Control Techniques for Increased Disturbance Rejection and Tracking Accuracy in Magnetic Disk Drives

by

Matthew Taylor White

B.S. (Michigan State University) 1990
M.S. (University of California, Berkeley) 1993

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy
in

Engineering–Mechanical Engineering

in the

GRADUATE DIVISION

of the

UNIVERSITY of CALIFORNIA, BERKELEY

Committee in charge:

Professor Masayoshi Tomizuka, Chair
Professor Roberto Horowitz
Professor Ronald S. Fearing

Spring 1997

The dissertation of Matthew Taylor White is approved:

Masayoshi Tomizuka May 16, 1997

Chair **Date**

Ronald S. Fearing May 16, 1997

Date

May 19, 1997

Date

University of California, Berkeley

Spring 1997

Control Techniques for Increased Disturbance Rejection and Tracking Accuracy in Magnetic Disk Drives

by

Matthew Taylor White

Abstract

Control Techniques for Increased Disturbance Rejection and Tracking Accuracy in Magnetic Disk Drives

by

Matthew Taylor White

Doctor of Philosophy in Engineering–Mechanical Engineering

University of California, Berkeley

Professor Masayoshi Tomizuka, Chair

As the density of data on magnetic disk drives increases, so does the need for more precise position control of the read/write head, especially in the presence of internal and external disturbances. This is achieved by two separate algorithms, acceleration feedforward control and a disturbance observer.

In the first algorithm, the acceleration of the drive is measured and fed forward to the actuator. By matching the electromechanical impedance between the disturbance and the position error, the feedforward controller can cancel the effects of the disturbance. Two techniques are presented for designing the feedforward controller. The first technique is an infinite impulse response filter that is designed off-line, and the second is a finite impulse response filter that is adapted on-line using the filtered-x LMS algorithm. Both feedforward techniques were tested through shake table experiments, resulting in reductions of the position error signal between 50% and 95% for vibration disturbances in the frequency range from 10 Hz to 500 Hz. Simulation of the system response to a shock disturbance resulted in a 89% reduction of the peak error.

The second algorithm does not require an external sensor. This is particularly relevant in products such as disk drives where cost is a major concern. The disturbance observer uses the position error signal and a nominal model of the plant to create an estimate of the disturbance. This estimate is then used to compensate for the disturbance effects. Vibration simulation indicated that the performance of the disturbance observer was comparable to the feedforward control in the 50 Hz to 500 Hz range, but offered significantly better results below 50 Hz. Experimental results verified that the disturbance observer was effective below 200 kHz, but noise in the system degraded performance between 200 Hz and 1 kHz. Simulation of shock testing with the disturbance observer was similar to the feedforward control results.

Masayoshi Tomizuka

Professor Masayoshi Tomizuka
Dissertation Committee Chair

Contents

List of Figures

List of Tables

List of Symbols

Acronyms

ADC	Analog-to-Digital Converter
ARMA	Auto-Regressive Moving-Average
DAC	Digital-to-Analog Converter
DSA	Dynamic Signal Analyzer
DSP	Digital Signal Processor
FEM	Finite Element Model
FIR	Finite Impulse Response
FPC	Flexible Printed Circuit
FSLQ	Frequency-Shaped Linear Quadratic
IIR	Infinite Impulse Response
LMS	Least-Mean-Square
LPF	Low-Pass Filter
LQG	Linear Quadratic Gaussian
LTR	Loop Transfer Recovery
NRRO	Non-Repeatable Run-Out
ODE	Ordinary Differential Equation
PAA	Parameter Adaptation Algorithm

PES	Position Error Signal
PID	Proportional-Integral-Derivative
PSD	Power Spectral Density
RAID	Redundant Array of Independent Disks
RMS	Root-Mean-Square
RRO	Repeatable Run-Out
SPR	Strictly Positive Real
VCM	Voice Coil Motor
ZPET	Zero Phase Error Tracking

Roman Symbols

$\mathbf{0}_{m\times n}$	Zero matrix with dimensions $m \times n$
$\mathbf{A}$	State matrix for the ODE technique
$\mathbf{A}_{\mathbf{d}}$	Submatrix of $\mathbf{A}$ associated with $\mathbf{d}$
$\hat{\mathbf{A}}_{\mathbf{d}}$	Submatrix of $\mathbf{A}$ associated with $\hat{\mathbf{d}}$
$\hat{\mathbf{A}}_{\mathbf{h}}$	Submatrix of $\mathbf{A}$ associated with $\hat{\mathbf{h}}$
$\mathbf{A}_{\mathbf{y}}$	General form of $\mathbf{A}$ submatrix associated with $\mathbf{y}$
A	Denominator of the nominal plant model
A_1	Magnitude of first acceleration signal
A_2	Magnitude of second acceleration signal
a_i	Nominal plant model denominator coefficients
B	Numerator of the nominal plant model
B^-	Uncancelable part of the nominal plant model numerator
B^+	Cancelable part of the nominal plant model numerator
$\mathbf{b}$	Input vector for the ODE technique

$\mathbf{b}_{\mathbf{d}}$	Subvector of $\mathbf{b}$ associated with $\mathbf{d}$
$\hat{\mathbf{b}}_{\mathbf{d}}$	Subvector of $\mathbf{b}$ associated with $\hat{\mathbf{d}}$
$\hat{\mathbf{b}}_{\mathbf{h}}$	Subvector of $\mathbf{b}$ associated with $\hat{\mathbf{h}}$
$\mathbf{b}_{\mathbf{y}}$	General form of $\mathbf{b}$ subvector associated with $\mathbf{y}$
b_h	Damping for the actuator pivot rotation mode
b_i	Nominal plant model numerator coefficients
b_r	Damping for the actuator pivot bending mode
b_s	Damping for the actuator arm bending mode
b_{sm}	Damping of the shock mounts
b_θ	Rotational damping of the actuator pivot
$\bar{\mathcal{D}}$	Compact subset of D_R
$\mathcal{D}_A$	Domain of attraction of the associated ODE
$\mathcal{D}_C$	Invariant set of the associated ODE
$\mathcal{D}_R$	Open, connected subset of D_S
$\bar{\mathcal{D}}_R$	Closed subset of D_R
$\mathcal{D}_S$	Set of $\mathbf{w}$ such that $\mathbf{A}$ is stable
D	General transfer function denominator
D_{out}	Denominator of G_{out}
$\hat{D}_{out}$	Denominator of $\hat{G}_{out}$
D_{ref}	Denominator of G_{ref}
$\mathbf{d}$	Vector of d values
$\hat{\mathbf{d}}$	Vector of $\hat{d}$ values
d	Desired output signal for adaptation
$\hat{d}$	Actual output signal for adaptation
d_i	General denominator coefficients
E	Expectation operator

$\bar{E}$	Expectation limit operator
$\mathbf{e}_m$	General form of $\mathbf{b}$ subvector with dimension m
e	General adaptive algorithm error
F_c	Force at the motor coil
F_{dist}	Disturbance force at the motor coil
F_h	Force at the read/write head
f	Function for the ODE technique
G	General transfer function
G_{bt}	Transfer function from the base acceleration to the track position
G_c	Feedback controller transfer function
G_{Fb}	Transfer function from the motor force to the base acceleration
G_{Fh}	Transfer function from the motor force to the read/write head position
G_{ff}	Feedforward controller transfer function
$G_{ff,nom}$	Nominal feedforward controller used for calculating G_{ref}
G_h	Transfer function of the actuator pivot rotational mode
G_n	Notch filter transfer functions
G_{out}	Output dynamics for the acceleration feedforward controller
$\hat{G}_{out}$	Estimate of the output dynamics
G_p	Actual plant for the disturbance observer
G_{pn}	Nominal plant model for the disturbance observer
G_r	Transfer function of the actuator pivot bending mode
G_{ref}	Model reference transfer function
G_s	Transfer function of the actuator arm bending mode
$G_{u\varepsilon}$	Transfer function from the control signal to the position error signal
G_w	Transfer function from the base acceleration to the position error signal with feedforward control

G_{wo}	Transfer function from the base acceleration to the position error signal without feedforward control
$G_{\delta\varepsilon}$	Transfer function from the input disturbance to the position error signal
$G_{\xi\varepsilon}$	Transfer function from the sensor noise to the position error signal
g	Output of the nominal plant model inverse
H	Ratio of the estimated output dynamics to the actual output dynamics
$\mathbf{h}$	Vector of h values
$\hat{\mathbf{h}}$	Vector of $\hat{h}$ values
h	Output of G_{out} for input x
$\hat{h}$	Output of $\hat{G}_{out}$ for input x
h_i	Parameters of H
$\mathbf{I}$	Identity matrix
$\hat{\mathbf{J}}$	Jordan form of $\hat{\mathbf{R}}$
J	Mass moment of inertia of the actuator arm about the actuator pivot
J_d	Mass moment of inertia of the disk drive about the actuator pivot
J_g	Mass moment of inertia of the actuator arm about the mass center
j	Square root of -1
K_a	Transconductance amplifier gain
K_F	Motor force constant
K_{pes}	Position error signal gain
k_h	Stiffness for the actuator pivot rotation mode
k_r	Stiffness for the actuator pivot bending mode
k_s	Stiffness for the actuator arm bending mode
k_{sm}	Stiffness of the shock mounts
k_θ	Rotational stiffness of the actuator pivot
$\mathbf{L}_{m\times n}$	General form of $\mathbf{A}$ submatrix with dimensions $m \times n$

$L + 1$	Number of feedforward controller taps
l_c	Distance from the motor coil to the actuator pivot
l_{gy}	Distance from the actuator pivot to the actuator arm mass center in the y-direction
l_h	Distance from the actuator pivot to the read/write head
M	General adaptive algorithm gain
m	Numerator order
m_d	Order of N_d
$\hat{m}_d$	Order of $\hat{N}_d$
m_g	Mass for the actuator pivot bending mode
m_h	Mass for the actuator pivot rotation mode
$\hat{m}_h$	Order of $\hat{N}_h$
m_{max}	Maximum of m_d, $\hat{m}_d$, and $\hat{m}_h$
m_s	Mass for the actuator arm bending mode
$\mathbf{N_d}$	Submatrix of $\mathbf{A}$ associated with $\mathbf{d}$
$\hat{\mathbf{N}}_\mathbf{d}$	Submatrix of $\mathbf{A}$ associated with $\hat{\mathbf{d}}$
$\hat{\mathbf{N}}_\mathbf{h}$	Submatrix of $\mathbf{A}$ associated with $\hat{\mathbf{h}}$
$\mathbf{N_y}$	General form of $\mathbf{A}$ submatrix associated with $\mathbf{y}$
$N + 1$	Length of the regression vector (Chapter 3)
N	General transfer function numerator (Appendix B)
N_{out}	Numerator of G_{out}
$\hat{N}_{out}$	Numerator of $\hat{G}_{out}$
N_{ref}	Numerator of G_{ref}
n	Denominator order
n_d	Order of D_d
$\hat{n}_d$	Order of $\hat{D}_d$

$\hat{n}_h$	Order of $\hat{D}_h$
n_i	General numerator coefficients
$\mathbf{p}$	Correlation vector of $\mathbf{h}$ and d
$\hat{\mathbf{p}}$	Correlation vector of $\hat{\mathbf{h}}$ and d
$\hat{\mathbf{Q}}$	Similarity transformation matrix for $\hat{\mathbf{R}}$
Q	Low-pass filter for the disturbance observer
q^{-1}	One-step delay operator
$\mathcal{R}^{L+1}$	Set of real vectors with dimension $L+1$
$\mathbf{R}$	Correlation matrix of $\mathbf{h}$
$\hat{\mathbf{R}}$	Correlation matrix of $\hat{\mathbf{h}}$ and $\mathbf{h}$
S	Sensitivity transfer function
s	Laplace transform variable
T	Complementary sensitivity transfer function
T_s	Sampling period
u	Output of the feedback controller
V	Lyapunov function
$\mathbf{v}$	Vector of translated feedforward controller tap weights
$\mathbf{v}'$	Vector of rotated feedforward controller tap weights
v	Model reference output
$\hat{v}$	Output of the feedforward controller
$\mathbf{w}$	Vector of feedforward controller tap weights
$\mathbf{w}^*$	Optimal feedforward controller tap weights
$\hat{\mathbf{w}}$	Converged feedforward controller tap weights
$\bar{\mathbf{w}}$	Constant value for feedforward controller tap weights
$\mathbf{w_D}$	Vector of feedforward tap weights in the time scale τ
$\mathbf{\Delta w}$	Change in feedforward controller tap weights (Section B.5.4)

$\Delta\mathbf{w}$	Difference between the converged and optimal feedforward controller tap weights (Section B.6)
w_i	Tap weights for the feedforward controller
$\mathbf{x}$	Vector of x values
$\mathbf{x}_m$	Vector of x values with dimension m
$\mathbf{x}_{m_d}$	Vector of x values with dimension m_d
$\mathbf{x}_{\hat{m}_d}$	Vector of x values with dimension $\hat{m}_d$
$\mathbf{x}_{\hat{m}_h}$	Vector of x values with dimension $\hat{m}_h$
x	Input to the feedforward filter
x_g	Position of the actuator arm mass center
x_h	Position of the read/write head
x_t	Track position
$\mathbf{y}$	Vector of y values
y	General output value
$\mathbf{z}$	General real vector
z	Z-transform variable
z^{-d}	Total delay for the disturbance observer
$z^{-\sigma}$	Delay due to uncancelable zeros of the nominal plant model
$z^{-\rho}$	Pure delay in the plant (including computations)

Greek Symbols

β_r	Coefficient for adjustment of the pivot bending model
γ	Diminishing adaptation gain
δ	Disturbance input for the disturbance observer formulation
$\hat{\delta}$	Estimate of the disturbance input from the disturbance observer

ε	Position error signal
ζ	Damping ratio for the notch filters
ζ_{sm}	Damping ratio of the shock mounts
θ	Actuator arm rotation
θ	Accelerometer rotation (Appendix A.1)
$\hat{\theta}$	General adaptive algorithm parameters
$\ddot{\theta}_{base}$	Base rotational acceleration
λ_i	Eigenvalues of $\hat{\mathbf{R}}$ and $\hat{\mathbf{J}}$
λ_{max}	Maximum eigenvalue of $\hat{\mathbf{R}}$ and $\hat{\mathbf{J}}$
μ	Adaptation gain for the filtered-x LMS algorithm
ξ	Noise (Chapters 3 and 5)
ξ	Mean-square error (Appendix B)
ξ_{min}	Minimum mean-square error
τ	Time scale for the ODE technique
τ_k	Time scale for the ODE technique associated with time k
$\mathbf{\Delta}\tau$	Change in the ODE time scale
$\Phi_{\phi\phi}$	Spectral density of ϕ
ϕ	General adaptive algorithm regression vector (Chapter 3)
ϕ	Inner product of $\mathbf{h}$ and $\mathbf{z}$ (Appendix B)
$\tilde{\phi}$	Inner product of $\mathbf{z}$ and $H(q^{-1})\mathbf{h}$
χ	Extended state vector for the ODE technique
ω	General frequency
ω_1	Natural frequency for the 2.1 kHz notch filter
ω_2	Natural frequency for the 4.1 kHz notch filter
ω_s	Natural frequency for the 20 kHz notch filter

Acknowledgments

First, I would like to thank my parents, Bill and Trish, and my sisters, Sara and Bekah, for all of their patience and support while I was thousands of miles away from home working on my degree.

In addition, I am grateful to Prof. Tomizuka for giving my project a clear direction and application. Despite an incredible workload, he always manages to find time for his students. His keen insight identified problems in my work on many occasions. My other committee members, Prof. Horowitz and Prof. Fearing, also deserve my appreciation for their time, comments, and suggestions while reviewing my dissertation.

To Mr. Yamaguchi and Mr. Shishida from Hitachi, Ltd., I would like to express my gratitude for their hospitality during my trip to Japan, and for taking time from their busy schedules to answer questions about the equipment that Hitachi graciously provided for this research. Similarly, I would like to thank Prof. Bogy and the Computer Mechanics Laboratory for the funding that I needed to complete this project, and Texas Instruments and PCB for their equipment donations.

Prof. Tomizuka has mentioned on several occasions that a graduate student should learn at least as much from his fellow students as he does from his advisor. I can definitely say that was the case with me. The Tomizuka Group is an amazing collection of personalities and intellects. Unfortunately, it is also too large to mention everyone by name, but they have all been of help to me at one time or another. Thanks to Carl, without whom I never could have gotten started on this work, to Lin for trading friendly barbs with me when I had been sitting at my desk too long, to Craig for his machining advice and endless supply of ideas, to Carlos and Victor for their Herculean efforts at keeping the computers running, and to Pushkar, Sujit, Chen, Mohammed, Weiguang, Erwin, *et al.* for making the office a fun place to work. Still more thanks to Sat and Dave for all of their help and suggestions on my experiments, and to Scott, Jim, and the rest of the gang for providing alternatives

to working on those experiments.

Special thanks to Kathy for her understanding during the dissertation-writing process, and for her help with proofreading the manuscript.

Finally to Tony and Mark, for taking the time to listen to me when I had a stupid question, and for always being around when it was time to take a break, even if all too often we just ended up talking about disk drives (or windsurfing).

Chapter 1

Introduction

1.1 General Problem Statement: Disturbance Effects in Disk Drive Systems

Despite amazing performance improvements in the recent past, the disk drive industry continues to search for ways to make its products smaller, faster, and able to hold more data. Commercial software is becoming more sophisticated, and therefore requires more disk space. Multimedia applications are also storage intensive.

Key elements in the search are increasing radial density, or the number of concentric data tracks stored on the disk surface, and decreasing access time, the time it takes to move from the current track to the desired track. Among other things, this requires more accurate positioning of the read/write head, often in the presence of significant disturbances. These disturbances can be in the form of shock or vibration while the drive is seeking or following a track.

Disk drives are being used in increasingly demanding environments. For example, laptop and hand-held computers are becoming even more popular. New applications such as cellular phones, digital video cameras, automobile navigation systems, hand-held medical devices, and palm-top organizers are being investigated [45, 34]. With the advancement of automation in industry, even traditional desktop computers are subjected to the rigors of the manufacturing floor. All of these situations elevate the level of external disturbances to which drives are subjected.

To compound this problem, many of the smaller form factors have abandoned the use of shock mounts that help to isolate the drives from such disturbances. Those drives that have retained the shock mounts are more susceptible to wind-up, the internal reaction force due to actuator motion, which is especially significant during seeks. To decrease seek time, drive engineers are pushing closer to time-optimal, or bang-bang, control. This type of control results in large reaction forces and excites modes of the drive dynamics that have been neglected in the past.

1.2 Previous Work

Pressure from consumers and other branches of the computer industry have fueled interest in disk drive research. Although the servo control problem is just one of many parts of the system, it is a critical component for increasing the storage capacity and access speed of the drives. This has made disk drives a very popular application for feedback control techniques. One of the major advantages of closed loop control is increased disturbance rejection, and a wide variety of such methods have been employed on magnetic disk drive systems. A comprehensive, although slightly out of date, description of the disk drive servo problem can be found in [18].

1.2.1 Disk Drive Servo Control

Some examples of the techniques employed on magnetic disk drives are fuzzy logic [8], linear quadratic Gaussian/loop transfer recovery (LQG/LTR), linear programming, and time optimal control [66, 22, 60], and H_∞ design [24]. An interesting comparison of Frequency-Shaped Linear Quadratic (FSLQ) control, LQG/LTR, and H_∞ control applied to disk drive servo control can be found in [33]. Additional control techniques such as repetitive and adaptive control will be discussed below.

The typical structure of a disk drive servo controller is actually composed of two controllers. One is optimized for track seeking, and the second for track following. The disk drive switches between these two controllers to perform these two tasks. Sometimes a third controller is implemented for settling, the transition from seeking to following. A technique that takes the switching nature of the disk drive controller explicitly into account is the initial value compensation method of Yamaguchi *et al.* [73].

The control signals are calculated from the position error signal (PES), which comes from position information recorded on the disks. A discussion of the PES generation can be found in [74]. In general, the PES is the only signal that the controller is able to access. This position information can either be recorded on a single disk that holds only position information (in a technique known as dedicated servo), or included in parts of each disk (in what is called embedded or sector servo).

From a control standpoint, dedicated servo is preferable because position information is available whenever it is desired. However, it is very costly to have a disk that is dedicated only to position information, and there are some problems that stem from the fact that the position measurement is not coming from the same disk as the data. Because of this, many drive manufacturers are adopting the sector servo method. This can create problems for the control engineer because the PES is not always available when it is needed. A technique to combat this problem called multi-rate control has been applied to disk drive systems by Chiang [12] and Phan [50].

A variety of error sources exist in the disk drive system. These are generally classified into repeatable run-out (RRO) and non-repeatable run-out (NRRO). RRO is synchronized with the rotation of the disks, and is caused by such factors as track eccentricity, bearing geometry and wear, motor geometry, disk warpage and slippage, and spindle tilt. NRRO is much more general. It can also be the result of bearing and spindle effects, or it can be due to external disturbance sources and internal sources like flexble printed circuits (FPC) and reaction forces. A popular technique for combatting RRO is the repetitive controller [35, 11, 42, 55].

One field that has become very popular in disk drive research is the use of microactuators and microsensors. Although these have not yet made it into production drives, they appear to be very promising, especially for meeting the requirements that are expected in the next decade. Currently, disk drive servo systems have a single actuator located at the base of the arm. It has been suggested that a second, microactuator could be placed at the tip of the arm to provide fine adjustments of the read/write head that would complement the course adjustments of the standard actuator [10, 44, 29]. Microsensors may also be used to provide more measurements of the drive behavior, similar to what will be shown in later sections for the acceleration feedforward technique.

1.2.2 Acceleration Feedforward Control

One of the techniques for disturbance rejection to be presented uses accelerometers to measure the motion of the drive, and then feeds this information forward to the actuator controller to coordinate the read/write head position with the desired track position.

A number of authors have suggested the use of accelerometers for disturbance rejection in the past. Aruga *et al.* [4] employed them on a dual, linear actuator system for a 10.5 inch form factor drive. They were concerned with the effects on the track following of one head as the second head performed a seek. Some dynamics of the drive were included, and robustness to modeling parameters was considered analytically. An experimental seek result was shown.

Davies and Sidman [13] developed a constant, single-parameter, acceleration feedforward controller with low-pass filtering on the accelerometer output. Again, an analysis of the robustness to parameter variations was presented, with experimental results for a seek test and an external impulse disturbance.

Kempf [33] used an accelerometer on a compact disk player to control the focus length of the reading lens. He considered the first bending mode of the disk with simple models for the motor and shock mounts. The filtered-x LMS (least-mean-squares) adaptation algorithm was applied to the controller parameters.

Although drive companies have included low-grade accelerometers as shock sensors for shutdown, until recently the cost of quality-signal accelerometers has prohibited their use for control. However, advancements in sensor technology have lowered prices to a level that makes including these accelerometers in a production drive feasible. Examples of such sensors were used in the experiments. Furthermore, the field of microsensors may soon provide low-cost accelerometers that, due to their small size, can be used in more sophisticated schemes.

1.2.3 Adaptive Control

Adaptive control is itself a very large and active area of research. A comprehensive discussion of this field of adaptive control is beyond the scope of this dissertation, but Åström and Wittenmark [2] and Goodwin and Sin [21] are excellent sources for a more thorough treatment of the subject. Typically, adaptive control is useful when there are unknown variations in the system dynamics or the character of the disturbances. However, in some cases it is more efficient to add intelligence to the controller in the form of adaptive control even when the operating conditions are reasonably well known. Adaptive controllers have been used in disk drive systems to reduce the effects of both RRO [11, 59] and NRRO [28, 27]. Adaptation will be used in conjunction with the acceleration feedforward technique using the filtered-x LMS algorithm described in [72].

1.2.4 Disturbance Observers

One of the drawbacks of the acceleration feedforward control system is that it requires the use of sensors that are not normally included in disk drives. Although the price of accelerometers is falling, and the sensors that were used in the experiments to be presented were extremely inexpensive, it is obviously cheaper to employ a method that does not require additional parts. One such method is called a disturbance observer. As the name suggests, the disturbance observer creates an estimate of the disturbance and uses this model to compensate for the effects of the disturbance. The disturbance observer was introduced by Ohnishi [47] and has been successfully applied to a wide range of systems such as dc servomotors [68], machine tools [36, 32], and robotics [6, 26]. A more complete description of the disturbance observer can be found in Chapter 5.

1.3 New Contributions

Although both acceleration feedforward control and the disturbance observer technique have been presented in previous work by other investigators, the results contained in this dissertation are unique in a variety of ways.

The acceleration feedforward control results presented here are the first to use low-cost accelerometers that employ sensor technology that has only recently been developed. Previous researchers used very expensive, high-quality sensors that were not feasible for actual implementation in a disk drive.

The work in this dissertation includes a more thorough model of the transfer function between the disturbance and the position of the read/write head, which led to increased performance of the control algorithms. The use of adaptation on the feedforward controller parameters has not been performed previously on a magnetic disk drive. Concurrent work in this area has been performed by Abramovitch [3], but an investigation of the controller parameters and the convergence properties of the adaptation algorithm was neglected.

This is also the first application of the disturbance observer to a hard disk drive, although Smith and Tomizuka [62] have performed concurrent investigations on removeable disk drives with shock disturbances. Finally, perhaps the most useful contribution of this research is that it includes a wide array of test results that allow a comparison of the feedforward control and the disturbance observer performance.

1.4 Summary of Chapters

Chapter 2 begins with a discussion of the disk drive system and introduces some standard assumptions used to create a model of the drive for use in the simulations. A short description of the test shaker and the experimental equipment is provided, although many of the details are reserved for Appendix A. The nominal feedback controller is also given, and the open loop system is used to verify the plant model structure and parameters.

Chapter 3 contains the development of the acceleration feedforward controllers. Two versions are presented. The first version uses a fixed-parameter, infinite impulse repsonse (IIR) filter. The second method is an adaptive, finite impulse response (FIR) filter. Details of the adaptive algorithm, the filtered-x LMS method, are given.

Chapter 4 presents the results of the two acceleration feedforward techniques for vibration and shock disturbances. All tests are performed in track-following mode. Results are given in both the time and frequency domains for simulations and experiments.

Chapter 5 provides an explanation of the disturbance observer method. This technique is then applied to the disk drive servo problem. Chapter 6 follows with the results of the disturbance observer simulations and experiments under the same conditions that were used to analyze the performance of the acceleration feedforward controller.

Chapter 7 gives a review of the major results. A comparison between acceleration feedforward control and the disturbance observer and the results achieved with these two methods is made, along with some conclusions.

As mentioned previously, Appendix A provides details on the experimental equipment, including information on the accelerometers used. Appendix B gives a convergence proof for the filtered-x LMS algorithm using the ODE (Ordinary Differential Equation) method for a diminishing adaptation gain. A second proof is provided for the case of constant adaptation gain.

Chapter 2

Disk Drive System

The disk drive used for experimentation was a 5.25 inch form factor with 2000 tracks per inch and a disk spindle speed of 3600 rpm. (See Figure 2.1 for a schematic of the drive.) The drive contained a stack of 8 disks. One disk side was dedicated to servo information to generate the PES between the read/write head and the desired track. The rotary actuator was composed of a voice coil motor (VCM) and the bearing-supported actuator arm with the read/write head and its suspension system. An FPC connected the actuator arm with the main disk drive electronics board located on the bottom of the drive.

The drive base was attached to a mounting frame with shock mounts. The drive frame was mounted on a rotary shaker system that provided the external disturbance. The drive was situated so that it rotated about the actuator axis. This configuration transmitted the maximum disturbance to the read/write head. As shown (exaggerated) in Figure 2.2, the disturbance to the base caused a relative displacement between the data track and the position of the head. It was this displacement that acceleration feedforward control was designed to minimize.

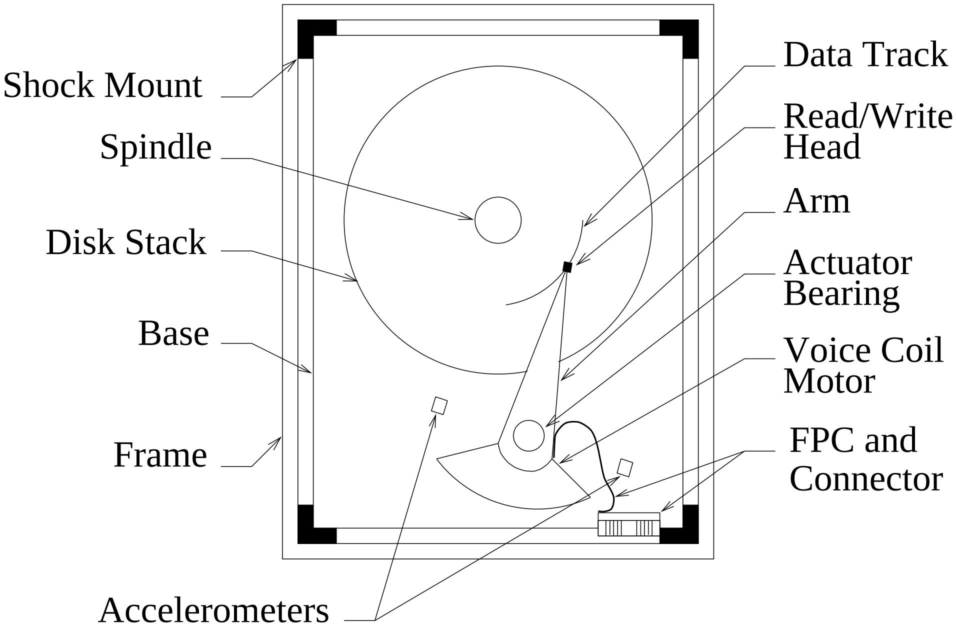


Figure 2.1: Disk Drive Components

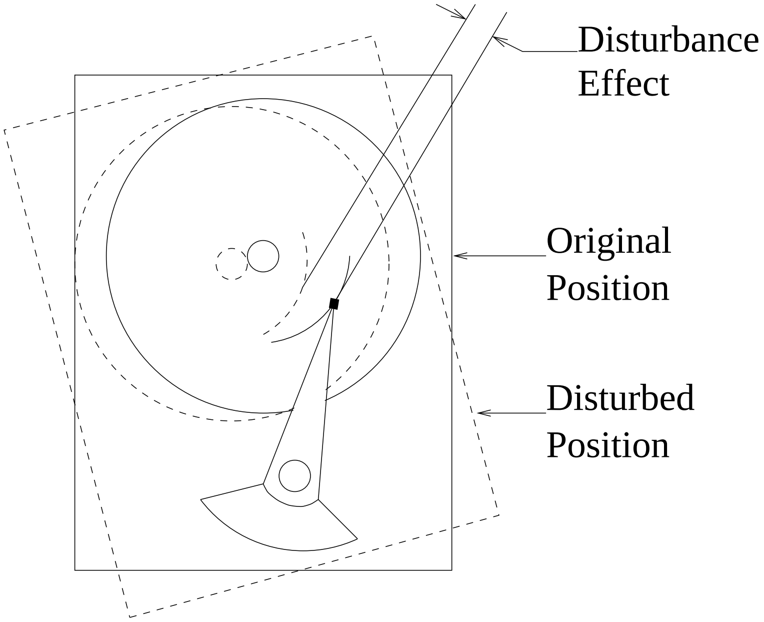


Figure 2.2: Effect of Disturbance on Track Misregistration

Variable	Explanation	Value
F_c	Actuator force at the coil	N
x_h	Linear displacement of the read/write head	m
x_g	Linear displacement of the mass center	m
θ	Rotation of the actuator	rad
l_h	Head moment arm from rotation center	70×10^{-3} m
l_c	Coil moment arm from rotation center	36×10^{-3} m
l_{gy}	y-distance from rotation center to mass center	0.096×10^{-3} m
J	Moment of inertia about the rotation center	0.593×10^{-6} kg m^2
b_θ	Rotational damping of the pivot	5.9×10^{-3} N s m
k_θ	Rotational stiffness of the pivot	5.29 N m
m_h	Mass for pivot rotation	0.0121 kg
b_h	Average damping coefficient for pivot rotation	1.07 N s/m
k_h	Average stiffness coefficient for pivot rotation	216 N/m
m_g	Actuator mass	116.3×10^{-3} kg
J_g	Moment of inertia about mass center	59.194×10^{-6} kg m^2
β_r	Experimental fit coefficient	7.0
b_r	Damping coefficient for pivot bending	36 N s/m
k_r	Stiffness coefficient for pivot bending	20×10^6 N/m
m_s	Mass for arm bending	25×10^{-6} kg
b_s	Damping coefficient for arm bending	7.0×10^{-3} N s/m
k_s	Stiffness coefficient for arm bending	17×10^3 N/m

Table 2.1: Explanation of Disk Drive Model Parameters

2.1 Disk Drive Actuator Model

Modeling information was obtained from the disk drive manufacturer and verified experimentally. Three modes were considered for the model of the actuator. Table 2.1 lists the parameter variables, their meanings, and their values for the three actuator modes.

The first mode included the damping and stiffness of the rotational motion of the actuator pivot bearing and the FPC, as shown in Figure 2.3. A torque balance about the center of rotation (the pivot bearing) gives

$$J\ddot{\theta} = F_c l_c - b_\theta \dot{\theta} - k_\theta \theta. \tag{2.1}$$

Using the small angle approximation $\sin\theta \approx \theta$ and substituting for

$$\begin{aligned} J &= m_h l_h^2 && (2.2) \\ \theta &= x_h / l_h && (2.3) \\ b_\theta &= b_h l_h^2 && (2.4) \\ k_\theta &= k_h l_h^2 && (2.5) \end{aligned}$$

results in

$$m_h l_h \ddot{x}_h = F_c l_c - b_h l_h \dot{x}_h - k_h l_h x_h. \tag{2.6}$$

Taking the Laplace transform gives

$$(m_h l_h s^2 + b_h l_h s + k_h l_h) x_h = F_c l_c \tag{2.7}$$

and therefore the transfer function from the force at the coil F_c to the displacement of the head x_h is

$$G_h(s) = \frac{x_h}{F_c} = \frac{l_c/l_h}{m_h s^2 + b_h s + k_h}. \tag{2.8}$$

The damping and stiffness are nonlinear, and are functions of the input magnitude. Figure 2.4 shows the outputs for a range of inputs. Constant values were used for the linear, time-invariant model. To determine the range of values, an experimental frequency response was calculated using a frequency sweep from low frequency to high frequency for a given input voltage to the VCM. A second sweep was performed from high frequency to low frequency. This process was repeated for a range of input voltages. For each frequency response, an effective damping coefficient and stiffness force were calculated. These results were used to determine average damping and stiffness values for all amplitudes of VCM motion. Results are shown in Figure 2.5 and Figure 2.6 for the damping coefficents and stiffness forces, respectively. The force at the head F_h displayed on the ordinate of Figure 2.6 is given by

$$F_h = (l_c/l_h) F_c \tag{2.9}$$

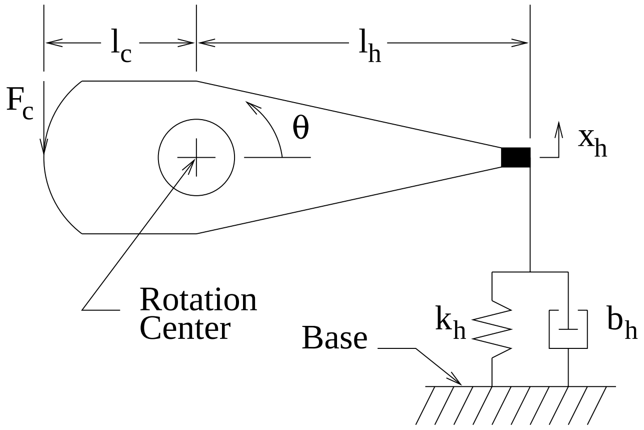


Figure 2.3: Schematic of the Actuator Pivot Bearing Rotational Mode

and then the stiffness k_h is calculated using

$$k_h = F_h / x_h. \tag{2.10}$$

The second mode resulted from the bending of the actuator pivot and its supporting structure, as shown in Figure 2.7. The natural frequency of this mode was much higher, at approximately 2.1 kHz. Consider the force balance in the x-direction

$$m_g \ddot{x}_g = F_c - b_r(\dot{x}_g + l_{gy}\dot{\theta}) - k_r(x_g + l_{gy}\theta) \tag{2.11}$$

and the torque balance about the mass center

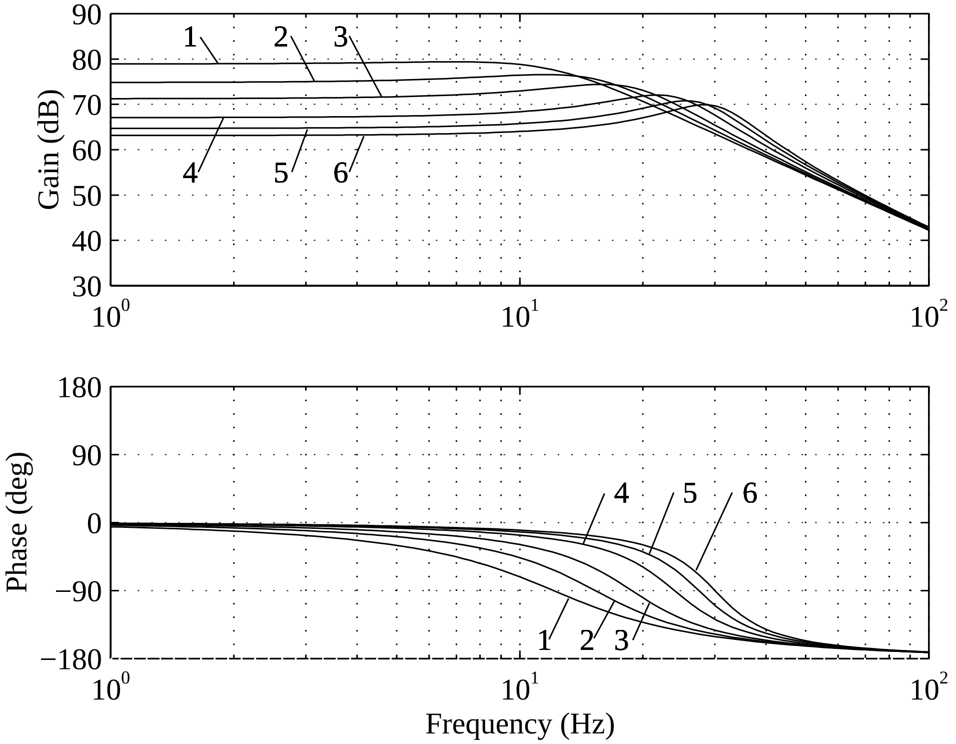
$$J_g \ddot{\theta} = F_c(l_c + l_{gy}) - b_r l_{gy}(\dot{x}_g + l_{gy}\dot{\theta}) - k_r l_{gy}(x_g + l_{gy}\theta). \tag{2.12}$$

Taking the Laplace transform and solving for the linear displacement of the head x_h in terms of the force input at the coil yields

$$G_r(s) = \frac{x_h}{F_c} = \frac{\{(l_h - l_{gy})(l_c + l_{gy})m_g - J_g\}s^2 + \beta_r b_r l_h l_c s + k_r l_h l_c}{\{m_g J_g s^2 + b_r(m_g l_{gy}^2 + J_g)s + k_r(m_g l_{gy}^2 + J_g)\}s^2} \tag{2.13}$$

where β_r is adjusted to match the experimental data.

