

Ben M. Chen
Tong H. Lee
Kemao Peng
Venkatakrisnan Venkataramanan

AIC

Advances in
Industrial Control

Hard Disk Drive Servo Systems

2nd Edition



Springer

Advances in Industrial Control

Other titles published in this Series:

Robust Control of Diesel Ship Propulsion
Nikolaos Xiros

Hydraulic Servo-systems
Mohieddine Jelali and Andreas Kroll

Strategies for Feedback Linearisation
Freddy Garces, Victor M. Becerra, Chandrasekhar Kambhampati and Kevin Warwick

Robust Autonomous Guidance
Alberto Isidori, Lorenzo Marconi and Andrea Serrani

Dynamic Modelling of Gas Turbines
Gennady G. Kulikov and Haydn A. Thompson (Eds.)

Control of Fuel Cell Power Systems
Jay T. Pukrushpan, Anna G. Stefanopoulou and Huei Peng

Fuzzy Logic, Identification and Predictive Control
Jairo Espinosa, Joos Vandewalle and Vincent Wertz

Optimal Real-time Control of Sewer Networks
Magdalene Marinaki and Markos Papageorgiou

Process Modelling for Control
Benoît Codrons

Computational Intelligence in Time Series Forecasting
Ajoy K. Palit and Dobrivoje Popovic

Modelling and Control of mini-Flying Machines
Pedro Castillo, Rogelio Lozano and Alejandro Dzul

Rudder and Fin Ship Roll Stabilization
Tristan Perez

Measurement, Control, and Communication Using IEEE 1588
John Eidson

Piezoelectric Transducers for Vibration Control and Damping
S.O. Reza Moheimani and Andrew J. Fleming
Publication due March 2006

Windup in Control
Peter Hippe
Publication due April 2006

Manufacturing Systems Control Design
Stjepan Bogdan, Frank L. Lewis, Zdenko Kovačić and José Mireles Jr.
Publication due May 2006

Practical Grey-box Process Identification
Torsten Bohlin
Publication due May 2006

Nonlinear H_2/H_∞ Constrained Feedback Control
Murad Abu-Khalaf, Jie Huang and Frank L. Lewis
Publication due May 2006

Ben M. Chen, Tong H. Lee, Kemao Peng
and Venkatakrisnan Venkataramanan

Hard Disk Drive Servo Systems

2nd Edition

With 124 Figures

 Springer

Ben M. Chen, PhD
Department of Electrical and Computer
Engineering
National University of Singapore
4 Engineering Drive 3
Singapore 117576

Tong H. Lee, PhD
Department of Electrical and Computer
Engineering
National University of Singapore
4 Engineering Drive 3
Singapore 117576

Kemao Peng, PhD
Department of Electrical and Computer
Engineering
National University of Singapore
4 Engineering Drive 3
Singapore 117576

Venkatkrishnan Venkataramanan, PhD
Mechatronics and Recording Channel
Division
Data Storage Institute
DSI Building, 5 Engineering Drive 1
Singapore 117608

British Library Cataloguing in Publication Data
Hard disk drive servo systems. - 2nd ed. - (Advances in
industrial control)
1.Servomechanisms 2.Data disk drives - Design 3.Hard disks
(Computer science)
I.Chen, Ben M., 1963-
629.8'323
ISBN-10: 1846283043

Library of Congress Control Number: 2006921170

Advances in Industrial Control series ISSN 1430-9491
ISBN-10: 1-84628-304-3 2nd edition e-ISBN 1-84628-305-1 2nd edition Printed on acid-free paper
ISBN-13: 978-1-84628-304-8 2nd edition
ISBN 1-85233-500-9 1st edition

© Springer-Verlag London Limited 2006

First published 2002
Second edition 2006

MATLAB® and Simulink® are registered trademarks of The MathWorks, Inc., 3 Apple Hill Drive Natick,
MA 01760-2098, U.S.A. <http://www.mathworks.com>

Apart from any fair dealing for the purposes of research or private study, or criticism or review, as
permitted under the Copyright, Designs and Patents Act 1988, this publication may only be reproduced,
stored or transmitted, in any form or by any means, with the prior permission in writing of the publishers,
or in the case of reprographic reproduction in accordance with the terms of licences issued by the
Copyright Licensing Agency. Enquiries concerning reproduction outside those terms should be sent to
the publishers.

The use of registered names, trademarks, etc. in this publication does not imply, even in the absence of a
specific statement, that such names are exempt from the relevant laws and regulations and therefore free
for general use.

The publisher makes no representation, express or implied, with regard to the accuracy of the information
contained in this book and cannot accept any legal responsibility or liability for any errors or omissions
that may be made.

Printed in Germany

9 8 7 6 5 4 3 2 1

Springer Science+Business Media
springer.com

Advances in Industrial Control

Series Editors

Professor Michael J. Grimble, Professor of Industrial Systems and Director
Professor Michael A. Johnson, Professor (Emeritus) of Control Systems
and Deputy Director

Industrial Control Centre
Department of Electronic and Electrical Engineering
University of Strathclyde
Graham Hills Building
50 George Street
Glasgow G1 1QE
United Kingdom

Series Advisory Board

Professor E.F. Camacho
Escuela Superior de Ingenieros
Universidad de Sevilla
Camino de los Descubrimientos s/n
41092 Sevilla
Spain

Professor S. Engell
Lehrstuhl für Anlagensteuerungstechnik
Fachbereich Chemietechnik
Universität Dortmund
44221 Dortmund
Germany

Professor G. Goodwin
Department of Electrical and Computer Engineering
The University of Newcastle
Callaghan
NSW 2308
Australia

Professor T.J. Harris
Department of Chemical Engineering
Queen's University
Kingston, Ontario
K7L 3N6
Canada

Professor T.H. Lee
Department of Electrical Engineering
National University of Singapore
4 Engineering Drive 3
Singapore 117576

Professor Emeritus O.P. Malik
Department of Electrical and Computer Engineering
University of Calgary
2500, University Drive, NW
Calgary
Alberta
T2N 1N4
Canada

Professor K.-F. Man
Electronic Engineering Department
City University of Hong Kong
Tat Chee Avenue
Kowloon
Hong Kong

Professor G. Olsson
Department of Industrial Electrical Engineering and Automation
Lund Institute of Technology
Box 118
S-221 00 Lund
Sweden

Professor A. Ray
Pennsylvania State University
Department of Mechanical Engineering
0329 Reber Building
University Park
PA 16802
USA

Professor D.E. Seborg
Chemical Engineering
3335 Engineering II
University of California Santa Barbara
Santa Barbara
CA 93106
USA

Doctor K.K. Tan
Department of Electrical Engineering
National University of Singapore
4 Engineering Drive 3
Singapore 117576

Doctor I. Yamamoto
Technical Headquarters
Nagasaki Research & Development Center
Mitsubishi Heavy Industries Ltd
5-717-1, Fukahori-Machi
Nagasaki 851-0392
Japan

To our families

Series Editors' Foreword

The series *Advances in Industrial Control* aims to report and encourage technology transfer in control engineering. The rapid development of control technology has an impact on all areas of the control discipline. New theory, new controllers, actuators, sensors, new industrial processes, computer methods, new applications, new philosophies..., new challenges. Much of this development work resides in industrial reports, feasibility study papers and the reports of advanced collaborative projects. The series offers an opportunity for researchers to present an extended exposition of such new work in all aspects of industrial control for wider and rapid dissemination.

Hard disk drive systems are ubiquitous in today's computer systems and the technology is still evolving. There is a review of hard disk drive technology and construction in the early pages of this monograph that looks at the characteristics of the disks and there it can be read that: "bit density... continues to increase at an amazing rate", "spindle speed... the move to faster and faster spindle speeds continue", "form factors... the trend...is downward... to smaller and smaller drives", "performance... factors are improving", "redundant arrays of inexpensive disks... becoming increasingly common, and is now seen in consumer desktop machines", "reliability... is improving slowly... it is very hard to improve the reliability of a product when it is changing rapidly" and finally "interfaces... continue to create new and improved standards... to match the increase in performance of the hard disks themselves". To match this forward drive in technology, control techniques need to progress too and that is the main reason why Professor Chen and his co-authors T.H. Lee, K. Peng and V. Venkataramanan have produced this second edition of their well-received *Advances in Industrial Control* monograph *Hard Disk Drive Servo Systems*.

The monograph opens with two chapters that create the historical context and the system modelling framework for hard disk drive systems. These chapters are followed by the control and applications content of the monograph. Hard disk drive systems are beset by nonlinear effects arising from friction, high-frequency mechanical resonances and actuator saturation so any control methods used have to be able to deal with these physical problems. Furthermore, there are two operational modes to contend with: track seeking and track following each with

different performance specifications. The type of control solution proposed by Professor Chen and his co-authors emerges from the interplay between the desire to mitigate the nonlinear effects and yet find a control strategy to unify the control of the two operational modes. To reveal the strategy developed in this Foreword would be like prematurely revealing the ending of a fascinating mystery story.

The monograph also has other valuable features: Chapter 3 contains succinct presentations of five different control methods with formulas given for both continuous and discrete forms. Two chapters on nonlinear control follow that covering linear control techniques. These chapters review classical time-optimal control and introduce the relatively new composite nonlinear feedback (CNF) control method. Again, presentations are given in both the continuous-time and discrete-time domains for completeness.

The second part of the monograph comprises five applications studies presented over five chapters. Whilst the first three of these chapters test out the control methods discussed in earlier chapters, the last two chapters introduce new applications hardware into the hard disk drive servo system problem: microdrive systems and piezoelectric actuators; nonlinear system effects are prominent in these new hardware systems.

Overall, it is an excellent monograph that exemplifies the topicality of control engineering problems today. Many lecturers will find invaluable material within this monograph with which to enthuse and motivate a new generation of control engineering students. Right at the end of this monograph, Professor Chen and his co-authors have extracted a benchmark control design problem for a typical hard disk drive system. The authors present their solution and “invite interested readers to challenge our design”, so happy reading and computing!

M.J. Grimble and M.A. Johnson
Industrial Control Centre
Glasgow, Scotland, U.K.

Preface

Nowadays, it is hard for us to imagine what life would be like without computers and what computers would be like without hard disks. Hard disks provide an important data-storage medium for computers and other data-processing systems. Many of us can still recall that the storage medium used on computers in the 1960s and 1970s was actually paper, which was later replaced by magnetic tapes. The key technological breakthrough that enabled the creation of the modern hard disk drives (HDDs) came in the 1950s, when a group of researchers and engineers in IBM made the very first production hard disk, IBM 305 RAMAC (random access method of accounting and control). The first generation of hard disks used in personal computers in the early 1980s had a capacity of 10 megabytes. Modern hard disks have a capacity of several hundred gigabytes.

In modern HDDs, rotating disks coated with a thin magnetic layer or recording medium are written with data that are arranged in concentric circles or tracks. Data are read or written with a read/write (R/W) head, which consists of a small horseshoe-shaped electromagnet. It is suggested that, on a disk surface, tracks should be written as closely spaced as possible so that we can maximize the usage of the disk surface. This means an increase in the track density, which subsequently means a more stringent requirement on the allowable variations of the position of the head from the true track center. The prevalent trend in hard disk design is towards smaller drives with increasingly larger capacities. This implies that the track width has to be smaller, leading to lower error tolerance in the positioning of the head. As such, it is necessary to introduce more advanced control techniques to achieve tighter regulation in the control of the HDD servomechanism.

The scope of this second edition remains the same. It is to provide a systematic treatment on the design of modern HDD servo systems. We particularly focus on the applications of some newly developed control theories, namely the robust and perfect tracking (RPT) control, and the composite nonlinear feedback (CNF) control. Emphasis is made on HDD servo systems with either a single-stage voice-coil-motor (VCM) actuator or a dual-stage actuator in which an additional microactuator is attached to a conventional VCM actuator to provide faster response and hence higher bandwidth in the track-following stage. New design considerations and techniques,

which have drastically improved the overall performance of our HDD servo systems, are introduced in this new edition. We also take this opportunity to extend the CNF control technique to systems with external disturbances and to include a comprehensive modeling and compensation of friction and nonlinearities as well as a complete servo system design of a microdrive.

The intended audience of this book includes practicing engineers in hard disk and CD-ROM drive industries and researchers in areas related to servo systems and engineering. An appropriate background for this monograph would be some senior level and/or first-year graduate level courses in linear systems and multivariable control. Some knowledge of control techniques for systems with actuator nonlinearities would certainly be helpful.

We have the benefit of the collaboration of several coworkers, from whom we have learned a great deal. Many of the results presented in this monograph are the results of our collaboration. Among these coworkers are Professor Chang C. Hang of the National University of Singapore, Dr Siri Weerasooriya, Dr Tony Huang, Mr Wei Guo and Dr Guoxiao Guo of the Data Storage Institute of Singapore. We are indebted to them for their contributions.

The authors of this monograph are particularly thankful to Guoyang Cheng for his help in proofreading the whole manuscript. The first two authors would also like to thank their current and former graduate students, especially Yi Guo, Xiaoping Hu, Lan Wang, Teck-Beng Goh, Kexiu Liu, Zhongming Li, Chen Lin and Guoyang Cheng, for their help and contributions.

We are grateful to Professor Zongli Lin of the University of Virginia, for his invaluable comments and discussions on the subject related to the composite nonlinear feedback control technique of Chapter 5. This technique, originally proposed by Zongli and his coworkers and later enhanced by us, has emerged as an effective tool in designing HDD servo systems. We are also indebted to Professor Iven Mareels of the University of Melbourne and Professor Frank Lewis of the University of Texas at Arlington, who were visiting our department here at the National University of Singapore, for many beneficial discussions on related subjects.

We would like to acknowledge the National University of Singapore for providing us with the funds for three research projects on the development of HDD servo systems. We are also grateful to people in the Design Technology Institute and the Data Storage Institute of Singapore for their support to our projects.

Last, but certainly not the least, we owe a debt of gratitude to our families for their sacrifice, understanding and encouragement during the course of preparing this monograph. It is very natural that we once again dedicate this second edition to our families.

Kent Ridge, Singapore
October 2005

*Ben M. Chen
Tong H. Lee
Kemao Peng
V. Venkataramanan*

Contents

Notation	xvii
-----------------------	------

Part I Introduction and Background Material

1 Introduction	3
1.1 Introduction	3
1.2 Historical Development	5
1.2.1 Chronological List of HDD History	6
1.2.2 Trends in Advances of HDD Systems	8
1.3 Overview of HDD Servo Systems	9
1.3.1 Mechanical Structure of an HDD	9
1.3.2 Issues on Control System Design	11
1.4 Implementation Setup	17
1.5 Preview of Each Chapter	18
2 System Modeling and Identification	21
2.1 Introduction	21
2.2 Time-domain Methods	22
2.2.1 Impulse Response Analysis	22
2.2.2 Step Response Analysis	24
2.3 Frequency-domain Methods	26
2.3.1 Prediction Error Identification Approach	26
2.3.2 Least Square Estimation Method	29
2.4 Physical Effect Approach with Monte Carlo Estimations	32
2.4.1 Structural Analysis of Physical Effects	32
2.4.2 Monte Carlo Estimations	33
2.4.3 Verification and Validation	33

3	Linear Systems and Control	37
3.1	Introduction	37
3.2	Structural Decomposition of Linear Systems	38
3.2.1	Interpretation	41
3.2.2	Properties	43
3.3	PID Control	47
3.3.1	Selection of Design Parameters	47
3.3.2	Sensitivity Functions	48
3.4	H_2 Optimal Control	49
3.4.1	Continuous-time Systems	50
3.4.2	Discrete-time Systems	59
3.5	H_∞ Control and Disturbance Decoupling	68
3.5.1	Continuous-time Systems	69
3.5.2	Discrete-time Systems	74
3.6	Robust and Perfect Tracking Control	76
3.6.1	Continuous-time Systems	76
3.6.2	Discrete-time Systems	84
3.7	Loop Transfer Recovery Technique	88
3.7.1	LTR at Input Point	88
3.7.2	LTR at Output Point	93
4	Classical Nonlinear Control	95
4.1	Introduction	95
4.2	Time-optimal Control	96
4.2.1	Open-loop Bang-bang Control	98
4.2.2	Closed-loop Bang-bang Control	99
4.3	Proximate Time-optimal Servomechanism	101
4.3.1	Continuous-time Systems	101
4.3.2	Discrete-time Systems	103
4.4	Mode-switching Control	104
4.4.1	Continuous-time Systems	104
4.4.2	Discrete-time Systems	109
5	Composite Nonlinear Feedback Control	119
5.1	Introduction	119
5.2	Continuous-time Systems	120
5.2.1	Systems without External Disturbances	121
5.2.2	Systems with External Disturbances	132
5.2.3	Selection of Nonlinear Feedback Parameters	139
5.2.4	An Illustrative Example	141
5.3	Discrete-time Systems	142
5.3.1	Systems without External Disturbances	142
5.3.2	Systems with External Disturbances	151
5.3.3	Selection of Nonlinear Feedback Parameters	158
5.3.4	An Illustrative Example	161

5.4	Can We Beat Time-optimal Control?	162
5.5	CNF Control Software Toolkit	164
5.5.1	Software Framework and User Guide	166
5.5.2	An Illustrative Example	172

Part II HDD Servo Systems Design

6	Track Following of a Single-stage Actuator	179
6.1	Introduction	179
6.2	VCM Actuator Model	180
6.3	Track-following Controller Design	181
6.4	Simulation and Implementation Results	188
6.4.1	Track-following Test	188
6.4.2	Frequency-domain Test	191
6.4.3	Runout Disturbance Test	191
6.4.4	Position Error Signal Test	198
7	Track Seeking of a Single-stage Actuator	201
7.1	Introduction	201
7.2	Track Seeking with PTOS Control	202
7.3	Track-seeking with MSC	203
7.4	Track Seeking with CNF Control	205
7.5	Simulation and Implementation Results	206
7.5.1	Track-seeking Test	206
7.5.2	Frequency-domain Test	209
8	Dual-stage Actuated Servo Systems	217
8.1	Introduction	217
8.2	Modeling of a Dual-stage Actuator	218
8.3	Dual-stage Servo System Design	220
8.4	Simulation and Implementation Results	224
8.4.1	Track-following Test	225
8.4.2	Frequency-domain Test	225
8.4.3	Runout Disturbance Test	225
8.4.4	Position Error Signal Test	239
9	Modeling and Design of a Microdrive System	243
9.1	Introduction	243
9.2	Modeling of the Microdrive Actuator	245
9.2.1	Structural Model of the VCM Actuator	245
9.2.2	Identification and Verification of Model Parameters	249
9.3	Microdrive Servo System Design	255

9.4	Simulation and Implementation Results	259
9.4.1	Track-following Test	259
9.4.2	Frequency-domain Test	259
10	Design of a Piezoelectric Actuator System	269
10.1	Introduction	269
10.2	Linearization of Nonlinear Hysteretic Dynamics	272
10.3	Almost Disturbance Decoupling Controller Design	275
10.4	Final Controller and Simulation Results	280
11	A Benchmark Problem	291
	References	297
	Index	307

Notation

We adopt the following notation and abbreviations throughout this monograph.

\mathbb{R}	the set of real numbers
\mathbb{C}	the entire complex plane
\mathbb{C}°	the set of complex numbers inside the unit circle
\mathbb{C}^{\otimes}	the set of complex numbers outside the unit circle
\mathbb{C}°	the unit circle in the complex plane
\mathbb{C}^{-}	the open left-half complex plane
\mathbb{C}^{+}	the open right-half complex plane
\mathbb{C}°	the imaginary axis in the complex plane
I	an identity matrix
I_k	an identity matrix of dimension $k \times k$
X'	the transpose of X
X^{H}	the complex conjugate transpose of X
$\text{Im}(X)$	the range space of X
$\text{Ker}(X)$	the null space of X
X^{\dagger}	the Moore–Penrose (pseudo) inverse of X
$\lambda(X)$	the set of eigenvalues of X
$\lambda_{\max}(X)$	the maximum eigenvalue of X
$\sigma_{\max}(X)$	the maximum singular value of X
$\ X\ $	the usual 2-norm of a matrix X
$\ G\ _2$	the H_2 -norm of a stable system $G(s)$ or $G(z)$
$\ g\ _2$	the l_2 -norm of a signal $g(t)$ or $g(k)$
L_2	the set of all functions whose l_2 norms are finite
$\ g\ _p$	the l_p -norm of a signal $g(t)$ or $g(k)$

L_p	the set of all functions whose l_p -norms are finite
$\ G\ _\infty$	the H_∞ -norm of a stable system $G(s)$ or $G(z)$
$\dim(\mathcal{X})$	the dimension of a subspace \mathcal{X}
\mathcal{X}^\perp	the orthogonal complement of a subspace \mathcal{X} of \mathbb{R}^n
ARE	algebraic Riccati equation
CNF	composite nonlinear feedback
DSA	digital signal analyzer
DSP	digital signal processor
GB	gigabytes
HDD	hard disk drive
LDV	laser Doppler vibrometer
LQG	linear quadratic Gaussian
LQR	linear quadratic regulator
LTR	loop transfer recovery
MB	megabytes
MSC	mode-switching control
N/RRO	non-/repeatable runouts
PES	position error signal
PID	proportional-integral-derivative
PTOS	proximate time-optimal servomechanism
RPT	robust and perfect tracking
R/W	read/write
TMR	track misregistration
TOC	time-optimal control
TPI	tracks per inch (kTPI = kilo TPI)
VCM	voice-coil-motor
ZOH	zero-order hold

Also, $C^{-1}\{\mathcal{X}\} := \{x \mid Cx \in \mathcal{X}\}$, where \mathcal{X} is a subspace and C is a matrix. Finally, we append a \diamond at the end of a proof or a result statement.

Introduction and Background Material

Introduction

1.1 Introduction

Hard disk drives (HDDs) provide an important data-storage medium for computers and other data-processing systems. In most commercial HDDs, rotating disks coated with a thin magnetic layer or recording medium are written with data that are arranged in concentric circles or tracks. Data are read or written with a read/write (R/W) head, which consists of a small horseshoe-shaped electromagnet. Figure 1.1 shows a simple illustration of a typical hard disk servo system with a voice-coil-motor (VCM) actuator.

The two main functions of the R/W head-positioning servomechanism in disk drives are track seeking and track following. Track seeking moves the R/W head from the present track to a specified destination track in minimum time using a bounded control effort. Track following maintains the head as close as possible to the destination track center while information is being read from or written to the disk. Track density is the reciprocal of the track width. It is suggested that, on a disk surface, tracks should be written as closely spaced as possible so that we can maximize the usage of the disk surface. This means an increase in the track density, which subsequently means a more stringent requirement on the allowable variations of the position of the heads from the true track center.

The prevalent trend in hard disk design is towards smaller hard disks with increasingly larger capacities. This implies that the track width has to be smaller, which leads to lower error tolerance in the positioning of the head. The controller for track following has to achieve tighter regulation in the control of the servomechanism. Basically, the servo system of an HDD can be divided into three stages, *i.e.* the track-seeking, track-settling and track-following stages (see Figure 1.2 for a detailed illustration). Current HDDs use a combination of classical control techniques, such as the proximate time-optimal control technique in the track-seeking stage, and lead-lag compensators, proportional-integral-derivative (PID) compensators in the track-following stage, plus some notch filters to reduce the effects of high-frequency resonance modes (see, *e.g.*, [1–16] and references cited therein). These classical methods can no longer meet the demand for HDDs of higher performance. Thus, many con-

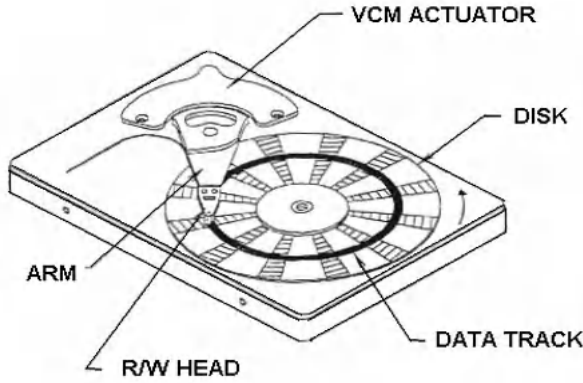


Figure 1.1. A typical HDD with a VCM actuator servo system

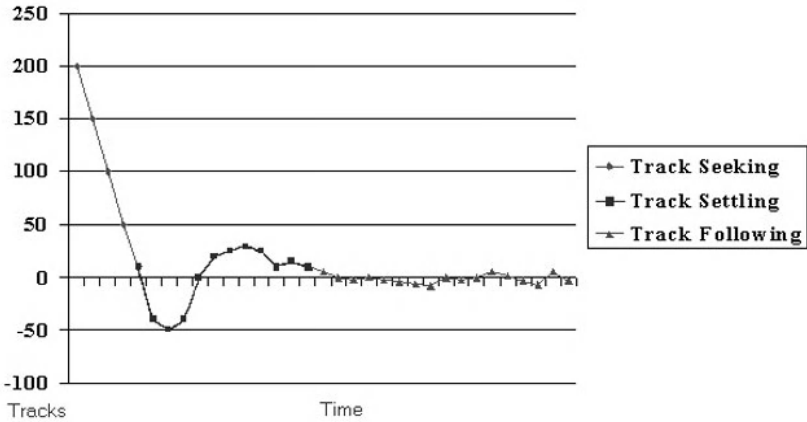


Figure 1.2. Track seeking and following of an HDD servo system

trol approaches have been tried, such as the linear quadratic Gaussian (LQG) with the loop transfer recovery (LTR) approach (see, *e.g.*, [17–19], H_∞ control approach (see, *e.g.*, [20–26], and adaptive control (see, *e.g.*, [27–30]) and so on. Although much work has been conducted to date, more studies need to be done to achieve better performance in HDDs.

The scope of this book is to provide a systematic treatment on the design of modern HDD servo systems. In particular, we focus on the applications of some newly developed results in control theory, *i.e.* robust and perfect tracking (RPT) control, which is suitable for track following, and composite nonlinear feedback (CNF) control, which is for track seeking and following. The emphasis is on HDD servo systems with either a single-stage VCM actuator or a dual-stage actuator in which an additional microactuator is attached to a conventional VCM actuator to provide

faster response and hence higher bandwidth in the track-following stage. Most of the results presented in this book are from research carried out by the authors and their coworkers over the last few years. The purpose of this book is to discuss various aspects of the subject under a single cover.

1.2 Historical Development

The first generation of hard disks used in PCs had a capacity of 10 megabytes (MB) and cost over \$100 per MB. Modern hard disks have capacities approaching 100 gigabytes (GB) and cost less than 1 cent per MB. This represents an improvement of 1000000% in less than 20 years and now it is cumulatively improving at 70% per year. At the same time, the speed of the hard disk and its interfaces has also increased dramatically.

Some of the very earliest computers had no storage at all. Each time a program had to be run it would have to be entered manually. It was realized then that to utilize the power of computers fully there was a need for permanent storage.

During the initial search for permanent storage, paper played a major role in human life. The computer scientists were also psychologically influenced by *paper*. This led to the use of paper as the first storage medium on computers, though magnetic storage had already gained momentum by that time. Programs and data were recorded using holes punched into paper tapes or punch cards to represent a “1”, and paper blocks to represent a “0” (or *vice versa*). This type of storage was used for many years until the creation of magnetic tapes. However, these tapes also lost their place when random access to the data was needed for quick and efficient usage of data stored. Thus, an improvement needed to be found. Disk drive development took an eventful spin when IBM announced, in May 1955, a product that offered unprecedented random-access storage to 5 million characters each of 7-bit.

These early prototypes had the heads of the hard disk in contact with the disk surface. This was done to allow the low-sensitivity electronics to be able to better read the magnetic fields on the disk surface. However, owing to the fact that manufacturing techniques were not nearly as sophisticated as they are now, it was not possible to produce a disk surface that was smooth enough for the head to slide smoothly over it at high speed while in contact with the surface. As a result, the heads and the magnetic coating on the surface of the disk would wear out over time. Thus the problem of reliability was not addressed.

IBM engineers working under R. Johnson at IBM in San Jose, California, between 1952 and 1954 realized that, with the proper design, the R/W heads could be suspended above the disk surface and read the bits as they passed underneath. This critical discovery, that contact with the surface of the disk was no longer necessary, was implemented as IBM 305 RAMAC (random access method of accounting and control), introduced on September 13, 1956. This early version stored 5 million characters on 50 disks, each 24" in diameter. The capacity was approximately 5 MB. Its bit density was about 2000 bits per square inch and the data transfer rate was an im-

pressive 8800 bytes per second. Over the succeeding years, the technology improved incrementally; bit density, capacity and performance all increased.

Next, we summarize the interesting history of the hard disk. In what follows, we present lists of some historical “firsts” and new trends in the development of HDDs. These lists are generated from the following sources on the net: www.pcguid.com, www.storage.ibm.com, www.storagereview.com and www.mkdata.dk [31, 32].

1.2.1 Chronological List of HDD History

There have been a number of important “firsts” in the world of hard disks over their first 40 years or so. The following is a list, in chronological order, of some of the products developed during the past half-century that introduced key or important technologies in HDDs.

- **FIRST HARD DISK (1956):** IBM 305 RAMAC was introduced. It had a capacity of about 5 MB, stored on fifty 24" disks. Its bit density was a mere 2000 bits per square inch and its data throughput was about 8800 bytes per second.
- **FIRST AIR-BEARING HEADS (1962):** IBM's model 1301 lowered the flying height of the R/W heads to 250 microinches. It had a 28-MB capacity with half as many heads as the original RAMAC, and increased both bit density and throughput by about 1000%.
- **FIRST REMOVABLE DISK DRIVE (1965):** IBM's model 2310 was the first disk drive with a removable disk pack. While many PC users think of removable hard disks as being a modern invention, in fact they were very popular in the 1960s and 1970s.
- **FIRST FERRITE HEADS (1966):** IBM's model 2314 was the first hard disk to use ferrite core heads, the first type later used on PC hard disks.
- **FIRST MODERN HARD DISK DESIGN (1973):** IBM's model 3340, nicknamed the *Winchester*, was introduced. With a capacity of 60 MB, it introduced several key technologies that led to it being considered by many as the ancestor of the modern disk drives.
- **FIRST THIN-FILM HEADS (1979):** IBM's model 3370 was the first with thin-film heads, which would for many years be the standard in the HDD industry.
- **FIRST 8" FORM FACTOR DISK DRIVE (1979):** IBM's model 3310 was the first disk drive with 8" platters, greatly reduced in size from the 14" that had been the standard for over a decade.
- **FIRST 5.25" FORM FACTOR DISK DRIVE (1980):** Seagate's ST-506 was the first drive in the 5.25" form factor, used in the earliest PCs.
- **FIRST 3.5" FORM FACTOR DISK DRIVE (1983):** Rodime introduced RO352, the first disk drive to use the 3.5" form factor, which became one of the most important industry standards.
- **FIRST EXPANSION CARD DISK DRIVE (1985):** Quantum introduced the Hard-card, a 10.5-MB hard disk mounted on an industry standard architecture (ISA) expansion card for PCs that were originally built without a hard disk. This product put Quantum “on the map” so to speak.

- FIRST VOICE-COIL ACTUATOR 3.5" DRIVE (1986): Conner Peripherals introduced CP340, the first disk drive to use a voice-coil actuator.
- FIRST "LOW-PROFILE" 3.5" DISK DRIVE (1988): Conner Peripherals introduced CP3022, which was the first 3.5" drive to use the reduced 1" height, now called *low profile* and the standard for modern 3.5" drives.
- FIRST 2.5" FORM FACTOR DISK DRIVE (1988): PrairieTek introduced a drive using 2.5" platters. This size later became a standard for portable computing.
- FIRST DRIVE WITH MR HEADS AND PARTIAL RESPONSE AND MAXIMUM LIKELIHOOD (PRML) DATA DECODING (1990): IBM's model 681 (Redwing), an 857 MB drive, was the first to use MR heads and PRML data decoding.
- FIRST THIN-FILM DISKS (1991): IBM's *Pacifica* mainframe drive was the first to replace oxide media with thin-film media on the platter surface.
- FIRST 1.8" FORM FACTOR DISK DRIVE (1991): Integral Peripherals' 1820 was the first hard disk with 1.8" platters, later used for PC-card disk drives.
- FIRST 1.3" FORM FACTOR DISK DRIVE (1992): Hewlett Packard's C3013A is the first 1.3" drive.
- FIRST 1" HIGH 1 GB DISK DRIVE (1993): IBM unveiled the world's first 1" high 1 GB disk drive, storing 354 million bits per square inch.
- FIRST 7200 RPM ULTRA ATA-INTERFACE DISK DRIVE (1997): Industry's first of this kind for desktop computers from Seagate Technology.
- FIRST 10000 RPM DISK DRIVE (1998): Seagate Technology introduced the first 10000 rpm drives, *i.e.* the 9.1-GB (ST19101) and 4.55-GB (ST34501) Cheetah family.
- FIRST ULTRA ATA/100 DISK DRIVES (2000): Seagate announced the first Ultra ATA/100 interface on its Barracuda ATA II disk drive, the industry's fastest desktop PC disk drive.
- LARGEST HDD (2000): At the time of the preparation of the first edition, Seagate's Barracuda 180 was the largest single drive in the world. It had a capacity of 180 GB.
- FIRST 100 GB/IN² DISK DRIVE (2001): Seagate Technology demonstrated a disk drive of more than 100 billion data bits per square inch.
- FIRST 60-GB/PLATTER DISK DRIVE (2002): Seagate Technology launched the Barracuda ATA V disc drive; it's the first hard drive to achieve 120 GB using only two discs.
- FIRST 2.5"/10,000 RPM ENTERPRISE DISK DRIVE (2004): Seagate Technology introduced Savvio, the world's first family of 2.5" enterprise-class hard disk drives.
- LARGEST 1" DISK DRIVE (2005): At the time of the preparation of this second edition, Seagate Technology produces the largest capacity 1" disk drive, which has a capacity of 8 GB.
- LARGEST HDD (2005): At the time of the presentation of this second edition, Barracuda 7200.9 of Seagate Technology is the largest drive in the world, which has a capacity of 500 GB.

1.2.2 Trends in Advances of HDD Systems

In spite of a slow change in the basic design of hard disks over the years, accelerated improvements in terms of their capacity, storage, reliability and other characteristics have been made. In what follows, the various trends are highlighted.

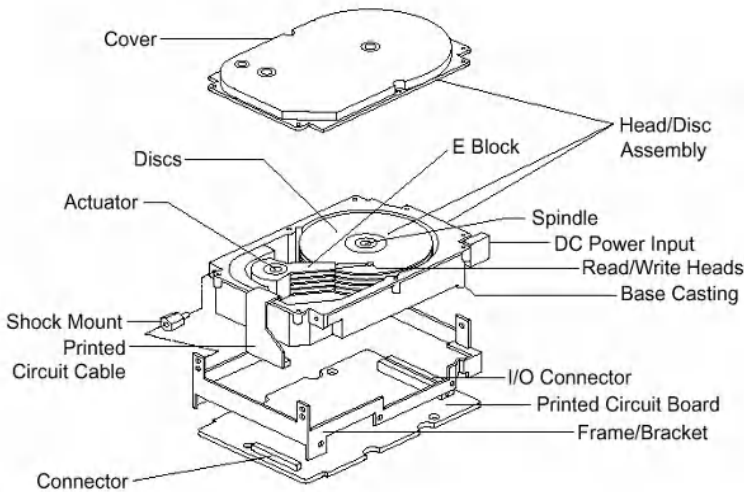
- **BIT DENSITY:** The bit density of hard disk platters continues to increase at an amazing rate, even exceeding some of the optimistic predictions of a few years ago. Densities in the laboratory are now approaching 1000 Gbits per square inch, and modern disks pack as much as 60 GB of data onto a single 3.5" platter.
- **CAPACITY:** Hard disk capacity continues to increase at an accelerating rate. From 10 MB in 1981, the normal capacity is now well over 400 GB. Consumer drives would most likely have a capacity of 1 TB within a couple of years.
- **SPINDLE SPEED:** The move to faster and faster spindle speeds continues. Since increasing the spindle speed improves both random access and sequential performance, this is likely to continue. 7200 rpm spindles are now standard on mainstream IDE/ATA drives. A 15000 rpm SCSI drive was announced by Seagate in 2000.
- **FORM FACTOR:** The trend in form factors is downward: to smaller and smaller drives. 5.25" drives have now all but disappeared from the mainstream market, with 3.5" drives dominating the desktop and server segment. In the mobile world, 2.5" drives are the standard, with smaller sizes becoming more prevalent; IBM in 1999 announced its Microdrive, a tiny 170 MB or 340 MB device, only 1" in diameter and less than 0.25" thick. Over the next few years, desktop and server drives are likely to make a transition to the 2.5" form factor as well. The primary reasons for this "shrinking trend" include the enhanced rigidity of smaller platters, reduction of mass to enable faster spin speeds, and improved reliability due to enhanced ease of manufacturing.
- **PERFORMANCE:** Both positioning and transfer performance factors are improving. The speed with which data can be pulled from the disk is increasing more rapidly than the improvement of positioning performance, suggesting that, over the next few years, addressing seek time and latency will be the areas of greatest value to hard disk engineers.
- **REDUNDANT ARRAYS OF INEXPENSIVE DISKS (RAID):** In the province of only high-end servers, the use of multiple disk arrays to improve performance and reliability is becoming increasingly common, and is now seen even in consumer desktop machines.
- **RELIABILITY:** The reliability of hard disks is improving slowly as manufacturers refine their processes and add new reliability-enhancing features, but this characteristic is not changing nearly as rapidly as the others above. It is simply very hard to improve the reliability of a product when it is changing rapidly.
- **INTERFACES:** Despite the introduction to the PC world of new interfaces, such as the IEEE-1394 and universal serial bus (USB), the mainstream interfaces are the same as they were through the 1990s: IDE/ATA and SCSI. The interfaces

themselves continue to create new and improved standards with higher maximum transfer rates, to match the increase in performance of the hard disks themselves.

1.3 Overview of HDD Servo Systems

1.3.1 Mechanical Structure of an HDD

The physical structure of a typical modern hard disk drive is depicted in Figure 1.3. The authors of this book are thankful to Seagate Technology for granting permission to use this figure in our work. A brief description (see also, *e.g.*, [33]) is given below:



Courtesy of Seagate Technology

Figure 1.3. Mechanical structure of a typical HDD

1. **DEVICE ENCLOSER.** This is the most important component, as it determines the reliability of the disk drives. It helps to keep the contamination low. With the aid of recirculation and a breather filter, it keeps out dust and other contamination that could enter between the R/W heads and the platters over which they float, and reduces the possibility of head crashes. The two major parts, the base casting and top cover, are sealed with a gasket. The base casting provides supports for the spindle, actuator, VCM yoke and electronics card.
2. **DISK.** Every hard disk has one or more flat rotating disks, each with two magnetic surfaces, called platters. These are made of either an Al-Mg alloy substrate material electroless plated with Ni-P, or a mixture of glass and ceramic. The magnetic material, to allow data storage, is applied as a thin coating on both

sides of each platter together with a carbon overcoat. The surfaces of each platter are precision machined and treated to remove any imperfections, and attention is paid during the manufacturing process to ensure a very smooth surface.

3. **ACTUATOR ASSEMBLY.** This consists of a VCM, data flex cable or printed circuit cable, actuator arms and crash-stops at both ends of travel. The data are read/written from/to the platters using the R/W heads mounted on the top and bottom surfaces of each platter. The heads are supported by the actuator arm. The actuator in HDDs, *i.e.* the VCM actuator, is so named as it works like a loud-speaker. The electrical input to the VCM is supplied through a flex cable. The coil of the VCM actuator extends between a yoke/magnets. The write-driver/pre-amplifier is often part of the actuator assembly, which is mounted on a data flex cable.
4. **HEAD/SUSPENSION ASSEMBLY.** The R/W heads are of ferrite, metal-in-gap, thin-film or magnetoresistive (MR) types. Older types, *i.e.* ferrite, metal-in-gap, and thin film, used the principle of electromagnetic induction, whereas the modern disk drive heads use MR heads, which use the principle of change of magnetoresistance. Both read and write operations in older disk drives were performed by a single head, but the modern HDDs use separate heads for read and write operations. These heads are positioned only microinches above the recording medium on an air-bearing surface, which is often referred to as a slider. A gimbal attaches the slider to a stainless steel suspension to allow for pitch and roll, and the suspension is attached to the arm of the actuator by a ball swaging.
5. **SPINDLE AND MOTOR ASSEMBLY.** These are responsible for turning the hard disk platters with stable, reliable and consistent turning power for thousands of hours of often continuous use. All hard disks use servo-controlled DC spindle motors and are configured for direct connection, *i.e.* there are no belts or gears used to connect them to the hard disk platter spindle. The critical component of the hard disk's spindle motor is the set of spindle motor bearings at each end of the spindle shaft. These bearings are used to turn the platters smoothly. The disk clamper and spacers are other important parts of this assembly.
6. **ELECTRONICS CARD:** This provides an interface to the host personal computer (PC). The most common interfaces used are the integrated drive electronics (IDE), the advanced technology attachment (ATA), and the small computer systems interface (SCSI), which all use integrated electronic circuits. These integrated circuits have a power driver for the spindle motor, VCM, R/W electronics, servo demodulator, controller chip for timing control and control of interface, microcontroller/digital signal processor (DSP) for servo control and control interface, and ROM and RAM for microcode and data transfer.

Lastly, we note that a fairly complete report on the basic mechanical and electrical structures of hard and floppy disk drives used in the 1970s and early 1980s can be found in Zhang [34].

1.3.2 Issues on Control System Design

For smaller drives with larger capacities, the control system has to achieve tighter regulation in the control of the servomechanism. To read (or write) the data reliably from (or to) the disk, the absolute track-following error with respect to the target track center, which is commonly called track misregistration (TMR), must be less than 10% of the track pitch. For example, for a 3.5" HDD with 25 kTPI, the track pitch is about 1 μm and its TMR must be less than 0.1 μm . Thus, for 70 kTPI, TMR must be less than 0.036 μm . This requires rigorous analysis of the sources of TMR and development of advanced techniques to overcome or eliminate these sources to meet the increasing demand for higher TPI. Figure 1.4 shows a typical disk drive servo channel indicating the various sources of disturbances and errors. Some of the larger components of TMR are due to the following error sources, listed roughly in the order of impact (see, *e.g.*, [35] and references cited therein):

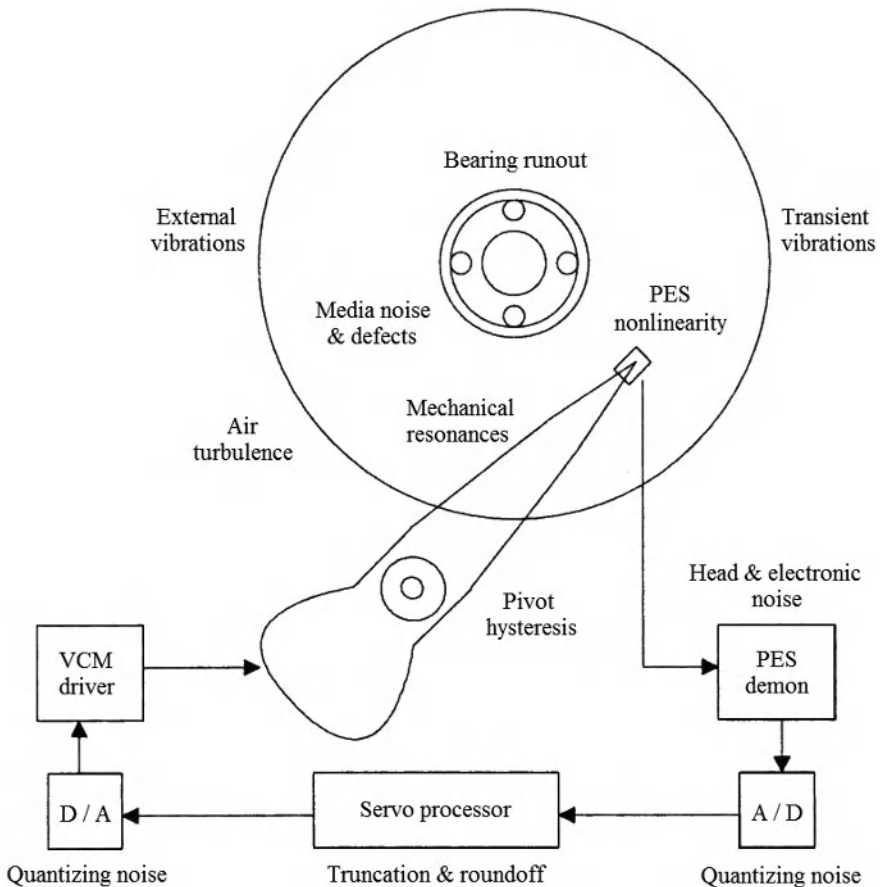


Figure 1.4. Sources of error in an HDD servo system

1. external shock and vibrations present in portable devices;
2. TMR caused by bearing hysteresis and poor velocity estimates during track-settling mode;
3. servo pattern nonlinearities and inaccuracies caused by head, media, and servo-writing effects;
4. mechanical resonances in suspension, actuator, disk, and housing;
5. electronic noise in recording channel entering the servo demodulator;
6. nonrepeatable spindle runout caused by bearings;
7. variations in RRO caused by thermal and other drifts.

A good HDD servo system has not only the desired track-seeking and following performance, but also the robustness to overcome all the above-listed disturbances and uncertainties. The following are robustness issues that one should consider in designing an HDD servo system:

Disturbance Rejection

As discussed earlier, higher TPI requires a tighter TMR, which is formally defined as three times the position error variance of the true position error signal (PES), *i.e.* $3\sigma_{\text{pes}}$. The sources of disturbances, which are the error sources contributing to $3\sigma_{\text{pes}}$, can be classified into three categories: input disturbance, output disturbance and measurement noise. The input disturbance is typically a color noise due to flexure, an electronic bias superimposed with selective energy arising from the natural frequencies of the various mechanical perturbations such as resonances, vibrations and friction. The output disturbance is also a color noise due to spindle rotation and its effects such as runout, windage and media noise. The measurement noise is a typical white noise due to the position-measurement techniques and/or sensors. These disturbances and noise can be modeled as in Figure 1.5. The objective would be to reject the effect of the disturbances and the measurement noise and achieve minimum position error variance.

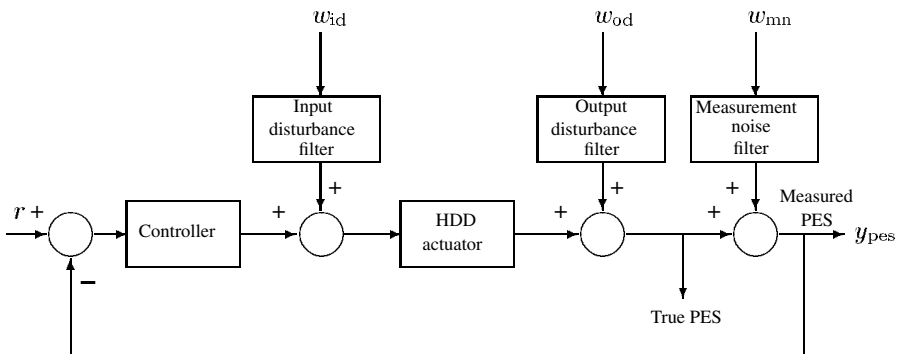


Figure 1.5. Modeling of disturbances in an HDD servo system

It has been shown [36] that, under some simplifications, σ_{pes} can be approximated as

$$\sigma_{\text{pes}} = \sqrt{\frac{1}{N-1} \sum_{i=1}^N y_{\text{pes}}^2(i)} \quad (1.1)$$

where N is the number of samples. Recently, Li *et al.* [37] gave a solution that minimizes σ_{pes} by converting the minimization of σ_{pes} into an equivalent H_2 optimal control problem for an auxiliary system, which contains the dynamics of the actuator, the input and output disturbances and the measurement noise, and which has a properly defined disturbance input and an output to be controlled. The problem to minimize σ_{pes} can then be solved using any appropriate H_2 optimal control methods (see, *e.g.*, Chapter 3).

Runout Compensation

Next, to understand runout disturbances, we recall the two main functions of HDD servo systems, *i.e.* track seeking and track following, which are usually achieved by two different controllers. The track-seeking controller moves the R/W head to the target track in minimum time; after this, when the control is switched to the track-following controller, it must make the R/W head follow the target track and keep the errors as small as possible. Thus, all HDDs must have a position-measurement mechanism. The position feedback signals in most HDD servomechanisms are derived through prerecorded position information recorded on one side of a disk surface at the time of manufacture using a servo writer. Ideally, servo tracks are perfect concentric circles. However, in the process of servo writing, the head that writes the signals cannot be kept perfectly still, due to, *e.g.*, the presence of vibration and NRRO effects, which result in tracks that are not perfect circles. This apparent track motion causes the R/W head to move in an attempt to minimize the position error, which results in positioning of the R/W head away from the real data track. Such an imperfection is termed a *runout*. This runout, depending upon its nature, can be classified as repeatable and nonrepeatable. In what follows we discuss these two types of runout and the methods available to compensate these disturbances collected from a literature survey.

Repeatable Runout. When the sampling frequency is equal to the spindle rotation frequency, or one of its multiples, the runout motion produced by the apparent track is repeated. This repeated runout, which is locked to the spindle rotation in both frequency and phase, is what we call a RRO. Thus, the major source of RROs is the eccentricity of the track. Other sources include the offset of the track center with respect to the spindle center, bearing geometry and wear, and motor geometry [38]. RROs caused by factors other than the eccentricity would cause a large amount of RROs at the rotational frequency of the spindle or its multiples, which is common to all tracks.

An RRO is a repetitive event in that both its amplitude and phase are locked to the rotation of the spindle. Therefore, this prior knowledge of an RRO can be used

as a feedforward signal to compensate the tracking error. Roughly, there exist three approaches to reject RRO: 1) repetitive control, 2) feedback control based on the internal model principle, and 3) identification and feedforward control. Many versions of the above compensation techniques for RROs have been reported, and some include adaptive feed-forward cancellation and repetitive control (see, *e.g.*, [38–41], PID with repetitive control (see, *e.g.*, [42]) and recurrent neural networks (see, *e.g.*, [43]). To be a little more specific, assuming that the RRO, $d(t)$, is a time-varying unknown disturbance consisting of a sum of n sinusoids of known frequencies, *i.e.*

$$d(t) = \sum_{i=1}^n \left[a_i(t) \cos(\omega_i t) + b_i(t) \sin(\omega_i t) \right] \quad (1.2)$$

the adaptive feedforward compensation approach is to design a control

$$u(t) = \hat{d}(t) = \sum_{i=1}^n \left[\hat{a}_i(t) \cos(\omega_i t) + \hat{b}_i(t) \sin(\omega_i t) + g_i k_i y(t) \right] \quad (1.3)$$

Next, an appropriate adaptive algorithm is used to adjust the estimates $\hat{a}_i(t)$ and $\hat{b}_i(t)$ so that these estimates are made equal to the nominal values, *i.e.*

$$\hat{a}_i(t) = a_i(t), \quad \hat{b}_i(t) = b_i(t) \quad (1.4)$$

The RRO disturbance approximated as in Equation 1.2 can then be canceled by the reproduced signal. This technique is also demonstrated by using neural networks in [43].

Nonrepeatable Runout. NRROs are a product of disk drive vibration and electrical noise in the measurement channel (see, *e.g.*, [1]). More specifically, the causes of NRRO are spindle-bearing defects, windage-induced disk flutter, electronics noise in the measurement channel, *etc.*, present during servo-track writing. NRROs can be minimized via improved servo writing, use of better bearings, and improved design of electronics. Since RROs are the harmonics of the motor rotational frequency in the frequency domain, an NRRO is the subtraction of the harmonics from the total indicated runout (TIR), which can be defined as the distance difference between the R/W head and the previously written track in an HDD, and hence an NRRO in the time domain can be easily constructed by the inverse Fourier transform of an NRRO in the frequency domain [44]. Alternatively, for a better understanding of the time trends of an NRRO, specifically a motor NRRO, Ohmi [45] proposed subtracting the averaged RRO from the experimentally obtained radial vibration of a rotor, the TIR. That is, an NRRO can be derived as follows [45]:

$$\text{RRO}(j) = \frac{1}{N} \sum_{i=1}^N \text{TIR}(i, j), \quad j = 1, 2, \dots, M \quad (1.5)$$

$$\text{NRRO}(i, j) = \text{TIR}(i, j) - \text{RRO}(j), \quad j = 1, 2, \dots, M \quad (1.6)$$

where N is the number of revolutions of the rotor, M is the number of samples per revolution, i represents the number of disk revolutions and j represents the number

of phases from a fixed point of the slit. An NRRO can be taken care of by the servo controller through improved loop bandwidth. However, the increase in servo bandwidth required to reject an NRRO is mainly determined by three factors, *i.e.* the servo sampling rate, the spectrum of the measurement noise, and the existence of plant resonance modes. The effect of resonance modes and their compensation is discussed in the following section. Recent research suggests the use of improved mechanical design, with a damped disk substrate and fluid-bearing spindles [46], which imposes less stress on the servo loop, to reject NRROs.

Resonance Compensation

The actuator and HDD structures, of course, are not perfectly rigid and have hundreds of flexible modes. This flexibility gives rise to vibrations, which results in a longer time to settle at the target track and amounts to a significant component of the TMR. The servo bandwidth of modern disk drives is approaching 2 kHz, and it has been proved in the literature that resonance modes that exist within a decade away from the servo-crossover frequency degrade the system performance. In short, in modern disk drives, resonance or vibration modes are the major sources of NRROs. Each resonance mode can be modeled as a second-order transfer function. The VCM actuator transfer function displaying multiple resonance modes can be modeled as [35],

$$G(s) = \frac{k_t}{Js^2} \prod_{i=1}^n \frac{\omega_i^2}{s^2 + 2\zeta_i\omega_i s + \omega_i^2} \quad (1.7)$$

where k_t is torque constant, J is inertia, ζ_i and ω_i are, respectively, the damping ratio and the natural frequency of the i th resonance mode. For simplicity, the frequency-response characteristics with a single resonance mode in actuator dynamics for a typical commercial drive are shown in Figure 1.6, in which the characteristics without resonance mode is shown by dashed lines. There is a significant difference in phase angle of the transfer function with the resonance mode. This tends to cause a loss of gain margin in the compensated loop and hence reduces its stability. Although there are hundreds of such resonances in an actual disk drive, many of the characteristics can be defined by considering only three or four modes, as other modes have an insignificant amplitude or are of too high a frequency to be of interest [1]. Some of the important resonance modes that must be considered in the design of high-density disk drives are the quasirigid body mode, the pivot bearing, the lateral elastic bending mode, and the vibrations of the individual disk platters.

Currently, the resonances of an HDD head actuator assembly caused by the pivot bearing have become a critical issue, since these resonances have been found to be the major design factor limiting the higher servo control bandwidth [47]. These resonances are excited during the track-seeking mode, and when the control is transferred to the track-following mode these vibrations result in an increased settling time. Recently, Mah *et al.* [48] have developed a novel moving-coil head actuator, which is designed deliberately to make sure that the force acting on the VCM is an orthogonal force so that there is no resulting force acting on the pivot bearing,

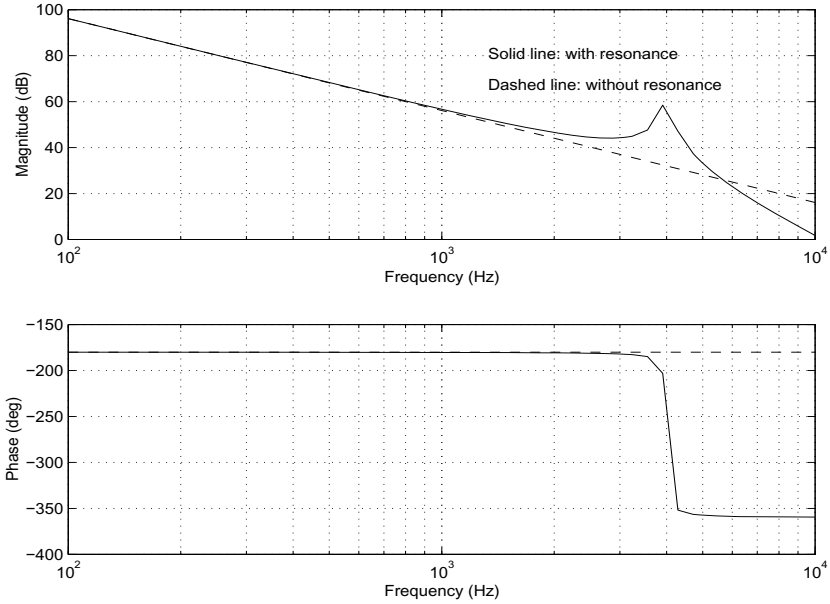


Figure 1.6. The ideal and “actual” frequency responses of HDD actuators

thereby minimizing residual vibrations. The rotational speeds of modern disk drives are progressively increasing, and hence the effect of the vibration of individual disk platters at their natural frequencies is a significant contributing factor to TMR in high-density disk drives. These resonances are driven primarily by internal windage excitation, and their behavior is dominated by the disk-material properties and geometry, and not by the spindle, enclosure, or structural design (see, *e.g.*, [49]). Use of alternate disk substrate materials can control these effects. Structural resonance modes can be compensated by using a notch filter as a precompensator. Since almost all structural resonance modes have lightly damped poles, the idea is to cancel lightly damped poles and place a pair of well-damped poles instead by using a notch filter. Hanselmann and Mortix [50] proposed the use of three notch filters to suppress the plant model resonance modes. These filters are preferred instead of low-pass filters because the sharper the cutoff in the magnitude of the frequency response, the lower the phase introduced in the loop. The transfer function of an analog notch filter is commonly chosen as (see, *e.g.*, [51])

$$G_n(s) = \frac{s^2 + (2\pi f_0)^2}{s^2 + \frac{2\pi f_0}{Q}s + (2\pi f_0)^2} \quad (1.8)$$

where f_0 is the center frequency and Q is the Q-factor. These notch filters can be realized by using switched capacitance filters. To use with digital control, digital notch filters can be realized using microprocessors or high-speed DSPs. Weaver and

Ehrlich [52] proposed the use of multirate filters to eliminate the resonance modes beyond the Nyquist frequency.

1.4 Implementation Setup

To make our work more complete, we have implemented almost all of our designs on actual HDDs with some highly advanced and accurate equipment. In what follows, we briefly summarize the key software and hardware tools used to obtain the simulation and implementation results.

1. **MATLAB[®] AND Simulink[®].** All offline computation and simulation of the results in this book are done using the well-known products from Mathworks, Inc., MATLAB[®] with its simulation package Simulink[®].
2. **SOFTWARE TOOLKITS.** Two software toolkits have been used in the development of results presented in this monograph. The linear systems toolkit [53], developed under a MATLAB[®] environment by the first author of the book and his coworkers, Zongli Lin and Xinmin Liu of the University of Virginia, collects a few tens of m-functions. These m-functions realize algorithms for computing linear system structures (such as the finite and infinite zero structures, invertibility structures, and many other properties) and algorithms for computing H_2 and H_∞ optimal controllers, as well as controllers that solve the H_∞ almost disturbance decoupling problem, and robust and perfect tracking problem. Another toolkit, the CNF control toolkit [54, 55], developed by our research team, has also been heavily used in designing track-seeking and track-following controllers in HDD servo systems presented in the coming chapters. Both toolkits are available for free at the website <http://hdd.ece.nus.edu.sg/~bmchen>.
3. **dSPACE DSP SYSTEM.** A dSpace DSP system is used in the actual implementation throughout the book. The system has the following main components:
 - *dSpace Add-on Card.* The main component of the dSpace DSP system is its add-on card, DS1102, which is built upon a Texas Instruments TMS320C31 floating-point DSP. The DSP has been supplemented by a set of analog-to-digital (A/D) and digital-to-analog (D/A) converters, a DSP microcontroller-based digital I/O subsystem and incremental sensor interfaces. Some major features of this add-on card are:
 - a) a TMS320C31 floating-point DSP;
 - b) two 16-bit 250 kHz and two 12-bit 800 kHz sampling A/D converters with input span of ± 10 V;
 - c) a quad 12-bit D/A converter with programmable output voltages;
 - d) a 16-bit fixed point digital I/O, a bit-selectable-parallel I/O port, four timers, six PWM circuits, and a serial interface.
 - *Real-time Interface (RTI) and Real-time Workshop (RTW).* The RTI acts as a link between Simulink[®] and the dSpace hardware. It has builtin hardware control functions and blocks for DS1102 add-on card based on Simulink[®]. This, together with the RTW, automatically generates real-time codes from

Simulink[®] offline models and implements these codes on the dSpace real-time hardware.

- *dSpace Control Desk*. This is a software platform that combines all the above tools of dSpace for controlling, monitoring, and automating the implementation process on the actual HDDs.
4. POLYTEC LASER DOPPLER VIBROMETER (LDV). The Polytec LDV is an optical instrument for accurately measuring velocity and displacement of vibrating surfaces completely without contact. The LDV system consists of two main components: 1) an optical sensor head or fiber optic unit (both are laser interferometers), which measures the dynamic Doppler shift from the vibrating object; and 2) a controller (processor), which provides power to the optics and demodulates the Doppler information using various types of Doppler signal decoder electronics, thereby producing an analog vibration signal (velocity and/or displacement) that can be viewed/measured by the customer using commercially available fast Fourier transform (FFT) analyzers and oscilloscopes. This instrument is used to measure the displacement and velocity of the R/W heads of HDDs.
 5. DYNAMIC SIGNAL ANALYZER (DSA). The HP35670A dynamic signal analyzer is a dynamic monitoring and measuring instrument that can be used for characterizing the performance and stability of a control system. Performance parameters, such as rise time, overshoot, and settling time, are generally specified in the time domain. Stability criteria, gain/phase margins, are generally specified in the frequency domain. The HP35670A DSA is capable of measuring in both the time and frequency domains. The instrument can also be used for system identification.
 6. VIBRATION-FREE TABLE. Since the success of the actual implementation depends largely on the accurate measurement of very small displacements of less than 1 μm , there is a need to isolate the HDD implementation setup from the external vibrations. A Vibraplane Model 9100/9200 series vibration-free workstation was used. These are designed and constructed to provide very effective isolation of vibrations at frequencies above 5 Hz and low amplification at low frequencies of 2–3 Hz. Hence, the use of this vibration-free table shows significant improvements in resolution and repeatability of the measurement.

The overall hardware setup in our laboratory at the National University of Singapore (NUS) is depicted in Figure 1.7.

1.5 Preview of Each Chapter

Since the publication of the first edition of this monograph [56], we have developed many new design methodologies and obtained many new theoretical and experimental results. We are to integrate all of these new developments and results in this second edition. This new edition can be naturally divided into two parts. The first part consists of Chapters 1 to 5, covering some background introductions to HDD servo



Figure 1.7. Implementation setup for HDD servo systems

systems, and some commonly used as well as newly developed linear and nonlinear control techniques. The second part, *i.e.* Chapters 6 to 10, deal with the design of specific single- and dual-stage actuated HDD servo systems.

More specifically, Chapter 2 recalls some commonly used system identification and modeling techniques, such as the prediction error identification and least squares estimation methods, applicable in the frequency domain, and the impulse response analysis and step response analysis in the time domain, as well as the physical effect approach together with Monte Carlo estimations. These techniques are employed to identify the models of VCM actuators and microactuators later in the book. Chapter 3 deals with linear systems and control techniques. In particular, we first recall a structural decomposition technique of linear systems, which has the distinct feature of displaying the finite and infinite zero structures as well as the invertibility structures of a given system, and plays a dominant role in the development of several linear control methods used in designing HDD servo systems. We also recall several commonly used linear control techniques, namely, the well-known classical PID control, H_2 optimal control, H_∞ control and almost disturbance decoupling, robust and perfect tracking (RPT) control, and the loop transfer recovery (LTR) technique. These methods are suitable for track-following control and have been used extensively in designing HDD servo systems in the literature. Chapter 4 focuses on some classical nonlinear control techniques such as the proximate time-optimal servomechanism (PTOS) and mode-switching control (MSC). PTOS is generally used to design a control law in the track-seeking stage of HDD servo systems, whereas the MSC design technique can be used to find a controller that is applicable for both track seeking and track following.

We devote Chapter 5 to cover the theory of so-called composite nonlinear feedback (CNF) control. The CNF control, which consists of a linear feedback law and a nonlinear feedback law, aims to improve the transient performance of the overall system. The linear feedback part is designed to yield a closed-loop system with a small damping ratio for quick response, whereas the nonlinear feedback part is used to increase the damping ratio of the closed-loop system as the system output approaches the target reference to reduce the overshoot. A software toolkit in MATLAB[®] is also presented for this new technique. It is available for free downloading at <http://hdd.ece.nus.edu.sg/~bmchen>.

In the second part, Chapters 6 and 7 focus on the design of HDD servo systems with a conventional single-stage VCM actuator. In particular, Chapter 6 deals with the modeling of the VCM actuator and design of track-following controllers using both the conventional PID control method and the recently developed RPT control method, whereas Chapter 7 deals with the design of track-seeking controllers. Three different designs using the PTOS, MSC and CNF approaches are presented in this chapter and their results are carefully compared. Likewise, Chapter 8 deals with the modeling and design of HDD servo systems with a dual-stage actuator. A complete HDD servo system with a dual-stage actuator is presented. Conventional HDDs with a single-stage VCM actuator usually have resonances in the positioning arm and low-frequency bearing effects. It is believed that the performances of such HDDs have been pushed almost to their limits. Dual-stage servo systems with high bandwidth and high accuracy control are a possible solution to overcome the problems associated with conventional HDDs.

Microdrives have become popular these days with high demand from many new applications. Unfortunately, many factors such as friction and nonlinearities, which can be safely neglected in normal drives, emerge as critical issues for microdrives. In Chapter 9, we present a comprehensive modeling and compensation of friction and nonlinearities as well as a complete servo system design of a typical microdrive. Chapter 10 considers a robust controller design for a piezoelectric bimorph nonlinear actuator using an H_∞ almost disturbance decoupling approach. Such an actuator has its own application domain including the use as a microactuator for the dual-stage actuated hard drives. Finally, we post in Chapter 11 a benchmark design problem for a typical HDD (Maxtor 51536U3) servo system, for which interested readers might try out to test their own techniques and designs.

System Modeling and Identification

2.1 Introduction

The purpose of this chapter is to revisit some basic theories and solutions of system identification, which are to be used later in the coming chapters to model various HDD systems. In general, the goal of system identification is to determine a mathematical model for a system or a process. Mathematical models may be developed either by use of “laws of nature”, commonly known as *modeling* or based on experimentation, which is known as *system identification* [57]. In order to achieve a certain desirable performance for a given plant, it is necessary to derive a model for the plant that is adequate for controller design. The conventional design techniques in linear control systems require either parametric or nonparametric models. For example, design methods via root locus or robust control technique require a transfer function or a state-space description of the plant to be controlled. The plant model is either described by the coefficients of certain polynomials or by the elements of state-space matrices. In either case, we call these polynomial coefficients or matrix elements the *parameters* of the model. The category of such models is a *parametric description* of the plant model. On the other hand, design based on Nyquist, Bode and Nichols methods requires curves of amplitude and phase of transfer function from input to the output as functions of real frequency ω . If we have experimental data from a typical frequency response test, then we will be able to obtain certain functional curves for the plant. These curves are called nonparametric models of the plant, as there is no finite set of numbers that describes it exactly (see, e.g., [1]).

Thus, for a given plant, the problem of system identification is to determine a system model from the relationship (either in the time or the frequency domain) between its input and output. The problem can be represented graphically as shown in Figure 2.1, in which $u(t)$ is the known input signal, $n(t)$ is the observation noise, and $y(t)$ is the measured output. A large variety of methods have been developed for solving such a problem (see, e.g., [58] and references cited therein). These methods include classical identification techniques (such as the impulse response analysis, step response analysis, frequency response identification) and equation error approaches and model adjustment techniques (such as the least square estimation, maximum

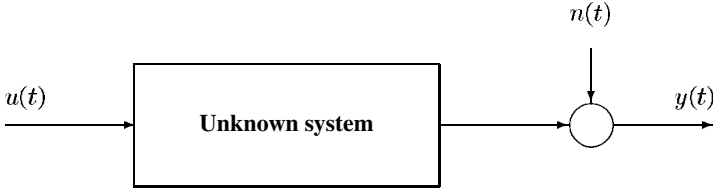


Figure 2.1. The unknown system to be identified

likelihood, and stochastic approximation, to name a few). The detailed derivations of these techniques can be found in a number of advanced texts devoted to system identification, *e.g.*, [57–61]. We note that the techniques are applicable only to linear systems. For systems with nonlinearities, other advanced modeling techniques have to be used. One of the key issues in modeling systems with nonlinearities is to determine the physical structures of the nonlinear components. Another issue is that there is generally no analytical solution for nonlinear differential equations, which causes trouble in identifying the model parameters. In the last section of this chapter, we present a so-called physical effect approach modeling method of [62] together with Monte Carlo estimations (see, *e.g.*, [63–65]). The technique will be utilized to identify a comprehensive model for a microdrive in Chapter 9, in which friction and nonlinearities associated with its VCM actuator are highly noticeable.

2.2 Time-domain Methods

In this section, we restrict our attention to identifying both parametric and nonparametric models through some commonly used time-domain techniques. Interested readers are referred to [57, 58, 66] for detailed materials on the identification through impulse and step response characteristics.

2.2.1 Impulse Response Analysis

Parametric Models. Parametric models are described by parameters of differential equations or transfer functions. From these analytic representations, plots or values of interest of frequency response can in general be generated without much difficulty, whereas the reverse process of deriving parameters from nonparametric model descriptions is much more difficult.

A fairly general parametric model of a single-input and single-output (SISO) system can be described by the following differential equation (see, *e.g.*, [66]),

$$\begin{aligned} a_n y^{(n)}(t) + \cdots + a_1 \dot{y}(t) + a_0 y(t) \\ = b_0 u(t - \tau_d) + b_1 \dot{u}(t - \tau_d) + \cdots + b_m u^{(m)}(t - \tau_d) \end{aligned} \quad (2.1)$$

Solving the differential equation for the input signal

$$u(t) = \delta(t) \quad (2.2)$$

with $\delta(t)$ being the unit impulse function, gives the impulse transfer function $h(t)$,

$$y(t) = h(t) \quad (2.3)$$

as the corresponding output function. Note that it is hard to generate an impulse input in the continuous-time domain, and hence this method is impractical.

Nonparametric Models. Consider again a SISO system as in Figure 2.1 with a scalar input signal $u(t)$ and a scalar output signal $y(t)$. Assume that the system is linear, time invariant and causal.

It is well known that a linear, time-invariant causal system can be described by its impulse response $g(\tau)$ as follows:

$$y(t) = \int_0^{\infty} g(\tau)u(t - \tau) d\tau + n(t) \quad (2.4)$$

Knowing $\{g(\tau)\}_{\tau=0}^{\infty}$ and knowing $u(s)$ for $s \leq t$, we can consequently compute the corresponding output $y(s)$, $s \leq t$ for any input. The impulse response is thus a complete characterization of the system.

The discrete equivalent of the output $y(t)$ can be written at the sampling instants $t_k = kT$, $k = 0, 1, 2, \dots$, as

$$y(t_k) = \int_0^{\infty} g(\tau)u(t_k - \tau) d\tau + n(t_k) \quad (2.5)$$

where T is the sampling period. Since, the input $u(t_k)$ is kept constant between the sampling instants:

$$u(t) = u(t_k) = u_k, \quad kT \leq t < (k+1)T \quad (2.6)$$

we can derive that

$$y(t_k) = \sum_{i=0}^{\infty} g(i)u(t_k - i) + n(t_k), \quad t_k = kT, \quad k = 0, 1, 2, \dots \quad (2.7)$$

Now, let $G(z)$ be the transfer function of the system from input to output with z being the usual forward shift operator, *i.e.*

$$G(z) = \sum_{k=0}^{\infty} g(k)z^{-k} \quad (2.8)$$

Then, Equation 2.7 can be written as

$$y(t_k) = G(z)u(t_k) + n(t_k) \quad (2.9)$$

If the system in Equation 2.9 is subjected to a pulse input

$$u(t_k) = \begin{cases} \alpha, & t_k = 0 \\ 0, & t_k \neq 0 \end{cases} \quad (2.10)$$

then the output is

$$y(t_k) = \alpha g(t_k) + n(t_k) \quad (2.11)$$

If the noise level is low, then the estimates of the coefficients of the impulse response $\{g(t_k)\}$ from an experiment will be

$$\hat{g}(t_k) = \frac{y(t_k)}{\alpha} \quad (2.12)$$

and the errors $n(t_k)/\alpha$. This simple analysis is called impulse response analysis. Unfortunately, many physical processes do not allow the error $n(t_k)/\alpha$ to be insignificant compared with the impulse response coefficients. Moreover, such an input could induce nonlinear effects that would disturb the linearized behavior of the model. As such, identification methods depended on impulse inputs are rarely used in practical situations.

2.2.2 Step Response Analysis

Parametric Models. Parametric models usually are described by their frequency response function,

$$G(\omega) = \frac{b_0 + b_1 j\omega + \dots b_m (j\omega)^m}{a_0 + a_1 j\omega + \dots a_n (j\omega)^n} \cdot \exp(-j\omega\tau) \quad (2.13)$$

where a_i , b_i , m , n and τ are parameters to be identified.

Many researchers proposed methods for determining parameter values from time functions of the process output provided that the process input is a well-determined signal. Of primary importance are step functions as input signals and step responses as output signals.

Almost all methods for the evaluation of step response functions use a small number of characteristic values of the response function. A first-order lag model with time delay and a frequency response

$$\hat{G}(\omega) = \frac{K}{1 + j\omega\tau_1} \cdot \exp(-j\omega\tau) \quad (2.14)$$

is widely used in step response analysis. In fact, there are quite a number of systems, especially in process control, that can be approximated by a first-order model with an appropriate delay. The parameters K , τ_1 and τ can be derived from the step response shown in Figure 2.2 with u_0 being the amplitude of the input step signal (see, e.g., [66]). In short, if the system model is of first order, one may need only obtain two pieces of information: (i) the steady-state response to the step input, and (ii) the time constant. The latter can be obtained either from the tangent with maximum slope of the step response or from the 10 to 90% rise time.

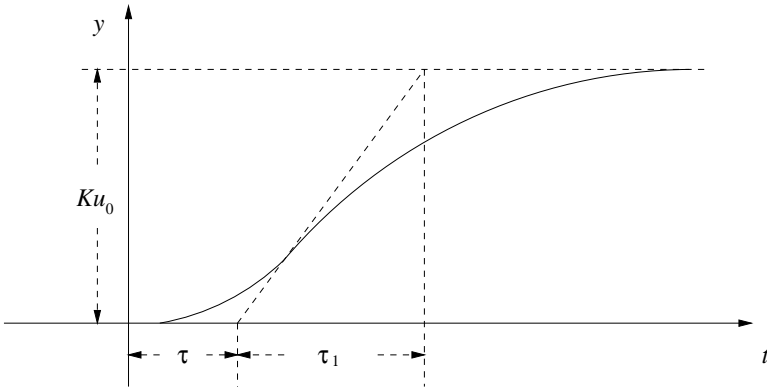


Figure 2.2. A typical step response

For a second-order system model (with two poles and no zero), there are two possible situations: 1) when the two poles are real and 2) when the poles are a complex conjugate pair. Formulae for finding these from measurements of the (a) steady-state response, (b) maximum overshoot, (c) time required to reach the first-peak, and (d) time required to reach 50% of the steady-state value (for overdamped systems) can be easily derived. For the general case of higher-order practical systems, it is perhaps best to use a gradient method to find the parameters of the model of a given order such that the integral of the square of the error is minimized (see, *e.g.*, [57]).

Nonparametric Models. Since the impulse response of a system is the derivative of the step response, the identification problem in this case may be regarded as the determination of the transfer function from the impulse response. Alternatively, a step function

$$u(t_k) = \begin{cases} \alpha, & t_k \geq 0 \\ 0, & t_k < 0 \end{cases} \quad (2.15)$$

when applied to Equation 2.9 gives the output

$$y(t_k) = \alpha \sum_{i=0}^k g(i) + n(t_k), \quad t_k = kT, \quad k = 0, 1, \dots \quad (2.16)$$

Then, the estimates can be obtained as

$$\hat{g}(t_k) = \frac{y(t_k) - y(t_{k-1})}{\alpha} \quad (2.17)$$

which has an error $[n(t_k) - n(t_{k-1})]/\alpha$. Hence, we would suffer again from large errors in most practical applications. But, if the goal is to determine some basic control-related characteristics, then the step responses from Equation 2.16 can very well furnish that information to a sufficient degree of accuracy. In fact, some well-known rules for tuning simple regulators, such as the Ziegler–Nichols rule, are based

on model information hidden in step responses. Based on plots of the step responses, some key characteristics of the system can be graphically constructed, which in turn can be used to determine system parameters.

2.3 Frequency-domain Methods

We recall in this section two identification methods in the frequency domain, namely, the prediction error identification approach and the least square estimation method. Both are particularly important to our studies in modeling the micro and VCM actuators in HDD servo systems in the forthcoming chapters. The theories behind these techniques can be found in various references (see, *e.g.*, [57, 67]).

2.3.1 Prediction Error Identification Approach

The prediction error approach is a black-box identification method. It includes the following three steps.

1. PARAMETER IDENTIFICATION. Suppose a system can be described as

$$y(k) = G(z^{-1})u(k) + H(z^{-1})n(k) \quad (2.18)$$

where $u(k)$ and $y(k)$ are its process input and output; $n(k)$ is noise input and supposed to be white; and

$$\left. \begin{aligned} G(z^{-1}) &= B(z^{-1})/A(z^{-1}) \\ H(z^{-1}) &= D(z^{-1})/A(z^{-1}) \\ A(z^{-1}) &= 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_{n_a}z^{-n_a} \\ B(z^{-1}) &= b_1z^{-1} + b_2z^{-2} + \dots + b_{n_b}z^{-n_b} \\ D(z^{-1}) &= d_1z^{-1} + d_2z^{-2} + \dots + d_{n_d}z^{-n_d} \end{aligned} \right\} \quad (2.19)$$

The predictor is:

$$\hat{y}(k|\theta) = [1 - H^{-1}(z^{-1}, \theta)]y(k) + H^{-1}(z^{-1}, \theta)G(z^{-1}, \theta)u(k) \quad (2.20)$$

where

$$\theta = \begin{pmatrix} a_1 \\ \vdots \\ a_{n_a} \\ b_1 \\ \vdots \\ b_{n_b} \\ d_1 \\ \vdots \\ d_{n_d} \end{pmatrix} \quad (2.21)$$

is the parameter vector of the system. Then, the prediction error given by a model is

$$e(k, \theta) = y(k) - \hat{y}(k|\theta) \quad (2.22)$$

Next, we define a loss function as

$$V_N(\theta, Z_N) = \frac{1}{N} \sum_{k=1}^N \ell(e(k, \theta)) \quad (2.23)$$

where $\ell(\cdot)$ is a scalar-valued positive function, and

$$Z_N = \left[y(1), \dots, y(N), u(1), \dots, u(N) \right] \quad (2.24)$$

is a set of input and output data from the experiment test. The desired system parameters can then be obtained by the minimization of this loss function, *i.e.*

$$\hat{\theta}_N = \arg \min V_N(\theta, Z_N) \quad (2.25)$$

2. DETERMINATION OF MODEL ORDER. The loss function $V_N(\theta, Z_N)$ can also be used to determine the order of a system. If the order of a model is lower than that of the system, then the value of the loss function will decrease significantly with the increase of the order of the model. However, when the order of the model is higher than that of the system, the increase of model order does not provide any more innovation for parameter identification, thus the value of $V_N(\theta, Z_N)$ will not decrease much. Therefore, the order of the system to be identified can be determined based on the decrease rate of $V_N(\theta, Z_N)$. Figure 2.3 shows a typical plot of the loss function versus identified model orders. It is clear from the plot that the order of the corresponding system to be identified is four.
3. MODEL VALIDATION. The last step of the prediction error identification method is to verify the correctness of the model obtained. It is clear that the residuals of the model can be obtained as

$$\begin{aligned} e(k, \theta) &= y(k) - \hat{y}(k|\theta) \\ &= \hat{H}(z^{-1}, \theta) \left[y(k) - \hat{G}(z^{-1}, \theta)u(k) \right] \end{aligned} \quad (2.26)$$

Obviously, if the model is correct, *i.e.*

$$\hat{G}(z^{-1}, \theta) = G_0(z^{-1}) \quad \text{and} \quad \hat{H}(z^{-1}, \theta) = H_0(z^{-1}) \quad (2.27)$$

the residual will tend to a white-noise sequence $n(k)$. However, the nonwhiteness of the residuals does not necessarily mean that the model is incorrect. In that case, the crosscorrelation of the input u and residuals e can be used to verify the model. If u and e are independent, this means that all information in the residuals is explained by the process model \hat{G} , and we can conclude that the estimate is correct. Otherwise the result is incorrect. The crosscorrelation of u and e is

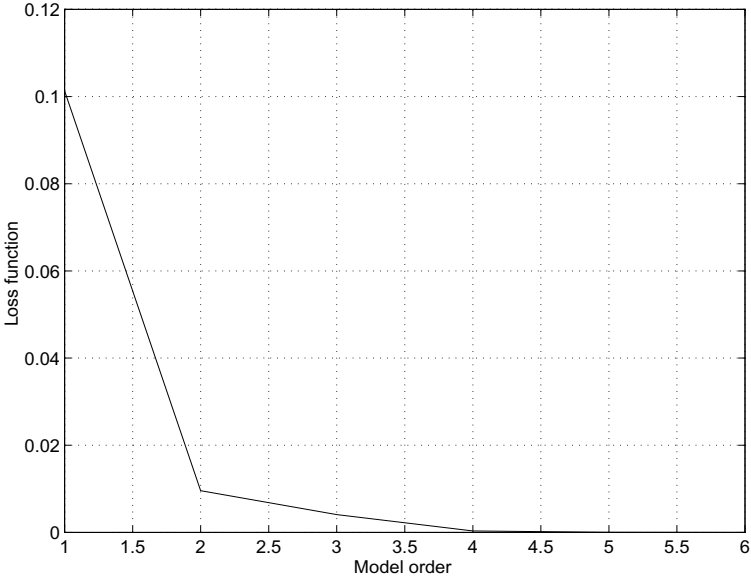


Figure 2.3. Values of loss function versus identified model orders

$$R_{eu}(\tau) = \mathbf{E}\{e(t + \tau)u(t)\} \quad (2.28)$$

where $\mathbf{E}\{\cdot\}$ is the expected value. If the residuals and input are independent, we have

$$\sqrt{N}R_{eu} \rightarrow \mathcal{N}(0, P) \quad \text{as } N \rightarrow \infty \quad (2.29)$$

where $P = \sum_{k=-\infty}^{\infty} R_e(k)R_u(k)$, and $\mathcal{N}(0, P)$ denotes the normal random distribution with zero mean and a variance P . Let N_α be the α -level of the $\mathcal{N}(0, P)$ such that

$$\mathbf{P} \left\{ |R_{eu}| \leq \sqrt{\frac{P}{N}} N_\alpha \right\} = \alpha \quad (2.30)$$

where $\mathbf{P}\{\cdot\}$ is the probability. Define the following null hypothesis:

$$H_0 : |R_{eu}| \leq \sqrt{\frac{P}{N}} N_\alpha \quad (2.31)$$

If H_0 is accepted, then we can say that the model is acceptable with a probability of $1 - \alpha$. Figure 2.4 shows a typical plot of the values of the crosscorrelation function between the input and the error residual. It can be seen that, for such a model, all the data are within the 95% confidence region. Hence, we can say that the corresponding identified model is acceptable with a probability of 95%.

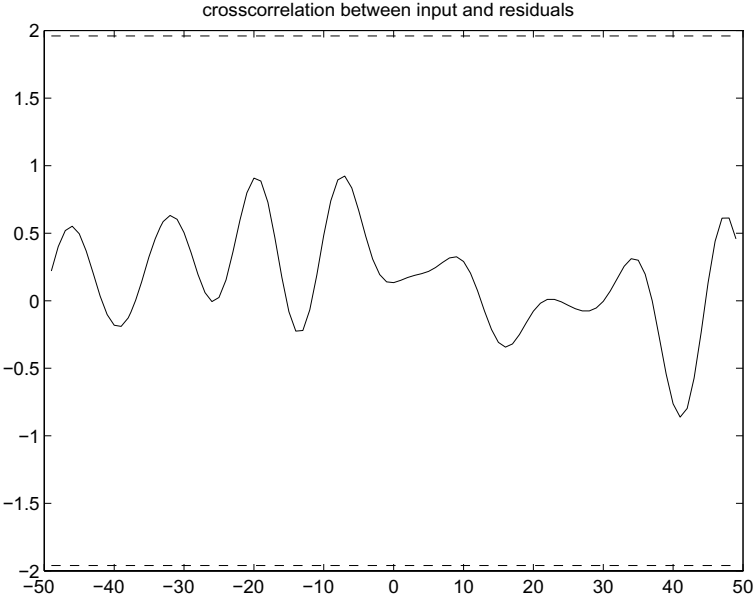


Figure 2.4. Model validation test

2.3.2 Least Square Estimation Method

We will utilize the frequency-response identification method (see, *e.g.*, [59]) to model our actuator. Such a method is applicable to minimum phase processes. We expect from the properties of the physical system that the VCM actuator is of minimum phase. The detailed procedure proceeds as follows: we first assume that the transfer function of a minimum phase plant is given by

$$G(s) = \frac{N(s)}{D(s)} = \frac{b_0 + b_1s + b_2s^2 + \cdots + b_ms^m}{1 + a_1s + a_2s^2 + \cdots + a_ns^n} \quad (2.32)$$

for some appropriate coefficients a_k , $k = 1, 2, \dots, n$, and b_k , $k = 0, 1, \dots, m$, with $n \geq m$. These parameters are to be identified. Then, its corresponding frequency response is given by

$$G(j\omega) = \frac{\alpha(\omega) + j\omega\beta(\omega)}{\sigma(\omega) + j\omega\tau(\omega)} = \frac{N(j\omega)}{D(j\omega)} \quad (2.33)$$

where

$$\left. \begin{aligned} \alpha(\omega_i) &= b_0 - b_2\omega_i^2 + b_4\omega_i^4 - \cdots \\ \beta(\omega_i) &= b_1 - b_3\omega_i^2 + b_5\omega_i^4 - \cdots \\ \sigma(\omega_i) &= 1 - a_2\omega_i^2 + a_4\omega_i^4 - \cdots \\ \tau(\omega_i) &= a_1 - a_3\omega_i^2 + a_5\omega_i^4 - \cdots \end{aligned} \right\} \quad (2.34)$$

Let $R(\omega)$ and $I(\omega)$ be the real and imaginary parts of the measured frequency response of the actuator system. The frequency response error between the model and the actual measurement data is given by

$$\mathcal{E}(j\omega) = [R(\omega) + jI(\omega)] - \frac{N(j\omega)}{D(j\omega)} \quad (2.35)$$

Thus, the parameters of the system can be obtained by minimizing the following index:

$$J = \sum_{i=1}^L |\mathcal{E}(j\omega_i)|^2 \quad (2.36)$$

where L is the total number of points of the measured data. Unfortunately, this is a nonlinear optimization problem, and it is difficult to solve. We then follow the results of [59] to modify the error norm as

$$J = \sum_{i=1}^L |D(j\omega_i)\mathcal{E}(j\omega_i)|^2 \quad (2.37)$$

The original problem now becomes a linear optimization problem. Using Equations 2.33 and 2.35, we can rewrite Equation 2.37 as follows

$$J = \sum_{i=1}^L \left\{ [X(\omega_i)]^2 + [Y(\omega_i)]^2 \right\} \quad (2.38)$$

where

$$X(\omega_i) = \sigma(\omega_i)R(\omega_i) - \omega_i\tau(\omega_i)I(\omega_i) - \alpha(\omega_i) \quad (2.39)$$

and

$$Y(\omega_i) = \omega_i\tau(\omega_i)R(\omega_i) + \sigma(\omega_i)I(\omega_i) - \omega_i\beta(\omega_i) \quad (2.40)$$

Therefore, J can be minimized by finding $\hat{b}_0, \hat{b}_1, \dots, \hat{b}_m$ and $\hat{a}_1, \hat{a}_2, \dots, \hat{a}_n$ such that

$$\left. \begin{aligned} \frac{\partial J}{\partial b_0} \Big|_{b_0=\hat{b}_0} &= \sum_{i=1}^L \left\{ 2X(\omega_i)(-1) \right\} \Big|_{b_0=\hat{b}_0} = 0 \\ \frac{\partial J}{\partial b_1} \Big|_{b_1=\hat{b}_1} &= \sum_{i=1}^L \left\{ 2Y(\omega_i)(-\omega_i) \right\} \Big|_{b_1=\hat{b}_1} = 0 \\ &\vdots \\ \frac{\partial J}{\partial a_1} \Big|_{a_1=\hat{a}_1} &= \sum_{i=1}^L 2\omega_i \left[Y(\omega_i)R(\omega_i) - X(\omega_i)I(\omega_i) \right] \Big|_{a_1=\hat{a}_1} = 0 \\ \frac{\partial J}{\partial a_2} \Big|_{a_2=\hat{a}_2} &= \sum_{i=1}^L -2\omega_i^2 \left[X(\omega_i)R(\omega_i) + Y(\omega_i)I(\omega_i) \right] \Big|_{a_2=\hat{a}_2} = 0 \\ &\vdots \end{aligned} \right\} \quad (2.41)$$

Rearranging the above equations, we obtain

$$\begin{pmatrix} \mathbf{A}_{11} & \mathbf{A}_{12} \\ \mathbf{A}_{21} & \mathbf{A}_{22} \end{pmatrix} \begin{pmatrix} \mathbf{b} \\ \mathbf{a} \end{pmatrix} = \begin{pmatrix} \mathbf{B}_1 \\ \mathbf{B}_2 \end{pmatrix} \quad (2.42)$$

where

$$\mathbf{A}_{11} = \begin{bmatrix} V_0 & 0 & -V_2 & 0 & V_4 & \cdots \\ 0 & V_2 & 0 & -V_4 & 0 & \cdots \\ V_2 & 0 & -V_4 & 0 & V_6 & \cdots \\ 0 & V_4 & 0 & -V_6 & 0 & \cdots \\ V_4 & 0 & -V_6 & 0 & V_8 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \mathbf{b} = \begin{bmatrix} \hat{b}_0 \\ \hat{b}_1 \\ \hat{b}_2 \\ \hat{b}_3 \\ \hat{b}_4 \\ \vdots \end{bmatrix} \quad (2.43)$$

$$\mathbf{A}_{12} = \begin{bmatrix} T_1 & S_2 & -T_3 & -S_4 & T_5 & \cdots \\ -S_2 & T_3 & S_4 & -T_5 & -S_6 & \cdots \\ T_3 & S_4 & -T_5 & -S_6 & T_7 & \cdots \\ -S_4 & T_5 & S_6 & -T_7 & -S_8 & \cdots \\ T_5 & S_6 & -T_7 & -S_8 & T_9 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \mathbf{a} = \begin{bmatrix} \hat{a}_1 \\ \hat{a}_2 \\ \hat{a}_3 \\ \hat{a}_4 \\ \hat{a}_5 \\ \vdots \end{bmatrix} \quad (2.44)$$

$$\mathbf{A}_{21} = \begin{bmatrix} T_1 & -S_2 & -T_3 & S_4 & T_5 & \cdots \\ S_2 & T_3 & -S_4 & -T_5 & S_6 & \cdots \\ T_3 & -S_4 & -T_5 & S_6 & T_7 & \cdots \\ S_4 & T_5 & -S_6 & -T_7 & S_8 & \cdots \\ T_5 & -S_6 & -T_7 & S_8 & T_9 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \mathbf{B}_1 = \begin{bmatrix} S_0 \\ T_1 \\ S_2 \\ T_3 \\ S_4 \\ \vdots \end{bmatrix} \quad (2.45)$$

$$\mathbf{A}_{22} = \begin{bmatrix} U_2 & 0 & -U_4 & 0 & U_6 & \cdots \\ 0 & U_4 & 0 & -U_6 & 0 & \cdots \\ U_4 & 0 & -U_6 & 0 & U_8 & \cdots \\ 0 & U_6 & 0 & -U_8 & 0 & \cdots \\ U_6 & 0 & -U_8 & 0 & U_{10} & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}, \quad \mathbf{B}_2 = \begin{bmatrix} 0 \\ U_2 \\ 0 \\ U_4 \\ 0 \\ \vdots \end{bmatrix} \quad (2.46)$$

and where

$$V_k = \sum_{i=1}^L \omega_i^k, \quad S_k = \sum_{i=1}^L \omega_i^k R(\omega_i) \quad (2.47)$$

$$T_k = \sum_{i=1}^L \omega_i^k I(\omega_i), \quad U_k = \sum_{i=1}^L \omega_i^k [R^2(\omega_i) + I^2(\omega_i)] \quad (2.48)$$

The desired parameters of the corresponding transfer function model can be obtained by solving the above equations.

We note that the methods recalled above are merely for the modeling and identification of the HDD VCM and microactuators in the forthcoming chapters. If systems to be identified are highly uncertain with disturbances, it would be more appropriate to use the methods reported in a recent monograph by Chen and Gu [68] to yield more accurate results.

2.4 Physical Effect Approach with Monte Carlo Estimations

Diverse methods have been proposed (see, *e.g.*, [13, 59, 69]) to identify the model of linear plants, in which the nonlinearities of the plants are assumed to be negligible. As such, these methods cannot be applied to identify the model of plants with significant nonlinearities. One of the key issues in identifying the model of a nonlinear plant is to configure its physical structure. Another issue is that there is generally no analytical solution for nonlinear differential equations, which causes problems in identifying the parameters in the model of nonlinear plants.

The so-called physical effect approach [62] with the Monte Carlo method (see, *e.g.*, [63–65]) can be applied to tackle these two problems efficiently. The approach is to examine and analyze physical effects that occur in or between physical parts of a nonlinear plant to configure a structured model and then to identify the parameters numerically in the obtained structured model with the Monte Carlo method. We proceed to summarize such an efficient parameter identification approach for nonlinear systems in the following.

2.4.1 Structural Analysis of Physical Effects

The structural analysis is to determine the physical structure of the nonlinear plants using analytical derivation based on some natural laws and theories, such as the well-known Newton's laws of motion, principles of circuits, electromagnetic effects, stochastic and probability theory, if applicable. It focuses on identifying the physical effects that occur in or between the components of the plants under consideration and on identifying their interconnection properties. Unfortunately, it is not always possible to capture all the properties of a physical plant. Instead, the structural analysis is more on the characterization of the major characteristics of physical plants based on some appropriate assumptions. For example, in HDD servo systems, Newton's laws can be applied to analyze the motion of the VCM actuator under the assumption that it is a rigid body. The permanent magnet associated with the coil is assumed to be constant for simplicity. Some other typical models of physical effects, such as the commonly used friction models and models for springs, can be applied to simplify the analysis. In principle, the structure of the nonlinearities has to be sufficient to characterize the main properties of the given system. One would generally obtain the structured model of a nonlinear plant in the following form:

$$\begin{cases} \dot{x} = f(x, u, P) \\ y = g(x, u, P) \end{cases} \quad (2.49)$$

where x , u and y are, respectively, the state, input and measurement output variables with appropriate dimensions. f and g are appropriate nonlinear functions. Lastly, P represents the set of all unknown parameters of the model, which is to be identified using some parameter-identification techniques.

2.4.2 Monte Carlo Estimations

In most situations, it is hard to find analytical solutions to the associated nonlinear differential equations. Thus, it is difficult to identify the parameters in nonlinear models. Alternatively, one has to search for solutions through numerical means. The Monte Carlo method (see, *e.g.*, [63–65]) is one such numerical technique, which can loosely be classified as a method purely based on numerical simulation. The essence of the Monte Carlo method is to employ certain stochastic techniques to identify the unknown parameters of a structured nonlinear model. The Monte Carlo method has been widely used from field to field even in the deterministic setting. It is very efficient in approximating solutions to various mathematical problems, for which their analytical forms are hard, if not impossible, to be determined.

Considering the structured model of Equation 2.49, the Monte Carlo estimation method is an optimization process in determining the set of the unknown parameters, P . It is to minimize the following performance index:

$$\min_P \left\{ \frac{1}{N} \sum_{k=1}^N [y_s(t_k) - y_e(t_k)]' W(t_k) [y_s(t_k) - y_e(t_k)] \right\} \quad (2.50)$$

subject to

$$\begin{cases} \dot{x} = f(x, u_e, P) \\ y_s = g(x, u_e, P) \end{cases} \quad (2.51)$$

where N is the number of the sampled data, $W(t_k)$, $k = 1, 2, \dots, N$, are appropriate weighting matrices, y_s and y_e are, respectively, the simulated and experimental measurement output responses corresponding to a set of preselected input signals u_e . Note that u_e should be carefully chosen to ensure that the unknown parameters in the structured model of the plant can be properly identified.

Various numerical optimization algorithms, such as genetic algorithm, simulated annealing and gradient algorithm, neural network and their combinations, can be adopted to solve the above problem. As usual, all these numerical methods might yield a locally optimal solution. To ensure obtaining a meaningful solution, one should have certain background knowledge on the problem to be solved. For example, in HDD servo systems, the ranges of certain portions of the system dynamics are known to us. Such information is useful and should be incorporated in the above optimization process.

Similarly, an optimization process in the frequency domain can also be formulated to identify certain parts of the system dynamics. In fact, it is effective to combine the results identified using both the time- and frequency-domain methods. In general, it would yield a more accurate solution. The overall Monte Carlo estimation processes in the time domain and frequency domain are, respectively, depicted in Figures 2.5 and 2.6.

2.4.3 Verification and Validation

Once a model with all its parameters is identified, it is necessary to perform a series of model verification and validation processes. Model verification is a series of

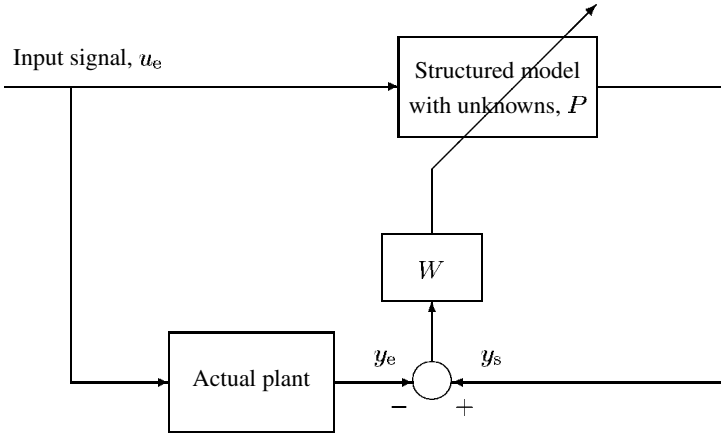


Figure 2.5. Monte Carlo estimation in the time-domain setting

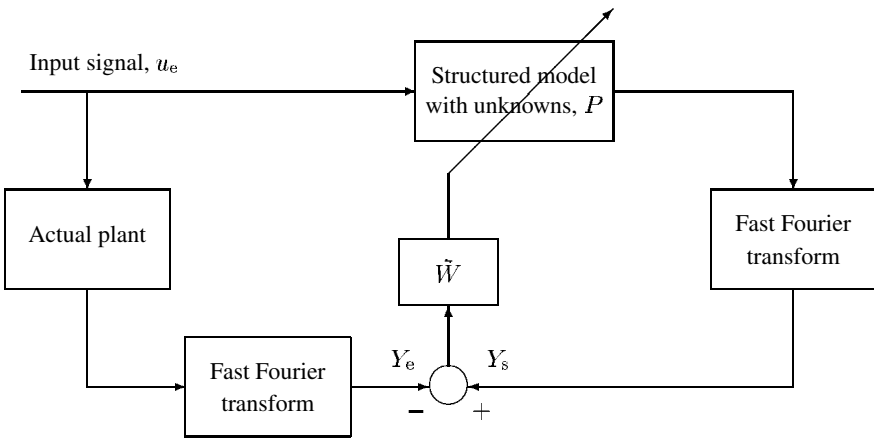


Figure 2.6. Monte Carlo estimation in the frequency-domain setting

quantitative examinations and comparisons between the actual experimental data and those generated from the identified model. It is to verify whether the identified model is a true representation of the real plants based on some intensive tests with various input-output responses other than those used in the identification process. On the other hand, validation is on qualitative examinations, which are to verify whether the features of the identified model are capable of displaying all of the essential characteristics of the actual plant. It is to recheck the process of the physical effect analysis, the correctness of the natural laws and theories used as well as the assumptions made.

In conclusion, verification and validation are two necessary steps that one needs to perform to ensure that the identified model is accurate and reliable. As mentioned earlier, the above technique will be utilized to identify the model of a commercial microdrive in Chapter 9.

Linear Systems and Control

3.1 Introduction

It is our belief that a good unambiguous understanding of linear system structures, *i.e.* the finite and infinite zero structures as well as the invertibility structures of linear systems, is essential for a meaningful control system design. As a matter of fact, the performance and limitation of an overall control system are primarily dependent on the structural properties of the given open-loop system. In our opinion, a control system engineer should thoroughly study the properties of a given plant before carrying out any meaningful design. Many of the difficulties one might face in the design stage may be avoided if the designer has fully understood the system properties or limitations. For example, it is well understood in the literature that a nonminimum phase zero would generally yield a poor overall performance no matter what design methodology is used. A good control engineer should try to avoid these kinds of problem at the initial stage by adding or adjusting sensors or actuators in the system. Sometimes, a simple rearrangement of existing sensors and/or actuators could totally change the system properties. We refer interested readers to the work by Liu *et al.* [70] and a recent monograph by Chen *et al.* [71] for details.

As such, we first recall in this chapter a structural decomposition technique of linear systems, namely the special coordinate basis of [72, 73], which has a unique feature of displaying the structural properties of linear systems. The detailed derivation and proof of such a technique can also be found in Chen *et al.* [71]. We then present some common linear control system design techniques, such as PID control, H_2 optimal control, H_∞ control, linear quadratic regulator (LQR) with loop transfer recovery design (LTR), together with some newly developed design techniques, such as the robust and perfect tracking (RPT) method. Most of these results will be intensively used later in the design of HDD servo systems, though some are presented here for the purpose of easy reference for general readers.

We have noticed that it is some kind of tradition or fashion in the HDD servo system research community in which researchers and practicing engineers prefer to carry out a control system design in the discrete-time setting. In this case, the designer would have to discretize the plant to be controlled (mostly using the ZOH

technique) first and then use some discrete-time control system design technique to obtain a discrete-time control law. However, in our personal opinion, it is easier to design a controller directly in the continuous-time setting and then use some continuous-to-discrete transformations, such as the bilinear transformation, to discretize it when it is to be implemented in the real system. The advantage of such an approach follows from the following fact that the bilinear transformation does not introduce unstable invariant zeros to its discrete-time counterpart. On the other hand, it is well known in the literature that the ZOH approach almost always produces some additional nonminimum-phase invariant zeros for higher-order systems with faster sampling rates. These nonminimum phase zeros cause some additional limitations on the overall performance of the system to be controlled. Nevertheless, we present both continuous-time and discrete-time versions of these control techniques for completeness. It is up to the reader to choose the appropriate approach in designing their own servo systems.

Lastly, we would like to note that the results presented in this chapter are well studied in the literature. As such, all results are quoted without detailed proofs and derivations. Interested readers are referred to the related references for details.

3.2 Structural Decomposition of Linear Systems

Consider a general proper linear time-invariant system Σ , which could be of either continuous- or discrete-time, characterized by a matrix quadruple (A, B, C, D) or in the state-space form

$$\Sigma : \begin{cases} \delta(x) = A x + B u \\ y = C x + D u \end{cases} \quad (3.1)$$

where $\delta(x) = \dot{x}(t)$ if Σ is a continuous-time system, or $\delta(x) = x(k+1)$ if Σ is a discrete-time system. Similarly, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$ are the state, input and output of Σ . They represent, respectively, $x(t)$, $u(t)$ and $y(t)$ if the given system is of continuous-time, or represent, respectively, $x(k)$, $u(k)$ and $y(k)$ if Σ is of discrete-time. Without loss of any generality, we assume throughout this section that both $[B' \ D']$ and $[C \ D]$ are of full rank. The transfer function of Σ is then given by

$$H(\varsigma) = C(\varsigma I - A)^{-1}B + D \quad (3.2)$$

where $\varsigma = s$, the Laplace transform operator, if Σ is of continuous-time, or $\varsigma = z$, the z -transform operator, if Σ is of discrete-time. It is simple to verify that there exist nonsingular transformations U and V such that

$$UDV = \begin{bmatrix} I_{m_0} & 0 \\ 0 & 0 \end{bmatrix} \quad (3.3)$$

where m_0 is the rank of matrix D . In fact, U can be chosen as an orthogonal matrix. Hence, hereafter, without loss of generality, it is assumed that the matrix D has the form given on the right-hand side of Equation 3.3. One can now rewrite system Σ of Equation 3.1 as

$$\begin{cases} \delta(x) = A x + [B_0 \ B_1] \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} \\ \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{bmatrix} C_0 \\ C_1 \end{bmatrix} x + \begin{bmatrix} I_{m_0} & 0 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} u_0 \\ u_1 \end{pmatrix} \end{cases} \quad (3.4)$$

where the matrices B_0 , B_1 , C_0 and C_1 have appropriate dimensions. Theorem 3.1 below on the special coordinate basis (SCB) of linear systems is mainly due to the results of Sannuti and Saberi [72, 73]. The proofs of all its properties can be found in Chen *et al.* [71] and Chen [74].

Theorem 3.1. *Given the linear system Σ of Equation 3.1, there exist*

1. *coordinate-free non-negative integers $n_a^-, n_a^0, n_a^+, n_b, n_c, n_d, m_d \leq m - m_0$ and $q_i, i = 1, \dots, m_d$, and*
2. *nonsingular state, output and input transformations Γ_s, Γ_o and Γ_i that take the given Σ into a special coordinate basis that displays explicitly both the finite and infinite zero structures of Σ .*

The special coordinate basis is described by the following set of equations:

$$x = \Gamma_s \tilde{x}, \quad y = \Gamma_o \tilde{y}, \quad u = \Gamma_i \tilde{u} \quad (3.5)$$

$$\tilde{x} = \begin{pmatrix} x_a \\ x_b \\ x_c \\ x_d \end{pmatrix}, \quad x_a = \begin{pmatrix} x_a^- \\ x_a^0 \\ x_a^+ \end{pmatrix}, \quad x_d = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{m_d} \end{pmatrix} \quad (3.6)$$

$$\tilde{y} = \begin{pmatrix} y_0 \\ y_d \\ y_b \end{pmatrix}, \quad y_d = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_{m_d} \end{pmatrix}, \quad \tilde{u} = \begin{pmatrix} u_0 \\ u_d \\ u_c \end{pmatrix}, \quad u_d = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{m_d} \end{pmatrix} \quad (3.7)$$

$$\delta(x_a^-) = A_{aa}^- x_a^- + B_{0a}^- y_0 + L_{ad}^- y_d + L_{ab}^- y_b \quad (3.8)$$

$$\delta(x_a^0) = A_{aa}^0 x_a^0 + B_{0a}^0 y_0 + L_{ad}^0 y_d + L_{ab}^0 y_b \quad (3.9)$$

$$\delta(x_a^+) = A_{aa}^+ x_a^+ + B_{0a}^+ y_0 + L_{ad}^+ y_d + L_{ab}^+ y_b \quad (3.10)$$

$$\delta(x_b) = A_{bb} x_b + B_{0b} y_0 + L_{bd} y_d, \quad y_b = C_b x_b \quad (3.11)$$

$$\begin{aligned} \delta(x_c) = & A_{cc} x_c + B_{0c} y_0 + L_{cb} y_b + L_{cd} y_d \\ & + B_c [E_{ca}^- x_a^- + E_{ca}^0 x_a^0 + E_{ca}^+ x_a^+] + B_c u_c \end{aligned} \quad (3.12)$$

$$y_0 = C_{0c} x_c + C_{0a}^- x_a^- + C_{0a}^0 x_a^0 + C_{0a}^+ x_a^+ + C_{0d} x_d + C_{0b} x_b + u_0 \quad (3.13)$$

and for each $i = 1, \dots, m_d$,

$$\begin{aligned} \delta(x_i) = & A_{q_i} x_i + L_{i0} y_0 + L_{id} y_d \\ & + B_{q_i} \left[u_i + E_{ia} x_a + E_{ib} x_b + E_{ic} x_c + \sum_{j=1}^{m_d} E_{ij} x_j \right] \end{aligned} \quad (3.14)$$

$$y_i = C_{q_i} x_i, \quad y_d = C_d x_d \quad (3.15)$$

Here the states $x_a^-, x_a^0, x_a^+, x_b, x_c$ and x_d are, respectively, of dimensions $n_a^-, n_a^0, n_a^+, n_b, n_c$ and $n_d = \sum_{i=1}^{m_d} q_i$, and x_i is of dimension q_i for each $i = 1, \dots, m_d$. The control vectors u_0, u_d and u_c are, respectively, of dimensions m_0, m_d and $m_c = m - m_0 - m_d$, and the output vectors y_0, y_d and y_b are, respectively, of dimensions $p_0 = m_0, p_d = m_d$ and $p_b = p - p_0 - p_d$. The matrices A_{q_i}, B_{q_i} and C_{q_i} have the following form:

$$A_{q_i} = \begin{bmatrix} 0 & I_{q_i-1} \\ 0 & 0 \end{bmatrix}, \quad B_{q_i} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad C_{q_i} = [1, 0, \dots, 0] \quad (3.16)$$

Assuming that $x_i, i = 1, 2, \dots, m_d$, are arranged such that $q_i \leq q_{i+1}$, the matrix L_{id} has the particular form

$$L_{id} = [L_{i1} \quad L_{i2} \quad \dots \quad L_{ii-1} \quad 0 \quad \dots \quad 0] \quad (3.17)$$

The last row of each L_{id} is identically zero. Moreover:

1. If Σ is a continuous-time system, then

$$\lambda(A_{aa}^-) \subset \mathbb{C}^-, \quad \lambda(A_{aa}^0) \subset \mathbb{C}^0, \quad \lambda(A_{aa}^+) \subset \mathbb{C}^+ \quad (3.18)$$

2. If Σ is a discrete-time system, then

$$\lambda(A_{aa}^-) \subset \mathbb{C}^\ominus, \quad \lambda(A_{aa}^0) \subset \mathbb{C}^\circ, \quad \lambda(A_{aa}^+) \subset \mathbb{C}^\otimes \quad (3.19)$$

Also, the pair (A_{cc}, B_c) is controllable and the pair (A_{bb}, C_b) is observable. \diamond

Note that a detailed procedure of constructing the above structural decomposition can be found in Chen *et al.* [71]. Its software realization can be found in Lin *et al.* [53], which is free for downloading at <http://linearsystemskit.net>.

