Linear and nonlinear compensation techniques for control of a single degree of freedom magnetic bearing

by

Claudio Salvatore

B.S. Mechanical Engineering
Illinois Institute of Technology
December 1995

Submitted to the Department of Mechanical Engineering
in partial fulfillment of the requirements for the degree of

Master of Science in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

June 1998

Author .....
Department of Mechanical Engineering
May 8, 1998

Certified by .....
David L. Trumper
Rockwell International Associate Professor of Mechanical Engineering
Thesis Supervisor

Accepted by .....
Ain A. Sonin
Chairman, Department Committee on Graduate Students

Linear and nonlinear compensation techniques for control of a single degree of freedom magnetic bearing

by

Claudio Salvatore

Submitted to the Department of Mechanical Engineering
on May 8, 1998, in partial fulfillment of the
requirements for the degree of
Master of Science in Mechanical Engineering

Abstract

Magnetic bearings have the desirable property that they allow frictionless relative motion between surfaces, which is greatly beneficial where precision motion control is required such as in photolithography stages. The test bed used in this thesis was built as a scaled-down version of a photolithography stage, but it is equally applicable as a test-bed for a wide variety of mechanical systems where positioning of a target is required. The system exhibits behaviours which can be found in any mechanical system namely resonances, cross-coupling between axes, actuator saturation and time delays, so the control strategies and limitations presented in this thesis can be extrapolated and used for other systems.

After using bond graphs and electromagnetic theory to model the system, various controllers are designed to close the loop on the position of the target. The electromagnetic actuators used introduce non-linearities in the system, while a linear capacitance probe is used as the sensor for the position of the target. At first a classical linear controller is used on the system and its limitations are discussed. Next a feedback linearizing controller is used to cancel the system nonlinearities to achieve, at least in theory, a linear system independent of the operating point. The third approach used is sliding control which both takes into account the non-linear nature of the system and by construction makes the system robust to parameter variation.

In the last two chapters we investigate, through simulations, two possible control strategies for meeting performance specifications on the stepping motion of a photolithography stage subject to ground vibrations, sensor quantization, actuator saturation and Coulomb friction. The first controller is designed using classical control theory after the plant has been decoupled. In the second part, optimal control is used to close the loop but since not all states are available for measurement a Kalman Filter is designed to obtain the state estimates.

Thesis Supervisor: David L. Trumper
Title: Rockwell International Associate Professor of Mechanical Engineering

Acknowledgments

I would like to start by thanking my advisor Prof. David L. Trumper for giving me the opportunity to join his lab. I consider this experience as being my formal introduction to the world of engineering and it has helped me see what I like and dislike about the path I've followed so far. He has set a very high academic standard for me to look up to, one that will be very hard to match.

I also wish to thank all my lab-partners in the Precision Motion Control Laboratory (PMC), whom I have often bothered with questions. They are undoubtedly some of the most dedicated and hard-working people I have encountered in my academic career and I wish them the success they deserve.

A special thank you goes to Prof. David E. Hardt for two great semesters where I had the privilege of serving as the Teaching Assistant for his graduate course "Control of Manufacturing Systems". Having been a student all my life I finally had the chance to see what goes on behind the scenes of grading homeworks and exams. It was then that I began to appreciate the importance of quality over quantity in written material. I am very grateful for the chances I got to lecture: seeing the entire class take notes on what I said and wrote on the board, made me realize the responsibility professors have, and their importance in shaping our lives.

Over the past couple years here at MIT I have met many people who have helped me get my mind off of work and remind me of the existence of other values in life. They are too many to name, but to them goes my deepest gratitude for all those enjoyable moments, deep talks, laughter and child-like playing.

During my undergraduate years at IIT in Chicago I was very fortunate to become friends with Prof. Serope Kalpakjian, himself an MIT alum. He has acted as a "life" advisor in addition to giving me academic advice and I am glad that his friendship continues to this day. I thank him for the countless hours spent chatting over equally countless glasses of wine (mostly red).

My dear mother and father, I will always be grateful for your undying support and faith in me. You were with me every step of the way, helping me up when I

stumbled; sorry for all the headaches I caused you. My hope for the future is that I shall be able to take full advantage of all the lessons you have taught me.

Claudio Salvatore

Cambridge, Massachusetts
May 1998

To my parents: Giuseppe and Elena Salvatore

Contents

List of Figures

Chapter 1

Introduction

Magnetic bearings are bearings whose suspension forces are generated magnetically without contact between surfaces. The name "magnetic bearing" will appear often throughout this thesis so it is important to distinguish it from "magnetic suspension" and "magnetic levitation" which although closely related have a somewhat narrower meaning. By definition both suspension and levitation must exert forces on the target in a direction opposite that of gravity, but while magnetic suspensions use repulsive forces, magnetic levitation uses attractive forces. Magnetic bearings, on the other hand, do not have a specified direction in which the magnetic force must act. Thus if the application requires the bearing to exert a vertical force to allow horizontal movement then it overlaps with the definition of magnetic suspension/levitation. But the direction of actuation can also be horizontal as is the case of the test bed used in this thesis, or it can be radial as in gyroscopes and bearings for rotating machinery.

Since their introduction in the 1930's magnetic bearings have been used in a wide variety of applications such as robotics, interferometry, photolithography, spacecraft attitude control, disk drives, power generators, and many more. Over the past three decades there has been a growing interest in magnetic levitation/suspension of transportation vehicles and together with bearings for rotating machinery these two applications are the most common use of magnetic bearings. Because they feature no physical contact between surfaces they are virtually frictionless and induce little heating. They do not wear or fail from metal fatigue and produce no audible vi-

bration noise. Likewise they do not require extensive lubrication systems, another source of mechanical bearing failure and machinery cost. Particulate contamination is avoided because normal working conditions typically use air gaps which are greater than $50\,\mu$m and minor precautions should be taken when the bearing will be made to operate at air gaps of the order of micrometers. The interest in this field has also been stirred up by advances in materials and electronics which have enabled designers to increase the force density of the actuators which has led to simpler, smaller, actuators. Most important is the fact that magnetic bearings can function both as precision actuators and as bearings; i.e., in addition to allowing frictionless relative movement they can also be modulated to exert activation forces on the target. Their properties can be changed online and this renders them particularly useful in vibration isolation [30, 6, 23]. The precision of magnetic bearings is limited only by sensor quantization and external noise, although if the actuators are driven to saturation precise control may become hard to achieve. The actuation bandwidth is usually limited only by the servo-control bandwidth.

In this thesis we focus in particular on the application of magnetic bearings to wafer steppers in photolithography. The test bed used was built as a scaled-down version of a photolithography stage however the control strategies derived, and the approaches adopted to deal with resonances, friction and so on, are quite general and can be extrapolated and used for other systems where accurate positioning of a target is required. The difficulties of achieving a stable system are highlighted by an examination of the inverse square law relating force to distance. Earnshaw [27], in his classic paper, shows mathematically that there is no stable equilibrium point for a pole placed in a static magnetic field, hence the need for feedback in order to stabilize the system.

The advent of quick, inexpensive, powerful computers and dedicated DSP chips has shifted automatic control from continuous-time implementations to digital implementations. The advantages offered by the latter are numerous, but the most important is probably the ease of altering the controller. Most control applications today use programmable computers and/or programmable logic chips to run the con-

trol algorithm. When the need arises for the controller to be altered one needs only to type in new parameters in the Graphical User Interface or maybe enter these new parameters in the control code, rather than having to re-wire the controller as might be required if the controller were made of op-amps and RLC elements. This allows for faster prototyping and experimentation. But use of digital computers has also greatly simplified operations that used to be hard to perform with continuous time controllers, and has made possible the implementation of control strategies which until then were considered to be purely of academic interest.

Control of magnetic bearings has been investigated quite extensively as indicated by the amount of literature on the subject. A lot of it presents continuous time controllers, but more and more is appearing in scientific journals detailing digital control approaches. The remainder of this section will give a brief overview of some of the recent applications and techniques used in magnetic bearings, which will help put the work done in this thesis into context.

A good place to start looking for information on magnetic bearings is the paper by Goodall [26]. Although this paper is directed towards Maglev, the analysis is strongly based upon a consideration of the suspension transfer functions and so many of the principles are applicable to actively controlled suspensions and bearings in general. In this regard we also mention the book by Jayawant [3] which, as well as providing a good historical introduction to magnetic suspensions/bearings, also details a wide variety of techniques for control. An interesting approach which is sometimes used for Maglev vehicles, and that we considered as a possible control strategy for our magnetic bearing, is flux feedback [25, 3]. The idea behind this technique is that by closing an inner loop on the flux density within the air gap the force-distance characteristics will be very nearly linear and independent of the operating point. This approach, however, was abandoned because deemed impractical for our setup.

The literature presents control techniques for suspensions having permanent magnets providing the bias force to balance out gravity at the nominal operating point. A lot of research has gone into this direction and many industrial applications of electromagnets do have a permanent magnet in the flux loop. The advantages with

this hybrid design is that the control currents (incremental currents) needed can be orders of magnitude less than the primary magnetic field and control currents in an all-electromagnetic setup. With the power consumption reduced by as much as 50 - 100 times, hybrid systems can be made much smaller and cost less to operate. The advent of high energy rare earth permanent-magnet materials (typically neodymium boron iron or samarium cobolt, depending on the operating temperatures) greatly increased the attractiveness of hybrid electromagnetic systems. Adding a bias force which is much greater than the incremental changes thereof, and allowing only incremental displacements from a nominal operating point, linearizes the system about that operating point and makes control more amenable using linear control theory. This is the most widely used compensation technique for magnetic suspension/bearings.

A very important industrial application in which electromagnets are called upon to increase the efficiency of the process, is in rotating machinery. As a way to eliminate friction in the rotor shaft radial bearings are being designed and integrated into the machines which allows for more benefits than backfitting these bearings into an existing piece of equipment. However, at rotational frequencies close to resonant frequencies of the shaft and/or structure, or under large loads, radial bearings run into stabilization problems because the flexibility of the shaft introduces self-excited vibrations. Finite Element Analysis, lumped parameter models, modal analysis and model reduction techniques have all been used to characterize these resonant states and radial magnetic bearing are being used both as bearings and as actuators to damp out these vibrations [9, 18, 16, 4].

The last application we will mention is the use of electromagnetic actuators as thrust bearings. Used in pairwise opposition where the target is typically a thrust disk, or runner on a shaft, they serve to prevent axial movement particularly of rotors used in turbomachinery [24]. It is this application which most resembles the setup used in this thesis, however while thrust bearings serve mostly to perform as regulators, the actuators in this thesis will be used to impart desired trajectories on the target.

1.1 Thesis Overview

We will begin chapter two by describing the hardware used and the software developed as part of the thesis. We will proceed to characterize the system by both estimating and then determining experimentally, flexure stiffness, actuator constants as well as in-plane and out-of-plane resonant frequencies. From the model of the electromagnets derived herein we will proceed to derive the state equations for the bearing which we use in subsequent chapters for controller design.

Before implementing nonlinear controllers, chapter three is devoted to testing the performance of linear controllers when used on the plant. Here we linearize the dynamics about an operating point and stabilize the system with a lead-lag compensator. Using a dynamic analyzer we are able to measure the loop transmission experimentally and verify our plant model. Several linear compensators are derived, some in discrete-time and others are mapped from continuous-time. The compensators are applied to the system and the responses compared to the simulated behaviour. We end the chapter by looking at limitations of these linear actuators and we show experimentally the effect of unmodeled mass and operating point variations.

In chapter four we present feedback linearization as a way to deal with the nonlinearities of the system. We derive the set of linearizing transformations and apply one of the controllers used in chapter three. We vary the operating point and introduce unmodeled mass variations to see the effects of these.

The control strategies presented thus far all suffer from lack of robustness to parametric uncertainties/variations. To overcome this, in chapter five we introduce sliding control which is a robust control technique. Experimentally we verify the closed loop system's robustness to parametric uncertainties.

The last two chapters are dedicated to designing controllers for a single degree of freedom wafer stepper. The stage is subject to performance specifications on the stepping motion and the control strategies must deal with ground vibrations, sensor quantization, actuator saturation and Coulomb friction. The first controller is designed using classical control theory after the plant has been decoupled. In the

second part, optimal control is used to close the loop but since not all states are available for measurement a Kalman Filter is designed to obtain the state estimates.

Chapter 2

Experimental hardware and modeling

This chapter introduces the experimental setup which is used for this thesis, and also illustrates the modeling procedure used to obtain a plant model useful for control design.

2.1 General overview of the system

Figure 2-1 shows a picture of the hardware setup with various components pointed out. The setup was originally built by a previous student, Sean Olson [28], for his Master thesis but the control electronics has greatly changed since then. The system is composed of the following parts :

1. Two electromagnetic actuators
2. Cantilevered arm with target on the end
3. Flexures
4. Capacitance gage including signal conditioning electronics
5. 16-bit A/D (with a 10 μs conversion time)

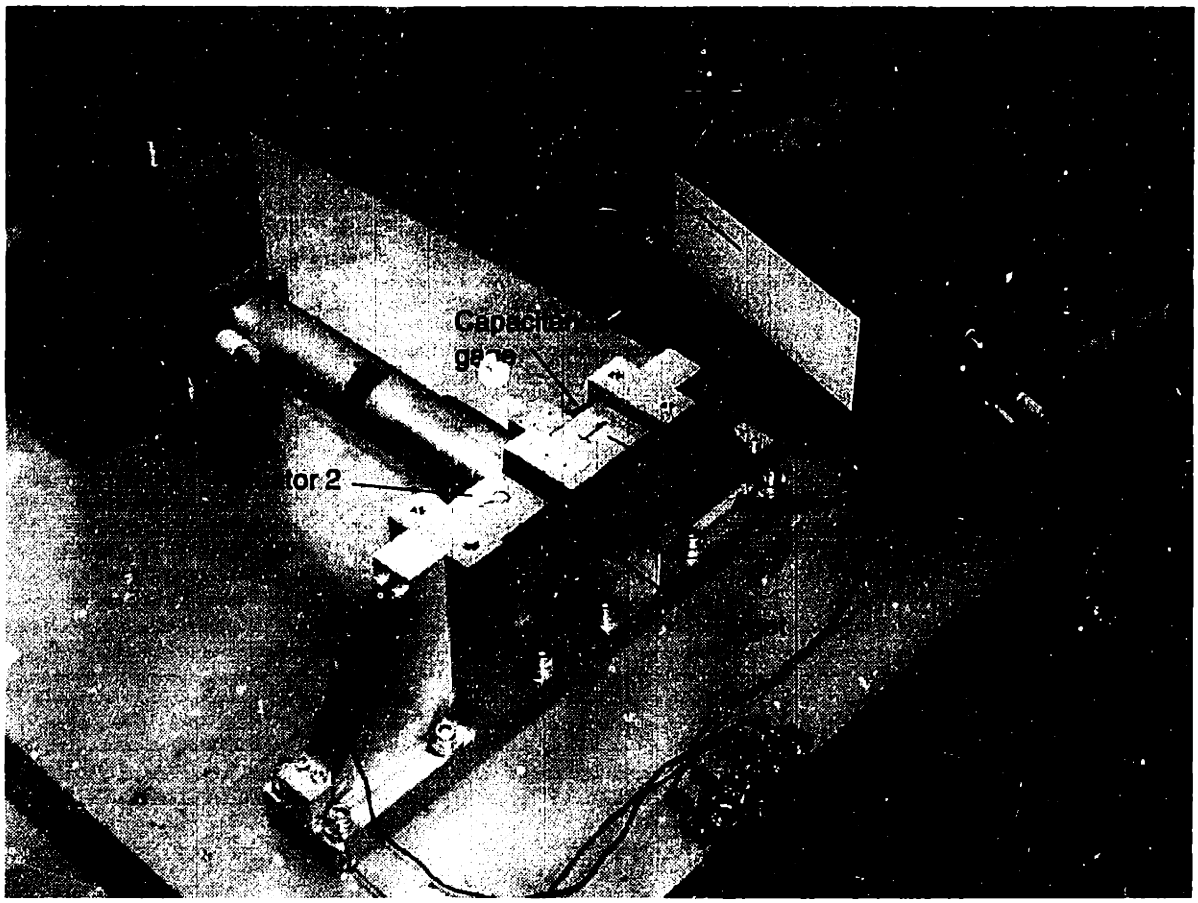

Figure 2-1: Photograph of the experimental setup.

6. 12-bit D/A converter (with a 15 μs average settling time)

7. Amplifier

8. Tiger 31/IP DSP board

9. Gateway 2000 P166 computer

10. HP signal generator

The rest of this chapter details each of these components and shows how they fit together. We also focus on the modeling of the system and analyze important system characteristics such as resonances, actuator constants and open loop instability.

To better understand, and expand upon, Figure 2-1 we show a diagram of the closed loop system in Figure 2-2. We note that the loop is closed by the capacitance gage from which we feedback position of the target.

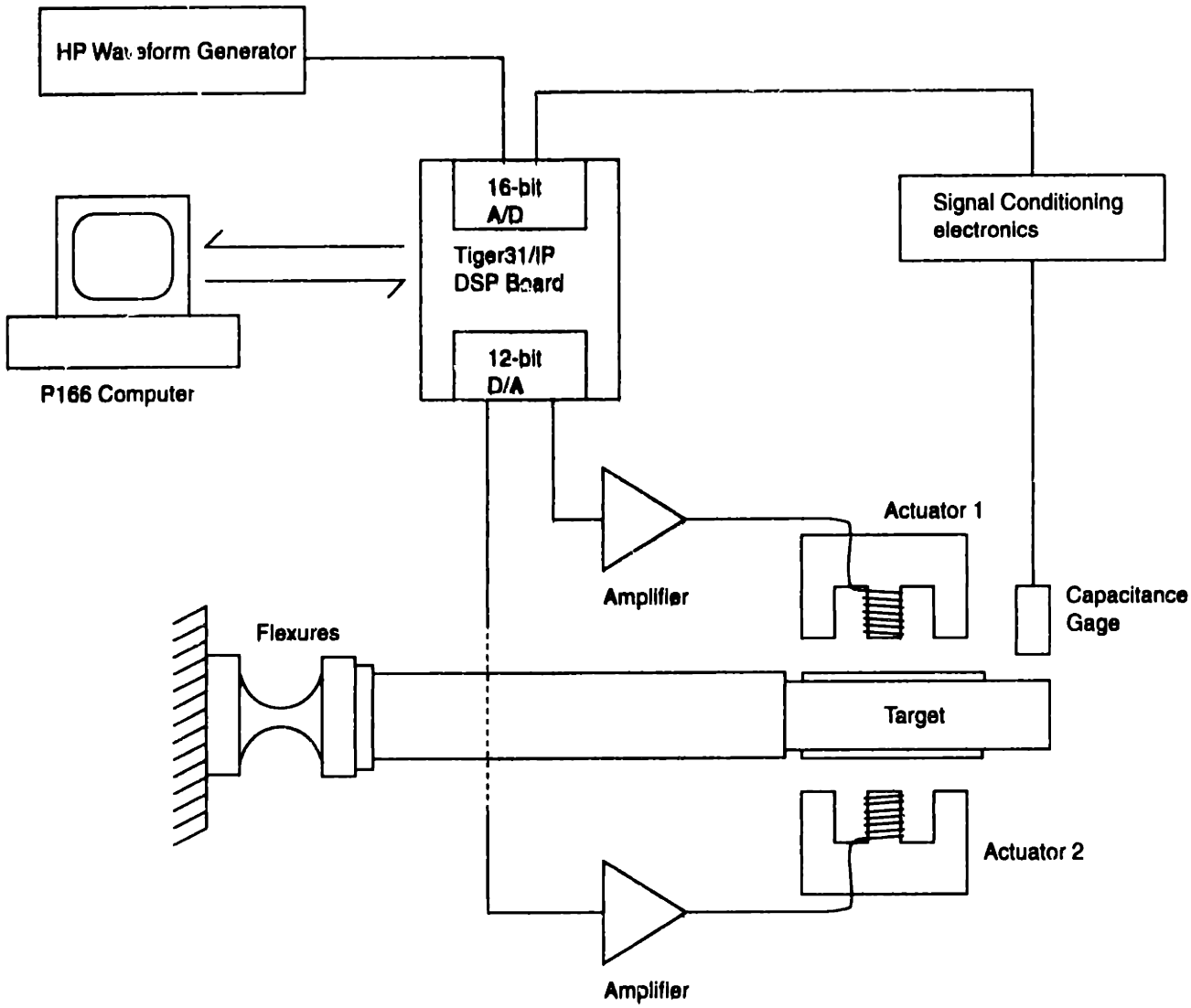

Figure 2-2: Diagram of the closed loop system.

The position signal is then sampled through the A/D and processed in the DSP before being output through the D/A. The effort signal is then amplified by the current amplifier which controls the current into the electromagnetic actuators. The plant has one output and we can command two plant inputs, namely the current to each actuator, thus the system is MISO. In the next two chapters we will present two ways to reduce the system to a SISO system.

2.2 Tiger 31/IP, computer and converters

A Gateway2000 Pentium 166 is used primarily for data storage and for a graphical user interface (GUI). All the control code and the sampling processes are executed on the DSP board which resides in the computer. The DSP board used is the Tiger 31/IP, designed by DSP Research Inc.[1] and based on the Texas Instrument TMS320C31 DSP chip. The Tiger 31 is a fully integrated 32-bit floating point DSP running at 48 MHz clock speed specifically designed for mathematical operations. A very attractive feature of this board which has contributed to our decision to adopt it for this project, is its modularity via the IP standard: there are four slots on the board itself into which it is possible to plug a wide variety of actuators, sensors and other devices. The IP-HiRes 16-bit A/D made by Wavetron[2] and the IP-DAC D/A by GreenSpring Computers[3] occupy two of these slots. Communication from the board to the plug-in modules occurs through an internal bus and therefore sampling and calculations are completely invisible to the computer. In addition, the DSP board has 512K of Static Random Access Memory (SRAM) and 4 Mb of Enhanced Dynamic Random Access Memory (EDRAM) which can be used for data storage while the system is running, and then the data can be dumped to the host computer and saved to disk. Using the EDRAM requires having to refresh the data stored in it at least every 64 ms. Altogether, the strengths of this DSP board can be summarized as follows :

- Powerful mathematical and I/O capabilities
- Independence from host computer
- IP plug-in hardware
- Optimizing compiler specific for the C31 chip
- Ease of operation

[1]DSP Research, Inc. 1095 East Duane Avenue, Suite 203 Sunnyvale, CA 94086
[2]Wavetron products are available through GreenSpring Computers.
[3]GreenSpring Computers, 1204 O'Brien Drive, Menlo Park, CA 94025.

With its on-board RAM and ROM the board could also be made to operate in stand-alone configuration if we provided a 12 V power supply. In this case we could run the control algorithm from the Boot ROM; but this approach was not taken as it does not allow for user interaction and easy change of controller parameters. In our setup the board sits inside the host computer from which it draws power and exchanges data. This arrangement requires two software programs to be written: one to run on the host and the other to run of the DSP. From now on, when it will be necessary to distinguish between the code running on the host and that running on the DSP we shall use the terms "host code" and "DSP code" respectively.

When the system is initialized, the first task of the host code is to initialize the board after which it downloads the control code to it. From then on the host code will act as the GUI from which the user can change the operation of the controller and save the data. Communication between the board and the computer occurs through a memory region called the Dual Port RAM (DPRAM). This memory region can be accessed simultaneously both by the computer and by the board, and is used primarily for data transfer.

A noticeable feature of this setup is that the control is entirely operating system independent. The control code is on the DSP and all transactions, with the exception of the data update on the screen, are initiated by the user through the GUI and therefore it is the computer which asserts the interrupts to the board when a transfer is needed and not the other way around. The update of the screen interface occurs at a rate of 10 Hz and is not interrupt-dependent. It simply occurs by having the interrupt service routine (ISR) on the DSP write the current data to the DPRAM every 0.1 seconds and using a timer in the host to read the DPRAM every 0.1 seconds or as soon after that as possible. We note that the data transfer could have been done by having the DSP write the data and assert a flag every 0.1 seconds while the host continuously monitored the flag and read the data when the flag went high. Either way, the screen update is non-real time but we are doing real-time control with the DSP board.

The host is running under Windows NT and this slows data display somewhat because Windows NT does not allow the user to access the hardware directly, therefore an extra layer of software had to be included. It was necessary to add the DriverX device-driver developed by Tetradyne Inc.[4] which enabled communication with the ports. The overall flow chart of the closed loop system for the process is shown in Figure 2-3.

2.3 Flexures

As part of the system characterization we briefly look at the flexures used to restrict the motion of the vibration shaft to one degree of freedom. A drawing of the flexure used is shown in Figure 2-4.

As the following section will detail, the system exhibits an open-loop resonance at 8 Hz. Since our aim is to achieve a closed loop bandwidth of the order of 100-150 Hz, the system is modeled as a free mass since at these frequencies the effect of the flexures and damping is almost negligible. However, to characterize the system, it is necessary to evaluate the spring constant of the flexures. For a theoretical estimate of this parameter refer to [15] which gives a simplified way of calculating the stiffness. For the single-degree-of-freedom flexure used in the project, the equations for the vertical and rotational spring constants are

$$\frac{\alpha_z}{F_y} \approx \frac{9\pi R^{\frac{3}{2}}}{2Ebt^{\frac{5}{2}}} \tag{2.1}$$

$$\frac{\Delta y}{F_y} \approx \frac{9\pi}{2Eb}\left(\frac{R}{t}\right)^{\frac{5}{2}}. \tag{2.2}$$

where E is the Young's Modulus of the material and the dimensions R, t and b are shown in Figure 2-4. The flexures constrain the arm to rotational motion about the flexure midpoint, however the length from the flexures to the end of the arm is 0.3 m while the chosen allowed motion of the arm on either side of the equilibrium position

[4]Tetradyne Software Inc. 2542 S. Bascom Ave, Suite 206 Campbell, CA 95008

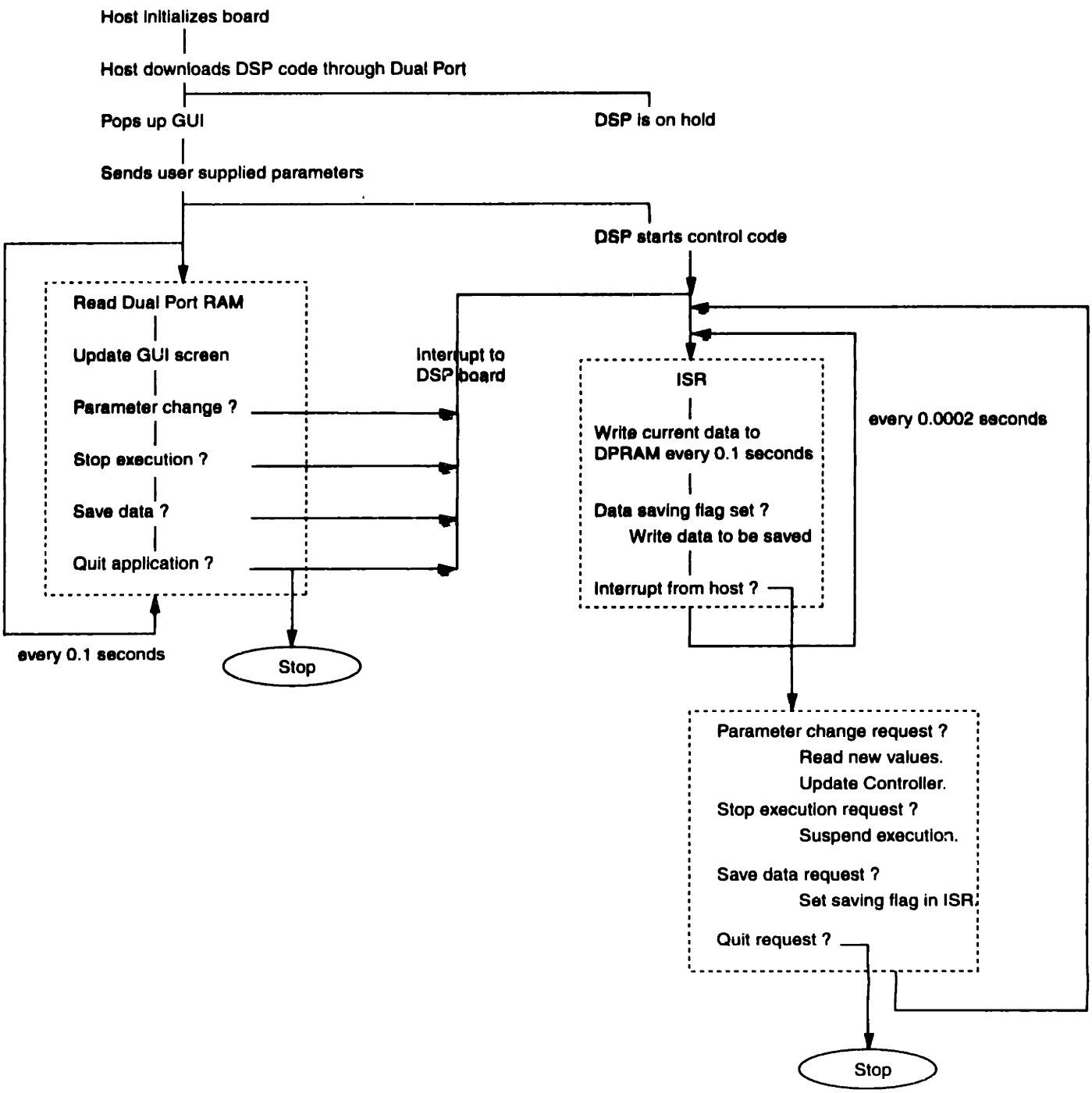

Figure 2-3: The flowchart for the operations performed in the closed-loop digital system.

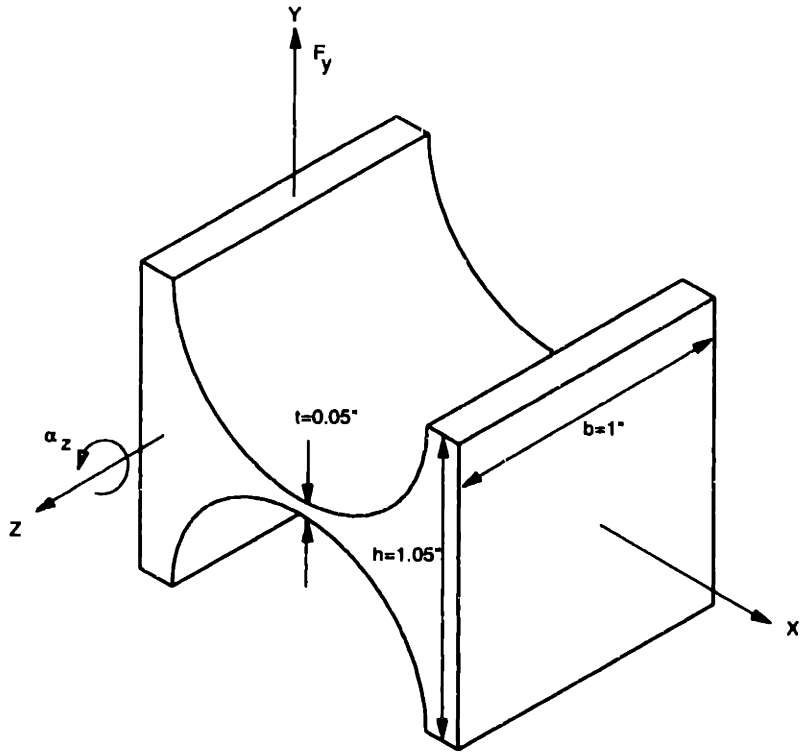

Figure 2-4: Drawing of one of the two flexures used to constrain the motion of the arm to one plane.

is $\pm 254\,\mu$m ($\pm 0.01''$) therefore it will be assumed that the target translates linearly.

The flexures are manufactured from 6061-T6 Aluminium with an alodine coating, so $E_{aluminium} = 75\,\text{GPa}$. Using the measurements shown in Figure 2-4 and using equation (2.2) for the linear stiffness, we get that the theoretical stiffness for a single flexure is, $k = 672\,\frac{\text{kN}}{\text{m}}$. The experimental determination of the stiffness was done by building a setup in which a micrometer capped off with a load cell was used to impose a displacement of the target. The load cell measures the reaction force exerted by the stiffness of the flexures and from the linear-motion assumption, the gradient of the displacement vs. force plot gives the stiffness of the flexures. The experimental value of the stiffness was determined to be 2218 N/m which does not compare at all well with the theoretical value given previously. The discrepancy has not been investigated further.

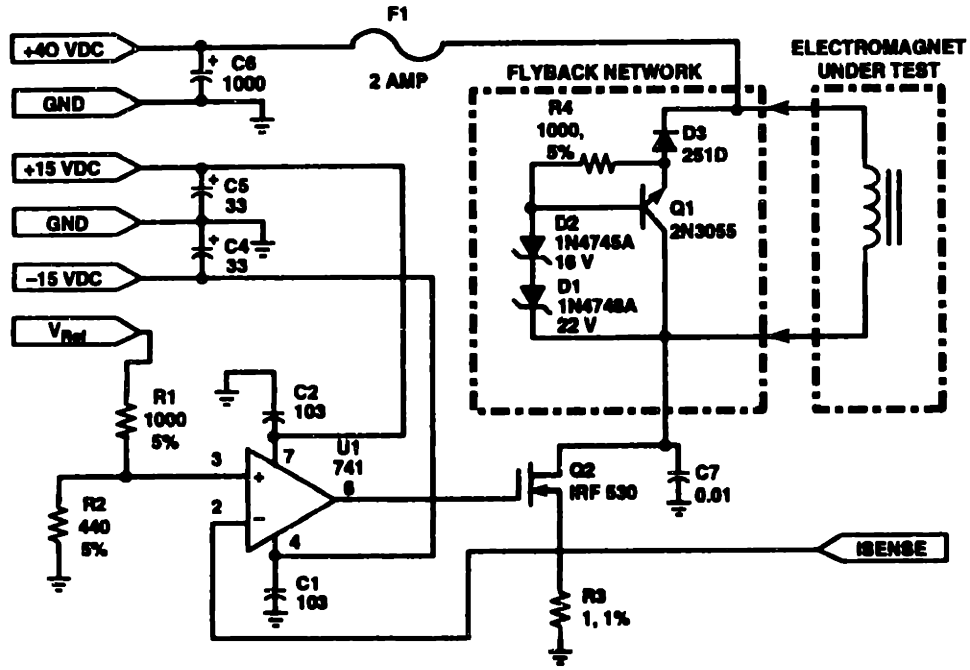

Figure 2-5: Circuit diagram of the current amplifier.

2.4 Amplifier

Before modeling the electromagnetics involved in the system, we will first make a short note on the amplifier used to amplify the D/A signal and drive the electromagnets. The design is borrowed from [29] and is notable because it allows rapid reduction of the coil current and has a high negative current slew rate capability. For reasons to be discussed in the following chapter, our desired bandwidth will be about 100 Hz, while as the amplifier has a 12 kHz bandwidth. From this we see that for the dynamic range of interest the amplifier can be approximated by a constant gain whose experimentally determined gain is $K_a = 0.127$ A/V. For completeness we show the circuit diagram of the amplifier in Figure 2-5.

2.5 Resonances

The system exhibits an open-loop resonance of 8 Hz which is clearly visible by connecting the capacitance gage to an oscilloscope and observing the arm moving freely. Using the experimental value of the spring constant given in section 2.3 and this

resonant frequency, we find that the equivalent translational mass of the target and arm is 1.15 kg. This is the values that will be used in the control design.

Although the double flexure arrangement constrains the motion of the arm to essentially one degree of freedom, we expect there to be some cross coupling between axes. It is therefore important to find the resonant frequency of the vertical mode in particular so that we can consider it in our controller design. The complete, closed form solution of the vibration equation for the system at hand is quite impractical and therefore we will use Rayleigh's Energy Method [22] which is an approximate technique. This energy method is based on the assumption that the system is conservative thus there is a non-dissipative interchange of energy between kinetic energy and potential energy. While no mechanical system is entirely conservative, this method serves to obtain an initial estimate which can then be refined by more accurate methods if needed [21]. Flexures are elements which tend to exhibit low energy dissipation and therefore the Rayleigh Method can potentially provide good results. The place where large errors are most easily introduced in the analysis is in the selection of trial mode shapes for the system. These assumed modes are used to calculate the potential and reference kinetic energies. The quality of the approximation of the natural frequency which results from using this technique, is highly coupled with the closeness between the actual mode shape and the assumed one.

The Rayleigh Energy Method states that the natural frequency is approximated by

$$\omega^2 \approx \frac{V_{max}}{T^*} \tag{2.3}$$

where V_{max} is the maximum potential energy of the system (calculated at its point of maximum deformation when the kinetic energy is zero), and T^* is the reference kinetic energy of the system.

For the horizontal dynamics, the compliance of the flexure is much greater than the compliance of the shaft to bending motion, hence the flexure will absorb virtually all of the rotation about the flexure center, while the shaft will remain undeflected.

We can assume a linear mode shape

$$\gamma = \frac{x}{L}$$

where x is the distance along the shaft and L is the length from the flexure midpoint to the end of the shaft. Using this assumed mode to calculate the energy of the system and then using (2.3) we get that

$$\omega^2 \approx \frac{\frac{1}{2}\int_0^L k\theta \, d\theta}{\frac{1}{2}\int_0^L m(x)\gamma^2 \, dx + \frac{1}{2}M\gamma^2(L)} \tag{2.4}$$

where k is the stiffness of the flexure, θ is the rotation of the shaft about the flexure midpoint, $m(x)$ is the approximate mass distribution of the shaft and target, and γ is the assumed mode shape.

In order to calculate the integral we must estimate the parameters involved. The length from the flexure midpoint to the tip of the target is $L = 0.3$ m (12"). The thickness of the tube making the shaft was taken to be $t = 1.73$ mm which is a standard tube thickness size; the target was weighed and its mass found to be $M = 0.45$ kg. Being an aluminium shaft the density is known to be $\rho = 2700 \frac{\text{kg}}{\text{m}^3}$. Without taking into account all the geometrical details of the setup, a quick estimate of the mass distribution is obtained by assuming the distribution to be uniform. For this we calculate $m(x) \approx \pi (r_o^2 - r_i^2)\rho$ where $r_o = 0.046$ m (1.81") is the outer diameter of the shaft and r_i is the inner diameter. This approximation gives the mass distribution as $m(x) = 0.746 \frac{\text{kg}}{\text{m}}$. Inserting these estimates into the above, the natural frequency is estimated to be 12 Hz. We note the discrepancy with the actual value of 8 Hz as partly due to estimation of $m(x)$ and partly due to the assumed mode shape.

For vertical dynamics, we re-write (2.3) as

$$\omega^2 \approx \frac{\frac{1}{2}\int_0^L EI\,(\gamma'')^2 \, dx \; + \; Mg\gamma(L)}{\frac{1}{2}\int_0^L m(x)\gamma^2 \, dx \; + \; \frac{1}{2}M\gamma^2(L)} \tag{2.5}$$

in which the potential energy contribution comes from the shaft bending as a cantilever beam. The expression is analogous in form to the equation for the in plane

motion, except that potential energy of the flexures has been replaced with the potential energy of the beam in bending for which the term EI represents the bending stiffness of the beam. A quadratic trial mode

$$\gamma = \frac{x^2}{L^2} \tag{2.6}$$

is assumed.

To determine the parameters to be used in the above integral we use a simplifying approximation and physical intuition: it is assumed that the deflection occurs only in the bending of the shaft and not in the flexures. This assumption can be qualitatively justified by noting that flexures exhibit a high stiffness to motion in directions perpendicular to the allowed direction of motion; in addition the two flexures are placed far apart in the vertical direction which greatly increases the opposing moment which they exert to motion in the vertical direction. The length to be used in this integral is different from the one used previously in the horizontal motion because it only accounts for the shaft overhang. From the base of the shaft to the tip of the target this length is $L = 0.267$ m (10.5"). Being fiber reinforced, the Young's modulus for the reinforced aluminium was approximated as $E = 80\,\text{GPa}$ and the density was given before as $\rho = 2700\,\frac{\text{kg}}{\text{m}^3}$. Similar to how we estimated the mass distribution, we can estimate the moment of inertia as $I \approx \frac{\pi}{4}(r_o^4 - r_i^4) = 6.7 \times 10^{-8}\,\text{m}^4$. Inserting these parameters into the equation above we get that $f_{vertical} = 241$ Hz. The actual value of the vertical resonance was experimentally determined to be around 270 Hz, and was measured using a geophone placed on the target. Geophones are devices which measure absolute velocity, and are essentially accelerometers with an integrator. The particular one used in this project is the McSeis 101LT by GeoSpace Corp[5] with a resonance of 5.5 Hz. Measurement of the shaft's vertical resonance was done by tapping the target and picking up the vertical vibrations with the geophone. Taking the DFT of the data, we see that the vertical resonance is approximately 270 Hz which is in good agreement with the result derived above. The spectrum of the vertical

[5]GeoSpace Corp. 7334 N. Gessner Houston, TX 77040

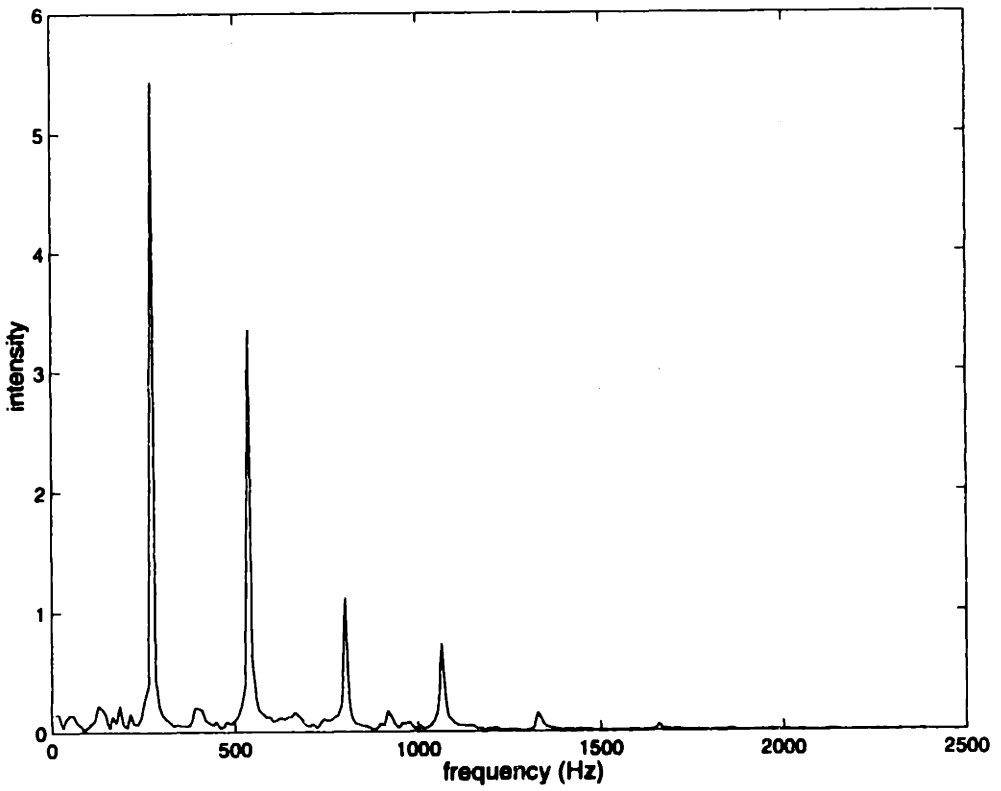

Figure 2-6: DFT of signal from vertical motion of target.

motion is shown in Figure 2-6.

2.6 Capacitance gage and RC circuit

As shown in Figure 2-2 the loop is closed around the position of the target as measured by the capacitance probe. The specific model used is the PX405HC made by Lion Precision[6] with a gain of 3.937×10^4 V/m (1000 V/in) over a range of $\pm 305\,\mu$m from the zero output position. This probe easily allows for sub-micron resolution provided the surface being measured is perpendicular to the sensor face; with good shielding the accuracy can be improved even more. The signal conditioning electronics requires a ± 15 V input and outputs a signal that varies linearly, (although towards the end of the range of operation, sensitivity starts rolling off), with position over a range of ± 12 V. The full range of the probe could not be used because the A/D only accepts an

[6]Lion Precision, A division of A.Q.T Inc., 563 Shoreview Park Road, St. Paul, MN 55126.

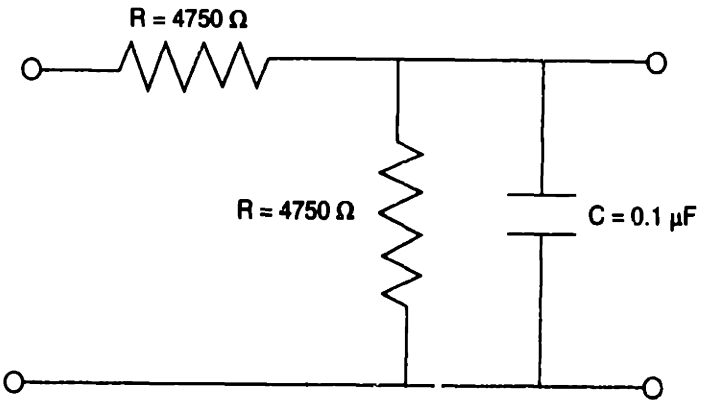

Figure 2-7: Divide-by-two passive filter.

input range of $\pm 5\,V$, thus we added a divide-by-two circuit on the input to the A/D. A passive RC filter was chosen over an active filter to avoid noise pickup problems; the filter is shown in Figure 2-7. Inserting a low-pass filter in front of the A/D has the added advantage of acting as an anti-alias filter for high frequency signals.

To make sure the filter dynamics does not interfere with that of the system, we designed it to have a bandwidth of 670 Hz. The steady-state gain of the filter is one half, so that it is able to keep the range of the probe to $\pm 6\,V$. Due to non-linearities in the probe design, the gain at large displacements will no longer be constant, thus the maximum displacement was chosen so as to produce a filtered output of ± 5 V into the A/D. The gain of the probe in series with the filter is now 1.9685×10^4 V/m and the linearly measured displacement of the target is bounded to $\pm 254\,\mu$m.

2.7 Electromagnetic actuators

Figure 2-8 shows a photo of one of the electromagnets used in the magnetic bearing. The E-core is composed of laminations made from 50% Ni-50% Fe alloy chosen because of its low hysteresis and reasonably high force characteristics.

Figure 2-9 shows the actuator diagram and how the flux path is closed.

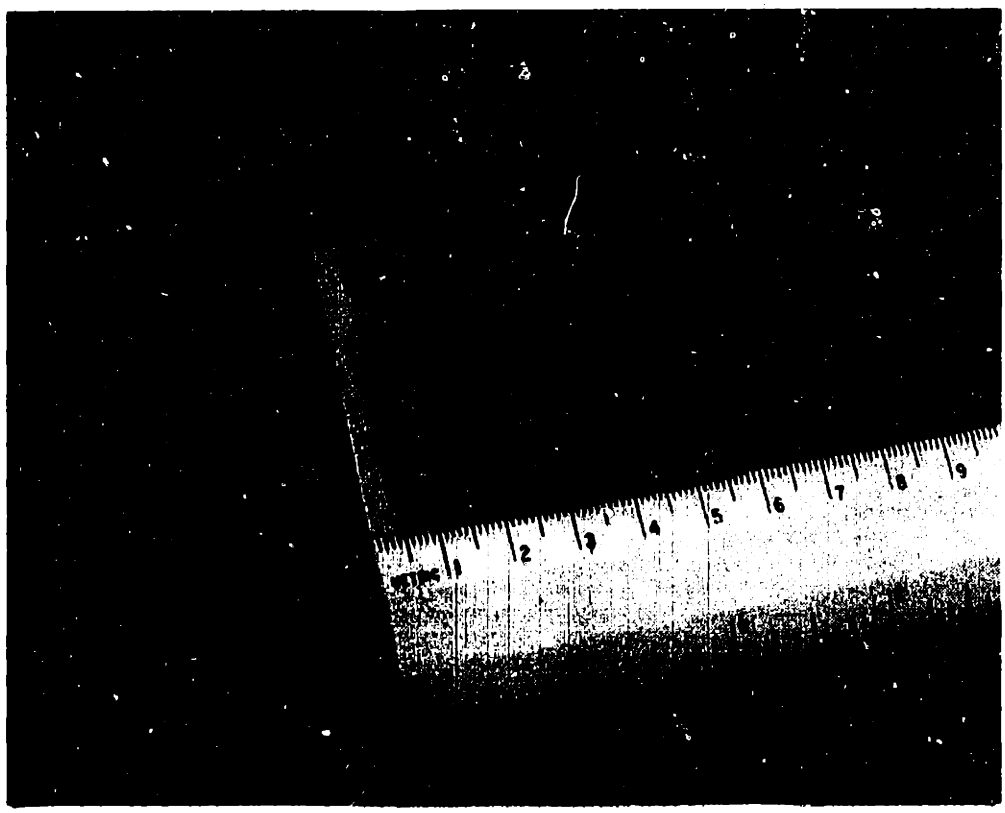

Figure 2-8: Picture of one of the electromagnetic actuators.