The third mode occurred at approximately 4.1 kHz, and was due to bending of the actuator arm. A schematic of this mode appears in Figure 2.8, and the transfer function is



Number	Input (v)	Head Displacement (μm)	Natural Frequency (Hz)
1	1.2	2.36	12.1
2	0.7	1.45	15.3
3	0.4	0.91	18.8
4	0.2	0.45	23.9
5	0.1	0.23	27.4
6	0.05	0.12	30.0

Figure 2.4: Frequency Response of the Actuator Pivot Bearing Rotational Mode for a Range of Input Magnitudes

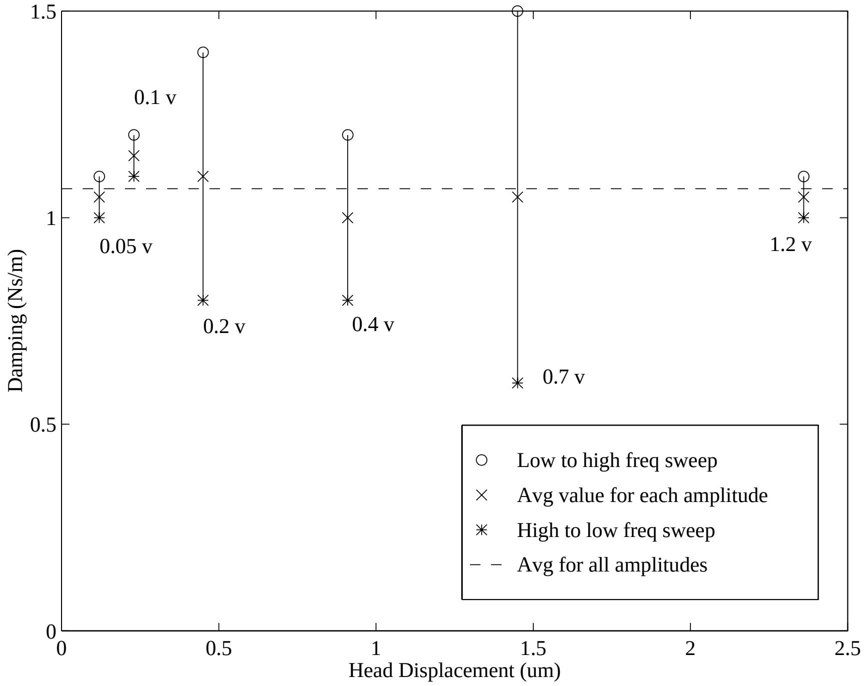


Figure 2.5: Experimental Damping Identification

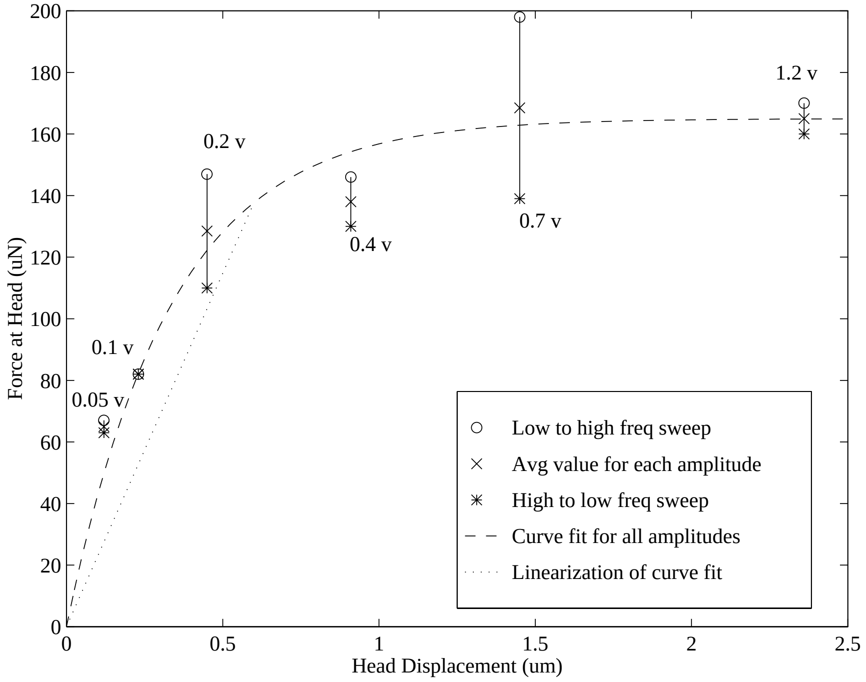


Figure 2.6: Experimental Stiffness Force Identification

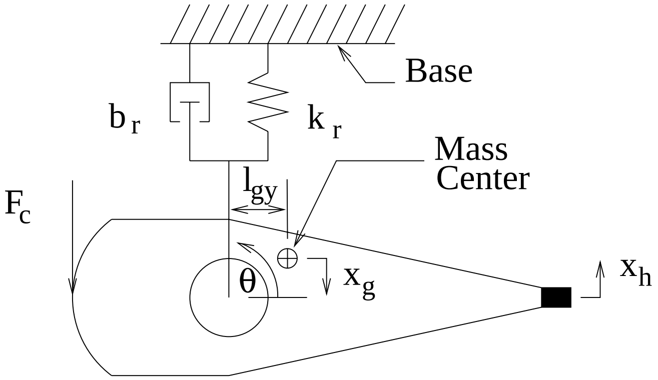


Figure 2.7: Schematic of the Actuator Pivot Bending Mode

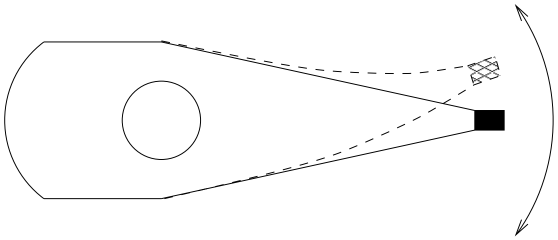


Figure 2.8: Schematic of the Actuator Arm Bending Mode

given by

$$G_s(s) = \frac{x_h}{F_c} = \frac{l_c/l_h}{m_s s^2 + b_s s + k_s}. \tag{2.14}$$

This can be derived from the force balance in the x-direction at the head

$$m_s \ddot{x}_h = \frac{l_c}{l_h} F_c - b_s \dot{x}_h - k_s x_h. \tag{2.15}$$

Combining the effects of the these three modes results in the total frequency response of the actuator from the motor force F_c to the displacement of the head x_h

$$G_{Fh}(s) = G_h(s) + G_r(s) + G_s(s). \tag{2.16}$$

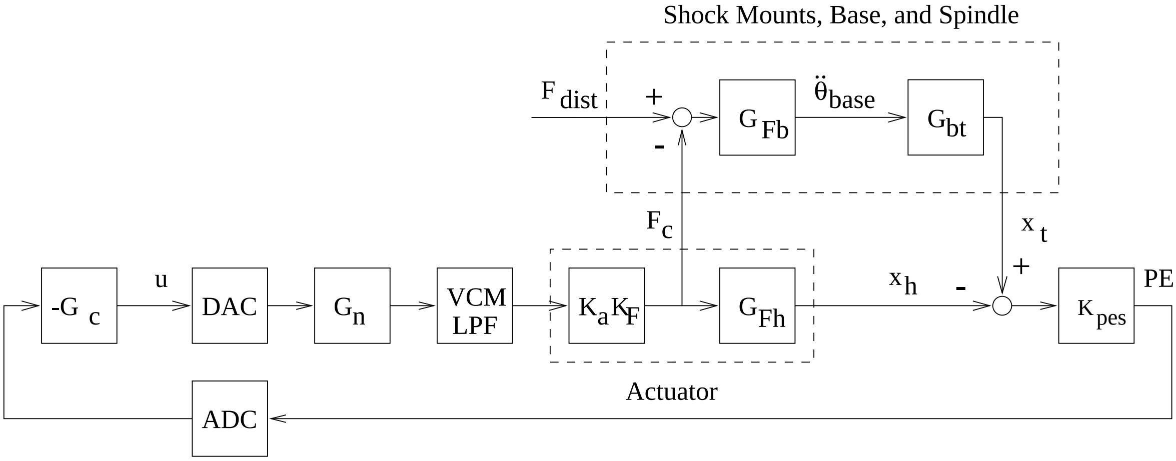


Figure 2.9: Block Diagram of Plant and Nominal Feedback Control System

2.2 Additional Dynamics and Nominal Feedback Controller

The nominal closed loop system consisted of the drive and the feedback controller, G_c. A block diagram of the nominal closed loop system is shown in Figure 2.9. The symbol F_{dist} is the force at the motor coil due to the disturbance at the base. The symbols ADC and DAC represent analog-to-digital and digital-to-analog converters, repsectively. The variables K_a, K_F, and K_{pes} represent the gains of the transconductance amplifier, motor force constant, and position error signal. A low-pass filter (LPF) was used to remove high-frequency signals from the input to the VCM.

The symbols G_{Fb} and G_{bt} represent the transfer functions from the disturbance to the base acceleration and from the base acceleration to the desired track position x_t, respectively. The dynamics of G_{Fb} are primarily determined by the shock mounts between the base and the frame. Recall that both F_c and F_{dist} are forces at the VCM coil, and are related to torques at the actuator pivot by the length l_c. The transfer function G_{Fb} is given by

$$G_{Fb} = \frac{l_c s^2}{J_d s^2 + b_{sm} s + k_{sm}} \tag{2.17}$$

where $J_d = 10.6 \times 10^{-3}$ kgm^2 is the mass moment of inertia of the drive. The values of the shock mount damping b_{sm} and stiffness k_{sm} were determined experimentally to be 500 Nsm and 6720 Nm. This corresponds to a natural frequency of 127 Hz.

The transfer function from the base acceleration to the desired track position involves the bending modes of the spindle and disks and the bearing dynamics. Equipment to perform system identification on the spindle and disks was not available, nor was this information avaialble from the disk drive manufacturer. Most previous research has focused on the NRRO of the spindle bearings [54, 9, 23], although some work has been done on the vibration modes of the spindle and disks using experimentation and finite element modeling (FEM) [41, 51, 19]. For more mathematical derivations of the spindle and disk dynamics, see [48, 58, 57].

The spindle on the drive under test was 2.5 inches high and held eight disks with diameters of 5.25 inches. It was supported at the top and bottom, and rotated at 3600 rpm. In the course of their investigations of NRRO, Richter and Talke [54] noticed resonances of the spindle and bearing at 20 Hz, 40 Hz, and 80 Hz for a 5.25 inch form factor drive spinning at 3600 rpm. A similar result was presented in Harrison and Talke [23], who also note that NRRO was significantly reduced in the case of top and bottom supports. Matsushita *et al.* [41] focused on the frequency region between 100 Hz and 600 Hz, and found as many as six resonances, with various effects on the PES. Frees [19] presents a clear analysis of the pitch (307 Hz), radial (895 Hz), and axial (913 Hz) modes of a 3.5 inch drive with four disks that is directed towards the effects of external vibrations.

Because an accurate model of the spindle dynamics was not available, and because both the adaptive acceleration feedforward technique and the disturbance observer technique provide some robustness to modeling errors, the spindle and disk dynamics were not considered explicitly for either method. This did not appear to affect the performance or the stability of the algorithms.

Compensation for the actuator bending modes of the pivot at $\omega_1 = 2.1$ kHz and

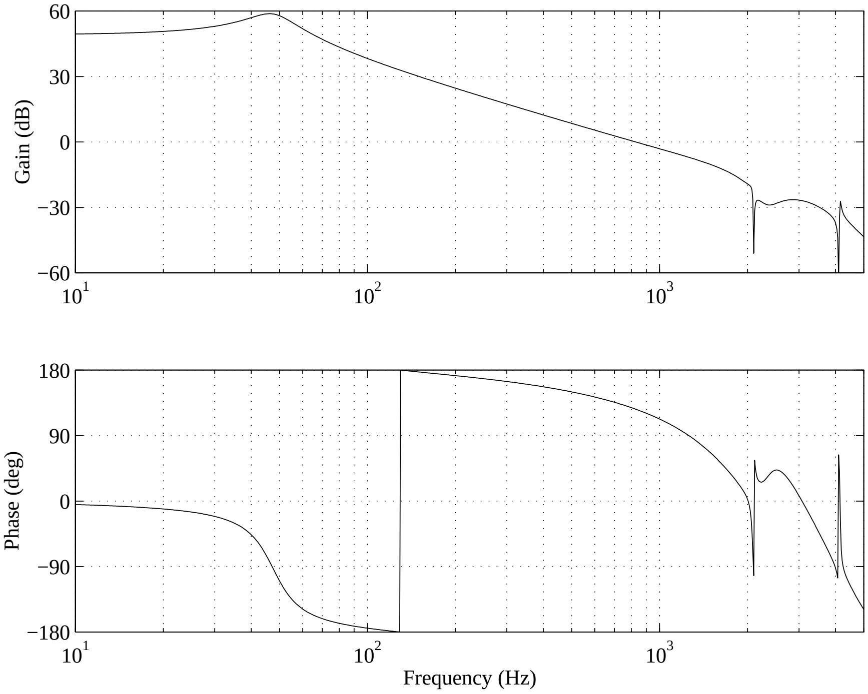


Figure 2.10: Disk Drive Actuator Frequency Response

the arm at $\omega_2 = 4.1$ kHz was made by analog notch filters. A third notch filter had a natural frequency equal to the sampling frequency, $\omega_s = 20$ kHz. The total notch filter transfer function was given by

$$G_n(s) = \frac{s^2+\omega_1^2}{s^2+2\zeta\omega_1 s+\omega_1^2} \times \frac{s^2+\omega_2^2}{s^2+2\zeta\omega_2 s+\omega_2^2} \times \frac{s^2+\omega_s^2}{s^2+2\zeta\omega_s s+\omega_s^2}. \tag{2.18}$$

A damping coefficient of $\zeta = 0.4$ was used for all of the notches. The combined actuator model, from the output of G_c to the PES, is shown in Figure 2.10.

The disk drive was equipped with a factory-installed analog feedback controller. Before experimentation, the analog controller was switched off and replaced with a digital controller. The digital feedback controller was a lead-lag filter designed in continuous time with the transfer function

$$G_c(s) = \frac{(s+2\pi 64)(s+2\pi 67)}{(s+2\pi 6.1)(s+2\pi 3280)}. \tag{2.19}$$

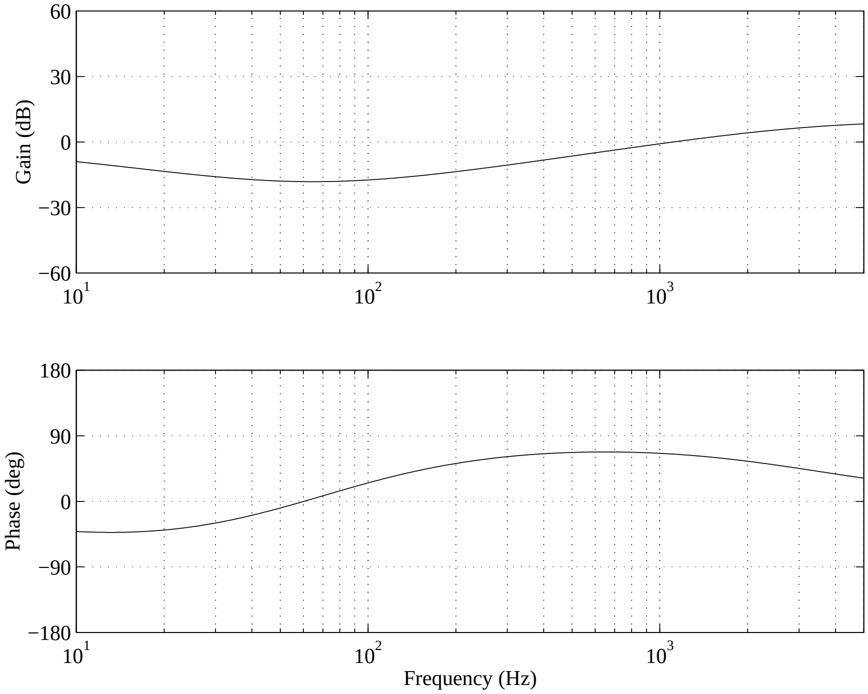


Figure 2.11: Nominal Feedback Controller Frequency Response

This transfer function was converted to discrete-time using the bilinear transformation

$$s = \frac{2(z-1)}{T_s(z+1)} \tag{2.20}$$

where T_s is the sampling time. The frequency response of the nominal controller G_c is shown in Figure 2.11.

2.3 Experimental Model Verification and Parameter Identification

The complete plant included the shock mounts, base, and actuator, as well as the effects of some of the disk drive hardware and software. A block diagram of the complete system and its explanation can be found in Appendix A. It was not possible to measure the plant frequency response directly, but the plant model was verified through the measurement of the open loop transfer function. The open loop frequency response was calculated from the known digital controller and the modeled plant, and compared to the measured open loop frequency response as shown in Figure 2.12. Frequency response measurements were obtained with a Hewlett-Packard dynamic signal analyzer (DSA) using a swept-sine noise input.

Application of the nominal controller to the plant resulted in an open loop gain crossover frequency of 635 Hz. The gain and phase margins were 5 dB and 25°, respectively. This is fairly standard for disk drive applications, *cf.* [64, 73, 74]. The damping and stiffness of the actuator bearing had prominent effects. Values of $b_h = 1.205$ Ns/m, $k_h = 1080$ N/m, and $\beta_r = 7.0$ were assigned based on the measured frequency response. These values of k_h and b_h differ from the averages calculated previously, but are within the expected range. A time delay of approximately 40 μs, due to the conversions between the analog and digital signals and the computations, contributed additional phase lag to the system.

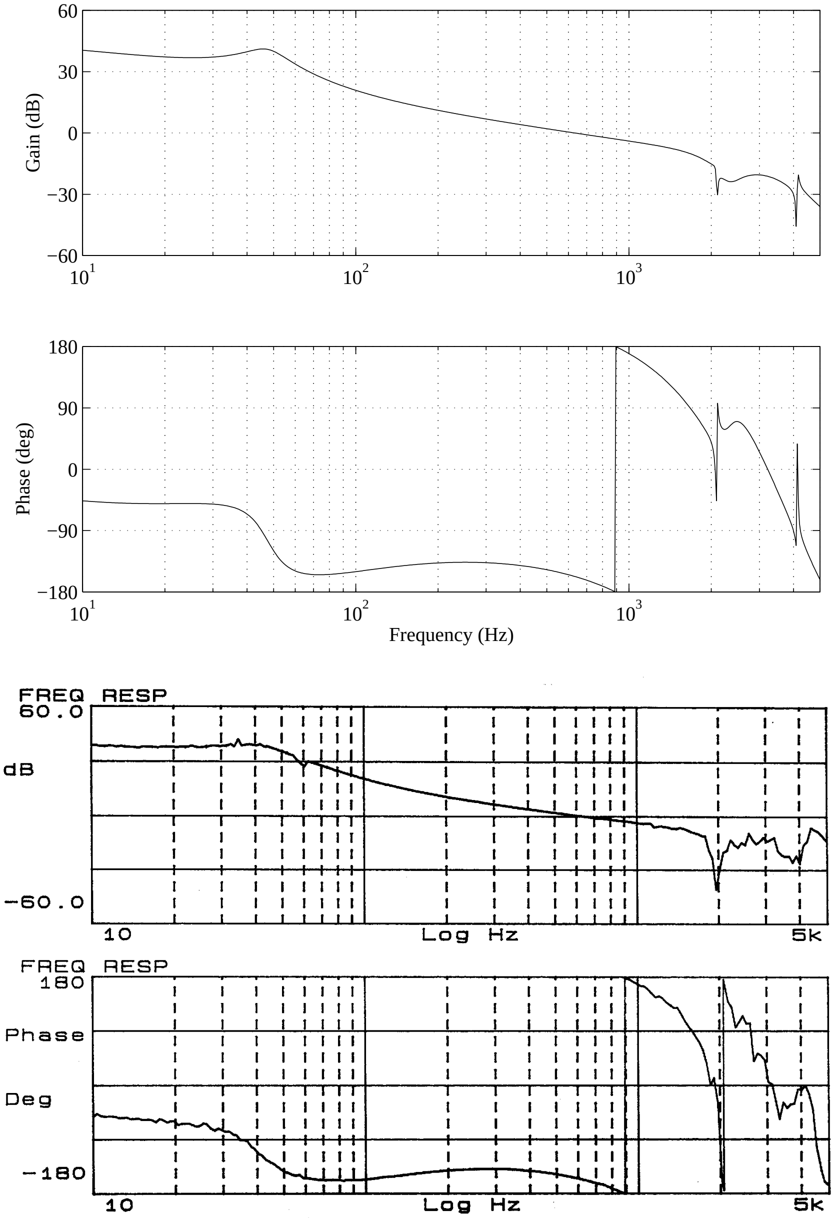


Figure 2.12: Comparison of the Modeled and Measured Open Loop Frequency Responses

Chapter 3

Acceleration Feedforward Algorithm

A block diagram of the nominal feedback system with the acceleration feedforward controller included is shown in Figure 3.1. As shown in Figure 2.1, two linear accelerometers were mounted on the drive base to measure the tangential components of the acceleration, and were used to calculate the angular acceleration of the drive base. The accelerometer signal was sent through a low-pass filter (LPF) with a -3 dB point of 1 kHz before it was received by the feedforward compensator G_{ff}. The feedforward controller calculated the additional control input $\hat{v}$ that was required due to the disturbance to the base, and this signal was added to the actuator input calculated by the feedback controller.

It should be noted that this technique was not pure feedforward. Because of the reaction force of the actuator, there was a feedback component of the output of the feedforward controller. However, due to the ratio of the masses of the actuator and the base, the feedforward component dominated, and the feedback component did not significantly affect stability. A frequency response of the open loop transfer function without and with the acceleration feedforward controller is given in Figure 3.2.

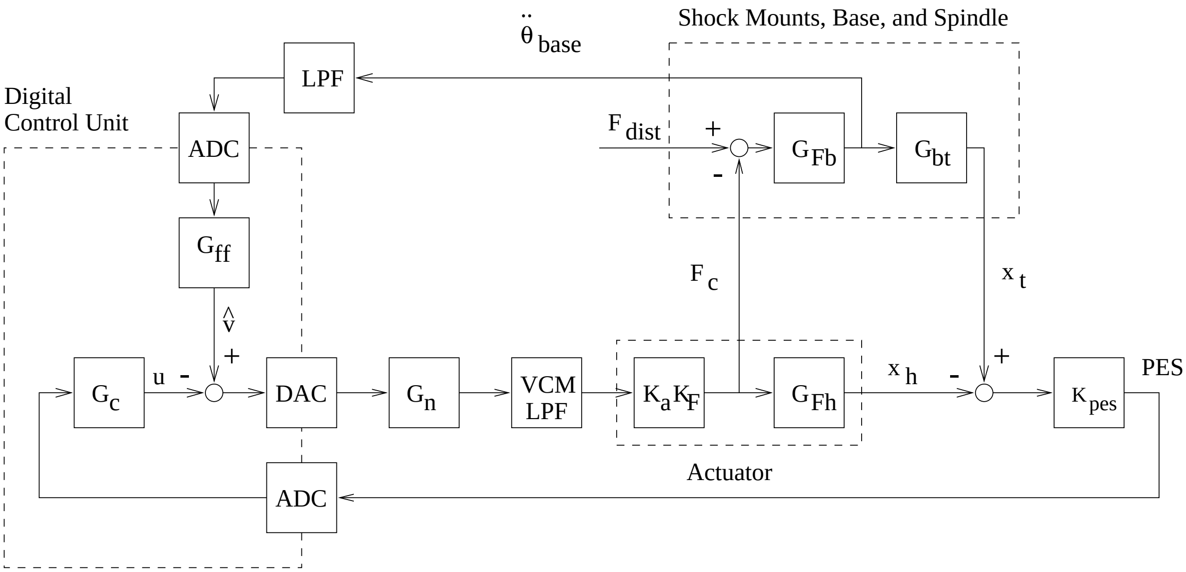


Figure 3.1: Block Diagram of Plant and Control System

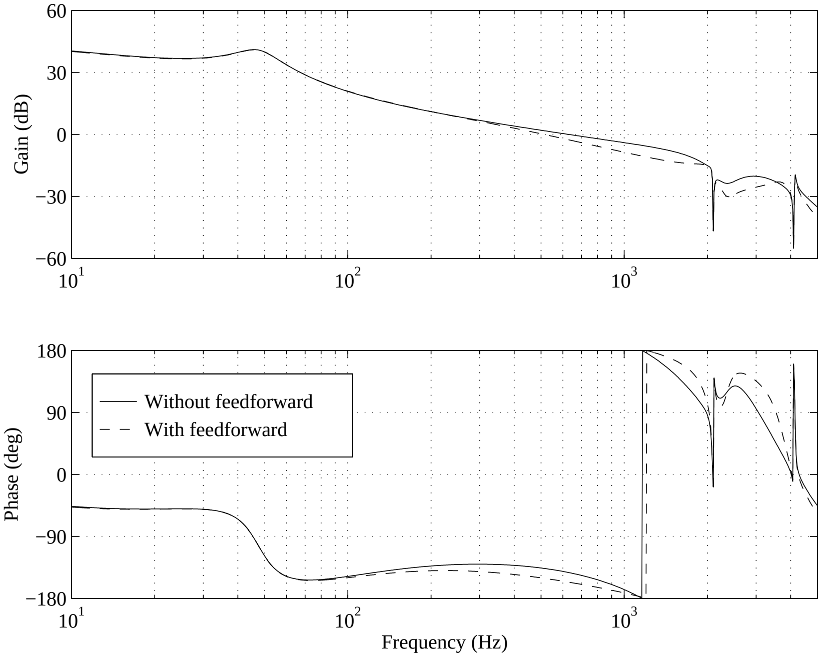


Figure 3.2: Open Loop Frequency Response without and with the Feedforward Controller

The feedforward controller was designed to match the electromechanical impedance between the base acceleration and the PES. By ignoring the reaction force feedback and VCM LPF effects, the block diagram of Figure 3.1 may be rearranged to appear as Figure 3.3. The reference transfer function G_{ref} represents the dynamics to be canceled by G_{ff}, and is given by

$$G_{ref} = \frac{G_{bt}}{K_a K_F G_{Fh}}. \tag{3.1}$$

The transfer function G_{out} represents additional dynamics before the PES,

$$G_{out} = \frac{K_a K_F K_{pes} G_{Fh}}{1 - K_a K_F K_{pes} G_{Fh} G_c}. \tag{3.2}$$

From Figure 3.3, it can be seen how the proper choice of G_{ff} can cancel the disturbance effects. Note that v is a mathematical artifact, not a physical signal in the system, and therefore could not be measured.

3.1 Fixed Parameter IIR Filter

The first controller design method tested was a simple fixed-parameter, infinite impulse response (IIR) filter, also known as a recursive filter or an auto-regressive moving-average (ARMA) filter [49]. The expected value of G_{ff} was determined through the use of the model information. This result was verified through experimental measurements. Recall that the signal v shown in Figure 3.3 is not measurable. Because of this, the frequency response of G_{ff} had to be determined indirectly. To begin, $G_{wo} = G_{ref} G_{out}$ was measured with the feedforward controller turned off. Then the frequency response G_w from the base acceleration $\ddot{\theta}_{base}$ to the PES was measured with a nominal feedforward controller $G_{ff,nom}$ in place. Thus, G_{out} could be calculated as $(G_{wo} - G_w)/G_{ff,nom}$. Finally, the desired G_{ff} was calculated as G_{wo}/G_{out}. The modeled and measured results for G_{ff} are shown in Figure 3.4. Note that the variations on the magnitude of the measured G_{ff} above 300 Hz are due to the measurement technique and the mathematical operations performed on the

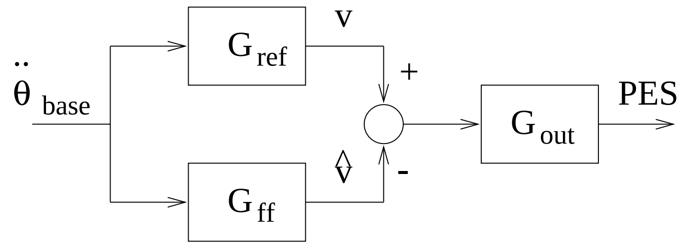


Figure 3.3: Impedance Matching for Feedforward Controller Design

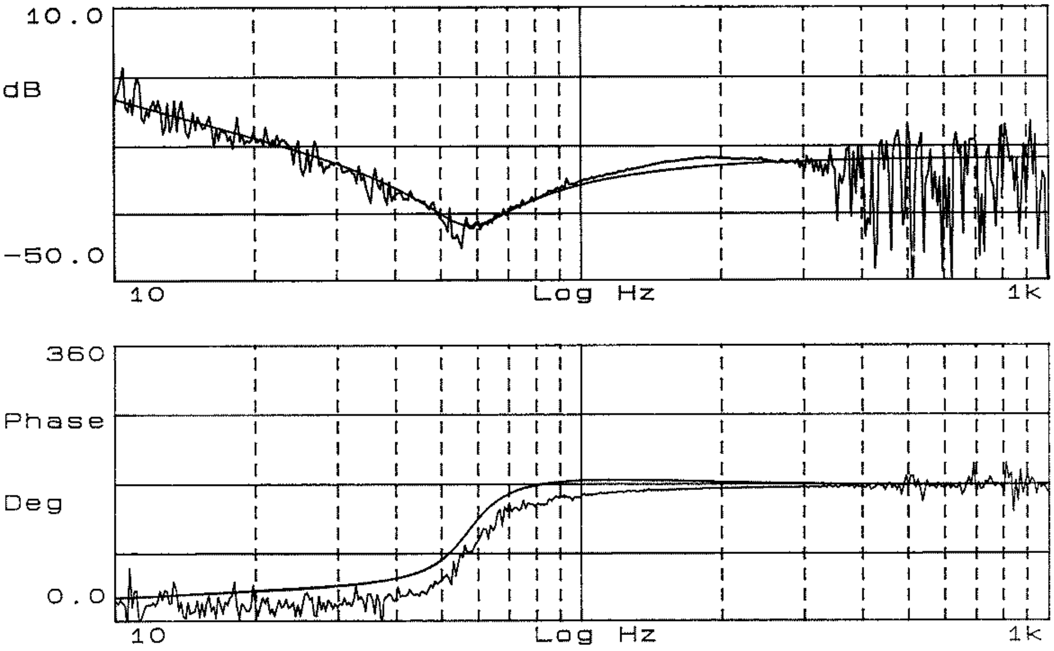


Figure 3.4: Comparison of the Measured and Modeled Feedforward Controller Frequency Responses

DSA in calculating G_{ff}, and are not real features of the system. As with G_c, the transfer function G_{ff} was designed in continuous time and converted to discrete time using the bilinear transformation. A third-order model was found to be adequate for G_{ff}. The results of applying the IIR feedforward controller are presented in Chapter 4.

3.2 Adaptive FIR Filter

Experimentation with the fixed-parameter design showed that performance of the feedforward controller was heavily dependent upon accurate modeling of the system. In mass-produced products such as disk drives, variations in system parameters between units are common. A typical variation is the position of the actuator arm center of mass. Additionally, the dynamics of a single drive are known to vary with age and use. These changes can be long-term due to extended wear or short-term due to such factors as thermal effects, although time scales are typically long enough that the system may be modeled as a time-invariant. To combat the problems associated with parameter variations, adaptation on the feedforward controller was applied using a model reference version of the filtered-x LMS technique described in [72]. The filtered-x LMS algorithm has been used extensively in the field of active noise control [20, 14, 53].

To develop the filtered-x LMS algorithm, begin with the standard recursive form (using the standard notation)

$$\hat{\theta}(k+1) = \hat{\theta}(k) + M(k)\phi(k)e(k+1) \tag{3.3}$$

where $\hat{\theta}(k)$ is the vector of parameters to be adapted, $M(k)$ is the adaptation gain, $\phi(k)$ is the regression vector, and $e(k)$ is the prediction error, all at time k. The least-mean-squares (LMS) algorithm is a simplification of the projection algorithm [21] with

$$\begin{aligned} M(k) &= \text{constant} && (3.4) \\ e(k+1) &= y(k+1) - \phi^T(k)\hat{\theta}(k) && (3.5) \end{aligned}$$

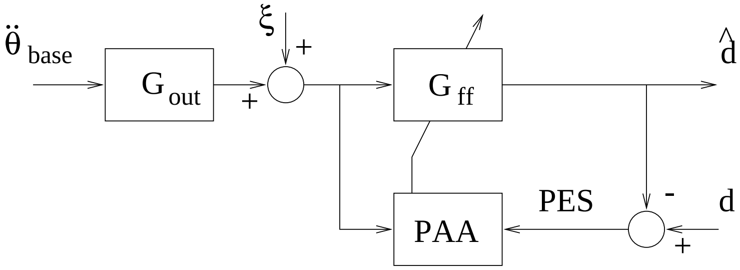


Figure 3.5: Least-Mean-Squares Adaptive Algorithm

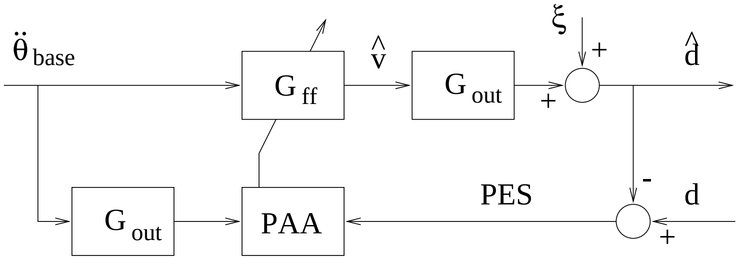


Figure 3.6: Filtered-x LMS Adaptive Algorithm

$$\phi(k) \quad = \quad [\ddot{\theta}_{base}(k), \ddot{\theta}_{base}(k-1), ..., \ddot{\theta}_{base}(k-N)]^T. \tag{3.6}$$

Note that with an finite impulse response (FIR) adaptive filter, the regression vector is composed only of the inputs $\ddot{\theta}_{base}(k)$, without the outputs $y(k)$.

For the disk drive system, the prediction error $e(k)$ is given by the PES $\varepsilon(k)$ and the adaptation gain is 2μ for the estimated parameters $\mathbf{w}$. The LMS algorithm applied to the disk drive system may be represented by the block diagram in Figure 3.5. In Figure 3.5, the signal d corresponds to the output y, and $\hat{d}$ is the estimated output $\phi^T\hat{\theta}$.

Graphically, the filtered-x LMS algorithm differs from the LMS algorithm in that the adaptive filter G_{ff} is placed before the plant dynamics G_{out}, as shown in Figure 3.6. Ignoring the noise ξ, these two algorithms are identical when the weight vectors of the two versions of the adaptive filter are equal, and G_{out} and G_{ff} commute. Two transfer functions can be shown to commute if they are both linear and time-invariant. Although the plant dynamics may be linear and time-invariant, the adaptive filter certainly is not. However, if the parameters of the adaptive filter are fixed, it is linear and time-invariant. Thus, under the conditions that the plant dynamics are linear and the parameters are slowly varying, the LMS and the filtered-x LMS algorithms are approximately equal. Interestingly, the filtered-x LMS algorithm has been shown to perform well even with rapid adaptation [72].

Note that there is a further obstacle before the filtered-x LMS algorithm can be implemented. Namely, an estimate of G_{out} is needed to generate the input signal to the parameter adaptation algorithm. However, if accurate modeling were available, adaptation would probably not be necessary. One option is to expand the adaptation to include an estimation of G_{out} in parallel with the actual plant dynamics, and then copy the adapted parameters to the estimate before the adaptation algorithm. This technique is shown in Figure 3.7. This increases the computation time and was not performed in the results presented here. Similarly, more complicated methods are available to estimate the expectation, but were not employed due to the extra computation requirements. Experience has shown that the estimate $\hat{G}_{out}$ does not need to be particularly accurate to achieve good behavior [14, 53, 72]. Given this fact, it is possible to use a rough *a priori* estimate of the plant dynamics and dispense with the additional adaptation. Just how rough the estimate can be is discussed in the convergence proof in Appendix B.

To apply this algorithm to the disk drive system, the signal d that is compared with $\hat{d}$ is generated by a reference model G_{ref} in series with the plant dynamics G_{out}. The appropriate transfer function for G_{ref} is the electromechanical impedance between the disturbance and the PES. Thus the system itself provides the reference model and its output d. This connection of the reference model to the filtered-x LMS algorithm is shown in Figure 3.8.

The filtered-x LMS algorithm was originally designed to reduce the effects of plant noise on the adaptation. In the LMS algorithm shown in Figure 3.5, the input to the adaptive filter G_{ff} contains the signal ξ, which is not correlated to the dynamics that G_{ff}

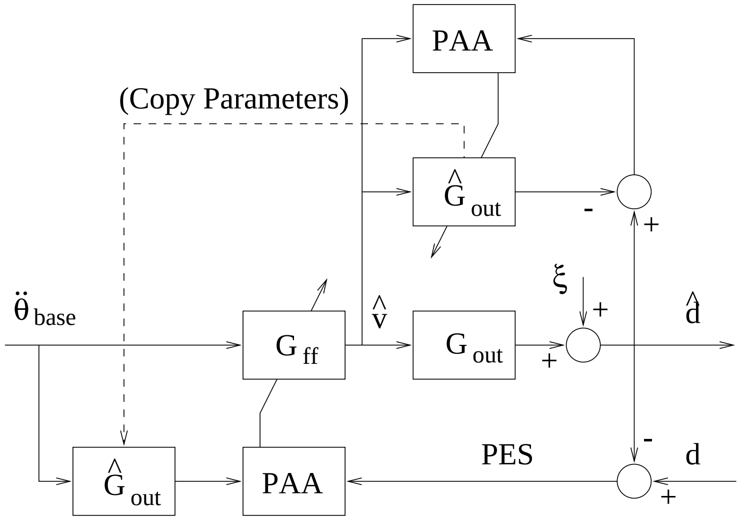


Figure 3.7: Filtered-x LMS Adaptive Algorithm with Extended Adaptation for G_{out}

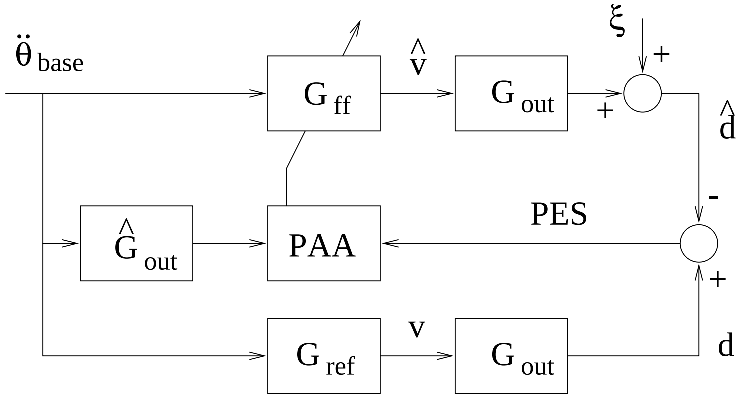


Figure 3.8: Model Reference Form of the Filtered-x LMS Algorithm

is trying to emulate. This significantly affects the results of the adaptation algorithm. By placing G_{ff} before the plant, the noise is no longer a direct input to the adaptive filter. Noise effects can still be seen in the PES, but this will not effect the converged solution of the adaptive filter [72]. Conveniently, with the extension of the filtered-x LMS algorithm to include a reference model, it is in the correct form for the disk drive feedforward adaptation.