We can rewrite the special coordinate basis of the quadruple (A, B, C, D) given by Theorem 3.1 in a more compact form:

$$\begin{aligned} \tilde{A} &= \Gamma_s^{-1} A \Gamma_s = A_s + B_0 C_0 \\ &= \begin{bmatrix} A_{aa}^- & 0 & 0 & L_{ab}^- C_b & 0 & L_{ad}^- C_d \\ 0 & A_{aa}^0 & 0 & L_{ab}^0 C_b & 0 & L_{ad}^0 C_d \\ 0 & 0 & A_{aa}^+ & L_{ab}^+ C_b & 0 & L_{ad}^+ C_d \\ 0 & 0 & 0 & A_{bb} & 0 & L_{bd} C_d \\ B_c E_{ca}^- & B_c E_{ca}^0 & B_c E_{ca}^+ & L_{cb} C_b & A_{cc} & L_{cd} C_d \\ B_d E_{da}^- & B_d E_{da}^0 & B_d E_{da}^+ & B_d E_{db} & B_d E_{dc} & A_{dd} \end{bmatrix} \\ &\quad + \begin{bmatrix} B_{0a}^- \\ B_{0a}^0 \\ B_{0a}^+ \\ B_{0b} \\ B_{0c} \\ B_{0d} \end{bmatrix} [C_{0a}^- \quad C_{0a}^0 \quad C_{0a}^+ \quad C_{0b} \quad C_{0c} \quad C_{0d}] \quad (3.20) \end{aligned}$$

$$\tilde{B} = \Gamma_s^{-1} B \Gamma_i = [B_0 \quad B_s] = \begin{bmatrix} B_{0a}^- & 0 & 0 \\ B_{0a}^0 & 0 & 0 \\ B_{0a}^+ & 0 & 0 \\ B_{0b} & 0 & 0 \\ B_{0c} & 0 & B_c \\ B_{0d} & B_d & 0 \end{bmatrix} \quad (3.21)$$

$$\tilde{C} = \Gamma_o^{-1} C \Gamma_s = \begin{bmatrix} C_0 \\ C_s \end{bmatrix} = \begin{bmatrix} C_{0a}^- & C_{0a}^0 & C_{0a}^+ & C_{0b} & C_{0c} & C_{0d} \\ 0 & 0 & 0 & 0 & 0 & C_d \\ 0 & 0 & 0 & C_b & 0 & 0 \end{bmatrix} \quad (3.22)$$

$$\tilde{D} = \Gamma_o^{-1} D \Gamma_i = D_s = \begin{bmatrix} I_{m_0} & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (3.23)$$

3.2.1 Interpretation

A block diagram of the structural decomposition of Theorem 3.1 is illustrated in Figure 3.1. In this figure, a signal given by a double-edged arrow is some linear combination of outputs y_i , $i = 0$ to m_d , whereas a signal given by the double-edged arrow with a solid dot is some linear combination of all the states.

$$B_{0a} := \begin{bmatrix} B_{0a}^- \\ B_{0a}^0 \\ B_{0a}^+ \end{bmatrix}, \quad L_{ab} := \begin{bmatrix} L_{ab}^- \\ L_{ab}^0 \\ L_{ab}^+ \end{bmatrix}, \quad L_{ad} := \begin{bmatrix} L_{ad}^- \\ L_{ad}^0 \\ L_{ad}^+ \end{bmatrix} \quad (3.24)$$

and

$$A_{aa} := \begin{bmatrix} A_{aa}^- & 0 & 0 \\ 0 & A_{aa}^0 & 0 \\ 0 & 0 & A_{aa}^+ \end{bmatrix}, \quad E_{ca} := [E_{ca}^- \quad E_{ca}^0 \quad E_{ca}^+] \quad (3.25)$$

Also, the block \triangleright is either an integrator if Σ is of continuous-time or a backward-shifting operator if Σ is of discrete-time. We note the following intuitive points.

1. The input u_i controls the output y_i through a stack of q_i integrators (or backward-shifting operators), whereas x_i is the state associated with those integrators (or backward-shifting operators) between u_i and y_i . Moreover, (A_{q_i}, B_{q_i}) and (A_{q_i}, C_{q_i}) , respectively, form controllable and observable pairs. This implies that all the states x_i are both controllable and observable.
2. The output y_b and the state x_b are not directly influenced by any inputs; however, they could be indirectly controlled through the output y_d . Moreover, (A_{bb}, C_b) forms an observable pair. This implies that the state x_b is observable.
3. The state x_c is directly controlled by the input u_c , but it does not directly affect any output. Moreover, (A_{cc}, B_c) forms a controllable pair. This implies that the state x_c is controllable.
4. The state x_a is neither directly controlled by any input nor does it directly affect any output.

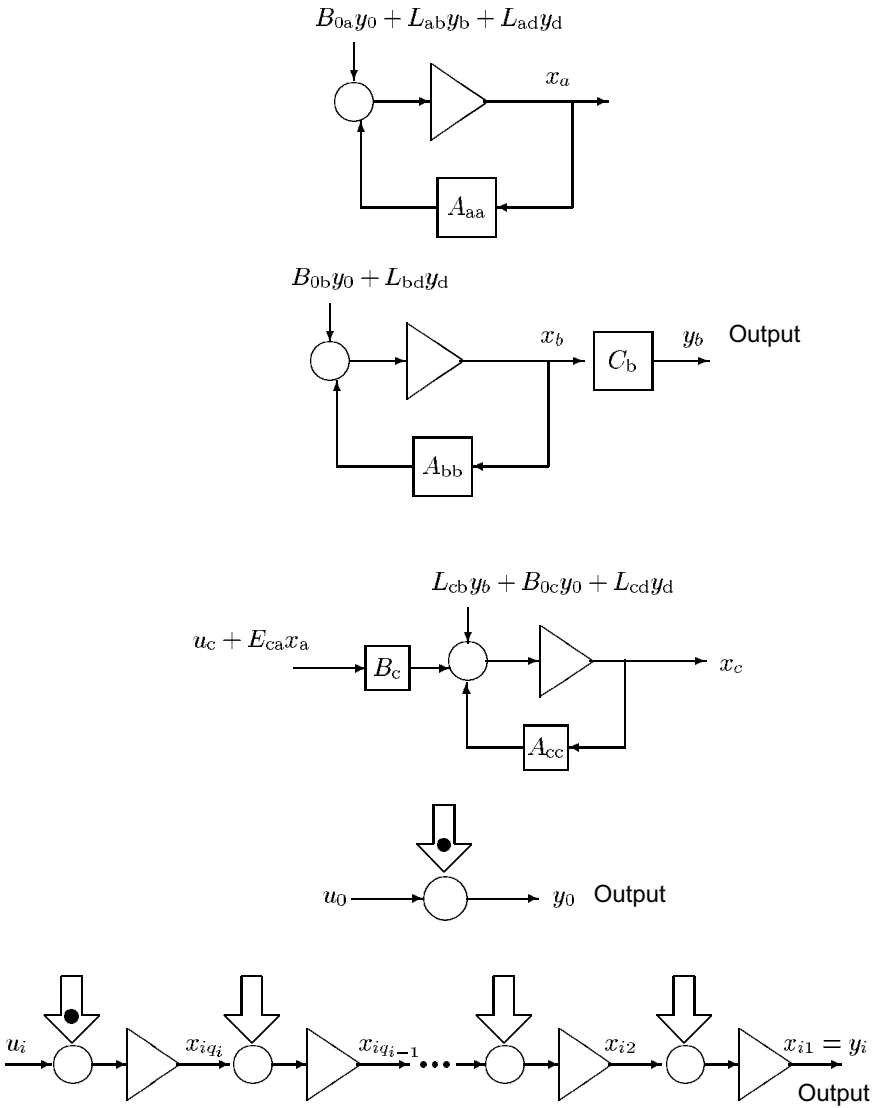


Figure 3.1. A block diagram representation of the special coordinate basis

3.2.2 Properties

In what follows, we state some important properties of the above special coordinate basis that are pertinent to our present work. As mentioned earlier, the proofs of these properties can be found in Chen *et al.* [71] and Chen [74].

Property 3.2. The given system Σ is observable (detectable) if and only if the pair $(A_{\text{obs}}, C_{\text{obs}})$ is observable (detectable), where

$$A_{\text{obs}} := \begin{bmatrix} A_{\text{aa}} & 0 \\ B_{\text{c}}E_{\text{ca}} & A_{\text{cc}} \end{bmatrix}, \quad C_{\text{obs}} := \begin{bmatrix} C_{0\text{a}} & C_{0\text{c}} \\ E_{\text{da}} & E_{\text{dc}} \end{bmatrix} \quad (3.26)$$

and where

$$E_{\text{da}} := [E_{\text{da}}^-, E_{\text{da}}^0, E_{\text{da}}^+], \quad C_{0\text{a}} := [C_{0\text{a}}^-, C_{0\text{a}}^0, C_{0\text{a}}^+] \quad (3.27)$$

Also, define

$$A_{\text{con}} := \begin{bmatrix} A_{\text{aa}} & L_{\text{ab}}C_{\text{b}} \\ 0 & A_{\text{bb}} \end{bmatrix}, \quad B_{\text{con}} := \begin{bmatrix} B_{0\text{a}} & L_{\text{ad}} \\ B_{0\text{b}} & L_{\text{bd}} \end{bmatrix} \quad (3.28)$$

Similarly, Σ is controllable (stabilizable) if and only if the pair $(A_{\text{con}}, B_{\text{con}})$ is controllable (stabilizable). \diamond

The invariant zeros of a system Σ characterized by (A, B, C, D) can be defined via the Smith canonical form of the (Rosenbrock) system matrix [75] of Σ :

$$P_{\Sigma}(\zeta) := \begin{bmatrix} \zeta I - A & -B \\ C & D \end{bmatrix} \quad (3.29)$$

We have the following definition for the invariant zeros (see also [76]).

Definition 3.3. (Invariant Zeros). A complex scalar $\alpha \in \mathbb{C}$ is said to be an invariant zero of Σ if

$$\text{rank} \{P_{\Sigma}(\alpha)\} < n + \text{normrank} \{H(\zeta)\} \quad (3.30)$$

where $\text{normrank} \{H(\zeta)\}$ denotes the normal rank of $H(\zeta)$, which is defined as its rank over the field of rational functions of ζ with real coefficients. \diamond

The special coordinate basis of Theorem 3.1 shows explicitly the invariant zeros and the normal rank of Σ . To be more specific, we have the following properties.

Property 3.4.

1. The normal rank of $H(\zeta)$ is equal to $m_0 + m_{\text{d}}$.
2. Invariant zeros of Σ are the eigenvalues of A_{aa} , which are the unions of the eigenvalues of $A_{\text{aa}}^-, A_{\text{aa}}^0$ and A_{aa}^+ . Moreover, the given system Σ is of minimum phase if and only if A_{aa} has only stable eigenvalues, marginal minimum phase if and only if A_{aa} has no unstable eigenvalue but has at least one marginally stable eigenvalue, and nonminimum phase if and only if A_{aa} has at least one unstable eigenvalue. \diamond

The special coordinate basis can also reveal the infinite zero structure of Σ . We note that the infinite zero structure of Σ can be either defined in association with root-locus theory or as Smith–McMillan zeros of the transfer function at infinity. For the sake of simplicity, we only consider the infinite zeros from the point of view of Smith–McMillan theory here. To define the zero structure of $H(\zeta)$ at infinity, one can use the familiar Smith–McMillan description of the zero structure at finite frequencies of a general not necessarily square but strictly proper transfer function matrix $H(\zeta)$. Namely, a rational matrix $H(\zeta)$ possesses an infinite zero of order k when $H(1/z)$ has a finite zero of precisely that order at $z = 0$ (see [75], [77–79]). The number of zeros at infinity, together with their orders, indeed defines an infinite zero structure. Owens [80] related the orders of the infinite zeros of the root-loci of a square system with a nonsingular transfer function matrix to the \mathcal{C}^* structural invariant indices list \mathcal{I}_4 of Morse [81]. This connection reveals that, even for general not necessarily strictly proper systems, the *structure at infinity is in fact the topology of inherent integrations between the input and the output variables*. The special coordinate basis of Theorem 3.1 explicitly shows this topology of inherent integrations. The following property pinpoints this.

Property 3.5. Σ has $m_0 = \text{rank}(D)$ infinite zeros of order 0. The infinite zero structure (of order greater than 0) of Σ is given by

$$S_{\infty}^*(\Sigma) = \{q_1, q_2, \dots, q_{m_a}\} \quad (3.31)$$

That is, each q_i corresponds to an infinite zero of Σ of order q_i . Note that for an SISO system Σ , we have $S_{\infty}^*(\Sigma) = \{q_1\}$, where q_1 is the *relative degree* of Σ . \diamond

The special coordinate basis can also exhibit the invertibility structure of a given system Σ . The formal definitions of right invertibility and left invertibility of a linear system can be found in [82]. Basically, for the usual case when $[B' \ D']$ and $[C \ D]$ are of maximal rank, the system Σ , or equivalently $H(\zeta)$, is said to be left invertible if there exists a rational matrix function, say $L(\zeta)$, such that

$$L(\zeta)H(\zeta) = I_m \quad (3.32)$$

Σ or $H(\zeta)$ is said to be right invertible if there exists a rational matrix function, say $R(\zeta)$, such that

$$H(\zeta)R(\zeta) = I_p \quad (3.33)$$

Σ is invertible if it is both left and right invertible, and Σ is degenerate if it is neither left nor right invertible.

Property 3.6. The given system Σ is right invertible if and only if x_b (and hence y_b) are nonexistent, left invertible if and only if x_c (and hence u_c) are nonexistent, and invertible if and only if both x_b and x_c are nonexistent. Moreover, Σ is degenerate if and only if both x_b and x_c are present. \diamond

By now it is clear that the special coordinate basis decomposes the state space into several distinct parts. In fact, the state-space \mathcal{X} is decomposed as

$$\mathcal{X} = \mathcal{X}_a^- \oplus \mathcal{X}_a^0 \oplus \mathcal{X}_a^+ \oplus \mathcal{X}_b \oplus \mathcal{X}_c \oplus \mathcal{X}_d \quad (3.34)$$

Here, \mathcal{X}_a^- is related to the stable invariant zeros, *i.e.* the eigenvalues of A_{aa}^- are the stable invariant zeros of Σ . Similarly, \mathcal{X}_a^0 and \mathcal{X}_a^+ are, respectively, related to the invariant zeros of Σ located in the marginally stable and unstable regions. On the other hand, \mathcal{X}_b is related to the right invertibility, *i.e.* the system is right invertible if and only if $\mathcal{X}_b = \{0\}$, whereas \mathcal{X}_c is related to left invertibility, *i.e.* the system is left invertible if and only if $\mathcal{X}_c = \{0\}$. Finally, \mathcal{X}_d is related to zeros of Σ at infinity.

There are interconnections between the special coordinate basis and various invariant geometric subspaces. To show these interconnections, we introduce the following geometric subspaces.

Definition 3.7. (Geometric Subspaces \mathcal{V}^x and \mathcal{S}^x). *The weakly unobservable subspaces of Σ , \mathcal{V}^x , and the strongly controllable subspaces of Σ , \mathcal{S}^x , are defined as follows:*

1. $\mathcal{V}^x(\Sigma)$ is the maximal subspace of \mathbb{R}^n that is $(A+BF)$ -invariant and contained in $\text{Ker}(C + DF)$ such that the eigenvalues of $(A + BF)|_{\mathcal{V}^x}$ are contained in $\mathbb{C}^x \subseteq \mathbb{C}$ for some constant matrix F .
2. $\mathcal{S}^x(\Sigma)$ is the minimal $(A + KC)$ -invariant subspace of \mathbb{R}^n containing the subspace $\text{Im}(B + KD)$ such that the eigenvalues of the map that is induced by $(A + KC)$ on the factor space $\mathbb{R}^n / \mathcal{S}^x$ are contained in $\mathbb{C}^x \subseteq \mathbb{C}$ for some constant matrix K .

Moreover, we let $\mathcal{V}^- = \mathcal{V}^x$ and $\mathcal{S}^- = \mathcal{S}^x$, if $\mathbb{C}^x = \mathbb{C}^- \cup \mathbb{C}^0$; $\mathcal{V}^+ = \mathcal{V}^x$ and $\mathcal{S}^+ = \mathcal{S}^x$, if $\mathbb{C}^x = \mathbb{C}^+$; $\mathcal{V}^\ominus = \mathcal{V}^x$ and $\mathcal{S}^\ominus = \mathcal{S}^x$, if $\mathbb{C}^x = \mathbb{C}^\ominus \cup \mathbb{C}^0$; $\mathcal{V}^\otimes = \mathcal{V}^x$ and $\mathcal{S}^\otimes = \mathcal{S}^x$, if $\mathbb{C}^x = \mathbb{C}^\otimes$; and finally $\mathcal{V}^* = \mathcal{V}^x$ and $\mathcal{S}^* = \mathcal{S}^x$, if $\mathbb{C}^x = \mathbb{C}$. \diamond

We have the following property.

Property 3.8.

1. $\mathcal{X}_a^- \oplus \mathcal{X}_a^0 \oplus \mathcal{X}_c$ spans $\begin{cases} \mathcal{V}^-(\Sigma), & \text{if } \Sigma \text{ is of continuous-time,} \\ \mathcal{V}^\ominus(\Sigma), & \text{if } \Sigma \text{ is of discrete-time.} \end{cases}$
2. $\mathcal{X}_a^+ \oplus \mathcal{X}_c$ spans $\begin{cases} \mathcal{V}^+(\Sigma), & \text{if } \Sigma \text{ is of continuous-time,} \\ \mathcal{V}^\otimes(\Sigma), & \text{if } \Sigma \text{ is of discrete-time.} \end{cases}$
3. $\mathcal{X}_a^- \oplus \mathcal{X}_a^0 \oplus \mathcal{X}_a^+ \oplus \mathcal{X}_c$ spans $\mathcal{V}^*(\Sigma)$.
4. $\mathcal{X}_a^+ \oplus \mathcal{X}_c \oplus \mathcal{X}_d$ spans $\begin{cases} \mathcal{S}^-(\Sigma), & \text{if } \Sigma \text{ is of continuous-time,} \\ \mathcal{S}^\ominus(\Sigma), & \text{if } \Sigma \text{ is of discrete-time.} \end{cases}$
5. $\mathcal{X}_a^- \oplus \mathcal{X}_a^0 \oplus \mathcal{X}_c \oplus \mathcal{X}_d$ spans $\begin{cases} \mathcal{S}^+(\Sigma), & \text{if } \Sigma \text{ is of continuous-time,} \\ \mathcal{S}^\otimes(\Sigma), & \text{if } \Sigma \text{ is of discrete-time.} \end{cases}$
6. $\mathcal{X}_c \oplus \mathcal{X}_d$ spans $\mathcal{S}^*(\Sigma)$. \diamond

Finally, for future development on deriving solvability conditions for H_∞ almost disturbance decoupling problems, we introduce two more subspaces of Σ . The original definitions of these subspaces were given by Scherer [83].

Definition 3.9. (Geometric Subspaces \mathcal{V}_λ and \mathcal{S}_λ). For any $\lambda \in \mathbb{C}$, we define

$$\mathcal{V}_\lambda(\Sigma) := \left\{ \zeta \in \mathbb{C}^n \mid \exists \omega \in \mathbb{C}^m : 0 = \begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix} \begin{pmatrix} \zeta \\ \omega \end{pmatrix} \right\} \quad (3.35)$$

and

$$\mathcal{S}_\lambda(\Sigma) := \left\{ \zeta \in \mathbb{C}^n \mid \exists \omega \in \mathbb{C}^{n+m} : \begin{pmatrix} \zeta \\ 0 \end{pmatrix} = \begin{bmatrix} A - \lambda I & B \\ C & D \end{bmatrix} \omega \right\} \quad (3.36)$$

$\mathcal{V}_\lambda(\Sigma)$ and $\mathcal{S}_\lambda(\Sigma)$ are associated with the so-called state zero directions of Σ if λ is an invariant zero of Σ . \diamond

These subspaces $\mathcal{S}_\lambda(\Sigma)$ and $\mathcal{V}_\lambda(\Sigma)$ can also be easily obtained using the special coordinate basis. We have the following new property of the special coordinate basis.

Property 3.10.

$$\mathcal{S}_\lambda(\Sigma) = \text{Im} \left\{ \Gamma_s \begin{bmatrix} \lambda I - A_{aa} & 0 & 0 & 0 \\ 0 & Y_{b\lambda} & 0 & 0 \\ 0 & 0 & I_{nc} & 0 \\ 0 & 0 & 0 & I_{nd} \end{bmatrix} \right\} \quad (3.37)$$

where

$$\text{Im} \{Y_{b\lambda}\} = \text{Ker} [C_b(A_{bb} + K_b C_b - \lambda I)^{-1}] \quad (3.38)$$

and where K_b is any appropriately dimensional matrix subject to the constraint that $A_{bb} + K_b C_b$ has no eigenvalue at λ . We note that such a K_b always exists, as (A_{bb}, C_b) is completely observable.

$$\mathcal{V}_\lambda(\Sigma) = \text{Im} \left\{ \Gamma_s \begin{bmatrix} X_{a\lambda} & 0 \\ 0 & 0 \\ 0 & X_{c\lambda} \\ 0 & 0 \end{bmatrix} \right\} \quad (3.39)$$

where $X_{a\lambda}$ is a matrix whose columns form a basis for the subspace,

$$\left\{ \zeta_a \in \mathbb{C}^{n_a} \mid (\lambda I - A_{aa})\zeta_a = 0 \right\} \quad (3.40)$$

and

$$X_{c\lambda} := (A_{cc} + B_c F_c - \lambda I)^{-1} B_c \quad (3.41)$$

with F_c being any appropriately dimensional matrix subject to the constraint that $A_{cc} + B_c F_c$ has no eigenvalue at λ . Again, we note that the existence of such an F_c is guaranteed by the controllability of (A_{cc}, B_c) . \diamond

Clearly, if $\lambda \notin \lambda(A_{aa})$, then we have $\mathcal{V}_\lambda(\Sigma) \subseteq \mathcal{V}^*(\Sigma)$ and $\mathcal{S}_\lambda(\Sigma) \supseteq \mathcal{S}^*(\Sigma)$. It is interesting to note that the subspaces $\mathcal{V}^*(\Sigma)$ and $\mathcal{S}^*(\Sigma)$ are dual in the sense that $\mathcal{V}^*(\Sigma^*) = \mathcal{S}^*(\Sigma)^\perp$, where Σ^* is characterized by the quadruple (A', C', B', D') . Also, $\mathcal{S}_\lambda(\Sigma) = \mathcal{V}_\lambda(\Sigma^*)^\perp$.

3.3 PID Control

PID control is the most popular technique used in industry because it is relatively easy and simple to design and implement. Most importantly, it works in most practical situations, although its performance is somewhat limited owing to its restricted structure. Nevertheless, in what follows, we recall this well-known classical control system design methodology for ease of reference.

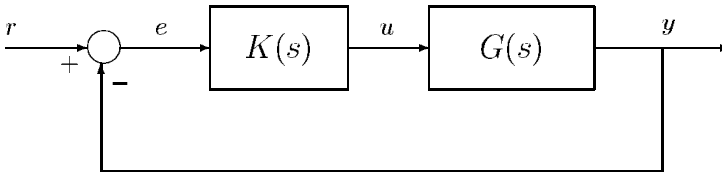


Figure 3.2. The typical PID control configuration

To be more specific, we consider the control system as depicted in Figure 3.2, in which $G(s)$ is the plant to be controlled and $K(s)$ is the PID controller characterized by the following transfer function

$$K(s) = K_p \left(1 + \frac{1}{T_i s} + T_d s \right) \quad (3.42)$$

The control system design is then to determine the parameters K_p , T_i and T_d such that the resulting closed-loop system yields a certain desired performance, *i.e.* it meets certain prescribed design specifications.

3.3.1 Selection of Design Parameters

Ziegler–Nichols tuning is one of the most common techniques used in practical situations to design an appropriate PID controller for the class of systems that can be exactly modeled as, or approximated by, the following first-order system:

$$G(s) = \frac{Y(s)}{U(s)} = \frac{K}{\tau s + 1} e^{-t_d s} \quad (3.43)$$

One of the methods proposed by Ziegler and Nichols ([84, 85]) is first to replace the controller $K(s)$ in Figure 3.2 by a simple proportional gain. We then increase this proportional gain to a value, say K_u , for which we observe continuous oscillations in its step response, *i.e.* the system becomes marginally stable. Assume that the corresponding oscillating frequency is ω_u . The PID controller parameters are then given as follows:

$$K_p = \frac{3K_u}{5}, \quad T_i = \frac{\pi}{\omega_u}, \quad T_d = \frac{\pi}{4\omega_u} \quad (3.44)$$

Experience has shown that such controller settings provide a good closed-loop response for many systems. Unfortunately, it will be seen shortly in the coming chapters that the typical model of a VCM actuator is actually a double integrator and thus Ziegler–Nichols tuning cannot be directly applied to design a servo system for the VCM actuator.

Another common way to design a PID controller is the pole assignment method, in which the parameters K_p , T_i and T_d are chosen such that the dominant roots of the closed-loop characteristic equation, *i.e.*

$$1 + K(s)G(s) = 0 \quad (3.45)$$

are assigned to meet certain desired specifications (such as overshoot, rise time, settling time, *etc.*), while its remaining roots are placed far away to the left on the complex plane (roughly three to four times faster compared with the dominant roots). The detailed procedure of this method can be found in most classical control engineering texts (see, *e.g.*, [86]). For the PID control of discrete-time systems, interested readers are referred to [1] for more information.

3.3.2 Sensitivity Functions

System stability margins such as gain margin and phase margin are also very important factors in designing control systems. These stability margins can be obtained from either the well-known Bode plot or Nyquist plot of the open-loop system, *i.e.* $K(s)G(s)$. For an HDD servo system with a large number of resonance modes, its Bode plot might have more than one gain and/or phase crossover frequencies. Thus, it would be necessary to double check these margins using its Nyquist plot. Sensitivity function and complementary sensitivity function are two other measures for a good control system design. The sensitivity function is defined as the closed-loop transfer function from the reference signal, r , to the tracking error, e , and is given by

$$S(s) = \frac{1}{1 + K(s)G(s)} \quad (3.46)$$

The complementary sensitivity function is defined as the closed-loop transfer function between the reference, r , and the system output, y , *i.e.*

$$T(s) = \frac{K(s)G(s)}{1 + K(s)G(s)} \quad (3.47)$$

Clearly, we have $S(s) + T(s) \equiv 1$. A good design should have a sensitivity function that is small at low frequencies for good tracking performance and disturbance rejection and is equal to unity at high frequencies. On the other hand, the complementary sensitivity function should be made unity at low frequencies. It must roll off at high frequencies to possess good attenuation of high-frequency noise.

Note that for a two-degrees-of-freedom control system with a precompensator in the feedforward path right after the reference signal (see, for example, Figure

3.3), the sensitivity and complementary sensitivity functions still remain the same as those in Equations 3.46 and 3.47, which represent, respectively, the closed-loop transfer function from the disturbance at the system output point, if any, to the system output, and the closed-loop transfer function from the measurement noise, if any, to the system output. Thus, a feedforward precompensator does not cause changes in the sensitivity and complementary sensitivity functions. It does, however, help in improving the system tracking performance.

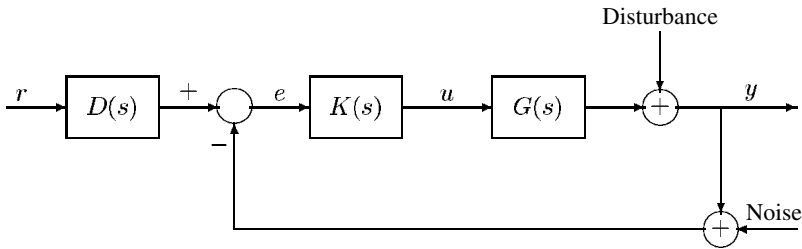


Figure 3.3. A two-degrees-of-freedom control system

3.4 H_2 Optimal Control

Most of the feedback design tools provided by the classical Nyquist–Bode frequency-domain theory are restricted to single-feedback-loop designs. Modern multivariable control theory based on state-space concepts has the capability to deal with multiple feedback-loop designs, and as such has emerged as an alternative to the classical Nyquist–Bode theory. Although it does have shortcomings of its own, a great asset of modern control theory utilizing the state-space description of systems is that the design methods derived from it are easily amenable to computer implementation. Owing to this, rapid progress has been made during the last two or three decades in developing a number of multivariable analysis and design tools using the state-space description of systems. One of the foremost and most powerful design tools developed in this connection is based on what is called linear quadratic Gaussian (LQG) control theory. Here, given a linear model of the plant in a state-space description, and assuming that the disturbance and measurement noise are Gaussian stochastic processes with known power spectral densities, the designer translates the design specifications into a quadratic performance criterion consisting of some state variables and control signal inputs. The object of design then is to minimize the performance criterion by using appropriate state or measurement feedback controllers while guaranteeing the closed-loop stability. A ubiquitous architecture for a measurement feedback controller has been observer based, wherein a state feedback control law is implemented by utilizing an estimate of the state. Thus, the design of a measurement feedback controller here is worked out in two stages. In the first stage, an

optimal internally stabilizing static state feedback controller is designed, and in the second stage a state estimator is designed. The estimator, otherwise called an observer or filter, is traditionally designed to yield the least mean square error estimate of the state of the plant, utilizing only the measured output, which is often assumed to be corrupted by an additive white Gaussian noise. The LQG control problem as described above is posed in a stochastic setting. The same can be posed in a deterministic setting, known as an H_2 optimal control problem, in which the H_2 norm of a certain transfer function from an exogenous disturbance to a pertinent controlled output of a given plant is minimized by appropriate use of an internally stabilizing controller.

Much research effort has been expended in the area of H_2 optimal control or optimal control in general during the last few decades (see, *e.g.*, Anderson and Moore [87], Fleming and Rishel [88], Kwakernaak and Sivan [89], and Saberi *et al.* [90], and references cited therein). In what follows, we focus mainly on the formulation and solution to both continuous- and discrete-time H_2 optimal control problems. Interested readers are referred to [90] for more detailed treatments of such problems.

3.4.1 Continuous-time Systems

We consider a generalized system Σ with a state-space description,

$$\Sigma : \begin{cases} \dot{x} = A x + B u + E w \\ y = C_1 x + D_{11} u + D_1 w \\ h = C_2 x + D_2 u + D_{22} w \end{cases} \quad (3.48)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, $w \in \mathbb{R}^q$ is the external disturbance input, $y \in \mathbb{R}^p$ is the measurement output, and $h \in \mathbb{R}^\ell$ is the controlled output of Σ . For the sake of simplicity in future development, throughout this chapter, we let Σ_p be the subsystem characterized by the matrix quadruple (A, B, C_2, D_2) and Σ_q be the subsystem characterized by (A, E, C_1, D_1) . Throughout this section, we assume that (A, B) is stabilizable and (A, C_1) is detectable.

Generally, we can assume that matrix D_{11} in Equation 3.48 is zero. This can be justified as follows: If $D_{11} \neq 0$, we define a new measurement output

$$y_{\text{new}} = y - D_{11}u = C_1x + D_1w \quad (3.49)$$

that does not have a direct feedthrough term from u . Suppose we carry on our control system design using this new measurement output to obtain a proper control law, say, $u = K(s)y_{\text{new}}$. Then, it is straightforward to verify that this control law is equivalent to the following one

$$u = [I + K(s)D_{11}]^{-1}K(s)y \quad (3.50)$$

provided that $[I + K(s)D_{11}]^{-1}$ is well posed, *i.e.* the inverse exists for almost all $s \in \mathbb{C}$. Thus, for simplicity, we assume that $D_{11} = 0$.

The standard H_2 optimal control problem is to find an internally stabilizing proper measurement feedback control law,

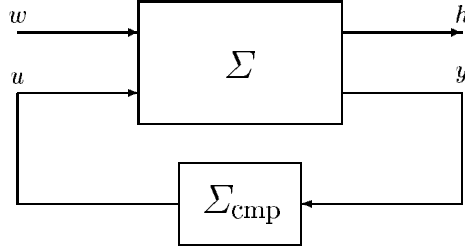


Figure 3.4. The typical control configuration in state-space setting

$$\Sigma_{\text{cmp}} : \begin{cases} \dot{v} = A_{\text{cmp}} v + B_{\text{cmp}} y \\ u = C_{\text{cmp}} v + D_{\text{cmp}} y \end{cases} \quad (3.51)$$

such that the H_2 -norm of the overall closed-loop transfer matrix function from w to h is minimized (see also Figure 3.4). To be more specific, we will say that the control law Σ_{cmp} of Equation 3.51 is internally stabilizing when applied to the system Σ of Equation 3.48, if the following matrix is asymptotically stable:

$$A_{\text{cl}} := \begin{bmatrix} A + BD_{\text{cmp}}C_1 & BC_{\text{cmp}} \\ B_{\text{cmp}}C_1 & A_{\text{cmp}} \end{bmatrix} \quad (3.52)$$

i.e. all its eigenvalues lie in the open left-half complex plane. It is straightforward to verify that the closed-loop transfer matrix from the disturbance w to the controlled output h is given by

$$T_{hw}(s) = C_e(sI - A_e)^{-1}B_e + D_e \quad (3.53)$$

where

$$\left. \begin{aligned} A_e &:= \begin{bmatrix} A + BD_{\text{cmp}}C_1 & BC_{\text{cmp}} \\ B_{\text{cmp}}C_1 & A_{\text{cmp}} \end{bmatrix} \\ B_e &:= \begin{bmatrix} E + BD_{\text{cmp}}D_1 \\ B_{\text{cmp}}D_1 \end{bmatrix} \\ C_e &:= [C_2 + D_2D_{\text{cmp}}C_1 \quad D_2C_{\text{cmp}}] \\ D_e &:= D_2D_{\text{cmp}}D_1 + D_{22} \end{aligned} \right\} \quad (3.54)$$

It is simple to note that if Σ_{cmp} is a static state feedback law, *i.e.* $u = Fx$, then the closed-loop transfer matrix from w to h is given by

$$T_{hw}(s) = (C_2 + D_2F)(sI - A - BF)^{-1}E + D_{22} \quad (3.55)$$

The H_2 -norm of a stable continuous-time transfer matrix, *e.g.*, $T_{hw}(s)$, is defined as follows:

$$\|T_{hw}\|_2 := \left(\frac{1}{2\pi} \text{trace} \left[\int_{-\infty}^{\infty} T_{hw}(j\omega) T_{hw}(j\omega)^H d\omega \right] \right)^{1/2} \quad (3.56)$$

By Parseval's theorem, $\|T_{hw}\|_2$ can equivalently be defined as

$$\|T_{hw}\|_2 = \left(\text{trace} \left[\int_0^\infty g(t)g(t)' dt \right] \right)^{1/2} \quad (3.57)$$

where $g(t)$ is the unit impulse response of $T_{hw}(s)$. Thus, $\|T_{hw}\|_2 = \|g\|_2$.

The H_2 optimal control is to design a proper controller Σ_{cmp} such that, when it is applied to the plant Σ , the resulting closed loop is asymptotically stable and the H_2 -norm of $T_{hw}(s)$ is minimized. For future use, we define

$$\gamma_2^* := \inf \left\{ \|T_{hw}(\Sigma \times \Sigma_{\text{cmp}})\|_2 \mid \Sigma_{\text{cmp}} \text{ internally stabilizes } \Sigma \right\} \quad (3.58)$$

Furthermore, a control law Σ_{cmp} is said to be an H_2 optimal controller for Σ if Equation 3.48 if its resulting closed-loop transfer function from w to h has an H_2 -norm equal to γ_2^* , i.e. $\|T_{hw}\|_2 = \gamma_2^*$.

It is clear to see from the definition of the H_2 -norm that, in order to have a finite $\|T_{hw}\|_2$, the following must be satisfied:

$$D_e = D_2 D_{\text{cmp}} D_1 + D_{22} = 0 \quad (3.59)$$

which is equivalent to the existence of a static measurement prefeedback law $u = Sy + u_{\text{new}}$ to the system in Equation 3.48 such that $D_2 S D_1 + D_{22} = 0$. We note that the minimization of $\|T_{hw}\|_2$ is meaningful only when it is finite. As such, it is without loss of any generality to assume that the feedforward matrix $D_{22} = 0$ hereafter in this section. In fact, in this case, $\|T_{hw}\|_2$ can be easily obtained. Solving either one of the following Lyapunov equations:

$$A_e' P_e + P_e A_e + C_e' C_e = 0, \quad A_e Q_e + Q_e A_e' + B_e B_e' = 0 \quad (3.60)$$

for P_e or Q_e , then the H_2 -norm of $T_{hw}(s)$ can be computed by

$$\|T_{hw}\|_2 = \sqrt{\text{trace} [B_e' P_e B_e]} = \sqrt{\text{trace} [C_e Q_e C_e']} \quad (3.61)$$

In what follows, we present solutions to the problem without detailed proofs. We start first with the simplest case, when the given system Σ satisfies the following assumptions of the so-called *regular case*:

1. Σ_p has no invariant zeros on the imaginary axis and D_2 is of maximal column rank.
2. Σ_Q has no invariant zeros on the imaginary axis and D_1 is of maximal row rank.

The problem is called the *singular case* if Σ does not satisfy these conditions.

The solution to the regular case of the H_2 optimal control problem is very simple. The optimal controller is given by (see, e.g., [91]),

$$\Sigma_{\text{cmp}} : \begin{cases} \dot{v} = (A + BF + KC_1) v - K y \\ u = F v \end{cases} \quad (3.62)$$

where

$$F = -(D_2' D_2)^{-1} (D_2' C_2 + B' P) \quad (3.63)$$

$$K = -(QC_1' + ED_1') (D_1 D_1')^{-1} \quad (3.64)$$

and where $P = P' \geq 0$ and $Q = Q' \geq 0$ are, respectively, the stabilizing solutions of the following Riccati equations:

$$A'P + PA + C_2' C_2 - (PB + C_2' D_2) (D_2' D_2)^{-1} (D_2' C_2 + B' P) = 0 \quad (3.65)$$

$$QA' + AQ + EE' - (QC_1' + ED_1') (D_1 D_1')^{-1} (D_1 E' + C_1 Q) = 0 \quad (3.66)$$

Moreover, the optimal value γ_2^* can be computed as follows:

$$\gamma_2^* = \left\{ \text{trace} [E' P E] + \text{trace} [(A' P + PA + C_2' C_2) Q] \right\}^{1/2} \quad (3.67)$$

We note that if all the states of Σ are available for feedback, then the optimal controller is reduced to a static law $u = Fx$ with F being given as in Equation 3.63.

Next, we present two methods that solve the singular H_2 optimal control problem. As a matter of fact, in the singular case, it is in general infeasible to obtain an optimal controller, although it is possible under certain restricted conditions (see, e.g., [90, 92]). The solutions to the singular case are generally suboptimal, and usually parameterized by a certain tuning parameter, say ε . A controller parameterized by ε is said to be suboptimal if there exists an $\varepsilon^* > 0$ such that for all $\varepsilon \leq \varepsilon^*$ the closed-loop system comprising the given plant and the controller is asymptotically stable, and the resulting closed-loop transfer function from w to h , which is obviously a function of ε , has an H_2 -norm arbitrarily close to γ_2^* as ε tends to 0.

The following is a so-called *perturbation approach* (see, e.g., [93]) that would yield a suboptimal controller for the general singular case. We note that such an approach is numerically unstable. The problem becomes very serious when the given system is ill-conditioned or has multiple time scales. In principle, the desired solution can be obtained by introducing some small perturbations to the matrices E , D_1 , C_2 and D_2 , i.e.

$$\tilde{E} := [E \quad \varepsilon I \quad 0], \quad \tilde{D}_1 := [D_1 \quad 0 \quad \varepsilon I] \quad (3.68)$$

and

$$\tilde{C}_2 := \begin{bmatrix} C_2 \\ \varepsilon I \\ 0 \end{bmatrix}, \quad \tilde{D}_2 := \begin{bmatrix} D_2 \\ 0 \\ \varepsilon I \end{bmatrix} \quad (3.69)$$

A full-order H_2 suboptimal output feedback controller is given by

$$\tilde{\Sigma}_{\text{cmp}}(\varepsilon) : \begin{cases} \dot{v} = A + B\tilde{F} + \tilde{K}C_1 v - \tilde{K} y \\ u = \tilde{F} v \end{cases} \quad (3.70)$$

where

$$\tilde{F} = -(\tilde{D}'_2 \tilde{D}_2)^{-1}(\tilde{D}'_2 \tilde{C}_2 + B' \tilde{P}) \quad (3.71)$$

$$\tilde{K} = -(\tilde{Q} C'_1 + \tilde{E} \tilde{D}'_1)(\tilde{D}_1 \tilde{D}'_1)^{-1} \quad (3.72)$$

and where $\tilde{P} = \tilde{P}' > 0$ and $\tilde{Q} = \tilde{Q}' > 0$ are respectively the solutions of the following Riccati equations:

$$A' \tilde{P} + \tilde{P} A + \tilde{C}'_2 \tilde{C}_2 - (\tilde{P} B + \tilde{C}'_2 \tilde{D}_2)(\tilde{D}'_2 \tilde{D}_2)^{-1}(\tilde{D}'_2 \tilde{C}_2 + B' \tilde{P}) = 0 \quad (3.73)$$

$$\tilde{Q} A' + A \tilde{Q} + \tilde{E} \tilde{E}' - (\tilde{Q} C'_1 + \tilde{E} \tilde{D}'_1)(\tilde{D}_1 \tilde{D}'_1)^{-1}(\tilde{D}_1 \tilde{E}' + C_1 \tilde{Q}) = 0 \quad (3.74)$$

Alternatively, one could solve the singular case by using numerically stable algorithms (see, *e.g.*, [90]) that are based on a careful examination of the structural properties of the given system. We separate the problem into three distinct situations: 1) the state feedback case, 2) the full-order measurement feedback case, and 3) the reduced-order measurement feedback case. The software realization of these algorithms in MATLAB[®] can be found in [53]. For simplicity, we assume throughout the rest of this subsection that both subsystems Σ_p and Σ_q *have no invariant zeros on the imaginary axis*. We believe that such a condition is always satisfied for most HDD servo systems. However, most servo systems can be represented as certain chains of integrators and thus could not be formulated as a regular problem without adding dummy terms. Nevertheless, interested readers are referred to the monograph [90] for the complete treatment of H_2 optimal control using the approach given below.

i. State Feedback Case. For the case when $y = x$ in the given system Σ of Equation 3.48, *i.e.* all the state variables of Σ are available for feedback, we have the following step-by-step algorithm that constructs an H_2 suboptimal static feedback control law $u = F(\varepsilon)x$ for Σ .

STEP 3.4.C.S.1: transform the system Σ_p into the special coordinate basis as given by Theorem 3.1. To all submatrices and transformations in the special coordinate basis of Σ_p , we append the subscript p to signify their relation to the system Σ_p . We also choose the output transformation Γ_{op} to have the following form:

$$\Gamma_{op} = \begin{bmatrix} I_{m_{op}} & 0 \\ 0 & \Gamma_{orp} \end{bmatrix} \quad (3.75)$$

where $m_{op} = \text{rank}(D_2)$. Next, define

$$A_{11p} = \begin{bmatrix} A_{aap}^+ & L_{abp}^+ C_{bp} \\ 0 & A_{bbp} \end{bmatrix}, \quad B_{11p} = \begin{bmatrix} B_{0ap}^+ \\ B_{0bp} \end{bmatrix}, \quad A_{13p} = \begin{bmatrix} L_{adp}^+ \\ L_{bdp} \end{bmatrix} \quad (3.76)$$

$$C_{21p} = \Gamma_{orp} \begin{bmatrix} 0 & 0 \\ 0 & C_{bp} \end{bmatrix}, \quad C_{23p} = \Gamma_{orp} \begin{bmatrix} C_{dcp} C'_{dp} \\ 0 \end{bmatrix} \quad (3.77)$$

$$A_{xp} = A_{11p} - A_{13p}(C'_{23p} C_{23p})^{-1} C'_{23p} C_{21p} \quad (3.78)$$

$$B_{xp} B'_{xp} = B_{11p} B'_{11p} + A_{13p}(C'_{23p} C_{23p})^{-1} A'_{13p} \quad (3.79)$$

$$C'_{xp} C_{xp} = C'_{21p} C_{21p} - C'_{21p} C_{21p}(C'_{23p} C_{23p})^{-1} C'_{23p} C_{21p} \quad (3.80)$$

STEP 3.4.C.S.2: solve the following algebraic matrix Riccati equation:

$$P_x A_{xp} + A'_{xp} P_x - P_x B_{xp} B'_{xp} P_x + C'_{xp} C_{xp} = 0 \quad (3.81)$$

for $P_x > 0$ and define

$$F_{11} := \begin{bmatrix} F_{a0}^+ & F_{b0} \\ F_{a1}^+ & F_{b1} \end{bmatrix} = \begin{bmatrix} B'_{11p} P_x \\ (C'_{23p} C_{23p})^{-1} (A'_{13p} P_x + C'_{23p} C_{21p}) \end{bmatrix} \quad (3.82)$$

Then, partition $[F_{a1}^+ \ F_{b1}]$ as

$$[F_{a1}^+ \ F_{b1}] = \begin{bmatrix} F_{a11}^+ & F_{b11} \\ F_{a12}^+ & F_{b12} \\ \vdots & \vdots \\ F_{a1m_{dp}}^+ & F_{b1m_{dp}} \end{bmatrix} \quad (3.83)$$

where F_{a1i}^+ and F_{b1i} are of dimensions $1 \times n_{ap}^+$ and $1 \times n_{bp}$, respectively.

STEP 3.4.C.S.3: let Δ_{cp} be any arbitrary $m_{cp} \times n_{cp}$ matrix subject to the constraint that

$$A_{ccp}^c = A_{ccp} - B_{cp} \Delta_{cp} \quad (3.84)$$

is a stable matrix. Note that the existence of such a Δ_{cp} is guaranteed by the property that (A_{ccp}, B_{cp}) is controllable.

STEP 3.4.C.S.4: this step makes use of subsystems, $i = 1$ to m_{dp} , represented by Equation 3.14. Let $\Lambda_i = \{ \lambda_{i1}, \lambda_{i2}, \dots, \lambda_{iq_i} \}$, $i = 1$ to m_{dp} , be the sets of q_i elements all in \mathbb{C}^- , which are closed under complex conjugation, where q_i and m_{dp} are as defined in Theorem 3.1 but associated with the special coordinate basis of Σ_p . Let $\Lambda_{dp} := \Lambda_1 \cup \Lambda_2 \cup \dots \cup \Lambda_{m_{dp}}$. For $i = 1$ to m_{dp} , we define

$$p_i(s) := \prod_{j=1}^{q_i} (s - \lambda_{ij}) = s^{q_i} + F_{i1} s^{q_i-1} + \dots + F_{iq_i-1} s + F_{iq_i} \quad (3.85)$$

and

$$\tilde{F}_i(\varepsilon) := \frac{1}{\varepsilon^{q_i}} [F_{iq_i}, \varepsilon F_{iq_i-1}, \dots, \varepsilon^{q_i-1} F_{i1}] \quad (3.86)$$

STEP 3.4.C.S.5: in this step, various gains calculated in Steps 3.4.C.S.2 to 3.4.C.S.4 are put together to form a composite state feedback gain for the given system Σ_p .

Let

$$\tilde{F}_{a1}^+(\varepsilon) := \begin{bmatrix} F_{a11}^+ F_{1q_1} / \varepsilon^{q_1} \\ F_{a12}^+ F_{2q_2} / \varepsilon^{q_2} \\ \vdots \\ F_{a1m_{dp}}^+ F_{m_{dp}q_{m_{dp}}} / \varepsilon^{q_{m_{dp}}} \end{bmatrix} \quad (3.87)$$

and

$$\tilde{F}_{b1}(\varepsilon) := \begin{bmatrix} F_{b11}F_{1q_1}/\varepsilon^{q_1} \\ F_{b12}F_{2q_2}/\varepsilon^{q_2} \\ \vdots \\ F_{b1m_{dp}}F_{m_{dp}q_{m_{dp}}}/\varepsilon^{q_{m_{dp}}} \end{bmatrix} \quad (3.88)$$

Then, the H_2 suboptimal state feedback gain is given by

$$F(\varepsilon) = -\Gamma_{\text{IP}} \left(\tilde{F}(\varepsilon) + \tilde{F}_0 \right) \Gamma_{\text{SP}}^{-1} \quad (3.89)$$

where

$$\tilde{F}(\varepsilon) = \begin{bmatrix} 0 & F_{a0}^+ & F_{b0} & 0 & 0 \\ 0 & \tilde{F}_{a1}^+(\varepsilon) & \tilde{F}_{b1}(\varepsilon) & 0 & \tilde{F}_d(\varepsilon) \\ 0 & 0 & 0 & \Delta_{\text{CP}} & 0 \end{bmatrix} \quad (3.90)$$

$$\tilde{F}_0 = \begin{bmatrix} C_{0aP}^- & C_{0aP}^+ & C_{0bP} & C_{0cP} & C_{0dP} \\ E_{daP}^- & E_{daP}^+ & E_{dbP} & E_{dcP} & E_{dP} \\ E_{caP}^- & E_{caP}^+ & E_{cbP} & 0 & 0 \end{bmatrix} \quad (3.91)$$

and where

$$E_{dP} = \begin{bmatrix} E_{11} & \cdots & E_{1m_{dp}} \\ \vdots & \ddots & \vdots \\ E_{m_{dp}1} & \cdots & E_{m_{dp}m_{dp}} \end{bmatrix} \quad (3.92)$$

$$\tilde{F}_d(\varepsilon) = \text{diag} \left[\tilde{F}_1(\varepsilon), \tilde{F}_2(\varepsilon), \dots, \tilde{F}_{m_{dp}}(\varepsilon) \right] \quad (3.93)$$

This completes the algorithm. \diamond

Theorem 3.11. Consider the given system in Equation 3.48 with $D_{22} = 0$ and $y = x$, i.e. all states are measurable. Assume that Σ_p has no invariant zeros on the imaginary axis. Then, the closed-loop system comprising that of Equation 3.48 and $u = F(\varepsilon)x$ with $F(\varepsilon)$ being given as in Equation 3.89 has the following properties:

1. it is internally stable for sufficiently small ε ;
2. the closed-loop transfer matrix from the disturbance w to the controlled output h , $T_{hw}(s, \varepsilon)$, possesses $\|T_{hw}(s, \varepsilon)\|_2 \rightarrow \gamma_2^*$ as $\varepsilon \rightarrow 0$.

Clearly, $u = F(\varepsilon)x$ is an H_2 suboptimal controller for the system given in Equation 3.48. \diamond

ii. Full-order Output Feedback Case. The following is a step-by-step algorithm for constructing a parameterized full-order output feedback controller that solves the general H_2 optimization problem.

STEP 3.4.C.F.1: (construction of the gain matrix $F_p(\varepsilon)$). Define an auxiliary system

$$\begin{cases} \dot{x} = A x + B u + E w \\ y = x \\ h = C_2 x + D_2 u \end{cases} \quad (3.94)$$

and then perform Steps 3.4.C.S.1 to 3.4.C.S.5 of the previous algorithm on the above system to obtain a parameterized gain matrix $F(\varepsilon)$. Let $F_p(\varepsilon) = F(\varepsilon)$.

STEP 3.4.C.F.2: (construction of the gain matrix $K_Q(\varepsilon)$). Define another auxiliary system

$$\begin{cases} \dot{x} = A' x + C_1' u + C_2' w \\ y = x \\ h = E' x + D_1' u \end{cases} \quad (3.95)$$

and then perform Steps 3.4.C.S.1 to 3.4.C.S.5 on the above system to get the parameterized gain matrix $F(\varepsilon)$. We let $K_Q(\varepsilon) = F(\varepsilon)'$.

STEP 3.4.C.F.3: (construction of the full-order controller Σ_{FC}). Finally, the parameterized full-order output feedback controller is given by

$$\Sigma_{FC} : \begin{cases} \dot{v} = A_{FC}(\varepsilon) v + B_{FC}(\varepsilon) y \\ u = C_{FC}(\varepsilon) v + D_{FC}(\varepsilon) y \end{cases} \quad (3.96)$$

where

$$\left. \begin{aligned} A_{FC}(\varepsilon) &:= A + BF_p(\varepsilon) + K_Q(\varepsilon)C_1 \\ B_{FC}(\varepsilon) &:= -K_Q(\varepsilon) \\ C_{FC}(\varepsilon) &:= F_p(\varepsilon) \\ D_{FC}(\varepsilon) &:= 0 \end{aligned} \right\} \quad (3.97)$$

This concludes the algorithm for constructing the full-order measurement feedback controller. \diamond

Theorem 3.12. *Consider the given system in Equation 3.48 with $D_{22} = 0$. Assume that Σ_p and Σ_Q have no invariant zeros on the imaginary axis. Then the closed-loop system comprising the given system and the full-order output feedback controller of Equation 3.96 has the following properties:*

1. *it is internally stable for sufficiently small ε ;*
2. *the closed-loop transfer matrix from the disturbance w to the controlled output h , $T_{hw}(s, \varepsilon)$, possesses $\|T_{hw}(s, \varepsilon)\|_2 \rightarrow \gamma_2^*$ as $\varepsilon \rightarrow 0$.*

By definition, Equation 3.96 is an H_2 suboptimal controller for the system given in Equation 3.48. \diamond

iii. Reduced-order Output Feedback Case. For the case when some measurement output channels are clean, *i.e.* they are not mixed with disturbances, then we can design an output feedback control law that has a dynamical order less than that of the given plant and yet has an identical performance compared with that of full-order control law. Such a control law is called the reduced-order output feedback controller. We note that the construction of a reduced-order controller was first reported by Chen *et al.* [94] for general linear systems, in which the direct feedthrough matrix from input is nonzero. It was shown in [94] that the reduced-order output feedback controller has the following advantages over the full-order counterpart:

1. the dynamical order of the reduced-order controller is generally smaller than that of the full-order counterpart;
2. the gain required for the same degree of performance for the reduced-order controller is smaller compared with that of the full-order counterpart.

We now proceed to design a reduced-order controller, which solves the general H_2 suboptimal problem. First, without loss of generality and for simplicity of presentation, we assume that the matrices C_1 and D_1 are already in the form

$$C_1 = \begin{bmatrix} 0 & C_{1,02} \\ I_k & 0 \end{bmatrix} \quad \text{and} \quad D_1 = \begin{bmatrix} D_{1,0} \\ 0 \end{bmatrix} \quad (3.98)$$

where $k = \ell - \text{rank}(D_1)$ and $D_{1,0}$ is of full rank. Then the given system in Equation 3.48 can be written as

$$\begin{cases} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u + \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} w \\ \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{bmatrix} 0 & C_{1,02} \\ I_k & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} D_{1,0} \\ 0 \end{bmatrix} w \\ h = [C_{2,1} \quad C_{2,2}] \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + D_2 u + D_{22} w \end{cases} \quad (3.99)$$

where the original state x is partitioned into two parts, x_1 and x_2 ; and y is partitioned into y_0 and y_1 with $y_1 \equiv x_1$. Thus, one needs to estimate only the state x_2 in the reduced-order controller design. Next, define an auxiliary subsystem Σ_{QR} characterized by a matrix quadruple $(A_{\text{R}}, E_{\text{R}}, C_{\text{R}}, D_{\text{R}})$, where

$$(A_{\text{R}}, E_{\text{R}}, C_{\text{R}}, D_{\text{R}}) = \left(A_{22}, E_2, \begin{bmatrix} C_{1,02} \\ A_{12} \end{bmatrix}, \begin{bmatrix} D_{1,0} \\ E_1 \end{bmatrix} \right) \quad (3.100)$$

The following is a step-by-step algorithm that constructs the reduced-order output feedback controller for the general H_2 optimization.

STEP 3.4.C.R.1: (construction of the gain matrix $F_{\text{p}}(\varepsilon)$). Define an auxiliary system

$$\begin{cases} \dot{x} = A x + B u + E w \\ y = x \\ h = C_2 x + D_2 u \end{cases} \quad (3.101)$$

and then perform Steps 3.4.C.S.1 to 3.4.C.S.5 on the above system to get the parameterized gain matrix $F(\varepsilon)$. We let $F_{\text{p}}(\varepsilon) = F(\varepsilon)$.

STEP 3.4.C.R.2: (construction of the gain matrix $K_{\text{R}}(\varepsilon)$). Define another auxiliary system

$$\begin{cases} \dot{x} = A'_{\text{R}} x + C'_{\text{R}} u + C'_{2,2} w \\ y = x \\ h = E'_{\text{R}} x + D'_{\text{R}} u \end{cases} \quad (3.102)$$

and then perform Steps 3.4.C.S.1 to 3.4.C.S.5 on the above system to get the parameterized gain matrix $F(\varepsilon)$. We let $K_{\text{R}}(\varepsilon) = F(\varepsilon)'$.

STEP 3.4.C.R.3: (construction of the reduced-order controller Σ_{RC}). Let us partition $F_{\text{P}}(\varepsilon)$ and $K_{\text{R}}(\varepsilon)$ as

$$F_{\text{P}}(\varepsilon) = [F_{\text{P1}}(\varepsilon) \quad F_{\text{P2}}(\varepsilon)], \quad K_{\text{R}}(\varepsilon) = [K_{\text{R0}}(\varepsilon) \quad K_{\text{R1}}(\varepsilon)] \quad (3.103)$$

in conformity with the partitions of $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $y = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}$ of Equation 3.99, respectively. Then define

$$G_{\text{R}}(\varepsilon) = [-K_{\text{R0}}(\varepsilon), \quad A_{21} + K_{\text{R1}}(\varepsilon)A_{11} - (A_{\text{R}} + K_{\text{R}}(\varepsilon)C_{\text{R}})K_{\text{R1}}(\varepsilon)]$$

Finally, the reduced-order output feedback controller is given by

$$\Sigma_{\text{RC}} : \begin{cases} \dot{v} = A_{\text{RC}}(\varepsilon) v + B_{\text{RC}}(\varepsilon) y \\ u = C_{\text{RC}}(\varepsilon) v + D_{\text{RC}}(\varepsilon) y \end{cases} \quad (3.104)$$

where

$$\left. \begin{aligned} A_{\text{RC}}(\varepsilon) &:= A_{\text{R}} + B_2 F_{\text{P2}}(\varepsilon) + K_{\text{R}}(\varepsilon)C_{\text{R}} + K_{\text{R1}}(\varepsilon)B_1 F_{\text{P2}}(\varepsilon) \\ B_{\text{RC}}(\varepsilon) &:= G_{\text{R}}(\varepsilon) + [B_2 + K_{\text{R1}}(\varepsilon)B_1] \\ &\quad \times [0, \quad F_{\text{P1}}(\varepsilon) - F_{\text{P2}}(\varepsilon)K_{\text{R1}}(\varepsilon)] \\ C_{\text{RC}}(\varepsilon) &:= F_{\text{P2}}(\varepsilon) \\ D_{\text{RC}}(\varepsilon) &:= [0, \quad F_{\text{P1}}(\varepsilon) - F_{\text{P2}}(\varepsilon)K_{\text{R1}}(\varepsilon)] \end{aligned} \right\} \quad (3.105)$$

This concludes the algorithm for constructing the reduced-order output feedback controller. \diamond

Theorem 3.13. *Consider the given system in Equation 3.48 with $D_{22} = 0$. Assume that Σ_{P} and Σ_{Q} have no invariant zeros on the imaginary axis. Then, the closed-loop system comprising the given system and the reduced-order output feedback controller in Equation 3.104 has the following properties:*

1. it is internally stable for sufficiently small ε ;
2. the closed-loop transfer matrix from the disturbance w to the controlled output h , $T_{hw}(s, \varepsilon)$, possesses $\|T_{hw}(s, \varepsilon)\|_2 \rightarrow \gamma_2^*$ as $\varepsilon \rightarrow 0$.

By definition, Equation 3.104 is an H_2 suboptimal controller for the system given in Equation 3.48. \diamond

3.4.2 Discrete-time Systems

We now consider a generalized discrete-time system Σ characterized by the following state-space equations

$$\Sigma : \begin{cases} x(k+1) = A x(k) + B u(k) + E w(k) \\ y(k) = C_1 x(k) + D_{11} u(k) + D_1 w(k) \\ h(k) = C_2 x(k) + D_2 u(k) + D_{22} w(k) \end{cases} \quad (3.106)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, $w \in \mathbb{R}^q$ is the external disturbance input, $y \in \mathbb{R}^p$ is the measurement output, and $h \in \mathbb{R}^\ell$ is the controlled output of Σ . As usual, we let Σ_p be the subsystem characterized by the matrix quadruple (A, B, C_2, D_2) and Σ_0 be the subsystem characterized by (A, E, C_1, D_1) . Without loss of any generality, we assume that $D_{11} = 0$, (A, B) is stabilizable and (A, C_1) is detectable.

The standard H_2 optimal control problem for a discrete-time system is to find an internally stabilizing proper measurement feedback control law,

$$\Sigma_{\text{cmp}} : \begin{cases} v(k+1) = A_{\text{cmp}} v(k) + B_{\text{cmp}} y(k) \\ u(k) = C_{\text{cmp}} v(k) + D_{\text{cmp}} y(k) \end{cases} \quad (3.107)$$

such that the H_2 -norm of the overall closed-loop transfer matrix function from w to h is minimized. To be more specific, we will say that the control law Σ_{cmp} of Equation 3.107 is internally stabilizing when applied to the system Σ of Equation 3.106, if the following matrix is asymptotically stable:

$$A_{\text{cl}} := \begin{bmatrix} A + BD_{\text{cmp}}C_1 & BC_{\text{cmp}} \\ B_{\text{cmp}}C_1 & A_{\text{cmp}} \end{bmatrix} \quad (3.108)$$

i.e. all its eigenvalues lie inside the open unit disc. The closed-loop transfer matrix from the disturbance w to the controlled output h is given by

$$T_{hw}(z) = C_e(zI - A_e)^{-1}B_e + D_e \quad (3.109)$$

where

$$A_e := \begin{bmatrix} A + BD_{\text{cmp}}C_1 & BC_{\text{cmp}} \\ B_{\text{cmp}}C_1 & A_{\text{cmp}} \end{bmatrix} \quad (3.110)$$

$$B_e := \begin{bmatrix} E + BD_{\text{cmp}}D_1 \\ B_{\text{cmp}}D_1 \end{bmatrix} \quad (3.111)$$

$$C_e := [C_2 + D_2D_{\text{cmp}}C_1 \quad D_2C_{\text{cmp}}] \quad (3.112)$$

$$D_e := D_2D_{\text{cmp}}D_1 + D_{22} \quad (3.113)$$

The H_2 -norm of a stable discrete-time transfer matrix, *e.g.*, $T_{hw}(z)$, is defined as follows:

$$\|T_{hw}\|_2 := \left(\frac{1}{2\pi} \text{trace} \left[\int_{-\pi}^{\pi} T_{hw}(e^{j\omega}) T_{hw}(e^{j\omega})^H d\omega \right] \right)^{1/2} \quad (3.114)$$

By Parseval's theorem, $\|T_{hw}\|_2$ can equivalently be defined as

$$\|T_{hw}\|_2 = \left(\text{trace} \left[\sum_{k=0}^{\infty} g(k)g(k)' dt \right] \right)^{1/2} \quad (3.115)$$

where $g(k)$ is the impulse response of $T_{hw}(k)$. Thus, $\|T_{hw}\|_2 = \|g\|_2$.

The H_2 optimal control for the discrete-time system of Equation 3.106 is to design a proper controller Σ_{cmp} such that, when it is applied to the plant Σ , the resulting closed loop is asymptotically stable and the H_2 -norm of $T_{hw}(z)$ is minimized. For future use, we define

$$\gamma_2^* := \inf \left\{ \|T_{hw}(\Sigma \times \Sigma_{\text{cmp}})\|_2 \mid \Sigma_{\text{cmp}} \text{ internally stabilizes } \Sigma \right\} \quad (3.116)$$

Again, a control law Σ_{cmp} is said to be an H_2 optimal controller for Σ of Equation 3.106 if its resulting closed-loop transfer function from w to h has an H_2 -norm equal to γ_2^* , i.e. $\|T_{hw}\|_2 = \gamma_2^*$.

For the case when $D_e = 0$, $\|T_{hw}\|_2$ can be computed by

$$\|T_{hw}\|_2 = \sqrt{\text{trace}[B_e' P_e B_e]} = \sqrt{\text{trace}[C_e Q_e C_e']} \quad (3.117)$$

where P_e and Q_e are, respectively, the solutions of the following Lyapunov equations:

$$A_e' P_e A - P_e + C_e' C_e = 0, \quad A_e Q_e A_e' - Q_e + B_e B_e' = 0 \quad (3.118)$$

We now present solutions to the discrete-time H_2 optimal control problem without detailed proofs. As usual, we assume that $D_{22} = 0$ for convenience. We start first with the simplest case when the given system Σ satisfies the following assumptions of the so-called *regular case*:

1. Σ_P has no invariant zeros on the imaginary axis and is left invertible;
2. Σ_Q has no invariant zeros on the imaginary axis and is right invertible.

The problem is called the *singular case* if Σ does not satisfy these conditions.

Again, the solution to the regular case of the discrete-time H_2 optimal control problem is very simple as well. The optimal controller is given by Σ_{cmp} :

$$\begin{cases} v(k+1) = (A + BF + KC_1 - BNC_1) v(k) + (BN - K) y(k) \\ u(k) = (F - NC_1) v(k) + N y(k) \end{cases} \quad (3.119)$$

where

$$F = -(B'PB + D_2'D_2)^{-1}(B'PA + D_2'C_2) \quad (3.120)$$

$$K = -(AQC_1' + ED_1')(D_1D_1' + C_1QC_1')^{-1} \quad (3.121)$$

and

$$N = -(B'PB + D_2'D_2)^{-1} \left[(B'PA + D_2'D_2)QC_1 + B'PED_1' \right] \times (D_1D_1' + C_1QC_1')^{-1} \quad (3.122)$$

and where $P = P' \geq 0$ and $Q = Q' \geq 0$ are, respectively, the stabilizing solutions of the following Riccati equations:

$$P = A'PA + C_2'C_2 - (C_2'D_2 + A'PB)(D_2'D_2 + B'PB)^{-1} \\ \times (D_2'C_2 + B'PA) \quad (3.123)$$

$$Q = AQA' + EE' - (ED_1' + AQC_1')(D_1D_1' + C_1QC_1')^{-1} \\ \times (D_1E' + C_1QA') \quad (3.124)$$

We note that the above discrete-time Riccati equations can be solved using the non-iterative method given in [74]. If all the states of Σ are available for feedback, then the optimal controller is reduced to a static law $u(k) = Fx(k)$ with F being given as in Equation 3.120.

Next, we present solutions to the singular H_2 optimal control problem. Similarly, the solutions to the singular case are generally suboptimal, and usually parameterized by a certain tuning parameter, say ε . Again, a discrete-time controller parameterized by ε is said to be suboptimal for the system in Equation 3.106 if there exists an $\varepsilon^* > 0$ such that for all $\varepsilon \leq \varepsilon^*$ the closed-loop system comprising the given plant and the controller is asymptotically stable, and the resulting closed-loop transfer function from w to h , which is obviously a function of ε , has an H_2 -norm arbitrarily close to γ_2^* as ε tends to 0.

The following *perturbation approach* would yield a suboptimal controller for the general discrete-time singular case. Given any $\varepsilon > 0$, define

$$\tilde{E} := [E \quad \varepsilon I \quad 0], \quad \tilde{D}_1 := [D_1 \quad 0 \quad \varepsilon I] \quad (3.125)$$

and

$$\tilde{C}_2 := \begin{bmatrix} C_2 \\ \varepsilon I \\ 0 \end{bmatrix}, \quad \tilde{D}_2 := \begin{bmatrix} D_2 \\ 0 \\ \varepsilon I \end{bmatrix} \quad (3.126)$$

A full-order H_2 suboptimal output feedback controller is given by

$$\begin{cases} v(k+1) = (A+B\tilde{F}+\tilde{K}C_1-B\tilde{N}C_1) v(k) + (B\tilde{N}-\tilde{K}) y(k) \\ u(k) = (\tilde{F}-\tilde{N}C_1) v(k) + \tilde{N} y(k) \end{cases} \quad (3.127)$$

where

$$\tilde{F} = -(B'\tilde{P}B + \tilde{D}_2'\tilde{D}_2)^{-1}(B'\tilde{P}A + \tilde{D}_2'\tilde{C}_2) \quad (3.128)$$

$$\tilde{K} = -(A\tilde{Q}C_1' + \tilde{E}\tilde{D}_1')(\tilde{D}_1\tilde{D}_1' + C_1\tilde{Q}C_1')^{-1} \quad (3.129)$$

and

$$\tilde{N} = -(B'\tilde{P}B + \tilde{D}_2'\tilde{D}_2)^{-1} \left[(B'\tilde{P}A + \tilde{D}_2'\tilde{D}_2)\tilde{Q}C_1 + B'\tilde{P}\tilde{E}\tilde{D}_1' \right] \\ \times (\tilde{D}_1\tilde{D}_1' + C_1\tilde{Q}C_1')^{-1} \quad (3.130)$$

and where $\tilde{P} = \tilde{P}' > 0$ and $\tilde{Q} = \tilde{Q}' > 0$ are, respectively, the stabilizing solutions of the following Riccati equations:

$$\begin{aligned} \tilde{P} = A'\tilde{P}A + \tilde{C}'_2\tilde{C}_2 - (\tilde{C}'_2\tilde{D}_2 + A'\tilde{P}B)(\tilde{D}'_2\tilde{D}_2 + B'\tilde{P}B)^{-1} \\ \times (\tilde{D}'_2\tilde{C}_2 + B'\tilde{P}A) \end{aligned} \quad (3.131)$$

$$\begin{aligned} \tilde{Q} = A\tilde{Q}A' + \tilde{E}\tilde{E}' - (\tilde{E}\tilde{D}'_1 + A\tilde{Q}C'_1)(\tilde{D}_1\tilde{D}'_1 + C_1\tilde{Q}C'_1)^{-1} \\ \times (\tilde{D}_1\tilde{E}' + C_1\tilde{Q}A') \end{aligned} \quad (3.132)$$

The following are alternative methods based on the structural decompositions of the given systems. Similarly, we separate the problem into three distinct situations: 1) the state feedback case, 2) the full-order measurement feedback case, and 3) the reduced-order measurement feedback case.

Similarly, for convenience, we assume throughout the rest of this subsection that both subsystems Σ_p and Σ_Q *have no invariant zeros on the unit circle*. The complete treatment of H_2 optimal control using the approach given below can be found in [90]. Interestingly, it turns out that for this case, although it is singular, we can always obtain a set of H_2 optimal controllers that need not be parameterized by any tuning scalar.

i. State Feedback Case. For the case when $y = x$ in the given system Σ of Equation 3.106, we have the following step-by-step algorithm that constructs an H_2 suboptimal static feedback control law $u = F(\varepsilon)x$ for Σ .

STEP 3.4.D.S.1: (decomposition of Σ_p). Transform the subsystem Σ_p , *i.e.* the matrix quadruple (A, B, C_2, D_2) , into the special coordinate basis as given by Theorem 3.1. Denote the state, output and input transformation matrices as Γ_{sp} , Γ_{op} and Γ_{ip} , respectively.

STEP 3.4.D.S.2: (gain matrix for the subsystem associated with \mathcal{X}_c). Let F_c be any constant matrix subject to the constraint that

$$A_{cc}^c = A_{cc} - B_c F_c \quad (3.133)$$

is a stable matrix. Note that the existence of such an F_c is guaranteed by the property of the special coordinate basis, *i.e.* (A_{cc}, B_c) is controllable.

STEP 3.4.D.S.3: (gain matrix for the subsystem associated with \mathcal{X}_a^+ , \mathcal{X}_b and \mathcal{X}_d).

Let

$$A_x = \begin{bmatrix} A_{aa}^+ & L_{ad}^+ C_d & L_{ab}^+ C_b \\ B_d E_{da}^+ & A_{dd} & B_d E_{db} \\ 0 & L_{bd} C_d & A_{bb} \end{bmatrix}, \quad E_x = \begin{bmatrix} E_a^+ \\ E_d \\ E_b \end{bmatrix} \quad (3.134)$$

$$B_x = \begin{bmatrix} B_{0a}^+ & 0 \\ B_{0d} & B_d \\ B_{0b} & 0 \end{bmatrix} \quad (3.135)$$

and

$$C_x = \Gamma_{op} \begin{bmatrix} 0 & 0 & 0 \\ 0 & C_d & 0 \\ 0 & 0 & C_b \end{bmatrix}, \quad D_x = \Gamma_{op} \begin{bmatrix} I & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (3.136)$$

Solve the following discrete-time Riccati equation

$$P_x = A'_x P_x A_x + C'_x C_x - (C'_x D_x + A'_x P_x B_x)(D'_x D_x + B'_x P_x B_x)^{-1} \times (D'_x C_x + B'_x P_x A_x) \quad (3.137)$$

for $P_x > 0$. Then partition

$$\begin{aligned} F_x &= (B'_x P_x B_x + D'_x D_x)^{-1} (B'_x P_x A_x + D'_x C_x) \\ &= \begin{bmatrix} F_{a0}^+ & F_{d0} & F_{b0} \\ F_{ad}^+ & F_{dd} & F_{bd} \end{bmatrix} \end{aligned} \quad (3.138)$$

STEP 3.4.D.S.4: (composition of gain matrix F). In this step, various gains calculated in the previous steps are put together to form a composite state feedback gain matrix F . It is given by

$$F = -\Gamma_{\text{IP}} \begin{bmatrix} C_{0a}^- & C_{0a}^+ + F_{a0}^+ & C_{0b} + F_{b0} & C_{0c} & C_{0d} + F_{d0} \\ E_{da}^- & F_{ad}^+ & F_{bd} & E_{dc} & F_{dd} \\ E_{ca}^- & E_{ca}^+ & 0 & F_c & 0 \end{bmatrix} \Gamma_{\text{SP}}^{-1} \quad (3.139)$$

This completes the algorithm. \diamond

Theorem 3.14. Consider the system given in Equation 3.106 with $D_{22} = 0$ and $y = x$, i.e. all states are measurable. Assume that Σ_p has no invariant zeros on the unit circle. Then, the closed-loop system comprising the given system and $u(k) = Fx(k)$ with F being given as in Equation 3.139 has the following properties:

1. it is internally stable;
2. the closed-loop transfer matrix from the disturbance w to the controlled output h , $T_{hw}(z)$, possesses $\|T_{hw}\|_2 = \gamma_2^*$.

Thus, $u(k) = Fx(k)$ is an H_2 optimal state feedback control law for the system in Equation 3.106. \diamond

ii. Full-order Output Feedback Case. The following is a step-by-step algorithm for constructing an H_2 optimal full-order output feedback controller.

STEP 3.4.D.F.1: (computation of N). Utilize the special coordinate basis properties to compute two constant matrices X and Y such that $\mathcal{V}^\ominus(\Sigma_p) = \text{Ker}(X)$ and $\mathcal{S}^\ominus(\Sigma_0) = \text{Im}(Y)$. Then, compute

$$\begin{aligned} N &= -(B'X'XB + D_2'D_2)^\dagger [B'X' \quad D_2'] \begin{bmatrix} XAY & XE \\ C_2Y & D_{22} \end{bmatrix} \\ &\quad \times \begin{bmatrix} Y'C_1' \\ D_1' \end{bmatrix} (C_1YY'C_1' + D_1D_1')^\dagger \end{aligned} \quad (3.140)$$

STEP 3.4.D.F.2: (construction of the gain matrix F_p). Define an auxiliary system

$$\begin{cases} x(k+1) = \tilde{A} x(k) + B u(k) + \tilde{E} w(k) \\ y(k) = x(k) \\ h(k) = \tilde{C}_2 x(k) + D_2 u(k) + 0 w(k) \end{cases} \quad (3.141)$$

where

$$\tilde{A} = A + BNC_1, \quad \tilde{E} = E + BND_1, \quad \tilde{C}_2 = C_2 + D_2NC_1 \quad (3.142)$$

and then perform Steps 3.4.D.S.1 to 3.4.D.S.4 of the previous algorithm on the above system in Equation 3.141 to obtain a gain matrix F . We let $F_p = F$.

STEP 3.4.D.F.3: (construction of the gain matrix K_Q). Define another auxiliary system

$$\begin{cases} x(k+1) = \tilde{A}' x(k) + C_1' u(k) + \tilde{C}_2' w(k) \\ y(k) = x(k) \\ h(k) = \tilde{E}' x(k) + D_1' u(k) + 0 w(k) \end{cases} \quad (3.143)$$

and then perform Steps 3.4.D.S.1 to 3.4.D.S.4 of the previous algorithm on the above system to get a gain matrix F . Similarly, we let $K_Q = F'$.

STEP 3.4.D.F.4: (construction of the full-order controller Σ_{FC}). Finally, the parameterized full-order output feedback controller is given by

$$\Sigma_{FC} : \begin{cases} v(k+1) = A_{FC} v(k) + B_{FC} y(k) \\ u(k) = C_{FC} v(k) + D_{FC} y(k) \end{cases} \quad (3.144)$$

where

$$\left. \begin{aligned} A_{FC} &:= A + BNC_1 + BF_p + K_Q C_1 \\ B_{FC} &:= -K_Q \\ C_{FC} &:= F_p \\ D_{FC} &:= N \end{aligned} \right\} \quad (3.145)$$

This concludes the algorithm for constructing the full-order measurement feedback controller. \diamond

Theorem 3.15. *Consider the system given in Equation 3.106 with $D_{22} = 0$. Assume that Σ_p and Σ_Q have no invariant zeros on the unit circle. Then the closed-loop system comprising the given system and the full-order output feedback controller of Equation 3.144 has the following properties:*

1. *it is internally stable;*
2. *the closed-loop transfer matrix from the disturbance w to the controlled output h , $T_{hw}(z)$, possesses $\|T_{hw}\|_2 = \gamma_2^*$.*

Hence, Equation 3.144 is an H_2 optimal control law for the system given in Equation 3.106. \diamond

iii. Reduced-order Output Feedback Case. We now follow the procedure as in the continuous-time case to design a reduced-order output feedback controller. For simplicity of presentation, we assume that the matrices C_1 and D_1 are already in the form

$$C_1 = \begin{bmatrix} 0 & C_{1,02} \\ I_k & 0 \end{bmatrix} \quad \text{and} \quad D_1 = \begin{bmatrix} D_{1,0} \\ 0 \end{bmatrix} \quad (3.146)$$

where $k = \ell - \text{rank}(D_1)$ and $D_{1,0}$ is of full rank. Next, we follow Steps 3.4.D.F.1 and 3.4.D.F.2 of the previous subsection to compute the constant matrix N , and form the following system:

$$\begin{cases} x(k+1) = \tilde{A} x(k) + B u(k) + \tilde{E} w(k) \\ y(k) = C_1 x(k) + D_1 w(k) \\ h(k) = \tilde{C}_2 x(k) + D_2 u(k) + 0 w(k) \end{cases} \quad (3.147)$$

where \tilde{A} , \tilde{E} and \tilde{C}_2 are defined as in Equation 3.142. Then, partition Equation 3.147 as follows:

$$\begin{cases} \begin{pmatrix} x_1(k+1) \\ x_2(k+1) \end{pmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u(k) + \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} w(k) \\ \begin{pmatrix} y_0(k) \\ y_1(k) \end{pmatrix} = \begin{bmatrix} 0 & C_{1,02} \\ I_k & 0 \end{bmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} + \begin{bmatrix} D_{1,0} \\ 0 \end{bmatrix} w(k) \\ h(k) = [C_{2,1} \quad C_{2,2}] \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} + D_2 u(k) + 0 w(k) \end{cases}$$

where the state x of Equation 3.147 is partitioned into two parts, x_1 and x_2 ; and y is partitioned to y_0 and y_1 with $y_1 \equiv x_1$. Thus, one needs to estimate only the state x_2 in the reduced-order controller design. Next, define an auxiliary subsystem Σ_{QR} characterized by a matrix quadruple $(A_{\text{R}}, E_{\text{R}}, C_{\text{R}}, D_{\text{R}})$, where

$$(A_{\text{R}}, E_{\text{R}}, C_{\text{R}}, D_{\text{R}}) = \left(A_{22}, E_2, \begin{bmatrix} C_{1,02} \\ A_{12} \end{bmatrix}, \begin{bmatrix} D_{1,0} \\ E_1 \end{bmatrix} \right) \quad (3.148)$$

The following is a step-by-step algorithm that constructs the reduced-order output feedback controller for the general H_2 optimization.

STEP 3.4.D.R.1: (construction of the gain matrix F_{p}). Define an auxiliary system

$$\begin{cases} x(k+1) = \tilde{A} x(k) + B u(k) + \tilde{E} w(k) \\ y(k) = x(k) \\ h(k) = \tilde{C}_2 x(k) + D_2 u(k) + 0 w(k) \end{cases} \quad (3.149)$$

and then perform Steps 3.4.D.S.1 to 3.4.D.S.4 of the previous algorithm on the above system to obtain a parameterized gain matrix F . Moreover, let $F_{\text{p}} = F$.

STEP 3.4.D.R.2: (construction of the gain matrix K_{R}). Define another auxiliary system

$$\begin{cases} x(k+1) = A'_R x(k) + C'_R u(k) + C'_{2,2} w(k) \\ y(k) = x(k) \\ h(k) = E'_R x(k) + D'_R u(k) + 0 w(k) \end{cases} \quad (3.150)$$

and then perform Steps 3.4.D.S.1 to 3.4.D.S.4 of the previous algorithm on the above system to obtain a parameterized gain matrix F . Similarly, let $K_R = F'$.

STEP 3.4.D.R.3: (construction of the reduced-order controller Σ_{RC}). Let us partition F_P and K_R as

$$F_P = [F_{P1} \ F_{P2}] \quad \text{and} \quad K_R = [K_{R0} \ K_{R1}] \quad (3.151)$$

in conformity with $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $y = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}$, respectively. Then define

$$G_R = [-K_{R0}, \ A_{21} + K_{R1}A_{11} - (A_R + K_R C_R)K_{R1}] \quad (3.152)$$

Finally, the parameterized reduced-order output feedback controller is given by

$$\Sigma_{RC} : \begin{cases} v(k+1) = A_{RC} v(k) + B_{RC} y(k) \\ u(k) = C_{RC} v(k) + D_{RC} y(k) \end{cases} \quad (3.153)$$

where

$$\left. \begin{aligned} A_{RC} &:= A_R + B_2 F_{P2} + K_R C_R + K_{R1} B_1 F_{P2} \\ B_{RC} &:= G_R + [B_2 + K_{R1} B_1] [0, \ F_{P1} - F_{P2} K_{R1}] \\ C_{RC} &:= F_{P2} \\ D_{RC} &:= [0, \ F_{P1} - F_{P2} K_{R1}] + N \end{aligned} \right\} \quad (3.154)$$

This concludes the algorithm for constructing the reduced-order output feedback controller. \diamond

Theorem 3.16. *Consider the system given in Equation 3.106 with $D_{22} = 0$. Assume that Σ_p and Σ_q have no invariant zeros on the unit circle. Then, the closed-loop system comprising the given system and the reduced-order output feedback controller in Equation 3.153 has the following properties:*

1. *it is internally stable;*
2. *the closed-loop transfer matrix from the disturbance w to the controlled output h , $T_{hw}(z)$, possesses $\|T_{hw}(z)\|_2 = \gamma_2^*$.*

Thus, Equation 3.153 is an H_2 optimal control law for the system given in Equation 3.106. \diamond

Lastly, we note that the result presented in this section, although it is not totally complete, is sufficient to obtain appropriate solutions for HDD servo systems and many engineering problems. We next move to H_∞ control and its related problems.

3.5 H_∞ Control and Disturbance Decoupling

The ultimate goal of a control system designer is to build a system that works in a real environment. Since the real environment may change and the operating conditions may vary from time to time, the control system must be able to withstand these variations. Even if the environment does not change, other factors of life are the model uncertainties, as well as noises. Any mathematical representation of a system often involves simplifying assumptions. Nonlinearities are either unknown, and hence unmodeled, or are modeled and later ignored in order to simplify analysis. High-frequency dynamics are often ignored at the design stage as well. In consequence, control systems designed based on simplified models may not work on real plants in real environments. The particular property that a control system must possess for it to operate properly in realistic situations is commonly called *robustness*. Mathematically, this means that the controller must perform satisfactorily not just for one plant, but for a family of plants. If a controller can be designed such that the whole system to be controlled remains stable when its parameters vary within certain expected limits, the system is said to possess robust stability. In addition, if it can satisfy performance specifications such as steady state tracking, disturbance rejection and speed of response requirements, it is said to possess robust performance. The problem of designing controllers that satisfy both robust stability and performance requirements is called robust control. H_∞ control theory is one of the cornerstones of modern control theory and was developed in an attempt to solve such a problem. Many robust control problems (such as the robust stability problem of unstructurally perturbed systems, the mixed-sensitivity problem, robust stabilization with additive and multiplicative perturbations, to name a few) can be cast into a standard H_∞ control problem (see, e.g., [74]).

Since the original formulation of the H_∞ control problem by Zames [95], a great deal of work has been done on finding the solution to this problem. Practically all the research results of the early years involved a mixture of time-domain and frequency-domain techniques, including the following: *Interpolation approach* (see, e.g., Limebeer and Anderson [96]); *Frequency domain approach* (see, e.g., Doyle [97], Francis [98] and Glover [99]); *Polynomial approach* (see, e.g., Kwakernaak [100]); and *J-spectral factorization approach* (see, e.g., Kimura [101]). Recently, considerable attention has been focused on purely *time-domain methods* based on algebraic Riccati equations and/or singular perturbation approach (see, e.g. [74, 91, 102] and references cited therein). Along this line of research, connections are also made between H_∞ optimal control and differential games (see, e.g., Başar and Bernhard [103]).

We also recall in this section the solutions to the problem of H_∞ almost disturbance decoupling with measurement feedback and internal stability. Although, in principle, it is a special case of the general H_∞ control problem, the problem of almost disturbance decoupling has a vast history behind it, occupying a central part of classical as well as modern control theory. Several important problems, such as robust control, decentralized control, non-interactive control, model reference or tracking control, H_2 and H_∞ optimal control problems can all be recast into an almost disturbance decoupling problem. Roughly speaking, the basic almost distur-

bance decoupling problem is to find an output feedback control law such that in the closed-loop system the disturbances are quenched, say in an L_p sense, up to any prespecified degree of accuracy while maintaining internal stability. Such a problem was originally formulated by Willems [104, 105] and termed almost disturbance decoupling problem with measurement feedback and internal stability (ADDPMS).

The formulation of H_∞ control is very similar to that of H_2 optimal control. In order to avoid unnecessary repetitions, we make use of some terms defined in the previous section, *e.g.*, the state-space equations of the given system and its subsystems Σ_p and Σ_o , the format of the control law and its corresponding closed-loop transfer matrix, as well as the definitions of the regular and singular problems.

3.5.1 Continuous-time Systems

We consider a continuous-time linear time-invariant system as given in Equation 3.48. For simplicity, we assume that (A, B) is stabilizable, (A, C_1) is detectable, $D_{11} = 0$ and $D_{22} = 0$. The standard H_∞ control problem for continuous-time systems is to find an internally stabilizing proper measurement feedback control law of the format in Equation 3.51 such that when it is applied to the given system the resulting closed-loop system is internally stable and the H_∞ -norm of the overall closed-loop transfer matrix function from w to h , *i.e.* $T_{hw}(s)$, is minimized. The H_∞ -norm of a stable continuous-time transfer matrix, *e.g.*, $T_{hw}(s)$, is defined as follows:

$$\|T_{hw}\|_\infty := \sup_{\omega \in [0, \infty)} \sigma_{\max} [T_{hw}(j\omega)] = \sup_{\|w\|_2=1} \frac{\|h\|_2}{\|w\|_2} \quad (3.155)$$

where w and h are, respectively, the input and output of $T_{hw}(s)$, and $\|\cdot\|_2$ is the l_2 -norm of the corresponding signal. It is clear that the H_∞ -norm of $T_{hw}(s)$ corresponds to the worst case gain from its input to its output. For future use, we define

$$\gamma_\infty^* := \inf \left\{ \|T_{hw}(\Sigma \times \Sigma_{\text{cmp}})\|_\infty \mid \Sigma_{\text{cmp}} \text{ internally stabilizes } \Sigma \right\} \quad (3.156)$$

We note that the determination of this γ_∞^* is rather tedious. For a fairly large class of systems, γ_∞^* can be exactly computed using some numerically stable algorithms. In general, an iterative scheme is required to determine γ_∞^* . We refer interested readers to the work of Chen [74] for a detailed treatment of this particular issue. For simplicity, we assume throughout this section that γ_∞^* has been determined and hence it is known.

For the case when $\gamma_\infty^* = 0$, the corresponding H_∞ control problem is commonly known in the literature as the problem of H_∞ almost disturbance decoupling with internal stability. It can be shown that such a problem is solvable for Σ of Equation 3.48 if and only if the following conditions hold (see, *e.g.*, [74, 83]):

1. (A, B) is stabilizable;
2. (A, C_1) is detectable;
3. $D_{22} + D_2 S D_1 = 0$, where $S = -(D_2' D_2)^\dagger D_2' D_{22} D_1' (D_1 D_1')^\dagger$;
4. $\text{Im}(E + B S D_1) \subset \mathcal{S}^+(\Sigma_p) \cap \{\cap_{\lambda \in \mathbb{C}^0} \mathcal{S}_\lambda(\Sigma_p)\}$;

5. $\text{Ker}(C_2 + D_2 S C_1) \supset \mathcal{V}^+(\Sigma_0) \cup \{\cup_{\lambda \in \mathbb{C}^0} \mathcal{V}_\lambda(\Sigma_0)\}$; and
6. $\mathcal{V}^+(\Sigma_0) \subset \mathcal{S}^+(\Sigma_p)$.

We note that if Σ_p is right invertible and of minimum phase, and Σ_0 is left invertible and of minimum phase, then conditions 4–6 are automatically satisfied.

It transpires that, for H_∞ control, it is almost impossible to find a control law with a finite gain to achieve the optimal performance, *i.e.* γ_∞^* . As such, we focus on designing H_∞ suboptimal controllers instead. To be more specific, given a scalar $\gamma > \gamma_\infty^*$, we focus on finding a control law that yields $\|T_{hw}\|_\infty < \gamma$, where $T_{hw}(s)$ is the corresponding closed-loop transfer matrix. Hereafter, we call a control law that possesses such a property an H_∞ γ -suboptimal controller.

Next, we proceed to construct a solution to the regular problem (for its definition see the previous section). Given a scalar $\gamma > \gamma_\infty^*$, we solve for positive semi-definite stabilizing solutions $P \geq 0$ and $Q \geq 0$ respectively to the following Riccati equations:

$$\begin{aligned} A'P + PA + C_2' C_2 + \gamma^{-2} P E E' P \\ - (PB + C_2' D_2)(D_2' D_2)^{-1}(B'P + D_2' C_2) = 0 \end{aligned} \quad (3.157)$$

and

$$\begin{aligned} AQ + QA' + EE' + \gamma^{-2} Q C_2' C_2 Q \\ - (QC_1' + ED_1')(D_1 D_1')^{-1}(C_1 Q + D_1 E') = 0 \end{aligned} \quad (3.158)$$

The H_∞ γ -suboptimal control law is given by (see also [91]),

$$\Sigma_{\text{cmp}} : \begin{cases} \dot{v} = A_{\text{cmp}} v + B_{\text{cmp}} y \\ u = C_{\text{cmp}} v + 0 y \end{cases} \quad (3.159)$$

where

$$A_{\text{cmp}} = A + \gamma^{-2} E E' P + B F + (I - \gamma^{-2} Q P)^{-1} K (C_1 + \gamma^{-2} D_1 E' P) \quad (3.160)$$

$$B_{\text{cmp}} = - (I - \gamma^{-2} Q P)^{-1} K \quad (3.161)$$

$$C_{\text{cmp}} = F \quad (3.162)$$

and where

$$F = -(D_2' D_2)^{-1}(D_2' C_2 + B' P), \quad K = -(QC_1' + ED_1')(D_1 D_1')^{-1} \quad (3.163)$$

Note that, for the state feedback case, the H_∞ γ -suboptimal control law is given by $u = Fx$ with F being given as in Equation 3.163.

For the singular case, the following perturbation method can be utilized. For $\gamma > \gamma_\infty^*$ and a positive scalar $\varepsilon > 0$, define \tilde{E} , \tilde{D}_1 , \tilde{C}_2 and \tilde{D}_2 as in Equations 3.68 and 3.69, and solve the following Riccati equations:

$$\begin{aligned} A' \tilde{P} + \tilde{P} A + \tilde{C}_2' \tilde{C}_2 + \gamma^{-2} \tilde{P} \tilde{E} \tilde{E}' \tilde{P} \\ - (\tilde{P} B + \tilde{C}_2' \tilde{D}_2)(\tilde{D}_2' \tilde{D}_2)^{-1}(B' \tilde{P} + \tilde{D}_2' \tilde{C}_2) = 0 \end{aligned} \quad (3.164)$$

and

$$\begin{aligned} A\tilde{Q} + \tilde{Q}A' + \tilde{E}\tilde{E}' + \gamma^{-2}\tilde{Q}\tilde{C}'_2\tilde{C}_2\tilde{Q} \\ - (\tilde{Q}C'_1 + \tilde{E}\tilde{D}'_1)(\tilde{D}_1\tilde{D}'_1)^{-1}(C_1\tilde{Q} + \tilde{D}_1\tilde{E}') = 0 \end{aligned} \quad (3.165)$$

for $\tilde{P} > 0$ and $\tilde{Q} > 0$. Then, it can be shown that there exists an $\varepsilon^* > 0$ such that for all $\varepsilon \in (0, \varepsilon^*]$ the following control law is an H_∞ γ -suboptimal for the given system:

$$\tilde{\Sigma}_{\text{cmp}} : \begin{cases} \dot{v} = \tilde{A}_{\text{cmp}} v + \tilde{B}_{\text{cmp}} y \\ u = \tilde{C}_{\text{cmp}} v + 0 y \end{cases} \quad (3.166)$$

where

$$\tilde{A}_{\text{cmp}} = A + \gamma^{-2}EE'\tilde{P} + B\tilde{F} + (I - \gamma^{-2}\tilde{Q}\tilde{P})^{-1}\tilde{K}(C_1 + \gamma^{-2}D_1E'\tilde{P}) \quad (3.167)$$

$$\tilde{B}_{\text{cmp}} = - (I - \gamma^{-2}\tilde{Q}\tilde{P})^{-1}\tilde{K} \quad (3.168)$$

$$\tilde{C}_{\text{cmp}} = \tilde{F} \quad (3.169)$$

and where

$$\tilde{F} = -(\tilde{D}'_2\tilde{D}_2)^{-1}(\tilde{D}'_2\tilde{C}_2 + B'\tilde{P}) \quad (3.170)$$

$$\tilde{K} = -(\tilde{Q}C'_1 + \tilde{E}\tilde{D}'_1)(\tilde{D}_1\tilde{D}'_1)^{-1} \quad (3.171)$$

Note that for the state feedback case, the H_∞ γ -suboptimal control law is given by $u = \tilde{F}x$ with \tilde{F} being given as in Equation 3.170.

Alternatively, the singular H_∞ control problem can also be solved using the singular perturbation approach as in the previous section for H_2 optimal control. In fact, we only need to modify the algorithms slightly for H_2 control to yield the required H_∞ γ -suboptimal controllers. We treat separately the state feedback case and the measurement feedback case. For simplicity, we assume that both subsystems Σ_p and Σ_q do not have invariant zeros on the imaginary axis.

i. State Feedback Case. Given a scalar $\gamma > \gamma_\infty^*$, the following algorithm yields an H_∞ γ -suboptimal state feedback gain matrix for Σ of Equation 3.48 with $y = x$.

STEP 3.5.C.S.1: transform the system Σ_p into the special coordinate basis as given by Theorem 3.1. To all submatrices and transformations in the special coordinate basis of Σ_p , we append the subscript p to signify their relation to the system Σ_p . We also choose the output transformation Γ_{op} to have the following form:

$$\Gamma_{\text{op}} = \begin{bmatrix} I_{m_{\text{op}}} & 0 \\ 0 & \Gamma_{\text{orp}} \end{bmatrix} \quad (3.172)$$

where $m_{\text{op}} = \text{rank}(D_2)$. Next, compute

$$\Gamma_{sp}^{-1}E = \begin{bmatrix} E_{ap}^- \\ E_{ap}^+ \\ E_{bp} \\ E_{cp} \\ E_{dp} \end{bmatrix} \quad (3.173)$$

and define A_{11p} , B_{11p} , A_{13p} , C_{21p} , C_{23p} , A_{xp} , $B_{xp}B'_{xp}$, and $C'_{xp}C_{xp}$ as in Equations 3.76–3.80. Finally, define

$$E_{xp} = \begin{bmatrix} E_{ap}^+ \\ E_{bp} \end{bmatrix} \quad (3.174)$$

STEP 3.5.C.S.2: solve the following algebraic matrix Riccati equation,

$$\begin{aligned} P_x A_{xp} + A'_{xp} P_x + P_x E_{xp} E'_{xp} P_x / \gamma^2 \\ - P_x B_{xp} B'_{xp} P_x + C'_{xp} C_{xp} = 0 \end{aligned} \quad (3.175)$$

for $P_x > 0$ and define

$$F_{11} := \begin{bmatrix} F_{a0}^+ & F_{b0} \\ F_{a1}^+ & F_{b1} \end{bmatrix} = \begin{bmatrix} B'_{11p} P_x \\ (C'_{23p} C_{23p})^{-1} (A'_{13p} P_x + C'_{23p} C_{21p}) \end{bmatrix} \quad (3.176)$$

Then, partition $[F_{a1}^+ \ F_{b1}]$ as

$$[F_{a1}^+ \ F_{b1}] = \begin{bmatrix} F_{a11}^+ & F_{b11} \\ F_{a12}^+ & F_{b12} \\ \vdots & \vdots \\ F_{a1m_{dp}}^+ & F_{b1m_{dp}} \end{bmatrix} \quad (3.177)$$

where F_{a1i}^+ and F_{b1i} are of dimensions $1 \times n_{ap}^+$ and $1 \times n_{bp}$, respectively.

STEP 3.5.C.S.3: follow exactly Steps 3.4.C.S.3 to 3.4.C.S.5 of the previous section to construct a state feedback gain matrix $F(\varepsilon)$. \diamond

We have the following result.

Theorem 3.17. *Consider the system given in Equation 3.48 with $D_{22} = 0$ and $y = x$, i.e. all states are measurable. Assume that Σ_p has no invariant zeros on the imaginary axis. Then, the closed-loop system comprising the given system and $u = F(\varepsilon)x$ with $F(\varepsilon)$ being given as in the above algorithm has the following properties:*

1. *it is internally stable for sufficiently small ε ;*
2. *the closed-loop transfer matrix from the disturbance w to the controlled output h , $T_{hw}(s, \varepsilon)$, possesses $\|T_{hw}(s, \varepsilon)\|_\infty < \gamma$, for sufficiently small ε .*

Clearly, $u = F(\varepsilon)x$ is an H_∞ γ -suboptimal control law for the system given in Equation 3.48. \diamond

ii. Measurement Feedback Case. Similarly, one can design either a full-order or a reduced-order output feedback control law that solves the H_∞ γ -suboptimal problem. Unfortunately, the reduced-order controller design for H_∞ control is quite different from its counterpart in H_2 control and is quite complicated. We only focus below on the design of a full-order H_∞ γ -suboptimal controller. Interested readers are referred to [74] for a more complete treatment. The following is a step-by-step algorithm to construct a full-order H_∞ γ -suboptimal control law for any $\gamma > \gamma_\infty^*$.

STEP 3.5.C.F.1: define an auxiliary full state feedback system

$$\begin{cases} \dot{x} = A x + B u + E w \\ y = x \\ h = C_2 x + D_2 u \end{cases} \quad (3.178)$$

and perform Steps 3.5.C.s.1 to 3.5.C.s.3 of the previous algorithm to obtain a gain matrix $F(\varepsilon)$. Also, define

$$P := (\Gamma_{\text{sp}}^{-1})' \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & P_x & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \Gamma_{\text{sp}}^{-1} \geq 0 \quad (3.179)$$

Note that $P_x > 0$ is the solution of Equation 3.175.

STEP 3.5.C.F.2: define another auxiliary full state feedback system as follows:

$$\begin{cases} \dot{x} = A' x + C_1' u + C_2' w \\ y = x \\ h = E' x + D_1' u \end{cases} \quad (3.180)$$

and perform Steps 3.5.C.s.1 to 3.5.C.s.3 of the previous algorithm, but for this auxiliary system, to obtain a gain matrix $F(\varepsilon)$. Let us define $K(\varepsilon) := F(\varepsilon)'$. Similarly, define a positive semi-definite matrix Q as in Equation 3.179, but for the current auxiliary system.

STEP 3.5.C.F.3: the full-order H_∞ γ -suboptimal controller is given by

$$\Sigma_{\text{cmp}} : \begin{cases} \dot{v} = A_{\text{cmp}} v + B_{\text{cmp}} y \\ u = C_{\text{cmp}} v + 0 y \end{cases} \quad (3.181)$$

where

$$\begin{aligned} A_{\text{cmp}} &= A + \gamma^{-2} E E' P + B F(\varepsilon) \\ &+ (I - \gamma^{-2} Q P)^{-1} \left[K(\varepsilon) (C_1 + \gamma^{-2} D_1 E' P) \right. \\ &+ \gamma^{-2} Q (A' P + P A + C_2' C_2 + \gamma^{-2} P E E' P) \\ &\left. + \gamma^{-2} Q (P B + C_2' D_2) F(\varepsilon) \right] \end{aligned} \quad (3.182)$$

and

$$B_{\text{cmp}} = -(I - \gamma^{-2} Q P)^{-1} K(\varepsilon), \quad C_{\text{cmp}} = F(\varepsilon) \quad (3.183)$$

This completes the algorithm. \diamond

We have the following theorem.

Theorem 3.18. *Consider the system given in Equation 3.48 with $D_{22} = 0$. Assume that Σ_p and Σ_q have no invariant zeros on the imaginary axis. Then, the closed-loop system comprising the given system and the measurement feedback controller in Equation 3.181 has the following properties:*

1. *it is internally stable for sufficiently small ε ;*
2. *the closed-loop transfer matrix from the disturbance w to the controlled output h , $T_{hw}(s, \varepsilon)$, possesses $\|T_{hw}(s, \varepsilon)\|_\infty < \gamma$, for sufficiently small ε .*

Hence, Equation 3.181 is an H_∞ γ -suboptimal control law for the system given in Equation 3.48. \diamond

Remark 3.19. For the case when $\gamma_\infty^* = 0$, i.e. the problem of H_∞ almost disturbance decoupling with internal stability for Σ of Equation 3.48 (with both subsystems Σ_p and Σ_q having no invariant zeros on the imaginary axis) is solvable, then either the above algorithm or the algorithm in the previous section for H_2 optimal control would yield desired solutions. We note that, in general, suboptimal solutions are nonunique. In fact, one could utilize the algorithm for constructing the reduced-order output feedback H_2 suboptimal controller in the previous section to construct a reduced-order solution for the H_∞ almost disturbance decoupling problem. \diamond

3.5.2 Discrete-time Systems

We now consider a discrete-time linear time-invariant system as given in Equation 3.106. Again, for simplicity, we assume that (A, B) is stabilizable, (A, C_1) is detectable, $D_{11} = 0$ and $D_{22} = 0$. The standard H_∞ control problem for discrete-time systems is to find an internally stabilizing proper measurement feedback control law of the format in Equation 3.107 such that, when it is applied to the system in Equation 3.106, the resulting closed-loop system is internally stable and the H_∞ -norm of the overall closed-loop transfer matrix function from w to h , i.e. $T_{hw}(z)$, is minimized. The H_∞ -norm of a stable discrete-time transfer matrix, e.g., $T_{hw}(z)$, is defined as follows:

$$\|T_{hw}\|_\infty := \sup_{\omega \in [0, 2\pi]} \sigma_{\max} [T_{hw}(e^{j\omega})] = \sup_{\|w\|_2=1} \frac{\|h\|_2}{\|w\|_2} \quad (3.184)$$

where w and h are, respectively, the input and output of $T_{hw}(s)$, and $\|\cdot\|_2$ is the l_2 -norm of the corresponding signal. Next, we define

$$\gamma_\infty^* := \inf \left\{ \|T_{hw}(\Sigma \times \Sigma_{\text{cmp}})\|_\infty \mid \Sigma_{\text{cmp}} \text{ internally stabilizes } \Sigma \right\} \quad (3.185)$$

Again, we refer interested readers to the work of Chen [74] for the computation of γ_∞^* . For simplicity, we assume throughout this section that γ_∞^* has been determined and hence it is known.

It can be shown that the problem of H_∞ almost disturbance decoupling with internal stability is solvable for Σ of Equation 3.106, i.e. $\gamma_\infty^* = 0$, if and only if the following conditions hold (see, e.g., Chen [74]):

1. (A, B) is stabilizable;
2. (A, C_1) is detectable;
3. $D_{22} + D_2 S D_1 = 0$, where $S = -(D_2' D_2)^\dagger D_2' D_{22} D_1' (D_1 D_1')^\dagger$;
4. $\text{Im}(E + B S D_1) \subset \left\{ \mathcal{V}^\odot(\Sigma_p) + B \text{Ker}(D_2) \right\} \cap \left\{ \bigcap_{|\lambda|=1} \mathcal{S}_\lambda(\Sigma_p) \right\}$;
5. $\text{Ker}(C_2 + D_2 S C_1) \supset \left\{ \mathcal{S}^\odot(\Sigma_Q) \cap C_1^{-1} \{ \text{Im}(D_1) \} \right\} \cup \left\{ \bigcup_{|\lambda|=1} \mathcal{V}_\lambda(\Sigma_Q) \right\}$;
6. $\mathcal{S}^\odot(\Sigma_Q) \subset \mathcal{V}^\odot(\Sigma_p)$.

We note that if Σ_p is right invertible and of minimum phase with no infinite zeros, and Σ_Q is left invertible and of minimum phase with no infinite zeros, then conditions 4–6 are automatically satisfied.

The problems of discrete-time H_∞ control and H_∞ almost disturbance decoupling can be solved explicitly in the discrete-time domain. Complete solutions to these problems have been reported by Chen [74]. However, by utilizing the bilinear transformations (see Chen *et al.* [71] for the mapping of structural properties under the bilinear transformations), we can convert these discrete-time problems into equivalent continuous-time problems, and thus all algorithms presented in the previous subsection can be readily applied to yield desired solutions. The procedure is fairly simple.

