_{\bar{c}o};
\end{aligned}$$

note that $\bar{x}_{\bar{c}o}(t)$ and $\bar{x}_{\bar{c}\bar{o}}(t)$ are not affected by the input u, while $\bar{x}_{c\bar{o}}(t)$ and $\bar{x}_{\bar{c}\bar{o}}(t)$ do not show up in the output y. Moreover, if the initial condition is zero, i.e., $\bar{x}(0) = 0$, then the output

$$y(t) = \int_0^t \bar{C}_{co}e^{\bar{A}_{co}(t-\tau)}\bar{B}_{co}u(\tau)d\tau.$$

However, if the initial state is not zero, then the response $\bar{x}_{\bar{c}o}(t)$ will show up in the output. In particular, if $\bar{A}_{\bar{c}o}$ is not stable, then the output $y(t)$ will grow without bound. The problems with uncontrollable and/or unobservable unstable modes are even more profound than what we have mentioned. Since the states are the internal signals of the dynamical system, any unbounded internal signal will eventually destroy the system. On the other hand, since it is impossible to make the initial states *exactly* zero, any uncontrollable and/or unobservable unstable mode will result in unacceptable system behavior. This issue will be exploited further in section 3.7. In the next section, we will consider how to place the system poles to achieve desired closed-loop behavior if the system is controllable.

3.4 Pole Placement and Canonical Forms

Consider a MIMO dynamical system described by

$$\begin{aligned}\dot{x} &= Ax + Bu \\ y &= Cx + Du,\end{aligned}$$

and let u be a state feedback control law given by

$$u = Fx + v.$$

This closed-loop system is as shown in Figure 3.1, and the closed-loop system equations are given by

$$\begin{aligned}\dot{x} &= (A + BF)x + Bv \\ y &= (C + DF)x + Dv.\end{aligned}$$

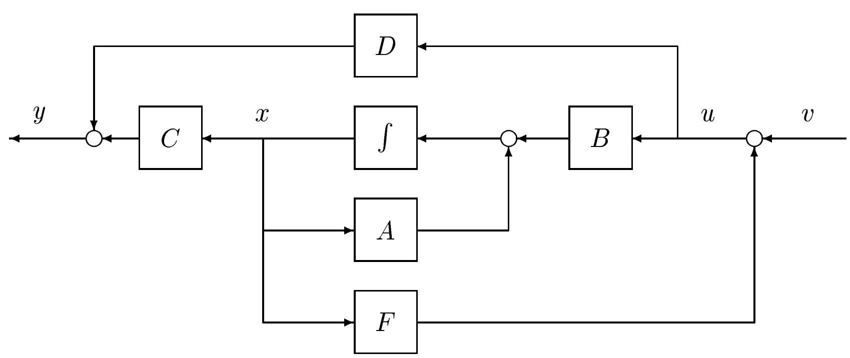


Figure 3.1: State Feedback

Then we have the following lemma in which the proof is simple and is left as an exercise to the reader.

Lemma 3.11 *Let F be a constant matrix with appropriate dimension; then (A, B) is controllable (stabilizable) if and only if $(A + BF, B)$ is controllable (stabilizable).*

However, the observability of the system may change under state feedback. For example, the following system

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|c} 1 & 1 & 1 \\ 1 & 0 & 0 \\ \hline 1 & 0 & 0 \end{array}\right]$$

is controllable and observable. But with the state feedback

$$u = Fx = \left[\begin{array}{cc} -1 & -1 \end{array}\right] x,$$

the system becomes

$$\left[\begin{array}{c|c} A + BF & B \\ \hline C + DF & D \end{array}\right] = \left[\begin{array}{cc|c} 0 & 0 & 1 \\ 1 & 0 & 0 \\ \hline 1 & 0 & 0 \end{array}\right]$$

and is not completely observable.

The dual operation of the dynamical system by

$$\dot{x} = Ax + Bu \longmapsto \dot{x} = Ax + Bu + Ly$$

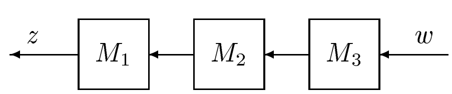
is called *output injection* which can be written as

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \longmapsto \left[\begin{array}{c|c} A + LC & B + LD \\ \hline C & D \end{array}\right].$$

By duality, the output injection does not change the system observability (detectability) but may change the system controllability (stabilizability).

Remark 3.1 We would like to call attention to the diagram and the signals flow convention used in this book. It may not be conventional to let signals flow from the right to the left, however, the reader will find that this representation is much more visually appealing than the traditional representation, and is consistent with the matrix manipulations. For example, the following diagram represents the multiplication of three matrices and will be very helpful in dealing with complicated systems:

$$z = M_1 M_2 M_3 w.$$



The conventional signal block diagrams, i.e., signals flowing from left to right, will also be used in this book. ♡

We now consider some special state space representations of the dynamical system described by equations (3.1) and (3.2). First, we will consider the systems with single inputs.

Assume that a single input and multiple output dynamical system is given by

$$G(s) = \left[\begin{array}{c|c} A & b \\ \hline C & d \end{array}\right], \quad b \in \mathbb{R}^n, \; C \in \mathbb{R}^{p \times n}, \; d \in \mathbb{R}^p$$

and assume that (A, b) is controllable. Let

$$\det(\lambda I - A) = \lambda^n + a_1 \lambda^{n-1} + \cdots + a_n,$$

and define

$$A_1 := \begin{bmatrix} -a_1 & -a_2 & \cdots & -a_{n-1} & -a_n \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \qquad b_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

and

$$\begin{aligned} \mathcal{C} &= \begin{bmatrix} b & Ab & \cdots & A^{n-1}b \end{bmatrix} \\ \mathcal{C}_1 &= \begin{bmatrix} b_1 & A_1 b_1 & \cdots & A_1^{n-1} b_1 \end{bmatrix}. \end{aligned}$$

Then it is easy to verify that both $\mathcal{C}$ and $\mathcal{C}_1$ are nonsingular. Moreover, the transformation

$$T_c = \mathcal{C}_1 \mathcal{C}^{-1}$$

will give the equivalent system representation

$$\left[\begin{array}{c|c} T_c A T_c^{-1} & T_c b \\ \hline C T_c^{-1} & d \end{array}\right] = \left[\begin{array}{c|c} A_1 & b_1 \\ \hline C T_c^{-1} & d \end{array}\right]$$

where

$$C T_c^{-1} = \begin{bmatrix} \beta_1 & \beta_2 & \cdots & \beta_{n-1} & \beta_n \end{bmatrix}$$

for some $\beta_i \in \mathbb{R}^p$. This state space representation is usually called *controllable canonical form* or *controller canonical form*. It is also easy to show that the transfer matrix is given by

$$G(s) = C(sI - A)^{-1} b + d = \frac{\beta_1 s^{n-1} + \beta_2 s^{n-2} + \cdots + \beta_{n-1} s + \beta_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1} s + a_n} + d,$$

which also shows that, given a column vector of transfer matrix $G(s)$, a state space representation of the transfer matrix can be obtained as above. The quadruple (A, b, C, d) is called a state space *realization* of the transfer matrix $G(s)$.

Now consider a dynamical system equation given by

$$\dot{x} = A_1 x + b_1 u$$

and a state feedback control law

$$u = Fx = \begin{bmatrix} f_1 & f_2 & \dots & f_{n-1} & f_n \end{bmatrix} x.$$

Then the closed-loop system is given by

$$\dot{x} = (A_1 + b_1 F)x$$

and $\det(\lambda I - (A_1 + b_1 F)) = \lambda^n + (a_1 - f_1)\lambda^{n-1} + \cdots + (a_n - f_n)$. It is clear that the zeros of $\det(\lambda I - (A_1 + b_1 F)$ can be made arbitrary for an appropriate choice of F provided that the complex zeros appear in conjugate pairs, thus showing that the eigenvalues of $A + bF$ can be freely assigned if (A, b) is controllable.

Dually, consider a multiple input and single output system

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline c & d \end{array}\right], \quad B \in \mathbb{R}^{n\times m}, \; c^* \in \mathbb{R}^n, \; d^* \in \mathbb{R}^m,$$

and assume that (c, A) is observable; then there is a transformation T_o such that

$$\left[\begin{array}{c|c} T_o A T_o^{-1} & T_o B \\ \hline c T_o^{-1} & d \end{array}\right] = \left[\begin{array}{ccccc|c} -a_1 & 1 & 0 & \cdots & 0 & \eta_1 \\ -a_2 & 0 & 1 & \cdots & 0 & \eta_2 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ -a_{n-1} & 0 & 0 & \cdots & 1 & \eta_{n-1} \\ -a_n & 0 & 0 & \cdots & 0 & \eta_n \\ \hline 1 & 0 & 0 & \cdots & 0 & d \end{array}\right], \quad \eta_i^* \in \mathbb{R}^m$$

and

$$G(s) = c(sI - A)^{-1}B + d = \frac{\eta_1 s^{n-1} + \eta_2 s^{n-2} + \cdots + \eta_{n-1}s + \eta_n}{s^n + a_1 s^{n-1} + \cdots + a_{n-1}s + a_n} + d.$$

This representation is called *observable canonical form* or *observer canonical form*.

Similarly, we can show that the eigenvalues of $A + Lc$ can be freely assigned if (c, A) is observable.

The pole placement problem for a multiple input system (A, B) can be converted into a simple single input pole placement problem. To describe the procedure, we need some preliminary results.

Lemma 3.12 *If an m input system pair (A, B) is controllable and if A is cyclic, then for almost all $v \in \mathbb{R}^m$, the single input pair (A, Bv) is controllable.*

Proof. Without loss of generality, assume that A is in the Jordan canonical form and that the matrix B is partitioned accordingly:

$$A = \begin{bmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_k \end{bmatrix} \quad B = \begin{bmatrix} B_1 \\ B_2 \\ \vdots \\ B_k \end{bmatrix}$$

where J_i is in the form of

$$J_i = \begin{bmatrix} \lambda_i & 1 & & & \\ & \lambda_i & \ddots & & \\ & & \ddots & \ddots & \\ & & & \lambda_i & 1 \\ & & & & \lambda_i \end{bmatrix}$$

and $\lambda_i \neq \lambda_j$ if $i \neq j$. By PBH test, the pair (A, B) is controllable if and only if, for each $i = 1, \ldots, k$, the last row of B_i is not zero. Let $b_i \in \mathbb{R}^m$ be the last row of B_i, and then we only need to show that, for almost all $v \in \mathbb{R}^m$, $b_i v \neq 0$ for each $i = 1, \ldots, k$ which is clear since for each i, the set $v \in \mathbb{R}^m$ such that $b_i v = 0$ has measure zero in $\mathbb{R}^m$ since $b_i \neq 0$. □

The cyclicity assumption in this theorem is essential. Without this assumption, the theorem does not hold. For example, the pair

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

is controllable but there is no $v \in \mathbb{R}^2$ such that (A, Bv) is controllable since A is not cyclic.

Since a matrix A with distinct eigenvalues is cyclic, by definition we have the following lemma.

Lemma 3.13 *If (A, B) is controllable, then for almost any $K \in \mathbb{R}^{m\times n}$, all the eigenvalues of $A + BK$ are distinct and, consequently, $A + BK$ is cyclic.*

A proof can be found in Brasch and Pearson [1970], Davison [1968], and Heymann [1968].

Now it is evident that given a multiple input controllable pair (A, B), there is a matrix $K \in \mathbb{R}^{m\times n}$ and a vector $v \in \mathbb{R}^m$ such that $A + BK$ is cyclic and $(A + BK, Bv)$ is controllable. Moreover, from the pole placement results for the single input system, there is a matrix $f^* \in \mathbb{R}^n$ so that the eigenvalues of $(A+BK)+(Bv)f$ can be arbitrarily assigned. Hence, the eigenvalues of $A + BF$ can be arbitrarily assigned by choosing a state feedback in the form of

$$u = Fx = (K + vf)x.$$

A dual procedure can be applied for the output injection $A + LC$.

The canonical form for single input or single output system can also be generalized to multiple input and multiple output systems at the expense of notation. The interested reader may consult Kailath [1980] or Chen [1984].

If a system is not completely controllable, then the Kalman controllable decomposition can be applied first and the above procedure can be used to assign the system poles corresponding to the controllable subspace.

3.5 Observers and Observer-Based Controllers

We have shown in the last section that if a system is controllable and, furthermore, if the system states are available for feedback, then the system closed loop poles can be assigned arbitrarily through a constant feedback. However, in most practical applications, the system states are not completely accessible and all the designer knows are the output y and input u. Hence, the estimation of the system states from the given output information y and input u is often necessary to realize some specific design objectives. In this section, we consider such an estimation problem and the application of this state estimation in feedback control.

Consider a plant modeled by equations (3.1) and (3.2). An observer is a dynamical system with input of (u, y) and output of, say $\hat{x}$, which asymptotically estimates the state x. More precisely, a (linear) *observer* is a system such as

$$\begin{aligned} \dot{q} &= Mq + Nu + Hy \\ \hat{x} &= Qq + Ru + Sy \end{aligned}$$

so that $\hat{x}(t) - x(t) \to 0$ as $t \to \infty$ for all initial states $x(0)$, $q(0)$ and for every input $u(\cdot)$.

Theorem 3.14 *An observer exists iff (C, A) is detectable. Further, if (C, A) is detectable, then a full order Luenberger observer is given by*

$$\dot{q} = Aq + Bu + L(Cq + Du - y) \tag{3.5}$$
$$\hat{x} = q \tag{3.6}$$

where L is any matrix such that $A + LC$ is stable.

Proof. We first show that the detectability of (C, A) is sufficient for the existence of an observer. To that end, we only need to show that the so-called Luenberger observer defined in the theorem is indeed an observer. Note that equation (3.5) for q is a simulation of the equation for x, with an additional forcing term $L(Cq + Du - y)$, which is a gain times the output error. Equivalent equations are

$$\begin{aligned} \dot{q} &= (A + LC)q + Bu + LDu - Ly \\ \hat{x} &= q. \end{aligned}$$

These equations have the form allowed in the definition of an observer. Define the error, $e := \hat{x} - x$, and then simple algebra gives

$$\dot{e} = (A + LC)e;$$

therefore $e(t) \to 0$ as $t \to \infty$ for every $x(0)$, $q(0)$, and $u(.)$.

To show the converse, assume that (C, A) is not detectable. Take the initial state $x(0)$ in the undetectable subspace and consider a candidate observer:

$$\begin{aligned} \dot{q} &= Mq + Nu + Hy \\ \hat{x} &= Qq + Ru + Sy. \end{aligned}$$

Take $q(0) = 0$ and $u(t) \equiv 0$. Then the equations for x and the candidate observer are

$$\begin{aligned} \dot{x} &= Ax \\ \dot{q} &= Mq + HCx \\ \hat{x} &= Qq + SCx. \end{aligned}$$

Since an unobservable subspace is an A-invariant subspace containing $x(0)$, it follows that $x(t)$ is in the unobservable subspace for all $t \geq 0$. Hence, $Cx(t) = 0$ for all $t \geq 0$, and, consequently, $q(t) \equiv 0$ and $\hat{x}(t) \equiv 0$. However, for some $x(0)$ in the undetectable subspace, $x(t)$ does not converge to zero. Thus the candidate observer does not have the required property, and therefore, no observer exists. □

The above Luenberger observer has dimension n, which is the dimension of the state x. It's possible to get an observer of lower dimension. The idea is this: since we can measure $y - Du = Cx$, we already know x modulo Ker C, so we only need to generate the part of x in Ker C. If C has full row rank and $p := \dim y$, then the dimension of Ker C equals $n - p$, so we might suspect that we can get an observer of order $n - p$. This is true. Such an observer is called a "minimal order observer". We will not pursue this issue further here. The interested reader may consult Chen [1984].

Recall that, for a dynamical system described by the equations (3.1) and (3.2), if (A, B) is controllable and state x is available for feedback, then there is a state feedback $u = Fx$ such that the closed-loop poles of the system can be arbitrarily assigned. Similarly, if (C, A) is observable, then the system observer poles can be arbitrarily placed so that the state estimator $\hat{x}$ can be made to approach x arbitrarily fast. Now let us consider what will happen if the system states are not available for feedback so that the estimated state has to be used. Hence, the controller has the following dynamics:

$$\begin{aligned} \dot{\hat{x}} &= (A + LC)\hat{x} + Bu + LDu - Ly \\ u &= F\hat{x}. \end{aligned}$$

Then the total system state equations are given by

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A & BF \\ -LC & A+BF+LC \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}.$$

Let $e := x - \hat{x}$, then the system equation becomes

$$\begin{bmatrix} \dot{e} \\ \dot{\hat{x}} \end{bmatrix} = \begin{bmatrix} A+LC & 0 \\ -LC & A+BF \end{bmatrix} \begin{bmatrix} e \\ \hat{x} \end{bmatrix},$$

and the closed-loop poles consist of two parts: the poles resulting from state feedback $\sigma(A+BF)$ and the poles resulting from the state estimation $\sigma(A+LC)$. Now if (A, B) is controllable and (C, A) is observable, then there exist F and L such that the eigenvalues of $A + BF$ and $A + LC$ can be arbitrarily assigned. In particular, they can be made to be stable. Note that a slightly weaker result can also result even if (A, B) and (C, A) are only stabilizable and detectable.

The controller given above is called an observer-based controller and is denoted as

$$u = K(s)y$$

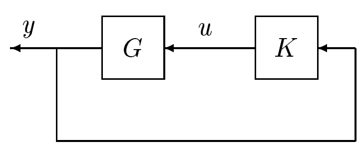
and

$$K(s) = \left[\begin{array}{c|c} A+BF+LC+LDF & -L \\ \hline F & 0 \end{array}\right].$$

Now denote the open loop plant by

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right];$$

then the closed-loop feedback system is as shown below:



In general, if a system is stabilizable through feeding back the output y, then it is said to be *output feedback stabilizable*. It is clear from the above construction that a system is output feedback stabilizable if (A, B) is stabilizable and (C, A) is detectable. The converse is also true and will be shown in Chapter 12.

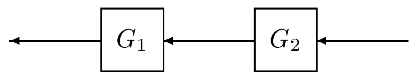
3.6 Operations on Systems

In this section, we present some facts about system interconnection. Since these proofs are straightforward, we will leave the details to the reader.

Suppose that G_1 and G_2 are two subsystems with state space representations:

$$G_1 = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] \qquad G_2 = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right].$$

Then the series or cascade connection of these two subsystems is a system with the output of the second subsystem as the input of the first subsystem as shown in the following diagram:



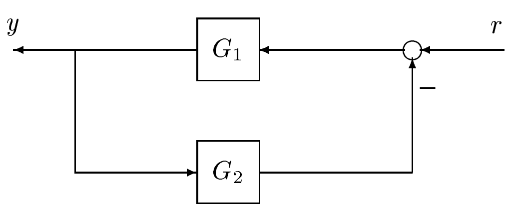
This operation in terms of the transfer matrices of the two subsystems is essentially the product of two transfer matrices. Hence, a representation for the cascaded system can be obtained as

$$\begin{aligned} G_1 G_2 &= \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right] \\ &= \left[\begin{array}{cc|c} A_1 & B_1 C_2 & B_1 D_2 \\ 0 & A_2 & B_2 \\ \hline C_1 & D_1 C_2 & D_1 D_2 \end{array}\right] = \left[\begin{array}{cc|c} A_2 & 0 & B_2 \\ B_1 C_2 & A_1 & B_1 D_2 \\ \hline D_1 C_2 & C_1 & D_1 D_2 \end{array}\right]. \end{aligned}$$

Similarly, the parallel connection or the addition of G_1 and G_2 can be obtained as

$$G_1 + G_2 = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] + \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & D_2 \end{array}\right] = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D_1 + D_2 \end{array}\right].$$

Next we consider a feedback connection of G_1 and G_2 as shown below:



Then the closed-loop transfer matrix from r to y is given by

$$T = \left[\begin{array}{cc|c} A_1 - B_1 D_2 R_{12}^{-1} C_1 & -B_1 R_{21}^{-1} C_2 & B_1 R_{21}^{-1} \\ B_2 R_{12}^{-1} C_1 & A_2 - B_2 D_1 R_{21}^{-1} C_2 & B_2 D_1 R_{21}^{-1} \\ \hline R_{12}^{-1} C_1 & -R_{12}^{-1} D_1 C_2 & D_1 R_{21}^{-1} \end{array}\right]$$

where $R_{12} = I + D_1 D_2$ and $R_{21} = I + D_2 D_1$. Note that these state space representations may not be necessarily controllable and/or observable even if the original subsystems G_1 and G_2 are.

For future reference, we shall also introduce the following definitions.

Definition 3.7 The *transpose* of a transfer matrix $G(s)$ or the *dual system* is defined as

$$G \longmapsto G^T(s) = B^*(sI - A^*)^{-1}C^* + D^*$$

or equivalently

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \longmapsto \left[\begin{array}{c|c} A^* & C^* \\ \hline B^* & D^* \end{array}\right].$$

Definition 3.8 The *conjugate* system of $G(s)$ is defined as

$$G \longmapsto G^{\sim}(s) := G^T(-s) = B^*(-sI - A^*)^{-1}C^* + D^*$$

or equivalently

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \longmapsto \left[\begin{array}{c|c} -A^* & -C^* \\ \hline B^* & D^* \end{array}\right].$$

In particular, we have $G^*(j\omega) := [G(j\omega)]^* = G^{\sim}(j\omega)$.

Definition 3.9 A real rational matrix $\hat{G}(s)$ is called a *right (left) inverse* of a transfer matrix $G(s)$ if $G(s)\hat{G}(s) = I$ ($\hat{G}(s)G(s) = I$). Moreover, if $\hat{G}(s)$ is both a right inverse and a left inverse of $G(s)$, then it is simply called the inverse of $G(s)$.

Lemma 3.15 *Let $D^\dagger$ denote a right (left) inverse of D if D has full row (column) rank. Then*

$$G^\dagger = \left[\begin{array}{c|c} A - BD^\dagger C & -BD^\dagger \\ \hline D^\dagger C & D^\dagger \end{array}\right]$$

is a right (left) inverse of G.

Proof. The right inverse case will be proven and the left inverse case follows by duality. Suppose $DD^\dagger = I$. Then

$$\begin{aligned} GG^\dagger &= \left[\begin{array}{cc|c} A & BD^\dagger C & BD^\dagger \\ 0 & A - BD^\dagger C & -BD^\dagger \\ \hline C & DD^\dagger C & DD^\dagger \end{array}\right] \\ &= \left[\begin{array}{cc|c} A & BD^\dagger C & BD^\dagger \\ 0 & A - BD^\dagger C & -BD^\dagger \\ \hline C & C & I \end{array}\right]. \end{aligned}$$

Performing similarity transformation $T = \begin{bmatrix} I & I \\ 0 & I \end{bmatrix}$ on the above system yields

$$\begin{aligned} GG^\dagger &= \left[\begin{array}{cc|c} A & 0 & 0 \\ 0 & A - BD^\dagger C & -BD^\dagger \\ \hline C & 0 & I \end{array}\right] \\ &= I. \end{aligned}$$

□

3.7 State Space Realizations for Transfer Matrices

In some cases, the natural or convenient description for a dynamical system is in terms of transfer matrices. This occurs, for example, in some highly complex systems for which the analytic differential equations are too hard or too complex to write down. Hence, certain engineering approximation or identification has to be carried out; for example, input and output frequency responses are obtained from experiments so that some transfer matrix approximating the system dynamics is available. Since the state space computation is most convenient to implement on the computer, some appropriate state space representation for the resulting transfer matrix is necessary.

In general, assume that $G(s)$ is a real-rational transfer matrix which is *proper*. Then we call a state-space model (A, B, C, D) such that

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right],$$

a *realization* of $G(s)$.

We note that if the transfer matrix is either single input or single output, then the formulas in Section 3.4 can be used to obtain a controllable or observable realization. The realization for a general MIMO transfer matrix is more complicated and is the focus of this section.

Definition 3.10 A state space realization (A, B, C, D) of $G(s)$ is said to be a *minimal realization* of $G(s)$ if A has the smallest possible dimension.

Theorem 3.16 *A state space realization* (A, B, C, D) *of* $G(s)$ *is minimal if and only if* (A, B) *is controllable and* (C, A) *is observable.*

Proof. We shall first show that if (A, B, C, D) is a minimal realization of $G(s)$, then (A, B) must be controllable and (C, A) must be observable. Suppose, on the contrary, that (A, B) is not controllable and/or (C, A) is not observable. Then from Kalman decomposition, there is a smaller dimensioned controllable and observable state space realization that has the same transfer matrix; this contradicts the minimality assumption. Hence (A, B) must be controllable and (C, A) must be observable.

Next we show that if an n-th order realization (A, B, C, D) is controllable and observable, then it is minimal. But supposing it is not minimal, let (A_m, B_m, C_m, D) be a minimal realization of $G(s)$ with order $k < n$. Since

$$G(s) = C(sI - A)^{-1}B + D = C_m(sI - A_m)^{-1}B_m + D,$$

we have

$$CA^iB = C_mA_m^iB_m, \quad \forall i \geq 0.$$

This implies that

$$\mathcal{O}\mathcal{C} = \mathcal{O}_m\mathcal{C}_m \tag{3.7}$$

where $\mathcal{C}$ and $\mathcal{O}$ are the controllability and observability matrices of (A, B) and (C, A), respectively, and

$$\mathcal{C}_m := \begin{bmatrix} B_m & A_m B_m & \cdots & A_m^{n-1} B_m \end{bmatrix}$$
$$\mathcal{O}_m := \begin{bmatrix} C_m \\ C_m A_m \\ \vdots \\ C_m A_m^{n-1} \end{bmatrix}.$$

By Sylvester's inequality,

$$\text{rank } \mathcal{C} + \text{rank } \mathcal{O} - n \leq \text{rank } (\mathcal{O}\mathcal{C}) \leq \min\{\text{rank } \mathcal{C}, \text{rank } \mathcal{O}\},$$

and, therefore, we have rank $(\mathcal{O}\mathcal{C}) = n$ since rank $\mathcal{C} =$ rank $\mathcal{O} = n$ by the controllability and observability assumptions. Similarly, since (A_m, B_m, C_m, D) is minimal, (A_m, B_m) is controllable and (C_m, A_m) is observable. Moreover,

$$\text{rank } \mathcal{O}_m \mathcal{C}_m = k < n,$$

which is a contradiction since rank $\mathcal{O}\mathcal{C} =$ rank $\mathcal{O}_m \mathcal{C}_m$ by equality (3.7). □

The following property of minimal realizations can also be verified, and this is left to the reader.

Theorem 3.17 *Let (A_1, B_1, C_1, D) and (A_2, B_2, C_2, D) be two minimal realizations of a real rational transfer matrix $G(s)$, and let $\mathcal{C}_1$, $\mathcal{C}_2$, $\mathcal{O}_1$, and $\mathcal{O}_2$ be the corresponding controllability and observability matrices, respectively. Then there exists a unique nonsingular T such that*

$$A_2 = T A_1 T^{-1}, \quad B_2 = T B_1, \quad C_2 = C_1 T^{-1}.$$

Furthermore, T can be specified as $T = (\mathcal{O}_2^* \mathcal{O}_2)^{-1} \mathcal{O}_2^* \mathcal{O}_1$ or $T^{-1} = \mathcal{C}_1 \mathcal{C}_2^* (\mathcal{C}_2 \mathcal{C}_2^*)^{-1}$.

We now describe several ways to obtain a state space realization for a given multiple input and multiple output transfer matrix $G(s)$. The simplest and most straightforward way to obtain a realization is by realizing each element of the matrix $G(s)$ and then combining all these individual realizations to form a realization for $G(s)$. To illustrate, let us consider a 2×2 (block) transfer matrix such as

$$G(s) = \begin{bmatrix} G_1(s) & G_2(s) \\ G_3(s) & G_4(s) \end{bmatrix}$$

and assume that $G_i(s)$ has a state space realization of

$$G_i(s) = \left[\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right], \quad i = 1, \ldots, 4.$$

Note that $G_i(s)$ may itself be a multiple input and multiple output transfer matrix. In particular, if $G_i(s)$ is a column or row vector of transfer functions, then the formulas in Section 3.4 can be used to obtain a controllable or observable realization for $G_i(s)$. Then a realization for $G(s)$ can be given by

$$G(s) = \left[\begin{array}{cccc|cc} A_1 & 0 & 0 & 0 & B_1 & 0 \\ 0 & A_2 & 0 & 0 & 0 & B_2 \\ 0 & 0 & A_3 & 0 & B_3 & 0 \\ 0 & 0 & 0 & A_4 & 0 & B_4 \\ \hline C_1 & C_2 & 0 & 0 & D_1 & D_2 \\ 0 & 0 & C_3 & C_4 & D_3 & D_4 \end{array}\right].$$

Alternatively, if the transfer matrix $G(s)$ can be factored into the product and/or the sum of several simply realized transfer matrices, then a realization for G can be obtained by using the cascade or addition formulas in the last section.

A problem inherited with these kinds of realization procedures is that a realization thus obtained will generally not be minimal. To obtain a minimal realization, a Kalman controllability and observability decomposition has to be performed to eliminate the uncontrollable and/or unobservable states. (An alternative numerically reliable method to eliminate uncontrollable and/or unobservable states is the *balanced realization* method which will be discussed later.)

We will now describe one factorization procedure that does result in a minimal realization by using partial fractional expansion (The resulting realization is sometimes called *Gilbert's realization* due to Gilbert).

Let $G(s)$ be a $p \times m$ transfer matrix and write it in the following form:

$$G(s) = \frac{N(s)}{d(s)}$$

with $d(s)$ a scalar polynomial. For simplicity, we shall assume that $d(s)$ has only real and distinct roots $\lambda_i \neq \lambda_j$ if $i \neq j$ and

$$d(s) = (s-\lambda_1)(s-\lambda_2)\cdots(s-\lambda_r).$$

Then $G(s)$ has the following partial fractional expansion:

$$G(s) = D + \sum_{i=1}^{r} \frac{W_i}{s-\lambda_i}.$$

Suppose

$$\operatorname{rank} W_i = k_i$$

and let $B_i \in \mathbb{R}^{k_i \times m}$ and $C_i \in \mathbb{R}^{p \times k_i}$ be two constant matrices such that

$$W_i = C_i B_i.$$

Then a realization for $G(s)$ is given by

$$G(s) = \left[\begin{array}{ccc|c} \lambda_1 I_{k_1} & & & B_1 \\ & \ddots & & \vdots \\ & & \lambda_r I_{k_r} & B_r \\ \hline C_1 & \cdots & C_r & D \end{array}\right].$$

It follows immediately from PBH tests that this realization is controllable and observable. Hence, it is minimal.

An immediate consequence of this minimal realization is that a transfer matrix with an r-th order polynomial denominator does not necessarily have an r-th order state space realization unless W_i for each i is a rank one matrix.

This approach can, in fact, be generalized to more complicated cases where $d(s)$ may have complex and/or repeated roots. Readers may convince themselves by trying some simple examples.

3.8 Lyapunov Equations

Testing stability, controllability, and observability of a system is very important in linear system analysis and synthesis. However, these tests often have to be done indirectly. In that respect, the Lyapunov theory is sometimes useful. Consider the following Lyapunov equation

$$A^*X + XA + Q = 0 \tag{3.8}$$

with given real matrices A and Q. It has been shown in Chapter 2 that this equation has a unique solution iff $\lambda_i(A) + \bar{\lambda}_j(A) \neq 0, \forall i, j$. In this section, we will study the relationships between the stability of A and the solution of X. The following results are standard.

Lemma 3.18 *Assume that A is stable, then the following statements hold:*

(i) $X = \int_0^\infty e^{A^*t} Q e^{At} dt$.

(ii) $X > 0$ *if* $Q > 0$ *and* $X \geq 0$ *if* $Q \geq 0$.

(iii) if $Q \geq 0$, *then* (Q, A) *is observable iff* $X > 0$.

An immediate consequence of part (iii) is that, given a stable matrix A, a pair (C, A) is observable if and only if the solution to the following Lyapunov equation is positive definite:

$$A^*L_o + L_oA + C^*C = 0.$$

The solution L_o is called the *observability Gramian*. Similarly, a pair (A, B) is controllable if and only if the solution to

$$AL_c + L_cA^* + BB^* = 0$$

is positive definite and L_c is called the *controllability Gramian*.

In many applications, we are given the solution of the Lyapunov equation and need to conclude the stability of the matrix A.

Lemma 3.19 *Suppose X is the solution of the Lyapunov equation (3.8), then*

(i) $Re\lambda_i(A) \leq 0$ *if* $X > 0$ *and* $Q \geq 0$.

(ii) A *is stable if* $X > 0$ *and* $Q > 0$.

(iii) A *is stable if* $X \geq 0$, $Q \geq 0$ *and* (Q, A) *is detectable.*

Proof. Let λ be an eigenvalue of A and $v \neq 0$ be a corresponding eigenvector, then $Av = \lambda v$. Pre-multiply equation (3.8) by v^* and postmultiply (3.8) by v to get

$$2\text{Re } \lambda(v^*Xv) + v^*Qv = 0.$$

Now if $X > 0$ then $v^*Xv > 0$, and it is clear that $\text{Re}\lambda \leq 0$ if $Q \geq 0$ and $\text{Re}\lambda < 0$ if $Q > 0$. Hence (i) and (ii) hold. To see (iii), we assume $\text{Re}\lambda \geq 0$. Then we must have $v^*Qv = 0$, i.e., $Qv = 0$. This implies that λ is an unstable and unobservable mode, which contradicts the assumption that (Q, A) is detectable. □

3.9 Balanced Realizations

Although there are infinitely many different state space realizations for a given transfer matrix, some particular realizations have proven to be very useful in control engineering and signal processing. Here we will only introduce one class of realizations for stable transfer matrices that are most useful in control applications. To motivate the class of realizations, we first consider some simple facts.

Lemma 3.20 *Let* $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *be a state space realization of a (not necessarily stable) transfer matrix $G(s)$. Suppose that there exists a symmetric matrix*

$$P = P^* = \begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix}$$

with P_1 nonsingular such that

$$AP + PA^* + BB^* = 0.$$

Now partition the realization (A, B, C, D) compatibly with P as

$$\left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right].$$

Then

$$\left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$$

is also a realization of G. *Moreover,* (A_{11}, B_1) *is controllable if* A_{11} *is stable.*

Proof. Use the partitioned P and (A, B, C) to get

$$0 = AP + PA^* + BB^* = \begin{bmatrix} A_{11}P_1 + P_1A_{11}^* + B_1B_1^* & P_1A_{21}^* + B_1B_2^* \\ A_{21}P_1 + B_2B_1^* & B_2B_2^* \end{bmatrix},$$

which gives $B_2 = 0$ and $A_{21} = 0$ since P_1 is nonsingular. Hence, part of the realization is not controllable:

$$\left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ 0 & A_{22} & 0 \\ \hline C_1 & C_2 & D \end{array}\right] = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right].$$

Finally, it follows from Lemma 3.18 that (A_{11}, B_1) is controllable if A_{11} is stable. □

We also have

Lemma 3.21 *Let* $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *be a state space realization of a (not necessarily stable) transfer matrix* $G(s)$. *Suppose that there exists a symmetric matrix*

$$Q = Q^* = \begin{bmatrix} Q_1 & 0 \\ 0 & 0 \end{bmatrix}$$

with Q_1 *nonsingular such that*

$$QA + A^*Q + C^*C = 0.$$

Now partition the realization (A, B, C, D) *compatibly with* Q *as*

$$\left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right].$$

Then

$$\left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$$

is also a realization of G. *Moreover,* (C_1, A_{11}) *is observable if* A_{11} *is stable.*

The above two lemmas suggest that to obtain a minimal realization from a stable non-minimal realization, one only needs to eliminate all states corresponding to the zero block diagonal term of the controllability Gramian P and the observability Gramian Q. In the case where P is not block diagonal, the following procedure can be used to eliminate non-controllable subsystems:

1. Let $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ be a stable realization.

2. Compute the controllability Gramian $P \geq 0$ from
$$AP + PA^* + BB^* = 0.$$

3. Diagonalize P to get $P = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \begin{bmatrix} \Lambda_1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} U_1 & U_2 \end{bmatrix}^*$ with $\Lambda_1 > 0$ and $\begin{bmatrix} U_1 & U_2 \end{bmatrix}$ unitary.

4. Then $G(s) = \left[\begin{array}{c|c} U_1^* A U_1 & U_1^* B \\ \hline C U_1 & D \end{array}\right]$ is a controllable realization.

A dual procedure can also be applied to eliminate non-observable subsystems.

Now assume that $\Lambda_1 > 0$ is diagonal and $\Lambda_1 = \text{diag}(\Lambda_{11}, \Lambda_{12})$ such that $\lambda_{max}(\Lambda_{12}) \ll \lambda_{min}(\Lambda_{11})$, then it is tempting to conclude that one can also discard those states corresponding to Λ_{12} without causing much error. However, this is not necessarily true as shown in the following example. Consider a stable transfer function
$$G(s) = \frac{3s+18}{s^2+3s+18}.$$
Then $G(s)$ has a state space realization given by
$$G(s) = \left[\begin{array}{cc|c} -1 & -4/\alpha & 1 \\ 4\alpha & -2 & 2\alpha \\ \hline -1 & 2/\alpha & 0 \end{array}\right]$$
where α is any nonzero number. It is easy to check that the controllability Gramian of the realization is given by
$$P = \begin{bmatrix} 0.5 & \\ & \alpha^2 \end{bmatrix}.$$
Since the last diagonal term of P can be made arbitrarily small by making α small, the controllability of the corresponding state can be made arbitrarily weak. If the state corresponding to the last diagonal term of P is removed, we get a transfer function
$$\hat{G} = \left[\begin{array}{c|c} -1 & 1 \\ \hline -1 & 0 \end{array}\right] = \frac{-1}{s+1}$$
which is not close to the original transfer function in any sense. The problem may be easily detected if one checks the observability Gramian Q, which is
$$Q = \begin{bmatrix} 0.5 & \\ & 1/\alpha^2 \end{bmatrix}.$$

Since $1/\alpha^2$ is very large if α is small, this shows that the state corresponding to the last diagonal term is strongly observable. This example shows that controllability (or observability) Gramian alone can not give an accurate indication of the dominance of the system states in the input/output behavior.

This motivates the introduction of a balanced realization which gives balanced Gramians for controllability and observability.

Suppose $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ is stable, i.e., A is stable. Let P and Q denote the controllability Gramian and observability Gramian, respectively. Then by Lemma 3.18, P and Q satisfy the following Lyapunov equations

$$AP + PA^* + BB^* = 0 \tag{3.9}$$

$$A^*Q + QA + C^*C = 0, \tag{3.10}$$

and $P \geq 0$, $Q \geq 0$. Furthermore, the pair (A, B) is controllable iff $P > 0$, and (C, A) is observable iff $Q > 0$.

Suppose the state is transformed by a nonsingular T to $\hat{x} = Tx$ to yield the realization

$$G = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] = \left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array}\right].$$

Then the Gramians are transformed to $\hat{P} = TPT^*$ and $\hat{Q} = (T^{-1})^*QT^{-1}$. Note that $\hat{P}\hat{Q} = TPQT^{-1}$, and therefore the eigenvalues of the product of the Gramians are invariant under state transformation.

Consider the similarity transformation T which gives the eigenvector decomposition

$$PQ = T^{-1}\Lambda T, \quad \Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n).$$

Then the columns of T^{-1} are eigenvectors of PQ corresponding to the eigenvalues $\{\lambda_i\}$. Later, it will be shown that PQ has a real diagonal Jordan form and that $\Lambda \geq 0$, which are consequences of $P \geq 0$ and $Q \geq 0$.

Although the eigenvectors are not unique, in the case of a minimal realization they can always be chosen such that

$$\hat{P} = TPT^* = \Sigma,$$

$$\hat{Q} = (T^{-1})^*QT^{-1} = \Sigma,$$

where $\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ and $\Sigma^2 = \Lambda$. This new realization with controllability and observability Gramians $\hat{P} = \hat{Q} = \Sigma$ will be referred to as a *balanced realization* (also called internally balanced realization). The decreasingly order numbers, $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n \geq 0$, are called the *Hankel singular values* of the system.

More generally, if a realization of a stable system is not minimal, then there is a transformation such that the controllability and observability Gramians for the transformed realization are diagonal and the controllable and observable subsystem is balanced. This is a consequence of the following matrix fact.

Theorem 3.22 *Let P and Q be two positive semidefinite matrices. Then there exists a nonsingular matrix T such that*

$$TPT^* = \begin{bmatrix} \Sigma_1 & & & \\ & \Sigma_2 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \quad (T^{-1})^* Q T^{-1} = \begin{bmatrix} \Sigma_1 & & & \\ & 0 & & \\ & & \Sigma_3 & \\ & & & 0 \end{bmatrix},$$

respectively, with Σ_1, Σ_2, Σ_3 diagonal and positive definite.

Proof. Since P is a positive semidefinite matrix, there exists a transformation T_1 such that

$$T_1 P T_1^* = \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix}.$$

Now let

$$(T_1^*)^{-1} Q T_1^{-1} = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^* & Q_{22} \end{bmatrix},$$

and there exists a unitary matrix U_1 such that

$$U_1 Q_{11} U_1^* = \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma_1 > 0.$$

Let

$$(T_2^*)^{-1} = \begin{bmatrix} U_1 & 0 \\ 0 & I \end{bmatrix},$$

and then

$$(T_2^*)^{-1}(T_1^*)^{-1} Q T_1^{-1} (T_2)^{-1} = \begin{bmatrix} \Sigma_1^2 & 0 & \hat{Q}_{121} \\ 0 & 0 & \hat{Q}_{122} \\ \hat{Q}_{121}^* & \hat{Q}_{122}^* & Q_{22} \end{bmatrix}.$$

But $Q \geq 0$ implies $\hat{Q}_{122} = 0$. So now let

$$(T_3^*)^{-1} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ -\hat{Q}_{121}^* \Sigma_1^{-2} & 0 & I \end{bmatrix},$$

giving

$$(T_3^*)^{-1}(T_2^*)^{-1}(T_1^*)^{-1} Q T_1^{-1} (T_2)^{-1} (T_3)^{-1} = \begin{bmatrix} \Sigma_1^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & Q_{22} - \hat{Q}_{121}^* \Sigma_1^{-2} \hat{Q}_{121} \end{bmatrix}.$$

Next find a unitary matrix U_2 such that

$$U_2 (Q_{22} - \hat{Q}_{121}^* \Sigma_1^{-2} \hat{Q}_{121}) U_2^* = \begin{bmatrix} \Sigma_3 & 0 \\ 0 & 0 \end{bmatrix}, \quad \Sigma_3 > 0.$$

Define

$$(T_4^*)^{-1} = \begin{bmatrix} \Sigma_1^{-1/2} & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & U_2 \end{bmatrix}$$

and let

$$T = T_4 T_3 T_2 T_1.$$

Then

$$TPT^* = \begin{bmatrix} \Sigma_1 & & & \\ & \Sigma_2 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \quad (T^*)^{-1}QT^{-1} = \begin{bmatrix} \Sigma_1 & & & \\ & 0 & & \\ & & \Sigma_3 & \\ & & & 0 \end{bmatrix}$$

with $\Sigma_2 = I$. □

Corollary 3.23 *The product of two positive semi-definite matrices is similar to a positive semi-definite matrix.*

Proof. Let P and Q be any positive semi-definite matrices. Then it is easy to see that with the transformation given above

$$TPQT^{-1} = \begin{bmatrix} \Sigma_1^2 & 0 \\ 0 & 0 \end{bmatrix}.$$

□

Corollary 3.24 *For any stable system* $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$, *there exists a nonsingular* T *such that* $G = \left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array}\right]$ *has controllability Gramian* P *and observability Gramian* Q *given by*

$$P = \begin{bmatrix} \Sigma_1 & & & \\ & \Sigma_2 & & \\ & & 0 & \\ & & & 0 \end{bmatrix}, \quad Q = \begin{bmatrix} \Sigma_1 & & & \\ & 0 & & \\ & & \Sigma_3 & \\ & & & 0 \end{bmatrix},$$

respectively, with Σ_1, Σ_2, Σ_3 *diagonal and positive definite.*

In the special case where $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ is a minimal realization, a balanced realization can be obtained through the following simplified procedure:

1. Compute the controllability and observability Gramians $P > 0, Q > 0$.
2. Find a matrix R such that $P = R^*R$.
3. Diagonalize RQR^* to get $RQR^* = U\Sigma^2U^*$.
4. Let $T^{-1} = R^*U\Sigma^{-1/2}$. Then $TPT^* = (T^*)^{-1}QT^{-1} = \Sigma$ and $\left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & D \end{array}\right]$ is balanced.

Assume that the Hankel singular values of the system is decreasingly ordered so that $\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n)$ and $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_n$ and suppose $\sigma_r \gg \sigma_{r+1}$ for some r then the balanced realization implies that those states corresponding to the singular values of $\sigma_{r+1}, \ldots, \sigma_n$ are less controllable and observable than those states corresponding to $\sigma_1, \ldots, \sigma_r$. Therefore, truncating those less controllable and observable states will not lose much information about the system. These statements will be made more concrete in Chapter 7 where error bounds will be derived for the truncation error.

Two other closely related realizations are called *input normal realization* with $P = I$ and $Q = \Sigma^2$, and *output normal realization* with $P = \Sigma^2$ and $Q = I$. Both realizations can be obtained easily from the balanced realization by a suitable scaling on the states.

3.10 Hidden Modes and Pole-Zero Cancelation

Another important issue associated with the realization theory is the problem of uncontrollable and/or unobservable unstable modes in the dynamical system. This problem is illustrated in the following example: Consider a series connection of two subsystems as shown in the following diagram

$$y \longleftarrow \boxed{\frac{1}{s-1}} \xleftarrow{\quad \xi \quad} \boxed{\frac{s-1}{s+1}} \xleftarrow{\quad u \quad}$$

The transfer function for this system,

$$g(s) = \frac{s-1}{s+1}\frac{1}{s-1} = \frac{1}{s+1},$$

is stable and has a first order minimal realization. On the other hand, let

$$\begin{aligned} x_1 &= y \\ x_2 &= u - \xi. \end{aligned}$$

Then a state space description for this dynamical system is given by

$$\begin{aligned} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} &= \begin{bmatrix} 1 & -1 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 1 \\ 2 \end{bmatrix} u \\ y &= \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \end{aligned}$$

and is a second order system. Moreover, it is easy to show that the unstable mode 1 is uncontrollable but observable. Hence, the output can be unbounded if the initial state $x_1(0)$ is not zero. We should also note that the above problem does not go away by changing the interconnection order:

$$y \longleftarrow \boxed{\frac{s-1}{s+1}} \xleftarrow{\quad \eta \quad} \boxed{\frac{1}{s-1}} \xleftarrow{\quad u \quad}$$

In the later case, the unstable mode 1 becomes controllable but unobservable. The unstable mode can still result in the internal signal η unbounded if the initial state $\eta(0)$ is not zero. Of course, there are fundamental differences between these two types of interconnections as far as control design is concerned. For instance, if the state is available for feedback control, then the latter interconnection can be stabilized while the former cannot be.

This example shows that we must be very careful in canceling unstable modes in the procedure of forming a transfer function in control designs; otherwise the results obtained may be misleading and those unstable modes become hidden modes waiting to blow. One observation from this example is that the problem is really caused by the unstable zero of the subsystem $\frac{s-1}{s+1}$. Although the zeros of an SISO transfer function are easy to see, it is not quite so for an MIMO transfer matrix. In fact, the notion of "system zero" cannot be generalized naturally from the scalar transfer function zeros. For example, consider the following transfer matrix

$$G(s) = \begin{bmatrix} \dfrac{1}{s+1} & \dfrac{1}{s+2} \\ \dfrac{2}{s+2} & \dfrac{1}{s+1} \end{bmatrix}$$

which is stable and each element of $G(s)$ has no finite zeros. Let

$$K = \begin{bmatrix} \dfrac{s+2}{s-\sqrt{2}} & -\dfrac{s+1}{s-\sqrt{2}} \\ 0 & 1 \end{bmatrix}$$

which is unstable. However,

$$KG = \begin{bmatrix} -\dfrac{s+\sqrt{2}}{(s+1)(s+2)} & 0 \\ \dfrac{2}{s+2} & \dfrac{1}{s+1} \end{bmatrix}$$

is stable. This implies that $G(s)$ must have an unstable zero at $\sqrt{2}$ that cancels the unstable pole of K. This leads us to the next topic: multivariable system poles and zeros.

3.11 Multivariable System Poles and Zeros

Let $\mathbb{R}[s]$ denote the polynomial ring with real coefficients. A matrix is called a polynomial matrix if every element of the matrix is in $\mathbb{R}[s]$. A square polynomial matrix is called a *unimodular matrix* if its determinant is a nonzero constant (not a polynomial of s). It is a fact that a square polynomial matrix is unimodular if and only if it is invertible in $\mathbb{R}[s]$, i.e., its inverse is also a polynomial matrix.

Definition 3.11 Let $Q(s) \in \mathbb{R}[s]$ be a $(p \times m)$ polynomial matrix. Then the *normal rank* of $Q(s)$, denoted *normalrank* $(Q(s))$, is the rank in $\mathbb{R}[s]$ or, equivalently, is the maximum dimension of a square submatrix of $Q(s)$ with nonzero determinant in $\mathbb{R}[s]$.

In short, sometimes we say that a polynomial matrix $Q(s)$ has rank$(Q(s))$ in $\mathbb{R}[s]$ when we refer to the normal rank of $Q(s)$.

To show the difference between the normal rank of a polynomial matrix and the rank of the polynomial matrix evaluated at certain point, consider

$$Q(s) = \begin{bmatrix} s & 1 \\ s^2 & 1 \end{bmatrix}.$$

Then $Q(s)$ has normal rank 2 since $\det Q(s) = s - s^2 \not\equiv 0$. However, $Q(0)$ has rank 1.

It is a fact in linear algebra that any polynomial matrix can be reduced to a so-called *Smith form* through some pre- and post- unimodular operations. [cf. Kailath, 1984, pp.391].

Lemma 3.25 (Smith form) *Let $P(s) \in \mathbb{R}[s]$ be any polynomial matrix, then there exist unimodular matrices $U(s), V(s) \in \mathbb{R}[s]$ such that*

$$U(s)P(s)V(s) = S(s) := \begin{bmatrix} \gamma_1(s) & 0 & \cdots & 0 & 0 \\ 0 & \gamma_2(s) & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \gamma_r(s) & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

and $\gamma_i(s)$ divides $\gamma_{i+1}(s)$.

$S(s)$ is called the *Smith form* of $P(s)$. It is also clear that r is the normal rank of $P(s)$. We shall illustrate the procedure of obtaining a Smith form by an example. Let

$$P(s) = \begin{bmatrix} s+1 & (s+1)(2s+1) & s(s+1) \\ s+2 & (s+2)(s^2+5s+3) & s(s+2) \\ 1 & 2s+1 & s \end{bmatrix}.$$

The polynomial matrix $P(s)$ has normal rank 2 since

$$\det(P(s)) \equiv 0, \quad \det \begin{bmatrix} s+1 & (s+1)(2s+1) \\ s+2 & (s+2)(s^2+5s+3) \end{bmatrix} = (s+1)^2(s+2)^2 \not\equiv 0.$$

First interchange the first row and the third row and use row elementary operation to zero out the $s+1$ and $s+2$ elements of $P(s)$. This process can be done by pre-multiplying a unimodular matrix U to $P(s)$:

$$U = \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & -(s+2) \\ 1 & 0 & -(s+1) \end{bmatrix}.$$

Then

$$P_1(s) := U(s)P(s) = \begin{bmatrix} 1 & 2s+1 & s \\ 0 & (s+1)(s+2)^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Next use column operation to zero out the $2s+1$ and s terms in P_1. This process can be done by post-multiplying a unimodular matrix V to $P_1(s)$:

$$V(s) = \begin{bmatrix} 1 & -(2s+1) & -s \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

and

$$P_1(s)V(s) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & (s+1)(s+2)^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then we have

$$S(s) = U(s)P(s)V(s) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & (s+1)(s+2)^2 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Similarly, let $\mathcal{R}_p(s)$ denote the set of rational proper transfer matrices.[1] Then any real rational transfer matrix can be reduced to a so-called *McMillan form* through some pre- and post- unimodular operations.

Lemma 3.26 (McMillan form) *Let $G(s) \in \mathcal{R}_p(s)$ be any proper real rational transfer matrix, then there exist unimodular matrices $U(s), V(s) \in \mathbb{R}[s]$ such that*

$$U(s)G(s)V(s) = M(s) := \begin{bmatrix} \frac{\alpha_1(s)}{\beta_1(s)} & 0 & \cdots & 0 & 0 \\ 0 & \frac{\alpha_2(s)}{\beta_2(s)} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & \frac{\alpha_r(s)}{\beta_r(s)} & 0 \\ 0 & 0 & \cdots & 0 & 0 \end{bmatrix}$$

and $\alpha_i(s)$ divides $\alpha_{i+1}(s)$ and $\beta_{i+1}(s)$ divides $\beta_i(s)$.

[1] It is obvious that a similar notion of normal rank can be extended to a matrix in $\mathcal{R}_p(s)$. In particular, let $G(s) \in \mathcal{R}_p(s)$ be a rational matrix and write $G(s) = N(s)/d(s)$ such that $d(s)$ is a polynomial and $N(s)$ is a polynomial matrix, then $G(s)$ is said to have normal rank r if $N(s)$ has normal rank r.

Proof. If we write the transfer matrix $G(s)$ as $G(s) = N(s)/d(s)$ such that $d(s)$ is a scalar polynomial and $N(s)$ is a $p \times m$ polynomial matrix and if let the Smith form of $N(s)$ be $S(s) = U(s)N(s)V(s)$, the conclusion follows by letting $M(s) = S(s)/d(s)$. $\Box$

Definition 3.12 The number $\sum_i \deg(\beta_i(s))$ is called the *McMillan degree* of $G(s)$ where $\deg(\beta_i(s))$ denotes the degree of the polynomial $\beta_i(s)$, i.e., the highest power of s in $\beta_i(s)$.

The McMillan degree of a transfer matrix is closely related to the dimension of a minimal realization of $G(s)$. In fact, it can be shown that the dimension of a minimal realization of $G(s)$ is exactly the McMillan degree of $G(s)$.

Definition 3.13 The roots of all the polynomials $\beta_i(s)$ in the McMillan form for $G(s)$ are called the *poles* of G.

Let (A, B, C, D) be a minimal realization of $G(s)$. Then it is fairly easy to show that a complex number is a pole of $G(s)$ if and only if it is an eigenvalue of A.

Definition 3.14 The roots of all the polynomials $\alpha_i(s)$ in the McMillan form for $G(s)$ are called the *transmission zeros* of $G(s)$. A complex number $z_0 \in \mathbb{C}$ is called a *blocking zero* of $G(s)$ if $G(z_0) = 0$.

It is clear that a blocking zero is a transmission zero. Moreover, for a scalar transfer function, the blocking zeros and the transmission zeros are the same.

We now illustrate the above concepts through an example. Consider a 3×3 transfer matrix:

$$G(s) = \begin{bmatrix} \dfrac{1}{(s+1)(s+2)} & \dfrac{2s+1}{(s+1)(s+2)} & \dfrac{s}{(s+1)(s+2)} \\ \dfrac{1}{(s+1)^2} & \dfrac{s^2+5s+3}{(s+1)^2} & \dfrac{s}{(s+1)^2} \\ \dfrac{1}{(s+1)^2(s+2)} & \dfrac{2s+1}{(s+1)^2(s+2)} & \dfrac{s}{(s+1)^2(s+2)} \end{bmatrix}.$$

Then $G(s)$ can be written as

$$G(s) = \frac{1}{(s+1)^2(s+2)} \begin{bmatrix} s+1 & (s+1)(2s+1) & s(s+1) \\ s+2 & (s+2)(s^2+5s+3) & s(s+2) \\ 1 & 2s+1 & s \end{bmatrix} := \frac{N(s)}{d(s)}.$$

Since $N(s)$ is exactly the same as the $P(s)$ in the previous example, it is clear that the $G(s)$ has the McMillan form

$$M(s) = U(s)G(s)V(s) = \begin{bmatrix} \frac{1}{(s+1)^2(s+2)} & 0 & 0 \\ 0 & \frac{s+2}{s+1} & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

and $G(s)$ has McMillan degree of 4. The poles of the transfer matrix are $\{-1, -1, -1, -2\}$ and the transmission zero is $\{-2\}$. Note that the transfer matrix has pole and zero at the same location $\{-2\}$; this is the unique feature of multivariable systems.

To get a minimal state space realization for $G(s)$, note that $G(s)$ has the following partial fractional expansion:

$$\begin{aligned} G(s) &= \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} + \frac{1}{s+1}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}\begin{bmatrix} 1 & -1 & -1 \end{bmatrix} + \frac{1}{s+1}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}\begin{bmatrix} 0 & 3 & 1 \end{bmatrix} \\ &+ \frac{1}{s+1}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}\begin{bmatrix} -1 & 3 & 2 \end{bmatrix} + \frac{1}{(s+1)^2}\begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & -1 \end{bmatrix} \\ &+ \frac{1}{s+2}\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & -3 & -2 \end{bmatrix} \end{aligned}$$

Since there are repeated poles at -1, the Gilbert's realization procedure described in the last section cannot be used directly. Nevertheless, a careful inspection of the fractional expansion results in a 4-th order minimal state space realization:

$$G(s) = \left[\begin{array}{cccc|ccc} -1 & 0 & 1 & 0 & 0 & 3 & 1 \\ 0 & -1 & 1 & 0 & -1 & 3 & 2 \\ 0 & 0 & -1 & 0 & 1 & -1 & -1 \\ 0 & 0 & 0 & -2 & 1 & -3 & -2 \\ \hline 0 & 0 & 1 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 \end{array}\right].$$

We remind readers that there are many different definitions of system zeros. The definitions introduced here are the most common ones and are useful in this book.

Lemma 3.27 *Let $G(s)$ be a $p \times m$ proper transfer matrix with full column normal rank. Then $z_0 \in \mathbb{C}$ is a transmission zero of $G(s)$ if and only if there exists a $0 \neq u_0 \in \mathbb{C}^m$ such that $G(z_0)u_0 = 0$.*

Proof. We shall outline a proof of this lemma. We shall first show that there is a vector $u_0 \in \mathbb{C}^m$ such that $G(z_0)u_0 = 0$ if $z_0 \in \mathbb{C}$ is a transmission zero. Without loss of generality, assume

$$G(s) = U_1(s) \begin{bmatrix} \frac{\alpha_1(s)}{\beta_1(s)} & 0 & \cdots & 0 \\ 0 & \frac{\alpha_2(s)}{\beta_2(s)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\alpha_m(s)}{\beta_m(s)} \\ 0 & 0 & \cdots & 0 \end{bmatrix} V_1(s)$$

for some unimodular matrices $U_1(s)$ and $V_1(s)$ and suppose z_0 is a zero of $\alpha_1(s)$, i.e., $\alpha_1(z_0) = 0$. Let

$$u_0 = V_1^{-1}(z_0)e_1 \neq 0$$

where $e_1 = [1, 0, 0, \ldots]^* \in \mathbb{R}^m$. Then it is easy to verify that $G(z_0)u_0 = 0$. On the other hand, suppose there is a $u_0 \in \mathbb{C}^m$ such that $G(z_0)u_0 = 0$. Then

$$U_1(z_0) \begin{bmatrix} \frac{\alpha_1(z_0)}{\beta_1(z_0)} & 0 & \cdots & 0 \\ 0 & \frac{\alpha_2(z_0)}{\beta_2(z_0)} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\alpha_m(z_0)}{\beta_m(z_0)} \\ 0 & 0 & \cdots & 0 \end{bmatrix} V_1(z_0)u_0 = 0.$$

Define

$$\begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_m \end{bmatrix} = V_1(z_0)u_0 \neq 0.$$

Then

$$\begin{bmatrix} \alpha_1(z_0)u_1 \\ \alpha_2(z_0)u_2 \\ \vdots \\ \alpha_m(z_0)u_m \end{bmatrix} = 0.$$

This implies that z_0 must be a root of one of polynomials $\alpha_i(s), i = 1, \ldots, m$. □

Note that the lemma may not be true if $G(s)$ does not have full column normal rank. This can be seen from the following example. Consider

$$G(s) = \frac{1}{s+1} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}, \quad u_0 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}.$$

It is easy to see that G has no transmission zero but $G(s)u_0 = 0$ for all s. It should also be noted that the above lemma applies even if z_0 is a pole of $G(s)$ although $G(z_0)$ is not defined. The reason is that $G(z_0)u_0$ may be well defined. For example,

$$G(s) = \begin{bmatrix} \frac{s-1}{s+1} & 0 \\ 0 & \frac{s+2}{s-1} \end{bmatrix}, \quad u_0 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Then $G(1)u_0 = 0$. Therefore, 1 is a transmission zero.

Similarly we have the following lemma:

Lemma 3.28 *Let $G(s)$ be a $p \times m$ proper transfer matrix with full row normal rank. Then $z_0 \in \mathbb{C}$ is a transmission zero of $G(s)$ if and only if there exists a $0 \neq \eta_0 \in \mathbb{C}^p$ such that $\eta_0^* G(z_0) = 0$.*

In the case where the transmission zero is not a pole of $G(s)$, we can give a useful alternative characterization of the transfer matrix transmission zeros. Furthermore, $G(s)$ is not required to be full column (or row) rank in this case.

The following lemma is easy to show from the definition of zeros.

Lemma 3.29 *Suppose $z_0 \in \mathbb{C}$ is not a pole of $G(s)$. Then z_0 is a transmission zero if and only if* $\operatorname{rank}(G(z_0)) < \operatorname{normalrank}(G(s))$.

Corollary 3.30 *Let $G(s)$ be a square $m \times m$ proper transfer matrix and* $\det G(s) \not\equiv 0$. *Suppose $z_o \in \mathbb{C}$ is not a pole of $G(s)$. Then $z_0 \in \mathbb{C}$ is a transmission zero of $G(s)$ if and only if* $\det G(z_0) = 0$.

Using the above corollary, we can confirm that the example in the last section does have a zero at $\sqrt{2}$ since

$$\det \begin{bmatrix} \dfrac{1}{s+1} & \dfrac{1}{s+2} \\ \dfrac{2}{s+2} & \dfrac{1}{s+1} \end{bmatrix} = \frac{2-s^2}{(s+1)^2(s+2)^2}.$$

Note that the above corollary may not be true if z_0 is a pole of G. For example,

$$G(s) = \begin{bmatrix} \frac{s-1}{s+1} & 0 \\ 0 & \frac{s+2}{s-1} \end{bmatrix}$$

has a zero at 1 which is not a zero of $\det G(s)$.

The poles and zeros of a transfer matrix can also be characterized in terms of its state space realizations. Let

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

be a state space realization of $G(s)$.

Definition 3.15 The eigenvalues of A are called the *poles* of the realization of $G(s)$.

To define zeros, let us consider the following system matrix

$$Q(s) = \begin{bmatrix} A - sI & B \\ C & D \end{bmatrix}.$$

Definition 3.16 A complex number $z_0 \in \mathbb{C}$ is called an *invariant zero* of the system realization if it satisfies

$$\operatorname{rank} \begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix} < normalrank \begin{bmatrix} A - sI & B \\ C & D \end{bmatrix}.$$

The invariant zeros are not changed by constant state feedback since

$$\operatorname{rank} \begin{bmatrix} A + BF - z_0 I & B \\ C + DF & D \end{bmatrix} = \operatorname{rank} \begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ F & I \end{bmatrix} = \operatorname{rank} \begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix}.$$

It is also clear that invariant zeros are not changed under similarity transformation.

The following lemma is obvious.

Lemma 3.31 *Suppose* $\begin{bmatrix} A - sI & B \\ C & D \end{bmatrix}$ *has full column normal rank. Then* $z_0 \in \mathbb{C}$ *is an invariant zero of a realization* (A, B, C, D) *if and only if there exist* $0 \neq x \in \mathbb{C}^n$ *and* $u \in \mathbb{C}^m$ *such that*

$$\begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = 0.$$

Moreover, if $u = 0$, *then* z_0 *is also a non-observable mode.*

Proof. By definition, z_0 is an invariant zero if there is a vector $\begin{bmatrix} x \\ u \end{bmatrix} \neq 0$ such that

$$\begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = 0$$

since $\begin{bmatrix} A - sI & B \\ C & D \end{bmatrix}$ has full column normal rank.

On the other hand, suppose z_0 is an invariant zero, then there is a vector $\begin{bmatrix} x \\ u \end{bmatrix} \neq 0$ such that

$$\begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix} \begin{bmatrix} x \\ u \end{bmatrix} = 0.$$

We claim that $x \neq 0$. Otherwise, $\begin{bmatrix} B \\ D \end{bmatrix} u = 0$ or $u = 0$ since $\begin{bmatrix} A - sI & B \\ C & D \end{bmatrix}$ has full column normal rank, i.e., $\begin{bmatrix} x \\ u \end{bmatrix} = 0$ which is a contradiction.

Finally, note that if $u = 0$, then

$$\begin{bmatrix} A - z_0 I \\ C \end{bmatrix} x = 0$$

and z_0 is a non-observable mode by PBH test. □

Lemma 3.32 *Suppose* $\begin{bmatrix} A - sI & B \\ C & D \end{bmatrix}$ *has full row normal rank. Then* $z_0 \in \mathbb{C}$ *is an invariant zero of a realization* (A, B, C, D) *if and only if there exist* $0 \neq y \in \mathbb{C}^n$ *and* $v \in \mathbb{C}^p$ *such that*

$$\begin{bmatrix} y^* & v^* \end{bmatrix} \begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix} = 0.$$

Moreover, if $v = 0$*, then* z_0 *is also a non-controllable mode.*

Lemma 3.33 $G(s)$ *has full column (row) normal rank if and only if* $\begin{bmatrix} A - sI & B \\ C & D \end{bmatrix}$ *has full column (row) normal rank.*

Proof. This follows by noting that

$$\begin{bmatrix} A - sI & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ C(A - sI)^{-1} & I \end{bmatrix} \begin{bmatrix} A - sI & B \\ 0 & G(s) \end{bmatrix}$$

and

$$\text{normalrank} \begin{bmatrix} A - sI & B \\ C & D \end{bmatrix} = n + \text{normalrank}(G(s)).$$

□

Theorem 3.34 *Let* $G(s)$ *be a real rational proper transfer matrix and let* (A, B, C, D) *be a corresponding minimal realization. Then a complex number* z_0 *is a transmission zero of* $G(s)$ *if and only if it is an invariant zero of the minimal realization.*

Proof. We will give a proof only for the case that the transmission zero is not a pole of $G(s)$. Then, of course, z_0 is not an eigenvalue of A since the realization is minimal. Note that

$$\begin{bmatrix} A - sI & B \\ C & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ C(A - sI)^{-1} & I \end{bmatrix} \begin{bmatrix} A - sI & B \\ 0 & G(s) \end{bmatrix}.$$

Since we have assumed that z_0 is not a pole of $G(s)$, we have

$$\text{rank} \begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix} = n + \text{rank } G(z_0).$$

Hence

$$\operatorname{rank} \begin{bmatrix} A - z_0 I & B \\ C & D \end{bmatrix} < \operatorname{normalrank} \begin{bmatrix} A - sI & B \\ C & D \end{bmatrix}$$

if and only if rank $G(z_0) <$ normalrank $G(s)$. Then the conclusion follows from Lemma 3.29. □

Note that the minimality assumption is essential for the converse statement. For example, consider a transfer matrix $G(s) = D$ (constant) and a realization of $G(s) = \left[\begin{array}{c|c} A & 0 \\ \hline C & D \end{array}\right]$ where A is any square matrix with any dimension and C is any matrix with compatible dimension. Then $G(s)$ has no poles or zeros but every eigenvalue of A is an invariant zero of the realization $\left[\begin{array}{c|c} A & 0 \\ \hline C & D \end{array}\right]$.

Nevertheless, we have the following corollary if a realization of the transfer matrix is not minimal.

Corollary 3.35 *Every transmission zero of a transfer matrix $G(s)$ is an invariant zero of all its realizations, and every pole of a transfer matrix $G(s)$ is a pole of all its realizations.*

Lemma 3.36 *Let $G(s) \in \mathcal{R}_p(s)$ be a $p \times m$ transfer matrix and let (A, B, C, D) be a minimal realization. If the system input is of the form $u(t) = u_0 e^{\lambda t}$, where $\lambda \in \mathbb{C}$ is not a pole of $G(s)$ and $u_0 \in \mathbb{C}^m$ is an arbitrary constant vector, then the output due to the input $u(t)$ and the initial state $x(0) = (\lambda I - A)^{-1} B u_0$ is $y(t) = G(\lambda) u_0 e^{\lambda t}$, $\forall t \geq 0$.*

Proof. The system response with respect to the input $u(t) = u_0 e^{\lambda t}$ and the initial condition $x(0) = (\lambda I - A)^{-1} B u_0$ is (in terms of Laplace transform):

$$\begin{aligned} Y(s) &= C(sI - A)^{-1} x(0) + C(sI - A)^{-1} B U(s) + D U(s) \\ &= C(sI - A)^{-1} x(0) + C(sI - A)^{-1} B u_0 (s - \lambda)^{-1} + D u_0 (s - \lambda)^{-1} \\ &= C(sI - A)^{-1} (x(0) - (\lambda I - A)^{-1} B u_0) + G(\lambda) u_0 (s - \lambda)^{-1} \\ &= G(\lambda) u_0 (s - \lambda)^{-1}. \end{aligned}$$

Hence $y(t) = G(\lambda) u_0 e^{\lambda t}$. □

Combining the above two lemmas, we have the following results that give a dynamical interpretation of a system's transmission zero.

Corollary 3.37 *Let $G(s) \in \mathcal{R}_p(s)$ be a $p \times m$ transfer matrix and let (A, B, C, D) be a minimal realization. Suppose that $z_0 \in \mathbb{C}$ is a transmission zero of $G(s)$ and is not a pole of $G(s)$. Then for any nonzero vector $u_0 \in \mathbb{C}^m$ the output of the system due to the initial state $x(0) = (z_0 I - A)^{-1} B u_0$ and the input $u = u_0 e^{z_0 t}$ is identically zero: $y(t) = G(z_0) u_0 e^{z_0 t} = 0$.*

The following lemma characterizes the relationship between zeros of a transfer function and poles of its inverse.

Lemma 3.38 *Suppose that* $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *is a square transfer matrix with* D *nonsingular, and suppose* z_0 *is not an eigenvalue of* A *(note that the realization is not necessarily minimal). Then there exists* x_0 *such that*

$$(A - BD^{-1}C)x_0 = z_0 x_0, \quad Cx_0 \neq 0$$

iff there exists $u_0 \neq 0$ *such that*

$$G(z_0)u_0 = 0.$$

Proof. $(\Leftarrow)$ $G(z_0)u_0 = 0$ implies that

$$G^{-1}(s) = \left[\begin{array}{c|c} A - BD^{-1}C & -BD^{-1} \\ \hline D^{-1}C & D^{-1} \end{array}\right]$$

has a pole at z_0 which is observable. Then, by definition, there exists x_0 such that

$$(A - BD^{-1}C)x_0 = z_0 x_0$$

and

$$Cx_0 \neq 0.$$

$(\Rightarrow)$ Set $u_0 = -D^{-1}Cx_0 \neq 0$. Then

$$(z_0 I - A)x_0 = -BD^{-1}Cx_0 = Bu_0.$$

Using this equality, one gets

$$G(z_0)u_0 = C(z_0 I - A)^{-1}Bu_0 + Du_0 = Cx_0 - Cx_0 = 0.$$

□

The above lemma implies that z_0 is a zero of an invertible $G(s)$ if and only if it is a pole of $G^{-1}(s)$.

3.12 Notes and References

Readers are referred to Brogan [1991], Chen [1984], Kailath [1980], and Wonham [1985] for the extensive treatment of the standard linear system theory. The balanced realization was first introduced by Mullis and Roberts [1976] to study the roundoff noise in digital filters. Moore [1981] proposed the balanced truncation method for model reduction which will be considered in Chapter 7.

4

Performance Specifications

The most important objective of a control system is to achieve certain performance specifications in addition to providing the internal stability. One way to describe the performance specifications of a control system is in terms of the size of certain signals of interest. For example, the performance of a tracking system could be measured by the size of the tracking error signal. In this chapter, we look at several ways of defining a signal's size, i.e., at several norms for signals. Of course, which norm is most appropriate depends on the situation at hand. For that purpose, we shall first introduce some normed spaces and some basic notions of linear operator theory, in particular, the Hardy spaces $\mathcal{H}_2$ and $\mathcal{H}_\infty$ are introduced. We then consider the performance of a system under various input signals and derive the worst possible outputs with the class of input signals under consideration. We show that $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms come out naturally as measures of the worst possible performance for many classes of input signals. Some state space methods of computing real rational $\mathcal{H}_2$ and $\mathcal{H}_\infty$ transfer matrix norms are also presented.

4.1 Normed Spaces

Let V be a vector space over $\mathbb{C}$ (or $\mathbb{R}$) and let $\|\cdot\|$ be a norm defined on V. Then V is a *normed space*. For example, the vector space $\mathbb{C}^n$ with any vector p-norm, $\|\cdot\|_p$, for $1 \leq p \leq \infty$, is a normed space. As another example, consider the linear vector space $C[a,b]$ of all bounded continuous functions on the real interval $[a,b]$. Then $C[a,b]$

becomes a normed space if a *supremum norm* is defined on the space:

$$\|f\|_\infty := \sup_{t\in[a,b]} |f(t)|.$$

A sequence $\{x_n\}$ in a normed space V is called a *Cauchy sequence* if $\|x_n - x_m\| \to 0$ as $n, m \to \infty$. A sequence $\{x_n\}$ is said to *converge* to $x \in V$, written $x_n \to x$, if $\|x_n - x\| \to 0$. A normed space V is said to be *complete* if every Cauchy sequence in V converges in V. A complete normed space is called a *Banach space*. For example, $\mathbb{R}^n$ and $\mathbb{C}^n$ with the usual spatial p-norm, $\|\cdot\|_p$ for $1 \le p \le \infty$, are Banach spaces. (One should not be confused with the notation $\|\cdot\|_p$ used here and the same notation used below for function spaces because usually the context will make the meaning clear.)

The following are some more examples of Banach spaces:

$l_p[0,\infty)$ spaces for $1 \le p < \infty$:

For each $1 \le p < \infty$, $l_p[0,\infty)$ consists of all sequences $x = (x_0, x_1, \ldots)$ such that $\sum_{i=0}^{\infty} |x_i|^p < \infty$. The associated norm is defined as

$$\|x\|_p := \left(\sum_{i=0}^{\infty} |x_i|^p\right)^{1/p}.$$

$l_\infty[0,\infty)$ space:

$l_\infty[0,\infty)$ consists of all bounded sequences $x = (x_0, x_1, \ldots)$, and the l_∞ norm is defined as

$$\|x\|_\infty := \sup_i |x_i|.$$

$\mathcal{L}_p(I)$ spaces for $1 \le p \le \infty$:

For each $1 \le p < \infty$, $\mathcal{L}_p(I)$ consists of all Lebesgue measurable functions $x(t)$ defined on an interval $I \subset \mathbb{R}$ such that

$$\|x\|_p := \left(\int_I |x(t)|^p dt\right)^{1/p} < \infty, \text{ for } \ 1 \le p < \infty$$

and

$$\|x(t)\|_\infty := \operatorname{ess}\sup_{t\in I} |x(t)|.$$

Some of these spaces, for example, $\mathcal{L}_2(-\infty, 0]$, $\mathcal{L}_2[0,\infty)$ and $\mathcal{L}_2(-\infty,\infty)$, will be discussed in more detail later on.

$C[a,b]$ space:

$C[a,b]$ consists of all continuous functions on the real interval $[a,b]$ with the norm defined as

$$\|x\|_\infty := \sup_{t\in[a,b]} |x(t)|.$$

Note that if each component or function is itself a vector or matrix, then the corresponding Banach space can also be formed by replacing the absolute value $|\cdot|$ of each component or function with its spatially normed component or function. For example, consider a vector space with all sequences in the form of

$$x = (x_0, x_1, \ldots)$$

where each component x_i is a $k \times m$ matrix and each element of x_i is bounded. Then x_i is bounded in any matrix norm, and the vector space becomes a Banach space if the following norm is defined

$$\|x\|_\infty := \sup_i \phi_i(x_i)$$

where $\phi_i(x_i) := \|x_i\|$ is any matrix norm. This space will also be denoted by l_∞.

Let V_1 and V_2 be two vector spaces and let T be an operator from $S \subset V_1$ into V_2. An operator T is said to be *linear* if for any $x_1, x_2 \in S$ and scalars $\alpha, \beta \in \mathbb{C}$, the following holds:

$$T(\alpha x_1 + \beta x_2) = \alpha(Tx_1) + \beta(Tx_2).$$

Moreover, let V_0 be a linear subspace in V_1. Then the operator $T_0 : V_0 \longmapsto V_2$ defined by $T_0 x = Tx$ for every $x \in V_0$ is called the *restriction* of T to V_0 and is denoted as $T|_{V_0} = T_0$. On the other hand, a linear operator $T : V_1 \longmapsto V_2$ coinciding with T_0 on $V_0 \subset V_1$ is called an *extension* of T_0. For example, let $V_0 := \left\{ \begin{bmatrix} x_1 \\ 0 \end{bmatrix} : x_1 \in \mathbb{C}^n \right\} \subset V_1 = \mathbb{C}^{n+m}$, and let

$$T = \begin{bmatrix} A_1 & A_2 \\ 0 & A_3 \end{bmatrix} \in \mathbb{C}^{(n+m)\times(n+m)}, \quad T_0 = \begin{bmatrix} A_1 & 0 \\ 0 & 0 \end{bmatrix} \in \mathbb{C}^{(n+m)\times(n+m)}.$$

Then $T|_{V_0} = T_0$.

Definition 4.1 Two normed spaces V_1 and V_2 are said to be *linearly isometric*, denoted by $V_1 \cong V_2$, if there exists a linear operator T of V_1 onto V_2 such that

$$\|Tx\| = \|x\|$$

for all x in V_1. In this case, the mapping T is said to be an *isometric isomorphism.*

4.2 Hilbert Spaces

Recall the inner product of vectors defined on a Euclidean space $\mathbb{C}^n$:

$$\langle x, y \rangle := x^* y = \sum_{i=1}^{n} \bar{x}_i y_i \quad \forall x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix},\ y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \in \mathbb{C}^n.$$

Note that many important metric notions and geometrical properties such as length, distance, angle, and the energy of physical systems can be deduced from this inner product. For instance, the length of a vector $x \in \mathbb{C}^n$ is defined as

$$\|x\| := \sqrt{\langle x, x\rangle},$$

and the angle between two vectors $x, y \in \mathbb{C}^n$ can be computed from

$$\cos\angle(x,y) = \frac{\langle x, y\rangle}{\|x\|\,\|y\|}, \quad \angle(x,y) \in [0, \pi].$$

The two vectors are said to be *orthogonal* if $\angle(x,y) = \frac{\pi}{2}$.

We now consider a natural generalization of the inner product on $\mathbb{C}^n$ to more general (possibly infinite dimensional) vector spaces.

Definition 4.2 Let V be a vector space over $\mathbb{C}$. An *inner product* on V is a complex valued function,

$$\langle \cdot, \cdot \rangle : V \times V \longmapsto \mathbb{C}$$

such that for any $x, y, z \in V$ and $\alpha, \beta \in \mathbb{C}$

(i) [1] $\langle x, \alpha y + \beta z\rangle = \alpha\langle x, y\rangle + \beta\langle x, z\rangle$

(ii) $\langle x, y\rangle = \overline{\langle y, x\rangle}$

(iii) $\langle x, x\rangle > 0$ if $x \neq 0$.

A vector space V with an inner product is called an *inner product space*.

It is clear that the inner product defined above induces a norm $\|x\| := \sqrt{\langle x, x\rangle}$, so that the norm conditions in Chapter 2 are satisfied. In particular, the distance between vectors x and y is $d(x,y) = \|x - y\|$.

Two vectors x and y in an inner product space V are said to be *orthogonal* if $\langle x, y\rangle = 0$, denoted $x \perp y$. More generally, a vector x is said to be orthogonal to a set $S \subset V$, denoted by $x \perp S$, if $x \perp y$ for all $y \in S$.

The inner product and the inner-product induced norm have the following familiar properties.

Theorem 4.1 *Let V be an inner product space and let $x, y \in V$. Then*

(i) $|\langle x, y\rangle| \leq \|x\|\,\|y\|$ *(Cauchy-Schwarz inequality). Moreover, the equality holds if and only if $x = \alpha y$ for some constant α or $y = 0$.*

(ii) $\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2$ *(Parallelogram law) .*

(iii) $\|x+y\|^2 = \|x\|^2 + \|y\|^2$ *if $x \perp y$.*

[1] This is the other way round to the usual mathematical convention since we want to have $\langle x, y\rangle = x^* y$ rather than $y^* x$ for $x, y \in \mathbb{C}^n$.

A *Hilbert space* is a complete inner product space with the norm induced by its inner product. Obviously, a Hilbert space is also a Banach space. For example, $\mathbb{C}^n$ with the usual inner product is a (finite dimensional) Hilbert space. More generally, it is straightforward to verify that $\mathbb{C}^{n\times m}$ with the inner product defined as

$$\langle A, B\rangle := \operatorname{Trace} A^*B = \sum_{i=1}^{n}\sum_{j=1}^{m} \bar{a}_{ij} b_{ij} \quad \forall A, B \in \mathbb{C}^{n\times m}$$

is also a (finite dimensional) Hilbert space.

Here are some examples of infinite dimensional Hilbert spaces:

$l_2(-\infty,\infty)$:

$l_2(-\infty,\infty)$ consists of the set of all real or complex square summable sequences

$$x = (\ldots, x_{-2}, x_{-1}, x_0, x_1, x_2, \ldots)$$

i.e,

$$\sum_{i=-\infty}^{\infty} |x_i|^2 < \infty$$

with the inner product defined as

$$\langle x, y\rangle := \sum_{i=-\infty}^{\infty} \bar{x}_i y_i$$

for $x, y \in l_2(-\infty,\infty)$. The subspaces $l_2(-\infty, 0)$ and $l_2[0,\infty)$ of $l_2(-\infty,\infty)$ are defined similarly and consist of sequences of the form $x = (\ldots, x_{-2}, x_{-1})$ and $x = (x_0, x_1, x_2, \ldots)$, respectively.

Note that we can also define a corresponding Hilbert space even if each component x_i is a vector or a matrix; in fact, the following inner product will suffice:

$$\langle x, y\rangle := \sum_{i=-\infty}^{\infty} \operatorname{Trace}(x_i^* y_i).$$

$\mathcal{L}_2(I)$ for $I \subset \mathbb{R}$:

$\mathcal{L}_2(I)$ consists of all square integrable and Lebesgue measurable functions defined on an interval $I \subset \mathbb{R}$ with the inner product defined as

$$\langle f, g\rangle := \int_I f(t)^* g(t) dt$$

for $f, g \in \mathcal{L}_2(I)$. Similarly, if the function is vector or matrix valued, the inner product is defined correspondingly as

$$\langle f, g\rangle := \int_I \operatorname{Trace}\left[f(t)^* g(t)\right] dt.$$

Some very often used spaces in this book are $\mathcal{L}_2[0,\infty), \mathcal{L}_2(-\infty,0], \mathcal{L}_2(-\infty,\infty)$. More precisely, they are defined as

$\mathcal{L}_2 = \mathcal{L}_2(-\infty,\infty)$: Hilbert space of matrix-valued functions on $\mathbb{R}$, with inner product

$$\langle f, g\rangle := \int_{-\infty}^{\infty} \operatorname{Trace}\left[f(t)^* g(t)\right] dt.$$

$\mathcal{L}_{2+} = \mathcal{L}_2[0,\infty)$: subspace of $\mathcal{L}_2(-\infty,\infty)$ with functions zero for $t < 0$.

$\mathcal{L}_{2-} = \mathcal{L}_2(-\infty,0]$: subspace of $\mathcal{L}_2(-\infty,\infty)$ with functions zero for $t > 0$.

Let $\mathcal{H}$ be a Hilbert space and $M \subset \mathcal{H}$ a subset. Then the *orthogonal complement* of M, denoted by $\mathcal{H} \ominus M$ or $M^\perp$, is defined as

$$M^\perp = \{x : \langle x, y\rangle = 0, \forall y \in M, x \in \mathcal{H}\}.$$

It can be shown that $M^\perp$ is closed (hence $M^\perp$ is also a Hilbert space). For example, let $M = \mathcal{L}_{2+} \subset \mathcal{L}_2$, then $M^\perp = \mathcal{L}_{2-}$ is a Hilbert space.

Let M and N be subspaces of a vector space V. V is said to be the *direct sum of M and N*, written $V = M \oplus N$, if $M \cap N = \{0\}$, and every element $v \in V$ can be expressed as $v = x + y$ with $x \in M$ and $y \in N$. If V is an inner product space and M and N are orthogonal, then V is said to be the *orthogonal direct sum of M and N*. As an example, it is easy to see that $\mathcal{L}_2$ is the orthogonal direct sum of $\mathcal{L}_{2-}$ and $\mathcal{L}_{2+}$. Similarly, $l_2(-\infty,\infty)$ is the orthogonal direct sum of $l_2(-\infty,0)$ and $l_2[0,\infty)$.

The following is a version of the so-called *orthogonal projection theorem*:

Theorem 4.2 *Let $\mathcal{H}$ be a Hilbert space, and let M be a closed subspace of $\mathcal{H}$. Then for each vector $v \in \mathcal{H}$, there exist unique vectors $x \in M$ and $y \in M^\perp$ such that $v = x + y$, i.e., $\mathcal{H} = M \oplus M^\perp$. Moreover, $x \in M$ is the unique vector such that $d(v, M) = \|v - x\|$.*

Let $\mathcal{H}_1$ and $\mathcal{H}_2$ be two Hilbert spaces, and let A be a bounded linear operator from $\mathcal{H}_1$ into $\mathcal{H}_2$. Then there exists a unique linear operator $A^* : \mathcal{H}_2 \longmapsto \mathcal{H}_1$ such that for all $x \in \mathcal{H}_1$ and $y \in \mathcal{H}_2$

$$\langle Ax, y\rangle = \langle x, A^* y\rangle.$$

A^* is called the *adjoint* of A. Furthermore, A is called *self-adjoint* if $A = A^*$.

Let $\mathcal{H}$ be a Hilbert space and $M \subset \mathcal{H}$ be a closed subspace. A bounded operator P mapping from $\mathcal{H}$ into itself is called the *orthogonal projection* onto M if

$$P(x + y) = x, \forall\, x \in M \quad \text{and} \quad y \in M^\perp.$$

4.3 Hardy Spaces $\mathcal{H}_2$ and $\mathcal{H}_\infty$

Let $S \subset \mathbb{C}$ be an open set, and let $f(s)$ be a complex valued function defined on S:

$$f(s) : S \longmapsto \mathbb{C}.$$

Then $f(s)$ is said to be *analytic at a point* z_0 in S if it is differentiable at z_0 and also at each point in some neighborhood of z_0. It is a fact that if $f(s)$ is analytic at z_0 then f has continuous derivatives of all orders at z_0. Hence, a function analytic at z_0 has a power series representation at z_0. The converse is also true, i.e., if a function has a power series at z_0, then it is analytic at z_0. A function $f(s)$ is said to be *analytic in* S if it has a derivative or is analytic at each point of S. A matrix valued function is analytic in S if every element of the matrix is analytic in S. For example, all real rational stable transfer matrices are analytic in the right-half plane and e^{-s} is analytic everywhere.

A well known property of the analytic functions is the so-called *Maximum Modulus Theorem.*

Theorem 4.3 *If $f(s)$ is defined and continuous on a closed-bounded set S and analytic on the interior of S, then the maximum of $|f(s)|$ on S is attained on the boundary of S, i.e.,*

$$\max_{s\in S} |f(s)| = \max_{s \in \partial S} |f(s)|$$

where ∂S denotes the boundary of S.

Next we consider some frequently used complex (matrix) function spaces.

$\mathcal{L}_2(j\mathbb{R})$ Space

$\mathcal{L}_2(j\mathbb{R})$ or simply $\mathcal{L}_2$ is a Hilbert space of matrix-valued (or scalar-valued) functions on $j\mathbb{R}$ and consists of all complex matrix functions F such that the integral below is bounded, i.e.,

$$\int_{-\infty}^{\infty} \operatorname{Trace}\left[F^*(j\omega)F(j\omega)\right] d\omega < \infty.$$

The inner product for this Hilbert space is defined as

$$\langle F, G\rangle := \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{Trace}\left[F^*(j\omega)G(j\omega)\right] d\omega$$

for $F, G \in \mathcal{L}_2$, and the inner product induced norm is given by

$$\|F\|_2 := \sqrt{\langle F, F\rangle}.$$

For example, all real rational strictly proper transfer matrices with no poles on the imaginary axis form a subspace (not closed) of $\mathcal{L}_2(j\mathbb{R})$ which is denoted by $\mathcal{RL}_2(j\mathbb{R})$ or simply $\mathcal{RL}_2$.

$\mathcal{H}_2$ Space[2]

$\mathcal{H}_2$ is a (closed) subspace of $\mathcal{L}_2(j\mathbb{R})$ with matrix functions $F(s)$ analytic in $Re(s) > 0$ (open right-half plane). The corresponding norm is defined as

$$\|F\|_2^2 := \sup_{\sigma>0}\left\{\frac{1}{2\pi}\int_{-\infty}^{\infty}\operatorname{Trace}\left[F^*(\sigma+j\omega)F(\sigma+j\omega)\right]d\omega\right\}.$$

It can be shown[3] that

$$\|F\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}\operatorname{Trace}\left[F^*(j\omega)F(j\omega)\right]d\omega.$$

Hence, we can compute the norm for $\mathcal{H}_2$ just as we do for $\mathcal{L}_2$. The real rational subspace of $\mathcal{H}_2$, which consists of all strictly proper and real rational stable transfer matrices, is denoted by $\mathcal{RH}_2$.

$\mathcal{H}_2^\perp$ Space

$\mathcal{H}_2^\perp$ is the orthogonal complement of $\mathcal{H}_2$ in $\mathcal{L}_2$, i.e., the (closed) subspace of functions in $\mathcal{L}_2$ that are analytic in the open left-half plane. The real rational subspace of $\mathcal{H}_2^\perp$, which consists of all strictly proper rational transfer matrices with all poles in the open right half plane, will be denoted by $\mathcal{RH}_2^\perp$. It is easy to see that if G is a strictly proper, stable, and real rational transfer matrix, then $G \in \mathcal{H}_2$ and $G^\sim \in \mathcal{H}_2^\perp$. Most of our study in this book will be focused on the real rational case.

The $\mathcal{L}_2$ spaces defined above in the frequency domain can be related to the $\mathcal{L}_2$ spaces defined in the time domain. Recall the fact that a function in $\mathcal{L}_2$ space in the time domain admits a bilateral Laplace (or Fourier) transform. In fact, it can be shown that this bilateral Laplace (or Fourier) transform yields an isometric isomorphism between the $\mathcal{L}_2$ spaces in the time domain and the $\mathcal{L}_2$ spaces in the frequency domain (this is what is called *Parseval's relations* or *Plancherel Theorem* in complex analysis):

$$\mathcal{L}_2(-\infty,\infty) \cong \mathcal{L}_2(j\mathbb{R})$$

$$\mathcal{L}_2[0,\infty) \cong \mathcal{H}_2$$

$$\mathcal{L}_2(-\infty,0] \cong \mathcal{H}_2^\perp.$$

As a result, if $g(t) \in \mathcal{L}_2(-\infty,\infty)$ and if its Fourier (or bilateral Laplace) transform is $G(j\omega) \in \mathcal{L}_2(j\mathbb{R})$, then

$$\|G\|_2 = \|g\|_2.$$

---

[2]The $\mathcal{H}_2$ space and $\mathcal{H}_\infty$ space defined below together with the $\mathcal{H}_p$ spaces, $p \geq 1$, which will not be introduced in this book, are usually called Hardy spaces named after the mathematician G. H. Hardy (hence the notation of $\mathcal{H}$).

[3]See Francis [1987].

Hence, whenever there is no confusion, the notation of functions in the time domain and in the frequency domain will be used interchangeably.

Define an orthogonal projection

$$P_+ : \mathcal{L}_2(-\infty, \infty) \longmapsto \mathcal{L}_2[0, \infty)$$

such that, for any function $f(t) \in \mathcal{L}_2(-\infty, \infty)$, we have $g(t) = P_+ f(t)$ with

$$g(t) := \begin{cases} f(t), & \text{for } t \geq 0; \\ 0, & \text{for } t < 0. \end{cases}$$

In this book, P_+ will also be used to denote the projection from $\mathcal{L}_2(j\mathbb{R})$ onto $\mathcal{H}_2$. Similarly, define P_- as another orthogonal projection from $\mathcal{L}_2(-\infty, \infty)$ onto $\mathcal{L}_2(-\infty, 0]$ (or $\mathcal{L}_2(j\mathbb{R})$ onto $\mathcal{H}_2^\perp$). Then the relationships between $\mathcal{L}_2$ spaces and $\mathcal{H}_2$ spaces can be shown as in Figure 4.1.

$$\begin{tikzcd}
\mathcal{L}_2[0,\infty) \arrow[rr, shift left, "\text{Laplace Transform}"] & & \mathcal{H}_2 \arrow[ll, shift left, "\text{Inverse Transform}"] \\
\mathcal{L}_2(-\infty,\infty) \arrow[u, "P_+"] \arrow[d, "P_-"'] \arrow[rr, shift left, "\text{Laplace Transform}"] & & \mathcal{L}_2(j\mathbb{R}) \arrow[u, "P_+"'] \arrow[d, "P_-"] \arrow[ll, shift left, "\text{Inverse Transform}"] \\
\mathcal{L}_2(-\infty,0] \arrow[rr, shift left, "\text{Laplace Transform}"] & & \mathcal{H}_2^\perp \arrow[ll, shift left, "\text{Inverse Transform}"]
\end{tikzcd}$$

Figure 4.1: Relationships among function spaces

Other classes of important complex matrix functions used in this book are those bounded on the imaginary axis.

$\mathcal{L}_\infty(j\mathbb{R})$ Space

$\mathcal{L}_\infty(j\mathbb{R})$ or simply $\mathcal{L}_\infty$ is a Banach space of matrix-valued (or scalar-valued) functions that are (essentially) bounded on $j\mathbb{R}$, with norm

$$\|F\|_\infty := \operatorname*{ess\,sup}_{\omega \in \mathbb{R}} \overline{\sigma}\left[F(j\omega)\right].$$

The rational subspace of $\mathcal{L}_\infty$, denoted by $\mathcal{RL}_\infty(j\mathbb{R})$ or simply $\mathcal{RL}_\infty$, consists of all rational proper transfer matrices with no poles on the imaginary axis.

$\mathcal{H}_\infty$ Space

$\mathcal{H}_\infty$ is a (closed) subspace in $\mathcal{L}_\infty$ with functions that are analytic in the open right-half plane and bounded on the imaginary axis. The $\mathcal{H}_\infty$ norm is defined as

$$\|F\|_\infty := \sup_{\mathrm{Re}(s)>0} \overline{\sigma}\left[F(s)\right] = \operatorname{ess}\sup_{\omega\in\mathbb{R}} \overline{\sigma}\left[F(j\omega)\right].$$

The second equality can be regarded as a generalization of the maximum modulus theorem for matrix functions. The real rational subspace of $\mathcal{H}_\infty$ is denoted by $\mathcal{RH}_\infty$ which consists of all proper and real rational stable transfer matrices.

$\mathcal{H}_\infty^-$ Space

$\mathcal{H}_\infty^-$ is a (closed) subspace in $\mathcal{L}_\infty$ with functions that are analytic in the open left-half plane and bounded on the imaginary axis. The $\mathcal{H}_\infty^-$ norm is defined as

$$\|F\|_\infty := \sup_{\mathrm{Re}(s)<0} \overline{\sigma}\left[F(s)\right] = \operatorname{ess}\sup_{\omega\in\mathbb{R}} \overline{\sigma}\left[F(j\omega)\right].$$

The real rational subspace of $\mathcal{H}_\infty^-$ is denoted by $\mathcal{RH}_\infty^-$ which consists of all proper rational transfer matrices with all poles in the open right half plane.

Definition 4.3 A transfer matrix $G(s) \in \mathcal{H}_\infty^-$ is usually said to be *antistable* or *anti-causal*.

Some facts about $\mathcal{L}_\infty$ and $\mathcal{H}_\infty$ functions are worth mentioning:

(i) if $G(s) \in \mathcal{L}_\infty$, then $G(s)\mathcal{L}_2 := \{G(s)f(s) : f(s) \in \mathcal{L}_2\} \subset \mathcal{L}_2$.

(ii) if $G(s) \in \mathcal{H}_\infty$, then $G(s)\mathcal{H}_2 := \{G(s)f(s) : f(s) \in \mathcal{H}_2\} \subset \mathcal{H}_2$.

(ii) if $G(s) \in \mathcal{H}_\infty^-$, then $G(s)\mathcal{H}_2^\perp := \{G(s)f(s) : f(s) \in \mathcal{H}_2^\perp\} \subset \mathcal{H}_2^\perp$.

Remark 4.1 The notation for $\mathcal{L}_\infty$ is somewhat unfortunate; it should be clear to the reader that the $\mathcal{L}_\infty$ space in the time domain and in the frequency domain denote completely different spaces. The $\mathcal{L}_\infty$ space in the time domain is usually used to denote signals, while the $\mathcal{L}_\infty$ space in the frequency domain is usually used to denote transfer functions and operators. ♡

Let $G(s) \in \mathcal{L}_\infty$ be a $p \times q$ transfer matrix. Then a multiplication operator is defined as

$$M_G : \mathcal{L}_2 \longmapsto \mathcal{L}_2$$

$$M_G f := Gf.$$

In writing the above mapping, we have assumed that f has a compatible dimension. A more accurate description of the above operator should be

$$M_G : \mathcal{L}_2^q \longmapsto \mathcal{L}_2^p$$

i.e., f is a q-dimensional vector function with each component in $\mathcal{L}_2$. However, we shall suppress all dimensions in this book and assume all objects have compatible dimensions.

It is easy to verify that the adjoint operator $M_G^* = M_{G^\sim}$.

A useful fact about the multiplication operator is that the norm of a matrix G in $\mathcal{L}_\infty$ equals the norm of the corresponding multiplication operator.

Theorem 4.4 *Let $G \in \mathcal{L}_\infty$ be a $p \times q$ transfer matrix. Then $\|M_G\| = \|G\|_\infty$.*

Remark 4.2 It is also true that this operator norm equals the norm of the operator restricted to $\mathcal{H}_2$ (or $\mathcal{H}_2^\perp$), i.e.,

$$\|M_G\| = \|M_G|_{\mathcal{H}_2}\| := \sup\{\|Gf\|_2 : \ f \in \mathcal{H}_2, \|f\|_2 \le 1\}.$$

This will be clear in the proof where an $f \in \mathcal{H}_2$ is constructed. ♡

Proof. By definition, we have

$$\|M_G\| = \sup\{\|Gf\|_2 : \ f \in \mathcal{L}_2, \|f\|_2 \le 1\}.$$

First we see that $\|G\|_\infty$ is an upper bound for the operator norm:

$$\begin{aligned} \|Gf\|_2^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} f^*(j\omega)G^*(j\omega)G(j\omega)f(j\omega)\ d\omega \\ &\le \|G\|_\infty^2 \frac{1}{2\pi}\int_{-\infty}^{\infty} \|f(j\omega)\|^2\ d\omega \\ &= \|G\|_\infty^2 \|f\|_2^2. \end{aligned}$$

To show that $\|G\|_\infty$ is the least upper bound, first choose a frequency ω_o where $\overline{\sigma}\,[G(j\omega)]$ is maximum, i.e.,

$$\overline{\sigma}\,[G(j\omega_0)] = \|G\|_\infty$$

and denote the singular value decomposition of $G(j\omega_0)$ by

$$G(j\omega_0) = \overline{\sigma} u_1(j\omega_0)v_1^*(j\omega_0) + \sum_{i=2}^{r} \sigma_i u_i(j\omega_0)v_i^*(j\omega_0)$$

where r is the rank of $G(j\omega_0)$ and u_i, v_i have unit length.

If $\omega_0 < \infty$, write $v_1(j\omega_0)$ as

$$v_1(j\omega_0) = \begin{bmatrix} \alpha_1 e^{j\theta_1} \\ \alpha_2 e^{j\theta_2} \\ \vdots \\ \alpha_q e^{j\theta_q} \end{bmatrix}$$

where $\alpha_i \in \mathbb{R}$ is such that $\theta_i \in [-\pi, 0)$ and q is the column dimension of G. Now let $\beta_i \geq 0$ be such that

$$\theta_i = \angle\left(\frac{\beta_i - j\omega_0}{\beta_i + j\omega_0}\right),$$

and let f be given by

$$f(s) = \begin{bmatrix} \alpha_1 \frac{\beta_1 - s}{\beta_1 + s} \\ \alpha_2 \frac{\beta_2 - s}{\beta_2 + s} \\ \vdots \\ \alpha_q \frac{\beta_q - s}{\beta_q + s} \end{bmatrix} \hat{f}(s)$$

where a scalar function $\hat{f}$ is chosen so that

$$|\hat{f}(j\omega)| = \begin{cases} c & \text{if } |\omega - \omega_o| < \epsilon \text{ or } |\omega + \omega_o| < \epsilon \\ 0 & \text{otherwise} \end{cases}$$

where ϵ is a small positive number and c is chosen so that $\hat{f}$ has unit 2-norm, i.e., $c = \sqrt{\pi/2\epsilon}$. This in turn implies that f has unit 2-norm. Then

$$\begin{aligned} \|Gf\|_2^2 &\approx \frac{1}{2\pi}\left[\overline{\sigma}\left[G(-j\omega_o)\right]^2 \pi + \overline{\sigma}\left[G(j\omega_o)\right]^2 \pi\right] \\ &= \overline{\sigma}\left[G(j\omega_o)\right]^2 = \|G\|_\infty^2. \end{aligned}$$

If $\omega_0 = \infty$, then $G(j\omega_0)$ is real and v_1 is real. In this case, the conclusion follows by letting $f(s) = v_1 \hat{f}(s)$. □

4.4 Power and Spectral Signals

In this section, we introduce two additional classes of signals that have been widely used in engineering. These classes of signals have some nice statistical and frequency domain representations. Let $u(t)$ be a function of time. Define its autocorrelation matrix as

$$R_{uu}(\tau) := \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} u(t+\tau)u^*(t)dt,$$

if the limit exists and is finite for all τ. It is easy to see from the definition that $R_{uu}(\tau) = R_{uu}^*(-\tau)$. Assume further that the Fourier transform of the signal's autocorrelation matrix function exists (and may contain impulses). This Fourier transform is called the *spectral density* of u, denoted $S_{uu}(j\omega)$:

$$S_{uu}(j\omega) := \int_{-\infty}^{\infty} R_{uu}(\tau)e^{-j\omega\tau}d\tau.$$

Thus $R_{uu}(\tau)$ can be obtained from $S_{uu}(j\omega)$ by performing an inverse Fourier transform:

$$R_{uu}(\tau) := \frac{1}{2\pi}\int_{-\infty}^{\infty} S_{uu}(j\omega)e^{j\omega\tau}\,d\omega.$$

We will call signal $u(t)$ a *power signal* if the autocorrelation matrix $R_{uu}(\tau)$ exists and is finite for all τ, and moreover, if the power spectral density function $S_{uu}(j\omega)$ exists (note that $S_{uu}(j\omega)$ need not be bounded and may include impulses).

The power of the signal is defined as

$$\|u\|_{\mathcal{P}} = \sqrt{\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T}\|u(t)\|^2\,dt} = \sqrt{\mathrm{Trace}\,[R_{uu}(0)]}$$

where $\|\cdot\|$ is the usual Euclidean norm and the capital script $\mathcal{P}$ is used to differentiate this power semi-norm from the usual Lebesgue $\mathcal{L}_p$ norm. The set of all signals having finite power will be denoted by $\mathcal{P}$.

The power semi-norm of a signal can also be computed from its spectral density function

$$\|u\|_{\mathcal{P}}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty}\mathrm{Trace}[S_{uu}(j\omega)]d\omega.$$

This expression implies that, if $u \in \mathcal{P}$, S_{uu} is strictly proper in the sense that $S_{uu}(\infty) = 0$. We note that if $u \in \mathcal{P}$ and $\|u(t)\|_\infty := \sup_t \|u(t)\| < \infty$, then $\|u\|_{\mathcal{P}} \le \|u\|_\infty$. However, not every signal having finite ∞-norm is a power signal since the limit in the definition may not exist. For example, let

$$u(t) = \begin{cases} 1 & 2^{2k} < t < 2^{2k+1}, \text{ for } k = 0,1,2,\ldots \\ 0 & \text{otherwise.}\end{cases}$$

Then $\lim_{T\to\infty}\frac{1}{2T}\int_{-T}^{T} u^2\,dt$ does not exist. Note also that power signals are persistent signals in time such as sines or cosines; clearly, a time-limited signal has zero power, as does an $\mathcal{L}_2$ signal. Thus $\|\cdot\|_{\mathcal{P}}$ is only a semi-norm, not a norm.

Now let G be a linear system transfer matrix with convolution kernel $g(t)$, input $u(t)$, and output $z(t)$. Then $R_{zz}(\tau) = g(\tau) * R_{uu}(\tau) * g^*(-\tau)$ and $S_{zz}(j\omega) = G(j\omega)S_{uu}(j\omega)G^*(j\omega)$. These properties are useful in establishing some input and output relationships in the next section.

A signal $u(t)$ is said to have bounded power spectral density if $\|S_{uu}(j\omega)\|_\infty < \infty$. The set of signals having bounded spectral density is denoted as

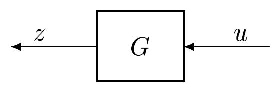
$$\mathcal{S} := \{u(t) \in \mathbb{R}^m \;:\; \|S_{uu}(j\omega)\|_\infty < \infty\}.$$

The quantity $\|u\|_s := \sqrt{\|S_{uu}(j\omega)\|}$ is called the spectral density norm of $u(t)$. The set $\mathcal{S}$ can be used to model signals with fixed spectral characteristics by passing white noise signals through a weighting filter. Similarly, $\mathcal{P}$ could be used to model signals whose spectrum is not known but which are finite in power.

4.5 Induced System Gains

Many control engineering problems involve keeping some system signals "small" under various conditions, for example, under a set of possible disturbances and system parameter variations. In this section we are interested in answering the following question: if we know how "big" the input (disturbance) is, how "big" is the output going to be for a given stable dynamical system?

Consider a q-input and p-output linear finite dimensional system as shown in the following diagram with input u, output z, and transfer matrix $G \in \mathcal{RH}_\infty$:



We will further assume that $G(s)$ is strictly proper, i.e, $G(\infty) = 0$ although most of the results derived here hold for the non-strictly proper case. In the time-domain an input-output model for such a system has the form of a convolution equation,

$$z = g * u$$

i.e.,

$$z(t) = \int_0^t g(t-\tau)u(\tau)\, d\tau$$

where the $p \times q$ real matrix $g(t)$ is the convolution kernel. Let the convolution kernel and the corresponding transfer matrix be partitioned as

$$g(t) = \begin{bmatrix} g_{11}(t) & \cdots & g_{1q}(t) \\ \vdots & & \vdots \\ g_{p1}(t) & \cdots & g_{pq}(t) \end{bmatrix} = \begin{bmatrix} g_1(t) \\ \vdots \\ g_p(t) \end{bmatrix},$$

$$G(s) = \begin{bmatrix} G_{11}(s) & \cdots & G_{1q}(s) \\ \vdots & & \vdots \\ G_{p1}(s) & \cdots & G_{pq}(s) \end{bmatrix} = \begin{bmatrix} G_1(s) \\ \vdots \\ G_p(s) \end{bmatrix}$$

where $g_i(t)$ is a q-dimensional row vector of the convolution kernel and $G_i(s)$ is a row vector of the transfer matrix. Now if G is considered as an operator from the input space to the output space, then a norm is induced on G, which, loosely speaking, measures the size of the output for a given input u. These norms can determine the achievable performance of the system for different classes of input signals

The various input/output relationships, given different classes of input signals, are summarized in two tables below. Table 4.1 summarizes the results for fixed input signals. Note that the $\delta(t)$ in this table denotes the unit impulse and $u_0 \in \mathbb{R}^q$ is a constant vector indicating the direction of the input signal.

We now prove Table 4.1.

Input $u(t)$	Output $z(t)$
$u(t) = u_0\delta(t), u_0 \in \mathbb{R}^q$	$\|z\|_2 = \|Gu_0\|_2 = \|gu_0\|_2$
	$\|z\|_\infty = \|gu_0\|_\infty$
	$\|z\|_{\mathcal{P}} = 0$
$u(t) = u_0 \sin\omega_0 t, u_0 \in \mathbb{R}^q$	$\|z\|_2 = \infty$
	$\limsup_{t\to\infty} \max_i \|z_i(t)\| = \|G(j\omega_0)u_0\|_\infty = \max_i \|G_i(j\omega_0)u_0\|$ $\limsup_{t\to\infty} \|z(t)\| = \|G(j\omega_0)u_0\|$
	$\|z\|_{\mathcal{P}} = \frac{1}{\sqrt{2}}\|G(j\omega_0)u_0\|$

Table 4.1: Output norms with fixed inputs

Proof.

$u = u_0\delta(t)$: If $u(t) = u_0\delta(t)$, then $z(t) = g(t)u_0$, so $\|z\|_2 = \|gu_0\|_2$ and $\|z\|_\infty = \|gu_0\|_\infty$. But by Parseval's relation, $\|gu_0\|_2 = \|Gu_0\|_2$. On the other hand,

$$\begin{aligned} \|z\|_{\mathcal{P}}^2 &= \lim_{T\to\infty} \frac{1}{2T}\int_{-T}^{T} u_0^* g^*(t)g(t)u_0 \; dt \\ &\le \lim_{T\to\infty} \frac{1}{2T}\int_{-\infty}^{\infty} \text{Trace}\{g^*(t)g(t)\} \; dt \; \|u_0\|^2 \\ &= \lim_{T\to\infty} \frac{1}{2T}\|G\|_2^2 \|u_0\|^2 = 0. \end{aligned}$$

$u = u_0 \sin\omega_0 t$: With the input $u(t) = u_0\sin(\omega_0 t)$, the i-th output as $t \to \infty$ is

$$z_i(t) = |G_i(j\omega_0)u_0| \sin[\omega_0 t + \arg\{G_i(j\omega_0)u_0\}]. \tag{4.1}$$

The 2-norm of this signal is infinite as long as $G_i(j\omega_0)u_0 \neq 0$, i.e., the system's transfer function does not have a zero in every channel in the input direction at the frequency of excitation.

The amplitude of the sinusoid (4.1) equals $|G_i(j\omega_0)u_0|$. Hence

$$\limsup_{t\to\infty} \max_i |z_i(t)| = \max_i |G_i(j\omega_0)u_0| = \|G(j\omega_0)u_0\|_\infty$$

and

$$\limsup_{t\to\infty} \|z(t)\| = \sqrt{\sum_{i=1}^{p} |G_i(j\omega_0)u_0|^2} = \|G(j\omega_0)u_0\|.$$

Finally, let $\phi_i := \arg G_i(j\omega)u_0$. Then

$$\begin{aligned}
\|z\|_{\mathcal{P}}^2 &= \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \sum_{i=1}^{p} |G_i(j\omega_0)u_0|^2 \sin^2(\omega_0 t + \phi_i)\, dt \\
&= \sum_{i=1}^{p} |G_i(j\omega_0)u_0|^2 \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \sin^2(\omega_0 t + \phi_i)\, dt \\
&= \sum_{i=1}^{p} |G_i(j\omega_0)u_0|^2 \lim_{T\to\infty} \frac{1}{2T} \int_{-T}^{T} \frac{1 - \cos 2(\omega_0 t + \phi_i)}{2}\, dt \\
&= \frac{1}{2} \sum_{i=1}^{p} |G_i(j\omega_0)u_0|^2 = \frac{1}{2} \|G(j\omega_0)u_0\|^2.
\end{aligned}$$

□

It is interesting to see what this table signifies from the control point of view. To focus our discussion, let us assume that $u(t) = u_0 \sin \omega_0 t$ is a disturbance or a command signal on a feedback system, and $z(t)$ is the tracking error. Then we say that the system has good tracking behavior if $z(t)$ is small in some sense, for instance, $\limsup_{t\to\infty} \|z(t)\|$ is small. Note that

$$\limsup_{t\to\infty} \|z(t)\| = \|G(j\omega_0)u_0\|$$

for any given ω_0 and $u_0 \in \mathbb{R}^q$. Now if we want to track signals from various channels, that is if u_0 can be chosen to be any direction, then we would require that $\bar{\sigma}(G(j\omega_0))$ be small. Furthermore, if, in addition, we want to track signals of many different frequencies, we then would require that $\bar{\sigma}(G(j\omega_0))$ be small at all those frequencies. This interpretation enables us to consider the control system in the frequency domain even though the specifications are given in the time domain.

Table 4.2 lists the maximum possible system gains when the input signal u is not required to be a fixed signal; instead it can be any signal in a unit ball in some function space.

Now we give a proof for Table 4.2. Note that the first row ($\mathcal{L}_2 \longmapsto \mathcal{L}_2$) has been shown in Section 4.3.

Input $u(t)$	Output $z(t)$	Signal Norms	Induced Norms
$\mathcal{L}_2$	$\mathcal{L}_2$	$\|u\|_2^2 = \int_0^\infty \|u\|^2 \, dt$	$\|G\|_\infty$
$\mathcal{S}$	$\mathcal{S}$	$\|u\|_{\mathcal{S}}^2 = \|S_{uu}\|_\infty$	$\|G\|_\infty$
$\mathcal{S}$	$\mathcal{P}$	$\|u\|_{\mathcal{P}}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{Trace}\{S_{uu}(j\omega)\}d\omega$	$\|G\|_2$
$\mathcal{P}$	$\mathcal{P}$	$\|u\|_{\mathcal{P}}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{Trace}\{S_{uu}(j\omega)\}d\omega$	$\|G\|_\infty$
$\mathcal{L}_\infty$	$\mathcal{L}_\infty$	$\|u\|_\infty = \sup_t \max_i \lvert u_i(t)\rvert$	$\max_i \|g_i\|_1$
		$\|u\|_\infty = \sup_t \|u(t)\|$	$\leq \int_0^\infty \|g(t)\| \, dt$

Table 4.2: Induced System Gains

Proof.

$\mathcal{S} \longmapsto \mathcal{S}$: If $u \in \mathcal{S}$, then

$$S_{zz}(j\omega) = G(j\omega)S_{uu}(j\omega)G^*(j\omega).$$

Now suppose

$$\overline{\sigma}[G(j\omega_0)] = \|G\|_\infty$$

and take a signal u such that $S_{uu}(j\omega_0) = I$. Then

$$\|S_{zz}(j\omega)\|_\infty = \|G\|_\infty^2 .$$

$\mathcal{S} \longmapsto \mathcal{P}$: By definition, we have

$$\|z\|_{\mathcal{P}}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{Trace}\{G(j\omega)S_{uu}(j\omega)G^*(j\omega)\} \, d\omega.$$

Now let u be white with unit spectral density, i.e., $S_{uu} = I$. Then

$$\|z\|_{\mathcal{P}}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{Trace}\{G(j\omega)G^*(j\omega)\}\, d\omega = \|G\|_2^2 .$$

$\mathcal{P} \longmapsto \mathcal{P}$:

If u is a power signal, then

$$\|z\|_{\mathcal{P}}^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \mathrm{Trace}\{G(j\omega)S_{uu}(j\omega)G^*(j\omega)\}\, d\omega,$$

and immediately, we get that

$$\|z\|_{\mathcal{P}} \le \|G\|_\infty \|u\|_{\mathcal{P}} .$$

To achieve the equality, assume that ω_0 is such that

$$\overline{\sigma}\left[G(j\omega_0)\right] = \|G\|_\infty$$

and denote the singular value decomposition of $G(j\omega_0)$ by

$$G(j\omega_0) = \overline{\sigma} u_1(j\omega_0)v_1^*(j\omega_0) + \sum_{i=2}^{r} \sigma_i u_i(j\omega_0)v_i^*(j\omega_0)$$

where r is the rank of $G(j\omega_0)$ and u_i, v_i have unit length.

If $\omega_0 < \infty$, write $v_1(j\omega_0)$ as

$$v_1(j\omega) = \begin{bmatrix} \alpha_1 e^{j\theta_1} \\ \alpha_2 e^{j\theta_2} \\ \vdots \\ \alpha_q e^{j\theta_q} \end{bmatrix}$$

where $\alpha_i \in \mathbb{R}$ is such that $\theta_i \in [-\pi, 0)$ and q is the column dimension of G. Now let $\beta_i \ge 0$ be such that

$$\theta_i = \angle\left(\frac{\beta_i - j\omega_0}{\beta_i + j\omega_0}\right)$$

and let the input u be generated from passing $\hat{u}$ through a filter:

$$u(t) = \begin{bmatrix} \alpha_1 \frac{\beta_1 - s}{\beta_1 + s} \\ \alpha_2 \frac{\beta_2 - s}{\beta_2 + s} \\ \vdots \\ \alpha_q \frac{\beta_q - s}{\beta_q + s} \end{bmatrix} \hat{u}(t)$$

where

$$\hat{u}(t) = \sqrt{2}\sin(\omega_o t).$$

Then $R_{\hat{u}\hat{u}}(\tau) = \cos(\omega_o \tau)$, so

$$\|\hat{u}\|_{\mathcal{P}} = R_{\hat{u}\hat{u}}(0) = 1.$$

Also,

$$S_{\hat{u}\hat{u}}(j\omega) = \pi\left[\delta(\omega - \omega_o) + \delta(\omega + \omega_o)\right].$$

Then

$$S_{uu}(j\omega) = \begin{bmatrix} \alpha_1 \frac{\beta_1 - j\omega}{\beta_1 + j\omega} \\ \alpha_2 \frac{\beta_2 - j\omega}{\beta_2 + j\omega} \\ \vdots \\ \alpha_q \frac{\beta_q - j\omega}{\beta_q + j\omega} \end{bmatrix} S_{\hat{u}\hat{u}}(j\omega) \begin{bmatrix} \alpha_1 \frac{\beta_1 - j\omega}{\beta_1 + j\omega} \\ \alpha_2 \frac{\beta_2 - j\omega}{\beta_2 + j\omega} \\ \vdots \\ \alpha_q \frac{\beta_q - j\omega}{\beta_q + j\omega} \end{bmatrix}^*$$

and it is easy to show

$$\|u\|_{\mathcal{P}} = 1.$$

Finally,

$$\begin{aligned} \|z\|_{\mathcal{P}}^2 &= \frac{1}{2}\overline{\sigma}\left[G(j\omega_o)\right]^2 + \frac{1}{2}\overline{\sigma}\left[G(-j\omega_o)\right]^2 \\ &= \overline{\sigma}\left[G(-j\omega_o)\right]^2 \\ &= \|G\|_\infty^2. \end{aligned}$$

Similarly, if $\omega_0 = \infty$, then $G(j\omega)$ is real. The conclusion follows by letting $u = v_1\hat{u}(t)$ and $\omega_o \to \infty$.

$\mathcal{L}_\infty \longmapsto \mathcal{L}_\infty$:

1. First of all, $\max_i \|g_i\|_1$ is an upper bound on the ∞-norm/∞-norm system gain:

$$\begin{aligned} |z_i(t)| &= \left| \int_0^t g_i(\tau)u(t-\tau)\,d\tau \right| \le \int_0^t |g_i(\tau)u(t-\tau)|\,d\tau \\ &\le \int_0^t \sum_{j=1}^q |g_{ij}(\tau)|\,d\tau\,\|u\|_\infty \le \int_0^\infty \sum_{j=1}^q |g_{ij}(\tau)|\,d\tau\,\|u\|_\infty \\ &= \|g_i\|_1 \|u\|_\infty \end{aligned}$$

and

$$\begin{aligned} \|z\|_\infty &:= \max_i \sup_t |z_i(t)| \\ &\le \max_i \|g_i\|_1 \|u\|_\infty . \end{aligned}$$

That $\max_i \|g_i\|_1$ is the least upper bound can be seen as follows: Assume that the maximum $\max_i \|g_i\|_1$ is achieved for $i = 1$. Let t_0 be given and set

$$u(t_0 - \tau) := \operatorname{sign}(g_1^*(\tau)), \ \forall \tau.$$

Note that since $g_1(\tau)$ is a vector function, the sign function $\operatorname{sign}(g_1^*(\tau))$ is a component-wise operation. Then $\|u\|_\infty = 1$ and

$$\begin{aligned} z_1(t_0) &= \int_0^{t_0} g_1(\tau) u(t_0 - \tau) \ d\tau \\ &= \int_0^{t_0} \sum_{j=1}^{q} |g_{1j}(\tau)| \ d\tau \\ &= \|g_1\|_1 - \int_{t_0}^{\infty} \sum_{j=1}^{q} |g_{1j}(\tau)| \ d\tau. \end{aligned}$$

Hence, let $t_0 \to \infty$, and we have $\|z\|_\infty = \|g_1\|_1$.

2. If $\|u(t)\|_\infty := \sup_t \|u(t)\|$, then

$$\begin{aligned} \|z(t)\| &= \left\| \int_0^t g(\tau) u(t-\tau) d\tau \right\| \\ &\le \int_0^t \|g(\tau)\| \, \|u(t-\tau)\| \, d\tau \\ &\le \int_0^t \|g(\tau)\| \, d\tau \ \ \|u\|_\infty . \end{aligned}$$

And, therefore, $\|z\|_\infty \le \int_0^\infty \|g(\tau)\| \, d\tau \ \ \|u\|_\infty$.

□

Next we shall derive some simple and useful bounds for the $\mathcal{H}_\infty$ norm and the $\mathcal{L}_1$ norm of a stable system. Suppose

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array} \right] \in \mathcal{RH}_\infty$$

is a balanced realization, i.e., there exists

$$\Sigma = \operatorname{diag}(\sigma_1, \sigma_2, \ldots, \sigma_n) \ge 0$$

with $\sigma_1 \ge \sigma_2 \ge \ldots \ge \sigma_n \ge 0$, such that

$$A\Sigma + \Sigma A^* + BB^* = 0 \qquad A^*\Sigma + \Sigma A + C^*C = 0.$$

Then we have the following theorem.

Theorem 4.5

$$\sigma_1 \le \|G\|_\infty \le \int_0^\infty \|g(t)\|\, dt \le 2\sum_{i=1}^n \sigma_i$$

where $g(t) = Ce^{At}B$.

Remark 4.3 It should be clear that the inequalities stated in the theorem do not depend on a particular state space realization of $G(s)$. However, use of the balanced realization does make the proof simple. ♡

Proof. The inequality $\sigma_1 \le \|G\|_\infty$ follows from the Nehari Theorem of Chapter 8. We will now show the other inequalities. Since

$$G(s) := \int_0^\infty g(t)e^{-st}dt, \ \mathrm{Re}(s) > 0,$$

by the definition of $\mathcal{H}_\infty$ norm, we have

$$\begin{aligned}
\|G\|_\infty &= \sup_{\mathrm{Re}(s)>0} \left\| \int_0^\infty g(t)e^{-st}dt \right\| \\
&\le \sup_{\mathrm{Re}(s)>0} \int_0^\infty \left\| g(t)e^{-st} \right\| dt \\
&\le \int_0^\infty \|g(t)\|\, dt.
\end{aligned}$$

To prove the last inequality, let u_i be the i^{th} unit vector. Then

$$u_i^* u_j = \delta_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \ne j \end{cases} \quad \text{and} \quad \sum_{i=1}^n u_i u_i^* = I.$$

Define $\gamma_i(t) = u_i^* e^{At/2}B$ and $\psi_i(t) = Ce^{At/2}u_i$. It is easy to verify that

$$\|\gamma_i\|_2^2 = \int_0^\infty u_i^* e^{At/2}BB^*e^{A^*t/2}u_i dt = 2u_i^*\Sigma u_i = 2\sigma_i$$

$$\|\psi_i\|_2^2 = \int_0^\infty u_i^* e^{A^*t/2}C^*Ce^{At/2}u_i dt = 2u_i^*\Sigma u_i = 2\sigma_i.$$

Using these two equalities, we have

$$\begin{aligned}
\int_0^\infty \|g(t)\|\, dt &= \int_0^\infty \left\| \sum_{i=1}^n \psi_i\gamma_i \right\| dt \le \sum_{i=1}^n \int_0^\infty \|\psi_i\gamma_i\|\, dt \\
&\le \sum_{i=1}^n \|\psi_i\|_2 \|\gamma_i\|_2 \le 2\sum_{i=1}^n \sigma_i.
\end{aligned}$$

□

It should be clear from the above two tables that many system performance criteria can be stipulated as requiring a certain closed loop transfer matrix have small $\mathcal{H}_2$ norm or $\mathcal{H}_\infty$ norm or L_1 norm. Moreover, if L_1 performance is satisfied, then the $\mathcal{H}_\infty$ norm performance is also satisfied. We will be most interested in $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance in this book.

4.6 Computing $\mathcal{L}_2$ and $\mathcal{H}_2$ Norms

Let $G(s) \in \mathcal{L}_2$ and recall that the $\mathcal{L}_2$ norm of G is defined as

$$\begin{aligned} \|G\|_2 &:= \sqrt{\frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{Trace}\{G^*(j\omega)G(j\omega)\}\, d\omega} \\ &= \|g\|_2 \\ &= \sqrt{\int_{-\infty}^{\infty} \operatorname{Trace}\{g^*(t)g(t)\}\, dt} \end{aligned}$$

where $g(t)$ denotes the convolution kernel of G.

It is easy to see that the $\mathcal{L}_2$ norm defined above is finite iff the transfer matrix G is strictly proper, i.e., $G(\infty) = 0$. Hence, we will generally assume that the transfer matrix is strictly proper whenever we refer to the $\mathcal{L}_2$ norm of G (of course, this also applies to $\mathcal{H}_2$ functions). One straightforward way of computing the $\mathcal{L}_2$ norm is to use contour integral. Suppose G is strictly proper, and then we have

$$\begin{aligned} \|G\|_2^2 &= \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{Trace}\{G^*(j\omega)G(j\omega)\}\, d\omega \\ &= \frac{1}{2\pi j}\oint \operatorname{Trace}\{G^\sim(s)G(s)\}\, ds. \end{aligned}$$

The last integral is a contour integral along the imaginary axis, and around an infinite semi-circle in the left half-plane; the contribution to the integral from this semi-circle equals zero because G is strictly proper. By the residue theorem, $\|G\|_2^2$ equals the sum of the residues of $\operatorname{Trace}\{G^\sim(s)G(s)\}$ at its poles in the left half-plane.

Although $\|G\|_2$ can, in principle, be computed from its definition or from the method suggested above, it is useful in many applications to have alternative characterizations and to take advantage of the state space representations of G. The computation of a $\mathcal{RH}_2$ transfer matrix norm is particularly simple.

Lemma 4.6 *Consider a transfer matrix*

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$$

with A stable. Then we have

$$\|G\|_2^2 = \text{trace}(B^* L_o B) = \text{trace}(C L_c C^*) \tag{4.2}$$

where L_o and L_c are observability and controllability Gramians which can be obtained from the following Lyapunov equations

$$A L_c + L_c A^* + B B^* = 0 \qquad A^* L_o + L_o A + C^* C = 0.$$

Proof. Since G is stable, we have

$$g(t) = \mathcal{L}^{-1}(G) = \begin{cases} Ce^{At}B, & t \geq 0 \\ 0, & t < 0 \end{cases}$$

and

$$\begin{aligned} \|G\|_2^2 &= \int_0^\infty \text{Trace}\{g^*(t)g(t)\}\, dt = \int_0^\infty \text{Trace}\{g(t)g(t)^*\}\, dt \\ &= \int_0^\infty \text{Trace}\{B^* e^{A^*t} C^* C e^{At} B\}\, dt = \int_0^\infty \text{Trace}\{C e^{At} B B^* e^{A^*t} C^*\}\, dt. \end{aligned}$$

The lemma follows from the fact that the controllability Gramian of (A, B) and the observability Gramian of (C, A) can be represented as

$$L_o = \int_0^\infty e^{A^*t} C^* C e^{At}\, dt, \quad L_c = \int_0^\infty e^{At} B B^* e^{A^*t}\, dt,$$

which can also be obtained from

$$A L_c + L_c A^* + B B^* = 0 \qquad A^* L_o + L_o A + C^* C = 0.$$

□

To compute the $\mathcal{L}_2$ norm of a rational transfer function, $G(s) \in \mathcal{L}_2$, using state space approach. Let $G(s) = [G(s)]_+ + [G(s)]_-$ with $G_+ \in \mathcal{RH}_2$ and $G_- \in \mathcal{RH}_2^\perp$. Then

$$\|G\|_2^2 = \|[G(s)]_+\|_2^2 + \|[G(s)]_-\|_2^2$$

where $\|[G(s)]_+\|_2$ and $\|[G(s)]_-\|_2 = \|[G(-s)]_+\|_2$ can be computed using the above lemma.

Still another useful characterization of the $\mathcal{H}_2$ norm of G is in terms of hypothetical input-output experiments. Let e_i denote the i^{th} standard basis vector of $\mathbb{R}^m$ where m is the input dimension of the system. Apply the impulsive input $\delta(t)e_i$ ($\delta(t)$ is the unit impulse) and denote the output by $z_i(t) (= g(t)e_i)$. Assume $D = 0$, and then $z_i \in \mathcal{L}_{2+}$ and

$$\|G\|_2^2 = \sum_{i=1}^m \|z_i\|_2^2.$$

Note that this characterization of the $\mathcal{H}_2$ norm can be appropriately generalized for nonlinear time varying systems, see Chen and Francis [1992] for an application of this norm in sampled-data control.

4.7 Computing $\mathcal{L}_\infty$ and $\mathcal{H}_\infty$ Norms

We shall first consider, as in the $\mathcal{L}_2$ case, how to compute the ∞ norm of an $\mathcal{L}_\infty$ transfer matrix. Let $G(s) \in \mathcal{L}_\infty$ and recall that the $\mathcal{L}_\infty$ norm of a transfer function G is defined as

$$\|G\|_\infty := \text{ess}\ \sup_\omega \overline{\sigma}\{G(j\omega)\}.$$

The computation of the $\mathcal{L}_\infty$ norm of G is complicated and requires a search. A control engineering interpretation of the infinity norm of a scalar transfer function G is the distance in the complex plane from the origin to the farthest point on the Nyquist plot of G, and it also appears as the peak value on the Bode magnitude plot of $|G(j\omega)|$. Hence the ∞ of a transfer function can in principle be obtained graphically.

To get an estimate, set up a fine grid of frequency points,

$$\{\omega_1, \cdots, \omega_N\}.$$

Then an estimate for $\|G\|_\infty$ is

$$\max_{1\le k\le N} \overline{\sigma}\{G(j\omega_k)\}.$$

This value is usually read directly from a Bode singular value plot. The $\mathcal{L}_\infty$ norm can also be computed in state space if G is rational.

Lemma 4.7 *Let $\gamma > 0$ and*

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RL}_\infty. \tag{4.3}$$

Then $\|G\|_\infty < \gamma$ if and only if $\overline{\sigma}(D) < \gamma$ and H has no eigenvalues on the imaginary axis where

$$H := \begin{bmatrix} A + BR^{-1}D^*C & BR^{-1}B^* \\ -C^*(I + DR^{-1}D^*)C & -(A + BR^{-1}D^*C)^* \end{bmatrix} \tag{4.4}$$

and $R = \gamma^2 I - D^*D$.

Proof. Let $\Phi(s) = \gamma^2 I - G^\sim(s)G(s)$. Then it is clear that $\|G\|_\infty < \gamma$ if and only if $\Phi(j\omega) > 0$ for all $\omega \in \mathbb{R}$. Since $\Phi(\infty) = R > 0$ and since $\Phi(j\omega)$ is a continuous function of ω, $\Phi(j\omega) > 0$ for all $\omega \in \mathbb{R}$ if and only if $\Phi(j\omega)$ is nonsingular for all $\omega \in \mathbb{R} \cup \{\infty\}$, i.e., $\Phi(s)$ has no imaginary axis zero. Equivalently, $\Phi^{-1}(s)$ has no imaginary axis pole. It is easy to compute by some simple algebra that

$$\Phi^{-1}(s) = \left[\begin{array}{c|c} H & \begin{bmatrix} BR^{-1} \\ -C^*DR^{-1} \end{bmatrix} \\ \hline \begin{bmatrix} R^{-1}D^*C & R^{-1}B^* \end{bmatrix} & R^{-1} \end{array}\right].$$

Thus the conclusion follows if the above realization has neither uncontrollable modes nor unobservable modes on the imaginary axis. Assume that $j\omega_0$ is an eigenvalue of H but not a pole of $\Phi^{-1}(s)$. Then $j\omega_0$ must be either an unobservable mode of $(\begin{bmatrix} R^{-1}D^*C & R^{-1}B^* \end{bmatrix}, H)$ or an uncontrollable mode of $(H, \begin{bmatrix} BR^{-1} \\ -C^*DR^{-1} \end{bmatrix})$. Now suppose $j\omega_0$ is an unobservable mode of $(\begin{bmatrix} R^{-1}D^*C & R^{-1}B^* \end{bmatrix}, H)$. Then there exists an $x_0 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq 0$ such that

$$Hx_0 = j\omega_0 x_0, \ \begin{bmatrix} R^{-1}D^*C & R^{-1}B^* \end{bmatrix} x_0 = 0.$$

These equations can be simplified to

$$\begin{aligned} (j\omega_0 I - A)x_1 &= 0 \\ (j\omega_0 I + A^*)x_2 &= -C^*Cx_1 \\ D^*Cx_1 + B^*x_2 &= 0. \end{aligned}$$

Since A has no imaginary axis eigenvalues, we have $x_1 = 0$ and $x_2 = 0$. This contradicts our assumption, and hence the realization has no unobservable modes on the imaginary axis.

Similarly, a contradiction will also be arrived if $j\omega_0$ is assumed to be an uncontrollable mode of $(H, \begin{bmatrix} BR^{-1} \\ -C^*DR^{-1} \end{bmatrix})$. □

Bisection Algorithm

Lemma 4.7 suggests the following bisection algorithm to compute $\mathcal{RL}_\infty$ norm:

(a) select an upper bound γ_u and a lower bound γ_l such that $\gamma_l \leq \|G\|_\infty \leq \gamma_u$;

(b) if $(\gamma_u - \gamma_l)/\gamma_l \leq$specified level, stop; $\|G\| \approx (\gamma_u + \gamma_l)/2$. Otherwise go to next step;

(c) set $\gamma = (\gamma_l + \gamma_u)/2$;

(d) test if $\|G\|_\infty < \gamma$ by calculating the eigenvalues of H for the given γ;

(e) if H has an eigenvalue on $j\mathbb{R}$ set $\gamma_l = \gamma$; otherwise set $\gamma_u = \gamma$; go back to step (b).

Of course, the above algorithm applies to $\mathcal{H}_\infty$ norm computation as well. Thus $\mathcal{L}_\infty$ norm computation requires a search, over either γ or ω, in contrast to $\mathcal{L}_2$ ($\mathcal{H}_2$) norm computation, which does not. A somewhat analogous situation occurs for constant matrices with the norms $\|M\|_2^2 = \text{trace}(M^*M)$ and $\|M\|_\infty = \overline{\sigma}[M]$. In principle, $\|M\|_2^2$ can be computed exactly with a finite number of operations, as can the test for whether $\overline{\sigma}(M) < \gamma$ (e.g. $\gamma^2 I - M^*M > 0$), but the value of $\overline{\sigma}(M)$ cannot. To compute $\overline{\sigma}(M)$, we must use some type of iterative algorithm.

Remark 4.4 It is clear that $\|G\|_\infty < \gamma$ iff $\left\|\gamma^{-1}G\right\|_\infty < 1$. Hence, there is no loss of generality to assume $\gamma = 1$. This assumption will often be made in the remainder of this book. It is also noted that there are other fast algorithms to carry out the above norm computation; nevertheless, this bisection algorithm is the simplest. ♡

The $\mathcal{H}_\infty$ norm of a stable transfer function can also be estimated experimentally using the fact that the $\mathcal{H}_\infty$ norm of a stable transfer function is the maximum magnitude of the steady-state response to all possible unit amplitude sinusoidal input signals.

4.8 Notes and References

The basic concept of function spaces presented in this chapter can be found in any standard functional analysis textbook, for instance, Naylor and Sell [1982] and Gohberg and Goldberg [1981]. The system theoretical interpretations of the norms and function spaces can be found in Desoer and Vidyasagar [1975]. The bisection $\mathcal{H}_\infty$ norm computational algorithm is first developed in Boyd, Balakrishnan, and Kabamba [1989]. A more efficient $\mathcal{L}_\infty$ norm computational algorithm is presented in Bruinsma and Steinbuch [1990].

5

Stability and Performance of Feedback Systems

This chapter introduces the feedback structure and discusses its stability and performance properties. The arrangement of this chapter is as follows: Section 5.1 discusses the necessity for introducing feedback structure and describes the general feedback configuration. In section 5.2, the well-posedness of the feedback loop is defined. Next, the notion of internal stability is introduced and the relationship is established between the state space characterization of internal stability and the transfer matrix characterization of internal stability in section 5.3. The stable coprime factorizations of rational matrices are also introduced in section 5.4. Section 5.5 considers feedback properties and discusses how to achieve desired performance using feedback control. These discussions lead to a loop shaping control design technique which is introduced in section 5.6. Finally, we consider the mathematical formulations of optimal $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems in section 5.7.

5.1 Feedback Structure

In designing control systems, there are several fundamental issues that transcend the boundaries of specific applications. Although they may differ for each application and may have different levels of importance, these issues are generic in their relationship to control design objectives and procedures. Central to these issues is the requirement to provide satisfactory performance in the face of modeling errors, system variations, and uncertainty. Indeed, this requirement was the original motivation for the development of feedback systems. Feedback is only required when system performance cannot be achieved because of uncertainty in system characteristics. The more detailed treatment of model uncertainties and their representations will be discussed in Chapter 9.

For the moment, assuming we are given a model including a representation of uncertainty which we believe adequately captures the essential features of the plant, the next step in the controller design process is to determine what structure is necessary to achieve the desired performance. Prefiltering input signals (or open loop control) can change the dynamic response of the model set but cannot reduce the effect of uncertainty. If the uncertainty is too great to achieve the desired accuracy of response, then a feedback structure is required. The mere assumption of a feedback structure, however, does not guarantee a reduction of uncertainty, and there are many obstacles to achieving the uncertainty-reducing benefits of feedback. In particular, since for any reasonable model set representing a physical system uncertainty becomes large and the phase is completely unknown at sufficiently high frequencies, the loop gain must be small at those frequencies to avoid destabilizing the high frequency system dynamics. Even worse is that the feedback system actually increases uncertainty and sensitivity in the frequency ranges where uncertainty is significantly large. In other words, because of the type of sets required to reasonably model physical systems and because of the restriction that our controllers be causal, we cannot use feedback (or any other control structure) to cause our closed-loop model set to be a proper subset of the open-loop model set. Often, what can be achieved with intelligent use of feedback is a significant reduction of uncertainty for certain signals of importance with a small increase spread over other signals. Thus, the feedback design problem centers around the tradeoff involved in reducing the overall impact of uncertainty. This tradeoff also occurs, for example, when using feedback to reduce command/disturbance error while minimizing response degradation due to measurement noise. To be of practical value, a design technique must provide means for performing these tradeoffs. We will discuss these tradeoffs in more detail later in section 5.5 and in Chapter 6.

To focus our discussion, we will consider the standard feedback configuration shown in Figure 5.1. It consists of the interconnected plant P and controller K forced by

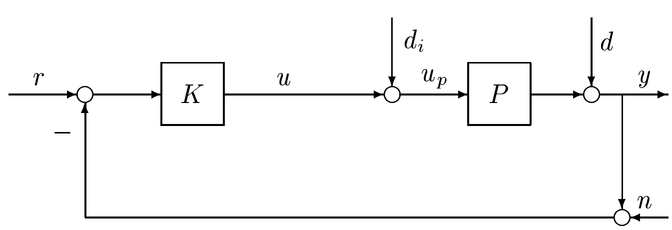


Figure 5.1: Standard Feedback Configuration

command r, sensor noise n, plant input disturbance d_i, and plant output disturbance d. In general, all signals are assumed to be multivariable, and all transfer matrices are assumed to have appropriate dimensions.

5.2 Well-Posedness of Feedback Loop

Assume that the plant P and the controller K in Figure 5.1 are fixed real rational proper transfer matrices. Then the first question one would ask is whether the feedback interconnection makes sense or is physically realizable. To be more specific, consider a simple example where

$$P = -\frac{s-1}{s+2}, \quad K = 1$$

are both proper transfer functions. However,

$$u = \frac{(s+2)}{3}(r - n - d) - \frac{s-1}{3} d_i$$

i.e., the transfer functions from the external signals $r - n - d$ and d_i to u are not proper. Hence, the feedback system is not physically realizable!

Definition 5.1 A feedback system is said to be *well-posed* if all closed-loop transfer matrices are well-defined and proper.

Now suppose that all the external signals r, n, d, and d_i are specified and that the closed-loop transfer matrices from them to u are respectively well-defined and proper. Then, y and all other signals are also well-defined and the related transfer matrices are proper. Furthermore, since the transfer matrices from d and n to u are the same and differ from the transfer matrix from r to u by only a sign, the system is well-posed if and only if the transfer matrix from $\begin{bmatrix} d_i \\ d \end{bmatrix}$ to u exists and is proper.

In order to be consistent with the notation used in the rest of the book, we shall denote

$$\hat{K} := -K$$

and regroup the external input signals into the feedback loop as w_1 and w_2 and regroup the input signals of the plant and the controller as e_1 and e_2. Then the feedback loop with the plant and the controller can be simply represented as in Figure 5.2 and the system is well-posed if and only if the transfer matrix from $\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ to e_1 exists and is proper.

Lemma 5.1 *The feedback system in Figure 5.2 is well-posed if and only if*

$$I - \hat{K}(\infty)P(\infty) \tag{5.1}$$

is invertible.

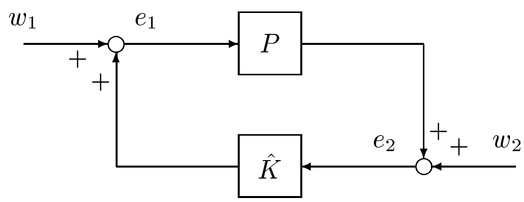


Figure 5.2: Internal Stability Analysis Diagram

Proof. The system in the above diagram can be represented in equation form as

$$\begin{aligned} e_1 &= w_1 + \hat{K} e_2 \\ e_2 &= w_2 + P e_1. \end{aligned}$$

Then an expression for e_1 can be obtained as

$$(I - \hat{K}P)e_1 = w_1 + \hat{K}w_2.$$

Thus well-posedness is equivalent to the condition that $(I - \hat{K}P)^{-1}$ exists and is proper. But this is equivalent to the condition that the constant term of the transfer function $I - \hat{K}P$ is invertible. $\square$

It is straightforward to show that (5.1) is equivalent to either one of the following two conditions:

$$\begin{bmatrix} I & -\hat{K}(\infty) \\ -P(\infty) & I \end{bmatrix} \text{ is invertible;} \tag{5.2}$$

$$I - P(\infty)\hat{K}(\infty) \text{ is invertible.}$$

The well-posedness condition is simple to state in terms of state-space realizations. Introduce realizations of P and $\hat{K}$:

$$P = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \tag{5.3}$$

$$\hat{K} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right]. \tag{5.4}$$

Then $P(\infty) = D$ and $\hat{K}(\infty) = \hat{D}$. For example, well-posedness in (5.2) is equivalent to the condition that

$$\begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix} \text{ is invertible.} \tag{5.5}$$

Fortunately, in most practical cases we will have $D = 0$, and hence well-posedness for most practical control systems is guaranteed.

5.3 Internal Stability

Consider a system described by the standard block diagram in Figure 5.2 and assume the system is well-posed. Furthermore, assume that the realizations for $P(s)$ and $\hat{K}(s)$ given in equations (5.3) and (5.4) are *stabilizable and detectable*.

Let x and $\hat{x}$ denote the state vectors for P and $\hat{K}$, respectively, and write the state equations in Figure 5.2 with w_1 and w_2 set to zero:

$$\begin{aligned} \dot{x} &= Ax + Be_1 && (5.6)\\ e_2 &= Cx + De_1 && (5.7)\\ \dot{\hat{x}} &= \hat{A}\hat{x} + \hat{B}e_2 && (5.8)\\ e_1 &= \hat{C}\hat{x} + \hat{D}e_2. && (5.9) \end{aligned}$$

Definition 5.2 The system of Figure 5.2 is said to be *internally stable* if the origin $(x, \hat{x}) = (0, 0)$ is asymptotically stable, i.e., the states $(x, \hat{x})$ go to zero from all initial states when $w_1 = 0$ and $w_2 = 0$.

Note that internal stability is a state space notion. To get a concrete characterization of internal stability, solve equations (5.7) and (5.9) for e_1 and e_2:

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}.$$

Note that the existence of the inverse is guaranteed by the well-posedness condition. Now substitute this into (5.6) and (5.8) to get

$$\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} = \tilde{A} \begin{bmatrix} x \\ \hat{x} \end{bmatrix}$$

where

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix} \begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix}.$$

Thus internal stability is equivalent to the condition that $\tilde{A}$ has all its eigenvalues in the open left-half plane. In fact, this can be taken as a definition of internal stability.

Lemma 5.2 *The system of Figure 5.2 with given stabilizable and detectable realizations for P and $\hat{K}$ is internally stable if and only if $\tilde{A}$ is a Hurwitz matrix.*

It is routine to verify that the above definition of internal stability depends only on P and $\hat{K}$, not on specific realizations of them as long as the realizations of P and $\hat{K}$ are both stabilizable and detectable, i.e., no extra unstable modes are introduced by the realizations.

The above notion of internal stability is defined in terms of state-space realizations of P and $\hat{K}$. It is also important and useful to characterize internal stability from the transfer matrix point of view. Note that the feedback system in Figure 5.2 is described, in term of transfer matrices, by

$$\begin{bmatrix} I & -\hat{K} \\ -P & I \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}. \tag{5.10}$$

Now it is intuitively clear that if the system in Figure 5.2 is internally stable, then for all bounded inputs (w_1, w_2), the outputs (e_1, e_2) are also bounded. The following lemma shows that this idea leads to a transfer matrix characterization of internal stability.

Lemma 5.3 *The system in Figure 5.2 is internally stable if and only if the transfer matrix*

$$\begin{bmatrix} I & -\hat{K} \\ -P & I \end{bmatrix}^{-1} = \begin{bmatrix} I + \hat{K}(I - P\hat{K})^{-1}P & \hat{K}(I - P\hat{K})^{-1} \\ (I - P\hat{K})^{-1}P & (I - P\hat{K})^{-1} \end{bmatrix} \tag{5.11}$$

from (w_1, w_2) to (e_1, e_2) belongs to $\mathcal{RH}_\infty$.

Proof. As above let $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ and $\left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right]$ be stabilizable and detectable realizations of P and $\hat{K}$, respectively. Let y_1 denote the output of P and y_2 the output of $\hat{K}$. Then the state-space equations for the system in Figure 5.2 are

$$\begin{aligned}
\begin{bmatrix} \dot{x} \\ \dot{\hat{x}} \end{bmatrix} &= \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \\
\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \begin{bmatrix} C & 0 \\ 0 & \hat{C} \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} D & 0 \\ 0 & \hat{D} \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \\
\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} &= \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}.
\end{aligned}$$

The last two equations can be rewritten as

$$\begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix} \begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix} \begin{bmatrix} x \\ \hat{x} \end{bmatrix} + \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

Now suppose that this system is internally stable. Then the well-posedness condition implies that $(I - D\hat{D}) = (I - P\hat{K})(\infty)$ is invertible. Hence, $(I - P\hat{K})$ is invertible. Furthermore, since the eigenvalues of

$$\tilde{A} = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix} + \begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix} \begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix}^{-1} \begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix}$$

are in the open left-half plane, it follows that the transfer matrix from (w_1, w_2) to (e_1, e_2) given in (5.11) is in $\mathcal{RH}_\infty$.

Conversely, suppose that $(I - P\hat{K})$ is invertible and the transfer matrix in (5.11) is in $\mathcal{RH}_\infty$. Then, in particular, $(I - P\hat{K})^{-1}$ is proper which implies that $(I - P\hat{K})(\infty) = (I - D\hat{D})$ is invertible. Therefore,

$$\begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix}$$

is nonsingular. Now routine calculations give the transfer matrix from $\begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ to $\begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$ in terms of the state space realizations:

$$\begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix}^{-1} \left\{ \begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix} + \begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix} (sI - \tilde{A})^{-1} \begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix} \right\} \begin{bmatrix} I & -\hat{D} \\ -D & I \end{bmatrix}^{-1}.$$

Since the above transfer matrix belongs to $\mathcal{RH}_\infty$, it follows that

$$\begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix} (sI - \tilde{A})^{-1} \begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix}$$

as a transfer matrix belongs to $\mathcal{RH}_\infty$. Finally, since (A, B, C) and $(\hat{A}, \hat{B}, \hat{C})$ are stabilizable and detectable,

$$\left(\tilde{A}, \begin{bmatrix} B & 0 \\ 0 & \hat{B} \end{bmatrix}, \begin{bmatrix} 0 & \hat{C} \\ C & 0 \end{bmatrix} \right)$$

is stabilizable and detectable. It then follows that the eigenvalues of $\tilde{A}$ are in the open left-half plane. $\square$

Note that to check internal stability, it is necessary (and sufficient) to test whether each of the four transfer matrices in (5.11) is in $\mathcal{RH}_\infty$. Stability cannot be concluded even if three of the four transfer matrices in (5.11) are in $\mathcal{RH}_\infty$. For example, let an interconnected system transfer function be given by

$$P = \frac{s-1}{s+1}, \qquad \hat{K} = -\frac{1}{s-1}.$$

Then it is easy to compute

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} = \begin{bmatrix} \dfrac{s+1}{s+2} & -\dfrac{s+1}{(s-1)(s+2)} \\ \dfrac{s-1}{s+2} & \dfrac{s+1}{s+2} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix},$$

which shows that the system is not internally stable although three of the four transfer functions are stable. This can also be seen by calculating the closed-loop A-matrix with any stabilizable and detectable realizations of P and $\hat{K}$.

Remark 5.1 It should be noted that internal stability is a basic requirement for a practical feedback system. This is because all interconnected systems may be unavoidably subject to some nonzero initial conditions and some (possibly small) errors, and it cannot be tolerated in practice that such errors at some locations will lead to unbounded signals at some other locations in the closed-loop system. Internal stability guarantees that all signals in a system are bounded provided that the injected signals (at any locations) are bounded. ♡

However, there are some special cases under which determining system stability is simple.

Corollary 5.4 *Suppose $\hat{K} \in \mathcal{RH}_\infty$. Then the system in Figure 5.2 is internally stable iff $(I - P\hat{K})^{-1}P \in \mathcal{RH}_\infty$.*

Proof. The necessity is obvious. To prove the sufficiency, it is sufficient to show that $(I - P\hat{K})^{-1} \in \mathcal{RH}_\infty$. But this follows from

$$(I - P\hat{K})^{-1} = I + (I - P\hat{K})^{-1}P\hat{K}$$

and $(I - P\hat{K})^{-1}P, \hat{K} \in \mathcal{RH}_\infty$. □

This corollary is in fact the basis for the classical control theory where the stability is checked only for one closed-loop transfer function with the implicit assumption that the controller itself is stable. Also, we have

Corollary 5.5 *Suppose $P \in \mathcal{RH}_\infty$. Then the system in Figure 5.2 is internally stable iff $\hat{K}(I - P\hat{K})^{-1} \in \mathcal{RH}_\infty$.*

Corollary 5.6 *Suppose $P \in \mathcal{RH}_\infty$ and $\hat{K} \in \mathcal{RH}_\infty$. Then the system in Figure 5.2 is internally stable iff $(I - P\hat{K})^{-1} \in \mathcal{RH}_\infty$.*

To study the more general case, define

$$\begin{aligned} n_c &:= \text{number of open rhp poles of } \hat{K}(s) \\ n_p &:= \text{number of open rhp poles of } P(s). \end{aligned}$$

Theorem 5.7 *The system is internally stable if and only if*

(i) the number of open rhp poles of $P(s)\hat{K}(s) = n_c + n_p$;

(ii) $\phi(s) := \det(I - P(s)\hat{K}(s))$ has all its zeros in the open left-half plane (i.e., $(I - P(s)\hat{K}(s))^{-1}$ is stable).

Proof. It is easy to show that $P\hat{K}$ and $(I - P\hat{K})^{-1}$ have the following realizations:

$$P\hat{K} = \left[\begin{array}{cc|c} A & B\hat{C} & B\hat{D} \\ 0 & \hat{A} & \hat{B} \\ \hline C & D\hat{C} & D\hat{D} \end{array}\right]$$

$$(I - P\hat{K})^{-1} = \left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline \bar{C} & \bar{D} \end{array}\right]$$

where

$$\begin{aligned} \bar{A} &= \begin{bmatrix} A & B\hat{C} \\ 0 & \hat{A} \end{bmatrix} + \begin{bmatrix} B\hat{D} \\ \hat{B} \end{bmatrix} (I - D\hat{D})^{-1} \begin{bmatrix} C & D\hat{C} \end{bmatrix} \\ \bar{B} &= \begin{bmatrix} B\hat{D} \\ \hat{B} \end{bmatrix} (I - D\hat{D})^{-1} \\ \bar{C} &= (I - D\hat{D})^{-1} \begin{bmatrix} C & D\hat{C} \end{bmatrix} \\ \bar{D} &= (I - D\hat{D})^{-1}. \end{aligned}$$

It is also easy to see that $\bar{A} = \tilde{A}$. Hence, the system is internally stable iff $\bar{A}$ is stable.

Now suppose that the system is internally stable, then $(I - P\hat{K})^{-1} \in \mathcal{RH}_\infty$. This implies that all zeros of $\det(I - P(s)\hat{K}(s))$ must be in the left-half plane. So we only need to show that given condition (ii), condition (i) is necessary and sufficient for the internal stability. This follows by noting that $(\bar{A}, \bar{B})$ is stabilizable iff

$$\left(\begin{bmatrix} A & B\hat{C} \\ 0 & \hat{A} \end{bmatrix}, \begin{bmatrix} B\hat{D} \\ \hat{B} \end{bmatrix} \right) \tag{5.12}$$

is stabilizable; and $(\bar{C}, \bar{A})$ is detectable iff

$$\left(\begin{bmatrix} C & D\hat{C} \end{bmatrix}, \begin{bmatrix} A & B\hat{C} \\ 0 & \hat{A} \end{bmatrix} \right) \tag{5.13}$$

is detectable. But conditions (5.12) and (5.13) are equivalent to condition (i), i.e., $P\hat{K}$ has no unstable pole/zero cancelations. $\square$

With this observation, the MIMO version of the Nyquist stability theorem is obvious.

Theorem 5.8 (Nyquist Stability Theorem) *The system is internally stable if and only if condition (i) in Theorem 5.7 is satisfied and the Nyquist plot of $\phi(j\omega)$ for $-\infty \leq \omega \leq \infty$ encircles the origin, $(0, 0)$, $n_k + n_p$ times in the counter-clockwise direction.*

Proof. Note that by SISO Nyquist stability theorem, $\phi(s)$ has all zeros in the open left-half plane if and only if the Nyquist plot of $\phi(j\omega)$ for $-\infty \leq \omega \leq \infty$ encircles the origin, $(0, 0)$, $n_k + n_p$ times in the counter-clockwise direction. $\square$

5.4 Coprime Factorization over $\mathcal{RH}_\infty$

Recall that two polynomials $m(s)$ and $n(s)$, with, for example, real coefficients, are said to be *coprime* if their greatest common divisor is 1 (equivalent, they have no common zeros). It follows from Euclid's algorithm[1] that two polynomials m and n are coprime iff there exist polynomials $x(s)$ and $y(s)$ such that $xm + yn = 1$; such an equation is called a Bezout identity. Similarly, two transfer functions $m(s)$ and $n(s)$ in $\mathcal{RH}_\infty$ are said to be *coprime over* $\mathcal{RH}_\infty$ if there exists $x, y \in \mathcal{RH}_\infty$ such that

$$xm + yn = 1.$$

The more primitive, but equivalent, definition is that m and n are coprime if every common divisor of m and n is invertible in $\mathcal{RH}_\infty$, i.e.,

$$h, mh^{-1}, nh^{-1} \in \mathcal{RH}_\infty \Longrightarrow h^{-1} \in \mathcal{RH}_\infty.$$

More generally, we have

Definition 5.3 Two matrices M and N in $\mathcal{RH}_\infty$ are *right coprime over* $\mathcal{RH}_\infty$ if they have the same number of columns and if there exist matrices X_r and Y_r in $\mathcal{RH}_\infty$ such that

$$\begin{bmatrix} X_r & Y_r \end{bmatrix} \begin{bmatrix} M \\ N \end{bmatrix} = X_r M + Y_r N = I.$$

Similarly, two matrices $\tilde{M}$ and $\tilde{N}$ in $\mathcal{RH}_\infty$ are *left coprime over* $\mathcal{RH}_\infty$ if they have the same number of rows and if there exist matrices X_l and Y_l in $\mathcal{RH}_\infty$ such that

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} \begin{bmatrix} X_l \\ Y_l \end{bmatrix} = \tilde{M} X_l + \tilde{N} Y_l = I.$$

Note that these definitions are equivalent to saying that the matrix $\begin{bmatrix} M \\ N \end{bmatrix}$ is left-invertible in $\mathcal{RH}_\infty$ and the matrix $\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}$ is right-invertible in $\mathcal{RH}_\infty$. These two equations are often called Bezout identities.

Now let P be a proper real-rational matrix. A *right-coprime factorization (rcf)* of P is a factorization $P = NM^{-1}$ where N and M are right-coprime over $\mathcal{RH}_\infty$. Similarly, a *left-coprime factorization (lcf)* has the form $P = \tilde{M}^{-1}\tilde{N}$ where $\tilde{N}$ and $\tilde{M}$ are left-coprime over $\mathcal{RH}_\infty$. A matrix $P(s) \in \mathcal{R}_p(s)$ is said to have *double coprime factorization* if there exist a right coprime factorization $P = NM^{-1}$, a left coprime factorization $P = \tilde{M}^{-1}\tilde{N}$, and X_r, Y_r, X_l, $Y_l \in \mathcal{RH}_\infty$ such that

$$\begin{bmatrix} X_r & Y_r \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & -Y_l \\ N & X_l \end{bmatrix} = I. \tag{5.14}$$

Of course implicit in these definitions is the requirement that both M and $\tilde{M}$ be square and nonsingular.

[1] See, e.g., [Kailath, 1980, pp. 140-141].

Theorem 5.9 *Suppose $P(s)$ is a proper real-rational matrix and*

$$P = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

is a stabilizable and detectable realization. Let F and L be such that $A+BF$ and $A+LC$ are both stable, and define

$$\begin{bmatrix} M & -Y_l \\ N & X_l \end{bmatrix} = \left[\begin{array}{c|cc} A+BF & B & -L \\ \hline F & I & 0 \\ C+DF & D & I \end{array}\right] \tag{5.15}$$

$$\begin{bmatrix} X_r & Y_r \\ -\tilde{N} & \tilde{M} \end{bmatrix} = \left[\begin{array}{c|cc} A+LC & -(B+LD) & L \\ \hline F & I & 0 \\ C & -D & I \end{array}\right]. \tag{5.16}$$

Then $P = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ are rcf *and* lcf*, respectively, and, furthermore, (5.14) is satisfied.*

Proof. The theorem follows by verifying the equation (5.14). □

Remark 5.2 Note that if P is stable, then we can take $X_r = X_l = I$, $Y_r = Y_l = 0$, $N = \tilde{N} = P$, $M = \tilde{M} = I$. ♡

Remark 5.3 The coprime factorization of a transfer matrix can be given a feedback control interpretation. For example, right coprime factorization comes out naturally from changing the control variable by a state feedback. Consider the state space equations for a plant P:

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du. \end{aligned}$$

Next, introduce a state feedback and change the variable

$$v := u - Fx$$

where F is such that $A + BF$ is stable. Then we get

$$\begin{aligned} \dot{x} &= (A+BF)x + Bv \\ u &= Fx + v \\ y &= (C+DF)x + Dv. \end{aligned}$$

Evidently from these equations, the transfer matrix from v to u is

$$M(s) = \left[\begin{array}{c|c} A+BF & B \\ \hline F & I \end{array}\right],$$

and that from v to y is

$$N(s) = \left[\begin{array}{c|c} A+BF & B \\ \hline C+DF & D \end{array}\right].$$

Therefore

$$u = Mv, \quad y = Nv$$

so that $y = NM^{-1}u$, i.e., $P = NM^{-1}$. ♡

We shall now see how coprime factorizations can be used to obtain alternative characterizations of internal stability conditions. Consider again the standard stability analysis diagram in Figure 5.2. We begin with any rcf's and lcf's of P and $\hat{K}$:

$$P = NM^{-1} = \tilde{M}^{-1}\tilde{N} \tag{5.17}$$

$$\hat{K} = UV^{-1} = \tilde{V}^{-1}\tilde{U}. \tag{5.18}$$

Lemma 5.10 *Consider the system in Figure 5.2. The following conditions are equivalent:*

1. *The feedback system is internally stable.*
2. $\begin{bmatrix} M & U \\ N & V \end{bmatrix}$ *is invertible in* $\mathcal{RH}_\infty$.
3. $\begin{bmatrix} \tilde{V} & -\tilde{U} \\ -\tilde{N} & \tilde{M} \end{bmatrix}$ *is invertible in* $\mathcal{RH}_\infty$.
4. $\tilde{M}V - \tilde{N}U$ *is invertible in* $\mathcal{RH}_\infty$.
5. $\tilde{V}M - \tilde{U}N$ *is invertible in* $\mathcal{RH}_\infty$.