With a few simplifications, Figure 3.8 can be reduced to Figure 3.9. Note that the signals v and $\hat{v}$ in Figure 3.8 are both inputs to G_{out}, resulting in the signals d and $\hat{d}$ that are compared to find the PES. The dynamics represented by G_{out} actually occur only once in the physical system, so it is appropriate to represent the difference of v and $\hat{v}$ as the input to a single G_{out} block, the output of which is the PES. Since $\hat{G}_{out}$ is an estimate, it may be included in the parameter adaptation algorithm (PAA) block. The result, neglecting the addition of noise, is Figure 3.9.

The adaptive feedforward control algorithm follows the same basic idea as the fixed-parameter algorithm. The goal is to adjust the feedforward controller G_{ff} so that it is equal to the transfer function G_{ref} of the physical system. For an FIR filter, G_{ff} takes the form

$$G_{ff}(k, q^{-1}) = w_0(k) + w_1(k)q^{-1} + ... + w_L(k)q^{-L} \tag{3.7}$$

where q^{-1} represents the one-step delay operator, and the parameters w_i, or tap weights, are adjusted on-line via the filtered-x LMS algorithm. The tap weights are adjusted in the direction that minimizes the expectation of the squared error with constant gain μ, or

$$\mathbf{w}(k+1) = \mathbf{w}(k) - \mu \frac{\partial}{\partial \mathbf{w}(k)}(E[\varepsilon^2(k)]) \tag{3.8}$$

where $\mathbf{w}(k)$ represents the vector of tap weights at time k, E denotes the expectation, and $\varepsilon(k)$ is the PES. Calculating the gradient $\frac{\partial}{\partial \mathbf{w}}$ results in

$$w_i(k+1) = w_i(k) + 2\mu E[\varepsilon(k)\{G_{out}(q^{-1})\ddot{\theta}_{base}(k-i)\}]. \tag{3.9}$$

This is the desired update law. However, the expectation is not known, nor is the transfer function $G_{out}(q^{-1})$. Thus, some approximations must be made. The transfer function $G_{out}(q^{-1})$ is replaced with an *a priori* estimate $\hat{G}_{out}(q^{-1})$, and the current value of $\varepsilon(k)\{\hat{G}_{out}(q^{-1})\ddot{\theta}_{base}(k)\}$ is used in place of the expected value. This results in an actual update law given by

$$w_i(k+1) = w_i(k) + 2\mu\varepsilon(k)\{\hat{G}_{out}(q^{-1})\ddot{\theta}_{base}(k-i)\}. \tag{3.10}$$

From Equation 3.10 it can be seen that the filtered-x LMS algorithm is of the standard recursive adaptation form with $\varepsilon(k)$ as the prediction error and $\{\hat{G}_{out}(q^{-1})\ddot{\theta}_{base}(k-i)\}$ as the regressor.

Figure 3.9 shows how the adaptation algorithm may be added to the feedforward technique shown in Figure 3.3. The filtered-x LMS algorithm is appealing for this application because it is simple and has very few computations. This is mainly due to the fact that it uses a constant adaptation gain. Disk drive control systems typically have little space or time to run complicated processes. In addition, the filtered-x LMS algorithm is designed to decrease the effects of noise. Using an FIR filter is also helpful in decreasing noise sensitivity [16], and FIR filters are always stable. The stability of IIR filters is dependent upon their parameters, which often means that the algorithm must include extra calculations to check the stability of the IIR filter before it is implemented. This is not necessary with an FIR filter.

There are some drawbacks to this algorithm. Because the adaptation gain is constant, this technique converges more slowly than some more complicated algorithms and there is no inherent signal normalization. Also, FIR filters typically require more parameters than IIR filters to describe the same frequency response.

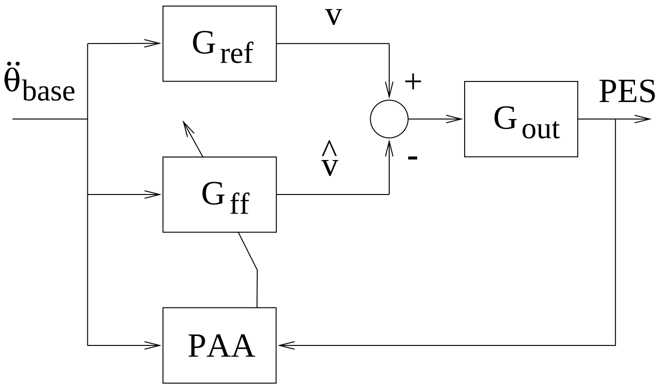


Figure 3.9: Impedance Matching for Feedforward Controller Design with Adaptive Feedforward Control

Chapter 4

Simulation and Experimental Results of Acceleration Feedforward Tests

4.1 Vibration Testing for Track Following

Vibration disturbances are a significant problem in RAID (Redundant Array of Independent Disks) applications. A RAID is a system of disk drives working together that is used for high-end server applications. The seeking motion and spindle vibration of each drive can cause disturbances in the other drives, typically in the frequency range from 5 Hz to 500 Hz. Vibration disturbances are also common in standard desktop and portable applications.

To test the efficacy of the feedforward control technique, comparisons of the position error signal with and without the acceleration information were made. For the experimental tests, the drive was put into track-following mode. Various vibration disturbances were generated by the shaker system, and the position errors were measured. For the simulation results, the model described in Chapter 2 was used. Computer code for the simulations was written in C language with the `dverk` integration routine. Post-processing and frequency response analyses were performed using MATLAB.

4.1.1 Frequency Response Results

To calculate the frequency response from the base acceleration to the position error, a swept-sine was sent to the shaker system that created an acceleration of approximately 1 G. This led to position errors on the order of 1 μm for the case with feedback control only, which is approximately the limit for reading and writing on a disk with 2000 tracks per inch. For higher density disks, this limit will obviously decrease.

Figure 4.1 shows the frequency response between the acceleration of the drive and the PES. Between 40 Hz and 400 Hz, the PES was reduced from 50% to 90% using the fixed-parameter IIR feedforward controller. The response above and below this range was relatively unaffected. Note that in all cases, the measured PES also contained repeatable runout components due to eccentricities of the disks that were not induced by the disturbance. Also note that the frequency responses contain a number of resonances that do not appear in the open loop frequency response in Figure 2.12. The disturbance generated by the shaker is exciting modes such as spindle bending that were not excited by the noise input used to generate the open loop frequency response.

The same experimental tests were run for the filtered-x LMS algorithm that were run for the fixed-parameter design. Figure 4.2 shows the magnitude of the frequency response from the disturbance to the PES with the adaptive feedforward compared to the case without feedforward. The PES is reduced between 55% and 95% over the range of 15 Hz to 400 Hz. Compare these results to those shown in Figure 4.1. Note that the performance has improved with the addition of the adaptation, especially in the low frequency region.

Even with adaptation, the improvement with acceleration feedforward control is

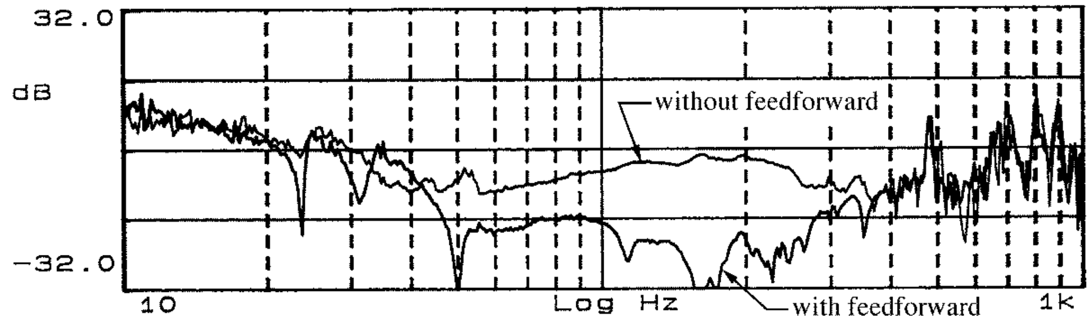


Figure 4.1: Magnitude of Experimental Frequency Response from Base Acceleration to PES without and with Fixed-Parameter IIR Feedforward Control

not as dramatic in the low frequency range. The signal strength of the acceleration measurement is very weak in the low frequency range. This is due to the accelerometers themselves (see the results of the accelerometer testing in Section A.1) and possibly a reduced disturbance amplitude from the shaker.

Although the FIR filtered-x LMS method is computationally efficient for an adaptation algorithm, it still requires more computations than the fixed-parameter scheme. The computation time for the adaptive algorithm is dependent upon the number of parameters included in the FIR filter. Experimentation showed that the optimal number of tap weights was approximately 40, so the sampling time for the adaptive case was increased to 100 μs. This is still quite reasonable for disk drive applications, as most sample in the 5 kHZ to 10 kHz range. Using fewer parameters at 50 μs sampling or more parameters, at sampling times of 100 μs and greater, did not significantly improve upon the results. Widrow *et al.* [71] have shown that increasing the number of parameters does not always improve performance. Sampling the fixed-parameter controller at 100 μs showed little change from the results presented previously.

The disk drive manufacturer's specifications for vibration disturbance rejection were 0.25 G during operation and 0.5 G during non-operation. This is somewhat lower than the levels used for the feedforward control tests. A frequency range was not included

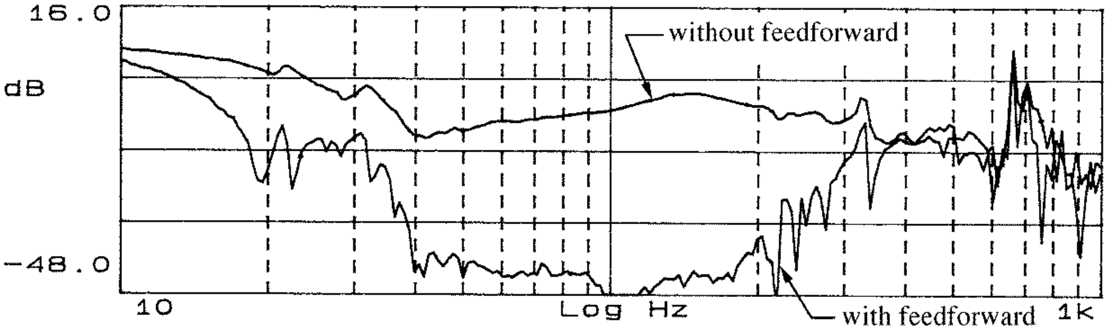


Figure 4.2: Magnitude of Experimental Frequency Response from Base Acceleration to PES without and with Adaptive FIR Feedforward Control

in the specifications, but a conversation with the manufacturer indicated that 5 Hz to 500 Hz is typical. This is consistent with specifications from other disk drive companies. Thus the frequency range that the feedforward controller was able to improve performance matches the desired frequency range very well.

4.1.2 Time Response Results: IIR Feedforward Control

Sample time traces are shown in Figures 4.3-4.6 for frequencies of 60 Hz and 200 Hz. At 60 Hz, the PES was 2.3 μm without feedforward control and 1.1 μm with the fixed-parameter IIR feedforward controller. For 200 Hz, the error amplitudes were 2.8 μm and 0.8 μm without and with the feedforward controller, respectively. These correspond to reductions of 52% and 71%, which agree with the transfer function results.

4.1.3 Time Response Results: Convergence of the Adaptive FIR Feedforward Controller

PES time responses are shown in Figures 4.7-4.10. Figure 4.7 and Figure 4.9 show the effects of the disturbance without feedforward control to be 1.8 μm at 40 Hz and 2.6 μm at 100 Hz. Figures 4.8 and 4.10 show the PES values for the same disturbances while the filtered-x LMS feedforward controller converges. Initially, the FIR filter parameters are set

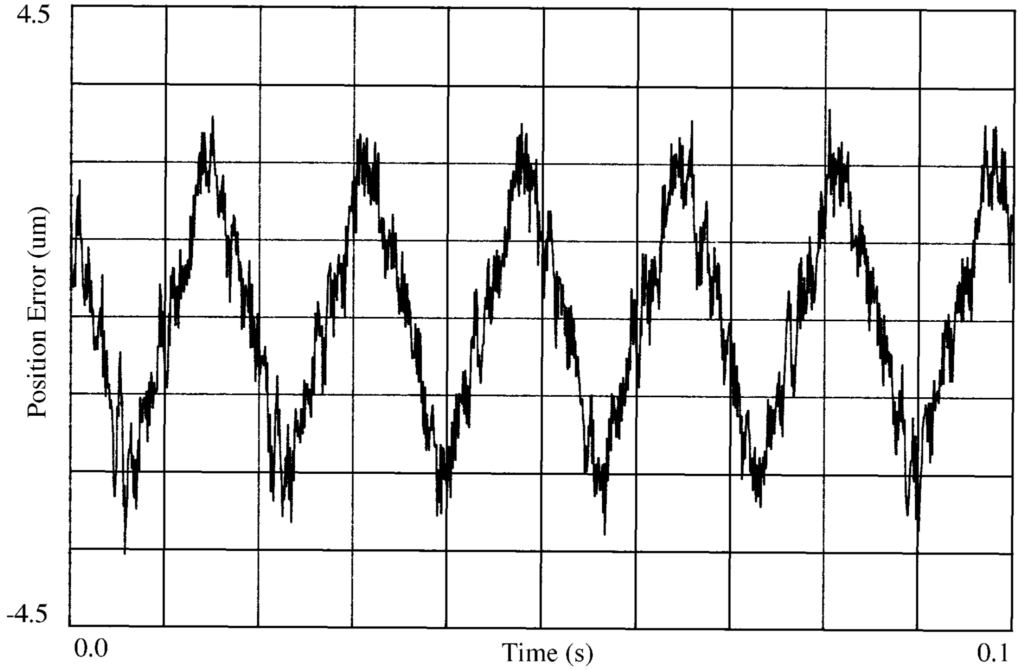


Figure 4.3: Experimental Time Trace without Feedforward Control at 60 Hz

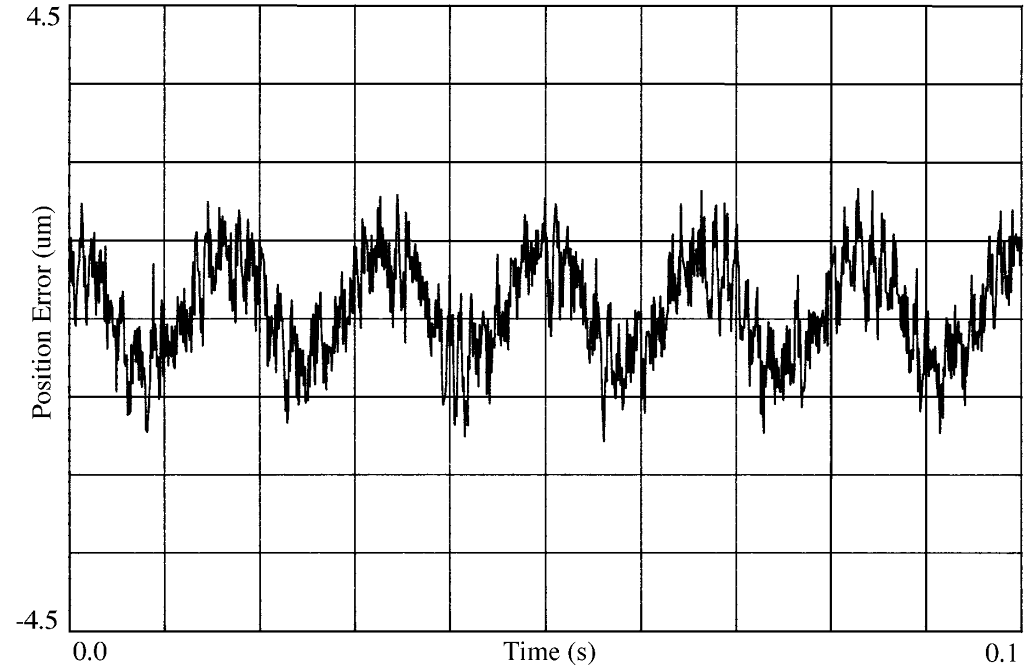


Figure 4.4: Experimental Time Trace with Fixed-Parameter IIR Feedforward Control at 60 Hz

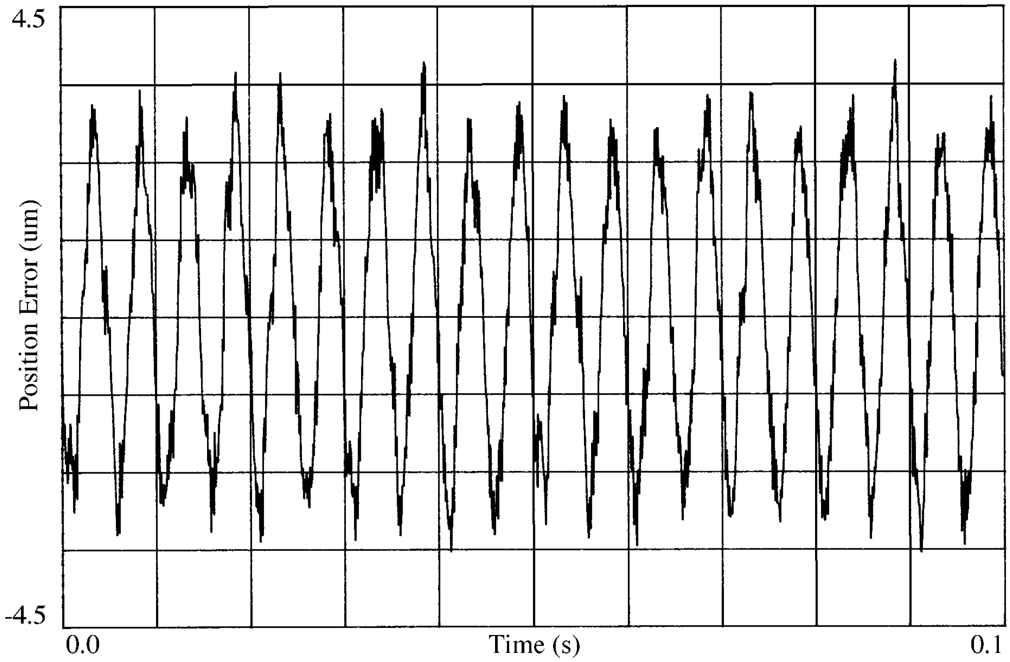


Figure 4.5: Experimental Time Trace without Feedforward Control at 200 Hz

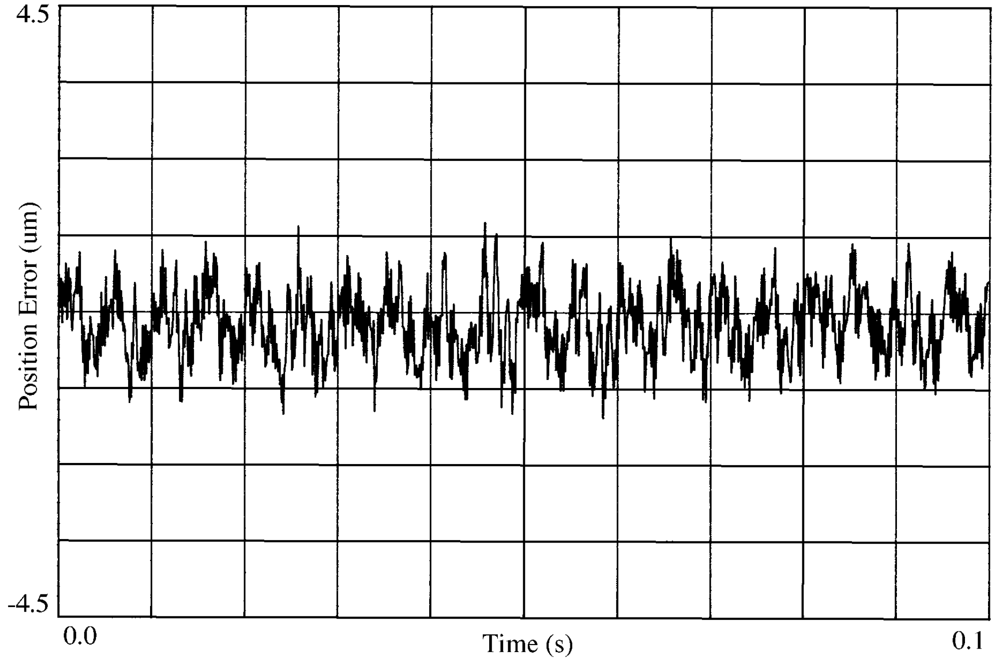


Figure 4.6: Experimental Time Trace with Fixed-Parameter IIR Feedforward Control at 200 Hz

equal to zero and there is no disturbance from the shaker. Note that the PES is nonzero due to error sources unrelated to the shaker disturbance. After the shaker disturbance begins, the feedforward parameters are allowed to adapt to their steady-state values. After the feedforward parameters have converged, the PES is reduced to approximately 0.6 μm at both 40 Hz and 100 Hz, which is about the level of the PES before the shaker disturbance was introduced. This is an improvement of 67% for the 40 Hz disturbance and 77% for the 100 Hz disturbance. The value of the adaptation gain, $\mu = 1.5 \times 10^{-7}$, used in the experiments shown was a fairly moderate value. The rate of convergence can be increased by increasing the value of μ, but at the expense of robustness to the magnitude of the disturbance. Limitations on the magnitude of μ for stability of the adaptation algorithm are discussed in Appendix B. The transient effects in Figures 4.8 and 4.10 are due to the convergence of the parameters. Once the parameters of the feedforward controller have converged, transients are negligible.

The same sort of results can be generated using a random disturbance. The ANSI C `rand` function was used to generate a random integer between 0 and 32,767 with a uniform distribution. The random integer was converted to floating point, normalized, and offset to give a random real number between the maximum and minimum acceleration levels. Figures 4.11 and 4.12 show the simulation results for the convergence of the adaptive algorithm for a random input, compared to the case without feedforward. The root-mean-square (RMS) value of the PES without feedforward control is 1.28 μm for an RMS base acceleration of 0.53 G. Note that the convergence time is somewhat longer for the random input compared to the single frequency vibration disturbances that were shown previously.

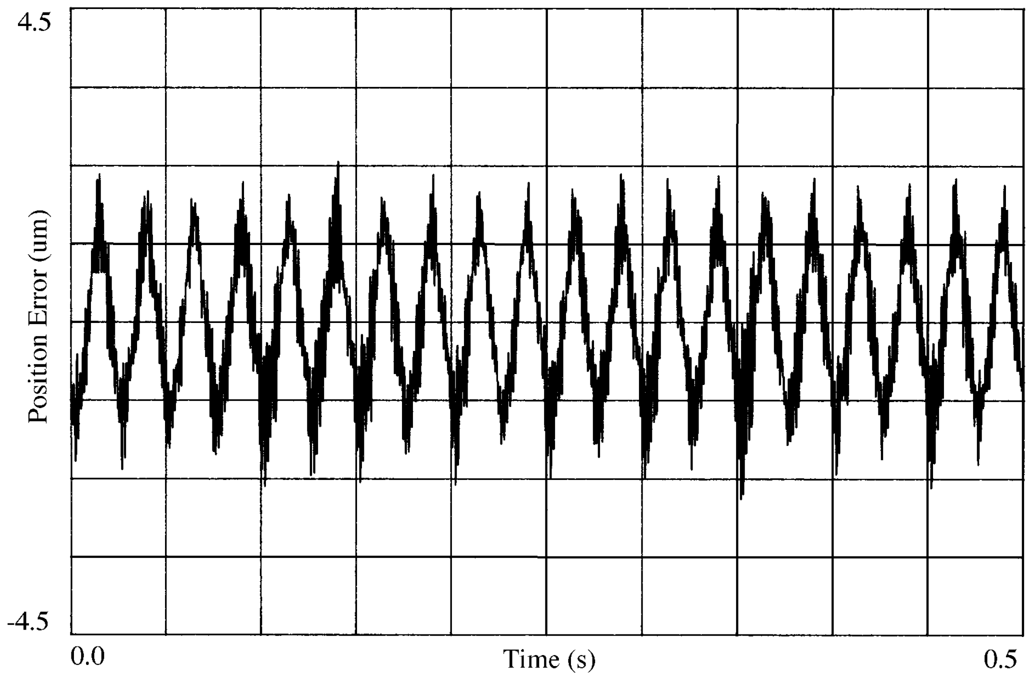


Figure 4.7: Experimental Time Trace without Feedforward Control at 40 Hz

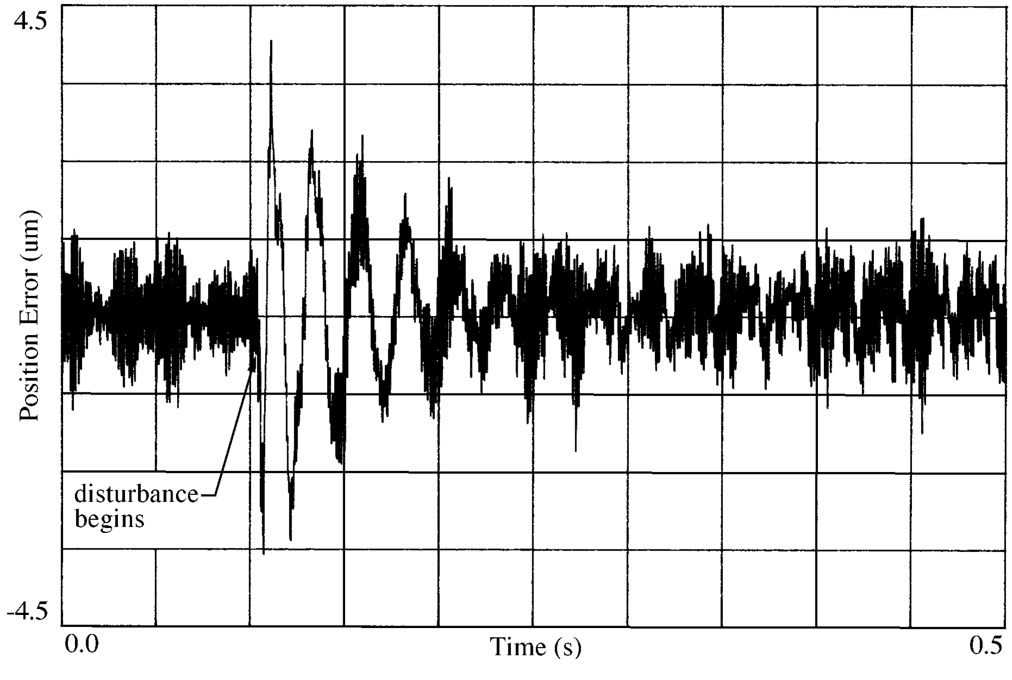


Figure 4.8: Experimental Time Trace with Adaptive FIR Feedforward Control at 40 Hz

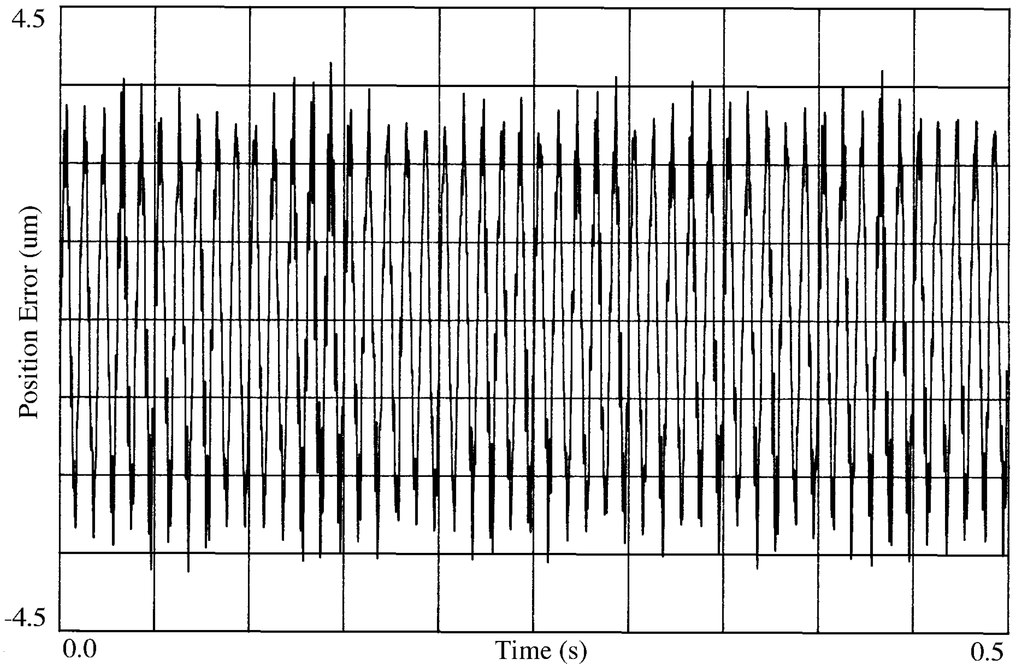


Figure 4.9: Experimental Time Trace without Feedforward Control at 100 Hz

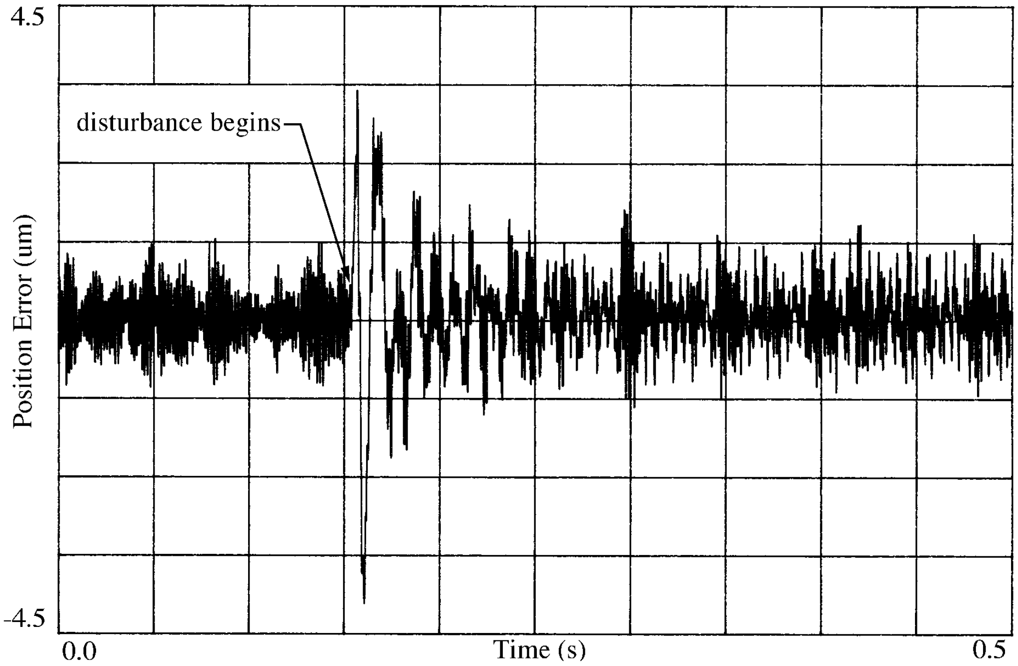


Figure 4.10: Experimental Time Trace with Adaptive FIR Feedforward Control at 100 Hz

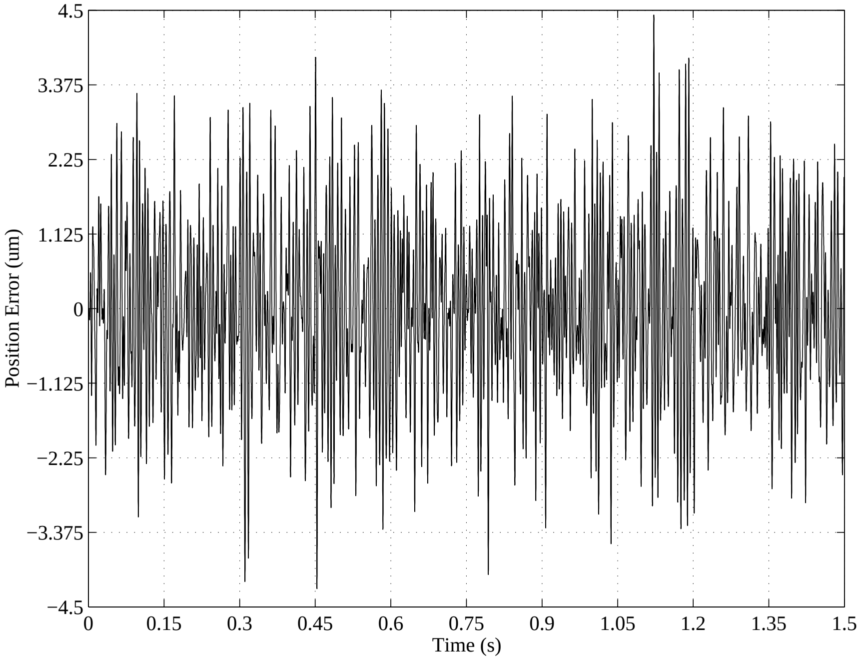


Figure 4.11: Simulation Time Trace without Feedforward Control for a Random Base Acceleration

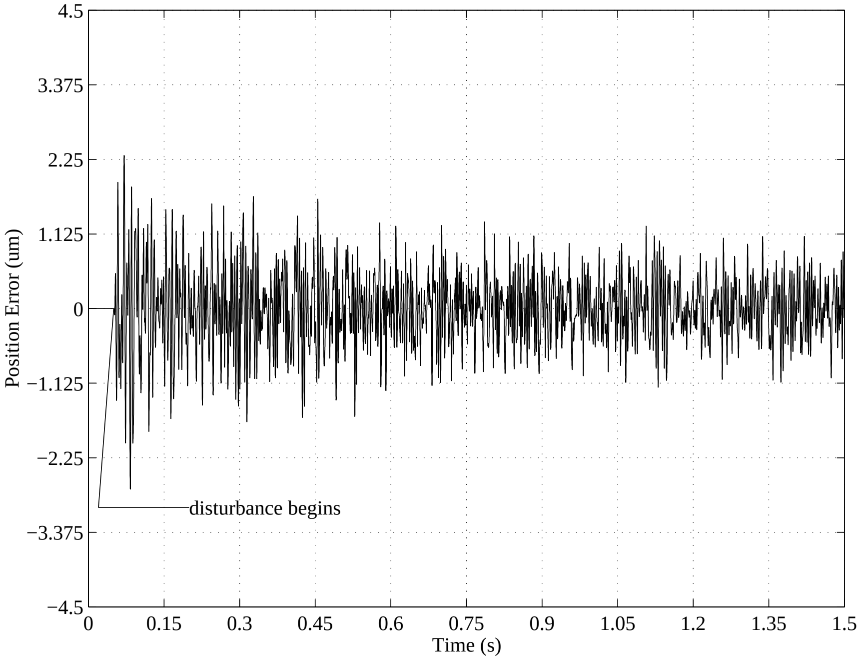


Figure 4.12: Simulation Time Trace with Adaptive FIR Feedforward Control for a Random Disturbance

4.1.4 Time Response Results Using Converged Values for the Feedforward Controller

The previous section demonstrated the convergence of the adaptive algorithm for a variety of inputs. Although the convergence of the algorithm was reasonably fast, the performance requirements of the disk drive will not allow the adaptation to start from a state of no information. Starting the adaptation with zero initial conditions may also result in transient peaks that are larger than the PES amplitude without feedforward control in some cases, *e.g.* Figure 4.8. In fact, it may be preferable not to run the adaptation algorithm during normal operation. This will save on computation. In this case, the parameters of the feedforward should be identified through adaptation when the drive is started up and then again whenever a change in parameters is expected, *e.g.* long enough after start-up for thermal effects to appear. Between adaptation sessions, the feedforward controller will be fixed.

In this section, the parameters were identified using the random input described above. These values were saved, and then used for each of the cases below without further adaptation. Figures 4.13 and 4.15 show simulation results without feedforward for the experimental cases presented in Figures 4.7 and 4.9. The PES amplitudes are 1.93 μm and 3.18 μm. The simulated values of the PES are slightly higher than the experimental values, but are still reasonable. Figures 4.14 and 4.16 show results using the converged FIR feedforward controller for the sinusoidal disturbances investigated previously. The PES is 0.22 μm at 40 Hz and 0.48 μm at 100 Hz. This represents reductions of 89% and 85%. Note that the reduction of the PES is significant, but does not approach zero as in the cases with continuous adaptation.