1. We first apply the well-known bilinear transformation to the given discrete-time system in Equation 3.106 to obtain an equivalent continuous-time counterpart, *i.e.*

$$\left. \begin{aligned} \dot{x} &= A_d x + B_d u + E_d w \\ y &= C_{1d} x + D_{11d} u + D_{1d} w \\ h &= C_{2d} x + D_{2d} u + D_{22d} w \end{aligned} \right\} \quad (3.186)$$

where

$$\left. \begin{aligned} A_d &= (A + I)^{-1}(A - I) \\ B_d &= \sqrt{2}(A + I)^{-1}B \\ E_d &= \sqrt{2}(A + I)^{-1}E \\ C_{1d} &= \sqrt{2}C_1(A + I)^{-1} \\ D_{11d} &= D_{11} - C_1(A + I)^{-1}B \\ D_{1d} &= D_1 - C_1(A + I)^{-1}E \\ C_{2d} &= \sqrt{2}C_2(A + I)^{-1} \\ D_{2d} &= D_2 - C_2(A + I)^{-1}B \\ D_{22d} &= D_{22} - C_2(A + I)^{-1}E \end{aligned} \right\} \quad (3.187)$$

Note that if A has eigenvalues at -1 , then one can apply some prefeedback laws to remove them.

2. Next, utilize any method of the previous subsection to the system in Equation 3.186 to find an appropriate H_∞ γ -suboptimal controller, say,

$$\left. \begin{aligned} \dot{v} &= A_c v + B_c y \\ u &= C_c v + D_c y \end{aligned} \right\} \quad (3.188)$$

Note that some prefeedback might be necessary to wash out D_{22d} .

3. Lastly, apply the inverse bilinear transformation to convert the controller in Equation 3.188 to a discrete-time equivalent system. It is known in the literature (see, e.g., [99]) that the discrete-time controller obtained is H_∞ γ -suboptimal to the original discrete-time system in Equation 3.106. \diamond

The above procedure is also applicable to finding discrete-time H_∞ almost disturbance decoupling controllers. As mentioned earlier, all results presented in the previous section on H_2 optimal control are applicable to solving the H_∞ almost disturbance decoupling problem when both Σ_p and Σ_q have no invariant zeros on the unit circle.

3.6 Robust and Perfect Tracking Control

We present in this section a robust and perfect tracking (RPT) problem that was proposed and solved by Liu *et al.* [106] for continuous-time systems and Chen *et al.* [107] for discrete-time systems (see also Chen [74]). The RPT problem is to design a controller such that the resulting closed-loop system is asymptotically stable and the controlled output almost perfectly tracks a given reference signal in the presence of any initial conditions and external disturbances. By almost perfect tracking we mean the ability of a controller to track a given reference signal with an arbitrarily fast settling time in the face of external disturbances and initial conditions. Of course, in real life, a certain tradeoff has to be made in order to design a physically implementable control law. The results of this section will be utilized heavily in later chapters to solve track-following problems in hard disk drive servo systems.

3.6.1 Continuous-time Systems

Consider the following continuous-time system:

$$\Sigma : \begin{cases} \dot{x} = A x + B u + E w, & x(0) = x_0, \\ y = C_1 x + D_1 w \\ h = C_2 x + D_2 u + D_{22} w \end{cases} \quad (3.189)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, $w \in \mathbb{R}^q$ is the external disturbance, $y \in \mathbb{R}^p$ is the measurement output, and $h \in \mathbb{R}^\ell$ is the output to be controlled. We also assume that the pair (A, B) is stabilizable and (A, C_1) is detectable. For future reference, we define Σ_p and Σ_q to be the subsystems characterized by the matrix quadruples (A, B, C_2, D_2) and (A, E, C_1, D_1) , respectively. Given the external disturbance $w \in L_p$, $p \in [1, \infty)$, and any reference signal vector $r \in \mathbb{R}^\ell$ with $r, \dot{r}, \dots, r^{(\kappa-1)}$, $\kappa \geq 1$, being available, and $r^{(\kappa)}$ being either a vector of delta functions or in L_p , the RPT problem for the system in Equation 3.189 is to find a parameterized dynamic measurement control law of the following form

$$\begin{cases} \dot{v} = A_{\text{cmp}}(\varepsilon)v + B_{\text{cmp}}(\varepsilon)y + G_0(\varepsilon)r + \cdots + G_{\kappa-1}(\varepsilon)r^{(\kappa-1)} \\ u = C_{\text{cmp}}(\varepsilon)v + D_{\text{cmp}}(\varepsilon)y + H_0(\varepsilon)r + \cdots + H_{\kappa-1}(\varepsilon)r^{(\kappa-1)} \end{cases} \quad (3.190)$$

such that when the controller of Equation 3.190 is applied to the system of Equation 3.189, we have

1. there exists an $\varepsilon^* > 0$ such that the resulting closed-loop system with $r = 0$ and $w = 0$ is asymptotically stable for all $\varepsilon \in (0, \varepsilon^*]$; and
2. let $h(t, \varepsilon)$ be the closed-loop controlled output response and let $e(t, \varepsilon)$ be the resulting tracking error, *i.e.* $e(t, \varepsilon) := h(t, \varepsilon) - r(t)$. Then, for any initial condition of the state, $x_0 \in \mathbb{R}^n$,

$$\|e\|_p = \left(\int_0^\infty |e(t)|^p dt \right)^{1/p} \rightarrow 0 \quad \text{as } \varepsilon \rightarrow 0 \quad (3.191)$$

We introduce in the above formulation some additional information besides the reference signal r , *i.e.* $\dot{r}, \ddot{r}, \dots, r^{(\kappa-1)}$, as additional controller inputs. The idea of utilizing the reference signal r together with its derivatives in tracking control is similar to the command-generator system introduced in Lewis [108]. Note that, in general, these additional signals can easily be generated without any extra costs. For example, if $r(t) = t^2 \cdot \mathbf{1}(t)$, where $\mathbf{1}(t)$ is a unit step function, then one can easily obtain its first-order derivative

$$\dot{r}(t) = 2t \cdot \mathbf{1}(t) + t^2 \cdot \delta(t) = 2t \cdot \mathbf{1}(t) \quad (3.192)$$

where $\delta(t)$ is a unit impulse function, and the second-order derivative

$$\ddot{r}(t) = 2 \cdot \mathbf{1}(t) \quad (3.193)$$

These $\dot{r}(t)$ and $\ddot{r}(t)$ can be used to improve the overall tracking performance, whereas $r^{(3)}(t) = 2 \cdot \delta(t)$ does not exist in the real world and hence cannot be used. We also note that our formulation covers all possible reference signals that have the form $r(t) = t^k$, $0 \leq k < \infty$. Thus, our method could be applied to track approximately those reference signals that have a Taylor series expansion at $t = 0$. This can be done by truncating the higher-order terms of the Taylor series of the given signal. Also, it is simple to see that, when $r(t) \equiv 0$, the proposed problem reduces to the well-known perfect regulation problem with measurement feedback.

It is shown that the RPT problem for the system in Equation 3.189 is solvable if and only if the following conditions hold:

1. (A, B) is stabilizable and (A, C_1) is detectable;
2. $D_{22} + D_2 S D_1 = 0$, where $S = -(D_2' D_2)^\dagger D_2' D_{22} D_1' (D_1 D_1')^\dagger$;
3. Σ_r , *i.e.* (A, B, C_2, D_2) , is right invertible and of minimum phase;
4. $\text{Ker}(C_2 + D_2 S C_1) \supset C_1^{-1} \{\text{Im}(D_1)\}$.

We assume throughout the rest of this subsection that the above conditions are satisfied, and move on to construct solutions to the RPT problem. As usual, we focus

on the following three cases: 1) the state feedback case; 2) the full-order measurement feedback case; and 3) the reduced-order measurement feedback case.

i. State Feedback Case. When all states of the plant are measured for feedback, the problem can be solved by a static control law. We construct in this subsection a parameterized state feedback control law,

$$u = F(\varepsilon)x + H_0(\varepsilon)r + \cdots + H_{\kappa-1}(\varepsilon)r^{(\kappa-1)} \quad (3.194)$$

that solves the RPT problem for the system in Equation 3.189. It is simple to note that we can rewrite the given reference in the following form:

$$\frac{d}{dt} \begin{pmatrix} r \\ \vdots \\ r^{(\kappa-2)} \\ r^{(\kappa-1)} \end{pmatrix} = \begin{bmatrix} 0 & I_\ell & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_\ell \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{pmatrix} r \\ \vdots \\ r^{(\kappa-2)} \\ r^{(\kappa-1)} \end{pmatrix} + \begin{bmatrix} 0 \\ \vdots \\ 0 \\ I_\ell \end{bmatrix} r^{(\kappa)} \quad (3.195)$$

Combining Equation 3.195 with the given system, we obtain the following augmented system:

$$\Sigma_{\text{AUG}} : \begin{cases} \dot{\mathbf{x}} = \mathbf{A} \mathbf{x} + \mathbf{B} u + \mathbf{E} \mathbf{w} \\ \mathbf{y} = \mathbf{x} \\ e = \mathbf{C}_2 \mathbf{x} + \mathbf{D}_2 u \end{cases} \quad (3.196)$$

where

$$\mathbf{w} := \begin{pmatrix} w \\ r^{(\kappa)} \end{pmatrix}, \quad \mathbf{x} := \begin{pmatrix} r \\ \vdots \\ r^{(\kappa-2)} \\ r^{(\kappa-1)} \\ x \end{pmatrix} \quad (3.197)$$

$$\mathbf{A} = \begin{bmatrix} 0 & I_\ell & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & I_\ell & 0 \\ 0 & 0 & \cdots & 0 & 0 \\ 0 & 0 & \cdots & 0 & A \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 0 \\ B \end{bmatrix}, \quad \mathbf{E} = \begin{bmatrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & I_\ell \\ E & 0 \end{bmatrix} \quad (3.198)$$

and

$$\mathbf{C}_2 = [-I_\ell \ 0 \ 0 \ \cdots \ 0 \ C_2], \quad \mathbf{D}_2 = D_2 \quad (3.199)$$

It is then straightforward to show that the subsystem from u to e in the augmented system of Equation 3.196, *i.e.* the quadruple $(\mathbf{A}, \mathbf{B}, \mathbf{C}_2, \mathbf{D}_2)$, is right invertible and has the same infinite zero structure as that of Σ_p . Furthermore, its invariant zeros contain those of Σ_p and $\ell \times \kappa$ extra ones at $s = 0$. We are now ready to present a step-by-step algorithm to construct the required control law of the form in Equation 3.194.

STEP 3.6.C.S.1: this step transforms the subsystem from u to e of the augmented system in Equation 3.196 into the special coordinate basis of Theorem 3.1, *i.e.*

finds nonsingular state, input and output transformations Γ_s , F_i and Γ_o to put the subsystem into the structural form of Theorem 3.1 and in a small variation of the compact form of Equations 3.20 to 3.23. It can be shown that the compact form of Equations 3.20 to 3.23 for the subsystem from u to e of Equation 3.196 can be written as

$$\tilde{A} = \begin{bmatrix} A_{aa}^0 & 0 & 0 & 0 \\ 0 & A_{aa}^- & 0 & L_{ad}^- C_d \\ B_c E_{ca}^0 & B_c E_{ca}^- & A_{cc} & L_{cd} C_d \\ B_d E_{da}^0 & B_d E_{da}^- & B_d E_{dc} & A_{dd} \end{bmatrix} \quad (3.200)$$

$$A_{aa}^0 = \begin{bmatrix} 0 & I_\ell & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & I_\ell \\ 0 & 0 & \cdots & 0 \end{bmatrix}, \quad \tilde{B} = \begin{bmatrix} 0 & 0 & 0 \\ B_{0a}^- & 0 & 0 \\ B_{0c} & 0 & B_c \\ B_{0d} & B_d & 0 \end{bmatrix} \quad (3.201)$$

and

$$\tilde{C} = \begin{bmatrix} C_{0a}^0 & C_{0a}^- & C_{0c} & C_{0d} \\ 0 & 0 & 0 & C_d \end{bmatrix}, \quad \tilde{D} = \begin{bmatrix} I_{m_0} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (3.202)$$

STEP 3.6.C.S.2: choose an appropriate dimensional matrix F_c such that

$$A_{cc}^c = A_{cc} - B_c F_c \quad (3.203)$$

is asymptotically stable. The existence of such an F_c is guaranteed by the property that (A_{cc}, B_c) is completely controllable.

STEP 3.6.C.S.3: for each x_i of x_d , which is associated with the infinite zero structure of Σ_p or the subsystem from u to e of Equation 3.196, we choose an F_i such that

$$p_i(s) = \prod_{j=1}^{q_i} (s - \lambda_{ij}) = s^{q_i} + F_{i1} s^{q_i-1} + \cdots + F_{iq_i-1} s + F_{iq_i} \quad (3.204)$$

with all λ_{ij} being in \mathbb{C}^- . Let

$$F_i = [F_{iq_i} \quad F_{iq_i-1} \quad \cdots \quad F_{i1}], \quad i = 1, \dots, m_d \quad (3.205)$$

STEP 3.6.C.S.4: next, we construct

$$F(\varepsilon) = -\Gamma_i \begin{bmatrix} C_{0a}^0 & C_{0a}^- & C_{0c} & C_{0d} \\ E_{da}^0 & E_{da}^- & E_{dc} & E_d + F_d(\varepsilon) \\ E_{ca}^0 & E_{ca}^- & F_c & 0 \end{bmatrix} \Gamma_s^{-1} \quad (3.206)$$

where

$$E_d = \begin{bmatrix} E_{11} & \cdots & E_{1m_d} \\ \vdots & \ddots & \vdots \\ E_{m_d 1} & \cdots & E_{m_d m_d} \end{bmatrix} \quad (3.207)$$

$$F_d(\varepsilon) = \text{blkdiag} \left\{ \frac{F_1}{\varepsilon^{q_1}} S_1(\varepsilon), \frac{F_2}{\varepsilon^{q_2}} S_2(\varepsilon), \dots, \frac{F_{m_d}}{\varepsilon^{q_{m_d}}} S_{m_d}(\varepsilon) \right\} \quad (3.208)$$

and where

$$S_i(\varepsilon) = \text{diag} \{1, \varepsilon, \varepsilon^2, \dots, \varepsilon^{q_i-1}\} \quad (3.209)$$

STEP 3.6.C.S.5: finally, we partition

$$F(\varepsilon) = [H_0(\varepsilon) \quad \dots \quad H_{\kappa-1}(\varepsilon) \quad F(\varepsilon)] \quad (3.210)$$

where $H_i(\varepsilon) \in \mathbb{R}^{m \times \ell}$ and $F(\varepsilon) \in \mathbb{R}^{m \times n}$. \diamond

We have the following result.

Theorem 3.20. *Consider the given system in Equation 3.189 with its external disturbance $w \in L_p$, $p \in [1, \infty)$, its initial condition $x(0) = x_0$. Assume that all its states are measured for feedback, i.e. $C_1 = I$ and $D_1 = 0$, and the RPT problem for the system in Equation 3.189 is solvable. Then, for any reference signal $r(t)$, which has all its i th-order derivatives, $i = 0, 1, \dots, \kappa - 1$, $\kappa \geq 1$, being available and $r^{(\kappa)}(t)$ being either a vector of delta functions or in L_p , the RPT problem is solved by the control law of Equation 3.194 with $F(\varepsilon)$ and $H_i(\varepsilon)$, $i = 0, 1, \dots, \kappa - 1$, as given in Equation 3.210. \diamond*

The following remark gives an alternative approach for solving the proposed RPT problem via full state feedback. We leave the proof of this method to readers as an exercise.

Remark 3.21. Note that the required gain matrices for the state feedback RPT problem might be computed by solving the following Riccati equation:

$$P\tilde{A} + \tilde{A}'P + \tilde{C}'_2\tilde{C}_2 - (PB + \tilde{C}'_2\tilde{D}_2)(\tilde{D}'_2\tilde{D}_2)^{-1}(PB + \tilde{C}'_2\tilde{D}_2)' = 0$$

for a positive-definite solution $P > 0$, where

$$\tilde{C}_2 = \begin{bmatrix} C_2 \\ \varepsilon I_{\kappa\ell+n} \\ 0 \end{bmatrix}, \quad \tilde{D}_2 = \begin{bmatrix} D_2 \\ 0 \\ \varepsilon I_m \end{bmatrix} \quad (3.211)$$

$$\tilde{A} = \begin{bmatrix} \tilde{A}_0 & 0 \\ 0 & A \end{bmatrix}, \quad \tilde{A}_0 = -\varepsilon I_{\kappa\ell} + \begin{bmatrix} 0 & I_\ell & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & I_\ell \\ 0 & 0 & \dots & 0 \end{bmatrix} \quad (3.212)$$

and where B , C_2 and D_2 are as defined in Equations 3.198 and 3.199. The required gain matrix is then given by

$$\begin{aligned} \tilde{F}(\varepsilon) &= -(\tilde{D}'_2\tilde{D}_2)^{-1}(PB + \tilde{C}'_2\tilde{D}_2)' \\ &= [H_0(\varepsilon) \quad \dots \quad H_{\kappa-1}(\varepsilon) \quad F(\varepsilon)] \end{aligned} \quad (3.213)$$

where $H_i(\varepsilon) \in \mathbb{R}^{m \times \ell}$ and $F(\varepsilon) \in \mathbb{R}^{m \times n}$. \diamond

Finally, we note that solutions to the Riccati equation in Remark 3.21 might have severe numerical problems as ε becomes smaller and smaller.

We consider two types of measurement feedback control laws: one is of full-order controllers whose dynamical order is equal to the order of the given system; the other is of reduced-order controllers with a dynamical order that is less than the order of the given system. Without loss of generality, we assume throughout this subsection that $D_{22} = 0$. If it is nonzero, it can always be washed out by the following preoutput feedback: $u = Sy$.

ii. Full-order Output Feedback Case. The following is a step-by-step algorithm for constructing a parameterized full-order measurement feedback controller, which solves the RPT problem.

STEP 3.6.C.F.1: for the given reference $r(t)$ and the given system in Equation 3.189, we first assume that all the state variables of Equation 3.189 are measurable and follow the procedures of the state feedback case to define an auxiliary system,

$$\begin{cases} \dot{x} = A x + B u + E w \\ y = x \\ e = C_2 x + D_2 u \end{cases} \quad (3.214)$$

Then, we follow Steps 3.6.C.S.1 to 3.6.C.S.5 of the algorithm of the state feedback case to construct a state feedback gain matrix:

$$F(\varepsilon) = [H_0(\varepsilon) \quad \cdots \quad H_{\kappa-1}(\varepsilon) \quad F(\varepsilon)] \quad (3.215)$$

STEP 3.6.C.F.2: let Σ_{Qaux} be characterized by a matrix quadruple

$$(A_{\text{Qaux}}, E_{\text{Qaux}}, C_{\text{Qaux}}, D_{\text{Qaux}}) := (A, [E \quad I_n], C_1, [D_1 \quad 0]) \quad (3.216)$$

This step is to transform this Σ_{Qaux} into the special coordinate basis of Theorem 3.1. Because of the special structure of the matrix E_{Qaux} , it is simple to show that Σ_{Qaux} is always right invertible and is free of invariant zeros. Utilize the results of Theorem 3.1 to find nonsingular state, input and output transformation Γ_{sq} , Γ_{iq} and Γ_{oq} such that

$$\Gamma_{\text{sq}}^{-1} A \Gamma_{\text{sq}} = \begin{bmatrix} A_{\text{ccq}} & L_{\text{cdq}} \\ E_{\text{dcq}} & A_{\text{ddq}} \end{bmatrix} + \begin{bmatrix} B_{\text{0cq}} \\ B_{\text{0dq}} \end{bmatrix} [C_{\text{0cq}} \quad 0] \quad (3.217)$$

$$\Gamma_{\text{sq}}^{-1} E_{\text{Qaux}} \Gamma_{\text{iq}} = \begin{bmatrix} B_{\text{0cq}} & 0 & I_{n-k} & 0 \\ B_{\text{0dq}} & I_k & 0 & 0 \end{bmatrix} \quad (3.218)$$

$$\Gamma_{\text{oq}}^{-1} C_1 \Gamma_{\text{sq}} = \begin{bmatrix} C_{\text{0cq}} & 0 \\ 0 & I_k \end{bmatrix} \quad (3.219)$$

$$\Gamma_{\text{oq}}^{-1} [D_1 \quad 0] \Gamma_{\text{iq}} = \begin{bmatrix} I_{p-k} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad (3.220)$$

where $k = p - \text{rank}(D_1)$. It can be verified that the pair (A, C_1) is detectable if and only if the pair

$$\left(A_{ccq}, \begin{bmatrix} C_{0cq} \\ E_{dcq} \end{bmatrix} \right) \quad (3.221)$$

is detectable.

STEP 3.6.C.F.3: let K_{cq} be an appropriate dimensional constant matrix such that the eigenvalues of the matrix

$$A_{ccq}^c = A_{ccq} - K_{cq} \begin{bmatrix} C_{0cq} \\ E_{dcq} \end{bmatrix} = A_{ccq} - [K_{cd0} \quad K_{cdq}] \begin{bmatrix} C_{0cq} \\ E_{dcq} \end{bmatrix} \quad (3.222)$$

are all in \mathbb{C}^- . Next, we define a parameterized observer gain matrix,

$$K(\varepsilon) = -\Gamma_{sq} \begin{bmatrix} B_{0cq} + K_{cd0} & L_{cdq} + K_{cdq}/\varepsilon \\ B_{0dq} & A_{ddq} + I_k/\varepsilon \end{bmatrix} \Gamma_{oo}^{-1} \quad (3.223)$$

STEP 3.6.C.F.4: finally, we obtain the following full-order measurement feedback control law:

$$\begin{cases} \dot{v} = A_{cmp}(\varepsilon)v - K(\varepsilon)y + BH_0(\varepsilon)r + \cdots + BH_{\kappa-1}(\varepsilon)r^{(\kappa-1)} \\ u = F(\varepsilon)v + H_0(\varepsilon)r + \cdots + H_{\kappa-1}(\varepsilon)r^{(\kappa-1)} \end{cases} \quad (3.224)$$

where $A_{cmp}(\varepsilon) = A + BF(\varepsilon) + K(\varepsilon)C_1$. This completes the construction of the full-order measurement feedback controller. \diamond

We have the following theorem.

Theorem 3.22. Consider the given system in Equation 3.189 with its external disturbance $w \in L_p$, $p \in [1, \infty)$, its initial condition $x(0) = x_0$. Assume that the RPT problem is solvable for the system in Equation 3.189. Then, for any reference signal $r(t)$, which has all its i th-order derivatives, $i = 0, 1, \dots, \kappa - 1$, $\kappa \geq 1$, being available and $r^{(\kappa)}(t)$ being either a vector of delta functions or in L_p , the RPT problem is solved by the parameterized full-order measurement feedback control laws as given in Equation 3.224. \diamond

The following remark yields an alternative way to compute the gain matrix $K(\varepsilon)$ in Step 3.6.C.F.3.

Remark 3.23. The gain matrix $K(\varepsilon)$ in Step 3.6.C.F.3 can also be computed by solving the following Riccati equation:

$$AQ + QA' + (EE' + I) - (QC'_1 + ED'_1)(D_1D'_1 + \varepsilon I)^{-1}(C_1Q + D_1E') = 0 \quad (3.225)$$

for a positive-definite solution $Q > 0$. The required gain matrix $K(\varepsilon)$ is then given by

$$K(\varepsilon) = -(QC'_1 + ED'_1)(D_1D'_1 + \varepsilon I)^{-1} \quad (3.226)$$

Again, this approach might have some numerical problems. \diamond

iii. Reduced-order Output Feedback Case. We now present solutions to the RPT problem via reduced-order measurement feedback control laws. For simplicity of presentation, we assume that matrices C_1 and D_1 have already been transformed into the following forms:

$$C_1 = \begin{bmatrix} 0 & C_{1,02} \\ I_k & 0 \end{bmatrix} \quad \text{and} \quad D_1 = \begin{bmatrix} D_{1,0} \\ 0 \end{bmatrix} \quad (3.227)$$

where $D_{1,0}$ is of full row rank. Before we present a step-by-step algorithm to construct a parameterized reduced-order measurement feedback controller, we first partition the following system

$$\begin{cases} \dot{x} = A x + B u + [E & I_n] \tilde{w} \\ y = C_1 x & + [D_1 & 0] \tilde{w} \end{cases} \quad (3.228)$$

in conformity with the structures of C_1 and D_1 in Equation 3.227, *i.e.*

$$\begin{cases} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u + \begin{bmatrix} E_1 & I_k & 0 \\ E_2 & 0 & I_{n-k} \end{bmatrix} \tilde{w} \\ \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{bmatrix} 0 & C_{1,02} \\ I_k & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} D_{1,0} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tilde{w} \end{cases} \quad (3.229)$$

where

$$\tilde{w} = \begin{pmatrix} w \\ x_0 \cdot \delta(t) \end{pmatrix} \quad (3.230)$$

Obviously, $y_1 = x_1$ is directly available and hence need not be estimated. Next, we define Σ_{QR} to be characterized by

$$(A_R, E_R, C_R, D_R) = \left(A_{22}, [E_2 \ 0 \ I_{n-k}], \begin{bmatrix} C_{1,02} \\ A_{12} \end{bmatrix}, \begin{bmatrix} D_{1,0} & 0 & 0 \\ E_1 & I_k & 0 \end{bmatrix} \right)$$

It is again straightforward to verify that Σ_{QR} is right invertible with no finite and infinite zeros. Moreover, (A_R, C_R) is detectable if and only if (A, C_1) is detectable. We are ready to present the following algorithm.

STEP 3.6.C.R.1: for the given reference $r(t)$ and the given system in Equation 3.189, we again assume that all the state variables of the given system are measurable and follow the procedures of the state feedback case to define an auxiliary system,

$$\begin{cases} \dot{x} = A x + B u + E w \\ y = x \\ e = C_2 x + D_2 u \end{cases} \quad (3.231)$$

Then, we follow Steps 3.6.C.S.1 to 3.6.C.S.5 of the algorithm of the state feedback case to construct a state feedback gain matrix

$$F(\varepsilon) = [H_0(\varepsilon) \ \cdots \ H_{\kappa-1}(\varepsilon) \ F(\varepsilon)] \quad (3.232)$$

Let us partition $F(\varepsilon)$ in conformity with x_1 and x_2 of Equation 3.229 as follows,

$$F(\varepsilon) = [F_1(\varepsilon) \ F_2(\varepsilon)] \quad (3.233)$$

STEP 3.6.C.R.2: let K_R be an appropriate dimensional constant matrix such that the eigenvalues of

$$A_R + K_R C_R = A_{22} + [K_{R0} \quad K_{R1}] \begin{bmatrix} C_{1,02} \\ A_{12} \end{bmatrix} \quad (3.234)$$

are all in \mathbb{C}^- . This can be done because (A_R, C_R) is detectable.

STEP 3.6.C.R.3: let

$$G_R = [-K_{R0}, \quad A_{21} + K_{R1}A_{11} - (A_R + K_R C_R)K_{R1}] \quad (3.235)$$

and

$$\left. \begin{aligned} A_{\text{cmp}}(\varepsilon) &= A_R + B_2 F_2(\varepsilon) + K_R C_R + K_{R1} B_1 F_2(\varepsilon) \\ B_{\text{cmp}}(\varepsilon) &= G_R + (B_2 + K_{R1} B_1) [0, \quad F_1(\varepsilon) - F_2(\varepsilon) K_{R1}] \\ C_{\text{cmp}}(\varepsilon) &= F_2(\varepsilon) \\ D_{\text{cmp}}(\varepsilon) &= [0, \quad F_1(\varepsilon) - F_2(\varepsilon) K_{R1}] \end{aligned} \right\} \quad (3.236)$$

STEP 3.6.C.R.4: finally, we obtain the following reduced-order measurement feedback control law:

$$\begin{cases} \dot{v} = A_{\text{cmp}}(\varepsilon)v + B_{\text{cmp}}(\varepsilon)y + G_0(\varepsilon)r + \cdots + G_{\kappa-1}(\varepsilon)r^{(\kappa-1)} \\ u = C_{\text{cmp}}(\varepsilon)v + D_{\text{cmp}}(\varepsilon)y + H_0(\varepsilon)r + \cdots + H_{\kappa-1}(\varepsilon)r^{(\kappa-1)} \end{cases} \quad (3.237)$$

where for $i = 0, 1, \dots, \kappa - 1$, $G_i(\varepsilon) = (B_2 + K_{R1} B_1)H_i(\varepsilon)$. This completes the constructing procedure. \diamond

Theorem 3.24. Consider the given system in Equation 3.189 with its external disturbance $w \in L_p$, $p \in [1, \infty)$, its initial condition $x(0) = x_0$. Assume that the RPT problem is solvable for the system in Equation 3.189. Then, for any reference signal $r(t)$, which has all its i th-order derivatives, $i = 0, 1, \dots, \kappa - 1$, $\kappa \geq 1$, being available and $r^{(\kappa)}(t)$ being either a vector of delta functions or in L_p , the RPT problem is solved by the parameterized reduced-order measurement feedback control laws of Equation 3.237. \diamond

3.6.2 Discrete-time Systems

We present in this subsection the RPT problem for the following discrete-time system:

$$\Sigma : \begin{cases} x(k+1) = A x(k) + B u(k) + E w(k), & x(0) = x_0 \\ y(k) = C_1 x(k) + D_1 w(k) \\ h(k) = C_2 x(k) + D_2 u(k) + D_{22} w(k) \end{cases} \quad (3.238)$$

where $x \in \mathbb{R}^n$ is the state, $u \in \mathbb{R}^m$ is the control input, $w \in \mathbb{R}^q$ is the external disturbance, $y \in \mathbb{R}^p$ is the measurement output, and $h \in \mathbb{R}^\ell$ is the output to be controlled.

We also assume that the pair (A, B) is stabilizable and (A, C_1) is detectable. For future reference, we define Σ_p and Σ_q to be the subsystems characterized by the matrix quadruples (A, B, C_2, D_2) and (A, E, C_1, D_1) respectively. Given the external disturbance $w \in L_p$, $p \in [1, \infty]$, and any reference signal vector $r \in \mathbb{R}^\ell$, the RPT problem for the discrete-time system in Equation 3.238 is to find a parameterized dynamic measurement feedback control law of the following form:

$$\begin{cases} v(k+1) = A_{\text{cmp}}(\varepsilon)v(k) + B_{\text{cmp}}(\varepsilon)y(k) + G(\varepsilon)r(k) \\ u(k) = C_{\text{cmp}}(\varepsilon)v(k) + D_{\text{cmp}}(\varepsilon)y(k) + H(\varepsilon)r(k) \end{cases} \quad (3.239)$$

such that, when the controller in Equation 3.239 is applied to the system in Equation 3.238,

1. there exists an $\varepsilon^* > 0$ such that the resulting closed-loop system with $r = 0$ and $w = 0$ is asymptotically stable for all $\varepsilon \in (0, \varepsilon^*]$; and
2. let $h(k, \varepsilon)$ be the closed-loop controlled output response and let $e(k, \varepsilon)$ be the resulting tracking error, *i.e.* $e(k, \varepsilon) := h(k, \varepsilon) - r(k)$. Then, for any initial condition of the state, $x_0 \in \mathbb{R}^n$, $\|e\|_p \rightarrow 0$ as $\varepsilon \rightarrow 0$.

It has been shown by Chen [74] that the above RPT problem is solvable for the system in Equation 3.238 if and only if the following conditions hold:

1. (A, B) is stabilizable and (A, C_1) is detectable;
2. $D_{22} + D_2SD_1 = 0$, where $S = -(D_2'D_2)^\dagger D_2'D_{22}D_1'(D_1D_1')^\dagger$;
3. Σ_p is right invertible and of minimum phase with no infinite zeros;
4. $\text{Ker}(C_2 + D_2SC_1) \supset C_1^{-1}\{\text{Im}(D_1)\}$.

It turns out that the control laws, which solve the RPT for the given plant in Equation 3.238 under the solvability conditions, need not be parameterized by any tuning parameter. Thus, Equation 3.239 can be replaced by

$$\begin{cases} v(k+1) = A_{\text{cmp}}v(k) + B_{\text{cmp}}y(k) + Gr(k) \\ u(k) = C_{\text{cmp}}v(k) + D_{\text{cmp}}y(k) + Hr(k) \end{cases} \quad (3.240)$$

and, furthermore, the resulting tracking error $e(k)$ can be made identically zero for all $k \geq 0$.

Assume that all the solvability conditions are satisfied. We present in the following solutions to the discrete-time RPT problem.

i. State Feedback Case. When all states of the plant are measured for feedback, the problem can be solved by a static control law. We construct in this subsection a state feedback control law,

$$u = Fx + Hr \quad (3.241)$$

that solves the RPT problem for the system in Equation 3.238. We have the following algorithm.

STEP 3.6.D.S.1: this step transforms the subsystem from u to h of the given system in Equation 3.238 into the special coordinate basis of Theorem 3.1, *i.e.* finds

nonsingular state, input and output transformations Γ_s , Γ_i and Γ_o to put it into the structural form of Theorem 3.1 as well as in the compact form of Equations 3.20 to 3.23, *i.e.*

$$\tilde{A} = \Gamma_s^{-1} A \Gamma_s = \begin{bmatrix} A_{aa}^- & 0 \\ B_c E_{ca}^- & A_{cc} \end{bmatrix} + B_0 C_{2,0} \quad (3.242)$$

$$\tilde{B} = \Gamma_s^{-1} B \Gamma_i = \Gamma_s^{-1} [B_0 \quad B_1] \Gamma_i = \begin{bmatrix} B_{0a}^- & 0 \\ B_{0c}^- & B_c \end{bmatrix} \quad (3.243)$$

$$\tilde{D}_2 = \Gamma_o^{-1} D_2 \Gamma_i = [I_{m_0} \quad 0] \quad (3.244)$$

$$\tilde{C}_2 = \Gamma_o^{-1} C_2 \Gamma_s = \Gamma_o^{-1} C_{2,0} \Gamma_s = [C_{2,0a}^- \quad C_{2,0c}] \quad (3.245)$$

STEP 3.6.D.S.2: choose an appropriate dimensional matrix F_c such that

$$A_{cc}^c = A_{cc} - B_c F_c \quad (3.246)$$

is asymptotically stable. The existence of such an F_c is guaranteed by the property that (A_{cc}, B_c) is completely controllable.

STEP 3.6.D.S.3: finally, we let

$$F = -\Gamma_i \begin{bmatrix} C_{2,0a}^- & C_{2,0c} \\ E_{ca}^- & F_c \end{bmatrix} \Gamma_s^{-1} \quad \text{and} \quad H = \Gamma_i \begin{bmatrix} I \\ 0 \end{bmatrix} \Gamma_o^{-1} \quad (3.247)$$

This ends the constructive algorithm. \diamond

We have the following result.

Theorem 3.25. *Consider the given discrete-time system in Equation 3.238 with any external disturbance $w(k)$ and any initial condition $x(0)$. Assume that all its states are measured for feedback, *i.e.* $C_1 = I$ and $D_1 = 0$, and the solvability conditions for the RPT problem hold. Then, for any reference signal $r(k)$, the proposed RPT problem is solved by the control law of Equation 3.241 with F and H as given in Equation 3.247. \diamond*

ii. Measurement Feedback Case. Without loss of generality, we assume throughout this subsection that matrix $D_{22} = 0$. If it is nonzero, it can always be washed out by the following preoutput feedback $u(k) = Sy(k)$. It turns out that, for discrete-time systems, the full-order observer-based control law is not capable of achieving the RPT performance, because there is a delay of one step in the observer itself. Thus, we focus on the construction of a reduced-order measurement feedback control law to solve the RPT problem. For simplicity of presentation, we assume that matrices C_1 and D_1 have already been transformed into the following forms,

$$C_1 = \begin{bmatrix} 0 & C_{1,02} \\ I_k & 0 \end{bmatrix} \quad \text{and} \quad D_1 = \begin{bmatrix} D_{1,0} \\ 0 \end{bmatrix} \quad (3.248)$$

where $D_{1,0}$ is of full row rank. Before we present a step-by-step algorithm to construct a reduced-order measurement feedback controller, we first partition the following system

$$\begin{cases} x(k+1) = A x(k) + B u(k) + [E \quad I_n] \tilde{w}(k) \\ y(k) = C_1 x(k) + [D_1 \quad 0] \tilde{w}(k) \end{cases} \quad (3.249)$$

in conformity with the structures of C_1 and D_1 in Equation 3.248, *i.e.*

$$\begin{cases} \begin{pmatrix} \delta(x_1) \\ \delta(x_2) \end{pmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u + \begin{bmatrix} E_1 & I_\kappa & 0 \\ E_2 & 0 & I_{n-\kappa} \end{bmatrix} \tilde{w} \\ \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{bmatrix} 0 & C_{1,02} \\ I_\kappa & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} D_{1,0} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \tilde{w} \end{cases}$$

where $\delta(x_1) = x_1(k+1)$ and $\delta(x_2) = x_2(k+1)$. Obviously, $y_1 = x_1$ is directly available and hence need not be estimated. Next, let Σ_{QR} be characterized by

$$(A_R, E_R, C_R, D_R) = \left(A_{22}, [E_2 \quad 0 \quad I_{n-\kappa}], \begin{bmatrix} C_{1,02} \\ A_{12} \end{bmatrix}, \begin{bmatrix} D_{1,0} & 0 & 0 \\ E_1 & I_\kappa & 0 \end{bmatrix} \right)$$

It is straightforward to verify that Σ_{QR} is right invertible with no finite and infinite zeros. Moreover, (A_R, C_R) is detectable if and only if (A, C_1) is detectable. We are ready to present the following algorithm.

STEP 3.6.D.R.1: for the given system in Equation 3.238, we again assume that all the state variables of the given system are measurable and then follow Steps 3.6.D.S.1 to 3.6.D.S.3 of the algorithm of the previous subsection to construct gain matrices F and H . We also partition F in conformity with x_1 and x_2 as follows:

$$F = [F_1 \quad F_2] \quad (3.250)$$

STEP 3.6.D.R.2: let K_R be an appropriate dimensional constant matrix such that the eigenvalues of

$$A_R + K_R C_R = A_{22} + [K_{R0} \quad K_{R1}] \begin{bmatrix} C_{1,02} \\ A_{12} \end{bmatrix} \quad (3.251)$$

are all in \mathbb{C}^\ominus . This can be done because (A_R, C_R) is detectable.

STEP 3.6.D.R.3: let

$$G_R = [-K_{R0}, \quad A_{21} + K_{R1}A_{11} - (A_R + K_R C_R)K_{R1}] \quad (3.252)$$

$$\left. \begin{aligned} A_{\text{cmp}} &= A_R + B_2 F_2 + K_R C_R + K_{R1} B_1 F_2 \\ B_{\text{cmp}} &= G_R + (B_2 + K_{R1} B_1) [0, \quad F_1 - F_2 K_{R1}] \\ C_{\text{cmp}} &= F_2 \\ D_{\text{cmp}} &= [0, \quad F_1 - F_2 K_{R1}] \end{aligned} \right\} \quad (3.253)$$

and

$$G = (B_2 + K_{R1} B_1) H \quad (3.254)$$

STEP 3.6.D.R.4: finally, we obtain the following reduced-order measurement feedback control law:

$$\begin{cases} v(k+1) = A_{\text{cmp}} v(k) + B_{\text{cmp}} y(k) + G r(k) \\ u(k) = C_{\text{cmp}} v(k) + D_{\text{cmp}} y(k) + H r(k) \end{cases} \quad (3.255)$$

This completes the algorithm. \diamond

Theorem 3.26. *Consider the given system in Equation 3.238 with any external disturbance $w(k)$ and any initial condition $x(0)$. Assume that the solvability conditions for the RPT problem hold. Then, for any reference signal $r(k)$, the proposed RPT problem is solved by the reduced-order measurement feedback control laws of Equation 3.255. \diamond*

3.7 Loop Transfer Recovery Technique

Another popular design methodology for multivariable systems, which is based on the ‘loop shaping’ concept, is linear quadratic Gaussian (LQG) with loop transfer recovery (LTR). It involves two separate designs of a state feedback controller and an observer or an estimator. The exact design procedure depends on the point where the unstructured uncertainties are modeled and where the loop is broken to evaluate the open-loop transfer matrices. Commonly, either the input point or the output point of the plant is taken as such a point. We focus on the case when the loop is broken at the input point of the plant. The required results for the output point can be easily obtained by appropriate dualization. Thus, in the two-step procedure of LQG/LTR, the first step of design involves loop shaping by a state feedback design to obtain an appropriate loop transfer function, called the target loop transfer function. Such a loop shaping is an engineering art and often involves the use of linear quadratic regulator (LQR) design, in which the cost matrices are used as free design parameters to generate the target loop transfer function, and thus the desired sensitivity and complementary sensitivity functions. However, when such a feedback design is implemented via an observer-based controller (or Kalman filter) that uses only the measurement feedback, the loop transfer function obtained, in general, is not the same as the target loop transfer function, unless proper care is taken in designing the observers. This is when the second step of LQG/LTR design philosophy comes into the picture. In this step, the required observer design is attempted so as to recover the loop transfer function of the full state feedback controller. This second step is known as LTR.

The topic of LTR was heavily studied in the 1980s. Major contributions came from [109–119]. We present in the following the methods of LTR design at both the input point and output point of the given plant.

3.7.1 LTR at Input Point

It turns out that it is very simple to formulate the LTR design technique for both continuous- and discrete-time systems into a single framework. Thus, we do it in one

shot. Let us consider a linear time-invariant multivariable system characterized by

$$\Sigma : \begin{cases} \delta(x) = A x + B u \\ y = C x + D u \end{cases} \quad (3.256)$$

where $\delta(x) = \dot{x}(t)$, if Σ is a continuous-time system, or $\delta(x) = x(k+1)$, if Σ is a discrete-time system. Similarly, $x \in \mathbb{R}^n$, $u \in \mathbb{R}^m$ and $y \in \mathbb{R}^p$ are the state, input and output of Σ . They represent, respectively, $x(t)$, $u(t)$ and $y(t)$ if the given system is of continuous-time, or represent, respectively, $x(k)$, $u(k)$ and $y(k)$ if Σ is of discrete-time. Without loss of any generality, we assume throughout this section that both $[B' \ D']$ and $[C \ D]$ are of full rank. The transfer function of Σ is then given by

$$P(\varsigma) = C(\varsigma I - A)^{-1}B + D \quad (3.257)$$

where $\varsigma = s$, the Laplace transform operator, if Σ is of continuous-time, or $\varsigma = z$, the z -transform operator, if Σ is of discrete-time.

As mentioned earlier, there are two steps involved in LQG/LTR design. In the first step, we assume that all state variables of the system in Equation 3.256 are available and design a full state feedback control law

$$u = -F x \quad (3.258)$$

such that

1. the closed-loop system is asymptotically stable, and
2. the open-loop transfer function when the loop is broken at the input point of the given system, *i.e.*

$$L_t(\varsigma) = F(\varsigma I - A)^{-1}B \quad (3.259)$$

meets some frequency-dependent specifications.

Arriving at an appropriate value for F is concerned with the issue of loop shaping, which often includes the use of LQR design in which the cost matrices are used as free design parameters to generate $L_t(\varsigma)$ that satisfies the given specifications.

To be more specific, if Σ is a continuous-time system, the target loop transfer function $L_t(s)$ can be generated by minimizing the following cost function:

$$J_c = \int_0^{\infty} (x' Q x + u' R u) dt \quad (3.260)$$

where $Q \geq 0$ and $R > 0$ are free design parameters provided that $(A, Q^{1/2})$ has no unobservable modes on the imaginary axis. The solution to the above problem is given by

$$F = R^{-1}B'P \quad (3.261)$$

where $P \geq 0$ is the stabilizing solution of the following algebraic Riccati equation (ARE):

$$PA + A'P - PBR^{-1}B'P + Q = 0 \quad (3.262)$$

It is known in the literature that a target loop transfer function $L_t(s)$ with F given as in Equation 3.261 has a phase margin greater than 60° and an infinite gain margin.

Similarly, if Σ is a discrete-time system, we can generate a target loop transfer function $L_t(z)$ by minimizing

$$J_D = \sum_{k=0}^{\infty} \left(x'(k)Qx(k) + u'(k)Ru(k) \right) \quad (3.263)$$

where $Q \geq 0$ and $R > 0$ are free design parameters provided that $(A, Q^{1/2})$ has no unobservable modes on the unit circle.

$$F = (R + B'PB)^{-1}B'PA \quad (3.264)$$

where $P \geq 0$ is the stabilizing solution of the following ARE:

$$P = A'PA - A'PB(R + B'PB)^{-1}B'PA + Q \quad (3.265)$$

Unfortunately, there are no guaranteed phase and gain margins for the target loop transfer function $L_t(z)$ resulting from the discrete-time linear quadratic regulator.

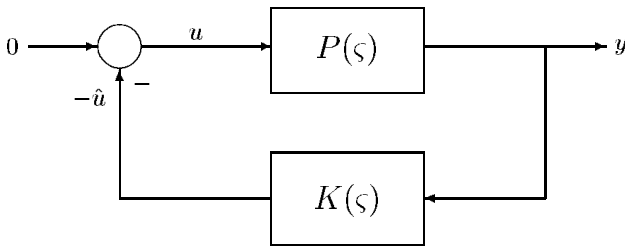


Figure 3.5. Plant-controller closed-loop configuration

Generally, it is unreasonable to assume that all the state variables of a given system can be measured. Thus, we have to implement the control law obtained in the first step by a measurement feedback controller. The technique of LTR is to design an appropriate measurement feedback control (see Figure 3.5) such that the resulting system is asymptotically stable and the achieved open-loop transfer function $L_a(s)$ from u to $-\hat{u}$ is either exactly or approximately matched with the target loop transfer function $L_t(s)$ obtained in the first step. In this way, all the nice properties associated with the target loop transfer function can be recovered by the measurement feedback controller. This is the so-called LTR design.

It is simple to observe that the achieved open-loop transfer function in the configuration of Figure 3.5 is given by

$$L_a(s) = K(s)P(s) \quad (3.266)$$

Let us define *recovery error* as

$$E(\varsigma) := L_t(\varsigma) - L_a(\varsigma) = F(\varsigma I - A)^{-1}B - K(\varsigma)P(\varsigma) \quad (3.267)$$

The LTR technique is to design an appropriate stabilizing $K(\varsigma)$ such that the recovery error $E(\varsigma)$ is either identically zero or small in a certain sense. As usual, two commonly used structures for $K(\varsigma)$ are: 1) the full-order observer-based controller, and 2) the reduced-order observer-based controller.

i. Full-order Observer-based Controller. The dynamic equations of a full-order observer-based controller are well known and are given by

$$\begin{cases} \delta(\hat{x}) = A\hat{x} + Bu + K_f(y - C\hat{x} - Du) \\ u = -F\hat{x} \end{cases} \quad (3.268)$$

where K_f is the full-order observer gain matrix and is the only free design parameter. It is chosen so that $A - K_f C$ is asymptotically stable. The transfer function of the full-order observer-based control is given by

$$K(\varsigma) = F(\varsigma I - A + BF + K_f C - K_f DF)^{-1}K_f \quad (3.269)$$

It has been shown [110, 117] that the recovery error resulting from the full-order observer-based controller can be expressed as

$$E(\varsigma) = M_f(\varsigma)[I + M_f(\varsigma)]^{-1}[I + F(\varsigma I - A)^{-1}B] \quad (3.270)$$

where

$$M_f(\varsigma) = F(\varsigma I - A + K_f C)^{-1}(B - K_f D) \quad (3.271)$$

Obviously, in order to render $E(\varsigma)$ to be zero or small, one has to design an observer gain K_f such that $M_f(\varsigma)$, or equivalently $M_f'(\varsigma)$, is zero or small (in a certain sense). Defining an auxiliary system,

$$\begin{cases} \dot{\tilde{x}} = A' \tilde{x} + C' \tilde{u} + F' \tilde{w} \\ \tilde{y} = \tilde{x} \\ \tilde{h} = B' \tilde{x} + D' \tilde{u} \end{cases} \quad (3.272)$$

with a state feedback control law,

$$\tilde{u} = -K_f' \tilde{x} \quad (3.273)$$

It is straightforward to verify that the closed-loop transfer matrix from \tilde{w} to \tilde{h} of the above system is equivalent to $M_f'(\varsigma)$. As such, any of the methods presented in Sections 3.4 and 3.5 for H_2 and H_∞ optimal control can be utilized to find K_f to minimize either the H_2 -norm or H_∞ -norm of $M_f'(\varsigma)$. In particular,

1. if the given plant Σ is a continuous-time system and if Σ is left invertible and of minimum phase,
2. if the given plant Σ is a discrete-time system and if Σ is left invertible and of minimum phase with no infinite zeros,

then either the H_2 -norm or H_∞ -norm of $M_f(\zeta)$ can be made arbitrarily small, and hence LTR can be achieved. If these conditions are not satisfied, the target loop transfer function $L_t(\zeta)$, in general, cannot be fully recovered!

For the case when the target loop transfer function can be approximately recovered, the following full-order Chen–Saber–Sannuti (CSS) architecture-based control law (see [111, 117]),

$$\begin{cases} \delta(v) = Av + K_f(y - Cv) \\ u = -Fv \end{cases} \quad (3.274)$$

which has a resulting recovery error,

$$E(\zeta) = M_f(\zeta) = F(\zeta I - A + K_f C)^{-1}(B - K_f D) \quad (3.275)$$

can be utilized to recover the target loop transfer function as well. In fact, when the same gain matrix K_f is used, the full-order CSS architecture-based controller would yield a much better recovery compared to that of the full order observer-based controller.

ii. Reduced-order Observer-based Controller. For simplicity, we assume that C and D have already been transformed into the form

$$C = \begin{bmatrix} 0 & C_{02} \\ I & 0 \end{bmatrix} \quad \text{and} \quad D = \begin{bmatrix} D_0 \\ 0 \end{bmatrix} \quad (3.276)$$

where D_0 is of full row rank. Then, the dynamic equations of Σ can be partitioned as follows:

$$\begin{cases} \begin{pmatrix} \delta(x_1) \\ \delta(x_2) \end{pmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} u \\ \begin{pmatrix} y_0 \\ y_1 \end{pmatrix} = \begin{bmatrix} 0 & C_{02} \\ I & 0 \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} D_0 \\ 0 \end{bmatrix} u \end{cases} \quad (3.277)$$

where x_1 is readily accessible. Let

$$A_r = A_{22}, \quad B_r = B_2, \quad C_r = \begin{bmatrix} C_{02} \\ A_{12} \end{bmatrix}, \quad D_r = \begin{bmatrix} D_0 \\ B_1 \end{bmatrix} \quad (3.278)$$

and the reduced-order observer gain matrix K_r be such that $A_r - K_r C_r$ is asymptotically stable. Next, we partition

$$F = [F_1 \quad F_2], \quad K_r = [K_{r0} \quad K_{r1}] \quad (3.279)$$

in conformity with the partitions of $x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$ and $y = \begin{pmatrix} y_0 \\ y_1 \end{pmatrix}$, respectively. Then, define

$$G_r = [K_{r0}, \quad A_{21} - K_{r1}A_{11} + (A_r - K_r C_r)K_{r1}] \quad (3.280)$$

The reduced-order observer-based controller is given by

$$\begin{cases} \delta(v) = (A_r - K_r C_r)v + (B_r - K_r D_r)u + G_r y \\ u = -F_2 v - [0, \quad F_1 + F_2 K_{r1}] y \end{cases} \quad (3.281)$$

It is again reported in [110, 117] that the recovery error resulting from the reduced-order observer-based controller can be expressed as

$$E(\varsigma) = M_r(\varsigma)[I + M_r(\varsigma)]^{-1}[I + F(\varsigma I - A)^{-1}B] \quad (3.282)$$

where

$$M_r(\varsigma) = F_2(\varsigma I - A_r + K_r C_r)^{-1}(B_r - K_r D_r) \quad (3.283)$$

Thus, making $E(\varsigma)$ zero or small is equivalent to designing a reduced-order observer gain K_r such that $M_r(\varsigma)$, or equivalently $M_r'(\varsigma)$, is zero or small. Following the same idea as in the full-order case, we define an auxiliary system

$$\begin{cases} \dot{\tilde{x}} = A_r' \tilde{x} + C_r' \tilde{u} + F_2' \tilde{w} \\ \tilde{y} = \tilde{x} \\ \tilde{h} = B_r' \tilde{x} + D_r' \tilde{u} \end{cases} \quad (3.284)$$

with a state feedback control law,

$$\tilde{u} = -K_r' \tilde{x} \quad (3.285)$$

Obviously, the closed-loop transfer matrix from \tilde{w} to \tilde{h} of the above system is equivalent to $M_r'(\varsigma)$. Hence, the methods of Sections 3.4 and 3.5 for H_2 and H_∞ optimal control again can be used to find K_r to minimize either the H_2 -norm or H_∞ -norm of $M_r'(\varsigma)$. In particular, for the case when Σ satisfies Condition 1 (for continuous-time systems) or Condition 2 (for discrete-time systems) stated in the full-order case, the target loop can be either exactly or approximately recovered. In fact, in this case, the following reduced-order CSS architecture-based controller

$$\begin{cases} \delta(v) = (A_r - K_r C_r) v + G_r y \\ u = -F_2 v - [0, F_1 + F_2 K_{r1}] y \end{cases} \quad (3.286)$$

which has a resulting recovery error,

$$E(\varsigma) = M_r(\varsigma) = F_2(\varsigma I - A_r + K_r C_r)^{-1}(B_r - K_r D_r) \quad (3.287)$$

can also be used to recover the given target loop transfer function. Again, when the same K_r is used, the reduced-order CSS architecture-based controller would yield a better recovery compared to that of the reduced-order observer-based controller (see [111, 117]).

3.7.2 LTR at Output Point

For the case when uncertainties of the given plant are modeled at the output point, the following dualization procedure can be used to find appropriate solutions. The basic idea is to convert the LTR design at the output point of the given plant into an equivalent LTR problem at the input point of an auxiliary system so that all the methods studied in the previous subsection can be readily applied.

1. Consider a plant Σ characterized by the quadruple (A, B, C, D) . Let us design a Kalman filter or an observer first with a Kalman filter or observer gain matrix K_f such that $A - K_f C$ is asymptotically stable and the resulting target loop

$$L_t(\zeta) = C(\zeta I - A)^{-1} K_f \quad (3.288)$$

meets all the design requirements specified at the output point. We are now seeking to design a measurement feedback controller $K(\zeta)$ such that all the properties of $L_t(\zeta)$ can be recovered.

2. Define a dual system Σ_{du} characterized by $(A_{du}, B_{du}, C_{du}, D_{du})$ where

$$A_{du} := A', \quad B_{du} := C', \quad C_{du} := B', \quad D_{du} := D' \quad (3.289)$$

Let $F_{du} := K_f'$ and let $L_{du}(\zeta)$ be defined as

$$L_{du}(\zeta) := L_t'(\zeta) = F_{du}(\zeta I - A_{du})^{-1} B_{du} \quad (3.290)$$

Let $L_{du}(\zeta)$ be considered as a target loop transfer function for Σ_{du} when the loop is broken at the input point of Σ_{du} . Let a measurement feedback controller $K_{du}(\zeta)$ be used for Σ_{du} . Here, the controller $K_{du}(\zeta)$ could be based either on a full- or a reduced-order observer or CSS architecture depending upon what $K(\zeta)$ is based on. Following the results given earlier for LTR at the input point to design an appropriate controller $K_{du}(\zeta)$, then the required controller for LTR at the output point of the original plant Σ is given by

$$K(\zeta) = K_{du}'(\zeta) \quad (3.291)$$

This concludes the LTR design for the case when the loop is broken at the output point of the plant.

Finally, we note that there are another type of loop transfer recovery techniques that have been proposed in the literature, *i.e.* in Chen *et al.* [120–122], in which the focus is to recover a closed-loop transfer function instead of an open-loop one as in the conventional LTR design studied in this section. Interested readers are referred to [120–122] for details.

Classical Nonlinear Control

4.1 Introduction

Every physical system in real life has nonlinearities and very little can be done to overcome them. Many practical systems are sufficiently nonlinear so that important features of their performance may be completely overlooked if they are analyzed and designed through linear techniques. In HDD servo systems, major nonlinearities are frictions, high-frequency mechanical resonances and actuator saturation nonlinearities. Among all these, the actuator saturation could be the most significant nonlinearity in designing an HDD servo system. When the actuator saturates, the performance of the control system designed will seriously deteriorate. Interested readers are referred to a recent monograph by Hu and Lin [123] for a fairly complete coverage of many newly developed results on control systems with actuator nonlinearities.

The actuator saturation in the HDD has seriously limited the performance of its overall servo system, especially in the track-seeking stage, in which the HDD R/W head is required to move over a wide range of tracks. It will be obvious in the forthcoming chapters that it is impossible to design a pure linear controller that would achieve a desired performance in the track-seeking stage. Instead, we have no choice but to utilize some sophisticated nonlinear control techniques in the design. The most popular nonlinear control technique used in the design of HDD servo systems is the so-called proximate time-optimal servomechanism (PTOS) proposed by Workman [30], which achieves near time-optimal performance for a large class of motion control systems characterized by a double integrator. The PTOS was actually modified from the well-known time-optimal control. However, it is made to yield a minimum variance with smooth switching from the track-seeking to track-following modes. We also introduce another nonlinear control technique, namely a mode-switching control (MSC). The MSC we present in this chapter is actually a combination of the PTOS and the robust and perfect tracking (RPT) control of Chapter 3. In particular, in the MSC scheme for HDD servo systems, the track-seeking mode is controlled by a PTOS and the track-following mode is controlled by a RPT controller. The MSC is a type of variable-structure control systems, but its switching is in only one direction.

4.2 Time-optimal Control

We recall the technique of the time-optimal control (TOC) in this section. Given a dynamic system characterized by

$$\dot{x} = h(x, u, t) \quad (4.1)$$

where x is the state variable and u is the control input, the objective of optimal control is to determine a control input u that causes a controlled process to satisfy the physical constraints and at the same time optimize a certain performance criterion,

$$J = \int_{t_0}^{t_f} f(x, u, t) dt \quad (4.2)$$

where t_0 and t_f are, respectively, initial time and final time of operation, and f is a scalar function. The TOC is a special class of optimization problems and is defined as the transfer of the system from an arbitrary initial state $x(t_0)$ to a specified target set point in minimum time. For simplicity, we let $t_0 = 0$. Hence, the performance criterion for the time-optimal problem becomes one of minimizing the following cost function with $f(x, u, t) = 1$, *i.e.*

$$J = \int_0^{t_f} dt = t_f \quad (4.3)$$

Let us now derive the TOC law using Pontryagin's principle and the calculus of variation (see, *e.g.*, [124]) for a simple dynamic system obeying Newton's law, *i.e.* for a double-integrator system represented by

$$\ddot{y}(t) = au(t) \quad (4.4)$$

where y is the position output, a is the acceleration constant and u is the input to the system. It will be seen later that the dynamics of the actuator of an HDD can be approximated as a double-integrator model. To start with, we rewrite Equation 4.4 as the following state-space model:

$$\dot{x}(t) = Ax(t) + Bu(t) \quad (4.5)$$

with

$$A = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ a \end{bmatrix}, \quad x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} y \\ v \end{pmatrix} = \begin{pmatrix} y \\ \dot{y} \end{pmatrix} \quad (4.6)$$

Note that $v = \dot{y}$ is the velocity of the system. Let the control input be constrained as follows:

$$|u(t)| \leq u_{\max} \quad (4.7)$$

Then, the Hamiltonian (see, *e.g.*, [124]) for such a problem is given by

$$H(x(t), u(t), \lambda(t)) = 1 + \lambda_1(t)x_2(t) + \lambda_2(t)au(t) \quad (4.8)$$

where $\lambda = (\lambda_1 \ \lambda_2)'$ is a vector of the time-varying Lagrange multipliers. Pontryagin's principle states that the Hamiltonian is minimized by the optimal control, or

$$H(x^*, u^*, \lambda^*) \leq H(x^*, u, \lambda^*) \quad (4.9)$$

where superscript * indicates optimality. Thus, from Equations 4.8 and 4.9, the optimal control is

$$u^*(t) = \left\{ \begin{array}{ll} -u_{\max}, & \text{for } \lambda_2^*(t) > 0 \\ +u_{\max}, & \text{for } \lambda_2^*(t) < 0 \end{array} \right\} := -\text{sgn}(\lambda_2^*(t))u_{\max} \quad (4.10)$$

The calculus of variation (see [124]) yields the following necessary condition for a time-optimal solution:

$$\dot{\lambda}^*(t) = -A'\lambda^*(t) \quad (4.11)$$

which is known as a costate equation in optimal control terminology. The solution to the costate equation is of the form

$$\left. \begin{array}{l} \lambda_1^*(t) = c_1 \\ \lambda_2^*(t) = -c_1 t + c_2 \end{array} \right\} \quad (4.12)$$

where c_1 and c_2 are constants of integration. Equation 4.12 indicates that λ_2^* and, therefore u^* can change sign at most once. Since there can be at most one switching, the optimal control for a specified initial state must be one of the following forms:

$$\left. \begin{array}{l} 1^\circ : \quad u^*(t) = +u_{\max}, \quad \forall t \in [t_0, t^*] \\ 2^\circ : \quad u^*(t) = \left\{ \begin{array}{l} +u_{\max}, \quad \forall t \in [t_0, t_1) \\ -u_{\max}, \quad \forall t \in [t_1, t^*] \end{array} \right. \\ 3^\circ : \quad u^*(t) = -u_{\max}, \quad \forall t \in [t_0, t^*] \\ 4^\circ : \quad u^*(t) = \left\{ \begin{array}{l} -u_{\max}, \quad \forall t \in [t_0, t_1) \\ +u_{\max}, \quad \forall t \in [t_1, t^*] \end{array} \right. \end{array} \right\} \quad (4.13)$$

Thus, the segment of optimal trajectories can be found by integrating Equation 4.5 with $u = \pm u_{\max}$ to obtain

$$x_2(t) = \pm a u_{\max} t + c_3 \quad (4.14)$$

$$x_1(t) = \pm \frac{1}{2} a u_{\max} t^2 + c_3 t + c_4 \quad (4.15)$$

where c_3 and c_4 are constants of integration. It is to be noted that if the initial state lies on the optimal trajectories defined by Equations 4.14 and 4.15 in the state plane, then the control will be either 1° or 3° in Equation 4.13 depending upon the direction of motion. In HDD servo systems, it will be shown later that the problem is of relative head-positioning control, and hence the initial and final states must be

$$x(0) = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad x(t^*) = \begin{pmatrix} r \\ 0 \end{pmatrix} \quad (4.16)$$

where r is the reference set point. Because of these kinds of initial state in HDD servo systems, the optimal control must be chosen from either 2° or 4° in Equation 4.13. Note that if the control input $+u_{\max}$ produces the acceleration α , then the input $-u_{\max}$ will produce a deceleration of the same magnitude.

Hence, the minimum time performance can be achieved either with maximum acceleration for half of the travel followed by maximum deceleration for an equal amount of time, or by first accelerating and then decelerating the system with maximum effort to follow some predefined optimal velocity trajectory to reach the final destination in minimum time. The former case results in an open-loop form of TOC that uses predetermined time-based acceleration and deceleration inputs, whereas the latter yields a closed-loop form of TOC. We note that if the area under acceleration, which is a function of time, is the same as the area under deceleration, there will be no net change in velocity after the input is removed. The final output velocity and the position will be in a steady state.

In general, the time-optimal performance can be achieved by switching the control between two extreme levels of the input, and we have shown that in the double-integrator system the number of switchings is at most equal to one, *i.e.* one less than the order of dynamics. Thus, if we extend the result to an n th-order system, it will need $n - 1$ switchings between maximum and minimum inputs to achieve a time-optimal performance. Since the control must be switched between two extreme values, the TOC is also known as *bang-bang* control.

In what follows, we discuss the bang-bang control in two versions, *i.e.* in the open-loop and in the closed-loop forms for the double-integrator model characterized by Equation 4.5 with the control constraint represented by Equation 4.7.

4.2.1 Open-loop Bang-bang Control

The open-loop method of bang-bang control uses maximum acceleration and maximum deceleration for a predetermined time period. Thus, the time required for the system to reach the target position in minimum time is predetermined from the above principles and the control input is switched between two extreme levels for this time period. We can precalculate the minimum time t^* for a specified reference set point r . Let the control be

$$u^*(t) = \begin{cases} +u_{\max}, & \text{for } t \in (0, \frac{t^*}{2}] \\ -u_{\max}, & \text{for } t \in (\frac{t^*}{2}, t^*] \end{cases} \quad (4.17)$$

We now solve Equations 4.14 and 4.15 for the accelerating phase with zero initial condition. For the accelerating phase, *i.e.* with $u = +u_{\max}$, we have

$$x_2(t) = a u_{\max} t, \quad x_1(t) = \frac{1}{2} a u_{\max} t^2 \quad (4.18)$$

At the end of the accelerating phase, *i.e.* at $t = t^*/2$,

$$x_{2a} = a u_{\max} \frac{t^*}{2}, \quad x_{1a} = \frac{1}{8} a u_{\max} (t^*)^2 \quad (4.19)$$

Similarly, at the end of decelerating phase, we can show that

$$x_{2d} = 0, \quad x_{1d} = \frac{1}{8} a u_{\max} (t^*)^2 \quad (4.20)$$

Obviously, the total displacement at the end of bang-bang control must reach the target, *i.e.* the reference set point r . Thus,

$$r = x_{1a} + x_{1d} = \frac{1}{4} a u_{\max} (t^*)^2 \quad (4.21)$$

which gives

$$t^* = \sqrt{\frac{4r}{a u_{\max}}} \quad (4.22)$$

the minimum time required to reach the target set point.

4.2.2 Closed-loop Bang-bang Control

In this method, the velocity of the plant is controlled to follow a predefined trajectory and more specifically the decelerating trajectory. These trajectories can be generated from the phase-plane analysis. This analysis is explained below for the system given by Equation 4.5 and can be extended to higher-order systems (see, *e.g.*, [124]). We will show later that this deceleration trajectory brings the system to the desired set point in finite time. We now move to find the deceleration trajectory.

First, eliminating t from Equations 4.14 and 4.15, we have

$$x_1(t) = \frac{1}{2 a u_{\max}} x_2^2(t) + c_5, \quad \text{for } u = +u_{\max} \quad (4.23)$$

$$x_1(t) = -\frac{1}{2 a u_{\max}} x_2^2(t) + c_6, \quad \text{for } u = -u_{\max} \quad (4.24)$$

where c_5 and c_6 are appropriate constants. Note that each of the above equations defines the family of parabolas. Let us define $e(t) := r - x_1(t)$ to be the positioning error with r being the desired final position. Then, if we consider the trajectories between $e(t)$ and $x_2(t)$, our desired final state in $e(t)$ and $x_2(t)$ plane must be

$$\begin{pmatrix} r - x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \quad (4.25)$$

In this case, the constants in the above trajectories are equal to zero. Moreover, both of the trajectories given by Equations 4.23 and 4.24 are the decelerating trajectories depending upon the direction of the travel. The mechanism of the TOC can be illustrated in a graphical form as given in Figure 4.1. Clearly, any initial state lying below the curve is to be driven by the positive accelerating force to bring the state to the

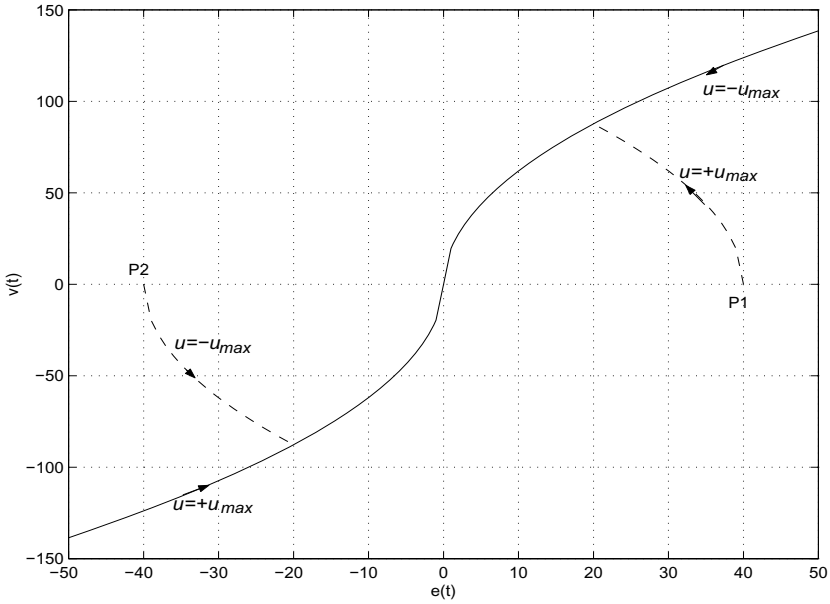


Figure 4.1. Deceleration trajectories for TOC

deceleration trajectory. On the other hand, any initial state lying above the curve is to be accelerated by the negative force to the deceleration trajectory.

Let

$$f_t(e) := \text{sgn}(e) \sqrt{2 a u_{\max} |e|} \tag{4.26}$$

The control law is then given by

$$u = u_{\max} \cdot \text{sgn}(f_t(e) - v) \tag{4.27}$$

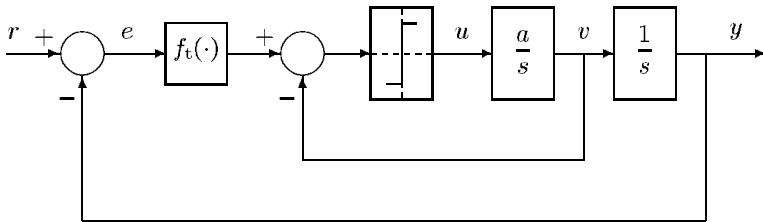


Figure 4.2. Typical scheme of TOC

A block diagram depicting the closed-loop method of bang-bang control is shown in Figure 4.2. Unfortunately, the control law given by Equation 4.27 for the system

shown in Figure 4.2, although time-optimal, is not practical. It applies maximum or minimum input to the plant to be controlled even for a small error. Moreover, this algorithm is not suited for disk drive applications for the following reasons:

1. even the smallest system process or measurement noise will cause control “chatter”. This will excite the high-frequency modes.
2. any error in the plant model, will cause limit cycles to occur.

As such, the TOC given above has to be modified to suit HDD applications. In the following section, we recall a modified version of the TOC proposed by Workman [30], *i.e.* the PTOS. Such a control scheme is widely used nowadays in designing HDD servo systems.

4.3 Proximate Time-optimal Servomechanism

The infinite gain of the signum function in the TOC causes control chatter, as seen in the previous section. Workman [30], in 1987, proposed a modification of this technique, *i.e.* the so-called PTOS, to overcome such a drawback. The PTOS essentially uses maximum acceleration where it is practical to do so. When the error is small, it switches to a linear control law. To do so, it replaces the signum function in TOC law by a saturation function. In the following sections, we revisit the PTOS method in continuous-time and in discrete-time domains.

4.3.1 Continuous-time Systems

The configuration of the PTOS is shown in Figure 4.3. The function $f_p(\cdot)$ is a finite-slope approximation to the switching function $f_t(\cdot)$ given by Equation 4.26. The PTOS control law for the system in Equation 4.5 is given by

$$u = u_{\max} \cdot \text{sat} \left(\frac{k_2[f_p(e) - v]}{u_{\max}} \right) \tag{4.28}$$

where $\text{sat}(x)$ is defined as

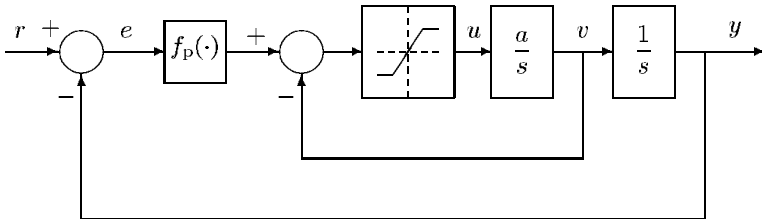


Figure 4.3. Continuous-time PTOS

$$\text{sat}(x) = \begin{cases} +1, & \text{if } x > 1 \\ x, & \text{if } -1 \leq x \leq 1 \\ -1, & \text{if } x < -1 \end{cases} \quad (4.29)$$

and the function $f_p(e)$ is given by

$$f_p(e) = \begin{cases} \frac{k_1}{k_2}e & \text{for } |e| \leq y_\ell \\ \text{sgn}(e) \left[\sqrt{2a u_{\max} \alpha |e|} - \frac{u_{\max}}{k_2} \right] & \text{for } |e| > y_\ell \end{cases} \quad (4.30)$$

Here we note that k_1 and k_2 are, respectively, the feedback gains for position and velocity, α is a constant between 0 and 1 and is referred to as the acceleration discount factor, and y_ℓ is the size of the linear region. Since the linear portion of the curve $f_p(\cdot)$ must connect the two disjoint halves of the nonlinear portion, we have constraints on the feedback gains and the linear region to guarantee the continuity of the function $f_p(\cdot)$. It was proved by Workman [30] that

$$y_\ell = \frac{u_{\max}}{k_1}, \quad k_2 = \sqrt{\frac{2k_1}{\alpha}} \quad (4.31)$$

The control zones in the PTOS are shown in Figure 4.4. The two curves bounding the switching curve (central curve) now redefine the control boundaries and it is termed a linear boundary. Let this region be \mathbf{U} . The region below the lower curve is

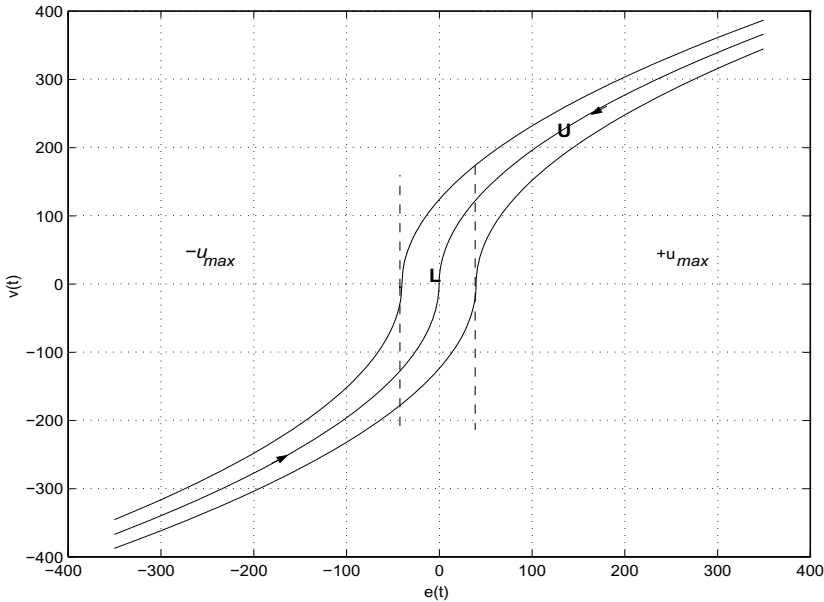


Figure 4.4. Control zones of a PTOS

the region where the control $u = +u_{\max}$, whereas the region above the upper curve is the region where the control $u = -u_{\max}$. It has been proved [30] that once the state trajectory enters the band \mathbf{U} in Figure 4.4 it remains within \mathbf{U} and the control signal is below the saturation. The region marked \mathbf{L} is the region where the linear control is applied.

The presence of the acceleration discount factor α allows us to accommodate uncertainties in the plant accelerating factor at the cost of increase in response time. By approximating the positioning time as the time that it takes the positioning error to be within the linear region, one can show that the percentage increase P in time taken by the PTOS over the time taken by the TOC is given by (see [30]):

$$P = \frac{1}{2} \left(\frac{1}{\sqrt{\alpha}} - 1 \right) \times 100\% \quad (4.32)$$

Clearly, larger values of α make the response closer to that of the TOC. As a result of changing the nonlinearity from $\text{sgn}(\cdot)$ to $\text{sat}(\cdot)$, the control chatter is eliminated.

4.3.2 Discrete-time Systems

The discrete-time PTOS can be derived from its continuous-time counterpart, but with some conditions on sample time to ensure stability. In his seminal work, Workman [30] extended the continuous-time PTOS to discrete-time control of a continuous-time double-integrator plant driven by a zero-order hold as shown in Figure 4.5. As in the continuous-time case, the states are defined as position and velocity. With insignificant calculation delay, the state-space description of the plant given by Equation 4.5 in the discrete-time domain is

$$x(k+1) = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} \alpha T_s^2 / 2 \\ \alpha T_s \end{bmatrix} u(k) \quad (4.33)$$

where T_s is the sampling period. The control structure is a discrete-time mapping of the continuous-time PTOS law, but with a constraint on the sampling period to

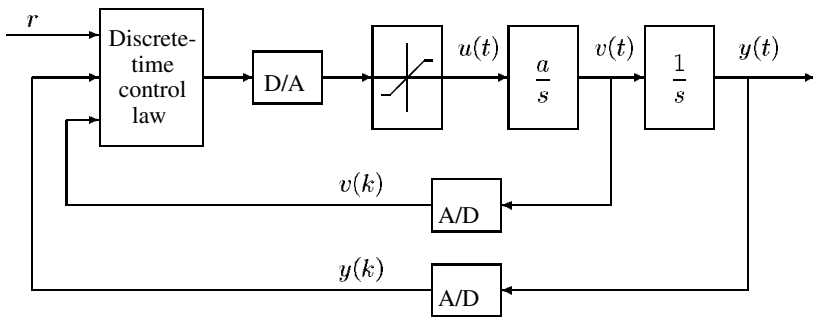


Figure 4.5. Discrete-time PTOS

guarantee that the control does not saturate during the deceleration phase to the target position and also to guarantee its stability. Thus, the mapped control law is

$$u(k) = u_{\max} \cdot \text{sat} \left(\frac{k_2 [f_p(e(k)) - v(k)]}{u_{\max}} \right) \quad (4.34)$$

with the following constraint on sampling frequency ω_s ,

$$\frac{\omega_s}{\omega_n} > 6.3 \quad (4.35)$$

where ω_n is the desired bandwidth of the closed-loop system.

4.4 Mode-switching Control

In this section, we present a mode-switching control (MSC) design technique for both continuous-time and discrete-time systems, which is a combination of the PTOS of the previous section and the RPT technique given in Chapter 4.

4.4.1 Continuous-time Systems

In this subsection, we follow the development of [125] to introduce the design of an MSC design for a system characterized by a double integrator or in the following state-space equation:

$$\dot{x} = Ax + Bu = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ a \end{bmatrix} u, \quad x = \begin{pmatrix} y \\ v \end{pmatrix} \quad (4.36)$$

where as usual x is the state, which consists of the displacement y and the velocity v ; u is the control input constrained by

$$|u(t)| \leq u_{\max} \quad (4.37)$$

As will be seen shortly in the forthcoming chapters, the VCM actuators of HDDs can generally be approximated by such a model with appropriate parameters a and u_{\max} . In HDD servo systems, in order to achieve both high-speed track seeking and highly accurate head positioning, multimode control designs are widely used. The two commonly used multimode control designs are MSC and sliding mode control. Both control techniques in fact belong to the category of variable-structure control. That is, the control is switched between two or more different controllers to achieve the two conflicting requirements. In this section, we propose an MSC scheme in which the seeking mode is controlled by a PTOS and the track-following mode is controlled by a RPT controller.

As noted earlier, the MSC (see, *e.g.*, [15]) is a type of variable structure control systems [126], but the switching is in only one direction. Figure 4.6 shows a basic schematic diagram of MSC. There are track seeking and track following modes.

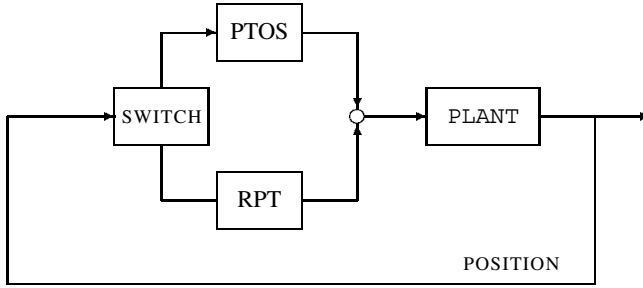


Figure 4.6. Basic schematic diagram of MSC

Each servo mode can be designed independently, and so the main issue in MSC is the design of the switching mechanism.

This design problem has not yet been completely resolved, and many heuristic approaches have been tried so far (see, *e.g.*, [127]). Several methods were proposed for mode switching from one controller to another. In [15], a method called initial value compensation is proposed. Note that, when the switch is transferred from the track-seeking mode to the track-following mode, the final states of the track-seeking controller become the initial states for the track-following controller, and hence affect the settling performance of track-following mode. In order to reduce the impact of these initial values during mode switching, some compensation must be worked out. Here, the initial values are referred to the values of the states during mode switching. However, the RPT controllers developed by Chen and coworkers [74, 106] (see also Chapter 3) have enough robustness against plant variations and are actually independent of initial values. Hence, the use of these controllers in the track-following mode eliminates the need for initial-value compensation during mode switching. Moreover, the RPT controllers in a track-following servo have been proved to be robust against resonance mode changes from disk to disk and work well against runout disturbances.

The MSC law that combines the PTOS and RPT controllers takes the following simple form:

$$u = \begin{cases} u_p, & t < t_1 \\ u_r, & t \geq t_1 \end{cases} \quad (4.38)$$

where u_p is a control signal generated by the PTOS control and is given as in Equation 4.28, and u_r is a signal generated by the reduced-order RPT control as given in Equation 3.237. Furthermore, t_1 , the time that MSC switches from one mode to the other, will be presented in the next subsection together with the stability analysis of the closed loop comprising the given plant and the MSC control law.

In what follows, we show the stability of the MSC and give a set of conditions for mode switching. First, we rewrite the given system in Equation 4.36 as follows:

$$\begin{pmatrix} \dot{e} \\ \dot{v} \end{pmatrix} = \begin{bmatrix} 0 & -1 \\ 0 & 0 \end{bmatrix} \begin{pmatrix} e \\ v \end{pmatrix} + \begin{bmatrix} 0 \\ a \end{bmatrix} u := A_p x_p + B_p u, \quad a > 0 \quad (4.39)$$

where $e = r - y$ is the tracking error with r being the target reference. In the HDD servo systems that we deal with in the forthcoming chapters, y is regarded as the displacement of an HDD R/W head, and v is its velocity. Recall the PTOS control law:

$$u_p = u_{\max} \cdot \text{sat} \left[\frac{k_2}{u_{\max}} (f(e) - v) \right] \quad (4.40)$$

where the u_{\max} used throughout this section has a saturation level equal to unity; the function $f(\cdot)$ and the feedback gain k_2 are as defined in the previous section. It has been shown [30] that the PTOS control law yields an asymptotically stable closed-loop system provided that the following conditions are satisfied:

1. $ak_2 > 0$;
2. $f(0) = 0$;
3. $f(e)e > 0$, for any nonzero e ;
4. $\lim_{e \rightarrow \infty} \int_0^e f(\delta) d\delta = \infty$;
5. $\dot{f}(e)$ exists everywhere; and
6. for any e , $u_{\max} \left[-a + \frac{1}{k_2} \right] \dot{f}(e) < -\dot{f}(e)f(e) < u_{\max} \left[a - \frac{1}{k_2} \dot{f}(e) \right]$.

Generally, as the velocity is not measurable, the PTOS control law has to be modified as follows if it is to be implemented in a real system:

$$\begin{cases} u_p = u_{\max} \text{sat} \left[\frac{k_2}{u_{\max}} (f(e) - \hat{v}) \right] \\ \dot{z} = -\kappa z + \kappa^2 e + a u_p \\ \hat{v} = z - \kappa e \end{cases} \quad (4.41)$$

where κ is the estimator feedback gain, and z is the estimator state. Next, we let $\tilde{z} = z - v - \kappa e$. Then, the dynamics of the closed-loop system with the above control law can be written as

$$\begin{cases} \dot{e} = -v \\ \dot{v} = a u_{\max} \text{sat} \left[\frac{k_2}{u_{\max}} (f(e) - \tilde{z} - v) \right] \\ \dot{\tilde{z}} = -\kappa \tilde{z} \end{cases} \quad (4.42)$$

It can be shown that the closed-loop system comprising the given plant and the modified PTOS control law, in which the velocity is replaced by the above estimation, is asymptotically stable, if conditions 1 to 5 above are satisfied and condition 6 is replaced by

$$a > \frac{1}{k_2} \dot{f}(e) > 0 \quad (4.43)$$

and

$$|\dot{f}(e)f(e)| < u_{\max} \left[a - \frac{1}{k_2} \dot{f}(e) \right] - \left[\kappa + \dot{f}(e) \right] \cdot |\tilde{z}(0)| \quad (4.44)$$

We can show that, under these new conditions, the closed-loop system is stable for the case when the control input is saturated, *i.e.* $|k_2[f(e) - (z - \kappa e)]| > u_{\max}$. For the case when $|k_2[f(e) - (z - \kappa e)]| \leq u_{\max}$, the closed-loop system in Equation 4.42 can be written as

$$\begin{cases} \dot{e} = -v \\ \dot{v} = ak_2[f(e) - \tilde{z} - v] \\ \dot{\tilde{z}} = -\kappa\tilde{z} \end{cases} \quad (4.45)$$

Following the result of [128], we propose the following Lyapunov function for the system in Equation 4.45:

$$V_p = \frac{v^2}{2k_2a} + \int_0^e f(\delta) d\delta + \frac{p_z}{2}\tilde{z}^2 \quad (4.46)$$

where $p_z > 0$ is a scalar constant. The derivative of the above Lyapunov function is given by

$$\dot{V}_p = -\left(v + \frac{\tilde{z}}{2}\right)^2 - \tilde{z}^2\left(p_z\kappa - \frac{1}{4}\right) \quad (4.47)$$

The last term is negative for all $p_z > \frac{1}{4\kappa}$. Thus, under this choice of p_z , $\dot{V}_p \leq 0$. It follows from LaSalle's Theorem [129] that the closed-loop system comprising the PTOS control law with the estimated velocity and the given plant is asymptotically stable.

It is obvious that the closed-loop system comprising the given plant in Equation 4.39 and the reduced-order RPT control law of Equation 3.237 is asymptotically stable when the control input is not saturated. For completeness, and for the analysis of the overall closed-loop system with the MSC scheme, we proceed to investigate the closed-loop system comprising the plant and the RPT controller, which can be written as

$$\begin{cases} \dot{z} = -\kappa z + \kappa^2 e + a u_{\max} \cdot \text{sat}\left(\frac{u_r}{u_{\max}}\right) \\ u_r = (f_1 - \kappa f_2)e + f_2 z \end{cases} \quad (4.48)$$

where κ is again the reduced-order observer gain, which is selected to be exactly the same as that used for the velocity estimation in the PTOS, and $F = [f_1 \ f_2]$ is the feedback gain obtained using the RPT technique of Chapter 3. Again, let $\tilde{z} = z - v - \kappa e$ and rewrite the RPT control law as

$$u_r = [f_1 \ f_2] \begin{pmatrix} e \\ v \end{pmatrix} + f_2 \tilde{z} = Fx_p + f_2 \tilde{z} \quad (4.49)$$

Let $W_x \in \mathbb{R}^{2 \times 2}$ be a positive-definite matrix and solve the following Lyapunov equation:

$$(A_p + B_p F)' P_x + P_x (A_p + B_p F) = -W_x \quad (4.50)$$

for $P_x > 0$. Such a P_x always exists as $A_p + B_p F$ is stable. Next, let

$$w_z > \max \left\{ \frac{1}{2}, f_2 B_p' P_x W_x^{-1} P_x B_p f_2 \right\} > 0 \quad (4.51)$$

and

$$p_z = \frac{w_z}{2\kappa} > \frac{1}{4\kappa} > 0 \quad (4.52)$$

Then, we define a set

$$\mathbf{X} := \left\{ \begin{pmatrix} x \\ z \end{pmatrix} : \begin{pmatrix} x \\ z \end{pmatrix}' \begin{bmatrix} P_x & 0 \\ 0 & p_z \end{bmatrix} \begin{pmatrix} x \\ z \end{pmatrix} \leq c \right\} \quad (4.53)$$

where $c > 0$ is the largest positive value such that

$$\begin{pmatrix} x \\ z \end{pmatrix} \in \mathbf{X} \Rightarrow \left| \begin{bmatrix} F & f_2 \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} \right| \leq u_{\max} \quad (4.54)$$

For all

$$\begin{pmatrix} x_p \\ z \end{pmatrix} \in \mathbf{X} \quad (4.55)$$

the resulting closed-loop system can then be written as

$$\begin{pmatrix} \dot{x}_p \\ \dot{\tilde{z}} \end{pmatrix} = \begin{bmatrix} A_p + B_p F & B_p f_2 \\ 0 & -\kappa \end{bmatrix} \begin{pmatrix} x_p \\ \tilde{z} \end{pmatrix} \quad (4.56)$$

Define a Lyapunov function,

$$V_r = \begin{pmatrix} x_p \\ \tilde{z} \end{pmatrix}' \begin{bmatrix} P_x & 0 \\ 0 & p_z \end{bmatrix} \begin{pmatrix} x_p \\ \tilde{z} \end{pmatrix} \quad (4.57)$$

and evaluate its derivative along the trajectories of the closed-loop system in Equation 4.56, *i.e.*

$$\begin{aligned} \dot{V}_r &= \begin{pmatrix} x_p \\ \tilde{z} \end{pmatrix}' \begin{bmatrix} -W_x & P_x B_p f_2 \\ f_2 B_p' P_x & -w_z \end{bmatrix} \begin{pmatrix} x_p \\ \tilde{z} \end{pmatrix} \\ &= \begin{pmatrix} \hat{x} \\ \tilde{z} \end{pmatrix}' \begin{bmatrix} -W_x & 0 \\ 0 & -\tilde{w}_z \end{bmatrix} \begin{pmatrix} \hat{x} \\ \tilde{z} \end{pmatrix} \leq 0 \end{aligned} \quad (4.58)$$

where

$$\hat{x} = x_p - W_x^{-1} P_x B_p f_2 \tilde{z}, \quad \tilde{w}_z = w_z - f_2 B_p' P_x W_x^{-1} P_x B_p f_2 > 0 \quad (4.59)$$

This shows that all trajectories of Equation 4.56 starting from \mathbf{X} remain there and converge asymptotically to zero. Hence, the closed-loop system comprising the plant and the reduced-order RPT control law is asymptotically stable provided that the control input is not saturated.

Next, we re-express Equation 4.46 using the Taylor expansion as follows:

$$V_p = \frac{v^2}{2k_2 a} + \frac{\dot{f}(\tau)}{2} e^2 + \frac{p_z}{2} \tilde{z}^2 = \begin{pmatrix} e \\ v \\ \tilde{z} \end{pmatrix}' \begin{bmatrix} \frac{\dot{f}(\tau)}{2} & 0 & 0 \\ 0 & \frac{1}{2ak_2} & 0 \\ 0 & 0 & \frac{p_z}{2} \end{bmatrix} \begin{pmatrix} e \\ v \\ \tilde{z} \end{pmatrix}$$

where τ is an appropriate scalar between 0 and e . Let

$$\sigma = \min \left\{ \frac{\dot{f}(\tau)}{2}, \frac{1}{2ak_2}, \frac{p_z}{2} \right\} / \max \{ \lambda_{\max}(P_x), p_z \} \quad (4.60)$$

The MSC scheme can be obtained as follows:

$$u(t) = \begin{cases} u_p, & t < t_1 \\ u_r, & t \geq t_1 \end{cases} \quad (4.61)$$

where t_1 is such that

$$\begin{pmatrix} x_p(t_1) \\ z(t_1) \end{pmatrix} \in \mathbf{X} \quad \text{and} \quad |e(t_1)| < y_\ell \quad (4.62)$$

and where y_ℓ is the size of the linear region of the PTOS control law. The Lyapunov function for the overall closed-loop system can be chosen as

$$V = V_p[1 - 1(t - t_1)] + \sigma V_r 1(t - t_1) \quad (4.63)$$

where $1(t)$ is the unit step function. It is simple to verify that

$$\dot{V} = \dot{V}_p[1 - 1(t - t_1)] + \sigma \dot{V}_r 1(t - t_1) + (\sigma V_r - V_p)\delta(t - t_1) \quad (4.64)$$

It has already been proved that the derivatives of the functions V_p and V_r are negative-definite. The last term is always negative in view of the definition of σ in Equation 4.60. Hence, $\dot{V} \leq 0$ and the resulting closed-loop system comprising the given plant and the mode-switching control law is stable. As such, Equation 4.62 can be regarded as the switching condition for the MSC design.

4.4.2 Discrete-time Systems

We now present an MSC design technique for a discrete-time system characterized by a double-integrator or by the following state-space form:

$$x(k+1) = \begin{bmatrix} 1 & T_s \\ 0 & 1 \end{bmatrix} x(k) + \begin{bmatrix} aT_s^2/2 \\ aT_s \end{bmatrix} u(k), \quad x = \begin{pmatrix} y \\ v \end{pmatrix} \quad (4.65)$$

where T_s is the sampling period, a is assumed to be a positive scalar for simplicity, y is the displacement and v is the velocity in the context of mechanical servo systems. The PTOS controller for such a discrete-time system is given in Equation 4.34, *i.e.*

$$u_p(k) = u_{\max} \cdot \text{sat} \left(\frac{k_2}{u_{\max}} \left[f(r - y(k)) - v(k) \right] \right) \quad (4.66)$$

where r is the target reference. To simplify our analysis, we introduce a new variable $\bar{v} = T_s v$ and rewrite the system of Equation 4.65 as follows

$$\begin{pmatrix} y(k+1) \\ \bar{v}(k+1) \end{pmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} y(k) \\ \bar{v}(k) \end{pmatrix} + \begin{bmatrix} \bar{a}/2 \\ \bar{a} \end{bmatrix} u(k) \quad (4.67)$$

and the corresponding PTOS controller as

$$u_p(k) = u_{\max} \cdot \text{sat} \left(\frac{\bar{k}_2}{u_{\max}} \left[\bar{f}(r - y(k)) - \bar{v}(k) \right] \right) \quad (4.68)$$

where

$$\bar{a} = T_s^2 a, \quad \bar{k}_2 = \frac{k_2}{T_s}, \quad \bar{f}(\cdot) = T_s f(\cdot) \quad (4.69)$$

The MSC law comprising the PTOS and RPT controllers is given as follows,

$$u(k) = \begin{cases} u_p(k), & k < k_s \\ u_r(k), & k \geq k_s \end{cases} \quad (4.70)$$

where $u_p(k)$ is a control signal generated by the PTOS controller and $u_r(k)$ is a control signal generated by the RPT controller (see Chapter 3 for details). k_s is the index that the MSC switches from one mode to the other.

In what follows, we proceed to present the detailed stability analysis and mode switching condition for the above MSC design. Noting that the tracking error is given by $e(k) = r - y(k)$, we can rewrite the system in Equation 4.67 as the following

$$\begin{pmatrix} e(k+1) \\ \bar{v}(k+1) \end{pmatrix} = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{pmatrix} e(k) \\ \bar{v}(k) \end{pmatrix} + \begin{bmatrix} -\bar{a}/2 \\ \bar{a} \end{bmatrix} u(k) \quad (4.71)$$

and the corresponding PTOS controller as follows,

$$u_p(k) = u_{\max} \cdot \text{sat} \left(\frac{\bar{k}_2}{u_{\max}} \left[\bar{f}(e(k)) - \bar{v}(k) \right] \right) \quad (4.72)$$

It has been shown in [30] that the discrete PTOS control law yields an asymptotically stable closed-loop system provided that the following conditions are satisfied:

1. $\bar{a}\bar{k}_2 \in (0, 2)$;
2. $f(0) = 0$;
3. $f(e)e > 0, \forall e \neq 0$;
4. $\lim_{e \rightarrow \infty} \int_0^e f(\delta) d\delta = \infty$;
5. $\dot{f}(e)$ exists everywhere;
6. For all e , $|\dot{f}(e)| < \frac{1}{2}T_s^{-1}$; and
7. For the unsaturated situation, $|f(e + \Delta e) - (v + \Delta v)| < \frac{u_{\max}}{k_2}$, where

$$\Delta e = -\bar{v} - \frac{1}{2}\bar{a}u_{\max} \cdot \text{sat} \left(\frac{k_2}{u_{\max}} [f(e) - v] \right) \quad (4.73)$$

and

$$\Delta \bar{v} = \bar{a}u_{\max} \cdot \text{sat} \left(\frac{k_2}{u_{\max}} [f(e) - v] \right), \quad \Delta v = \frac{\Delta \bar{v}}{T_s} \quad (4.74)$$

As the velocity \bar{v} is generally not measurable, the PTOS control law has to be modified as follows,

$$\begin{cases} u_p(k) = u_{\max} \cdot \text{sat} \left(\frac{\bar{k}_2}{u_{\max}} [\bar{f}(e(k)) - \hat{v}(k)] \right) \\ x_v(k+1) = (1 - \kappa)x_v(k) + \kappa^2 e(k) + (1 - \kappa/2)\bar{a}u_p(k) \\ \hat{v}(k) = x_v(k) - \kappa e(k) \end{cases} \quad (4.75)$$

where $\kappa \in (0, 1)$ is the feedback gain of the estimator and $x_v(k)$ is the state variable of the estimator. Next, let

$$\tilde{x}_v(k) = x_v(k) - \bar{v}(k) - \kappa e(k) \quad (4.76)$$

Then, the closed-loop system comprising Equations 4.71 and 4.75 can be rewritten as

$$\begin{cases} e(k+1) = e(k) - \bar{v}(k) - \frac{1}{2}\bar{a}u_p(k) \\ \bar{v}(k+1) = \bar{v}(k) + \bar{a}u_p(k) \\ \tilde{x}_v(k+1) = (1 - \kappa)\tilde{x}_v(k) \\ u_p(k) = u_{\max} \text{sat} \left(\frac{\bar{k}_2}{u_{\max}} [\bar{f}(e(k)) - \bar{v}(k) - \tilde{x}_v(k)] \right) \end{cases} \quad (4.77)$$

For such a case, Condition 7 above has to be replaced by

$$|\bar{f}(e + \Delta e) - (x_v + \Delta x_v) + \kappa(e + \Delta e)| < \frac{u_{\max}}{\bar{k}_2} \quad (4.78)$$

for the unsaturated situation, where

$$\Delta e = -x_v + \kappa e + \tilde{x}_v - \frac{1}{2}\bar{a}u_{\max} \text{sat} \left(\frac{\bar{k}_2}{u_{\max}} [\bar{f}(e) - x_v + \kappa e] \right)$$

and

$$\Delta x_v = -\kappa x_v + \kappa^2 e + \left(1 - \frac{1}{2}\kappa\right) \bar{a}u_{\max} \cdot \text{sat} \left(\frac{\bar{k}_2}{u_{\max}} [\bar{f}(e) - x_v + \kappa e] \right)$$

Following similar lines of reasoning as in [30], we can show that, under the above new conditions, the closed-loop system is stable for the situation when the control input is saturated, *i.e.* $|\bar{k}_2[\bar{f}(e(k)) - \bar{v}(k) - \tilde{x}_v(k)]| > u_{\max}$. When the control input does not exceed the saturation level, $|\bar{k}_2[\bar{f}(e(k)) - \bar{v}(k) - \tilde{x}_v(k)]| \leq u_{\max}$, the closed-loop system in Equation 4.77 can be written as

$$\begin{cases} e(k+1) = e(k) - \bar{v}(k) - \frac{1}{2}\bar{a}u_p(k) \\ \bar{v}(k+1) = \bar{v}(k) + \bar{a}u_p(k) \\ \tilde{x}_v(k+1) = (1 - \kappa)\tilde{x}_v(k) \\ u_p(k) = \bar{k}_2 [\bar{f}(e(k)) - \bar{v}(k) - \tilde{x}_v(k)] \end{cases} \quad (4.79)$$

Again, following the result of [30], we propose a Lyapunov function for the system in Equation 4.79:

$$V_p(k) = p_v \bar{v}^2(k) + \int_0^{e(k)} \bar{f}(\delta) d\delta + p_s \tilde{x}_v^2(k) \quad (4.80)$$

where p_v and p_s are appropriate positive scalars to be selected later. The increase of the Lyapunov function in Equation 4.80 along the trajectory of the closed-loop system in Equation 4.79 is given by

$$\begin{aligned} \Delta V_p(k) &= V_p(k+1) - V_p(k) \\ &= p_v [\bar{v}^2(k+1) - \bar{v}^2(k)] + \int_{e(k)}^{e(k+1)} \bar{f}(\delta) d\delta + p_s [\tilde{x}_v^2(k+1) - \tilde{x}_v^2(k)] \\ &= p_v [\bar{v}^2(k+1) - \bar{v}^2(k)] + p_s (\kappa^2 - 2\kappa) \tilde{x}_v^2(k) \\ &\quad + \bar{f}(e(k))[e(k+1) - e(k)] + \frac{1}{2} \dot{\bar{f}}(\xi)[e(k+1) - e(k)]^2 \\ &\leq p_v [\bar{v}^2(k+1) - \bar{v}^2(k)] + p_s (\kappa^2 - 2\kappa) \tilde{x}_v^2(k) \\ &\quad + \bar{f}(e(k))[e(k+1) - e(k)] + c_\xi [e(k+1) - e(k)]^2 \end{aligned} \quad (4.81)$$

where

$$c_\xi := \sup \left| \frac{1}{2} \dot{\bar{f}}(\xi) \right|$$

for ξ between $e(k)$ and $e(k+1)$, $k = 0, 1, \dots$, and the integration is expressed using the Taylor expansion as follows,

$$\begin{aligned} \int_{e(k)}^{e(k+1)} \bar{f}(\delta) d\delta &= \bar{f}(e(k))[e(k+1) - e(k)] + \frac{1}{2} \dot{\bar{f}}(\xi)[e(k+1) - e(k)]^2 \\ &\leq \bar{f}(e(k))[e(k+1) - e(k)] + c_\xi [e(k+1) - e(k)]^2 \end{aligned} \quad (4.82)$$

By Condition 6, $c_\xi \in (0, 1/4)$. Letting $\vartheta = \bar{a}\bar{k}_2$, we have

$$\bar{a}u_p(k) = \vartheta[\bar{f}(e(k)) - \bar{v}(k) - \tilde{x}_v(k)] \quad (4.83)$$

$$\begin{aligned} \bar{v}^2(k+1) - \bar{v}^2(k) &= (\vartheta^2 - 2\vartheta)\bar{v}^2(k) + \vartheta^2 \bar{f}^2(e(k)) + \vartheta^2 \tilde{x}_v^2(k) \\ &\quad + (2\vartheta - 2\vartheta^2)\bar{v}(k)\bar{f}(e(k)) - 2\vartheta^2 \tilde{x}_v(k)\bar{f}(e(k)) \\ &\quad + (2\vartheta^2 - 2\vartheta)\bar{v}(k)\tilde{x}_v(k) \end{aligned} \quad (4.84)$$

$$\begin{aligned} \bar{f}(e(k))[e(k+1) - e(k)] &= -\frac{1}{2}\vartheta \bar{f}^2(e(k)) + \left(\frac{1}{2}\vartheta - 1\right)\bar{v}(k)\bar{f}(e(k)) \\ &\quad + \frac{1}{2}\vartheta \tilde{x}_v(k)\bar{f}(e(k)) \end{aligned} \quad (4.85)$$

and

$$\begin{aligned}
[e(k+1) - e(k)]^2 &= \left(1 - \frac{1}{2}\vartheta\right)^2 \bar{v}^2(k) + \frac{1}{4}\vartheta^2 \bar{f}^2(e(k)) + \frac{1}{4}\vartheta^2 \tilde{x}_v^2(k) \\
&\quad + \left(\vartheta - \frac{1}{2}\vartheta^2\right) \bar{v}(k)\bar{f}(e(k)) - \frac{1}{2}\vartheta^2 \tilde{x}_v(k)\bar{f}(e(k)) \\
&\quad + \left(\frac{1}{2}\vartheta^2 - \vartheta\right) \bar{v}(k)\tilde{x}_v(k)
\end{aligned} \tag{4.86}$$

Thus,

$$\begin{aligned}
\Delta V_p(k) &\leq \left[p_v(\vartheta^2 - 2\vartheta) + c_\xi \left(1 - \frac{1}{2}\vartheta\right)^2 \right] \bar{v}^2(k) \\
&\quad + \left(p_v\vartheta^2 - \frac{1}{2}\vartheta + \frac{1}{4}c_\xi\vartheta^2 \right) \bar{f}^2(e(k)) \\
&\quad + \left[p_v(2\vartheta - 2\vartheta^2) + \left(\frac{1}{2}\vartheta - 1\right) + c_\xi \left(\vartheta - \frac{1}{2}\vartheta^2\right) \right] \bar{v}(k)\bar{f}(e(k)) \\
&\quad + \left[p_v\vartheta^2 + \frac{1}{4}c_\xi\vartheta^2 + p_s(\kappa^2 - 2\kappa) \right] \tilde{x}_v^2(k) \\
&\quad + \left(-2p_v\vartheta^2 + \frac{1}{2}\vartheta - \frac{1}{2}c_\xi\vartheta^2 \right) \tilde{x}_v(k)\bar{f}(e(k)) \\
&\quad + \left[p_v(2\vartheta^2 - 2\vartheta) + c_\xi \left(\frac{1}{2}\vartheta^2 - \vartheta\right) \right] \bar{v}(k)\tilde{x}_v(k) \\
&= \begin{pmatrix} \bar{v}(k) \\ \tilde{x}_v(k) \end{pmatrix}' P \begin{pmatrix} \bar{v}(k) \\ \tilde{x}_v(k) \end{pmatrix} + \begin{pmatrix} \bar{v}(k) \\ \tilde{x}_v(k) \end{pmatrix}' M \bar{f}(e(k)) + N \bar{f}^2(e(k))
\end{aligned} \tag{4.87}$$

where

$$\begin{aligned}
P &= \begin{bmatrix} p_v(\vartheta^2 - 2\vartheta) + c_\xi \left(1 - \frac{1}{2}\vartheta\right)^2 & \frac{1}{2} \left\{ p_v(2\vartheta^2 - 2\vartheta) + c_\xi \left(\frac{1}{2}\vartheta^2 - \vartheta\right) \right\} \\ \frac{1}{2} \left\{ p_v(2\vartheta^2 - 2\vartheta) + c_\xi \left(\frac{1}{2}\vartheta^2 - \vartheta\right) \right\} & p_v\vartheta^2 + \frac{1}{4}c_\xi\vartheta^2 + p_s(\kappa^2 - 2\kappa) \end{bmatrix} \\
M &= \begin{bmatrix} p_v(2\vartheta - 2\vartheta^2) + \left(\frac{1}{2}\vartheta - 1\right) + c_\xi \left(\vartheta - \frac{1}{2}\vartheta^2\right) \\ -2p_v\vartheta^2 + \frac{1}{2}\vartheta - \frac{1}{2}c_\xi\vartheta^2 \end{bmatrix} = \begin{bmatrix} m_1 \\ m_2 \end{bmatrix}
\end{aligned}$$

and

$$N = p_v\vartheta^2 - \frac{1}{2}\vartheta + \frac{1}{4}c_\xi\vartheta^2$$

Obviously, in order to ensure the right-hand side of Equation 4.87 to be negative-definite, we need to select p_v and p_s such that

$$P < 0, \quad N - \frac{1}{4}M'P^{-1}M \leq 0 \tag{4.88}$$

In particular,

$$P = \begin{bmatrix} p_{11} & p_{12} \\ p_{12} & p_{22} \end{bmatrix} < 0$$

is equivalent to

$$p_{11} < 0, \quad p_{11}p_{22} - p_{12}^2 > 0$$

for which the first inequality implies

$$0 < \vartheta < 2, \quad p_v > \frac{2 - \vartheta}{4\vartheta} c_\xi \quad (4.89)$$

and the second one is guaranteed if p_s is chosen to satisfy the following condition:

$$p_s > \frac{1}{1 - (1 - \kappa)^2} \left(-\frac{p_{12}^2}{p_{11}} + p_v \vartheta^2 + \frac{1}{4} \vartheta^2 c_\xi \right) \quad (4.90)$$

To ensure $N - M'P^{-1}M/4 \leq 0$, it requires that

$$N \leq 0, \quad (4Np_{11} - m_1^2)(4Np_{22} - m_2^2) \geq (4Np_{12} - m_1m_2)^2 \quad (4.91)$$

Noting that

$$\begin{aligned} 4Np_{11} - m_1^2 &= - \left(2p_v \vartheta - 1 + \frac{1}{2} \vartheta \right)^2 \\ 4Np_{22} - m_2^2 &= 4Np_s (\kappa^2 - 2\kappa) - \frac{1}{4} \vartheta^2 \end{aligned}$$

and

$$(4Np_{12} - m_1m_2)^2 = \frac{1}{4} \vartheta^2 \left(2p_v \vartheta - 1 + \frac{1}{2} \vartheta \right)^2$$

the inequalities in Equation 4.91 imply

$$Np_s (\kappa^2 - 2\kappa) \left(2p_v \vartheta - 1 + \frac{1}{2} \vartheta \right)^2 \leq 0 \quad (4.92)$$

Recall that $\kappa \in (0, 1)$, which implies $\kappa^2 - 2\kappa < 0$, and $p_s > 0$, $N \leq 0$. It is clear that Equation 4.92 can only be satisfied when its left-hand side is equal to 0, which implies either $N = 0$ or

$$2p_v \vartheta - 1 + \frac{1}{2} \vartheta = 0 \quad (4.93)$$

i.e.

$$p_v = \frac{2 - \vartheta}{4\vartheta} \quad (4.94)$$

Noting that Condition 1 implies $\vartheta \in (0, 2)$ and Condition 6 implies $c_\xi \in (0, 1/4)$, we have

$$p_v = \frac{2 - \vartheta}{4\vartheta} > \frac{2 - \vartheta}{4\vartheta} c_\xi \quad (4.95)$$

Hence, by the proper choices of p_v and p_s as in Equations 4.95 and 4.90, respectively, we have $\Delta V_p(k) \leq 0$. As in its continuous-time counterpart, it is straightforward to verify the closed-loop system comprising the modified PTOS control law and the given plant is asymptotically stable.

