Design, Fabrication and Mechanical Optimization of A Flexural High Speed Nanopositioning Imaging Stage

by

Robert M. Panas

Sc.B. Mechanical Engineering
Sc.B. Physics
Massachusetts Institute of Technology, 2007

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 2009

© 2009 Massachusetts Institute of Technology
All rights reserved.

ARCHIVES

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
JUN 16 2009
LIBRARIES

Signature of Author.....
Department of Mechanical Engineering
May 14, 2009

Certified by.....
Martin L. Culpepper
Associate Professor of Mechanical Engineering
Thesis Supervisor

Accepted by.....
David E. Hardt
Professor of Mechanical Engineering
Graduate Officer

Design, Fabrication and Mechanical Optimization of A
Flexural High Speed Nanopositioning Imaging Stage

by

Robert M. Panas

Submitted to the Department of Mechanical Engineering
on May 14, 2009 in Partial Fulfillment of the
Requirements for the Degree of Master of Science in
Mechanical Engineering

ABSTRACT

The intent of this research is to generate the knowledge required to design, fabricate and operate a device capable of high speed nano-scale vertical positioning of microscopy samples. The high speed focusing device (HSFD) created during this research utilizes a new combination of technologies for the purpose of imaging: Lorentz coil actuation, flexural bearings and strain gage sensing. The application of the technologies combined with precision design principles, as used in the HSFD, result in a demonstrated combination of performance and cost gains over a measured commercially available system. The HSFD is able to perform steps with 8 ms 95% settling time, 2% dynamic accuracy, and 0.005% static accuracy while operating with a resolution of 10.5 nm (1σ) over a range of 500 μm at a cost of about $1400. This performance is $3x$ faster stepping, $2x$ better dynamic accuracy, $\approx 100x$ better static accuracy, equivalent resolution and range to the top of the line commercial devices at less than half of the cost. The reduced cost is envisioned to enable greater distribution and use of nano-positioning imaging stages, while the increased performance is envisioned to enable faster, more benign (in the case of biological sciences) and more precise imaging. The increased use and data gathering ability of the new device are envisioned to enable fields of research such as biology and materials science to extend their bounds further into the micro/nano-scale as well as further along the time scale for both high speed and low speed processes.

Thesis Supervisor: Martin L. Culpepper
Title: Associate Professor of Mechanical Engineering

ACKNOWLEDGEMENTS

This work was made possible through the topic suggestion and financial support of Prior Scientific. I would like to thank the individuals at Prior, including Jim Hunt and Tom Freda in particular, for giving me both the freedom to work on this project in the directions I thought necessary, and the time to bring the HSFD to life. I subjected Tom and Jim to innumerable extraordinarily long presentations over the two years of this project. That alone deserves thanks for their patience through these meetings. I look forward to working with Prior to bring the HSFD into realization as a commercial product.

Briefly, I would like to thank my Professor, Martin Culpepper for his insight into difficult problems, his trust in me to carry out my work as I believed necessary and his willingness to allow me into the PCSL about 6 years ago now as a UROP, despite things getting off to a rocky start. I owe Prof. Culpepper a few pages worth of thanks for all of the opportunities and advice he has given me, however this paragraph will have to suffice until my Ph.D. acknowledgements.

I would also like to thank my fellow graduate students, Chris Dibiasio, Michael Cullinan, A.J. Schrauth, Aaron Mazzeo, Kevin Miu, Matthew Dircxz, Melinda Hale Terianne Hall and Alex Slocum for listening to my complaints, lending me equipment when necessary, and taking the time to offer excellent advice in a great number of fields which I can only hope to know half as well as them someday. In particular, Aaron Mazzeo, Kevin Miu each kindly volunteered a number of hours helping me gain sufficient insight into the topic of control systems to get my thesis functioning. Terianne Hall worked exhaustively with me over the last month of the project to refine the control loop into a scheme which drove the system to the limits of its performance envelope. My thanks to Alex Slocum for volunteering a number of hours to help correct thesis drafts.

Finally, it is unquestionably necessary to give grateful thanks for the personal support of my fiancée and soon to be fellow graduate student, Cynthia Walker. As well to my mother and father and everyone else who got earfuls of largely senseless technical details pertaining to my weekly victories and defeats as I worked my way through this- sometimes overwhelming- project for the last two years.

CONTENTS

FIGURES

TABLES

CHAPTER

1

INTRODUCTION

1.1 Motivation

The intent of this research is to generate the knowledge required to design, fabricate and operate a precision machine capable of high speed nano-scale positioning of microscopy samples for volumetric imaging. This machine, or high speed focusing device (HSFD), shown in Figure 1.1, utilizes a new combination of technologies that differ from the technology sets commonly used in existing focusing devices [1][2][3][4] to enable a combination of performance and cost gains. This new set of technologies includes electro-magnetic Lorentz coil actuation, flexural bearings and strain gage sensing.

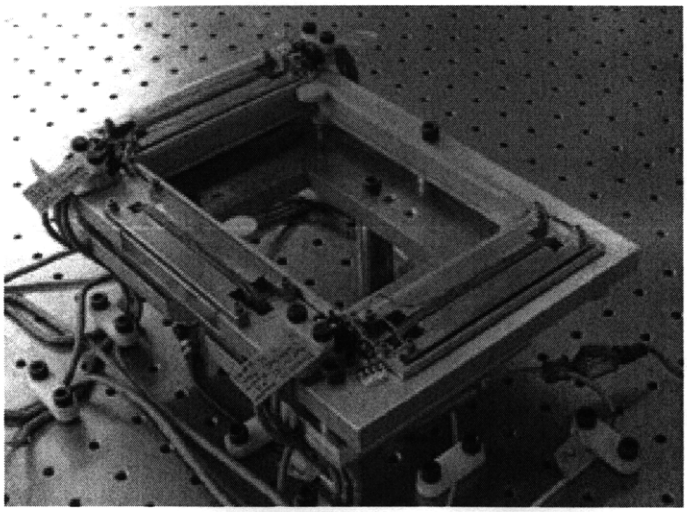

Figure 1.1: Completed HSFD Prototype.

The focusing device designed through this research will enable advances in the performance of volumetric imaging systems such as Confocal and Two-Photon microscopy. New Confocal systems can already image at 200 Hz [5] and above, roughly 4*x* faster than present focusing devices are able to operate, leaving the focusing device as the system bottleneck. The performance of the entire volumetric imaging system can be improved by advancing the speed and reliability of the focusing device. The HSFD is shown to perform steps with 8 ms 95% settling time, 2% dynamic accuracy, and 0.005% static accuracy while operating with a resolution of 10.5 nm (1σ) over a range of 500 μm at a cost of about $1400. This performance is 3*x* faster stepping, 2*x* better dynamic accuracy, ≈100*x* better static accuracy, with equivalent resolution and range to the top of the line commercial devices at less than half of the cost. Microscopy as a research tool can be further distributed among the research community by lowering the cost of these research devices. Volumetric imaging, such as Confocal microscopy, is an important tool in biological and material sciences [6][7][8][9][10] and would directly benefit from the ability to more rapidly and accurately adjust the focal depth in the sample. The result of this improvement would be both the ability to rapidly step to image faster processes and hold a samples location for imaging slower processes. Improvements in performance will result in faster, more benign (in the case of biological sciences) and more precise imaging. The increased data gathering ability of the new device, combined with increased use in laboratories, will enable fields of research such as biology and materials science to extend their analytical and imaging capabilities further into the micro/nano-scale as well as further along the time scale for both high speed and low speed processes.

1.2 Background

1.2.1 Confocal Microscopy

Confocal microscopy was first outlined by Marvin Minsky in 1961 [11] as a method o non-invasive volumetric imaging. The sample is viewed over a series of stacked focal planes which are reconstructed to form a volumetric (3-D) image of the sample. Each plane is scanned sequentially using the particular optical property of Confocal microscopy which enables light from a single point in 3-D space to be measured. This method has several advantages over the standard practice of optical microscopy, where only the surface of the sample can easily be

discerned. In order for traditional optical microscopy to construct volumetric, or even sub-surface images, slices have to be physically removed from the sample. Non-destructive imaging enables in-vivo studies of samples, potentially gathering much more information about the object of interest as well as preserving the sample's structure, which is sometimes key to obtaining an accurate diagnosis. Confocal microscopy also simplifies the imaging of structures with features extending along the Z-axis, such as cells [7][9] and 3D micro-materials [6][12][13].

A Confocal microscope operates through the use of two apertures, as shown in Figure 1.2. The first (Illumination) aperture turns the illumination into a point source that has a set distance from the focusing lens. This ensures the illumination is focused to a particular height (Focal Plane) in the sample. The return image is diverted using a partially reflecting mirror through a second aperture (Image Aperture). The image aperture is placed at an equivalent optical distance from the focusing lens as the illumination aperture. In such a configuration the two apertures are said to be 'confocal'. Additionally, this configuration ensures that the only image that is not attenuated by the image aperture is the one which was emitted from the illuminated point on the focal plane.

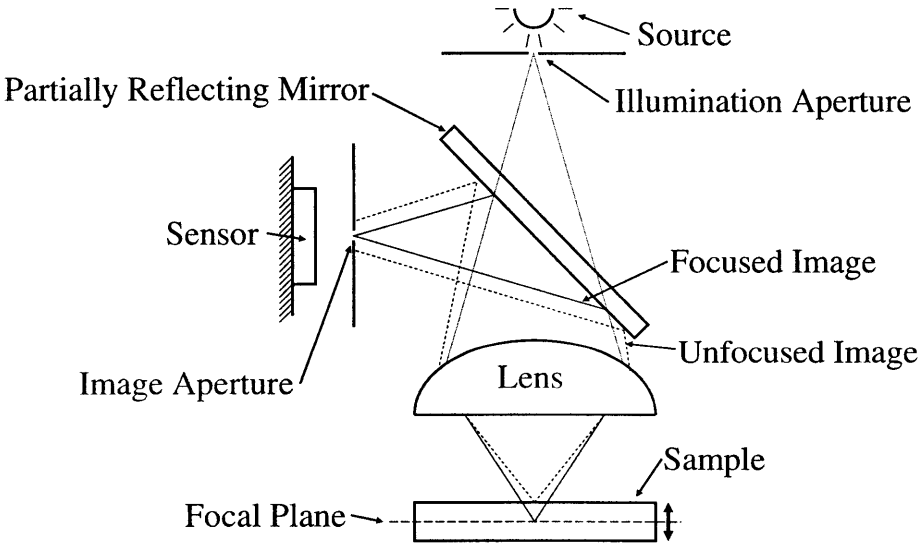

Figure 1.2: Schematic of Confocal microscope.

The desired image from the illuminated point in the focal plane is, in theory, able to pass unaltered through the image aperture and is captured by the imaging sensor. Focusing the illumination to exactly the point which is imaged maximizes the image quality, however the apertures must be aligned for best performance otherwise the image quality or intensity is

degraded. This system is only capable of examining one 3D spot at a time on the focal plane, thus in order to collect images, the imaging point must be scanned over the focal plane. One method for doing this is through the use of Nipkow disks for the aperture [14], which have holes in spiral patterns cut through them. By rotating the disk, the imaging point is rapidly translated over the focal plane, as shown in Figure 1.3.

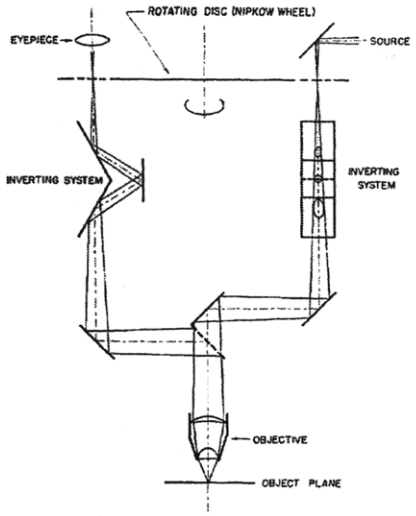

Figure 1.3: Schematic of Confocal microscope utilizing Nipkow disk [15]. © 1968 OSA.

Confocal microscopy has proven to be an extremely useful tool for science over the years beginning with its commercial introduction in the 1980's, and has gained widespread use throughout a range of fields from chemistry to biology to micro-electronics [19]. It first gained widespread use in biology as a method to observe structures within living cells, which was not feasible with earlier technologies due to the tendency to kill the cell during processing of the sample for imaging. Sample preparation and special chemicals were not required for Confocal imaging and systems could be observed in full 3-D. This revolutionized light microscopy in cell biology and pathology [16]. Confocal imaging has expanded into materials science and is now used to study a wide range of micro- and nano-scale structures such as those shown below.

The propagation of electric fields through liquid crystals used for displays [6] was studied using Confocal microscopy with sub-micron resolution to better understand the response of the material to applied fields, as shown in Figure 1.4.

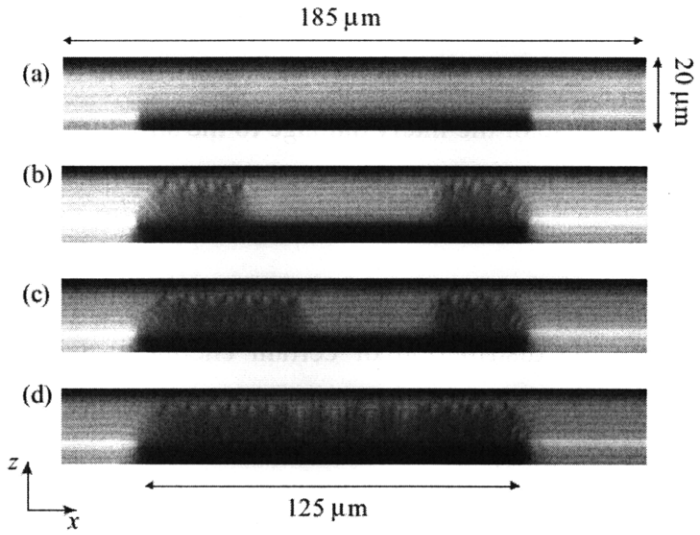

Figure 1.4: Confocal image of liquid crystals in an electric field [20]. © 2008 SPIE.

This direct observation and study of the micro-scale behavior of liquid crystals enables a better understanding of the physics underlying the liquid crystals. An improvement in the theory of liquid crystal enables improvements in liquid crystal display performance and device efficiency.

New biological structures, called bio-nanotubes, were recently experimentally observed through Confocal imaging of cells [7], as shown in Figure 1.5.

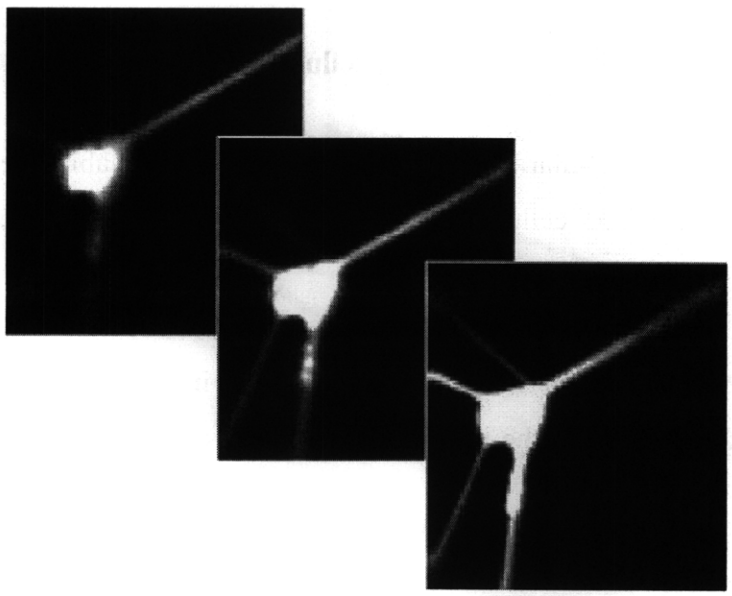

Figure 1.5: Confocal image of bio-nanotubes [21]. © 2008 ACS.