Proof. As we saw in Lemma 5.3, internal stability is equivalent to

$$\begin{bmatrix} I & -\hat{K} \\ -P & I \end{bmatrix}^{-1} \in \mathcal{RH}_\infty$$

or, equivalently,

$$\begin{bmatrix} I & \hat{K} \\ P & I \end{bmatrix}^{-1} \in \mathcal{RH}_\infty. \tag{5.19}$$

Now

$$\begin{bmatrix} I & \hat{K} \\ P & I \end{bmatrix} = \begin{bmatrix} I & UV^{-1} \\ NM^{-1} & I \end{bmatrix} = \begin{bmatrix} M & U \\ N & V \end{bmatrix} \begin{bmatrix} M^{-1} & 0 \\ 0 & V^{-1} \end{bmatrix}$$

so that

$$\begin{bmatrix} I & \hat{K} \\ P & I \end{bmatrix}^{-1} = \begin{bmatrix} M & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} M & U \\ N & V \end{bmatrix}^{-1}.$$

Since the matrices

$$\begin{bmatrix} M & 0 \\ 0 & V \end{bmatrix}, \begin{bmatrix} M & U \\ N & V \end{bmatrix}$$

are right-coprime (this fact is left as an exercise for the reader), (5.19) holds iff

$$\begin{bmatrix} M & U \\ N & V \end{bmatrix}^{-1} \in \mathcal{RH}_\infty.$$

This proves the equivalence of conditions 1 and 2. The equivalence of 1 and 3 is proved similarly.

The conditions 4 and 5 are implied by 2 and 3 from the following equation:

$$\begin{bmatrix} \tilde{V} & -\tilde{U} \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & U \\ N & V \end{bmatrix} = \begin{bmatrix} \tilde{V}M - \tilde{U}N & 0 \\ 0 & \tilde{M}V - \tilde{N}U \end{bmatrix}.$$

Since the left hand side of the above equation is invertible in $\mathcal{RH}_\infty$, so is the right hand side. Hence, conditions 4 and 5 are satisfied. We only need to show that either condition 4 or condition 5 implies condition 1. Let us show condition $5 \rightarrow 1$; this is obvious since

$$\begin{aligned} \begin{bmatrix} I & \hat{K} \\ P & I \end{bmatrix}^{-1} &= \begin{bmatrix} I & \tilde{V}^{-1}\tilde{U} \\ NM^{-1} & I \end{bmatrix}^{-1} \\ &= \begin{bmatrix} M & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} \tilde{V}M & \tilde{U} \\ N & I \end{bmatrix}^{-1} \begin{bmatrix} \tilde{V} & 0 \\ 0 & I \end{bmatrix} \in \mathcal{RH}_\infty \end{aligned}$$

if $\begin{bmatrix} \tilde{V}M & \tilde{U} \\ N & I \end{bmatrix}^{-1} \in \mathcal{RH}_\infty$ or if condition 5 is satisfied. □

Combining Lemma 5.10 and Theorem 5.9, we have the following corollary.

Corollary 5.11 *Let P be a proper real-rational matrix and $P = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ be corresponding rcf and lcf over $\mathcal{RH}_\infty$. Then there exists a controller*

$$\hat{K}_0 = U_0 V_0^{-1} = \tilde{V}_0^{-1}\tilde{U}_0$$

with $U_0, V_0, \tilde{U}_0$, and $\tilde{V}_0$ in $\mathcal{RH}_\infty$ such that

$$\begin{bmatrix} \tilde{V}_0 & -\tilde{U}_0 \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}. \tag{5.20}$$

Furthermore, let F and L be such that $A+BF$ and $A+LC$ are stable. Then a particular set of state space realizations for these matrices can be given by

$$\begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} = \left[\begin{array}{c|cc} A+BF & B & -L \\ \hline F & I & 0 \\ C+DF & D & I \end{array}\right] \tag{5.21}$$

$$\begin{bmatrix} \tilde{V}_0 & -\tilde{U}_0 \\ -\tilde{N} & \tilde{M} \end{bmatrix} = \left[\begin{array}{c|cc} A+LC & -(B+LD) & L \\ \hline F & I & 0 \\ C & -D & I \end{array}\right]. \tag{5.22}$$

Proof. The idea behind the choice of these matrices is as follows. Using the observer theory, find a controller $\hat{K}_0$ achieving internal stability; for example

$$\hat{K}_0 := \left[\begin{array}{c|c} A+BF+LC+LDF & -L \\ \hline F & 0 \end{array}\right]. \tag{5.23}$$

Perform factorizations

$$\hat{K}_0 = U_0 V_0^{-1} = \tilde{V}_0^{-1}\tilde{U}_0$$

which are analogous to the ones performed on P. Then Lemma 5.10 implies that each of the two left-hand side block matrices of (5.20) must be invertible in $\mathcal{RH}_\infty$. In fact, (5.20) is satisfied by comparing it with the equation (5.14). □

5.5 Feedback Properties

In this section, we discuss the properties of a feedback system. In particular, we consider the benefit of the feedback structure and the concept of design tradeoffs for conflicting objectives – namely, how to achieve the benefits of feedback in the face of uncertainties.

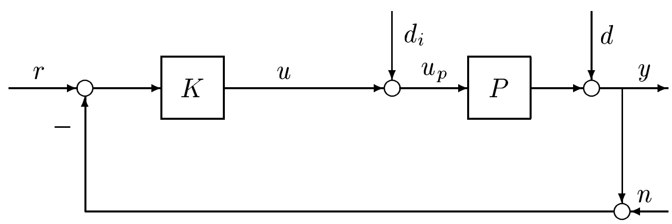


Figure 5.3: Standard Feedback Configuration

Consider again the feedback system shown in Figure 5.1. For convenience, the system diagram is shown again in Figure 5.3. For further discussion, it is convenient to define the *input loop transfer matrix*, L_i, and *output loop transfer matrix*, L_o, as

$$L_i = KP, \quad L_o = PK,$$

respectively, where L_i is obtained from breaking the loop at the input (u) of the plant while L_o is obtained from breaking the loop at the output (y) of the plant. The *input sensitivity* matrix is defined as the transfer matrix from d_i to u_p:

$$S_i = (I+L_i)^{-1}, \quad u_p = S_i d_i.$$

And the *output sensitivity* matrix is defined as the transfer matrix from d to y:

$$S_o = (I+L_o)^{-1}, \quad y = S_o d.$$

The *input* and *output complementary sensitivity* matrices are defined as

$$T_i = I - S_i = L_i(I + L_i)^{-1}$$

$$T_o = I - S_o = L_o(I + L_o)^{-1},$$

respectively. (The word *complementary* is used to signify the fact that T is the complement of S, $T = I - S$.) The matrix $I + L_i$ is called *input return difference matrix* and $I + L_o$ is called *output return difference matrix.*

It is easy to see that the closed-loop system, if it is internally stable, satisfies the following equations:

$$\begin{aligned} y &= T_o(r-n) + S_oPd_i + S_od && (5.24)\\ r-y &= S_o(r-d) + T_on - S_oPd_i && (5.25)\\ u &= KS_o(r-n) - KS_od - T_id_i && (5.26)\\ u_p &= KS_o(r-n) - KS_od + S_id_i. && (5.27)\end{aligned}$$

These four equations show the fundamental benefits and design objectives inherent in feedback loops. For example, equation (5.24) shows that the effects of disturbance d on the plant output can be made "small" by making the output sensitivity function S_o small. Similarly, equation (5.27) shows that the effects of disturbance d_i on the plant input can be made small by making the input sensitivity function S_i small. The notion of smallness for a transfer matrix in a certain range of frequencies can be made explicit using frequency dependent singular values, for example, $\overline{\sigma}(S_o) < 1$ over a frequency range would mean that the effects of disturbance d at the plant output are effectively desensitized over that frequency range.

Hence, good disturbance rejection at the plant output (y) would require that

$$\begin{aligned} \overline{\sigma}(S_o) &= \overline{\sigma}\left((I+PK)^{-1}\right) = \frac{1}{\underline{\sigma}(I+PK)}, && \text{(for disturbance at plant output, } d)\\ \overline{\sigma}(S_oP) &= \overline{\sigma}\left((I+PK)^{-1}P\right) = \overline{\sigma}(PS_i), && \text{(for disturbance at plant input, } d_i)\end{aligned}$$

be made small and good disturbance rejection at the plant input (u_p) would require that

$$\begin{aligned} \overline{\sigma}(S_i) &= \overline{\sigma}\left((I+KP)^{-1}\right) = \frac{1}{\underline{\sigma}(I+KP)}, && \text{(for disturbance at plant input, } d_i)\\ \overline{\sigma}(S_iK) &= \overline{\sigma}\left(K(I+PK)^{-1}\right) = \overline{\sigma}(KS_o), && \text{(for disturbance at plant output, } d)\end{aligned}$$

be made small, particularly in the low frequency range where d and d_i are usually significant.

Note that

$$\begin{aligned} \underline{\sigma}(PK) - 1 &\leq \underline{\sigma}(I+PK) \leq \underline{\sigma}(PK) + 1\\ \underline{\sigma}(KP) - 1 &\leq \underline{\sigma}(I+KP) \leq \underline{\sigma}(KP) + 1\end{aligned}$$

then

$$\frac{1}{\underline{\sigma}(PK)+1} \leq \overline{\sigma}(S_o) \leq \frac{1}{\underline{\sigma}(PK)-1}, \quad \text{if} \quad \underline{\sigma}(PK) > 1$$
$$\frac{1}{\underline{\sigma}(KP)+1} \leq \overline{\sigma}(S_i) \leq \frac{1}{\underline{\sigma}(KP)-1}, \quad \text{if} \quad \underline{\sigma}(KP) > 1.$$

These equations imply that

$$\overline{\sigma}(S_o) \ll 1 \iff \underline{\sigma}(PK) \gg 1$$
$$\overline{\sigma}(S_i) \ll 1 \iff \underline{\sigma}(KP) \gg 1.$$

Now suppose P and K are invertible, then

$$\underline{\sigma}(PK) \gg 1 \text{ or } \underline{\sigma}(KP) \gg 1 \iff \overline{\sigma}(S_oP) = \overline{\sigma}\left((I+PK)^{-1}P\right) \approx \overline{\sigma}(K^{-1}) = \frac{1}{\underline{\sigma}(K)}$$
$$\underline{\sigma}(PK) \gg 1 \text{ or } \underline{\sigma}(KP) \gg 1 \iff \overline{\sigma}(KS_o) = \overline{\sigma}\left(K(I+PK)^{-1}\right) \approx \overline{\sigma}(P^{-1}) = \frac{1}{\underline{\sigma}(P)}.$$

Hence good performance at plant output (y) requires in general large output loop gain $\underline{\sigma}(L_o) = \underline{\sigma}(PK) \gg 1$ in the frequency range where d is significant for desensitizing d and large enough controller gain $\underline{\sigma}(K) \gg 1$ in the frequency range where d_i is significant for desensitizing d_i. Similarly, good performance at plant input (u_p) requires in general large input loop gain $\underline{\sigma}(L_i) = \underline{\sigma}(KP) \gg 1$ in the frequency range where d_i is significant for desensitizing d_i and large enough plant gain $\underline{\sigma}(P) \gg 1$ in the frequency range where d is significant, which can not changed by controller design, for desensitizing d. (It should be noted that in general $S_o \neq S_i$ unless K and P are square and diagonal which is true if P is a scalar system. Hence, small $\overline{\sigma}(S_o)$ does not necessarily imply small $\overline{\sigma}(S_i)$; in other words, good disturbance rejection at the output does not necessarily mean good disturbance rejection at the plant input.)

Hence, *good multivariable feedback loop design boils down to achieving high loop (and possibly controller) gains in the necessary frequency range.*

Despite the simplicity of this statement, feedback design is by no means trivial. This is true because loop gains cannot be made arbitrarily high over arbitrarily large frequency ranges. Rather, they must satisfy certain performance tradeoff and design limitations. A major performance tradeoff, for example, concerns commands and disturbance error reduction versus stability under the model uncertainty. Assume that the plant model is perturbed to $(I + \Delta)P$ with Δ stable, and assume that the system is nominally stable, i.e., the closed-loop system with $\Delta = 0$ is stable. Now the perturbed closed-loop system is stable if

$$\det\left(I + (I+\Delta)PK\right) = \det(I+PK)\det(I+\Delta T_o)$$

has no right-half plane zero. This would in general amount to requiring that $\|\Delta T_o\|$ be small or that $\bar{\sigma}(T_o)$ be small at those frequencies where Δ is significant, typically at high frequency range, which in turn implies that the loop gain, $\overline{\sigma}(L_o)$, should be small at those frequencies.

Still another tradeoff is with the sensor noise error reduction. The conflict between the disturbance rejection and the sensor noise reduction is evident in equation (5.24). Large $\underline{\sigma}(L_o(j\omega))$ values over a large frequency range make errors due to d small. However, they also make errors due to n large because this noise is "passed through" over the same frequency range, i.e.,

$$y = T_o(r-n) + S_oPd_i + S_od \approx (r-n).$$

Note that n is typically significant in the high frequency range. Worst still, large loop gains outside of the bandwidth of P, i.e., $\underline{\sigma}(L_o(j\omega)) \gg 1$ or $\underline{\sigma}(L_i(j\omega)) \gg 1$ while $\overline{\sigma}(P(j\omega)) \ll 1$, can make the control activity (u) quite unacceptable, which may cause the saturation of actuators. This follows from

$$u = KS_o(r-n-d) - T_id_i = S_iK(r-n-d) - T_id_i \approx P^{-1}(r-n-d) - d_i.$$

Here, we have assumed P to be square and invertible for convenience. The resulting equation shows that disturbances and sensor noise are actually amplified at u whenever the frequency range significantly exceeds the bandwidth of P, since for ω such that $\overline{\sigma}(P(j\omega)) \ll 1$, we have

$$\underline{\sigma}[P^{-1}(j\omega)] = \frac{1}{\overline{\sigma}[P(j\omega)]} \gg 1.$$

Similarly, the controller gain, $\overline{\sigma}(K)$, should also be kept not too large in the frequency range where the loop gain is small in order to not saturate the actuators. This is because for small loop gain $\overline{\sigma}(L_o(j\omega)) \ll 1$ or $\overline{\sigma}(L_i(j\omega)) \ll 1$

$$u = KS_o(r-n-d) - T_id_i \approx K(r-n-d).$$

Therefore, it is desirable to keep $\overline{\sigma}(K)$ not too large when the loop gain is small.

To summarize the above discussion, we note that good performance requires in some frequency range, typically some low frequency range $(0, \omega_l)$:

$$\underline{\sigma}(PK) \gg 1, \quad \underline{\sigma}(KP) \gg 1, \quad \underline{\sigma}(K) \gg 1$$

and good robustness and good sensor noise rejection require in some frequency range, typically some high frequency range (ω_h, ∞)

$$\overline{\sigma}(PK) \ll 1, \quad \overline{\sigma}(KP) \ll 1, \quad \overline{\sigma}(K) \le M$$

where M is not too large. These design requirements are shown graphically in Figure 5.4. The specific frequencies ω_l and ω_h depend on the specific applications and the knowledge one has on the disturbance characteristics, the modeling uncertainties, and the sensor noise levels.

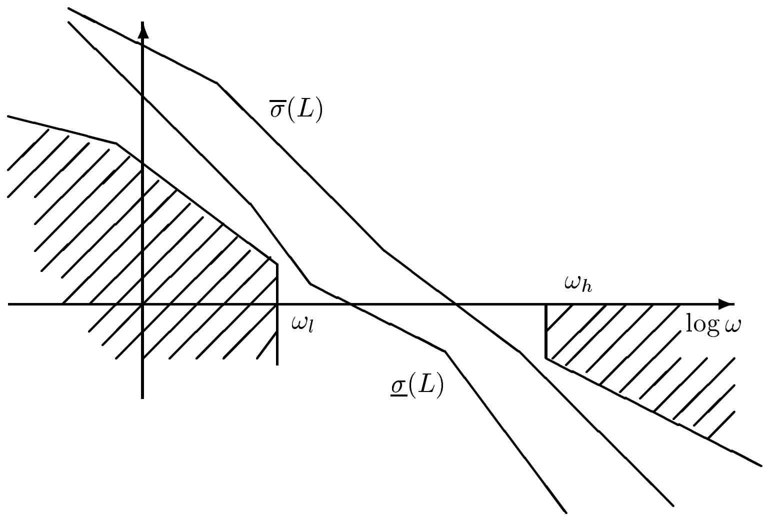


Figure 5.4: Desired Loop Gain

5.6 The Concept of Loop Shaping

The analysis in the last section motivates a conceptually simple controller design technique: loop shaping. Loop shaping controller design involves essentially finding a controller K that shapes the loop transfer function L so that the loop gains, $\underline{\sigma}(L)$ and $\overline{\sigma}(L)$, clear the boundaries specified by the performance requirements at low frequencies and by the robustness requirements at high frequencies as shown in Figure 5.4.

In the SISO case, the loop shaping design technique is particularly effective and simple since $\overline{\sigma}(L) = \underline{\sigma}(L) = |L|$. The design procedure can be completed in two steps:

SISO Loop Shaping

(1) Find a rational strictly proper transfer function L which contains all the right half plane poles and zeros of P such that $|L|$ clears the boundaries specified by the performance requirements at low frequencies and by the robustness requirements at high frequencies as shown in Figure 5.4.

L must also be chosen so that $1+L$ has all zeros in the open left half plane, which can usually be guaranteed by making L well-behaved in the crossover region, i.e., L should not be decreasing too fast in the frequency range of $|L(j\omega)| \approx 1$.

(2) The controller is given by $K = L/P$.

The loop shaping for MIMO system can be done similarly if the singular values of the loop transfer functions are used for the loop gains.

MIMO Loop Shaping

(1) Find a rational strictly proper transfer function L which contains all the right half plane poles and zeros of P so that the product of P and $P^{-1}L$ (or LP^{-1}) has no unstable poles and/or zeros cancelations, and $\underline{\sigma}(L)$ clears the boundary specified by the performance requirements at low frequencies and $\overline{\sigma}(L)$ clears the boundary specified by the robustness requirements at high frequencies as shown in Figure 5.4.

L must also be chosen so that $\det(I + L)$ has all zeros in the open left half plane. (This is not easy for MIMO systems.)

(2) The controller is given by $K = P^{-1}L$ if L is the output loop transfer function (or $K = LP^{-1}$ if L is the input loop transfer function).

The loop shaping design technique can be quite useful especially for SISO control system design. However, there are severe limitations when it is used for MIMO system design.

Limitations of MIMO Loop Shaping

Although the above loop shaping design can be effective in some of applications, there are severe intrinsic limitations. Some of these limitations are listed below:

- The loop shaping technique described above can only effectively deal with problems with uniformly specified performance and robustness specifications. More specifically, the method can not effectively deal with problems with different specifications in different channels and/or problems with different uncertainty characteristics in different channels without introducing significant conservatism. To illustrate this difficulty, consider an uncertain dynamical system

  $$P_\Delta = (I + \Delta)P$$

  where P is the nominal plant and Δ is the multiplicative modeling error. Assume that Δ can be written in the following form

  $$\Delta = \tilde{\Delta} W_t, \quad \overline{\sigma}(\tilde{\Delta}) < 1.$$

  Then, for robust stability, we would require $\overline{\sigma}(\Delta T_o) = \overline{\sigma}(\tilde{\Delta} W_t T_o) < 1$ or $\overline{\sigma}(W_t T_o) \leq 1$. If a uniform bound is required on the loop gain to apply the loop shaping technique, we would need to overbound $\overline{\sigma}(W_t T_o)$:

  $$\overline{\sigma}(W_t T_o) \leq \overline{\sigma}(W_t)\overline{\sigma}(T_o) \leq \overline{\sigma}(W_t)\frac{\overline{\sigma}(L_o)}{1 - \overline{\sigma}(L_o)}, \quad \text{if} \quad \overline{\sigma}(L_o) < 1$$

  and the robust stability requirement is implied by

  $$\overline{\sigma}(L_o) \leq \frac{1}{\overline{\sigma}(W_t) + 1} \approx \frac{1}{\overline{\sigma}(W_t)}, \quad \text{if} \quad \overline{\sigma}(L_o) < 1.$$

Similarly, if the performance requirements, say output disturbance rejection, are not uniformly specified in all channels but by a weighting matrix W_s such that $\overline{\sigma}(W_s S_o) \leq 1$, then it is also necessary to overbound $\overline{\sigma}(W_s S_o)$ in order to apply the loop shaping techniques:

$$\overline{\sigma}(W_s S_o) \leq \overline{\sigma}(W_s)\overline{\sigma}(S_o) \leq \frac{\overline{\sigma}(W_s)}{\underline{\sigma}(L_o) - 1}, \quad \text{if} \quad \underline{\sigma}(L_o) > 1$$

and the performance requirement is implied by

$$\underline{\sigma}(L_o) \geq \overline{\sigma}(W_s) + 1 \approx \overline{\sigma}(W_s), \quad \text{if} \quad \underline{\sigma}(L_o) > 1.$$

It is possible that the bounds for the loop shape may contradict each other at some frequency range, as shown in the figure 5.5. However, this does not imply that there is no controller that will satisfy both nominal performance and robust stability except for SISO systems. This contradiction happens because the bounds

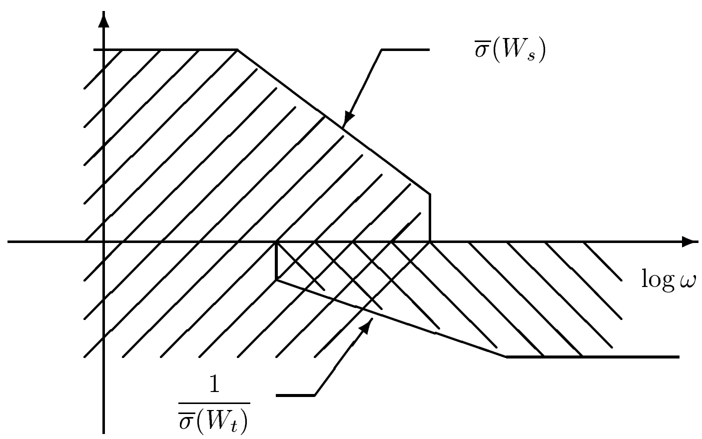


Figure 5.5: Conflict Requirements

do not utilize the structure of weights, W_s and W_t; and the bounds are only sufficient conditions for robust stability and nominal performance. This possibility can be further illustrated by the following example: Assume that a two-input and two-output system transfer matrix is given by

$$P(s) = \frac{1}{s+1} \begin{bmatrix} 1 & \alpha \\ 0 & 1 \end{bmatrix},$$

and suppose the weighting matrices are given by

$$W_s = \begin{bmatrix} \frac{1}{s+1} & \frac{\alpha}{(s+1)(s+2)} \\ 0 & \frac{1}{s+1} \end{bmatrix}, \quad W_t = \begin{bmatrix} \frac{s+2}{s+10} & \frac{\alpha(s+1)}{s+10} \\ 0 & \frac{s+2}{s+10} \end{bmatrix}.$$

It is easy to show that for large α, the weighting functions are as shown in Figure 5.5, and thus the above loop shaping technique cannot be applied. However, it is also easy to show that the system with controller

$$K = I_2$$

gives

$$W_s S = \begin{bmatrix} \frac{1}{s+2} & 0 \\ 0 & \frac{1}{s+2} \end{bmatrix}, \quad W_t T = \begin{bmatrix} \frac{1}{s+10} & 0 \\ 0 & \frac{1}{s+10} \end{bmatrix}$$

and, therefore, the robust performance criterion is satisfied.

- Even if all of the above problems can be avoided, it may still be difficult to find a *matrix function* L_o so that $K = P^{-1} L_o$ is stabilizing. This becomes much harder if P is non-minimum phase and/or unstable.

Hence some new methodologies have to be introduced to solve complicated problems. The so-called LQG/LTR (Linear Quadratic Gaussian/Loop Transfer Recovery) procedure developed first by Doyle and Stein [1981] and extended later by various authors can solve some of the problems, but it is essentially limited to nominally minimum phase and output multiplicative uncertain systems. For these reasons, it will not be introduced here. This motivates us to consider the closed-loop performance directly in terms of the closed-loop transfer functions instead of open loop transfer functions. The following section considers some simple closed-loop performance problem formulations.

5.7 Weighted $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Performance

In this section, we consider how to formulate some performance objectives into mathematically tractable problems. As shown in section 5.5, the performance objectives of a feedback system can usually be specified in terms of requirements on the sensitivity functions and/or complementary sensitivity functions or in terms of some other closed-loop transfer functions. For instance, the performance criteria for a scalar system may be specified as requiring

$$\begin{cases} |s(j\omega)| \le \alpha < 1 & \forall \omega \le \omega_0, \\ |s(j\omega)| \le \beta > 1 & \forall \omega > \omega_0 \end{cases}$$

where $s(j\omega) = \frac{1}{1+p(j\omega)k(j\omega)}$. However, it is much more convenient to reflect the system performance objectives by choosing appropriate weighting functions. For example, the above performance objective can be written as

$$|w_s(j\omega)s(j\omega)| \le 1, \quad \forall \omega$$

with

$$|w_s(j\omega)| = \begin{cases} \alpha^{-1} & \forall \omega \le \omega_0, \\ \beta^{-1} & \forall \omega > \omega_0. \end{cases}$$

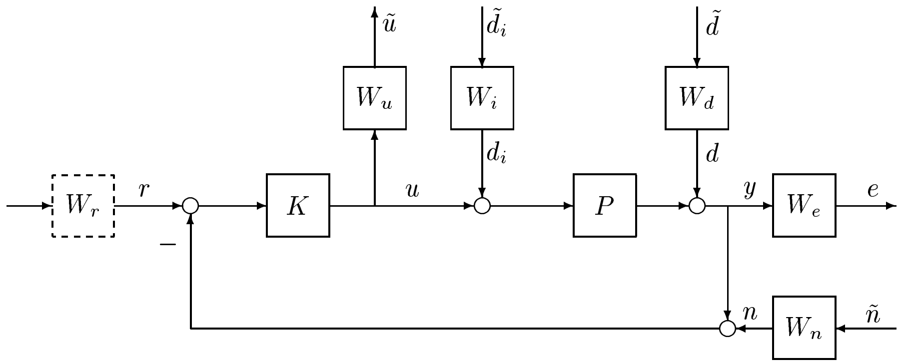


Figure 5.6: Standard Feedback Configuration with Weights

In order to use w_s in control design, a rational transfer function w_s is usually used to approximate the above frequency response.

The advantage of using weighted performance specifications is obvious in multivariable system design. First of all, some components of a vector signal are usually more important than others. Secondly, each component of the signal may not be measured in the same metric; for example, some components of the output error signal may be measured in terms of length, and the others may be measured in terms of voltage. Therefore, weighting functions are essential to make these components comparable. Also, we might be primarily interested in rejecting errors in a certain frequency range (for example low frequencies), hence some frequency dependent weights must be chosen.

In general, we shall modify the standard feedback diagram in Figure 5.3 into Figure 5.6. The weighting functions in Figure 5.6 are chosen to reflect the design objectives and knowledge on the disturbances and sensor noise. For example, W_d and W_i may be chosen to reflect the frequency contents of the disturbances d and d_i or they may be used to model the disturbance power spectrum depending on the nature of signals involved in the practical systems. The weighting matrix W_n is used to model the frequency contents of the sensor noise while W_e may be used to reflect the requirements on the shape of certain closed-loop transfer functions, for example, the shape of the output sensitivity function. Similarly, W_u may be used to reflect some restrictions on the control or actuator signals, and the dashed precompensator W_r is an optional element used to achieve deliberate command shaping or to represent a non-unity feedback system in equivalent unity feedback form.

It is, in fact, essential that some appropriate weighting matrices be used in order to utilize the optimal control theory discussed in this book, i.e., $\mathcal{H}_2$ and $\mathcal{H}_\infty$ theory. So a very important step in controller design process is to choose the appropriate weights, W_e, W_d, W_u, and possibly W_n, W_i, and W_r. The appropriate choice of weights for a particular practical problem is not trivial. In many occasions, as in the scalar case, the weights are chosen purely as a design parameter without any physical bases, so these weights may be treated as tuning parameters which are chosen by the designer to achieve the best compromise between the conflicting objectives. The selection of the weighting matrices should be guided by the expected system inputs and the relative importance of the outputs.

Hence, control design may be regarded as a process of choosing a controller K such that certain weighted signals are made small in some sense. There are many different ways to define the smallness of a signal or transfer matrix, as we have discussed in the last chapter. Different definitions lead to different control synthesis methods, and some are much harder than others. A control engineer should make a judgment of the mathematical complexity versus engineering requirements.

Below, we introduce two classes of performance formulations: $\mathcal{H}_2$ and $\mathcal{H}_\infty$ criteria. For the simplicity of presentation, we shall assume $d_i = 0$ and $n = 0$.

$\mathcal{H}_2$ Performance

Assume, for example, that the disturbance $\tilde{d}$ can be approximately modeled as an impulse with random input direction, i.e.,

$$\tilde{d}(t) = \eta\delta(t)$$

and

$$E(\eta\eta^*) = I$$

where E denotes the expectation. We may choose to minimize the expected energy of the error e due to the disturbance $\tilde{d}$:

$$E\left\{\|e\|_2^2\right\} = E\left\{\int_0^\infty \|e\|^2\, dt\right\} = \|W_e S_o W_d\|_2^2 .$$

Alternatively, if we suppose that the disturbance $\tilde{d}$ can be approximately modeled as white noise with $S_{\tilde{d}\tilde{d}} = I$, then

$$S_{ee} = (W_e S_o W_d) S_{\tilde{d}\tilde{d}} (W_e S_o W_d)^*,$$

and we may chose to minimize the power of e:

$$\|e\|_P^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} \operatorname{Trace} S_{ee}(j\omega)\, d\omega = \|W_e S_o W_d\|_2^2 .$$

In general, a controller minimizing only the above criterion can lead to a very large control signal u that could cause saturation of the actuators as well as many other undesirable problems. Hence, for a realistic controller design, it is necessary to include the control signal u in the penalty function. Thus, our design criterion would usually be something like this

$$E\left\{\|e\|_2^2 + \rho^2\|\tilde{u}\|_2^2\right\} = \left\| \begin{bmatrix} W_e S_o W_d \\ \rho W_u K S_o W_d \end{bmatrix} \right\|_2^2$$

with some appropriate choice of weighting matrix W_u and scalar ρ. The parameter ρ clearly defines the tradeoff we discussed earlier between good disturbance rejection at the output and control effort (or disturbance and sensor noise rejection at the actuators). Note that ρ can be set to $\rho = 1$ by an appropriate choice of W_u. This problem can be viewed as minimizing the *energy* consumed by the system in order to reject the disturbance d.

This type of problem was the dominant paradigm in the 1960's and 1970's and is usually referred to as Linear Quadratic Gaussian Control or simply as LQG. (They will also be referred to as $\mathcal{H}_2$ mixed sensitivity problems for the consistency with the $\mathcal{H}_\infty$ problems discussed next.) The development of this paradigm stimulated extensive research efforts and is responsible for important technological innovation, particularly in the area of estimation. The theoretical contributions include a deeper understanding of linear systems and improved computational methods for complex systems through state-space techniques. The major limitation of this theory is the lack of formal treatment of uncertainty in the plant itself. By allowing only additive noise for uncertainty, the stochastic theory ignored this important practical issue. Plant uncertainty is particularly critical in feedback systems.

$\mathcal{H}_\infty$ Performance

Although the $\mathcal{H}_2$ norm (or $\mathcal{L}_2$ norm) may be a meaningful performance measure and although LQG theory can give efficient design compromises under certain disturbance and plant assumptions, the $\mathcal{H}_2$ norm suffers a major deficiency. This deficiency is due to the fact that the tradeoff between disturbance error reduction and sensor noise error reduction is not the only constraint on feedback design. The problem is that these performance tradeoffs are often overshadowed by a second limitation on high loop gains – namely, the requirement for tolerance to uncertainties. Though a controller may be designed using FDLTI models, the design must be implemented and operated with a real physical plant. The properties of physical systems, in particular the ways in which they deviate from finite-dimensional linear models, put strict limitations on the frequency range over which the loop gains may be large.

A solution to this problem would be to put explicit constraints on the loop gain in the penalty function. For instance, one may chose to minimize

$$\sup_{\|\tilde{d}\|_2 \le 1} \|e\|_2 = \|W_e S_o W_d\|_\infty ,$$

subject to some restrictions on the control energy or control bandwidth:

$$\sup_{\|\tilde{d}\|_2 \le 1} \|\tilde{u}\|_2 = \|W_u K S_o W_d\|_\infty .$$

Or more frequently, one may introduce a parameter ρ and a mixed criterion

$$\sup_{\|\tilde{d}\|_2 \le 1} \left\{ \|e\|_2^2 + \rho^2 \|\tilde{u}\|_2^2 \right\} = \left\| \begin{bmatrix} W_e S_o W_d \\ \rho W_u K S_o W_d \end{bmatrix} \right\|_\infty^2 .$$

This problem can also be regarded as minimizing the maximum power of the error subject to all bounded power disturbances: let

$$\hat{e} := \begin{bmatrix} e \\ \tilde{u} \end{bmatrix}$$

then

$$S_{\hat{e}\hat{e}} = \begin{bmatrix} W_e S_o W_d \\ W_u K S_o W_d \end{bmatrix} S_{\tilde{d}\tilde{d}} \begin{bmatrix} W_e S_o W_d \\ W_u K S_o W_d \end{bmatrix}^*$$

and

$$\sup_{\|\tilde{d}\|_P \leq 1} \|\hat{e}\|_P^2 = \sup_{\|\tilde{d}\|_P \leq 1} \left\{ \frac{1}{2\pi} \int_{-\infty}^{\infty} \operatorname{Trace} S_{\hat{e}\hat{e}}(j\omega)\, d\omega \right\} = \left\| \begin{bmatrix} W_e S_o W_d \\ W_u K S_o W_d \end{bmatrix} \right\|_\infty^2 .$$

Alternatively, if the system robust stability margin is the major concern, the weighted complementary sensitivity has to be limited. Thus the whole cost function may be

$$\left\| \begin{bmatrix} W_e S_o W_d \\ \rho W_1 T_o W_2 \end{bmatrix} \right\|_\infty$$

where W_1 and W_2 are the frequency dependent uncertainty scaling matrices. These design problems are usually called $\mathcal{H}_\infty$ mixed sensitivity problems. For a scalar system, an $\mathcal{H}_\infty$ norm minimization problem can also be viewed as minimizing the maximum magnitude of the system's steady-state response with respect to the worst case sinusoidal inputs.

5.8 Notes and References

The presentation of this chapter is based primarily on Doyle [1984]. The discussion of internal stability and coprime factorization can also be found in Francis [1987] and Vidyasagar [1985]. The loop shaping design is well known for SISO systems in the classical control theory. The idea was extended to MIMO systems by Doyle and Stein [1981] using LQG design technique. The limitations of the loop shaping design are discussed in detail in Stein and Doyle [1991]. Chapter 18 presents another loop shaping method using $\mathcal{H}_\infty$ control theory which has the potential to overcome the limitations of the LQG/LTR method.

6

Performance Limitations

This chapter introduces some multivariable versions of the Bode's sensitivity integral relations and Poisson integral formula. The sensitivity integral relations are used to study the design limitations imposed by bandwidth constraints and the open-loop unstable poles, while the Poisson integral formula is used to study the design constraints imposed by the non-minimum phase zeros. These results display that the design limitations in multivariable systems are dependent on the directionality properties of the sensitivity function as well as those of the poles and zeros, in addition to the dependence upon pole and zero locations which is known in single-input single-output systems. These integral relations are also used to derive lower bounds on the singular values of the sensitivity function which display the design tradeoffs.

6.1 Introduction

One important problem that arises frequently is concerned with the level of performance that can be achieved in feedback design. It has been shown in the previous chapters that the feedback design goals are inherently conflicting, and a tradeoff must be performed among different design objectives. It is also known that the fundamental requirements such as stability and robustness impose inherent limitations upon the feedback properties irrespective of design methods, and the design limitations become more severe in the presence of right-half plane zeros and poles in the open-loop transfer function.

An important tool that can be used to quantify feedback design constraints is furnished by the Bode's sensitivity integral relation and the Poisson integral formula. These integral formulae express design constraints directly in terms of the system's sensitivity and complementary sensitivity functions. A well-known theorem due to Bode states that for single-input single-output open-loop stable systems with more than one pole-zero excess, the integral of the logarithmic magnitude of the sensitivity function over all frequencies must equal to zero. This integral relation therefore suggests that in the presence of bandwidth constraints, the desirable property of sensitivity reduction in one frequency range must be traded off against the undesirable property of sensitivity increase at other frequencies. A result by Freudenberg and Looze [1985] further extends Bode's theorem to open-loop unstable systems, which shows that the presence of open-loop unstable poles makes the sensitivity tradeoff a more difficult task. In the same reference, the limitations imposed by the open loop non-minimum phase zeros upon the feedback properties were also quantified using the Poisson integral. The results presented here are some multivariable extensions of the above mentioned integral relations.

We shall now consider a linear time-invariant feedback system with an $n \times n$ loop transfer matrix L. Let $S(s)$ and $T(s)$ be the sensitivity function and the complementary sensitivity function, respectively

$$S(s) = (I + L(s))^{-1}, \quad T(s) = L(s)(I + L(s))^{-1}. \tag{6.1}$$

Before presenting the multivariable integral relations, recall that a point $z \in \mathbb{C}$ is a transmission zero of $L(s)$, which has full normal rank and a minimal state space realization (A, B, C, D), if there exist vectors ζ and η such that the relation

$$\begin{bmatrix} A - zI & B \\ C & D \end{bmatrix} \begin{bmatrix} \zeta \\ \eta \end{bmatrix} = 0$$

holds, where $\eta^*\eta = 1$, and η is called the *input zero direction* associated with z. Analogously, a transmission zero z of $L(s)$ satisfies the relation

$$\begin{bmatrix} x^* & w^* \end{bmatrix} \begin{bmatrix} A - zI & B \\ C & D \end{bmatrix} = 0,$$

where x and w are some vectors with w satisfying the condition $w^*w = 1$. The vector w is called the *output zero direction* associated with z. Note also that $p \in \mathbb{C}$ is a pole of $L(s)$ if and only if it is a zero of $L^{-1}(s)$. By a slight abuse of terminology, we shall call the input and output zero directions of $L^{-1}(s)$ the input and output pole directions of $L(s)$, respectively.

In the sequel we shall preclude the possibility that z is both a zero and pole of $L(s)$. Then, by Lemma 3.27 and 3.28, z is a zero of $L(s)$ if and only if $L(z)\eta = 0$ for some vector η, $\eta^*\eta = 1$, or $w^*L(z) = 0$ for some vector w, $w^*w = 1$. Similarly, p is a pole of $L(s)$ if and only if $L^{-1}(p)\eta = 0$ for some vector η, $\eta^*\eta = 1$, or $w^*L^{-1}(p) = 0$ for some vector w, $w^*w = 1$.

It is well-known that a non-minimum phase transfer function admits a factorization that consists of a minimum phase part and an all-pass factor. Let $z_i \in \mathbb{C}_+$, $i = 1, \cdots, k$, be the non-minimum phase zeros of $L(s)$ and let $\eta_i, \eta_i^*\eta_i = 1$, be the input directions generated from the following iterative procedure

- Let (A, B, C, D) be a minimal realization of $L(s)$ and $B^{(0)} := B$;
- Repeat for $i = 1$ to k

$$\begin{bmatrix} A - z_i I & B^{(i-1)} \\ C & D \end{bmatrix} \begin{bmatrix} \zeta_i \\ \eta_i \end{bmatrix} = 0$$

$$B^{(i)} := B^{(i-1)} - 2(\mathrm{Re} z_i)\zeta_i \eta_i^*.$$

Then, the *input factorization* of $L(s)$ is given by

$$L(s) = L_m(s) B_k(s) \cdots B_1(s)$$

where $L_m(s)$ denotes the minimum phase factor of $L(s)$, and $B_i(s)$ corresponds to the all-pass factor associated with z_i:

$$B_i(s) = I - \frac{2\mathrm{Re} z_i}{s + \bar{z}_i} \eta_i \eta_i^*. \tag{6.2}$$

In fact, $L_m(s)$ can be written as

$$L_m(s) := \left[\begin{array}{c|c} A & B^{(k)} \\ \hline C & D \end{array}\right].$$

For example, suppose $z \in \mathbb{C}$ is a zero of $L(s)$. Then it is easy to verify using state space calculation that $L(s)$ can be factorized as

$$L(s) = \left[\begin{array}{c|c} A & B - 2(\mathrm{Re} z)\zeta\eta^* \\ \hline C & D \end{array}\right] \left(I - \frac{2\mathrm{Re} z}{s + \bar{z}} \eta\eta^* \right).$$

Note that a non-minimum phase transfer function admits an *output factorization* analogous to the input factorization, and the subsequent results can be applied to both types of factorizations.

6.2 Integral Relations

In this section we provide extensions of the Bode's sensitivity integral relations and Poisson integral relations to multivariable systems. Consider a unit feedback system with a loop transfer function L. The following assumptions are needed:

Assumption 6.1 *The closed-loop system is stable.*

Assumption 6.2

$$\lim_{R \to \infty} \sup_{\substack{s \in \overline{\mathbb{C}}_+ \\ |s| \geq R}} R\, \bar{\sigma}(L(s)) = 0$$

Each of these assumptions has important implications. Assumption 6.1 implies that the sensitivity function $S(s)$ is analytic in $\overline{\mathbb{C}}_+$. Assumption 6.2 states that the open-loop transfer function has a rolloff rate of more than one pole-zero excess. Note that most of practical systems require a rolloff rate of more than one pole-zero excess in order to maintain a sufficient stability margin. One instance for this assumption to hold is that each element of L has a rolloff rate of more than one pole-zero excess.

Suppose that the open-loop transfer function $L(s)$ has poles p_i in the open right-half plane with input pole directions η_i, $i = 1, \cdots, k$, which are obtained through a similar iterative procedure as in the last section.

Lemma 6.1 *Let $L(s)$ and $S(s)$ be defined by (6.1). Then $p \in \mathbb{C}$ is a zero of $S(s)$ with zero direction η if and only if it is a pole of $L(s)$ with pole direction η.*

Proof. Let p be a pole of $L(s)$. Then there exists a vector η such that $\eta^*\eta = 1$ and $L^{-1}(p)\eta = 0$. However, $S(s) = (I + L(s))^{-1} = \left(I + L^{-1}(s)\right)^{-1} L^{-1}(s)$. Hence, $S(p)\eta = 0$. This establishes the sufficiency part. The proof for necessity follows by reversing the above procedure. □

Then the sensitivity function $S(s)$ can be factorized as

$$S(s) = S_m(s)B_1(s)B_2(s)\cdots B_k(s) \tag{6.3}$$

where $S_m(s)$ has no zeros in $\overline{\mathbb{C}}_+$, and $B_i(s)$ is given by

$$B_i(s) = I - \frac{2\mathrm{Re}p_i}{s + \overline{p}_i}\eta_i\eta_i^*.$$

Theorem 6.2 *Let $S(s)$ be factorized in (6.3) and suppose that Assumptions 6.1-6.2 hold. Then*

$$\int_0^\infty \ln \overline{\sigma}(S(j\omega))d\omega \geq \pi\lambda_{\max}\left(\sum_{j=1}^{k}(\mathrm{Re}p_j)\eta_j\eta_j^*\right). \tag{6.4}$$

It is also instructive to examine the following two extreme cases.

(i) If $\eta_i^*\eta_j = 0$ for all $i, j = 1, \cdots, k$, $i \neq j$, then

$$\int_0^\infty \ln \overline{\sigma}(S(j\omega))d\omega \geq \pi \max \mathrm{Re}p_i. \tag{6.5}$$

(ii) If $\eta_i = \eta$ for all $i = 1, \cdots, k$, then

$$\int_0^\infty \ln \overline{\sigma}(S(j\omega))d\omega \geq \pi \sum_{i=1}^{k} \mathrm{Re}p_i. \tag{6.6}$$

Note also that the following equality holds

$$\int_0^\infty \ln |S(j\omega)| d\omega = \pi \sum_{i=1}^{k} \mathrm{Re} p_i$$

if $S(s)$ is a scalar function.

Theorem 6.2 has an important implication toward the limitations imposed by the open-loop unstable poles on sensitivity properties. It shows that there will exist a frequency range over which the largest singular value of the sensitivity function exceeds one if it is to be kept below one at other frequencies. In the presence of bandwidth constraint, this imposes a sensitivity tradeoff in different frequency ranges. Furthermore, this result suggests that unlike in single-input single-output systems, the limitations imposed by open-loop unstable poles in multivariable systems are related not only to the locations, but also to the directions of poles and their relative interaction. The case (i) of this result corresponds to the situation where the pole directions are mutually orthogonal, for which the integral corresponding to each singular value is related solely to one unstable pole with a corresponding distance to the imaginary axis, as if each channel of the system is decoupled from the others in effects of sensitivity properties. The case (ii) corresponds to the situation where all the pole directions are parallel, for which the unstable poles affect only the integral corresponding to the largest singular value, as if the channel corresponding to the largest singular value contains all unstable poles. Clearly, these phenomena are unique to multivariable systems. The following result further strengthens these observations and shows that the interaction between open-loop unstable poles plays an important role toward sensitivity properties.

Corollary 6.3 *Let $k = 2$ and let the Assumptions 6.1-6.2 hold. Then*

$$\int_0^\infty \ln \overline{\sigma}(S(j\omega)) d\omega \geq \frac{\pi}{2}\left(\mathrm{Re}(p_1+p_2) + \sqrt{(\mathrm{Re}(p_1-p_2))^2 + 4\mathrm{Re}p_1 \mathrm{Re}p_2 \cos^2 \angle(\eta_1, \eta_2)}\right). \tag{6.7}$$

Proof. Note that

$$\begin{aligned}
\lambda_{\max}\left((\mathrm{Re}p_1)\eta_1\eta_1^* + (\mathrm{Re}p_2)\eta_2\eta_2^*\right) &= \lambda_{\max}\left(\begin{bmatrix} \eta_1 & \eta_2 \end{bmatrix} \begin{bmatrix} \mathrm{Re}p_1 & 0 \\ 0 & \mathrm{Re}p_2 \end{bmatrix} \begin{bmatrix} \eta_1^* \\ \eta_2^* \end{bmatrix}\right) \\
&= \lambda_{\max}\left(\begin{bmatrix} \mathrm{Re}p_1 & 0 \\ 0 & \mathrm{Re}p_2 \end{bmatrix} \begin{bmatrix} \eta_1^* \\ \eta_2^* \end{bmatrix} \begin{bmatrix} \eta_1 & \eta_2 \end{bmatrix}\right) \\
&= \lambda_{\mathrm{amx}}\left(\begin{bmatrix} \mathrm{Re}p_1 & 0 \\ 0 & \mathrm{Re}p_2 \end{bmatrix} \begin{bmatrix} 1 & \eta_1^*\eta_2 \\ \eta_2^*\eta_1 & 1 \end{bmatrix}\right) \\
&= \lambda_{\max}\left(\begin{bmatrix} \mathrm{Re}p_1 & (\mathrm{Re}p_1)\eta_1^*\eta_2 \\ (\mathrm{Re}p_2)\eta_2^*\eta_1 & \mathrm{Re}p_2 \end{bmatrix}\right).
\end{aligned}$$

A straightforward calculation then gives

$$\lambda_{\max}\left((\mathrm{Re}p_1)\eta_1\eta_1^* + (\mathrm{Re}p_2)\eta_2\eta_2^*\right) = \frac{1}{2}\mathrm{Re}(p_1+p_2)$$

$$+\frac{1}{2}\sqrt{(\mathrm{Re}(p_1 - p_2))^2 + 4\mathrm{Re}p_1\mathrm{Re}p_2\cos^2\angle(\eta_1, \eta_2)}.$$

The proof is now completed. □

The utility of this corollary is clear. This result fully characterizes the limitation imposed by a pair of open-loop unstable poles on the sensitivity reduction properties. This limitation depends not only on the relative distances of the poles to the imaginary axis, but also on the principal angle between the two pole directions.

Next, we investigate the design constraints imposed by open-loop non-minimum phase zeros upon sensitivity properties. The results below may be considered to be a matrix extension of the Poisson integral relation.

Theorem 6.4 *Let $S(s) \in \mathcal{H}_\infty$ be factorized as in (6.3) and assume that*

$$\lim_{R\to\infty} \max_{\phi\in[-\pi/2,\pi/2]} \frac{\left|\ln\overline{\sigma}(S(Re^{j\phi}))\right|}{R} = 0. \tag{6.8}$$

Then, for any non-minimum phase zero $z = x_0 + jy_0 \in \mathbb{C}_+$ of $L(s)$ with output direction w, $w^*w = 1$,

$$\int_{-\infty}^{\infty} \ln\overline{\sigma}(S(j\omega))\frac{x_0}{x_0^2 + (\omega - y_0)^2}d\omega \geq \pi\ln\overline{\sigma}(S_m(z)) \geq \pi\ln\overline{\sigma}(S(z)). \tag{6.9}$$

$$\int_{-\infty}^{\infty} \ln\overline{\sigma}(S(j\omega))\frac{x_0}{x_0^2 + (\omega - y_0)^2}d\omega \geq \pi\ln\left\|w^*B_k^{-1}(z)\cdots B_1^{-1}(z)\right\|. \tag{6.10}$$

Note that the condition (6.8) is satisfied if $L(s)$ is a proper rational transfer matrix.

Furthermore, for single-input single-output systems ($n = 1$),

$$\left\|w^*\prod_{i=1}^{k}B_i^{-1}(z)\right\| = \prod_{i=1}^{k}\left|\frac{z+\overline{p}_i}{z-p_i}\right|$$

and

$$\int_{-\infty}^{\infty} \ln|S(j\omega)|\frac{x_0}{x_0^2 + (\omega - y_0)^2}d\omega = \pi\ln\prod_{i=1}^{k}\left|\frac{z+\overline{p}_i}{z-p_i}\right|.$$

The more interesting result, however, is the inequality (6.10). This result again suggests that the multivariable sensitivity properties are closely related to the pole and zero directions. This result implies that the sensitivity reduction ability of the system may be severely limited by the open-loop unstable poles and non-minimum phase zeros, especially when these poles and zeros are close to each other and the angles between their directions are small.

6.3 Design Limitations and Sensitivity Bounds

The integral relations derived in the preceding section are now used to analyze the design tradeoffs and the limitations imposed by the bandwidth constraint and right-half plane poles and zeros upon sensitivity reduction properties. Similar to their counterparts for single-input single-output systems, these integral relations show that there will necessarily exist frequencies at which the sensitivity function exceeds one if it is to be kept below one over other frequencies, hence exhibiting a tradeoff between the reduction of the sensitivity over one frequency range against its increase over another frequency range.

Suppose that the feedback system is designed such that the level of sensitivity reduction is given by

$$\overline{\sigma}(S(j\omega)) \le M_L < 1, \quad \forall \omega \in [-\omega_L,\ \omega_L], \tag{6.11}$$

where $M_L > 0$ is a given constant. Let $z = x_0 + jy_0 \in \mathbb{C}_+$ be an open right-half plane zero of $L(s)$ with output direction w. Define also

$$\theta(z) := \int_{-\omega_L}^{\omega_L} \frac{x_0}{x_0^2 + (\omega - y_0)^2} d\omega.$$

The following lower bound on the maximum sensitivity displays a limitation due to the open right-half plane zeros upon the sensitivity reduction properties.

Corollary 6.5 *Let the assumption in Theorem 6.4 holds. In addition, suppose that the condition (6.11) is satisfied. Then, for each open right-half plane zero $z \in \mathbb{C}_+$ of $L(s)$ with output direction w,*

$$\|S(s)\|_\infty \ge \left(\frac{1}{M_L}\right)^{\frac{\theta(z)}{\pi-\theta(z)}} \left\|w^* B_k^{-1}(z) \cdots B_1^{-1}(z)\right\|^{\frac{\pi}{\pi-\theta(z)}}, \tag{6.12}$$

and

$$\|S(s)\|_\infty \ge \left(\frac{1}{M_L}\right)^{\frac{\theta(z)}{\pi-\theta(z)}} \left(\overline{\sigma}(S(z))\right)^{\frac{\pi}{\pi-\theta(z)}}. \tag{6.13}$$

Proof. Note that

$$\int_{-\infty}^{\infty} \ln \overline{\sigma}(S(j\omega)) \frac{x_0}{x_0^2 + (\omega - y_0)^2} d\omega \le (\pi - \theta(z)) \ln \|S(j\omega)\|_\infty + \theta(z) \ln(M_L).$$

Then the inequality (6.12) follows by applying inequality (6.10) and inequality (6.13) follows by applying inequality (6.9) □

The interpretation of Corollary 6.5 is similar to that in single-input single-output systems. Roughly stated, this result shows that for a non-minimum phase system, its sensitivity must increase beyond one at certain frequencies if the sensitivity reduction is to be achieved at other frequencies. Of particular importance here is that the sensitivity function will in general exhibit a larger peak in multivariable systems than in single-input single-output systems, due to the fact that $\overline{\sigma}(S(z)) \geq 1$.

The design tradeoffs and limitations on the sensitivity reduction which arise from bandwidth constraints as well as open-loop unstable poles can be studied using the extended Bode integral relations. However, these integral relations by themselves do not mandate a meaningful tradeoff between the sensitivity reduction and the sensitivity increase, since the sensitivity function can be allowed to exceed one by an arbitrarily small amount over an arbitrarily large frequency range so as not to violate the Bode integral relations. However, bandwidth constraints in feedback design typically require that the open-loop transfer function be small above a specified frequency, and that it roll off at a rate of more than one pole-zero excess above that frequency. These constraints are commonly needed to ensure stability robustness despite the presence of modeling uncertainty in the plant model, particularly at high frequencies. One way of quantifying such bandwidth constraints is by requiring the open-loop transfer function to satisfy

$$\overline{\sigma}(L(j\omega)) \leq \frac{M_H}{\omega^{1+k}} \leq \epsilon < 1, \quad \forall \omega \in [\omega_H,\ \infty) \tag{6.14}$$

where $\omega_H > \omega_L$, and $M_H > 0$, $k > 0$ are some given constants. With the bandwidth constraint given as such, the following result again shows that the sensitivity reduction specified by (6.11) can be achieved only at the expense of increasing the sensitivity at certain frequencies.

Corollary 6.6 *Suppose the Assumptions 6.1-6.2 hold. In addition, suppose that the conditions (6.11) and (6.14) are satisfied for some ω_H and ω_L such that $\omega_H > \omega_L$. Then*

$$\max_{\omega \in [\omega_L, \omega_H]} \overline{\sigma}(S(j\omega)) \geq e^{\alpha} \left(\frac{1}{M_L}\right)^{\frac{\omega_L}{\omega_H - \omega_L}} (1-\epsilon)^{\frac{\omega_H}{k(\omega_H - \omega_L)}} \tag{6.15}$$

where

$$\alpha = \frac{\pi \lambda_{\max}\left(\sum_{i=1}^{k} (Re p_i)\, \eta_i \eta_i^*\right)}{\omega_H - \omega_L}.$$

Proof. Note first that for $\omega \geq \omega_H$,

$$\overline{\sigma}(S(j\omega)) = \frac{1}{\underline{\sigma}(I + L(j\omega))} \leq \frac{1}{1 - \overline{\sigma}(L(j\omega))} \leq \frac{1}{1 - \frac{M_H}{\omega^{1+k}}}$$

and

$$-\int_{\omega_H}^{\infty} \ln\left(1 - \frac{M_H}{\omega^{1+k}}\right) d\omega \quad = \quad \sum_{i=1}^{\infty} \int_{\omega_H}^{\infty} \frac{1}{i} \left(\frac{M_H}{\omega^{1+k}}\right)^{i} d\omega$$

$$
\begin{aligned}
&= \sum_{i=1}^{\infty} \frac{1}{i} \frac{\omega_H}{i(1+k)-1} \left(\frac{M_H}{\omega_H^{1+k}} \right)^i \\
&\leq \frac{\omega_H}{k} \sum_{i=1}^{\infty} \frac{1}{i} \left(\frac{M_H}{\omega_H^{1+k}} \right)^i = -\frac{\omega_H}{k} \ln \left(1 - \frac{M_H}{\omega_H^{1+k}} \right) \\
&\leq -\frac{\omega_H}{k} \ln(1-\epsilon).
\end{aligned}
$$

Then

$$
\int_0^\infty \ln \overline{\sigma}(S(j\omega))d\omega = \int_0^{\omega_L} \ln \overline{\sigma}(S(j\omega))d\omega + \int_{\omega_L}^{\omega_H} \ln \overline{\sigma}(S(j\omega))d\omega + \int_{\omega_H}^{\infty} \ln \overline{\sigma}(S(j\omega))d\omega
$$

$$
\begin{aligned}
&\leq \omega_L \ln M_L + (\omega_H - \omega_L) \max_{\omega \in [\omega_L, \omega_H]} \ln \overline{\sigma}(S(j\omega)) - \int_{\omega_H}^{\infty} \ln \left(1 - \frac{M_H}{\omega^{1+k}} \right) d\omega \\
&\leq \omega_L \ln M_L + (\omega_H - \omega_L) \max_{\omega \in [\omega_L, \omega_H]} \ln \overline{\sigma}(S(j\omega)) - \frac{\omega_H}{k} \ln(1-\epsilon)
\end{aligned}
$$

and the result follows from applying (6.4). □

The above lower bound shows that the sensitivity can be very significant in the transition band.

6.4 Bode's Gain and Phase Relation

In the classical feedback theory, the Bode's gain-phase integral relation (see Bode [1945]) has been used as an important tool to express design constraints in scalar systems. The following is an extended version of the Bode's gain and phase relationship for an open-loop stable scalar system with possible right-half plane zeros, see Freudenberg and Looze [1988] and Doyle, Francis, and Tannenbaum [1992]:

$$
\angle L(j\omega_0) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{d \ln |L|}{d\nu} \ln \coth \frac{|\nu|}{2} d\nu + \angle \prod_{i=1}^{k} \frac{j\omega_0 + \overline{z}_i}{j\omega_0 - z_i}
$$

where z_i's are assumed to be the right-half plane zeros of $L(s)$ and $\nu := \ln(\omega/\omega_0)$. Note that $\dfrac{d \ln |L(j\omega)|}{d\nu}$ is the slope of the Bode plot which is almost always negative. It follows that $\angle L(j\omega_0)$ will be large if the gain L attenuates slowly and small if it attenuates rapidly. The behavior of $\angle L(j\omega)$ is particularly important near the crossover frequency ω_c where $|L(j\omega_c)| = 1$ since $\pi + \angle L(j\omega_c)$ is the phase margin of the feedback system, and further the return difference is given by

$$
|1 + L(j\omega_c)| = |1 + L^{-1}(j\omega_c)| = 2 \left| \sin \frac{\pi + \angle L(j\omega_c)}{2} \right|
$$

which must not be too small for good stability robustness. If $\pi + \angle L(j\omega_c)$ is forced to be very small by rapid gain attenuation, the feedback system will amplify disturbances and exhibit little uncertainty tolerance at and near ω_c. Since $\angle \dfrac{j\omega_0 + \bar{z}_i}{j\omega_0 - z_i} \leq 0$ for each i, a non-minimum phase zero contributes an additional phase lag and imposes limitations upon the rolloff rate of the open-loop gain. The conflict between attenuation rate and loop quality near crossover is thus clearly evident. A thorough discussion of the limitations these relations impose upon feedback control design is given by Bode [1945], Horowitz [1963], and Freudenberg and Looze [1988]. See also Freudenberg and Looze [1988] for some multivariable generalizations.

In the classical feedback theory, it has been common to express design goals in terms of the "shape" of the open-loop transfer function. A typical design requires that the open-loop transfer function have a high gain at low frequencies and a low gain at high frequencies while the transition should be well-behaviored. The same conclusion applies to multivariable system where the singular value plots should be well-behaviored between the transition band.

6.5 Notes and References

The results presented in this chapter are based on Chen [1992a, 1992b, 1995]. Some related results can be found in Boyd and Desoer [1985] and Freudenberg and Looze [1988]. The related results for scalar systems can be found in Bode [1945], Horowitz [1963], Doyle, Francis, and Tannenbaum [1992], and Freudenberg and Looze [1988].

The study of analytic functions, harmonic functions[1], and various integral relations in the scalar case can be found in Garnett [1981] and Hoffman [1962].

[1] A function $f: \mathbb{C} \longmapsto \mathbb{R}$ is said to be a harmonic function (subharmonic function) if $\bigtriangledown^2 f(s) = 0$ ($\bigtriangledown^2 f(s) \geq 0$) where the symbol $\bigtriangledown^2 f(s)$ with $s = x + jy$ denotes the Laplacian of $f(s)$ and is defined by

$$\bigtriangledown^2 f(s) := \frac{\partial^2 f(s)}{\partial x^2} + \frac{\partial^2 f(s)}{\partial y^2}.$$

7

Model Reduction by Balanced Truncation

In this chapter we consider the problem of reducing the order of a linear multivariable dynamical system. There are many ways to reduce the order of a dynamical system. However, we shall study only two of them: balanced truncation method and Hankel norm approximation method. This chapter focuses on the balanced truncation method while the next chapter studies the Hankel norm approximation method.

A model order reduction problem can in general be stated as follows: Given a full order model $G(s)$, find a lower order model, say, an r-th order model G_r, such that G and G_r are close in some sense. Of course, there are many ways to define the closeness of an approximation. For example, one may desire that the reduced model be such that

$$G = G_r + \Delta_a$$

and Δ_a is small in some norm. This model reduction is usually called an *additive* model reduction problem. On the other hand, one may also desire that the approximation be in relative form

$$G_r = G(I + \Delta_{rel})$$

so that Δ_{rel} is small in some norm. This is called a *relative* model reduction problem. We shall be only interested in $\mathcal{L}_\infty$ norm approximation in this book. Once the norm is chosen, the additive model reduction problem can be formulated as

$$\inf_{\deg(G_r) \leq r} \|G - G_r\|_\infty$$

and the relative model reduction problem can be formulated as

$$\inf_{\deg(G_r)\leq r} \left\| G^{-1}(G - G_r) \right\|_\infty$$

if G is invertible. In general, a practical model reduction problem is inherently frequency weighted, i.e., the requirement on the approximation accuracy at one frequency range can be drastically different from the requirement at another frequency range. These problems can in general be formulated as frequency weighted model reduction problems

$$\inf_{\deg(G_r)\leq r} \| W_o(G - G_r)W_i \|_\infty$$

with appropriate choice of W_i and W_o. We shall see in this chapter how the balanced realization can give an effective approach to the above model reduction problems.

7.1 Model Reduction by Balanced Truncation

Consider a stable system $G \in \mathcal{RH}_\infty$ and suppose $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ is a balanced realization, i.e., its controllability and observability Gramians are equal and diagonal. Denote the balanced Gramians by Σ, then

$$A\Sigma + \Sigma A^* + BB^* = 0 \tag{7.1}$$

$$A^*\Sigma + \Sigma A + C^*C = 0. \tag{7.2}$$

Now partition the balanced Gramian as $\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$ and partition the system accordingly as

$$G = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right].$$

Then (7.1) and (7.2) can be written in terms of their partitioned matrices as

$$\begin{aligned}
A_{11}\Sigma_1 + \Sigma_1 A_{11}^* + B_1B_1^* &= 0 && (7.3)\\
\Sigma_1 A_{11} + A_{11}^*\Sigma_1 + C_1^*C_1 &= 0 && (7.4)\\
A_{21}\Sigma_1 + \Sigma_2 A_{12}^* + B_2B_1^* &= 0 && (7.5)\\
\Sigma_2 A_{21} + A_{12}^*\Sigma_1 + C_2^*C_1 &= 0 && (7.6)\\
A_{22}\Sigma_2 + \Sigma_2 A_{22}^* + B_2B_2^* &= 0 && (7.7)\\
\Sigma_2 A_{22} + A_{22}^*\Sigma_2 + C_2^*C_2 &= 0. && (7.8)
\end{aligned}$$

The following theorem characterizes the properties of these subsystems.

Theorem 7.1 *Assume that Σ_1 and Σ_2 have no diagonal entries in common. Then both subsystems (A_{ii}, B_i, C_i), $i = 1, 2$ are asymptotically stable.*

Proof. It is clearly sufficient to show that A_{11} is asymptotically stable. The proof for the stability of A_{22} is similar.

By Lemma 3.20 or Lemma 3.21, Σ_1 can be assumed to be positive definite without loss of generality. Then it is obvious that $\lambda_i(A_{11}) \leq 0$ by Lemma 3.19. Assume that A_{11} is not asymptotically, then there exists an eigenvalue at $j\omega$ for some ω. Let V be a basis matrix for $\mathrm{Ker}(A_{11} - j\omega I)$. Then we have

$$(A_{11} - j\omega I)V = 0 \tag{7.9}$$

which gives

$$V^*(A_{11}^* + j\omega I) = 0.$$

Equations (7.3) and (7.4) can be rewritten as

$$(A_{11} - j\omega I)\Sigma_1 + \Sigma_1(A_{11}^* + j\omega I) + B_1 B_1^* = 0 \tag{7.10}$$

$$\Sigma_1(A_{11} - j\omega I) + (A_{11}^* + j\omega I)\Sigma_1 + C_1^* C_1 = 0. \tag{7.11}$$

Multiplication of (7.11) from the right by V and from the left by V^* gives $V^* C_1^* C_1 V = 0$, which is equivalent to

$$C_1 V = 0.$$

Multiplication of (7.11) from the right by V now gives

$$(A_{11}^* + j\omega I)\Sigma_1 V = 0.$$

Analogously, first multiply (7.10) from the right by $\Sigma_1 V$ and from the left by $V^*\Sigma_1$ to obtain

$$B_1^* \Sigma_1 V = 0.$$

Then multiply (7.10) from the right by $\Sigma_1 V$ to get

$$(A_{11} - j\omega I)\Sigma_1^2 V = 0.$$

It follows that the columns of $\Sigma_1^2 V$ are in $\mathrm{Ker}(A_{11} - j\omega I)$. Therefore, there exists a matrix $\bar{\Sigma}_1$ such that

$$\Sigma_1^2 V = V\bar{\Sigma}_1^2.$$

Since $\bar{\Sigma}_1^2$ is the restriction of Σ_1^2 to the space spanned by V, it follows that it is possible to choose V such that $\bar{\Sigma}_1^2$ is diagonal. It is then also possible to choose $\bar{\Sigma}_1$ diagonal and such that the diagonal entries of $\bar{\Sigma}_1$ are a subset of the diagonal entries of Σ_1.

Multiply (7.5) from the right by $\Sigma_1 V$ and (7.6) by V to get

$$\begin{aligned} A_{21}\Sigma_1^2 V + \Sigma_2 A_{12}^* \Sigma_1 V &= 0 \\ \Sigma_2 A_{21} V + A_{12}^* \Sigma_1 V &= 0 \end{aligned}$$

which gives

$$(A_{21}V)\bar{\Sigma}_1^2 = \Sigma_2^2(A_{21}V).$$

This is a Sylvester equation in $(A_{21}V)$. Because $\bar{\Sigma}_1^2$ and Σ_2^2 have no diagonal entries in common it follows from Lemma 2.7 that

$$A_{21}V = 0 \tag{7.12}$$

is the unique solution. Now (7.12) and (7.9) implies that

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} V \\ 0 \end{bmatrix} = j\omega \begin{bmatrix} V \\ 0 \end{bmatrix}$$

which means that the A-matrix of the original system has an eigenvalue at $j\omega$. This contradicts the fact that the original system is asymptotically stable. Therefore A_{11} must be asymptotically stable. □

Corollary 7.2 *If Σ has distinct singular values, then every subsystem is asymptotically stable.*

The stability condition in Theorem 7.1 is only sufficient. For example,

$$\frac{(s-1)(s-2)}{(s+1)(s+2)} = \left[\begin{array}{cc|c} -2 & -2.8284 & -2 \\ 0 & -1 & -1.4142 \\ \hline 2 & 1.4142 & 1 \end{array}\right]$$

is a balanced realization with $\Sigma = I$ and every subsystem of the realization is stable. On the other hand,

$$\frac{s^2 - s + 2}{s^2 + s + 2} = \left[\begin{array}{cc|c} -1 & 1.4142 & 1.4142 \\ -1.4142 & 0 & 0 \\ \hline -1.4142 & 0 & 1 \end{array}\right]$$

is also a balanced realization with $\Sigma = I$ but one of the subsystems is not stable.

Theorem 7.3 *Suppose $G(s) \in \mathcal{RH}_\infty$ and*

$$G(s) = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right]$$

is a balanced realization with Gramian $\Sigma = \mathrm{diag}(\Sigma_1, \Sigma_2)$

$$\begin{aligned} \Sigma_1 &= diag(\sigma_1 I_{s_1}, \sigma_2 I_{s_2}, \ldots, \sigma_r I_{s_r}) \\ \Sigma_2 &= diag(\sigma_{r+1} I_{s_{r+1}}, \sigma_{r+2} I_{s_{r+2}}, \ldots, \sigma_N I_{s_N}) \end{aligned}$$

and

$$\sigma_1 > \sigma_2 > \cdots > \sigma_r > \sigma_{r+1} > \sigma_{r+2} > \cdots > \sigma_N$$

where σ_i has multiplicity s_i, $i = 1, 2, \ldots, N$ and $s_1 + s_2 + \cdots + s_N = n$. Then the truncated system

$$G_r(s) = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$$

is balanced and asymptotically stable. Furthermore

$$\|G(s) - G_r(s)\|_\infty \leq 2(\sigma_{r+1} + \sigma_{r+2} + \cdots + \sigma_N)$$

and the bound is achieved if $r = N - 1$, i.e.,

$$\|G(s) - G_{N-1}(s)\|_\infty = 2\sigma_N.$$

Proof. The stability of G_r follows from Theorem 7.1. We shall now prove the error bound for the case $s_i = 1$ for all i. The case where the multiplicity of σ_i is not equal to one is more complicated and an alternative proof is given in the next chapter. Hence, we assume $s_i = 1$ and $N = n$.

Let

$$\begin{array}{rcl} \phi(s) & := & (sI - A_{11})^{-1} \\ \psi(s) & := & sI - A_{22} - A_{21}\phi(s)A_{12} \\ \tilde{B}(s) & := & A_{21}\phi(s)B_1 + B_2 \\ \tilde{C}(s) & := & C_1\phi(s)A_{12} + C_2 \end{array}$$

then using the partitioned matrix results of section 2.3,

$$\begin{array}{rcl} G(s) - G_r(s) & = & C(sI - A)^{-1}B - C_1\phi(s)B_1 \\ & = & \left[\begin{array}{cc} C_1 & C_2 \end{array}\right] \left[\begin{array}{cc} sI - A_{11} & -A_{12} \\ -A_{21} & sI - A_{22} \end{array}\right]^{-1} \left[\begin{array}{c} B_1 \\ B_2 \end{array}\right] - C_1\phi(s)B_1 \\ & = & \tilde{C}(s)\psi^{-1}(s)\tilde{B}(s) \end{array}$$

computing this quantity on the imaginary axis to get

$$\overline{\sigma}\left[G(j\omega) - G_r(j\omega)\right] = \lambda_{\max}^{1/2}\left[\psi^{-1}(j\omega)\tilde{B}(j\omega)\tilde{B}^*(j\omega)\psi^{-*}(j\omega)\tilde{C}^*(j\omega)\tilde{C}(j\omega)\right]. \quad (7.13)$$

Expressions for $\tilde{B}(j\omega)\tilde{B}^*(j\omega)$ and $\tilde{C}^*(j\omega)\tilde{C}(j\omega)$ are obtained by using the partitioned form of the internally balanced Gramian equations (7.3)–(7.8).

An expression for $\tilde{B}(j\omega)\tilde{B}^*(j\omega)$ is obtained by using the definition of $B(s)$, substituting for $B_1B_1^*$, $B_1B_2^*$ and $B_2B_2^*$ from the partitioned form of the Gramian equations (7.3)–(7.5), we get

$$\tilde{B}(j\omega)\tilde{B}^*(j\omega) = \psi(j\omega)\Sigma_2 + \Sigma_2\psi^*(j\omega).$$

The expression for $\tilde{C}^*(j\omega)\tilde{C}(j\omega)$ is obtained analogously and is given by

$$\tilde{C}^*(j\omega)\tilde{C}(j\omega) = \Sigma_2\psi(j\omega) + \psi^*(j\omega)\Sigma_2.$$

These expressions for $\tilde{B}(j\omega)\tilde{B}^*(j\omega)$ and $\tilde{C}^*(j\omega)\tilde{C}(j\omega)$ are then substituted into (7.13) to obtain

$$\overline{\sigma}\left[G(j\omega) - G_r(j\omega)\right] = \lambda_{\max}^{1/2}\left\{\left[\Sigma_2 + \psi^{-1}(j\omega)\Sigma_2\psi^*(j\omega)\right]\left[\Sigma_2 + \psi^{-*}(j\omega)\Sigma_2\psi(j\omega)\right]\right\}.$$

Now consider one-step order reduction, i.e., $r = n-1$, then $\Sigma_2 = \sigma_n$ and

$$\overline{\sigma}\left[G(j\omega) - G_r(j\omega)\right] = \sigma_n\lambda_{\max}^{1/2}\left\{\left[1 + \Theta^{-1}(j\omega)\right]\left[1 + \Theta(j\omega)\right]\right\} \tag{7.14}$$

where $\Theta := \psi^{-*}(j\omega)\psi(j\omega) = \Theta^{-*}$ is an "all pass" scalar function. (This is the only place we need the assumption of $s_i = 1$) Hence $|\Theta(j\omega)| = 1$.

Using triangle inequality we get

$$\overline{\sigma}\left[G(j\omega) - G_r(j\omega)\right] \leq \sigma_n\left[1 + |\Theta(j\omega)|\right] = 2\sigma_n. \tag{7.15}$$

This completes the bound for $r = n-1$.

The remainder of the proof is achieved by using the order reduction by one step results and by noting that $G_k(s) = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$ obtained by the "k-th" order partitioning is internally balanced with balanced Gramian given by

$$\Sigma_1 = \operatorname{diag}(\sigma_1 I_{s_1}, \sigma_2 I_{s_2}, \ldots, \sigma_k I_{s_k}).$$

Let $E_k(s) = G_{k+1}(s) - G_k(s)$ for $k = 1, 2, \ldots, N-1$ and let $G_N(s) = G(s)$. Then

$$\overline{\sigma}\left[E_k(j\omega)\right] \leq 2\sigma_{k+1}$$

since $G_k(s)$ is a reduced order model obtained from the internally balanced realization of $G_{k+1}(s)$ and the bound for one-step order reduction, (7.15) holds.

Noting that

$$G(s) - G_r(s) = \sum_{k=r}^{N-1} E_k(s)$$

by the definition of $E_k(s)$, we have

$$\overline{\sigma}\left[G(j\omega) - G_r(j\omega)\right] \leq \sum_{k=r}^{N-1}\overline{\sigma}\left[E_k(j\omega)\right] \leq 2\sum_{k=r}^{N-1}\sigma_{k+1}.$$

This is the desired upper bound.

To see that the bound is actually achieved when $r = N-1$, we note that $\Theta(0) = I$. Then the right hand side of (7.14) is $2\sigma_N$ at $\omega = 0$. □

The singular values σ_i are called the Hankel singular values. A useful consequence of the above theorem is the following corollary.

Corollary 7.4 *Let $\sigma_i, i = 1, \ldots, N$ be the Hankel singular values of $G(s) \in \mathcal{RH}_\infty$. Then*

$$\|G(s) - G(\infty)\|_\infty \leq 2(\sigma_1 + \ldots + \sigma_N).$$

The above bound can be tight for some systems. For example, consider an n-th order transfer function

$$G(s) = \sum_{j=1}^{n} \frac{\alpha^{2j}}{s + \alpha^{2j}},$$

with $\alpha > 0$. Then $\|G(s)\|_\infty = G(0) = n$ and $G(s)$ has the following state space realization

$$G = \left[\begin{array}{cccc|c} -\alpha^2 & & & & \alpha \\ & -\alpha^4 & & & \alpha^2 \\ & & \ddots & & \vdots \\ & & & -\alpha^{2n} & \alpha^n \\ \hline \alpha & \alpha^2 & \cdots & \alpha^n & 0 \end{array}\right]$$

and the controllability and observability Gramians of the realization are given by

$$P = Q = \left[\frac{\alpha^{i+j}}{\alpha^{2i} + \alpha^{2j}}\right]$$

and $P = Q \to \frac{1}{2} I_n$ as $\alpha \to \infty$. So the Hankel singular values $\sigma_j \to \frac{1}{2}$ and $2(\sigma_1 + \sigma_2 + \cdots + \sigma_n) \to n = \|G(s)\|_\infty$ as $\alpha \to \infty$.

The model reduction bound can also be loose for systems with Hankel singular values close to each other. For example, consider the balanced realization of a fourth order system

$$G(s) = \left[\begin{array}{cccc|c} -19.9579 & -5.4682 & 9.6954 & 0.9160 & -6.3180 \\ 5.4682 & 0 & 0 & 0.2378 & 0.0020 \\ -9.6954 & 0 & 0 & -4.0051 & -0.0067 \\ 0.9160 & -0.2378 & 4.0051 & -0.0420 & 0.2893 \\ \hline -6.3180 & -0.0020 & 0.0067 & 0.2893 & 0 \end{array}\right]$$

with Hankel singular values given by

$$\sigma_1 = 1, \quad \sigma_2 = 0.9977, \quad \sigma_3 = 0.9957, \sigma_4 = 0.9952.$$

The approximation errors and the estimated bounds are listed in the following table. The table shows that the actual error for an r-th order approximation is almost the same as $2\sigma_{r+1}$ which would be the estimated bound if we regard $\sigma_{r+1} = \sigma_{r+2} = \cdots = \sigma_4$. In general, it is not hard to construct an n-th order system so that the r-th order balanced model reduction error is approximately $2\sigma_{r+1}$ but the error bound is arbitrarily close to $2(n-r)\sigma_{r+1}$. One method to construct such system is as follows: Let $G(s)$ be a stable all-pass function, i.e., $G(s)^\sim G(s) = I$, then there is a balanced realization for G so that the controllability and observability Gramians are $P = Q = I$. Next make a very small perturbation to the balanced realization then the perturbed system has a balanced realization with distinct singular values and $P = Q \approx I$. This perturbed system will have the desired properties and this is exactly how the above example is constructed.

r	0	1	2	3
$\|G - G_r\|_\infty$	2	1.996	1.991	1.9904
Bounds: $2\sum_{i=r+1}^{4} \sigma_i$	7.9772	5.9772	3.9818	1.9904
$2\sigma_{r+1}$	2	1.9954	1.9914	1.9904

7.2 Frequency-Weighted Balanced Model Reduction

This section considers the extension of the balanced truncation method to frequency weighted case. Given the original full order model $G \in \mathcal{RH}_\infty$, the input weighting matrix $W_i \in \mathcal{RH}_\infty$ and the output weighting matrix $W_o \in \mathcal{RH}_\infty$, our objective is to find a lower order model G_r such that

$$\|W_o(G - G_r)W_i\|_\infty$$

is made as small as possible. Assume that G, W_i, and W_o have the following state space realizations

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right], \quad W_i = \left[\begin{array}{c|c} A_i & B_i \\ \hline C_i & D_i \end{array}\right], \quad W_o = \left[\begin{array}{c|c} A_o & B_o \\ \hline C_o & D_o \end{array}\right]$$

with $A \in \mathbb{R}^{n\times n}$. Note that there is no loss of generality in assuming $D = G(\infty) = 0$ since otherwise it can be eliminated by replacing G_r with $D + G_r$.

Now the state space realization for the weighted transfer matrix is given by

$$W_o G W_i = \left[\begin{array}{ccc|c} A & 0 & BC_i & BD_i \\ B_oC & A_o & 0 & 0 \\ 0 & 0 & A_i & B_i \\ \hline D_oC & C_o & 0 & 0 \end{array}\right] =: \left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline \bar{C} & 0 \end{array}\right].$$

Let $\bar{P}$ and $\bar{Q}$ be the solutions to the following Lyapunov equations

$$\bar{A}\bar{P} + \bar{P}\bar{A}^* + \bar{B}\bar{B}^* = 0 \qquad (7.16)$$

$$\bar{Q}\bar{A} + \bar{A}^*\bar{Q} + \bar{C}^*\bar{C} = 0. \qquad (7.17)$$

Then the input weighted Gramian P and the output weighted Gramian Q are defined by

$$P := \left[\begin{array}{cc} I_n & 0 \end{array}\right] \bar{P} \left[\begin{array}{c} I_n \\ 0 \end{array}\right], \quad Q := \left[\begin{array}{cc} I_n & 0 \end{array}\right] \bar{Q} \left[\begin{array}{c} I_n \\ 0 \end{array}\right].$$

It can be shown easily that P and Q satisfy the following lower order equations

$$\begin{bmatrix} A & BC_i \\ 0 & A_i \end{bmatrix}\begin{bmatrix} P & P_{12} \\ P_{12}^* & P_{22} \end{bmatrix}+\begin{bmatrix} P & P_{12} \\ P_{12}^* & P_{22} \end{bmatrix}\begin{bmatrix} A & BC_i \\ 0 & A_i \end{bmatrix}^*+\begin{bmatrix} BD_i \\ B_i \end{bmatrix}\begin{bmatrix} BD_i \\ B_i \end{bmatrix}^*=0 \tag{7.18}$$

$$\begin{bmatrix} Q & Q_{12} \\ Q_{12}^* & Q_{22} \end{bmatrix}\begin{bmatrix} A & 0 \\ B_oC & A_o \end{bmatrix}+\begin{bmatrix} A & 0 \\ B_oC & A_o \end{bmatrix}^*\begin{bmatrix} Q & Q_{12} \\ Q_{12}^* & Q_{22} \end{bmatrix}+\begin{bmatrix} C^*D_o^* \\ C_o^* \end{bmatrix}\begin{bmatrix} C^*D_o^* \\ C_o^* \end{bmatrix}^*=0. \tag{7.19}$$

The computation can be further reduced if $W_i = I$ or $W_o = I$. In the case of $W_i = I$, P can be obtained from

$$PA^* + AP + BB^* = 0 \tag{7.20}$$

while in the case of $W_o = I$, Q can be obtained from

$$QA + A^*Q + C^*C = 0. \tag{7.21}$$

Now let T be a nonsingular matrix such that

$$TPT^* = (T^{-1})^*QT^{-1} = \begin{bmatrix} \Sigma_1 & \\ & \Sigma_2 \end{bmatrix}$$

(i.e., balanced) with $\Sigma_1 = \mathrm{diag}(\sigma_1 I_{s_1}, \ldots, \sigma_r I_{s_r})$ and $\Sigma_2 = \mathrm{diag}(\sigma_{r+1} I_{s_{r+1}}, \ldots, \sigma_n I_{s_n})$ and partition the system accordingly as

$$\left[\begin{array}{c|c} TAT^{-1} & TB \\ \hline CT^{-1} & 0 \end{array}\right] = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & 0 \end{array}\right].$$

Then a reduced order model G_r is obtained as

$$G_r = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & 0 \end{array}\right].$$

Unfortunately, there is generally no known a priori error bound for the approximation error and the reduced order model G_r is not guaranteed to be stable either.

7.3 Relative and Multiplicative Model Reductions

A very special frequency weighted model reduction problem is the relative error model reduction problem where the objective is to find a reduced order model G_r so that

$$G_r = G(I + \Delta_{rel})$$

and $\|\Delta_{rel}\|_\infty$ is made as small as possible. Δ_{rel} is usually called the *relative error*. In the case where G is square and invertible, this problem can be simply formulated as

$$\min_{\deg G_r \le r} \left\| G^{-1}(G - G_r) \right\|_\infty .$$

Of course the dual approximation problem

$$G_r = (I + \Delta_{rel})G$$

can be obtained by taking the transpose of G. We will show below that, as a bonus, the approximation G_r obtained below also serves as a multiplicative approximation:

$$G = G_r(I + \Delta_{mul})$$

where Δ_{mul} is usually called the *multiplicative error*.

Theorem 7.5 *Let $G, G^{-1} \in \mathcal{RH}_\infty$ be an n-th order square transfer matrix with a state space realization*

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

Let $W_i = I$ and $W_o = G^{-1}(s) = \left[\begin{array}{c|c} A - BD^{-1}C & -BD^{-1} \\ \hline D^{-1}C & D^{-1} \end{array}\right]$.

(a) Then the weighted Gramians P and Q for the frequency weighted balanced realization of G can be obtained as

$$PA^* + AP + BB^* = 0 \tag{7.22}$$

$$Q(A - BD^{-1}C) + (A - BD^{-1}C)^*Q + C^*(D^{-1})^*D^{-1}C = 0. \tag{7.23}$$

(b) Suppose the realization for G is weighted balanced, i.e.,

$$P = Q = \operatorname{diag}(\sigma_1 I_{s_1}, \ldots, \sigma_r I_{s_r}, \sigma_{r+1} I_{s_{r+1}}, \ldots, \sigma_N I_{s_N}) = \operatorname{diag}(\Sigma_1, \Sigma_2)$$

with $\sigma_1 > \sigma_2 > \ldots > \sigma_N \ge 0$ and let the realization of G be partitioned compatibly with Σ_1 and Σ_2 as

$$G(s) = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right].$$

Then

$$G_r(s) = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$$

is stable and minimum phase. Furthermore

$$\|\Delta_{rel}\|_\infty \le \prod_{i=r+1}^{N} \left(1 + 2\sigma_i(\sqrt{1+\sigma_i^2} + \sigma_i)\right) - 1$$

$$\|\Delta_{mul}\|_\infty \le \prod_{i=r+1}^{N} \left(1 + 2\sigma_i(\sqrt{1+\sigma_i^2} + \sigma_i)\right) - 1.$$

Proof. Since the input weighting matrix $W_i = I$, it is obvious that the input weighted Gramian is given by P. Now the output weighted transfer matrix is given by

$$G^{-1}(G-D) = \left[\begin{array}{cc|c} A & 0 & B \\ -BD^{-1}C & A-BD^{-1}C & 0 \\ \hline D^{-1}C & D^{-1}C & 0 \end{array}\right] =: \left[\begin{array}{c|c} \bar{A} & \bar{B} \\ \hline \bar{C} & 0 \end{array}\right].$$

It is easy to verify that

$$\bar{Q} := \begin{bmatrix} Q & Q \\ Q & Q \end{bmatrix}$$

satisfies the following Lyapunov equation

$$\bar{Q}\bar{A} + \bar{A}^*\bar{Q} + \bar{C}^*\bar{C} = 0.$$

Hence Q is the output weighted Gramian.

The proof for part (b) is more involved and needs much more work. We refer readers to the references at the end of the chapter for details. □

In the above theorem, we have assumed that the system is square, we shall now extend the results to include non-square case. Let $G(s)$ be a minimum phase transfer matrix with a minimal realization

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

and assume that D has full row rank. Without loss of generality, we shall also assume that D is normalized such that $DD^* = I$. Let $D_\perp$ be a matrix with full row rank such that $\begin{bmatrix} D \\ D_\perp \end{bmatrix}$ is square and unitary.

Lemma 7.6 *A complex number $z \in \mathbb{C}$ is a zero of $G(s)$ if and only if z is an uncontrollable mode of $(A - BD^*C, BD_\perp^*)$.*

Proof. Since D has full row rank and $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ is a minimal realization, z is a transmission zero of $G(s)$ if and only if

$$\begin{bmatrix} A - zI & B \\ C & D \end{bmatrix}$$

does not have full row rank. Now note that

$$\begin{bmatrix} A - zI & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ 0 & [D^*, D_\perp^*] \end{bmatrix} \begin{bmatrix} I & 0 & 0 \\ -C & I & 0 \\ 0 & 0 & I \end{bmatrix}$$

$$= \begin{bmatrix} A - BD^*C - zI & BD^* & BD_\perp^* \\ 0 & I & 0 \end{bmatrix}.$$

Then it is clear that

$$\begin{bmatrix} A - zI & B \\ C & D \end{bmatrix}$$

does not have full row rank if and only if

$$\begin{bmatrix} A - BD^*C - zI & BD_\perp^* \end{bmatrix}$$

does not have full row rank. By PBH test, this implies that z is a zero of $G(s)$ if and only if it is an uncontrollable mode of $(A - BD^*C, BD_\perp^*)$. □

Corollary 7.7 *There exists a matrix $\tilde{C}$ such that the augmented system*

$$G_a := \left[\begin{array}{c|c} A & B \\ \hline C_a & D_a \end{array}\right] = \left[\begin{array}{c|c} A & B \\ \hline C & D \\ \tilde{C} & D_\perp \end{array}\right] = \begin{bmatrix} G(s) \\ \tilde{G}(s) \end{bmatrix}$$

is minimum phase.

Proof. Note that the zeros of G_a are given by the eigenvalues of

$$A - B \begin{bmatrix} D \\ D_\perp \end{bmatrix}^{-1} \begin{bmatrix} C \\ \tilde{C} \end{bmatrix} = A - BD^*C - BD_\perp^*\tilde{C}.$$

Hence $\tilde{C}$ can always be chosen so that $A - BD^*C - BD_\perp^*\tilde{C}$ is stable. □

If the previous model reduction algorithms are applied to the augmented system G_a, the corresponding P and Q equations are given by

$$PA^* + AP + BB^* = 0$$

$$Q(A - BD_a^{-1}C_a) + (A - BD_a^{-1}C_a)^*Q + C_a^*(D_a^{-1})^*D_a^{-1}C_a = 0.$$

Moreover, we have

$$\begin{bmatrix} \hat{G}(s) \\ \hat{\tilde{G}}(s) \end{bmatrix} = \begin{bmatrix} G(s) \\ \tilde{G}(s) \end{bmatrix}(I + \Delta_{rel}), \quad \begin{bmatrix} G(s) \\ \tilde{G}(s) \end{bmatrix} = \begin{bmatrix} \hat{G}(s) \\ \hat{\tilde{G}}(s) \end{bmatrix}(I + \Delta_{mul})$$

and

$$\hat{G}(s) = G(s)(I + \Delta_{rel}), \quad G(s) = \hat{G}(s)(I + \Delta_{mul}).$$

However, there are in general infinitely many choices of $\tilde{C}$ and the model reduction results will in general depend on the specific choice. Hence an appropriate choice of $\tilde{C}$ is important. To motivate our choice, note that the equation for Q can be rewritten as

$$Q(A-BD^*C)+(A-BD^*C)^*Q-QBD_\perp^*D_\perp B^*Q+C^*C+(\tilde{C}-D_\perp B^*Q)^*(\tilde{C}-D_\perp B^*Q) = 0$$

A natural choice might be

$$\tilde{C} = D_\perp B^* Q.$$

The existence of a solution Q to the following so-called algebraic Riccati equation

$$Q(A - BD^*C) + (A - BD^*C)^*Q - QBD_\perp^* D_\perp B^* Q + C^*C = 0$$

such that

$$A - BD_a^{-1}C_a = A - BD^*C - BD_\perp^*\tilde{C} = A - BD^*C - BD_\perp^* D_\perp B^* Q$$

is stable will be shown in Chapter 13.

In the case where the model is not stable and /or is not minimum phase, the following procedure can be used: Let $G(s)$ be factorized as $G(s) = G_{ap}(s)G_{mp}(s)$ such that G_{ap} is an all-pass, i.e., $G_{ap}^\sim G_{ap} = I$, and G_{mp} is stable and minimum phase. Let $\hat{G}_{mp}$ be a relative/multiplicative reduced model of G_{mp} such that

$$\hat{G}_{mp} = G_{mp}(I + \Delta_{\text{rel}})$$

and

$$G_{mp} = \hat{G}_{mp}(I + \Delta_{\text{mul}}).$$

Then $\hat{G} := G_{ap}\hat{G}_{mp}$ has exactly the same right half plane poles and zeros as that of G and

$$\hat{G} = G(I + \Delta_{\text{rel}})$$

$$G = \hat{G}(I + \Delta_{\text{mul}}).$$

Unfortunately, this approach may be conservative if the transfer matrix has many non-minimum phase zeros or unstable poles.

An alternative relative/multiplicative model reduction approach, which does not require that the transfer matrix be minimum phase but does require solving an algebraic Riccati equation, is the so-called Balanced Stochastic Truncation (BST) method. Let $G(s) \in \mathcal{RH}_\infty$ be a square transfer matrix with a state space realization

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

and $\det(D) \neq 0$. Let $W(s) \in \mathcal{RH}_\infty$ be a minimum phase left spectral factor of $G(s)G^\sim(s)$, i.e,

$$W^\sim(s)W(s) = G(s)G^\sim(s).$$

Then $W(s)$ can be obtained as

$$W(s) = \left[\begin{array}{c|c} A & B_W \\ \hline C_W & D^* \end{array}\right]$$

with

$$\begin{aligned} B_W &= PC^* + BD^* \\ C_W &= D^{-1}(C - B_W^* X) \end{aligned}$$

where P is the controllability Gramian given by

$$AP + PA^* + BB^* = 0 \tag{7.24}$$

and X is the solution of a Riccati equation

$$XA_W + A_W^* X + XB_W(DD^*)^{-1}B_W^* X + C^*(DD^*)^{-1}C = 0 \tag{7.25}$$

with $A_W := A - B_W(DD^*)^{-1}C$ such that $A_W + B_W(DD^*)^{-1}B_W^* X$ is stable. The realization G is said to be a *balanced stochastic realization* if

$$P = X = \begin{bmatrix} \mu_1 I_{s_1} & & & \\ & \mu_2 I_{s_2} & & \\ & & \ddots & \\ & & & \mu_n I_{s_n} \end{bmatrix}$$

with $\mu_1 > \mu_2 > \ldots > \mu_n \geq 0$. μ_i is in fact the i-th Hankel singular value of the so-called "phase matrix" $(W^\sim(s))^{-1}G(s)$.

Theorem 7.8 *Let $G(s) \in \mathcal{RH}_\infty$ have the following balanced stochastic realization*

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right]$$

with $\det(D) \neq 0$ *and* $P = X = \mathrm{diag}(M_1, M_2)$ *where*

$$M_1 = \mathrm{diag}(\mu_1 I_{s_1}, \ldots, \mu_r I_{s_r}),\ M_2 = \mathrm{diag}(\mu_{r+1} I_{s_{r+1}}, \ldots, \mu_n I_{s_n}).$$

Then

$$\hat{G} = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$$

is stable and

$$\left\| G^{-1}(G - \hat{G}) \right\|_\infty \leq \prod_{i=r+1}^{n} \frac{1+\mu_i}{1-\mu_i} - 1$$

$$\left\| \hat{G}^{-1}(G - \hat{G}) \right\|_\infty \leq \prod_{i=r+1}^{n} \frac{1+\mu_i}{1-\mu_i} - 1.$$

It can be shown that the balanced stochastic realization and the self-weighted balanced realization in Theorem 7.5 are the same and $\mu_i = \sigma_i/\sqrt{1+\sigma_i^2}$ if $G(s)$ is minimum phase. To illustrate, consider a third order stable and minimum phase transfer function

$$G(s) = \frac{s^3 + 2s^2 + 3s + 4}{s^3 + 3s^2 + 4s + 4}.$$

It is easy to show that the Hankel singular values of the phase function is given by

$$\mu_1 = 0.55705372196966, \quad \mu_2 = 0.53088390857035, \quad \mu_3 = 0.03715882438832,$$

and a first order BST approximation is given by

$$\hat{G} = \frac{s + 0.00375717470515}{s + 0.00106883052786}.$$

The relative approximation error is 2.51522024045904 and the error bound is

$$\prod_{i=2}^{3} \frac{1+\mu_i}{1-\mu_i} - 1 = 2.51522024046226.$$

7.4 Notes and References

The balanced model reduction method was first introduced by Moore [1981]. The stability properties of the reduced order model were shown by Pernebo and Silverman [1982]. The error bound for the balanced model reduction was shown by Enns [1984] and Glover [1984] subsequently gave an independent proof. The frequency weighted balanced model reduction method was also introduced by Enns [1984] from a somewhat different perspective. The error bounds for the relative and multiplicative approximations using the self-weighted balanced realization were shown by Zhou [1993]. The Balanced Stochastic Truncation (BST) method was proposed by Desai and Pal [1984] and generalized by Green [1988a,1988b] and many other people. The relative error bound for the Balanced Stochastic Truncation method was obtained by Green [1988a] and the multiplicative error bound for the BST was obtained by Wang and Safonov [1992]. It was also shown in Zhou [1993] that the frequency self-weighted method and the BST method are the same if the transfer matrix is minimum phase. Improved error bounds for the BST reduction were reported in Wang and Safonov [1990,1992] where it was claimed that the following error bounds hold

$$\left\| G^{-1}(G - \hat{G}) \right\|_\infty \le 2 \sum_{i=r+1}^{n} \frac{\mu_i}{1-\mu_i}$$

$$\left\| \hat{G}^{-1}(G - \hat{G}) \right\|_\infty \le 2 \sum_{i=r+1}^{n} \frac{\mu_i}{1-\mu_i}.$$

However, the example in the last section gives

$$2\sum_{i=2}^{3}\frac{\mu_i}{1-\mu_i} = 2.34052280324021$$

which is smaller than the actual error. Other weighted model reduction methods can be found in Al-Saggaf and Franklin [1988], Glover [1986,1989], Glover, Limebeer and Hung [1992], Hung and Glover [1986] and references therein. Discrete time balance model reduction results can be found in Al-Saggaf and Franklin [1987], Hinrichsen and Pritchard [1990], and references therein.

8

Hankel Norm Approximation

This chapter is devoted to the study of optimal Hankel norm approximation and its applications in $\mathcal{L}_\infty$ norm model reduction. The Hankel operator is introduced first together with some time domain and frequency domain characterizations. The optimal Hankel norm approximation problem can be stated as follows: Given $G(s)$ of McMillan degree n, find $\hat{G}(s)$ of McMillan degree $k < n$ such that $\left\|G(s) - \hat{G}(s)\right\|_H$ is minimized. The solution to this approximation problem relies on the all-pass dilation result of a square transfer function which will be given for a general class of transfer functions. The all-pass dilation results are then specialized to obtain the optimal Hankel norm approximations, which gives

$$\inf \left\|G(s) - \hat{G}(s)\right\|_H = \sigma_{k+1}$$

where $\sigma_1 > \sigma_2 \ldots > \sigma_{k+1} \ldots > \sigma_n$ are the Hankel singular values of $G(s)$. Moreover, we show that a square stable transfer function $G(s)$ can be represented as

$$G(s) = D_0 + \sigma_1 E_1(s) + \sigma_2 E_2(s) + \ldots + \sigma_n E_n(s)$$

where $E_k(s)$ are all-pass functions and the partial sum $D_0 + \sigma_1 E_1(s) + \sigma_2 E_2(s) + \ldots + \sigma_k E_k(s)$ have McMillan degrees k. This representation is obtained by reducing the order one dimension at a time via optimal Hankel norm approximations. This representation also gives that

$$\|G\|_\infty \leq 2(\sigma_1 + \ldots + \sigma_n)$$

and further that there exists a constant D_0 such that

$$\|G(s) - D_0\|_\infty \leq (\sigma_1 + \ldots + \sigma_n).$$

The above bounds are then used to show that the k-th order optimal Hankel norm approximation, $\hat{G}(s)$, together with some constant matrix D_0 satisfies

$$\left\| G(s) - \hat{G}(s) - D_0 \right\|_\infty \le (\sigma_{k+1} + \ldots + \sigma_n).$$

We shall also provide an alternative proof for the error bounds derived in the last chapter for the truncated balanced realizations using the results obtained in this chapter.

Finally we consider the Hankel operator in discrete time and offer some alternative proof of the well-known Nehari's theorem.

8.1 Hankel Operator

Let $G(s) \in \mathcal{L}_\infty$ be a matrix function. The *Hankel operator* associated with G will be denoted by Γ_G and is defined as

$$\Gamma_G : \ \mathcal{H}_2^\perp \longmapsto \mathcal{H}_2$$

$$\Gamma_G f := (P_+ M_G) f = P_+(Gf), \quad \text{for } f \in \mathcal{H}_2^\perp$$

i.e., $\Gamma_G = P_+ M_G|_{\mathcal{H}_2^\perp}$. This is shown in the following diagram:

$$\begin{tikzcd} & \mathcal{L}_2 \arrow[d, "P_+"] \\ \mathcal{H}_2^\perp \arrow[ur, "M_G"] \arrow[r, "\Gamma_G"'] & \mathcal{H}_2 \end{tikzcd}$$

There is a corresponding Hankel operator in the time domain. Let $g(t)$ denote the inverse (bilateral) Laplace transform of $G(s)$. Then the time domain Hankel operator is

$$\Gamma_g : \ \mathcal{L}_2(-\infty, 0] \longmapsto \mathcal{L}_2[0, \infty)$$

$$\Gamma_g f := P_+(g * f), \quad \text{for } f \in \mathcal{L}_2(-\infty, 0].$$

Thus

$$(\Gamma_g f)(t) = \begin{cases} \int_{-\infty}^0 g(t-\tau) f(\tau) d\tau, & t \ge 0; \\ 0, & t < 0. \end{cases}$$

Because of the isometric isomorphism property between the $\mathcal{L}_2$ spaces in the time domain and in the frequency domain, we have

$$\|\Gamma_g\| = \|\Gamma_G\|.$$

Hence, in this book we will use the time domain and the frequency domain descriptions for Hankel operators interchangeably.

For the interest of this book, we will now further restrict G to be rational, i.e., $G(s) \in \mathcal{RL}_\infty$. Then G can be decomposed into strictly causal part and anticausal part, i.e., there are $G_s(s) \in \mathcal{RH}_2$ and $G_u(s) \in \mathcal{RH}_2^\perp$ such that

$$G(s) = G_s(s) + G(\infty) + G_u(s).$$

Now for any $f \in \mathcal{H}_2^\perp$, it is easy to see that

$$\Gamma_G f = P_+(Gf) = P_+(G_s f).$$

Hence, the Hankel operator associated with $G \in \mathcal{RL}_\infty$ depends only on the strictly causal part of G. In particular, if G is antistable, i.e., $G^\sim(s) \in \mathcal{RH}_\infty$, then $\Gamma_G = 0$. Therefore, there is no loss of generality in assuming $G \in \mathcal{RH}_\infty$ and strictly proper.

The adjoint operator of Γ_G can be computed easily from the definition as below: let $f \in \mathcal{H}_2^\perp, g \in \mathcal{H}_2$, then

$$\begin{aligned}
\langle \Gamma_G f, g\rangle &:= \langle P_+ Gf, g\rangle \\
&= \langle P_+ Gf, g\rangle + \langle P_- Gf, g\rangle \quad \text{(since } P_- Gf \text{ and } g \text{ are orthogonal)} \\
&= \langle Gf, g\rangle \\
&= \langle f, G^\sim g\rangle \\
&= \langle f, P_+ G^\sim g\rangle + \langle f, P_- G^\sim g\rangle \\
&= \langle f, P_- G^\sim g\rangle \quad \text{(since } f \text{ and } P_+ G^\sim g \text{ are orthogonal).}
\end{aligned}$$

Hence $\Gamma_G^* g = P_-(G^\sim g) : \ \mathcal{H}_2 \longmapsto \mathcal{H}_2^\perp$ or $\Gamma_G^* = P_- M_{G^\sim}|_{\mathcal{H}_2}$.

Now suppose $G \in \mathcal{RH}_\infty$ has a state space realization as given below:

$$\begin{aligned}
\dot{x} &= Ax + Bu \\
y &= Cx
\end{aligned}$$

with A stable and $x(-\infty) = 0$. Then the Hankel operator Γ_g can be written as

$$\Gamma_g u(t) = \int_{-\infty}^0 Ce^{A(t-\tau)} Bu(\tau) d\tau, \quad \text{for } t \geq 0$$

and has the interpretation of the system future output

$$y(t) = \Gamma_g u(t), \quad t \geq 0$$

based on the past input $u(t), t \leq 0$.

In the state space representation, the Hankel operator can be more specifically decomposed as the composition of maps from the past input to the initial state and then

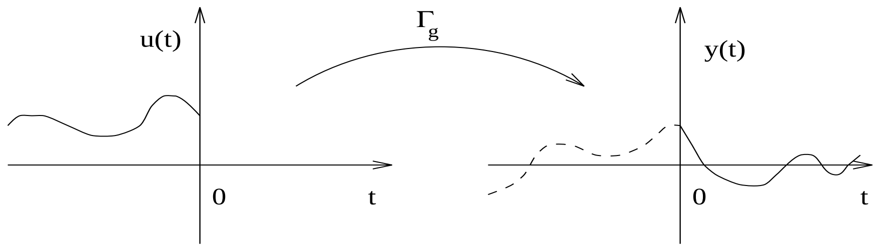


Figure 8.1: System theoretical interpretation of Hankel operators

from the initial state to the future output. These two operators will be called the *controllability operator*, Ψ_c, and the *observability operator*, Ψ_o, respectively, and are defined as

$$\Psi_c : \ \mathcal{L}_2(-\infty, 0] \longmapsto \mathbb{C}^n$$
$$\Psi_c u := \int_{-\infty}^{0} e^{-A\tau} B u(\tau) d\tau$$

and

$$\Psi_o : \ \mathbb{C}^n \longmapsto \mathcal{L}_2[0, \infty)$$
$$\Psi_o x_0 := C e^{At} x_0, \ t \geq 0.$$

(If all the data are real, then the two operators become $\Psi_c : \ \mathcal{L}_2(-\infty, 0] \longmapsto \mathbb{R}^n$ and $\Psi_o : \ \mathbb{R}^n \longmapsto \mathcal{L}_2[0, \infty)$.) Clearly, $x_0 = \Psi_c u(t)$ for $u(t) \in \mathcal{L}_2(-\infty, 0]$ is the system state at $t = 0$ due to the past input and $y(t) = \Psi_o x_0, \ t \geq 0$, is the future output due to the initial state x_0 with the input set to zero.

It is easy to verify that

$$\Gamma_g = \Psi_o \Psi_c.$$

$$\begin{tikzcd}
 & \mathbb{C}^n \arrow[dr, "\Psi_o"] & \\
\mathcal{L}_2(-\infty, 0] \arrow[ur, "\Psi_c"] \arrow[rr, "\Gamma_g"'] & & \mathcal{L}_2[0, \infty)
\end{tikzcd}$$

The adjoint operators of Ψ_c and Ψ_o can also be obtained easily from their definitions as follows: let $u(t) \in \mathcal{L}_2(-\infty, 0]$, $x_0 \in \mathbb{C}^n$, and $y(t) \in \mathcal{L}_2[0, \infty)$, then

$$\langle \Psi_c u, x_0 \rangle_{\mathbb{C}^n} = \int_{-\infty}^{0} u^*(\tau) B^* e^{-A^*\tau} x_0 d\tau = \langle u, B^* e^{-A^*\tau} x_0 \rangle_{\mathcal{L}_2(-\infty,0]} = \langle u, \Psi_c^* x_0 \rangle_{\mathcal{L}_2(-\infty,0]}$$

and

$$\langle \Psi_o x_0, y\rangle_{\mathcal{L}_2[0,\infty)} = \int_0^\infty x^*(t)e^{A^*t}C^*y(t)dt = \langle x_0, \int_0^\infty e^{A^*t}C^*y(t)dt\rangle_{\mathbb{C}^n} = \langle x_0, \Psi_o^* y\rangle_{\mathbb{C}^n}$$

where $\langle\cdot,\cdot\rangle_X$ denotes the inner product in the Hilbert space X. Therefore, we have

$$\Psi_c^*: \ \mathbb{C}^n \longmapsto \mathcal{L}_2(-\infty, 0]$$
$$\Psi_c^* x_0 = B^* e^{-A^*\tau}x_0, \ \tau \le 0$$

and

$$\Psi_o^*: \ \mathcal{L}_2[0,\infty) \longmapsto \mathbb{C}^n$$
$$\Psi_o^* y(t) = \int_0^\infty e^{A^*t}C^*y(t)dt.$$

This also gives the adjoint of Γ_g:

$$\Gamma_g^* = (\Psi_o\Psi_c)^* = \Psi_c^*\Psi_o^*: \ \mathcal{L}_2[0,\infty) \longmapsto \mathcal{L}_2(-\infty,0]$$
$$\Gamma_g^* y = \int_0^\infty B^* e^{A^*(t-\tau)}C^*y(t)dt, \ \tau \le 0.$$

Let L_c and L_o be the controllability and observability Gramians of the system, i.e.,

$$\begin{aligned} L_c &= \int_0^\infty e^{At}BB^*e^{A^*t}dt \\ L_o &= \int_0^\infty e^{A^*t}C^*Ce^{At}dt. \end{aligned}$$

Then we have

$$\begin{aligned} \Psi_c\Psi_c^* x_0 &= L_c x_0 \\ \Psi_o^*\Psi_o x_0 &= L_o x_0 \end{aligned}$$

for every $x_0 \in \mathbb{C}^n$. Thus L_c and L_o are the matrix representations of the operators $\Psi_c\Psi_c^*$ and $\Psi_o^*\Psi_o$.

Theorem 8.1 *The operator $\Gamma_g^*\Gamma_g$ (or $\Gamma_G^*\Gamma_G$) and the matrix L_cL_o have the same nonzero eigenvalues. In particular $\|\Gamma_g\| = \sqrt{\rho(L_cL_o)}$.*

Proof. Let $\sigma^2 \ne 0$ be an eigenvalue of $\Gamma_g^*\Gamma_g$, and let $0 \ne u \in \mathcal{L}_2(-\infty,0]$ be a corresponding eigenvector. Then by definition

$$\Gamma_g^*\Gamma_g u = \Psi_c^*\Psi_o^*\Psi_o\Psi_c u = \sigma^2 u. \tag{8.1}$$

Pre-multiply (8.1) by Ψ_c and define $x = \Psi_c u \in \mathbb{C}^n$ to get

$$L_cL_o x = \sigma^2 x. \tag{8.2}$$

Note that $x = \Psi_c u \neq 0$ since otherwise $\sigma^2 u = 0$ from (8.1) which is impossible. So σ^2 is an eigenvalue of $L_c L_o$.

On the other hand, suppose $\sigma^2 \neq 0$ and $x \neq 0$ are an eigenvalue and a corresponding eigenvector of $L_c L_o$. Pre-multiply (8.2) by $\Psi_c^* L_o$ and define $u = \Psi_c^* L_o x$ to get (8.1). It is easy to see that $u \neq 0$ since $\Psi_c u = \Psi_c \Psi_c^* L_o x = L_c L_o x = \sigma^2 x \neq 0$. Therefore σ^2 is an eigenvalue of $\Gamma_g^* \Gamma_g$.

Finally, since $G(s)$ is rational, $\Gamma_g^* \Gamma_g$ is compact and self-adjoint and has only discrete spectrum. Hence $\|\Gamma_g\|^2 = \|\Gamma_g^* \Gamma_g\| = \rho(L_c L_o)$. □

Remark 8.1 Let $\sigma^2 \neq 0$ be an eigenvalue of $\Gamma_g^* \Gamma_g$ and $0 \neq u \in \mathcal{L}_2(-\infty, 0]$ be a corresponding eigenvector. Define

$$v := \frac{1}{\sigma} \Gamma_g u \in \mathcal{L}_2[0, \infty).$$

Then (u, v) satisfy

$$\begin{aligned} \Gamma_g u &= \sigma v \\ \Gamma_g^* v &= \sigma u. \end{aligned}$$

This pair of vectors (u, v) are called a *Schmidt pair* of Γ_g. The proof given above suggests a way to construct this pair: find the eigenvalues and eigenvectors of $L_c L_o$, i.e., σ_i^2 and x_i such that

$$L_c L_o x_i = \sigma_i^2 x_i.$$

Then the pairs (u_i, v_i) given below are the corresponding Schmidt pairs:

$$u_i = \Psi_c^* \left(\frac{1}{\sigma_i} L_o x_i\right) \in \mathcal{L}_2(-\infty, 0], \quad v_i = \Psi_o x_i \in \mathcal{L}_2[0, \infty).$$

♡

Remark 8.2 As seen in various literature, there are some alternative ways to write a Hankel operator. For comparison, let us examine some of the alternatives below:

(i) Let $v(t) = u(-t)$ for $u(t) \in \mathcal{L}_2(-\infty, 0]$, and then $v(t) \in \mathcal{L}_2[0, \infty)$. Hence, the Hankel operator can be written as

$$\Gamma_g : \mathcal{L}_2[0, \infty) \longmapsto \mathcal{L}_2[0, \infty) \quad \text{or} \quad \Gamma_G : \mathcal{H}_2 \longmapsto \mathcal{H}_2$$

$$\begin{aligned} (\Gamma_g v)(t) &= \begin{cases} \int_0^\infty g(t+\tau) v(\tau) d\tau, & t \geq 0; \\ 0, & t < 0 \end{cases} \\ &= \int_0^\infty C e^{A(t+\tau)} B v(\tau) d\tau, \quad \text{for } t \geq 0. \end{aligned}$$

(ii) In some applications, it is more convenient to work with an anticausal operator G and view the Hankel operator associated with G as the mapping from the future input to the past output. It will be clear in later chapters that this operator is closely related to the problem of approximating an anticausal function by a causal function, which is the problem at the heart of the $\mathcal{H}_\infty$ Control Theory.

Let $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ be an antistable transfer matrix, i.e., all the eigenvalues of A have positive real parts. Then the Hankel operator associated with $G(s)$ can be written as

$$\hat{\Gamma}_g : \ \mathcal{L}_2[0,\infty) \longmapsto \mathcal{L}_2(-\infty,0]$$

$$\begin{aligned} (\hat{\Gamma}_g v)(t) &= \begin{cases} \int_0^\infty g(t-\tau)v(\tau)d\tau, & t \le 0; \\ 0, & t > 0 \end{cases} \\ &= \int_0^\infty Ce^{A(t-\tau)}Bv(\tau)d\tau, \quad \text{for } t \le 0 \end{aligned}$$

or in the frequency domain

$$\hat{\Gamma}_G = P_- M_G|_{\mathcal{H}_2} : \mathcal{H}_2 \longmapsto \mathcal{H}_2^\perp$$

$$\hat{\Gamma}_G v = P_-(Gv), \quad \text{for } v \in \mathcal{H}_2.$$

Now for any $v \in \mathcal{H}_2$ and $u \in \mathcal{H}_2^\perp$, we have

$$\langle P_-(Gv), u\rangle = \langle Gv, u\rangle = \langle v, G^\sim u\rangle = \langle v, P_+(G^\sim u)\rangle.$$

Hence, $\hat{\Gamma}_G = \Gamma^*_{G^\sim}$.

♡

8.2 All-pass Dilations

This section considers the dilation of a given transfer function to an all-pass transfer function. This transfer function dilation is the key to the optimal Hankel norm approximation in the next section. But first we need some preliminary results and some state space characterizations of all-pass functions.

Definition 8.1 *The inertia of a general complex, square matrix A denoted* $\mathrm{In}(A)$ *is the triple* $(\pi(A), \nu(A), \delta(A))$ *where*

$\pi(A) =$ *number of eigenvalues of A in the open right half-plane.*

$\nu(A) =$ *number of eigenvalues of A in the open left half-plane.*

$\delta(A) =$ *number of eigenvalues of A on the imaginary axis.*

Theorem 8.2 *Given complex $n \times n$ and $n \times m$ matrices A and B, and hermitian matrix $P = P^*$ satisfying*

$$AP + PA^* + BB^* = 0 \tag{8.3}$$

then

(1) If $\delta(P) = 0$ then $\pi(A) \le \nu(P), \nu(A) \le \pi(P)$.

(2) If $\delta(A) = 0$ then $\pi(P) \le \nu(A), \nu(P) \le \pi(A)$.

Proof. (1) If $\delta(P) = 0$ then observe that (8.3) implies

$$AP + PA^* \le 0.$$

Now suppose $\pi(A) > \nu(P)$ then there is an eigenvalue of A, $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > 0$, and a corresponding eigenvector, $x \in \mathbb{C}^n$, such that $x^*A = \lambda x^*$ and $x^*Px > 0$, which implies

$$x^*(AP + PA^*)x = (\lambda + \bar{\lambda})x^*Px \le 0$$

i.e., $\mathrm{Re}(\lambda) \le 0$, a contradiction. Hence $\pi(A) \le \nu(P)$. Similarly, we can show that $\nu(A) \le \pi(P)$.

(2) Assume $\delta(A) = 0$ and that $P = U \begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix} U^*$ with $\delta(P_1) = 0, U^*U = I$, and define

$$\tilde{A} = U^*AU = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \ \tilde{B} = U^*B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}.$$

Then $U^*(8.3)U$ gives

$$\begin{aligned}
\tilde{A} \begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} P_1 & 0 \\ 0 & 0 \end{bmatrix} \tilde{A} + \tilde{B}\tilde{B}^* &= 0 && (8.4)\\
(8.4) \Rightarrow B_2B_2^* = 0 \Rightarrow B_2 &= 0 && (8.5)\\
(8.4), (8.5) \Rightarrow A_{21}P_1 = 0 \Rightarrow A_{21} &= 0 && (8.6)\\
(8.4) \Rightarrow A_{11}P_1 + P_1A_{11}^* + B_1B_1^* &= 0 && (8.7)\\
\Rightarrow \text{(by part (1))} \ \nu(A_{11}) &\le \pi(P_1)\\
\pi(A_{11}) &\le \nu(P_1)
\end{aligned}$$

but since $\delta(A_{11}) = \delta(P_1) = 0$

$$\begin{aligned}
\pi(P_1) = \nu(A_{11}) &\le \nu(A)\\
\nu(P_1) = \pi(A_{11}) &\le \pi(A).
\end{aligned}$$

□

Theorem 8.3 *Given a realization (A, B, C) (not necessarily stable) with $A \in \mathbb{C}^{n \times n}$, $B \in \mathbb{C}^{n \times m}$, $C \in \mathbb{C}^{m \times n}$, then*

(1) If (A, B, C) is completely controllable and completely observable the following two statements are equivalent:

(a) there exists a D such that $GG^\sim = \sigma^2 I$ where $G(s) = D + C(sI - A)^{-1}B$.

(b) there exist $P, Q \in \mathbb{C}^{n\times n}$ such that

(i) $P = P^*$, $Q = Q^*$

(ii) $AP + PA^* + BB^* = 0$

(iii) $A^*Q + QA + C^*C = 0$

(iv) $PQ = \sigma^2 I$

(2) Given that part (1b) is satisfied then there exists a D satisfying

$$\begin{aligned} D^*D &= \sigma^2 I \\ D^*C + B^*Q &= 0 \\ DB^* + CP &= 0 \end{aligned}$$

and any such D will satisfy part (1a) (note, observability and controllability are not assumed).

Proof. Any systems satisfying part (1a) or (1b) can be transformed to the case $\sigma = 1$ by $\hat{B} = B/\sqrt{\sigma}$, $\hat{C} = C/\sqrt{\sigma}$, $\hat{D} = D/\sigma$, $\hat{P} = P/\sigma$, $\hat{Q} = Q/\sigma$. Hence, without loss of generality the proof will be given for the case $\sigma = 1$ only.

$(1a) \Rightarrow (1b)$ This is proved by constructing P and Q to satisfy $(1b)$ as follows. Given (1a), $G(\infty) = D \Rightarrow DD^* = I$. Also $GG^\sim = I \Rightarrow G^\sim = G^{-1}$, i.e.,

$$\begin{aligned} G^{-1}(s) &= \left[\begin{array}{c|c} A - BD^{-1}C & -BD^{-1} \\ \hline D^{-1}C & D^{-1} \end{array}\right] = \left[\begin{array}{c|c} A - BD^*C & -BD^* \\ \hline D^*C & D^* \end{array}\right] \\ &= G^\sim = \left[\begin{array}{c|c} -A^* & -C^* \\ \hline B^* & D^* \end{array}\right]. \end{aligned}$$

These two transfer functions are identical and both minimal (since (A, B, C) is assumed to be minimal), and hence there exists a similarity transformation T relating the state-space descriptions, i.e.,

$$-A^* = T(A - BD^*C)T^{-1} \tag{8.8}$$

$$C^* = TBD^* \tag{8.9}$$

$$B^* = D^*CT^{-1}. \tag{8.10}$$

Further

$$(8.9) \Rightarrow B^* = D^*C(T^*)^{-1} \tag{8.11}$$

$$(8.10) \Rightarrow C^* = T^*BD^* \tag{8.12}$$

$$\begin{aligned} (8.8) \Rightarrow -A^* &= -C^*DB^* + (T^{-1}A^*T)^* \\ &= T^*(A - (T^*)^{-1}C^*DB^*T^*)(T^*)^{-1} \end{aligned}$$

$$(8.9) \text{ and } (8.10) \Rightarrow \quad = T^*(A - BD^*C)(T^*)^{-1}. \tag{8.13}$$

Hence, T and T^* satisfy identical equations, (8.8) to (8.10) and (8.11) to (8.13), and minimality implies these have a unique solution and hence $T = T^*$.

Now setting

$$
\begin{aligned}
Q &= -T && (8.14)\\
P &= -T^{-1} && (8.15)
\end{aligned}
$$

clearly satisfies part (1b), equations (i) and (iv). Further, (8.8) and (8.9) imply

$$TA + A^*T - C^*C = 0 \tag{8.16}$$

which verifies (1b), equation (iii). Also (8.16) implies

$$AT^{-1} + T^{-1}A^* - T^{-1}C^*CT^{-1} = 0 \tag{8.17}$$

which together with (8.10) implies part (1b), equation (ii).

$(1b) \Rightarrow (1a)$ This is proved by first constructing D according to part (2) and then verifying part (1a) by calculation. Firstly note that since $Q = P^{-1}$, $Q \times ((1b), \text{equation } (ii)) \times Q$ gives

$$QA + A^*Q + QBB^*Q = 0 \tag{8.18}$$

which together with part (1b), equation (iii) implies that

$$QBB^*Q = C^*C \tag{8.19}$$

and hence by Lemma 2.14 there exists a D such that $D^*D = I$ and

$$
\begin{aligned}
DB^*Q &= -C && (8.20)\\
DB^* &= -CQ^{-1} = -CP. && (8.21)
\end{aligned}
$$

Equations (8.20) and (8.21) imply that the conditions of part (2) are satisfied. Now note that

$$
\begin{aligned}
BB^* &= (sI - A)P + P(-sI - A^*)\\
\Rightarrow C(sI - A)^{-1}BB^*(-sI - A^*)^{-1}C^* &= CP(-sI - A^*)^{-1}C^* + C(sI - A)^{-1}PC^*\\
(8.21) \Rightarrow \quad &= -DB^*(-sI - A^*)^{-1}C^* - C(sI - A)^{-1}BD^*.
\end{aligned}
$$

Hence, on expanding $G(s)G^{\sim}(s)$ we get

$$G(s)G^{\sim} = I.$$

Part (2) follows immediately from the proof of $(1b) \Rightarrow (1a)$ above. □

The following theorem dilates a given transfer function to an all-pass function.

Theorem 8.4 *Let* $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *with* $A \in \mathbb{C}^{n\times n}, B \in \mathbb{C}^{n\times m}, C \in \mathbb{C}^{m\times n}, D \in \mathbb{C}^{m\times m}$ *satisfy*

$$\begin{aligned} AP + PA^* + BB^* &= 0 && (8.22) \\ A^*Q + QA + C^*C &= 0 && (8.23) \end{aligned}$$

for

$$\begin{aligned} P = P^* &= \operatorname{diag}(\Sigma_1, \sigma I_r) && (8.24) \\ Q = Q^* &= \operatorname{diag}(\Sigma_2, \sigma I_r) && (8.25) \end{aligned}$$

with Σ_1 *and* Σ_2 *diagonal,* $\sigma \neq 0$ *and* $\delta(\Sigma_1\Sigma_2 - \sigma^2 I) = 0$.

Partition (A, B, C) *conformably with* P, *as*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix} \tag{8.26}$$

and define $W(s) := \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right]$ *with*

$$\begin{aligned} \hat{A} &= \Gamma^{-1}(\sigma^2 A_{11}^* + \Sigma_2 A_{11}\Sigma_1 - \sigma C_1^* U B_1^*) && (8.27) \\ \hat{B} &= \Gamma^{-1}(\Sigma_2 B_1 + \sigma C_1^* U) && (8.28) \\ \hat{C} &= C_1\Sigma_1 + \sigma U B_1^* && (8.29) \\ \hat{D} &= D - \sigma U && (8.30) \end{aligned}$$

where U *is a unitary matrix satisfying*

$$B_2 = -C_2^* U \tag{8.31}$$

and

$$\Gamma = \Sigma_1\Sigma_2 - \sigma^2 I. \tag{8.32}$$

Also define the error system

$$E(s) = G(s) - W(s) = \left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & D_e \end{array}\right]$$

with

$$A_e = \begin{bmatrix} A & 0 \\ 0 & \hat{A} \end{bmatrix}, \quad B_e = \begin{bmatrix} B \\ \hat{B} \end{bmatrix}, \quad C_e = \begin{bmatrix} C & -\hat{C} \end{bmatrix}, \quad D_e = D - \hat{D}. \tag{8.33}$$

Then

(1) (A_e, B_e, C_e) *satisfy*

$$\begin{aligned} A_eP_e + P_eA_e^* + B_eB_e^* &= 0 \quad (8.34)\\ A_e^*Q_e + Q_eA_e + C_e^*C_e &= 0 \quad (8.35)\end{aligned}$$

with

$$P_e = \begin{bmatrix} \Sigma_1 & 0 & I \\ 0 & \sigma I_r & 0 \\ I & 0 & \Sigma_2\Gamma^{-1} \end{bmatrix} \quad (8.36)$$

$$Q_e = \begin{bmatrix} \Sigma_2 & 0 & -\Gamma \\ 0 & \sigma I_r & 0 \\ -\Gamma & 0 & \Sigma_1\Gamma \end{bmatrix} \quad (8.37)$$

$$P_eQ_e = \sigma^2 I \quad (8.38)$$

(2) $E(s)E^{\sim}(s) = \sigma^2 I$.

(3) If $\delta(A) = 0$ *then*

(a) $\delta(\hat{A}) = 0$

(b) If $\delta(\Sigma_1\Sigma_2) = 0$ *then*

$$\mathrm{In}(\hat{A}) = \mathrm{In}(-\Sigma_1\Gamma) = \mathrm{In}(-\Sigma_2\Gamma)$$

(c) If $P > 0, Q > 0$ *then the McMillan degree of the stable part of* $(\hat{A}, \hat{B}, \hat{C})$ *equals* $\pi(\Sigma_1\Gamma) = \pi(\Sigma_2\Gamma)$.

(d) If either (i) $\Sigma_1\Gamma > 0$ *and* $\Sigma_2\Gamma > 0$ *or (ii)* $\Sigma_1\Gamma < 0$ *and* $\Sigma_2\Gamma < 0$ *then* $(\hat{A}, \hat{B}, \hat{C})$ *is a minimal realization.*

Proof. For notational convenience it will be assumed that $\sigma = 1$ and this can be done without loss of generality since B, C and Σ can be simply rescaled to give $\sigma = 1$.

It is first necessary to verify that there exists a unitary matrix U satisfying (8.31). The (2,2) blocks of (8.22) and (8.23) give

$$\begin{aligned} A_{22} + A_{22}^* + B_2B_2^* &= 0 \quad (8.39)\\ A_{22}^* + A_{22} + C_2^*C_2 &= 0 \quad (8.40)\end{aligned}$$

and hence $B_2B_2^* = C_2^*C_2$ and by Lemma 2.14 there exists a unitary U satisfying (8.31).

(1) The proof of equations (8.34) to (8.38) is by a straightforward calculation, as follows. To verify (8.34) and (8.36) we need (8.22) which is assumed, together with

$$\begin{aligned} A_{11} + \hat{A}^* + B_1\hat{B}^* &= 0 \quad (8.41)\\ A_{21} + B_2\hat{B}^* &= 0 \quad (8.42)\\ \hat{A}\Sigma_2\Gamma^{-1} + \Sigma_2\Gamma^{-1}\hat{A}^* + \hat{B}\hat{B}^* &= 0 \quad (8.43)\end{aligned}$$

which will now be verified.

$$
\begin{aligned}
B_2\hat{B}^* &= B_2(B_1^*\Sigma_2 + U^*C_1)\Gamma^{-1} && (8.44)\\
(8.31) \Rightarrow &= (B_2B_1^*\Sigma_2 - C_2^*C_1)\Gamma^{-1} && (8.45)\\
&= ((-A_{21}\Sigma_1 - A_{12}^*)\Sigma_2 + A_{12}^*\Sigma_2 + A_{21})\Gamma^{-1} && (8.46)\\
&= -A_{21} \Rightarrow (8.42)
\end{aligned}
$$

where $B_2B_1^*$ and $C_2^*C_1$ were substituted in (8.45) using the (2,1) blocks of (8.22) and (8.23), respectively. To verify (8.41)

$$
\begin{aligned}
B_1\hat{B}^* &= (B_1B_1^*\Sigma_2 + B_1U^*C_1)\Gamma^{-1} && (8.47)\\
&= (-A_{11}\Sigma_1\Sigma_2 - \Sigma_1A_{11}^*\Sigma_2 + B_1U^*C_1)\Gamma^{-1} && (8.48)\\
&= -A_{11} - \hat{A}^* \Rightarrow (8.41)
\end{aligned}
$$

where $B_1B_1^*$ was substituted using the $(1,1)$ block of (8.22) and (8.27) substituting in (8.48). Finally to verify (8.43) consider

$$
\begin{aligned}
\Gamma\hat{A}\Sigma_2 + \Sigma_2\hat{A}^*\Gamma &= (A_{11}^* + \Sigma_2A_{11}\Sigma_1 - C_1^*UB_1^*)\Sigma_2 + \Sigma_2(A_{11} + \Sigma_1A_{11}^*\Sigma_2 - B_1U^*C_1) && (8.49)\\
&= -(\Sigma_2B_1 + C_1^*U)(B_1^*\Sigma_2 + U^*C_1) + (A_{11}^*\Sigma_2 + \Sigma_2A_{11} + C_1^*C_1)\\
&\qquad + \Sigma_2(A_{11}\Sigma_1 + \Sigma_1A_{11}^* + B_1B_1^*)\Sigma_2 && (8.50)\\
&= -\Gamma\hat{B}\hat{B}^*\Gamma \Rightarrow (8.43)
\end{aligned}
$$

where (8.27)→(8.49) and (8.50) is a rearrangement of (8.49) and finally, the $(1,1)$ blocks of (8.22) and (8.23) are used. Equations (8.34) and (8.36) are hence verified.

Similarly in order to verify (8.35) and (8.37) we need (8.23) which is assumed together with

$$
\begin{aligned}
A_{11}^*(-\Gamma) + (-\Gamma)\hat{A} - C_1^*\hat{C} &= 0 && (8.51)\\
A_{12}^*(-\Gamma) - C_2^*\hat{C} &= 0 && (8.52)\\
\hat{A}^*\Sigma_1\Gamma + \Sigma_1\Gamma\hat{A} + \hat{C}^*\hat{C} &= 0. && (8.53)
\end{aligned}
$$

Equations (8.51) to (8.53) are now verified in an analogous manner to equations (8.41) to (8.43)

$$
\begin{aligned}
C_2^*\hat{C} &= C_2^*(C_1\Sigma_1 + UB_1^*)\\
&= (-A_{12}^*\Sigma_2 - A_{21})\Sigma_1 - B_2B_1^*\\
&= -A_{12}^*\Gamma \Rightarrow (8.52)\\
C_1^*\hat{C} &= C_1^*C_1\Sigma_1 + C_1^*UB_1^*\\
&= -A_{11}^*\Sigma_2\Sigma_1 - \Sigma_2A_{11}\Sigma_1 + C_1^*UB_1^*
\end{aligned}
$$

$$\begin{aligned}
&= -A_{11}^*\Gamma - \Gamma\hat{A} \Rightarrow (8.51)\\
\hat{A}^*\Sigma_1\Gamma + \Sigma_1\Gamma\hat{A} &= (A_{11} + \Sigma_1 A_{11}^*\Sigma_2 - B_1U^*C_1)\Sigma_1 + \Sigma_1(A_{11}^* + \Sigma_2 A_{11}\Sigma_1 - C_1^*UB_1)\\
&= -(\Sigma_1C_1^* + B_1U^*)(C_1\Sigma_1 + UB_1^*)\\
&\quad +\Sigma_1(A_{11}^*\Sigma_2 + \Sigma_2A_{11} + C_1^*C_1)\Sigma_1 + (A_{11}\Sigma_1 + \Sigma_1A_{11}^* + B_1B_1^*)\\
&= -\hat{C}^*\hat{C} \Rightarrow (8.53)
\end{aligned}$$

Therefore, (8.35) and (8.37) have been verified, (8.38) is immediate, and the proof of part (1) is complete.

(2) Equations (8.34), (8.35) and (8.38) ensure the conditions of Theorem 8.3, part (1b) are satisfied and Theorem 8.3, part (2) can be used to show that the D_e given in (8.33) makes $E(s)$ all-pass. (Note it is still assumed that $\sigma = 1$.) We hence need to verify that

$$\begin{aligned}
D_e^*D_e &= I && (8.54)\\
D_e^*C_e + B_e^*Q_e &= 0 && (8.55)\\
D_eB_e^* + C_eP_e &= 0. && (8.56)
\end{aligned}$$

Equation (8.54) is immediate, (8.55) follows by substituting the definitions of $\hat{B}, \hat{C}, D_e$ and Q, and (8.56) follows from $D_e \times (8.55) \times P_e$.

(3) (a) To show that $\delta(\hat{A}) = 0$ if $\delta(A) = 0$ we will assume that there exists $x \in \mathbb{C}^{n-r}$ and $\lambda \in \mathbb{C}$ such that $\hat{A}x = \lambda x$ and $\lambda + \bar{\lambda} = 0$, and show that this implies $x = 0$. From $x^*(8.53)x$,

$$\begin{aligned}
(\lambda + \bar{\lambda})x^*\Sigma_1\Gamma x + x^*\hat{C}^*\hat{C}x &= 0 && (8.57)\\
\Rightarrow \hat{C}x &= 0. && (8.58)
\end{aligned}$$

Now $(8.51)x$ gives

$$\begin{aligned}
-A_{11}^*\Gamma x - \Gamma\lambda x + C_1^*\hat{C}x &= 0\\
\Rightarrow x^*\Gamma A_{11} &= -\bar{\lambda}x^*\Gamma. && (8.59)
\end{aligned}$$

Also $(8.52)x$ and (8.58) give

$$A_{12}^*\Gamma x = 0. \quad (8.60)$$

Equations (8.59) and (8.60) imply that $(x^*\Gamma, 0)A = -\bar{\lambda}(x^*\Gamma, 0)$ but since it is assumed that $\delta(A) = 0$, $\lambda + \bar{\lambda} = 0$ and Γ^{-1} exists this implies that $x = 0$ and $\delta(\hat{A}) = 0$ is proven.

(b) Since $\delta(\hat{A}) = 0$ has been proved and $\delta(\Sigma_1\Sigma_2) = 0$ is assumed $(\Rightarrow \delta(\Sigma_2\Gamma^{-1}) = \delta(\Sigma_1\Gamma) = 0)$ Theorem 8.2 can be applied since equations (8.43) and (8.53) have been verified. Hence

$$\mathrm{In}(\hat{A}) = \mathrm{In}(-\Sigma_1\Gamma^{-1}) = \mathrm{In}(-\Sigma_1\Gamma) = \mathrm{In}(-\Sigma_2\Gamma)$$

(c) Assume that there exists $x \neq 0 \in \mathbb{C}^{n-r}$ and $\lambda \in \mathbb{C}$ such that $\hat{A}x = \lambda x$ and $\hat{C}x = 0$ (i.e., $(\hat{C}, \hat{A})$ is not completely observable). Then $(8.51)x$ and $(8.52)x$ give

$$\begin{aligned}
-A_{11}^*\Gamma x - \Gamma\lambda x &= 0\\
-A_{12}^*\Gamma x &= 0
\end{aligned}$$

hence $(-\bar{\lambda})$ is an eigenvalue of A since $\Gamma x \neq 0$. However, since $P > 0$ and $\delta(A) = 0$ are assumed then $\mathrm{In}(A) = (0, n, 0)$ and all the unobservable modes must be in the open right half plane. Similarly, if it is assumed that $(\hat{A}, \hat{B})$ is not completely controllable then (8.41) and (8.42) will give the analogous conclusion and therefore all the modes in the left half-plane are controllable and observable, and the condition in (3b) gives their number.

(d) (i) If $\Sigma_1\Gamma > 0$ or $\Sigma_2\Gamma > 0$ then by (3b) $\mathrm{In}(\hat{A}) = (0, n-r, 0)$ and by (3c) the McMillan degree of $(\hat{A}, \hat{B}, \hat{C})$ is $n - r$ and the result is proven.

(ii) Assume there exists x such that $\hat{A}x = \lambda x$ and $\hat{C}x = 0$. Then $x^*(8.53)x$ gives

$$(\lambda + \bar{\lambda})x^*\Sigma_1\Gamma x = 0$$

but $(\lambda + \bar{\lambda}) \neq 0$ by (3a) and $\Sigma_1\Gamma < 0$ is assumed so that $x = 0$. Hence $(\hat{C}, \hat{A})$ is completely observable. Similarly (8.43) gives $(\hat{A}, \hat{B})$ completely controllable. $\square$

Example 8.1 Take

$$G(s) = \frac{39s^2 + 105s + 250}{(s+2)(s+5)^2}.$$

This has a balanced realization given by

$$A = \begin{bmatrix} -2 & 4 & -4 \\ -4 & -1 & -4 \\ -4 & 4 & -9 \end{bmatrix}, \quad B = \begin{bmatrix} 2 \\ 1 \\ 6 \end{bmatrix}, \quad C = \begin{bmatrix} 2 & -1 & 6 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 2 \end{bmatrix}$$

Now using the above construction with $\Sigma_1 = \begin{bmatrix} 1 & 0 \\ 0 & \frac{1}{2} \end{bmatrix}$, $\sigma = 2$ gives

$$\hat{A} = \frac{1}{3}\begin{bmatrix} 2 & 10 \\ -8 & 5 \end{bmatrix}, \quad \hat{B} = \frac{1}{3}\begin{bmatrix} 2 \\ -2 \end{bmatrix}, \quad \hat{C} = \begin{bmatrix} -2 & -5/2 \end{bmatrix}, \quad \hat{D} = 2$$

$$W(s) = \frac{6s^2 - 13s + 90}{3s^2 - 7s + 30}, \quad G(s) - W(s) = \frac{2(-s+2)(-s+5)^2(3s^2 + 7s + 30)}{(s+2)(s+5)^2(3s^2 - 7s + 30)}$$

$W(s)$ is an optimal anticausal approximation to $G(s)$ with $\mathcal{L}_\infty$ error of 2. $\diamond$

Example 8.2 Let us also illustrate Theorem 8.4 when $(\Sigma_1^2 - \sigma^2 I)$ is indefinite. Take $G(s)$ as in the above example and permute the first and third states of the balanced realization so that $\Sigma = \mathrm{diag}(2, \frac{1}{2}, 1)$, $\Sigma_1 = \mathrm{diag}(2, \frac{1}{2})$, $\sigma = 1$. The construction of Theorem 8.4 now gives

$$\hat{A} = \begin{bmatrix} -3 & 2 \\ 8 & 3 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 2 \\ -2 \end{bmatrix}, \quad \hat{C} = \begin{bmatrix} 6 & -3/2 \end{bmatrix}, \quad \hat{D} = 1.$$

Theorem 8.4, part (2b) implies that

$$\mathrm{In}(\hat{A}) = \mathrm{In}(-\Sigma_1(\Sigma_1^2 - \sigma^2 I)) = \mathrm{In}\begin{bmatrix} -6 & 0 \\ 0 & \frac{3}{8} \end{bmatrix} = (1, 1, 0)$$

which is verified by noting that $\hat{A}$ has eigenvalues of 5 and -5.

$$W(s) = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] = \frac{s+20}{s+5}$$

and we note that the stable part of $W(s)$ has McMillan degree 1 as predicted by Theorem 8.4, part (3c). However, this example has been constructed to show that $(\hat{A}, \hat{B}, \hat{C})$ itself may not be minimal when the conditions of part (3d) are not satisfied, and in this case the unstable pole at $+5$ is both uncontrollable and unobservable. $\frac{s+20}{s+5}$ is in fact an optimal Hankel norm approximation to $G(s)$ of degree 1 and

$$E(s) = \frac{(-s+2)(-s+5)^2}{(s+2)(s+5)^2}.$$

In general the error $E(j\omega)$ will have modulus equal to σ but $E(s)$ will contain unstable poles. $\diamond$

Example 8.3 Let us finally complete the analysis of this $G(s)$ by permuting the second and third states in the balanced realization of the last example to obtain $\Sigma_1 = \mathrm{diag}(2, 1)$, $\sigma = \frac{1}{2}$. We will find

$$\hat{A} = \begin{bmatrix} -15 & -4 \\ -20 & -6 \end{bmatrix}, \quad \hat{B} = \begin{bmatrix} 4 \\ 4 \end{bmatrix}, \quad \hat{C} = \begin{bmatrix} 15 & 3 \end{bmatrix}, \quad \hat{D} = -\frac{1}{2}$$

$$W(s) = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] = \frac{1}{2}\frac{(-s^2+123s+110)}{(s^2+21s+10)}$$

$$E(s) = G(s) - W(s) = -\frac{1}{2}\frac{(-s+2)(-s+5)^2(s^2-21s+10)}{(s+2)(s+5)^2(s^2+21s+10)}.$$

Note that $-\Sigma_1\Gamma = \mathrm{diag}(-15/2, -3/4)$ so that $\hat{A}$ is stable by Theorem 8.4, part (3b). $|E(j\omega)| = \frac{1}{2}$ by Theorem 8.4, part (2), $(\hat{A}, \hat{B}, \hat{C})$ is minimal by Theorem 8.4, part (3d). $W(s)$ is in fact an optimal second-order Hankel norm approximation to $G(s)$. $\diamond$

8.3 Optimal Hankel Norm Approximation

We are now ready to give a solution to the optimal Hankel norm approximation problem based on Theorem 8.4. The following Lemma gives a lower bound on the achievable Hankel norm of the error

$$\inf_{\hat{G}} \left\| G(s) - \hat{G}(s) \right\|_H \geq \sigma_{k+1}(G)$$

and then Theorem 8.6 shows that the construction of Theorem 8.4 can be used to achieve this lower bound.

Lemma 8.5 *Given a stable, rational, $p \times m$, transfer function matrix $G(s)$ with Hankel singular values $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_k \geq \sigma_{k+1} \geq \sigma_{k+2} \ldots \geq \sigma_n > 0$, then for all $\hat{G}(s)$ stable and of McMillan degree $\leq k$*

$$\begin{aligned} \sigma_i(G(s) - \hat{G}(s)) &\geq \sigma_{i+k}(G(s)), \quad i = 1, \ldots, n-k, && (8.61) \\ \sigma_{i+k}(G(s) - \hat{G}(s)) &\leq \sigma_i(G(s)), \quad i = 1, \ldots, n. && (8.62) \end{aligned}$$

In particular,

$$\left\| G(s) - \hat{G}(s) \right\|_H \geq \sigma_{k+1}(G(s)). \tag{8.63}$$

Proof. We shall prove (8.61) only and the inequality (8.62) follows from (8.61) by setting

$$G(s) = (G(s) - \hat{G}(s)) - (-\hat{G}(s)).$$

Let $(\hat{A}, \hat{B}, \hat{C})$ be a minimal state space realization of $\hat{G}(s)$, then (A_e, B_e, C_e) given by (8.33) will be a state space realization of $G(s) - \hat{G}(s)$. Now let $P = P^*$ and $Q = Q^*$ satisfy (8.34) and (8.35) respectively (but not necessary (8.36) and (8.37) and write

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^* & P_{22} \end{bmatrix}, \quad Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{12}^* & Q_{22} \end{bmatrix}, P_{11}, Q_{11} \in \mathbb{R}^{n \times n}.$$

Since $P \geq 0$ it can be factorized as

$$P = RR^*$$

where

$$R = \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix}$$

with

$$R_{22} = P_{22}^{1/2}, \quad R_{12} = P_{12} P_{22}^{-1/2}, \quad R_{11} R_{11}^* = P_{11} - R_{12} R_{12}^*$$

($P_{22} > 0$ since $(\hat{A}, \hat{B}, \hat{C})$ is a minimal realization.)

$$\begin{aligned} \sigma_i(G(s) - \hat{G}(s)) &= \lambda_i(PQ) = \lambda_i(RR^*Q) = \lambda_i(R^*QR) \\ &\geq \lambda_i\left(\begin{bmatrix} I_n & 0 \end{bmatrix} R^*QR \begin{bmatrix} I_n \\ 0 \end{bmatrix} \right) \\ &= \lambda_i\left(\begin{bmatrix} R_{11}^* & 0 \end{bmatrix} Q \begin{bmatrix} R_{11} \\ 0 \end{bmatrix} \right) \\ &= \lambda_i(R_{11}^* Q_{11} R_{11}) = \lambda_i(Q_{11} R_{11} R_{11}^*) \\ &= \lambda_i(Q_{11}(P_{11} - R_{12} R_{12}^*)) \\ &= \lambda_i(Q_{11}^{1/2} P_{11} Q_{11}^{1/2} - XX^*) \quad \text{where} \quad X = Q_{11}^{1/2} R_{12} \\ &\geq \lambda_{i+k}(Q_{11}^{1/2} P_{11} Q_{11}^{1/2}) && (8.64) \\ &= \lambda_{i+k}(P_{11} Q_{11}) = \sigma_{i+k}^2(G) \end{aligned}$$

where (8.64) follows from the fact that X is an $n \times k$ matrix ($\Rightarrow$ rank$(XX^*) \leq k$). $\square$

We can now give a solution to the optimal Hankel norm approximation problem for square transfer functions.

Theorem 8.6 *Given a stable, rational, $m \times m$, transfer function $G(s)$ then*

(1) $\sigma_{k+1}(G(s)) = \inf_{\hat{G}\in\mathcal{H}_\infty, F\in\mathcal{H}_\infty^-} \left\| G(s) - \hat{G}(s) - F(s) \right\|_\infty$*, McMillan degree* $(\hat{G}) \leq k$.

(2) If $G(s)$ has Hankel singular values $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_k > \sigma_{k+1} = \sigma_{k+2} \ldots = \sigma_{k+r} > \sigma_{k+r+1} \geq \ldots \geq \sigma_n > 0$ then $\hat{G}(s)$ of McMillan degree k is an optimal Hankel norm approximation to $G(s)$ if and only if there exists $F(s) \in \mathcal{H}_\infty^-$ (whose McMillan degree can be chosen $\leq n + k - 1$) such that $E(s) := G(s) - \hat{G}(s) - F(s)$ satisfies

$$E(s)E^\sim(s) = \sigma_{k+1}^2 I. \tag{8.65}$$

In which case

$$\left\| G(s) - \hat{G}(s) \right\|_H = \sigma_{k+1}. \tag{8.66}$$

(3) Let $G(s)$ be as in (2) above, then an optimal Hankel norm approximation of McMillan degree k, $\hat{G}(s)$, can be constructed as follows. Let (A, B, C) be a balanced realization of $G(s)$ with corresponding

$$\Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_k, \sigma_{k+r+1}, \ldots, \sigma_n, \sigma_{k+1}, \ldots, \sigma_{k+r}),$$

and define $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ from equations (8.26) to (8.31). Then

$$\hat{G}(s) + F(s) = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] \tag{8.67}$$

where $\hat{G}(s) \in \mathcal{H}_\infty$ and $F(s) \in \mathcal{H}_\infty^-$ with the McMillan degree of $\hat{G}(s) = k$ and the McMillan degree of $F(s) = n - k - r$.

Proof. By the definition of $\mathcal{L}_\infty$ norm, for all $F(s) \in \mathcal{H}_\infty^-$, and $\hat{G}(s)$ of McMillan degree k

$$\begin{aligned}
\left\| G(s) - \hat{G}(s) - F(s) \right\|_\infty &\geq \sup_{f\in\mathcal{H}_2^\perp, \|f\|_2\leq 1} \left\| (G(s) - \hat{G}(s) - F(s))f \right\|_2 \\
&\geq \sup_{f\in\mathcal{H}_2^\perp, \|f\|_2\leq 1} \left\| P_+(G - \hat{G} - F)f \right\|_2 \\
&= \sup_{f\in\mathcal{H}_2^\perp, \|f\|_2\leq 1} \left\| P_+(G - \hat{G})f \right\|_2 \\
&= \left\| G - \hat{G} \right\|_H \\
&\geq \sigma_{k+1}(G(s))
\end{aligned} \tag{8.68}$$

where (8.68) follows from Lemma 8.5.

Now define $\hat{G}(s)$ and $F(s)$ via equation (8.26), then Theorem 8.4, part (2) implies that (8.65) holds and hence

$$\|E(s)\|_\infty = \sigma_{k+1}. \tag{8.69}$$

Also from Theorem 8.4, part (3b)

$$\mathrm{In}(\hat{A}) = \mathrm{In}(-\Sigma_1(\Sigma_1^2 - \sigma_{k+1}^2 I)) = (n-k-r, k, 0). \tag{8.70}$$

Hence, $\hat{G}$ has McMillan degree k and it in the correct class, and therefore (8.69) implies that the inequalities in (8.68) becomes equalities, and part (1) is proven, as in part (3). Clearly the sufficiency of part (2) can be similarly verified by noting that (8.65) implies that (8.68) is satisfied with equality.

To show the necessity of part (2) suppose that $\hat{G}(s)$ is an optimal Hankel norm approximation to $G(s)$ of McMillan degree k, i.e, equation (8.66) holds. Now Theorem 8.4 can be applied to $G(s) - \hat{G}(s)$ to produce an optimal anticausal approximation $F(s)$, such that $(G(s) - \hat{G}(s) - F(s))/\sigma_{k+1}(G)$ is all-pass since $\sigma_{k+1}(G) = \sigma_1(G - \hat{G})$. Further, the McMillan degree of this $F(s)$ will be, the McMillan degree of $(G(s) - \hat{G}(s))$ minus the multiplicity of $\sigma_1(G - \hat{G})$, $\leq n + k - 1$. □

The following corollary gives the solution to the well-known Nehari's problem.

Corollary 8.7 *Let $G(s)$ be a stable, rational, $m \times m$, transfer function of McMillan degree n such that $\sigma_1(G)$ has multiplicity r_1. Also let $F(s)$ be an optimal anticausal approximation of degree $n - r_1$ given by the construction of Theorem 8.4. Then*

(1) $(G(s) - F(s))/\sigma_1$ is all-pass.

(2) $\sigma_{i-r_1}(F(-s)) = \sigma_i(G(s)), i = r_1 + 1, \ldots, n$.

Proof. (1) is proved in Theorem 8.4, part (2). (2) is obtained from the forms of P_e and Q_e in Theorem 8.4, part (1). $F(-s)$ is used since it will be stable and have well-defined Hankel singular values. □

The optimal Hankel norm approximation for non-square case can be obtained by first augmenting the function to form a square function. For example, consider a stable, rational, $p \times m$ $(p < m)$, transfer function $G(s)$. Let $G_a = \begin{bmatrix} G \\ 0 \end{bmatrix}$ be an augmented square transfer function and let $\hat{G}_a = \begin{bmatrix} \hat{G} \\ \hat{G}_2 \end{bmatrix}$ be the optimal Hankel norm approximation of G_a such that

$$\left\| G_a - \hat{G}_a \right\|_H = \sigma_{k+1}(G_a).$$

Then

$$\sigma_{k+1}(G) \leq \left\| G - \hat{G} \right\|_H \leq \left\| G_a - \hat{G}_a \right\|_H = \sigma_{k+1}(G_a) = \sigma_{k+1}(G)$$

i.e., $\hat{G}$ is an optimal Hankel norm approximation of $G(s)$.

8.4 $\mathcal{L}_\infty$ Bounds for Hankel Norm Approximation

The natural question that arise now is, does the Hankel norm being small imply that any other more familiar norms are also small? We shall have a definite answer in this section.

Lemma 8.8 *Let an $m \times m$ transfer matrix $E = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ satisfy $E(s)E^{\sim}(s) = \sigma^2 I$ and all equations of Theorem 8.3 and let A have dimension $n_1 + n_2$ with n_1 eigenvalues strictly in the left half plane and $n_2 < n_1$ eigenvalues strictly in the right half plane. If $E = G_s + F$ with $G_s \in \mathcal{RH}_\infty$ and $F \in \mathcal{RH}_\infty^-$ then,*

$$\sigma_i(G_s) = \begin{cases} \sigma & i = 1, 2, \ldots, n_1 - n_2 \\ \sigma_{i-n_1+n_2}(F(-s)) & i = n_1 - n_2 + 1, \ldots, n_1 \end{cases}$$

Proof. Firstly let the realization be transformed to,

$$E = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & A_2 & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right], \quad \mathrm{Re}\lambda_i(A_1) < 0, \quad \mathrm{Re}\lambda_i(A_2) > 0,$$

in which case $G = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D \end{array}\right]$, $F = \left[\begin{array}{c|c} A_2 & B_2 \\ \hline C_2 & 0 \end{array}\right]$. The equations of Theorem 8.3 *(i)-(iv)* are then satisfied by a transformed P and Q, partitioned as,

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^* & P_{22} \end{bmatrix}, \quad Q = \begin{bmatrix} Q_{11} & Q_{21}^* \\ Q_{21} & Q_{22} \end{bmatrix}$$

$PQ = \sigma^2 I$ implies that,

$$\begin{aligned} \det(\lambda I - P_{11}Q_{11}) &= \det(\lambda I - (\sigma^2 I - P_{12}Q_{21})) \\ &= \det((\lambda - \sigma^2)I + P_{12}Q_{21}) \\ &= (\lambda - \sigma^2)^{n_1 - n_2} \det((\lambda - \sigma^2)I + Q_{21}P_{12}) \\ &= (\lambda - \sigma^2)^{n_1 - n_2} \det(\lambda I - Q_{22}P_{22}). \end{aligned}$$

The result now follows on observing that $\sigma_i(G(s)) = \lambda_i(P_{11}Q_{11})$ and $\sigma_i^2(F(-s)) = \lambda_i(Q_{22}P_{22})$. □

Corollary 8.9 *Let $E(s) = G(s) - \hat{G}(s) - F(s)$ be as defined in part (3) of Theorem 8.6 with $G(s), \hat{G}(s) \in \mathcal{RH}_\infty$ and $F(s) \in \mathcal{RH}_\infty^-$. Then for $i = 1, 2, ..., 2k + r$,*

$$\sigma_i(G - \hat{G}) = \sigma_{k+1}(G),$$

and for $i = 1, 2, ..., n - k - r$,

$$\sigma_{i+3k+r}(G) \leq \sigma_i(F(-s)) = \sigma_{i+2k+r}(G - \hat{G}) \leq \sigma_{i+k+r}(G)$$

Proof. The construction of $E(s)$ ensures that the all-pass equations are satisfied and an inertia argument easily establishes that the A-matrix has precisely $n_1 = n + k$ eigenvalues in the open left half plane and $n_2 = n - k - r$ in the open right half plane. Hence Lemma 8.8 can be applied to give the equalities. The inequalities follow from Lemma 8.5. $\square$

The following lemma gives the properties of certain optimal Hankel norm approximations when the degree is reduced by the multiplicity of σ_n. In this case some precise statements on the error and the approximating system can be made.

Lemma 8.10 *Let $G(s)$ be a stable, rational $m \times m$, transfer function of McMillan degree n and such that $\sigma_n(G)$ has multiplicity r. Also let $\hat{G}(s)$ be an optimal Hankel norm approximation of degree $n - r$ given by Theorem 8.6, part (3) (with $F(s) \equiv 0$) then*

(1) $(G(s) - \hat{G}(s))/\sigma_n(G(s))$ *is all-pass.*

(2) $\sigma_i(\hat{G}(s)) = \sigma_i(G(s)), i = 1, \ldots, n - r.$

Proof. Theorem 8.6 gives that $\hat{A} \in \mathbb{R}^{(n-r)\times(n-r)}$ is stable and hence $F(s)$ can be chosen to be zero and therefore $(G(s) - \hat{G}(s))/\sigma_n(G)$ is all-pass. The $\sigma_i(\hat{G}(s))$ are obtained from Lemma 8.8. $\square$

Applying the above reduction procedure again on $\hat{G}(s)$ and repeating until $\hat{G}(s)$ has zero McMillan degree gives the following new representation of stable systems.

Theorem 8.11 *Let $G(s)$ be a stable, rational $m \times m$, transfer function with Hankel singular values $\sigma_1 > \sigma_2 \ldots > \sigma_N$ where σ_i has multiplicity r_i and $r_1 + r_2 + \ldots + r_N = n$. Then there exists a representation of $G(s)$ as*

$$G(s) = D_0 + \sigma_1 E_1(s) + \sigma_2 E_2(s) + \ldots + \sigma_N E_N(s) \tag{8.71}$$

where

(1) $E_k(s)$ *are all-pass and stable for all* k.

(2) For $k = 1, 2, \ldots, N$

$$\hat{G}_k(s) := D_0 + \sum_{i=1}^{k} \sigma_i E_i(s)$$

has McMillan degree $r_1 + r_2 + \ldots + r_k$.

Proof. Let $\hat{G}_k(s)$ be the optimal Hankel norm approximation to $\hat{G}_{k+1}(s)$ (given by Lemma 8.10) of degree $r_1 + r_2 + \ldots + r_k$, with $\hat{G}_N(s) := G(s)$. Lemma 8.10 (2) applied at each step then gives that the Hankel singular values of $\hat{G}_k(s)$ will be $\sigma_1, \sigma_2, \ldots, \sigma_k$ with multiplicities $r_1, r_2, \ldots, r_k$, respectively. Hence Lemma 8.10 (1) gives that $\hat{G}_k(s) - \hat{G}_{k-1}(s) = \sigma_k E_k(s)$ for some stable, all-pass $E_k(s)$. Note also that Theorem 8.4, part (3d), relation (i) also ensures that each $\hat{G}_k(s)$ will have McMillan degree $r_1 + r_2 + \ldots + r_k$. Finally taking $D_0 = \hat{G}_0(s)$ which will be a constant and combining the steps gives the result. □

Note that the construction of Theorem 8.11 immediately gives an approximation algorithm that will satisfy $\left\| G(s) - \hat{G}(s) \right\|_\infty \le \sigma_{k+1} + \sigma_{k+2} + \ldots + \sigma_N$. This will not be an optimal Hankel norm approximation in general, but would involve less computation since the decomposition into $\hat{G}(s) = F(s)$ need not be done, and at each step a balanced realization of $\hat{G}_k(s)$ is given by $(\hat{A}_k, \hat{B}_k, \hat{C}_k)$ with a diagonal scaling.

An upper-bound on the $\mathcal{L}_\infty$ norm of $G(s)$ is now obtained as an immediate consequence of Theorem 8.11.

Corollary 8.12 *Let $G(s)$ be a stable, rational $p \times m$, transfer function with Hankel singular values $\sigma_1 > \sigma_2 > \ldots > \sigma_N$, where each σ_i has multiplicity r_i, and such that $G(\infty) = 0$. Then*

(1) $\|G(s)\|_\infty \le 2(\sigma_1 + \sigma_2 + \ldots + \sigma_N)$

(2) there exists a constant D_0 such that

$$\|G(s) - D_0\|_\infty \le \sigma_1 + \sigma_2 + \ldots + \sigma_N.$$

Proof. For $p = m$ consider the representation of $G(s)$ given by Theorem 8.11 then

$$\begin{aligned} \|G(s) - D_0\|_\infty &= \|\sigma_1 E_1(s) + \sigma_2 E_2(s) + \ldots + \sigma_N E_N(s)\|_\infty \\ &\le \sigma_1 + \sigma_2 + \ldots + \sigma_N \end{aligned}$$

since $E_k(s)$ are all-pass. Further setting $s = \infty$, since $G(\infty) = 0$, gives

$$\begin{aligned} \|D_0\| &\le \sigma_1 + \sigma_2 + \ldots + \sigma_N \\ \Rightarrow \|G(s)\|_\infty &\le 2(\sigma_1 + \sigma_2 + \ldots + \sigma_N). \end{aligned}$$

For the case $p \ne m$ just augment $G(s)$ by zero rows or columns to make it square, but will have the same $\mathcal{L}_\infty$ norm, then the above argument gives upper bounds on the $\mathcal{L}_\infty$ norm of this augmented system. □

Theorem 8.13 *Given a stable, rational, $m \times m$, transfer function $G(s)$ with Hankel singular values $\sigma_1 \ge \sigma_2 \ldots \ge \sigma_k > \sigma_{k+1} = \sigma_{k+2} \ldots = \sigma_{k+r} > \sigma_{k+r+1} \ge \ldots \ge \sigma_n > 0$*

and let $\hat{G}(s) \in \mathcal{RH}_\infty$ of McMillan degree k be an optimal Hankel norm approximation to $G(s)$ obtained in Theorem 8.6, then there exists a D_0 such that

$$\sigma_{k+1}(G) \le \|G - \hat{G} - D_0\|_\infty \le \sigma_{k+1}(G) + \sum_{i=1}^{n-k-r} \sigma_{i+k+r}(G).$$

Proof. The theorem follows from Corollary 8.9 and Corollary 8.12. □

It should be noted that if $\hat{G}$ is an optimal Hankel norm approximation of G then $\hat{G} + D$ for any constant matrix D is also an optimal Hankel norm approximation. Hence the constant term of $\hat{G}$ can not be determined from Hankel norm.