Next, we employ a reduced-order RPT control law in our MSC framework. The detailed design procedure of the RPT controller is given in Chapter 3. It is given by

$$\begin{cases} x_v(k+1) = (1-\kappa)x_v(k) + \kappa^2 e(k) + (1-\kappa/2)\bar{a}u_{\max} \cdot \text{sat}\left(\frac{u_r(k)}{u_{\max}}\right) \\ u_r(k) = (f_1 - \kappa f_2)e(k) + f_2 x_v(k) \end{cases} \quad (4.96)$$

where κ is the reduced-order observer gain and $F = [f_1 \quad f_2]$ is the state feedback gain obtained using the RPT technique. For simplicity, we rewrite the system of Equation 4.71 as

$$x_p(k+1) = A_p x_p(k) + B_p u(k) \quad (4.97)$$

where

$$x_p(k) = \begin{pmatrix} e(k) \\ \bar{v}(k) \end{pmatrix}, \quad A_p = \begin{bmatrix} 1 & -1 \\ 0 & 1 \end{bmatrix}, \quad B_p = \begin{bmatrix} -\bar{a}/2 \\ \bar{a} \end{bmatrix}$$

Letting $\tilde{x}_v(k) = x_v(k) - \bar{v}(k) - \kappa e(k)$, then the RPT controller can be expressed as

$$\tilde{x}_v(k+1) = (1-\kappa)\tilde{x}_v(k) \quad (4.98)$$

and

$$u_r(k) = F x_p(k) + f_2 \tilde{x}_v(k) \quad (4.99)$$

Let $W_x \in \mathbb{R}^{2 \times 2}$ be a positive-definite matrix and solve the following Lyapunov equation:

$$P_x = (A_p + B_p F)' P_x (A_p + B_p F) + W_x \quad (4.100)$$

for $P_x > 0$. Note that such a P_x always exists as $A_p + B_p F$ is asymptotically stable. Next, letting

$$w_s > \max \left\{ -\frac{p_{12}^2}{p_{11}} + p_v \vartheta^2 + \frac{1}{4} \vartheta^2 c_\xi \right. \\ \left. f_2 (B_p' P_x B_p + B_p' P_x (A_p + B_p F) W_x^{-1} (A_p + B_p F)' P_x B_p) f_2 \right\} \quad (4.101)$$

and

$$p_s = \frac{w_s}{1 - (1-\kappa)^2} > \frac{1}{1 - (1-\kappa)^2} \left[-\frac{p_{12}^2}{p_{11}} + p_v \vartheta^2 + \frac{1}{4} \vartheta^2 c_\xi \right] \quad (4.102)$$

we define a set

$$\mathbf{X} := \left\{ \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix} : \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix}' \begin{bmatrix} P_x & 0 \\ 0 & p_s \end{bmatrix} \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix} \leq c \right\} \quad (4.103)$$

where $c > 0$ is the largest positive scalar such that

$$\begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix} \in \mathbf{X} \Rightarrow \left| [F \quad f_2] \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix} \right| \leq u_{\max} \quad (4.104)$$

For all $\begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix} \in X$, the resulting closed-loop system can then be written as

$$\begin{pmatrix} x_p(k+1) \\ \tilde{x}_v(k+1) \end{pmatrix} = \begin{bmatrix} A_p + B_p F & B_p f_2 \\ 0 & 1 - \kappa \end{bmatrix} \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix} \quad (4.105)$$

which is clearly an asymptotically stable system. Define a Lyapunov function,

$$V_R(k) = \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix}' \begin{bmatrix} P_x & 0 \\ 0 & p_s \end{bmatrix} \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix} \quad (4.106)$$

and evaluate its increase along the trajectories of the closed-loop system in Equation 4.105 as follows,

$$\begin{aligned} \Delta V_R(k) &= \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix}' \begin{bmatrix} -W_x & (A_p + B_p F)' P_x B_p f_2 \\ f_2 B_p' P_x (A_p + B_p F) & -w_s + f_2 B_p' P_x B_p f_2 \end{bmatrix} \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix} \\ &= \begin{pmatrix} \hat{x}(k) \\ \tilde{x}_v(k) \end{pmatrix}' \begin{bmatrix} -W_x & 0 \\ 0 & -\hat{w}_s \end{bmatrix} \begin{pmatrix} \hat{x}(k) \\ \tilde{x}_v(k) \end{pmatrix} \\ &\leq 0 \end{aligned} \quad (4.107)$$

where

$$\hat{x}(k) = x_p(k) - W_x^{-1} (A_p + B_p F)' P_x B_p f_2 \tilde{x}_v(k)$$

and

$$\hat{w}_s = w_s - f_2 \left[B_p' P_x B_p + B_p' P_x (A_p + B_p F) W_x^{-1} (A_p + B_p F)' P_x B_p \right] f_2 > 0$$

Hence, all trajectories of the closed-loop system in Equation 4.105 starting from \mathbf{X} remain in the set and converge asymptotically to the origin. Hence, the closed-loop system comprising the reduced-order RPT controller is asymptotically stable provided that the control input is not saturated.

Next, we re-express Equation 4.80 using the Taylor expansion as follows,

$$\begin{aligned} V_p(k) &= p_v \bar{v}^2(k) + \frac{1}{2} \tilde{f}(\tau) e^2(k) + p_s \tilde{x}_v^2(k) \\ &= \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix}' \begin{bmatrix} \tilde{f}(\tau)/2 & 0 & 0 \\ 0 & p_v & 0 \\ 0 & 0 & p_s \end{bmatrix} \begin{pmatrix} x_p(k) \\ \tilde{x}_v(k) \end{pmatrix} \end{aligned} \quad (4.108)$$

where τ is an appropriate scalar between 0 and $e(k)$. Let

$$\sigma = \min \left\{ \frac{1}{2} \tilde{f}(\tau), p_v, p_s \right\} / \max \{ \lambda_{\max}(P_x), p_s \} \quad (4.109)$$

The MSC law is given by

$$u(k) = \begin{cases} u_p(k), & k < k_s \\ u_r(k), & k \geq k_s \end{cases} \quad (4.110)$$

where k_s is such that

$$\begin{pmatrix} x_p(k_s) \\ \tilde{x}_v(k_s) \end{pmatrix} \in \mathbf{X} \quad \text{and} \quad |e(k_s)| \leq y_\ell \quad (4.111)$$

where y_ℓ is the size of the linear region of the PTOS control law. The Lyapunov function for the overall closed-loop system can be chosen as

$$V(k) = V_p(k)[1 - \mathbf{1}(k - k_s)] + \sigma V_r(k)\mathbf{1}(k - k_s) \quad (4.112)$$

where

$$\mathbf{1}(k - k_s) = \begin{cases} 0, & k < k_s \\ 1, & k \geq k_s \end{cases}$$

It is simple to verify that

$$\begin{aligned} \Delta V(k) &= \Delta V_p(k)[1 - \mathbf{1}(k + 1 - k_s)] + \sigma \Delta V_r(k)\mathbf{1}(k + 1 - k_s) \\ &\quad + (\sigma V_r(k) - V_p(k))[\mathbf{1}(k + 1 - k_s) - \mathbf{1}(k - k_s)] \end{aligned} \quad (4.113)$$

It has already been proved that the increase of the Lyapunov function $V_p(k)$ and $V_r(k)$ are negative-definite when they are effective, respectively. The last item is always negative in view of the definition of σ in Equation 4.109. Hence, $\Delta V(k) \leq 0$ and the resulting closed-loop system comprising the given plant and the MSC control law is asymptotically stable. Furthermore, Equation 4.111 gives the mode-switching condition for the proposed MSC scheme.

Composite Nonlinear Feedback Control

5.1 Introduction

The PTOS and MSC schemes discussed in the previous chapter have two controllers, *i.e.* a nonlinear controller and a linear controller, operating in two different time stages. The switching elements in these control schemes generally yield a system response not as smooth as what we would expect and thus limit the overall performance of the corresponding servo systems. Furthermore, both schemes do not pay any special attention to the transient performance of the overall design, which is in fact one of the important issues in tracking control problems. In general, quick response and small overshoot are desirable in most practical situations. However, it is well understood that quick response is often accompanied by a large overshoot. Thus, most of the design schemes have to make a tradeoff between these two transient performance indices. Inspired by a recent study of Lin *et al.* [130], which was introduced to improve the tracking performance under state feedback laws for a class of second-order systems subject to actuator saturation, we have developed in this chapter a nonlinear control technique, the so-called composite nonlinear feedback (CNF) control, for a more general class of systems with measurement feedback. It is applicable to systems with or without external disturbances.

The CNF control consists of a linear feedback law and a nonlinear feedback law without any switching element. Unlike the MSC and PTOS control laws, both the linear and nonlinear controllers in CNF are in operation all the time. The linear feedback part is designed to yield a closed-loop system with a small damping ratio for a quick response, while at the same time not exceeding the actuator limits for the desired command input levels. The nonlinear feedback law is used to increase the damping ratio of the closed-loop system as the system output approaches the target reference to reduce the overshoot caused by the linear part. We would like to note that the design philosophy of the CNF technique is very different from the commonly used antiwindup technique. The CNF control applies more control effort to the system around the final stage when the controlled output is approaching to the target reference. It focuses on reducing overshoot and speeding up its settling time.

On the other hand, the antiwindup design focuses primarily on the initial stage to reduce the effect of saturation.

We show by an example that such a technique could yield a better performance compared to that of the time-optimal control in asymptotic tracking. It is noted that the new control scheme can be utilized to design servo systems that deal with asymptotic target tracking or “point-and-shoot” fast targeting. As will be seen soon in the forthcoming chapters, this new control method has improved the performance of the overall servo system by a great deal.

Since the initiation of CNF in Lin *et al.* [130] for a class of second-order systems, there have been efforts to generalize it to more general systems. For example, Turner *et al.* [131] extended the results of [130] to higher-order and multiple-input systems. This extension was made under a restrictive assumption on the system that excludes many systems including those originally considered in [130]. Also, as in [130], only state feedback is considered in [131]. A fairly complete study dealing with general linear continuous-time systems with measurement feedback has been reported in Chen *et al.* [132], whereas its discrete-time counterpart has been reported in Venkataramanan *et al.* [133]. The technique has recently been extended to solve general multivariable systems with measurement feedback by He *et al.* [134, 135], and to solve a class of nonlinear systems in [136, 137]. The result for tackling systems with external disturbances has recently been reported by Peng *et al.* [138].

Although the CNF technique is fairly mature and complete for solving general linear systems, we would, however, restrict our attention to single-input systems in this chapter for simplicity. The results presented in this chapter are sufficient in designing HDD servo systems in the forthcoming chapters. The results for multi-input systems and nonlinear systems are rather involved. We refer interested readers to the references cited above.

The outline of this chapter is as follows: In Section 5.2, we present the technique of CNF control for general single-input continuous-time systems without and with external disturbances. The discrete-time version is given in Section 5.3. Section 5.4 shows in a numerical example that the CNF control is capable of outperforming the time-optimal control in asymptotic tracking situations. Finally, in Section 5.5, we present a user-friendly software toolkit for the CNF control design for continuous-time systems with and without disturbances. The toolkit is written in MATLAB® together with its simulation package Simulink®. It is available for free downloading at the website <http://hdd.ece.nus.edu.sg/~bmchen>.

5.2 Continuous-time Systems

We first present the result for systems without external disturbances and then extend it to the case when the given system has external disturbances. For the latter, an integrator is introduced in the controller to remove or minimize the steady-state bias due to unknown constant disturbances. The selection of nonlinear design parameters is also presented together with an interpretation of the nonlinear gain introduced in the CNF design under the well-known framework of the classical root-locus theory.

5.2.1 Systems without External Disturbances

Consider a linear continuous-time system Σ with an amplitude-constrained actuator characterized by

$$\begin{cases} \dot{x} = A x + B \text{sat}(u), & x(0) = x_0 \\ y = C_1 x \\ h = C_2 x \end{cases} \quad (5.1)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}$, $y \in \mathbb{R}^p$ and $h \in \mathbb{R}$ are, respectively, the state, control input, measurement output and controlled output of Σ . A , B , C_1 and C_2 are appropriate dimensional constant matrices, and $\text{sat}: \mathbb{R} \rightarrow \mathbb{R}$ represents the actuator saturation defined as

$$\text{sat}(u) = \text{sgn}(u) \min\{u_{\max}, |u|\} \quad (5.2)$$

with u_{\max} being the saturation level of the input. The following assumptions on the system matrices are required:

1. (A, B) is stabilizable,
2. (A, C_1) is detectable, and
3. (A, B, C_2) is invertible and has no invariant zeros at $s = 0$.

The objective here is to design a CNF control law that causes the output to track a high-amplitude step input rapidly without experiencing large overshoot and without the adverse actuator saturation effects. This is done through the design of a linear feedback law with a small closed-loop damping ratio and a nonlinear feedback law through an appropriate Lyapunov function to cause the closed-loop system to be highly damped as the system output approaches the command input to reduce the overshoot. Similar to the design of linear controllers discussed in Chapter 4, we separate the CNF controller design into three distinct situations 1) the state feedback case, 2) the full-order measurement feedback case, and 3) the reduced-order measurement feedback case.

i. State Feedback Case. We proceed to develop a composite nonlinear feedback control technique for the case when all the states of the plant Σ are measurable, *i.e.* $y = x$. It is done in three steps. In the first step, a linear feedback control law is designed; in the second step, the design of nonlinear feedback control is carried out. Lastly, in the third step, the linear and nonlinear feedback laws are combined to give a composite nonlinear feedback control law. Again, we note that the procedure given for this case follows closely from that reported in Lin *et al.* [130], although our result is applicable to a much larger class of systems.

STEP 5.C.S.1: design a linear feedback law,

$$u_c = Fx + Gr \quad (5.3)$$

where F is chosen such that 1) $A + BF$ is an asymptotically stable matrix, and 2) the closed-loop system $C_2(sI - A - BF)^{-1}B$ has certain desired properties, *e.g.*, having a small damping ratio. We note that such an F can be designed using

methods such as the H_2 and H_∞ optimization approaches, as well as the RPT technique given in Chapter 3. Furthermore, G is a scalar and is given by

$$G = - [C_2(A + BF)^{-1}B]^{-1} \quad (5.4)$$

and r is a step command input. Here we note that G is well defined because $A + BF$ is stable, and the triple (A, B, C_2) is invertible and has no invariant zeros at $s = 0$.

The following lemma determines the magnitude of r that can be tracked by such a control law without exceeding the control limits.

Lemma 5.1. *Given a positive-definite matrix $W \in \mathbb{R}^{n \times n}$, let $P > 0$ be the solution of the following Lyapunov equation:*

$$(A + BF)'P + P(A + BF) = -W \quad (5.5)$$

Such a P exists since $A + BF$ is asymptotically stable. For any $\delta \in (0, 1)$, let $c_\delta > 0$ be the largest positive scalar satisfying the following condition:

$$|Fx| \leq u_{\max}(1 - \delta), \quad \forall x \in \mathbf{X}_\delta := \left\{ x : x'Px \leq c_\delta \right\} \quad (5.6)$$

Also, let

$$H := [1 - F(A + BF)^{-1}B]G \quad (5.7)$$

and

$$x_e := G_e r := -(A + BF)^{-1}BG r \quad (5.8)$$

Then, the control law of Equation 5.3 is capable of driving the controlled output h to track asymptotically a step command input r , provided that the initial state x_0 and r satisfy:

$$\tilde{x}_0 := (x_0 - x_e) \in \mathbf{X}_\delta \quad \text{and} \quad |Hr| \leq \delta u_{\max} \quad (5.9)$$

Proof. Let us first define a new state variable

$$\tilde{x} = x - x_e \quad (5.10)$$

It is simple to verify that the linear feedback control law of Equation 5.3 can be rewritten as

$$u_c(t) = F\tilde{x}(t) + [1 - F(A + BF)^{-1}B]Gr = F\tilde{x}(t) + Hr \quad (5.11)$$

and hence for all $\tilde{x} \in \mathbf{X}_\delta$ and, provided that $|Hr| \leq \delta u_{\max}$, the closed-loop system is linear and is given by

$$\dot{\tilde{x}} = (A + BF)\tilde{x} + Ax_e + B Hr \quad (5.12)$$

Noting that

$$\begin{aligned}
Ax_e + B Hr &= \left\{ B[1 - F(A + BF)^{-1}B]G - A(A + BF)^{-1}BG \right\} r \\
&= \left\{ [I - BF(A + BF)^{-1}]BG - A(A + BF)^{-1}BG \right\} r \\
&= \left\{ I - BF(A + BF)^{-1} - A(A + BF)^{-1} \right\} BGr = 0 \quad (5.13)
\end{aligned}$$

the closed-loop system in Equation 5.12 can then be simplified as

$$\dot{\tilde{x}} = (A + BF)\tilde{x} \quad (5.14)$$

Next, we define a Lyapunov function:

$$V(\tilde{x}) = \tilde{x}' P \tilde{x} \quad (5.15)$$

Along the trajectories of the closed-loop system in Equation 5.14, $V(\tilde{x})$ satisfies

$$\dot{V}(\tilde{x}) = -\tilde{x}' W \tilde{x} \quad (5.16)$$

which implies that $V(\tilde{x})$ is a monotonically decreasing function with respect to t along the trajectories of Equation 5.14. Thus, all trajectories of Equation 5.14 starting from \mathbf{X}_δ remain there and converge asymptotically to the origin. For an initial state x_0 and the step command input r that satisfy Equation 5.9, we have

$$\lim_{t \rightarrow \infty} x(t) = x_e \quad (5.17)$$

and hence

$$\lim_{t \rightarrow \infty} h(t) = \lim_{t \rightarrow \infty} C_2 x(t) = C_2 x_e = -C_2(A + BF)^{-1} BGr = r \quad (5.18)$$

This completes the proof of Lemma 5.1. \diamond

Remark 5.2. We would like to note that for the case when $x_0 = 0$, any step command of amplitude r can be asymptotically tracked if

$$|r| \leq [c_\delta (G_e' P G_e)^{-1}]^{1/2} \quad \text{and} \quad |Hr| \leq \delta \cdot u_{\max} \quad (5.19)$$

Clearly, the trackable amplitudes of reference inputs by the linear feedback control law can be increased by increasing δ and/or decreasing $G_e' P G_e$ through the choice of W . However, the change in gain F indeed affects the damping ratio of the closed-loop system and hence its rise time. \diamond

STEP 5.C.S.2: the nonlinear feedback control law $u_N(t)$ is given by

$$u_N = \rho(r, y) B' P (x - x_e) \quad (5.20)$$

where $\rho(r, y)$ is any nonpositive function locally Lipschitz in y , which is used to change the system closed-loop damping ratio as the output approaches the step command input. The choice of such a $\rho(r, y)$ is to be discussed later.

STEP 5.C.S.3: the linear and nonlinear feedback laws derived in the previous steps are now combined to form a CNF controller:

$$u = u_L + u_N = Fx + Gr + \rho(r, y)B'P(x - x_e) \quad (5.21)$$

We now move to prove that the closed-loop system comprising the given plant in Equation 5.1 and the CNF control law of Equation 5.21 is asymptotically stable.

Theorem 5.3. *Consider the given system in Equation 5.1. Then, for any nonpositive function $\rho(r, y)$, locally Lipschitz in y , the composite nonlinear feedback law in Equation 5.21 internally stabilizes the given plant and drive the system controlled output $h(t)$ to track asymptotically the step command input of amplitude r from an initial state x_0 , provided that x_0 and r satisfy Equation 5.9. \diamond*

Proof. Again, let $\tilde{x} = x - x_e$. Then, the closed-loop system comprising the given plant in Equation 5.1 and the CNF control law of Equation 5.21 can be expressed as

$$\dot{\tilde{x}} = (A + BF)\tilde{x} + Bw \quad (5.22)$$

where

$$w = \text{sat}(F\tilde{x} + Hr + u_N) - F\tilde{x} - Hr \quad (5.23)$$

Clearly, for the given x_0 satisfying Equation 5.9, we have $\tilde{x}_0 = (x_0 - x_e) \in \mathbf{X}_\delta$. Using the Lyapunov function

$$V = \tilde{x}'P\tilde{x} \quad (5.24)$$

we can evaluate the derivative of V along the trajectories of the closed-loop system in Equation 5.22, *i.e.*

$$\begin{aligned} \dot{V} &= \dot{\tilde{x}}'P\tilde{x} + \tilde{x}'P\dot{\tilde{x}} \\ &= \tilde{x}'(A + BF)'P\tilde{x} + \tilde{x}'P(A + BF)\tilde{x} + 2\tilde{x}'PBw \\ &= -\tilde{x}'W\tilde{x} + 2\tilde{x}'PBw \end{aligned} \quad (5.25)$$

Note that for all

$$\tilde{x} \in \mathbf{X}_\delta = \{\tilde{x} : \tilde{x}'P\tilde{x} \leq c_\delta\} \Rightarrow |F\tilde{x}| \leq u_{\max}(1 - \delta) \quad (5.26)$$

We next calculate \dot{V} for three different values of saturation function.

Case 1. If $|F\tilde{x} + Hr + u_N| \leq u_{\max}$, then $w = u_N = \rho B'P\tilde{x}$ and thus

$$\dot{V} = -\tilde{x}'W\tilde{x} + 2\rho\tilde{x}'PBB'P\tilde{x} \leq -\tilde{x}'W\tilde{x} \quad (5.27)$$

Case 2. If $F\tilde{x} + Hr + u_N > u_{\max}$, and by construction $|F\tilde{x} + Hr| \leq u_{\max}$, we have

$$0 < w = u_{\max} - F\tilde{x} - Hr < u_N = \rho B'P\tilde{x} \quad (5.28)$$

which implies that $\tilde{x}'PB < 0$ and hence $\dot{V} = -\tilde{x}'W\tilde{x} + 2\tilde{x}'PBw \leq -\tilde{x}'W\tilde{x}$.

Case 3. Finally, if $F\tilde{x} + Hr + u_N < -u_{\max}$, we have

$$\rho B' P \tilde{x} = u_N < w = -u_{\max} - F\tilde{x} - Hr < 0 \quad (5.29)$$

implying $\tilde{x}'PB > 0$ and hence $\dot{V} \leq -\tilde{x}'W\tilde{x}$.

In conclusion, we have shown that

$$\dot{V} \leq -\tilde{x}'W\tilde{x}, \quad \tilde{x} \in \mathbf{X}_\delta \quad (5.30)$$

which implies that \mathbf{X}_δ is an invariant set of the closed-loop system in Equation 5.22 and all trajectories of Equation 5.22 starting from inside \mathbf{X}_δ converge to the origin. This, in turn, indicates that, for all initial states x_0 and the step command input of amplitude r that satisfy Equation 5.9,

$$\lim_{t \rightarrow \infty} x(t) = x_e \quad (5.31)$$

Therefore,

$$\lim_{t \rightarrow \infty} h(t) = \lim_{t \rightarrow \infty} C_2 x(t) = C_2 x_e = r \quad (5.32)$$

This completes the proof of Theorem 5.3. \diamond

Remark 5.4. Theorem 5.3 shows that the additional nonlinear feedback control law u_N , as given in Equation 5.20, does not affect the ability of the closed-loop system to track the command input. Any command input that can be asymptotically tracked by the linear feedback law of Equation 5.3 can also be asymptotically tracked by the CNF control law in Equation 5.21. However, this additional term u_N in the CNF control law can be used to improve the performance of the overall closed-loop system. This is the key property of the CNF control technique. \diamond

ii. Full-order Measurement Feedback Case. The assumption that all the states of Σ are measurable is, in general, not practical. For example, in HDD servo systems, the velocity of the actuator is not usually directly measurable. Thus, one has to implement the PTOS controller or the control law obtained in the previous case via a certain velocity estimation. In what follows, we proceed to develop a CNF design using only measurement information.

STEP 5.C.F.1: we first construct a linear full-order measurement feedback control law:

$$\Sigma_F : \begin{cases} \dot{x}_v = (A + KC_1)x_v - Ky + B \text{sat}(u_L) \\ u_L = F(x_v - x_e) + Hr \end{cases} \quad (5.33)$$

where r is the reference input and $x_v \in \mathbb{R}^n$ is the state of the controller. As usual, F and K are gain matrices and are designed such that $A + BF$ and $A + KC_1$ are asymptotically stable and the resulting closed-loop system has the desired properties. As defined in Equations 5.7 and 5.8,

$$H = [1 - F(A + BF)^{-1}B]G, \quad G = -[C_2(A + BF)^{-1}B]^{-1} \quad (5.34)$$

and $x_e = -(A + BF)^{-1}BGr$.

We have the following result.

Lemma 5.5. *Given a positive-definite matrix $W_p \in \mathbb{R}^{n \times n}$, let $P > 0$ be the solution to the Lyapunov equation*

$$(A + BF)'P + P(A + BF) = -W_p \quad (5.35)$$

Given another positive-definite matrix $W_q \in \mathbb{R}^{n \times n}$ with

$$W_q > F' B' P W_p^{-1} P B F \quad (5.36)$$

let $Q > 0$ be the solution to the Lyapunov equation

$$(A + KC_1)'Q + Q(A + KC_1) = -W_q \quad (5.37)$$

Note that such P and Q exist as $A + BF$ and $A + KC_1$ are asymptotically stable. For any $\delta \in (0, 1)$, let c_δ be the largest positive scalar such that for all

$$\left(\begin{array}{c} x \\ x_v \end{array} \right) \in \mathbf{X}_{r\delta} := \left\{ \left(\begin{array}{c} x \\ x_v \end{array} \right) : \left(\begin{array}{c} x \\ x_v \end{array} \right)' \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \left(\begin{array}{c} x \\ x_v \end{array} \right) \leq c_\delta \right\} \quad (5.38)$$

we have

$$\left| \begin{bmatrix} F & F \end{bmatrix} \left(\begin{array}{c} x \\ x_v \end{array} \right) \right| \leq u_{\max}(1 - \delta) \quad (5.39)$$

The linear control law in Equation 5.33 drives the system controlled output $h(t)$ to track asymptotically a step command input of amplitude r from an initial state x_0 , provided that $x_0, x_{v0} = x_v(0)$ and r satisfy:

$$|Hr| \leq \delta \cdot u_{\max} \quad \text{and} \quad \left(\begin{array}{c} x_0 - x_e \\ x_{v0} - x_0 \end{array} \right) \in \mathbf{X}_{r\delta} \quad (5.40)$$

Proof. Let us transform the system coordinate by defining

$$\tilde{x} = x - x_e \quad \text{and} \quad \tilde{x}_v = x_v - x \quad (5.41)$$

Then, the linear control law of Equation 5.33 can be written as

$$\dot{\tilde{x}}_v = (A + KC_1)\tilde{x}_v, \quad u_L = \begin{bmatrix} F & F \end{bmatrix} \left(\begin{array}{c} \tilde{x} \\ \tilde{x}_v \end{array} \right) + Hr \quad (5.42)$$

Hence, for all states

$$\left(\begin{array}{c} \tilde{x} \\ \tilde{x}_v \end{array} \right) \in \mathbf{X}_{r\delta} \Rightarrow \left| \begin{bmatrix} F & F \end{bmatrix} \left(\begin{array}{c} \tilde{x} \\ \tilde{x}_v \end{array} \right) \right| \leq u_{\max}(1 - \delta) \quad (5.43)$$

and for any r satisfying

$$|Hr| \leq \delta \cdot u_{\max} \quad (5.44)$$

we have

$$|u_L| = \left| \begin{bmatrix} F & F \end{bmatrix} \left(\begin{array}{c} \tilde{x} \\ \tilde{x}_v \end{array} \right) + Hr \right| \leq \left| \begin{bmatrix} F & F \end{bmatrix} \left(\begin{array}{c} \tilde{x} \\ \tilde{x}_v \end{array} \right) \right| + |Hr| = u_{\max} \quad (5.45)$$

Thus, for all \tilde{x} and \tilde{x}_v satisfying the condition as given in Equation 5.43, the closed-loop system comprising the given plant and the linear control law of Equation 5.33 can be rewritten as

$$\begin{pmatrix} \dot{\tilde{x}} \\ \dot{\tilde{x}}_v \end{pmatrix} = \begin{bmatrix} A + BF & BF \\ 0 & A + KC_1 \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \quad (5.46)$$

Next, we define a Lyapunov function for the closed-loop system in Equation 5.46:

$$V = \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix}' \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \quad (5.47)$$

Along the trajectories of the closed-loop system in Equation 5.46 the derivative of the Lyapunov function is given by

$$\begin{aligned} \dot{V} &= \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix}' \begin{bmatrix} -W_p & PBF \\ F'B'P & -W_q \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \\ &= \begin{pmatrix} \hat{x} \\ \tilde{x}_v \end{pmatrix}' \begin{bmatrix} -W_p & 0 \\ 0 & -\tilde{W}_q \end{bmatrix} \begin{pmatrix} \hat{x} \\ \tilde{x}_v \end{pmatrix} \end{aligned} \quad (5.48)$$

where

$$\hat{x} = \tilde{x} - W_p^{-1}PBF\tilde{x}_v, \quad \tilde{W}_q = W_q - F'B'PW_p^{-1}PBF \quad (5.49)$$

With the choice of W_q satisfying Equation 5.36, it is obvious that $\dot{V} \leq 0$. This shows that $\mathbf{X}_{r\delta}$ is an invariant set of the closed-loop system in Equation 5.46 and all trajectories starting from the set converge asymptotically to the origin. Thus, for the initial states of x_0 and x_{v0} and step command inputs that satisfy Equation 5.40,

$$\lim_{t \rightarrow \infty} \tilde{x}_v(t) = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} x(t) = x_e \quad (5.50)$$

which imply

$$\lim_{t \rightarrow \infty} h(t) = \lim_{t \rightarrow \infty} C_2 x(t) = C_2 x_e = r \quad (5.51)$$

This completes the proof of Lemma 5.5. \diamond

STEP 5.C.F.2: as in the state feedback case, the linear control law of Equation 5.33 obtained in the above step is to be combined with a nonlinear control law to form the following CNF controller:

$$\begin{cases} \dot{x}_v = (A + KC_1)x_v - Ky + B \text{sat}(u) \\ u = F(x_v - x_e) + Hr + \rho(r, y)B'P(x_v - x_e) \end{cases} \quad (5.52)$$

where $\rho(r, y)$ is a nonpositive scalar function, locally Lipschitz in y , and is to be chosen to improve the performance of the closed-loop system.

It turns out that, for the measurement feedback case, the choice of $\rho(r, y)$, the nonpositive scalar function, is not totally free. It is subject to certain constraints. We have the following theorem.

Theorem 5.6. Consider the given system in Equation 5.1. Then, there exists a scalar $\rho^* > 0$ such that for any nonpositive function $\rho(r, y)$, locally Lipschitz in y and $|\rho(r, y)| \leq \rho^*$, the CNF control law of Equation 5.52 internally stabilizes the given plant and drive the system controlled output $h(t)$ to track asymptotically the step command input of amplitude r from an initial state x_0 , provided that x_0 , x_{v0} and r satisfy Equation 5.40. \diamond

Proof. Again, let $\tilde{x} = x - x_e$ and $\tilde{x}_v = x_v - x$. For simplicity, we drop r and h in $\rho(r, y)$ throughout the rest of the proof of this theorem. Then, the closed-loop system with the CNF control law of Equation 5.52 can be expressed as

$$\begin{pmatrix} \dot{\tilde{x}} \\ \dot{\tilde{x}}_v \end{pmatrix} = \begin{bmatrix} A + BF & BF \\ 0 & A + KC_1 \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} w \quad (5.53)$$

where

$$w = \text{sat} \left[\begin{bmatrix} F & F \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + Hr + \rho \begin{bmatrix} B'P & B'P \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \right] - \begin{bmatrix} F & F \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} - Hr \quad (5.54)$$

Clearly, for the given x_0 and x_{v0} satisfying Equation 5.40, we have

$$\begin{pmatrix} \tilde{x}(0) \\ \tilde{x}_v(0) \end{pmatrix} \in \mathbf{X}_{F\delta} \quad (5.55)$$

Using the following Lyapunov function:

$$V = \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix}' \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \quad (5.56)$$

we evaluate the derivative of V along the trajectories of the closed-loop system in Equation 5.53, *i.e.*

$$\dot{V} = \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix}' \begin{bmatrix} -W_p & PBF \\ F'B'P & -W_q \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + 2\tilde{x}'PBw \quad (5.57)$$

Note that for all

$$\begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \in \mathbf{X}_{F\delta} \Rightarrow \left| \begin{bmatrix} F & F \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \right| \leq u_{\max}(1 - \delta) \quad (5.58)$$

Again, as is done in the full state feedback case, let us find the above derivative of V for three different cases.

Case 1. If

$$\left| \begin{bmatrix} F & F \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + Hr + \rho \begin{bmatrix} B'P & B'P \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \right| \leq u_{\max} \quad (5.59)$$

then

$$w = \rho[B'P \quad B'P] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \quad (5.60)$$

which implies

$$\begin{aligned} \dot{V} &= \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix}' \begin{bmatrix} -W_p & PB(F + \rho B'P) \\ (F + \rho B'P)'B'P & -W_q \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + 2\rho\tilde{x}'PBB'P\tilde{x} \\ &\leq \begin{pmatrix} \hat{x} \\ \tilde{x}_v \end{pmatrix}' \begin{bmatrix} -W_p & 0 \\ 0 & -\tilde{W}_q \end{bmatrix} \begin{pmatrix} \hat{x} \\ \tilde{x}_v \end{pmatrix} \end{aligned} \quad (5.61)$$

where

$$\hat{x} = \tilde{x} - W_p^{-1}PB(F + \rho B'P)\tilde{x}_v, \quad (5.62)$$

$$\tilde{W}_q = W_q - (F + \rho B'P)'B'PW_p^{-1}PB(F + \rho B'P) \quad (5.63)$$

Noting Equation 5.36, *i.e.* $W_q > F'B'PW_p^{-1}PBF$, and $\rho(r, y)$ is locally Lipschitz, it is clear that there exists a $\rho_1^* > 0$ such that for any scalar function with $|\rho(r, y)| \leq \rho_1^*$ we have $\tilde{W}_q > 0$ and hence $\dot{V} \leq 0$.

Case 2. If

$$[F \quad F] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + Hr + \rho[B'P \quad B'P] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} > u_{\max} \quad (5.64)$$

then for the trajectories inside $\mathbf{X}_{r\delta}$,

$$\left| [F \quad F] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + Hr \right| \leq u_{\max} \quad (5.65)$$

which implies that

$$0 < w < \rho[B'P \quad B'P] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \quad (5.66)$$

Next, let us express

$$w = q\rho[B'P \quad B'P] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \quad (5.67)$$

for an appropriate positive piecewise continuous function $q(t)$, bounded by 1 for all t . In this case, the derivative of V becomes

$$\begin{aligned} \dot{V} &= \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix}' \begin{bmatrix} -W_p & PB(F + q\rho B'P) \\ (F + q\rho B'P)'B'P & -W_q \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + 2q\rho\tilde{x}'PBB'P\tilde{x} \\ &\leq \begin{pmatrix} \hat{x}_+ \\ \tilde{x}_v \end{pmatrix}' \begin{bmatrix} -W_p & 0 \\ 0 & -\tilde{W}_{q+} \end{bmatrix} \begin{pmatrix} \hat{x}_+ \\ \tilde{x}_v \end{pmatrix} \end{aligned} \quad (5.68)$$

where

$$\hat{x}_+ = \tilde{x} - W_p^{-1}PB(F + q\rho B'P)\tilde{x}_v \quad (5.69)$$

$$\tilde{W}_{Q+} = W_Q - (F + q\rho B'P)'B'PW_p^{-1}PB(F + q\rho B'P) \quad (5.70)$$

Again, noting Equation 5.36, it can be shown that there exists a ρ_2^* such that for any $\rho(r, y)$ satisfying $|\rho(r, y)| \leq \rho_2^*$ we have $\tilde{W}_{Q+} > 0$ and hence $\dot{V} \leq 0$.

Case 3. Similarly, for the case when

$$[F \quad F] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + Hr + \rho[B'P \quad B'P] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} < -u_{\max} \quad (5.71)$$

we can show that there exists a $\rho_3^* > 0$ such that for any $\rho(r, y)$ satisfying $|\rho(r, y)| \leq \rho_3^*$, we have $\dot{V} \leq 0$ for all the trajectories in $\mathbf{X}_{F\delta}$.

Finally, let $\rho^* = \min\{\rho_1^*, \rho_2^*, \rho_3^*\}$. Then, we have for any nonpositive scalar function $\rho(r, y)$ satisfying $|\rho(r, y)| \leq \rho^*$,

$$\dot{V} \leq 0, \quad \forall \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} \in \mathbf{X}_{F\delta} \quad (5.72)$$

Thus, $\mathbf{X}_{F\delta}$ is an invariant set of the closed-loop system in Equation 5.53, and all trajectories starting from $\mathbf{X}_{F\delta}$ remain inside and asymptotically converge to the origin. This, in turn, indicates that, for the initial state of the given system x_0 , the initial state of the controller x_{v0} , and step command input r that satisfy Equation 5.40,

$$\lim_{t \rightarrow \infty} \tilde{x}_v(t) = 0 \quad \text{and} \quad \lim_{t \rightarrow \infty} x(t) = x_e \quad (5.73)$$

and hence

$$\lim_{t \rightarrow \infty} h(t) = \lim_{t \rightarrow \infty} C_2 x(t) = C_2 x_e = r \quad (5.74)$$

This completes the proof of Theorem 5.6. \diamond

iii. Reduced-order Measurement Feedback Case. For the given system in Equation 5.1, it is clear that there are p states of the system measurable if C_1 is of maximal rank. Thus, in general, it is not necessary to estimate these measurable states in measurement feedback laws. As such, we design a dynamic controller that has a dynamical order less than that of the given plant. We now proceed to construct such a control law under the CNF control framework.

For simplicity of presentation, we assume that C_1 is already in the form

$$C_1 = [I_p \quad 0] \quad (5.75)$$

Then, the system in Equation 5.1 can be rewritten as

$$\begin{cases} \begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \text{sat}(u), & x_0 = \begin{pmatrix} x_{10} \\ x_{20} \end{pmatrix} \\ y = [I_p \quad 0] \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ h = C_2 \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \end{cases} \quad (5.76)$$

where the original state x is partitioned into two parts, x_1 and x_2 with $y \equiv x_1$. Thus, we only need to estimate x_2 in the reduced-order measurement feedback design. Next, we let F be chosen such that 1) $A + BF$ is asymptotically stable, and 2) $C_2(sI - A - BF)^{-1}B$ has the desired properties, and let K_R be chosen such that $A_{22} + K_R A_{12}$ is asymptotically stable. Here, we note that it was shown in Chen [110] that (A_{22}, A_{12}) is detectable if and only if (A, C_1) is detectable. Thus, there exists a stabilizing K_R . Again, such F and K_R can be designed using any of the linear control techniques presented in Chapter 3. We then partition F in conformity with x_1 and x_2 :

$$F = [F_1 \quad F_2] \quad (5.77)$$

As defined in Equations 5.7 and 5.8,

$$H = [1 - F(A + BF)^{-1}B]G, \quad G = -[C_2(A + BF)^{-1}B]^{-1} \quad (5.78)$$

and $x_e = -(A + BF)^{-1}BGr$.

The reduced-order CNF controller is given by

$$\begin{aligned} \dot{x}_v = & (A_{22} + K_R A_{12})x_v + [A_{21} + K_R A_{11} - (A_{22} + K_R A_{12})K_R]y \\ & + (B_2 + K_R B_1) \text{sat}(u) \end{aligned} \quad (5.79)$$

and

$$u = F \left[\begin{pmatrix} y \\ x_v - K_R y \end{pmatrix} - x_e \right] + Hr + \rho(r, y) B' P \left[\begin{pmatrix} y \\ x_v - K_R y \end{pmatrix} - x_e \right] \quad (5.80)$$

where $\rho(r, y)$ is a nonpositive scalar function locally Lipschitz in y subject to certain constraints to be discussed later.

Next, given a positive-definite matrix $W_p \in \mathbb{R}^{n \times n}$, let $P > 0$ be the solution to the Lyapunov equation

$$(A + BF)'P + P(A + BF) = -W_p \quad (5.81)$$

Given another positive-definite matrix $W_R \in \mathbb{R}^{(n-p) \times (n-p)}$ with

$$W_R > F_2' B' P W_p^{-1} P B F_2 \quad (5.82)$$

let $Q_R > 0$ be the solution to the Lyapunov equation

$$(A_{22} + K_R A_{12})' Q_R + Q_R (A_{22} + K_R A_{12}) = -W_R \quad (5.83)$$

Note that such P and Q_R exist as $A + BF$ and $A_{22} + K_R A_{12}$ are asymptotically stable. For any $\delta \in (0, 1)$, let c_δ be the largest positive scalar such that for all

$$\left(\begin{pmatrix} x \\ x_v \end{pmatrix} \in \mathbf{X}_{R\delta} := \left\{ \begin{pmatrix} x \\ x_v \end{pmatrix} : \begin{pmatrix} x \\ x_v \end{pmatrix}' \begin{bmatrix} P & 0 \\ 0 & Q_R \end{bmatrix} \begin{pmatrix} x \\ x_v \end{pmatrix} \leq c_\delta \right\} \right) \quad (5.84)$$

we have

$$\left| [F \quad F_2] \begin{pmatrix} x \\ x_v \end{pmatrix} \right| \leq u_{\max}(1 - \delta) \quad (5.85)$$

We have the following theorem.

Theorem 5.7. Consider the given system in Equation 5.1. Then, there exists a scalar $\rho^* > 0$ such that for any nonpositive function $\rho(r, y)$, locally Lipschitz in y satisfying $|\rho(r, y)| \leq \rho^*$, the reduced-order CNF control law given by Equations 5.79 and 5.80 internally stabilizes the given plant and drives the system controlled output $h(t)$ to track asymptotically the step command input of amplitude r from an initial state x_0 , provided that x_0 , x_{v0} and r satisfy

$$\begin{pmatrix} x_0 - x_e \\ x_{v0} - x_{20} - K_R x_{10} \end{pmatrix} \in \mathbf{X}_{R\delta} \quad \text{and} \quad |Hr| \leq \delta \cdot u_{\max} \quad (5.86)$$

Proof. Let $\tilde{x} = x - x_e$ and $\tilde{x}_v = x_v - x_2 - K_R x_1$. Then, the closed-loop system comprising the given plant in Equation 5.1 and the reduced-order CNF control law of Equations 5.79 and 5.80 can be expressed as

$$\begin{pmatrix} \dot{\tilde{x}} \\ \dot{\tilde{x}}_v \end{pmatrix} = \begin{bmatrix} A + BF & BF_2 \\ 0 & A_{22} + K_R A_{12} \end{bmatrix} \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + \begin{bmatrix} B \\ 0 \end{bmatrix} w \quad (5.87)$$

where

$$w = \text{sat} \left\{ \begin{aligned} & [F \quad F_2] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} + Hr + \rho(r, y) B' P \left[\tilde{x} + \begin{pmatrix} 0 \\ \tilde{x}_v \end{pmatrix} \right] \\ & - [F \quad F_2] \begin{pmatrix} \tilde{x} \\ \tilde{x}_v \end{pmatrix} - Hr \end{aligned} \right\} \quad (5.88)$$

The rest of the proof follows along similar lines to the reasoning given in the full-order measurement feedback case. \diamond

5.2.2 Systems with External Disturbances

We introduce in this subsection an enhanced version of the CNF control design, which is capable of removing constant bias in servo systems. When the given system has disturbances, the resulting system output generally does not asymptotically match the target reference without knowing *a priori* the level of bias. A common approach for removing bias resulting from constant disturbances is to add an integrator to the controller. In what follows, we propose an enhanced CNF design scheme by introducing an additional integration action in the design. The new approach retains the fast rise time property of the original CNF control and at the same time has an additional capacity of eliminating steady-state bias due to disturbances.

Specifically, we consider a linear system with an amplitude constrained actuator, characterized by

$$\begin{cases} \dot{x} = Ax + B \text{sat}(u) + Ew, & x(0) = x_0 \\ y = C_1 x \\ h = C_2 x \end{cases} \quad (5.89)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}$, $y \in \mathbb{R}^p$, $h \in \mathbb{R}$ and $w \in \mathbb{R}$ are, respectively, the state, control input, measurement output, controlled output and disturbance input of the system. A ,

B , C_1 , C_2 and E are appropriate dimensional constant matrices. The function, $\text{sat}: \mathbb{R} \rightarrow \mathbb{R}$, represents the actuator saturation defined as

$$\text{sat}(u) = \text{sgn}(u) \min\{u_{\max}, |u|\} \quad (5.90)$$

with u_{\max} being the saturation level of the input. The assumptions on the given system are made:

1. (A, B) is stabilizable,
2. (A, C_1) is detectable,
3. (A, B, C_2) is invertible and has no invariant zero at $s = 0$,
4. w is bounded unknown constant disturbance, and
5. h is part of y , *i.e.* h is also measurable.

Note that all these assumptions are fairly standard for tracking control. We aim to design an enhanced CNF control law for the system with disturbances such that the resulting controlled output would track a target reference (set point), say r , as fast and as smooth as possible without having steady-state error. We first follow the usual practice to augment an integrator into the given system. Such an integrator eventually becomes part of the final control law. To be more specific, we define an auxiliary state variable,

$$\dot{x}_i := \kappa_i e := \kappa_i (h - r) = \kappa_i C_2 x - \kappa_i r \quad (5.91)$$

which is implementable as h is assumed to be measurable, and κ_i is a positive scalar to be selected to yield an appropriate integration action. The augmented system is then given as follows,

$$\begin{cases} \dot{\bar{x}} = \bar{A} \bar{x} + \bar{B} \text{sat}(u) + \bar{B}_r r + \bar{E} w \\ \bar{y} = \bar{C}_1 \bar{x} \\ h = \bar{C}_2 \bar{x} \end{cases} \quad (5.92)$$

where

$$\bar{x} = \begin{pmatrix} x_i \\ x \end{pmatrix}, \quad \bar{x}_0 = \begin{pmatrix} 0 \\ x_0 \end{pmatrix}, \quad \bar{y} = \begin{pmatrix} x_i \\ y \end{pmatrix} \quad (5.93)$$

$$\bar{A} = \begin{bmatrix} 0 & \kappa_i C_2 \\ 0 & A \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} 0 \\ B \end{bmatrix}, \quad \bar{B}_r = \begin{bmatrix} -\kappa_i \\ 0 \end{bmatrix} \quad (5.94)$$

and

$$\bar{E} = \begin{bmatrix} 0 \\ E \end{bmatrix}, \quad \bar{C}_1 = \begin{bmatrix} 1 & 0 \\ 0 & C_1 \end{bmatrix}, \quad \bar{C}_2 = [0 \quad C_2] \quad (5.95)$$

It is straightforward to show that Assumptions 1 and 3 imply that the pair (\bar{A}, \bar{B}) is stabilizable.

Next, we proceed to carry out the design of enhanced CNF control laws for two different cases, *i.e.* the state feedback case and the reduced-order measurement feedback case. The full-order measurement feedback case is straightforward once the result for the reduced-order case is established. We note that the procedure for designing the enhanced CNF control laws is a natural extension of that given in the previous subsection.

i. State Feedback Case. We first investigate the case when all the state variables of the plant in Equation 5.92 are measurable, *i.e.* $\bar{y} = \bar{x}$. The procedure that generates an enhanced CNF state feedback law is done in three steps. That is, in the first step, a linear feedback control law is designed, in the second step, the design of nonlinear feedback control is carried out, and lastly, in the final step, the linear and nonlinear feedback laws are combined to form an enhanced CNF control law.

STEP 5.C.W.S.1: design a linear feedback control law,

$$u_L = F\bar{x} + Gr \quad (5.96)$$

where F is chosen such that 1) $\bar{A} + \bar{B}F$ is an asymptotically stable matrix, and 2) the closed-loop system $\bar{C}_2(sI - \bar{A} - \bar{B}F)^{-1}\bar{B}$ has certain desired properties. Let us partition $F = [F_1 \quad F_x]$ in conformity with x_1 and x . The general guideline in the design of such an F is to place the closed-loop pole of $\bar{A} + \bar{B}F$ corresponding to the integration mode, x_1 , to be sufficiently closer to the imaginary axis compared to the rest eigenvalues, which implies that F_1 is a relatively small scalar. The remaining closed-loop poles of $\bar{A} + \bar{B}F$ are placed to have a dominating pair with a small damping ratio, which in turn would yield a fast rise time in the closed-loop system response. Finally, G is chosen as

$$G = -[C_2(A + BF_x)^{-1}B]^{-1} \quad (5.97)$$

which is well defined as (A, B, C_2) is assumed to have no invariant zeros at $s = 0$ and $A + BF_x$ is nonsingular whenever $\bar{A} + \bar{B}F$ is stable and F_1 is relatively small.

STEP 5.C.W.S.2: given a positive-definite matrix $W \in \mathbb{R}^{(n+1) \times (n+1)}$, we solve the following Lyapunov equation:

$$(\bar{A} + \bar{B}F)'P + P(\bar{A} + \bar{B}F) = -W \quad (5.98)$$

for $P > 0$. Such a solution is always existent as $\bar{A} + \bar{B}F$ is asymptotically stable. The nonlinear feedback portion of the enhanced CNF control law, u_N , is given by

$$u_N = \rho(e)\bar{B}'P(\bar{x} - \bar{x}_e) \quad (5.99)$$

where $\rho(e)$, with $e = h - r$ being the tracking error, is a smooth and nonpositive function of $|e|$, and tends to a constant as $t \rightarrow \infty$. It is used to gradually change the system closed-loop damping ratio to yield a better tracking performance. The choices of the design parameters, $\rho(e)$ and W , will be discussed later. Next, we define

$$G_e := \begin{bmatrix} 0 \\ -(A + BF_x)^{-1}BG \end{bmatrix} \quad \text{and} \quad \bar{x}_e := G_e r \quad (5.100)$$

STEP 5.C.W.S.3: the linear feedback control law and nonlinear feedback portion derived in the previous steps are now combined to form an enhanced CNF control law,

$$u = u_L + u_N = F\bar{x} + Gr + \rho(e)\bar{B}'P(\bar{x} - \bar{x}_e) \quad (5.101)$$

We have the following result.

Theorem 5.8. Consider the given system in Equation 5.89 with $y = x$ and the disturbance w being bounded by a non-negative scalar τ_w , i.e. $|w| \leq \tau_w$. Let

$$\gamma := 2\tau_w\lambda_{\max}(PW^{-1}) (\bar{E}'P\bar{E})^{1/2} \quad (5.102)$$

Then, for any $\rho(e)$, which is a smooth and nonpositive function of $|e|$ and tends to a constant as $t \rightarrow \infty$, the enhanced CNF control law in Equation 5.101 internally stabilizes the given plant and drives the system controlled output h to track the step reference of amplitude r from an initial state \bar{x}_0 asymptotically without steady-state error, provided that the following conditions are satisfied:

1. There exist positive scalars $\delta \in (0, 1)$ and $c_\delta > \gamma^2$ such that

$$\forall \bar{x} \in \mathbf{X}(F, c_\delta) := \{ \bar{x} : \bar{x}'P\bar{x} \leq c_\delta \} \Rightarrow |F\bar{x}| \leq (1 - \delta)u_{\max} \quad (5.103)$$

2. The initial condition, \bar{x}_0 , satisfies

$$\bar{x}_0 - \bar{x}_e \in \mathbf{X}(F, c_\delta) \quad (5.104)$$

3. The level of the target reference, r , satisfies

$$|Hr| \leq \delta u_{\max} \quad (5.105)$$

where $H := FG_e + G$.

Proof. For simplicity, we drop the variable e in $\rho(e)$ throughout this proof. Noting that

$$\begin{aligned} \bar{A}\bar{x}_e + \bar{B}Hr + \bar{B}_r r &= \bar{A}\bar{x}_e + \bar{B}(FG_e + G)r + \bar{B}_r r \\ &= (\bar{A} + \bar{B}F)\bar{x}_e + \bar{B}_r r + \bar{B}Gr \\ &= \begin{bmatrix} 0 & \kappa_i C_2 \\ BF_i & A + BF_x \end{bmatrix} \begin{bmatrix} 0 \\ -(A + BF_x)^{-1}BG \end{bmatrix} r + \begin{bmatrix} -\kappa_i \\ 0 \end{bmatrix} r + \begin{bmatrix} 0 \\ BG \end{bmatrix} r \\ &= \begin{bmatrix} -\kappa_i C_2 (A + BF_x)^{-1}BG \\ -BG \end{bmatrix} r + \begin{bmatrix} -\kappa_i \\ BG \end{bmatrix} r = 0 \end{aligned} \quad (5.106)$$

and letting $\tilde{x} = \bar{x} - \bar{x}_e$, then the dynamics equation of the augmented plant in Equation 5.92 can be expressed as,

$$\begin{aligned} \dot{\tilde{x}} &= \dot{\bar{x}} = \bar{A}(\tilde{x} + \bar{x}_e) + \bar{B} \text{sat}(u) + \bar{B}_r r + \bar{E}w \\ &= (\bar{A} + \bar{B}F)\tilde{x} + \bar{A}\bar{x}_e - \bar{B}F\tilde{x} + \bar{B} \text{sat}(u) + \bar{B}_r r + \bar{E}w \\ &= (\bar{A} + \bar{B}F)\tilde{x} + \bar{B}[\text{sat}(u) - F\tilde{x} - Hr] + (\bar{A}\bar{x}_e + \bar{B}Hr + \bar{B}_r r) + \bar{E}w \\ &= (\bar{A} + \bar{B}F)\tilde{x} + \bar{B}v + \bar{E}w \end{aligned} \quad (5.107)$$

where $v := \text{sat}(u) - F\tilde{x} - Hr$. Following the similar lines of reasoning as those in the previous subsection, we can show that for $\tilde{x} \in \mathbf{X}(F, c_\delta)$ and $|Hr| \leq \delta u_{\max}$,

v can be expressed as $v = q\rho\bar{B}'P\tilde{x}$, for some non-negative variable $q \in [0, 1]$. Thus, for the case when $\tilde{x} \in \mathbf{X}(F, c_\delta)$ and $|Hr| \leq \delta u_{\max}$, the closed-loop system comprising the augmented plant in Equation 5.92 and the enhanced CNF control law in Equation 5.101 can be expressed as follows:

$$\dot{\tilde{x}} = (\bar{A} + \bar{B}F + q\rho\bar{B}\bar{B}'P)\tilde{x} + \bar{E}w \quad (5.108)$$

In what follows, we show that the system in Equation 5.108 is stable provided that the initial condition, \tilde{x}_0 , the target reference, r , and the disturbance, w , satisfy those conditions listed in the theorem. Let us define a Lyapunov function

$$V = \tilde{x}'P\tilde{x} \quad (5.109)$$

For easy derivation, we introduce a matrix S such that $P = S'S$. We then obtain the derivative of V calculated along the trajectory of the system in Equation 5.108,

$$\begin{aligned} \dot{V} &= \dot{\tilde{x}}'P\tilde{x} + \tilde{x}'P\dot{\tilde{x}} \\ &= \tilde{x}'[(\bar{A} + \bar{B}F)'P + P(\bar{A} + \bar{B}F)]\tilde{x} + 2q\rho\tilde{x}'P\bar{B}\bar{B}'P\tilde{x} + 2\tilde{x}'P\bar{E}w \\ &\leq -\tilde{x}'W\tilde{x} + 2\tilde{x}'P\bar{E}w \\ &= -\tilde{x}'S'SP^{-1}WP^{-1}S'\tilde{x} + 2\tilde{x}'S'S\bar{E}w \\ &\leq -\lambda_{\min}(SP^{-1}WP^{-1}S')\tilde{x}'S'S\tilde{x} + 2\|S\tilde{x}\| \cdot \|S\bar{E}\|\tau_w \\ &= -\lambda_{\min}(P^{-1}W)\tilde{x}'P\tilde{x} + 2\tau_w(\tilde{x}'P\tilde{x})^{1/2}(\bar{E}'P\bar{E})^{1/2} \\ &= -\lambda_{\min}(P^{-1}W)(\tilde{x}'P\tilde{x})^{1/2} \left[(\tilde{x}'P\tilde{x})^{1/2} - 2\tau_w\lambda_{\max}(PW^{-1})(\bar{E}'P\bar{E})^{1/2} \right] \\ &= -\lambda_{\min}(P^{-1}W)(\tilde{x}'P\tilde{x})^{1/2} \left[(\tilde{x}'P\tilde{x})^{1/2} - \gamma \right] \end{aligned} \quad (5.110)$$

We note that we have used the following matrix properties: i) $z'Xz \geq \lambda_{\min}(X)z'z$, where X is a symmetric matrix; ii) $\lambda(XY) = \lambda(YX)$, if both X and Y are square matrices; and iii) $\lambda(XY) > 0$ if $X > 0$ and $Y > 0$. Clearly, the closed-loop system in the absence of the disturbance, w , has $\dot{V} < 0$ and thus is asymptotically stable.

With the presence of the unknown constant disturbance, w , and with the initial condition $\tilde{x}(0) = \tilde{x}_0 - \tilde{x}_e \in \mathbf{X}(F, c_\delta)$, where $c_\delta > \gamma^2$, the corresponding trajectory of Equation 5.108 remains in $\mathbf{X}(F, c_\delta)$ and converges to a point on a ball $\{\tilde{x} : \tilde{x}'P\tilde{x} \leq \tilde{\gamma}^2\}$ with $\tilde{\gamma} \leq \gamma$. Since $x_i(t) = \int_0^t \kappa_i e(\tau) d\tau$ converges to a constant, it is clear that the tracking error $e(t) \rightarrow 0$ as $t \rightarrow \infty$. This completes the proof of Theorem 5.8. \diamond

ii. Measurement Feedback Case. Next, we proceed to design an enhanced CNF control law using only information measurable from the plant. In principle, we can design either a full-order measurement feedback control law, for which its dynamical order is identical to that of the given plant, or a reduced-order measurement feedback control law, in which we make a full use of the measurement output and estimate only the unknown part of the state variable. As such, the dynamical order of the controller is reduced. It is more feasible to implement controllers with smaller dynamical order. The procedure below on the enhanced CNF control using

reduced-order measurement feedback follows closely from that given in the previous subsection.

For simplicity of presentation, we assume that C_1 in the measurement output of the given plant in Equation 5.89 is already in the form,

$$C_1 = [I_p \quad 0] \quad (5.111)$$

The augmented plant in Equation 5.92 can then be partitioned as follows:

$$\left\{ \begin{array}{l} \begin{pmatrix} \dot{x}_i \\ \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{bmatrix} 0 & \kappa_i C_{21} & \kappa_i C_{22} \\ 0 & A_{11} & A_{12} \\ 0 & A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x_i \\ x_1 \\ x_2 \end{pmatrix} + \begin{bmatrix} 0 \\ B_1 \\ B_2 \end{bmatrix} \text{sat}(u) + \begin{bmatrix} -\kappa_i \\ 0 \\ 0 \end{bmatrix} r + \begin{bmatrix} 0 \\ E_1 \\ E_2 \end{bmatrix} w \\ \bar{y} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & I_p & 0 \end{bmatrix} \begin{pmatrix} x_i \\ x_1 \\ x_2 \end{pmatrix} \\ h = [0 \quad C_{21} \quad C_{22}] \begin{pmatrix} x_i \\ x_1 \\ x_2 \end{pmatrix} \end{array} \right. \quad (5.112)$$

where

$$\begin{pmatrix} x_i \\ x_1 \\ x_2 \end{pmatrix} = \bar{x}, \quad \begin{pmatrix} x_i(0) \\ x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} 0 \\ x_{10} \\ x_{20} \end{pmatrix} = \bar{x}_0 \quad (5.113)$$

and

$$\bar{y} = \begin{pmatrix} x_i \\ y \end{pmatrix} = \begin{pmatrix} x_i \\ x_1 \end{pmatrix} \quad (5.114)$$

Clearly, x_i and x_1 are readily available and need not be estimated. We only need to estimate x_2 . There are two main step in designing a reduced-order measurement feedback control laws: i) the construction of a full state feedback gain matrix F ; and ii) the construction of a reduced-order observer gain matrix K_R . The construction of the gain matrix F is totally identical to that given in the previous subsection, which can be partitioned in conformity with x_i , x_1 and x_2 , as follows:

$$F = [F_i \quad F_1 \quad F_2] \quad (5.115)$$

The reduced-order observer gain matrix K_R is chosen such that the closed-loop poles of $A_{22} + K_R A_{12}$ are placed in appropriate locations in the open-left half plane.

The reduced-order enhanced CNF control law is then given by,

$$\begin{aligned} \dot{x}_v = (A_{22} + K_R A_{12})x_v + (B_2 + K_R B_1) \text{sat}(u) \\ + [A_{21} + K_R A_{11} - (A_{22} + K_R A_{12})K_R]y \end{aligned} \quad (5.116)$$

and

$$u = F \begin{pmatrix} x_i \\ x_1 \\ x_v - K_R y \end{pmatrix} + Gr + \rho(e) \bar{B}' P \left[\begin{pmatrix} x_i \\ x_1 \\ x_v - K_R y \end{pmatrix} - \bar{x}_e \right] \quad (5.117)$$

where G is as defined in Equation 5.97 and $\rho(e)$ is the smooth, nonpositive and nondecreasing function of $|e|$, to be chosen to yield a desired performance.

Next, given a positive-definite matrix $W \in \mathbb{R}^{(n+1) \times (n+1)}$, let $P > 0$ be the solution to the Lyapunov equation

$$(\bar{A} + \bar{B}F)'P + P(\bar{A} + \bar{B}F) = -W \quad (5.118)$$

Given another positive-definite matrix $W_R \in \mathbb{R}^{(n-p) \times (n-p)}$ with

$$W_R > F_2' \bar{B}' P W^{-1} P \bar{B} F_2 \quad (5.119)$$

let $Q_R > 0$ be the solution to the Lyapunov equation

$$(A_{22} + K_R A_{12})' Q_R + Q_R (A_{22} + K_R A_{12}) = -W_R \quad (5.120)$$

Note that such P and Q_R exist as $\bar{A} + \bar{B}F$ and $A_{22} + K_R A_{12}$ are both asymptotically stable. We have the following result.

Theorem 5.9. *Consider the given system in Equation 5.89 with w being bounded by a scalar $\tau_w \geq 0$, i.e. $|w| \leq \tau_w$. Let*

$$\begin{aligned} \gamma_R := & 2\tau_w \lambda_{\max} \left(\begin{bmatrix} P & 0 \\ 0 & Q_R \end{bmatrix} \begin{bmatrix} W & -P \bar{B} F_2 \\ -F_2' \bar{B}' P & W_R \end{bmatrix}^{-1} \right) \\ & \times [\bar{E}' P \bar{E} + (E_2 + K_R E_1)' Q_R (E_2 + K_R E_1)]^{1/2} \end{aligned} \quad (5.121)$$

Then, there exists a scalar $\rho^* > 0$ such that for any $\rho(e)$, a smooth and nonpositive function of $|e|$ with $|\rho(e)| \leq \rho^*$ and tending to a constant as $t \rightarrow \infty$, the enhanced reduced-order CNF control law of Equations 5.116 and 5.117 internally stabilizes the given plant and drives the system controlled output h to track the step reference of amplitude r asymptotically without steady-state error, provided that the following conditions are satisfied:

1. There exist positive scalars $\delta \in (0, 1)$ and $c_{R\delta} > \gamma_R^2$ such that

$$\begin{aligned} \forall \bar{x} \in \mathbf{X}(F, c_{R\delta}) := & \left\{ \bar{x} : \bar{x}' \begin{bmatrix} P & 0 \\ 0 & Q_R \end{bmatrix} \bar{x} \leq c_{R\delta} \right\} \\ \Rightarrow & |[F \ F_2] \bar{x}| \leq (1 - \delta) u_{\max} \end{aligned} \quad (5.122)$$

2. The initial conditions, \bar{x}_0 and $x_{v0} = x_v(0)$, satisfy

$$\begin{pmatrix} \bar{x}_0 - \bar{x}_e \\ x_{v0} - x_{20} - K_R x_{10} \end{pmatrix} \in \mathbf{X}(F, c_{R\delta}) \quad (5.123)$$

3. The level of the target reference, r , satisfies

$$|Hr| \leq \delta u_{\max} \quad (5.124)$$

where H is the same as that defined in Theorem 5.8.

Proof. The result follows from similar lines of reasoning as those in Theorem 5.8 and those for the measurement feedback case in the previous subsection. \diamond

5.2.3 Selection of Nonlinear Feedback Parameters

Basically, the freedom to choose the function ρ in either the usual CNF design or the enhanced CNF design is used to tune the control laws so as to improve the performance of the closed-loop system as the controlled output, h , approaches the set point, r . Since the main purpose of adding the nonlinear part to the CNF or the enhanced CNF controllers is to shorten the settling time, or equivalently to contribute a significant value to the control input when the tracking error, e , is small. The nonlinear part, in general, is set in action when the control signal is far away from its saturation level, and thus it does not cause the control input to hit its limits. For simplicity, we now focus our attention on the case when the given system has external disturbances. The following analysis is equally applicable to the case when the given system does not have disturbances. Under such circumstances, the closed-loop system comprising the augmented plant in Equation 5.92 and the enhanced CNF control law can be expressed as:

$$\dot{\tilde{x}} = (\bar{A} + \bar{B}F + \rho\bar{B}\bar{B}'P)\tilde{x} + \bar{E}w \quad (5.125)$$

We note that the additional term ρ does not affect the stability of the estimators. It is now clear that eigenvalues of the closed-loop system in Equation 5.125 can be changed by the function ρ . Such a mechanism can be interpreted using the classical feedback control concept as shown in Figure 5.1, where the auxiliary system $G_{\text{aux}}(s)$ is defined as:

$$G_{\text{aux}}(s) := C_{\text{aux}}(sI - A_{\text{aux}})^{-1}B_{\text{aux}} := \bar{B}'P(sI - \bar{A} - \bar{B}F)^{-1}\bar{B} \quad (5.126)$$

$G_{\text{aux}}(s)$ has the following interesting properties.

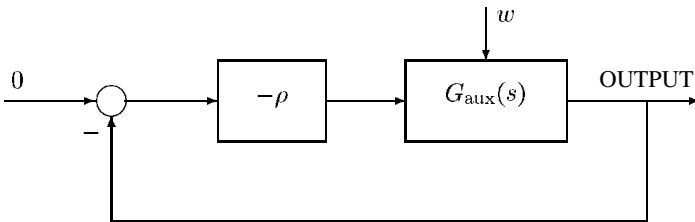


Figure 5.1. Interpretation of the nonlinear function $\rho(e)$

Theorem 5.10. *The auxiliary system $G_{\text{aux}}(s)$ defined in Equation 5.126 is stable and invertible with a relative degree equal to 1, and is of minimum phase with n stable invariant zeros.* \diamond

Proof. First, it is obvious to see that $G_{\text{aux}}(s)$ is stable since $\bar{A} + \bar{B}F$ is a stable matrix. Next, since $P > 0$ and $\bar{B} \neq 0$, we have

$$C_{\text{aux}}B_{\text{aux}} = \bar{B}'P\bar{B} > 0 \quad (5.127)$$

which implies that $G_{\text{aux}}(s)$ is invertible and has a relative degree equal to 1 (or an infinite zero of order 1). Furthermore, $G_{\text{aux}}(s)$ has n invariant zeros, as it is a SISO system.

The last property of $G_{\text{aux}}(s)$, *i.e.* the invariant zeros of $G_{\text{aux}}(s)$ are stable and hence it is of minimum phase, can be shown by using the well-known classical root-locus theory. Observing the block diagram in Figure 5.1, it follows from the classical feedback control theory (see, *e.g.*, [1]) that the poles of the closed-loop system of Equation 5.125, which are the functions of the tuning parameter ρ , start from the open-loop poles, *i.e.* the eigenvalues of $\bar{A} + \bar{B}F$, when $\rho = 0$, and end up at the open-loop zeros (including the zero at the infinity) as $|\rho| \rightarrow \infty$. It then follows from the proof of Theorem 5.3 that the closed-loop system remains asymptotically stable for any nonpositive ρ , which implies that all the invariant zeros of the open-loop system, *i.e.* $G_{\text{aux}}(s)$, must be stable. \diamond

It is clear from Theorem 5.10 and its proof that the invariant zeros of $G_{\text{aux}}(s)$ play an important role in selecting the poles of the closed-loop system of Equation 5.125. The poles of the closed-loop system approach the locations of the invariant zeros of $G_{\text{aux}}(s)$ as $|\rho|$ becomes larger and larger. We would like to note that there is freedom in preselecting the locations of these invariant zeros. This can actually be done by selecting an appropriate $W > 0$ in Equation 5.98. In general, we should select the invariant zeros of $G_{\text{aux}}(s)$, which are corresponding to the closed-loop poles for larger $|\rho|$, such that the dominated ones have a large damping ratio, which in turn yields a smaller overshoot. The following procedure can be used as a guideline for the selection of such a W :

1. Given the pair $(A_{\text{aux}}, B_{\text{aux}})$ and the desired locations of the invariant zeros of $G_{\text{aux}}(s)$, we follow the result of Chen and Zheng [139] (see also Chapter 9 of Chen *et al.* [71]) on finite and infinite zero assignment to obtain an appropriate matrix C_{aux} such that the resulting matrix triple $(A_{\text{aux}}, B_{\text{aux}}, C_{\text{aux}})$ has the desired relative degree and invariant zeros.
2. Solve $C_{\text{aux}} = \bar{B}'P$ for a $P = P' > 0$. In general, the solution is nonunique as there are $n(n+1)/2$ elements in P available for selection. However, if the solution does not exist, we go back to the previous step to reselect the invariant zeros.
3. Calculate W using Equation 5.98 and check if W is positive-definite. If W is not positive-definite, we go back to the previous step to choose another solution of P or go to the first step to reselect the invariant zeros.

Generally, the above procedure would yield a desired result. The selection of the nonlinear function ρ is relatively simple once the desired invariant zeros of $G_{\text{aux}}(s)$ are obtained. Assuming the tracking error e is available, the following choice of $\rho(e)$ is a smooth and nonpositive function of $|e|$:

$$\rho(e) = -\beta \left| e^{-\alpha|e|} - e^{-\alpha|h(0)-r|} \right| \quad (5.128)$$

where α and β are appropriate positive scalars that can be chosen to yield a desired performance, *i.e.* fast settling time and small overshoot. This function $\rho(e)$ changes from 0 to $\rho_0 = -\beta |1 - e^{-\alpha|h(0)-r}|$ as the tracking error approaches zero. At the initial stage, when the controlled output, h , is far away from the final set point, $\rho(e)$ is small and the effect of the nonlinear part on the overall system is very limited. When the controlled output, h , approaches the set point, $\rho(e) \approx \rho_0$, and the nonlinear control law becomes effective. In general, the parameter ρ_0 is chosen such that the poles of $\bar{A} + \bar{B}F + \rho_0 \bar{B} \bar{B}' P$ are in the desired locations, *e.g.*, the dominated poles have a large damping ratio, which would reduce the overshoot of the output response. Note that the choice of $\rho(e)$ is nonunique. Any function would work so long as it has similar properties of that given in Equation 5.128.

5.2.4 An Illustrative Example

We illustrate the enhanced CNF control technique for continuous-time systems in the following example. We consider a continuous-time system of Equation 5.89 with

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad E = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad x_0 = \begin{pmatrix} 0 \\ 0.5 \\ 0 \end{pmatrix} \quad (5.129)$$

$$C_1 = I_3, \quad C_2 = [1 \ 0 \ 0] \quad (5.130)$$

and $u_{\max} = 2$. The disturbance w is unknown. For simulation purpose, we assume $w = -0.1$. Our goal is to design an enhanced CNF state feedback control law that would yield a good transient performance in tracking a target reference $r = 1$.

Following the procedure given in the previous subsection, we select an integration gain $\kappa_i = 1$ and obtain an appropriate augmented system. After a few tries, we found that the following state feedback gain to the augmented system would yield a good performance for our problem:

$$F = [-0.6 \quad -2.8 \quad -3.8 \quad -3.2] \quad (5.131)$$

which places the poles of $\bar{A} + \bar{B}F$ at $-0.2, -3, -0.5 \pm j0.866$. We note that the first one corresponds to the integrator. Both the linear state feedback control and enhanced CNF control share the same integration dynamics:

$$\dot{x}_i = h - r \quad (5.132)$$

The linear state feedback control law is given by

$$u = [-2.8 \quad -3.8 \quad -3.2]x - 0.6x_i + 3.8r \quad (5.133)$$

Letting $W = \text{diag}\{0.04, 0.1, 5, 0.1\}$, we obtain a positive-definite solution P for Equation 5.98, which is given by

$$P = \begin{bmatrix} 0.2403 & 0.6694 & 0.2684 & 0.0333 \\ 0.6694 & 3.8665 & 1.5752 & 0.1893 \\ 0.2684 & 1.5752 & 4.4039 & 0.8490 \\ 0.0333 & 0.1893 & 0.8490 & 0.2140 \end{bmatrix} \quad (5.134)$$

and an enhanced CNF state feedback law:

$$u = [-2.8 \quad -3.8 \quad -3.2]x - 0.6x_1 + 3.8r + \rho(e) \\ \times [0.0333 \quad 0.1893 \quad 0.8490 \quad 0.2140] \left[\begin{pmatrix} x_1 \\ x \end{pmatrix} - \begin{pmatrix} 0 \\ 2 \\ 0 \\ 0 \end{pmatrix} \right] \quad (5.135)$$

where $\rho(e)$ is as given in (5.128) with $\alpha = 1$ and $\beta = 8$. The simulation results given in Figures 5.2 and 5.3 clearly show that the CNF control has outperformed the linear control.

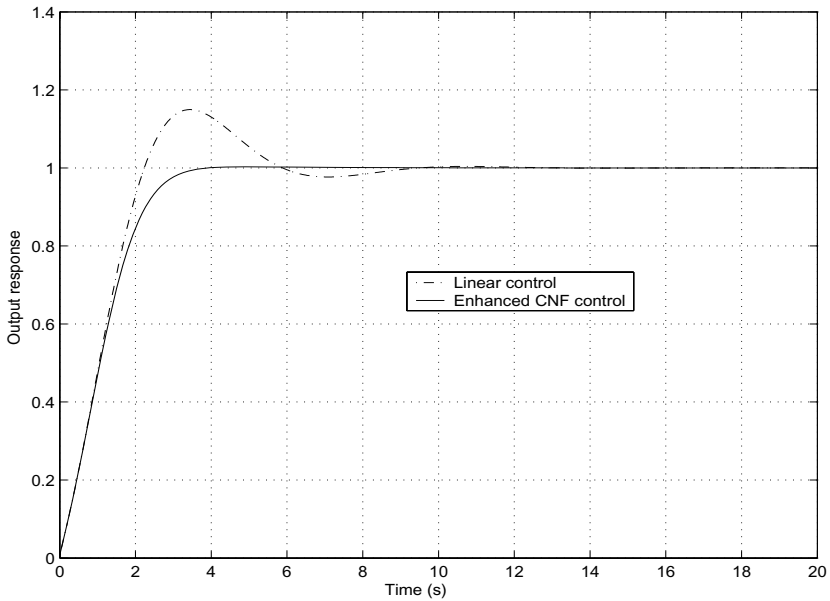


Figure 5.2. Output responses of the enhanced CNF control and linear control

5.3 Discrete-time Systems

As in the continuous-time case, we present in this section the CNF design technique for systems without and with external disturbances. Selection and interpretation of nonlinear gain design parameters are also discussed.

5.3.1 Systems without External Disturbances

Let us now consider a linear discrete-time system Σ with an amplitude-constrained actuator characterized by

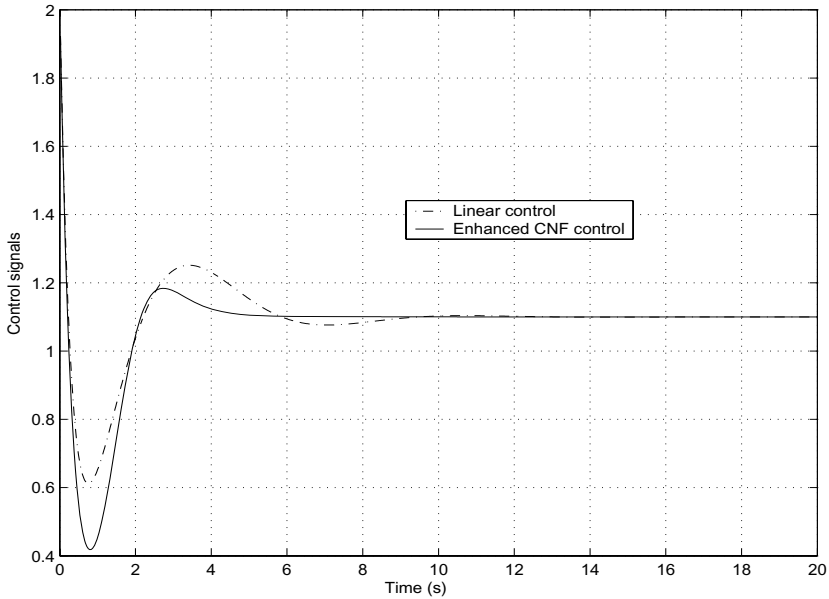


Figure 5.3. Control signals of the enhanced CNF control and linear control

$$\begin{cases} x(k+1) = A x(k) + B \text{sat}(u(k)), & x(0) = x_0 \\ y(k) = C_1 x(k) \\ h(k) = C_2 x(k) \end{cases} \quad (5.136)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}$, $y \in \mathbb{R}^p$ and $h \in \mathbb{R}$ are, respectively, the state, control input, measurement output and controlled output of Σ . A , B , C_1 and C_2 are appropriate dimensional constant matrices, and $\text{sat}: \mathbb{R} \rightarrow \mathbb{R}$ represents the actuator saturation defined as

$$\text{sat}(u) = \text{sgn}(u) \min\{u_{\max}, |u|\} \quad (5.137)$$

with u_{\max} being the saturation level of the input. The following assumptions on the system matrices are required:

1. (A, B) is stabilizable,
2. (A, C_1) is detectable, and
3. (A, B, C_2) is invertible and has no invariant zeros at $z = 1$.

We now extend the results of the continuous-time composite nonlinear control method to the discrete-time system in Equation 5.136. Thus, the objective here is to design a discrete-time CNF control law that causes the output to track a high-amplitude step input rapidly without experiencing large overshoot and without the adverse actuator saturation effects. This can be done through the design of a discrete-time linear feedback law with a small closed-loop damping ratio and a nonlinear feedback law through an appropriate Lyapunov function to cause the closed-loop

system to be highly damped as system output approaches the command input to reduce the overshoot. The result of this discrete-time version is analogous to that of its continuous-time counterpart. Here, we again separate the design of discrete-time CNF control into three distinct situations, *i.e.* 1) the state feedback case, 2) the full-order measurement case, and 3) the reduced-order measurement feedback case.

i. State Feedback Case. We consider the case when $y = x$, *i.e.* all the state variables of Σ of Equation 5.136 are available for feedback.

STEP 5.D.S.1: design a linear feedback law,

$$u_l(k) = Fx(k) + Gr \quad (5.138)$$

where r is the input command, and F is chosen such that $A + BF$ has all its eigenvalues in \mathbb{C}^\ominus and the closed-loop system $C_2(zI - A - BF)^{-1}B$ meets certain design specifications. We note again that such an F can be designed using any of the techniques reported in Chapter 3. Furthermore,

$$G = [C_2(I - A - BF)^{-1}B]^{-1} \quad (5.139)$$

We note that G is well defined because $A + BF$ has all its eigenvalues in \mathbb{C}^\ominus , and (A, B, C_2) is invertible and has no invariant zeros at $z = 1$.

The following lemma determines the magnitude of r that can be tracked by such a control law without exceeding the control limits.

Lemma 5.11. *Given a positive-definite matrix $W \in \mathbb{R}^{n \times n}$, let $P > 0$ be the solution of the following Lyapunov equation:*

$$P = (A + BF)'P(A + BF) + W. \quad (5.140)$$

Such a P exists as $A + BF$ is asymptotically stable. For any $\delta \in (0, 1)$, let $c_\delta > 0$ be the largest positive scalar such that

$$|Fx(k)| \leq u_{\max}(1 - \delta), \quad \forall x(k) \in \mathbf{X}_\delta := \left\{ x : x'Px \leq c_\delta \right\} \quad (5.141)$$

Also, let

$$H := [1 + F(I - A - BF)^{-1}B]G \quad (5.142)$$

and

$$x_e := G_e r := (I - A - BF)^{-1}BG r \quad (5.143)$$

Then, the control law in Equation 5.138 is capable of driving the system controlled output $h(k)$ to track asymptotically a step command input of amplitude r , provided that the initial state x_0 and r satisfy:

$$(x_0 - x_e) \in \mathbf{X}_\delta \quad \text{and} \quad |Hr| \leq \delta \cdot u_{\max} \quad (5.144)$$

Proof. Let $\tilde{x} = x - x_e$. Then, the linear feedback control law u_L can be rewritten as

$$u_L(k) = F\tilde{x}(k) + [1 + F(I - A - BF)^{-1}B]Gr = F\tilde{x}(k) + Hr \quad (5.145)$$

Hence, for all

$$\tilde{x}(k) \in \mathbf{X}_\delta \Rightarrow |F\tilde{x}(k)| \leq u_{\max}(1 - \delta) \quad (5.146)$$

and for any r satisfying

$$|Hr| \leq \delta \cdot u_{\max} \quad (5.147)$$

the linear control law can be written as

$$|u_L(k)| = |F\tilde{x}(k) + Hr| \leq |F\tilde{x}(k)| + |Hr| \leq u_{\max} \quad (5.148)$$

which indicates that the control signal $u_L(k)$ never exceeds the saturation. Next, let us move to verify the asymptotic stability of the closed-loop system comprising the given plant in Equation 5.136 and the linear feedback law in Equation 5.138, which can be expressed as follows:

$$\tilde{x}(k+1) = (A + BF)\tilde{x}(k) \quad (5.149)$$

Let us define a Lyapunov function for the closed-loop system in Equation 5.149 as

$$V(k) = \tilde{x}'(k)P\tilde{x}(k) \quad (5.150)$$

Along the trajectories of the closed-loop system in Equation 5.149 the increment of the Lyapunov function in Equation 5.150 is given by

$$\begin{aligned} \Delta V(k) &= \tilde{x}'(k+1)P\tilde{x}(k+1) - \tilde{x}'(k)P\tilde{x}(k) \\ &= \tilde{x}'(k)(A + BF)'P(A + BF)\tilde{x}(k) - \tilde{x}'(k)P\tilde{x}(k) \\ &= -\tilde{x}'(k)W\tilde{x}(k) \\ &\leq 0 \end{aligned} \quad (5.151)$$

This shows that \mathbf{X}_δ is an invariant set of the the closed-loop system in Equation 5.149 and all trajectories of Equation 5.149 starting from \mathbf{X}_δ converge to the origin. Thus, for any initial state x_0 and the step command input r that satisfy Equation 5.144, we have

$$\lim_{k \rightarrow \infty} x(k) = x_e \quad (5.152)$$

and hence

$$\lim_{k \rightarrow \infty} h(k) = \lim_{k \rightarrow \infty} C_2x(k) = C_2x_e = r \quad (5.153)$$

This completes the proof of Lemma 5.11. \diamond

Remark 5.12. We would like to note that, for the case when $x_0 = 0$, any step command of amplitude r can be tracked asymptotically provided that

$$|r| \leq [c_\delta(G_e'PG_e)^{-1}]^{1/2} \quad \text{and} \quad |Hr| \leq \delta \cdot u_{\max} \quad (5.154)$$

This input command amplitude can be increased by increasing δ and/or decreasing $G_e'PG_e$ through the choice of W . However, the change in F affects the damping ratio of the closed-loop system and hence its rise time. \diamond

STEP 5.D.S.2: the nonlinear feedback control law $u_N(k)$ is given by

$$u_N(k) = \rho(r, y)B'P(A + BF)\left[x(k) - x_e\right] \quad (5.155)$$

where $\rho(r, y)$ is a nonpositive scalar function, locally Lipschitz in y , and is to be used to change the system closed-loop damping ratio as the output approaches the step command input. The choice of $\rho(r, y)$ will be discussed later in detail.

STEP 5.D.S.3: the linear and nonlinear components derived above are now combined to form a discrete-time CNF control law:

$$u(k) = u_L(k) + u_N(k) \quad (5.156)$$

We have the following result.

Theorem 5.13. *Consider the discrete-time system in Equation 5.136. Then, for any nonpositive $\rho(r, y)$, locally Lipschitz in y and $|\rho(r, y)| \leq \rho^* := 2(B'PB)^{-1}$, the CNF control law in Equation 5.156 is capable of stabilizing the given plant and driving the system controlled output $h(k)$ to track the step command input of amplitude r from an initial state x_0 , provided that x_0 and r satisfy the properties in Equation 5.144. \diamond*

Proof. Let $\tilde{x} = x - x_e$. Then, the closed-loop system can be written as

$$\tilde{x}(k+1) = (A + BF)\tilde{x}(k) + Bw(k) \quad (5.157)$$

where

$$w(k) = \text{sat}\left[F\tilde{x}(k) + Hr + u_N(k)\right] - F\tilde{x}(k) - Hr \quad (5.158)$$

Equation 5.144 implies that $\tilde{x}_0 \in \mathbf{X}_\delta$. Define a Lyapunov function

$$V(k) = \tilde{x}'(k)P\tilde{x}(k) \quad (5.159)$$

Noting that

$$\tilde{x}(k) \in \mathbf{X}_\delta \Rightarrow |F\tilde{x}(k)| \leq u_{\max}(1 - \delta) \quad (5.160)$$

we can evaluate the increment of $V(k)$ along the trajectories of the closed-loop system in Equation 5.157 as follows:

$$\begin{aligned} \Delta V(k) &= \tilde{x}'(k+1)P\tilde{x}(k+1) - \tilde{x}'(k)P\tilde{x}(k) \\ &= \tilde{x}'(k)(A + BF)'P(A + BF)\tilde{x}(k) - \tilde{x}'(k)P\tilde{x}(k) + w'(k)B'PBw(k) \\ &\quad + 2\tilde{x}'(k)(A + BF)'PBw(k) \\ &= -\tilde{x}'(k)W\tilde{x}(k) + 2\tilde{x}'(k)(A + BF)'PBw(k) + w'(k)B'PBw(k) \end{aligned} \quad (5.161)$$

Next, we proceed to find the increment of $V(k)$ for three different cases, as is done in continuous-time systems.

If $|F\tilde{x}(k) + Hr + u_N(k)| \leq u_{\max}$, then

$$w(k) = u_N(k) = \rho B'P(A + BF)\tilde{x}(k) \quad (5.162)$$

Thus,

$$\begin{aligned} \Delta V(k) &= -\tilde{x}'(k)W\tilde{x}(k) + 2\tilde{x}'(k)(A + BF)'PBw(k) + w'(k)B'PBw(k) \\ &= -\tilde{x}'(k)W\tilde{x}(k) \\ &\quad + \rho\tilde{x}'(k)(A + BF)'PB(2 + \rho B'PB)B'P(A + BF)\tilde{x}(k) \end{aligned} \quad (5.163)$$

For any nonpositive $\rho(r, y)$ with $|\rho(r, y)| \leq \rho^*$, it is clear that the increment $\Delta V(k) \leq -\tilde{x}'(k)W\tilde{x}(k) \leq 0$

If $F\tilde{x}(k) + Hr + u_N(k) > u_{\max}$, then $|F\tilde{x}(k) + Hr| \leq u_{\max}$ implies that $0 < w(k) < u_N(k)$ and $\rho(r, y) < 0$. Hence,

$$\begin{aligned} \Delta V(k) &= -\tilde{x}'(k)W\tilde{x}(k) + 2\tilde{x}'(k)(A + BF)'PBw(k) + w'(k)B'PBw(k) \\ &= -\tilde{x}'(k)W\tilde{x}(k) + w'(k)[2\rho^{-1}u_N(k) + B'PBw(k)] \\ &< -\tilde{x}'(k)W\tilde{x}(k) + w'(k)[2\rho^{-1}u_N(k) + B'PBu_N(k)] \\ &= -\tilde{x}'(k)W\tilde{x}(k) + w'(k)(2\rho^{-1} + B'PB)u_N(k) \end{aligned} \quad (5.164)$$

Thus, for all $-\rho^* \leq \rho(r, y) < 0$, we have $2\rho^{-1} + B'PB \leq 0$, and hence

$$\Delta V(k) \leq -\tilde{x}'(k)W\tilde{x}(k) \leq 0 \quad (5.165)$$

Similarly, for the case when $F\tilde{x}(k) + Hr + u_N(k) < -u_{\max}$, it can be shown that $\Delta V(k) \leq -\tilde{x}'(k)W\tilde{x}(k) \leq 0$.

Thus, \mathbf{X}_δ is an invariant set of the closed-loop system in Equation 5.157 and all trajectories of Equation 5.157 starting from \mathbf{X}_δ remain there and converge to the origin. This, in turn, indicates that, for all initial states x_0 and the step command input of amplitude r that satisfy Equation 5.144,

$$\lim_{k \rightarrow \infty} x(k) = x_e \quad (5.166)$$

and

$$\lim_{k \rightarrow \infty} y(k) = \lim_{k \rightarrow \infty} C_2x(k) = C_2x_e = r \quad (5.167)$$

This completes the proof of Theorem 5.13. \diamond

Remark 5.14. Theorem 5.13 shows that the addition of the nonlinear feedback control law u_N as given in Equation 5.155 does not affect the ability to track the class of command inputs. Any command input that can be tracked by the linear feedback law in Equation 5.138 can also be tracked by the CNF control law in Equation 5.156. The composite feedback law in Equation 5.156 does not reduce the level of the trackable command input for any choice of the function $\rho(r, y)$. This freedom can be used to improve the performance of the overall system. The choice of $\rho(r, y)$ will be discussed in the forthcoming subsection. \diamond

ii. Full-order Measurement Feedback Case. We proceed to construct a discrete-time full-order CNF control law in the following.

STEP 5.D.F.1: we first construct a linear full-order measurement feedback control law

$$\begin{cases} x_v(k+1) = (A + KC_1)x_v(k) - Ky(k) + B \text{sat}(u_l(k)) \\ u_l(k) = F[x_v(k) - x_e] + Hr \end{cases} \quad (5.168)$$

where r is the command input, $x_v \in \mathbb{R}^n$ is the state of the controller, F and K are chosen such that $A + BF$ and $A + KC_1$ have all their eigenvalues in \mathbb{C}^\ominus , i.e. both are stable matrices, and, furthermore, the resulting closed-loop system has met certain design specifications. As usual, we let

$$x_e = (I - A - BF)^{-1}B \cdot [C_2(I - A - BF)^{-1}B]^{-1}r \quad (5.169)$$

and

$$H = [1 + F(I - A - BF)^{-1}B] \cdot [C_2(I - A - BF)^{-1}B]^{-1} \quad (5.170)$$

We note that both x_e and H are well defined.

Lemma 5.15. *Given a positive-definite matrix $W_p \in \mathbb{R}^{n \times n}$, let $P > 0$ be the solution to the Lyapunov equation*

$$P = (A + BF)'P(A + BF) + W_p \quad (5.171)$$

Given another positive-definite matrix $W_q \in \mathbb{R}^{n \times n}$ with

$$W_q > F'B'P(A + BF)W_p^{-1}(A + BF)'PBF \quad (5.172)$$

let $Q > 0$ be the solution to the Lyapunov equation

$$Q = (A + KC_1)'Q(A + KC_1) + W_q \quad (5.173)$$

Note that such P and Q exist as $A + BF$ and $A + KC_1$ are asymptotically stable. For any $\delta \in (0, 1)$, let c_δ be the largest positive scalar such that for all

$$\begin{pmatrix} x \\ x_v \end{pmatrix} \in \mathbf{X}_{F\delta} := \left\{ \begin{pmatrix} x \\ x_v \end{pmatrix} : \begin{pmatrix} x \\ x_v \end{pmatrix}' \begin{bmatrix} P & 0 \\ 0 & Q \end{bmatrix} \begin{pmatrix} x \\ x_v \end{pmatrix} \leq c_\delta \right\} \quad (5.174)$$

we have

$$\left| [F \quad F] \begin{pmatrix} x \\ x_v \end{pmatrix} \right| \leq u_{\max}(1 - \delta) \quad (5.175)$$

The linear control law in Equation 5.168 drives the system controlled output $h(k)$ to track asymptotically a step command input of amplitude r from an initial state x_0 , provided that $x_0, x_{v0} = x_v(0)$ and r satisfy:

$$|Hr| \leq \delta \cdot u_{\max} \quad \text{and} \quad \begin{pmatrix} x_0 - x_e \\ x_{v0} - x_0 \end{pmatrix} \in \mathbf{X}_{F\delta} \quad (5.176)$$

Proof. This follows along similar lines to the reasoning given in the proofs of Lemmas 5.5 and 5.11. \diamond

STEP 5.D.F.2: the discrete-time full-order measurement composite nonlinear feedback control law is given by

$$x_v(k+1) = (A + KC_1)x_v(k) - Ky(k) + B\text{sat}(u(k)) \quad (5.177)$$

and

$$u(k) = F[x_v(k) - x_e] + Hr + \rho(r, y)B'P(A + BF)[x_v(k) - x_e] \quad (5.178)$$

where $\rho(r, y)$ is a nonpositive scalar function, locally Lipschitz in y , and is to be chosen to improve the performance of the closed-loop system.

We have the following result.

Theorem 5.16. *Consider the given discrete-time system in Equation 5.136. Then, there exists a scalar $0 < \rho^* \leq 2(B'PB)^{-1}$ such that for any nonpositive function $\rho(r, y)$, locally Lipschitz in y and $|\rho(r, y)| \leq \rho^*$, the discrete-time CNF control law in Equations 5.177 and 5.178 internally stabilizes the given plant and drive the system controlled output $h(k)$ to track asymptotically the step command input of amplitude r from an initial state x_0 , provided that x_0 , x_{v0} and r satisfy the conditions in Equation 5.176.* \diamond

Proof. The proof of this theorem follows along similar lines to the reasoning given in Theorems 5.6 and 5.13. \diamond

iii. Reduced-order Measurement Feedback Case. As in its continuous-time counterpart, we now proceed to design a reduced-order measurement feedback controller. For the given system in Equation 5.136, it is clear that p states of the system are measurable if C_1 is of maximal rank. As such, we could design a dynamic controller that has a dynamical order less than that of the given plant. We now proceed to construct such a control law under the CNF control framework.

For simplicity of presentation, we assume that C_1 is already in the form

$$C_1 = [I_p \quad 0] \quad (5.179)$$

Then, the system in Equation 5.136 can be rewritten as

$$\left\{ \begin{array}{l} \begin{pmatrix} x_1(k+1) \\ x_2(k+1) \end{pmatrix} = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} + \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} \text{sat}(u(k)) \\ y(k) = [I_p \quad 0] \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} \\ h(k) = C_2 \begin{pmatrix} x_1(k) \\ x_2(k) \end{pmatrix} \end{array} \right. \quad (5.180)$$

and

$$x_0 = \begin{pmatrix} x_{10} \\ x_{20} \end{pmatrix} \quad (5.181)$$

where the original state x is partitioned into two parts, x_1 and x_2 with $y \equiv x_1$. Thus, we only need to estimate x_2 in the reduced-order measurement feedback design. Next, we let F be chosen such that 1) $A + BF$ is asymptotically stable, and 2) $C_2(sI - A - BF)^{-1}B$ has the desired properties, and let K_R be chosen such that $A_{22} + K_R A_{12}$ is asymptotically stable. Again, it follows from Chen [110] that (A_{22}, A_{12}) is detectable if and only if (A, C_1) is detectable. Thus, there exists a stabilizing K_R . Again, such F and K_R can be designed using any of the linear control techniques presented in Chapter 3. We then partition F in conformity with x_1 and x_2 :

$$F = [F_1 \quad F_2] \quad (5.182)$$

As defined in Equations 5.169 and 5.169, we let

$$x_e = (I - A - BF)^{-1}B \cdot [C_2(I - A - BF)^{-1}B]^{-1}r \quad (5.183)$$

and

$$H = [1 + F(I - A - BF)^{-1}B] \cdot [C_2(I - A - BF)^{-1}B]^{-1} \quad (5.184)$$

The reduced-order CNF controller is given by

$$\begin{aligned} x_v(k+1) &= (A_{22} + K_R A_{12})x_v(k) + (B_2 + K_R B_1) \text{sat}(u(k)) \\ &\quad + [A_{21} + K_R A_{11} - (A_{22} + K_R A_{12})K_R]y(k) \end{aligned} \quad (5.185)$$

and

$$\begin{aligned} u(k) &= F \left[\begin{pmatrix} y \\ x_v(k) - K_R y(k) \end{pmatrix} - x_e \right] + Hr \\ &\quad + \rho(r, y)B'P(A + BF) \left[\begin{pmatrix} y(k) \\ x_v - K_R y(k) \end{pmatrix} - x_e \right] \end{aligned} \quad (5.186)$$

where $\rho(r, y)$ is a nonpositive scalar function locally Lipschitz in y subject to certain constraints to be discussed later.

Next, given a positive-definite matrix $W_p \in \mathbb{R}^{n \times n}$, let $P > 0$ be the solution to the Lyapunov equation

$$P = (A + BF)'P(A + BF) + W_p \quad (5.187)$$

Given another positive-definite matrix $W_r \in \mathbb{R}^{(n-p) \times (n-p)}$ with

$$W_r > F_2' B' P (A + BF) W_p^{-1} (A + BF)' P B F_2 \quad (5.188)$$

let $Q_r > 0$ be the solution to the Lyapunov equation

$$Q_r = (A_{22} + K_R A_{12})' Q_r (A_{22} + K_R A_{12}) + W_r \quad (5.189)$$

Note that such P and Q_R exist as $A + BF$ and $A_{22} + K_R A_{12}$ are asymptotically stable. For any $\delta \in (0, 1)$, let c_δ be the largest positive scalar such that for all

$$\begin{pmatrix} x(k) \\ x_v(k) \end{pmatrix} \in \mathbf{X}_{r\delta} := \left\{ \begin{pmatrix} x \\ x_v \end{pmatrix} : \begin{pmatrix} x \\ x_v \end{pmatrix}' \begin{bmatrix} P & 0 \\ 0 & Q_R \end{bmatrix} \begin{pmatrix} x \\ x_v \end{pmatrix} \leq c_\delta \right\} \quad (5.190)$$

we have

$$\left| \begin{bmatrix} F & E_2 \end{bmatrix} \begin{pmatrix} x(k) \\ x_v(k) \end{pmatrix} \right| \leq u_{\max}(1 - \delta) \quad (5.191)$$

We have the following theorem.

Theorem 5.17. *Consider the system given in Equation 5.1. Then, there exists a scalar $0 < \rho^* \leq 2(B'PB)^{-1}$ such that for any nonpositive function $\rho(r, y)$, locally Lipschitz in y and $|\rho(r, y)| \leq \rho^*$, the reduced-order CNF control law given by Equations 5.185 and 5.186 internally stabilizes the given plant and drives the system controlled output $h(k)$ to track asymptotically the step command input of amplitude r from an initial state x_0 , provided that x_0 , x_{v0} and r satisfy*

$$\begin{pmatrix} x_0 - x_e \\ x_{v0} - x_{20} - K_R x_{10} \end{pmatrix} \in \mathbf{X}_{r\delta} \quad \text{and} \quad |Hr| \leq \delta \cdot u_{\max} \quad (5.192)$$

Proof. Again, the proof of this theorem is similar to those given earlier. \diamond

5.3.2 Systems with External Disturbances

We consider a linear discrete-time system with actuator saturation and disturbances characterized by

$$\begin{cases} x(k+1) = Ax(k) + B \text{sat}(u(k)) + Ew, & x(0) = x_0 \\ y(k) = C_1 x(k) \\ h(k) = C_2 x(k) \end{cases} \quad (5.193)$$

where $x \in \mathbb{R}^n$, $u \in \mathbb{R}$, $y \in \mathbb{R}^p$, $h \in \mathbb{R}$ and $w \in \mathbb{R}$ are, respectively, the state, control input, measurement output, controlled output and disturbance input of the system. A , B , C_1 , C_2 and E are appropriate dimensional constant matrices. The function, $\text{sat}: \mathbb{R} \rightarrow \mathbb{R}$, represents the actuator saturation defined as

$$\text{sat}(u(k)) = \text{sgn}(u(k)) \min\{u_{\max}, |u(k)|\} \quad (5.194)$$

with u_{\max} being the input saturation level. The following assumptions on the given system are made:

1. (A, B) is stabilizable,
2. (A, C_1) is detectable,
3. (A, B, C_2) is invertible with no invariant zero at $z = 1$,
4. w is bounded unknown constant disturbance, and

5. $h(k)$ is part of $y(k)$, *i.e.* $h(k)$ is also measurable.

We aim to design a discrete enhanced CNF control law for the system with input saturation and disturbances to track a step reference, say r , neither violating the input saturation nor having steady-state bias. An equivalent discrete integration, which eventually becomes part of the final control law, is defined as follows,

$$x_i(k+1) = x_i(k) + \kappa_i e(k) = x_i(k) + \kappa_i C_2 x(k) - \kappa_i r \quad (5.195)$$

where the tracking error $e(k) := h(k) - r$ is available for feedback as $h(k)$ is assumed to be measurable and κ_i is a positive scalar to be selected to yield an appropriate integration speed. By integrating Equation 5.195 into the given system, we obtain the following augmented system

$$\begin{cases} \bar{x}(k+1) = \bar{A} \bar{x}(k) + \bar{B} \text{sat}(u(k)) + \bar{B}_r r + \bar{E} w \\ \bar{y}(k) = \bar{C}_1 \bar{x}(k) \\ h(k) = \bar{C}_2 \bar{x}(k) \end{cases} \quad (5.196)$$

where

$$\bar{x}(k) = \begin{pmatrix} x_i(k) \\ x(k) \end{pmatrix}, \quad \bar{x}_0 = \begin{pmatrix} 0 \\ x_0 \end{pmatrix}, \quad \bar{y}(k) = \begin{pmatrix} x_i(k) \\ y(k) \end{pmatrix} \quad (5.197)$$

$$\bar{A} = \begin{bmatrix} 1 & \kappa_i C_2 \\ 0 & A \end{bmatrix}, \quad \bar{B} = \begin{bmatrix} 0 \\ B \end{bmatrix}, \quad \bar{B}_r = \begin{bmatrix} -\kappa_i \\ 0 \end{bmatrix} \quad (5.198)$$

and

$$\bar{E} = \begin{bmatrix} 0 \\ E \end{bmatrix}, \quad \bar{C}_1 = \begin{bmatrix} 1 & 0 \\ 0 & C_1 \end{bmatrix}, \quad \bar{C}_2 = [0 \quad C_2] \quad (5.199)$$

We note that under Assumptions 1 and 3, it is straightforward to verify that the pair (\bar{A}, \bar{B}) is stabilizable.

In what follows, we proceed to design an enhanced CNF control laws for the given system for two different cases, *i.e.* the state feedback case and the reduced-order measurement feedback case. The full-order measurement feedback case can be solved in a straightforward manner once the result for the reduced-order case is established.

i. State Feedback Case. We consider in the following the situation when all the state variables of the given system in Equation 5.193 are measurable, *i.e.* $y = x$. The procedure that generates an enhanced CNF state feedback law is done in three steps. That is, in the first step, a linear feedback control law with appropriate properties is designed, then in the second step, the design of nonlinear feedback portion is carried out, and lastly, in the final step, the linear and nonlinear feedback laws are combined to form an enhanced CNF control law.

STEP 5.D.W.S.1: Design a linear feedback control law,

$$u_l(k) = F \bar{x}(k) + G r \quad (5.200)$$

where F is chosen such that i) $\bar{A} + \bar{B}F$ is asymptotically stable, and ii) the closed-loop system $\bar{C}_2(zI - \bar{A} - \bar{B}F)^{-1}\bar{B}$ has certain desired properties. Let us partition $F = [F_1 \quad F_x]$ in conformity with $x_1(k)$ and $x(k)$. The general guideline in designing such a state feedback gain F is to place the closed-loop pole of $\bar{A} + \bar{B}F$ corresponding to $x_1(k)$ to be sufficiently closer to $z = 1$ compared to the other eigenvalues, which implies that F_1 is a relatively small scalar. The remaining closed-loop poles of $\bar{A} + \bar{B}F$ are placed to have a dominating pair with a small damping ratio, which in turn would yield a fast rise time in the closed-loop system response. Finally, G is chosen as

$$G = [C_2(I - A - BF_x)^{-1}B]^{-1} \quad (5.201)$$

which is well defined as (A, B, C_2) is assumed to have no invariant zeros at $z = 1$ and $A + BF_x$ is nonsingular whenever $\bar{A} + \bar{B}F$ is stable and F_1 is relatively small.

STEP 5.D.W.S.2: Given an appropriate positive-definite matrix $W \in \mathbb{R}^{(n+1) \times (n+1)}$, we solve the following Lyapunov equation:

$$P = (\bar{A} + \bar{B}F)'P(\bar{A} + \bar{B}F) + W \quad (5.202)$$

for $P > 0$. Such a solution is always existent as $\bar{A} + \bar{B}F$ is asymptotically stable. The nonlinear feedback portion of the enhanced CNF control law, $u_n(k)$, is then given by

$$u_n(k) = \rho(e(k))\bar{B}'P(\bar{A} + \bar{B}F)[\bar{x}(k) - \bar{x}_e] \quad (5.203)$$

where $\rho(e(k))$, with $e = h - r$, is a nonpositive function of $|e|$ and tends to a finite scalar as $k \rightarrow \infty$. It is to be used to gradually change the system closed-loop damping ratio to yield a better tracking performance. The choices of the design parameters, $\rho(e(k))$ and W , will be discussed later. Next, we define

$$G_e := \begin{bmatrix} 0 \\ (I - A - BF_x)^{-1}BG \end{bmatrix}, \quad \bar{x}_e := G_e r \quad (5.204)$$

STEP 5.D.W.S.3: the linear and nonlinear feedback control laws derived in the previous steps are now combined to form an enhanced CNF control law,

$$u(k) = F\bar{x}(k) + Gr + \rho(e(k))\bar{B}'P(\bar{A} + \bar{B}F)[\bar{x}(k) - \bar{x}_e] \quad (5.205)$$

We have the following result.

Theorem 5.18. Consider the given system in Equation 5.193 with $y = x$ and the disturbance w being bounded by a non-negative scalar τ_w , i.e. $|w| \leq \tau_w$. Let

$$\gamma := \left[1 - (1 - \lambda_{\min}(WP^{-1}))^{1/2}\right]^{-1} (\bar{E}'P\bar{E})^{1/2} \tau_w \quad (5.206)$$

Then, for any $\rho(e(k)) \in [-2(\bar{B}'P\bar{B})^{-1}, 0]$, which is a nonpositive function of $|e|$ and tends to a constant as $k \rightarrow \infty$, the enhanced CNF control law in Equation 5.205 internally stabilizes the given plant and drives the system controlled output $h(k)$ to track the step reference of amplitude r from an initial state \bar{x}_0 asymptotically without steady-state error, provided that the following conditions are satisfied:

1. There exist scalars $\delta \in (0, 1)$ and $c_\delta > \gamma^2$ such that

$$\begin{aligned} \forall \bar{x}(k) \in \mathbf{X}(F, c_\delta) &:= \{\bar{x}(k) : \bar{x}'(k)P\bar{x}(k) \leq c_\delta\} \\ &\Rightarrow |F\bar{x}(k)| \leq (1 - \delta)u_{\max} \end{aligned} \quad (5.207)$$

2. The initial condition, \bar{x}_0 , satisfies

$$\bar{x}_0 - \bar{x}_e \in \mathbf{X}(F, c_\delta) \quad (5.208)$$

3. The level of the target reference, r , satisfies

$$|Hr| \leq \delta u_{\max} \quad (5.209)$$

where $H := FG_e + G$. Note that $\lambda_{\min}(WP^{-1}) \in (0, 1)$.