These structures have not previously been observed due to their extreme geometry. The extraordinarily small cross section of the tubes makes it difficult to observe them using 2-D slices of the sample, both because of the likely damage to the structure during sample preparation and the small cross-section. Only when observed in-situ and through 3-D imaging do they become apparent.

Intracellular structures have been studied with Confocal imaging, resulting in both the measurements of real-time 3D distribution of certain chemicals within the cell [8] and the geometric layout of sub-cellular organelles [9][22], such as chloroplasts as shown in Figure 1.6.

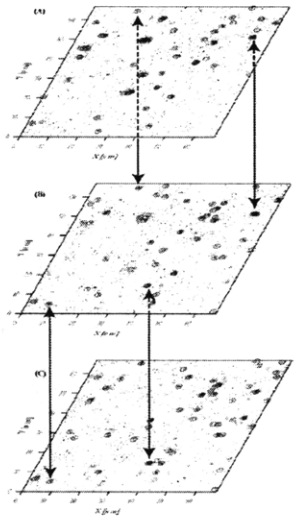

Figure 1.6: Confocal image of intracellular structure [23]. © 2008 ACS.

Direct observation of intracellular chemistries and organelles enables cellular models to be quantitatively fit to observed cellular properties, making it possible to better predict and understand cellular mechanics.

Confocal imaging has been used to study the properties of micro and nano-scale materials such as fibers [12] and silicon nanorods [13]. This imaging technology enables direct observation and quantification of the nano-scale geometry of the materials, as shown in Figure 1.7.

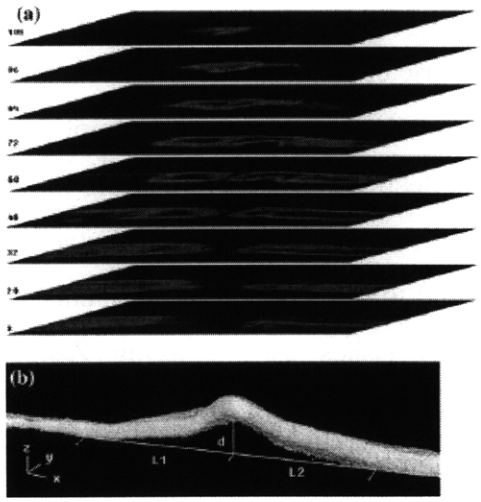

Figure 1.7: Confocal image of micro-fiber [24]. © 2008 JMS.

The 3D images of micro- and nano-scale structures allow for study of their material properties, and evaluation of their fabrication. All of this is a result of Confocal microscopy technology.

1.2.2 Positioning Stage

In order to obtain volumetric images, the focal plane of the microscope must be moved relative to the sample. The process of collecting these planar images is referred to as Z-stacking [16]. There are two methods to do this: realign optics to change the focal plane while the sample remains stationary, or realign the sample height while the focal plane remains unchanged.

Realigning the sample is generally the better choice for several reasons which are linked to the complexity of the optics. By altering the optical system through motion of the components, the possibility of damage (thermal warping, creep, fatigue, etc) is increased exactly where it is most difficult to repair, inside the tightly controlled environment of the optical column. Additionally, the optical components are calibrated to exact locations and angles; motion of these components makes continued assurance of this calibration difficult. In comparison, the sample is often of low mass and resistant to motion based damage, thus the small amount introduced by physically displacing it does not affect the sample or the images significantly.

In-plane motion can be carried out through several methods: shifting the imaging point with a Nipkow disk-like aperture mechanism as described previously, by adjusting the angle of mirrors or an Acousto-Optic Deflector (AOD) directing the light to the sample [5] or by moving

the sample. The AOD adjusts the optical path of the microscope through slight variations in the index of refraction of a piezoelectric crystal. These variations are generated by a standing sound wave which is created in the crystal. Variation in the amplitude and frequency are generally used to control the amount of deflection of the light passing through the AOD. The case of in-plane translation of the sample differs in comparison to Z-stacking in that the scanning mirrors are low mass and sturdy compared to other optical components. Motion of these mirrors or the AOD does not cause the same issues as attempting Z-stacking using only the optics. Additionally, the frequency of motion is much higher in plane than out-of-plane, as the focal plane scanning for video-rate imaging must be carried out over each plane before stepping up or down. This makes low mass mirrors or AOD the general solution for scanning.

1.3 Prior Art

1.3.1 Technologies

A wide range of technologies have been used in a variety of permutations to produce Z-axis positioning stages. Technologies used for actuation include leadscrews [2][11], linear actuators [3][25][26], and piezoelectrics [1][4][27][28]. Technologies used for bearings include sliding [3][11][26], rolling [3][26], and flexural [1][2][4][25][27]. Technologies used for position sensing (when it is used at all) include- strain gages [25], optical [1], and piezoelectric [4]. A few selected designs encompassing the various important permutations of technologies listed above have been examined in greater detail in the sections below.

1.3.2 Precision Microscope Stage

George described a design for a microscope stage utilizing flexural bearings and leadscrew actuation, as laid out in his 1972 patent [2]. This design is shown in Figure 1.8.

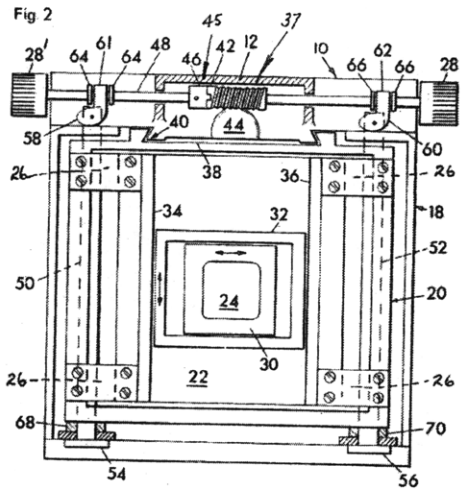

Figure 1.8: Diagram of precision microscope stage [2].

The design utilized flexural bearings (element 26 in the figure) to set the kinematics of the outer stage to only enable vertical deflections. This bearing was then driven through a worm-gear like actuator (44) which acted as a displacement actuator. The inner nested stages (30, 32) enabled X and Y motion through some un-specified actuation and bearing.

The design lacks the closed loop operation which is crucial to precision devices for rejecting disturbances such as material property creep, thermal error, loading errors and a whole host of other minor parameters which cannot be held precisely constant. This design is intended to be manually actuated, and this actuation is coarse, due to the use of rubbing contact to drive motion. Backlash and static friction will become major problems for this device if high accuracy and resolution positioning is desired.

1.3.3 Mechanical Scan System for Microscopic Applications

Marsman et al. described a design for a microscope stage utilizing flexural bearings, piezoelectric actuation, piezoelectric sensing and closed loop control as laid out in his 1983 paper [28]. This design is shown in Figure 1.9.

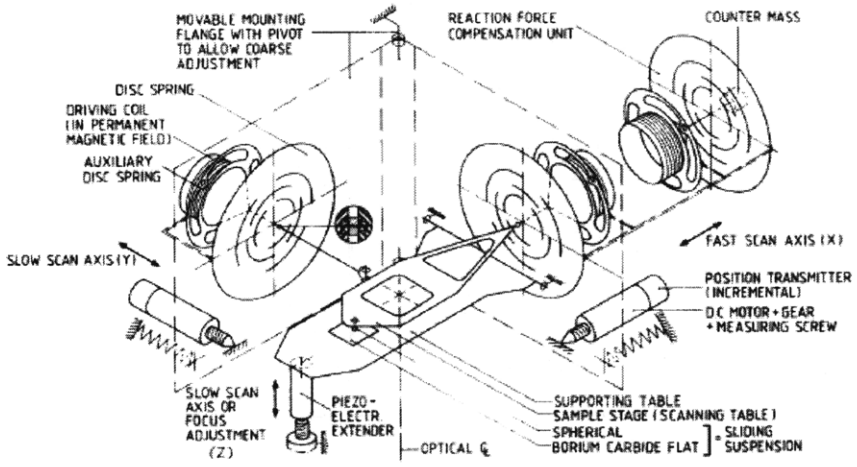

Figure 1.9: Diagram of mechanical scan system [29]. © 1983 AIP.

The voice-coil type actuation is used for in-plane actuation, and small sapphire balls are used as bearings for the in-plane motion. A combination of flexural bearing and piezoelectric actuation are used for the vertical axis. The actuation is not done from directly below the sample, but rather from a lever arm where the bearing is the rotary joint. Thus, the vertical motion contains parasitic in-plane motion due to its inherently rotary nature. The device operates in a closed loop form, and has three actuators for three degree-of-freedom (DOF) motion

The design has extremely limited vertical displacement (5 um), limited sample holding capabilities and a large profile, requiring custom fitting to Confocal microscopes. It would be difficult to turn this imaging stage into a commercial product due to its complexity and size.

1.3.4 Piezoelectric Z-axis Stage

Rondeau et al. described a design for a microscope stage utilizing flexural bearings, optical sensing, piezoelectric actuation and closed loop control, as laid out in his 2007 patent [1]. The design is shown in Figure 1.10.

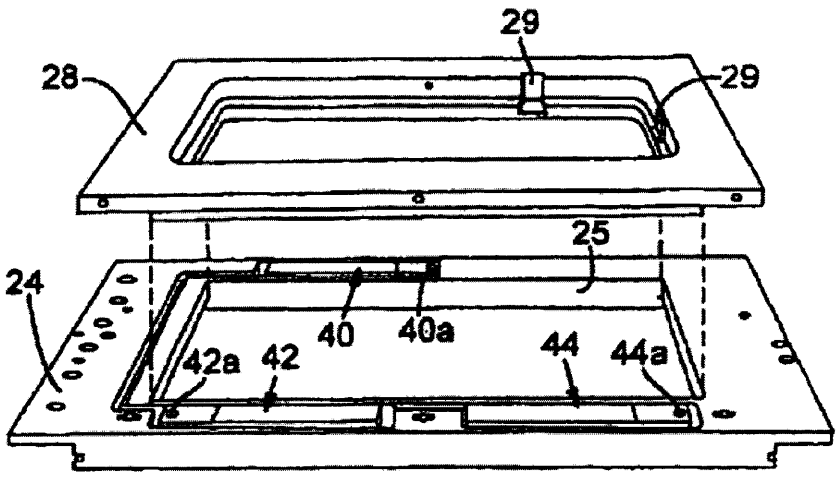

Figure 1.10: Diagram of piezoelectric Z-axis stage [1].

This design is radically simpler those previous described due to the repeated use of a combined bearing/actuation set (40, 42, 44) for producing vertical motion. This set is composed of a piezoelectric actuator nested within a flexure designed to magnify displacements through a mechanical amplification structure. Three of these sets were used to ensure kinematic control of the central moving stage (28). Optical measurements of the stage location are suggested for position sensing. This would be carried out through the use of the microscope imaging the sample- the relative motion of the stage could be observed and controlled out through the actuators.

The design does not specifically address the problem of unwanted rotations in the stage during vertical motion, especially through the slow drift of piezoelectric actuator properties. These could easily lead to large undesired and uncontrolled rotations of the central stage that would be difficult to capture with optical control. The displacement based actuation means three points of contact are made with the center stage, making it likely that some non-repeatable sliding occurs at these points unless the actuators are allowed freedom to move in-plane. Piezoelectric actuators have a limited range for their size. They also have large problems with accuracy over time, due to a slow relaxation even under voltage command. The drift would require a high resolution control loop to remove, however the suggested optical system is extremely complex, relatively low resolution and prone to grounding loop issues (such as motion of the camera being interpreted as stage motion).

1.4 Deficiencies

A number of limitations on present Z-axis positioning stage technology limit the further use of these devices for research. Most Z-axis positioning stages have difficulty achieving large displacements (>100 μm) due to their use of piezoelectric actuation [1]. This actuation technology is also expensive, prone to drifts and requires high voltage control equipment to drive correctly. Position sensing, and thus precise motion control, is not common on commercial devices. This has the added problem of making the devices sensitive to time and thermal variations in actuator/bearing performance.

The technologies presented and the sets used in the particular cases studied indicate the possibility of a new combination of technologies which could produce performance superior to any of the described cases. Recent advances in optics have made it possible to obtain single nanometer level Z-axis resolution in Confocal imaging [30]. Improvements in electronics and optics have enables imaging at rates of 200 Hz and above [5]. This means advances in the nano-positioning stages in resolution, accuracy and speed can be fully utilized by the image capture technology. With stable actuation, high resolution bearings and precise control, nanometer scale position control should be possible, enabling a host of new insights in biology, chemistry and materials science.

1.5 Requirements

The requirements for a new competitive Z-axis positioning stage, the HSFD, are laid out in Table 1.1.

Table 1.1: HSFD requirements.

Functional Requirement	Required	Desired	Units
Range	100	500	μm
Repeatability	40	1	nm
Resolution	100	10	nm
Bandwidth	100	200	Hz
Accuracy	0.5	0.05	% of travel
Maximum Load	500	500	g
Natural Frequency	>1	>1	kHz
In-plane Tilt	<10	<10	μrad
Temperature range	5 to 50	5 to 50	°C
Target system cost	<2500	<2500	$
Material	Any	Al	---
Power Draw	<30	<10	W
Stage Temp	<5	<2	°C (ΔT)
Aperture (x,y)	≥85*x*65	165*x*112	mm
Envelope (x,y,z)	≤260*x*165*x*26	<260*x*165*x*26	mm

The range, repeatability, resolution, bandwidth, accuracy, in-plane tilt and cost are set to either compete with or exceed that of existing top-of-the-line systems utilizing piezoelectric actuation such as those produced by Physik Instrument [31]. The natural frequency constraint ensures that the device settling time is rapid enough to enable sequential stepping for Z-stacking. The temperature range is the typical range set for biological systems, which the system must endure. While any material may be used, aluminum is preferred due to its machinability and low creep. The power draw constraint ensures that normal power electronics can be used to drive the device. The stage temperature constraints ensure the safety of the sample during operation. Finally, the aperture and envelope geometry specify the operating space available for the device such that it easily operates within commercial Confocal microscopes and can carry common sample holders such as 96 well plates or slides.

1.6 Scope

This thesis describes the methods and reasons underlying the engineering decisions used for design, presents the characteristics of the fabricated prototype, and describes further work which may be done to finalize the design of the HSFD.

Chapter 2 presents the qualitative design process through which the technology set for the HSFD was determined, as well as the general structure of the device. Chapter 3 describes the quantitative design process, including theory and results of the design process. Chapter 4 describes the fabrication and assembly process required to create the prototype without damaging the components. Chapter 5 presents the device characteristics, including calibrations and performance. Methods for calibration are also discussed. Chapter 6 summarizes the work done in this thesis and provides a discussion of remaining work that may be done on the HSFD to optimize its performance.

CHAPTER

2

CONCEPT COMPARISON

2.1 Introduction

The first step of the design process, qualitative analysis, is described in this chapter. Qualitative design concepts are generated through evaluation of available technologies as part of this analysis. These design concepts are compared over a range of relevant characteristics and the qualitative design with the greatest potential is chosen for further quantitative analysis.