Figure 4.18 shows the time response with the FIR feedforward controller, again with the adaptation turned off. The RMS PES is 0.23 μm for an RMS base acceleration of 0.53 G. This gives a reduction of 82% when compared to the result given in Figure 4.17

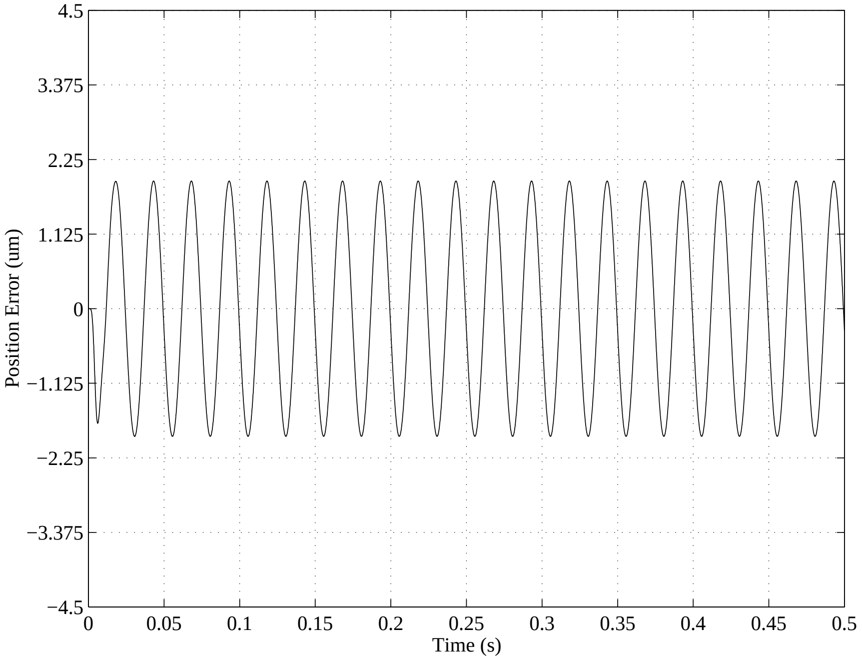


Figure 4.13: Simulation Time Trace without Feedforward Control for a Disturbance of 40 Hz

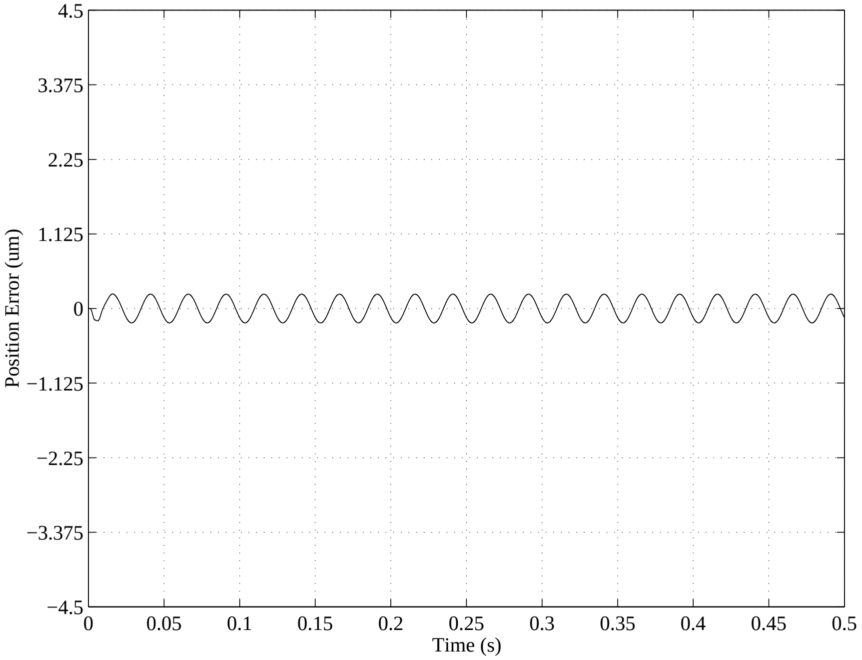


Figure 4.14: Simulation Time Trace with Adaptive FIR Feedforward Control for a Disturbance of 40 Hz

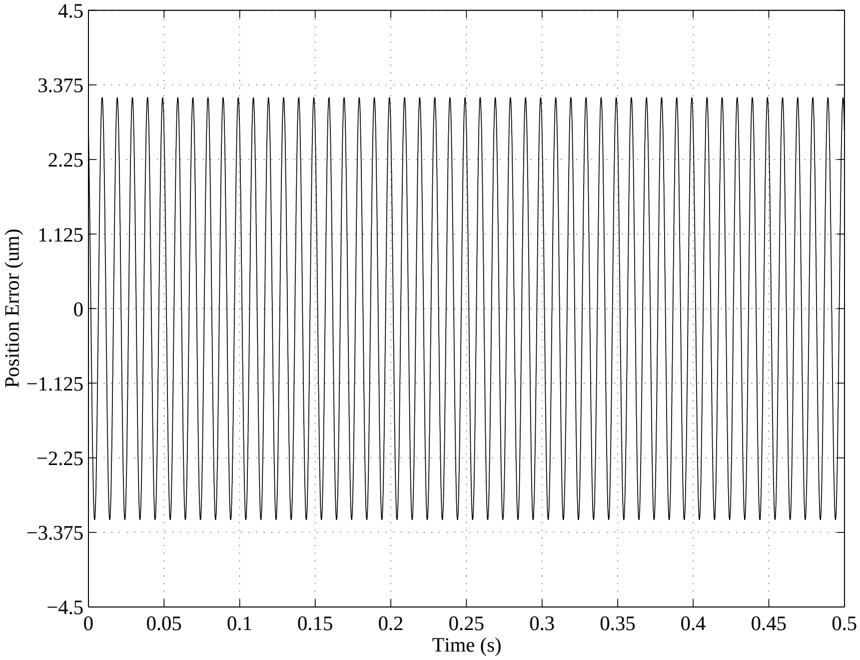


Figure 4.15: Simulation Time Trace without Feedforward Control for a Disturbance of 100 Hz

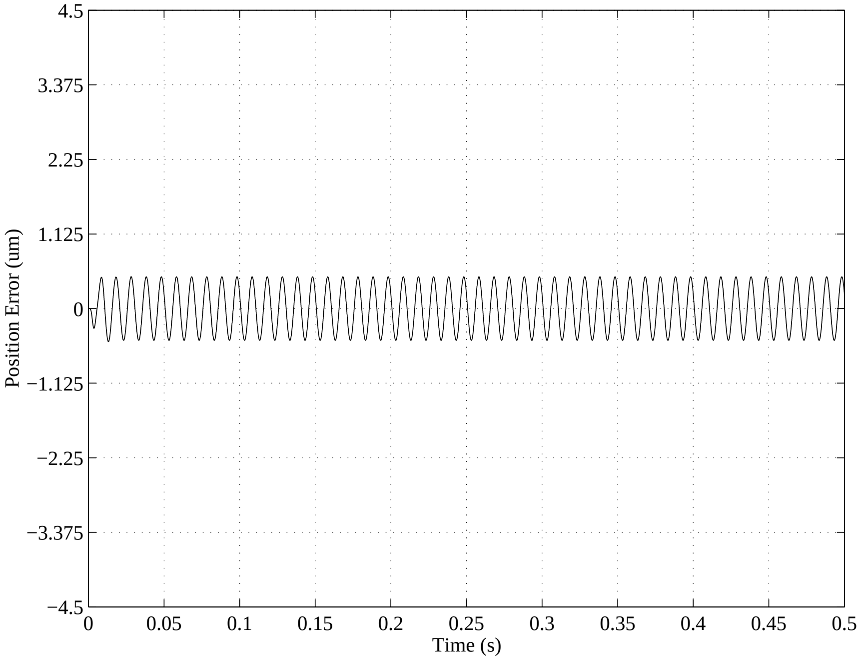


Figure 4.16: Simulation Time Trace with Adaptive FIR Feedforward Control for a Disturbance of 100 Hz

without feedforward control. Note the change in the time axes from Figures 4.11 and 4.12. Figures 4.19 and 4.20 show the power spectral densities (PSD) of the base acceleration and the PES, respectively.

4.2 Shock Testing for Track Following

As mentioned previously, vibration disturbances are common in RAID systems. Shock disturbances, on the other hand, are more common in portable applications. Specifications for both types of disturbances are typically given for all disk drives, for both operating and non-operating conditions. The standard industry shock test is a half sine wave of a given amplitude and duration.

This sort of test was used to determine the shock performance of the disturbance observer in simulation. The operating shock specification for the disk drive under test was 2 G, with a non-operating shock specification of 20 G. No duration was given in the specifications, but a conversation with the manufacturer indicated that 10 ms was standard. This is consistent with specifications from other manufacturers.

Given these specifications, a shock disturbance of 10 ms duration, or a half sine wave of 50 Hz, with an amplitude of 2 G was given to the system. Figure 4.21 shows the simulated PES time trace without feedforward control for the shock disturbance. The peak error is 4.81 μm. Using the adapted parameters for the FIR feedforward controller results in a peak error of 0.51 μm, a reduction of 89% when compared to the case without feedforward control. This result is given in Figure 4.22.

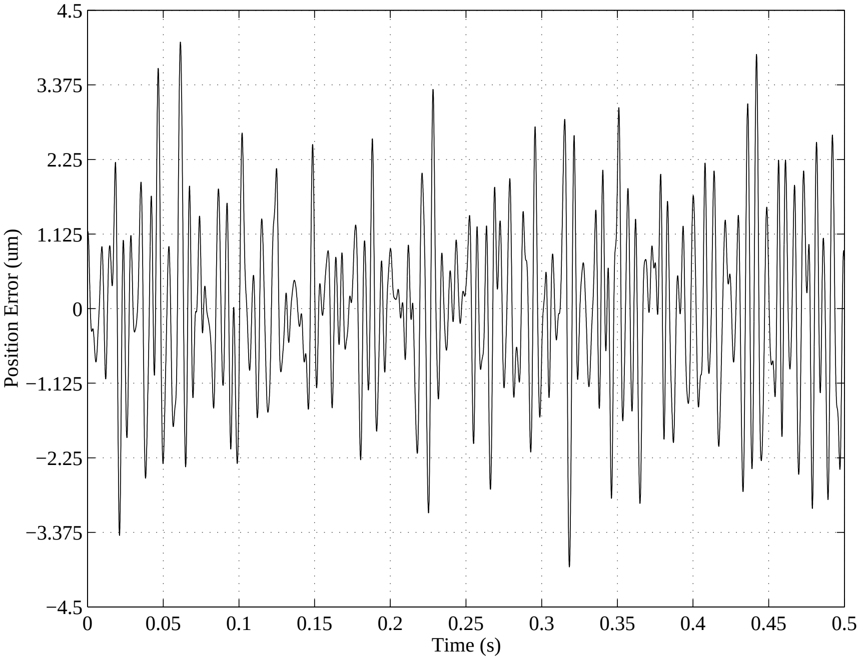


Figure 4.17: Simulation Time Trace without Adaptive FIR Feedforward Control for a Random Base Acceleration

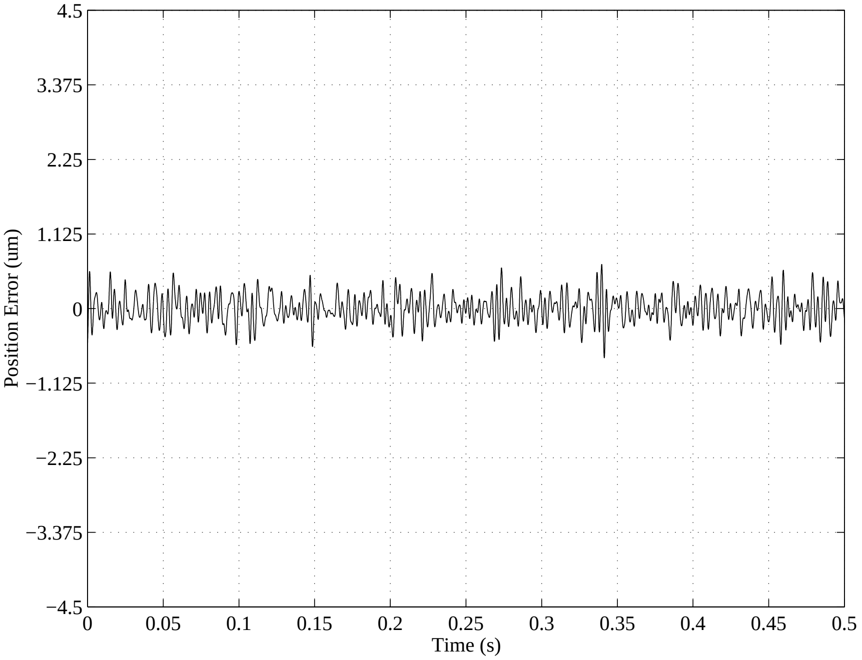


Figure 4.18: Simulation Time Trace with Adaptive FIR Feedforward Control for a Random Base Acceleration

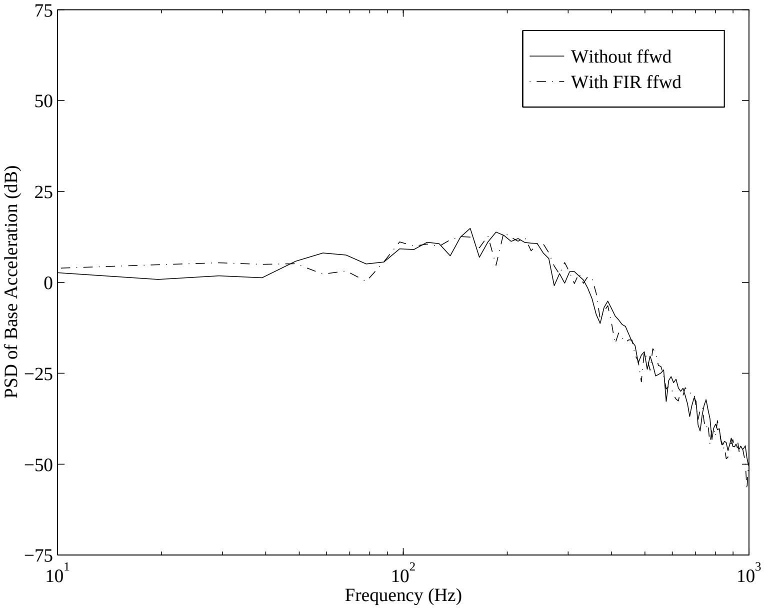


Figure 4.19: Simulated Power Spectral Densities of the Base Accelerations

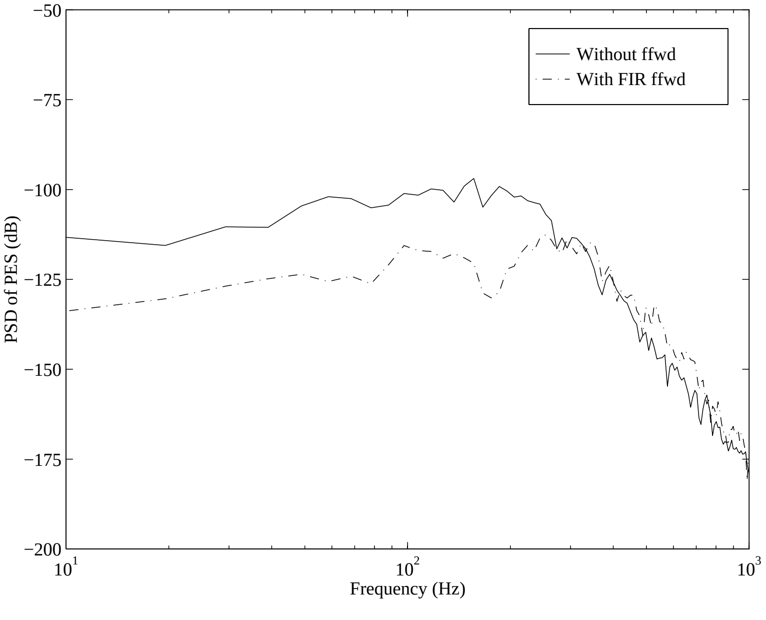


Figure 4.20: Simulated Power Spectral Densities of the PES without and with the Adaptive Feedforward Controller

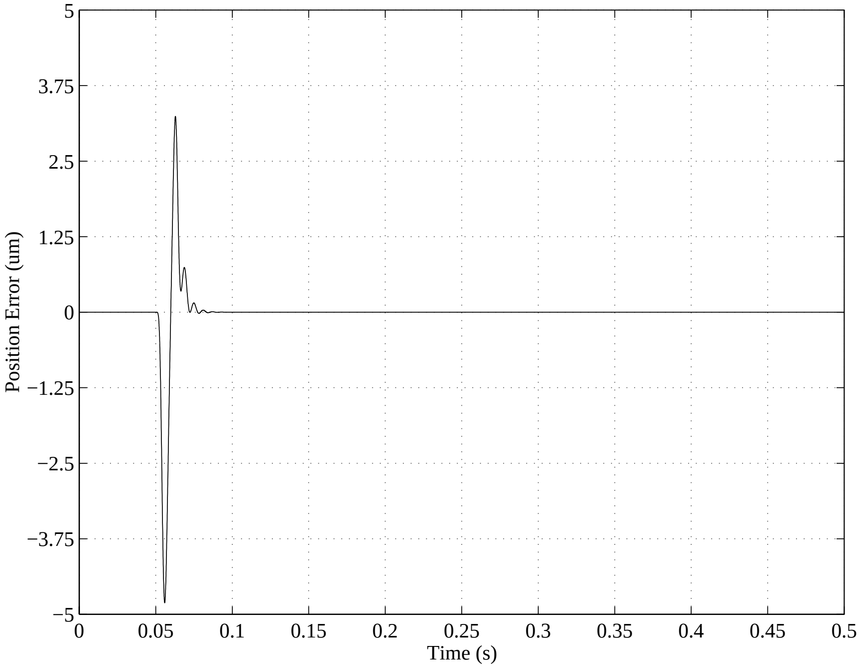


Figure 4.21: Simulation Time Trace without Feedforward Control for a Shock Disturbance

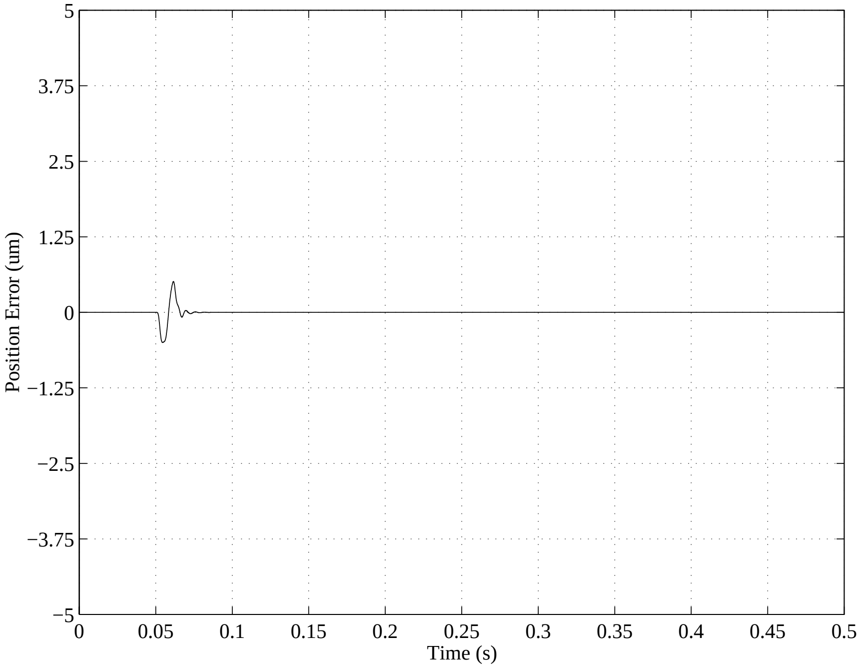


Figure 4.22: Simulation Time Trace with FIR Feedforward Control for a Shock Disturbance

Chapter 5

Disturbance Observer Algorithm

The use of acceleration feedforward control to measure and then compensate for disturbances was shown to be effective in Chapters 3 and 4. However, this technique does require some additional hardware. There are control methods that are designed specifically to reject disturbances that do not require sensors. For example, the disturbance observer algorithm uses an estimate of the disturbance instead of a measured value. The trade-off is that since the disturbance observer is a feedback technique that relies only on the PES, it cannot be expected to respond to a disturbance as quickly as the acceleration feedforward method. The disturbance effects must first appear in the PES before the disturbance observer can compensate for them. The acceleration feedforward controller can measure the disturbance and compensate for it before it appears in the PES.

The disturbance observer algorithm is designed to enforce robust input/output behavior by canceling the effects of disturbances and modeling errors [6]. It was originally proposed by Ohnishi [47], with significant work by Umeno and Hori following shortly thereafter [67, 68]. In addition to those systems listed in Section 1.2.4, applications have included wire bonding [15], steel rolling mills [25], and floppy disk drives [62]. Bickel and Tomizuka [6] found the performance of a disturbance observer to be comparable to adaptive feedback control in a robotic system.

5.1 Algorithm Explanation

Consider the block diagram shown in Figure 5.1. The feedback controller is represented by the transfer function G_c. The disk drive transfer function from the control input u to the PES is given by G_p. Assume that all disturbances into the system are lumped into an equivalent disturbance δ at the input to G_p, and that there is additional measurement noise ξ on the PES. These elements make up the nominal feedback loop.

The disturbance observer consists of the delay block z^{-d}, the inverted plant estimate G_{pn}^{-1}, and the filter Q. The disturbance observer takes the noisy PES signal and passes it through G_{pn}^{-1}. The result is an estimate of the sum of the control signal and the disturbance, plus some noise effects. The combination of $G_p G_{pn}^{-1}$ has some delay. This delay is matched by the value of the exponent d, and includes computational delay. Thus, the output of the delay block is subtracted from the signal g to give a noisy estimate of the disturbance. This estimate of the disturbance is filtered by Q, and subtracted from the control input to offset the addition of the actual disturbance.

To gain a further appreciation of the disturbance observer effects, and to motivate the design of the filter Q, it is instructive to examine the transfer functions of Figure 5.1. The transfer function $G_{\delta\varepsilon}$ from the disturbance to the PES, the transfer function $G_{u\varepsilon}$ from the control input to the PES, and the transfer function $G_{\xi\varepsilon}$ from the noise to the PES are given by

$$G_{\delta\varepsilon} = \frac{(1 - Qz^{-d})G_p G_{pn}}{G_{pn} + Q(G_p - G_{pn} z^{-d})} \tag{5.1}$$

$$G_{u\varepsilon} = \frac{G_p G_{pn}}{G_{pn} + Q(G_p - G_{pn} z^{-d})} \tag{5.2}$$

$$G_{\xi\varepsilon} = \frac{-G_p Q}{G_{pn} + Q(G_p - G_{pn} z^{-d})}. \tag{5.3}$$

If the Q filter is set to unity and the delay is small so that $\{(1 - z^{-d})G\} \approx 0$ for a transfer function G, then

$$\begin{aligned} G_{\delta\varepsilon} &\approx 0 && (5.4) \\ G_{u\varepsilon} &\approx G_{pn} && (5.5) \\ G_{\xi\varepsilon} &\approx -1. && (5.6) \end{aligned}$$

Thus, the disturbance is canceled and the feedback controller sees a transfer function that behaves like the nominal plant, but the sensor noise is unattenuated. Conversely, if $Q = 0$, then

$$\begin{aligned} G_{\delta\varepsilon} &= G_p && (5.7) \\ G_{u\varepsilon} &= G_p && (5.8) \\ G_{\xi\varepsilon} &= 0 && (5.9) \end{aligned}$$

so both the disturbances and the feedback controller see the true plant transfer function, but the sensor noise is eliminated. By designing Q to be a low-pass filter with unity dc gain, it is possible to take advantage of the best of both of these situations. Sensor noise is typically high frequency, so it is attenuated where it is strongest and allowed to pass in those frequencies that it has little power. In the low frequency range where the performance is typically more critical and disturbances are concentrated, the effects of the disturbance and parameter mismatch are canceled as desired. At the higher frequencies, the disturbances can still be expected to be canceled somewhat if G_p is low-pass. Since $G_{u\varepsilon}$ is also close to G_p at high frequencies, robustness to parameter variations cannot be expected in this frequency range. However, since this is typically above the bandwidth of G_c, not much performance can be expected in this region even with a perfect model.

Murakami and Ohnishi [46] provide a further discussion of the trade-offs involved in the choice of the cut-off frequency of Q. Examination of the transfer functions reveals that $G_{\delta\varepsilon}$ correpsonds to G_pS where S is the sensitivity function, and $-G_{\xi\varepsilon}$ corresponds to the complementary sensitivity function T. As T decreases, the stability of the system

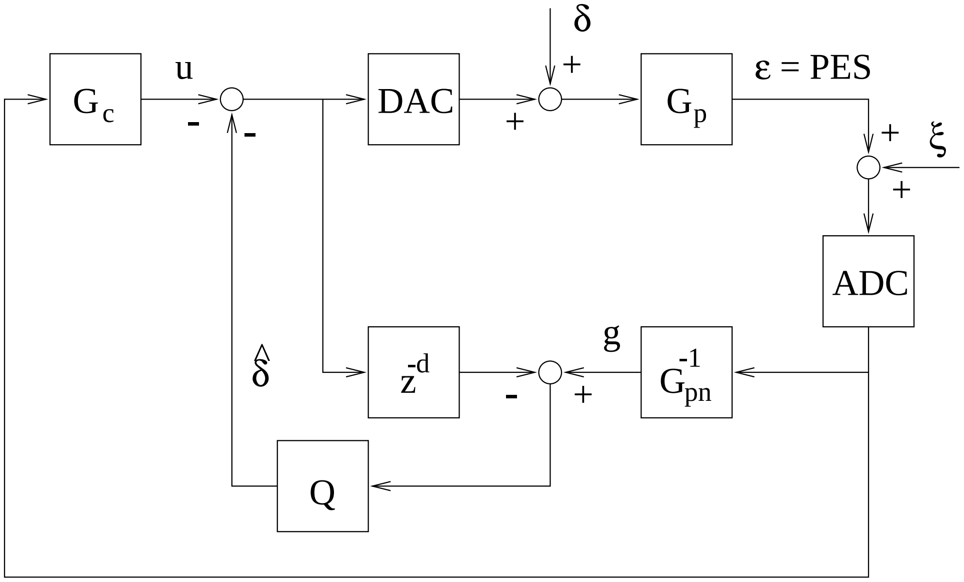


Figure 5.1: Disturbance Observer Block Diagram

increases, and as S decreases, the robustness of the system to disturbances and parameter variations increases. Thus, it is desirable to make both T and S small. Unfortunately, the sensitivity and the complementary sensitivity are related by

$$T + S = 1 \tag{5.10}$$

so that they cannot both be made small. The choice of Q determines in what frequency ranges T is small and in what frequency ranges S is small.

5.2 Other Configurations

The block diagram shown in Figure 5.1 is just one implementation configuration. Designs may also be made in continuous time, and if desired, discretized using the bilinear transformation. For the continuous time design in Figure 5.2, the delay block does not exist. Although the transfer functions remain the same, the block diagram must be manipulated to ensure realizability of the transfer functions as shown in Figure 5.3. In this case, the

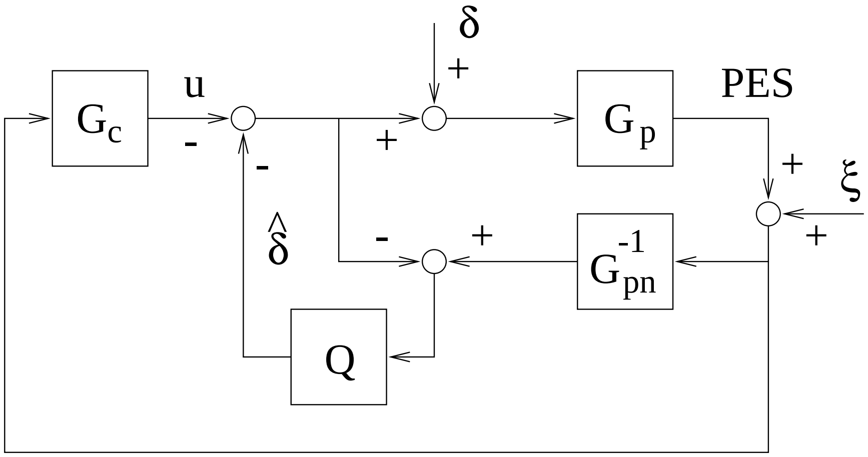


Figure 5.2: Continuous Time Disturbance Observer Block Diagram

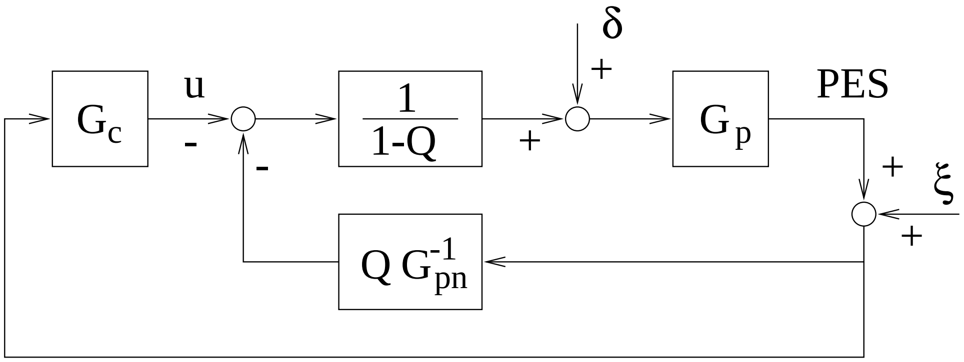


Figure 5.3: Block Diagram of the Implementable Version of the Continuous Time Disturbance Observer

Q filter has the additional function of compensating for the relative order of the inverted plant. A more thorough discussion of this technique may be found in [6].

It is also possible to include the feedback controller within the transfer function G_p, *i.e.* to swap disturbance observer and feedback loops. An example of this technique is given in [62]. The feedback controller may not be completely omitted, however, as it is critical to guarantee the performance and stability of the system.

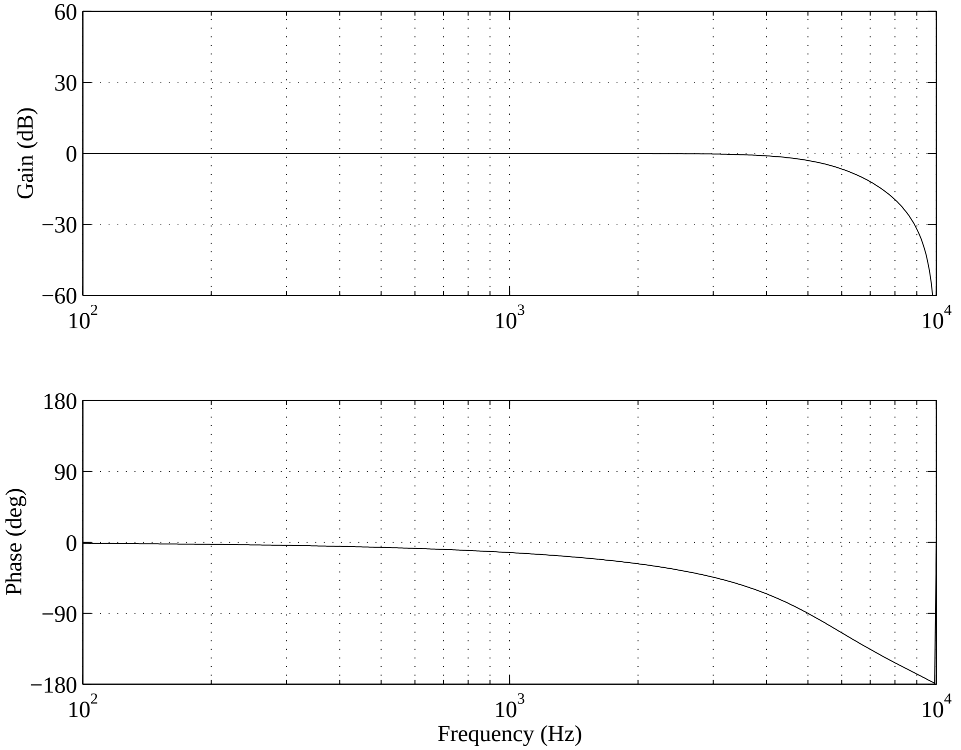


Figure 5.4: Frequency Response of the Low-Pass Q Filter

5.3 Implementation Issues

The design of the Q filter has a major influence upon the performance of the disturbance observer. As mentioned above, in the continuous time case, the relative degree of the Q filter must be set so that the transfer function QG_{pn}^{-1} is realizable. This is not a problem for the discrete-time version that was used to create the results in Chapter 6. However, the cut-off frequency and the rate at which the Q filter rolls off determine the frequency where the trade-offs between the rejection of disturbances and parameter mismatch and the attenuation of sensor noise occur. A second-order Butterworth filter with a break frequency of 5 kHz was chosen for this case, as shown in Figure 5.4.

The disturbance observer requires an inverted model of the plant G_{pn}. Obviously,

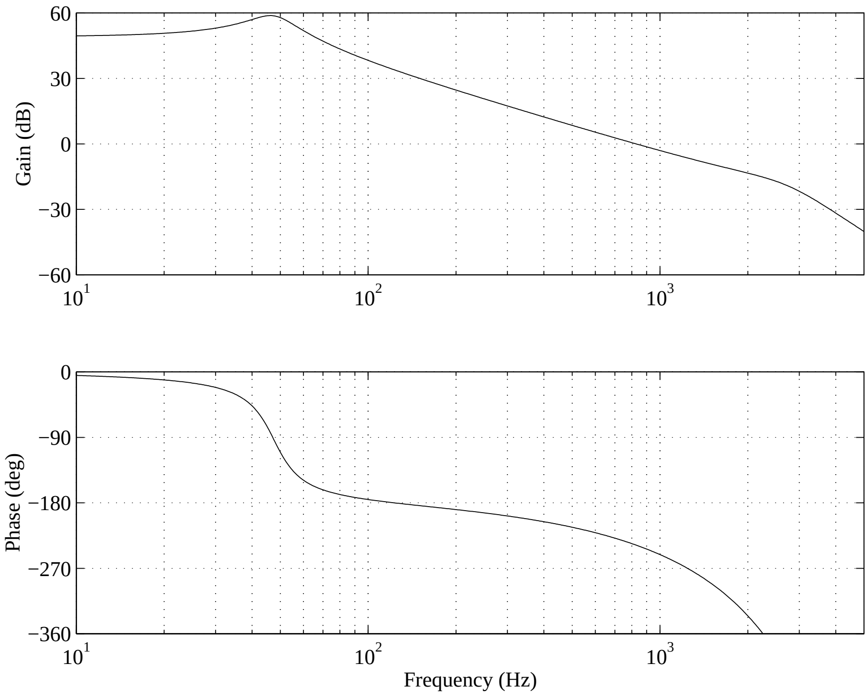


Figure 5.5: Frequency Response of the Nominal Plant Model

unstable zeros cannot be inverted. It is well-known [1] that discretization of a plant, especially with fast sampling times such as those used here, can lead to unstable zeros even when the continuous time plant is minimum-phase. Often, it is undesirable to invert even lightly-damped stable zeros.

The nominal plant was chosen to to be a double integrator with the actuator bearing mode, the VCM LPF, and 150 μs of additional delay to match the effects of the high frequency resonances and the notch filters. The frequency response of G_{pn} is shown in Figure 5.5. Compare this to the full model given in Figure 2.10. Note that the high frequency resonances and the notch filters were not included in the nominal plant.

The discretized version of the plant model had one unstable zero and no stable non-invertible zeros ($\sigma = 1$), and three pure delays ($\rho = 3$). To compensate for the noninvertible zeros, an estimate of the inverted plant model was created using the ideas contained in the zero phase error tracking (ZPET) controller of Tomizuka [65]. To generate the ZPET-like inverted plant, define the numerator and denominator of the discretized plant as

$$G_{pn}(z^{-1}) = \frac{z^{-\rho}B(z^{-1})}{A(z^{-1})} \tag{5.11}$$

where

$$\begin{aligned} B(z^{-1}) &= b_0 + b_1 z^{-1} + ... + b_m z^{-m} && (5.12) \\ A(z^{-1}) &= 1 + a_1 z^{-1} + ... + a_n z^{-n} && (5.13) \end{aligned}$$

and $b_0 \neq 0$. The term $z^{-\rho}$ represents the number of pure delays in the plant. Divide the numerator $B(z^{-1})$ into two parts

$$B(z^{-1}) = B^{-}(z^{-1})B^{+}(z^{-1}) \tag{5.14}$$

where $B^{+}(z^{-1})$ contains the cancellable zeros, and $B^{-}(z^{-1})$ contains the uncancellable zeros. The inverted nominal plant is then given as

$$G_{pn}^{-1}(z^{-1}) = \frac{z^{\rho+\sigma}A(z^{-1})B^{-}(z)}{B^{+}(z^{-1})(B^{-}(1))^2} \tag{5.15}$$

where $(B^{-}(1))^2$ is included to match the dc gain of the true inverted plant, and the exponent σ is the number of uncanceled zeros. Note that because of the $z^{\rho+\sigma}$ term, the transfer function of Equation 5.15 is unrealizable. In a tracking system with a known desired trajectory, the effects of $z^{\rho+\sigma}$ can be compensated for using preview information. In the case of the disturbance observer, the disturbance is unknown and no preview is available. Thus, the signals must be delayed $d = \rho + \sigma$ samples before the disturbance estimate can be calculated. From the analysis of the disturbance observer transfer functions above, it was noted that delays decreased the performance. As the exponent d increases, the transfer function $\{(1 - z^{-d})G\}$ cannot be approximated as zero, and the approximations given in Equations 5.4-5.6 are no longer valid. Thus, unstable zeros have a very direct impact upon the performance of the technique. Figure 5.6 shows a plot of the unstable, inverted plant

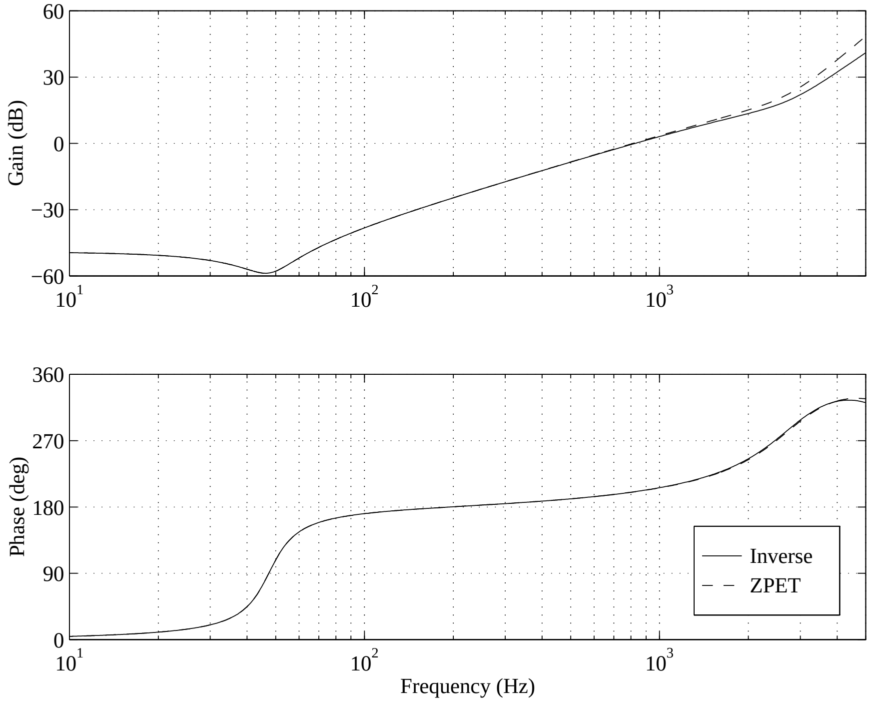


Figure 5.6: Comparison of the Inverted Plant and the ZPET Approximation

model and the ZPET-like approximation that was actually implemented. This approximation holds very well until about 1 kHz.

5.4 Stability of the Disturbance Observer

Bickel and Tomizuka [7] have shown that for a first-order Q filter, the disturbance observer is equivalent to the passivity-based control techniques of Sadegh and Horowitz [56] and Slotine and Li [61]. This is significant in nonlinear systems like robotics for proving stability. Although there are some nonlinear effects in the bearings of the actuator pivot, these effects are relatively small for track following. Thus, the nonlinearities may be considered disturbances to a linear model. Since the disturbance observer is a linear, time-invariant feedback technique, standard linear analysis methods may be used to determine the effect of the disturbance observer on the stability of the system.

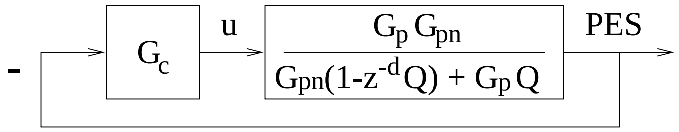


Figure 5.7: Block Diagram of the Disturbance Observer Included in the Plant Transfer Function

Referring to Figure 5.1, there are two ways to consider the disturbance observer. The first is to assume that the disturbance observer is part of a modified plant, such as shown in the block diagram of Figure 5.7 and the frequency response of Figure 5.8. This approach is consistent with the analysis presented in Section 5.1, which describes the effects that the disturbance observer has on the disturbance rejection, robustness to model uncertainties, and sensor noise attenuation. This approach also leads to a simple approach for designing the nominal feedback controller, since the nominal plant may be used in place of the true dynamics.

The second approach is to consider the disturbance observer as part of the total feedback system. See Figure 5.9 for a block diagram of this approach. This outlook may facilitate the comparison of the disturbance observer with other design techniques such as PID (Proportional-Integral-Derivative) control, H_∞ control, and optimal control techniques. Compare the nominal feedback controller in Figure 2.11 with the feedback controller augmented with the disturbance observer in Figure 5.10.

Either approach will give the same open loop transfer function that can be used to analyze the stability of the system, as shown in Figure 5.11 along with the open loop frequency response without the disturbance observer. The gain and phase margins and crossover frequency are virtually unchanged. However, the high-frequency resonances have been attenuated since they were neglected in the nominal plant.