An appropriate constant term D_0 in Theorem 8.13 can be obtained in many ways. We shall mention three of them:

- Apply Corollary 8.12 to $G(s) - \hat{G}(s)$. This is usually complicated since $(G(s) - \hat{G}(s))$ in general has McMillan degree of $n + k$.
- An alternative is to use the unstable part of the optimal Hankel norm approximation in Theorem 8.6. Let $\hat{G} + F$ be obtained from Theorem 8.6, part (3) such that $F(s) \in \mathcal{RH}_\infty^-$ has McMillan degree $\le n - k - r$ then

  $$\|G - \hat{G} - D_0\|_\infty \le \left\|G - \hat{G} - F\right\|_\infty + \|F - D_0\|_\infty = \sigma_{k+1}(G) + \|F - D_0\|_\infty$$

  Now Corollary 8.12 can be applied to $F(-s)$ to obtain a D_0 such that

  $$\|F - D_0\|_\infty \le \sum_{i=1}^{n-k-r} \sigma_i(F(-s)) \le \sum_{i=1}^{n-k-r} \sigma_{i+k+r}(G).$$

  since by Corollary 8.9,

  $$\sigma_i(F(-s)) \le \sigma_{i+k+r}(G)$$

  for $i = 1, 2, ..., n - k - r$.
- D_0 can of course be obtained using any standard convex optimization algorithm:

  $$D_0 = \arg\min_{D_0} \left\|G - \hat{G} - D_0\right\|_\infty .$$

Note that Theorem 8.13 can also be applied to non-square systems. For example, consider a stable, rational, $p \times m$ $(p < m)$, transfer function $G(s)$. Let $G_a = \begin{bmatrix} G \\ 0 \end{bmatrix}$ be an augmented square transfer function and let $\hat{G}_a = \begin{bmatrix} \hat{G} \\ \hat{G}_2 \end{bmatrix}$ be the optimal Hankel norm approximation of G_a such that

$$\sigma_{k+1}(G_a) \le \left\|G_a - \hat{G}_a - D_a\right\|_\infty \le \sigma_{k+1}(G_a) + \sum_{i=1}^{n-k-r} \sigma_{i+k+r}(G_a)$$

with $D_a = \begin{bmatrix} D_0 \\ D_2 \end{bmatrix}$. Then

$$\sigma_{k+1}(G) \leq \left\| G - \hat{G} - D_0 \right\|_\infty \leq \sigma_{k+1}(G) + \sum_{i=1}^{n-k-r} \sigma_{i+k+r}(G)$$

since $\sigma_i(G) = \sigma_i(G_a)$ and $\left\| G - \hat{G} - D_0 \right\|_\infty \leq \left\| G_a - \hat{G}_a - D_a \right\|_\infty$.

A less tighter error bound can be obtained without computing the appropriate D_0.

Corollary 8.14 *Given a stable, rational, $m \times m$, strictly proper transfer function $G(s)$, with Hankel singular values $\sigma_1 \geq \sigma_2 \ldots \geq \sigma_k > \sigma_{k+1} = \sigma_{k+2} \ldots = \sigma_{k+r} > \sigma_{k+r+1} \geq \ldots \geq \sigma_n > 0$ and let $\hat{G}(s) \in \mathcal{RH}_\infty$ of McMillan degree k be a strictly proper optimal Hankel norm approximation to $G(s)$ obtained in Theorem 8.6, then*

$$\sigma_{k+1}(G) \leq \|G - \hat{G}\|_\infty \leq 2(\sigma_{k+1}(G) + \sum_{i=1}^{n-k-r} \sigma_{i+k+r}(G))$$

Proof. The result follows from Theorem 8.13. □

8.5 Bounds for Balanced Truncation

Very similar techniques to those of Theorem 8.11 can be used to bound the error obtained by truncating a balanced realization. We will first need a lemma that gives a perhaps surprising relationship between a truncated balanced realization of degree $(n - r_N)$ and an optimal Hankel norm approximation of the same degree.

Lemma 8.15 *Let (A, B, C) be a balanced realization of the stable, rational, $m \times m$ transfer function $G(s)$, and let*

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \sigma I \end{bmatrix}$$

Let $(\hat{A}, \hat{B}, \hat{C}, \hat{D})$ be defined by equations (8.27) to (8.32) (where $\Sigma_2 = \Sigma_1$) and define

$$\begin{aligned} G_b &:= \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & 0 \end{array}\right] \\ G_h &:= \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] \end{aligned}$$

then

(1) $(G_b(s) - G_h(s))/\sigma$ *is all-pass.*

(2) $\|G(s) - G_b(s)\|_\infty \le 2\sigma$.

(3) If $\Sigma_1 > \sigma I$ *then* $\|G(s) - G_b(s)\|_H \le 2\sigma$.

Proof. (1) In order to prove that $(G_b(s) - G_h(s))/\sigma$ is all-pass we note that

$$G_b(s) - G_h(s) = \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & \tilde{D} \end{array}\right]$$

where

$$\tilde{A} = \begin{bmatrix} A_{11} & 0 \\ 0 & \hat{A} \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} B_1 \\ \hat{B} \end{bmatrix}, \quad \tilde{C} = \begin{bmatrix} C_1 & -\hat{C} \end{bmatrix}, \quad \tilde{D} = -\hat{D}.$$

Now Theorem 8.4, part (1) gives that the solutions to Lyapunov equations

$$\tilde{A}\tilde{P} + \tilde{P}\tilde{A}^* + \tilde{B}\tilde{B}^* = 0 \tag{8.72}$$

$$\tilde{A}^*\tilde{Q} + \tilde{Q}\tilde{A} + \tilde{C}^*\tilde{C} = 0 \tag{8.73}$$

are

$$\tilde{P} = \begin{bmatrix} \Sigma_1 & I \\ I & \Sigma_1\Gamma^{-1} \end{bmatrix}, \quad \tilde{Q} = \begin{bmatrix} \Sigma_1 & -\Gamma \\ -\Gamma & \Sigma_1\Gamma \end{bmatrix} \tag{8.74}$$

(This is verified by noting that the blocks of equations (8.72) and (8.73) are also blocks of equations (8.34) and (8.35) for P_e and Q_e.) Hence $\tilde{P}\tilde{Q} = \sigma^2 I$ and by Theorem 8.3 there exists $\tilde{D}$ such that $(G_b(s) - G_h(s))/\sigma$ is all-pass. That $\tilde{D} = -\hat{D}$ is an appropriate choice is verified from equations (8.54) to (8.56) and Theorem 8.3, part (2).

(2) $(G(s) - G_b(s))/\sigma = (G(s) - G_h(s))/\sigma + (G_b(s) - G_h(s))/\sigma$ but the first term on the right hand side is all-pass by Theorem 8.4, part (2) and the second term is all-pass by part (1) above. Hence $\|G(s) - G_b(s)\|_\infty \le 2\sigma$.

(3) Similarly using the fact that all-pass functions have unity Hankel norms gives that

$$\|G(s) - G_b(s)\|_H \le \|G(s) - G_h(s)\|_H + \|G_b(s) - G_h(s)\|_H = 2\sigma$$

(Note that $G_h(s)$ is stable if $\Sigma_1 > \sigma I$.) □

Given the results of Lemma 8.15 bounds on the error in a truncated balanced realization are easily proved as follows.

Theorem 8.16 *Let* $G(s)$ *be a stable, rational,* $p \times m$, *transfer function with Hankel singular values* $\sigma_1 > \sigma_2 \ldots > \sigma_N$, *where each* σ_i *has multiplicity* r_i *and let* $\tilde{G}_k(s)$ *be obtained by truncating the balanced realization of* $G(s)$ *to the first* $(r_1 + r_2 + \ldots + r_k)$ *states. Then*

(1) $\left\|G(s) - \tilde{G}_k(s)\right\|_\infty \le 2(\sigma_{k+1} + \sigma_{k+2} + \ldots + \sigma_N)$.

(2) $\left\|G(s) - \tilde{G}_k(s)\right\|_H \le 2(\sigma_{k+1} + \sigma_{k+2} + \ldots + \sigma_N)$.

Proof. If $p \neq m$ then augmenting B or C by zero columns or rows, respectively, will still give a balanced realization and the same argument is valid. Hence assume $p = m$. Notice that since truncation of balanced realization are also balanced, satisfying the truncated Lyapunov equations, the Hankel singular values of $\tilde{G}_i(s)$ will be $\sigma_1, \sigma_2, \ldots, \sigma_i$ with multiplicities $r_1, r_2, \ldots, r_i$, respectively. Also $\tilde{G}_i(s)$ can be obtained by truncating the balanced realization of $\tilde{G}_{i+1}(s)$ and hence $\left\| \tilde{G}_{i+1}(s) - \tilde{G}_i(s) \right\| \leq 2\sigma_{i+1}$ for both $\mathcal{L}_\infty$ and Hankel norms. Hence $(G_N(s) := G(s))$

$$\left\| G(s) - \tilde{G}_k(s) \right\| = \left\| \sum_{i=k}^{N-1} (\tilde{G}_{i+1}(s) - \tilde{G}_i(s)) \right\| \leq 2(\sigma_{k+1} + \sigma_{k+2} + \ldots + \sigma_N)$$

for both norms, and the proof is complete. □

8.6 Toeplitz Operators

In this section, we consider another operator. Again let $G(s) \in \mathcal{L}_\infty$. Then a *Toeplitz operator* associated with G is denoted by T_G and is defined as

$$T_G : \mathcal{H}_2 \longmapsto \mathcal{H}_2$$

$$T_G f := (P_+ M_G) f = P_+(Gf), \quad \text{for } f \in \mathcal{H}_2$$

i.e., $T_G = P_+ M_G|_{\mathcal{H}_2}$. In particular if $G(s) \in \mathcal{H}_\infty$ then $T_G = M_G|_{\mathcal{H}_2}$.

$$\begin{tikzcd} & \mathcal{L}_2 \arrow[d, "P_+"] \\ \mathcal{H}_2 \arrow[ur, "M_G"] \arrow[r, "T_G"'] & \mathcal{H}_2 \end{tikzcd}$$

Analogous to the Hankel operator, there is also a corresponding time domain description for the Toeplitz operator:

$$T_g : \mathcal{L}_2[0, \infty) \longmapsto \mathcal{L}_2[0, \infty)$$

$$T_g f := P_+(g * f) = \int_0^\infty g(t - \tau) f(\tau) d\tau, \quad t \geq 0$$

for $f(t) \in \mathcal{L}_2[0, \infty)$.

It is seen that the multiplication operator (in frequency domain) or the convolution operator (in time domain) plays an important role in the development of the Hankel and Toeplitz operators. In fact, a multiplication operator can be decomposed as several Hankel and Toeplitz operators: let $\mathcal{L}_2 = \mathcal{H}_2^\perp \oplus \mathcal{H}_2$ and $G \in \mathcal{L}_\infty$. Then the multiplication operator associated with G can be written as

$$M_G: \ \mathcal{H}_2^\perp \oplus \mathcal{H}_2 \longmapsto \mathcal{H}_2^\perp \oplus \mathcal{H}_2$$

$$M_G = \begin{bmatrix} P_- M_G|_{\mathcal{H}_2^\perp} & P_- M_G|_{\mathcal{H}_2} \\ P_+ M_G|_{\mathcal{H}_2^\perp} & P_+ M_G|_{\mathcal{H}_2} \end{bmatrix} = \begin{bmatrix} P_- M_G|_{\mathcal{H}_2^\perp} & \Gamma^*_{G^\sim} \\ \Gamma_G & T_G \end{bmatrix}.$$

Note that $P_- M_G|_{\mathcal{H}_2^\perp}$ is actually a Toeplitz operator; however, we will not discuss it further here. A fact worth mentioning is that if G is causal, i.e., $G \in \mathcal{H}_\infty$, then $\Gamma^*_{G^\sim} = 0$ and M_G is a lower block triangular operator. In fact, it can be shown that G is causal if and only if $\Gamma^*_{G^\sim} = 0$, i.e., the future input does not affect the past output. This is yet another characterization of causality.

8.7 Hankel and Toeplitz Operators on the Disk*

It is sometimes more convenient and insightful to study operators on the disk. From the control system point of view, some operators are much easier to interpret and compute in discrete time. The objective of this section is to examine the Hankel and Toeplitz operators in discrete time (i.e., on the disk) and, hopefully, to give the readers some intuitive ideas about these operators since they are very important in the $\mathcal{H}_\infty$ control theory. To start with, it is necessary to introduce some function spaces in respect to a unit disk.

Let $\mathbb{D}$ denote the unit disk :

$$\mathbb{D} := \{\lambda \in \mathbb{C} : \ |\lambda| < 1\},$$

and let $\overline{\mathbb{D}}$ and $\partial\mathbb{D}$ denote its closure and boundary, respectively:

$$\overline{\mathbb{D}} := \{\lambda \in \mathbb{C} : \ |\lambda| \leq 1\}$$

$$\partial\mathbb{D} := \{\lambda \in \mathbb{C} : \ |\lambda| = 1\}.$$

Let $\mathcal{L}_2(\partial\mathbb{D})$ denote the Hilbert space of matrix valued functions defined on the unit circle $\partial\mathbb{D}$ as

$$\mathcal{L}_2(\partial\mathbb{D}) = \left\{F(\lambda) : \frac{1}{2\pi}\int_0^{2\pi} \text{Trace}\left[F^*(e^{j\theta})F(e^{j\theta})\right] d\theta < \infty\right\}$$

with inner product

$$\langle F, G\rangle := \frac{1}{2\pi}\int_0^{2\pi} \text{Trace}\left[F^*(e^{j\theta})G(e^{j\theta})\right] d\theta.$$

Furthermore, let $\mathcal{H}_2(\partial\mathbb{D})$ be the (closed) subspace of $\mathcal{L}_2(\partial\mathbb{D})$ with matrix functions $F(\lambda)$ analytic in $\mathbb{D}$, i.e.,

$$\mathcal{H}_2(\partial\mathbb{D}) = \left\{ F(\lambda) \in \mathcal{L}_2(\partial\mathbb{D}) : \frac{1}{2\pi}\int_0^{2\pi} F(e^{j\theta})e^{jn\theta}d\theta = 0, \quad \text{for all } n > 0 \right\},$$

and let $\mathcal{H}_2^{\perp}(\partial\mathbb{D})$ be the (closed) subspace of $\mathcal{L}_2(\partial\mathbb{D})$ with matrix functions $F(\lambda)$ analytic in $\mathbb{C}\backslash\overline{\mathbb{D}}$.

It can be shown that the Fourier transform (or bilateral $\mathcal{Z}$ transform) gives the following isometric isomorphism:

$$\begin{aligned} \mathcal{L}_2(\partial\mathbb{D}) &\cong l_2(-\infty,\infty) \\ \mathcal{H}_2(\partial\mathbb{D}) &\cong l_2[0,\infty) \\ \mathcal{H}_2^{\perp}(\partial\mathbb{D}) &\cong l_2(-\infty,0). \end{aligned}$$

Remark 8.3 It should be kept in mind that, in contrast to the variable z in the standard Z-transform, λ here denotes $\lambda = z^{-1}$. ♡

Analogous to the space in the half plane, $\mathcal{L}_\infty(\partial\mathbb{D})$ is used to denote the Banach space of essentially bounded matrix functions with norm

$$\|F\|_\infty = \operatorname*{ess\,sup}_{\theta\in[0,2\pi]} \overline{\sigma}\left[F(e^{j\theta})\right].$$

The Hardy space $\mathcal{H}_\infty(\partial\mathbb{D})$ is the closed subspace of $\mathcal{L}_\infty(\partial\mathbb{D})$ with functions analytic in $\mathbb{D}$ and is defined as

$$\mathcal{H}_\infty(\partial\mathbb{D}) = \left\{ F(\lambda) \in \mathcal{L}_\infty(\partial\mathbb{D}) : \int_0^{2\pi} F(e^{j\theta})e^{jn\theta}d\theta = 0, \quad \text{for all } n > 0 \right\}.$$

The $\mathcal{H}_\infty$ norm is defined as

$$\|F\|_\infty := \sup_{\lambda\in\mathbb{D}} \overline{\sigma}\left[F(\lambda)\right] = \operatorname*{ess\,sup}_{\theta\in[0,2\pi]} \overline{\sigma}\left[F(e^{j\theta})\right].$$

It is easy to see that $\mathcal{L}_\infty(\partial\mathbb{D}) \subset \mathcal{L}_2(\partial\mathbb{D})$ and $\mathcal{H}_\infty(\partial\mathbb{D}) \subset \mathcal{H}_2(\partial\mathbb{D})$. (However, it should be pointed out that these inclusions are not true for functions in the half planes or for continuous time functions.)

Example 8.4 Let $F(\lambda) \in \mathcal{L}_2(\partial\mathbb{D})$, and let $F(\lambda)$ have a power series representation as follows:

$$F(\lambda) = \sum_{i=-\infty}^{\infty} F_i\lambda^i.$$

Then $F(\lambda) \in \mathcal{H}_2(\partial\mathbb{D})$ if and only if $F_i = 0$ for all $i < 0$ and $F(\lambda) \in \mathcal{H}_2^{\perp}(\partial\mathbb{D})$ if and only if $F_i = 0$ for all $i \geq 0$. ◇

Now let P_- and P_+ denote the orthogonal projections:

$$P_+ : \ l_2(-\infty,\infty) \longmapsto l_2[0,\infty) \ \text{ or } \ \mathcal{L}_2(\partial\mathbb{D}) \longmapsto \mathcal{H}_2(\partial\mathbb{D})$$

$$P_- : \ l_2(-\infty,\infty) \longmapsto l_2(-\infty,0) \ \text{ or } \ \mathcal{L}_2(\partial\mathbb{D}) \longmapsto \mathcal{H}_2^{\perp}(\partial\mathbb{D}).$$

Suppose $G_d(\lambda) \in \mathcal{L}_\infty(\partial\mathbb{D})$. Then the Hankel operator associated with $G_d(\lambda)$ is defined as

$$\Gamma_{G_d} : \ l_2(-\infty,0) \longmapsto l_2[0,\infty) \ \text{ or } \ \mathcal{H}_2^{\perp}(\partial\mathbb{D}) \longmapsto \mathcal{H}_2(\partial\mathbb{D})$$

$$\Gamma_{G_d} = P_+ M_{G_d}|_{\mathcal{H}_2^{\perp}(\partial\mathbb{D})}.$$

Similarly, a Toeplitz operator associated with $G_d(\lambda)$ is defined as

$$T_{G_d} : \ l_2[0,\infty) \longmapsto l_2[0,\infty) \ \text{ or } \ \mathcal{H}_2(\partial\mathbb{D}) \longmapsto \mathcal{H}_2(\partial\mathbb{D})$$

$$T_{G_d} = P_+ M_{G_d}|_{\mathcal{H}_2(\partial\mathbb{D})}.$$

Now let $G_d(\lambda) = \sum_{i=-\infty}^{\infty} G_i \lambda^i \in \mathcal{L}_\infty(\partial\mathbb{D})$ be a linear system transfer matrix, $u(\lambda) = \sum_{i=-\infty}^{\infty} u_i \lambda^i \in \mathcal{L}_2(\partial\mathbb{D})$ be the system input in the frequency domain, and u_i be the input at time $t = i$. Then the system output in the frequency domain is given by

$$y(\lambda) = G_d(\lambda)u(\lambda) = \sum_{i=-\infty}^{\infty} \sum_{j=-\infty}^{\infty} G_i u_j \lambda^{i+j} =: \sum_{i=-\infty}^{\infty} y_i \lambda^i$$

where y_i is the system time response at time $t = i$, and consequently, we have

$$\left[\begin{array}{c} \bullet \\ \bullet \\ y_2 \\ y_1 \\ y_0 \\ \hdashline y_{-1} \\ y_{-2} \\ y_{-3} \\ \bullet \\ \bullet \end{array}\right] = \left[\begin{array}{ccccc:ccccc} \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & G_0 & G_1 & G_2 & G_3 & G_4 & G_5 & \bullet & \bullet \\ \bullet & \bullet & G_{-1} & G_0 & G_1 & G_2 & G_3 & G_4 & \bullet & \bullet \\ \bullet & \bullet & G_{-2} & G_{-1} & G_0 & G_1 & G_2 & G_3 & \bullet & \bullet \\ \hdashline \bullet & \bullet & G_{-3} & G_{-2} & G_{-1} & G_0 & G_1 & G_2 & \bullet & \bullet \\ \bullet & \bullet & G_{-4} & G_{-3} & G_{-2} & G_{-1} & G_0 & G_1 & \bullet & \bullet \\ \bullet & \bullet & G_{-5} & G_{-4} & G_{-3} & G_{-2} & G_{-1} & G_0 & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \end{array}\right] \left[\begin{array}{c} \bullet \\ \bullet \\ u_2 \\ u_1 \\ u_0 \\ \hdashline u_{-1} \\ u_{-2} \\ u_{-3} \\ \bullet \\ \bullet \end{array}\right]. \tag{8.75}$$

Equation (8.75) can be rewritten as

$$\left[\begin{array}{c} y_0 \\ y_1 \\ y_2 \\ \bullet \\ \bullet \\ \hdashline y_{-1} \\ y_{-2} \\ y_{-3} \\ \bullet \\ \bullet \end{array}\right] = \left[\begin{array}{ccccc:ccccc} G_0 & G_{-1} & G_{-2} & \bullet & \bullet & G_1 & G_2 & G_3 & \bullet & \bullet \\ G_1 & G_0 & G_{-1} & \bullet & \bullet & G_2 & G_3 & G_4 & \bullet & \bullet \\ G_2 & G_1 & G_0 & \bullet & \bullet & G_3 & G_4 & G_5 & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \hdashline G_{-1} & G_{-2} & G_{-3} & \bullet & \bullet & G_0 & G_1 & G_2 & \bullet & \bullet \\ G_{-2} & G_{-3} & G_{-4} & \bullet & \bullet & G_{-1} & G_0 & G_1 & \bullet & \bullet \\ G_{-3} & G_{-4} & G_{-5} & \bullet & \bullet & G_{-2} & G_{-1} & G_0 & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet & \bullet \end{array}\right] \left[\begin{array}{c} u_0 \\ u_1 \\ u_2 \\ \bullet \\ \bullet \\ \hdashline u_{-1} \\ u_{-2} \\ u_{-3} \\ \bullet \\ \bullet \end{array}\right]$$

$$=: \left[\begin{array}{c:c} T_1 & H_1 \\ \hdashline H_2 & T_2 \end{array}\right] \left[\begin{array}{c} u_0 \\ u_1 \\ u_2 \\ \bullet \\ \bullet \\ \hdashline u_{-1} \\ u_{-2} \\ u_{-3} \\ \bullet \\ \bullet \end{array}\right].$$

Matrices like T_1 and T_2 are called (block) *Toeplitz matrices* and matrices like H_1 and H_2 are called (block) *Hankel matrices*. In fact, H_1 is the matrix representation of the Hankel operator Γ_{G_d} and T_1 is the matrix representation of the Toeplitz operator T_{G_d}, and so on. Thus these operator norms can be computed from the matrix norms of their corresponding matrix representations.

Lemma 8.17 $\|G_d(\lambda)\|_\infty = \left\| \begin{bmatrix} T_1 & H_1 \\ H_2 & T_2 \end{bmatrix} \right\|$, *$\|\Gamma_{G_d}\| = \|H_1\|$, and $\|T_{G_d}\| = \|T_1\|$.*

Example 8.5 Let $G_d(\lambda) \in \mathcal{RH}_\infty(\partial\mathbb{D})$ and $G_d(\lambda) = C(\lambda^{-1}I - A)^{-1}B + D$ be a state space realization of $G_d(\lambda)$ and let A have all the eigenvalues in $\mathbb{D}$. Then

$$G_d(\lambda) = D + \sum_{i=0}^{\infty} \lambda^{i+1} C A^i B = \sum_{i=0}^{\infty} G_i \lambda^i$$

with $G_0 = D$ and $G_i = CA^{i-1}B, \forall i \geq 1$. The state space equations are given by

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k, \quad x_{-\infty} = 0, \\ y_k &= Cx_k + Du_k. \end{aligned}$$

Then for $u \in l_2(-\infty, 0)$, we have $x_0 = \sum_{i=1}^{\infty} A^{i-1}Bu_{-i}$, which defines the controllability operator

$$x_0 = \Psi_c u = \begin{bmatrix} B & AB & A^2B & \cdots \end{bmatrix} \begin{bmatrix} u_{-1} \\ u_{-2} \\ u_{-3} \\ \vdots \end{bmatrix} \in \mathbb{R}^n.$$

On the other hand, given x_0 and $u_k = 0, i \geq 0$, $x \in \mathbb{R}^n$, the output can be computed as

$$\Psi_o x_0 = \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \end{bmatrix} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} x_0 \in l_2[0, \infty)$$

which defines the observability operator. Of course, the adjoint operators of Ψ_c and Ψ_o can be computed easily as

$$\Psi_c^* x_0 = \begin{bmatrix} B^* \\ B^*A^* \\ B^*(A^*)^2 \\ \vdots \end{bmatrix} x_0 \in l_2(-\infty, 0)$$

and

$$\Psi_o^* y = \begin{bmatrix} C^* & A^*C^* & (A^*)^2C^* & \cdots \end{bmatrix} \begin{bmatrix} y_0 \\ y_1 \\ y_2 \\ \vdots \end{bmatrix} \in \mathbb{R}^n.$$

Hence, the Hankel operator has the following matrix representation:

$$H = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \begin{bmatrix} B & AB & A^2B & \cdots \end{bmatrix} = \begin{bmatrix} G_1 & G_2 & G_3 & \cdots \\ G_2 & G_3 & G_4 & \cdots \\ G_3 & G_4 & G_5 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

◇

It is interesting to establish some connections between a Hankel operator defined on the unit disk and the one defined on the half plane. First define the map as

$$\lambda = \frac{s-1}{s+1}, \quad s = \frac{1+\lambda}{1-\lambda}, \tag{8.76}$$

which maps the closed right-half plane $\mathrm{Re}(s) \geq 0$ onto the closed unit disk, $\overline{\mathbb{D}}$.

Suppose $G(s) \in \mathcal{H}_\infty(j\mathbb{R})$ and define

$$G_d(\lambda) := G(s)|_{s=\frac{1+\lambda}{1-\lambda}}. \tag{8.77}$$

Since $G(s)$ is analytic in $\mathrm{Re}(s) > 0$, including the point at ∞, $G_d(\lambda)$ is analytic in $\mathbb{D}$, i.e., $G_d(\lambda) \in \mathcal{H}_\infty(\partial\mathbb{D})$.

Lemma 8.18 $\left\|\Gamma_{G(s)}\right\| = \left\|\Gamma_{G_d(\lambda)}\right\|$.

Proof. Define the function

$$\phi(s) = \frac{\sqrt{2}}{s+1}.$$

The relation between a point $j\omega$ on the imaginary axis and the corresponding point $e^{j\theta}$ on the unit circle is, from (8.76),

$$e^{j\theta} = \frac{j\omega - 1}{j\omega + 1}.$$

This yields

$$d\theta = -\frac{2}{\omega^2+1}d\omega = -|\phi(j\omega)|^2 d\omega,$$

which implies that the mapping

$$f_d(\lambda) \longmapsto \phi(s)f(s) : \ \mathcal{H}_2(\partial\mathbb{D}) \longmapsto \mathcal{H}_2(j\mathbb{R})$$

where $f(s) = f_d(\lambda)|_{\lambda=\frac{s-1}{s+1}}$ is an isomorphism. Similarly,

$$f_d(\lambda) \longmapsto \phi(s)f(s) : \ \mathcal{H}_2^{\perp}(\partial\mathbb{D}) \longmapsto \mathcal{H}_2^{\perp}(j\mathbb{R})$$

is an isomorphism; note that if $f_d(\lambda) \in \mathcal{H}_2^{\perp}(\partial\mathbb{D})$, then $f_d(\infty) = 0$, so that $f(-1) = 0$, and hence $\phi(s)f(s)$ is analytic in $\mathrm{Re}(s) < 0$.

The lemma now follows from the commutative diagram

$$\begin{tikzcd}
\mathcal{H}_2^{\perp}(\partial\mathbb{D}) \arrow[r, "\Gamma_{G_d(\lambda)}"] \arrow[d] & \mathcal{H}_2(\partial\mathbb{D}) \arrow[d] \\
\mathcal{H}_2^{\perp}(j\mathbb{R}) \arrow[r, "\Gamma_{G(s)}"'] & \mathcal{H}_2(j\mathbb{R})
\end{tikzcd}$$

□

The above isomorphism between $\mathcal{H}_2(\partial\mathbb{D})$ and $\mathcal{H}_2(j\mathbb{R})$ has also the implication that every discrete time $\mathcal{H}_2$ control problem can be converted to an equivalent continuous time $\mathcal{H}_2$ control problem. It should be emphasized that the bilinear transformation is not an isomorphism between $\mathcal{H}_2(\partial\mathbb{D})$ and $\mathcal{H}_2(j\mathbb{R})$.

8.8 Nehari's Theorem*

As we have mentioned at the beginning of this chapter, a Hankel operator is closely related to an analytic approximation problem. In this section, we shall be concerned with approximating G by an anticausal transfer matrix, i.e., one analytic in $\mathrm{Re}(s) < 0$ (or $|\lambda| > 1$), where the approximation is with respect to the $\mathcal{L}_\infty$ norm. For convenience, let $\mathcal{H}_\infty^-(\partial\mathbb{D})$ denote the subspace of $\mathcal{L}_\infty(\partial\mathbb{D})$ with functions analytic in $|\lambda| > 1$. So $G(\lambda) \in \mathcal{H}_\infty^-(\partial\mathbb{D})$ if and only if $G(\lambda^{-1}) \in \mathcal{H}_\infty(\partial\mathbb{D})$.

Minimum Distance Problem

In the following, we shall establish that the distance in $\mathcal{L}_\infty$ from G to the nearest matrix in $\mathcal{H}_\infty^-$ equals $\|\Gamma_G\|$. This is the classical Nehari Theorem. Some applications and explicit construction of the approximation for the matrix case will be considered in the later chapters.

Theorem 8.19 *Suppose $G \in \mathcal{L}_\infty$, then*

$$\inf_{Q \in \mathcal{H}_\infty^-} \|G - Q\|_\infty = \|\Gamma_G\|$$

and the infimum is achieved.

Remark 8.4 Note that from the mapping (8.76) and Lemma 8.18, the above theorem is the same whether the problem is on the unit disk or on the half plane. Hence, we will only give the proof on the unit disk. ♡

Proof. We shall give the proof on the unit disk. First, it is easy to establish that the Hankel norm is the lower bound: for fixed $Q \in \mathcal{H}_\infty^-(\partial\mathbb{D})$, we have

$$\begin{aligned} \|G - Q\|_\infty &= \sup_{f \in \mathcal{H}_2^\perp(\partial\mathbb{D})} \frac{\|(G-Q)f\|_2}{\|f\|_2} \\ &\geq \sup_{f \in \mathcal{H}_2^\perp(\partial\mathbb{D})} \frac{\|P_+(G-Q)f\|_2}{\|f\|_2} \\ &= \sup_{f \in \mathcal{H}_2^\perp(\partial\mathbb{D})} \frac{\|P_+ G f\|_2}{\|f\|_2} \\ &= \|\Gamma_G\|. \end{aligned}$$

Next we shall construct a $Q(\lambda) \in \mathcal{H}_\infty^-(\partial\mathbb{D})$ such that

$$\|G - Q\|_\infty = \|\Gamma_G\|. \tag{8.78}$$

Let the power series expansion of $G(\lambda)$ and $Q(\lambda)$ be

$$\begin{aligned} G(\lambda) &= \sum_{i=-\infty}^{\infty} \lambda^i G_i \\ Q(\lambda) &= \sum_{i=-\infty}^{0} \lambda^i Q_i. \end{aligned}$$

The left-hand side of (8.78) equals the norm of the operator

$$f \longmapsto (G(\lambda) - Q(\lambda))f(\lambda) : \ \mathcal{H}_2^\perp(\partial\mathbb{D}) \longmapsto \mathcal{L}_2(\partial\mathbb{D}).$$

From the discussion of the last section, there is a matrix representation of this operator:

$$\begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ G_3 & G_4 & G_5 & \bullet & \bullet \\ G_2 & G_3 & G_4 & \bullet & \bullet \\ G_1 & G_2 & G_3 & \bullet & \bullet \\ G_0 - Q_0 & G_1 & G_2 & \bullet & \bullet \\ G_{-1} - Q_{-1} & G_0 - Q_0 & G_1 & \bullet & \bullet \\ G_{-2} - Q_{-2} & G_{-1} - Q_{-1} & G_0 - Q_0 & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \end{bmatrix}. \tag{8.79}$$

The idea in the construction of a Q to satisfy (8.78) is to select Q_0, Q_{-1}, $\ldots$, in turn to minimize the matrix norm of (8.79). First choose Q_0 to minimize

$$\left\| \left[\begin{array}{c|cccc} \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ G_3 & G_4 & G_5 & \bullet & \bullet \\ G_2 & G_3 & G_4 & \bullet & \bullet \\ G_1 & G_2 & G_3 & \bullet & \bullet \\ \hline G_0 - Q_0 & G_1 & G_2 & \bullet & \bullet \end{array} \right] \right\|.$$

By Parrott's Theorem (i.e., matrix dilation theory presented in Chapter 2), the minimum equals the norm of the following matrix

$$\begin{bmatrix} \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ G_3 & G_4 & G_5 & \bullet & \bullet \\ G_2 & G_3 & G_4 & \bullet & \bullet \\ G_1 & G_2 & G_3 & \bullet & \bullet \end{bmatrix}$$

which is the rotated Hankel matrix H_1. Hence, the minimum equals the Hankel norm $\|\Gamma_G\|$. Next, choose Q_{-1} to minimize

$$\left\| \left[\begin{array}{c|cccc} \bullet & \bullet & \bullet & \bullet & \bullet \\ \bullet & \bullet & \bullet & \bullet & \bullet \\ G_3 & G_4 & G_5 & \bullet & \bullet \\ G_2 & G_3 & G_4 & \bullet & \bullet \\ G_1 & G_2 & G_3 & \bullet & \bullet \\ G_0 - Q_0 & G_1 & G_2 & \bullet & \bullet \\ \hline G_{-1} - Q_{-1} & G_0 - Q_0 & G_1 & \bullet & \bullet \end{array} \right] \right\|.$$

Again, the minimum equals $\|\Gamma_G\|$. Continuing in this way gives a suitable Q. $\Box$

As we have mentioned earlier we might be interested in approximating an $\mathcal{L}_\infty$ function by an $\mathcal{H}_\infty$ function. Then we have the following corollary.

Corollary 8.20 *Suppose $G \in \mathcal{L}_\infty$, then*

$$\inf_{Q \in \mathcal{H}_\infty} \|G - Q\|_\infty = \left\|\hat{\Gamma}_G\right\|$$

and the infimum is achieved.

Proof. This follows from the fact that

$$\inf_{Q \in \mathcal{H}_\infty} \|G - Q\|_\infty = \inf_{Q^\sim \in \mathcal{H}_\infty^-} \|G^\sim - Q^\sim\|_\infty = \|\Gamma_{G^\sim}\| = \left\|\hat{\Gamma}_G\right\|.$$

□

Let $G \in \mathcal{RH}_\infty$ and g be the impulse response of G and let σ_1 be the largest Hankel singular value of G, i.e., $\|\Gamma_G\| = \sigma_1$. Suppose $u \in \mathcal{L}_2(-\infty, 0]$ (or $l_2(-\infty, 0)$) and $v \in \mathcal{L}_2[0, \infty)$ (or $l_2[0, \infty)$) are the corresponding Schmidt pairs:

$$\begin{aligned} \Gamma_g u &= \sigma_1 v \\ \Gamma_g^* v &= \sigma_1 u. \end{aligned}$$

Now denote the Laplace transform (or Z-transform) of u and v as $U \in \mathcal{RH}_2^\perp$ and $V \in \mathcal{RH}_2$.

Lemma 8.21 *Let $G \in \mathcal{RH}_\infty$ and σ_1 be the largest Hankel singular value of G. Then*

$$\inf_{Q \in \mathcal{H}_\infty^-} \|G - Q\|_\infty = \sigma_1$$

and

$$(G - Q)U = \sigma_1 V.$$

Moreover, if G is a scalar function, then $Q = G - \sigma_1 V/U$ is the unique solution and $G - Q$ is all-pass.

Proof. Let $H := (G - Q)U$ and note that $\Gamma_G U \in \mathcal{RH}_2$ and $P_+ H = P_+(GU) = \Gamma_G U$. Then

$$\begin{aligned} 0 &\le \|H - \Gamma_G U\|_2^2 \\ &= \|H\|_2^2 + \|\Gamma_G U\|_2^2 - \langle H, \Gamma_G U\rangle - \langle \Gamma_G U, H\rangle \\ &= \|H\|_2^2 + \|\Gamma_G U\|_2^2 - \langle P_+ H, \Gamma_G U\rangle - \langle \Gamma_G U, P_+ H\rangle \\ &= \|H\|_2^2 - \langle \Gamma_G U, \Gamma_G U\rangle \\ &= \|H\|_2^2 - \langle U, \Gamma_G^* \Gamma_G U\rangle \\ &= \|H\|_2^2 - \sigma_1^2 \langle U, U\rangle \\ &= \|H\|_2^2 - \sigma_1^2 \|U\|_2^2 \\ &\le \|G - Q\|_\infty^2 \|U\|_2^2 - \sigma_1^2 \|U\|_2^2 \\ &= 0. \end{aligned}$$

Hence, $H = \Gamma_G U$, i.e., $(G - Q)U = \Gamma_G U = \sigma_1 V$.

Now if G is a scalar function, then Q is uniquely determined and $Q = G - \sigma_1 V/U$. We shall prove through explicit construction below that V/U is an all-pass. □

Formulas for Continuous Time

Let $G(s) \in \mathcal{RH}_\infty$ and let

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right].$$

Let L_c and L_o be the corresponding controllability and observability Gramians:

$$AL_c + L_c A^* + BB^* = 0$$

$$A^* L_o + L_o A + C^* C = 0.$$

And let σ_1^2, η be the largest eigenvalue and the corresponding eigenvector of $L_c L_o$:

$$L_c L_o \eta = \sigma_1^2 \eta.$$

Define

$$\xi := \frac{1}{\sigma_1} L_o \eta.$$

Then

$$\begin{aligned} u &= \Psi_c^* \xi = B^* e^{-A^* \tau} \xi \in \mathcal{L}_2(-\infty, 0] \\ v &= \Psi_o \eta = C e^{At} \eta \in \mathcal{L}_2[0, \infty) \end{aligned}$$

are the Schmidt pair and

$$\begin{aligned} U(s) &= \left[\begin{array}{c|c} -A^* & \xi \\ \hline -B^* & 0 \end{array}\right] \in \mathcal{RH}_2^\perp \\ V(s) &= \left[\begin{array}{c|c} A & \eta \\ \hline C & 0 \end{array}\right] \in \mathcal{RH}_2 \quad . \end{aligned}$$

It is easy to show that if G is a scalar, then U, V are scalar transfer functions and $U^\sim U = V^\sim V$. Hence, V/U is an all-pass. The details are left as an exercise for the reader.

Formulas for Discrete Time

Similarly, let G be a discrete time transfer matrix and let

$$G(\lambda) = C(\lambda^{-1} I - A)^{-1} B = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right].$$

Let L_c and L_o be the corresponding controllability and observability Gramians:

$$AL_cA^* - L_c + BB^* = 0$$

$$A^*L_oA - L_o + C^*C = 0.$$

And let σ_1^2, η be the largest eigenvalue and a corresponding eigenvector of L_cL_o:

$$L_cL_o\eta = \sigma_1^2\eta.$$

Define

$$\xi := \frac{1}{\sigma_1}L_o\eta.$$

Then

$$\begin{aligned} u &= \Psi_c^*\xi = \begin{bmatrix} B^* \\ B^*A^* \\ B^*(A^*)^2 \\ \vdots \end{bmatrix} \xi \in l_2(-\infty, 0) \\ v &= \Psi_o\eta = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \end{bmatrix} \eta \in l_2[0, \infty) \end{aligned}$$

are the Schmidt pair and

$$\begin{aligned} U(\lambda) &= \sum_{i=1}^{\infty} B^*(A^*)^{i-1}\xi\lambda^{-i} = \left[\begin{array}{c|c} A & B \\ \hline \xi^* & 0 \end{array}\right]^{\sim} \in \mathcal{RH}_2^{\perp} \\ V(\lambda) &= \sum_{i=0}^{\infty} CA^i\eta\lambda^i = \left[\begin{array}{c|c} A & \eta \\ \hline CA & C\eta \end{array}\right] \in \mathcal{RH}_2 \end{aligned}$$

where $W(\lambda)^{\sim} := W^T(\lambda^{-1})$.

Alternatively, since the Hankel operator has a matrix representation H, u and v can be obtained from a "singular value decomposition" (possibly infinite size): let

$$G(\lambda) = \sum_{i=1}^{\infty} G_i\lambda^i, \quad G_i \in \mathbb{R}^{p\times q}$$

(in fact, $G_i = CA^{i-1}B$ if a state space realization of G is available) and let

$$H = \begin{bmatrix} G_1 & G_2 & G_3 & \cdots \\ G_2 & G_3 & G_4 & \cdots \\ G_3 & G_4 & G_5 & \cdots \\ \vdots & \vdots & \vdots & \vdots \end{bmatrix}.$$

Suppose H has a "singular value decomposition"

$$\begin{aligned} H &= U\Sigma V^* \\ \Sigma &= \operatorname{diag}\{\sigma_1, \sigma_2, \cdots\} \\ U &= \begin{bmatrix} u_1 & u_2 & \cdots \end{bmatrix} \\ V &= \begin{bmatrix} v_1 & v_2 & \cdots \end{bmatrix} \end{aligned}$$

with $U^*U = UU^* = I$ and $V^*V = VV^* = I$. Then

$$u = u_1 = \begin{bmatrix} u_{11} \\ u_{12} \\ \vdots \end{bmatrix}, v = v_1 = \begin{bmatrix} v_{11} \\ v_{12} \\ \vdots \end{bmatrix}$$

where u and v are partitioned such that $u_{1i} \in \mathbb{R}^p$ and $v_{1i} \in \mathbb{R}^q$. Finally, $U(\lambda)$ and $V(\lambda)$ can be obtained as

$$\begin{aligned} U(\lambda) &= \sum_{i=1}^{\infty} u_{1i}\lambda^{-i} \in \mathcal{RH}_2^{\perp} \\ V(\lambda) &= \sum_{i=1}^{\infty} v_{1i}\lambda^{i-1} \in \mathcal{RH}_2 \quad . \end{aligned}$$

In particular, if $G(\lambda)$ is an n-th order matrix polynomial, then matrix H has only a finite number of nonzero elements and

$$H = \begin{bmatrix} H_n & 0 \\ 0 & 0 \end{bmatrix}$$

with

$$H_n = \begin{bmatrix} G_1 & G_2 & \cdots & G_{n-1} & G_n \\ G_2 & G_3 & \cdots & G_n & 0 \\ G_3 & G_4 & \cdots & 0 & 0 \\ \vdots & \vdots & & \vdots & \vdots \\ G_n & 0 & \cdots & 0 & 0 \end{bmatrix}.$$

Hence, u and v can be obtained from a standard SVD of H_n.

8.9 Notes and References

The study of the Hankel operators and of the optimal Hankel norm approximation theory can be found in Adamjan, Arov, and Krein [1978], Bettayeb, Silverman, and Safonov [1980], Kung and Lin [1981], Power [1982], Francis [1987], Kavranoğlu and Bettayeb [1994] and references therein. The presentation of this chapter is based on Francis [1987] and Glover [1984,1989].

9

Model Uncertainty and Robustness

In this chapter we briefly describe various types of uncertainties which can arise in physical systems, and we single out "unstructured uncertainties" as generic errors which are associated with all design models. We obtain robust stability tests for systems under various model uncertainty assumptions through the use of the small gain theorem. And we also obtain some sufficient conditions for robust performance under unstructured uncertainties. The difficulty associated with MIMO robust performance design and the role of plant condition numbers for systems with skewed performance and uncertainty specifications are revealed. We show by examples that the classical gain margin and phase margin are insufficient indicators for the system robustness. A simple example is also used to indicate the fundamental difference between the robustness of an SISO system and that of an MIMO system. In particular, we show that applying the SISO analysis/design method to an MIMO system may lead to erroneous results.

9.1 Model Uncertainty

Most control designs are based on the use of a design model. The relationship between models and the reality they represent is subtle and complex. A mathematical model provides a map from inputs to responses. The quality of a model depends on how closely its responses match those of the true plant. Since no single fixed model can respond exactly like the true plant, we need, at the very least, a set of maps. However, the modeling problem is much deeper – the universe of mathematical models from which a model set is chosen is distinct from the universe of physical systems. Therefore, a model set which includes the true physical plant can never be constructed. It is necessary for the engineer to make a leap of faith regarding the applicability of a particular design based on a mathematical model. To be practical, a design technique must help make this leap small by accounting for the inevitable inadequacy of models. A good model should be simple enough to facilitate design, yet complex enough to give the engineer confidence that designs based on the model will work on the true plant.

The term *uncertainty* refers to the differences or errors between models and reality, and whatever mechanism is used to express these errors will be called a *representation of uncertainty*. Representations of uncertainty vary primarily in terms of the amount of structure they contain. This reflects both our knowledge of the physical mechanisms which cause differences between the model and the plant and our ability to represent these mechanisms in a way that facilitates convenient manipulation. For example, consider the problem of bounding the magnitude of the effect of some uncertainty on the output of a nominally fixed linear system. A useful measure of uncertainty in this context is to provide a bound on the spectrum of the output's deviation from its nominal response. In the simplest case, this spectrum is assumed to be independent of the input. This is equivalent to assuming that the uncertainty is generated by an additive noise signal with a bounded spectrum; the uncertainty is represented as additive noise. Of course, no physical system is linear with additive noise, but some aspects of physical behavior are approximated quite well using this model. This type of uncertainty received a great deal of attention in the literature during the 1960's and 1970's, and the attention is probably due more to the elegant theoretical solutions that are yielded (e.g., white noise propagation in linear systems, Wiener and Kalman filtering, LQG) than to the great practical significance offered.

Generally, the deviation's spectrum of the true output from the nominal will depend significantly on the input. For example, an additive noise model is entirely inappropriate for capturing uncertainty arising from variations in the material properties of physical plants. The actual construction of model sets for more general uncertainty can be quite difficult. For example, a set membership statement for the parameters of an otherwise known FDLTI model is a highly-structured representation of uncertainty. It typically arises from the use of linear incremental models at various operating points, e.g., aerodynamic coefficients in flight control vary with flight environment and aircraft configurations, and equation coefficients in power plant control vary with aging, slag buildup, coal composition, etc. In each case, the amounts of variation and any known relationships between parameters can be expressed by confining the parameters to appropriately defined subsets of parameter space. However, for certain classes of signals (e.g., high frequency), the parameterized FDLTI model fails to describe the plant because the plant will always have dynamics which are not represented in the fixed order model.

In general, we are forced to use not just a single parameterized model but model sets that allow for plant dynamics which are not explicitly represented in the model structure.

A simple example of this involves using frequency-domain bounds on transfer functions to describe a model set. To use such sets to describe physical systems, the bounds must roughly grow with frequency. In particular, at sufficiently high frequencies, phase is completely unknown, i.e., $\pm 180^o$ uncertainties. This is a consequence of dynamic properties which inevitably occur in physical systems. This gives a less structured representation of uncertainty.

Examples of less-structured representations of uncertainty are direct set membership statements for the transfer function matrix of the model. For instance, the statement

$$P_\Delta(s) = P(s) + W_1(s)\Delta(s)W_2(s), \quad \overline{\sigma}[\Delta(j\omega)] < 1,\ \forall \omega \geq 0, \tag{9.1}$$

where W_1 and W_2 are stable transfer matrices that characterize the spatial and frequency structure of the uncertainty, confines the matrix P_Δ to a neighborhood of the nominal model P. In particular, if $W_1 = I$ and $W_2 = w(s)I$ where $w(s)$ is a scalar function, then P_Δ describes a disk centered at P with radius $w(j\omega)$ at each frequency as shown in Figure 9.1. The statement does not imply a mechanism or structure which gives rise to Δ. The uncertainty may be caused by parameter changes, as mentioned above or by neglected dynamics or by a host of other unspecified effects. An alternative statement to (9.1) is the so-called multiplicative form:

$$P_\Delta(s) = (I + W_1(s)\Delta(s)W_2(s))P(s). \tag{9.2}$$

This statement confines P_Δ to a normalized neighborhood of the nominal model P. An advantage of (9.2) over (9.1) is that in (9.2) compensated transfer functions have the same uncertainty representation as the raw model (i.e., the weighting functions apply to PK as well as P). Some other alternative set membership statements will be discussed later.

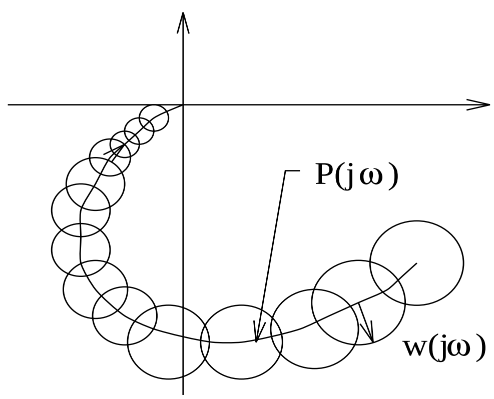


Figure 9.1: Nyquist Diagram of an Uncertain Model

The best choice of uncertainty representation for a specific FDLTI model depends, of course, on the errors the model makes. In practice, it is generally possible to represent some of these errors in a highly-structured parameterized form. These are usually the low frequency error components. There are always remaining higher frequency errors, however, which cannot be covered this way. These are caused by such effects as infinite-dimensional electro-mechanical resonance, time delays, diffusion processes, etc. Fortunately, the less-structured representations, e.g., (9.1) or (9.2), are well suited to represent this latter class of errors. Consequently, (9.1) and (9.2) have become widely used "generic" uncertainty representations for FDLTI models. An important point is that the construction of the weighting matrices W_1 and W_2 for multivariable systems is not trivial.

Motivated from these observations, we will focus for the moment on the multiplicative description of uncertainty. We will assume that P_Δ in (9.2) remains a strictly proper FDLTI system for all Δ. More general perturbations (e.g., time varying, infinite dimensional, nonlinear) can also be covered by this set provided they are given appropriate "conic sector" interpretations via Parseval's theorem. This connection is developed in [Safonov, 1980] and [Zames, 1966] and will not be pursued here.

When used to represent the various high frequency mechanisms mentioned above, the weighting functions in (9.2) commonly have the properties illustrated in Figure 9.2. They are small ($\ll 1$) at low frequencies and increase to unity and above at higher frequencies. The growth with frequency inevitably occurs because phase uncertainties eventually exceed ± 180 degrees and magnitude deviations eventually exceed the nominal transfer function magnitudes. Readers who are skeptical about this reality are encouraged to try a few experiments with physical devices.

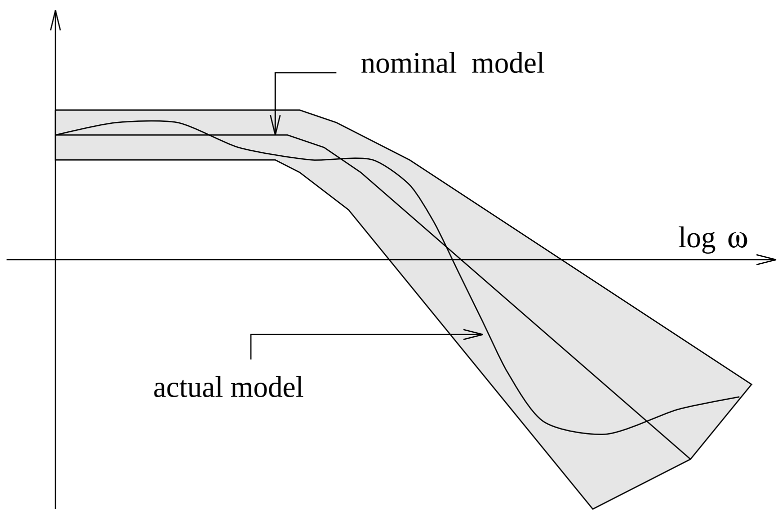


Figure 9.2: Typical Behavior of Multiplicative Uncertainty: $p_\delta(s) = [1 + w(s)\delta(s)]p(s)$

Also note that the representation of uncertainty in (9.2) can be used to include perturbation effects that are in fact certain. A nonlinear element, may be quite accurately modeled, but because our design techniques cannot effectively deal with the nonlinearity, it is treated as a conic linearity[1]. As another example, we may deliberately choose to ignore various known dynamic characteristics in order to achieve a simple nominal design model. This is the model reduction process discussed in the previous chapters.

The following terminologies are used in this book.

Definition 9.1 Given the description of an uncertainty model set $\mathbf{\Pi}$ and a set of performance objectives, suppose $P \in \mathbf{\Pi}$ is the nominal design model and K is the resulting controller. Then the closed-loop feedback system is said to have

Nominal Stability (NS): if K internally stabilizes the nominal model P.

Robust Stability (RS): if K internally stabilizes every plant belong to $\mathbf{\Pi}$.

Nominal Performance (NP): if the performance objectives are satisfied for the nominal plant P.

Robust Performance (RP): if the performance objectives are satisfied for every plant belong to $\mathbf{\Pi}$.

The nominal stability and performance can be easily checked using various standard techniques. The conditions for which the robust stability and robust performance are satisfied under various assumptions on the uncertainty set $\mathbf{\Pi}$ will be considered in the following sections.

9.2 Small Gain Theorem

This section and the next section consider the stability test of a nominally stable system under unstructured perturbations. The basis for the robust stability criteria derived in the sequel is the so-called *small gain theorem.*

Consider the interconnected system shown in Figure 9.3 with $M(s)$ a stable $p \times q$ transfer matrix.

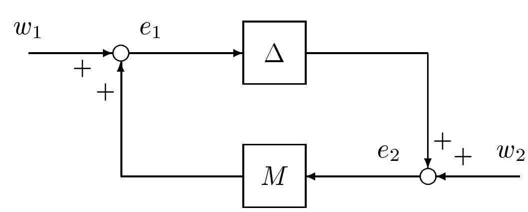


Figure 9.3: Small Gain Theorem

[1]See, for example, Safonov [1980] and Zames [1966].

Theorem 9.1 (Small Gain Theorem) *Suppose $M \in \mathcal{RH}_\infty$. Then the interconnected system shown in Figure 9.3 is well-posed and internally stable for all $\Delta(s) \in \mathcal{RH}_\infty$ with*

(a) $\|\Delta\|_\infty \leq 1/\gamma$ if and only if $\|M(s)\|_\infty < \gamma$;

(b) $\|\Delta\|_\infty < 1/\gamma$ if and only if $\|M(s)\|_\infty \leq \gamma$.

Proof. We shall only prove part (a). The proof for part (b) is similar. Without loss of generality, assume $\gamma = 1$.

(Sufficiency) It is clear that $M(s)\Delta(s)$ is stable since both $M(s)$ and $\Delta(s)$ are stable. Thus by Theorem 5.7 (or Corollary 5.6) the closed-loop system is stable if $\det(I - M\Delta)$ has no zero in the closed right-half plane for all $\Delta \in \mathcal{RH}_\infty$ and $\|\Delta\|_\infty \leq 1$. Equivalently, the closed-loop system is stable if

$$\inf_{s \in \overline{\mathbb{C}}_+} \underline{\sigma}\left(I - M(s)\Delta(s)\right) \neq 0$$

for all $\Delta \in \mathcal{RH}_\infty$ and $\|\Delta\|_\infty \leq 1$. But this follows from

$$\inf_{s \in \overline{\mathbb{C}}_+} \underline{\sigma}\left(I - M(s)\Delta(s)\right) \geq 1 - \sup_{s \in \overline{\mathbb{C}}_+} \bar{\sigma}\left(M(s)\Delta(s)\right) = 1 - \|M(s)\Delta(s)\|_\infty \geq 1 - \|M(s)\|_\infty > 0.$$

(Necessity) This will be shown by contradiction. Suppose $\|M\|_\infty \geq 1$. We will show that there exists a $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty \leq 1$ such that $\det(I - M(s)\Delta(s))$ has a zero on the imaginary axis, so the system is unstable. Suppose $\omega_0 \in \mathbb{R}_+ \cup \{\infty\}$ is such that $\bar{\sigma}(M(j\omega_0)) \geq 1$. Let $M(j\omega_0) = U(j\omega)\Sigma(j\omega_0)V^*(j\omega_0)$ be a singular value decomposition with

$$U(j\omega_0) = \begin{bmatrix} u_1 & u_2 & \cdots & u_p \end{bmatrix}$$
$$V(j\omega_0) = \begin{bmatrix} v_1 & v_2 & \cdots & v_q \end{bmatrix}$$
$$\Sigma(j\omega_0) = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix}$$

and $\|M\|_\infty = \overline{\sigma}\left(M(j\omega_0)\right) = \sigma_1$. To obtain a contradiction, it now suffices to construct a $\Delta \in \mathcal{RH}_\infty$ such that $\Delta(j\omega_0) = \frac{1}{\sigma_1} v_1 u_1^*$ and $\|\Delta\|_\infty \leq 1$. Indeed, for such $\Delta(s)$,

$$\det(I - M(j\omega_0)\Delta(j\omega_0)) = \det(I - U\Sigma V^* v_1 u_1^*/\sigma_1) = 1 - u_1^* U \Sigma V^* v_1/\sigma_1 = 0$$

and thus the closed-loop system is either not well-posed (if $\omega_0 = \infty$) or unstable (if $\omega \in \mathbb{R}$). There are two different cases:

(1) $\omega_0 = 0$ or ∞: then U and V are real matrices. In this case, $\Delta(s)$ can be chosen as

$$\Delta = \frac{1}{\sigma_1} v_1 u_1^* \in \mathbb{R}^{q \times p}.$$

(2) $0 < \omega_0 < \infty$: write u_1 and v_1 in the following form:

$$u_1 = \begin{bmatrix} u_{11}e^{j\theta_1} \\ u_{12}e^{j\theta_2} \\ \vdots \\ u_{1p}e^{j\theta_p} \end{bmatrix}, \quad v_1 = \begin{bmatrix} v_{11}e^{j\phi_1} \\ v_{12}e^{j\phi_2} \\ \vdots \\ v_{1q}e^{j\phi_q} \end{bmatrix}$$

where $u_{1i} \in \mathbb{R}$ and $v_{1j} \in \mathbb{R}$ are chosen so that $\theta_i, \phi_j \in [-\pi, 0)$ for all i, j.

Choose $\beta_i \geq 0$ and $\alpha_j \geq 0$ so that

$$\angle\left(\frac{\beta_i - j\omega_0}{\beta_i + j\omega_0}\right) = \theta_i, \quad \angle\left(\frac{\alpha_j - j\omega_0}{\alpha_j + j\omega_0}\right) = \phi_j$$

for $i = 1, 2, \ldots, p$ and $j = 1, 2, \ldots, q$. Let

$$\Delta(s) = \frac{1}{\sigma_1}\begin{bmatrix} v_{11}\frac{\alpha_1 - s}{\alpha_1 + s} \\ \vdots \\ v_{1q}\frac{\alpha_q - s}{\alpha_q + s} \end{bmatrix}\begin{bmatrix} u_{11}\frac{\beta_1 - s}{\beta_1 + s} & \cdots & u_{1p}\frac{\beta_p - s}{\beta_p + s} \end{bmatrix} \in \mathcal{RH}_\infty.$$

Then $\|\Delta\|_\infty = 1/\sigma_1 \leq 1$ and $\Delta(j\omega_0) = \frac{1}{\sigma_1}v_1u_1^*$.

□

The theorem still holds even if Δ and M are infinite dimensional. This is summarized as the following corollary.

Corollary 9.2 *The following statements are equivalent:*

(i) The system is well-posed and internally stable for all $\Delta \in \mathcal{H}_\infty$ *with* $\|\Delta\|_\infty < 1/\gamma$;

(ii) The system is well-posed and internally stable for all $\Delta \in \mathcal{RH}_\infty$ *with* $\|\Delta\|_\infty < 1/\gamma$;

(iii) The system is well-posed and internally stable for all $\Delta \in \mathbb{C}^{q\times p}$ *with* $\|\Delta\| < 1/\gamma$;

(iv) $\|M\|_\infty \leq \gamma$.

Remark 9.1 It can be shown that the small gain condition is sufficient to guarantee internal stability even if Δ is a nonlinear and time varying "stable" operator with an appropriately defined stability notion, see Desoer and Vidyasagar [1975]. ♡

The following lemma shows that if $\|M\|_\infty > \gamma$, there exists a destabilizing Δ with $\|\Delta\|_\infty < 1/\gamma$ such that the closed-loop system has poles in the open right half plane. (This is stronger than what is given in the proof of Theorem 9.1.)

Lemma 9.3 *Suppose* $M \in \mathcal{RH}_\infty$ *and* $\|M\|_\infty > \gamma$. *Then there exists a* $\sigma_0 > 0$ *such that for any given* $\sigma \in [0, \sigma_0]$ *there exists a* $\Delta \in \mathcal{RH}_\infty$ *with* $\|\Delta\|_\infty < 1/\gamma$ *such that* $\det(I - M(s)\Delta(s))$ *has a zero on the axis* $\mathrm{Re}(s) = \sigma$.

Proof. Without loss of generality, assume $\gamma = 1$. Since $M \in \mathcal{RH}_\infty$ and $\|M\|_\infty > 1$, there is a sufficiently small $\sigma_0 > 0$ such that $\|M(\sigma_0 + s)\|_\infty := \sup_{Re(s)>0} \|M(\sigma_0 + s)\| > 1$. Now let $\sigma \in [0, \sigma_0]$. Then there exists a $0 < \omega_0 < \infty$ such that $\|M(\sigma + j\omega_0)\| > 1$.

Let $M(\sigma + j\omega_0) = U\Sigma V^*$ be a singular value decomposition with

$$U = \begin{bmatrix} u_1 & u_2 & \cdots & u_p \end{bmatrix}$$

$$V = \begin{bmatrix} v_1 & v_2 & \cdots & v_q \end{bmatrix}$$

$$\Sigma = \begin{bmatrix} \sigma_1 & & \\ & \sigma_2 & \\ & & \ddots \end{bmatrix}.$$

Write u_1 and v_1 in the following form:

$$u_1 = \begin{bmatrix} u_{11}e^{j\theta_1} \\ u_{12}e^{j\theta_2} \\ \vdots \\ u_{1p}e^{j\theta_p} \end{bmatrix}, \quad v_1 = \begin{bmatrix} v_{11}e^{j\phi_1} \\ v_{12}e^{j\phi_2} \\ \vdots \\ v_{1q}e^{j\phi_q} \end{bmatrix}$$

where $u_{1i} \in \mathbb{R}$ and $v_{1j} \in \mathbb{R}$ are chosen so that $\theta_i, \phi_j \in [-\pi, 0)$ for all i, j.

Choose $\beta_i \geq 0$ and $\alpha_j \geq 0$ so that

$$\angle\left(\frac{\beta_i - \sigma - j\omega_0}{\beta_i + \sigma + j\omega_0}\right) = \theta_i, \quad \angle\left(\frac{\alpha_j - \sigma - j\omega_0}{\alpha_j + \sigma + j\omega_0}\right) = \phi_j$$

for $i = 1, 2, \ldots, p$ and $j = 1, 2, \ldots, q$. Let

$$\Delta(s) = \frac{1}{\sigma_1}\begin{bmatrix} v_{11}\frac{\alpha_1 - s}{\alpha_1 + s} \\ \vdots \\ v_{1q}\frac{\alpha_q - s}{\alpha_q + s} \end{bmatrix} \begin{bmatrix} u_{11}\frac{\beta_1 - s}{\beta_1 + s} & \cdots & u_{1p}\frac{\beta_p - s}{\beta_p + s} \end{bmatrix} \in \mathcal{RH}_\infty.$$

Then $\|\Delta\|_\infty = 1/\sigma_1 < 1$ and $\det(I - M(\sigma + j\omega_0)\Delta(\sigma + j\omega_0)) = 0$. Hence $s = \sigma + j\omega_0$ is a zero for the transfer function $\det(I - M(s)\Delta(s))$. □

The above lemma plays a key role in the necessity proofs of many robust stability tests in the sequel.

9.3 Stability under Stable Unstructured Uncertainties

The small gain theorem in the last section will be used here to derive robust stability tests under various assumptions of model uncertainties. The modeling error Δ will again be assumed to be stable. (Most of the robust stability tests discussed in the sequel can be easily generalized to unstable Δ case with some mild assumptions on the number of unstable poles of the uncertain model, we encourage readers to fill in the details.) In addition, we assume that the modeling error Δ is suitably scaled with weighting functions W_1 and W_2, i.e., the uncertainty can be represented as $W_1 \Delta W_2$.

We shall consider the standard setup shown in Figure 9.4, where $\mathbf{\Pi}$ is the set of uncertain plants with $P \in \mathbf{\Pi}$ as the nominal plant and with K as the internally stabilizing controller for P. The sensitivity and complementary sensitivity matrix functions are defined as usual as

$$S_o = (I + PK)^{-1}, \quad T_o = I - S_o$$

and

$$S_i = (I + KP)^{-1}, \quad T_i = I - S_i.$$

Recall that the closed-loop system is well-posed and internally stable if and only if

$$\begin{bmatrix} I & K \\ -\Pi & I \end{bmatrix}^{-1} = \begin{bmatrix} (I + K\Pi)^{-1} & -K(I + \Pi K)^{-1} \\ (I + \Pi K)^{-1}\Pi & (I + \Pi K)^{-1} \end{bmatrix} \in \mathcal{RH}_\infty.$$

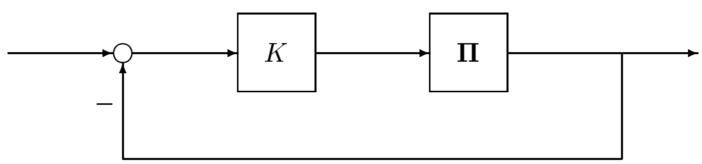


Figure 9.4: Unstructured Robust Stability Analysis

9.3.1 Additive Uncertainty

We assume that the model uncertainty can be represented by an additive perturbation:

$$\Pi = P + W_1 \Delta W_2.$$

Theorem 9.4 *Let* $\Pi = \{P + W_1 \Delta W_2 : \ \Delta \in \mathcal{RH}_\infty\}$ *and let* K *be a stabilizing controller for the nominal plant* P*. Then the closed-loop system is well-posed and internally stable for all* $\|\Delta\|_\infty < 1$ *if and only if* $\|W_2 K S_o W_1\|_\infty \leq 1$.

Proof. Let $\Pi = P + W_1 \Delta W_2$. Then

$$\begin{bmatrix} I & K \\ -\Pi & I \end{bmatrix}^{-1} = \begin{bmatrix} (I + KS_o W_1 \Delta W_2)^{-1} S_i & -KS_o(I + W_1 \Delta W_2 K S_o)^{-1} \\ (I + S_o W_1 \Delta W_2 K)^{-1} S_o (P + W_1 \Delta W_2) & S_o(I + W_1 \Delta W_2 K S_o)^{-1} \end{bmatrix}$$

is well-posed and internally stable if $(I + \Delta W_2 K S_o W_1)^{-1} \in \mathcal{RH}_\infty$ since

$$\begin{aligned} \det(I + KS_o W_1 \Delta W_2) &= \det(I + W_1 \Delta W_2 K S_o) = (I + S_o W_1 \Delta W_2 K) \\ &= \det(I + \Delta W_2 K S_o W_1). \end{aligned}$$

But $(I+\Delta W_2 K S_o W_1)^{-1} \in \mathcal{RH}_\infty$ is guaranteed if $\|\Delta W_2 K S_o W_1\|_\infty < 1$. Hence $\|W_2 K S_o W_1\|_\infty \leq 1$ is sufficient for robust stability.

To show the necessity, note that robust stability implies that

$$K(I+\Pi K)^{-1} = KS_o(I+W_1\Delta W_2 KS_o)^{-1} \in \mathcal{RH}_\infty$$

for all admissible Δ. This in turn implies that

$$\Delta W_2 K(I+\Pi K)^{-1} W_1 = I - (I+\Delta W_2 K S_o W_1)^{-1} \in \mathcal{RH}_\infty$$

for all admissible Δ. By small gain theorem, this is true for all $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty < 1$ only if $\|W_2 K S_o W_1\|_\infty \leq 1$. □

Similarly, it is easy to show that the closed-loop system is stable for all $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty \leq 1$ if and only if $\|W_2 K S_o W_1\|_\infty < 1$.

9.3.2 Multiplicative Uncertainty

In this section, we assume that the system model is described by the following set of multiplicative perturbation

$$P_\Delta = (I + W_1 \Delta W_2)P$$

with $W_1, W_2, \Delta \in \mathcal{RH}_\infty$. Consider the feedback system shown in the Figure 9.5.

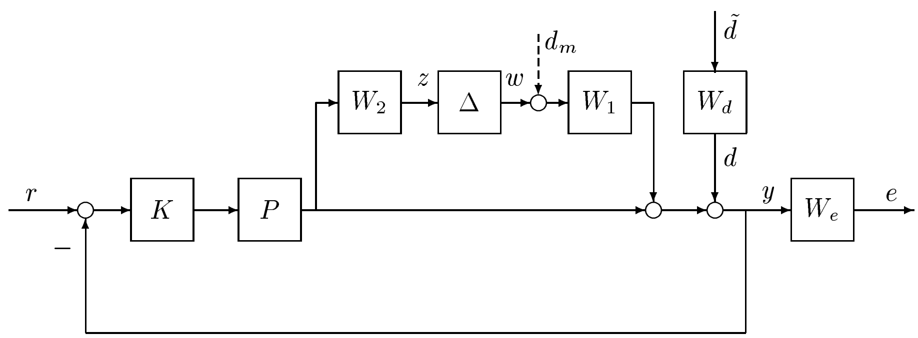


Figure 9.5: Output Multiplicative Perturbed Systems

Theorem 9.5 *Let* $\Pi = \{(I+W_1\Delta W_2)P : \quad \Delta \in \mathcal{RH}_\infty\}$ *and let* K *be a stabilizing controller for the nominal plant* P*. Then*

(i) the closed-loop system is well-posed and internally stable for all $\Delta \in \mathcal{RH}_\infty$ *with* $\|\Delta\|_\infty < 1$ *if and only if* $\|W_2 T_o W_1\|_\infty \leq 1$.

(ii) the closed-loop system is well-posed and internally stable for all $\Delta \in \mathcal{RH}_\infty$ *with* $\|\Delta\|_\infty \leq 1$ *if* $\|W_2 T_o W_1\|_\infty < 1$.

(iii) *the robust stability of the closed-loop system for all* $\Delta \in \mathcal{RH}_\infty$ *with* $\|\Delta\|_\infty \le 1$ *does not necessarily imply* $\|W_2 T_o W_1\|_\infty < 1$.

(iv) *the closed-loop system is well-posed and internally stable for all* $\Delta \in \mathcal{RH}_\infty$ *with* $\|\Delta\|_\infty \le 1$ *only if* $\|W_2 T_o W_1\|_\infty \le 1$.

(v) *In addition, assume that neither* P *nor* K *has poles on the imaginary axis. Then the closed-loop system is well-posed and internally stable for all* $\Delta \in \mathcal{RH}_\infty$ *with* $\|\Delta\|_\infty \le 1$ *if and only if* $\|W_2 T_o W_1\|_\infty < 1$.

Proof. We shall first prove that the condition in (i) is necessary for robust stability. Suppose $\|W_2 T_o W_1\|_\infty > 1$. Then by Lemma 9.3, for any given sufficiently small $\sigma > 0$, there is a $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty < 1$ such that $(I + \Delta W_2 T_o W_1)^{-1}$ has poles on the axis $\mathrm{Re}(s) = \sigma$. This implies

$$(I + \Pi K)^{-1} = S_o (I + W_1 \Delta W_2 T_o)^{-1}$$

has poles on the axis $\mathrm{Re}(s) = \sigma$ since σ can always be chosen so that the unstable poles are not cancelled by the zeros of S_o. Hence $\|W_2 T_o W_1\|_\infty \le 1$ is necessary for robust stability. In fact, we have also proven part (iv). The sufficiency parts of (i), (ii), and (v) follow from the small gain theorem.

To show the necessity part of (v), suppose $\|W_2 T_o W_1\|_\infty = 1$. From the proof of the small gain theorem, there is a $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty \le 1$ such that $(I + \Delta W_2 T_o W_1)^{-1}$ has poles on the imaginary axis. This implies

$$(I + \Pi K)^{-1} = S_o (I + W_1 \Delta W_2 T_o)^{-1}$$

has poles on the imaginary axis since the imaginary axis poles of $(I + W_1 \Delta W_2 T_o)^{-1}$ are not cancelled by the zeros of S_o, which are the poles of P and K. Hence $\|W_2 T_o W_1\|_\infty < 1$ is necessary for robust stability.

The proof of part (iii) is given below by exhibiting an example with $\|W_2 T_o W_1\|_\infty = 1$ but there is no destabilizing $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty \le 1$. □

Example 9.1 Let $P(s) = \frac{1}{s}$, $K(s) = 1$, and $W_1 = W_2 = 1$. It is easy to check that K stabilizes P. We have

$$T_o = \frac{1}{s+1}, \quad \|T_o\|_\infty = 1$$

and

$$(I + \Pi K)^{-1} = (I + K\Pi)^{-1} = K(I + \Pi K)^{-1} = \frac{s}{s+1} \frac{1}{1 + \frac{1}{s+1}\Delta}$$

$$(I + \Pi K)^{-1}\Pi = \frac{1}{s+1} \frac{1 + \Delta}{1 + \frac{1}{s+1}\Delta}.$$

Since $|T_o(s)| = \left|\frac{1}{s+1}\right| < 1$ for all $0 \neq s \in \mathbb{C}$ and $Re(s) \geq 0$, $1 + \frac{1}{s+1}\Delta \neq 0$ for all $0 \neq s \in \mathbb{C}$, $Re(s) \geq 0$, $\Delta \in \mathcal{RH}_\infty$ and $\|\Delta\| \leq 1$. The only point where $1 + \frac{1}{s+1}\Delta = 0$ in the closed right half plane is $s = 0$. Then $\Delta(0) = -1$. By assumption, Δ is analytic in a neighborhood of the origin since $\Delta \in \mathcal{RH}_\infty$. Hence, we can write

$$\Delta(s) = -1 + \sum_{i=1}^{\infty} a_n s^n, \quad a_i \in \mathbb{R}.$$

We now claim that $a_1 \geq 0$. Otherwise, $\left.\frac{d\Delta(s)}{ds}\right|_{s=0} = a_1 < 0$, along with $\Delta(0) = -1$, implies $\|\Delta\|_\infty > 1$. Hence $\Delta(s) = -1 + a_1 s + s^2 g(s)$ for some $g(s)$ and $a_1 \geq 0$ and

$$(I + \Pi K)^{-1} = (I + K\Pi)^{-1} = K(I + \Pi K)^{-1} = \frac{s}{s+1}\frac{1}{1 + \frac{1}{s+1}\Delta} = \frac{1}{1 + a_1 + sg}$$

$$(I + \Pi K)^{-1}\Pi = \frac{1}{s+1}\frac{1+\Delta}{1 + \frac{1}{s+1}\Delta} = \frac{a_1 + sg}{1 + a_1 + sg}$$

both of which are bounded in the neighborhood of the origin. Hence there is no destabilizing $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty \leq 1$. $\diamond$

The gap between the necessity and the sufficiency of the above robust stability conditions for the closed ball of uncertainty is only technique and will not be considered in the sequel. The reason for the existence of a such gap may be attributed to the fact that in the multiplicative perturbed case, the signals z, w, and d_m in Figure 9.5 are artificial and they are not physical signals. Indeed, $\Pi = (I + W_1 \Delta W_2)P$ is a single system, the internal stability of the closed-loop system does not necessarily imply the boundedness of the artificial signals z or w with respect to the artificial disturbance d_m. This is the case for the above example where $\Pi = (1 + \Delta)P = (a_1 s + s^2 g(s))/s = a_1 + sg(s)$ and the pole $s = 0$ is cancelled. This cancelation is artificial and is caused by the particular *model representation* (i.e., there is really no cancelation in the physical system.) Thus the closed-loop system is robustly stable although the transfer function from d_m to z is unstable.

9.3.3 Coprime Factor Uncertainty

As another example, consider a left coprime factor perturbed plant described in Figure 9.6.

Theorem 9.6 *Let*

$$\Pi = (\tilde{M} + \tilde{\Delta}_M)^{-1}(\tilde{N} + \tilde{\Delta}_N)$$

with $\tilde{M}, \tilde{N}, \tilde{\Delta}_M$, $\tilde{\Delta}_N \in \mathcal{RH}_\infty$. *The transfer matrices* $(\tilde{M}, \tilde{N})$ *are assumed to be a stable left coprime factorization of* P *(i.e.,* $P = \tilde{M}^{-1}\tilde{N}$*), and* K *internally stabilizes*

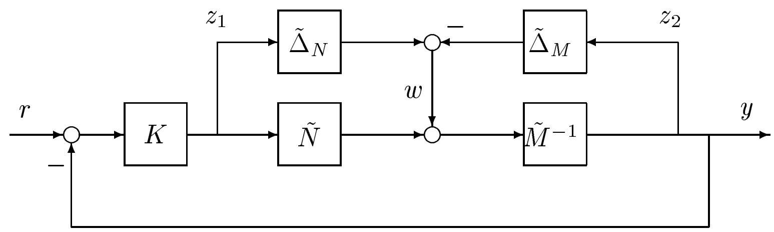


Figure 9.6: Left Coprime Factor Perturbed Systems

the nominal system P. Define $\Delta := \begin{bmatrix} \tilde{\Delta}_N & \tilde{\Delta}_M \end{bmatrix}$. Then the closed-loop system is well-posed and internally stable for all $\|\Delta\|_\infty < 1$ if and only if

$$\left\| \begin{bmatrix} K \\ I \end{bmatrix} (I + PK)^{-1} \tilde{M}^{-1} \right\|_\infty \le 1.$$

Proof. Let $K = UV^{-1}$ be a right coprime factorization over $\mathcal{RH}_\infty$. By Lemma 5.10, the closed-loop system is internally stable if and only if

$$\left((\tilde{N} + \tilde{\Delta}_N)U + (\tilde{M} + \tilde{\Delta}_M)V \right)^{-1} \in \mathcal{RH}_\infty. \tag{9.3}$$

Since K stabilizes P, $(\tilde{N}U + \tilde{M}V)^{-1} \in \mathcal{RH}_\infty$. Hence (9.3) holds if and only if

$$\left(I + (\tilde{\Delta}_N U + \tilde{\Delta}_M V)(\tilde{N}U + \tilde{M}V)^{-1} \right)^{-1} \in \mathcal{RH}_\infty.$$

By the small gain theorem, the above is true for all $\|\Delta\|_\infty < 1$ if and only if

$$\left\| \begin{bmatrix} U \\ V \end{bmatrix} (\tilde{N}U + \tilde{M}V)^{-1} \right\|_\infty = \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I + PK)^{-1} \tilde{M}^{-1} \right\|_\infty \le 1.$$

□

Similarly, one can show that the closed-loop system is well-posed and internally stable for all $\|\Delta\|_\infty \le 1$ if and only if

$$\left\| \begin{bmatrix} K \\ I \end{bmatrix} (I + PK)^{-1} \tilde{M}^{-1} \right\|_\infty < 1.$$

9.3.4 Unstructured Robust Stability Tests

Table 9.1 summaries robust stability tests on the plant uncertainties under various assumptions. All of the tests pertain to the standard setup shown in Figure 9.4, where $\mathbf{\Pi}$ is the set of uncertain plants with $P \in \mathbf{\Pi}$ as the nominal plant and with K as the internally stabilizing controller of P.

Table 9.1 should be interpreted as

UNSTRUCTURED ANALYSIS THEOREM
Given **NS & Perturbed Model Sets**
Then **Closed-Loop Robust Stability**
if and only if **Robust Stability Tests**

The table also indicates representative types of physical uncertainties which can be usefully represented by cone bounded perturbations inserted at appropriate locations. For example, the representation $P_\Delta = (I + W_1 \Delta W_2)P$ in the first row is useful for output errors at high frequencies (HF), covering such things as unmodeled high frequency dynamics of sensors or plant, including diffusion processes, transport lags, electro-mechanical resonances, etc. The representation $P_\Delta = P(I + W_1 \Delta W_2)$ in the second row covers similar types of errors at the inputs. Both cases should be contrasted with the third and the fourth rows which treat $P(I+W_1\Delta W_2)^{-1}$ and $(I+W_1\Delta W_2)^{-1}P$. These representations are more useful for variations in modeled dynamics, such as low frequency (LF) errors produced by parameter variations with operating conditions, with aging, or across production copies of the same plant. Discussion of still other cases is left to the table.

Note from the table that the stability requirements on Δ do not limit our ability to represent variations in either the number or locations of rhp singularities as can be seen from some simple examples.

Example 9.2 Suppose an uncertain *system with changing numbers of right-half plane poles* is described by

$$P_\Delta = \left\{ \frac{1}{s-\delta} : \ \delta \in \mathbb{R}, \ |\delta| \le 1 \right\}.$$

Then $P_1 = \frac{1}{s-1} \in P_\Delta$ has one right-half plane pole and $P_2 = \frac{1}{s+1} \in P_\Delta$ has no right-half plane pole. Nevertheless, the set of P_Δ can be covered by a set of feedback uncertain plants:

$$P_\Delta \subset \mathbf{\Pi} := \left\{ P(1+\delta P)^{-1} : \delta \in \mathcal{RH}_\infty, \ \ \|\delta\|_\infty \le 1 \right\}$$

with $P = \frac{1}{s}$. ◇

Example 9.3 As another example, consider the following set of plants:

$$P_\Delta = \frac{s+1+\alpha}{(s+1)(s+2)}, \ |\alpha| \le 2.$$

$W_1 \in \mathcal{RH}_\infty \quad W_2 \in \mathcal{RH}_\infty \quad \Delta \in \mathcal{RH}_\infty \quad \|\Delta\|_\infty < 1$		
Perturbed Model Sets $\mathbf{\Pi}$	Representative Types of Uncertainty Characterized	Robust Stability Tests
$(I + W_1 \Delta W_2)P$	output (sensor) errors neglected HF dynamics changing # of rhp zeros	$\|W_2 T_o W_1\|_\infty \le 1$
$P(I + W_1 \Delta W_2)$	input (actuators) errors neglected HF dynamics changing # of rhp zeros	$\|W_2 T_i W_1\|_\infty \le 1$
$(I + W_1 \Delta W_2)^{-1} P$	LF parameter errors changing # of rhp poles	$\|W_2 S_o W_1\|_\infty \le 1$
$P(I + W_1 \Delta W_2)^{-1}$	LF parameter errors changing # of rhp poles	$\|W_2 S_i W_1\|_\infty \le 1$
$P + W_1 \Delta W_2$	additive plant errors neglected HF dynamics uncertain rhp zeros	$\|W_2 K S_o W_1\|_\infty \le 1$
$P(I + W_1 \Delta W_2 P)^{-1}$	LF parameter errors uncertain rhp poles	$\|W_2 S_o P W_1\|_\infty \le 1$
$(\tilde{M} + \tilde{\Delta}_M)^{-1}(\tilde{N} + \tilde{\Delta}_N)$ $P = \tilde{M}^{-1}\tilde{N}$ $\Delta = \begin{bmatrix} \tilde{\Delta}_N & \tilde{\Delta}_M \end{bmatrix}$	LF parameter errors neglected HF dynamics uncertain rhp poles & zeros	$\left\\| \begin{bmatrix} K \\ I \end{bmatrix} S_o \tilde{M}^{-1} \right\\|_\infty \le 1$
$(N + \Delta_N)(M + \Delta_M)^{-1}$ $P = NM^{-1}$ $\Delta = \begin{bmatrix} \Delta_N \\ \Delta_M \end{bmatrix}$	LF parameter errors neglected HF dynamics uncertain rhp poles & zeros	$\left\\| M^{-1} S_i [K \ I] \right\\|_\infty \le 1$

Table 9.1: Unstructured Robust Stability Tests (# stands for 'the number'.)

This set of plants have *changing numbers of right-half plane zeros* since the plant has no right-half plane zero when $\alpha = 0$ and has one right-half plane zero when $\alpha = -2$. The uncertain plant can be covered by a set of multiplicative perturbed plants:

$$P_\Delta \subset \mathbf{\Pi} := \left\{ \frac{1}{s+2}(1 + \frac{2\delta}{s+1}), \; \delta \in \mathcal{RH}_\infty, \quad \|\delta\|_\infty \leq 1 \right\}.$$

It should be noted that this covering can be quite conservative. $\diamond$

9.3.5 Equivalence: Robust Stability vs. Nominal Performance

A robust stability problem can be viewed as another nominal performance problem. For example, the output multiplicative perturbed robust stability problem can be treated as a sensor noise rejection problem shown in Figure 9.7 and vise versa. It is clear that the system with output multiplicative uncertainty as shown in Figure 9.5 is robustly stable for $\|\Delta\|_\infty < 1$ iff the $\mathcal{H}_\infty$ norm of the transfer function from w to z, T_{zw}, is no greater than 1. Since $T_{zw} = T_{e\tilde{n}}$, hence $\|T_{zw}\|_\infty \leq 1$ iff $\sup_{\|\tilde{n}\|_2 \leq 1} \|e\|_2 = \|W_2 T_o W_1\|_\infty \leq 1$.

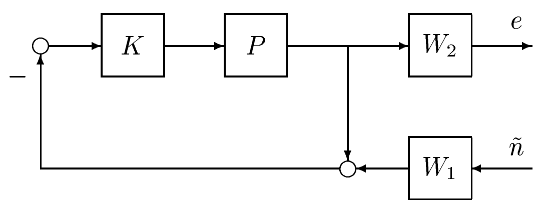


Figure 9.7: Equivalence Between Robust Stability With Output Multiplicative Uncertainty and Nominal Performance With Sensor Noise Rejection

There is, in fact, a much more general result along this line of equivalence: any robust stability problem (with open ball of uncertainty) can be regarded as an equivalent performance problem. This will be considered in Chapter 11.

9.4 Unstructured Robust Performance

Consider the perturbed system shown in Figure 9.8 with the set of perturbed models described by a set $\mathbf{\Pi}$. Suppose the weighting matrices $W_d, W_e \in \mathcal{RH}_\infty$ and the performance criterion is to keep the error e as small as possible in some sense for all possible models belong to the set $\mathbf{\Pi}$. In general, the set $\mathbf{\Pi}$ can be either parameterized set or unstructured set such as those described in Table 9.1. The performance specifications are usually specified in terms of the magnitude of each component e in time domain, i.e., $\mathcal{L}_\infty$ norm, with respect to bounded disturbances, or alternatively and more conveniently some requirements on the closed-loop frequency response of the transfer matrix between $\tilde{d}$ and e, say, integral of square error or the magnitude of the steady state error with

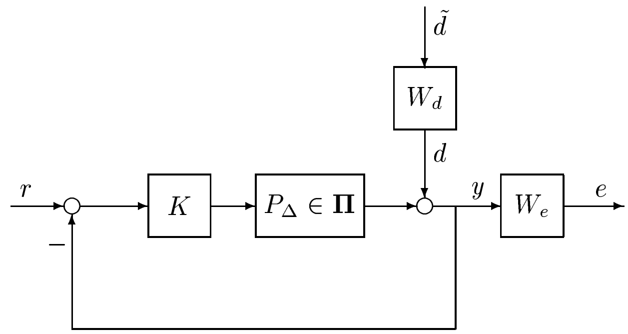


Figure 9.8: Diagram for Robust Performance Analysis

respect to sinusoidal disturbances. The former design criterion leads to the so-called $\mathcal{L}_1$-optimal control framework and the latter leads to $\mathcal{H}_2$ and $\mathcal{H}_\infty$ design frameworks, respectively. In this section, we will focus primarily on the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance objectives with unstructured model uncertainty descriptions. The performance under structured uncertainty will be considered in Chapter 11.

9.4.1 Robust $\mathcal{H}_2$ Performance

Although nominal $\mathcal{H}_2$ performance analysis is straightforward and involves only the computation of $\mathcal{H}_2$ norms, the $\mathcal{H}_2$ performance analysis with $\mathcal{H}_\infty$ norm bounded model uncertainty is much harder and little studied. Nevertheless, this problem is sensible and important for the reason that performance is sometimes more appropriately specified in terms of the $\mathcal{H}_2$ norm than the $\mathcal{H}_\infty$ norm. On the other hand, model uncertainties are more conveniently described using the $\mathcal{H}_\infty$ norm bounded sets and arise naturally in identification processes.

Let $T_{e\tilde{d}}^{\Delta}$ denote the transfer matrix between $\tilde{d}$ and e, then

$$T_{e\tilde{d}}^{\Delta} = W_e(I + P_\Delta K)^{-1} W_d, \quad P_\Delta \in \mathbf{\Pi}. \tag{9.4}$$

And the robust $\mathcal{H}_2$ performance analysis problem can be stated as finding

$$\sup_{P_\Delta \in \mathbf{\Pi}} \left\| T_{e\tilde{d}}^{\Delta} \right\|_2 .$$

To simplify our analysis, consider a scalar system with $W_d = 1$, $W_e = w_s$, $P = p$, and assume that the model is given by a multiplicative uncertainty set

$$\mathbf{\Pi} = \{(1 + w_t\delta)p : \quad \delta \in \mathcal{RH}_\infty, \ \|\delta\|_\infty < 1\}.$$

Assume further that the system is robustly stabilized by a controller k. Then

$$\sup_{P_\Delta \in \mathbf{\Pi}} \left\| T_{e\tilde{d}}^{\Delta} \right\|_2 = \sup_{\|\delta\|_\infty < 1} \left\| \frac{w_s \varepsilon}{1 + \tau w_t \delta} \right\|_2 = \left\| \frac{w_s \varepsilon}{1 - |\tau w_t|} \right\|_2$$

where $\varepsilon = (1+pk)^{-1}$ and $\tau = 1 - \varepsilon$.

The exact analysis for the matrix case is harder to determine although some upper bounds can be derived as we shall do for the $\mathcal{H}_\infty$ case below. However, the upper bounds do not seem to be insightful in the $\mathcal{H}_2$ setting, and, therefore, are omitted. It should be pointed out that synthesis for robust $\mathcal{H}_2$ performance is much harder even in the scalar case although synthesis for nominal performance is relatively easy and will be considered in Chapter 14.

9.4.2 Robust $\mathcal{H}_\infty$ Performance with Output Multiplicative Uncertainty

Suppose the performance criterion is to keep the energy of error e as small as possible, i.e.,

$$\sup_{\|\tilde{d}\|_2 \leq 1} \|e\|_2 \leq \epsilon$$

for some small ϵ. By scaling the error e (i.e., by properly selecting W_e) we can assume without loss of generality that $\epsilon = 1$. Then the robust performance criterion in this case can be described as requiring that the closed-loop system be robustly stable and that

$$\left\|T^{\Delta}_{e\tilde{d}}\right\|_\infty \leq 1, \quad \forall P_\Delta \in \mathbf{\Pi}. \tag{9.5}$$

More specifically, an output multiplicatively perturbed system will be analyzed first. The analysis for other classes of models can be done analogously. The perturbed model can be described as

$$\mathbf{\Pi} := \{(I + W_1 \Delta W_2)P : \quad \Delta \in \mathcal{RH}_\infty, \ \|\Delta\|_\infty < 1\} \tag{9.6}$$

with $W_1, W_2 \in \mathcal{RH}_\infty$. The explicit system diagram is as shown in Figure 9.5. For this class of models, we have

$$T^{\Delta}_{e\tilde{d}} = W_e S_o (I + W_1 \Delta W_2 T_o)^{-1} W_d,$$

and the robust performance is satisfied iff

$$\|W_2 T_o W_1\|_\infty \leq 1$$

and

$$\left\|T^{\Delta}_{e\tilde{d}}\right\|_\infty \leq 1, \ \forall \Delta \in \mathcal{RH}_\infty, \ \|\Delta\|_\infty < 1.$$

The exact analysis for this robust performance problem is not trivial and will be given in Chapter 11. However, some sufficient conditions are relatively easy to obtain by bounding these two inequalities, and they may shed some light on the nature of these problems. It will be assumed throughout that the controller K internally stabilizes the nominal plant P.

Theorem 9.7 *Suppose* $P_\Delta \in \{(I + W_1\Delta W_2)P : \ \Delta \in \mathcal{RH}_\infty, \ \|\Delta\|_\infty < 1\}$ *and* K *internally stabilizes* P*. Then the system robust performance is guaranteed if either one of the following conditions is satisfied*

(i) for each frequency ω

$$\overline{\sigma}(W_d)\overline{\sigma}(W_e S_o) + \overline{\sigma}(W_1)\overline{\sigma}(W_2 T_o) \leq 1; \tag{9.7}$$

(ii) for each frequency ω

$$\kappa(W_1^{-1}W_d)\overline{\sigma}(W_e S_o W_d) + \overline{\sigma}(W_2 T_o W_1) \leq 1 \tag{9.8}$$

where W_1 *and* W_d *are assumed to be invertible and* $\kappa(W_1^{-1}W_d)$ *is the condition number.*

Proof. It is obvious that both condition (9.7) and condition (9.8) guarantee that $\|W_2 T_o W_1\|_\infty \leq 1$. So it is sufficient to show that $\left\|T_{e\tilde{d}}^\Delta\right\|_\infty \leq 1, \forall\Delta \in \mathcal{RH}_\infty, \|\Delta\|_\infty < 1$. Now for any frequency ω, it is easy to see that

$$\begin{aligned}
\overline{\sigma}(T_{e\tilde{d}}^\Delta) &\leq \overline{\sigma}(W_e S_o)\overline{\sigma}[(I + W_1\Delta W_2 T_o)^{-1}]\overline{\sigma}(W_d) \\
&= \frac{\overline{\sigma}(W_e S_o)\overline{\sigma}(W_d)}{\underline{\sigma}(I + W_1\Delta W_2 T_o)} \\
&\leq \frac{\overline{\sigma}(W_e S_o)\overline{\sigma}(W_d)}{1 - \overline{\sigma}(W_1\Delta W_2 T_o)} \\
&\leq \frac{\overline{\sigma}(W_e S_o)\overline{\sigma}(W_d)}{1 - \overline{\sigma}(W_1)\overline{\sigma}(W_2 T_o)\overline{\sigma}(\Delta)}.
\end{aligned}$$

Hence condition (9.7) guarantees $\overline{\sigma}(T_{e\tilde{d}}^\Delta) \leq 1$ for all $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty < 1$ at all frequencies.

Similarly, suppose W_1 and W_d are invertible; write

$$T_{e\tilde{d}}^\Delta = W_e S_o W_d (W_1^{-1}W_d)^{-1}(I + \Delta W_2 T_o W_1)^{-1}(W_1^{-1}W_d),$$

and then

$$\overline{\sigma}(T_{e\tilde{d}}^\Delta) \leq \frac{\overline{\sigma}(W_e S_o W_d)\kappa(W_1^{-1}W_d)}{1 - \overline{\sigma}(W_2 T_o W_1)\overline{\sigma}(\Delta)}.$$

Hence by condition (9.8), $\overline{\sigma}(T_{e\tilde{d}}^\Delta) \leq 1$ is guaranteed for all $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty < 1$ at all frequencies. □

Remark 9.2 It is not hard to show that either one of the conditions in the theorem is also necessary for scalar valued systems.

Remark 9.3 Suppose $\kappa(W_1^{-1}W_d) \approx 1$ (weighting matrices satisfying this condition are usually called round weights). This is particularly the case if $W_1 = w_1(s)I$ and $W_d = w_d(s)I$. Recall that $\overline{\sigma}(W_e S_o W_d) \leq 1$ is the necessary and sufficient condition for nominal performance and that $\overline{\sigma}(W_2 T_o W_1) \leq 1$ is the necessary and sufficient condition for robust stability. Hence it is clear that in this case nominal performance plus robust stability almost guarantees robust performance, i.e., NP + RS $\approx$ RP. However, this is not necessarily true for other types of uncertainties as will be shown later. ♡

Remark 9.4 Note that in light of the equivalence relation between the robust stability and nominal performance, it is reasonable to conjecture that the above robust performance problem is equivalent to the robust stability problem in Figure 9.4 with the uncertainty model set given by

$$\mathbf{\Pi} := (I + W_d \Delta_e W_e)^{-1}(I + W_1 \Delta W_2)P$$

and $\|\Delta_e\|_\infty < 1$, $\|\Delta\|_\infty < 1$, as shown in Figure 9.9. This conjecture is indeed true; however, the equivalent model uncertainty is structured, and the exact stability analysis for such systems is not trivial and will be studied in Chapter 11. ♡

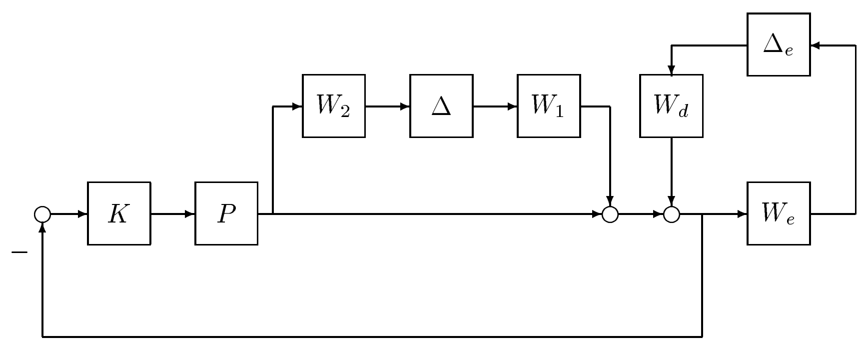


Figure 9.9: Robust Performance with Unstructured Uncertainty vs. Robust Stability with Structured Uncertainty

Remark 9.5 Note that if W_1 and W_d are invertible, then $T_{e\tilde{d}}^{\Delta}$ can also be written as

$$T_{e\tilde{d}}^{\Delta} = W_e S_o W_d \left[I + (W_1^{-1}W_d)^{-1}\Delta W_2 T_o W_1 (W_1^{-1}W_d)\right]^{-1}.$$

So another alternative sufficient condition for robust performance can be obtained as

$$\overline{\sigma}(W_e S_o W_d) + \kappa(W_1^{-1}W_d)\overline{\sigma}(W_2 T_o W_1) \leq 1.$$

A similar situation also occurs in the skewed case below. We will not repeat all these variations. ♡

9.4.3 Skewed Specifications and Plant Condition Number

We now consider the system with skewed specifications, i.e., the uncertainty and performance are not measured at the same location. For instance, the system performance is still measured in terms of output sensitivity, but the uncertainty model is in input multiplicative form:

$$\mathbf{\Pi} := \{P(I + W_1 \Delta W_2) : \ \Delta \in \mathcal{RH}_\infty, \ \|\Delta\|_\infty < 1\}.$$

For systems described by this class of models, see Figure 9.8, the robust stability condition becomes

$$\|W_2 T_i W_1\|_\infty \leq 1,$$

and the nominal performance condition becomes

$$\|W_e S_o W_d\|_\infty \leq 1.$$

To consider the robust performance, let $\tilde{T}^{\Delta}_{e\tilde{d}}$ denote the transfer matrix from $\tilde{d}$ to e. Then

$$\begin{aligned} \tilde{T}^{\Delta}_{e\tilde{d}} &= W_e S_o (I + P W_1 \Delta W_2 K S_o)^{-1} W_d \\ &= W_e S_o W_d \left[I + (W_d^{-1} P W_1) \Delta (W_2 T_i W_1)(W_d^{-1} P W_1)^{-1} \right]^{-1}. \end{aligned}$$

The last equality follows if W_1, W_d, and P are invertible and, if W_2 is invertible, can also be written as

$$\tilde{T}^{\Delta}_{e\tilde{d}} = W_e S_o W_d (W_1^{-1} W_d)^{-1} \left[I + (W_1^{-1} P W_1) \Delta (W_2 P^{-1} W_2^{-1})(W_2 T_o W_1) \right]^{-1} (W_1^{-1} W_d).$$

Then the following results follow easily.

Theorem 9.8 *Suppose $P_\Delta \in \mathbf{\Pi} = \{P(I + W_1 \Delta W_2) : \ \Delta \in \mathcal{RH}_\infty, \|\Delta\|_\infty < 1\}$ and K internally stabilizes P. Assume that P, W_1, W_2, and W_d are square and invertible. Then the system robust performance is guaranteed if either one of the following conditions is satisfied*

(i) for each frequency ω

$$\overline{\sigma}(W_e S_o W_d) + \kappa(W_d^{-1} P W_1) \overline{\sigma}(W_2 T_i W_1) \leq 1; \tag{9.9}$$

(ii) for each frequency ω

$$\kappa(W_1^{-1} W_d) \overline{\sigma}(W_e S_o W_d) + \overline{\sigma}(W_1^{-1} P W_1) \overline{\sigma}(W_2 P^{-1} W_2^{-1}) \overline{\sigma}(W_2 T_o W_1) \leq 1. \tag{9.10}$$

Remark 9.6 If the appropriate invertibility conditions are not satisfied, then an alternative sufficient condition for robust performance can be given by

$$\overline{\sigma}(W_d) \overline{\sigma}(W_e S_o) + \overline{\sigma}(P W_1) \overline{\sigma}(W_2 K S_o) \leq 1.$$

Similar to the previous case, there are many different variations of sufficient conditions although (9.10) may be the most useful one. ♡

Remark 9.7 It is important to note that in this case, the robust stability condition is given in terms of $L_i = KP$ while the nominal performance condition is given in terms of $L_o = PK$. These classes of problems are called *skewed problems* or problems with skewed specifications.[2] Since, in general, $PK \neq KP$, the robust stability margin or tolerances for uncertainties at the plant input and output are generally not the same. ♡

Remark 9.8 It is also noted that the robust performance condition is related to the condition number of the weighted nominal model. So in general if the weighted nominal model is ill-conditioned at the range of critical frequencies, then the robust performance condition may be far more restrictive than the robust stability condition and the nominal performance condition together. For simplicity, assume $W_1 = I$, $W_d = I$ and $W_2 = w_t I$ where $w_t \in \mathcal{RH}_\infty$ is a scalar function. Further, P is assumed to be invertible. Then robust performance condition (9.10) can be written as

$$\overline{\sigma}(W_e S_o) + \kappa(P)\overline{\sigma}(w_t T_o) \leq 1, \forall \omega.$$

Comparing these conditions with those obtained for non-skewed problems shows that the condition related to robust stability is scaled by the condition number of the plant[3]. Since $\kappa(P) \geq 1$, it is clear that the skewed specifications are much harder to satisfy if the plant is not well conditioned. This problem will be discussed in more detail in section 11.3.3 of Chapter 11. ♡

Remark 9.9 Suppose K is invertible, then $\tilde{T}^{\Delta}_{e\tilde{d}}$ can be written as

$$\tilde{T}^{\Delta}_{e\tilde{d}} = W_e K^{-1}(I + T_i W_1 \Delta W_2)^{-1} S_i K W_d.$$

Assume further that $W_e = I, W_d = w_s I, W_2 = I$ where $w_s \in \mathcal{RH}_\infty$ is a scalar function. Then a sufficient condition for robust performance is given by

$$\kappa(K)\overline{\sigma}(S_i w_s) + \overline{\sigma}(T_i W_1) \leq 1, \forall \omega,$$

with $\kappa(K) := \overline{\sigma}(K)\overline{\sigma}(K^{-1})$. This is equivalent to treating the input multiplicative plant uncertainty as the output multiplicative controller uncertainty. ♡

These skewed specifications also create problems for MIMO loop shaping design which has been discussed briefly in Chapter 5. The idea of loop shaping design is based on the fact that robust performance is guaranteed by designing a controller with a sufficient nominal performance margin and a sufficient robust stability margin. For example, if $\kappa(W_1^{-1} W_d) \approx 1$, the output multiplicative perturbed robust performance is guaranteed by designing a controller with twice the required nominal performance and robust stability margins.

[2] See Stein and Doyle [1991].

[3] Alternative condition can be derived so that the condition related to nominal performance is scaled by the condition number.

The fact that the condition number appeared in the robust performance test for skewed problems can be given another interpretation by considering two sets of plants $\mathbf{\Pi}_1$ and $\mathbf{\Pi}_2$ as shown below.

$$\begin{aligned}
\mathbf{\Pi}_1 &:= \{P(I + w_t\Delta) : \ \Delta \in \mathcal{RH}_\infty, \ \|\Delta\|_\infty < 1\} \\
\mathbf{\Pi}_2 &:= \{(I + \tilde{w}_t\Delta)P : \ \Delta \in \mathcal{RH}_\infty, \ \|\Delta\|_\infty < 1\}.
\end{aligned}$$

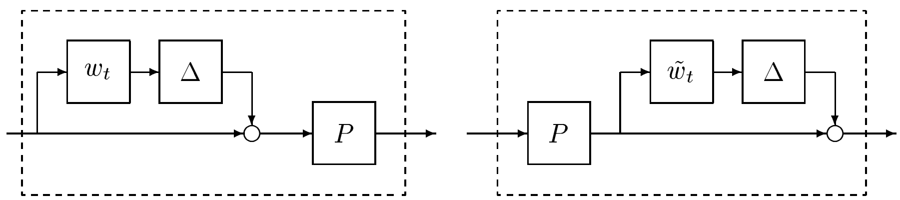


Figure 9.10: Converting Input Uncertainty to Output Uncertainty

Assume that P is invertible, then

$$\mathbf{\Pi}_2 \supseteq \mathbf{\Pi}_1 \quad \text{if} \quad |\tilde{w}_t| \geq |w_t|\kappa(P) \quad \forall\omega$$

since $P(I + w_t\Delta) = (I + w_tP\Delta P^{-1})P$.

The condition number of a transfer matrix can be very high at high frequency which may significantly limit the achievable performance. The example below, taken from the textbook [Franklin, Powell, and Workman, 1990], shows that the condition number shown in Figure 9.11 may increase with the frequency:

$$P(s) = \left[\begin{array}{ccc|cc} -0.2 & 0.1 & 1 & 0 & 1 \\ -0.05 & 0 & 0 & 0 & 0.7 \\ 0 & 0 & -1 & 1 & 0 \\ \hline 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{array}\right] = \frac{1}{a(s)} \begin{bmatrix} s & (s+1)(s+0.07) \\ -0.05 & 0.7(s+1)(s+0.13) \end{bmatrix}$$

where $a(s) = (s+1)(s+0.1707)(s+0.02929)$.

9.5 Gain Margin and Phase Margin

In this section, we show that the gain margin and phase margin defined in classical control theory may not be sufficient indicators of a system's robustness. Let $L(s)$ be a scalar transfer function and consider a unit feedback system such as the one shown in the following diagram:

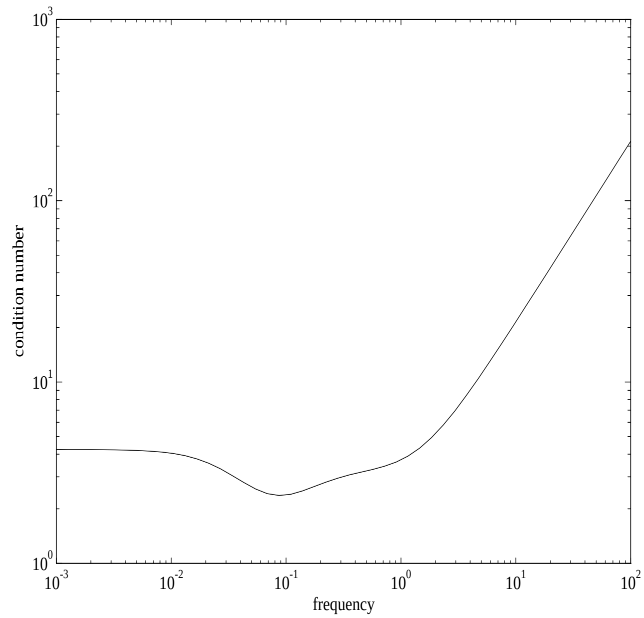
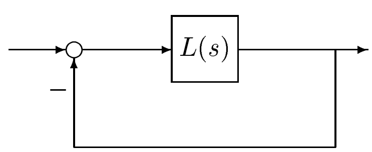


Figure 9.11: Condition Number $\kappa(\omega) = \bar{\sigma}(P(j\omega))/\underline{\sigma}(P(j\omega))$



Suppose that the above closed-loop feedback system with $L(s) = L_0(s)$ is stable. Then the system is said to have

Gain Margin $k_{\min}$ *and* $k_{\max}$ if the closed-loop system is stable for all $L(s) = kL_0(s)$ with $k_{\min} < k < k_{\max}$ but unstable for $L(s) = k_{max}L_0(s)$ and for $L(s) = k_{min}L_0(s)$ where $0 \le k_{\min} \le 1$ and $k_{\max} \ge 1$.

Phase Margin $\phi_{\min}$ *and* $\phi_{\max}$ if the closed-loop system is stable for all $L(s) = e^{-j\phi}L_0(s)$ with $\phi_{\min} < \phi < \phi_{\max}$ but unstable for $L(s) = e^{-j\phi_{\max}}L_0(s)$ and for $L(s) = e^{-j\phi_{\min}}L_0(s)$ where $-\pi \le \phi_{\min} \le 0$ and $0 \le \phi_{\max} \le \pi$.

These margins can be easily read from the open-loop system Nyquist diagram as shown in Figure 9.12 where $k_{\max}$ and $k_{\min}$ represent how much the loop gain can be increased and decreased, respectively, without causing instability. Similarly $\phi_{\max}$ and $\phi_{\min}$ represent how much loop phase lag and lead can be tolerated without causing instability.

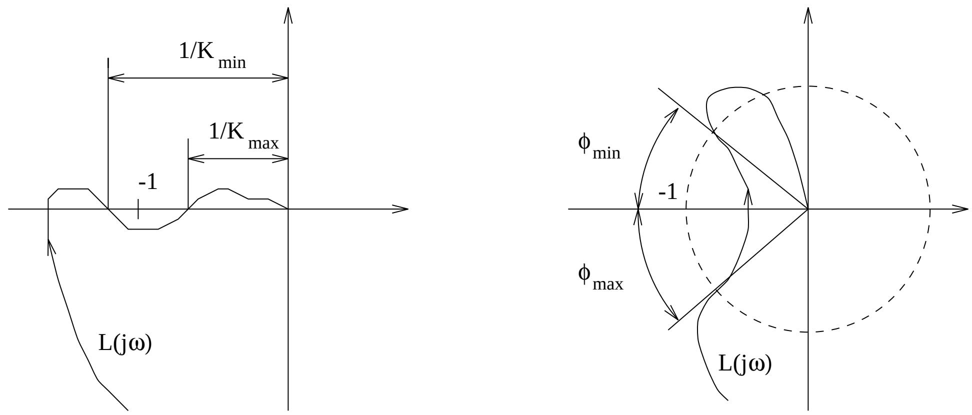


Figure 9.12: Gain Margin and Phase Margin of A Scalar System

However, gain margin or phase margin alone may not be a sufficient indicator of a system's robustness. To be more specific, consider a simple dynamical system

$$P = \frac{a-s}{as-1}, \quad a > 1$$

with a stabilizing controller K. Now let $L = PK$ and consider a controller

$$K = \frac{b+s}{bs+1}, \quad b > 0.$$

It is easy to show that the closed-loop system is stable for any

$$\frac{1}{a} < b < a.$$

To compute the stability margins, consider three cases:

(i) $b = 1$: in this case, $K = 1$ and the stability margins can be easily shown to be

$$k_{\min} = \frac{1}{a}, \quad k_{\max} = a, \quad \phi_{\min} = -\pi, \quad \phi_{\max} = \sin^{-1}\left(\frac{a^2-1}{a^2+1}\right) =: \theta.$$

It is easy to see that both gain margin and phase margin are very large for large a.

(ii) $\frac{1}{a} < b < a$ and $b \to a$: in this case

$$k_{\min} = \frac{1}{ab} \to \frac{1}{a^2}, \quad k_{\max} = ab \to a^2, \quad \phi_{\min} = -\pi, \quad \phi_{\max} \to 0$$

i.e., very large gain margin but arbitrarily small phase margin.

(iii) $\frac{1}{a} < b < a$ and $b \to \frac{1}{a}$: in this case

$$k_{\min} = \frac{1}{ab} \to 1, \quad k_{\max} = ab \to 1, \quad \phi_{\min} = -\pi, \quad \phi_{\max} \to 2\theta$$

i.e., very large phase margin but arbitrarily small gain margin.

The open-loop frequency response of the system is shown in Figure 9.13 for $a = 2$ and $b = 1, b = 1.9$ and $b = 0.55$, respectively.

Sometimes, gain margin and phase margin together may still not be enough to indicate the true robustness of a system. For example, it is possible to construct a (complicated) controller such that

$$k_{\min} < \frac{1}{a}, \quad k_{\max} > a, \quad \phi_{\min} = -\pi, \quad \phi_{\max} > \theta$$

but the Nyquist plot is arbitrarily close to $(-1, 0)$. Such a controller is complicated to construct; however, the following controller should give the reader a good idea of its construction:

$$K_{\text{bad}} = \frac{s+3.3}{3.3s+1} \, \frac{s+0.55}{0.55s+1} \, \frac{1.7s^2+1.5s+1}{s^2+1.5s+1.7}.$$

The open-loop frequency response of the system with this controller is also shown in Figure 9.13 by the dotted line. It is easy to see that the system has at least the same gain margin and phase margin as the system with controller $K = 1$, but the Nyquist plot is closer to $(-1, 0)$. Therefore this system is less robust with respect to the simultaneous gain and phase perturbations. The problem is that the gain margin and phase margin do not give the correct indication of the system's robustness when the gain and phase are varying simultaneously.

9.6 Deficiency of Classical Control for MIMO Systems

In this section, we show through an example that the classical control theory may not be reliable when it is applied to MIMO system design.

Consider a symmetric spinning body with torque inputs, T_1 and T_2, along two orthogonal transverse axes, x and y, as shown in Figure 9.14. Assume that the angular velocity of the spinning body with respect to the z axis is constant, Ω. Assume further that the inertia of the spinning body with respect to the x, y, and z axes are I_1, $I_2 = I_1$, and I_3, respectively. Denote by ω_1 and ω_2 the angular velocities of the body with respect to the x and y axes, respectively. Then the Euler's equation of the spinning body is given by

$$\begin{aligned} I_1\dot{\omega}_1 - \omega_2\Omega(I_1 - I_3) &= T_1 \\ I_1\dot{\omega}_2 - \omega_1\Omega(I_3 - I_1) &= T_2. \end{aligned}$$

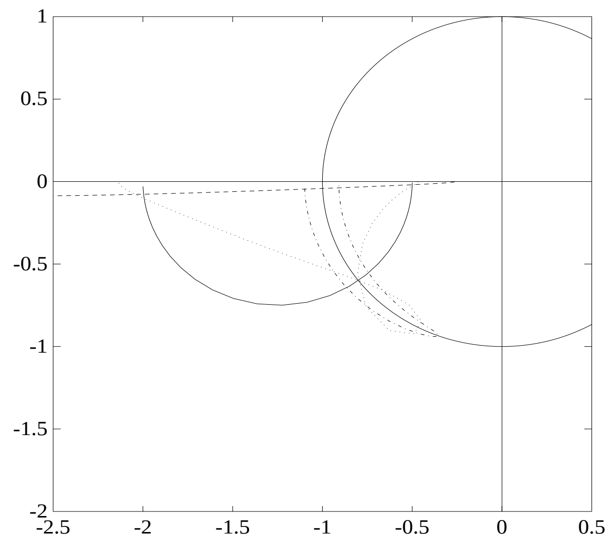


Figure 9.13: Nyquist Plots of L with $a = 2$ and $b = 1$(solid), $b = 1.9$(dashed), $b = 0.55$(dashdot) and with $K_{\rm bad}$ (dotted)

Define

$$\begin{bmatrix} u_1 \\ u_2 \end{bmatrix} := \begin{bmatrix} T_1/I_1 \\ T_2/I_1 \end{bmatrix}, \ a := (1 - I_3/I_1)\Omega.$$

Then the system dynamical equations can be written as

$$\begin{bmatrix} \dot\omega_1 \\ \dot\omega_2 \end{bmatrix} = \begin{bmatrix} 0 & a \\ -a & 0 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}.$$

Now suppose that the angular rates ω_1 and ω_2 are measured in scaled and rotated coordinates:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \frac{1}{\cos\theta} \begin{bmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \begin{bmatrix} 1 & a \\ -a & 1 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix}$$

where $\tan\theta := a$. (There is no specific physical meaning for the measurements of y_1 and y_1 but they are assumed here only for the convenience of discussion.) Then the transfer matrix for the spinning body can be computed as

$$Y(s) = P(s)U(s)$$

with

$$P(s) = \frac{1}{s^2 + a^2} \begin{bmatrix} s - a^2 & a(s+1) \\ -a(s+1) & s - a^2 \end{bmatrix}.$$

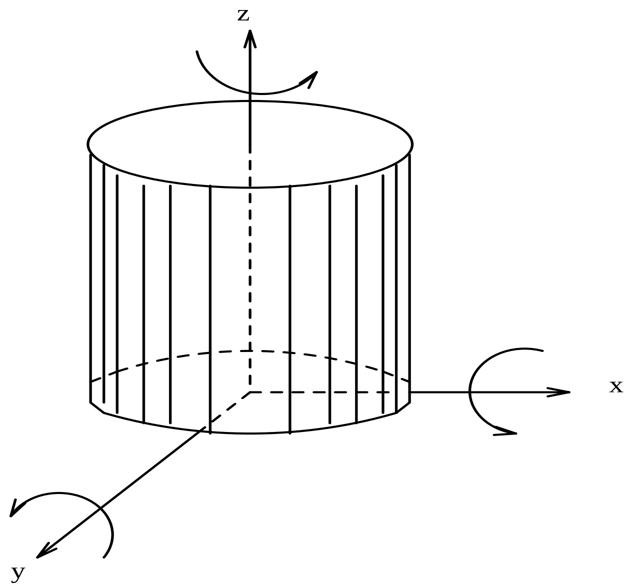


Figure 9.14: Spinning Body

Suppose the control law is chosen as

$$u = K_1 r - y$$

where

$$K_1 = \frac{1}{1+a^2}\begin{bmatrix} 1 & -a \\ a & 1 \end{bmatrix}.$$

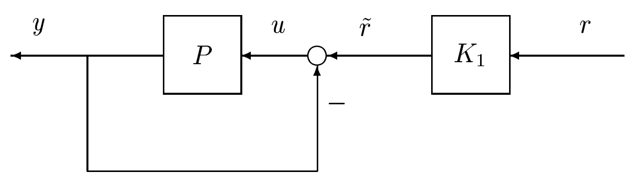


Figure 9.15: Closed-loop with a "Bizarre" Controller

Then the closed-loop transfer function is given by

$$Y(s) = \frac{1}{s+1}\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} R(s)$$

and the sensitivity function and the complementary sensitivity function are given by

$$S = (I+P)^{-1} = \frac{1}{s+1}\begin{bmatrix} s & -a \\ a & s \end{bmatrix}, \quad T = P(I+P)^{-1} = \frac{1}{s+1}\begin{bmatrix} 1 & a \\ -a & 1 \end{bmatrix}.$$

It is noted that this controller design has the property of decoupling the loops. Furthermore, each single loop has the open-loop transfer function as

$$\frac{1}{s}$$

so each loop has phase margin $\phi_{\max} = -\phi_{\min} = 90^o$ and gain margin $k_{\min} = 0, k_{\max} = \infty$.

Suppose one loop transfer function is perturbed, as shown in Figure 9.16.

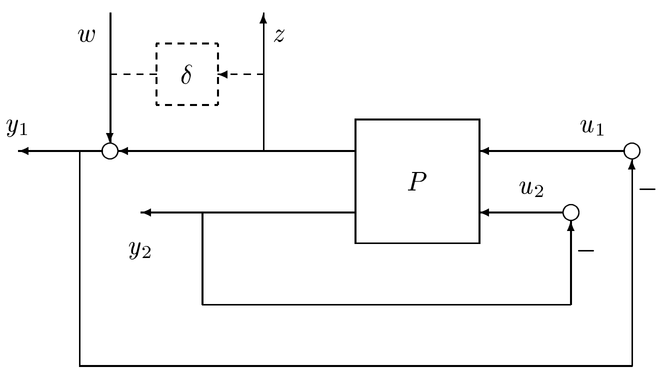


Figure 9.16: One-Loop-At-A-Time Analysis

Denote

$$\frac{z(s)}{w(s)} = -T_{11} = -\frac{1}{s+1}.$$

Then the maximum allowable perturbation is given by

$$\|\delta\|_\infty < \frac{1}{\|T_{11}\|_\infty} = 1$$

which is independent of a. Similarly the maximum allowable perturbation on the other loop is also 1 by symmetry. However, if both loops are perturbed at the same time, then the maximum allowable perturbation is much smaller, as shown below.

Consider a multivariable perturbation, as shown in Figure 9.17, i.e., $P_\Delta = (I+\Delta)P$, with

$$\Delta = \begin{bmatrix} \delta_{11} & \delta_{12} \\ \delta_{21} & \delta_{22} \end{bmatrix} \in \mathcal{RH}_\infty$$

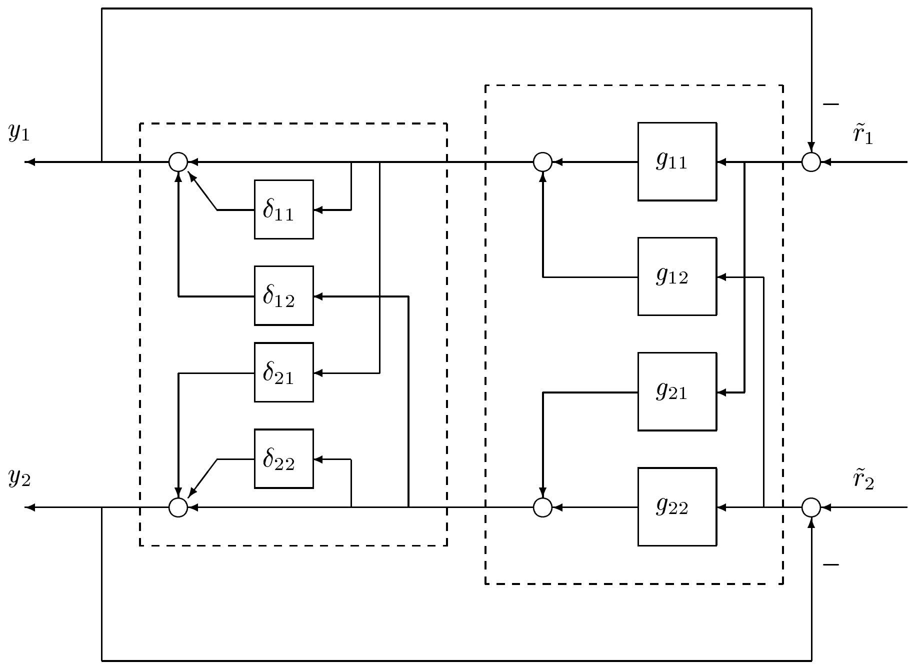


Figure 9.17: Simultaneous Perturbations

a 2×2 transfer matrix such that $\|\Delta\|_\infty < \gamma$. Then by the small gain theorem, the system is robustly stable for every such Δ iff

$$\gamma \leq \frac{1}{\|T\|_\infty} = \frac{1}{\sqrt{1+a^2}} \quad (\ll 1 \text{ if } a \gg 1).$$

In particular, consider

$$\Delta = \Delta_d = \begin{bmatrix} \delta_{11} & \\ & \delta_{22} \end{bmatrix} \in \mathbb{R}^{2\times 2}.$$

Then the closed-loop system is stable for every such Δ iff

$$\det(I + T\Delta_d) = \frac{1}{(s+1)^2}\left(s^2 + (2 + \delta_{11} + \delta_{22})s + 1 + \delta_{11} + \delta_{22} + (1+a^2)\delta_{11}\delta_{22}\right)$$

has no zero in the closed right-half plane. Hence the stability region is given by

$$\begin{aligned} 2 + \delta_{11} + \delta_{22} &> 0 \\ 1 + \delta_{11} + \delta_{22} + (1+a^2)\delta_{11}\delta_{22} &> 0. \end{aligned}$$

It is easy to see that the system is unstable with

$$\delta_{11} = -\delta_{22} = \frac{1}{\sqrt{1+a^2}}.$$

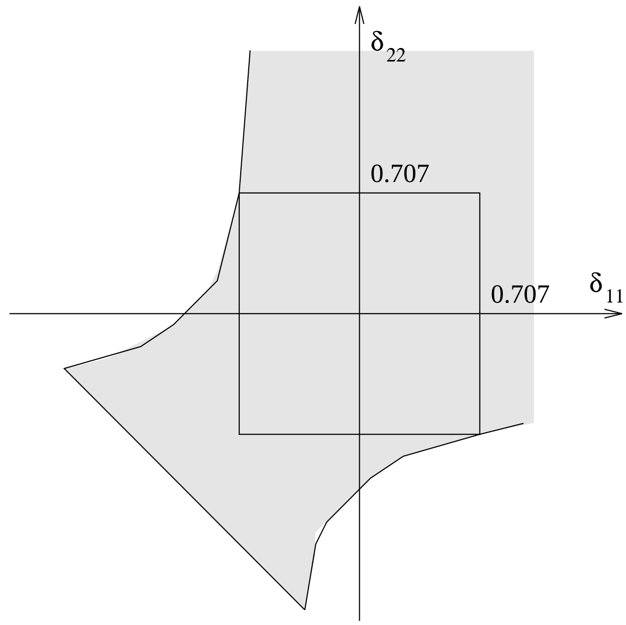


Figure 9.18: Stability Region for $a = 1$

The stability region is drawn in the Figure 9.18. This clearly shows that the analysis of an MIMO system using SISO methods can be misleading and can even give erroneous results. Hence an MIMO method has to be used.

9.7 Notes and References

The small gain theorem was first presented by Zames [1966]. The book by Desoer and Vidyasagar [1975] contains a quite extensive treatment and applications of this theorem in various forms. Robust stability conditions under various uncertainty assumptions are discussed in Doyle, Wall, and Stein [1982]. It was first shown in Kishore and Pearson [1992] that the small gain condition may not be necessary for robust stability with closed-ball perturbed uncertainties. In the same reference, some extensions of stability and performance criteria under various structured/unstructured uncertainties are given. Some further extensions are also presented in Tits and Fan [1994].

10

Linear Fractional Transformation

This chapter introduces a new matrix function: linear fractional transformation (LFT). We show that many interesting control problems can be formulated in an LFT framework and thus can be treated using the same technique.

10.1 Linear Fractional Transformations

This section introduces the matrix linear fractional transformations. It is well known from the one complex variable function theory that a mapping $F : \mathbb{C} \mapsto \mathbb{C}$ of the form

$$F(s) = \frac{a + bs}{c + ds}$$

with a, b, c and $d \in \mathbb{C}$ is called a *linear fractional transformation*. In particular, if $c \neq 0$ then $F(s)$ can also be written as

$$F(s) = \alpha + \beta s(1 - \gamma s)^{-1}$$

for some α, β and $\gamma \in \mathbb{C}$. The linear fractional transformation described above for scalars can be generalized to the matrix case.

Definition 10.1 Let M be a complex matrix partitioned as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \in \mathbb{C}^{(p_1+p_2)\times(q_1+q_2)},$$

and let $\Delta_\ell \in \mathbb{C}^{q_2 \times p_2}$ and $\Delta_u \in \mathbb{C}^{q_1 \times p_1}$ be two other complex matrices. Then we can formally define a *lower LFT* with respect to Δ_ℓ as the map

$$\mathcal{F}_\ell(M, \bullet): \ \mathbb{C}^{q_2 \times p_2} \mapsto \mathbb{C}^{p_1 \times q_1}$$

with

$$\mathcal{F}_\ell(M, \Delta_\ell) \triangleq M_{11} + M_{12}\Delta_\ell(I - M_{22}\Delta_\ell)^{-1}M_{21}$$

provided that the inverse $(I - M_{22}\Delta_\ell)^{-1}$ exists. We can also define an *upper LFT* with respect to Δ_u as

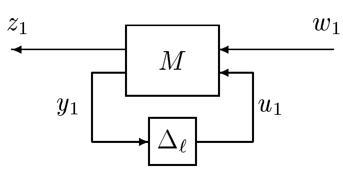
$$\mathcal{F}_u(M, \bullet): \ \mathbb{C}^{q_1 \times p_1} \mapsto \mathbb{C}^{p_2 \times q_2}$$

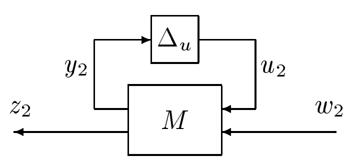
with

$$\mathcal{F}_u(M, \Delta_u) = M_{22} + M_{21}\Delta_u(I - M_{11}\Delta_u)^{-1}M_{12}$$

provided that the inverse $(I - M_{11}\Delta_u)^{-1}$ exists.

The matrix M in the above LFTs is called the *coefficient matrix*. The motivation for the terminologies of *lower* and *upper* LFTs should be clear from the following diagram representations of $\mathcal{F}_\ell(M, \Delta_\ell)$ and $\mathcal{F}_u(M, \Delta_u)$:





The diagram on the left represents the following set of equations:

$$\begin{aligned} \begin{bmatrix} z_1 \\ y_1 \end{bmatrix} &= M \begin{bmatrix} w_1 \\ u_1 \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ u_1 \end{bmatrix}, \\ u_1 &= \Delta_\ell \, y_1 \end{aligned}$$

while the diagram on the right represents

$$\begin{aligned} \begin{bmatrix} y_2 \\ z_2 \end{bmatrix} &= M \begin{bmatrix} u_2 \\ w_2 \end{bmatrix} = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \begin{bmatrix} u_2 \\ w_2 \end{bmatrix}, \\ u &= \Delta_u \, y_2. \end{aligned}$$

It is easy to verify that the mapping defined on the left diagram is equal to $\mathcal{F}_\ell(M, \Delta_\ell)$ and the mapping defined on the right diagram is equal to $\mathcal{F}_u(M, \Delta_u)$. So from the above diagrams, $\mathcal{F}_\ell(M, \Delta_\ell)$ is a transformation obtained from closing the *lower* loop on the left diagram; similarly $\mathcal{F}_u(M, \Delta_u)$ is a transformation obtained from closing the *upper* loop on the right diagram. In most cases, we will use the general term *LFT* in referring to both upper and lower LFTs and assume that the contents will distinguish the situations. The reason for this is that one can use either of these notations to express a given object. Indeed, it is clear that $\mathcal{F}_u(N,\Delta) = \mathcal{F}_\ell(M,\Delta)$ with $N = \begin{bmatrix} M_{22} & M_{21} \\ M_{12} & M_{11} \end{bmatrix}$. It is usually not crucial which expression is used; however, it is often the case that one expression is more convenient than the other for a given problem. It should also be clear to the reader that in writing $\mathcal{F}_\ell(M,\Delta)$ (or $\mathcal{F}_u(M,\Delta)$) it is implied that Δ has compatible dimensions.

A useful interpretation of an LFT, e.g., $\mathcal{F}_\ell(M,\Delta)$, is that $\mathcal{F}_\ell(M,\Delta)$ has a nominal mapping, M_{11}, and is perturbed by Δ, while M_{12}, M_{21}, and M_{22} reflect a prior knowledge as to how the perturbation affects the nominal map, M_{11}. A similar interpretation can be applied to $\mathcal{F}_u(M,\Delta)$. This is why LFT is particularly useful in the study of perturbations, which is the focus of the next chapter.

The physical meaning of an LFT in control science is obvious if we take M as a proper transfer matrix. In that case, the LFTs defined above are simply the closed-loop transfer matrices from $w_1 \mapsto z_1$ and $w_2 \mapsto z_2$, respectively, i.e.,

$$T_{zw1} = \mathcal{F}_\ell(M,\Delta_\ell), \qquad T_{zw2} = \mathcal{F}_u(M,\Delta_u)$$

where M may be the controlled plant and Δ may be either the system model uncertainties or the controllers.

Definition 10.2 An LFT, $\mathcal{F}_\ell(M,\Delta)$, is said to be *well defined (or well-posed)* if $(I - M_{22}\Delta)$ is invertible.

Note that this definition is consistent with the well-posedness definition of the feedback system, which requires the corresponding transfer matrix be invertible in $\mathcal{R}_p(s)$. It is clear that the study of an LFT that is not well-defined is meaningless, hence throughout the book, whenever an LFT is invoked, it will be assumed implicitly to be well defined. It is also clear from the definition that, for any M, $\mathcal{F}_\ell(M,0)$ is well defined; hence any function that is not well defined at the origin cannot be expressed as an LFT in terms of its variables. For example, $f(\delta) = 1/\delta$ is not an LFT of δ.

In some literature, LFT is used to refer to the following matrix functions:

$$(A+BQ)(C+DQ)^{-1} \qquad \textit{or} \qquad (C+QD)^{-1}(A+QB)$$

where C is usually assumed to be invertible due to practical consideration. The following is true.

Lemma 10.1 *Suppose C is invertible. Then*

$$\begin{aligned} (A+BQ)(C+DQ)^{-1} &= \mathcal{F}_\ell(M,Q) \\ (C+QD)^{-1}(A+QB) &= \mathcal{F}_\ell(N,Q) \end{aligned}$$

with

$$M = \begin{bmatrix} AC^{-1} & B - AC^{-1}D \\ C^{-1} & -C^{-1}D \end{bmatrix}, \quad N = \begin{bmatrix} C^{-1}A & C^{-1} \\ B - DC^{-1}A & -DC^{-1} \end{bmatrix}.$$

The converse also holds if M satisfies certain conditions.

Lemma 10.2 *Let $\mathcal{F}_\ell(M,Q)$ be a given LFT with $M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$, then*

(a) if M_{21} is invertible,

$$\mathcal{F}_\ell(M,Q) = (A+BQ)(C+DQ)^{-1}$$

with

$$A = M_{11}M_{21}^{-1}, \quad B = M_{12} - M_{11}M_{21}^{-1}M_{22}, \quad C = M_{21}^{-1}, \quad D = -M_{21}^{-1}M_{22}.$$

(b) if M_{12} is invertible,

$$\mathcal{F}_\ell(M,Q) = (C+QD)^{-1}(A+QB)$$

with

$$A = M_{12}^{-1}M_{11}, \quad B = M_{21} - M_{22}M_{12}^{-1}M_{11}, \quad C = M_{12}^{-1}, \quad D = -M_{22}M_{12}^{-1}.$$

However, for an arbitrary LFT $\mathcal{F}_\ell(M,Q)$, neither M_{21} nor M_{12} is necessarily square and invertible; therefore, the alternative fractional formula is more restrictive.

It should be pointed out that some seemingly similar functions do not have simple LFT representations. For example,

$$(A+QB)(I+QD)^{-1}$$

cannot always be written in the form of $\mathcal{F}_\ell(M,Q)$ for some M; however, it can be written as

$$(A+QB)(I+QD)^{-1} = \mathcal{F}_\ell(N,\Delta)$$

with

$$N = \left[\begin{array}{c:cc} A & I & A \\ \hdashline -B & 0 & -B \\ D & 0 & D \end{array}\right], \qquad \Delta = \begin{bmatrix} Q & \\ & Q \end{bmatrix}.$$

Note that the dimension of Δ is twice as many as Q.

The following lemma summarizes some of the algebraic properties of *LFT*s.

Lemma 10.3 *Let M, Q, and G be suitably partitioned matrices:*

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}, \quad Q = \begin{bmatrix} Q_{11} & Q_{12} \\ Q_{21} & Q_{22} \end{bmatrix}, \quad G = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{bmatrix}.$$

(i) $\mathcal{F}_u(M,\Delta) = \mathcal{F}_l(N,\Delta)$ *with*

$$N = \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} M \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} = \begin{bmatrix} M_{22} & M_{21} \\ M_{12} & M_{11} \end{bmatrix}$$

where the dimensions of identity matrices are compatible with the partitions of M *and* N.

(ii) Suppose $\mathcal{F}_u(M,\Delta)$ *is square and well-defined and* M_{22} *is nonsingular. Then the inverse of* $\mathcal{F}_u(M,\Delta)$ *exists and is also an LFT with respect to* Δ*:*

$$(\mathcal{F}_u(M,\Delta))^{-1} = \mathcal{F}_u(N,\Delta)$$

with N *given by*

$$N = \begin{bmatrix} M_{11} - M_{12}M_{22}^{-1}M_{21} & -M_{12}M_{22}^{-1} \\ M_{22}^{-1}M_{21} & M_{22}^{-1} \end{bmatrix}.$$

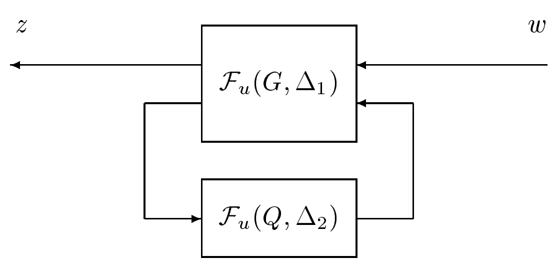
(iii) $\mathcal{F}_u(M,\Delta_1) + \mathcal{F}_u(Q,\Delta_2) = \mathcal{F}_u(N,\Delta)$ *with*

$$N = \left[\begin{array}{cc:c} M_{11} & 0 & M_{12} \\ 0 & Q_{11} & Q_{12} \\ \hdashline M_{21} & Q_{21} & M_{22}+Q_{22} \end{array}\right], \qquad \Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix}.$$

(iv) $\mathcal{F}_u(M,\Delta_1)\mathcal{F}_u(Q,\Delta_2) = \mathcal{F}_u(N,\Delta)$ *with*

$$N = \left[\begin{array}{cc:c} M_{11} & M_{12}Q_{21} & M_{12}Q_{22} \\ 0 & M_{11} & Q_{12} \\ \hdashline M_{21} & M_{22}Q_{21} & M_{22}Q_{22} \end{array}\right], \qquad \Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix}.$$

(v) Consider the following interconnection structure where the dimensions of Δ_1 *are compatible with* A*:*



Then the mapping from w *to* z *is given by*

$$\mathcal{F}_l\left(\mathcal{F}_u(G,\Delta_1), \mathcal{F}_u(Q,\Delta_2)\right) = \mathcal{F}_u\left(\mathcal{F}_l\left(G, \mathcal{F}_u(Q,\Delta_2)\right), \Delta_1\right) = \mathcal{F}_u(N,\Delta)$$

$$N = \left[\begin{array}{cc|c} A + B_2Q_{22}L_1C_2 & B_2L_2Q_{21} & B_1 + B_2Q_{22}L_1D_{21} \\ Q_{12}L_1C_2 & Q_{11} + Q_{12}L_1D_{22}Q_{21} & Q_{12}L_1D_{21} \\ \hline C_1 + D_{12}L_2Q_{22}C_2 & D_{12}L_2Q_{21} & D_{11} + D_{12}Q_{22}L_1D_{21} \end{array}\right]$$

where $L_1 := (I - D_{22}Q_{22})^{-1}$, $L_2 := (I - Q_{22}D_{22})^{-1}$, *and* $\Delta = \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix}$.

Proof. These properties can be straightforwardly verified by the definition of LFT, so the proofs are omitted. □

Property (v) shows that if the open-loop system parameters are LFTs of some variables, then the closed-loop system parameters are also LFTs of the same variables. This property is particularly useful in perturbation analysis and in building the interconnection structure. Similar results can be stated for lower LFT. It is usually convenient to interpret an LFT $\mathcal{F}_u(M, \Delta)$ as a state space realization of a generalized system with frequency structure Δ. In fact, all the above properties can be reduced to the standard transfer matrix operations if $\Delta = \frac{1}{s}I$.

The following proposition is concerned with the algebraic properties of LFTs in the general control setup.

Lemma 10.4 *Let* $P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}$ *and let K be rational transfer function matrices and let* $G = \mathcal{F}_\ell(P, K)$. *Then*

(a) *G is proper if P and K are proper with* $\det(I - P_{22}K)(\infty) \neq 0$.

(b) $\mathcal{F}_\ell(P, K_1) = \mathcal{F}_\ell(P, K_2)$ *implies that* $K_1 = K_2$ *if P_{12} and P_{21} have normal full column and row rank in* $\mathcal{R}_p(s)$, *respectively.*

(c) *If P and G are proper,* $\det P(\infty) \neq 0$, $\det\left(P - \begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix}\right)(\infty) \neq 0$ *and P_{12} and P_{21} are square and invertible for almost all s, then K is proper and*

$$K = \mathcal{F}_u(P^{-1}, G).$$

Proof.

(a) is immediate from the definition of $\mathcal{F}_\ell(P, K)$ (or well-posedness condition).

(b) follows from the identity

$$\mathcal{F}_\ell(P, K_1) - \mathcal{F}_\ell(P, K_2) = P_{12}(I - K_2P_{22})^{-1}(K_1 - K_2)(I - P_{22}K_1)^{-1}P_{21}.$$

(c) it is sufficient to show that the formula for K is well-posed and K thus obtained is proper. Let $Q = P^{-1}$, which will be proper since $\det P(\infty) \neq 0$, and define

$$K = \mathcal{F}_u(Q, G) = Q_{22} + Q_{21}G(I - Q_{11}G)^{-1}Q_{12}.$$

This expression is well-posed and proper since at $s = \infty$

$$\begin{aligned}\det(I - Q_{11}G) &= \det\left(I - \begin{bmatrix} I & 0 \end{bmatrix} P^{-1} \begin{bmatrix} I \\ 0 \end{bmatrix} G\right) \\ &= \det\left[P^{-1}\left(P - \begin{bmatrix} G & 0 \\ 0 & 0 \end{bmatrix}\right)\right] \neq 0.\end{aligned}$$

We also need to ensure that $\mathcal{F}_\ell(P, K)$ is well-posed:

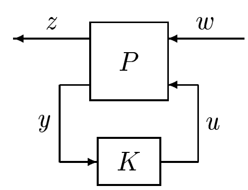
$$\begin{aligned}I - P_{22}K &= (I - P_{22}Q_{22}) - P_{22}Q_{21}G(I - Q_{11}G)^{-1}Q_{12} \\ &= P_{21}Q_{12} + P_{21}Q_{11}G(I - Q_{11}G)^{-1}Q_{12} \\ &= P_{21}(I - Q_{11}G)^{-1}Q_{12}.\end{aligned}$$

The above form is obtained by using the fact that $PQ = I$. Then $\det(I - P_{22}K) \neq 0$ since P_{21}^{-1} exists and $Q_{12}^{-1} = P_{21} - P_{22}P_{12}^{-1}P_{11}$. Hence the LFTs are both well-posed and we immediately obtain that $\mathcal{F}_\ell(P, K) = G$ as required upon substituting for K and $(I - P_{22}K)$, as shown above.

□

Remark 10.1 This lemma shows that under certain conditions, an LFT of transfer matrices is a bijective map between two sets of proper and real rational matrices. When given proper transfer matrices P and G with compatible dimensions which satisfy conditions in (c), there exists a unique proper K such that $G = \mathcal{F}_l(P, K)$. On the other hand, the conditions of part *(c)* show that the feedback systems are well-posed. ♡

Remark 10.2 A simple interpretation of the result (c) is given by considering the signals in the feedback systems,



assuming this structure is well-posed. And we have

$$\begin{aligned}\begin{bmatrix} z \\ y \end{bmatrix} &= P \begin{bmatrix} w \\ u \end{bmatrix}, \qquad u = Ky \\ \Rightarrow z &= \mathcal{F}_\ell(P, K)w \;=\; Gw;\end{aligned}$$

hence

$$\begin{aligned}\begin{bmatrix} w \\ u \end{bmatrix} &= P^{-1} \begin{bmatrix} z \\ y \end{bmatrix}, \qquad z = Gw \\ \Rightarrow u &= \mathcal{F}_u(P^{-1}, G)y, \;\; \text{or} \;\; K = \mathcal{F}_u(P^{-1}, G).\end{aligned}$$

♡

10.2 Examples of LFTs

LFT is a very convenient tool to formulate many mathematical objects. In this section and the sections to follow, some commonly encountered control or mathematical objects are given new perspectives, i.e., they will be examined from the LFT point of view.

Polynomials

A very commonly encountered object in control and mathematics is a polynomial function. For example,

$$p(\delta) = a_0 + a_1\delta + \cdots + a_n\delta^n$$

with indeterminate δ. It is easy to verify that $p(\delta)$ can be written in the following LFT form:

$$p(\delta) = \mathcal{F}_\ell(M, \delta I_n)$$

with

$$M = \left[\begin{array}{c:cccc} a_0 & a_1 & \cdots & a_{n-1} & a_n \\ \hdashline 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ 0 & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{array}\right].$$

Hence every polynomial is a linear fraction of their indeterminates. More generally, any multivariate (matrix) polynomials are also LFTs in their indeterminates; for example,

$$p(\delta_1, \delta_2) = a_1\delta_1^2 + a_2\delta_2^2 + a_3\delta_1\delta_2 + a_4\delta_1 + a_5\delta_2 + a_6.$$

Then

$$p(\delta_1, \delta_2) = \mathcal{F}_\ell(N, \Delta)$$

with

$$N = \left[\begin{array}{c:cccc} a_6 & 1 & 0 & 1 & 0 \\ \hdashline a_4 & 0 & a_1 & 0 & a_3 \\ 1 & 0 & 0 & 0 & 0 \\ a_5 & 0 & 0 & 0 & a_2 \\ 1 & 0 & 0 & 0 & 0 \end{array}\right], \qquad \Delta = \begin{bmatrix} \delta_1 I_2 & \\ & \delta_2 I_2 \end{bmatrix}.$$

It should be noted that these representations or realizations of polynomials are neither unique nor necessarily minimal. Here a minimal realization refers to a realization with the smallest possible dimension of Δ. As commonly known, in multidimensional systems and filter theory, it is usually very hard, if not impossible, to find a minimal realization for even a two variable polynomial. In fact, the minimal dimension of Δ depends also on the field (real, complex, etc.) of the realization. More detailed discussion of this issue is beyond the scope of this book, the interested readers should consult the references in 2-d or n-d systems or filter theory.

Rational Functions

As another example of LFT representation, we consider a rational matrix function (not necessarily proper), $F(\delta_1, \delta_2, \cdots, \delta_m)$, with a finite value at the origin: $F(0, 0, \cdots, 0)$ is finite. Then $F(\delta_1, \delta_2, \cdots, \delta_m)$ can be written as an LFT in $(\delta_1, \delta_2, \cdots, \delta_m)$ (some δ_i may be repeated). To see that, write

$$F(\delta_1, \delta_2, \cdots, \delta_m) = \frac{N(\delta_1, \delta_2, \cdots, \delta_m)}{d(\delta_1, \delta_2, \cdots, \delta_m)} = N(\delta_1, \delta_2, \cdots, \delta_m) \quad (d(\delta_1, \delta_2, \cdots, \delta_m)I)^{-1}$$

where $N(\delta_1, \delta_2, \cdots, \delta_m)$ is a multivariate matrix polynomial and $d(\delta_1, \delta_2, \cdots, \delta_m)$ is a scalar multivariate polynomial with $d(0, 0, \cdots, 0) \neq 0$. Both N and dI can be represented as LFTs, and, furthermore, since $d(0, 0, \cdots, 0) \neq 0$, the inverse of dI is also an LFT as shown in Lemma 10.3. Now the conclusion follows by the fact that the product of LFTs is also an LFT. (Of course, the above LFT representation problem is exactly the problem of state space realization for a multidimensional discrete transfer matrix.) However, this is usually not a nice way to get an LFT representation for a rational matrix since this approach usually results in a much higher dimensioned Δ than required. For example,

$$f(\delta) = \frac{\alpha + \beta\delta}{1 + \gamma\delta} = \mathcal{F}_\ell(M, \delta)$$

with

$$M = \left[\begin{array}{c|c} \alpha & \beta - \alpha\gamma \\ \hline 1 & -\gamma \end{array}\right].$$

By using the above approach, we would end up with

$$f(\delta) = \mathcal{F}_\ell(N, \delta I_2)$$

and

$$N = \left[\begin{array}{c|cc} \alpha & \beta & -\alpha\gamma \\ \hline 1 & 0 & -\gamma \\ 1 & 0 & -\gamma \end{array}\right].$$

Although the latter can be reduced to the former, it is not easy to see how to carry out such reduction for a complicated problem, even if it is possible.

State Space Realizations

We can use the LFT formulae to establish the relationship between transfer matrices and their state space realizations. A system with a state space realization as

$$\begin{aligned} \dot{x} &= Ax + Bu \\ y &= Cx + Du \end{aligned}$$

has a transfer matrix of

$$G(s) = D + C(sI - A)^{-1}B = \mathcal{F}_u\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \frac{1}{s}I\right).$$

Now take $\Delta = \frac{1}{s}I$, the transfer matrix can be written as

$$G(s) = \mathcal{F}_u(\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \Delta).$$

More generally, consider a discrete time 2-D (or MD) system realized by the first-order state space equation

$$\begin{aligned} x_1(k_1+1, k_2) &= A_{11}x_1(k_1,k_2) + A_{12}x_2(k_1,k_2) + B_1u(k_1,k_2) \\ x_2(k_1, k_2+1) &= A_{21}x_1(k_1,k_2) + A_{22}x_2(k_1,k_2) + B_2u(k_1,k_2) \\ y(k_1,k_2) &= C_1x_1(k_1,k_2) + C_2x_2(k_1,k_2) + Du(k_1,k_2). \end{aligned}$$

In the same way, take

$$\Delta = \begin{bmatrix} z_1^{-1}I & 0 \\ 0 & z_2^{-1}I \end{bmatrix} =: \begin{bmatrix} \delta_1 I & 0 \\ 0 & \delta_2 I \end{bmatrix}$$

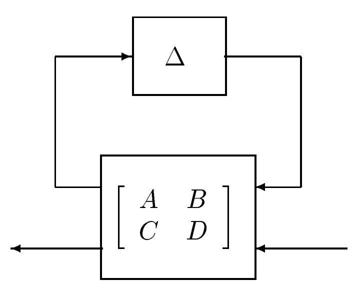
where z_i denotes the forward shift operator, and let

$$A \triangleq \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}, \quad B \triangleq \begin{bmatrix} B_1 \\ B_2 \end{bmatrix}, \quad C \triangleq \begin{bmatrix} C_1 & C_2 \end{bmatrix}$$

then its transfer matrix is

$$\begin{aligned} G(z_1,z_2) &= D + C(\begin{bmatrix} z_1 I & 0 \\ 0 & z_2 I \end{bmatrix} - A)^{-1}B = D + C\Delta(I - \Delta A)^{-1}B \\ &=: \mathcal{F}_u(\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \Delta). \end{aligned}$$

Both formulations can correspond to the following diagram:



The following notation for a transfer matrix has already been adopted in the previous chapters:

$$\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] := \mathcal{F}_u(\begin{bmatrix} A & B \\ C & D \end{bmatrix}, \Delta).$$

It is easy to see that this notation can be adopted for general dynamical systems, e.g., multidimensional systems, as far as the structure Δ is specified. This notation means that the transfer matrix can be expressed as an LFT of Δ with the coefficient matrix $\begin{bmatrix} A & B \\ C & D \end{bmatrix}$. In this special case, we say the parameter matrix Δ is the *frequency structure* of the system state space realization. This notation is deliberately somewhat ambiguous and can be viewed as both a transfer matrix and one of its realizations. The ambiguity is benign and convenient and can always be resolved from the context.

Frequency Transformation

The bilinear transformation between the z-domain and s-domain

$$s = \frac{z+1}{z-1}$$

transforms the unit disk to the left-half plane and is the simplest example of an LFT. We may rewrite it in our standard form as

$$\frac{1}{s}I = I - \sqrt{2}I \quad z^{-1}I \quad (I + z^{-1}I)^{-1} \; \sqrt{2}I = \mathcal{F}_u(N, z^{-1}I)$$

where

$$N = \begin{bmatrix} I & \sqrt{2}I \\ -\sqrt{2}I & -I \end{bmatrix}.$$

Now consider a continuous system

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \mathcal{F}_u(M, \frac{1}{s}I)$$

where

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix};$$

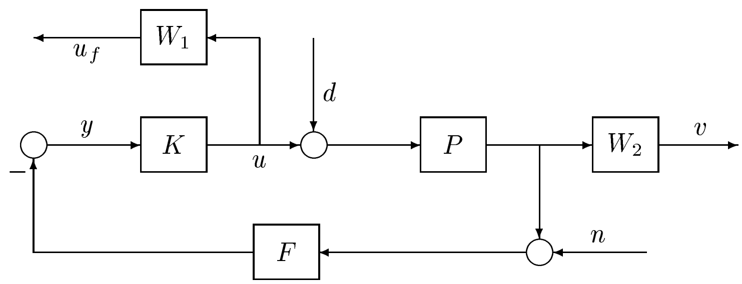
then the corresponding discrete time system realization is given by

$$\tilde{G}(z) = \mathcal{F}_u(M, \frac{z-1}{z+1}I) = \mathcal{F}_u(M, \mathcal{F}_u(N, z^{-1}I)) = \mathcal{F}_u(\tilde{M}, z^{-1}I)$$

with

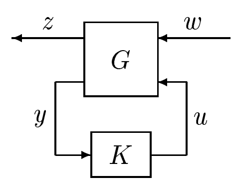
$$\tilde{M} = \begin{bmatrix} -(I-A)^{-1}(I+A) & -\sqrt{2}(I-A)^{-1}B \\ \sqrt{2}C(I-A)^{-1} & C(I-A)^{-1}B + D \end{bmatrix}.$$

The transformation from the z-domain to the s-domain can be obtained similarly.



Simple Block Diagrams

A feedback system with the following block diagram can be rearranged as an LFT:



with

$$w = \begin{pmatrix} d \\ n \end{pmatrix} \qquad\qquad z = \begin{pmatrix} v \\ u_f \end{pmatrix}$$

and

$$G = \left[\begin{array}{cc:c} W_2P & 0 & W_2P \\ 0 & 0 & W_1 \\ \hdashline -FP & -F & -FP \end{array}\right].$$

Constrained Structure Synthesis

Using the properties of LFTs, we can show that constrained structure control synthesis problems can be converted to constrained structure *constant* output feedback problems. Consider the synthesis structure in the last example and assume

$$G = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \qquad\qquad K = \left[\begin{array}{c|c} A_K & B_K \\ \hline C_K & D_K \end{array}\right].$$

Then it is easy to show that

$$\mathcal{F}_\ell(G, K) = \mathcal{F}_\ell(M(s), F)$$

where

$$M(s) = \left[\begin{array}{cc|c:cc} A & 0 & B_1 & 0 & B_2 \\ 0 & 0 & 0 & I & 0 \\ \hline C_1 & 0 & D_{11} & 0 & D_{12} \\ \hdashline 0 & I & 0 & 0 & 0 \\ C_2 & 0 & D_{21} & 0 & D_{22} \end{array}\right]$$

and

$$F = \begin{bmatrix} A_K & B_K \\ C_K & D_K \end{bmatrix}.$$

Note that F is a *constant matrix*, not a system matrix. Hence if the controller structure is fixed (or constrained), then the corresponding control problem becomes a constant (constrained) output feedback problem.

Parametric Uncertainty: A Mass/Spring/Damper System

One natural type of uncertainty is unknown coefficients in a state space model. To motivate this type of uncertainty description, we will begin with a familiar mass/spring/damper system, shown below in Figure 10.1.

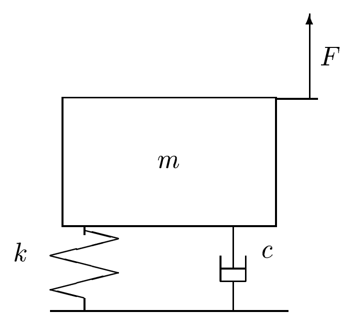


Figure 10.1: Mass/Spring/Damper System

The dynamical equation of the system motion can be described by

$$\ddot{x} + \frac{c}{m}\dot{x} + \frac{k}{m}x = \frac{F}{m}.$$

Suppose that the 3 physical parameters m, c, and k are not known exactly, but are believed to lie in known intervals. In particular, the actual mass m is within 10% of a nominal mass, $\bar{m}$, the actual damping value c is within 20% of a nominal value of $\bar{c}$, and the spring stiffness is within 30% of its nominal value of $\bar{k}$. Now introducing perturbations δ_m, δ_c, and δ_k, which are assumed to be unknown but lie in the interval $[-1, 1]$, the block diagram for the dynamical system can be shown in Figure 10.2.

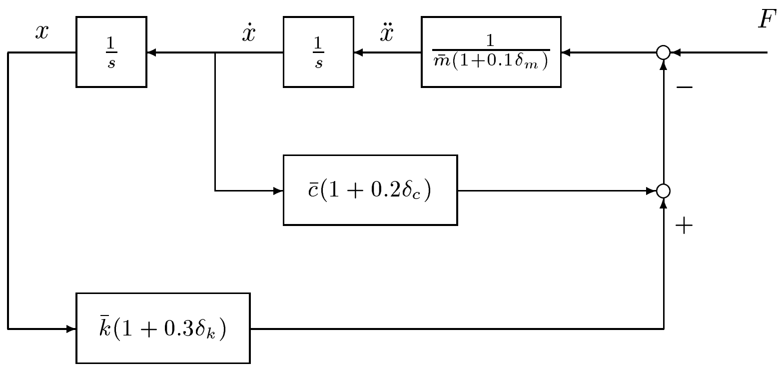


Figure 10.2: Block Diagram of Mass/Spring/Damper Equation

It is easy to check that $\frac{1}{m}$ can be represented as an LFT in δ_m:

$$\frac{1}{m} = \frac{1}{\bar{m}(1+0.1\delta_m)} = \frac{1}{\bar{m}} - \frac{0.1}{\bar{m}}\delta_m(1+0.1\delta_m)^{-1} = \mathcal{F}_\ell\left(M_1, \delta_m\right).$$

with $M_1 = \begin{bmatrix} \frac{1}{\bar{m}} & -\frac{0.1}{\bar{m}} \\ 1 & -0.1 \end{bmatrix}$. Suppose that the input signals of the dynamical system are selected as $x_1 = x, x_2 = \dot{x}, F$, and the output signals are selected as $\dot{x}_1$ and $\dot{x}_2$. To represent the system model as an LFT of the natural uncertainty parameters δ_m, δ_c and δ_k, we shall first isolate the uncertainty parameters and denote the inputs and outputs of δ_k, δ_c and δ_m as y_k, y_c, y_m and u_k, u_c, u_m, respectively, as shown in Figure 10.3. Then

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ y_k \\ y_c \\ y_m \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{\bar{k}}{\bar{m}} & -\frac{\bar{c}}{\bar{m}} & \frac{1}{\bar{m}} & -\frac{1}{\bar{m}} & -\frac{1}{\bar{m}} & -\frac{0.1}{\bar{m}} \\ 0.3\bar{k} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.2\bar{c} & 0 & 0 & 0 & 0 \\ -\bar{k} & -\bar{c} & 1 & -1 & -1 & -0.1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ F \\ u_k \\ u_c \\ u_m \end{bmatrix}, \quad \begin{bmatrix} u_k \\ u_c \\ u_m \end{bmatrix} = \Delta \begin{bmatrix} y_k \\ y_c \\ y_m \end{bmatrix}$$

i.e.,

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \mathcal{F}_\ell(M, \Delta) \begin{bmatrix} x_1 \\ x_2 \\ F \end{bmatrix}$$

where

$$M = \left[\begin{array}{ccc:ccc} 0 & 1 & 0 & 0 & 0 & 0 \\ -\frac{\bar{k}}{\bar{m}} & -\frac{\bar{c}}{\bar{m}} & \frac{1}{\bar{m}} & -\frac{1}{\bar{m}} & -\frac{1}{\bar{m}} & -\frac{0.1}{\bar{m}} \\ \hdashline 0.3\bar{k} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.2\bar{c} & 0 & 0 & 0 & 0 \\ -\bar{k} & -\bar{c} & 1 & -1 & -1 & -0.1 \end{array}\right], \quad \Delta = \begin{bmatrix} \delta_k & 0 & 0 \\ 0 & \delta_c & 0 \\ 0 & 0 & \delta_m \end{bmatrix}.$$

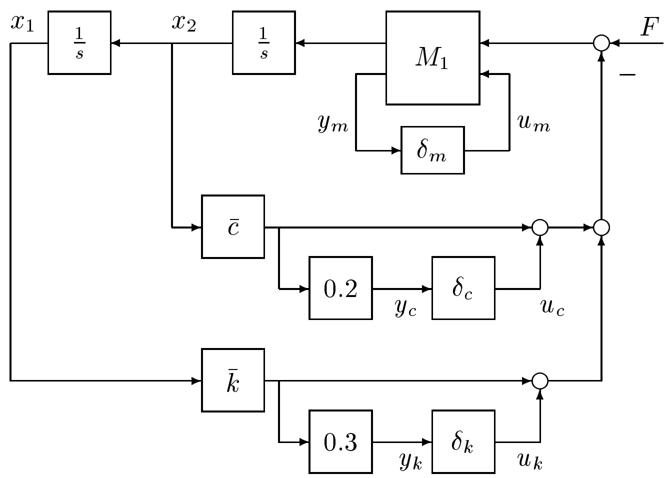


Figure 10.3: Mass/Spring/Damper System

General Affine State-Space Uncertainty

We will consider a special class of state space models with unknown coefficients and show how this type of uncertainty can be represented via the LFT formulae with respect to an uncertain parameter matrix so that the perturbations enter the system in a feedback form. This type of modeling will form the basic building blocks for components with *parametric* uncertainty.