2.2 Bearing

2.2.1 Functional Requirements

The bearing functional requirements as extrapolated from the general device requirements described in Table 1.1 previously are shown in Table 2.1.

Table 2.1: Bearing functional requirements.

Bearing Property	Required	Desired	Units
DOF	1	3	---
Cost	<1000	<100	$
Geometry	Must fit between aperture and envelope		
Resolution	<100	1	nm
Natural Frequency	1	>1	kHz

The bearing's number of DOF is determined by the basic requirement of the Z-axis positioning stage to move along the optical axis (1 DOF). Very likely the motion will not be perfect and parasitic rotations around the two in-plane axes will occur while the device is

undergoing this translation, thereby resulting in positioning error across the face of the stage. To control this, the device should be able to actively move in these added DOF in order to counteract the parasitic motions. This results in desired three DOF for the bearing. For the two rotational DOF, only a small range of motion is required. The cost is largely an order of magnitude estimate made by assuming that bearing, actuator and sensor will roughly distribute the cost quoted in Table 1.1 evenly, and is set down simply as an extreme cutoff for possible exotic and extremely expensive technologies that would otherwise meet the requirements. In order to achieve the resolution desired for the HSFD device as a whole system, all components- actuator, sensor and bearing must be capable of achieving this resolution. The geometric and natural frequency requirements are translated directly from Table 1.1.

A range of bearing technologies were considered, however an in-depth evaluation was restricted to a select range of technologies. A cursory examination with regards to simultaneously meeting cost, geometry, and resolution requirements were generally sufficient to exempt most technologies. Non-contact bearings such as: magnetic, hydrostatic and aerostatic all required significant operational support equipment which adds significant cost, complexity and size to the system. Contact bearings such as sliding, roller and flexural do not require support equipment, and thus may be fit to a compact design. Roller bearings were discarded due to an inability in the HSFD to capitalize on the main advantage this type of bearings: long range. The remaining choices were sliding and flexural bearings [32].

2.2.2 Concept: Sliding Bearing

A sliding bearing is composed of two essentially rigid bodies that are constrained through surface contact to only allow relative motion in a particular direction as shown in Figure 2.1. The sliding contact of the surfaces restrains motion in the direction normal to the surface, while allowing motion coplanar to the surfaces. The degree of constraint (DOC) is characterized by the stiffness of the 'rigid' bodies, while the DOF is characterized by the combination of static and kinetic friction. The fundamental differences between the physics characterizing these degrees enables extremely large differences in the force range associated with motion along the DOC versus the DOF.

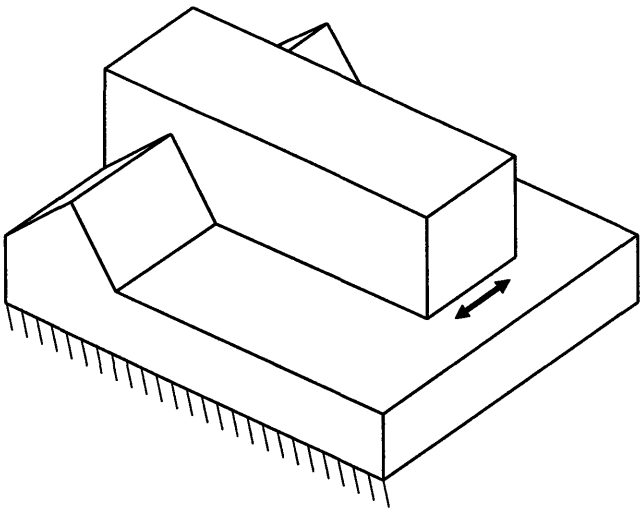

Figure 2.1: Sliding bearing.

The bearing surfaces are usually coated with a lubricant to reduce both static and kinetic friction. Friction both limits the minimum step and produces heat, which may thermally strain the device, reducing accuracy. Specialty coatings have been deposited on the bearing surfaces to further attenuate these effects in high end sliding bearings; however the coatings are expensive and complex to produce in high quality [32].

2.2.3 Concept: Flexural Bearing

A flexural bearing is composed of a structure directly linking the moving body to ground which is capable of deforming over the range of the bearing due to loading. The compliance of the flexural bearing sets the relation between the actuation force and displacement of the bearing. The flexural bearing is modeled as compliant in all directions, with the magnitude of the stiffness determining the attenuation of input forces. Generally, DOCs are characterized by stiffnesses which are two to four orders of magnitude higher than the stiffness of the DOFs.

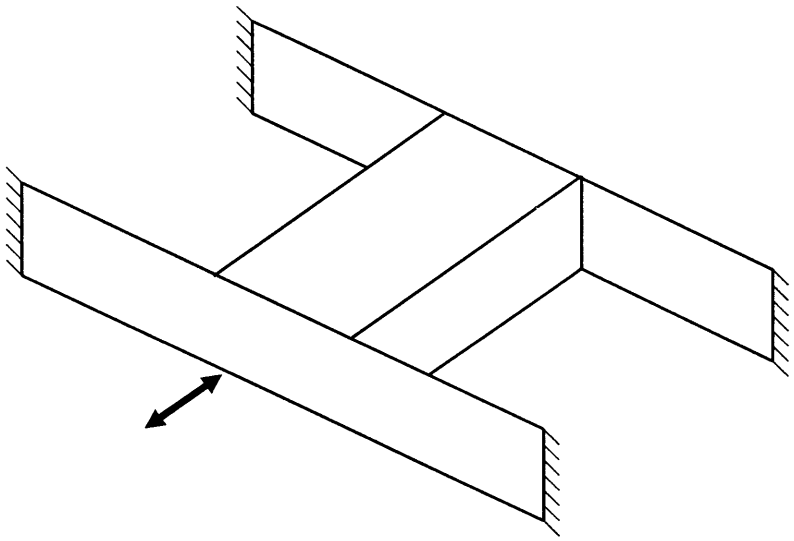

Figure 2.2: Flexural bearing.

The resolution of a flexural bearing is not limited by the flexures, but rather by the actuation equipment. Flexural bearings have a low displacement per size ratio, on the order of 0.01-0.1 due to the low strain limits of commonly used materials. The deterministic relation between applied load and resulting displacement of the bearing is extremely useful for high precision equipment [32].

2.2.4 Concept Comparison

The bearing concepts described above were qualitatively compared to determine the design best suited for use in the HSFD. Bearing properties which differed significantly among the concepts are shown in Table 2.2 and discussed in detail below.

Table 2.2: Bearing concepts Pugh chart.

Bearing Property	Sliding Bearing	Flexural Bearing
Design Freedom	0	-
Kinematics	0	+
DOF	0	+
Resolution	0	+
Friction	0	+
DOC stiffness	0	--
Complexity	0	+

The design freedom of the flexural bearing is somewhat more constrained than that of the sliding bearing due to the need to ensure the DOCs are sufficiently rigid. The sliding bearing will largely ensure this without specific design effort. The desired kinematics of the HSFD will be difficult to achieve with the sliding bearing, particularly given the low profile. The requirement of long contact area versus separation between bearing surfaces to avoid locking the bearing will be difficult to achieve in the available space. The resolution of the sliding bearing may be sufficient if the surface is properly fabricated and treated, however this will be expensive and difficult as described earlier [32]. The friction generated by the sliding contact will generate heat exactly at the location where thermal growth is least desired, likely causing accuracy issues. The flexural bearing will have difficulty reaching the desired rejection of motion along the constrained directions compared to the sliding bearing. The fundamentally different principles setting the two systems' DOC/DOF ratios (stiffnesses) mean that the sliding bearing has a significant advantage. Conversely, the lower DOC/DOF stiffness ratio of the flexures makes it easier to control the undesired motions. The rigid sliding bearing is totally un-adjustable once fabricated, while the flexural bearing's motion can be adjusted during operation. The sliding bearing will have a simpler interface, however the required kinematics and difficulty of fabrication suggest that it will be significantly more complex to implement than the flexural bearing.

2.2.5 Concept Selection

In nearly every category, the flexural bearing was determined to be the better choice, especially when it is noted that the categories in which it did not exceed the sliding bearing are those which may be accounted for through proper design. The selection process determined that flexural bearing technology should be used, however the type (geometry) of the bearing remained undecided.

2.2.6 Flexural Concept: Single Level Flexural Bearing

The single level flexural bearing concept is composed of a number of coplanar discrete flexures as shown in Figure 2.3.

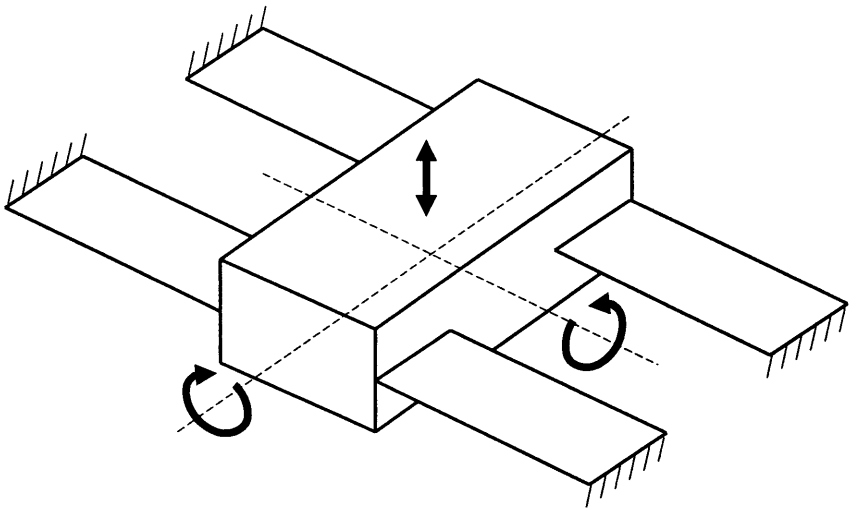

Figure 2.3: Single level flexural bearing.

The out-of-plane bending mode of a flexure is assumed to be its DOF, while the in-plane deflections are assumed to be its DOCs. The single level flexural setup will not have any significant constraint rejecting the two rotations around the in-plane axes [33], which are shown in Figure 2.3. Thus it will have low stiffness in three DOFs. It is desired that the bearing have a low stiffness only in the vertical DOF, and have moderate to high stiffness in the other two DOFs to minimize the magnitude of the rotations.

2.2.7 Flexural Concept: Multi-Level Sheet Flexural Bearing

The multi-level sheet flexural bearing concept is composed of two parallel planes of flexures as shown in Figure 2.4.

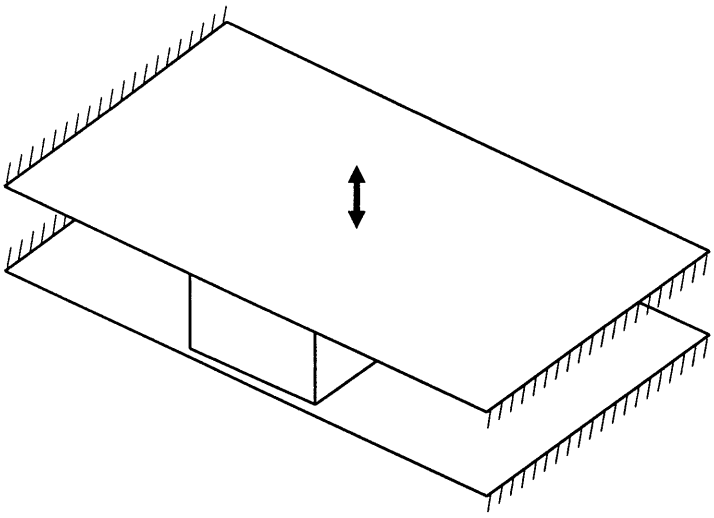

Figure 2.4: Multi-level sheet flexural bearing.

The flexures are no longer on the same plane for this concept, therefore the center of rotation is now somewhere between the planes. All of the rotations described in the single level flexural bearing will now translate into in-plane motions for the sheet flexures, resulting in high stiffnesses for these rotations [33]. The multi-level sheet flexural bearing will have only one DOF that is adjustable within a reasonable force range.

2.2.8 Flexural Concept: Multi-Level Discrete Blade Flexural Bearing

The multi-level discrete blade flexural bearing concept is composed of two parallel planes of flexure blades as shown in Figure 2.5.

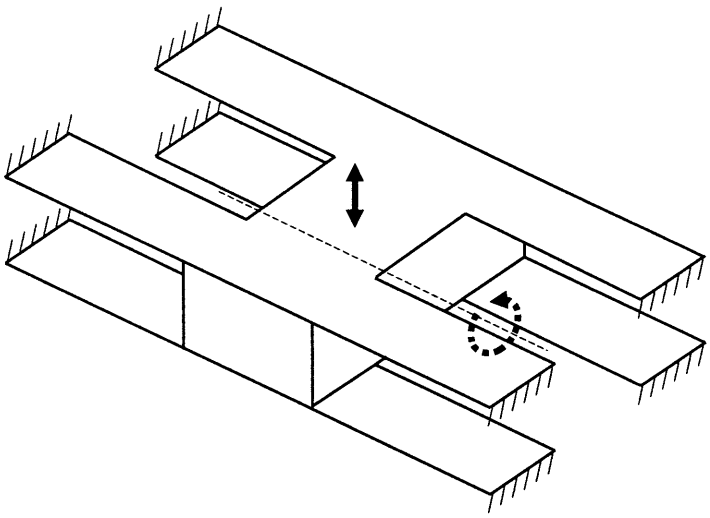

Figure 2.5: Multi-level discrete blade flexural bearing.

The multi-level discrete blade flexural bearing concept is related to the multi-level sheet flexure concept, but sections of the sheet flexure have been removed to regain some compliance in the rotational DOF. The multi-level discrete flexure blades will still have a high stiffness in the rotational DOF; however this effect will be dependent on the bearing geometry for the rotational axis shown in Figure 2.5. The flexures do not attach to the center stage directly above the axis of rotation; therefore the instantaneous direction of motion for the flexures is not necessarily largely in-plane (the DOC for flexures). As the center stage is reduced in height and increased in length (flattened), the direction of motion for the discrete blades comes closer to approximating purely vertical motion, with its correspondingly low stiffness. Thus, the stiffness for the pictured rotational motion is somewhere between the vertical DOF and the rotational DOC.

2.2.9 Flexural Concept Comparison

The flexural concepts described above were qualitatively compared to determine the design best suited for use in the HSFD. Flexural bearing properties which differed significantly among the concepts are shown in Table 2.3 and discussed in detail.

Table 2.3: Flexural concepts Pugh chart.

Bearing Property	SLDF	MLDF	MLSF
DOF	0	+	++
Clamping	0	0	-
Space requirements	0	0	-
Complexity	0	0	+
Design freedom	0	0	-

All three designs technically have three DOF, however it is desired that the two rotational DOF have much higher stiffnesses than the vertical DOF. The single level discrete flexure (SLDF) bearing has the lowest ratio of stiffnesses for rotational vs. vertical motions. The multi-level sheet flexure (MLSF) has the highest ratio. The MLSF will present fabrication difficulties due to the large area that must be clamped for the flexure to be repeatable. The low profile geometry of the HSFD requires all the components to be stacked together in as small a space as possible. The MLSF will make this difficult because it occupies a large area. The complexity of both the SLDF and the multi-level discrete flexure (MLDF) bearings are greater than the MLSF, owing to the detailed geometry, versus the simple sheet. The MLSF has significantly less design freedom, making it harder to adjust the bearing properties to desired values.