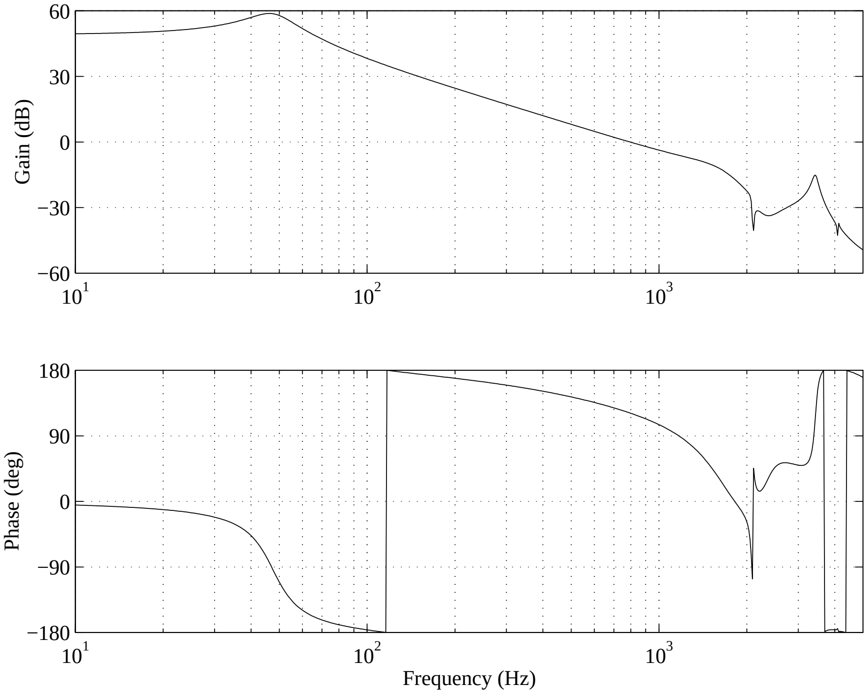


Figure 5.8: Frequency Response of the Disturbance Observer Included in the Plant Transfer Function

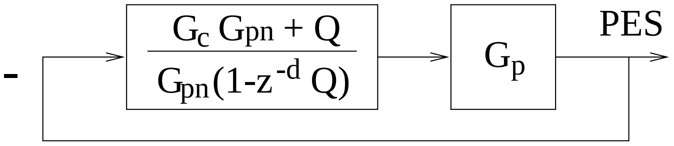


Figure 5.9: Block Diagram of the Disturbance Observer Included in the Controller Transfer Function

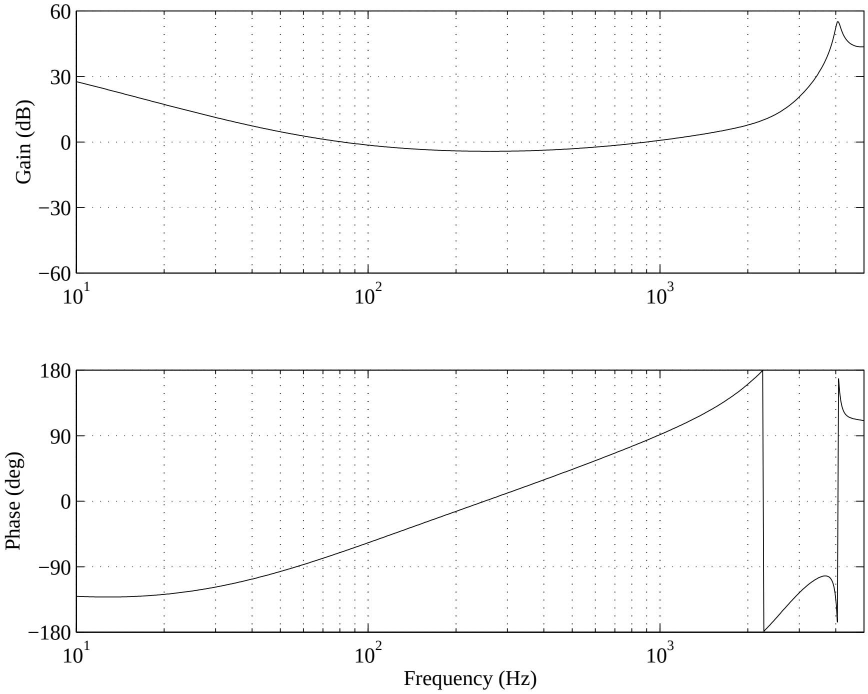


Figure 5.10: Frequency Response of the Disturbance Observer Included in the Controller Transfer Function

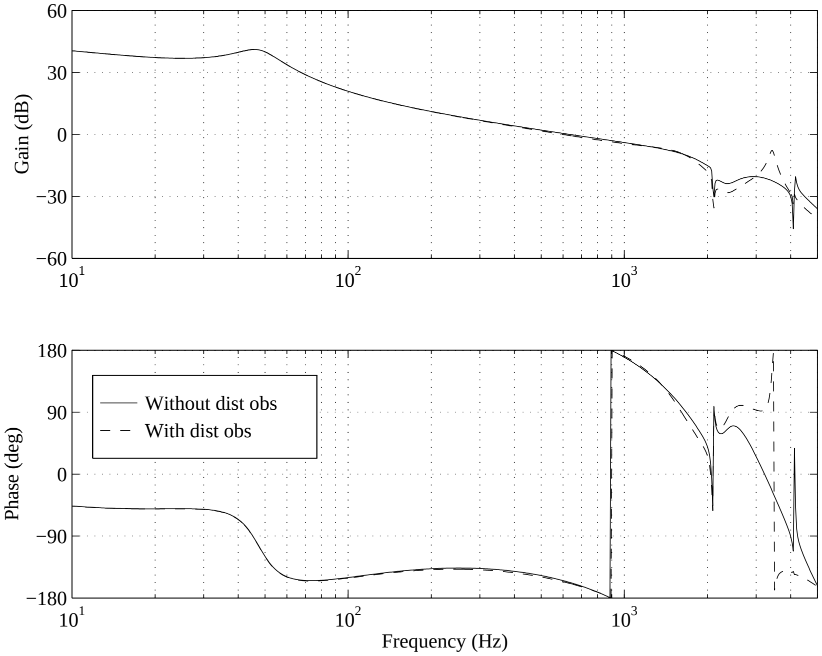


Figure 5.11: Frequency Response of the Open Loop System without and with the Disturbance Observer

Chapter 6

Simulation and Experimental Results of Disturbance Observer Tests

The same techniques were used for the disturbance observer tests that were used for the acceleration feedforward control simulations and experiments. For purposes of comparison, the input levels and frequency ranges were maintained as well.

6.1 Vibration Testing for Track Following: Simulations

6.1.1 Frequency Response Results

By running the simulation program for a number of different sinusoidal frequencies, the magnitude of the PES generated by the base disturbances can be determined. These data points are shown in Figure 6.1. This figure should be compared to the results from the acceleration feedforward control in Figures 4.1 and 4.2. In the higher frequency range, above 500 Hz, there is very little improvement. In the middle frequency range, the disturbance

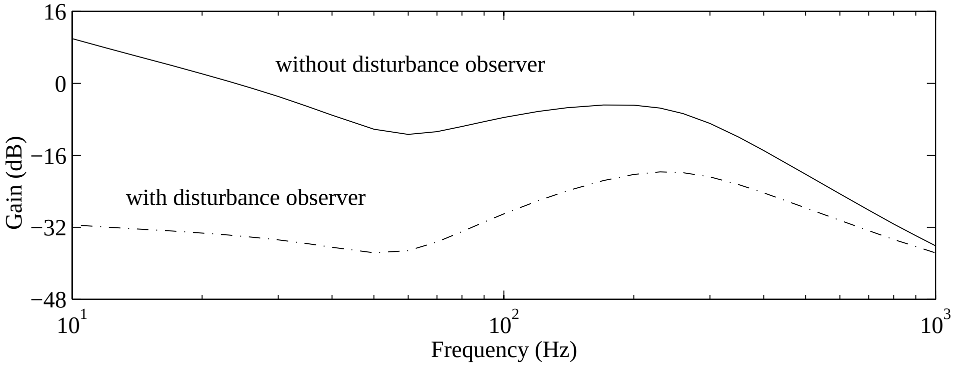


Figure 6.1: Simulated Frequency Response from the Base Acceleration to the PES without and with the Disturbance Observer

observer performs approximately as well as the fixed-parameter feedforward controller, but not as well as the adaptive feedforward controller. The reductions in the magnitude of the PES due to the disturbance observer compared to the case with only feedback control are approximately 25 dB to 7 dB (94% to 55%) between 50 Hz and 500 Hz. In the low-frequency range, the disturbance observer performs significantly better than either of the feedforward controller designs. Reductions in the PES are from 40 dB to 25 dB (99% to 94%) between 10 Hz and 50 Hz.

This behavior is to be expected. For low frequencies, the delay involved with the disturbance observer is less significant, so the approximations in Equations 5.4-5.6 hold well. At higher frequencies, the delays are more significant, and the effects of the low-pass Q filter start to appear.

6.1.2 Time Response Results

Time traces of the simulations without and with the disturbance observer are given in Figures 6.2-6.9 for the same frequencies that were investigated for the feedforward controllers in Section 4.1. The PES amplitudes for the various disturbance frequencies are given in Table 6.1. These results are consistent with the trends of the frequency response results, showing excellent performance in the low-frequency range that diminishes as the

	Without Dist Obs		With Dist Obs		
Frequency (Hz)	Figure	PES (μm)	Figure	PES (μm)	Reduction (%)
40	6.2	1.93	6.3	0.07	96
60	6.4	2.39	6.5	0.12	95
100	6.6	3.19	6.7	0.28	91
200	6.8	3.63	6.9	0.62	83

Table 6.1: Results of the Disturbance Observer Vibration Simulations

frequency of the disturbance is increased.

A random input was also used to test the performance of the disturbance observer. The PSD of the base acceleration and the PES without and with the disturbance observer are shown in Figures 6.10 and 6.11, respectively. Representative time responses of the PES without and with the disturbance observer for a random disturbance input are shown in Figures 6.12 and 6.13. The case without the disturbance observer had an RMS value of the base acceleration of 0.48 G and an RMS value of the PES of 1.16 μm. With the disturbance observer, the RMS base acceleration was 0.56 G and the RMS PES was 0.19 μm. Normalized for the corresponding base accelerations, the disturbance observer resulted in a reduction of the RMS value of the PES of 86%.

6.2 Vibration Testing for Track Following: Experiments

6.2.1 Frequency Response Results

Implementation of the disturbance observer on the experimental system exposed some of the problems associated with the technique. As mentioned in Chapter 5, the disturbance observer relies on a clear distinction between the frequency range where the disturbance is expected and the frequency range where the sensor noise is expected. Unfortunately, the noise on the PES overlapped the frequency range where disturbance rejection was desired. Because of this, the frequency of the Q filter had to be decreased significantly,

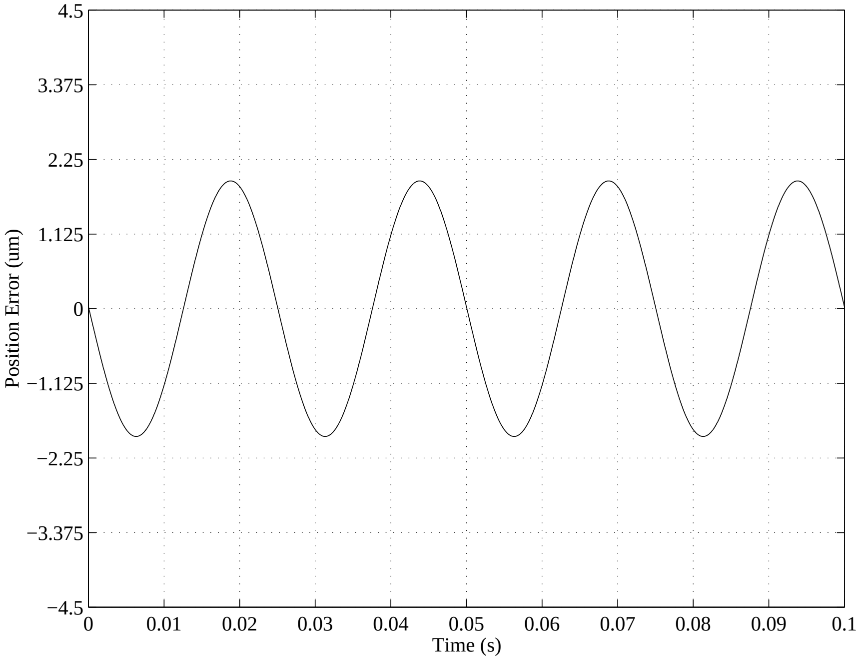


Figure 6.2: Simulation Time Trace without the Disturbance Observer at 40 Hz

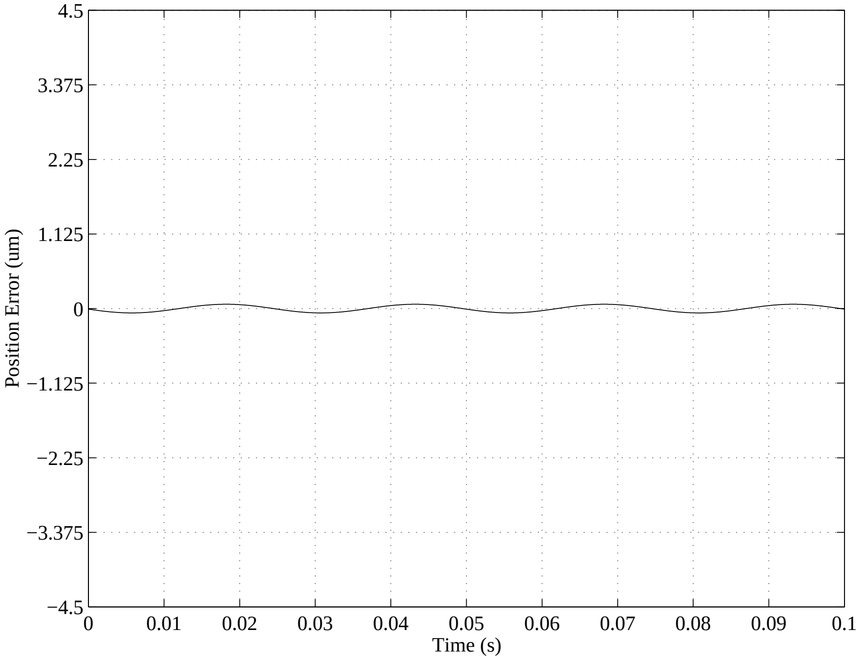


Figure 6.3: Simulation Time Trace with the Disturbance Observer at 40 Hz

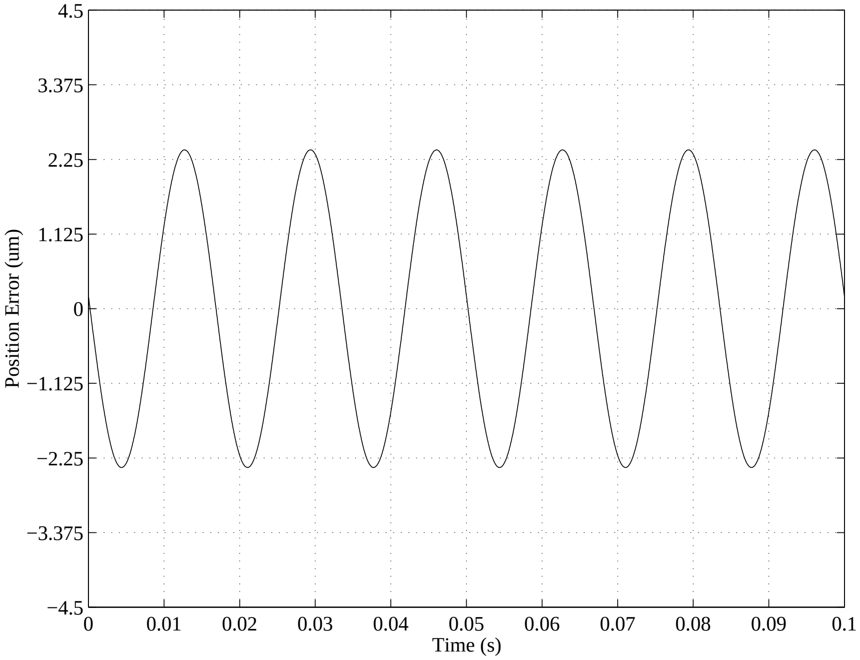


Figure 6.4: Simulation Time Trace without the Disturbance Observer at 60 Hz

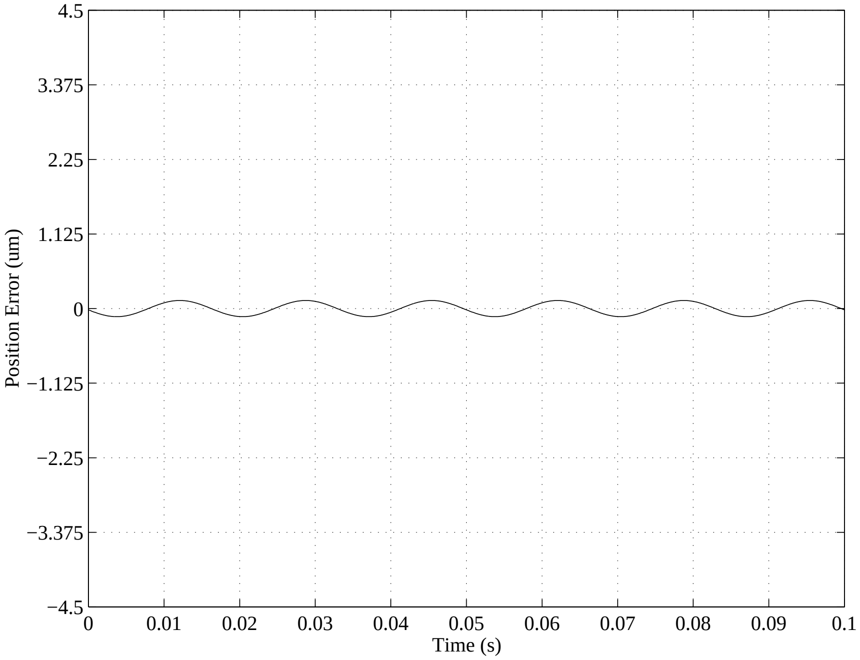


Figure 6.5: Simulation Time Trace with the Disturbance Observer at 60 Hz

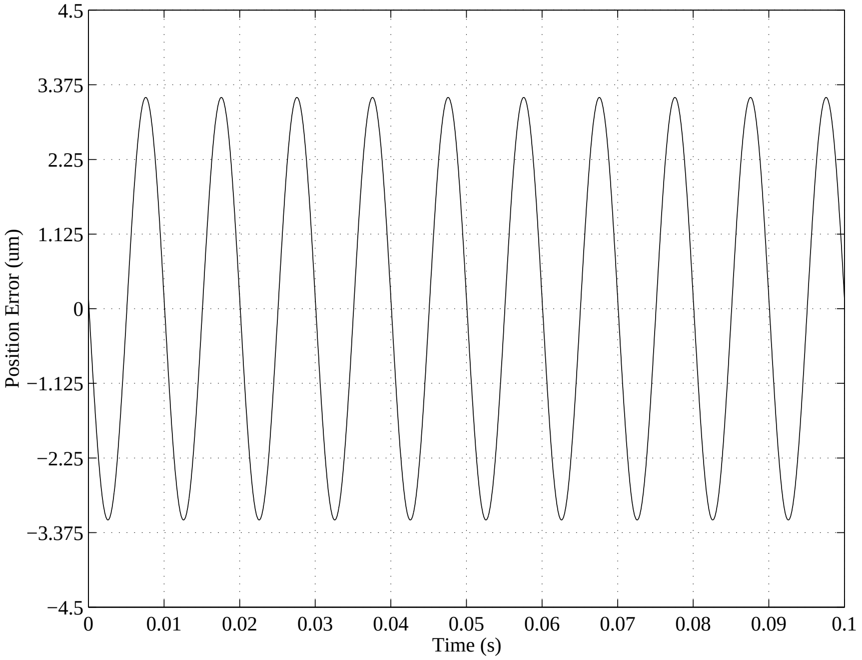


Figure 6.6: Simulation Time Trace without the Disturbance Observer at 100 Hz

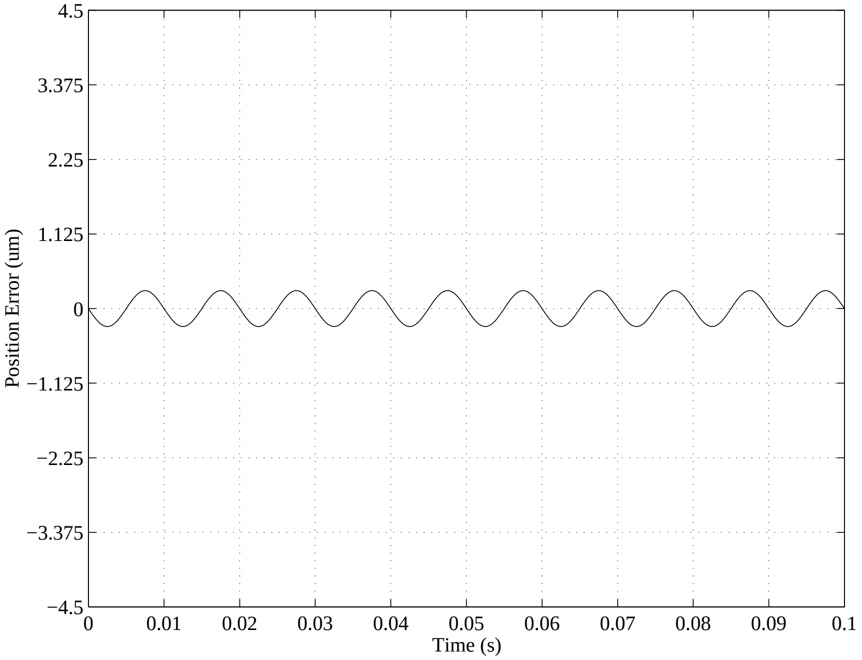


Figure 6.7: Simulation Time Trace with the Disturbance Observer at 100 Hz

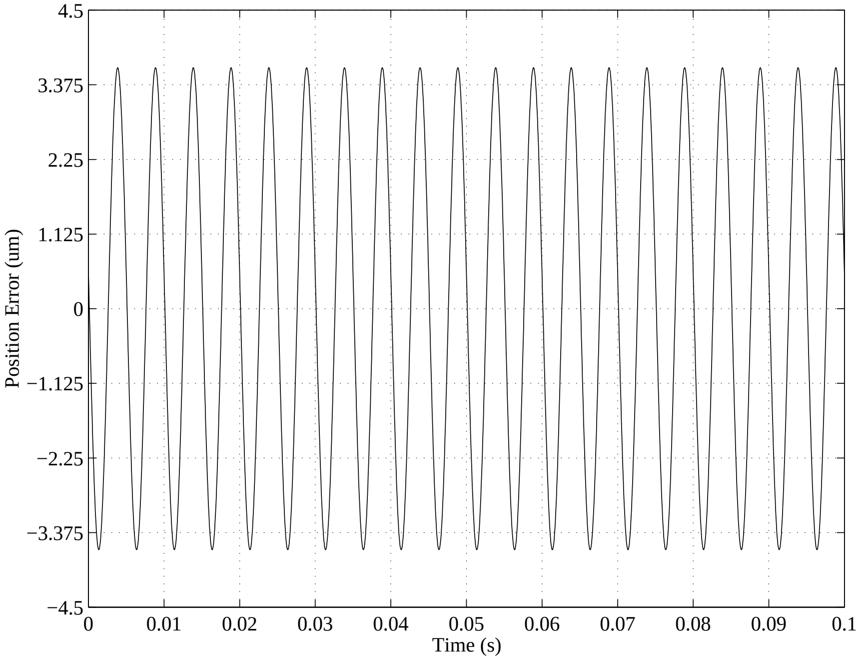


Figure 6.8: Simulation Time Trace without the Disturbance Observer at 200 Hz

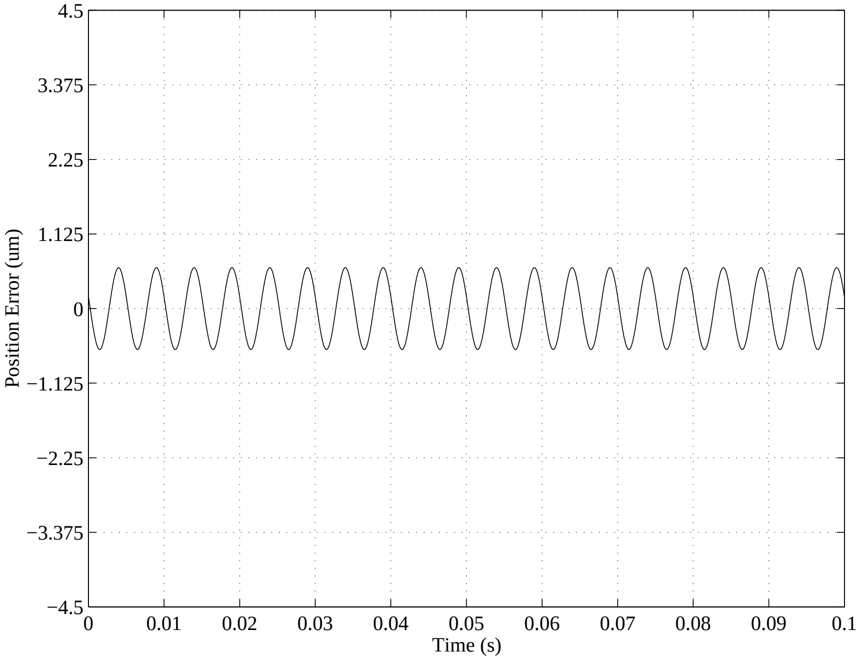


Figure 6.9: Simulation Time Trace with the Disturbance Observer at 200 Hz

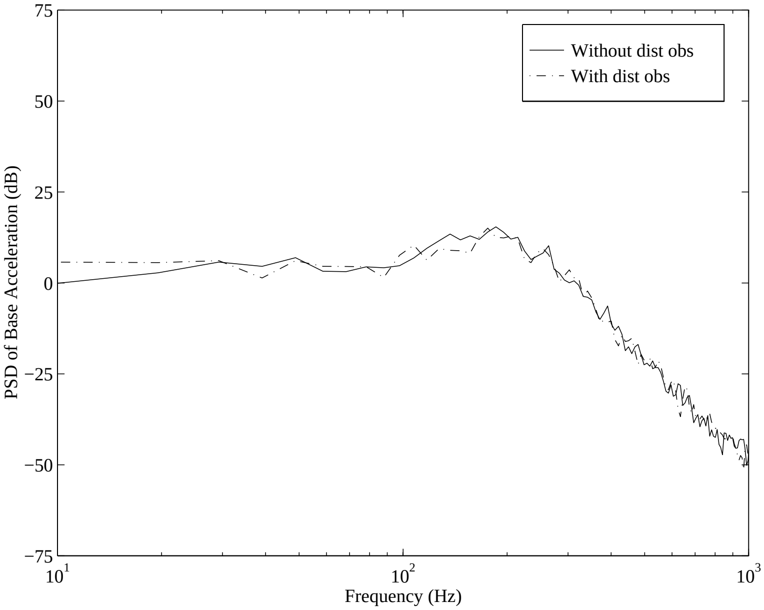


Figure 6.10: Power Spectral Densities of the Base Accelerations

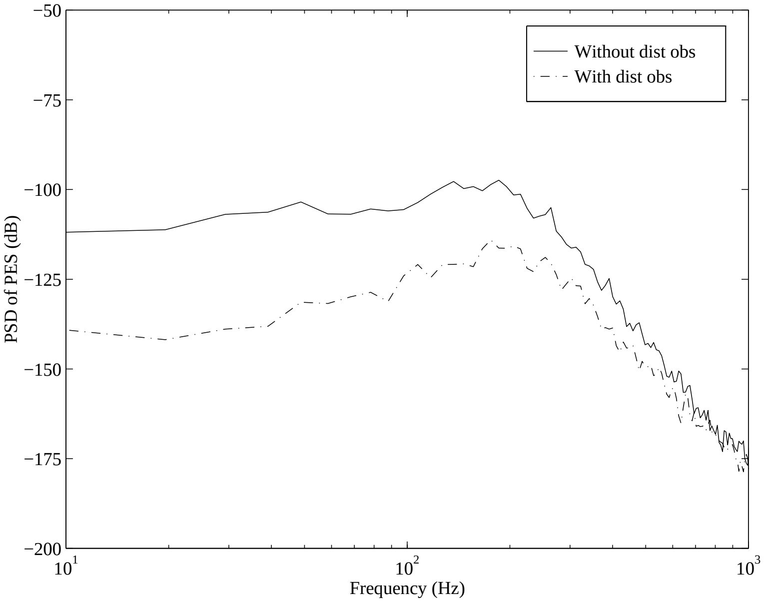


Figure 6.11: Power Spectral Densities of the PES without and with the Disturbance Observer

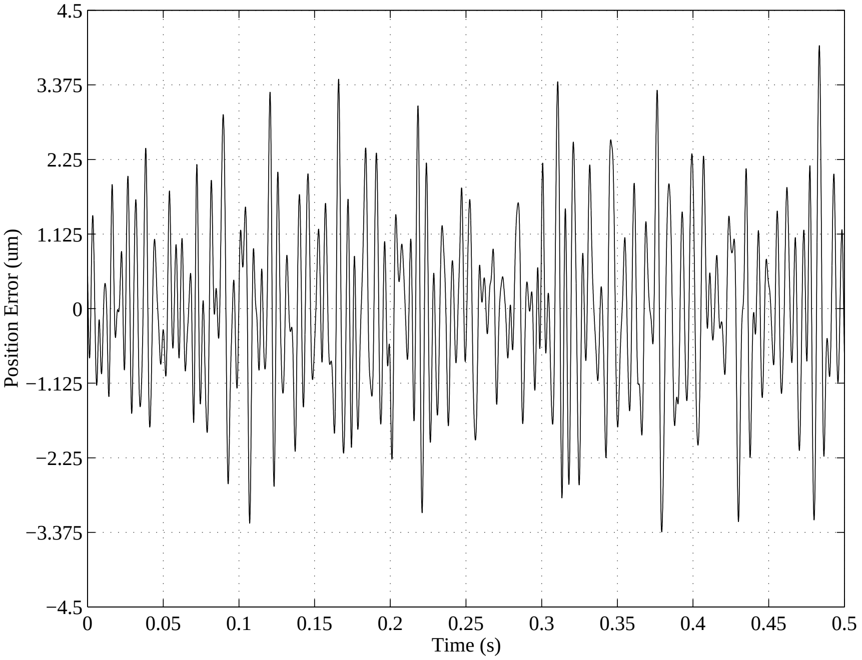


Figure 6.12: Simulation Time Trace without the Disturbance Observer for a Random Base Acceleration

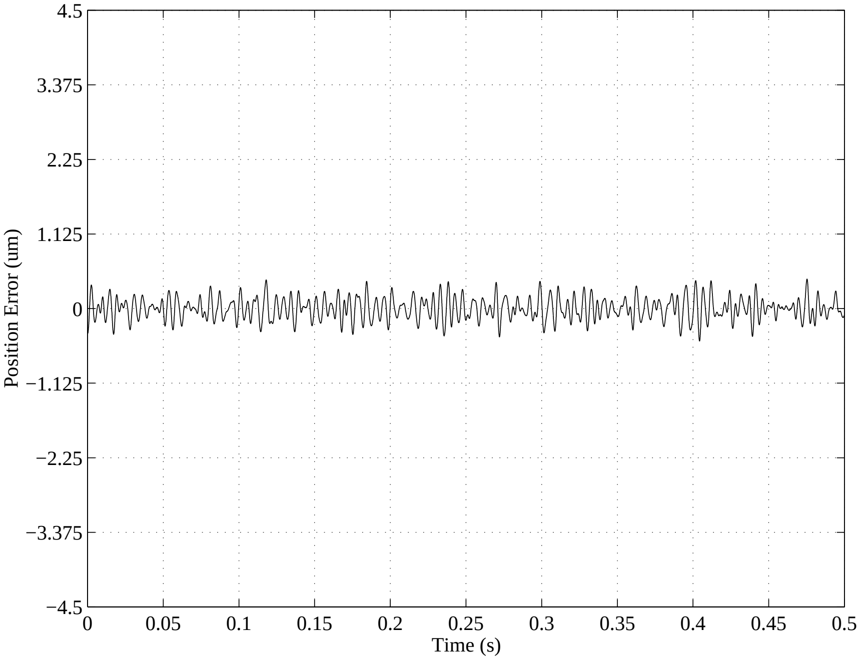


Figure 6.13: Simulation Time Trace with the Disturbance Observer for a Random Base Acceleration

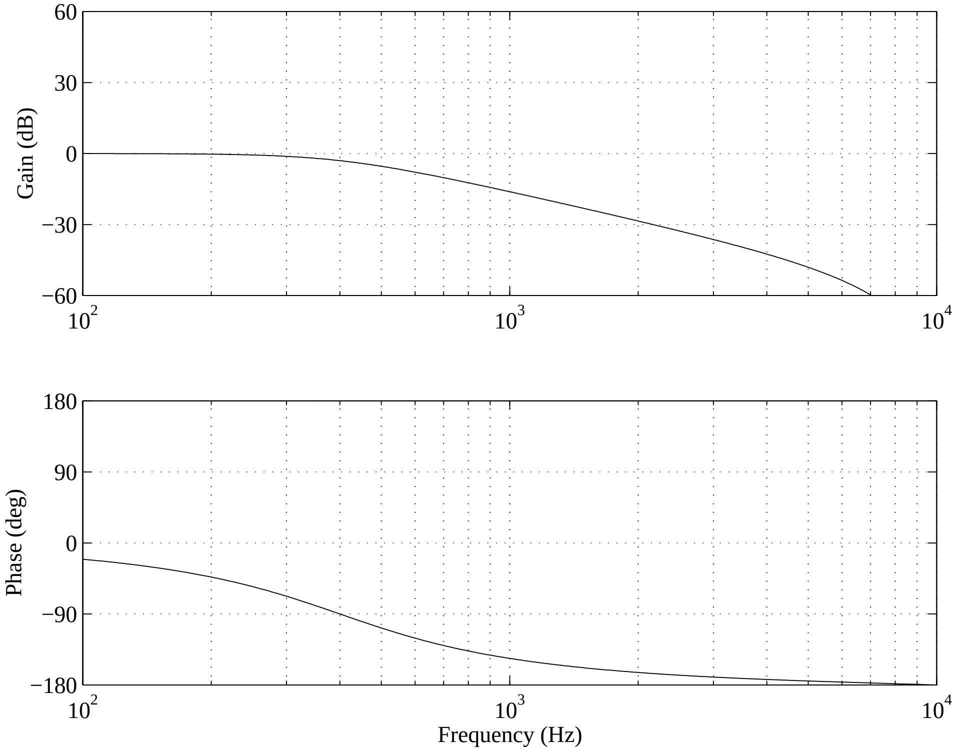


Figure 6.14: Frequency Response of the Redesigned Q Filter for Use in the Experiments

from the 5 kHz used in the simulations in Section 6.1 to 400 Hz, as shown in Figure 6.14. This reduced the disturbance rejection properties. Sugiura and Hori ran into similar problems in their investigation of steel rolling mills [63].

Figure 6.15 shows the experimental frequency response from a disturbance at the input of G_p (as opposed to a disturbance to the base) to the PES without and with the disturbance observer for the redesigned Q filter. The trends of the frequency response are the same as described for the simulations. The disturbance observer does a good job in the low frequency range, with reductions in the PES from 22 dB to 8 dB (96% to 60%) below 100 Hz. Performance in the middle-frequency range is degraded, with the disturbance observer actually increasing the PES slightly between 200 Hz and 1 kHz. Since the phase

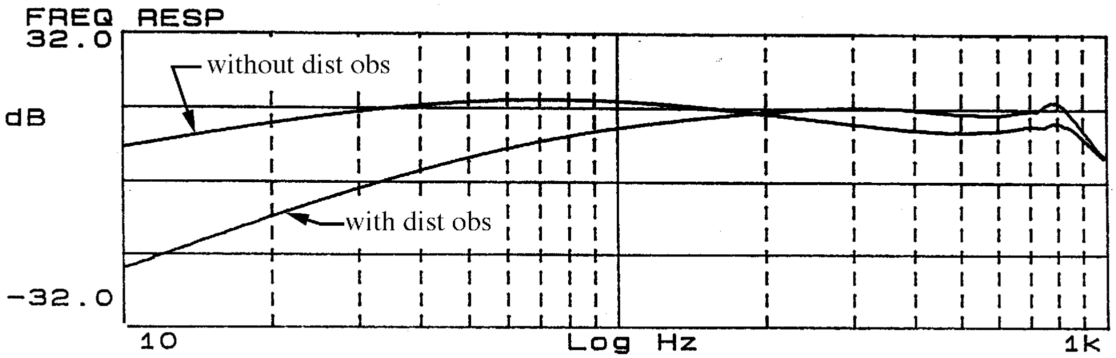


Figure 6.15: Experimental Frequency Response from the Input Disturbance to the PES without and with the Disturbance Observer

of the Q filter is decreasing from $0°$ to $-180°$, it is actually adding to the input disturbance instead of subtracting from it. This effect disappears as the magnitude of the Q filter decreases. Above 1 kHz, there is little difference in the disturbance rejection with and without the disturbance observer.

6.2.2 Time Response Results

Experimental time traces of the PES for disturbances to the base without and with the disturbance observer are given in Figures 6.16-6.23. These should be compared to the simulation results in Section 6.1 and the feedforward results in Chapter 4. Note that Figures 6.16-6.23 are plotted on the longer time axis that was required to show the convergence of the adaptive algorithm in Figures 4.7-4.10, while the simulation results of Figures 6.2-6.9 used the shorter time axis of the IIR feedforward controller results in Figures 4.3-4.6.

At 40 Hz and 60 Hz the disturbance observer is able to decrease the PES significantly, by 16 dB and 10 dB respectively. At 100 Hz, the benefits of the disturbance observer are less evident, with a reduction of 8 dB. At 200 Hz, there is virtually no difference between the PES without and with the disturbance observer. The PES amplitudes for the various disturbance frequencies are given in Table 6.2.