Consider a linear system $G_\delta(s)$ that is parameterized by k uncertain parameters, $\delta_1, \dots, \delta_k$, and has the realization

$$G_\delta(s) = \left[\begin{array}{c|c} A + \sum_{i=1}^{k} \delta_i \hat{A}_i & B + \sum_{i=1}^{k} \delta_i \hat{B}_i \\ \hline C + \sum_{i=1}^{k} \delta_i \hat{C}_i & D + \sum_{i=1}^{k} \delta_i \hat{D}_i \end{array}\right].$$

Here $A, \hat{A}_i \in \mathbb{R}^{n\times n}, B, \hat{B}_i \in \mathbb{R}^{n\times n_u}, C, \hat{C}_i \in \mathbb{R}^{n_y\times n}$, and $D, \hat{D}_i \in \mathbb{R}^{n_y\times n_u}$.

The various terms in these state equations are interpreted as follows: the nominal system description $G(s)$, given by known matrices A, B, C, and D, is (A, B, C, D) and the parametric uncertainty in the nominal system is reflected by the k scalar uncertain parameters $\delta_1, \ldots, \delta_k$, and we can specify them, say by $\delta_i \in [-1, 1]$. The structural knowledge about the uncertainty is contained in the matrices $\hat{A}_i, \hat{B}_i, \hat{C}_i$, and $\hat{D}_i$. They reflect how the i'th uncertainty, δ_i, affects the state space model.

Now, we consider the problem of describing the perturbed system via the LFT formulae so that all the uncertainty can be represented as a nominal system with the unknown parameters entering it as the feedback gains. This is shown in Figure 10.4.

Since $G_\delta(s) = \mathcal{F}_u(M_\delta, \frac{1}{s}I)$ where

$$M_\delta \triangleq \begin{bmatrix} A + \sum_{i=1}^{k} \delta_i \hat{A}_i & B + \sum_{i=1}^{k} \delta_i \hat{B}_i \\ C + \sum_{i=1}^{k} \delta_i \hat{C}_i & D + \sum_{i=1}^{k} \delta_i \hat{D}_i \end{bmatrix},$$

we need to find an LFT representation for the matrix M_δ with respect to

$$\Delta_p = \operatorname{diag}\{\delta_1 I, \delta_2 I, \ldots \delta_k I\}.$$

To achieved this with the smallest possibly size of repeated blocks, let q_i denote the rank of the matrix

$$P_i \triangleq \begin{bmatrix} \hat{A}_i & \hat{B}_i \\ \hat{C}_i & \hat{D}_i \end{bmatrix} \in \mathbb{R}^{(n+n_y)\times(n+n_u)}$$

for each i. Then P_i can be written as

$$P_i = \begin{bmatrix} L_i \\ W_i \end{bmatrix} \begin{bmatrix} R_i \\ Z_i \end{bmatrix}^*$$

where $L_i \in \mathbb{R}^{n\times q_i}$, $W_i \in \mathbb{R}^{n_y\times q_i}$, $R_i \in \mathbb{R}^{n\times q_i}$ and $Z_i \in \mathbb{R}^{n_u\times q_i}$. Hence, we have

$$\delta_i P_i = \begin{bmatrix} L_i \\ W_i \end{bmatrix} [\delta_i I_{q_i}] \begin{bmatrix} R_i \\ Z_i \end{bmatrix}^*,$$

and M_δ can be written as

$$M_\delta = \overbrace{\begin{bmatrix} A & B \\ C & D \end{bmatrix}}^{M_{11}} + \overbrace{\begin{bmatrix} L_1 & \cdots & L_k \\ W_1 & \cdots & W_k \end{bmatrix}}^{M_{12}} \overbrace{\begin{bmatrix} \delta_1 I_{q_1} & & \\ & \ddots & \\ & & \delta_k I_{q_k} \end{bmatrix}}^{\Delta_p} \overbrace{\begin{bmatrix} R_1^* & Z_1^* \\ \vdots & \vdots \\ R_k^* & Z_k^* \end{bmatrix}}^{M_{21}}$$

i.e.

$$M_\delta = \mathcal{F}_\ell(\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & 0 \end{bmatrix}, \Delta_p).$$

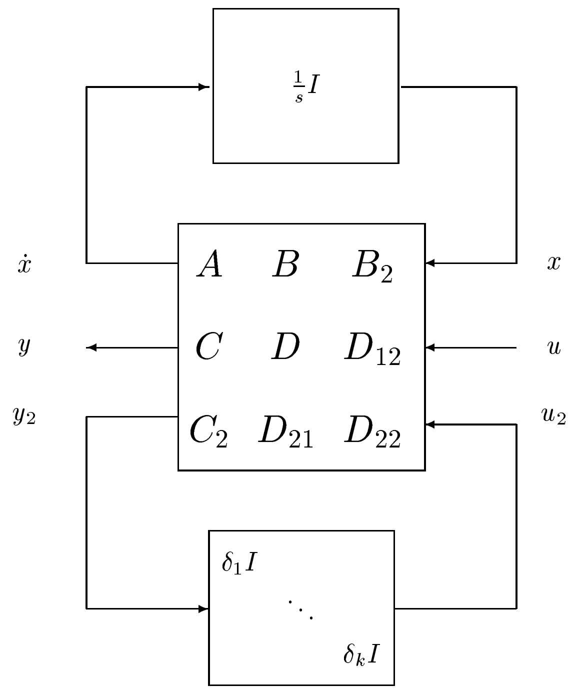


Figure 10.4: LFT Representation of State Space Uncertainty

Therefore, the matrices B_2, C_2, D_{12}, D_{21}, and D_{22} in the diagram are

$$\begin{aligned} B_2 &= \begin{bmatrix} L_1 & L_2 & \cdots & L_k \end{bmatrix} \\ D_{12} &= \begin{bmatrix} W_1 & W_2 & \cdots & W_k \end{bmatrix} \\ C_2 &= \begin{bmatrix} R_1 & R_2 & \cdots & R_k \end{bmatrix}^* \\ D_{21} &= \begin{bmatrix} Z_1 & Z_2 & \cdots & Z_k \end{bmatrix}^* \\ D_{22} &= 0 \end{aligned}$$

and

$$G_\delta(\Delta) = \mathcal{F}_u(\mathcal{F}_\ell(\begin{bmatrix} M_{11} & M_{12} \\ M_{21} & 0 \end{bmatrix}, \Delta_p), \frac{1}{s}I).$$

10.3 Basic Principle

We have studied several simple examples of the use of LFTs and, in particular, their role in modeling uncertainty. The basic principle at work here in writing a matrix LFT is often referred to as *"pulling out the Δs"*. We will try to illustrate this with another picture (inspired by Boyd and Barratt [1991]). Consider a structure with four substructures interconnected in some known way as shown in Figure 10.5.

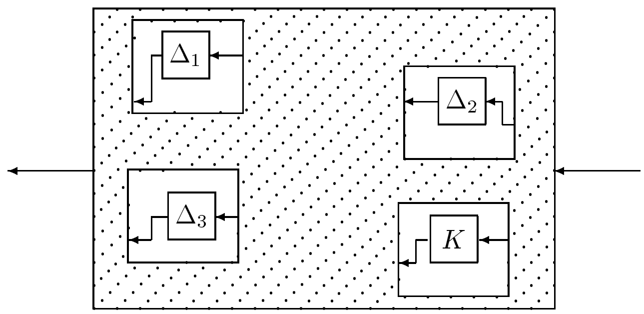


Figure 10.5: Multiple Source of Uncertain Structure

This diagram can be redrawn as a standard one via "pulling out the Δs" in Figure 10.6. Now the matrix "M" of the LFT can be obtained by computing the corresponding transfer matrix in the shadowed box.

We shall illustrate the above principle with an example. Consider an input/output relation

$$z = \frac{a + b\delta_2 + c\delta_1\delta_2^2}{1 + d\delta_1\delta_2 + e\delta_1^2} w =: Gw$$

where a, b, c, d and e are given constants or transfer functions. We would like to write G as an LFT in terms of δ_1 and δ_2. We shall do this in three steps:

1. Draw a block diagram for the input/output relation with each δ separated as shown in Figure 10.7.
2. Mark the inputs and outputs of the δ's as y's and u's, respectively. (This is essentially *pulling out the δs*).
3. Write z and y's in terms of w and u's with all δ's taken out. (This step is equivalent to computing the transformation in the shadowed box in Figure 10.6.)

$$\begin{bmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ z \end{bmatrix} = M \begin{bmatrix} u_1 \\ u_2 \\ u_3 \\ u_4 \\ w \end{bmatrix}$$

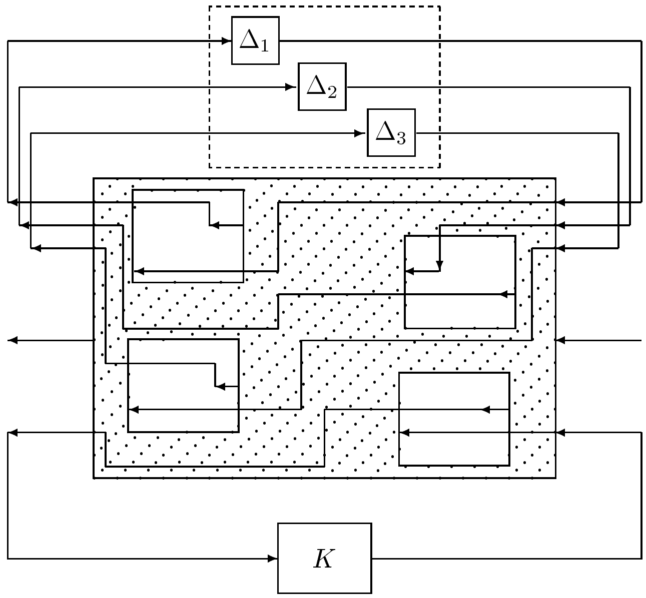


Figure 10.6: Pulling out the Δs

where

$$M = \left[\begin{array}{cccc:c} 0 & e & d & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & be & bd+c & 0 & b \\ \hdashline 0 & ae & ad & 1 & a \end{array}\right].$$

Then

$$z = \mathcal{F}_u(M, \Delta)w, \quad \Delta = \begin{bmatrix} \delta_1 I_2 & 0 \\ 0 & \delta_2 I_2 \end{bmatrix}.$$

All LFT examples in the last section can be obtained following the above steps.

10.4 Redheffer Star-Products

The most important property of LFTs is that any interconnection of LFTs is again an LFT. This property is by far the most often used and is the heart of LFT machinery. Indeed, it is not hard to see that most of the interconnection structures discussed early, e.g., feedback and cascade, can be viewed as special cases of the so-called *star product*.

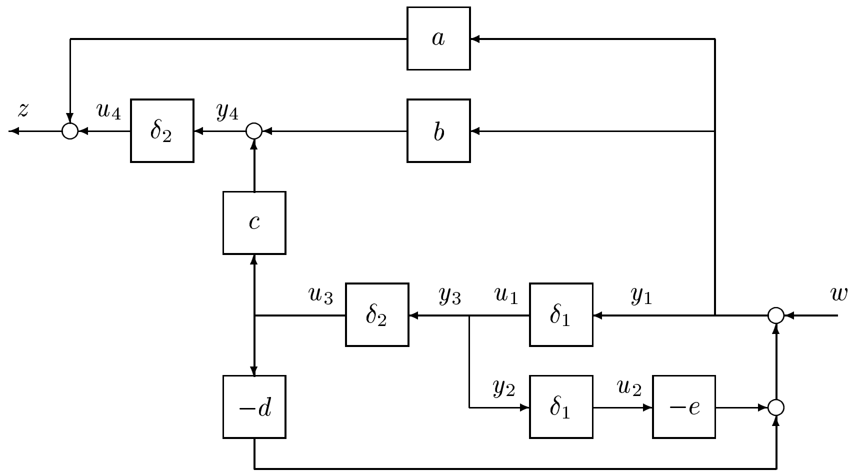


Figure 10.7: Block diagram for G

Suppose that P and K are compatibly partitioned matrices

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix}, \quad K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$$

such that the matrix product $P_{22}K_{11}$ is well defined and square, and assume further that $I - P_{22}K_{11}$ is invertible. Then the *star product of* P *and* K *with respect to this partition* is defined as

$$\mathcal{S}(P,K) := \begin{bmatrix} F_l(P, K_{11}) & P_{12}(I - K_{11}P_{22})^{-1} K_{12} \\ K_{21}(I - P_{22}K_{11})^{-1} P_{21} & F_u(K, P_{22}) \end{bmatrix}.$$

Note that this definition is dependent on the partitioning of the matrices P and K above. In fact, this star product may be well defined for one partition and not well defined for another; however, we will not explicitly show this dependence because it is always clear from the context. In a block diagram, this dependence appears, as shown in Figure 10.8.

Now suppose that P and K are transfer matrices with state space representations:

$$P = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] \qquad K = \left[\begin{array}{c|cc} A_K & B_{K1} & B_{K2} \\ \hline C_{K1} & D_{K11} & D_{K12} \\ C_{K2} & D_{K21} & D_{K22} \end{array}\right].$$

Then the transfer matrix

$$\mathcal{S}(P,K) : \begin{bmatrix} w \\ \hat{w} \end{bmatrix} \mapsto \begin{bmatrix} z \\ \hat{z} \end{bmatrix}$$

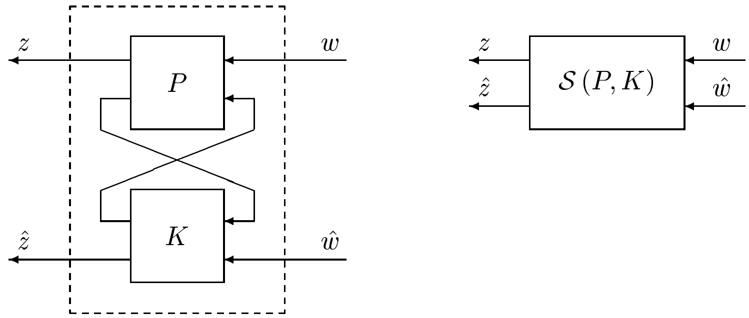


Figure 10.8: Interconnection of LFTs

has a representation

$$\mathcal{S}(P,K)=\left[\begin{array}{c|cc}\bar{A} & \bar{B}_1 & \bar{B}_2\\ \hline \bar{C}_1 & \bar{D}_{11} & \bar{D}_{12}\\ \bar{C}_2 & \bar{D}_{21} & \bar{D}_{22}\end{array}\right]=\left[\begin{array}{c|c}\bar{A} & \bar{B}\\ \hline \bar{C} & \bar{D}\end{array}\right]$$

where

$$\begin{aligned}
\bar{A} &= \begin{bmatrix} A+B_2\tilde{R}^{-1}D_{K11}C_2 & B_2\tilde{R}^{-1}C_{K1}\\ B_{K1}R^{-1}C_2 & A_K+B_{K1}R^{-1}D_{22}C_{K1}\end{bmatrix}\\
\bar{B} &= \begin{bmatrix} B_1+B_2\tilde{R}^{-1}D_{K11}D_{21} & B_2\tilde{R}^{-1}D_{K12}\\ B_{K1}R^{-1}D_{21} & B_{K2}+B_{K1}R^{-1}D_{22}D_{K12}\end{bmatrix}\\
\bar{C} &= \begin{bmatrix} C_1+D_{12}D_{K11}R^{-1}C_2 & D_{12}\tilde{R}^{-1}C_{K1}\\ D_{K21}R^{-1}C_2 & C_{K2}+D_{K21}R^{-1}D_{22}C_{K1}\end{bmatrix}\\
\bar{D} &= \begin{bmatrix} D_{11}+D_{12}D_{K11}R^{-1}D_{21} & D_{12}\tilde{R}^{-1}D_{K12}\\ D_{K21}R^{-1}D_{21} & D_{K22}+D_{K21}R^{-1}D_{22}D_{K12}\end{bmatrix}
\end{aligned}$$

$$R=I-D_{22}D_{K11},\quad \tilde{R}=I-D_{K11}D_{22}.$$

In fact, it is easy to show that

$$\begin{aligned}
\bar{A} &= \mathcal{S}\left(\begin{bmatrix} A & B_2\\ C_2 & D_{22}\end{bmatrix},\begin{bmatrix} D_{K11} & C_{K1}\\ B_{K1} & A_K\end{bmatrix}\right),\\
\bar{B} &= \mathcal{S}\left(\begin{bmatrix} B_1 & B_2\\ D_{21} & D_{22}\end{bmatrix},\begin{bmatrix} D_{K11} & D_{K12}\\ B_{K1} & B_{K2}\end{bmatrix}\right),
\end{aligned}$$

$$\begin{aligned}
\bar{C} &= \mathcal{S}\left(\begin{bmatrix} C_1 & D_{12} \\ C_2 & D_{22} \end{bmatrix}, \begin{bmatrix} D_{K11} & C_{K1} \\ D_{K21} & C_{K2} \end{bmatrix}\right), \\
\bar{D} &= \mathcal{S}\left(\begin{bmatrix} D_{11} & D_{12} \\ D_{21} & D_{22} \end{bmatrix}, \begin{bmatrix} D_{K11} & D_{K12} \\ D_{K21} & D_{K22} \end{bmatrix}\right).
\end{aligned}$$

10.5 Notes and References

This chapter is based on the lecture notes by Packard [1991] and the paper by Doyle, Packard, and Zhou [1991].

11

Structured Singular Value

It is noted that the robust stability and robust performance criteria derived in Chapter 9 vary with the assumptions on the uncertainty descriptions and performance requirements. We will show in this chapter that they can all be treated in a unified framework using the LFT machinery introduced in the last chapter and the structured singular value to be introduced in this chapter. This, of course, does not mean that those special problems and their corresponding results are not important; on the contrary, they are sometimes very enlightening to our understanding of complex problems such as those in which complex problems are formed from simple problems. On the other hand, a unified approach may relieve the mathematical burden of dealing with specific problems repeatedly. Furthermore, the unified framework introduced here will enable us to treat exactly the robust stability and robust performance problems for systems with multiple sources of uncertainties, which is a formidable problem in the standing point of Chapter 9, in the same fashion as single unstructured uncertainty. Indeed, if a system is subject to multiple sources of uncertainties, in order to use the results in Chapter 9 for unstructured cases, it is necessary to reflect all sources of uncertainties from their known point of occurrence to a single reference location in the loop. Such reflected uncertainties invariably have a great deal of structure which must then be "covered up" with a large, arbitrarily more conservative perturbation in order to maintain a simple cone bounded representation at the reference location. Readers might have already had some idea about the conservativeness in such reflection from the skewed specification problem, where an input multiplicative uncertainty of the plant is reflected at the output and the size of the reflected uncertainty is proportional to the condition number of the plant. In general, the reflected uncertainty may be proportional to the condition number of the transfer matrix between its original location and the reflected location. Thus it is highly desirable to treat the uncertainties as they are and where they are. The structured singular value is defined exactly for that purpose.

11.1 General Framework for System Robustness

As we have illustrated in the last chapter, any interconnected system may be rearranged to fit the general framework in Figure 11.1. Although the interconnection structure can become quite complicated for complex systems, many software packages, such as SIMULINK[1] and μ-TOOLS[2], are available which could be used to generate the interconnection structure from system components. Various modeling assumptions will be considered and the impact of these assumptions on analysis and synthesis methods will be explored in this general framework.

Note that uncertainty may be modeled in two ways, either as external inputs or as perturbations to the nominal model. The performance of a system is measured in terms of the behavior of the outputs or errors. The assumptions which characterize the uncertainty, performance, and nominal models determine the analysis techniques which must be used. The models are assumed to be FDLTI systems. The uncertain inputs are assumed to be either filtered white noise or weighted power or weighted $\mathcal{L}_p$ signals. Performance is measured as weighted output variances, or as power, or as weighted output $\mathcal{L}_p$ norms. The perturbations are assumed to be themselves FDLTI systems which are norm-bounded as input-output operators. Various combinations of these assumptions form the basis for all the standard linear system analysis tools.

Given that the nominal model is an FDLTI system, the interconnection system has the form

$$P(s) = \begin{bmatrix} P_{11}(s) & P_{12}(s) & P_{13}(s) \\ P_{21}(s) & P_{22}(s) & P_{23}(s) \\ P_{31}(s) & P_{32}(s) & P_{33}(s) \end{bmatrix}$$

and the closed-loop system is an LFT on the perturbation and the controller given by

$$\begin{aligned} z &= \mathcal{F}_u\left(\mathcal{F}_\ell(P, K), \Delta\right) w \\ &= \mathcal{F}_\ell\left(\mathcal{F}_u(P, \Delta), K\right) w. \end{aligned}$$

We will focus our discussion in this section on analysis methods; therefore, the controller may be viewed as just another system component and absorbed into the interconnection structure. Denote

$$M(s) = \mathcal{F}_\ell\left(P(s), K(s)\right) = \begin{bmatrix} M_{11}(s) & M_{12}(s) \\ M_{21}(s) & M_{22}(s) \end{bmatrix},$$

[1]SIMULINK is a trademark of The MathWorks, Inc.
[2]μ-TOOLS is a trademark of MUSYN Inc.

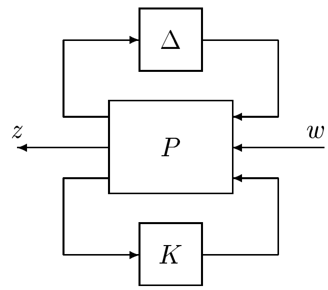


Figure 11.1: General Framework

and then the general framework reduces to Figure 11.2, where

$$z = \mathcal{F}_u(M, \Delta)w = \left[M_{22} + M_{21}\Delta(I - M_{11}\Delta)^{-1}M_{12}\right] w.$$

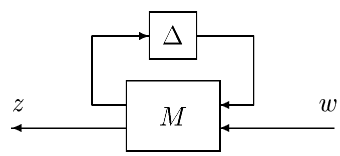


Figure 11.2: Analysis Framework

Suppose $K(s)$ is a stabilizing controller for the nominal plant P. Then $M(s) \in \mathcal{RH}_\infty$. In general, the stability of $\mathcal{F}_u(M, \Delta)$ does not necessarily imply the internal stability of the closed-loop feedback system. However, they can be made equivalent with suitably chosen w and z. For example, consider again the multiplicatively perturbed system shown in Figure 11.3. Now let

$$w := \begin{bmatrix} d_1 \\ d_2 \end{bmatrix}, \quad z := \begin{bmatrix} e_1 \\ e_2 \end{bmatrix}$$

then the system is robustly stable for all $\Delta(s) \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty < 1$ if and only if $\mathcal{F}_u(M, \Delta) \in \mathcal{RH}_\infty$ for all admissible Δ, which is guaranteed by $\|M_{11}\|_\infty \leq 1$. (Note that this is not necessarily equivalent to $(I - M_{11}\Delta)^{-1} \in \mathcal{RH}_\infty$ if Δ belongs to a closed ball as shown in Theorem 9.5.)

The analysis results presented in the previous chapters together with the associated synthesis tools are summarized in Table 11.1 with various uncertainty modeling assumptions.

However, the analysis is not so simple for systems with multiple sources of model uncertainties, including the robust performance problem for systems with unstructured

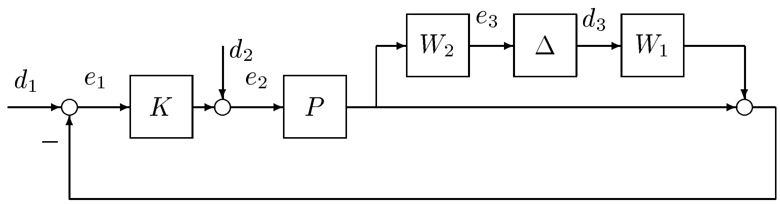


Figure 11.3: Multiplicatively Perturbed Systems

uncertainty. As we have shown in the last chapter, if a system is built from components which are themselves uncertain, then, in general, the uncertainty in the system level is structured involving typically a large number of real parameters. The stability analysis involving real parameters is much more involved and is beyond the scope of this book and that, instead, we shall simply cover the real parametric uncertainty with norm bounded dynamical uncertainty. Moreover, the interconnection model M can always be chosen so that $\Delta(s)$ is block diagonal, and, by absorbing any weights, $\|\Delta\|_\infty < 1$. Thus we shall assume that $\Delta(s)$ takes the form of

$$\Delta(s) = \{\text{diag }[\delta_1 I_{r_1}, \ldots, \delta_s I_{r_S}, \Delta_1, \ldots, \Delta_F] : \quad \delta_i(s) \in \mathcal{RH}_\infty, \quad \Delta_j \in \mathcal{RH}_\infty\}$$

with $\|\delta_i\|_\infty < 1$ and $\|\Delta_j\|_\infty < 1$. Then the system is robustly stable iff the interconnected system in Figure 11.4 is stable.

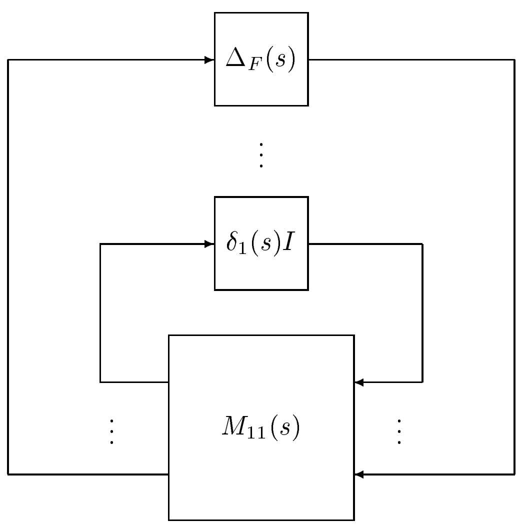


Figure 11.4: Robust Stability Analysis Framework

The results of Table 11.1 can be applied to the analyses of the system's robust stability in two ways:

<table>
<tr><th>Input Assumptions</th><th>Performance Specifications</th><th>Perturbation Assumptions</th><th>Analysis Tests</th><th>Synthesis Methods</th></tr>
<tr><td>$E(w(t)w(t)^*) = I$</td><td>$E(z(t)^* z(t)) \leq 1$</td><td rowspan="2">$\Delta = 0$</td><td rowspan="2">$\|M_{22}\|_2 \leq 1$</td><td rowspan="2">LQG
Wiener-Hopf
$\mathcal{H}_2$</td></tr>
<tr><td>$w = U_0 \delta(t)$
$E(U_0 U_0^*) = I$</td><td>$E(\|z\|_2^2) \leq 1$</td></tr>
<tr><td>$\|w\|_2 \leq 1$</td><td>$\|z\|_2 \leq 1$</td><td>$\Delta = 0$</td><td>$\|M_{22}\|_\infty \leq 1$</td><td rowspan="2">Singular Value Loop Shaping
$\mathcal{H}_\infty$</td></tr>
<tr><td>$\|w\|_2 \leq 1$</td><td>Internal Stability</td><td>$\|\Delta\|_\infty < 1$</td><td>$\|M_{11}\|_\infty \leq 1$</td></tr>
</table>

Table 11.1: General Analysis for Single Source of Uncertainty

(1) $\|M_{11}\|_\infty \leq 1$ implies stability, but not conversely, because this test ignores the known block diagonal structure of the uncertainties and is equivalent to regarding Δ as unstructured. This can be arbitrarily conservative[3] in that stable systems can have arbitrarily large $\|M_{11}\|_\infty$.

(2) Test for each δ_i (Δ_j) individually (assuming no uncertainty in other channels). This test can be arbitrarily optimistic because it ignores interaction between the δ_i (Δ_j). This optimism is also clearly shown in the spinning body example.

The difference between the stability margins(or bounds on Δ) obtained in (1) and (2) can be arbitrarily far apart. Only when the margins are close can conclusions be made about the general case with structured uncertainty.

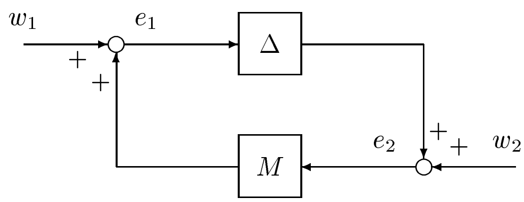
The exact stability and performance analysis for systems with structured uncertainty requires a new matrix function called the structured singular value (SSV) which is denoted by μ.

[3]By "arbitrarily conservative," we mean that examples can be constructed where the degree of conservatism is arbitrarily large. Of course, other examples exist where it is quite reasonable, see for example the spinning body example.

11.2 Structured Singular Value

11.2.1 Basic Concept

Conceptually, the structured singular value is nothing but a straightforward generalization of the singular values for constant matrices. To be more specific, it is instructive at this point to consider again the robust stability problem of the following standard feedback interconnection with stable $M(s)$ and $\Delta(s)$.



One important question one might ask is how large Δ (in the sense of $\|\Delta\|_\infty$) can be without destabilizing the feedback system. Since the closed-loop poles are given by $\det(I - M\Delta) = 0$, the feedback system becomes unstable if $\det(I - M(s)\Delta(s)) = 0$ for some $s \in \overline{\mathbb{C}}_+$. Now let $\alpha > 0$ be a sufficiently small number such that the closed-loop system is stable for all stable $\|\Delta\|_\infty < \alpha$. Next increase α until α_{max} so that the closed-loop system becomes unstable. So α_{max} is the robust stability margin. By small gain theorem,

$$\frac{1}{\alpha_{max}} = \|M\|_\infty := \sup_{s \in \overline{\mathbb{C}}_+} \overline{\sigma}(M(s)) = \sup_\omega \overline{\sigma}(M(j\omega))$$

if Δ is unstructured. Note that for any fixed $s \in \overline{\mathbb{C}}_+$, $\overline{\sigma}(M(s))$ can be written as

$$\overline{\sigma}(M(s)) = \frac{1}{\min\left\{\overline{\sigma}(\Delta) : \det\left(I - M(s)\Delta\right) = 0,\ \Delta \text{ is unstructured}\right\}}. \tag{11.1}$$

In other words, the reciprocal of the *largest singular value* of M is a measure of the smallest unstructured Δ that causes instability of the feedback system.

To quantify the smallest destabilizing *structured* Δ, the concept of singular values needs to be generalized. In view of the characterization of the largest singular value of a matrix $M(s)$ given by (11.1), we shall define

$$\mu_\Delta(M(s)) = \frac{1}{\min\left\{\overline{\sigma}(\Delta) : \det\left(I - M(s)\Delta\right) = 0,\ \Delta \text{ is structured}\right\}} \tag{11.2}$$

as the largest structured singular value of $M(s)$ with respect to the structured Δ. Then it is obvious that the robust stability margin of the feedback system with structured uncertainty Δ is

$$\frac{1}{\alpha_{max}} = \sup_{s \in \overline{\mathbb{C}}_+} \mu_\Delta(M(s)) = \sup_\omega \mu_\Delta(M(j\omega)).$$

The last equality follows from the following lemma.

Lemma 11.1 *Let Δ be a structured set and $M(s) \in \mathcal{RH}_\infty$. Then*

$$\sup_{s \in \overline{\mathbb{C}}_+} \mu_\Delta(M(s)) = \sup_{s \in \mathbb{C}_+} \mu_\Delta(M(s)) = \sup_\omega \mu_\Delta(M(j\omega)).$$

Proof. It is clear that

$$\sup_{s \in \overline{\mathbb{C}}_+} \mu_\Delta(M(s)) = \sup_{s \in \mathbb{C}_+} \mu_\Delta(M(s)) \geq \sup_\omega \mu_\Delta(M(j\omega)).$$

Now suppose $\sup_{s \in \mathbb{C}_+} \mu_\Delta(M(s)) > 1/\alpha$, then by the definition of μ, there is an $s_o \in \overline{\mathbb{C}}_+ \cup \{\infty\}$ and a complex structured Δ such that $\overline{\sigma}(\Delta) < \alpha$ and $\det(I - M(s_o)\Delta) = 0$. This implies that there is a $0 \leq \hat{\omega} \leq \infty$ and $0 < \beta \leq 1$ such that $\det(I - M(j\hat{\omega})\beta\Delta) = 0$. This in turn implies that $\mu_\Delta(M(j\hat{\omega})) > 1/\alpha$ since $\overline{\sigma}(\beta\Delta) < \alpha$. In other words, $\sup_{s \in \mathbb{C}_+} \mu_\Delta(M(s)) \leq \sup_\omega \mu_\Delta(M(j\omega))$. The proof is complete. □

The formal definition and characterization of the structured singular value of a constant matrix will be given below.

11.2.2 Definitions of μ

This section is devoted to defining the structured singular value, a matrix function denoted by $\mu(\cdot)$. We consider matrices $M \in \mathbb{C}^{n \times n}$. In the definition of $\mu(M)$, there is an underlying structure $\mathbf{\Delta}$ (a prescribed set of block diagonal matrices) on which everything in the sequel depends. For each problem, this structure is, in general, different; it depends on the *uncertainty* and *performance objectives* of the problem. Defining the structure involves specifying three things: the type of each block, the total number of blocks, and their dimensions.

There are two types of blocks: *repeated scalar* and *full* blocks. Two nonnegative integers, S and F, represent the number of *repeated scalar* blocks and the number of *full* blocks, respectively. To bookkeep their dimensions, we introduce positive integers $r_1, \ldots, r_S$; $m_1, \ldots, m_F$. The i'th repeated scalar block is $r_i \times r_i$, while the j'th full block is $m_j \times m_j$. With those integers given, we define $\mathbf{\Delta} \subset \mathbb{C}^{n \times n}$ as

$$\mathbf{\Delta} = \left\{ \text{diag}\ [\delta_1 I_{r_1}, \ldots, \delta_s I_{r_S}, \Delta_1, \ldots, \Delta_F] : \delta_i \in \mathbb{C}, \Delta_j \in \mathbb{C}^{m_j \times m_j} \right\}. \qquad (11.3)$$

For consistency among all the dimensions, we must have

$$\sum_{i=1}^{S} r_i + \sum_{j=1}^{F} m_j = n.$$

Often, we will need norm bounded subsets of $\mathbf{\Delta}$, and we introduce the following notation:

$$\mathbf{B\Delta} = \{\Delta \in \mathbf{\Delta} : \overline{\sigma}(\Delta) \leq 1\} \qquad (11.4)$$

$$\mathbf{B}^o\mathbf{\Delta} = \{\Delta \in \mathbf{\Delta} : \overline{\sigma}(\Delta) < 1\} \qquad (11.5)$$

where the superscript "o" symbolizes the open ball. To keep the notation as simple as possible in (11.3), we place all of the repeated scalar blocks first; in actuality, they can come in any order. Also, the full blocks do not have to be square, but restricting them as such saves a great deal in terms of notation.

Definition 11.1 *For $M \in \mathbb{C}^{n\times n}$, $\mu_{\boldsymbol{\Delta}}(M)$ is defined as*

$$\mu_{\boldsymbol{\Delta}}(M) := \frac{1}{\min\left\{\overline{\sigma}(\Delta) : \Delta \in \boldsymbol{\Delta}, \det(I - M\Delta) = 0\right\}} \tag{11.6}$$

unless no $\Delta \in \boldsymbol{\Delta}$ makes $I - M\Delta$ singular, in which case $\mu_{\boldsymbol{\Delta}}(M) := 0$.

Remark 11.1 Without a loss in generality, the full blocks in the minimal norm Δ can each be chosen to be dyads (rank = 1). To see this, assume $S = 0$, i.e., all blocks are full blocks. Suppose that $I - M\Delta$ is singular for some $\Delta \in \boldsymbol{\Delta}$. Then there is an $x \in \mathbb{C}^n$ such that $M\Delta x = x$. Now partition x conformably with Δ:

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_F \end{bmatrix}, \quad x_i \in \mathbb{C}^{m_i}, i = 1, \ldots, F$$

and define

$$D = \operatorname{diag}(d_1 I_{m_1}, \ldots, d_F I_{m_F})$$

where $d_i = 1$ if $x_i = 0$ and $d_i = 1/\left\|x_i\right\|$ if $x_i \neq 0$. Next define

$$\tilde{x} = \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \\ \vdots \\ \tilde{x}_F \end{bmatrix} := Dx = \begin{bmatrix} d_1 x_1 \\ d_2 x_2 \\ \vdots \\ d_m x_F \end{bmatrix}$$

and

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_F \end{bmatrix} := \Delta\tilde{x} = \begin{bmatrix} \Delta_1\tilde{x}_1 \\ \Delta_2\tilde{x}_2 \\ \vdots \\ \Delta\tilde{x}_F \end{bmatrix}.$$

It follows that $\|\tilde{x}_i\| = 1$ if $x_i \neq 0$, $\|\tilde{x}_i\| = 0$ if $x_i = 0$, and $y \neq 0$. Hence, define a new perturbation $\tilde{\Delta} \in \mathbb{C}^{n\times n}$ as

$$\tilde{\Delta} := \operatorname{diag}\left[y_1\tilde{x}_1^*, \ldots, y_F\tilde{x}_F^*\right].$$

Obviously, $\overline{\sigma}(\tilde{\Delta}) \leq \overline{\sigma}(\Delta)$ and $y = \tilde{\Delta}\tilde{x}$. Note that $D\Delta = \Delta D$ and $D\tilde{\Delta} = \tilde{\Delta}D$, we have

$$M\Delta x = x \implies MD^{-1}\Delta\tilde{x} = x \implies MD^{-1}y = x$$
$$\implies MD^{-1}\tilde{\Delta}\tilde{x} = x \implies M\tilde{\Delta}x = x$$

i.e., $I - M\tilde{\Delta}$ is also singular. Hence we have replaced a general perturbation Δ which satisfies the singularity condition with a perturbation $\tilde{\Delta}$ that is no larger (in the $\overline{\sigma}(\cdot)$ sense) and has rank 1 for each blocks but still satisfies the singularity condition. ♡

An alternative expression for $\mu_{\mathbf{\Delta}}(M)$ follows from the definition.

Lemma 11.2 $\mu_{\mathbf{\Delta}}(M) = \max_{\Delta \in \mathbf{B\Delta}} \rho(M\Delta)$

In view of this lemma, continuity of the function $\mu : \mathbb{C}^{n\times n} \to \mathbb{R}$ is apparent. In general, though, the function $\mu : \mathbb{C}^{n\times n} \to \mathbb{R}$ is not a norm, since it doesn't satisfy the triangle inequality; however, for any $\alpha \in \mathbb{C}$, $\mu(\alpha M) = |\alpha|\mu(M)$, so in some sense, it is related to how "big" the matrix is.

We can relate $\mu_{\mathbf{\Delta}}(M)$ to familiar linear algebra quantities when $\mathbf{\Delta}$ is one of two extreme sets.

- If $\mathbf{\Delta} = \{\delta I : \delta \in \mathbb{C}\}$ $(S = 1, F = 0, r_1 = n)$, then $\mu_{\mathbf{\Delta}}(M) = \rho(M)$, the spectral radius of M.

  Proof. The only Δ's in $\mathbf{\Delta}$ which satisfy the $\det(I - M\Delta) = 0$ constraint are reciprocals of nonzero eigenvalues of M. The smallest one of these is associated with the largest (magnitude) eigenvalue, so, $\mu_{\mathbf{\Delta}}(M) = \rho(M)$. □

- If $\mathbf{\Delta} = \mathbb{C}^{n\times n}$ $(S=0, F=1, m_1=n)$, then $\mu_{\mathbf{\Delta}}(M) = \overline{\sigma}(M)$.

  Proof. If $\overline{\sigma}(\Delta) < \frac{1}{\overline{\sigma}(M)}$, then $\overline{\sigma}(M\Delta) < 1$, so $I - M\Delta$ is nonsingular. Applying equation (11.6) implies $\mu_{\mathbf{\Delta}}(M) \leq \overline{\sigma}(M)$. On the other hand, let u and v be unit vectors satisfying $Mv = \overline{\sigma}(M)u$, and define $\Delta := \frac{1}{\overline{\sigma}(M)}vu^*$. Then $\overline{\sigma}(\Delta) = \frac{1}{\overline{\sigma}(M)}$ and $I - M\Delta$ is obviously singular. Hence, $\mu_{\mathbf{\Delta}}(M) \geq \overline{\sigma}(M)$. □

Obviously, for a general $\mathbf{\Delta}$ as in (11.3) we must have

$$\{\delta I_n : \delta \in \mathbb{C}\} \subset \mathbf{\Delta} \subset \mathbb{C}^{n\times n}. \tag{11.7}$$

Hence directly from the definition of μ and from the two special cases above, we conclude that

$$\rho(M) \leq \mu_{\mathbf{\Delta}}(M) \leq \overline{\sigma}(M). \tag{11.8}$$

These bounds alone are not sufficient for our purposes because the gap between ρ and $\overline{\sigma}$ can be arbitrarily large. For example, suppose $\Delta = \begin{bmatrix} \delta_1 & 0 \\ 0 & \delta_2 \end{bmatrix}$ and consider

(1) $M = \begin{bmatrix} 0 & \beta \\ 0 & 0 \end{bmatrix}$ for any $\beta > 0$. Then $\rho(M) = 0$ and $\overline{\sigma}(M) = \beta$. But $\det(I - M\Delta) = 1$ so $\mu(M) = 0$.

(2) $M = \begin{bmatrix} -1/2 & 1/2 \\ -1/2 & 1/2 \end{bmatrix}$. Then $\rho(M) = 0$ and $\overline{\sigma}(M) = 1$. Since $\det(I - M\Delta) = 1 + \frac{\delta_1 - \delta_2}{2}$, it is easy to see that $\min\left\{\max_i |\delta_i|, 1 + \frac{\delta_1-\delta_2}{2} = 0\right\} = 1$, so $\mu(M) = 1$.

Thus neither ρ nor $\overline{\sigma}$ provide useful bounds even in simple cases. The only time they do provide reliable bounds is when $\rho \approx \overline{\sigma}$.

However, the bounds can be refined by considering transformations on M that *do not affect* $\mu_{\mathbf{\Delta}}(M)$, but *do affect* ρ and $\overline{\sigma}$. To do this, define the following two subsets of $\mathbb{C}^{n\times n}$:

$$\mathcal{U} = \{U \in \mathbf{\Delta} : \ UU^* = I_n\} \tag{11.9}$$

$$\mathcal{D} = \left\{ \begin{array}{c} \operatorname{diag}\left[D_1, \ldots, D_S, d_1 I_{m_1}, \ldots, d_{F-1} I_{m_{F-1}}, I_{m_F}\right] : \\ D_i \in \mathbb{C}^{r_i \times r_i}, D_i = D_i^* > 0, d_j \in \mathbb{R}, d_j > 0 \end{array} \right\}. \tag{11.10}$$

Note that for any $\Delta \in \mathbf{\Delta}, U \in \mathcal{U}$, and $D \in \mathcal{D}$,

$$U^* \in \mathcal{U} \qquad U\Delta \in \mathbf{\Delta} \qquad \Delta U \in \mathbf{\Delta} \qquad \overline{\sigma}(U\Delta) = \overline{\sigma}(\Delta U) = \overline{\sigma}(\Delta) \tag{11.11}$$

$$D\Delta = \Delta D. \tag{11.12}$$

Consequently,

Theorem 11.3 *For all $U \in \mathcal{U}$ and $D \in \mathcal{D}$*

$$\mu_{\mathbf{\Delta}}(MU) = \mu_{\mathbf{\Delta}}(UM) = \mu_{\mathbf{\Delta}}(M) = \mu_{\mathbf{\Delta}}\left(DMD^{-1}\right). \tag{11.13}$$

Proof. For all $D \in \mathcal{D}$ and $\Delta \in \mathbf{\Delta}$,

$$\det(I - M\Delta) = \det\left(I - MD^{-1}\Delta D\right) = \det\left(I - DMD^{-1}\Delta\right)$$

since D commutes with Δ. Therefore $\mu_{\mathbf{\Delta}}(M) = \mu_{\mathbf{\Delta}}\left(DMD^{-1}\right)$. Also, for each $U \in \mathcal{U}$, $\det(I - M\Delta) = 0$ if and only if $\det(I - MUU^*\Delta) = 0$. Since $U^*\Delta \in \mathbf{\Delta}$ and $\overline{\sigma}(U^*\Delta) = \overline{\sigma}(\Delta)$, we get $\mu_{\mathbf{\Delta}}(MU) = \mu_{\mathbf{\Delta}}(M)$ as desired. The argument for UM is the same. □

Therefore, the bounds in (11.8) can be tightened to

$$\max_{U\in\mathcal{U}} \rho(UM) \le \max_{\Delta \in \mathbf{B\Delta}} \rho(\Delta M) = \mu_{\mathbf{\Delta}}(M) \le \inf_{D\in\mathcal{D}} \overline{\sigma}\left(DMD^{-1}\right) \tag{11.14}$$

where the equality comes from Lemma 11.2. Note that the last element in the D matrix is normalized to 1 since for any nonzero scalar γ, $DMD^{-1} = (\gamma D)\, M\, (\gamma D)^{-1}$.

Remark 11.2 Note that the scaling set $\mathcal{D}$ in Theorem 11.3 and in the inequality (11.14) does not necessarily be restricted to Hermitian. In fact, they can be replaced by any set of nonsingular matrices that satisfy (11.12). However, enlarging the set of scaling matrices does not improve the upper bound in inequality (11.14). This can be shown as follows: Let D be any nonsingular matrix such that $D\Delta = \Delta D$. Then there exist a Hermitian matrix $0 < R = R^* \in \mathcal{D}$ and a unitary matrix U such that $D = UR$ and

$$\inf_D \overline{\sigma}\left(DMD^{-1}\right) = \inf_D \overline{\sigma}\left(URMR^{-1}U^*\right) = \inf_{R\in\mathcal{D}} \overline{\sigma}\left(RMR^{-1}\right).$$

Therefore, there is no loss of generality in assuming $\mathcal{D}$ to be Hermitian. ♡

11.2.3 Bounds

In this section we will concentrate on the bounds

$$\max_{U\in\mathcal{U}} \rho\left(UM\right) \leq \mu_{\mathbf{\Delta}}(M) \ \leq \inf_{D\in\mathcal{D}} \overline{\sigma}\left(DMD^{-1}\right).$$

The lower bound is always an equality [Doyle, 1982].

Theorem 11.4 $\max_{U\in\mathcal{U}} \rho(MU) = \mu(M)$.

Unfortunately, the quantity $\rho\left(UM\right)$ can have multiple local maxima which are not global. Thus local search cannot be guaranteed to obtain μ, but can only yield a lower bound. For computation purposes one can derive a slightly different formulation of the lower bound as a power algorithm which is reminiscent of power algorithms for eigenvalues and singular values [Packard, Fan, and Doyle, 1988]). While there are open questions about convergence, the algorithm usually works quite well and has proven to be an effective method to compute μ.

The upper bound can be reformulated as a convex optimization problem, so the global minimum can, in principle, be found. Unfortunately, the upper bound is not always equal to μ. For block structures $\mathbf{\Delta}$ satisfying $2S + F \leq 3$, the upper bound is always equal to $\mu_{\mathbf{\Delta}}(M)$, and for block structures with $2S + F > 3$, there exist matrices for which μ is less than the infimum. This can be summarized in the following diagram, which shows for which cases the upper bound is guaranteed to be equal to μ.

Theorem 11.5 $\mu(M) = \inf_{D\in\mathcal{D}} \overline{\sigma}(DMD^{-1})$ *if* $2S + F \leq 3$

F= S=	*0*	*1*	*2*	*3*	*4*
0		*yes*	*yes*	*yes*	*no*
1	*yes*	*yes*	*no*	*no*	*no*
2	*no*	*no*	*no*	*no*	*no*

Several of the boxes have connections with standard results.

- $S = 0$, $F = 1$: $\mu_{\mathbf{\Delta}}(M) = \overline{\sigma}(M)$.

- $S = 1$, $F = 0$: $\mu_{\mathbf{\Delta}}(M) = \rho(M) = \inf_{D \in \mathcal{D}} \overline{\sigma}\left(DMD^{-1}\right)$. This is a standard result in linear algebra. In fact, without a loss in generality, the matrix M can be assumed in Jordan Canonical form. Now let

$$J_1 = \begin{bmatrix} \lambda & 1 & & & \\ & \lambda & 1 & & \\ & & \ddots & \ddots & \\ & & & \lambda & 1 \\ & & & & \lambda \end{bmatrix}, \quad D_1 = \begin{bmatrix} 1 & & & & \\ & k & & & \\ & & \ddots & & \\ & & & k^{n_1-2} & \\ & & & & k^{n_1-1} \end{bmatrix} \in \mathbb{C}^{n_1 \times n_1}.$$

  Then $\inf_{D_1 \in \mathbb{C}^{n_1 \times n_1}} \overline{\sigma}(D_1 J_1 D_1^{-1}) = \lim_{k \to \infty} \overline{\sigma}(D_1 J_1 D_1^{-1}) = |\lambda|$. (Note that by Remark 11.2, the scaling matrix does not need to be Hermitian.) The conclusion follows by applying this result to each Jordan block.

  It is also equivalent to the fact that Lyapunov asymptotic stability and exponential stability are equivalent for discrete time systems. This is because $\rho(M) < 1$ (exponential stability of a discrete time system matrix M) implies for some nonsingular $D \in \mathbb{C}^{n \times n}$

$$\overline{\sigma}(DMD^{-1}) < 1 \quad \text{or} \quad (D^{-1})^* M^* D^* DMD^{-1} - I < 0$$

  which in turn is equivalent to the existence of a $P = D^* D > 0$ such that

$$M^* PM - P < 0$$

  (Lyapunov asymptotic stability).

- $S = 0$, $F = 2$: This case was studied by Redheffer [1959].

- $S = 1$, $F = 1$: This is equivalent to a state space characterization of the $\mathcal{H}_\infty$ norm of a discrete time transfer function, see Chapter 21.

- $S = 2$, $F = 0$: This is equivalent to the fact that for multidimensional systems (2-d, in fact), exponential stability is not equivalent to Lyapunov stability.

- $S = 0$, $F \geq 4$: For this case, the upper bound is not always equal to μ. This is important, as these are the cases that arise most frequently in applications. Fortunately, the bound seems to be close to μ. The worst known example has a ratio of μ over the bound of about .85, and most systems are close to 1.

The above bounds are much more than just computational schemes. They are also theoretically rich and can unify a number of apparently quite different results in linear systems theory. There are several connections with Lyapunov asymptotic stability, two of which were hinted at above, but there are further connections between the upper bound scalings and solutions to Lyapunov and Riccati equations. Indeed, many major theorems in linear systems theory follow from the upper bounds and from some results of Linear Fractional Transformations. The lower bound can be viewed as a natural generalization of the maximum modulus theorem.

Of course one of the most important uses of the upper bound is as a computational scheme when combined with the lower bound. For reliable use of the μ theory, it is essential to have upper and lower bounds. Another important feature of the upper bound is that it can be combined with H_∞ controller synthesis methods to yield an ad-hoc μ-synthesis method. Note that the upper bound when applied to transfer functions is simply a scaled H_∞ norm. This is exploited in the $D-K$ iteration procedure to perform approximate μ-synthesis (Doyle[1982]), which will be briefly introduced in section 11.4.

11.2.4 Well Posedness and Performance for Constant LFTs

Let M be a complex matrix partitioned as

$$M = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} \tag{11.15}$$

and suppose there are two defined block structures, $\mathbf{\Delta}_1$ and $\mathbf{\Delta}_2$, which are compatible in size with M_{11} and M_{22}, respectively. Define a third structure $\mathbf{\Delta}$ as

$$\mathbf{\Delta} = \left\{ \begin{bmatrix} \Delta_1 & 0 \\ 0 & \Delta_2 \end{bmatrix} : \Delta_1 \in \mathbf{\Delta}_1, \Delta_2 \in \mathbf{\Delta}_2 \right\}. \tag{11.16}$$

Now, we may compute μ with respect to three structures. The notations we use to keep track of these computations are as follows: $\mu_1(\cdot)$ is with respect to $\mathbf{\Delta}_1$, $\mu_2(\cdot)$ is with respect to $\mathbf{\Delta}_2$, and $\mu_{\mathbf{\Delta}}(\cdot)$ is with respect to $\mathbf{\Delta}$. In view of these notations, $\mu_1(M_{11})$, $\mu_2(M_{22})$ and $\mu_{\mathbf{\Delta}}(M)$ all make sense, though, for instance, $\mu_1(M)$ does not.

This section is interested in following constant matrix problems:

- determine whether the LFT $F_l(M, \Delta_2)$ is well posed for all $\Delta_2 \in \mathbf{\Delta}_2$ with $\overline{\sigma}(\Delta_2) \leq \beta$ $(< \beta)$, and,
- if so, then determine how "large" $F_l(M, \Delta_2)$ can get for this norm-bounded set of perturbations.

Let $\Delta_2 \in \mathbf{\Delta}_2$. Recall that $F_l(M, \Delta_2)$ is *well posed* if $I - M_{22}\Delta_2$ is invertible. The first theorem is nothing more than a restatement of the definition of μ.

Theorem 11.6 *The linear fractional transformation $F_l(M, \Delta_2)$ is well posed* *(a) for all* $\Delta_2 \in \mathbf{B\Delta}_2$ *if and only if* $\mu_2(M_{22}) < 1$.

(b) for all $\Delta_2 \in \mathbf{B}^\circ\mathbf{\Delta}_2$ *if and only if* $\mu_2(M_{22}) \leq 1$.

As the "perturbation" Δ_2 deviates from zero, the matrix $F_l(M, \Delta_2)$ deviates from M_{11}. The range of values that $\mu_1(F_l(M, \Delta_2))$ takes on is intimately related to $\mu_{\mathbf{\Delta}}(M)$, as shown in the following theorem:

Theorem 11.7 (MAIN LOOP THEOREM) *The following are equivalent:*

$$\mu_{\mathbf{\Delta}}(M) < 1 \qquad \Longleftrightarrow \qquad \begin{cases} \mu_2(M_{22}) < 1, \quad \text{and} \\ \\ \max\limits_{\Delta_2 \in \mathbf{B\Delta}_2} \mu_1(F_l(M, \Delta_2)) < 1. \end{cases}$$

$$\mu_{\mathbf{\Delta}}(M) \leq 1 \qquad \Longleftrightarrow \qquad \begin{cases} \mu_2(M_{22}) \leq 1, \quad \text{and} \\ \\ \max\limits_{\Delta_2 \in \mathbf{B}^\circ\mathbf{\Delta}_2} \mu_1(F_l(M, \Delta_2)) \leq 1. \end{cases}$$

Proof. We shall only prove the first part of the equivalence. The proof for the second part is similar.

$\Leftarrow$ Let $\Delta_i \in \mathbf{\Delta}_i$ be given, with $\overline{\sigma}(\Delta_i) \leq 1$, and define $\Delta = \text{diag}\;[\Delta_1, \Delta_2]$. Obviously $\Delta \in \mathbf{\Delta}$. Now

$$\det(I - M\Delta) = \det \begin{bmatrix} I - M_{11}\Delta_1 & -M_{12}\Delta_2 \\ -M_{21}\Delta_1 & I - M_{22}\Delta_2 \end{bmatrix}. \tag{11.17}$$

By hypothesis $I - M_{22}\Delta_2$ is invertible, and hence, $\det(I - M\Delta)$ becomes

$$\det(I - M_{22}\Delta_2) \det\left(I - M_{11}\Delta_1 - M_{12}\Delta_2 (I - M_{22}\Delta_2)^{-1} M_{21}\Delta_1\right).$$

Collecting the Δ_1 terms leaves

$$\det(I - M\Delta) = \det(I - M_{22}\Delta_2) \det(I - F_l(M, \Delta_2)\Delta_1). \tag{11.18}$$

But, $\mu_1(F_l(M, \Delta_2)) < 1$ and $\Delta_1 \in \mathbf{B\Delta}_1$, so $I - F_l(M, \Delta_2)\Delta_1$ must be nonsingular. Therefore, $I - M\Delta$ is nonsingular and, by definition, $\mu_{\mathbf{\Delta}}(M) < 1$.

$\Rightarrow$ Basically, the argument above is reversed. Again let $\Delta_1 \in \mathbf{B\Delta}_1$ and $\Delta_2 \in \mathbf{B\Delta}_2$ be given, and define $\Delta = \text{diag}\;[\Delta_1, \Delta_2]$. Then $\Delta \in \mathbf{B\Delta}$ and, by hypothesis, $\det(I - M\Delta) \neq 0$. It is easy to verify from the definition of μ that (always)

$$\mu(M) \geq \max\left\{\mu_1(M_{11}),\; \mu_2(M_{22})\right\}.$$

We can see that $\mu_2(M_{22}) < 1$, which gives that $I - M_{22}\Delta_2$ is also nonsingular. Therefore, the expression in (11.18) is valid, giving

$$\det(I - M_{22}\Delta_2) \det(I - F_l(M, \Delta_2)\Delta_1) = \det(I - M\Delta) \neq 0.$$

Obviously, $I - F_l(M, \Delta_2)\Delta_1$ is nonsingular for all $\Delta_i \in \mathbf{B\Delta}_i$, which indicates that the claim is true. $\square$

Remark 11.3 This theorem forms the basis for all uses of μ in linear system robustness analysis, whether from a state-space, frequency domain, or Lyapunov approach. $\heartsuit$

The role of the block structure $\mathbf{\Delta}_2$ in the MAIN LOOP theorem is clear - it is the structure that the perturbations come from; however, the role of the perturbation structure $\mathbf{\Delta}_1$ is often misunderstood. Note that $\mu_1(\cdot)$ appears on the right hand side of the theorem, so that the set $\mathbf{\Delta}_1$ defines what particular property of $F_l(M, \Delta_2)$ is considered. As an example, consider the theorem applied with the two simple block structures considered right after Lemma 11.2. Define $\mathbf{\Delta}_1 := \{\delta_1 I_n : \delta_1 \in \mathbb{C}\}$. Hence, for $A \in \mathbb{C}^{n\times n}$, $\mu_1(A) = \rho(A)$. Likewise, define $\mathbf{\Delta}_2 = \mathbb{C}^{m\times m}$; then for $D \in \mathbb{C}^{m\times m}$, $\mu_2(D) = \overline{\sigma}(D)$. Now, let $\mathbf{\Delta}$ be the diagonal augmentation of these two sets, namely

$$\mathbf{\Delta} := \left\{ \begin{bmatrix} \delta_1 I_n & 0_{n\times m} \\ 0_{m\times n} & \Delta_2 \end{bmatrix} : \delta_1 \in \mathbb{C}, \Delta_2 \in \mathbb{C}^{m\times m} \right\} \subset \mathbb{C}^{(n+m)\times(n+m)}.$$

Let $A \in \mathbb{C}^{n\times n}, B \in \mathbb{C}^{n\times m}, C \in \mathbb{C}^{m\times n}$, and $D \in \mathbb{C}^{m\times m}$ be given, and interpret them as the state space model of a discrete time system

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k \\ y_k &= Cx_k + Du_k. \end{aligned}$$

And let $M \in \mathbb{C}^{(n+m)\times(n+m)}$ be the block state space matrix of the system

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}.$$

Applying the theorem with this data gives that the following are equivalent:

- The spectral radius of A satisfies $\rho(A) < 1$, and
$$\max_{\substack{\delta_1 \in \mathbb{C} \\ |\delta_1| \le 1}} \overline{\sigma}\left(D + C\delta_1 (I - A\delta_1)^{-1} B\right) < 1. \tag{11.19}$$

- The maximum singular value of D satisfies $\overline{\sigma}(D) < 1$, and
$$\max_{\substack{\Delta_2 \in \mathbb{C}^{m\times m} \\ \overline{\sigma}(\Delta_2) \le 1}} \rho\left(A + B\Delta_2 (I - D\Delta_2)^{-1} C\right) < 1. \tag{11.20}$$

- The structured singular value of M satisfies
$$\mu_{\mathbf{\Delta}}(M) < 1. \tag{11.21}$$

The first condition is recognized by two things: the system is stable, and the $||\cdot||_\infty$ norm on the transfer function from u to y is less than 1 (by replacing δ_1 with $\frac{1}{z}$)

$$\|G\|_\infty := \max_{\substack{z\in\mathbb{C}\\|z|\geq 1}} \overline{\sigma}\left(D + C\left(zI-A\right)^{-1}B\right) = \max_{\substack{\delta_1\in\mathbb{C}\\|\delta_1|\leq 1}} \overline{\sigma}\left(D + C\delta_1\left(I - A\delta_1\right)^{-1}B\right).$$

The second condition implies that $(I - D\Delta_2)^{-1}$ is well defined for all $\overline{\sigma}(\Delta_2) \leq 1$ and that a robust stability result holds for the uncertain difference equation

$$x_{k+1} = \left(A + B\Delta_2\left(I - D\Delta_2\right)^{-1}C\right)x_k$$

where Δ_2 is any element in $\mathbb{C}^{m\times m}$ with $\overline{\sigma}(\Delta_2) \leq 1$, but otherwise unknown.

This equivalence between the small gain condition, $\|G\|_\infty < 1$, and the stability robustness of the uncertain difference equation is well known. This is the small gain theorem, in its necessary and sufficient form for linear, time invariant systems with one of the components norm-bounded, but otherwise unknown. What is important to note is that both of these conditions are equivalent to a condition involving the structured singular value of the state space matrix. Already we have seen that special cases of μ are the spectral radius and the maximum singular value. Here we see that other important linear system properties, namely robust stability and input-output gain, are also related to a particular case of the structured singular value.

11.3 Structured Robust Stability and Performance

11.3.1 Robust Stability

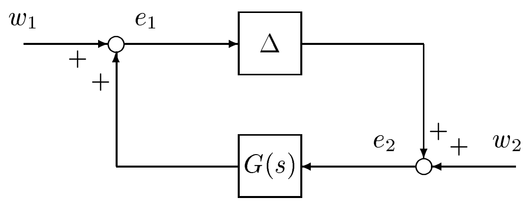
The most well-known use of μ as a robustness analysis tool is in the frequency domain. Suppose $G(s)$ is a stable, multi-input, multi-output transfer function of a linear system. For clarity, assume G has q_1 inputs and p_1 outputs. Let $\mathbf{\Delta}$ be a block structure, as in equation (11.3), and assume that the dimensions are such that $\mathbf{\Delta} \subset \mathbb{C}^{q_1\times p_1}$. We want to consider feedback perturbations to G which are themselves dynamical systems with the block-diagonal structure of the set $\mathbf{\Delta}$.

Let $\mathcal{M}(\mathbf{\Delta})$ denote the set of all block diagonal and stable rational transfer functions that have block structures such as $\mathbf{\Delta}$.

$$\mathcal{M}(\mathbf{\Delta}) := \left\{\Delta(\cdot) \in \mathcal{RH}_\infty : \Delta(s_o) \in \mathbf{\Delta} \text{ for all } s_o \in \overline{\mathbb{C}}_+\right\}$$

Theorem 11.8 *Let $\beta > 0$. The loop shown below is well-posed and internally stable for all $\Delta(\cdot) \in \mathcal{M}(\mathbf{\Delta})$ with $\|\Delta\|_\infty < \frac{1}{\beta}$ if and only if*

$$\sup_{\omega\in\mathbb{R}} \mu_{\mathbf{\Delta}}(G(j\omega)) \leq \beta$$



Proof. ($\Longleftarrow$) By Lemma 11.1, $\sup_{s \in \overline{\mathbb{C}}_+} \mu_{\mathbf{\Delta}}(G(s)) = \sup_{\omega \in \mathbb{R}} \mu_{\mathbf{\Delta}}(G(j\omega)) \leq \beta$. Hence $\det(I - G(s)\Delta(s)) \neq 0$ for all $s \in \overline{\mathbb{C}}_+ \cup \{\infty\}$ whenever $\|\Delta\|_\infty < 1/\beta$, i.e., the system is robustly stable.

($\Longrightarrow$) Suppose $\sup_{\omega \in \mathbb{R}} \mu_{\mathbf{\Delta}}(G(j\omega)) > \beta$. Then there is a $0 < \omega_o < \infty$ such that $\mu_{\mathbf{\Delta}}(G(j\omega_o)) > \beta$. By Remark 11.1, there is a complex $\Delta_c \in \mathbf{\Delta}$ that each full block has rank 1 and $\overline{\sigma}(\Delta_c) < 1/\beta$ such that $I - G(j\omega_o)\Delta_c$ is singular. Next, using the same construction used in the proof of the small gain theorem (Theorem 9.1), one can find a rational $\Delta(s)$ such that $\|\Delta(s)\|_\infty = \overline{\sigma}(\Delta_c) < 1/\beta$, $\Delta(j\omega_o) = \Delta_c$, and $\Delta(s)$ destabilizes the system. □

Hence, the peak value on the μ plot of the frequency response determines the size of perturbations that the loop is robustly stable against.

Remark 11.4 The internal stability with closed ball of uncertainties is more complicated. The following example is shown in Tits and Fan [1994]. Consider

$$G(s) = \frac{1}{s+1} \begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}$$

and $\Delta = \delta(s) I_2$. Then

$$\sup_{\omega \in \mathbb{R}} \mu_{\mathbf{\Delta}}(G(j\omega)) = \sup_{\omega \in \mathbb{R}} \frac{1}{|j\omega + 1|} = \mu_{\mathbf{\Delta}}(G(j0)) = 1.$$

On the other hand, $\mu_{\mathbf{\Delta}}(G(s)) < 1$ for all $s \neq 0, s \in \overline{\mathbb{C}}_+$, and the only matrices in the form of $\Gamma = \gamma I_2$ with $|\gamma| \leq 1$ for which

$$\det(I - G(0)\Gamma) = 0$$

are the *complex* matrices $\pm j I_2$. Thus, clearly, $(I - G(s)\Delta(s))^{-1} \in \mathcal{RH}_\infty$ for all real rational $\Delta(s) = \delta(s) I_2$ with $\|\delta\|_\infty \leq 1$ since $\Delta(0)$ must be real. This shows that $\sup_{\omega \in \mathbb{R}} \mu_{\mathbf{\Delta}}(G(j\omega)) < 1$ is not necessary for $(I - G(s)\Delta(s))^{-1} \in \mathcal{RH}_\infty$ with the closed ball of structured uncertainty $\|\Delta\|_\infty \leq 1$. Similar examples with no repeated blocks are generated by setting $G(s) = \frac{1}{s+1} M$ where M is any real matrix with $\mu_{\mathbf{\Delta}}(M) = 1$ for which there is no real $\Delta \in \mathbf{\Delta}$ with $\overline{\sigma}(\Delta) = 1$ such that $\det(I - M\Delta) = 0$. For example, let

$$M = \begin{bmatrix} 0 & \beta \\ \gamma & \alpha \\ \gamma & -\alpha \end{bmatrix} \begin{bmatrix} -\beta & \alpha & \alpha \\ 0 & -\gamma & \gamma \end{bmatrix}, \quad \mathbf{\Delta} = \left\{ \begin{bmatrix} \delta_1 & & \\ & \delta_2 & \\ & & \delta_3 \end{bmatrix}, \ \delta_i \in \mathbb{C} \right\}$$

with $\gamma^2 = \frac{1}{2}$ and $\beta^2 + 2\alpha^2 = 1$. Then it is shown in Packard and Doyle [1993] that $\mu_{\mathbf{\Delta}}(M) = 1$ and all $\Delta \in \mathbf{\Delta}$ with $\overline{\sigma}(\Delta) = 1$ that satisfy $\det(I - M\Delta) = 0$ must be complex. ♡

Remark 11.5 Let $\Delta \in \mathcal{RH}_\infty$ be a structured uncertainty and

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \in \mathcal{RH}_\infty$$

then $F_u(G, \Delta) \in \mathcal{RH}_\infty$ does not necessarily imply $(I - G_{11}\Delta)^{-1} \in \mathcal{RH}_\infty$ whether Δ is in an open ball or is in a closed ball. For example, consider

$$G(s) = \left[\begin{array}{cc|c} \frac{1}{s+1} & 0 & 1 \\ 0 & \frac{10}{s+1} & 0 \\ \hline 1 & 0 & 0 \end{array}\right]$$

and $\Delta = \begin{bmatrix} \delta_1 & \\ & \delta_2 \end{bmatrix}$ with $\|\Delta\|_\infty < 1$. Then $F_u(G, \Delta) = \dfrac{1}{1 - \delta_1 \frac{1}{s+1}} \in \mathcal{RH}_\infty$ for all admissible Δ ($\|\Delta\|_\infty < 1$) but $(I - G_{11}\Delta)^{-1} \in \mathcal{RH}_\infty$ is true only for $\|\Delta\|_\infty < 0.1$. ♡

11.3.2 Robust Performance

Often, stability is not the only property of a closed-loop system that must be robust to perturbations. Typically, there are exogenous disturbances acting on the system (wind gusts, sensor noise) which result in tracking and regulation errors. Under perturbation, the effect that these disturbances have on error signals can greatly increase. In most cases, long before the onset of instability, the closed-loop performance will degrade to the point of unacceptability, hence the need for a "robust performance" test. Such a test will indicate the worst-case level of performance degradation associated with a given level of perturbations.

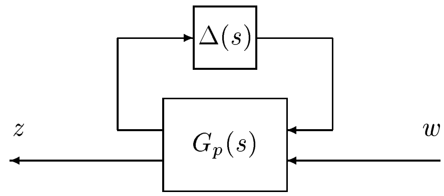
Assume G_p is a stable, real-rational, proper transfer function with $q_1 + q_2$ inputs and $p_1 + p_2$ outputs. Partition G_p in the obvious manner

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

so that G_{11} has q_1 inputs and p_1 outputs, and so on. Let $\mathbf{\Delta} \subset \mathbb{C}^{q_1 \times p_1}$ be a block structure, as in equation (11.3). Define an augmented block structure

$$\mathbf{\Delta}_P := \left\{ \begin{bmatrix} \Delta & 0 \\ 0 & \Delta_f \end{bmatrix} : \Delta \in \mathbf{\Delta}, \Delta_f \in \mathbb{C}^{q_2 \times p_2} \right\}.$$

The setup is to theoretically address the robust performance questions about the loop shown below



The transfer function from w to z is denoted by $F_u(G_p, \Delta)$.

Theorem 11.9 *Let $\beta > 0$. For all $\Delta(s) \in \mathcal{M}(\mathbf{\Delta})$ with $\|\Delta\|_\infty < \frac{1}{\beta}$, the loop shown above is well-posed, internally stable, and $\|F_u(G_p, \Delta)\|_\infty \leq \beta$ if and only if*

$$\sup_{\omega \in \mathbb{R}} \mu_{\mathbf{\Delta}_P}(G_p(j\omega)) \leq \beta.$$

Note that by internal stability, $\sup_{\omega \in \mathbb{R}} \mu_{\mathbf{\Delta}}(G_{11}(j\omega)) \leq \beta$, then the proof of this theorem is exactly along the lines of the earlier proof for Theorem 11.8, but also appeals to Theorem 11.7. This is a remarkably useful theorem. It says that a robust performance problem is equivalent to a robust stability problem with augmented uncertainty Δ as shown in Figure 11.5.

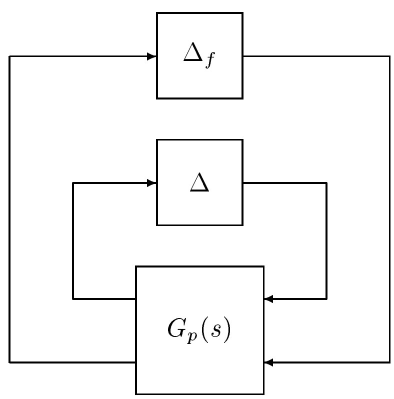


Figure 11.5: Robust Performance vs Robust Stability

11.3.3 Two Block μ: Robust Performance Revisited

Suppose that the uncertainty block is given by

$$\Delta = \begin{bmatrix} \Delta_1 & \\ & \Delta_2 \end{bmatrix} \in \mathcal{RH}_\infty$$

with $\|\Delta\|_\infty < 1$ and that the interconnection model G is given by

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} \in \mathcal{RH}_\infty.$$

Then the closed-loop system is well-posed and internally stable iff $\sup_\omega \mu_{\mathbf{\Delta}}(G(j\omega)) \leq 1$. Let

$$D_\omega = \begin{bmatrix} d_\omega I & \\ & I \end{bmatrix}, \quad d_\omega \in \mathbb{R}_+$$

then

$$D_\omega G(j\omega) D_\omega^{-1} = \begin{bmatrix} G_{11}(j\omega) & d_\omega G_{12}(j\omega) \\ \frac{1}{d_\omega} G_{21}(j\omega) & G_{22}(j\omega) \end{bmatrix}.$$

Hence by Theorem 11.5, at each frequency ω

$$\mu_{\mathbf{\Delta}}(G(j\omega)) = \inf_{d_\omega \in \mathbb{R}_+} \overline{\sigma} \left(\begin{bmatrix} G_{11}(j\omega) & d_\omega G_{12}(j\omega) \\ \frac{1}{d_\omega} G_{21}(j\omega) & G_{22}(j\omega) \end{bmatrix} \right). \qquad (11.22)$$

Since the minimization is convex in $\log d_\omega$ [see, Doyle, 1982], the optimal d_ω can be found by a search; however, two approximations to d_ω can be obtained easily by approximating the right hand side of (11.22):

(1) From Lemma 2.10, we have

$$\mu_{\mathbf{\Delta}}(G(j\omega)) \leq \inf_{d_\omega \in \mathbb{R}_+} \overline{\sigma} \left(\begin{bmatrix} \|G_{11}(j\omega)\| & d_\omega \|G_{12}(j\omega)\| \\ \frac{1}{d_\omega} \|G_{21}(j\omega)\| & \|G_{22}(j\omega)\| \end{bmatrix} \right)$$

$$\begin{aligned} &\leq \sqrt{\inf_{d_\omega \in \mathbb{R}_+} \left(\|G_{11}(j\omega)\|^2 + d_\omega^2 \|G_{12}(j\omega)\|^2 + \frac{1}{d_\omega^2} \|G_{21}(j\omega)\|^2 + \|G_{22}(j\omega)\|^2 \right)} \\ &= \sqrt{\|G_{11}(j\omega)\|^2 + \|G_{22}(j\omega)\|^2 + 2\|G_{12}(j\omega)\| \|G_{21}(j\omega)\|} \end{aligned}$$

with the minimizing d_ω given by

$$\hat{d}_\omega = \begin{cases} \sqrt{\frac{\|G_{21}(j\omega)\|}{\|G_{12}(j\omega)\|}} & \text{if } G_{12} \neq 0 \ \& \ G_{21} \neq 0, \\ 0 & \text{if } G_{21} = 0, \\ \infty & \text{if } G_{12} = 0. \end{cases} \qquad (11.23)$$

(2) Alternative approximation can be obtained by using the Frobenius norm

$$\mu_{\mathbf{\Delta}}(G(j\omega)) \leq \inf_{d_\omega \in \mathbb{R}_+} \left\| \begin{bmatrix} G_{11}(j\omega) & d_\omega G_{12}(j\omega) \\ \frac{1}{d_\omega} G_{21}(j\omega) & G_{22}(j\omega) \end{bmatrix} \right\|_F$$

$$= \sqrt{\inf_{d_\omega \in \mathbb{R}_+} \left(\|G_{11}(j\omega)\|_F^2 + d_\omega^2 \|G_{12}(j\omega)\|_F^2 + \frac{1}{d_\omega^2} \|G_{21}(j\omega)\|_F^2 + \|G_{22}(j\omega)\|_F^2 \right)}$$

$$= \sqrt{\|G_{11}(j\omega)\|_F^2 + \|G_{22}(j\omega)\|_F^2 + 2\|G_{12}(j\omega)\|_F \|G_{21}(j\omega)\|_F}$$

with the minimizing d_ω given by

$$\tilde{d}_\omega = \begin{cases} \sqrt{\frac{\|G_{21}(j\omega)\|_F}{\|G_{12}(j\omega)\|_F}} & \text{if } G_{12} \neq 0 \ \& \ G_{21} \neq 0, \\ 0 & \text{if } G_{21} = 0, \\ \infty & \text{if } G_{12} = 0. \end{cases} \tag{11.24}$$

It can be shown that the approximations for the scalar d_ω obtained above are exact for a 2×2 matrix G. For higher dimensional G, the approximations for d_ω are still reasonably good. Hence an approximation of μ can be obtained as

$$\mu_{\mathbf{\Delta}}(G(j\omega)) \leq \overline{\sigma}\left(\begin{bmatrix} G_{11}(j\omega) & \hat{d}_\omega G_{12}(j\omega) \\ \frac{1}{\hat{d}_\omega} G_{21}(j\omega) & G_{22}(j\omega) \end{bmatrix} \right) \tag{11.25}$$

or, alternatively, as

$$\mu_{\mathbf{\Delta}}(G(j\omega)) \leq \overline{\sigma}\left(\begin{bmatrix} G_{11}(j\omega) & \tilde{d}_\omega G_{12}(j\omega) \\ \frac{1}{\tilde{d}_\omega} G_{21}(j\omega) & G_{22}(j\omega) \end{bmatrix} \right). \tag{11.26}$$

We can now see how these approximated μ tests are compared with the sufficient conditions obtained in Chapter 9.

Example 11.1 Consider again the robust performance problem of a system with output multiplicative uncertainty in Chapter 9 (see Figure 9.5):

$$P_\Delta = (I + W_1 \Delta W_2)P, \quad \|\Delta\|_\infty < 1.$$

Then it is easy to show that the problem can be put in the general framework by selecting

$$G(s) = \begin{bmatrix} -W_2 T_o W_1 & -W_2 T_o W_d \\ W_e S_o W_1 & W_e S_o W_d \end{bmatrix}$$

and that the robust performance condition is satisfied if and only if

$$\|W_2 T_o W_1\|_\infty \leq 1 \tag{11.27}$$

and

$$\|\mathcal{F}_u(G, \Delta)\|_\infty \leq 1 \tag{11.28}$$

for all $\Delta \in \mathcal{RH}_\infty$ with $\|\Delta\|_\infty < 1$. But (11.27) and (11.28) are satisfied iff for each frequency ω

$$\mu_{\mathbf{\Delta}}(G(j\omega)) = \inf_{d_\omega \in \mathbb{R}_+} \overline{\sigma}\left(\begin{bmatrix} -W_2T_oW_1 & -d_\omega W_2T_oW_d \\ \frac{1}{d_\omega}W_eS_oW_1 & W_eS_oW_d \end{bmatrix}\right) \leq 1.$$

Note that, in contrast to the sufficient condition obtained in Chapter 9, this condition is an exact test for robust performance. To compare the μ test with the criteria obtained in Chapter 9, some upper bounds for μ can be derived. Let

$$d_\omega = \sqrt{\frac{\|W_eS_oW_1\|}{\|W_2T_oW_d\|}}.$$

Then, using the first approximation for μ, we get

$$\begin{aligned}
\mu_{\mathbf{\Delta}}(G(j\omega)) &\leq \sqrt{\|W_2T_oW_1\|^2 + \|W_eS_oW_d\|^2 + 2\,\|W_2T_oW_d\|\,\|W_eS_oW_1\|} \\
&\leq \sqrt{\|W_2T_oW_1\|^2 + \|W_eS_oW_d\|^2 + 2\kappa(W_1^{-1}W_d)\,\|W_2T_oW_1\|\,\|W_eS_oW_d\|} \\
&\leq \|W_2T_oW_1\| + \kappa(W_1^{-1}W_d)\,\|W_eS_oW_d\|
\end{aligned}$$

where W_1 is assumed to be invertible in the last two inequalities. The last term is exactly the sufficient robust performance criteria obtained in Chapter 9. It is clear that any term preceding the last forms a tighter test since $\kappa(W_1^{-1}W_d) \geq 1$. Yet another alternative sufficient test can be obtained from the above sequence of inequalities:

$$\mu_{\mathbf{\Delta}}(G(j\omega)) \leq \sqrt{\kappa(W_1^{-1}W_d)}(\|W_2T_oW_1\| + \|W_eS_oW_d\|).$$

Note that this sufficient condition is not easy to get from the approach taken in Chapter 9 and is potentially less conservative than the bounds derived there. ◇

Next we consider the skewed specification problem, but first the following lemma is needed in the sequel.

Lemma 11.10 *Suppose* $\overline{\sigma} = \sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_m = \underline{\sigma} > 0$, *then*

$$\inf_{d \in \mathbb{R}_+} \max_i \left\{(d\sigma_i)^2 + \frac{1}{(d\sigma_i)^2}\right\} = \frac{\overline{\sigma}}{\underline{\sigma}} + \frac{\underline{\sigma}}{\overline{\sigma}}.$$

Proof. Consider a function $y = x + 1/x$, then y is a convex function and the maximization over a closed interval is achieved at the boundary of the interval. Hence for any fixed d

$$\max_i \left\{(d\sigma_i)^2 + \frac{1}{(d\sigma_i)^2}\right\} = \max\left\{(d\overline{\sigma})^2 + \frac{1}{(d\overline{\sigma})^2},\ \ (d\underline{\sigma})^2 + \frac{1}{(d\underline{\sigma})^2}\right\}.$$

Then the minimization over d is obtained iff

$$(d\overline{\sigma})^2 + \frac{1}{(d\overline{\sigma})^2} = (d\underline{\sigma})^2 + \frac{1}{(d\underline{\sigma})^2}$$

which gives $d^2 = \frac{1}{\overline{\sigma}\,\underline{\sigma}}$. The result then follows from substituting d. $\square$

Example 11.2 As another example, consider again the skewed specification problem from Chapter 9. Then the corresponding G matrix is given by

$$G = \begin{bmatrix} -W_2 T_i W_1 & -W_2 K S_o W_d \\ W_e S_o P W_1 & W_e S_o W_d \end{bmatrix}.$$

So the robust performance specification is satisfied iff

$$\mu_{\mathbf{\Delta}}(G(j\omega)) = \inf_{d_\omega \in \mathbb{R}_+} \overline{\sigma}\left(\begin{bmatrix} -W_2 T_i W_1 & -d_\omega W_2 K S_o W_d \\ \frac{1}{d_\omega} W_e S_o P W_1 & W_e S_o W_d \end{bmatrix}\right) \leq 1$$

for all $\omega \geq 0$. As in the last example, an upper bound can be obtained by taking

$$d_\omega = \sqrt{\frac{\|W_e S_o P W_1\|}{\|W_2 K S_o W_d\|}}.$$

Then

$$\mu_{\mathbf{\Delta}}(G(j\omega)) \leq \sqrt{\kappa(W_d^{-1} P W_1)}(\|W_2 T_i W_1\| + \|W_e S_o W_d\|).$$

In particular, this suggests that the robust performance margin is inversely proportional to the square root of the plant condition number if $W_d = I$ and $W_1 = I$. This can be further illustrated by considering a plant-inverting control system.

To simplify the exposition, we shall make the following assumptions:

$$W_e = w_s I,\; W_d = I,\; W_1 = I,\; W_2 = w_t I,$$

and P is stable and has a stable inverse (i.e., minimum phase) (P can be strictly proper). Furthermore, we shall assume that the controller has the form

$$K(s) = P^{-1}(s) l(s)$$

where $l(s)$ is a scalar loop transfer function which makes $K(s)$ proper and stabilizes the closed-loop. This compensator produces diagonal sensitivity and complementary sensitivity functions with identical diagonal elements, namely

$$S_o = S_i = \frac{1}{1+l(s)} I, \quad T_o = T_i = \frac{l(s)}{1+l(s)} I.$$

Denote

$$\varepsilon(s) = \frac{1}{1+l(s)}, \quad \tau(s) = \frac{l(s)}{1+l(s)}$$

and substitute these expressions into G; we get

$$G = \begin{bmatrix} -w_t\tau I & -w_t\tau P^{-1} \\ w_s\varepsilon P & w_s\varepsilon I \end{bmatrix}.$$

The structured singular value for G at frequency ω can be computed by

$$\mu_{\mathbf{\Delta}}(G(j\omega)) = \inf_{d\in\mathbb{R}_+} \overline{\sigma}\left(\begin{bmatrix} -w_t\tau I & -w_t\tau(dP)^{-1} \\ w_s\varepsilon dP & w_s\varepsilon I \end{bmatrix}\right).$$

Let the singular value decomposition of $P(j\omega)$ at frequency ω be

$$P(j\omega) = U\Sigma V^*, \quad \Sigma = \mathrm{diag}(\sigma_1, \sigma_2, \ldots, \sigma_m)$$

with $\sigma_1 = \overline{\sigma}$ and $\sigma_m = \underline{\sigma}$ where m is the dimension of P. Then

$$\mu_{\mathbf{\Delta}}(G(j\omega)) = \inf_{d\in\mathbb{R}_+} \overline{\sigma}\left(\begin{bmatrix} -w_t\tau I & -w_t\tau(d\Sigma)^{-1} \\ w_s\varepsilon d\Sigma & w_s\varepsilon I \end{bmatrix}\right)$$

since unitary operations do not change the singular values of a matrix. Note that

$$\begin{bmatrix} -w_t\tau I & -w_t\tau(d\Sigma)^{-1} \\ w_s\varepsilon d\Sigma & w_s\varepsilon I \end{bmatrix} = P_1 \mathrm{diag}(M_1, M_2, \ldots, M_m) P_2$$

where P_1 and P_2 are permutation matrices and where

$$M_i = \begin{bmatrix} -w_t\tau & -w_t\tau(d\sigma_i)^{-1} \\ w_s\varepsilon d\sigma_i & w_s\varepsilon \end{bmatrix}.$$

Hence

$$\begin{aligned}
\mu_{\mathbf{\Delta}}(G(j\omega)) &= \inf_{d\in\mathbb{R}_+} \max_i \overline{\sigma}\left(\begin{bmatrix} -w_t\tau & -w_t\tau(d\sigma_i)^{-1} \\ w_s\varepsilon d\sigma_i & w_s\varepsilon \end{bmatrix}\right) \\
&= \inf_{d\in\mathbb{R}_+} \max_i \overline{\sigma}\left(\begin{bmatrix} -w_t\tau \\ w_s\varepsilon d\sigma_i \end{bmatrix} \begin{bmatrix} 1 & (d\sigma_i)^{-1} \end{bmatrix}\right) \\
&= \inf_{d\in\mathbb{R}_+} \max_i \sqrt{(1+|d\sigma_i|^{-2})(|w_s\varepsilon d\sigma_i|^2 + |w_t\tau|^2)} \\
&= \inf_{d\in\mathbb{R}_+} \max_i \sqrt{|w_s\varepsilon|^2 + |w_t\tau|^2 + |w_s\varepsilon d\sigma_i|^2 + \left|\frac{w_t\tau}{d\sigma_i}\right|^2}.
\end{aligned}$$

Using Lemma 11.10, it is easy to show that the maximum is achieved at either $\overline{\sigma}$ or $\underline{\sigma}$ and that optimal d is given by

$$d^2 = \frac{|w_t\tau|}{|w_s\varepsilon|\underline{\sigma}\overline{\sigma}},$$

so the structured singular value is

$$\mu_{\mathbf{\Delta}}(G(j\omega)) = \sqrt{|w_s\varepsilon|^2 + |w_t\tau|^2 + |w_s\varepsilon||w_t\tau|[\kappa(P) + \frac{1}{\kappa(P)}]}. \tag{11.29}$$

Note that if $|w_s\varepsilon|$ and $|w_t\tau|$ are not too large, which are guaranteed if the nominal performance and robust stability conditions are satisfied, then the structured singular value is proportional to the square root of the plant condition number:

$$\mu_{\mathbf{\Delta}}(G(j\omega)) \approx \sqrt{|w_s\varepsilon||w_t\tau|\kappa(P)}\ . \tag{11.30}$$

◇

This confirms our intuition that an ill-conditioned plant with skewed specifications is hard to control.

11.3.4 Approximation of Multiple Full Block μ

The approximations given in the last subsection can be generalized to the multiple block μ problem by assuming that M is partitioned consistently with the structure of

$$\Delta = \text{diag}(\Delta_1, \Delta_2, \ldots, \Delta_F)$$

so that

$$M = \begin{bmatrix} M_{11} & M_{12} & \cdots & M_{1F} \\ M_{21} & M_{22} & \cdots & M_{2F} \\ \vdots & \vdots & & \vdots \\ M_{F1} & M_{F2} & \cdots & M_{FF} \end{bmatrix}$$

and

$$D = \text{diag}(d_1 I, \ldots, d_{F-1} I, I).$$

Now

$$DMD^{-1} = \left[M_{ij}\frac{d_i}{d_j}\right],\ d_F := 1.$$

And hence

$$\mu_{\mathbf{\Delta}}(M) \quad \leq \quad \inf_{D\in\mathcal{D}} \overline{\sigma}(DMD^{-1}) = \inf_{D\in\mathcal{D}} \overline{\sigma}\left[M_{ij}\frac{d_i}{d_j}\right]$$

$$\leq \inf_{D\in\mathcal{D}} \overline{\sigma}\left[\|M_{ij}\| \frac{d_i}{d_j}\right] \leq \inf_{D\in\mathcal{D}} \sqrt{\sum_{i=1}^{F}\sum_{j=1}^{F} \|M_{ij}\|^2 \frac{d_i^2}{d_j^2}}$$
$$\leq \inf_{D\in\mathcal{D}} \sqrt{\sum_{i=1}^{F}\sum_{j=1}^{F} \|M_{ij}\|_F^2 \frac{d_i^2}{d_j^2}}.$$

An approximate D can be found by solving the following minimization problem:

$$\inf_{D\in\mathcal{D}} \sum_{i=1}^{F}\sum_{j=1}^{F} \|M_{ij}\|^2 \frac{d_i^2}{d_j^2}$$

or, more conveniently, by minimizing

$$\inf_{D\in\mathcal{D}} \sum_{i=1}^{F}\sum_{j=1}^{F} \|M_{ij}\|_F^2 \frac{d_i^2}{d_j^2}$$

with $d_F = 1$. The optimal d_i minimizing of the above two problems satisfy, respectively,

$$d_k^4 = \frac{\sum_{i\neq k} \|M_{ik}\|^2 d_i^2}{\sum_{j\neq k} \|M_{kj}\|^2 / d_j^2}, \quad k = 1, 2, \ldots, F-1 \tag{11.31}$$

and

$$d_k^4 = \frac{\sum_{i\neq k} \|M_{ik}\|_F^2 d_i^2}{\sum_{j\neq k} \|M_{kj}\|_F^2 / d_j^2}, \quad k = 1, 2, \ldots, F-1. \tag{11.32}$$

Using these relations, d_k can be obtained by iterations.

Example 11.3 Consider a 3×3 complex matrix

$$M = \begin{bmatrix} 1+j & 10-2j & -20j \\ 5j & 3+j & -1+3j \\ -2 & j & 4-j \end{bmatrix}$$

with structured $\Delta = \mathrm{diag}(\delta_1, \delta_2, \delta_3)$. The largest singular value of M is $\overline{\sigma}(M) = 22.9094$ and the structured singular value of M computed using the μ-TOOLS is equal to its upper bound:

$$\mu_{\mathbf{\Delta}}(M) = \inf_{D\in\mathcal{D}} \overline{\sigma}(DMD^{-1}) = 11.9636$$

with the optimal scaling $D_{opt} = \mathrm{diag}(0.3955, 0.6847, 1)$. The optimal D minimizing

$$\inf_{D\in\mathcal{D}} \sum_{i=1}^{F}\sum_{j=1}^{F} \|M_{ij}\|^2 \frac{d_i^2}{d_j^2}$$

is $D_{subopt} = \text{diag}(0.3212, 0.4643, 1)$ which is solved from equation (11.31). Using this D_{subopt}, we obtain another upper bound for the structured singular value:

$$\mu_{\mathbf{\Delta}}(M) \leq \overline{\sigma}(D_{subopt} M D_{subopt}^{-1}) = 12.2538.$$

One may also use this D_{subopt} as an initial guess for the exact optimization. $\diamond$

11.4 Overview on μ Synthesis

This section briefly outlines various synthesis methods. The details are somewhat complicated and are treated in the other parts of the book. At this point, we simply want to point out how the analysis theory discussed in the previous sections leads naturally to synthesis questions.

From the analysis results, we see that each case eventually leads to the evaluation of

$$\|M\|_{\alpha} \qquad \alpha = 2,\ \infty,\ \text{or}\ \ \mu \tag{11.33}$$

for some transfer matrix M. Thus when the controller is put back into the problem, it involves only a simple linear fractional transformation as shown in Figure 11.6 with

$$M = \mathcal{F}_\ell(G, K) = G_{11} + G_{12}K(I - G_{22}K)^{-1}G_{21}.$$

where $G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}$ is chosen, respectively, as

- nominal performance only ($\Delta = 0$): $G = \begin{bmatrix} P_{22} & P_{23} \\ P_{32} & P_{33} \end{bmatrix}$
- robust stability only: $G = \begin{bmatrix} P_{11} & P_{13} \\ P_{31} & P_{33} \end{bmatrix}$
- robust performance: $G = P = \left[\begin{array}{cc:c} P_{11} & P_{12} & P_{13} \\ P_{21} & P_{22} & P_{23} \\ \hdashline P_{31} & P_{32} & P_{33} \end{array}\right]$.

Each case then leads to the synthesis problem

$$\min_K \|\mathcal{F}_\ell(G, K)\|_{\alpha} \quad \text{for } \alpha = 2,\ \infty,\ \text{or}\ \ \mu \tag{11.34}$$

which is subject to the internal stability of the nominal.

The solutions of these problems for $\alpha = 2$ and ∞ are the focus of the rest of this book. The solutions presented in this book unify the two approaches in a common synthesis framework. The $\alpha = 2$ case was already known in the 1960's, and the results

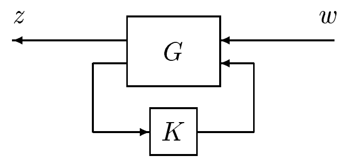


Figure 11.6: Synthesis Framework

are simply a new interpretation. The two Riccati solutions for the $\alpha = \infty$ case were new products of the late 1980's.

The synthesis for the $\alpha = \mu$ case is not yet fully solved. Recall that μ may be obtained by scaling and applying $\|\cdot\|_\infty$ (for $F \leq 3$ and $S = 0$), a reasonable approach is to "solve"

$$\min_K \inf_{D, D^{-1} \in \mathcal{H}_\infty} \left\| D \mathcal{F}_\ell(G, K) D^{-1} \right\|_\infty \tag{11.35}$$

by iteratively solving for K and D. This is the so-called *D-K Iteration*. The stable and minimum phase scaling matrix $D(s)$ is chosen such that $D(s)\Delta(s) = \Delta(s)D(s)$ (Note that $D(s)$ is not necessarily belong to $\mathcal{D}$ since $D(s)$ is not necessarily Hermitian, see Remark 11.2). For a fixed scaling transfer matrix D, $\min_K \left\| D\mathcal{F}_\ell(G,K)D^{-1} \right\|_\infty$ is a standard $\mathcal{H}_\infty$ optimization problem which will be solved in the later part of the book. For a given stabilizing controller K, $\inf_{D,D^{-1}\in\mathcal{H}_\infty} \left\| D\mathcal{F}_\ell(G,K)D^{-1} \right\|_\infty$ is a standard convex optimization problem and it can be solved pointwise in the frequency domain:

$$\sup_\omega \inf_{D_\omega \in \mathcal{D}} \overline{\sigma} \left[D_\omega \mathcal{F}_\ell(G,K)(j\omega) D_\omega^{-1} \right].$$

Indeed,

$$\inf_{D, D^{-1} \in \mathcal{H}_\infty} \left\| D\mathcal{F}_\ell(G,K)D^{-1} \right\|_\infty = \sup_\omega \inf_{D_\omega \in \mathcal{D}} \overline{\sigma} \left[D_\omega \mathcal{F}_\ell(G,K)(j\omega) D_\omega^{-1} \right].$$

This follows intuitively from the following arguments: the left hand side is always no smaller than the right hand side, and, on the other hand, given the minimizing D_ω from the right hand side across the frequency, there is always a rational function $D(s)$ uniformly approximating the magnitude frequency response D_ω.

Note that when $S = 0$, (no scalar blocks)

$$D_\omega = \operatorname{diag}(d_1^\omega I, \ldots, d_{F-1}^\omega I, I) \in \mathcal{D},$$

which is a block-diagonal scaling matrix applied pointwise across frequency to the frequency response $\mathcal{F}_\ell(G,K)(j\omega)$.

D-K Iterations proceed by performing this two-parameter minimization in sequential fashion: first minimizing over K with D_ω fixed, then minimizing pointwise over D_ω with K fixed, then again over K, and again over D_ω, etc. Details of this process are summarized in the following steps:

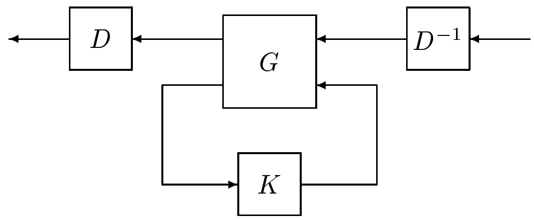


Figure 11.7: μ-Synthesis via Scaling

(i) Fix an initial estimate of the scaling matrix $D_\omega \in \mathcal{D}$ pointwise across frequency.

(ii) Find scalar transfer functions $d_i(s), d_i^{-1}(s) \in \mathcal{RH}_\infty$ for $i = 1, 2, \ldots, (F-1)$ such that $|d_i(j\omega)| \approx d_i^\omega$. This step can be done using the interpolation theory [Youla and Saito, 1967]; however, this will usually result in very high order transfer functions, which explains why this process is currently done mostly by graphical matching using lower order transfer functions.

(iii) Let
$$D(s) = \operatorname{diag}(d_1(s)I, \ldots, d_{F-1}(s)I, I).$$
Construct a state space model for system
$$\hat{G}(s) = \begin{bmatrix} D(s) & \\ & I \end{bmatrix} G(s) \begin{bmatrix} D^{-1}(s) & \\ & I \end{bmatrix}.$$

(iv) Solve an $\mathcal{H}_\infty$-optimization problem to minimize
$$\left\| \mathcal{F}_\ell(\hat{G}, K) \right\|_\infty$$
over all stabilizing K's. Note that this optimization problem uses the scaled version of G. Let its minimizing controller be denoted by $\hat{K}$.

(v) Minimize $\overline{\sigma}[D_\omega \mathcal{F}_\ell(G, \hat{K}) D_\omega^{-1}]$ over D_ω, pointwise across frequency.[4] Note that this evaluation uses the minimizing $\hat{K}$ from the last step, but that G is unscaled. The minimization itself produces a new scaling function. Let this new function be denoted by $\hat{D}_\omega$.

(vi) Compare $\hat{D}_\omega$ with the previous estimate D_ω. Stop if they are close, but, otherwise, replace D_ω with $\hat{D}_\omega$ and return to step (ii).

With either K or D fixed, the global optimum in the other variable may be found using the μ and $\mathcal{H}_\infty$ solutions. Although the joint optimization of D and K is not convex and

[4] The approximate solutions given in the last section may be used.

the global convergence is not guaranteed, many designs have shown that this approach works very well [see e.g. Balas, 1990]. In fact, this is probably the most effective design methodology available today for dealing with such complicated problems. The detailed treatment of μ analysis is given in Packard and Doyle [1991]. The rest of this book will focus on the $\mathcal{H}_\infty$ optimization which is a fundamental tool for μ synthesis.

11.5 Notes and References

This chapter is partially based on the lecture notes given by Doyle [1984] in Honeywell and partially based on the lecture notes by Packard [1991] and the paper by Doyle, Packard, and Zhou [1991]. Parts of section 11.3.3 come from the paper by Stein and Doyle [1991]. The small μ theorem for systems with non-rational plants and uncertainties is proven in Tits [1994]. Other results on μ can be found in Fan and Tits [1986], Fan, Tits, and Doyle [1991], Packard and Doyle [1993], Packard and Pandey [1993], Young [1993], and references therein.

12

Parameterization of Stabilizing Controllers

The basic configuration of the feedback systems considered in this chapter is an LFT as shown in Figure 12.1, where G is the generalized plant with two sets of inputs: the exogenous inputs w, which include disturbances and commands, and control inputs u. The plant G also has two sets of outputs: the measured (or sensor) outputs y and the regulated outputs z. K is the controller to be designed. A control problem in this setup is either to analyze some specific properties, e.g., stability or performance, of the closed-loop or to design the feedback control K such that the closed-loop system is stable in some appropriate sense and the error signal z is specified, i.e., some performance condition is satisfied. In this chapter we are only concerned with the basic internal stabilization problems. We will see again that this setup is very convenient for other general control synthesis problems in the coming chapters.

Suppose that a given feedback system is feedback stabilizable. In this chapter, the problem we are mostly interested in is parameterizing all controllers that stabilize the system. The *parameterization* of all internally stabilizing controllers was first introduced by Youla et al [1976]; in their parameterization, the *coprime factorization* technique is used. All of the existing results are mainly in the frequency domain although they can also be transformed to state-space descriptions. In this chapter, we consider this issue in the general setting and directly in state space without adopting coprime factorization technique. The construction of the controller parameterization is done via considering a sequence of special problems, which are so-called *full information (FI)* problems, *disturbance feedforward (DF)* problems, *full control* (FC) problems and *output estimation*

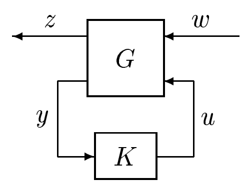


Figure 12.1: General System Interconnection

(OE) problems. On the other hand, these special problems are also of importance in their own right.

In addition to presenting the controller parameterization, this chapter also aims at introducing the synthesis machinery, which is essential in some control syntheses (the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control in Chapter 14 and Chapter 16), and at seeing how it works in the controller parameterization problem. The structure of this chapter is as follows: in section 12.1, the conditions for the existence of a stabilizing controller are examined. In section 12.2, we shall examine the stabilization of different special problems and establish the relations among them. In section 12.3, the construction of controller parameterization for the general output feedback problem will be considered via the special problems FI, DF, FC and OE. In section 12.4, the structure of the controller parameterization is displayed. Section 12.5 shows the closed-loop transfer matrix in terms of the parameterization. Section 12.6 considers an alternative approach to the controller parameterization using coprime factorizations and establishes the connections with the state space approach. This section can be either studied independently of all the preceding sections or skipped over without loss of continuity.

12.1 Existence of Stabilizing Controllers

Consider a system described by the standard block diagram in Figure 12.1. Assume that $G(s)$ has a *stabilizable and detectable* realization of the form

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]. \tag{12.1}$$

The stabilization problem is to find feedback mapping K such that the closed-loop system is internally stable; the well-posedness is required for this interconnection. This general synthesis question will be called the *output feedback (OF)* problem.

Definition 12.1 A proper system G is said to be *stabilizable* through output feedback if there exists a proper controller K internally stabilizing G in Figure 12.1. Moreover, a proper controller $K(s)$ is said to be *admissible* if it internally stabilizes G.

The following result is standard and follows from Chapter 3.

Lemma 12.1 *There exists a proper K achieving internal stability iff (A, B_2) is stabilizable and (C_2, A) is detectable. Further, let F and L be such that $A + B_2F$ and $A + LC_2$ are stable, then an observer-based stabilizing controller is given by*

$$K(s) = \left[\begin{array}{c|c} A + B_2F + LC_2 + LD_{22}F & -L \\ \hline F & 0 \end{array}\right].$$

Proof. ($\Leftarrow$) By the stabilizability and detectability assumptions, there exist F and L such that $A + B_2F$ and $A + LC_2$ are stable. Now let $K(s)$ be the observer-based controller given in the lemma, then the closed-loop A-matrix is given by

$$\tilde{A} = \begin{bmatrix} A & B_2F \\ -LC_2 & A + B_2F + LC_2 \end{bmatrix}.$$

It is easy to check that this matrix is similar to the matrix

$$\begin{bmatrix} A + LC_2 & 0 \\ -LC_2 & A + B_2F \end{bmatrix}.$$

Thus the spectrum of $\tilde{A}$ equals the union of the spectra of $A + LC_2$ and $A + B_2F$. In particular, $\tilde{A}$ is stable.

($\Rightarrow$) If (A, B_2) is not stabilizable or if (C_2, A) is not detectable, then there are some eigenvalues of $\tilde{A}$ which are fixed in the right half-plane, no matter what the compensator is. The details are left as an exercise. □

The stabilizability and detectability conditions of (A, B_2, C_2) are assumed throughout the remainder of this chapter[1]. It follows that the realization for G_{22} is stabilizable and detectable, and these assumptions are enough to yield the following result:

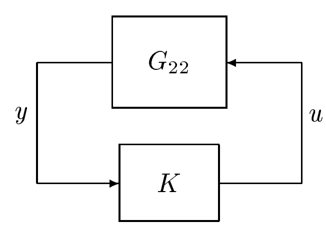


Figure 12.2: Equivalent Stabilization Diagram

[1]It should be clear that the stabilizability and detectability of a realization for G do not guarantee the stabilizability and/or detectability of the corresponding realization for G_{22}.

Lemma 12.2 *Suppose* (A, B_2, C_2) *is stabilizable and detectable and* $G_{22} = \left[\begin{array}{c|c} A & B_2 \\ \hline C_2 & D_{22} \end{array}\right]$. *Then the system in Figure 12.1 is internally stable iff the one in Figure 12.2 is internally stable.*

In other words, $K(s)$ internally stabilizes $G(s)$ if and only if it internally stabilizes G_{22}.

Proof. The necessity follows from the definition. To show the sufficiency, it is sufficient to show that the system in Figure 12.1 and that in Figure 12.2 share the same A-matrix, which is obvious. □

From Lemma 12.2, we see that the stabilizing controller for G depends only on G_{22}. Hence all stabilizing controllers for G can be obtained by using only G_{22}, which is how it is usually done in the conventional Youla parameterization. However, it will be shown that the general setup is very convenient and much more useful since any closed-loop system information can also be considered in the same framework.

Remark 12.1 There should be no confusion between a given realization for a transfer matrix G_{22} and the inherited realization from G where G_{22} is a submatrix. A given realization for G_{22} may be stabilizable and detectable while the inherited realization may be not. For instance,

$$G_{22} = \frac{1}{s+1} = \left[\begin{array}{c|c} -1 & 1 \\ \hline 1 & 0 \end{array}\right]$$

is a minimal realization but the inherited realization of G_{22} from

$$\begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} -1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ \hline 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{array}\right]$$

is

$$G_{22} = \left[\begin{array}{cc|c} -1 & 0 & 1 \\ 0 & 1 & 0 \\ \hline 1 & 0 & 0 \end{array}\right] \quad \left(= \frac{1}{s+1}\right)$$

which is neither stabilizable nor detectable. ♡

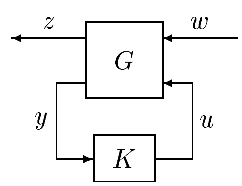
12.2 Duality and Special Problems

In this section, we will discuss four problems from which the output feedback solutions are constructed via a separation argument. These special problems are fundamental to the approach taken for synthesis in this book, and, as we shall see, they are also of importance in their own right.

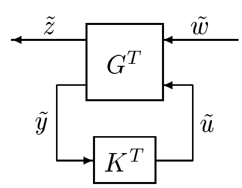
12.2.1 Algebraic Duality and Special Problems

Before we get into the study of the algebraic structure of control systems, we now introduce the concept of algebraic duality which will play an important role. It is well known that the concepts of controllability (stabilizability) and observability (detectability) of a system (C, A, B) are dual because of the duality between (C, A, B) and (B^T, A^T, C^T). So, to deal with the issues related to a system's controllability and/or observability, we only need to examine the issues related to the observability and/or controllability of its dual system, respectively. The notion of duality can be generalized to a general setting.

Consider a standard system block diagram



where the plant G and controller K are assumed to be linear time invariant. Now consider another system shown below



whose plant and controller are obtained by transposing G and K. We can check easily that $T_{zw}^T = [\mathcal{F}_\ell(G, K)]^T = \mathcal{F}_\ell(G^T, K^T) = T_{\tilde{z}\tilde{w}}$. It is not difficult to see that K internally stabilizes G iff K^T internally stabilizes G^T. And we say that these two control structures are *algebraically dual*, especially, G^T and K^T which are dual objects of G and K, respectively. So as far as stabilization or other synthesis problems are concerned, we can obtain the results for G^T from the results for its dual object G if they are available.

Now, we consider some special problems which are related to the general OF problems stated in the last section and which are important in constructing the results for OF problems. The special problems considered here all pertain to the standard block diagram, but to different structures than G. The problems are labeled as

FI. Full information, with the corresponding plant

$$G_{FI} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}\right].$$

FC. Full control, with the corresponding plant

$$G_{FC} = \left[\begin{array}{c|cc} A & B_1 & \begin{bmatrix} I & 0 \end{bmatrix} \\ \hline C_1 & D_{11} & \begin{bmatrix} 0 & I \end{bmatrix} \\ C_2 & D_{21} & \begin{bmatrix} 0 & 0 \end{bmatrix} \end{array}\right].$$

DF. Disturbance feedforward, with the corresponding plant

$$G_{DF} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & I & 0 \end{array}\right].$$

OE. Output estimation, with the corresponding plant

$$G_{OE} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & I \\ C_2 & D_{21} & 0 \end{array}\right].$$

The motivations for these special problems will be given later when they are considered. There are also two additional structures which are standard and which will not be considered in this chapter; they are

SF. State feedback

$$G_{SF} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ I & 0 & 0 \end{array}\right].$$

OI. Output injection

$$G_{OI} = \left[\begin{array}{c|cc} A & B_1 & I \\ \hline C_1 & D_{11} & 0 \\ C_2 & D_{21} & 0 \end{array}\right].$$

Here we assume that all physical variables have compatible dimensions. We say that these special problems are special cases of OF problems in the sense that their structures are specified in comparison to OF problems.

The structure of transfer matrices shows clearly that FC, OE (and OI) are duals of FI, DF (and SF), respectively. These relationships are shown in the following diagram:

$$\begin{tikzcd}[column sep=huge, row sep=huge]
\text{FI} \arrow[r, leftrightarrow, "\textit{dual}"] \arrow[d, leftrightarrow, "\text{equivalent}"'] & \text{FC} \arrow[d, leftrightarrow, "\text{equivalent}"] \\
\text{DF} \arrow[r, leftrightarrow, "\textit{dual}"] & \text{OE}
\end{tikzcd}$$

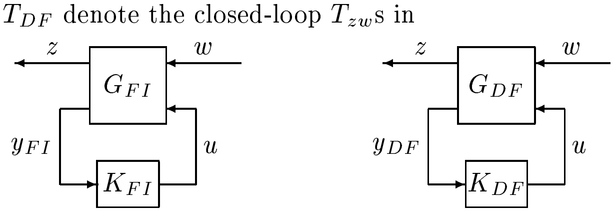
The precise meaning of "equivalent" in this diagram will be explained below.

12.2.2 Full Information and Disturbance Feedforward

In the FI problem, the controller is provided with *Full Information* since $y = \begin{bmatrix} x \\ w \end{bmatrix}$. For the FI problem, we only need to assume that (A, B_2) is stabilizable to guarantee the solvability. It is clear that if any output feedback control problem is to be solvable then the corresponding FI problem has to be solvable, provided FI is of the same structure with OF except for the specified parts.

To motivate the name *Disturbance Feedforward*, consider the special case with $C_2 = 0$. Then there is no feedback and the measurement is exactly w, where w is generally regarded as disturbance to the system. Only the disturbance, w, is fed through directly to the output. As we shall see, the feedback caused by $C_2 \neq 0$ does not affect the transfer function from w to the output z, but it does affect internal stability. In fact, the conditions for the solvability of the DF problem are that (A, B_2) is stabilizable and (C_2, A) is detectable.

Now we examine the connection between the DF problem and the FI problem and show the meaning of their equivalence. Suppose that we have controllers K_{FI} and K_{DF} and let T_{FI} and T_{DF} denote the closed-loop T_{zw}s in

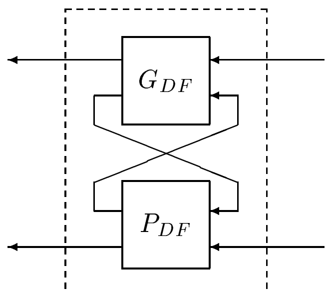


The question of interest is as follows: given either the K_{FI} or the K_{DF} controller, can we construct the other in such a way that $T_{FI} = T_{DF}$? The answer is positive. Actually, we have the following:

Lemma 12.3 *Let G_{FI} and G_{DF} be given as above. Then*

(i) $G_{DF}(s) = \begin{bmatrix} I & 0 & 0 \\ 0 & C_2 & I \end{bmatrix} G_{FI}(s).$

(ii) $G_{FI} = \mathcal{S}(G_{DF}, P_{DF})$ *(where $\mathcal{S}(\cdot,\cdot)$ denotes the star-product)*



$$P_{DF}(s) = \left[\begin{array}{c|cc} A - B_1C_2 & B_1 & B_2 \\ \hline 0 & 0 & I \\ \begin{bmatrix} I \\ -C_2 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}\right].$$

Proof. Part (i) is obvious. Part (ii) follows from applying the star product formula. Nevertheless, we will give a direct proof to show the system structure. Let x and $\hat{x}$ denote the state of G_{DF} and P_{DF}, respectively. Take $e := x - \hat{x}$ and $\hat{x}$ as the states of the resulting interconnected system; then its realization is

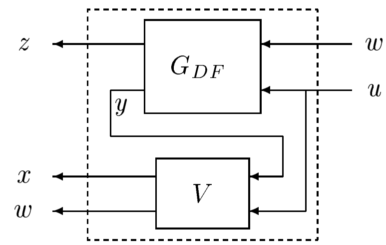
$$\left[\begin{array}{cc|cc} A - B_1C_2 & 0 & 0 & 0 \\ B_1C_2 & A & B_1 & B_2 \\ \hline C_1 & C_1 & D_{11} & D_{12} \\ \begin{bmatrix} 0 \\ C_2 \end{bmatrix} & \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}\right] = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}\right]$$

which is exactly G_{FI}, as claimed. □

Remark 12.2 There is an alternative way to see part (ii). The fact is that in the DF problems, the disturbance w and the system states x can be solved in terms of y and u:

$$\begin{bmatrix} x \\ w \end{bmatrix} = \left[\begin{array}{c|cc} A - B_1C_2 & B_1 & B_2 \\ \hline I & 0 & 0 \\ -C_2 & I & 0 \end{array}\right] \begin{bmatrix} y \\ u \end{bmatrix} =: V \begin{bmatrix} y \\ u \end{bmatrix}.$$

Now connect V up with G_{DF} as shown below



Then it is easy to show that the transfer function from (w, u) to (z, x, w) is G_{FI}, and, furthermore, that the internal stability is not changed if $A - B_1C_2$ is stable. ♡

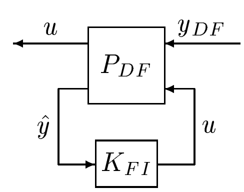
The following theorem follows immediately:

Theorem 12.4 *Let G_{FI}, G_{DF}, and P_{DF} be given as above.*

(i) $K_{FI} := K_{DF}\begin{bmatrix} C_2 & I \end{bmatrix}$ internally stabilizes G_{FI} if K_{DF} internally stabilizes G_{DF}. Furthermore,

$$\mathcal{F}_\ell(G_{FI}, K_{DF}\begin{bmatrix} C_2 & I \end{bmatrix}) = \mathcal{F}_\ell(G_{DF}, K_{DF}).$$

(ii) Suppose $A - B_1C_2$ is stable. Then $K_{DF} := \mathcal{F}_\ell(P_{DF}, K_{FI})$ as shown below



internally stabilizes G_{DF} if K_{FI} internally stabilizes G_{FI}. Furthermore,

$$\mathcal{F}_\ell(G_{DF}, \mathcal{F}_\ell(P_{DF}, K_{FI})) = \mathcal{F}_\ell(G_{FI}, K_{FI}).$$

Remark 12.3 This theorem shows that if $A - B_1C_2$ is stable, then problems FI and DF are equivalent in the above sense. Note that the transfer function from w to y_{DF} is

$$G_{21}(s) = \left[\begin{array}{c|c} A & B_1 \\ \hline C_2 & I \end{array}\right].$$

Hence this stability condition implies that $G_{21}(s)$ has neither right half plane invariant zeros nor hidden unstable modes. ♡

12.2.3 Full Control and Output Estimation

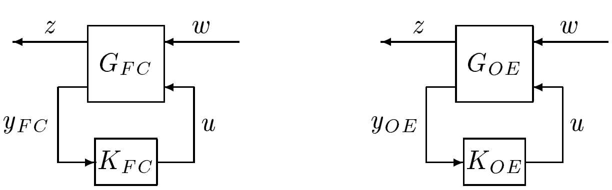
For the FC problem, the term *Full Control* is used because the controller has full access to both the state through output injection and the output z. The only restriction on the controller is that it must work with the measurement y. This problem is dual to the FI case and has the dual solvability condition to the FI problem, which is also guaranteed by the assumptions on OF problems. The solutions to this kind of control problem can be obtained by first transposing G_{FC}, and solving the corresponding FI problem, and then transposing back.

On the other hand, problem OE is dual to DF. Thus the discussion of the DF problem is relevant here, when appropriately dualized. And the solvability conditions for the OE problem are that (A, B_2) is stabilizable and (C_2, A) is detectable. To examine the physical meaning of *output estimation*, first note that

$$z = C_1 x + D_{11} w + u$$

where z is to be controlled by an appropriately designed control u. In general, our control objective will be to specify z in some well-defined mathematical sense. To put it in other words, it is desired to find a u that will estimate $C_1 x + D_{11} w$ in such defined mathematical sense. So this kind of control problem can be regarded as an estimation problem. We are focusing on this particular estimation problem because it is the one that arises in solving the output feedback problem. A more conventional estimation problem would be the special case where no internal stability condition is imposed and $B_2 = 0$. Then the problem would be that of estimating the output z given the measurement y. This special case motivates the term *output estimation* and can be obtained immediately from the results obtained for the general case.

The following discussion will explain the meaning of equivalence between FC and OE problems. Consider the following FC and OE diagrams:



We have similar results to the ones in the last subsection:

Lemma 12.5 *Let G_{FC} and G_{OE} be given as above. Then*

(i) $G_{OE}(s) = G_{FC}(s) \begin{bmatrix} I & 0 \\ 0 & B_2 \\ 0 & I \end{bmatrix}$

(ii) $G_{FC} = \mathcal{S}(G_{OE}, P_{OE})$, *where* P_{OE} *is given by*

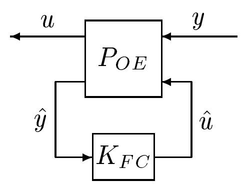
$$P_{OE}(s) = \left[\begin{array}{c|c|c} A - B_2C_1 & 0 & \begin{bmatrix} I & -B_2 \end{bmatrix} \\ \hline C_1 & 0 & \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \\ C_2 & I & \end{array}\right].$$

Theorem 12.6 *Let* G_{FC}, G_{OE}, *and* P_{OE} *be given as above.*

(i) $K_{FC} := \begin{bmatrix} B_2 \\ I \end{bmatrix} K_{OE}$ *internally stabilizes* G_{FC} *if* K_{OE} *internally stabilizes* G_{OE}. *Furthermore,*

$$\mathcal{F}_\ell(G_{FC}, \begin{bmatrix} B_2 \\ I \end{bmatrix} K_{OE}) = \mathcal{F}_\ell(G_{OE}, K_{OE}).$$

(ii) Suppose $A - B_2C_1$ *is stable. Then* $K_{OE} := \mathcal{F}_\ell(P_{OE}, K_{FC})$, *as shown below*



internally stabilizes G_{OE} *if* K_{FC} *internally stabilizes* G_{FC}. *Furthermore,*

$$\mathcal{F}_\ell(G_{OE}, \mathcal{F}_\ell(P_{OE}, K_{FC})) = \mathcal{F}_\ell(G_{FC}, K_{FC}).$$

Remark 12.4 It is seen that if $A - B_2C_1$ is stable, then FC and OE problems are equivalent in the above sense. This condition implies that the transfer matrix $G_{12}(s)$ from u to z has neither right half-plane invariant zeros nor hidden unstable modes, which indicates that it has a stable inverse. ♡

12.3 Parameterization of All Stabilizing Controllers

12.3.1 Problem Statement and Solution

Consider again the standard system block diagram in Figure 12.1 with

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix}.$$

Suppose (A, B_2) is stabilizable and (C_2, A) is detectable. In this section we discuss the following problem:

Given a plant G, parameterize all controllers K that internally stabilize G.

This parameterization for all stabilizing controllers is usually called Youla parameterization. As we have mentioned early, the stabilizing controllers for G will depend only on G_{22}. However, it is more convenient to consider the problem in the general framework as will be shown. The parameterization of all stabilizing controllers is easy when the plant itself is stable.

Theorem 12.7 *Suppose $G \in \mathcal{RH}_\infty$; then the set of all stabilizing controllers can be described as*

$$K = Q(I + G_{22}Q)^{-1} \tag{12.2}$$

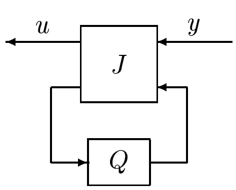
for any $Q \in \mathcal{RH}_\infty$ and $I + D_{22}Q(\infty)$ nonsingular.

Remark 12.5 This result is very natural considering Corollary 5.5, which says that a controller K stabilizes a stable plant G_{22} iff $K(I - G_{22}K)^{-1}$ is stable. Now suppose $Q = K(I - G_{22}K)^{-1}$ is a stable transfer matrix, then K can be solved from this equation which gives exactly the controller parameterization in the above theorem. ♡

Proof. Note that $G_{22}(s)$ is stable by the assumptions on G. Now use straightforward algebra to verify that the controllers given above stabilize G_{22}. On the other hand, suppose K_0 is a stabilizing controller; then $Q_0 := K_0(I - G_{22}K_0)^{-1} \in \mathcal{RH}_\infty$, so K_0 can be expressed as $K_0 = Q_0(I + G_{22}Q_0)^{-1}$. Note that the invertibility in the last equation is guaranteed by the well posedness of the interconnected system with controller K_0 since $I + D_{22}Q_0(\infty) = (I - D_{22}K_0(\infty))^{-1}$. □

However, if G is not stable, the parameterization is much more complicated. The results can be more conveniently stated using state space representations.

Theorem 12.8 *Let F and L be such that $A + LC_2$ and $A + B_2F$ are stable, and then all controllers that internally stabilize G can be parameterized as the transfer matrix from y to u below*



$$J = \left[\begin{array}{c|cc} A + B_2F + LC_2 + LD_{22}F & -L & B_2 + LD_{22} \\ \hline F & 0 & I \\ -(C_2 + D_{22}F) & I & -D_{22} \end{array}\right]$$

with any $Q \in \mathcal{RH}_\infty$ and $I + D_{22}Q(\infty)$ nonsingular.

A non-constructive proof of the theorem can be given by using the same argument as in the proof of Theorem 12.7, i.e., first verify that any controller given by the formula indeed stabilizes the system G, and then show that we can construct a stable Q for any given stabilizing controller K. This approach, however, does not give much insight into the controller construction and thus can not be generalized to other synthesis problems.

The conventional Youla approach to this problem is via coprime factorization [Youla et al, 1976, Vidyasagar, 1985, Desoer et al, 1982], which will be adopted in the later part of this chapter as an alternative approach.

In the following sections, we will present a novel approach to this problem without adopting coprime factorizations. The idea of this approach is to reduce the output feedback problem into some simpler problems, such as FI and OE or FC and DF which admit simple solutions, and then to solve the output feedback problem by the separation argument. The advantages of this approach are that it is simple and that many other synthesis problems, such as $\mathcal{H}_2$ and $\mathcal{H}_\infty$ optimal control problems in Chapters 14 and 16, can be solved by using the same machinery.

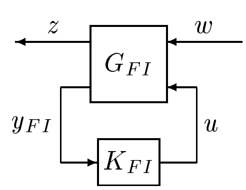
Readers should bear in mind that our objective here is to find all admissible controllers for the OF problem. So at first, we will try to build up enough tools for this objective by considering the special problems. We will see that it is not necessary to parameterize all stabilizing controllers for these special problems to get the required tools. Instead, we only parameterize some equivalent classes of controllers which generate the *same control action.*

Definition 12.2 Two controllers K and $\hat{K}$ are of *equivalent control actions* if their corresponding closed loop transfer matrices are identical, i.e. $\mathcal{F}_l(G, K) = \mathcal{F}_l(G, \hat{K})$, written as $K \cong \hat{K}$.

Algebraically, the controller equivalence is an *equivalence relation.* We will see that for different special problems we have different refined versions of this relation. We will also see that the characterizations of equivalent classes of stabilizing controllers for special problems are good enough to construct the parameterization of all stabilizing controllers for the OF problem. In the next two subsections, we mainly consider the stabilizing controller characterizations for special problems. Also, we use the solutions to these special problems and the approach provided in the last section to characterize all stabilizing controllers of OF problems.

12.3.2 Stabilizing Controllers for FI and FC Problems

In this subsection, we first examine the FI structure



where the transfer matrix G_{FI} is given in section 12.2. The purpose of this subsection is to characterize the equivalent classes of stabilizing controllers K_{FI} that stabilize internally G_{FI} and to build up enough tools to be used later. For this problem, we say *two controllers K_{FI} and $\hat{K}_{FI}$ are equivalent if they produce the same closed-loop transfer function from w to u.* Obviously, this also guarantees that $\mathcal{F}_l(G_{FI}, K_{FI}) = \mathcal{F}_l(G_{FI}, \hat{K}_{FI})$. The characterization of equivalent classes of controllers can be suitably called the *control parameterization* in contrast with controller parameterization. Note that the same situation will occur in FC problems by duality.

Since we have full information for feedback, our controller will have the following general form:

$$K_{FI} = \begin{bmatrix} K_1(s) & K_2(s) \end{bmatrix}$$

with $K_1(s)$ stabilizing internally $(sI - A)^{-1}B_2$ and arbitrary $K_2(s) \in \mathcal{RH}_\infty$. Note that the stability of K_2 is required to guarantee the internal stability of the whole system since w is fed directly through K_2.

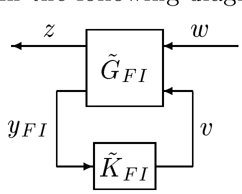
Lemma 12.9 *Let F be a constant matrix such that $A + B_2F$ is stable. Then all stabilizing controllers, in the sense of generating all admissible closed-loop transfer functions, for FI can be parameterized as*

$$K_{FI} \cong \begin{bmatrix} F & Q \end{bmatrix}$$

with any $Q \in \mathcal{RH}_\infty$.

Note that for the parameter matrix $Q \in \mathcal{RH}_\infty$, it is reasonable to assume that *the realization of $Q(s)$ is stabilizable and detectable.*

Proof. It is easy to see that the controller given in the above formula stabilizes the system G_{FI}. Hence we only need to show that the given set of controllers parameterizes all stabilizing control action, u, i.e., there is a choice of $Q \in \mathcal{RH}_\infty$ such that the transfer functions from w to u for any stabilizing controller $K_{FI} = \begin{bmatrix} K_1(s) & K_2(s) \end{bmatrix}$ and for $K_{FI}^0 = \begin{bmatrix} F & Q \end{bmatrix}$ are the same. To show this, make a change of control variable as $v = u - Fx$, where x denotes the state of the system G_{FI}; then the system with the controller K_{FI} will be as shown in the following diagram:



with

$$\tilde{G}_{FI} = \left[\begin{array}{c|cc} A + B_2F & B_1 & B_2 \\ \hline C_1 + D_{12}F & D_{11} & D_{12} \\ \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}\right], \quad \tilde{K}_{FI} = K_{FI} - [F \; 0].$$

Let Q be the transfer matrix from w to v; Q belongs to $\mathcal{RH}_\infty$ by internal stability. Then $u = Fx + v = Fx + Qw$, so $K_{FI} \cong \begin{bmatrix} F & Q \end{bmatrix}$. □

Remark 12.6 The equivalence of $K_{FI} \cong K_{FI}^0$ in the above equation can actually be shown by directly computing Q from equating the transfer matrices from w to u for the cases of K_{FI} and K_{FI}^0. In fact, the transfer matrices from w to u with K_{FI} and K_{FI}^0 are given by

$$\left[I - K_1(sI-A)^{-1}B_2\right]^{-1} K_1(sI-A)^{-1}B_1 + \left[I - K_1(sI-A)^{-1}B_2\right]^{-1} K_2 \tag{12.3}$$

and

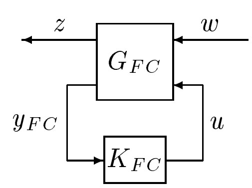
$$\left[I - F(sI-A)^{-1}B_2\right]^{-1} F(sI-A)^{-1}B_1 + \left[I - F(sI-A)^{-1}B_2\right]^{-1} Q, \tag{12.4}$$

respectively. We can verify that

$$Q = K_2 + (K_1 - F)(sI - A - B_2K_1)^{-1}(B_2K_2 + B_1)$$

is stable and makes the formulas in (12.3) and (12.4) equal. ♡

Now we consider the dual FC problem; the system diagram pertinent to this case is



Dually, we say *controllers K_{FC} and $\hat{K}_{FC}$ are equivalent in the sense that the same injection inputs y_{FC}'s produce the same outputs z's.* This also guarantees the identity of their resulting closed-loop transfer matrices from w to z. And we also have

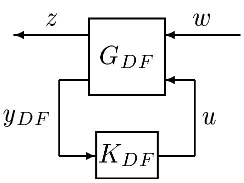
Lemma 12.10 *Let L be a constant matrix such that $A + LC_2$ is stable. Then the set of equivalent classes of all stabilizing controllers for FC in the above sense can be parameterized as*

$$K_{FC} \cong \begin{bmatrix} L \\ Q \end{bmatrix}$$

with any $Q \in \mathcal{RH}_\infty$.

12.3.3 Stabilizing Controllers for DF and OE Problems

In the DF case we have the following system diagram:



The transfer matrix is given as in section 12.2.1. We will further assume that $A - B_1C_2$ is stable in this subsection. It should be pointed out that the existence of a stabilizing controller for this system is guaranteed by the stabilizability of (A, B_2) and detectability of (C_2, A). Hence this assumption is not necessary for our problem to be solvable; however, it does simplify the solution.

We will now parameterize stabilizing controllers for G_{DF} by invoking the relationship between the FI problem and DF problem established in section 12.2. We say that *the controllers K_{DF} and $\hat{K}_{DF}$ are equivalent for the DF problem if the two transfer matrices from w to u in the above diagram are the same.* Of course, the resulting two closed-loop transfer matrices from w to z are identical.

Remark 12.7 By the equivalence between FI and DF problems, it is easy to show that if $K_{DF} \cong \hat{K}_{DF}$ in the DF structure, then $K_{DF}\begin{bmatrix} C_2 & I \end{bmatrix} \cong \hat{K}_{DF}\begin{bmatrix} C_2 & I \end{bmatrix}$ in the corresponding FI structure. We also have that if $K_{FI} \cong \hat{K}_{FI}$, then $\mathcal{F}_l(P_{DF}, K_{FI}) \cong \mathcal{F}_l(P_{DF}, \hat{K}_{FI})$. ♡

Now we construct the parameterization of equivalent classes of stabilizing controllers in DF structure via the tool we have developed in section 12.2.

Let $K_{DF}(s)$ be a stabilizing control for DF; then $K_{FI}(s) = K_{DF}(s)\begin{bmatrix} C_2 & I \end{bmatrix}$ stabilizes the corresponding G_{FI}. Assume $K_{FI} \cong \hat{K}_{FI} = \begin{bmatrix} F & Q \end{bmatrix}$ for some $Q \in \mathcal{RH}_\infty$; then $\hat{K}_{FI}$ stabilizes G_{FI} and $\mathcal{F}_\ell(J_{DF}, Q) = \mathcal{F}_\ell(P_{DF}, \hat{K}_{FI})$ where

$$J_{DF} = \left[\begin{array}{c|cc} A + B_2F - B_1C_2 & B_1 & B_2 \\ \hline F & 0 & I \\ -C_2 & I & 0 \end{array}\right]$$

with F such that $A + B_2F$ is stable. Hence by Theorem 12.4, $\hat{K}_{DF} := \mathcal{F}_\ell(J_{DF}, Q)$ stabilizes G_{DF} for any $Q \in \mathcal{RH}_\infty$. Since $K_{FI} \cong \hat{K}_{FI}$, by Remarks 12.7 we have $K_{DF} \cong \hat{K}_{DF} = \mathcal{F}_\ell(J_{DF}, Q)$. This equation characterizes an equivalent class of all controllers for the DF problem by the equivalence of FI and DF.

In fact, we have the following lemma which shows that the above construction of parameterization characterizes all stabilizing controllers for the DF problem.

Lemma 12.11 *All stabilizing controllers for the DF problem can be characterized by* $K_{DF} = \mathcal{F}_\ell(J_{DF}, Q)$ *with* $Q \in \mathcal{RH}_\infty$*, where* J_{DF} *is given as above.*

Proof. We have already shown that the controller $K_{DF} = \mathcal{F}_\ell(J_{DF}, Q)$ for any given $Q \in \mathcal{RH}_\infty$ does internally stabilize G_{DF}. Now let K_{DF} be any stabilizing controller for G_{DF}; then $\mathcal{F}_\ell(\hat{J}_{DF}, K_{DF}) \in \mathcal{RH}_\infty$ where

$$\hat{J}_{DF} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline -F & 0 & I \\ C_2 & I & 0 \end{array}\right].$$

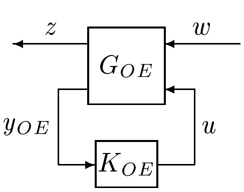
($\hat{J}_{DF}$ is stabilized by K_{DF} since it has the same 'G_{22}' matrix as G_{DF}.)

Let $Q_0 := \mathcal{F}_\ell(\hat{J}_{DF}, K_{DF}) \in \mathcal{RH}_\infty$; then $\mathcal{F}_\ell(J_{DF}, Q_0) = \mathcal{F}_\ell(J_{DF}, \mathcal{F}_\ell(\hat{J}_{DF}, K_{DF})) =: \mathcal{F}_\ell(J_{tmp}, K_{DF})$, where J_{tmp} can be obtained by using the state space star product formula given in Chapter 10:

$$\begin{aligned} J_{tmp} &= \left[\begin{array}{cc|cc} A - B_1C_2 + B_2F & -B_2F & B_1 & B_2 \\ -B_1C_2 & A & B_1 & B_2 \\ \hline F & -F & 0 & I \\ -C_2 & C_2 & I & 0 \end{array}\right] \\ &= \left[\begin{array}{cc|cc} A - B_1C_2 & -B_2F & B_1 & B_2 \\ 0 & A + B_2F & 0 & 0 \\ \hline 0 & -F & 0 & I \\ 0 & C_2 & I & 0 \end{array}\right] \\ &= \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}. \end{aligned}$$

Hence $\mathcal{F}_\ell(J_{DF}, Q_0) = \mathcal{F}_\ell(J_{tmp}, K_{DF}) = K_{DF}$. This shows that any stabilizing controller can be expressed in the form of $\mathcal{F}_\ell(J_{DF}, Q_0)$ for some $Q_0 \in \mathcal{RH}_\infty$. □

Now we turn to the dual OE case. The corresponding system diagram is shown as below:



We will assume that $A - B_2C_1$ is stable. Again this assumption is made only for the simplicity of the solution, it is not necessary for the stabilization problem to be solvable. The parameterization of an equivalent class of stabilizing controllers for OE can be obtained by invoking Theorem 12.6 and the equivalent classes of controllers for FC. Here we say *controllers K_{OE} and $\hat{K}_{OE}$ are equivalent if the transfer matrices from y_{OE} to z are the same.* This also guarantees that the resulting closed-loop transfer matrices are identical.

Now we construct the parameterization for the OE structure as a dual case to DF. Assume that K_{OE} is an admissible controller for OE; then $K_{FC} = \begin{bmatrix} B_2 \\ I \end{bmatrix} K_{OE} \cong \begin{bmatrix} L \\ Q \end{bmatrix} =: \hat{K}_{FC}$ for some $Q \in \mathcal{RH}_\infty$, and $\hat{K}_{FC}$ stabilizes G_{FC} and $\mathcal{F}_\ell(J_{OE}, Q) = \mathcal{F}_\ell(P_{OE}, \hat{K}_{FC})$, where

$$J_{OE} = \left[\begin{array}{c|cc} A - B_2C_1 + LC_2 & L & -B_2 \\ \hline C_1 & 0 & I \\ C_2 & I & 0 \end{array}\right]$$

with L such that $A + LC_2$ is stable. Hence by Theorem 12.6, $\hat{K}_{OE} = \mathcal{F}_\ell(J_{OE}, Q)$ stabilizes G_{OE} for any $Q \in \mathcal{RH}_\infty$. Since $K_{OE} \cong \hat{K}_{OE}$, $K_{OE} \cong \mathcal{F}_\ell(J_{OE}, Q)$. In fact, we have the following lemma.

Lemma 12.12 *All admissible controllers for the OE problem can be characterized as $\mathcal{F}_\ell(J_{OE}, Q_0)$ with any $Q_0 \in \mathcal{RH}_\infty$, where J_{OE} is defined as above.*

Proof. The controllers in the form as stated in the theorem are admissible since the corresponding FC controllers internally stabilize resulting G_{FC}.

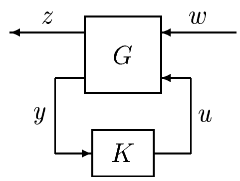
Now assume K_{OE} is any stabilizing controller for G_{OE}; then $\mathcal{F}_\ell(\hat{J}_{OE}, K_{OE}) \in \mathcal{RH}_\infty$ where

$$\hat{J}_{OE} = \left[\begin{array}{c|cc} A & -L & B_2 \\ \hline C_1 & 0 & I \\ C_2 & I & 0 \end{array}\right].$$

Let $Q_0 := \mathcal{F}_\ell(\hat{J}_{OE}, K_{OE}) \in \mathcal{RH}_\infty$. Then $\mathcal{F}_\ell(J_{OE}, Q_0) = \mathcal{F}_\ell(J_{OE}, \mathcal{F}_\ell(\hat{J}_{OE}, K_{OE})) = K_{OE}$, by using again the state space star product formula given in Chapter 10. This shows that any stabilizing controller can be expressed in the form of $\mathcal{F}_\ell(J_{OE}, Q_0)$ for some $Q_0 \in \mathcal{RH}_\infty$. □

12.3.4 Output Feedback and Separation

We are now ready to give a complete proof for Theorem 12.8. We will assume the results of the special problems and show how to construct all admissible controllers for the OF problem from them. And we can also observe the separation argument as the byproduct; this essentially involves reducing the OF problem to the combination of the simpler FI and FC problems. Moreover, we can see from the construction why the stability conditions of $A - B_1C_2$ and $A - B_2C_1$ in DF and OE problems were reasonably assumed and are automatically guaranteed in this case. Again we assume that the system has the following standard system block diagram:



with

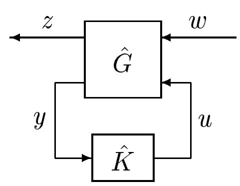
$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

and that (A, B_2) is stabilizable and (C_2, A) is detectable.

Proof of Theorem 12.8. Without loss of generality, we shall assume $D_{22} = 0$. For more general cases, i.e. $D_{22} \neq 0$, the mapping

$$\hat{K}(s) = K(s)(I - D_{22}K(s))^{-1}$$

is well defined if the closed-loop system is assumed to be well posed. Then the system in terms of $\hat{K}$ has the structure



where

$$\hat{G}(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right].$$

Now we construct the controllers for the OF problem with $D_{22} = 0$. Denote x the state of system G; then the open-loop system can be written as

$$\begin{bmatrix} \dot{x} \\ z \\ y \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{bmatrix} \begin{bmatrix} x \\ w \\ u \end{bmatrix}.$$

Since (A, B_2) is stabilizable, there is a constant matrix F such that $A + B_2F$ is stable. Note that $\begin{bmatrix} F & 0 \end{bmatrix}$ is actually a special FI stabilizing controller. Now let

$$v = u - Fx.$$

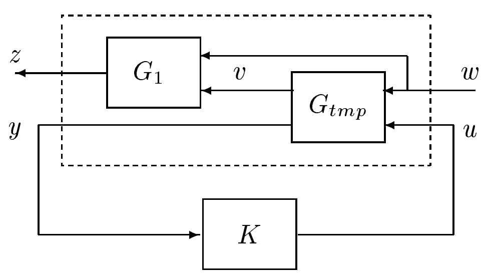
Then the system can be broken into two subsystems:

$$\begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} A + B_2F & B_1 & B_2 \\ C_1 + D_{12}F & D_{11} & D_{12} \end{bmatrix} \begin{bmatrix} x \\ w \\ v \end{bmatrix}$$

and

$$\begin{bmatrix} \dot{x} \\ v \\ y \end{bmatrix} = \begin{bmatrix} A & B_1 & B_2 \\ -F & 0 & I \\ C_2 & D_{21} & 0 \end{bmatrix} \begin{bmatrix} x \\ w \\ u \end{bmatrix}.$$

This can be shown pictorially below:



with

$$G_1 = \left[\begin{array}{c|cc} A + B_2F & B_1 & B_2 \\ \hline C_1 + D_{12}F & D_{11} & D_{12} \end{array}\right] \in \mathcal{RH}_\infty$$

and

$$G_{tmp} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline -F & 0 & I \\ C_2 & D_{21} & 0 \end{array}\right].$$

Obviously, K stabilizes G if and only if K stabilizes G_{tmp}; however, G_{tmp} is of OE structure. Now let L be such that $A + LC_2$ is stable. Then by Lemma 12.12 all controllers stabilizing G_{tmp} are given by

$$K = \mathcal{F}_\ell(J, Q)$$

where

$$J = \left[\begin{array}{c|cc} A + B_2F + LC_2 & L & -B_2 \\ \hline -F & 0 & I \\ C_2 & I & 0 \end{array}\right] = \left[\begin{array}{c|cc} A + B_2F + LC_2 & -L & B_2 \\ \hline F & 0 & I \\ -C_2 & I & 0 \end{array}\right].$$

This concludes the proof. □

Remark 12.8 We can also get the same result by applying the dual procedure to the above construction, i.e., first use an output injection to reduce the OF problem to a DF problem. The separation argument is obvious since the synthesis of the OF problem can be reduced to FI and FC problems, i.e. the latter two problems can be designed independently. ♡

Remark 12.9 Theorem 12.8 shows that any stabilizing controller $K(s)$ can be characterized as an LFT of a parameter matrix $Q \in \mathcal{RH}_\infty$, i.e., $K(s) = \mathcal{F}_\ell(J, Q)$. Moreover, using the same argument as in the proof of Lemma 12.11, a realization of $Q(s)$ in terms of K can be obtained as

$$Q := \mathcal{F}_\ell(\hat{J}, K)$$

where

$$\hat{J} = \left[\begin{array}{c|cc} A & -L & B_2 \\ \hline -F & 0 & I \\ C_2 & I & D_{22} \end{array}\right]$$

and where $K(s)$ has the stabilizable and detectable realization. ♡

Now we can reconsider the characterizations of all stabilizing controllers for the special problems with some reasonable assumptions, i.e. the stability conditions of $A - B_1C_2$ and $A - B_2C_1$ for DF and OE problems which were assumed in the last section can be dropped.

If we specify

$$C_2 = \begin{bmatrix} I \\ 0 \end{bmatrix} \quad D_{21} = \begin{bmatrix} 0 \\ I \end{bmatrix} \quad D_{22} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

the OF problem, in its general setting, becomes the FI problem. We know that the solvability conditions for the FI problem are reduced, because of its special structure, to (A, B_2) as stabilizable. By assuming this, we can get the following result from the OF problem.

Corollary 12.13 *Let L_1 and F be such that $A + L_1$ and $A + B_2F$ are stable; then all controllers that stabilize G_{FI} can be characterized as $\mathcal{F}_\ell(J_{FI}, Q)$ with any $Q \in \mathcal{RH}_\infty$,*

where

$$J_{FI} = \left[\begin{array}{c|cc} A + B_2F + L_1 & L & -B_2 \\ \hline -F & 0 & I \\ \begin{bmatrix} I \\ 0 \end{bmatrix} & I & 0 \end{array}\right]$$

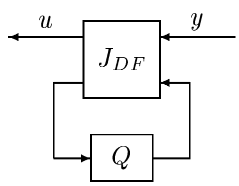
and where $L = (L_1 \ L_2)$ *is the injection matrix for any* L_2 *with compatible dimensions.*

In the same way, we can consider the FC problem as the special OF problem by specifying

$$B_2 = \begin{bmatrix} I & 0 \end{bmatrix} \quad D_{12} = \begin{bmatrix} 0 & I \end{bmatrix} \quad D_{22} = \begin{bmatrix} 0 & 0 \end{bmatrix}.$$

The $DF(OE)$ problem can also be considered as the special case of OF by simply setting $D_{21} = I(D_{12} = I)$ and $D_{22} = 0$.

Corollary 12.14 *Consider the DF problem, and assume that* (C_2, A, B_2) *is stabilizable and detectable. Let* F *and* L *be such that* $A + LC_2$ *and* $A + B_2F$ *are stable, and then all controllers that internally stabilize* G *can be parameterized as* $\mathcal{F}_l(J_{DF}, Q)$ *for some* $Q \in \mathcal{RH}_\infty$*, i.e. the transfer function from* y *to* u *is shown as below*



$$J_{DF} = \left[\begin{array}{c|cc} A + B_2F + LC_2 & -L & B_2 \\ \hline F & 0 & I \\ -C_2 & I & 0 \end{array}\right].$$

Remark 12.10 It would be interesting to compare this result with Lemma 12.11. It can be seen that Lemma 12.11 is a special case of this corollary. The condition that $A - B_1C_2$ is stable, which is required in Lemma 12.11, provides the natural injection matrix $L = -B_1$ which satisfies a partial condition in this corollary. ♡

12.4 Structure of Controller Parameterization

Let us recap what we have done. We begin with a stabilizable and detectable realization of G_{22}

$$G_{22} = \left[\begin{array}{c|c} A & B_2 \\ \hline C_2 & D_{22} \end{array}\right].$$

We choose F and L so that $A + B_2F$ and $A + LC_2$ are stable. Define J by the formula in Theorem 12.8. Then the proper K's achieving internal stability are precisely those representable in Figure 12.3 and $K = \mathcal{F}_\ell(J, Q)$ where $Q \in \mathcal{RH}_\infty$ and $I + D_{22}Q(\infty)$ is invertible.

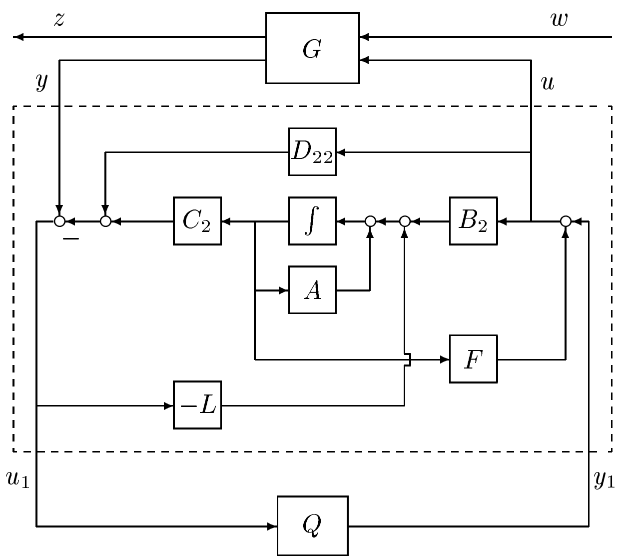


Figure 12.3: Structure of Stabilizing Controllers

It is interesting to note that the system in the dashed box is an observer-based stabilizing controller for G (or G_{22}). Furthermore, it is easy to show that the transfer function between (y, y_1) and (u, u_1) is J, i.e.,

$$\begin{bmatrix} u \\ u_1 \end{bmatrix} = J \begin{bmatrix} y \\ y_1 \end{bmatrix}.$$

It is also easy to show that the transfer matrix from y_1 to u_1 is zero.

This diagram of all stabilizing controller parameterization also suggests an interesting interpretation: every internal stabilization amounts to adding stable dynamics to the plant and then stabilizing the extended plant by means of an observer. The precise statement is as follows: for simplicity of the formulas, only the cases of strictly proper G_{22} and K are treated.

Theorem 12.15 *Assume that G_{22} and K are strictly proper and the system is Figure 12.1 is internally stable. Then G_{22} can be embedded in a system*

$$\left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & 0 \end{array}\right]$$

where

$$A_e = \begin{bmatrix} A & 0 \\ 0 & A_a \end{bmatrix}, \quad B_e = \begin{bmatrix} B_2 \\ 0 \end{bmatrix}, \quad C_e = \begin{bmatrix} C_2 & 0 \end{bmatrix} \tag{12.5}$$

and where A_a is stable, such that K has the form

$$K = \left[\begin{array}{c|c} A_e + B_e F_e + L_e C_e & -L_e \\ \hline F_e & 0 \end{array}\right] \tag{12.6}$$

where $A_e + B_e F_e$ and $A_e + L_e C_e$ are stable.

Proof. K is representable as in Figure 12.3 for some Q in $\mathcal{RH}_\infty$. For K to be strictly proper, Q must be strictly proper. Take a minimal realization of Q:

$$Q = \left[\begin{array}{c|c} A_a & B_a \\ \hline C_a & 0 \end{array}\right].$$

Since $Q \in \mathcal{RH}_\infty$, A_a is stable. Let x and x_a denote state vectors for J and Q, respectively, and write the equations for the system in Figure 12.3:

$$\begin{aligned} \dot{x} &= (A + B_2F + LC_2)x - Ly + B_2 y_1 \\ u &= Fx + y_1 \\ u_1 &= -C_2 x + y \\ \dot{x}_a &= A_a x_a + B_a u_1 \\ y_1 &= C_a x_a \end{aligned}$$

These equations yield

$$\begin{aligned} \dot{x}_e &= (A_e + B_e F_e + L_e C_e)x_e - L_e y \\ u &= F_e x_e \end{aligned}$$

where

$$x_e := \begin{bmatrix} x \\ x_a \end{bmatrix}, \; F_e := \begin{bmatrix} F & C_a \end{bmatrix}, \; L_e := \begin{bmatrix} L \\ -B_a \end{bmatrix}$$

and where A_e, B_e, C_e are as in (12.5). □

12.5 Closed-Loop Transfer Matrix

Recall that the closed-loop transfer matrix from w to z is a linear fractional transformation $\mathcal{F}_\ell(G, K)$ and that K stabilizes G if and only if K stabilizes G_{22}. Elimination of the signals u and y in Figure 12.3 leads to Figure 12.4 for a suitable transfer matrix T. Thus all closed-loop transfer matrices are representable as in Figure 12.4.

$$z = \mathcal{F}_\ell(G, K)w = \mathcal{F}_\ell(G, \mathcal{F}_\ell(J, Q))w = \mathcal{F}_\ell(T, Q)w. \tag{12.7}$$

It remains to give a realization of T.

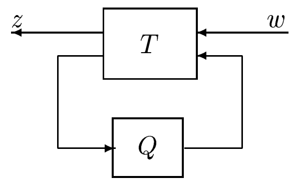


Figure 12.4: Closed loop system

Theorem 12.16 *Let F and L be such that $A+BF$ and $A+LC$ are stable. Then the set of all closed-loop transfer matrices from w to z achievable by an internally stabilizing proper controller is equal to*

$$\mathcal{F}_\ell(T,Q) = \{T_{11} + T_{12}QT_{21} : \ Q \in \mathcal{RH}_\infty, \ I + D_{22}Q(\infty) \text{ invertible}\}$$

where T is given by

$$T = \begin{bmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} A+B_2F & -B_2F & B_1 & B_2 \\ 0 & A+LC_2 & B_1+LD_{21} & 0 \\ \hline C_1+D_{12}F & -D_{12}F & D_{11} & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right].$$

Proof. This is straightforward by using the state space star product formula and follows from some tedious algebra. Hence it is left for the readers to verify. □

An important point to note is that the closed-loop transfer matrix is simply an affine function of the controller parameter matrix Q since $T_{22} = 0$.

12.6 Youla Parameterization via Coprime Factorization*

In this section, all stabilizing controller parameterization will be derived using the conventional coprime factorization approach. Readers should be familiar with the results presented in Section 5.4 of Chapter 5 before proceeding further.

Theorem 12.17 *Let $G_{22} = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ be the rcf and lcf of G_{22} over $\mathcal{RH}_\infty$, respectively. Then the set of all proper controllers achieving internal stability is parameterized either by*

$$K = (U_0 + MQ_r)(V_0 + NQ_r)^{-1}, \quad \det(I + V_0^{-1}NQ_r)(\infty) \neq 0 \tag{12.8}$$

for $Q_r \in \mathcal{RH}_\infty$ or by

$$K = (\tilde{V}_0 + Q_l\tilde{N})^{-1}(\tilde{U}_0 + Q_l\tilde{M}), \quad \det(I + Q_l\tilde{N}\tilde{V}_0^{-1})(\infty) \neq 0 \tag{12.9}$$

for $Q_l \in \mathcal{RH}_\infty$ *where* $U_0, V_0, \tilde{U}_0, \tilde{V}_0 \in \mathcal{RH}_\infty$ *satisfy the Bezout identities:*

$$\tilde{V}_0 M - \tilde{U}_0 N = I, \qquad \tilde{M} V_0 - \tilde{N} U_0 = I.$$

Moreover, if $U_0, V_0, \tilde{U}_0$*, and* $\tilde{V}_0$ *are chosen such that* $U_0 V_0^{-1} = \tilde{V}_0^{-1} \tilde{U}_0$*, i.e.,*

$$\begin{bmatrix} \tilde{V}_0 & -\tilde{U}_0 \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}.$$

Then

$$\begin{aligned} K &= (U_0 + M Q_y)(V_0 + N Q_y)^{-1} \\ &= (\tilde{V}_0 + Q_y \tilde{N})^{-1} (\tilde{U}_0 + Q_y \tilde{M}) \\ &= \mathcal{F}_\ell(J_y, Q_y) \end{aligned} \tag{12.10}$$

where

$$J_y := \begin{bmatrix} U_0 V_0^{-1} & \tilde{V}_0^{-1} \\ V_0^{-1} & -V_0^{-1} N \end{bmatrix} \tag{12.11}$$

and where Q_y *ranges over* $\mathcal{RH}_\infty$ *such that* $(I + V_0^{-1} N Q_y)(\infty)$ *is invertible*

Proof. We shall prove the parameterization given in (12.8) first. Assume that K has the form indicated, and define

$$U := U_0 + M Q_r, \; V := V_0 + N Q_r.$$

Then

$$\tilde{M} V - \tilde{N} U = \tilde{M}(V_0 + N Q_r) - \tilde{N}(U_0 + M Q_r) = \tilde{M} V_0 - \tilde{N} U_0 + (\tilde{M} N - \tilde{N} M) Q_r = I.$$

Thus K achieves internal stability by Lemma 5.10.

Conversely, suppose K is proper and achieves internal stability. Introduce an rcf of K over $\mathcal{RH}_\infty$ as $K = U V^{-1}$. Then by Lemma 5.10, $Z := \tilde{M} V - \tilde{N} U$ is invertible in $\mathcal{RH}_\infty$. Define Q_r by the equation

$$U_0 + M Q_r = U Z^{-1}, \tag{12.12}$$

so

$$Q_r = M^{-1}(U Z^{-1} - U_0).$$

Then using the Bezout identity, we have

$$\begin{aligned} V_0 + N Q_r &= V_0 + N M^{-1}(U Z^{-1} - U_0) \\ &= V_0 + \tilde{M}^{-1} \tilde{N}(U Z^{-1} - U_0) \\ &= \tilde{M}^{-1}(\tilde{M} V_0 - \tilde{N} U_0 + \tilde{N} U Z^{-1}) \\ &= \tilde{M}^{-1}(I + \tilde{N} U Z^{-1}) \\ &= \tilde{M}^{-1}(Z + \tilde{N} U) Z^{-1} \\ &= \tilde{M}^{-1} \tilde{M} V Z^{-1} \\ &= V Z^{-1}. \end{aligned} \tag{12.13}$$

Thus,

$$\begin{aligned} K &= UV^{-1} \\ &= (U_0 + MQ_r)(V_0 + NQ_r)^{-1}. \end{aligned}$$

To see that Q_r belongs to $\mathcal{RH}_\infty$, observe first from (12.12) and then from (12.13) that both MQ_r and NQ_r belong to $\mathcal{RH}_\infty$. Then

$$Q_r = (\tilde{V}_0 M - \tilde{U}_0 N)Q_r = \tilde{V}_0(MQ_r) - \tilde{U}_0(NQ_r) \in \mathcal{RH}_\infty.$$

Finally, since V and Z evaluated at $s = \infty$ are both invertible, so is $V_0 + NQ_r$ from (12.13), hence so is $I + V_0^{-1}NQ_r$.

Similarly, the parameterization given in (12.9) can be obtained.

To show that the controller can be written in the form of equation (12.10), note that

$$(U_0 + MQ_y)(V_0 + NQ_y)^{-1} = U_0V_0^{-1} + (M - U_0V_0^{-1}N)Q_y(I + V_0^{-1}NQ_y)^{-1}V_0^{-1}$$

and that $U_0V_0^{-1} = \tilde{V}_0^{-1}\tilde{U}_0$. We have

$$(M - U_0V_0^{-1}N) = (M - \tilde{V}_0^{-1}\tilde{U}_0 N) = \tilde{V}_0^{-1}(\tilde{V}_0 M - \tilde{U}_0 N) = \tilde{V}_0^{-1}$$

and

$$K = U_0V_0^{-1} + \tilde{V}_0^{-1}Q_y(I + V_0^{-1}NQ_y)^{-1}V_0^{-1}. \tag{12.14}$$

□

Corollary 12.18 *Given an admissible controller K with coprime factorizations $K = UV^{-1} = \tilde{V}^{-1}\tilde{U}$, the free parameter $Q_y \in \mathcal{RH}_\infty$ in Youla parameterization is given by*

$$Q_y = M^{-1}(UZ^{-1} - U_0)$$

where

$$Z := \tilde{M}V - \tilde{N}U.$$

Next, we shall establish the precise relationship between the above all stabilizing controller parameterization and the parameterization obtained in the previous sections via LFT framework.

Theorem 12.19 *Let the doubly coprime factorizations of G_{22} be chosen as*

$$\begin{bmatrix} M & U_0 \\ N & V_0 \end{bmatrix} = \left[\begin{array}{c|cc} A + B_2F & B_2 & -L \\ \hline F & I & 0 \\ C_2 + D_{22}F & D_{22} & I \end{array}\right]$$

$$\begin{bmatrix} \tilde{V}_0 & -\tilde{U}_0 \\ -\tilde{N} & \tilde{M} \end{bmatrix} = \left[\begin{array}{c|cc} A + LC_2 & -(B_2 + LD_{22}) & L \\ \hline F & I & 0 \\ C_2 & -D_{22} & I \end{array}\right]$$

where F and L are chosen such that $A + B_2F$ and $A + LC_2$ are both stable.

Then J_y can be computed as

$$J_y = \left[\begin{array}{c|cc} A + B_2F + LC_2 + LD_{22}F & -L & B_2 + LD_{22} \\ \hline F & 0 & I \\ -(C_2 + D_{22}F) & I & -D_{22} \end{array}\right].$$

Proof. This follows from some tedious algebra. □

Remark 12.11 Note that J_y is exactly the same as the J in Theorem 12.8 and that $K_0 := U_0V_0^{-1}$ is an observer-based stabilizing controller with

$$K_0 := \left[\begin{array}{c|c} A + B_2F + LC_2 + LD_{22}F & -L \\ \hline F & 0 \end{array}\right].$$

♡

12.7 Notes and References

The special problems FI, DF, FC, and OE were first introduced in Doyle, Glover, Khargonekar, and Francis [1989] for solving the $\mathcal{H}_\infty$ problem, and they have been since used in many other papers for different problems. The new derivation of all stabilizing controllers was reported in Lu, Zhou, and Doyle [1991]. The paper by Moore *et al* [1990] contains some other related interesting results. The conventional Youla parameterization can be found in Youla *et al* [1976], Desoer *et al* [1980], Doyle [1984], Vidyasagar [1985], and Francis [1987]. The parameterization of all two-degree-of-freedom stabilizing controllers is given in Youla and Bongiorno [1985] and Vidyasagar [1985].

13

Algebraic Riccati Equations

We have studied the Lyapunov equation in Chapter 3 and have seen the roles it played in some applications. A more general equation than the Lyapunov equation in control theory is the so-called *Algebraic Riccati Equation* or ARE for short. Roughly speaking, Lyapunov equations are most useful in system analysis while AREs are most useful in control system synthesis; particularly, they play the central roles in $\mathcal{H}_2$ and $\mathcal{H}_\infty$ optimal control.

Let A, Q, and R be real $n \times n$ matrices with Q and R symmetric. Then an algebraic Riccati equation is the following matrix equation:

$$A^*X + XA + XRX + Q = 0. \tag{13.1}$$

Associated with this Riccati equation is a $2n \times 2n$ matrix:

$$H := \begin{bmatrix} A & R \\ -Q & -A^* \end{bmatrix}. \tag{13.2}$$

A matrix of this form is called a *Hamiltonian matrix*. The matrix H in (13.2) will be used to obtain the solutions to the equation (13.1). It is useful to note that $\sigma(H)$ (the spectrum of H) is symmetric about the imaginary axis. To see that, introduce the $2n \times 2n$ matrix:

$$J := \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}$$

having the property $J^2 = -I$. Then

$$J^{-1}HJ = -JHJ = -H^*$$

so H and $-H^*$ are similar. Thus λ is an eigenvalue iff $-\bar{\lambda}$ is.

This chapter is devoted to the study of this algebraic Riccati equation and related problems: the properties of its solutions, the methods to obtain the solutions, and some applications.

In Section 13.1, we will study all solutions to (13.1). The word "all" means that any X, which is not necessarily real, not necessarily hermitian, not necessarily nonnegative, and not necessarily stabilizing, satisfies equation (13.1). The conditions for a solution to be hermitian, real, and so on, are also given in this section. The most important part of this chapter is Section 13.2 which focuses on the stabilizing solutions. This section is designed to be essentially self-contained so that readers who are only interested in the stabilizing solution may go to Section 13.2 directly without any difficulty. Section 13.3 presents the extreme (i.e., maximal or minimal) solutions of a Riccati equation and their properties. The relationship between the stabilizing solution of a Riccati equation and the spectral factorization of some frequency domain function is established in Section 13.4. Positive real functions and inner functions are introduced in Section 13.5 and 13.6. Some other special rational matrix factorizations, e.g., inner-outer factorizations and normalized coprime factorization, are given in Sections 13.7-13.8.