Proof. For simplicity, we drop $e(k)$ in the nonlinear function ρ throughout the following proof. First, it is straightforward to verify that

$$\bar{A}\bar{x}_e + \bar{B}Hr + \bar{B}_r r = \bar{x}_e. \quad (5.210)$$

Letting $\tilde{x}(k) = \bar{x}(k) - \bar{x}_e$, the augmented system in Equation 5.196 can be expressed as

$$\tilde{x}(k+1) = (\bar{A} + \bar{B}F)\tilde{x}(k) + \bar{B}v(k) + \bar{E}w \quad (5.211)$$

where

$$v(k) := \text{sat}(u(k)) - F\tilde{x}(k) - Hr \quad (5.212)$$

and the control law in Equation 5.205 can be rewritten as

$$u(k) = F\tilde{x}(k) + Hr + \rho\bar{B}'P(\bar{A} + \bar{B}F)\tilde{x}(k) \quad (5.213)$$

Next, for $\tilde{x}(k) \in \mathbf{X}(F, c_\delta)$ and $|Hr| \leq \delta u_{\max}$, we have

$$|F\tilde{x}(k) + Hr| \leq |F\tilde{x}(k)| + |Hr| \leq u_{\max} \quad (5.214)$$

which implies

$$v(k) = \rho\bar{B}'P(\bar{A} + \bar{B}F)\tilde{x}(k) \quad (5.215)$$

if $|u(k)| \leq u_{\max}$, or

$$0 < v(k) < \rho\bar{B}'P(\bar{A} + \bar{B}F)\tilde{x}(k) \quad (5.216)$$

if $u(k) > u_{\max}$, or

$$\rho\bar{B}'P(\bar{A} + \bar{B}F)\tilde{x}(k) < v(k) < 0 \quad (5.217)$$

if $u(k) < -u_{\max}$. Obviously, for all possible situations, $v(k)$ can be written as

$$v(k) = q\rho\bar{B}'P(\bar{A} + \bar{B}F)\tilde{x}(k) \quad (5.218)$$

with some appropriate $q \in [0, 1]$. Thus, for $\tilde{x}(k) \in \mathbf{X}(F, c_\delta)$ and $|Hr| \leq \delta u_{\max}$, the closed-loop system comprising the augmented system in Equation 5.196 and the CNF control law in Equation 5.205 can be expressed as follows

$$\tilde{x}(k+1) = [\bar{A} + \bar{B}F + q\rho\bar{B}\bar{B}'P(\bar{A} + \bar{B}F)]\tilde{x}(k) + \bar{E}w \quad (5.219)$$

Defining a discrete-time Lyapunov function, $V(k) = \tilde{x}'(k)P\tilde{x}(k)$, and factoring $P > 0$ as $P = S'S$, the increment of $V(k)$ along the trajectory of the system in Equation 5.219 can be calculated as

$$\begin{aligned} \Delta V(k) &= -\tilde{x}'(k)P\tilde{x}(k) + \tilde{x}'(k)(\bar{A} + \bar{B}F)'P(\bar{A} + \bar{B}F)\tilde{x}(k) \\ &\quad + \tilde{x}'(k)(\bar{A} + \bar{B}F)'P\bar{B}(2q\rho + q^2\rho^2\bar{B}'P\bar{B})\bar{B}'P(\bar{A} + \bar{B}F)\tilde{x}(k) \\ &\quad + w'\bar{E}'P\bar{E}w + 2\tilde{x}'(k)(\bar{A} + \bar{B}F)'(I + q\rho\bar{B}\bar{B}'P)'P\bar{E}w \\ &\leq -\tilde{x}'(k)P\tilde{x}(k) + \tilde{x}'(k)(P - W)\tilde{x}(k) + \bar{E}'P\bar{E}\tau_w^2 \\ &\quad + 2\|\tilde{x}'(k)(\bar{A} + \bar{B}F)'(I + q\rho\bar{B}\bar{B}'P)'S'\| \cdot \|S\bar{E}\|\tau_w \end{aligned} \quad (5.220)$$

Noting that

$$\begin{aligned} &\|\tilde{x}'(k)(\bar{A} + \bar{B}F)'(I + q\rho\bar{B}\bar{B}'P)'S'\| \\ &\leq \left[\tilde{x}'(k)(\bar{A} + \bar{B}F)'P(\bar{A} + \bar{B}F)\tilde{x}(k) \right]^{1/2} \end{aligned} \quad (5.221)$$

for $\rho \in [-2(\bar{B}'P\bar{B})^{-1}, 0]$, we have

$$\begin{aligned} \Delta V(k) &\leq -\tilde{x}'(k)P\tilde{x}(k) + \tilde{x}'(k)(P - W)\tilde{x}(k) + \bar{E}'P\bar{E}\tau_w^2 \\ &\quad + 2\left[\tilde{x}'(k)(\bar{A} + \bar{B}F)'P(\bar{A} + \bar{B}F)\tilde{x}(k) \right]^{1/2} (\bar{E}'P\bar{E})^{1/2}\tau_w \\ &= -\left\{ (\tilde{x}'(k)P\tilde{x}(k))^{1/2} + [\tilde{x}'(k)(P - W)\tilde{x}(k)]^{1/2} + (\bar{E}'P\bar{E})^{1/2}\tau_w \right\} \\ &\quad \times \left\{ (\tilde{x}'(k)P\tilde{x}(k))^{1/2} - [\tilde{x}'(k)(P - W)\tilde{x}(k)]^{1/2} - (\bar{E}'P\bar{E})^{1/2}\tau_w \right\} \\ &= -\left\{ (\tilde{x}'(k)P\tilde{x}(k))^{1/2} + [\tilde{x}'(k)(P - W)\tilde{x}(k)]^{1/2} + (\bar{E}'P\bar{E})^{1/2}\tau_w \right\} \\ &\quad \times \left\{ -(\bar{E}'P\bar{E})^{1/2}\tau_w + (\tilde{x}'(k)P\tilde{x}(k))^{1/2} \left[1 - \left(1 - \frac{\tilde{x}'(k)W\tilde{x}(k)}{\tilde{x}'(k)P\tilde{x}(k)} \right)^{1/2} \right] \right\} \\ &\leq -\left\{ (\tilde{x}'(k)P\tilde{x}(k))^{1/2} + [\tilde{x}'(k)(P - W)\tilde{x}(k)]^{1/2} + (\bar{E}'P\bar{E})^{1/2}\tau_w \right\} \\ &\quad \times \left\{ -(\bar{E}'P\bar{E})^{1/2}\tau_w + (\tilde{x}'(k)P\tilde{x}(k))^{1/2} \left[1 - (1 - \lambda_{\min}(WP^{-1}))^{1/2} \right] \right\} \\ &= -\left\{ (\tilde{x}'(k)P\tilde{x}(k))^{1/2} + [\tilde{x}'(k)(P - W)\tilde{x}(k)]^{1/2} + (\bar{E}'P\bar{E})^{1/2}\tau_w \right\} \\ &\quad \times \left[1 - (1 - \lambda_{\min}(WP^{-1}))^{1/2} \right] \left[(\tilde{x}'(k)P\tilde{x}(k))^{1/2} - \gamma \right] \end{aligned} \quad (5.222)$$

Note that we have used the following property:

$$\lambda_{\min}(WP^{-1}) = \min_{x \neq 0} \frac{x'Wx}{x'Px} \quad (5.223)$$

as both W and P are positive-definite matrices. Clearly, the closed-loop system in the absence of the disturbance, w , has $\Delta V(k) < 0$ and thus is asymptotically stable.

With the presence of the disturbance, w , and with $\tilde{x}(0) = \bar{x}_0 - \bar{x}_e \in \mathbf{X}(F, c_\delta)$, where $c_\delta > \gamma^2$, the corresponding trajectory of Equation 5.219 remains in $\mathbf{X}(F, c_\delta)$ and converges to a ball characterized by $\{\tilde{x} : \tilde{x}'P\tilde{x} \leq \tilde{\gamma}^2\}$ with $\tilde{\gamma} \leq \gamma$.

Note that ρ is chosen such that it tends to a constant as $k \rightarrow \infty$. Also, for large k , the control signal $u(k)$ is under its saturation level. Thus, the closed-loop system in Equation 5.219 becomes an almost linear time-invariant system. We can conclude that the corresponding trajectory of Equation 5.219 converges asymptotically to a point, *i.e.* $\tilde{x}(k)$ tends to a constant. Thus,

$$\lim_{k \rightarrow \infty} e(k) = \lim_{k \rightarrow \infty} \frac{1}{\kappa_1} [x_i(k+1) - x_i(k)] = 0 \quad (5.224)$$

This completes the proof of Theorem 5.18. \diamond

ii. Measurement Feedback Case. Next, we consider the general measurement feedback situation, in which there is only part of state variables available for feedback. As usual, for such a situation, one could either design a full-order or a reduced-order measurement feedback control law. In what follows, we focus on designing a reduced-order controller. Without loss of generality, we assume that C_1 in the measurement output of the given plant in Equation 5.193 is in the form:

$$C_1 = [I_p \quad 0] \quad (5.225)$$

The augmented plant in Equation 5.196 can then be partitioned as

$$\begin{pmatrix} x_i(k+1) \\ x_1(k+1) \\ x_2(k+1) \end{pmatrix} = \begin{bmatrix} 1 & \kappa_1 C_{21} & \kappa_1 C_{22} \\ 0 & A_{11} & A_{12} \\ 0 & A_{21} & A_{22} \end{bmatrix} \begin{pmatrix} x_i(k) \\ x_1(k) \\ x_2(k) \end{pmatrix} + \begin{bmatrix} 0 \\ B_1 \\ B_2 \end{bmatrix} \text{sat}(u(k)) \\ + \begin{bmatrix} -\kappa_1 \\ 0 \\ 0 \end{bmatrix} r + \begin{bmatrix} 0 \\ E_1 \\ E_2 \end{bmatrix} w \quad (5.226)$$

$$\bar{y}(k) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & I_p & 0 \end{bmatrix} \begin{pmatrix} x_i \\ x_1 \\ x_2 \end{pmatrix} (k) \quad (5.227)$$

and

$$h(k) = [0 \quad C_{21} \quad C_{22}] \begin{pmatrix} x_i \\ x_1 \\ x_2 \end{pmatrix} (k) \quad (5.228)$$

with

$$\begin{pmatrix} x_i(0) \\ x_1(0) \\ x_2(0) \end{pmatrix} = \begin{pmatrix} 0 \\ x_{10} \\ x_{20} \end{pmatrix} = \bar{x}_0 \quad (5.229)$$

We only need to estimate $x_2(k)$. There are two main steps in designing a reduced-order measurement feedback control law: i) the construction of a full state feedback gain matrix F , which is totally identical to that given in the previous subsection; and

ii) the reduced-order observer gain matrix K_R , which is chosen such that the poles of $A_{22} + K_R A_{12}$ are placed at appropriate locations inside the unit circle. To derive the reduced-order measurement feedback controller, we partitioned $F = [F_1 \ F_1 \ F_2]$ in conformity with $x_1(k)$, $x_1(k)$ and $x_2(k)$. The reduced-order enhanced CNF control law is then given by,

$$\begin{aligned} x_v(k+1) &= (A_{22} + K_R A_{12})x_v(k) + (B_2 + K_R B_1)\text{sat}(u(k)) \\ &\quad + [A_{21} + K_R A_{11} - (A_{22} + K_R A_{12})K_R]y(k) \end{aligned} \quad (5.230)$$

and

$$u(k) = Fx_r(k) + Gr + \rho(e(k))\bar{B}'P(\bar{A} + \bar{B}F)[x_r(k) - \bar{x}_e] \quad (5.231)$$

where

$$x_r(k) = \begin{pmatrix} x_1(k) \\ x_1(k) \\ x_v(k) - K_R y(k) \end{pmatrix}$$

G is as defined in Equation 5.201, \bar{x}_e is as defined in Equation 5.204 and $\rho(e(k))$ is the nonpositive function of $|e(k)|$, which tends to a constant as $k \rightarrow \infty$.

Next, given a positive-definite matrix $W \in \mathbb{R}^{(n+1) \times (n+1)}$, let $P > 0$ be the solution to the Lyapunov equation

$$P = (\bar{A} + \bar{B}F)'P(\bar{A} + \bar{B}F) + W \quad (5.232)$$

Let $W_R \in \mathbb{R}^{(n-p) \times (n-p)}$ be a positive-definite matrix such that

$$W_R > F_2'[\bar{B}'P\bar{B} + \bar{B}'P(\bar{A} + \bar{B}F)W^{-1}(\bar{A} + \bar{B}F)'P\bar{B}]F_2 \quad (5.233)$$

and let $Q_R > 0$ be the solution to the Lyapunov equation

$$Q_R = (A_{22} + K_R A_{12})'Q_R(A_{22} + K_R A_{12}) + W_R \quad (5.234)$$

Note that such P and Q_R exist as $\bar{A} + \bar{B}F$ and $A_{22} + K_R A_{12}$ are both asymptotically stable. Next, let

$$\lambda_R := \lambda_{\min} \left\{ \begin{bmatrix} P & 0 \\ 0 & Q_R \end{bmatrix}^{-1} \begin{bmatrix} W & -(\bar{A} + \bar{B}F)'P\bar{B}F_2 \\ -F_2'\bar{B}'P(\bar{A} + \bar{B}F) & W_R - F_2'\bar{B}'P\bar{B}F_2 \end{bmatrix} \right\} \quad (5.235)$$

It is noted that $\lambda_R \in (0, 1)$. We have the following result.

Theorem 5.19. *Consider the given system in Equation 5.193 with the disturbance w being bounded by a scalar $\tau_w > 0$, i.e. $|w| \leq \tau_w$. Let*

$$\gamma_R := \left[1 - (1 - \lambda_R)^{1/2} \right]^{-1} \left[\bar{E}'P\bar{E} + (E_2 + K_R E_1)'Q_R(E_2 + K_R E_1) \right]^{1/2} \tau_w \quad (5.236)$$

Then, there exists a $\rho^ \in (0, 2(\bar{B}'P\bar{B})^{-1}]$ such that for any $\rho(e(k)) \in [-\rho^*, 0]$, which is a nonpositive function of $|e(k)|$ and tends to a constant as $k \rightarrow \infty$, the*

reduced-order enhanced CNF control law of Equations 5.230 and 5.231 internally stabilizes the given plant and drives the system controlled output $h(k)$ to track the step reference of amplitude r asymptotically without steady-state error, provided that the following conditions are satisfied:

1. There exist positive scalars $\delta \in (0, 1)$ and $c_{r\delta} > \gamma_R^2$ such that

$$\begin{aligned} \forall \bar{x}(k) \in \mathbf{X}(F, c_{r\delta}) &:= \left\{ \bar{x}(k) : \bar{x}'(k) \begin{bmatrix} P & 0 \\ 0 & Q_R \end{bmatrix} \bar{x}(k) \leq c_{r\delta} \right\} \\ &\Rightarrow |[F \quad F_2] \bar{x}(k)| \leq (1 - \delta)u_{\max} \end{aligned} \quad (5.237)$$

2. The initial conditions, \bar{x}_0 and $x_{v0} = x_v(0)$, satisfy

$$\begin{pmatrix} \bar{x}_0 - \bar{x}_e \\ x_{v0} - x_{20} - K_R x_{10} \end{pmatrix} \in \mathbf{X}(F, c_{r\delta}) \quad (5.238)$$

3. The level of the target reference, r , satisfies

$$|Hr| \leq \delta u_{\max} \quad (5.239)$$

where H is the same as that defined in Theorem 5.18.

Proof. It follows from the similar lines of reasoning as those in Theorem 5.18 and the similar arguments for the measurement feedback case reported in its continuous-time counterpart in the previous section. \diamond

5.3.3 Selection of Nonlinear Feedback Parameters

As in its continuous-time counterpart, we focus our attention on the situation for the case when the given system has external disturbances. Since the main purpose of adding the nonlinear part to the CNF controller is to shorten the settling time and to reduce the overshoot, or equivalently to contribute a significant value to the control input when the tracking error, $h - r$, is small, it is appropriate for us to select a nonlinear gain matrix such that the nonlinear part is in action when the control signal is far from its saturation level, and thus it does not cause the control input to hit its limits. Under such a circumstance, it is straightforward to verify that the closed-loop system comprising the augmented plant in Equation 5.196 and the CNF control law in Equation 5.205 can be expressed as

$$\tilde{x}(k+1) = (\bar{A} + \bar{B}F)\tilde{x}(k) + \rho \bar{B}\bar{B}'P(\bar{A} + \bar{B}F)\tilde{x}(k) + \bar{E}w \quad (5.240)$$

As in its continuous-time counterpart, the eigenvalues of the closed-loop system in Equation 5.240 can be changed by the nonlinear function ρ . Assuming that h is available, we propose the following nonlinear gain function

$$\rho(e(k)) = -\beta(\bar{B}'P\bar{B})^{-1} \frac{2}{\pi} \arctan \left(\alpha \left| |h(k) - r| - |h(0) - r| \right| \right) \quad (5.241)$$

with $0 \leq \beta \leq 2$. We note that ρ starts from 0 and gradually decreases to a constant $-\beta(\bar{B}'P\bar{B})^{-1}2\arctan(\alpha|h(0)-r|)/\pi > -\beta(\bar{B}'P\bar{B})^{-1}$ as h approaches the target reference r . The parameter α is used to determine the speed of change in ρ .

To examine the behavior of the closed-loop system in Equation 5.240 more explicitly, we define an auxiliary system $G_{\text{aux}}(z)$ characterized by

$$G_{\text{aux}}(z) := C_{\text{aux}}(zI - A_{\text{aux}})^{-1}B_{\text{aux}} := \bar{B}'P(zI - \bar{A} - \bar{B}F)^{-1}\bar{B} \quad (5.242)$$

Clearly, $G_{\text{aux}}(z)$ is stable. Note that $C_{\text{aux}}B_{\text{aux}} = \bar{B}'P\bar{B} > 0$, which implies $G_{\text{aux}}(z)$ is a square, invertible and uniform rank system with a relative degree of 1 and with n invariant zeros. We show that this auxiliary system is in fact of minimum phase, *i.e.* all its invariant zeros are stable. We note that for such a system, it follows from the result reported in Chapter 5 of Chen *et al.* [71] that there exist nonsingular transformations $\Gamma_s \in \mathbb{R}^{(n+1) \times (n+1)}$, $\Gamma_i \in \mathbb{R}$ and $\Gamma_o \in \mathbb{R}$ such that the transformed system has the following special form,

$$\left(\Gamma_s^{-1}A_{\text{aux}}\Gamma_s, \Gamma_s^{-1}B_{\text{aux}}\Gamma_i, \Gamma_o^{-1}C_{\text{aux}}\Gamma_s \right) = \left(\begin{bmatrix} A_{\text{aa}} & L_{\text{ad}} \\ E_{\text{da}} & A_{\text{dd}} \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \end{bmatrix}, [0 \quad 1] \right)$$

where the eigenvalues of A_{aa} are the invariant zeros of the auxiliary system $G_{\text{aux}}(z)$, L_{ad} , E_{da} and A_{dd} are some constant matrices. Next, we proceed to show that all the eigenvalues of A_{aa} are inside the unit circle and thus $G_{\text{aux}}(z)$ is of minimum phase. We note that at the steady state when $h = r$, the nonlinear function matrix ρ of Equation 5.241 with an appropriately chosen β can be set to $\rho = -(\bar{B}'P\bar{B})^{-1}$ and the closed-loop system of Equation 5.240 can be expressed as

$$\begin{aligned} \tilde{x}(k+1) &= (\bar{A} + \bar{B}F)\tilde{x}(k) - \bar{B}(\bar{B}'P\bar{B})^{-1}\bar{B}'P(\bar{A} + \bar{B}F)\tilde{x}(k) \\ &= [I - \bar{B}(\bar{B}'P\bar{B})^{-1}\bar{B}'P](\bar{A} + \bar{B}F)\tilde{x}(k) \\ &= [I - B_{\text{aux}}(C_{\text{aux}}B_{\text{aux}})^{-1}C_{\text{aux}}]A_{\text{aux}}\tilde{x}(k) \\ &= \left[I - \Gamma_s \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Gamma_i^{-1} \left(\Gamma_o [0 \quad 1] \Gamma_s^{-1} \Gamma_s \begin{bmatrix} 0 \\ 1 \end{bmatrix} \Gamma_i^{-1} \right)^{-1} \Gamma_o [0 \quad 1] \Gamma_s^{-1} \right] \\ &\quad \times \Gamma_s \begin{bmatrix} A_{\text{aa}} & L_{\text{ad}} \\ E_{\text{da}} & A_{\text{dd}} \end{bmatrix} \Gamma_s^{-1} \tilde{x}(k) \\ &= \left(\Gamma_s \begin{bmatrix} A_{\text{aa}} & L_{\text{ad}} \\ 0 & 0 \end{bmatrix} \Gamma_s^{-1} \right) \tilde{x}(k) \end{aligned} \quad (5.243)$$

Clearly, the closed-loop system has n eigenvalues at $\lambda(A_{\text{aa}})$ and one at 0. Thus, the stability of the closed-loop system with $\rho = -(\bar{B}'P\bar{B})^{-1}$ implies the eigenvalues of A_{aa} are all inside the unit circle. This shows that $G_{\text{aux}}(z)$ is indeed of minimum phase.

It is noted that there is freedom in preselecting the locations of these invariant zeros by choosing an appropriate W in Equation 5.202. In general, we should select the invariant zeros of $G_{\text{aux}}(z)$, which correspond to the closed-loop poles of Equation 5.240 for the steady-state nonlinear gain matrix, with dominating ones having a large damping ratio, which in turn generally yield a smaller overshoot. The following procedure might be used for such a purpose.

1. Given a set of n self-conjugated complex scalars, which include all the uncontrollable modes, if any, of (\bar{A}, \bar{B}) , we are to determine a $W > 0$ such that the resulting auxiliary system $G_{\text{aux}}(z)$ has its invariant zeros placed exactly at the locations given in the set.

First, use the singular value decomposition technique to find a unitary matrix $U \in \mathbb{R}^{n \times n}$ and a nonsingular matrix $T_i \in \mathbb{R}^{m \times m}$ such that

$$\tilde{B}_{\text{aux}} = U' B_{\text{aux}} T_i = U' \bar{B} T_i = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

and partition accordingly

$$\tilde{A}_{\text{aux}} = U' A_{\text{aux}} U = U' (\bar{A} + \bar{B} F) U = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

It is straightforward to verify that the stabilizability of (\bar{A}, \bar{B}) implies the stabilizability of (A_{11}, A_{12}) . In fact, their uncontrollable modes, if any, are identical. Next, for determining an appropriate matrix $P = P' > 0$, we partition it accordingly as follows

$$\tilde{P} = U' P U = \begin{bmatrix} P_{11} & P'_{21} \\ P_{21} & P_{22} \end{bmatrix} \quad (5.244)$$

Then, C_{aux} can be expressed as

$$\begin{aligned} C_{\text{aux}} &= \bar{B}' P = (T_i^{-1})' [0 \quad 1] U' U \begin{bmatrix} P_{11} & P'_{21} \\ P_{21} & P_{22} \end{bmatrix} U' \\ &= (T_i^{-1})' [P_{21} \quad P_{22}] U' \\ &= [(T_i^{-1})' P_{22}] [P_{22}^{-1} P_{21} \quad 1] U' \\ &:= T_o [P_{22}^{-1} P_{21} \quad 1] U' \end{aligned} \quad (5.245)$$

Using the results of Chen *et al.* [71] (see *e.g.*, Chapters 8 and 9), we can show that the invariant zeros of the auxiliary system $G_{\text{aux}}(z)$ are given by the eigenvalues of $A_{11} - A_{12} P_{22}^{-1} P_{21}$. Since (A_{11}, A_{12}) is stabilizable and the given set of complex scalars include all its uncontrollable modes, there exists a constant matrix, say F_* , such that $A_{11} - A_{12} F_*$ has its eigenvalues placed exactly at the locations given in the set. Obviously, we can select P_{22} and P_{21} such that $P_{22}^{-1} P_{21} = F_*$

2. Select appropriately dimensional matrices $P_{22} = P'_{22} > 0$, $P_{21} = P_{22} F_*$, and $P_{11} = P'_{11} > P'_{21} P_{22}^{-1} P_{21}$ to ensure that

$$P = U \begin{bmatrix} P_{11} & P'_{21} \\ P_{21} & P_{22} \end{bmatrix} U' > 0 \quad (5.246)$$

3. Compute

$$W = P - (\bar{A} + \bar{B} F)' P (\bar{A} + \bar{B} F) \quad (5.247)$$

If W is not positive-definite, we go back to Step 2 to choose another solution of P or go to the first step to reselect another set of desired invariant zeros.

Another method for selecting W is based on a trial and error approach by limiting the choice of W to be a diagonal matrix and adjusting its diagonal weights through simulation. Generally, such an approach would yield a satisfactory result as well.

5.3.4 An Illustrative Example

We illustrate the enhanced CNF control technique for discrete-time systems in the following example. We consider a discrete-time system of Equation 5.193 with

$$A = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & -3 & 3 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad E = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \quad x_0 = \begin{pmatrix} 0 \\ 0 \\ 0.5 \end{pmatrix} \quad (5.248)$$

$$C_1 = I_3, \quad C_2 = [1 \ 0 \ 0] \quad (5.249)$$

and $u_{\max} = 1$. The disturbance w is unknown. For simulation purpose, we assume $w = -0.1$. Our goal is to design a CNF state feedback control law that would yield a good transient performance in tracking a target reference $r = 2$.

Following the procedure given in the previous subsection, we select an integration gain $\kappa_i = 0.3$ and obtain an appropriate augmented system. After a few tries, we found that the following state feedback gain to the augmented system would yield a good performance for our problem:

$$F = [-0.1 \quad -0.85 \quad 2.14 \quad -1.7] \quad (5.250)$$

which places the poles of $\bar{A} + \bar{B}F$ at $0.9, 0.4, 0.5 \pm j0.5$. We note that the first one corresponds to the integrator. Both the linear state feedback control and enhanced CNF control share the same integration dynamics:

$$x_i(k+1) = x_i(k) + 0.3[h(k) - r] \quad (5.251)$$

The linear state feedback control law is given by

$$u(k) = [-0.85 \quad 2.14 \quad -1.7]x(k) - 0.1x_i(k) + 0.41r \quad (5.252)$$

Letting $W = \text{diag}\{0.1, 1, 1, 1\}$, we obtain a positive-definite solution P for Equation 5.202, which is given by

$$P = \begin{bmatrix} 2.6201 & 0.7258 & 0.6543 & 1.1678 \\ 0.7258 & 1.6414 & -1.4825 & 2.1586 \\ 0.6543 & -1.4825 & 8.8061 & -7.3158 \\ 1.1678 & 2.1586 & -7.3158 & 13.3551 \end{bmatrix} \quad (5.253)$$

and a CNF state feedback law:

$$u(k) = [-0.85 \quad 2.14 \quad -1.7]x(k) - 0.1x_i(k) + 0.41r + \rho(e(k)) \\ \times [-0.1678 \quad 2.3536 \quad -9.3268 \quad 10.0458] \left[\begin{pmatrix} x_i(k) \\ x(k) \end{pmatrix} - \begin{pmatrix} 0 \\ 2 \\ 2 \\ 2 \end{pmatrix} \right] \quad (5.254)$$

where ρ is as given in (5.241) with $\alpha = 20$ and $\beta = 1.5$. The simulation results given in Figures 5.4 and 5.5 clearly show that the CNF control has outperformed the linear control.

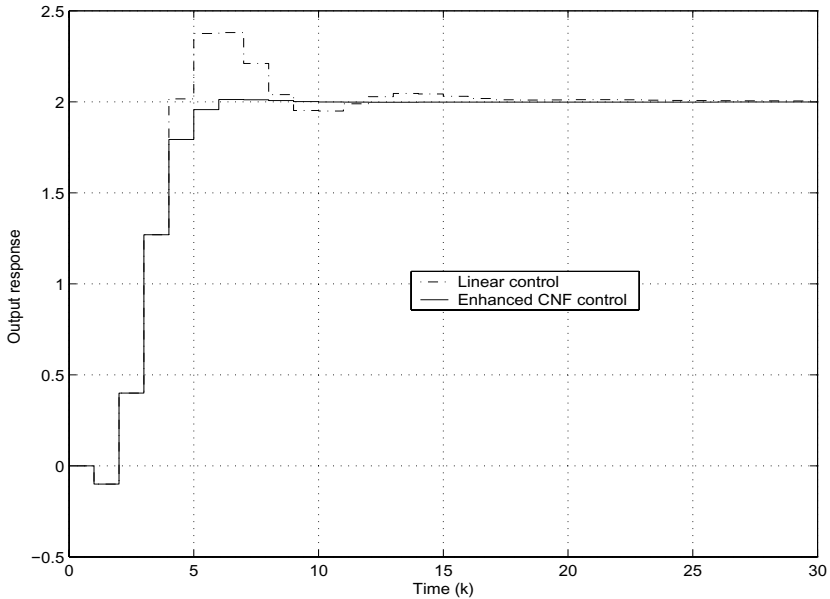


Figure 5.4. Output responses of the enhanced CNF control and linear control

5.4 Can We Beat Time-optimal Control?

So far, we have presented quite a number of control techniques in Chapters 4 and 5 that can be used to design control laws to track certain target references for systems with actuator saturations. The TOC technique is believed to be nonrobust to system uncertainties and noise, and thus cannot be used in tackling real problems. Unfortunately, it has also been regarded as a method that would, at least theoretically, yield the best performance in terms of settling time.

Can we design a control system that would beat the performance of the TOC? Obviously, the answer to this question is no if it is required to have a precise point-to-point tracking, *i.e.* to track a target reference precisely from a given initial point. However, surprisingly, the answer would be yes if we consider an asymptotic tracking situation, *i.e.* if we consider the settling time to be the total time that the controlled system output takes to get from its initial position to reach a predetermined neighborhood of the target reference. The reason that we are interested in this issue is that asymptotic tracking is widely used in almost all practical situations.

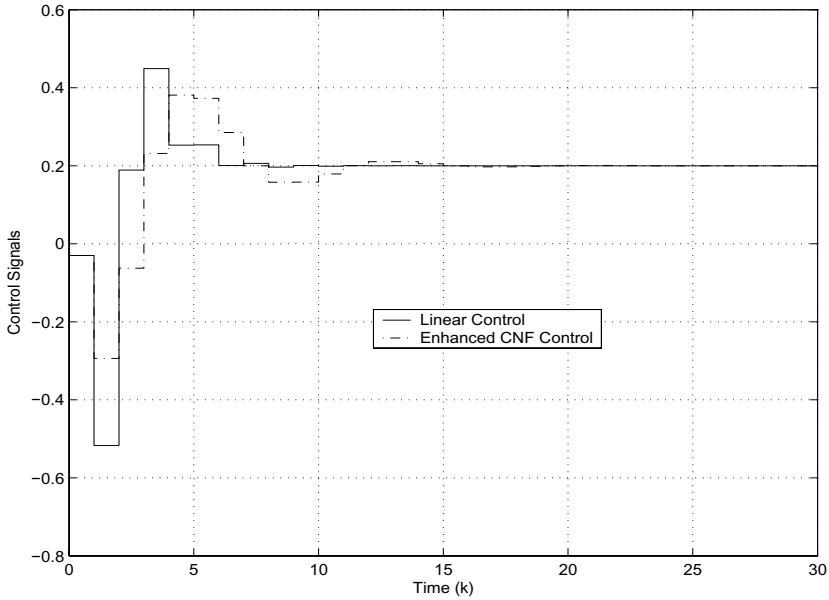


Figure 5.5. Control signals of the enhanced CNF control and linear control

In what follows, we show the above observation in an example. Let us consider a system characterized by a double integrator, *i.e.*

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 1 \end{bmatrix} \text{sat}(u), \quad y = x, \quad h = [1 \quad 0]x \quad (5.255)$$

where as usual x is the state, u is the input, and y and h are, respectively, the measurement and controlled outputs. Moreover, we assume that

$$\text{sat}(u) = \text{sgn}(u) \cdot \min\{1, |u|\} \quad (5.256)$$

Let the initial state $x(0) = 0$ and the target reference $r = 1$. Then, it follows from Equation 4.22 that the minimum time required for the controlled output to reach precisely the target reference under TOC control is exactly 2 s. Let us now consider an asymptotic tracking situation instead. As is commonly accepted in the literature (see, *e.g.*, [86]), we define the settling time to be the total time that it takes for the control output h to enter the $\pm 1\%$ region of the target reference. The following control law, obtained from a variation form of the CNF control technique, would give a faster settling time than that of the TOC,

$$u = [-6.5 \quad -1]x + 6.5r - \left(e^{-|1-h|} - 0.36788 \right) [1.4481 \quad 10.8609] \left(x - \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right) \quad (5.257)$$

Figures 5.6 and 5.7, respectively, show the resulting controlled output responses and the control signals of the TOC and the modified CNF control. The resulting output

response of the modified CNF control has an overshoot of less than 1%. However, if we zoom in on the output responses (see Figure 5.8), we will see that the modified CNF control clearly has a faster settling time than that of the TOC when it enters the target region, *i.e.* $0.99 \leq h \leq 1.01$. It can be computed that the modified CNF control has a settling time of 1.8453 s whereas the TOC has a settling time of 1.8586 s. Although the difference is not much, since we have not tried to optimize the solution of the modified CNF control, it is, however, significant enough to address one interesting issue: *there are control laws that can achieve a faster settling time than that of the TOC in asymptotic tracking situations*. It can also be shown that, no matter how small the target region is, say $1 \pm \varepsilon$ for any small $\varepsilon > 0$, we can always find a suitable control law that beats the TOC in settling time. Nonetheless, we believe that it would be interesting to carry out some further studies in this subject.

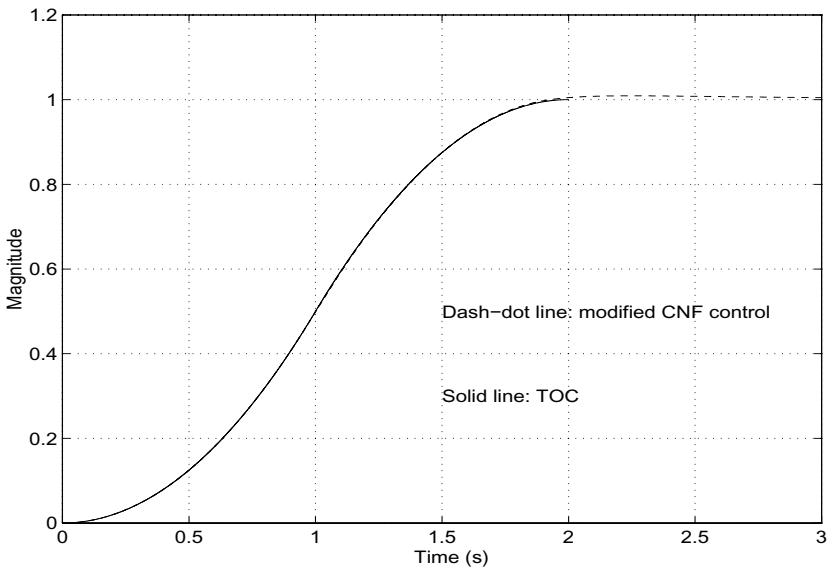


Figure 5.6. Controlled output responses of TOC and modified CNF control

5.5 CNF Control Software Toolkit

The CNF design involves selecting quite a number of design parameters, especially those parameters for forming the nonlinear feedback law. Some tuning and retuning are needed in order to obtain the best possible solutions. In what follows, we present a MATLAB[®] toolkit with a user-friendly graphical interface for the CNF control system design. With the help of the rich collection of m-functions in MATLAB[®]

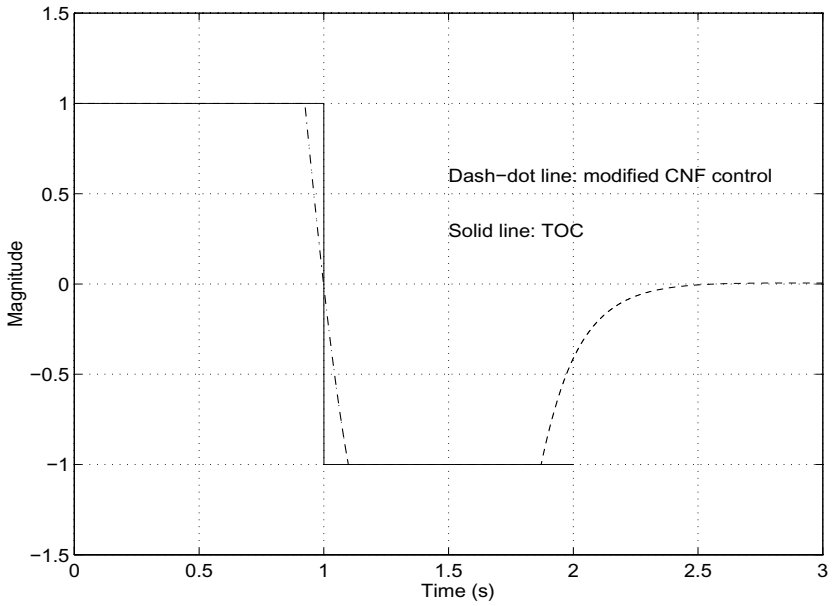


Figure 5.7. Control signals of TOC and modified CNF control

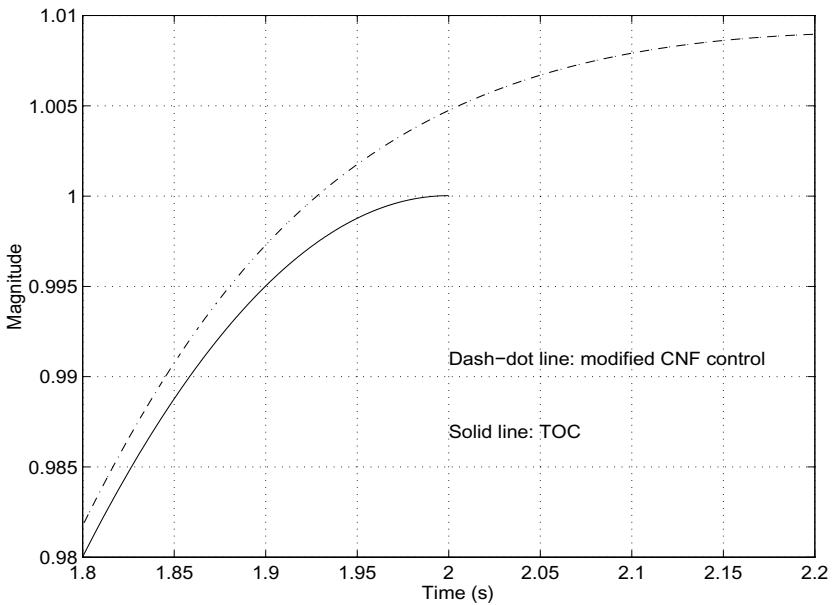


Figure 5.8. Controlled output responses around the target reference

and the powerful simulation capacity of its Simulink[®], we develop a toolkit for the CNF control technique. The toolkit can be utilized to design fast and smooth tracking controllers for general SISO linear systems and a class of nonlinear systems (see, for example, [136] for details) with actuator saturation and other nonlinearities such as friction, as well as external disturbances. The toolkit can display both time-domain and frequency-domain responses on its main panel, and generate three different types of control laws, namely, the state feedback, the full-order measurement feedback and the reduced-order measurement feedback controllers. The usage and design procedure of the toolkit are illustrated by an example on the design of an HDD servo system. The toolkit has been reported earlier in Cheng *et al.* [54, 55]. It can be downloaded at the website <http://hdd.ece.nus.edu.sg/~bmchen>.

5.5.1 Software Framework and User Guide

The CNF control toolkit is developed under MATLAB[®] together with Simulink[®]. The software toolkit fully utilizes the graphical user interface (GUI) resources of MATLAB[®] and provides a user-friendly graphical interface. The main interface of the toolkit consists of three panels, the panel for conducting simulation, the panel for setting up system data and the panel for specifying an appropriate controller. We illustrate the design procedure using our CNF control toolkit in the following:

STAGE 1: INITIALIZATION. Once the toolkit is properly executed, a main panel as shown in Figure 5.9 will be generated in a popup window. Users have to first enter required data for a system to be controlled and then specify an appropriate controller structure before running simulation on this panel.

STAGE 2: PLANT MODEL SETUP. To enter system data, users need to click on the box labeled with `PLANT` to open the plant model setup panel as shown in Figure 5.10. In addition to the state-space model of Equation 5.89, the toolkit also allows users to specify resonance modes of the plant on this panel. We note that high-frequency resonance modes are existent in almost all mechanical systems. Because of the complexity of resonance modes, they are generally ignored or simplified in the controller design stage. However, these resonance modes have to be included in the simulation and evaluation of the overall control system design.

Each time a plant model is keyed in or modified on the panel, the toolkit automatically runs a checkup on the system stabilizability, detectability, invertibility, and other requirements. For a nonlinear system, the toolkit also checks the stability of its nonlinear dynamics. Users will be warned if the solvability conditions for the CNF tracking control are not satisfied and users have to revise the model before proceeding to controller design.

STAGE 3: CONTROLLER SETUP. As the core of this toolkit, the CNF controller design is to be proceeded in a configurable and convenient fashion. A controller setup panel as shown in Figure 5.11 is opened when the user activates the box marked with `CONTROLLER` in the main panel. This panel carries a block diagram for an adjustable controller configuration, which automatically refreshes when the user makes any change or reselection on the controller structure. Users need to decide

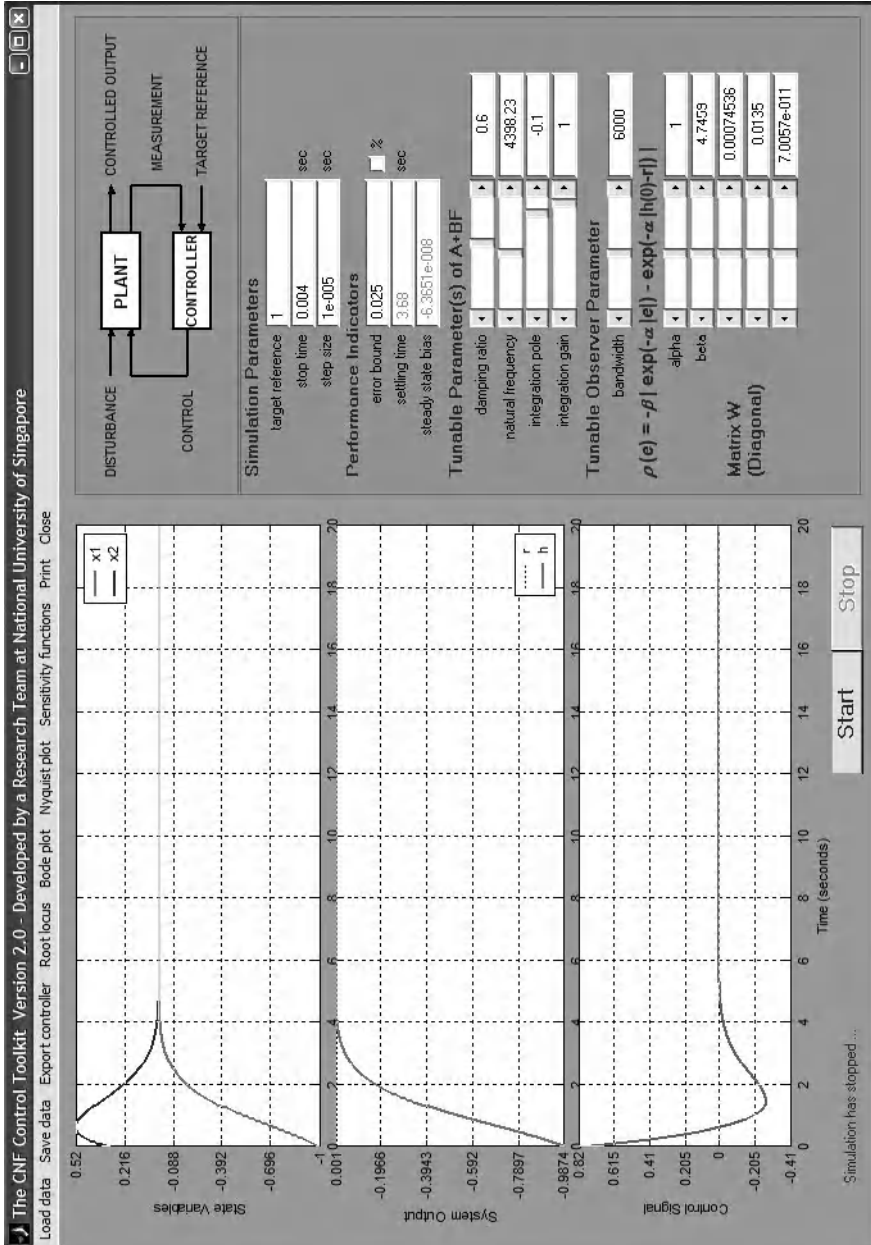


Figure 5.9. The main panel of the CNF control toolkit

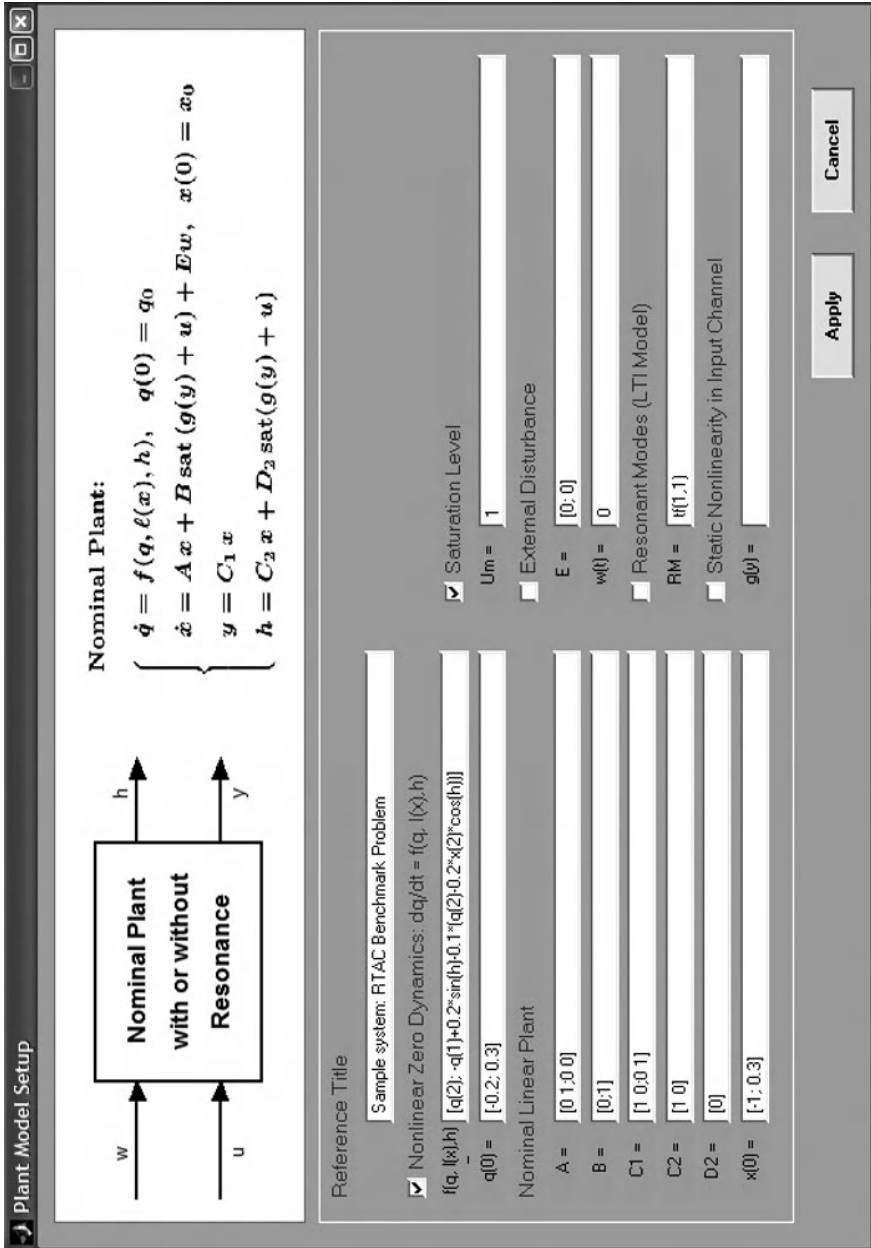


Figure 5.10. The panel for the plant model setup

a controller structure first before proceeding to specify the corresponding controller parameters. The following options are available for the controller in the toolkit:

1. If the given plant in Equation 5.89 has some nonlinearities, users can choose the precompensation option and enter an appropriate nonlinear function to cancel as many nonlinearities as possible.
2. If the given plant in Equation 5.89 has some unknown constant disturbances or other types of disturbances, users can select a controller structure with integrator to remove the steady-state bias.
3. If the given plant in Equation 5.89 is the nominal model of a noisy plant, which has high-frequency resonance modes, users might enter a predesigned lowpass or notch filter to minimize their effects on the overall performance.
4. Based on the properties of the given plant in Equation 5.89 and personal interest, users can then choose a controller with either one of the following options:
 - a) State feedback;
 - b) Full-order measurement feedback;
 - c) Reduced-order measurement feedback.

When it comes to the design of the state feedback gain F , users have three choices. Users can either specify an explicit matrix (obtained through any other design methodologies such as H_∞ control) for F , or employ the H_2 control technique (provided on the panel), or use the pole-placement method characterized by the damping ratio and natural frequency of the dominant poles as well as integration pole and gain (if applicable). For the pole-placement method, the remaining closed-loop poles are placed three times faster than the dominant pair with a Butterworth pattern.

In the case of measurement feedback control, users also have three choices for the design of the observer gain K . Users can directly enter a predesigned solution for K , or design it online using the H_2 control technique, or using the pole-placement method to organize the observer poles into a Butterworth pattern with an appropriate bandwidth. In any case, the corresponding parameters can be further tuned on the main panel to obtain a satisfactory performance for the overall system.

STAGE 4: DESIGN AND SIMULATION. Once the plant model setup and the controller setup are completed, users can then specify directly on the main panel the simulation parameters, such as the setpoint for the target reference, the duration of simulation and the step size. Users can also define the tracking performance indicator and obtain the result for settling time and steady-state bias of the controlled output response. For example, in Figure 5.9, the settling time is defined as the time when the controlled output of the closed-loop system enters the neighborhood of ± 0.025 of the target reference. Alternatively, one can define such a neighborhood in terms of the percentage of final target instead of the absolute error bound.

As mentioned in the previous section, the CNF controller consists of two parts, a linear part and a nonlinear part. On the main panel, users are able to tune the properties of the linear part by selecting appropriate values of the damping ratio and natural frequency of the dominant modes of the linear state feedback dynamical matrix $\bar{A} + \bar{B}F$. The remaining eigenvalues of $\bar{A} + \bar{B}F$ are placed three times faster

CNF Controller Setup

Controller Structure

- Using Nonlinearity Pre-compensation
- Using Integral Compensation
- Using an Additional (Notch) Filter
- State Feedback
- Full Order Measurement Feedback
- Reduced Order Measurement Feedback

Design State Feedback Gain Matrix F using

- A specific F
- H2 control with a preliminary tuning parameter epsilon (tunable)
- Pole placement with preliminary damping ratio and natural frequency (tunable)

Design State Feedback Gain Matrix F using

A specific F

H2 technique with a tuning parameter epsilon (tunable)

Pole placement with a preliminary tuning parameter epsilon (tunable)

Design State Feedback Gain Matrix F using

A specific F

H2 technique with a tuning parameter epsilon (tunable)

Desired bandwidth of observer dynamics in Butterworth pattern (tunable)

The Transfer Function of the Additional (Notch) Filter

Nonlinear Pre-compensation Function $g(y)$

CNF Nonlinear Feedback

The diagonal vector of matrix W (tunable)

Nonlinear function ρ (parameters tunable at simulation panel).

Design Observer Gain Matrix K using

- A specific K
- H2 technique with a tuning parameter epsilon (tunable)
- Desired bandwidth of observer dynamics in Butterworth pattern (tunable)

The Transfer Function of the Additional (Notch) Filter

Nonlinear Pre-compensation Function $g(y)$

Apply Cancel

Figure 5.11. The panel for the CNF controller setup

than the dominant modes with a Butterworth pattern. Such an arrangement is purely for the simplification of the control system design. The pole corresponding to the integration part, if applicable, can be tuned on the main panel as well. Alternatively, users are allowed to apply the H_2 design with a tunable parameter ε , or directly specify a state feedback gain matrix (obtained using any design method) in the controller setup panel without any tunable parameters.

Design parameters W and $\rho(e)$ for the nonlinear part of the CNF controller can also be tuned online on the main panel. In particular, the parameters α and β for the nonlinear function $\rho(e)$ of Equation 5.128 can be easily adjusted using the computer mouse or by directly keying their values onto the spaces provided. In the current version of the CNF control toolkit, the design parameter W is restricted to a diagonal matrix for the simplicity of software implementation.

There are three windows on the main panel for displaying the system state variables, the controlled output response and the control input signal, together with a block diagram showing the structure of the overall control system. Using the right button of the computer mouse to click on the window displaying the state variables, output response and control signal, users are prompted by a small text window showing options to redraw the plots on a new popup window or export the simulation data to the MATLAB[®] workspace.

Finally, the following commands and functions are also implemented on the main panel for saving and loading data as well as for evaluating the frequency-domain properties of the overall control system:

1. Load data: This function is used to load data previously saved in the toolkit.
2. Save data: This function is to save the system and controller data for future use.
3. Export controller: This function is to export the data of the CNF controller obtained to the MATLAB[®] workspace. The controller data are given by

$$\left. \begin{aligned} \dot{z} &= \kappa_i(h - r) \\ \dot{x}_v &= A_{\text{cmp}} x_v + B_{\text{ycmp}} y + B_{\text{ucmp}} \text{sat}(\bar{u}) \\ \dot{\hat{x}} &= C_{\text{cmp}} x_v + D_{\text{cmp}} y \\ \bar{u} &= F \begin{pmatrix} z \\ \hat{x} \end{pmatrix} + G r + \rho(e) F_n \left[\begin{pmatrix} z \\ \hat{x} \end{pmatrix} - G_e r \right] \\ u &= H(s)\bar{u} - g(y) \end{aligned} \right\} \quad (5.258)$$

where A_{cmp} , B_{ycmp} , B_{ucmp} , C_{cmp} , D_{cmp} , F , F_n , G , G_e are constant vectors or matrices, and $\rho(e)$ and $g(y)$ are scalar functions. If a filter is used to reduce the effects of noise or high-frequency resonance modes of a physical plant, $H(s)$ will represent the transfer function of such a filter. All these parameters can be saved under a structured workspace variable (specified by the user) in the command window.

4. Root locus: This function is to generate the root locus of the control system with the CNF controller with respect to the change of the nonlinear function $\rho(e)$.
5. Bode plot: This function is to generate the Bode magnitude and phase responses of the open-loop system comprising the plant and the controller in the steady-

state situation when the nonlinear gain is converging to a constant. This function can be used to evaluate the frequency-domain properties of the control system, such as gain and phase margins.

6. Nyquist plot: Similar to the Bode plot given above. Note that for both the Bode and Nyquist plots, users are allowed to specify a frequency range of interest.
7. Sensitivity functions: This function is to plot the sensitivity and complementary sensitivity functions of the overall design with a prespecified frequency range.
8. Print: This function is to print the items shown on the main panel.
9. Close: This command is to close up the CNF control toolkit.

Although the CNF control toolkit is meant to design a composite nonlinear feedback controller, users have the option to choose only the linear portion of the CNF control law for their design. This option is particularly useful for the comparison of performances of the CNF controller and the linear controller.

5.5.2 An Illustrative Example

We illustrate the CNF control toolkit using a practical example on a micro hard disk drive (HDD) servo system design. The following dynamical model of an IBM microdrive (DMDM-10340) has been recently reported in Peng *et al.* [138] and will be further studied in Chapter 9:

$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 2.35 \times 10^8 \end{bmatrix} [\text{sat}(g(y) + u) + w] \\ y = h = [1 \quad 0] x \end{cases} \quad (5.259)$$

where the state variable x , consists of the displacement and velocity of the read/write (R/W) head of the microdrive; y and h are simply the displacement (in μm) of the R/W head of the drive; w represents friction and some unknown disturbances of the system, and is set to be -1 mV for this demonstration; u is the control input, limited by ± 3 V; and finally the nonlinearity $g(y) = -0.02887 \text{atan}(0.5886y)$ is generated by the data flex cable in the drive. The high-frequency resonance modes of the microdrive as given in Chapter 9 are included in simulation.

For hard disk drive servo systems, the task of the controller is to move the actuator R/W head to a desired track as fast as possible and to maintain the R/W head as close as possible to the target track center when data are being read or written. To ensure reliable data reading and writing, it is required that the deviation of the R/W head from the target track center should not exceed 5% of the track pitch. In this demonstration, we focus our attention on designing a servo system that yields an optimal performance for short-distance tracking. In particular, we set the target reference $r = 1 \mu\text{m}$. Thus, the settling time is defined as the total duration for the R/W head to enter the $\pm 0.05 \mu\text{m}$ region of the target track.

Using the CNF control toolkit of [55] with few online adjustments on the design parameters, we manage to obtain a reduced-order CNF controller that yields a good performance. The dynamics equations of the controller is given by:

$$\begin{pmatrix} \dot{x}_i \\ \dot{x}_v \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -6000 \end{bmatrix} \begin{pmatrix} x_i \\ x_v \end{pmatrix} + \begin{bmatrix} 0 \\ 2.35 \times 10^8 \end{bmatrix} \text{sat}(u) - \begin{bmatrix} -10 \\ 3.6 \times 10^7 \end{bmatrix} y - \begin{bmatrix} 10 \\ 0 \end{bmatrix} r \quad (5.260)$$

$$u = \rho(e) [0.02398 \quad 0.02646 \quad 2.5472 \times 10^{-5}] \begin{pmatrix} x_i \\ y - r \\ x_v + 6000y \end{pmatrix} - [2.0851 \quad 0.20911 \quad 6.3830 \times 10^{-6}] \begin{pmatrix} x_i \\ y - r \\ x_v + 6000y \end{pmatrix} - g(y), \quad (5.261)$$

with $\rho(e) = -4.4 |e^{-4|e|} - e^{-4}|$. Figure 5.12 shows the controlled output response of the servo system with the obtained CNF controller. The resulting 2.5% settling time is about 0.27 ms. The frequency-domain properties of the CNF control servo system are, respectively, shown in Figures 5.13 to 5.15.

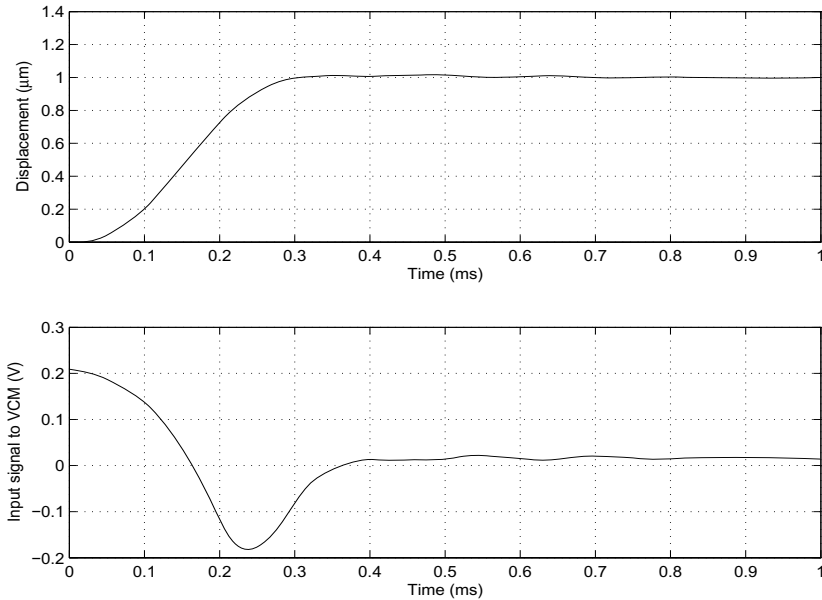
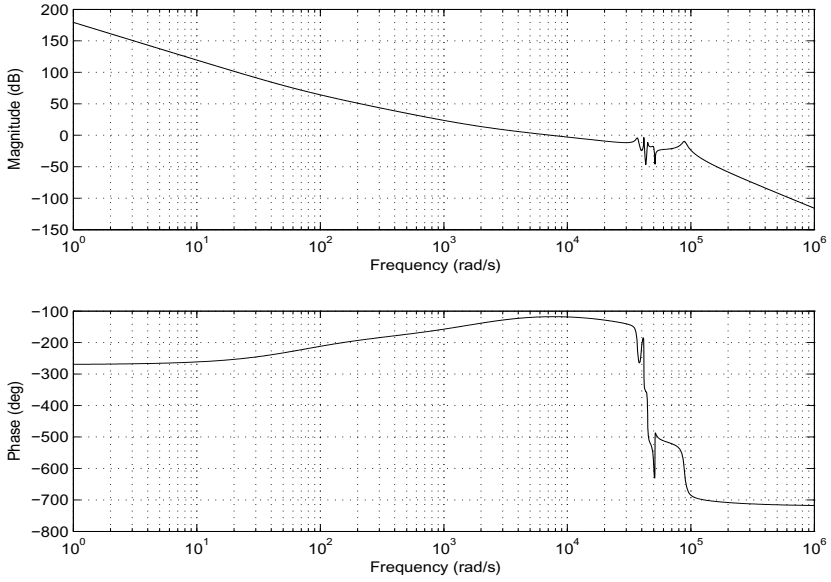
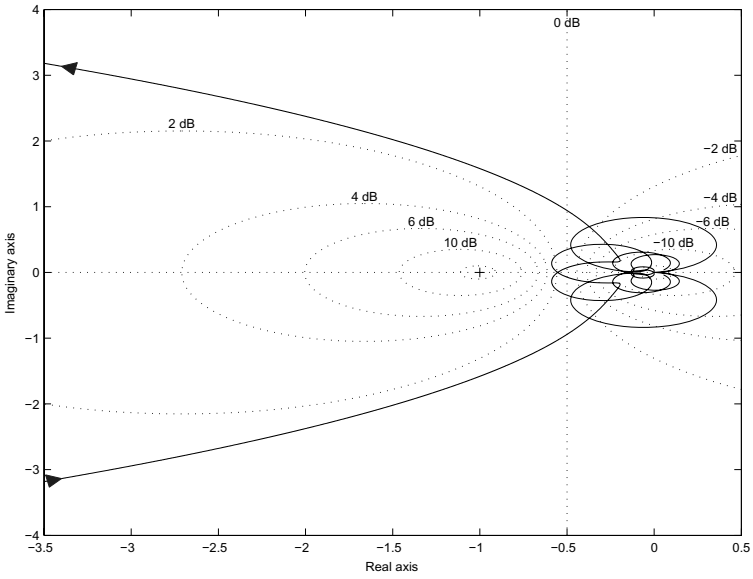


Figure 5.12. Output response and control signal of the CNF control system



(a) Bode plot



(b) Nyquist plot

Figure 5.13. Bode and Nyquist plots of the CNF control system

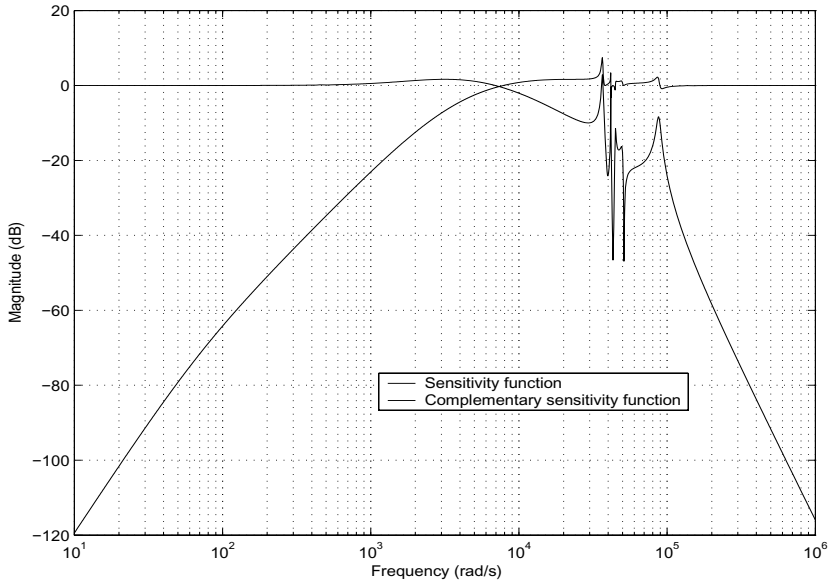


Figure 5.14. Sensitivity and complementary sensitivity functions of the CNF control system

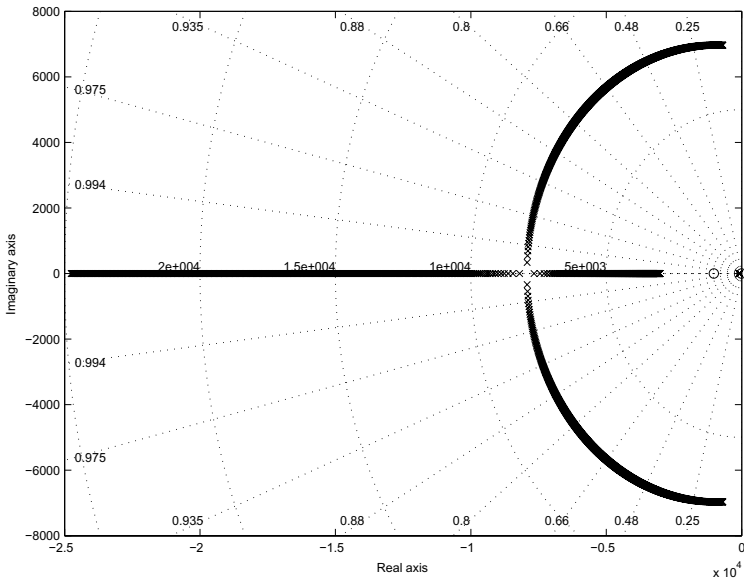


Figure 5.15. Root locus of the closed-loop system versus the nonlinear function $\rho(e)$

HDD Servo Systems Design

Track Following of a Single-stage Actuator

6.1 Introduction

The prevalent trend in hard disk design is towards smaller hard disks with increasingly larger capacities. This implies that the track width has to be smaller, leading to lower error tolerance in the positioning of the head. The controller for track following has to achieve tighter regulation in the control of the servomechanism. Current HDDs use a combination of classical control techniques, such as lead-lag compensators, PID compensators, and notch filters. These classical methods can no longer meet the demand for HDDs of higher performance. Thus, many control approaches have been tried, such as LQG and/or LTR approach (see, *e.g.*, [18, 19]), and adaptive control (see, *e.g.*, [28]) and so on.

The purpose of this chapter is to use the results of the RPT control method of Chapter 3 and the CNF control technique of Chapter 5 to carry out the design of track-following controllers for an HDD with a single VCM actuator. We first obtain a model of the VCM actuator and then cast the overall track-following control system design into an RPT design framework. A second-order dynamic measurement feedback controller is then designed to achieve robust and perfect tracking for any step reference. Our controller is theoretically capable of making the L_p -norm of the resulting tracking error arbitrarily small in faces of external disturbances and initial conditions. Some tradeoffs are then made in order for the RPT controller to be implementable using the existing hardware setup and to meet physical constraints such as sampling rates and the limit of control of the system. The CNF control technique and the classical PID control technique is also employed to design track-following controllers for the same drive. The simulation and implementation results of these controllers are done in both time domain and frequency domain and compared. The result on the RPT design presented in this chapter are rooted in earlier work reported in [3] (see also [74]). However, it is obtained using a new HDD, namely a Maxtor (Model 51536U3) HDD. Furthermore, we introduce a notch filter to minimize the effect of the resonance modes and follow the development of a recent work by Cheng *et al.* [140] to add an additional integrator to the RPT controller. The new design is capable of eliminating steady-state bias.

6.2 VCM Actuator Model

In this section, we present the modeling of the VCM actuator that is well known in the research community of the HDD servo systems to have a characteristic of a double integrator cascaded with some high-frequency resonance, which can reduce the system stability if neglected. There are some bias forces in the HDD system that cause steady-state errors in tracking performance. Moreover, there are also some nonlinearities in the system at low frequencies, which are primarily due to the pivot-bearing friction. All these factors have to be taken into consideration when considering the design of a controller for the VCM actuator. For the purpose of developing a model, we have to compromise between accuracy and simplicity. In this section, a relatively simplified model of the VCM actuator is identified and presented.

The dynamics of an ideal VCM actuator can be formulated as a second-order state-space model as follows:

$$\begin{pmatrix} \dot{y} \\ \dot{v} \end{pmatrix} = \begin{bmatrix} 0 & k_y \\ 0 & 0 \end{bmatrix} \begin{pmatrix} y \\ v \end{pmatrix} + \begin{pmatrix} 0 \\ k_v \end{pmatrix} u \quad (6.1)$$

where u is the actuator input (in volts), y and v are the position (in tracks) and the velocity of the R/W head, k_y is the position measurement gain and $k_v = k_t/m$, with k_t being the current–force conversion coefficient and m being the mass of the VCM actuator. Thus, the transfer function of an ideal VCM actuator model appears to be a double integrator, *i.e.*

$$G_{v1}(s) = \frac{k_v k_y}{s^2} \quad (6.2)$$

However, if we also consider the high-frequency resonance modes, a more realistic model for the VCM actuator will be

$$G_v(s) = \frac{k_v k_y}{s^2} \prod_{i=1}^N G_{r,i}(s) \quad (6.3)$$

where N is the number of significant resonance modes in the frequency range of interest, and $G_{r,i}(s)$, $i = 1, 2, \dots, N$, are the transfer functions for the resonance modes. The frequency characteristics of the Maxtor (Model 51536U3) HDD have been obtained using an LDV and an HP Dynamic Signal Analyzer. The actual frequency response is shown in Figure 6.1. Applying the least squares estimation identification method given in Chapter 2 (see also [13, 59]) to the measured data from the actual system, we obtain a tenth-order model for the actuator:

$$G_v(s) = \frac{6.4013 \times 10^7}{s^2} \prod_{i=1}^4 G_{r,i}(s) \quad (6.4)$$

with

$$G_{r,1}(s) = \frac{0.912s^2 + 457.4s + 1.433 \times 10^8}{s^2 + 359.2s + 1.433 \times 10^8} \quad (6.5)$$

$$G_{r,2}(s) = \frac{0.7586s^2 + 962.2s + 2.491 \times 10^8}{s^2 + 789.1s + 2.491 \times 10^8} \quad (6.6)$$

$$G_{r,3}(s) = \frac{9.917 \times 10^8}{s^2 + 1575s + 9.917 \times 10^8} \quad (6.7)$$

and

$$G_{r,4}(s) = \frac{2.731 \times 10^9}{s^2 + 2613s + 2.731 \times 10^9} \quad (6.8)$$

Figure 6.1 shows that the frequency response of the identified model matches the measured data very well for the frequency range from 0 to 1.8 kHz, which far exceeds the working range of the VCM actuator.

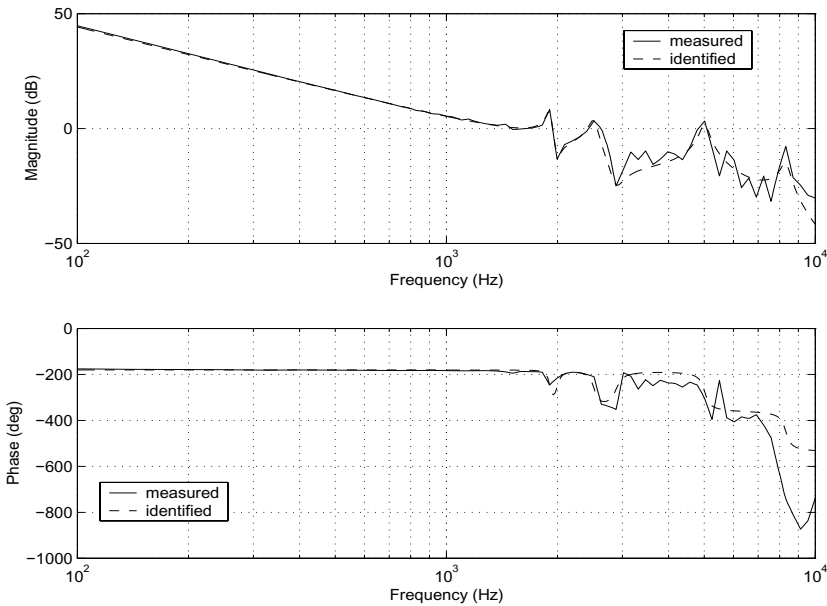


Figure 6.1. Frequency responses of the actual and identified VCM actuator models

6.3 Track-following Controller Design

We now present the control system design for the actuator identified in the previous section. Basically, the majority of commercially available HDD servo systems to date are designed using a conventional PID approach. For drives with a single VCM actuator, designers would encounter problems if they wished to push up the track-following speed. Usually, there are some huge overshoot peaks in step response. Thus, in practice, one would have to make tradeoffs between the track-following

speed and the overshoot by selecting appropriate PID controller gains. These drawbacks can be overcome by using our newly developed control techniques.

We design an HDD servo system that meets the following design constraints and specifications:

1. the control input does not exceed ± 3 V owing to physical constraints on the actual VCM actuator;
2. the overshoot and undershoot of the step response are kept to less than 5% as the R/W head can start to read or write within $\pm 5\%$ of the target;
3. the 5% settling time in the step response is as short as possible;
4. the gain margin and phase margin of the overall design are, respectively, greater than 6 dB and 30° ;
5. the maximum peaks of the sensitivity and complementary sensitivity functions are less than 6 dB; and
6. the sampling frequency in implementing the actual controller is 20 kHz.

In order to minimize the effect of the high-frequency resonance modes, we add a notch filter to the plant to cancel as much as possible of the unwanted responses. For the plant considered in this chapter, we introduce the following notch filter

$$G_{\text{notch}}(s) = \left(\frac{s^2 + 238.8s + 1.425 \times 10^8}{s^2 + 2388s + 1.425 \times 10^8} \right) \times \left(\frac{s^2 + 314.2s + 2.467 \times 10^8}{s^2 + 3142s + 2.467 \times 10^8} \right) \times \left(\frac{s^2 + 628.3s + 9.87 \times 10^8}{s^2 + 12570s + 9.87 \times 10^8} \right) \quad (6.9)$$

whose frequency response is given in Figure 6.2. The overall response of the plant together with the notch filter is given in Figure 6.3, which shows that the effect of the first two resonance modes are indeed reduced. This same notch filter will be used in Chapters 7 and 8 as well.

Thus, we only consider a second-order model for the VCM actuator at this stage. We will then put the resonance modes back when we are to evaluate the performance of the overall design. Thus, in our design, we first use the following simplified model of the VCM actuator:

$$\dot{x} = Ax + Bu = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 6.4013 \times 10^7 \end{bmatrix} u \quad (6.10)$$

and

$$y = C_1 x = [1 \quad 0] x \quad (6.11)$$

Next, we define the output to be controlled as

$$h = y = C_2 x + D_2 u = [1 \quad 0] x \quad (6.12)$$

The overall control system is depicted as in Figure 6.4.

RPT Control System Design. Consider the reference $r(t)$ to be a step function with a magnitude α , *i.e.* $r(t) = \alpha \cdot \mathbf{1}(t)$, where $\mathbf{1}(t)$ is a unit step function. Then, we have

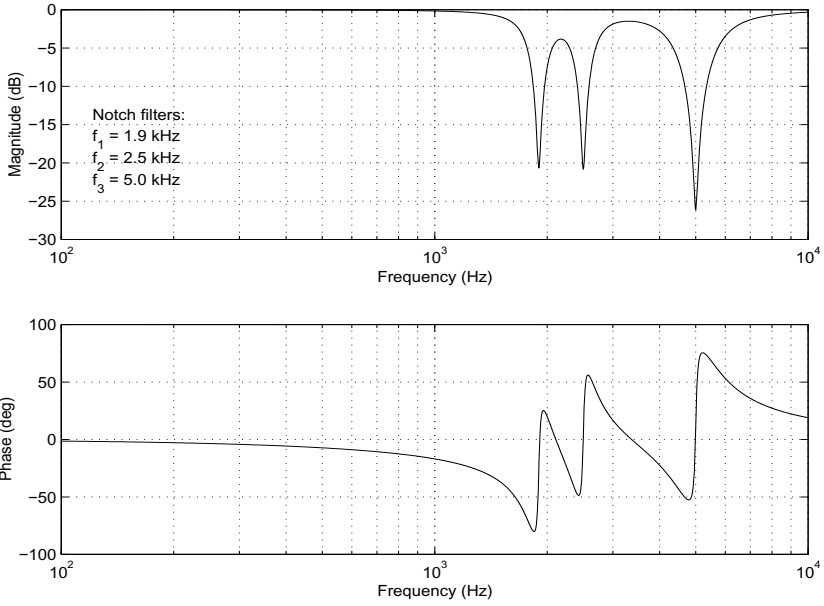


Figure 6.2. Frequency responses of the notch filter

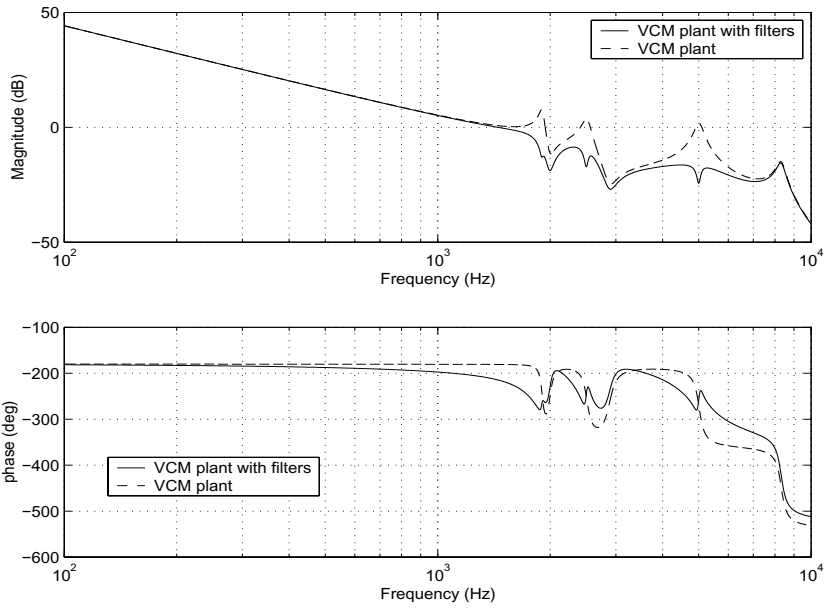


Figure 6.3. Frequency responses of the plant with the notch filter

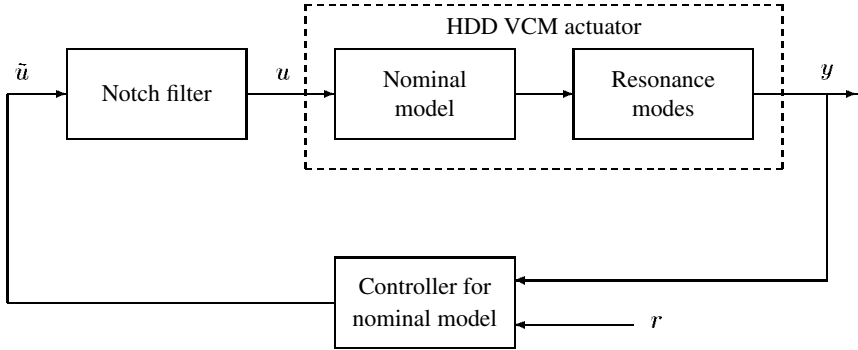


Figure 6.4. Control system configuration for the HDD VCM actuator

$$\dot{r}(t) = \alpha \cdot \delta(t) \tag{6.13}$$

where $\delta(t)$ is a unit impulse function. Following the results of Chapter 3, we obtain a corresponding auxiliary system:

$$\begin{cases} \dot{\mathbf{x}} = \begin{bmatrix} 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ 6.4013 \times 10^7 \end{bmatrix} u + \begin{bmatrix} 0 \\ 1 \\ 0 \\ 0 \end{bmatrix} \mathbf{w} \\ \mathbf{y} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \mathbf{x} + \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \mathbf{w} \\ e = [0 \quad -1 \quad 1 \quad 0] \mathbf{x} + 0 \quad u \end{cases} \tag{6.14}$$

where

$$\mathbf{x} = \begin{pmatrix} \int e \\ r \\ x \end{pmatrix}, \quad \mathbf{w} = \alpha \cdot \delta(t), \quad \mathbf{y} = \begin{pmatrix} \int e \\ r \\ y \end{pmatrix}, \quad e = h - r \tag{6.15}$$

We note that the additional integrator added is to compensate the steady-state bias in the implementation. It is simple to see that (A, B, C_2, D_2) is invertible and free of invariant zeros, and $\text{Ker}(C_1) = \text{Ker}(C_2)$. Hence, it follows from the result of Chapter 3 that the RPT performance is achievable. Actually, there exists a family of measurement feedback control laws, parameterized by a tuning parameter ε , such that when it is applied to the given VCM actuator:

1. the resulting closed-loop system is asymptotically stable for sufficiently small ε ;
2. for any given initial condition x_0 and any $p \in [1, \infty)$, the l_p -norm of the resulting tracking error e has the property $\|e\|_p \rightarrow 0$, as $\varepsilon \rightarrow 0$.

Following the procedure for the reduced-order RPT controller given in Chapter 3, we obtain a parameterized measurement feedback control law of the form

$$\dot{x}_v = A_{rc}(\varepsilon)x_v + B_{rc}(\varepsilon) \begin{pmatrix} r \\ y \end{pmatrix}, \quad \tilde{u} = C_{rc}(\varepsilon)x_v + D_{rc}(\varepsilon) \begin{pmatrix} r \\ y \end{pmatrix} \quad (6.16)$$

with

$$\left. \begin{aligned} A_{rc}(\varepsilon) &= \begin{bmatrix} 0 & 0 \\ -696.4/\varepsilon^3 & -7229.5/\varepsilon \end{bmatrix} \\ B_{rc}(\varepsilon) &= \begin{bmatrix} -1 & 1 \\ 69.962/\varepsilon^2 & -359.14/\varepsilon^2 \end{bmatrix} \\ C_{rc}(\varepsilon) &= [-1.0879/\varepsilon^3 \quad -5.0451/\varepsilon] \\ D_{rc}(\varepsilon) &= [0.10929/\varepsilon^2 \quad -0.3111/\varepsilon^2] \end{aligned} \right\} \quad (6.17)$$

The actual control input to the HDD system is then given by

$$u = G_{\text{notch}}(s)\tilde{u} \quad (6.18)$$

where $G_{\text{notch}}(s)$ is as given in Equation 6.9. Figure 6.5 clearly shows that the RPT problem is solved as we tune the tuning parameter ε to be smaller and smaller. Unfortunately, owing to the constraints of the physical system, *i.e.* the limits in control inputs and sampling rates, as well as resonance modes, it is impossible to implement a controller that tracks the reference with zero time. We would thus have to make some compromises in the track-following speed because of these limitations. After several trials, we found that the controller parameters of Equation 6.17 with $\varepsilon = 1$ would give us a satisfactory performance.

We then discretize it using a ZOH transformation with a sampling frequency of 20 kHz. The discretized controller is given by

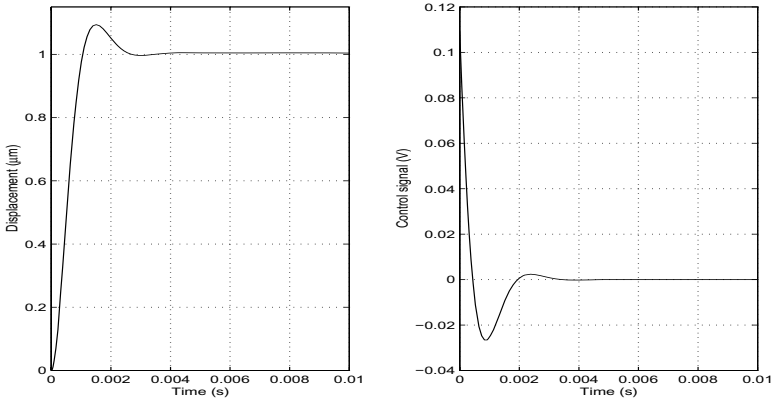
$$\begin{cases} x_v(k+1) = \begin{bmatrix} 1 & 0 \\ -0.14611 & 0.69665 \end{bmatrix} x_v(k) + \begin{bmatrix} -1 & 1 \\ 293.64 & -1507 \end{bmatrix} \times 10^{-5} \begin{pmatrix} r(k) \\ y(k) \end{pmatrix} \\ \tilde{u}(k) = [-5.4395 \quad -5.0451] x_v(k) + [0.10929 \quad -0.3111] \begin{pmatrix} r(k) \\ y(k) \end{pmatrix} \end{cases} \quad (6.19)$$

and

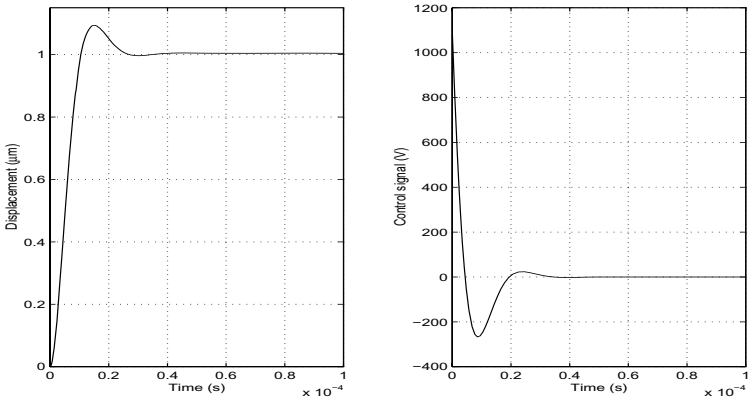
$$u(k) = G_{\text{notch}}(s) \Big|_{\text{ZOH}} \cdot \tilde{u}(k) \quad (6.20)$$

CNF Control System Design. Following the procedure given in Chapter 5 for the CNF design for systems with disturbances, in which an integrator is added to the controller to eliminate steady-state bias, we choose an integration gain $\kappa_i = 1$ and obtain a corresponding augmented plant:

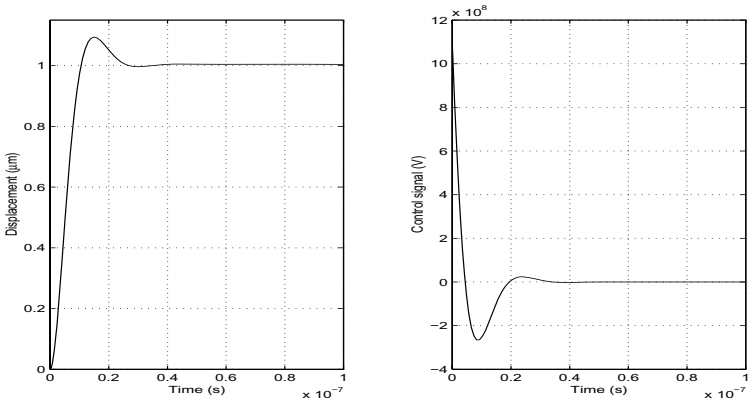
$$\begin{cases} \dot{\bar{x}} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix} \bar{x} + \begin{bmatrix} 0 \\ 0 \\ 6.4013 \times 10^7 \end{bmatrix} \text{sat}(u) + \begin{bmatrix} -1 \\ 0 \\ 0 \end{bmatrix} r \\ y = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \bar{x} \\ h = [0 \quad 1 \quad 0] \bar{x} \end{cases} \quad (6.21)$$



(a) $\varepsilon = 1$



(b) $\varepsilon = 0.01$



(c) $\varepsilon = 10^{-5}$

Figure 6.5. Responses of the closed-loop systems with RPT controller

For track following, $r = 1$. Following a few simulation tries, we obtain a state feedback gain matrix,

$$F = -[9.8676 \times 10^{-5} \quad 0.098676 \quad 2.3557 \times 10^{-5}] \quad (6.22)$$

which places the closed-loop poles at $-801.11 \pm j2547.4$ and -0.001 . Next, we choose W to be a diagonal matrix with diagonal elements being 1.95×10^{-4} , 0.0195 and 3×10^{-10} , respectively. Solving the Lyapunov equation in Equation 5.98, we obtain

$$P = \begin{bmatrix} 0.09751 & 2.3284 \times 10^{-5} & 1.5436 \times 10^{-8} \\ 2.3284 \times 10^{-5} & 9.4272 \times 10^{-6} & 1.5473 \times 10^{-9} \\ 1.5436 \times 10^{-8} & 1.5473 \times 10^{-9} & 1.1255 \times 10^{-12} \end{bmatrix} > 0 \quad (6.23)$$

The reduced-order observer gain matrix is selected as $K_r = -4000$, which places the observer pole at -4000 , and the nonlinear gain function is selected as follows:

$$\rho(e) = -1.5029 \left| e^{-|e|} - 0.3679 \right| \quad (6.24)$$

Finally, the reduced-order enhanced CNF control law for the servo system is given by

$$\begin{pmatrix} \dot{x}_i \\ \dot{x}_v \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -4000 \end{bmatrix} \begin{pmatrix} x_i \\ x_v \end{pmatrix} - \begin{bmatrix} -1 \\ 1.6 \times 10^7 \end{bmatrix} y + \begin{bmatrix} 0 \\ 6.4013 \times 10^7 \end{bmatrix} \tilde{u} - \begin{pmatrix} r \\ 0 \end{pmatrix} \quad (6.25)$$

$$\begin{aligned} \tilde{u} = & \rho(e) [0.98808 \quad 0.099044 \quad 7.2048 \times 10^{-5}] \begin{pmatrix} x_i \\ y - r \\ x_v + 4000y \end{pmatrix} \\ & - [9.8676 \times 10^{-5} \quad 0.098676 \quad 2.3557 \times 10^{-5}] \begin{pmatrix} x_i \\ y - r \\ x_v + 4000y \end{pmatrix} \end{aligned} \quad (6.26)$$

and

$$u = G_{\text{notch}}(s) \cdot \tilde{u} \quad (6.27)$$

The above CNF controller is discretized using the ZOH method with a sampling frequency of 20 kHz when implemented onto the actual HDD hardware.

PID Control System Design. Following the result of Chapter 3, we obtain a finely tuned PID controller, which has the following transfer function:

$$\tilde{u}(k) = \frac{0.50495z^2 - 0.99147z + 0.48653}{z^2 - 1.09091z + 0.09091} \cdot \frac{0.2696}{z - 0.7304} \cdot [r(k) - y(k)] \quad (6.28)$$

where the second term is a low-pass filter added to improve the implementation performance of the control system. In our experimental tests, we find that it is necessary to add the low-pass filter to the PID control law in order to obtain a meaningful

experimental performance. The above discrete-time PID controller is obtained from a continuous-time counterpart using the ZOH method with a sampling frequency of 20 kHz. Once again, we note that the above PID controller is tuned to meet the requirements on the gain and phase margins, and the design specifications on the sensitivity and complementary sensitivity functions. Although the PID control has the simplest structure, its dynamical order, which is 3, is higher than that of the RPT and CNF controllers. As expected, the complete control input is given by

$$u(k) = G_{\text{notch}}(s) \Big|_{\text{ZOH}} \cdot \tilde{u}(k) \quad (6.29)$$

6.4 Simulation and Implementation Results

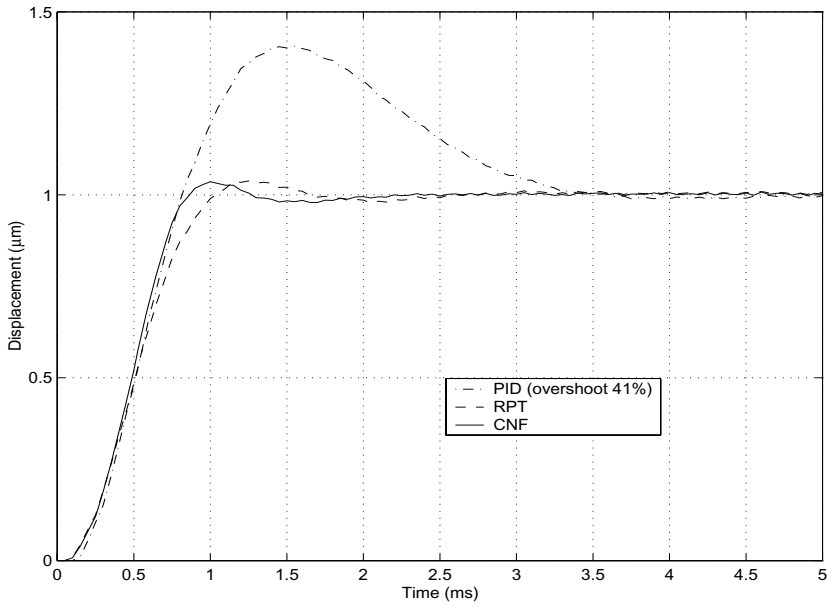
In this section we present the simulation and actual implementation results of our designs and their comparison. The following tests are presented: i) the track-following test of the closed-loop systems, ii) the frequency-domain test including the Bode and Nyquist plots as well as the plots of the resulting sensitivity and complementary sensitivity functions, iii) the runout disturbance test, and lastly iv) the PES test. Our controller was implemented on an open HDD with a sampling rate of 20 kHz. Closed-loop actuation tests were performed using an LDV to measure the R/W head position. The resolution used for LDV was 2 $\mu\text{m/V}$. This displacement output is then fed into the DSP, which would then generate the necessary control signal to the VCM actuator. The actual implementation setup is as depicted in Figure 1.7.

6.4.1 Track-following Test

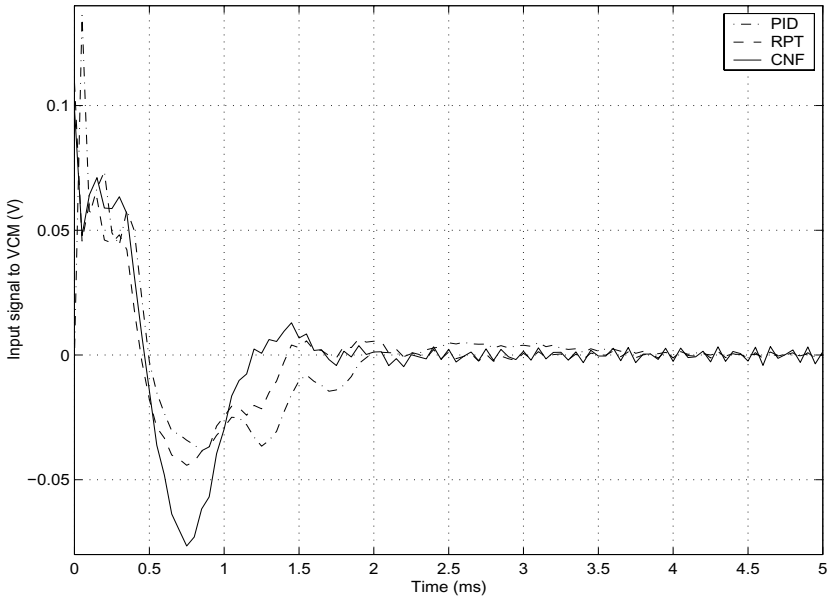
The simulation result and actual implementation result of the closed-loop responses for the control systems are, respectively, shown in Figures 6.6 and 6.7. It is noted that the PID control generates large overshoots in both simulation and implementation, while the systems with the RPT and CNF control have very little overshoot. We summarize the resulting 5% settling time, which is commonly used in the HDD research community, in Table 6.1. Clearly, the CNF control gives the best performance in the time domain compared to those of the other two systems.

Table 6.1. Performances of the track-following controllers

	Settling time (ms)		
	PID control	RPT control	CNF control
Simulation	3.10	0.95	0.80
Implementation	2.65	1.05	0.85

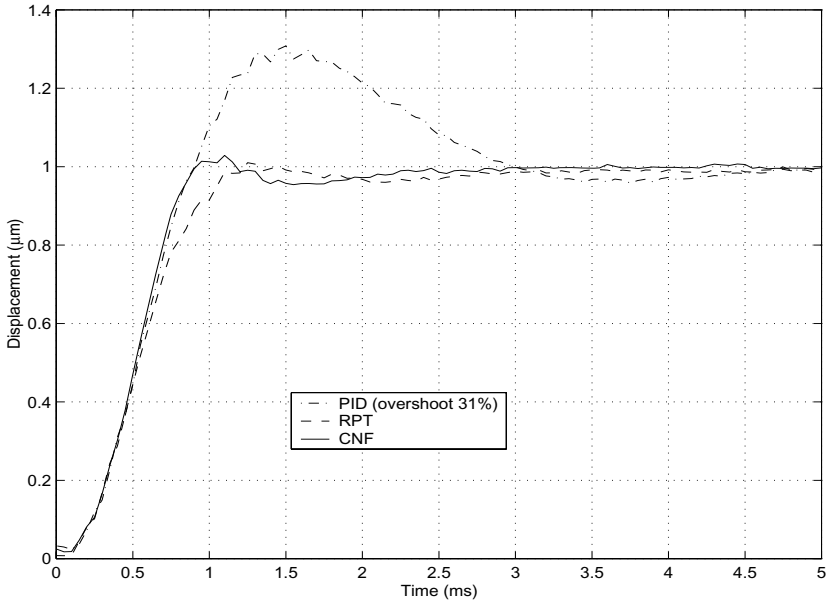


(a) Output responses

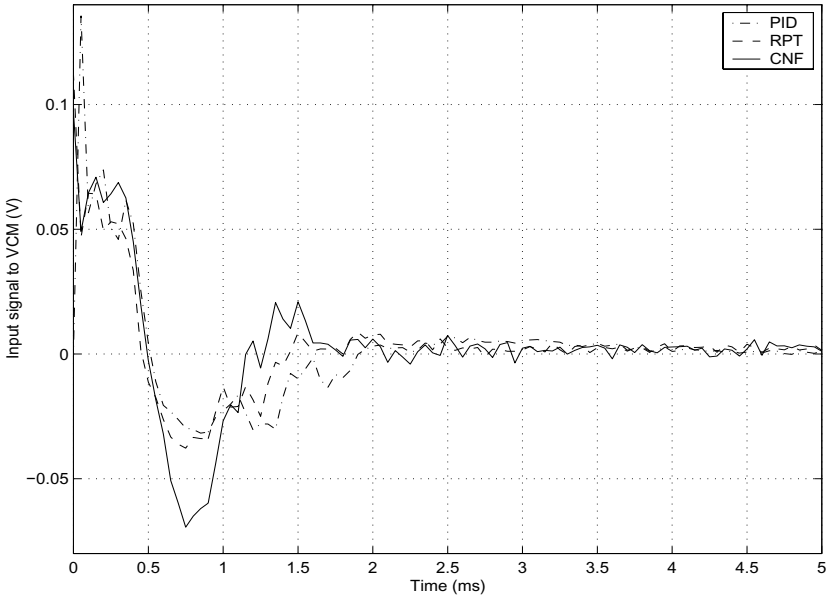


(b) Control signals

Figure 6.6. Simulation result: step responses with PID, RPT and CNF control



(a) Output responses



(b) Control signals

Figure 6.7. Implementation result: step responses with PID, RPT and CNF control

We believe that the shortcoming of the PID control is mainly due to its structure, *i.e.* it only feeds in the error signal, $y - r$, instead of feeding in both y and r independently. We trust that the same problem might be present in other control methods if the only signal fed is $y - r$. The PID control structure might well be as simple as most researchers and engineers have claimed. However, the RPT controller is even simpler, but it and the CNF controller have fully utilized all available information associated with the actual system.

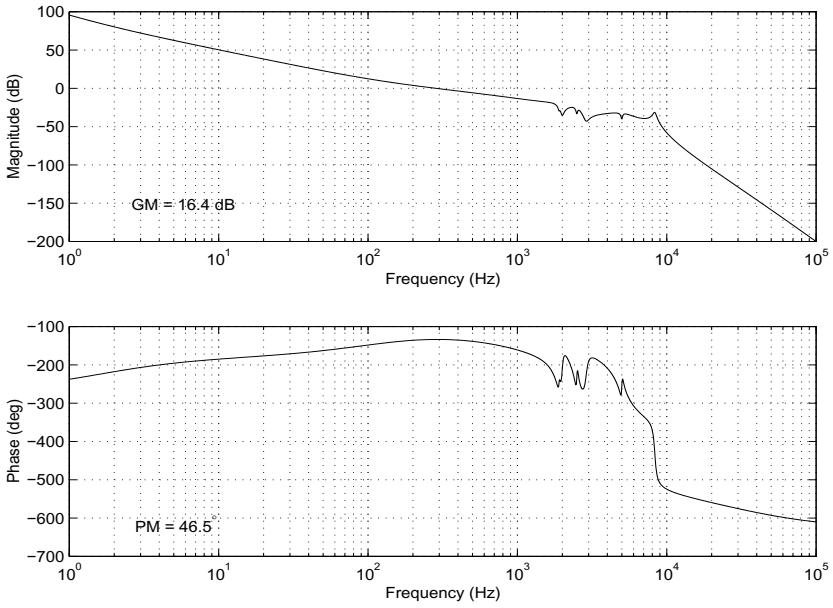
Unfortunately, we could not compare our results with those in the literature. Most of the references we found in the open literature contained only simulation results in this regard. Some of the implementation results we found were, however, very different in nature. For example, Hanselmann and Engelke [18] reported an implementation result of a disk drive control system design using the LQG approach with a sampling frequency of 34 kHz. The overall step response in [18] with a higher-order LQG controller and higher sampling frequency is worse than that of ours.

6.4.2 Frequency-domain Test

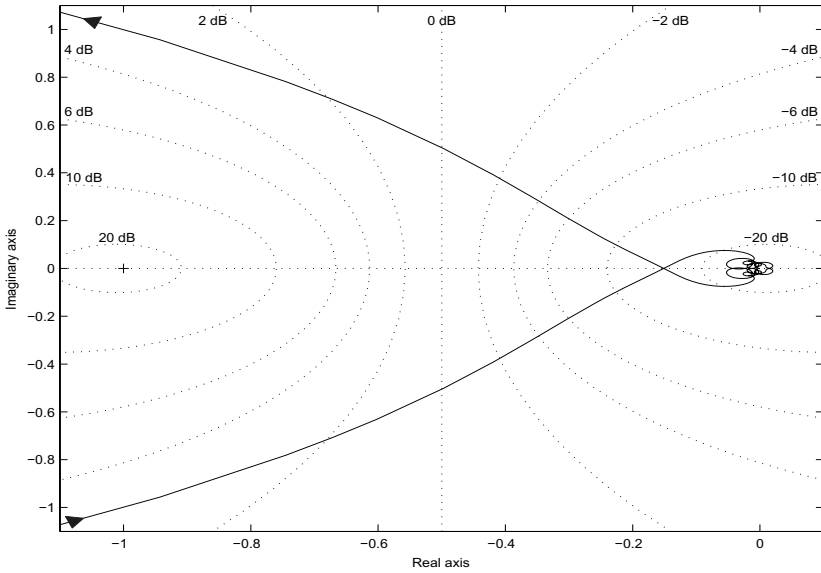
For practical consideration, it is important and necessary to examine the frequency-domain properties of control system design, which include the results of gain and phase margins and the plots of sensitivity and complementary sensitivity functions. Traditionally, gain and phase margins can be obtained through the Bode plot of the open-loop transfer function comprising the given plant and the controller. However, for the HDD system considered in our design, which has additional high-frequency resonance modes, the corresponding Bode plots might have more than one gain and/or phase crossover frequencies. Thus, it is important to verify the stability margins obtained from the associated Nyquist plots. Figures 6.8 to 6.13, respectively, show the Bode plot, the Nyquist plot, and the sensitivity and complementary sensitivity functions, as well as the closed-loop transfer functions (from the reference input r to the controlled output $h = y$) of the resulting control systems. For the CNF design, which is a nonlinear controller, its frequency-domain functions are calculated at the steady-state situation for which the nonlinear gain function $\rho(e)$ has approached its final constant value. The results show that all these designs meet the frequency-domain specifications and have about the same closed-loop bandwidth.