2.2.10 Flexural Concept Selection

The initial choice for the bearing was the MLSF, as little weight was given to the clamping or space requirement property. A proof-of-concept prototype (A) was fabricated and is shown in Figure 2.6. Several problems were immediately noted with the design. The clamping and space requirements were re-weighted to be much more important and the required thickness of the flexure sheet was low enough that the sheet was very susceptible to damage. The MLDF bearing was re-evaluated to be the bearing choice because of the insights gained from prototype A.

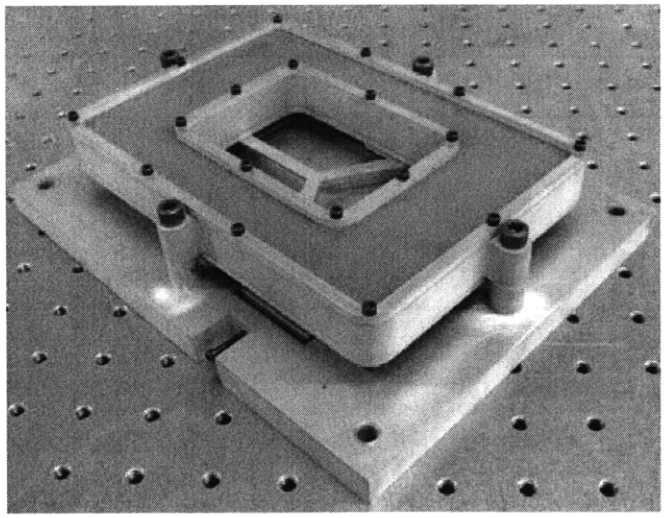

Figure 2.6: Multi-level sheet flexural bearing prototype.

2.3 Actuator

2.3.1 Functional Requirements

The actuator functional requirements as extrapolated from the general device requirements described in Table 1.1 are shown in Table 2.4.

Table 2.4: Actuator functional requirements.

Actuator Property	Required	Desired	Units
DOF	1	3	---
Cost	<1000	<100	$
Geometry	Must fit between aperture and envelope		
Resolution	1	100	nm
Power	<30	<10	W
Max force	>1	>10	N
Response Time	<10	<1	ms

The actuator functional requirements match those of the bearing for DOF, cost, geometry and resolution because the actuator must function with the bearing to produce the full system DOF, must fit in the device and must itself also have a high resolution to enable the whole system to achieve high resolution. The power requirement is drawn from Table 1.1. This requirement is intended to ensure that the HSFD can be driven by common power electronics,

rather than necessitating an expensive high-power custom setup. The maximum force requirement provides a force estimate upon which the bearing and actuator models later converge. The response time requirement ensures that the actuator is able to drive the system at or above the required bandwidth in Table 1.1.

A range of actuation technologies were considered, and as with the bearing qualitative analysis, the in-depth evaluation was restricted to a select range of technologies. The simultaneous requirements for high resolution and low power exempted most technologies, such as steppers, leadscrews, etc. The remaining technologies were piezoelectric and electromagnetic such as reluctance and Lorentz actuators [32].

2.3.2 Concept: Piezoelectric Actuator

A piezoelectric actuator is composed of stacks of disks, each of which expands when voltage is applied across it. In order to magnify the effect, many disks are stacked in series, as shown in Figure 2.7.

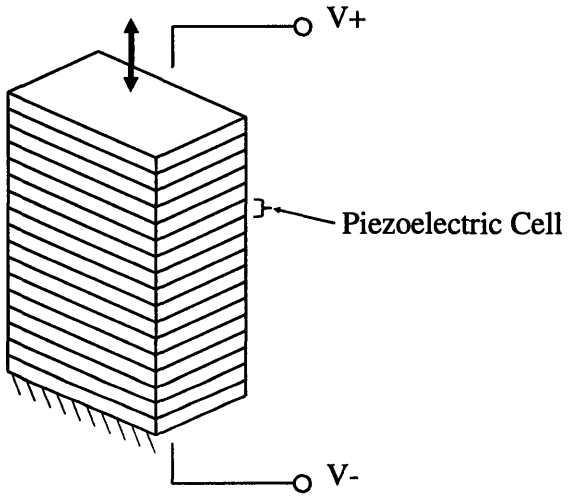

Figure 2.7: Piezoelectric actuator.

The piezoelectric effect requires high voltages but low current, thus it draws little power and generates little heat. It is also high speed because the physics of the effect are based on an atomic realignment rather than a diffusion effect. Piezoelectric actuators act as displacement sources due to their relatively high stiffness. Hysteresis and drift are significant factors when dealing with these actuators [32].

2.3.3 Concept: EM- Reluctance Actuator

A reluctance actuator is composed of a magnetic field generator (Coil) and a magnetically permeable component (Plunger) which is drawn into the magnetic circuit to lower the net circuit resistance and correspondingly lower the net circuit energy, when the coil is activated. Often, the actuator efficiency will be increased through the use of the magnetic flux guiding path (Core), as shown in Figure 2.8.

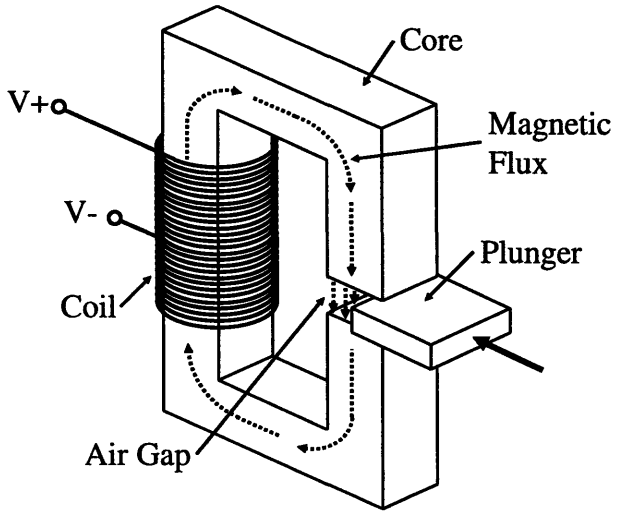

Figure 2.8: Reluctance actuator.

Reluctance actuators can produce extremely high forces because of the large difference in energy as the plunger is drawn into the magnetic circuit. In the case of reluctance actuators that utilize cores, the dominant energy storage location for the circuit is the volume of the gap (Air Gap) into which the plunger is pulled. The replacement of the air resistor with the magnetically permeable plunger (generally a ferromagnetic material like steel) reduces the magnetic resistance of the gap, lowering the energy stored in the system. Reluctance actuators have a relatively slow response because of the need for the actuator to charge up the gap magnetic field before a force is felt by the plunger [32].

2.3.4 Concept: EM- Lorentz Actuator

A Lorentz actuator is composed of a magnetic field generator (Permanent Magnet) and a coil which pushes against the magnetic field through the Lorentz force. The magnetic field is

usually guided with a magnetically permeable core, similarly to the reluctance actuator, as shown in Figure 2.9.

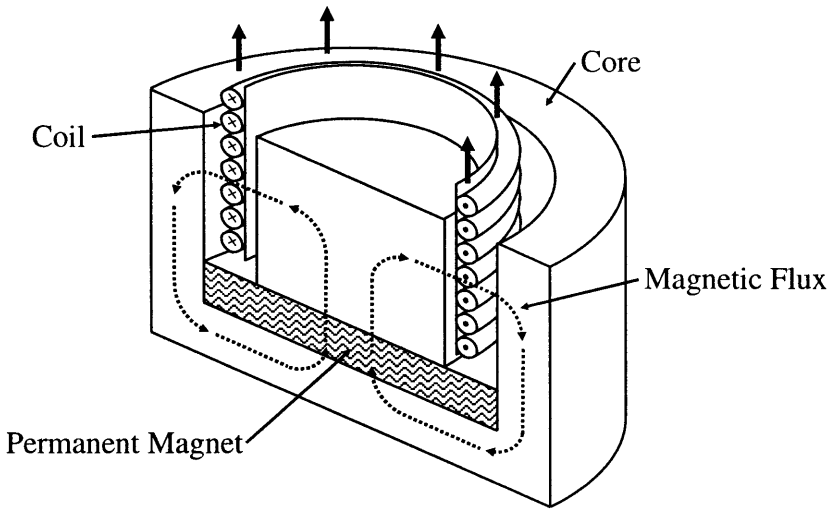

Figure 2.9: Lorentz actuator.

Once the static magnetic field is generated, which can be done using a separate coil, current is driven through the main coil to create a force. Unlike the reluctance actuator, the main coil does not have to create a magnetic field to produce a force. The Lorentz actuator coil simply pushes against an existing magnetic field, making it possible to achieve a much higher operating bandwidth than the reluctance actuator [32].

2.3.5 Concept Comparison

The actuator concepts described above were qualitatively compared to determine the design best suited for use in the HSFD. Actuator properties which differed significantly among the concepts are shown in Table 2.5 and discussed in detail below.

Table 2.5: Actuator concepts Pugh chart.

Actuator Property	Piezoelectric	Reluctance	Lorentz
Range	0	++	++
Geometry layout	0	-	0
Response time	0	--	0
Force	0	-	-
Control Equipment	0	+	+
Repeatability	0	0	+

The range of the piezoelectric actuator is limited to a small percentage of its length, while the EM actuators have a range limit of approximately the length of the plunger or coil. This requires the piezoelectric actuator to either be large or use mechanical amplification. The shape of the actuators must work with the low profile of the HSFD. The reluctance actuator design is least compatible with this, owing to its dispersed component layout. The response time of the reluctance actuator is significantly worse then the other two actuation methods due to the intrinsic requirement of charging the air gap magnetic field to create the actuation force. The piezoelectric actuator has a significant advantage in force output because of its displacement-based actuation and high stiffness. The physics of the piezoelectric actuator require high voltages (>100 V) to extend over its full range, meaning that control electronics must be specialized and are more expensive than the single digit voltage and current draw of the EM designs. The repeatability of the Lorentz actuator is likely the best among the concepts because of drift in the piezoelectric actuator and the latent magnetization of the reluctance actuator core.

2.3.6 Concept Selection

The initial thought was to use the reluctance actuator; however early modeling suggested the response time and heat dissipation were significantly higher than the functional requirement limits. The secondary choice of Lorentz actuation was therefore implemented. The geometric layout of the actuator is determined next.

2.3.7 Lorentz Concept: Ring Actuator

The ring actuator concept is composed of a single large coil wrapped in-plane around the center stage, which interacts with a magnetic field that is produced by a magnet distributed over its entire length, as shown in Figure 2.10.

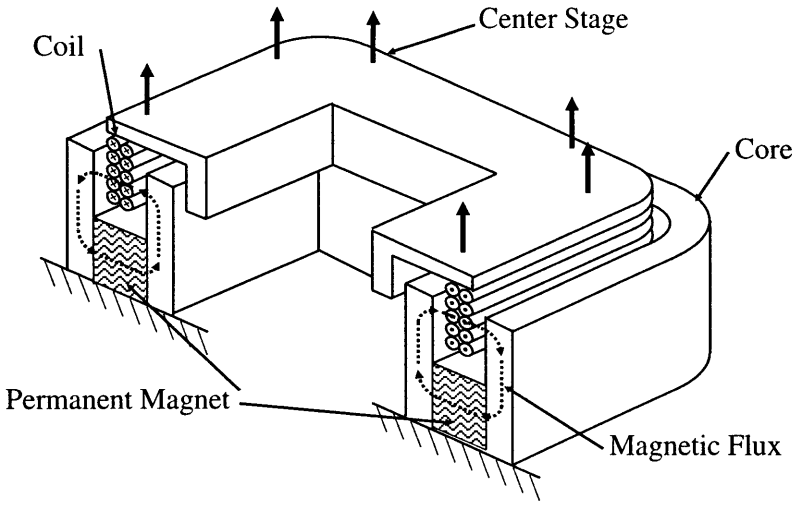

Figure 2.10: Lorentz ring actuator.

Thin walls of magnetically permeable material focus the flux from the magnet over the coil, resulting in a single net force that is pointed vertically and distributed over the entire coil.

2.3.8 Lorentz Concept: Voice Coil Actuator

The voice coil actuator concept is composed of discrete, tightly packed Lorentz coils where the magnetic flux is driven radially across the cylindrical coil for maximal packing efficiency of the actuator, as shown in Figure 2.11.

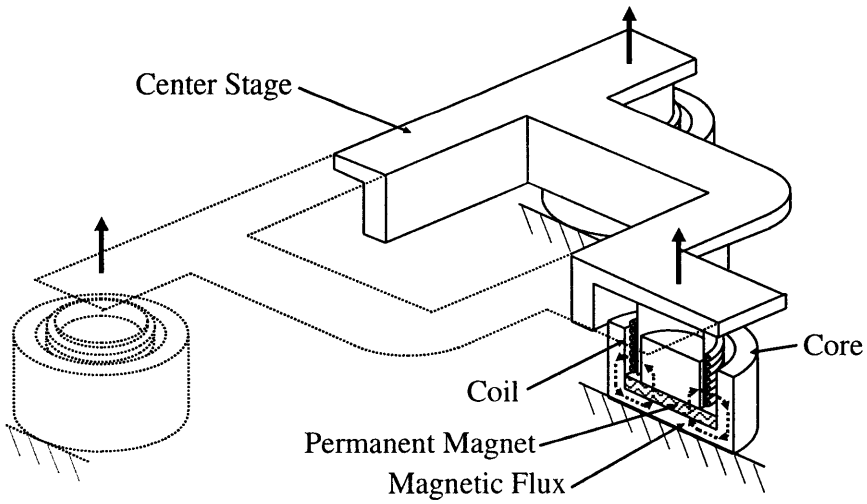

Figure 2.11: Lorentz voice coil actuator.

Each coil is independently adjustable, making it possible to command motions over more than simply the vertical DOF. The placement of the coils as shown above is done to locate their center of force near the center of the stage as well as to maintain symmetry.

2.3.9 Lorentz Concept: Compact Ring Actuator

The compact ring actuator concept is composed of discrete Lorentz actuators as with the previous design, however now the actuators are altered to conform to the low profile and rectangular geometry working volume of the HSFD, as shown in Figure 2.12.

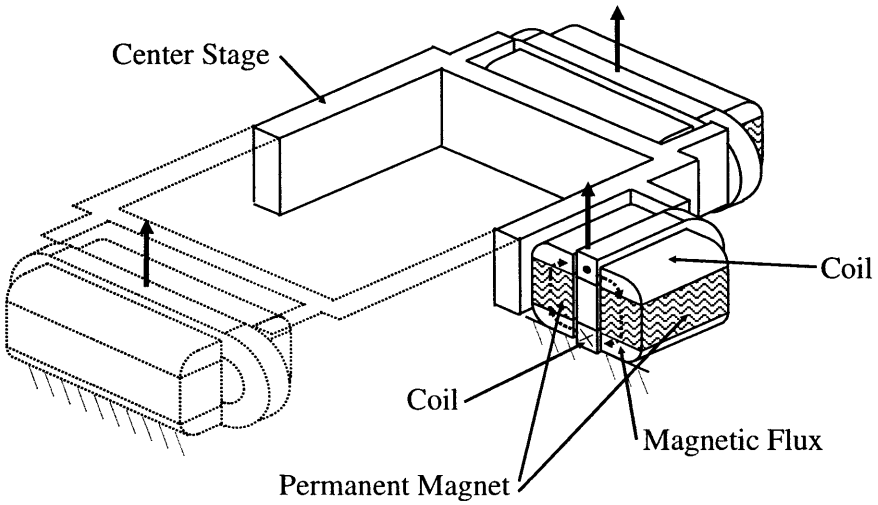

Figure 2.12: Lorentz compact ring actuator.