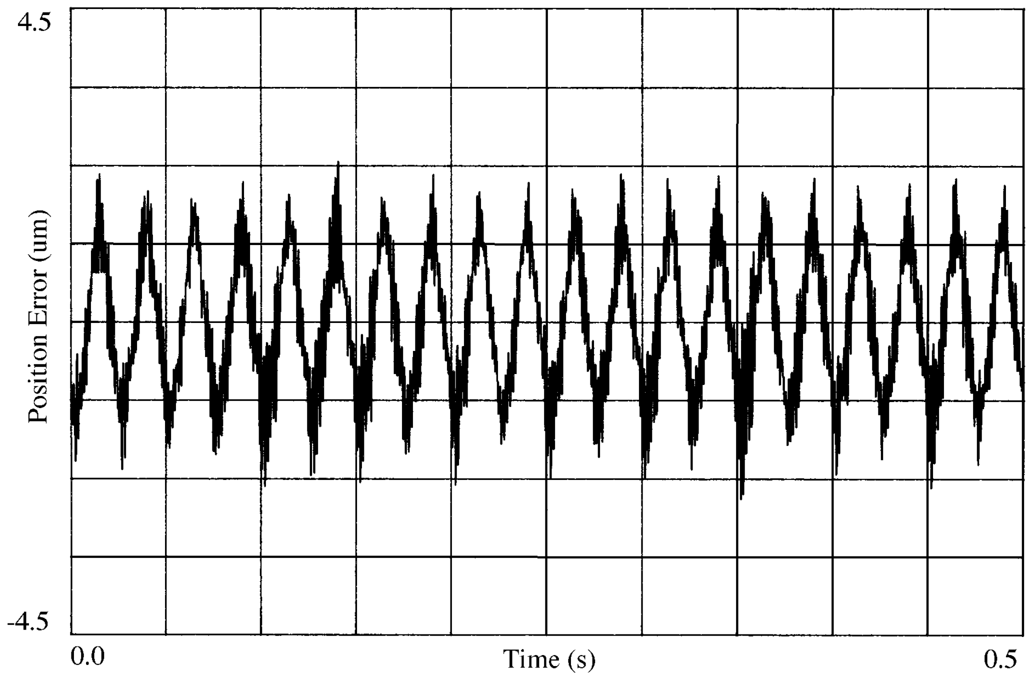


Figure 6.16: Experimental Time Trace without the Disturbance Observer at 40 Hz

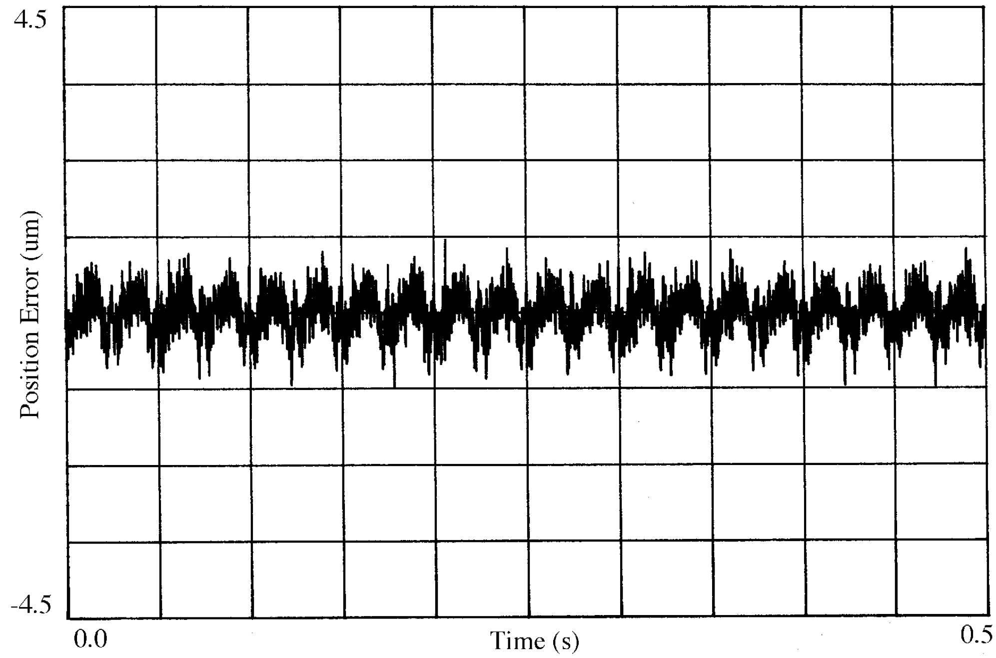


Figure 6.17: Experimental Time Trace with the Disturbance Observer at 40 Hz

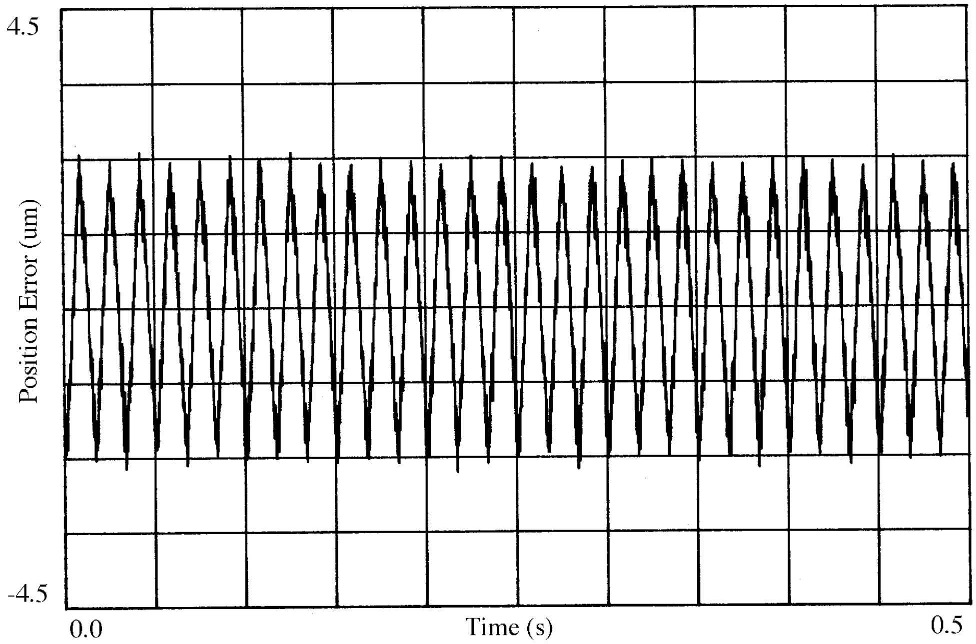


Figure 6.18: Experimental Time Trace without the Disturbance Observer at 60 Hz

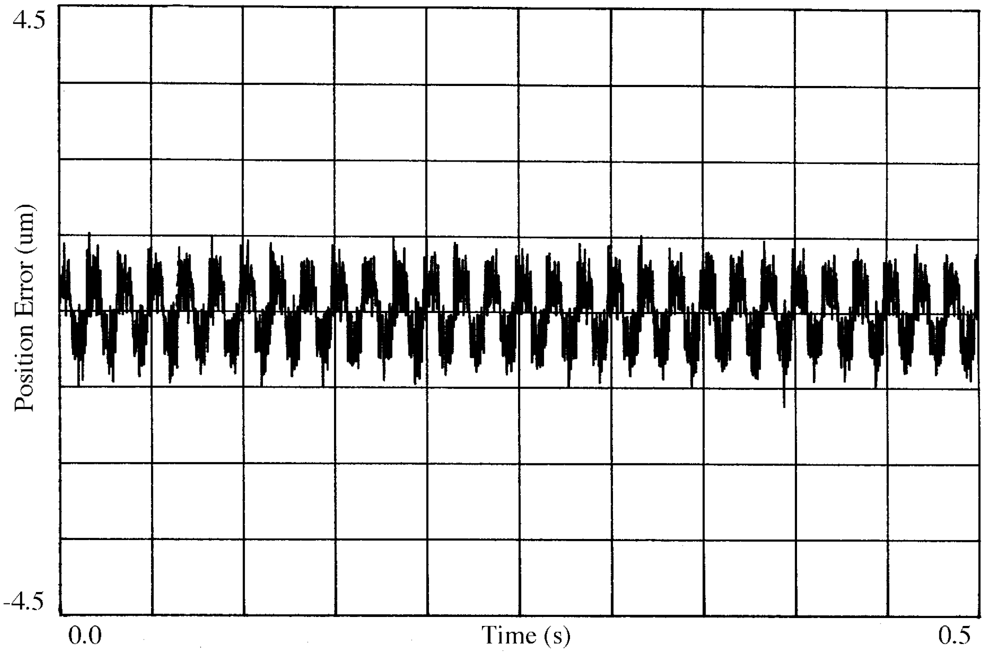


Figure 6.19: Experimental Time Trace with the Disturbance Observer at 60 Hz

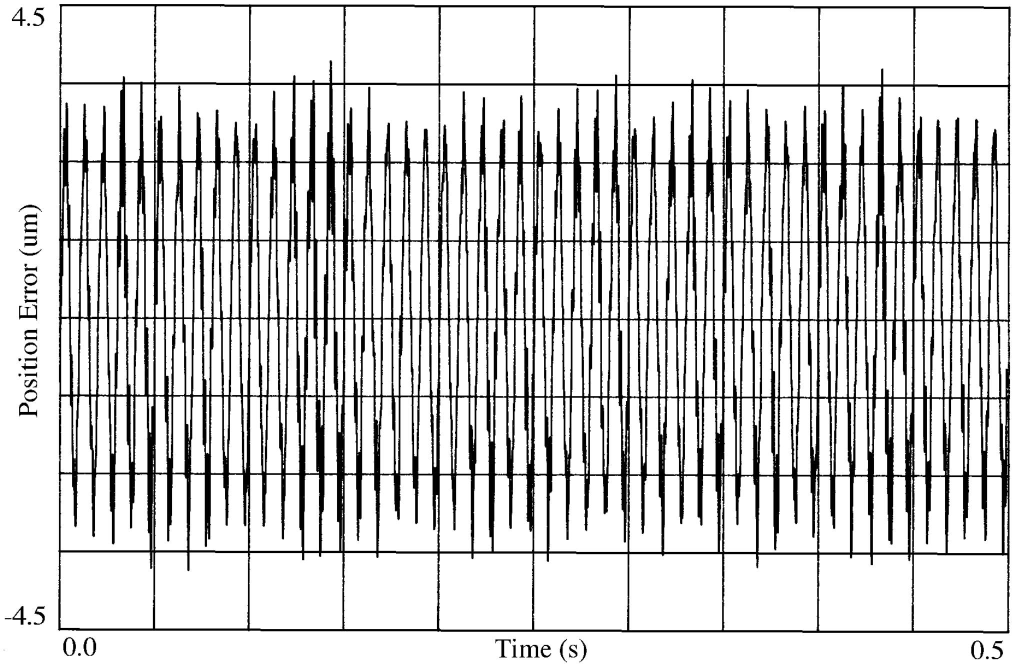


Figure 6.20: Experimental Time Trace without the Disturbance Observer at 100 Hz

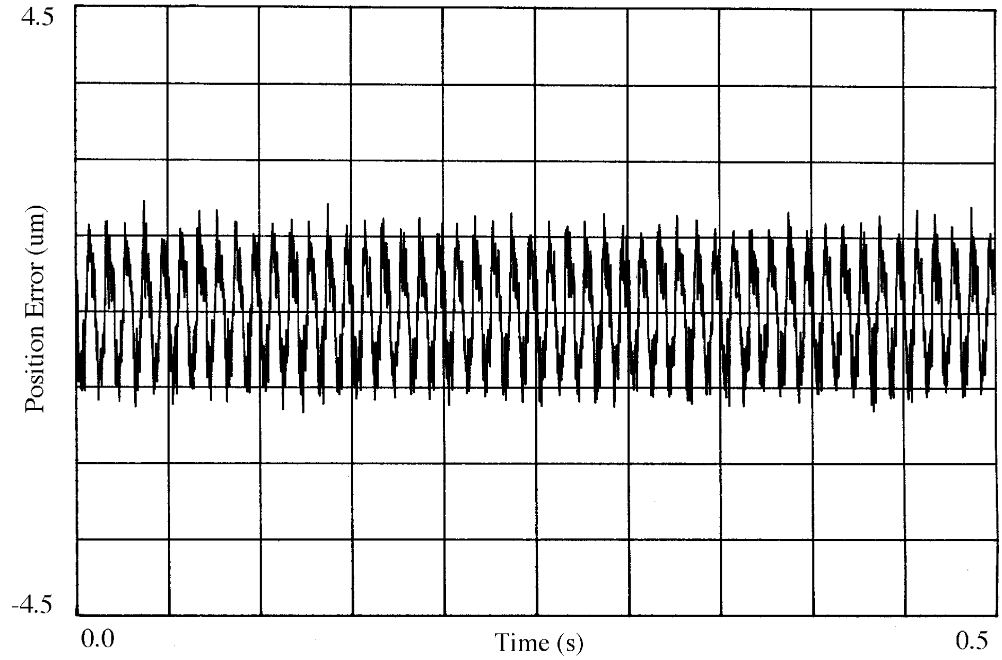


Figure 6.21: Experimental Time Trace with the Disturbance Observer at 100 Hz

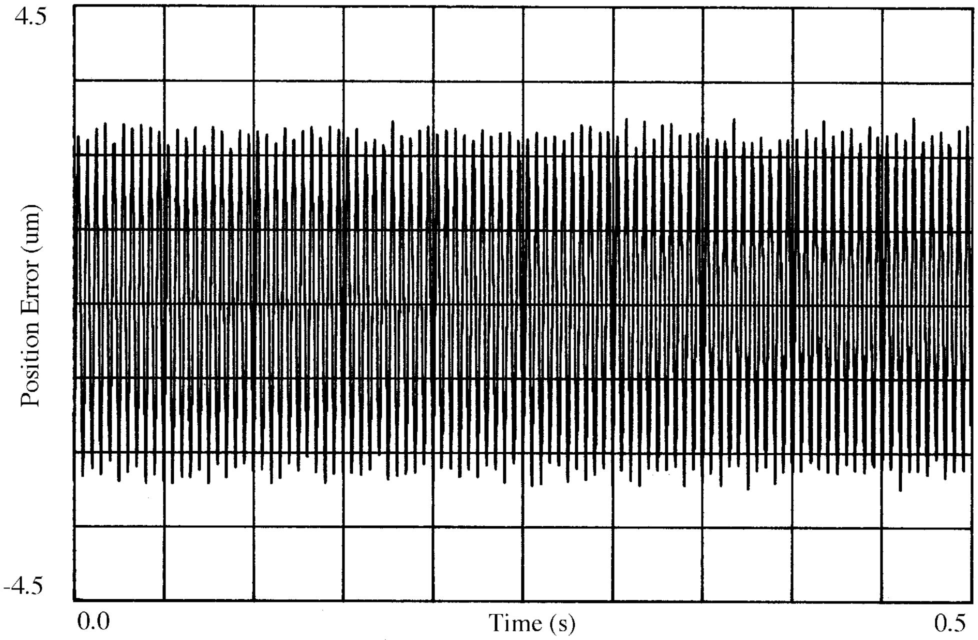


Figure 6.22: Experimental Time Trace without the Disturbance Observer at 200 Hz

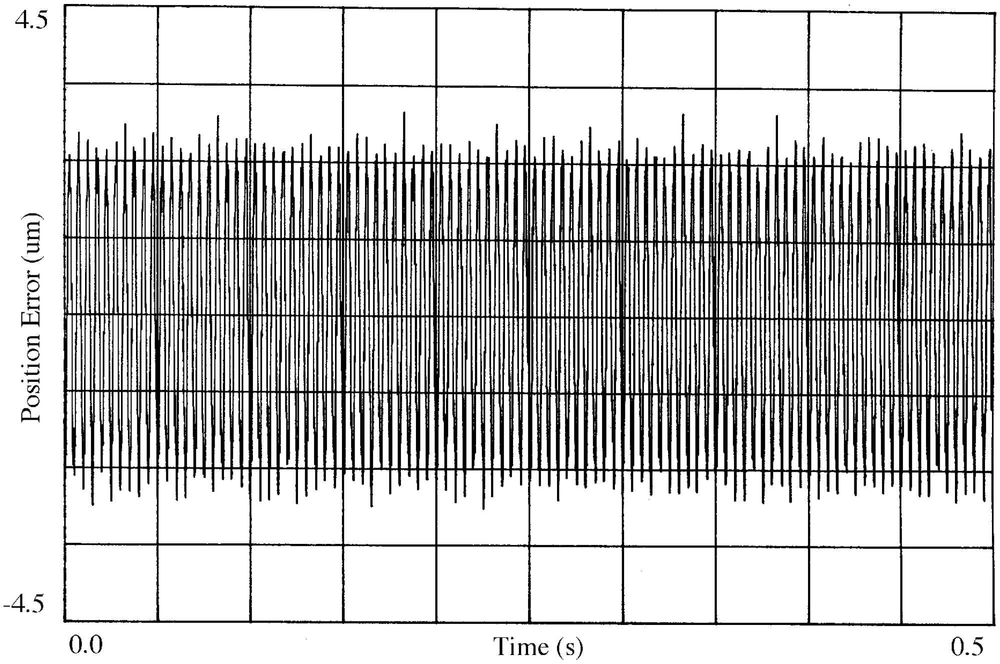


Figure 6.23: Experimental Time Trace with the Disturbance Observer at 200 Hz

	Without Dist Obs		With Dist Obs		
Frequency (Hz)	Figure	PES (μm)	Figure	PES (μm)	Reduction (%)
40	6.16	1.8	6.17	0.3	83
60	6.18	1.9	6.19	0.6	68
100	6.20	2.6	6.21	1.0	61
200	6.22	2.3	6.23	2.2	4

Table 6.2: Results of the Disturbance Observer Vibration Experiments

Implementation of a disturbance observer typically imposes a very light computational burden. The effectiveness of the disturbance observer at low frequencies, coupled with the good results of the filtered-x LMS acceleration feedforward technique in the middle-frequency range, suggest that it may be possible to combine these two techniques to improve performance over a wider range of frequencies.

6.3 Shock Testing for Track Following

As in the case for the feedforward controller, the effectiveness of the disturbance observer for reducing the position error due to a shock disturbance was examined. Again, a shock consisting of a half sine wave of 10 ms duration and an amplitude of 2 G was used. Simulation results of the time responses are shown in Figures 6.24 and 6.25 for the cases without and with the disturbance observer, respectively. The disturbance observer does an excellent job of canceling the disturbance. The peak value was reduced from a magnitude of 4.8 μm to 0.26 μm, a change of 95%.

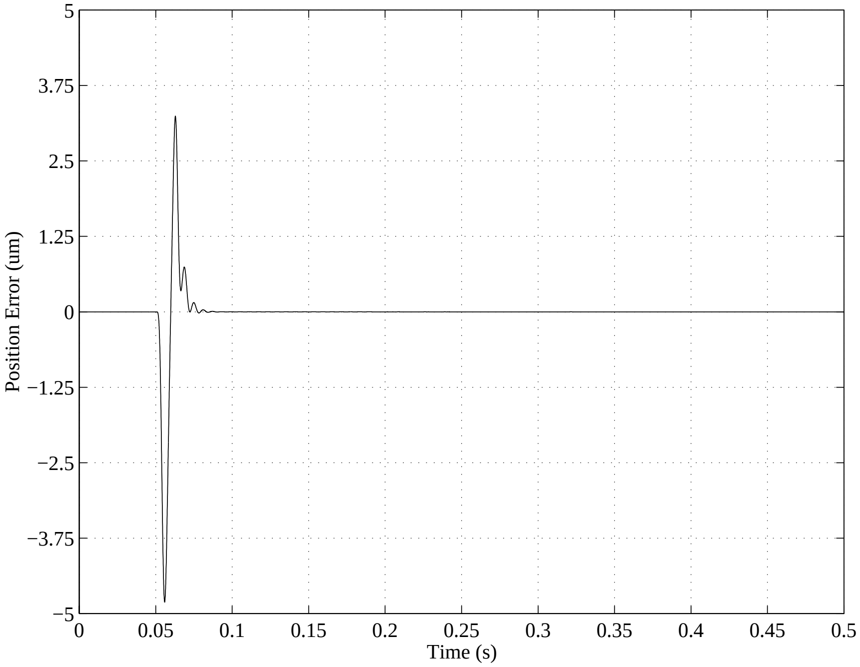


Figure 6.24: Simulation Time Trace without the Disturbance Observer for a Shock Disturbance

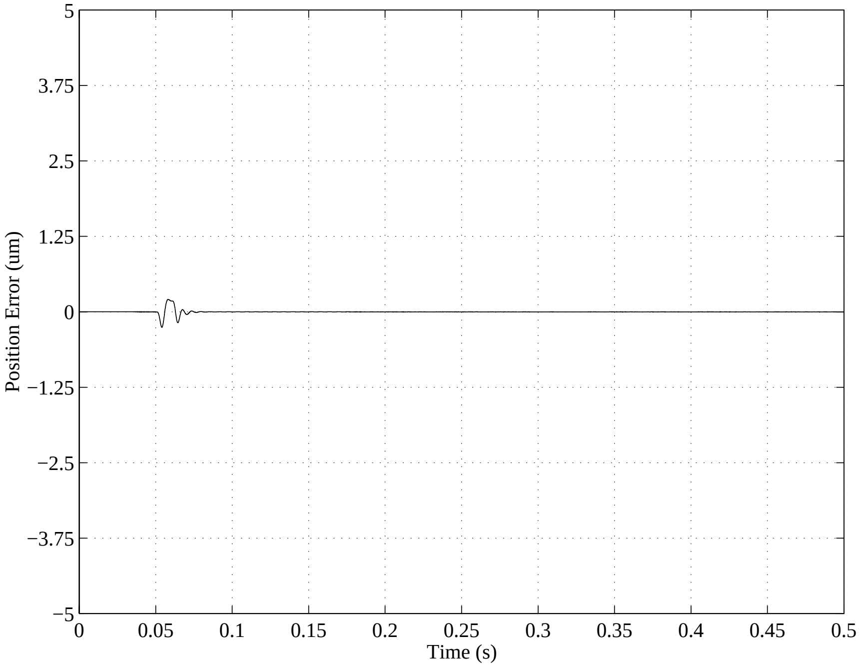


Figure 6.25: Simulation Time Trace with the Disturbance Observer for a Shock Disturbance

Chapter 7

Conclusions

7.1 Summary of Results

Disk drive servo control is a very active area of research, and some of the previous techniques that have been employed to reduce tracking errors were presented in the Introduction. The magnetic disk drive system was explained, as well as how disturbances affect disk drive performance. Two control methods to reduce these effects were described, acceleration feedforward control and the disturbance observer. Both techniques require some knowledge of the system model, so a theoretical derivation of the system equations was presented in Chapter 2. Also included were an experimental verification of the analytical model and a description of the feedback controller that was used for both control methods.

7.1.1 Acceleration Feedforward Control

The first control technique used accelerometers to measure the acceleration of the drive and a feedforward filter to send a canceling signal to the drive actuator. Two different implementations of the feedforward filter were used. The first was a fixed-parameter IIR filter. The second was an FIR adaptive filter using the filtered-x LMS adaptive algorithm. These techniques were implemented experimentally using a digital signal processor to calculate the control, and a shake table to provide the disturbance. Results demonstrated that acceleration feedforward was effective in reducing the position error of the drive actuator due to disturbances. The PES was reduced from 50% to 90% over the frequency range from 40 Hz to 400 Hz with the fixed-parameter filter and vibration disturbances.

Simulation and experimental results showed that performance of the feedforward controller was heavily dependent upon the controller parameters. This motivated the use of an adaptation algorithm to adjust the feedforward controller parameter filters. Vibration test results from the adaptive technique showed a reduction of the PES of 55% to 95% over the frequency range of 15 Hz to 400 Hz. Tests were performed with single sinusoids and random inputs. A comparison between the two feedforward controllers showed that the adaptive controller gave better results, and can adapt to variations in the drive parameters, but at the cost of increased computation.

Shock testing was performed in simulation using an industry-standard half sine wave input to the base. The feedforward controller was found to be effective in rejecting this type of disturbance as well. With the feedforward controller, the peak error was reduced by 89%.

7.1.2 Disturbance Observer

Although the acceleration feedforward controllers were extremely effective in reducing the effects of disturbances, they do require the use of an additional sensor. Despite the level of technology involved in producing a disk drive, the market price is actually quite low. Even low-cost sensors such as the ones used in the experiments may be too expensive for some applications. In this case, it is preferable to employ a technique that does not require any sensors. The disturbance observer does this by creating an estimate of the disturbance using the PES, the control input, and a model of the plant. Because the disturbance observer requires the disturbance effects to appear in the PES before they can be compensated for, it cannot be expected to react as quickly to disturbances as the feedforward method.

The same tests were conducted for the disturbance observer that were used for the acceleration feedforward controller. Vibration simulation results using the disturbance observer were comparable in the frequency range between 50 Hz and 500 Hz, with reductions in the PES of 55% to 94%. The disturbance observer did a much better job in the low-frequency range, decreasing the PES by 94% to 99% between 10 Hz and 50 Hz. Neither technique offered much performance improvement above 500 Hz. Because of noise in the system, the break frequency of the Q filter had to be set lower in the experimental tests than the value that was used in the simulations. This decreased its effectiveness. It was still able to perform very well in the low-frequency range, but it was not able to reject disturbances above 200 Hz. Rejection of shock disturbances using the disturbance observer was similar to the performance of the feedforward controller. Simulation showed a peak position error reduction of 95% with the disturbance observer.

7.2 Future Work

Although the disturbance rejection techniques have been shown to be effective in experimentation, there are several remaining practical questions to be addressed. As discussed previously, there is often very little spare computation time available in disk drives. The filtered-x LMS algorithm was chosen to minimize computation time. However, it would be preferable to run the algorithm only when the drive first boots and then again only when the dynamics of the drive change significantly. These changes in dynamics may be due to thermal effects, or even the position of the actuator arm. If the dynamics of the drive are found to change quickly or often, the use of multi-rate techniques may be useful to update the parameters at a slower sampling time than the feedback controller. This would reduce computations while maintaining the benefits of the adaptation at all times.

In conjunction with the decision of when to implement the adaptation algorithm is the question of input generation for parameter identification. Is it possible to generate a sufficiently rich signal from arm motion only, or is it necessary to include some sort of noise generation electronics in the drive? Certainly with the high accuracy required in disk drives, it would be impractical to introduce noise into the system while the system was trying to read or write.

The implementations of the algorithms were made on a 32-bit floating point digital signal processor (DSP) with 10-bit analog-to-digital (ADC) and digital-to-analog (DAC) converters. Typically, disk drives currently use fixed-point DSPs, and may have low precision on the DSPs and the ADCs and DACs. These effects should be investigated further. The filtered-x LMS algorithm has been implemented successfully for acceleration feedforward control on a compact disk player with a 16-bit fixed-point DSP with 10-bit ADCs and DACs [33].

The simulation and experimental results considered only rotational disturbances in the plane of the disks. Thus, coupling of the dynamics and disturbances in other coordinates were not considered. Obviously, actual implementations will encounter a variety of disturbance types. Compensating for this may require additional sensors and/or a more careful consideration of their placement.

Theoretically, both the acceleration feedforward controller and the disturbance observer can be employed in conjunction with a track seeking controller. In this case, however, the dynamics of the drive may become more significant, especially with the settling phase. This appears to be a very interesting area for research. Ishikawa and Tomizuka [30] have some preliminary results using a disturbance observer designed with Kalman filter theory. The VCM current and measured acceleration of the actuator arm are compared to compensate for pivot friction in a disk drive.

Bibliography

[1] K.J. Åström, P. Hagander, and J. Sternby. Zeros of sampled systems. *Automatica*, 20(1):31–38, 1984.

[2] K.J. Åström and B. Wittenmark. *Adaptive Control.* Addison-Wesley, Menlo Park, CA, 1995.

[3] D.Y. Abramovitch. Rejecting rotational disturbances on small disk drives using rotational accelerometers. In *Proceedings of the 13th World Congress of IFAC*, volume O, pages 483–488, San Francisco, July 1996.

[4] K. Aruga, Y. Mizoshita, M. Iwatsubo, and T. Hatagami. Acceleration feedforward control for head positioning in magnetic disk drives. *JSME International Journal*, 33(1):35–41, 1990. series III.

[5] A. Benveniste and G. Ruget. A measure of the tracking capability of recursive stochastic algorithms with constant gains. *IEEE Transactions on Automatic Control*, AC-27(3):639–649, June 1982.

[6] R. Bickel and M. Tomizuka. Disturbance observer based hybrid impedance control. In *Proceedings of the American Control Conference*, volume 1, pages 729–733, Seattle, June 1995.

[7] R. Bickel and M. Tomizuka. Passivity based versus disturbance observer based robust control: Equivalence and stability. In *Proceedings of the 13th World Congress of IFAC*, volume A, pages 223–228, San Francisco, July 1996.

[8] B.E. Boling, M.A. Marra, B.L. Walcott, and S.A. Williams. Fuzzy logic control of voice coil motor hard disk drives. In *Proceedings of the 13th World Congress of IFAC*, volume C, pages 55–60, San Francisco, July 1996.

[9] G. Bouchard, L. Lau, and F.E. Talke. An investigation of non-repeatable spindle runout. *IEEE Transactions on Magnetics*, MAG-23(5):3687–3689, September 1987.

[10] P. Cheung, R. Horowitz, and R. Howe. Design, fabrication, position sensing, and control of an electrostatically-driven polysilicon microactuator. In *Proceedings of the ASME IMECE*, San Francisco, November 1995.

[11] K.K. Chew and M. Tomizuka. Digital control of repetitive errors in disk drive systems. In *Proceedings of the American Control Conference*, volume 1, pages 540–548, Pittsburgh, June 1989.

[12] W.-W. Chiang. Multirate state-space digital controller for sector servo systems. In *Proceedings of the 29th IEEE Conference on Decision and Control*, pages 1902–1907, Honolulu, HI, December 1990.

[13] D.B. Davies and M.D. Sidman. Active compensation of shock, vibration, and wind-up in disk drives. In *Advances in Information Storage Systems*, volume 5, pages 5–20. ASME Press, 1993. B. Bhushan, editor.

[14] S.J. Elliott, I.M. Stothers, and P.A. Nelson. A multiple error LMS algorithm and its application to the active control of sound and vibration. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-35(10):1423–1434, October 1987.

[15] S. Endo, M. Tomizuka, and Y. Hori. Robust digital tracking controller design for high-speed positioning systems. In *Proceedings of the American Control Conference*, pages 2494–2498, San Francisco, June 1993.

[16] P. Eykhoff. *System Identification: Parameter and State Estimation.* Wiley, New York, 1974.

[17] D.C. Farden. Tracking properties of adaptive signal processing algorithms. *IEEE Transactions on Acoustics, Speech, and Signal Processing*, ASSP-29(3):439–446, June 1981.

[18] G.F Franklin, J.D. Powell, and M.L. Workman. *Digital Control of Dynamic Systems.* Addison-Wesley, Menlo Park, CA, 2nd edition, 1992.

[19] G. Frees. Disk drive spindle dynamics - analysis and measurement. In *Advances in Information Storage Systems*, volume 6. ASME Press, 1994. B. Bhushan, editor.

[20] C.R. Fuller and A.H. von Flotow. Active control of sounds and vibration. *IEEE Control Systems*, pages 9–19, December 1995.

[21] G.C. Goodwin and K.S. Sin. *Adaptive Filtering Prediction and Control.* Prentice Hall, Englewood Cliffs, NJ, 1984.

[22] H. Hanselmann and A. Engelke. LQG-control of a highly resonant disk drive head positioning actuator. *IEEE Transactions on Industrial Electronics*, 35(1):100–104, February 1988.

[23] J.C. Harrison and F.E. Talke. Non-repeatable runout of cantilever and doubly supported $5\frac{1}{4}$" disc drive spindles. *Precision Engineering*, 13(1):33–40, January 1991.

[24] M. Hirata, K.-Z. Liu, T. Mita, and T. Yamaguchi. Head positioning control of a hard disk drive using H-infinity control theory. *SICE*, 29(1):71–77, 1993. In Japanese.

[25] Y. Hori. A review of torsional vibration control methods and a proposal of disturbance observer-based new techniques. In *Proceedings of the 13th World Congress of IFAC*, volume C, pages 7–12, San Francisco, July 1996.

[26] Y. Hori, K. Shimura, and M. Tomizuka. Position/force control of multi-axis robot manipulator based on the TDOF robust servo controller for each joint. In *Proceedings of the American Control Conference*, volume 1, pages 753–757, Chicago, June 1992.

[27] R. Horowitz and B. Li. Design and implementation of adaptive non-repetitive track-following disk file servos. In *Proceedings of the American Control Conference*, volume 6, pages 4151–4155, Seattle, June 1995.

[28] R. Horowitz and J. McCormick. A self-tuning control scheme for disk file servos. *IEEE Transactions on Magnetics*, 27(6):4490–4495, November 1991.

[29] T. Imamura, S. Hasegawa, K. Takaishi, and Y. Mizoshita. Piezoelectric microactuator compensating for off-track errors in magnetic disk drives. In *Advances in Information Storage Systems*, volume 5, pages 119–126. ASME Press, New York, 1993. B. Bhushan, editor.

[30] J. Ishikawa and M. Tomizuka. Pivot friction compensation using an acceleromter and a disturbance observer for hard disk drives. In *Advances in Information Storage*

and Processing Systems, Proceedings of the ASME IMECE, volume 3, Dallas, TX, November 1997. To appear.

[31] T. Kailath. *Linear Systems.* Prentice Hall, Englewood Cliffs, NJ, 1980.

[32] S. Kawaji, Y. Suenaga, T. Maeda, N. Matsunaga, N. Matsunaga, and T. Sasaoka. Control of cutting torque in the drilling process using disturbance observer. In *Proceedings of the American Control Conference*, volume 1, pages 723–728, Seattle, June 1995.

[33] C.J. Kempf. *Design of Servo Systems for Disturbance Rejection and Applications to Disk File Storage Systems.* PhD thesis, University of California, Berkeley, 1994.

[34] E. Korczynski. IBM mulling 1-in. hard drive. *Data Storage*, page 17, November/December 1996.

[35] H.S. Lee and H. Ferrier. Open-loop compensation of repeatable runouts in discrete-track magnetic disks. In *Advances in Information Storage and Processing Systems, Proceedings of the ASME IMECE*, pages 221–226, San Francisco, November 1995.

[36] H.S. Lee and M. Tomizuka. Robust tracking controller design for high-speed machining. In *Proceedings of the American Control Conference*, volume 1, pages 215–219, Seattle, June 1995.

[37] L. Ljung. Analysis of recursive stochastic algorithms. *IEEE Transactions on Automatic Control*, AC-22(4):551–575, August 1977.

[38] L. Ljung. *System Identification: Theory for the User.* Prentice-Hall, Englewood Cliffs, NJ, 1987.

[39] L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification.* MIT Press, Cambridge, MA, 1983.

[40] O. Macchi and E. Eweda. Second-order convergence analysis of stochastic adaptive linear filtering. *IEEE Transactions on Automatic Control*, AC-28(1):76–85, January 1983.

[41] O. Matsushita, T. Sonoda, H. Ohta, J. Naruse, Y. Inoue, and I. Koromegawa. Vibration analysis of disk-spindle due to ball bearing excitation. *Bulletin of the JSME*, 29(256):3510–3519, October 1986.

[42] W. Messner and C. Kempf. A comparison of four discrete-time repetitive control algorithms. Technical report, University of California, Berkeley, Mechanical Engineering Department, Computer Mechanics Laboratory, December 1991.

[43] M. Metivier and P. Priouret. Applications of a Kushner and Clark lemma to general classes of stochastic algorithms. *IEEE Transactions on Information Theory*, IT-30(2):140–151, March 1984.

[44] D.K. Miu and Y.-C. Tai. Silicon microstructures and microactuators for compact computer disk drives. In *Proceedings of the American Control Conference*, volume 6, pages 4136–4140, Seattle, June 1995.

[45] J. Morehouse and J. Blagaila. The little disk drives that could. *Data Storage*, pages 51–56, October 1996.

[46] T. Murakami and K. Ohnishi. Advanced motion control in mechatronics - a tutorial. In *Proceedings of the IEEE International Workshop on Intelligent Control*, volume 1, pages SL9–SL17, Istanbul, Turkey, August 1990.

[47] K. Ohnishi. A new servo method in mechatronics. *Transaction of the Japanese Society of Electrical Engineers*, 107-D:83–86, 1987.

[48] R.G. Parker and C.D. Mote, Jr. Vibration and coupling phenomena in asymmetric disk-spindle systems. In *Advances in Information Storage and Processing Systems, Proceedings of the ASME IMECE*, pages 293–302, San Francisco, November 1995.

[49] T.W. Parks and C.S. Burrus. *Digital Filter Design.* Wiley, New York, 1987.

[50] D.T. Phan. The design and modeling of multirate digital control systems for disk drive applications. In *Proceedings of the 1993 Asia-Pacific Workshop on Advances in Motion Control*, pages 189–205, Singapore, July 1993.

[51] H.R. Radwan and R. Whaley. Servo-structure interaction in disk drives using finite element analysis. In *Advances in Information Storage Systems*, volume 5, pages 101–118. ASME Press, 1993. B. Bhushan, editor.

[52] J. Reason and W. Ren. Estimating the optimal adaptation gain for the LMS algorithm. In *Proceedings of the 32nd IEEE Conference on Decision and Control*, pages 1587–1588, San Antonio, December 1993.

[53] W. Ren and P.R. Kumar. Adaptive active noise control: Structures, algorithms and convergence analysis. In *Proceedings of the 1989 International Conference on Noise Control Engineering*, volume 1, pages 435–440, Newport Beach, CA, December 1989. G.C. Maling, Jr., editor.

[54] W.O. Richter and F.E. Talke. Nonrepeatable radial and axial runout of $5\frac{1}{4}$” disk drive spindles. *IEEE Transactions on Magnetics*, 24(6):2760–2762, November 1988.

[55] A.H. Sacks, M. Bodson, and W. Messner. Advanced methods for repeatable runout compensation. *IEEE Transactions on Magnetics*, 31(2):1031–1036, March 1995.

[56] N. Sadegh and R. Horowitz. Stability and robustness analysis for a class of adaptive controllers for robotic manipulators. *International Journal of Robotics Research*, 9(3):74–92, 1990.

[57] I.Y. Shen and C.-P. Roger Ku. On the rocking motion of a rotating flexible-disk/rigid-shaft system. In *Advances in Information Storage and Processing Systems, Proceedings of the ASME IMECE*, pages 271–282, San Francisco, November 1995.

[58] I.Y. Shen and C.-P. Roger Ku. On the vibration analysis of multiple rotating flexible disks. In *Advances in Information Storage and Processing Systems, Proceedings of the ASME IMECE*, pages 259–270, San Francisco, November 1995.

[59] M.D. Sidman. Convergence properties of an adaptive runout correction system for disk drives. In *Advances in Information Storage Systems*, volume 1, pages 61–78. ASME Press, New York, 1991. B. Bhushan, editor.

[60] M.D. Sidman. Time optimal voltage-constrained deceleration velocity profile design for disk drives. In *Proceedings of the American Control Conference*, volume 6, pages 4146–4150, Seattle, June 1995.

[61] J.J.E. Slotine and W. Li. On the adaptive control of robotic manipulators. *International Journal of Robotics Research*, 6(3):49–59, 1987.

[62] C. Smith and M. Tomizuka. Shock rejection for repetitive control using a disturbance observer. In *Proceedings of the 35th IEEE Conference on Decision and Control*, pages 2503–2504, Kobe, Japan, December 1996.

[63] K. Sugiura and Y. Hori. Vibration suppression in 2- and 3-mass system based on the feedback of imperfect derivative of the estimated torsional torque. *IEEE Transactions on Industrial Electronics*, 43(1):56–64, February 1996.

[64] S. Tohyama, T. Yoshida, H. Otsuki, and T. Yamaguchi. Fast access servo method using two-degree-of-freedom control for magnetic disk drives. In *Advances in Information Storage and Processing Systems, Proceedings of the ASME IMECE*, pages 207–212, San Francisco, November 1995.

[65] M. Tomizuka. Zero-phase error tracking controller for digital control. *ASME Journal of Dynamic Systems, Measurement and Control*, 109:65–68, 1987.

[66] B.P. Tremaine. Optimal control of a VCM motor with vibro-acoustic constraints. In *Advances in Information Storage and Processing Systems, Proceedings of the ASME IMECE*, pages 197–205, San Francisco, November 1995.

[67] T. Umeno and Y. Hori. Two degrees of freedom controllers for robust servomechanism. In *IEEE International Workshop on Advanced Motion Control*, pages 47–56, Yokohama, Japan, March 1990.

[68] T. Umeno and Y. Hori. Robust speed control of dc servomotors using modern two degrees-of-freedom controller design. *IEEE Transactions on Industrial Electronics*, 38(5):363–368, October 1991.

[69] B. Widrow. Adaptive filters. In *Aspects of Network and System Theory*, pages 563–587. Holt, Rinehart and Winston, New York, 1970. R.E. Kalman and N. DeClaris, editors.