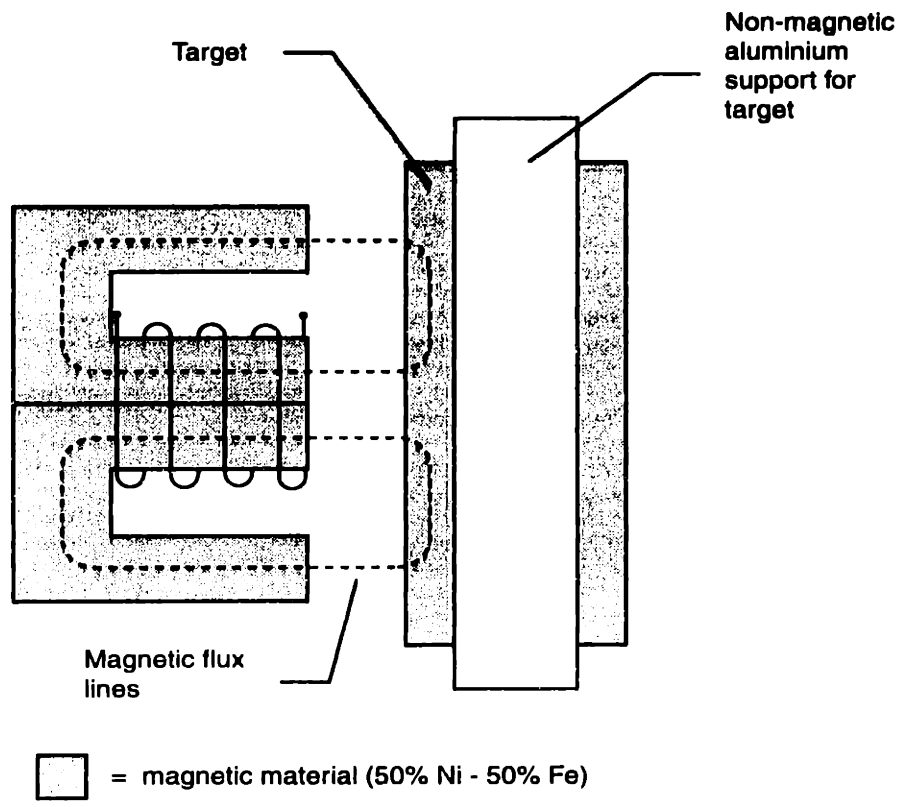

Figure 2-9: Diagram of the electromagnetic actuator and the magnetic circuit.

2.7.1 Classical modeling

This section illustrates how to model the actuator using basic principles from classical electromagnetic theory. This presentation is tailored after [12, 25].

We start from Maxwell's equation

$$\oint_c \mathbf{H} \cdot dl = NI \tag{2.7}$$

which, if integrated around one of the two flux paths, gives

$$H_{core}\, l_{core} + 2\, H_{gap}\, x + H_{target}\, l_{target} = NI \tag{2.8}$$

where H is the magnetic field intensity, N is the number of turns of the coil, I is the current through the coil and x is the variable gap size. We can now use Gauss' Law to express the continuity of the flux around the flux path as

$$\oint_s \mathbf{B} \cdot dA = 0 \tag{2.9}$$

where B is the magnetic flux density and A is the cross-sectional area of the actuator face. Assuming no fringing in the field as well as a linear constitutive relation $B = \mu H$ of the material, we find that

$$\mu_{core}\, H_{core}\, A = \mu_{gap}\, H_{gap}\, A = \mu_{target}\, H_{target}\, A \equiv \Phi \tag{2.10}$$

where μ is the permittivity of the part cited in the subscript and Φ is the magnetic flux. We can solve for H_{core} and H_{target} and insert them into (2.8) to get an equation for H_{gap} as

$$H_{gap} = \frac{NI}{\left(\frac{\mu_{gap}\, l_{core}}{\mu_{core}} + 2x + \frac{\mu_{gap}\, l_{target}}{\mu_{target}}\right)}. \tag{2.11}$$

In order to find the force exerted, we can either find the potential energy in the gap and take the partial derivative with respect to the displacement x, or more simply we

can utilize Maxwell's stress tensor

$$\mathbf{F} = \frac{\mathbf{B}^2}{2\,\mu_o} \cdot \mathbf{A} \tag{2.12}$$

where $\mathbf{A}$ is the area of one of the flux paths in the gap. As Figure 2-9 shows, the E-core magnet sets up two magnetic loops thus the total area of the flux is twice the area of one of the legs of the E-core. Use of the above equation on one of the loops only gives half the total force exerted by the actuator so the result must be multiplied by two. Using Faraday's Law we can also say that for this system

$$\Phi = \frac{Ni}{\mathcal{R}} \tag{2.13}$$

where $\mathcal{R}$ is the reluctance of the flux path given by

$$\mathcal{R} = \left(\frac{2x}{\mu_{gap}\,A} + \frac{l_{core}}{\mu_{core}\,A} + \frac{l_{target}}{\mu_{target}\,A} \right). \tag{2.14}$$

Inserting (2.10), (2.11), (2.13) and (2.14) into (2.12) the force exerted by the actuator is

$$F = \frac{N^2 i^2}{\mu_{gap}\,A} \frac{1}{\left(\frac{2x}{\mu_{gap}\,A} + \frac{l_{core}}{\mu_{core}\,A} + \frac{l_{target}}{\mu_{target}\,A} \right)^2} \tag{2.15}$$

We note that for the choice of materials used in the actuators,

$$\mu_{gap} \ll \mu_{target} \tag{2.16}$$

$$\mu_{gap} \ll \mu_{core}$$

so with these physical considerations it is possible to simplify the expression for the force to

$$F = -\frac{\mu_{gap} N^2 A}{4} \left(\frac{i}{x} \right)^2. \tag{2.17}$$

We will see in the section dedicated to actuator calibration that we can use this simplified version to model the actuators, but we will need to add a term to account for the finite permeability of the core and target.

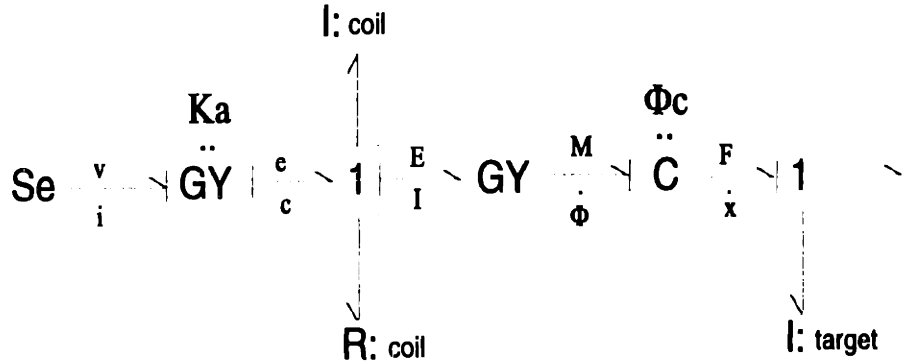

Figure 2-10: Bond graph model of the electromagnetic actuator shown in Figure 2-9.

2.7.2 Bond graph modeling of the electromagnetic actuators

In this section we briefly present an alternative way of modeling the actuators based on the bond graph method [5].

Figure 2-10 shows the bond graph model for the actuator shown in Figure 2-9. The first gyrator represents the current amplifier used in changing the voltage signal, Se, exiting the D/A to the current used to drive the actuators. As we noted before, the bandwidth of the amplifier is much higher than the bandwidth which we are aiming for, therefore it can be modeled as a constant gain, K_a, without the need to add parasitic dynamics. The second gyrator represents the conversion between energy domains occurring inside the actuator. The input to the latter are the electrical power variables (voltage, E, and current, I) while as the output from the actuator are the magnetic power variables (magnetomotive force, M, and flux rate, $\dot{\Phi}$).

Unlike the analogy with Ohm's law that is often made in elementary books on electromagnetics the gap described by (2.18), where $\mathcal{R}$ is the reluctance, is in fact a capacitor which stores energy in the magnetic field. Equation (2.18) shows this because it relates the a generalized displacement (the flux) to a generalized effort (the magnetomotive force).

$$M = \mathcal{R}\Phi \tag{2.18}$$

The power balance on the generalized capacitance is

$$P = M\dot{\Phi} + F\dot{x} \tag{2.19}$$

which integrated gives the potential energy, E_p.

$$\Delta E_p = \int_{t_0}^{t} \left(M\dot{\Phi} + F\dot{x} \right) dt \tag{2.20}$$

Since potential energy is independent of the path taken, we can select the path of integration that allows for a simpler integral; an example of which is a path for which $F = 0$ always.

$$\Delta E_p = \int_{\Phi_0}^{\Phi} M \, d\Phi \tag{2.21}$$

Use of (2.13) allows this integral to be evaluated. From the potential energy stored in the gap, the force exerted by the magnetic field is

$$\mathbf{F} = -\nabla E_p \tag{2.22}$$

We can thus use (2.13), (2.18) and (2.21) in (2.22) and get the force as

$$F = \frac{N^2 i^2}{\mu_{gap} A} \frac{1}{\left(\frac{2x}{\mu_{gap} A} + \frac{l_{core}}{\mu_{core} A} + \frac{l_{target}}{\mu_{target} A} \right)^2} \tag{2.23}$$

which we see is equivalent to (2.15).

2.8 Actuator calibration

Having derived a model for the actuator we can, as a notational simplification, group together constant parameters and define a constant C as

$$F = \frac{C\, i^2}{(g_0 + x)^2} \qquad \text{where} \qquad C = \frac{\mu_{gap} N^2 A}{4} \tag{2.24}$$

and x is the distance from the pole face; g_0 is added as a bias displacement to account

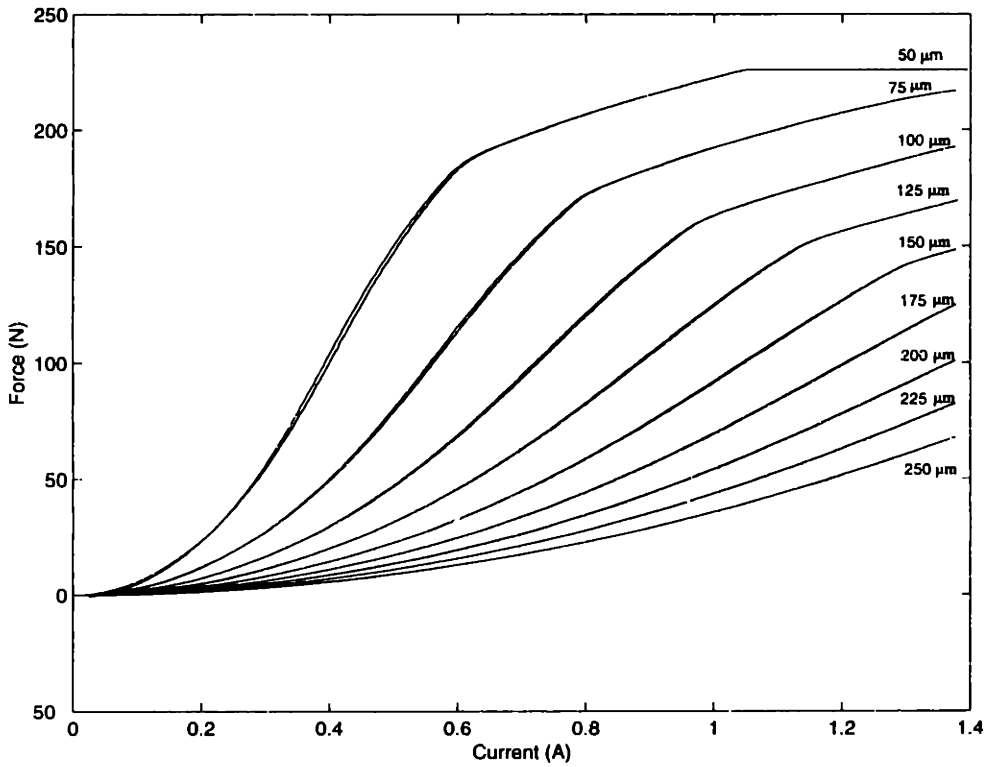

Figure 2-11: Experimental force-current-displacement plots of actuator one.

for the finite permeability of the cores. Equation (2.16) still applies, but in a real system $\mu_{core} < \infty$ and $\mu_{target} < \infty$ and a term must be introduced to account for this whose value is determined to be $\approx 20\,\mu$m. This value was obtained by fitting a quadratic curve to the calibration data for the actuator and seeing what value of g_0 would give the best correlation between the actual and theoretical data.

Each actuator has N=230 coils and a cross-sectional area of A=0.013 m $\times$ 0.013 m (0.51" $\times$ 0.51"), thus with $\mu_{gap} = \mu_{free\,space} \equiv 4\pi \times 10^{-7}$ H/m we get a theoretical value of $C = 2.789 \times 10^{-6}\,\frac{\mathrm{N\,m^2}}{\mathrm{A}}$. To measure this parameter experimentally we use the calibration fixture built by Poovey et al. [29]. We shall not detail the calibration fixture, so the reader is referred to the appropriate reference. Figures 2-11 and 2-12 show the force-current-displacement plots obtained by using this calibration fixture.

The graphs show the saturation characteristics of the actuators. The plot for an air gap of $50\,\mu$m shows a double knee in the curve, but the upper knee is due to saturation of the A/D in the calibration fixture which can be inferred from the fact

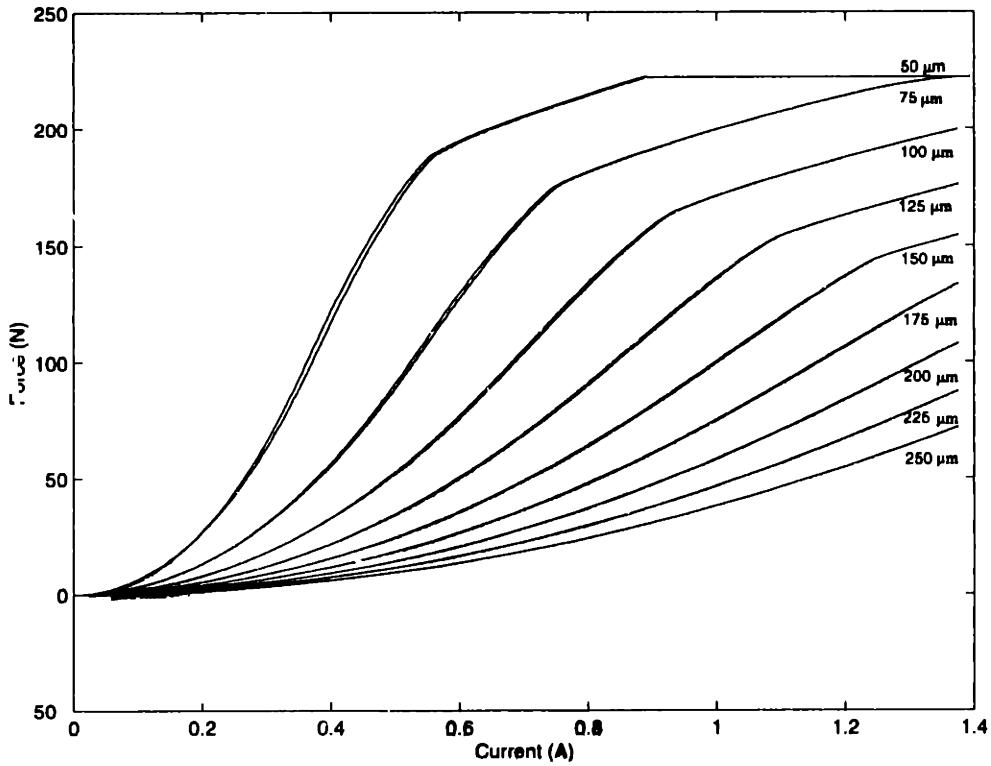

Figure 2-12: Experimental force-current-displacement plots of actuator two.

that the response curve becomes flat. The first knee of this curve, as well as the knee displayed by the other curves, is due to the preferential saturation of the core due to the grain orientation within the material. The electromagnets used are made of E-core laminations which are obtained by cutting an E-shape from a sheet of material which had been previously rolled to the desired thickness. The rolling process orients the grains in the material along the rolling direction so that upon magnetization, the flux will in some parts be parallel to this orientation and in others it will be perpendicular to it as shown in Figure 2-13. This leads to preferential magnetization and saturation.

We also note the relatively high force exerted by the actuators, and the low hysteresis. In order to obtain the experimental value for the actuator constant, C, a least-squares best fit is done on the part of the data which exhibits no saturation and we get that the constants for actuators one and two are $2.6 \times 10^{-6} \frac{\mathrm{Nm}^2}{\mathrm{A}}$ and $2.79 \times 10^{-6} \frac{\mathrm{Nm}^2}{\mathrm{A}}$ respectively which are in excellent agreement with the predicted

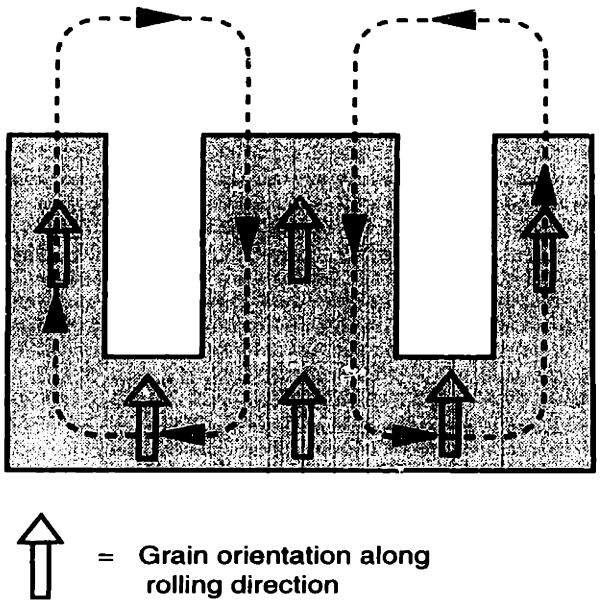

Figure 2-13: Effect of grain orientation on the saturation characteristics of the core

value.

2.9 State variable derivation

Equations (2.15) and (2.23) expressed the forces exerted by the actuators on the target. The bearing is set up so that the electromagnets are in pairwise opposition and attract the target in opposite directions; the net force is therefore the difference between the two forces. The spring constant of the flexures gives such a small resonant frequency with respect to the bandwidth at which we wish to close the loop that its effect will be negligible. In addition, the flexures provide little damping to the system especially at low amplitude motion, so that effect too can be neglected above resonance. Overall the plant can be modeled as a free mass subject to an input force equal in magnitude to the difference between the actuator forces. For the mechanical side of the system, we select the position of the target and the velocity of the latter as the state variables of interest. The state equations for the amplifier are not included for reasons described earlier, namely that the amplifier acts as a constant gain for the frequencies of interest. Therefore, the state equations are only those for the

mechanical system driven by the nonlinear electromagnetic forces.

$$\dot{x_1} = x_2 \tag{2.25}$$

$$\dot{x_2} = \frac{C_2}{M}\left(\frac{i_2}{g_0 + x_{20} - x}\right)^2 - \frac{C_1}{M}\left(\frac{i_1}{g_0 + x_{10} + x}\right)^2 + \frac{d}{M} \tag{2.26}$$

where x is the motion from the equilibrium operating point, x_{10} and x_{20} are the equilibrium gaps between the target and the pole faces of actuators one and two respectively and d is a disturbance force.

While as the setup is oriented in the horizontal plane and motion is thus perpendicular to gravity, the setup could have easily been rotated such that motion is in the direction of gravity. The latter arrangement is what most would envision as a magnetic bearing setup, however gravity enters as a simple bias in the state equations. As a matter of fact, gravity bias can be included in the generalized disturbance term in the above state equations so that the exact same relationship applies. Viewed in the sliding control framework which will be presented in chapter 5, the same exact equations which will be developed then, apply with the minor modification of having to change the disturbance bounds.

Chapter 3

Linear Control

Having derived the plant model in the previous chapter, this chapter will be dedicated to the design and implementation of several digital linear controllers. The state equations presented at the end of the last chapter show that the plant is non-linear in i^2 and the coefficient of the input is state dependent in x^{-2}. The state equations also show that the system is two-input, single-output. As for most non-linear systems, before approaching the complexity of non-linear control, it is good to check the applicability and validity of linearizing the plant about an operating point and using a linear controller on it. This chapter thus focuses on designing linear controllers for the magnetic bearing and checking the performance both through simulations and through digital implementation of the controller on the actual hardware.

3.1 Linearization

We show again the schematic of the actuators is pairwise opposition in Figure 3-1. If we select a nominal operating point around which the magnetic bearing is to operate, for small displacements around this operating point we can write

$$i_1 = I_1 + \delta i_1$$

$$i_2 = I_2 + \delta i_2$$

$$x = X + \delta x$$

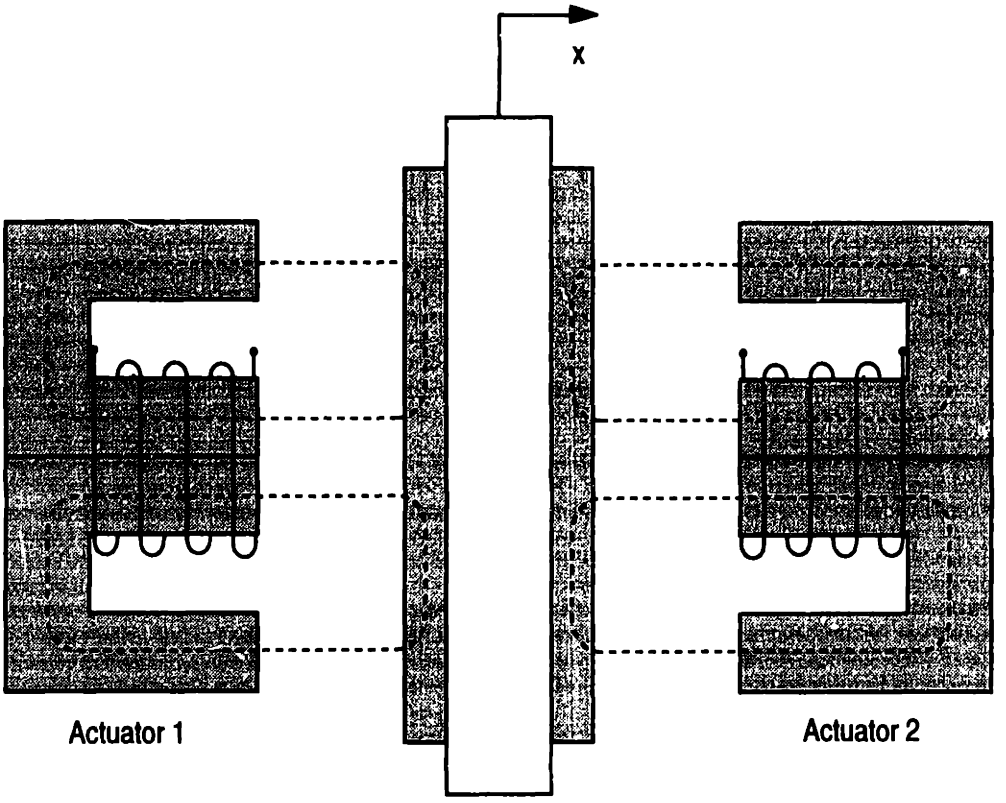

Figure 3-1: Schematic of the actuators in pairwise opposition

$$d = D + \delta d$$

where i_1 and i_2 are the currents to actuators one and two respectively and δi_1 and δi_2 are the changes in current from the respective bias values I_1 and I_2 [19]. The assumed disturbance force d is also taken as being composed of a bias component, D, and a change from the latter, δd.

We can now use a Taylor series expansion and retain terms only up to the first derivative, to get a linear model of the system around an equilibrium or operating point as

$$f(x, i_1, i_2, d) = f_e(x_e, i_{1e}, i_{2e}, d_e) + \left.\frac{\partial f}{\partial x}\right|_e \delta x + \left.\frac{\partial f}{\partial i_1}\right|_e \delta i_1 + \left.\frac{\partial f}{\partial i_2}\right|_e \delta i_2 + \left.\frac{\partial f}{\partial d}\right|_e \delta d. \quad (3.1)$$

The subscript e next to the partial derivatives indicates that the partial is to be evaluated at the operating point.

To avoid treating the problem as a multiple input single output system, let us

constrain the actuators such that the change in current into each actuator is equal in magnitude but opposite in direction. Let us further apply equal biases to each of the actuators so that the control effort will be more evenly distributed amongst the two electromagnets. We express this as

$$\delta i_2 = -\delta i_1 = \delta i \tag{3.2}$$

$$I_1 = I_2 = I \ .$$

With these constraints, we can linearize the plant according to (3.1) and re-write the equations in matrix notation. The partial derivatives evaluated at the point about which the plant is being linearized will become the constant coefficients for the incremental quantities δi_1, δi_2, δx and δd so we can define the increments as the new state variables. For convenience we drop the δ, but for the remainder of this chapter the state variables will be the incremental quantities measured from the point about which the system was linearized; they are *not* the absolute values of the states.

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ \frac{2C_2}{m}\frac{I_2^2}{(g_0+x_{20})^3} + \frac{2C_1}{m}\frac{I_1^2}{(g_0+x_{10})^3} & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + \begin{pmatrix} 0 \\ \frac{2C_2}{m}\frac{I_2}{(g_0+x_{20})^2} + \frac{2C_1}{m}\frac{I_1}{(g_0+x_{10})^2} \end{pmatrix} \delta i \tag{3.3}$$

$$y = [1\ 0]\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + [0\ 0]\begin{pmatrix} i \\ d \end{pmatrix} \tag{3.4}$$

$(g_0 + x_{10})$ and $(g_0 + x_{20})$ are the operating point displacements from the two actuators respectively plus the $g_0 = 20\,\mu$m to account for the finite permeability of the target. The system has now been put in the standard form $\dot{\mathbf{x}} = \mathbf{Ax} + \mathbf{Bu}$ for controller design. We note that if the flexure stiffness had been included in the model it would enter as $-\frac{k}{m}$ in the lower left entry of the **A** matrix in (3.3). However, by substituting parameter values into the equation, we see that the effect of the stiffness is very small compared to the coefficient already there; the latter represents the negative stiffness of the magnetic bearing. From the state space representation we get the transfer

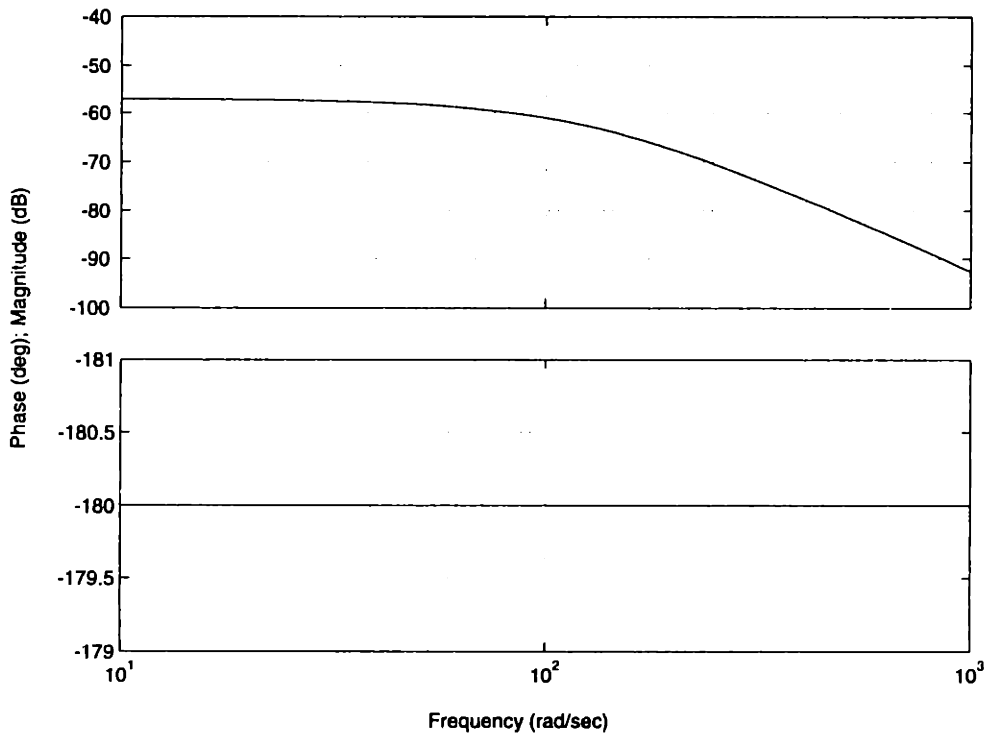

Figure 3-2: Bode plot of plant transfer function

function

$$\frac{X_1(s)}{I(s)} = \frac{23.913}{s^2 - 17080.745} \cdot \tag{3.5}$$

The transfer function clearly shows the existence of a right hand plane pole which makes the linearized plant open-loop unstable. Figure 3-2 shows the bode plot of the plant from which we see the existence of an unstable pole because the magnitude breaks but the phase remains constant at -180^o. Lyapunov's first method can now be invoked to claim that the non-linear system is also open-loop unstable. This is Earnshaw's theorem and is a property of all magnetic suspensions. It makes intuitive sense if we take the simple example of a ball levitator. Applying a constant current to the electromagnet coils will generate a force which will tend to draw the ball towards the magnet's face. There exists only one equilibrium point for this system and that is where the weight of the ball exactly balances the attractive force of the electromagnet, assuming initial conditions are zeroed. However, this equilibrium point is unstable because any infinitesimal displacement thereof will result in the ball rising to touch the

actuator or dropping to the ground, depending on the direction of the displacement. Stability can only be achieved by modulating the current in the coils and thus the need for feedback control.

For the magnetic bearing setup, we aim for a reasonably fast settling time, no steady state error, and a bandwidth of around 100 Hz. The latter specification comes from not wanting to excite the vertical dynamics too much since its effects have not been modeled. In Chapter 5 we will see that a rule of thumb to start designing a sliding mode controller is to set the bandwidth of the system to 1/3 the first unmodeled mode of the system [14], and we shall adopt that rule for the linear design. Experimentation has shown that this is indeed a valid crossover to aim for, and the reason for this shall be detailed later in this chapter. To check the validity of the model we would need to compare the loop transmission of the model with that of the actual system, but in order to obtain the experimental loop transmission we must first stabilize the system.

3.2 Model verification

The system was modeled as a free mass and thus has no damping. Although this is not strictly true in practice, we stated in the previous chapter that the assumption is justified because of the low damping of the flexures. This lack of damping calls for a lead compensator to be included in the system in order to raise the phase margin. Also, in order to increase low frequency response of the system as a way to achieve better disturbance rejection, especially in view of the 8 Hz natural frequency, we include a lag compensator. The actual design of the controller can be done in continuous time and then mapped by emulation techniques, or it can be performed in discrete time directly. Out of the linear controllers to be presented some are mapped to discrete time and some are designed in discrete time.

Damping is primarily a function of the lead while as crossover is adjusted using the gain. Since the lag is usually designed to affect low frequencies and have almost no effect around crossover, the lead and the lag can be designed separately. To have maximum phase increase at the desired crossover, we select the lead such that

the desired crossover is the geometric mean of its pole and zero. To express this mathematically, for a lead whose general form is given by

$$G_{lead}(s) = K_{lead}\frac{s-\delta}{s-\gamma}\ , \tag{3.6}$$

where $\delta < \gamma$, the maximum phase bump occurs at a frequency given by

$$\omega_{max\,phase} = \sqrt{\delta\gamma}$$

The separation of the pole and zero has a direct influence on the amount of phase increase that is achievable. Ideally it is desirable to have a very large pole-zero separation so as to increase the phase as much as possible, but the limitation is that for plants whose response above crossover is falling with increasing frequency, we will generally loose gain margin. The reason is that in between the break points of the lead the relative order of the system is decreased by one and thus, especially if crossover occurs in this region, the rate of change of magnitude will increase by 20 dB/decade. Another practical, and more intuitive, limitation of selecting a very large pole-zero separation is that a lead is a high pass filter and therefore raises the noise response of the system. A separation factor of 15 between the lead zero and the lead pole will ideally give about 60^o of phase increase [19], but depending on where the lead is placed and what poles/zeros are breaking around it, this phase increase me be substantially lower. A separation factor of 15 is chosen for all of the linear controllers designed here. In order to place the phase bump at 100 Hz, we thus solve the following equations:

$$\sqrt{\delta\gamma} = 2 \times \pi \times 100$$
$$\frac{\gamma}{\delta} = 15$$

which give $\delta = 162.231$ and $\gamma = 2433.467$. The bode plot of the lead compensator is shown in Figure 3-3 from which we see that phase bump is indeed at 100 Hz. The gain of the lead will be combined with that of the lag and will be used to set the

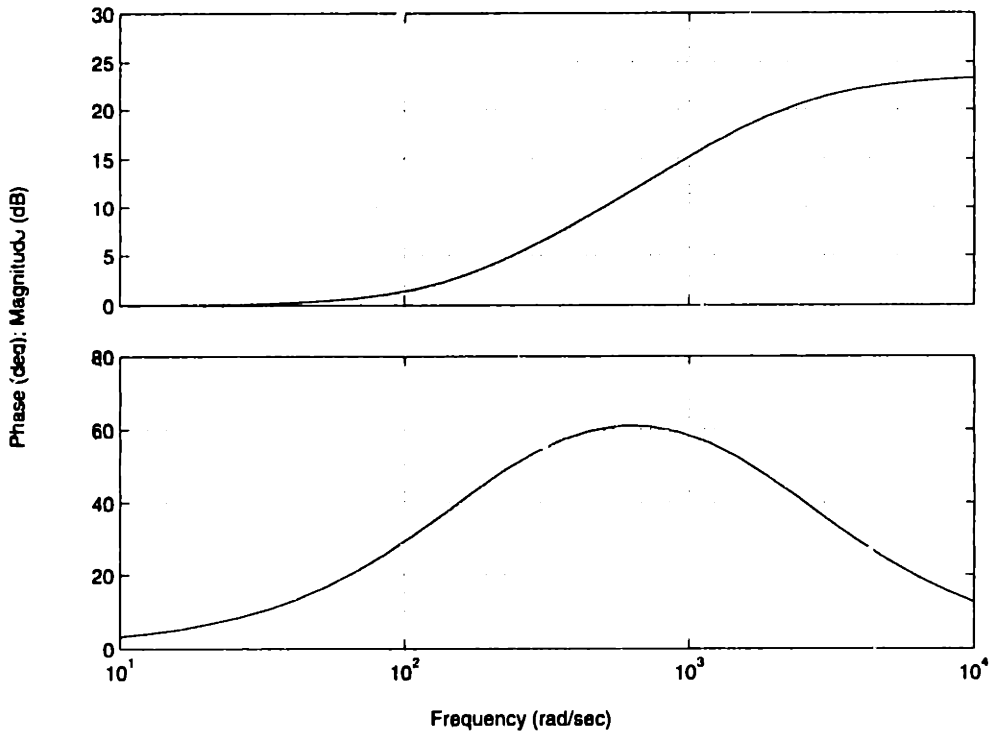

Figure 3-3: Lead compensator designed to produce a phase bump at 100 Hz.

crossover.

Designing the lag is analogous to the procedure detailed above, but the problem is in selecting where to place the zero. In order to keep the lag from affecting the phase at crossover too much, the break point was first chosen to be a decade below crossover, i.e., at 10 Hz. The lag is thus given by

$$G_{lag}(s) = K_{lag}\frac{s + 62.831}{s} .$$

Simulating the lead-lag compensated plant with the gain of the controller set to 1, we find the magnitude of the loop transmission at 100 Hz and then use the inverse of that value as the gain of our controller. Figure 3-4 shows the loop transmission of the continuous-time and discrete-time lead-lag compensated system.

In order to map the controller to discrete time so that it can be implemented on the DSP board, we must select a sampling rate for the digital system. Franklin et al. [11] suggest using sampling rates as low as six times the crossover, but it is also

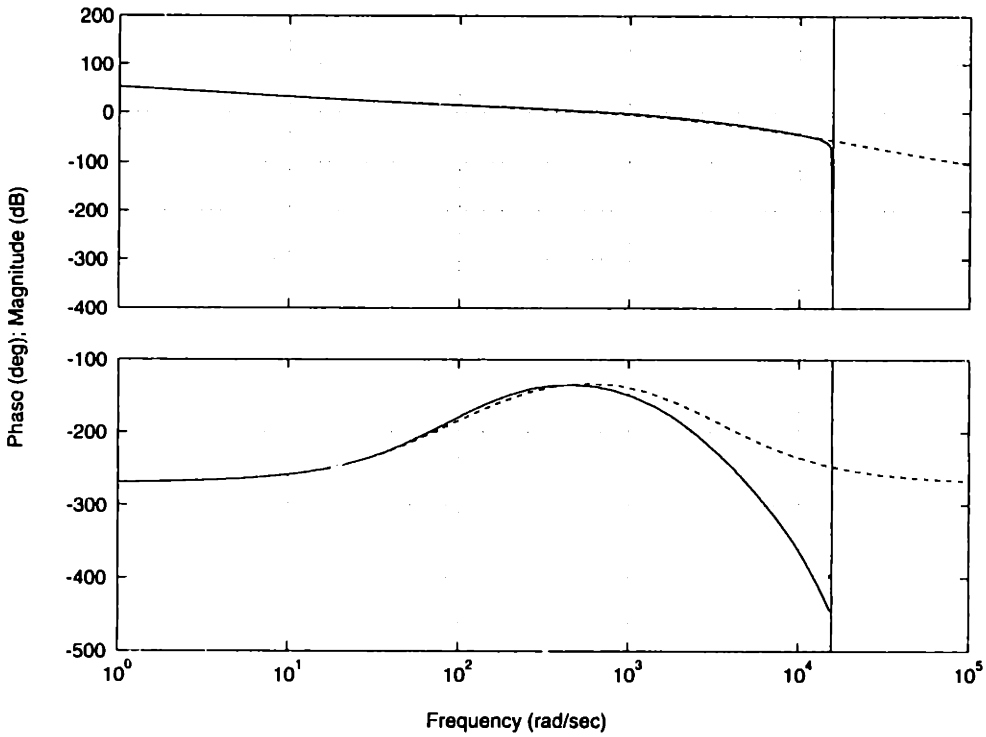

Figure 3-4: Loop transmission of lead-lag compensated system

stated that a value of 20 times or more should be implemented where possible. We have the possibility of sampling much faster than that and in general the faster the better, so we select a sampling rate of 5 kHz. Mapping the controller with the ZOH equivalent gives

$$G_{leadlag}(z) = 21.02\,\frac{z-0.9743}{z-0.6147}\,\frac{z-0.9874}{z-1} \tag{3.7}$$

In order to simulate the plant in discrete time it is also necessary to take the ZOH equivalent of the plant

$$\frac{X_1(z)}{I(z)} = \frac{4.783\times10^{-7}z + 4.783\times10^{-7}}{z^2 - 2.001\,z + 1} \tag{3.8}$$

where I is the input current into the plant.

3.3 Digital implementation

To reduce excessive control effort we implement the controller with the lead placed in the feedback path as shown in Figure 3-5, where the general form of the lead is (3.6) (with z substituted for s) and that of the lag is

$$G_{lag} = K_{lag}\frac{z-\alpha}{z-1} \ . \tag{3.9}$$

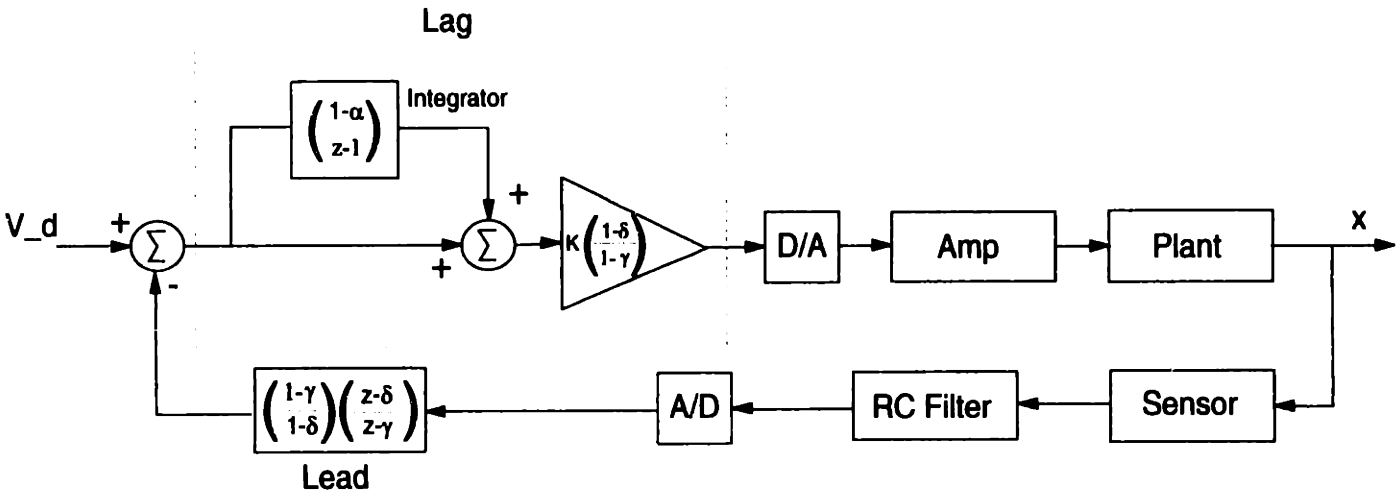

Figure 3-5: General structure of lead-lag compensated system

The lag is divided into two parts one being the proportional part, the other being the integrator. This functional splitting allows us direct access to the integral term so that we can implement antiwindup measures [7]. The latter are included in the control algorithm in the form of a software switch which limits the value of the integral to the range of operation of the D/A, i.e. ± 10 V. In order to avoid steady-state errors it in necessary to normalize the DC gain of the lead compensator to unity. Therefore we use part of the controller gain to normalize the lead and use the remainder of the controller gain as the gain of the lag compensator [11, 19].

The controller designed above stabilized the system and allowed the loop transmission to be measured experimentally using a dynamic analyzer. We created a summing junction for injecting a swept sine signal into the A/D and then measured the return signal out of the capacitance gage electronics. The setup is shown schematically

in Figure 3-6. By superposition, we can write the output voltage of the summing junction as

$$V_o = \frac{R_1}{R_1 + R_2} V_i + \frac{R_2}{R_1 + R_2} V_s \tag{3.10}$$

which, for the chosen resistors is

$$V_o = \frac{1}{11} V_i + \frac{10}{11} V_s \ . \tag{3.11}$$

The output of the junction, which is what is being fed to the RC filter and A/D, is almost the same values as the source voltage being applied, while as the return signal from the capacitance gage is attenuated by a factor of 11. In this configuration the plant is being driven mostly by the swept sine wave voltage source. Figure 3-7 shows the negative of the loop transmission, and on the same graph we have also plotted the bode plot of the discrete time compensated plant model.

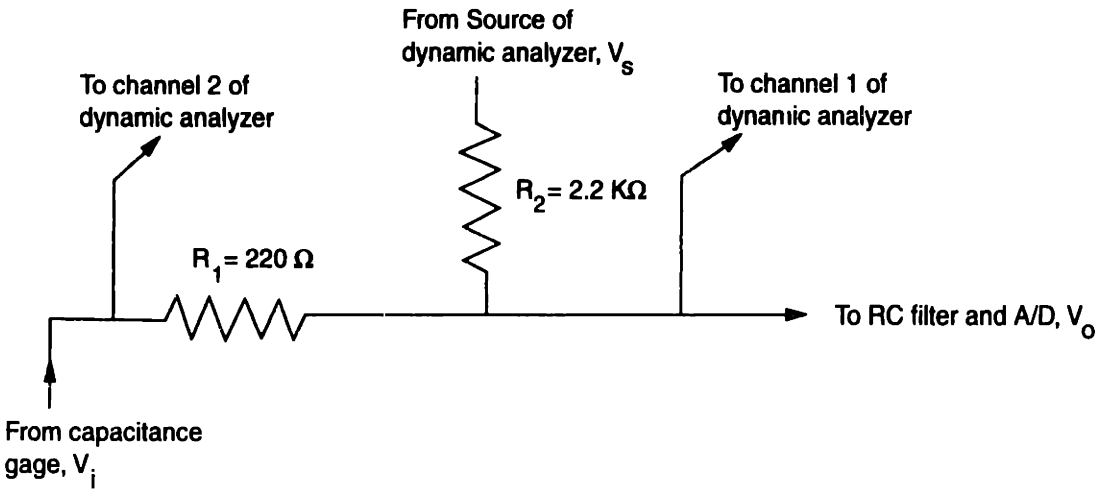

Figure 3-6: Setup for measuring the loop transmission experimentally.

We note the very good correspondence between the actual and theoretical model for the magnitude plot. On the other hand the phase plots show good agreement up to 60 Hz or so and then diverge quite rapidly. The actual system is losing phase more rapidly than the model predicts. The way the phase keeps falling off would seem to be the effect of neglected time delays however good fit between model and experimental data could not be obtained even if delays were added to the model. For

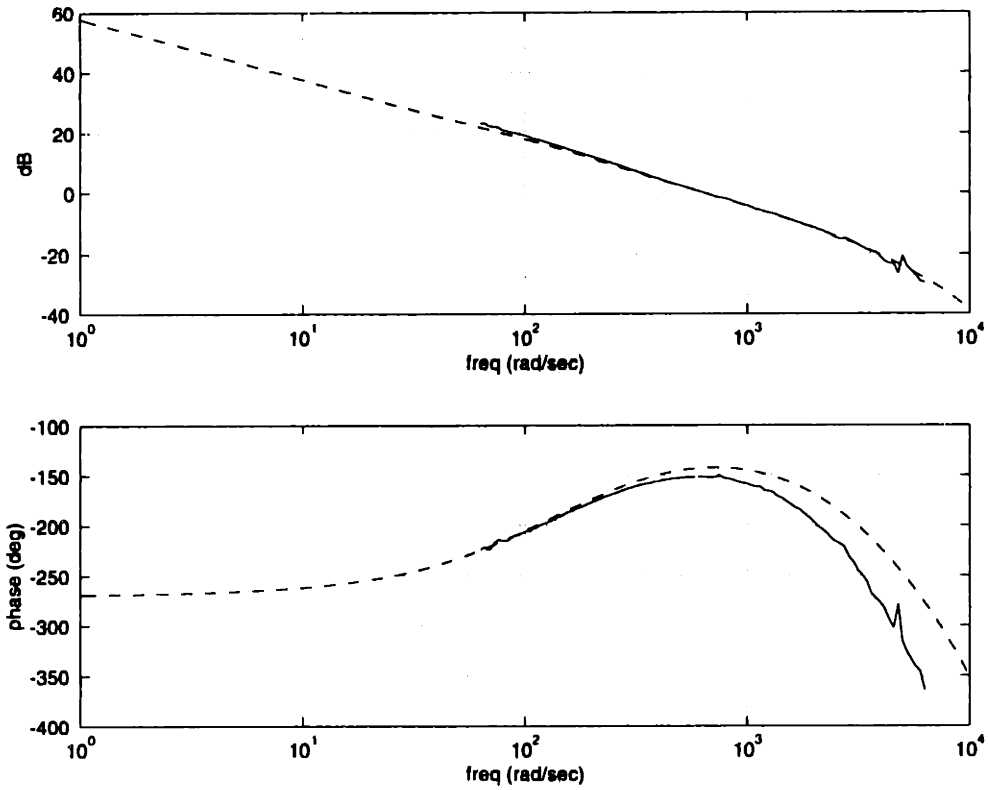

Figure 3-7: Experimental plot of the negative of the discrete loop transmission (solid) overlayed with bode plot of the discrete time compensated plant model (dashed).

instance, including a time delay of 300 μs with a Pade approximation could make the phase coincide well until around 90 Hz after which the experimental data and the model would diverge. We note though that 300 μs is equivalent to 1.5 sampling intervals which is an extremely long time delay considering the speed of the DSP board; in addition, when we simulate a step response for the system with this delay the overshoot is much greater than the one measured on the hardware.