6.4.3 Runout Disturbance Test

Although we do not consider the effects of runout disturbances in our problem formulation, it turns out that our controllers are capable of rejecting the repeatable runout disturbances, which are mainly due to the imperfectness of the data tracks and the spindle motor speeds, and commonly have frequencies at the multiples of the spindle speed, which is about 55 Hz. We simulate these runout effects by injecting a sinusoidal signal into the measurement output, *i.e.* the new measurement output is the sum of the actuator output and the runout disturbance. Figure 6.14 shows the implementation result of the output responses of the overall control system comprising the tenth-order model of the VCM actuator model and the controllers together

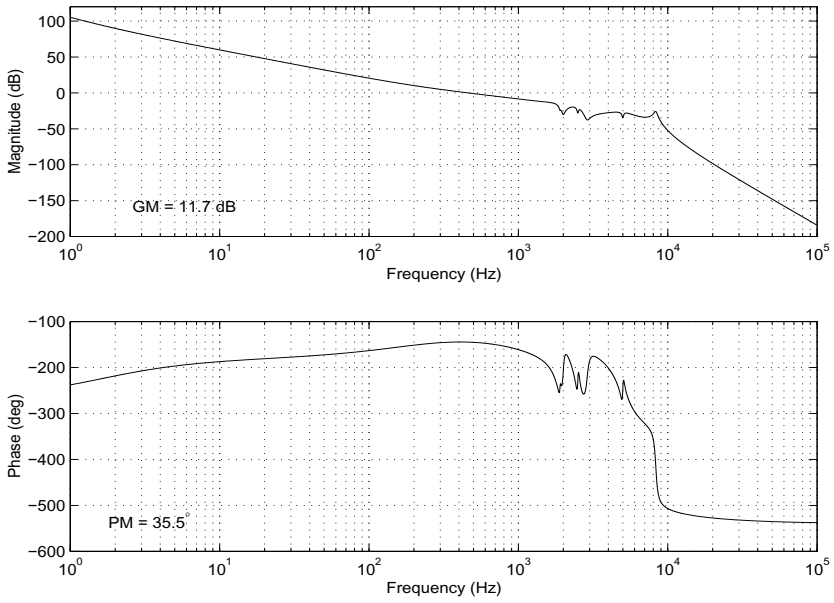


(a) Bode plot

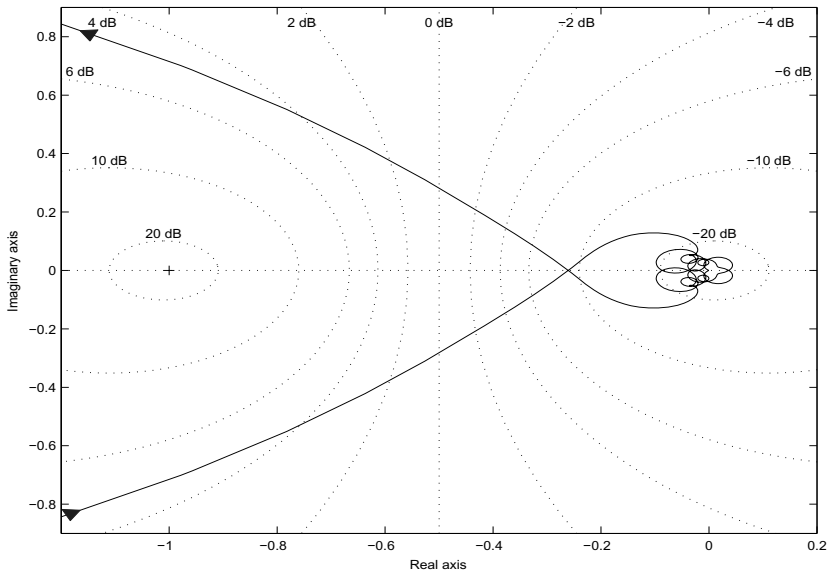


(b) Nyquist plot

Figure 6.8. Bode and Nyquist plots of the PID control system

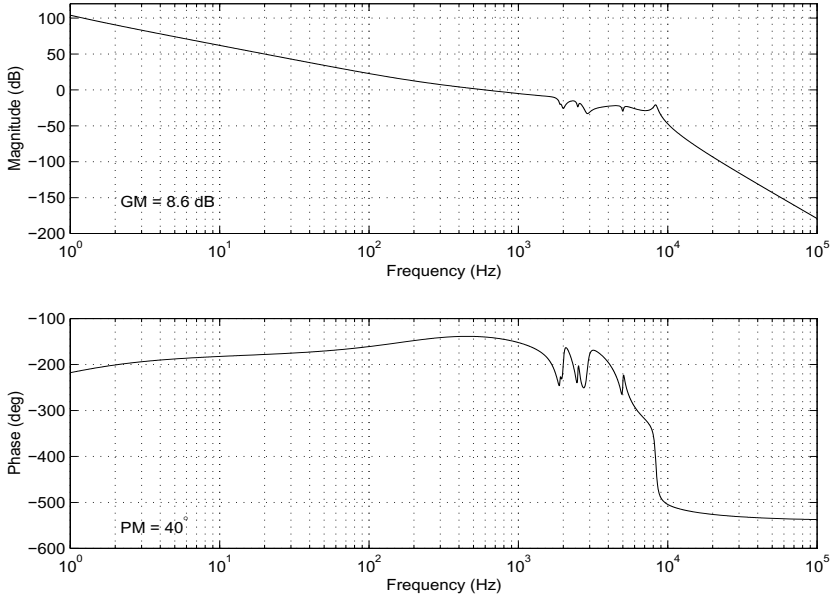


(a) Bode plot

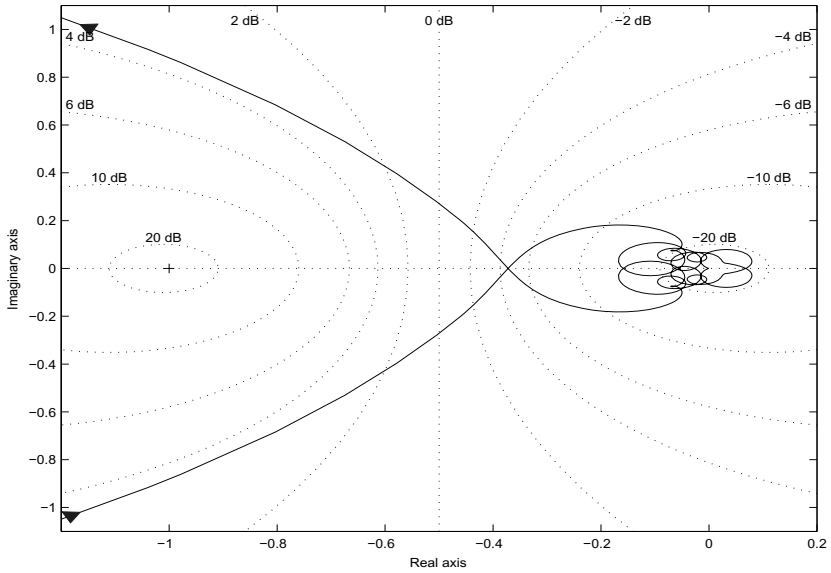


(b) Nyquist plot

Figure 6.9. Bode and Nyquist plots of the RPT control system

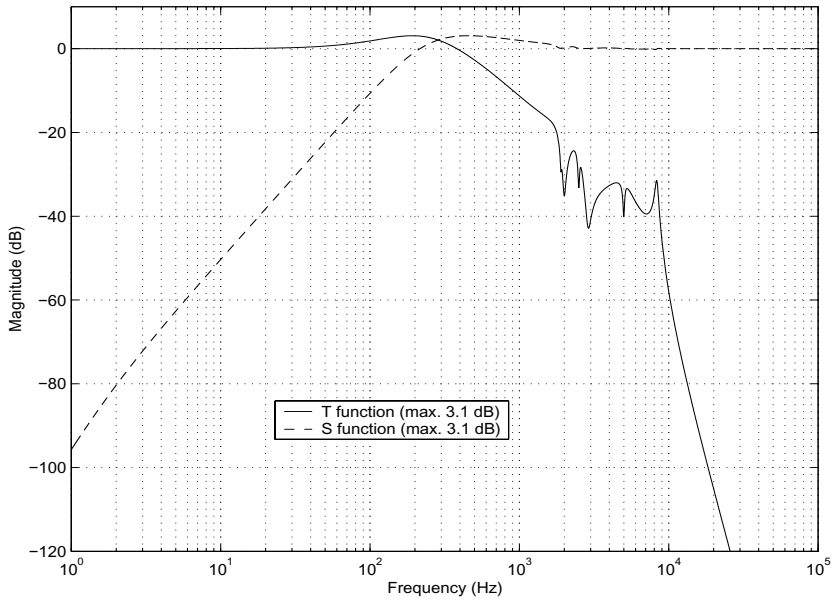


(a) Bode plot

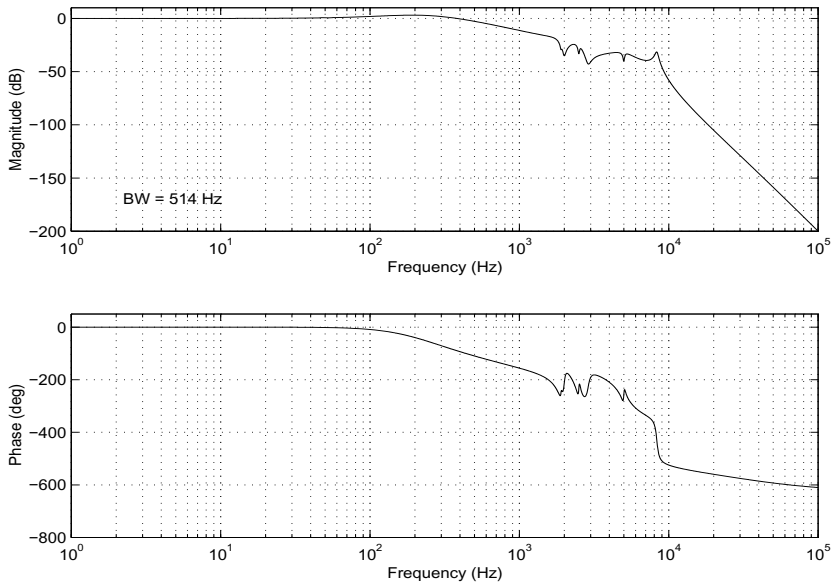


(b) Nyquist plot

Figure 6.10. Bode and Nyquist plots of the CNF control system

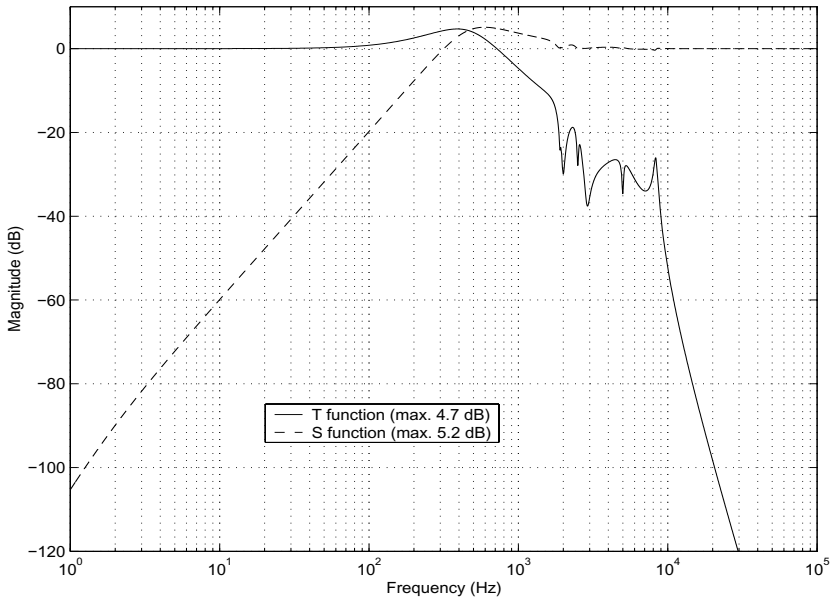


(a) Sensitivity and complementary sensitivity functions

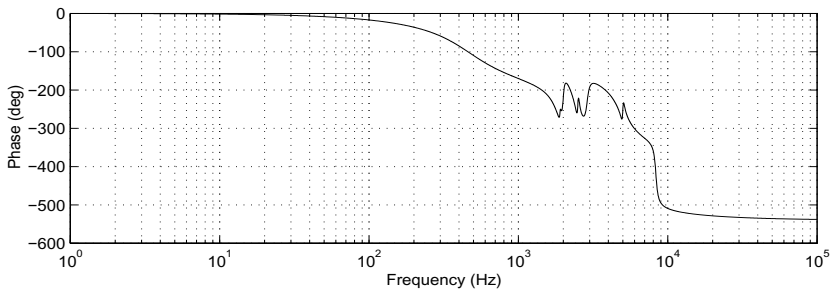
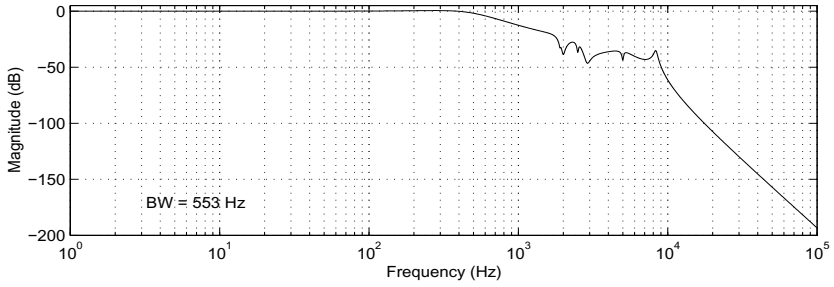


(b) Closed-loop response

Figure 6.11. Sensitivity functions and closed-loop transfer function of the PID control system

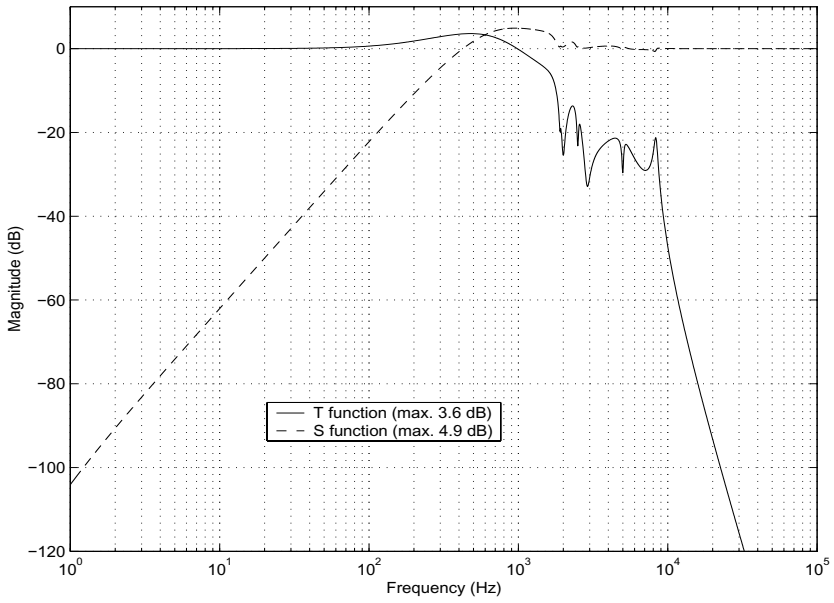


(a) Sensitivity and complementary sensitivity functions

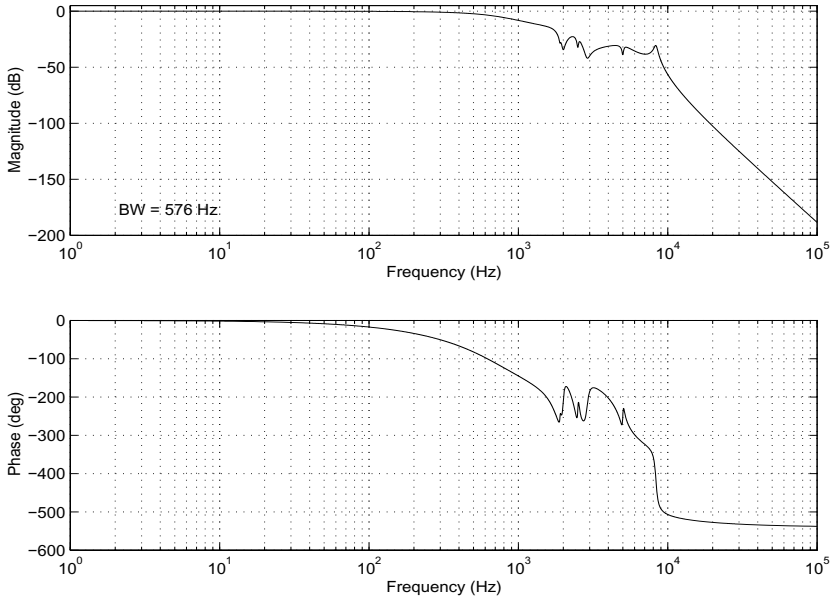


(b) Closed-loop response

Figure 6.12. Sensitivity functions and closed-loop transfer function of the RPT control system



(a) Sensitivity and complementary sensitivity functions



(b) Closed-loop response

Figure 6.13. Sensitivity functions and closed-loop transfer function of the CNF control system

with a fictitious runout disturbance injection

$$\tilde{w}(t) = 0.5 + 0.1 \cos(110\pi t) + 0.05 \sin(220\pi t) + 0.02 \sin(440\pi t) + 0.01 \sin(880\pi t) \quad (6.30)$$

and a zero reference $r(t)$. The result shows that the RPT and CNF controllers again have better performance and the effects of such a disturbance on the overall response under CNF control are minimal. A more comprehensive test on runout disturbances, *i.e.* the position error signal (PES) test on the actual system will be presented in the next section.

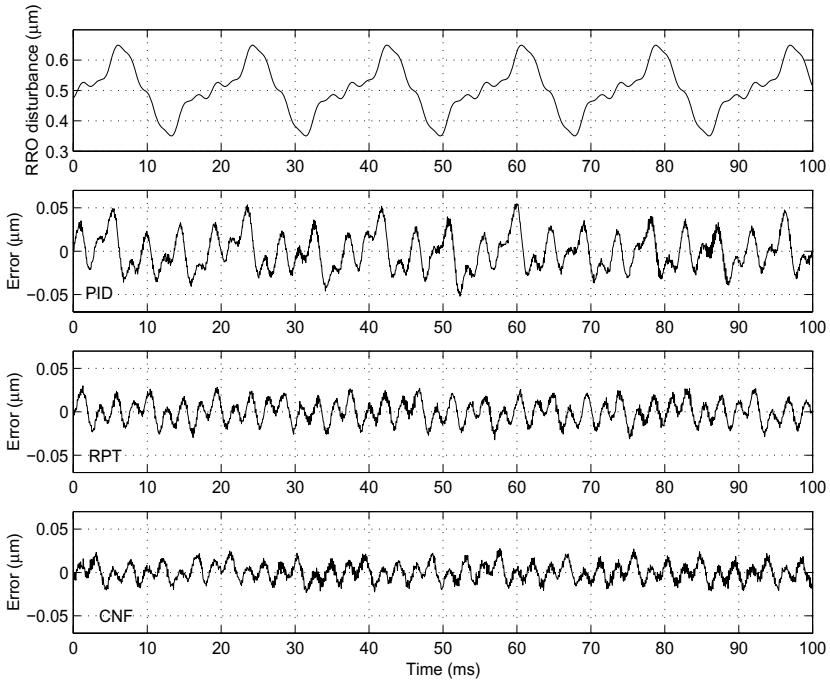


Figure 6.14. Closed-loop output responses due to a runout disturbance

6.4.4 Position Error Signal Test

The disturbances in a real HDD are usually considered as a lumped disturbance at the plant output, also known as runouts. Repeatable runouts (RROs) and nonrepeatable runouts (NRROs) are the major sources of track-following errors. RROs are caused by the rotation of the spindle motor and consists of frequencies that are multiples of the spindle frequency. NRROs can be perceived as coming from three main sources: vibration shocks, mechanical disturbance and electrical noise. Static force due to

flex cable bias, pivot-bearing friction and windage are all components of the vibration shock disturbance. Mechanical disturbances include spindle motor variations, disk flutter and slider vibrations. Electrical noises include quantization errors, media noise, servo demodulator noise and power amplifier noise. NRROs are usually random and unpredictable by nature, unlike repeatable runouts. They are also of a lower magnitude (see, *e.g.*, [1]). A perfect servo system for HDDs has to reject both the RROs and NRROs.

In our experiment, we have simplified the system somewhat by removing many sources of disturbances, especially that of the spinning magnetic disk. Therefore, we actually have to add the runouts and other disturbances into the system manually. Based on previous experiments, we know that the runouts in real disk drives are composed mainly of RROs, which are basically sinusoidal with a frequency of about 55 Hz, equivalent to the spin rate of the spindle motor. By manually adding this “noise” to the output while keeping the reference signal at zero, we can then read off the subsequent position signal as the expected PES in the presence of runouts. For actual drives, prewritten PES data might be estimated at high sampling rates using servo sector measurements (see, for example, [141]). In disk drive applications, the variation in the position of the R/W head from the center of the track during track following, which can be directly read off as the PES, is very important. Track-following servo systems have to ensure that the PES is kept to a minimum. Having deviations that are above the tolerance of the disk drive would result in too many read or write errors, making the disk drive unusable. A suitable measure is the standard deviation of the readings, σ_{pes} . A useful guideline is to make the $3\sigma_{\text{pes}}$ value less than 10% of the track pitch, which is about $0.1 \mu\text{m}$ for a track density of 25 kTPI.

Figure 6.15 shows the histograms of the tracking errors of the respective control systems under the disturbance of the runouts. The $3\sigma_{\text{pes}}$ values of the PES test are summarized in Table 6.2. Again, the CNF control yields the best performance in the PES test.

Table 6.2. The $3\sigma_{\text{pes}}$ values of the PES test

	PID control	RPT control	CNF control
$3\sigma_{\text{pes}}$ (μm)	0.0615	0.0375	0.0288

In conclusion, the RPT and CNF controllers have much better performance in track following and in the PES tests compared with that of the PID controller. We note that the results can be further improved if we used a better VCM actuator and arm assembly (such as those used in minidrives and microdrives) with a higher resonance frequency. We will carry out a detailed study on the servo system of a microdrive later in Chapter 9.

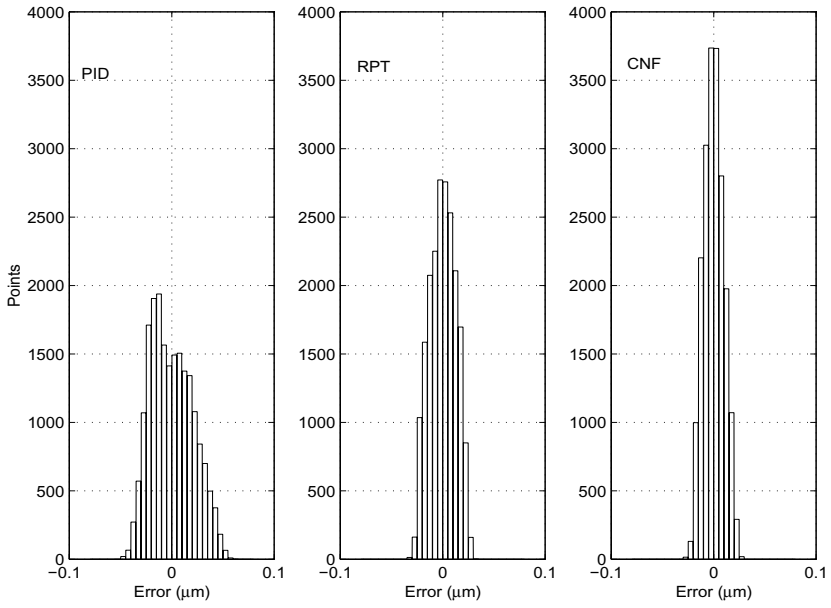


Figure 6.15. Implementation result: histograms of the PES tests

Track Seeking of a Single-stage Actuator

7.1 Introduction

In this chapter, we proceed to design track-seeking controllers for a single-stage actuated HDD that would give high-speed seeking performance. We utilize the nonlinear control techniques reported in Chapters 4 and 5 as well as the linear techniques reported in Chapter 3 to carry out the design of three different types of track-seeking controllers for a Maxtor HDD with a single VCM actuator. More specifically, we design the servo systems using the conventional PTOS approach, the CNF control technique, and the MSC system with PTOS and RPT controllers.

As in Chapter 6, a Maxtor (Model 51536U3) HDD is used to implement our design. The actual frequency response and the identified model are given Figure 6.1. The frequency-domain model has been identified earlier in Chapter 6 and is given in Equations 6.4–6.8. The same notch filter as in Equation 6.9 is again utilized for track seeking. With such a formulation, it is safe to approximate the VCM actuator model with the notch filter as a double integrator with an appropriate gain. Such an approximation simplifies the overall design procedure a great deal. Most importantly, it works very well. However, in order to make our design more realistic, all our simulation results are done using the tenth-order model. The final implementation is, of course, to be carried out on the actual system.

The following state-space model is then used throughout our design of track-seeking controllers:

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 6.4013 \times 10^7 \end{bmatrix} \text{sat}(u), \quad x = \begin{pmatrix} y \\ v \end{pmatrix} \quad (7.1)$$

where y and v are, respectively, the position of the VCM actuator head in micrometers and velocity in micrometers per second, and u is the control input in volts. In general, the velocity of the VCM actuator in the actual system is not available, and thus y is the only measurable state variable. For this particular system, the controlled output is also the measurement output, *i.e.*

$$h = y = [1 \quad 0]x \quad (7.2)$$

Our objective is to design a servo controller that meets the following physical constraints and design specifications:

1. the control input does not exceed ± 3 V owing to physical constraints on the actual VCM actuator;
2. the overshoot and undershoot of track seeking are kept to less than $0.5 \mu\text{m}$, the limit of our measurement device for large displacement. As such, the settling time used in this chapter is defined as the time taken for the R/W head to reach the $\pm 0.5 \mu\text{m}$ of the target track from its initial point.
3. the gain margin and phase margin of the overall design are, respectively, greater than 6 dB and 30° .

As mentioned earlier, three different approaches, namely the PTOS method, the MSC method and the CNF method, are presented in the following to design appropriate servo systems for the given HDD. We carry out control system design for each method first. All simulation and implementation results, as well as their comparison are to be discussed in the last section.

7.2 Track Seeking with PTOS Control

We present in this section the design and implementation of an HDD servo system using the PTOS approach (see Chapter 4). The first step is to find the state feedback gains k_1 and k_2 in the PTOS control law based on the design specifications. To get specifications in terms of required closed-loop poles we need the natural frequency ω_n and the damping ratio ζ . Let us choose the natural frequency to be 3141.6 rad/s , *i.e.* 500 Hz, and the damping factor to be 0.7255 so as to have an acceleration discount factor α of 0.95, which yields a reasonably good performance for seek lengths up to $300 \mu\text{m}$. It follows from Equation 4.32 that a PTOS control law with such a discount factor only increases the total tracking time by about 1.6% from that required in the TOC. Clearly, the performance of the PTOS control is pushed very closely to its limit, the TOC. Interested readers are referred to [30, 142, 143] for detailed information on the selection of these parameters. Note that the relation between the damping ratio ζ and the acceleration discount factor in PTOS control law is given by (see [30])

$$\alpha = \frac{1}{2\zeta^2} \quad (7.3)$$

Then, the corresponding s -plane closed-loop poles are

$$s_{1,2} = -2279.2 \pm j2162.2 \quad (7.4)$$

Using the m-function `acker` in MATLAB[®], we obtain the following feedback gains

$$k_1 = 0.15418 \quad \text{and} \quad k_2 = 7.1209 \times 10^{-5} \quad (7.5)$$

and the length of the linear region in PTOS can be found from Equation 4.31 and is given by $y_\ell = 19.458 \mu\text{m}$. Thus, the PTOS control law for our disk drive is as follows:

$$\tilde{u}_p = u_{\max} \cdot \text{sat} \left(\frac{k_2 [f_p(e) - v]}{u_{\max}} \right) \quad (7.6)$$

where $e = r - y$ with r being the target reference, and

$$f_p(e) = \begin{cases} \frac{k_1}{k_2} e, & \text{for } |e| \leq y_\ell \\ \text{sgn}(e) \left[\sqrt{2u_{\max} a \alpha} |e| - \frac{u_{\max}}{k_2} \right], & \text{for } |e| > y_\ell \end{cases} \quad (7.7)$$

with

$$a = 6.4013 \times 10^7, \quad k_1 = 0.15418, \quad k_2 = 7.1209 \times 10^{-5} \quad (7.8)$$

and

$$y_\ell = 19.458, \quad \alpha = 0.95 \quad (7.9)$$

The advantage of this control scheme is that it is quite simple to understand. The implementation of such a controller requires an estimation of the VCM actuator velocity (with the estimator pole being placed at -4000). More precisely, the following velocity estimator is used:

$$\dot{x}_c = -4000x_c + 6.4013 \times 10^7 \tilde{u}_p - 1.6 \times 10^7 y \quad (7.10)$$

and

$$\hat{v} = x_c + 4000y \quad (7.11)$$

In the actual simulation and implementation of the PTOS controller of Equation 7.6, v is replaced with \hat{v} of Equation 7.11 and

$$u = G_{\text{notch}}(s) \cdot \tilde{u}_p \quad (7.12)$$

where $G_{\text{notch}}(s)$ is as given in Equation 6.9. The simulation and implementation results of the above design will be given later in Section 7.5.

7.3 Track-seeking with MSC

In this section, we apply the MSC method of Chapter 4 to the disk drive given earlier. The MSC scheme uses the proximate time-optimal controller in the track-seeking mode, and the RPT controller in the track-following mode. We note that in MSC, initially, the plant is controlled by the seeking controller and at the end of the seeking mode a switch changes it to a track-following controller. In [127], the mode switching was done after finding the optimal mode-switching conditions such that the impact of the initial values on settling performance was minimized. But the impact of the resulting control signal on the resonance modes was not considered. It has been shown [74, 106] that the RPT controller is independent of these initial values. The optimal mode-switching conditions in our scheme can just be set such that the control signal is small enough so as not to excite the resonance vibrations. The

problem of unmodeled mechanical resonance can be treated more rigorously either by minimizing the jerk as defined by $\|du/dt\|_2$ as reported in [144] or by using a method developed in the frequency domain in [145]. However, by utilizing the features of RPT control, such as it works for a wide range of resonance frequencies (see Chapter 6), the mode-switching conditions can be determined in a very simple way (see Chapter 4).

We now move to present an MSC controller for the HDD with a single VCM actuator. The control law in track-seeking mode (here we label its control signal as u_p) is given in Equation 7.6, as this mode uses the PTOS control. The control law in the track-following mode, *i.e.* the reduced-order measurement feedback RPT control law, is given by

$$\begin{pmatrix} \dot{x}_i \\ \dot{x}_c \end{pmatrix} = A_{RC} \begin{pmatrix} x_i \\ x_c \end{pmatrix} + B_{RC} \begin{pmatrix} r \\ y \end{pmatrix}, \quad \tilde{u}_R = C_{RC} \begin{pmatrix} x_i \\ x_c \end{pmatrix} + D_{RC} \begin{pmatrix} r \\ y \end{pmatrix} \quad (7.13)$$

with r being the target reference and

$$\left. \begin{aligned} A_{RC} &= \begin{bmatrix} 0 & 0 \\ -1.0268 \times 10^3 & -8.8166 \times 10^3 \end{bmatrix} \\ B_{RC} &= \begin{bmatrix} -1 & 1 \\ 1.0316 \times 10^2 & -4.5583 \times 10^2 \end{bmatrix} \\ C_{RC} &= [-1.6041 \quad -7.5245] \\ D_{RC} &= [0.16116 \quad -0.46214] \end{aligned} \right\} \quad (7.14)$$

Next we find the mode-switching conditions as defined in Equation 4.62. Using the RPT controller parameters, and following the results of Chapter 4, the mode-switching conditions can be determined as $|e(t_1)| \leq 2 \mu\text{m} < y_\ell = 19.458 \mu\text{m}$ and $|v(t_1)| \leq 5000 \mu\text{m/s}$. We select the MSC law

$$\tilde{u} = \begin{cases} \tilde{u}_p, & t < t_1 \\ \tilde{u}_R, & t \geq t_1 \end{cases} \quad (7.15)$$

in which \tilde{u}_p is as given in Equation 7.6 and t_1 is chosen such that

$$|y(t_1) - r| = 2 \mu\text{m} \quad \text{and} \quad |\hat{v}(t_1)| \leq 5000 \mu\text{m/s} \quad (7.16)$$

As in the PTOS case, the actual control signal is generated by

$$u = G_{\text{notch}}(s) \cdot \tilde{u} \quad (7.17)$$

The overall closed-loop system comprising the given VCM actuated HDD and the MSC control law is asymptotically stable. For easy comparison, the simulation and implementation of the overall system with the MSC control law will again be presented in Section 7.5.

7.4 Track Seeking with CNF Control

We now move to the design of a reduced-order continuous-time composite nonlinear control law as given by Equations 5.79 and 5.80 for the commercial hard disk model shown in Figure 6.1. As the CNF control law depends on the size of the step command input, we derive, for our HDD model given in Equation 7.1, the following parameterized state feedback gain $F(\varepsilon)$:

$$F(\varepsilon) = -\frac{1}{6.4013 \times 10^7} \left[\frac{4\pi^2 f^2 s_0}{\kappa_1 \varepsilon^3} \quad \frac{4\pi^2 f^2 + 4\pi f \zeta s_0}{\varepsilon^2} \quad \frac{4\pi f \zeta + s_0}{\varepsilon} \right] \quad (7.18)$$

which places the eigenvalues of $\bar{A} + \bar{B}F(\varepsilon)$ exactly at $(-\zeta \pm j\sqrt{1-\zeta^2})2\pi f/\varepsilon$ and $-s_0/\varepsilon$. The latter is the pole associated with the integration dynamics. Following the design procedure given in Chapter 5 and the physical properties of the given system, for $r = 300 \mu\text{m}$, which is to be used in the next section for simulation and implementation, we choose a damping ratio of $\zeta = 0.52$ and $f = 200 \text{ Hz}$, which corresponds roughly to the normal working frequency range of the linear part of the CNF control law with $\varepsilon = 1$. Selecting $\kappa_1 = 1$, $s_0 = 0.001$ and

$$W = \begin{bmatrix} 1.2 \times 10^{-5} & 0 & 0 \\ 0 & 10^{-3} & 0 \\ 0 & 0 & 10^{-20} \end{bmatrix} \quad (7.19)$$

we obtain

$$P = \begin{bmatrix} 6.0005 \times 10^{-3} & 4.9660 \times 10^{-6} & 3.7995 \times 10^{-9} \\ 4.9660 \times 10^{-6} & 8.0005 \times 10^{-7} & 3.1977 \times 10^{-10} \\ 3.7995 \times 10^{-9} & 3.1977 \times 10^{-10} & 2.4468 \times 10^{-13} \end{bmatrix} \quad (7.20)$$

and a reduced-order CNF control law as follows:

$$\dot{x}_1 = y - r \quad (7.21)$$

with r being the target reference,

$$\dot{x}_c = -4000x_c + 6.4013 \times 10^7 \tilde{u} - 1.6 \times 10^7 y \quad (7.22)$$

and

$$\tilde{u} = \text{sat} \left\{ \left[K_1 + \rho(e)K_2 \right] \begin{pmatrix} x_1 \\ y - r \\ x_c + 4000y \end{pmatrix} \right\} \quad (7.23)$$

where

$$K_1 = [-2.4669 \times 10^{-5} \quad -2.4669 \times 10^{-2} \quad -2.0416 \times 10^{-5}] \quad (7.24)$$

$$K_2 = [0.24322 \quad 2.0470 \times 10^{-2} \quad 1.5663 \times 10^{-5}] \quad (7.25)$$

and

$$\rho(e) = -2.4287 \left| e^{-2|e|/r} - e^{-2} \right| \quad (7.26)$$

Again, the actual control signal is generated as follows:

$$u = G_{\text{notch}}(s) \cdot \tilde{u} \quad (7.27)$$

Note that in both simulation and implementation, the initial condition of x_1 is set to zero at $t = 0$ and reset to zero when the R/W head of the actuator reaches the point that is $2 \mu\text{m}$ from the target to reinforce the integration action. Again, the simulation and implementation results of the servo system with the CNF control law will be presented in the next section for an easy comparison.

7.5 Simulation and Implementation Results

Now, we are ready to present the simulation and implementation results for all three servo systems discussed in the previous sections and do a full-scale comparison on the performances of these methods. In particular, we study the following tests:

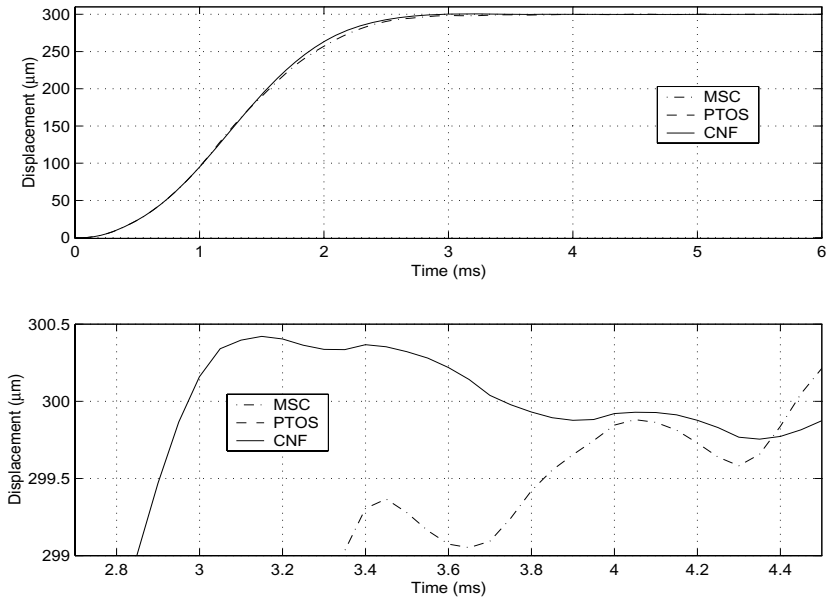
1. track-seeking test;
2. frequency-domain test;
3. runout disturbance test; and
4. PES test.

All simulation results presented in this section have been obtained using Simulink[®] in MATLAB[®] and all implementation results are carried out using our own experimental setup as described in Chapter 1. The sampling frequency for actual implementation is chosen as 20 kHz. Here, we note that all our controllers are discretized using the ZOH technique.

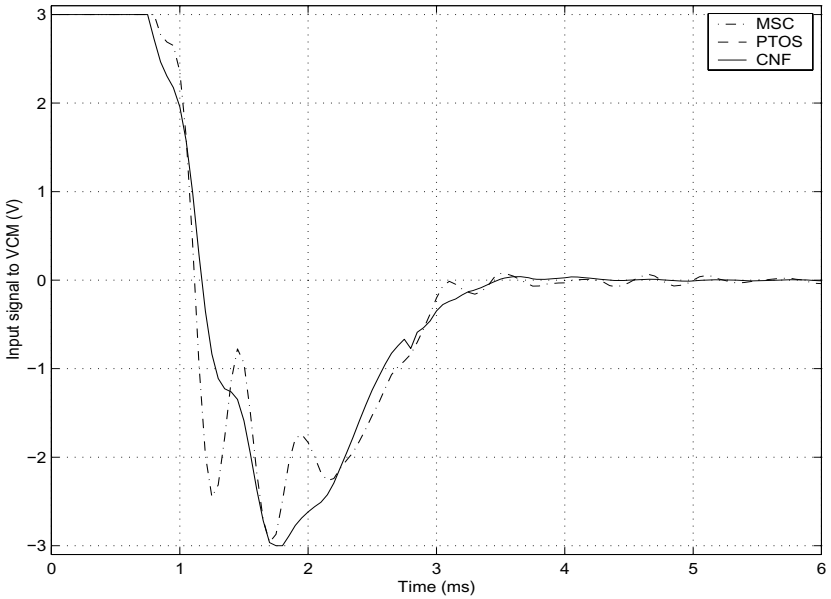
7.5.1 Track-seeking Test

In our simulation and implementation, we use a track pitch of $1 \mu\text{m}$ for the HDD. In what follows, we present results for a track seek length of $300 \mu\text{m}$. Unfortunately, owing to the capacity of the LDV that has been used to measure the displacement of the R/W head of the VCM actuator, the absolute errors of our implementation results given below are $0.5 \mu\text{m}$. As such, the settling time for both implementation and simulation results is defined as the total time required for the R/W head to move from its initial position to the entrance of the region of the final target with plus and minus the absolute error. This is the best we can do with our current experimental setup. Nonetheless, the results we obtain here are sufficient to illustrate our design ideas and philosophy. The simulation and implementation results of the track-seeking performances of the obtained servo systems are, respectively, shown in Figures 7.1 and 7.2. We summarize the overall results on settling times in Table 7.1.

Clearly, the simulation and implementation results show that the servo system with the CNF controller has the best performance. We believe that this is due to the fact that the CNF control law unifies the nonlinear and linear components without switching, whereas the other two servo systems involve switching elements between the nonlinear and linear parts, which degrades the overall performance.

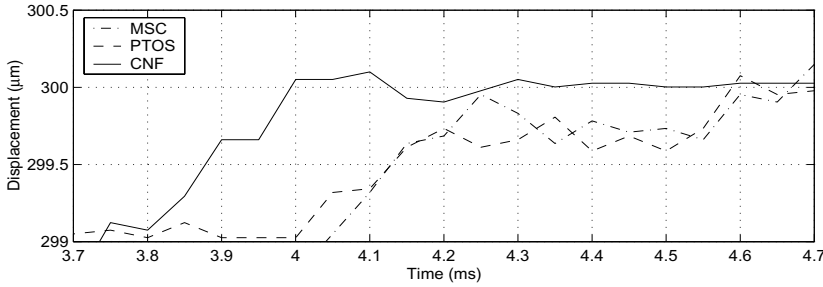
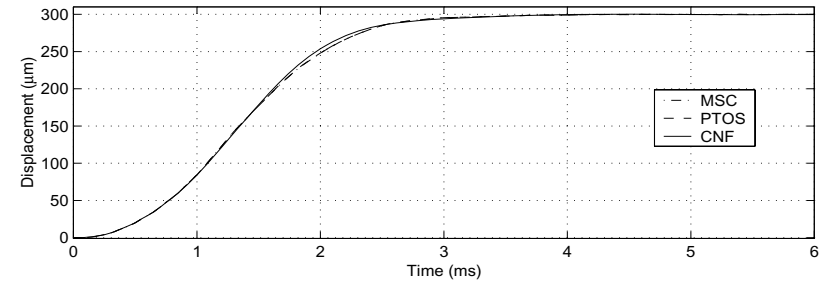


(a) Output responses

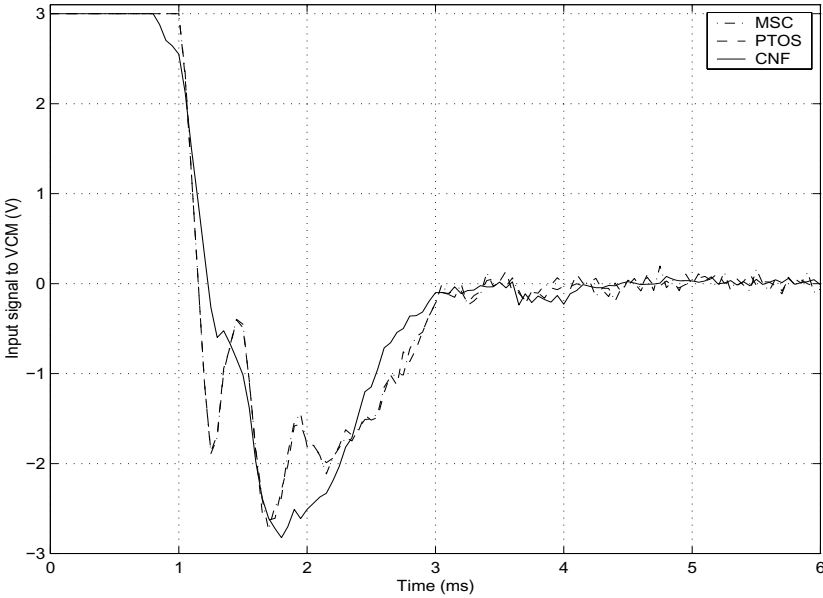


(b) Control signals

Figure 7.1. Simulation result: response and control of the track-seeking systems



(a) Output responses



(b) Control signals

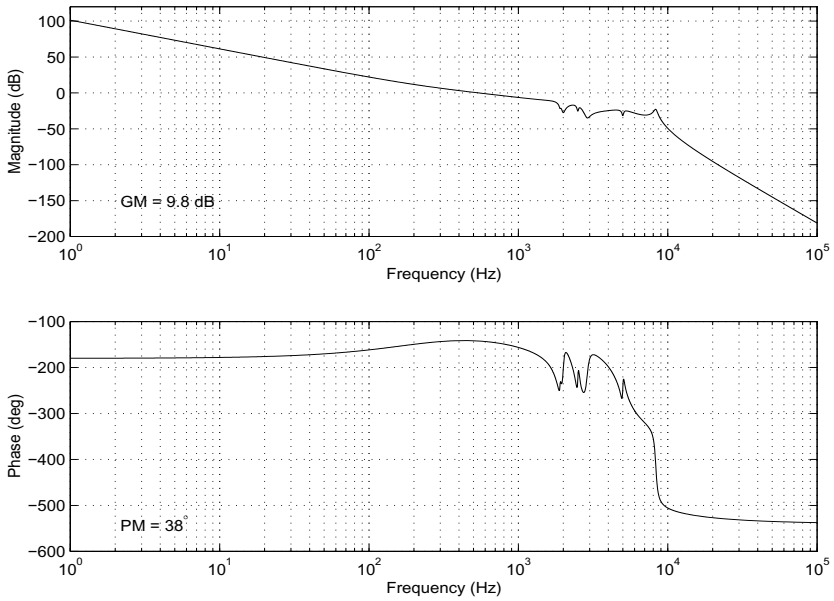
Figure 7.2. Implementation result: response and control of the track-seeking systems

Table 7.1. Simulation and implementation: settling time of the track-seeking systems

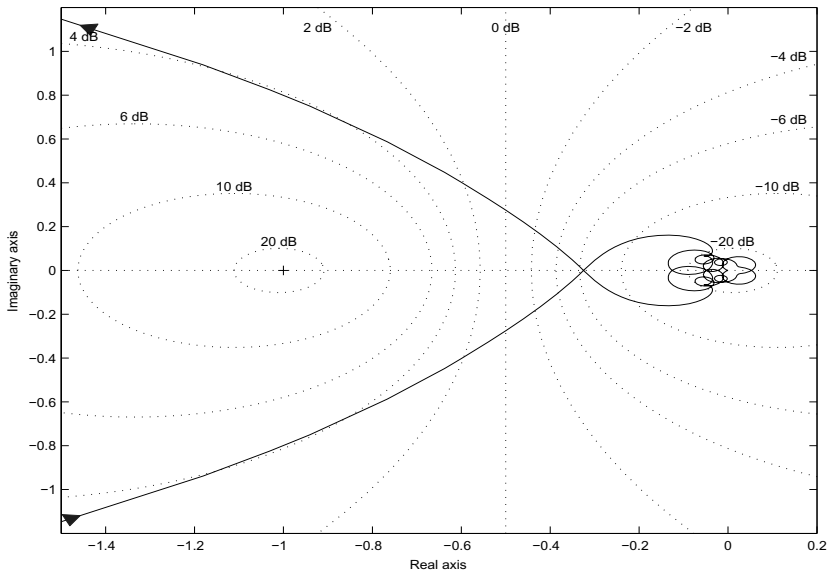
Settling time (ms)					
Simulation			Implementation		
PTOS	MSC	CNF	PTOS	MSC	CNF
3.85	3.85	2.95	4.15	4.15	3.90

7.5.2 Frequency-domain Test

As seen in the previous sections, all three track-seeking controllers that we have designed are nonlinear in nature. The frequency-domain properties of nonlinear track-seeking controllers are not well defined and in fact not as important as those of track-following controllers. For completeness and for comparison, we define the frequency responses of the open and closed-loop systems with a nonlinear controller as those corresponding to the steady-state situation when the nonlinear action of the controller has become a constant or vanished. The Bode plot, Nyquist plot, sensitivity and complementary sensitivity functions as well as the closed-loop frequency responses for the system with each control technique are, respectively, shown in Figures 7.3 to 7.8. All three control systems meet the design specifications in the frequency domain. Finally, we note that for track-seeking controllers, the PES test is not necessary.

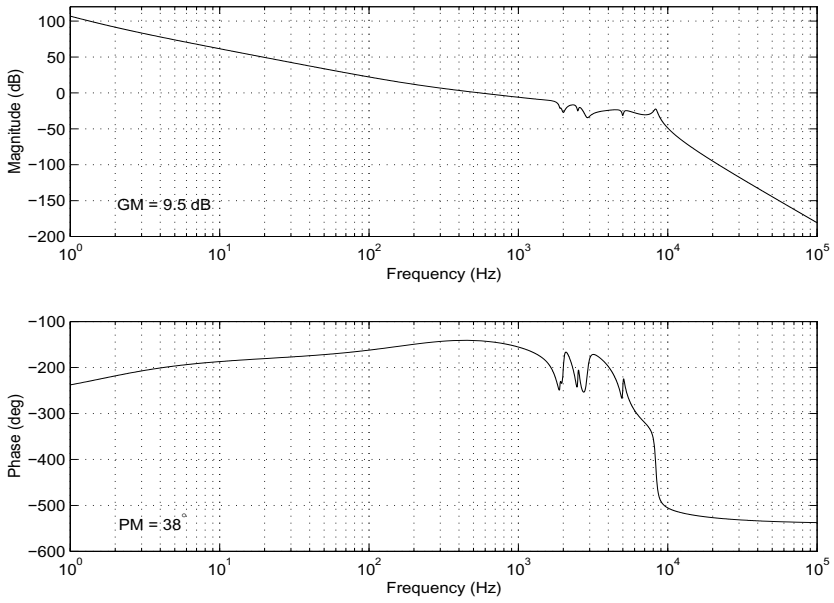


(a) Bode plot

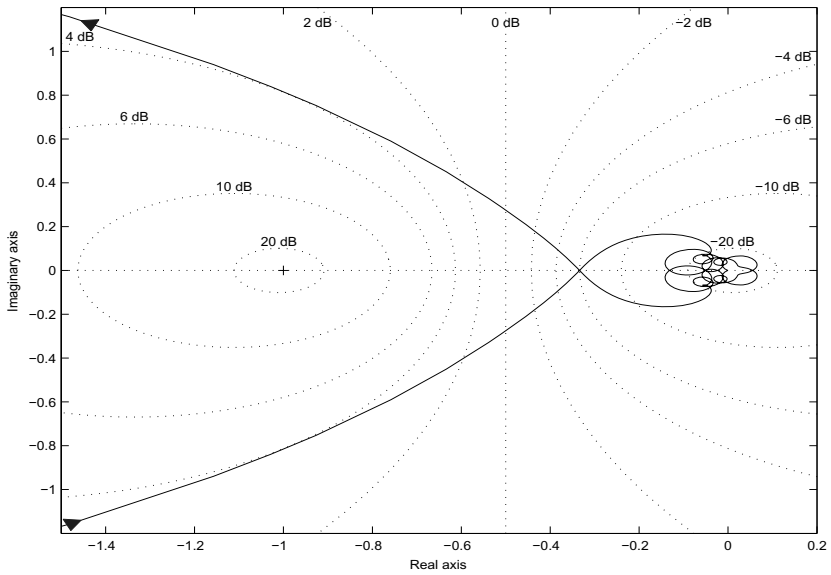


(b) Nyquist plot

Figure 7.3. Bode and Nyquist plots of the PTOS control system

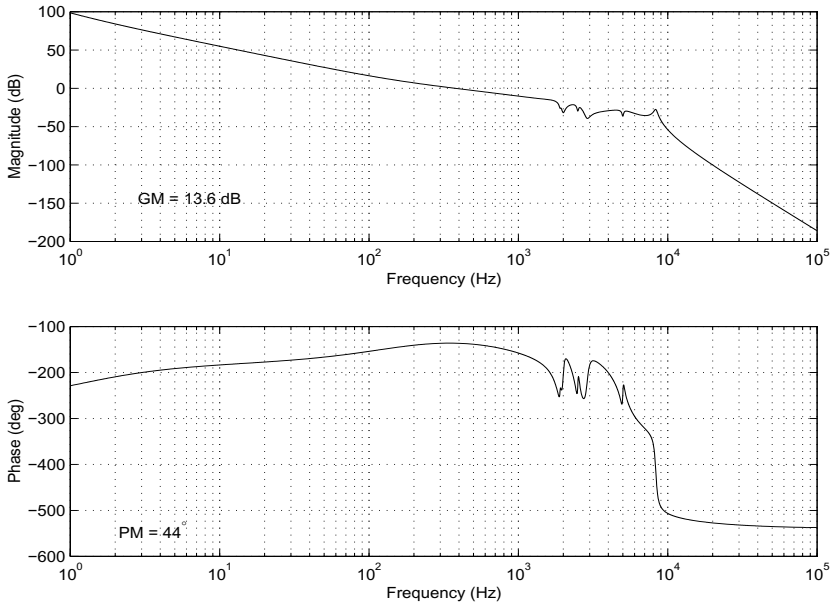


(a) Bode plot

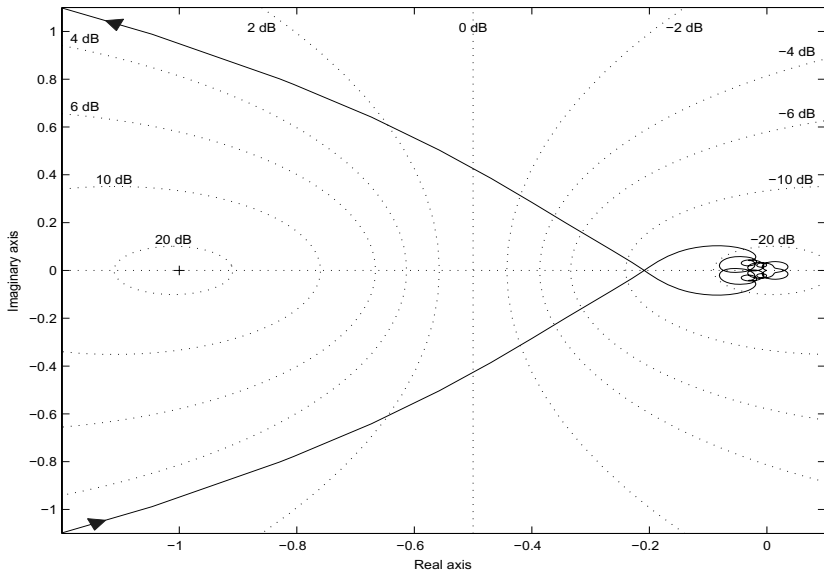


(b) Nyquist plot

Figure 7.4. Bode and Nyquist plots of the MSC control system

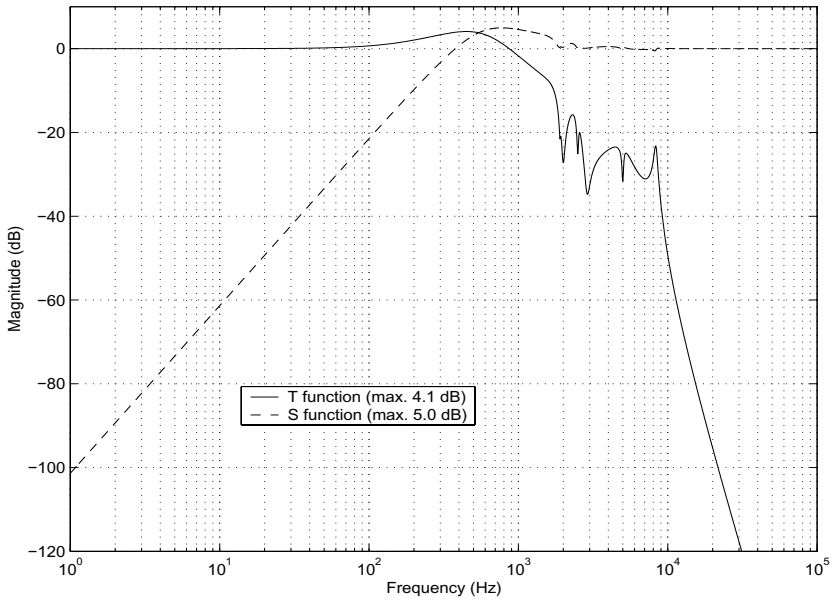


(a) Bode plot

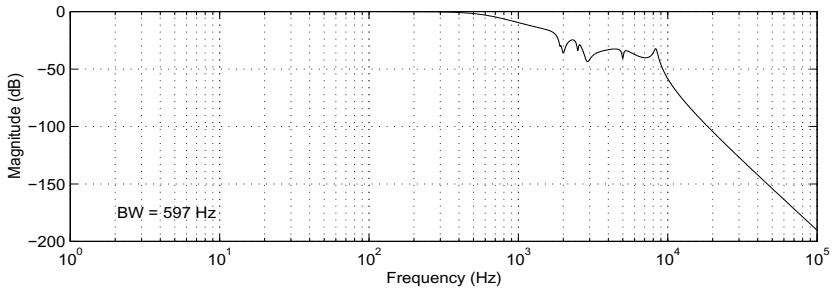


(b) Nyquist plot

Figure 7.5. Bode and Nyquist plots of the CNF control system

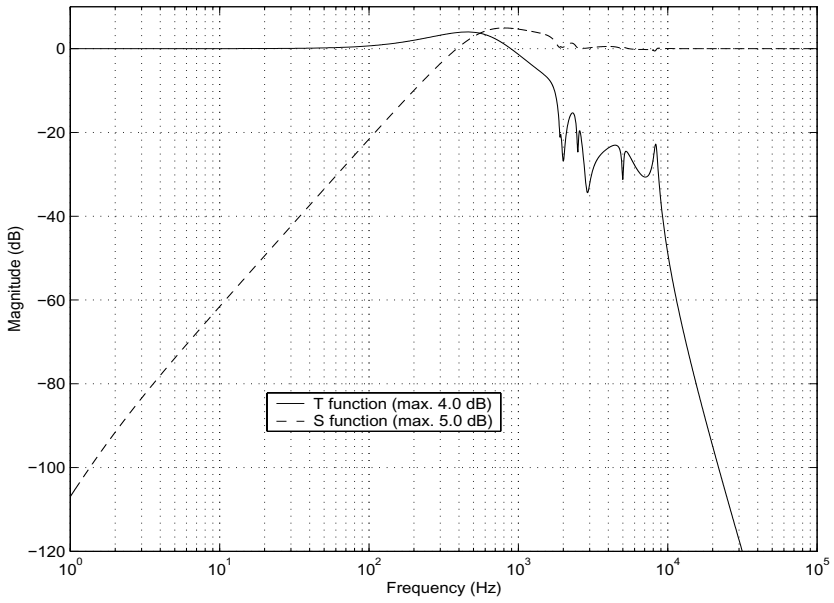


(a) Sensitivity and complementary sensitivity functions

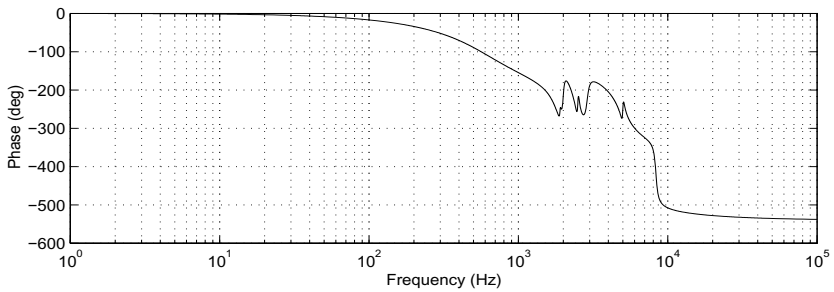
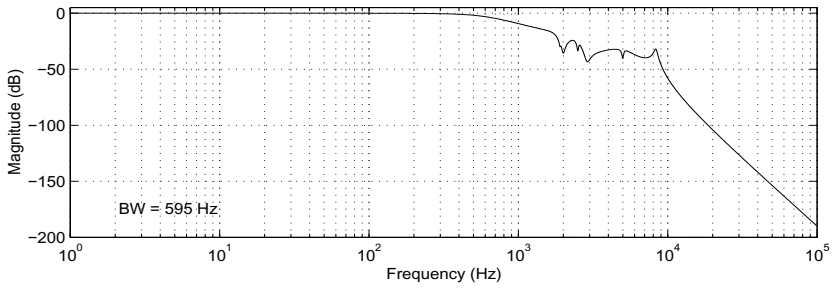


(b) Closed-loop response

Figure 7.6. Sensitivity functions and closed-loop transfer function of the PTOS control system

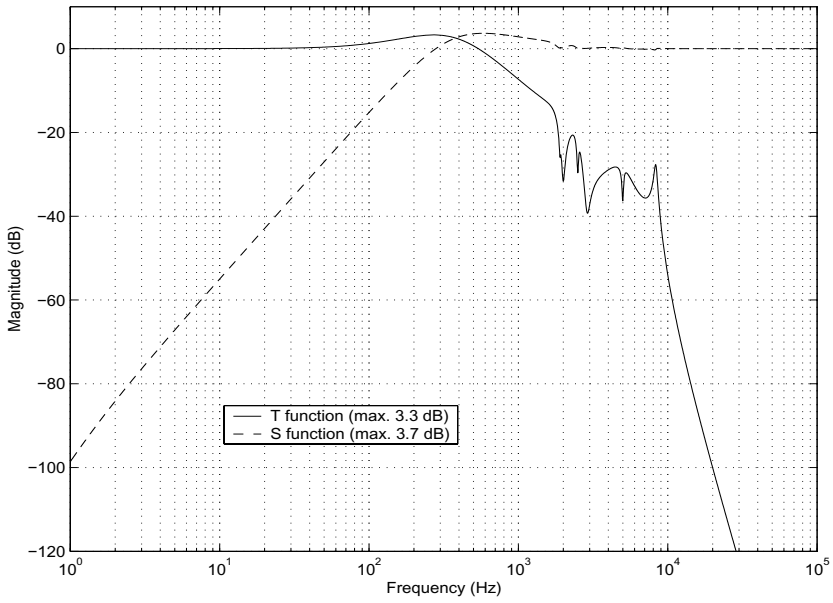


(a) Sensitivity and complementary sensitivity functions

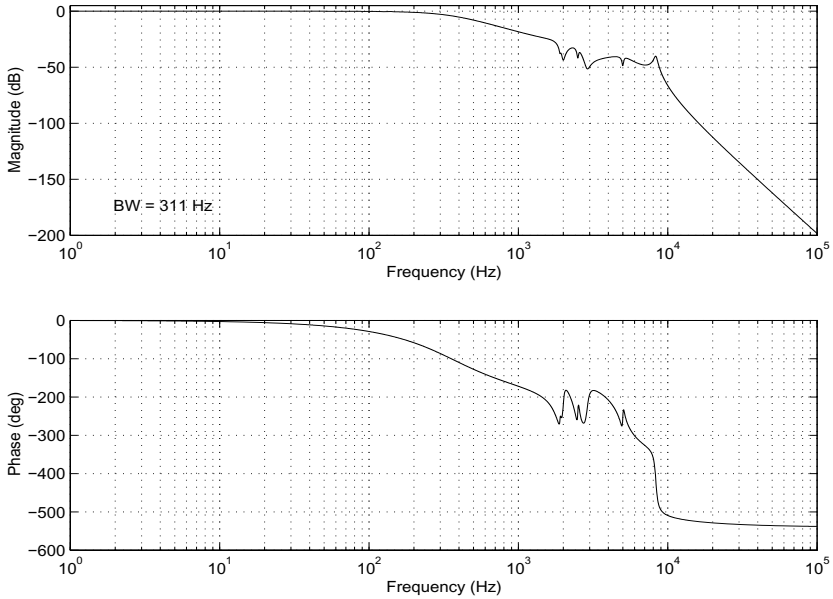


(b) Closed-loop response

Figure 7.7. Sensitivity functions and closed-loop transfer function of the MSC control system



(a) Sensitivity and complementary sensitivity functions



(b) Closed-loop response

Figure 7.8. Sensitivity functions and closed-loop transfer function of the CNF control system

Dual-stage Actuated Servo Systems

8.1 Introduction

The present demand for large-capacity disk drives is leading to an increase in areal density at a rate of 100% per year. This requires a positioning accuracy of the order of a few nanometers. The servo bandwidth of the current disk drive actuators makes it very hard to achieve this. The VCM actuator used in conventional disk drives has hundreds of flexible resonances at high frequencies (see, *e.g.*, [1, 30]), which limits the increase of bandwidth and hence the positioning accuracy. In order to develop high bandwidth (track-following) servo systems, dual-stage actuation has been proposed as a possible solution. Dual-stage actuator refers to the fact that there is a microactuator mounted on a large conventional VCM actuator. A dual-stage actuated HDD was successfully demonstrated by Tsuchiura *et al.* [146] of Hitachi. In [146], a fine positioner based on a piezoelectric structure was mounted at the end of a primary VCM stage to form the dual actuator. The higher bandwidth of the fine positioner allowed the R/W heads to be positioned accurately. There have been other instances where electromagnetic [147] and electrostatic [148] microactuators have been used for fine positioning of R/W heads. The two most fundamental choices in a dual-stage system are the actuator configuration and the control algorithm. There have been proposals for electromagnetic, electrostatic, piezoelectric, shape memory and rubber microactuators, *etc.*, each with their own advantages and disadvantages. Many research studies have been done and reported in the literature (see, *e.g.*, [20, 148–167] just to name a few). In this chapter, we focus on the design of complete HDD servo systems with a dual-stage actuator with a piezoelectric actuator in its second stage (see Figure 8.1).

Diverse control strategies and methods have been reported in the design of HDD servo systems with a dual-stage actuator (see, *e.g.*, [150, 154, 156, 157, 159, 161–163]. Guo *et al.* [154] have proposed four control strategies to design the dual-stage actuator control system: the so-called parallel loop, master–slave loop, dual feedback loop and master-slave with decoupling methods. Hu *et al.* [159] and Guo *et al.* [156] have also utilized the well-known LQG/LTR method to design the dual-stage actuator control system. These studies have accelerated the progress to improve HDD

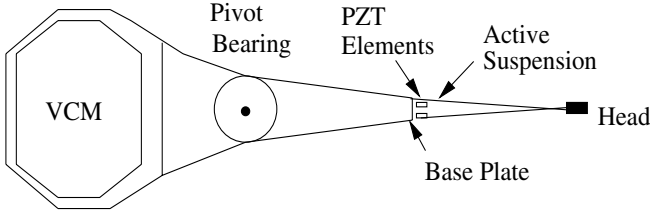


Figure 8.1. A dual-stage HDD actuator

servo system performances, but more studies need to be done before such dual-stage actuated HDDs can be considered for commercialization.

We present in this chapter the design of track-following controllers for the dual-stage actuated HDD with three different approaches, *i.e.* the PID control, the RPT control method, and the CNF control technique. In each design, the VCM actuator is controlled by a control law obtained using one of these three approaches and the microactuator is controlled by a proportional gain together with an appropriate filter.

8.2 Modeling of a Dual-stage Actuator

In this section we develop a model for the dual-stage actuator. Since the VCM actuator and the microactuator are decoupled, we need to identify two separate models for the VCM actuator and for the microactuator. A Maxtor dual-stage HDD is used in our design and implementation. It has a fine positioner based on a piezoelectric suspension mounted at the end of a primary VCM arm (see Figure 8.1), and the microactuators produce the relative motion of the R/W head along the radial direction. Here we note that only the displacement of the R/W head is available as the measurement output. Also, the VCM arm in this HDD is quite similar to that in the HDD studied in Chapters 6 and 7. Figures 8.2 and 8.3 respectively show the frequency-response characteristics of the VCM actuator and the microactuator.

Using the data measured from the actual system, and the identification algorithm given in Chapter 2 (see also [13, 59]), we obtain a tenth-order model for the actuator (see also Chapter 6):

$$G_v(s) = \frac{6.4013 \times 10^7}{s^2} \prod_{i=1}^4 G_{v,r,i}(s) \quad (8.1)$$

with

$$G_{v,r,1}(s) = \frac{0.912s^2 + 457.4s + 1.433 \times 10^8}{s^2 + 359.2s + 1.433 \times 10^8} \quad (8.2)$$

$$G_{v,r,2}(s) = \frac{0.7586s^2 + 962.2s + 2.491 \times 10^8}{s^2 + 789.1s + 2.491 \times 10^8} \quad (8.3)$$

$$G_{v,r,3}(s) = \frac{9.917 \times 10^8}{s^2 + 1575s + 9.917 \times 10^8} \quad (8.4)$$

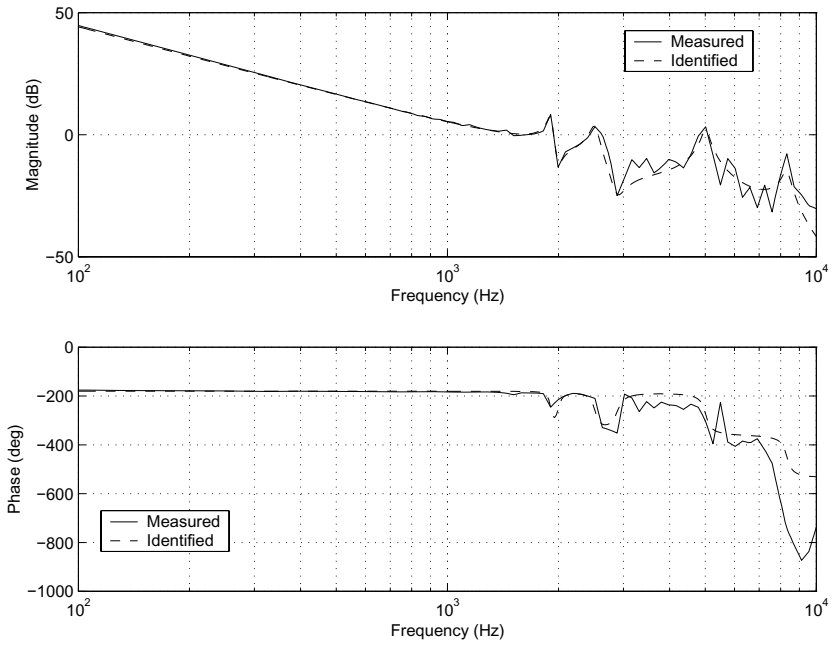


Figure 8.2. Frequency responses of the actual and identified VCM actuator models

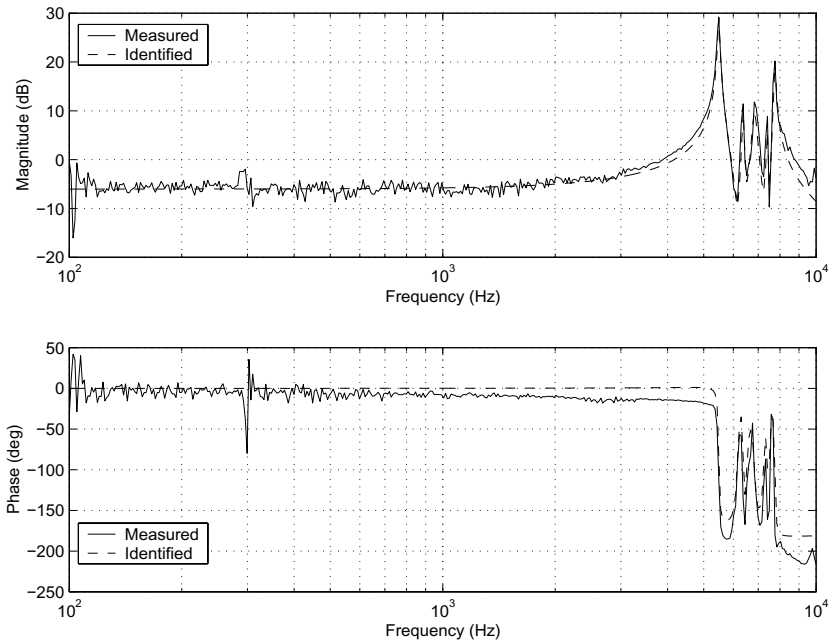


Figure 8.3. Frequency responses of the actual and identified microactuator models

and

$$G_{v,r,4}(s) = \frac{2.731 \times 10^9}{s^2 + 2613s + 2.731 \times 10^9} \quad (8.5)$$

and a tenth-order model for the microactuator,

$$G_m(s) = 0.5 \prod_{i=1}^5 G_{m,r,i}(s) \quad (8.6)$$

with

$$G_{m,r,1}(s) = \frac{0.7938s^2 + 767.9s + 1.189 \times 10^9}{s^2 + 344.8s + 1.189 \times 10^9} \quad (8.7)$$

$$G_{m,r,2}(s) = \frac{0.955s^2 + 978.6s + 1.605 \times 10^9}{s^2 + 400.6s + 1.605 \times 10^9} \quad (8.8)$$

$$G_{m,r,3}(s) = \frac{0.8912s^2 + 1013s + 1.843 \times 10^9}{s^2 + 1073s + 1.843 \times 10^9} \quad (8.9)$$

$$G_{m,r,4}(s) = \frac{0.9772s^2 + 460.1s + 2.167 \times 10^9}{s^2 + 465.5s + 2.167 \times 10^9} \quad (8.10)$$

and

$$G_{m,r,5}(s) = \frac{2.376 \times 10^9}{s^2 + 487.4s + 2.376 \times 10^9} \quad (8.11)$$

We note here that the inputs to both actuators are voltages (in volts) and their outputs are displacements (in micrometers). It can be seen clearly from Figures 8.2 and 8.3 that the actual responses and those of the identified models are matched very well at the frequencies of interest.

8.3 Dual-stage Servo System Design

We now carry out the design of servo systems for the HDD with a dual-stage actuator. Similarly, we would like to design our servo systems to the following constraints and requirements.

1. The control input to the VCM actuator does not exceed ± 3 V, whereas the control input to the microactuator is within ± 2 V.
2. The displacement of the microactuator does not exceed 1 μm . Moreover, it has to settle down to zero in the steady state so that the microactuator can be further used for the next move.
3. The overshoot and undershoot of the step response are kept less than 0.05 μm , *i.e.* 5% of one track pitch. As pointed out earlier, the R/W head of the HDD servo system can start writing data onto the disk when it is within 5% of one track pitch of the target.
4. the gain margin and phase margin of the overall design are respectively greater than 6 dB and 30° ;

5. the maximum peaks of the sensitivity and complementary sensitivity functions are less than 6 dB; and
6. the sampling frequency in implementing the actual controller is 20 kHz.

Unfortunately, the only available measurement in the dual-stage actuated HDD is the displacement of the R/W head, which is a combination of the displacement of the VCM actuator and that of the microactuator. Practically, we have to control both actuators using the same measurement, which makes the servo system design very difficult. Observing that the frequency response of the microactuator is as in Figure 8.3, we find that it is nothing more than a constant gain at low frequencies with a gain of 0.5. This property is valid so long as we do not push the speed of the microactuator too high. As such, we propose in Figure 8.4 a control configuration for the dual-stage actuated servo system. We note that the filter in the feedforward path of the microactuator is the combination of an appropriate notch filter, which is to attenuate the high-frequency resonance modes, and a lead-lag compensator, which is to compensate some phase losses resulted in the notch filter.

To be more specific, we estimate the displacement of the microactuator \hat{y}_m directly from its input u_m and then the estimation of the displacement of the VCM actuator can be obtained as $y - \hat{y}_m$. It can be observed from the configuration in Figure 8.4 that y_m , the displacement of the microactuator, settles down to zero as the tracking error approaches zero. As pointed out earlier, such a feature would enable the microactuator to be used for the next move. Since the microactuator can only produce a maximum displacement of 1 μm , it is helpful in track seeking. As such, we only focus on the design of controllers for track following, in which the microactuator is expected to contribute significantly.

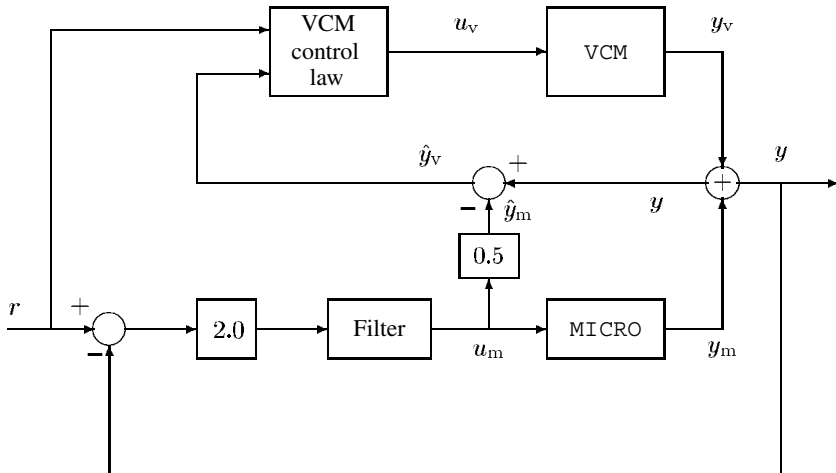


Figure 8.4. The schematic representation of a dual-stage actuator control

To reduce the effects of the resonance modes in the VCM actuator, as in Chapter 6, we introduce a notch filter whose transfer function is given as in Equation 6.9. We can then approximate the models of the VCM actuator as follows:

$$\Sigma_v : \begin{cases} \dot{x}_v = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x_v + \begin{bmatrix} 0 \\ 6.4013 \times 10^7 \end{bmatrix} u_v \\ y_v = [1 \ 0] x_v \end{cases} \quad (8.12)$$

In order to minimize the effects of the resonance modes in the microactuator, we introduce the following filter in the feedforward path:

$$G_{\text{filter}}(s) = \frac{N_{\text{filter}}(s)}{D_{\text{filter}}(s)} \quad (8.13)$$

where

$$\begin{aligned} N_{\text{filter}}(s) = & 1.164 \times 10^5 s^5 + 1.177 \times 10^9 s^4 + 3.99 \times 10^{14} s^3 \\ & + 3.681 \times 10^{18} s^2 + 3.112 \times 10^{23} s + 2.617 \times 10^{27} \end{aligned} \quad (8.14)$$

and

$$\begin{aligned} D_{\text{filter}}(s) = & 0.0002449 s^7 + 47.92 s^6 + 3.99 \times 10^6 s^5 + 1.953 \times 10^{11} s^4 \\ & + 5.71 \times 10^{15} s^3 + 1.058 \times 10^{20} s^2 + 8.387 \times 10^{23} s + 1.309 \times 10^{27} \end{aligned} \quad (8.15)$$

It is a combination of a notch filter and a low-pass filter. The frequency responses of the filter in Equations 8.13–8.15 and the compensated microactuator system are given, respectively, in Figures 8.5 and 8.6. Clearly, the dynamics of the microactuator can then be safely approximated by a static equation:

$$y_m = 0.5 u_m \quad (8.16)$$

Note that the control input to the VCM actuator is constrained within ± 3 V, which is the same as the one used in Chapters 6 and 7, and the control input to the microactuator has to be kept within ± 2 V. Moreover, the maximum displacement that can be generated by the microactuator cannot exceed 1 μm , which is equivalent to one track pitch of the HDD servo system. Clearly, we are having a plant with both sensor and actuator nonlinearities.

As in Chapter 6, we present in the following three different types of controllers for the VCM actuator, *i.e.* PID, RPT and CNF control.

1. The PID control law (with the notch filter) for the VCM actuator is identical to that given in Chapter 6. It is given by

$$\tilde{u}(k) = \frac{0.50495z^2 - 0.99147z + 0.48653}{z^2 - 1.09091z + 0.09091} \cdot \frac{0.2696}{z - 0.7304} \cdot [r(k) - \hat{y}_v(k)] \quad (8.17)$$

and

$$u_v(k) = G_{\text{notch}}(s) \Big|_{\text{ZOH}} \cdot \tilde{u}(k) \quad (8.18)$$

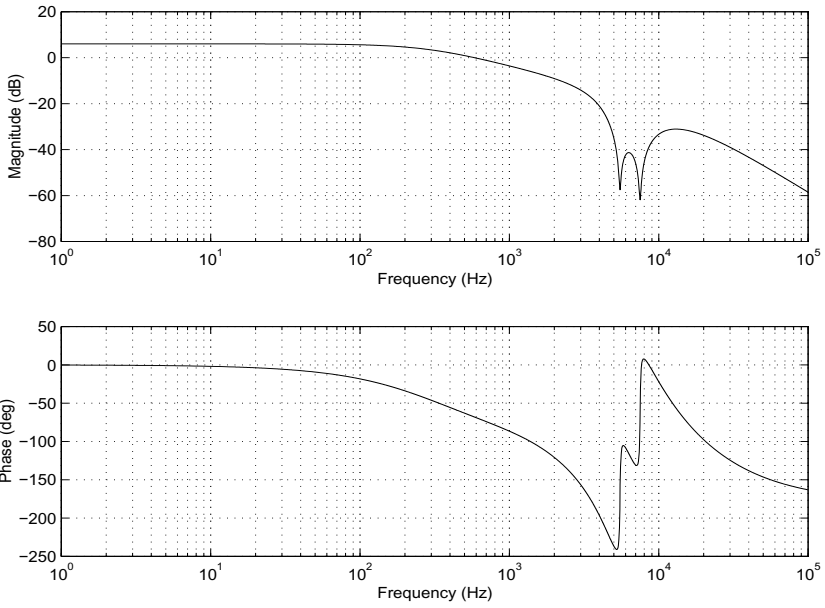


Figure 8.5. Frequency response of the microactuator filter

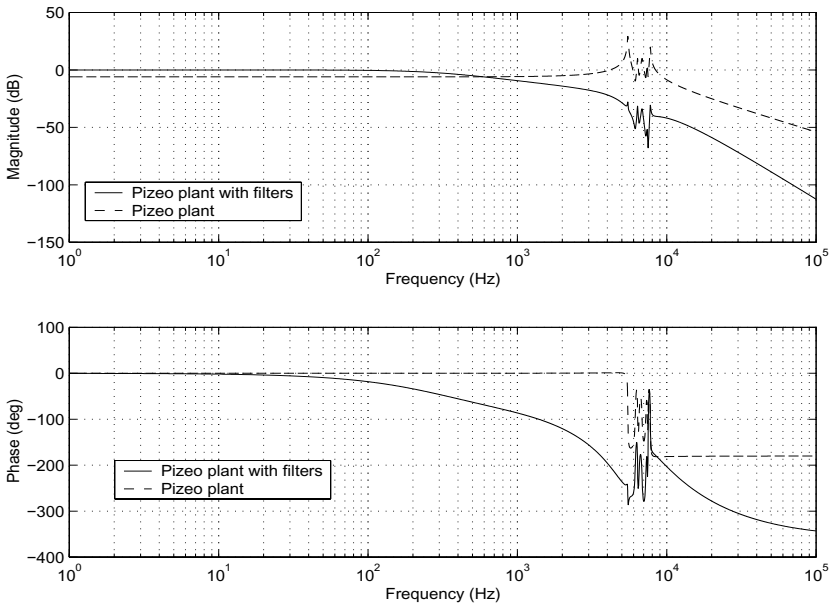


Figure 8.6. Frequency response of the microactuator with the filter

2. The RPT controller is slightly different from the one given in Chapter 6 and is given by

$$\dot{x}_v = \begin{bmatrix} 0 & 0 \\ -696.40 & -7704.5 \end{bmatrix} x_v + \begin{bmatrix} -1 & 1 \\ 70.009 & -378.19 \end{bmatrix} \begin{pmatrix} r \\ \hat{y}_v \end{pmatrix} \quad (8.19)$$

$$\tilde{u} = [-1.0879 \quad -5.7871] x_v + [0.10937 \quad -0.34085] \begin{pmatrix} r \\ \hat{y}_v \end{pmatrix} \quad (8.20)$$

and

$$u_v = G_{\text{notch}}(s) \cdot \tilde{u} \quad (8.21)$$

where $G_{\text{notch}}(s)$ is the transfer function of the notch filter given in Equation 6.9.

3. Finally, the CNF control law is given as follows:

$$\begin{pmatrix} \dot{x}_i \\ \dot{x}_v \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -4000 \end{bmatrix} \begin{pmatrix} x_i \\ x_v \end{pmatrix} - \begin{bmatrix} -1 \\ 1.6 \times 10^7 \end{bmatrix} \hat{y}_v + \begin{bmatrix} 0 \\ 6.4013 \times 10^7 \end{bmatrix} \tilde{u} - \begin{pmatrix} r \\ 0 \end{pmatrix} \quad (8.22)$$

$$\begin{aligned} \tilde{u} = & \rho(\hat{y}_v - r) [0.7555 \quad 0.07555 \quad 1.2597 \times 10^{-4}] \begin{pmatrix} x_i \\ \hat{y}_v - r \\ x_v + 4000\hat{y}_v \end{pmatrix} \\ & - [7.5549 \times 10^{-5} \quad 0.07555 \quad 1.0306 \times 10^{-5}] \begin{pmatrix} x_i \\ \hat{y}_v - r \\ x_v + 4000\hat{y}_v \end{pmatrix} \end{aligned} \quad (8.23)$$

with

$$\rho(\hat{y}_v - r) = -0.0791 \left| e^{-|\hat{y}_v - r|} - 0.3679 \right| - 0.35 \quad (8.24)$$

and

$$u_v = G_{\text{notch}}(s) \cdot \tilde{u} \quad (8.25)$$

8.4 Simulation and Implementation Results

We now present the simulation and implementation results of the servo systems obtained in the previous section. Simulations are done with the continuous-time plant models of Equations 8.1 to 8.11 with controllers being discretized with the ZOH method with a sampling frequency of 20 kHz. Implementations are carried out at a sampling frequency of 20 kHz. The results of the dual-stage actuated HDD servo systems will then be compared with those of the servo systems with a single-stage actuator. The latter are done on the same drive by keeping the microactuator inactive throughout the whole implementation process. The controller parameters for the single-stage actuated systems are identical to those given in Chapter 6.

8.4.1 Track-following Test

The simulated output responses of the servo systems with $r = 1 \mu\text{m}$ are given in Figure 8.7. Once again, the RPT and CNF control systems yield much better performances compared to that of the PID control system. For comparison, we present the responses and control signals of each individual approach for the dual-stage actuated servo system together with those of its single-stage counterpart in Figures 8.8 to 8.10, respectively. The implementation results of the corresponding servo systems are then presented in Figures 8.11 to 8.13. Finally, we summarize the resulting settling time of each control system in Table 8.1.

Table 8.1. Performances of the servo systems

(a) Simulation results

	Settling time (ms)		
	PID Control	RPT Control	CNF Control
Single-stage	3.10	0.95	0.80
Dual-stage	2.15	0.40	0.40

(b) Implementation results

	Settling time (ms)		
	PID Control	RPT Control	CNF Control
Single-stage	2.70	1.05	0.85
Dual-stage	1.80	0.30	0.30

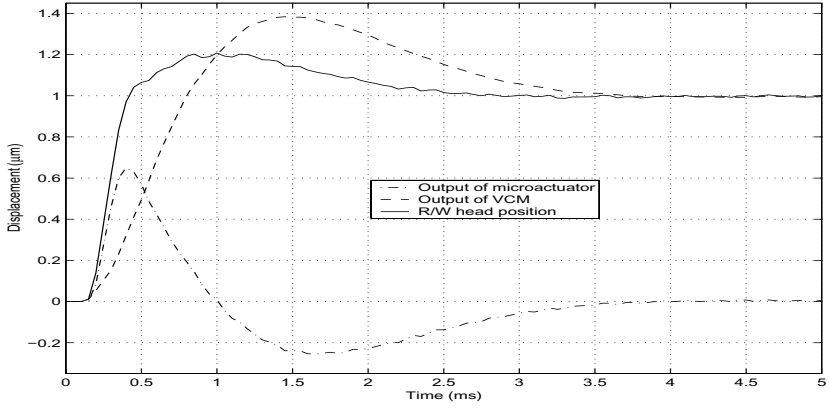
8.4.2 Frequency-domain Test

To verify the frequency-domain properties of our designs, Figures 8.14 to 8.19, respectively, give the Bode plot, the Nyquist plot, and the sensitivity and complementary sensitivity functions, as well as the closed-loop transfer functions (from the reference input r to the controlled output $h = y$) of the resulting control systems. For the CNF design, once again, its frequency-domain functions are calculated at the steady-state situation for which the nonlinear gain function $\rho(e)$ has approached its final constant value. The results show that all these designs meet the frequency-domain design specifications.

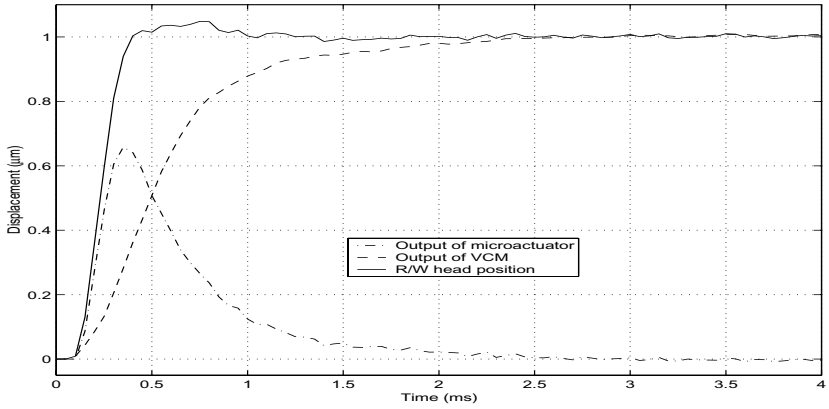
8.4.3 Runout Disturbance Test

As in the previous chapters, we artificially add a runout disturbance

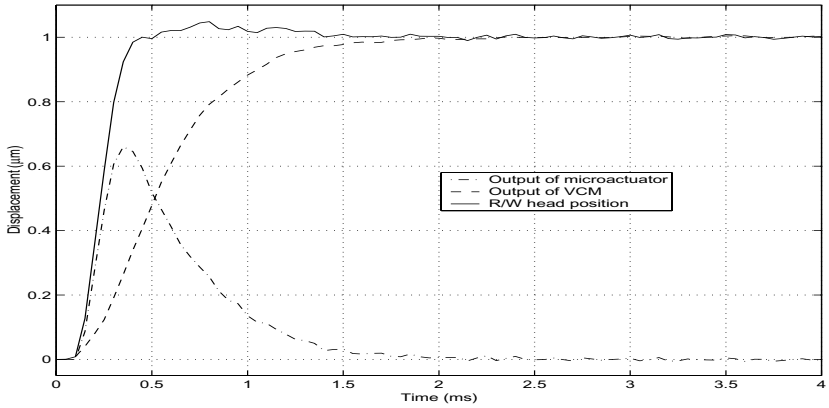
$$w(t) = 0.5 + 0.1\cos(100\pi t) + 0.05\sin(220\pi t) + 0.02\sin(440\pi t) + 0.01\sin(880\pi t) \quad (8.26)$$



(a) PID control

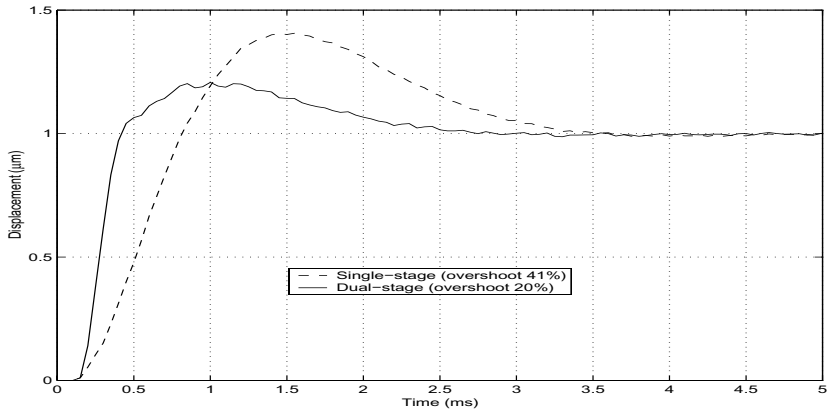


(b) RPT control

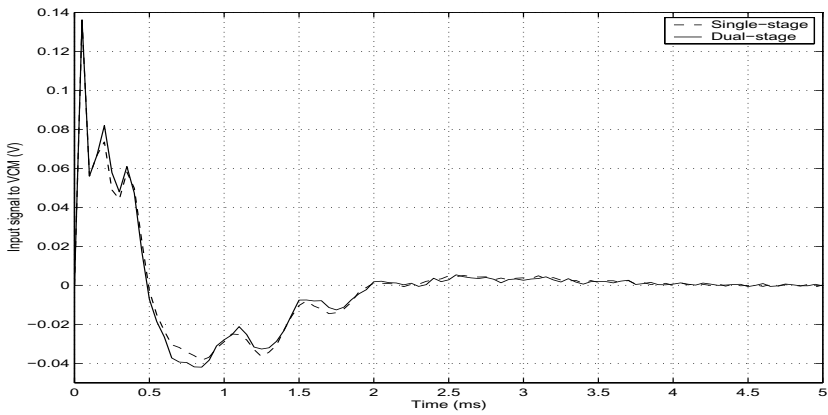


(c) CNF control

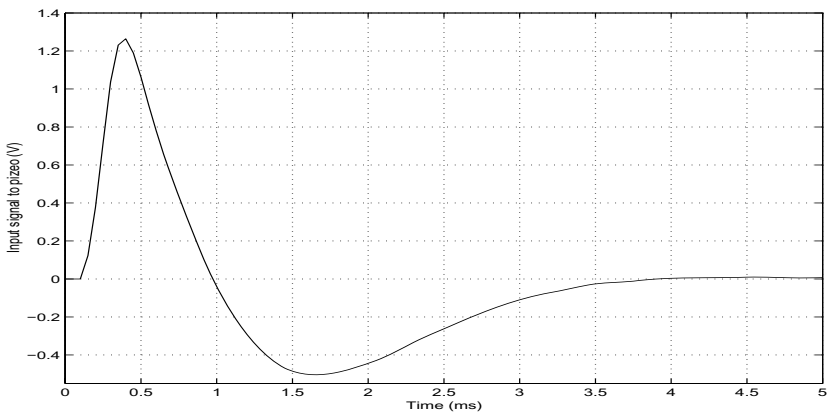
Figure 8.7. Simulation: output responses of the dual-stage actuated systems



(a) Output responses

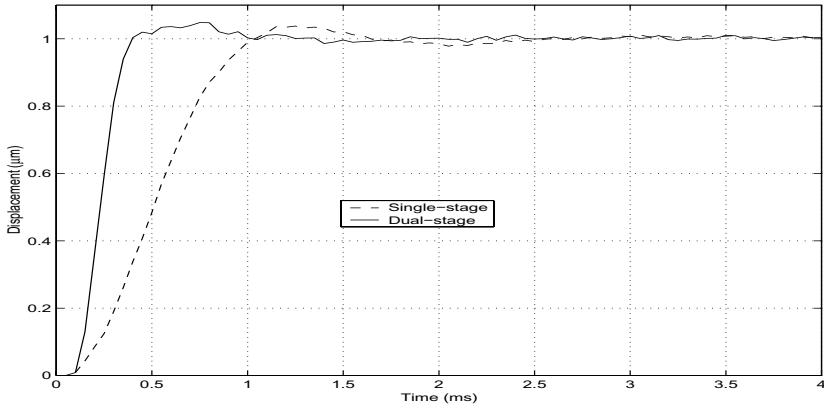


(b) Control signals to VCM actuator

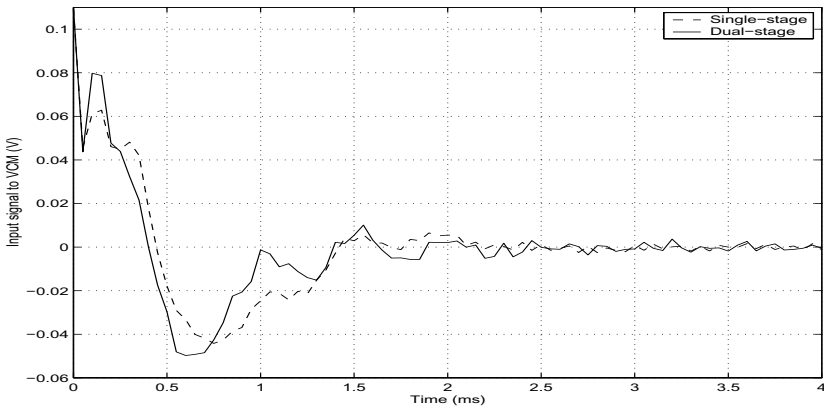


(c) Control signal to microactuator

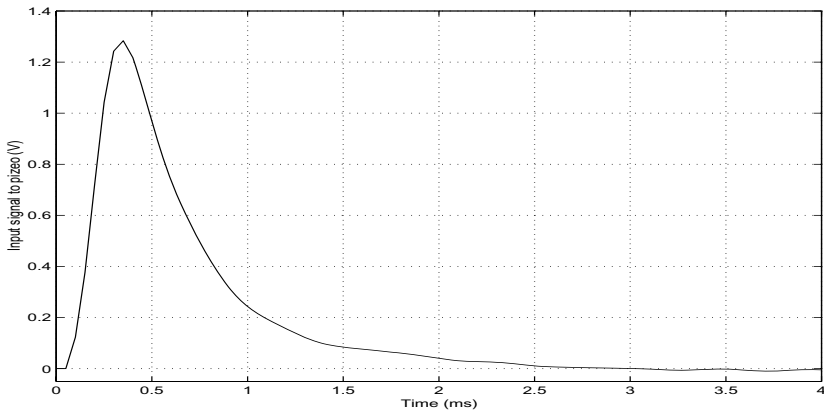
Figure 8.8. Simulation results of single- and dual-stage servo systems with PID control



(a) Output responses

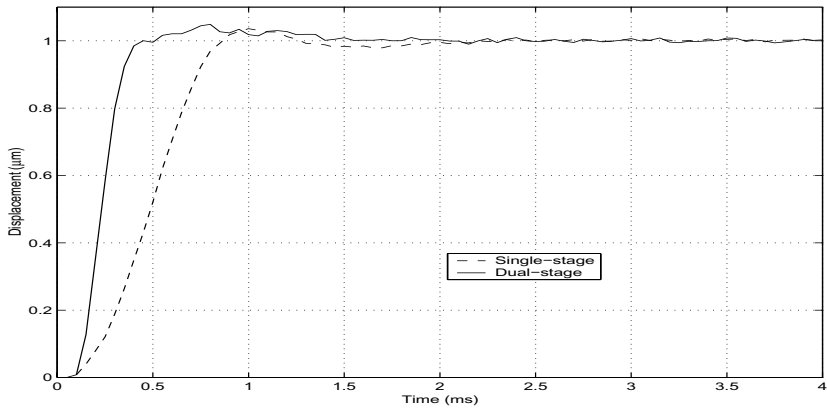


(b) Control signals to VCM actuator

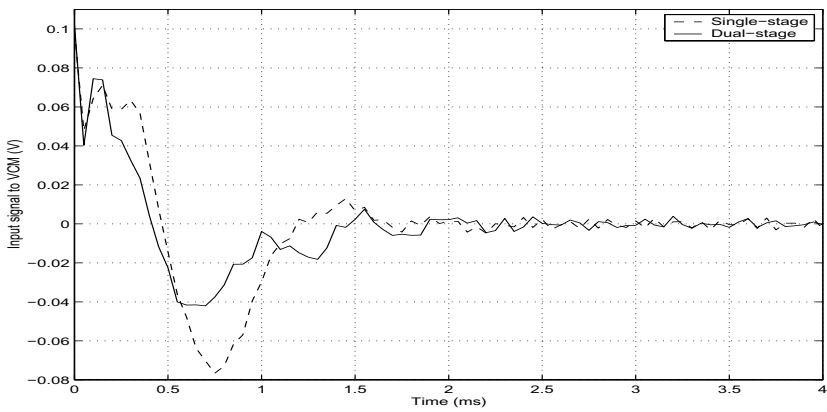


(c) Control signal to microactuator

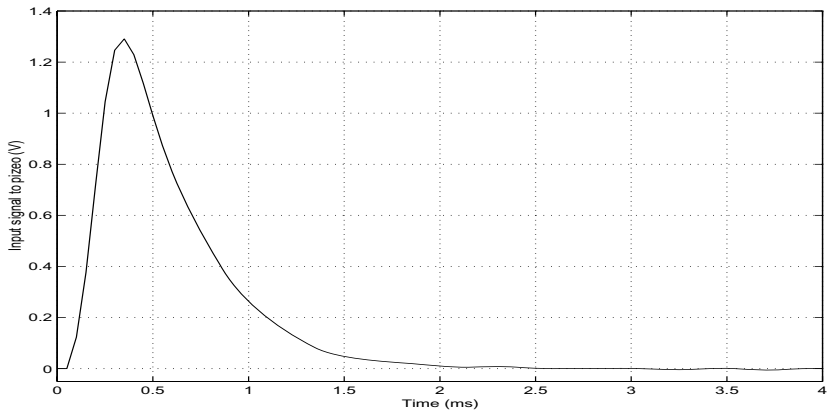
Figure 8.9. Simulation results of single- and dual-stage servo systems with RPT control



(a) Output responses

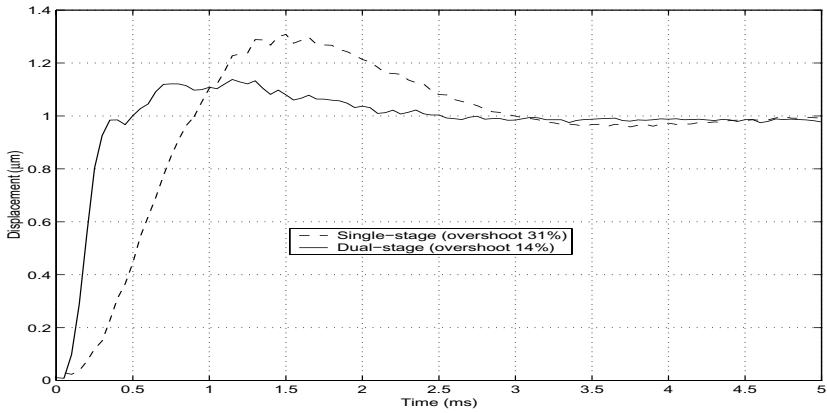


(b) Control signals to VCM actuator

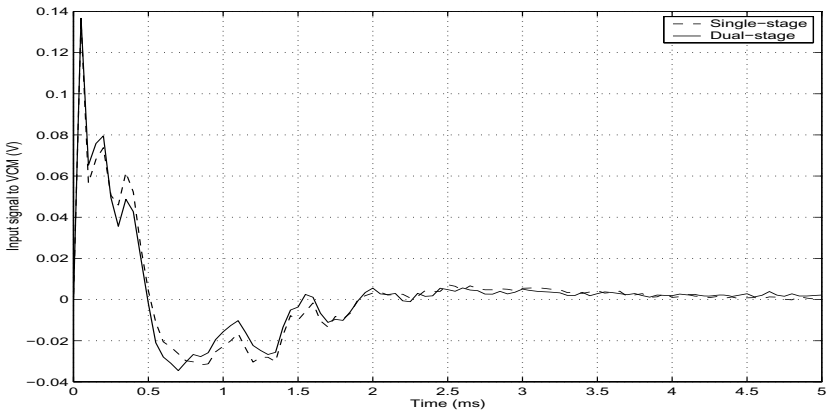


(c) Control signal to microactuator

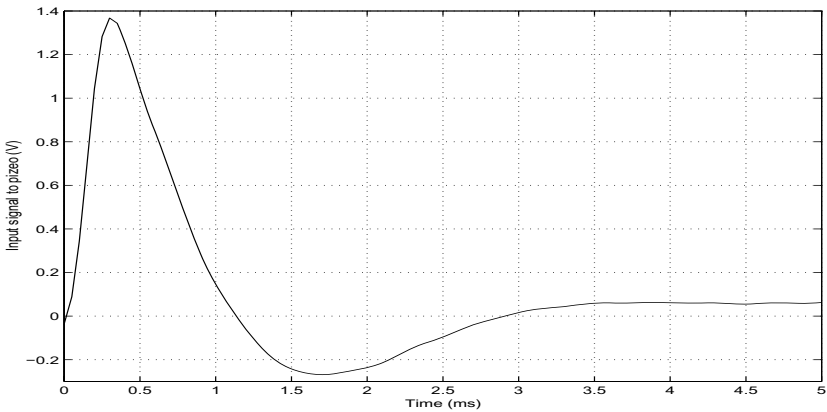
Figure 8.10. Simulation results of single- and dual-stage servo systems with CNF control



(a) Output responses

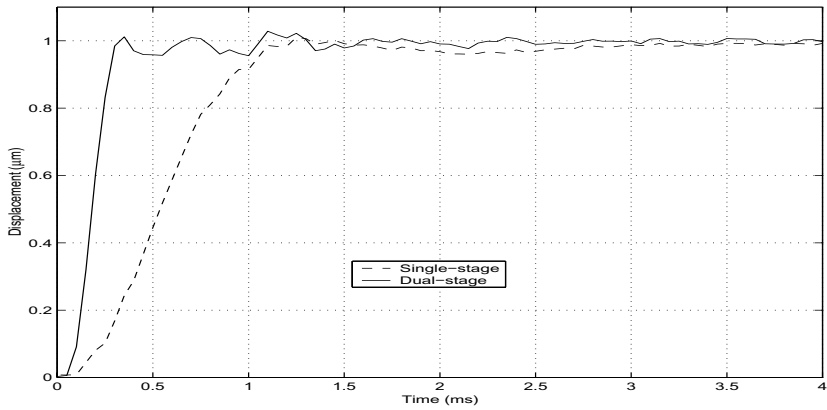


(b) Control signals to VCM actuator

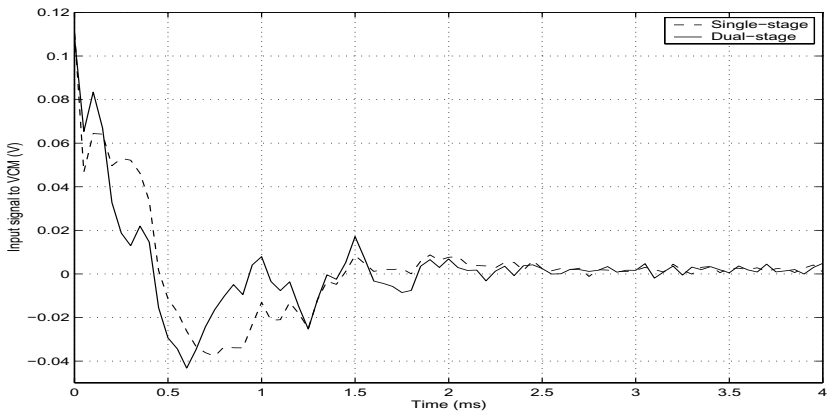


(c) Control signal to microactuator

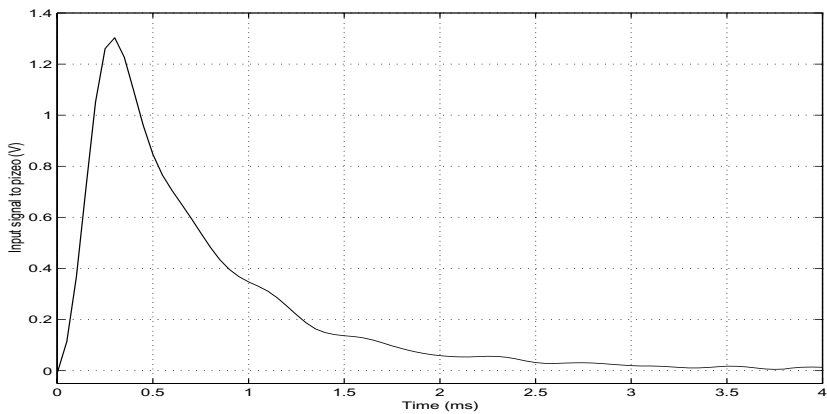
Figure 8.11. Implementation results of single- and dual-stage servo systems with PID control



(a) Output responses

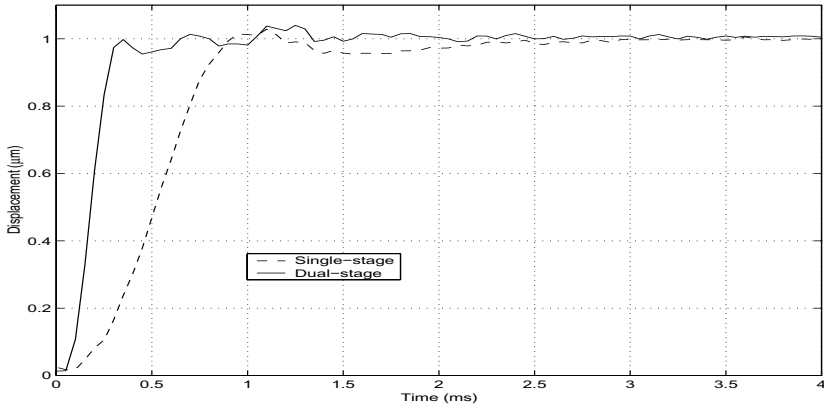


(b) Control signals to VCM actuator

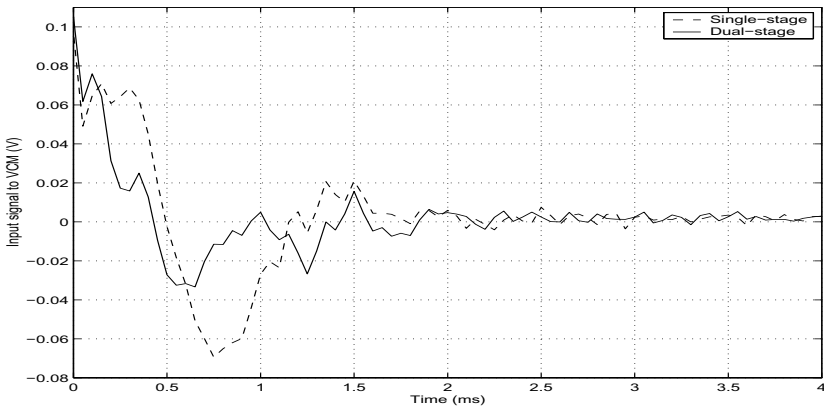


(c) Control signal to microactuator

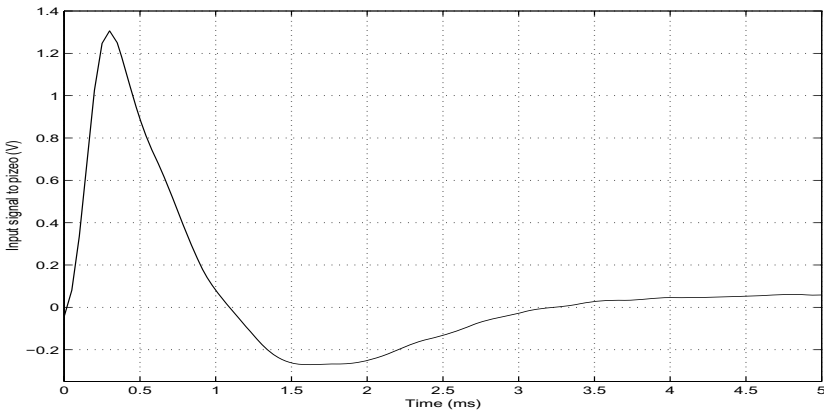
Figure 8.12. Implementation results of single- and dual-stage servo systems with RPT control



(a) Output responses

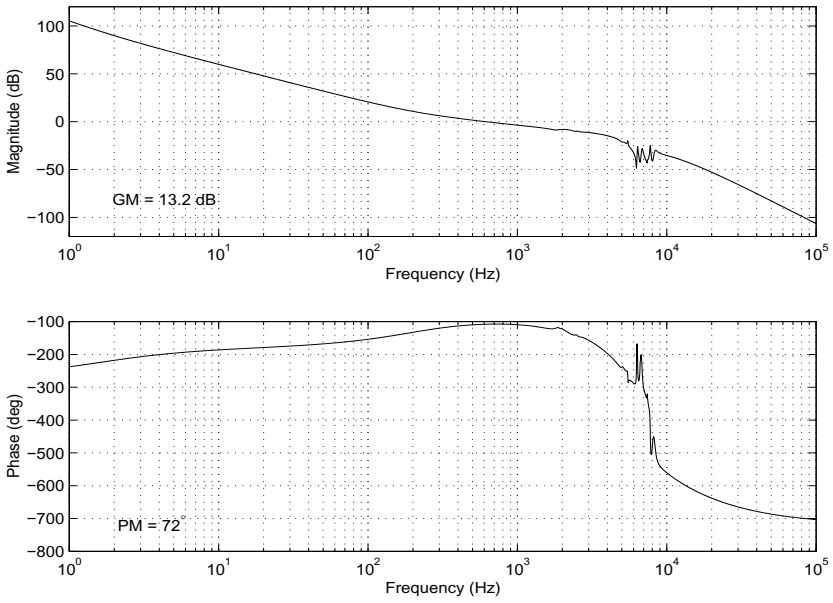


(b) Control signals to VCM actuator

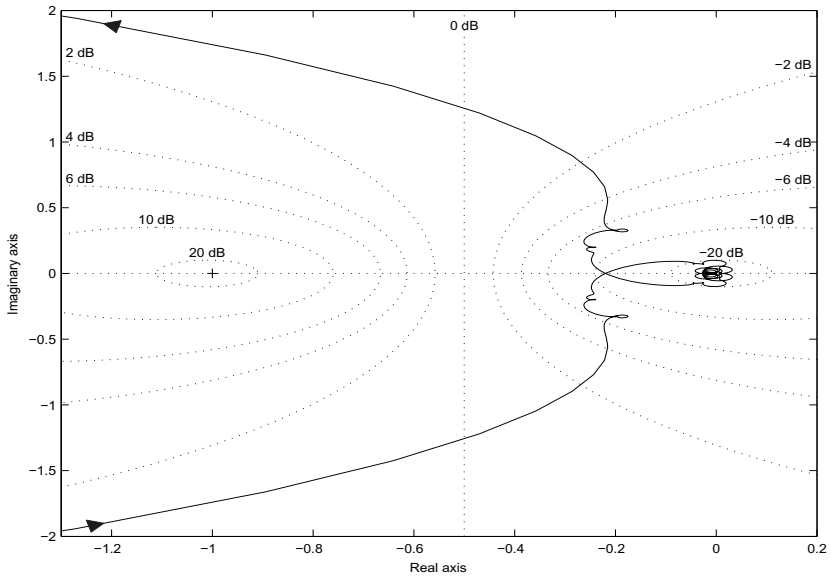


(c) Control signal to microactuator

Figure 8.13. Implementation results of single- and dual-stage servo systems with CNF control