[70] B. Widrow, J.R. Glover, Jr., J.M. McCool, J. Kaunitz, C.S. Williams, R.H. Hearn, J.R. Zeidler, E. Dong, Jr., and R.C. Goodlin. Adaptive noise canceling: Principles and applications. *Proceedings of the IEEE*, 63(12):1692–1716, December 1975.

[71] B. Widrow, J.M. McCool, M.G. Larimore, and C.R. Johnson, Jr. Stationary and nonstationary learning characteristics of the LMS adaptive filter. *Proceedings of the IEEE*, 64(8):1151–1162, August 1976.

[72] B. Widrow and S.D. Stearns. *Adaptive Signal Processing.* Prentice Hall, Englewood Cliffs, NJ, 1985.

[73] T. Yamaguchi, Y. Soyama, H. Hosokawa, K. Tsuneta, and H. Hirai. Improvement of settling response of disk drive head positioning servo using mode switching control with initial value compensation. *IEEE Transactions on Magnetics*, 32(3):1767–1772, May 1996.

[74] T. Yamaguchi, Y. Soyama, H. Hosokawa, K. Tsuneta, and H. Hirai. Mode switching control design and its application to head positioning control on magnetic disk drives. *IEEE Transactions on Industrial Electronics*, 43(1):65–73, February 1996.

Appendix A

Disk Drive Test Equipment

A.1 Accelerometer Testing

The accelerometers used in the experiments were pre-production models (356M02) from PCB Piezotronics, Inc. They were designed to be extremely low-cost, *i.e.* on the order of a few dollars. Since these were test models, no specifications were available from the manufacturer. The consistency of the signals, the off-axis sensitivity, and the frequency range of the accelerometers were tested to determine whether they were suitable for the acceleration feedforward technique.

A diagram of the shaker system for testing of the accelerometers is shown in Figure A.1. A Brüel & Kjær mini-shaker provided translational, oscillatory motion. The shaker was connected to the centerplate by a thin nylon rod called a stinger. The opposite end of the centerplate was attached to a rod that slid in a linear rotary bearing, in-line with the shaker motion. Two accelerometers were soldered to identical copper-clad printed circuit boards. The printed circuit boards were mounted to the top and bottom of the centerplate. The printed circuit boards were slotted so that they could be rotated with respect to the shaker axis. Also mounted on the centerplate was a mirror that allowed an optical sensor to measure the position of the system. The position information was used in a feedback loop to control the shaker motion.

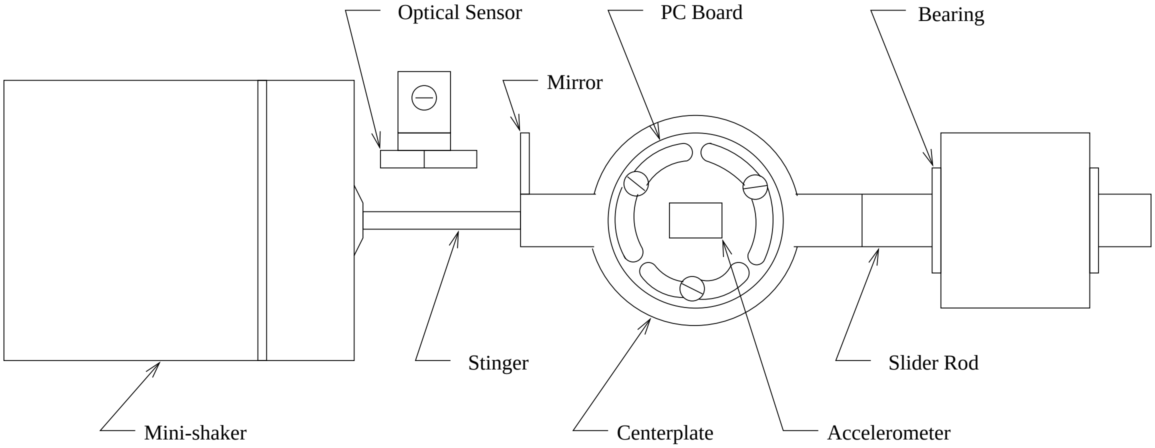


Figure A.1: Accelerometer Test System

A.1.1 Measurement Techniques

A single experimental trial consisted of shaking the two accelerometers at a fixed frequency. The first accelerometer was rotated by an angle θ from the shaker axis. The second accelerometer was kept in-line with the shaker motion. Signals from the accelerometers were sent through two PCB ICP power units to the two input channels of the dynamic signal analyzer (DSA), which allowed them to be compared. Signals from the y-axes of the accelerometers were analyzed in both the time and frequency domains.

In the time domain, time traces 1.00 second long were captured by the DSA. These time traces were nominally sinusoids with amplitudes A_1 and A_2. Due to the rotation of the first accelerometer, the amplitudes were related by $A_1 = A_2 \cos\theta$. The percent error was therefore defined as

$$\text{percent error} = \frac{|A_1 - A_2 \cos\theta|}{A_2} * 100, \tag{A.1}$$

and was calculated using features on the DSA.

In the frequency domain, the transfer function from the second accelerometer to the first accelerometer was calculated by the DSA and displayed as a Bode plot. The expected result was a constant magnitude of $\cos\theta$ with zero phase. To compare the results from the two methods, magnitude values at the frequencies used for the time domain analysis were extracted from the Bode plots and used to calculate the percent error.

A.1.2 Results

The percent errors calculated from the time domain analysis ranged from 1.56 % to 7.87 % with a mean of 4.25 % and a standard deviation of 1.92 %. Frequency domain values ranged from 0.10 % to 8.17 % with a mean of 3.37 % and a standard deviation of 2.22 %. The percent error did not appear to be dependent upon the value of θ. This suggested that off-axis sensitivity was not a significant factor. Measurements made using the x-axes of the two accelerometers and by rotating the second accelerometer instead of the first gave similar results.

In general, the frequency responses were well-behaved between 20 Hz and 1 kHz, although poorer responses were obtained for greater θ values as the signal level from the first accelerometer decreased. This is approximately the frequency range of interest for the disk drive experiments. Below 20 Hz, the acceleration levels from the experimental system were too small to generate recognizable signals. Above 1 kHz, the differences in the power units appeared to be a problem.

A.1.3 Mounting

For the disk drive experiments, the accelerometers were attached to the disk drive base with double-sided mounting strips. Shake testing of this method (courtesy of Sat Pannu) found it satisfactory up to approximately 1 kHz. This is similar to the frequency range where wax mounting is effective, although not as good as more permanent methods such as adhesives or screws.

A.2 Hitachi Test Equipment

Miscellaneous specifications for the disk drive used in the experimentation are given in Table A.2. During testing, the disk drive was connected to a digital control unit that replaced the analog controller circuitry that was built into the drive. (See Figure A.2 for a diagram of the experimental system.) The digital control unit included a TMS320C30 digital signal processor (DSP) that was used to calculate the feedback, feedforward, and disturbance observer control inputs for the drive. The TMS320C30 is a 32-bit floating point DSP. Also included in the digital control unit were 10-bit analog-to-digital converters (ADC) for the accelerometer signal and PES, 10-bit digital-to-analog converters (DAC) for

Specification	Value
Form factor	5.25 inch
Formatted capacity	1.321 Gbytes
Capacity per sector	512 bytes
Sectors per track	77
Number of cylinders	2235
Number of disks	8
Number of data heads	15
Number of servo heads	1
Average seek time	14 s
Maximum seek time	30 ms
Minimum seek time	3 ms
Average latency	8.3 ms
Disk speed	3600 rpm
Data transfer rate	2.75 Mbytes/s
Recording density	44,060 bpi
Track density	2000 tpi
Height	82.5 mm
Width	146 mm
Depth	203 mm
Mass	3.5 kg
Shock (operating)	2 G
Shock (nonoperating)	20 G
Vibration (operating)	0.25 G
Vibration (nonoperating)	0.5 G

Table A.1: Miscellaneous Specifications of the Experimental Disk Drive

the control inputs to the drive and signals for analysis, and a second-order LPF with a break frequency of 2.8 kHz for smoothing the VCM command signal.

A personal computer was used to run an interactive system control program that allowed the user to specify the operating mode of the drive. This included a variety of seeking exercises with timing capabilities, bias force calculation, and the default track following mode. A second computer ran the DSP debugger software with an XDS 510 emulator. An external hardware interrupt signaled the controller sampling period of 50 μs. Programs for the two personal computers were written in C and assembly languages.

A block diagram of the experimental equipment is shown in Figure A.3. This shows more explicitly the shaker, drive and digital control unit. The shaker signal was generated by the DSA through a power amplifier. The system hardware gains that appear

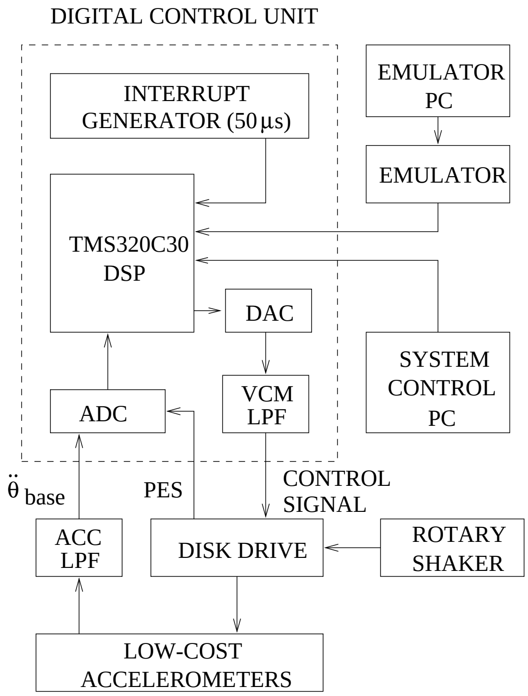


Figure A.2: Schematic of the Experimental System

in Figure A.3 and their values are given in Table A.2. Explanations for the transfer functions are presented in Table A.2. Figures A.4 and A.5 show the disk drive and the rotary shaker from two different views.

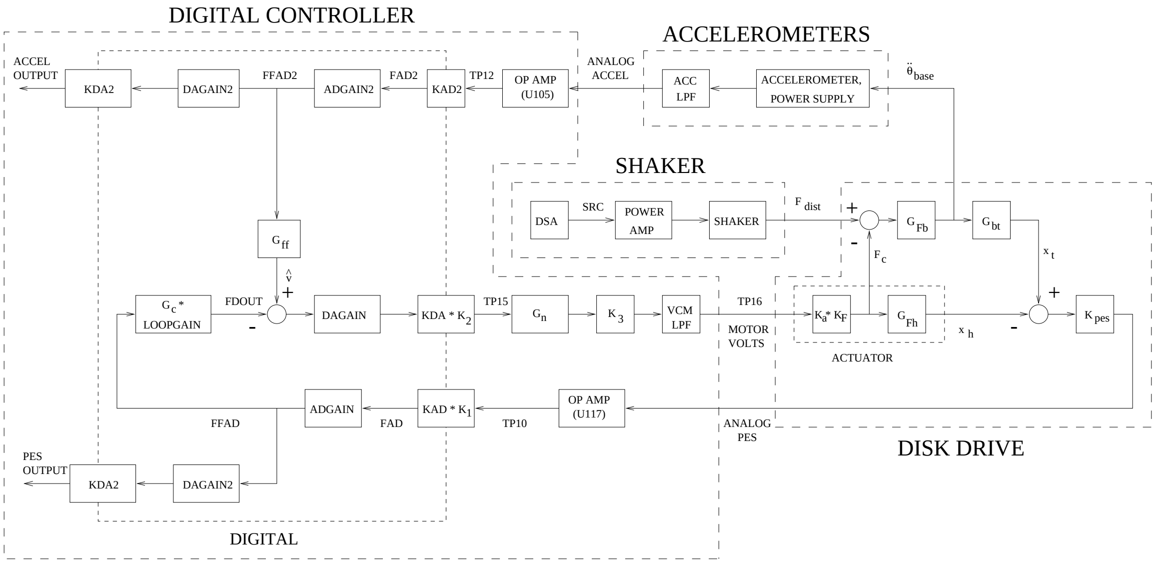


Figure A.3: Block Diagram of the Complete Disk Drive System and the Nominal Feedback Controller

Variable	Explanation	Value
K_{pes}	PES gain	0.11 v/μm
K_1	Gain adjustment	4.0 v/v
KAD	Analog-to-digital converter	1024/10 count/v
ADGAIN	Gain compensation for KAD	10/1024 count/count
LOOPGAIN	Gain adjustment	-3.1060218
KDA	Digital-to-analog converter	5/1024 v/count
DAGAIN	Gain compensation for KDA	1024/5 count/count
K_2	Gain adjustment	0.125 v/v
K_3	Gain adjustment	1.0 v/v
K_a	Transconductance amplifier	2.0 A/v
K_F	Motor force constant	2.8 N/A
KAD2	Second analog-to-digital converter	4096/5 count/v
ADGAIN2	Gain compensation for KAD2	5/4096 count/count
KDA2	Second digital-to-analog converter	5/256 v/count
DAGAIN2	Gain compensation for KDA2	256/5 count/count

Table A.2: Hardware Specifications of the Experimental Disk Drive

Symbol	Explanation
G_{Fb}	From force to base position
G_{bt}	From base position to track position
G_{Fh}	From torque to head position
G_n	Notch filters
ACC LPF	Low-pass filter for accelerometer signal
VCM LPF	Low-pass filter for voice coil motor input

Table A.3: Transfer Functions in Block Diagram

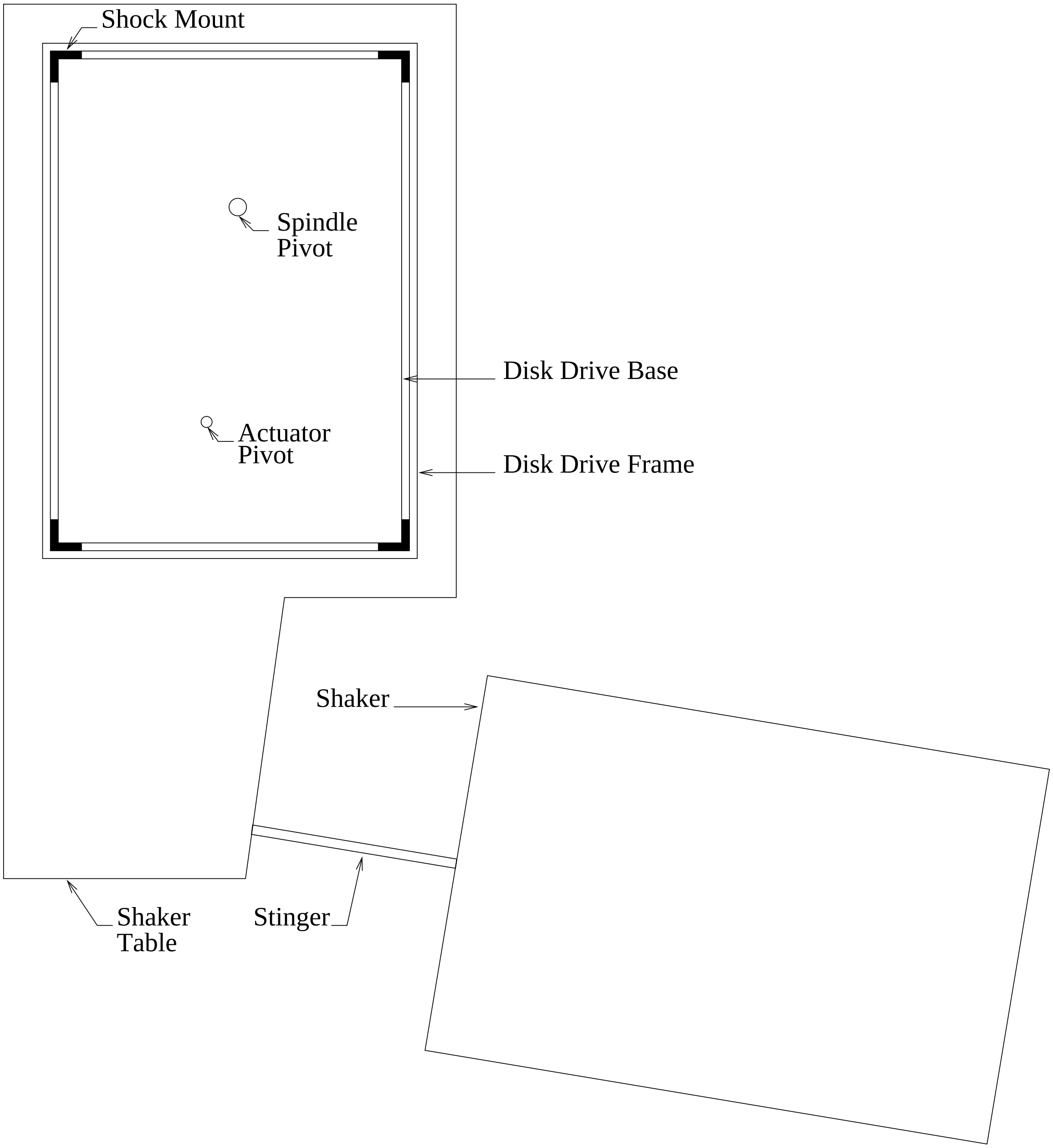


Figure A.4: Top View of the Rotary Shaker and the Disk Drive

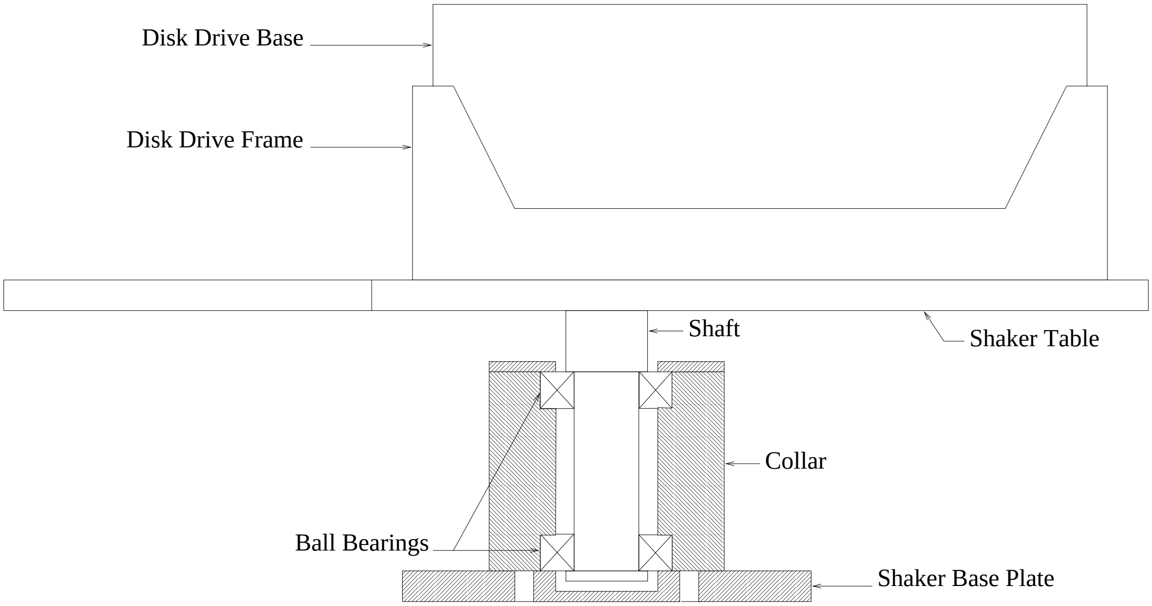


Figure A.5: Side View of the Rotary Shaker Table and the Disk Drive

Appendix B

Convergence Proofs for the Filtered-x LMS Adaptive Algorithm

B.1 Algorithm and Notation

Consider the filtered-x LMS algorithm as shown in Figure B.1. G_{ff} is the adaptive filter, G_{out} represents the plant dynamics, and $\hat{G}_{out}$ is an *a priori* estimate of G_{out}. If G_{ff} is a finite impulse response (FIR) filter, then

$$G_{ff}(k, q^{-1}) = w_0(k) + w_1(k)q^{-1} + ... + w_L(k)q^{-L} \tag{B.1}$$

where $w_i(k)$ are the filter parameters, or tap weights, at time k and q^{-1} is the one-step delay operator. For more compact notation, let $x(k)$ represent the input, which for the acceleration feedforward controller is the base acceleration $\ddot{\theta}_{base}(k)$. Then $\hat{v}(k)$, the output of G_{ff} for input $x(k)$, may be written as

$$\begin{aligned} \hat{v}(k) &= \mathbf{w}^T(k)\mathbf{x}(k) = \mathbf{x}^T(k)\mathbf{w}(k) && \text{(B.2)} \\ \mathbf{w}^T(k) &= [w_0(k), w_1(k), ..., w_L(k)] && \text{(B.3)} \\ \mathbf{x}^T(k) &= [x(k), x(k-1), ..., x(k-L)] && \text{(B.4)} \end{aligned}$$

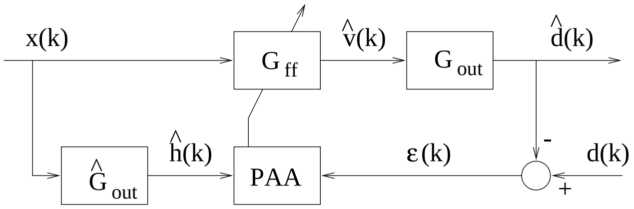


Figure B.1: Filtered-x LMS Algorithm

and $\hat{d}(k)$, the output of G_{out}, is

$$\hat{d}(k) = G_{out}(q^{-1})\hat{v}(k). \tag{B.5}$$

A key assumption of the filtered-x LMS algorithm is that $G_{out}(q^{-1})$ and $G_{ff}(k, q^{-1})$ commute, at least approximately, which can be justified in the case of small adaptation gain. This implies that

$$\hat{d}(k) = G_{out}(q^{-1})\left\{\mathbf{x}^T(k)\mathbf{w}(k)\right\} = \left\{G_{out}(q^{-1})\mathbf{x}^T(k)\right\}\mathbf{w}(k). \tag{B.6}$$

Now define the error $\varepsilon(k)$ and the mean-square error $\xi(k)$ as

$$\varepsilon(k) = d(k) - \hat{d}(k) \tag{B.7}$$

$$\xi(k) = E[\varepsilon^2(k)] \tag{B.8}$$

where $d(k)$ is the desired value and E denotes the expectation. For the model reference case discussed in Chapter 3, the desired value is given by

$$d(k) = G_{out}(q^{-1})G_{ref}(q^{-1})x(k). \tag{B.9}$$

Let $\varepsilon^2(k)$ serve as the estimate of its expectation $\xi(k)$. Thus the estimate of $\nabla(k)$, the gradient of $\xi(k)$, is

$$\hat{\nabla}(k) = \frac{\partial\varepsilon^2(k)}{\partial\mathbf{w}(k)} \tag{B.10}$$

$$= 2\varepsilon(k)\frac{\partial\varepsilon(k)}{\partial\mathbf{w}(k)} \tag{B.11}$$

$$= 2\varepsilon(k)\frac{\partial(d(k) - \hat{d}(k))}{\partial\mathbf{w}(k)} \tag{B.12}$$

$$= -2\varepsilon(k)\frac{\partial\hat{d}(k)}{\partial\mathbf{w}(k)} \tag{B.13}$$

$$\begin{aligned} &= -2\varepsilon(k)\frac{\partial(\{G_{out}(q^{-1})\mathbf{x}^T(k)\}\mathbf{w}(k))}{\partial \mathbf{w}(k)} && \text{(B.14)} \\ &= -2\varepsilon(k)\{G_{out}(q^{-1})\mathbf{x}(k)\} && \text{(B.15)} \\ &= -2\{G_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k). && \text{(B.16)} \end{aligned}$$

However, the vector $G_{out}(q^{-1})\mathbf{x}(k)$ is not available. Instead, use the modeling approximation $\hat{G}_{out}(q^{-1})\mathbf{x}(k)$. This results in

$$\hat{\nabla}(k) = -2\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k). \tag{B.17}$$

The update law for the parameters of $G_{ff}(k, q^{-1})$ is made in the direction of the negative gradient

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) - \mu\nabla(k) && \text{(B.18)} \\ &\approx \mathbf{w}(k) + 2\mu\{G_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k) && \text{(B.19)} \\ &\approx \mathbf{w}(k) + 2\mu\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k). && \text{(B.20)} \end{aligned}$$

B.2 Solution for the Optimal Filter Weights

Before analyzing the convergence of the algorithm, it is useful to determine the value of the optimal filter weights. To this end, begin with the error equation

$$\begin{aligned} \varepsilon(k) &= d(k) - \hat{d}(k) && \text{(B.21)} \\ &= d(k) - \{G_{out}(q^{-1})\mathbf{x}(k)\}^T\mathbf{w}(k) && \text{(B.22)} \end{aligned}$$

where the assumption that $G_{ff}(k, q^{-1})$ and $G_{out}(q^{-1})$ commute has been applied. Squaring Equation B.22 results in

$$\begin{aligned} \varepsilon^2(k) &= d^2(k) + \mathbf{w}^T(k)\{G_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T\mathbf{w}(k) \\ &\quad - 2d(k)\{G_{out}(q^{-1})\mathbf{x}(k)\}^T\mathbf{w}(k). && \text{(B.23)} \end{aligned}$$

Assume that $\varepsilon(k)$ is stationary and that $d(k)$ and $\mathbf{x}(k)$ are jointly stationary. Then taking the expectation of Equation B.23 leads to

$$\begin{aligned} E[\varepsilon^2(k)] &= E[d^2(k)] + E[\mathbf{w}^T(k)\{G_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T\mathbf{w}(k)] \\ &\quad - E[2d(k)\{G_{out}(q^{-1})\mathbf{x}(k)\}^T\mathbf{w}(k)]. && \text{(B.24)} \end{aligned}$$

Now assume that $\mathbf{w}$ is constant, then

$$\begin{aligned} E[\varepsilon^2(k)] &= E[d^2(k)] + \mathbf{w}^T E[\{G_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T]\mathbf{w} \\ &\quad - 2E[d(k)\{G_{out}(q^{-1})\mathbf{x}(k)\}^T]\mathbf{w}. \end{aligned} \tag{B.25}$$

Define

$$\mathbf{R} = E[\{G_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T] \tag{B.26}$$

or

$$\mathbf{R} = E\begin{bmatrix} \{G_{out}(q^{-1})x(k)\}^2 & \dots & \{G_{out}(q^{-1})x(k)\}\{G_{out}(q^{-1})x(k-L)\} \\ \vdots & \ddots & \vdots \\ \{G_{out}(q^{-1})x(k-L)\}\{G_{out}(q^{-1})x(k)\} & \dots & \{G_{out}(q^{-1})x(k-L)\}^2 \end{bmatrix} \tag{B.27}$$

and

$$\mathbf{p} = E[d(k)\{G_{out}(q^{-1})\mathbf{x}(k)\}] = E[\{G_{out}(q^{-1})\mathbf{x}(k)\}d(k)]. \tag{B.28}$$

Note that by construction, $\mathbf{R}$ is real and symmetric. Furthermore, $\mathbf{R}$ is positive semi-definite, since for any $\mathbf{z} \in \mathcal{R}^{L+1}$, the set of real $(L+1)$ vectors,

$$\mathbf{z}^T\mathbf{R}\mathbf{z} = \mathbf{z}^T E[\{G_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T]\mathbf{z} \tag{B.29}$$

$$= E[\Big(\mathbf{z}^T\{G_{out}(q^{-1})\mathbf{x}(k)\}\Big)\Big(\{G_{out}(q^{-1})\mathbf{x}(k)\}^T\mathbf{z}\Big)] \tag{B.30}$$

$$= E[\Big(\mathbf{z}^T\{G_{out}(q^{-1})\mathbf{x}(k)\}\Big)^2] \tag{B.31}$$

$$\geq 0 \;\; if \;\; \mathbf{z} \neq \mathbf{0} \tag{B.32}$$

$$= 0 \;\; if \;\; \mathbf{z} = \mathbf{0}. \tag{B.33}$$

The eigenvalues of $\mathbf{R}$ are real and non-negative. Typically, $\mathbf{R}$ will be positive definite, in which case the eigenvalues are positive. Substituting for $\mathbf{R}$ and $\mathbf{p}$, Equation B.25 becomes

$$\xi = E[\varepsilon^2(k)] \tag{B.34}$$

$$= E[d^2(k)] + \mathbf{w}^T\mathbf{R}\mathbf{w} - 2\mathbf{p}^T\mathbf{w} \tag{B.35}$$

and the gradient can be calculated as

$$\nabla = \frac{\partial \xi}{\partial \mathbf{w}} \tag{B.36}$$

$$= \frac{\partial}{\partial \mathbf{w}}(E[d^2(k)] + \mathbf{w}^T\mathbf{R}\mathbf{w} - 2\mathbf{p}^T\mathbf{w}) \tag{B.37}$$

$$= 2\mathbf{R}\mathbf{w} - 2\mathbf{p}. \tag{B.38}$$

Note that due to the stationarity assumptions, $\mathbf{R}$, $\mathbf{p}$, ξ, and ∇ are all constant with respect to k. Setting $\nabla = \mathbf{0}$ and solving for the optimal weights $\mathbf{w} = \mathbf{w}^*$ results in

$$\mathbf{w}^* = \mathbf{R}^{-1}\mathbf{p} \tag{B.39}$$

assuming that $\mathbf{R}$ is positive definite, and therefore nonsingular. Substituting these values into Equation B.35 gives the minimum mean-square error

$$\begin{aligned}
\xi &= E[d^2(k)] + \mathbf{w}^{*T}\mathbf{R}\mathbf{w}^* - 2\mathbf{p}^T\mathbf{w}^* && \text{(B.40)}\\
&= E[d^2(k)] + (\mathbf{R}^{-1}\mathbf{p})^T\mathbf{R}\mathbf{R}^{-1}\mathbf{p} - 2\mathbf{p}^T\mathbf{R}^{-1}\mathbf{p} && \text{(B.41)}\\
&= E[d^2(k)] - \mathbf{p}^T\mathbf{R}^{-1}\mathbf{p} && \text{(B.42)}\\
&= E[d^2(k)] - \left(E[d(k)\mathbf{h}^T(k)]\right)\left(E[\mathbf{h}(k)\mathbf{h}^T(k)]\right)^{-1}\left(E[d(k)\mathbf{h}(k)]\right) && \text{(B.43)}\\
&= \xi_{min} && \text{(B.44)}
\end{aligned}$$

where

$$\mathbf{h}(k) = G_{out}(q^{-1})\mathbf{x}(k). \tag{B.45}$$

B.3 Solution for the Converged Filter Weights

Define

$$\hat{\mathbf{R}} = E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T] \tag{B.46}$$

or

$$\hat{\mathbf{R}} = E\begin{bmatrix} \{\hat{G}_{out}(q^{-1})x(k)\}\{G_{out}(q^{-1})x(k)\} & \dots & \{\hat{G}_{out}(q^{-1})x(k)\}\{G_{out}(q^{-1})x(k-L)\} \\ \vdots & \ddots & \vdots \\ \{\hat{G}_{out}(q^{-1})x(k-L)\}\{G_{out}(q^{-1})x(k)\} & \dots & \{\hat{G}_{out}(q^{-1})x(k-L)\}\{G_{out}(q^{-1})x(k-L)\} \end{bmatrix} \tag{B.47}$$

and

$$\hat{\mathbf{p}} = E[d(k)\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}] = E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}d(k)]. \tag{B.48}$$

Assume that

$$H(q^{-1}) = \frac{\hat{G}_{out}(q^{-1})}{G_{out}(q^{-1})} \tag{B.49}$$

is strictly positive real (SPR), *i.e.*

$$\mathrm{Re}(H(e^{-j\omega})) > 0 \qquad \forall\,\omega \tag{B.50}$$

then $\hat{\mathbf{R}}$ is positive semi-definite. To see this, begin with Equation B.46

$$\begin{aligned}
\hat{\mathbf{R}} &= E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T] && \text{(B.51)}\\
&= E[\left\{\frac{\hat{G}_{out}(q^{-1})}{G_{out}(q^{-1})}G_{out}(q^{-1})\mathbf{x}(k)\right\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T] && \text{(B.52)}\\
&= E[\{H(q^{-1})\mathbf{h}(k)\}\mathbf{h}^T(k)]. && \text{(B.53)}
\end{aligned}$$

Consider any $\mathbf{z} \in \mathcal{R}^{L+1}$, then

$$\begin{aligned}
\mathbf{z}^T\hat{\mathbf{R}}\mathbf{z} &= \mathbf{z}^T E[\{H(q^{-1})\mathbf{h}(k)\}\mathbf{h}^T(k)]\mathbf{z} && \text{(B.54)}\\
&= E[\mathbf{z}^T\{H(q^{-1})\mathbf{h}(k)\}\mathbf{h}^T(k)\mathbf{z}] && \text{(B.55)}\\
&= E[\tilde{\phi}(k)\phi(k)] && \text{(B.56)}\\
&= \frac{1}{2\pi}\int_{-\pi}^{\pi} H(e^{-j\omega})\Phi_{\phi\phi}(\omega)\,d\omega && \text{(B.57)}
\end{aligned}$$

where $\phi(k) = \mathbf{h}^T(k)\mathbf{z}$ and $\tilde{\phi}(k) = \mathbf{z}^T\{H(q^{-1})\mathbf{h}(k)\}$ are scalars, and $\Phi_{\phi\phi}(\omega)$ is the spectral density of $\phi(k)$. Since $\Phi_{\phi\phi}(\omega)$ and $E[\tilde{\phi}(k)\phi(k)]$ are real

$$\mathbf{z}^T\hat{\mathbf{R}}\mathbf{z} = \frac{1}{2\pi}\int_{-\pi}^{\pi} \mathrm{Re}(H(e^{-j\omega}))\Phi_{\phi\phi}(\omega)\,d\omega \tag{B.58}$$

but $\mathrm{Re}(H(e^{-j\omega})) > 0$, and $\Phi_{\phi\phi}(\omega) \geq 0$, therefore

$$\mathbf{z}^T\hat{\mathbf{R}}\mathbf{z} \geq 0 \tag{B.59}$$

and $\hat{\mathbf{R}}$ is positive semi-definite [31], although $\hat{\mathbf{R}}$ is not necessarily symmetric. The eigenvalues of $\hat{\mathbf{R}}$ are real and non-negative. Typically, $\hat{\mathbf{R}}$ will be positive definite, in which case the eigenvalues are positive and $\hat{\mathbf{R}}$ is nonsingular. Note that due to the stationarity assumptions, $\hat{\mathbf{R}}$ and $\hat{\mathbf{p}}$ are constant with respect to k. Assuming that $\hat{\mathbf{R}}$ is positive definite, define

$$\hat{\mathbf{w}} = \hat{\mathbf{R}}^{-1}\hat{\mathbf{p}}. \tag{B.60}$$

Note that if $\hat{G}_{out}(q^{-1}) = G_{out}(q^{-1})$, then $\hat{\mathbf{R}} = \mathbf{R}$, $\hat{\mathbf{p}} = \mathbf{p}$, and $\hat{\mathbf{w}} = \mathbf{w}^*$.