We re-design the controller to make the system faster. To do this we push the lead further into the left half plane and center the phase bump beyond 100 Hz. We do this because in view of the asymmetric phase drop, the maximum increase in phase will occur before the geometric mean of the lead break points. Furthermore, to increase response we also bring the lag zero closer to the lead zero so that we get more gain at all frequencies. The down side of this last step is that it detracts some phase at crossover thus making the system more oscillatory. With these considerations we design the following compensator which we implemented on the magnetic bearing and

obtained the step response shown in Figure 3-8 when a 25 μm step was commanded. On this plot we have also shown the step response of the model with a 300 μs time delay. The negative of the loop transmission for this revised system is shown in Figure 3-9 for comparison.

$$G_{lead-lag} = 25\,\frac{z-0.9601}{z-1}\,\frac{z-0.965508}{0.481909} \tag{3.12}$$

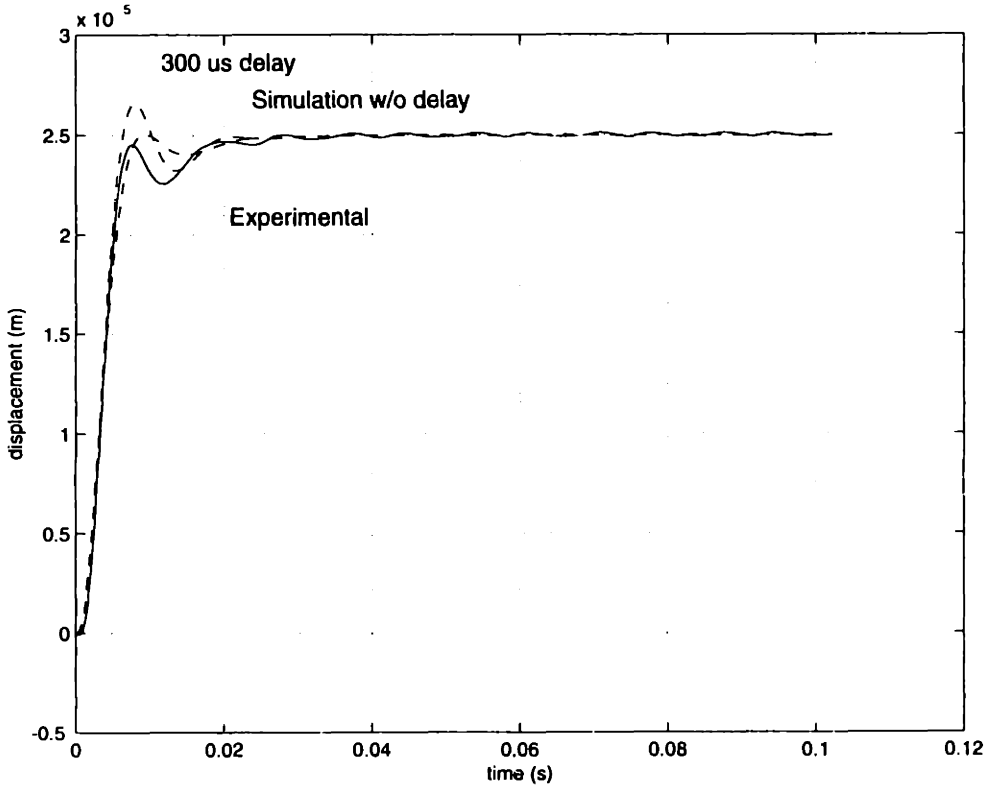

Figure 3-8: 25 μm step response of system compensated with (3.12) with and without the time delay.

The unmodeled phase roll-off reduced phase margin, which accounts for the slightly more oscillatory response seen in the real system but we see from Figure 3-8 that the time delay is still not sufficient to account for the discrepancies between the simulated and experimental responses. Pin-pointing the exact nature of this phenomenon is one of the recommended areas of future work. We note however that the settling time of the model and that of the actual system are equal which enables us to make time predictions using the model. The rate at which the controller drives the trajectory

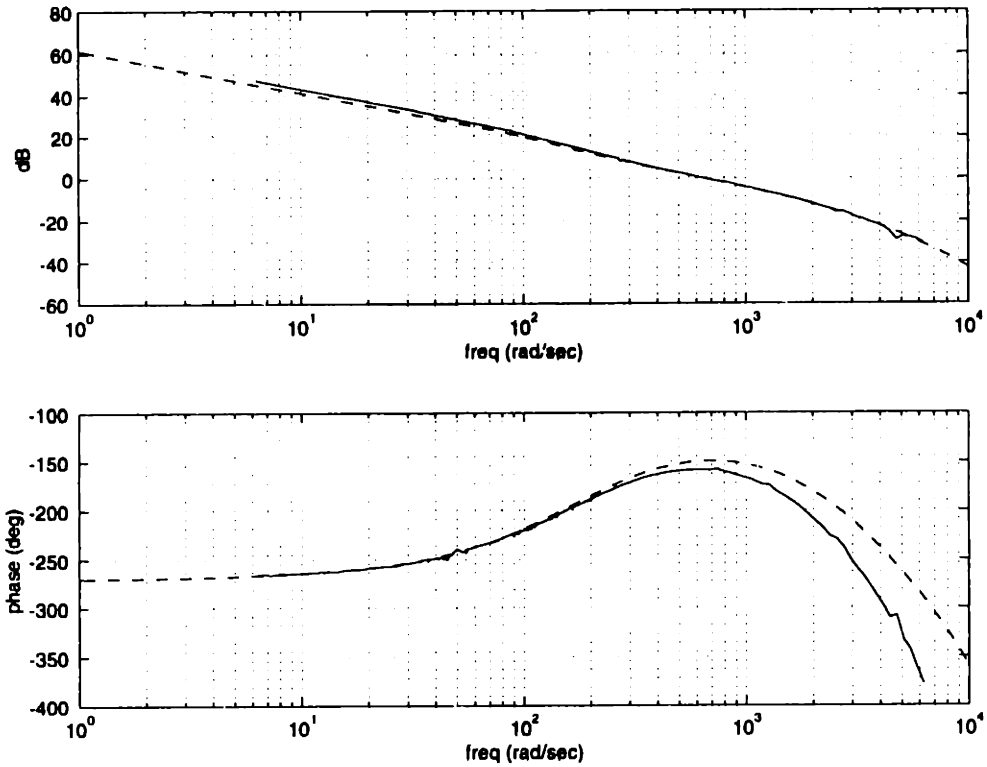

Figure 3-9: Experimental (solid) and theoretical (dashed) bode plots using (3.12).

to the reference is thought to be a major cause and contributor to the difference in the observed behaviour. In fact using the controller in (3.13), which still has a bandwidth of 100 Hz but whose lag zero is well below the lead pole, we get much better agreement in the step response. We show this in Figure 3-10.

$$G_{lead\,lag} = 34.8557\,\frac{z - 0.9513307}{z - 1}\,\frac{z - 0.9811504}{z - 0.26996} \tag{3.13}$$

The results just presented indicate that using a linear controller on the system is a valid approach provided the target remains close to the operating point about which the plant was linearized. In order to test the performance of the system at other operating points, we command 25 μm steps starting at different gap spacings. The results are shown in Figure 3-11 where the direction of the positive gap is towards actuator two.

As expected, the performance of the system gets progressively worse as the target moves away from the point used in the linearization. Instability sets in at a dis-

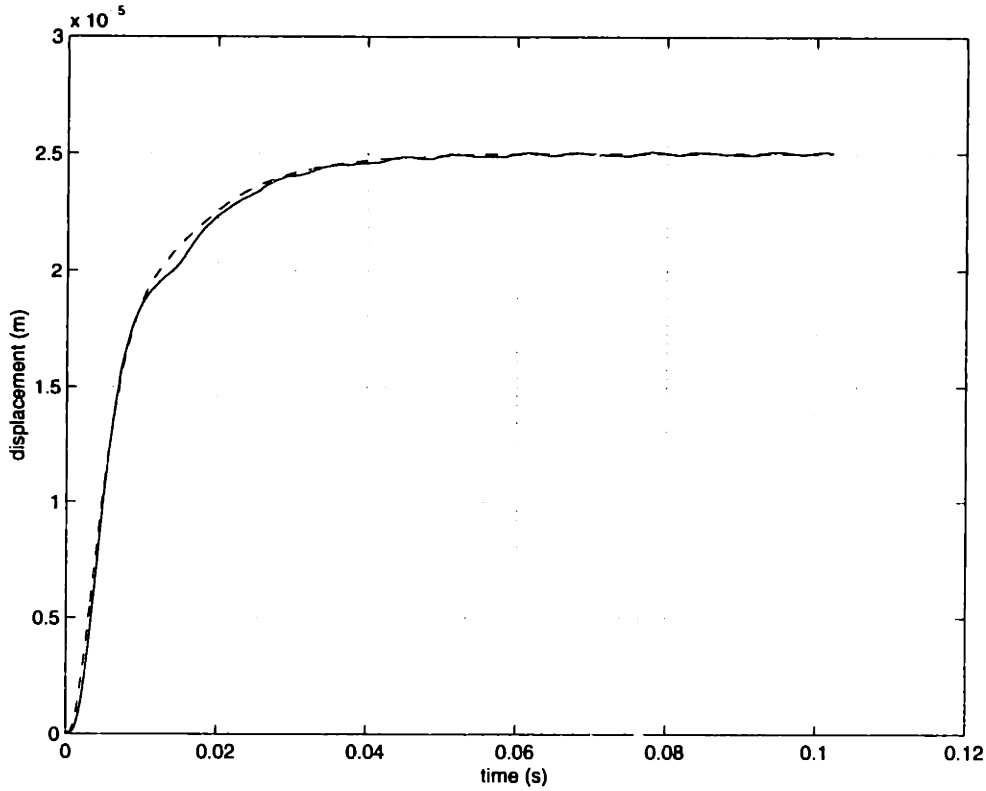

Figure 3-10: Experimental (solid) and simulated (dashed) step response of (3.13)

placement of 175 μm from the position in which the target is centered between the actuators. The same applies for steps taken in the negative direction, (i.e. towards actuator one), so the corresponding plot is not shown.

In the linearization process it is assumed that the system parameters such as the mass and the actuator constants are known exactly. A change in any of the parameters would require re-computation of the linearized system if we are to have a valid model, hence we see that linearization calls for some form of gain scheduling if the system is to operate over a large range. To test the "robustness" of the plant model (3.3) controlled by the lead-lag compensator (3.12), we input a 40 Hz sinusoidal trajectory having an amplitude of 50 μm centered at the midpoint between actuators, and then add weights to the target. The experimental responses with and without an unmodeled mass of 100 g are shown in Figure 3-12. The linearized controller is not able to accommodate unmodeled parameter variations and a mass variation of $\approx$ 8.7% produces an amplitude change of $\approx$ 8%.

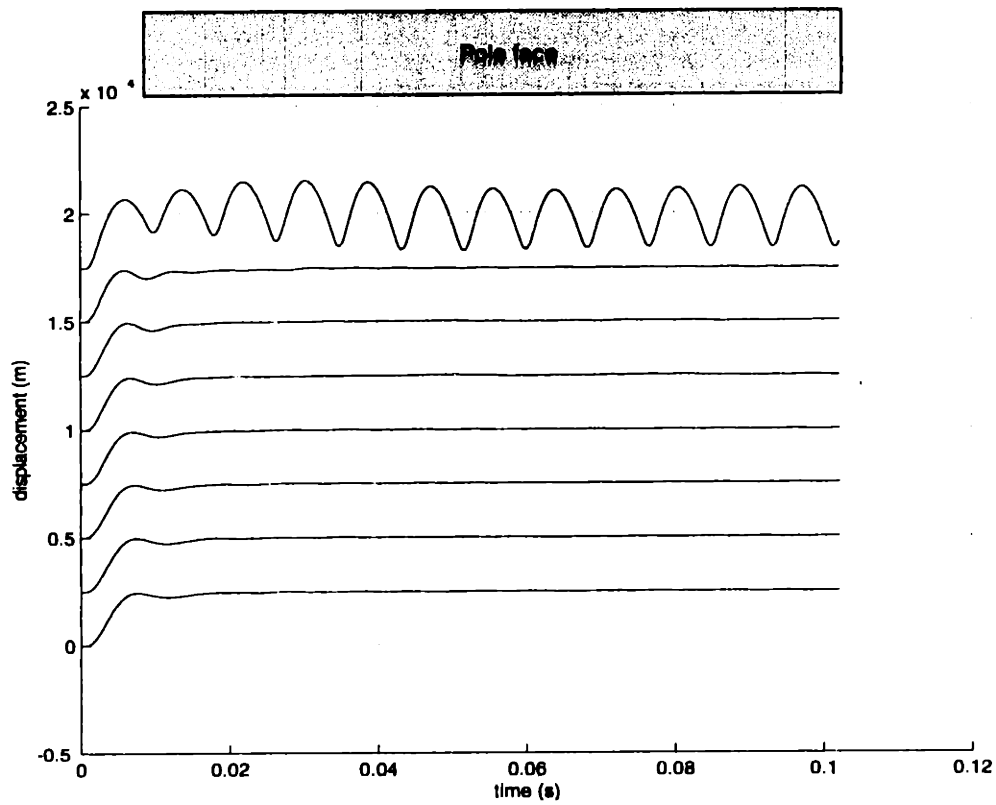

Figure 3-11: Step responses of linearized system starting at different operating points.

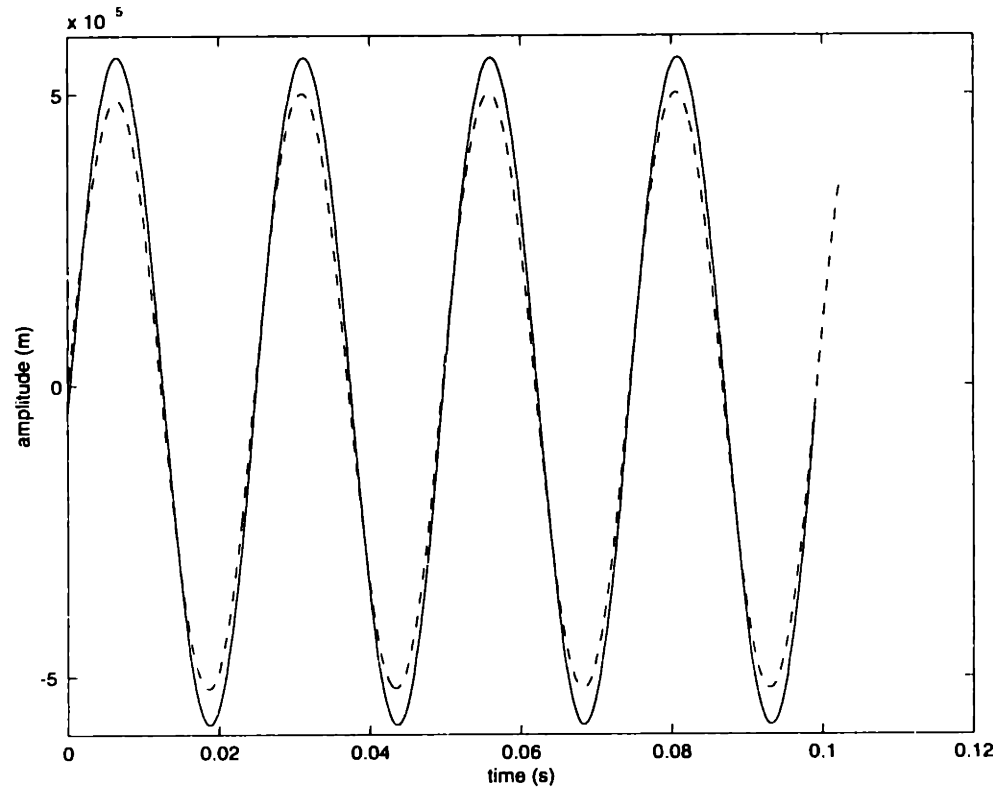

Figure 3-12: Experimental response of linearized system with (solid) and without (dashed) a 100 g mass added to the target.

Chapter 4

Feedback Linearization

4.1 Introduction

In the last chapter we have seen how a linear controller can be used on the system and we also showed that we can get good performance if the system is translating close to the point about which it was linearized. If, however, the trajectory takes the target well away from the nominal operating point, instability may result. In this chapter we will develop a first approach to nonlinear control design using a method whose intuitive nature makes it an attractive starting point in the design of more complicated controllers. In fact we shall use this technique again in the next chapter in conjunction with sliding mode control.

The attractiveness of feedback linearization lies in its ease of application to a certain category of systems. The central idea behind this approach is to transform a nonlinear system into a linear one by canceling nonlinearities through either input or state transformations (depending on the system). Successful completion of this step will leave behind a linear system which can be compensated using linear control theory [14]. The difference between this and linearization using the Jacobian of the system is that in the former the controller is made up of two parts : a set of input/state linearizing transformations and a linear controller, so that the overall controller still retains the nonlinear characteristics of the real plant. In the case of using the Jacobian, all nonlinearities are lost upon linearization.

4.2 Application of Feedback Linearization

Feedback linearization involves canceling nonlinearities and can be very directly applied to systems described by

$$x^{(n)} = f(\mathbf{x}) + b(\mathbf{x})u \tag{4.1}$$

where $x^{(n)}$ is the n^{th} derivative of the output of interest, u is the control input and $\mathbf{x}$ is the state vector. The general form of (4.1) is called Companion Form or Controllability Canonical Form and shall be important also for the development of sliding control in the next chapter. A noticeable feature of this form is that it contains no derivatives of the control input. In state space notation, we can re-write (4.1) as

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \\ .. \\ x^{(n-1)} \end{pmatrix} = \begin{pmatrix} x_2 \\ x_3 \\ .. \\ f(\mathbf{x}) + b(\mathbf{x})u \end{pmatrix} \tag{4.2}$$

which we recognize as analogous to the Controllability Canonical Form in linear control theory.

Assuming the plant is perfectly known so that $f(\mathbf{x})$ and $b(\mathbf{x})$ are fully determined we can define a transformation (4.3), where v is the new input, and insert it in (4.1) to get a multiple integrator linear system, i.e.,

$$u = \frac{1}{b(x)}[v - f(x)] \tag{4.3}$$

$$x^{(n)} = v\ . \tag{4.4}$$

Under this transformation we have canceled the nonlinearity and reduced the system to a linear one to which we can easily apply linear techniques. For example it is possible to use pole-placement in which case we could select a new input having

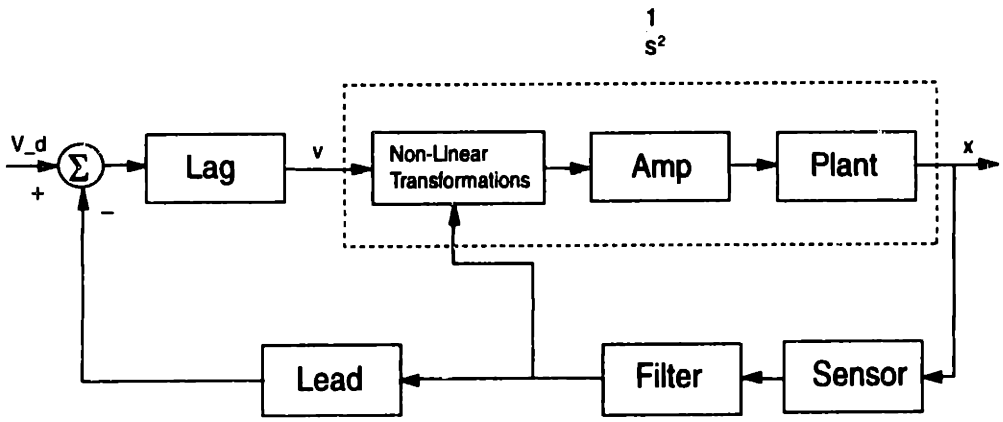

Figure 4-1: Diagram of the closed loop setup for general system with feedback linearization.

the form

$$v = -k_0 x - k_1 \dot{x} - ..... - k_{n-1} x^{(n-1)}$$

for which we could chose gains k_i such that all the poles are in the left half plane. This is possible because there are n constants and n degrees of freedom. Using full state feedback is attractive, however we do not have access to all states and therefore this approach requires the use of an estimator. The issue of estimating states shall be dealt with in the next chapter where it will be necessary to access all states in order to implement sliding control. In this chapter we will follow a classical control approach and adopt the same controller we used in the previous chapter, modifying only the gain to achieve a bandwidth of 100 Hz. Our objective is to achieve a closed-loop arrangement like the one shown in Figure 4-1.

We see clearly from the figure that there are two loops involved. The inner one carries out the linearization and establishes v as the new input to the feedback linearized plant, while the outer one stabilizes the resulting linear system. We also begin to see from the figure, and we'll see it more afterwards, that for feedback linearization we rely heavily on the plant model both for designing the controller and for computing the linearizing transformations. Hence if the model contains uncertainties these will lead to incomplete cancelations of non-linearities. These concepts can be expanded to a broader class of non-linear systems using concepts from differential geometry and topology [14]. Such generalizations, however, are not needed for the system at hand

and shall not be detailed. The state equations for the magnetic bearing were given in chapter 2 and are written again here for reference.

$$\dot{x_1} = x_2 \tag{4.5}$$

$$\dot{x_2} = \frac{C_2}{M}\left(\frac{i_2}{g_0 + x_{20} - x}\right)^2 - \frac{C_1}{M}\left(\frac{i_1}{g_0 + x_{10} + x}\right)^2 + \frac{d}{M} \tag{4.6}$$

Comparison with the general form of equation (4.1) shows that we can rewrite these equations as (4.2) if, at any one time, we drive the plant with one current and set the other to zero. With this control strategy, and seeing that there are no derivatives of the input, the plant is in companion form. We also note that with this "switching" current, the equations are directly input-state linearized. To show that the plant is also input-output linearized we note that the output of the plant is the position of the target so

$$y = x \tag{4.7}$$

from which we can say

$$\ddot{y} = \ddot{x} = f(\mathbf{x}) + b(\mathbf{x})u \ . \tag{4.8}$$

The output is then directly controlled by the input.

To transform it into a SISO system in the new input v, we can use the following transformations.

$$\text{for v} \geq 0 \ \Rightarrow \quad i_1 = 0 \qquad i_2 = (g_0 + x_0 - x_1)\sqrt{\frac{M\,v}{C_2}} \tag{4.9}$$

$$\text{for v} < 0 \ \Rightarrow \quad i_1 = (g_0 + x_0 + x_1)\sqrt{\frac{-M\,v}{C_1}} \qquad i_2 = 0$$

Inserting these transformations into the state equations, we see that the plant will be linearized to

$$\dot{x_1} = x_2$$

$$\dot{x_2} = v$$

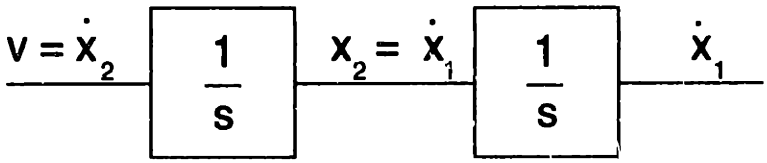

Figure 4-2: A perfectly linearized system is equivalent to a series of integrators.

which is a double integrator as shown in Figure 4-2 [7].

We implement the same controller used in the previous chapter, equation (3.12), with the difference that the gain is modified to set the bandwidth at 100 Hz. The controller is

$$G_c = 75\,\frac{z-0.9601}{z-1}\,\frac{z-0.965508}{z-0.481909} \tag{4.10}$$

Figure 4-3 shows response of commanding a 25 μm step in displacement from the equilibrium position. The response shows the same deviation from the predicted response as was seen in the last chapter. Using the same setup as Figure 3-6 we measure the loop transmission; this is shown in Figure 4-4.

We showed that the linearized plant should be equivalent to a series of cascaded integrators and we check this claim by dividing out the controller from the experimental loop transmission to recover the plant transfer function shown in Figure 4-5. As expected the magnitude plots are in excellent agreement; the phase plots however display the same inconsistency highlighted previously. The loop transmission plots show the existence of unmodeled resonance around 4800 rad/s. The magnitude plot would indicate a set of complex conjugate zeros breaking before a set of complex conjugate poles which we speculate as being due the vertical resonance of the shaft. Referring to Figure 2-6 we see that there is a vertical resonance at this frequency, and since the controller has virtually no effect this far above the bandwidth, it will not compensate for it.

For the linear controller of the previous chapter, performance degrades when the operating point moves away from the one about which the plant was linearized. In feedback linearization the transformations are functions of the displacement, x, and are therefore dynamically compensating for the target position; this ideally makes the

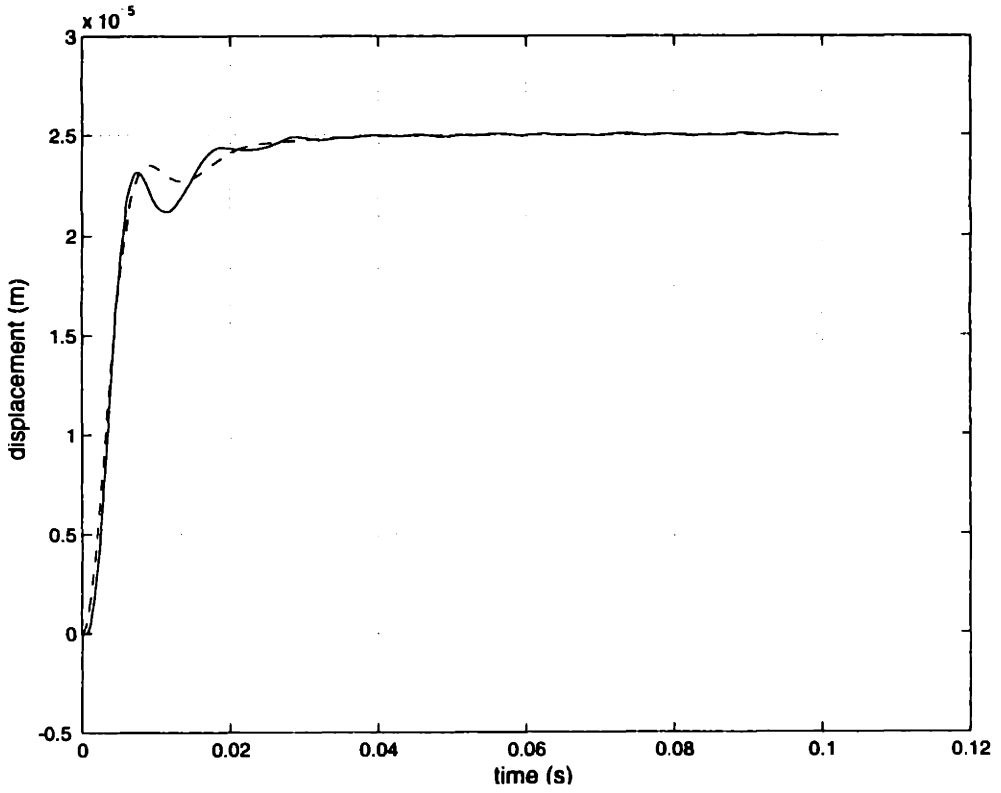

Figure 4-3: Experimental (solid) and theoretical (dashed) response of system controlled with (4.10) subject to a 25 μm step in displacement

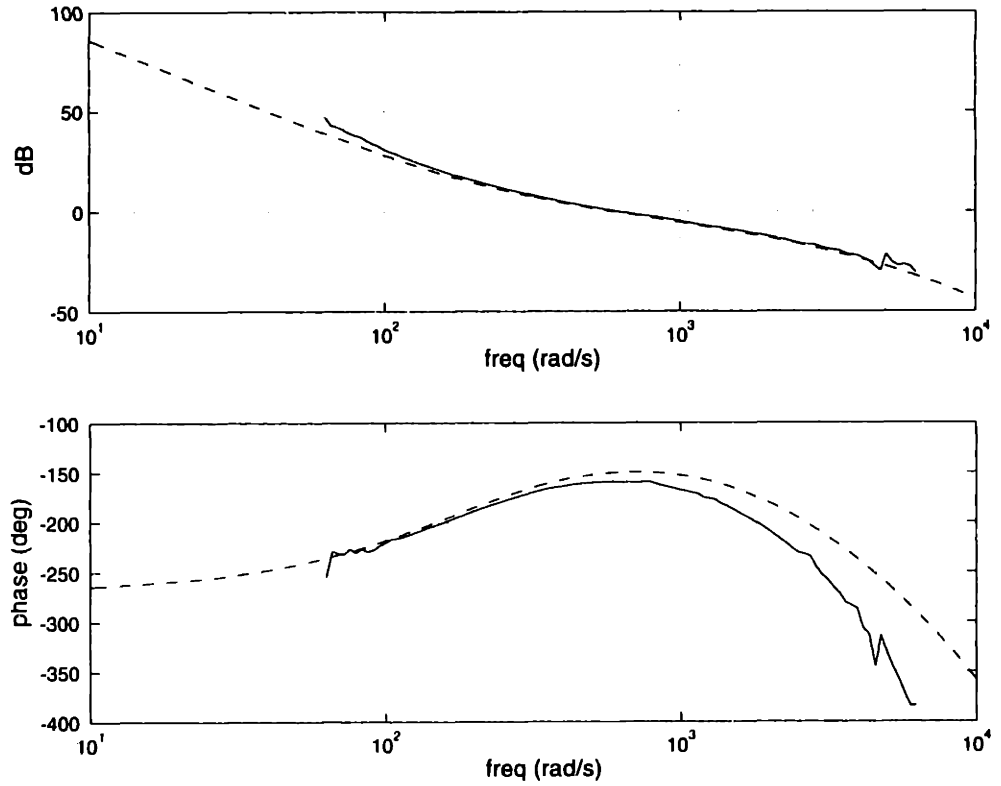

Figure 4-4: Exerimental (solid) and theoretical (dashed) bode plot for the lead-lag compensated system

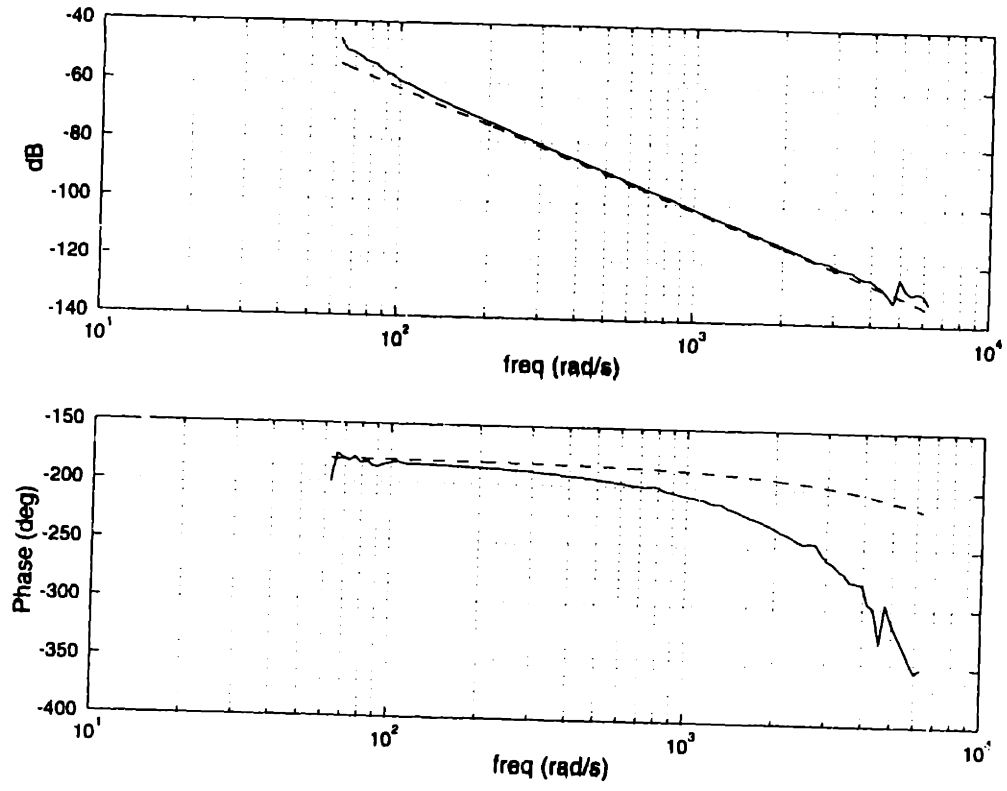

Figure 4-5: Exerimental (solid) and theoretical (dashed) bode plot of the feedback linearized plant

system independent of the operating point. We test this by commanding 25 μm step displacements at various air-gaps. Unlike the case where we used the linear controller alone, using feedback linearization allows the target to be controlled over the full range of motion until it sticks to the pole face of one of the electromagnets. Figure 4-6 shows this. With perfect cancelation of the nonlinearities the responses should be identical regardless of the operating point, we see however that as the target approaches the pole face the responses become more oscillatory. This behaviour is expected since we have seen how the model is not exact, and thus perfect linearization is not being achieved. This is a good example of the non-robust nature of this technique. A similar analysis was carried out in Subrahmanyan et al. [7] however in that case the plant model was known almost exactly and close to total cancelation of the nonlinearities was achieved. To check the effect of unmodeled mass variation Figure 4-7 shows the response of adding 100 g to the target and referencing a 40 Hz, 50 μm amplitude, sinusoidal reference trajectory.

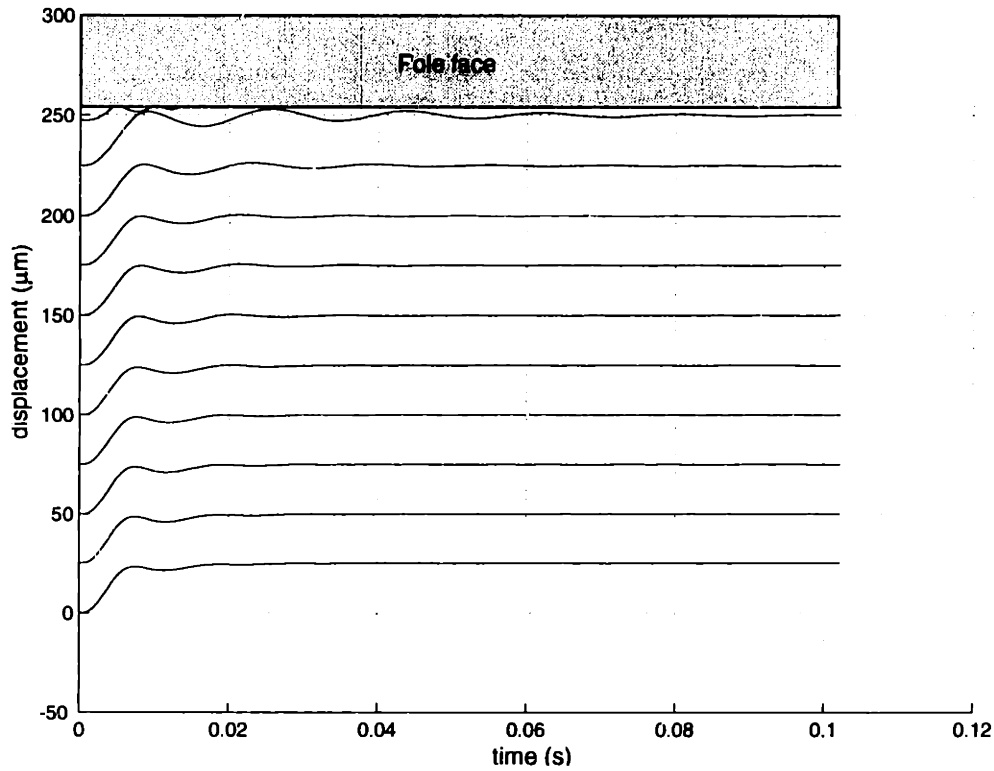

Figure 4-6: 25 μm step responses for various air gaps.

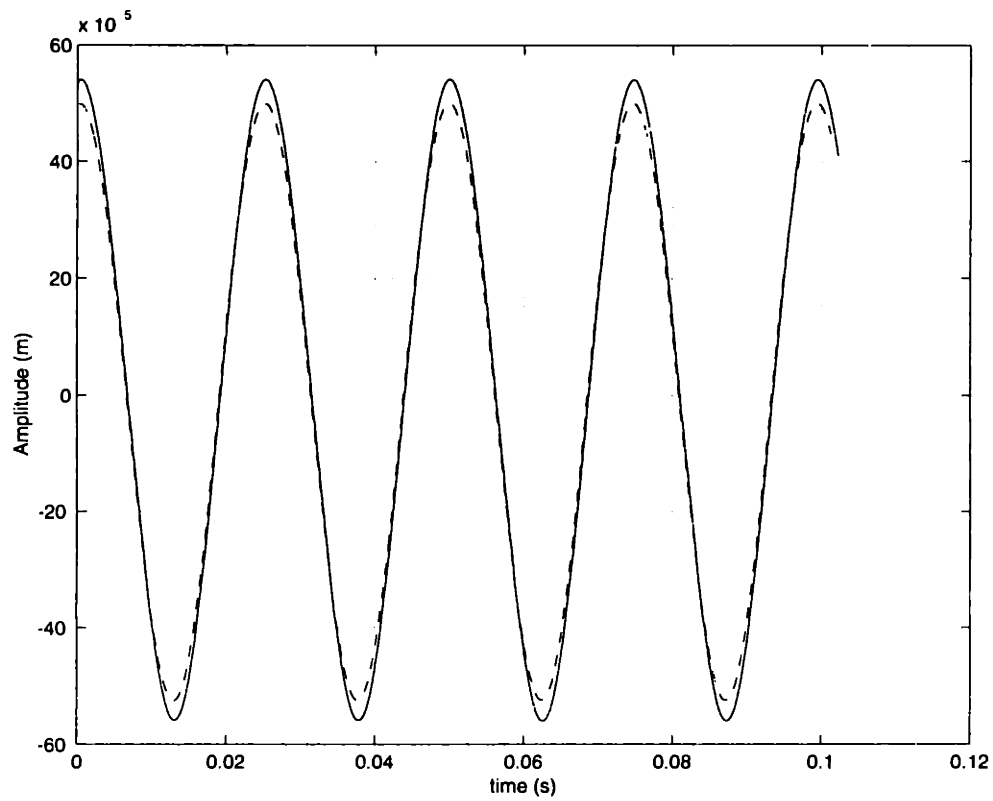

Figure 4-7: Sinusoidal response with (solid) and without (dashed) the unmodeled mass of 100 g.

Chapter 5

Sliding Controller

5.1 Introduction to sliding control

We have seen in the previous chapters that we can achieve good performance using both simple linear controllers and feedback linearizing controllers, but it was also shown that both of these are subject to limitations on robustness to parameter, and operating point, variations as well as on parametric uncertainties. To overcome these shortcomings, in this chapter we use a technique from robust control theory which allows us to include our estimated levels of parametric uncertainties in the controller. The resulting controller will then guarantee stability and robustness at least with respect to the uncertainties included in the controller design. This chapter is not meant to be a course in sliding control, rather it is aimed at presenting the main results which will then be used on the magnetic bearing. For more information on sliding control methodologies, consult references such as [14].

A fact that is easily encountered in linear control is that first order systems are inherently simpler to control than are higher order ones. The difference is mostly mathematical, but low order systems also allow a certain degree of physical insight which may not be there when dealing with complicated higher order dynamics. The idea behind sliding control is quite simply to reduce the system to a stabilization problem in one dimension by means of a particular choice of a variable called the "sliding variable". We shall also see that the sliding variable is not only a form of

state variable, but it is also a dynamics towards which the system is directed.

Sliding control in its basic form can ideally deliver perfect performance by the use of a discontinuous control law, however this approach is hardly ever adopted because the discontinuities are directly related to the parametric uncertainty levels which are being built into the plant model, and thus can be quite large. The result is that the system will reach the desired dynamics, then enter a limit cycle in which chattering occurs leading thus to extremely high control activity and excitation of unmodeled resonant modes of the structure. For this reason a smoothed version of the switching controller is preferred which renders the control more feasible but has the disadvantage of diminishing the accuracy of the controller. The approach to designing the smoothed controller involves first designing the switching controller and then smoothing it out it in a boundary layer around the desired dynamics.

5.2 Switching Control

$$x^{(n)} = f(\mathbf{x}, t) + b(\mathbf{x}, t)u \tag{5.1}$$

Let us consider a system described by equation (5.1) where $f(\mathbf{x}, t)$ and $b(\mathbf{x}, t)$ are generally non-linear functions of the state vector, $\mathbf{x}$, and possibly time. The output of interest is the scalar x and u is the control input applied to the system. This is not the only form to which sliding control can be applied, but it is the one applicable to the magnetic bearing being used. The state vector is as follows :

$$\mathbf{x} = [x \quad \dot{x} \ .... \ x^{(n-1)}]^T \tag{5.2}$$

where n is the order of the system. Analogous to the above, the state vector of desired states is defined from the desired reference trajectory, x_d, and contains the derivatives of the desired trajectory.

$$\mathbf{x_d} = [x_d \quad \dot{x}_d \ .... \ x_d^{(n-1)}]^T \tag{5.3}$$

While as the exact non-linear parameters may not be known, it is assumed that bounds can be placed on them either from physical principles, intuition or by other

means. These bounds are expressed as

$$|f(\mathbf{x},t) - \hat{f}(\mathbf{x},t)| \leq F \tag{5.4}$$

$$0 < b_{min} \leq b(\mathbf{x},t) \leq b_{max} \tag{5.5}$$

where $\hat{f}$ is the estimate of the $f(\mathbf{x},t)$ and b_{min}, b_{max} are the upper and lower bounds of $b(\mathbf{x},t)$.

As a new state variable, define the sliding variable to be a weighted sum of errors in the states

$$s \equiv \left(\frac{d}{dt} + \lambda\right)^{(n-1)} \tilde{x} \tag{5.6}$$

where the weighing parameter, λ, is strictly positive and can be shown to be the equivalent bandwidth of the controller [13]; $\tilde{x}$ is the scalar output error

$$\tilde{x} = x - x_d \ . \tag{5.7}$$

With this definition of the sliding variable we see that since s is a weighted sum of the state errors, it represents a measure of the deviation between the desired state vector and the actual state vector. We can now define a sliding surface, S(**x**,t), as

$$S(\mathbf{x},t) \ : \quad s(\mathbf{x},t) = 0 \tag{5.8}$$

Perfect tracking requires that all the state errors vanish, so from (5.6) this is seen to be analogous to requiring s to go to zero. We also notice that s is a dynamic system because it is state and time dependent. When the system has reached $s = 0$ the system trajectory will be determined by the sliding surface itself

$$\left(\frac{d}{dt} + \lambda\right)^{(n-1)} \tilde{x} = 0 \tag{5.9}$$

Differentiating (5.6) with respect to time gives

$$\dot{s} = h(\dot{x}, \ddot{x}, ..., x^{(n)}) = h(\dot{x}, \ddot{x}, ..., f(\mathbf{x}, t) + b(\mathbf{x}, t)u) \tag{5.10}$$

where h represents the function

$$\frac{d}{dt}\left[\left(\frac{d}{dt} + \lambda\right)^{(n-1)} \tilde{x}\right] . \tag{5.11}$$

The input has appeared directly and the problem has been reduced to a stabilization problem of a first order differential equation in s.

Perfect tracking is analogous to steering the system towards the sliding surface and therefore we can chose a control law that aims at minimizing the square error in the distance to the surface. The sliding condition from which we derive the control law can therefore be chosen to ensure that the gradient of the square of the distance to the surface is always negative. This ensures that the sliding surface is an attractive invariant set so that all system trajectories outside this surface will point towards it. Formally we express it as

$$\frac{1}{2}\frac{d}{dt}s^2 \leq -\eta|s| \tag{5.12}$$

where the constant η is strictly positive and determines the power of attraction of the surface and hence the "transient" time required by the dynamics to reach the surface in the event in which the initial conditions differ from the desired ones. Given an initial state error, the trajectory will reach the sliding surface in a time interval $t_{s=0}$ given by

$$t_{s=0} \leq \frac{|s(0)|}{\eta} . \tag{5.13}$$

From the sliding condition we calculate the control law and from the first order stabilization problem (5.10) the best estimate of the control effort required to bring $\dot{s}$ to zero is

$$\hat{u} = -\hat{f}(\mathbf{x}, t) - \sum_{p=1}^{n-1} \binom{n-1}{p} \lambda^p \tilde{x}^{(n-p)} . \tag{5.14}$$

where $\hat{u}$ is our best estimate of the actual control effort u required to reach the sliding

surface, and $\hat{f}(\mathbf{x},t)$ is our best estimate of $f(\mathbf{x},t)$.

If the plant were perfectly known then this would be the exact effort required, but to account for the uncertainties in the model, a correction factor of magnitude k is introduced which is discontinuous across the sliding surface. This gives the full control effort as

$$u = -\hat{f}(\mathbf{x},t) - \sum_{p=1}^{n-1} \binom{n-1}{p} \lambda^p \tilde{x}^{(n-p)} - k\, sgn(s) \tag{5.15}$$

where the $sgn(s)$ function is defined mathematically as

$$sgn(s) = \begin{cases} +1 & : \quad s \geq 0 \\ -1 & : \quad s < 0 \end{cases} \tag{5.16}$$

The lower bound on k which will guarantee attractiveness of the sliding surface is obtained by satisfying the sliding condition. We note that we cannot derive a unique solution for k because for some of the parameters in the system we know only the bounds, and not the exact values. Substituting (5.15) back into (5.12) and solving, gives the resulting inequality for k as a function of β, η, F, $\hat{u}$, D. More compactly the control law (5.15) can be written as

$$u = \hat{u} - k\, sgn(s)\ . \tag{5.17}$$

where $\hat{u}$ was presented earlier as the best estimate of the control effort required to bring the trajectory to the sliding surface, and $k\, sgn(s)$ is the discontinuous term which compensates for model uncertainties.

This discontinuous control law will give perfect performance at the price of extremely high control activity, the latter being due to chattering induced by the discontinuous term. In Figure 5-1 we show a representative example of a second order system sliding surface (which is a line in the phase plane defined by the two state variables) and how a trajectory is attracted to it. Overshooting the surface on either side will lead to the chattering phenomenon showed.

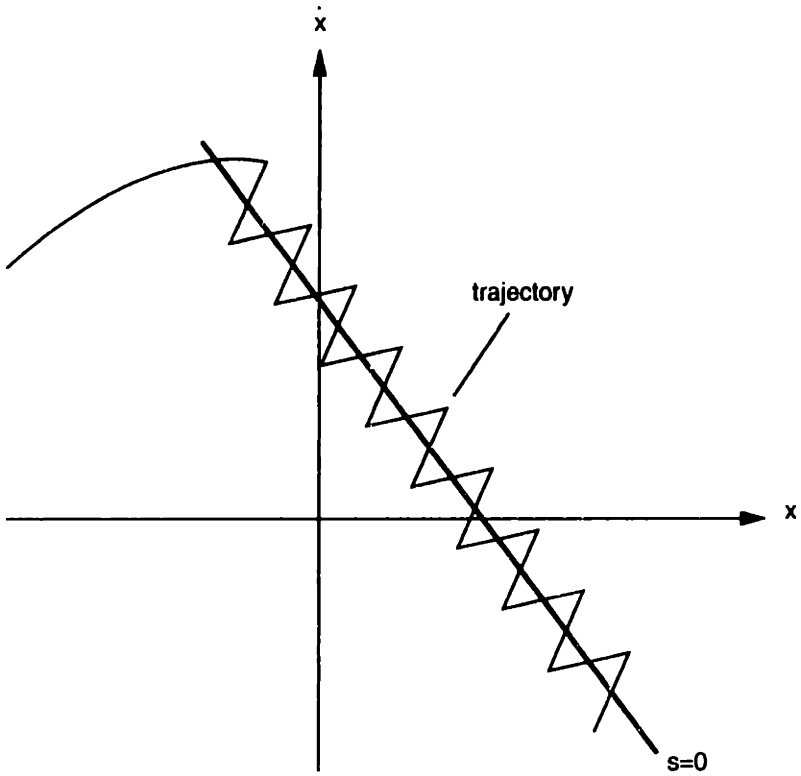

Figure 5-1: Chattering in a second order system induced by the discontinuous control law.

For the magnetic bearing studied in this thesis, the control algorithm on the DSP will be switching between actuators at very high frequencies possibly driving the D/A to saturation every time, which makes the control unfeasible and impractical on a real system. This effect will be shown later in this chapter.

5.3 Sliding Control

In the previous section we presented the switching controller and detailed the steps required to design one. At the end of the section we also stated that the practical problems encountered in applying this controller force us to seek ways to smoothen out the control discontinuities and obtain a continuous control effort. In this section we present the "continuous approximation to sliding control", otherwise simply referred to as sliding control, which is an extension of the switching controller. The fundamental difference between the two is that while the switching controller is dras-

tic in commanding the effort as the sliding variable passes from $s > 0$ to $s < 0$ or vice-versa, the sliding controller provides an effort which is not discontinuous. It achieves this by using a different control law when the sliding variable is within a thin boundary layer around the sliding surface.

When s is outside the boundary layer the same switching control laws derived in the previous section apply and this makes the boundary layer an attractive invariant set. However, when s enters the boundary layer we substitute the switching law with a linear interpolation, to calculate the effort. The effect of this can be shown to be the same as adding another lowpass filter to the dynamics of s. It is precisely this lowpass action that eliminates the chattering. It suffices to say that λ is normally chosen as the bandwidth of this filtering action, so while we would want to increase λ to get faster response times, increasing λ also reduces attenuation of chattering. Later in this section we shall present some guidelines for selecting λ.