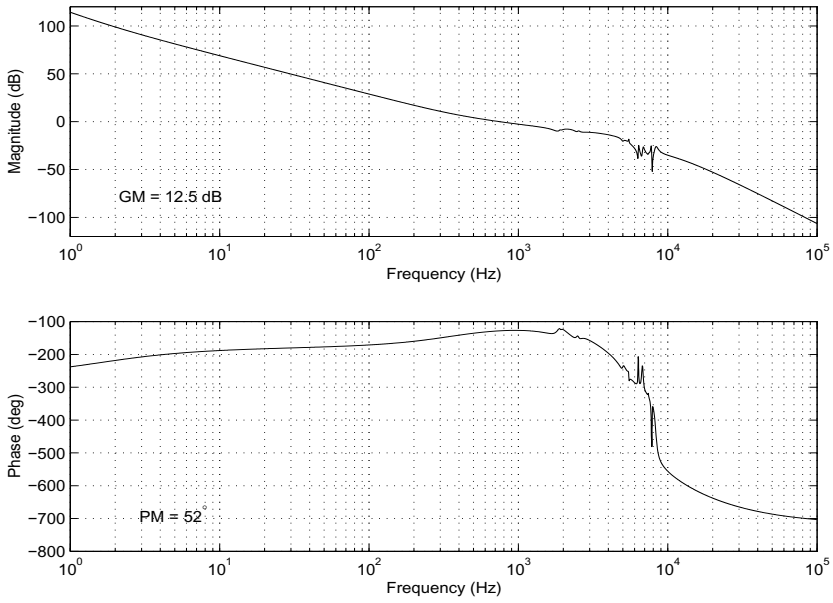


(a) Bode plot

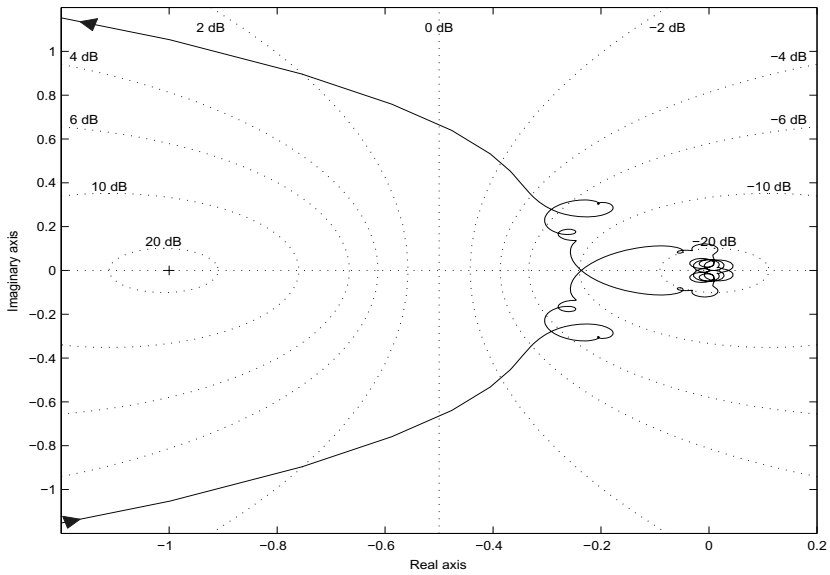


(b) Nyquist plot

Figure 8.14. Bode and Nyquist plots of the PID control system

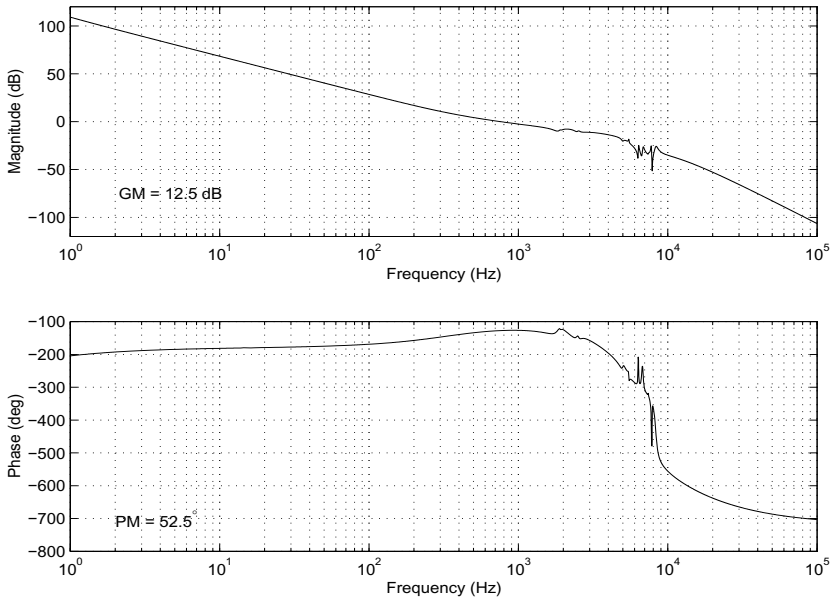


(a) Bode plot

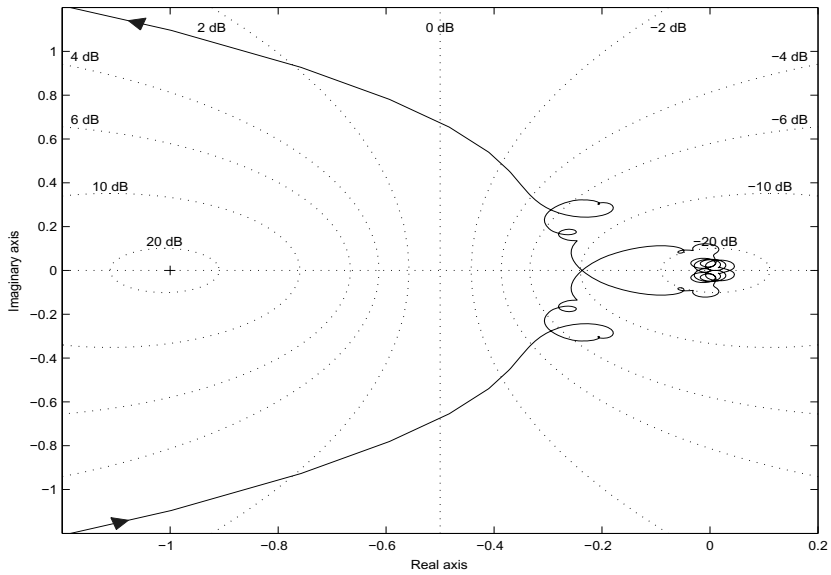


(b) Nyquist plot

Figure 8.15. Bode and Nyquist plots of the RPT control system

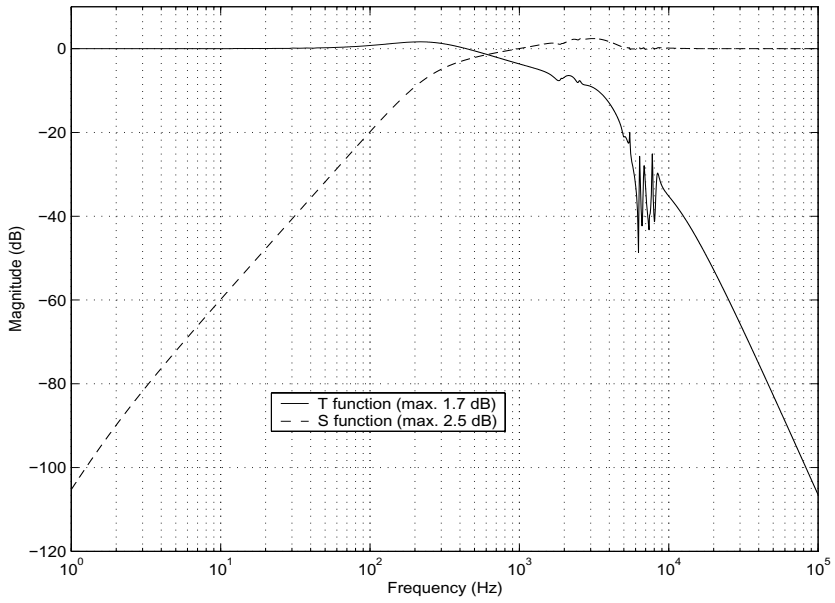


(a) Bode plot

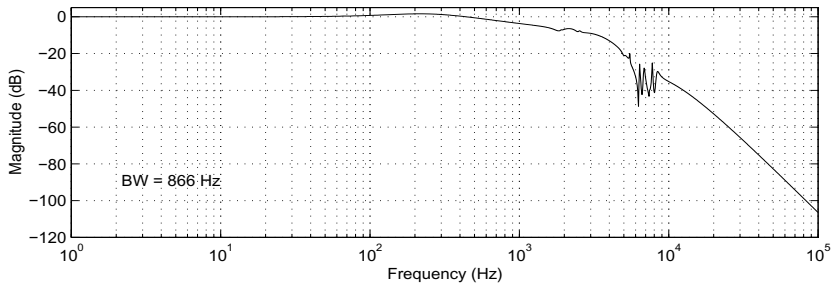


(b) Nyquist plot

Figure 8.16. Bode and Nyquist plots of the CNF control system

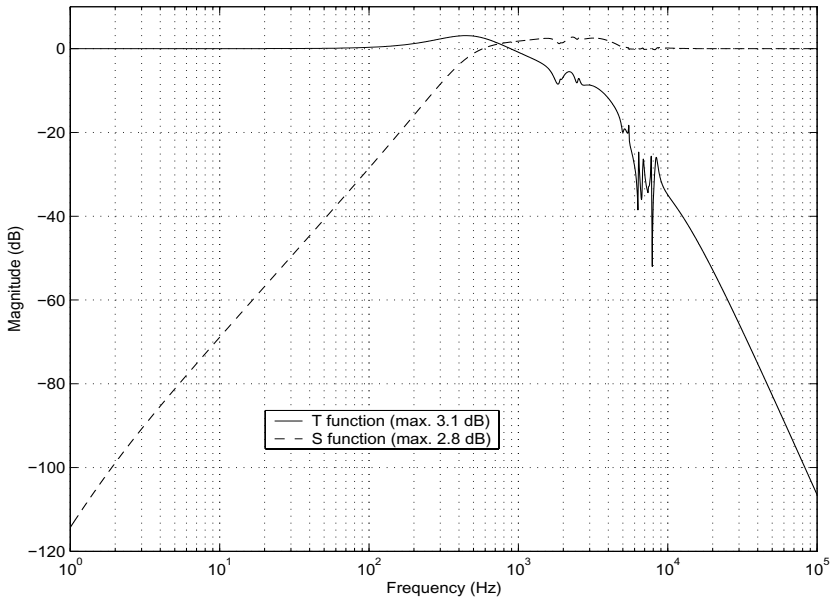


(a) Sensitivity and complementary sensitivity functions

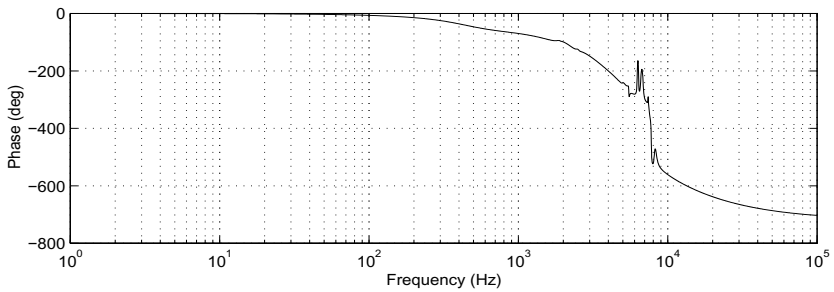
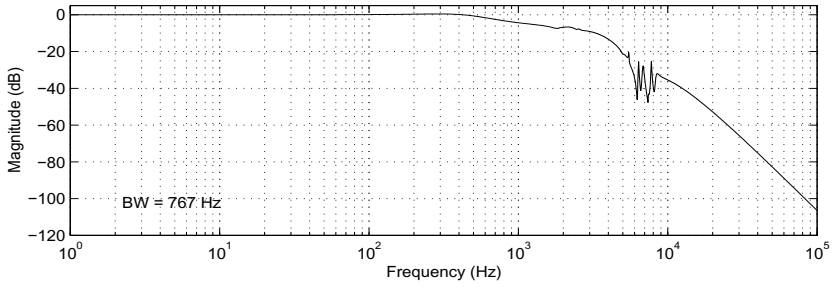


(b) Closed-loop response

Figure 8.17. Sensitivity functions and closed-loop transfer function of the PID control system

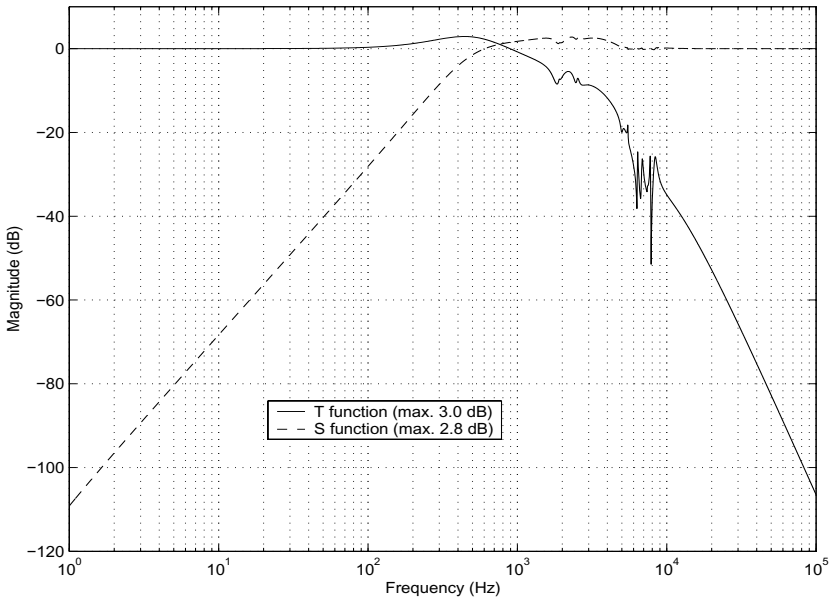


(a) Sensitivity and complementary sensitivity functions

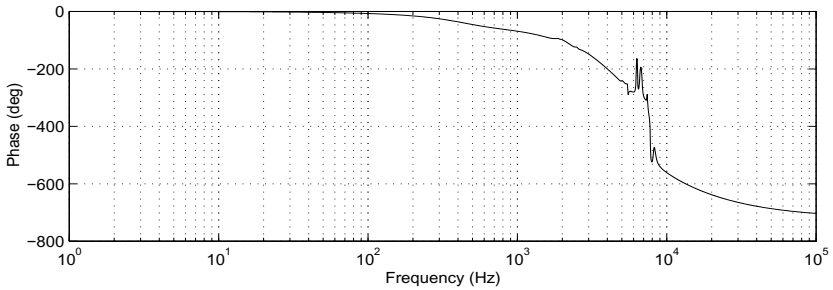
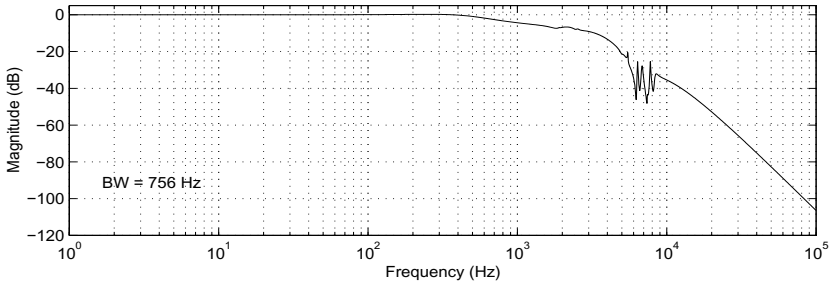


(b) Closed-loop response

Figure 8.18. Sensitivity functions and closed-loop transfer function of the RPT control system



(a) Sensitivity and complementary sensitivity functions



(b) Closed-loop response

Figure 8.19. Sensitivity functions and closed-loop transfer function of the CNF control system

which is the same as those in the previous chapters, to our servo systems. The implementation results of the corresponding responses are respectively shown in Figures 8.20 to 8.22.

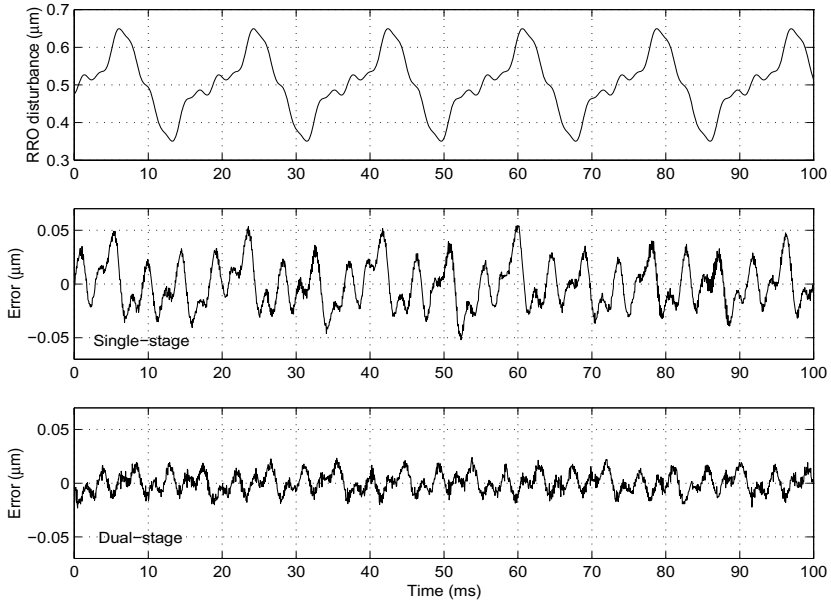


Figure 8.20. Implementation results: Responses to a runout disturbance (PID)

8.4.4 Position Error Signal Test

Lastly, as were done in Chapters 6 and 7, we conduct the PES tests for the complete single- and dual-stage actuated servo systems. The results, *i.e.* the histograms of the PES tests, are given in Figures 8.23 to 8.25. The $3\sigma_{\text{pes}}$ values of the PES tests, which are a measure of track misregistration (TMR) in HDDs and that are closely related to the maximum achievable track density, are summarized in Table 8.2.

Table 8.2. The $3\sigma_{\text{pes}}$ values of the PES tests

	$3\sigma_{\text{pes}}$ (μm)		
	PID control	RPT control	CNF control
Single-stage	0.0615	0.0375	0.0288
Dual-stage	0.0273	0.0204	0.0195

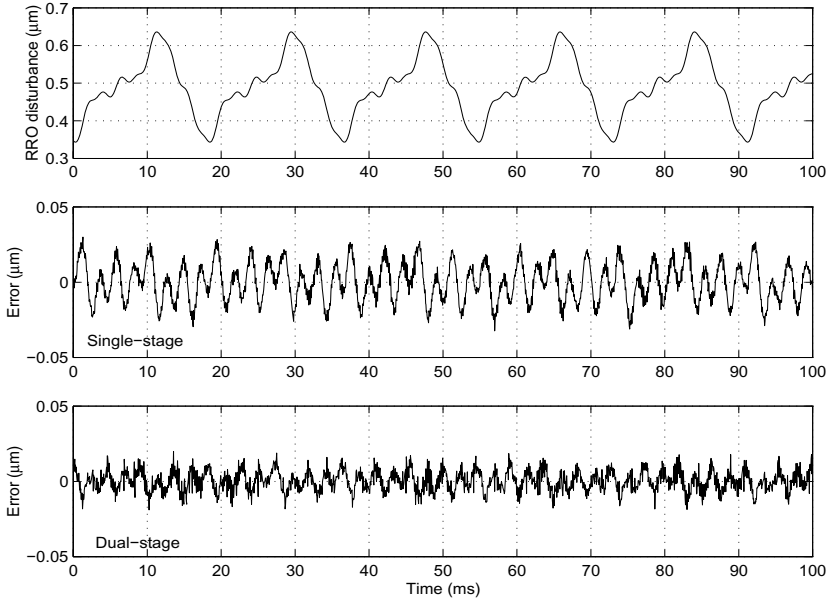


Figure 8.21. Implementation results: Responses to a runout disturbance (RPT)

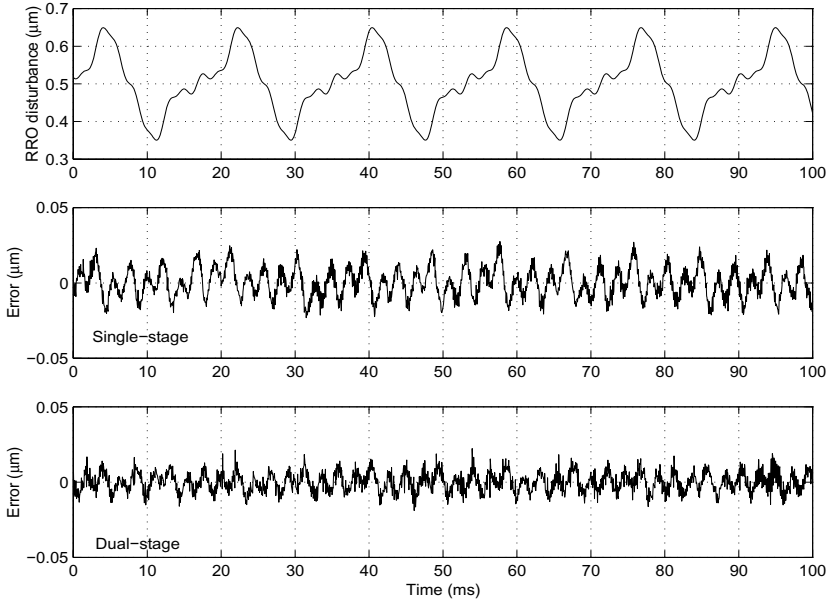


Figure 8.22. Implementation results: Responses to a runout disturbance (CNF)

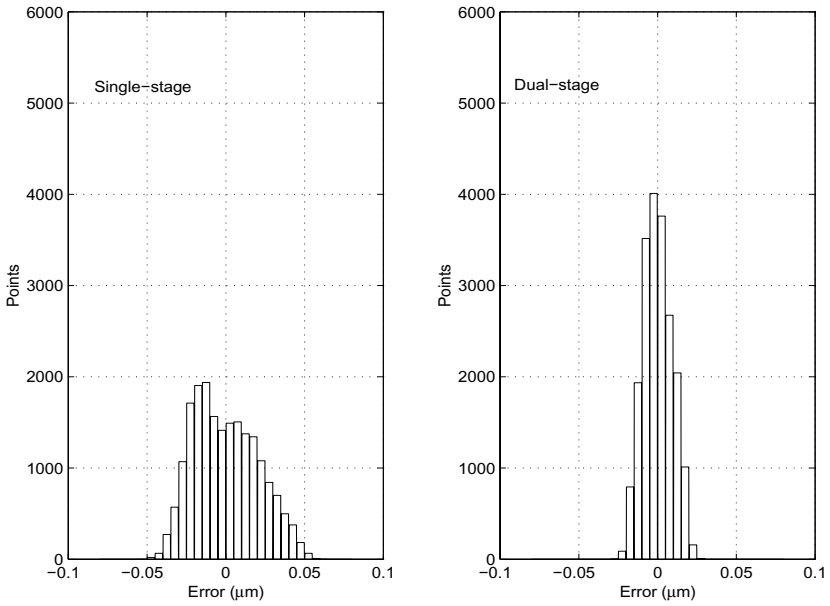


Figure 8.23. Implementation results: PES test histograms (PID)

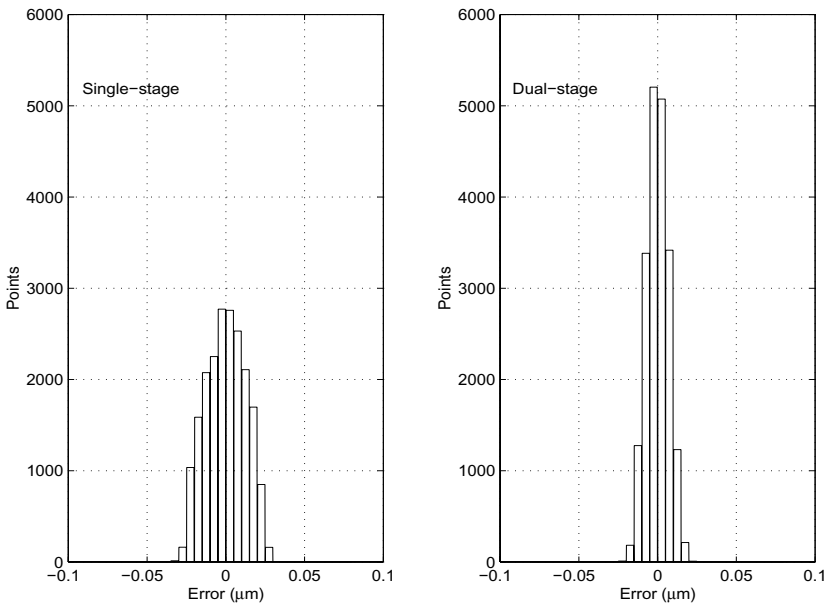


Figure 8.24. Implementation results: PES test histograms (RPT)

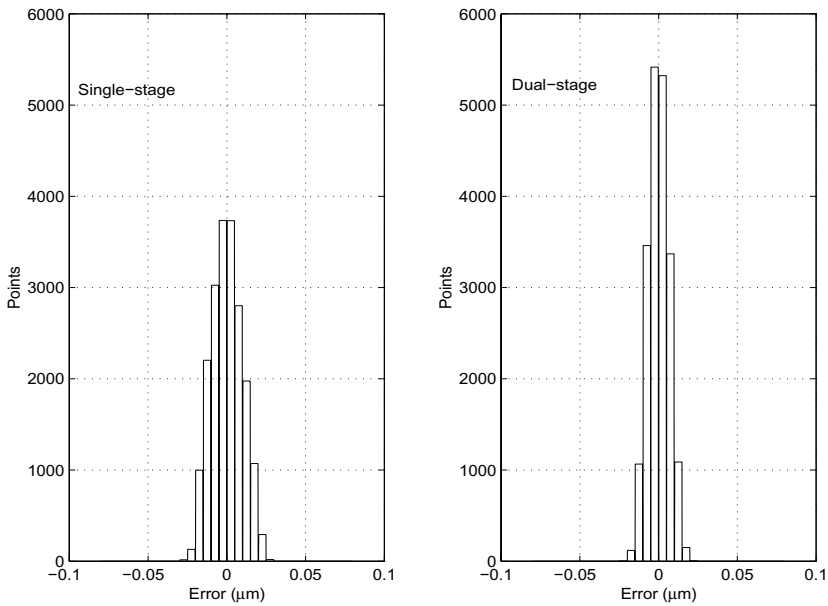


Figure 8.25. Implementation results: PES test histograms (CNF)

It can be easily observed from the results obtained that the dual-stage actuated servo systems do provide a faster settling time and better positioning accuracy compared with those of the single-stage actuated counterparts. The improvement in the track-following stage turns out to be very noticeable. This was actually the original purpose of introducing the microactuator into HDD servo systems. However, we personally feel that the price we have paid (*i.e.* by adding an expensive and delicate piezoelectric actuator to the system) for such an improvement is too high.

Modeling and Design of a Microdrive System

9.1 Introduction

Chapters 6 to 8 focus on the design of single- and dual-stage actuated hard drive servo systems. The hard drives considered are those used in normal desktop computers. As mentioned earlier in the introduction chapter, microdrives have become popular these days because of high demand from many new applications. Many factors such as frictional forces and nonlinearities, which are negligible for normal drives and thus ignored in servo systems given in Chapters 6 to 8, emerge as critical issues for microdrives. It can be observed that nonlinearities from friction in the actuator rotary pivot bearing and data flex cable in the VCM actuator (see Figure 9.1) generate large residual errors and deteriorate the performance of head positioning of HDD servo systems, which is much more severe in the track-following stage when the R/W head is moving from the current track to its neighborhood tracks. The desire to fully understand the behaviors of nonlinearities and friction in microdrives is obvious. Actually, this motivates us to carry out a complete study and modeling of friction and nonlinearities for the VCM-actuated HDD servo systems.

Friction is hard nonlinear and may result in residual errors, limit cycles and poor performance (see, *e.g.*, [168–171]). Friction exists in almost all servomechanisms, behaves in features of the Stribeck effect, hysteresis, stiction and varying break-away force, occurs in all mechanical systems and appears at the physical interface between two contact surfaces moving relative to each other. The features of friction have been extensively studied (see, *e.g.*, [168–174]), but there are significant differences among diverse systems. There has been a significantly increased interest in friction in the industry, which is driven by strong engineering needs in a wide range of industries and availability of both precise measurement and advanced control techniques.

The HDD industry persists in the need for companies to come up with devices that are cheaper and able to store more data and retrieve or write to them at faster speed. Decreasing the HDD track width is a feasible idea to achieve these objectives. But, the presence of friction in the rotary actuator pivot bearing results in large residual errors and high-frequency oscillations, which may produce a larger positioning error signal to hold back the further decreasing of the track width and to degrade



Figure 9.1. An HDD with a VCM actuator

the performance of the servo systems. This issue becomes more noticeable for small drives and is one of the challenges to design head positioning servo systems for small HDDs. Much effort has been put into the research on mitigation of the friction in the pivot bearing in the HDD industry in the last decade (see, *e.g.*, [8, 175–178]). It is still ongoing in the disk drive industry (see, *e.g.*, [69, 179, 180]).

Diverse modeling methods had been proposed (see, *e.g.*, [59, 69]) based on linear systems, where nonlinearities of plants are assumed to be tiny and can be neglected. As such, these methods cannot be directly applied to model plants with significant nonlinearities. Instead, in the first part of this chapter, we utilize the physical effect approach given in Chapter 2 to determine the structures of nonlinearities and friction associated with the VCM actuator in a typical HDD servo system. This is done by carefully examining and analyzing physical effects that occur in or between electromechanical parts. Then, we employ a Monte Carlo process (see Chapter 2) to identify the parameters in the structured model. We note that Monte Carlo methods are very effective in approximating solutions to a variety of mathematical problems, for which their analytical solutions are hard, if not impossible, to determine. Our simulation and experimental results show that the identified model of friction and nonlinearities using such approaches matches very well the behavior of the actual system.

The second part of this chapter focuses on the controller design for the HDD servo system. Our philosophy of designing servo systems is rather simple. Once the model of the friction and nonlinearities of the VCM actuator is obtained, we will try to cancel as much as possible all these unwanted elements in the servo system. As it is impossible to have perfect models for friction and nonlinearities, a perfect cancellation of these elements is unlikely to happen in the real world. We then formulate our design by treating the uncompensated portion as external disturbances. The PID

and RPT control techniques of Chapter 3 and the CNF control technique of Chapter 5 are to be used to carry out our servo system design. We note that some of the results presented in this chapter have been reported earlier in [138].

9.2 Modeling of the Microdrive Actuator

The physical structure of a typical VCM actuator is shown in Figure 9.2. The motion of the coil is driven by a permanent magnet similar to typical DC motors. The stator of the VCM is built of a permanent magnet. The rotor consists of a coil, a pivot and a metal arm on which the R/W head is attached. A data flex cable is connected with the R/W head through the metal arm to transfer data read from or written to the HDD disc via the R/W head. Typically, the rotor has a deflected angle, α in rad, ranging up to 0.5 rad in commercial disk drives. We are particularly interested in the modeling of the friction and nonlinearities for the actuator in the track-following stage, in which the R/W head movement is within the neighborhood of its current track and thus $\alpha \ll 1$. An IBM microdrive (DMDM-10340) is used throughout for illustration.

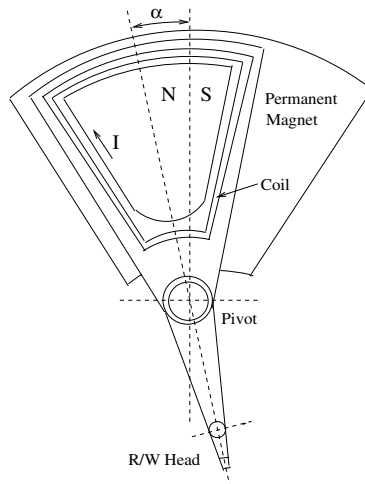


Figure 9.2. The mechanical structure of a typical VCM actuator

9.2.1 Structural Model of the VCM Actuator

We first adopt the physical effect analysis of Chapter 2 to determine the structures of nonlinearities in the VCM actuator. It is to analyze the effects between the components of the actuator, such as the stator, rotor and support plane as well as the VCM driver. The VCM actuator is designed to position the R/W head fast and precisely

onto the target track, and is driven by a VCM driver, a full bridge power amplifier, which converts an input voltage into an electric current. The electrical circuit of a typical VCM driver is shown in Figure 9.3, where Z represents the coil of the VCM actuator and the external input voltage is exerted directly into the VCM driver to drive the coil.

In order to simplify our analysis, we assume that the physical system has the following properties: i) the permanent magnet is constant; and ii) the coil is assembled strictly along the radius and concentric circle of the pivot; Furthermore, we assume that the friction of a mechanical object consists of Coloumb friction and viscous damping, and is characterized by a typical friction function as follows:

$$f(F_N, F_e, v, \mu_d, \mu_s, \mu_v) = \begin{cases} -\mu_d F_N \operatorname{sgn}(v) - \mu_v v, & v \neq 0 \\ -F_e, & v = 0, |F_e| \leq \mu_s F_N \\ -\mu_s F_N \operatorname{sgn}(F_e), & v = 0, |F_e| > \mu_s F_N \end{cases} \quad (9.1)$$

where f is the friction force, F_N is the normal force, *i.e.* the force perpendicular to the contacted surfaces of the objects, F_e is the external force applied to the object, v is the relative moving speed between two contact surfaces, and $\mu_s F_N$ is the breakaway force. Furthermore, μ_d , μ_s and μ_v are, respectively, the dynamic, static and viscous coefficients of friction.

Through a detailed analysis of the VCM driver circuit in Figure 9.3, it is straightforward to verify that the relationship between the driver input voltage and the current and voltage of the VCM coil is given by

$$\frac{R_s + (1 + A_c)(1 + A_b)R}{R_s A_c (1 + A_c)(1 + A_b)R} \cdot \frac{R_2}{R_1} u = \frac{1 + s \left[R_3 + \frac{R_2}{A_c(1 + A_b)} \right] C_3}{1 + s R_3 C_3} I + \mathcal{E}(s) u_c \quad (9.2)$$

where

$$\mathcal{E}(s) = \frac{A_b}{(1 + A_c)(1 + A_b)R} \cdot \frac{1 + s \left[R_3 + \frac{R_2}{A_c A_b} \left(1 + \frac{R(1 + A_c)}{R_s} \right) \right] C_3}{1 + s R_3 C_3} \quad (9.3)$$

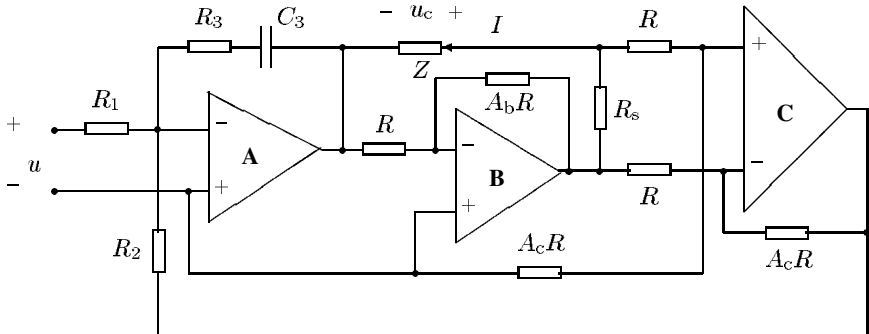


Figure 9.3. The electrical circuit of a typical VCM driver

u is the input voltage to the VCM driver, I and u_c are, respectively, the VCM coil current and voltage. For the IBM microdrive (DMDM-10340) used in our experiment, $R_1 = 8.2 \text{ k}\Omega$, $R_2 = 1 \text{ k}\Omega$, $R_3 = 270 \text{ k}\Omega$, $R_s = 1 \text{ }\Omega$, $R = 16 \text{ k}\Omega$, $C_3 = 270 \text{ pF}$, and the amplifier gains $A_c = 2$ and $A_b = 4.7$. For such a drive, we have

$$\mathcal{E}(s) = \frac{1.7178 \times 10^{-5}(1 + 1.45 \times 10^{-3}s)}{1 + 7.29 \times 10^{-5}s} \quad (9.4)$$

which has a magnitude response ranging from -95 dB (for frequency less than 110 Hz) to -69 dB (for frequency greater than 2.2 kHz), and

$$\frac{1 + s \left[R_3 + \frac{R_2}{A_c(1 + A_b)} \right] C_3}{1 + sR_3C_3} \approx 1, \quad \text{for all } s \quad (9.5)$$

Such a property generally holds for all commercial disk drives. As such, it is safe to approximate the relationship of u and I of the VCM driver as

$$I = \frac{R_2}{A_c R_1 R_s} u = k_{vd} u \quad (9.6)$$

For the IBM microdrive used in this work, $k_{vd} = 0.061 \text{ }\Omega^{-1}$.

Next, it is straightforward to derive that the torque T_m , relative to the center of the pivot and that moving anticlockwise is positive and produced by the permanent magnet B_{coil} in the coil with the electric current, is given by

$$T_m = (r_1^2 - r_2^2) B_{\text{coil}} n_{\text{coil}} k_{vd} u \quad (9.7)$$

r_1 and r_2 are, respectively, the outside and inside radius of the coil to the center of the pivot, and n_{coil} is the number of windings of the coil. The total external torque T_e applied to the VCM actuator is given as follows:

$$T_e = T_m - T_c(\alpha) \quad (9.8)$$

where $T_c(\alpha)$ is the spring torque produced by the data flex cable and is a function of the deflection angle α or the displacement of the R/W head. The friction torque in the VCM actuator comes from two major sources: One is the friction in the pivot bearing and the other is between the pivot bearing and the support plane. The friction torque in the pivot bearing can be characterized as

$$T_{f1} = f(F_{N1}, F_{e1}, r_3 \dot{\alpha}, \mu_{d1}, \mu_{s1}, \mu_{v1}) r_3 \quad (9.9)$$

where F_{e1} is the external force, μ_{d1} , μ_{s1} and μ_{v1} are the related friction coefficients as defined in Equation 9.1, r_3 is the radius of the pivot to its center, and

$$F_{N1} = |F_r + m_r(\dot{\alpha})^2| \quad (9.10)$$

is the normal force, which consists of the centrifugal force of the rotor and the diametrical force, F_r . Furthermore, m_r is a constant dependent on the mass distribution of the rotor, and

$$F_r = 2(r_1 - r_2)B_{\text{coil}}n_{\text{coil}}k_{\text{vd}}u\alpha \quad (9.11)$$

is the force along the radius of the pivot bearing produced in the coil by the permanent magnet.

The friction torque between the pivot bearing and the support plane can be characterized as:

$$T_{f2} = f(F_{N2}, F_{e2}, r_3\dot{\alpha}, \mu_{d2}, \mu_{s2}, \mu_{v2})r_3 \quad (9.12)$$

where F_{e2} is the external force, μ_{d2} , μ_{s2} and μ_{v2} are the related friction coefficients as defined in Equation 9.1, and

$$F_{N2} = |T_0|r_3^{-1} \quad (9.13)$$

is the normal force resulted from a static balance torque of the rotor, T_0 . Thus, the total friction torque T_f presented in the VCM actuator is given by

$$T_f = T_{f1} + T_{f2} = \begin{cases} -T_v, & \dot{\alpha} \neq 0 \\ -T_e, & \dot{\alpha} = 0 \text{ and } |T_e| \leq T_s \\ -T_s \operatorname{sgn}(T_e), & \dot{\alpha} = 0 \text{ and } |T_e| > T_s \end{cases} \quad (9.14)$$

where

$$T_v = \left[\mu_{d1} \left| 2(r_1 - r_2)B_{\text{coil}}n_{\text{coil}}k_{\text{vd}}u\alpha + m_r(\dot{\alpha})^2 \right| r_3 + \mu_{d2}|T_0| \right] \operatorname{sgn}(\dot{\alpha}) \\ + (\mu_{v1} + \mu_{v2})r_3^2\dot{\alpha} \quad (9.15)$$

and

$$T_s = 2\mu_{s1} |(r_1 - r_2)B_{\text{coil}}n_{\text{coil}}k_{\text{vd}}u_0\alpha_0| r_3 + \mu_{s2}|T_0| \quad (9.16)$$

is the breakaway torque, and where u_0 and α_0 are, respectively, the corresponding input voltage and the deflection angle for the situation when $\dot{\alpha} = 0$.

Lastly, it is simple to verify that the relative displacement of the R/W head, y , is given by

$$y = 2r_4 \sin \frac{\alpha}{2} \approx r_4\alpha, \quad |\alpha| \ll 1 \quad (9.17)$$

where r_4 is the length from the R/W head to the center of the pivot. Following Newton's law of motion, $J\ddot{\alpha} = T_e + T_f$, where J is the moment of inertia of the VCM rotor, we have

$$\ddot{y} = -\tilde{T}_c + \tilde{T}_f + bu \quad (9.18)$$

where

$$\tilde{T}_f = \begin{cases} -[d_1buy + d_2(\dot{y})^2 + d_3] \operatorname{sgn}(\dot{y}) - d_0\dot{y}, & \dot{y} \neq 0 \\ -\tilde{T}_e, & \dot{y} = 0, |\tilde{T}_e| \leq \tilde{T}_s \\ -\tilde{T}_s \operatorname{sgn}(\tilde{T}_e), & \dot{y} = 0, |\tilde{T}_e| > \tilde{T}_s \end{cases} \quad (9.19)$$

where

$$\left. \begin{aligned}
 \tilde{T}_e &= -\tilde{T}_c(y) + bu \\
 \tilde{T}_s &= d_4 b |u_0 y_0| + d_5 \\
 \tilde{T}_c &= J^{-1} r_4 T_c(\alpha) \\
 b &= J^{-1} r_4 (r_1^2 - r_2^2) B_{\text{coil}} n_{\text{coil}} k_{\text{vd}} \\
 d_0 &= J^{-1} (\mu_{\text{v1}} + \mu_{\text{v2}}) r_3^2 \\
 d_1 &= 2r_3 \mu_{\text{d1}} [r_4 (r_1 + r_2)]^{-1} \\
 d_2 &= r_3 m_r \mu_{\text{d1}} (J r_4)^{-1} \\
 d_3 &= J^{-1} r_4 \mu_{\text{d2}} |T_0| \\
 d_4 &= 2r_3 \mu_{\text{s1}} [r_4 (r_1 + r_2)]^{-1} \\
 d_5 &= J^{-1} r_4 \mu_{\text{s2}} |T_0|
 \end{aligned} \right\} \quad (9.20)$$

with u_0 and y_0 being, respectively, the corresponding input voltage and the displacement for the case when $\dot{y} = 0$. It is clear now that the expressions in Equations 9.18–9.20 give a complete structure of the VCM model including friction and nonlinearities from the data flex cable. Our next task is to identify all these parameters for the IBM microdrive (DMDM-10340).

9.2.2 Identification and Verification of Model Parameters

We proceed to identify the parameters of the VCM actuator model given in Equations 9.18–9.20. We note that there are results available in the literature (see, *e.g.*, [168]) to estimate friction parameters for typical DC motors for which both velocity and displacement are measurable and without constraint. Unfortunately, for the VCM actuator studied in this chapter, it is impossible to measure the time responses in constant-velocity motions and only the relative displacement of the R/W head is measurable. As such, the method of [168] cannot be adopted to solve our problems. Instead, we employ the popular Monte Carlo method of Chapter 2 (see also, [63–65]), which has been widely used in solving engineering problems and is capable of providing good numerical solutions.

First, it is simple to obtain from Equation 9.18 at a steady state when $\dot{y} = 0$ and $\ddot{y} = 0$,

$$|u_0 - b^{-1} \tilde{T}_c(y_0)| = b^{-1} |\tilde{T}_f| \leq b^{-1} \tilde{T}_s = d_4 |u_0 y_0| + b^{-1} d_5 \quad (9.21)$$

Our experimental results show that the right hand side of Equation 9.21 is very insignificant for small input signal u_0 and small displacement y_0 . This will be verified later when the model parameters are fully identified. Thus, we have

$$u_0 \approx b^{-1} \tilde{T}_c(y_0) := \tilde{T}_{\text{bc}}(y_0) \quad (9.22)$$

which is used to identify $\tilde{T}_c(y)$ or equivalently $T_c(\alpha)$, the spring torque produced by the data flex cable. Next, for the small neighborhood of (u_0, y_0) , we can rewrite the dynamic expression of Equation 9.18 as

$$\ddot{y} = - \left. \partial \tilde{T}_c / \partial y \right|_{y=y_0} (y - y_0) + b(u - u_0) + \tilde{T}_f \quad (9.23)$$

For small signals, and by omitting the nonlinear terms in \tilde{T}_f , the system dynamics in Equation 9.23 can be approximated by a second-order linear system with a transfer function from $u - u_0$ to $y - y_0$:

$$G_{y_0, u_0}(s) = \frac{b}{s^2 + d_0 s + \omega_p^2} \quad (9.24)$$

The natural frequency of the above transfer function (or roughly its peak frequency), ω_p , is given by

$$\omega_p = \left(\left. \frac{\partial \tilde{T}_c}{\partial y} \right|_{y=y_0} \right)^{\frac{1}{2}} \quad (9.25)$$

and its static gain is given by $k_{dc} = b\omega_p^{-2}$, which implies that

$$b = k_{dc}\omega_p^2 = 4\pi^2 k_{dc} f_p^2 \quad (9.26)$$

where $f_p = \omega_p/(2\pi)$. The expression in Equation 9.26 will be used to estimate the parameter b . More specifically, the parameters of the dynamic models of the VCM actuator will be identified using the following procedure:

1. The nonlinear characteristics of the data flex cable or equivalently $\tilde{T}_{bc}(y)$ will be initially determined using Equation 9.22 with a set of input signal, u_0 , and its corresponding output displacement, y_0 . It will be fine tuned later using the Monte Carlo method.
2. The parameter b will be initially computed using measured static gains and peak frequencies as in Equation 9.26, resulting from the dynamical responses of the actuator to a set of small input signals. Again, the identified parameter will be fine tuned later using the Monte Carlo method.
3. All system parameters will then be identified using the Monte Carlo method to match the frequency response to small input signal;
4. The high-frequency resonance modes of the actuator, which have not been included in either Equation 9.18 or 9.23, will be determined from frequency responses to input signals at high frequencies.

The above procedure will yield a complete and comprehensive model including nominal dynamics, high-frequency resonance modes, friction and nonlinearities of the VCM actuator. In our experiments, the relative displacement of the R/W head is the only measurable output and is measured using a laser Doppler vibrometer (LDV). A dynamic signal analyzer (DSA) (Model SRS 785) is used to measure the frequency responses of the VCM actuator. The DSA is also used to record both input and output signals of time-domain responses. Square waves are generated with a dSpace DSP board installed in a personal computer.

The time-domain response of the VCM actuator to a typical square input signal about 1 Hz is shown in Figure 9.4. With a group of time-domain responses to a set of square input signals, we obtain the corresponding measurement data for the nonlinear function, $\tilde{T}_{bc}(y_0)$, which can be matched nicely by an arctan function (see Figure 9.5) as follows:

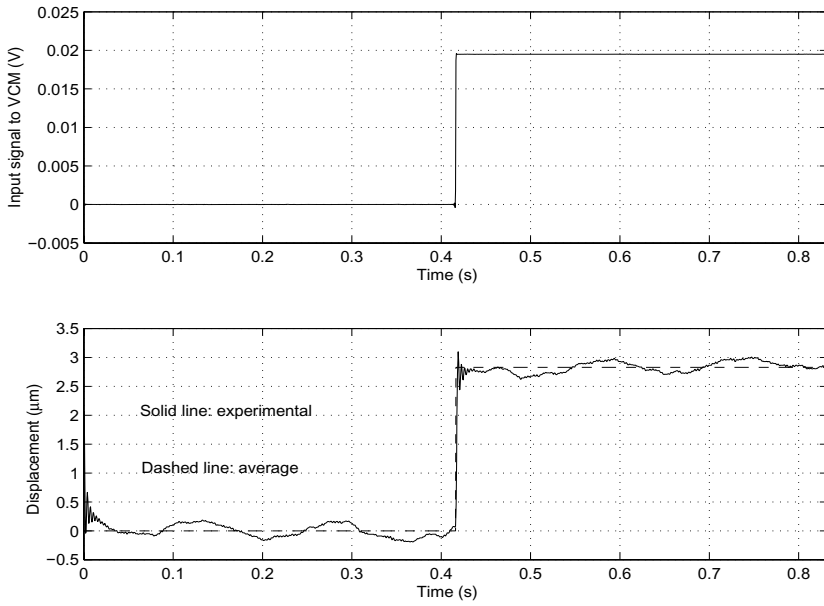


Figure 9.4. Time-domain response of the VCM actuator to a square wave input

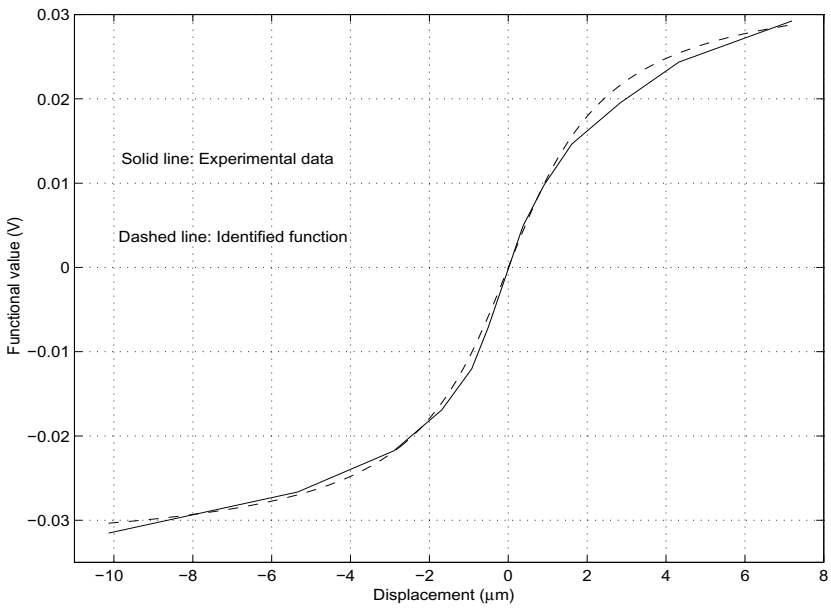


Figure 9.5. Nonlinear characteristics of the data flex cable

$$\tilde{T}_{bc}(y) = a \arctan(cy) \tag{9.27}$$

where $a = 21.8748 \times 10^{-3}$ V and $c = 0.5351$ (μm)⁻¹. These parameters will be further fine tuned later in the Monte Carlo process.

Next, by fixing a particular input offset point u_0 and by injecting on top of u_0 a sweep of small sinusoidal signals with an amplitude of 1 mV, we are able to obtain a corresponding frequency response within the range of interest. It then follows from Equation 9.26 that the values of the static gain, k_{dc} , and peak frequency, f_p , of the frequency response can be used to estimate the parameter, b . Figure 9.6 shows the frequency response of the system for the pair $(u_0, y_0) = (0, 0)$, which gives a static gain of 63.71 and a peak frequency of 305.24 Hz. In order to obtain a more accurate result, we repeat the above experimental tests for several pairs (u_0, y_0) and the results are shown in Table 9.1. The parameter, b , can then be more accurately determined from these data using a least square fitting,

$$\min_b \sum_{i=1}^7 [b - k_{dc,i} \omega_{p,i}^2]^2 \tag{9.28}$$

which gives an optimal solution $b = 2.28 \times 10^8$ $\mu\text{m} / (\text{Vs}^2)$. Nonetheless, this parameter will again be fine tuned later in the Monte Carlo process.

Lastly, we apply a Monte Carlo process to identify all other parameters of our VCM actuator model and to fine tune those parameters, which have previously been identified. Monte Carlo processes are known as numerical simulation methods

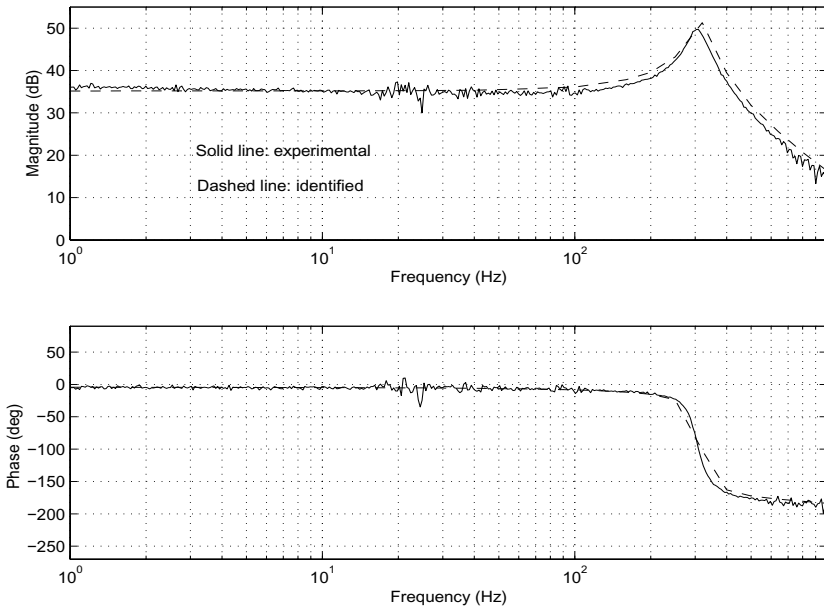


Figure 9.6. Frequency response to small signals at the steady state with $u_0 = 0$

Table 9.1. Static gains and peak frequencies of the actuator for small inputs

u_0 (mV)	-10	-5	0	5	10	15	20
k_{dc}	60.73	59.06	63.71	62.43	63.72	65.12	65.50
f_p (Hz)	310.30	310.63	305.24	303.88	299.56	296.99	295.43

that make use of random numbers and probability statistics to solve some complicated mathematical problems. The detailed treatments of Monte Carlo methods vary widely from field to field. Originally, a Monte Carlo experiment means to use random numbers to examine some stochastic problems. The idea can be extended to deterministic problems by presetting some parameters and conditions of the problems. The use of Monte Carlo methods for modeling physical systems allows us to solve more complicated problems, and provides approximate solutions to a variety of mathematical problems, whose analytical solutions are hard, if not impossible, to derive. In what follows, a Monte Carlo process is utilized to obtain time-domain responses of the VCM actuator model in Equation 9.18 with a set of preset parameters ($b, a, c, d_0, d_1, d_2, d_3, d_4, d_5$) and input signals. The corresponding frequency responses are obtained through Fourier transformation of the obtained time-domain responses. Our idea of using the Monte Carlo process is to minimize the differences between simulated frequency responses and the experimental ones by iteratively adjusting the parameters of the physical model in Equation 9.18. The input signals in our simulations are again a combination of an offset u_0 and sinusoidal signals with a small amplitude 1 mV and several frequencies ranging from 1 Hz to 1 kHz.

Although Monte Carlo methods can only give locally minimal solutions, in our problem, however, the predetermined nonlinear characteristics of the data flex cable and the parameter, b , have given us a rough idea on what the true solution should be. The solution within the neighborhood of the previously identified parameters are given by

$$\left. \begin{aligned}
 b &= 2.35 \times 10^8 \mu\text{m} / (\text{V}\cdot\text{s}^2) \\
 a &= 0.02887 \text{ V} \\
 c &= 0.5886 (\mu\text{m})^{-1} \\
 d_0 &= 282.6 \text{ s}^{-1} \\
 d_1 &= 0.005 (\mu\text{m})^{-1} \\
 d_2 &= 0.01 (\mu\text{m})^{-1} \\
 d_3 &= 1.5 \times 10^4 \mu\text{m} / \text{s}^2 \\
 d_4 &= 0.0055 (\mu\text{m})^{-1} \\
 d_5 &= 1.65 \times 10^4 \mu\text{m} / \text{s}^2
 \end{aligned} \right\} \quad (9.29)$$

These parameters will be used for further verifications using the experimental setup of the actual system.

So far, we have only focused on the low-frequency components of the VCM actuator model. In fact, there are many high-frequency resonance modes, which are

crucial to the overall performance of HDD servo systems. The high-frequency resonance modes of the VCM actuator can be obtained from frequency responses of the system in the high-frequency region (see Figure 9.7). The transfer function that matches the frequency responses given in Figure 9.7 is identified using the standard least square estimation method of Chapter 2 and is characterized by

$$G(s) = \frac{2.35 \times 10^8}{s^2} G_{r.m.1}(s) G_{r.m.2}(s) G_{r.m.3}(s) G_{r.m.4}(s) G_{r.m.5}(s) \quad (9.30)$$

with the resonance modes being given as

$$G_{r.m.1}(s) = \frac{0.8709s^2 + 1726s + 1.369 \times 10^9}{s^2 + 1480s + 1.369 \times 10^9} \quad (9.31)$$

$$G_{r.m.2}(s) = \frac{0.9332s^2 - 805.8s + 1.739 \times 10^9}{s^2 + 125.1s + 1.739 \times 10^9} \quad (9.32)$$

$$G_{r.m.3}(s) = \frac{1.072s^2 + 925.1s + 1.997 \times 10^9}{s^2 + 536.2s + 1.997 \times 10^9} \quad (9.33)$$

$$G_{r.m.4}(s) = \frac{0.9594s^2 + 98.22s + 2.514 \times 10^9}{s^2 + 1805s + 2.514 \times 10^9} \quad (9.34)$$

and

$$G_{r.m.5}(s) = \frac{7.877 \times 10^9}{s^2 + 6212s + 7.877 \times 10^9} \quad (9.35)$$

Finally, for easy reference, we conclude this section by explicitly expressing the identified rigid model of VCM actuator:

$$\ddot{y} = 2.35 \times 10^8 u - 6.78445 \times 10^6 \arctan(0.5886y) + \tilde{T}_f \quad (9.36)$$

where

$$\tilde{T}_f = \begin{cases} -\tilde{T}_e, & \dot{y} = 0, |\tilde{T}_e| \leq \tilde{T}_s \\ -\tilde{T}_s \operatorname{sgn}(\tilde{T}_e), & \dot{y} = 0, |\tilde{T}_e| > \tilde{T}_s \\ -\left[1.175 \times 10^6 u y + (0.1\dot{y})^2 + 15000\right] \operatorname{sgn}(\dot{y}) - 282.6\dot{y}, & \dot{y} \neq 0 \end{cases} \quad (9.37)$$

and where

$$\tilde{T}_e = 2.35 \times 10^8 \left[-0.02887 \arctan(0.5886y) + u\right] \quad (9.38)$$

$$\tilde{T}_s = 1.293 \times 10^6 |u_0 y_0| + 1.65 \times 10^4 \quad (9.39)$$

with u_0 and y_0 being, respectively, the corresponding input voltage and the displacement for the case when $\dot{y} = 0$. Note that in the above model, the input signal u is in voltage and the output displacement y is in micrometers. Together with the high-frequency resonance modes of Equations 9.31–9.35, the above model presents a comprehensive characterization of the VCM actuator studied. This model will be further verified using experimental tests on the actual system.

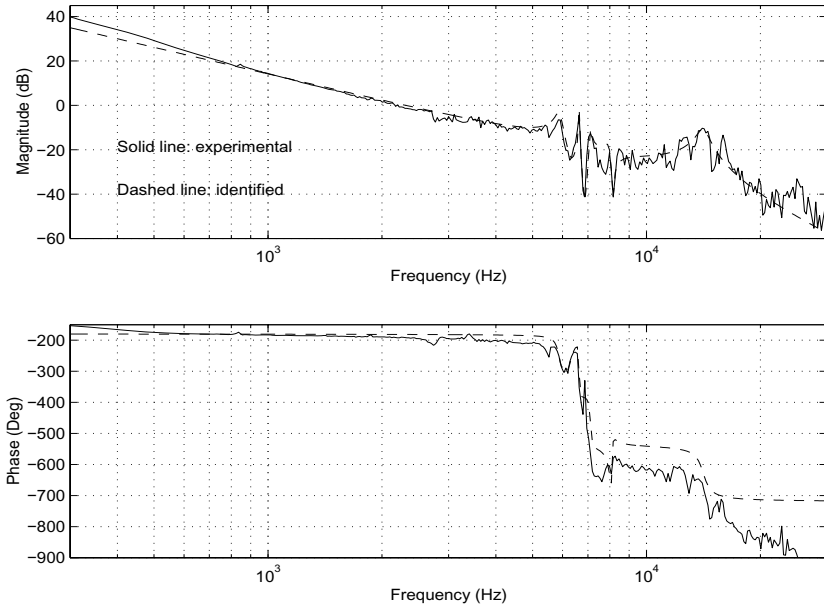


Figure 9.7. Frequency responses of the VCM actuator in the high-frequency region

In order to verify the validity of the established model of the VCM actuator, we carry out a series of comparisons between the experimental results and computed results of the time-domain responses and frequency-domain responses of the actuator. The comparison of the frequency responses between the experimental result and the identified result for inputs consisting of $u_0 = 5$ mV and sine waves with amplitude of 1 mV is shown in Figure 9.8. It clearly shows that the result of the identified model matches well with the experimental result. The comparison of the time-domain responses for an input signal consisting of $u_0 = -5$ mV and a sine wave with an amplitude of 5 mV is given in Figure 9.9. It shows that the simulation results match the trends and values of those obtained from experiments. The noises associated with experimental results in Figures 9.8 and 9.9 are drift noises caused by the LDV and/or DSA. The comparisons of both frequency-domain and time-domain responses demonstrate that the identified model of the VCM actuator indeed describes the features of the actuator.

9.3 Microdrive Servo System Design

We proceed to design a servo system for the microdrive identified in Section 9.2. As mentioned earlier, our design philosophy is rather simple. We make full use of the obtained model of the friction and nonlinearities of the VCM actuator to design a precompensator, which would cancel as much as possible all the unwanted elements in the servo system. As it is impossible to have perfect models for friction

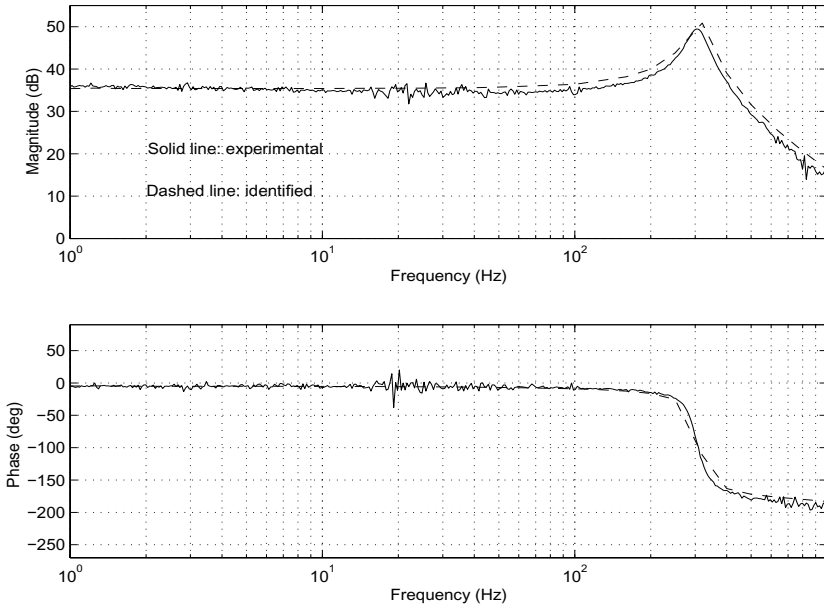


Figure 9.8. Comparison of frequency responses to small signals of actuator with $u_0 = 5 \text{ mV}$

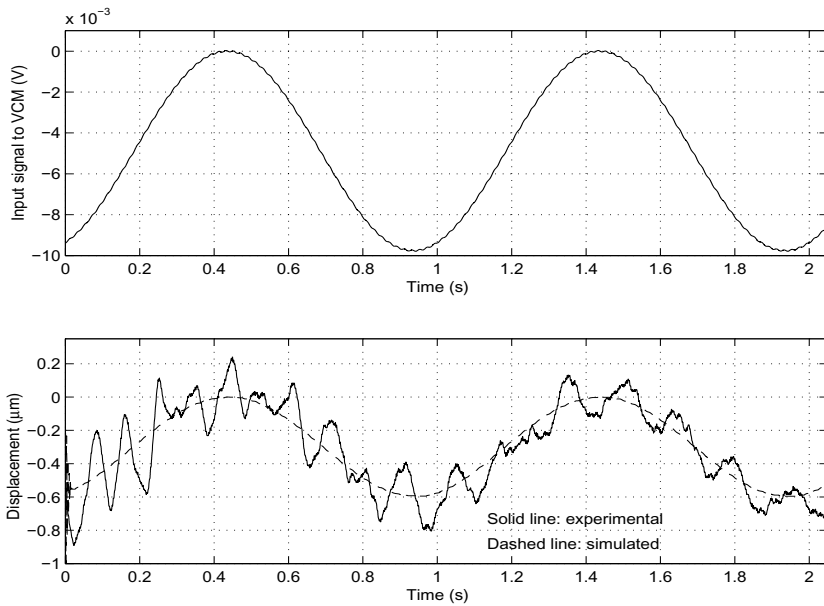


Figure 9.9. Comparison of time-domain responses of the VCM actuator

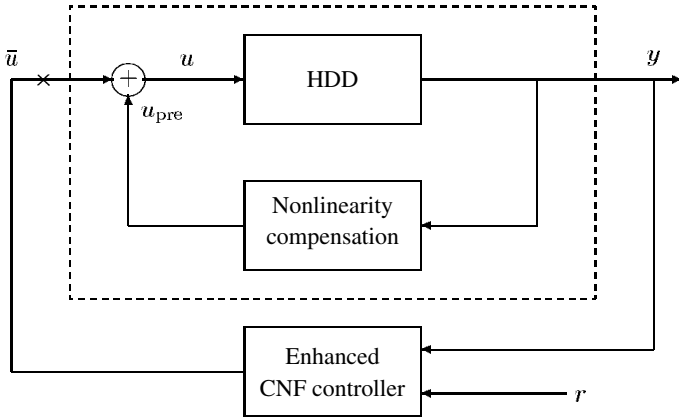


Figure 9.10. Control scheme for the HDD servo system

and nonlinearities, a perfect cancellation of these elements is unlikely to happen in the real world. We then formulate our design by treating the uncompensated portion as external disturbances. The enhanced CNF control technique of Chapter 5 is then employed to design an effective tracking controller. The overall control scheme for the servo system is depicted in Figure 9.10. Although we focus our attention here on HDD, it is our belief that such an approach can be adopted to solve other servo problems.

Examining the model of Equation 9.36, it is easy to obtain a precompensation,

$$u_{pre} = u - \bar{u} = 0.0288737 \arctan(0.5886 y) \tag{9.40}$$

which would eliminate the majority of nonlinearities in the data flex cable. The HDD model of Equation 9.18 can then be simplified as follows:

$$\begin{cases} \dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 2.35 \times 10^8 \end{bmatrix} \text{sat}(\bar{u}) + \begin{bmatrix} 0 \\ 2.35 \times 10^8 \end{bmatrix} w \\ y = h = [1 \quad 0] x \end{cases} \tag{9.41}$$

where the disturbance, w , represents uncompensated nonlinearities, and $y = h$ is the relative displacement of the R/W head (in micrometers). The control input, \bar{u} , is to be limited within $\pm \bar{u}_{max}$ with $\bar{u}_{max} = 3$ V.

We design a microdrive servo system that meets the following design constraints and specifications:

1. the control input does not exceed ± 3 V owing to physical constraints on the actual VCM actuator;
2. the overshoot and undershoot of the step response are kept to less than 5% as the R/W head can start to read or write within $\pm 5\%$ of the target;

3. the 5% settling time in the step response is as short as possible;
4. the gain margin and phase margin of the overall design are, respectively, greater than 6 dB and 30°;
5. the maximum peaks of the sensitivity and complementary sensitivity functions are less than 6 dB; and
6. the sampling frequency in implementing the actual controller is 20 kHz.

It turns out that for the microdrive its resonance modes are at very high frequencies that are far above the working range of the drive. It is thus not necessary to add a notch filter to minimize their effects. As usual, we consider a second-order nominal model of Equation 9.41 for the VCM actuator. The resonance modes and the notch filter will be put back to evaluate the performance of the overall design. As in Chapters 6 and 8, we design our servo system using, respectively, PID, RPT and CNF control.

1. The PID control law (discretized with a sampling frequency of 20 kHz) is given by

$$\bar{u}(k) = \frac{0.3145z^2 - 0.6095z + 0.2951}{z^2 - 1.091z + 0.09091} \cdot \frac{0.4665}{z - 0.5335} [r(k) - y(k)] \quad (9.42)$$

2. The RPT controller is given by

$$\dot{x}_v = \begin{bmatrix} 0 & 0 \\ -2220.7 & -12911 \end{bmatrix} x_v + \begin{bmatrix} -1 & 1 \\ 222.56 & -1255.4 \end{bmatrix} \begin{pmatrix} r \\ y \end{pmatrix} \quad (9.43)$$

and

$$\bar{u} = [-0.94496 \quad -2.0897] x_v + [0.094705 \quad -0.26188] \begin{pmatrix} r \\ y \end{pmatrix} \quad (9.44)$$

3. Finally, the CNF control law is given as follows:

$$\begin{pmatrix} \dot{x}_i \\ \dot{x}_v \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -8000 \end{bmatrix} \begin{pmatrix} x_i \\ x_v \end{pmatrix} - \begin{bmatrix} -1 \\ 6.4 \times 10^7 \end{bmatrix} y + \begin{bmatrix} 0 \\ 2.35 \times 10^8 \end{bmatrix} \text{sat}(\bar{u}) - \begin{pmatrix} r \\ 0 \end{pmatrix} \quad (9.45)$$

and

$$\begin{aligned} \bar{u} = \rho(e) [& 1.3607 \quad 0.13607 \quad 2.4063 \times 10^{-5}] \begin{pmatrix} z \\ y - r \\ x_v + 8000y \end{pmatrix} \\ & - [1.3607 \times 10^{-4} \quad 0.13607 \quad 2.6470 \times 10^{-5}] \begin{pmatrix} z \\ y - r \\ x_v + 8000y \end{pmatrix} \end{aligned} \quad (9.46)$$

where

$$\rho(e) = -0.3164 \left| e^{-|e|} - 0.3679 \right| - 0.3 \quad (9.47)$$

9.4 Simulation and Implementation Results

As in the early chapters, the following tests are performed on the obtained microdrive servo systems: i) the track-following test of the closed-loop systems, ii) the frequency-domain test including the Bode and Nyquist plots as well as the plots of the resulting sensitivity and complementary sensitivity functions, iii) the runout disturbance test, and lastly iv) the PES test. Our controller was implemented on an open microdrive with a sampling rate of 20 kHz with the actual implementation setup as depicted in Figure 1.7.

9.4.1 Track-following Test

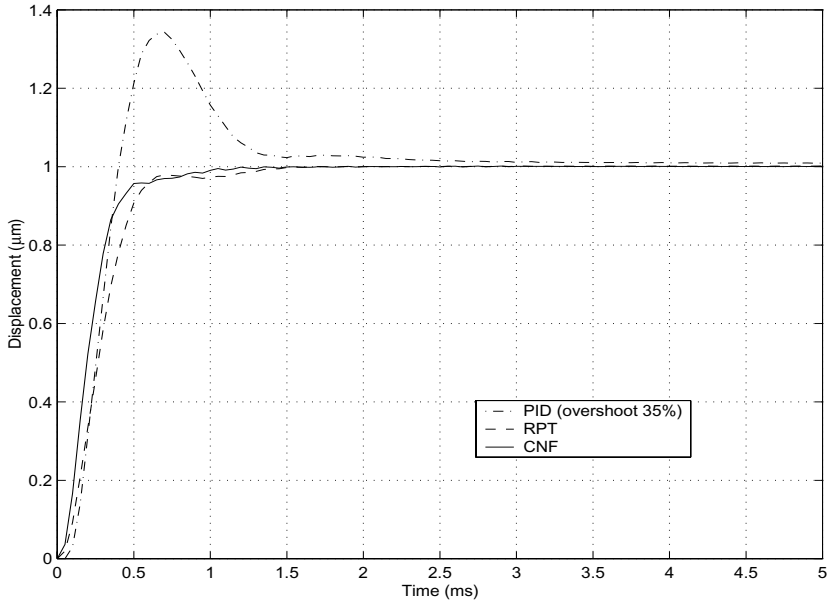
The simulation result and actual implementation result of the closed-loop responses for the control systems are, respectively, shown in Figures 9.11 and 9.12. As expected, the PID control generates large overshoots in both simulation and implementation, while the systems with the RPT and CNF control have very little overshoot. We summarize the resulting 5% settling time in Table 9.2. Clearly, the CNF control once again gives the best performance in the time domain compared to those of the other two systems.

Table 9.2. Performances of the track-following controllers

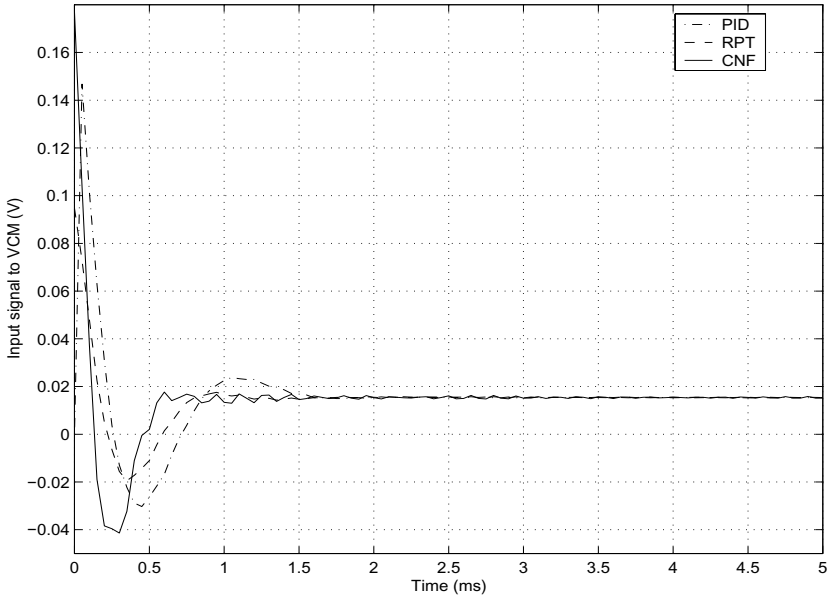
	Settling time (ms)		
	PID control	RPT control	CNF control
Simulation	1.25	0.60	0.50
Implementation	1.50	0.65	0.60

9.4.2 Frequency-domain Test

To verify the frequency-domain properties of our design, we present in Figures 9.13 to 9.18, respectively, the Bode plot, the Nyquist plot, and the sensitivity and complementary sensitivity functions, as well as the closed-loop transfer functions of the resulting control systems. The results show that all these designs meet the frequency-domain specifications.

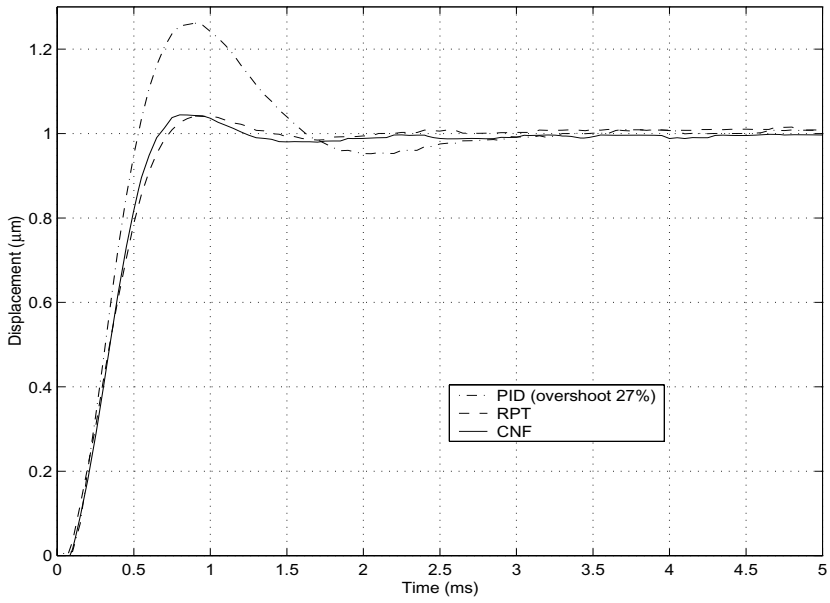


(a) Output responses

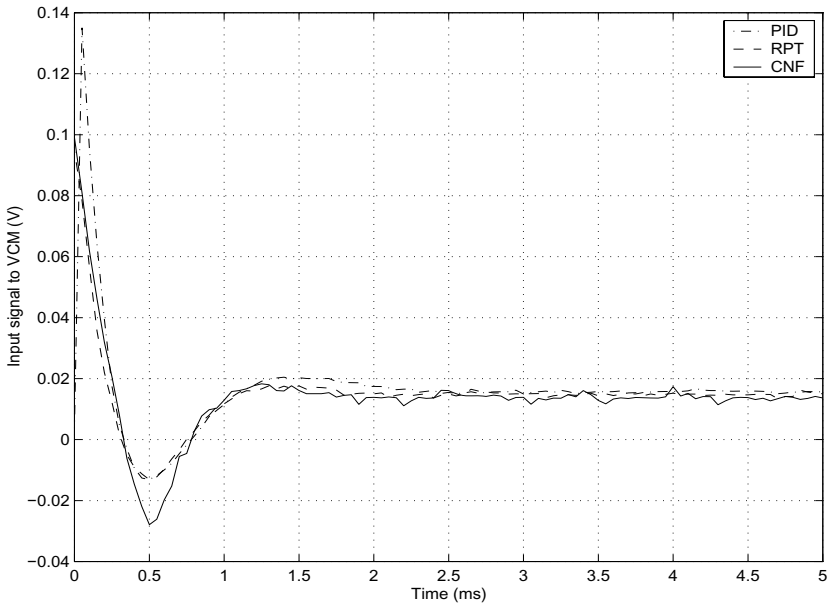


(b) Control signals

Figure 9.11. Simulation result: step responses with PID, RPT and CNF control

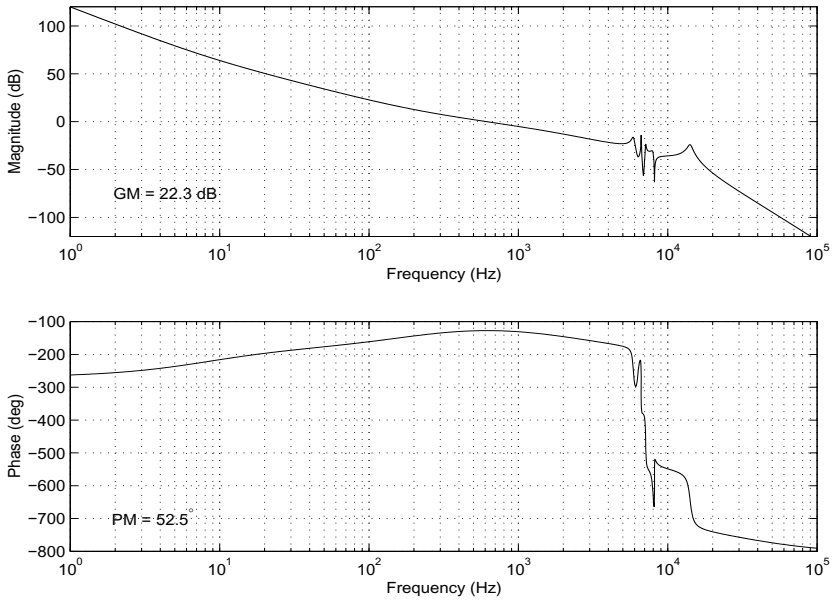


(a) Output responses

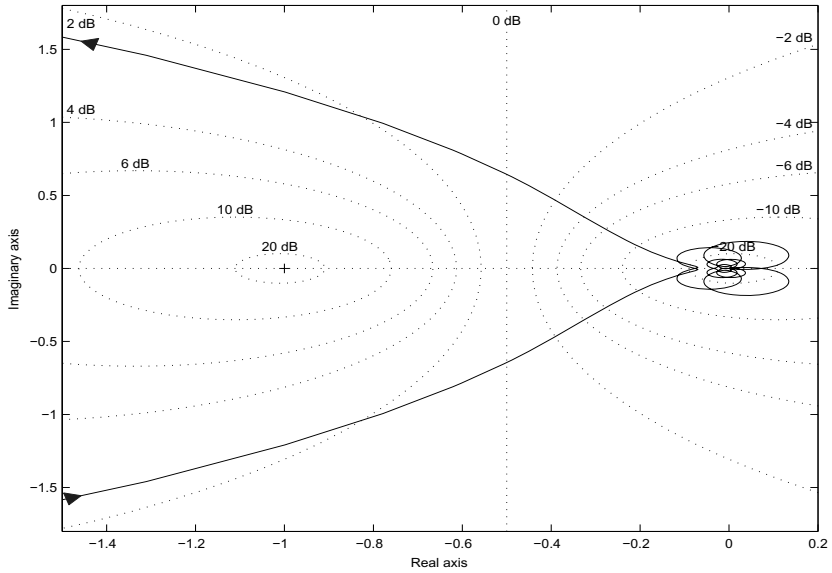


(b) Control signals

Figure 9.12. Implementation result: step responses with PID, RPT and CNF control

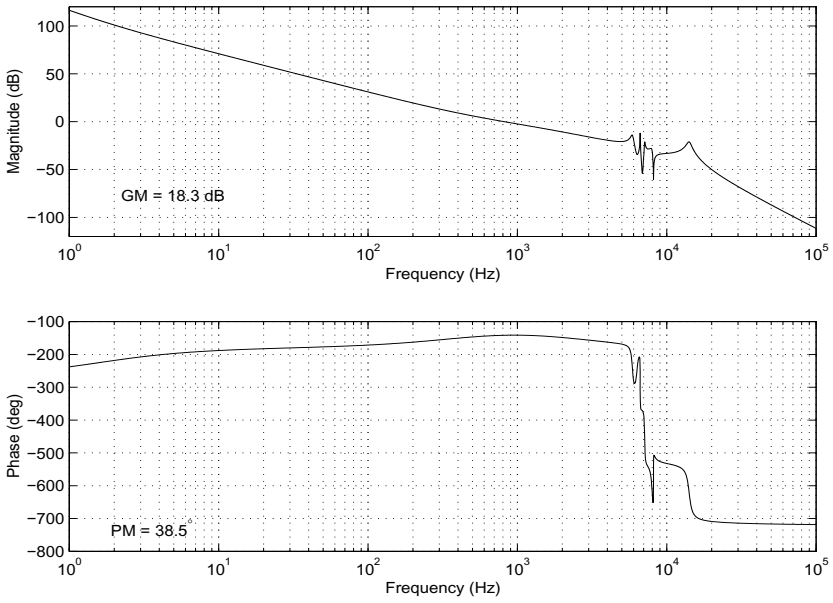


(a) Bode plot

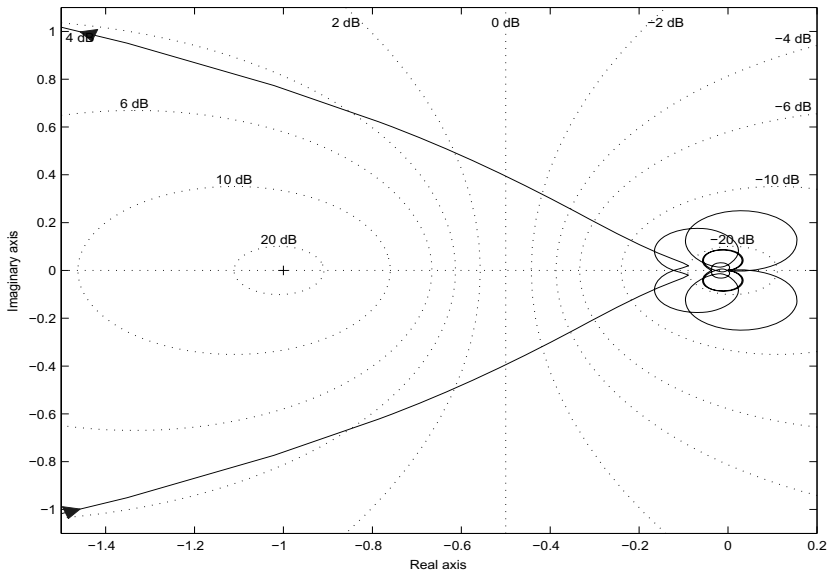


(b) Nyquist plot

Figure 9.13. Bode and Nyquist plots of the PID control system

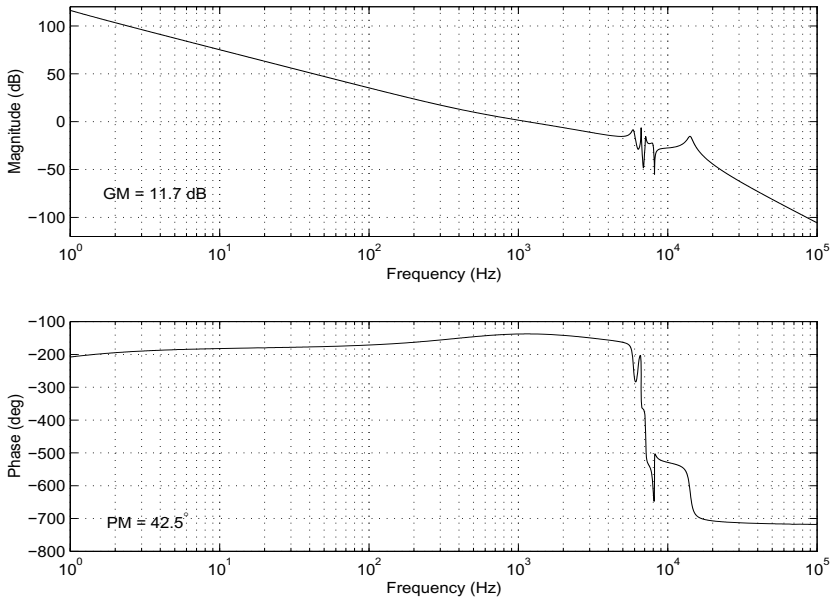


(a) Bode plot

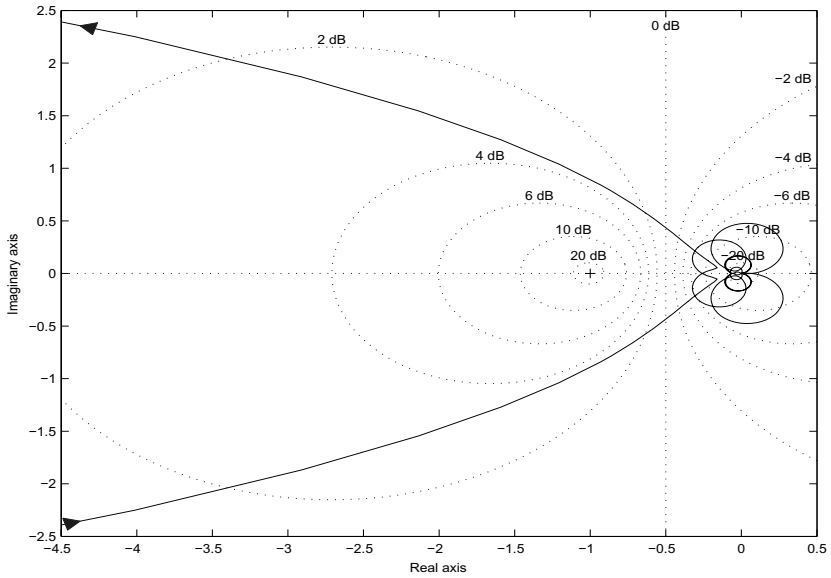


(b) Nyquist plot

Figure 9.14. Bode and Nyquist plots of the RPT control system

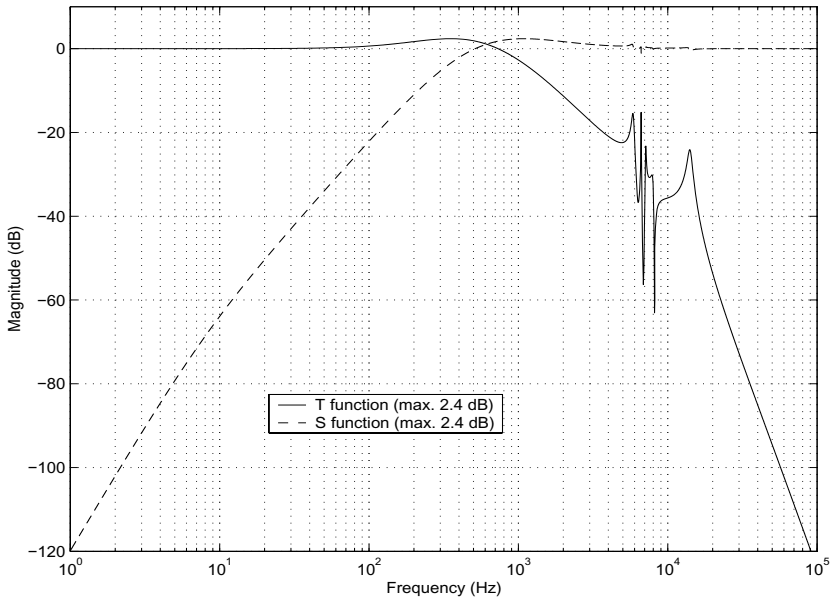


(a) Bode plot

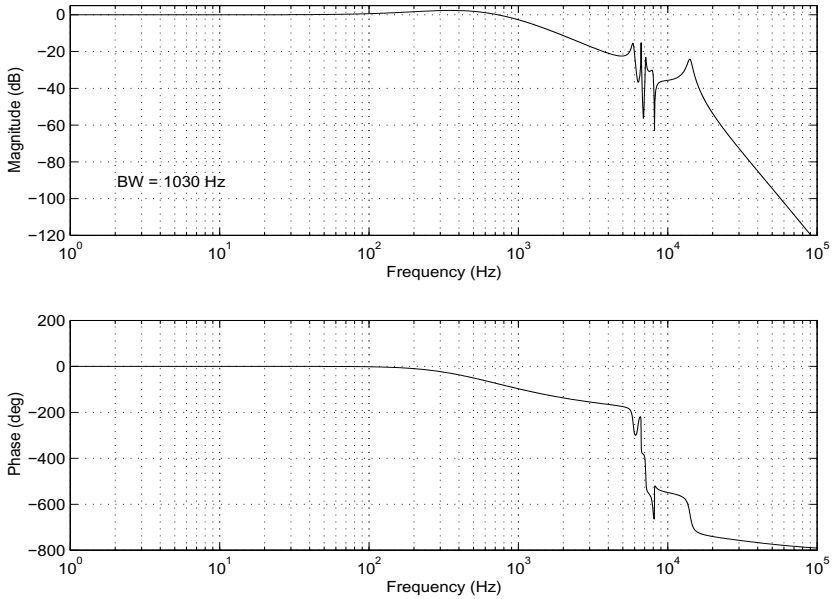


(b) Nyquist plot

Figure 9.15. Bode and Nyquist plots of the CNF control system

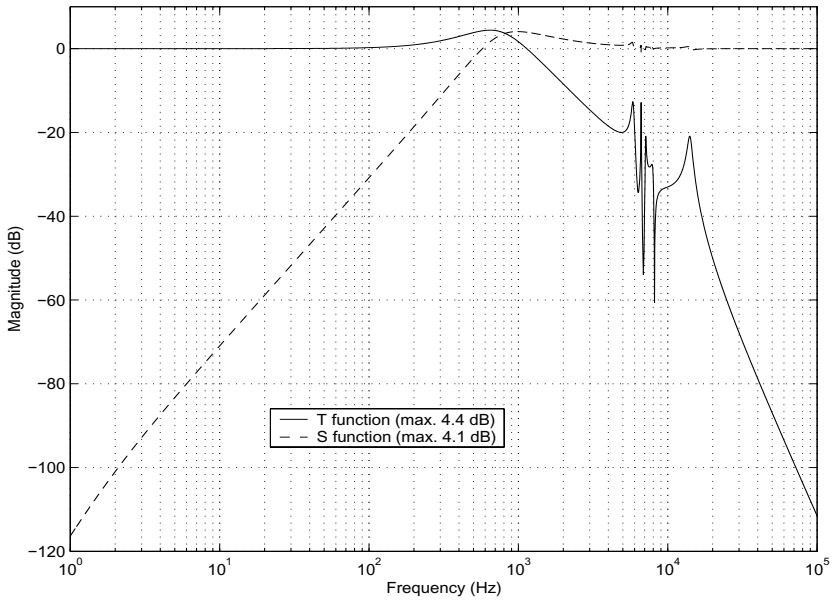


(a) Sensitivity and complementary sensitivity functions

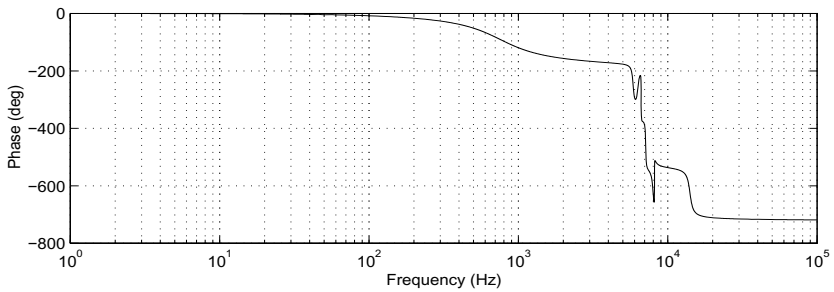
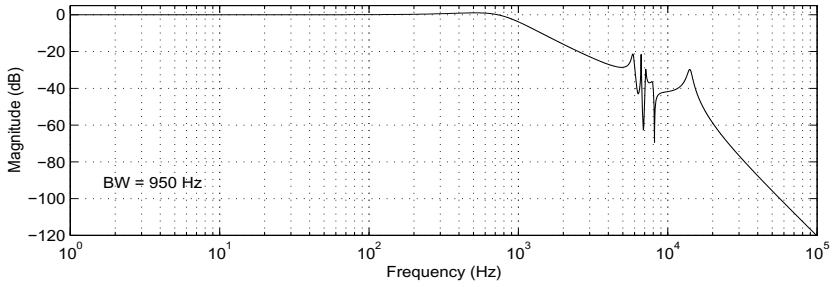


(b) Closed-loop response

Figure 9.16. Sensitivity functions and closed-loop transfer function of the PID control system

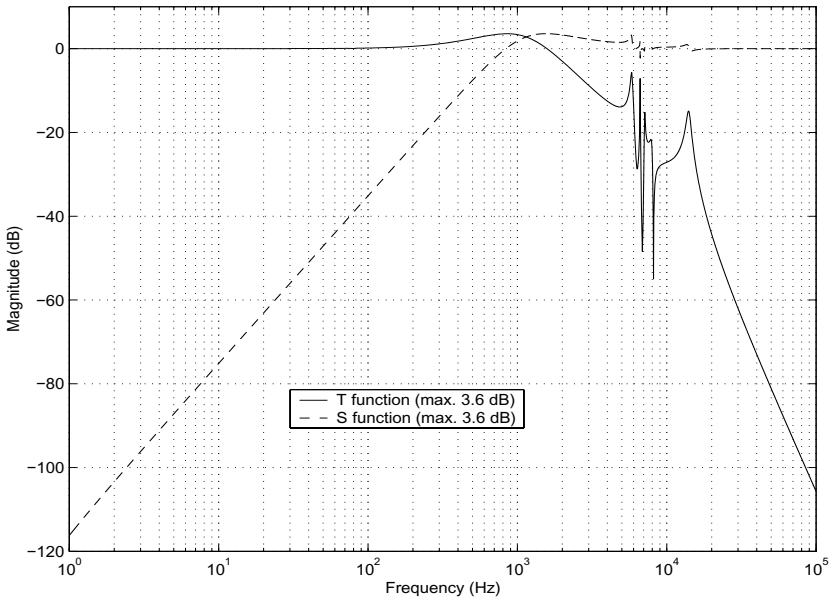


(a) Sensitivity and complementary sensitivity functions

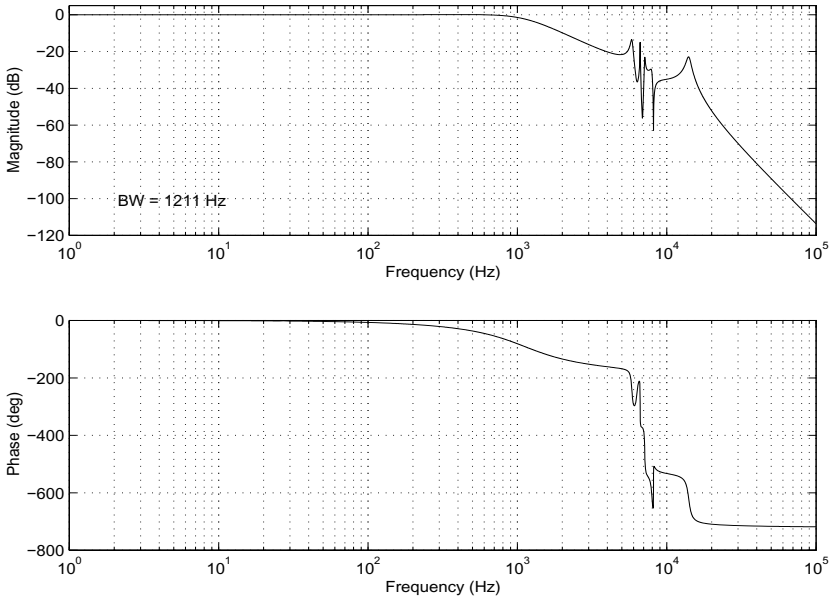


(b) Closed-loop response

Figure 9.17. Sensitivity functions and closed-loop transfer function of the RPT control system



(a) Sensitivity and complementary sensitivity functions



(b) Closed-loop response

Figure 9.18. Sensitivity functions and closed-loop transfer function of the CNF control system

Design of a Piezoelectric Actuator System

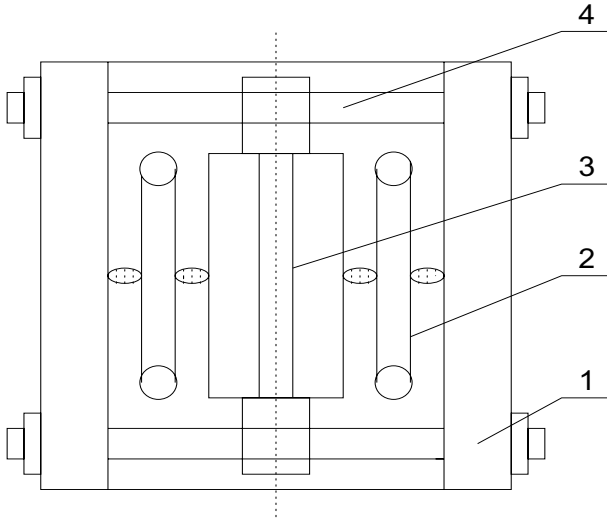
10.1 Introduction

We present in this chapter a case study on a piezoelectric bimorph actuator control system design using an H_∞ optimization approach, which was originally reported by Chen *et al.* [20]. Piezoelectricity is a fundamental process in electromechanical energy conversion. It relates electric polarization to mechanical stress/strain in piezoelectric materials. Under the direct piezoelectric effect, an electric charge can be observed when the materials are deformed. The converse, or the reciprocal piezoelectric effect, is that the application of an electric field can cause mechanical stress/strain in the piezo materials. There are numerous piezoelectric materials available today, including PZT (lead zirconate titanate), PLZT (lanthanum-modified lead zirconate titanate), and PVDF (piezoelectric polymeric polyvinylidene fluoride), to name a few (see Low and Guo [181]).

Piezoelectric structures are widely used in applications that require electrical to mechanical energy conversion coupled with size limitations, precision, and speed of operation. Typical examples are microsensors, micropositioners, speakers, medical diagnostics, shutters and impact print hammers. In most applications, bimorph or stack piezoelectric structures are used because of the relatively high stress/strain to input electric field ratio (see Low and Guo [181]).

The work was motivated by the possibility of applying piezoelectric microactuators in magnetic recording. The focus of this chapter is on the control issues involved in dealing with the nonlinear hysteresis behavior displayed by most piezoelectric actuators. More specifically, we consider a robust controller design for a piezoelectric bimorph actuator as depicted in Figure 10.1. A scaled-up model of this piezoelectric actuator, which is targeted for use in the secondary stage of a dual-stage actuator for magnetic recording, was actually built and modeled by Low and Guo [181]. It has two pairs of bimorph beams that are subjected to bipolar excitation. The dynamics of the actuator were identified in [181] as a second-order linear model coupled with a hysteresis. The linear model is given by

$$m\ddot{x}_1 + b\dot{x}_1 + kx_1 = k(du - z) \quad (10.1)$$



1-base; 2-piezoelectric bimorph beams; 3-moving plate; and 4-guides

Figure 10.1. Structure of the piezoelectric bimorph actuator

where m , b , k and d are the tangent mass, damping, stiffness and effective piezoelectric coefficients; u is the input voltage that generates excitation forces in the actuator system. The variable x_1 is the displacement of the actuator and is also the only measurement we can have in this system. It is noted that the working range of the displacement of this actuator is within $\pm 1 \mu\text{m}$. The variable z arises from hysteretic nonlinear dynamics [181] and is governed by

$$\dot{z} = \alpha \dot{u} - \beta |\dot{u}| z - \gamma \dot{u} |z| \tag{10.2}$$

where α , β and γ are some constants that control the shape of the hysteresis. For the actuator system that we are considering here, the above coefficients are identified as follows:

$$\left. \begin{aligned} m &= 0.01595 \text{ kg} \\ b &= 1.169 \text{ N s/m} \\ k &= 4385 \text{ N/m} \\ d &= 8.209 \times 10^{-7} \text{ m/V} \\ \alpha &= 0.4297 \\ \beta &= 0.03438 \\ \gamma &= -0.002865 \end{aligned} \right\} \tag{10.3}$$

For a more detailed description of this piezoelectric actuator system and the identifications of the above parameters, we refer interested readers to the work of Low and Guo [181]. Our goal in this chapter is to design a robust controller, as in Figure 10.2, that meets the following design specifications.

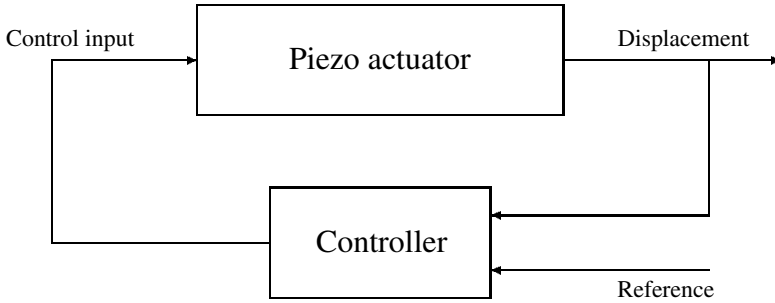


Figure 10.2. Piezoelectric bimorph actuator plant with controller

1. The steady-state tracking error of the displacement is less than 1% for any input reference signals that have frequencies ranging from 0 to 30 Hz, as the actuator is to be used to track certain colored noise types of signal in disk drive systems.
2. The 1% settling time is as short as possible (we are able to achieve a 1% settling time of less than 0.003 s in our design).
3. The control input signal $u(t)$ does not exceed 112.5 V because of the physical limitations of the piezoelectric materials.

Our approach is as follows: we first use the stochastic equivalent linearization method proposed by Chang [182] to obtain a linearized model for the nonlinear hysteretic dynamics. Then we reformulate our design into an H_∞ almost disturbance decoupling problem in which the disturbance inputs are the reference input and the error between the hysteretic dynamics and that of its linearized model, and where the controlled output is simply the double integration of the tracking error. Thus, our task boils down to designing a controller such that, when it is applied to the piezoelectric actuator, the overall system is asymptotically stable, and the controlled output, which corresponds to the tracking error, is as small as possible and decays as fast as possible. Such an approach is quite innovative and can be adopted to tackle similar problems for other types of piezo materials and applications. Ever since its original publication in [20], the work has been frequently cited in the literature (see, for example, [3, 132, 133, 183–201]).

The outline of this chapter is as follows. In Section 10.2 a first-order linearized model is obtained for the nonlinear hysteresis using the stochastic equivalent linearization method. A simulation result is also given to show the match between the nonlinear and linearized models. In Section 10.3 we formulate our controller design into a standard almost disturbance decoupling problem by properly defining the disturbance input and the controlled output. Two integrators are augmented into the original plant to enhance the performance of the overall system. Then a robust controller that is explicitly parameterized by a certain tuning parameter, and that solves the proposed almost disturbance decoupling problem, is carried out using a so-called asymptotic time scale and eigenstructure assignment technique. In Section 10.4 we

present the final controller and simulation results of our overall control system using MATLAB[®] and Simulink[®]. We also obtain an explicit relationship between the peak values of the control signal and the tuning parameter of the controller, as well as an explicit linear relationship of the maximum trackable frequency, for which the corresponding tracking error can be settled to 1%, versus the tuning parameter of the controller.