B.4 Filtered-x LMS Convergence Proof: Constant Adaptation Gain

The convergence proof presented in this section follows the proof for the LMS algorithm given in Widrow and Stearns [72], but has been modified for the filtered-x LMS algorithm. The major difference between the two techniques is the introduction of the estimated transfer function $\hat{G}_{out}$ for the filtered-x LMS algorithm. To analyze the convergence of $\mathbf{w}(k)$, take the expectation of the parameter update equation using the expectation and modeling approximations, Equation B.20.

$$\begin{aligned} E[\mathbf{w}(k+1)] &= E[\mathbf{w}(k)] + 2\mu E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k)] && \text{(B.61)} \\ &= E[\mathbf{w}(k)] + 2\mu E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}d(k)] \\ &\quad - 2\mu E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\hat{d}(k)] && \text{(B.62)} \\ &= E[\mathbf{w}(k)] + 2\mu E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}d(k)] \\ &\quad - 2\mu E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}^T(k)\}\mathbf{w}(k)] && \text{(B.63)} \end{aligned}$$

where the commutivity assumption has again been used. Assume as well that the stationarity conditions remain and that $\mathbf{x}(k)$ and $\mathbf{w}(k)$ are uncorrelated. A sufficient condition for $\mathbf{x}(k)$ and $\mathbf{w}(k)$ to be uncorrelated is that successive input vectors are independent over time [72]. Although this is a standard assumption, it is not always justifiable. Convergence can be shown for certain cases with correlated data, but this complicates the analysis considerably. Furthermore, techniques that allow correlated data typically have shortcomings of their own. See [17, 40] for further discussion on this topic. For the sake of clarity, and since results predicted using the assumption of uncorrelated data have been shown to match experimental results with correlated data [71, 70, 69], the assumption that $\mathbf{x}(k)$ and $\mathbf{w}(k)$ are uncorrelated will be retained. It follows that

$$\begin{aligned} & E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T\mathbf{w}(k)] \\ &\qquad = E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T]E[\mathbf{w}(k)] && \text{(B.64)} \\ &\qquad = \hat{\mathbf{R}}E[\mathbf{w}(k)]. && \text{(B.65)} \end{aligned}$$

Substituting $\hat{\mathbf{R}}$ and $\hat{\mathbf{p}}$ into Equation B.63,

$$\begin{aligned} E[\mathbf{w}(k+1)] &= E[\mathbf{w}(k)] + 2\mu\left\{\hat{\mathbf{p}} - \hat{\mathbf{R}}E[\mathbf{w}(k)]\right\} && \text{(B.66)} \\ &= (\mathbf{I} - 2\mu\hat{\mathbf{R}})E[\mathbf{w}(k)] + 2\mu\hat{\mathbf{p}} && \text{(B.67)} \\ &= (\mathbf{I} - 2\mu\hat{\mathbf{R}})E[\mathbf{w}(k)] + 2\mu\hat{\mathbf{R}}\hat{\mathbf{w}} && \text{(B.68)} \end{aligned}$$

where $\mathbf{I}$ is the identity matrix. Let $\hat{\mathbf{Q}}$ be the similarity transformation matrix such that

$$\hat{\mathbf{Q}}^{-1}\hat{\mathbf{R}}\hat{\mathbf{Q}} = \hat{\mathbf{J}} \tag{B.69}$$

where $\hat{\mathbf{J}}$ is the Jordan form of $\hat{\mathbf{R}}$. Now, transform Equation B.68 to the principal coordinate system, first by the translation $\mathbf{w}(k) = \mathbf{v}(k) + \hat{\mathbf{w}}$,

$$\begin{aligned} E[\mathbf{v}(k+1) + \hat{\mathbf{w}}] &= (\mathbf{I} - 2\mu\hat{\mathbf{R}})E[\mathbf{v}(k) + \hat{\mathbf{w}}] + 2\mu\hat{\mathbf{R}}\hat{\mathbf{w}} && \text{(B.70)} \\ \Rightarrow E[\mathbf{v}(k+1)] + \hat{\mathbf{w}} &= (\mathbf{I} - 2\mu\hat{\mathbf{R}})E[\mathbf{v}(k)] + (\mathbf{I} - 2\mu\hat{\mathbf{R}})\hat{\mathbf{w}} + 2\mu\hat{\mathbf{R}}\hat{\mathbf{w}} && \text{(B.71)} \\ \Rightarrow E[\mathbf{v}(k+1)] &= (\mathbf{I} - 2\mu\hat{\mathbf{R}})E[\mathbf{v}(k)] && \text{(B.72)} \end{aligned}$$

and then by the rotation $\mathbf{v}(k) = \hat{\mathbf{Q}}\mathbf{v}'(k)$,

$$\begin{aligned} E[\hat{\mathbf{Q}}\mathbf{v}'(k+1)] &= (\mathbf{I} - 2\mu\hat{\mathbf{R}})E[\hat{\mathbf{Q}}\mathbf{v}'(k)] && \text{(B.73)} \\ \Rightarrow E[\mathbf{v}'(k+1)] &= \hat{\mathbf{Q}}^{-1}(\mathbf{I} - 2\mu\hat{\mathbf{R}})\hat{\mathbf{Q}}E[\mathbf{v}'(k)] && \text{(B.74)} \\ \Rightarrow E[\mathbf{v}'(k+1)] &= (\mathbf{I} - 2\mu\hat{\mathbf{J}})E[\mathbf{v}'(k)]. && \text{(B.75)} \end{aligned}$$

Since the eigenvalues λ_i of $\hat{\mathbf{J}}$ are the diagonal elements, the eigenvalues of $(\mathbf{I} - 2\mu\hat{\mathbf{J}})$ also lie on the diagonal. Recall that the eigenvalues of $\hat{\mathbf{R}}$, and therefore $\hat{\mathbf{J}}$, are positive. Thus, the system described by Equation B.75 is stable if

$$|1 - 2\mu\lambda_i| < 1 \tag{B.76}$$

for $i = 1, ..., (L+1)$ or

$$0 < \mu < \frac{1}{\lambda_{max}} \tag{B.77}$$

where λ_{max} is the maximum eigenvalue of $\hat{\mathbf{J}}$, which equals the maximum eigenvalue of $\hat{\mathbf{R}}$. Note that

$$\lambda_{max} \leq tr[\hat{\mathbf{J}}] = tr[\hat{\mathbf{R}}] \tag{B.78}$$

which implies that

$$\frac{1}{\lambda_{max}} \geq \frac{1}{tr[\hat{\mathbf{J}}]} = \frac{1}{tr[\hat{\mathbf{R}}]}. \tag{B.79}$$

Thus, convergence of the mean is assured if

$$0 < \mu < \frac{1}{tr[\hat{\mathbf{R}}]} \tag{B.80}$$

since this implies the following

$$\begin{aligned} \lim_{k\to\infty} E[\mathbf{v}'(k)] &= \mathbf{0} && \text{(B.81)} \\ \lim_{k\to\infty} E[\hat{\mathbf{Q}}^{-1}\mathbf{v}(k)] &= \mathbf{0} && \text{(B.82)} \\ \lim_{k\to\infty} E[\mathbf{v}(k)] &= \mathbf{0} && \text{(B.83)} \\ \lim_{k\to\infty} E[\mathbf{w}(k) - \hat{\mathbf{w}}] &= \mathbf{0}. && \text{(B.84)} \end{aligned}$$

Since this is a single-input system and $x(k)$ is assumed to be stationary,

$$\begin{aligned} &E[\{\hat{G}_{out}(q^{-1})x(k)\}\{G_{out}(q^{-1})x(k)\}] = \\ &\quad E[\{\hat{G}_{out}(q^{-1})x(k-i)\}\{G_{out}(q^{-1})x(k-i)\}] \end{aligned} \tag{B.85}$$

for all $i = 0, ..., L$. Thus

$$\begin{aligned} tr[\hat{\mathbf{R}}] &= \sum_{i=0}^{L} E[\{\hat{G}_{out}(q^{-1})x(k-i)\}\{G_{out}(q^{-1})x(k-i)\}] && \text{(B.86)} \\ &= (L+1)E[\{\hat{G}_{out}(q^{-1})x(k)\}\{G_{out}(q^{-1})x(k)\}] && \text{(B.87)} \end{aligned}$$

or $(L+1)$ times the filtered input signal power. This is a convenient, if possibly conservative, estimate of λ_{max}. Since the signal $G_{out}(q^{-1})x(k)$ is not available, use the average value of $\hat{h}^2(k) = \{\hat{G}_{out}(q^{-1})x(k)\}^2$ for the estimate of $E[\{\hat{G}_{out}(q^{-1})x(k)\}\{G_{out}(q^{-1})x(k)\}]$. Widrow and Stearns [72] suggest a value of μ approximately equal to one-tenth of the maximum value given by Equation B.80. Note that due to the trade-offs between fast convergence and sensitivity to noise, the maximum value of μ is rarely the optimal value [5, 52].

Using $\hat{G}_{out}(q^{-1})$ in place of $G_{out}(q^{-1})$ for calculating the bound on μ limits the magnitude of the allowable modeling error in $\hat{G}_{out}(q^{-1})$. Considering Equations B.80 and

B.87, it is evident that there is an inverse relationship between the accuracy of the estimate of the bound on μ and the size of the $\hat{G}_{out}(q^{-1})$ magnitude error. To set μ close to its maximum value, an accurate estimate of the magnitude of $G_{out}(q^{-1})$ is required. However, as stated above, it is typical to use a value of μ significantly lower than the maximum. This relaxes the requirement on the estimate of the magnitude of $G_{out}(q^{-1})$.

The other requirement on the estimate $\hat{G}_{out}(q^{-1})$ is the SPR condition stated previously. The SPR condition limits the phase error in $\hat{G}_{out}(q^{-1})$, since it is equivalent to the condition

$$-90^\circ < \angle(G_{out}(e^{-j\omega})) - \angle(\hat{G}_{out}(e^{-j\omega})) < 90^\circ \qquad \forall\,\omega \tag{B.88}$$

where $\angle(G(e^{-j\omega}))$ denotes the phase angle of $G(e^{-j\omega})$. Thus, there is a requirement for the magnitude estimate of $G_{out}(q^{-1})$ and a requirement for the phase estimate of $G_{out}(q^{-1})$ to achieve convergence of the mean. The SPR condition of Equation B.50 is not terribly restrictive, and if μ is chosen to be near one-tenth of its maximum, neither is the requirement on the $\hat{G}_{out}(q^{-1})$ magnitude error.

B.5 Filtered-x LMS Convergence Proof: Vanishing Adaptation Gain

B.5.1 Derivation and Motivation of the Associated ODE

Consider the filtered-x LMS adaptive algorithm

$$\begin{aligned} \mathbf{w}(k+1) &= \mathbf{w}(k) + \gamma(k)\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k) && \text{(B.89)} \\ \varepsilon(k) &= d(k) - G_{out}(q^{-1})\{\mathbf{w}^T(k)\mathbf{x}(k)\} && \text{(B.90)} \end{aligned}$$

which differs from the typical filtered-x LMS in that the constant adaptation gain 2μ has been replaced with $\gamma(k)$ such that $\gamma(k) \to 0$ as $k \to \infty$. Assume as before that

- $G_{out}(q^{-1})$ and $G_{ff}(k, q^{-1})$ commute

- $\varepsilon(k)$ is stationary, and $d(k)$ and $x(k)$ are jointly stationary
- $H(q^{-1}) = \frac{\hat{G}_{out}(q^{-1})}{G_{out}(q^{-1})}$ is SPR
- $x(k)$ and $\mathbf{w}(k)$ are uncorrelated.

Ljung proposed the ODE method as a general technique for analyzing the convergence of adaptive algorithms. The ODE method determines the asymptotic behavior of the adaptive algorithm by associating it with an ordinary differential equation (ODE), and then determining the behavior of the ODE [37, 38]. It will be shown that the filtered-x LMS algorithm with vanishing adaptation gain satisfies the requirements of the ODE method, and then by applying the ODE method, that the tap weights of the filtered-x LMS algorithm converge with probability one.

For the filtered-x LMS algorithm, the associated ODE is given by

$$\begin{aligned} \frac{d}{d\tau}\mathbf{w}_D(\tau) &= \bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\varepsilon(n, \mathbf{w}_D(\tau))] \quad &(B.91) \\ &= f(\mathbf{w}_D(\tau)) \quad &(B.92) \end{aligned}$$

where τ is a new time scale and $\mathbf{w}_D(\tau)$ is the vector of tap weights as a function of the new time scale. The derivation of τ and $\mathbf{w}_D(\tau)$ are explained in the following development. The expectation $\bar{E}$ is defined as

$$\begin{aligned} &\bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\varepsilon(n, \mathbf{w}_D(\tau))] \\ &\qquad = \lim_{N\to\infty} \frac{1}{N} \sum_{n=1}^{N} E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\varepsilon(n, \mathbf{w}_D(\tau))] \quad (B.93) \end{aligned}$$

where n has replaced τ in $x(\tau)$ and $\varepsilon(\tau)$ to emphasize that $\mathbf{w}_D(\tau)$ is held constant over the summation from $n = 1$ to $n = N$. To derive the associated ODE and link it heuristically to the adaptive algorithm, begin by noting that

$$\Delta\mathbf{w}(k+1) \sim \{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k) \quad (B.94)$$

where $\sim$ denotes proportionality. For small values of $\gamma(k)$, any significant changes in $\mathbf{w}(k)$ can be assumed to occur in the direction of the expectation, *i.e.*

$$\Delta\mathbf{w}(k+1) \sim E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k)]. \quad (B.95)$$

Let $\mathbf{w}(k_0) = \bar{\mathbf{w}}$, a constant, and choose $k > k_0$ so that $\mathbf{w}(k)$ and $\mathbf{w}(k_0)$ are nearly equal. Therefore,

$$\begin{aligned} \mathbf{w}(k) - \mathbf{w}(k_0) &= \sum_{m=k_0}^{k} \gamma(m)\{\hat{G}_{out}(q^{-1})\mathbf{x}(m)\}\varepsilon(m) && \text{(B.96)} \\ &\approx \sum_{m=k_0}^{k} \gamma(m)\{\hat{G}_{out}(q^{-1})\mathbf{x}(m)\}\varepsilon(m, \bar{\mathbf{w}}) && \text{(B.97)} \end{aligned}$$

since $\mathbf{w}(k) \approx \bar{\mathbf{w}}$. Substituting for the value at time m with the average value $f(\bar{\mathbf{w}})$ defined by

$$f(\bar{\mathbf{w}}) = \bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\varepsilon(n, \bar{\mathbf{w}})] \tag{B.98}$$

leads to

$$\begin{aligned} \mathbf{w}(k) - \mathbf{w}(k_0) &\approx \sum_{m=k_0}^{k} \gamma(m)\bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\varepsilon(n, \bar{\mathbf{w}})] && \text{(B.99)} \\ &= \left(\sum_{m=k_0}^{k} \gamma(m)\right)\left(\bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\varepsilon(n, \bar{\mathbf{w}})]\right) && \text{(B.100)} \end{aligned}$$

since the average value is independent of m. Substituting the definition

$$\Delta\tau = \sum_{m=k_0}^{k} \gamma(m) \tag{B.101}$$

results in

$$\mathbf{w}(k) - \mathbf{w}(k_0) \approx \Delta\tau\bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\varepsilon(n, \bar{\mathbf{w}})]. \tag{B.102}$$

Now define a change of time scale from k to τ_k by

$$\begin{aligned} \sum_{m=1}^{k} \gamma(m) &\leftrightarrow \tau_k && \text{(B.103)} \\ \mathbf{w}(k) &\leftrightarrow \mathbf{w}_D(\tau_k). && \text{(B.104)} \end{aligned}$$

Then Equation B.102 becomes

$$\mathbf{w}_D(\tau_k) - \mathbf{w}_D(\tau_{k_0}) = (\tau_k - \tau_{k_0})\bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\varepsilon(n, \mathbf{w}_D(\tau_{k_0}))] \tag{B.105}$$

which is Euler's method for solving Equation B.91, and therefore links the associated ODE to the adaptation algorithm, assuming that $\gamma(k)$ is small. A more rigorous justification of the use of the associated ODE for analyzing the adaptive algorithm can be found in [37].

B.5.2 Stability of the Equilibrium Point $\hat{\mathbf{w}}$

Once the associated ODE has been derived, its behavior is analyzed to determine the asymptotic behavior of the adaptation algorithm. For the filtered-x LMS algorithm, a Lyapunov analysis is suitable. To find the equilibrium point $\hat{\mathbf{w}}$, set $\frac{d}{d\tau}\mathbf{w}_D(\tau)$ equal to zero so that

$$\begin{aligned}
\mathbf{0} &= \bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\varepsilon(n, \mathbf{w}_D)] && \text{(B.106)}\\
&= \bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\left(d(n) - G_{out}(q^{-1})\{\mathbf{x}^T(n)\mathbf{w}_D\}\right)] && \text{(B.107)}\\
&= \bar{E}[\{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}d(n) - \{\hat{G}_{out}(q^{-1})\mathbf{x}(n)\}\{G_{out}(q^{-1})\mathbf{x}(n)\}^T\mathbf{w}_D] && \text{(B.108)}\\
&= \hat{\mathbf{p}} - \hat{\mathbf{R}}\mathbf{w}_D. && \text{(B.109)}
\end{aligned}$$

Therefore

$$\mathbf{w}_D = \hat{\mathbf{R}}^{-1}\hat{\mathbf{p}} = \hat{\mathbf{w}} \tag{B.110}$$

is an equilibrium point. Choose a Lyapunov function

$$V = \frac{1}{2}(\mathbf{w}_D - \hat{\mathbf{w}})^T(\mathbf{w}_D - \hat{\mathbf{w}}). \tag{B.111}$$

Note that V is positive definite and radially unbounded. Determine the derivative of V along the trajectories of the associated ODE

$$\begin{aligned}
\frac{dV}{d\tau} &= (\mathbf{w}_D - \hat{\mathbf{w}})^T\frac{d}{d\tau}(\mathbf{w}_D - \hat{\mathbf{w}}) && \text{(B.112)}\\
&= (\mathbf{w}_D - \hat{\mathbf{w}})^T\frac{d}{d\tau}\mathbf{w}_D && \text{(B.113)}\\
&= (\mathbf{w}_D - \hat{\mathbf{w}})^T(\hat{\mathbf{p}} - \hat{\mathbf{R}}\mathbf{w}_D) && \text{(B.114)}\\
&= (\mathbf{w}_D - \hat{\mathbf{w}})^T(\hat{\mathbf{R}}\hat{\mathbf{R}}^{-1})(\hat{\mathbf{p}} - \hat{\mathbf{R}}\mathbf{w}_D) && \text{(B.115)}\\
&= (\mathbf{w}_D - \hat{\mathbf{w}})^T\hat{\mathbf{R}}(\hat{\mathbf{R}}^{-1}\hat{\mathbf{p}} - \hat{\mathbf{R}}^{-1}\hat{\mathbf{R}}\mathbf{w}_D) && \text{(B.116)}\\
&= (\mathbf{w}_D - \hat{\mathbf{w}})^T\hat{\mathbf{R}}(\hat{\mathbf{w}} - \mathbf{w}_D) && \text{(B.117)}\\
&= -(\mathbf{w}_D - \hat{\mathbf{w}})^T\hat{\mathbf{R}}(\mathbf{w}_D - \hat{\mathbf{w}}) && \text{(B.118)}
\end{aligned}$$

so $\frac{dV}{d\tau}$ is negative definite. If the set $\mathcal{D}_C$ is defined as $\mathcal{D}_C = \{\hat{\mathbf{w}}\}$, then the Lyapunov analysis shows that trajectories of the associated ODE will converge to $\mathcal{D}_C$ as $k \to \infty$, and $\hat{\mathbf{w}}$ is a globally asymptotically stable equilibrium point of the associated ODE.

B.5.3 Statement and Satisfaction of the Sufficient Conditions

In [37], several conditions are described that allow the results of the associated ODE analysis to be applied to the adaptive algorithm. In addition to the update law given by Equation B.89, Ljung refers to the system

$$\chi(k) = \mathbf{A}(\mathbf{w}(k))\chi(k-1) + \mathbf{b}(\mathbf{w}(k))x(k) \tag{B.119}$$

in the statement and proof of the ODE Theorem. Its derivation is presented to show that the structure of the filtered-x LMS algorithm is suitable for application of the ODE Theorem, and to show that the conditions of the ODE Theorem are satisfied.

Recall that

$$d(k) = G_{out}(q^{-1})G_{ref}(q^{-1})x(k) \tag{B.120}$$

$$\hat{d}(k) = G_{out}(q^{-1})G_{ff}(k,q^{-1})x(k) \tag{B.121}$$

$$\hat{h}(k) = \hat{G}_{out}(q^{-1})x(k) \tag{B.122}$$

where

$$G_{ff}(k,q^{-1}) = w_0(k) + w_1(k)q^{-1} + \ldots + w_L(k)q^{-L}. \tag{B.123}$$

Define the following numerator and denominator polynomials

$$G_{out}(q^{-1}) = \frac{N_{out}(q^{-1})}{D_{out}(q^{-1})} \tag{B.124}$$

$$G_{ref}(q^{-1}) = \frac{N_{ref}(q^{-1})}{D_{ref}(q^{-1})} \tag{B.125}$$

$$\hat{G}_{out}(q^{-1}) = \frac{\hat{N}_{out}(q^{-1})}{\hat{D}_{out}(q^{-1})} \tag{B.126}$$

so that

$$D_{out}(q^{-1})D_{ref}(q^{-1})d(k) = N_{out}(q^{-1})N_{ref}(q^{-1})x(k) \tag{B.127}$$

$$D_{out}(q^{-1})\hat{d}(k) = N_{out}(q^{-1})G_{ff}(k,q^{-1})x(k) \tag{B.128}$$

$$\hat{D}_{out}(q^{-1})\hat{h}(k) = \hat{N}_{out}(q^{-1})x(k). \tag{B.129}$$

Assume, without loss of generality, that $\hat{D}_{out}(q^{-1})$, $D_{out}(q^{-1})$, and $D_{ref}(q^{-1})$ are monic. Thus, Equations B.127-B.129 are of the form

$$D(q^{-1})y(k) = N(q^{-1})x(k) \tag{B.130}$$

or

$$\begin{aligned}\left[1 + d_1q^{-1} + d_2q^{-2} + ... + d_nq^{-n}\right] y(k) = \\ [n_0 + n_1q^{-1} + n_2q^{-2} + ... + n_mq^{-m}]x(k).\end{aligned} \tag{B.131}$$

Therefore

$$\begin{aligned} y(k) &= -[d_1q^{-1} + d_2q^{-2} + ... + d_nq^{-n}]y(k) \\ &\quad + [n_1q^{-1} + n_2q^{-2} + ... + n_mq^{-m}]x(k) + n_0x(k) \end{aligned} \tag{B.132}$$

which can be put in the state space form

$$\left[\begin{array}{c} \mathbf{y}(k) \\ \hline \mathbf{x}_m(k) \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A_y} & \mathbf{N_y} \\ \hline \mathbf{0}_{m\times n} & \mathbf{L}_{m\times m} \end{array}\right] \left[\begin{array}{c} \mathbf{y}(k-1) \\ \hline \mathbf{x}_m(k-1) \end{array}\right] + \left[\begin{array}{c} \mathbf{b_y} \\ \hline \mathbf{e}_m \end{array}\right] x(k). \tag{B.133}$$

The vectors of Equation B.133 are defined as

$$\mathbf{y}(k) = [y(k), y(k-1), ..., y(k-n+1)]^T \tag{B.134}$$

$$\mathbf{x}_m(k) = [x(k), x(k-1), ..., x(k-m+1)]^T \tag{B.135}$$

$$\mathbf{b_y} = [n_0, 0, ..., 0]^T \tag{B.136}$$

$$\mathbf{e}_m = [1, 0, ..., 0]^T \tag{B.137}$$

where $\mathbf{b_y}$ is length n and $\mathbf{e}_m$ is length m. The submatrices of Equation B.133 are defined as

$$\mathbf{A_y} = \begin{bmatrix} -d_1 & -d_2 & -d_3 & \cdots & -d_{n-2} & -d_{n-1} & -d_n \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \tag{B.138}$$

$$\mathbf{N_y} = \begin{bmatrix} n_1 & n_2 & n_3 & \cdots & n_{m-2} & n_{m-1} & n_m \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{bmatrix} \tag{B.139}$$

$$\mathbf{L}_{m\times m} = \begin{bmatrix} 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 1 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 1 & 0 & \cdots & 0 & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & 0 & 0 \\ 0 & 0 & 0 & \ddots & 1 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 1 & 0 \end{bmatrix} \tag{B.140}$$

$$\tag{B.141}$$

and $\mathbf{0}_{m\times n}$ is a $m \times n$ matrix of zeros. In particular,

$$\left[\begin{array}{c} \mathbf{d}(k) \\ \hline \mathbf{x}_{m_d}(k) \end{array}\right] = \left[\begin{array}{c|c} \mathbf{A_d} & \mathbf{N_d} \\ \hline \mathbf{0}_{m_d\times n_d} & \mathbf{L}_{m_d\times m_d} \end{array}\right] \left[\begin{array}{c} \mathbf{d}(k-1) \\ \hline \mathbf{x}_{m_d}(k-1) \end{array}\right] + \left[\begin{array}{c} \mathbf{b_d} \\ \hline \mathbf{e}_{m_d} \end{array}\right] x(k) \tag{B.142}$$

$$\left[\begin{array}{c} \hat{\mathbf{d}}(k) \\ \hline \mathbf{x}_{\hat{m}_d}(k) \end{array}\right] = \left[\begin{array}{c|c} \hat{\mathbf{A}}_\mathbf{d} & \hat{\mathbf{N}}_\mathbf{d} \\ \hline \mathbf{0}_{\hat{m}_d\times \hat{n}_d} & \mathbf{L}_{\hat{m}_d\times \hat{m}_d} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{d}}(k-1) \\ \hline \mathbf{x}_{\hat{m}_d}(k-1) \end{array}\right] + \left[\begin{array}{c} \hat{\mathbf{b}}_\mathbf{d} \\ \hline \mathbf{e}_{\hat{m}_d} \end{array}\right] x(k) \tag{B.143}$$

$$\left[\begin{array}{c} \hat{\mathbf{h}}(k) \\ \hline \mathbf{x}_{\hat{m}_h}(k) \end{array}\right] = \left[\begin{array}{c|c} \hat{\mathbf{A}}_\mathbf{h} & \hat{\mathbf{N}}_\mathbf{h} \\ \hline \mathbf{0}_{\hat{m}_h\times \hat{n}_h} & \mathbf{L}_{\hat{m}_h\times \hat{m}_h} \end{array}\right] \left[\begin{array}{c} \hat{\mathbf{h}}(k-1) \\ \hline \mathbf{x}_{\hat{m}_h}(k-1) \end{array}\right] + \left[\begin{array}{c} \hat{\mathbf{b}}_\mathbf{h} \\ \hline \mathbf{e}_{\hat{m}_h} \end{array}\right] x(k) \tag{B.144}$$

where

$$\mathbf{d}(k) = [d(k), d(k-1), ..., d(k-n_d+1)]^T \tag{B.145}$$

$$\hat{\mathbf{d}}(k) = [\hat{d}(k), \hat{d}(k-1), ..., \hat{d}(k-\hat{n}_d+1)]^T \tag{B.146}$$

$$\hat{\mathbf{h}}(k) = [\hat{h}(k), \hat{h}(k-1), ..., \hat{h}(k-\hat{n}_h+1)]^T. \tag{B.147}$$

Equations B.142-B.144 can be combined into the complete system

$$\chi(k) = \mathbf{A}(\mathbf{w}(k))\chi(k-1) + \mathbf{b}(\mathbf{w}(k))x(k) \tag{B.148}$$

where

$$\chi(k) = \left[\begin{array}{c} \mathbf{d}(k) \\ \hline \hat{\mathbf{d}}(k) \\ \hline \hat{\mathbf{h}}(k) \\ \hline \mathbf{x}_{m_{max}}(k) \end{array}\right] \tag{B.149}$$

$$\mathbf{A}(\mathbf{w}(k)) = \left[\begin{array}{c|c|c|cc} \mathbf{A_d} & \mathbf{0}_{m_d \times \hat{m}_d} & \mathbf{0}_{m_d \times \hat{m}_h} & \mathbf{N_d} & \mathbf{0}_{m_d \times (m_{max} - m_d)} \\ \hline \mathbf{0}_{\hat{m}_d \times m_d} & \hat{\mathbf{A}}_\mathbf{d} & \mathbf{0}_{\hat{m}_d \times \hat{m}_h} & \hat{\mathbf{N}}_\mathbf{d}(\mathbf{w}(k)) & \mathbf{0}_{\hat{m}_d \times (m_{max} - \hat{m}_d)} \\ \hline \mathbf{0}_{\hat{m}_h \times m_d} & \mathbf{0}_{\hat{m}_h \times \hat{m}_d} & \hat{\mathbf{A}}_\mathbf{h} & \hat{\mathbf{N}}_\mathbf{h} & \mathbf{0}_{\hat{m}_h \times (m_{max} - \hat{m}_h)} \\ \hline \mathbf{0}_{m_{max} \times m_d} & \mathbf{0}_{m_{max} \times \hat{m}_d} & \mathbf{0}_{m_{max} \times \hat{m}_h} & \multicolumn{2}{c}{\mathbf{L}_{m_{max} \times m_{max}}} \end{array}\right] \tag{B.150}$$

$$\mathbf{b}(\mathbf{w}(k)) = \left[\begin{array}{c} \mathbf{b_d} \\ \hline \hat{\mathbf{b}}_\mathbf{d}(\mathbf{w}(k)) \\ \hline \hat{\mathbf{b}}_\mathbf{h} \\ \hline \mathbf{e}_{m_{max}} \end{array}\right] \tag{B.151}$$

and the value of m_{max} is the maximum of m_d, $\hat{m}_d$, and $\hat{m}_h$. This system definition will be used in the proof that follows. Note that the eigenvalues of $\mathbf{A}(\mathbf{w}(k))$ are the eigenvalues of $\mathbf{A_d}$, $\hat{\mathbf{A}}_\mathbf{d}$, $\hat{\mathbf{A}}_\mathbf{h}$ and $\mathbf{L}$. Furthermore, the eigenvalues of $\mathbf{A_d}$ equal the eigenvalues of $G_{out}(q^{-1})G_{ref}(q^{-1})$, the eigenvalues of $\hat{\mathbf{A}}_\mathbf{d}$ equal the eigenvalues of $G_{out}(q^{-1})G_{ff}(k, q^{-1})$, and the eigenvalues of $\hat{\mathbf{A}}_\mathbf{h}$ equal the eigenvalues of $\hat{G}_{out}(q^{-1})$. Since $G_{ff}(k, q^{-1})$ is an FIR filter, it does not affect the eigenvalues. Therefore, the eigenvalues of $G_{out}(q^{-1})G_{ff}(k, q^{-1})$ are just the eigenvalues of $G_{out}(q^{-1})$, which implies that the eigenvalues of $\mathbf{A}(\mathbf{w}(k))$ are independent of the estimated parameters $\mathbf{w}(k)$. The eigenvalues of $\mathbf{L}$ are all zero.

Let the set $\mathcal{D}_S$ be defined as $\mathcal{D}_S = \{\mathbf{w} \mid \mathbf{A}(\mathbf{w}) \text{ is stable}\}$. Therefore, $\mathcal{D}_S$ is the set of all finite values of $\mathbf{w}$ if $\hat{G}_{out}(q^{-1})$, $G_{out}(q^{-1})$, and $G_{ref}(q^{-1})$ are stable. The stability of $\hat{G}_{out}(q^{-1})$, $G_{out}(q^{-1})$, and $G_{ref}(q^{-1})$ will be assumed in the following. The set $\mathcal{D}_R$ is defined to be an open, connected subset of $\mathcal{D}_S$ such that the Conditions 1 through 9 hold within $\mathcal{D}_R$. Sufficient conditions for applying the ODE Theorem are

1. $x(k)$ is a sequence of independent random variables

2. $\|x(k)\| <$, with probability one for all k and for some , > 0
3. The function $\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k)$ is continuously differentiable with respect to $\mathbf{w}$ and χ for $\mathbf{w} \in \mathcal{D}_R$. The derivatives are bounded in k for fixed $\mathbf{w}$ and χ.
4. The matrix functions $\mathbf{A}(\mathbf{w})$ and $\mathbf{b}(\mathbf{w})$ are Lipschitz continuous in $\mathcal{D}_R$.
5. $\lim_{k\to\infty} E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k)]$ exists for $\mathbf{w} \in \mathcal{D}_R$, where the expectation is over $x(k)$
6. $\sum_{k=1}^{\infty} \gamma(k) = \infty$
7. $\sum_{k=1}^{\infty} \gamma^p(k) < \infty$ for some $p > 0$
8. $\gamma(k)$ is a decreasing sequence
9. $\lim_{k\to\infty} \sup(\frac{1}{\gamma(k)} - \frac{1}{\gamma(k-1)}) < \infty$

Condition 1 requires that the input be stochastic. Although Condition 2 is reasonable in practice, it does exclude Gaussian signals. Recall that $\hat{G}_{out}(q^{-1})$, $G_{out}(q^{-1})$, $G_{ref}(q^{-1})$, and $G_{ff}(k, q^{-1})$ represent stable, linear systems, then Condition 3 follows. Since $\mathbf{A}(\mathbf{w})$ and $\mathbf{b}(\mathbf{w})$ are linear in $\mathbf{w}$, Condition 4 is met. Condition 5 holds because $x(k)$ is assumed to be stationary. Conditions 6 through 9 can be satisfied by choosing $\gamma(k) = \beta k^{-\alpha}$ for $\beta > 0$ and $0 < \alpha \leq 1$.

Ljung [37] lists a second set of conditions that can be applied if a wider range of stochastic inputs is desired, including Gaussian. These conditions are similar to those listed above, but replace Conditions 2 and 3 with four, more-complicated conditions on $\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k)$ and $x(k)$. A third set of conditions exists for the case where $x(k)$ is deterministic. A more thorough discussion of these sets of conditions and their applicability for various adaptation algorithms may be found in [39].

B.5.4 Statement and Application of the ODE Convergence Theorem

Once the associated ODE has been derived and analyzed, and the conditions satisfied, the following theorem may be applied to determine the behavior of the adaptive algorithm.

ODE Theorem Consider the algorithm of Equations B.89 and B.148, subject to Conditions 1 through 9. Assume

A. there is a compact subset $\bar{\mathcal{D}}$ of $\mathcal{D}_R$ such that the trajectories of the associated ODE of Equation B.91 that start in $\bar{\mathcal{D}}$ remain in a closed subset $\bar{\mathcal{D}}_R$ of $\mathcal{D}_R$ for $\tau > 0$

B. there is a random variable C such that $\mathbf{w}(k) \in \bar{\mathcal{D}}$ and $\|\chi(k)\| < C$ infinitely often with probability one

C. the associated ODE has an invariant set $\mathcal{D}_C$ with a domain of attraction $\mathcal{D}_A \supset \bar{\mathcal{D}}$.

Then $\mathbf{w}(k) \to \mathcal{D}_C$ with probability one as $k \to \infty$.

For the proof of this theorem, see [37]. By the Lyapunov analysis, the sets $\bar{\mathcal{D}}$ and $\bar{\mathcal{D}}_R$ exist and Condition A is satisfied. Furthermore, $\hat{\mathbf{w}}$ is an equilibrium point, and therefore an invariant set, and since it is globally asymptotically stable, its domain of attraction is all of $\mathcal{R}^{L+1}$. Thus $\bar{\mathcal{D}} \subset \mathcal{D}_A$, and Condition C holds. Since $x(k)$ is assumed to be bounded by Condition 2, and $\hat{G}_{out}(q^{-1})$, $G_{out}(q^{-1})$, $G_{ff}(k, q^{-1})$, and $G_{ref}(q^{-1})$ are stable, Condition B is satisfied. Additionally, as discussed in [37], if $\mathcal{D}_R = \mathcal{D}_A = \mathcal{R}^{L+1}$, then Condition B may be met by introducing saturation in the update law. Ljung even suggests that this may be accomplished automatically by computer overflow. Furthermore, Metivier and Priouret [43] have shown that this boundedness condition may be dropped in algorithms that are linear in the estimated parameters, such as the filtered-x LMS algorithm.

Both the ODE Theorem and the technique of Widrow and Stearns show convergence of the filtered-x LMS algorithm tap weights to $\hat{\mathbf{w}}$. The ODE result is stronger, with convergence with probability one compared to convergence of the mean for the Widrow and Stearns technique. However, the ODE method requires, among other conditions, that the adaptation gain $\gamma(k) \to 0$ as $k \to \infty$, which the Widrow and Stearns technique does not. Although the ODE method is for the case of vanishing adaptation gain, Ljung suggests that it is reasonable to assume that the behavior of the adaptive algorithm for small, nonzero, adaptation gains is similar [37]. Recall that the Widrow and Stearns technique places a limit on the magnitude of the adaptation gain, and that the justification for the commutivity of $G_{out}(q^{-1})$ and $G_{ff}(k, q^{-1})$ is that the controller parameters are slowly varying.

B.6 Analysis of the Mean-Square Error

Since the expected values of the tap weights have been shown to converge to $\hat{\mathbf{w}} = \hat{\mathbf{R}}^{-1}\hat{\mathbf{p}}$, it is of interest to see how this value differs from the optimal, and how it effects the mean-square error. The mean-square error with the converged values of the filtered-x LMS algorithm is Equation B.35 with

$$\mathbf{w} = \hat{\mathbf{w}} = \mathbf{w}^* + (\hat{\mathbf{w}} - \mathbf{w}^*) = \mathbf{w}^* + \boldsymbol{\Delta} w \tag{B.152}$$

or

$$\begin{aligned}
\xi &= E[d^2(k)] + (\mathbf{w}^* + \boldsymbol{\Delta} w)^T \mathbf{R}(\mathbf{w}^* + \boldsymbol{\Delta} w) - 2\mathbf{p}^T(\mathbf{w}^* + \boldsymbol{\Delta} w) && \text{(B.153)} \\
&= E[d^2(k)] + \mathbf{w}^{*T}\mathbf{R}\mathbf{w}^* + 2\mathbf{w}^{*T}\mathbf{R}\boldsymbol{\Delta} w + \boldsymbol{\Delta} w^T \mathbf{R}\boldsymbol{\Delta} w \\
&\quad - 2\mathbf{p}^T\mathbf{w}^* - 2\mathbf{p}^T\boldsymbol{\Delta} w && \text{(B.154)} \\
&= E[d^2(k)] + \left(\mathbf{w}^{*T}\mathbf{R}\mathbf{w}^* - 2\mathbf{p}^T\mathbf{w}^*\right) \\
&\quad + \left(2\mathbf{w}^{*T}\mathbf{R}\boldsymbol{\Delta} w + \boldsymbol{\Delta} w^T\mathbf{R}\boldsymbol{\Delta} w - 2\mathbf{p}^T\boldsymbol{\Delta} w\right) && \text{(B.155)} \\
&= \xi_{min} + \boldsymbol{\Delta} w^T\mathbf{R}\boldsymbol{\Delta} w + 2\left(\mathbf{w}^{*T}\mathbf{R} - \mathbf{p}^T\right)\boldsymbol{\Delta} w. && \text{(B.156)}
\end{aligned}$$

As mentioned previously, if $\hat{G}_{out}(q^{-1}) = G_{out}(q^{-1})$, then $\boldsymbol{\Delta} w = \mathbf{0}$ and $\xi = \xi_{min}$. Ren and Kumar [53] suggest a second situation where $\xi = \xi_{min}$. From Equation B.39, the optimal parameters $\mathbf{w}^*$ satisfy

$$\mathbf{R}\mathbf{w}^* - \mathbf{p} = \mathbf{0}. \tag{B.157}$$

Let $H(q^{-1})$ be defined as before and let h_i be such that

$$H(q^{-1}) = \sum_{i=0}^{\infty} h_i q^{-i}. \tag{B.158}$$

Then

$$\mathbf{R}\hat{\mathbf{w}} - \mathbf{p} = \sum_{i=1}^{\infty} \frac{h_i}{h_0} E[\{G_{out}(q^{-1})x(k-i)\varepsilon(k, \hat{\mathbf{w}})] \tag{B.159}$$

since from Equation B.60

$$\hat{\mathbf{R}}\hat{\mathbf{w}} = \hat{\mathbf{p}} \tag{B.160}$$

$$E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T\hat{\mathbf{w}}] = E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}d(k)] \tag{B.161}$$

$$E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\left(\{G_{out}(q^{-1})\mathbf{x}(k)\}^T\hat{\mathbf{w}} - d(k)\right)] = \mathbf{0} \tag{B.162}$$

$$E[\{\hat{G}_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k, \hat{\mathbf{w}})] = \mathbf{0} \tag{B.163}$$

$$H(q^{-1})E[\{G_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k, \hat{\mathbf{w}})] = \mathbf{0} \tag{B.164}$$

$$\sum_{i=0}^{\infty} h_i q^{-i} E[\{G_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k, \hat{\mathbf{w}})] = \mathbf{0} \tag{B.165}$$

and splitting the summation results in

$$h_0 E[\{G_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k, \hat{\mathbf{w}})] = -\sum_{i=1}^{\infty} h_i q^{-i} E[\{G_{out}(q^{-1})\mathbf{x}(k)\}\varepsilon(k, \hat{\mathbf{w}})] \tag{B.166}$$

$$\begin{aligned} &h_0 E[\{G_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T\hat{\mathbf{w}} - \{G_{out}(q^{-1})\mathbf{x}(k)\}d(k)] \\ &\qquad = \sum_{i=1}^{\infty} h_i E[\{G_{out}(q^{-1})\mathbf{x}(k-i)\}\varepsilon(k, \hat{\mathbf{w}})] \end{aligned} \tag{B.167}$$

$$\begin{aligned} &E[\{G_{out}(q^{-1})\mathbf{x}(k)\}\{G_{out}(q^{-1})\mathbf{x}(k)\}^T]\hat{\mathbf{w}} - E[\{G_{out}(q^{-1})\mathbf{x}(k)\}d(k)] \\ &\qquad = \sum_{i=1}^{\infty} \frac{h_i}{h_0} E[\{G_{out}(q^{-1})\mathbf{x}(k-i)\}\varepsilon(k, \hat{\mathbf{w}})] \end{aligned} \tag{B.168}$$

which is the desired result. Note that if $\hat{G}_{out}(q^{-1}) = G_{out}(q^{-1})$, then $H(q^{-1}) = h_0 = 1$ and the right hand side of Equation B.159 is zero. Solving for $\hat{\mathbf{w}}$ results in

$$\hat{\mathbf{w}} = \mathbf{R}^{-1}\mathbf{p} = \mathbf{w}^* \tag{B.169}$$

as expected. The second condition suggested by Ren and Kumar occurs when $\mathbf{w}^*$ achieves complete cancellation of the error component that is correlated with the input $x(k)$. In this case,

$$E[\varepsilon(k, \mathbf{w}^*)x(k')] = 0 \tag{B.170}$$

for all $k \neq k'$. Then

$$E[\varepsilon(k, \mathbf{w}^*)\mathbf{x}(k-i)] = \mathbf{0} \tag{B.171}$$

for all $i \neq 0$, which implies

$$E[\varepsilon(k, \mathbf{w}^*)\{G_{out}(q^{-1})\mathbf{x}(k-i)\}] = \mathbf{0} \tag{B.172}$$

for all $i \neq 0$, so $\mathbf{w}^*$ is a solution of Equation B.159 even if $\hat{G}_{out}(q^{-1}) \neq G_{out}(q^{-1})$. Since $\mathbf{R}$ is assumed to be positive definite, Equation B.159 has a unique solution, and therefore $\hat{\mathbf{w}}$ must equal $\mathbf{w}^*$.