Let us define a boundary layer, $B(t)$, as

$$B(t) = \{x, |s(\mathbf{x}, t)| \leq \Phi(t)\} \qquad \Phi(t) > 0 \tag{5.18}$$

where $\Phi(t)$ is the time varying boundary layer thickness and is strictly positive. Use of the switching control laws when $|s| > \Phi$ ensures that all trajectories will converge to the boundary layer, however once they reach it, the objective is now to keep the trajectory inside the boundary layer and aiming toward $s = 0$. Sastry and Slotine [13], and many others since, have shown that we can achieve this by substituting the discontinuous $k\,sgn(s)$ term in (5.15) with a linear interpolator $\bar{k}(\mathbf{x})\frac{s}{\Phi}$ inside the boundary layer. $\bar{k}$ is a new gain term modulated according to the rate of change of the boundary layer. Intuitively, if $\dot{\Phi} < 0$, so that the boundary layer is shrinking, $\bar{k}$ must increase in order to maintain the trajectory within the layer; if instead $\dot{\Phi} > 0$ then we can relax the control effort by decreasing $\bar{k}$ and still remain within the boundary layer. This smoothed implementation is achieved by substituting the term $k\,sgn(s)$

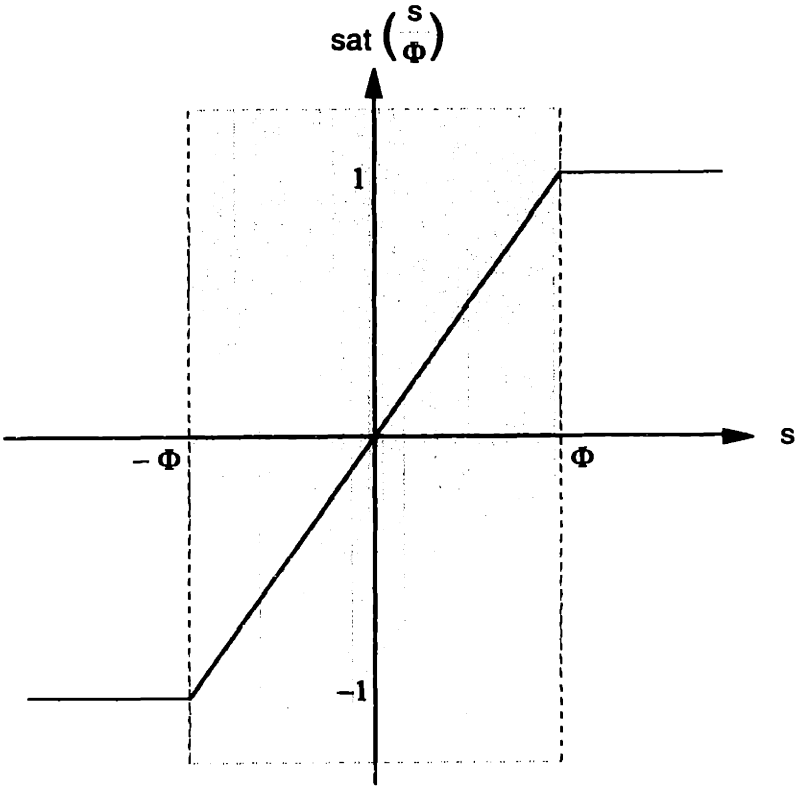

Figure 5-2: Plot of $sat\left(\frac{s}{\Phi}\right)$.

with the term $\bar{k}\, sat\left(\frac{s}{\Phi}\right)$, from which we re-write the complete control law as

$$u = -\hat{f}(\mathbf{x}, t) - \sum_{p=1}^{n-1} \binom{n-1}{p} \lambda^p \tilde{x}^{(n-p)} - \bar{k}\, sat(\frac{s}{\Phi}) \ . \tag{5.19}$$

The saturation function is defined mathematically as

$$sat(s) = \begin{cases} sgn(s) & : \ |s| > 1 \\ s & : \ |s| \leq 1 \end{cases} \tag{5.20}$$

so we see that if s positive and outside the boundary layer, $(s > \Phi)$, it follows that $\frac{s}{\Phi} > 1$ and therefore $sat\left(\frac{s}{\Phi}\right) = sgn\left(\frac{s}{\Phi}\right) = 1$. Similarly if $s < -\Phi$ then $sat\left(\frac{s}{\Phi}\right) = -1$. If instead $|s| < \Phi$ then the discontinuous part of the control law is replaced altogether. We see therefore that the only discontinuity that could have occurred was if the sgn function was valid across $s = 0$, but it's not. Figure 5-2 shows the plot of $sat\left(\frac{s}{\Phi}\right)$.

This approach requires us to derive a relation for the gain $\bar{k}$ as well as deriving the differential equations governing the boundary layer. Again constraining the rate of change of the square of the distance to the sliding surface to be negative, we can redefine the sliding condition as

$$\frac{1}{2}\frac{d}{dt}s^2 \leq (\dot{\Phi} - \eta)|s| \tag{5.21}$$

where the power of attraction is now $(\dot{\Phi} - \eta)$. We do this for the same reason we modulate $\bar{k}$: during boundary layer contraction the power of attraction needs to increase to ensure that the trajectory stays within the boundary layer. It has been shown [13] that the new sliding condition, (5.21), is satisfied by the following set of equations referred to as the balance conditions.

$$k(\mathbf{x}_d) \geq \frac{\lambda\Phi}{\beta_d} \quad \Rightarrow \quad \dot{\Phi} + \lambda\Phi = \beta_d k(\mathbf{x}_d) \qquad \bar{k}(\mathbf{x}) = k(\mathbf{x}) - \frac{\dot{\Phi}}{\beta} \tag{5.22}$$

$$k(\mathbf{x}_d) \leq \frac{\lambda\Phi}{\beta_d} \quad \Rightarrow \quad \dot{\Phi} + \frac{\lambda\Phi}{\beta_d^2} = \frac{k(\mathbf{x}_d)}{\beta_d} \qquad \bar{k}(\mathbf{x}) = k(\mathbf{x}) - \dot{\Phi}\beta \tag{5.23}$$

where the subscript d refers to the value of that parameter evaluated along the desired trajectory and the gain margin β is defined as

$$\beta = \sqrt{\frac{b_{max}}{b_{min}}} \tag{5.24}$$

so that bound (5.5) can be re-written as

$$\beta^{-1} \leq \frac{\hat{b}}{b} \leq \beta \tag{5.25}$$

where $\hat{b}$ is the geometric mean of the bounds on b. In the balance conditions the gain k is the same one derived for the switching controller, and the gain $k(\mathbf{x}_d)$ is the gain of the switching controller evaluated along the desired trajectory which is found from $k(\mathbf{x}$ by setting $\tilde{x} = \dot{\tilde{x}} = \ddot{\tilde{x}} = ... = 0$ because on the desired trajectory there are no

errors. The initial condition for the boundary layer can be shown to be

$$\Phi(0) = \frac{\beta_d k(\mathbf{x}_d(0))}{\lambda} \tag{5.26}$$

The sliding controlled is thus defined by (5.19), (5.22) and (5.23).

A key step in the implementation of sliding control is the necessity to decide on a suitable value for λ. As we said before this parameter represents the equivalent bandwidth of the controller. Ideally it is desirable that λ be large, but practical limitations pose limits on the values this parameter can take. Slotine and Li [14] give typical limitations as due to three effects :

- structural resonant modes

  $$\lambda \leq \frac{2\pi}{3}\nu_R$$

  where ν_R is the first unmodeled resonant frequency of the structure, measured in Hertz. In most practical cases this criterion will be the one limiting the choice of λ, however this bound can be changed as a result of experimental evaluation and engineering judgment. Back in chapter two, we used this to decide on the desired bandwidth.

- neglected time delays

  $$\lambda \leq \frac{1}{3T_A}$$

  where T_A is largest unmodeled time-delay which in the case of the magnetic bearing is due to the execution of the ISR.

- sampling rate

  $$\lambda \leq \frac{1}{5}\nu_{sampling}$$

  where $\nu_{sampling}$ is the sampling rate.

5.4 Sliding Control with mass uncertainty

The magnetic bearing system contains two components whose characteristic values introduce the most variability. These are the equivalent mass of the target and the actuator constants. While as the mass of the target is easily determined by simply weighing it, the equivalent translational mass of the target and arm is not easily determinable to a high degree of accuracy. As for the actuators, calibration gives a good estimate of the actuator constant but the true value is still not known especially if the actuators are driven to saturation. In any case the aim is to make the system robust to parameter changes, so we must assume that some or all the parameters are not well known.

In this section only the mass uncertainty is accounted for, and we assume that the actuator constants are accurately known. In the next section both mass and actuator uncertainties will be dealt with.

Sliding control is used here in conjunction with feedback linearization where the latter is used to produce an equivalent linear system as was described in chapter four. Unlike using feedback linearization alone where the closed loop system is highly model-dependent, using sliding control on the feedback linearized system allows us to specifically take the model uncertainties into consideration when designing the controller and thus make the closed loop system robust at least with respect to the modeled uncertainties.

We write here again the plant equations for the feedback linearized system, where the non-linear transformations are slightly different to the ones presented in the last chapter: namely, the mass and the actuator constants are not included in the linearizing transformations and are therefore left in the resulting plant equation.

$$u \geq 0 \quad \Rightarrow i_2 = (g_0 + x_0 - x)\sqrt{u} \quad \text{and} \quad i_1 = 0 \tag{5.27}$$

$$u \leq 0 \quad \Rightarrow i_1 = (g_0 + x_0 + x)\sqrt{-u} \quad \text{and} \quad i_2 = 0 \tag{5.28}$$

The plant equation is re-written as

$$\frac{m}{C}\ddot{x} = u + \frac{d}{C} \ . \tag{5.29}$$

A single value for actuator constant, C, is used to make the plant model conform to (5.1). Comparing (5.29) to (5.1) we get that for the magnetic bearing with mass uncertainty

$$f(\mathbf{x}, t) = 0 \quad \text{and} \quad b(\mathbf{x}, t) = \frac{C}{m} \tag{5.30}$$

The value of C used in this section will be the geometric mean of the actual values as measured by the calibration procedure: $C = \sqrt{C_1 C_2} = 2.69 \times 10^{-6} \frac{\text{Nm}^2}{\text{A}}$. In designing the continuous sliding control, it is necessary to first design the switching controller.

To characterize the mass uncertainty, a range for the actual value of the mass must be specified, and since the mass enters the plant equation in a multiplicative fashion (see equation (5.29)), it seems reasonable to define the estimate of the mass as the geometric mean of its uncertainty range.

$$\text{Uncertainty :} \qquad m_{min} \leq m \leq m_{max} \ \Rightarrow \ 1.035 \ \leq \ m \ \leq \ 1.265 \text{ kg} \tag{5.31}$$

This range represents a $\pm 10\%$ deviation from the nominal value of 1.15 kg used in the previous chapters. This level of uncertainty is picked quite arbitrarily and a higher level could have been used, but we feel that with this choice the real mass uncertainty is covered.

$$\text{Estimate of mass :} \quad \hat{m} = \sqrt{m_{min}\, m_{max}} \tag{5.32}$$

$$\text{Estimate of b :} \quad \hat{b} = \frac{C}{\sqrt{m_{min}\, m_{max}}} \tag{5.33}$$

It is necessary to bound the mass uncertainty as well as the disturbance uncertainty,

so we can postulate

$$|m - \hat{m}| \leq M \tag{5.34}$$

$$|d - \hat{d}| \leq D \tag{5.35}$$

where $\hat{d}$ is the geometric mean of the estimated disturbance given by an equation analogous to (5.32). The gain margin for the plant was given in (5.24) which for the case of only mass uncertainty can be re-written using (5.30) as

$$\beta = \sqrt{\frac{m_{max}}{m_{min}}} \quad . \tag{5.36}$$

Referring to (5.6), the sliding surface for the second order plant ($n = 2$), is

$$\dot{\tilde{x}} + \lambda \tilde{x} = 0 \tag{5.37}$$

which implies that after convergence to the sliding surface, the tracking error decreases exponentially with a time constant of $\frac{1}{\lambda}$. This fortifies our interpretation of λ as the equivalent bandwidth of the controller.

Differentiating (5.37) with respect to time we get the equivalent stabilization problem in the sliding variable

$$\dot{s} = \ddot{\tilde{x}} + \lambda \dot{\tilde{x}} \ . \tag{5.38}$$

We can now use (5.7) to rewrite the above as a first order equation in s containing the control input, just as we showed in (5.10):

$$\dot{s} = b(\mathbf{x}, t)u + \frac{d}{m} - \ddot{x}_d + \lambda \dot{\tilde{x}} \ . \tag{5.39}$$

The estimate of the control effort required to bring the sliding variable to the surface is found using (5.14).

$$\hat{u} = -\frac{\hat{d}}{\hat{m}} + \ddot{x}_d - \lambda \dot{\tilde{x}} \tag{5.40}$$

and the discontinuous control effort will then be given by (5.17) but with a correction

factor in front to account for the uncertainty in b.

$$u = \hat{b}^{-1}[\hat{u} - k\, sgn(s)] \tag{5.41}$$

Inserting (5.40) and (5.41) into (5.39)

$$\dot{s} = \frac{b}{\hat{b}}\left(-\frac{\hat{d}}{\hat{m}} + \ddot{x}_d - \lambda\dot{\tilde{x}} - k\, sgn(s)\right) + \frac{d}{m} - \ddot{x}_d + \lambda\dot{\tilde{x}} \tag{5.42}$$

Using (5.30) we can substitute $\frac{b}{\hat{b}} = \frac{\hat{m}}{m}$ and then substitute (5.42) into the sliding condition (5.12)

$$s\left[\left(\frac{\hat{m}}{m} - 1\right)\ddot{x}_d + \left(1 - \frac{\hat{m}}{m}\right)\lambda\dot{\tilde{x}} + \frac{d - \hat{d}}{m} - \frac{k}{m}\hat{m}\, sgn(s)\right] \leq -\eta|s| \ . \tag{5.43}$$

Taking advantage of the fact that $s\, sgn(s) = |s|$, (5.43) can be rearranged to find k.

$$k\,|s| \geq \eta\frac{m}{\hat{m}}|s| + \frac{m}{\hat{m}}s\left[\left(\frac{\hat{m}}{m} - 1\right)\ddot{x}_d + \left(1 - \frac{\hat{m}}{m}\right)\lambda\dot{\tilde{x}} + \frac{d - \hat{d}}{m}\right]$$

$$k \geq \eta\frac{m}{\hat{m}} + \frac{m}{\hat{m}}\, sgn(s)\left[\left(\frac{\hat{m}}{m} - 1\right)\ddot{x}_d + \left(1 - \frac{\hat{m}}{m}\right)\lambda\dot{\tilde{x}} + \frac{d - \hat{d}}{m}\right]$$

Being an inequality there is not a unique solution, so the bound will still be satisfied if we set

$$k \geq \eta\frac{m}{\hat{m}} + \frac{m}{\hat{m}}\left|\left(\frac{\hat{m}}{m} - 1\right)\ddot{x}_d + \left(1 - \frac{\hat{m}}{m}\right)\lambda\dot{\tilde{x}} + \frac{d - \hat{d}}{m}\right| \ .$$

We can further manipulate the inequality by noting that

$$(\beta - 1) \geq \left(\frac{\hat{m}}{m} - 1\right)$$

and that

$$|a + b| \leq |a| + |b|$$

which allow us to take the disturbance uncertainty term out of the modulus and get

a relation for the discontinuous gain k

$$k \geq \beta\eta + (\beta - 1)|-\ddot{x}_d + \lambda\dot{\tilde{x}}| + \frac{D}{\hat{m}} \tag{5.44}$$

The switching controller is therefore defined by equations (5.32), (5.33), (5.40), (5.41) and (5.44).

We note that $k \neq k(\tilde{x})$ and $\hat{u} \neq \hat{u}(\tilde{x})$ thus the controller may not deliver excellent performance in response to slow varying dynamics even though $\tilde{x}$ may be large. Equation (5.44) shows that the bounds on k depend upon $\ddot{x}_d$ which makes it impossible for us to apply a step reference to the plant because the second derivative of a step is undefined. We cannot, therefore, obtain step response plots like the ones shown in previous chapters.

It was previously stated that a switching controller is hardly ever used as it is because of the chattering problems it leads to. We can see this by simulating the response of the system to the switching controller just derived. Figure 5-3 shows how the voltage output from channel one of the D/A would look like if we command a 40 Hz sinusoidal trajectory having an amplitude of 50 μm. The D/A would saturate at almost every cycle; to avoid this we now design and implement a sliding controller

For the magnetic bearing the sliding controller for the case of only mass uncertainty is given by the following set of equations.

$$u = \frac{\sqrt{m_{min}\, m_{max}}}{C}\left(\hat{u} - \bar{k}\, sat\left(\frac{s}{\Phi}\right)\right)$$

$$\hat{u} = \ddot{x}_d - \lambda\dot{\tilde{x}} - \frac{\hat{d}}{\hat{m}}$$

$$k \geq \beta\eta + (\beta - 1)|-\ddot{x}_d + \lambda\dot{\tilde{x}}| + \frac{D}{\hat{m}}$$

$$\left.\begin{matrix}\bar{k} \\ \Phi\end{matrix}\right\} \text{ from Balance Conditions (5.22), (5.23)}$$

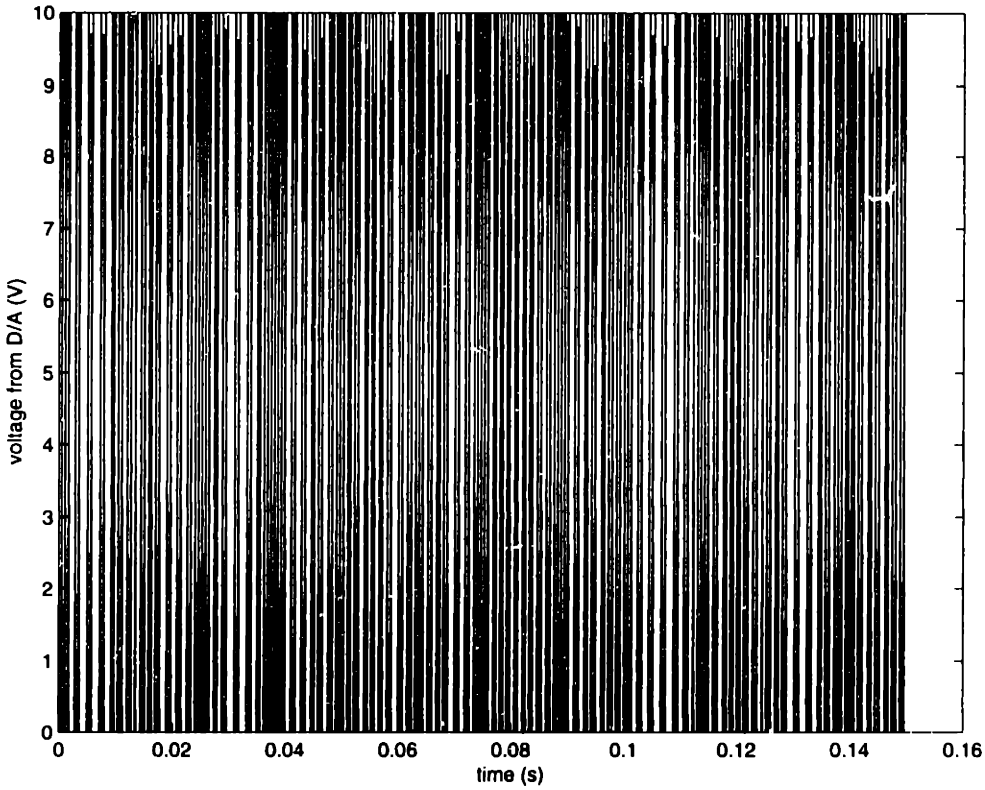

Figure 5-3: Voltage output from channel one of the D/A using the discontinuous control law for the case of only mass uncertainty.

5.4.1 Digital implementation

The sliding controller is implemented digitally on the computer, with a sampling rate of 5 kHz as in the previous chapters. Unlike the linear controller and the feedback linearizing controller where we used the position signal as the only feedback element, sliding control requires all the states to be available. We can see this in (5.14), for example, where the estimate of the required control effort is a linear sum of state errors which implies that we first need to know the state in order to calculate the error. Just like in state feedback methodologies used in linear control, we could design an observer to provide an estimate of the system states. Use of a Luenberger observer for a non-linear system will degrade performance since these observers are model-based and the whole purpose of designing a robust controller is to compensate for parametric uncertainties and/or parameter changes neither of which will be taken into account by this type of observer. Ideally a non-linear observer is required for a non-linear

system and Hedrick et al. [8] provide a good summary of some non-linear observer techniques. The paper by Walcott and Zac [2] provides another approach to non-linear estimation in the case of imperfect knowledge of the plant non-linearities. The dual of a sliding controller is a sliding observer where similar parameter bounds can be incorporated into the observer. However for the purpose of these experiments we use the backward difference approach. This suffices since only first order derivatives need to be estimated; if higher order derivatives were required more accurate methods may need to be used. Experimentation showed that, as expected, this approach led to noise which limited how much we could increase the system bandwidth, but the results clearly show the system robustness. We further note that decreasing sampling time, T, has a detrimental effect on the amount of noise generated by the backward differentiation.

Backward differentiation is a linear system

$$v(k+1) = \frac{x(k+1) - x(k)}{T} \tag{5.45}$$

whose transfer function is given by

$$\frac{V(z)}{X(z)} = \frac{z-1}{Tz} \tag{5.46}$$

where V is the z-transform of the velocity and X is that of position. From Figure 5-4, which shows the bode diagram of the above transfer function for sampling rates of 5 kHz and 1 kHz, we see that the higher the input frequency into the differentiator the more noise we are introducing into the system.

In deciding on a suitable value for λ the trial bounds given in the previous section limit the allowable values as follows:

- $\lambda \leq \frac{2\pi}{3} 270\,\text{Hz} = 565\,\text{rad/s}$
- $\lambda \leq \frac{1}{3 \times 30\,\mu s} = 1.1 \times 10^4\,\text{rad/s}$
- $\lambda \leq \frac{1}{5} 5000\,\text{Hz} = 1000\,\text{rad/s}$

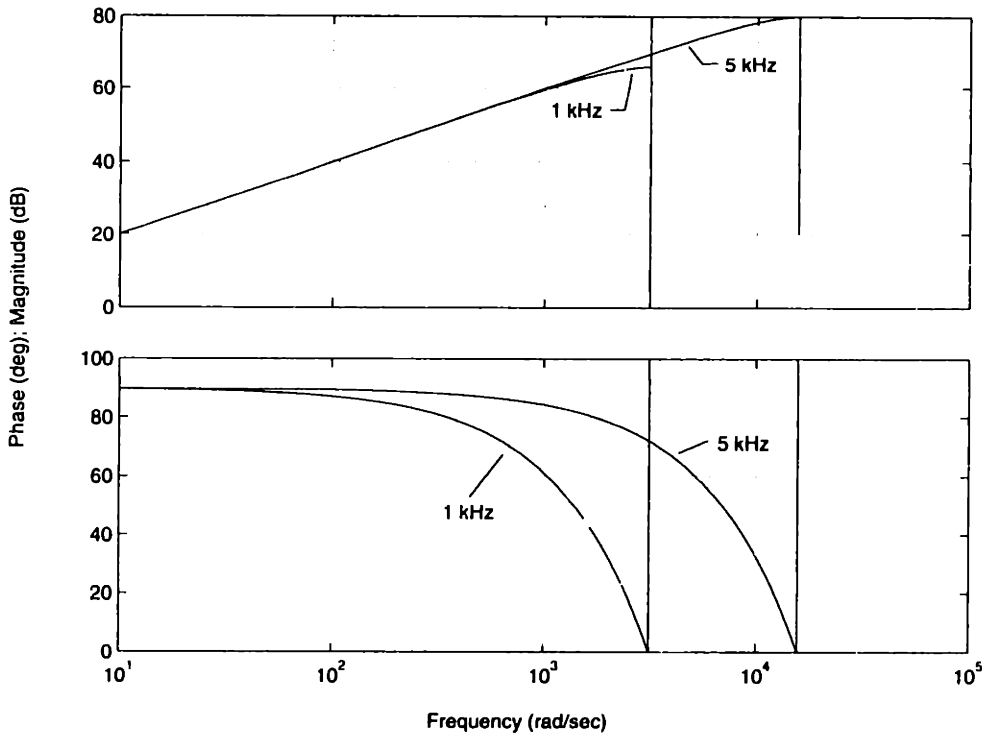

Figure 5-4: Bode plot of backward difference for different sampling rates.

According to these bounds, the choice of λ is limited by structural resonances, however, experiments show that it is possible to use higher values of λ which lead to improved performance because of the higher bandwidth.

When direct backward difference is used the equivalent bandwidth of the controller, λ, is found to be limited to 600 rad/s. To get better performance a slightly different interpolator is used which has the effect of better smoothing the estimated velocity signal. Rather than estimating the instantaneous velocity using (5.45) we can estimate it by finding the mean of the current velocity and the velocity at the previous sampling instant:

$$v_{avg}(k) = \frac{v(k) - v(k-1)}{2} = \frac{x(k) - x(k-2)}{2T} \tag{5.47}$$

With this choice of interpolator we can push the bandwidth to a little over 750 rad/s which is better, but there is still ample margin for improvement. We see from Figure 5-5, which depicts the bode plot of (5.47), that the high frequency noise is

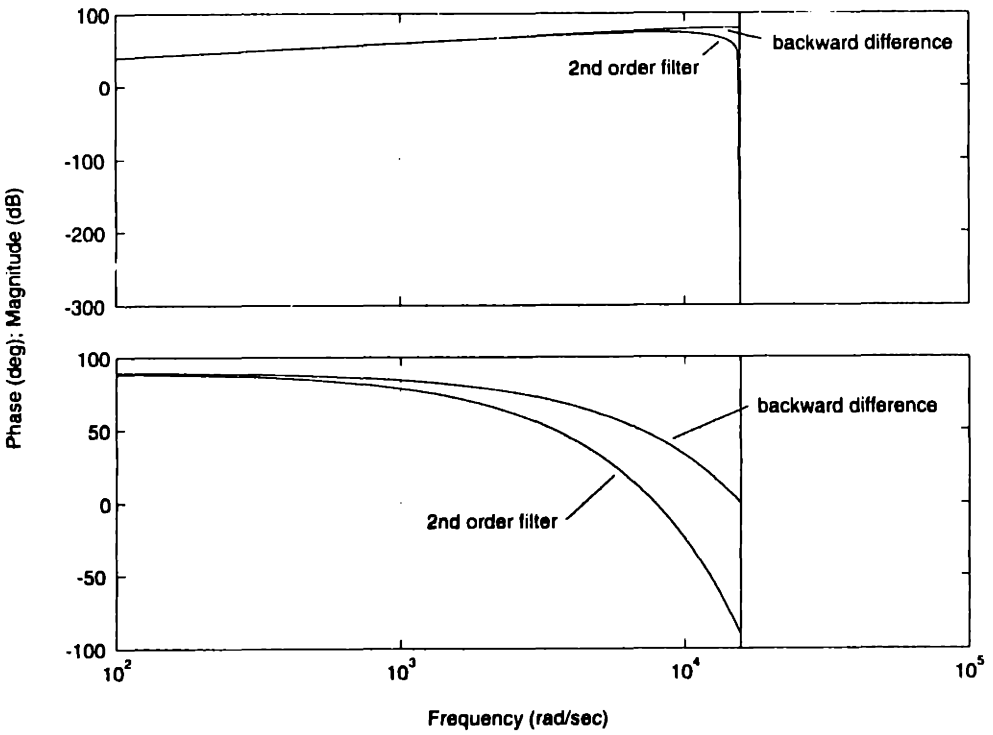

Figure 5-5: Bode plots of filters (5.46) and (5.47).

smoothed more.

The sliding controller also requires the boundary layer to be updated in real time. From the balance conditions (5.22) and (5.23) the boundary layer is a first order dynamic system which requires computation of $\dot{\Phi}$. For this we have no trouble in using backward difference since the boundary layer exists numerically only inside the computer. Equations (5.49) - (5.52) are how the balance conditions are implemented in the Interrupt Service Routine on the DSP board, where we have estimated $\dot{\Phi}$ as

$$\dot{\Phi}(k) = \frac{\Phi(k) - \Phi(k-1)}{T} \tag{5.48}$$

For $k(\mathbf{x}_d) \geq \frac{\lambda \Phi(k-1)}{\beta}$

$$\Phi(k) = \frac{T\beta\, k(\mathbf{x}_d) + \Phi(k-1)}{1 + \lambda T} \tag{5.49}$$

$$\bar{k}(\mathbf{x}) = k(\mathbf{x}) - \frac{\Phi(k) - \Phi(k-1)}{T\beta} \tag{5.50}$$

For $k(\mathbf{x}_d) \leq \frac{\lambda \Phi(k-1)}{\beta}$

$$\Phi(k) = \frac{T\beta\, k(\mathbf{x}_d) + \beta^2\, \Phi(k-1)}{\beta^2 + \lambda T} \tag{5.51}$$

$$\bar{k}(\mathbf{x}) = k(\mathbf{x}) - \frac{\beta[\Phi(k) - \Phi(k-1)]}{T} \tag{5.52}$$

Interrupt Service Routine

The following is the algorithm implemented in the ISR:

1. Initiate conversion of A/D.
2. Read data from A/D and reset status bits.
3. Convert A/D counts to displacement of target.
4. Calculate desired states x_d, $\dot{x}_d$, $\ddot{x}_d$.
5. Calculate best estimate of effort (5.40).
6. Use digital Balance Conditions, (5.49)-(5.52), to find Φ and $\bar{k}$.
7. Use (5.19) to determine required effort.
8. Pass effort, u, through linearizing transformations, (5.27).
9. Output effort through D/A.
10. Update parameters.

5.4.2 Experimental results

The desired trajectory applied to the system is a 40 Hz sinusoid with a 50 μm amplitude generated internally by the DSP board; its derivatives are programmed in the ISR as well. Figure 5-6 shows the experimental results of applying the sliding control with mass uncertainty to the magnetic bearing, overlayed with the simulated response. We note the good correspondence between the two responses. Figure 5-7 shows the experimental time history of the sliding variable. We note that the sliding

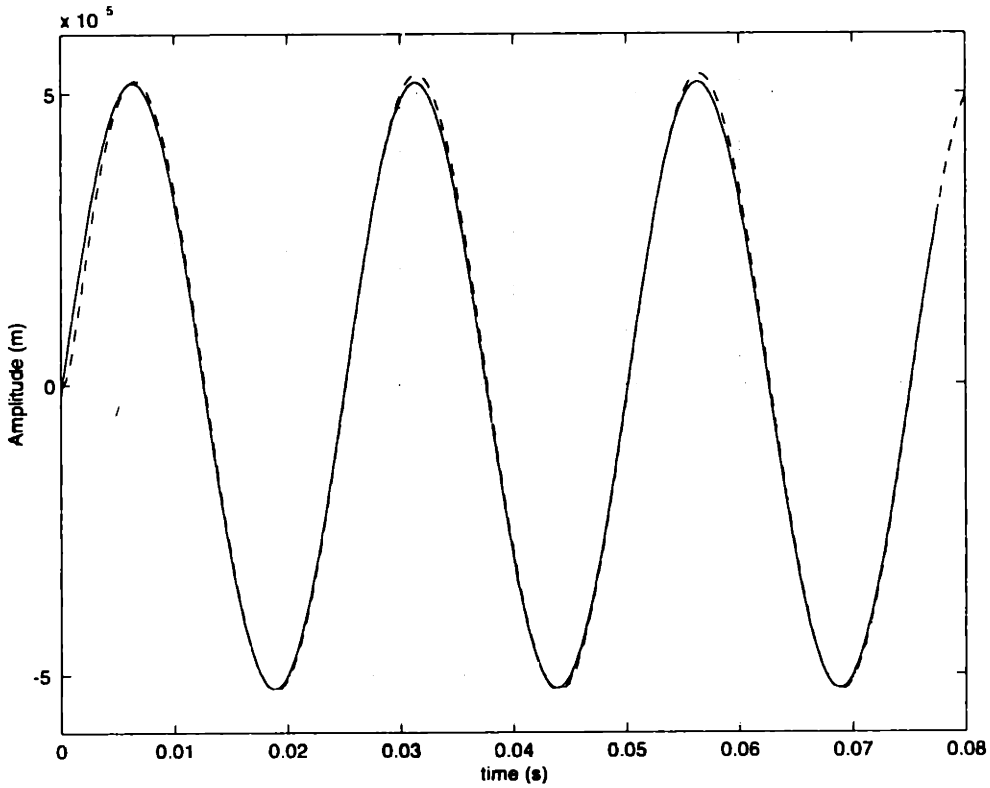

Figure 5-6: Experimental and simulated response of applying sliding control to magnetic bearing ($\eta = 10$ $\lambda = 600$)

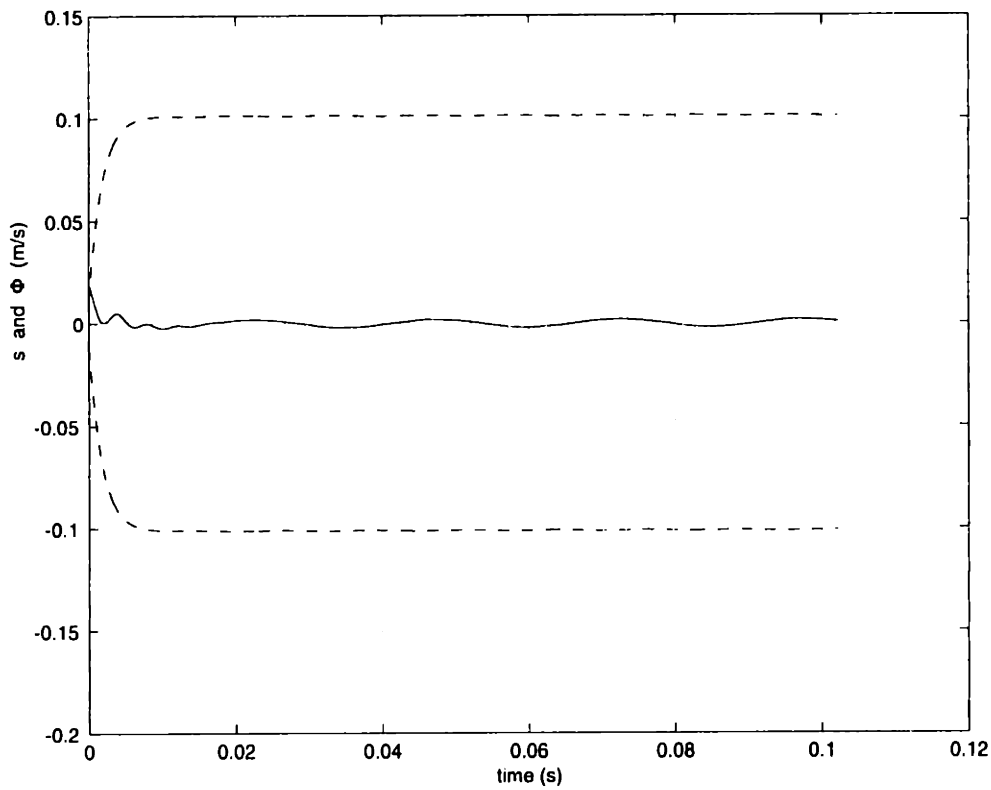

Figure 5-7: Experimental time history of the sliding variable (solid) and the boundary layer (dashed)

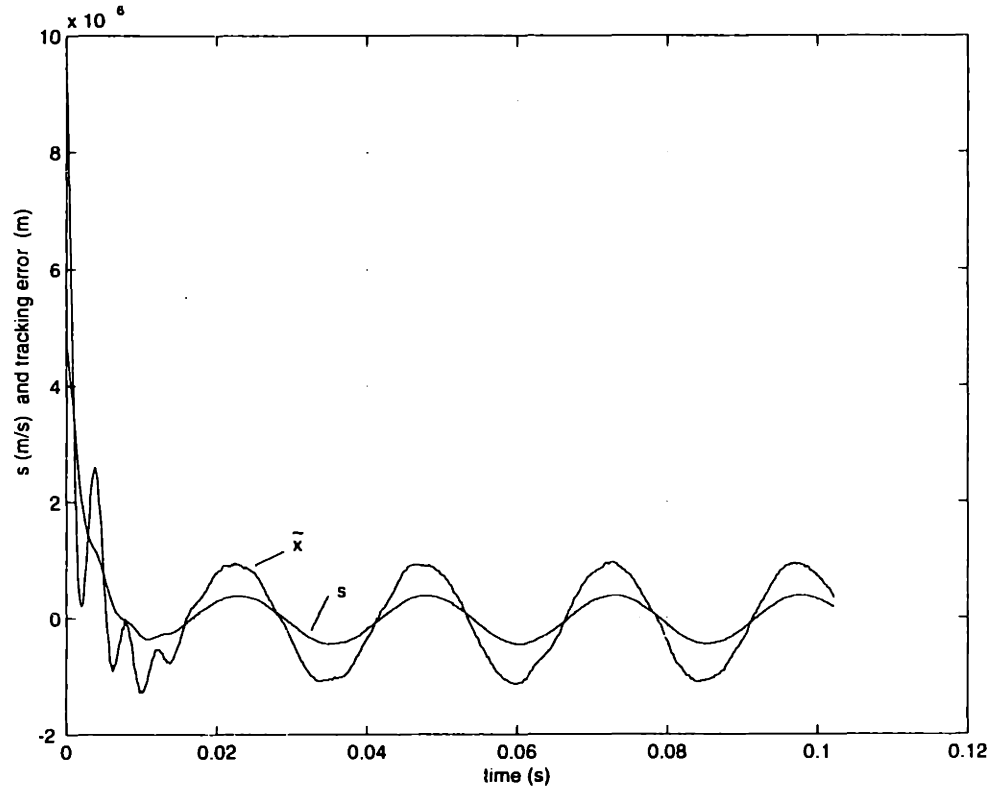

Figure 5-8: Tracking error, $\tilde{x}$, overlayed with the sliding variable, s, scaled by 200.

variable is always inside the boundary layer which is a result of the boundary layer being an attractive invariant set. Figure 5-8 shows the plot of the trajectory error plotted against the sliding variable (scaled by 200 so that it is of comparable magnitude to the error) for the experiment. What are the effects of η and λ ? Having ascertained that simulations and experimental results agree very closely, we can vary the parameters in the simulation to see the resulting effects. In particular, varying η leads to the following behaviour:

- Increasing η ensures faster convergence to the boundary layer if the sliding variable is outside.
- The biggest effect of η is on the boundary layer : the larger η the larger the boundary layer.
- Equation (5.44) shows that k depends additively on $\beta\eta$.
- The control effort during the initial transient increases with η because the power

of attraction of the boundary layer increases.

- η has less effect on $\tilde{x}$ than does λ.

Similarly the effects of increasing λ are :

- Smaller boundary layer. This is easy to see from the Balance Conditions since λ enters directly and multiplicatively in the denominator of the equation for the boundary layer.

- More control effort. λ is interpreted as the equivalent bandwidth of the controller, thus if we increase the bandwidth we expect an increase in control effort. The simulations shown in Figure 5-11, however, showed that the control effort does not change very much.

- $\tilde{x}$ decreases more than due to a corresponding change in η.

- If the values of λ decreases too much the controller will be unable to keep up with the reference trajectory and this leads to instability.

Based on the observations made above, we show the effect of λ by simulating the response of the sliding controller with $\lambda = 1000$; the results are shown in Figures 5-9 through 5-11.

We test robustness of the feedback system by adding weights on the target to alter the mass and run the system with the same 40 Hz desired trajectory. Figure 5-12 shows the experimental data from having added 100 g and 400 g weights. We can see how the influence of the added mass is mitigated by the controller so that the performance of the system is virtually unchanged even with an added mass of 35% of the nominal value. The change of amplitude with the 400 g added mass is less than 2 μm.

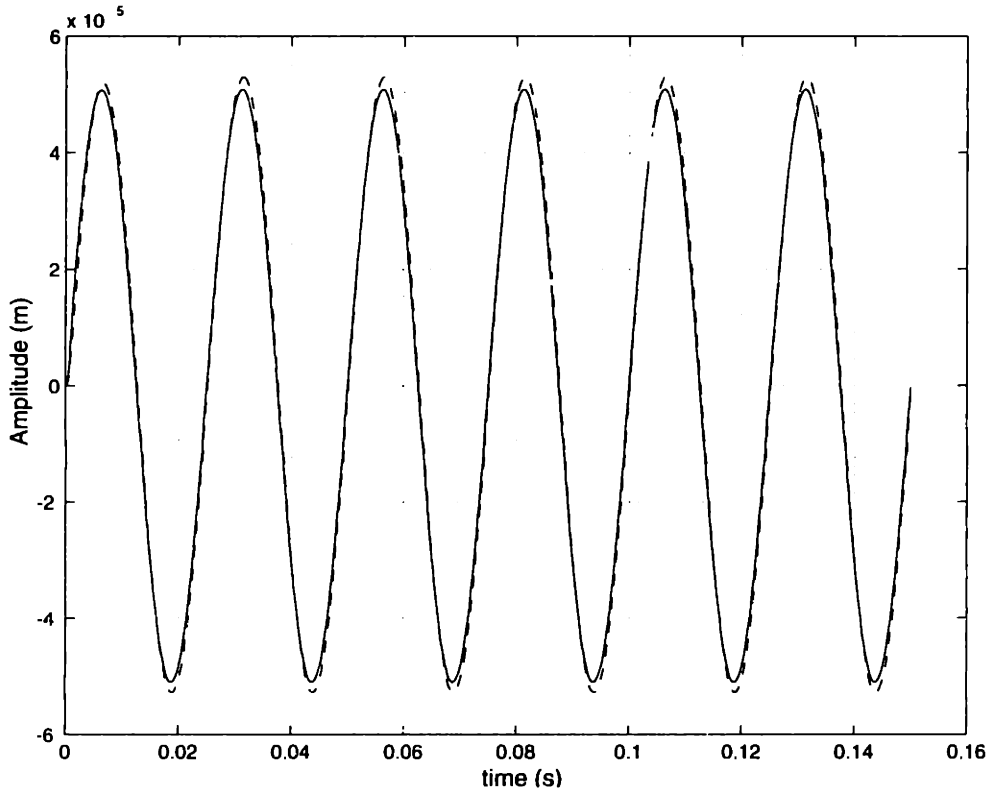

Figure 5-9: Simulated plots of applying sliding control with $\lambda = 1000$ (solid) and $\lambda = 600$ (dashed).

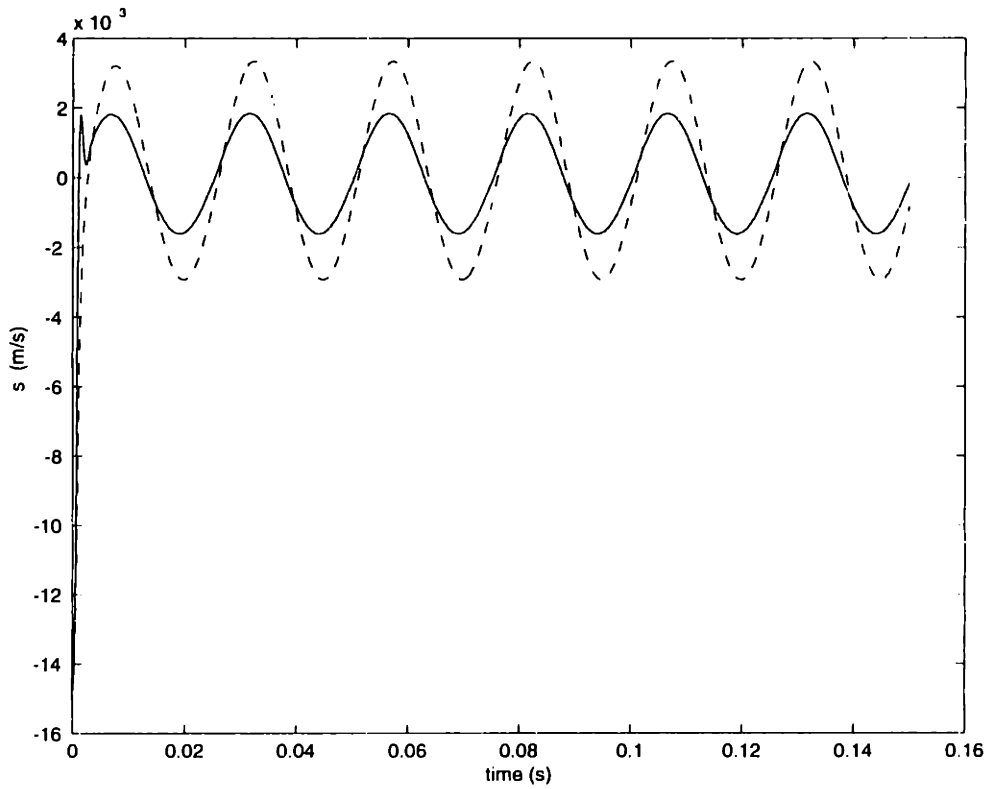

Figure 5-10: Sliding variable with $\lambda = 1000$ (solid) and $\lambda = 600$ (dashed).

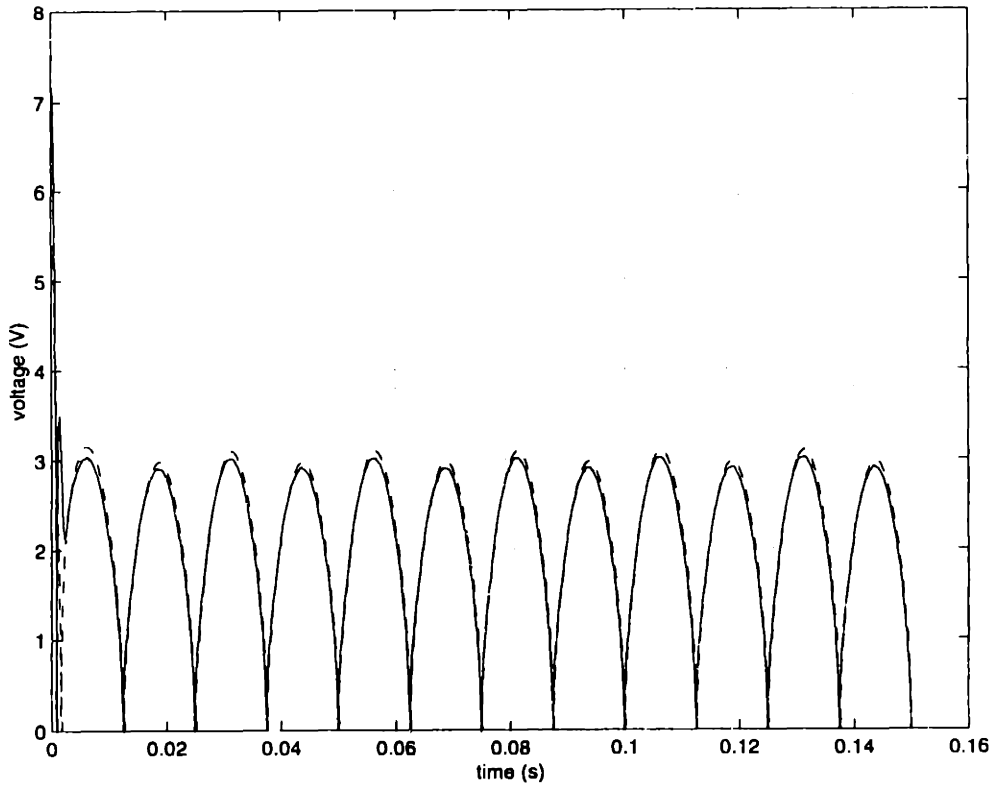

Figure 5-11: Control effort with $\lambda = 1000$ (solid) and $\lambda = 600$ (dashed)

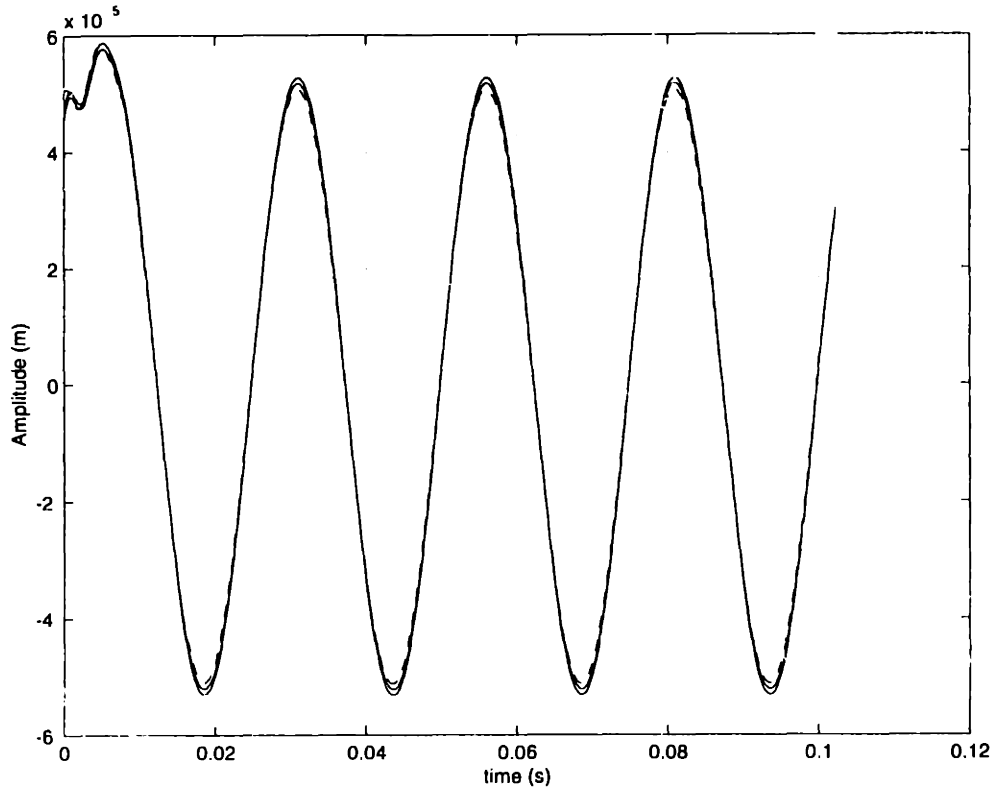

Figure 5-12: Experimental plots with added masses of : no added mass (dashed), 100 g, 400 g (solid).