13.1 All Solutions of A Riccati Equation

The following theorem gives a way of constructing solutions to (13.1) in terms of invariant subspaces of H.

Theorem 13.1 *Let $\mathcal{V} \subset \mathbb{C}^{2n}$ be an n-dimensional invariant subspace of H, and let $X_1, X_2 \in \mathbb{C}^{n\times n}$ be two complex matrices such that*

$$\mathcal{V} = \operatorname{Im} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}.$$

If X_1 is invertible, then $X := X_2 X_1^{-1}$ is a solution to the Riccati equation (13.1) and $\sigma(A + RX) = \sigma(H|_{\mathcal{V}})$. Furthermore, the solution X is independent of a specific choice of bases of $\mathcal{V}$.

Proof. Since $\mathcal{V}$ is an H invariant subspace, there is a matrix $\Lambda \in \mathbb{C}^{n\times n}$ such that

$$\begin{bmatrix} A & R \\ -Q & -A^* \end{bmatrix} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \Lambda.$$

Postmultiply the above equation by X_1^{-1} to get

$$\begin{bmatrix} A & R \\ -Q & -A^* \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} X_1 \Lambda X_1^{-1}. \tag{13.3}$$

Now pre-multiply (13.3) by $\begin{bmatrix} -X & I \end{bmatrix}$ to get

$$\begin{aligned} 0 &= \begin{bmatrix} -X & I \end{bmatrix} \begin{bmatrix} A & R \\ -Q & -A^* \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} \\ &= -XA - A^*X - XRX - Q, \end{aligned}$$

which shows that X is indeed a solution of (13.1). Equation (13.3) also gives

$$A + RX = X_1 \Lambda X_1^{-1};$$

therefore, $\sigma(A + RX) = \sigma(\Lambda)$. But, by definition, Λ is a matrix representation of the map $H|_{\mathcal{V}}$, so $\sigma(A + RX) = \sigma(H|_{\mathcal{V}})$. Finally note that any other basis spanning $\mathcal{V}$ can be represented as

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} P = \begin{bmatrix} X_1 P \\ X_2 P \end{bmatrix}$$

for some nonsingular matrix P. The conclusion follows from the fact $(X_2P)(X_1P)^{-1} = X_2X_1^{-1}$. □

As we would expect, the converse of the theorem also holds.

Theorem 13.2 *If $X \in \mathbb{C}^{n\times n}$ is a solution to the Riccati equation (13.1), then there exist matrices $X_1, X_2 \in \mathbb{C}^{n\times n}$, with X_1 invertible, such that $X = X_2X_1^{-1}$ and the columns of $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ form a basis of an n-dimensional invariant subspace of H.*

Proof. Define $\Lambda := A + RX$. Multiplying this by X gives

$$X\Lambda = XA + XRX = -Q - A^*X$$

with the second equality coming from the fact that X is a solution to (13.1). Write these two relations as

$$\begin{bmatrix} A & R \\ -Q & -A^* \end{bmatrix} \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} \Lambda.$$

Hence, the columns of $\begin{bmatrix} I \\ X \end{bmatrix}$ span an n-dimensional invariant subspace of H, and defining $X_1 := I$, and $X_2 := X$ completes the proof. □

Remark 13.1 It is now clear that to obtain solutions to the Riccati equation, it is necessary to be able to construct bases for those invariant subspaces of H. One way of constructing those invariant subspaces is to use eigenvectors and generalized eigenvectors of H. Suppose λ_i is an eigenvalue of H with multiplicity k (then $\lambda_{i+j} = \lambda_i$ for all $j = 1, \ldots, k-1$), and let v_i be a corresponding eigenvector and $v_{i+1}, \ldots, v_{i+k-1}$ be the corresponding generalized eigenvectors associated with v_i and λ_i. Then v_j are related by

$$\begin{aligned}
(H - \lambda_i I)v_i &= 0 \\
(H - \lambda_i I)v_{i+1} &= v_i \\
&\vdots \\
(H - \lambda_i I)v_{i+k-1} &= v_{i+k-2},
\end{aligned}$$

and the span$\{v_j, j = i, \ldots, i+k-1\}$ is an invariant subspace of H. ♡

Example 13.1 Let

$$A = \begin{bmatrix} -3 & 2 \\ -2 & 1 \end{bmatrix} \quad R = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

The eigenvalues of H are $1, 1, -1, -1$, and the corresponding eigenvector and generalized eigenvector of 1 are

$$v_1 = \begin{bmatrix} 1 \\ 2 \\ 2 \\ -2 \end{bmatrix}, \quad v_2 = \begin{bmatrix} -1 \\ -3/2 \\ 1 \\ 0 \end{bmatrix}.$$

The corresponding eigenvector and generalized eigenvector of -1 are

$$v_3 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}, \quad v_4 = \begin{bmatrix} 1 \\ 3/2 \\ 0 \\ 0 \end{bmatrix}.$$

All solutions of the Riccati equation under various combinations are given below:

- span$\{v_1, v_2\}$ is H-invariant: let $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = [v_1 \; v_2]$, then $X = X_2 X_1^{-1} = \begin{bmatrix} -10 & 6 \\ 6 & -4 \end{bmatrix}$ is a solution and $\sigma(A + RX) = \{1, 1\}$;
- span$\{v_1, v_3\}$ is H-invariant: let $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = [v_1 \; v_3]$, then $X = X_2 X_1^{-1} = \begin{bmatrix} -2 & 2 \\ 2 & -2 \end{bmatrix}$ is also a solution and $\sigma(A + RX) = \{1, -1\}$;

- $\text{span}\{v_3, v_4\}$ is H-invariant: let $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = [v_3 \; v_4]$, then $X = 0$ is a solution and $\sigma(A + RX) = \{-1, -1\}$;
- $\text{span}\{v_1, v_4\}$, $\text{span}\{v_2, v_3\}$, and $\text{span}\{v_2, v_4\}$ are not H-invariant subspaces. Readers can verify that the matrices constructed from those vectors are not solutions to the Riccati equation (13.1).

$\diamond$

Up to this point, we have said nothing about the structure of the solutions given by Theorem 13.1 and 13.2. The following theorem gives a sufficient condition for a Riccati solution to be hermitian (not necessarily real symmetric).

Theorem 13.3 *Let $\mathcal{V}$ be an n-dimensional H-invariant subspace and let $X_1, X_2 \in \mathbb{C}^{n\times n}$ be such that*

$$\mathcal{V} = \text{Im} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}.$$

Then $\lambda_i + \bar{\lambda}_j \neq 0$ for all $i, j = 1, \ldots, n$, $\lambda_i, \lambda_j \in \sigma(H|_\mathcal{V})$ implies that $X_1^* X_2$ is hermitian, i.e., $X_1^* X_2 = (X_1^* X_2)^*$. Furthermore, if X_1 is nonsingular, then $X = X_2 X_1^{-1}$ is hermitian.

Proof. Since $\mathcal{V}$ is an invariant subspace of H, there is a matrix representation Λ for $H|_\mathcal{V}$ such that $\sigma(\Lambda) = \sigma(H|_\mathcal{V})$ and

$$H \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \Lambda.$$

Pre-multiply this equation by $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}^* J$ to get

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}^* JH \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}^* J \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \Lambda. \tag{13.4}$$

Note that JH is hermitian (actually symmetric since H is real); therefore, the left-hand side of (13.4) is hermitian as well as the right-hand side:

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}^* J \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} \Lambda = \Lambda^* \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}^* J^* \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = -\Lambda^* \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}^* J \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

i.e.,

$$(-X_1^* X_2 + X_2^* X_1)\Lambda = -\Lambda^*(-X_1^* X_2 + X_2^* X_1).$$

This is a Lyapunov equation. Since $\lambda_i + \bar{\lambda}_j \neq 0$, the equation has a unique solution:

$$-X_1^* X_2 + X_2^* X_1 = 0.$$

This implies that $X_1^* X_2$ is hermitian.

That X is hermitian is easy to see by noting that $X = (X_1^{-1})^* (X_1^* X_2) X_1^{-1}$. □

Remark 13.2 It is clear from Example 13.1 that the condition $\lambda_i + \bar{\lambda}_j \neq 0$ is not necessary for the existence of a hermitian solution. ♡

The following theorem gives necessary and sufficient conditions for a solution to be real.

Theorem 13.4 *Let $\mathcal{V}$ be an n-dimensional H-invariant subspace, and let $X_1, X_2 \in \mathbb{C}^{n \times n}$ be such that X_1 is nonsingular and the columns of $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ form a basis of $\mathcal{V}$. Then $X := X_2 X_1^{-1}$ is real if and only if $\mathcal{V}$ is conjugate symmetric, i.e., $v \in \mathcal{V}$ implies that $\bar{v} \in \mathcal{V}$.*

Proof. ($\Leftarrow$) Since $\mathcal{V}$ is conjugate symmetric, there is a nonsingular matrix $P \in \mathbb{C}^{n \times n}$ such that

$$\begin{bmatrix} \bar{X}_1 \\ \bar{X}_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} P$$

where the over bar $\bar{X}$ denotes the complex conjugate. Therefore, $\bar{X} = \bar{X}_2 \bar{X}_1^{-1} = X_2 P (X_1 P)^{-1} = X_2 X_1^{-1} = X$ is real as desired.

($\Rightarrow$) Define $X := X_2 X_1^{-1}$. By assumption, $X \in \mathbb{R}^{n \times n}$ and

$$\operatorname{Im} \begin{bmatrix} I \\ X \end{bmatrix} = \mathcal{V};$$

therefore, $\mathcal{V}$ is conjugate symmetric. □

Example 13.2 This example is intended to show that there are non-real, non-hermitian solutions to equation (13.1). It is also designed to show that the condition $\lambda_i + \bar{\lambda}_j \neq 0, \forall i, j$, which excludes the possibility of having imaginary axis eigenvalues since if $\lambda_l = j\omega$ then $\lambda_l + \bar{\lambda}_l = 0$, is not necessary for the existence of a hermitian solution. Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 1 & 0 \end{bmatrix}, \quad R = \begin{bmatrix} -1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & -4 \end{bmatrix}, \quad Q = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

Then H has eigenvalues $\lambda_1 = 1 = -\lambda_2, \lambda_3 = \lambda_5 = j = -\lambda_4 = -\lambda_6$. It is easy to show that

$$X = \begin{bmatrix} 0 & 0.5+0.5j & 0.5-0.5j \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

satisfies equation (13.1) and that X is neither real nor hermitian. On the other hand,

$$X = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

is a real symmetric nonnegative definite solution corresponding to the eigenvalues -1, j, $-j$. $\diamond$

13.2 Stabilizing Solution and Riccati Operator

In this section, we discuss when a solution is stabilizing, i.e., $\sigma(A+RX) \subset \mathbb{C}_-$ and the properties of such solutions. This section is the central part of this chapter and is designed to be self-contained. Hence some of the material appearing in this section may be similar to that seen in the previous section.

Assume H has no eigenvalues on the imaginary axis. Then it must have n eigenvalues in Re $s < 0$ and n in Re $s > 0$. Consider the two n-dimensional spectral subspaces, $\mathcal{X}_-(H)$ and $\mathcal{X}_+(H)$: the former is the invariant subspace corresponding to eigenvalues in Re $s < 0$ and the latter corresponds to eigenvalues in Re $s > 0$. By finding a basis for $\mathcal{X}_-(H)$, stacking the basis vectors up to form a matrix, and partitioning the matrix, we get

$$\mathcal{X}_-(H) = \mathrm{Im} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

where $X_1, X_2 \in \mathbb{C}^{n\times n}$. ($X_1$ and X_2 can be chosen to be real matrices.) If X_1 is nonsingular or, equivalently, if the two subspaces

$$\mathcal{X}_-(H), \quad \mathrm{Im} \begin{bmatrix} 0 \\ I \end{bmatrix} \tag{13.5}$$

are complementary, we can set $X := X_2 X_1^{-1}$. Then X is uniquely determined by H, i.e., $H \longmapsto X$ is a function, which will be denoted *Ric*. We will take the domain of *Ric*, denoted *dom*(*Ric*), to consist of Hamiltonian matrices H with two properties: H has no eigenvalues on the imaginary axis and the two subspaces in (13.5) are complementary. For ease of reference, these will be called the *stability property* and the *complementarity* *property*, respectively. This solution will be called the *stabilizing solution*. Thus, $X = Ric(H)$ and

$$Ric: \ dom(Ric) \subset \mathbb{R}^{2n\times 2n} \longmapsto \mathbb{R}^{n\times n}.$$

The following well-known results give some properties of X as well as verifiable conditions under which H belongs to $dom(Ric)$.

Theorem 13.5 *Suppose $H \in dom(Ric)$ and $X = Ric(H)$. Then*

(i) X is real symmetric;

(ii) X satisfies the algebraic Riccati equation

$$A^*X + XA + XRX + Q = 0;$$

(iii) $A + RX$ is stable .

Proof. (i) Let X_1, X_2 be as above. It is claimed that

$$X_1^* X_2 \text{ is symmetric.} \tag{13.6}$$

To prove this, note that there exists a stable matrix H_- in $\mathbb{R}^{n\times n}$ such that

$$H\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} H_- \quad .$$

(H_- is a matrix representation of $H|_{\mathcal{X}_-(H)}$.) Pre-multiply this equation by

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}^* J$$

to get

$$\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}^* JH \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}^* J \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} H_- \quad . \tag{13.7}$$

Since JH is symmetric, so is the left-hand side of (13.7) and so is the right-hand side:

$$\begin{aligned} (-X_1^*X_2 + X_2^*X_1)H_- &= H_-^*(-X_1^*X_2 + X_2^*X_1)^* \\ &= -H_-^*(-X_1^*X_2 + X_2^*X_1). \end{aligned}$$

This is a Lyapunov equation. Since H_- is stable, the unique solution is

$$-X_1^*X_2 + X_2^*X_1 = 0.$$

This proves (13.6).

Since X_1 is nonsingular and $X = (X_1^{-1})^*(X_1^*X_2)X_1^{-1}$, X is symmetric.

(ii) Start with the equation

$$H \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} H_-$$

and post-multiply by X_1^{-1} to get

$$H \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} X_1 H_- X_1^{-1}. \tag{13.8}$$

Now pre-multiply by $[X \quad -I]$:

$$[X \quad -I] H \begin{bmatrix} I \\ X \end{bmatrix} = 0.$$

This is precisely the Riccati equation.

(iii) Pre-multiply (13.8) by $[I \quad 0]$ to get

$$A + RX = X_1 H_- X_1^{-1}.$$

Thus $A + RX$ is stable because H_- is. □

Now, we are going to state one of the main theorems of this section which gives the necessary and sufficient conditions for the existence of a unique stabilizing solution of (13.1) under certain restrictions on the matrix R.

Theorem 13.6 *Suppose H has no imaginary eigenvalues and R is either positive semi-definite or negative semi-definite. Then $H \in dom(Ric)$ if and only if (A, R) is stabilizable.*

Proof. ($\Leftarrow$) To prove that $H \in dom(Ric)$, we must show that

$$\mathcal{X}_-(H), \quad \mathrm{Im} \begin{bmatrix} 0 \\ I \end{bmatrix}$$

are complementary. This requires a preliminary step. As in the proof of Theorem 13.5 define X_1, X_2, H_- so that

$$\mathcal{X}_-(H) = \mathrm{Im} \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$$

$$H \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix} H_- \quad . \tag{13.9}$$

We want to show that X_1 is nonsingular, i.e., Ker $X_1 = 0$. First, it is claimed that Ker X_1 is H_--invariant. To prove this, let $x \in \operatorname{Ker} X_1$. Pre-multiply (13.9) by $[I \quad 0]$ to get

$$AX_1 + RX_2 = X_1 H_- \quad . \tag{13.10}$$

Pre-multiply by $x^* X_2^*$, post-multiply by x, and use the fact that $X_2^* X_1$ is symmetric (see (13.6)) to get

$$x^* X_2^* R X_2 x = 0.$$

Since R is semidefinite, this implies that $RX_2 x = 0$. Now post-multiply (13.10) by x to get $X_1 H_- x = 0$, i.e. $H_- x \in \operatorname{Ker} X_1$. This proves the claim.

Now to prove that X_1 is nonsingular, suppose, on the contrary, that Ker $X_1 \neq 0$. Then $H_-|_{\operatorname{Ker} X_1}$ has an eigenvalue, λ, and a corresponding eigenvector, x:

$$H_- x = \lambda x \tag{13.11}$$

$$\operatorname{Re} \lambda < 0, \quad 0 \neq x \in \operatorname{Ker} X_1.$$

Pre-multiply (13.9) by $[0 \quad I]$:

$$-QX_1 - A^* X_2 = X_2 H_- \quad . \tag{13.12}$$

Post-multiply the above equation by x and use (13.11):

$$(A^* + \lambda I) X_2 x = 0.$$

Recall that $RX_2 x = 0$, we have

$$x^* X_2^* [A + \overline{\lambda} I \quad R] = 0.$$

Then stabilizability of (A, R) implies $X_2 x = 0$. But if both $X_1 x = 0$ and $X_2 x = 0$, then $x = 0$ since $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ has full column rank, which is a contradiction.

($\Rightarrow$) This is obvious since $H \in \mathit{dom}(\mathit{Ric})$ implies that X is a stabilizing solution and that $A + RX$ is asymptotically stable. It also implies that (A, R) must be stabilizable. □

Theorem 13.7 *Suppose H has the form*

$$H = \begin{bmatrix} A & -BB^* \\ -C^*C & -A^* \end{bmatrix}.$$

Then $H \in \mathit{dom}(\mathit{Ric})$ iff (A, B) is stabilizable and (C, A) has no unobservable modes on the imaginary axis. Furthermore, $X = \mathit{Ric}(H) \geq 0$ if $H \in \mathit{dom}(\mathit{Ric})$, and $\operatorname{Ker}(X) = 0$ *if and only if (C, A) has no stable unobservable modes.*

Note that $\operatorname{Ker}(X) \subset \operatorname{Ker}(C)$, so that the equation $XM = C^*$ always has a solution for M, and a minimum F-norm solution is given by $X^\dagger C^*$.

Proof. It is clear from Theorem 13.6 that the stabilizability of (A, B) is necessary, and it is also sufficient if H has no eigenvalues on the imaginary axis. So we only need to show that, assuming (A, B) is stabilizable, H has no imaginary eigenvalues iff (C, A) has no unobservable modes on the imaginary axis. Suppose that $j\omega$ is an eigenvalue and $0 \neq \begin{bmatrix} x \\ z \end{bmatrix}$ is a corresponding eigenvector. Then

$$Ax - BB^* z = j\omega x$$

$$-C^*Cx - A^*z = j\omega z.$$

Re-arrange:

$$(A - j\omega I)x = BB^*z \tag{13.13}$$

$$-(A - j\omega I)^* z = C^*Cx. \tag{13.14}$$

Thus

$$\langle z, (A - j\omega I)x\rangle = \langle z, BB^*z\rangle = \|B^*z\|^2$$

$$-\langle x, (A - j\omega I)^*z\rangle = \langle x, C^*Cx\rangle = \|Cx\|^2$$

so $\langle x, (A - j\omega I)^*z\rangle$ is real and

$$-\|Cx\|^2 = \langle (A - j\omega I)x, z\rangle = \overline{\langle z, (A - j\omega I)x\rangle} = \|B^*z\|^2.$$

Therefore $B^*z = 0$ and $Cx = 0$. So from (13.13) and (13.14)

$$(A - j\omega I)x = 0$$

$$(A - j\omega I)^*z = 0.$$

Combine the last four equations to get

$$z^*[A - j\omega I \quad B] = 0$$

$$\begin{bmatrix} A - j\omega I \\ C \end{bmatrix} x = 0.$$

The stabilizability of (A, B) gives $z = 0$. Now it is clear that $j\omega$ is an eigenvalue of H iff $j\omega$ is an unobservable mode of (C, A).

Next, set $X := Ric(H)$. We'll show that $X \geq 0$. The Riccati equation is

$$A^*X + XA - XBB^*X + C^*C = 0$$

or equivalently

$$(A - BB^*X)^*X + X(A - BB^*X) + XBB^*X + C^*C = 0. \tag{13.15}$$

Noting that $A - BB^*X$ is stable (Theorem 13.5), we have

$$X = \int_0^\infty e^{(A-BB^*X)^*t}(XBB^*X + C^*C)e^{(A-BB^*X)t}dt. \tag{13.16}$$

Since $XBB^*X + C^*C$ is positive semi-definite, so is X.

Finally, we'll show that Ker$X$ is non-trivial if and only if (C, A) has stable unobservable modes. Let $x \in \text{Ker}X$, then $Xx = 0$. Pre-multiply (13.15) by x^* and post-multiply by x to get

$$Cx = 0.$$

Now post-multiply (13.15) again by x to get

$$XAx = 0.$$

We conclude that Ker(X) is an A-invariant subspace. Now if Ker$(X) \neq 0$, then there is a $0 \neq x \in \text{Ker}(X)$ and a λ such that $\lambda x = Ax = (A - BB^*X)x$ and $Cx = 0$. Since $(A - BB^*X)$ is stable, Re$\lambda < 0$; thus λ is a stable unobservable mode. Conversely, suppose (C, A) has an unobservable stable mode λ, i.e., there is an x such that $Ax = \lambda x, Cx = 0$. By pre-multiplying the Riccati equation by x^* and post-multiplying by x, we get

$$2\text{Re}\lambda x^*Xx - x^*XBB^*Xx = 0.$$

Hence $x^*Xx = 0$, i.e., X is singular. □

Example 13.3 This example shows that the observability of (C, A) is not necessary for the existence of a positive definite stabilizing solution. Let

$$A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}, \; B = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \; C = \begin{bmatrix} 0 & 0 \end{bmatrix}.$$

Then (A, B) is stabilizable, but (C, A) is not detectable. However,

$$X = \begin{bmatrix} 18 & -24 \\ -24 & 36 \end{bmatrix} > 0$$

is the stabilizing solution. ◇

Corollary 13.8 *Suppose that (A, B) is stabilizable and (C, A) is detectable. Then the Riccati equation*

$$A^*X + XA - XBB^*X + C^*C = 0$$

has a unique positive semidefinite solution. Moreover, the solution is stabilizing.

Proof. It is obvious from the above theorem that the Riccati equation has a unique stabilizing solution and that the solution is positive semidefinite. Hence we only need to show that any positive semidefinite solution $X \geq 0$ must also be stabilizing. Then by the uniqueness of the stabilizing solution, we can conclude that there is only one positive semidefinite solution. To achieve that goal, let us assume that $X \geq 0$ satisfies the Riccati equation but that it is not stabilizing. First rewrite the Riccati equation as

$$(A - BB^*X)^*X + X(A - BB^*X) + XBB^*X + C^*C = 0 \tag{13.17}$$

and let λ and x be an unstable eigenvalue and the corresponding eigenvector of $A - BB^*X$, respectively, i.e.,

$$(A - BB^*X)x = \lambda x.$$

Now pre-multiply and postmultiply equation (13.17) by x^* and x, respectively, and we have

$$(\bar{\lambda} + \lambda)x^*Xx + x^*(XBB^*X + CC)x = 0.$$

This implies

$$B^*Xx = 0, \quad Cx = 0$$

since $\mathrm{Re}(\lambda) \geq 0$ and $X \geq 0$. Finally, we arrive at

$$Ax = \lambda x, \quad Cx = 0$$

i.e., (C, A) is not detectable, which is a contradiction. Hence $\mathrm{Re}(\lambda) < 0$, i.e., $X \geq 0$ is the stabilizing solution. □

Lemma 13.9 *Suppose D has full column rank and let $R = D^*D > 0$; then the following statements are equivalent:*

(i) $\begin{bmatrix} A - j\omega I & B \\ C & D \end{bmatrix}$ *has full column rank for all ω.*

(ii) $\big((I - DR^{-1}D^*)C,\ A - BR^{-1}D^*C\big)$ *has no unobservable modes on $j\omega$-axis.*

Proof. Suppose $j\omega$ is an unobservable mode of $\big((I - DR^{-1}D^*)C,\ A - BR^{-1}D^*C\big)$; then there is an $x \neq 0$ such that

$$(A - BR^{-1}D^*C)x = j\omega x, \quad (I - DR^{-1}D^*)Cx = 0$$

i.e.,

$$\begin{bmatrix} A - j\omega I & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ -R^{-1}D^*C & I \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = 0.$$

But this implies that

$$\begin{bmatrix} A - j\omega I & B \\ C & D \end{bmatrix} \tag{13.18}$$

does not have full column rank. Conversely, suppose (13.18) does not have full column rank for some ω; then there exists $\begin{bmatrix} u \\ v \end{bmatrix} \neq 0$ such that

$$\begin{bmatrix} A - j\omega I & B \\ C & D \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = 0.$$

Now let

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} I & 0 \\ -R^{-1}D^*C & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Then

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} I & 0 \\ R^{-1}D^*C & I \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \neq 0$$

and

$$(A - BR^{-1}D^*C - j\omega I)x + By = 0 \tag{13.19}$$

$$(I - DR^{-1}D^*)Cx + Dy = 0. \tag{13.20}$$

Pre-multiply (13.20) by D^* to get $y = 0$. Then we have

$$(A - BR^{-1}D^*C)x = j\omega x, \quad (I - DR^{-1}D^*)Cx = 0$$

i.e., $j\omega$ is an unobservable mode of $\big((I - DR^{-1}D^*)C,\, A - BR^{-1}D^*C\big)$. □

Remark 13.3 If D is not square, then there is a $D_\perp$ such that $\begin{bmatrix} D_\perp & DR^{-1/2} \end{bmatrix}$ is unitary and that $D_\perp D_\perp^* = I - DR^{-1}D^*$. Hence, in some cases we will write the condition (ii) in the above lemma as $(D_\perp^* C, A - BR^{-1}D^*C)$ having no imaginary unobservable modes. Of course, if D is square, the condition is simplified to $A - BR^{-1}D^*C$ with no imaginary eigenvalues. Note also that if $D^*C = 0$, condition (ii) becomes (C, A) with no imaginary unobservable modes. ♡

Corollary 13.10 *Suppose D has full column rank and denote $R = D^*D > 0$. Let H have the form*

$$\begin{aligned} H &= \begin{bmatrix} A & 0 \\ -C^*C & -A^* \end{bmatrix} - \begin{bmatrix} B \\ -C^*D \end{bmatrix} R^{-1} \begin{bmatrix} D^*C & B^* \end{bmatrix} \\ &= \begin{bmatrix} A - BR^{-1}D^*C & -BR^{-1}B^* \\ -C^*(I - DR^{-1}D^*)C & -(A - BR^{-1}D^*C)^* \end{bmatrix}. \end{aligned}$$

Then $H \in dom(Ric)$ *iff* (A, B) *is stabilizable and* $\begin{bmatrix} A - j\omega I & B \\ C & D \end{bmatrix}$ *has full column rank for all* ω. *Furthermore,* $X = Ric(H) \geq 0$ *if* $H \in dom(Ric)$, *and* $Ker(X) = 0$ *if and only if* $(D_\perp^* C, A - BR^{-1}D^*C)$ *has no stable unobservable modes.*

Proof. This is the consequence of the Lemma 13.9 and Theorem 13.7. □

Remark 13.4 It is easy to see that the detectability (observability) of $(D_\perp^* C, A - BR^{-1}D^*C)$ implies the detectability (observability) of (C, A); however, the converse is in general not true. Hence the existence of a stabilizing solution to the Riccati equation in the above corollary is not guaranteed by the stabilizability of (A, B) and detectability of (C, A). Furthermore, even if a stabilizing solution exists, the positive definiteness of the solution is not guaranteed by the observability of (C, A) unless $D^*C = 0$. As an example, consider

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \; B = \begin{bmatrix} 0 \\ -1 \end{bmatrix}, \; C = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \; D = \begin{bmatrix} 1 \\ 0 \end{bmatrix}.$$

Then (C, A) is observable, (A, B) is controllable, and

$$A - BD^*C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad D_\perp^* C = \begin{bmatrix} 0 & 0 \end{bmatrix}.$$

A Riccati equation with the above data has a nonnegative definite stabilizing solution since $(D_\perp^* C, A - BR^{-1}D^*C)$ has no unobservable modes on the imaginary axis. However, the solution is not positive definite since $(D_\perp^* C, A - BR^{-1}D^*C)$ has a stable unobservable mode. On the other hand, if the B matrix is changed to

$$B = \begin{bmatrix} 0 \\ 1 \end{bmatrix},$$

then the corresponding Riccati equation has no stabilizing solution since, in this case, $(A - BD^*C)$ has eigenvalues on the imaginary axis although (A, B) is controllable and (C, A) is observable. ♡

13.3 Extreme Solutions and Matrix Inequalities

We have shown in the previous sections that given a Riccati equation, there are generally many possible solutions. Among all solutions, we are most interested in those which are real, symmetric, and, in particular, the stabilizing solutions. There is another class of solutions which are interesting; they are called *extreme* (maximal or minimal) solutions. Some properties of the extreme solutions will be studied in this section. The connections between the extreme solutions and the stabilizing solutions will also be established in this section. To illustrate the idea, let us look at an example.

Example 13.4 Let

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}, \; B = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \; C = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}.$$

The eigenvalues of matrix $H = \begin{bmatrix} A & -BB^* \\ -C^*C & -A^* \end{bmatrix}$ are

$$\lambda_1 = 1, \; \lambda_2 = -1, \; \lambda_3 = 2, \; \lambda_4 = -2, \; \lambda_5 = -3, \; \lambda_6 = 3,$$

and their corresponding eigenvectors are

$$v_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_2 = \begin{bmatrix} 3 \\ 2 \\ -3 \\ 6 \\ 0 \\ 0 \end{bmatrix}, v_3 = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_4 = \begin{bmatrix} 4 \\ 3 \\ -12 \\ 0 \\ 12 \\ 0 \end{bmatrix}, v_5 = \begin{bmatrix} 0 \\ 0 \\ 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}, v_6 = \begin{bmatrix} -3 \\ -6 \\ -1 \\ 0 \\ 0 \\ 6 \end{bmatrix}.$$

There are four distinct nonnegative definite symmetric solutions depending on the chosen invariant subspaces:

- $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = [v_1 \; v_3 \; v_5]; \; Y_1 = X_2 X_1^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};$
- $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = [v_2 \; v_3 \; v_5]; \; Y_2 = X_2 X_1^{-1} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix};$
- $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = [v_1 \; v_4 \; v_5]; \; Y_3 = X_2 X_1^{-1} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 0 \end{bmatrix};$
- $\begin{bmatrix} X_1 \\ X_2 \end{bmatrix} = [v_2 \; v_4 \; v_5]; \; Y_4 = X_2 X_1^{-1} = \begin{bmatrix} 18 & -24 & 0 \\ -24 & 36 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$

These solutions can be ordered as $Y_4 \geq Y_i \geq Y_1$, $i = 2, 3$. Of course, this is only a partial ordering since Y_2 and Y_3 are not comparable. Note also that only Y_4 is a stabilizing solution, i.e., $A - BB^*Y_4$ is stable. Furthermore, Y_4 and Y_1 are the "maximal" and "minimal" solutions, respectively. $\diamond$

The partial ordering concept shown in Example 13.4 can be stated in a much more general setting. To do that, again consider the Riccati equation (13.1). We shall call a hermitian solution X_+ of (13.1) a *maximal solution* if $X_+ \geq X$ for all hermitian solutions X of (13.1). Similarly, we shall call a hermitian solution X_- of (13.1) a *minimal solution* if $X_- \leq X$ for all hermitian solutions X of (13.1). Clearly, maximal and minimal solutions are unique if they exist.

To study the properties of the maximal and minimal solutions, we shall introduce the following quadratic matrix:

$$\mathcal{Q}(X) := A^*X + XA + XRX + Q. \tag{13.21}$$

Theorem 13.11 *Assume $R \leq 0$ and assume there is a hermitian matrix $X = X^*$ such that $\mathcal{Q}(X) \geq 0$.*

(i) If (A, R) is stabilizable, then there exists a unique maximal solution X_+ to the Riccati equation (13.1). Furthermore,

$$X_+ \geq X, \ \forall X \text{ such that } \mathcal{Q}(X) \geq 0$$

and $\sigma(A + RX_+) \subset \overline{\mathbb{C}}_-$.

(ii) If $(-A, R)$ is stabilizable, then there exists a unique minimal solution X_- to the Riccati equation (13.1). Furthermore,

$$X_- \leq X, \ \forall X \text{ such that } \mathcal{Q}(X) \geq 0$$

and $\sigma(A + RX_-) \subset \overline{\mathbb{C}}_+$.

(iii) If (A, R) is controllable, then both X_+ and X_- exist. Furthermore, $X_+ > X_-$ iff $\sigma(A + RX_+) \subset \mathbb{C}_-$ iff $\sigma(A + RX_-) \subset \mathbb{C}_+$. In this case,

$$\begin{aligned} X_+ - X_- &= \left(\int_0^\infty e^{(A+RX_+)t} R e^{(A+RX_+)^*t} dt \right)^{-1} \\ &= \left(\int_{-\infty}^0 e^{(A+RX_-)t} R e^{(A+RX_-)^*t} dt \right)^{-1}. \end{aligned}$$

(iv) If $\mathcal{Q}(X) > 0$, the results in (i) and (ii) can be respectively strengthened to $X_+ > X$, $\sigma(A + RX_+) \subset \mathbb{C}_-$, and $X_- < X$, $\sigma(A + RX_-) \subset \mathbb{C}_+$.

Proof. Let $R = -BB^*$ for some B. Note the fact that (A, R) is stabilizable (controllable) iff (A, B) is.

(i): Let X be such that $\mathcal{Q}(X) \geq 0$. Since (A, B) is stabilizable, there is an F_0 such that

$$A_0 := A + BF_0$$

is stable. Now let X_0 be the unique solution to the Lyapunov equation

$$X_0 A_0 + A_0^* X_0 + F_0^* F_0 + Q = 0.$$

Then X_0 is hermitian. Define

$$\hat{F}_0 := F_0 + B^* X,$$

and we have the following equation:

$$(X_0 - X)A_0 + A_0^*(X_0 - X) = -\hat{F}_0^* \hat{F}_0 - \mathcal{Q}(X) \leq 0.$$

The stability of A_0 implies that

$$X_0 \geq X.$$

Starting with X_0, we shall define a non-increasing sequence of hermitian matrices $\{X_i\}$. Associated with $\{X_i\}$, we shall also define a sequence of stable matrices $\{A_i\}$ and a sequence of matrices $\{F_i\}$. Assume inductively that we have already defined matrices $\{X_i\}$, $\{A_i\}$, and $\{F_i\}$ for i up to $n-1$ such that X_i is hermitian and

$$X_0 \geq X_1 \geq \cdots \geq X_{n-1} \geq X,$$

$$A_i = A + BF_i, \text{ is stable}, \ i = 0, \ldots, n-1;$$

$$F_i = -B^* X_{i-1}, \ i = 1, \ldots, n-1;$$

$$X_i A_i + A_i^* X_i = -F_i^* F_i - Q, \ i = 0, 1, \ldots, n-1. \tag{13.22}$$

Next, introduce

$$F_n = -B^* X_{n-1},$$

$$A_n = A + BF_n.$$

First we show that A_n is stable. Then, using (13.22), with $i = n$, we define a hermitian matrix X_n with $X_{n-1} \geq X_n \geq X$. Now using (13.22), with $i = n-1$, we get

$$X_{n-1} A_n + A_n^* X_{n-1} + Q + F_n^* F_n + (F_n - F_{n-1})^*(F_n - F_{n-1}) = 0. \tag{13.23}$$

Let

$$\hat{F}_n := F_n + B^* X;$$

then

$$(X_{n-1} - X)A_n + A_n^*(X_{n-1} - X) = -\mathcal{Q}(X) - \hat{F}_n^* \hat{F}_n - (F_n - F_{n-1})^*(F_n - F_{n-1}). \tag{13.24}$$

Now assume that A_n is not stable, i.e., there exists an λ with $\mathrm{Re}\lambda \geq 0$ and $x \neq 0$ such that $A_n x = \lambda x$. Then pre-multiply (13.24) by x^* and postmultiply by x, and we have

$$2\mathrm{Re}\lambda x^*(X_{n-1} - X)x = -x^*\{\mathcal{Q}(X) + \hat{F}_n^* \hat{F}_n + (F_n - F_{n-1})^*(F_n - F_{n-1})\}x.$$

Since it is assumed $X_{n-1} \geq X$, each term on the right-hand side of the above equation has to be zero. So we have

$$x^*(F_n - F_{n-1})^*(F_n - F_{n-1})x = 0.$$

This implies

$$(F_n - F_{n-1})x = 0.$$

But now

$$A_{n-1}x = (A + BF_{n-1})x = (A + BF_n)x = A_n x = \lambda x,$$

which is a contradiction with the stability of A_{n-1}. Hence A_n is stable as well.

Now we introduce X_n as the unique solution of the Lyapunov equation

$$X_n A_n + A^* X_n = -F_n^* F_n - Q. \tag{13.25}$$

Then X_n is hermitian. Next, we have

$$(X_n - X)A_n + A_n^*(X_n - X) = -\mathcal{Q}(X) - \hat{F}_n^* \hat{F}_n \leq 0,$$

and, by using (13.23),

$$(X_{n-1} - X_n)A_n + A_n^*(X_{n-1} - X_n) = -(F_n - F_{n-1})^*(F_n - F_{n-1}) \leq 0.$$

Since A_n is stable, we have

$$X_{n-1} \geq X_n \geq X.$$

We have a non-increasing sequence $\{X_i\}$, and the sequence is bounded below by $X_i \geq X$. Hence the limit

$$X_f := \lim_{n \to \infty} X_n$$

exists and is hermitian, and we have $X_f \geq X$. Passing the limit $n \to \infty$ in (13.25), we get $\mathcal{Q}(X_f) = 0$. So X_f is a solution of (13.1). Since X is an arbitrary element satisfying $\mathcal{Q}(X) \geq 0$ and X_f is independent of the choice of X, we have

$$X_f \geq X, \ \forall \ X \text{ such that } \mathcal{Q}(X) \geq 0.$$

In particular, X_f is the maximal solution of the Riccati equation (13.1), i.e., $X_f = X_+$.

To establish the stability property of the maximal solution, note that A_n is stable for any n. Hence, in the limit, the eigenvalues of

$$A - BB^* X_f$$

will have non-positive real parts. The uniqueness follows from the fact that the maximal solution is unique.

(ii): The results follow by the following substitutions in the proof of part (i):

$$A \leftarrow -A; \ X \leftarrow -X; \ X_+ \leftarrow -X_-.$$

(iii): The existence of X_+ and X_- follows from (i) and (ii). Let $A_+ := A + RX_+$; we now show that $X_+ - X_- > 0$ iff $\sigma(A_+) \subset \mathbb{C}_-$. It is easy to verify that

$$A_+^*(X_+ - X_-) + (X_+ - X_-)A_+ - (X_+ - X_-)R(X_+ - X_-) = 0.$$

($\Rightarrow$): Suppose $X_+ - X_- > 0$; then $(A_+, R(X_+ - X_-))$ is controllable. Therefore, from Lyapunov theorem, we conclude that A_+ is stable.

($\Leftarrow$): Assuming now that A_+ is stable and that X is any other solution to the Riccati equation (13.1),

$$A_+^*(X_+ - X) + (X_+ - X)A_+ - (X_+ - X)R(X_+ - X) = 0.$$

This equation has an invertible solution

$$X_+ - \bar{X} = \left(-\int_0^\infty e^{(A+RX_+)t} R e^{(A+RX_+)^* t} dt\right)^{-1} > 0.$$

A simple rearrangement of terms gives

$$(A + R\bar{X})^*(X_+ - \bar{X}) + (X_+ - \bar{X})(A + R\bar{X}) + (X_+ - \bar{X})R(X_+ - \bar{X}) = 0.$$

Since $X_+ - \bar{X} > 0$, we conclude $\sigma(A + R\bar{X}) \subset \mathbb{C}_+$. This in turn implies $\bar{X} = X_-$. Therefore, $X_+ > X_-$. That $X_+ - X_- > 0$ iff $\sigma(A + RX_-) \subset \mathbb{C}_+$ follows by analogy.

(iv): We shall only show the case for X_+; the case for X_- follows by analogy. Note that from (i) we have

$$(X_+ - X)A_+ + A_+^*(X_+ - X) = -\mathcal{Q}(X) + (X_+ - X)R(X_+ - X) < 0 \tag{13.26}$$

and $X_+ - X \geq 0$. Now suppose $X_+ - X$ is singular and there is an $x \neq 0$ such that $(X_+ - X)x = 0$. By pre-multiplying (13.26) by x^* and post-multiplying by x, we get $x^*\mathcal{Q}(X)x = 0$, a contradiction. The stability of $A + RX_+$ then follows from the Lyapunov theorem. □

Remark 13.5 The proof given above also gives an iterative procedure to compute the maximal and minimal solutions. For example, to find the maximal solution, the following procedures can be used:

(i) find F_0 such that $A_0 = A + BF_0$ is stable;

(ii) solve X_i: $X_iA_i + A_i^*X_i + F_i^*F_i + Q = 0$;

(iii) if $\|X_i - X_{i-1}\| \leq \epsilon$ =specified accuracy, stop. Otherwise go to (iv);

(iv) let $F_{i+1} = -B^*X_i$ and $A_{i+1} = A + BF_{i+1}$ go to (ii).

This procedure will converge to the stabilizing solution if the solution exists. ♡

Corollary 13.12 *Let $R \leq 0$ and suppose (A, R) is controllable and X_1, X_2 are two solutions to the Riccati equation (13.1). Then $X_1 > X_2$ implies that $X_+ = X_1$, $X_- = X_2$, $\sigma(A + RX_1) \subset \mathbb{C}_-$ and that $\sigma(A + RX_2) \subset \mathbb{C}_+$.*

The following example illustrates that the stabilizability of (A, R) is not sufficient to guarantee the existence of a minimal hermitian solution of equation (13.1).

Example 13.5 Let

$$A = \begin{bmatrix} 0 & 0 \\ 0 & -1 \end{bmatrix}, \; R = \begin{bmatrix} -1 & 0 \\ 0 & 0 \end{bmatrix}, \; Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Then it can be shown that all the hermitian solutions of (13.1) are given by

$$\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \; \begin{bmatrix} -1 & \alpha \\ \bar{\alpha} & -\frac{1}{2}|\alpha|^2 \end{bmatrix}, \; \alpha \in \mathbb{C}.$$

The maximal solution is clearly

$$X_+ = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix};$$

however, there is no minimal solution. $\diamond$

The Riccati equation appeared in $\mathcal{H}_\infty$ control, which will be considered in the later part of this book, often has $R \geq 0$. However, these conditions are only a dual of the above theorem.

Corollary 13.13 *Assume that $R \geq 0$ and that there is a hermitian matrix $X = X^*$ such that $\mathcal{Q}(X) \leq 0$.*

(i) If (A, R) is stabilizable, then there exists a unique minimal solution X_- to the Riccati equation (13.1). Furthermore,

$$X_- \leq X, \; \forall X \text{ such that } \mathcal{Q}(X) \leq 0$$

and $\sigma(A + RX_-) \subset \overline{\mathbb{C}}_-$.

(ii) If $(-A, R)$ is stabilizable, then there exists a unique maximal solution X_+ to the Riccati equation (13.1). Furthermore,

$$X_+ \geq X, \; \forall X \text{ such that } \mathcal{Q}(X) \leq 0$$

and $\sigma(A + RX_+) \subset \overline{\mathbb{C}}_+$.

(iii) If (A, R) is controllable, then both X_+ and X_- exist. Furthermore, $X_+ > X_-$ iff $\sigma(A + RX_-) \subset \mathbb{C}_-$ iff $\sigma(A + RX_+) \subset \mathbb{C}_+$. In this case,

$$\begin{aligned} X_+ - X_- &= \left(\int_0^\infty e^{(A+RX_-)t} R e^{(A+RX_-)^* t} dt\right)^{-1} \\ &= \left(\int_{-\infty}^0 e^{(A+RX_+)t} R e^{(A+RX_+)^* t} dt\right)^{-1}. \end{aligned}$$

(iv) If $\mathcal{Q}(X) < 0$, the results in (i) and (ii) can be respectively strengthened to $X_+ > X$, $\sigma(A + RX_+) \subset \mathbb{C}_+$, and $X_- < X$, $\sigma(A + RX_-) \subset \mathbb{C}_-$.

Proof. The proof is similar to the proof for Theorem 13.11. □

Theorem 13.11 can be used to derive some comparative results for some Riccati equations. More specifically, let

$$H_s := \begin{bmatrix} A - BR_s^{-1}S^* & -BR_s^{-1}B^* \\ -P + SR_s^{-1}S^* & -(A - BR_s^{-1}S^*)^* \end{bmatrix}$$

and

$$\tilde{H}_s := \begin{bmatrix} \tilde{A} - \tilde{B}\tilde{R}_s^{-1}\tilde{S}^* & -\tilde{B}\tilde{R}_s^{-1}\tilde{B}^* \\ -\tilde{P} + \tilde{S}\tilde{R}_s^{-1}\tilde{S}^* & -(\tilde{A} - \tilde{B}\tilde{R}_s^{-1}\tilde{S}^*)^* \end{bmatrix}$$

where $P, \tilde{P}, R_s$, and $\tilde{R}_s$ are real symmetric and $R_s > 0$, $\tilde{R}_s > 0$. We shall also make use of the following matrices:

$$T := \begin{bmatrix} P & S \\ S^* & R_s \end{bmatrix}, \ \tilde{T} := \begin{bmatrix} \tilde{P} & \tilde{S} \\ \tilde{S}^* & \tilde{R}_s \end{bmatrix}.$$

We denote by X_+ and $\tilde{X}_+$ the maximal solution to the Riccati equation associated with H_s and $\tilde{H}_s$, respectively:

$$(A - BR_s^{-1}S^*)^*X + X(A - BR_s^{-1}S^*) - XBR_s^{-1}B^*X + (P - SR_s^{-1}S^*) = 0 \quad (13.27)$$

$$(\tilde{A} - \tilde{B}\tilde{R}_s^{-1}\tilde{S}^*)^*\tilde{X} + \tilde{X}(\tilde{A} - \tilde{B}\tilde{R}_s^{-1}\tilde{S}^*) - \tilde{X}\tilde{B}\tilde{R}_s^{-1}\tilde{B}^*\tilde{X} + (\tilde{P} - \tilde{S}\tilde{R}_s^{-1}\tilde{S}^*) = 0. \quad (13.28)$$

Recall that JH_s and $J\tilde{H}_s$ are hermitian where J is defined as

$$J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}.$$

Let

$$\begin{aligned} K := JH_s &= \begin{bmatrix} P - SR_s^{-1}S^* & (A - BR_s^{-1}S^*)^* \\ A - BR_s^{-1}S^* & -BR_s^{-1}B^* \end{bmatrix} \\ \tilde{K} := J\tilde{H}_s &= \begin{bmatrix} \tilde{P} - \tilde{S}\tilde{R}_s^{-1}\tilde{S}^* & (\tilde{A} - \tilde{B}\tilde{R}_s^{-1}\tilde{S}^*)^* \\ \tilde{A} - \tilde{B}\tilde{R}_s^{-1}\tilde{S}^* & -\tilde{B}\tilde{R}_s^{-1}\tilde{B}^* \end{bmatrix}. \end{aligned}$$

Theorem 13.14 *Suppose (A, B) and $(\tilde{A}, \tilde{B})$ are stabilizable.*

(i) Assume that (13.28) has a hermitian solution and that $K \geq \tilde{K}$ ($K > \tilde{K}$); then X_+ and $\tilde{X}_+$ exist, and $X_+ \geq \tilde{X}_+$ ($X_+ > \tilde{X}_+$).

(ii) Let $A = \tilde{A}$, $B = \tilde{B}$, and $T \geq \tilde{T}$ ($T > \tilde{T}$). Assume that (13.28) has a hermitian solution. Then X_+ and $\tilde{X}_+$ exist, and $X_+ \geq \tilde{X}_+$ ($X_+ > \tilde{X}_+$).

(iii) If $T \geq 0$ ($T > 0$), then X_+ exists, and $X_+ \geq 0$ ($X_+ > 0$).

Proof. (i): Let X be a hermitian solution of (13.28); then we have

$$\begin{bmatrix} I \\ X \end{bmatrix}^* \tilde{K} \begin{bmatrix} I \\ X \end{bmatrix} = 0.$$

Then, since $K \geq \tilde{K}$, we have

$$\mathcal{Q}_s(X) := \begin{bmatrix} I \\ X \end{bmatrix}^* K \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix}^* (K - \tilde{K}) \begin{bmatrix} I \\ X \end{bmatrix} \geq 0.$$

Now we use Theorem 13.11 to obtain the existence of X_+, and $X_+ \geq X$. Since X is arbitrary, we get $X_+ \geq X$ for all hermitian solutions X of (13.28). The existence of $\tilde{X}_+$ follows immediately by applying Theorem 13.11. Moreover,

$$X_+ \geq \tilde{X}_+.$$

(ii): Let $A = \tilde{A}$, $B = \tilde{B}$ and let X be a hermitian solution of (13.28). Denote $L = -R_s^{-1}(S^* + B^*X)$ and $\tilde{L} = -\tilde{R}_s^{-1}(\tilde{S}^* + B^*X)$. Then

$$\mathcal{Q}_s(X) = \begin{bmatrix} I \\ X \end{bmatrix}^* K \begin{bmatrix} I \\ X \end{bmatrix} = A^*X + XA - L^*R_sL + P$$

while (13.28) becomes

$$A^*X + XA - \tilde{L}^*\tilde{R}_s\tilde{L} + \tilde{P} = 0.$$

It is easy to show that

$$\begin{aligned} \mathcal{Q}_s(X) &= \begin{bmatrix} I \\ X \end{bmatrix}^* (K - \tilde{K}) \begin{bmatrix} I \\ X \end{bmatrix} \\ &= P - \tilde{P} - L^*R_sL + \tilde{L}^*\tilde{R}_s\tilde{L} \\ &= \begin{bmatrix} I \\ L \end{bmatrix}^* (T - \tilde{T}) \begin{bmatrix} I \\ L \end{bmatrix} + (L - \tilde{L})^*\tilde{R}_s(L - \tilde{L}) \geq 0. \end{aligned}$$

Now as in part (i), there exist X_+ and $\tilde{X}_+$, and $X_+ \geq \tilde{X}_+$.

(iii): The condition $T \geq 0$ implies $P - SR_s^{-1}S^* \geq 0$, so we have $\mathcal{Q}_s(0) = P - SR_s^{-1}S^* \geq 0$. Apply Theorem 13.11 to get the existence of X_+, and $X_+ \geq 0$. □

13.4 Spectral Factorizations

Let A, B, P, S, R be real matrices of compatible dimensions such that $P = P^*$, $R = R^*$, and define a parahermitian rational matrix function

$$\begin{aligned}\Phi(s) &= R + S^*(sI - A)^{-1}B + B^*(-sI - A^*)^{-1}S + B^*(-sI - A^*)^{-1}P(sI - A)^{-1}B \\ &= \begin{bmatrix} B^*(-sI - A^*)^{-1} & I \end{bmatrix} \begin{bmatrix} P & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix}. \qquad (13.29)\end{aligned}$$

Lemma 13.15 *Suppose R is nonsingular and either one of the following assumptions is satisfied:*

(A1) A has no eigenvalues on $j\omega$-axis;

(A2) P is sign definite, i.e., $P \geq 0$ or $P \leq 0$, (A, B) has no uncontrollable modes on $j\omega$-axis, and (P, A) has no unobservable modes on the $j\omega$-axis.

Then the following statements are equivalent:

(i) $\Phi(j\omega_0)$ *is singular for some* $0 \leq \omega_0 \leq \infty$.

(ii) The Hamiltonian matrix

$$H = \begin{bmatrix} A - BR^{-1}S^* & -BR^{-1}B^* \\ -(P - SR^{-1}S^*) & -(A - BR^{-1}S^*)^* \end{bmatrix}$$

has an eigenvalue at $j\omega_0$.

Proof. $(i) \Rightarrow (ii)$: Let

$$\Phi(s) = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right] := \left[\begin{array}{cc|c} A & 0 & B \\ -P & -A^* & -S \\ \hline S^* & B^* & R \end{array}\right].$$

Then $H = \hat{A} - \hat{B}\hat{D}^{-1}\hat{C}$ and

$$\Phi^{-1}(s) = \left[\begin{array}{c|c} H & \begin{bmatrix} -BR^{-1} \\ SR^{-1} \end{bmatrix} \\ \hline \begin{bmatrix} R^{-1}S^* & R^{-1}B^* \end{bmatrix} & R^{-1} \end{array}\right].$$

If $\Phi(j\omega_0)$ is singular, then $j\omega_0$ is a zero of $\Phi(s)$. Hence $j\omega_0$ is a pole of $\Phi^{-1}(s)$, and $j\omega_0$ is an eigenvalue of H.

(*ii*) $\Rightarrow$ (*i*): Suppose $j\omega_0$ is an eigenvalue of H but is not a pole of $\Phi^{-1}(s)$. Then $j\omega_0$ must be either an unobservable mode of $(\begin{bmatrix} R^{-1}S^* & R^{-1}B^* \end{bmatrix}, H)$ or an uncontrollable mode of $(H, \begin{bmatrix} -BR^{-1} \\ SR^{-1} \end{bmatrix})$. Now suppose $j\omega_0$ is an unobservable mode of $(\begin{bmatrix} R^{-1}S^* & R^{-1}B^* \end{bmatrix}, H)$. Then there exists an $x_0 = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \neq 0$ such that

$$Hx_0 = j\omega_0 x_0, \quad \begin{bmatrix} R^{-1}S^* & R^{-1}B^* \end{bmatrix} x_0 = 0.$$

These equations can be simplified to

$$(j\omega_0 I - A)x_1 = 0 \tag{13.30}$$

$$(j\omega_0 I + A^*)x_2 = -Px_1 \tag{13.31}$$

$$S^* x_1 + B^* x_2 = 0. \tag{13.32}$$

We now consider two cases under the different assumptions:

(a) If assumption (A1) is true, then $x_1 = 0$ from (13.30), and this, in turn, implies $x_2 = 0$ from (13.31), which is a contradiction.

(b) If assumption (A2) is true, since (13.30) implies $x_1^*(j\omega_0 I + A^*) = 0$, from (13.31) we have $x_1^* P x_1 = 0$. This gives $Px_1 = 0$ since P is sign definite ($P \geq 0$ or $P \leq 0$). This implies, along with (13.30) that $j\omega_0$ is an unobservable mode of (P, A) if $x_1 \neq 0$. On the other hand, if $x_1 = 0$, then (13.31) and (13.32) imply that (A, B) has an uncontrollable mode at $j\omega_0$, again a contradiction.

Similarly, a contradiction will also be derived if $j\omega_0$ is assumed to be an uncontrollable mode of $(H, \begin{bmatrix} -BR^{-1} \\ SR^{-1} \end{bmatrix})$. □

Corollary 13.16 *Suppose $R > 0$ and either one of the assumptions (A1) or (A2) defined in Lemma 13.15 is true. Then the following statements are equivalent:*

(i) $\Phi(j\omega) > 0$ *for all* $0 \leq \omega \leq \infty$.

(ii) The Hamiltonian matrix

$$H = \begin{bmatrix} A - BR^{-1}S^* & -BR^{-1}B^* \\ -(P - SR^{-1}S^*) & -(A - BR^{-1}S^*)^* \end{bmatrix}$$

has no eigenvalue on $j\omega$-axis.

Proof. $(i) \Rightarrow (ii)$ follows easily from Lemma 13.15. To prove $(ii) \Rightarrow (i)$, note that $\Phi(j\infty) = R > 0$ and $\det \Phi(j\omega) \neq 0$ for all ω from Lemma 13.15. Then the continuity of $\Phi(j\omega)$ gives $\Phi(j\omega) > 0$. □

Lemma 13.17 *Suppose A is stable and $P \leq 0$. Then*

(i) the matrix

$$\Phi_0(s) = \begin{bmatrix} B^*(-sI - A^*)^{-1} & I \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix}$$

satisfies

$$\Phi_0(j\omega) > 0, \quad \textit{for all } 0 \leq \omega \leq \infty$$

if and only if there exists a unique real $X = X^* \leq 0$ such that

$$A^*X + XA - XBB^*X + P = 0$$

and $\sigma(A - BB^*X) \subset \mathbb{C}_-$.

(ii) $\Phi_0(j\omega) \geq 0$, for all $0 \leq \omega \leq \infty$ if and only if there exists a unique real $X = X^* \leq 0$ such that

$$A^*X + XA - XBB^*X + P = 0 \tag{13.33}$$

and $\sigma(A - BB^*X) \subset \overline{\mathbb{C}}_-$.

Proof. (i): $(\Rightarrow)$ From Corollary 13.16, $\Phi_0(j\omega) > 0$ implies that a Hamiltonian matrix

$$\begin{bmatrix} A & -BB^* \\ -P & -A^* \end{bmatrix}$$

has no eigenvalues on the imaginary axis. This in turn implies from Theorem 13.6 that (13.33) has a stabilizing solution. Furthermore, since

$$\begin{bmatrix} A & -BB^* \\ -P & -A^* \end{bmatrix} = \begin{bmatrix} -I & \\ & I \end{bmatrix} \begin{bmatrix} A & BB^* \\ P & -A^* \end{bmatrix} \begin{bmatrix} -I & \\ & I \end{bmatrix},$$

we have

$$Ric \begin{bmatrix} A & -BB^* \\ -P & -A^* \end{bmatrix} = -Ric \begin{bmatrix} A & BB^* \\ P & -A^* \end{bmatrix} \leq 0.$$

$(\Leftarrow)$ The sufficiency proof is omitted here and is a special case of $(b) \Rightarrow (a)$ of Theorem 13.19 below.

(ii): $(\Rightarrow)$ Let $P = -C^*C$ and $G(s) := C(sI - A)^{-1}B$. Then $\Phi_0(s) = I - G^\sim(s)G(s)$. Let $0 < \gamma < 1$ and define

$$\Psi_\gamma(s) := I - \gamma^2 G^\sim(s)G(s).$$

Then $\Psi_\gamma(j\omega) > 0, \quad \forall \omega$. Thus from part (i), there is an $X_\gamma = X_\gamma^* \leq 0$ such that $A - BB^*X_\gamma$ is stable and

$$A^*X_\gamma + X_\gamma A - X_\gamma BB^*X_\gamma - \gamma^2 C^*C = 0.$$

It is easy to see from Theorem 13.14 that X_γ is monotone-decreasing with γ, i.e., $X_{\gamma_1} \geq X_{\gamma_2}$ if $\gamma_1 \leq \gamma_2$. To show that $\lim_{\gamma\to 1} X_\gamma$ exists, we need to show that X_γ is bounded below for all $0 < \gamma < 1$.

In the following, it will be assumed that (A, B) is controllable. The controllability assumption will be removed later.

Let Y be the destabilizing solution to the following Riccati equation:

$$A^*Y + YA - YBB^*Y = 0$$

with $\sigma(A - BB^*Y) \subset \mathbb{C}_+$ (note that the existence of such a solution is guaranteed by the controllability assumption). Then it is easy to verify that

$$(A - BB^*Y)^*(X_\gamma - Y) + (X_\gamma - Y)(A - BB^*Y) - (X_\gamma - Y)BB^*(X_\gamma - Y) - \gamma^2 C^*C = 0.$$

This implies that

$$X_\gamma - Y = \int_0^\infty e^{-(A-BB^*Y)^*t}[(X_\gamma - Y)BB^*(X_\gamma - Y) + \gamma^2 C^*C]e^{-(A-BB^*Y)t}dt \geq 0.$$

Thus X_γ is bounded below with Y as the lower bound, and $\lim_{\gamma\to 1} X_\gamma$ exists. Let $X := \lim_{\gamma\to 1} X_\gamma$; then from continuity argument, X satisfies the Riccati equation

$$A^*X + XA - XBB^*X + P = 0$$

and $\sigma(A - BB^*X) \subset \overline{\mathbb{C}}_-$.

Now suppose (A, B) is not controllable, and then assume without loss of generality that

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix}, \quad B = \begin{bmatrix} B_1 \\ 0 \end{bmatrix}, \quad C = \begin{bmatrix} C_1 & C_2 \end{bmatrix}$$

so that (A_{11}, B_1) is controllable and A_{11} and A_{22} are stable. Then the Riccati equation for

$$X_\gamma = \begin{bmatrix} X_{11}^\gamma & X_{12}^\gamma \\ (X_{12}^\gamma)^* & X_{22}^\gamma \end{bmatrix}$$

can be written as three equations

$$A_{11}^*X_{11}^\gamma + X_{11}^\gamma A_{11} - X_{11}^\gamma B_1B_1^*X_{11}^\gamma - \gamma^2 C_1^*C_1 = 0$$

$$(A_{11} - B_1B_1X_{11}^\gamma)^*X_{12}^\gamma + X_{12}^\gamma A_{22} + X_{11}^\gamma A_{12} - \gamma^2 C_1^*C_2 = 0$$

$$A_{22}^*X_{22}^\gamma + X_{22}^\gamma A_{22} + (X_{12}^\gamma)^*A_{12} + A_{12}^*X_{12}^\gamma - (X_{12}^\gamma)^*B_1B_1^*X_{12}^\gamma - \gamma^2 C_2^*C_2 = 0$$

and

$$A - BB^*X = \begin{bmatrix} A_{11} - B_1B_1^*X_{11}^\gamma & A_{12} - B_1B_1^*X_{12}^\gamma \\ 0 & A_{22} \end{bmatrix}$$

is stable.

Let Y_{11} be the anti-stabilizing solution to

$$A_{11}^*Y_{11} + Y_{11}A_{11} - Y_{11}B_1B_1^*Y_{11} = 0$$

with $\sigma(A_{11} - B_1B_1^*Y_{11}) \subset \mathbb{C}_+$. Then it is clear that $X_{11}^\gamma - Y_{11} \geq 0$. Moreover, $X_{11} = \lim_{\gamma \to 1} X_{11}^\gamma$ exists and satisfies the following Riccati equation:

$$A_{11}^*X_{11} + X_{11}A_{11} - X_{11}B_1B_1^*X_{11} - C_1^*C_1 = 0$$

with $\sigma(A_{11} - B_1B_1^*X_{11}) \subset \overline{\mathbb{C}}_-$. Consequently, the following Sylvester equation

$$(A_{11} - B_1B_1X_{11})^*X_{12} + X_{12}A_{22} + X_{11}A_{12} - C_1^*C_2 = 0$$

has a unique solution X_{12} since $\lambda_i(A_{11} - B_1B_1X_{11}) + \lambda_j(A_{22}) \neq 0,\ \forall i, j$. Furthermore, the following Lyapunov equation has a unique nonnegative definite solution X_{22}:

$$A_{22}^*X_{22} + X_{22}A_{22} + X_{12}^*A_{12} + A_{12}^*X_{12} - X_{12}^*B_1B_1^*X_{12} - C_2^*C_2 = 0.$$

We have proven that there exists a unique X such that

$$A^*X + XA - XBB^*X - C^*C = 0$$

and $\sigma(A - BB^*X) \subset \overline{\mathbb{C}}_-$.

($\Leftarrow$) same as in part (i). □

Lemma 13.18 *Let $R > 0$,*

$$\Phi(s) = \begin{bmatrix} B^*(-sI - A^*)^{-1} & I \end{bmatrix} \begin{bmatrix} P & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix}$$

and

$$\hat{\Phi}(s) = \begin{bmatrix} \hat{B}^*(-sI - \hat{A}^*)^{-1} & I \end{bmatrix} \begin{bmatrix} \hat{P} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} (sI - \hat{A})^{-1}\hat{B} \\ I \end{bmatrix}$$

where

$$\hat{A} := A - BR^{-1}S^*,\ \hat{B} := BR^{-1/2},\ \hat{P} := P - SR^{-1}S^*.$$

Then $\Phi(j\omega) \geq 0$ iff $\hat{\Phi}(j\omega) \geq 0$.

Proof. Note that

$$\begin{bmatrix} P & S \\ S^* & R \end{bmatrix} = \begin{bmatrix} I & SR^{-1/2} \\ 0 & R^{1/2} \end{bmatrix} \begin{bmatrix} \hat{P} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} I & 0 \\ R^{-1/2}S^* & R^{1/2} \end{bmatrix}.$$

Hence the function $\Phi(s)$ can be written as

$$\Phi(s) = \begin{bmatrix} B^*(-sI - A^*)^{-1} & \varphi^{\sim}(s) \end{bmatrix} \begin{bmatrix} \hat{P} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}B \\ \varphi(s) \end{bmatrix}$$

with $\varphi(s) = R^{1/2} + R^{-1/2}S^*(sI - A)^{-1}B$. It is easy to verify that

$$\hat{\Phi}(s) = [\varphi^{-1}(s)]^{\sim}\Phi(s)\varphi^{-1}(s).$$

Hence the conclusion follows. $\square$

Now we are ready to state and prove one of the main results of this section. The following theorem and corollary characterize the relations among spectral factorizations, Riccati equations, and decomposition of Hamiltonians.

Theorem 13.19 *Let A, B, P, S, R be matrices of compatible dimensions such that $P = P^*$, $R = R^* > 0$, with (A, B) stabilizable. Suppose either one of the following assumptions is satisfied:*

(A1) A has no eigenvalues on $j\omega$-axis;

(A2) P is sign definite, i.e., $P \geq 0$ or $P \leq 0$ and (P, A) has no unobservable modes on the $j\omega$-axis.

Then

(I) The following statements are equivalent:

(a) The parahermitian rational matrix

$$\Phi(s) = \begin{bmatrix} B^*(-sI - A^*)^{-1} & I \end{bmatrix} \begin{bmatrix} P & S \\ S^* & R \end{bmatrix} \begin{bmatrix} (sI - A)^{-1}B \\ I \end{bmatrix}$$

satisfies

$$\Phi(j\omega) > 0 \text{ for all } 0 \leq \omega \leq \infty.$$

(b) There exists a unique real $X = X^*$ such that

$$(A - BR^{-1}S^*)^*X + X(A - BR^{-1}S^*) - XBR^{-1}B^*X + P - SR^{-1}S^* = 0$$

and that $A - BR^{-1}S^* - BR^{-1}B^*X$ is stable.

(c) The Hamiltonian matrix

$$H = \begin{bmatrix} A - BR^{-1}S^* & -BR^{-1}B^* \\ -(P - SR^{-1}S^*) & -(A - BR^{-1}S^*)^* \end{bmatrix}$$

has no $j\omega$-axis eigenvalues.

(II) The following statements are also equivalent:

(d) $\Phi(j\omega) \geq 0$ for all $0 \leq \omega \leq \infty$.

(e) There exists a unique real $X = X^*$ such that

$$(A - BR^{-1}S^*)^* X + X(A - BR^{-1}S^*) - XBR^{-1}B^*X + P - SR^{-1}S^* = 0$$

and that $\sigma(A - BR^{-1}S^* - BR^{-1}B^*X) \subset \overline{\mathbb{C}}_-$.

The following corollary is usually referred to as the spectral factorization theory.

Corollary 13.20 *If either one of the conditions, (a)–(e), in Theorem 13.19 is satisfied, then there exists an $M \in \mathcal{R}_p$ such that*

$$\Phi = M^\sim RM.$$

A particular realization of one such M is

$$M = \left[\begin{array}{c|c} A & B \\ \hline -F & I \end{array}\right]$$

where $F = -R^{-1}(S^* + B^*X)$. Furthermore, if X is the stabilizing solution, then $M^{-1} \in \mathcal{RH}_\infty$.

Remark 13.6 If the stabilizability of (A, B) is changed into the stabilizability of $(-A, B)$, then the theorem still holds except that the solutions X in (b) and (e) are changed into the destabilizing solution ($\sigma(A - BR^{-1}S^* - BR^{-1}B^*X) \subset \mathbb{C}_+$) and the weakly destabilizing solution ($\sigma(A - BR^{-1}S^* - BR^{-1}B^*X) \subset \overline{\mathbb{C}}_+$), respectively. ♡

Proof. $(a) \Rightarrow (c)$ follows from Corollary 13.16.

$(c) \Rightarrow (b)$ follows from Theorem 13.6 and Theorem 13.5.

$(b) \Rightarrow (a)$ Suppose $\exists X = X^*$ such that $A - BR^{-1}S^* - BR^{-1}B^*X = A - BR^{-1}(S^* + B^*X)$ is stable. Let $F = -R^{-1}(S^* + B^*X)$ and

$$M = \left[\begin{array}{c|c} A & B \\ \hline -F & I \end{array}\right].$$

It is easily verified by use of the Riccati equation for X and by routine algebra that $\Phi = M^{\sim} R M$. Since

$$M^{-1} = \left[\begin{array}{c|c} A + BF & B \\ \hline F & I \end{array}\right],$$

$M^{-1} \in \mathcal{RH}_\infty$. Thus $M(s)$ has no zeros on the imaginary axis and $\Phi(j\omega) > 0$.

$(e) \Rightarrow (d)$ follows the same procedure as the proof of $(b) \Rightarrow (a)$.

$(d) \Rightarrow (e)$: Assume $S = 0$ and $R = I$; otherwise use Lemma 13.18 first to get a new function with such properties. Let $P = C_1^* C_1 - C_2^* C_2$ with C_1 and C_2 square nonsingular. Note that this decomposition always exists since $P = \alpha I - (\alpha I - P)$, with $\alpha > 0$ sufficiently large, defines one such possibility. Let X_1 be the positive definite solution to

$$A^* X_1 + X_1 A - X_1 B B^* X_1 + C_1^* C_1 = 0.$$

By Theorem 13.7, X_1 indeed exists and is stabilizing, i.e., $A_1 := A - BB^* X_1$ is stable. Let $\Delta = X - X_1$. Then the equation in Δ becomes

$$A_1^* \Delta + \Delta A_1 - \Delta B B^* \Delta - C_2^* C_2 = 0.$$

To show that this equation has a solution, recall Lemma 13.17 and note that A_1 is stable; then it is sufficient to show that

$$I - B^*(-j\omega I - A_1^*)^{-1} C_2^* C_2 (j\omega I - A_1)^{-1} B$$

is positive semi-definite.

Notice first that

$$C_2(sI - A + BB^* X_1)^{-1} B = C_2 (sI - A)^{-1} B \left[I + B^* X_1 (sI - A)^{-1} B\right]^{-1}.$$

From the definition of X_1 and Corollary 13.20, we also have that

$$\begin{aligned} & I + B^*(-sI - A^*)^{-1} C_1^* C_1 (sI - A)^{-1} B \\ & = \left[I + B^* X_1 (sI - A)^{-1} B\right]^{\sim} \left[I + B^* X_1 (sI - A)^{-1} B\right]. \end{aligned}$$

Now, by assumption

$$\begin{aligned} \Phi(j\omega) = {} & I + B^*(-j\omega I - A^*)^{-1} C_1^* C_1 (j\omega I - A)^{-1} B \\ & - B^*(-j\omega I - A^*)^{-1} C_2^* C_2 (j\omega I - A)^{-1} B \geq 0. \end{aligned}$$

It follows that

$$\begin{aligned} I - & \left\{ \left[I + B^*(-j\omega I - A^*)^{-1} X_1 B\right]^{-1} B^*(-j\omega I - A^*)^{-1} C_2^* \right\} \\ & \left\{ C_2 (j\omega I - A)^{-1} B \left[I + B^* X_1 (j\omega I - A)^{-1} B\right]^{-1} \right\} \geq 0. \end{aligned}$$

Consequently,

$$I - B^*(-j\omega I - A_1^*)^{-1} C_2^* C_2 (j\omega I - A_1)^{-1} B \geq 0.$$

We may now apply Lemma 13.17 to the Δ equation. Consequently, there exists a unique solution Δ such that $\sigma(A_1 - BB^*\Delta) = \sigma(A - BB^*X) \subset \overline{\mathbb{C}}_-$. This shows the existence and uniqueness of X.

□

We shall now illustrate the proceeding results through a simple example. Note in particular that the function can have poles on the imaginary axis.

Example 13.6 Let $A = 0, B = 1, R = 1, S = 0$ and $P = 1$. Then $\Phi(s) = 1 - \frac{1}{s^2}$ and $\Phi(j\omega) > 0$. It is easy to check that the Hamiltonian matrix

$$H = \begin{bmatrix} 0 & -1 \\ -1 & 0 \end{bmatrix}$$

does not have eigenvalues on the imaginary axis and $X = 1$ is the stabilizing solution to the corresponding Riccati equation and the spectral factor is given by

$$M(s) = \left[\begin{array}{c|c} 0 & 1 \\ \hline 1 & 1 \end{array}\right] = \frac{s+1}{s}.$$

◇

Some frequently used special spectral factorizations are now considered.

Corollary 13.21 *Assume that* $G(s) := \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RL}_\infty$ *is a stabilizable and detectable realization and* $\gamma > \|G(s)\|_\infty$*. Then, there exists a transfer matrix* $M \in \mathcal{RL}_\infty$ *such that* $M^\sim M = \gamma^2 I - G^\sim G$ *and* $M^{-1} \in \mathcal{RH}_\infty$*. A particular realization of* M *is*

$$M(s) = \left[\begin{array}{c|c} A & B \\ \hline -R^{1/2}F & R^{1/2} \end{array}\right]$$

where

$$\begin{aligned} R &= \gamma^2 I - D^*D \\ F &= R^{-1}(B^*X + D^*C) \\ X &= Ric \begin{bmatrix} A + BR^{-1}D^*C & BR^{-1}B^* \\ -C^*(I + DR^{-1}D^*)C & -(A + BR^{-1}D^*C)^* \end{bmatrix} \end{aligned}$$

and $X \geq 0$ *if* A *is stable.*

Proof. This is a special case of Theorem 13.19. In fact, the theorem follows by letting $P = -C^*C$, $S = -C^*D$, $R = \gamma^2 I - D^*D$ in Theorem 13.19 and by using the fact that

$$\mathrm{Ric}\begin{bmatrix} A + BR^{-1}D^*C & BR^{-1}B^* \\ -C^*(I + DR^{-1}D^*)C & -(A + BR^{-1}D^*C)^* \end{bmatrix} =$$

$$-\mathrm{Ric}\begin{bmatrix} A + BR^{-1}D^*C & -BR^{-1}B^* \\ C^*(I + DR^{-1}D^*)C & -(A + BR^{-1}D^*C)^* \end{bmatrix}.$$

□

The spectral factorization for the dual case is also often used and follows by taking the transpose of the corresponding transfer matrices.

Corollary 13.22 *Assume* $G(s) := \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RL}_\infty$ *and* $\gamma > \|G(s)\|_\infty$. *Then, there exists a transfer matrix* $M \in \mathcal{RL}_\infty$ *such that* $MM^\sim = \gamma^2 I - GG^\sim$ *and* $M^{-1} \in \mathcal{RH}_\infty$. *A particular realization of* M *is*

$$M(s) = \left[\begin{array}{c|c} A & -LR^{1/2} \\ \hline C & R^{1/2} \end{array}\right]$$

where

$$\begin{aligned} R &= \gamma^2 I - DD^* \\ L &= (YC^* + BD^*)R^{-1} \\ Y &= Ric\begin{bmatrix} (A + BD^*R^{-1}C)^* & C^*R^{-1}C \\ -B(I + D^*R^{-1}D)B^* & -(A + BD^*R^{-1}C) \end{bmatrix} \end{aligned}$$

and $Y \geq 0$ *if* A *is stable.*

For convenience, we also include the following spectral factorization results which are again special cases of Theorem 13.19.

Corollary 13.23 *Let* $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *be a stabilizable and detectable realization.*

(a) Suppose $G^\sim(j\omega)G(j\omega) > 0$ *for all* ω *or* $\begin{bmatrix} A - j\omega & B \\ C & D \end{bmatrix}$ *has full column rank for all* ω. *Let*

$$X = Ric\begin{bmatrix} A - BR^{-1}D^*C & -BR^{-1}B^* \\ -C^*(I - DR^{-1}D^*)C & -(A - BR^{-1}D^*C)^* \end{bmatrix}$$

with $R := D^*D > 0$. *Then we have the following spectral factorization*

$$W^\sim W = G^\sim G$$

where $W^{-1} \in \mathcal{RH}_\infty$ *and*

$$W = \left[\begin{array}{c|c} A & B \\ \hline R^{-1/2}(D^*C + B^*X) & R^{1/2} \end{array}\right].$$

(b) Suppose $G(j\omega)G^\sim(j\omega) > 0$ *for all* ω *or* $\begin{bmatrix} A - j\omega & B \\ C & D \end{bmatrix}$ *has full row rank for all* ω. *Let*

$$Y = Ric \begin{bmatrix} (A - BD^*\tilde{R}^{-1}C)^* & -C^*\tilde{R}^{-1}C \\ -B(I - D^*\tilde{R}^{-1}D)B^* & -(A - BD^*\tilde{R}^{-1}C) \end{bmatrix}$$

with $\tilde{R} := DD^* > 0$. *Then we have the following spectral factorization*

$$\tilde{W}\tilde{W}^\sim = GG^\sim$$

where $\tilde{W}^{-1} \in \mathcal{RH}_\infty$ *and*

$$\tilde{W} = \left[\begin{array}{c|c} A & (BD^* + YC^*)\tilde{R}^{-1/2} \\ \hline C & \tilde{R}^{1/2} \end{array}\right].$$

Theorem 13.19 also gives some additional characterizations of a transfer matrix $\mathcal{H}_\infty$ norm.

Corollary 13.24 *Let* $\gamma > 0$, $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$ *and*

$$H := \begin{bmatrix} A + BR^{-1}D^*C & BR^{-1}B^* \\ -C^*(I + DR^{-1}D^*)C & -(A + BR^{-1}D^*C)^* \end{bmatrix}$$

where $R = \gamma^2 I - D^*D$. *Then the following conditions are equivalent:*

(i) $\|G\|_\infty < \gamma$.

(ii) $\bar{\sigma}(D) < \gamma$ *and* H *has no eigenvalues on the imaginary axis.*

(iii) $\bar{\sigma}(D) < \gamma$ *and* $H \in dom(Ric)$.

(iv) $\bar{\sigma}(D) < \gamma$ *and* $H \in dom(Ric)$ *and* $Ric(H) \geq 0$ *(*$Ric(H) > 0$ *if* (C, A) *is observable).*

Proof. This follows from the fact that $\|G\|_\infty < \gamma$ is equivalent to that the following function is positive definite for all ω:

$$\Phi(j\omega) := \gamma^2 I - G^T(-j\omega)G(j\omega)$$

$$= \begin{bmatrix} B^*(-j\omega I - A^*)^{-1} & I \end{bmatrix} \begin{bmatrix} -C^*C & -C^*D \\ -D^*C & \gamma^2 I - D^*D \end{bmatrix} \begin{bmatrix} (j\omega I - A)^{-1}B \\ I \end{bmatrix} > 0,$$

and the fact that

$$\begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} H \begin{bmatrix} -I & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} A + BR^{-1}D^*C & -BR^{-1}B^* \\ C^*(I + DR^{-1}D^*)C & -(A + BR^{-1}D^*C)^* \end{bmatrix}.$$

□

The equivalence between (i) and (iv) in the above corollary is usually referred as *Bounded Real Lemma.*

13.5 Positive Real Functions

A square $(m \times m)$ matrix function $G(s) \in \mathcal{RH}_\infty$ is said to be *positive real (PR)* if $G(j\omega) + G^*(j\omega) \geq 0$ for all finite ω, i.e., $\omega \in \mathbb{R}$, and $G(s)$ is said to be *strictly positive real (SPR)* if $G(j\omega) + G^*(j\omega) > 0$ for all $\omega \in \mathbb{R}$.

Theorem 13.25 *Let* $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *be a state space realization of* $G(s)$ *with* A *stable (not necessarily a minimal realization). Suppose there exist an* $X \geq 0$, Q, *and* W *such that*

$$XA + A^*X = -Q^*Q \tag{13.34}$$
$$B^*X + W^*Q = C \tag{13.35}$$
$$D + D^* = W^*W, \tag{13.36}$$

Then $G(s)$ *is positive real and*

$$G(s) + G^\sim(s) = M^\sim(s)M(s)$$

with $M(s) = \left[\begin{array}{c|c} A & B \\ \hline Q & W \end{array}\right]$. *Furthermore, if* $M(j\omega)$ *has full column rank for all* $\omega \in \mathbb{R}$, *then* $G(s)$ *is strictly positive real.*

Proof.

$$G(s)+G^{\sim}(s) = \left[\begin{array}{cc|c} A & 0 & B \\ 0 & -A^* & -C^* \\ \hline C & B^* & D+D^* \end{array}\right] = \left[\begin{array}{cc|c} A & 0 & B \\ 0 & -A^* & -(XB+Q^*W) \\ \hline B^*X+W^*Q & B^* & W^*W \end{array}\right].$$

Apply a similarity transformation $\begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$ to the last realization to get

$$\begin{aligned} G(s)+G^{\sim}(s) &= \left[\begin{array}{cc|c} A & 0 & B \\ XA+A^*X & -A^* & -Q^*W \\ \hline W^*Q & B^* & W^*W \end{array}\right] = \left[\begin{array}{cc|c} A & 0 & B \\ -Q^*Q & -A^* & -Q^*W \\ \hline W^*Q & B^* & W^*W \end{array}\right] \\ &= \left[\begin{array}{c|c} -A^* & -Q^* \\ \hline B^* & W^* \end{array}\right] \left[\begin{array}{c|c} A & B \\ \hline Q & W \end{array}\right]. \end{aligned}$$

This implies that

$$G(j\omega)+G^*(j\omega) = M^*(j\omega)M(j\omega) \geq 0$$

i.e., $G(s)$ is positive real. Finally note that if $M(j\omega)$ has full column rank for all $\omega \in \mathbb{R}$, then $M(j\omega)x \neq 0$ for all $x \in \mathbb{C}^m$ and $\omega \in \mathbb{R}$. Thus $G(s)$ is strictly positive real. □

Theorem 13.26 *Suppose (A,B,C,D) is a minimal realization of $G(s)$ with A stable and $G(s)$ is positive real. Then there exist an $X \geq 0$, Q, and W such that*

$$\begin{aligned} XA+A^*X &= -Q^*Q \\ B^*X+W^*Q &= C \\ D+D^* &= W^*W. \end{aligned}$$

and

$$G(s)+G^{\sim}(s) = M^{\sim}(s)M(s)$$

with $M(s) = \left[\begin{array}{c|c} A & B \\ \hline Q & W \end{array}\right]$. *Furthermore, if $G(s)$ is strictly positive real, then $M(j\omega)$ given above has full column rank for all $\omega \in \mathbb{R}$.*

Proof. Since $G(s)$ is assumed to be positive real, there exists a transfer matrix $M(s) = \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right]$ with A_1 stable such that

$$G(s)+G^{\sim}(s) = M^{\sim}(s)M(s)$$

where A and A_1 have the same dimensions. Now let $X_1 \geq 0$ be the solution of the following Lyapunov equation:

$$X_1 A_1 + A_1^* X_1 = -C_1^* C_1.$$

Then

$$\begin{aligned}
M^{\sim}(s)M(s) &= \left[\begin{array}{c|c} -A_1^* & -C_1^* \\ \hline B_1^* & D_1^* \end{array}\right] \left[\begin{array}{c|c} A_1 & B_1 \\ \hline C_1 & D_1 \end{array}\right] = \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ -C_1^* C_1 & -A_1^* & -C_1^* D_1 \\ \hline D_1^* C_1 & B_1^* & D_1^* D_1 \end{array}\right] \\
&= \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ X_1 A_1 + A_1^* X_1 & -A_1^* & -C_1^* D_1 \\ \hline D_1^* C_1 & B_1^* & D_1^* D_1 \end{array}\right] \\
&= \left[\begin{array}{cc|c} A_1 & 0 & B_1 \\ 0 & -A_1^* & -(X_1 B_1 + C_1^* D_1) \\ \hline B_1^* X + D_1^* C_1 & B_1^* & D_1^* D_1 \end{array}\right] \\
&= D_1^* D_1 + \left[\begin{array}{c|c} A_1 & B_1 \\ \hline B_1^* X + D_1^* C_1 & 0 \end{array}\right] + \left[\begin{array}{c|c} -A_1^* & -(B_1^* X + D_1^* C_1)^* \\ \hline B_1^* & 0 \end{array}\right].
\end{aligned}$$

But the realization for $G(s) + G^{\sim}(s)$ is given by

$$G(s) + G^{\sim}(s) = D + D^* + \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] + \left[\begin{array}{c|c} -A^* & -C^* \\ \hline B^* & 0 \end{array}\right].$$

Since the realization for $G(s)$ is minimal, there exists a nonsingular matrix T such that

$$A = TA_1T^{-1}, \quad B = TB_1, \quad C = (B_1^* X + D_1^* C_1)T^{-1}, \quad D + D^* = D_1^* D_1.$$

Now the conclusion follows by defining

$$X = (T^{-1})^* X_1 T^{-1}, \quad W = D_1, \quad Q = C_1 T^{-1}.$$

□

Corollary 13.27 *Let* $\left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *be a state space realization of* $G(s) \in \mathcal{RH}_\infty$ *with* A *stable and* $R := D + D^* > 0$. *Then* $G(s)$ *is strictly positive real if and only if there exists a stabilizing solution to the following Riccati equation:*

$$X(A - BR^{-1}C) + (A - BR^{-1}C)^* X + XBR^{-1}B^* X + C^* R^{-1} C = 0.$$

Moreover, $M(s) = \left[\begin{array}{c|c} A & B \\ \hline R^{-\frac{1}{2}}(C - B^* X) & R^{\frac{1}{2}} \end{array}\right]$ *is minimal phase and*

$$G(s) + G^{\sim}(s) = M^{\sim}(s)M(s).$$

Proof. This follows from Theorem 13.19 and from the fact that

$$G(j\omega) + G^*(j\omega) > 0$$

for all ω including ∞. □

The above corollary also leads to the following special spectral factorization.

Corollary 13.28 *Let* $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$ *with* D *full row rank and* $G(j\omega)G^*(j\omega) > 0$ *for all* ω. *Let* P *be the controllability grammian of* (A, B)*:*

$$PA^* + AP + BB^* = 0.$$

Define

$$B_W = PC^* + BD^*.$$

Then there exists an $X \geq 0$ *such that*

$$XA + A^*X + (C - B_W^*X)^*(DD^*)^{-1}(C - B_W^*X) = 0.$$

Furthermore, there is an $M(s) \in \mathcal{RH}_\infty$ *such that* $M^{-1}(s) \in \mathcal{RH}_\infty$ *and*

$$G(s)G^\sim(s) = M^\sim(s)M(s)$$

where $M(s) = \left[\begin{array}{c|c} A & B_W \\ \hline C_W & D_W \end{array}\right]$ *with a square matrix* D_W *such that*

$$D_W^*D_W = DD^*$$

and

$$C_W = D_W(DD^*)^{-1}(C - B_W^*X).$$

Proof. This corollary follows from Corollary 13.27 and the fact that $G(j\omega)G^*(j\omega) > 0$ and

$$G(s)G^\sim(s) = \left[\begin{array}{c|c} A & B_W \\ \hline C & 0 \end{array}\right] + \left[\begin{array}{c|c} -A^* & -C^* \\ \hline B_W^* & 0 \end{array}\right] + DD^*.$$

□

A dual spectral factorization can also be obtained easily.

13.6 Inner Functions

A transfer function N is called *inner* if $N \in \mathcal{RH}_\infty$ and $N^\sim N = I$ and *co-inner* if $N \in \mathcal{RH}_\infty$ and $NN^\sim = I$. Note that N need not be square. Inner and co-inner are dual notions, i.e., N is an inner iff N^T is a co-inner. A matrix function $N \in \mathcal{RL}_\infty$ is called *all-pass* if N is square and $N^\sim N = I$; clearly a square inner function is all-pass. We will focus on the characterizations of inner functions here and the properties of co-inner functions follow by duality.

Note that N inner implies that N has at least as many rows as columns. For N inner and any $q \in \mathbb{C}^m$, $v \in \mathcal{L}_2$, $\|N(j\omega)q\| = \|q\|$, $\forall\omega$ and $\|Nv\|_2 = \|v\|_2$ since $N(j\omega)^*N(j\omega) = I$ for all ω. Because of these norm preserving properties, inner matrices will play an important role in the control synthesis theory in this book. In this section, we present a state-space characterization of inner transfer functions.

Lemma 13.29 *Suppose* $N = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$ *and* $X = X^* \geq 0$ *satisfies*

$$A^*X + XA + C^*C = 0. \tag{13.37}$$

Then

(a) $D^*C + B^*X = 0$ *implies* $N^\sim N = D^*D$.

(b) (A, B) *controllable, and* $N^\sim N = D^*D$ *implies* $D^*C + B^*X = 0$.

Proof. Conjugating the states of

$$N^\sim N = \left[\begin{array}{cc|c} A & 0 & B \\ -C^*C & -A^* & -C^*D \\ \hline D^*C & B^* & D^*D \end{array}\right]$$

by $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}$ on the left and $\begin{bmatrix} I & 0 \\ -X & I \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ X & I \end{bmatrix}$ on the right yields

$$\begin{aligned} N^\sim N &= \left[\begin{array}{cc|c} A & 0 & B \\ -(A^*X + XA + C^*C) & -A^* & -(XB + C^*D) \\ \hline B^*X + D^*C & B^* & D^*D \end{array}\right] \\ &= \left[\begin{array}{cc|c} A & 0 & B \\ 0 & -A^* & -(XB + C^*D) \\ \hline B^*X + D^*C & B^* & D^*D \end{array}\right]. \end{aligned}$$

Then (a) and (b) follow easily. □

This lemma immediately leads to one characterization of inner matrices in terms of their state space representations. Simply add the condition that $D^*D = I$ to Lemma 13.29 to get $N^\sim N = I$.

Corollary 13.30 *Suppose* $N = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *is stable and minimal, and* X *is the observability grammian. Then* N *is inner if and only if*

(a) $D^*C + B^*X = 0$

(b) $D^*D = I$.

A transfer matrix $N_\perp$ is called a *complementary inner factor (CIF)* of N if $[N\, N_\perp]$ is square and inner. The dual notion of the complementary co-inner factor is defined in the obvious way. Given an inner N, the following lemma gives a construction of its CIF. The proof of this lemma follows from straightforward calculation and from the fact that $CX^\dagger X = C$ since $\mathrm{Im}(I - X^+X) \subset \mathrm{Ker}(X) \subset \mathrm{Ker}(C)$.

Lemma 13.31 *Let* $N = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *be an inner and* X *be the observability grammian. Then a CIF* $N_\perp$ *is given by*

$$N_\perp = \left[\begin{array}{c|c} A & -X^\dagger C^* D_\perp \\ \hline C & D_\perp \end{array}\right]$$

where $D_\perp$ *is an orthogonal complement of* D *such that* $[D\, D_\perp]$ *is square and orthogonal.*

13.7 Inner-Outer Factorizations

In this section, some special form of coprime factorizations will be developed. In particular, explicit realizations are given for coprime factorizations $G = NM^{-1}$ with inner numerator N and inner denominator M, respectively. The former factorization in the case of $G \in \mathcal{RH}_\infty$ will give an inner-outer factorization[1]. The results will be proven for the right coprime factorizations, while the results for left coprime factorizations follow by duality.

Let $G \in \mathcal{R}_p$ be a $p \times m$ transfer matrix and denote $R^{1/2}R^{1/2} = R$. For a given full column rank matrix D, let $D_\perp$ denote for any orthogonal complement of D so that $\begin{bmatrix} DR^{-1/2} & D_\perp \end{bmatrix}$ (with $R = D^*D > 0$) is square and orthogonal. To obtain an *rcf* of G with N inner, we note that if NM^{-1} is an rcf, then $(NZ_r)(MZ_r)^{-1}$ is also an rcf for any nonsingular real matrix Z_r. We simply need to use the formulas in Theorem 5.9 to solve for F and Z_r.

[1] A $p \times m$ ($p \le m$) transfer matrix $G_o \in \mathcal{RH}_\infty$ is called an *outer* if $G_o(s)$ has full row rank in the open right half plane, i.e., $G_o(s)$ has full row normal rank and has no open right half plane zeros.

Theorem 13.32 *Assume* $p \geq m$. *Then there exists an* rcf $G = NM^{-1}$ *such that* N *is* inner *if and only if* $G^{\sim}G > 0$ *on the* $j\omega$*-axis, including at* ∞. *This factorization is unique up to a constant unitary multiple. Furthermore, assume that the realization of* $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *is stabilizable and that* $\begin{bmatrix} A - j\omega I & B \\ C & D \end{bmatrix}$ *has full column rank for all* $\omega \in \mathbb{R}$. *Then a particular realization of the desired coprime factorization is*

$$\begin{bmatrix} M \\ N \end{bmatrix} := \left[\begin{array}{c|c} A+BF & BR^{-1/2} \\ \hline F & R^{-1/2} \\ C+DF & DR^{-1/2} \end{array}\right] \in \mathcal{RH}_\infty$$

where

$$R = D^*D > 0$$

$$F = -R^{-1}(B^*X + D^*C)$$

and

$$X = Ric \begin{bmatrix} A - BR^{-1}D^*C & -BR^{-1}B^* \\ -C^*(I - DR^{-1}D^*)C & -(A - BR^{-1}D^*C)^* \end{bmatrix} \geq 0.$$

Moreover, a complementary inner factor can be obtained as

$$N_\perp = \left[\begin{array}{c|c} A+BF & -X^\dagger C^* D_\perp \\ \hline C+DF & D_\perp \end{array}\right]$$

if $p > m$.

Proof. ($\Rightarrow$) Suppose $G = NM^{-1}$ is an *rcf* and $N^{\sim}N = I$. Then $G^{\sim}G = (M^{-1})^{\sim}M^{-1} > 0$ on the $j\omega$-axis since $M \in \mathcal{RH}_\infty$.

($\Leftarrow$) This can be shown by using Corollary 13.20 first to get a factorization $G^{\sim}G = (M^{-1})^{\sim}(M^{-1})$ and then to compute $N = GM$. The following proof is more direct. It will be proven by showing that the definition of the inner and coprime factorization formula given in Theorem 5.9 lead directly to the above realization of the *rcf* of G with an inner numerator. That $G = NM^{-1}$ is an *rcf* follows immediately once it is established that F is a stabilizing state feedback. Now suppose

$$N = \left[\begin{array}{c|c} A+BF & BZ_r \\ \hline C+DF & DZ_r \end{array}\right]$$

and let F and Z_r be such that

$$(DZ_r)^*(DZ_r) = I \tag{13.38}$$

$$(BZ_r)^*X + (DZ_r)^*(C+DF) = 0 \tag{13.39}$$

$$(A+BF)^*X + X(A+BF) + (C+DF)^*(C+DF) = 0. \tag{13.40}$$

Clearly, we have that $Z_r = R^{-1/2}U$ where $R = D^*D > 0$ and where U is any orthogonal matrix. Take $U = I$ and solve (13.39) for F to get

$$F = -R^{-1}(B^*X + D^*C).$$

Then substitute F into (13.40) to get

$$\begin{aligned} 0 &= (A+BF)^*X + X(A+BF) + (C+DF)^*(C+DF) \\ &= (A - BR^{-1}D^*C)^*X + X(A - BR^{-1}D^*C) - XBR^{-1}B^*X + C^*D_\perp D_\perp^* C \end{aligned}$$

where $D_\perp D_\perp^* = I - DR^{-1}D^*$. To show that such choices indeed make sense, we need to show that $H \in dom(Ric)$, where

$$H = \begin{bmatrix} A - BR^{-1}D^*C & -BR^{-1}B^* \\ -C^*D_\perp D_\perp^* C & -(A - BR^{-1}D^*C)^* \end{bmatrix}$$

so $X = Ric(H)$. However, by Theorem 13.19, $H \in dom(Ric)$ is guaranteed by the fact that $\begin{bmatrix} A - j\omega & B \\ C & D \end{bmatrix}$ has full column rank (or $G^\sim(j\omega)G(j\omega) > 0$.

The uniqueness of the factorization follows from coprimeness and N inner. Suppose that $G = N_1M_1^{-1} = N_2M_2^{-1}$ are two right coprime factorizations and that both numerators are inner. By coprimeness, these two factorizations are unique up to a right multiple which is a unit[2] in $\mathcal{RH}_\infty$. That is, there exists a unit $\Theta \in \mathcal{RH}_\infty$ such that $\begin{bmatrix} M_1 \\ N_1 \end{bmatrix}\Theta = \begin{bmatrix} M_2 \\ N_2 \end{bmatrix}$. Clearly, Θ is inner since $\Theta^\sim\Theta = \Theta^\sim N_1^\sim N_1 \Theta = N_2^\sim N_2 = I$. The only inner units in $\mathcal{RH}_\infty$ are constant matrices, and thus the desired uniqueness property is established. Note that the non-uniqueness is contained entirely in the choice of a particular square root of R.

Finally, the formula for $N_\perp$ follows from Lemma 13.31. □

Note that the important inner-outer factorization formula can be obtained from this inner numerator coprime factorization if $G \in \mathcal{RH}_\infty$.

Corollary 13.33 *Suppose $G \in \mathcal{RH}_\infty$; then the denominator matrix M in Theorem 13.32 is an outer. Hence, the factorization $G = N(M^{-1})$ given in Theorem 13.32 is an inner-outer factorization.*

Remark 13.7 It is noted that the above inner-outer factorization procedure does not apply to the strictly proper transfer matrix even if the factorization exists. For example, $G(s) = \frac{s-1}{s+1}\frac{1}{s+2}$ has inner-outer factorizations but the above procedure cannot be used. The inner-outer factorization for the general transfer matrices can be done using the method adopted in Section 6.1 of Chapter 6. ♡

[2] A function Θ is called a unit in $\mathcal{RH}_\infty$ if $\Theta, \Theta^{-1} \in \mathcal{RH}_\infty$.

Suppose that the system G is not stable; then a coprime factorization with an inner denominator can also be obtained by solving a special Riccati equation. The proof of this result is similar to the inner numerator case and is omitted.

Theorem 13.34 *Assume that* $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{R}_p$ *and* (A, B) *is stabilizable. Then there exists a right coprime factorization* $G = NM^{-1}$ *such that* $M \in \mathcal{RH}_\infty$ *is inner if and only if* G *has no poles on* $j\omega$*-axis. A particular realization is*

$$\begin{bmatrix} M \\ N \end{bmatrix} := \left[\begin{array}{c|c} A + BF & B \\ \hline F & I \\ C + DF & D \end{array}\right] \in \mathcal{RH}_\infty$$

where

$$F = -B^* X$$

$$X = Ric \begin{bmatrix} A & -BB^* \\ 0 & -A^* \end{bmatrix} \geq 0.$$

Dual results can be obtained when $p \leq m$ by taking the transpose of the transfer function matrix. In these factorizations, output injection using the dual Riccati solution replaces state feedback to obtain the corresponding left factorizations.

Theorem 13.35 *Assume* $p \leq m$. *Then there exists an* lcf $G = \tilde{M}^{-1}\tilde{N}$ *such that* $\tilde{N}$ *is a* co-inner *if and only if* $GG^\sim > 0$ *on the* $j\omega$*-axis, including at* ∞. *This factorization is unique up to a constant unitary multiple. Furthermore, assume that the realization of* $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *is detectable and that* $\begin{bmatrix} A - j\omega I & B \\ C & D \end{bmatrix}$ *has full row rank for all* $\omega \in \mathbb{R}$. *Then a particular realization of the desired coprime factorization is*

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} := \left[\begin{array}{c|cc} A + LC & L & B + LD \\ \hline \tilde{R}^{-1/2}C & \tilde{R}^{-1/2} & \tilde{R}^{-1/2}D \end{array}\right] \in \mathcal{RH}_\infty$$

where

$$\tilde{R} = DD^* > 0$$

$$L = -(BD^* + YC^*)\tilde{R}^{-1}$$

and

$$Y = Ric \begin{bmatrix} (A - BD^*\tilde{R}^{-1}\tilde{C})^* & -C^*\tilde{R}^{-1}C \\ -B(I - D^*\tilde{R}^{-1}D)B & -(A - BD^*\tilde{R}^{-1}C) \end{bmatrix} \geq 0.$$

Moreover, a complementary co-inner factor can be obtained as

$$\tilde{N}_\perp = \left[\begin{array}{c|c} A + LC & B + LD \\ \hline -\tilde{D}_\perp B^* Y^\dagger & \tilde{D}_\perp \end{array}\right]$$

if $p < m$, *where* $\tilde{D}_\perp$ *is a full row rank matrix such that* $\tilde{D}_\perp^*\tilde{D}_\perp = I - D^*\tilde{R}^{-1}D$.

Theorem 13.36 *Assume that* $G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{R}_p$ *and* (C, A) *is detectable. Then there exists a left coprime factorization* $G = \tilde{M}^{-1}\tilde{N}$ *such that* $\tilde{M} \in \mathcal{RH}_\infty$ *is inner if and only if* G *has no poles on* $j\omega$*-axis. A particular realization is*

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} := \left[\begin{array}{c|cc} A + LC & L & B + LD \\ \hline C & I & D \end{array}\right] \in \mathcal{RH}_\infty$$

where

$$L = -YC^*$$
$$Y = Ric \begin{bmatrix} A^* & -C^*C \\ 0 & -A \end{bmatrix} \geq 0.$$

13.8 Normalized Coprime Factorizations

A right coprime factorization of $G = NM^{-1}$ with $N, M \in \mathcal{RH}_\infty$ is called a *normalized right coprime factorization* if

$$M^\sim M + N^\sim N = I$$

i.e., if $\begin{bmatrix} M \\ N \end{bmatrix}$ is an inner. Similarly, an *lcf* $G = \tilde{M}^{-1}\tilde{N}$ is called a *normalized left coprime factorization* if $\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}$ is a co-inner.

The normalized coprime factorization is easy to obtain from the definition. The following theorem can be proven using the same procedure as in the proof for the coprime factorization with inner numerator. In this case, the proof involves choosing F and Z_r such that $\begin{bmatrix} M \\ N \end{bmatrix}$ is an inner.

Theorem 13.37 *Let a realization of* G *be given by*

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

and define

$$R = I + D^*D > 0, \qquad \tilde{R} = I + DD^* > 0.$$

(a) Suppose (A, B) *is stabilizable and* (C, A) *has no unobservable modes on the imaginary axis. Then there is a normalized right coprime factorization* $G = NM^{-1}$

$$\begin{bmatrix} M \\ N \end{bmatrix} := \left[\begin{array}{c|c} A + BF & BR^{-1/2} \\ \hline F & R^{-1/2} \\ C + DF & DR^{-1/2} \end{array}\right] \in \mathcal{RH}_\infty$$

where

$$F = -R^{-1}(B^*X + D^*C)$$

and

$$X = Ric\begin{bmatrix} A - BR^{-1}D^*C & -BR^{-1}B^* \\ -C^*\tilde{R}^{-1}C & -(A - BR^{-1}D^*C)^* \end{bmatrix} \geq 0.$$

(b) *Suppose (C, A) is detectable and (A, B) has no uncontrollable modes on the imaginary axis. Then there is a normalized left coprime factorization $G = \tilde{M}^{-1}\tilde{N}$*

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} := \left[\begin{array}{c|cc} A + LC & L & B + LD \\ \hline \tilde{R}^{-1/2}C & \tilde{R}^{-1/2} & \tilde{R}^{-1/2}D \end{array}\right]$$

where

$$L = -(BD^* + YC^*)\tilde{R}^{-1}$$

and

$$Y = Ric\begin{bmatrix} (A - BD^*\tilde{R}^{-1}C)^* & -C^*\tilde{R}^{-1}C \\ -BR^{-1}B^* & -(A - BD^*\tilde{R}^{-1}C) \end{bmatrix} \geq 0.$$

(c) *The controllability grammian P and observability grammian Q of $\begin{bmatrix} M \\ N \end{bmatrix}$ are given by*

$$P = (I + YX)^{-1}Y, \quad Q = X$$

while the controllability grammian $\tilde{P}$ and observability grammian $\tilde{Q}$ of $\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}$ are given by

$$\tilde{P} = Y, \quad \tilde{Q} = (I + XY)^{-1}X.$$

Proof. We shall only prove the first part of (c). It is obvious that $Q = X$ since the Riccati equation for X can be written as

$$X(A + BF) + (A + BF)^*X + \begin{bmatrix} F \\ C + DF \end{bmatrix}^* \begin{bmatrix} F \\ C + DF \end{bmatrix} = 0$$

while the controllability grammian solves the Lyapunov equation

$$(A + BF)P + P(A + BF)^* + BR^{-1}B^* = 0$$

or equivalently

$$P = \mathrm{Ric}\begin{bmatrix} (A + BF)^* & 0 \\ -BR^{-1}B^* & -(A + BF) \end{bmatrix}.$$

Now let $T = \begin{bmatrix} I & X \\ 0 & I \end{bmatrix}$; then

$$\begin{bmatrix} (A+BF)^* & 0 \\ -BR^{-1}B^* & -(A+BF) \end{bmatrix} = T \begin{bmatrix} A - BD^*\tilde{R}^{-1}C & -C^*\tilde{R}^{-1}C \\ -BR^{-1}B^* & -(A-BD^*\tilde{R}^{-1}C)^* \end{bmatrix} T^{-1}.$$

This shows that the stable invariant subspaces for these two Hamiltonian matrices are related by

$$\mathcal{X}_- \begin{bmatrix} (A+BF)^* & 0 \\ -BR^{-1}B^* & -(A+BF) \end{bmatrix} = T\ \mathcal{X}_- \begin{bmatrix} A - BD^*\tilde{R}^{-1}C & -C^*\tilde{R}^{-1}C \\ -BR^{-1}B^* & -(A-BD^*\tilde{R}^{-1}C)^* \end{bmatrix}$$

or

$$\operatorname{Im} \begin{bmatrix} I \\ P \end{bmatrix} = T \operatorname{Im} \begin{bmatrix} I \\ Y \end{bmatrix} = \operatorname{Im} \begin{bmatrix} I + XY \\ Y \end{bmatrix} = \operatorname{Im} \begin{bmatrix} I \\ Y(I+XY)^{-1} \end{bmatrix}.$$

Hence we have $P = Y(I+XY)^{-1}$. □

13.9 Notes and References

The general solutions of a Riccati equation are given by Martensson [1971]. The iterative procedure for solving ARE was first introduced by Kleinman [1968] for a special case and was further developed by Wonham [1968]. It was used by Coppel [1974], Ran and Vreugdenhil [1988], and many others for the proof of the existence of maximal and minimal solutions. The comparative results were obtained in Ran and Vreugdenhil [1988]. The paper by Wimmer[1985] also contains comparative results for some special cases. The paper by Willems [1971] contains a comprehensive treatment of ARE and the related optimization problems. Some matrix factorization results are given in Doyle [1984]. Numerical methods for solving ARE can be found in Arnold and Laub [1984], Dooren [1981], and references therein. The state space spectral factorization for functions singular at ∞ or on imaginary axis is considered in Clements and Glover [1989] and Clements [1993].

14

$\mathcal{H}_2$ Optimal Control

In this chapter we treat the optimal control of linear time-invariant systems with a quadratic performance criterion. The material in this chapter is standard, but the treatment is somewhat novel and lays the foundation for the subsequent chapters on $\mathcal{H}_\infty$-optimal control.

14.1 Introduction to Regulator Problem

Consider the following dynamical system:

$$\dot{x} = Ax + B_2u, \quad x(t_0) = x_0 \tag{14.1}$$

where x_0 is given but arbitrary. Our objective is to find a control function $u(t)$ defined on $[t_0, T]$ which can be a function of the state $x(t)$ such that the state $x(t)$ is driven to a (small) neighborhood of origin at time T. This is the so-called *Regulator Problem*. One might suggest that this regulator problem can be trivially solved for any $T > t_0$ if the system is controllable. This is indeed the case if the controller can provide arbitrarily large amounts of energy since, by the definition of controllability, one can immediately construct a control function that will drive the state to zero in an arbitrarily short time. However, this is not practical since any physical system has the energy limitation, i.e., the actuator will eventually saturate. Furthermore, large control action can easily drive the system out of the region where the given linear model is valid. Hence certain limitations have to be imposed on the control in practical engineering implementation.

The constraints on control u may be measured in many different ways; for example,

$$\int_{t_0}^{T} \|u\|\,dt, \quad \int_{t_0}^{T} \|u\|^2\,dt, \quad \sup_{t\in[t_0,T]} \|u\|$$

i.e., in terms of $\mathcal{L}_1$-norm, $\mathcal{L}_2$-norm, and $\mathcal{L}_\infty$-norm, or more generally, weighted $\mathcal{L}_1$-norm, $\mathcal{L}_2$-norm, and $\mathcal{L}_\infty$-norm

$$\int_{t_0}^{T} \|W_u u\|\,dt, \quad \int_{t_0}^{T} \|W_u u\|^2\,dt, \quad \sup_{t\in[t_0,T]} \|W_u u\|$$

for some constant weighting matrix W_u.

Similarly, one might also want to impose some constraints on the transient response $x(t)$ in a similar fashion

$$\int_{t_0}^{T} \|W_x x\|\,dt, \quad \int_{t_0}^{T} \|W_x x\|^2\,dt, \quad \sup_{t\in[t_0,T]} \|W_x x\|$$

for some weighting matrix W_x. Hence the regulator problem can be posed as an optimal control problem with certain combined performance index on u and x, as given above. In this chapter, we shall be concerned exclusively with the $\mathcal{L}_2$ performance problem or quadratic performance problem. Moreover, we will focus on the infinite time regulator problem, i.e., $T \to \infty$, and, without loss of generality, we shall assume $t_0 = 0$. In this case, our problem is as follows: find a control $u(t)$ defined on $[0, \infty)$ such that the state $x(t)$ is driven to the origin at $t \to \infty$ and the following performance index is minimized:

$$\min_u \int_0^\infty \begin{bmatrix} x(t) \\ u(t) \end{bmatrix}^* \begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \begin{bmatrix} x(t) \\ u(t) \end{bmatrix} dt \tag{14.2}$$

for some $Q = Q^*$, S, and $R = R^* > 0$. This problem is traditionally called a *Linear Quadratic Regulator* problem or simply an LQR problem. Here we have assumed $R > 0$ to emphasis that the control energy has to be finite, i.e., $u(t) \in \mathcal{L}_2[0, \infty)$. So this is the space over which the integral is minimized. Moreover, it is also generally assumed that

$$\begin{bmatrix} Q & S \\ S^* & R \end{bmatrix} \geq 0. \tag{14.3}$$

Since R is positive definite, it has a square root, $R^{1/2}$, which is also positive-definite. By the substitution

$$u \leftarrow R^{1/2} u,$$

we may as well assume at the start that $R = I$. In fact, we can even assume $S = 0$ by using a pre-state feedback $u = -S^* x + v$ provided some care is exercised; however, this will not be assumed in the sequel. Since the matrix in (14.3) is positive semi-definite with $R = I$, it can be factored as

$$\begin{bmatrix} Q & S \\ S^* & I \end{bmatrix} = \begin{bmatrix} C_1^* \\ D_{12}^* \end{bmatrix} \begin{bmatrix} C_1 & D_{12} \end{bmatrix}.$$

And (14.2) can be rewritten as

$$\min_{u \in \mathcal{L}_2[0,\infty)} \|C_1 x + D_{12} u\|_2^2 .$$

In fact, the LQR problem is posed traditionally as the minimization problem

$$\min_{u \in \mathcal{L}_2[0,\infty)} \|C_1 x + D_{12} u\|_2^2 \tag{14.4}$$

$$\dot{x} = Ax + Bu, \quad x(0) = x_0 \tag{14.5}$$

without explicitly mentioning the condition that the control should drive the state to the origin. Instead some assumptions are imposed on Q, S, and R (or equivalently C_1 and D_{12}) to ensure that the optimal control law u has this property. To see what assumption one needs to make in order to ensure that the minimization problem formulated in (14.4) and (14.5) has a sensible solution, let us consider a simple example with $A = 1, B = 1, Q = 0, S = 0$, and $R = 1$:

$$\min_{u \in \mathcal{L}_2[0,\infty)} \int_0^\infty u^2 dt, \quad \dot{x} = x + u, \quad x(0) = x_0.$$

It is clear that $u = 0$ is the optimal solution. However, the system with $u = 0$ is unstable and $x(t)$ diverges exponentially to infinity, $x(t) = e^t x_0$. The problem with this example is that the performance index does not "see" the unstable state x. This is true in general, and the proof of this fact is left as an exercise to the reader. Hence in order to ensure that the minimization problem in (14.4) and (14.5) is sensible, we must assume that all unstable states can be "seen" from the performance index, i.e., (C_1, A) must be detectable. This will be called a *standard LQR problem*.

On the other hand, if the closed-loop stability is imposed on the above minimization, then it can be shown that $\min_{u \in \mathcal{L}_2[0,\infty)} \int_0^\infty u^2 dt = 2x_0^2$ and $u(t) = -2x(t)$ is the optimal control. This can also be generalized to a more general case where (C_1, A) is not necessarily detectable. This problem will be referred to as an *Extended LQR problem*.

14.2 Standard LQR Problem

In this section, we shall consider the LQR problem as traditionally formulated.

Standard LQR Problem

Let a dynamical system be described by

$$\begin{aligned} \dot{x} &= Ax + B_2 u, \quad x(0) = x_0 \text{ given but arbitrary} & (14.6) \\ z &= C_1 x + D_{12} u & (14.7) \end{aligned}$$

and suppose that the system parameter matrices satisfy the following assumptions:

(A1) (A, B_2) is stabilizable;

(A2) D_{12} has full column rank with $\begin{bmatrix} D_{12} & D_\perp \end{bmatrix}$ unitary;

(A3) (C_1, A) is detectable;

(A4) $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all ω.

Find an optimal control law $u \in \mathcal{L}_2[0,\infty)$ such that the performance criterion $\|z\|_2^2$ is minimized.

Remark 14.1 Assumption (A1) is clearly necessary for the existence of a stabilizing control function u. The assumption (A2) is made for simplicity of notation and is actually a restatement that $R = D_{12}^* D_{12} = I$. Note also that $D_\perp$ drops out when D_{12} is square. It is interesting to point out that (A3) is not needed in the Extended LQR problem. The assumption (A3) enforces that the unconditional optimization problem will result in a stabilizing control law. In fact, the assumption (A3) together with (A1) guarantees that the input/output stability implies the internal stability, i.e., $u \in \mathcal{L}_2$ and $z \in \mathcal{L}_2$ imply $x \in \mathcal{L}_2$, which will be shown in Lemma 14.1. Finally note that (A4) is equivalent to the condition that $(D_\perp^* C_1, A - B_2 D_{12}^* C_1)$ has no unobservable modes on the imaginary axis and is weaker than the popular assumption of detectability of $(D_\perp^* C_1, A - B_2 D_{12}^* C_1)$. (A4), together with the stabilizability of (A, B_2), guarantees by Corollary 13.10 that the following Hamiltonian matrix belongs to $dom(Ric)$ and that $X = Ric(H) \geq 0$:

$$\begin{aligned} H &= \begin{bmatrix} A & 0 \\ -C_1^* C_1 & -A^* \end{bmatrix} - \begin{bmatrix} B_2 \\ -C_1^* D_{12} \end{bmatrix} \begin{bmatrix} D_{12}^* C_1 & B_2^* \end{bmatrix} \\ &= \begin{bmatrix} A - B_2 D_{12}^* C_1 & -B_2 B_2^* \\ -C_1^* D_\perp D_\perp^* C_1 & -(A - B_2 D_{12}^* C_1)^* \end{bmatrix}. \qquad (14.8) \end{aligned}$$

Note also that if $D_{12}^* C_1 = 0$, then (A4) is implied by the detectability of (C_1, A), while the detectability of (C_1, A) is implied by the detectability of $(D_\perp^* C_1, A - B_2 D_{12}^* C_1)$. The above implication is not true if $D_{12}^* C_1 \neq 0$, for example,

$$A = \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix}, \quad B_2 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad C_1 = \begin{bmatrix} 1 & 0 \end{bmatrix}, \quad D_{12} = 1.$$

Then (C_1, A) is detectable and $A - B_2 D_{12}^* C_1 = \begin{bmatrix} 1 & 0 \\ 0 & -2 \end{bmatrix}$ has no eigenvalue on the imaginary axis but is not stable. ♡

Note also that the Riccati equation corresponding to (14.8) is

$$(A - B_2 D_{12}^* C_1)^* X + X(A - B_2 D_{12}^* C_1) - X B_2 B_2^* X + C_1^* D_\perp D_\perp^* C_1 = 0. \tag{14.9}$$

Now let X be the corresponding stabilizing solution and define

$$F := -(B_2^* X + D_{12}^* C_1). \tag{14.10}$$

Then $A + B_2 F$ is stable. Denote

$$A_F := A + B_2 F, \quad C_F := C_1 + D_{12} F$$

and re-arrange equation (14.9) to get

$$A_F^* X + X A_F + C_F^* C_F = 0. \tag{14.11}$$

Thus X is the observability Gramian of (C_F, A_F).

Consider applying the control law $u = Fx$ to the system (14.6) and (14.7). The controlled system is

$$\begin{aligned} \dot{x} &= A_F x, \quad x(0) = x_0 \\ z &= C_F x \end{aligned}$$

or equivalently

$$\begin{aligned} \dot{x} &= A_F x + x_0 \delta(t), \quad x(0_-) = 0 \\ z &= C_F x. \end{aligned}$$

The associated transfer matrix is

$$G_c(s) = \left[\begin{array}{c|c} A_F & I \\ \hline C_F & 0 \end{array}\right]$$

and

$$\|G_c x_0\|_2^2 = x_0^* X x_0.$$

The proof of the following theorem requires a preliminary result about internal stability given input-output stability.

Lemma 14.1 *If $u, z \in \mathcal{L}_p[0, \infty)$ for $p \geq 1$ and (C_1, A) is detectable in the system described by equations (14.6) and (14.7), then $x \in \mathcal{L}_p[0, \infty)$. Furthermore, if $p < \infty$, then $x(t) \to 0$ as $t \to \infty$.*

Proof. Since (C_1, A) is detectable, there exists L such that $A + LC_1$ is stable. Let $\hat{x}$ be the state estimate of x given by

$$\dot{\hat{x}} = (A + LC_1)\hat{x} + (LD_{12} + B_2)u - Lz.$$

Then $\hat{x} \in \mathcal{L}_p[0, \infty)$ since z and u are in $\mathcal{L}_p[0, \infty)$. Now let $e = x - \hat{x}$; then

$$\dot{e} = (A + LC_1)e$$

and $e \in \mathcal{L}_p[0, \infty)$. Therefore, $x = e + \hat{x} \in \mathcal{L}_p[0, \infty)$. It is easy to see that $e(t) \to 0$ as $t \to \infty$ for any initial condition $e(0)$. Finally, $x(t) \to 0$ follows from the fact that if $p < \infty$ then $\hat{x} \to 0$. □

Theorem 14.2 *There exists a unique optimal control for the LQR problem, namely $u = Fx$. Moreover,*

$$\min_{u \in \mathcal{L}_2[0,\infty)} \|z\|_2 = \|G_c x_0\|_2 .$$

Note that the optimal control strategy is constant gain state feedback, and this gain is independent of the initial condition x_0.

Proof. With the change of variable $v = u - Fx$, the system can be written as

$$\begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} A_F & B_2 \\ C_F & D_{12} \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}, \qquad x(0) = x_0. \tag{14.12}$$

Now if $v \in \mathcal{L}_2[0, \infty)$, then $x, z \in \mathcal{L}_2[0, \infty)$ and $x(\infty) = 0$ since A_F is stable. Hence $u = Fx + v \in \mathcal{L}_2[0, \infty)$. Conversely, if $u, z \in \mathcal{L}_2[0, \infty)$, then from Lemma 14.1 $x \in \mathcal{L}_2[0, \infty)$. So $v \in \mathcal{L}_2[0, \infty)$. Thus the mapping $v = u - Fx$ between $v \in \mathcal{L}_2[0, \infty)$ and those $u \in \mathcal{L}_2[0, \infty)$ that make $z \in \mathcal{L}_2[0, \infty)$ is one-to-one and onto. Therefore,

$$\min_{u \in \mathcal{L}_2[0,\infty)} \|z\|_2 = \min_{v \in \mathcal{L}_2[0,\infty)} \|z\|_2 .$$

By differentiating $x(t)^* X x(t)$ with respect to t along a solution of the differential equation (14.12) and by using (14.9) and the fact that $C_F^* D_{12} = -XB_2$, we see that

$$\begin{aligned} \frac{d}{dt} x^* X x &= \dot{x}^* X x + x^* X \dot{x} \\ &= x^*(A_F^* X + X A_F)x + 2x^* X B_2 v \\ &= -x^* C_F^* C_F x + 2x^* X B_2 v \\ &= -(C_F x + D_{12} v)^*(C_F x + D_{12} v) + 2x^* C_F^* D_{12} v + v^* v + 2x^* X B_2 v \\ &= -\|z\|^2 + \|v\|^2 . \end{aligned} \tag{14.13}$$

Now integrate (14.13) from 0 to ∞ to get

$$\|z\|_2^2 = x_0^* X x_0 + \|v\|_2^2 .$$

Clearly, the unique optimal control is $v = 0$, i.e., $u = Fx$. $\square$

This method of proof, involving change of variables and the completion of the square, is a standard technique and variants of it will be used throughout this book. An alternative proof can be given in frequency domain. To do that, let us first note the following fact:

Lemma 14.3 *Let a transfer matrix be defined as*

$$U := \left[\begin{array}{c|c} A_F & B_2 \\ \hline C_F & D_{12} \end{array}\right] \in \mathcal{RH}_\infty .$$

Then U is inner and $U^\sim G_c \in \mathcal{RH}_2^\perp$.

Proof. The proof uses standard manipulations of state space realizations. From U we get

$$U^\sim(s) = \left[\begin{array}{c|c} -A_F^* & -C_F^* \\ \hline B_2^* & D_{12}^* \end{array}\right].$$

Then it is easy to compute

$$U^\sim U = \left[\begin{array}{cc|c} -A_F^* & -C_F^* C_F & -C_F^* D_{12} \\ 0 & A_F & B_2 \\ \hline B_2^* & D_{12}^* C_F & I \end{array}\right], \quad U^\sim G_c = \left[\begin{array}{cc|c} -A_F^* & -C_F^* C_F & 0 \\ 0 & A_F & I \\ \hline B_2^* & D_{12}^* C_F & 0 \end{array}\right].$$

Now do the similarity transformation

$$\begin{bmatrix} I & -X \\ 0 & I \end{bmatrix}$$

on the states of the transfer matrices and use (14.11) to get

$$U^\sim U = \left[\begin{array}{cc|c} -A_F^* & 0 & 0 \\ 0 & A_F & B_2 \\ \hline B_2^* & 0 & I \end{array}\right] = I$$

$$U^\sim G_c = \left[\begin{array}{cc|c} -A_F^* & 0 & -X \\ 0 & A_F & I \\ \hline B_2^* & 0 & 0 \end{array}\right] = \left[\begin{array}{c|c} -A_F^* & -X \\ \hline B_2^* & 0 \end{array}\right] \in \mathcal{RH}_2^\perp .$$

□

An alternative proof of Theorem 14.2 We have in the frequency domain

$$z = G_c x_0 + Uv.$$

Let $v \in \mathcal{H}_2$. By Lemma 14.3, $G_c x_0$ and Uv are orthogonal. Hence

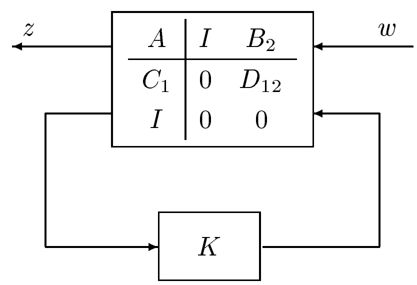
$$\|z\|_2^2 = \|G_c x_0\|_2^2 + \|Uv\|_2^2 .$$

Since U is inner, we get

$$\|z\|_2^2 = \|G_c x_0\|_2^2 + \|v\|_2^2 .$$

This equation immediately gives the desired conclusion. □

Remark 14.2 It is clear that the LQR problem considered above is essentially equivalent to minimizing the 2-norm of z with the input $w = x_0\delta(t)$ in the following diagram:



But this problem is a special $\mathcal{H}_2$ norm minimization problem considered in a later section. ♡

14.3 Extended LQR Problem

This section considers the extended LQR problem where no detectability assumption is made for (C_1, A).

Extended LQR Problem

Let a dynamical system be given by

$$\begin{aligned} \dot{x} &= Ax + B_2 u, \quad x(0) = x_0 \ \text{ given but arbitrary} \\ z &= C_1 x + D_{12} u \end{aligned}$$

with the following assumptions:

(A1) (A, B_2) is stabilizable;

(A2) D_{12} has full column rank with $\begin{bmatrix} D_{12} & D_\perp \end{bmatrix}$ unitary;

(A3) $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all ω.

Find an optimal control law $u \in \mathcal{L}_2[0,\infty)$ such that the system is internally stable, i.e., $x \in \mathcal{L}_2[0,\infty)$ and the performance criterion $\|z\|_2^2$ is minimized.

Assume the same notation as above, and we have

Theorem 14.4 *There exists a unique optimal control for the extended LQR problem, namely $u = Fx$. Moreover,*

$$\min_{u \in \mathcal{L}_2[0,\infty)} \|z\|_2 = \|G_c x_0\|_2 .$$

Proof. The proof of this theorem is very similar to the proof of the standard LQR problem except that, in this case, the input/output stability may not necessarily imply the internal stability. Instead, the internal stability is guaranteed by the way of choosing control law.

Suppose that $u \in \mathcal{L}_2[0,\infty)$ is such a control law that the system is stable, i.e., $x \in \mathcal{L}_2[0,\infty)$. Then $v = u - Fx \in \mathcal{L}_2[0,\infty)$. On the other hand, let $v \in \mathcal{L}_2[0,\infty)$ and consider

$$\begin{bmatrix} \dot{x} \\ z \end{bmatrix} = \begin{bmatrix} A_F & B_2 \\ C_F & D_{12} \end{bmatrix} \begin{bmatrix} x \\ v \end{bmatrix}, \quad x(0) = x_0.$$

Then $x, z \in \mathcal{L}_2[0,\infty)$ and $x(\infty) = 0$ since A_F is stable. Hence $u = Fx + v \in \mathcal{L}_2[0,\infty)$. Again the mapping $v = u - Fx$ between $v \in \mathcal{L}_2[0,\infty)$ and those $u \in \mathcal{L}_2[0,\infty)$ that make $z \in \mathcal{L}_2[0,\infty)$ and $x \in \mathcal{L}_2[0,\infty)$ is one to one and onto. Therefore,

$$\min_{u \in \mathcal{L}_2[0,\infty)} \|z\|_2 = \min_{v \in \mathcal{L}_2[0,\infty)} \|z\|_2 .$$

Using the same technique as in the proof of the standard LQR problem, we have

$$\|z\|_2^2 = x_0^* X x_0 + \|v\|_2^2 .$$

And the unique optimal control is $v = 0$, i.e., $u = Fx$. □

14.4 Guaranteed Stability Margins of LQR

Now we will consider the system described by equation (14.6) with the LQR control law $u = Fx$. The closed-loop block diagram is as shown in Figure 14.1.

The following result is the key to stability margins of an LQR control law.

Lemma 14.5 *Let $F = -(B_2^* X + D_{12}^* C_1)$ and define $G_{12} = D_{12} + C_1(sI - A)^{-1}B_2$. Then*

$$\left(I - B_2^*(-sI - A^*)^{-1}F^*\right)\left(I - F(sI - A)^{-1}B_2\right) = G_{12}^\sim(s) G_{12}(s).$$

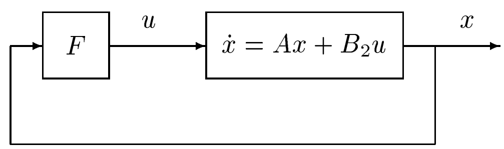


Figure 14.1: LQR closed-loop system

Proof. Note that the Riccati equation (14.9) can be written as

$$XA + A^*X - F^*F + C_1^*C_1 = 0.$$

Add and subtract sX to the equation to get

$$-X(sI - A) - (-sI - A^*)X - F^*F + C_1^*C_1 = 0.$$

Now multiply the above equation from the left by $B_2^*(-sI - A^*)^{-1}$ and from the right by $(sI - A)^{-1}B_2$ to get

$$\begin{aligned} -B_2^*(-sI - A^*)^{-1}XB_2 - B_2^*X(sI - A)^{-1}B_2 - B_2^*(-sI - A^*)^{-1}F^*F(sI - A)^{-1}B_2 \\ +B_2^*(-sI - A^*)^{-1}C_1^*C_1(sI - A)^{-1}B_2 = 0. \end{aligned}$$

Using $-B_2^*X = F + D_{12}^*C_1$ in the above equation, we have

$$\begin{aligned} B_2^*(-sI - A^*)^{-1}F^* + F(sI - A)^{-1}B_2 - B_2^*(-sI - A^*)^{-1}F^*F(sI - A)^{-1}B_2 \\ +B_2^*(-sI - A^*)^{-1}C_1^*D_{12} + D_{12}^*C_1(sI - A)^{-1}B_2 \\ +B_2^*(-sI - A^*)^{-1}C_1^*C_1(sI - A)^{-1}B_2 = 0. \end{aligned}$$

Then the result follows from completing the square and from the fact that $D_{12}^*D_{12} = I$. □

Corollary 14.6 *Suppose* $D_{12}^*C_1 = 0$. *Then*

$$\left(I - B_2^*(-sI - A^*)^{-1}F^*\right)\left(I - F(sI - A)^{-1}B_2\right) = I + B_2^*(-sI - A^*)^{-1}C_1^*C_1(sI - A)^{-1}B_2.$$

In particular,

$$\left(I - B_2^*(-j\omega I - A^*)^{-1}F^*\right)\left(I - F(j\omega I - A)^{-1}B_2\right) \geq I \tag{14.14}$$

and

$$\left(I + B_2^*(-j\omega I - A^* - F^*B_2^*)^{-1}F^*\right)\left(I + F(j\omega I - A - B_2F)^{-1}B_2\right) \leq I. \tag{14.15}$$

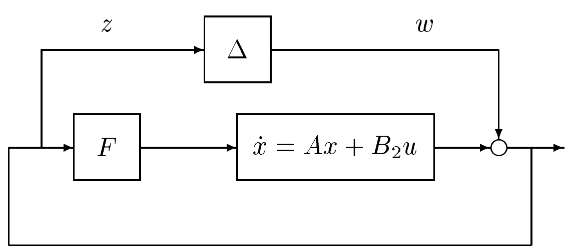
Note that the inequality (14.15) follows from taking the inverse of inequality (14.14).

Define $G(s) = -F(sI - A)^{-1}B_2$ and assume for the moment that the system is single input. Then the inequality (14.14) shows that the open-loop Nyquist diagram of the system $G(s)$ in Figure 14.1 never enters the unit disk centered at $(-1, 0)$ of the complex plane. Hence the system has at least the following stability margins:

$$k_{\min} \leq \frac{1}{2}, \quad k_{\max} = \infty, \quad \phi_{\min} \leq -60^o, \quad \phi_{\max} \geq 60^o$$

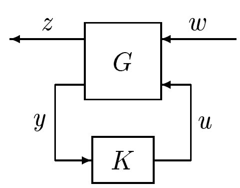
i.e., the system has at least a $6dB$ $(= 20 \log 2)$ gain margin and a 60^o phase margin in both directions. A similar interpretation may be generalized to multiple input systems.

Next, it is noted that the inequality (14.15) can also be given some robustness interpretation. In fact, it implies that the closed-loop system in Figure 14.1 is stable even if the open-loop system $G(s)$ is perturbed additively by a $\Delta \in \mathcal{RH}_\infty$ as long as $\|\Delta\|_\infty < 1$. This can be seen from the following block diagram and small gain theorem where the transfer matrix from w to z is exactly $I + F(j\omega I - A - B_2F)^{-1}B_2$.



14.5 Standard $\mathcal{H}_2$ Problem

The system considered in this section is described by the following standard block diagram:



The realization of the transfer matrix G is taken to be of the form

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right].$$

Notice the special off-diagonal structure of D: D_{22} is assumed to be zero so that G_{22} is strictly proper[1]; also, D_{11} is assumed to be zero in order to guarantee that the $\mathcal{H}_2$ problem properly posed.[2] The case for $D_{11} \neq 0$ will be discussed in Section 14.7.

The following additional assumptions are made for the output feedback $\mathcal{H}_2$ problem in this chapter:

(i) (A, B_2) is stabilizable and (C_2, A) is detectable;

(ii) D_{12} has full column rank with $\begin{bmatrix} D_{12} & D_\perp \end{bmatrix}$ unitary, and D_{21} has full row rank with $\begin{bmatrix} D_{21} \\ \tilde{D}_\perp \end{bmatrix}$ unitary;

(iii) $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all ω;

(iv) $\begin{bmatrix} A - j\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ has full row rank for all ω.

The first assumption is for the stabilizability of G by output feedback, and the third and the fourth assumptions together with the first guarantee that the two Hamiltonian matrices associated with the $\mathcal{H}_2$ problem below belong to $dom(Ric)$. The rank assumptions (ii) are necessary to guarantee that the $\mathcal{H}_2$ optimal controller is a finite dimensional linear time invariant one, while the unitary assumptions are made for the simplicity of the final solution; they are not restrictions (see e.g., Chapter 17).

> $\mathcal{H}_2$ **Problem** *The $\mathcal{H}_2$ control problem is to find a proper, real-rational controller K which stabilizes G internally and minimizes the $\mathcal{H}_2$-norm of the transfer matrix T_{zw} from w to z.*

In the following discussions we shall assume that we have state models of G and K. Recall that a controller is said to be admissible if it is internally stabilizing and proper.

We now state the solution of the problem and then take up its derivation in the next several sections. By Corollary 13.10 the two Hamiltonian matrices

$$H_2 := \begin{bmatrix} A & 0 \\ -C_1^*C_1 & -A^* \end{bmatrix} - \begin{bmatrix} B_2 \\ -C_1^*D_{12} \end{bmatrix} \begin{bmatrix} D_{12}^*C_1 & B_2^* \end{bmatrix}$$

$$J_2 := \begin{bmatrix} A^* & 0 \\ -B_1B_1^* & -A \end{bmatrix} - \begin{bmatrix} C_2^* \\ -B_1D_{21}^* \end{bmatrix} \begin{bmatrix} D_{21}B_1^* & C_2 \end{bmatrix}$$

[1] As we have discussed in Section 12.3.4 of Chapter 12 there is no loss of generality in making this assumption since the controller for D_{22} nonzero case can be recovered from the zero case.

[2] Recall that a rational proper stable transfer function is an $\mathcal{RH}_2$ function iff it is strictly proper.

belong to $dom(Ric)$, and, moreover, $X_2 := Ric(H_2) \geq 0$ and $Y_2 := Ric(J_2) \geq 0$. Define

$$F_2 := -(B_2^* X_2 + D_{12}^* C_1), \quad L_2 := -(Y_2 C_2^* + B_1 D_{21}^*)$$

and

$$A_{F_2} := A + B_2 F_2, \quad C_{1F_2} := C_1 + D_{12} F_2$$

$$A_{L_2} := A + L_2 C_2, \quad B_{1L_2} := B_1 + L_2 D_{21}$$

$$\hat{A}_2 := A + B_2 F_2 + L_2 C_2$$

$$G_c(s) := \left[\begin{array}{c|c} A_{F_2} & I \\ \hline C_{1F_2} & 0 \end{array}\right], \qquad G_f(s) := \left[\begin{array}{c|c} A_{L_2} & B_{1L_2} \\ \hline I & 0 \end{array}\right].$$

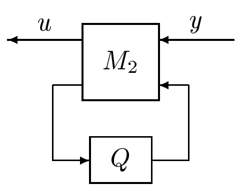
Theorem 14.7 *There exists a unique optimal controller*

$$K_{opt}(s) := \left[\begin{array}{c|c} \hat{A}_2 & -L_2 \\ \hline F_2 & 0 \end{array}\right].$$

Moreover, $\min \|T_{zw}\|_2^2 = \|G_c B_1\|_2^2 + \|F_2 G_f\|_2^2 = \|G_c L_2\|_2^2 + \|C_1 G_f\|_2^2$.

The controller K_{opt} has the well-known separation structure, which will be discussed in more detail in Section 14.9. For comparison with the $\mathcal{H}_\infty$ results, it is useful to describe all suboptimal controllers.

Theorem 14.8 *The family of all admissible controllers such that* $\|T_{zw}\|_2 < \gamma$ *equals the set of all transfer matrices from* y *to* u *in*



$$M_2(s) = \left[\begin{array}{c|cc} \hat{A}_2 & -L_2 & B_2 \\ \hline F_2 & 0 & I \\ -C_2 & I & 0 \end{array}\right]$$

where $Q \in \mathcal{RH}_2$, $\|Q\|_2^2 < \gamma^2 - (\|G_c B_1\|_2^2 + \|F_2 G_f\|_2^2)$.

Thus, the suboptimal controllers are parameterized by a fixed (independent of γ) linear-fractional transformation with a free parameter Q. With $Q = 0$, we recover K_{opt}. It is worth noting that the parameterization in Theorem 14.8 makes T_{zw} affine in Q and yields the Youla parameterization of all stabilizing controllers when the conditions on Q are replaced by $Q \in \mathcal{RH}_\infty$.

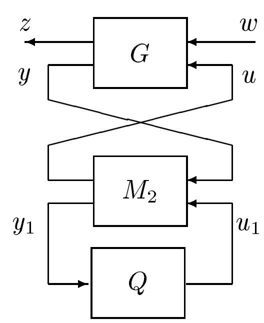
14.6 Optimal Controlled System

In this section, we look at the controller,

$$K(s) = \mathcal{F}_\ell(M_2, Q), \quad Q \in \mathcal{RH}_\infty$$

connected to G. (Keep in mind that all admissible controllers are parameterized by the above formula). We will give a brief analysis of the closed-loop system. It will be seen that a direct consequence from this analysis is the results of Theorem 14.7 and 14.8. The proof given here is not our emphasis. The reason is that this approach does not generalize nicely to other control problems and is often very involved. An alternative proof will be given in the later part of this chapter by using the FI and OE results discussed in section 14.8 and the separation argument. The idea of separation is the main theme for synthesis. We shall now analyze the system structure under the control of a such controller. In particular, we will compute explicitly $\|T_{zw}\|_2^2$.

Consider the following system diagram with controller $K(s) = \mathcal{F}_\ell(M_2, Q)$:



Then $T_{zw} = \mathcal{F}_\ell(N, Q)$ with

$$N = \left[\begin{array}{cc|cc} A_{F_2} & -B_2F_2 & B_1 & B_2 \\ 0 & A_{L_2} & B_{1L_2} & 0 \\ \hline C_{1F_2} & -D_{12}F_2 & 0 & D_{12} \\ 0 & C_2 & D_{21} & 0 \end{array}\right].$$

Define

$$U = \left[\begin{array}{c|c} A_{F_2} & B_2 \\ \hline C_{1F_2} & D_{12} \end{array}\right], \quad V = \left[\begin{array}{c|c} A_{L_2} & B_{1L_2} \\ \hline C_2 & D_{21} \end{array}\right].$$

We have

$$T_{zw} = G_cB_1 - UF_2G_f + UQV.$$

It follows from Lemma 14.3 that G_cB_1 and U are orthogonal. Thus

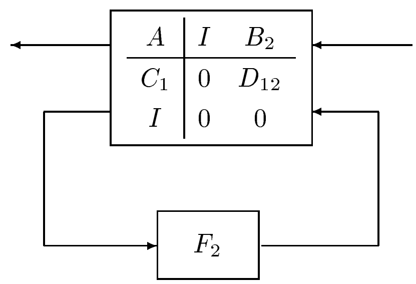
$$\|T_{zw}\|_2^2 = \|G_cB_1\|_2^2 + \|UF_2G_f - UQV\|_2^2 = \|G_cB_1\|_2^2 + \|F_2G_f - QV\|_2^2.$$

It can also be shown easily by duality that G_f and V are orthogonal, i.e., $G_f V^{\sim} \in \mathcal{RH}_2^{\perp}$, and V is a co-inner, so we have

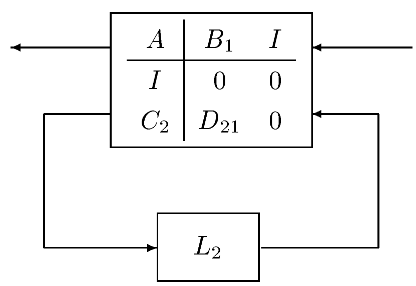
$$\|T_{zw}\|_2^2 = \|G_c B_1\|_2^2 + \|F_2 G_f - QV\|_2^2 = \|G_c B_1\|_2^2 + \|F_2 G_f\|_2^2 + \|Q\|_2^2 .$$

This shows clearly that $Q = 0$ gives the unique optimal control, so $K = \mathcal{F}_\ell(M_2, 0)$ is the unique optimal controller. Note also that $\|T_{zw}\|_2$ is finite if and only if $Q \in \mathcal{RH}_2$. Hence Theorem 14.7 and 14.8 follow easily.

It is interesting to examine the structure of G_c and G_f. First of all the transfer matrix G_c can be represented as a fixed system with the feedback matrix F_2 wrapped around it:



F_2 is, in fact, an optimal LQR controller and minimizes the $\mathcal{H}_2$ norm of G_c. Similarly, G_f can be represented as



and L_2 minimizes the $\mathcal{H}_2$ norm of G_f and solves a special filtering problem.

14.7 $\mathcal{H}_2$ Control with Direct Disturbance Feedforward*

Let us consider the generalized system structure again with D_{11} not necessarily zero:

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right]$$

We shall consider the following question: what will happen and under what condition will the $\mathcal{H}_2$ optimal control problem make sense if $D_{11} \neq 0$?

Recall that $\mathcal{F}_\ell(M_2, Q)$ with $Q \in \mathcal{RH}_\infty$ parameterizes all stabilizing controllers for G regardless of $D_{11} = 0$ or not. Now again consider the closed loop transfer matrix with the controller $K = \mathcal{F}_\ell(M_2, Q)$; then

$$T_{zw} = G_c B_1 - U F_2 G_f + UQV + D_{11}$$

and

$$T_{zw}(\infty) = D_{12} Q(\infty) D_{21} + D_{11}.$$

Hence the $\mathcal{H}_2$ optimal control problem will make sense, i.e., having finite $\mathcal{H}_2$ norm, if and only if there is a constant $Q(\infty)$ such that

$$D_{12} Q(\infty) D_{21} + D_{11} = 0.$$

This requires that

$$Q(\infty) = -D_{12}^* D_{11} D_{21}^*$$

and that

$$-D_{12} D_{12}^* D_{11} D_{21}^* D_{21} + D_{11} = 0. \tag{14.16}$$

Note that the equation (14.16) is a very restrictive condition. For example, suppose

$$D_{12} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad D_{21} = \begin{bmatrix} 0 & I \end{bmatrix}$$

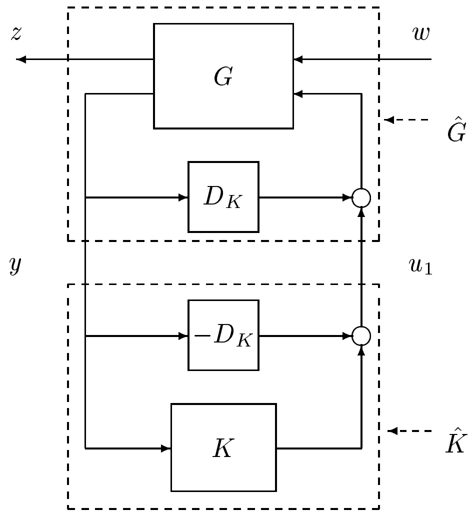
and D_{11} is partitioned accordingly

$$D_{11} = \begin{bmatrix} D_{1111} & D_{1112} \\ D_{1121} & D_{1122} \end{bmatrix}.$$

Then equation (14.16) implies that

$$\begin{bmatrix} D_{1111} & D_{1112} \\ D_{1121} & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

and that $Q(\infty) = -D_{1122}$. So only D_{1122} can be nonzero for a sensible $\mathcal{H}_2$ problem. Hence from now on in this section we shall assume that (14.16) holds and denotes $D_K := -D_{12}^* D_{11} D_{21}^*$. To find the optimal control law for the system G with $D_{11} \neq 0$, let us consider the following system configuration:



Then

$$\hat{G} = \left[\begin{array}{c|cc} A + B_2 D_K C_2 & B_1 + B_2 D_K D_{21} & B_2 \\ \hline C_1 + D_{12} D_K C_2 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right]$$

and

$$K = D_K + \hat{K}.$$

It is easy to check that the system $\hat{G}$ satisfies all assumptions in Section 14.5; hence the controller formula in Section 14.5 can be used. A little bit of algebra will show that

$$\hat{K} = \left[\begin{array}{c|c} \hat{A}_2 - B_2 D_K C_2 & -(L_2 - B_2 D_K) \\ \hline F_2 - D_K C_2 & 0 \end{array}\right]$$

is the $\mathcal{H}_2$ optimal controller for $\hat{G}$. Hence the controller K for the original system G will be given by

$$K = \left[\begin{array}{c|c} \hat{A}_2 - B_2 D_K C_2 & -(L_2 - B_2 D_K) \\ \hline F_2 - D_K C_2 & D_K \end{array}\right] = \mathcal{F}_\ell(M_2, D_K).$$

14.8 Special Problems

In this section we look at various $\mathcal{H}_2$-optimization problems from which the output feedback solutions of the previous sections will be constructed via a separation argument. All the special problems in this section are to find K stabilizing G and minimizing the $\mathcal{H}_2$-norm from w to z in the standard setup, but with different structures for G. As in Chapter 12, we shall call these special problems, respectively, state feedback (SF), output injection (OI), full information (FI), full control (FC), disturbance feedforward (DF), and output estimation (OE). OI, FC, and OE are natural duals of SF, FI, and DF, respectively. The output feedback solutions will be constructed out of the FI and OE results.

The special problems SF, OI, FI, and FC are not, strictly speaking, special cases of the output feedback problem since they do not satisfy all of the assumptions for output feedback (while DF and OE do). Each special problem inherits some of the assumptions (i)-(iv) from the output feedback as appropriate. The assumptions will be discussed in the subsections for each problem.

In each case, the results are summarized as a list of three items; (in all cases, K must be admissible)

1. the minimum of $\|T_{zw}\|_2$;
2. the unique controller minimizing $\|T_{zw}\|_2$;
3. the family of all controllers such that $\|T_{zw}\|_2 < \gamma$, where γ is greater than the minimum norm.

Warning: we will be more specific below about what we mean about the uniqueness and all controllers in the second and third item. In particular, the controllers characterized here for SF, OI, FI and FC problems are neither unique nor all-inclusive. Once again we regard all controllers that give the same control signal u, i.e., having the same transfer function from w to u, as an equivalence class. In other words, if K_1 and K_2 generate the same control signal u, we will regard them as the same, denoted as $K_1 \cong K_2$. Hence the "unique controller" here means one of the controllers from the unique equivalence class. The same comments apply to the "all" situation. This will be much clearer in section 14.8.1 when we consider the state feedback problem. In that case we actually give a parameterization of all the elements in the equivalence class of the "unique" optimal controllers. Thus the unique controller is really not unique. We chose not to give the parameterization of all the elements in an equivalence class in this book since it is very messy, as can be seen in section 14.8.1, and not very useful. However, it will be seen that this equivalence class problem will not occur in the general output feedback case including DF and OE problems.

14.8.1 State Feedback

Consider an open-loop system transfer matrix

$$G_{SF}(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ I & 0 & 0 \end{array}\right]$$

with the following assumptions:

(i) (A, B_2) is stabilizable;

(ii) D_{12} has full column rank with $\begin{bmatrix} D_{12} & D_\perp \end{bmatrix}$ unitary;

(iii) $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all ω.

This is very much like the LQR problem except that we require from the start that u be generated by state feedback and that the detectability of (C_1, A) is not imposed since the controllers are restricted to providing internal stability. The controller is allowed to be dynamic, but it turns out that dynamics are not necessary.

State Feedback:

1. $\min \|T_{zw}\|_2 = \|G_c B_1\|_2 = (\mathrm{trace}(B_1^* X_2 B_1))^{1/2}$

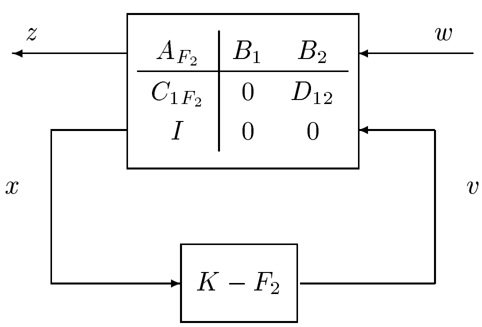
2. $K(s) \cong F_2$

Remark 14.3 The class of all suboptimal controllers for state feedback are messy and are not very useful in this book, so they are omitted, as are the OI problems. ♡

Proof. Let K be a stabilizing controller, $u = K(s)x$. Change control variables by defining $v := u - F_2 x$ and then write the system equations as

$$\begin{bmatrix} \dot{x} \\ z \\ v \end{bmatrix} = \begin{bmatrix} A_{F_2} & B_1 & B_2 \\ C_{1F_2} & 0 & D_{12} \\ (K - F_2) & 0 & 0 \end{bmatrix} \begin{bmatrix} x \\ w \\ v \end{bmatrix}.$$

The block diagram is



Let T_{vw} denote the transfer matrix from w to v. Notice that $T_{vw} \in \mathcal{RH}_2$ because K stabilizes G. Then

$$T_{zw} = G_c B_1 + U T_{vw}$$

where $U = \left[\begin{array}{c|c} A_{F_2} & B_2 \\ \hline C_{1F_2} & D_{12} \end{array}\right]$, and by Lemma 14.3 U is inner and $U^{\sim}G_c$ is in $\mathcal{RH}_2^{\perp}$. We get

$$\|T_{zw}\|_2^2 = \|G_cB_1\|_2^2 + \|T_{vw}\|_2^2 .$$

Thus $\min \|T_{zw}\|_2^2 = \|G_cB_1\|_2^2$ and the minimum is achieved iff $T_{vw} = 0$. Furthermore, $K = F_2$ is a controller achieving this minimum, and any other controllers achieving minimum are in the equivalence class of F_2. □

Note that the above proof actually yields a much stronger result than what is needed. The proof that the optimal T_{vw} is $T_{vw} = 0$ does not depend on the restriction that the controller measures just the state. We only require that the controller produce v as a causal stable function T_{vw} of w. This means that the optimal state feedback is also optimal for the full information problem as well.

We now give some further explanation about the uniqueness of the optimal controller that we commented on before. The important observation for this issue is that the controllers making $T_{vw} = 0$ are not unique. The controller given above, F_2, is only one of them. We will now try to find all of those controllers that stabilize the system and give $T_{vw} = 0$, i.e., all $K(s) \cong F_2$.

Proposition 14.9 *Let V_c be a matrix whose columns form a basis for $KerB_1^*$ ($V_c^*B_1 = 0$). Then all $\mathcal{H}_2$ optimal state feedback controllers can be parameterized as $K_{opt} = \mathcal{F}_\ell(M_{sf}, \Theta)$ with $\Theta \in \mathcal{RH}_2$ and*

$$M_{sf} = \begin{bmatrix} F_2 & I \\ V_c^*(sI - A_{F_2}) & -V_c^*B_2 \end{bmatrix}.$$

Proof. Since

$$T_{vw} = \left(I - \left[\begin{array}{c|c} A_{F_2} & B_2 \\ \hline K - F_2 & 0 \end{array}\right]\right)^{-1} \left[\begin{array}{c|c} A_{F_2} & B_1 \\ \hline K - F_2 & 0 \end{array}\right] = 0,$$

we get

$$(K - F_2)(sI - A_{F_2})^{-1}B_1 = 0 \tag{14.17}$$

which is achieved if $K = F_2$. Clearly, this is the only solution if B_1 is square and nonsingular or if K is restricted to be constant and (A_{F_2}, B_1) is controllable.

To parameterize all elements in the equivalence class ($T_{vw} = 0$) to which $K = F_2$ belongs, let

$$P_c(s) := \left[\begin{array}{c|c} A_{F_2} & B_2 \\ \hline I & 0 \end{array}\right].$$

Then all state feedback controllers stabilizing G can be parameterized as

$$K(s) = F_2 + (I + QP_c)^{-1}Q, \quad Q(s) \in \mathcal{RH}_\infty$$

where Q is free. Substitute K in equation (14.17), and we get

$$Q(s)(sI - A_{F_2})^{-1}B_1 = 0. \tag{14.18}$$

Hence we have

$$Q(s)(sI - A_{F_2})^{-1} = \Theta(s)V_c^*, \quad \Theta(s) \in \mathcal{RH}_2.$$

Thus all $Q(s) \in \mathcal{RH}_\infty$ satisfying (14.18) can be written as

$$Q(s) = \Theta(s)V_c^*(sI - A_{F_2}), \quad \Theta(s) \in \mathcal{RH}_2. \tag{14.19}$$

Therefore,

$$\begin{aligned} K_{opt}(s) &= F_2 + (I + \Theta(s)V_c^*(sI - A_{F_2})P_c)^{-1}\,\Theta(s)V_c^*(sI - A_{F_2}) \\ &= F_2 + (I + \Theta(s)V_c^*B_2)^{-1}\,\Theta(s)V_c^*(sI - A_{F_2}), \quad \Theta(s) \in \mathcal{RH}_2 \end{aligned}$$

parameterizes the equivalence class of $K = F_2$. □

14.8.2 Full Information and Other Special Problems

We shall consider the FI problem first.

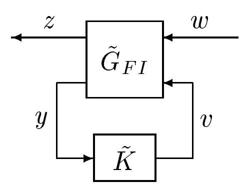
$$G_{FI}(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}\right]$$

The assumptions relevant to the FI problem are the same as the state feedback problem. This is similar to the state feedback problem except that the controller now has more information (w). However, as was pointed out in the discussion of the state feedback problem, this extra information is not used by the optimal controller.

Full Information:

1. $\min \|T_{zw}\|_2 = \|G_cB_1\|_2 = (\text{trace}(B_1^*X_2B_1))^{1/2}$
2. $K(s) \cong \begin{bmatrix} F_2 & 0 \end{bmatrix}$
3. $K(s) \cong \begin{bmatrix} F_2 & Q(s) \end{bmatrix}$, *where* $Q \in \mathcal{RH}_2$, $\|Q\|_2^2 < \gamma^2 - \|G_cB_1\|_2^2$

Proof. Items 1 and 2 follow immediately from the proof of the state feedback results because the argument that $T_{vw} = 0$ did not depend on the restriction to state feedback only. Thus we only need to prove item 3. Let K be an admissible controller such that $\|T_{zw}\|_2 < \gamma$. As in the SF proof, define a new control variable $v = u - F_2 x$; then the closed-loop system is as shown below



with

$$\tilde{G}_{FI} = \left[\begin{array}{c|cc} A_{F_2} & B_1 & B_2 \\ \hline C_{1F_2} & 0 & D_{12} \\ \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}\right], \quad \tilde{K} = K - [F_2\ 0].$$

Denote by Q the transfer matrix from w to v; it belongs to $\mathcal{RH}_2$ by internal stability and the fact that D_{12} has full column rank and T_{zw} with $z = C_{1F_2}x + D_{12}Qw$ has finite $\mathcal{H}_2$ norm. Then $u = F_2 x + v = F_2 x + Qw = \begin{bmatrix} F_2 & Q \end{bmatrix} y$ so $K \cong \begin{bmatrix} F_2 & Q \end{bmatrix}$, and $\|T_{zw}\|_2^2 = \|G_c B_1 + UQ\|_2^2 = \|G_c B_1\|_2^2 + \|Q\|_2^2$; hence,

$$\|Q\|_2^2 = \|T_{zw}\|_2^2 - \|G_c B_1\|_2^2 < \gamma^2 - \|G_c B_1\|_2^2.$$

Likewise, one can show that every controller of the form given in item no.3 is admissible and suboptimal. $\square$

The results for DF, OI, FC, and OE follow from the parallel development of Chapter 12.

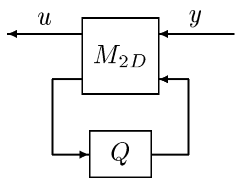
Disturbance Feedforward:

$$G_{DF}(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & I & 0 \end{array}\right]$$

This problem inherits the same assumptions (i)-(iii) as in the state feedback problem, in addition to the stability condition of $A - B_1 C_2$.

1. $\min \|T_{zw}\|_2 = \|G_c B_1\|_2$
2. $K(s) = \left[\begin{array}{c|c} A + B_2 F_2 - B_1 C_2 & B_1 \\ \hline F_2 & 0 \end{array}\right]$

3. the set of all transfer matrices from y *to* u *in*



$$M_{2D}(s) = \left[\begin{array}{c|cc} A + B_2F_2 - B_1C_2 & B_1 & B_2 \\ \hline F_2 & 0 & I \\ -C_2 & I & 0 \end{array}\right]$$

where $Q \in \mathcal{RH}_2$, $\|Q\|_2^2 < \gamma^2 - \|G_cB_1\|_2^2$

Output Injection:

$$G_{OI}(s) = \left[\begin{array}{c|cc} A & B_1 & I \\ \hline C_1 & 0 & 0 \\ C_2 & D_{21} & 0 \end{array}\right]$$

with the following assumptions:

(i) (C_2, A) *is detectable;*

(ii) D_{21} *has full row rank with* $\begin{bmatrix} D_{21} \\ \tilde{D}_\perp \end{bmatrix}$ *unitary;*

(iii) $\begin{bmatrix} A - j\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ *has full row rank for all* ω.

1. $\min \|T_{zw}\|_2 = \|C_1G_f\|_2 = (\mathrm{trace}(C_1Y_2C_1^*))^{1/2}$
2. $K(s) \cong \begin{bmatrix} L_2 \\ 0 \end{bmatrix}$

Full Control:

$$G_{FC}(s) = \left[\begin{array}{c|c|c} A & B_1 & \begin{bmatrix} I & 0 \end{bmatrix} \\ \hline C_1 & 0 & \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix} \\ C_2 & D_{21} & \end{array}\right]$$

with the same assumptions as an output injection problem.

1. $\min \|T_{zw}\|_2 = \|C_1G_f\|_2 = (\mathrm{trace}(C_1Y_2C_1^*))^{1/2}$
2. $K(s) \cong \begin{bmatrix} L_2 \\ 0 \end{bmatrix}$

3. $K(s) \cong \begin{bmatrix} L_2 \\ Q(s) \end{bmatrix}$, *where* $Q \in \mathcal{RH}_2$, $\|Q\|_2^2 < \gamma^2 - \|C_1 G_f\|_2^2$

Output Estimation:

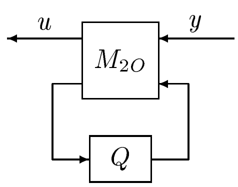
$$G_{OE}(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & I \\ C_2 & D_{21} & 0 \end{array}\right]$$

The assumptions are taken to be those in the output injection problem plus an additional assumption that $A - B_2C_1$ *is stable.*

1. $\min \|T_{zw}\|_2 = \|C_1 G_f\|_2$

2. $K(s) = \left[\begin{array}{c|c} A + L_2C_2 - B_2C_1 & L_2 \\ \hline C_1 & 0 \end{array}\right]$

3. the set of all transfer matrices from y *to* u *in*



$$M_{2O}(s) = \left[\begin{array}{c|cc} A + L_2C_2 - B_2C_1 & L_2 & -B_2 \\ \hline C_1 & 0 & I \\ C_2 & I & 0 \end{array}\right]$$

where $Q \in \mathcal{RH}_2$, $\|Q\|_2^2 < \gamma^2 - \|C_1 G_f\|_2^2$

14.9 Separation Theory

Given the results for the special problems, we can now prove Theorem 14.7 using separation arguments. This essentially involves reducing the output feedback problem to a combination of the Full Information and the Output Estimation problems.

14.9.1 $\mathcal{H}_2$ Controller Structure

Recall that the unique $\mathcal{H}_2$ optimal controller is

$$K_2(s) := \left[\begin{array}{c|c} \hat{A}_2 & -L_2 \\ \hline F_2 & 0 \end{array}\right] = \left[\begin{array}{c|c} A + B_2F_2 + L_2C_2 & Y_2C_2^* \\ \hline -B_2^*X_2 & 0 \end{array}\right]$$

and $$\min \|T_{zw}\|_2^2 = \|G_cB_1\|_2^2 + \|F_2G_f\|_2^2$$

where $X_2 := Ric(H_2)$ and $Y_2 := Ric(J_2)$ and the min is over all stabilizing controllers. Note that F_2 is the optimal state feedback in the Full Information problem and L_2 is the optimal output injection in the Full Control case. The well-known separation property of the $\mathcal{H}_2$ solution is reflected in the fact that K_2 is exactly the optimal output estimate of $F_2 x$ and can be obtained by setting $C_1 = F_2$ in OE.2. Also, the minimum cost is the sum of the FI cost (FI.1) and the OE cost for estimating $F_2 x$ (OE.1).

The controller equations can be written in standard observer form as

$$\begin{aligned} \dot{\hat{x}} &= A\hat{x} + B_2 u + L_2(C_2\hat{x} - y) \\ u &= F_2\hat{x} \end{aligned}$$

where $\hat{x}$ is the optimal estimate of x.

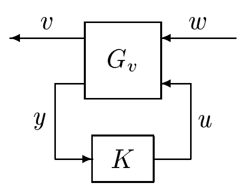
14.9.2 Proof of Theorem 14.7

As before we define a new control variable, $v := u - F_2 x$, and the transfer function to z becomes

$$z = \left[\begin{array}{c|cc} A_{F_2} & B_1 & B_2 \\ \hline C_{1F_2} & 0 & D_{12} \end{array}\right] \begin{bmatrix} w \\ v \end{bmatrix} = G_c B_1 w + U v \tag{14.20}$$

where $G_c(s) := \left[\begin{array}{c|c} A_{F_2} & I \\ \hline C_{1F_2} & 0 \end{array}\right]$ and $U(s) := \left[\begin{array}{c|c} A_{F_2} & B_2 \\ \hline C_{1F_2} & D_{12} \end{array}\right]$. Furthermore, U is inner (i.e., $U^\sim U = I$) and $U^\sim G_c$ belongs to $\mathcal{RH}_2^\perp$ from Lemma 14.3.