The extended geometry of the actuators makes them less efficient at utilizing the coil, because there is some coil out of the field at the edges. The lower efficiency is mitigated by the conforming rectangular geometry, which enables larger magnets and coils to be used, compared to the ring coils. The net result is a more powerful actuator.

2.3.10 Lorentz Concept Comparison

The Lorentz actuator concepts described above were qualitatively compared to determine the design with best suited for use in the HSFD. Lorentz actuator properties which significantly differed among the concepts are shown in Table 2.6 and discussed in detail below.

Table 2.6: Lorentz concepts Pugh chart.

Actuator Property	Ring	Voice Coil	Compact Ring
DOF	0	+	+
Geometry layout	0	-	+
Flux efficiency	0	+	+

The ring actuator concept only has one DOF for its commands as it is composed of a single coil. This does not allow the actuation force to be adjusted, only scaled. The other two designs are able to achieve three DOF for actuation. The geometric layout of the actuators is worst for the voice coil concept due to its need for large circular volumes for each actuator. The ring concept would nest well with the center stage, but would require a large volume of magnets and magnetically permeable material. The compact ring concept would fit in presently unused space in the HSFD and make maximal use of the permanent magnets used in it. As a measure of use of the available permanent magnet, the flux efficiency is lowest for the ring concept because the coil would require much core and permanent magnet material.

2.3.11 Lorentz Concept Selection

The compact Lorentz actuator was judged to be the best design concept for the HSFD actuation owing to its all around strengths compared to the other designs.

2.4 Sensor

2.4.1 Requirements

The sensor functional requirements as extrapolated from the general device requirements described in Table 1.1 previously are shown in Table 2.7.

Table 2.7: Sensor functional requirements.

Sensor Property	Required	Desired	Units
DOF	1	3	---
Geometry layout	Must fit between aperture and envelope		
Resolution	100	1	nm
Cost	<1000	<100	$
Response time	1	<0.1	ms

The sensor functional requirements match those of the bearing and actuator for DOF, cost, geometry and resolution. The remaining, unique functional requirement is the sensor response time, which should be greater than 10x faster than the mechanical bandwidth of the device. Preferably, the sensing system has a bandwidth greater than 50*x* faster, such that the digital sampling of the system (carried out at about 30*x* the operating bandwidth) is not affected by sensor dynamics.

A range of sensor technologies were considered, and as with the bearing qualitative analysis, the in-depth evaluation was restricted to a select range of technologies. The simultaneous requirements for extremely low space requirement and low cost exempted most technologies. Piezoelectric sensing was rejected due to the inherent dynamics of the sensor which cause it to act similarly to a high-pass filter. The remaining technologies were Linear Variable Differential Transformers (LVDT) and strain gage sensors [32].

2.4.2 Concept: LVDT

An LVDT sensor is composed of three coils, one (Reference Coil) which drives a sinusoidally varying magnetic flux through a simple magnetic circuit (Magnetic Flux), exciting the two sensing coils (Secondary Coil) proportionally to the location of the central moving component (Core) as shown in Figure 2.13.

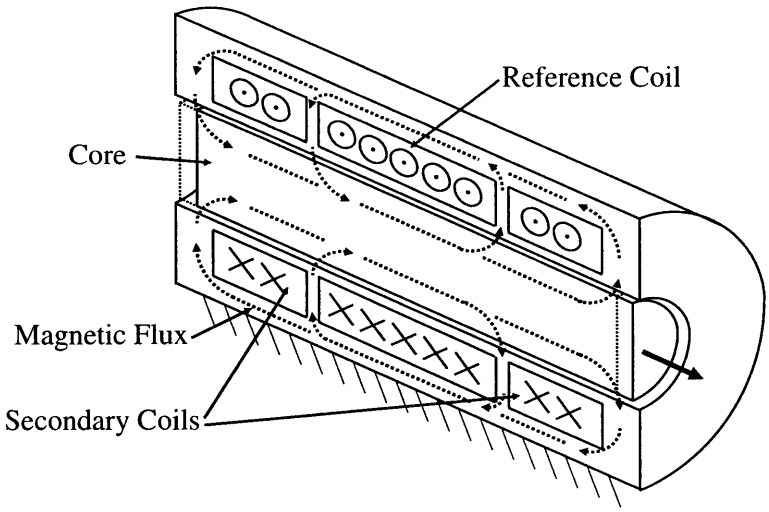

Figure 2.13: LVDT sensor.

The location of the core determines the magnetic resistance of the circuits in which the secondary coils lie. Motion of the core simultaneously increases the resistance in one secondary coil's circuit while decreasing the resistance in the other secondary coil's circuit. This difference alters the total flux passing through the secondary coils, thus affecting their responding voltage to the time varying magnetic field passing through them. The secondary coils are wired to measure the difference voltage between them. The net result is an output signal amplitude proportional to the core distance from center equilibrium. The overlap of the outer magnetically permeable shell beyond the core length ensures that the total flux driven by the reference coil is not altered, thus the moving core sees no forces generated by the LVDT. The core does not have to be in contact with the outer shell, as the magnetic field can jump the small air gap. LVDT sensors generally have a resolution on the scale of about 1 μm or greater [32].

2.4.3 Concept: Strain Gage

A strain gage sensor is composed of a thin resistance wire adhered to the surface of the material undergoing strain. Current is applied over this resistor, and the small variation in resistance with strain is read as a voltage change. The general form of a strain gage is shown in Figure 2.14.

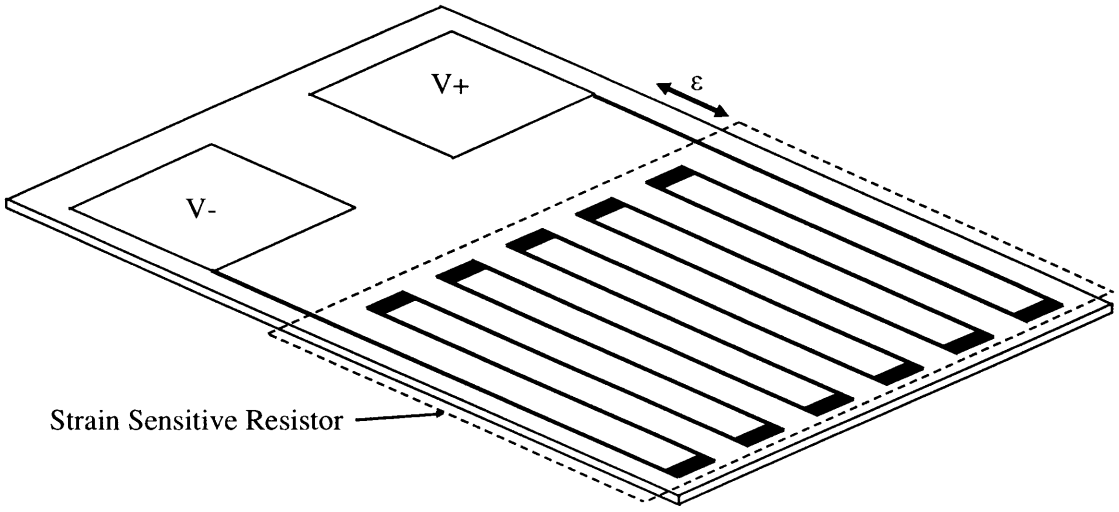

Figure 2.14: Strain gage sensor.

Strain gage sensors rely on underlying structural design of the substrate material such that the measured strain is somehow linked to the desired parameter to be measured. In the case of the HSFD, this means the strain must be linked to the displacement of the device, a requirement which is met on the flexures in the bearing. The strain dependent change in the sensor is extremely small in magnitude compared to the steady output, additional signal processing is usually done to isolate and boost the signal. Generally, this is done through hardware such as a Wheatstone bridge, as described in section 3.5.3 later. A strain gage has a theoretically unlimited inherent bandwidth, because the resistance change dynamics are nearly immeasurably fast. Strain gages have no inherent drift exempting thermal variation, which is discussed further in Section 3.5.2. Thermal compensation is generally carried out with proper hardware design. The resolution of these sensors is limited only by the electronic noise of the sensing circuitry [32].

2.4.4 Concept Comparison

The sensor concepts described above were qualitatively compared to determine the design with best suited for use in the HSFD. Sensor properties which significantly differed among the concepts are shown in Table 2.8 and discussed in detail.

Table 2.8: Sensor concepts Pugh chart.

Sensor Property	LDVT	Strain Gage
Geometry layout	0	+
Cost	0	+
Resolution	0	++

The geometry of the strain gage sensor is more easily integrated into the composite HSFD design, as it mounts directly to the surface of the flexural bearings, occupying negligible volume. LVDT sensors would require a mounting point on both the rigid base and movable stage and would have to be aligned vertically between these points. The LVDT sensor must be about on the scale of 12 mm or less in length for use in the HSFD, which would likely require a custom sensing solution using LVDT technology. The cost of such a sensor would be significantly more than strain gages, which are commercially available for a few dollars. The resolution of strain gage sensors far exceeds that of LVDT sensors due to the inherent limitations in the LVDT.

2.4.5 Concept Selection

The strain gage sensor was determined to be the best design concept for the HSFD sensing for three reasons. The extra signal processing electronics required for the LVDT sensor presented significantly more complexity then that required for a strain gage sensor. The easy integration into the present bearing design of the strain gage sensor allowed for a simpler geometric design. Finally, the lack of any inherent resolution limits to the strain gage sensor made it possible to reach the desired resolution range for the HSFD.

CHAPTER

3

ANALYTICAL DESIGN

3.1 Introduction

In this chapter, the analytical design process for the HSFD is described with the explanations and descriptions of the engineering models created. A single overall model is created for the HSFD containing sub-models of the main components including bearing, actuator, sensor and electronics. Using this full model, the exact properties (geometry, components, materials, etc.) of the HSFD are determined.

This design process is characterized by a focus on closed-form mathematical modeling of the device parameters of dominant importance. Simplicity is favored in the design process, as it generally enables a closer match between the model predictions and device performance. Accuracy in the modeling when carried out in closed form enables the characteristic equations to be inverted to determine device geometry and other properties as functions of the desired performance. Secondary methods such as finite element modeling are used to predict system performance and iterate to improvements in this performance when it is not possible to make closed form models.

3.2 Prototyping

Several prototypes were made during the design process to investigate various designs and models. Prototype A was created to explore the multi-level sheet flexure design. Prototype B was designed to explore the performance of reluctance actuation, Prototype C was created to explore the control issues with the HSFD. Prototype D was designed and fabricated using the

insights from C for optimal mechanical performance for bearings and sensors. Prototype E was designed but not fabricated as a possible modification for D with increased actuator performance.

3.3 Bearing Design

3.3.1 Method

The bearing design was carried out by generating closed-formed models to predict bearing properties such as stiffnesses, and range limitations. These models were used to explore the available design space as bounded by the various constraints set on the HSFD, such as geometrical constraints and functional requirements. The desired properties were set and where possible, the models were inverted to determine the geometry which would result in those properties.

3.3.2 Choice of Materials

A number of factors must be considered when determining the bearing material composition. Investigation of thin sheet flexures through the fabrication of Prototype A suggested that the thinnest workable material would be around 127 μm (0.005 in.) to ensure the flexural bearing was not damaged during fabrication. Order of magnitude estimates suggests that flexures with thicknesses over 1 mm (0.040 in.) will not fit within the work volume. A range of materials were studied to determine which could sustain the greatest yield strain, which would result in the highest displacement per size. A low Young's Modulus is desired to reduce force output for the actuators. Non-magnetic materials are required to avoid interactions with the actuators. It is also desired that the material is easily machinable, commonly available and has minimal creep. The materials that meet the criteria above are listed in Table 3.1 with their associated properties [34]. The thickness of the materials are listed in the common units of the material supplier, thousandths of an inch.

Table 3.1: Flexural materials comparison.

Material	E GPa	σ_y MPa	ν	G GPa	ε_y με	Thickness Thousandths (inches)
7075-T6 Al	71.7	503	0.33	27	7000	32-40
2024-T4 Al	73.1	345	0.33	28	4700	16, 20, 25, 32, 40
6061-T651 Al	68.9	255	0.33	26	3700	16, 25, 32, 40
1100-H18 Al	68.9	150	0.33	26	2200	5, 7, 10, 20, 32, 40
3003-H12 Al	68.9	124	0.33	25	1800	15, 16, 20, 25, 30, 32, 40
316 Steel	193	230	0.29	80	1200	5, 7, 10, 12, 15, 20, 25, 30, 40

The materials are sorted by yield strain, that being the most important parameter for performance. While 7075-T6 has the highest yield strain, it is fairly limited in its range of thicknesses. Conversely, steel has the lowest yield strain but the greatest range of available thicknesses. Aluminum occupies a middle ground between these two extremes. The materials and their associated thicknesses are used as inputs, starting from the top of the chart, to the finalized models during the bearing optimization process as described in the optimization section below.

3.3.3 General Flexural Design Equations

The modeling of the flexural bearing is based on a differential analysis of the flexural beam, as shown in Figure 3.1.

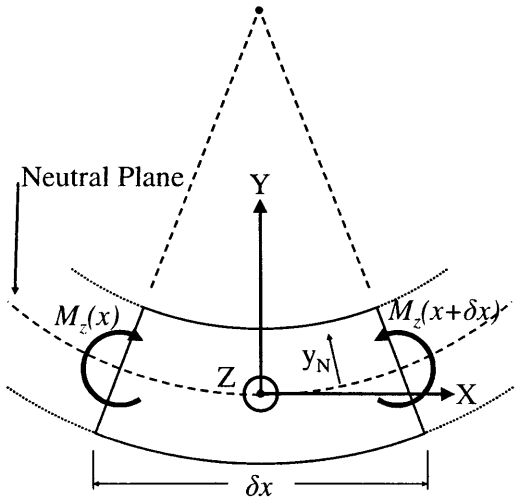

Figure 3.1: Differential segment of beam under pure moment loading.

The flexure is assumed to be undergoing pure moment loading $M_z(x)$, to be thin, isotropic, linear elastic and to be undergoing small deflections. Finally, the initial beam geometry is assumed to be flat and symmetric around the neutral plane. This enables the beam deformation, $y(x)$ to be described using the Bernoulli beam-bending expression [35]:

$$\frac{\delta^2 y(x)}{\delta x^2} = \frac{M_z(x)}{EI_z}, \tag{3.1}$$

where E is the material Young's Modulus and I_z is the second moment of area of the beam. For a beam with rectangular cross-section of thickness (Y-axis in Figure 3.1) h, and width (Z-axis) b, the second moment is

$$I_z = \frac{bh^3}{12}. \tag{3.2}$$

For the flexural bearing design discussed in Chapter 2, the flexures have clamped-guided boundary conditions, as shown in Figure 3.2. The guided end condition simulates the purely vertical displacement of the center stage set at the end of the beam a distance L from the base.