5.5 Sliding control with mass and actuator uncertainty

In the previous section we have seen how mass uncertainty can be taken into account by incorporating it in the controller. In this section we expand our treatment of uncertainties to include the uncertainty of the actuator constants as well. We can do this is two ways, one will be essentially identical to the development presented in the previous section and the other will show another approach to the problem. The values used in this section are

$$1.035 \leq m \leq 1.265 \text{ kg}$$

$$2.43 \times 10^{-6} \leq C \leq 2.97 \times 10^{-6} \frac{\text{N m}^2}{\text{A}^2}$$

both of which represent a $\pm 10\%$ deviation from the nominal values.

5.5.1 Mass and actuator constants included in the linearizing transformations

The difference in how we chose to incorporate actuator uncertainty into the system comes from how we chose the linearizing transformations. If we chose the transformations such that they include our estimates of the mass and actuator constants, $\hat{m}$ and $\hat{C}$, then we can rewrite the transformations as

$$u \geq 0 \quad \Rightarrow \quad i_2 = (g_0 + x_0 - x)\left(\frac{\hat{m}u}{\hat{C}}\right)^{0.5} \quad \text{and} \quad i_1 = 0 \tag{5.53}$$

$$u \leq 0 \quad \Rightarrow \quad i_1 = (g_0 + x_0 + x)\left(\frac{-\hat{m}u}{\hat{C}}\right)^{0.5} \quad \text{and} \quad i_2 = 0 \tag{5.54}$$

which are the same as (4.9) except that the estimates of the parameters are used. In the presence of parametric uncertainties feedback linearization does not guarantee

good performance nor stability. With these transformations, the plant reduces to

$$\ddot{x} = bu + \frac{d}{m}$$

where the new coefficient of the effort is now

$$b = \frac{C\,\hat{m}}{\hat{C}\,m}$$

As before, taking the estimate $\hat{b}$ to be the geometric mean of b_{min} and b_{max} the estimate will be

$$\hat{b} = 1 \ .$$

We can now use (5.24) to find the new expression for β

$$\beta = \left(\frac{C_{max}\,m_{max}}{C_{min}\,m_{min}}\right)^{0.5} \ . \tag{5.55}$$

Having characterized the uncertainties and noted that the plant equation has the same general form as that used in the preceding section, it follows that the development of the sliding controller is completely analogous. Without repeating the previous derivation we will simply state the results.

$$k \geq (\beta - 1)|-\ddot{x} + \lambda\dot{\tilde{x}}| + \frac{\hat{C}D}{C_{min}\hat{m}} + \left(1 - \frac{\hat{C}}{C_{max}}\right)\frac{\hat{d}}{\hat{m}} + \eta\frac{\hat{C}m_{max}}{C_{min}\hat{m}} \tag{5.56}$$

$$\left.\begin{matrix}\Phi \\ \bar{k}\end{matrix}\right\} \text{ from Balance Conditions (5.22), (5.23)} \tag{5.57}$$

$$\hat{u} = \ddot{x} - \lambda\dot{\tilde{x}} - \frac{\hat{d}}{\hat{m}} \tag{5.58}$$

$$u = \hat{b}^{-1}\left(\hat{u} - \bar{k}\,sat\left(\frac{s}{\Phi}\right)\right) \tag{5.59}$$

In implementing this new controller digitally, the steps performed in the ISR remain the same as those presented previously in this chapter. When applied to the magnetic bearing this new controller produces an output which was similar to that produced

by just using the mass uncertainty. The trajectory is shown in Figure 5-13.

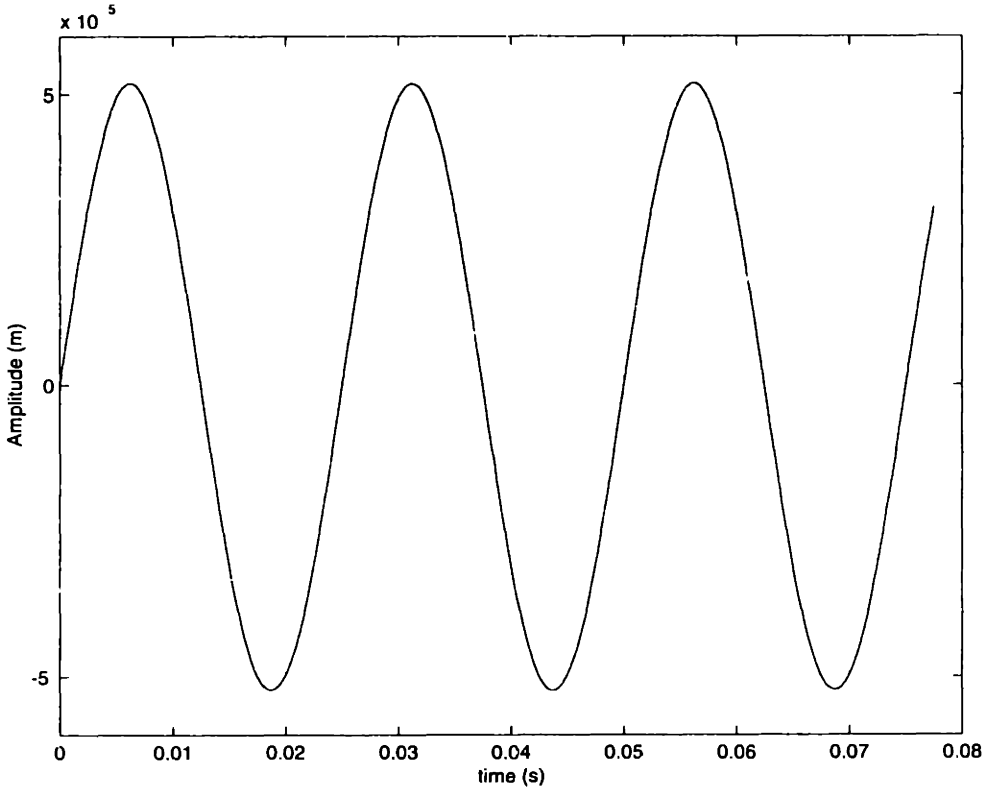

Figure 5-13: Experimental trajectory using controller with mass and actuator uncertainties.

Again we test the robustness of the controller to mass uncertainty and the results plotted in Figure 5-14 show a marked improvement over those obtained using only mass uncertainty. This was expected because when accounting for only mass uncertainty we did not take into account the difference in actuator constants between the two electromagnets.

5.5.2 Mass and actuator constants not included in the linearizing transformations

We have seen how, by choosing the linearizing transformations as (5.53) and (5.54), the procedure of designing the sliding controller is identical to that illustrated with the mass uncertainty only. In this section we develop a sliding controller for the case of mass and actuator uncertainty by using a slightly different approach. The linearizing

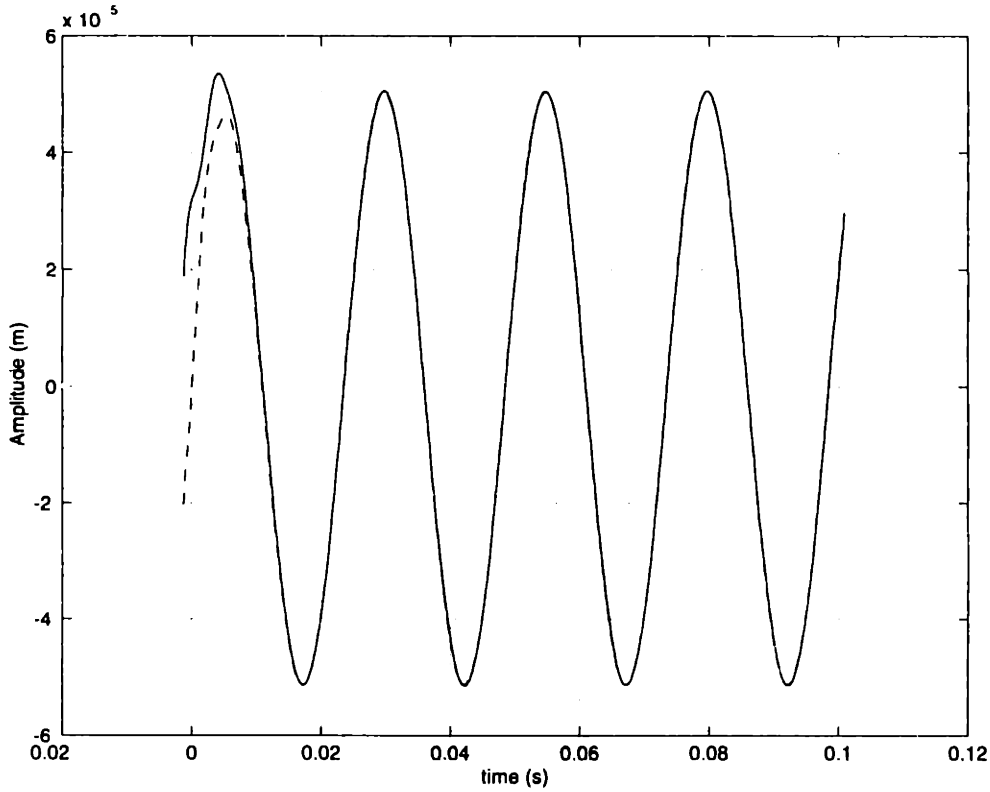

Figure 5-14: Trajectory with (solid) and without (dashed) an added mass of 100 g.

transformations are written without the mass and actuator constant estimates, just as we did in (5.27). Using these transformations the plant will be the same as (5.29).

For notational simplicity, we will multiply (5.12) on both sides by the mass and use it as the new sliding condition. Following the procedure leading to (5.40), we get that

$$\hat{u} = \frac{1}{\hat{C}}\hat{m}\ddot{x}_d - \hat{m}\lambda\dot{\tilde{x}} - \hat{d} \tag{5.60}$$

For the switching controller we apply (5.15) directly and this will be the effort required by the switching controller.

Proceeding as before we differentiate the sliding condition with respect to time and insert (5.15) and (5.60) in it to get the inequality from which we calculate the gain, k. After simplifying the resulting inequality we get that

$$k \geq \frac{m\eta}{C} + sgn\left[\left(\frac{\hat{m}}{\hat{C}} - \frac{m}{C}\right)\ddot{x}_d + \left(\frac{m}{C} - \frac{\hat{m}}{\hat{C}}\right)\lambda\dot{\tilde{x}} + \left(\frac{d}{C} - \frac{\hat{d}}{\hat{C}}\right)\right] \quad .$$

We see that the above expression still includes the true actuator constant, C, the true mass, m, and the true disturbance, d, all of which are parameters which we do not know accurately. We can thus make the following substitutions which will allow us to take the mass and disturbance bounds out of the *sgn* function.

$$d = (d - \hat{d}) + \hat{d}$$
$$m = (m - \hat{m}) + \hat{m}$$

Substituting for the above and using (5.34) and (5.35) to simplify the inequality, the desired expression for k in terms of known parameters is

$$k = \frac{m_{max}\eta}{C_{min}} + \left(\frac{1}{C_{min}} - \frac{1}{\hat{C}}\right)\left|-\hat{m}\ddot{x}_d + \hat{m}\lambda\dot{\tilde{x}} + \hat{d}\right| + \frac{M}{C_{min}}\left|\ddot{x}_d + \lambda\dot{\tilde{x}}\right| + \frac{D}{C_{min}} \qquad (5.61)$$

We have finished designing the switching controller, and must now smooth it out with the boundary layer approach. The effort will still be given by (5.19) but the uncertainty propagates in a way which requires us to change the balance conditions.

We can show that the following conditions will satisfy (5.21).

$$\bar{k} = \begin{cases} k - \frac{m_{min}}{C_{max}}\dot{\Phi} & \dot{\Phi} \geq 0 \\ k - \frac{m_{max}}{C_{min}}\dot{\Phi} & \dot{\Phi} \leq 0 \end{cases} \qquad (5.62)$$

In order to find the governing equations for the boundary layer, we look at when the trajectory is inside the boundary layer in which case $sat\left(\frac{s}{\Phi}\right) = \frac{s}{\Phi}$. We again make the analogy with a first order low-pass filter in s and if we set the bandwidth of this filter as λ then we derive

$$\frac{C\bar{k}}{m\Phi} = \lambda \quad \Rightarrow \quad \frac{\hat{C}\bar{k}_d}{\hat{m}\Phi} \approx \lambda$$

which we rearrange to give

$$\bar{k}(\mathbf{x}_d) \approx \frac{\lambda\Phi\hat{m}}{\hat{C}} \; . \qquad (5.63)$$

Where we have used the approximation that within the boundary layer $k \approx k(\mathbf{x}_d)$.

Inserting this into (5.62) we also conclude that $\bar{k} \approx \bar{k}(\mathbf{x}_d)$. Placing the latter approximations and (5.63) into (5.62) the new equations for the boundary layer are :

For $\dot{\Phi} \geq 0$

$$\dot{\Phi} + \frac{C_{max}\hat{m}}{\hat{C}m_{min}}\lambda\Phi = \frac{C_{max}}{m_{min}}k(\mathbf{x}_d) \tag{5.64}$$

For $\dot{\Phi} \leq 0$

$$\dot{\Phi} + \frac{C_{min}\hat{m}}{\hat{C}m_{max}}\lambda\Phi = \frac{C_{min}}{m_{max}}k(\mathbf{x}_d) \tag{5.65}$$

In conclusion, the sliding controller designed in this section is represented by Equations (5.60), (5.61), (5.62), (5.64), and (5.65). These equations were implemented in the computer as explained previously. Figure 5-15 shows the experimental results for the controller just derived when the desired trajectory is, again, a 40 Hz sinusoid of amplitude 50 μm.

As we did before we add an unmodeled 100 g mass to the target, the trajectory is shown in Figure 5-16. We see that performance is very similar to that for the case with mass and actuator constant uncertainty presented in the previous section. In both cases the results are better than those obtained with the controller designed only for mass uncertainty, and are much better than the results obtained with the linear or feedback-linearized controllers. For the latter two controllers, parameter variation was not accounted for and thus any deviation from the nominal values for which they were designed, lead to substantial degradation in performance.

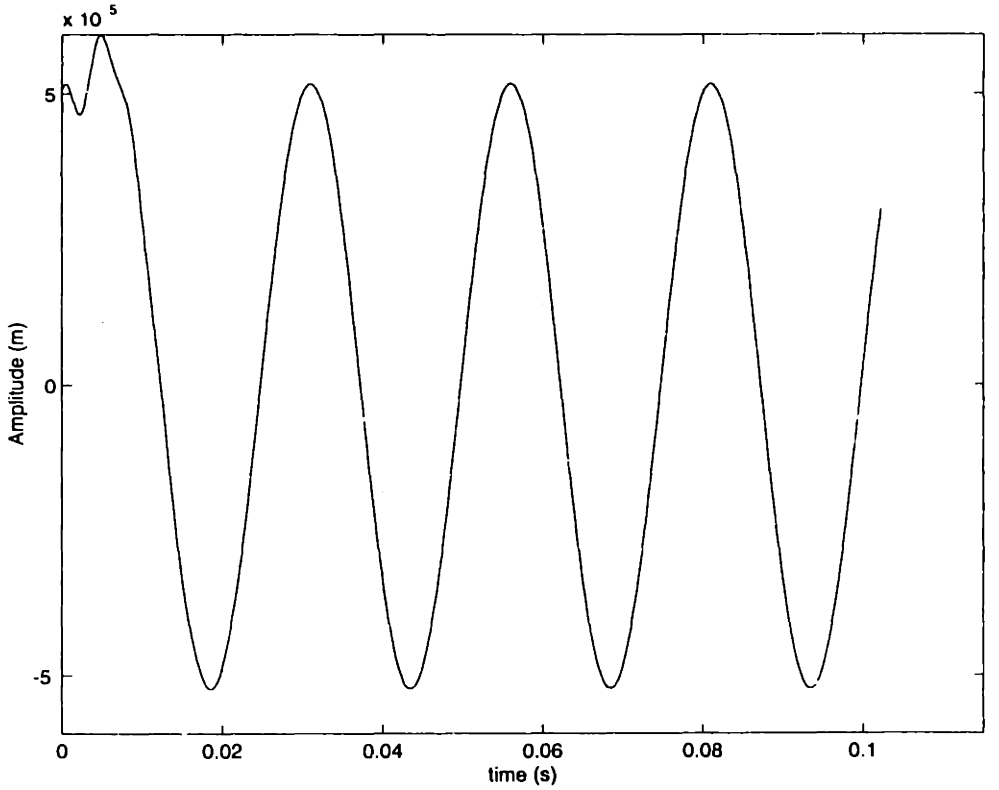

Figure 5-15: Experimental plots controller with mass and actuator uncertainties

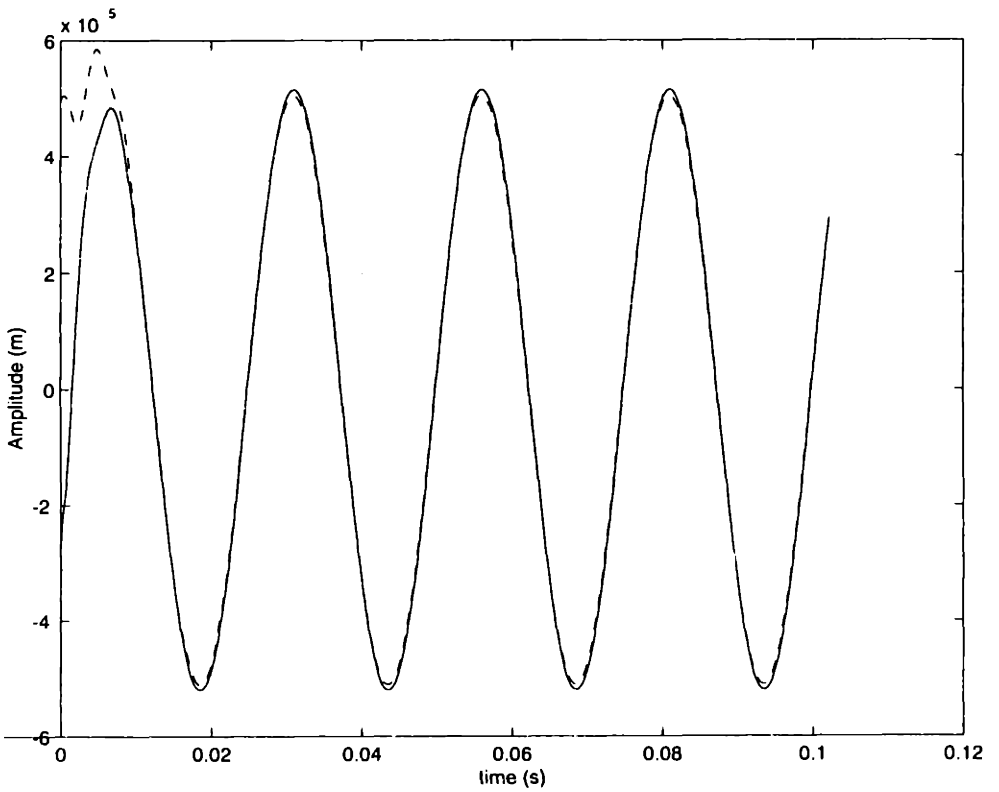

Figure 5-16: Experimental trajectory with (solid) and without (dashed) 100 g mass added

Chapter 6

Conclusions and Future Work for the Single Degree of Freedom Magnetic Bearing

The goal of this thesis is the control of a single degree of freedom magnetic bearing. Proceeding in order of increasing complexity we start by designing several linear lead-lag controllers; we then retain one of these controllers while adding to the feedback system an inner loop which linearizes the model. Having tested these techniques we design and implement sliding control.

The data presented demonstrates the superior performance of sliding control when applied to this test setup, over application of the other controllers. The linear compensators provide good performance for small excursions of the target from the nominal operating point, but will become unstable at large displacements. We overcome this problem by using feedback linearization. The target can now move the full range of motion of the bearing and still be stable. However, model imperfections prevent us from achieving perfect cancelation of the nonlinearities: the loop transmission of the plant shows that our model predicts the magnitudes very well, but is unable to explain phase roll-off above 90 Hz. We venture to postulate that this is the result of neglected time delays, but even by including these in our model we are not able to explain the behaviour.

Both of the techniques implemented above suffer from lack of robustness to parametric uncertainties/variations. We find sliding control to be invaluable in making the closed loop system virtually insensitive to parametric variations having known bounds. Of the methods presented here sliding control is to be preferred where large target displacements are required and/or in the presence of parametric uncertainties.

As future work we propose first of all to investigate the unmodeled phase roll off. Also, for future experiments on the existing setup, we recommend stiffening the arm holding the target; raising the resonant frequencies of the vertical vibrations will enable us to design controllers having higher bandwidths.

The purpose of having a long arm is to be able to approximate rotational motion as linear motion. As a way to reduce the size of the assembly, we suggest replacing the arm and the notch hinges with leaf spring flexures mounted below the target. For small displacements this type of flexures provide quasi-linear motion.

Looking at the "bigger picture" i.e., beyond this setup and in industry, the work done in this document provides a core analysis substantiated by simulations and experimentation, which can now be tailored to a specific application. In particular, a survey of previous work in this area revealed that vibration isolation is often an engineering solution that is not properly included in early conceptual designs of machinery. This provides ample opportunities for integration of magnetic bearing technology. A natural extension of the work done would be to increase the number of controlled degrees of freedom to actively mitigate cross-axis vibrations and/or rotational motion of the machine or subassembly. This leads to active vibration isolation of the payload from vibrations produced by other subsystems, and to isolation from ground vibrations in ultraprecision machines. This is the trend in the photolithography industry where magnetic bearings are being designed into the base to damp out vibrations entering through the ground and reaction forces from the moving stage.

Reduction of critical dimensions as a result of industry's trend towards miniaturization poses increasingly stringent specifications and requirements on machines whose dynamics is not known, or exactly modeled. Researching new technologies, manufacturing solutions, etc. are radical ways to circumvent the obstacle when it

may be possible to adjust the control loop to get desired performance. Robust control methods are particularly attractive in this regard, and we suggest studying techniques such as H_2-control, H_∞-control and LTR on test-beds as general as the one used in this thesis.

Chapter 7

Design of a controller for a single degree of freedom wafer stepper using the classical approach to design

7.1 Problem specification

We stated in the first chapter that the target application for this projet was the control of a photolithography stage, although the hardware used can serve as a testbed for a wide variety of mechanical systems. In this chapter and the next we focus our attention on the lithographic process in particular and introduce two design approaches for designing a control system for a photolithography stage subject to specified constraints on accuracy, movement and sensors.

The system which will be studied is shown in Figure 7-1 and is a version of the stage presented in [10]. The problem statement which follows has been adapted from a take home design project for the Digital Control course at MIT, taught by Prof. Trumper. The model is one-dimensional in the sense that we model all of the masses as able to move in only the x-direction. Springs and dampers connected between

elements are sensitive to the relative x-motion of their ends, and forces are applied along the x-axis.

The wafer is carried on the mirror M_4, which is mounted to the fine stage M_3 via the spring k_4 and damper b_4. These elements model the relatively stiff kinematic coupling which connects the mirror to the fine stage. The fine stage moves on relatively soft flexures k_3 and b_3 with respect to the coarse stage M_2 as driven by the input force F_2. In the real system this force is supplied by a voice coil actuator, which is not modeled in any further detail herein. The coarse stage rides on plain bearings which will be assumed to have Coulomb friction acting on them. The coarse stage is driven by force F_1 with respect to the machine base M_1. In the real system, this force is supplied by a rotary motor through a rack-and-pinion drive, but this mechanism is not modeled here. The machine base is supported on spring k_1 and damper b_1. These model the vibration isolation legs which are present on the real stepper. In order to isolate well, these elements are made to be compliant, however, this introduces the problem that the base is driven by the stage input forces.

The two system outputs are y_1, which is the distance between the coarse and fine stages, and y_2 which is the distance between the machine base and the mirror. Output y_1 is measured with a linear differential variable transformer (LVDT) and y_2 is measured with a laser interferometer. Since the system lens is modeled as rigidly attached to the machine base, and the wafer is carried on the mirror, the output y_2 is the critical output for control purposes. The details of the lithographic process require that the errors in y_2 from its setpoint be less than 10 nanometers RMS in a 500 msec exposure interval after each step; this is the design specification.

The main task of the stage system is to take repeated 20 mm steps while maintaining an exposure window of 500 msec during which time the RMS position error specification is maintained. Under the constraint described below, a controller must be designed which meets the specifications.

A note on numerics: This system is 8 th order, and with the controllers the order will increase. There are also widely spaced singularities (i.e., it's "stiff"), so attention must be paid to the numerics of any simulations. As a general rule in all systems,

especially those which are stiff, it is important to carry out reasonableness tests on any of the data to be sure that the results are not simply numerical artifacts.

The system model is linear and the plain bearings experience a friction force $F_{friction} = 10\, sgn(q_4)$ which opposes the control force F_1. Nonlinearities arise in that force actuator saturation occurs, specifically F_1 is limited to $\pm F_{1max}$ and that F_2 is limited to $\pm F_{2max}$ where $F_{1max} = 200$ N and $F_{2max} = 50$ N. A further requirement of the system performance is that the fine stage relative excursion y_1 must be maintained at less than ± 200 μm at all times. Furthermore, sensors exhibit quantization and the entire stage is subject to floor vibrations described as a set of two sinusoids as follows:

$$\ddot{x}_0(t) = e_1 \cos \omega_1 t + e_2 \cos \omega_2 t \tag{7.1}$$

Where the parameters $e_1 = 3 \times 10^{-7}$m/sec^2 and $e_2 = 1 \times 10^{-4}$m/sec^2 are constant amplitudes while $\omega_1 = 1$ rad/sec and $\omega_2 = 20$ rad/sec are the frequencies of the main two components of ground vibrations. In a real setting, a model for ground vibrations can be obtained by sampling the output of an accelerometer placed on the ground next to the stage and taking the Fourier Series of the signal. The two components shown above are assumed to be the greatest contributors to the vibrations. For the quantization, the LVDT output y_1 is quantized at the 25 nm level (this represents approximately a 14-bit A/D converter on this channel) and the interferometer output y_2 is quantized at the 2.5 nm level (this is a typical resolution for a laser interferometer system).

The main task of the stage is to take repeated ± 20 mm steps and between each step there must be a 500 msec exposure window during which time the error in y_2 is held to less than 10 nm RMS. The controller has direct access to the measurements y_1 and y_2 but cannot use any other system outputs and the controller output will drive the two input forces F_1 and F_2. While the model will be perfectly known in simulations, one cannot take unreasonable advantage of this fact for example by using notch filters to exactly cancel the plant modes, or inverting the plant to obtain exact feedforward blocks. As part of the controller design, we will also generate an appropriately shaped

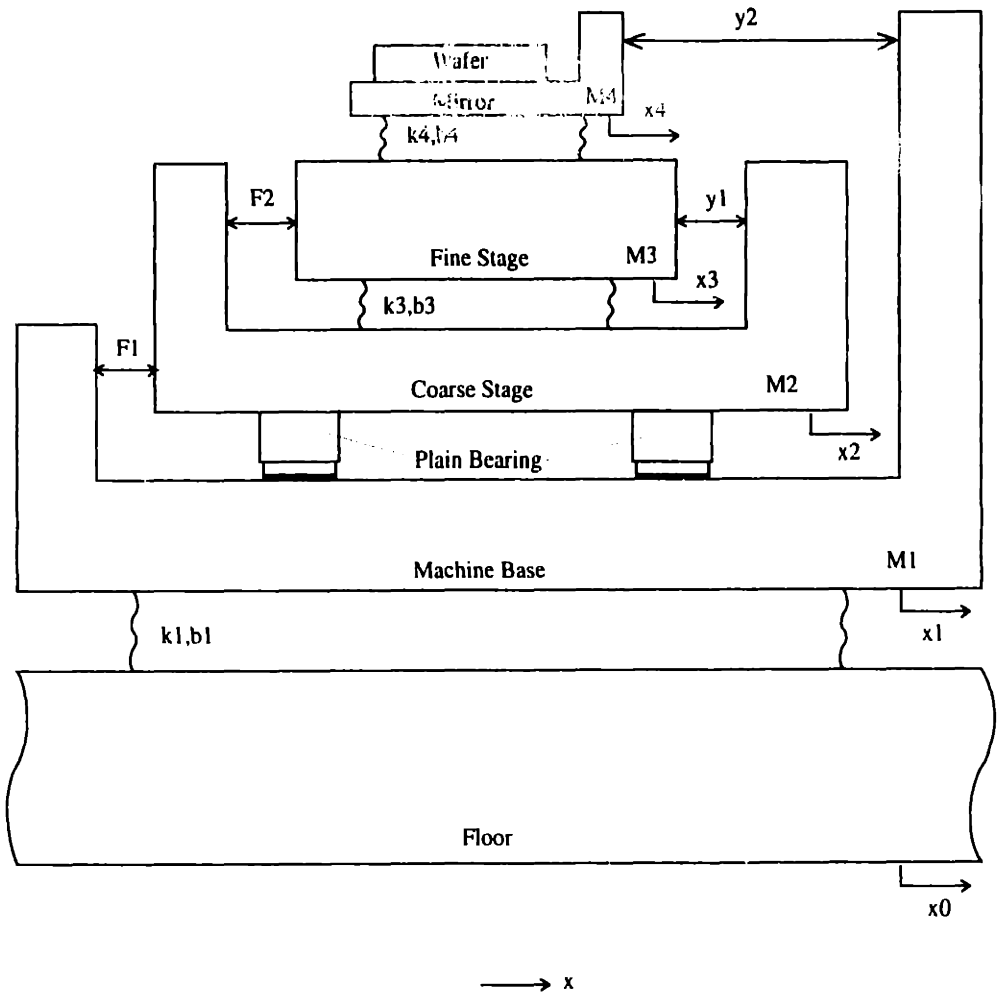

Figure 7-1: Stage system model

reference trajectory which will be necessary to meet specifications.

7.2 RMS error requirement

The RMS specifications are based upon the requirements for the lithographic progress. When the wafer is exposed to the light source, the RMS requirement allows for some vibration in the stage, but limits it to a level that the pattern being impressed is not smeared to the extent that it is un-usable. Overlapping patterns is a major source of concern especially since the trend is towards circuits which have an increasing

component density. For a continuous-time signal $e(t)$, the RMS value is defined as:

$$e_{RMS} = \sqrt{\frac{1}{T}\int_0^T e^2(t)\mathrm{dt}} \tag{7.2}$$

When using Matlab to solve the equations, the system will be solved in discrete fashion and hence values will be quantized. For this reason the integral given above becomes a summation over time.

$$\sqrt{\frac{1}{window}\sum_{i=1}^{window} e_i^2} \tag{7.3}$$

where `window` refers to the number of time intervals in the required time window. For example, if the required window is 0.5 s and the time step used is 1 ms, then `window` = 500. Note that a fixed time interval is used in the calculations.

As a point of reference in understanding RMS, a Gaussian random waveform will have a peak-to-peak value of about 6 times its RMS value (99% of values from such a distribution will fall within this limit), whereas a pure sinusoid has a peak-to-peak value of $2\sqrt{2}$ times its RMS value (The power lines rated at 110 V actually have a peak voltage of $\pm 160\,V$).

7.3 Plant linear model

In developing a model for the system, the first step is to decide upon the state-variables to use. The preferred set of states to be utilized in this problem is one which uses relative displacements between elements. This is a natural choice since all the dynamics are associated with relative displacements and relative velocities. Furthermore, if absolute displacements are used, the base vibration enters the system as x_0, $\dot{x}_0$ and $\ddot{x}_0$ which may cause problems since only $\ddot{x}_0$ is known. The linear dynamics for the system shown in Figure 7-1 may be expressed in standard state space representation as

$$\dot{\mathbf{q}} = \mathbf{A}\,\mathbf{q} + \mathbf{B}\,\mathbf{u} \tag{7.4}$$

$$\mathbf{y} = \mathbf{C}\mathbf{q} + \mathbf{D}\mathbf{u} \tag{7.5}$$

where the eight state variables are

$$\begin{aligned} q_1 &= x_1 - x_0 \\ q_2 &= \dot{q}_1 \\ q_3 &= x_2 - x_1 \\ q_4 &= \dot{q}_3 \\ q_5 &= x_3 - x_2 \\ q_6 &= \dot{q}_5 \\ q_7 &= x_4 - x_3 \\ q_8 &= \dot{q}_7 \end{aligned}$$

The following table lists the parameters for the mechanical elements in the system

	1	2	3	4
Masses (kg)	$m_1 = 500$	$m_2 = 50$	$m_3 = 10$	$m_4 = 5$
Spring Constants	$k_1 = 9000$	$k_2 = 0$	$k_3 = 1500$	$k_4 = 5e6$
Damping Constant	$b_1 = 800$	$b_2 = 0$	$b_3 = 15$	$b_4 = 700$

To derive the state equations, we first use Newton's Laws of motion applied to each element. With this we get

$$m_1\ddot{x}_1 = -k_1 q_1 - b_1 q_2 - F_1 \tag{7.6}$$

$$m_2\ddot{x}_2 = k_3 q_5 + b_3 q_6 + F_1 - F_2 \tag{7.7}$$

$$m_3\ddot{x}_3 = -k_3 q_5 - b_3 q_6 + k_4 q_7 + b_4 q_8 + F_2 \tag{7.8}$$

$$m_4\ddot{x}_4 = -k_4 q_7 - b_4 q_8 \tag{7.9}$$

Now substitute

$$\begin{aligned}
\ddot{x_1} &= \dot{q_2} + \ddot{x_0} \\
\ddot{x_2} &= \dot{q_4} + \ddot{x_1} \\
\ddot{x_3} &= \dot{q_6} + \ddot{x_2} \\
\ddot{x_4} &= \dot{q_8} + \ddot{x_3}
\end{aligned}$$

To get

$$m_1\dot{q_2} = -k_1q_1 - b_1q_2 - F_1 - m_1\ddot{x_0} \tag{7.10}$$

$$m_2\dot{q_4} = k_3q_5 + b_3q_6 + F_1 - F_2 - \frac{m_2}{m_1}[-k_1q_1 - b_1q_2 - F_1] \tag{7.11}$$

$$m_3\dot{q_6} = -k_3q_5 - b_3q_6 + k_4q_7 + b_4q_8 + F_2 - \frac{m_3}{m_2}[k_3q_5 + b_3q_6 + F_1 - F_2] \tag{7.12}$$

$$m_4\dot{q_8} = -k_4 + 4q_7 - b_4q_8 - \frac{m_4}{m_3}[-k_3q_5 - b_3q_6 + k_4q_7 + b_4q_8 + F_2] \tag{7.13}$$

The system outputs are $y = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} -q_5 \\ -q_3 - q_5 - q_7 \end{pmatrix}$

while as the inputs are $u = \begin{pmatrix} F_1 \\ F_2 \\ \ddot{x_0} \end{pmatrix}$

These equations can be written in matrix form as in (7.5) where

$$\mathbf{A} = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{k_1}{m_1} & -\frac{b_1}{m_1} 0 & 0 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{k_1}{m_1} & \frac{b_1}{m_1} & 0 & 0 & \frac{k_3}{m_2} & \frac{b_3}{m_2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 01 & 0 & 0 & \\ 0 & 0 & 0 & 0 & -k_3[\frac{1}{m_3} + \frac{1}{m_2}] & -b_3[\frac{1}{m_3} + \frac{1}{m_2}] & \frac{k_4}{m_3} & \frac{b_4}{m_3} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & \frac{k_3}{m_3} & \frac{b_3}{m_3} & -k_4[\frac{1}{m_4} + \frac{1}{m_3}] & -b_4[\frac{1}{m_4} + \frac{1}{m_3}] \end{pmatrix} \tag{7.14}$$

$$\mathbf{B} = \begin{pmatrix} 0 & 0 & 0 \\ -\frac{1}{m_1} & 0 & -1 \\ 0 & 0 & 0 \\ \frac{1}{m_2} + \frac{1}{m_1} & -\frac{1}{m_2} & 0 \\ 0 & 0 & 0 \\ -\frac{1}{m_2} & \frac{1}{m_2} + \frac{1}{m_3} & 0 \\ 0 & 0 & 0 \\ 0 & -\frac{1}{m_3} & 0 \end{pmatrix} \tag{7.15}$$

$$\mathbf{C} = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & -1 & 0 & -1 & 0 \end{pmatrix} \tag{7.16}$$

$$\mathbf{D} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \tag{7.17}$$

The real plant is represented using the above **C** matrix, but for simulations an augmented **C** matrix must be used which outputs q_4 so that the latter can be used to calculate the friction.

7.4 Controller Design

There are multiple ways in which this design problem can be approached. Before deciding which method to utilize, it is more fundamental to decide which domain the design is going to be carried out in initially. If a digital controller is needed, one can design in continuous time and then use emulation techniques to map the results over to discrete time, or it is possible to design in discrete time from the start. The first stage of the latter will require ZOH transformation of the state equations given above. To show the design procedure involved in controller design, in this problem a continuous-time controller will be derived. This approach was chosen because in controller design the major challenge is to find a controller that is able to perform the desired function without the need to worry about sampling. The transition to discrete time can then be made by methods such as the Tustin approximation, pole-zero mapping or ZOH. Sampling fast enough will ensure that the discrete controller will function very similarly to the continuous-time controller, but practical limitations on the sampling frequency may lead to the need to make some changes to the discrete equivalent.

There are several approaches which can be used to tackle the problem. One way is to use Optimal Control for the stage. Appropriate weight matrices will need to be found, and this will be the longest part of adopting this design approach because unfortunately no straightforward technique is available to help in the selection of weight matrices. Optimal control requires all states to be available so an observer

will need to be implemented; a Kalman filter is the natural choice for such an observer. The task of designing an estimator is simplified by the Separation Principle which states that the controller and estimator can be designed separately and implemented together. This approach will be taken on in the next chapter.

Another design approach, the one which is developed in this chapter, is based on the classical approach to design; i.e. output feedback. While as this approach is not restricted to SISO systems, its ease of use is very much increased if the system at hand is an SISO linear system. However, there are three inputs (two of which are controllable, F_1 and F_2 and one is uncontrollable, $\ddot{x}_0$) while as there are only two outputs thus it is necessary to make some assumptions in order to be able to decouple the system. The final goal is to have one force controlling one output.

We check the effect of the ground vibrations on the outputs by finding the transfer functions between $\ddot{x}_0$ and y_1 and between $\ddot{x}_0$ and y_2. Manual calculations of these transfer functions are rather long and very prone to errors because the resulting characteristic equation is eighth order; a computer solution is highly recommended. The relations are the following

$$\frac{Y_1}{\ddot{X}_0} = \frac{0}{s^8 + 213.4\,s^7 + 1.5\,s^6 + 4.4 \times 10^6\,s^5 + 2.3 \times 10^8\,s^4 + 3.5 \times 10^8\,s^3 + 3.5 \times 10^9\,s^2} \tag{7.18}$$

$$\frac{Y_2}{\ddot{X}_0} = \frac{1.6\,s^5 + 356.9\,s^4 + 2.41^6\,s^3 + 3.0 \times 10^7\,s^2 + 3.5 \times 10^8\,s + 3.5 \times 10^9}{s^8 + 213.4\,s^7 + 1.5 \times 10^6\,s^6 + 4.4 \times 10^6\,s^5 + 2.3 \times 10^8\,s^4 + 3.5 \times 10^8\,s^3 + 3.5 \times 10^9\,s^2} \tag{7.19}$$

The Bode plot of Equation (7.18) is shown in Figure 7-2. The plot shows that at input frequencies of $w_1 = 1$ rad/sec and $w_2 = 20$ rad/sec the magnitudes of the transfer function from $\ddot{x}_0$ to y_2 are 1.06 m and 2.39×10^{-4} m respectively, showing that the slower sinusoid has more impact on the output and passes through un-attenuated. The faster sinusoidal component, on the other hand, is attenuated in magnitude by a factor of 2.39×10^{-4}. The entire system is linear thus the output is scaled linearly by the magnitude of the input hence, with the values of e_1 and e_2 stated in the problem, the slower sinusoid will contribute at most $\pm 3 \times 10^{-7}$ m to the output, while as

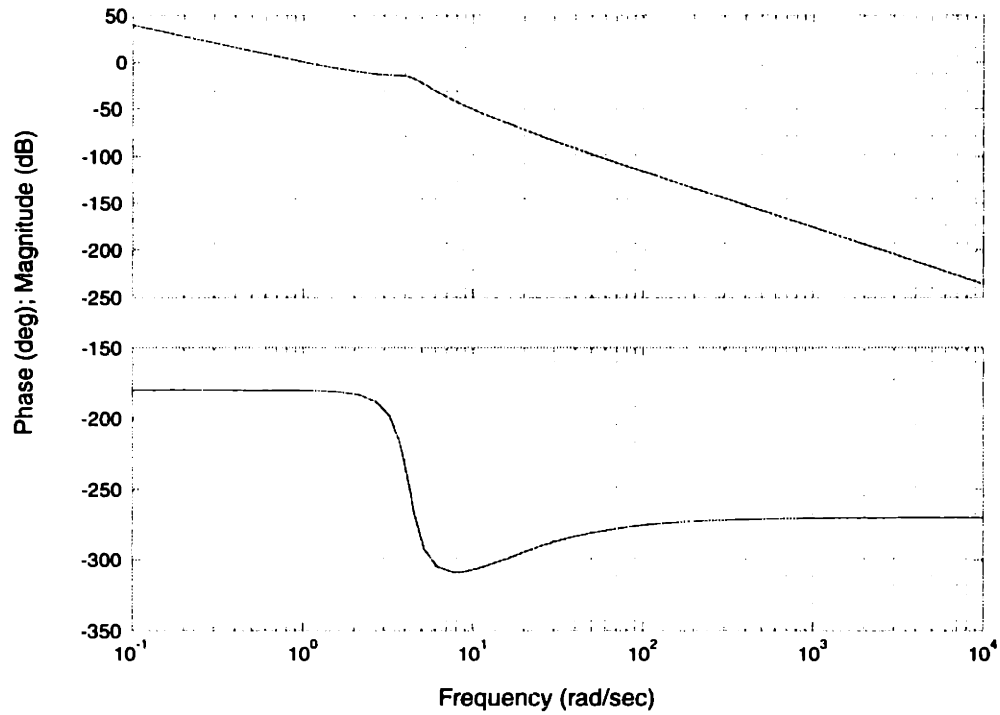

Figure 7-2: Bode plot of the transfer function from disturbance to y_2

the faster sinusoid will contribute at most $\pm 2.39 \times 10^{-8}$ m. Added together, the maximum effect of the disturbance on the output is $\pm 3.2 \times 10^{-7}$m. The effect of the disturbance is very small and will be reduced to the nanometer level with a controller gain of 100 or more at w = 1 rad/sec. Note that the gain of the controller enters the closed loop disturbance transfer function in the denominator (because it's in the disturbance feedback loop) therefore it will act to increase disturbance rejection.

The first assumption is thus to design the controller without the disturbance; this problem is now better suited to decoupling because there are two significant inputs and two outputs.

7.4.1 Decoupling Approach

Consider a system consisting of the mirror, m_4 and the fine-stage, m_3, coupled by a spring and damper and with the actuating force, F_2. The system is shown in Figure 7-3.

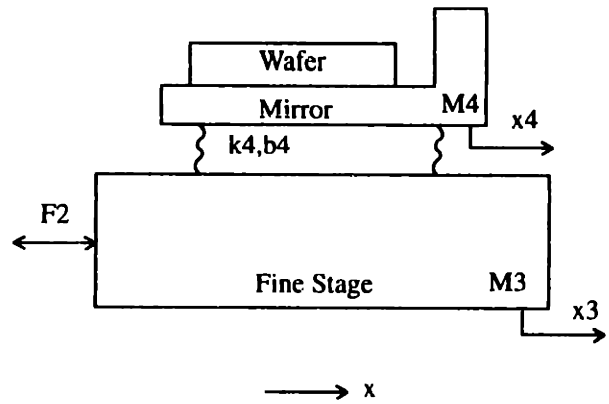

Figure 7-3: Coupled masses

Writing the equations of motion for the masses

$$F_2 + b_4(\dot{x}_4 - \dot{x}_3) + k_4(x_4 - x_3) = m_3\ddot{x}_3 \tag{7.20}$$

$$m_4\ddot{x}_4 + b_4\dot{x}_4 + k_4x_4 = b_4\dot{x}_3 + k_4x_3 \tag{7.21}$$

Taking the Laplace transforms of these and re-arranging, the following transfer functions are obtained.

$$\frac{X_4}{F_2} = \frac{b_4\,s + k_4}{s^2\,(m_3m_4\,s^2 + (m_3 + m_4)b_4\,s + (m_3 + m_4)k_4)} \tag{7.22}$$

$$\frac{X_3}{F_2} = \frac{m_4\,s^2 + b_4\,s + k_4}{s^2\,(m_3m_4\,s^2 + (m_3 + m_4)b_4\,s + (m_3 + m_4)k_4)} \tag{7.23}$$

The Bode plots for this system are shown in Figure 7-4

There is a resonance at 1225 rad/sec for both systems, and there is also an anti-resonance at 1000 rad/sec for the $\frac{X_3}{F_2}$ transfer function. The anti-resonance is due to a complex zero in the transfer function and it is interpreted as the steady-state input dynamics for which m_3 doesn't move. This is identical to the natural frequency when m_3 is clamped. In other words it can be shown that if the mass m_3 were held fixed then this anti-resonance becomes a resonance for the clamped system (i.e. a complex pole). The plots further show that below the resonances, the plants are coupled above after the resonances, the subsystems become essentially decoupled. This can be seen

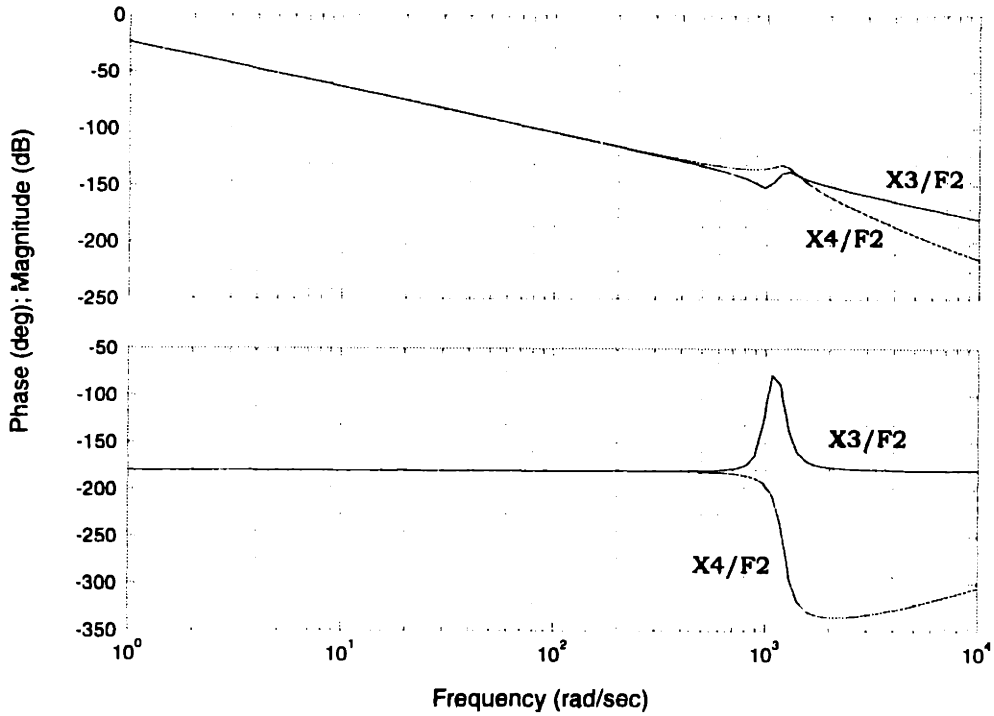

Figure 7-4: Bode plots of the coupled m3-m4 subsystem

by the fact that the magnitude plot for $\frac{X_3}{F_2}$ rolls off with a -40 dB slope, while as the $\frac{X_3}{F_2}$ plot rolls off with a -60 dB slope. Two systems are coupled when their dynamics depend on each other. In this problem, the force F_2 applied to one stage will transmit through to the other stage as is shown by the magnitude plot of the bode diagram. In particular we see that below 1000 rad/sec the input force affects x_3 and x_4 equally hence the two masses will move together as a rigid body.