10.2 Linearization of Nonlinear Hysteretic Dynamics

We proceed to linearize the nonlinear hysteretic dynamics of Equation 10.2 in this section. As pointed out by Chang [182], there are basically three methods available in the literature to linearize the hysteretic types of nonlinear system. These are (i) the Fokker–Planck equation approach (see, *e.g.*, Caughey [202]), (ii) the perturbation techniques (see, *e.g.*, Crandall [203] and Lyon [204]) and (iii) the stochastic linearization approach. All of them have certain advantages and limitations. However, the stochastic linearization technique has the widest range of applications compared with the other methods. This method is based on the concept of replacing the nonlinear system with an “equivalent” linear system in such a way that the “difference” between these two systems is minimized in a certain sense. The technique was initiated by Booton [205]. In this chapter, we just follow the stochastic linearization method given by Chang [182] to obtain a linear model of the following form:

$$\dot{z} = k_1\dot{u} + k_2z \tag{10.4}$$

for the hysteretic dynamics of Equation 10.2, where k_1 and k_2 are the linearization coefficients and are to be determined. The procedure is quite straightforward and proceeds as follows. First, we introduce a so-called “difference” function e between \dot{z} of Equation 10.2 and \dot{z} of Equation 10.4:

$$e(k_1, k_2) = \alpha\dot{u} - \beta|\dot{u}|z - \gamma\dot{u}|z| - (k_1\dot{u} + k_2z) \tag{10.5}$$

Then minimizing $\mathbf{E}[e^2]$, where \mathbf{E} is the expectation operator, with respect to k_1 and k_2 , we obtain

$$\frac{\partial \mathbf{E}[e^2]}{\partial k_1} = \frac{\partial \mathbf{E}[e^2]}{\partial k_2} = 0 \tag{10.6}$$

from which the stochastic linearization coefficients k_1 and k_2 are determined. It turns out that if z and \dot{u} are of zero means and jointly Gaussian, then k_1 and k_2 can be easily obtained. Let us assume that z and \dot{u} have a joint probability density function

$$f_{\dot{u}z}(\dot{u}, z) = \frac{1}{2\pi\sigma_{\dot{u}}\sigma_z\sqrt{1-\rho_{\dot{u}z}^2}} \exp \left\{ -\frac{\sigma_{\dot{u}}^2z^2 - 2\sigma_{\dot{u}}\sigma_z\rho_{\dot{u}z}\dot{u}z + \sigma_z^2\dot{u}^2}{2\sigma_{\dot{u}}^2\sigma_z^2(1-\rho_{\dot{u}z}^2)} \right\}$$

where $\rho_{\dot{u}z}$ is the normalized covariance of \dot{u} and z , and $\sigma_{\dot{u}}$ and σ_z are the standard deviations of \dot{u} and z , respectively. Then the linearization coefficients k_1 and k_2 can be expressed as follows:

$$k_1 = \alpha d - \beta c_1 - \gamma c_2 \quad (10.7)$$

and

$$k_2 = -\beta c_3 - \gamma c_4 \quad (10.8)$$

where c_1 , c_2 , c_3 and c_4 are given by

$$c_1 = 0.79788456\sigma_z \cos \left[\tan^{-1} \left(\frac{\sqrt{1 - \rho_{\dot{u}z}^2}}{\rho_{\dot{u}z}} \right) \right] \quad (10.9)$$

$$c_2 = 0.79788456\sigma_z, \quad c_4 = 0.79788456\rho_{\dot{u}z}\sigma_{\dot{u}} \quad (10.10)$$

and

$$c_3 = 0.79788456\sigma_{\dot{u}} \left\{ 1 - \rho_{\dot{u}z}^2 + \rho_{\dot{u}z} \cos \left[\tan^{-1} \left(\frac{\sqrt{1 - \rho_{\dot{u}z}^2}}{\rho_{\dot{u}z}} \right) \right] \right\} \quad (10.11)$$

After a few iterations, we found that a sinusoidal excitation \dot{u} with frequencies ranging from 0 to 100 Hz (the expected working frequency range) and peak magnitude of 50 V, which has a standard deviation of $\sigma_{\dot{u}} = 35$, would yield a suitable linearized model for Equation 10.2. For this excitation, we obtain $\sigma_z = 5 \times 10^{-7}$, $\rho_{\dot{u}z} = 5 \times 10^{-3}$

$$c_1 = 1.9947 \times 10^{-9}, \quad c_2 = 3.9894 \times 10^{-7}, \quad c_3 = 27.9260 \quad (10.12)$$

and

$$c_4 = 0.1396, \quad k_1 = 3.5382 \times 10^{-7}, \quad k_2 = -0.9597 \quad (10.13)$$

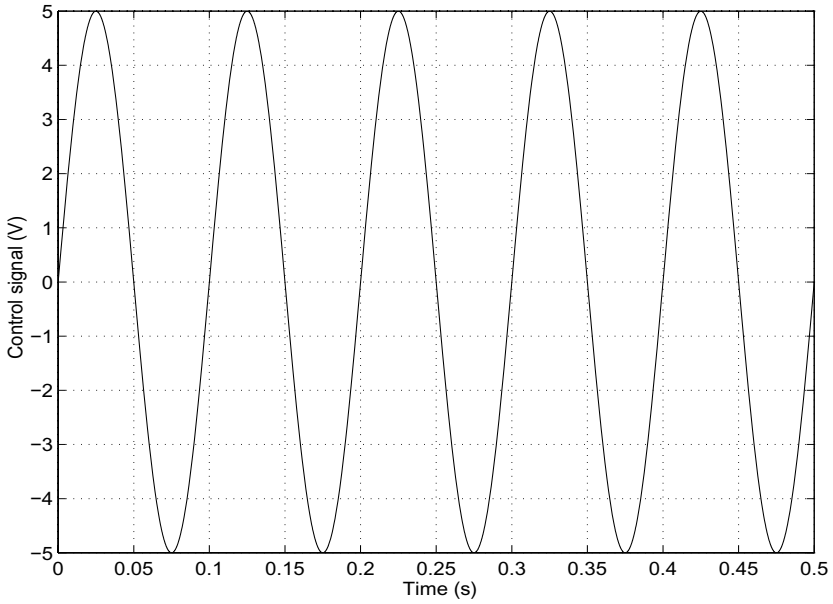
The stochastic linearization model of the given nonlinear hysteretic dynamics of Equation 10.2 is then given by

$$\dot{\hat{z}} = k_1 \dot{u} + k_2 \hat{z} = 3.5382 \times 10^{-7} \dot{u} - 0.9597 \hat{z} \quad (10.14)$$

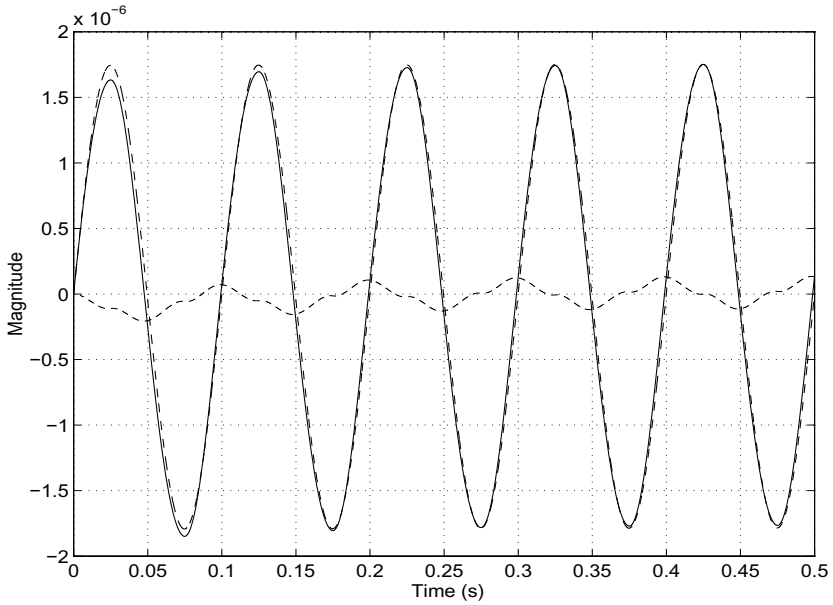
For future use, let us define the linearization error as

$$e_z = z - \hat{z} \quad (10.15)$$

Figure 10.3 shows the open-loop simulation results of the nonlinear hysteresis and its linearized model, as well as the error for a typical sine wave input signal u . The results are quite satisfactory. Here we note that, because of the nature of our approach in controller design later in the next section, the variation of the linearized model within a certain range, which might result in larger linearization error e_z , does not much affect the overall performance of the closed-loop system. We formulate e_z as a disturbance input and our controller automatically rejects it from the output response.



(a) Control input signal, u



(b) z (solid), \hat{z} (dashed), and e_z (dash-dotted)

Figure 10.3. Responses of hysteresis and its linearized model to a sine input

10.3 Almost Disturbance Decoupling Controller Design

This section is the heart of this chapter. We first formulate our control system design for the piezoelectric bimorph actuator into a standard H_∞ almost disturbance decoupling problem, and then apply the results of Chapter 3 to check the solvability of the proposed problem. Finally, we utilize the results in Chapter 3 to find an internally stabilizing controller that solves the proposed almost disturbance decoupling problem. Of course, most importantly, the resulting closed-loop system and its responses have to meet all the design specifications as stated at the beginning of this chapter. To do this, we have to convert the dynamic model of Equation 10.1 with the linearized model of the hysteresis into a state-space form. Let us first define a new state variable

$$v = \hat{z} - k_1 u \quad (10.16)$$

Then, from Equation 10.14, we have

$$\dot{v} = \dot{\hat{z}} - k_1 \dot{u} = k_2 \hat{z} = k_2 v + k_1 k_2 u \quad (10.17)$$

Substituting Equations 10.15 and 10.16 into Equation 10.1, we obtain

$$\ddot{x}_1 + \frac{b}{m} \dot{x}_1 + \frac{k}{m} x_1 + \frac{k}{m} v = \frac{k(d - k_1)}{m} u - \frac{k}{m} e_z \quad (10.18)$$

The overall controller structure of our approach is then depicted in Figure 10.4. Note that in Figure 10.4 we have augmented two integrators after e , the tracking error between the displacement x_1 and the reference input signal r . We have observed a very interesting property of this problem, *i.e.* the more integrators that we augment after the tracking error e , the smaller the tracking error we can achieve for the same level of control input u . Because our control input u is limited to the range from -112.5 to 112.5 V, it turns out that two integrators are needed in order to meet all the design specifications. It is clear to see that the augmented system has an order of five. Next, let us define the state of the augmented system as

$$x = (x_1 \quad \dot{x}_1 \quad v \quad x_4 \quad x_5)' \quad (10.19)$$

and the measurement output

$$y = \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_4 \\ x_5 \end{pmatrix} \quad (10.20)$$

i.e. the original measurement of displacement x_1 plus two augmented states. The auxiliary disturbance input is

$$w = \begin{pmatrix} e_z \\ r \end{pmatrix} \quad (10.21)$$

and the output to be controlled, h , is simply the double integration of the tracking error. The state-space model of the overall augmented system is then given by

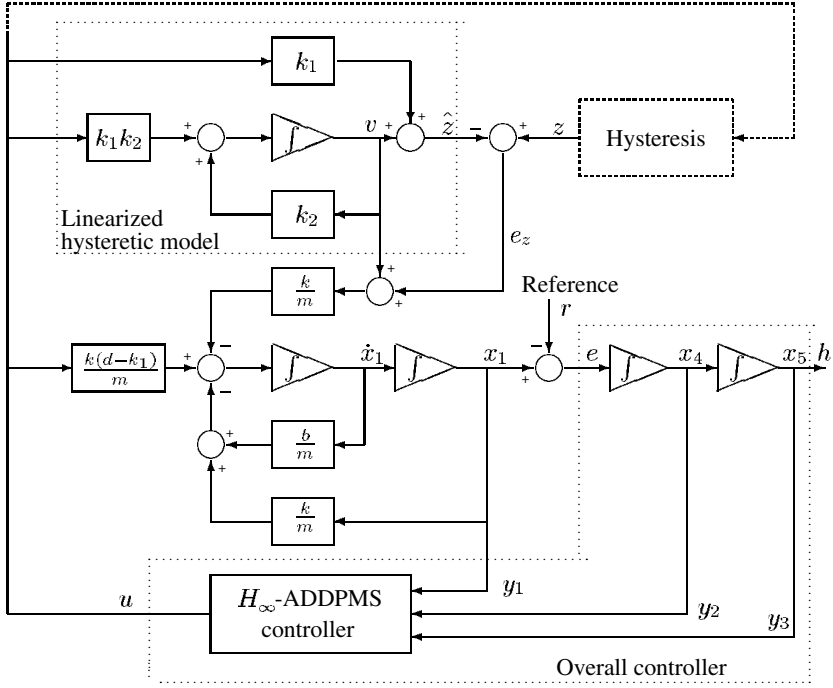


Figure 10.4. Augmented linearized model with controller

$$\Sigma : \begin{cases} \dot{x} = A x + B u + E w \\ y = C_1 x + D_1 w \\ h = C_2 x + D_2 u \end{cases} \quad (10.22)$$

with

$$C_1 = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \quad D_1 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix} \quad (10.23)$$

$$C_2 = [0 \ 0 \ 0 \ 0 \ 1], \quad D_2 = 0 \quad (10.24)$$

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -k/m & -b/m & -k/m & 0 & 0 \\ 0 & 0 & k_2 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 \\ -274921.63 & -73.2915 & -274921.63 & 0 & 0 \\ 0 & 0 & -0.9597 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \end{bmatrix} \quad (10.25)$$

$$B = \begin{bmatrix} 0 \\ k(d - k_1)/m \\ k_1 k_2 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0.12841 \\ -3.39561 \times 10^{-7} \\ 0 \\ 0 \end{bmatrix} \quad (10.26)$$

$$E = \begin{bmatrix} 0 & 0 \\ -k/m & 0 \\ 0 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ -274921.63 & 0 \\ 0 & 0 \\ 0 & -1 \\ 0 & 0 \end{bmatrix} \quad (10.27)$$

For the problem that we are considering here, it is simple to verify that the system Σ of Equation 10.22 has the following properties.

1. The subsystem (A, B, C_2, D_2) is invertible and of minimum phase with one invariant zero at -1.6867 . It also has one infinite zero of order four.
2. The subsystem (A, E, C_1, D_1) is left invertible and of minimum phase with one invariant zero at -0.9597 and two infinite zeros of orders one and two.

Then, it follows from the results of Section 3.5 that the H_∞ -ADDPMS for the system in Equation 10.22 is solvable. In fact, one can design either a full-order observer-based controller or a reduced-order observer-based controller to solve this problem. For the full-order observer-based controller, the order of the disturbance decoupling controller (see Figure 10.4) is five and the order of the final overall controller (again see Figure 10.4) is seven (the disturbance decoupling controller plus two integrators). On the other hand, if we use a reduced-order observer in the disturbance decoupling controller, the total order of the resulting final overall controller will be reduced to four. From the practical point of view, the latter is much more desirable than the former. Thus, in what follows we only focus on the controller design based on a reduced-order observer. We can separate our controller design into two steps:

1. In the first step, we assume that all five states of Σ in Equation 10.22 are available and then design a static and parameterized state feedback control law,

$$u = F(\varepsilon)x \quad (10.28)$$

such that it solves the almost disturbance decoupling problem for the state feedback case, *i.e.* $y = x$, by adjusting the tuning parameter ε to an appropriate value;

2. In the second step, we design a reduced-order observer-based controller. It has a parameterized reduced-order observer gain matrix $K_2(\varepsilon)$ that can be tuned to recover the performance achieved by the state feedback control law in the first step.

We use the structural decomposition approach of Chapter 3 to construct both the state feedback law and the reduced-order observer gain. We would like to note that, in principle, one can also apply the ARE-based H_∞ optimization technique (see, *e.g.*, Zhou and Khargonekar [93]) to solve this problem. However, because the numerical

conditions of our system Σ are very poor, we are unable to obtain any satisfactory solution from the ARE approach. We cannot get any meaningful solution for the associated H_∞ continuous-time ARE in MATLAB[®]. The following is a closed-form solution of the static state feedback parameterized gain matrix obtained using the method given in Chapter 3.

$$F(\varepsilon) = \begin{bmatrix} (2.1410 \times 10^6 - 62.3004/\varepsilon^2) & (570.7619 - 31.1502/\varepsilon) \\ 2.1410 \times 10^6 & -62.3004/\varepsilon^3 & -31.1502/\varepsilon^4 \end{bmatrix} \quad (10.29)$$

where ε is the tuning parameter that can be adjusted to achieve almost disturbance decoupling. It can be verified that the closed-loop system matrix, $A + BF(\varepsilon)$ is asymptotically stable for all $0 < \varepsilon < \infty$ and the closed-loop transfer function from the disturbance w to the controlled output h , $T_{hw}(\varepsilon, s)$, satisfying

$$\|T_{hw}(\varepsilon, s)\|_\infty = \|[C_2 + D_2F(\varepsilon)][sI - A - BF(\varepsilon)]^{-1}E\|_\infty \rightarrow 0 \quad (10.30)$$

as $\varepsilon \rightarrow 0$.

The next step is to design a reduced-order observer-based controller that recovers the performance of the above state feedback control law. First, let us perform the following nonsingular (permutation) state transformation to the system Σ of Equation 10.22,

$$x = T\tilde{x} \quad (10.31)$$

where

$$T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} \quad (10.32)$$

such that the transformed measurement matrix has the form of

$$C_1T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \end{bmatrix} = [I_3 \quad 0] \quad (10.33)$$

Clearly, the first three states of the transformed system, or x_1 , x_4 and x_5 of the original system Σ in Equation 10.22, need not be estimated as they are already available from the measurement output. Let us now partition the transformed system as follows:

$$\begin{aligned} T^{-1}AT &= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\ &= \left[\begin{array}{ccc|cc} 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ \hline -274921.63 & 0 & 0 & -73.2915 & -274921.63 \\ 0 & 0 & 0 & 0 & -0.9597 \end{array} \right] \quad (10.34) \end{aligned}$$

$$T^{-1}B = \begin{bmatrix} B_1 \\ B_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ \hline 0.12841 \\ -3.39561 \times 10^{-7} \end{bmatrix} \quad (10.35)$$

$$T^{-1}E = \begin{bmatrix} E_1 \\ E_2 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -1 \\ 0 & 0 \\ \hline -274921.63 & 0 \\ 0 & 0 \end{bmatrix} \quad (10.36)$$

Also, we partition

$$\begin{aligned} F(\varepsilon)T &= [F_1(\varepsilon) | F_2(\varepsilon)] \quad (10.37) \\ &= \left[\begin{array}{ccc|c} (2.14 \times 10^6 - 62.3/\varepsilon^2) & -62.3/\varepsilon^3 & -31.15/\varepsilon^4 & \\ \hline (570.76 - 31.15/\varepsilon) & 2.14 \times 10^6 & & \end{array} \right] \quad (10.38) \end{aligned}$$

Then the reduced-order observer-based controller (see Chapter 3) is given in the form of

$$\Sigma_{\text{cmp}} : \begin{cases} \dot{v} = A_{\text{cmp}}(\varepsilon) v + B_{\text{cmp}}(\varepsilon) y \\ u = C_{\text{cmp}}(\varepsilon) v + D_{\text{cmp}}(\varepsilon) y \end{cases} \quad (10.39)$$

with

$$A_{\text{cmp}}(\varepsilon) = A_{22} + K_2(\varepsilon)A_{12} + B_2F_2(\varepsilon) + K_2(\varepsilon)B_1F_2(\varepsilon) \quad (10.40)$$

$$\begin{aligned} B_{\text{cmp}}(\varepsilon) &= A_{21} + K_2(\varepsilon)A_{11} - [A_{22} + K_2(\varepsilon)A_{12}]K_2(\varepsilon) \\ &\quad + [B_2 + K_2(\varepsilon)B_1][F_1(\varepsilon) - F_2(\varepsilon)K_2(\varepsilon)] \end{aligned} \quad (10.41)$$

$$C_{\text{cmp}}(\varepsilon) = F_2(\varepsilon) \quad (10.42)$$

$$D_{\text{cmp}}(\varepsilon) = F_1(\varepsilon) - F_2(\varepsilon)K_2(\varepsilon) \quad (10.43)$$

where $K_2(\varepsilon)$ is the parameterized reduced-order observer gain matrix and is to be designed such that $A_{22} + K_2(\varepsilon)A_{12}$ is asymptotically stable for sufficiently small ε and also

$$\| [sI - A_{22} - K_2(\varepsilon)A_{12}]^{-1} [E_2 + K_2(\varepsilon)E_1] \|_{\infty} \rightarrow 0 \quad (10.44)$$

as $\varepsilon \rightarrow 0$. Again, using the software package of [53], we obtained the following parameterized reduced-order observer gain matrix

$$K_2(\varepsilon) = \begin{bmatrix} 73.2915 - 1/\varepsilon & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad (10.45)$$

Then the explicitly parameterized matrices of the state-space model of the reduced-order observer-based controller are given by

$$A_{\text{cmp}}(\varepsilon) = \begin{bmatrix} 73.2915 - 4/\varepsilon - 1/\varepsilon & 0 \\ -1.9381 \times 10^{-4} + 1.0577 \times 10^{-5}/\varepsilon & -1.6867 \end{bmatrix}$$

$$C_{\text{cmp}}(\varepsilon) = [570.7619 - 31.1502/\varepsilon \quad 2140967]$$

$$D_{\text{cmp}}(\varepsilon) = [2099135.4 + 2853.81/\varepsilon - 93.45/\varepsilon^2 \quad -62.3/\varepsilon^3 \quad -31.15/\varepsilon^4]$$

$$B_{\text{cmp}}(\varepsilon) = \begin{bmatrix} \psi_1 & -8/\varepsilon^3 & -4/\varepsilon^4 \\ \psi_2 & 2.1155 \times 10^{-5}/\varepsilon^3 & 1.0577 \times 10^{-5}/\varepsilon^4 \end{bmatrix}$$

where

$$\psi_1 = -5731.6533 - 13/\varepsilon^2 + 439.7492/\varepsilon \quad (10.46)$$

and

$$\psi_2 = -0.7128 + 3.1732 \times 10^{-5}/\varepsilon^2 - 9.6904 \times 10^{-4}/\varepsilon \quad (10.47)$$

The overall closed-loop system comprising the system Σ of Equation 10.22 and the above controller would be asymptotically stable as long as $\varepsilon \in (0, \infty)$. In fact, the closed-loop poles are exactly located at -1.6867 , two pairs at $-1/\varepsilon \pm j1/\varepsilon$, -0.9597 and $-1/\varepsilon$. The plots of the maximum singular values of the closed-loop transfer function matrix from the disturbance w to the controlled output h , namely $T_{hw}(\varepsilon, s)$, for several values of ε , *i.e.* $\varepsilon = 1/100$, $\varepsilon = 1/400$ and $\varepsilon = 1/3000$, in Figure 10.5 show that as ε tends to smaller values, the H_∞ -norm of $T_{hw}(\varepsilon, s)$ also becomes smaller. Hence, almost disturbance decoupling is indeed achieved. These are the properties of our control system in the frequency domain. In the next section we will address its time-domain properties, which are, of course, much more important, as all the design specifications are in the time domain.

10.4 Final Controller and Simulation Results

In this section we put our design of the previous section into a final controller as depicted in Figure 10.2. It is simple to derive the state-space model of the final overall controller by observing its interconnection with the disturbance decoupling controller Σ_{cmp} of Equation 10.39 (see Figure 10.4). We also present simulation results of the responses of the overall design to several different types of reference input signal. They clearly show that all the design specifications are successfully achieved. Furthermore, because our controller is explicitly parameterized by a tuning parameter, it is very easy to adjust it to meet other design specifications without going through it all over again from the beginning. This is also discussed next.

As mentioned earlier, the final overall controller of our design is of the order of four, of which two are from the disturbance decoupling controller and two from the augmented integrators. It has two inputs: one is the displacement x_1 and the other is the reference signal r . It is straightforward to verify that the state-space model of the final overall controller is given by

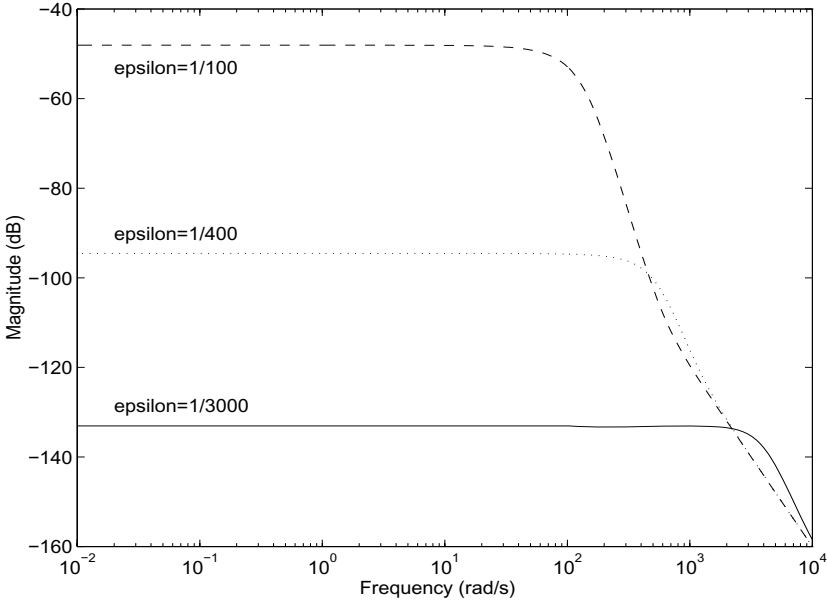


Figure 10.5. Maximum singular values of closed-loop transfer function $T_{hw}(\varepsilon, s)$

$$\Sigma_{oc}(\varepsilon) : \begin{cases} \dot{v} = A_{oc}(\varepsilon) v + B_{oc}(\varepsilon) x_1 + G_{oc} r \\ u = C_{oc}(\varepsilon) v + D_{oc}(\varepsilon) x_1 \end{cases} \quad (10.48)$$

where

$$A_{oc}(\varepsilon) = \begin{bmatrix} 73.29 - 5/\varepsilon & 0 & -8/\varepsilon^3 & -4/\varepsilon^4 \\ 1.06 \times 10^{-5}/\varepsilon & -1.69 & 2.12 \times 10^{-5}/\varepsilon^3 & 1.08 \times 10^{-5}/\varepsilon^4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (10.49)$$

$$G_{oc} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad B_{oc}(\varepsilon) = \begin{bmatrix} \psi_1 \\ \psi_2 \\ 1 \\ 0 \end{bmatrix} \quad (10.50)$$

with ψ_1 and ψ_2 given by Equations 10.46 and 10.47, respectively,

$$C_{oc}(\varepsilon) = [570.76 - 31.15/\varepsilon \quad 2140967 \quad -62.3/\varepsilon^3 \quad -31.15/\varepsilon^4] \quad (10.51)$$

and

$$D_{oc}(\varepsilon) = 2099135.4 - 93.45/\varepsilon^2 + 2853.81/\varepsilon \quad (10.52)$$

There are some very interesting and useful properties of this parameterized controller. After repeatedly simulating the overall design, we found that the maximum peak values of the control signal u are independent of the frequencies of the reference signals. They are only dependent on the initial error between displacement x_1

and the reference r . The larger the initial error, the bigger is the peak that occurs in u . Because the working range of our actuator is within $\pm 1 \mu\text{m}$, we assume that the largest magnitude of the initial error in any situation is not larger than $1 \mu\text{m}$. This assumption is reasonable, as we can always reset our displacement x_1 to zero before the system is to track any reference and hence the magnitude of initial tracking error can never be larger than $1 \mu\text{m}$. Let us consider the worst case, *i.e.* that the magnitude of the initial error is $1 \mu\text{m}$. Then, interestingly, we are able to obtain a clear relationship between the tuning parameter $1/\varepsilon$ and the maximum peak of u . The result is plotted in Figure 10.6. We also found that the tracking error is independent of initial errors. It only depends on the frequencies of the references, *i.e.* the larger the frequency that the reference signal r has, the larger the tracking error that occurs. Again, we can obtain a simple and linear relationship between the tuning parameter ε and the maximum frequency that a reference signal can have such that the corresponding tracking error is no larger than 1%, which is one of our main design specifications. The result is plotted in Figure 10.7.

Clearly, from Figure 10.6, we know that owing to the constraints on the control input, *i.e.* it must be kept within $\pm 112.5 \text{ V}$, we have to select our controller with $\varepsilon > 1/3370$. From Figure 10.7, we know that in order to meet the first design specification, *i.e.* the steady-state tracking errors are less than 1% for reference inputs that have frequencies up to 30 Hz, we have to choose our controller with $\varepsilon < 1/2680$. Hence, the final controller as given in Equations 10.48 to 10.52 meets all the design goals for our piezoelectric actuator system of Equations 10.1 and 10.2, for all $\varepsilon \in (1/3370, 1/2680)$. Let us choose $\varepsilon = 1/3000$. We obtain the overall controller as in the form of Equation 10.48 with

$$A_{\text{oc}} = \begin{bmatrix} -14926.71 & 0 & -2.16 \times 10^{11} & -3.24 \times 10^{14} \\ 0.0315 & -1.69 & 5.71 \times 10^5 & 8.57 \times 10^8 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \quad (10.53)$$

$$B_{\text{oc}} = \begin{bmatrix} -1.16 \times 10^8 \\ 281.97 \\ 1 \\ 0 \end{bmatrix}, \quad G_{\text{oc}} = \begin{bmatrix} 0 \\ 0 \\ -1 \\ 0 \end{bmatrix}, \quad D_{\text{oc}} = -8.3 \times 10^8 \quad (10.54)$$

$$C_{\text{oc}} = [-92879.9 \quad 2140967 \quad -1.68 \times 10^{12} \quad -2.52 \times 10^{15}] \quad (10.55)$$

The simulation results presented in the following are done using the Simulink® package in MATLAB®, which is widely available these days. Two different reference inputs are simulated using the Runge–Kutta 5 method in Simulink® with a minimum step size of $10 \mu\text{s}$ and a maximum step size of $100 \mu\text{s}$ as well as a tolerance of 10^{-5} . These references are: 1) a cosine signal with a frequency of 30 Hz and peak magnitude of $1 \mu\text{m}$, and 2) a sine signal with a frequency of 34 Hz and peak magnitude of $1 \mu\text{m}$. The results for the cosine signal are given in Figures 10.8 to 10.10. In Figure 10.8, the solid-line curve is x_1 and the dash-dotted curve is the reference. The tracking error and the control signal corresponding to this reference are given in Figures 10.9 and 10.10, respectively. Similarly, Figures 10.11 to 10.13

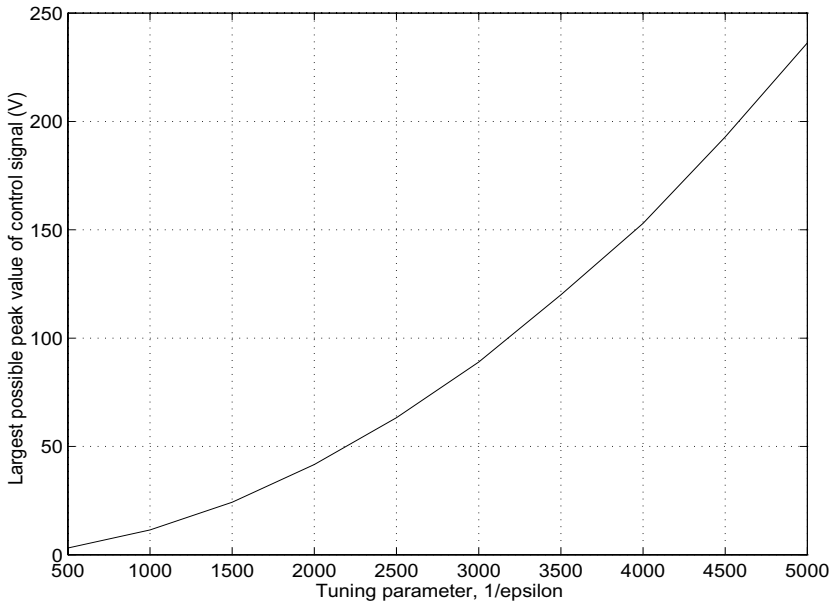


Figure 10.6. Parameter $1/\epsilon$ versus maximum peaks of u in worst initial errors

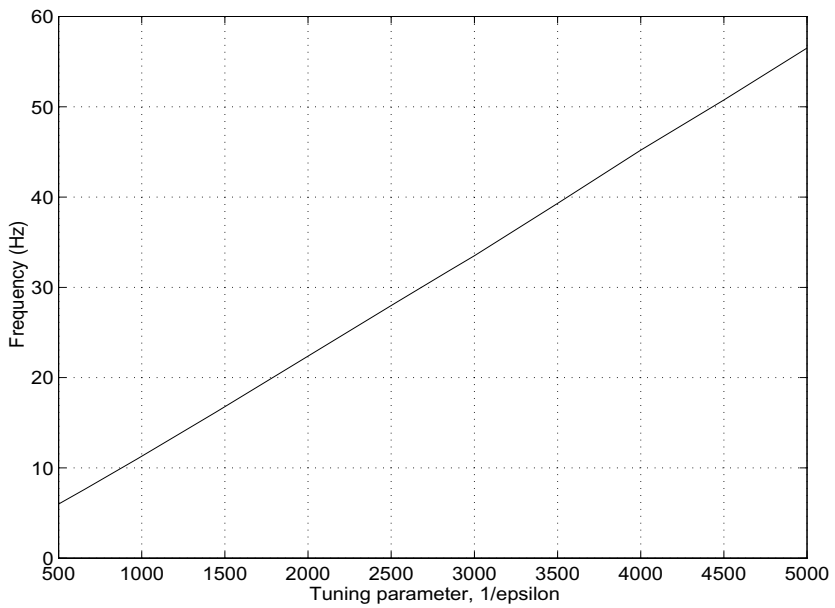


Figure 10.7. Parameter $1/\epsilon$ versus maximum frequency of r that has 1% tracking error

are the results corresponding to the sine signal. All these results show that our design goals are fully achieved. To be more specific, the tracking error for a 30-Hz cosine wave reference is about 0.8%, which is better than the specification, and the worst peak magnitude of the control signal is less than 90 V, which is of course less than the saturated level, *i.e.* 112.5 V. Furthermore, the 1% tracking error settling times for both cases are less than 0.003 s. It is interesting to note that the performance of the actual closed-loop system with the nonlinear hysteretic dynamics is even better than that of its linearized counterpart.

Although we do not consider frequency-domain specifications in our design, the sensitivity and complementary sensitivity functions, the Bode plot and Nyquist plot (near the origin for verification) given, respectively, in Figures 10.14 and 10.15 show that the overall control system turns out to have impressive robustness with respect to disturbances and measurement noise, and impressive gain and phase margins. In particular, the system has a lower gain margin of 0.487 and an infinite upper gain margin, and a phase margin of 30.58° .

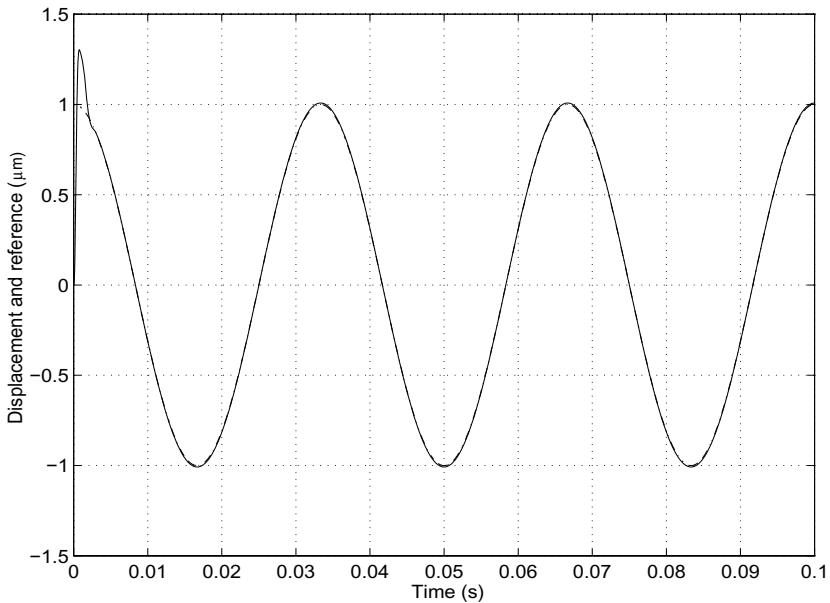
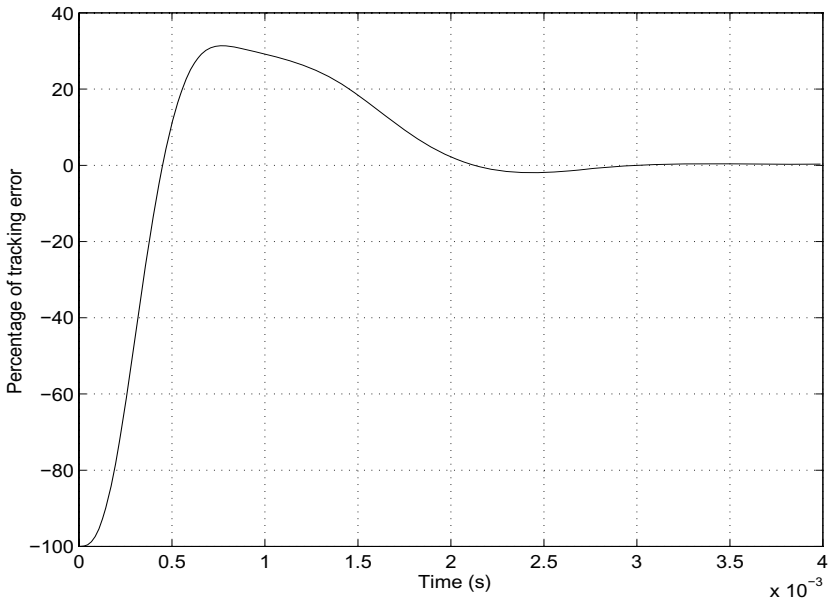
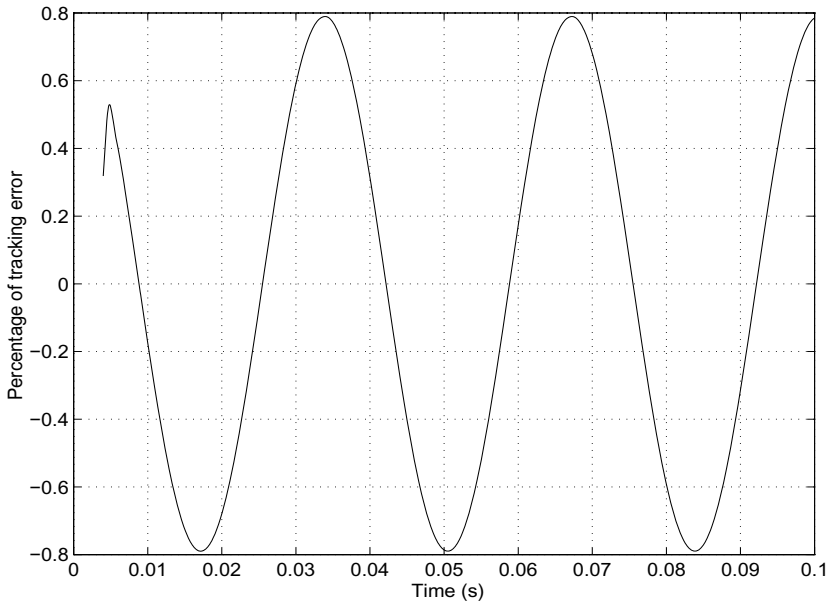


Figure 10.8. Responses of the displacement and the 30-Hz cosine reference

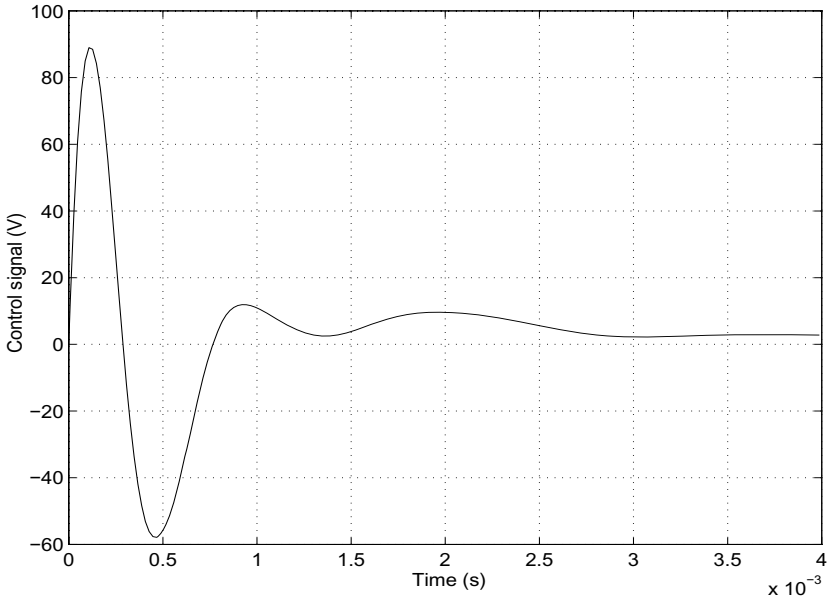


(a) Tracking error from 0 to 0.004 s

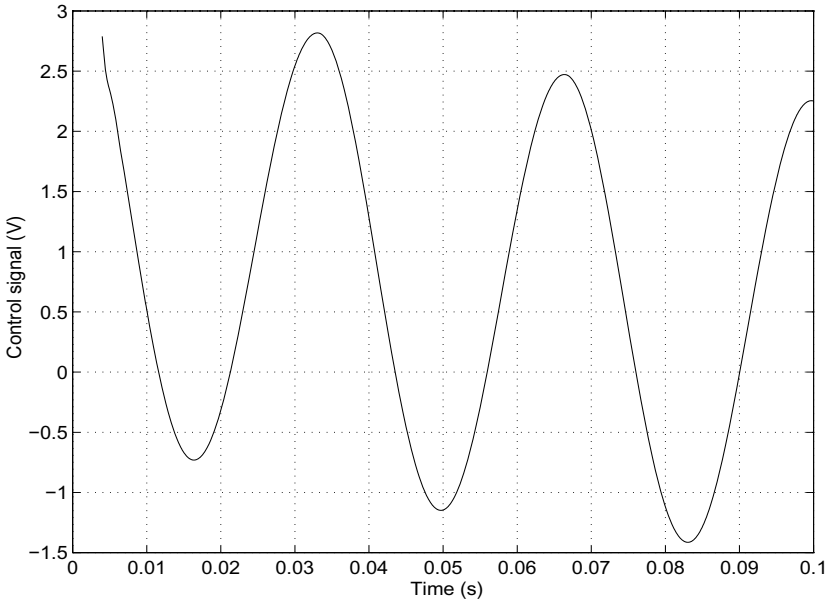


(b) Tracking error from 0.004 to 0.1 s

Figure 10.9. Tracking error for the 30-Hz cosine reference



(a) Control signal from 0 to 0.004 s



(b) Control signal from 0.004 to 0.1 s

Figure 10.10. Control signal for the 30-Hz cosine reference

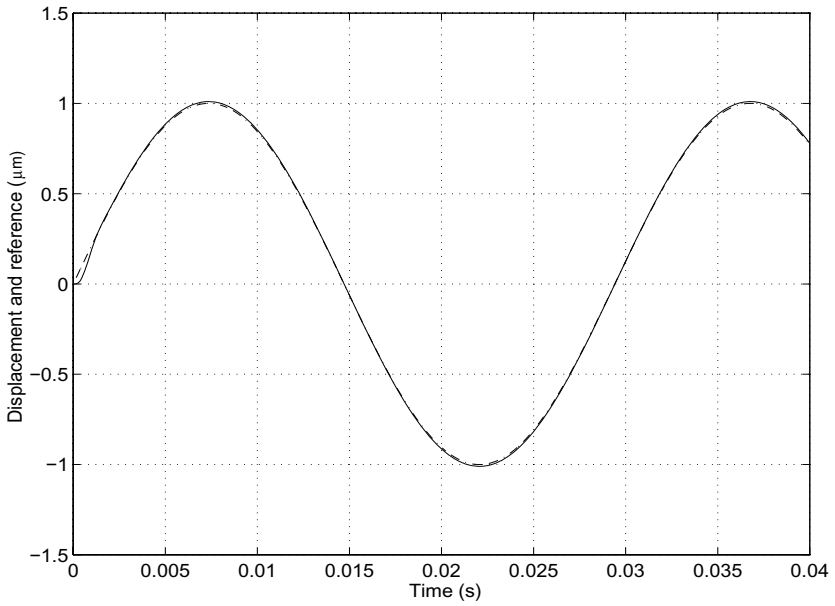


Figure 10.11. Responses of the displacement and the 34-Hz sine reference

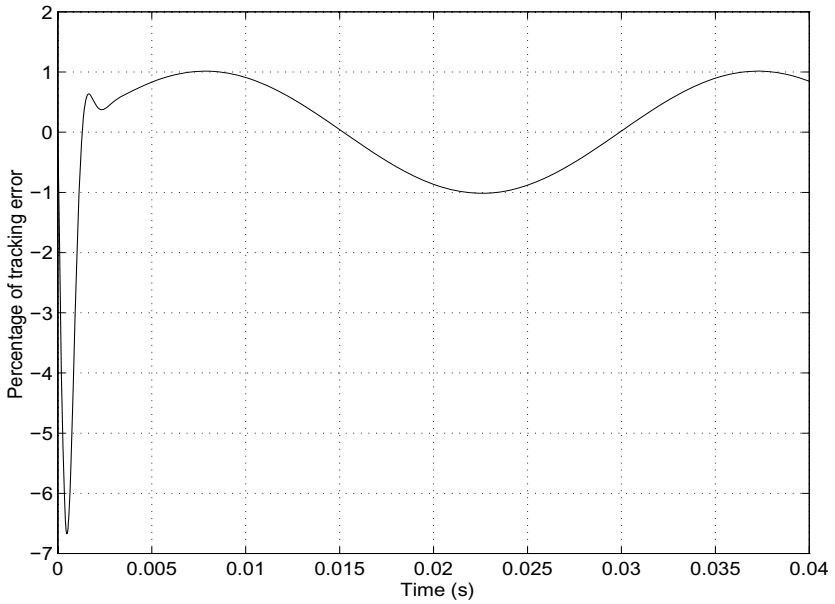


Figure 10.12. Tracking error for the 34-Hz sine reference

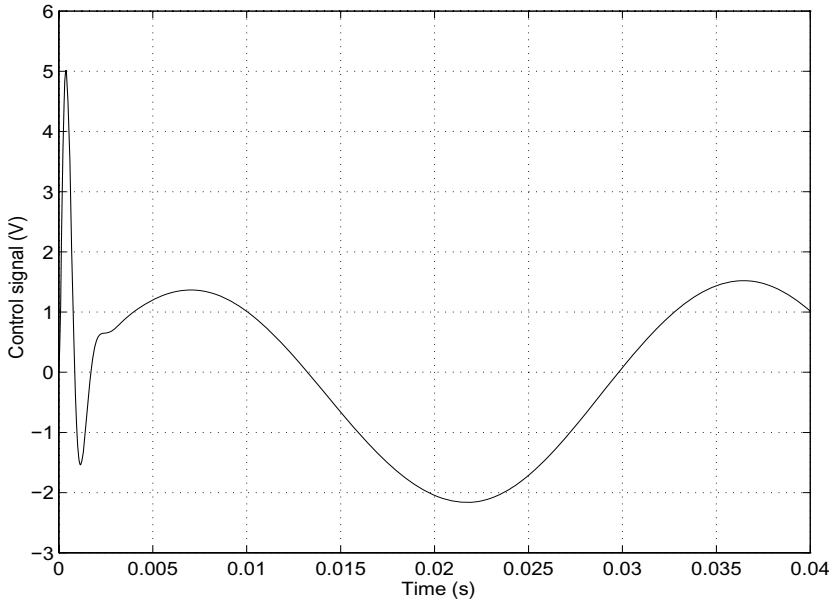


Figure 10.13. Control signal for the 34-Hz sine reference

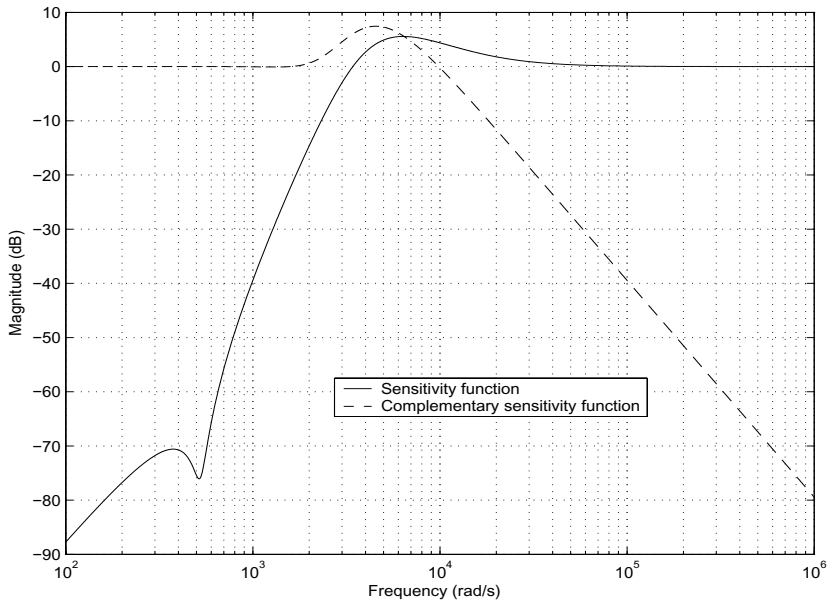
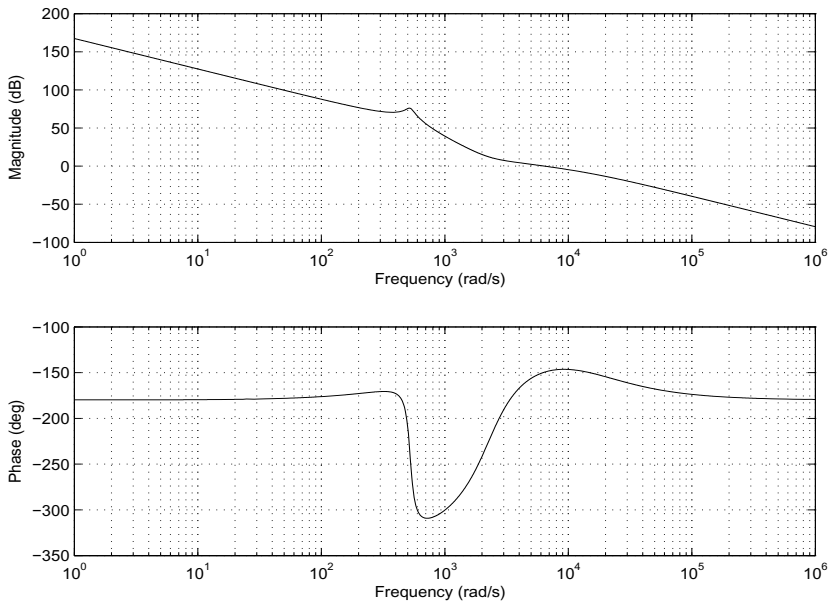
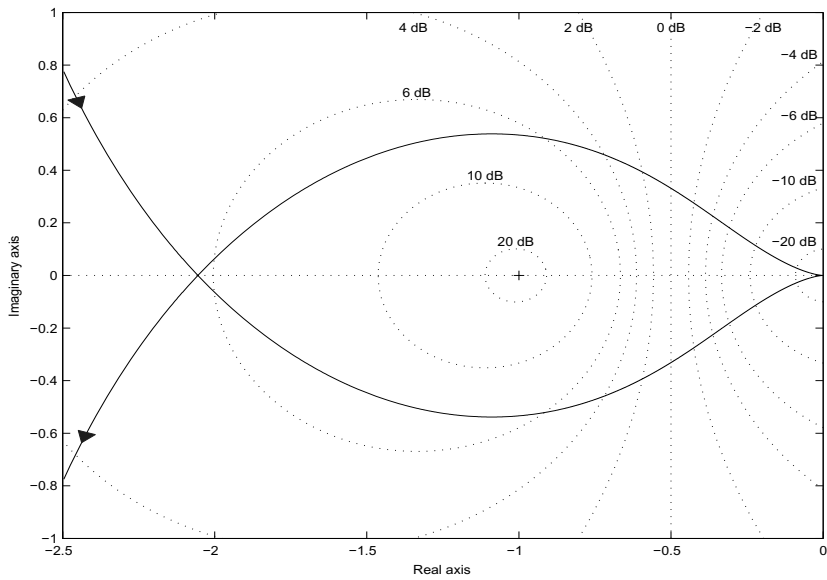


Figure 10.14. Sensitivity and complementary sensitivity functions of overall control system



(a) Bode plot



(b) Nyquist plot

Figure 10.15. Bode and Nyquist plots of the overall control system

A Benchmark Problem

Before ending this book, we post in this chapter a typical HDD servo control design problem. The problem has been tackled in the previous chapters using several design methods, such as PID, RPT, CNF, PTOS and MSC control. We feel that it can serve as an interesting and excellent benchmark example for testing other linear and nonlinear control techniques.

We recall that the complete dynamics model of a Maxtor (Model 51536U3) hard drive VCM actuator can be depicted as in Figure 11.1:

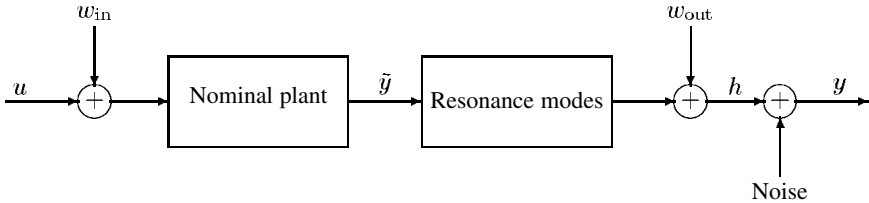


Figure 11.1. Block diagram of the dynamical model of the hard drive VCM actuator

The nominal plant of the HDD VCM actuator is characterized by the following second-order system:

$$\dot{x} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 6.4013 \times 10^7 \end{bmatrix} (\text{sat}(u) + w_{\text{in}}) \quad (11.1)$$

and

$$\tilde{y} = [1 \quad 0] x \quad (11.2)$$

where the control input u is limited within ± 3 V and w_{in} is an unknown input disturbance with $|w_{\text{in}}| \leq 3$ mV. For simplicity and for simulation purpose, we assume that the unknown disturbance $w_{\text{in}} = -3$ mV. The measurement output available for

control, *i.e.* y (in μm), is the measured displacement of the VCM R/W head and is given by

$$y = \left[\prod_{i=1}^4 G_{r,i}(s) \right] \tilde{y} + w_{\text{out}} + \text{Noise} \quad (11.3)$$

where the transfer functions of the resonance modes are given by

$$\left. \begin{aligned} G_{r,1}(s) &= \frac{0.912s^2 + 457.4s + 1.433(1 + \delta) \times 10^8}{s^2 + 359.2s + 1.433(1 + \delta) \times 10^8} \\ G_{r,2}(s) &= \frac{0.7586s^2 + 962.2s + 2.491(1 + \delta) \times 10^8}{s^2 + 789.1s + 2.491(1 + \delta) \times 10^8} \\ G_{r,3}(s) &= \frac{9.917(1 + \delta) \times 10^8}{s^2 + 1575s + 9.917(1 + \delta) \times 10^8} \\ G_{r,4}(s) &= \frac{2.731(1 + \delta) \times 10^9}{s^2 + 2613s + 2.731(1 + \delta) \times 10^9} \end{aligned} \right\} \quad (11.4)$$

with $-20\% \leq \delta \leq 20\%$ represents the variation of the resonance modes of the actual actuators whose resonant dynamics change from time to time and also from disk to disk in a batch of million drives. Note that many new hard drives in the market nowadays might have resonance modes at much higher frequencies (such as those for the IBM microdrives studied in Chapter 9). But, structurewise, they are almost the same. The output disturbance (in μm), which is mainly the repeatable runouts, is given by

$$w_{\text{out}} = 0.1 \sin(110\pi t) + 0.05 \sin(220\pi t) + 0.02 \sin(440\pi t) + 0.01 \sin(880\pi t) \quad (11.5)$$

and the measurement noise is assumed to be a zero-mean Gaussian white noise with a variance $\sigma_n^2 = 9 \times 10^{-6} (\mu\text{m})^2$.

The problem is to design a controller such that when it is applied to the VCM actuator system, the resulting closed-loop system is asymptotically stable and the actual displacement of the actuator, *i.e.* h , tracks a reference $r = 1 \mu\text{m}$. The overall design has to meet the following specifications:

1. the overshoot of the actual actuator output is less than 5%;
2. the mean of the steady-state error is zero;
3. the gain margin and phase margin of the overall design are, respectively, greater than 6 dB and 30° ; and
4. the maximum peaks of the sensitivity and complementary sensitivity functions are less than 6 dB.

The results of Chapter 6 show that the 5% settling times of our design using the CNF control technique are, respectively, 0.80 ms in simulation and 0.85 ms in actual hardware implementation. We note that the simulation result can be further improved if we do not consider actual hardware constraints in our design. For example, the

CNF control law given below meets all design specifications and achieves a 5% settling time of 0.68 ms. It is obtained by using the toolkit of [55] under the option of the pole-placement method with a damping ratio of 0.1 and a natural frequency of 2800 rad/sec together with a diagonal matrix $W = \text{diag}\{1.5, 0.01, 2 \times 10^{-10}\}$. The dynamic equation of the control law is given by

$$\begin{pmatrix} \dot{x}_i \\ \dot{x}_v \end{pmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & -4000 \end{bmatrix} \begin{pmatrix} x_i \\ x_v \end{pmatrix} - \begin{bmatrix} -10 \\ 1.6 \times 10^7 \end{bmatrix} y + \begin{bmatrix} 0 \\ 6.4013 \times 10^7 \end{bmatrix} \text{sat}(\tilde{u}) - \begin{pmatrix} 10r \\ 0 \end{pmatrix} \quad (11.6)$$

$$\begin{aligned} \tilde{u} = \rho(e) [0.61237 \quad 0.049449 \quad 8.4622 \times 10^{-5}] & \begin{pmatrix} x_i \\ y - r \\ x_v + 4000y \end{pmatrix} \\ & - [1.2248 \quad 0.12335 \quad 1.0310 \times 10^{-5}] \begin{pmatrix} x_i \\ y - r \\ x_v + 4000y \end{pmatrix} \end{aligned} \quad (11.7)$$

where

$$\rho(e) = -2.65 \left| e^{-0.7|e|} - e^{-0.7} \right| \quad (11.8)$$

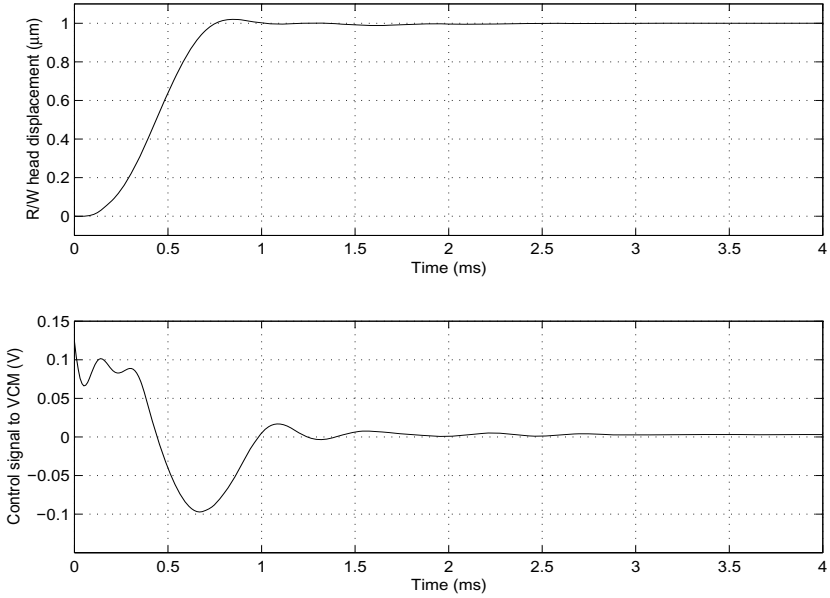
and

$$u = G_{\text{notch}}(s) \cdot \tilde{u} \quad (11.9)$$

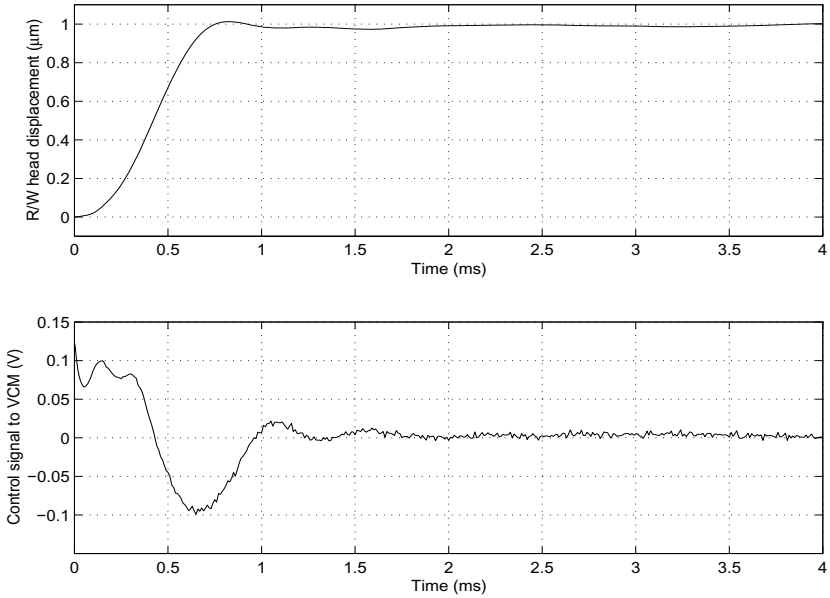
with $G_{\text{notch}}(s)$ being given as in Equation 6.9.

The simulation results obtained with $\delta = 0$ given in Figures 11.2 to 11.4 show that all the design specifications have been achieved. In particular, the resulting 5% settling time is 0.68 ms, the gain margin is 7.85 dB and the phase margin is 44.7° , and finally, the maximum values of the sensitivity and complementary sensitivity functions are less than 5 dB. The overall control system can still produce a satisfactory result and satisfy all the design specifications by varying the resonance modes with the value of δ changing from -20% to 20% .

Nonetheless, we invite interested readers to challenge our design. Noting that for the track-following case, *i.e.* when $r = 1 \mu\text{m}$, the control signal is far below its saturation level. Because of the bandwidth constraint of the overall system, it is not possible (and not necessary) to utilize the full scale of the control input to the actuator in the track-following stage. However, in the track-seeking case or equivalently by setting a larger target reference, say $r = 500 \mu\text{m}$, the very problem can serve as a good testbed for control techniques developed for systems with actuator saturation. Interested readers are referred to Chapter 7 for more information on track seeking of HDD servo systems.

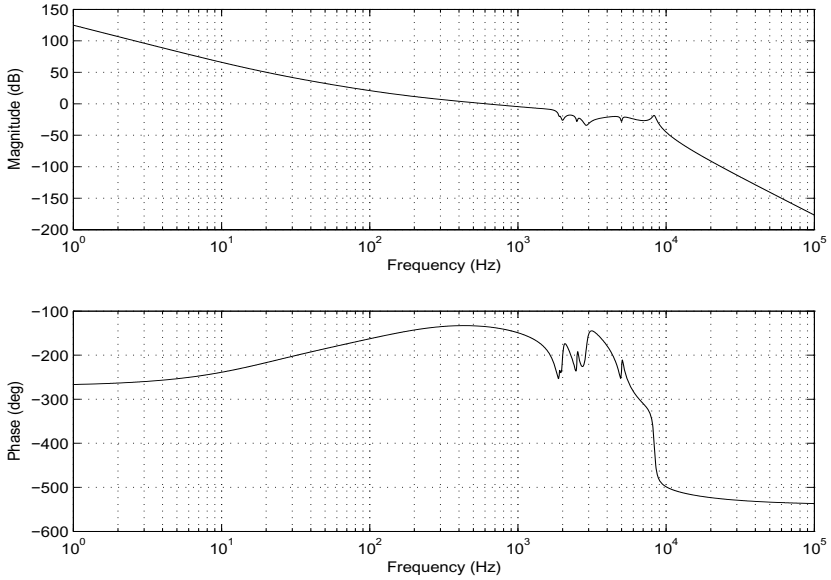


(a) h and u for the system without output disturbance and noise

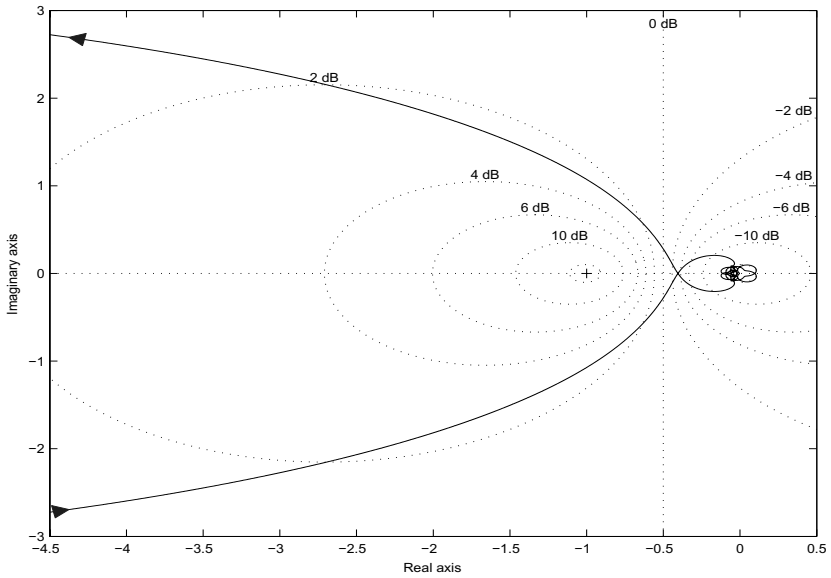


(b) h and u for the system with output disturbance and noise

Figure 11.2. Output responses and control signals of the CNF control system



(a) Bode plot



(b) Nyquist plot

Figure 11.3. Bode and Nyquist plots of the CNF control system

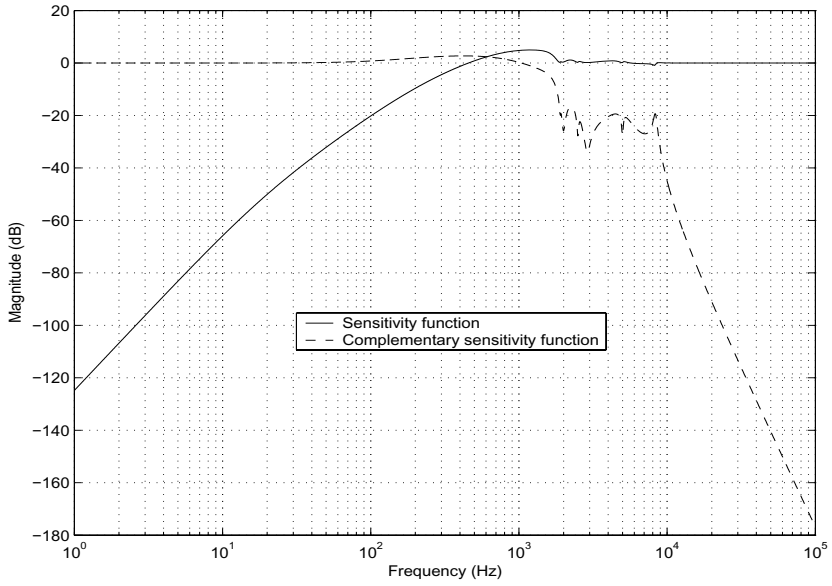


Figure 11.4. Sensitivity and complementary sensitivity functions with the CNF control

References

1. Franklin GF, Powell JD, Workman ML. Digital control of dynamic systems. 3rd edn Reading (MA): Addison-Wesley; 1998.
2. Fujimoto H, Hori Y, Yarnaguchi T, Nakagawa S. Proposal of seeking control of hard disk drives based on perfect tracking control using multirate feedforward control. Proc 6th Int Workshop Adv Motion Contr; Nagoya, Japan; 2000. p. 74–9.
3. Goh TB, Li Z, Chen BM, Lee TH, Huang T. Design and implementation of a hard disk drive servo system using robust and perfect tracking approach. IEEE Trans Contr Syst Technol 2001; 9:221–33.
4. Gu Y, Tomizuka M. Digital redesign and multi-rate control for motion control – a general approach and application to hard disk drive servo system. Proc 6th Int Workshop Adv Motion Contr; Nagoya, Japan; 2000. p. 246–51.
5. Hara S, Hara T, Yi L, Tomizuka M. Two degree-of-freedom controllers for hard disk drives with novel reference signal generation. Proc American Contr Conf; San Diego, CA; 1999. p. 4132–6.
6. Huang Y, Messner WC, Steele J. Feed-forward algorithms for time-optimal settling of hard disk drive servo systems. Proc 23rd Int Conf Ind Electron Contr Instrum; New Orleans, LA; 1997. p. 52–7.
7. Ho HT. Fast bang-bang servo control. IEEE Trans Magn 1997; 33:4522–7.
8. Ishikawa J, Tomizuka M. A novel add-on compensator for cancellation of pivot nonlinearities in hard disk drives. IEEE Trans on Magn 1998; 34:1895–7.
9. Iwashiro M, Yatsu M, Suzuki H. Time optimal track-to-track seek control by model following deadbeat control. IEEE Trans Magn 1999; 35:904–9.
10. Pao LY, Franklin GF. Proximate time-optimal control of third-order servomechanisms. IEEE Trans Automat Contr 1993; 38:560–80.
11. Patten WN, Wu HC, White L. A minimum time seek controller for a disk drive. IEEE Trans Magn 1995; 31:2380–7.
12. Takakura S. Design of a tracking system using n -delay two-degree-of-freedom control and its application to hard disk drives. Proc 1999 IEEE Int Conf Contr Appl; Kohala Coast, HI; 1999. p. 170–5.
13. Wang L, Yuan L, Chen BM, Lee TH. Modeling and control of a dual actuator servo system for hard disk drives. Proc 1998 Int Conf Mechatron Technol; Hsinchu, Taiwan; 1998. p. 533–8.
14. Weerasooriya S, Low TS, Mamun AA. Design of a time optimal variable structure controller for a disk drive actuator. Proc Int Conf Ind Electron Contr Instrum; Hawaii; 1993. p. 2161–5.

15. Yamaguchi T, Soyama Y, Hosokawa H, Tsuneta K, Hirai H. Improvement of settling response of disk drive head positioning servo using mode switching control with initial value compensation. *IEEE Trans Magn* 1996; 32:1767–72.
16. Zhang DQ, Guo GX. Discrete-time sliding mode proximate time optimal seek control of hard disk drives. *IEE Proc Contr Theory Appl* 2000; 147:440–6.
17. Chang JK, Ho HT. LQG/LTR frequency loop shaping to improve TMR budget. *IEEE Trans Magn* 1999; 35:2280–82.
18. Hanselmann H, Engelke A. LQG-control of a highly resonant disk drive head positioning actuator. *IEEE Trans Ind Electron* 1988; 35:100–4.
19. Weerasooriya S, Phan DT. Discrete-time LQG/LTR design and modeling of a disk drive actuator tracking servo system. *IEEE Trans Ind Electron* 1995; 42:240–7.
20. Chen BM, Lee TH, Hang CC, Guo Y, Weerasooriya S. An H_∞ almost disturbance decoupling robust controller design for a piezoelectric bimorph actuator with hysteresis. *IEEE Trans Contr Syst Technol* 1997; 7:160–74.
21. Hirata M, Atsumi T, Murase A, Nonami K. Following control of a hard disk drive by using sampled-data H_∞ control. *Proc 1999 IEEE Int Conf Contr Appl*; Kohala Coast, HI; 1999. p. 182–6.
22. Li Y, Tomizuka M. Two degree-of-freedom control with adaptive robust control for hard disk servo systems. *IEEE/ASME Trans Mechatron* 1999; 4:17–24.
23. Hirata M, Liu KZ, Mita T, Yamaguchi T. Head positioning control of a hard disk drive using H_∞ theory. *Proc 31st IEEE Conf Dec Contr*; Tucson, AZ; 1992. p. 2460–1.
24. Kim BK, Chung WK, Lee HS, Choi HT, Suh IH, Chang YH. Robust time optimal controller design for hard disk drives. *IEEE Trans Magn* 1999; 35:3598–607.
25. Teo YT, Tay TT. Application of the l_1 optimal regulation strategy to a hard disk servo system. *IEEE Trans Contr Syst Technol* 1996; 4:467–72.
26. Du CL, Xie LH, Teoh JN, Guo GX. An improved mixed H_2/H_∞ control design for hard disk drives. *IEEE Trans Contr Syst Technol* 2005; 13:832–9.
27. Chen R, Guo G, Huang T, Low TS. Adaptive multirate control for embedded HDD servo systems. *Proc 24th Int Conf Ind Electron Contr Instrum*; Aachen, Germany; 1998. p. 1716–20.
28. McCormick J, Horowitz R. A direct adaptive control scheme for disk file servos. *Proc 1993 American Contr Conf*; San Francisco, CA; 1993. p. 346–51.
29. Weerasooriya S, Low TS. Adaptive sliding mode control of a disk drive actuator. *Proc Asia-Pacific Workshop Adv Motion Contr*; Singapore, 1993. p. 177–82.
30. Workman ML. Adaptive proximate time optimal servomechanisms [PhD diss]. Stanford University; 1987.
31. Internet websites: www.storagereview.com; www.pcguide.com; www.storage.ibm.com; www.mkdata.dk; 2001.
32. Porter J. Disk drives' evolution. Presented at 100th Ann Conf Magn Rec Info Stor; Santa Clara, CA; 1998.
33. Mamun AA. Servo engineering. Lecture Notes; Dept of Electrical & Computer Engineering, National University of Singapore; 2000.
34. Zhang JL. Electronic computer magnetic storing devices. Beijing: Military Industry Publishing House; 1981 (in Chinese).
35. Li Q, Hong OE, Hui Z, Mannan MA, Weerasooriya S, Ann MY, Chen SX, Wood R. Analysis of the dynamics of 3.5" hard disk drive actuators. Technical Report. Data Storage Institute (Singapore); 1997.
36. Chang JK, Weerasooriya S, Ho HT. Improved TMR through a frequency shaped servo design. *Proc 23rd Int Conf Ind Electron Contr Instrum*; New Orleans, LA; 1997. p. 47–51.

37. Li Z, Guo G, Chen BM, Lee TH. Optimal control design to achieve highest track-per-inch in hard disk drives. *J Inform Stor Proc Syst* 2001; 3:27–41.
38. Sacks AH, Bodson M, Messner W. Advanced methods for repeatable runout compensation. *IEEE Trans Magn* 1995; 31:1031–6.
39. Bodson M, Sacks A, Khosla P. Harmonic generation in adaptive feedforward cancellation schemes. *IEEE Trans Automat Contr* 1994; 39:1939–44.
40. Guo L, Tomizuka M. High speed and high precision motion control with an optimal hybrid feedforward controller. *IEEE/ASME Trans Mechatron* 1997; 2:110–22.
41. Kempf C, Messner W, Tomizuka M, Horowitz R. A comparison of four discrete-time repetitive control algorithms. *IEEE Contr Syst Mag* 1993; 13:48–54.
42. Guo L. A new disturbance rejection scheme for hard disk drive control. *Proc American Contr Conf; Philadelphia, PA; 1998*. p. 1553–7.
43. Weerasooriya S. Learning and compensation for repeatable runout of a disk drive servo using a recurrent neural network. Technical Report; Magnetics Technology Center, National University of Singapore; 1995.
44. Jamg G, Kim D, Oh JE. New frequency domain method of nonrepeatable runout measurement in a hard disk drive spindle motor. *IEEE Trans Magn* 1999; 35:833–8.
45. Ohmi T. Non-repeatable runout of ball-bearing spindle-motor for 2.5" HDD. *IEEE Trans Magn* 1996; 32:1715–20.
46. Abramovitch D, Hurst T, Henze D. An overview of the PES Pareto method for decomposing baseline noise sources in hard disk position error signals. *IEEE Trans Magn* 1998; 34:17–23.
47. Zeng S, Lin RM, Xu LM. Novel method for minimizing track seeking residual vibrations of hard disk drives. *IEEE Trans Magn* 2001; 37:1146–56.
48. Mah YA, Lin H, Li QH. Design of a high bandwidth moving-coil actuator with force couple actuation. *IEEE Trans Magn* 1999; 35:874–8.
49. McAllister JS. The effect of disk platter resonances on track misregistration in 3.5" disk drives. *IEEE Trans Magn* 1996; 32:1762–6.
50. Hanselmann H, Mortix W. High-bandwidth control of the head positioning mechanism in a winchester disk drive. *IEEE Contr Syst Mag* 1987; 7:15–9.
51. Guo G. Lecture notes in servo engineering. Dept of Electrical and Computer Engineering, National University of Singapore; 1998.
52. Weaver PA, Ehrlich RM. The use of multirate notch filters in embedded servo disk drives. *Proc American Contr Conf; San Francisco, CA; 1993*. p. 4156–60.
53. Lin Z, Chen BM, Liu X. Linear systems toolkit. Technical Report, Dept of Electrical and Computer Engineering, University of Virginia, Charlottesville (VA); 2004.
54. Cheng G, Chen BM, Peng K, Lee TH. A Matlab toolkit for composite nonlinear feedback control. *Proc 8th Int Conf Contr Automat Robot Vision; Kunming, China; 2004*. p. 878–83.
55. Cheng G, Chen BM, Peng K, Lee TH. A Matlab toolkit for composite nonlinear feedback control. Available online at <http://hdd.ece.nus.edu.sg/bmchen/>.
56. Chen BM, Lee TH, Venkataraman V. Hard disk drive servo systems. New York (NY): Springer; 2002.
57. Ljung L. System identification: theory for the user. Englewood Cliffs (NJ): Prentice Hall; 1987.
58. Sinha NK, Kuszta B. Modeling and identification of dynamic systems. New York (NY): Van Nostrand Reinhold Company; 1983.
59. Eykhoff P. System identification – parameter and state estimation. New York (NY): John Wiley; 1981.

60. Hsia TC. System identification. Lexington (MA): Lexington Books; 1977.
61. Sage AP, Melsa JL. System identification. New York (NY): Academic Press; 1971.
62. Dammers D, Binet P, Pelz G, Vobkarper LM. Motor modeling based on physical effect models. Proc 2001 IEEE Int Workshop on Behav Model Simulation; Santa Rosa, CA; 2001. p. 78–83.
63. Dubi A. Monte Carlo applications in systems engineering. New York (NY): John Wiley; 2000.
64. Evans M, Swartz T. Approximating integrals via Monte Carlo and deterministic methods. London: Oxford University Press; 2000.
65. Mikhailov GA. Parametric estimates by the Monte Carlo method. Utrecht, The Netherlands: V.S.P. International Science; 1999.
66. Rake H. Step response and frequency response methods. Automatica 1980; 16:519–26.
67. Eykhoff P. Trends and progress in system identification. New York (NY): Pergamon Press; 1981.
68. Chen J, Gu G. Control oriented system identification: an H_∞ approach. New York (NY): John Wiley & Sons; 2000.
69. Gong JQ, Guo L, Lee HS, Yao B. Modeling and cancellation of pivot nonlinearity in hard disk drives IEEE Trans Magn 2002; 38:3560–5.
70. Liu X, Chen BM, Lin Z. On the problem of general structural assignments of linear systems through sensor/actuator selection. Automatica 2003; 39:233–41.
71. Chen BM, Lin Z, Shamash Y. Linear systems theory: a structural decomposition approach. Boston: Birkhäuser; 2004.
72. Sannuti P, Saberi A. A special coordinate basis of multivariable linear systems – Finite and infinite zero structure, squaring down and decoupling. Int J Contr 1987; 45:1655–704.
73. Saberi A, Sannuti P. Squaring down of non-strictly proper systems. Int J Contr 1990; 51:621–9.
74. Chen BM. Robust and H_∞ Control. London: Springer; 2000.
75. Rosenbrock HH. State-space and Multivariable Theory. New York (NY): John Wiley; 1970.
76. MacFarlane AGJ, Karcanias N. Poles and zeros of linear multivariable systems: a survey of the algebraic, geometric and complex variable theory. Int J Contr 1976; 24:33–74.
77. Commault C, Dion JM. Structure at infinity of linear multivariable systems: a geometric approach. IEEE Trans Automat Contr 1982; 27:693–6.
78. Pugh AC, Ratcliffe PA. On the zeros and poles of a rational matrix. Int J Contr 1979; 30:213–27.
79. Verghese G. Infinite frequency behavior in generalized dynamical systems [PhD diss]. Stanford University; 1978.
80. Owens DH. Invariant zeros of multivariable systems: a geometric analysis. Int J Contr 1978; 28:187–98.
81. Morse AS. Structural invariants of linear multivariable systems. SIAM J Contr 1973; 11:446–65.
82. Moylan P. Stable inversion of linear systems. IEEE Trans Automat Contr 1977; 22:74–8.
83. Scherer C. H_∞ -optimization without assumptions on finite or infinite zeros. SIAM J Contr Optimiz 1992; 30:143–66.
84. Ziegler JG, Nichols NB. Optimum settings for automatic controllers. Trans ASME 1942; 64:759–68.
85. Ziegler JG, Nichols NB. Process lags in automatic control circuits. Trans ASME 1943; 65:433–44.

86. Franklin GF, Powell JD, Emami-Naeini A. Feedback control of dynamic systems. 3rd edn Reading (MA): Addison-Wesley; 1994.
87. Anderson BDO, Moore JB. Optimal control: linear quadratic methods. Englewood Cliffs (NJ): Prentice Hall; 1989.
88. Fleming WH, Rishel RW. Deterministic and stochastic optimal control. New York (NY): Springer-Verlag; 1975.
89. Kwakernaak H, Sivan R. Linear optimal control systems. New York (NY): John Wiley; 1972.
90. Saberi A, Sannuti P, Chen BM. H_2 optimal control. London: Prentice Hall; 1995.
91. Doyle J, Glover K, Khargonekar PP, Francis BA. State space solutions to standard H_2 and H_∞ control problems. IEEE Trans Automat Contr 1989; 34:831–47.
92. Chen BM, Saberi A, Sannuti P, Shamash Y. Construction and parameterization of all static and dynamic H_2 -optimal state feedback solutions, optimal fixed modes and fixed decoupling zeros. IEEE Trans Automat Contr 1993; 38:248–61.
93. Zhou K, Khargonekar P. An algebraic Riccati equation approach to H_∞ -optimization. Syst Contr Lett 1988; 11:85–91.
94. Chen BM, Saberi A, Bingularac S, Sannuti P. Loop transfer recovery for non-strictly proper plants. Contr Theory Adv Technol 1990; 6:573–94.
95. Zames G. Feedback and optimal sensitivity: Model reference transformations, multiplicative seminorms, and approximate inverses. IEEE Trans Automat Contr 1981; 26:301–20.
96. Limebeer DJN, Anderson BDO. An interpolation theory approach to H_∞ controller degree bounds. Linear Algebra Appl 1988; 98:347–86.
97. Doyle JC. Lecture notes in advances in multivariable control. ONR-Honeywell Workshop; 1984.
98. Francis BA. A course in H_∞ control theory. Berlin: Springer; 1987.
99. Glover K. All optimal Hankel-norm approximations of linear multivariable systems and their \mathcal{L}_∞ error bounds. Int J Contr 1984; 39:1115–93.
100. Kwakernaak H. A polynomial approach to minimax frequency domain optimization of multivariable feedback systems. Int J Contr 1986; 41:117–56.
101. Kimura H. Chain-scattering approach to H_∞ -control. Boston: Birkhäuser; 1997.
102. Zhou K, Doyle J, Glover K. Robust and optimal control. Englewood Cliffs (NJ): Prentice Hall; 1996.
103. Başar T, Bernhard P. H_∞ optimal control and related minimax design problems: a dynamic game approach. 2nd edn Boston: Birkhäuser; 1995.
104. Willems JC. Almost invariant subspaces: an approach to high gain feedback design – part I: almost controlled invariant subspaces. IEEE Trans Automat Contr 1981; 26:235–52.
105. Willems JC. Almost invariant subspaces: an approach to high gain feedback design – part II: almost conditionally invariant subspaces. IEEE Trans Automat Contr 1982; 27:1071–85.
106. Liu K, Chen BM, Lin Z. On the problem of robust and perfect tracking for linear systems with external disturbances. Int J Contr 2001; 74:158–74.
107. Chen BM, Lin Z, Liu K. Robust and perfect tracking of discrete-time systems. Automatica 2002; 38:293–9.
108. Lewis FL. Applied optimal control and estimation. Englewood Cliffs (NJ): Prentice Hall; 1992.
109. Athans M. A tutorial on LQG/LTR methods. Proc American Contr Conf; Seattle, WA; 1986. p. 1289–96.

110. Chen BM. Theory of loop transfer recovery for multivariable linear systems [PhD diss]. Pullman (WA): Washington State University; 1991.
111. Chen BM, Saberi A, Sannuti P. A new stable compensator design for exact and approximate loop transfer recovery. *Automatica* 1991; 27:257–80.
112. Doyle JC, Stein G. Multivariable feedback design: concepts for a classical/modern synthesis. *IEEE Trans Automat Contr* 1981; 26:4–16.
113. Goodman GC. The LQG/LTR method and discrete-time control systems. Technical report. MIT (MA): Report No.: LIDS-TH-1392; 1984.
114. Kwakernaak H. Optimal low sensitivity linear feedback systems. *Automatica* 1969; 5:279–85.
115. Matson CL, Maybeck PS. On an assumed convergence result in the LQG/LTR technique. *Proc 26th IEEE Conf Dec Contr*; Los Angeles, CA; 1987. p. 951–2.
116. Niemann HH, Sogaard-Andersen P, Stoustrup J. Loop Transfer Recovery: Analysis and Design for General Observer Architecture. *Int J Contr* 1991; 53:1177–203.
117. Saberi A, Chen BM, Sannuti P. Loop transfer recovery: analysis and design. London: Springer; 1993.
118. Stein G, Athans M. The LQG/LTR procedure for multivariable feedback control design. *IEEE Tran Automat Contr* 1987; 32:105–14.
119. Zhang Z, Freudenberg JS. Loop transfer recovery for nonminimum phase plants. *IEEE Trans Automat Contr* 1990; 35:547–53.
120. Chen BM, Saberi A, Ly U. Closed loop transfer recovery with observer based controllers: analysis. in *Contr Dynam Syst* (ed. Leondes CT). San Diego (CA): Academic Press; 1992; 51:247–93.
121. Chen BM, Saberi A, Ly U. Closed loop transfer recovery with observer based controllers: design. in *Contr Dynam Syst* (ed. Leondes CT). San Diego (CA): Academic Press; 1992; 56:295–348.
122. Chen BM, Saberi A, Berg MC, Ly U. Closed loop transfer recovery for discrete time systems. in *Contr Dynam Syst* (ed. Leondes CT). San Diego (CA): Academic Press; 1993; 56:443–81.
123. Hu T, Lin Z. Control systems with actuator saturation: analysis and design. Boston: Birkhäuser; 2001.
124. Kirk DE. Optimal control theory. Englewood Cliffs (NJ): Prentice Hall; 1970.
125. Venkataramanan V, Chen BM, Lee TH, Guo G. A new approach to the design of mode switching control in hard disk drive servo systems. *Contr Eng Prac* 2002; 10:925–39.
126. Itkis U. Control systems of variable structure. New York (NY): Wiley; 1976.
127. Yamaguchi T, Numasato H, Hirai H. A mode-switching control for motion control and its application to disk drives: Design of optimal mode-switching conditions. *IEEE/ASME Trans Mechatron* 1998; 3:202–9.
128. Salle JL, Lefschetz S. Stability by Liapunov's direct method. New York (NY): Academic Press; 1961.
129. LaSalle J. Stability by Liapunov's direct method with applications. New York (NY): Academic Press; 1961.
130. Lin Z, Pachter M, Banda S. Toward improvement of tracking performance – nonlinear feedback for linear systems. *Int J Contr* 1998; 70:1–11.
131. Turner MC, Postlethwaite I, Walker DJ. Nonlinear tracking control for multivariable constrained input linear systems. *Int J Contr* 2000; 73:1160–72.
132. Chen BM, Lee TH, Peng K, Venkataramanan V. Composite nonlinear feedback control: theory and an application. *IEEE Trans Automat Contr* 2003; 48:427–39.

133. Venkataramanan V, Peng K, Chen BM, Lee TH. Discrete-time composite nonlinear feedback control with an application in design of a hard disk drive servo system. *IEEE Trans Contr Syst Technol* 2003; 11:16–23.
134. He Y, Chen BM, Wu C. Composite nonlinear control with state and measurement feedback for general multivariable systems with input saturation. *Syst Contr Lett* 2005; 54:455–69.
135. He Y, Chen BM, Wu C. Composite nonlinear feedback control for general discrete-time multivariable systems with actuator nonlinearities. *Proc 5th Asian Contr Conf; Melbourne, Australia; 2004. p. 539–44.*
136. Lan W, Chen BM, He Y. Improving transient performance in tracking control for a class of nonlinear systems with input saturation. *Syst Contr Lett* 2006; 55:132–8.
137. He Y, Chen BM, Lan W. Improving transient performance in tracking control for a class of nonlinear discrete-time systems with input saturation. *Proc 44th IEEE Conf Dec Contr; Seville, Spain; 2005. p. 8094–9.*
138. Peng K, Chen BM, Cheng G, Lee TH. Modeling and compensation of nonlinearities and friction in a micro hard disk drive servo system with nonlinear feedback control. *IEEE Trans Contr Syst Technol* 2005; 13:708–21.
139. Chen BM, Zheng D. Simultaneous finite and infinite zero assignments of linear systems. *Automatica* 1995; 31:643–8.
140. Cheng G, Peng K, Chen BM, Lee TH. A microdrive track following controller design using robust and perfect tracking control with nonlinear compensation. *Mechatronics* 2005; 15:933–48.
141. Iannou PA, Kosmatopoulos EB, Despain AM. Position error signal estimation at high sampling rates using data and servo sector measurements. *IEEE Trans Contr Syst Technol* 2003; 11:325–34.
142. Weerasooriya S, Low TS, Huang YH. Adaptive time optimal control of a disk drive actuator. *IEEE Trans Magn* 1994; 30:4224–6.
143. Xiong Y, Weerasooriya S, Low TS. Improved discrete proximate time optimal controller of a disk drive actuator. *IEEE Trans Magn* 1996; 32:4010–2.
144. Mizoshita Y, Hasegawa S, Takaishi K. Vibration minimized access control for disk drives. *IEEE Trans Magn* 1996; 32:1793–8.
145. Yamaguchi T, Nakagawa S. Recent control technologies for fast and precise servo system of hard disk drives. *Proc 6th Int Workshop Adv Motion Contr; Nagoya, Japan; 2000. p. 69–73.*
146. Tsuchiura KM, Tsukuba HH, Toride HO, Takahashi T. Disk system with sub-actuators for fine head displacement. *US Patent No: 5189578; 1993.*
147. Miu DK, Tai YC. Silicon micromachined SCALED technology. *IEEE Trans Ind Electron* 1995; 42:234–9.
148. Fan LS, Ottesen HH, Reiley TC, Wood RW. Magnetic recording head positioning at very high track densities using a microactuator based, two stage servo system. *IEEE Trans Ind Electron* 1995; 42:222–33.
149. Aggarwal SK, Horsley DA, Horowitz R, Pisano AP. Microactuators for high density disk drives. *Proc American Contr Conf; Albuquerque, NM; 1997. p. 3979–84.*
150. Ding, J, Tomizuka M, Numasato H. Design and robustness analysis of dual stage servo system. *Proc American Contr Conf; Chicago, IL; 2000. p. 2605–09.*
151. Evans RB, Griesbach JS, Messner WC. Piezoelectric microactuator for dual stage control. *IEEE Trans Magn* 1999; 35:977–82.
152. Fan LS, Hirano T, Hong J, Webb PR, Juan WH, Lee WY, *et al.* Electrostatic microactuator and design considerations for HDD application. *IEEE Trans Magn* 1999; 35:1000–5.

153. Guo L, Chang JK, Hu X. Track-following and seek/settle control schemes for high density disk drives with dual-stage actuators. Proc 2001 IEEE/ASME Int Conf Adv Intell Mechatron; Como, Italy; 2001. p. 1136–41.
154. Guo L, Martin D, Brunnett D. Dual-stage actuator servo control for high density disk drives. Proc 1999 IEEE/ASME Int Conf Adv Intell Mechatron; Atlanta, GA; 1999. p. 132–7.
155. Guo W, Weerasooriya S, Goh TB, Li QH, Bi C, Chang KT, *et al.* Dual stage actuators for high density rotating memory devices. IEEE Trans Magn 1998; 34:450–5.
156. Guo W, Yuan L, Wang L, Guo G, Huang T, Chen BM, *et al.* Linear quadratic optimal dual-stage servo control systems for hard disk drives. Proc 24th IEEE Ind Electron Soc Ann Conf; Aachen, Germany; 1998. p. 1405–10.
157. Hernandez D, Park SS, Horowitz R, Packard A. Dual-stage track-following design for hard disk drives. Proc American Contr Conf; San Diego, CA; 1999. p. 4116–21.
158. Horsley DA, Hernandez D, Horowitz R, Packard AK, Pisano AP. Closed-loop control of a microfabricated actuator for dual-stage hard disk drive servo systems. Proc American Contr Conf; Philadelphia, PA; 1998. p. 3028–32.
159. Hu X, Guo W, Huang T, Chen BM. Discrete time LQG/LTR dual-stage controller design and implementation for high track density HDDs. Proc American Contr Conf; San Diego, CA; 1999. p. 4111–5.
160. Kobayashi M, Horowitz R. Track seek control for hard disk dual-stage servo systems. IEEE Trans Magn 2001; 37:949–54.
161. Li Y, Horowitz R. Track-following controller design of MEMS based dual-stage servos in magnetic hard disk drives. Proc 2000 IEEE Int Conf Robot Automat; San Francisco, CA; 2000. p. 953–8.
162. Mori K, Munemoto T, Otsuki H, Yamaguchi Y, Akagi K. A dual-stage magnetic disk drive actuator using a piezoelectric device for a high track density. IEEE Trans Magn 1991; 27:5298–300.
163. Schroeck SJ, Messner WC. On controller design for linear time-invariant dual-input single-output systems. Proc American Contr Conf; San Diego, CA; 1999. p. 4122–6.
164. Semba T, Hirano T, Hong J, Fan LS. Dual-stage servo controller for HDD using MEMS microactuator. IEEE Trans Magn 1999; 35:2271–3.
165. Suthasan T, Mareels I, Mamun AA. System identification and control design for dual actuated disk drive. Contr Eng Prac 2002; 12:665–76.
166. Takaishi K, Imamura T, Mizasgita Y, Hasegawa S, Ueno T, Yamada T. Microactuator control for disk drive. IEEE Trans Magn 1996; 32:1863–6.
167. Du CL, Guo GX. Lowering the hump of sensitivity functions for discrete-time dual-stage systems. IEEE Trans Contr Syst Technol 2005; 13:791–7.
168. Canudas de Wit C, Lischinsky P. Adaptive friction compensation with partially known dynamic friction model. Int J Adapt Contr Signal Proc 1997; 11:65–80.
169. Canudas de Wit C, Olsson H, Aström KJ, Lischinsky P. A new model for control of systems with friction. IEEE Trans Automat Contr 1995; 40:419–25.
170. Canudas de Wit C, Olsson H, Aström KJ, Lischinsky P. Dynamic friction models and control design. Proc American Contr Conf; San Francisco, CA; 1993. p. 1920–6.
171. Olsson H, Aström KJ. Observer-based friction compensation. Proc 35th IEEE Conf Dec Contr; Kobe, Japan; 1996. p. 4345–50.
172. Dahl PR. Solid friction damping of mechanical vibrations. AIAA J 1976; 14:1675–82.
173. Ge SS, Lee TH, Ren SX. Adaptive friction compensation of servomechanisms. Int J Sys Sci 2001; 32:523–32.
174. Maria HA, Abrahams ID. Active control of friction-driven oscillations. J Sound Vibra 1996; 193:417–26.

175. Abramovitch D, Wang F, Franklin G. Disk drive pivot nonlinearity modeling – Part I: frequency domain. *Proc American Contr Conf*; Baltimore, MD; 2004. p. 2600–3.
176. Ishikawa J, Tomizuka M. Pivot friction compensation using an accelerometer and a disturbance observer for hard disk drives. *IEEE/ASME Trans Mechatron* 1998; 3:194–201.
177. Wang F, Abramovitch D, Franklin G. A method for verifying measurements and models of linear and nonlinear systems. *Proc American Contr Conf*; San Francisco, CA; 1993. p. 93–7.
178. Wang F, Hurst T, Abramovitch D, Franklin G. Disk drive pivot nonlinearity modeling – Part II: time domain. *Proc American Contr Conf*; Baltimore, MD; 1994. p. 2604–7.
179. Chang HS, Baek SE, Park JH, Byun YK. Modeling of pivot friction using relay function and estimation of its functional parameters. *Proc American Contr Conf*; San Francisco, CA; 1999. p. 3784–9.
180. Liu X, Liu JC. Analysis and measurement of torque hysteresis of pivot bearing in hard disk drive applications. *Tribology Int* 1999; 32:125–30.
181. Low TS, Guo W. Modeling of a three-layer piezoelectric bimorph beam with hysteresis. *J Microelectromech Syst* 1995; 4:230–7.
182. Chang TP. Seismic response analysis of nonlinear structures using the stochastic equivalent linearization technique [PhD diss]. New York (NY): Columbia University; 1985.
183. Peng K, Venkataraman V, Chen BM, Lee TH. Design and implementation of a dual-stage actuator HDD servo system via composite nonlinear feedback approach. *Mechatron* 2004; 14:965–88.
184. Hwang CL, Lin CH. A discrete-time multivariable neuro-adaptive control for nonlinear unknown dynamic systems. *IEEE Trans Syst Man Cyb B* 2000; 30:865–77.
185. Adriaens HJMTA, de Koning WL, Banning R. Modeling piezoelectric actuators. *IEEE/ASME Trans Mechatron* 2000; 5:331–41.
186. Cruz-Hernandez JM, Hayward V. Phase control approach to hysteresis reduction. *IEEE Trans Contr Syst Technol* 2001; 9:17–26.
187. Cheng HM, Ewe MTS, Chiu GTC, Bashir R. Modeling and control of piezoelectric cantilever beam micro-mirror and micro-laser arrays to reduce image banding in electrophotographic processes. *J Micromech Microeng* 2001; 11:487–98.
188. Guo G, Chen R, Low TS, Wang Y. Optimal control design for hard disk drive servosystems. *IEE Proc–Contr Theor Appl* 2002; 149:237–42.
189. Ewe MTS, Grice JM, Chiu GTC, Allebach JP, Chan CS, Foote W. Banding reduction in electrophotographic processes using a piezoelectric actuated laser beam deflection device. *J Imaging Sci Technol* 2002; 46:433–42.
190. Lin CL, Jan HY, Shieh NC. GA-based multiobjective PID control for a linear brushless DC motor. *IEEE/ASME Trans Mechatron* 2003; 8:56–65.
191. Hwang CL, Jan C. Optimal and reinforced robustness designs of fuzzy variable structure tracking control for a piezoelectric actuator system. *IEEE Trans Fuzzy Syst* 2003; 11:507–17.
192. Jan C, Hwang CL. A nonlinear observer-based sliding-mode control for piezoelectric actuator systems: Theory and experiments. *J Chinese Inst Engr* 2004; 27:9–22.
193. Huang YC, Lin DY. Ultra-fine tracking control on piezoelectric actuated motion stage using piezoelectric hysteretic model. *Asian J Contr* 2004; 6:208–16.
194. Hwang CL, Jan C. Nano trajectory control of multilayer low-voltage PZT render actuator systems. *Asian J Contr* 2004; 6:187–98.
195. Hwang CL, Chen YM. Discrete sliding-mode tracking control of high-displacement piezoelectric actuator systems. *J Dynam Syst–Trans ASME* 2004; 126:721–31.

196. Hwang CL, Chen YM, Jan C. Trajectory tracking of large-displacement piezoelectric actuators using a nonlinear observer-based variable structure control. *IEEE Trans Contr Syst Technol* 2005; 13:56–66.
197. Ikhouane F, Manosa V, Rodellar J. Adaptive control of a hysteretic structural system. *Automatica* 2005; 41:225–31.
198. Cruz-Hernandez JM, Hayward V. Position stability for phase control of the Preisach hysteresis model. *Trans Canadian Soc Mech Eng* 2005; 29:129–42.
199. Hwang CL, Jan C. State-estimator-based feedback control for a class of piezoelectric systems with hysteretic nonlinearity. *IEEE Trans Syst Man Cyb A* 2005; 35:654–64.
200. Ikhouane F, Rodellar J. On the hysteretic Bouc–Wen model. *Nonlinear Dynam* 2005; 42:63–78.
201. Moheimani SOR, Vautier BJB. Resonant control of structural vibration using charge-driven piezoelectric actuators. *IEEE Trans Contr Syst Technol* 2005; 13:1021–35.
202. Caughey TK. Derivation and application of the Fokker–Planck equation to discrete nonlinear dynamic systems subjected to white random excitation. *J Acoust Soc Am* 1963; 35:1683–92.
203. Crandall ST. Perturbation techniques for random vibration of nonlinear systems. *J Acoust Soc Am* 1963; 35:1700–05.
204. Lyon RH. Response of a nonlinear string to random excitation. *J Acoust Soc Am* 1960; 32:953–60.
205. Booton, Jr., RC. Nonlinear control systems with random inputs. *IRE Trans Circuit Theory* 1954; CT-1:9–18.

Index

- Almost disturbance decoupling, 68, 70, 74, 275
 - applications, 275
 - continuous-time, 70
 - discrete-time, 74
 - solvability conditions, 70, 74
- Bang-bang control, 98, 99
- Benchmark problem, 291
- Bilinear transformations
 - H_∞ control, 76
- Canonical forms of linear systems
 - special coordinate basis, 38
- CNF control toolkit, 164
- Complementary sensitivity functions, 48
 - two-degrees-of-freedom control, 49
- Composite nonlinear feedback control
 - continuous-time, 120
 - design parameter selection, 139, 158
 - discrete-time, 142
 - full-order output feedback, 125, 147
 - HDD servo systems, 205, 206, 225
 - interpretation, 139
 - Lyapunov functions, 123, 125, 127, 145, 147
 - microdrive servo systems, 258
 - nonlinear tuning function, 123
 - reduced-order output feedback, 130, 149
 - root locus, 139, 173
 - software toolkit, 164
 - state feedback, 121, 144
 - systems with disturbances, 132, 151
 - systems without disturbances, 121, 142
- Data flex cables, 245
- Digital signal processor, 17
- Disturbances, 11, 225
 - decoupling, 271
 - modeling, 13
 - rejection, 12
- Dual-stage actuators, 218
 - control configuration, 221
 - dynamical models, 220
 - frequency responses, 218
 - modeling, 218
 - physical configuration, 218
 - position error signal test, 239
 - runout disturbances, 225
 - sensitivity functions, 225
 - servo systems, 220
 - track following, 225
- Dynamic signal analyzer, 18
- Experimental setup, 17
- Finite zero structure of linear systems, 39, 43
- Friction
 - compensation, 257
 - model, 246
 - modeling, 245
- Gain margins, 48, 191, 209, 225, 259
- Geometric subspaces of linear systems, 45
 - \mathcal{S}_λ , 46
 - \mathcal{V}_λ , 46
 - strongly controllable subspaces, 45
 - weakly unobservable subspaces, 45

- H_2 control, 49
 - configuration, 50
 - continuous-time, 50
 - discrete-time, 59
 - full-order output feedback, 56, 64
 - optimal values, 52, 53, 61
 - perturbation approach, 53, 62
 - reduced-order output feedback, 57, 66
 - regular case, 52, 61, 62
 - Riccati equations, 53, 54, 61–63
 - singular case, 52, 53, 61, 62
 - state feedback, 54, 63
 - structural decomposition approach, 54, 56, 57, 63, 64, 66
- H_∞ control, 68
 - almost disturbance decoupling, 70, 74, 277
 - bilinear transformation, 76
 - configuration, 50
 - continuous-time, 69
 - discrete-time, 74
 - measurement feedback, 73
 - optimal values, 69, 74
 - perturbation approach, 70
 - regular case, 70
 - Riccati equations, 70
 - singular case, 70
 - state feedback, 71
 - structural decomposition approach, 71, 73
 - suboptimal controller, 70
- Hamiltonian, 97
- Hard disk drives
 - actuator assembly, 10
 - composite nonlinear feedback control, 205, 206, 225
 - data flex cable, 245
 - disturbance modeling, 13
 - disturbances, 11, 12
 - dual-stage actuated, 217
 - experimental setup, 17
 - first disk, 6
 - friction, 245
 - future trends, 8
 - historical development, 5, 6
 - mechanical structure, 3, 9
 - microdrive, 243
 - mode-switching control, 203, 206
 - modeling, 245
 - nonlinearities, 245
 - proximate time-optimal control, 202, 206
 - resonance compensation, 11
 - resonance modes, 180, 220, 255
 - robust and perfect tracking, 203
 - servo systems, 201, 217, 220, 255
 - single-stage actuated, 201
 - sources of errors, 12
 - spindle motor assembly, 10
 - suspension assembly, 10
 - track following, 3, 225
 - track misregistration, 11, 239
 - track seeking, 3, 206
 - track settling, 3
 - VCM actuators, 3, 201
- Hysteresis, 270
- Infinite zero structure of linear systems, 39, 44
- Invariant zeros of linear systems, 43
- Invertibility of linear systems, 44
 - degenerate, 44
 - invertible, 44
 - left invertible, 44
 - right invertible, 44
- Laser Doppler vibrometer, 18
- Least square estimation, 29
- Linear quadratic regulator
 - Riccati equations, 90
 - solutions, 90
- Linear systems toolkit, 40
- Loop transfer recovery, 88
 - achieved loop, 90
 - at input point, 88
 - at output point, 94
 - closed-loop recovery, 94
 - control configuration, 90
 - CSS architecture based, 92
 - duality, 94
 - full-order output feedback, 91
 - observer based, 91
 - recovery error, 90, 92, 93
 - reduced-order output feedback, 92
 - target loop, 89
- Lyapunov functions
 - composite nonlinear feedback control, 123, 125, 127, 145, 147
 - mode-switching control, 107, 109
 - proximate time-optimal control, 107

- Microactuators, 218, 269
 - control, 220
 - dual-stage actuator, 218
 - frequency responses, 218
 - modeling, 218
 - piezoelectric, 269
- Microdrives, 243
 - dynamic model, 249, 255
 - friction, 246
 - modeling, 245
 - nonlinearities, 249
 - resonance modes, 255
 - sensitivity functions, 259
- Mode-switching control, 104
 - configuration, 105
 - control law, 105
 - HDD servo systems, 203, 206
 - Lyapunov functions, 107, 109
 - stability analysis, 105
 - switching conditions, 109
- Modeling and identification, 21
 - confidence region, 28
 - dual-stage actuator, 220
 - impulse analysis, 22
 - least square method, 28
 - loss function, 27
 - microdrive, 245
 - model order, 27
 - model validation, 27, 33
 - Monte Carlo estimation, 32, 244, 249, 250
 - physical effect approach, 32
 - prediction error method, 26
 - step analysis, 24
 - VCM actuator, 180
- Monte Carlo estimation, 33, 244, 249, 250
- Normal rank of linear systems, 43
- Norms
 - l_p -norm, 77
 - H_2 -norm, 52, 60
 - H_∞ -norm, 69, 74
- Notch filters, 17, 182, 201, 258
- Phase margins, 48, 191, 209, 225, 259
- PID control, 47
 - configuration, 47
 - gain selection, 48
 - sensitivity functions, 48
 - Ziegler–Nichols tuning, 48
- Piezoelectric actuator system, 269
 - design formulation, 275
 - design specifications, 270
 - dynamical model, 269
 - hysteretic model, 270, 272
 - introduction, 269
 - simulations, 280
 - zero structures, 277
- Pontryagin's principle, 97
- Position error signal tests, 198, 239
 - dual-stage actuators, 239
 - dual-stage servo systems, 239
 - VCM actuators, 198
- Proximate time-optimal control, 101, 105
 - configurations, 101, 103
 - continuous-time, 101
 - control laws, 101, 104
 - control zones, 102
 - discrete-time, 103
 - HDD servo systems, 202, 206
 - Lyapunov functions, 107
 - sampling frequency, 104
- Relative degree of linear systems, 44
- Resonance modes
 - compensation, 11, 15
 - microactuator, 220
 - microdrive, 255
 - VCM actuator, 180
- Riccati equations
 - H_2 control, 53, 54, 61–63
 - H_∞ control, 70
 - linear quadratic regulator, 90
 - robust and perfect tracking, 80, 82
- Robust and perfect tracking, 76, 184
 - continuous systems, 76
 - continuous-time, 76
 - controller structures, 76, 85
 - discrete systems, 84
 - discrete-time, 84
 - full-order output feedback, 81, 83
 - hard disk drives, 203
 - measurement feedback, 86
 - perturbation approach, 81
 - Riccati equations, 80, 82
 - solvability conditions, 77, 85
 - state feedback, 78, 85
 - structural decomposition approach, 78, 81, 83, 85, 86

- Rosenbrock system matrix, 43
- Runout disturbances, 11, 191, 225
 - dual-stage actuators, 225
 - nonrepeatable runout, 14
 - repeatable runout, 13
 - VCM actuators, 191
- Sensitivity functions, 48, 191, 209, 225, 259
 - two-degrees-of-freedom control, 49
- Software toolkits, 17
 - CNF control, 17, 164
 - linear systems, 17, 40
- Special coordinate basis, 38, 78
 - block diagram, 42
 - compact form, 40
 - properties, 43–45
 - state-space decomposition, 45
 - transformations, 39
- Stability margins, 48
- Time-optimal control, 96, 163
 - closed-loop, 99
 - control scheme, 100
 - control signals, 97, 99
 - deceleration trajectories, 100
 - Hamiltonian, 97
 - minimum time, 99
 - open-loop, 98
 - optimal trajectories, 97
 - Pontryagin's principle, 97
- Track misregistration, 11, 239
 - dual-stage servo systems, 239
- Two-degrees-of-freedom control system, 49
- VCM actuators, 3, 179, 245
 - design specifications, 182, 258
 - driver, 246
 - dynamical models, 180, 181, 201, 220
 - frequency responses, 181, 201
 - implementation, 198, 259
 - microdrive, 243
 - modeling, 180, 245
 - position error signal tests, 198
 - runout disturbances, 191
 - sensitivity functions, 191, 259
 - servo systems, 201
 - track following, 188, 259
 - track seeking, 206
- Vibration-free table, 18
- Zero placement, 140, 159
- Ziegler–Nichols PID tuning, 47