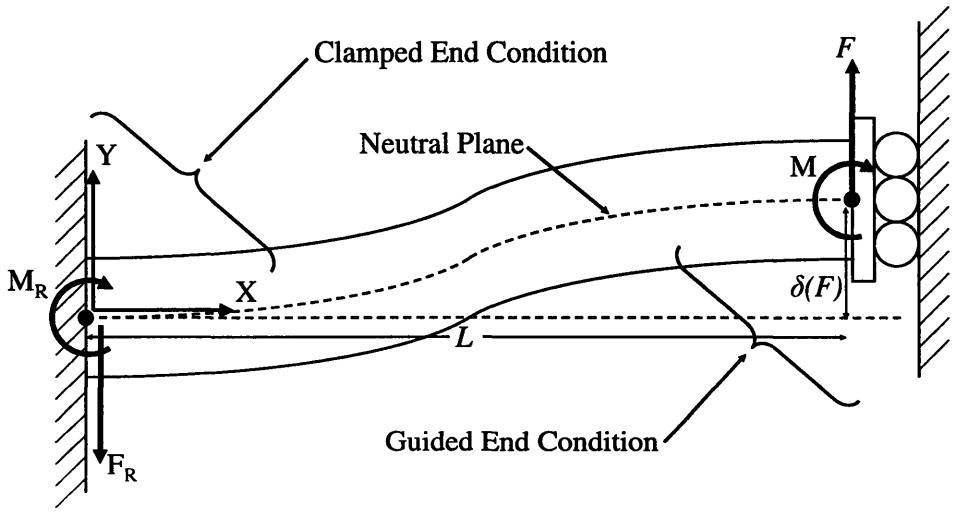

Figure 3.2: Clamped guided boundary conditions.

The boundary conditions of the beam can be characterized by three equations which capture the zero slope behavior of the beam at either end as well as the zero displacement at the base of the flexure:

$$\begin{aligned} y(0) &= 0 \quad \dot{y}(L) = 0 \\ \dot{y}(0) &= 0 \end{aligned} \tag{3.3}$$

Integrating Equation (3.1) along the beam and solving for the integration constants through Equations (3.2) and (3.3) results in a displacement profile for the beam. The displacement $\delta(F)$ at the beam end ($x = L$) is a linear function of the transverse load on the beam F, material properties and the beam geometry:

$$\delta(F) = \frac{FL^3}{12EI_Z}. \tag{3.4}$$

From Equation (3.1) and the differential model in Figure 3.1, the stress $\sigma(x,y_N)$ in the beam, where y_N is distance from the neutral axis, can be defined using the geometric effects of the beam curvature:

$$\sigma(x, y_N) = \frac{M_Z(x) y_N}{I_Z}. \tag{3.5}$$

The max stress σ_{max} will occur at the ends of the flexure, on its surface. This value can be related to the yield stress σ_y through the safety fraction SF, which is derived from Equation (3.5) and the moment distribution within the clamped-guided beam as:

$$SF = \frac{\sigma_{max}}{\sigma_y} = \frac{\frac{FL}{2}\frac{h}{2}}{I_Z}. \tag{3.6}$$

For wide plates ($w/L \geq 1$), the beam observes clamped boundary conditions along the Z-axis in Figure 3.1, which leaves the beam unable to reduce inner Poisson effect stress generated by the tension/compression occurring along the X-axis. The net result is an effective stiffening of the beam which can be characterized [35] by:

$$E_{eff} = \frac{E}{1-\nu^2}, \tag{3.7}$$

where E_{eff} is the material's effective Young's Modulus and ν is the material's Poisson ratio. This is only true when the beam width is on the same order as the beam length.

3.3.4 Linear Stiffness

The vertical linear stiffness of the flexural bearing is calculated through finding the stiffness of the discrete flexures, then summing up the flexures in parallel. The flexures split into two sets, 1 & 2, as shown in Figure 3.3.

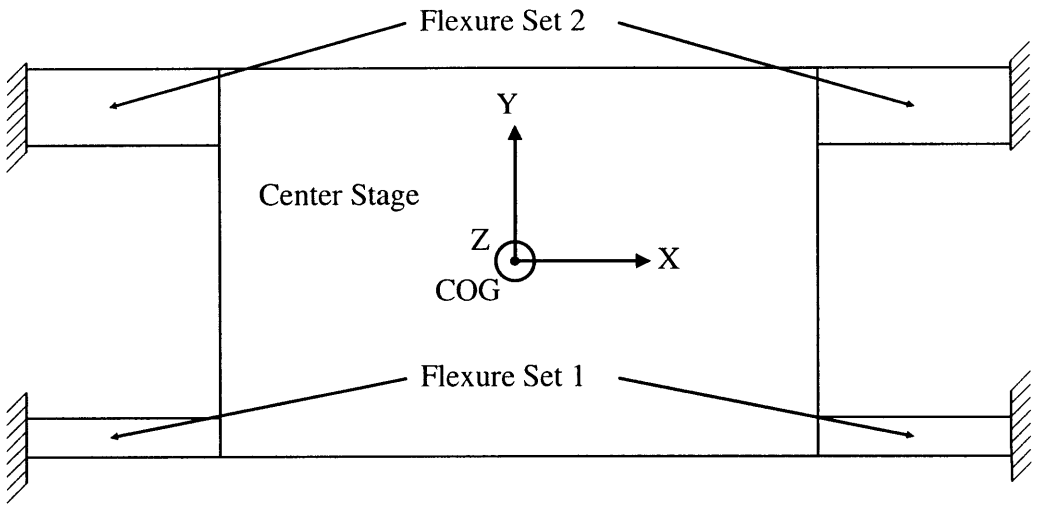

Figure 3.3: Flexural bearing geometric layout.

The center of geometry (COG) is located at the geometric center of the stage, and is used as the origin of the coordinate frame in which the bearing is analyzed. Several assumptions are made in order to calculate the vertical stiffness,: first that the center stage is rigid, second that the outer base is rigid, and third that all of the Bernoulli beam-bending assumptions are met.

Using the expression for the transverse stiffness of a clamped guided beam as shown in Equation (3.4), the vertical stiffness of the center stage k_Z can be written as

$$k_Z = 4\frac{12EI_{Z1}}{L_{f1}^{\ 3}} + 4\frac{12EI_{Z2}}{L_{f2}^{\ 3}}, \tag{3.8}$$

where I_{Z1}, I_{Z2} are the second moment of area and L_{f1}, L_{f2} are the lengths of flexures in sets 1 and 2 respectively. The center stage stiffness in the Y-axis, k_Y, ss calculated through assuming shear to be the dominant mode of deflection, as suggested by the ratio of bending to shear stiffness, resulting in the expression

$$k_Y = 4\frac{(b_{f1}h_{f1})G}{L_{f1}} + 4\frac{(b_{f2}h_{f2})G}{L_{f2}}, \tag{3.9}$$

where G is the material's shear modulus, b_{f1}, b_{f2} are the widths and h_{f1}, h_{f2} are the thicknesses of flexures in sets 1 and 2 respectively. The center stage stiffness in the X-axis, k_X, is calculated through assuming axial stress to be the dominant mode of deflection, as suggested by the ratio of bending to axial stiffness, resulting in the expression

$$k_X = 4\frac{(b_{f1}h_{f1})E}{L_{f1}} + 4\frac{(b_{f2}h_{f2})E}{L_{f2}}. \tag{3.10}$$

3.3.5 Rotational Stiffness

The rotational stiffness of the center stage is calculated through determining the differential in- and out-of-plane motions of the discrete flexures associated with the rotations, then summing up the effective stiffnesses to find the total value.

The method for determining the rotational stiffness for the θ_X-axis is described below, and extrapolated to the θ_Y-axis. The cross section of the center stage with the flexural bearing is shown in Figure 3.4 to illustrate the geometry determining the bearing response to rotations. These rotations are modeled as occurring around an axis passing through the center of stiffness (COS), which sits a distance y_{COS} from the COG.

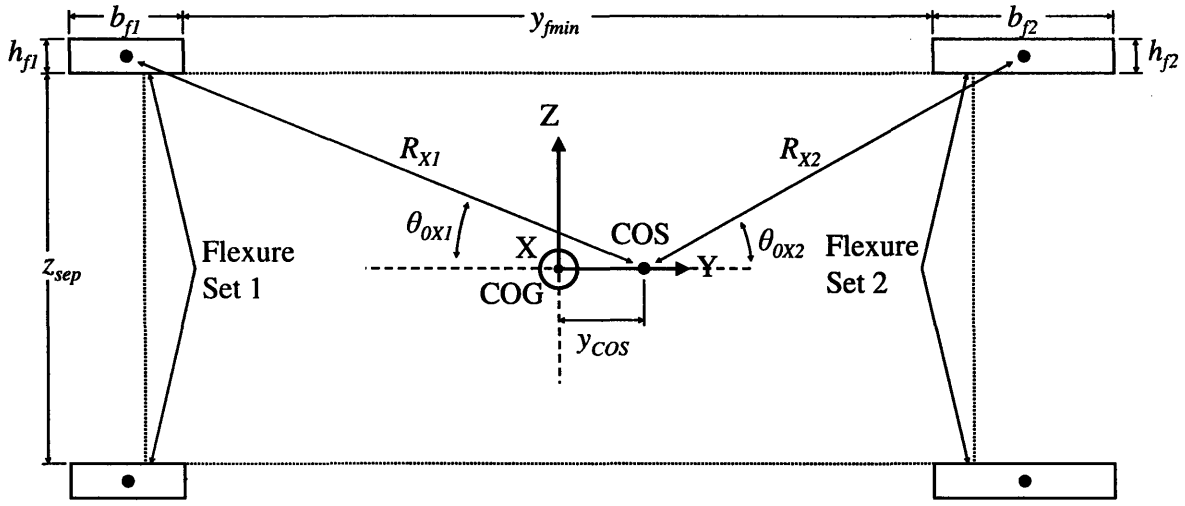

Figure 3.4: Cross section of flexural bearing.

The equilibrium angle of the flexures to the COS, θ_{0X1} & θ_{0X2}, can be calculated from Figure 3.4 above, where z_{sep} is the thickness of the center stage, and y_{fmin} is the separation of the flexures:

$$\theta_{0X1} = tan^{-1}\left[\frac{z_{sep}+h_{f1}}{\left(y_{fmin}+b_{f1}\right)+2y_{COS}}\right]$$
$$\theta_{0X2} = tan^{-1}\left[\frac{z_{sep}+h_{f2}}{\left(y_{fmin}+b_{f2}\right)-2y_{COS}}\right]. \quad (3.11)$$

The distance of each of the flexures from the COS, R_{X1} & R_{X2}, are calculated in the same fashion:

$$R_{X1} = \sqrt{\left(\frac{z_{sep}+h_{f1}}{2}\right)^2 + \left(\frac{y_{fmin}+b_{f1}}{2}+y_{COS}\right)^2}$$
$$R_{X2} = \sqrt{\left(\frac{z_{sep}+h_{f2}}{2}\right)^2 + \left(\frac{y_{fmin}+b_{f1}}{2}-y_{COS}\right)^2}. \quad (3.12)$$

The differential motion of the flexures as a result of small rotations is depicted in Figure 3.5. For small rotations δ_θ, the motion can be decomposed into linear displacements along the Z-axis, δ_Y, and Y-axis, δ_Z.

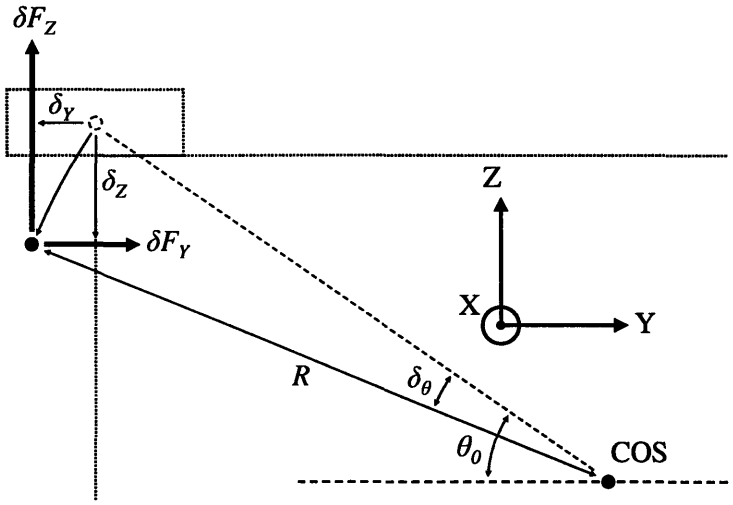

Figure 3.5: Differential model of stage rotation.

The flexural stiffness for each of the displaced directions defines the force, δF_Y & δF_Z, produced by the flexures in set *i* in response to the displacement,

$$\delta F_{Yi} = k_{Yi}\delta_Y = k_{Yi}\cos(\theta_{0Xi})\delta_\theta R_{Xi}$$
$$\delta F_{Zi} = k_{Zi}\delta_Z = k_{Zi}\sin(\theta_{0Xi})\delta_\theta R_{Xi}, \quad (3.13)$$

The forces produced by each flexure can be translated into an effective torque, δT, resulting from the small rotation around the COS:

$$\delta T_i = \delta F_{Yi}\sin(\theta_{0Xi})R_{Xi} + \delta F_{Zi}\cos(\theta_{0Xi})R_{Xi}. \quad (3.14)$$

Combining Equations (3.13) and (3.14) results in the effective rotational stiffness $k_{\theta X}$, for each flexure:

$$k_{\theta Xi} = k_{Yi}\left(\sin(\theta_{0Xi})R_{Xi}\right)^2 + k_{Zi}\left(\cos(\theta_{0Xi})R_{Xi}\right)^2. \quad (3.15)$$

The net center stage rotational stiffness is found by summing up the effective rotational stiffnesses of the four flexures in set 1 and four in set 2, and assuming both center stage and outer base rigidity:

$$k_{\theta X} = 4k_{\theta X1} + 4k_{\theta X2}. \tag{3.16}$$

The rotational stiffness of the θ_Y-axis is defined using the same modeling technique.

3.3.6 Range

The maximum vertical displacement of the center stage, δ_{Zmax}, is calculated from the superposition of the required device range, δ_{Zrange}, and the displacement caused by the settling of the stage under its maximum weight. The weight is defined as the inherent mass of the stage m_{stage}, which is iteratively calculated using CAD models throughout the design process, combined with the maximum sample mass m_{load}, defined in Table 1.1. The center stage operating range is centered on equilibrium to minimize actuator force output, resulting in the maximum displacement expression:

$$\delta_{Zmax} = \frac{\delta_{Zrange}}{2} + \frac{\left(m_{stage} + m_{load}\right) g}{k_Z}. \tag{3.17}$$

3.3.7 Center of Stiffness and Rotation

The flexural bearing has two centers which are important to the design process. The COS and center of rotation (COR) are depicted in Figure 3.6, along with the layout of the actuators. Due to the device symmetry around the Y-axis, the centers are along the Y-axis.

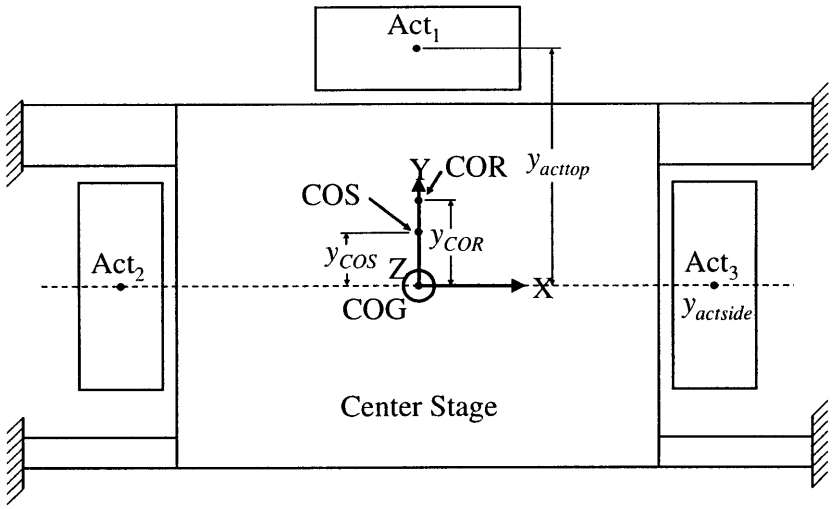

Figure 3.6: Geometric layout of bearing centers.