The decoupling approach presented in the preceding paragraph is true provided that the inputs are pure sinusoids. In reality, though, the actuation forces are not pure sinusoids and will contain an entire spectrum of frequencies. As was seen above, for frequencies above 1000 rad/sec the decoupling argument will not hold because the mirror and the fine stage will become progressively more uncoupled as higher frequency content is introduced in the input. Nevertheless, by selecting a smooth trajectory, the high frequency content of the input can be substantially reduced so the decoupling approach will be valid. Under the assumption that the dominant

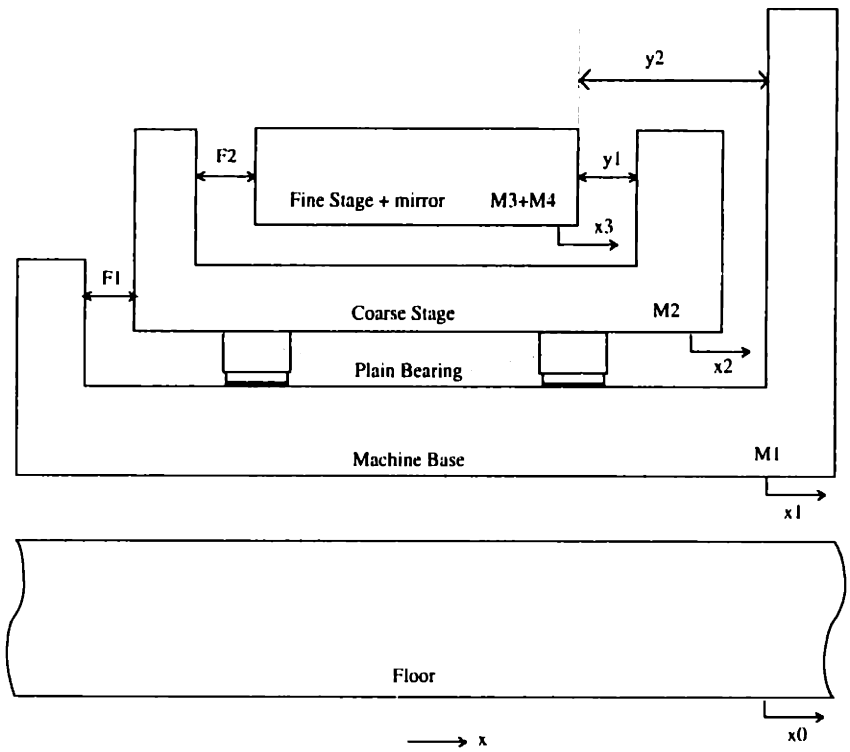

Figure 7-5: Decoupled stepper stage

frequencies are below 1000 rad/sec the two masses m_3 and m_4 would be perfectly coupled and would move as a rigid body of mass $(m_3 + m_4)$.

This canonical example showing how it may be possible to decouple masses if the input frequency is above resonance, can be extended to the other subsystems of the stage, in particular the $m_2 - m_3$ subsystem. Thus when operating between the resonant frequencies of the entire plant, it may be possible to make the statement that the entire stage can be approximated as a system of only masses. The $m_3 - m_4$ subsystem resonance was found to be $\sqrt{\frac{m_3+m_4}{m_3\,m_4}\,k_4} = 1225$ rad/sec and a similar analysis shows that the resonance for the $m_2 - m_3$ subsystem is $\sqrt{\frac{m_2+m_3+m_4}{m_2\,(m_3+m_4)}\,k_3} = 11.4$ rad/sec. The latter is the lower resonance and therefore, between approximately $10 < \omega < 1000$ rad/sec, the decoupled mass assumption (remember that the $m_3 - m_4$ subsystem is coupled in this frequency range, and is acting as a single mass) may be applicable. In this frequency range the system would look like Figure 7-1 without the springs and dampers and with m_3 and m_4 rigidly connected. This new model is shown in Figure 7-5. For this system the equations become

$$(m_3 + m_4)\, s^2 X_3(s) = F_2 \tag{7.24}$$

$$m_2\, s^2 X_2(s) = -F_2 + F_1 \tag{7.25}$$

$$m_1\, s^2 X_1(s) = -F_1 \tag{7.26}$$

With m_3 and m_4 lumped together it would be preferable to keep y_1 almost zero in order not to exceed the 200 μm limit imposed by physical constraints on y_1. Figure 7-5 shows that

$$y_2 = y_1 + q_3$$

Thus trying to keep y_1 almost zero suggests using y_2 and q_3 as outputs, and feeding the same reference trajectory into them. A linear transformation, T_0, can be carried out on the system outputs y_1 and y_2 so that the new outputs to be fed back are y_2 and q_3. The linear transformation thus makes the following changes of variables

$$y_a = y_2$$

$$y_b = y_2 - y_1$$

Hence

$$T_0 = \begin{pmatrix} 0 & 1 \\ -1 & 1 \end{pmatrix} \tag{7.27}$$

With this new choice of outputs the free-mass system equations (7.24), (7.25), (7.26) and (7.27) can be written in matrix notation as

$$\begin{pmatrix} y_a \\ y_b \end{pmatrix} = \frac{1}{s^2} \begin{pmatrix} -\frac{1}{m_1} & -\frac{1}{m_3+m_4} \\ -\frac{1}{m_1} - \frac{1}{m_2} & \frac{1}{m_2} \end{pmatrix} \begin{pmatrix} F_1 \\ F_2 \end{pmatrix} \tag{7.28}$$

For the sake of future reference, we shall define

$$T_i = \begin{pmatrix} -\frac{1}{m_1} & -\frac{1}{m_3+m_4} \\ -\frac{1}{m_1} - \frac{1}{m_2} & \frac{1}{m_2} \end{pmatrix}^{-1}$$

In order to decouple the system, it is convenient to define a new set of inputs

$\begin{pmatrix} F_a \\ F_b \end{pmatrix}$ such that the transfer matrix between inputs and outputs is diagonal; this can be achieved by defining new outputs as

$$\begin{pmatrix} F_1 \\ F_2 \end{pmatrix} = T_i \begin{pmatrix} F_a \\ F_b \end{pmatrix} \tag{7.29}$$

substituting (7.28) into (7.29) results in the following matrix equation which shows the perfect theoretical decoupling.

$$\begin{pmatrix} y_a \\ y_b \end{pmatrix} = \begin{pmatrix} \frac{1}{s^2} & 0 \\ 0 & \frac{1}{s^2} \end{pmatrix} \begin{pmatrix} F_a \\ F_b \end{pmatrix} \tag{7.30}$$

We note again that (7.30) is only valid in the range in which the input frequency is $10 < \omega < 1000$ rad/sec.

Provided that the assumptions on the input frequency are satisfied, we have seen in (7.30) that the system is almost decoupled and equivalent to two separate masses acted on by forces. To verify the validity of this modeling assumption, the following page shows the bode plots of the entire plant with the linear transformations. In other words the bode plots are for the system : $[T_o] \times [Plant] \times [T_i]$

Figure 7-6 and Figure 7-7 show how, between the resonances, the two systems are separated by as much as three decades. While not a perfect decoupling, it is enough to warrant proceeding with this approach.

For completeness we note that the problem could have been approached by feeding back from y_1 and y_2 rather than y_2 and q_3. The same reasoning applies in this case, but there will be no T_o matrix. As for the reference trajectories, the one for y_2 will be analogous to the one which will be presented later in this module, and the one for y_1 should be zero.

Continuing with the decoupled system, two free masses have no damping and hence require the use of a lead compensator. Given the limits on F_1 and F_2, it is necessary to avoid excessive control usage hence the lead will be placed in the feedback loop. A lag compensator is also needed to add low frequency stiffness to the system

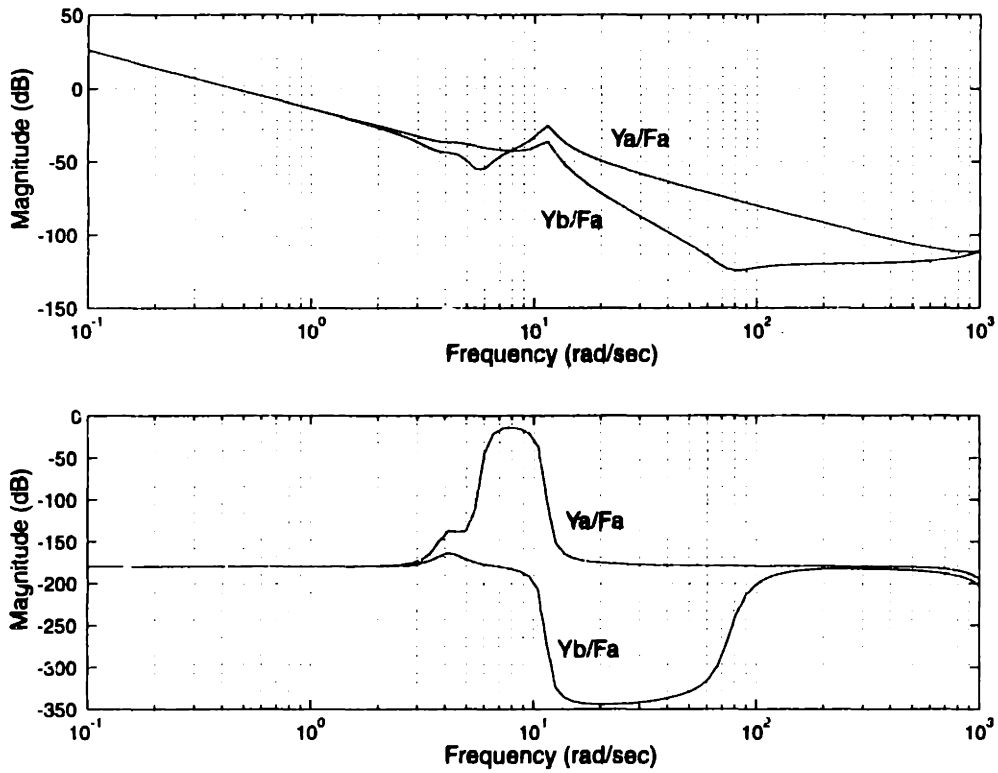

Figure 7-6: Bode plots of transformed stepper stage alone (Input is Fa)

as well as to help decouple the dynamics. The reasoning is that at frequencies in the range from 10 rad/sec to 1000 rad/sec the linear transformations will be the major factors in the decoupling. On the other hand at low frequencies this decoupling is no longer valid because below resonance, dampers and springs cannot be neglected. We are thus counting on the controller gains to create decoupling at low frequencies. Altogether the compensated system will look like Figure 7-8

We have added a lag compensator on the coarse stage although we are aware that it may cause hunting because of friction. We wished to test it out and possibly remove it if performance was bad, however simulations showed that we could achieve good performance with it there, so it was kept.

It is important to note that the lead networks in the feedback paths are designed for unity DC gain so as to maintain accuracy.

The lead compensator is chosen for cross-over around 300 rad/sec. At this frequency the plant is almost decoupled (see Figure 7-6), so the decoupling is maintained.

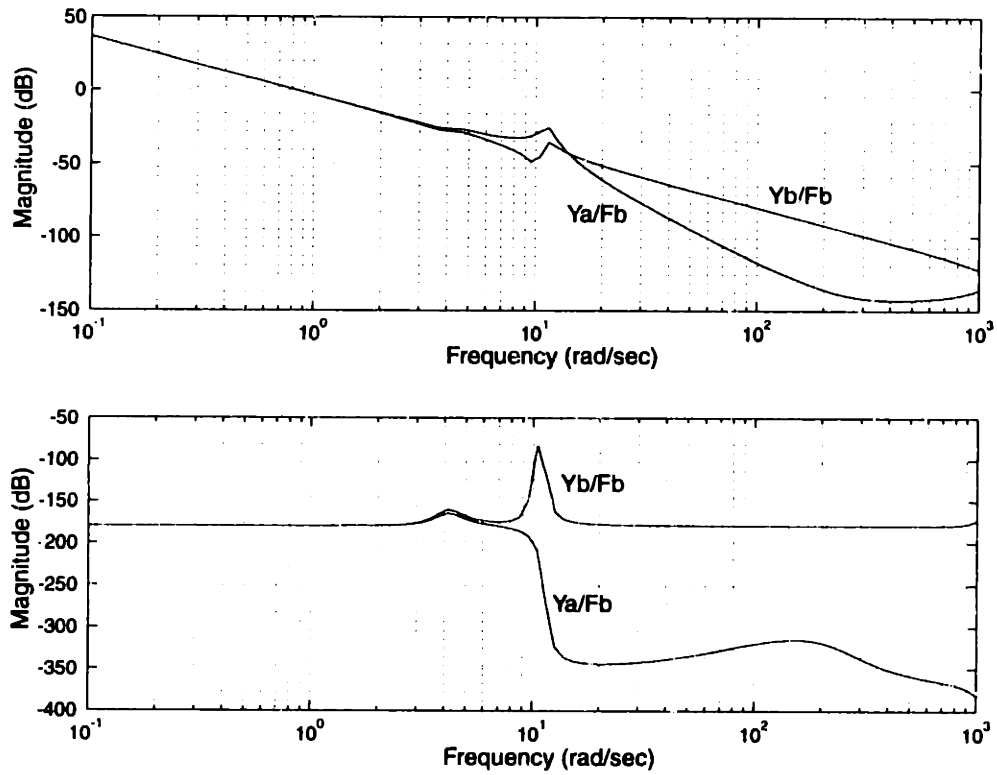

Figure 7-7: Bode plots of transformed stepper stage alone (Input is Fb)

Crossover at 100 rad/sec would decouple the plant more but the bandwidth will be rather low and hence the dynamics will be slow. Although bandwidth and crossover are different concepts, as a rough approximation their values are similar.

7.4.2 Lead - Lag compensator

The controllers used are of the form

$$G_{lead} = K_{lead}\left(\frac{\tau s + 1}{\alpha\tau s + 1}\right)$$

$$G_{lag} = K_{lag}\left(\frac{\gamma s + 1}{s}\right)$$

Combining the two controllers gives the following form for the lead-lag compensator

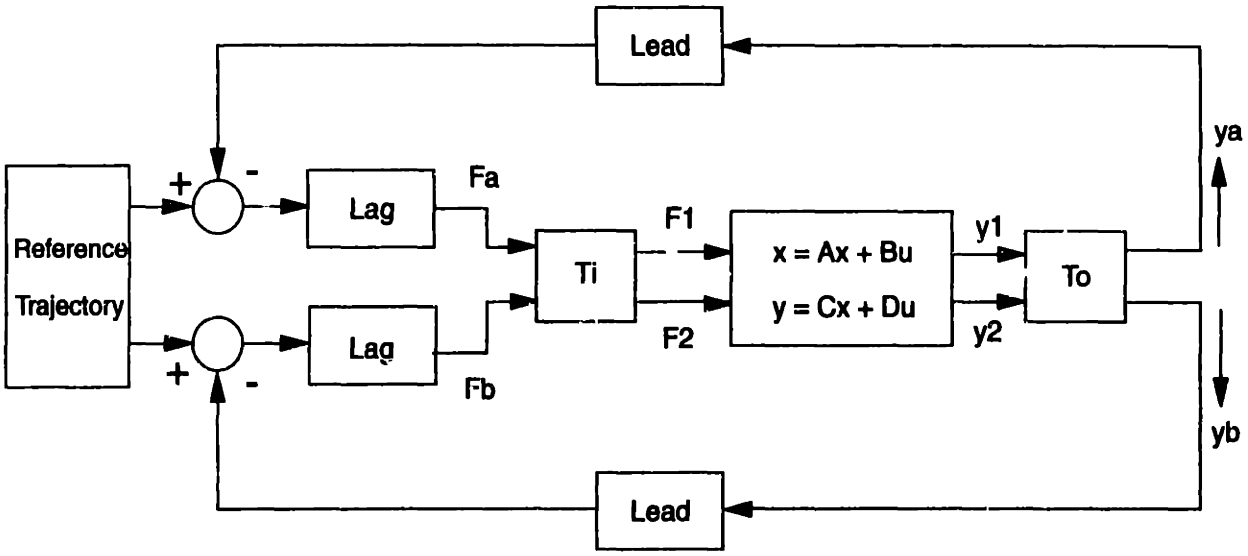

Figure 7-8: Compensated wafer stage

$$G_c = K_c \left(\frac{\tau s + 1}{\alpha \tau s + 1} \right) \left(\frac{\gamma s + 1}{s} \right) \tag{7.31}$$

Lead portion

It was seen from Figure 7-6 that 300 rad/sec was a reasonable frequency for crossover. In order to find the controller parameters which will place the crossover there we can use the following properties of the lead compensator. It can be shown that the peak of the phase bump is at

$$\omega_m = \frac{1}{\sqrt{\alpha}\,\tau}$$

while as the height of the phase bump is given as

$$\sin \phi_m = \frac{1 - \alpha}{1 + \alpha}$$

With $\alpha = 0.1 \quad \Rightarrow \quad \phi_m = 54.9\,\text{deg.}$

Setting $\omega_m = 300 \quad \Rightarrow \quad \tau = 1.05 \times 10^{-2}\text{sec.}$

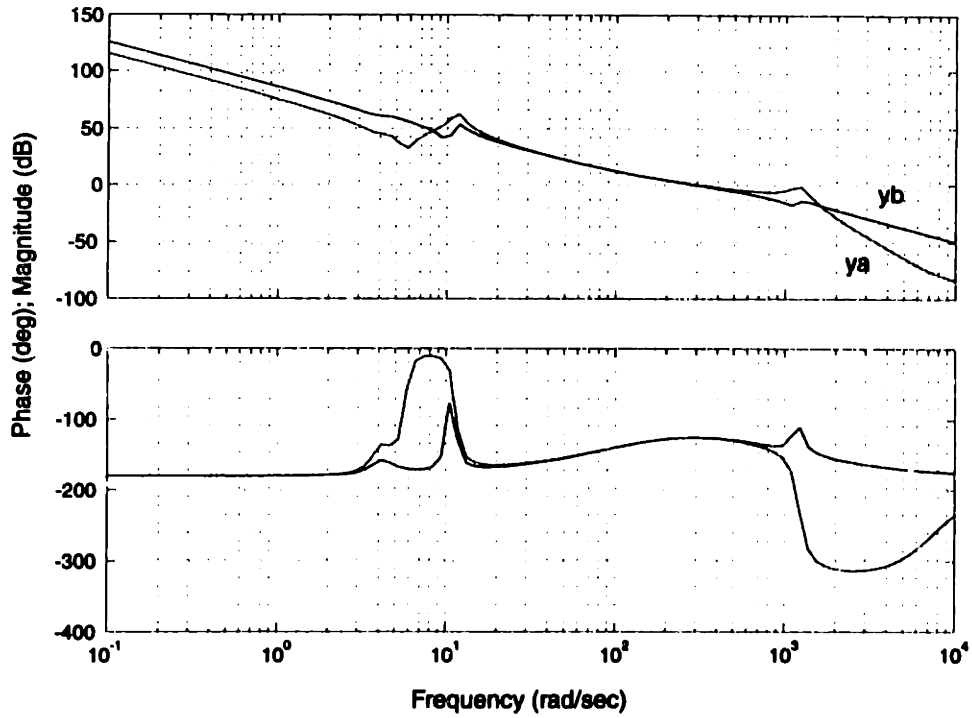

Figure 7-9: Lead compensated loop transmissions for y_a and y_b

To find the gain K_{lead} that will allow a crossover at 300 rad/sec we solve for

$$\left| \frac{K_{lead}}{s^2} \left(\frac{\tau s + 1}{\alpha \tau s + 1} \right) \right| = 1 \ .$$

This gives $K_{lead} = 28551$.

In order to decide where to place the lag compensator, let us plot the loop transmission for the system with only the lead compensator and see where would be a good location for the lag. Figure 7-9 a plot of the loop transmission for the lead compensated y_a loop with the y_b loop open, and a plot of the lead compensated y_b loop with y_a open. We see that the phase margin is about 55 degrees, and we also note that the resonant peak at 1225 rad/sec is close to crossing over unity gain once more. This is a problem as will be detailed later. To be conservative, we could keep the phase margin around 40 degrees, but aiming for around 30 degrees of phase margin is suitable provided we do not anticipate too much disturbance.

Lag portion

The upper break point of the lag is chosen at $\omega = 150$ rad/sec. With this break point, the lag controller will decrease the phase by about 20 degrees at crossover, but the system will greatly benefit from the added low frequency stiffness obtained. Remember that we are counting on the controller gain to achieve decoupling at low frequencies because our model without springs and dampers is no longer valid.

At crossover, the magnitude of the loop transmission is unity, thus in order to find K_{lag} we must solve

$$|G_{lag}| = 1 \quad .$$

This requires

$$K_{lag} = 134 \qquad \text{(for break} - \text{point at 150 rad/sec)} \quad .$$

The full controller is thus

$$G_c = 3843250 \left(\frac{1.05 \times 10^{-2} s + 1}{1.05 \times 10^{-3} s + 1} \right) \left(\frac{0.00667 s + 1}{s} \right) \quad . \tag{7.32}$$

The bode plots which follow, depict various aspects of the compensated system. Figure 7-10 and Figure 7-11 show the loop transmissions for the system with one loop closed and the other open. As part of a design having multiple loops, final stability is best viewed as a sequential loop closing problem in which loops are closed one at the time, and stability checked at each stage. Normally the faster loop is closed before the slowest loop (such as a tachometer feedback in a motor controller) but the way the problem has been addressed until now, the two loops have essentially the same bandwidth and crossover. Therefore there is no preference as to which loop is closed first. The two loop transmissions were calculated as follows

$$\begin{pmatrix} y_a \\ y_b \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix} \begin{pmatrix} F_a \\ F_b \end{pmatrix} \tag{7.33}$$

Where α_i = the corresponding transfer function. Closing the loop around y_a we get

$F_a = -Lead \times Lag \times y_a$. Therefore the loop path from y_a back to y_a is given by $-\alpha_1 \times Lead \times Lag$. Similarly the path around the loop from y_b back to y_b is $-\alpha_4 \times Lead \times Lag$. The negative of these functions (i.e. negative of the loop transmission) is what is plotted.

We note that both loops are stable because their phase margins are positive and so are their gain margins. Note, though, that for the loop around y_a the resonant peak at 1225 rad/sec is extremely close to crossing over unity gain again. At that point the phase is less than -180 deg which means that if it were to crossover, the system would be unstable. This issue of multiple crossover is typical of systems with resonances. Trying to solve the problem by adding a pole which breaks before the resonance will lower the dangerous peak, but it will kill the phase margin bringing the plant closer to instability. A possible alternative to overcome this problem is to add physical damping to the system, i.e. add pads of vibration absorbing material or whatever is suitable for the application at hand.

Figure 7-12 and Figure 7-13 depict the loop transmission around y_a and y_b respectively, when both loops are closed. As an example, the loop transmission around y_a was found as follows

$$\begin{pmatrix} y_a \\ y_b \end{pmatrix} = \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix} \begin{pmatrix} F_a \\ F_b \end{pmatrix}$$

$$= \begin{pmatrix} \alpha_1 & \alpha_2 \\ \alpha_3 & \alpha_4 \end{pmatrix} \begin{pmatrix} -Lead \times Lag \times y_a \\ -Lead \times Lag \times y_b \end{pmatrix}$$

Expanding these equations we get the following

$$y_a = -(\alpha_1 \times Lead \times Lag)\, y_a - (\alpha_2 \times Lead \times Lag)\, y_b \qquad (7.34)$$

$$y_b = -(\alpha_3 \times Lead \times Lag)\, y_a - (\alpha_4 \times Lead \times Lag)\, y_b \qquad (7.35)$$

Eliminating y_b from (7.35) the loop transmission becomes

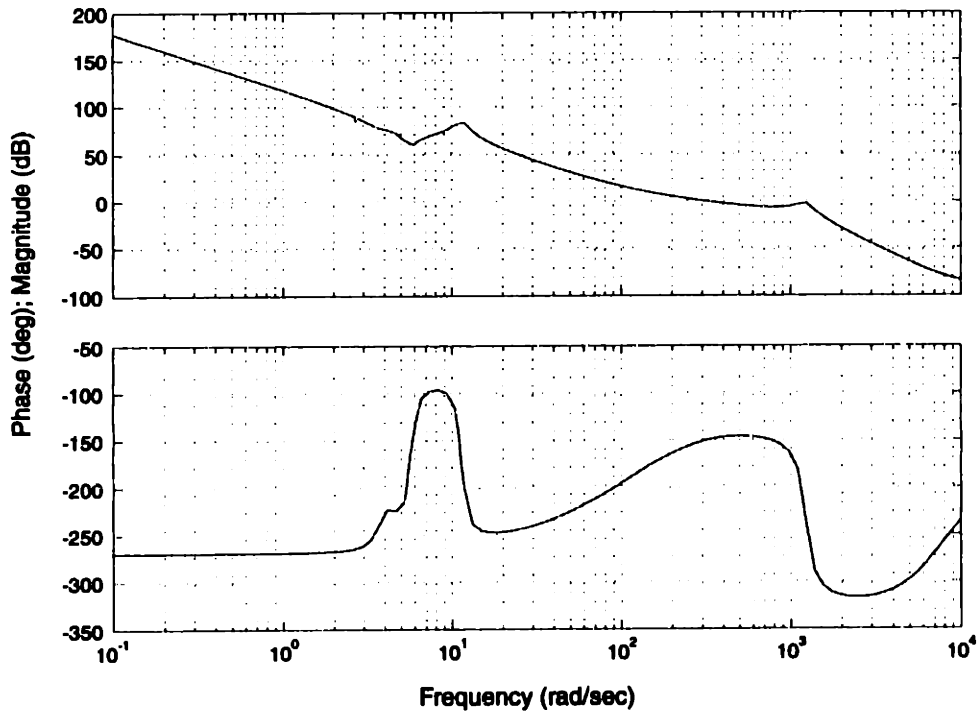

Figure 7-10: Loop transmissions for y_a with y_b loop open

$$y_a = \left(-\alpha_1 \times Lead \times Lag + \frac{\alpha_2 \times \alpha_3 \times (Lead \times Lag)^2}{1 + \alpha_4 \times Lead \times Lag} \right) y_a \tag{7.36}$$

The plots in Figure 7-12 and Figure 7-13 show that the phase margin is 29.4 deg and 28.1 deg respectively, although the resonance in the mirror stage makes the loop around y_a almost crossover again, as pointed out before.

The closed loop transfer function plotted in Figure 7-14 and Figure 7-15 are the relation between $\mathbf{y_{ref}}$ and $\mathbf{y}$. The decoupling which is achieved is apparent in these plots and it shows that one reference trajectory has control over one output far more than it has over the other output. More specifically it is clearly seen that y_{aref} guides y_a via F_a and y_{bref} guides y_b via F_b.

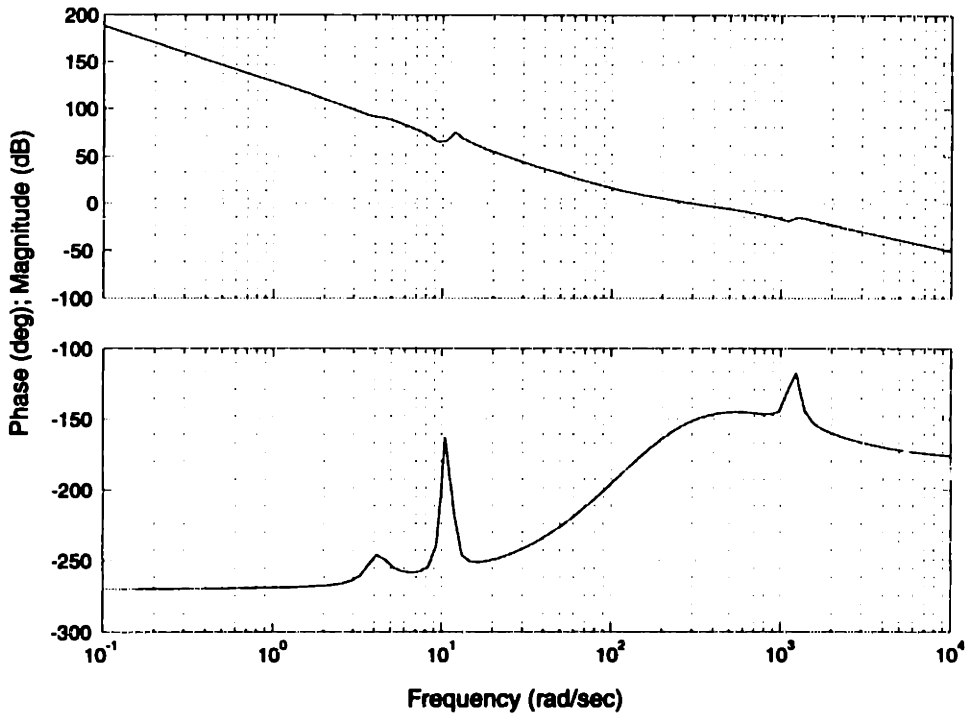

Figure 7-11: Loop transmissions for y_b with y_a loop open

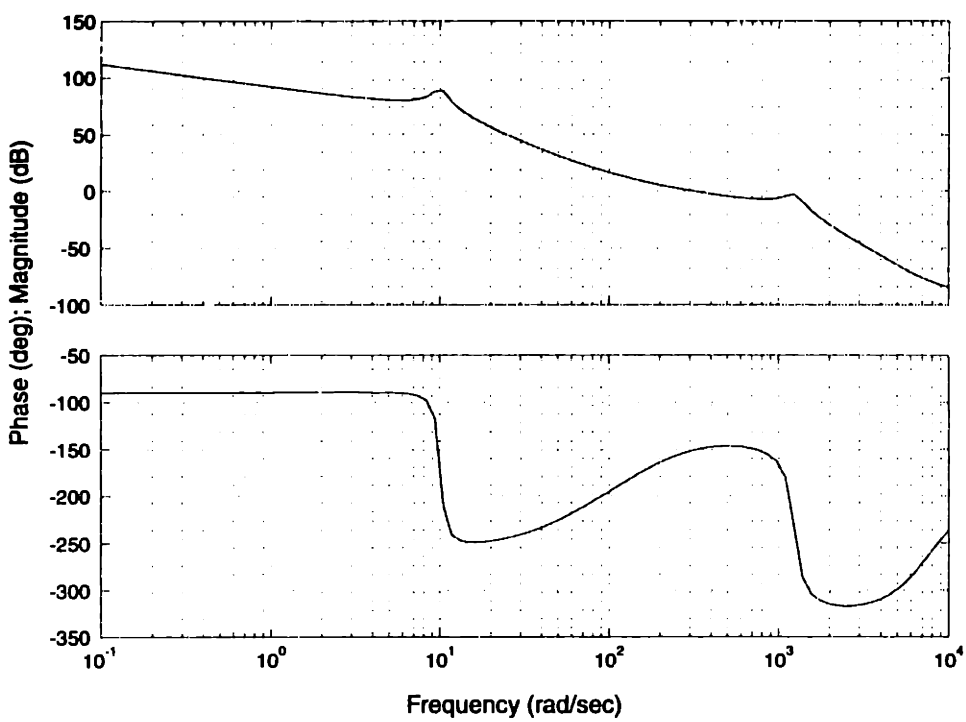

Figure 7-12: Loop transmission for y_a with both loops closed

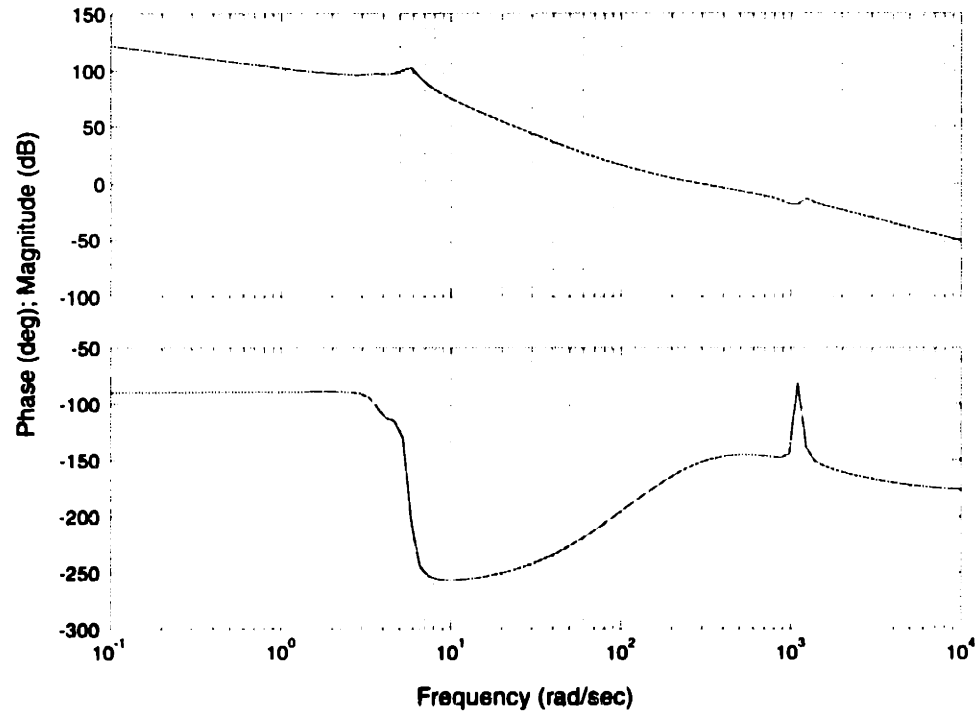

Figure 7-13: Loop transmission for y_b with both loops closed

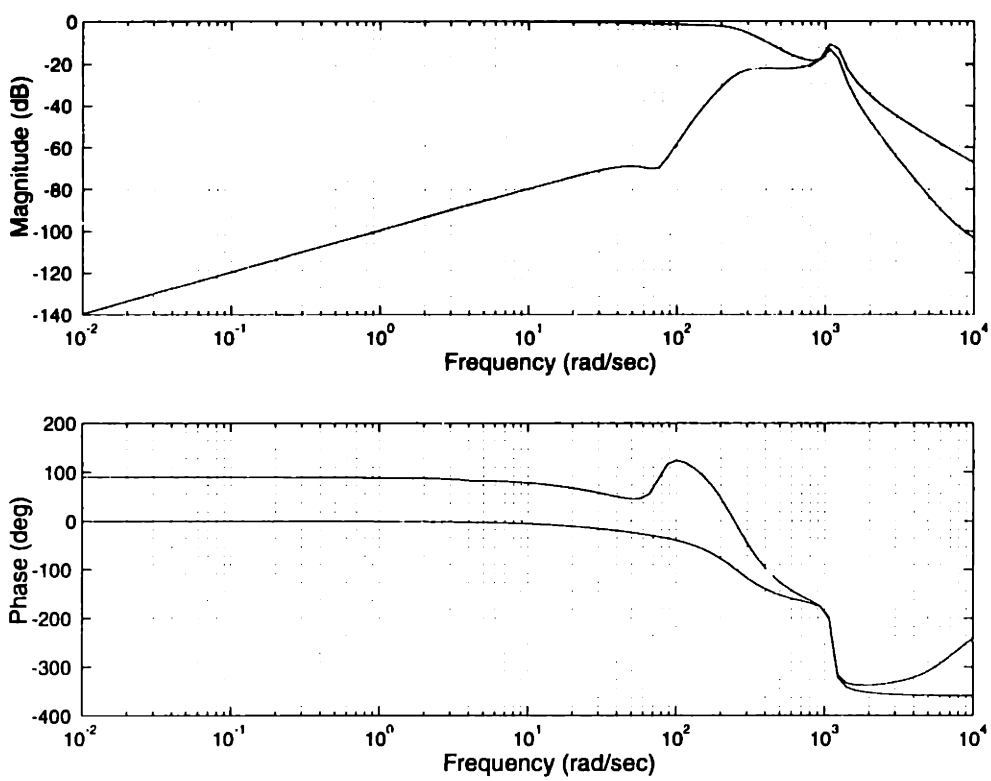

Figure 7-14: Bode plots of the closed loop dynamics for the compensated system (Input y_{aref})

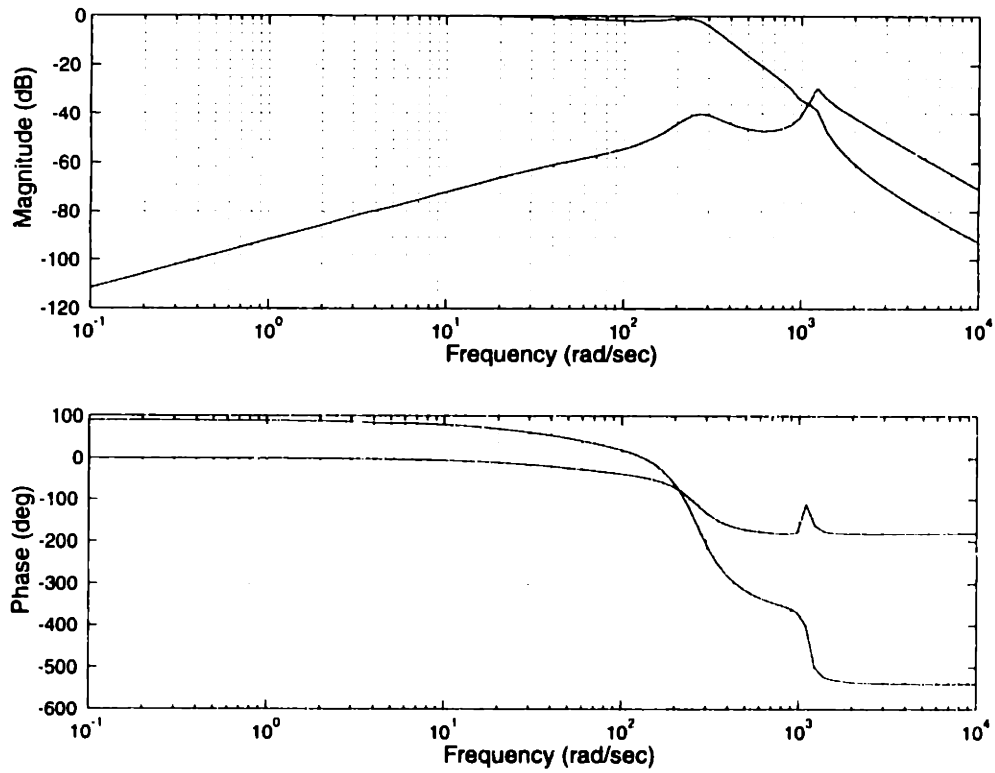

Figure 7-15: Bode plots of the closed loop dynamics for the compensated system (Input y_{bref})

7.5 Reference Trajectory

While as in simulations a step function can be easily implemented, in a real mechanical system applying such a drastic input is hardly a good idea. The high frequency content of the step will excite all sorts of unmodeled dynamics of the system, and will more easily lead to things such as fatigue damage, and so on. For mechanical systems with large inertial loads, this is even more true. For this reason the trajectory which is supplied to the plant should be such that it is relatively smooth while as still achieving the desired position. It is possible to filter a step function and feed this as the input trajectory, but it must be kept in mind that filters may only be asymptotic in the way that their output approaches the final value, i.e., the final value is only reached as $t \rightarrow \infty$. A better approach is to fit a trajectory which reaches the desired final value in a finite, pre-determined time. The approach which was taken in developing the figures which follow, was to select a cubic trajectory profile which would reach

the desired set-point in a user-determined time and then be fixed at this value until the stage needed to step again.

Figure 7-16 shows the trajectory used.

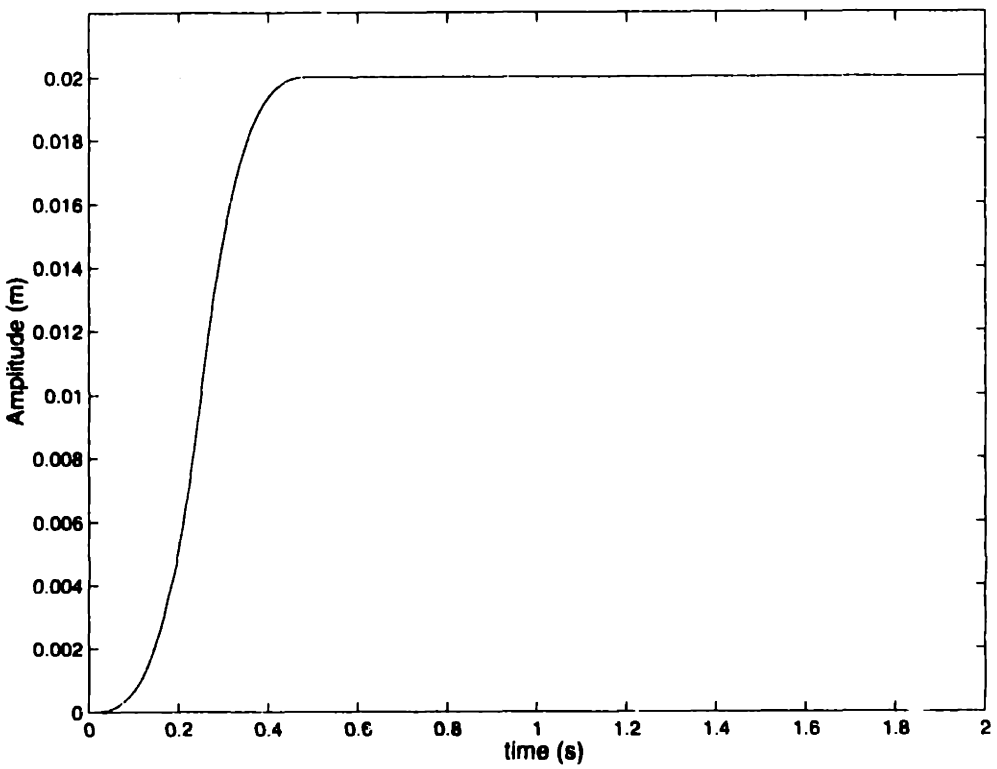

Figure 7-16: Cubic reference trajectory

This reference trajectory was generated based on three user-defined parameters :

1. stepping time, `step_time`, (the desired interval between steps.)
2. rise time, t_s, (time at which the trajectory reaches the set point of 0.02 m.)
3. time interval, or, sampling time, T (a necessary parameter because even though the trajectory should be continuous in time, the computer generates it by discrete points).

Using the general form for a cubic equation with no second order or first order terms, i.e. an equation of the form $y = kt^3$, the constant k was determined by setting y = 0.01 m at t=ts/2 s. The cubic trajectory was plotted in the time interval $-ts/2 \rightarrow ts/2$ and then the part defined in the negative time interval was cut and

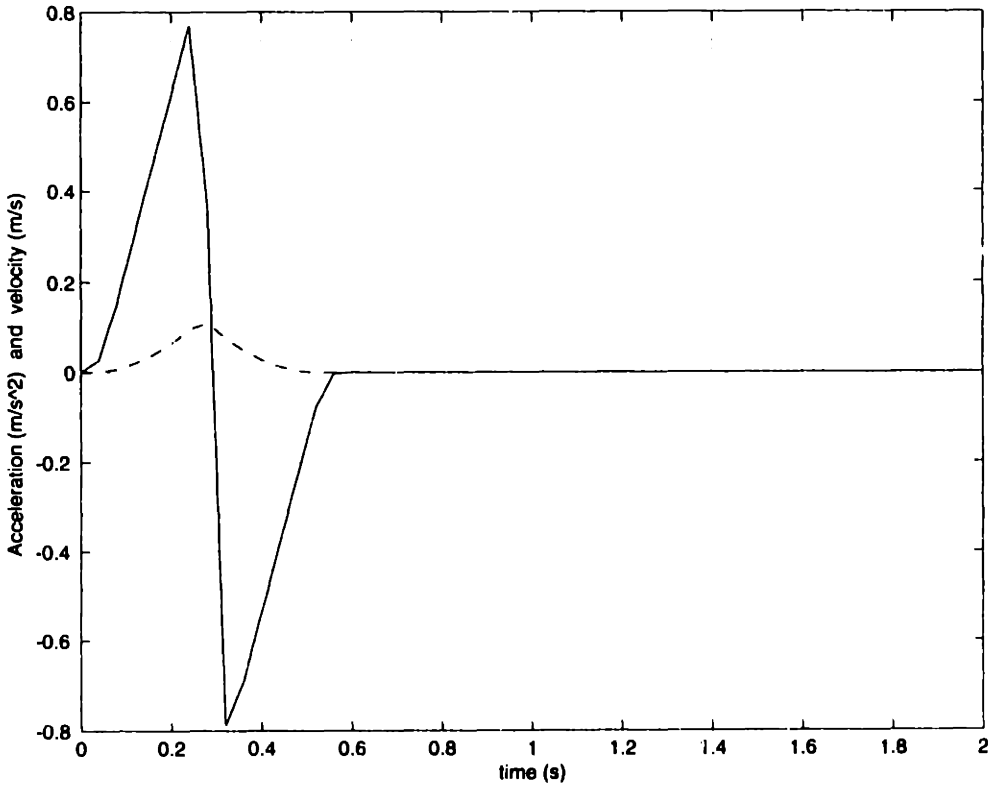

Figure 7-17: Velocity (dashed) and acceleration (solid) profiles for the reference trajectory of Figure 7-16.

attached after the part defined over the positive time interval. This procedure gives the first `ts` seconds of motion. The remaining trajectory, namely that over the time interval `ts` $\rightarrow$ `step_time`, was filled in with a constant vector of 0.02. The entire trajectory looks like Figure 7-16. Note that the trajectory as it stands is only defined in the time interval 0 $\rightarrow$ `step_time` sec, but it can be replicated by concatenation so as to achieve a continuous reference trajectory for as long a period as desired. Figure 7-17, instead, shows the velocity and acceleration profiles for this trajectory. In particular we note the sharp change in acceleration which will appear directly in the control effort as will be discussed later on.

The results of the simulations are shown in the following two sections. The first section shows the response of the continuous-time system and the second section shows the response when the system is implemented in discrete-time as it would be in a real setup. The two results are very similar and the controller did not even have to be adjusted when mapped to discrete time. The controllers were designed

without taking into account the friction, and the good performance shows that with a high enough gain the disturbance rejection of the system will reduce the effect of the disturbance to the extent that it will affect performance very little. Quantization will be seen to be more problematic than ground vibrations.

7.6 Simulations of the controller implemented in continuous-time

Figure 7-18 shows the output responses y_a and y_b of the system. The two responses are essentially identical so only one curve is apparent.

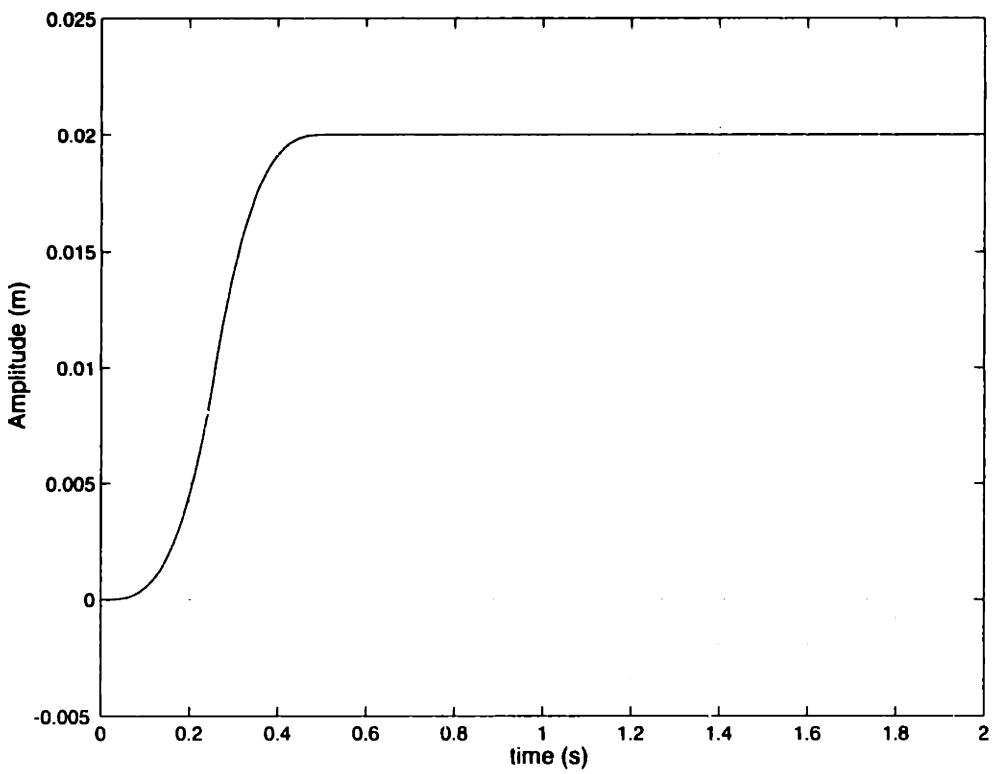

Figure 7-18: Time response of y_a and y_b

The Simulink block diagram used to generate the plots is shown if Figure 7-19 and is of the same topology as the complete, compensated system shown in Figure 7-8.

Figure 7-20 shows an expanded section of the output response graph, Figure 7-18. We see that the nonlinear nature of quantization introduces fast dynamics.

Figure 7-21 shows the time response of the output y_1. As can be seen, the maximum displacement in y_1 is 2.75 μm which is well within the limit of the allowable 200 μm set by the constraints. Keeping y_1 so low was possible because the same input trajectory was supplied to both y_a and y_b of the transformed system (as was explained earlier). A striking phenomena of this plot is the limit cycle which has been set up around the amplitude = 0 m point. This limit cycle is yet another result of the quantization in y_1. any movement greater than the quantization level will result in a jump in the output which is fed back. Hence in trying to set y_1 to zero, the controller is continuously "kicking" the output around the desired position. Zooming in on the limit cycle we see that the spikes are 25 nm in height.

Also quantized at 25 nm is the output shown in Figure 7-22. In this figure note that as well as the LVDT quantization, the interferometer quantization is also visible (the smaller transitions). This is because $y_b = y_2 - y_1$ and since y_a is quantized at 2.5 nm and y_b is quantized at 25 nm, both quantizations appear together.