Let K be any admissible controller and notice how v is generated:



$$G_v = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline -F_2 & 0 & I \\ C_2 & D_{21} & 0 \end{array}\right]$$

Note that K stabilizes G iff K stabilizes G_v (the two closed-loop systems have identical A-matrices) and that G_v has the form of the Output Estimation problem. From (14.20) and the properties of U we have that

$$\min \|T_{zw}\|_2^2 = \|G_c B_1\|_2^2 + \min \|T_{vw}\|_2^2. \tag{14.21}$$

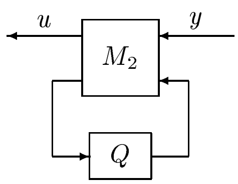
But from item OE.2, $\|T_{vw}\|_2$ is minimized by the controller

$$\left[\begin{array}{c|c} A + B_2F_2 + L_2C_2 & -L_2 \\ \hline F_2 & 0 \end{array}\right],$$

and then from OE.1 $\min \|T_{vw}\|_2 = \|F_2 G_f\|_2$.

14.9.3 Proof of Theorem 14.8

Continuing with the development in the previous proof, we see that the set of all suboptimal controllers equals the set of all K's such that $\|T_{vw}\|_2^2 < \gamma^2 - \|G_c B_1\|_2^2$. Apply item OE.3 to get that such K's are parameterized by



$$M_2(s) = \left[\begin{array}{c|cc} \hat{A}_2 & -L_2 & B_2 \\ \hline F_2 & 0 & I \\ -C_2 & I & 0 \end{array}\right]$$

with $Q \in \mathcal{RH}_2$, $\|Q\|_2^2 < \gamma^2 - \|G_c B_1\|_2^2 - \|F_2 G_f\|_2^2$.

14.10 Stability Margins of $\mathcal{H}_2$ Controllers

We have shown that the system with LQR controller has at least 60^o phase margin and $6dB$ gain margin. However, it is not clear whether these stability margins will be preserved if the states are not available and the output feedback $\mathcal{H}_2$ (or LQG) controller has to be used. The answer is provided here through a counterexample: there are no guaranteed stability margins for a $\mathcal{H}_2$ controller.

Consider a single input and single output two state generalized dynamical system:

$$G(s) = \left[\begin{array}{c|cc} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} & \begin{bmatrix} \sqrt{\sigma} & 0 \\ \sqrt{\sigma} & 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \hline \begin{bmatrix} \sqrt{q} & \sqrt{q} \\ 0 & 0 \end{bmatrix} & 0 & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ \begin{bmatrix} 1 & 0 \end{bmatrix} & \begin{bmatrix} 0 & 1 \end{bmatrix} & 0 \end{array}\right].$$

It can be shown analytically that

$$X_2 = \begin{bmatrix} 2\alpha & \alpha \\ \alpha & \alpha \end{bmatrix}, \quad Y_2 = \begin{bmatrix} 2\beta & \beta \\ \beta & \beta \end{bmatrix}$$

and

$$F_2 = -\alpha \begin{bmatrix} 1 & 1 \end{bmatrix}, \quad L_2 = -\beta \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

where

$$\alpha = 2 + \sqrt{4+q}\,, \quad \beta = 2 + \sqrt{4+\sigma}.$$

Then the optimal output $\mathcal{H}_2$ controller is given by

$$K_{opt} = \left[\begin{array}{cc|c} 1-\beta & 1 & \beta \\ -(\alpha+\beta) & 1-\alpha & \beta \\ \hline -\alpha & -\alpha & 0 \end{array}\right].$$

Suppose that the resulting closed-loop controller (or plant G_{22}) has a scalar gain k with a nominal value $k = 1$. Then the controller implemented in the system is actually

$$K = kK_{opt},$$

and the closed-loop system A-matrix becomes

$$\tilde{A} = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & -k\alpha & -k\alpha \\ \beta & 0 & 1-\beta & 1 \\ \beta & 0 & -\alpha-\beta & 1-\alpha \end{bmatrix}.$$

It can be shown that the characteristic polynomial has the form

$$\det(sI - \tilde{A}) = a_4 s^4 + a_3 s^3 + a_2 s^2 + a_1 s + a_0$$

with

$$a_1 = \alpha + \beta - 4 + 2(k-1)\alpha\beta, \quad a_0 = 1 + (1-k)\alpha\beta.$$

Note that for closed-loop stability it is necessary to have $a_0 > 0$ and $a_1 > 0$. Note also that $a_0 \approx (1-k)\alpha\beta$ and $a_1 \approx 2(k-1)\alpha\beta$ for sufficiently large α and β if $k \neq 1$. It is easy to see that for sufficiently large α and β (or q and σ), the system is unstable for arbitrarily small perturbations in k in either direction. Thus, by choice of q and σ, the gain margins may be made arbitrarily small.

It is interesting to note that the margins deteriorate as control weight $(1/q)$ gets small (large q) and/or system deriving noise gets large (large σ). In modern control folklore, these have often been considered ad hoc means of improving sensitivity.

It is also important to recognize that vanishing margins are not only associated with open-loop unstable systems. It is easy to construct minimum phase, open-loop stable counterexamples for which the margins are arbitrarily small.

The point of these examples is that $\mathcal{H}_2$ (LQG) solutions, unlike LQR solutions, provide no global system-independent guaranteed robustness properties. Like their more classical colleagues, modern LQG designers are obliged to test their margins for each specific design.

It may, however, be possible to improve the robustness of a given design by relaxing the optimality of the filter (or FC controller) with respect to error properties. A successful approach in this direction is the so called LQG loop transfer recovery (LQG/LTR) design technique. The idea is to design a filtering gain (or FC control law) in such way so that the LQG (or $\mathcal{H}_2$) control law will approximate the loop properties of the regular LQR control. This will not be explored further here; interested reader may consult related references.

14.11 Notes and References

The detailed treatment of $\mathcal{H}_2$ related theory, LQ optimal control, Kalman filtering, etc., can be found in Anderson and Moore [1990] or Kwakernaak and Sivan [1972].

15

Linear Quadratic Optimization

This chapter considers time domain characterizations of Hankel operators and Toeplitz operators by means of some related quadratic optimizations. These characterizations will be used to prove a max-min problem which is the key to the $\mathcal{H}_\infty$ theory considered in the next chapter.

15.1 Hankel Operators

Let $G(s)$ be a stable real rational transfer matrix with a state space realization

$$\begin{aligned} \dot{x} &= Ax + Bw \\ z &= Cx + Dw. \end{aligned} \tag{15.1}$$

Consider first the problem of using an input $w \in \mathcal{L}_{2-}$ to maximize $\|P_+ z\|_2^2$. This is exactly the standard problem of computing the Hankel norm of G, i.e., the induced norm of the Hankel operator

$$P_+ M_G : \mathcal{H}_2^\perp \to \mathcal{H}_2,$$

and the norm can be expressed in terms of the controllability Gramian L_c and observability Gramian L_o:

$$AL_c + L_c A^* + BB^* = 0 \qquad A^* L_o + L_o A + C^* C = 0.$$

Although this result is well-known, we will include a time-domain proof similar in technique to the proofs of the optimal $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control results.

Lemma 15.1 $\inf_{w \in \mathcal{L}_{2-}} \left\{ \|w\|_2^2 \;\middle|\; x(0) = x_0 \right\} = x_0^* y_0$ *where* y_0 *solves* $L_c y_0 = x_0$.

Proof. Assume (A, B) is controllable; otherwise, factor out the uncontrollable subspace. Then L_c is invertible and $y_0 = L_c^{-1} x_0$. Moreover, $w \in \mathcal{L}_{2-}$ can be used to produce any $x(0) = x_0$ given $x(-\infty) = 0$. We need to show

$$\inf_{w \in \mathcal{L}_{2-}} \left\{ \|w\|_2^2 \;\middle|\; x(0) = x_0 \right\} = x_0^* L_c^{-1} x_0. \tag{15.2}$$

To show this, we can differentiate $x(t)^* L_c^{-1} x(t)$ along the solution of (15.1) for any given input w as follows:

$$\frac{d}{dt}(x^* L_c^{-1} x) = \dot{x}^* L_c^{-1} x + x^* L_c^{-1} \dot{x} = x^* (A^* L_c^{-1} + L_c^{-1} A) x + 2\langle w, B^* L_c^{-1} x \rangle.$$

Using L_c equation to substitute for $A^* L_c^{-1} + L_c^{-1} A$ and completion of the squares gives

$$\frac{d}{dt}(x^* L_c^{-1} x) = \|w\|^2 - \|w - B^* L_c^{-1} x\|^2.$$

Integration from $t = -\infty$ to $t = 0$ with $x(-\infty) = 0$ and $x(0) = x_0$ gives

$$x_0^* L_c^{-1} x_0 = \|w\|_2^2 - \|w - B^* L_c^{-1} x\|_2^2 \le \|w\|_2^2.$$

If $w(t) = B^* e^{-A^* t} L_c^{-1} x_0 = B^* L_c^{-1} e^{(A + BB^* L_c^{-1})t} x_0$ on $(-\infty, 0]$, then $w \in \mathcal{L}_{2-}$, $w = B^* L_c^{-1} x$ and equality is achieved, thus proving (15.2). $\square$

Lemma 15.2 $\sup_{w \in \mathcal{BL}_{2-}} \|P_+ z\|_2^2 = \sup_{w \in \mathcal{BH}_2^\perp} \|P_+ M_G w\|_2^2 = \rho(L_o L_c)$.

Proof. Given $x(0) = x_0$ and $w = 0$, for $t \ge 0$ the norm of $z(t) = Ce^{At} x_0$ can be found from

$$\|P_+ z\|_2^2 = \int_0^\infty x_0^* e^{A^* t} C^* C e^{At} x_0 \, dt = x_0^* L_o x_0.$$

Combine this result with Lemma 15.1 to give

$$\sup_{w \in \mathcal{BL}_{2-}} \|P_+ z\|_2^2 = \sup_{0 \ne w \in \mathcal{L}_{2-}} \frac{\|P_+ z\|_2^2}{\|w\|_2^2} = \max_{x_0 \ne 0} \frac{x_0^* L_o x_0}{x_0^* L_c^{-1} x_0} = \rho(L_o L_c).$$

$\square$

Remark 15.1 Another useful way to characterize the Hankel norm is to examine the following quadratic optimization with initial condition $x(-\infty) = 0$:

$$\sup_{w \in \mathcal{L}_{2-}} \left\{ \|P_+ z\|_2^2 - \beta^2 \|w\|_2^2 \right\}.$$

It is easy to see from the definition of the Hankel norm that

$$\sup_{0 \neq w \in \mathcal{L}_{2-}} \frac{\|P_+ z\|_2}{\|w\|_2} \leq \beta$$

iff

$$\sup_{0 \neq w \in \mathcal{L}_{2-}} \left\{ \|P_+ z\|_2^2 - \beta^2 \|w\|_2^2 \right\} \leq 0.$$

So the Hankel norm is equal to the smallest β such that the above inequality holds. Now

$$\begin{aligned} \sup_{w \in \mathcal{L}_{2-}} \left\{ \|P_+ z\|_2^2 - \beta^2 \|w\|_2^2 \right\} &= \sup_{x_0 \in \mathbb{R}^n} (x_0^* L_o x_0 - \beta^2 x_0^* L_c^{-1} x_0) \\ &= \begin{cases} 0, & \rho(L_o L_c) \leq \beta^2; \\ +\infty, & \rho(L_o L_c) > \beta^2. \end{cases} \end{aligned}$$

Hence the Hankel norm is equal to the square root of $\rho(L_o L_c)$. ♡

15.2 Toeplitz Operators

If a transfer matrix $G \in \mathcal{RH}_\infty$ and $\|G\|_\infty < 1$, then by Corollary 13.24, the Hamiltonian matrix

$$H = \begin{bmatrix} A + BR^{-1}D^*C & BR^{-1}B^* \\ -C^*(I + DR^{-1}D^*)C & -(A + BR^{-1}D^*C)^* \end{bmatrix}, \qquad R = \gamma^2 I - D^*D$$

with $\gamma = 1$ is in $dom(Ric)$, $X = Ric(H) \geq 0$, $A + BR^{-1}(B^*X + D^*C)$ is stable and

$$A^*X + XA + (XB + C^*D)R^{-1}(B^*X + D^*C) + C^*C = 0. \tag{15.3}$$

The following lemma offers yet another consequence of $\|G\|_\infty < 1$. (Recall that the $\mathcal{H}_\infty$ norm of a stable matrix is the Toeplitz operator norm.)

Lemma 15.3 *Suppose $\|G\|_\infty < 1$ and $x(0) = x_0$. Then*

$$\sup_{w \in \mathcal{L}_{2+}} (\|z\|_2^2 - \|w\|_2^2) = x_0^* X x_0$$

and the sup is achieved.

Proof. We can differentiate $x(t)^*Xx(t)$ as in the last section, use the Riccati equation (15.3) to substitute for $A^*X + XA$, and complete the squares to get

$$\frac{d}{dt}(x^*Xx) = -\|z\|^2 + \|w\|^2 - \|R^{-1/2}[Rw - (B^*X + D^*C)x]\|^2.$$

If $w \in \mathcal{L}_{2+}$, then $x \in \mathcal{L}_{2+}$, so integrating from $t = 0$ to $t = \infty$ gives

$$\|z\|_2^2 - \|w\|_2^2 = x_0^*Xx_0 - \|R^{-1/2}[Rw - (B^*X + D^*C)x]\|_2^2 \leq x_0^*Xx_0. \tag{15.4}$$

If we let $w = R^{-1}(B^*X + D^*C)x = R^{-1}(B^*X + D^*C)e^{[A+BR^{-1}(B^*X+D^*C)]t}x_0$, then $w \in \mathcal{L}_{2+}$ because $A + BR^{-1}(B^*X + D^*C)$ is stable. Thus the inequality in (15.4) can be made an equality and the proof is complete. Note that the sup is achieved for a w which is a linear function of the state. $\square$

As a direct consequence of this lemma, we have the following

Corollary 15.4 *Let $x(0) = x_0$ and $X_0 = X_0^* > 0$.*

(i) if $\|G\|_\infty < \gamma$ and $X = Ric(H)$. Then

$$\begin{aligned}
&\sup_{0\neq(x_0,w)\in\mathbb{R}^n\times\mathcal{L}_2[0,\infty)} \left\{\|z\|_2^2 - \gamma^2(\|w\|_2^2 + x_0^*X_0x_0)\right\} \\
&= \sup_{0\neq x_0\in\mathbb{R}^n} \left\{x_0^*Xx_0 - \gamma^2 x_0^*X_0x_0\right\} \\
&\begin{cases} < 0, & \lambda_{max}(X - \gamma^2 X_0) < 0; \\ = 0, & \lambda_{max}(X - \gamma^2 X_0) = 0; \\ = +\infty, & \lambda_{max}(X - \gamma^2 X_0) > 0. \end{cases}
\end{aligned}$$

(ii) $\displaystyle\sup_{0\neq(x_0,w)\in\mathbb{R}^n\times\mathcal{L}_2[0,\infty)} \frac{\|P_+z\|_2^2}{\|w\|_2^2 + x_0^*X_0x_0} < \gamma^2$ *if and only if $\bar{\sigma}(D) < \gamma$, $H \in dom(Ric)$, and $\lambda_{max}(X - \gamma^2 X_0) < 0$.*

Remark 15.2 The matrix X_0 has the interpretation of the confidence on the initial condition x_0. So if $\underline{\sigma}(X_0)$ is small, then the initial condition is probably not known very well. In that case γ_{min} will be large where γ_{min} denotes the smallest γ such that $\lambda_{max}(X - \gamma^2 X_0) \leq 0$. On the other hand, a large $\underline{\sigma}(X_0)$ implies that the initial condition is known very well and that γ_{min} is determined essentially by the condition $H \in dom(Ric)$. ♡

A dual version of Lemma 15.3 can also be obtained and is useful in characterizing the so-called 2×2 block mixed Hankel-Toeplitz operator. To do that, we first note that[1]

$$G^T(s) = \left[\begin{array}{c|c} A^* & C^* \\ \hline B^* & D^* \end{array}\right]$$

[1]Note that since the system matrices are real, $A^T = A^*, B^T = B^*$, etc. The conjugate transpose is used here for the transpose for the sake of consistency in notation.

and $\|G\|_\infty < 1$ iff $\|G^T\|_\infty < 1$. Let J denote the following Hamiltonian matrix

$$J = \begin{bmatrix} A^* & 0 \\ -BB^* & -A \end{bmatrix} + \begin{bmatrix} C^* \\ -BD^* \end{bmatrix} \tilde{R}^{-1} \begin{bmatrix} DB^* & C \end{bmatrix}$$

where $\tilde{R} := I - DD^*$. Then $J \in dom(Ric)$, $Y = Ric(J) \geq 0$, $A + (YC^* + BD^*)\tilde{R}^{-1}C$ is stable and

$$AY + YA^* + (YC^* + BD^*)\tilde{R}^{-1}(CY + DB^*) + BB^* = 0 \tag{15.5}$$

if $\|G\|_\infty < 1$. For simplicity, we shall assume that (A, B) is controllable; hence, $Y > 0$. The case in which Y is singular can be obtained by restricting $x_0 \in \mathrm{Im}(Y)$ and replacing Y^{-1} by Y^+.

Lemma 15.5 *Suppose $\|G\|_\infty < 1$ and (A, B) is controllable. Then*

$$\sup_{w \in \mathcal{L}_{2-}} \left\{ \|P_- z\|_2^2 - \|w\|_2^2 \;\middle|\; x(0) = x_0 \right\} = -x_0^* Y^{-1} x_0$$

and the sup is achieved.

Proof. Analogous to the proof of Lemma 15.3, we can differentiate $x(t)^* Y^{-1} x(t)$, use the Riccati equation (15.5) to substitute for $AY + YA^*$, and complete the squares to get

$$\frac{d}{dt}(x^* Y^{-1} x) = -\|z\|^2 + \|w\|^2 - \|R^{-1/2}[Rw - (B^* Y^{-1} + D^* C)x]\|^2$$

where $R = I - D^*D > 0$. If $w \in \mathcal{L}_{2-}$, then $x \in \mathcal{L}_{2-}$ and $x(-\infty) = 0$; so integrating from $t = -\infty$ to $t = 0$ gives

$$\|z\|_2^2 - \|w\|_2^2 = -x_0^* Y^{-1} x_0 - \|R^{-1/2}[Rw - (B^* Y^{-1} + D^* C)x]\|_2^2 \leq -x_0^* Y^{-1} x_0. \tag{15.6}$$

If we let $w = R^{-1}(B^* Y^{-1} + D^* C)x = R^{-1}(B^* Y^{-1} + D^* C)e^{[A + BR^{-1}(B^* Y^{-1} + D^* C)]t} x_0$, then $w \in \mathcal{L}_{2-}$ because $A + BR^{-1}(B^* Y^{-1} + D^* C) = -Y\{A + (YC^* + BD^*)\tilde{R}^{-1}C\}^* Y^{-1}$ and $A + (YC^* + BD^*)\tilde{R}^{-1}C$ is stable. Thus the inequality can be made an equality and the proof is complete. $\square$

15.3 Mixed Hankel-Toeplitz Operators

Now suppose that the input is partitioned so that $B = \begin{bmatrix} B_1 & B_2 \end{bmatrix}$, $D = \begin{bmatrix} D_1 & D_2 \end{bmatrix}$,

$$G(s) = \begin{bmatrix} G_1(s) & G_2(s) \end{bmatrix} \in \mathcal{RH}_\infty,$$

and w is partitioned conformably. Then $\|G_2\|_\infty < 1$ iff $\bar{\sigma}(D_2) < 1$, and

$$H_W := \begin{bmatrix} A & 0 \\ -C^*C & -A^* \end{bmatrix} + \begin{bmatrix} B_2 \\ -C^*D_2 \end{bmatrix} R_2^{-1} \begin{bmatrix} D_2^*C & B_2^* \end{bmatrix}$$

is in $dom(Ric)$ where $R_2 := I - D_2^*D_2 > 0$. For $H_W \in dom(Ric)$, define $W = Ric(H_W)$. Let

$$w \in \mathcal{W} := \mathcal{H}_2^\perp \oplus \mathcal{L}_2 = \left\{ \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \middle| w_1 \in \mathcal{H}_2^\perp, w_2 \in \mathcal{L}_2 \right\}. \tag{15.7}$$

We are interested in a test for $\sup_{w \in \mathcal{BW}} \|P_+ z\|_2 < 1$, or, equivalently,

$$\sup_{w \in \mathcal{BW}} \|\Gamma w\|_2 < 1 \tag{15.8}$$

where $\Gamma = P_+[M_{G_1}\ M_{G_2}] : \mathcal{W} \to \mathcal{H}_2$ is a mixed Hankel-Toeplitz operator:

$$\begin{aligned} \Gamma \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} &= P_+ \begin{bmatrix} G_1 & G_2 \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \qquad w_1 \in \mathcal{H}_2^\perp, \quad w_2 \in \mathcal{L}_2 \\ &= P_+ \begin{bmatrix} G_1 & G_2 \end{bmatrix} P_- w + P_+ G_2 P_+ w_2. \end{aligned}$$

Thus Γ is the sum of the Hankel operator $P_+ M_G : \mathcal{H}_2^\perp \oplus \mathcal{H}_2^\perp \to \mathcal{H}_2$ and the Toeplitz operator $P_+ M_{G_2} : \mathcal{H}_2 \to \mathcal{H}_2$. The following lemma generalizes Corollary 13.24 ($B_1 = 0, D_1 = 0$) and Lemma 15.2 ($B_2 = 0, D_2 = 0$).

Lemma 15.6 $\sup_{w \in \mathcal{BW}} \|\Gamma w\|_2 < 1$ *iff the following two conditions hold:*

(i) $\bar{\sigma}(D_2) < 1$ *and* $H_W \in dom(Ric)$*;*

(ii) $\rho(WL_c) < 1$.

Proof. By Corollary 13.24, condition (i) is necessary for (15.8), so we will prove that given condition (i), (15.8) holds iff condition (ii) holds. We will do this by showing, equivalently, that $\rho(WL_c) \geq 1$ iff $\sup_{w \in \mathcal{BW}} \|\Gamma w\|_2 \geq 1$. By definition of $\mathcal{W}$, if $w \in \mathcal{W}$ then

$$\|P_+ z\|_2^2 - \|w\|_2^2 = \|P_+ z\|_2^2 - \|P_+ w_2\|_2^2 - \|P_- w\|_2^2.$$

Note that the last term only contributes to $\|P_+ z\|_2^2$ through $x(0)$. Thus if L_c is invertible, then Lemma 15.1 and 15.3 yield

$$\sup_{w \in \mathcal{W}} \left\{ \|P_+ z\|_2^2 - \|w\|_2^2 \;\middle|\; x(0) = x_0 \right\} \;=\; x_0^* W x_0 - x_0^* L_c^{-1} x_0 \tag{15.9}$$

and the supremum is achieved for some $w \in \mathcal{W}$ that can be constructed from the previous lemmas. Since $\rho(WL_c) \geq 1$ iff $\exists\ x_0 \neq 0$ such that the right-hand side of

(15.9) is ≥ 0, we have, by (15.9), that $\rho(WL_c) \geq 1$ iff $\exists\ w \in \mathcal{W}$, $w \neq 0$ such that $\|P_+ z\|_2^2 \geq \|w\|_2^2$. But this is true iff $\sup_{w \in \mathcal{BW}} \|\Gamma w\|_2 \geq 1$.

If L_c is not invertible, we need only restrict x_0 in (15.9) to $\mathrm{Im}(L_c)$, and then the above argument generalizes in a straightforward way. □

In the differential game problem considered later and in the $\mathcal{H}_\infty$ optimal control problem, we will make use of the adjoint $\Gamma^* : \mathcal{H}_2 \to \mathcal{W}$, which is given by

$$\Gamma^* z = \begin{bmatrix} P_-(G_1^\sim z) \\ G_2^\sim z \end{bmatrix} = \begin{bmatrix} P_- G_1^\sim \\ G_2^\sim \end{bmatrix} z \tag{15.10}$$

where $P_- G z := P_-(Gz) = (P_- M_G) z$. That the expression in (15.10) is actually the adjoint of Γ is easily verified from the definition of the inner product on vector-valued $\mathcal{L}_2$ space. The adjoint of $\Gamma : \mathcal{W} \to \mathcal{H}_2$ is the operator $\Gamma^* : \mathcal{H}_2 \to \mathcal{W}$ such that $< z, \Gamma w > = < \Gamma^* z, w >$ for all $w \in \mathcal{W}$, $z \in \mathcal{H}_2$. By definition, we have

$$\begin{aligned} < z, \Gamma w > &= < z, P_+(G_1 w_1 + G_2 w_2) > = < z, G_1 w_1 > + < z, G_2 w_2 > \\ &= < P_-(G_1^\sim z), w_1 > + < G_2^\sim z, w_2 > \\ &= < \Gamma^* z, w > . \end{aligned}$$

The mixed Hankel-Toeplitz operator just studied is the so-called 2×1-block mixed Hankel-Toeplitz operator. There is a 2×2-block generalization.

15.4 Mixed Hankel-Toeplitz Operators: The General Case*

Historically, the mixed Hankel-Toeplitz operators have played important roles in $\mathcal{H}_\infty$ theory, so it is interesting to consider the 2×2-block generalization of Lemma 15.6. In fact, the whole $\mathcal{H}_\infty$ control theory can be developed using these tools. See Section 17.7 in Chapter 17. The proof of Lemma 15.7 below is completely straightforward and fairly short, given the other results in the previous sections. Suppose that

$$G(s) = \begin{bmatrix} G_{11}(s) & G_{12}(s) \\ G_{21}(s) & G_{22}(s) \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] =: \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty .$$

Denote

$$D_{\bullet 2} := \begin{bmatrix} D_{12} \\ D_{22} \end{bmatrix} \qquad D_{2\bullet} := \begin{bmatrix} D_{21} & D_{22} \end{bmatrix}$$

$$R_x := I - D_{\bullet 2}^* D_{\bullet 2} \qquad R_y := I - D_{2\bullet} D_{2\bullet}^*$$

$$H_X := \begin{bmatrix} A & 0 \\ -C^*C & -A^* \end{bmatrix} + \begin{bmatrix} B_2 \\ -C^*D_{\bullet 2} \end{bmatrix} R_x^{-1} \begin{bmatrix} D_{\bullet 2}^*C & B_2^* \end{bmatrix}$$

$$H_Y := \begin{bmatrix} A^* & 0 \\ -BB^* & -A \end{bmatrix} + \begin{bmatrix} C_2^* \\ -BD_{2\bullet}^* \end{bmatrix} R_y^{-1} \begin{bmatrix} D_{2\bullet}B^* & C_2 \end{bmatrix}.$$

Define $\mathcal{W} = \mathcal{H}_2^\perp \oplus \mathcal{L}_2$, $\mathcal{Z} = \mathcal{H}_2 \oplus \mathcal{L}_2$, and $\Gamma : \mathcal{W} \to \mathcal{Z}$ as

$$\Gamma \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = \begin{bmatrix} P_+ & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}.$$

Lemma 15.7 $\sup_{w\in\mathcal{BW}} \|\Gamma w\|_2 < 1$ *holds iff the following three conditions hold:*

(i) $\bar\sigma(D_{\bullet 2}) < 1$ *and* $H_X \in$ *dom(Ric);*

(ii) $\bar\sigma(D_{2\bullet}) < 1$ *and* $H_Y \in$ *dom(Ric);*

(iii) $\rho(XY) < 1$ *for* $X = Ric(H_X)$ *and* $Y = Ric(H_Y)$.

Proof. The mixed Hankel-Toeplitz operator can be written as

$$\Gamma \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = P_+ G \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} + \begin{bmatrix} 0 \\ P_- \begin{bmatrix} G_{21} & G_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \end{bmatrix}.$$

Hence

$$\|\Gamma w\|_2^2 = \|P_+Gw\|_2^2 + \|P_- \begin{bmatrix} G_{21} & G_{22} \end{bmatrix} w\|_2^2.$$

So $\sup_{w\in\mathcal{BW}} \|\Gamma w\| < 1$ implies

$$\sup_{w\in\mathcal{BW}} \|P_+Gw\| < 1$$

and

$$\sup_{w\in\mathcal{BW}} \|P_- \begin{bmatrix} G_{21} & G_{22} \end{bmatrix} w\| < 1.$$

But

$$\left\| \begin{bmatrix} G_{12} \\ G_{22} \end{bmatrix} \right\|_\infty = \sup_{w_2\in\mathcal{H}_2} \left\| P_+ G \begin{bmatrix} 0 \\ w_2 \end{bmatrix} \right\|_2 \le \sup_{w\in\mathcal{BW}} \|P_+Gw\| < 1$$

and

$$\left\| \begin{bmatrix} G_{21} & G_{22} \end{bmatrix} \right\|_\infty = \sup_{w\in\mathcal{H}_2^\perp} \left\| P_- \begin{bmatrix} G_{21} & G_{22} \end{bmatrix} w \right\|_2 \le \sup_{w\in\mathcal{BW}} \|P_- \begin{bmatrix} G_{21} & G_{22} \end{bmatrix} w\| < 1.$$

These two inequalities then imply that (i) and (ii) are necessary. Analogous to the proof of Lemma 15.6, we will show that given conditions (i) and (ii), $\sup_{w\in\mathcal{BW}} \|\Gamma w\|_2 < 1$

holds iff condition (iii) holds. We will do this by showing, equivalently, that $\rho(XY) \geq 1$ iff $\sup_{w\in\mathcal{BW}} \|\Gamma w\|_2 \geq 1$. By definition of $\mathcal{W}$, if $w \in \mathcal{W}$ then

$$\|\Gamma w\|_2^2 - \|w\|_2^2 = \left(\|P_+ z\|_2^2 - \|P_+ w_2\|_2^2\right) + \left(\|P_- \begin{bmatrix} G_{21} & G_{22} \end{bmatrix} w\|_2^2 - \|P_- w\|_2^2\right).$$

Thus if Y is invertible, then Lemma 15.3 and 15.5 yield

$$\sup_{w\in\mathcal{W}} \left\{ \|\Gamma w\|_2^2 - \|w\|_2^2 \;\middle|\; x(0) = x_0 \right\} \;=\; x_0^* X x_0 - x_0^* Y^{-1} x_0.$$

Now the same arguments as in the proof of Lemma 15.6 give the desired conclusion. □

15.5 Linear Quadratic Max-Min Problem

Consider the dynamical system

$$\dot{x} = Ax + B_1 w + B_2 u \tag{15.11}$$
$$z = C_1 x + D_{12} u \tag{15.12}$$

with the following assumptions:

(i) (C_1, A) is observable;

(ii) (A, B_2) is stabilizable;

(iii) $D_{12}^* \begin{bmatrix} C_1 & D_{12} \end{bmatrix} = \begin{bmatrix} 0 & I \end{bmatrix}$.

In this section, we are interested in answering the following question: when

$$\sup_{w\in\mathcal{BL}_{2+}} \min_{u\in\mathcal{L}_{2+}} \|z\|_2 < 1?$$

Remark 15.3 This max-min problem is a game problem in the sense that u is chosen to minimize the quadratic norm of z and that w is chosen to maximize the norm. In other words, inputs u and w act as "opposing players". This linear quadratic max-min problem can be reformulated in the traditional fashion as in the previous sections:

$$\sup_{w\in\mathcal{L}_{2+}} \min_{u\in\mathcal{L}_{2+}} \int_0^\infty \left(\|z\|^2 - \|w\|^2\right) dt = \sup_{w\in\mathcal{L}_{2+}} \min_{u\in\mathcal{L}_{2+}} \left(\|z\|_2^2 - \|w\|_2^2\right).$$

A conventional game problem setup would be to consider the min-max problem, i.e., switching the order of sup and min. However, it will be seen that they are equivalent and that a saddle point exists. By saying that, we would like to warn readers that this may not be true in the general case where $z = C_1 x + D_{11} w + D_{12} u$ and $D_{11} \neq 0$. In that case, it is possible that $\sup_w \inf_u < \inf_u \sup_w$. This will be elaborated in Chapter 17.

It should also be pointed out that the results presented here still hold, subject to some minor modifications, if the assumptions (i) and (iii) on the dynamical system are relaxed to:

(i)' $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all $\omega \in \mathbb{R}$, and

(iii)' D_{12} has full column rank. ♡

It is clear from the assumptions that $H_2 \in dom(Ric)$ and $X_2 = Ric(H_2) > 0$, where

$$H_2 = \begin{bmatrix} A & -B_2 B_2^* \\ -C_1^* C_1 & -A^* \end{bmatrix}.$$

Let $F_2 = -B_2^* X_2$ and $D_\perp$ be such that $[D_{12}\ D_\perp]$ is an orthogonal matrix. Define

$$G_c(s) := \left[\begin{array}{c|c} A_{F_2} & I \\ \hline C_{1F_2} & 0 \end{array}\right], \quad U(s) := \left[\begin{array}{c|c} A_{F_2} & B_2 \\ \hline C_{1F_2} & D_{12} \end{array}\right]$$

and

$$U_\perp = \left[\begin{array}{c|c} A_{F_2} & -X_2^{-1} C_1^* D_\perp \\ \hline C_{1F_2} & D_\perp \end{array}\right] \tag{15.13}$$

where $A_{F_2} = A + B_2 F_2$ and $C_{1F_2} = C_1 + D_{12} F_2$. The following is easily proven using Lemma 13.29 by obtaining a state-space realization and by eliminating uncontrollable states using a little algebra involving the Riccati equation for X_2.

Lemma 15.8 $[U\ U_\perp]$ *is square and inner and a realization for* $G_c^\sim \begin{bmatrix} U & U_\perp \end{bmatrix}$ *is*

$$G_c^\sim \begin{bmatrix} U & U_\perp \end{bmatrix} = \left[\begin{array}{c|cc} A_{F_2} & -B_2 & X_2^{-1} C_1^* D_\perp \\ \hline X_2 & 0 & 0 \end{array}\right] \in \mathcal{RH}_2. \tag{15.14}$$

This implies that U and $U_\perp$ are each inner and that both $U_\perp^\sim G_c$ and $U^\sim G_c$ are in $\mathcal{RH}_2^\perp$. To answer our earlier question, define a Hamiltonian matrix H_∞ and the associated Riccati equation as

$$H_\infty := \begin{bmatrix} A & B_1 B_1^* - B_2 B_2^* \\ -C_1^* C_1 & -A^* \end{bmatrix}$$

$$A^* X_\infty + X_\infty A + X_\infty B_1 B_1^* X_\infty - X_\infty B_2 B_2^* X_\infty + C_1^* C_1 = 0.$$

Theorem 15.9 $\sup\limits_{w \in \mathcal{BL}_{2+}} \min\limits_{u \in \mathcal{L}_{2+}} \|z\|_2 < 1$ *if and only if* $H_\infty \in dom(Ric)$ *and* $X_\infty = Ric(H_\infty) > 0$. *Furthermore, if the condition is satisfied, then* $u = F_\infty x$ *with* $F_\infty := -B_2^* X_\infty$ *is an optimal control.*

Proof. ($\Rightarrow$) As in Chapter 14, define $\nu := u - F_2 x$ to get

$$z = G_c B_1 w + U\nu.$$

Then the hypothesis implies that

$$\sup_{w\in\mathcal{BH}_2} \min_{\nu\in\mathcal{H}_2} \|z\|_2 < 1. \tag{15.15}$$

Since by Lemma 15.8 $[U\ U_\perp]$ is square and inner, $\|z\|_2 = \|[U\ U_\perp]^\sim z\|_2$, and

$$\begin{bmatrix} U & U_\perp \end{bmatrix}^\sim z = \begin{bmatrix} U^\sim G_c B_1 w + \nu \\ U_\perp^\sim G_c B_1 w \end{bmatrix} = \begin{bmatrix} P_-\left(U^\sim G_c B_1 w\right) + P_+\left(U^\sim G_c B_1 w + \nu\right) \\ U_\perp^\sim G_c B_1 w \end{bmatrix}.$$

Thus

$$\sup_{w\in\mathcal{BH}_2} \min_{\nu\in\mathcal{H}_2} \|z\|_2 = \sup_{w\in\mathcal{BH}_2} \min_{\nu\in\mathcal{H}_2} \left\| \begin{bmatrix} P_-\left(U^\sim G_c B_1 w\right) + P_+\left(U^\sim G_c B_1 w + \nu\right) \\ U_\perp^\sim G_c B_1 w \end{bmatrix} \right\|_2,$$

and the right hand of the above equation is minimized by $\nu = -P_+\left(U^\sim G_c B_1 w\right)$; we have

$$\begin{aligned} \sup_{w\in\mathcal{BH}_2} \min_{\nu\in\mathcal{H}_2} \|z\|_2 \quad &= \quad \sup_{w\in\mathcal{BH}_2} \left\| \begin{bmatrix} P_-\left(U^\sim G_c B_1 w\right) \\ U_\perp^\sim G_c B_1 w \end{bmatrix} \right\|_2 \\ &=: \quad \sup_{w\in\mathcal{BH}_2} \|\Gamma^* w\|_2 < 1 \end{aligned}$$

where $\Gamma^* : \mathcal{L}_{2+} \to \mathcal{W}$ is defined as

$$\Gamma^* w = \begin{bmatrix} P_-(U^\sim G_c B_1 w) \\ U_\perp^\sim G_c B_1 w \end{bmatrix} = \begin{bmatrix} P_- U^\sim \\ U_\perp^\sim \end{bmatrix} G_c B_1 w$$

with

$$\mathcal{W} := \left\{ \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \middle| q_1 \in \mathcal{H}_2^\perp, q_2 \in \mathcal{L}_2 \right\}.$$

Note that from equation (15.10) the adjoint operator $\Gamma : \mathcal{W} \to \mathcal{H}_2$ is given by

$$\Gamma \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} = P_+(B_1^* G_c^\sim (U q_1 + U_\perp q_2)) = P_+ B_1^* G_c^\sim \begin{bmatrix} U & U_\perp \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix}$$

where $G_c^\sim [U\ U_\perp] \in \mathcal{RH}_2$ has the realization in (15.14). So we have

$$\sup_{q\in\mathcal{BW}} \|\Gamma q\|_2 < 1.$$

This is just the condition (15.8), so from Lemma 15.6 and equation (15.14) we have that

$$H_W := \begin{bmatrix} A_{F_2} & X_2^{-1}C_1^*C_1X_2^{-1} \\ -X_2B_1B_1^*X_2 & -A_{F_2}^* \end{bmatrix} \in dom(Ric)$$

and $W = Ric(H_W) \geq 0$. Note that the observability of (C_1, A) implies $X_2 > 0$. Furthermore, the controllability Gramian for the system (15.14) is X_2^{-1} since

$$A_{F_2}X_2^{-1} + X_2^{-1}A_{F_2}^* + B_2B_2^* + X_2^{-1}C_1^*C_1X_2^{-1} = 0.$$

Lemma 15.6 also implies $\rho(WX_2^{-1}) < 1$ or, equivalently, $X_2 > W$. Using the Riccati equation for X_2, one can verify that $T := \begin{bmatrix} -I & X_2^{-1} \\ -X_2 & 0 \end{bmatrix}$ provides a similarity transformation between H_∞ and H_W, i.e., $H_\infty = TH_WT^{-1}$. Then

$$\mathcal{X}_-(H_\infty) = T\mathcal{X}_-(H_W) = T \text{ Im} \begin{bmatrix} I \\ W \end{bmatrix} = \text{Im} \begin{bmatrix} I - X_2^{-1}W \\ X_2 \end{bmatrix},$$

so $H_\infty \in dom(Ric)$ and $X_\infty = X_2(X_2 - W)^{-1}X_2 > 0$.

($\Leftarrow$) If $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) > 0$, $A + B_1B_1^*X_\infty - B_2B_2^*X_\infty$ is stable. Define

$$A_{F_\infty} := A + B_2F_\infty, \quad C_{1F_\infty} := C_1 + D_{12}F_\infty.$$

Then the Riccati equation can be written as

$$A_{F_\infty}^*X_\infty + X_\infty A_{F_\infty} + C_{1F_\infty}^*C_{1F_\infty} + X_\infty B_1B_1^*X_\infty = 0.$$

We conclude from Lyapunov theory that A_{F_∞} is stable since $(A_{F_\infty}, B_1^*X_\infty)$ is detectable and $X_\infty > 0$. Now with the given control law $u = F_\infty x$, the dynamical system becomes

$$\begin{aligned} \dot{x} &= A_{F_\infty}x + B_1w \\ z &= C_{1F_\infty}x. \end{aligned}$$

So by Corollary 13.24, $\|T_{zw}\|_\infty < 1$, i.e., $\sup_{w\in\mathcal{L}_{2+}} \|z\|_2 < 1$ for the given control law. □

Theorem 15.9 will be used in the next chapter to solve the FI $\mathcal{H}_\infty$ control problem.

15.6 Notes and References

Differential game is a well-studied topic, see e.g., Bryson and Ho [1975]. The paper by Mageirou and Ho [1977] is one of the early papers that are relevant to the topics covered in this chapter. The current setup and proof are taken from Doyle, Glover, Khargonekar, and Francis [1989]. The application of game theoretic results to $\mathcal{H}_\infty$ problems can be found in Başar and Bernhard [1991], Khargonekar, Petersen, and Zhou [1990], Limebeer, Anderson, Khargonekar, and Green [1992], and references therein.

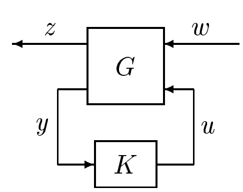
16

$\mathcal{H}_\infty$ Control: Simple Case

In this chapter we consider $\mathcal{H}_\infty$ control theory. Specifically, we formulate the optimal and suboptimal $\mathcal{H}_\infty$ control problems in section 16.1. However, we will focus on the suboptimal case in this book and discuss why we do so. In section 16.2 all suboptimal controllers are characterized for a class of simplified problems while leaving the more general problems to the next chapter. Some preliminary analysis is given in section 16.3 for the output feedback results. The analysis suggested the need for solving the Full Information and Output Estimation problems, which are the topics of sections 16.4-16.7. Section 16.8 discusses the $\mathcal{H}_\infty$ separation theory and presents the proof of the output feedback results. The behavior of the $\mathcal{H}_\infty$ controller as a function of performance level γ is considered in section 16.9. The optimal controllers are also briefly considered in this section. Some other interpretations of the $\mathcal{H}_\infty$ controllers are given in section 16.10. Finally, section 16.11 presents the formulas for an optimal $\mathcal{H}_\infty$ controller.

16.1 Problem Formulation

Consider the system described by the block diagram



where the plant G and controller K are assumed to be real-rational and proper. It will be assumed that state space models of G and K are available and that their realizations are assumed to be stabilizable and detectable. Recall again that a controller is said to be *admissible* if it internally stabilizes the system. Clearly, stability is the most basic requirement for a practical system to work. Hence any sensible controller has to be admissible.

> **Optimal $\mathcal{H}_\infty$ Control:** *find all admissible controllers $K(s)$ such that $\|T_{zw}\|_\infty$ is minimized.*

It should be noted that the optimal $\mathcal{H}_\infty$ controllers as defined above are generally not unique for MIMO systems. Furthermore, finding an optimal $\mathcal{H}_\infty$ controller is often both numerically and theoretically complicated, as shown in Glover and Doyle [1989]. This is certainly in contrast with the standard $\mathcal{H}_2$ theory, in which the optimal controller is unique and can be obtained by solving two Riccati equations without iterations. Knowing the achievable optimal (minimum) $\mathcal{H}_\infty$ norm may be useful theoretically since it sets a limit on what we can achieve. However, in practice it is often not necessary and sometimes even undesirable to design an optimal controller, and it is usually much cheaper to obtain controllers that are very close in the norm sense to the optimal ones, which will be called *suboptimal controllers*. A suboptimal controller may also have other nice properties over optimal ones, e.g., lower bandwidth.

> **Suboptimal $\mathcal{H}_\infty$ Control:** *Given $\gamma > 0$, find all admissible controllers $K(s)$ if there is any such that $\|T_{zw}\|_\infty < \gamma$.*

For the reasons mentioned above, we focus our attention in this book on suboptimal control. When appropriate, we briefly discuss what will happen when γ approaches the optimal value.

16.2 Output Feedback $\mathcal{H}_\infty$ Control

16.2.1 Internal Stability and Input/output Stability

Now suppose K is a stabilizing controller for the system G. Then the internal stability guarantees $T_{zw} = \mathcal{F}_\ell(G,K) \in \mathcal{RH}_\infty$, but the latter does not necessarily imply the internal stability. The following lemma provides the additional (mild) conditions to the equivalence of $T_{zw} = \mathcal{F}_\ell(G,K) \in \mathcal{RH}_\infty$ and internal stability of the closed-loop system. To state the lemma, we shall assume that G and K have the following stabilizable and detectable realizations, respectively:

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right], \qquad K(s) = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & \hat{D} \end{array}\right].$$

Lemma 16.1 *Suppose that the realizations for G and K are both stabilizable and detectable. Then the feedback connection $T_{zw} = \mathcal{F}_\ell(G,K)$ of the realizations for G and K is*

(a) detectable if $\begin{bmatrix} A-\lambda I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ *has full column rank for all* $Re\lambda \geq 0$;

(b) stabilizable if $\begin{bmatrix} A-\lambda I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ *has full row rank for all* $Re\lambda \geq 0$.

Moreover, if (a) and (b) hold, then K is an internally stabilizing controller iff $T_{zw} \in \mathcal{RH}_\infty$.

Proof. The state-space equations for the closed-loop are:

$$\begin{aligned} \mathcal{F}_\ell(G,K) &= \left[\begin{array}{cc|c} A + B_2\hat{D}L_1C_2 & B_2L_2\hat{C} & B_1 + B_2\hat{D}L_1D_{21} \\ \hat{B}L_1C_2 & \hat{A}+\hat{B}L_1D_{22}\hat{C} & \hat{B}L_1D_{21} \\ \hline C_1 + D_{12}L_2\hat{D}C_2 & D_{12}L_2\hat{C} & D_{11} + D_{12}\hat{D}L_1D_{21} \end{array}\right] \\ &=: \left[\begin{array}{c|c} A_c & B_c \\ \hline C_c & D_c \end{array}\right] \end{aligned}$$

where $L_1 := (I - D_{22}\hat{D})^{-1}$, $L_2 := (I - \hat{D}D_{22})^{-1}$.

Suppose $\mathcal{F}_\ell(G,K)$ has undetectable state $(x', y')'$ and mode $\mathrm{Re}\lambda \geq 0$; then the PBH test gives

$$\begin{bmatrix} A_c - \lambda I \\ C_c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = 0.$$

This can be simplified as

$$\begin{bmatrix} A-\lambda I & B_2 \\ C_1 & D_{12} \end{bmatrix} \begin{bmatrix} x \\ \hat{D}L_1C_2x + L_2\hat{C}y \end{bmatrix} = 0$$

and

$$\hat{B}L_1(C_2x + D_{22}\hat{C}y) + \hat{A}y - \lambda y = 0.$$

Now if

$$\begin{bmatrix} A-\lambda I & B_2 \\ C_1 & D_{12} \end{bmatrix}$$

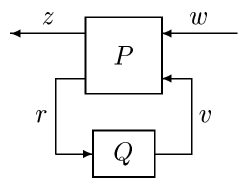
has full column rank, then $x = 0$ and $\hat{C}y = 0$. This implies $\hat{A}y = \lambda y$. Since $(\hat{C}, \hat{A})$ is detectable, we get $y = 0$, which is a contradiction. Hence part (a) is proven, and part (b) is a dual result. □

These relations will be used extensively below to simplify our development and to enable us to focus on input/output stability only.

16.2.2 Contraction and Stability

One of the keys to the entire development of $\mathcal{H}_\infty$ theory is the fact that the contraction and internal stability is preserved under an inner linear fractional transformation.

Theorem 16.2 *Consider the following feedback system:*



$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} \in \mathcal{RH}_\infty$$

Suppose that $P^\sim P = I$, $P_{21}^{-1} \in \mathcal{RH}_\infty$ and that Q is a proper rational matrix. Then the following are equivalent:

(a) The system is internally stable, well-posed, and $\|T_{zw}\|_\infty < 1$.

(b) $Q \in \mathcal{RH}_\infty$ and $\|Q\|_\infty < 1$.

Proof. (b) $\Rightarrow$ (a). Note that since $P, Q \in \mathcal{RH}_\infty$, the system internal stability is guaranteed if $(I - P_{22}Q)^{-1} \in \mathcal{RH}_\infty$. Therefore, internal stability and well-posedness follow easily from $\|P_{22}\|_\infty \le 1$, $\|Q\|_\infty < 1$, and a small gain argument. (Note that $\|P_{22}\|_\infty \le 1$ follows from the fact that P_{22} is a compression of P.)

To show that $\|T_{zw}\|_\infty < 1$, consider the closed-loop system at any frequency $s = j\omega$ with the signals fixed as complex constant vectors. Let $\|Q\|_\infty =: \epsilon < 1$ and note that $T_{wr} = P_{21}^{-1}(I - P_{22}Q) \in \mathcal{RH}_\infty$. Also let $\kappa := \|T_{wr}\|_\infty$. Then $\|w\| \le \kappa\|r\|$, and P inner implies that $\|z\|^2 + \|r\|^2 = \|w\|^2 + \|v\|^2$. Therefore,

$$\|z\|^2 \le \|w\|^2 + (\epsilon^2 - 1)\|r\|^2 \le [1 - (1 - \epsilon^2)\kappa^{-2}]\|w\|^2$$

which implies $\|T_{zw}\|_\infty < 1$.

(a) $\Rightarrow$ (b). To show that $\|Q\|_\infty < 1$, suppose there exist a (real or infinite) frequency ω and a constant nonzero vector r such that at $s = j\omega$, $\|Qr\| \ge \|r\|$. Then setting $w = P_{21}^{-1}(I - P_{22}Q)r$, $v = Qr$ gives $v = T_{vw}w$. But as above, P inner implies that $\|z\|^2 + \|r\|^2 = \|w\|^2 + \|v\|^2$ and, hence, that $\|z\|^2 \ge \|w\|^2$, which is impossible since $\|T_{zw}\|_\infty < 1$. It follows that $\sigma_{max}(Q(j\omega)) < 1$ for all ω, i.e., $\|Q\|_\infty < 1$ since Q is rational.

To show $Q \in \mathcal{RH}_\infty$, let $Q = NM^{-1}$ with $N, M \in \mathcal{RH}_\infty$ be a right coprime factorization, i.e., there exist $X, Y \in \mathcal{RH}_\infty$ such that $XN + YM = I$. We shall show that $M^{-1} \in \mathcal{RH}_\infty$. By internal stability we have

$$Q(I - P_{22}Q)^{-1} = N(M - P_{22}N)^{-1} \in \mathcal{RH}_\infty$$

and

$$(I - P_{22}Q)^{-1} = M(M - P_{22}N)^{-1} \in \mathcal{RH}_\infty.$$

Thus

$$XQ(I - P_{22}Q)^{-1} + Y(I - P_{22}Q)^{-1} = (M - P_{22}N)^{-1} \in \mathcal{RH}_\infty.$$

This implies that the winding number of $\det(M - P_{22}N)$, as s traverses the Nyquist contour, equals zero. Now note the fact that, for all $s = j\omega$, $\det M^{-1} \neq 0$, $\det(I - \alpha P_{22}Q) \neq 0$ for all α in [0,1] (this uses the fact that $\|P_{22}\|_\infty \leq 1$ and $\|Q\|_\infty < 1$). Also, $\det(I - \alpha P_{22}Q) = \det(M - \alpha P_{22}N) \det M^{-1}$, and we have $\det(M - \alpha P_{22}N) \neq 0$ for all α in [0,1] and all $s = j\omega$. We conclude that the winding number of $\det M$ also equals zero. Therefore, $Q \in \mathcal{RH}_\infty$, and the proof is complete. □

16.2.3 Simplifying Assumptions

In this chapter, we discuss a simplified version of $\mathcal{H}_\infty$ theory. The general case will be considered in the next chapter. The main reason for doing so is that the general case has its unique features but is much more involved algebraically. Involved algebra may distract attention from the essential ideas of the theory and therefore lose insight into the problem. Nevertheless, the problem considered below contains the essential features of the $\mathcal{H}_\infty$ theory.

The realization of the transfer matrix G is taken to be of the form

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right].$$

The following assumptions are made:

(i) (A, B_1) is stabilizable and (C_1, A) is detectable;

(ii) (A, B_2) is stabilizable and (C_2, A) is detectable;

(iii) $D_{12}^*\begin{bmatrix} C_1 & D_{12} \end{bmatrix} = \begin{bmatrix} 0 & I \end{bmatrix}$;

(iv) $\begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} D_{21}^* = \begin{bmatrix} 0 \\ I \end{bmatrix}$.

Assumption (i) is made for a technical reason: together with (ii) it guarantees that the two Hamiltonian matrices (H_2 and J_2 in Chapter 14) in the $\mathcal{H}_2$ problem belong to *dom*(*Ric*). This assumption simplifies the theorem statements and proofs, but if it is relaxed, the theorems and proofs can be modified so that the given formulae are still correct, as will be seen in the next chapter. An important simplification that is a consequence of the assumption (i) is that internal stability is essentially equivalent to input-output stability ($T_{zw} \in \mathcal{RH}_\infty$). This equivalence enables us to focus only on input/output stability and is captured in Corollary 16.3 below. Of course, assumption (ii) is necessary and sufficient for G to be internally stabilizable, but is not needed to prove the equivalence of internal stability and $T_{zw} \in \mathcal{RH}_\infty$. (Readers should be clear that this does not mean that the realization for G need not be stabilizable and detectable. In point of fact, the internal stability and input-output stability can never be equivalent if either G or K has unstabilizable or undetectable modes.)

Corollary 16.3 *Suppose that assumptions (i), (iii), and (iv) hold. Then a controller K is admissible iff $T_{zw} \in \mathcal{RH}_\infty$.*

Proof. The realization for plant G is stabilizable and detectable by assumption (i). We only need to verify that the rank conditions of the two matrices in Lemma 16.1 are satisfied. Now suppose assumptions (i) and (iii) are satisfied and let $D_\perp$ be such that $\begin{bmatrix} D_{12} & D_\perp \end{bmatrix}$ is a unitary matrix. Then

$$\begin{aligned} \operatorname{rank}\begin{bmatrix} A-\lambda I & B_2 \\ C_1 & D_{12} \end{bmatrix} &= \operatorname{rank}\begin{bmatrix} I & 0 \\ 0 & \begin{bmatrix} D_{12}^* \\ D_\perp^* \end{bmatrix} \end{bmatrix}\begin{bmatrix} A-\lambda I & B_2 \\ C_1 & D_{12} \end{bmatrix} \\ &= \operatorname{rank}\begin{bmatrix} A-\lambda I & B_2 \\ \begin{bmatrix} 0 \\ D_\perp^* C_1 \end{bmatrix} & \begin{bmatrix} I \\ 0 \end{bmatrix} \end{bmatrix}. \end{aligned}$$

So

$$\begin{bmatrix} A-\lambda I & B_2 \\ C_1 & D_{12} \end{bmatrix}$$

has full column rank for all $\mathrm{Re}\lambda \geq 0$ iff

$$\begin{bmatrix} A-\lambda I \\ D_\perp^* C_1 \end{bmatrix}$$

has full column rank. However, the last matrix has full rank for all $\mathrm{Re}\lambda \geq 0$ iff $(D_\perp^* C_1, A)$ is detectable. Since $D_\perp(D_\perp^* C_1) = (I - D_{12}D_{12}^*)C_1 = C_1$, $(D_\perp^* C_1, A)$ is detectable iff (C_1, A) is detectable. The rank condition for the other matrix follows by duality. □

Two additional assumptions that are implicit in the assumed realization for $G(s)$ are that $D_{11} = 0$ and $D_{22} = 0$. As we have mentioned many times, $D_{22} \neq 0$ does not pose any problem since it is easy to form an equivalent problem with $D_{22} = 0$ by a linear fractional transformation on the controller $K(s)$. However, relaxing the assumption $D_{11} = 0$ complicates the formulae substantially, as will be seen in the next chapter.

16.2.4 Suboptimal $\mathcal{H}_\infty$ Controllers

In this subsection, we present the necessary and sufficient conditions for the existence of an admissible controller $K(s)$ such that $\|T_{zw}\|_\infty < \gamma$ for a given γ, and, furthermore, if the necessary and sufficient conditions are satisfied, we characterize all admissible controllers that satisfy the norm condition. The proofs of these results will be given in the later sections. Let $\gamma_{opt} := \min \{\|T_{zw}\|_\infty : \ K(s) \text{ admissible}\}$, i.e., the optimal level. Then, clearly, γ must be greater than γ_{opt} for the existence of suboptimal $\mathcal{H}_\infty$ controllers. In Section 16.9 we will briefly discuss how to find an admissible K to minimize $\|T_{zw}\|_\infty$. Optimal $\mathcal{H}_\infty$ controllers are more difficult to characterize than suboptimal ones, and this is one major difference between the $\mathcal{H}_\infty$ and $\mathcal{H}_2$ results. Recall that similar differences arose in the norm computation problem as well.

The $\mathcal{H}_\infty$ solution involves the following two Hamiltonian matrices:

$$H_\infty := \begin{bmatrix} A & \gamma^{-2}B_1B_1^* - B_2B_2^* \\ -C_1^*C_1 & -A^* \end{bmatrix}, \qquad J_\infty := \begin{bmatrix} A^* & \gamma^{-2}C_1^*C_1 - C_2^*C_2 \\ -B_1B_1^* & -A \end{bmatrix}.$$

The important difference here from the $\mathcal{H}_2$ problem is that the (1,2)-blocks are not sign definite, so we cannot use Theorem 13.7 in Chapter 13 to guarantee that $H_\infty \in dom(Ric)$ or $Ric(H_\infty) \geq 0$. Indeed, these conditions are intimately related to the existence of $\mathcal{H}_\infty$ suboptimal controllers. Note that the (1,2)-blocks are a suggestive combination of expressions from the $\mathcal{H}_\infty$ norm characterization in Chapter 4 (or bounded real ARE in Chapter 13) and from the $\mathcal{H}_2$ synthesis of Chapter 14. It is also clear that if γ approaches infinity, then these two Hamiltonian matrices become the corresponding $\mathcal{H}_2$ control Hamiltonian matrices. The reasons for the form of these expressions should become clear through the discussions and proofs for the following theorem.

Theorem 16.4 *There exists an admissible controller such that $\|T_{zw}\|_\infty < \gamma$ iff the following three conditions hold:*

(i) $H_\infty \in$ *dom(Ric) and* $X_\infty := Ric(H_\infty) \geq 0$;

(ii) $J_\infty \in$ *dom(Ric) and* $Y_\infty := Ric(J_\infty) \geq 0$;

(iii) $\rho(X_\infty Y_\infty) < \gamma^2$.

Moreover, when these conditions hold, one such controller is

$$K_{sub}(s) := \left[\begin{array}{c|c} \hat{A}_\infty & -Z_\infty L_\infty \\ \hline F_\infty & 0 \end{array}\right]$$

where

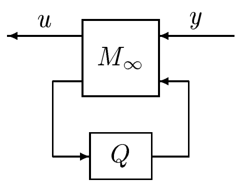
$$\hat{A}_\infty := A + \gamma^{-2}B_1B_1^*X_\infty + B_2F_\infty + Z_\infty L_\infty C_2$$

$$F_\infty := -B_2^*X_\infty, \qquad L_\infty := -Y_\infty C_2^*, \qquad Z_\infty := (I - \gamma^{-2}Y_\infty X_\infty)^{-1}.$$

The $\mathcal{H}_\infty$ controller displayed in Theorem 16.4, which is often called the *central controller* or *minimum entropy controller*, has certain obvious similarities to the $\mathcal{H}_2$ controller as well as some important differences. Although not as apparent as in the $\mathcal{H}_2$ case, the $\mathcal{H}_\infty$ controller also has an interesting separation structure. Furthermore, each of the conditions in the theorem can be given a system-theoretic interpretation in terms of this separation. These interpretations, given in Section 16.8, require the filtering and full information (i.e., state feedback) results in sections 16.7 and 16.4. The proof of Theorem 16.4 is constructed out of these results as well. The term central controller will be obvious from the parameterization of all suboptimal controllers given below, while the meaning of minimum entropy will be discussed in Section 16.10.1.

The following theorem parameterizes the controllers that achieve a suboptimal $\mathcal{H}_\infty$ norm less than γ.

Theorem 16.5 *If conditions (i) to (iii) in Theorem 16.4 are satisfied, the set of all admissible controllers such that $\|T_{zw}\|_\infty < \gamma$ equals the set of all transfer matrices from y to u in*



$$M_\infty(s) = \left[\begin{array}{c|cc} \hat{A}_\infty & -Z_\infty L_\infty & Z_\infty B_2 \\ \hline F_\infty & 0 & I \\ -C_2 & I & 0 \end{array}\right]$$

where $Q \in \mathcal{RH}_\infty$, $\|Q\|_\infty < \gamma$.

As in the $\mathcal{H}_2$ case, the suboptimal controllers are parameterized by a fixed linear-fractional transformation with a free parameter Q. With $Q = 0$ (at the "center" of the set $\|Q\|_\infty < \gamma$), we recover the central controller $K_{sub}(s)$.

16.3 Motivation for Special Problems

Although the proof for output feedback results will be given later, we shall now try to give some ideas for approaching the problem. Specifically, we try to motivate the study of the OE (and hence other special problems) and show how this problem arises naturally in proving the output feedback results. The key is to use the fact that contraction and internal stability are preserved under an inner linear fractional transformation, which is Theorem 16.2. Assuming that X_∞ exists, we will show that the general output feedback problem boils down to an output estimation problem which can be solved easily if a state feedback or full information control problem can be solved.

Suppose $X_\infty := Ric(H_\infty)$ exists. Then X_∞ satisfies the following Riccati equation:

$$A^* X_\infty + X_\infty A + C_1^* C_1 + \gamma^{-2} X_\infty B_1 B_1^* X_\infty - X_\infty B_2 B_2^* X_\infty = 0. \tag{16.1}$$

Let x denote the state of G with respect to a given input w, and then we can differentiate $x(t)^* X_\infty x(t)$:

$$\frac{d}{dt}(x^* X_\infty x) = \dot{x}^* X_\infty x + x^* X_\infty \dot{x}$$

$$= x^*(A^*X_\infty + X_\infty A)x + 2\langle w, B_1^*X_\infty x\rangle + 2\langle u, B_2^*X_\infty x\rangle.$$

Using the Riccati equation for X_∞ to substitute in for $A^*X_\infty + X_\infty A$ gives

$$\frac{d}{dt}(x^*X_\infty x) = -\|C_1x\|^2 - \gamma^{-2}\|B_1^*X_\infty x\|^2 + \|B_2^*X_\infty x\|^2 + 2\langle w, B_1^*X_\infty x\rangle + 2\langle u, B_2^*X_\infty x\rangle.$$

Finally, completion of the squares along with orthogonality of C_1x and $D_{12}u$ gives the key equation

$$\frac{d}{dt}(x^*X_\infty x) = -\|z\|^2 + \gamma^2\|w\|^2 - \gamma^2\|w - \gamma^{-2}B_1^*X_\infty x\|^2 + \|u + B_2^*X_\infty x\|^2. \quad (16.2)$$

Assume $x(0) = x(\infty) = 0$, $w \in \mathcal{L}_{2+}$, and integrate (16.2) from $t = 0$ to $t = \infty$:

$$\|z\|_2^2 - \gamma^2\|w\|_2^2 = \|u + B_2^*X_\infty x\|_2^2 - \gamma^2\|w - \gamma^{-2}B_1^*X_\infty x\|_2^2 = \|v\|_2^2 - \gamma^2\|r\|_2^2 \quad (16.3)$$

where

$$v := u + B_2^*X_\infty x, \qquad r := w - \gamma^{-2}B_1^*X_\infty x. \quad (16.4)$$

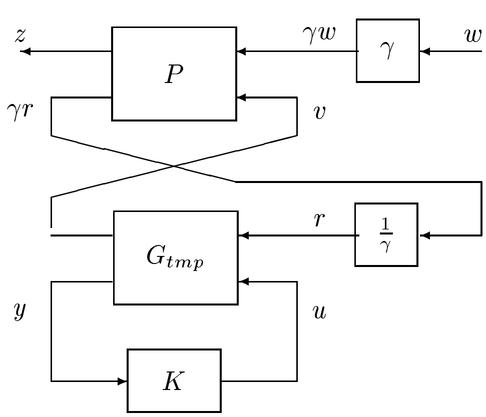
With these new defined variables, the closed-loop system can be expressed as two interconnected subsystems below:

$$\begin{bmatrix} \dot{x} \\ z \\ \gamma r \end{bmatrix} = \begin{bmatrix} A_{F_\infty} & \gamma^{-1}B_1 & B_2 \\ C_{1F_\infty} & 0 & D_{12} \\ -\gamma^{-1}B_1^*X_\infty & I & 0 \end{bmatrix} \begin{bmatrix} x \\ \gamma w \\ v \end{bmatrix} \qquad \begin{array}{ll} A_{F_\infty} & := A + B_2F_\infty \\ C_{1F_\infty} & := C_1 + D_{12}F_\infty \end{array}$$

and

$$\begin{bmatrix} \dot{x} \\ v \\ y \end{bmatrix} = \begin{bmatrix} A_{tmp} & B_1 & B_2 \\ -F_\infty & 0 & I \\ C_2 & D_{21} & 0 \end{bmatrix} \begin{bmatrix} x \\ r \\ u \end{bmatrix} \qquad A_{tmp} := A + \gamma^{-2}B_1B_1^*X_\infty$$

where F_∞ is defined as in Section 16.2.4. This is shown in the following diagram:



where

$$P := \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A_{F_\infty} & \gamma^{-1}B_1 & B_2 \\ \hline C_{1F_\infty} & 0 & D_{12} \\ -\gamma^{-1}B_1^* X_\infty & I & 0 \end{array}\right] \tag{16.5}$$

and

$$G_{tmp} = \left[\begin{array}{c|cc} A_{tmp} & B_1 & B_2 \\ \hline -F_\infty & 0 & I \\ C_2 & D_{21} & 0 \end{array}\right]. \tag{16.6}$$

The equality (16.3) motivates the change of variables to r and v as in (16.4), and these variables provide the connection between T_{zw} and T_{vr}. Note that $T_{z(\gamma w)} = \gamma^{-1}T_{zw}$ and $T_{v(\gamma r)} = \gamma^{-1}T_{vr}$. It is immediate from equality (16.3) that $\|T_{zw}\|_\infty \leq \gamma$ iff $\|T_{vr}\|_\infty \leq \gamma$. While this is the basic idea behind the proof of Lemma 16.8 below, the details needed for strict inequality and internal stability require a bit more work.

Note that $w_{worst} := \gamma^{-2}B_1^*X_\infty x$ is the worst disturbance input in the sense that it maximizes the quantity $\|z\|_2^2 - \gamma^2\|w\|_2^2$ in (16.3) for the minimizing value of $u = -B_2^*X_\infty x$; that is, the u making $v = 0$ and the w making $r = 0$ are values satisfying a saddle point condition. (See Section 17.8 for more interpretations.) It is also interesting to note that w_{worst} is the optimal strategy for w in the corresponding LQ game problem (see the differential game problem in the last chapter). Equation (16.3) also suggests that $u = -B_2^*X_\infty x$ is a suboptimal control for a full information (FI) problem if the state x is available. This will be shown later. In terms of the OE problem for G_{tmp}, the output being estimated is the optimal FI control input $F_\infty x$ and the new disturbance r is offset by the "worst case" FI disturbance input w_{worst}.

Notice the structure of G_{tmp}: it is an OE problem. We will show below that the output feedback can indeed be transformed into the OE problem. To show this we first need to prove some preliminary facts.

Lemma 16.6 *Suppose $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty)$. Then $A_{F_\infty} = A + B_2F_\infty$ is stable iff $X_\infty \geq 0$.*

Proof. Re-arrange the Riccati equation for X_∞ and use the definition of F_∞ and C_{1F_∞} to get

$$A_{F_\infty}^* X_\infty + X_\infty A_{F_\infty} + \begin{bmatrix} C_{1F_\infty} \\ -\gamma^{-1}B_1^*X_\infty \end{bmatrix}^* \begin{bmatrix} C_{1F_\infty} \\ -\gamma^{-1}B_1^*X_\infty \end{bmatrix} = 0. \tag{16.7}$$

Since $H_\infty \in dom(Ric)$, $(A_{F_\infty} + \gamma^{-2}B_1B_1^*X_\infty)$ is stable and hence $(B_1^*X_\infty, A_{F_\infty})$ is detectable. Then from standard results on Lyapunov equations (see Lemma 3.19), A_{F_∞} is stable iff $X_\infty \geq 0$. □

Equation (16.3) can be written as

$$\|z\|_2^2 + \|\gamma r\|_2^2 = \|\gamma w\|_2^2 + \|v\|_2^2 .$$

This suggests that P might be inner when $X_\infty \geq 0$, which is verified by the following lemma.

Lemma 16.7 *If $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) \geq 0$, then P in (16.5) is in $\mathcal{RH}_\infty$ and inner, and $P_{21}^{-1} \in \mathcal{RH}_\infty$.*

Proof. By Lemma 16.6, A_{F_∞} is stable. So $P \in \mathcal{RH}_\infty$. That P is inner ($P^\sim P = I$) follows from Lemma 13.29 upon noting that the observability Gramian of P is X_∞ (see (16.7)) and

$$\begin{bmatrix} 0 & I \\ D_{12}^* & 0 \end{bmatrix} \begin{bmatrix} C_{1F_\infty} \\ -\gamma^{-1}B_1^*X_\infty \end{bmatrix} + \begin{bmatrix} \gamma^{-1}B_1^* \\ B_2^* \end{bmatrix} X_\infty = 0.$$

Finally, the state matrix for P_{21}^{-1} is $(A_{F_\infty} + \gamma^{-2}B_1B_1^*X_\infty)$, which is stable by definition. Thus, $P_{21}^{-1} \in \mathcal{RH}_\infty$. □

The following lemma connects these two systems T_{zw} and T_{vr}, which is the central part of the separation argument in Section 16.8.2. Recall that internal and input-output stability are equivalent for admissibility of K in the output feedback problem by Corollary 16.3 .

Lemma 16.8 *Assume $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) \geq 0$. Then K is admissible for G and $\|T_{zw}\|_\infty < \gamma$ iff K is admissible for G_{tmp} and $\|T_{vr}\|_\infty < \gamma$.*

Proof. We may assume without loss of generality that the realization of K is stabilizable and detectable. Recall from Corollary 16.3 that internal stability for T_{zw} is equivalent to $T_{zw} \in \mathcal{RH}_\infty$. Similarly since

$$\begin{bmatrix} A_{tmp} - \lambda I & B_1 \\ C_2 & D_{21} \end{bmatrix} = \begin{bmatrix} A - \lambda I & B_1 \\ C_2 & D_{21} \end{bmatrix} \begin{bmatrix} I & 0 \\ \gamma^{-2}B_1^*X_\infty & I \end{bmatrix}$$

has full row rank for all $Re(\lambda) \geq 0$ and since

$$\det \begin{bmatrix} A_{tmp} - \lambda I & B_2 \\ -F_\infty & I \end{bmatrix} = \det(A_{tmp} + B_2F_\infty - \lambda I) \neq 0$$

for all $Re(\lambda) \geq 0$ by the stability of $A_{tmp} + B_2F_\infty$, we have that

$$\begin{bmatrix} A_{tmp} - \lambda I & B_2 \\ -F_\infty & I \end{bmatrix}$$

has full column rank. Hence by Lemma 16.1 the internal stability of T_{vr}, i.e., the internal stability of the subsystem consisting of G_{tmp} and controller K, is also equivalent to $T_{vr} \in \mathcal{RH}_\infty$. Thus internal stability is equivalent to input-output stability for both

G and G_{tmp}. This shows that K is an admissible controller for G if and only if it is admissible for G_{tmp}. Now it follows from Theorem 16.2 and Lemma 16.7 along with the above block diagram that $\left\|T_{z(\gamma w)}\right\|_\infty < 1$ iff $\left\|T_{v(\gamma r)}\right\|_\infty < 1$ or, equivalently, $\|T_{zw}\|_\infty < \gamma$ iff $\|T_{vr}\|_\infty < \gamma$. □

From the previous analysis, it is clear that to solve the output feedback problem we need to show

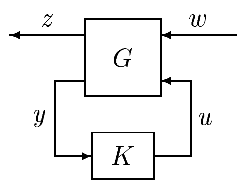
(a) $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) \geq 0$;

(b) $\|T_{vr}\|_\infty < \gamma$.

To show (a), we need to solve a FI problem. The problem (b) is an OE problem which can be solved by using the relationship between FC and OE problems in Chapter 12, while the FC problem can be solved by using the FI solution through duality. So in the sections to follow, we will focus on these special problems.

16.4 Full Information Control

Our system diagram in this section is standard as before



with

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}\right].$$

The $\mathcal{H}_\infty$ problem corresponding to this setup again is not, strictly speaking, a special case of the output feedback problem because it does not satisfy all of the assumptions. In particular, it should be noted that for the FI (and FC in the next section) problem, internal stability is not equivalent to $T_{zw} \in \mathcal{RH}_\infty$ since

$$\begin{bmatrix} A - \lambda I & B_1 \\ \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} \end{bmatrix}$$

can never have full row rank, although this presents no difficulties in solving this problem. We simply must remember that in the FI case, K admissible means internally stabilizing, not just $T_{zw} \in \mathcal{RH}_\infty$.

We have seen that in the $\mathcal{H}_2$ FI case, the optimal controller uses just the state x even though the controller is provided with full information. We will show below that, in the $\mathcal{H}_\infty$ case, a suboptimal controller exists which also uses just x. This case could have been restricted to state feedback, which is more traditional, but we believe that, once one gets outside the pure $\mathcal{H}_2$ setting, the full information problem is more fundamental and more natural than the state feedback problem.

One setting in which the full information case is more natural occurs when the parameterization of all suboptimal controllers is considered. It is also appropriate when studying the general case when $D_{11} \neq 0$ in the next chapter or when $\mathcal{H}_\infty$ optimal (not just suboptimal) controllers are desired. Even though the optimal problem is not studied in detail in this book, we want the methods to extend to the optimal case in a natural and straightforward way.

The assumptions relevant to the FI problem which are inherited from the output feedback problem are

(i) (C_1, A) is detectable;

(ii) (A, B_2) is stabilizable;

(iii) $D_{12}^* \begin{bmatrix} C_1 & D_{12} \end{bmatrix} = \begin{bmatrix} 0 & I \end{bmatrix}$.

Assumptions (iv) and the second part of (ii) for the general output feedback case have been effectively strengthened because of the assumed structure for C_2 and D_{21}.

Theorem 16.9 *There exists an admissible controller $K(s)$ for the FI problem such that $\|T_{zw}\|_\infty < \gamma$ if and only if $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) \geq 0$. Furthermore, if these conditions are satisfied, then the equivalence class of all admissible controllers satisfying $\|T_{zw}\|_\infty < \gamma$ can be parameterized as*

$$K(s) \cong \begin{bmatrix} F_\infty - \gamma^{-2} Q(s) B_1^* X_\infty & Q(s) \end{bmatrix} \tag{16.8}$$

where $Q \in \mathcal{RH}_\infty$, $\|Q\|_\infty < \gamma$.

It is easy to see by comparing the $\mathcal{H}_\infty$ solution with the corresponding $\mathcal{H}_2$ solution that a fundamental difference between $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers is that the $\mathcal{H}_\infty$ controller depends on the disturbance through B_1 whereas the $\mathcal{H}_2$ controller does not. This difference is essentially captured by the necessary and sufficient conditions for the existence of a controller given in Theorem 16.9. Note that these conditions are the same as condition (i) in Theorem 16.4.

The two individual conditions in Theorem 16.9 may each be given their own interpretations. The condition that $H_\infty \in dom(Ric)$ implies that $X_\infty := Ric(H_\infty)$ exists and $K(s) = \begin{bmatrix} F_\infty & 0 \end{bmatrix}$ gives T_{zw} as

$$T_{zw} = \left[\begin{array}{c|c} A_{F_\infty} & B_1 \\ \hline C_{1F_\infty} & 0 \end{array}\right] \qquad \begin{aligned} A_{F_\infty} &= A + B_2 F_\infty \\ C_{1F_\infty} &= C_1 + D_{12} F_\infty . \end{aligned} \tag{16.9}$$

Furthermore, since $T_{zw} = \gamma P_{11}$ and $P_{11}^\sim P_{11} = I - P_{21}^\sim P_{21}$ by $P^\sim P = I$, we have $\|P_{11}\|_\infty < 1$ and $\|T_{zw}\|_\infty = \gamma \|P_{11}\|_\infty < \gamma$. The further condition that $X_\infty \geq 0$ is equivalent, by Lemma 16.6, to this K stabilizing T_{zw}.

Remark 16.1 Given that $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) \geq 0$, there is an intuitive way to see why all the equivalence classes of controllers can be parameterized in the form of (16.8). Recall the following equality from equation (16.3):

$$\|z\|_2^2 - \gamma^2 \|w\|_2^2 = \|u + B_2^* X_\infty x\|_2^2 - \gamma^2 \|w - \gamma^{-2} B_1^* X_\infty x\|_2^2 = \|v\|_2^2 - \gamma^2 \|r\|_2^2 .$$

$\|T_{zw}\|_\infty < \gamma$ means that $\|z\|_2^2 - \gamma^2 \|w\|_2^2 < 0$ for $w \neq 0$. This implies that $r \neq 0$ and $\|v\|_2^2 - \gamma^2 \|r\|_2^2 < 0$. Now all v satisfying this inequality can be written as $v = Qr$ for $Q \in \mathcal{RH}_\infty$ and $\|Q\|_\infty < \gamma$ or as $u + B_2^* X_\infty x = Q(w - \gamma^{-2} B_1^* X_\infty x)$. This gives equation (16.8). ♡

We shall now prove the theorem.

Proof. ($\Rightarrow$) For simplicity, in the proof to follow we assume that the system is normalized such that $\gamma = 1$. Further, we will show that we can, without loss of generality, strengthen the assumption on (C_1, A) from detectable to observable. Suppose there exists an admissible controller $\hat{K} = \left[\begin{array}{c|cc} \hat{A} & \hat{B}_1 & \hat{B}_2 \\ \hline \hat{C} & \hat{D}_1 & \hat{D}_2 \end{array}\right]$ such that $\|T_{zw}\|_\infty < 1$. If (C_1, A) is detectable but not observable, then change coordinates for the state of G to $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ with x_2 unobservable, (C_{11}, A_{11}) observable, and A_{22} stable, giving the following closed-loop state equations:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{\hat{x}} \\ z \\ u \end{bmatrix} = \begin{bmatrix} A_{11} & 0 & 0 & B_{11} & B_{21} \\ A_{21} & A_{22} & 0 & B_{12} & B_{22} \\ \hat{B}_{11} & \hat{B}_{12} & \hat{A} & \hat{B}_2 & 0 \\ C_{11} & 0 & 0 & 0 & D_{12} \\ \hat{D}_{11} & \hat{D}_{12} & \hat{C} & \hat{D}_2 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \hat{x} \\ w \\ u \end{bmatrix} .$$

If we take a new plant G_{obs} with state x_1 and output z and group the rest of the equations as a new controller K_{obs} with the state made up of x_2 and $\hat{x}$, then

$$G_{obs}(s) = \left[\begin{array}{c|cc} A_{11} & B_{11} & B_{21} \\ \hline C_{11} & 0 & D_{12} \\ \begin{bmatrix} I \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} 0 \\ 0 \end{bmatrix} \end{array}\right]$$

still satisfies the assumptions of the FI problem and is stabilized by K_{obs} with the closed-loop $\mathcal{H}_\infty$ norm $\|T_{zw}\|_\infty < 1$ where

$$K_{obs} = \left[\begin{array}{cc|cc} A_{22} + B_{22}\hat{D}_{12} & B_{22}\hat{C} & A_{21} + B_{22}\hat{D}_{11} & B_{12} + B_{22}\hat{D}_2 \\ \hat{B}_{12} & \hat{A} & \hat{B}_{11} & \hat{B}_2 \\ \hline \hat{D}_{12} & \hat{C} & \hat{D}_{11} & \hat{D}_2 \end{array}\right].$$

If we now show that there exists $\hat{X}_\infty > 0$ solving the H_∞ Riccati equation for G_{obs}, i.e.,

$$\hat{X}_\infty = Ric \begin{bmatrix} A_{11} & B_{11}B_{11}^* - B_{21}B_{21}^* \\ -C_{11}^*C_{11} & -A_{11}^* \end{bmatrix},$$

then

$$Ric(H_\infty) = X_\infty = \begin{bmatrix} \hat{X}_\infty & 0 \\ 0 & 0 \end{bmatrix} \geq 0$$

exists for G. We can therefore assume without loss of generality that (C_1, A) is observable.

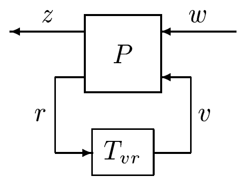
We shall suppose that there exists an admissible controller such that $\|T_{zw}\|_\infty < 1$. But note that the existence of an admissible controller such that $\|T_{zw}\|_\infty < 1$ is equivalent to that the admissible controller makes $\sup_{w \in \mathcal{BL}_{2+}} \|z\|_2 < 1$; hence, it is necessary that

$$\sup_{w \in \mathcal{BL}_{2+}} \min_{u \in \mathcal{L}_{2+}} \|z\|_2 < 1$$

since the latter is always no greater than the former by the fact that the set of signals u generated from admissible controllers is a subset of $\mathcal{L}_{2+}$. But from Theorem 15.9, the latter is true if and only if $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) > 0$. Hence the necessity is proven.

($\Leftarrow$) Suppose $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) \geq 0$ and suppose $K(s)$ is an admissible controller such that $\|T_{zw}\|_\infty < 1$. Again change variables to $v := u - F_\infty x$

and $r := w - B_1^* X_\infty x$, so that the closed-loop system is as shown below:



$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{21} & P_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A_{F_\infty} & B_1 & B_2 \\ \hline C_{1F_\infty} & 0 & D_{12} \\ -B_1^* X_\infty & I & 0 \end{array}\right]$$

By Lemma 16.7, P is inner and $P_{21}^{-1} \in \mathcal{RH}_\infty$. By Theorem 16.2 the system is internally stable and $\|T_{zw}\|_\infty < 1$ iff $T_{vr} \in \mathcal{RH}_\infty$ and $\|T_{vr}\|_\infty < 1$. Now denote $Q := T_{vr}$, then $v = Qr$ and

$$\begin{aligned} v &= u - F_\infty x = \left(K - \begin{bmatrix} F_\infty & 0 \end{bmatrix}\right) \begin{bmatrix} x \\ w \end{bmatrix} \\ r &= w - B_1^* X_\infty x = \begin{bmatrix} -B_1^* X_\infty & I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix}. \end{aligned}$$

Hence we have

$$\left(K - \begin{bmatrix} F_\infty & 0 \end{bmatrix}\right) \begin{bmatrix} x \\ w \end{bmatrix} = Q \begin{bmatrix} -B_1^* X_\infty & I \end{bmatrix} \begin{bmatrix} x \\ w \end{bmatrix}$$

or

$$K - \begin{bmatrix} F_\infty & 0 \end{bmatrix} \cong Q \begin{bmatrix} -B_1^* X_\infty & I \end{bmatrix}.$$

Thus

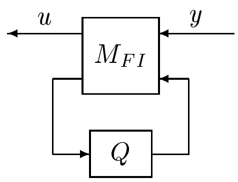
$$K(s) \cong \begin{bmatrix} F_\infty - Q(s) B_1^* X_\infty & Q(s) \end{bmatrix}, \quad Q \in \mathcal{RH}_\infty, \|Q\|_\infty < 1.$$

Hence all suboptimal controllers are in the stated equivalence class. □

Remark 16.2 It should be emphasized that the set of controllers given above does not parameterize all controllers although it is sufficient for the purpose of deriving the output feedback results, and that is why "equivalence class" is used. It is clear that there is a suboptimal controller $K_1 = \begin{bmatrix} F_1 & 0 \end{bmatrix}$ with $F_1 \neq F_\infty$; however, there is no choice of Q such that K_1 belongs to the set. Nevertheless, this problem will not occur in output feedback case. ♡

The following theorem gives all full information controllers.

Theorem 16.10 *Suppose the condition in Theorem 16.9 is satisfied; then all admissible controllers satisfying* $\|T_{zw}\|_\infty < \gamma$ *can be parameterized as* $K = \mathcal{F}_\ell(M_{FI}, Q)$*:*



$$M_{FI}(s) = \left[\begin{array}{c|cc} A + B_2F_\infty & \begin{bmatrix} 0 & B_1 \end{bmatrix} & B_2 \\ \hline 0 & \begin{bmatrix} F_\infty & 0 \end{bmatrix} & I \\ \begin{bmatrix} -I \\ 0 \end{bmatrix} & \begin{bmatrix} I & 0 \\ -\gamma^{-2}B_1^*X_\infty & I \end{bmatrix} & 0 \end{array}\right]$$

where $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \in \mathcal{RH}_\infty$ *and* $\|Q_2\|_\infty < \gamma$.

Remark 16.3 It is simple to verify that for $Q_1 = 0$, we have

$$K = \begin{bmatrix} F_\infty - \gamma^{-2}Q_2B_1^*X_\infty & Q_2 \end{bmatrix},$$

which is the parameterization given in Theorem 16.9. The parameterization of all suboptimal FC controllers follows by duality and, therefore, are omitted. ♡

Proof. We only need to show that $\mathcal{F}_\ell(M_{FI}, Q)$ with $\|Q_2\|_\infty < \gamma$ parameterizes all FI $\mathcal{H}_\infty$ suboptimal controllers. To show that, we shall make a change of variables as before:

$$v = u + B_2^*X_\infty x, \qquad r = w - \gamma^{-2}B_1^*X_\infty x.$$

Then the system equations can be written as follows:

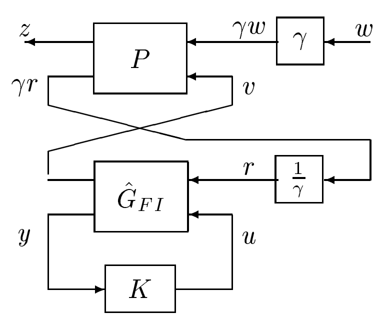
$$\begin{bmatrix} \dot{x} \\ z \\ \gamma r \end{bmatrix} = \begin{bmatrix} A_{F_\infty} & \gamma^{-1}B_1 & B_2 \\ C_{1F_\infty} & 0 & D_{12} \\ -\gamma^{-1}B_1^*X_\infty & I & 0 \end{bmatrix} \begin{bmatrix} x \\ \gamma w \\ v \end{bmatrix}$$

and

$$\begin{bmatrix} \dot{x} \\ v \\ y \end{bmatrix} = \begin{bmatrix} A_{tmp} & B_1 & B_2 \\ -F_\infty & 0 & I \\ \begin{bmatrix} I \\ \gamma^{-2}B_1^*X_\infty \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & 0 \end{bmatrix} \begin{bmatrix} x \\ r \\ u \end{bmatrix}$$

where $A_{tmp} := A + \gamma^{-2}B_1B_1^*X_\infty$.

This is shown pictorially in the following diagram:



where P is as given in equation (16.5) and

$$\hat{G}_{FI} = \left[\begin{array}{c|cc} A_{tmp} & B_1 & B_2 \\ \hline -F_\infty & 0 & I \\ \begin{bmatrix} I \\ \gamma^{-2}B_1^*X_\infty \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & 0 \end{array}\right].$$

So from Theorem 16.2 and Lemma 16.8, we conclude that K is an admissible controller for G and $\|T_{zw}\|_\infty < \gamma$ iff K is an admissible controller for $\hat{G}_{FI}$ and $\|T_{vr}\|_\infty < \gamma$. Now let $L = \begin{bmatrix} B_2F_\infty & -B_1 \end{bmatrix}$; then $A_{tmp} + L\begin{bmatrix} I \\ \gamma^{-2}B_1^*X_\infty \end{bmatrix} = A + B_2F_\infty$ is stable. Also note that $A_{tmp} + B_2F_\infty$ is stable. Then all controllers that stabilize $\hat{G}_{FI}$ can be parameterized as $K = \mathcal{F}_\ell(M, \Phi)$, $\Phi = \begin{bmatrix} \Phi_1 & \Phi_2 \end{bmatrix} \in \mathcal{RH}_\infty$ where

$$M = \left[\begin{array}{c|cc} A_{tmp} + B_2F_\infty + L\begin{bmatrix} I \\ \gamma^{-2}B_1^*X_\infty \end{bmatrix} & -L & B_2 \\ \hline F_\infty & 0 & I \\ -\begin{bmatrix} I \\ \gamma^{-2}B_1^*X_\infty \end{bmatrix} & I & 0 \end{array}\right].$$

With this parameterization of all controllers, the transfer matrix from r to v is $T_{vr} = \mathcal{F}_\ell(\hat{G}_{FI}, \mathcal{F}_\ell(M, \Phi)) =: \mathcal{F}_\ell(N, \Phi)$. It is easy to show that

$$N = \begin{bmatrix} 0 & I \\ \begin{bmatrix} 0 \\ I \end{bmatrix} & 0 \end{bmatrix}$$

and $T_{vr} = \mathcal{F}_\ell(N, \Phi) = \Phi_2$. Hence $\|T_{vr}\|_\infty < \gamma$ if and only if $\|\Phi_2\|_\infty < \gamma$. This implies that all FI $\mathcal{H}_\infty$ controllers can be parameterized as $K = \mathcal{F}_\ell(M, \Phi)$ with $\Phi = \begin{bmatrix} \Phi_1 & \Phi_2 \end{bmatrix} \in \mathcal{RH}_\infty$, $\|\Phi_2\|_\infty < \gamma$ and

$$M = \left[\begin{array}{c|cc} A + 2B_2F_\infty & -\begin{bmatrix} B_2F_\infty & -B_1 \end{bmatrix} & B_2 \\ \hline F_\infty & 0 & I \\ -\begin{bmatrix} I \\ \gamma^{-2}B_1^*X_\infty \end{bmatrix} & I & 0 \end{array}\right].$$

Now let

$$\Phi_1 = F_\infty - \gamma^{-2}Q_2B_1^*X_\infty + Q_1, \qquad \Phi_2 = Q_2.$$

Then it is easy to show that $\mathcal{F}_\ell(M, \Phi) = \mathcal{F}_\ell(M_{FI}, Q)$. □

16.5 Full Control

$$G(s) = \left[\begin{array}{c|c|c} A & B_1 & \begin{bmatrix} I & 0 \end{bmatrix} \\ \hline C_1 & 0 & \begin{bmatrix} 0 & I \end{bmatrix} \\ C_2 & D_{21} & \begin{bmatrix} 0 & 0 \end{bmatrix} \end{array}\right]$$

This problem is dual to the Full Information problem. The assumptions that the FC problem inherits from the output feedback problem are just the dual of those in the FI problem:

(i) (A, B_1) is stabilizable;

(ii) (C_2, A) is detectable;

(iv) $\begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} D_{21}^* = \begin{bmatrix} 0 \\ I \end{bmatrix}$.

Theorem 16.11 *There exists an admissible controller $K(s)$ for the FC problem such that $\|T_{zw}\|_\infty < \gamma$ if and only if $J_\infty \in dom(Ric)$ and $Y_\infty = Ric(J_\infty) \geq 0$. Moreover, if these conditions are satisfied then an equivalence class of all admissible controllers satisfying $\|T_{zw}\|_\infty < \gamma$ can be parameterized as*

$$K(s) \cong \begin{bmatrix} L_\infty - \gamma^{-2} Y_\infty C_1^* Q(s) \\ Q(s) \end{bmatrix}$$

where $Q \in \mathcal{RH}_\infty$, $\|Q\|_\infty < \gamma$.

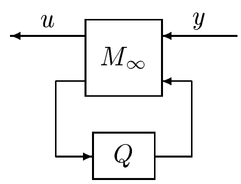
As expected, the condition in Theorem 16.11 is the same as that in (ii) of Theorem 16.4.

16.6 Disturbance Feedforward

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & I & 0 \end{array}\right]$$

This problem inherits the same assumptions (i)-(iii) as in the FI problem, but for internal stability we shall add that $A - B_1 C_2$ is stable. With this assumption, it is easy to check that the condition in Lemma 16.1 is satisfied so that the internal stability is again equivalent to $T_{zw} \in \mathcal{RH}_\infty$, as in the output feedback case.

Theorem 16.12 *There exists an admissible controller for the DF problem such that* $\|T_{zw}\|_\infty < \gamma$ *if and only if* $H_\infty \in dom(Ric)$ *and* $X_\infty = Ric(H_\infty) \geq 0$. *Moreover, if these conditions are satisfied then all admissible controllers satisfying* $\|T_{zw}\|_\infty < \gamma$ *can be parameterized as the set of all transfer matrices from* y *to* u *in*



$$M_\infty(s) = \left[\begin{array}{c|cc} A + B_2F_\infty - B_1C_2 & B_1 & B_2 \\ \hline F_\infty & 0 & I \\ -C_2 - \gamma^{-2}B_1^*X_\infty & I & 0 \end{array}\right]$$

with $Q \in \mathcal{RH}_\infty$, $\|Q\|_\infty < \gamma$

Proof. Suppose there is a controller K_{DF} solving the above problem, i.e., with $\|T_{zw}\|_\infty < \gamma$. Then by Theorem 12.4, the controller $K_{FI} = K_{DF} \begin{bmatrix} C_2 & I \end{bmatrix}$ solves the corresponding $\mathcal{H}_\infty$ FI problem. Hence the conditions $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) \geq 0$ are necessarily satisfied. On the other hand, if these conditions are satisfied then FI is solvable. It is easy to verify that $\mathcal{F}_\ell(M_\infty, Q) = \mathcal{F}_\ell(P_{DF}, K_{FI})$ with $K_{FI} = \begin{bmatrix} F_\infty - \gamma^{-2}Q(s)B_1^*X_\infty & Q(s) \end{bmatrix}$ where P_{DF} is as defined in section 12.2 of Chapter 12. So again by Theorem 12.4, the controller $\mathcal{F}_\ell(M_\infty, Q)$ solves the DF problem.

To show that $\mathcal{F}_\ell(M_\infty, Q)$ with $\|Q\|_\infty < \gamma$ parameterizes all DF $\mathcal{H}_\infty$ suboptimal controllers, we shall make a change of variables as in equation (16.4):

$$v = u + B_2^*X_\infty x, \qquad r = w - \gamma^{-2}B_1^*X_\infty x.$$

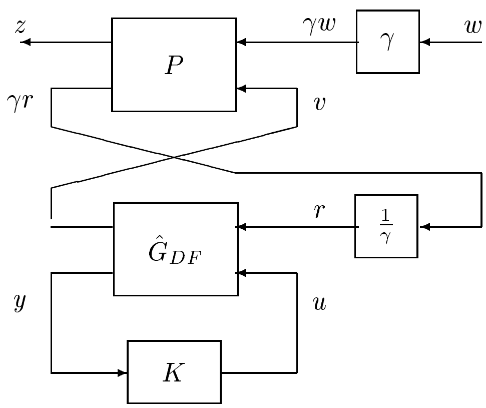
Then the system equations can be written as follows:

$$\begin{bmatrix} \dot{x} \\ z \\ \gamma r \end{bmatrix} = \begin{bmatrix} A_{F_\infty} & \gamma^{-1}B_1 & B_2 \\ C_{1F_\infty} & 0 & D_{12} \\ -\gamma^{-1}B_1^*X_\infty & I & 0 \end{bmatrix} \begin{bmatrix} x \\ \gamma w \\ v \end{bmatrix}$$

and

$$\begin{bmatrix} \dot{x} \\ v \\ y \end{bmatrix} = \begin{bmatrix} A_{tmp} & B_1 & B_2 \\ -F_\infty & 0 & I \\ C_2 + \gamma^{-2}B_1^*X_\infty & I & 0 \end{bmatrix} \begin{bmatrix} x \\ r \\ u \end{bmatrix}.$$

This is shown pictorially in the following diagram:



where

$$P = \left[\begin{array}{c|cc} A_{F_\infty} & \gamma^{-1}B_1 & B_2 \\ \hline C_{1F_\infty} & 0 & D_{12} \\ -\gamma^{-1}B_1^* X_\infty & I & 0 \end{array}\right]$$

and

$$\hat{G}_{DF} = \left[\begin{array}{c|cc} A_{tmp} & B_1 & B_2 \\ \hline -F_\infty & 0 & I \\ C_2 + \gamma^{-2}B_1^* X_\infty & I & 0 \end{array}\right].$$

Since $A_{tmp} - B_1(C_2 + \gamma^{-2}B_1^* X_\infty) = A - B_1C_2$ and $A_{tmp} + B_2F_\infty$ are stable, the rank conditions of Lemma 16.1 for system $\hat{G}_{DF}$ are satisfied. So from Theorem 16.2 and Lemma 16.7, we conclude that K is an admissible controller for G and $\|T_{zw}\|_\infty < \gamma$ iff K is an admissible controller for $\hat{G}_{DF}$ and $\|T_{vr}\|_\infty < \gamma$. Now it is easy to see by comparing this formula with the controller parameterization in Theorem 12.8 that $\mathcal{F}_\ell(M_\infty, Q)$ with $Q \in \mathcal{RH}_\infty$ (no norm constraint) parameterizes all stabilizing controllers for $\hat{G}_{DF}$; however, simple algebra shows that $T_{vr} = \mathcal{F}_\ell(\hat{G}_{DF}, \mathcal{F}_\ell(M_\infty, Q)) = Q$. So $\|T_{vr}\|_\infty < \gamma$ iff $\|Q\|_\infty < \gamma$, and $\mathcal{F}_\ell(M_\infty, Q)$ with $Q \in \mathcal{RH}_\infty$ and $\|Q\|_\infty < \gamma$ parameterizes all suboptimal controllers for G. □

16.7 Output Estimation

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & I \\ C_2 & D_{21} & 0 \end{array}\right]$$

This problem is dual to DF, just as FC was to FI. Thus the discussion of the DF problem is relevant here, when appropriately dualized. The OE assumptions are

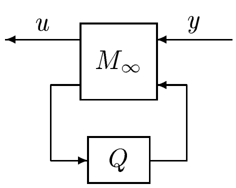
(i) (A, B_1) is stabilizable and $A - B_2C_1$ is stable;

(ii) (C_2, A) is detectable;

(iv) $\begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} D_{21}^* = \begin{bmatrix} 0 \\ I \end{bmatrix}$.

Assumption (i), together with (iv), imply that internal stability is again equivalent to $T_{zw} \in \mathcal{RH}_\infty$, as in the output feedback case.

Theorem 16.13 *There exists an admissible controller for the OE problem such that $\|T_{zw}\|_\infty < \gamma$ if and only if $J_\infty \in dom(Ric)$ and $Ric(J_\infty) \geq 0$. Moreover, if these conditions are satisfied then all admissible controllers satisfying $\|T_{zw}\|_\infty < \gamma$ can be parameterized as the set of all transfer matrices from y to u in*



$$M_\infty(s) = \left[\begin{array}{c|cc} A + L_\infty C_2 - B_2 C_1 & L_\infty & -B_2 - \gamma^{-2} Y_\infty C_1^* \\ \hline C_1 & 0 & I \\ C_2 & I & 0 \end{array}\right]$$

with $Q \in \mathcal{RH}_\infty$, $\|Q\|_\infty < \gamma$.

It is interesting to compare $\mathcal{H}_\infty$ and $\mathcal{H}_2$ in the context of the OE problem, even though, by duality, the essence of these remarks was made before. Both optimal estimators are observers with the observer gain determined by $Ric(J_\infty)$ and $Ric(J_2)$. Optimal $\mathcal{H}_2$ output estimation consists of multiplying the optimal state estimate by the output map C_1. Thus optimal $\mathcal{H}_2$ estimation depends only trivially on the output z that is being estimated, and *state* estimation is the fundamental problem. In contrast, the $\mathcal{H}_\infty$ estimation problem depends very explicitly and importantly on the output being estimated. This will have implications for the separation properties of the $\mathcal{H}_\infty$ output feedback controller.

16.8 Separation Theory

If we assume the results of the special problems, which are proven in the previous sections, we can now prove Theorems 16.4 and 16.5 using separation arguments. This essentially involves reducing the output feedback problem to a combination of the Full Information and the Output Estimation problems. The separation properties of the $\mathcal{H}_\infty$ controller are more complicated than the $\mathcal{H}_2$ controller, although they are no less interesting. The notation and assumptions for this section are as in Section 16.2.

16.8.1 $\mathcal{H}_\infty$ Controller Structure

The $\mathcal{H}_\infty$ controller formulae from Theorem 16.4 are

$$K_{sub}(s) := \left[\begin{array}{c|c} \hat{A}_\infty & -Z_\infty L_\infty \\ \hline F_\infty & 0 \end{array}\right]$$

$$\hat{A}_\infty := A + \gamma^{-2}B_1B_1^*X_\infty + B_2F_\infty + Z_\infty L_\infty C_2$$

$$F_\infty := -B_2^*X_\infty, \qquad L_\infty := -Y_\infty C_2^*, \qquad Z_\infty := (I - \gamma^{-2}Y_\infty X_\infty)^{-1}$$

where $X_\infty := Ric(H_\infty)$ and $Y_\infty := Ric(J_\infty)$. The necessary and sufficient conditions for the existence of an admissible controller such that $\|T_{zw}\|_\infty < \gamma$ are

(i) $H_\infty \in dom(Ric)$ and $X_\infty := Ric(H_\infty) \geq 0$;

(ii) $J_\infty \in dom(Ric)$ and $Y_\infty := Ric(J_\infty \geq 0$;

(iii) $\rho(X_\infty Y_\infty) < \gamma^2$.

We have seen that condition (i) corresponds to the Full Information condition and that (ii) corresponds to the Full Control condition. It is easily shown that, given the FI and FC results, these conditions are necessary for the output feedback case as well.

Lemma 16.14 *Suppose there exists an admissible controller making $\|T_{zw}\|_\infty < \gamma$. Then conditions (i) and (ii) hold.*

Proof. Let K be an admissible controller for which $\|T_{zw}\|_\infty < \gamma$. The controller $K[C_2 \quad D_{21}]$ solves the FI problem; hence from Theorem 16.9, $H_\infty \in dom(Ric)$ and $X_\infty := Ric(H_\infty) \geq 0$. Condition (ii) follows by the dual argument. □

We would also expect some condition beyond these two, and that is provided by (iii), which is an elegant combination of elements from FI and FC. Note that all the conditions of Theorem 16.4 are symmetric in H_∞, J_∞, X_∞, and Y_∞, but the formula for the controller is not. Needless to say, there is a dual form that can be obtained by inspection from the above formula. For a symmetric formula, the state equations above can be multiplied through by Z_∞^{-1} and put in descriptor form. A simple substitution from the Riccati equation for X_∞ will then yield a symmetric, though more complicated, formula:

$$(I - \gamma^{-2}Y_\infty X_\infty)\dot{\hat{x}} = A_s\hat{x} - L_\infty y \tag{16.10}$$

$$u = F_\infty \hat{x} \tag{16.11}$$

where $A_s := A + B_2F_\infty + L_\infty C_2 + \gamma^{-2}Y_\infty A^*X_\infty + \gamma^{-2}B_1B_1^*X_\infty + \gamma^{-2}Y_\infty C_1^*C_1$.

To emphasize its relationship to the $\mathcal{H}_2$ controller formulae, the $\mathcal{H}_\infty$ controller can be written as

$$\dot{\hat{x}} = A\hat{x} + B_1\hat{w}_{worst} + B_2u + Z_\infty L_\infty(C_2\hat{x} - y)$$

$$u = F_\infty \hat{x}, \qquad \hat{w}_{worst} = \gamma^{-2} B_1^* X_\infty \hat{x}.$$

These equations have the structure of an observer-based compensator. The obvious questions that arise when these formulae are compared with the $\mathcal{H}_2$ formulae are

1) Where does the term $B_1 \hat{w}_{worst}$ come from?

2) Why $Z_\infty L_\infty$ instead of L_∞?

3) Is there a separation interpretation of these formulae analogous to that for $\mathcal{H}_2$?

The proof of Theorem 16.4 reveals that there is a very well-defined separation interpretation of these formulae and that $w_{worst} := \gamma^{-2} B_1^* X_\infty x$ is, in some sense, a worst-case input for the Full Information problem. Furthermore, $Z_\infty L_\infty$ is actually the optimal filter gain for estimating $F_\infty x$, which is the optimal Full Information control input, in the presence of this worst-case input. It is therefore not surprising that $Z_\infty L_\infty$ should enter in the controller equations instead of L_∞. The term $\hat{w}_{worst}$ may be thought of loosely as an estimate for w_{worst}.

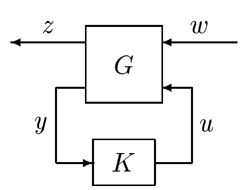
16.8.2 Proof of Theorem 16.4

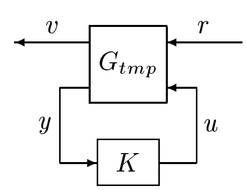
It has been shown from Lemma 16.14 that conditions (i) and (ii) are necessary for $\|T_{zw}\|_\infty < \gamma$. Hence we only need to show that if conditions (i) and (ii) are satisfied, condition (iii) is necessary and sufficient for $\|T_{zw}\|_\infty < \gamma$. As in section 16.3, we define new disturbance and control variables

$$r := w - \gamma^{-2} B_1^* X_\infty x, \qquad v := u + B_2^* X_\infty x.$$

Then

$$\begin{bmatrix} v \\ y \end{bmatrix} = \left[\begin{array}{c|cc} A_{tmp} & B_1 & B_2 \\ \hline -F_\infty & 0 & I \\ C_2 & D_{21} & 0 \end{array}\right] \begin{bmatrix} r \\ u \end{bmatrix} = G_{tmp} \begin{bmatrix} r \\ u \end{bmatrix} \qquad A_{tmp} := A + \gamma^{-2} B_1 B_1^* X_\infty.$$





Recall from Lemma 16.8 that K is admissible for G and $\|T_{zw}\|_\infty < \gamma$ iff K is admissible for G_{tmp} and $\|T_{vr}\|_\infty < \gamma$.

While G_{tmp} has the form required for the OE problem, to actually use the OE results, we will need to verify that G_{tmp} satisfies the following assumptions for the OE problem:

(i) (A_{tmp}, B_1) is stabilizable and $A_{tmp} + B_2F_\infty$ is stable;

(ii) (C_2, A_{tmp}) is detectable;

(iv) $\begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} D_{21}^* = \begin{bmatrix} 0 \\ I \end{bmatrix}$.

Assumption (iv) and that (A_{tmp}, B_1) is stabilizable follow immediately from the corresponding assumptions for Theorem 16.4. The stability of $A_{tmp} + B_2F_\infty$ follows from the definition of $H_\infty \in dom(Ric)$. The following lemma gives conditions for assumption (ii) to hold. Of course, the existence of an admissible controller for G_{tmp} immediately implies that assumption (ii) holds. Note that the OE Hamiltonian matrix for G_{tmp} is

$$J_{tmp} := \begin{bmatrix} A_{tmp}^* & \gamma^{-2}F_\infty^* F_\infty - C_2^*C_2 \\ -B_1B_1^* & -A_{tmp} \end{bmatrix}.$$

Lemma 16.15 *If $J_{tmp} \in dom(Ric)$ and $Y_{tmp} := Ric(J_{tmp}) \geq 0$, then (C_2, A_{tmp}) is detectable.*

Proof. The lemma follows from the dual to Lemma 16.6, which gives that $(A_{tmp} - Y_{tmp}C_2^*C_2)$ is stable. □

Proof of Theorem 16.4 (Sufficiency) Assume the conditions (i) through (iii) in the theorem statement hold. Using the Riccati equation for X_∞, one can easily verify that $T := \begin{bmatrix} I & -\gamma^{-2}X_\infty \\ 0 & I \end{bmatrix}$ provides a similarity transformation between J_{tmp} and J_∞, i.e., $T^{-1}J_{tmp}T = J_\infty$. So

$$\mathcal{X}_-(J_{tmp}) = T\mathcal{X}_-(J_\infty) = T \ \mathrm{Im} \begin{bmatrix} I \\ Y_\infty \end{bmatrix} = \mathrm{Im} \begin{bmatrix} I - \gamma^{-2}X_\infty Y_\infty \\ Y_\infty \end{bmatrix}$$

and $\rho(X_\infty Y_\infty) < \gamma^2$ implies that $J_{tmp} \in dom(Ric)$ and $Y_{tmp} := Ric(J_{tmp}) = Y_\infty(I - \gamma^{-2}X_\infty Y_\infty)^{-1} = Z_\infty Y_\infty \geq 0$. Thus by Lemma 16.15 the OE assumptions hold for G_{tmp}, and by Theorem 16.13 the OE problem is solvable. From Theorem 16.13 with $Q = 0$, one solution is

$$\left[\begin{array}{c|c} A + \gamma^{-2}B_1B_1^*X_\infty - Y_{tmp}C_2^*C_2 + B_2F_\infty & Y_{tmp}C_2^* \\ \hline F_\infty & 0 \end{array}\right]$$

but this is precisely K_{sub} defined in Theorem 16.4. We conclude that K_{sub} stabilizes G_{tmp} and that $\|T_{vr}\|_\infty < \gamma$. Then by Lemma 16.8, K_{sub} stabilizes G and that $\|T_{zw}\|_\infty < \gamma$.

(Necessity) Let K be an admissible controller for which $\|T_{zw}\|_\infty < \gamma$. By Lemma 16.14, $H_\infty \in dom(Ric)$, $X_\infty := Ric(H_\infty) \geq 0$, $J_\infty \in dom(Ric)$, and $Y_\infty := Ric(J_\infty) \geq 0$. From Lemma 16.8, K is admissible for G_{tmp} and $\|T_{vr}\|_\infty < \gamma$. This implies that the OE assumptions hold for G_{tmp} and that the OE problem is solvable. Therefore, from Theorem 16.13 applied to G_{tmp}, we have that $J_{tmp} \in dom(Ric)$ and $Y_{tmp} = Ric(J_{tmp}) \geq 0$. Using the same similarity transformation formula as in the sufficiency part, we get that $Y_{tmp} = (I - \gamma^{-2}Y_\infty X_\infty)^{-1}Y_\infty \geq 0$. We shall now show that $Y_{tmp} \geq 0$ implies that $\rho(X_\infty Y_\infty) < \gamma^2$. We shall consider two cases:

- Y_∞ is nonsingular: in this case $Y_{tmp} \geq 0$ implies that $I - \gamma^{-2}Y_\infty^{1/2}X_\infty Y_\infty^{1/2} > 0$. So $\rho(Y_\infty^{1/2}X_\infty Y_\infty^{1/2}) < \gamma^2$ or $\rho(X_\infty Y_\infty) < \gamma^2$.
- Y_∞ is singular: there is a unitary matrix U such that

$$Y_\infty = U^* \begin{bmatrix} Y_{11} & 0 \\ 0 & 0 \end{bmatrix} U$$

with $Y_{11} > 0$. Let $UX_\infty U^*$ be partitioned accordingly,

$$UX_\infty U^* = \begin{bmatrix} X_{11} & X_{12} \\ X_{21} & X_{22} \end{bmatrix}.$$

Then by the same argument as in the Y_∞ nonsingular case,

$$Y_{tmp} = U^* \begin{bmatrix} (I - \gamma^{-2}Y_{11}X_{11})^{-1}Y_{11} & 0 \\ 0 & I \end{bmatrix} U \geq 0$$

implies that $\gamma^2 > \rho(X_{11}Y_{11})$ $(= \rho(X_\infty Y_\infty))$. □

We now see exactly why the term involving $\hat{w}_{worst}$ appears and why the "observer" gain is $Z_\infty L_\infty$. Both terms are consequences of estimating the optimal Full Information (i.e., state feedback) control gain. While an analogous output estimation problem arises in the $\mathcal{H}_2$ output feedback problem, the resulting equations are much simpler. This is because there is no "worst-case" disturbance for the $\mathcal{H}_2$ Full Information problem and because the problem of estimating any output, including the optimal state feedback, is equivalent to state estimation.

We now present a separation interpretation for $\mathcal{H}_\infty$ suboptimal controllers. It will be stated in terms of the central controller, but similar interpretations could be made for the parameterization of all suboptimal controllers (see the proofs of Theorems 16.4 and 16.5).

> *The $\mathcal{H}_\infty$ output feedback controller is the output estimator of the full information control law in the presence of the "worst-case" disturbance w_{worst}.*

Note that the same statement holds for the $\mathcal{H}_2$ optimal controller, except that $w_{worst} = 0$.

16.8.3 Proof of Theorem 16.5

From Lemma 16.8, the set of all admissible controllers for G such that $\|T_{zw}\|_\infty < \gamma$ equals the set of all admissible controllers for G_{tmp} such that $\|T_{vr}\|_\infty < \gamma$. Apply Theorem 16.13. □

16.9 Optimality and Limiting Behavior

In this section, we will discuss, without proof, the behavior of the $\mathcal{H}_\infty$ suboptimal solution as γ varies, especially as γ approaches the infimal achievable norm, denoted by γ_{opt}. Since Theorem 16.4 gives necessary and sufficient conditions for the existence of an admissible controller such that $\|T_{zw}\|_\infty < \gamma$, γ_{opt} is the infimum over all γ such that conditions (i)-(iii) are satisfied. Theorem 16.4 does not give an explicit formula for γ_{opt}, but, just as for the $\mathcal{H}_\infty$ norm calculation, it can be computed as closely as desired by a search technique.

Although we have not focused on the problem of $\mathcal{H}_\infty$ *optimal* controllers, the assumptions in this book make them relatively easy to obtain in most cases. In addition to describing the qualitative behavior of suboptimal solutions as γ varies, we will indicate why the descriptor version of the controller formulae from Section 16.8.1 can usually provide formulae for the optimal controller when $\gamma = \gamma_{opt}$. Most of these results can be obtained relatively easily using the machinery that is developed in the previous sections. The reader interested in filling in the details is encouraged to begin by strengthening assumption (i) to controllable and observable and considering the Hamiltonians for X_∞^{-1} and Y_∞^{-1}.

As $\gamma \to \infty$, $H_\infty \to H_2$, $X_\infty \to X_2$, etc., and $K_{sub} \to K_2$. This fact is the result of the particular choice for the central controller ($Q = 0$) that was made here. While it could be argued that K_{sub} is a natural choice, this connection with $\mathcal{H}_2$ actually hints at deeper interpretations. In fact, K_{sub} is the minimum entropy solution (see next section) as well as the minimax controller for $\|z\|_2^2 - \gamma^2\|w\|_2^2$.

If $\gamma_2 \geq \gamma_1 > \gamma_{opt}$, then $X_\infty(\gamma_1) \geq X_\infty(\gamma_2)$ and $Y_\infty(\gamma_1) \geq Y_\infty(\gamma_2)$. Thus X_∞ and Y_∞ are decreasing functions of γ, as is $\rho(X_\infty Y_\infty)$. At $\gamma = \gamma_{opt}$, anyone of the three conditions in Theorem 16.4 can fail. If only condition (iii) fails, then it is relatively straightforward to show that the descriptor formulae for $\gamma = \gamma_{opt}$ are optimal, i.e., the optimal controller is given by

$$(I - \gamma_{opt}^{-2} Y_\infty X_\infty)\dot{\hat{x}} = A_s \hat{x} - L_\infty y \tag{16.12}$$
$$u = F_\infty \hat{x} \tag{16.13}$$

where $A_s := A + B_2 F_\infty + L_\infty C_2 + \gamma_{opt}^{-2} Y_\infty A^* X_\infty + \gamma_{opt}^{-2} B_1 B_1^* X_\infty + \gamma_{opt}^{-2} Y_\infty C_1^* C_1$. See the example below.

The formulae in Theorem 16.4 are not well-defined in the optimal case because the term $(I - \gamma_{opt}^{-2} X_\infty Y_\infty)$ is not invertible. It is possible but far less likely that conditions (i) or (ii) would fail before (iii). To see this, consider (i) and let γ_1 be the largest γ

for which H_∞ fails to be in $dom(Ric)$ because the H_∞ matrix fails to have either the stability property or the complementarity property. The same remarks will apply to (ii) by duality.

If complementarity fails at $\gamma = \gamma_1$, then $\rho(X_\infty) \to \infty$ as $\gamma \to \gamma_1$. For $\gamma < \gamma_1$, H_∞ may again be in $dom(Ric)$, but X_∞ will be indefinite. For such γ, the controller $u = -B_2^* X_\infty x$ would make $\|T_{zw}\|_\infty < \gamma$ but would not be stabilizing. See part 1) of the example below. If the stability property fails at $\gamma = \gamma_1$, then $H_\infty \notin dom(Ric)$ but Ric can be extended to obtain X_∞ and the controller $u = -B_2^* X_\infty x$ is stabilizing and makes $\|T_{zw}\|_\infty = \gamma_1$. The stability property will also not hold for any $\gamma \le \gamma_1$, and no controller whatsoever exists which makes $\|T_{zw}\|_\infty < \gamma_1$. In other words, if stability breaks down first, then the infimum over stabilizing controllers equals the infimum over all controllers, stabilizing or otherwise. See part 2) of the example below. In view of this, we would typically expect that complementarity would fail first.

Complementarity failing at $\gamma = \gamma_1$ means $\rho(X_\infty) \to \infty$ as $\gamma \to \gamma_1$, so condition (iii) would fail at even larger values of γ, unless the eigenvectors associated with $\rho(X_\infty)$ as $\gamma \to \gamma_1$ are in the null space of Y_∞. Thus condition (iii) is the most likely of all to fail first. If condition (i) or (ii) fails first because the stability property fails, the formulae in Theorem 16.4 as well as their descriptor versions are optimal at $\gamma = \gamma_{opt}$. This is illustrated in the example below for the output feedback. If the complementarity condition fails first, (but (iii) does not fail), then obtaining formulae for the optimal controllers is a more subtle problem.

Example 16.1 Let an interconnected dynamical system realization be given by

$$G(s) = \left[\begin{array}{c|cc} a & \begin{bmatrix} 1 & 0 \end{bmatrix} & b_2 \\ \hline \begin{bmatrix} 1 \\ 0 \end{bmatrix} & 0 & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \\ c_2 & \begin{bmatrix} 0 & 1 \end{bmatrix} & 0 \end{array}\right]$$

with $|c_2| \ge |b_2| > 0$. Then all assumptions for the output feedback problem are satisfied and

$$H_\infty = \begin{bmatrix} a & \frac{1-b_2^2\gamma^2}{\gamma^2} \\ -1 & -a \end{bmatrix}, \quad J_\infty = \begin{bmatrix} a & \frac{1-c_2^2\gamma^2}{\gamma^2} \\ -1 & -a \end{bmatrix}.$$

The eigenvalues of H_∞ and J_∞ are given, respectively, by

$$\sigma(H_\infty) = \left\{ \pm \frac{\sqrt{(a^2+b_2^2)\gamma^2 - 1}}{\gamma} \right\}, \quad \sigma(J_\infty) = \left\{ \pm \frac{\sqrt{(a^2+c_2^2)\gamma^2 - 1}}{\gamma} \right\}.$$

If $\gamma^2 > \dfrac{1}{a^2+b_2^2} \left(\ge \dfrac{1}{a^2+c_2^2} \right)$, then $\mathcal{X}_-(H_\infty)$ and $\mathcal{X}_-(J_\infty)$ exist and

$$\mathcal{X}_-(H_\infty) = \operatorname{Im} \begin{bmatrix} \frac{\sqrt{(a^2+b_2^2)\gamma^2-1} - a\gamma}{\gamma} \\ 1 \end{bmatrix}$$

$$\mathcal{X}_-(J_\infty) = \operatorname{Im}\begin{bmatrix} \frac{\sqrt{(a^2+c_2^2)\gamma^2-1}-a\gamma}{\gamma} \\ 1 \end{bmatrix}.$$

We shall consider two cases:

1) $a > 0$: In this case, the complementary property of $dom(Ric)$ will fail before the stability property fails since

$$\sqrt{(a^2+b_2^2)\gamma^2-1} - a\gamma = 0$$

when $\gamma^2 = \dfrac{1}{b_2^2} \quad \left(> \dfrac{1}{a^2+b_2^2}\right)$.

Nevertheless, if $\gamma^2 > \dfrac{1}{a^2+b_2^2}$ and $\gamma^2 \neq \dfrac{1}{b_2^2}$, then $H_\infty \in dom(Ric)$ and

$$X_\infty = \frac{\gamma}{\sqrt{(a^2+b_2^2)\gamma^2-1}-a\gamma} = \begin{cases} >0; & \text{if } \gamma^2 > \frac{1}{b_2^2} \\ <0; & \text{if } \frac{1}{a^2+b_2^2} < \gamma^2 < \frac{1}{b_2^2}. \end{cases}$$

Let $F_\infty = -B_2^* X_\infty$; then

$$A+B_2F_\infty = -\frac{a+b_2^2\gamma\sqrt{(a^2+b_2^2)\gamma^2-1}}{b_2^2\gamma^2-1} = \begin{cases} <0 \text{ (stable)}; & \text{if } \gamma^2 > \frac{1}{b_2^2} \\ >0 \text{ (unstable)}; & \text{if } \frac{1}{a^2+b_2^2} < \gamma^2 < \frac{1}{b_2^2}. \end{cases}$$

– Suppose full information (or states) are available for feedback and let

$$u = F_\infty x.$$

Then the closed-loop transfer matrix is given by

$$T_{zw} = \left[\begin{array}{c|c} A+B_2F_\infty & B_1 \\ \hline C_1+D_{12}F_\infty & 0 \end{array}\right] = \left[\begin{array}{c|c} -\frac{a+b_2^2\gamma\sqrt{(a^2+b_2^2)\gamma^2-1}}{b_2^2\gamma^2-1} & \begin{bmatrix} 1 & 0 \end{bmatrix} \\ \hline \begin{bmatrix} 1 \\ \frac{-b_2\gamma}{\sqrt{(a^2+b_2^2)\gamma^2-1}-a\gamma} \end{bmatrix} & 0 \end{array}\right],$$

and T_{zw} is stable for all $\gamma^2 > \dfrac{1}{b_2^2}$ and is not stable for $\dfrac{1}{a^2+b_2^2} < \gamma^2 < \dfrac{1}{b_2^2}$. Furthermore, it can be shown that $\|T_{zw}\| < \gamma$ for all $\gamma^2 > \dfrac{1}{a^2+b_2^2}$ and $\gamma^2 \neq \dfrac{1}{b_2^2}$.

It is clear that the optimal $\mathcal{H}_\infty$ norm is $\frac{1}{b_2^2}$ but is not achievable.

– Suppose the states are not available; then output feedback must be considered. Note that if $\gamma^2 > \dfrac{1}{b_2^2}$, then $H_\infty \in dom(Ric)$, $J_\infty \in dom(Ric)$, and

$$X_\infty = \frac{\gamma}{\sqrt{(a^2+b_2^2)\gamma^2 - 1} - a\gamma} > 0$$

$$Y_\infty = \frac{\gamma}{\sqrt{(a^2+c_2^2)\gamma^2 - 1} - a\gamma} > 0.$$

Hence conditions (i) and (ii) in Theorem 16.4 are satisfied, and need to check condition (iii). Since

$$\rho(X_\infty Y_\infty) = \frac{\gamma^2}{(\sqrt{(a^2+b_2^2)\gamma^2-1} - a\gamma)(\sqrt{(a^2+c_2^2)\gamma^2-1} - a\gamma)},$$

it is clear that $\rho(X_\infty Y_\infty) \to \infty$ when $\gamma^2 \to \dfrac{1}{b_2^2}$. So condition (iii) will fail before condition (i) or (ii) fails.

2) $a < 0$: In this case, complementary property is always satisfied, and, furthermore, $H_\infty \in dom(Ric)$, $J_\infty \in dom(Ric)$, and

$$X_\infty = \frac{\gamma}{\sqrt{(a^2+b_2^2)\gamma^2 - 1} - a\gamma} > 0$$

$$Y_\infty = \frac{\gamma}{\sqrt{(a^2+c_2^2)\gamma^2 - 1} - a\gamma} > 0$$

for $\gamma^2 > \dfrac{1}{a^2+b_2^2}$.

However, for $\gamma^2 \le \dfrac{1}{a^2+b_2^2}$, $H_\infty \notin dom(Ric)$ since stability property fails. Nevertheless, in this case, if $\gamma_0^2 = \dfrac{1}{a^2+b_2^2}$, we can extend the $dom(Ric)$ to include those matrices H_∞ with imaginary axis eigenvalues as

$$\overline{\mathcal{X}_-}(H_\infty) = \operatorname{Im}\begin{bmatrix} -a \\ 1 \end{bmatrix}$$

such that $X_\infty = -\dfrac{1}{a}$ is a solution to the Riccati equation

$$A^*X_\infty + X_\infty A + C_1^*C_1 + \gamma_0^{-2}X_\infty B_1B_1^*X_\infty - X_\infty B_2B_2^*X_\infty = 0$$

and $A + \gamma_0^{-2}B_1B_1^*X_\infty - B_2B_2^*X_\infty = 0$.

– For $\gamma = \gamma_0$ and $F_\infty = -B_2^* X_\infty$, then $A + B_2 F_\infty = a + \dfrac{b_2^2}{a} < 0$. So if states are available for feedback and $u = F_\infty x$, we have

$$T_{zw} = \left[\begin{array}{c|c} a + \frac{b_2^2}{a} & \begin{bmatrix} 1 & 0 \end{bmatrix} \\ \hline \begin{bmatrix} 1 \\ \frac{b_2}{a} \end{bmatrix} & 0 \end{array}\right] \in \mathcal{RH}_\infty$$

and $\|T_{zw}\|_\infty = \dfrac{1}{\sqrt{a^2 + b_2}} = \gamma_0$. Hence the optimum is achieved.

– If states are not available, the output feedback is considered, and $|b_2| = |c_2|$, then it can be shown that

$$\rho(X_\infty Y_\infty) = \frac{\gamma^2}{(\sqrt{(a^2 + b_2^2)\gamma^2 - 1} - a\gamma)^2} < \gamma^2$$

is satisfied if and only if

$$\gamma > \frac{\sqrt{a^2 + 2b_2^2} + a}{b_2^2} \quad \left(> \frac{1}{\sqrt{a^2 + b_2^2}} \right).$$

So condition (iii) of Theorem 16.4 will fail before either (i) or (ii) fails.

In both $a > 0$ and $a < 0$ cases, the optimal γ for the output feedback is given by

$$\gamma_{opt} = \frac{\sqrt{a^2 + 2b_2^2} + a}{b_2^2}$$

if $|b_2| = |c_2|$; and the optimal controller given by the descriptor formula in equations (16.12) and (16.13) is a constant. In fact,

$$u_{opt} = -\frac{\gamma_{opt}}{\sqrt{(a^2 + b_2^2)\gamma_{opt}^2 - 1} - a\gamma_{opt}}\, y.$$

For instance, let $a = -1$ and $b_2 = 1 = c_2$. Then $\gamma_{opt} = \sqrt{3} - 1 = 0.7321$ and $u_{opt} = -0.7321\, y$. Further,

$$T_{zw} = \left[\begin{array}{c|cc} -1.7321 & 1 & -0.7321 \\ \hline 1 & 0 & 0 \\ -0.7321 & 0 & -0.7321 \end{array}\right].$$

It is easy to check that $\|T_{zw}\|_\infty = 0.7321$. ◇

16.10 Controller Interpretations

This section considers some additional connections with the minimum entropy solution and the work of Whittle and will be of interest primarily to readers already familiar with them. The connection with the Q-parameterization approach will be considered in the next chapter for the general case.

Section 16.10.2 gives another separation interpretation of the central $\mathcal{H}_\infty$ controller of Theorem 16.4 in the spirit of Whittle (1981). It has been shown in Glover and Doyle [1988] that the central controller corresponds exactly to the steady state version of the optimal risk sensitive controller derived by [Whittle, 1981], who also derives a separation result and a certainty equivalence principle (see also [Whittle, 1986]).

16.10.1 Minimum Entropy Controller

Let T be a transfer matrix with $\|T\|_\infty < \gamma$. Then the entropy of $T(s)$ is defined by

$$I(T,\gamma) = -\frac{\gamma^2}{2\pi}\int_{-\infty}^{\infty} \ln\left|\det\left(I - \gamma^{-2}T^*(j\omega)T(j\omega)\right)\right| d\omega.$$

It is easy to see that

$$I(T,\gamma) = -\frac{\gamma^2}{2\pi}\int_{-\infty}^{\infty} \sum_i \ln\left|1 - \gamma^{-2}\sigma_i^2\left(T(j\omega)\right)\right| d\omega$$

and $I(T,\gamma) \geq 0$, where $\sigma_i\left(T(j\omega)\right)$ is the ith singular value of $T(j\omega)$. It is also easy to show that

$$\lim_{\gamma\to\infty} I(T,\gamma) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \sum_i \sigma_i^2\left(T(j\omega)\right) d\omega = \|T\|_2^2.$$

Thus the entropy $I(T,\gamma)$ is in fact a performance index measuring the tradeoff between the $\mathcal{H}_\infty$ optimality ($\gamma \to \|T\|_\infty$) and the $\mathcal{H}_2$ optimality ($\gamma \to \infty$).

It has been shown in Glover and Mustafa [1989] that the central controller given in Theorem 16.4 is actually the controller that satisfies the norm condition $\|T_{zw}\|_\infty < \gamma$ and minimizes the following entropy:

$$-\frac{\gamma^2}{2\pi}\int_{-\infty}^{\infty} \ln\left|\det\left(I - \gamma^{-2}T_{zw}^*(j\omega)T_{zw}(j\omega)\right)\right| d\omega.$$

Therefore, the central controller is also called the minimum entropy controller (maximum entropy controller if the entropy is defined as $\tilde{I}(T,\gamma) = -I(T,\gamma)$).

16.10.2 Relations with Separation in Risk Sensitive Control

Although [Whittle, 1981] treats a finite horizon, discrete time, stochastic control problem, his separation result has a clear interpretation for the present infinite horizon, continuous time, deterministic control problem, as given below; and it is an interesting exercise to compare the two separation statements. This discussion will be entirely in the time-domain.

We will consider the system at time, $t = 0$, and evaluate the *past stress*, $\mathcal{S}_-$, and *future stress*, $\mathcal{S}_+$, as functions of the current state, x. First define the future stress as

$$\mathcal{S}_+(x) := \sup_w \inf_u (\|P_+ z\|_2^2 - \gamma^2 \|P_+ w\|_2^2);$$

then by the completion of the squares and by the saddle point argument of Section 16.3, where u is not constrained to be a function of the measurements (FI case), we obtain

$$\mathcal{S}_+(x) = x^* X_\infty x.$$

The *past stress*, $\mathcal{S}_-(x)$, is a function of the past inputs and observations, $u(t), y(t)$ for $-\infty < t < 0$, and of the present state, x, and is produced by the worst case disturbance, w, that is consistent with the given data:

$$\mathcal{S}_-(x) := \sup(\|P_- z\|_2^2 - \gamma^2 \|P_- w\|_2^2).$$

In order to evaluate $\mathcal{S}_-$ we see that w can be divided into two components, $D_{21} w$ and $D_{21}^\perp w$, with x dependent only on $D_{21}^\perp w$ (since $B_1 D_{21}^* = 0$) and $D_{21} w = y - C_2 x$. The past stress is then calculated by a completion of the square and in terms of a filter output. In particular, let $\bar{x}$ be given by the stable differential equation

$$\dot{\bar{x}} = A\bar{x} + B_2 u + L_\infty (C_2 \bar{x} - y) + Y_\infty C_1^* C_1 \bar{x} \quad \text{with} \quad \bar{x}(-\infty) = 0.$$

Then it can be shown that the worst case w is given by

$$D_{21}^\perp w = D_{21}^\perp B_1^* Y_\infty^{-1} (x(t) - \bar{x}(t)) \quad \text{for } t < 0$$

and that this gives, with $e := x - \bar{x}$,

$$\mathcal{S}_-(x) = -\gamma^2 e(0)^* Y_\infty^{-1} e(0) - \gamma^2 \|P_-(y - C_2 \bar{x})\|_2^2 + \|P_-(C_1 \bar{x})\|_2^2 + \|P_- u\|_2^2.$$

The worst case disturbance will now reach the value of x to maximize the total stress, $\mathcal{S}_-(x) + \mathcal{S}_+(x)$, and this is easily shown to be achieved at the current state of

$$\hat{x} = Z_\infty \bar{x}(0).$$

The definitions of X_∞ and Y_∞ can be used to show that the state equations for the central controller can be rewritten with the state $\bar{x} := Z_\infty \hat{x}$ and with $\bar{x}$ as defined above. The control signal is then

$$u = F_\infty \hat{x} = F_\infty Z_\infty \bar{x}.$$

The *separation* is between the evaluation of future stress, which is a control problem with an unconstrained input, and the past stress, which is a filtering problem with known control input. The central controller then combines these evaluations to give a worst case estimate, $\hat{x}$, and the control law acts as if this were the perfectly observed state.

16.11 An Optimal Controller

To offer a general idea about the appearance of an optimal controller, we shall give in the following without proof the conditions under which an optimal controller exists and an explicit formula for an optimal controller.

Theorem 16.16 *There exists an admissible controller such that $\|T_{zw}\|_\infty \leq \gamma$ iff the following three conditions hold:*

(i) there exists a full column rank matrix

$$\begin{bmatrix} X_{\infty 1} \\ X_{\infty 2} \end{bmatrix} \in \mathbb{R}^{2n \times n}$$

such that

$$H_\infty \begin{bmatrix} X_{\infty 1} \\ X_{\infty 2} \end{bmatrix} = \begin{bmatrix} X_{\infty 1} \\ X_{\infty 2} \end{bmatrix} T_X, \quad \operatorname{Re} \lambda_i(T_X) \leq 0 \ \forall i$$

and

$$X_{\infty 1}^* X_{\infty 2} = X_{\infty 2}^* X_{\infty 1};$$

(ii) there exists a full column rank matrix

$$\begin{bmatrix} Y_{\infty 1} \\ Y_{\infty 2} \end{bmatrix} \in \mathbb{R}^{2n \times n}$$

such that

$$J_\infty \begin{bmatrix} Y_{\infty 1} \\ Y_{\infty 2} \end{bmatrix} = \begin{bmatrix} Y_{\infty 1} \\ Y_{\infty 2} \end{bmatrix} T_Y, \quad \operatorname{Re} \lambda_i(T_Y) \leq 0 \ \forall i$$

and

$$Y_{\infty 1}^* Y_{\infty 2} = Y_{\infty 2}^* Y_{\infty 1};$$

(iii)

$$\begin{bmatrix} X_{\infty 2}^* X_{\infty 1} & \gamma^{-1} X_{\infty 2}^* Y_{\infty 2} \\ \gamma^{-1} Y_{\infty 2}^* X_{\infty 2} & Y_{\infty 2}^* Y_{\infty 1} \end{bmatrix} \geq 0.$$

Moreover, when these conditions hold, one such controller is

$$K_{opt}(s) := C_K (sE_K - A_K)^+ B_K$$

where

$$\begin{aligned} E_K &:= Y_{\infty 1}^* X_{\infty 1} - \gamma^{-2} Y_{\infty 2}^* X_{\infty 2} \\ B_K &:= Y_{\infty 2}^* C_2^* \\ C_K &:= -B_2^* X_{\infty 2} \\ A_K &:= E_K T_X - B_K C_2 X_{\infty 1} = T_Y^* E_K + Y_{\infty 1}^* B_2 C_K. \end{aligned}$$

Remark 16.4 It is simple to show that if $X_{\infty 1}$ and $Y_{\infty 1}$ are nonsingular and if $X_\infty = X_{\infty 2} X_{\infty 1}^{-1}$ and $Y_\infty = Y_{\infty 2} Y_{\infty 1}^{-1}$, then condition (iii) in the above theorem is equivalent to $X_\infty \geq 0$, $Y_\infty \geq 0$, and $\rho(Y_\infty X_\infty) \leq \gamma^2$. So in this case, the conditions for the existence of an optimal controller can be obtained from "taking the limit" of the corresponding conditions in Theorem 16.4. Moreover, the controller given above is reduced to the descriptor form given in equations (16.12) and (16.13). ♡

16.12 Notes and References

This chapter is based on Doyle, Glover, Khargonekar, Francis [1989], and Zhou [1992]. The minimum entropy controller is studied in detail in Glover and Mustafa [1989] and Mustafa and Glover [1990]. The risk sensitivity problem is treated in Whittle [1981, 1986]. The connections between the risk sensitivity controller and the central $\mathcal{H}_\infty$ controller are explored in Doyle and Glover [1988]. The complete characterization of optimal controllers for the general setup can be found in Glover, Limebeer, Doyle, Kasenally, and Safonov [1991].

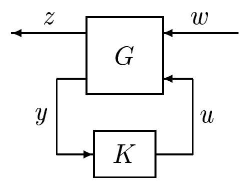
17

$\mathcal{H}_\infty$ Control: General Case

In this chapter we will consider again the standard $\mathcal{H}_\infty$ control problem but with some assumptions in the last chapter relaxed. Since the proof techniques in the last chapter can be applied to this general case except with some more involved algebra, the detailed proof for the general case will not be given; only the formulas are presented. However, some procedures to carry out the proof will be outlined together with some alternative approaches to solve the standard $\mathcal{H}_\infty$ problem and some interpretations of the solutions. We will also indicate how the assumptions in the general case can be relaxed further to accommodate other more complicated problems. More specifically, Section 17.1 presents the solutions to the general $\mathcal{H}_\infty$ problem. Section 17.2 discusses the techniques to transform a general problem to a standard problem which satisfies the assumptions in the last chapter. The problems associated with relaxing the assumptions for the general standard problems and techniques for dealing with them will be considered in Section 17.3. Section 17.4 considers the integral control in the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ theory and Section 17.5 considers how the general $\mathcal{H}_\infty$ solution can be used to solve the $\mathcal{H}_\infty$ filtering problem. Section 17.6 considers an alternative approach to the standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ problems using Youla controller parameterizations, and Section 17.7 gives an 2×2 Hankel-Toeplitz operator interpretations of the $\mathcal{H}_\infty$ solutions presented here and in the last chapter. Finally, the general state feedback $\mathcal{H}_\infty$ control problem and its relations with full information control and differential game problems are discussed in section 17.8 and 17.9.

17.1 General $\mathcal{H}_\infty$ Solutions

Consider the system described by the block diagram



where, as usual, G and K are assumed to be real rational and proper with K constrained to provide internal stability. The controller is said to be admissible if it is real-rational, proper, and stabilizing. Although we are taking everything to be real, the results presented here are still true for the complex case with some obvious modifications. We will again only be interested in characterizing all suboptimal $\mathcal{H}_\infty$ controllers.

The realization of the transfer matrix G is taken to be of the form

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

which is compatible with the dimensions $z(t) \in \mathbb{R}^{p_1}$, $y(t) \in \mathbb{R}^{p_2}$, $w(t) \in \mathbb{R}^{m_1}$, $u(t) \in \mathbb{R}^{m_2}$, and the state $x(t) \in \mathbb{R}^n$. The following assumptions are made:

(A1) (A, B_2) is stabilizable and (C_2, A) is detectable;

(A2) $D_{12} = \begin{bmatrix} 0 \\ I \end{bmatrix}$ and $D_{21} = \begin{bmatrix} 0 & I \end{bmatrix}$;

(A3) $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all ω;

(A4) $\begin{bmatrix} A - j\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ has full row rank for all ω.

Assumption (A1) is necessary for the existence of stabilizing controllers. The assumptions in (A2) mean that the penalty on $z = C_1 x + D_{12} u$ includes a nonsingular, normalized penalty on the control u, that the exogenous signal w includes both plant disturbance and sensor noise, and that the sensor noise weighting is normalized and nonsingular. Relaxation of (A2) leads to singular control problems; see Stroorvogel [1990]. For those problems that have D_{12} full column rank and D_{21} full row rank but do not satisfy assumption (A2), a normalizing procedure is given in the next section so that an equivalent new system will satisfy this assumption.

Assumptions (A3) and (A4) are made for a technical reason: together with (A1) they guarantee that the two Hamiltonian matrices in the corresponding $\mathcal{H}_2$ problem belong to $dom(Ric)$, as we have seen in Chapter 14. It is tempting to suggest that (A3) and (A4) can be dropped, but they are, in some sense, necessary for the methods presented in the last chapter to be applicable. A further discussion of the assumptions and their possible relaxation will be discussed in Section 17.3.

The main result is now stated in terms of the solutions of the X_∞ and Y_∞ Riccati equations together with the "state feedback" and "output injection" matrices F and L.

$$\begin{aligned}
R &:= D_{1\bullet}^* D_{1\bullet} - \begin{bmatrix} \gamma^2 I_{m_1} & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{where} \quad D_{1\bullet} := [D_{11} \ \ D_{12}] \\
\tilde{R} &:= D_{\bullet 1} D_{\bullet 1}^* - \begin{bmatrix} \gamma^2 I_{p_1} & 0 \\ 0 & 0 \end{bmatrix}, \quad \text{where} \quad D_{\bullet 1} := \begin{bmatrix} D_{11} \\ D_{21} \end{bmatrix} \\
H_\infty &:= \begin{bmatrix} A & 0 \\ -C_1^* C_1 & -A^* \end{bmatrix} - \begin{bmatrix} B \\ -C_1^* D_{1\bullet} \end{bmatrix} R^{-1} \begin{bmatrix} D_{1\bullet}^* C_1 & B^* \end{bmatrix} \\
J_\infty &:= \begin{bmatrix} A^* & 0 \\ -B_1 B_1^* & -A \end{bmatrix} - \begin{bmatrix} C^* \\ -B_1 D_{\bullet 1}^* \end{bmatrix} \tilde{R}^{-1} \begin{bmatrix} D_{\bullet 1} B_1^* & C \end{bmatrix}
\end{aligned}$$

$$X_\infty := Ric(H_\infty) \qquad Y_\infty := Ric(J_\infty)$$

$$\begin{aligned}
F &:= \begin{bmatrix} F_{1\infty} \\ F_{2\infty} \end{bmatrix} := -R^{-1} \left[D_{1\bullet}^* C_1 + B^* X_\infty \right] \\
L &:= \begin{bmatrix} L_{1\infty} & L_{2\infty} \end{bmatrix} := -[B_1 D_{\bullet 1}^* + Y_\infty C^*] \tilde{R}^{-1}
\end{aligned}$$

Partition D, $F_{1\infty}$, and $L_{1\infty}$ are as follows:

$$\left[\begin{array}{c|c} & F' \\ \hline L' & D \end{array}\right] = \left[\begin{array}{c|ccc} & F_{11\infty}^* & F_{12\infty}^* & F_{2\infty}^* \\ \hline L_{11\infty}^* & D_{1111} & D_{1112} & 0 \\ L_{12\infty}^* & D_{1121} & D_{1122} & I \\ L_{2\infty}^* & 0 & I & 0 \end{array}\right].$$

Remark 17.1 In the above matrix partitioning, some matrices may not exist depending on whether D_{12} or D_{21} is square. This issue will be discussed further later. For the time being, we shall assume all matrices in the partition exist. ♡

Theorem 17.1 *Suppose G satisfies the assumptions (A1)–(A4).*

(a) There exists an admissible controller $K(s)$ such that $\|\mathcal{F}_\ell(G, K)\|_\infty < \gamma$ (i.e. $\|T_{zw}\|_\infty < \gamma$) if and only if

(i) $\gamma > max(\overline{\sigma}[D_{1111}, D_{1112},], \overline{\sigma}[D_{1111}^*, D_{1121}^*])$;

(ii) $H_\infty \in dom(Ric)$ *with* $X_\infty = Ric(H_\infty) \geq 0$;

(iii) $J_\infty \in dom(Ric)$ *with* $Y_\infty = Ric(J_\infty) \geq 0$;

(iv) $\rho(X_\infty Y_\infty) < \gamma^2$.

(b) Given that the conditions of part (a) are satisfied, then all rational internally stabilizing controllers $K(s)$ *satisfying* $||\mathcal{F}_\ell(G, K)||_\infty < \gamma$ *are given by*

$$K = \mathcal{F}_\ell(M_\infty, Q) \qquad \text{for arbitrary } Q \in \mathcal{RH}_\infty \qquad \text{such that} \quad \|Q\|_\infty < \gamma$$

where

$$M_\infty = \left[\begin{array}{c|cc} \hat{A} & \hat{B}_1 & \hat{B}_2 \\ \hline \hat{C}_1 & \hat{D}_{11} & \hat{D}_{12} \\ \hat{C}_2 & \hat{D}_{21} & 0 \end{array}\right]$$

$$\hat{D}_{11} = -D_{1121}D_{1111}^*(\gamma^2 I - D_{1111}D_{1111}^*)^{-1}D_{1112} - D_{1122},$$

$\hat{D}_{12} \in \mathbb{R}^{m_2 \times m_2}$ *and* $\hat{D}_{21} \in \mathbb{R}^{p_2 \times p_2}$ *are any matrices (e.g. Cholesky factors) satisfying*

$$\hat{D}_{12}\hat{D}_{12}^* = I - D_{1121}(\gamma^2 I - D_{1111}^* D_{1111})^{-1}D_{1121}^*,$$
$$\hat{D}_{21}^*\hat{D}_{21} = I - D_{1112}^*(\gamma^2 I - D_{1111}D_{1111}^*)^{-1}D_{1112},$$

and

$$\hat{B}_2 = Z_\infty(B_2 + L_{12\infty})\hat{D}_{12},$$
$$\hat{C}_2 = -\hat{D}_{21}(C_2 + F_{12\infty}),$$
$$\hat{B}_1 = -Z_\infty L_{2\infty} + \hat{B}_2\hat{D}_{12}^{-1}\hat{D}_{11},$$
$$\hat{C}_1 = F_{2\infty} + \hat{D}_{11}\hat{D}_{21}^{-1}\hat{C}_2,$$
$$\hat{A} = A + BF + \hat{B}_1\hat{D}_{21}^{-1}\hat{C}_2,$$

where

$$Z_\infty = (I - \gamma^{-2}Y_\infty X_\infty)^{-1}.$$

(Note that if $D_{11} = 0$ then the formulae are considerably simplified.)

Some Special Cases:

Case 1: $D_{12} = I$

In this case

1. in part (a), (i) becomes $\gamma > \overline{\sigma}(D_{1121})$.
2. in part (b)

$$\hat{D}_{11} = -D_{1122}$$
$$\hat{D}_{12}\hat{D}_{12}^* = I - \gamma^{-2}D_{1121}D_{1121}^*$$
$$\hat{D}_{21}^*\hat{D}_{21} = I.$$

Case 2: $D_{21} = I$

In this case

1. in part (a), (i) becomes $\gamma > \overline{\sigma}(D_{1112})$.
2. in part (b)

$$\hat{D}_{11} = -D_{1122}$$
$$\hat{D}_{12}\hat{D}_{12}^* = I$$
$$\hat{D}_{21}^*\hat{D}_{21} = I - \gamma^{-2}D_{1112}^*D_{1112}.$$

Case 3: $D_{12} = I$ & $D_{21} = I$

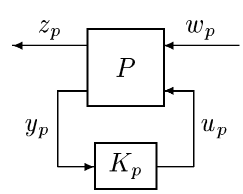
In this case

1. in part (a), (i) drops out.
2. in part (b)

$$\hat{D}_{11} = -D_{1122}$$
$$\hat{D}_{12}\hat{D}_{12}^* = I$$
$$\hat{D}_{21}^*\hat{D}_{21} = I.$$

17.2 Loop Shifting

Let a given problem have the following diagram where $z_p(t) \in \mathbb{R}^{p_1}$, $y_p(t) \in \mathbb{R}^{p_2}$, $w_p(t) \in \mathbb{R}^{m_1}$, and $u_p(t) \in \mathbb{R}^{m_2}$:



The plant P has the following state space realization with D_{p12} *full column rank and* D_{p21} *full row rank:*

$$P(s) = \left[\begin{array}{c|cc} A_p & B_{p1} & B_{p2} \\ \hline C_{p1} & D_{p11} & D_{p12} \\ C_{p2} & D_{p21} & D_{p22} \end{array}\right]$$

The objective is to find all rational proper controllers $K_p(s)$ which stabilize P and $\|\mathcal{F}_\ell(P, K_p)\|_\infty < \gamma$. To solve this problem we first transfer it to the standard one treated in the last section. *Note that the following procedure can also be applied to the* $\mathcal{H}_2$ *problem (except the procedure for the case* $D_{11} \neq 0$*).*

Normalize D_{12} and D_{21}

Perform singular value decompositions to obtain the matrix factorizations

$$D_{p12} = U_p \begin{bmatrix} 0 \\ I \end{bmatrix} R_p, \quad D_{p21} = \tilde{R}_p \begin{bmatrix} 0 & I \end{bmatrix} \tilde{U}_p$$

such that U_p and $\tilde{U}_p$ are square and unitary. Now let

$$z_p = U_p z, \quad w_p = \tilde{U}_p^* w, \quad y_p = \tilde{R}_p y, \quad u_p = R_p u$$

and

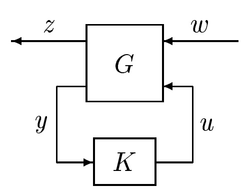
$$K(s) = R_p K_p(s) \tilde{R}_p$$

$$\begin{aligned} G(s) &= \begin{bmatrix} U_p^* & 0 \\ 0 & \tilde{R}_p^{-1} \end{bmatrix} P(s) \begin{bmatrix} \tilde{U}_p^* & 0 \\ 0 & R_p^{-1} \end{bmatrix} \\ &= \left[\begin{array}{c|cc} A_p & B_{p1}\tilde{U}_p^* & B_{p2}R_p^{-1} \\ \hline U_p^* C_{p1} & U_p^* D_{p11}\tilde{U}_p^* & U_p^* D_{p12} R_p^{-1} \\ \tilde{R}_p^{-1} C_{p2} & \tilde{R}_p^{-1} D_{p21}\tilde{U}_p^* & \tilde{R}_p^{-1} D_{p22} R_p^{-1} \end{array}\right] \\ &=: \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right] = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]. \end{aligned}$$

Then

$$D_{12} = \begin{bmatrix} 0 \\ I \end{bmatrix} \quad D_{21} = \begin{bmatrix} 0 & I \end{bmatrix},$$

and the new system is shown below:



Furthermore, $||\mathcal{F}_\ell(P, K_p)||_\alpha = ||U_p \mathcal{F}_\ell(G, K)\tilde{U}_p||_\alpha = \|\mathcal{F}_\ell(G, K)\|_\alpha$ for $\alpha = 2$ or ∞ since U_p and $\tilde{U}_p$ are unitary.

Remove the Assumption $D_{22} = 0$

Suppose $K(s)$ is a controller for G with D_{22} set to zero. Then the controller for $D_{22} \neq 0$ is $K(I + D_{22}K)^{-1}$. Hence there is no loss of generality in assuming $D_{22} = 0$.

Remove the Assumption $D_{11} = 0$

We can even assume that $D_{11} = 0$. In fact, Theorem 17.1 can be shown by first transforming the general problem to the standard problem considered in the last chapter using Theorem 16.2. This transformation is called *loop-shifting*. Before we go into the detailed description of the transformation, let us first consider a simple unitary matrix dilation problem.

Lemma 17.2 *Suppose D is a constant matrix such that $\|D\| < 1$; then*

$$N = \begin{bmatrix} -D & (I - DD^*)^{1/2} \\ (I - D^*D)^{1/2} & D^* \end{bmatrix}$$

is a unitary matrix, i.e., $N^*N = I$.

This result can be easily verified, and the matrix N is called a unitary dilation of D^* (of course, there are many other ways to dilate a matrix to a unitary matrix).

Consider again the feedback system shown at the beginning of this chapter; without loss of generality, we shall assume that the system G has the following realization:

$$G = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & \begin{bmatrix} D_{1111} & D_{1112} \\ D_{1121} & D_{1122} \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} \\ C_2 & \begin{bmatrix} 0 & I \end{bmatrix} & 0 \end{array}\right]$$

with

$$D_{12} = \begin{bmatrix} 0 \\ I \end{bmatrix}, \; D_{21} = \begin{bmatrix} 0 & I \end{bmatrix}.$$

Suppose there is a stabilizing controller K such that $\|\mathcal{F}_\ell(G,K)\|_\infty < 1$. (Suppose we have normalized γ to 1). In the following, we will show how to construct a new system transfer matrix $M(s)$ and a new controller $\tilde{K}$ such that the D_{11} matrix for $M(s)$ is zero, and, furthermore, $\|\mathcal{F}_\ell(G,K)\|_\infty < 1$ if and only if $\left\|\mathcal{F}_\ell(M,\tilde{K})\right\|_\infty < 1$. To begin with, note that $\|\mathcal{F}_\ell(G,K)(\infty)\| < 1$ by the assumption

$$\|\mathcal{F}_\ell(G,K)(\infty)\| = \left\| \begin{matrix} D_{1111} & D_{1112} \\ D_{1121} & D_{1122} + K(\infty) \end{matrix} \right\|.$$

Let

$$D_\infty \in \left\{ X : \left\| \begin{matrix} D_{1111} & D_{1112} \\ D_{1121} & D_{1122} + X \end{matrix} \right\| < 1 \right\}.$$

For example, let $D_\infty = -D_{1122} - D_{1121}(I - D_{1111}^* D_{1111})^{-1} D_{1111}^* D_{1112}$ and define

$$\tilde{D}_{11} := \begin{bmatrix} D_{1111} & D_{1112} \\ D_{1121} & D_{1122} + D_\infty \end{bmatrix};$$

then $\left\|\tilde{D}_{11}\right\| < 1$. Let

$$K(s) = \tilde{K}(s) + D_\infty$$

to get

$$\mathcal{F}_\ell(G, K) = \mathcal{F}_\ell(G, \tilde{K} + D_\infty) = \mathcal{F}_\ell(\tilde{G}, \tilde{K})$$

where

$$\begin{aligned}\tilde{G} &= \left[\begin{array}{c|cc} A + B_2 D_\infty C_2 & B_1 + B_2 D_\infty D_{21} & B_2 \\ \hline C_1 + D_{12} D_\infty C_2 & \begin{bmatrix} D_{1111} & D_{1112} \\ D_{1121} & D_{1122} + D_\infty \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} \\ C_2 & \begin{bmatrix} 0 & I \end{bmatrix} & 0 \end{array}\right] \\ &=: \left[\begin{array}{c|cc} \tilde{A} & \tilde{B}_1 & \tilde{B}_2 \\ \hline \tilde{C}_1 & \tilde{D}_{11} & \tilde{D}_{12} \\ \tilde{C}_2 & \tilde{D}_{21} & 0 \end{array}\right].\end{aligned}$$

Let

$$N = \begin{bmatrix} -\tilde{D}_{11} & (I - \tilde{D}_{11}\tilde{D}_{11}^*)^{1/2} \\ (I - \tilde{D}_{11}^*\tilde{D}_{11})^{1/2} & \tilde{D}_{11}^* \end{bmatrix}$$

and then $N^*N = I$. Furthermore, by Theorem 16.2, we have that K stabilizes G and that $\|\mathcal{F}_\ell(G, K)\|_\infty < 1$ if and only if $\mathcal{F}_\ell(G, K)$ internally stabilizes N and

$$\|\mathcal{F}_\ell(N, \mathcal{F}_\ell(G, K))\|_\infty < 1.$$

Note that

$$\mathcal{F}_\ell(N, \mathcal{F}_\ell(G, K)) = \mathcal{F}_\ell(M, \tilde{K})$$

with

$$M(s) = \left[\begin{array}{c|cc} \tilde{A} + \tilde{B}_1 R_1^{-1} \tilde{D}_{11}^* \tilde{C}_1 & \tilde{B}_1 R_1^{-1/2} & \tilde{B}_2 + \tilde{B}_1 R_1^{-1} \tilde{D}_{11}^* \tilde{D}_{12} \\ \hline \tilde{R}_1^{-1/2} \tilde{C}_1 & 0 & \tilde{R}_1^{-1/2} \tilde{D}_{12} \\ \tilde{C}_2 + \tilde{D}_{21} R_1^{-1} \tilde{D}_{11}^* \tilde{C}_1 & \tilde{D}_{21} R_1^{-1/2} & \tilde{D}_{21} R_1^{-1} \tilde{D}_{11}^* \tilde{D}_{12} \end{array}\right]$$

where $R_1 = I - \tilde{D}_{11}^*\tilde{D}_{11}$ and $\tilde{R}_1 = I - \tilde{D}_{11}\tilde{D}_{11}^*$. In summary, we have the following lemma.

Lemma 17.3 *There is a controller K that internally stabilizes G and $\|\mathcal{F}_\ell(G, K)\|_\infty < 1$ if and only if there is a $\tilde{K}$ that stabilizes M and $\left\|\mathcal{F}_\ell(M, \tilde{K})\right\|_\infty < 1$.*

Corollary 17.4 *Let $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$. Then $\|G(s)\|_\infty < 1$ if and only if $\|D\| < 1$,*

$$M(s) = \left[\begin{array}{c|c} A + B(I - D^*D)^{-1}D^*C & B(I - D^*D)^{-1/2} \\ \hline (I - DD^*)^{-1/2}C & 0 \end{array}\right] \in \mathcal{RH}_\infty,$$

and $\|M(s)\|_\infty < 1$.

17.3 Relaxing Assumptions

In this section, we indicate how the results of section 17.1 can be extended to more general cases.

Relaxing A3 and A4

Suppose that

$$G = \left[\begin{array}{c|cc} 0 & 0 & 1 \\ \hline 0 & 0 & 1 \\ 1 & 1 & 0 \end{array}\right]$$

which violates both A3 and A4 and corresponds to the robust stabilization of an integrator. If the controller $u = -\epsilon x$ where $\epsilon > 0$ is used, then

$$T_{zw} = \frac{-\epsilon s}{s+\epsilon}, \quad \text{with } \|T_{zw}\|_\infty = \epsilon.$$

Hence the norm can be made arbitrarily small as $\epsilon \to 0$, but $\epsilon = 0$ is not admissible since it is not stabilizing. This may be thought of as a case where the $\mathcal{H}_\infty$-optimum is not achieved on the set of admissible controllers. Of course, for this system, $\mathcal{H}_\infty$ *optimal* control is a silly problem, although the suboptimal case is not obviously so.

If one simply drops the requirement that controllers be admissible and removes assumptions A3 and A4, then the formulae presented above will yield $u = 0$ for both the optimal controller and the suboptimal controller with $\Phi = 0$. This illustrates that assumptions A3 and A4 are necessary for the techniques used in the last chapter to be directly applicable. An alternative is to develop a theory which maintains the same notion of admissibility but relaxes A3 and A4. The easiest way to do this would be to pursue the suboptimal case introducing ϵ perturbations so that A3 and A4 are satisfied.

Relaxing A1

If assumption A1 is violated, then it is obvious that no admissible controllers exist. Suppose A1 is relaxed to allow unstabilizable and/or undetectable modes on the $j\omega$ axis and internal stability is also relaxed to also allow closed-loop $j\omega$ axis poles, but A2-A4 is still satisfied. It can be easily shown that under these conditions the closed-loop $\mathcal{H}_\infty$ norm cannot be made finite and, in particular, that the unstabilizable and/or undetectable modes on the $j\omega$ axis must show up as poles in the closed-loop system, see Lemma 16.1.

Violating A1 *and* either or both of A3 and A4

Sensible control problems can be posed which violate A1 *and* either or both of A3 and A4. For example, cases when A has modes at $s = 0$ which are unstabilizable through B_2 and/or undetectable through C_2 arise when an integrator is included in a weight on a disturbance input or an error term. In these cases, either A3 or A4 are also violated, or the closed-loop $\mathcal{H}_\infty$ norm cannot be made finite. In many applications, such problems can be reformulated so that the integrator occurs inside the loop (essentially using the internal model principle) and is hence detectable and stabilizable. We will show this process in the next section.

An alternative approach to such problems which could potentially avoid the problem reformulation would be to pursue the techniques in the last chapter but to relax internal stability to the requirement that all closed-loop modes be in the closed left half plane. Clearly, to have finite $\mathcal{H}_\infty$ norms, these closed-loop modes could not appear as poles in T_{zw}. The formulae given in this chapter will often yield controllers compatible with these assumptions. The user would then have to decide whether closed-loop poles on the imaginary axis were due to weights and hence acceptable or due to the problem being poorly posed, as in the above example.

A third alternative is to again introduce ϵ perturbations so that A1, A3, and A4 are satisfied. Roughly speaking, this would produce sensible answers for sensible problems, but the behavior as $\epsilon \to 0$ could be problematic.

Relaxing A2

In the cases that either D_{12} is not full column rank or that D_{21} is not full row rank, improper controllers can give a bounded $\mathcal{H}_\infty$-norm for T_{zw}, although the controllers will not be admissible by our definition. Such singular filtering and control problems have been well-studied in $\mathcal{H}_2$ theory and many of the same techniques go over to the $\mathcal{H}_\infty$-case (e.g. Willems [1981], Willems, Kitapci, and Silverman [1986], and Hautus and Silverman [1983]). In particular, the structure algorithm of Silverman [1969] could be used to make the terms D_{12} and D_{21} full rank by the introduction of suitable differentiators in the controller. A complete solution to the singular problem can be found in [Stroorvogel, 1990].