The COS is the point at which a load applied purely in the Z-axis will produce a displacement only in the Z-axis, with no rotations around the X- or Y-axis. This point is the geometric average of the stiffnesses of the discrete flexures. All of the flexures are the same thickness, thus the stiffnesses can be divided out of the weighted average, resulting in the expression for the Y-axis coordinate of the COS, y_{COS}, as:

$$y_{COS} = \frac{1}{b_{f1} + b_{f2}} \left[\frac{b_{f1}}{2} \left(-y_{fmin} - b_{f1} \right) + \frac{b_{f2}}{2} \left(y_{fmin} + b_{f2} \right) \right]. \tag{3.18}$$

The COR is the point at which resonance rotations of the center stage in both rotational axes cause the least motion along the Z-axis. This point lies at the intersections of the axes around which the center stage rotationally resonates, as shown in Figure 3.7.

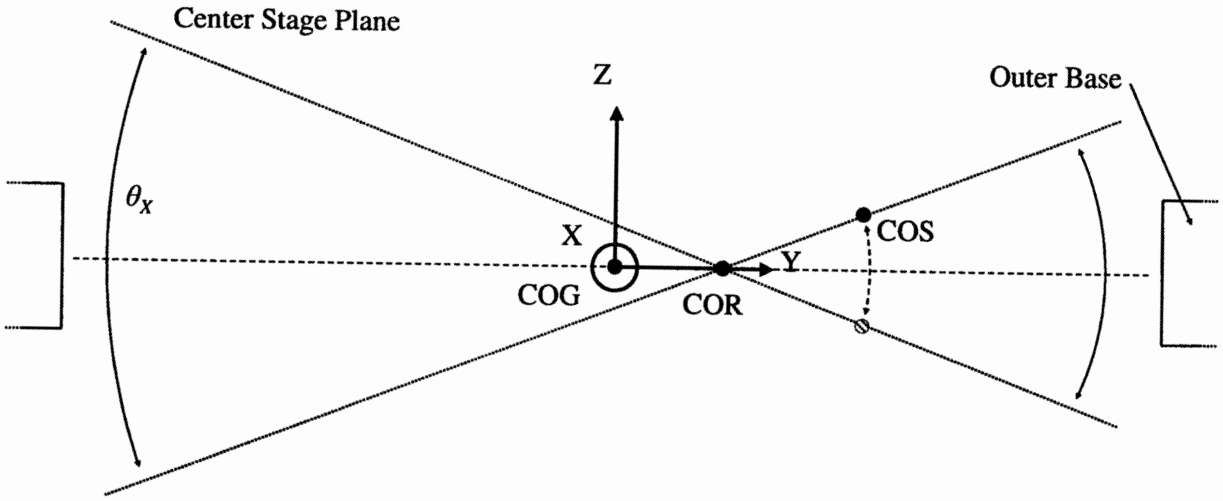

Figure 3.7: Schematic of COR.

3.3.8 Aligning Center of Stiffness and Rotation

The COS and COR are able to interact when dynamic oscillations of the center stage around the COR causes the COS to oscillate along the Z-axis. This vertical motion 'siphons' energy into the vertical axis mass/spring system, causing the Z-axis to begin resonating on its own. The misalignment of these two centers results in an energy leak to other axes, generating parasitic dynamics. If the COS is perfectly aligned with the COR, the only activated energy storage mechanisms during rotational oscillation are the rotational spring and mass of the driven axis, exactly as desired for decoupled behavior of the axes.

It was decided to align the COS and COR of the bearing for best dynamic performance. Initially, it was believed that the COR was a function only of mass distribution, therefore the COS could be adjusted to align with the COR, once the COR location was measured. Alignment of less than 1 mm was determined to be the cutoff for the alignment process. Four iterations were carried out in all, one as prototype C and three as prototype D. The measured locations for the COS and COR were recorded for each iteration and are shown in Figure 3.8.

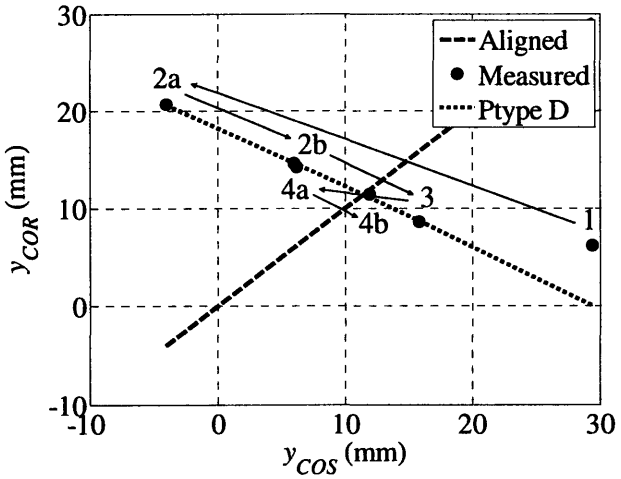

Figure 3.8: COR-COS alignment process.

The goal of the alignment was to produce a flexural bearing whose centers fell within the 1 mm tolerance of the 'Aligned' line. The iteration 1 point was measured on intermediate flexural geometry (prototype C), and first indicated the alignment issue. Iterations 2-4 were measured on the final center stage structure used in prototype D. The added points (2a, 2b, etc.) indicate multiple measurements of the same iteration, differing through incorrect flexural clamping methods (2a, 4a), which had a significant effect on both the COS and COR, and was fixed for later measurements (2b, 4b).

It was observed that the COS and COR are coupled in their location, which explained the inability to align after the first iteration. The fit line passing through all of the measurements of prototype D shows a strong link between the designed COS and the measured COR. The fit line corresponds to a balance where changes in COS are roughly halved and mimicked in reverse by changes in COR around a center point at 11.36 mm. Estimates of the center of mass (COMa) using the CAD model of the device place it at about 13 mm, near to the center point for the COS and COR. The COS and COR vary around what is believed to be the COMa as shown in Figure 3.8. The location of the COS appears to dictate the energy transfer between potential energy (stored in the springs effectively located at the COS) and kinetic energy (stored in the mass effectively located at the COMa), thus setting the COR.

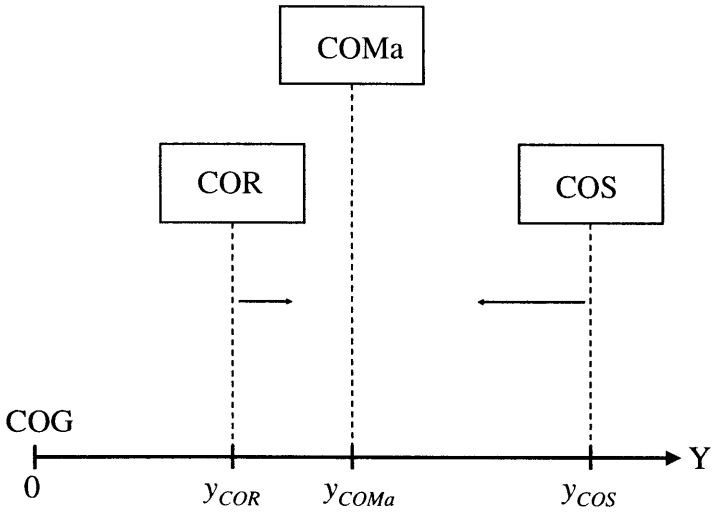

Figure 3.9: Schematic of COR COS coupling.

The desired location of the COS is set to be the true COMa, 11.36 mm. The location of the COS is predicted using Equation (3.18), however this predicted value is consistently less than the location of the actual COS. Observations of the COS show a consistent scale factor of predicted/measured location of 0.847. This was used to scale the scale the design value down to 9.45mm to cancel out the effect. It is believed that this effect is due to the assumption of perfect boundary conditions on the flexures not being met with the present clamping method.

3.3.9 Natural Frequency

The bearing natural frequency in the Z-axis, ω_{nZ}, is calculated in Hertz through the simple expression:

$$\omega_{nZ} = \frac{1}{2\pi}\sqrt{\frac{k_Z}{m_{stage}}}. \tag{3.19}$$

Two constraints are set of the device stiffness. First, one of the HSFD requirements is that the natural frequency be >1 kHz as described in Table 1.1, setting a minimum stiffness. Second, the maximum stiffness is calculated as the ratio of the estimated maximum force output of the actuators (30 N) over half the required stage range. CAD modeling of the design estimated the mass of the center stage at 0.25 kg, which resulted in the first constraint requiring a minimum bearing stiffness on the order of 10^7 while the second constraint sets the maximum bearing stiffness on the order of 10^5. The orders of magnitude separation between these constraints required one of the constraints to be discarded.

The first constraint was determined by the need of the device to operate at greater than 100 Hz, open loop. With closed-loop control, the same system response is possible with a far lower plant natural frequency, allowing the first constraint to be relaxed. Throughout the design process, the mass of the center stage was minimized whenever possible to keep the bearing natural frequency high.

3.3.10 Flexure Anchoring

At large displacements, the kinematics of a clamped-free flexure display parasitic motions corresponding to the flexure tracing out a circular path. The clamped-guided end condition on the HSFD bearing flexures does not allow for any axial translation of the beam end, with the result that the flexures are non-linearly placed under axial loading as the displacement is increased. This effect is akin to the cosine effect length reduction for small angles. It is extremely difficult to model this effect in closed-form, therefore a simple FEA model is used to generate an order of magnitude estimate of the maximum axial loading on the bearing flexures. The model was updated with the flexural geometry throughout the design process. The axial force is about 1.1 N, suggesting that with an aluminum-on-aluminum friction coefficient of about 1 [36] a clamping force of just over 1 N should be sufficient to avoid slip. This value is sufficiently low to suggest that the flexure anchoring may be assumed to be non-slip.

3.3.11 Design Optimization Method

The flexural bearing geometry is dependent on six variables: the length, width and height for each of the flexure sets. These variables, in conjunction with material properties, determine the bearing properties, as laid out in the design equations above. The closed form equations are inverted and solved for the flexural geometry as functions of the desired bearing performance according to the method laid out at the beginning of the chapter. Six equations are required to fully define the six geometry variables. Generally the equations are initially chosen to define all bearing performance parameters of interest. Initial guesses for cross domain variables are set, then the system properties are predicted and constraints are checked to see if the design results lie within the design space. Bearing-performance-setting equations are removed and replaced with an expression forcing the bearing to meet the conflicting constraint as design space restrictions are encountered. The bearing model is linked through the cross-domain variables to

the other models (actuators, sensors, electronics) once the results are determined to be within the design space (or on its boundary).

The set of design equations for the bearing is: flexure thickness set to the stock thickness h_{stock} (2x equations), equal max strain in the flexures (1x equation), COS/COR alignment using Equation (3.18) (1x equation,), and bearing stiffness tuned to actuator force output F_{max} (1x equation),

$$\begin{aligned} h_{f1} &= h_{stock} \\ h_{f2} &= h_{f1} \\ L_{f2} &= L_{f1} \\ y_{COS} &= y_{COR} \\ k_Z &= F_{max}\frac{\delta_{Zrange}}{2} \end{aligned} \quad . \tag{3.20}$$

The max strain equality reduces Equation (3.6) and the definition of Young's Modulus to an equivalence of flexure length, owing to the equal thicknesses of the flexures.

The 6th equation for the design may be any one of three depending on which constraint met; the safety fraction set to maximum safe strain, the length of the flexure set to the length of the attached strain gages, L_{min}, or the width of the flexure set to the width of the attached strain gages, b_{min},

$$\begin{aligned} SF &= 0.25 \\ L_{f1} &= 2L_{min} \\ b_{f1} &= b_{min} \end{aligned} \quad . \tag{3.21}$$

The bearing strain performance is set at its optimal value if the safety fraction equation is used, because high strain readings are desired for the best performance strain gage position sensing. The other two equations define limits to the design space which keep the bearing design from reaching the optimal performance. The flexure needs to be at least twice the length, and at least as wide as the strain gages in order to attach a full bridge for sensing. Either of these conditions become the 6th equation setting the flexural bearing design if they are not inherently met by the design.

3.4 Actuator Design

3.4.1 Method

The actuator design is carried out by generating closed-formed models to predict actuator properties such as max force output, current and power draw. These models are used to explore the available design space as bounded by the various constraints set on the HSFD, such as geometrical constraints and functional requirements. The desired properties are set and where possible, the models are inverted to determine the geometry which results in these properties.

3.4.2 Lorentz Force

Lorentz actuators exploit the interaction between electrons moving through a magnetic field to create a force. This phenomenon scales with the number of electrons exposed to the field, resulting in a force, $\boldsymbol{F}_{Lorentz}$, which is exerted perpendicular to both the current, $\boldsymbol{I}$, flowing along a path of length L through the magnetic field, with magnetic flux density $\boldsymbol{B}$ [37]:

$$\boldsymbol{F}_{Lorentz} = \int_L (\boldsymbol{B} \times \boldsymbol{I}) \cdot \delta l. \tag{3.22}$$

The magnetic flux and current should be made mutually perpendicular to maximize the effect, and the flux density should be driven up as high as possible. This will increase the actuator efficiency in transforming current into force. The magnetic flux should be generated from a permanent magnet rather than another coil to minimize power draw.

3.4.3 Magnetic Circuit Model

A model of the magnetic circuit within the actuator is used to predict and optimize the performance of the actuator. Both the magnets and coils produce magnetic flux, which is analogous to current in electric systems. Gauss's Law of magnetism [38] ensures the conservation of flux over an enclosed surface, S:

$$\oint_S \boldsymbol{B} \cdot \delta \boldsymbol{a} = 0. \tag{3.23}$$

The flux density is not zero over the whole surface, but rather the net value is zero, meaning there are no net magnetic charges, no magnetic monopoles. Some areas $\boldsymbol{a}$ may have an outward flux, however this is balanced by the inward flux over other parts of the surface. The magnetic flux density is driven by a magnetic field, equivalent in concept to current density driven by an

electric field. The flux density is proportional and parallel to the magnetic field intensity $\boldsymbol{H}$, related by the magnetic permeability μ of the material through which the field flows:

$$\boldsymbol{B} = \mu \boldsymbol{H}. \tag{3.24}$$

The magnetic field is governed by Ampere's law [38] which states that closed loop integral, L, of the magnetic field intensity must be equal to the net enclosed current which passes through the (not enclosed) surface S bounded by the loop,

$$\oint_L \boldsymbol{H} \cdot \delta \boldsymbol{l} = \int_S \boldsymbol{J} \cdot \delta \boldsymbol{a} = nI, \tag{3.25}$$

where $\boldsymbol{J}$ is the current density, n is the number of conductors and I is the current through each of the conductors. The integral of the magnetic field intensity over a loop produces a quantity described as the magneto-motive force (mmf), the magnetic equivalent of the electro-motive force, otherwise known as voltage. The discrete form of the current integral is the more commonly used form in engineering, where the mmf for separate segments of the loop are calculated as constant field intensity times the component length, then summed to find the net mmf for the whole loop. This form of Ampere's Law is a simplification from the full form, following the assumption that the electric field is essentially steady in time, therefore there is no displacement current [37].

Equations (3.23), (3.24), and (3.25) can be used like Kirchoff's Current Law, Ohms Law and Kirchoff's Voltage Law in electrical circuit analysis to analyze the system as a magnetic circuit. This enables a relation to be defined between the actuator geometries and the actuator performance.