The rise time for the chosen reference trajectory is 0.5 s and all the following plots show that at half this rise time, i.e. at the point where the second derivative of this trajectory changes sign, the dynamics of the system receives a jolt. This is not surprising since the second derivative of the trajectory represents the acceleration, and by Newton's Second Law the latter is related to the force applied to the system.

Figure 7-23 shows how the outputs y_a and y_b differ from the reference trajectory. In the time between the start and half the rise time, the output trajectories continue diverging from the reference trajectory. This is due to the fact that the reference trajectory is a cubic and therefore there are not enough free integrators in the system to make the output converge without steady state errors. In fact if the cubic trajectory was allowed to continue to infinity, the error depicted in Figure 7-23 would increase without bounds. Between half the rise time and the rise time, the situation is reversed and the problem is perfectly symmetrical. Finally in the plateau between the rise time and the stepping time the reference trajectory is a constant and thus the output is able to follow it without steady state errors.

Figure 7-24 shows that all the actuation forces are well within the allowable lim-

its. Note the large jump in forces which occurs at 0.25 s. As explained before, this transition is due to the discontinuity in the acceleration profile associated with the reference trajectory; and acceleration is directly related to forces by Newton's Second Law of motion.

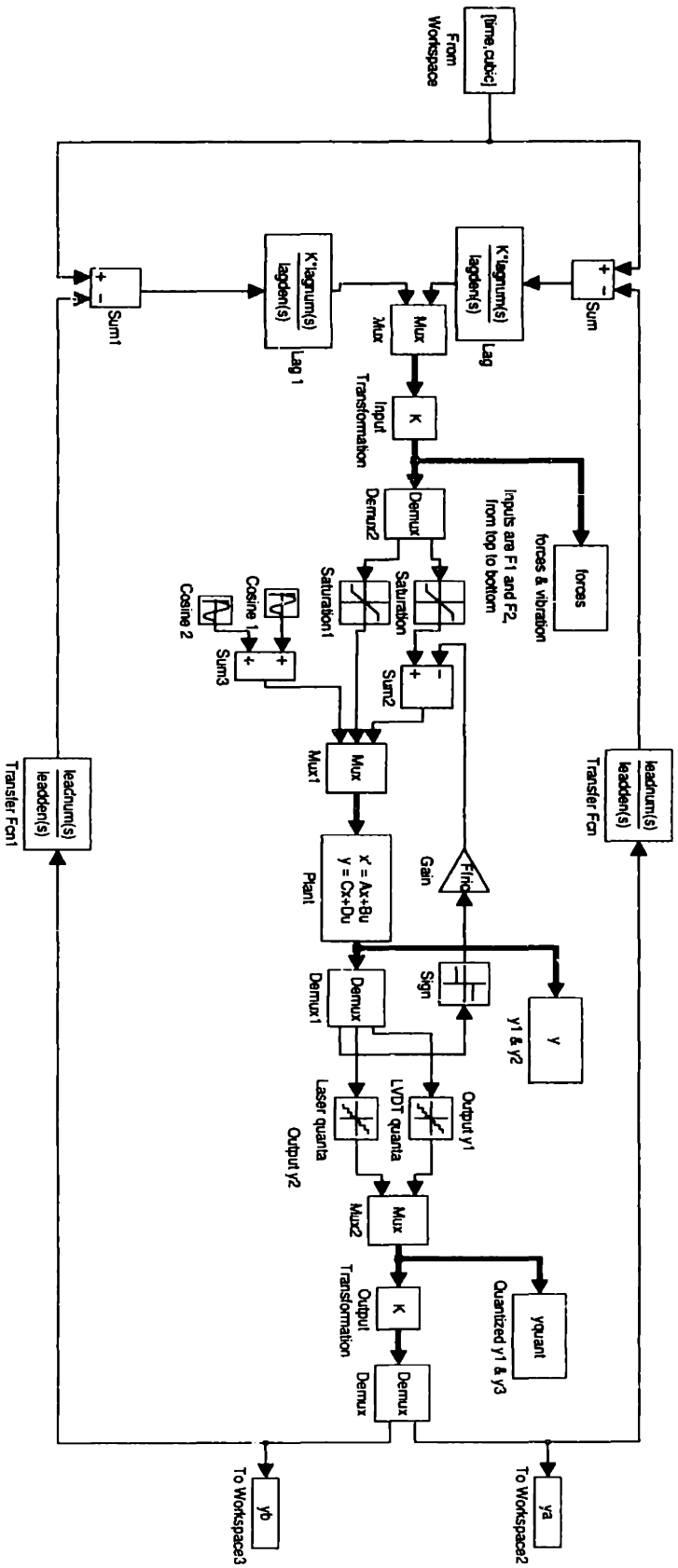

Figure 7-19: Simulink block diagram for continuous time controlled system

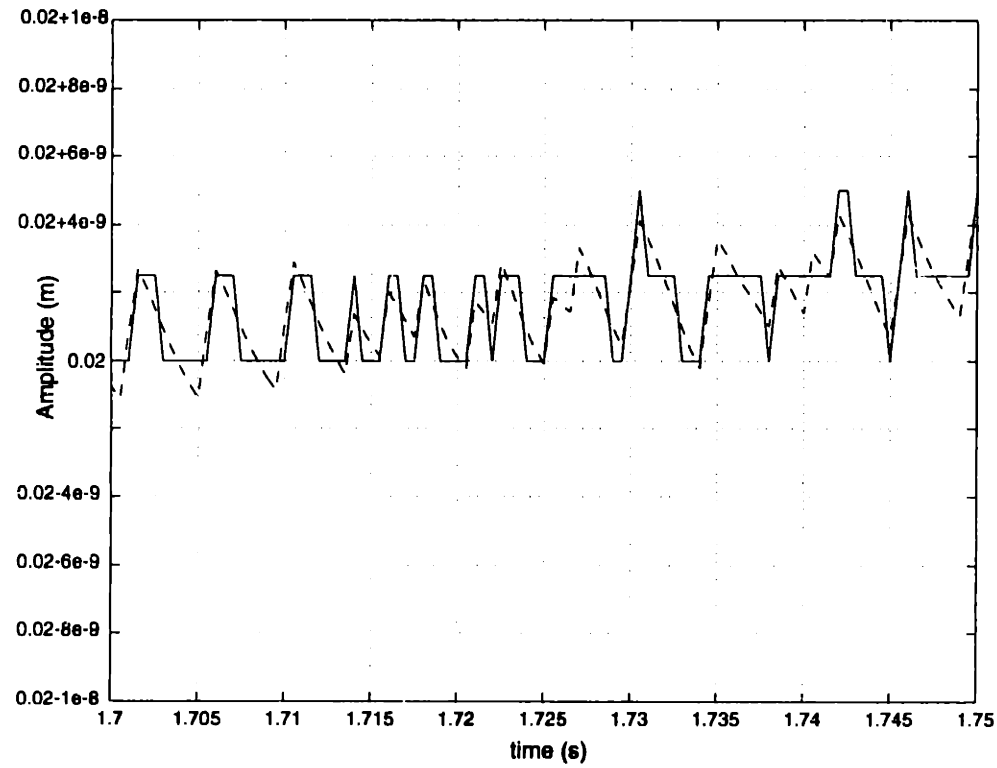

Figure 7-20: Zoom in of Figure 7-18 showing non-quantized output y_2 (dashed) and quantized y_a (solid).

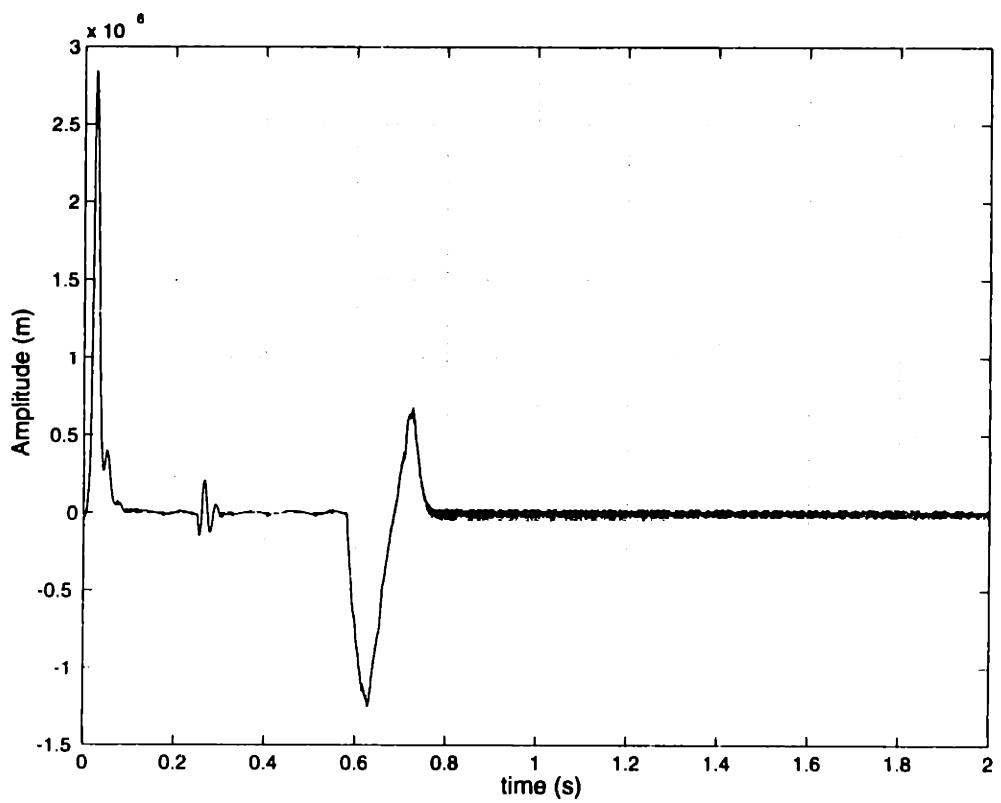

Figure 7-21: Output response y_1

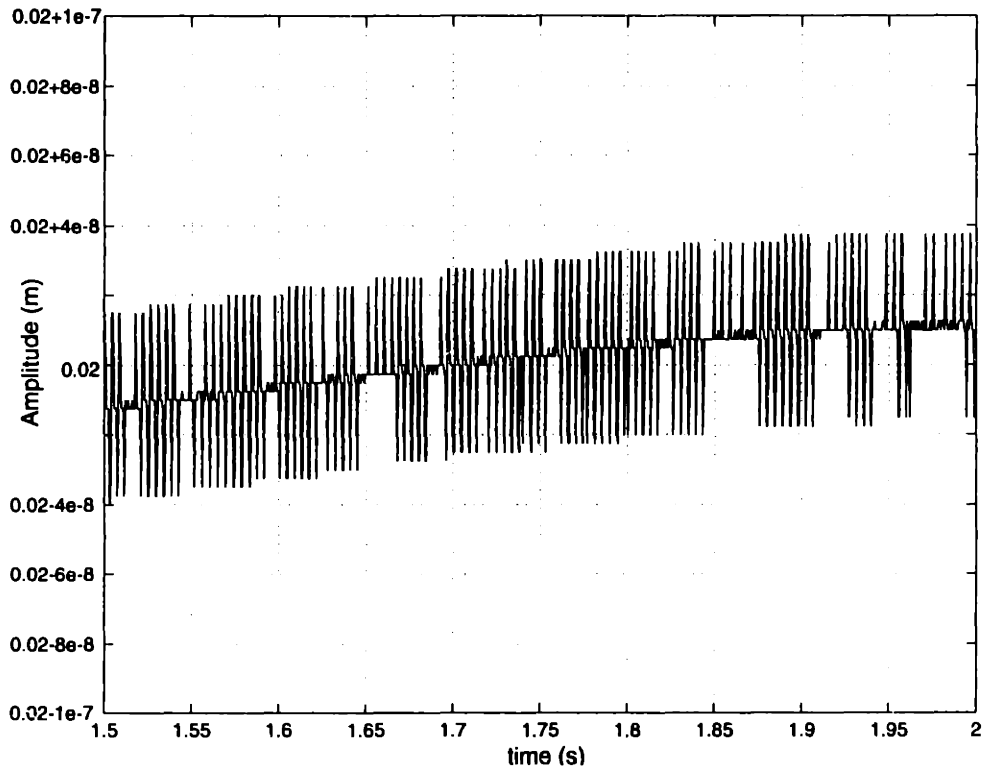

Figure 7-22: Output response of y_b.

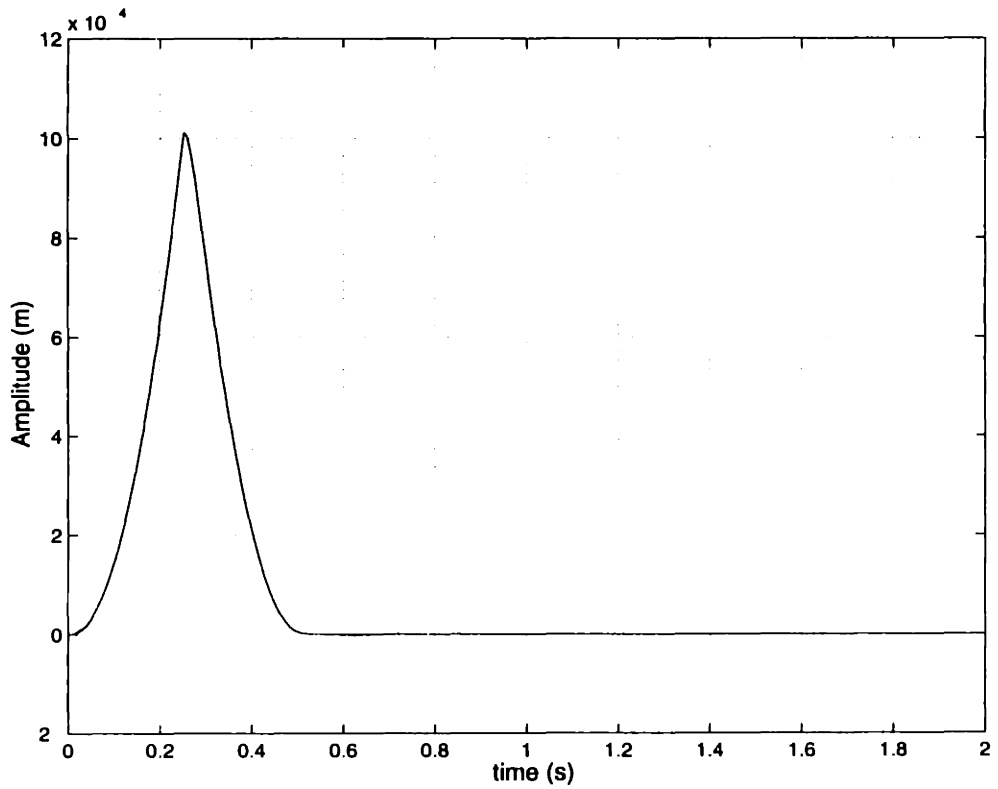

Figure 7-23: Discrepancy between the reference trajectory and the output responses

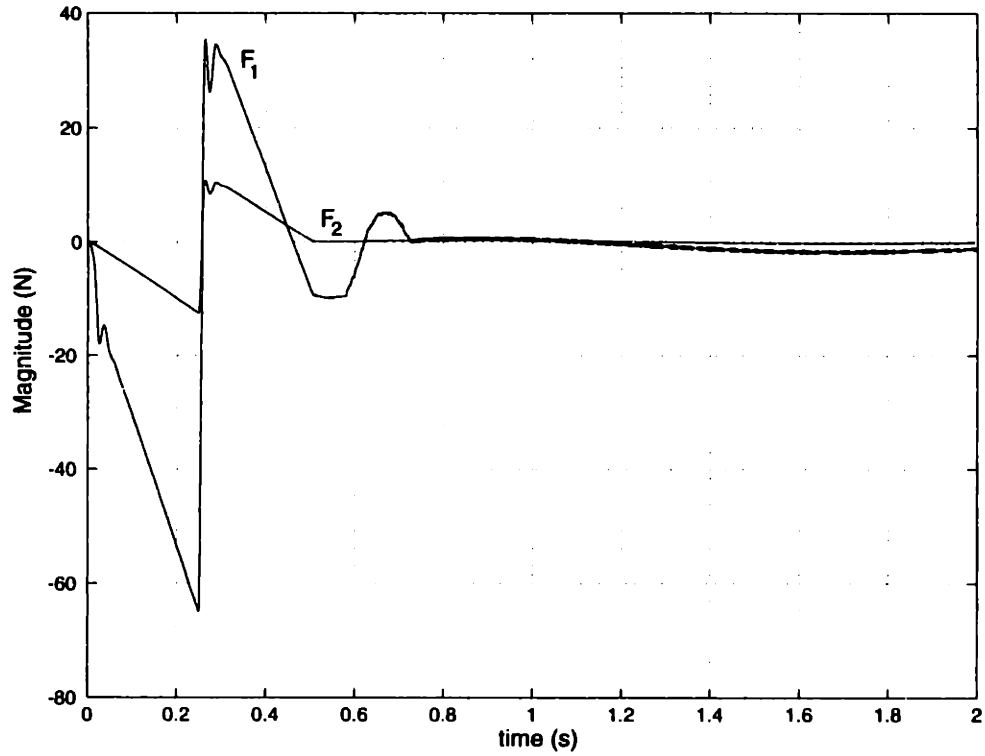

Figure 7-24: Actuation forces applied to the system

7.7 Digital equivalent of the continuous-time controller

Now that we have designed a continuous-time controller and checked its validity we can use emulation techniques to find its discrete time equivalent. Continuous time is the limiting case of discrete-time so we expect the performance of a discrete time controller to converge to that of its continuous-time equivalent as the sampling interval is reduced. The problem statement does not specify any lower bound on the sampling interval but an A/D or D/A conversion takes on average between 5 μs and 20 μs. The ones which were used on the magnetic bearing took around 10 μs. If we allow for some computation time, a sampling interval of $40 - 50\,\mu$s is quite feasible and adequate. Using these values and a sampling rate of 2KHz, the ratio of the time delay to the sampling interval is around 5. Ideally the smaller this number the better,

but a value of 5 is acceptable. The emulation technique which we shall adopt will be the Pole-Zero Mapping. This method uses the relation which maps the poles from one domain to the other, and applies it to the zeros as well. By the definition of the Z-transform, it is known that

$$z = e^{sT} \tag{7.37}$$

To map the controller gain, we choose to keep the same DC gain as in the continuous design so that equilibrium is not affected. The following table shows the equivalent discrete pole locations for the compensators.

	Continuous	Discrete
Lead zero	-95.2	0.954
Lead pole	-952.4	0.622
Lag zero	-150	0.928
Lag pole	0	1

The compensators are as follows :

$$G_{lead}(z) = \frac{8.153\,z - 7.775}{z - 0.6223} \tag{7.38}$$

$$G_{lag}(z) = \frac{2.651 \times 10^4\,z - 2.459 \times 10^4}{z - 1} \tag{7.39}$$

Using these compensators in the simulation the performance is entirely unchanged: the RMS for the continuous controller was 7.6759×10^{-9} m and that of the digital controller is 7.671×10^{-9} m. The plots were almost identical to the ones shown in the previous section and are therefore not repeated.

Chapter 8

Design of a controller for a single degree of freedom wafer stepper using LQG

8.1 Introduction

The previous chapter presented a control design based on a system representation which allowed the dynamics to be decoupled. The feature that accomplished this was the realization that the effect of springs and dampers is reduced as frequency increases. This is particularly true for frequencies above resonance; the latter thus supplies us with a convenient breakpoint which allows us to re-derive the system model making certain simplifying assumptions. Such simplifications were shown to be valid through simulations before any further use was made of the simplified model, and it must not be assumed that these simplifications are necessarily applicable to all mechanical systems. Decoupling the dynamics was important because classical control design is best suited to SISO systems because it relies heavily on inferences and analysis of visual aids such as Bode plots which only represent one input-output relation between parameters. In the case of a MIMO system, Bode plots of diagonal I/O terms (i.e. $y_1/u_1, y_2/u_2$, etc) as well as Bode plots of cross-terms (i.e. $y_1/u_2, y_2/u_1$, etc.), would

have to be analyzed simultaneously, rendering this approach rather inconvenient.

In order to be able to work with a MIMO system directly, it is more convenient to adopt state space methods whose matrix representations, although less intuitive, allow MIMO systems to be supported far more easily. In this regard, the procedure which is adopted in this chapter is the design of a controller using optimal control methods, also referred to as Linear Quadratic control (LQ). The rest of the chapter will cover the steps involved in designing an optimal controller and will point out the heuristic steps involved in its design. An LQR system is based upon feedback of all the system states, i.e. full state feedback, but this is not available to us, as the problem statement clearly specifies. An observer must therefore be included in the system design to estimate the states which are not directly accessible. The chosen observer is the Kalman filter also referred to as the optimal estimator because its design parallels that of the LQR. The combined LQR-Kalman Filter system is referred to as the Linear Quadratic Gaussian (LQG); the term Gaussian entering because the Kalman filter assumes a Gaussian model of noise for its rejection properties.

8.2 Discrete Plant

In the previous chapter, the controller was designed in continuous time and then mapped to discrete time by emulation techniques. This last step was important because in a real setting the controller will most likely be implemented on a digital computer. In this chapter we take another approach and carry out the controller design in discrete-time from the start. We will use a 2 kHz sampling frequency since we showed in the previous chapter that with this choice we will be able to capture all the resonances. Taking the ZOH of (7.14)-(7.17) we can re-write the plant in the discrete time form

$$\mathbf{q}(k+1) = \mathbf{\Phi}\,\mathbf{q}(k) + \mathbf{\Gamma}\,\mathbf{u}(k) \tag{8.1}$$

$$\mathbf{y}(k) = \mathbf{C}\,\mathbf{q} + \mathbf{D}\,\mathbf{u}(k) \tag{8.2}$$

where the matrices $\mathbf{\Phi}, \mathbf{\Gamma}, \mathbf{C}, \mathbf{D}$ are the new matrices defining the plant.

8.3 Properties of LQR

LQR design is another form of state feedback design in which the goal is to find a set of gains which will place the poles of the closed loop system in specific locations. However, unlike pole placement techniques in which the designer is faced with the dilemma of having to decide upon n (n being the order of the system) desired pole location, LQR techniques are attractive in that the procedures systematically set the poles at locations which are in a sense optimal. Optimal in this sense does not refer to global optimality, it refers instead to the optimal solution of minimizing the cost function, J. This represents a quadratic expression in the states (or alternatively the outputs since the two are linearly related for a linear system) and inputs normally of the form

$$J = \frac{1}{2} \sum_{k=0}^{N} \left[\mathbf{x}^T(k)\mathbf{Q}\mathbf{x}(k) + \mathbf{u}^T(k)\mathbf{R}\mathbf{u}(k) \right] \tag{8.3}$$

In the above equation N is the discrete time indicator, $\mathbf{x}$ is the state vector, $\mathbf{u}$ is the input vector, and the matrices $\mathbf{Q}$ and $\mathbf{R}$ are the state and input weighting matrices respectively. It is required that $\mathbf{Q}$ be at least positive semidefinite while as $\mathbf{R}$ be positive definite.

There are several important properties of LQR which are worth mentioning although for complete details refer to [20]. First of all, and most fundamental, it can be shown either through defining a Lyapunov function or other means that the gain matrix $\mathbf{K}$ derived from the steady state solution of the Algebraic Riccati Equation (ARE) yields an asymptotically stable system. It can also be shown that LQR design guarantees a certain level of stability robustness, most noticeably that the phase margin is at least $\pm 60°$. On a Nyquist diagram this is represented by having a unit circle centered around the -1 point and the plot of the controlled system will never enter that region. Another robustness guarantee is that the controller derived from the LQR procedure will remain stable if its gains are increased by any factor > 1 (this is referred to as having a positive gain margin of ∞) and still be stable if the

gains are reduced by a factor of up to $\frac{1}{2}$ (this is referred to as having a negative gain margin of $\frac{1}{2}$). In either of these cases, stability is maintained but performance guarantees are lost and performance is obviously suboptimal for the given cost function. For completeness we also note that being a pole placement technique, using gains to change pole locations does not affect the order of the system.

To carry out LQR design, several assumptions need to be satisfied. It is assumed that the entire state vector is available for feedback. For the wafer stepper design at hand, this is not the case since access is limited to the two outputs y_1 and y_2 which calls for an observer to be introduced in the system. Adding the observer requires that the plant be detectable otherwise states could not be reconstructed. The observer design will be discussed later in the chapter. Another requirement is that the plant be stabilizable which is a formal way of proving that the closed loop system poles can be placed arbitrarily in any location with the appropriate selection of gains, and the poles that cannot be placed are stable. Stabilizability is a subset of reachability and detectability is a subset of observability. Under these assumptions the LQR is the optimal input

$$\mathbf{u}(t) = -\mathbf{K}\,\mathbf{x} \tag{8.4}$$

which minimizes the cost function (8.3) subject to the dynamic constraints of the system (8.1).

In order to proceed with this design approach, it is necessary to show first that the plant is both detectable and stabilizable and hence the system is minimal. We start by proving reachability of the system. For the system defined by (8.1), complete state controllability can be proved by first finding the modal matrix, $\mathbf{S}$, such that

$$\mathbf{S}^{-1}\mathbf{A}\,\mathbf{S} = \mathbf{J_o}$$

where $\mathbf{J_o}$ is the Jordan canonical representation for the system. The eigenvalues for

the discrete plant are

$$\begin{pmatrix} 1 \\ 1 \\ 0.9996 + 0.0020824i \\ 0.9996\text{-}0.0020824i \\ 0.77757\text{+}0.54359i \\ 0.77757\text{-}0.54359i \\ 0.99966\text{+}0.0056896i \\ 0.99966\text{-}0.0056896i \end{pmatrix} \tag{8.5}$$

The complex Jordan form is a diagonal matrix of these eigenvalues, and the modal matrix is the matrix of the right eigenvectors of the system. Note that the system is very stiff so the modal matrix is close to singular. Having determined these, the system is state controllable if and only if

- no two Jordan blocks in $\mathbf{J_o}$ are associated with the same eigenvalues
- the elements of any row of $\mathbf{S^{-1}\Gamma}$ that correspond to the last row of each Jordan block are not all zero
- the elements of each row of $\mathbf{S^{-1}\Gamma}$ that correspond to distinct eigenvalues are not all zero.

The plant passes these three conditions and is therefore reachable. In order to test for observability we test for the following properties :

- no two Jordan blocks are associated with the same eigenvalue
- no columns of **CS** that correspond to the first row of each Jordan block consist of zero elements
- no columns of **CS** that correspond to distinct eigenvalues consist of zero elements.

Again we find that the plant passes these conditions.

8.4 Regulator Design

The order for the model-based controller with the observer will be twice the order of the plant alone since the state variables of both the plant, $\mathbf{x}$, and the observer, $\hat{\mathbf{x}}$, are required to describe the dynamics of the system. For an eighth order system with two outputs this means selecting 32 gains. The task can be reduced by invoking the Separation Principle which states that the closed loop dynamics matrix of the system with compensator and observer will be block upper triangular and thus the eigenvalues of the closed loop system will be those of the controller as well as those of the observer. This allows us to design the controller and compensator independently. In this section we shall focus on the controller design only and the goal is to form a setup whose general form is shown in Figure 8-1.

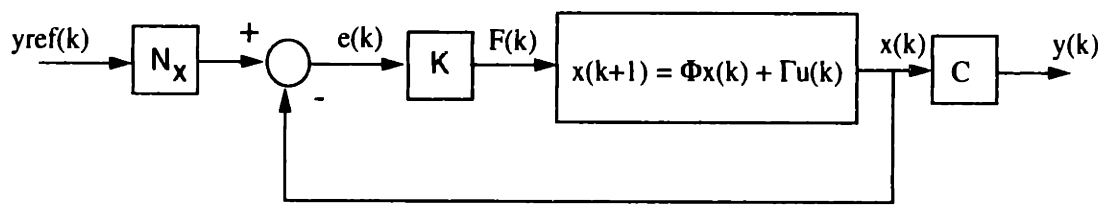

Figure 8-1: LQR setup

The matrix $\mathbf{N_x}$ is the State Command Matrix and the structure of the closed loop setup is the State Command Structure which derives its name from the fact that the reference input is the direct reference to the states (or a linear combination of the latter). There are two inputs to the plant and there are also two outputs; the case of equal number of inputs and outputs is the only case which has a unique solution. As was done in the previous chapter, it is known that

$$y_1 = -q_5 \qquad \text{and} \qquad y_2 = -q_3 - q_5 - q_7$$

and physical constraints require $y_1 \leq 200\,\mu$m. We therefore reference $q_5 = 0$, and we also reference $q_7 = 0$ since we can't, and don't want to, control the relative motion within the kinematic coupling. This coupling is very stiff which would require a high bandwidth controller as well as a large force input; but more importantly the

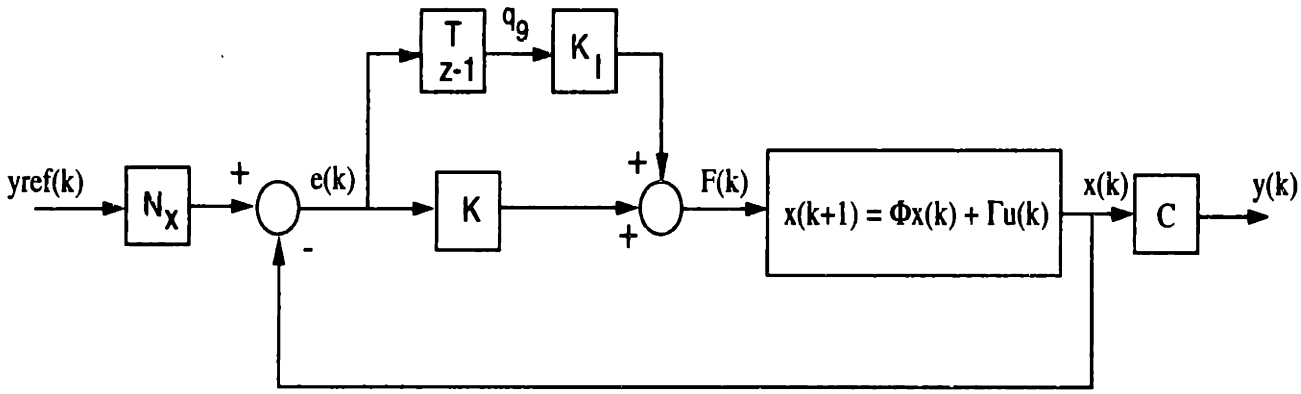

Figure 8-2: Block diagram for integral control with full-state feedback

actuators do not act directly on q_7. We therefore conclude that the desired reference is

$$y_2 = -q_3$$

and thus

$$\mathbf{N_x} = (0\ 0\ -1\ 0\ 0\ 0\ 0\ 0\ 0)^T \ .$$

Note that $\mathbf{N_x}$ commands nine states in anticipation of what is discussed below. Having decided on the control strategy to adopt, the choice is now on what control structure to use. There are several possibilities: simple state feedback structure, integral controller, feedforward controller, just to name a few. A simple structure would seem adequate for the task however in an attempt to increase performance of the system it was decided that an integrator should be used in the system to help drive q_3 to the desired trajectory as fast and efficiently as possible. With this state augmentation the dynamics becomes ninth order and the compensated setup is shown in Figure 8-2. Referring to this figure the new state is defined as

$$q_9(k+1) = q_9(k) + T(q_{3ref} - q_3) \quad . \tag{8.6}$$

where T is the sampling interval of T=0.0005 s. This requires the system matrices for the integral model to be redefined as in (8.7)-(8.10).

$$\mathbf{\Phi_I} = \begin{pmatrix} \mathbf{\Phi} & 0 \\ \mathbf{\Psi} & 1 \end{pmatrix} \tag{8.7}$$

where $\mathbf{\Psi} = \begin{bmatrix} 0 & 0 & -T & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$ and $\mathbf{\Phi}$ is the matrix for the unaugmented plant.

$$\mathbf{\Gamma_I} = \begin{pmatrix} \mathbf{\Gamma} \\ 0 \end{pmatrix} \tag{8.8}$$

$$\mathbf{C_I} = (\mathbf{C} \;\; \mathbf{0}) \tag{8.9}$$

$$\mathbf{D_I} = \mathbf{D} \tag{8.10}$$

The feedback control law for the regulator setup is

$$\mathbf{u}(k) = -\begin{bmatrix} \mathbf{K} & -\mathbf{K_I} \end{bmatrix} \begin{pmatrix} \mathbf{q}(k) \\ q_9(k) \end{pmatrix} \tag{8.11}$$

which is composed of two parts : $\mathbf{u}_{physical\ states} = -\mathbf{K}\,\mathbf{q}$ and $\mathbf{u}_{integral\ state} = \mathbf{K_I}\,q_9$ where $\mathbf{u}$ is the vector of inputs (forces) to the plant and the subscripts are self explanatory. There are two inputs to command, F_1 and F_2, therefore $\mathbf{K}$ will be a 2×8 matrix and $\mathbf{K_I}$ will be a 2×1 vector.

A fundamental distinction must be made between the integral plant model and the physical plant model: the actual physical plant does *not* contain the integral state. The latter is added in a control scheme aimed at improving performance but is a purely fictitious state which exists only inside the computer.

The integral model has now been put in the general form of (8.1) which can be used in LQR design. In order to minimize (8.3) it remains to chose the weighting matrices $\mathbf{Q}$ and $\mathbf{R}$. The latter are only weakly connected to performance and experimentation with different sets of gains showed that the system is extremely stiff with respect to changes in the weighting matrices. More than the absolute values of the weighting gains, the performance was more a function of the ratio between the entries. There

is no universal way of directly correlating the weight matrices to the performance of the system so LQR inevitably leads to an iterative procedure guided both by trial and error, and by intuition. Weighting coefficients are constants, (although in the literature there are LQR techniques based on frequency weighted cost functions for which this is not true [1, 20]), thus in minimizing the quadratic expression the optimizing algorithm will minimize the corresponding input or state. Intuition leads us to think, and this is indeed the case, that the larger the weights the more that particular combination of states or inputs is penalized. Therefore to decrease the excursion of a specific state or input we would increase the gain associated with its quadratic form. It is this basic notion which guided the iterative procedure of coming up with appropriate weight matrices.

As a starting point for the selection of the weights, we adopt the guidelines commonly referred to as the Bryson Rules which suggest that the weight matrices should be chosen to be diagonal with entries of

$$\frac{1}{(\text{max. allowable deviation})^2}$$

Using symmetric weighting matrices and a quadratic cost function of the states, (not the outputs), would leave us with 39 gains to select. Quite a formidable task since gain selection is fundamentally a heuristic approach. Furthermore, while as we have a physical intuition for the effects of diagonal terms in the weighting matrices, the effect of off-diagonal terms is much harder to get a feeling for. We will restrict the choice of weighting matrices to diagonal matrices so that the entries will be directly related to the penalization of the corresponding state or input. This step reduces the number of weights to chose to 10. A way of reducing the choice of weights even more is to use an output transformation to transform a 2×2 weight matrix for the outputs, $\bar{\mathbf{Q}}$, into a 9×9 weight matrix for the states.

$$\mathbf{Q} = \mathbf{C_I}^T \bar{\mathbf{Q}} \mathbf{C_I} \tag{8.12}$$

where $\mathbf{C_I}$ is the output matrix (8.9). Choice of the gains in this fashion is attractive

because the weights enter the cost function directly through the important outputs. This is good since the performance requirement which we're trying to meet is based on one of the system outputs. Use of the Bryson Rules for this case would suggest the following weighting matrices.

$$\bar{\mathbf{Q}} = \begin{pmatrix} \frac{1}{(10^{-9})^2} & 0 \\ 0 & \frac{1}{(10^{-9})^2} \end{pmatrix}$$

$$\mathbf{R} = \begin{pmatrix} \frac{1}{200^2} & 0 \\ 0 & \frac{1}{50^2} \end{pmatrix}$$

Using these matrices lead to poor performance but they were used as the starting point for the iterative procedure. Numerous unsuccessful iterations finally convinced us that the approach had to be altered somewhat because it did not take into account certain physical considerations of the setup. First, in view of the large (with respect to the specifications) quantization noise of the LVDT we cannot rely entirely on our estimate of q_3 or q_5. Also, from the point of view of control effort, penalizing q_3 very much means asking for faster performance and hence greater control effort. This is not desired because q_3 is directly related to motion of the heavy coarse stage and high bandwidth control of it requires much more effort than high bandwidth control of the smaller masses. These considerations are not taken into account if (8.12) is used; in fact use of this equation would penalize q_3 the most because this state is common to both y_1 and y_2 and it thus considered the most important state. It was therefore decided that a diagonal weight matrix for the states should be used so that each entry could be independently adjusted after each iteration. The following weight matrices were found to lead to good performance, and we note that they are substantially

different from the ones suggested by the Bryson Rules.

$$\mathbf{Q} = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{(10^{-6})^2} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{(10^{-7})^2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{(10^{-9})^2} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{1}{(10^{-10})^2} \end{pmatrix} \tag{8.13}$$

$$\mathbf{R} = \begin{pmatrix} 2 & 0 \\ 0 & 1.2 \end{pmatrix} \tag{8.14}$$

Matrix $\mathbf{Q}$ penalizes q_3 by $(10^6)^2$, q_5 by $(10^7)^2$, q_7 by $(10^9)^2$ and q_9 by $(10^{10})^2$. The integral state has been penalized the most since it is a non-physical state so we don't need to worry about it getting too high (provided, of course, that it remains within the working limits of the computer). State q_3 is weighed the least because it is directly related to friction: if we penalized this state too much we are asking for the controller to track the reference on q_3 very closely and in the presence of friction this will lead to chattering and very high control activity. The weight on q_5 was also chosen to be rather loose because of the large quantization level, so penalizing this state more would lead to chattering introduced by quantization. Finally, the weighting of q_7 was made more stringent because the kinematic coupling is stiff and only allows nanometer level excursions in q_7. We note again that while these were the guidelines we used in searching for appropriate weights, many iterations were required to come up with the values presented above.

Attempts were made to plot the locus of the closed loop poles as a function of the weighting matrices with the idea of checking the asymptotic behaviours of LQR. The idea was abandoned because the scaling of the system matrices led to singularity problems owing to the finite precision of the computer.

8.5 Estimator design

It was previously shown that the plant was observable so in this section we focus on the design of an observer. The natural choice is to use a Kalman Filter (KF) whose derivation is the dual of the LQR procedure, in that the optimal estimator gain matrix is the outcome of minimizing another cost function. The overall structure of a KF is the same as that of a current estimator, (8.15), so Kalman filtering boils down to a method of selecting the gains of the current estimator, $\hat{x}$.

$$\hat{\mathbf{x}}(k) = \bar{\mathbf{x}}(k) + \mathbf{K_e}\left[\mathbf{y}(k) - \bar{\mathbf{y}}(\mathbf{k})\right] \tag{8.15}$$

where $\bar{\mathbf{x}}$ is the prediction estimate from the previous time interval.

$$\bar{\mathbf{x}}(k+1) = \mathbf{\Phi}\hat{\mathbf{x}}(k) + \mathbf{\Gamma}\mathbf{u}(k) \tag{8.16}$$

We can rewrite the above equation grouping terms together as follows

$$\hat{\mathbf{x}}(k+1) = \left[\mathbf{\Phi} - \mathbf{K_e}\,\mathbf{C}\,\mathbf{\Phi}\right]\hat{\mathbf{x}}(k) + \left[\mathbf{\Gamma} - \mathbf{K_e}\,\mathbf{C}\,\mathbf{\Gamma}\right]\mathbf{u}(k) + \mathbf{K_e}\,\mathbf{y}(k+1) \tag{8.17}$$

The plant is assumed to obey a stochastic difference equation having the general form

$$\mathbf{x}(k+1) = \mathbf{\Phi}\,\mathbf{x}(k) + \mathbf{\Gamma}\,\mathbf{u}(k) + \mathbf{\Lambda}\,\mathbf{w}(k) \tag{8.18}$$

where $\mathbf{w}$ is the vector of process noise disturbance entering the plant and is assumed to have certain statistical properties corresponding to a stationary continuous time white Gaussian noise with zero mean. $\mathbf{\Lambda}$ is the matrix showing how this disturbance enters the dynamics.

In the classical Kalman filter formulation, it is assumed that measurements can only be made in the presence of additive white noise

$$\mathbf{y}(k) = \mathbf{C}\,\mathbf{x}(k) + \mathbf{v}(k) \tag{8.19}$$

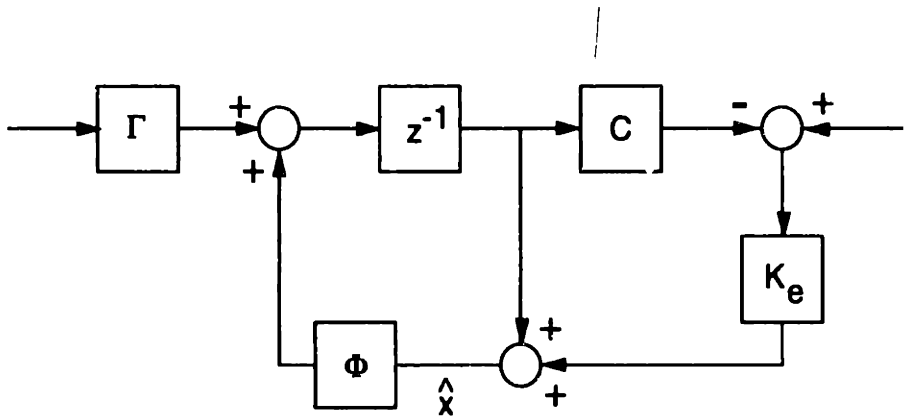

Figure 8-3: Block diagram for a current estimator

where $\mathbf{v}(k)$ is the process measurement noise assumed to be a continuous time white Gaussian random process independent of $\mathbf{w}$, with zero mean.

For the wafer stepper setup we identify the disturbances as the ground vibrations, friction force and quantization of the sensors, which we can divide into driving noise and measurement noise.

$$\mathbf{w}(k) = \begin{pmatrix} \text{Friction} \\ \text{Floor vibr.} \end{pmatrix}$$

$$\mathbf{v}(k) = \begin{pmatrix} \text{Quantization in } y_1 \\ \text{Quantization in } y_2 \end{pmatrix}$$

From the setup we know that friction enters the same way as $-F_1$, while as ground vibration enters through $\dot{q}_2$.

8.6 Current estimator

The current estimator uses measurement up to time $t = k$ to produce a state estimate at the same sampling interval. The block diagram of this estimator is shown in Figure 8-3. A useful fact about the current estimator can be proved by defining the state error vector as

$$\tilde{\mathbf{x}} = \mathbf{x} - \hat{\mathbf{x}}$$

This definition then leads to the state error dynamics for the plant with noise

$$\tilde{\mathbf{x}}(k+1) = [\mathbf{\Phi} - \mathbf{K_e C \Phi}]\tilde{\mathbf{x}}(k) + [\mathbf{\Lambda} - \mathbf{K_e C \Lambda}]\,\mathbf{w}(k) - \mathbf{K_e}\,\mathbf{v}(k+1)$$

which is a stable dynamic system if

$$|\lambda_i[\mathbf{\Phi} - \mathbf{K_e C \Phi}]| < 1$$

where λ_i is the i^{th} eigenvalue of $\mathbf{\Phi} - \mathbf{K_e C \Phi}$. It can be shown that if the system defined by ($\mathbf{\Phi}$,$\mathbf{C}$) is detectable then it is guaranteed that there is at least one filter gain matrix $\mathbf{K_e}$ such that the filter is stable.

The KF enjoys most of the same properties of as LQR, foremost that the procedure guarantees a stable filter. However, while as both LQR and KF have robustness properties if each is analyzed independently the combined system, LQG, does not have robustness guarantees; the only guarantee that can be made is stability. During the iteration procedure, several LQG designs were tried which led to unstable closed loop systems, but the instabilities were due to numerical limitations of the computer.

The weighting matrices for the estimator cost function are the covariance matrices of the process and measurement noise. Under the assumption that the process noises are stationary and hence arise from a time invariant parent distribution, the mean of the noises for random processes $x(t)$ and $y(t)$ are μ_x and μ_y both of which are independent of time.

For arbitrary means of the noises, the covariance functions are related to the autocorrelation functions by the following equations [17].

$$C_{xx} = R_{xx}(\tau) - \mu_x^2 \tag{8.20}$$

$$C_{yy} = R_{yy}(\tau) - \mu_y^2$$

$$C_{xy} = R_{xy}(\tau) - \mu_x \mu_y$$

We associate the random variable x with friction and y with ground vibrations. The latter is a sum of two sinusoids each centered about zero therefore it has a zero mean value, $\mu_y = 0$. From the above relations and from the definition of autocorrelation

we redefine $y_1 = y(t)$ and $y_2 = y(t+\tau)$

$$C_{yy} = R_{yy}(\tau) = \int^{\infty} \int_{-\infty} y_1\, y_2\, p(y_1, y_2)\, dy_1\, dy_2 \tag{8.21}$$

For a random process of ground vibration having the form

$$y(t) = a_1\phi_1(t) + a_2\phi_2(t)$$

where $a_1 = e_1$ and $a_2 = e_2$ are constant amplitudes and $\phi_1 = \cos(\omega_1 t + \theta)$, $\phi_2 = \cos(\omega_2 t + \theta)$ are themselves independent random processes in the variable θ

$$R_{yy} = a_1^2 R_{\phi_1\phi_1} + a_1\, a_2[R_{\phi_1\phi_2} + R_{\phi_2\phi_1}] + a_2^2\, R_{\phi_2\phi_2}$$

however, the process is deterministic because ground vibrations are assumed to be exactly known and thus the pdf of θ is non-zero only at the origin :

$$p(\theta) = \begin{cases} 1 & \theta = 0 \\ 0 & \text{elsewhere} \end{cases} \tag{8.22}$$

Inserting (8.22) into (8.21) gives the covariance of the ground vibrations as

$$R_{yy} = e_1^2 \cos(\omega_1\tau) + e_1\, e_2 \cos(\omega_1\tau) \cos(\omega_2\tau) + e_2^2 \cos(\omega_2\tau) \tag{8.23}$$

We can find an upper bound for (8.23) by utilizing

$$|R_{\phi_1\phi_2}(\tau)|^2 \leq R_{\phi_1\phi_1}(0)\, R_{\phi_2\phi_2}(0) \tag{8.24}$$

$$|R_{\phi_1\phi_1}(\tau)| \leq R_{\phi_1\phi_1}(0) \tag{8.25}$$

$$|R_{\phi_2\phi_2}(\tau)| \leq R_{\phi_2\phi_2}(0) \tag{8.26}$$

Therefore, inserting (8.24)-(8.26) in (8.23) the ground vibration covariance becomes

$$R_{yy} \leq e_1^2 + e_1 e_2 + e_2^2$$

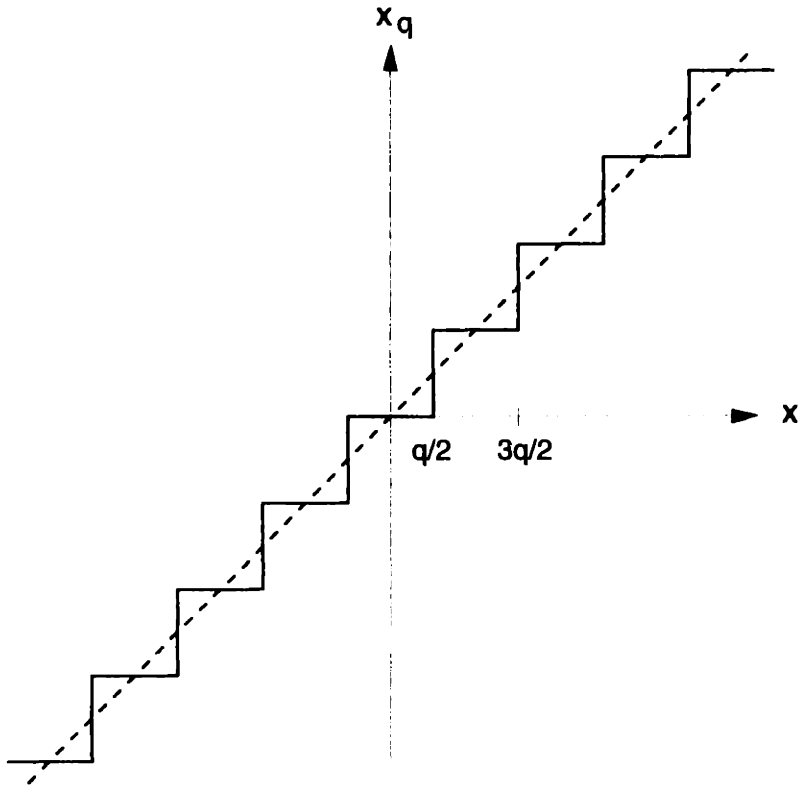

Figure 8-4: Plot of variable versus quantized value

which we can simplify even further by noting that the problem statement specifies $e_2 \gg e_1$. Therefore as a starting point in the design of the KF we can say

$$R_{yy} \approx e_2^2 = 10^{-8} \tag{8.27}$$

For friction we cannot perform a similar analysis because we cannot say a priori how the friction is going to oscillate between ± 10 N. We can, nonetheless say that $R_{xx} \propto c^2$ where c is the magnitude of the Coulomb friction. This leads to an initial trial value of

$$R_{xx} = (10)^2 \tag{8.28}$$

Having calculated starting estimates for driving noise covariance we now look for estimates of the measurement noise covariance. Measurement noise is taken as deriving solely from quantization of the sensors and not from external interference. The plot of the actual value of a variable versus its quantized value, x_q, is shown in

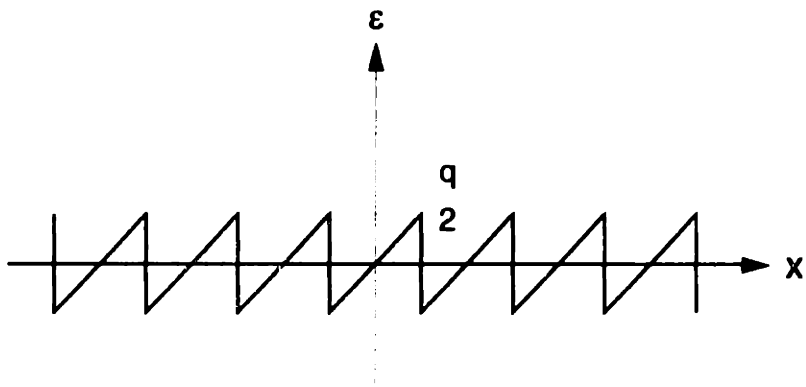

Figure 8-5: Round-off error

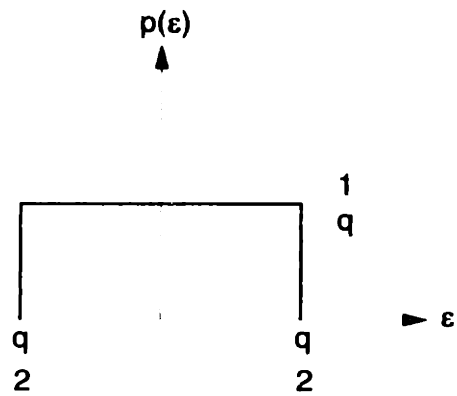

Figure 8-6: Uniform pdf

Figure 8-4 while Figure 8-5 shows the round-off error, ε, of the quantization.