17.4 $\mathcal{H}_2$ and $\mathcal{H}_\infty$ Integral Control

It is interesting to note that the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ design frameworks do not in general produce integral control. In this section we show how to introduce integral control into the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ design framework through a simple disturbance rejection problem. We consider a feedback system shown in Figure 17.1. We shall assume that the frequency contents of the disturbance w are effectively modeled by the weighting $W_d \in \mathcal{RH}_\infty$ and the constraints on control signal are limited by an appropriate choice of $W_u \in \mathcal{RH}_\infty$. In order to let the output y track the reference signal r, we require K contain an integral, i.e., $K(s)$ has a pole at $s = 0$. (In general, K is required to have poles on the imaginary axis.)

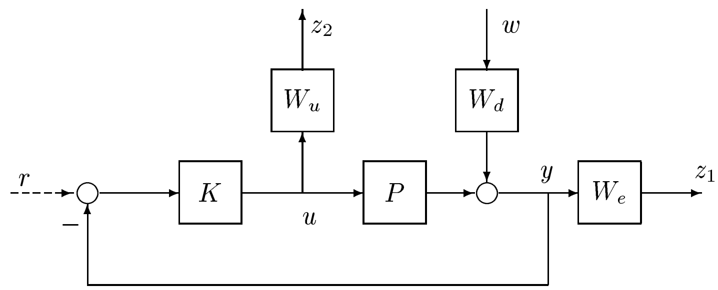


Figure 17.1: A Simple Disturbance Rejection Problem

There are several ways to achieve the integral design. One approach is to introduce an integral in the performance weight W_e. Then the transfer function between w and z_1 is given by

$$z_1 = W_e(I + PK)^{-1} W_d w.$$

Now if the resulting controller K stabilizes the plant P and makes the norm (2-norm or ∞-norm) between w and z_1 finite, then K must have a pole at $s = 0$ which is the zero of the sensitivity function (Assuming W_d has no zeros at $s = 0$). (This follows from the well-known internal model principle.) The problem with this approach is that the $\mathcal{H}_2$ (or $\mathcal{H}_\infty$) control theory presented in this chapter and in the previous chapters can not be applied to this problem formulation directly because the pole $s = 0$ of W_e becomes an uncontrollable pole of the feedback system and the very first assumption for the $\mathcal{H}_2$ (or $\mathcal{H}_\infty$) theory is violated.

However, the obstacles can be overcome by appropriately reformulating the problem. Suppose W_e can be factorized as follows

$$W_e = \tilde{W}_e(s) M(s)$$

where $M(s)$ is proper, containing all the imaginary axis poles of W_e, and $M^{-1}(s) \in \mathcal{RH}_\infty$, $\tilde{W}_e(s)$ is stable and minimum phase. Now suppose there exists a controller $K(s)$ which contains the same imaginary axis poles that achieves the performance. Then without loss of generality, K can be factorized as

$$K(s) = -\hat{K}(s) M(s)$$

where there is no unstable pole/zero cancelation in forming the product $\hat{K}(s)M(s)$. Now the problem can be reformulated as in Figure 17.2. Figure 17.2 can be put in the general LFT framework as in Figure 17.3.

Let $\tilde{W}_e, W_u, W_d, M$, and P have the following stabilizable and detectable state space

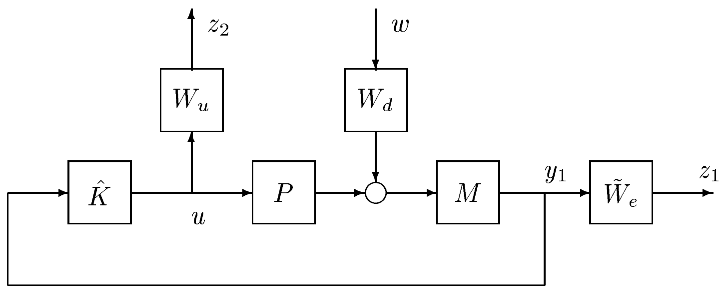


Figure 17.2: Disturbance Rejection with Imaginary Axis Poles

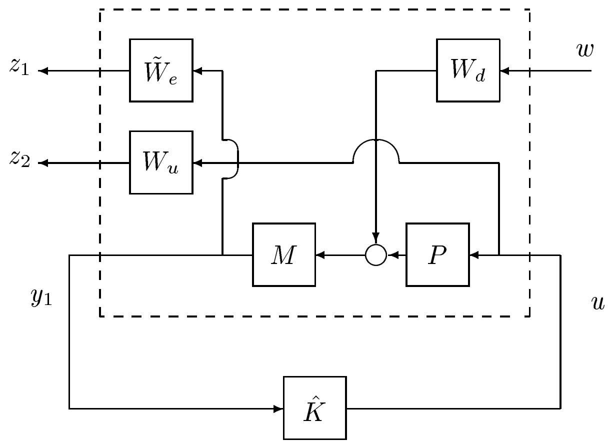


Figure 17.3: LFT Framework for the Disturbance Rejection Problem

realizations:

$$\tilde{W}_e = \left[\begin{array}{c|c} A_e & B_e \\ \hline C_e & D_e \end{array}\right], \quad W_u = \left[\begin{array}{c|c} A_u & B_u \\ \hline C_u & D_u \end{array}\right], \quad W_d = \left[\begin{array}{c|c} A_d & B_d \\ \hline C_d & D_d \end{array}\right]$$

$$M = \left[\begin{array}{c|c} A_m & B_m \\ \hline C_m & D_m \end{array}\right], \quad P = \left[\begin{array}{c|c} A_p & B_p \\ \hline C_p & D_p \end{array}\right].$$

Then a realization for the generalized system is given by

$$G(s) = \begin{bmatrix} \begin{bmatrix} \tilde{W}_e M W_d \\ 0 \end{bmatrix} & \begin{bmatrix} \tilde{W}_e M P \\ W_u \end{bmatrix} \\ M W_d & M P \end{bmatrix}$$

$$= \left[\begin{array}{ccccc|cc} A_e & 0 & B_eC_m & B_eD_mC_d & B_eD_mC_p & B_eD_mD_d & B_eD_mD_p \\ 0 & A_u & 0 & 0 & 0 & 0 & B_u \\ 0 & 0 & A_m & B_mC_d & B_mC_p & B_mD_d & B_mD_p \\ 0 & 0 & 0 & A_d & 0 & B_d & 0 \\ 0 & 0 & 0 & 0 & A_p & 0 & B_p \\ \hline C_e & 0 & D_eC_m & D_eD_mC_d & D_eD_mC_p & D_eD_mD_d & D_eD_mD_p \\ 0 & C_u & 0 & 0 & 0 & 0 & D_u \\ 0 & 0 & C_m & D_mC_d & D_mC_p & D_mD_d & D_mD_p \end{array}\right].$$

We shall illustrate the above design through a simple numerical example. Let

$$P = \frac{s-2}{(s+1)(s-3)} = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ 3 & 2 & 1 \\ \hline -2 & 1 & 0 \end{array}\right], \quad W_d = 1,$$

$$W_u = \frac{s+10}{s+100} = \left[\begin{array}{c|c} -100 & -90 \\ \hline 1 & 1 \end{array}\right], \quad W_e = \frac{1}{s}.$$

Then we can choose without loss of generality that

$$M = \frac{s+\alpha}{s}, \quad \tilde{W}_e = \frac{1}{s+\alpha}, \quad \alpha > 0.$$

This gives the following generalized system

$$G(s) = \left[\begin{array}{ccccc|cc} -\alpha & 0 & 1 & -2 & 1 & 1 & 0 \\ 0 & -100 & 0 & 0 & 0 & 0 & -90 \\ 0 & 0 & 0 & -2\alpha & \alpha & \alpha & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 3 & 2 & 0 & 1 \\ \hline 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & -2 & 1 & 1 & 0 \end{array}\right]$$

and suboptimal $\mathcal{H}_\infty$ controller $\hat{K}_\infty$ for each α can be computed easily as

$$\hat{K}_\infty = \frac{-2060381.4(s+1)(s+\alpha)(s+100)(s-0.1557)}{(s+\alpha)^2(s+32.17)(s+262343)(s-19.89)}$$

which gives the closed-loop ∞ norm 7.854. Hence the controller $K_\infty = -\hat{K}_\infty(s)M(s)$ is given by

$$K_\infty(s) = \frac{2060381.4(s+1)(s+100)(s-0.1557)}{s(s+32.17)(s+262343)(s-19.89)} \approx \frac{7.85(s+1)(s+100)(s-0.1557)}{s(s+32.17)(s-19.89)}$$

which is independent of α as expected. Similarly, we can solve an optimal $\mathcal{H}_2$ controller

$$\hat{K}_2 = \frac{-43.487(s+1)(s+\alpha)(s+100)(s-0.069)}{(s+\alpha)^2(s^2+30.94s+411.81)(s-7.964)}$$

and

$$K_2(s) = -\hat{K}_2(s)M(s) = \frac{43.487(s+1)(s+100)(s-0.069)}{s(s^2+30.94s+411.81)(s-7.964)}.$$

An approximate integral control can also be achieved without going through the above process by letting

$$W_e = \tilde{W}_e = \frac{1}{s+\epsilon}, \quad M(s) = 1$$

for a sufficiently small $\epsilon > 0$. For example, a controller for $\epsilon = 0.001$ is given by

$$K_\infty = \frac{316880(s+1)(s+100)(s-0.1545)}{(s+0.001)(s+32)(s+40370)(s-20)} \approx \frac{7.85(s+1)(s+100)(s-0.1545)}{s(s+32)(s-20)}$$

which gives the closed-loop $\mathcal{H}_\infty$ norm of 7.85. Similarly, an approximate $\mathcal{H}_2$ integral controller is obtained as

$$K_2 = \frac{43.47(s+1)(s+100)(s-0.0679)}{(s+0.001)(s^2+30.93s+411.7)(s-7.9718)}.$$

17.5 $\mathcal{H}_\infty$ Filtering

In this section we show how the filtering problem can be solved using the $\mathcal{H}_\infty$ theory developed early. Suppose a dynamic system is described by the following equations

$$\begin{aligned} \dot{x} &= Ax + B_1 w(t), \qquad x(0) = 0 & (17.1)\\ y &= C_2 x + D_{21} w(t) & (17.2)\\ z &= C_1 x + D_{11} w(t) & (17.3) \end{aligned}$$

The filtering problem is to find an estimate $\hat{z}$ of z in some sense using the measurement of y. The restriction on the filtering problem is that the filter has to be causal so that it can be realized, i.e., $\hat{z}$ has to be generated by a causal system acting on the measurements. We will further restrict our filter to be *unbiased*, i.e., given $T > 0$ the estimate $\hat{z}(t) = 0\ \forall t \in [0, T]$ if $y(t) = 0,\ \forall t \in [0, T]$. Now we can state our $\mathcal{H}_\infty$ filtering problem.

> $\mathcal{H}_\infty$ **Filtering**: Given a $\gamma > 0$, find a causal filter $F(s) \in \mathcal{RH}_\infty$ if it exists such that
>
> $$J := \sup_{w \in \mathcal{L}_2[0,\infty)} \frac{\|z - \hat{z}\|_2^2}{\|w\|_2^2} < \gamma^2$$
>
> with $\hat{z} = F(s)y$.

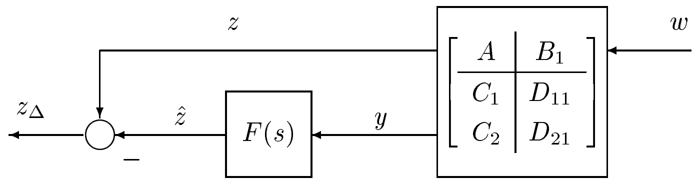
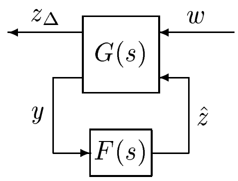


Figure 17.4: Filtering Problem Formulation

A diagram for the filtering problem is shown in Figure 17.4.

The above filtering problem can also be formulated in an LFT framework: given a system shown below



$$G(s) = \left[\begin{array}{c|cc} A & B_1 & 0 \\ \hline C_1 & D_{11} & -I \\ C_2 & D_{21} & 0 \end{array}\right]$$

find a filter $F(s) \in \mathcal{RH}_\infty$ such that

$$\sup_{w \in \mathcal{L}_2} \frac{\|z_\Delta\|_2^2}{\|w\|_2^2} < \gamma^2. \tag{17.4}$$

Hence the filtering problem can be regarded as a special $\mathcal{H}_\infty$ problem. However, comparing with control problems there is no internal stability requirement in the filtering problem. Hence the solution to the above filtering problem can be obtained from the $\mathcal{H}_\infty$ solution in the previous sections by setting $B_2 = 0$ and dropping the internal stability requirement.

Theorem 17.5 *Suppose (C_2, A) is detectable and*

$$\begin{bmatrix} A - j\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$$

has full row rank for all ω. Let D_{21} be normalized and D_{11} partitioned conformably as

$$\begin{bmatrix} D_{11} \\ D_{21} \end{bmatrix} = \begin{bmatrix} D_{111} & D_{112} \\ 0 & I \end{bmatrix}.$$

Then there exists a causal filter $F(s) \in \mathcal{RH}_\infty$ such that $J < \gamma^2$ if and only if $\overline{\sigma}(D_{111}) < \gamma$ and $J_\infty \in dom(Ric)$ with $Y_\infty = Ric(J_\infty) \geq 0$ where

$$\tilde{R} := \begin{bmatrix} D_{11} \\ D_{21} \end{bmatrix} \begin{bmatrix} D_{11} \\ D_{21} \end{bmatrix}^* - \begin{bmatrix} \gamma^2 I & 0 \\ 0 & 0 \end{bmatrix}$$

$$J_\infty := \begin{bmatrix} A^* & 0 \\ -B_1B_1^* & -A \end{bmatrix} - \begin{bmatrix} C_1^* & C_2^* \\ -B_1D_{11}^* & -B_1D_{21}^* \end{bmatrix} \tilde{R}^{-1} \begin{bmatrix} D_{11}B_1^* & C_1 \\ D_{21}B_1^* & C_2 \end{bmatrix}.$$

Moreover, if the above conditions are satisfied, then a rational causal filter $F(s)$ satisfying $J < \gamma^2$ is given by

$$\hat{z} = F(s)y = \left[\begin{array}{c|c} A + L_{2\infty}C_2 + L_{1\infty}D_{112}C_2 & -L_{2\infty} - L_{1\infty}D_{112} \\ \hline C_1 - D_{112}C_2 & D_{112} \end{array}\right] y$$

where

$$\begin{bmatrix} L_{1\infty} & L_{2\infty} \end{bmatrix} := -\begin{bmatrix} B_1D_{11}^* + Y_\infty C_1^* & B_1D_{21}^* + Y_\infty C_2^* \end{bmatrix} \tilde{R}^{-1}.$$

In the case where $D_{11} = 0$ and $B_1D_{21}^* = 0$ the filter becomes much simpler:

$$\hat{z} = \left[\begin{array}{c|c} A - Y_\infty C_2^*C_2 & Y_\infty C_2^* \\ \hline C_1 & 0 \end{array}\right] y$$

where Y_∞ is the stabilizing solution to

$$Y_\infty A^* + AY_\infty + Y_\infty(\gamma^{-2}C_1^*C_1 - C_2^*C_2)Y_\infty + B_1B_1^* = 0.$$

17.6 Youla Parameterization Approach*

In this section, we shall briefly outline an alternative approach to solving the standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems using the Q parameterization (Youla parameterization) approach. This approach was the main focus in the early 1980's and is still very useful in solving many interesting problems. We will see that this approach may suggest additional interpretations for the results presented in the last chapter and applies to both $\mathcal{H}_2$ and $\mathcal{H}_\infty$ problems. The $\mathcal{H}_2$ problem is very simple and involves a projection. While the $\mathcal{H}_\infty$ problem is much more difficult, it bears some similarities to the constant matrix dilation problem but with the restriction of internal stability. Nevertheless, we have built enough tools to give a fairly complete solution using this Q parameterization approach.

Assume again that the G has the realization

$$G = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

with the same assumptions as before. But for convenience, we will allow the assumption (A2) to be relaxed to the following

(A2′) D_{12} is full column rank with $\begin{bmatrix} D_{12} & D_\perp \end{bmatrix}$ unitary, and D_{21} is full row rank with $\begin{bmatrix} D_{21} \\ \tilde{D}_\perp \end{bmatrix}$ unitary.

Next, we will outline the steps required to solve the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems. Because of the similarity between the $\mathcal{H}_2$ and $\mathcal{H}_\infty$ problems, they are developed in parallel below.

Parameterization: Recall that all controllers stabilizing the plant G can be expressed as

$$K = \mathcal{F}_\ell(M_2, Q), \ I + D_{22}Q(\infty) \text{ invertible}$$

with $Q \in \mathcal{RH}_\infty$ and

$$M_2 = \left[\begin{array}{c|cc} A + B_2F_2 + L_2C_2 + L_2D_{22}F_2 & -L_2 & B_2 + L_2D_{22} \\ \hline F_2 & 0 & I \\ -(C_2 + D_{22}F_2) & I & -D_{22} \end{array}\right]$$

where

$$\begin{aligned} F_2 &= -(B_2^*X_2 + D_{12}^*C_1) \\ L_2 &= -(Y_2C_2^* + B_1D_{21}^*) \\ X_2 &= \operatorname{Ric}\begin{bmatrix} A - B_2D_{12}^*C_1 & -B_2B_2^* \\ -C_1^*D_\perp D_\perp^* C_1 & -(A - B_2D_{12}^*C_1)^* \end{bmatrix} \geq 0 \\ Y_2 &= \operatorname{Ric}\begin{bmatrix} (A - B_1D_{21}^*C_2)^* & -C_2^*C_2 \\ -B_1\tilde{D}_\perp^*\tilde{D}_\perp B_1^* & -(A - B_1D_{21}^*C_2) \end{bmatrix} \geq 0. \end{aligned}$$

Then the closed-loop transfer matrix from w to z can be written as

$$\mathcal{F}_\ell(G, K) = T_o + UQV, \qquad I + D_{22}Q(\infty) \text{ invertible}$$

where

$$T_o = \left[\begin{array}{cc|c} A_{F_2} & -B_2F_2 & B_1 \\ 0 & A_{L_2} & B_{1L_2} \\ \hline C_{1F_2} & -D_{12}F_2 & D_{11} \end{array}\right] \in \mathcal{RH}_\infty$$

$$U = \left[\begin{array}{c|c} A_{F_2} & B_2 \\ \hline C_{1F_2} & D_{12} \end{array}\right], \ V = \left[\begin{array}{c|c} A_{L_2} & B_{1L_2} \\ \hline C_2 & D_{21} \end{array}\right]$$

and

$$A_{F_2} := A + B_2F_2, \quad A_{L_2} := A + L_2C_2$$
$$C_{1F_2} = C_1 + D_{12}F_2, \quad B_{1L_2} = B_1 + L_2D_{21}.$$

It is easy to show that U is an inner, $U^\sim U = I$, and V is a co-inner, $VV^\sim = I$.

Unitary Invariant: There exist $U_\perp$ and $V_\perp$ such that $\begin{bmatrix} U & U_\perp \end{bmatrix}$ and $\begin{bmatrix} V \\ V_\perp \end{bmatrix}$ are square and inner:

$$U_\perp = \left[\begin{array}{c|c} A_{F_2} & -X_2^+ C_1^* D_\perp \\ \hline C_{1F_2} & D_\perp \end{array}\right], \quad V_\perp = \left[\begin{array}{c|c} A_{L_2} & B_{1L_2} \\ \hline -\tilde{D}_\perp B_1^* Y_2^+ & \tilde{D}_\perp \end{array}\right].$$

Since the multiplication of square inner matrices do not change $\mathcal{H}_2$ and $\mathcal{H}_\infty$ norms, we have for $\alpha = 2$ or ∞

$$\begin{aligned} \|\mathcal{F}_\ell(P,K)\|_\alpha &= \|T_o + UQV\|_\alpha \\ &= \left\| \begin{bmatrix} U & U_\perp \end{bmatrix}^\sim (T_o + UQV) \begin{bmatrix} V \\ V_\perp \end{bmatrix}^\sim \right\|_\alpha \\ &= \left\| \begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\alpha \end{aligned}$$

where

$$R = \begin{bmatrix} U^\sim \\ U_\perp^\sim \end{bmatrix} T_o \begin{bmatrix} V^\sim & V_\perp^\sim \end{bmatrix}$$

with the obvious partitioning. It can be shown through some long and tedious algebra that R is antistable and has the following representation:

$$R = \left[\begin{array}{cc|cc} -A_{F_2}^* & EB_{1L2}^* & -ED_{21}^* & -E\tilde{D}_\perp^* \\ 0 & -A_{L_2}^* & C_2^* & -Y_2^+ B_1 \tilde{D}_\perp^* \\ \hline B_2^* & F_2 Y_2 - D_{12}^* D_{11} B_{1L_2}^* & D_{12}^* D_{11} D_{21}^* & D_{12}^* D_{11} \tilde{D}_\perp^* \\ -D_\perp^* C_1 X_2^+ & -D_\perp^* D_{11} B_{1L_2}^* & D_\perp^* D_{11} D_{21}^* & D_\perp^* D_{11} \tilde{D}_\perp^* \end{array}\right]$$

where $E := X_2 B_1 + C_{1F_2}^* D_{11}$.

Projection/Dilation: At this point the $\alpha = 2$ and $\alpha = \infty$ cases have to be treated separately.

$\mathcal{H}_2$ **case:** In this case, we will assume $D_\perp^* D_{11} = 0$ and $D_{11}\tilde{D}_\perp^* = 0$; otherwise, the 2-norm of the closed loop transfer matrix will be unbounded since

$$R(\infty) = \begin{bmatrix} D_{12}^* D_{11} D_{21}^* & D_{12}^* D_{11} \tilde{D}_\perp^* \\ D_\perp^* D_{11} D_{21}^* & D_\perp^* D_{11} \tilde{D}_\perp^* \end{bmatrix}.$$

Now from the definition of 2-norm, the problem can be rewritten as

$$\|\mathcal{F}_\ell(G,K)\|_2^2 = \left\| \begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_2^2$$

$$= \quad \|R_{11}+Q\|_2^2 + \left\| \begin{bmatrix} 0 & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_2^2 .$$

Furthermore,

$$\|R_{11}+Q\|_2^2 = \|R_{11} - D_{12}^* D_{11} D_{21}^*\|_2^2 + \|Q + D_{12}^* D_{11} D_{21}^*\|_2^2$$

since $(R_{11} - D_{12}^* D_{11} D_{21}^*) \in \mathcal{H}_2^\perp$ and $(Q + D_{12}^* D_{11} D_{21}^*) \in \mathcal{H}_\infty$. In fact, $(Q + D_{12}^* D_{11} D_{21}^*)$ has to be in $\mathcal{H}_2$ to guarantee that the 2-norm be finite. Hence the unique optimal solution is given by a projection to $\mathcal{H}_\infty$:

$$Q_{opt} = [-R_{11}]_+ = -D_{12}^* D_{11} D_{21}^* .$$

The optimal controller is given by

$$K_{opt} = \mathcal{F}_\ell(M_2, Q_{opt}).$$

In particular, if $D_{11} = 0$ then $K_{opt} = \mathcal{F}_\ell(M_2, 0)$.

$\mathcal{H}_\infty$ **case:** Recall the definition of the ∞-norm:

$$\left\| \begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty = \sup_\omega \overline{\sigma} \left(\begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right) (j\omega).$$

Consider the minimization problem with respect to Q frequency by frequency; it becomes a constant matrix dilation problem if no causality restrictions are imposed on Q. The $\mathcal{H}_\infty$ optimal control follows this idea. However, it should be noted that since Q is restricted to be an $\mathcal{RH}_\infty$ matrix, the term $R_{11}+Q$ cannot be made into an arbitrary matrix. Therefore, the problem will be significantly different from the constant matrix case, and, generically,

$$\min_{Q \in \mathcal{H}_\infty} \left\| \begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty > \gamma_0$$

where

$$\gamma_0 := \max \left\{ \left\| \begin{bmatrix} R_{21} & R_{22} \end{bmatrix} \right\|_\infty , \left\| \begin{bmatrix} R_{12} \\ R_{22} \end{bmatrix} \right\|_\infty \right\} .$$

It is convenient to use the characterization in Corollary 2.23. Recall that

$$\left\| \begin{bmatrix} R_{11}+Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty \le \gamma \ (< \gamma), \text{ for } \gamma > \gamma_0$$

iff

$$\left\| (I - WW^\sim)^{-1/2} (Q + R_{11} + W R_{22}^\sim Z)(I - Z^\sim Z)^{-1/2} \right\|_\infty \le \gamma \ (< \gamma)$$

where

$$\begin{aligned} W &= R_{12}(\gamma^2 I - R_{22}^{\sim} R_{22})^{-1/2} \\ Z &= (\gamma^2 I - R_{22} R_{22}^{\sim})^{-1/2} R_{21}. \end{aligned}$$

The key here is to find the spectral factors $(I - WW^{\sim})^{1/2}$ and $(I - Z^{\sim}Z)^{1/2}$ with stable inverses such that if $M = (I - WW^{\sim})^{1/2}$ and $N = (I - Z^{\sim}Z)^{1/2}$, then $M, M^{-1}, N, N^{-1} \in \mathcal{H}_\infty$; $MM^{\sim} = I - WW^{\sim}$, and $N^{\sim}N = (I - Z^{\sim}Z)$. Now let

$$\hat{Q} := M^{-1}QN^{-1}, \quad G := M^{-1}(R_{11} + WR_{22}^{\sim}Z)N^{-1};$$

then the problem is reduced to finding $\hat{Q} \in \mathcal{H}_\infty$ such that

$$\left\| G + \hat{Q} \right\|_\infty \leq \gamma \ (< \gamma). \tag{17.5}$$

Note that $Q = M\hat{Q}N \in \mathcal{H}_\infty$ iff $\hat{Q} \in \mathcal{H}_\infty$.

The final step in the $\mathcal{H}_\infty$ problem involves solving (17.5) for $\hat{Q} \in \mathcal{H}_\infty$. This is a standard Nehari problem and is solved in Chapter 8. The optimal control law will be given by

$$K_{opt} = \mathcal{F}_\ell(J, Q_{opt}).$$

17.7 Connections

This section considers the connections between the Youla parameterization approach and the current state space approach discussed in the last chapter. The key is Lemma 15.7 and its connection with the formulae given in the last section.

To see how Lemma 15.7 might be used in the last section to prove Theorems 16.4 and 16.5 or Theorem 17.1, we begin with G having state dimension n. For simplicity, we assume further that $D_{11} = 0$. Then from the last section, we have

$$\begin{aligned} \|T_{zw}\|_\infty &= \left\| R + \begin{bmatrix} Q & 0 \\ 0 & 0 \end{bmatrix} \right\|_\infty \\ &= \left\| \begin{bmatrix} R_{11} + Q & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \right\|_\infty \\ &= \left\| R^{\sim} + \begin{bmatrix} Q^{\sim} & 0 \\ 0 & 0 \end{bmatrix} \right\|_\infty \end{aligned}$$

where R has state dimension $2n$. Now define

$$\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} := R^{\sim}.$$

Then

$$\begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} = \left[\begin{array}{cc|cc} A_{F_2} & 0 & B_2 & -X_2^+ C_1^* D_\perp \\ -B_{1L2}B_1^* X_2 & A_{L_2} & Y_2 F_2^* & 0 \\ \hline D_{21}B_1^* X_2 & -C_2 & 0 & 0 \\ \tilde{D}_\perp B_1^* X_2 & \tilde{D}_\perp B_1^* Y_2^+ & 0 & 0 \end{array}\right] \in \mathcal{RH}_\infty, \tag{17.6}$$

and the $\mathcal{H}_\infty$ problem becomes to find an antistable $Q^\sim$ such that

$$\left\| \begin{bmatrix} N_{11} + Q^\sim & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \right\|_\infty < \gamma.$$

Let $w = \begin{bmatrix} w_1 \\ w_2 \end{bmatrix}$ and note that

$$\begin{aligned} \left\| \begin{bmatrix} N_{11} + Q^\sim & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \right\|_\infty^2 &:= \sup_{w \in \mathcal{BL}_2} \left\| \begin{bmatrix} N_{11} + Q^\sim & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\|_2^2 \\ &= \sup_{\{w \in \mathcal{BL}_2\} \cap \{w_1 \in \mathcal{H}_2^\perp\}} \left\| \begin{bmatrix} N_{11} + Q^\sim & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\|_2^2 . \end{aligned}$$

The right hand side of the above equation can be written as

$$\sup_{\{w \in \mathcal{BL}_2\} \cap \{w_1 \in \mathcal{H}_2^\perp\}} \left\{ \left\| \begin{bmatrix} P_+ & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\|_2^2 + \| Q^\sim w_1 + P_-(N_{11}w_1 + N_{12}w_2) \|_2^2 \right\}.$$

It is clear that the last term can be made zero for an appropriate choice of an antistable $Q^\sim$. Hence the $\mathcal{H}_\infty$ problem has a solution if and only if

$$\sup_{\{w \in \mathcal{BL}_2\} \cap \{w_1 \in \mathcal{H}_2^\perp\}} \left\| \begin{bmatrix} P_+ & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix} \begin{bmatrix} w_1 \\ w_2 \end{bmatrix} \right\|_2 < \gamma.$$

But this is exactly the same operator as the one in Lemma 15.7. Lemma 15.7 may be applied to derive the solvability conditions and some additional arguments to construct a $Q \in \mathcal{RH}_\infty$ from X and Y such that $\|T_{zw}(Q)\|_\infty < \gamma$. In fact, it turns out that X in Lemma 15.7 for N is exactly W in the FI proof or in the differential game problem.

The final step is to obtain the controller from M_2 and Q. Since M_2 has state dimension n and Q has $2n$, the apparent state dimension of K is $3n$, but some tedious state space manipulations produce cancelations resulting in the n dimensional controller formulae in Theorems 16.4 and 16.5. This approach is exactly the procedure used in Doyle [1984] and Francis [1987] with Lemma 15.7 used to solve the general distance problem. Although this approach is conceptually straightforward and was, in fact, used to obtain the first proof of the current state space results in this chapter, it seems unnecessarily cumbersome and indirect.

17.8 State Feedback and Differential Game

It has been shown in Chapters 15 and 16 that a (central) suboptimal full information $\mathcal{H}_\infty$ control law is actually a pure state feedback if $D_{11} = 0$. However, this is not true in general if $D_{11} \neq 0$, as will be shown below. Nevertheless, the state feedback $\mathcal{H}_\infty$ control is itself a very interesting problem and deserves special treatment. This section and the section to follow are devoted to the study of this state feedback $\mathcal{H}_\infty$ control problem and its connections with full information control and differential game.

Consider a dynamical system

$$\dot{x} = Ax + B_1 w + B_2 u \tag{17.7}$$

$$z = C_1 x + D_{11} w + D_{12} u \tag{17.8}$$

where $z(t) \in \mathbb{R}^{p_1}$, $y(t) \in \mathbb{R}^{p_2}$, $w(t) \in \mathbb{R}^{m_1}$, $u(t) \in \mathbb{R}^{m_2}$, and $x(t) \in \mathbb{R}^n$. The following assumptions are made:

(AS1) (A, B_2) is stabilizable;

(AS2) There is a matrix $D_\perp$ such that $\begin{bmatrix} D_{12} & D_\perp \end{bmatrix}$ is unitary;

(AS3) $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all ω.

We are interested in the following two related quadratic min-max problems: given $\gamma > 0$, check if

$$\sup_{w \in B\mathcal{L}_2[0,\infty)} \min_{u \in \mathcal{L}_2[0,\infty)} \|z\|_2 < \gamma$$

and

$$\min_{u \in \mathcal{L}_2[0,\infty)} \sup_{w \in B\mathcal{L}_2[0,\infty)} \|z\|_2 < \gamma.$$

The first problem can be regarded as a full information control problem since the control signal u can be a function of the disturbance w and the system state x. On the other hand, the optimal control signal in the second problem cannot depend on the disturbance w (the worst disturbance w can be a function of u and x). In fact, it will be shown that the control signal in the latter problem depends only on the system state; hence, it is equivalent to a state feedback control.

Theorem 17.6 *Let $\gamma > 0$ be given and define*

$$R := D_{1\bullet}^* D_{1\bullet} - \begin{bmatrix} \gamma^2 I_{m_1} & 0 \\ 0 & 0 \end{bmatrix}, \quad \textit{where} \quad D_{1\bullet} := [D_{11} \;\; D_{12}]$$

$$H_\infty := \begin{bmatrix} A & 0 \\ -C_1^* C_1 & -A^* \end{bmatrix} - \begin{bmatrix} B \\ -C_1^* D_{1\bullet} \end{bmatrix} R^{-1} \begin{bmatrix} D_{1\bullet}^* C_1 & B^* \end{bmatrix}$$

where $B := \begin{bmatrix} B_1 & B_2 \end{bmatrix}$.

(a) $\sup_{w \in B\mathcal{L}_2[0,\infty)} \min_{u \in \mathcal{L}_2[0,\infty)} \|z\|_2 < \gamma$ *if and only if*

$$\overline{\sigma}(D_\perp^* D_{11}) < \gamma, \quad H_\infty \in dom(Ric), \quad X_\infty = Ric(H_\infty) \geq 0.$$

Moreover, an optimal u is given by

$$u = -D_{12}^* D_{11} w + \begin{bmatrix} D_{12}^* D_{11} & I \end{bmatrix} Fx,$$

and a worst w_{fiworst} is given by

$$w_{\text{fiworst}} = F_{1\infty} x$$

where

$$F := \begin{bmatrix} F_{1\infty} \\ F_{2\infty} \end{bmatrix} := -R^{-1} \left[D_{1\bullet}^* C_1 + B^* X_\infty \right].$$

(b) $\min_{u \in \mathcal{L}_2[0,\infty)} \sup_{w \in B\mathcal{L}_2[0,\infty)} \|z\|_2 < \gamma$ *if and only if*

$$\overline{\sigma}(D_{11}) < \gamma, \quad H_\infty \in dom(Ric), \quad X_\infty = Ric(H_\infty \geq 0.$$

Moreover, an optimal u is given by

$$u = F_{2\infty} x = -D_{12}^* D_{11} w_{\text{fiworst}} + \begin{bmatrix} D_{12}^* D_{11} & I \end{bmatrix} Fx,$$

and the worst w_{sfworst} is given by

$$w_{\text{sfworst}} = (\gamma^2 I - D_{11}^* D_{11})^{-1} \left\{ (D_{11}^* C_1 + B_1^* X_\infty) x + D_{11}^* D_{12} u \right\}.$$

Proof. The condition for part (a) can be shown in the same way as in Chapter 15 and is, in fact, the solution to the FI problem for the general case. We now prove the condition for part (b). It is not hard to see that $\overline{\sigma}(D_{11}) < \gamma$ is necessary since control u cannot feed back w directly, so the D_{11} term cannot be (partially) eliminated as in the FI problem. The conditions $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty \geq 0$ can be easily seen as necessary since

$$\sup_{w \in B\mathcal{L}_2[0,\infty)} \min_{u \in \mathcal{L}_2[0,\infty)} \|z\|_2 \leq \min_{u \in \mathcal{L}_2[0,\infty)} \sup_{w \in B\mathcal{L}_2[0,\infty)} \|z\|_2.$$

It is easy to verify directly that the optimal control and the worst disturbance can be chosen in the given form. □

Define

$$\begin{aligned} R_0 &= I - D_{11}^* D_\perp D_\perp^* D_{11}/\gamma^2 \\ \tilde{R}_0 &= I + D_{12}^* D_{11}(\gamma^2 I - D_{11}^* D_{11})^{-1} D_{11}^* D_{12} \\ \hat{R}_0 &= I - D_{11}^* D_{11}/\gamma^2. \end{aligned}$$

Then it is easy to show that

$$\begin{aligned} \|z\|^2 - \gamma^2 \|w\|^2 + \frac{d}{dt}(x^* X x) &= \left\| u + D_{12}^* D_{11} w - \left[\begin{array}{cc} D_{12}^* D_{11} & I \end{array} \right] F x \right\|^2 \\ & \quad -\gamma^2 \left\| R_0^{1/2}(w - F_{1\infty} x) \right\|^2 \end{aligned}$$

if conditions in (a) are satisfied. On the other hand, we have

$$\|z\|^2 - \gamma^2 \|w\|^2 + \frac{d}{dt}(x^* X x) = \left\| \tilde{R}_0^{1/2}(u - F_{2\infty} x) \right\|^2 - \gamma^2 \left\| \hat{R}_0^{1/2}(w - w_{\text{sfworst}}) \right\|^2$$

if conditions in (b) are satisfied. Integrating both equations from $t = 0$ to ∞ with $x(0) = x(\infty) = 0$ gives

$$\|z\|_2^2 - \gamma^2 \|w\|_2^2 = \left\| u + D_{12}^* D_{11} w - \left[\begin{array}{cc} D_{12}^* D_{11} & I \end{array} \right] F x \right\|_2^2 - \gamma^2 \left\| R_0^{1/2}(w - F_{1\infty} x) \right\|_2^2$$

if conditions in (a) are satisfied, and

$$\|z\|_2^2 - \gamma^2 \|w\|_2^2 = \left\| \tilde{R}_0^{1/2}(u - F_{2\infty} x) \right\|_2^2 - \gamma^2 \left\| \hat{R}_0^{1/2}(w - w_{\text{sfworst}}) \right\|_2^2$$

if conditions in (b) are satisfied. These relations suggest that an optimal control law and a worst disturbance for problem (a) are

$$u = -D_{12}^* D_{11} w + \left[\begin{array}{cc} D_{12}^* D_{11} & I \end{array} \right] F x, \quad w = F_{1\infty} x$$

and that an optimal control law and a worst disturbance for problem (b) are

$$u = F_{2\infty} x \quad w = w_{\text{sfworst}}.$$

Moreover, if problem (b) has a solution for a given γ, then the corresponding differential game problem

$$\min_{u\in\mathcal{L}_2[0,\infty)} \sup_{w\in\mathcal{L}_2[0,\infty)} \left\{\|z\|_2^2 - \gamma^2\|w\|_2^2\right\}$$

has a saddle point, i.e,

$$\sup_{w\in\mathcal{L}_2[0,\infty)} \min_{u\in\mathcal{L}_2[0,\infty)} \left\{\|z\|_2^2 - \gamma^2\|w\|_2^2\right\} = \min_{u\in\mathcal{L}_2[0,\infty)} \sup_{w\in\mathcal{L}_2[0,\infty)} \left\{\|z\|_2^2 - \gamma^2\|w\|_2^2\right\}$$

since the solvability of problem (b) implies the solvability of problem (a). However, the converse may not be true. In fact, it is easy to construct an example so that the problem (a) has a solution for a given γ and problem (b) does not. On the other hand, the problems (a) and (b) are equivalent if $D_{11} = 0$.

17.9 Parameterization of State Feedback $\mathcal{H}_\infty$ Controllers

In this section, we shall consider the parameterization of all state feedback control laws. We shall first assume for simplicity that $D_{11} = 0$ and show later how to reduce the general $D_{11} \neq 0$ case to an equivalent problem with $D_{11} = 0$. We shall assume

$$G = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ I & 0 & 0 \end{array}\right].$$

Note that the state feedback $\mathcal{H}_\infty$ problem is *not* a special case of the output feedback problem since $D_{21} = 0$. Hence the parameterization cannot be obtained from the output feedback.

Theorem 17.7 *Suppose that the assumptions $(AS1)-(AS3)$ are satisfied and that B_1 has the following SVD:*

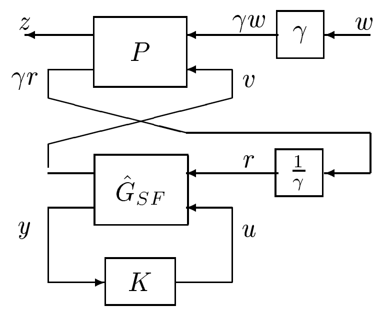
$$B_1 = U \begin{bmatrix} \Sigma & 0 \\ 0 & 0 \end{bmatrix} V^*, \quad UU^* = I_n, \quad V^*V = I_{m_1}, \quad 0 < \Sigma \in \mathbb{R}^{r\times r}.$$

There exists an admissible controller $K(s)$ for the SF problem such that $\|T_{zw}\|_\infty < \gamma$ if and only if $H_\infty \in dom(Ric)$ and $X_\infty = Ric(H_\infty) \geq 0$. Furthermore, if these conditions are satisfied, then all admissible controllers satisfying $\|T_{zw}\|_\infty < \gamma$ can be parameterized as

$$K = F_\infty + \left\{ I_{m_2} + Q \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & I_{m_1-r} \end{bmatrix} U^{-1}B_2 \right\}^{-1} Q \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & I_{m_1-r} \end{bmatrix} U^{-1}(sI - \hat{A})$$

where $F_\infty = -(D_{12}^*C_1 + B_2^*X_\infty)$, $\hat{A} = A + \gamma^{-2}B_1B_1^*X_\infty + B_2F_\infty$, and $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \in \mathcal{RH}_2$ with $\|Q_1\|_\infty < \gamma$. The dimensions of Q_1 and Q_2 are $m_2 \times r$ and $m_2 \times (m_1 - r)$, respectively.

Proof. The conditions for the existence of the state feedback control law have been shown in the last section and in Chapter 15. We only need to show that the above parameterization indeed satisfies the $\mathcal{H}_\infty$ norm condition and gives all possible state feedback $\mathcal{H}_\infty$ control laws. As in the proof of the FI problem in Chapter 16, we make the same change of variables to get $\hat{G}_{SF}$ instead of $\hat{G}_{FI}$:



$$\hat{G}_{SF} = \left[\begin{array}{c|cc} A_{tmp} & B_1 & B_2 \\ \hline -F_\infty & 0 & I \\ I & 0 & 0 \end{array}\right].$$

So again from Theorem 16.2 and Lemma 16.8, we conclude that K is an admissible controller for G and $\|T_{zw}\|_\infty < \gamma$ iff K is an admissible controller for $\hat{G}_{SF}$ and $\|T_{vr}\|_\infty < \gamma$. Now let $L = B_2 F_\infty$, and then $A_{tmp} + L = A_{tmp} + B_2 F_\infty$ is stable. All controllers that stabilize $\hat{G}_{SF}$ can be parameterized as $K = \mathcal{F}_\ell(M_{tmp}, \Phi)$, $\Phi \in \mathcal{RH}_\infty$ where

$$M_{tmp} = \left[\begin{array}{c|cc} A_{tmp} + B_2 F_\infty + L & -L & B_2 \\ \hline F_\infty & 0 & I \\ -I & I & 0 \end{array}\right].$$

Then $T_{vr} = \mathcal{F}_\ell(\hat{G}_{SF}, \mathcal{F}_\ell(M_{tmp}, \Phi)) =: \mathcal{F}_\ell(N_{tmp}, \Phi)$. It is easy to show that

$$N_{tmp} = \left[\begin{array}{c|cc} A_{tmp} + B_2 F_\infty & B_1 & 0 \\ \hline -F_\infty & 0 & I \\ I & 0 & 0 \end{array}\right].$$

Now let $\Phi = F_\infty + \hat{\Phi}$, and we have

$$\mathcal{F}_\ell(N_{tmp}, \Phi) = \hat{\Phi} \left[\begin{array}{c|c} A_{tmp} + B_2 F_\infty & B_1 \\ \hline I & 0 \end{array}\right].$$

Let

$$\hat{\Phi} = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \begin{bmatrix} \Sigma^{-1} & 0 \\ 0 & I \end{bmatrix} U^{-1}(sI - \hat{A}).$$

Then the mapping from $Q = \begin{bmatrix} Q_1 & Q_2 \end{bmatrix} \in \mathcal{RH}_2$ to $\hat{\Phi} \in \mathcal{RH}_\infty$ is one-to-one. Hence we have

$$\mathcal{F}_\ell(N_{tmp}, \Phi) = \begin{bmatrix} Q_1 & 0 \end{bmatrix} V^*$$

and $\|\mathcal{F}_\ell(N_{tmp}, \Phi)\|_\infty = \|Q_1\|_\infty$, which in turn implies that $\|T_{vr}\|_\infty = \|\mathcal{F}_\ell(N_{tmp}, \Phi)\|_\infty < \gamma$ if and only if $\|Q_1\|_\infty < \gamma$. Finally, substituting $\Phi = F_\infty + \hat{\Phi}$ into $K = \mathcal{F}_\ell(M_{tmp}, \Phi)$, we get the desired controller parameterization. $\square$

The controller parameterization for the general case can also be obtained:

$$G_g(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ I & 0 & 0 \end{array}\right].$$

Let

$$N = \begin{bmatrix} -D_{11} & (I - D_{11}D_{11}^*)^{1/2} \\ (I - D_{11}^* D_{11})^{1/2} & D_{11}^* \end{bmatrix},$$

and then $N^*N = I$. Furthermore, by Theorem 16.2, we have that K stabilizes G_g and $\|\mathcal{F}_\ell(G_g, K)\|_\infty < 1$ if and only if $\mathcal{F}_\ell(G_g, K)$ internally stabilizes N and

$$\|\mathcal{F}_\ell(N, \mathcal{F}_\ell(G_g, K))\|_\infty < 1.$$

Note that

$$\mathcal{F}_\ell(N, \mathcal{F}_\ell(G_g, K)) = \mathcal{F}_\ell(M, \tilde{K})$$

with

$$M(s) = \left[\begin{array}{c|cc} A + B_1 D_{11}^* R_1^{-1} C_1 & B_1(I - D_{11}^* D_{11})^{-1/2} & B_2 + B_1 D_{11}^* R_1^{-1} D_{12} \\ \hline R_1^{-1/2} C_1 & 0 & R_1^{-1/2} D_{12} \\ I & 0 & 0 \end{array}\right]$$

where $R_1 := I - D_{11}D_{11}^*$. In summary, we have the following lemma.

Lemma 17.8 *There is a K that internally stabilizes G_g and $\|\mathcal{F}_\ell(G_g, K)\|_\infty < 1$ if and only if $\|D_{11}\| < 1$ and there is a $\tilde{K}$ that stabilizes M and $\left\|\mathcal{F}_\ell(M, \tilde{K})\right\|_\infty < 1$.*

Now Theorem 17.7 can be applied to the new system $M(s)$ to obtain the controller parameterization for the general problem with $D_{11} \neq 0$.

17.10 Notes and References

The detailed derivation of the $\mathcal{H}_\infty$ solution for the general case is treated in Glover and Doyle [1988, 1989]. The loop-shifting procedures are given in Safonov, Limebeer, and Chiang [1989]. The idea is also used in Zhou and Khargonekar [1988] for state feedback problems. A fairly complete solution to the singular $\mathcal{H}_\infty$ problem is obtained in Stoorvogel [1992]. The $\mathcal{H}_\infty$ filtering and smoothing problems are considered in detail in Nagpal and Khargonekar [1991]. The Youla parameterization approach is treated very extensively in Doyle [1984] and Francis [1987] and in the references therein. The presentation of the state feedback $\mathcal{H}_\infty$ control in this chapter is based on Zhou [1992].

18

$\mathcal{H}_\infty$ Loop Shaping

This chapter introduces a design technique which incorporates loop shaping methods to obtain performance/robust stability tradeoffs, and a particular $\mathcal{H}_\infty$ optimization problem to guarantee closed-loop stability and a level of robust stability at all frequencies. The proposed technique uses only the basic concept of loop shaping methods and then a robust stabilization controller for the normalized coprime factor perturbed system is used to construct the final controller. This chapter is arranged as follows: The $\mathcal{H}_\infty$ theory is applied to solve the stabilization problem of a normalized coprime factor perturbed system in Section 18.1. The loop shaping design procedure is described in Section 18.2. The theoretical justification for the loop shaping design procedure is given in Section 18.3.

18.1 Robust Stabilization of Coprime Factors

In this section, we use the $\mathcal{H}_\infty$ control theory developed in the previous chapters to solve the robust stabilization of a left coprime factor perturbed plant given by

$$P_\Delta = (\tilde{M} + \tilde{\Delta}_M)^{-1}(\tilde{N} + \tilde{\Delta}_N)$$

with $\tilde{M}, \tilde{N}, \tilde{\Delta}_M, \tilde{\Delta}_N \in \mathcal{RH}_\infty$ and $\left\| \begin{bmatrix} \tilde{\Delta}_N & \tilde{\Delta}_M \end{bmatrix} \right\|_\infty < \epsilon$. The transfer matrices $(\tilde{M}, \tilde{N})$ are assumed to be a left coprime factorization of P (i.e., $P = \tilde{M}^{-1}\tilde{N}$), and K internally stabilizes the nominal system.

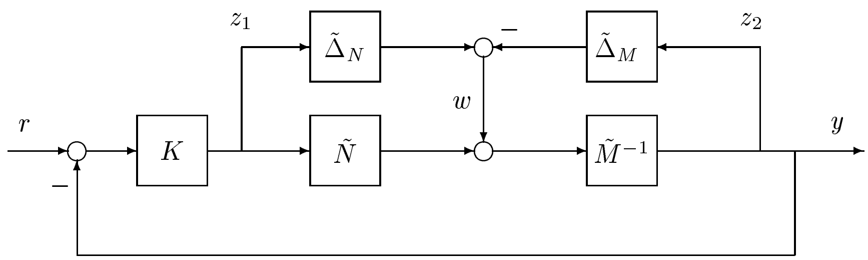


Figure 18.1: Left Coprime Factor Perturbed Systems

It has been shown in Chapter 9 that the system is robustly stable iff

$$\left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1}\tilde{M}^{-1} \right\|_\infty \le 1/\epsilon.$$

Finding a controller such that the above norm condition holds is an $\mathcal{H}_\infty$ norm minimization problem which can be solved using $\mathcal{H}_\infty$ theory developed in the previous chapters.

Suppose P has a stabilizable and detectable state space realization given by

$$P = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

and let L be a matrix such that $A+LC$ is stable then a left coprime factorization of $P = \tilde{M}^{-1}\tilde{N}$ is given by

$$\begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} = \left[\begin{array}{c|cc} A+LC & B+LD & L \\ \hline C & D & I \end{array}\right].$$

Denote

$$\hat{K} = -K$$

then the system diagram can be put in an LFT form as in Figure 18.2 with the generalized plant

$$G(s) = \begin{bmatrix} \begin{bmatrix} 0 \\ \tilde{M}^{-1} \end{bmatrix} & \begin{bmatrix} I \\ P \end{bmatrix} \\ \tilde{M}^{-1} & P \end{bmatrix} = \left[\begin{array}{c|cc} A & -L & B \\ \hline \begin{bmatrix} 0 \\ C \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} & \begin{bmatrix} I \\ D \end{bmatrix} \\ C & I & D \end{array}\right]$$

$$=: \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right].$$

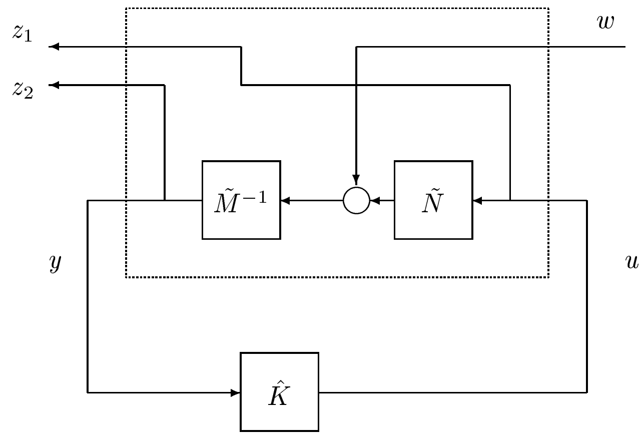


Figure 18.2: LFT Diagram for Coprime Factor Stabilization

To apply the $\mathcal{H}_\infty$ control formulae in Chapter 17, we need to normalize the "D_{12}" matrix first. Note that

$$\begin{bmatrix} I \\ D \end{bmatrix} = U \begin{bmatrix} 0 \\ I \end{bmatrix} (I + D^*D)^{\frac{1}{2}}, \quad \text{where } U = \begin{bmatrix} D^*(I + DD^*)^{-\frac{1}{2}} & I + D^*D)^{-\frac{1}{2}} \\ -(I + DD^*)^{-\frac{1}{2}} & D(I + D^*D)^{-\frac{1}{2}} \end{bmatrix}$$

and U is a unitary matrix. Let

$$\begin{aligned} \hat{K} &= (I + D^*D)^{-\frac{1}{2}} \tilde{K} \\ \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} &= U \begin{bmatrix} \hat{z}_1 \\ \hat{z}_2 \end{bmatrix}. \end{aligned}$$

Then $\|T_{zw}\|_\infty = \|U^* T_{zw}\|_\infty = \|T_{\hat{z}w}\|_\infty$ and the problem becomes of finding a controller $\tilde{K}$ so that $\|T_{\hat{z}w}\|_\infty < \gamma$ with the following generalized plant

$$\hat{G} = \begin{bmatrix} U^* & 0 \\ 0 & I \end{bmatrix} G \begin{bmatrix} I & 0 \\ 0 & (I + D^*D)^{-\frac{1}{2}} \end{bmatrix}$$

$$= \left[\begin{array}{c|cc} A & -L & B \\ \hline \begin{bmatrix} -(I + DD^*)^{-\frac{1}{2}} C \\ (I + D^*D)^{-\frac{1}{2}} D^* C \end{bmatrix} & \begin{bmatrix} -(I + DD^*)^{-\frac{1}{2}} \\ (I + D^*D)^{-\frac{1}{2}} D^* \end{bmatrix} & \begin{bmatrix} 0 \\ I \end{bmatrix} \\ C & I & D(I + D^*D)^{-\frac{1}{2}} \end{array}\right].$$

Now the formulae in Chapter 17 can be applied to $\hat{G}$ to obtain a controller $\tilde{K}$ and then the K can be obtained from $K = -(I + D^*D)^{-\frac{1}{2}} \tilde{K}$. We shall leave the detail to the reader. In the sequel, we shall consider the case $D = 0$. In this case, we have $\gamma > 1$ and

$$X_\infty(A - \frac{LC}{\gamma^2 - 1}) + (A - \frac{LC}{\gamma^2 - 1})^* X_\infty - X_\infty(BB^* - \frac{LL^*}{\gamma^2 - 1})X_\infty + \frac{\gamma^2 C^* C}{\gamma^2 - 1} = 0$$

$$Y_\infty(A + LC)^* + (A + LC)Y_\infty - Y_\infty C^* C Y_\infty = 0.$$

It is clear that $Y_\infty = 0$ is the stabilizing solution. Hence by the formulae in Chapter 17 we have

$$\begin{bmatrix} L_{1\infty} & L_{2\infty} \end{bmatrix} = \begin{bmatrix} 0 & L \end{bmatrix}$$

and

$$Z_\infty = I, \quad \hat{D}_{11} = 0, \quad \hat{D}_{12} = I, \quad \hat{D}_{21} = \frac{\sqrt{\gamma^2 - 1}}{\gamma} I.$$

The results are summarized in the following theorem.

Theorem 18.1 *Let $D = 0$ and let L be such that $A + LC$ is stable then there exists a controller K such that*

$$\left\| \begin{bmatrix} K \\ I \end{bmatrix} (I + PK)^{-1} \tilde{M}^{-1} \right\|_\infty < \gamma$$

iff $\gamma > 1$ and there exists a stabilizing solution $X_\infty \geq 0$ solving

$$X_\infty(A - \frac{LC}{\gamma^2 - 1}) + (A - \frac{LC}{\gamma^2 - 1})^* X_\infty - X_\infty(BB^* - \frac{LL^*}{\gamma^2 - 1})X_\infty + \frac{\gamma^2 C^* C}{\gamma^2 - 1} = 0.$$

Moreover, if the above conditions hold a central controller is given by

$$K = \left[\begin{array}{c|c} A - BB^* X_\infty + LC & L \\ \hline -B^* X_\infty & 0 \end{array} \right].$$

It is clear that the existence of a robust stabilizing controller depends upon the choice of the stabilizing matrix L, i.e., the choice of the coprime factorization. Now let $Y \geq 0$ be the stabilizing solution to

$$AY + YA^* - YC^*CY + BB^* = 0$$

and let $L = -YC^*$. Then the left coprime factorization $(\tilde{M}, \tilde{N})$ given by

$$\begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} = \left[\begin{array}{c|cc} A - YC^*C & B & -YC^* \\ \hline C & 0 & I \end{array} \right]$$

is a normalized left coprime factorization (see Chapter 13).

Corollary 18.2 *Let $D = 0$ and $L = -YC^*$ where $Y \geq 0$ is the stabilizing solution to*

$$AY + YA^* - YC^*CY + BB^* = 0.$$

Then $P = \tilde{M}^{-1}\tilde{N}$ is a normalized left coprime factorization and

$$\begin{aligned}\inf_{K \ stabilizing} \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I + PK)^{-1}\tilde{M}^{-1} \right\|_\infty &= \frac{1}{\sqrt{1 - \lambda_{max}(YQ)}} \\ &= \left(1 - \left\| \begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} \right\|_H^2 \right)^{-1/2} =: \gamma_{min}\end{aligned}$$

where Q is the solution to the following Lyapunov equation

$$Q(A - YC^*C) + (A - YC^*C)^*Q + C^*C = 0.$$

Moreover, if the above conditions hold then for any $\gamma > \gamma_{min}$ a controller achieving

$$\left\| \begin{bmatrix} K \\ I \end{bmatrix} (I + PK)^{-1}\tilde{M}^{-1} \right\|_\infty < \gamma$$

is given by

$$K(s) = \left[\begin{array}{c|c} A - BB^*X_\infty - YC^*C & -YC^* \\ \hline -B^*X_\infty & 0 \end{array}\right]$$

where

$$X_\infty = \frac{\gamma^2}{\gamma^2 - 1} Q \left(I - \frac{\gamma^2}{\gamma^2 - 1} YQ\right)^{-1}.$$

Proof. Note that the Hamiltonian matrix associated with X_∞ is given by

$$H_\infty = \begin{bmatrix} A + \frac{1}{\gamma^2 - 1} YC^*C & -BB^* + \frac{1}{\gamma^2-1} YC^*CY \\ -\frac{\gamma^2}{\gamma^2-1} C^*C & -(A + \frac{1}{\gamma^2-1} YC^*C)^* \end{bmatrix}.$$

Straightforward calculation shows that

$$H_\infty = \begin{bmatrix} I & -\frac{\gamma^2}{\gamma^2-1} Y \\ 0 & \frac{\gamma^2}{\gamma^2-1} I \end{bmatrix} H_q \begin{bmatrix} I & -\frac{\gamma^2}{\gamma^2-1} Y \\ 0 & \frac{\gamma^2}{\gamma^2-1} I \end{bmatrix}^{-1}$$

where

$$H_q = \begin{bmatrix} A - YC^*C & 0 \\ -C^*C & -(A - YC^*C)^* \end{bmatrix}.$$

It is clear that the stable invariant subspace of H_q is given by

$$\mathcal{X}_-(H_q) = \mathrm{Im} \begin{bmatrix} I \\ Q \end{bmatrix}$$

and the stable invariant subspace of H_∞ is given by

$$\mathcal{X}_-(H_\infty) = \begin{bmatrix} I & -\frac{\gamma^2}{\gamma^2-1}Y \\ 0 & \frac{\gamma^2}{\gamma^2-1}I \end{bmatrix} \mathcal{X}_-(H_q) = \mathrm{Im} \begin{bmatrix} I - \frac{\gamma^2}{\gamma^2-1}YQ \\ \frac{\gamma^2}{\gamma^2-1}Q \end{bmatrix}.$$

Hence there is a nonnegative definite stabilizing solution to the algebraic Riccati equation of X_∞ if and only if

$$I - \frac{\gamma^2}{\gamma^2-1}YQ > 0$$

or

$$\gamma > \frac{1}{\sqrt{1-\lambda_{max}(YQ)}}$$

and the solution, if it exists, is given by

$$X_\infty = \frac{\gamma^2}{\gamma^2-1}Q\left(I - \frac{\gamma^2}{\gamma^2-1}YQ\right)^{-1}.$$

Note that Y and Q are the controllability Gramian and the observability Gramian of $\begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix}$ respectively. Therefore, we also have that the Hankel norm of $\begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix}$ is $\sqrt{\lambda_{max}(YQ)}$. □

Corollary 18.3 *Let $P = \tilde{M}^{-1}\tilde{N}$ be a normalized left coprime factorization and*

$$P_\Delta = (\tilde{M} + \tilde{\Delta}_M)^{-1}(\tilde{N} + \tilde{\Delta}_N)$$

with

$$\left\| \begin{bmatrix} \tilde{\Delta}_N & \tilde{\Delta}_M \end{bmatrix} \right\|_\infty < \epsilon.$$

Then there is a robustly stabilizing controller for P_Δ if and only if

$$\epsilon \le \sqrt{1-\lambda_{max}(YQ)} = \sqrt{1 - \left\| \begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} \right\|_H^2}.$$

The solutions to the normalized left coprime factorization stabilization problem are also solutions to a related $\mathcal{H}_\infty$ problem which is shown in the following lemma.

Lemma 18.4 *Let $P = \tilde{M}^{-1}\tilde{N}$ be a normalized left coprime factorization. Then*

$$\left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1}\tilde{M}^{-1} \right\|_\infty = \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1} \begin{bmatrix} I & P \end{bmatrix} \right\|_\infty.$$

Proof. Since $(\tilde{M}, \tilde{N})$ is a normalized left coprime factorization of P, we have

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} \begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}^{\sim} = I$$

and

$$\left\| \begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} \right\|_\infty = \left\| \begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}^{\sim} \right\|_\infty = 1.$$

Using these equations, we have

$$\begin{aligned} & \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1}\tilde{M}^{-1} \right\|_\infty \\ = & \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1}\tilde{M}^{-1} \begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} \begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}^{\sim} \right\|_\infty \\ \leq & \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1}\tilde{M}^{-1} \begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} \right\|_\infty \left\| \begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}^{\sim} \right\|_\infty \\ = & \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1} \begin{bmatrix} I & P \end{bmatrix} \right\|_\infty \\ \leq & \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1}\tilde{M}^{-1} \right\|_\infty \left\| \begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} \right\|_\infty \\ = & \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1}\tilde{M}^{-1} \right\|_\infty . \end{aligned}$$

This implies

$$\left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1}\tilde{M}^{-1} \right\|_\infty = \left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1} \begin{bmatrix} I & P \end{bmatrix} \right\|_\infty .$$

□

Corollary 18.5 *A controller solves the normalized left coprime factor robust stabilization problem if and only if it solves the following $\mathcal{H}_\infty$ control problem*

$$\left\| \begin{bmatrix} I \\ K \end{bmatrix} (I+PK)^{-1} \begin{bmatrix} I & P \end{bmatrix} \right\|_\infty < \gamma$$

and

$$\begin{aligned} \inf_{K \ stabilizing} \left\| \begin{bmatrix} I \\ K \end{bmatrix} (I+PK)^{-1} \begin{bmatrix} I & P \end{bmatrix} \right\|_\infty & = \frac{1}{\sqrt{1-\lambda_{max}(YQ)}} \\ & = \left(1 - \left\| \begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} \right\|_H^2 \right)^{-1/2} . \end{aligned}$$

The solution Q can also be obtained in other ways. Let $X \geq 0$ be the stabilizing solution to

$$XA + A^*X - XBB^*X + C^*C = 0$$

then it is easy to verify that

$$Q = (I + XY)^{-1}X.$$

Hence

$$\gamma_{min} = \frac{1}{\sqrt{1 - \lambda_{max}(YQ)}} = \left(1 - \left\| \begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} \right\|_H^2 \right)^{-1/2} = \sqrt{1 + \lambda_{max}(XY)}.$$

Similar results can be obtained if one starts with a normalized right coprime factorization. In fact, a rather strong relation between the normalized left and right coprime factorization problems can be established using the following matrix fact.

Lemma 18.6 *Let M and N be any compatibly dimensioned complex matrices such that $MM = M$, $NN = N$, and $M + N = I$. Then $\sigma_i(M) = \sigma_i(N)$ for all i such that $0 < \sigma_i(M) \neq 1$.*

Proof. We first show that the eigenvalues of M and N are either 0 or 1 and M and N are diagonalizable. In fact, assume that λ is an eigenvalue of M and x is a corresponding eigenvector, then $\lambda x = Mx = MMx = M(Mx) = \lambda Mx = \lambda^2 x$, i.e., $\lambda(1-\lambda)x = 0$. This implies that either $\lambda = 0$ or $\lambda = 1$. To show that M is diagonalizable, assume $M = TJT^{-1}$ where J is a Jordan canonical form, it follows immediately that J must be diagonal by the condition $M = MM$. The proof for N is similar.

Next, assume that M is diagonalized by a nonsingular matrix T such that

$$M = T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T^{-1}.$$

Then

$$N = I - M = T \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix} T^{-1}.$$

Define

$$\begin{bmatrix} A & B \\ B^* & D \end{bmatrix} := T^*T$$

and assume $0 < \lambda \neq 1$. Then $A > 0$ and

$$\begin{aligned} & \det(M^*M - \lambda I) = 0 \\ \Leftrightarrow \quad & \det(\begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} T^*T \begin{bmatrix} I & 0 \\ 0 & 0 \end{bmatrix} - T^*T) = 0 \end{aligned}$$

$$
\begin{aligned}
\Leftrightarrow \quad & \det \begin{bmatrix} (1-\lambda)A & -\lambda B \\ -\lambda B^* & -\lambda D \end{bmatrix} = 0 \\
\Leftrightarrow \quad & \det(-\lambda D - \frac{\lambda^2}{1-\lambda} B^* A^{-1} B) = 0 \\
\Leftrightarrow \quad & \det(\frac{1-\lambda}{\lambda} D + B^* A^{-1} B) = 0 \\
\Leftrightarrow \quad & \det \begin{bmatrix} -\lambda A & -\lambda B \\ -\lambda B^* & (1-\lambda)D \end{bmatrix} = 0 \\
\Leftrightarrow \quad & \det(N^* N - \lambda I) = 0.
\end{aligned}
$$

This implies that all nonzero eigenvalues of M^*M and N^*N that are not equal to 1 are equal, i.e., $\sigma_i(M) = \sigma_i(N)$ for all i such that $0 < \sigma_i(M) \neq 1$. □

Using this matrix fact, we have the following corollary.

Corollary 18.7 *Let K and P be any compatibly dimensioned complex matrices. Then*

$$
\left\| \begin{bmatrix} I \\ K \end{bmatrix} (I+PK)^{-1} \begin{bmatrix} I & P \end{bmatrix} \right\| = \left\| \begin{bmatrix} I \\ P \end{bmatrix} (I+KP)^{-1} \begin{bmatrix} I & K \end{bmatrix} \right\|.
$$

Proof. Define

$$
M = \begin{bmatrix} I \\ K \end{bmatrix} (I+PK)^{-1} \begin{bmatrix} I & P \end{bmatrix}, \quad N = \begin{bmatrix} -P \\ I \end{bmatrix} (I+KP)^{-1} \begin{bmatrix} -K & I \end{bmatrix}.
$$

Then it is easy to verify that $MM = M$, $NN = N$ and $M + N = I$. By Lemma 18.6, we have $\|M\| = \|N\|$. The corollary follows by noting that

$$
\begin{bmatrix} I \\ P \end{bmatrix} (I+KP)^{-1} \begin{bmatrix} I & K \end{bmatrix} = \begin{bmatrix} 0 & I \\ -I & 0 \end{bmatrix} N \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix}.
$$

□

Corollary 18.8 *Let $P = \tilde{M}^{-1}\tilde{N} = NM^{-1}$ be respectively the normalized left and right coprime factorizations. Then*

$$
\left\| \begin{bmatrix} K \\ I \end{bmatrix} (I+PK)^{-1} \tilde{M}^{-1} \right\|_\infty = \left\| M^{-1} (I+KP)^{-1} \begin{bmatrix} I & K \end{bmatrix} \right\|_\infty.
$$

Proof. This follows from Corollary 18.7 and the fact that

$$\left\| M^{-1}(I+KP)^{-1} \begin{bmatrix} I & K \end{bmatrix} \right\|_\infty = \left\| \begin{bmatrix} I \\ P \end{bmatrix} (I+KP)^{-1} \begin{bmatrix} I & K \end{bmatrix} \right\|_\infty .$$

□

This corollary says that any $\mathcal{H}_\infty$ controller for the normalized left coprime factorization is also an $\mathcal{H}_\infty$ controller for the normalized right coprime factorization. Hence one can work with either factorization.

18.2 Loop Shaping Using Normalized Coprime Stabilization

This section considers the $\mathcal{H}_\infty$ loop shaping design. The objective of this approach is to incorporate the simple performance/robustness tradeoff obtained in the loop shaping, with the guaranteed stability properties of $\mathcal{H}_\infty$ design methods. Recall from Section 5.5 of Chapter 5 that good performance controller design requires that

$$\overline{\sigma}\left((I+PK)^{-1}\right), \quad \overline{\sigma}\left((I+PK)^{-1}P\right), \quad \overline{\sigma}\left((I+KP)^{-1}\right), \quad \overline{\sigma}\left(K(I+PK)^{-1}\right) \quad (18.1)$$

be made small, particularly in some low frequency range. And good robustness requires that

$$\overline{\sigma}\left(PK(I+PK)^{-1}\right), \quad \overline{\sigma}\left(KP(I+KP)^{-1}\right) \quad (18.2)$$

be made small, particularly in some high frequency range. These requirements in turn imply that good controller design boils down to achieving the desired loop (and controller) gains in the appropriate frequency range:

$$\underline{\sigma}(PK) \gg 1, \quad \underline{\sigma}(KP) \gg 1, \quad \underline{\sigma}(K) \gg 1$$

in some low frequency range and

$$\overline{\sigma}(PK) \ll 1, \quad \overline{\sigma}(KP) \ll 1, \quad \overline{\sigma}(K) \leq M$$

in some high frequency range where M is not too large.

The design procedure is stated below.

Loop Shaping Design Procedure

(1) Loop Shaping: Using a precompensator W_1 and/or a postcompensator W_2, the singular values of the nominal plant are shaped to give a desired open-loop shape. The nominal plant G and the shaping functions W_1, W_2 are combined to form the shaped plant, G_s where $G_s = W_2 G W_1$. We assume that W_1 and W_2 are such that G_s contains no hidden modes.

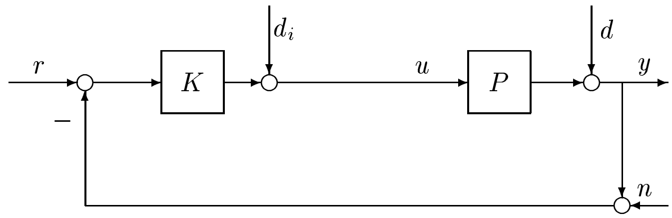


Figure 18.3: Standard Feedback Configuration

(2) Robust Stabilization: a) Calculate ϵ_{max}, where

$$\begin{aligned}\epsilon_{max} &= \left(\inf_{K \ stabilizing} \left\| \begin{bmatrix} I \\ K \end{bmatrix} (I + G_s K)^{-1} \tilde{M}_s^{-1} \right\|_\infty \right)^{-1} \\ &= \sqrt{1 - \left\| \begin{bmatrix} \tilde{N}_s & \tilde{M}_s \end{bmatrix} \right\|_H^2} < 1\end{aligned}$$

and $\tilde{M}_s, \tilde{N}_s$ define the normalized coprime factors of G_s such that $G_s = \tilde{M}_s^{-1}\tilde{N}_s$ and

$$\tilde{M}_s \tilde{M}_s^\sim + \tilde{N}_s \tilde{N}_s^\sim = I.$$

If $\epsilon_{max} \ll 1$ return to (1) and adjust W_1 and W_2.

b) Select $\epsilon \leq \epsilon_{max}$, then synthesize a stabilizing controller K_∞, which satisfies

$$\left\| \begin{bmatrix} I \\ K_\infty \end{bmatrix} (I + G_s K_\infty)^{-1} \tilde{M}_s^{-1} \right\|_\infty \leq \epsilon^{-1}.$$

(3) The final feedback controller K is then constructed by combining the $\mathcal{H}_\infty$ controller K_∞ with the shaping functions W_1 and W_2 such that

$$K = W_1 K_\infty W_2.$$

A typical design works as follows: the designer inspects the open-loop singular values of the nominal plant, and shapes these by pre- and/or postcompensation until nominal performance (and possibly robust stability) specifications are met. (Recall that the open-loop shape is related to closed-loop objectives.) A feedback controller K_∞ with associated stability margin (for the shaped plant) $\epsilon \leq \epsilon_{max}$, is then synthesized. If ϵ_{max} is small, then the specified loop shape is incompatible with robust stability requirements, and should be adjusted accordingly, then K_∞ is reevaluated.

In the above design procedure we have specified the desired loop shape by $W_2 G W_1$. But, after Stage (2) of the design procedure, the actual loop shape achieved is in fact

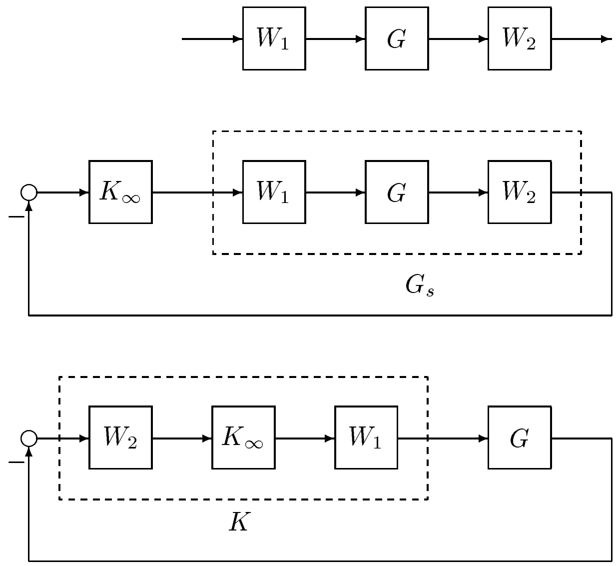


Figure 18.4: The Loop Shaping Design Procedure

given by $W_1 K_\infty W_2 G$ at plant input and $G W_1 K_\infty W_2$ at plant output. It is therefore possible that the inclusion of K_∞ in the open-loop transfer function will cause deterioration in the open-loop shape specified by G_s. In the next section, we will show that the degradation in the loop shape caused by the $\mathcal{H}_\infty$ controller K_∞ is limited at frequencies where the specified loop shape is sufficiently large or sufficiently small. In particular, we show in the next section that ϵ can be interpreted as an indicator of the success of the loop shaping in addition to providing a robust stability guarantee for the closed-loop systems. A small value of ϵ_{max} ($\epsilon_{max} \ll 1$) in Stage (2) always indicates incompatibility between the specified loop shape, the nominal plant phase, and robust closed-loop stability.

Remark 18.1 Note that, in contrast to the classical loop shaping approach, the loop shaping here is done without explicit regard for the nominal plant phase information. That is, closed-loop stability requirements are disregarded at this stage. Also, in contrast with conventional $\mathcal{H}_\infty$ design, the robust stabilization is done without frequency weighting. The design procedure described here is both simple and systematic, and only assumes knowledge of elementary loop shaping principles on the part of the designer. ♡

Remark 18.2 The above robust stabilization objective can also be interpreted as the more standard $\mathcal{H}_\infty$ problem formulation of minimizing the $\mathcal{H}_\infty$ norm of the frequency weighted gain from disturbances on the plant input and output to the controller input and output as follows.

$$\begin{aligned}
\left\| \begin{bmatrix} I \\ K_\infty \end{bmatrix} (I+G_sK_\infty)^{-1}\tilde{M}_s^{-1} \right\|_\infty &= \left\| \begin{bmatrix} I \\ K_\infty \end{bmatrix} (I+G_sK_\infty)^{-1} \begin{bmatrix} I & G_s \end{bmatrix} \right\|_\infty \\
&= \left\| \begin{bmatrix} W_2 \\ W_1^{-1}K \end{bmatrix} (I+GK)^{-1} \begin{bmatrix} W_2^{-1} & GW_1 \end{bmatrix} \right\|_\infty \\
&= \left\| \begin{bmatrix} I \\ G_s \end{bmatrix} (I+K_\infty G_s)^{-1} \begin{bmatrix} I & K_\infty \end{bmatrix} \right\|_\infty \\
&= \left\| \begin{bmatrix} W_1^{-1} \\ W_2G \end{bmatrix} (I+KG)^{-1} \begin{bmatrix} W_1 & GW_2^{-1} \end{bmatrix} \right\|_\infty
\end{aligned}$$

This shows how all the closed-loop objectives in (18.1) and (18.2) are incorporated. ♡

18.3 Theoretical Justification for $\mathcal{H}_\infty$ Loop Shaping

The objective of this section is to provide justification for the use of parameter ϵ as a design indicator. We will show that ϵ is a measure of both closed-loop robust stability and the success of the design in meeting the loop shaping specifications.

We first examine the possibility of loop shape deterioration at frequencies of high loop gain (typically low frequency). At low frequency (in particular, $\omega \in (0, \omega_l)$), the deterioration in loop shape at plant output can be obtained by comparing $\underline{\sigma}(GW_1K_\infty W_2)$ to $\underline{\sigma}(G_s) = \underline{\sigma}(W_2GW_1)$. Note that

$$\underline{\sigma}(GK) = \underline{\sigma}(GW_1K_\infty W_2) = \underline{\sigma}(W_2^{-1}W_2GW_1K_\infty W_2) \geq \underline{\sigma}(W_2GW_1)\underline{\sigma}(K_\infty)/\kappa(W_2) \tag{18.3}$$

where $\kappa(\cdot)$ denotes condition number. Similarly, for loop shape deterioration at plant input, we have

$$\underline{\sigma}(KG) = \underline{\sigma}(W_1K_\infty W_2G) = \underline{\sigma}(W_1K_\infty W_2GW_1W_1^{-1}) \geq \underline{\sigma}(W_2GW_1)\underline{\sigma}(K_\infty)/\kappa(W_1). \tag{18.4}$$

In each case, $\underline{\sigma}(K_\infty)$ is required to obtain a bound on the deterioration in the loop shape at low frequency. Note that the condition numbers $\kappa(W_1)$ and $\kappa(W_2)$ are selected by the designer.

Next, recalling that G_s denotes the shaped plant, and that K_∞ robustly stabilizes the normalized coprime factorization of G_s with stability margin ϵ, then we have

$$\left\| \begin{bmatrix} I \\ K_\infty \end{bmatrix} (I+G_sK_\infty)^{-1}\tilde{M}_s^{-1} \right\|_\infty \leq \epsilon^{-1} := \gamma \tag{18.5}$$

where $(\tilde{N}_s, \tilde{M}_s)$ is a normalized left coprime factorization of G_s, and the parameter γ is defined to simplify the notation to follow. The following result shows that $\underline{\sigma}(K_\infty)$

is explicitly bounded by functions of ϵ and $\underline{\sigma}(G_s)$, the minimum singular value of the shaped plant, and hence by (18.3) and (18.4) K_∞ will only have a limited effect on the specified loop shape at low frequency.

Theorem 18.9 *Any controller K_∞ satisfying (18.5), where G_s is assumed square, also satisfies*

$$\underline{\sigma}(K_\infty(j\omega)) \geq \frac{\underline{\sigma}(G_s(j\omega)) - \sqrt{\gamma^2 - 1}}{\sqrt{\gamma^2 - 1}\underline{\sigma}(G_s(j\omega)) + 1}$$

for all ω such that

$$\underline{\sigma}(G_s(j\omega)) > \sqrt{\gamma^2 - 1}.$$

Furthermore, if $\underline{\sigma}(G_s) \gg \sqrt{\gamma^2 - 1}$, then $\underline{\sigma}(K_\infty(j\omega)) \gtrapprox 1/\sqrt{\gamma^2 - 1}$, where $\gtrapprox$ denotes asymptotically greater than or equal to as $\underline{\sigma}(G_s) \to \infty$.

Proof. First note that $\underline{\sigma}(G_s) > \sqrt{\gamma^2 - 1}$ implies that

$$I + G_s G_s^* > \gamma^2 I.$$

Further since $(\tilde{N}_s, \tilde{M}_s)$ is a normalized left coprime factorization of G_s, we have

$$\tilde{M}_s \tilde{M}_s^* = I - \tilde{N}_s \tilde{N}_s^* = I - \tilde{M}_s G_s G_s^* \tilde{M}_s^*.$$

Then

$$\tilde{M}_s^* \tilde{M}_s = (I + G_s G_s^*)^{-1} < \gamma^{-2} I.$$

Now

$$\left\| \begin{bmatrix} I \\ K_\infty \end{bmatrix} (I + G_s K_\infty)^{-1} \tilde{M}_s^{-1} \right\|_\infty \leq \gamma$$

can be rewritten as

$$(I + K_\infty^* K_\infty) \leq \gamma^2 (I - K_\infty^* G_s^*)(\tilde{M}_s^* \tilde{M}_s)(I - G_s K_\infty). \tag{18.6}$$

We will next show that K_∞ is invertible. Suppose that there exists an x such that $K_\infty x = 0$, then $x^* \times (18.6) \times x$ gives

$$\gamma^{-2} x^* x \leq x^* \tilde{M}_s^* \tilde{M}_s x$$

which implies that $x = 0$ since $\tilde{M}_s^* \tilde{M}_s < \gamma^{-2} I$, and hence K_∞ is invertible. Equation (18.6) can now be written as

$$(K_\infty^{-*} K_\infty^{-1} + I) \leq \gamma^2 (K_\infty^{-*} - G_s^*) \tilde{M}_s^* \tilde{M}_s (K_\infty^{-1} - G_s). \tag{18.7}$$

Define W such that

$$(WW^*)^{-1} = I - \gamma^2 \tilde{M}_s^* \tilde{M}_s = I - \gamma^2 (I + G_s G_s^*)^{-1}$$

and completing the square in (18.7) with respect to K_∞^{-1} yields

$$(K_\infty^{-*} + N^*)(WW^*)^{-1}(K_\infty^{-1} + N) \leq (\gamma^2 - 1)R^*R$$

where

$$\begin{aligned} N &= \gamma^2 G_s((1-\gamma^2)I + G_s^* G_s)^{-1} \\ R^*R &= (I + G_s^* G_s)((1-\gamma^2)I + G_s^* G_s)^{-1}. \end{aligned}$$

Hence we have

$$R^{-*}(K_\infty^{-*} + N^*)(WW^*)^{-1}(K_\infty^{-1} + N)R^{-1} \leq (\gamma^2 - 1)I$$

and

$$\overline{\sigma}\left(W^{-1}(K_\infty^{-1} + N)R^{-1}\right) \leq \sqrt{\gamma^2 - 1}.$$

Use $\overline{\sigma}\left(W^{-1}(K_\infty^{-1} + N)R^{-1}\right) \geq \underline{\sigma}(W^{-1})\overline{\sigma}(K_\infty^{-1} + N)\underline{\sigma}(R^{-1})$ to get

$$\overline{\sigma}(K_\infty^{-1} + N) \leq \sqrt{\gamma^2 - 1}\,\overline{\sigma}(W)\overline{\sigma}(R)$$

and use $\underline{\sigma}(K_\infty^{-1} + N) \geq \underline{\sigma}(K_\infty) - \overline{\sigma}(N)$ to get

$$\underline{\sigma}(K_\infty) \geq \left\{(\gamma^2 - 1)^{1/2}\overline{\sigma}(W)\overline{\sigma}(R) + \overline{\sigma}(N)\right\}^{-1}. \tag{18.8}$$

Next, note that the eigenvalues of WW^*, N^*N, and R^*R can be computed as follows

$$\lambda(WW^*) = \frac{1 + \lambda(G_s G_s^*)}{1 - \gamma^2 + \lambda(G_s G_s^*)}$$

$$\lambda(N^*N) = \frac{\gamma^4 \lambda(G_s G_s^*)}{(1 - \gamma^2 + \lambda(G_s G_s^*))^2}$$

$$\lambda(R^*R) = \frac{1 + \lambda(G_s G_s^*)}{1 - \gamma^2 + \lambda(G_s G_s^*)}$$

therefore

$$\overline{\sigma}(W) = \sqrt{\lambda_{max}(WW^*)} = \left(\frac{1 + \lambda_{min}(G_s G_s^*)}{1 - \gamma^2 + \lambda_{min}(G_s G_s^*)}\right)^{1/2} = \left(\frac{1 + \underline{\sigma}^2(G_s)}{1 - \gamma^2 + \underline{\sigma}^2(G_s)}\right)^{1/2}$$

$$\overline{\sigma}(N) = \sqrt{\lambda_{max}(N^*N)} = \frac{\gamma^2\sqrt{\lambda_{min}(G_s G_s^*)}}{1 - \gamma^2 + \lambda_{min}(G_s G_s^*)} = \frac{\gamma^2 \underline{\sigma}(G_s)}{1 - \gamma^2 + \underline{\sigma}^2(G_s)}$$

$$\overline{\sigma}(R) = \sqrt{\lambda_{max}(R^*R)} = \left(\frac{1 + \lambda_{min}(G_s G_s^*)}{1 - \gamma^2 + \lambda_{min}(G_s G_s^*)}\right)^{1/2} = \left(\frac{1 + \underline{\sigma}^2(G_s)}{1 - \gamma^2 + \underline{\sigma}^2(G_s)}\right)^{1/2}.$$

Substituting these formulas into (18.8), we have

$$\underline{\sigma}(K_\infty) \geq \left\{ \frac{(\gamma^2-1)^{1/2}(1+\underline{\sigma}^2(G_s)) + \gamma^2\underline{\sigma}(G_s)}{\underline{\sigma}^2(G_s) - (\gamma^2-1)} \right\}^{-1} = \frac{\underline{\sigma}(G_s)) - \sqrt{\gamma^2-1}}{\sqrt{\gamma^2-1}\underline{\sigma}(G_s)+1}.$$

□

The main implication of Theorem 18.9 is that the bound on $\underline{\sigma}(K_\infty)$ depends only on the selected loop shape, and the stability margin of the shaped plant. The value of $\gamma(=\epsilon^{-1})$ directly determines the frequency range over which this result is valid–a small γ (large ϵ) is desirable, as we would expect. Further, G_s has a sufficiently large loop gain, then so also will $G_s K_\infty$ provided $\gamma(=\epsilon^{-1})$ is sufficiently small.

In an analogous manner, we now examine the possibility of deterioration in the loop shape at high frequency due to the inclusion of K_∞. Note that at high frequency (in particular, $\omega \in (\omega_h, \infty)$) the deterioration in plant output loop shape can be obtained by comparing $\overline{\sigma}(GW_1K_\infty W_2)$ to $\overline{\sigma}(G_s) = \overline{\sigma}(W_2GW_1)$. Note that, analogously to (18.3) and (18.4) we have

$$\overline{\sigma}(GK) = \overline{\sigma}(GW_1K_\infty W_2) \leq \overline{\sigma}(W_2GW_1)\overline{\sigma}(K_\infty)\kappa(W_2).$$

Similarly, the corresponding deterioration in plant input loop shape is obtained by comparing $\overline{\sigma}(W_1K_\infty W_2G)$ to $\overline{\sigma}(W_2GW_1)$ where

$$\overline{\sigma}(KG) = \overline{\sigma}(W_1K_\infty W_2G) \leq \overline{\sigma}(W_2GW_1)\overline{\sigma}(K_\infty)\kappa(W_1).$$

Hence, in each case, $\overline{\sigma}(K_\infty)$ is required to obtain a bound on the deterioration in the loop shape at high frequency. In an identical manner to Theorem 18.9, we now show that $\overline{\sigma}(K_\infty)$ is explicitly bounded by functions of γ, and $\overline{\sigma}(G_s)$, the maximum singular value of the shaped plant.

Theorem 18.10 *Any controller K_∞ satisfying (18.5) also satisfies*

$$\underline{\sigma}(K_\infty(j\omega)) \leq \frac{\sqrt{\gamma^2-1} + \overline{\sigma}(G_s(j\omega))}{1 - \sqrt{\gamma^2-1}\overline{\sigma}(G_s(j\omega))}$$

for all ω such that

$$\overline{\sigma}(G_s(j\omega)) < \frac{1}{\sqrt{\gamma^2-1}}.$$

Furthermore, if $\overline{\sigma}(G_s) \ll 1/\sqrt{\gamma^2-1}$, then $\underline{\sigma}(K_\infty(j\omega)) \lessapprox \sqrt{\gamma^2-1}$, where $\lessapprox$ denotes asymptotically less than or equal to as $\overline{\sigma}(G_s) \to 0$.

Proof. The proof of Theorem 18.10 is similar to that of Theorem 18.9, and is only sketched here: As in the proof of Theorem 18.9, we have $\tilde{M}_s^*\tilde{M}_s = (I + G_sG_s^*)^{-1}$ and

$$(I + K_\infty^* K_\infty) \leq \gamma^2(I - K_\infty^* G_s^*)(\tilde{M}_s^*\tilde{M}_s)(I - G_sK_\infty). \tag{18.9}$$

Since $\overline{\sigma}(G_s) < \frac{1}{\sqrt{\gamma^2-1}}$,

$$I - \gamma^2 G_s^*(I + G_s G_s^*)^{-1} G_s > 0$$

and there exists a spectral factorization

$$V^*V = I - \gamma^2 G_s^*(I + G_s G_s^*)^{-1} G_s.$$

Now completing the square in (18.9) with respect to K_∞ yields

$$(K_\infty^* + M^*)V^*V(K_\infty + M) \leq (\gamma^2 - 1)Y^*Y$$

where

$$\begin{aligned} M &= \gamma^2 G_s^*(I + (1-\gamma^2)G_s G_s^*)^{-1} \\ Y^*Y &= (\gamma^2 - 1)(I + G_s G_s^*)(I + (1-\gamma^2)G_s G_s^*)^{-1}. \end{aligned}$$

Hence we have

$$\overline{\sigma}\left(V(K_\infty + M)Y^{-1}\right) \leq \sqrt{\gamma^2 - 1}$$

which implies

$$\overline{\sigma}(K_\infty) \leq \frac{\sqrt{\gamma^2-1}}{\underline{\sigma}(V)\underline{\sigma}(Y^{-1})} + \overline{\sigma}(M). \tag{18.10}$$

As in the proof of Theorem 18.9, it is easy to show that

$$\underline{\sigma}(V) = \underline{\sigma}(Y^{-1}) = \left(\frac{1 - (\gamma^2-1)\overline{\sigma}^2(G_s)}{1 + \overline{\sigma}^2(G_s)}\right)^{1/2}$$

$$\overline{\sigma}(M) = \frac{\gamma^2\overline{\sigma}(G_s)}{1 - (\gamma^2-1)\overline{\sigma}^2(G_s)}.$$

Substituting these formulas into (18.10), we have

$$\overline{\sigma}(K_\infty) \leq \frac{(\gamma^2-1)^{1/2}(1 + \overline{\sigma}^2(G_s)) + \gamma^2\overline{\sigma}(G_s)}{1 - (\gamma^2-1)\overline{\sigma}^2(G_s)} = \frac{\sqrt{\gamma^2-1} + \overline{\sigma}(G_s)}{1 - \sqrt{\gamma^2-1}\,\overline{\sigma}(G_s)}.$$

□

The results in Theorem 18.9 and 18.10 confirm that γ (alternatively ϵ) indicates the compatibility between the specified loop shape and closed-loop stability requirements.

Theorem 18.11 *Let G be the nominal plant and let $K = W_1 K_\infty W_2$ be the associated controller obtained from loop shaping design procedure in the last section. Then if*

$$\left\| \begin{bmatrix} I \\ K_\infty \end{bmatrix} (I + G_s K_\infty)^{-1} \tilde{M}_s^{-1} \right\|_\infty \leq \gamma$$

we have

$$\begin{aligned}
\overline{\sigma}\left(K(I+GK)^{-1}\right) &\leq \gamma\overline{\sigma}(\tilde{M}_s)\overline{\sigma}(W_1)\overline{\sigma}(W_2) && (18.11)\\
\overline{\sigma}\left((I+GK)^{-1}\right) &\leq \min\left\{\gamma\overline{\sigma}(\tilde{M}_s)\kappa(W_2),\ 1+\gamma\overline{\sigma}(N_s)\kappa(W_2)\right\} && (18.12)\\
\overline{\sigma}\left(K(I+GK)^{-1}G\right) &\leq \min\left\{\gamma\overline{\sigma}(\tilde{N}_s)\kappa(W_1),\ 1+\gamma\overline{\sigma}(M_s)\kappa(W_1)\right\} && (18.13)\\
\overline{\sigma}\left((I+GK)^{-1}G\right) &\leq \frac{\gamma\overline{\sigma}(\tilde{N}_s)}{\underline{\sigma}(W_1)\underline{\sigma}(W_2)} && (18.14)\\
\overline{\sigma}\left((I+KG)^{-1}\right) &\leq \min\left\{1+\gamma\overline{\sigma}(\tilde{N}_s)\kappa(W_1),\ \gamma\overline{\sigma}(M_s)\kappa(W_1)\right\} && (18.15)\\
\overline{\sigma}\left(G(I+KG)^{-1}K\right) &\leq \min\left\{1+\gamma\overline{\sigma}(\tilde{M}_s)\kappa(W_2),\ \gamma\overline{\sigma}(N_s)\kappa(W_2)\right\} && (18.16)
\end{aligned}$$

where

$$\overline{\sigma}(\tilde{N}_s) = \overline{\sigma}(N_s) = \left(\frac{\overline{\sigma}^2(W_2GW_1)}{1+\overline{\sigma}^2(W_2GW_1)}\right)^{1/2} \qquad (18.17)$$

$$\overline{\sigma}(\tilde{M}_s) = \overline{\sigma}(M_s) = \left(\frac{1}{1+\overline{\sigma}^2(W_2GW_1)}\right)^{1/2} \qquad (18.18)$$

and $(\tilde{N}_s, \tilde{M}_s)$*, respectively,* (N_s, M_s)*, is a normalized left coprime factorization, respectively, right coprime factorization, of* $G_s = W_2GW_1$*.*

Proof. Note that

$$\tilde{M}_s^*\tilde{M}_s = (I+G_sG_s^*)^{-1}$$

and

$$\tilde{M}_s\tilde{M}_s^* = I - \tilde{N}_s\tilde{N}_s^*.$$

Then

$$\overline{\sigma}^2(\tilde{M}_s) = \lambda_{max}(\tilde{M}_s^*\tilde{M}_s) = \frac{1}{1+\lambda_{max}(G_sG_s^*)} = \frac{1}{1+\overline{\sigma}^2(G_s)}$$

$$\overline{\sigma}^2(\tilde{N}_s) = 1-\overline{\sigma}^2(M_s) = \frac{\overline{\sigma}^2(G_s)}{1+\overline{\sigma}^2(G_s)}.$$

The proof for the normalized right coprime factorization is similar. All other inequalities follow from noting

$$\left\|\begin{bmatrix} I \\ K_\infty \end{bmatrix}(I+G_sK_\infty)^{-1}\tilde{M}_s^{-1}\right\|_\infty \leq \gamma$$

and

$$\begin{aligned}
\left\|\begin{bmatrix} I \\ K_\infty \end{bmatrix}(I+G_sK_\infty)^{-1}\tilde{M}_s^{-1}\right\|_\infty &= \left\|\begin{bmatrix} W_2 \\ W_1^{-1}K \end{bmatrix}(I+GK)^{-1}\begin{bmatrix} W_2^{-1} & GW_1 \end{bmatrix}\right\|_\infty \\
&= \left\|\begin{bmatrix} W_1^{-1} \\ W_2G \end{bmatrix}(I+KG)^{-1}\begin{bmatrix} W_1 & GW_2^{-1} \end{bmatrix}\right\|_\infty
\end{aligned}$$

□

This theorem shows that all closed-loop objectives are guaranteed to have bounded magnitude and the bounds depend only on γ, W_1, W_2, and G.

18.4 Notes and References

The $\mathcal{H}_\infty$ loop shaping using normalized coprime factorization was developed by McFarlane and Glover [1990, 1992]. In the same references, some design examples were also shown. The method has been applied to the design of scheduled controllers for a VSTOL aircraft in Hyde and Glover [1993]. The robust stabilization of normalized coprime factors is closely related to the robustness in the gap metric and graph topology, see El-Sakkary [1985], Georgiou and Smith [1990], Glover and McFarlane [1989], McFarlane, Glover, and Vidyasagar [1990], Qiu and Davison [1992a, 1992b] Vinnicombe [1993], Vidyasagar [1984, 1985], Zhu [1989], and references therein.

19

Controller Order Reduction

We have shown in the previous chapters that the $\mathcal{H}_\infty$ control theory and μ synthesis can be used to design robust performance controllers for highly complex uncertain systems. However, since a great many physical plants are modeled as high order dynamical systems, the controllers designed with these methodologies typically have orders comparable to those of the plants. Simple linear controllers are normally preferred over complex linear controllers in control system designs for some obvious reasons: they are easier to understand and computationally less demanding; they are also easier to implement and have higher reliability since there are fewer things to go wrong in the hardware or bugs to fix in the software. Therefore, a lower order controller should be sought whenever the resulting performance degradation is kept within an acceptable magnitude. There are usually three ways in arriving at a lower order controller. A seemingly obvious approach is to design lower order controllers directly based on the high order models. However, this is still largely an open research problem. The Lagrange multiplier method developed in the next chapter is potentially useful for some problems. Another approach is to first reduce the order of a high order plant, and then based on the reduced plant model a lower order controller may be designed accordingly. A potential problem associated with this approach is that such a lower order controller may not even stabilize the full order plant since the error information between the full order model and the reduced order model is not considered in the design of the controller. On the other hand, one may seek to design first a high order, high performance controller and subsequently proceed with a reduction of the designed controller. This approach is usually referred to as controller reduction. A crucial consideration in controller order reduction is to take into account the closed-loop so that the closed-loop stability is guaranteed and the performance degradation is minimized with the reduced order controllers. The purpose of this chapter is to introduce several controller reduction methods that can guarantee the closed-loop stability and possibly the closed-loop performance as well.

19.1 Controller Reduction with Stability Criteria

We consider a closed-loop system shown in Figure 19.1 where the n-th order generalized plant G is given by

$$G = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix} = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right]$$

and $G_{22} = \left[\begin{array}{c|c} A & B_2 \\ \hline C_2 & D_{22} \end{array}\right]$ is a $p \times q$ transfer matrix. Suppose K is an m-th order controller which stabilizes the closed-loop system. We are interested in investigating controller reduction methods that can preserve the closed-loop stability and minimize the performance degradation of the closed-loop systems with reduced order controllers.

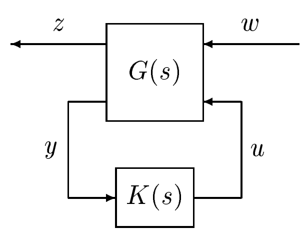


Figure 19.1: Closed-loop System Diagram

Let $\hat{K}$ be a reduced order controller and assume for the sake of argument that $\hat{K}$ has the same number of right half plane poles. Then it is obvious that the closed-loop stability is guaranteed and the closed-loop performance degradation is limited if $\left\|K - \hat{K}\right\|_\infty$ is *sufficiently* small. Hence a trivial controller reduction approach is to apply the model reduction procedure to the full order controller K. Unfortunately, this approach has only limited applications. One simple reason is that a reduced order controller that stabilizes the closed-loop system and gives satisfactory performance does not necessarily make the error $\bar{\sigma}(K - \hat{K})(j\omega)$ sufficiently small *uniformly* over all frequencies. Therefore, the approximation error only has to be made small over those critical frequency ranges that affect the closed-loop stability and performance. Since stability is the most basic requirement for a feedback system, we shall first derive controller reduction methods that guarantee this property.

19.1.1 Additive Reduction

The following lemma follows from small gain theorem.

Lemma 19.1 *Let K be a stabilizing controller and $\hat{K}$ be a reduced order controller. Suppose $\hat{K}$ and K have the same number of right half plane poles and define*

$$\Delta := \hat{K} - K, \quad W_a := (I - G_{22}K)^{-1}G_{22}.$$

Then the closed-loop system with $\hat{K}$ is stable if either

$$\|W_a\Delta\|_\infty < 1 \tag{19.1}$$

or

$$\|\Delta W_a\|_\infty < 1. \tag{19.2}$$

Proof. Since

$$I - G_{22}\hat{K} = I - G_{22}K - G_{22}\Delta = (I - G_{22}K)(I - (I - G_{22}K)^{-1}G_{22}\Delta)$$

by small gain theorem, the system is stable if $\|W_a\Delta\|_\infty < 1$. On the other hand,

$$I - \hat{K}G_{22} = I - KG_{22} - \Delta G_{22} = (I - \Delta(I - G_{22}K)^{-1}G_{22})(I - KG_{22})$$

so the system is stable if $\|\Delta W_a\|_\infty < 1$. □

Now suppose K has the following state space realization

$$K = \left[\begin{array}{c|c} A_k & B_k \\ \hline C_k & D_k \end{array}\right].$$

Then

$$\begin{aligned} W_a &= \mathcal{S}\left(\begin{bmatrix} 0 & I_p \\ I_q & K \end{bmatrix}, G_{22}\right) = \mathcal{S}\left(\left[\begin{array}{c|cc} A_k & 0 & B_k \\ \hline 0 & 0 & I_p \\ C_k & I_q & D_k \end{array}\right], \left[\begin{array}{c|c} A & B_2 \\ \hline C_2 & D_{22} \end{array}\right]\right) \\ &= \left[\begin{array}{cc|c} A + B_2D_kR^{-1}C_2 & B_2\tilde{R}^{-1}C_k & B_2\tilde{R}^{-1} \\ B_kR^{-1}C_2 & A_k + B_kD_{22}\tilde{R}^{-1}C_k & B_kD_{22}\tilde{R}^{-1} \\ \hline R^{-1}C_2 & R^{-1}D_{22}C_k & D_{22}\tilde{R}^{-1} \end{array}\right] \end{aligned}$$

where $R := I - D_{22}D_k$ and $\tilde{R} := I - D_kD_{22}$. Hence in general the order of W_a is equal to the sum of the orders of G and K.

In view of the above lemma, the controller K should be reduced in such a way so that the weighted error $\left\|W_a(K - \hat{K})\right\|_\infty$ or $\left\|(K - \hat{K})W_a\right\|_\infty$ is small and $\hat{K}$ and K have the same number of unstable poles. Suppose K is unstable, then in order to make sure that $\hat{K}$ and K have the same number of right half plane poles, K is usually separated into stable and unstable parts as

$$K = K_+ + K_-$$

where K_+ is stable and a reduction is done on K_+ to obtain a reduced $\hat{K}_+$, the final reduced order controller is given by $\hat{K} = \hat{K}_+ + K_-$.

We shall illustrate the above procedure through a simple example.

Example 19.1 Consider a system with

$$A = \begin{bmatrix} -1 & 0 & 4 \\ 0 & 2 & 0 \\ 0 & 0 & -3 \end{bmatrix}, \quad B_1 = C_1^* = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_2 = C_2^* = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$D_{11} = 0, \quad D_{12} = D_{21}^* = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad D_{22} = 0.$$

A controller minimizing $\|T_{zw}\|_2$ is given by

$$K = -\frac{148.79(s+1)(s+3)}{(s+31.74)(s+3.85)(s-9.19)}$$

with $\|T_{zw}\|_2 = 55.09$. Since K is unstable, we need to separate K into stable part and antistable part $K = K_+ + K_-$ with

$$K_+ = -\frac{114.15(s+3.61)}{(s+31.74)(s+3.85)}, \quad K_- = \frac{-34.64}{s-9.19}.$$

Next apply the frequency weighted balanced model reduction in Chapter 7 to

$$\left\| W_a(K_+ - \hat{K}_+) \right\|_\infty,$$

we have

$$\hat{K}_+ = -\frac{117.085}{s+34.526}$$

and

$$\hat{K} := \hat{K}_+ + K_- = -\frac{151.72(s+0.788)}{(s+34.526)(s-9.19)}.$$

The $\|T_{zw}\|_2$ with the reduced order controller $\hat{K}$ is 61.69. On the other hand, if K_+ is reduced directly by balanced truncation without stability weighting W_a, then the reduced order controller does not stabilize the closed-loop system. The results are summarized in Table 19.1 where both weighted and unweighted errors are listed.

Methods	$\left\|K-\hat{K}\right\|_\infty$	$\left\|W_a(K-\hat{K})\right\|_\infty$	$\|T_{zw}\|_2$
BT	0.1622	2.5295	unstable
WBT	0.1461	0.471	61.69

Table 19.1: BT: Balance and Truncate, WBT: Weighted Balance and Truncate

It may seem somewhat strange that the unweighted error $\left\|K-\hat{K}\right\|_\infty$ resulted from weighted balanced reduction is actually smaller than that from unweighted balanced reduction. This happens because the balanced model reduction is not optimal in $\mathcal{L}_\infty$ norm. We should also point out that the stability criterion $\left\|W_a(K-\hat{K})\right\|_\infty < 1$ (or $\left\|(K-\hat{K})W_a\right\|_\infty < 1$) is only sufficient. Hence having $\left\|W_a(K-\hat{K})\right\|_\infty \geq 1$ does not necessarily imply that $\hat{K}$ is not a stabilizing controller. $\diamond$

19.1.2 Coprime Factor Reduction

It is clear that the additive controller reduction method in the last section is somewhat restrictive, in particular, this method can not be used to reduce the controller order if the controller is totally unstable, i.e., K has all poles in the right half plane. This motivates the following coprime factorization reduction approach.

Let G_{22} and K have the following left and right coprime factorizations, respectively

$$G_{22} = \tilde{M}^{-1}\tilde{N} = NM^{-1}, \qquad K = \tilde{V}^{-1}\tilde{U} = UV^{-1}$$

and define

$$\begin{bmatrix} \tilde{N}_n & \tilde{M}_n \end{bmatrix} := (\tilde{M}V - \tilde{N}U)^{-1}\begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} = V^{-1}(I - G_{22}K)^{-1}\begin{bmatrix} G_{22} & I \end{bmatrix}$$

and

$$\begin{bmatrix} N_n \\ M_n \end{bmatrix} := \begin{bmatrix} N \\ M \end{bmatrix}(\tilde{V}M - \tilde{U}N)^{-1} = \begin{bmatrix} G_{22} \\ I \end{bmatrix}(I - KG_{22})^{-1}\tilde{V}^{-1}.$$

Note that $\tilde{M}_n, \tilde{N}_n, M_n, N_n$ do not depend upon the specific coprime factorizations of G_{22}.

Lemma 19.2 *Let $\hat{U}, \hat{V} \in \mathcal{RH}_\infty$ be the reduced right coprime factors of U and V. Then $\hat{K} := \hat{U}\hat{V}^{-1}$ stabilizes the system if*

$$\left\|\begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix}\left(\begin{bmatrix} U \\ V \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix}\right)\right\|_\infty < 1. \tag{19.3}$$

Similarly, let $\hat{\tilde{U}}, \hat{\tilde{V}} \in \mathcal{RH}_\infty$ *be the reduced left coprime factors of* $\tilde{U}$ *and* $\tilde{V}$. *Then* $\hat{K} := \hat{\tilde{V}}^{-1}\hat{\tilde{U}}$ *stabilizes the system if*

$$\left\| \left(\begin{bmatrix} \tilde{U} & \tilde{V} \end{bmatrix} - \begin{bmatrix} \hat{\tilde{U}} & \hat{\tilde{V}} \end{bmatrix} \right) \begin{bmatrix} -N_n \\ M_n \end{bmatrix} \right\|_\infty < 1. \tag{19.4}$$

Proof. We shall only show the results for the right coprime controller reduction. The case for the left coprime factorization is analogous. It is well known that $\hat{K} := \hat{U}\hat{V}^{-1}$ stabilizes the system if and only if $(\tilde{M}\hat{V} - \tilde{N}\hat{U})^{-1} \in \mathcal{RH}_\infty$. Since

$$\tilde{M}\hat{V} - \tilde{N}\hat{U} = (\tilde{M}V - \tilde{N}U)\left[I - \begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix} \begin{bmatrix} U - \hat{U} \\ V - \hat{V} \end{bmatrix} \right]$$

the stability of the closed-loop system is guaranteed if

$$\left\| \begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix} \begin{bmatrix} U - \hat{U} \\ V - \hat{V} \end{bmatrix} \right\|_\infty < 1.$$

□

Now suppose $\hat{U}$ and $\hat{V}$ have the following state space realizations

$$\begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C}_1 & \hat{D}_1 \\ \hat{C}_2 & \hat{D}_2 \end{array} \right] \tag{19.5}$$

and suppose $\hat{D}_2$ is nonsingular. Then the reduced order controller is given by

$$\hat{K} = \left[\begin{array}{c|c} \hat{A} - \hat{B}\hat{D}_2^{-1}\hat{C}_2 & \hat{B}\hat{D}_2^{-1} \\ \hline \hat{C}_1 - \hat{D}_1\hat{D}_2^{-1}\hat{C}_2 & \hat{D}_1\hat{D}_2^{-1} \end{array} \right]. \tag{19.6}$$

Similarly, suppose $\hat{\tilde{U}}$ and $\hat{\tilde{V}}$ have the following state space realizations

$$\begin{bmatrix} \hat{\tilde{U}} & \hat{\tilde{V}} \end{bmatrix} = \left[\begin{array}{c|cc} \hat{A} & \hat{B}_1 & \hat{B}_2 \\ \hline \hat{C} & \hat{D}_1 & \hat{D}_2 \end{array} \right] \tag{19.7}$$

and suppose $\hat{D}_2$ is nonsingular. Then the reduced order controller is given by

$$\hat{K} = \left[\begin{array}{c|c} \hat{A} - \hat{B}_2\hat{D}_2^{-1}\hat{C} & \hat{B}_1 - \hat{B}_2\hat{D}_2^{-1}\hat{D}_1 \\ \hline \hat{D}_2^{-1}\hat{C} & \hat{D}_2^{-1}\hat{D}_1 \end{array} \right]. \tag{19.8}$$

It is clear from this lemma that the coprime factors of the controller should be reduced so that the weighted errors in (19.3) and (19.4) are small. Note that there is no restriction on the right half plane poles of the controller. In particular, K and $\hat{K}$ may have different number of right half plane poles. It is also not hard to see that the additive reduction method in the last subsection may be regarded as a special case of the coprime factor reduction if the controller K is stable by taking $V = I$ or $\tilde{V} = I$.

Let L, F, L_k and F_k be any matrices such that $A + LC_2, A + B_2F, A_k + L_kC_k$ and $A_k + B_kF_k$ are stable. Define

$$\begin{bmatrix} \tilde{N} & \tilde{M} \end{bmatrix} = \left[\begin{array}{c|cc} A + LC_2 & B_2 + LD_{22} & L \\ \hline C_2 & D_{22} & I \end{array}\right], \quad \begin{bmatrix} N \\ M \end{bmatrix} = \left[\begin{array}{c|c} A + B_2F & B_2 \\ \hline C_2 + D_{22}F & D_{22} \\ F & I \end{array}\right],$$

$$\begin{bmatrix} U \\ V \end{bmatrix} = \left[\begin{array}{c|c} A_k + B_kF_k & B_k \\ \hline C_k + D_kF_k & D_k \\ F_k & I \end{array}\right], \quad \begin{bmatrix} \tilde{U} & \tilde{V} \end{bmatrix} = \left[\begin{array}{c|cc} A_k + L_kC_k & B_k + L_kD_k & L_k \\ \hline C_k & D_k & I \end{array}\right].$$

Then $K = UV^{-1} = \tilde{V}^{-1}\tilde{U}$ and $G_{22} = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ are right and left coprime factorizations over $\mathcal{RH}_\infty$, respectively. Moreover,

$$\begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix} = \left[\begin{array}{cc|cc} A + B_2D_kR^{-1}C_2 & -B_2\tilde{R}^{-1}C_k & B_2\tilde{R}^{-1} & -B_2D_kR^{-1} \\ -B_kR^{-1}C_2 & A_k + B_kR^{-1}D_{22}C_k & -B_kR^{-1}D_{22} & B_kR^{-1} \\ \hline -R^{-1}C_2 & R^{-1}D_{22}C_k - F_k & -R^{-1}D_{22} & R^{-1} \end{array}\right]$$

$$\begin{bmatrix} -N_n \\ M_n \end{bmatrix} = \left[\begin{array}{cc|c} A_k + B_kD_{22}\tilde{R}^{-1}C_k & -B_kR^{-1}C_2 & -L_k + B_kD_{22}\tilde{R}^{-1} \\ -B_2\tilde{R}^{-1}C_k & A + B_2\tilde{R}^{-1}D_kC_2 & -B_2\tilde{R}^{-1} \\ \hline -D_{22}\tilde{R}^{-1}C_k & R^{-1}C_2 & -D_{22}\tilde{R}^{-1} \\ \tilde{R}^{-1}C_k & -\tilde{R}^{-1}D_kC_2 & \tilde{R}^{-1} \end{array}\right]$$

where $R := I - D_{22}D_k$ and $\tilde{R} := I - D_kD_{22}$. Note that the orders of the weighting matrices $\begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix}$ and $\begin{bmatrix} -N_n \\ M_n \end{bmatrix}$ are in general equal to the sum of the orders of G and K. However, if K is an observer-based controller, i.e.,

$$K = \left[\begin{array}{c|c} A + B_2F + LC_2 + LD_{22}F & -L \\ \hline F & 0 \end{array}\right],$$

letting $F_k = -(C_2 + D_{22}F)$ and $L_k = -(B_2 + LD_{22})$, we get

$$\begin{bmatrix} U \\ V \end{bmatrix} = \left[\begin{array}{c|c} A + B_2F & -L \\ \hline F & 0 \\ C_2 + D_{22}F & I \end{array}\right], \quad \begin{bmatrix} \tilde{U} & \tilde{V} \end{bmatrix} = \left[\begin{array}{c|cc} A + LC_2 & -L & -(B_2 + LD_{22}) \\ \hline F & 0 & I \end{array}\right],$$

$$\tilde{M}V - \tilde{N}U = I, \quad \tilde{V}M - \tilde{U}N = I.$$

Therefore, we can chose $\begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix} = \begin{bmatrix} -\tilde{N} & \tilde{M} \end{bmatrix}$ and $\begin{bmatrix} -N_n \\ M_n \end{bmatrix} = \begin{bmatrix} -N \\ M \end{bmatrix}$ which have the same orders as that of the plant G.

We shall also illustrate the above procedure through a simple example.

Example 19.2 Consider a system with

$$A = \begin{bmatrix} -1 & 0 & 4 \\ 0 & -2 & 0 \\ 0 & 0 & -3 \end{bmatrix}, \quad B_1 = C_1^* = \begin{bmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad B_2 = C_2^* = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

$$D_{11} = 0, \quad D_{12} = D_{21}^* = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad D_{22} = 0.$$

A controller minimizing $\|T_{zw}\|_2$ is given by

$$K = \left[\begin{array}{ccc|c} -1 & -8.198 & 4 & 0 \\ -8.198 & -18.396 & -8.198 & 8.198 \\ 0 & -8.198 & -3 & 0 \\ \hline 0 & -8.198 & 0 & 0 \end{array}\right] = -\frac{67.2078(s+1)(s+3)}{(s+23.969)(s+3.7685)(s-5.3414)}$$

with $\|T_{zw}\|_2 = 37.02$. Since the controller is an observer-based controller, a natural coprime factorization of K is given by

$$\begin{bmatrix} U \\ V \end{bmatrix} = \left[\begin{array}{ccc|c} -1 & -8.198 & 4 & 0 \\ 0 & -10.198 & 0 & 8.198 \\ 0 & -8.198 & -3 & 0 \\ \hline 0 & -8.198 & 0 & 0 \\ 1 & 1 & 1 & 1 \end{array}\right].$$

Furthermore, we have

$$\begin{bmatrix} -\tilde{N} & \tilde{M} \end{bmatrix} = \left[\begin{array}{ccc|cc} -1 & 0 & 4 & -1 & 0 \\ -8.198 & -10.198 & -8.198 & -1 & -8.198 \\ 0 & 0 & -3 & -1 & 0 \\ \hline 1 & 1 & 1 & 0 & 0 \end{array}\right].$$

Applying the frequency weighted balanced model reduction in Chapter 7 to the weighted coprime factors, we obtain a first order approximation

$$\begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} = \left[\begin{array}{c|c} -0.1814 & 1.0202 \\ \hline 1.2244 & 0 \\ 6.504 & 1 \end{array}\right]$$

which gives

$$\left\| \begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix} \left(\begin{bmatrix} U \\ V \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right) \right\|_\infty = 2.0928 > 1,$$

$$\hat{K} = \frac{1.2491}{s + 6.8165}, \quad \|T_{zw}\|_2 = 52.29$$

and a second order approximation

$$\begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} = \left[\begin{array}{cc|c} -0.1814 & 14.5511 & 1.0202 \\ -0.5288 & -4.1434 & -1.2642 \\ \hline 1.2244 & 29.5182 & 0 \\ 6.504 & -1.3084 & 1 \end{array}\right]$$

which gives

$$\left\| \begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix} \left(\begin{bmatrix} U \\ V \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right) \right\|_\infty = 0.024 < 1,$$

$$\hat{K} = -\frac{36.069(s + 1.1102)}{(s + 17.3741)(s - 4.7601)}, \quad \|T_{zw}\|_2 = 39.14.$$

Note that the first order reduced controller does not have the same number of right half plane poles as that of the full order controller. Moreover, the sufficient stability condition is not satisfied nevertheless the controller is a stabilizing controller. It is also interesting to note that the unstable pole of the second order reduced controller is not at the same location as that of the full order controller. $\diamond$

19.2 $\mathcal{H}_\infty$ Controller Reductions

In this section, we consider $\mathcal{H}_\infty$ performance preserving controller order reduction problem. Again we consider the feedback system shown in Figure 19.1 with a generalized plant realization given by

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & D_{22} \end{array}\right].$$

The following assumptions are made:

(A1) (A, B_2) is stabilizable and (C_2, A) is detectable;

(A2) D_{12} has full column rank and D_{21} has full row rank;

(A3) $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all ω;

(A4) $\begin{bmatrix} A - j\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ has full row rank for all ω.

It is shown in Chapter 17 that all stabilizing controllers satisfying $\|T_{zw}\|_\infty < \gamma$ can be parameterized as

$$K = \mathcal{F}_\ell(M_\infty, Q), \quad Q \in \mathcal{RH}_\infty, \quad \|Q\|_\infty < \gamma \tag{19.9}$$

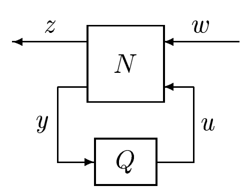
where M_∞ is of the form

$$M_\infty = \begin{bmatrix} M_{11}(s) & M_{12}(s) \\ M_{21}(s) & M_{22}(s) \end{bmatrix} = \left[\begin{array}{c|cc} \hat{A} & \hat{B}_1 & \hat{B}_2 \\ \hline \hat{C}_1 & \hat{D}_{11} & \hat{D}_{12} \\ \hat{C}_2 & \hat{D}_{21} & \hat{D}_{22} \end{array}\right]$$

such that $\hat{D}_{12}$ and $\hat{D}_{21}$ are invertible and $\hat{A} - \hat{B}_2\hat{D}_{12}^{-1}\hat{C}_1$ and $\hat{A} - \hat{B}_1\hat{D}_{21}^{-1}\hat{C}_2$ are both stable, i.e., M_{12}^{-1} and M_{21}^{-1} are both stable.

The problem to be considered here is to find a controller $\hat{K}$ with a minimal possible order such that the $\mathcal{H}_\infty$ performance requirement $\left\|\mathcal{F}_\ell(G, \hat{K})\right\|_\infty < \gamma$ is satisfied. This is clearly equivalent to finding a Q so that it satisfies the above constraint and the order of $\hat{K}$ is minimized. Instead of choosing Q directly, we shall approach this problem from a different perspective. The following lemma is useful in the subsequent development and can be regarded as a special case of Theorem 11.7 (main loop theorem).

Lemma 19.3 *Consider a feedback system shown below*



where N is a suitably partitioned transfer matrix

$$N(s) = \begin{bmatrix} N_{11} & N_{12} \\ N_{21} & N_{22} \end{bmatrix}.$$

Then, the closed-loop transfer matrix from w to z is given by

$$T_{zw} = \mathcal{F}_\ell(N, Q) = N_{11} + N_{12}Q(I - N_{22}Q)^{-1}N_{21}.$$

Assume that the feedback loop is well-posed, i.e., $\det(I - N_{22}(\infty)Q(\infty)) \neq 0$*, and either $N_{21}(j\omega)$ has full row rank for all $\omega \in \mathbb{R} \cup \infty$ or $N_{12}(j\omega)$ has full column rank for all $\omega \in \mathbb{R} \cup \infty$ and $\|N\|_\infty \leq 1$ then $\|\mathcal{F}_\ell(N, Q)\|_\infty < 1$ if $\|Q\|_\infty < 1$.*

Proof. We shall assume N_{21} has full row rank. The case when N_{12} has full column rank can be shown in the same way.

To show that $\|T_{zw}\|_\infty < 1$, consider the closed-loop system at any frequency $s = j\omega$ with the signals fixed as complex constant vectors. Let $\|Q\|_\infty =: \epsilon < 1$ and note that $T_{wy} = N_{21}^+(I - N_{22}Q)$ where N_{21}^+ is a right inverse of N_{21}. Also let $\kappa := \|T_{wy}\|_\infty$. Then $\|w\|_2 \le \kappa\|y\|_2$, and $\|N\|_\infty \le 1$ implies that $\|z\|_2^2 + \|y\|_2^2 \le \|w\|_2^2 + \|u\|_2^2$. Therefore,

$$\|z\|_2^2 \le \|w\|_2^2 + (\epsilon^2 - 1)\|y\|_2^2 \le [1 - (1 - \epsilon^2)\kappa^{-2}]\|w\|_2^2$$

which implies $\|T_{zw}\|_\infty < 1$. □

19.2.1 Additive Reduction

Consider the class of (reduced order) controllers that can be represented in the form

$$\hat{K} = K_0 + W_2 \Delta W_1,$$

where K_0 may be interpreted as a nominal, higher order controller, Δ is a stable perturbation, with stable, minimum phase, and invertible weighting functions W_1 and W_2. Suppose that $\left\|\mathcal{F}_\ell(G, K_0)\right\|_\infty < \gamma$. A natural question is whether it is possible to obtain a reduced order controller $\hat{K}$ in this class such that the $\mathcal{H}_\infty$ performance bound remains valid when $\hat{K}$ is in place of K_0. Note that this is somewhat a special case of the above general problem; the specific form of $\hat{K}$ restricts that $\hat{K}$ and K_0 must possess the same right half plane poles, thus to a certain degree limiting the set of attainable reduced order controllers.

Suppose $\hat{K}$ is a suboptimal $\mathcal{H}_\infty$ controller, i.e., there is a $Q \in \mathcal{RH}_\infty$ with $\left\|Q\right\|_\infty < \gamma$ such that $\hat{K} = \mathcal{F}_\ell(M_\infty, Q)$. It follows from simple algebra that

$$Q = \mathcal{F}_\ell(\bar{K}_a^{-1}, \hat{K})$$

where

$$\bar{K}_a^{-1} := \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix} M_\infty^{-1} \begin{bmatrix} 0 & I \\ I & 0 \end{bmatrix}.$$

Furthermore, it follows from straightforward manipulations that

$$\begin{aligned} \left\|Q\right\|_\infty < \gamma \quad &\Longleftrightarrow \quad \left\|\mathcal{F}_\ell(\bar{K}_a^{-1}, \hat{K})\right\|_\infty < \gamma \\ &\Longleftrightarrow \quad \left\|\mathcal{F}_\ell(\bar{K}_a^{-1}, K_0 + W_2\Delta W_1)\right\|_\infty < \gamma \\ &\Longleftrightarrow \quad \left\|\mathcal{F}_\ell(\tilde{R}, \Delta)\right\|_\infty < 1 \end{aligned}$$

where

$$\tilde{R} = \begin{bmatrix} \gamma^{-1/2} I & 0 \\ 0 & W_1 \end{bmatrix} \begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} \begin{bmatrix} \gamma^{-1/2} I & 0 \\ 0 & W_2 \end{bmatrix}$$

and R is given by the star product

$$\begin{bmatrix} R_{11} & R_{12} \\ R_{21} & R_{22} \end{bmatrix} = \mathcal{S}(\bar{K}_a^{-1}, \begin{bmatrix} K_o & I \\ I & 0 \end{bmatrix}).$$

It is easy to see that $\tilde{R}_{12}$ and $\tilde{R}_{21}$ are both minimum phase and invertible, and hence have full column and full row rank, respectively for all $\omega \in \mathbb{R} \cup \infty$. Consequently, by invoking Lemma 19.3, we conclude that if $\tilde{R}$ is a contraction and $\|\Delta\|_\infty < 1$ then $\left\|\mathcal{F}_\ell(\tilde{R}, \Delta)\right\|_\infty < 1$. This guarantees the existence of a Q such that $\|Q\|_\infty < \gamma$, or equivalently, the existence of a $\hat{K}$ such that $\left\|\mathcal{F}_\ell(G, \hat{K})\right\|_\infty < \gamma$. This observation leads to the following theorem.

Theorem 19.4 *Suppose W_1 and W_2 are stable, minimum phase and invertible transfer matrices such that $\tilde{R}$ is a contraction. Let K_0 be a stabilizing controller such that $\|\mathcal{F}_\ell(G, K_0)\|_\infty < \gamma$. Then $\hat{K}$ is also a stabilizing controller such that $\left\|\mathcal{F}_\ell(G, \hat{K})\right\|_\infty < \gamma$ if*

$$\|\Delta\|_\infty = \left\|W_2^{-1}(\hat{K} - K_0)W_1^{-1}\right\|_\infty < 1.$$

Since $\tilde{R}$ can always be made contractive for sufficiently small W_1 and W_2, there are infinite many W_1 and W_2 that satisfy the conditions in the theorem. It is obvious that to make $\left\|W_2^{-1}(\hat{K} - K_0)W_1^{-1}\right\|_\infty < 1$ for some $\hat{K}$, one would like to select the "largest" W_1 and W_2.

Lemma 19.5 *Assume $\|R_{22}\|_\infty < \gamma$ and define*

$$L = \begin{bmatrix} L_1 & L_2 \\ L_2^\sim & L_3 \end{bmatrix} = \mathcal{F}_\ell(\left[\begin{array}{cc:cc} 0 & -R_{11} & 0 & R_{12} \\ -R_{11}^\sim & 0 & R_{21}^\sim & 0 \\ \hdashline 0 & R_{21} & 0 & -R_{22} \\ R_{12}^\sim & 0 & -R_{22}^\sim & 0 \end{array}\right], \gamma^{-1} I).$$

Then $\tilde{R}$ is a contraction if W_1 and W_2 satisfy

$$\begin{bmatrix} (W_1^\sim W_1)^{-1} & 0 \\ 0 & (W_2 W_2^\sim)^{-1} \end{bmatrix} \geq \begin{bmatrix} L_1 & L_2 \\ L_2^\sim & L_3 \end{bmatrix}.$$

Proof. See Goddard and Glover [1993]. □

An algorithm that maximizes $\det(W_1^\sim W_1)\det(W_2 W_2^\sim)$ has been developed by Goddard and Glover [1993]. The procedure below, devised directly from the above theorem, can be used to generate a required reduced order controller which will preserve the closed-loop $\mathcal{H}_\infty$ performance bound $\left\|\mathcal{F}_\ell(G, \hat{K})\right\|_\infty < \gamma$.

1. Let K_0 be a full order controller such that $\|\mathcal{F}_\ell(G, K_0)\|_\infty < \gamma$;
2. Compute W_1 and W_2 so that $\tilde{R}$ is a contraction;
3. Using model reduction method to find a $\hat{K}$ so that $\left\| W_2^{-1}(\hat{K} - K_0)W_1^{-1} \right\|_\infty < 1$.

19.2.2 Coprime Factor Reduction

The $\mathcal{H}_\infty$ controller reduction problem can also be considered in the coprime factor framework. For that purpose, we need the following alternative representation of all admissible $\mathcal{H}_\infty$ controllers.

Lemma 19.6 *The family of all admissible controllers such that $\|T_{zw}\|_\infty < \gamma$ can also be written as*

$$\begin{aligned} K(s) = \mathcal{F}_\ell(M_\infty, Q) &= (\Theta_{11}Q + \Theta_{12})(\Theta_{21}Q + \Theta_{22})^{-1} := UV^{-1} \\ &= (Q\tilde{\Theta}_{12} + \tilde{\Theta}_{22})^{-1}(Q\tilde{\Theta}_{11} + \tilde{\Theta}_{21}) := \tilde{V}^{-1}\tilde{U} \end{aligned}$$

where $Q \in \mathcal{RH}_\infty$, $\|Q\|_\infty < \gamma$, and UV^{-1} and $\tilde{V}^{-1}\tilde{U}$ are respectively right and left coprime factorizations over $\mathcal{RH}_\infty$, and

$$\Theta = \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} = \left[\begin{array}{c|cc} \hat{A} - \hat{B}_1\hat{D}_{21}^{-1}\hat{C}_2 & \hat{B}_2 - \hat{B}_1\hat{D}_{21}^{-1}\hat{D}_{22} & \hat{B}_1\hat{D}_{21}^{-1} \\ \hline \hat{C}_1 - \hat{D}_{11}\hat{D}_{21}^{-1}\hat{C}_2 & \hat{D}_{12} - \hat{D}_{11}\hat{D}_{21}^{-1}\hat{D}_{22} & \hat{D}_{11}\hat{D}_{21}^{-1} \\ -\hat{D}_{21}^{-1}\hat{C}_2 & -\hat{D}_{21}^{-1}\hat{D}_{22} & \hat{D}_{21}^{-1} \end{array}\right]$$

$$\tilde{\Theta} = \begin{bmatrix} \tilde{\Theta}_{11} & \tilde{\Theta}_{12} \\ \tilde{\Theta}_{21} & \tilde{\Theta}_{22} \end{bmatrix} = \left[\begin{array}{c|cc} \hat{A} - \hat{B}_2\hat{D}_{12}^{-1}\hat{C}_1 & \hat{B}_1 - \hat{B}_2\hat{D}_{12}^{-1}\hat{D}_{11} & -\hat{B}_2\hat{D}_{12}^{-1} \\ \hline \hat{C}_2 - \hat{D}_{22}\hat{D}_{12}^{-1}\hat{C}_1 & \hat{D}_{21} - \hat{D}_{22}\hat{D}_{12}^{-1}\hat{D}_{11} & -\hat{D}_{22}\hat{D}_{12}^{-1} \\ \hat{D}_{12}^{-1}\hat{C}_1 & \hat{D}_{12}^{-1}\hat{D}_{11} & \hat{D}_{12}^{-1} \end{array}\right]$$

$$\Theta^{-1} = \left[\begin{array}{c|cc} \hat{A} - \hat{B}_2\hat{D}_{12}^{-1}\hat{C}_1 & \hat{B}_2\hat{D}_{12}^{-1} & \hat{B}_1 - \hat{B}_2\hat{D}_{12}^{-1}\hat{D}_{11} \\ \hline -\hat{D}_{12}^{-1}\hat{C}_1 & \hat{D}_{12}^{-1} & -\hat{D}_{12}^{-1}\hat{D}_{11} \\ \hat{C}_2 - \hat{D}_{22}\hat{D}_{12}^{-1}\hat{C}_1 & \hat{D}_{22}\hat{D}_{12}^{-1} & \hat{D}_{21} - \hat{D}_{22}\hat{D}_{12}^{-1}\hat{D}_{11} \end{array}\right]$$

$$\tilde{\Theta}^{-1} = \left[\begin{array}{c|cc} \hat{A} - \hat{B}_1\hat{D}_{21}^{-1}\hat{C}_2 & -\hat{B}_1\hat{D}_{21}^{-1} & \hat{B}_2 - \hat{B}_1\hat{D}_{21}^{-1}\hat{D}_{22} \\ \hline \hat{D}_{21}^{-1}\hat{C}_2 & \hat{D}_{21}^{-1} & \hat{D}_{21}^{-1}\hat{D}_{22} \\ \hat{C}_1 - \hat{D}_{11}\hat{D}_{21}^{-1}\hat{C}_2 & -\hat{D}_{11}\hat{D}_{21}^{-1} & \hat{D}_{12} - \hat{D}_{11}\hat{D}_{21}^{-1}\hat{D}_{22} \end{array}\right].$$

Proof. Note that all admissible $\mathcal{H}_\infty$ controllers are given by

$$K(s) = \mathcal{F}_\ell(M_\infty, Q)$$

where $M_\infty = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix}$ is given in the last subsection. Next note that from Lemma 10.2, we have

$$\Theta = \begin{bmatrix} \Theta_{11} & \Theta_{12} \\ \Theta_{21} & \Theta_{22} \end{bmatrix} = \begin{bmatrix} M_{12} - M_{11}M_{21}^{-1}M_{22} & M_{11}M_{21}^{-1} \\ -M_{21}^{-1}M_{22} & M_{21}^{-1} \end{bmatrix}$$

$$\tilde{\Theta} = \begin{bmatrix} \tilde{\Theta}_{11} & \tilde{\Theta}_{12} \\ \tilde{\Theta}_{21} & \tilde{\Theta}_{22} \end{bmatrix} = \begin{bmatrix} M_{21} - M_{22}M_{12}^{-1}M_{11} & -M_{22}M_{12}^{-1} \\ M_{12}^{-1}M_{11} & M_{12}^{-1} \end{bmatrix}$$

$$\Theta^{-1} = \begin{bmatrix} M_{12}^{-1} & -M_{12}^{-1}M_{11} \\ M_{22}M_{12}^{-1} & M_{21} - M_{22}M_{12}^{-1}M_{11} \end{bmatrix}$$

$$\tilde{\Theta}^{-1} = \begin{bmatrix} M_{21}^{-1} & M_{21}^{-1}M_{22} \\ -M_{11}M_{21}^{-1} & M_{12} - M_{11}M_{21}^{-1}M_{22} \end{bmatrix}.$$

Then the results follow from some algebra. □

Theorem 19.7 *Let $K_0 = \Theta_{12}\Theta_{22}^{-1}$ be a central $\mathcal{H}_\infty$ controller such that $\|\mathcal{F}_\ell(G, K_0)\|_\infty < \gamma$ and let $\hat{U}, \hat{V} \in \mathcal{RH}_\infty$ be such that*

$$\left\| \begin{bmatrix} \gamma^{-1}I & 0 \\ 0 & I \end{bmatrix} \Theta^{-1} \left(\begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right) \right\|_\infty < 1/\sqrt{2}. \tag{19.10}$$

Then $\hat{K} = \hat{U}\hat{V}^{-1}$ is also a stabilizing controller such that $\|\mathcal{F}_\ell(G, \hat{K})\|_\infty < \gamma$.

Proof. Note that by Lemma 19.6, K is an admissible controller such that $\|T_{zw}\|_\infty < \gamma$ if and only if there exists a $Q \in \mathcal{RH}_\infty$ with $\|Q\|_\infty < \gamma$ such that

$$\begin{bmatrix} U \\ V \end{bmatrix} := \begin{bmatrix} \Theta_{11}Q + \Theta_{12} \\ \Theta_{21}Q + \Theta_{22} \end{bmatrix} = \Theta \begin{bmatrix} Q \\ I \end{bmatrix} \tag{19.11}$$

and

$$K = UV^{-1}.$$

Hence, to show that $\hat{K} = \hat{U}\hat{V}^{-1}$ with $\hat{U}$ and $\hat{V}$ satisfying equation (19.10) is also a stabilizing controller such that $\|\mathcal{F}_\ell(G, \hat{K})\|_\infty < \gamma$, we need to show that there is another coprime factorization for $\hat{K} = UV^{-1}$ and a $Q \in \mathcal{RH}_\infty$ with $\|Q\|_\infty < \gamma$ such that equation (19.11) is satisfied.

Define

$$\Delta := \begin{bmatrix} \gamma^{-1}I & 0 \\ 0 & I \end{bmatrix} \Theta^{-1} \left(\begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right)$$

and partition Δ as

$$\Delta := \begin{bmatrix} \Delta_U \\ \Delta_V \end{bmatrix}.$$

Then

$$\begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} = \begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \Theta \begin{bmatrix} \gamma I & 0 \\ 0 & I \end{bmatrix} \Delta = \Theta \begin{bmatrix} -\gamma\Delta_U \\ I - \Delta_V \end{bmatrix}$$

and

$$\begin{bmatrix} \hat{U}(I-\Delta_V)^{-1} \\ \hat{V}(I-\Delta_V)^{-1} \end{bmatrix} = \Theta \begin{bmatrix} -\gamma\Delta_U(I-\Delta_V)^{-1} \\ I \end{bmatrix}.$$

Define $U := \hat{U}(I-\Delta_V)^{-1}$, $V := \hat{V}(I-\Delta_V)^{-1}$ and $Q := -\gamma\Delta_U(I-\Delta_V)^{-1}$. Then UV^{-1} is another coprime factorization for $\hat{K}$. To show that $\hat{K} = UV^{-1} = \hat{U}\hat{V}^{-1}$ is a stabilizing controller such that $\|\mathcal{F}_\ell(G,\hat{K})\|_\infty < \gamma$, we need to show that $\left\|\gamma\Delta_U(I-\Delta_V)^{-1}\right\|_\infty < \gamma$, or equivalently $\left\|\Delta_U(I-\Delta_V)^{-1}\right\|_\infty < 1$. Now

$$\begin{aligned} \Delta_U(I-\Delta_V)^{-1} &= \begin{bmatrix} I & 0 \end{bmatrix} \Delta \left(I - \begin{bmatrix} 0 & I \end{bmatrix} \Delta\right)^{-1} \\ &= \mathcal{F}_\ell\left(\begin{bmatrix} 0 & \begin{bmatrix} I & 0 \end{bmatrix} \\ I/\sqrt{2} & \begin{bmatrix} 0 & I/\sqrt{2} \end{bmatrix} \end{bmatrix}, \sqrt{2}\Delta \right) \end{aligned}$$

and by Lemma 19.3 $\left\|\Delta_U(I-\Delta_V)^{-1}\right\|_\infty < 1$ since

$$\begin{bmatrix} 0 & \begin{bmatrix} I & 0 \end{bmatrix} \\ I/\sqrt{2} & \begin{bmatrix} 0 & I/\sqrt{2} \end{bmatrix} \end{bmatrix}$$

is a contraction and $\left\|\sqrt{2}\Delta\right\|_\infty < 1$. □

Similarly, we have the following theorem.

Theorem 19.8 *Let $K_0 = \tilde{\Theta}_{22}^{-1}\tilde{\Theta}_{21}$ be a central $\mathcal{H}_\infty$ controller such that $\|\mathcal{F}_\ell(G,K_0)\|_\infty < \gamma$ and let $\hat{\tilde{U}}, \hat{\tilde{V}} \in \mathcal{RH}_\infty$ be such that*

$$\left\| \left(\begin{bmatrix} \tilde{\Theta}_{21} \\ \tilde{\Theta}_{22} \end{bmatrix} - \begin{bmatrix} \hat{\tilde{U}} \\ \hat{\tilde{V}} \end{bmatrix} \right) \tilde{\Theta}^{-1} \begin{bmatrix} \gamma^{-1} I & 0 \\ 0 & I \end{bmatrix} \right\|_\infty < 1/\sqrt{2}.$$

Then $\hat{K} = \hat{\tilde{V}}^{-1}\hat{\tilde{U}}$ is also a stabilizing controller such that $\|\mathcal{F}_\ell(G,\hat{K})\|_\infty < \gamma$.

The above two theorems show that the $\mathcal{H}_\infty$ controller reduction problem is equivalent to a frequency weighted $\mathcal{H}_\infty$ model reduction problem.

$\mathcal{H}_\infty$ Controller Reduction Procedures

(i) Let $K_0 = \Theta_{12}\Theta_{22}^{-1}(= \tilde{\Theta}_{22}^{-1}\tilde{\Theta}_{21})$ be a suboptimal $\mathcal{H}_\infty$ central controller ($Q = 0$) such that $\|T_{zw}\|_\infty < \gamma$.

(ii) Find a reduced order controller $\hat{K} = \hat{U}\hat{V}^{-1}$ (or $\hat{\tilde{V}}^{-1}\hat{\tilde{U}}$) such that the following frequency weighted $\mathcal{H}_\infty$ error

$$\left\| \begin{bmatrix} \gamma^{-1}I & 0 \\ 0 & I \end{bmatrix} \Theta^{-1} \left(\begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right) \right\|_\infty < 1/\sqrt{2}$$

or

$$\left\| \left(\begin{bmatrix} \tilde{\Theta}_{21} \\ \tilde{\Theta}_{22} \end{bmatrix} - \begin{bmatrix} \hat{\tilde{U}} \\ \hat{\tilde{V}} \end{bmatrix} \right) \tilde{\Theta}^{-1} \begin{bmatrix} \gamma^{-1}I & 0 \\ 0 & I \end{bmatrix} \right\|_\infty < 1/\sqrt{2}.$$

(iii) The closed-loop system with the reduced order controller $\hat{K}$ is stable and the performance is maintained with the reduced order controller, i.e.,

$$\|T_{zw}\|_\infty = \left\| \mathcal{F}_\ell(G, \hat{K}) \right\|_\infty < \gamma.$$

19.3 Frequency-Weighted $\mathcal{L}_\infty$ Norm Approximations

We have shown in the previous sections that controller reduction problems are equivalent to frequency weighted model reduction problems. To that end, the frequency weighted balanced model reduction approach in Chapter 7 can be applied. In this section, we propose another method based on the frequency weighted Hankel norm approximation method.

Theorem 19.9 *Let $W_1(s) \in \mathcal{RH}_\infty^-$ and $W_2(s) \in \mathcal{RH}_\infty^-$ with minimal state space realizations*

$$W_1(s) = \left[\begin{array}{c|c} A_{1w} & B_{1w} \\ \hline C_{1w} & D_{1w} \end{array}\right], \quad W_2(s) = \left[\begin{array}{c|c} A_{2w} & B_{2w} \\ \hline C_{2w} & D_{2w} \end{array}\right]$$

and let $G(s) \in \mathcal{RH}_\infty$. Suppose that $\hat{G}_1(s) = \left[\begin{array}{c|c} \hat{A}_1 & \hat{B}_1 \\ \hline \hat{C}_1 & \hat{D}_1 \end{array}\right] \in \mathcal{RH}_\infty$ is an r-th order optimal Hankel norm approximation of $[W_1 G W_2]_+$, i.e.,

$$\hat{G}_1 = \arg \inf_{\deg Q \le r} \left\| [W_1 G W_2]_+ - Q \right\|_H$$

and assume

$$\begin{bmatrix} A_{1w} - \lambda I & B_{1w} \\ C_{1w} & D_{1w} \end{bmatrix}, \quad \begin{bmatrix} A_{2w} - \lambda I & B_{2w} \\ C_{2w} & D_{2w} \end{bmatrix}$$

have respectively full row rank and full column rank for all $\lambda = \lambda_i(\hat{A}_1)$, $i = 1, \ldots, r$. *Then there exist matrices* X, Y, Q, *and* Z *such that*

$$\begin{aligned} A_{1w}X - X\hat{A}_1 + B_{1w}Y &= 0 && (19.12)\\ C_{1w}X + D_{1w}Y &= \hat{C}_1 && (19.13)\\ QA_{2w} - \hat{A}_1 Q + ZC_{2w} &= 0 && (19.14)\\ QB_{2w} + ZD_{2w} &= \hat{B}_1. && (19.15) \end{aligned}$$

Furthermore, $G_r := \left[\begin{array}{c|c} \hat{A}_1 & Z \\ \hline Y & 0 \end{array}\right]$ *is the frequency weighted optimal Hankel norm approximation, i.e.,*

$$\inf_{\deg \hat{G} \le r} \left\| W_1(G - \hat{G})W_2 \right\|_H = \|W_1(G - G_r)W_2\|_H = \sigma_{r+1}\left([W_1 G W_2]_+\right).$$

Proof. We shall assume $W_2 = I$ for simplicity. The general case can be proven similarly. Assume without loss of generality that $\hat{A}_1$ has a diagonal form

$$\hat{A}_1 = \operatorname{diag}[\lambda_1, \lambda_2, \ldots, \lambda_r].$$

(The proof can be easily modified if $\hat{A}_1$ has a general Jordan canonical form). Partition X, Y, and $\hat{C}_1$ as

$$X = [X_1, X_2, \ldots, X_r], \quad Y = [Y_1, Y_2, \ldots, Y_r], \quad \hat{C}_1 = [\hat{C}_{11}, \hat{C}_{12}, \ldots, \hat{C}_{1r}].$$

Then the equations (19.12) and (19.13) can be rewritten as

$$\begin{bmatrix} A_{1w} - \lambda_i I & B_{1w} \\ C_{1w} & D_{1w} \end{bmatrix} \begin{bmatrix} X_i \\ Y_i \end{bmatrix} = \begin{bmatrix} 0 \\ \hat{C}_{1i} \end{bmatrix}, \quad i = 1, 2, \ldots, r.$$

By the assumption, the matrix

$$\begin{bmatrix} A_{1w} - \lambda_i I & B_{1w} \\ C_{1w} & D_{1w} \end{bmatrix}$$

has full row rank for all i and thus the existence of X and Y is guaranteed. Let

$$\hat{W}_1 = \left[\begin{array}{c|c} A_{1w} & -X\hat{B}_1 \\ \hline C_{1w} & -\hat{D}_1 \end{array}\right] \in \mathcal{RH}_\infty^-.$$

Then using equations (19.12) and (19.13), we get

$$W_1 G_r = \left[\begin{array}{cc|c} A_{1w} & B_{1w}Y & 0 \\ 0 & \hat{A}_1 & \hat{B}_1 \\ \hline C_{1w} & D_{1w}Y & 0 \end{array}\right] = \left[\begin{array}{cc|c} A_{1w} & A_{1w}X - X\hat{A}_1 + B_{1w}Y & -X\hat{B}_1 \\ 0 & \hat{A}_1 & \hat{B}_1 \\ \hline C_{1w} & C_{1w}X + D_{1w}Y & 0 \end{array}\right]$$

$$= \left[\begin{array}{cc|c} A_{1w} & 0 & -X\hat{B}_1 \\ 0 & \hat{A}_1 & \hat{B}_1 \\ \hline C_{1w} & \hat{C}_1 & 0 \end{array}\right] = \hat{W}_1 + \hat{G}_1.$$

Using this expression, we have

$$\begin{aligned} \|W_1(G-G_r)\|_H &= \left\|[W_1G]_+ + [W_1G]_- - \hat{W}_1 - \hat{G}_1\right\|_H = \left\|[W_1G]_+ - \hat{G}_1\right\|_H \\ &= \sigma_{r+1}\left([W_1G]_+\right) \leq \inf_{\deg \hat{G} \leq r} \left\|W_1(G-\hat{G})\right\|_H \leq \|W_1(G-G_r)\|_H . \end{aligned}$$

□

Note that $Y = \hat{C}_1$ if $W_1 = I$, $Z = \hat{B}_1$ if $W_2 = I$, and the rank conditions in the above theorem are actually equivalent to the statements that the poles of $\hat{G}_1$ are not zeros of W_1 and W_2. These conditions will of course be satisfied automatically if $W_1(s)$ and $W_2(s)$ have all zeros in the right half plane. Numerical experience shows that if the weighted Hankel approximation is used to obtain a $\mathcal{L}_\infty$ norm approximation, then choosing $W_1(s)$ and $W_2(s)$ to have all poles and zeros in the right half plane may reduce the $\mathcal{L}_\infty$ norm approximation error significantly.

Corollary 19.10 *Let $W_1(s) \in \mathcal{RH}_\infty^-, W_2(s) \in \mathcal{RH}_\infty^-$ and $G(s) \in \mathcal{RH}_\infty$. Then*

$$\inf_{\deg \hat{G} \leq r} \left\|W_1(G-\hat{G})W_2\right\|_\infty \geq \inf_{\deg \hat{G} \leq r} \left\|W_1(G-\hat{G})W_2\right\|_H = \sigma_{r+1}\left([W_1GW_2]_+\right).$$

The lower bound in the above corollary is not necessarily achievable. To make the ∞-norm approximation error as small as possible, a suitable constant matrix D_r should be chosen so that $\left\|W_1(G-\hat{G}-D_r)W_2\right\|_\infty$ is made as small as possible. This D_r can usually be obtained using any standard convex optimization algorithm. To further reduce the approximation error, the following optimization is suggested.

Weighted $\mathcal{L}_\infty$ Model Reduction Procedures

Let W_1 and W_2 be any antistable transfer matrices with all zeros in the right half plane.

(i) Let

$$\hat{G}_1 = \left[\begin{array}{c|c} \hat{A}_1 & Z \\ \hline Y & 0 \end{array}\right]$$

be a weighted optimal Hankel norm approximation of G.

(ii) Let the reduced order model $\hat{G}$ be parameterized as

$$\hat{G}(\theta) = \left[\begin{array}{c|c} \hat{A}_1 & B_\theta \\ \hline Y & D_\theta \end{array}\right], \quad \text{or} \quad \hat{G}(\theta) = \left[\begin{array}{c|c} \hat{A}_1 & Z \\ \hline C_\theta & D_\theta \end{array}\right].$$

(iii) Find C_θ (or B_θ) and D_θ from the following convex optimization:

$$\min_{\theta \in \mathbb{R}^m} \left\| W_1(G - \hat{G}(\theta))W_2 \right\|_\infty .$$

It is noted that the weighted Hankel singular values can be used to predict the approximation error and hence to determine the order of the reduced model as in the unweighted Hankel approximation problem although we do not have an explicit $\mathcal{L}_\infty$ norm error bound in the weighted case.

If the given W_1 and W_2 do not have all poles and zeros in the right half plane, factorizations must be performed first to obtain the equivalent $\bar{W}_1(s)$ and $\bar{W}_2(s)$ so that $\bar{W}_1(s)$ and $\bar{W}_2(s)$ have all poles and zeros in the right half plane and

$$W_1^\sim(s)W_1(s) = \bar{W}_1^\sim(s)\bar{W}_1(s), \quad W_2(s)W_2^\sim(s) = \bar{W}_2(s)\bar{W}_2^\sim(s)$$

Then we have

$$\left\| W_1(G - \hat{G})W_2 \right\|_\infty = \left\| \bar{W}_1(G - \hat{G})\bar{W}_2 \right\|_\infty .$$

These factorizations can be easily done using Corollary 13.28 if W_1 and W_2 are stable and $W_1(\infty)$ and $W_2(\infty)$ have respectively full column rank and full row rank. For example, assume $W_2 = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$ with D full row rank and $W_2(j\omega)W_2^*(j\omega) > 0$ for all ω. Then there is a $M(s) = \left[\begin{array}{c|c} A & B_W \\ \hline C_W & (DD^*)^{1/2} \end{array}\right] \in \mathcal{RH}_\infty$ such that $M^{-1}(s) \in \mathcal{RH}_\infty$ and

$$W_2(s)W_2^\sim(s) = M^\sim(s)M(s)$$

where

$$B_W = PC^* + BD^*, \quad C_W = (DD^*)^{-1/2}(C - B_W^* X)$$

and

$$PA^* + AP + BB^* = 0$$

$$XA + A^*X + (C - B_W^*X)^*(DD^*)^{-1}(C - B_W^*X) = 0.$$

Finally, take $\bar{W}_2(s) = M^\sim(s)$. Then $\bar{W}_2(s)$ has all the poles and zeros in the right half plane and

$$\left\| W_1(G - \hat{G})W_2 \right\|_\infty = \left\| W_1(G - \hat{G})\bar{W}_2 \right\|_\infty .$$

In the case where W_1 and W_2 are not necessarily stable, the following procedures can be applied to accomplish this task.

Spectral Factorization Procedures

Let $W_1 \in \mathcal{L}_\infty$ and $W_2 \in \mathcal{L}_\infty$.

(i) Let $W_{1n} := W_1(-s)$ and $W_{2n} := W_{2n}(-s)$.

(ii) Let $W_{1n} = M_1^{-1} N_1$ and $W_{2n} = N_2 M_2^{-1}$ be respectively the left and right coprime factorizations such that M_1 and M_2 are inners. (This step can be done using Theorem 13.34.)

(iii) Perform the following spectral factorizations

$$N_1^{\sim} N_1 = V_1^{\sim} V_1, \quad N_2 N_2^{\sim} = V_2 V_2^{\sim}$$

so that V_1 and V_2 have all zeros in the left half plane. (Corollary 13.23 may be used here if $N_1(\infty)$ has full column rank and $N_2(\infty)$ has full row rank. Otherwise, the factorization in Section 6.1 may be used to factor out the undesirable poles and zeros in the weights.)

(iv) Let $\bar{W}_1(s) = V_1(-s)$ and $\bar{W}_2(s) = V_2(-s)$.

We shall summarize the state space formulas for the above factorizations as a lemma.

Lemma 19.11 *Let* $W(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{L}_\infty$ *be a controllable and observable realization.*

(a) Suppose $W^{\sim}(j\omega)W(j\omega) > 0$ *for all* ω *or* $\begin{bmatrix} A - j\omega & B \\ C & D \end{bmatrix}$ *has full column rank for all* ω*. Let*

$$Y = Ric \begin{bmatrix} -A^* & -C^*C \\ 0 & A \end{bmatrix} \geq 0$$

$$X = Ric \begin{bmatrix} -(A - BR^{-1}D^*C) & -BR^{-1}B^* \\ -C^*(I - DR^{-1}D^*)C & (A - BR^{-1}D^*C)^* \end{bmatrix} \geq 0$$

with $R := D^*D > 0$*. Then we have the following spectral factorization*

$$W^{\sim}W = \bar{W}^{\sim}\bar{W}$$

where $\bar{W}, \bar{W}^{-1} \in \mathcal{RH}_\infty^-$ *and*

$$\bar{W} = \left[\begin{array}{c|c} A + YC^*C & B + YC^*D \\ \hline R^{-1/2}(D^*C - B^*X)(I + YX)^{-1} & R^{1/2} \end{array}\right].$$

(b) Suppose $W(j\omega)W^{\sim}(j\omega) > 0$ *for all* ω *or* $\begin{bmatrix} A - j\omega & B \\ C & D \end{bmatrix}$ *has full row rank for all* ω*. Let*

$$X = Ric \begin{bmatrix} -A & -BB^* \\ 0 & A^* \end{bmatrix} \geq 0$$

$$Y = Ric \begin{bmatrix} -(A - BD^*\tilde{R}^{-1}C)^* & -C^*\tilde{R}^{-1}C \\ -B(I - D^*\tilde{R}^{-1}D)B^* & (A - BD^*\tilde{R}^{-1}C) \end{bmatrix} \geq 0$$

with $\tilde{R} := DD^* > 0$. Then we have the following spectral factorization

$$WW^\sim = \bar{W}\bar{W}^\sim$$

where $\bar{W}, \bar{W}^{-1} \in \mathcal{RH}_\infty^-$ and

$$\bar{W} = \left[\begin{array}{c|c} A + BB^*X & (I + YX)^{-1}(BD^* - YC^*)\tilde{R}^{-1/2} \\ \hline C + DB^*X & \tilde{R}^{1/2} \end{array}\right].$$

19.4 An Example

We consider a four-disk control system studied by Enns [1984]. We shall set up the dynamical system in the standard linear fractional transformation form

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= \begin{bmatrix} \sqrt{q_1}H \\ 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ I \end{bmatrix} u \\ y &= C_2 x + \begin{bmatrix} 0 & I \end{bmatrix} w \end{aligned}$$

where $q_1 = 1 \times 10^{-6}, q_2 = 1$ and

$$A = \begin{bmatrix} -0.161 & -6.004 & -0.58215 & -9.9835 & -0.40727 & -3.982 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad B_2 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}$$

$$B_1 = \begin{bmatrix} \sqrt{q_2}B_2 & 0 \end{bmatrix}, \quad H = \begin{bmatrix} 0 & 0 & 0 & 0 & 0.55 & 11 & 1.32 & 18 \end{bmatrix}$$

$$C_2 = \begin{bmatrix} 0 & 0 & 6.4432 \times 10^{-3} & 2.3196 \times 10^{-3} & 7.1252 \times 10^{-2} & 1.0002 & 0.10455 & 0.99551 \end{bmatrix}.$$

The optimal $\mathcal{H}_\infty$ norm for T_{zw} is $\gamma_{opt} = 1.1272$. We choose $\gamma = 1.2$ to compute an 8th order suboptimal controller K_o. The coprime factorizations of K_o are obtained using Lemma 19.6 as $K_o = \Theta_{12}\Theta_{22}^{-1} = \tilde{\Theta}_{22}^{-1}\tilde{\Theta}_{21}$. The controller is reduced using several methods and the results are listed in Table 19.2 where the following abbreviations are made:

UWA Unweighted additive reduction:

$$\left\| K_o - \hat{K} \right\|_\infty$$

UWRCF Unweighted right coprime factor reduction:

$$\left\| \begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right\|_\infty$$

UWLCF Unweighted left coprime factor reduction:

$$\left\| \begin{bmatrix} \tilde{\Theta}_{21} & \tilde{\Theta}_{22} \end{bmatrix} - \begin{bmatrix} \hat{\tilde{U}} & \hat{\tilde{V}} \end{bmatrix} \right\|_\infty$$

SWA Stability weighted additive reduction:

$$\left\| W_a(K_o - \hat{K}) \right\|_\infty$$

SWRCF Stability weighted right coprime factor reduction:

$$\left\| \begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix} \left(\begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right) \right\|_\infty$$

SWLCF Stability weighted left coprime factor reduction:

$$\left\| \left(\begin{bmatrix} \tilde{\Theta}_{12} & \tilde{\Theta}_{22} \end{bmatrix} - \begin{bmatrix} \hat{\tilde{U}} & \hat{\tilde{V}} \end{bmatrix} \right) \begin{bmatrix} -N_n \\ M_n \end{bmatrix} \right\|_\infty$$

PWA Performance weighted additive reduction:

$$\left\| W_2^{-1}(K_o - \hat{K})W_1^{-1} \right\|_\infty$$

PWRCF Performance weighted right coprime factor reduction:

$$\left\| \begin{bmatrix} \gamma^{-1}I & 0 \\ 0 & I \end{bmatrix} \Theta^{-1} \left(\begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right) \right\|_\infty$$

PWLCF Performance weighted left coprime factor reduction:

$$\left\| \left(\begin{bmatrix} \tilde{\Theta}_{12} & \tilde{\Theta}_{22} \end{bmatrix} - \begin{bmatrix} \hat{\tilde{U}} & \hat{\tilde{V}} \end{bmatrix} \right) \tilde{\Theta}^{-1} \begin{bmatrix} \gamma^{-1}I & 0 \\ 0 & I \end{bmatrix} \right\|_\infty$$

B Balance reduction with or without weighting

H/O Hankel/convex optimization reduction with or without weighting

Table 19.3 lists the performance weighted right coprime factor reduction errors and their lower bounds obtained using Corollary 19.10. The ϵ in Table 19.3 is defined as

$$\epsilon := \left\| \begin{bmatrix} \gamma^{-1}I & 0 \\ 0 & I \end{bmatrix} \Theta^{-1} \left(\begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right) \right\|_\infty .$$

By Corollary 19.10, $\epsilon \geq \sigma_{r+1}$ if the McMillan degree of $\begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix}$ is no greater than r.

Similarly, Table 19.4 lists the stability weighted right coprime factor reduction errors

Order of $\hat{K}$		7	6	5	4	3	2	1	0
PWA	B	1.196	1.196	1.199	1.197	U	4.99	U	U
	H/O	1.196	1.197	1.195	1.199	2.73	1.94	U	U
PWRCF	B	1.2	1.196	1.207	1.195	2.98	1.674	U	U
	H/O	1.196	1.198	1.196	1.199	2.036	1.981	U	U
PWLCF	B	1.197	1.196	U	1.197	U	U	U	U
	H/O	1.197	1.197	1.198	1.198	1.586	2.975	3.501	U
UWA	B	U	1.321	U	U	U	U	U	U
	H/O	23.15	U	U	U	U	U	U	U
UWRCF	B	1.198	1.196	1.199	1.196	U	U	U	U
	H/O	1.197	1.197	1.282	1.218	U	U	U	U
UWLCF	B	1.985	1.258	27.04	5.059	U	U	U	U
	H/O	5.273	U	U	U	U	U	U	U
SWA	B	1.327	1.199	2.27	1.47	23.5	U	U	U
	H/O	1.375	2.503	2.802	4.341	1.488	15.12	2.467	U
SWRCF	B	1.236	1.197	1.251	1.201	13.91	1.415	U	U
	H/O	2.401	1.893	1.612	1.388	2.622	3.527	U	U
SWLCF	B	1.417	1.217	48.04	3.031	U	U	U	U
	H/O	1.267	1.485	2.259	1.849	4.184	27.965	3.251	U

Table 19.2: $\mathcal{F}_\ell(G, \hat{K})$ with reduced order controller: U–closed-loop system is unstable

and their lower bounds obtained using Corollary 19.10. The ϵ_s and ϵ_u in Table 19.3 are defined as

$$\epsilon_s := \left\| \begin{bmatrix} -\tilde{N}_n & \tilde{M}_n \end{bmatrix} \left(\begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right) \right\|_\infty$$

$$\epsilon_u := \left\| \begin{bmatrix} \Theta_{12} \\ \Theta_{22} \end{bmatrix} - \begin{bmatrix} \hat{U} \\ \hat{V} \end{bmatrix} \right\|_\infty$$

where ϵ_u is obtained by taking the same $\hat{U}$ and $\hat{V}$ as in ϵ_s, not from the unweighted model reduction.

Table 19.2 shows that performance weighted controller reduction methods work very well. In particular, the PWRCF and PWLCF are easy to use and effective and there is in general no preference of using either the right coprime method or left coprime method. Although some unweighted reduction methods and some stability weighted reduction methods do give reasonable results in some cases, their performances are very hard to predict. What is the worst is that the small approximation error for the reduction criterion may have little relevance to the closed-loop $\mathcal{H}_\infty$ performance. For example, Table 19.4 shows that the 7-th order weighted approximation error ϵ_s using the Hankel/Optimization method is very small, however, the $\mathcal{H}_\infty$ performance is very far away from the desired level.