In both plots q is the quantization step. If we imagine a random variable $x(k)$ that takes on values in a scattered way at successive sampling instants then it is reasonable to suppose that the errors will be scattered over the range $-q/2 \rightarrow q/2$. Furthermore, there is no reason to believe that the error has a greater probability of taking on certain values rather than other values and since the signal will jump several quantization steps in a sampling interval there is reason to assume that the signal is uncorrelated. It therefore seems reasonable to say that the error is equally likely to be anywhere in the range $-q/2 \leq \varepsilon \leq q/2$ and that its pdf is hence a uniform distribution as shown in Figure 8-6. The covariance of a uniform, zero mean distribution is easily found to be

$$C_{\varepsilon\varepsilon}(k) = \frac{q^2}{12} .$$

Knowing that the quantization step of the LVDT is 25 nm and that of the interferometer is 2.5 nm we get the following covariances.

$$C_{w_1 w_1} = 5.2 \times 10^{-17} \tag{8.29}$$

$$C_{w_2 w_2} = 5.2 \times 10^{-19} \tag{8.30}$$

We now have all the components to design the estimator: use the following matrices to find $\mathbf{K_e}$ and use this (8.17).

$$\mathbf{\Phi_e} = \mathbf{\Phi} \qquad (not\ \mathbf{\Phi_I})$$

$$\mathbf{\Lambda} = \begin{pmatrix} 2.5 \times 10^{-10} & -1.25 \times 10^{-7} \\ 1 \times 10^{-6} & -5 \times 10^{-4} \\ -2.75 \times 10^{-9} & 3.3 \times 10^{-11} \\ -1.1 \times 10^{-5} & -2 \times 10^{-7} \\ 2.5 \times 10^{-9} & 0 \\ 1 \times 10^{-5} & 0 \\ 6 \times 10^{-13} & 0 \\ 3.6 \times 10^{-9} & 0 \end{pmatrix} \tag{8.31}$$

$$\mathbf{C_e} = \mathbf{C}$$

Using the weighting matrices derived in this section together with (8.13) and (8.14) gave a reasonable staring performance: the RMS was 3.02×10^{-7} however actuator one saturated since very fast excursions of 500 N were required. Adjusting the entries in the covariance matrices and simulating the system we were able to see the effects of parameter changes on the performance, and used this to guide the selection of the weighting matrices. It was noticed that entry $\mathbf{Q_e}(1,1)$, where the matrix $\mathbf{Q_e}$ is the driving noise covariance matrix, had the most impact on the estimate of q_3 as well as the excursions of q_3 and on the control effort F_1. This fact is intuitive because this particular entry in the matrix is associated with how the mathematics of the minimization process deals with the friction in the system. Entry $\mathbf{Q_e}(2,2)$ was also

found to have a very large impact on the forces. We note however that too large of a ratio $[\mathbf{Q_e}(1,1)]/[\mathbf{Q_e}(2,2)]$ led to an unstable closed loop system. As was mentioned before this was due not to failure of LQG to provide a stable system, rather to numerical problems owing to the numerical stiffness of the system and finite precision of the computer. Similar instability resulted if the ratio $[\mathbf{R_e}(1,1)]/[\mathbf{R_e}(2,2)]$ was too high, where $\mathbf{R_e}$ is the measurement noise covariance matrix. On the whole, increasing all the parameters from their initial estimates was seen to lead to better performance, and good results were obtained with the following set of weighting matrices.

$$\mathbf{Q_e} = \begin{pmatrix} 100 & 0 \\ 0 & 1 \end{pmatrix} \tag{8.32}$$

$$\mathbf{R_e} = \begin{pmatrix} 5.2 \times 10^{-13} & 0 \\ 0 & 5.2 \times 10^{-15} \end{pmatrix} \tag{8.33}$$

8.7 Simulation Results

In this section we present the results of the using matrices (8.13), (8.14), (8.32), (8.33) in the LQG procedure. Supplying a reference trajectory for state q_3 gave results which were very similar to the ones presented in this section. However, slightly better performance was found to be obtainable if we supplied a reference trajectory for both q_3 and q_4 where the latter is the derivative of the former. The controller design technique does not change at all, and the gain matrices $\mathbf{K}$ and $\mathbf{K_I}$ are obtained in the exact same way as was described in the previous sections; the same goes for the observer gains. We note also that introduction of a reference trajectory does not alter the poles of the system it only changes the closed loop zeros [11].

Figure 8-7 shows how the system was implemented in Simulink. The RMS value for this design was 7.41×10^{-9} m which amply met the specifications requested. Simulations also showed that the result was quite insensitive with respect to changes in the covariance matrices, however it was quite susceptible to the controller weighting matrices.

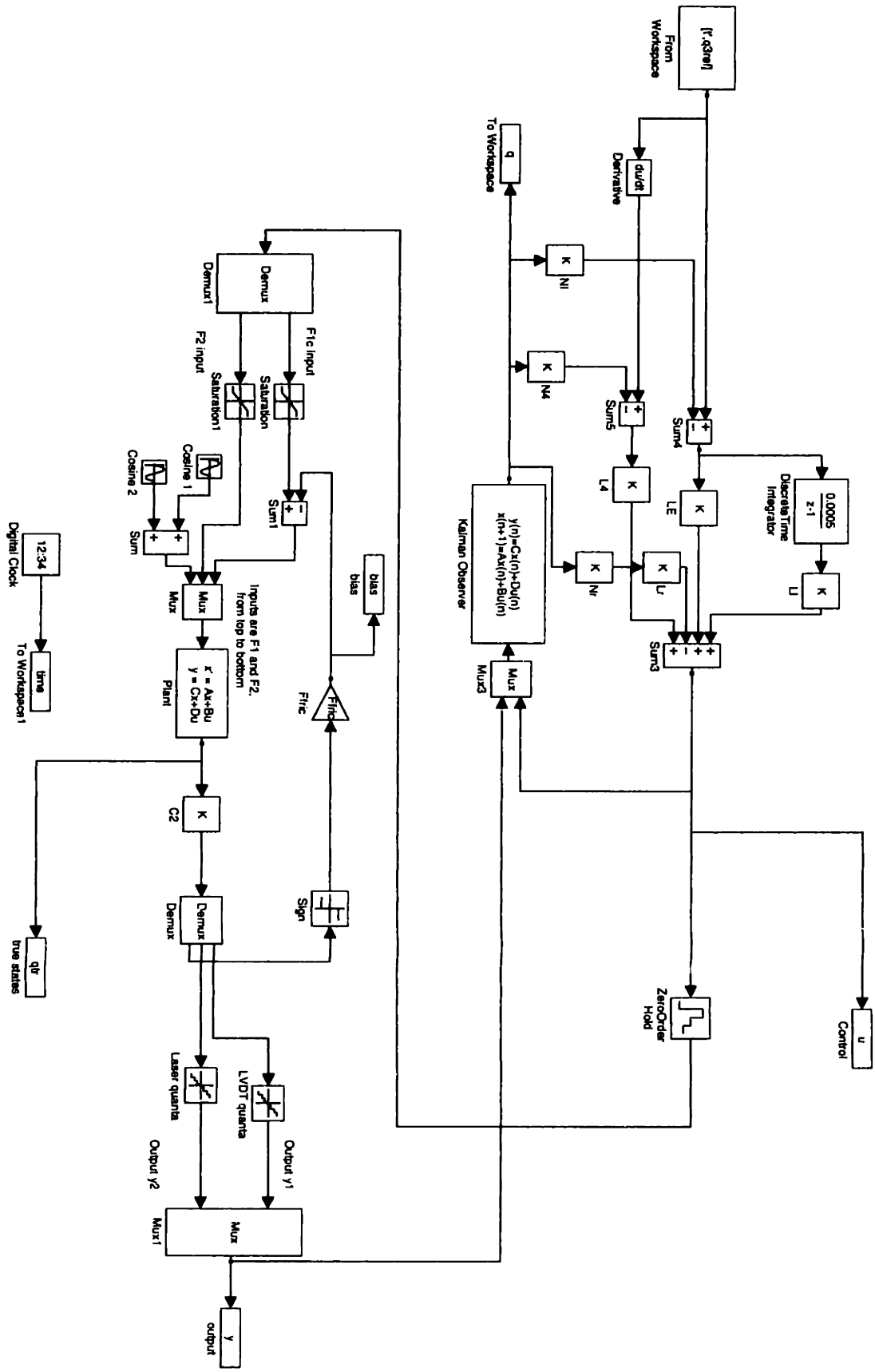

Figure 8-7: Simulink block diagram used to simulate the design.

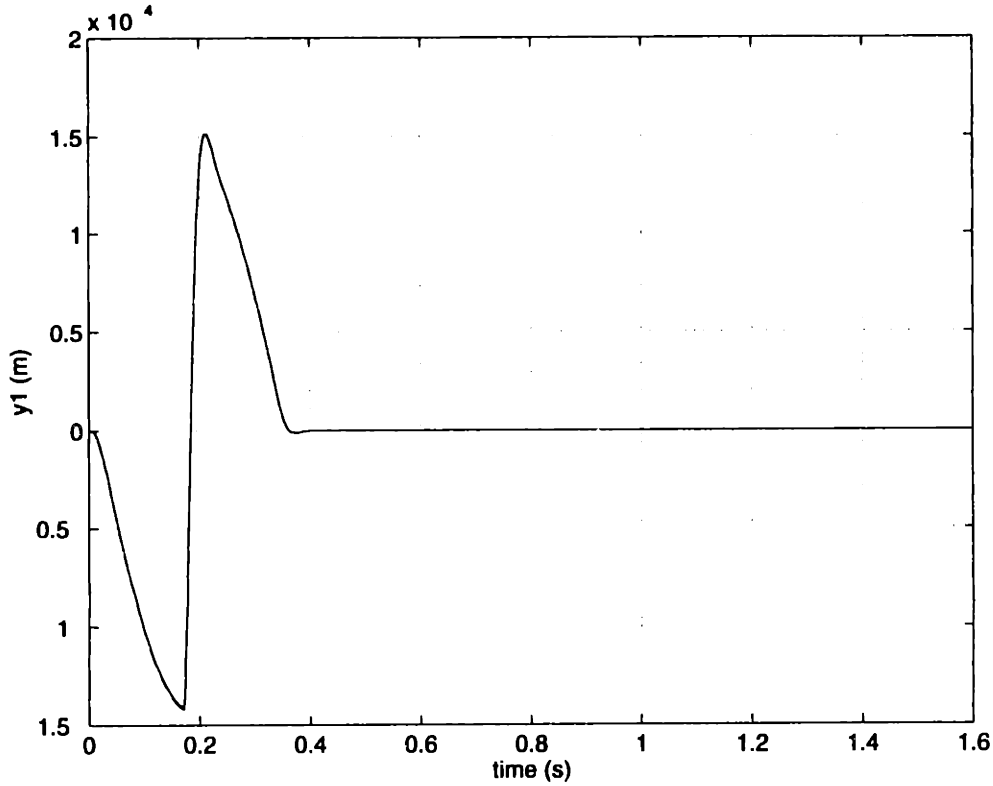

Figure 8-8: Response of output y_1.

Figures 8-8 and 8-9 show the outputs, with y_2 being the output of interest. The physical constraints on movement of the flexures holding the fine stage are met. Owing to the quantization of the LVDT, simulation showed that if in the matrix (8.13) the weight given to q_5 was increased, it would lead to the control signal becoming far more oscillatory and hence necessitating a much higher bandwidth controller. This is intuitively obvious because by penalizing q_5 more we are requesting that it be controlled more stringently which is not possible because the quantization steps are so large.

Figures 8-10 and 8-12 show that the actuators are well below their saturation levels. The bandwidth required by actuator one would be something to look more into depending on the physical hardware available, while as use of a voice coil setup like the one used for the magnetic bearing presented in this thesis can be shown to have more than enough slew rate (bandwidth) for this design. Earlier we noted the effect of increasing the weight of q_5 in (8.13). We show this in Figure 8-11 where

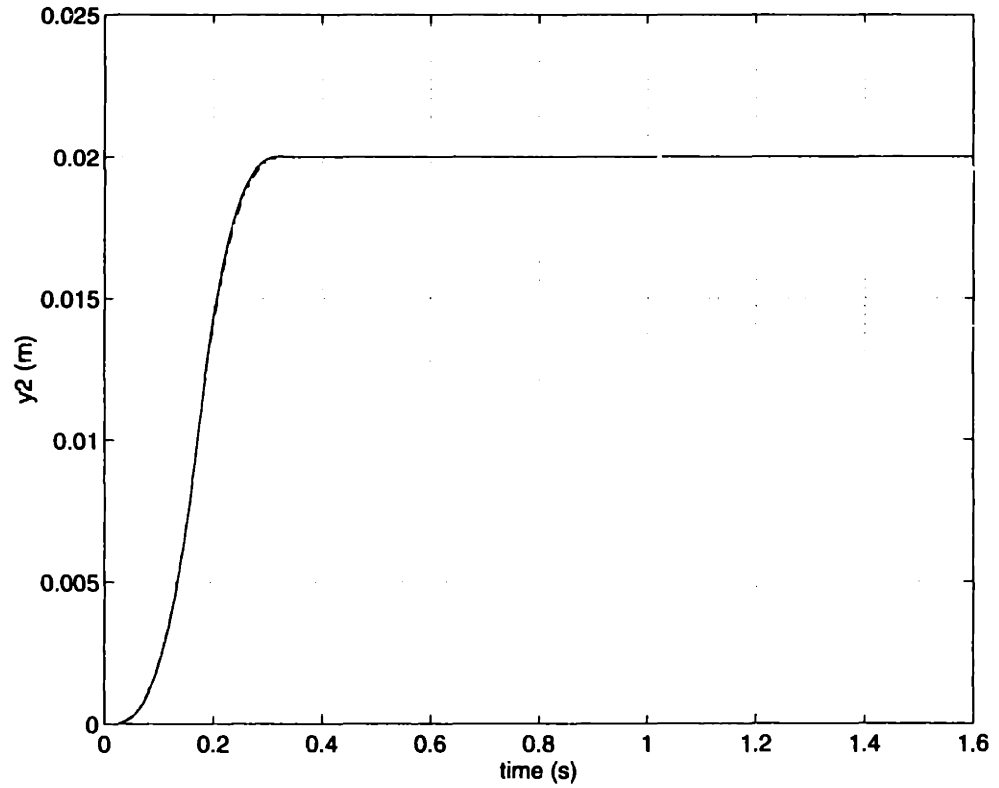

Figure 8-9: Referenced (dashed) and actual (solid) response of output y_2.

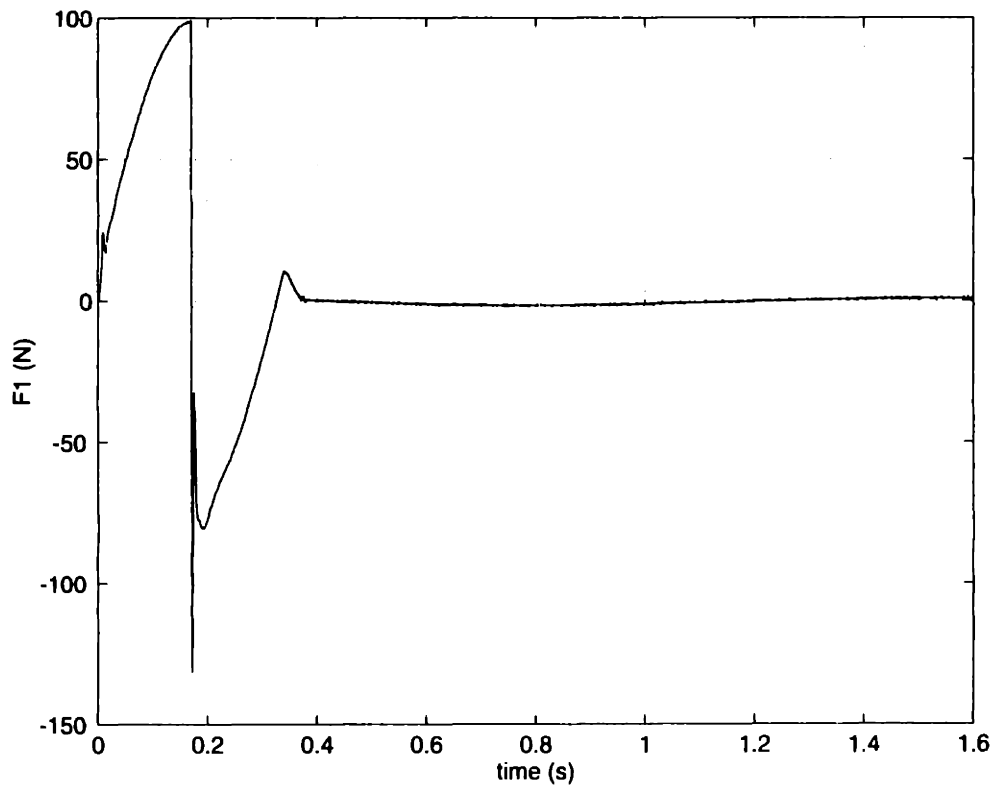

Figure 8-10: Required actuation force F_1.

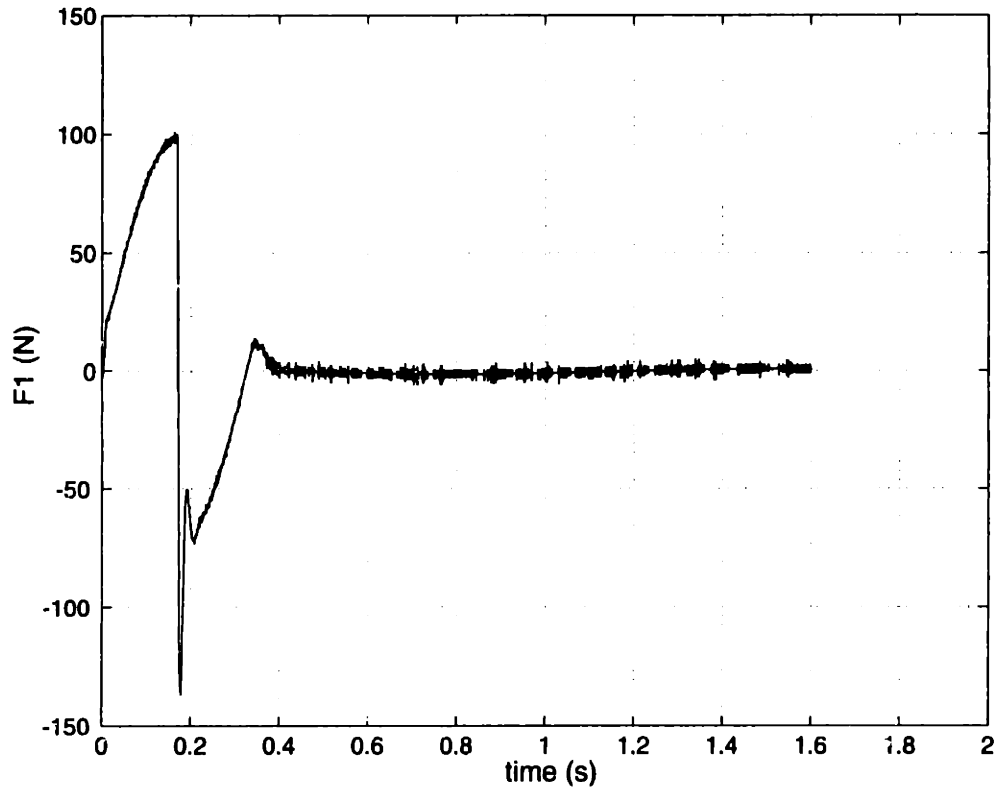

Figure 8-11: Required force F_1 if weight of q_5 is increased to $1/(10^{-9})^2$.

the observer used in exactly the same, but a larger weight of $1/(10^{-9})^2$ is used in (8.13) (all other weights remain the same) to calculate the controller gains. This plot highlights the excessive bandwidth required.

Figures 8-14 - 8-17 show the estimation errors of how the Kalman Filter estimates the states. The last plot, Figure 8-18 plots the Coulomb friction which after a transient of ≈ 0.35 s, (0.34 s being the time which the reference trajectory takes to reach the desired displacement value of 0.02 m), starts switching at virtually every sampling interval. A noticeable phenomenon brought out by these plots is the rapid convergence of the state estimates at the onset of the rapidly changing friction. Note how in the initial transient when friction is not oscillating and is fixed at +10 N, the KF converges to estimates which have a constant bias. From Figures 8-14 and 8-15 q_4 is estimated with a constant bias of -4.5×10^{-4} m/s and q_6 with a bias of 4×10^{-4} m/s both from the actual values of the states. Similarly from Figures 8-16 and 8-17 the estimate of q_5 converges with an error of 7.2×10^{-7} m/s, and for q_3 the error is a

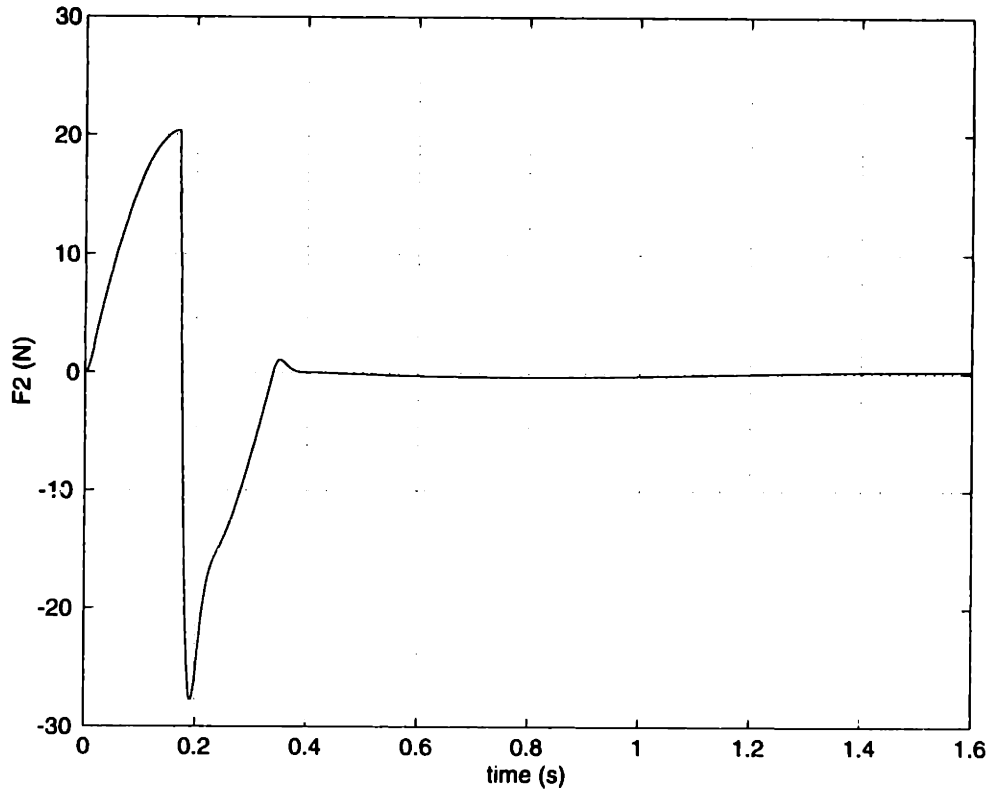

Figure 8-12: Required actuation force F_2.

constant -6.8×10^{-7} m/s. However at the onset of oscillating friction the estimates rapidly converge to the real values of the states. The explanation for the convergence to the wrong estimates lies in some of the assumption of the Kalman Filter not being satisfied. The derivation assumes zero mean white Gaussian noise with a positive definite covariance matrix, $\mathbf{Q_e}$. From (8.21) we find that the mean and autocorrelation for constant friction of +10 N are $\mu_x = 10$ N and $R_{xx} = 100\text{m}^2$ respectively, which shows that the zero mean assumption is being violated. Furthermore, inserting these values in (8.20) shows that $C_{xx} = 0$ which makes $\mathbf{Q_e}$ positive semi-definite. However when friction starts oscillating the mean value of the square signal gets closer and closer to zero with time, and the covariance becomes $\propto c^2$ which is what we assumed in our design.

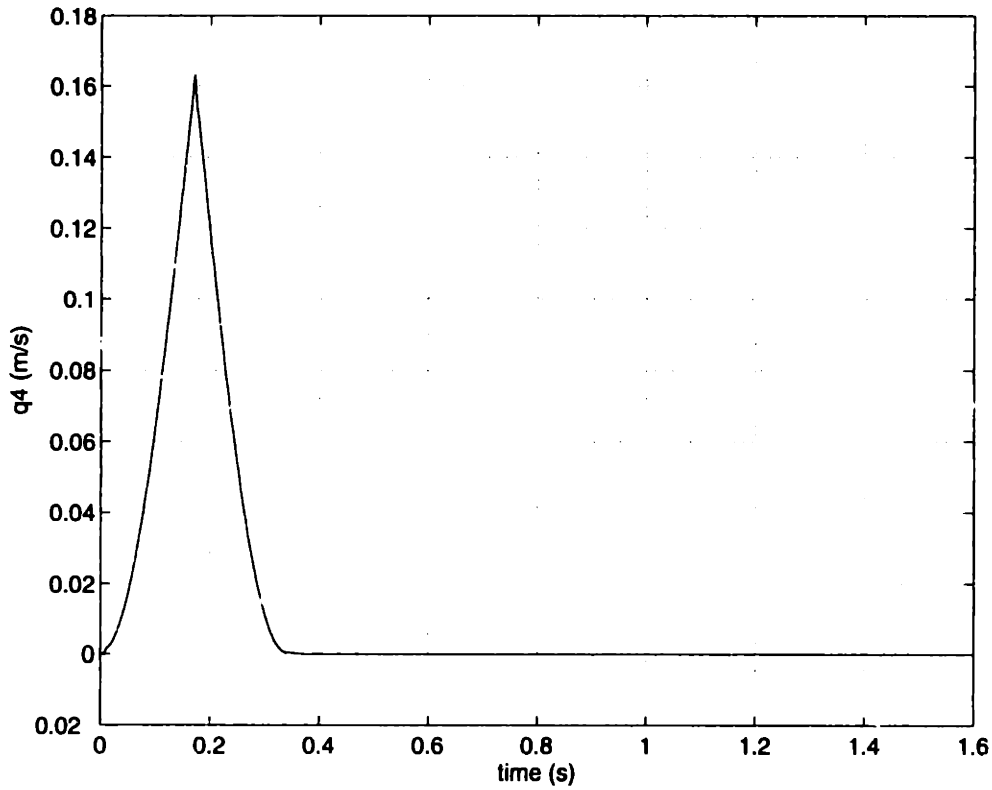

Figure 8-13: Referenced (dashed) and simulated (solid) plot of q_4.

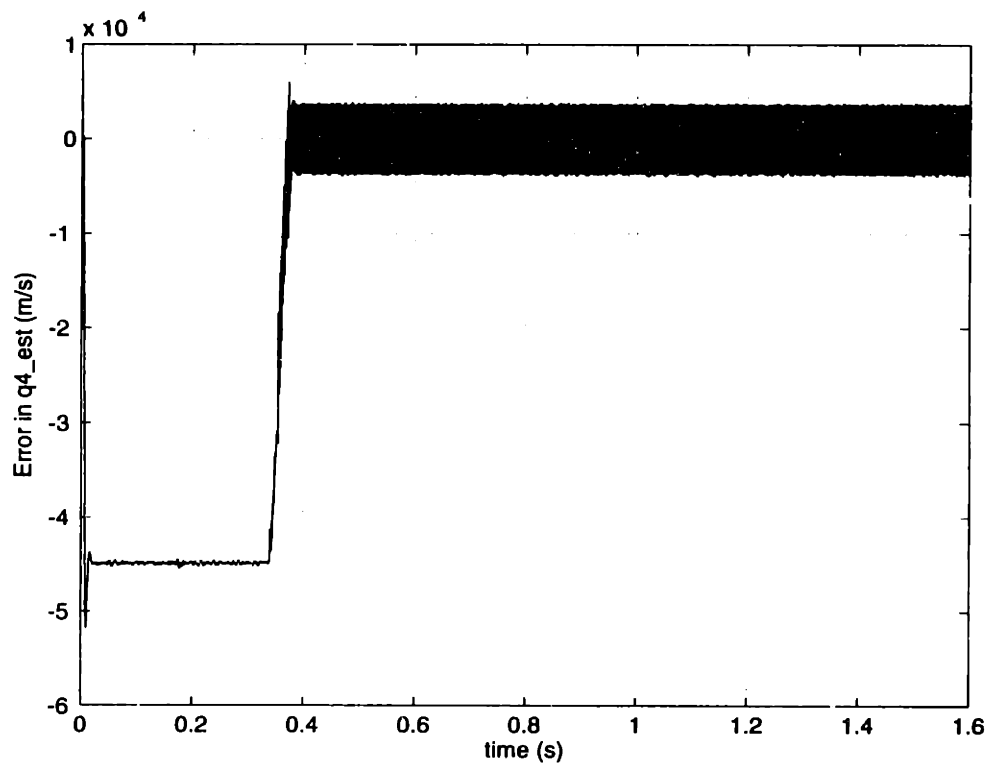

Figure 8-14: State estimate error for q_4.

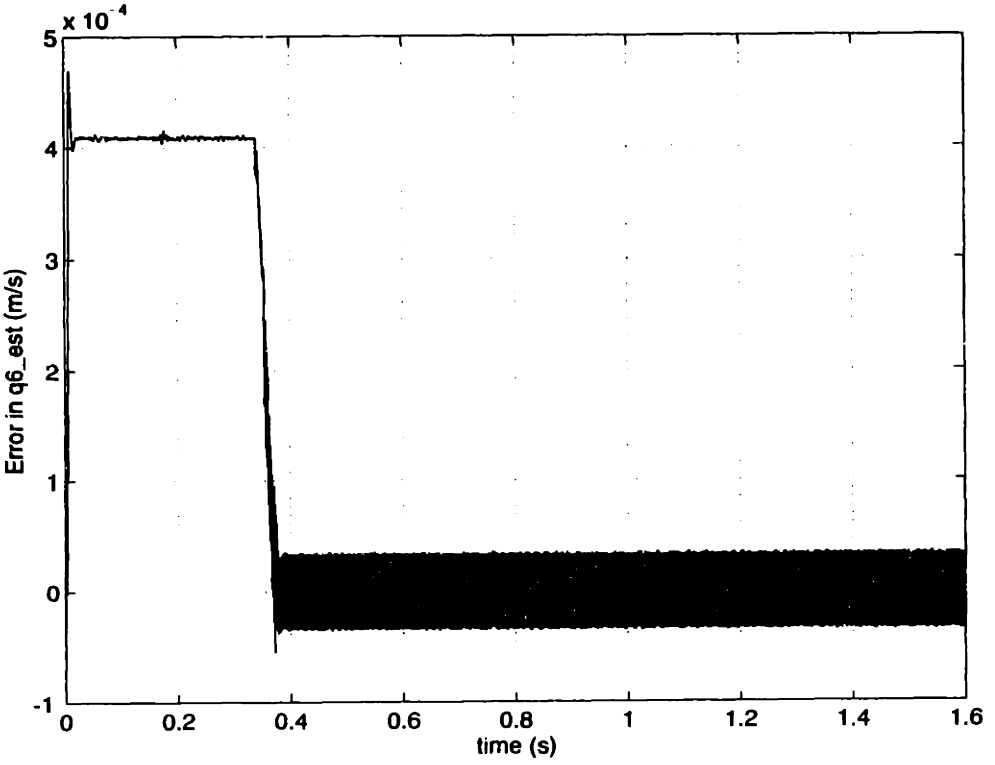

Figure 8-15: State estimate error of q_6.

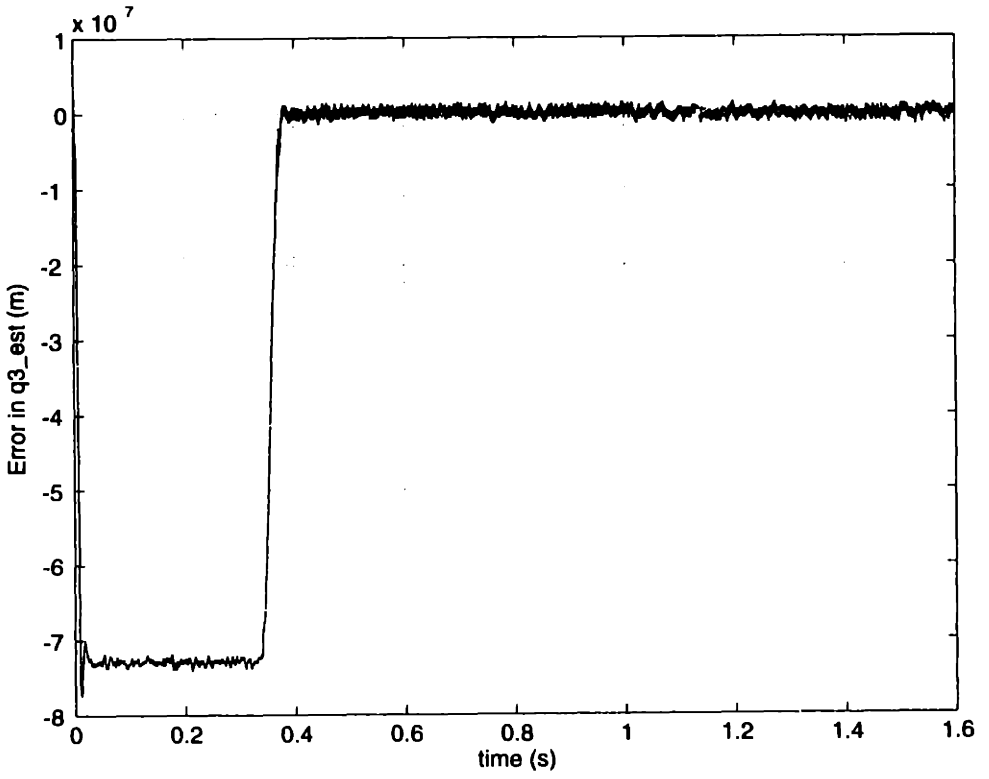

Figure 8-16: State estimate error of q_3.

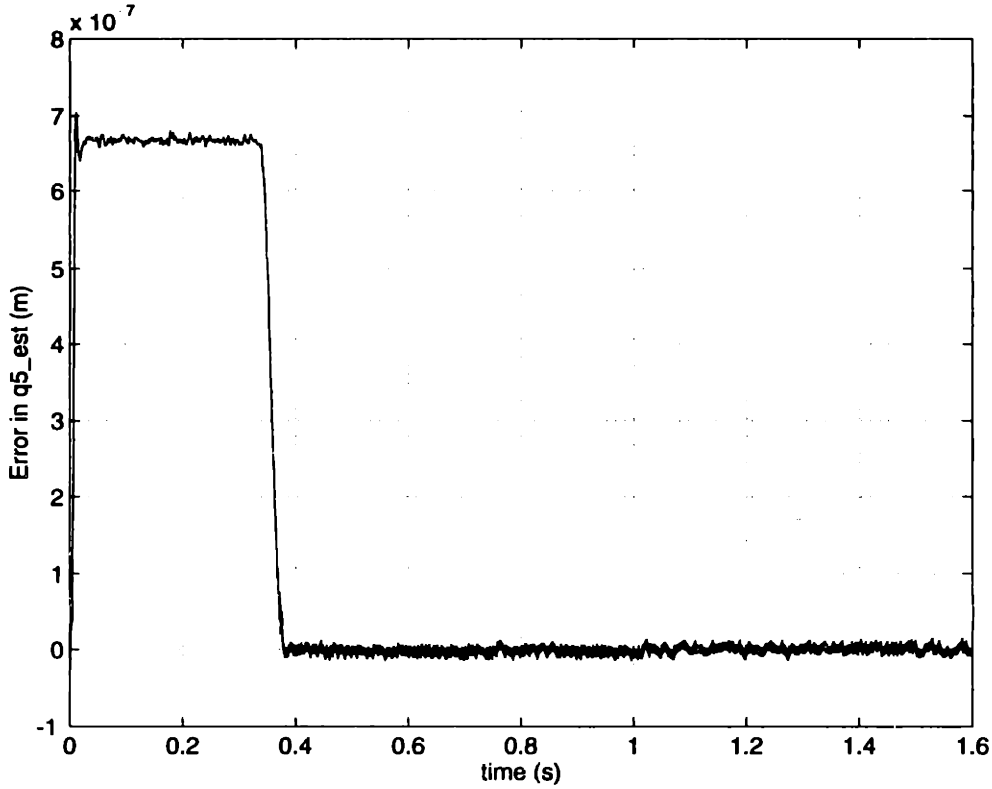

Figure 8-17: State estimate error of q_5.

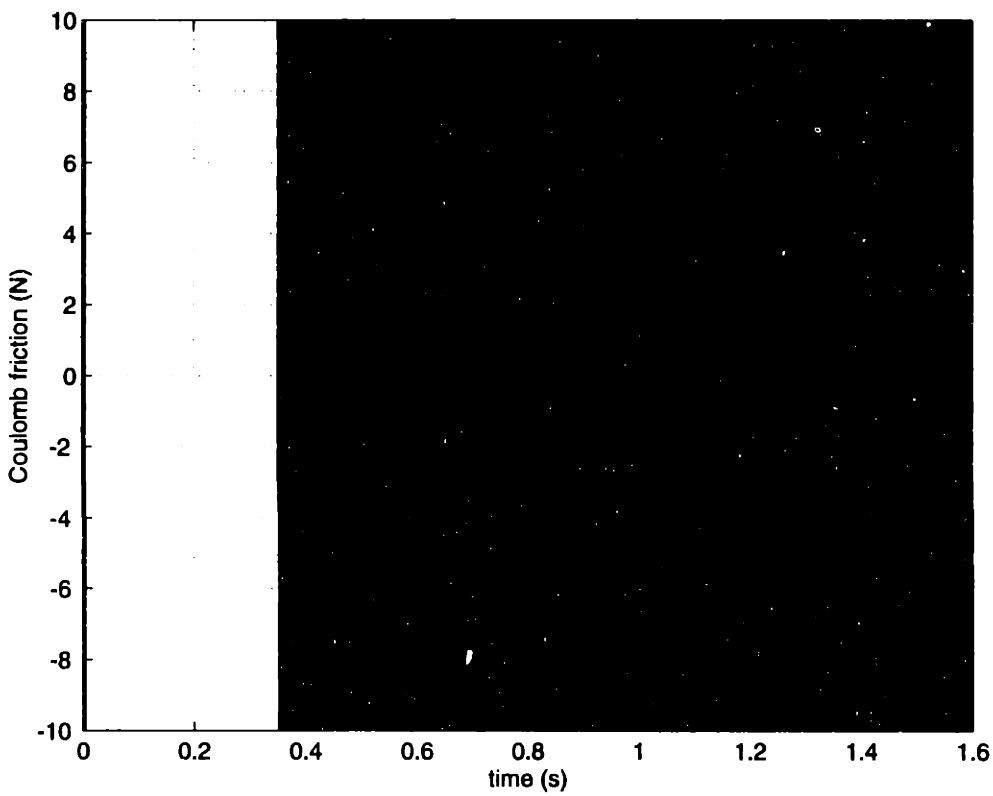

Figure 8-18: Time history of Coulomb friction in the system.

Bibliography

[1] Weinmann A. *Uncertain Models and Robust Control.* Springer Verlag, New York, 1991.

[2] Walcott B. L. and Zak S. H. State observation of nonlinear uncertain dynamical systems. *IEEE Transactions on Automatic Control*, 32(2):166–170, February 1987.

[3] Jayawant B. V. *Electromagnetic Levitation and Suspension Techniques.* Edward Arnold, London, United Kingdom, first edition, 1981.

[4] Wai C. C. H. and Morel J. Application of an active bearing for vibration control near critical speeds in large rotating machines. *Rotating Machinery Dynamics*, 1:151–158, 1987.

[5] Karnopp D. C., Margolis D. L., and Rosenberg R. C. *System Dynamics : A Unified Approach.* John Wiley and Sons, New York, New York, second edition, 1990.

[6] Havenhill D. D. and Kral K. D. Payload isolation using magnetic suspension. *Proceedings AAS Guidance and Control Conference, Keystone, CO*, February 1985.

[7] Trumper D. L., Olson S. M., and Subrahmanyan P. K. Linearizing control of magnetic suspension systems. *IEEE Transactions on Control Systems Technology*, 5(4):427–438, 1997.

[8] Misawa E. A. and Hedrick J. K. Nonlinear observers - a state of the art survey. *Transactions of the ASME*, 111:344–352, 1989.

[9] Matsumura F., Fujita M., and Oida C. Theory and experiment of magnetic bearing combining radial control and thrust control. *IEEE Transactions on Magnetics*, pages 2581–2583, 1987.

[10] Biernson G. *Principles of Feedback Control*, volume 2, chapter 15. John Wiley and Sons, New York, New York, 1988.

[11] Franklin G., Powell D. J., and Workman M. L. *Digital Control of Dynamic Systems.* Addison-Wesley, Reading, Massachusetts, second edition, 1994.

[12] Woodson H. H. and Melcher J. R. *Electromechanical dynamics.* Krieger Publishing Company, Malabar, Florida, 1990.

[13] Slotine J. J. E. Sliding controller design for non-linear systems. *Int. J. Control*, 40(2):421–434, 1984.

[14] Slotine J. J. E. and Li W. *Applied Nonlinear Control.* Prentice Hall, Englewood Cliffs, New Jersey, first edition, 1991.

[15] Paros J. M. and Weisbord L. How to design flexure hinges. *Machine Design*, pages 151–156, November 1965.

[16] Salm J. R. Active electromagnetic suspension of an elastic rotor: Modeling control and experimental results. *Rotating Machinery Dynamics*, 1:141–149, 1987.

[17] Bendat J. S. and Piersol A.G. *Random Data; Analysis and Measurement Procedures.* John Wiley and Sons, New York, New York, second edition, 1986.

[18] Nonami K. *Magnetic Bearings*, pages 177–186. Springer-Verlag, 1988.

[19] Ogata K. *Modern Control Engineering.* Prentice Hall, Upper Saddle River, New Jersey, third edition.

[20] Lublin L. and Athans M. *The Control Handbook.* CRC Press Inc., 1996.

[21] Meirovitch L. *Analytical Methods in Vibrations.* The Macmillan Co., New York, New York, 1967.

[22] Meirovitch L. *Elements of Vibration Analysis.* McGraw-Hill, New York, New York, second edition, 1986.

[23] Williams M., Subramanyan P., and Trumper D. L. Six axis active vibration isolation and payload reaction force compensation system. *Proceedings of the* 12^{th} *Annual ASPE Meeting, Norfolk, VA*, pages 494–497, October 1997.

[24] Allaire P. E., Maslen E. H., Lewis D. W., and Flack R. D. Magnetic thrust bearing operation and industrial pump application. *Transactions of the ASME*, 119:168–173, 1997.

[25] Sinha P. K. *Electromagnetic Suspension - Dynamics and Control.* Peter Peregrinus, London, United Kingdom, 1987.

[26] Goodall R. Dynamic characteristics in the design of a maglev suspension. *Journal of Rail and Rapid Transit*, pages 33–41, 1994.

[27] Earnshaw S. On the nature of the molecular forces which regulate the constitution of the luminferous ether. *Trans. Camb. Phil. Soc.*, pages 97–112, 1842.

[28] Olson S. M. Nonlinear compensation of a single degree of freedom magnetic suspension system. Master's thesis, Massachusetts Institute of Technology, Mechanical Engineering Department, 1994.

[29] Poovey T., Holmes M., and Trumper D. L. A kinematically-coupled magnetic bearing calibration fixture. *Precision Engineering*, 16(2):99–108, November 1994.

[30] Geng Z. J. and Haynes L. S. Six degree-of-freedom active vibration control using the stewart platforms. *IEEE Transactions on Control Systems Technology*, 2(1):45–53, 1994.

[21] Meirovitch L. *Analytical Methods in Vibrations.* The Macmillan Co., New York, New York, 1967.

[22] Meirovitch L. *Elements of Vibration Analysis.* McGraw-Hill, New York, New York, second edition, 1986.

[23] Williams M., Subramanyan P., and Trumper D. L. Six axis active vibration isolation and payload reaction force compensation system. *Proceedings of the 12^{th} Annual ASPE Meeting, Norfolk, VA*, pages 494–497, October 1997.

[24] Allaire P. E., Maslen E. H., Lewis D. W., and Flack R. D. Magnetic thrust bearing operation and industrial pump application. *Transactions of the ASME*, 119:168–173, 1997.

[25] Sinha P. K. *Electromagnetic Suspension - Dynamics and Control.* Peter Peregrinus, London, United Kingdom, 1987.

[26] Goodall R. Dynamic characteristics in the design of a maglev suspension. *Journal of Rail and Rapid Transit*, pages 33–41, 1994.

[27] Earnshaw S. On the nature of the molecular forces which regulate the constitution of the luminferous ether. *Trans. Camb. Phil. Soc.*, pages 97–112, 1842.

[28] Olson S. M. Nonlinear compensation of a single degree of freedom magnetic suspension system. Master's thesis, Massachusetts Institute of Technology, Mechanical Engineering Department, 1994.

[29] Poovey T., Holmes M., and Trumper D. L. A kinematically-coupled magnetic bearing calibration fixture. *Precision Engineering*, 16(2):99–108, November 1994.

[30] Geng Z. J. and Haynes L. S. Six degree-of-freedom active vibration control using the stewart platforms. *IEEE Transactions on Control Systems Technology*, 2(1):45–53, 1994.

[21] Meirovitch L. *Analytical Methods in Vibrations.* The Macmillan Co., New York, New York, 1967.

[22] Meirovitch L. *Elements of Vibration Analysis.* McGraw-Hill, New York, New York, second edition, 1986.

[23] Williams M., Subramanyan P., and Trumper D. L. Six axis active vibration isolation and payload reaction force compensation system. *Proceedings of the* 12^{th} *Annual ASPE Meeting, Norfolk, VA*, pages 494–497, October 1997.

[24] Allaire P. E., Maslen E. H., Lewis D. W., and Flack R. D. Magnetic thrust bearing operation and industrial pump application. *Transactions of the ASME*, 119:168–173, 1997.

[25] Sinha P. K. *Electromagnetic Suspension - Dynamics and Control.* Peter Peregrinus, London, United Kingdom, 1987.

[26] Goodall R. Dynamic characteristics in the design of a maglev suspension. *Journal of Rail and Rapid Transit*, pages 33–41, 1994.

[27] Earnshaw S. On the nature of the molecular forces which regulate the constitution of the luminferous ether. *Trans. Camb. Phil. Soc.*, pages 97–112, 1842.

[28] Olson S. M. Nonlinear compensation of a single degree of freedom magnetic suspension system. Master's thesis, Massachusetts Institute of Technology, Mechanical Engineering Department, 1994.

[29] Poovey T., Holmes M., and Trumper D. L. A kinematically-coupled magnetic bearing calibration fixture. *Precision Engineering*, 16(2):99–108, November 1994.

[30] Geng Z. J. and Haynes L. S. Six degree-of-freedom active vibration control using the stewart platforms. *IEEE Transactions on Control Systems Technology*, 2(1):45–53, 1994.