Although the unweighted right coprime factor reduction method gives very good results for this example, one should not be led to conclude that the unweighted right coprime factor method will work well in general. If this is true, then one can easily conclude that the unweighted left coprime factor method will do equally well by considering a dual problem. Of course, Table 19.2 shows that this is not true because the unweighted left coprime factor method does not give good results. The only reason for the good performance of the right coprime factor method for this example is the special data structure of this example, i.e., the relationship between the B_1 matrix and B_2 matrix:

$$B_1 = \begin{bmatrix} \sqrt{q_2}B_2 & 0 \end{bmatrix}.$$

The interested reader may want to explore this special structure further.

Order of $\hat{K}$	r	7	6	5	4	3	2	1	0
Lower bounds	σ_{r+1}	0.295	0.303	0.385	0.405	0.635	0.668	0.687	0.702
ϵ	B	1.009	0.626	4.645	0.750	71.8	6.59	127.2	2.029
	H/O	0.295	0.323	0.389	0.658	0.960	1	1	1

Table 19.3: PWRCF: ϵ and lower bounds

19.5 Notes and References

The stability oriented controller reduction criterion is first proposed by Enns [1984]. The weighted and unweighted coprime factor controller reduction methods are proposed by Liu and Anderson [1986, 1990], Liu, Anderson, and Ly [1990], Anderson and Liu [1989], and Anderson [1993]. For normalized $\mathcal{H}_\infty$ controller, Mustafa and Glover [1991] have proposed a controller reduction method with a prior performance bounds. Normalized coprime factors have been used in McFarlane and Glover [1990] for controller order reductions. Lenz, Khargonekar and Doyle [1987] have also proposed another $\mathcal{H}_\infty$ controller reduction method with guaranteed performance for a class of $\mathcal{H}_\infty$ problems. The main results presented in this chapter are based on the work of Goddard and Glover [1993,1994].

We should note that a satisfactory solution to the general frequency weighted $\mathcal{L}_\infty$ norm model reduction problem remains unavailable and this problem has a crucial implication toward controller reduction with preserving closed-loop $\mathcal{H}_\infty$ performance as its objective. The frequency weighted Hankel norm approximation is considered in Latham and Anderson [1986], Hung and Glover [1986], and Zhou [1993]. The $\mathcal{L}_\infty$ model reduction procedures discussed in this chapter are due to Zhou [1993].

Order of $\hat{K}$	r	7	6	5	4
Lower bounds of ϵ_s	σ_{r+1}	1.1×10^{-6}	1.2×10^{-6}	1.9×10^{-6}	1.9×10^{-6}
ϵ_s	B	0.8421	0.5048	2.5439	0.5473
	H/O	0.0001	0.2092	0.3182	0.3755
ϵ_u	B	254.28	9.7018	910.01	21.444
	H/O	185.9	30.85	305.3	15.38
Order of $\hat{K}$	r	3	2	1	0
Lower bounds of ϵ_s	σ_{r+1}	9×10^{-6}	6.23×10^{-5}	1.66×10^{-4}	2.145×10^{-4}
ϵ_s	B	11.791	1.3164	9.1461	1.5341
	H/O	0.5403	0.7642	1	1
ϵ_u	B	2600.2	365.45	3000.6	383.277
	H/O	397.9	288.1	384.3	384.3

Table 19.4: SWRCF: ϵ_s and the corresponding ϵ_u

20

Structure Fixed Controllers

In this chapter we focus on the problem of designing optimal controllers with controller structures restricted; for instance, the controller may be limited to be a state feedback or a constant output feedback or a fixed order dynamic controller. We shall be interested in deriving some explicit necessary conditions that an optimal fixed structure controller ought to satisfy. The fundamental idea is to formulate our optimal control problems as some constrained minimization problems. Then the first-order Lagrange multiplier necessary conditions for optimality are applied to derive our optimal controller formulae. Readers should keep in mind that our purpose here is to introduce the method, not to try to solve as many problems as possible. Hence, we will try to be concise but clear. In section 20.1, we will review some Lagrange multiplier optimization methods. Then these tools will be used in section 20.2 to solve a fixed order $\mathcal{H}_2$ optimal controller problem.

20.1 Lagrange Multiplier Method

In this section, we consider the constrained minimization problem. The results quoted below are standard and can be found in any references given at the end of the chapter.

Let $f(x) := f(x_1, x_2, \ldots, x_n) \in \mathbb{R}$ be a real valued function defined on a set $S \subset \mathbb{R}^n$. A point $x_0 \in \mathbb{R}^n$ in S is said to be a (global) *minimum point* of f on S if

$$f(x) \geq f(x_0)$$

for all points $x \in S$. A point $x_0 \in S$ is said to be a *local minimum point* of f on S if there is a neighborhood N of x_0 such that $f(x) \geq f(x_0)$ for all points $x \in N$.

We will be particularly interested in the case where the set S is described by a set of functions, $h_i(x) = 0, i = 1, 2, \ldots, m$ and $m < n$ or equivalently

$$H(x) := \begin{bmatrix} h_1(x) & h_2(x) & \ldots & h_m(x) \end{bmatrix} = 0.$$

Hence we will focus on the local necessary (and sufficient) conditions for the following problem:

$$\begin{cases} \text{minimize}\, f(x) \\ \text{subject to}\, H(x) = 0. \end{cases} \tag{20.1}$$

We shall assume that $f(x)$ and $h_i(x)$ are differentiable and denote

$$\nabla f(x) := \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$$

$$\nabla H(x) := \begin{bmatrix} \nabla h_1(x) & \nabla h_2(x) & \cdots & \nabla h_m(x) \end{bmatrix} = \begin{bmatrix} \frac{\partial h_1}{\partial x_1} & \frac{\partial h_2}{\partial x_1} & \cdots & \frac{\partial h_m}{\partial x_1} \\ \frac{\partial h_1}{\partial x_2} & \frac{\partial h_2}{\partial x_2} & \cdots & \frac{\partial h_m}{\partial x_2} \\ \vdots & \vdots & & \vdots \\ \frac{\partial h_1}{\partial x_n} & \frac{\partial h_2}{\partial x_n} & \cdots & \frac{\partial h_m}{\partial x_n} \end{bmatrix}.$$

Definition 20.1 A point $x_0 \in \mathbb{R}^n$ satisfying the constraints $H(x_0) = 0$ is said to be a *regular point* of the constraints if $\nabla H(x_0)$ has full column rank (m); equivalently, let $\phi(x) := H(x)z$ for $z \in \mathbb{R}^m$, and then $\nabla\phi(x_0) = \nabla H(x_0)z = 0$ has the unique solution $z = 0$.

Theorem 20.1 *Suppose that $x_0 \in \mathbb{R}^n$ is a local minimum of the $f(x)$ subject to the constraints $H(x) = 0$ and suppose further that x_0 is a regular point of the constraints. Then there exists a unique multiplier*

$$\Lambda = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_m \end{bmatrix} \in \mathbb{R}^m$$

such that, if we set $F(x) = f(x) + H(x)\Lambda$, then $\nabla F(x_0) = 0$, i.e.,

$$\nabla F(x_0) = \nabla f(x_0) + \nabla H(x_0)\Lambda = 0.$$

In the case where the regular point conditions are either not satisfied or hard to verify, we have the following alternative.

Theorem 20.2 *Suppose that $x_0 \in \mathbb{R}^n$ is a local minimum of $f(x)$ subject to the constraints $H(x) = 0$. Then there exists*

$$\begin{bmatrix} \lambda_0 \\ \Lambda \end{bmatrix} \in \mathbb{R}^{m+1}$$

such that $\lambda_0 \nabla f(x_0) + \nabla H(x_0)\Lambda = 0$.

Remark 20.1 Although some second order necessary and sufficient conditions for local minimality can be given, they are usually not very useful in our applications for the reason that they are very hard to verify except in some very special cases. Furthermore, even if some sufficient conditions can be verified, it is still not clear whether the minima is global. It is a common practice in many applications that the first-order necessary conditions are used to derive some necessary conditions for the existence of a minima. Then find a solution (or solutions) from these necessary conditions and check if the solution(s) satisfies our objectives regardless of the solution(s) being a global minima or not. ♡

In most of the control applications, constraints are given by a symmetric matrix function, and in this case, we have the following lemma.

Lemma 20.3 *Let $T(x) = T(x)^* \in \mathbb{R}^{l\times l}$ be a symmetric matrix function and let $x_0 \in \mathbb{R}^n$ be such that $T(x_0) = 0$. Then x_0 is a regular point of the constraints $T(x) = T(x)^* = 0$ if, for $P = P^* \in \mathbb{R}^{l\times l}$, $\nabla \operatorname{Trace}(T(x_0)P) = 0$ has the unique solution $P = 0$.*

Proof. Since $T(x) = [t_{ij}(x)]$ is a symmetric matrix, $t_{ij} = t_{ji}$ and the effective constraints for $T(x) = 0$ are given by the $l(l+1)/2$ equations, $t_{ij}(x) = 0$ for $i = 1, 2, \ldots, l$ and $i \le j \le l$. By definition, x_0 is a regular point for the effective constraints ($t_{ij}(x) = 0$ for $i = 1, 2, \ldots, l$ and $i \le j \le l$) if the following equation has the unique solution $p_{ij} = 0$ for $i = 1, 2, \ldots, l$ and $i \le j \le l$:

$$\psi(x_0) := \sum_{i=1}^{l} \nabla t_{ii} p_{ii} + \sum_{i=1}^{l} \sum_{j=i+1}^{l} 2\nabla t_{ij} p_{ji} = 0.$$

Now the result follows by defining $p_{ij} := p_{ji}$ and by noting that $\psi(x_0)$ can be written as

$$\psi(x_0) = \nabla \left(\sum_{i=1}^{l} \sum_{j=1}^{l} t_{ij} p_{ji} \right) = \nabla \operatorname{Trace}(T(x_0)P)$$

with $P = P^*$. □

Corollary 20.4 *Suppose that $x_0 \in \mathbb{R}^n$ is a local minimum of $f(x)$ subject to the constraints $T(x) = 0$ where $T(x) = T(x)^* \in \mathbb{R}^{l\times l}$ and suppose further that x_0 is a regular* *point of the constraints. Then there exists a unique multiplier $P = P^* \in \mathbb{R}^{l \times l}$ such that if we set $F(x) = f(x) + \mathrm{Trace}(T(x)P)$, then $\nabla F(x_0) = 0$, i.e.,*

$$\nabla F(x_0) = \nabla f(x_0) + \nabla \, \mathrm{Trace}(T(x_0)P) = 0.$$

In general, in the case where a local minimal point x_0 is not necessarily a regular point, we have the following corollary.

Corollary 20.5 *Suppose that $x_0 \in \mathbb{R}^n$ is a local minimum of $f(x)$ subject to the constraints $T(x) = 0$ where $T(x) = T(x)^* \in \mathbb{R}^{l \times l}$. Then there exist $0 \neq (\lambda_0, P) \in \mathbb{R} \times \mathbb{R}^{l \times l}$ with $P = P^*$ such that*

$$\lambda_0 \nabla f(x_0) + \nabla \, \mathrm{Trace}(T(x_0)P) = 0.$$

Remark 20.2 We shall also note that the variable $x \in \mathbb{R}^n$ may be more conveniently given in terms of a matrix $X \in \mathbb{R}^{k \times q}$, i.e., we have

$$x = \mathrm{Vec} X := \begin{bmatrix} x_{11} \\ \vdots \\ x_{k1} \\ x_{12} \\ \vdots \\ x_{k2} \\ \vdots \\ x_{kq} \end{bmatrix}.$$

Then

$$\nabla F(x) := \begin{bmatrix} \frac{\partial F(x)}{\partial x_{11}} \\ \vdots \\ \frac{\partial F(x)}{\partial x_{kq}} \end{bmatrix} = 0$$

is equivalent to

$$\frac{\partial F(x)}{\partial X} := \begin{bmatrix} \frac{\partial F(x)}{\partial x_{11}} & \frac{\partial F(x)}{\partial x_{12}} & \cdots & \frac{\partial F(x)}{\partial x_{1q}} \\ \frac{\partial F(x)}{\partial x_{21}} & \frac{\partial F(x)}{\partial x_{22}} & \cdots & \frac{\partial F(x)}{\partial x_{2q}} \\ \vdots & \vdots & & \vdots \\ \frac{\partial F(x)}{\partial x_{k1}} & \frac{\partial F(x)}{\partial x_{k2}} & \cdots & \frac{\partial F(x)}{\partial x_{kq}} \end{bmatrix} = 0.$$

This later expression will be used throughout in the sequel. ♡

As an example, let us consider the following $\mathcal{H}_2$ norm minimization with constant state feedback: the dynamic system is given by

$$\begin{aligned} \dot{x} &= Ax + B_1 w + B_2 u \\ z &= C_1 x + D_{12} u, \end{aligned}$$

and the feedback $u = Fx$ is chosen so that $A + B_2F$ is stable and

$$J_0 = \|T_{zw}\|_2^2$$

is minimized. For simplicity, we shall assume that $D_{12}^* D_{12} = I$ and $D_{12}^* C_1 = 0$. It is routine to verify that $J_0 = \text{Trace}(B_1 B_1^* X)$ where $X = X^* \geq 0$ satisfies

$$T(X,F) := X(A + B_2F) + (A + B_2F)^* X + (C_1 + D_{12}F)^*(C_1 + D_{12}F) = 0.$$

Hence the optimal control problem becomes a constrained minimization problem, and we can use the Lagrange multipliers method outlined above. (Note that in this case, the variable x takes the form $x = \begin{bmatrix} \text{Vec}X \\ \text{Vec}F \end{bmatrix}$). Let

$$J(X,F) := J_0 + \text{Trace}(T(X,F)P)$$

with $P = P^*$. We first verify the regularity conditions: the equation

$$\nabla\, \text{Trace}(T(X,F)P) = 0$$

or, equivalently,

$$\begin{bmatrix} \frac{\partial\, \text{Trace}(T(X,F)P)}{\partial X} \\ \frac{\partial\, \text{Trace}(T(X,F)P)}{\partial F} \end{bmatrix} = \begin{bmatrix} P(A + B_2F)^* + (A + B_2F)P \\ 2(B_2^* X + F)P \end{bmatrix} = 0$$

has a unique solution $P = 0$ since $A + B_2F$ is assumed to be stable at the minimum point. Hence regularity conditions are satisfied. Now the necessary condition for local optimum can be applied:

$$\frac{\partial J(X,F)}{\partial X} = B_1 B_1^* + P(A + B_2F)^* + (A + B_2F)P = 0 \tag{20.2}$$

$$\frac{\partial J(X,F)}{\partial F} = 2(B_2^* X + F)P = 0 \tag{20.3}$$

$$\frac{\partial J(X,F)}{\partial P} = X(A + B_2F) + (A + B_2F)^* X + (C_1 + D_{12}F)^*(C_1 + D_{12}F) = 0. \tag{20.4}$$

It should be pointed out that, in general, we cannot conclude from equation (20.3) that $B_2^* X + F = 0$. Care must be exercised to arrive at such a conclusion. For example, if we assume that B_1 is square and nonsingular, then we know that if F is such that $A + B_2F$ is stable, then $P > 0$. Hence we have

$$F = -B_2^* X,$$

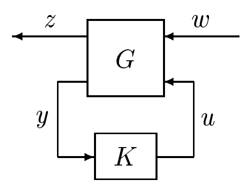
and we substitute this relation into equation (20.4) and get the familiar Riccati equation:

$$XA + A^* X - XB_2 B_2^* X + C_1^* C_1 = 0.$$

This Riccati equation has a stabilizing solution if (A, B_2) is stabilizable and if (C_1, A) has no unobservable modes on the imaginary axis. Indeed, in this case, the controller thus obtained is a global optimal control law.

20.2 Fixed Order Controllers

In this section, we shall use the Lagrange multiplier method to derive some necessary conditions for a fixed-order controller that minimizes an $\mathcal{H}_2$ performance. We shall again consider a standard system setup



where the system G is n-th order with a realization given by

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right].$$

For simplicity, we shall assume that

(i) (A, B_1) is stabilizable and (C_1, A) is detectable;

(ii) (A, B_2) is stabilizable and (C_2, A) is detectable;

(iii) $D_{12}^* \begin{bmatrix} C_1 & D_{12} \end{bmatrix} = \begin{bmatrix} 0 & I \end{bmatrix}$;

(iv) $\begin{bmatrix} B_1 \\ D_{21} \end{bmatrix} D_{21}^* = \begin{bmatrix} 0 \\ I \end{bmatrix}$.

We shall be interested in the following fixed order $\mathcal{H}_2$ optimal controller problem: *given an integer $n_c \le n$, find an n_c-th order controller*

$$K = \left[\begin{array}{c|c} A_c & B_c \\ \hline C_c & 0 \end{array}\right]$$

that internally stabilizes the system G and minimizes the $\mathcal{H}_2$ norm of the transfer matrix T_{zw}. For technical reasons, we will further assume that the realization of the controller is minimal, i.e., (A_c, B_c) *is controllable and* (C_c, A_c) *is observable.*

Suppose a such controller exists, and then the closed loop transfer matrix T_{zw} can be written as

$$T_{zw} = \left[\begin{array}{cc|c} A & B_2 C_c & B_1 \\ B_c C_2 & A_c & B_c D_{21} \\ \hline C_1 & D_{12} C_c & 0 \end{array}\right] =: \left[\begin{array}{c|c} \tilde{A} & \tilde{B} \\ \hline \tilde{C} & 0 \end{array}\right]$$

with $\tilde{A}$ stable. Moreover,

$$\|T_{zw}\|_2^2 = \text{Trace}(\tilde{B}\tilde{B}^*\tilde{X}) \tag{20.5}$$

where $\tilde{X}$ is the observability Gramian of T_{zw}:

$$\tilde{X}\tilde{A} + \tilde{A}^*\tilde{X} + \tilde{C}^*\tilde{C} = 0. \tag{20.6}$$

Theorem 20.6 *Suppose (A_c, B_c, C_c) is a controllable and observable triple and $K = \left[\begin{array}{c|c} A_c & B_c \\ \hline C_c & 0 \end{array}\right]$ internally stabilizes the system G and minimizes the norm T_{zw}. Then there exist $n \times n$ nonnegative definite matrices X, Y, $\hat{X}$, and $\hat{Y}$ such that A_c, B_c, and C_c are given by*

$$\begin{aligned} A_c &= \Gamma(A - B_2B_2^*X - YC_2^*C_2)\Pi^* && (20.7)\\ B_c &= \Gamma YC_2^* && (20.8)\\ C_c &= -B_2^*X\Pi^* && (20.9) \end{aligned}$$

for some factorization

$$\hat{Y}\hat{X} = \Pi^* M\Gamma, \quad \Gamma\Pi^* = I_{n_c}$$

with M positive-semisimple[1] *and such that with $\tau := \Pi^*\Gamma$ and $\tau_\perp := I_n - \tau$ the following conditions are satisfied:*

$$\begin{aligned} 0 &= A^*X + XA - XB_2B_2^*X + \tau_\perp^* XB_2B_2^*X\tau_\perp + C_1^*C_1 && (20.10)\\ 0 &= AY + YA^* - YC_2^*C_2Y + \tau_\perp YC_2^*C_2Y\tau_\perp^* + B_1B_1^* && (20.11)\\ 0 &= (A - YC_2^*C_2)^*\hat{X} + \hat{X}(A - YC_2^*C_2) + XB_2B_2^*X - \tau_\perp^* XB_2B_2^*X\tau_\perp && (20.12)\\ 0 &= (A - B_2B_2^*X)\hat{Y} + \hat{Y}(A - B_2B_2^*X)^* + YC_2^*C_2Y - \tau_\perp YC_2^*C_2Y\tau_\perp^* && (20.13) \end{aligned}$$

$$\text{rank } \hat{X} = \text{rank } \hat{Y} = \text{rank } \hat{Y}\hat{X} = n_c.$$

Proof. The problem can be viewed as a constrained minimization problem with the objective function given by equation (20.5) and with constraints given by equation (20.6). Let $\tilde{Y} = \tilde{Y}^* \in \mathbb{R}^{(n+n_c)\times(n+n_c)}$ and denote

$$J_1 = \text{Trace}\left\{(\tilde{X}\tilde{A} + \tilde{A}^*\tilde{X} + \tilde{C}^*\tilde{C})\tilde{Y}\right\}.$$

Then

$$\frac{\partial J_1}{\partial \tilde{X}} = \tilde{Y}\tilde{A}^* + \tilde{A}\tilde{Y} = 0$$

has the unique solution $\tilde{Y} = 0$ since $\tilde{A}$ is assumed to be stable. Hence the regularity conditions are satisfied and the Lagrange multiplier method can be applied. Form the Lagrange function J as

$$J := \text{Trace}(\tilde{B}\tilde{B}^*\tilde{X}) + J_1$$

[1] A matrix M is called *positive semisimple* if it is similar to a positive definite matrix.

and partition $\tilde{X}$ and $\tilde{Y}$ as

$$\tilde{X} = \begin{bmatrix} X_{11} & X_{12} \\ X_{12}^* & X_{22} \end{bmatrix}, \quad \tilde{Y} = \begin{bmatrix} Y_{11} & Y_{12} \\ Y_{12}^* & Y_{22} \end{bmatrix}.$$

The necessary conditions for (A_c, B_c, C_c) to be a local minima are

$$\begin{aligned} \frac{\partial J}{\partial A_c} &= 2(X_{12}^* Y_{12} + X_{22} Y_{22}) = 0 && (20.14) \\ \frac{\partial J}{\partial B_c} &= 2(X_{22} B_c + X_{12}^* Y_{11} C_2^* + X_{22} Y_{12}^* C_2^*) = 0 && (20.15) \\ \frac{\partial J}{\partial C_c} &= 2(B_2^* X_{11} Y_{12} + B_2^* X_{12} Y_{22} + C_c Y_{22}) = 0 && (20.16) \\ \frac{\partial J}{\partial \tilde{Y}} &= \tilde{X}\tilde{A} + \tilde{A}^*\tilde{X} + \tilde{C}^*\tilde{C} = 0 && (20.17) \\ \frac{\partial J}{\partial \tilde{X}} &= \tilde{Y}\tilde{A}^* + \tilde{A}\tilde{Y} + \tilde{B}^*\tilde{B} = 0. && (20.18) \end{aligned}$$

It is clear that $\tilde{X} \geq 0$ and $\tilde{Y} \geq 0$ since $\tilde{A}$ is stable. Equations (20.17) and (20.18) can be written as

$$\begin{aligned} 0 &= X_{11}A + A^*X_{11} + X_{12}B_cC_2 + C_2^*B_c^*X_{12}^* + C_1^*C_1 && (20.19) \\ 0 &= X_{12}A_c + A^*X_{12} + X_{11}B_2C_c + C_2^*B_c^*X_{22} && (20.20) \\ 0 &= X_{22}A_c + A_c^*X_{22} + X_{12}^*B_2C_c + C_c^*B_2^*X_{12} + C_c^*C_c && (20.21) \\ 0 &= AY_{11} + Y_{11}A^* + B_2C_cY_{12}^* + Y_{12}C_c^*B_2^* + B_1B_1^* && (20.22) \\ 0 &= AY_{12} + Y_{12}A_c^* + B_2C_cY_{22} + Y_{11}C_2^*B_c^* && (20.23) \\ 0 &= A_cY_{22} + Y_{22}A_c^* + B_cC_2Y_{12} + Y_{12}^*C_2^*B_c^* + B_cB_c^*. && (20.24) \end{aligned}$$

For clarity, we shall present the proof in three steps:

1. $X_{22} > 0$ and $Y_{22} > 0$: We show first that $X_{22} > 0$. Since $\tilde{X} \geq 0$, we have $X_{22} \geq 0$ and $X_{22}X_{22}^+X_{12}^* = X_{12}^*$ by Lemma 2.16. Hence equation (20.21) can be written as

   $$0 = X_{22}(A_c + X_{22}^+X_{12}^*B_2C_c) + (A_c + X_{22}^+X_{12}^*B_2C_c)^*X_{22} + C_c^*C_c.$$

   Since (C_c, A_c) is observable by assumption, $(C_c, A_c + X_{22}^+X_{12}^*B_2C_c)$ is also observable. Now it follows from the Lyapunov theorem (Lemma 3.18) that $X_{22} > 0$. $Y_{22} > 0$ follows by a similar argument.

2. formula for A_c, B_c, C_c, Γ, and Π: given $X_{22} > 0$ and $Y_{22} > 0$, we define

   $$\begin{aligned} \Gamma &:= -X_{22}^{-1}X_{12}^* && (20.25) \\ \Pi &:= Y_{22}^{-1}Y_{12}^* && (20.26) \\ \tau &:= \Pi^*\Gamma && (20.27) \end{aligned}$$

$$\hat{X} \quad := \quad X_{12}X_{22}^{-1}X_{12}^* \geq 0 \tag{20.28}$$
$$\hat{Y} \quad := \quad Y_{12}Y_{22}^{-1}Y_{12}^* \geq 0 \tag{20.29}$$
$$X \quad := \quad X_{11} - X_{12}X_{22}^{-1}X_{12}^* \tag{20.30}$$
$$Y \quad := \quad Y_{11} - Y_{12}Y_{22}^{-1}Y_{12}^*. \tag{20.31}$$

Then it follows from equation (20.14) that

$$\Gamma\Pi^* = I$$

and $\tau^2 = \Pi^*\Gamma\Pi^*\Gamma = \Pi^*\Gamma = \tau$. It also follows from $\tilde{X} \geq 0$ and $\tilde{Y} \geq 0$ that $X \geq 0$ and $Y \geq 0$. Moreover, from equation (20.14), we have

$$n_c = \text{rank } X_{22} \leq \text{rank } X_{12} \leq n_c.$$

This implies that rank $X_{12} = n_c =$ rank Y_{12} and

$$\text{rank } \hat{X} = \text{rank } X_{12} = n_c = \text{rank } \hat{Y} = \text{rank } \hat{Y}\hat{X}.$$

The product of $\hat{Y}\hat{X}$ can be factored as

$$\hat{Y}\hat{X} = \Pi^*(-Y_{12}^*X_{12})\Gamma = \Pi^*Y_{22}X_{22}\Gamma,$$

and

$$M := Y_{22}X_{22} = Y_{22}^{1/2}(Y_{22}^{1/2}X_{22}Y_{22}^{1/2})Y_{22}^{-1/2}$$

is a positive semisimple matrix.

Using these formulae in equations (20.15) and (20.16), we get

$$\begin{aligned} B_c &= -(X_{22}^{-1}X_{12}^*Y_{11} + Y_{12}^*)C_2^* = (\Gamma Y_{11} - (\Gamma\Pi^*)Y_{12}^*)C_2^* \\ &= \Gamma(Y_{11} - \Pi^*Y_{12}^*)C_2^* \\ &= \Gamma Y C_2^* \end{aligned}$$

and

$$\begin{aligned} C_c &= -B_2^*(X_{11}Y_{12}Y_{22}^{-1} + X_{12}) = -B_2^*(X_{11} + X_{12}\Gamma)\Pi^* \\ &= -B_2^*X\Pi^*. \end{aligned}$$

The formula for A_c follows from (20.24)-$\Gamma\times$(20.23) and some algebra.

3. Equations for X, Y, $\hat{X}$, and $\hat{Y}$: Equations (20.12) and (20.13) follow by substituting A_c, B_c, and C_c into (20.20)$\times\Gamma$ and (20.23)$\times\Pi$ and by using the fact that $\hat{X} = \tau\hat{X}$ and $\hat{Y} = \tau\hat{Y}$. Finally, equations (20.10) and (20.11) follow from equations (20.19) and (20.22) with some tedious algebra.

□

Remark 20.3 It is interesting to note that if the full order controller $n_c = n$ is considered, then we have $\tau = I$ and $\tau_\perp = 0$. In that case, equations (20.10) and (20.11) become standard $\mathcal{H}_2$ Riccati equations:

$$\begin{aligned} 0 &= A^*X + XA - XB_2B_2^*X + C_1^*C_1 \\ 0 &= AY + YA^* - YC_2^*C_2Y + B_1B_1^* \end{aligned}$$

Moreover, there exist unique stabilizing solutions $X \geq 0$ and $Y \geq 0$ to these two Riccati equations such that $A - B_2B_2^*X$ and $A - YC_2^*C_2$ are stable. Using these facts, we get that equations (20.12) and (20.13) have unique solutions:

$$\begin{aligned} \hat{X} &= \int_0^\infty e^{(A-YC_2^*C_2)^*t} XB_2B_2^*X e^{(A-YC_2^*C_2)t} dt \geq 0 \\ \hat{Y} &= \int_0^\infty e^{(A-B_2B_2^*X)t} YC_2^*C_2Y e^{(A-B_2B_2^*X)^*t} dt \geq 0. \end{aligned}$$

It is a fact that $\hat{X}$ is nonsingular iff $(B_2^*X, A - YC_2^*C_2)$ is observable and that $\hat{Y}$ is nonsingular iff $(A - B_2B_2^*X, YC_2^*)$ is controllable, or equivalently iff

$$K_o := \left[\begin{array}{c|c} A - B_2B_2^*X - YC_2^*C_2 & YC_2^* \\ \hline -B_2^*X & 0 \end{array}\right]$$

is controllable and observable. (Note that K_o is known to be the optimal $\mathcal{H}_2$ controller from Chapter 14). Furthermore, if $\hat{X}$ and $\hat{Y}$ are nonsingular, we can indeed find Γ and Π such that $\Gamma\Pi^* = I_n$. In fact, in this case, Γ and Π are both square and $\Gamma = (\Pi^*)^{-1}$. Hence, we have

$$\begin{aligned} K &= \left[\begin{array}{c|c} A_c & B_c \\ \hline C_c & 0 \end{array}\right] = \left[\begin{array}{c|c} \Gamma(A - B_2B_2^*X - YC_2^*C_2)\Gamma^{-1} & \Gamma YC_2^* \\ \hline -B_2^*X\Gamma^{-1} & 0 \end{array}\right] \\ &= \left[\begin{array}{c|c} A - B_2B_2^*X - YC_2^*C_2 & YC_2^* \\ \hline -B_2^*X & 0 \end{array}\right] = K_o \end{aligned}$$

i.e., if $\hat{X}$ and $\hat{Y}$ are nonsingular or, equivalently, if optimal controller K_o is controllable and observable as we assumed, then Theorem 20.6 generates the optimal controller. However, in general, K_o is not necessarily minimal; hence Theorem 20.6 will not be applicable.

It is possible to derive some similar results to Theorem 20.6 without assuming the minimality of the optimal controller by using pseudo-inverse in the derivations, but that, in general, is much more complicated. An alternative solution to this dilemma would be simply by direct testing: if a given n_c does not generate a controllable and observable controller, then lower n_c and try again. ♡

Remark 20.4 We should also note that although we have the necessary conditions for a reduced order optimal controller, it is generally hard to solve these coupled equations although some ad hoc homotopy algorithm might be used to find a local minima. ♡

Remark 20.5 This method can also be used to derive the $\mathcal{H}_\infty$ results presented in the previous chapters. The interested reader should consult the references for details. It should be pointed out that this method suffers a severe deficiency: global results are hard to find. This is due to (a) only first order necessary conditions can be relatively easily derived; (b) the controller order must be fixed; hence even if a fixed-order optimal controller can be found, it may not be optimal over all stabilizing controllers. ♡

20.3 Notes and References

Optimization using the Lagrange multiplier can be found in any standard optimization textbook. In particular, the book by Hestenes [1975] contains the finite dimensional case, and the one by Luenberger [1969] contains both finite and infinite dimensional cases. The Lagrange multiplier method has been used extensively by Hyland and Bernstein [1984], Bernstein and Haddard [1989], and Skelton [1988] in control applications. Theorem 20.6 was originally shown in Hyland and Bernstein [1984].

21

Discrete Time Control

In this chapter we discuss discrete time Riccati equations and some of their applications in discrete time control. A simpler form of a Riccati equation is the so-called Lyapunov equation. Hence we will start from the solutions of a discrete Lyapunov equation which are given in section 21.1. Section 21.2 presents the basic property of a Riccati equation solution as well as the necessary and sufficient conditions for the existence of a solution to the LQR problem related Riccati equation. Various different characterizations of a bounded real transfer matrix are presented in section 21.3. The key is the relationship between the existence of a solution to a Riccati equation and the norm bound of a stable transfer matrix. Section 21.4 collects some useful matrix function factorizations and characterizations. In particular, state space criteria are stated (mostly without proof) for a transfer matrix to be inner, for the existence of coprime factorizations, inner-outer factorizations, normalized coprime factorizations, and spectral factorizations. The discrete time $\mathcal{H}_2$ optimal control will be considered briefly in Section 21.5. Finally, the discrete time balanced model reduction is considered in section 21.6 and 21.7.

21.1 Discrete Lyapunov Equations

Let A, B, and Q be real matrices with appropriate dimensions, and consider the following linear equation:

$$AXB - X + Q = 0. \tag{21.1}$$

Lemma 21.1 *The equation (21.1) has a unique solution if and only if $\lambda_i(A)\lambda_j(B) \neq 1$ for all i, j.*

Proof. Analogous to the continuous case. □

Remark 21.1 If $\lambda_i(A)\lambda_j(B) = 1$ for some i, j, then the equation (21.1) has either no solution or more than one solution depending on the specific data given. If $B = A^*$ and $Q = Q^*$, then the equation is called the discrete Lyapunov equation. ♡

The following results are analogous to the corresponding continuous time cases, so they will be stated without proof.

Lemma 21.2 *Let Q be a symmetric matrix and consider the following Lyapunov equation:*

$$AXA^* - X + Q = 0$$

1. *Suppose that A is stable, and then the following statements hold:*
   (a) $X = \sum_{i=0}^{\infty} A^i Q (A^*)^i$ *and* $X \geq 0$ *if* $Q \geq 0$.
   (b) *if* $Q \geq 0$, *then* (Q, A) *is observable iff* $X > 0$.

2. *Suppose that X is the solution of the Lyapunov equation; then*
   (a) $|\lambda_i(A)| \leq 1$ *if* $X > 0$ *and* $Q \geq 0$.
   (b) A *is stable if* $X \geq 0$, $Q \geq 0$ *and* (Q, A) *is detectable.*

21.2 Discrete Riccati Equations

This section collects some basic results on the discrete Riccati equations. So the presentation of this section and the sections to follow will be very much like the corresponding sections in Chapter 13. Just as the continuous time Riccati equations play the essential roles in continuous $\mathcal{H}_2$ and $\mathcal{H}_\infty$ theories, the discrete time Riccati equations play the essential roles in discrete time $\mathcal{H}_2$ and $\mathcal{H}_\infty$ theories.

Let a matrix $S \in \mathbb{R}^{2n \times 2n}$ be partitioned into four $n \times n$ blocks as

$$S := \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix},$$

and let $J = \begin{bmatrix} 0 & -I \\ I & 0 \end{bmatrix} \in \mathbb{R}^{2n \times 2n}$; then S is called *simplectic* if $J^{-1} S^* J = S^{-1}$. A simplectic matrix has no eigenvalues at the origin, and, furthermore, it is easy to see that if λ is an eigenvalue of a simplectic matrix S, then $\bar{\lambda}, 1/\lambda$, and $1/\bar{\lambda}$ are also eigenvalues of S.

Let A, Q, and G be real $n \times n$ matrices with Q and G symmetric and A nonsingular. Define a $2n \times 2n$ matrix:

$$S := \begin{bmatrix} A + G(A^*)^{-1}Q & -G(A^*)^{-1} \\ -(A^*)^{-1}Q & (A^*)^{-1} \end{bmatrix}.$$

Then S is a simplectic matrix. Assume that S has no eigenvalues on the unit circle. Then it must have n eigenvalues in $|z| < 1$ and n in $|z| > 1$. Consider the two n-dimensional spectral subspaces $\mathcal{X}_-(S)$ and $\mathcal{X}_+(S)$: the former is the invariant subspace corresponding to eigenvalues in $|z| < 1$, and the latter corresponds to eigenvalues in $|z| > 1$. After finding a basis for $\mathcal{X}_-(S)$, stacking the basis vectors up to form a matrix, and partitioning the matrix, we get

$$\mathcal{X}_-(S) = \operatorname{Im} \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$$

where $T_1, T_2 \in \mathbb{R}^{n\times n}$. If T_1 is nonsingular or, equivalently, if the two subspaces

$$\mathcal{X}_-(S), \quad \operatorname{Im} \begin{bmatrix} 0 \\ I \end{bmatrix}$$

are complementary, we can set $X := T_2 T_1^{-1}$. Then X is uniquely determined by S, i.e., $S \mapsto X$ is a function which will be denoted Ric; thus, $X = Ric(S)$. As in the continuous time case, we make the following definition.

Definition 21.1 The domain of Ric, denoted by $dom(Ric)$, consists of all $(2n \times 2n)$ simplectic matrices S such that S has no eigenvalues on the unit circle and the two subspaces $\mathcal{X}_-(S)$ and $\operatorname{Im} \begin{bmatrix} 0 \\ I \end{bmatrix}$ are complementary.

Theorem 21.3 *Suppose $S \in dom(Ric)$ and $X = Ric(S)$. Then*

(a) X is unique and symmetric;

(b) $I + XG$ is invertible and X satisfies the algebraic Riccati equation

$$A^*XA - X - A^*XG(I+XG)^{-1}XA + Q = 0; \tag{21.2}$$

(c) $A - G(I+XG)^{-1}XA = (I+GX)^{-1}A$ is stable.

Note that the discrete Riccati equation in (21.2) can also be written as

$$A^*(I+XG)^{-1}XA - X + Q = 0.$$

Remark 21.2 In the case that A is singular, all results presented in this chapter will still be true if the eigenvalue problem of S is replaced by the following generalized eigenvalue problem:

$$\lambda \begin{bmatrix} I & G \\ 0 & A^* \end{bmatrix} - \begin{bmatrix} A & 0 \\ -Q & I \end{bmatrix}$$

and $\mathcal{X}_-(S)$ is taken to be the subspace spanned by the *generalized principal vectors* corresponding to those generalized eigenvalues in $|z| < 1$. Here the generalized principal vectors corresponding to a generalized eigenvalue λ are referred to the set of vectors $\{x_1, \ldots, x_k\}$ satisfying

$$\begin{aligned} \begin{bmatrix} A & 0 \\ -Q & I \end{bmatrix} x_1 &= \lambda \begin{bmatrix} I & G \\ 0 & A^* \end{bmatrix} x_1 \\ \left(\begin{bmatrix} A & 0 \\ -Q & I \end{bmatrix} - \lambda \begin{bmatrix} I & G \\ 0 & A^* \end{bmatrix} \right) x_i &= \begin{bmatrix} I & G \\ 0 & A^* \end{bmatrix} x_{i-1}, \quad i = 2, \ldots, k. \end{aligned}$$

See Dooren [1981] and Arnold and Laub [1984] for details. ♡

Proof. (a): Since $S \in \mathrm{dom(Ric)}$, $\exists T = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$ such that

$$\mathcal{X}_-(S) = \mathrm{Im} \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$$

and T_1 is invertible. Let $X := T_2 T_1^{-1}$, then

$$\mathcal{X}_-(S) = \mathrm{Im} \begin{bmatrix} T_1 \\ T_2 \end{bmatrix} = \mathrm{Im} \begin{bmatrix} I \\ X \end{bmatrix} T_1 = \mathrm{Im} \begin{bmatrix} I \\ X \end{bmatrix}.$$

Obviously, X is unique since

$$\mathrm{Im} \begin{bmatrix} I \\ X_1 \end{bmatrix} = \mathrm{Im} \begin{bmatrix} I \\ X_2 \end{bmatrix}$$

iff $X_1 = X_2$. Now let us show that X is symmetric. Since

$$XT_1 = T_2, \tag{21.3}$$

pre-multiply by T_1^* to get

$$T_1^* X T_1 = T_1^* T_2.$$

We only need to show that $T_1^* T_2$ is symmetric.

Since $\mathcal{X}_-(S)$ is a stable invariant subspace, there is a stable $n \times n$ matrix S_- such that

$$ST = TS_-. \tag{21.4}$$

Pre-multiply by $S_-^* T^* J$ and note that $S^* J S = J$; we get

$$(S_-^* T^* J) T S_- = (S_-^* T^* J) S T = (S_-^* T^*) J S T = T^* S^* J S T = T^* J T$$

$$S_-^* (T^* J T) S_- - T^* J T = 0.$$

This is a Lyapunov equation and S_- is stable. Hence there is a unique solution

$$T^* J T = 0$$

i.e.,

$$T_2^* T_1 = T_1^* T_2.$$

Thus we have $X = X^*$ since T_1 is nonsingular.

(b): To show that X is the solution of the Riccati equation, we note that equation (21.4) can be written as

$$S \begin{bmatrix} I \\ X \end{bmatrix} = \begin{bmatrix} I \\ X \end{bmatrix} T_1 S_- T_1^{-1}. \tag{21.5}$$

Pre-multiply (21.5) by $\begin{bmatrix} -X & I \end{bmatrix}$ to get

$$\begin{bmatrix} -X & I \end{bmatrix} S \begin{bmatrix} I \\ X \end{bmatrix} = 0.$$

Equivalently, we get

$$-XA + (I + XG)(A^*)^{-1}(X - Q) = 0. \tag{21.6}$$

We now show that $I + XG$ is invertible. Suppose $I + XG$ is not invertible, then $\exists\, v \in \mathbb{R}^n$ such that

$$v^*(I + XG) = 0. \tag{21.7}$$

Pre-multiply (21.6) by v^*, and we get $v^* X A = 0$. Since A is assumed to be invertible, it follows that $v^* X = 0$. This, in turn, implies $v = 0$ by (21.7). Hence $I + XG$ is invertible and the equation (21.6) can be written as

$$A^*(I + XG)^{-1} X A - X + Q = 0.$$

This is the Riccati equation (21.2).

(c): To show that X is the stabilizing solution, pre-multiply (21.5) by $\begin{bmatrix} I & 0 \end{bmatrix}$ to get

$$A + G(A^*)^{-1} Q - G(A^*)^{-1} X = T_1 S_- T_1^{-1}.$$

The above equation can be simplified by using Riccati equation to get

$$(I + GX)^{-1} A = T_1 S_- T_1^{-1}$$

which is stable. □

Remark 21.3 Let $\mathcal{X}_+(S) = \operatorname{Im} \begin{bmatrix} \tilde{T}_1 \\ \tilde{T}_2 \end{bmatrix}$ and suppose that $\tilde{T}_1$ is nonsingular. Then the Riccati equation has an anti-stabilizing solution $\tilde{X} = \tilde{T}_2 \tilde{T}_1^{-1}$ such that $(I + G\tilde{X})^{-1}A$ is antistable. ♡

Lemma 21.4 *Suppose that G and Q are positive semi-definite and that S has no eigenvalues on unit circle. Let $\mathcal{X}_-(S) = \operatorname{Im} \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$; then $T_1^* T_2 \geq 0$.*

Proof. Let $T = \begin{bmatrix} T_1 \\ T_2 \end{bmatrix}$ and S_- be such that

$$ST = TS_-, \tag{21.8}$$

and S_- has all eigenvalues inside of the unit disc. Let

$$U_k = TS_-^k, \; k = 0, 1, \ldots .$$

Then $U_{k+1} = SU_k$ with $U_0 = T$. Defining $V = \begin{bmatrix} 0 & I \\ 0 & 0 \end{bmatrix}$, we get $T_1^* T_2 = U_0^* V U_0$. Further define

$$\begin{aligned} Y_k &:= -U_k^* V U_k + U_0^* V U_0 \\ &= -\sum_{i=0}^{k-1} (U_{i+1}^* V U_{i+1} - U_i^* V U_i) \\ &= -\sum_{i=0}^{k-1} U_i^* (S^* V S - V) U_i. \end{aligned}$$

Now

$$S^* V S - V = \begin{bmatrix} I & -Q \\ 0 & I \end{bmatrix} \begin{bmatrix} -Q & 0 \\ 0 & -A^{-1} G (A^*)^{-1} \end{bmatrix} \begin{bmatrix} I & 0 \\ Q & I \end{bmatrix} \leq 0$$

since G and Q are assumed to be positive semi-definite. So $Y_k \geq 0$ for all $k \geq 0$. Note that $U_k \to 0$ as $k \to \infty$ since S_- has all eigenvalues inside the unit disc. Therefore $T_1^* T_2 = \lim_{k \to \infty} Y_k \geq 0$. □

Lemma 21.5 *Suppose that G and Q are positive semi-definite. Then $S \in dom(Ric)$ iff (A, G) is stabilizable and S has no eigenvalues on the unit circle.*

Proof. The necessary part is obvious. We now show that the stabilizability of (A, G) and S having no eigenvalues on the unit circle are, in fact, sufficient. To show this, we only need to show that T_1 is nonsingular, i.e. Ker $T_1 = 0$. First, it is claimed that KerT_1 is S_--invariant. To prove this, let $x \in \operatorname{Ker} T_1$. Rewrite (21.8) as

$$\begin{aligned} (A + G(A^*)^{-1}Q)T_1 - G(A^*)^{-1}T_2 &= T_1 S_- && (21.9) \\ -(A^*)^{-1}QT_1 + (A^*)^{-1}T_2 &= T_2 S_-. && (21.10) \end{aligned}$$

Substitute (21.10) into (21.9) to get

$$AT_1 - GT_2 S_- = T_1 S_-, \tag{21.11}$$

and pre-multiply the above equation by $x^* S_-^* T_2^*$ and post-multiply x to get

$$-x^* S_-^* T_2^* G T_2 S_- x = x^* S_-^* T_2^* T_1 S_- x.$$

According to Lemma 21.4 $T_2^* T_1 \geq 0$, we have

$$GT_2 S_- x = 0.$$

This in turn implies that $T_1 S_- x = 0$ from (21.11). Hence Ker T_1 is invariant under S_-.

Now to prove that T_1 is nonsingular, suppose on the contrary that Ker $T_1 \neq 0$. Then $S_-|_{\operatorname{Ker} T_1}$ has an eigenvalue, λ, and a corresponding eigenvector, x:

$$S_- x = \lambda x \tag{21.12}$$

$$|\lambda| < 1, \quad 0 \neq x \in \operatorname{Ker} T_1.$$

Post-multiply (21.10) by x to get

$$(A^*)^{-1} T_2 x = T_2 S_- x = \lambda T_2 x.$$

Now if $T_2 x \neq 0$, then $1/\bar{\lambda}$ is an eigenvalue of A and, furthermore, $0 = GT_2 S_- x = \lambda G T_2 x$, so $GT_2 x = 0$, which is contradictory to the stabilizability assumption of (A, G). Hence we must have $T_2 x = 0$. But this would imply $\begin{bmatrix} T_1 \\ T_2 \end{bmatrix} x = 0$, which is impossible. □

Lemma 21.6 *Let G and Q be positive semi-definite matrices and*

$$S = \begin{bmatrix} A + G(A^*)^{-1}Q & -G(A^*)^{-1} \\ -(A^*)^{-1}Q & (A^*)^{-1} \end{bmatrix}.$$

Then S has no eigenvalues on the unit circle iff (A, G) has no uncontrollable modes and (Q, A) has no unobservable modes on the unit circle.

Proof. ($\Leftarrow$) Suppose, on the contrary, that S has an eigenvalue $e^{j\theta}$. Then

$$S\begin{bmatrix} x \\ y \end{bmatrix} = e^{j\theta}\begin{bmatrix} x \\ y \end{bmatrix}$$

for some $\begin{bmatrix} x \\ y \end{bmatrix} \neq 0$, i.e.,

$$\begin{aligned} (A + G(A^*)^{-1}Q)x - G(A^*)^{-1}y &= e^{j\theta}x \\ -(A^*)^{-1}Qx + (A^*)^{-1}y &= e^{j\theta}y. \end{aligned}$$

Multiplying the second equation by G and adding it to the first one give

$$Ax - e^{j\theta}Gy = e^{j\theta}x \tag{21.13}$$

$$-Qx + y = e^{j\theta}A^*y. \tag{21.14}$$

Pre-multiplying equation (21.13) by $e^{-j\theta}y^*$ and equation (21.14) by x^* yield

$$\begin{aligned} e^{-j\theta}y^*Ax &= y^*Gy + y^*x \\ -x^*Qx + x^*y &= e^{j\theta}x^*A^*y. \end{aligned}$$

Thus

$$-y^*Gy - x^*Qx = 0.$$

It follows that

$$\begin{aligned} y^*G &= 0 \\ Qx &= 0. \end{aligned}$$

Substitute these relationships into (21.13) and (21.14) to get

$$\begin{aligned} Ax &= e^{j\theta}x \\ e^{j\theta}A^*y &= y. \end{aligned}$$

Since x and y cannot be zero simultaneously, $e^{j\theta}$ is either an unobservable mode of (Q, A) or an uncontrollable mode of (A, G), a contradiction.

($\Rightarrow$): Suppose that S has no eigenvalue on the unit circle but $e^{j\theta}$ is an unobservable mode of (Q, A) and x is a corresponding eigenvector. Then it is easy to verify that

$$S\begin{bmatrix} x \\ 0 \end{bmatrix} = e^{j\theta}\begin{bmatrix} x \\ 0 \end{bmatrix},$$

so $e^{j\theta}$ is an eigenvalue of S, again a contradiction. The case for (A, G) having uncontrollable mode on the unit circle can be proven similarly. $\square$

Theorem 21.7 *Let G and Q be positive semi-definite matrices and*

$$S = \begin{bmatrix} A + G(A^*)^{-1}Q & -G(A^*)^{-1} \\ -(A^*)^{-1}Q & (A^*)^{-1} \end{bmatrix}.$$

Then $S \in dom(Ric)$ iff (A, G) is stabilizable and (Q, A) has no unobservable modes on the unit circle. Furthermore, $X = Ric(S) \geq 0$ if $S \in dom(Ric)$ and $X > 0$ if and only if (Q, A) has no unobservable stable modes.

Proof. Let $Q = C^*C$ for some matrix C. The first half of the theorem follows from Lemmas 21.5 and 21.6. Now rewrite the discrete Riccati equation as

$$A^*(I + XG)^{-1}X(I + GX)^{-1}A - X + A^*X(I + GX)^{-2}GXA + C^*C = 0 \tag{21.15}$$

and note that by definition $(I+GX)^{-1}A$ is stable and $A^*X(I+GX)^{-2}GXA+C^*C \geq 0$. Thus $X \geq 0$ by Lyapunov theorem. To show that the kernel of X has the refereed property, suppose $x \in \text{Ker}X$, pre-multiply (21.15) by x^*, and post-multiply by x to get

$$XAx = 0, \; Cx = 0. \tag{21.16}$$

This implies that $\text{Ker}X$ is an A-invariant subspace. If $\text{Ker}X \neq 0$, then there is an $0 \neq x \in \text{Ker}X$, so $Cx = 0$, such that $Ax = \lambda x$. But for $x \in \text{Ker}X$, $(I + GX)^{-1}Ax = \lambda(I + GX)^{-1}x = \lambda x$, so $|\lambda| < 1$ since $(I + GX)^{-1}A$ is stable. Thus λ is an unobservable stable mode of (Q, A).

On the other hand, suppose that $|\lambda| < 1$ is a stable unobservable mode of (Q, A). Then there exists a $x \in \mathbb{C}^n$ such that $Ax = \lambda x$ and $Cx = 0$; do the same pre- and post-multiplications on (21.15) as before to get

$$|\lambda|^2 x^*(XG + I)^{-1}Xx - x^*Xx = 0.$$

This can be rewritten as

$$x^*X^{1/2}[|\lambda|^2(I + X^{1/2}GX^{1/2})^{-1} - I]X^{1/2}x = 0.$$

Now $X \geq 0$, $G \geq 0$, and $|\lambda| < 1$ imply that $|\lambda|^2(I + X^{1/2}GX^{1/2})^{-1} - I < 0$. Hence $Xx = 0$, i.e., X is singular. □

Lemma 21.8 *Suppose that D has full column rank and let $R = D^*D > 0$; then the following statements are equivalent:*

(i) $\begin{bmatrix} A - e^{j\theta}I & B \\ C & D \end{bmatrix}$ *has full column rank for all $\theta \in [0, 2\pi]$.*

(ii) $\left((I - DR^{-1}D^*)C,\, A - BR^{-1}D^*C\right)$ *has no unobservable modes on the unit circle, or, equivalently, $(D_\perp^*C, A - BR^{-1}D^*C)$ has no unobservable modes on the unit circle.*

Proof. Suppose that $e^{j\theta}$ is an unobservable mode of $\big((I-DR^{-1}D^*)C,\, A-BR^{-1}D^*C\big)$; then there is an $x \neq 0$ such that

$$(A-BR^{-1}D^*C)x = e^{j\theta}x, \quad (I-DR^{-1}D^*)Cx = 0$$

i.e.,

$$\begin{bmatrix} A-e^{j\theta}I & B \\ C & D \end{bmatrix} \begin{bmatrix} I & 0 \\ -R^{-1}D^*C & I \end{bmatrix} \begin{bmatrix} x \\ 0 \end{bmatrix} = 0.$$

But this implies that

$$\begin{bmatrix} A-e^{j\theta}I & B \\ C & D \end{bmatrix} \tag{21.17}$$

does not have full column rank. Conversely, suppose that (21.17) does not have full column rank for some θ; then there exists $\begin{bmatrix} u \\ v \end{bmatrix} \neq 0$ such that

$$\begin{bmatrix} A-e^{j\theta}I & B \\ C & D \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} = 0.$$

Now let

$$\begin{bmatrix} u \\ v \end{bmatrix} = \begin{bmatrix} I & 0 \\ -R^{-1}D^*C & I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}.$$

Then

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} I & 0 \\ R^{-1}D^*C & I \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} \neq 0$$

and

$$(A-BR^{-1}D^*C-e^{j\theta}I)x + By = 0 \tag{21.18}$$

$$(I-DR^{-1}D^*)Cx + Dy = 0. \tag{21.19}$$

Pre-multiply (21.19) by D^* to get $y=0$. Then we have

$$(A-BR^{-1}D^*C)x = e^{j\theta}x, \quad (I-DR^{-1}D^*)Cx = 0$$

i.e., $e^{j\theta}$ is an unobservable mode of $\big((I-DR^{-1}D^*)C,\, A-BR^{-1}D^*C\big)$. □

Corollary 21.9 *Suppose that D has full column rank and denote $R = D^*D > 0$. Let S have the form*

$$S = \begin{bmatrix} E+G(E^*)^{-1}Q & -G(E^*)^{-1} \\ -(E^*)^{-1}Q & (E^*)^{-1} \end{bmatrix}$$

where $E = A-BR^{-1}D^*C$, $G = BR^{-1}B^*$, $Q = C^*(I-DR^{-1}D^*)C$, and E is assumed to be invertible. Then $S \in dom(Ric)$ iff (A,B) is stabilizable and $\begin{bmatrix} A-e^{j\theta}I & B \\ C & D \end{bmatrix}$ has full column rank for all $\theta \in [0, 2\pi]$. Furthermore, $X = Ric(S) \geq 0$.

Note that the Riccati equation corresponding to the simplectic matrix in Corollary 21.9 is

$$E^*XE - X - E^*XG(I + XG)^{-1}XE + Q = 0.$$

This equation can also be written as

$$A^*XA - X - (B^*XA + D^*C)^*(D^*D + B^*XB)^{-1}(B^*XA + D^*C) + C^*C = 0.$$

21.3 Bounded Real Functions

Let a real rational transfer matrix be given by

$$M(z) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

where again A is assumed to be nonsingular and the realization is assumed to have no uncontrollable and no unobservable modes on the unit circle. Note again that all results hold for A singular case with the same modification as in the last section. Define $M^{\sim}(z) := M^T(z^{-1})$. Then

$$M^{\sim}(z) = \left[\begin{array}{c|c} (A^*)^{-1} & -(A^*)^{-1}C^* \\ \hline B^*(A^*)^{-1} & D^* - B^*(A^*)^{-1}C^* \end{array}\right].$$

Lemma 21.10 *Let* $M(z) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \in \mathcal{RL}_\infty$ *and let S be a simplectic matrix defined by*

$$S := \begin{bmatrix} A - BB^*(A^*)^{-1}C^*C & BB^*(A^*)^{-1} \\ -(A^*)^{-1}C^*C & (A^*)^{-1} \end{bmatrix}.$$

Then the following statements are equivalent:

(i) $\|M(z)\|_\infty < 1$;

(ii) *S has no eigenvalues on the unit circle and* $\|C(I - A)^{-1}B\| < 1$.

Proof. It is easy to compute that

$$[I - M^{\sim}(z)M(z)]^{-1} = \left[\begin{array}{cc|c} A - BB^*(A^*)^{-1}C^*C & -BB^*(A^*)^{-1} & B \\ (A^*)^{-1}C^*C & (A^*)^{-1} & 0 \\ \hline -B^*(A^*)^{-1}C^*C & -B^*(A^*)^{-1} & I \end{array}\right].$$

It is claimed that $[I - M^{\sim}(z)M(z)]^{-1}$ has no uncontrollable and/or unobservable modes on the unit circle.

To show that, suppose that $\lambda = e^{j\theta}$ is an uncontrollable mode of $[I - M^{\sim}(z)M(z)]^{-1}$. Then $\exists\, q = \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} \in \mathbb{C}^{2n}$ such that

$$q^* \begin{bmatrix} A - BB^*(A^*)^{-1}C^*C & -BB^*(A^*)^{-1} \\ (A^*)^{-1}C^*C & (A^*)^{-1} \end{bmatrix} = e^{j\theta} q^*, \quad q^* \begin{bmatrix} B \\ 0 \end{bmatrix} = 0.$$

Hence $q_1^* B = 0$ and

$$\begin{bmatrix} q_1^* A + q_2^*(A^*)^{-1}C^*C & q_2^*(A^*)^{-1} \end{bmatrix} = e^{j\theta} \begin{bmatrix} q_1^* & q_2^* \end{bmatrix}.$$

There are two possibilities:

1. $q_2 \neq 0$. Then we have $q_2^*(A^*)^{-1} = e^{j\theta} q_2^*$, i.e., $A q_2 = e^{-j\theta} q_2$. This implies $e^{-j\theta}$ is an eigenvalue of A. This is a contradiction since $M(z) \in \mathcal{RL}_\infty$.
2. $q_2 = 0$. Then $q_1^* A = e^{j\theta} q_1^*$, which again implies that $M(z)$ has a mode on the unit circle if $q_1 \neq 0$, again a contradiction.

Similar proof can be done for observability, hence the claim is true.

Now note that

$$S = \begin{bmatrix} -I & \\ & I \end{bmatrix} \begin{bmatrix} A - BB^*(A^*)^{-1}C^*C & -BB^*(A^*)^{-1} \\ (A^*)^{-1}C^*C & (A^*)^{-1} \end{bmatrix} \begin{bmatrix} -I & \\ & I \end{bmatrix}.$$

Hence S does not have eigenvalues on the unit circle. It is clear that we have already proven that S has no eigenvalues on the unit circle iff $(I - M^{\sim}M)^{-1} \in \mathcal{RL}_\infty$. So it is sufficient to show that

$$\|M(z)\|_\infty < 1 \quad \Leftrightarrow (I - M^{\sim}M)^{-1} \in \mathcal{RL}_\infty \text{ and} \|M(1)\| < 1.$$

It is obvious that the right hand side is necessary. To show that it is also sufficient, suppose $\|M(z)\|_\infty \geq 1$, then $\sigma_{max}(M(e^{j\theta})) = 1$ for some $\theta \in [0,\, 2\pi]$, since $\sigma_{max}(M(1)) < 1$ and $M(e^{j\theta})$ is continuous in θ. This implies that 1 is an eigenvalue of $M^*(e^{-j\theta})M(e^{j\theta})$, so $I - M^*(e^{-j\theta})M(e^{j\theta})$ is singular. This contradicts to $(I - M^{\sim}M)^{-1} \in \mathcal{RL}_\infty$. □

In the above Lemma, we have assumed that the transfer matrix is strictly proper. We shall now see how to handle non-strictly proper case. For that purpose we shall focus our attention on the stable system, and we shall give an example below to show why this restriction is sometimes necessary for the technique to work.

We first note that $\mathcal{H}_\infty$-norm of a stable system is defined as

$$\|M(z)\|_\infty = \sup_{|z| \geq 1} \overline{\sigma}\left(M(z)\right).$$

Then it is clear that $\|M(z)\|_\infty \geq \overline{\sigma}\left(M(\infty)\right) = \|D\|$. Thus in particular if $\|M(z)\|_\infty < 1$ then $I - D^*D > 0$.

On the other hand, if a function is only known to be in $\mathcal{RL}_\infty$, the above condition may not be true if the system is not stable.

Example 21.1 Let $0 < \alpha < 1/2$ and let

$$M_1(z) = \frac{z}{z - 1/\alpha} = \left[\begin{array}{c|c} 1/\alpha & \alpha \\ \hline 1 & 1 \end{array}\right] \in \mathcal{RL}_\infty.$$

Then $\|M_1(z)\|_\infty = \frac{\alpha}{1-\alpha} < 1$, but $1 - D^*D = 0$. In general, if $M \in \mathcal{RL}_\infty$ and $\|M\|_\infty < 1$, then $I - D^*D$ can be indefinite. $\diamond$

Lemma 21.11 *Let* $M(z) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$. *Then* $\|M(z)\|_\infty < 1$ *if and only if* $N(z) \in \mathcal{RH}_\infty$ *and* $\|N(z)\|_\infty < 1$ *where*

$$N(z) = \left[\begin{array}{c|c} A + B(I - D^*D)^{-1}D^*C & B(I - D^*D)^{-1/2} \\ \hline (I - DD^*)^{-1/2}C & 0 \end{array}\right] =: \left[\begin{array}{c|c} E & \hat{B} \\ \hline \hat{C} & 0 \end{array}\right].$$

Proof. This is exactly the analogy of Corollary 17.4. $\square$

Theorem 21.12 *Let* $M(z) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$ *and define*

$$\begin{aligned} E &:= A + B(I - D^*D)^{-1}D^*C \\ G &:= -B(I - D^*D)^{-1}B^* \\ Q &:= C^*(I - DD^*)^{-1}C. \end{aligned}$$

Suppose that E is nonsingular and define a simplectic matrix as

$$S := \begin{bmatrix} E + G(E^*)^{-1}Q & -G(E^*)^{-1} \\ -(E^*)^{-1}Q & (E^*)^{-1} \end{bmatrix}.$$

Then the following statements are equivalent:

(a) $\|M(z)\|_\infty < 1$;

(b) S has no eigenvalues on the unit circle and $\|C(I - A)^{-1}B + D\| < 1$;

(c) $\exists X \geq 0$ *such that* $I - D^*D - B^*XB > 0$ *and*

$$E^*XE - X - E^*XG(I + XG)^{-1}XE + Q = 0$$

and $(I + GX)^{-1}E$ is stable. Moreover, $X > 0$ if (C, A) is observable.

(d) $\exists X > 0$ *such that* $I - D^*D - B^*XB > 0$ *and*

$$E^*XE - X - E^*XG(I + XG)^{-1}XE + Q < 0;$$

(e) $\exists X > 0$ *such that*

$$\begin{bmatrix} A & B \\ C & D \end{bmatrix}^* \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} - \begin{bmatrix} X & 0 \\ 0 & I \end{bmatrix} < 0;$$

(f) $\exists T$ *nonsingular such that*

$$\overline{\sigma}\left(\begin{bmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{bmatrix}\right) = \left\| \begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} T & 0 \\ 0 & I \end{bmatrix}^{-1} \right\| < 1;$$

(g) $\mu_\Delta\left(\begin{bmatrix} A & B \\ C & D \end{bmatrix}\right) < 1$ *with* $\Delta \in \mathbf{\Delta}$ *and*

$$\mathbf{\Delta} := \left\{ \begin{bmatrix} \delta_1 I_n & 0 \\ 0 & \Delta_2 \end{bmatrix} : \delta_1 \in \mathbb{C}, \Delta_2 \in \mathbb{C}^{m\times p} \right\} \subset \mathbb{C}^{(n+m)\times(n+p)}$$

(assuming that $M(s)$ *is a* $p \times m$ *matrix and* A *is an* $n \times n$ *matrix).*

Note that the Riccati equation in (c) can also be written as

$$A^*XA - X + (B^*XA + D^*C)^*(I - D^*D - B^*XB)^{-1}(B^*XA + D^*C) + C^*C = 0.$$

Proof. (a)$\Leftrightarrow$(b) follows by Lemma 21.10

(a)$\Rightarrow$(g) follows from Theorem 11.7.

(g)$\Rightarrow$(f) follows from Theorem 11.5.

(f)$\Rightarrow$(e) follows by letting $X = T^*T$.

(e)$\Rightarrow$(d) follows by Schur complementary formula.

(d)$\Rightarrow$(c) can be shown in the same way as in the proof of Theorem 13.11.

(c)$\Rightarrow$(a) We shall only give the proof for $D = 0$ case, the case $D \neq 0$ can be transformed to the zero case by Lemma 21.11. Hence in the following we have $E = A$, $G = -BB^*$, and $Q = C^*C$.

Assuming (c) is satisfied by some $X \geq 0$ and considering the obvious relation with $z := e^{j\theta}$

$$\begin{aligned} (z^{-1}I - A^*)X(zI - A) + (z^{-1}I - A^*)XA + A^*X(zI - A) &= X - A^*XA \\ &= C^*C + A^*XB(I - B^*XB)^{-1}B^*XA. \end{aligned}$$

The last equality is obtained from substituting in Riccati equation. Now pre-multiply the above equation by $B^*(z^{-1}I-A)^{-1}$ and post-multiply by $(zI-A)^{-1}B$ to get

$$I - M^*(z^{-1})M(z) = W^*(z^{-1})W(z)$$

where

$$W(z) = \left[\begin{array}{c|c} A & B \\ \hline (I-B^*XB)^{-1/2}B^*XA & -(I-B^*XB)^{1/2} \end{array}\right].$$

Suppose $W(e^{j\theta})v = 0$ for some θ and v; then $e^{j\theta}$ is a zero of $W(z)$ if $v \neq 0$. However, all the zeros of $W(z)$ are given by the eigenvalues of

$$A + B(I-B^*XB)^{-1}B^*XA = (I-BB^*X)^{-1}A$$

that are all inside of the unit circle. Hence $e^{j\theta}$ cannot be a zero of W.

Therefore, we get $I - M^*(e^{-j\theta})M(e^{j\theta}) > 0$ for all $\theta \in [0, 2\pi]$, i.e., $\|M\|_\infty < 1$. □

The following more general results can be proven easily following the same procedure as in the proof of (c)⇒(a).

Corollary 21.13 *Let* $M(z) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right] \in \mathcal{RL}_\infty$ *and suppose* $\exists\, X = X^*$ *such that*

$$A^*XA - X + A^*XB(I-B^*XB)^{-1}B^*XA + C^*C = 0.$$

Then

$$I - M^*(z^{-1})M(z) = W^*(z^{-1})(I-B^*XB)W(z)$$

where

$$W(z) = \left[\begin{array}{c|c} A & B \\ \hline (I-B^*XB)^{-1}B^*XA & -I \end{array}\right].$$

Moreover, the following statements hold:

(1) if $I - B^*XB > 0\,(<0)$*, then* $\|M(z)\|_\infty \leq 1\,(\geq 1)$ *;*

(2) if $I - B^*XB > 0\,(<0)$ *and* $|\lambda_i\{(I-BB^*X)^{-1}A\}| \neq 1$*, then* $\|M(z)\|_\infty < 1\,(>1)$*.*

Remark 21.4 As in the continuous time case, the equivalence between (a) and (b) in Theorem 21.12 can be used to compute the $\mathcal{H}_\infty$ norm of a discrete time transfer matrix. ♡

21.4 Matrix Factorizations

21.4.1 Inner Functions

A transfer matrix $N(z)$ is called *inner* if $N(z)$ is stable and $N^*(z)N(z) = I$ for all $z = e^{j\theta}$. Note that if N has no poles at the origin, then $N^*(e^{j\theta}) = N^{\sim}(e^{j\theta})$. A transfer matrix is called *outer* if all its transmission zeros are stable (i.e., inside of the unit disc).

Lemma 21.14 *Let* $N = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$ *and suppose that* $X = X^*$ *satisfies*

$$A^*XA - X + C^*C = 0.$$

Then

(a) $D^*C + B^*XA = 0$ *implies* $N^{\sim}N = (D - CA^{-1}B)^*D$*;*

(b) $X \geq 0$*,* (A, B) *controllable, and* $N^{\sim}N = (D-CA^{-1}B)^*D$ *implies* $D^*C+B^*XA = 0$.

Proof. The results follow by noting the following equality:

$$\begin{aligned} N^{\sim}(z)N(z) &= \left[\begin{array}{cc|c} A & 0 & B \\ (A^*)^{-1}C^*C & (A^*)^{-1} & (A^*)^{-1}C^*D \\ \hline (D - CA^{-1}B)^*C & -B^*(A^*)^{-1} & (D - CA^{-1}B)^*D \end{array}\right] \\ &= \left[\begin{array}{cc|c} A & 0 & B \\ 0 & (A^*)^{-1} & XB + (A^*)^{-1}C^*D \\ \hline D^TC + B^*XA & -B^*(A^*)^{-1} & (D - CA^{-1}B)^*D \end{array}\right]. \end{aligned}$$

□

The following corollary is a special case of this lemma which gives the necessary and sufficient conditions of a discrete inner transfer matrix with the state-space representation.

Corollary 21.15 *Suppose that* $N(z) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$ *is a controllable realization; then* $N(z)$ *is* inner *if and only if there exists a matrix* $X = X^* \geq 0$ *such that*

(a) $A^*XA - X + C^*C = 0$

(b) $D^*C + B^*XA = 0$

(c) $(D - CA^{-1}B)^*D = D^*D + B^*XB = I$.

The following alternative characterization of the inner transfer matrix is often useful and insightful.

Corollary 21.16 *Let* $N(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$ *and assume that there exists a* T *nonsingular such that*

$$P = \begin{bmatrix} TAT^{-1} & TB \\ CT^{-1} & D \end{bmatrix} \text{ and } P^*P = I.$$

Then $N(z)$ *is an inner. Furthermore, if the realization of* N *is minimal, then such* T *exists.*

Proof. Rewrite $P^*P = I$ as

$$\begin{bmatrix} A^* & C^* \\ B^* & D^* \end{bmatrix} \begin{bmatrix} T^*T & \\ & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} = \begin{bmatrix} T^*T & \\ & I \end{bmatrix}.$$

Let $X = T^*T$, and then

$$\begin{bmatrix} A^*XA - X + C^*C & A^*XB + C^*D \\ B^*XA + D^*C & B^*XB + D^*D - I \end{bmatrix} = 0.$$

This is the desired equation, so N is inner. On the other hand, if the realization is minimal, then $X > 0$. This implies that T exists.

□

In a similar manner, Corollary 21.15 can be used to derive the state-space representation of the complementary inner factor (*CIF*).

Lemma 21.17 *Suppose that a* $p \times m$ $(p > m)$ *transfer matrix* $N(z)$ *(minimal) is inner; then there exists a* $p \times (p-m)$ CIF $N_\perp \in \mathcal{RH}_\infty$ *such that the matrix* $\begin{bmatrix} N & N_\perp \end{bmatrix}$ *is square and inner. A particular realization is*

$$N_\perp(z) = \left[\begin{array}{c|c} A & Y \\ \hline C & Z \end{array}\right]$$

where Y *and* Z *satisfy*

$$A^*XY + C^*Z = 0 \tag{21.20}$$

$$B^*XY + D^*Z = 0 \tag{21.21}$$

$$Z^*Z + Y^*XY = I. \tag{21.22}$$

Proof. Note that $\begin{bmatrix} N & N_\perp \end{bmatrix} = \left[\begin{array}{c|cc} A & B & Y \\ \hline C & D & Z \end{array}\right]$ is inner. Now it is easy to prove the results by using the formula in Corollary 21.15. □

21.4.2 Coprime Factorizations

Recall that two transfer matrices $M(z)$, $N(z) \in \mathcal{RH}_\infty$ are said to be right coprime if $\begin{bmatrix} M \\ N \end{bmatrix}$ is left invertible in $\mathcal{RH}_\infty$ i.e., $\exists\, U, V \in \mathcal{RH}_\infty$ such that

$$U(z)N(z) + V(z)M(z) = I.$$

The left coprime is defined analogously. A plant $G(z) \in \mathcal{RL}_\infty$ is said to have double coprime factorization if $\exists$ a right coprime factorization $G = NM^{-1}$, a left coprime factorization $G = \tilde{M}^{-1}\tilde{N}$, and $U, V, \tilde{U}, \tilde{V} \in \mathcal{RH}_\infty$ such that

$$\begin{bmatrix} V & U \\ -\tilde{N} & \tilde{M} \end{bmatrix} \begin{bmatrix} M & -\tilde{U} \\ N & \tilde{V} \end{bmatrix} = I. \tag{21.23}$$

The state space formulae for discrete time transfer matrix coprime factorization are the same as for the continuous time. They are given by the following theorem.

Theorem 21.18 *Let* $G(z) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RL}_\infty$ *be a stabilizable and detectable realization. Choose* F *and* L *such that* $A + BF$ *and* $A + LC$ *are both stable. Let* $U, V, \tilde{U}, \tilde{V}, N, M, \tilde{N}$, *and* $\tilde{M}$ *be given as follows*

$$\begin{bmatrix} M \\ N \end{bmatrix} := \left[\begin{array}{c|c} A + BF & BZ_r \\ \hline F & Z_r \\ C + DF & DZ_r \end{array}\right]$$

$$\begin{bmatrix} \tilde{U} \\ \tilde{V} \end{bmatrix} := \left[\begin{array}{c|c} A + BF & LZ_l^{-1} \\ \hline F & 0 \\ -(C + DF) & Z_l^{-1} \end{array}\right]$$

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} := \left[\begin{array}{c|cc} A + LC & L & B + LD \\ \hline Z_l C & Z_l & Z_l D \end{array}\right]$$

$$\begin{bmatrix} U & V \end{bmatrix} := \left[\begin{array}{c|cc} A + LC & L & -(B + LD) \\ \hline Z_l^{-1} F & 0 & Z_l^{-1} \end{array}\right]$$

where Z_r *and* Z_l *are any nonsingular matrices. Then* $G = NM^{-1} = \tilde{M}^{-1}\tilde{N}$ *are* rcf *and* lcf*, respectively, and (21.23) is satisfied.*

Some coprime factorizations are particularly interesting, for example, the *coprime factorization with inner numerator.* This factorization in the case of $G(z) \in \mathcal{RH}_\infty$ yields an *inner-outer factorization.*

Theorem 21.19 *Assume that* (A, B) *is stabilizable,* $\begin{bmatrix} A - e^{j\theta}I & B \\ C & D \end{bmatrix}$ *has full column rank for all* $\theta \in [0, 2\pi]$*, and* D *has full column rank. Then there exists a right coprime factorization* $G = NM^{-1}$ *such that* N *is* inner*. Furthermore, a particular realization is given by*

$$\begin{bmatrix} M \\ N \end{bmatrix} := \left[\begin{array}{c|c} A + BF & BR^{-1/2} \\ \hline F & R^{-1/2} \\ C + DF & DR^{-1/2} \end{array}\right]$$

where

$$R = D^*D + B^*XB$$

$$F = -R^{-1}(B^*XA + D^*C)$$

and $X = X^* \geq 0$ *is the unique stabilizing solution*

$$A_D^*X(I + B(D^*D)^{-1}B^*X)^{-1}A_D - X + C^*D_\perp D_\perp^* C = 0$$

where $A_D := A - B(D^*D)^{-1}D^*C$*.*

Using Lemma 21.17, the complementary inner factor of N in Theorem 21.19 can be obtained as follows

$$N_\perp = \left[\begin{array}{c|c} A + BF & Y \\ \hline C + DF & Z \end{array}\right]$$

where Y and Z satisfy

$$A^*XY + C^*Z = 0$$

$$B^*XY + D^*Z = 0$$

$$Z^*Z + Y^*XY = I.$$

Note that Y and Z are only related to F implicitly through X.

Remark 21.5 If $G(z) \in \mathcal{RH}_\infty$, then the denominator matrix M in Theorem 21.19 is an outer. Hence, the factorization $G = N(M^{-1})$ is an inner-outer factorization. ♡

Suppose that the system $G(z)$ is not stable; then a *coprime factorization with inner denominator* can also be obtained by solving a special Riccati equation.

Theorem 21.20 *Assume that* $G(z) := \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RL}_\infty$ *and that* (A, B) *is stabilizable. Then there exists a right coprime factorization* $G = NM^{-1}$ *such that* M *is* inner *if and only if* G *has no poles on the unit circle. A particular realization is*

$$\begin{bmatrix} M \\ N \end{bmatrix} := \left[\begin{array}{c|c} A + BF & BR^{-1/2} \\ \hline F & R^{-1/2} \\ C + DF & DR^{-1/2} \end{array}\right]$$

where

$$R = I + B^*XB$$
$$F = -R^{-1}B^*XA$$

and $X = X^* \geq 0$ *is the unique stabilizing solution to*

$$A^*X(I + BB^*X)^{-1}A - X = 0.$$

Another special coprime factorization is called *normalized coprime factorization* which has found applications in many control problems such as model reduction controller design and gap metric characterization.

Recall that a right coprime factorization of $G = NM^{-1}$ with $N, M \in \mathcal{RH}_\infty$ is called a *normalized right coprime factorization* if

$$M^\sim M + N^\sim N = I$$

i.e., if $\begin{bmatrix} M \\ N \end{bmatrix}$ is an inner.

Similarly, an *lcf* $G = \tilde{M}^{-1}\tilde{N}$ is called a *normalized left coprime factorization* if $\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}$ is a co-inner. Then the following results follow in the same way as for the continuous time case.

Theorem 21.21 *Let a realization of* G *be given by*

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

and define

$$R = I + D^*D > 0, \qquad \tilde{R} = I + DD^* > 0.$$

(a) *Suppose that* (A, B) *is stabilizable and that* (C, A) *has no unobservable modes on the imaginary axis. Then there is a normalized right coprime factorization* $G = NM^{-1}$

$$\begin{bmatrix} M \\ N \end{bmatrix} := \left[\begin{array}{c|c} A + BF & BZ^{-1/2} \\ \hline F & Z^{-1/2} \\ C + DF & DZ^{-1/2} \end{array}\right] \in \mathcal{RH}_\infty$$

where

$$Z = R + B^*XB$$
$$F = -Z^{-1}(B^*XA + D^*C)$$

and $X = X^* \geq 0$ *is the unique stabilizing solution*

$$A_n^*X(I + BR^{-1}B^*X)^{-1}A_n - X + C^*\tilde{R}^{-1}C = 0$$

where $A_n := A - BR^{-1}D^*C$.

(b) *Suppose that* (C, A) *is detectable and that* (A, B) *has no uncontrollable modes on the imaginary axis. Then there is a normalized left coprime factorization* $G = \tilde{M}^{-1}\tilde{N}$

$$\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix} := \left[\begin{array}{c|cc} A + LC & L & B + LD \\ \hline \tilde{Z}^{-1/2}C & \tilde{Z}^{-1/2} & \tilde{Z}^{-1/2}D \end{array}\right]$$

where

$$\tilde{Z} = \tilde{R} + CYC^*$$
$$L = -(BD^* + AYC^*)\tilde{Z}^{-1}$$

and $Y = Y^* \geq 0$ *is the unique stabilizing solution*

$$\tilde{A}_nY(I + C^*\tilde{R}^{-1}CY)^{-1}\tilde{A}_n^* - Y + BR^{-1}B^* = 0$$

where $\tilde{A}_n := A - BD^*\tilde{R}^{-1}C$.

(c) *The controllability Gramian* P *and observability Gramian* Q *of* $\begin{bmatrix} M \\ N \end{bmatrix}$ *are given by*

$$P = (I + YX)^{-1}Y, \quad Q = X$$

while the controllability Gramian $\tilde{P}$ *and observability Gramian* $\tilde{Q}$ *of* $\begin{bmatrix} \tilde{M} & \tilde{N} \end{bmatrix}$ *are given by*

$$\tilde{P} = Y, \quad \tilde{Q} = (I + XY)^{-1}X.$$

21.4.3 Spectral Factorizations

The following theorem gives a solution to a special class of spectral factorization problems.

Theorem 21.22 *Assume* $G(z) := \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$ *and* $\gamma > \|G(z)\|_\infty$. *Then, there exists a transfer matrix* $M \in \mathcal{RH}_\infty$ *such that* $M^*M = \gamma^2 I - G^*G$ *and* $M^{-1} \in \mathcal{RH}_\infty$. *A particular realization of* M *is*

$$M(z) = \left[\begin{array}{c|c} A & B \\ \hline -R^{1/2}F & R^{1/2} \end{array}\right]$$

where

$$R_D = \gamma^2 I - D^* D$$

$$R = R_D - B^* X B$$

$$F = (R_D - B^* X B)^{-1}(B^* X A + D^* C)$$

and $X = X^* \geq 0$ *is the stabilizing solution of*

$$A_s^* X (I - B R_D^{-1} B^* X)^{-1} A_s - X + C^*(I + D R_D^{-1} D^*) C = 0$$

where $A_s := A + B R_D^{-1} D^* C$.

Similar to the continuous time, we have the following theorem.

Theorem 21.23 *Let* $G(z) := \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty$ *with* D *full row rank and* $G(e^{j\theta})G^*(e^{j\theta}) > 0$ *for all* θ*. Then, there exists a transfer matrix* $M \in \mathcal{RH}_\infty$ *such that* $M^* M = G G^*$*. A particular realization of* M *is*

$$M(z) = \left[\begin{array}{c|c} A & B_W \\ \hline C_W & D_W \end{array}\right]$$

where

$$\begin{aligned} B_W &= APC^* + BD^* \\ D_W^* D_W &= DD^* \\ C_W &= D_W (DD^*)^{-1}(C - B_W^* X A) \end{aligned}$$

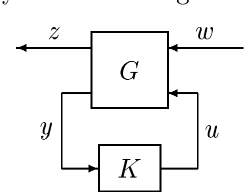
and

$$APA^* - P + BB^* = 0$$

$$A^* X A - X + (C - B_W^* X A)^* (DD^*)^{-1} (C - B_W^* X A) = 0.$$

21.5 Discrete Time $\mathcal{H}_2$ Control

Consider the system described by the block diagram



The realization of the transfer matrix G is taken to be of the form

$$G(z) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & D_{11} & D_{12} \\ C_2 & D_{21} & 0 \end{array}\right] =: \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right].$$

The following assumptions are made:

(A1) (A, B_2) is stabilizable and (C_2, A) is detectable;

(A2) D_{12} is full column rank with $\begin{bmatrix} D_{12} & D_\perp \end{bmatrix}$ unitary and D_{21} is full row rank with $\begin{bmatrix} D_{21} \\ \tilde{D}_\perp \end{bmatrix}$ unitary;

(A3) $\begin{bmatrix} A - e^{j\theta}I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all $\theta \in [0, 2\pi]$;

(A4) $\begin{bmatrix} A - e^{j\theta}I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ has full row rank for all $\theta \in [0, 2\pi]$.

The problem in this section is to find an admissible controller K which minimizes $\|T_{zw}\|_2$.

Denote

$$A_x := A - B_2 D_{12}^* C_1, \quad A_y := A - B_1 D_{21}^* C_2.$$

Let $X_2 \geq 0$ and $Y_2 \geq 0$ be the stabilizing solutions to the following Riccati equations:

$$A_x^* (I + X_2 B_2 B_2^*)^{-1} X_2 A_x - X_2 + C_1^* D_\perp D_\perp^* C_1 = 0$$

and

$$A_y (I + Y_2 C_2^* C_2)^{-1} Y_2 A_y^* - Y_2 + B_1 \tilde{D}_\perp^* \tilde{D}_\perp B_1^* = 0.$$

Note that the stabilizing solutions exist by the assumptions (A3) and (A4). Note also that if A_x and A_y are nonsingular, the solutions can be obtained through the following two simplectic matrices:

$$H_2 := \begin{bmatrix} A_x + B_2 B_2^* (A_x^*)^{-1} C_1^* D_\perp D_\perp^* C_1 & -B_2 B_2^* (A_x^*)^{-1} \\ -(A_x^*)^{-1} C_1^* D_\perp D_\perp^* C_1 & (A_x^*)^{-1} \end{bmatrix}$$

$$J_2 := \begin{bmatrix} A_y^* + C_2^* C_2 A_y^{-1} B_1 \tilde{D}_\perp^* \tilde{D}_\perp B_1^* & -C_2^* C_2 A_y^{-1} \\ -A_y^{-1} B_1 \tilde{D}_\perp^* \tilde{D}_\perp B_1^* & A_y^{-1} \end{bmatrix}.$$

Define

$$\begin{aligned}
R_b &:= I + B_2^* X_2 B_2 \\
F_2 &:= -(I + B_2^* X_2 B_2)^{-1}(B_2^* X_2 A + D_{12}^* C_1) \\
F_0 &:= -(I + B_2^* X_2 B_2)^{-1}(B_2^* X_2 B_1 + D_{12}^* D_{11}) \\
L_2 &:= -(A Y_2 C_2^* + B_1 D_{21}^*)(I + C_2 Y_2 C_2^*)^{-1} \\
L_0 &:= (F_2 Y_2 C_2^* + F_0 D_{21}^*)(I + C_2 Y_2 C_2^*)^{-1}
\end{aligned}$$

and

$$A_{F_2} := A + B_2 F_2, \quad C_{1F_2} := C_1 + D_{12} F_2$$

$$A_{L_2} := A + L_2 C_2, \quad B_{1L_2} := B_1 + L_2 D_{21}$$

$$\hat{A}_2 := A + B_2 F_2 + L_2 C_2$$

$$G_c(z) := \left[\begin{array}{c|c} A_{F_2} & B_1 + B_2 F_0 \\ \hline C_{1F_2} & D_{11} + D_{12} F_0 \end{array}\right]$$

$$G_f(z) := \left[\begin{array}{c|c} A_{L_2} & B_{1L_2} \\ \hline R_b^{1/2}(L_0 C_2 - F_2) & R_b^{1/2}(L_0 D_{21} - F_0) \end{array}\right].$$

Theorem 21.24 *The unique optimal controller is*

$$K_{opt}(z) := \left[\begin{array}{c|c} \hat{A}_2 - B_2 L_0 C_2 & -(L_2 - B_2 L_0) \\ \hline F_2 - L_0 C_2 & L_0 \end{array}\right].$$

Moreover, $\min \|T_{zw}\|_2^2 = \|G_c\|_2^2 + \|G_f\|_2^2$.

Remark 21.6 Note that for a discrete time transfer matrix $G(z) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_2$, its $\mathcal{H}_2$ norm can be computed as

$$\|G(z)\|_2^2 = \mathrm{Trace}\{D^* D + B^* L_o B\} = \mathrm{Trace}\{D D^* + C L_c C^*\}$$

where L_c and L_o are the controllability and observability Gramians

$$A L_c A^* - L_c + B B^* = 0$$

$$A^* L_o A - L_o + C^* C = 0.$$

Using the above formula, we can compute $\min \|T_{zw}\|_2^2$ by noting that X_2 and Y_2 satisfy the equations

$$A_{F_2}^* X_2 A_{F_2} - X_2 + C_{1F_2}^* C_{1F_2} = 0$$

$$A_{L_2} Y_2 A_{L_2}^* - Y_2 + B_{1L_2} B_{1L_2}^* = 0.$$

For example,

$$\|G_c\|_2^2 = \text{Trace}\,\{(D_{11}+D_{12}F_0)^*(D_{11}+D_{12}F_0) + (B_1+D_2F_0)X_2(B_1+D_2F_0)\}$$

and

$$\|G_f\|_2^2 = \text{Trace}\,R_b\,\{(L_0D_{21}-F_0)(L_0D_{21}-F_0)^* + (L_0C_2-F_2)Y_2(L_0C_2-F_2)^*\}\,.$$

♡

Proof. Let x denote the states of the system G. Then the system can be written as

$$\begin{aligned} \dot{x} &= Ax + B_1w + B_2u & (21.24)\\ z &= C_1x + D_{11}w + D_{12}u & (21.25)\\ y &= C_2x + D_{21}w. & (21.26)\end{aligned}$$

Define $\nu := u - F_2x - F_0w$; then the transfer function from w, ν to z becomes

$$z = \left[\begin{array}{c|cc} A_{F_2} & B_1+B_2F_0 & B_2 \\ \hline C_{1F_2} & D_{11}+D_{12}F_0 & D_{12}\end{array}\right]\left[\begin{array}{c} w \\ \nu \end{array}\right] = G_cw + UR_b^{1/2}\nu$$

where

$$U(s) := \left[\begin{array}{c|c} A_{F_2} & B_2R_b^{-1/2} \\ \hline C_{1F_2} & D_{12}R_b^{-1/2}\end{array}\right].$$

It is easy to shown that U is an inner and that $U^{\sim}G_c \in \mathcal{RH}_2^{\perp}$. Now denote the transfer function from w to ν by $T_{\nu w}$. Then

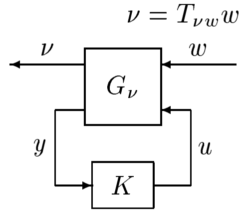
$$T_{zw} = G_c + UR_b^{1/2}T_{\nu w}$$

and

$$\|T_{zw}\|_2^2 = \|G_c\|_2^2 + \left\|R_b^{1/2}T_{\nu w}\right\|_2^2 \geq \|G_c\|_2^2$$

for any given stabilizing controller K. Hence if the states (x) and the disturbance (w) are both available for feedback (i.e., full information control) and $u = F_2x + F_0w$, then $T_{\nu w} = 0$ and $\|T_{zw}\|_2 = \|G_c\|_2$. Therefore, $u = F_2x + F_0w$ is an optimal full information control law. Note that

$$\nu = T_{\nu w}w, \quad T_{\nu w} = \mathcal{F}_\ell(G_\nu, K)$$



$$G_\nu = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline -F_2 & -F_0 & I \\ C_2 & D_{21} & 0 \end{array}\right]$$

Since $A - B_2(-F_2) = A + B_2F_2$ is stable, from the relationship between an output estimation (OE) problem and a full control (FC) problem, all admissible controllers for G_ν (hence for the output feedback problem) can be written as

$$K = \mathcal{F}_\ell(M_t, K_{FC}), \quad M_t = \left[\begin{array}{c|cc} A + B_2F_2 & 0 & \begin{bmatrix} I & -B_2 \end{bmatrix} \\ \hline -F_2 & 0 & \begin{bmatrix} 0 & I \end{bmatrix} \\ C_2 & I & \begin{bmatrix} 0 & 0 \end{bmatrix} \end{array}\right]$$

where K_{FC} is an internally stabilizing controller for the following system:

$$\hat{G}_\nu = \left[\begin{array}{c|c|c} A & B_1 & \begin{bmatrix} I & 0 \end{bmatrix} \\ \hline -F_2 & -F_0 & \begin{bmatrix} 0 & I \end{bmatrix} \\ C_2 & D_{21} & \begin{bmatrix} 0 & 0 \end{bmatrix} \end{array}\right].$$

Furthermore, $T_{\nu w} = \mathcal{F}_\ell(G_\nu, \mathcal{F}_\ell(M_t, K_{FC})) = \mathcal{F}_\ell(\hat{G}_\nu, K_{FC})$. Hence

$$\min_K \|T_{zw}\|_2^2 = \|G_cB_1\|_2^2 + \min_{K_{FC}} \|R_b^{1/2}\mathcal{F}_\ell(\hat{G}_\nu, K_{FC})\|_2^2.$$

Next define

$$\begin{aligned} \tilde{G}_\nu &:= \begin{bmatrix} R_b^{1/2} & 0 \\ 0 & I \end{bmatrix} \hat{G}_\nu \left[\begin{array}{c:cc} I & 0 & 0 \\ \hdashline 0 & I & 0 \\ 0 & 0 & R_b^{-1/2} \end{array}\right] \\ &= \left[\begin{array}{c|c|c} A & B_1 & \begin{bmatrix} I & 0 \end{bmatrix} \\ \hline -R_b^{1/2}F_2 & -R_b^{1/2}F_0 & \begin{bmatrix} 0 & I \end{bmatrix} \\ C_2 & D_{21} & \begin{bmatrix} 0 & 0 \end{bmatrix} \end{array}\right] \end{aligned}$$

and

$$\tilde{K}_{FC} := \begin{bmatrix} I & 0 \\ 0 & R_b^{1/2} \end{bmatrix} K_{FC}.$$

Then it is easy to see that

$$R_b^{1/2}\mathcal{F}_\ell(\hat{G}_\nu, K_{FC}) = \mathcal{F}_\ell(\tilde{G}_\nu, \tilde{K}_{FC})$$

and

$$\min_{K_{FC}} \|R_b^{1/2}\mathcal{F}_\ell(\hat{G}_\nu, K_{FC})\|_2 = \min_{\tilde{K}_{FC}} \|\mathcal{F}_\ell(\tilde{G}_\nu, \tilde{K}_{FC})\|_2.$$

A controller minimizing $\|\mathcal{F}_\ell(\tilde{G}_\nu, \tilde{K}_{FC})\|_2$ is $\tilde{K}_{FC} = \begin{bmatrix} L_2 \\ R_b^{1/2} L_0 \end{bmatrix}$ since the transpose (or dual) of $\mathcal{F}_\ell(\tilde{G}_\nu, \tilde{K}_{FC})$ is a full information feedback problem considered at the beginning of the proof. Hence we get

$$\mathcal{F}_\ell(\tilde{G}_\nu, \begin{bmatrix} L_2 \\ R_b^{1/2} L_0 \end{bmatrix}) = G_f$$

and

$$\min_K \|T_{zw}\|_2^2 = \|G_c\|_2^2 + \|G_f\|_2^2.$$

Finally, the optimal output feedback controller is given by

$$K = \mathcal{F}_\ell(M_t, \begin{bmatrix} L_2 \\ L_0 \end{bmatrix}) = \left[\begin{array}{c|c} A + B_2F_2 + L_2C_2 - B_2L_0C_2 & L_2 - B_2L_0 \\ \hline -F_2 + L_0C_2 & L_0 \end{array}\right].$$

The proof of uniqueness is similar to the continuous time case, and hence omitted. □

It should be noted that in contrast with the continuous time the full information optimal control problem in the discrete time is not a state feedback even when $D_{11} = 0$.

The discrete time $\mathcal{H}_\infty$ control problem is much more involved and it is probably more effective to obtain the discrete solution by using a bilinear transformation.

21.6 Discrete Balanced Model Reduction

In this section, we will show another application of the LFT machinery in discrete balanced model reduction. We will show an elegant proof of the balanced truncation error bound.

Consider a stable discrete time system $G(z)$ and assume that the transfer matrix has the following realization:

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] \in \mathcal{RH}_\infty.$$

Let P and Q be two positive semi-definite symmetric matrices such that

$$APA^* - P + BB^* \leq 0 \tag{21.27}$$

$$A^*QA - Q + C^*C \leq 0. \tag{21.28}$$

Without loss of generality, we shall assume that

$$P = Q = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$$

with

$$\begin{aligned}\Sigma_1 &= \operatorname{diag}(\sigma_1 I_{s_1}, \sigma_2 I_{s_2}, \ldots, \sigma_r I_{s_r}) \geq 0\\ \Sigma_2 &= \operatorname{diag}(\sigma_{r+1} I_{s_{r+1}}, \sigma_{r+2} I_{s_{r+2}}, \ldots, \sigma_n I_{s_n}) \geq 0\end{aligned}$$

where s_i denotes the multiplicity of σ_i. (Note that the singular values are not necessarily ordered.) Moreover, the realization for $G(z)$ is partitioned conformably with P and Q:

$$G(s) = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right].$$

Theorem 21.25 *Suppose $\Sigma_1 > 0$. Then A_{11} is stable.*

Proof. We shall first assume $\Sigma_2 > 0$. From equation (21.28), we have

$$A_{11}^*\Sigma_1 A_{11} - \Sigma_1 + A_{21}^*\Sigma_2 A_{21} + C_1^* C_1 \leq 0. \tag{21.29}$$

Assume that λ is an eigenvalue of A_{11}; then there is an $x \neq 0$ such that

$$A_{11}x = \lambda x. \tag{21.30}$$

Now pre-multiply x^* and post-multiply x to equation (21.29) to get

$$(|\lambda|^2 - 1)x^*\Sigma_1 x + x^* A_{21}^*\Sigma_2 A_{21}x + x^* C_1^* C_1 x \leq 0.$$

It is clear that $|\lambda| \leq 1$. However, if $|\lambda| = 1$, say $\lambda = e^{j\theta}$ for some θ, we have

$$A_{21}x = 0, \quad C_1 x = 0.$$

These equations together with equation (21.30) imply that

$$\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}\begin{bmatrix} x \\ 0 \end{bmatrix} = e^{j\theta}\begin{bmatrix} x \\ 0 \end{bmatrix}$$

i.e., $e^{j\theta}$ is an eigenvalue of A. This contradicts the stability assumption of A, so A_{11} is stable.

Now assume that Σ_2 is singular, and we will show that we can remove all those states corresponding to the zero singular values without changing the system stability. For that purpose, we assume $\Sigma_2 = 0$. Then the inequality (21.27) can be written as

$$\begin{bmatrix} A_{11}\Sigma_1 A_{11}^* - \Sigma_1 + B_1 B_1^* & A_{11}\Sigma_1 A_{21}^* + B_1 B_2^* \\ A_{21}\Sigma_1 A_{11}^* + B_2 B_1^* & A_{21}\Sigma_1 A_{21}^* + B_2 B_2^* \end{bmatrix} \leq 0.$$

This implies that

$$A_{11}\Sigma_1 A_{11}^* - \Sigma_1 + B_1 B_1^* \leq 0 \tag{21.31}$$

and

$$A_{21}\Sigma_1 A_{21}^* + B_2 B_2^* \leq 0. \tag{21.32}$$

But inequality (21.32) implies that

$$A_{21} = 0, \quad B_2 = 0.$$

Hence we have

$$A = \begin{bmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{bmatrix},$$

and the stability of A_{11} and A_{22} are ensured.

Substitute this A matrix into the inequality (21.28), and we obtain

$$A_{11}^*\Sigma_1 A_{11} - \Sigma_1 + C_1^* C_1 \leq 0.$$

The subsystem with A_{11} still satisfies inequalities (21.27) and (21.28) with $\Sigma_1 > 0$. This proves that we can assume without loss of generality that $\Sigma_2 > 0$. □

Remark 21.7 It is important to note that the realization for the truncated subsystem

$$G_r = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$$

is still *balanced* in some sense[1] since the system parameters satisfy the following equations:

$$A_{11}\Sigma_1 A_{11}^* - \Sigma_1 + A_{12}\Sigma_2 A_{12}^* + B_1 B_1^* \leq 0$$
$$A_{11}^*\Sigma_1 A_{11} - \Sigma_1 + A_{21}^*\Sigma_2 A_{21} + C_1^* C_1 \leq 0.$$

But these equations imply that

$$A_{11}\Sigma_1 A_{11}^* - \Sigma_1 + B_1 B_1^* \leq 0$$

$$A_{11}^*\Sigma_1 A_{11} - \Sigma_1 + C_1^* C_1 \leq 0$$

hold. ♡

Theorem 21.26 *Suppose* $G_r = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$. *Then* $\|G - G_r\|_\infty \leq 2\sum_{i=r+1}^{n} \sigma_i$.

In particular, $\|G\|_\infty \leq \|D\| + 2\sum_{i=1}^{n} \sigma_i$.

[1] Balanced in the sense that the same inequalities as (21.27) and (21.28) are satisfied.

Without loss of generality, we shall assume $\sigma_n = 1$. We will prove that for $\Sigma_2 = \sigma_n I = I$, we have

$$\|G - G_r\|_\infty \leq 2, \; r = n - 1.$$

Then the theorem follows immediately by scaling and recursively applying this result since the reduced system G_r is still balanced.

It will be seen that it is more convenient to set $\Lambda = \Sigma_1^{1/2}$. The proof of the theorem will follow from the following two lemmas and the bounded real lemma which establishes the relationship between the $\mathcal{H}_\infty$ norm of a transfer matrix and its realizations. (Note that in the following, a constant matrix X is said to be *contractive* or *a contraction* if $\|X\| \leq 1$ and *strictly contractive* if $\|X\| < 1$).

The lemma below shows that for any stable system there is a realization such that $\begin{bmatrix} A & B \end{bmatrix}$ is a contraction, and, similarly, there is another realization such that $\begin{bmatrix} A \\ C \end{bmatrix}$ is a contraction.

Lemma 21.27 *Suppose that a realization of the transfer matrix G satisfies $P = Q = \mathrm{diag}\{\Lambda^2, I\}$; then*

$$\begin{bmatrix} \Lambda^{-1}A_{12} & \Lambda^{-1}A_{11}\Lambda & \Lambda^{-1}B_1 \\ A_{22}\Lambda & A_{21} & B_2 \end{bmatrix}$$

and

$$\begin{bmatrix} A_{21}\Lambda^{-1} & A_{22} \\ \Lambda A_{11}\Lambda^{-1} & \Lambda A_{12} \\ C_1\Lambda^{-1} & C_2 \end{bmatrix}$$

are contractive.

Proof. Since $P = \begin{bmatrix} \Lambda^2 & 0 \\ 0 & I \end{bmatrix}$ satisfies the inequality (21.27),

$$\begin{bmatrix} A & B \end{bmatrix} \begin{bmatrix} P & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A^* \\ B^* \end{bmatrix} \leq P$$

i.e, $\begin{bmatrix} P^{-1/2}AP^{1/2} & P^{-1/2}B \end{bmatrix}$ is a contraction. But

$$\begin{bmatrix} P^{-1/2}AP^{1/2} & P^{-1/2}B \end{bmatrix} = \begin{bmatrix} \Lambda^{-1}A_{12} & \Lambda^{-1}A_{11}\Lambda & \Lambda^{-1}B_1 \\ A_{22}\Lambda & A_{21} & B_2 \end{bmatrix} \begin{bmatrix} 0 & I & 0 \\ I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}.$$

Hence

$$\begin{bmatrix} \Lambda^{-1}A_{12} & \Lambda^{-1}A_{11}\Lambda & \Lambda^{-1}B_1 \\ A_{22}\Lambda & A_{21} & B_2 \end{bmatrix}$$

is also a contraction. The other part follows by a similar argument. □

Lemma 21.28 *Suppose that* $X = \begin{bmatrix} X_{11} & X_{12} \\ Z & X_{22} \end{bmatrix}$ *and* $Y = \begin{bmatrix} Y_{11} & Z \\ Y_{21} & Y_{22} \end{bmatrix}$ *are contractive (strictly contractive). Then*

$$M \triangleq \begin{bmatrix} 0 & \frac{1}{\sqrt{2}}X_{11} & X_{12} \\ \frac{1}{\sqrt{2}}Y_{11} & Z & \frac{1}{\sqrt{2}}X_{22} \\ Y_{21} & \frac{1}{\sqrt{2}}Y_{22} & 0 \end{bmatrix}$$

is also contractive (strictly contractive).

Proof. Dilate M to the following matrix:

$$M_d \triangleq \left[\begin{array}{ccc:c} 0 & \frac{1}{\sqrt{2}}X_{11} & X_{12} & \frac{1}{\sqrt{2}}X_{11} \\ \frac{1}{\sqrt{2}}Y_{11} & Z & \frac{1}{\sqrt{2}}X_{22} & 0 \\ Y_{21} & \frac{1}{\sqrt{2}}Y_{22} & 0 & -\frac{1}{\sqrt{2}}Y_{22} \\ \hdashline \frac{1}{\sqrt{2}}Y_{11} & 0 & -\frac{1}{\sqrt{2}}X_{22} & -Z \end{array}\right].$$

Considering X and Y are contractive, we can easily verify that $M_d^* M_d \leq I$, i.e, M_d is a contraction. □

We can now prove the theorem.

Proof of Theorem 21.26. Note that

$$G_r = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right] = \left[\begin{array}{cc|c} A_{11} & 0 & B_1 \\ 0 & 0 & 0 \\ \hline C_1 & 0 & D \end{array}\right].$$

Hence

$$\frac{1}{2}(G - G_r) = \left[\begin{array}{cccc|c} A_{11} & 0 & 0 & 0 & \frac{1}{\sqrt{2}}B_1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & A_{11} & A_{12} & \frac{1}{\sqrt{2}}B_1 \\ 0 & 0 & A_{21} & A_{22} & \frac{1}{\sqrt{2}}B_2 \\ \hline -\frac{1}{\sqrt{2}}C_1 & 0 & \frac{1}{\sqrt{2}}C_1 & \frac{1}{\sqrt{2}}C_2 & 0 \end{array}\right].$$

Now apply the similarity transformation

$$T = \begin{bmatrix} -\Lambda & 0 & \Lambda & 0 \\ 0 & -I & 0 & I \\ \Lambda^{-1} & 0 & \Lambda^{-1} & 0 \\ 0 & I & 0 & I \end{bmatrix}$$

to the realization of $\frac{1}{2}(G - G_r)$ to get

$$\frac{1}{2}(G - G_r) = \left[\begin{array}{cccc|c} \Lambda A_{11}\Lambda^{-1} & \frac{1}{2}\Lambda A_{12} & 0 & \frac{1}{2}\Lambda A_{12} & 0 \\ \frac{1}{2}A_{21}\Lambda^{-1} & \frac{1}{2}A_{22} & \frac{1}{2}A_{21}\Lambda & \frac{1}{2}A_{22} & \frac{1}{2}B_2 \\ 0 & \frac{1}{2}\Lambda^{-1}A_{12} & \Lambda^{-1}A_{11}\Lambda & \frac{1}{2}\Lambda^{-1}A_{12} & \Lambda^{-1}B_1 \\ \frac{1}{2}A_{21}\Lambda^{-1} & \frac{1}{2}A_{22} & \frac{1}{2}A_{21}\Lambda & \frac{1}{2}A_{22} & \frac{1}{2}B_2 \\ \hline C_1\Lambda^{-1} & \frac{1}{2}C_2 & 0 & \frac{1}{2}C_2 & 0 \end{array}\right].$$

It is easy to verify that *as a constant matrix* the right hand side of the above realization for $\frac{1}{2}(G - G_r)$ can be written as

$$\begin{bmatrix} 0 & 0 & I & 0 \\ 0 & \frac{1}{\sqrt{2}}I & 0 & 0 \\ I & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}I & 0 & 0 \\ 0 & 0 & 0 & I \end{bmatrix} \hat{M}_d \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\sqrt{2}}I & 0 & \frac{1}{\sqrt{2}}I & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix} \tag{21.33}$$

where

$$\hat{M}_d \triangleq \begin{bmatrix} 0 & \frac{1}{\sqrt{2}}\Lambda^{-1}A_{12} & \Lambda^{-1}A_{11}\Lambda & \Lambda^{-1}B_1 \\ \frac{1}{\sqrt{2}}A_{21}\Lambda^{-1} & A_{22} & \frac{1}{\sqrt{2}}A_{21}\Lambda & \frac{1}{\sqrt{2}}B_2 \\ \Lambda A_{11}\Lambda^{-1} & \frac{1}{\sqrt{2}}\Lambda A_{12} & 0 & 0 \\ C_1\Lambda^{-1} & \frac{1}{\sqrt{2}}C_2 & 0 & 0 \end{bmatrix}. \tag{21.34}$$

According to Theorem 21.12 (a) and (g), the theorem follows if we can show that the realization for $\frac{1}{2}(G - G_r)$ as a constant matrix is a contraction. However, this is guaranteed if $\hat{M}_d$ is a contraction since both the right and left hand matrices in (21.33) are contractive.

Finally, the contractiveness of $\hat{M}_d$ follows from Lemmas 21.27 and 21.28 by identifying

$$Z = A_{22},\ X_{11} = \frac{1}{\sqrt{2}}\Lambda^{-1}A_{12},\ Y_{11} = \frac{1}{\sqrt{2}}A_{21}\Lambda^{-1}$$

and

$$\begin{bmatrix} X_{12} \\ X_{22} \end{bmatrix} = \begin{bmatrix} \Lambda^{-1}A_{11}\Lambda & \Lambda^{-1}B_1 \\ \frac{1}{\sqrt{2}}A_{21}\Lambda & \frac{1}{\sqrt{2}}B_2 \end{bmatrix} \quad \begin{bmatrix} Y_{21} & Y_{22} \end{bmatrix} = \begin{bmatrix} \Lambda A_{11}\Lambda^{-1} & \frac{1}{\sqrt{2}}\Lambda A_{12} \\ C_1\Lambda^{-1} & \frac{1}{\sqrt{2}}C_2 \end{bmatrix}.$$

□

21.7 Model Reduction Using Coprime Factors

In this section, we consider lower order controller design using coprime factor reduction. We shall only consider the special case where the normalized right coprime factors are used.

Suppose that a dynamic system is given by

$$G = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

and assume that the realization is stabilizable and detectable. Recall from Theorem 21.21 that there exists a normalized right coprime factorization $G = NM^{-1}$ such that $\begin{bmatrix} M \\ N \end{bmatrix}$ is inner.

Lemma 21.29 *Let* $\nu_i = \sqrt{\lambda_i(YX)}$. *Then the Hankel singular values of* $\begin{bmatrix} M \\ N \end{bmatrix}$ *are given by*

$$\sigma_i = \frac{\nu_i}{\sqrt{1+\nu_i^2}} < 1.$$

Proof. This is obvious since

$$\sigma_i^2 = \lambda_i(PQ) = \frac{\lambda_i(YX)}{1+\lambda_i(YX)}$$

and $\lambda_i(YX) \geq 0$. □

It is known that there exists a transformation such that X and Y are balanced:

$$X = Y = \Pi = \begin{bmatrix} \Pi_1 & 0 \\ 0 & \Pi_2 \end{bmatrix}$$

with $\Pi_1 = \operatorname{diag}[\nu_1 I_{s_1}, \ldots, \nu_r I_{s_r}] > 0$.

Now partitioning the system G and matrix F accordingly,

$$G = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right] \qquad F = \begin{bmatrix} F_1 & F_2 \end{bmatrix}.$$

Then the reduced coprime factors

$$\begin{bmatrix} \hat{M} \\ \hat{N} \end{bmatrix} := \left[\begin{array}{c|c} A_{11}+B_1F_1 & B_1Z^{-1/2} \\ \hline F_1 & Z^{-1/2} \\ C_1+DF_1 & DZ^{-1/2} \end{array}\right] \in \mathcal{RH}_\infty$$

satisfy the following error bound.

Lemma 21.30 $$\left\| \begin{bmatrix} M \\ N \end{bmatrix} - \begin{bmatrix} \hat{M} \\ \hat{N} \end{bmatrix} \right\|_\infty \leq 2 \sum_{i \geq r+1} \sigma_i = 2 \sum_{i \geq r+1} \frac{\nu_i}{\sqrt{1+\nu_i^2}}.$$

Proof. Analogous to the continuous time case. □

Remark 21.8 It should be understood that the reduced model can be obtained by directly computing X and P and by obtaining a balanced model without solving the Riccati equation for Y. ♡

This reduced coprime factors combined with the robust or $\mathcal{H}_\infty$ controller design methods can be used to design lower order controllers so that the system is robustly stable and some specified performance criteria are satisfied. We will leave the readers to explore the utility of this model reduction method. However, we would like to point out that unlike the continuous time case, the reduced coprime factors in discrete time *may not* be normalized. In fact, we can prove a more general result.

Lemma 21.31 *Let a realization for $N(z) \in \mathcal{RH}_\infty$ be given by*

$$N(z) = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right]$$

with A stable. Suppose that there exists an $X = X^* \geq 0$ such that

$$\begin{bmatrix} A^*XA - X + C^*C & A^*XB + C^*D \\ B^*XA + D^*C & B^*XB + D^*D - I \end{bmatrix} = 0. \tag{21.35}$$

Then N is an inner, i.e., $N^\sim N = I$. Moreover, if the realization for

$$N = \left[\begin{array}{c|c} A & B \\ \hline C & D \end{array}\right] = \left[\begin{array}{cc|c} A_{11} & A_{12} & B_1 \\ A_{21} & A_{22} & B_2 \\ \hline C_1 & C_2 & D \end{array}\right]$$

is also balanced with

$$X = \Sigma = \begin{bmatrix} \Sigma_1 & 0 \\ 0 & \Sigma_2 \end{bmatrix}$$

and

$$A\Sigma A^* - \Sigma + BB^* = 0$$

with $\Sigma_1 > 0$, then the truncated system

$$N_r = \left[\begin{array}{c|c} A_{11} & B_1 \\ \hline C_1 & D \end{array}\right]$$

is stable and contractive, i.e., $N_r^\sim N_r \leq I$.

Proof. Pre-multiply equation (21.35) by

$$U = \begin{bmatrix} \begin{bmatrix} I & 0 \end{bmatrix} & 0 \\ 0 & I \end{bmatrix}$$

and post-multiply equation (21.35) by U^* to get

$$\begin{bmatrix} A_{11}^*\Sigma_1 A_{11} - \Sigma_1 + C_1^*C_1 + A_{21}^*\Sigma_2 A_{21} & A_{11}^*\Sigma_1 B_1 + C_1^*D + A_{21}^*\Sigma_2 B_2 \\ B_1^*\Sigma_1 A_{11} + D^*C_1 + B_2^*\Sigma_2 A_{21} & B_1^*\Sigma_1 B_1 + D^*D - I + B_2^*\Sigma_2 B_2 \end{bmatrix} = 0$$

or

$$\begin{bmatrix} A_{11}^*\Sigma_1 A_{11} - \Sigma_1 + C_1^*C_1 & A_{11}^*\Sigma_1 B_1 + C_1^*D \\ B_1^*\Sigma_1 A_{11} + D^*C_1 & B_1^*\Sigma_1 B_1 + D^*D - I \end{bmatrix} = -\begin{bmatrix} A_{21}^* \\ B_2^* \end{bmatrix} \Sigma_2 \begin{bmatrix} A_{21}^* \\ B_2^* \end{bmatrix}^*.$$

This gives

$$\begin{bmatrix} A_{11} & B_1 \\ C_1 & D \end{bmatrix}^* \begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} A_{11} & B_1 \\ C_1 & D \end{bmatrix}^* - \begin{bmatrix} \Sigma_1 & 0 \\ 0 & I \end{bmatrix} \leq 0.$$

Now let $T = \Sigma_1^{1/2}$, and then

$$\left\| \begin{bmatrix} TA_{11}T^{-1} & TB_1 \\ C_1T^{-1} & D \end{bmatrix} \right\| \leq 1$$

i.e., $\|N_r\|_\infty \leq 1$. It is clear that N_r is inner if $\begin{bmatrix} A_{21}^* \\ B_2^* \end{bmatrix} \Sigma_2 = 0$ (although the converse may not be true). □

21.8 Notes and References

The results for the discrete Riccati equation are based mostly on the work of Kucera [1972] and Molinari [1975]. Numerical algorithms for solving discrete ARE with singular A matrix can be found in Arnold and Laub [1984], Dooren [1981], and references therein. The matrix factorizations are obtained by Chu [1988]. The normalized coprime factorizations are obtained by Meyer [1990] and Walker[1990]. The detailed treatment of discrete time $\mathcal{H}_\infty$ control can be found in Stoorvogel [1990], Limebeer, Green, and Walker [1989], and Iglesias and Glover [1991].

Bibliography

[1] Adamjan, V.M., D.Z. Arov, and M.G. Krein (1971). "Analytic properties of Schmidt pairs for a Hankel operator and the generalized Schur-Takagi problem," *Math. USSR Sbornik*, Vol. 15, pp. 31-73.

[2] Adamjan, V.M., D.Z. Arov, and M.G. Krein (1978). "Infinite block Hankel matrices and related extension problems," *AMS Transl.*, vol. 111, pp. 133-156.

[3] Al-Saggaf, U. M. and G. F. Franklin (1987). "An error bound for a discrete reduced order model of a linear multivariable system," *IEEE Trans. Automat. Contr.*, Vol.AC-32, pp. 815-819.

[4] Al-Saggaf, U. M. and G. F. Franklin (1988). "Model reduction via balanced realizations: an extension and frequency weighting techniques," *IEEE Trans. Automat. Contr.*, Vol.AC-33, No. 7, pp. 687-692.

[5] Anderson, B. D. O. (1967). "A system theory criterion for positive real matrices," *SIAM J. Contr. Optimiz.*, Vol. 6, No. 2, pp. 171-192.

[6] Anderson, B.D.O. (1967). "An algebraic solution to the spectral factorization problem," *IEEE Trans. Auto. Control*, Vol. AC-12, pp. 410-414.

[7] Anderson, B. D. O. (1993). "Bode Prize Lecture (Control design: moving from theory to practice)," *IEEE Control Systems,* Vol. 13, No.4, pp. 16-25.

[8] Anderson, B.D.O., P.Agathoklis, E.I.Jury and M.Mansour (1986). "Stability and the Matrix Lyapunov Equation for Discrete 2-Dimensional Systems," *IEEE Trans.*, Vol.CAS-33, pp.$261 - 266$.

[9] Anderson, B. D. O. and Y. Liu (1989). "Controller reduction: concepts and approaches, " *IEEE Trans. Automat. Contr.* , Vol. AC-34, No. 8, pp. 802-812.

[10] Anderson, B. D. O. and J. B. Moore (1979). *Optimal Filtering.* Prentice-Hall, Englewood Cliffs, New Jersey.

[11] Anderson, B. D. O. and J. B. Moore (1989). *Optimal Control: Linear Quadratic Methods.* Prentice-Hall, Englewood Cliffs, New Jersey.

[12] Anderson, B. D. O. and S. Vongpanitlerd (1973). *Network Analysis and Synthesis: A Modern System Theory Approach.* Prentice-Hall, Englewood-Cliffs, New Jersey.

[13] Arnold, W. F. and A. J. Laub (1984). "Generalized Eigenproblem algorithms and software for algebraic Riccati equations," *Proceedings of the IEEE*, Vol. 72, No. 12, pp. 1746-1754.

[14] Balas, G. (1990). *Robust Control of Flexible Structures: Theory and Experiments*, PhD. Thesis, Caltech.

[15] Balas, G., J. C. Doyle, K. Glover, A. Packard, and R. Smith (1991). *μ-Analysis and Synthesis Toolbox*, MUSYN Inc. and The MathWorks, Inc.

[16] Ball, J.A. and N. Cohen (1987). "Sensitivity minimization in an $\mathcal{H}_\infty$ norm: parameterization of all suboptimal solutions," *Int. J. Control*, vol. 46, pp. 785-816.

[17] Ball, J.A. and J.W. Helton (1983). "A Beurling-Lax theorem for the Lie group U(m,n) which contains most classical interpolation theory," *J. Op. Theory*, vol. 9, pp. 107-142.

[18] Başar, T. and P. Bernhard (1991). *$\mathcal{H}_\infty$-Optimal Control and Related Minimax Design Problems: A Dynamic Game Approach.* Systems and Control: Foundations and Applications. Birkhäuser, Boston.

[19] Başar, T. and G. J. Olsder (1982). *Dynamic Noncooperative Game Theory*, Academic Press.

[20] Bernstein, D.S. and W.M. Haddard (1989). "LQG control with an $\mathcal{H}_\infty$ performance bound: A Riccati equation approach," *IEEE Trans. Auto. Control*, vol. AC-34, pp. 293-305.

[21] Bettayeb, M., L. M. Silverman, and M. G. Safonov (1980). "Optimal approximation of continuous-time systems." *Proc. IEEE Conf. Dec. Contr.*, Albuquerque, pp. 195-198.

[22] Bode, H. W. (1945). *Network Analysis and Feedback Amplifier Design*, D. Van Nostrand, Princeton.

[23] Boyd, S.P., V. Balakrishnan, C. H. Barratt, N. M. Khraishi, X. Li, D. G. Meyer, S. A. Norman (1988). "A New CAD Method and Associated Architectures for Linear Controllers," *IEEE Trans. Auto. Contr.*, Vol.AC-33, pp.$268-283$.

[24] Boyd, S., V. Balakrishnan, and P. Kabamba (1989). "A bisection method for computing the $\mathcal{H}_\infty$ norm of a transfer matrix and related problems," *Math. Control, Signals, and Systems*, vol. 2, No. 3, pp. 207-220.

[25] Boyd, S. and C. Barratt (1991). *Linear Controller Design – Limits of Performance*, Prentice Hall, Inc.

[26] Boyd, S. and C.A. Desoer (1985). "Subharmonic functions and performance bounds in linear time-invariant feedback systems," *IMA J. Math. Contr. and Info.*, vol. 2, pp. 153-170.

[27] Boyd, S. and Q. Yang (1989). "Structured and simultaneous Lyapunov functions for system stability problems," *Int. J. Control*, vol. 49, No. 6, pp. 2215-2240.

[28] Brasch, F. M. and J. B. Pearson (1970). "Pole placement using dynamical compensators," *IEEE Trans. Auto. Contr.*, Vol. 15, pp. 34-43.

[29] Brogan, W. L. (1991). *Modern Control Theory, 3rd Ed.*, Prentice Hall, Englewood Cliffs, New Jersey.

[30] Bruinsma, N. A., and M. Steinbuch (1990). "A fast algorithm to compute the $\mathcal{H}_\infty$-norm of a transfer function matrix", *Systems and Control letters*, Vol. 14, pp. 287-293.

[31] Bryson Jr., A. E. and Y-C. Ho (1975). *Applied Optimal Control*, Hemisphere Publishing Corporation.

[32] Chen, C. T. (1984). *Linear System Theory and Design.* Holt, Rinehart and Winston.

[33] Chen, J. (1992a). "Sensitivity integral relations and design tradeoffs in linear multivariable feedback systems," submitted to *IEEE Trans. Auto. Contr.*.

[34] Chen, J. (1992b). "Multivariable gain-phase and sensitivity integral relations and design tradeoffs," submitted to *IEEE Trans. Auto. Contr.*.

[35] Chen, J. (1995). Personal communication.

[36] Chen, M. J. and C.A. Desoer (1982). "Necessary and sufficient condition for robust stability of linear distributed feedback systems," *Int. J. Control*, vol. 35, no. 2, pp. 255-267.

[37] Chen, T. and B. A. Francis (1992). " $\mathcal{H}_2$-optimal sampled-data control," *IEEE Trans. Auto. Contr.*, Vol. 36, No. 4, pp. 387-397.

[38] Chu, C. C. (1985). *$\mathcal{H}_\infty$ optimization and robust multivariable control.* PhD Thesis, University of Minnesota.

[39] Chu, C. C. (1988). "On discrete inner-outer and spectral factorization," *Proc. Amer. Contr. Conf.*.

[40] Clements, D. J. (1993). "Rational spectral factorization using state-space methods," *Systems and Control Letters*, Vol. 20, pp. 335-343.

[41] Clements, D. J. and K. Glover (1989). "Spectral factorization by Hermitian pencils," *Linear Algebra and its Applications*, Vol. 122-124, pp. 797-846.

[42] Copple, W. A. (1974). "Matrix quadratic equations," *Bull. Austral. Math. Soc.*, Vol. 10, pp. 377 $\sim$ 401.

[43] Dahleh, M. A. and J. B. Pearson (1987). "ℓ_1 optimal feedback controllers for MIMO discrete time systems," *IEEE Trans. Auto. Contr.*, Vol. AC-32, No. 4, pp. 314 $\sim$ 322.

[44] Daniel, R. W., B. Kouvaritakis, and H. Latchman (1986). "Principal direction alignment: a geometric framework for the complete solution to the μ problem," *IEE Proceedings*, vol. 133, Part D, No. 2, pp. 45-56.

[45] Davis, C., W. M. Kahan, and H. F. Weinberger (1982). "Norm-preserving dilations and their applications to optimal error bounds," *SIAM J. Numer. Anal.,* vol 19, pp.445 $\sim$ 469.

[46] Davison, E. J. (1968). "On pole assignment in multivariable linear systems," *IEEE Trans. Auto. Contr.*, Vol. 13, No. 6, pp. 747-748.

[47] Delsarte, P., Y. Genin, and Y. Kamp (1979). "The Navanlinna-Pick problem for matrix valued functions," *SIAM J. Applied Math.*, Vol. 36, pp. 47-61.

[48] Desai, U. B. and D. Pal (1984). "A transformation approach to stochastic model reduction," *IEEE Trans. Automat. Contr.,* Vol 29, NO. 12, pp. 1097-1100.

[49] Desoer, C. A., R. W. Liu, J. Murray, and R. Saeks (1980). "Feedback system design: the fractional representation approach to analysis and synthesis," *IEEE Trans. Auto. Contr.*, Vol. AC-25, No. 6, pp. 399 $\sim$ 412.

[50] Desoer, C. A. and M. Vidyasagar (1975). *Feedback Systems: Input-Output Properties. Academic Press*, New York.

[51] Dooren, P. V. (1981). "A generalized eigenvalue approach for solving Riccati equations," *SIAM J. Sci. Stat. Comput.*, Vol. 2, No. 2, pp. 121-135.

[52] Doyle, J. C. (1978). "Guaranteed Margins for LQG Regulators," *IEEE Trans. Auto. Contr.*, Vol. AC-23, No. 4, pp. 756 $\sim$ 757.

[53] Doyle, J.C. (1982). "Analysis of feedback systems with structured uncertainties," *IEE Proceedings*, Part D, Vol.133, pp.$45-56$.

[54] Doyle, J. C. (1985). "Structured uncertainty in control system design", IEEE CDC, Ft. Lauderdale.

[55] Doyle, J. C. (1984). "Lecture notes in advances in multivariable control," *ONR/Honeywell Workshop*, Minneapolis.

[56] Doyle, J. C., B. Francis, and A. Tannenbaum (1992). *Feedback Control Theory*, Macmillan Publishing Company.

[57] Doyle, J.C., K. Glover, P.P. Khargonekar, B.A. Francis (1989). " State-space solutions to standard $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control problems," *IEEE Trans. Auto. Control*, vol. AC-34, No. 8, pp. 831-847. Also see 1988 American Control Conference, Atlanta, June, 1988.

[58] Doyle, J., K. Lenz, and A. Packard (1986). "Design Examples using μ synthesis: Space Shuttle Lateral Axis FCS during reentry," *IEEE CDC*, December 1986, pp. 2218-2223.

[59] Doyle, J.C., A.Packard and K. Zhou (1991). "Review of LFTs, LMIs and μ," *Proc. 30th IEEE Conf. Dec. Contr.*, England, pp. 1227-1232

[60] Doyle, J.C., R.S. Smith, and D.F. Enns (1987). "Control of Plants with Input Saturation Nonlinearities," *American Control Conf.*, pp 1034-1039.

[61] Doyle, J. C. and G. Stein (1981). "Multivariable feedback design: Concepts for a classical/modern synthesis," *IEEE Trans. Auto. Control*, vol. AC-26, pp. 4-16, Feb. 1981.

[62] Doyle, J. C., J. Wall and G. Stein (1982). "Performance and robustness analysis for structured uncertainty," in *Proc. 21st IEEE Conf. Decision Contr.*, pp. 629-636.

[63] Doyle, J. C., K. Zhou, and B. Bodenheimer (1989). "Optimal Control with mixed $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance objective," Proc. of American Control Conference, Pittsburgh, Pennsylvania, pp. 2065-2070.

[64] Doyle, J. C., K. Zhou, K. Glover, and B. Bodenheimer (1994). "Mixed $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance objectives II: optimal control," *IEEE Trans. on Automat. Contr.*, vol. 39, no. 8, pp. 1575-1587.

[65] El-Sakkary, A. (1985). "The gap metric: Robustness of stabilization of feedback systems," *IEEE Trans. Automat. Contr.*, Vol. 30, pp. 240-247.

[66] Enns, D. (1984a). *Model Reduction for Control System Design*, Ph.D. Dissertation, Department of Aeronautics and Astronautics, Stanford University, Stanford, California.

[67] Enns, D. (1984b) "Model reduction with balanced realizations: An error bound and a frequency weighted generalization," *Proc. 23rd Conf. Dec. Contr.* , Las Vegas, NV.

[68] Fan, M. K. H. and A.L. Tits (1986). "Characterization and efficient computation of the structured singular value," *IEEE Trans. Auto. Control*, vol. AC-31, no. 8, pp. 734-743.

[69] Fan, M. K. H. and A.L. Tits, and J. C. Doyle (1991). "Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics," *IEEE Trans. Auto. Control*, vol. AC-36, no. 1, pp. 25-38.

[70] Foias, C. and A. Tannenbaum (1988). "On the four-block problem, I," *Operator Theory: Advances and Applications*, vol. 32 (1988), pp. 93-112 and "On the four block problem, II: the singular system," *Integral Equations and Operator Theory*, vol. 11 (1988), pp. 726-767.

[71] Francis, B. A. (1987). *A course in $\mathcal{H}_\infty$ control theory*, Lecture Notes in Control and Information Sciences, vol. 88.

[72] Francis, B.A. and J.C. Doyle (1987). "Linear control theory with an $\mathcal{H}_\infty$ optimality criterion," *SIAM J. Control Opt.*, vol. 25, pp. 815-844.

[73] Freudenberg, J.S. (1989). "Analysis and design for ill-conditioned plants, part 2: directionally uniform weightings and an example," *Int. J. Contr.*, vol. 49, no. 3, pp. 873-903.

[74] Freudenberg, J.S. and D.P. Looze (1985). "Right half plane zeros and poles and design tradeoffs in feedback systems," *IEEE Trans. Auto. Contr.*, vol. AC-30, no. 6, pp. 555-565.

[75] Freudenberg, J.S. and D.P. Looze (1988). *Frequency Domain Properties of Scalar and Multivariable Feedback Systems*, Lecture Notes in Contr. and Info. Science, vol. 104, Berlin: Springer-Verlag, 1988.

[76] Freudenberg, J. S., D.P. Looze and J.B. Cruz (1982). "Robustness analysis using singular value sensitivities", *Int. J. Control*, vol. 35, no. 1, pp. 95-116.

[77] Garnett, J. B. (1981). *Bounded Analytic Functions*, Academic Press.

[78] Georgiou, T. T. (1988). "On the computation of the gap metric," *Syst. Contr. Lett.*, Vol. 11, pp. 253-257.

[79] Georgiou, T. T. and M. C. Smith (1990). "Optimal robustness in the gap metric," *IEEE Trans. Automat. Contr.*, Vol.AC-35, No. 6, pp. 673-686.

[80] Glover, K. (1984). "All optimal Hankel-norm approximations of linear multivariable systems and their $\mathcal{L}_\infty$-error bounds," *Int. J. Control*, vol. 39, pp. 1115-1193, 1984.

[81] Glover, K. (1986). "Robust stabilization of linear multivariable systems: Relations to approximation," *Int. J. Contr.*, Vol. 43, No. 3, pp. 741-766.

[82] Glover, K. (1989). "A tutorial on Hankel-norm Approximation," in *Data to Model*, J. C. Willems, Ed. New York: Springer-Verlag.

[83] Glover, K. (1986). "Multiplicative approximation of linear multivariable systems with L_∞ error bounds, " *Proc. Amer. Contr. Conf.*, Seattle, WA, pp. 1705-1709.

[84] Glover, K. and J. Doyle (1988). "State-space formulae for all stabilizing controllers that satisfy an $\mathcal{H}_\infty$ norm bound and relations to risk sensitivity," *Systems and Control Letters*, vol. 11, pp. 167-172.

[85] Glover, K. and J. C. Doyle (1989). "A state space approach to $\mathcal{H}_\infty$ optimal control," in *Three Decades of Mathematical Systems Theory: A Collection of Surveys at the Occasion of the 50th Birthday of Jan C. Willems*, H. Nijmeijer and J. M. Schumacher (Eds.), Springer-Verlag, Lecture Notes in Control and Information Sciences, vol. 135, 1989.

[86] Glover, K., D.J.N. Limebeer, J.C. Doyle, E.M. Kasenally, and M.G. Safonov (1991). "A characterization of all solutions to the four block general distance problem," *SIAM J. Optimization and Control*, Vol. 29, No. 2, pp. 283 $\sim$ 324.

[87] Glover, K., D. J. N. Limebeer, and Y. S. Hung (1992). "A structured approximation problem with applications to frequency weighted model reduction, " *IEEE Trans. Automat. Contr.*, vol.AC-37, No. 4, pp. 447-465.

[88] Glover, K. and D. McFarlane (1989). "Robust stabilization of normalized coprime factor plant descriptions with $\mathcal{H}_\infty$ bounded uncertainty," *IEEE Trans. Automat. Contr.*, vol. 34, no. 8, pp.821-830.

[89] Glover, K. and D. Mustafa (1989). "Derivation of the maximum entropy $\mathcal{H}_\infty$ controller and a state space formula for its entropy," *Int. J. Control,* vol. 50, pp. 899.

[90] Goddard, P. J. and K. Glover (1993). "Controller reduction: weights for stability and performance preservation," *Proceedings of the 32nd Conference on Decision and Control*, San Antonio, Texas, December 1993, pp. 2903-2908.

[91] Goddard, P. J. and K. Glover (1994). "Performance preserving frequency weighted controller approximation: a coprime factorization approach," *Proceedings of the 33nd Conference on Decision and Control*, Orlando, Florida, December 1994.

[92] Gohberg, I., and S. Goldberg (1981). *Basic Operator Theory.* Birkhauser, Boston.

[93] Gohberg, I., P. Lancaster, and L. Rodman (1986). "On hermitian solutions of the symmetric algebraic Riccati equation," *SIAM J. Control and Optimization*, Vol. 24, No. 6, November 1986, pp. 1323-1334.

[94] Golub, G. H., and C. F. Van Loan (1983). *Matrix Computations.* Baltimore, Md.: Johns Hopkind Univ. Press.

[95] Green, M. (1988a). "A relative error bound for balanced stochastic truncation," *IEEE Trans. Automat. Contr.*, Vol.AC-33, No. 10, pp. 961-965.

[96] Green, M (1988b). "Balanced stochastic realizations," *Journal of Linear Algebra and its Applications,* Vol. 98, pp. 211-247.

[97] Green, M., K. Glover, D.J.N. Limebeer and J.C. Doyle (1988), "A J-spectral factorization approach to H_∞ control," *SIAM J. Control and Optimization*, Vol. 28, pp. 1350-1371.

[98] Hautus, M. L. J. and L. M. Silverman (1983). "System structure and singular control," *Linear Algebra and its Applications*, Vol. 50, pp. 369-402.

[99] Helton, J. W. (1981). "Broadband gain equalization directly from data," *IEEE Trans. Circuits Syst.*, Vol. CAS-28, pp. 1125-1137.

[100] Hestenes, M. (1975). *Optimization Theory: the finite dimensional case*, John Wiley & Sons.

[101] Heymann, M. (1968). "Comments "On pole assignment in multi-input controllable linear systems," " *IEEE Trans. Auto. Contr.*, Vol. 13, No. 6, pp. 748-749.

[102] Hinrichsen, D., and A. J. Pritchard (1990). "An improved error estimate for reduced-order models of discrete time system," *IEEE Transactions on Automatic Control*, AC-35, pp. 317-320.

[103] Hoffman, K. (1962). *Banach Spaces of Analytic Functions*, Prentice-hall, Inc., Englewood Cliffs, N.J.

[104] Horn, R. A. and C. R. Johnson (1990). *Matrix Analysis*, Cambridge University Press, First Published in 1985.

[105] Horn, R. A. and C. R. Johnson (1991). *Topics in Matrix Analysis*, Cambridge University Press.

[106] Horowitz, I. M. (1963). *Synthesis of Feedback Systems*, Academic Press, London.

[107] Hung, Y. S. (1989)."$\mathcal{H}_\infty$ -optimal control- Part I: model matching, - Part II: solution for controllers," *Int. J. Control*, Vol. 49, pp. 1291-1359.

[108] Hung, Y. S. and K. Glover (1986). "Optimal Hankel-norm approximation of stable systems with first order stable weighting functions," *Syst. Contr. Lett.*, Vol. 7 pp. 165-172.

[109] Hung, Y. S. and A. G. J. MacFarlane (1982). *Multivariable feedback: A quasi-classical approach*, Springer-Verlag.

[110] Hyde, R. A. and K. Glover (1993). "The Application of Scheduled $\mathcal{H}_\infty$ Controllers to a VSTOL Aircraft," *IEEE Trans. Automatic Control*, Vol. 38, No. 7, pp. 1021-1039.

[111] Hyland, D. C. and D. S. Bernstein (1984). "The optimal projection equations for fixed-order dynamic compensation," *IEEE Trans. Auto. Contr.*, Vol. AC-29, N0. 11, pp. 1034 $\sim$ 1037.

[112] Iglesias, P. A. and K. Glover (1991). "State-space approach to discrete-time $\mathcal{H}_\infty$ control," *Int. J. Control*, Vol. 54, No. 5, pp. 1031-1073.

[113] Iglesias, P. A., D. Mustafa, and K. Glover (1990). "Discrete time $\mathcal{H}_\infty$ controllers satisfying minimum entropy criterion," *Syst. Contr. Lett.*, Vol. 14, pp. 275-286.

[114] Jonckheere, E.A. and J.C. Juang (1987). "Fast computation of achievable performance in mixed sensitivity $\mathcal{H}_\infty$ design," *IEEE Trans. on Auto. Control*, vol. AC-32, pp. 896-906.

[115] Jonckheere, E. A. and L. M. Silverman (1978). "Spectral theory of the linear quadratic optimal control problem: discrete-time single-input case," *IEEE Trans. Circuits. Syst.*, Vol. CAS-25, No. 9, pp. 810-825.

[116] Kailath, T. (1980). *Linear Systems.* Prentice-Hall, Inc., Englewood Cliffs, N.J. 07632.

[117] Kalman, R. E. (1964). "When is a control system optimal?" *ASME Trans. Series D: J. Basic Engineering*, Vol. 86, pp. 1-10.

[118] Kalman, R. E. and R. S. Bucy (1960). "New results in linear filtering and prediction theory," *ASME Trans. Series D: J. Basic Engineering*, Vol. 83, pp. 95-108.

[119] Kavranoğlu, D. and M. Bettayeb (1994). "Characterization and computation of the solution to the optimal $\mathcal{L}_\infty$ approximation problem," *IEEE Trans. Auto. Contr.*, Vol. 39, No. 9, pp. 1899-1904.

[120] Khargonekar, P.P., I.R. Petersen, and M.A. Rotea (1988). "$\mathcal{H}_\infty$- optimal control with state-feedback," *IEEE Trans. on Auto. Control*, vol. AC-33, pp. 786-788.

[121] Khargonekar, P.P., I.R. Petersen, and K. Zhou (1990). "Robust stabilization and $\mathcal{H}_\infty$-optimal control," *IEEE Trans. Auto. Contr.*, Vol. 35, No. 3, pp. 356 $\sim$ 361.

[122] Khargonekar, P. P. and M. A. Rotea (1991). "Mixed $\mathcal{H}_2\mathcal{H}_\infty$ control: a convex optimization approach," *IEEE Trans. Automat. Contr.*, vol. 36, no. 7, pp. 824-737.

[123] Khargonekar, P. P. and E. Sontag (1982). "On the relation between stable matrix fraction factorizations and regulable realizations of linear systems over rings," *IEEE Trans. Automat. Contr.*, Vol. 27, pp. 627-638.

[124] Kimura, H. (1988). "Synthesis of $\mathcal{H}_\infty$ Controllers based on Conjugation," *Proc. IEEE Conf. Decision and Control* , Austin, Texas.

[125] Kishore, A. P. and J. B. Pearson (1992). "Uniform stability and performance in $\mathcal{H}_\infty$," *Proc. IEEE Conf. Dec. Contr.*, Tucson, Arizona, pp. 1991-1996.

[126] Kleinman, D. L. (1968). "On an iterative technique for Riccati equation computation," *IEEE Transactions on Automatic Control*, Vol. 13, pp. 114-115.

[127] Kung, S. K., B. Levy, M. Morf and T. Kailath (1977). "New results in 2-D Systems Theory, 2-D State-Space Models – Realization and the Notions of Controllability, Observability and Minimality," *Proc. IEEE*, pp. 945 $\sim$ 961.

[128] Kung, S. K. and D. W. Lin (1981). "Optimal Hankel norm model reduction: Multivariable systems." *IEEE Trans. Automat. Contr.*, Vol. AC-26, pp. 832-852.

[129] Kwakernaak, H. (1986). "A polynomial approach to minimax frequency domain optimization of multivariable feedback systems," *Int. J. Control,* pp. 117-156.

[130] Kwakernaak, H. and R. Sivan (1972). *Linear Optimal Control Systems*, Wiley-Interscience, New York.

[131] Kucera, V. (1972). "A contribution to matrix quadratic equations," *IEEE Trans. Auto. Control*, AC-17, No. 3, 344-347.

[132] Kucera, V. (1979). *Discrete Linear Control*, John Wiley and Sons, New York.

[133] Lancaster, P. and M. Tismenetsky (1985). *The Theory of Matrices: with applications*, 2nd Ed., Academic Press.

[134] Latham, G. A. and B. D. O. Anderson (1986). "Frequency weighted optimal Hankel-norm approximation of stable transfer function," *Syst. Contr. Lett.*, Vol. 5 pp. 229-236.

[135] Laub, S. J. (1980). "Computation of 'balancing' transformations," *Proc. 1980 JACC*, Session FA8-E.

[136] Lenz, K. E., P. P. Khargonekar, and John C. Doyle (1987). "Controller order reduction with guaranteed stability and performance, " *American Control Conference*, pp. 1697-1698.

[137] Limebeer, D. J. N., B. D. O. Anderson, P. P. Khargonekar, and M. Green (1992). "A game theoretic approach to $\mathcal{H}_\infty$ control for time varying systems," *SIAM J. Control and Optimization*, Vol. 30, No. 2, pp. 262-283.

[138] Limebeer, D. J. N., M. Green, and D. Walker (1989). "Discrete time $\mathcal{H}_\infty$ control," *Proc. IEEE Conf. Dec. Contr.*, Tampa, Florida, pp. 392-396.

[139] Limebeer, D.J.N. and G.D. Halikias (1988). "A controller degree bound for $\mathcal{H}_\infty$-optimal control problems of the second kind," *SIAM J. Control Opt.*, vol. 26, no. 3, pp. 646-677.

[140] Limebeer, D.J.N. and Y.S. Hung (1987). "An analysis of pole-zero cancelations in $\mathcal{H}_\infty$-optimal control problems of the first kind," *SIAM J. Control Opt.*, vol.25, pp. 1457-1493.

[141] Liu, K. Z., T. Mita, and R. Kawtani (1990). "Parameterization of state feedback $\mathcal{H}_\infty$ controllers," *Int. J. Contr.*, Vol.51, No. 3, pp. 535-551.

[142] Liu, K. Z. and T. Mita (1992). "Generalized $\mathcal{H}_\infty$ control theory," *Proc. American Control Conference.*

[143] Liu, Y. and B. D. O. Anderson (1986). "Controller reduction via stable factorization and balancing," *Int. J. Control*, Vol. 44, pp. 507-531.

[144] Liu, Y. and B. D. O. Anderson (1989). "Singular perturbation approximation of balanced systems," *Int. J. Control*, Vol. 50, No. 4, pp. 1379-1405.

[145] Liu, Y. and B. D. O. Anderson (1990). "Frequency weighted controller reduction methods and loop transfer recovery," *Automatica*, Vol. 26, No. 3, pp. 487-497.

[146] Liu, Y., B. D. O. Anderson, and U. Ly (1990). "Coprime factorization controller reduction with Bezout identity induced frequency weighting," *Automatica*, Vol. 26, No. 2, pp. 233-249, 1990.

[147] Lu, W. M., K. Zhou, and J. C. Doyle (1991). "Stabilization of LFT systems," *Proc. 30th IEEE Conf. Dec. Contr.*, Brighton, England, pp. 1239-1244.

[148] Luenberger, D. G. (1969). *Optimization by Vector Space Methods*, John Wiley & Sons, Inc.

[149] Luenberger, D. G. (1971). "An introduction tp observers," *IEEE Trans. Automat. Contr.*, Vol. 16, No. 6, pp. 596-602.

[150] MacFarlane, A.G.J. and J.S. Hung (1983). "Analytic Properties of the singular values of a rational matrix," *Int. J. Contr.*, vol. 37, no. 2, pp. 221-234.

[151] Maciejowski, J. M. (1989). *Multivariable Feedback Design,* Wokingham: Addison-Wesley.

[152] Mageirou, E. F. and Y. C. Ho (1977). "Decentralized stabilization via game theoretic methods," *Automatica*, Vol. 13, pp. 393-399.

[153] Mariton, M. and R. Bertrand (1985). "A homotopy algorithm for solving coupled Riccati equations," *Optimal. Contr. Appl. Meth.*, vol. 6, pp. 351-357.

[154] Martensson, K. (1971). "On the matrix Riccati equation," *Information Sciences*, Vol. 3, pp. 17-49.

[155] McFarlane, D. C., K. Glover, and M. Vidyasagar (1990). "Reduced-order controller design using coprime factor model reduction," *IEEE Trans. Automat. Contr.*, Vol. 35, No. 3, pp. 369-373.

[156] McFarlane, D. C. and K. Glover (1990). *Robust Controller Design Using Normalized Coprime Factor Plant Descriptions*, vol. 138, Lecture Notes in Control and Information Sciences. Springer-Verlag.

[157] McFarlane, D. C. and K. Glover (1992). "A loop shaping design procedure using $\mathcal{H}_\infty$ synthesis," *IEEE Trans. Automat. Contr.*, Vol. 37, No. 6, pp. 759-769.

[158] Meyer, D. G. (1990). "The construction of special coprime factorizations in discrete time," *IEEE Trans. Automat. Contr.*, Vol. 35, No. 5, pp. 588-590.

[159] Meyer, D. G. (1990). "Fractional balanced reduction: model reduction via fractional representation," *IEEE Trans. Automat. Contr.*, Vol. 35, No. 12, pp. 1341-1345.

[160] Meyer, D. G. and G. Franklin (1987). "A connection between normalized coprime factorizations and linear quadratic regulator theory," *IEEE Trans. Automat. Contr.*, Vol. 32, pp. 227-228.

[161] Molinari, B. P. (1975). "The stabilizing solution of the discrete algebraic Riccati equation," *IEEE Trans. Auto. Contr.*, June 1975, pp. 396 $\sim$ 399.

[162] Moore, B. C. (1981). "Principal component analysis in linear systems: controllability, observability, and model reduction," *IEEE Trans. Auto. Contr.*, Vol. 26, No.2.

[163] Moore, J. B., K. Glover, and A. Telford (1990). "All stabilizing controllers as frequency shaped state estimate feedback," *IEEE Trans. Auto. Contr.*, Vol. AC-35, pp. 203 $\sim$ 208.

[164] Mullis, C. T. and R. A. Roberts (1976). "Synthesis of minimum roundoff noise fixed point digital filters," *IEEE Trans. Circuits and Systems*, Vol. 23, pp. 551-562.

[165] Mustafa, D. (1989). "Relations between maximum entropy / $\mathcal{H}_\infty$ control and combined $\mathcal{H}_\infty/LQG$ control," *Systems and Control Letters,* vol. 12, no. 3, p. 193.

[166] Mustafa, D. and K. Glover (1990). *Minimum Entropy $\mathcal{H}_\infty$ Control*, Lecture notes in control and information sciences, Springer-Verlag.

[167] Mustafa, D. and K. Glover (1991). "Controller Reduction by $\mathcal{H}_\infty$- Balanced Truncation", *IEEE Trans. Automat. Contr.*, vol.AC-36, No. 6, pp. 669-682.

[168] Naylor, A. W., and G. R. Sell (1982). *Linear Operator Theory in Engineering and Science*, Springer-Verlag.

[169] Nagpal, K. M. and P. P. Khargonekar (1991). "Filtering and smoothing in an $\mathcal{H}_\infty$ setting," *IEEE Trans. Automat. Contr.*, Vol.AC-36, No. 2, pp. 152-166.

[170] Nehari, Z. (1957). "On bounded bilinear forms," *Annals of Mathematics*, Vol. 15, No. 1, pp. 153-162.

[171] Nesterov, Y. and A. Nemirovski (1994). *Interior-point polynomial algorithms in convex programming*, SIAM.

[172] Nett, C.N., C.A.Jacobson and N.J.Balas (1984). "A Connection between State-Space and Doubly Coprime Fractional Representations," *IEEE Trans.*, Vol.AC-29, pp.831 − 832.

[173] Osborne, E.E. (1960). "On Preconditioning of Matrices," *J. Assoc. Comp. Mach.*, 7, pp. 338-345.

[174] Packard, A. (1991). *Notes on* μ_Δ. Unpublished lecture notes.

[175] Packard, A. and J. C. Doyle (1988). "Structured singular value with repeated scalar blocks," 1988 ACC, Atlanta.

[176] Packard, A. and J. C. Doyle (1988). *Robust Control of Multivariable and Large Scale Systems*, Final Technical Report for Air Force Office of Scientific Research.

[177] Packard, A. and J. C. Doyle (1993), "The complex structured singular value," *Automatica*, Vol. 29, pp. 71-109.

[178] Packard, A. and P. Pandey (1993). Continuity properties of the real/complex structured singular value," *IEEE Trans. Automat. Contr.*, Vol. 38, No. 3, pp. 415-428.

[179] Packard, A., K. Zhou, P. Pandey, J. Leonhardson, G. Balas (1992). "Optimal, Constant I/O similarity scaling for full information and state feedback control problems," *Syst. Contr. Lett.*, Vol. 19, No. 4, pp. 271-280.

[180] Parrott, S. (1978). "On a quotient norm and the Sz-Nagy Foias lifting theorem," *Journal of Functional Analysis*, vol. 30, pp. 311 $\sim$ 328.

[181] Pernebo, L. and L. M. Silverman (1982). "Model Reduction via Balanced State Space Representation," *IEEE Trans. Automat. Contr.*, vol.AC-27, No. 2, pp. 382-387.

[182] Petersen, I.R. (1987). "Disturbance attenuation and $\mathcal{H}_\infty$ optimization: a design method based on the algebraic Riccati equation," *IEEE Trans. Auto. Contr.*, vol. AC-32, pp. 427-429.

[183] Postlethwaite, I., J. M. Edmunds, and A. G. J. MacFarlane (1981). "Principal gains and Principal phases in the analysis of linear multivariable feedback systems," *IEEE Trans. Auto. Contr.*, Vol. AC-26, No. 1, pp. 32-46.

[184] Postlethwaite, I. and A. G. J. MacFarlane (1979). *A complex variable approach to the analysis of linear multivariable feedback systems.* Berlin: Springer.

[185] Power, S. C. (1982). *Hankel operators on Hilbert space.* Pitman Advanced Publishing Program.

[186] Ran, A. C. M. and R. Vreugdenhil (1988). "Existence and comparison theorems for algebraic Riccati equations for continuous- and discrete- time systems," *Linear Algebra and its Applications*, Vol. 99, pp. 63-83.

[187] Redheffer, R. M. (1959). "Inequalities for a matrix Riccati equation," *Journal of Mathematics and Mechanics*, vol. 8, no. 3.

[188] Redheffer, R.M. (1960). "On a certain linear fractional transformation," *J. Math. and Physics*, vol. 39, pp. 269-286.

[189] Richter, S.L. (1987). "A homotopy algorithm for solving the optimal projection equations for fixed-order dynamic compensation: existence, convergence and global optimality," American Control Conference, Minneapolis.

[190] Rockafellar, R. T. (1970). *Convex Analysis* . Princeton University Press, Princeton, New Jersey.

[191] Roesser, R.P. (1975). "A Discrete State Space Model for Linear Image Processing," *IEEE Trans. Automat. Contr.*, Vol.AC-20, No. 2, pp.1 $\sim$ 10.

[192] Rotea, M. A., and P. P. Khargonekar (1991). "$\mathcal{H}_2$ optimal control with an $\mathcal{H}_\infty$ constraint: the state feedback case," *Automatica*, Vol. 27, No. 2, pp. 307-316.

[193] Rudin, W. (1966). *Real and Complex Analysis.* McGraw Hill Book Company, New York.

[194] Qiu, L. and E. J. Davison (1992a). "Feedback stability under simultaneous gap metric uncertainties in plant and controller," *Syst. Contr. Lett.*, Vol. 18, pp. 9-22.

[195] Qiu, L. and E. J. Davison (1992b). "Pointwise gap metrics on transfer matrices," *IEEE Trans. Automat. Contr.*, Vol. 37, No. 6, pp. 770-780.

[196] Safonov, M.G. (1980). *Stability and Robustness of Multivariable Feedback Systems. MIT Press.*

[197] Safonov, M. G. (1982). "Stability margins of diagonally perturbed multivariable feedback systems," *Proc. IEE, Part D*, Vol. 129, No. 6, pp. 251-256.

[198] Safonov, M.G. (1984). "Stability of interconnected systems having slope-bounded nonlinearities," 6th Int. Conf. on Analysis and Optimization of Systems, Nice, France.

[199] Safonov, M.G. and J.C. Doyle (1984). "Minimizing conservativeness of robustness singular values", in *Multivariable Control*, S.G. Tzafestas, Ed. New York: Reidel.

[200] Safonov, M. G., A. J. Laub, and G. L. Hartmann (1981). "Feedback properties of multivariable systems: the role and use of the return difference matrix," *IEEE Trans. Auto. Contr.*, Vol. AC-26, No. 1, pp. 47 $\sim$ 65.

[201] Safonov, M.G. and D.J.N. Limebeer, and R. Y. Chiang (1990). "Simplifying the $\mathcal{H}_\infty$ Theory via Loop Shifting, matrix pencil and descriptor concepts," *Int. J. contr.*, Vol 50, pp. 2467 $\sim$ 2488.

[202] Sarason, D. (1967). "Generalized interpolation in $\mathcal{H}_\infty$," *Trans. AMS.*, vol. 127, pp. 179-203.

[203] Scherer, C. (1992). "$\mathcal{H}_\infty$ control for plants with zeros on the imaginary axis," *SIAM J. Contr. Optimiz.*, Vol. 30, No. 1, pp. 123-142.

[204] Scherer, C. (1992). "$\mathcal{H}_\infty$ optimization without assumptions on finite or infinite zeros," *SIAM J. Contr. Optimiz.*, Vol. 30, No. 1, pp. 143-166.

[205] Silverman, L. M. (1969). "Inversion of multivariable linear systems," *IEEE Trans. Automat. Contr.*, Vol. AC-14, pp. 270-276.

[206] Skelton, R. E. (1988). *Dynamic System Control*, John Willey & Sons.

[207] Stein, G. and J. C. Doyle (1991). "Beyond singular values and loop shapes," *AIAA Journal of Guidance and Control*, vol. 14, no.1.

[208] Stein, G., G. L. Hartmann, and R. C. Hendrick (1977). "Adaptive control laws for F-8c flight test," *IEEE Trans. Auto. Contr.*

[209] Stoorvogel, A. A. (1992). *The $\mathcal{H}_\infty$ Control Problem: A State Space Approach*, Prentice-Hall, Englewood Cliffs.

[210] Tadmor, G. (1990). "Worst-case design in the time domain: the maximum principle and the standard $\mathcal{H}_\infty$ problem," *Mathematics of Control, Systems and Signal Processing*, pp. 301-324.

[211] Tits, A. L. (1994). "The small μ theorem with non-rational uncertainty," preprint.

[212] Tits, A. L. and M. K. H. Fan (1994). "On the small μ theorem," preprint.

[213] Vidyasagar, M. (1984). "The graph metric for unstable plants and robustness estimates for feedback stability," *IEEE Trans. Automatic Control*, Vol. 29, pp.403-417.

[214] Vidyasagar, M. (1985). *Control System Synthesis: A Factorization Approach.* MIT Press, Cambridge, MA.

[215] Vidyadagar, M. (1988). "Normalized coprime factorizations for non-strictly proper systems," *IEEE Trans. Automatic Control*, Vol. 33, No. 3, pp. 300-301.

[216] Vidyadagar, M. and H. Kimura (1986). "Robust controllers for uncertain linear multivariable systems," *Automatica*, Vol. 22, pp. 85-94.

[217] Vinnicombe, G. (1993). "Frequency domain uncertainty and the graph topology," *IEEE Trans. Automatic Control*, Vol. 38, No. 9, pp. 1371-1383.

[218] Walker, D. J. (1990). "Robust stabilizability of discrete-time systems with normalized stable factor perturbation," *Int. J. Control*, Vol. 52, No.2, pp.441-455.

[219] Wall, J.E., Jr., J.C. Doyle, and C.A. Harvey (1980). "Tradeoffs in the design of multivariable feedback systems," *Proc. 18th Allerton Conf.*, pp. 715-725.

[220] Wang, W., J.C.Doyle and K.Glover (1991). "Stability and Model Reduction of Linear Fractional Transformation System," *1991 IEEE CDC*, England.

[221] Wang, W. and M. G. Safonov (1990). "A tighter relative-error bound for balanced stochastic truncation," *Syst. Contr. Lett.*, Vol 14, pp. 307-317.

[222] Wang, W. and M. G. Safonov (1992). "Multiplicative-error bound for balanced stochastic truncation model reduction," *IEEE Trans. Automat. Contr.*, vol.AC-37, No. 8, pp. 1265-1267.

[223] Whittle, P. (1981). "Risk-sensitive LQG control," *Adv. Appl. Prob.*, vol. 13, pp. 764-777.

[224] Whittle, P. (1986). "A risk-sensitive certainty equivalence principle," in *Essays in Time Series and Allied Processes* (Applied Probability Trust, London), pp 383-388.

[225] Willems, J.C. (1971). "Least Squares Stationary Optimal Control and the Algebraic Riccati Equation", *IEEE Trans. Auto. Control*, vol. AC-16, no. 6, pp. 621-634.

[226] Willems, J. C. (1981). "Almost invariant subspaces: an approach to high gain feedback design – Part I: almost controlled invariant subspaces," *IEEE Trans. Auto. Control*, vol. AC-26, pp. 235-252.

[227] Willems, J. C., A. Kitapci and L. M. Silverman (1986). "Singular optimal control: a geometric approach," *SIAM J. Control Optimiz.*, Vol. 24, pp. 323-337.

[228] Wimmer, H. K. (1985). "Monotonicity of Maximal solutions of algebraic Riccati equations," *Systems and Control Letters*, Vol. 5, April 1985, pp. 317-319.

[229] Wonham, W. M. (1968). "On a matrix Riccati equation of stochastic control," *SIAM J. Control*, Vol. 6, No. 4, pp. 681-697.

[230] Wonham, W.M. (1985). *Linear Multivariable Control: A Geometric Approach*, third edition, Springer-Verlag, New York.

[231] Youla, D.C. and J.J. Bongiorno, Jr. (1985). "A feedback theory of two-degree-of-freedom optimal Wiener-Hopf design," *IEEE Trans. Auto. Control*, vol. AC-30, No. 7, pp. 652-665.

[232] Youla, D.C., H.A. Jabr, and J.J. Bongiorno (1976). "Modern Wiener-Hopf design of optimal controllers: part I," *IEEE Trans. Auto. Control*, vol. AC-21, pp. 3-13.

[233] Youla, D.C., H.A. Jabr, and J.J. Bongiorno (1976). "Modern Wiener-Hopf design of optimal controllers: part II," *IEEE Trans. Auto. Control*, vol. AC-21, pp. 319-338.

[234] Youla, D.C., H.A. Jabr, and C. N. Lu (1974). "Single-loop feedback stabilization of linear multivariable dynamical plants," *Automatica*, Vol. 10, pp. 159-173.

[235] Youla, D. C. and M. Saito (1967). "Interpolation with positive-real functions", *Journal of The Franklin Institute*, Vol. 284, No. 2, pp. 77-108.

[236] Young, P. M. (1993). *Robustness with Parametric and Dynamic Uncertainty*, PhD Thesis, California Institute of Technology.

[237] Yousuff, A. and R. E. Skelton (1984). "A note on balanced controller reduction," *IEEE Trans. Automat. Contr.*, vol.AC-29, No. 3, pp. 254-257.

[238] Zames, G. (1966). "On the input-output stability of nonlinear time-varying feedback systems, parts I and II," *IEEE Trans. Auto. Control*, vol. AC-11, pp. 228 and 465.

[239] Zames, G. (1981). "Feedback and optimal sensitivity: model reference transformations, multiplicative seminorms, and approximate inverses," *IEEE Trans. Auto. Control*, vol. AC-26, pp. 301-320.

[240] Zames, G. and A. K. El-Sakkary (1980). "Unstable systems and feedback: The gap metric," *Proc. Allerton Conf.*, pp. 380-385.

[241] Zhou, K. (1992). "Comparison between $\mathcal{H}_2$ and $\mathcal{H}_\infty$ controllers," *IEEE Trans. on Automatic Control,* Vol. 37, No. 8, pp. 1261-1265.

[242] Zhou, K. (1992). "On the parameterization of $\mathcal{H}_\infty$ controllers," *IEEE Trans. on Automatic Control,* Vol. 37, No. 9, pp. 1442-1445.

[243] Zhou, K. (1993). "Frequency weighted model reduction with $\mathcal{L}_\infty$ error bounds," *Syst. Contr. Lett.*, Vol. 21, 115-125, 1993.

[244] Zhou, K. (1993). "Frequency weighted $\mathcal{L}_\infty$ norm and optimal Hankel norm model reduction," submitted to *IEEE Trans. on Automatic Control* .

[245] Zhou, K. and P. P. Khargonekar (1988). "An algebraic Riccati equation approach to $\mathcal{H}_\infty$ optimization," *Systems and Control Letters*, vol. 11, pp. 85-92, 1988.

[246] Zhou, K., J. C. Doyle, K. Glover, and B. Bodenheimer (1990). "Mixed $\mathcal{H}_2$ and $\mathcal{H}_\infty$ control," Proc. of the 1990 American Control Conference, San Diego, California, pp. 2502-2507.

[247] Zhou, K., K. Glover, B. Bodenheimer, and J. Doyle (1994). "Mixed $\mathcal{H}_2$ and $\mathcal{H}_\infty$ performance objectives I: robust performance analysis," *IEEE Trans. on Automat. Contr.*, vol. 39, no. 8, pp. 1564-1574.

[248] Zhu, S. Q. (1989). "Graph topology and gap topology for unstable systems," *IEEE Trans. Automat. Contr.*, Vol. 34, NO. 8, pp. 848-855.

Index