The geometric layout for the Lorentz coil is designed to guide and constrain the flux as much as possible within magnetically permeable material. This material is referred to as the core, in reference to the common terminology of magnetic design. Unconstrained flux is essentially inefficient, as it does not pass through the current bearing coil and does not contribute to force generation. To predict the performance of the Lorentz coil actuator, a model was created with geometries labeled as shown in Figure 3.10.

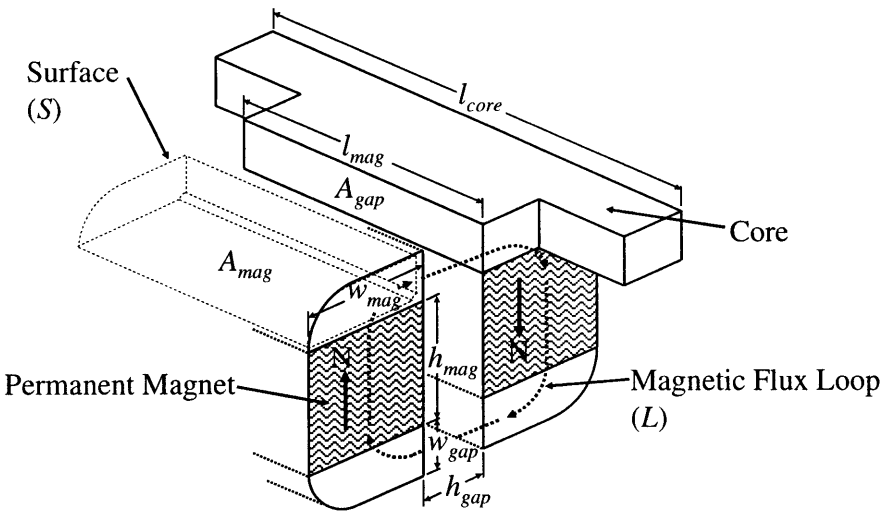

Figure 3.10: Cross section of Lorentz actuator.

The design equation equivalent of Gauss's Law is created by applying Equation (3.23) to a surface, *S*, encompassing one of the cores, assuming no leakage flux:

$$\oint_S \boldsymbol{B}\bullet\delta\boldsymbol{a} = \sum_i B_i A_i = B_{gap}A_{gap} - B_{mag}A_{mag} = 0, \tag{3.26}$$

where B_{gap}, B_{mag} are the magnetic flux densities, and A_{gap}, A_{mag} are the cross sectional areas of the gap and permanent magnet respectively. The design equation equivalent of Ampere's Law is created by applying Equation (3.25) to a closed loop, *L*, and summing up the discrete segments:

$$\oint_M \boldsymbol{H}\bullet\delta\boldsymbol{l} = \sum_i H_i l_i = 2H_{mag}h_{mag} + 2H_{gap}h_{gap} + 4H_{core}h_{core} = 0, \tag{3.27}$$

where H_{mag}, H_{gap} & H_{core} are the magnetic field intensities, h_{mag} h_{gap} & h_{core} are the lengths of the segments as seen by the loop for the permanent magnet, gap and core respectively.

Magnetic flux passes more easily through ferromagnetic materials like the steel core ($\mu \approx 900x10^{-6}$) than air ($\mu = 4\pi x10^{-6}$) for a given magnetic field intensity [37]. The permeability of each section can be observed directly in the design equation by replacing the magnetic field intensity terms in Equation (3.27) using Equation (3.24). The rest of the terms besides the permeabilities are roughly equal. The length of the loop path through the cores is roughly equivalent in magnitude to the length of the loop path through the air gap, and the magnetic flux densities are same order of magnitude. The different permeability of the two components enables the core to be ignored as an insignificant term, reducing Equation (3.27) to the simpler expression:

$$2H_{mag}h_{mag} - 2H_{gap}h_{gap} = 0. \tag{3.28}$$

The final equation needed to determine the magnetic performance of the actuator is the expression describing the behavior of the permanent magnet.

3.4.4 Permanent Magnet Model

The intended use for the permanent magnet is to produce a high-value magnetic flux density, without drawing power, for the actuator coil to push against. The ideal operation of the magnet is to generate the highest magnetic flux density over the largest volume of space, to enable the most coils to be exposed to the flux. This translates to an effective maximum magnetic power output, the magnetic equivalent of maximizing current times voltage. The magnetic flux density over the differential area determines the flux (magnetic current) while the magnetic field intensity over the differential length of the volume determines the field (magnetic voltage) for a differential volume of the air gap, as shown in Figure 3.11.

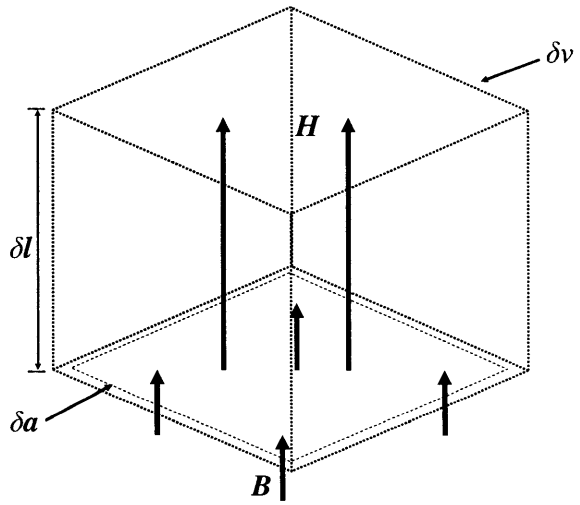

Figure 3.11: Differential model of magnetic power density.

The product of these two terms is referred to as the magnetic energy product [37] and occurs in the form of energy per volume, which should be maximized to make the most effective use of the permanent magnet. The total energy, E_{mag}, stored in the magnetic circuit can be found through integrating the above differential volume over the entire circuit enclosed in volume V:

$$E_{mag} = \int_V \mu |\boldsymbol{B}|^2 \, \delta v. \tag{3.29}$$

The relation between the magnetic field and flux is not the linear one described by Equation (3.24) but rather an 'S' shaped hysteresis curve due to significant remnant flux

generation of the permanent magnetic material. The permanent magnet is magnetized by driving up the magnetic field intensity through the magnet to the far right tip of the hysteresis curve, such that the magnet can not drive any higher flux density through itself- it is saturated. The permanent magnet relaxes back to outputting its remnant flux density, B_r, as shown in Figure 3.12 when the external field is removed ($\boldsymbol{H} \approx 0$). The application of Ampere's Law as in Equation (3.27) shows that the magnetic field intensity of the magnet is drawn negative in order to drive flux through a magnetic circuit. This forces the magnet to operate in the 2nd quadrant of the hysteresis curve defining the B-H relationship in the magnetic material. The slope of the curve, μ_{mag}, describes the permeability of the magnetic material.

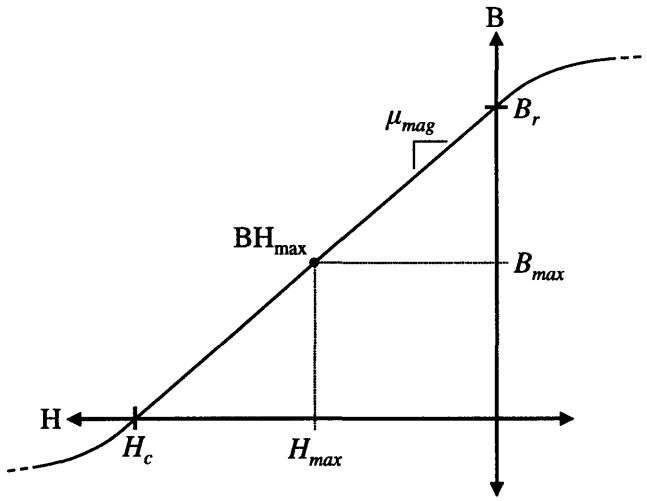

Figure 3.12: B-H curve for rare earth permanent magnet.

The permanent magnet puts out no flux at a sufficiently high magnetic field intensity, which is defined as the coercivity H_c,. The hysteresis curve extends asymptotically in the third quadrant down to a minimum magnetic flux. This 3rd quadrant behavior mimics the 1st quadrant. Upon return to zero magnetic field intensity, the remnant flux is negative. The energy product as described by Equation (3.29) reaches a maximum on the B-H curve at the point BH_{max}, with flux density B_{max}, and field intensity H_{max}. The permanent magnet is operating at maximum magnetic power at this point and produces the most magnetic energy per volume of magnetic material used.

Each type of magnetic material has a characteristic B-H curve. For instance, steel magnets have high remnant flux densities, yet low coercivity, resulting in low max energy

products. Other types of permanent magnets like Alnico have higher coercivities, and thus can put out larger, stronger fields. Rare earth magnets have an energy product which significantly surpasses other permanent magnet materials and have the added benefit of a linear B-H curve which appears as shown in Figure 3.12 [37]. This linear curve can be described by a simple expression relating the magnetic flux density of the permanent magnet, $B_{mag}(H_{mag})$, to the magnetic field intensity, H_{mag}, in the magnet,

$$B_{mag}(H_{mag}) = H_{mag}\frac{B_r}{H_c} + B_r. \tag{3.30}$$

This simple relation enables easy integration of the permanent magnet into closed-form models.

Magnets with the highest energy product were chosen to produce an actuator with the best performance. The best magnets commonly available are rare earth Neodymium-Iron-Boron magnets, type NE52. The type number is a measure of the energy produce in units of mega-Gauss-Oersteds.

3.4.5 Magnetic Circuit Analysis

The magnetic circuit is fully described when the permanent magnet equation (3.30) is combined with the two magnetic circuit design equations (3.26), (3.28). The ideal gap geometry, $w_{gapideal}$ and $h_{gapideal}$, can be written as a function of the desired gap flux density if the permanent magnet is set to run at maximum energy product,:

$$\begin{aligned} w_{gapideal} &= \frac{B_{max} w_{mag}}{B_{gap}} \\ h_{gapideal} &= \frac{\mu_0 h_{mag} H_{max}}{B_{gap}} \end{aligned}. \tag{3.31}$$

The gap flux density can also be described as a function of the circuit geometry:

$$B_{gap} = B_r \frac{1}{\frac{w_{gap}}{w_{mag}} + \frac{B_r}{\mu_0 H_c}\frac{h_{gap}}{h_{mag}}}. \tag{3.32}$$

3.4.6 Multi-Flux Path Model

All possible paths were added to the magnetic circuit model, which is shown in Figure 3.13, to improve its accuracy. These new paths account for leakage flux, both across the gap and back to the other side of the magnet without crossing the gap.

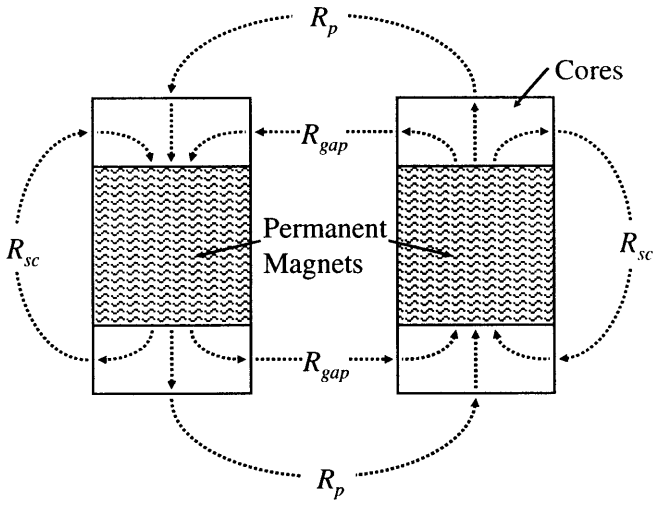

Figure 3.13: Cross section of actuator model with multiple flux paths.

The single flux path model contains only the gap path, R_g. A parallel path, R_p, contains the magnetic flux that does not interact with the coil yet crosses the gap to the core on the other side. A short circuit path, R_{sc}, contains the magnetic flux that jumps back on the main loop to cycle back through the permanent magnet without ever crossing the gap. The physical model described in Figure 3.13 can be rewritten in its electrical equivalent as shown in Figure 3.14. The mmf of the magnet can be considered to be a voltage, which drives current (flux) through the magnetic paths.

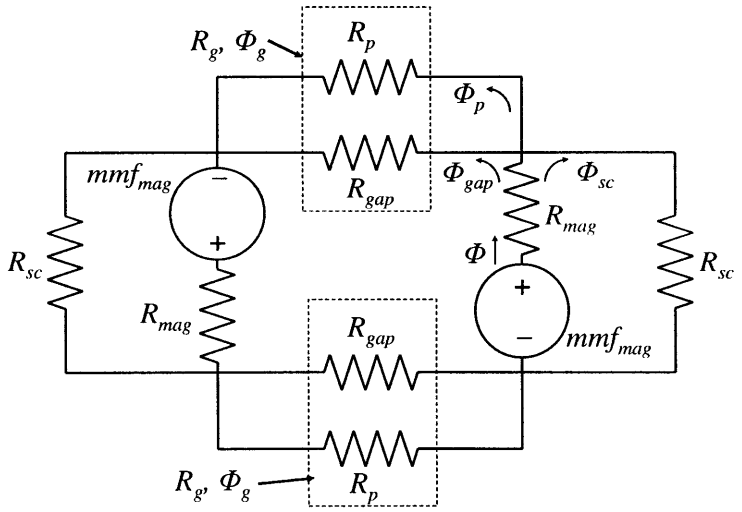

Figure 3.14: Electrical circuit equivalent of multiple flux path model.

The model enables the total flux, Φ, driven by the magnet to be split among the multiple paths, Φ_{gap} , Φ_p , Φ_{sc}, corresponding to the gap, parallel and short circuit paths. Each path can be characterized by a magnetic resistance [37] which is similar to electrical resistance in definition,

$$R = \frac{L}{\mu A}, \tag{3.33}$$

where L is the length of the path, μ is the permeability of the material through which the path travels, and A is the effective cross section of the path. The gap and parallel paths can be combined through summing the resistors into a single equivalent resistor R_g through which Φ_g flux flows.

The permanent magnet is modeled as a battery with a voltage source and internal resistance. These properties can be observed by redefining the permanent magnet equation to be in terms of the magnet mmf and total flux. This is done by multiplying both sides of Equation (3.30) by A_{mag}/l_{mag}, and solving for the total mmf over the magnet. The resulting expression contains a constant mmf term, mmf_{mag}, defined by magnet parameters to be:

$$mmf_{mag} = h_{mag} H_c. \tag{3.34}$$

The expression also contains a flux dependent term, R_{mag}, which acts as an internal resistor:

$$R_{mag} = \frac{H_c h_{mag}}{B_r A_{mag}}. \tag{3.35}$$

Standard circuit analysis enables the electrical circuit diagram in Figure 3.14 to be represented by a system of three equations:

$$\begin{aligned} \Phi - \Phi_{sc} - \Phi_g &= 0 \\ mmf_{mag} - \Phi R_{mag} - \Phi_{sc} R_{sc} &= 0. \\ mmf_{mag} - \Phi R_{mag} - \Phi_g R_g &= 0 \end{aligned} \tag{3.36}$$

Once the three fluxes are found from the system of equations, the gap flux density is calculated through a resistive current divider expression (for resistors in parallel) and dividing the gap flux by area:

$$B_{gap} = \frac{R_p}{A_{gap} R_g} \cdot \frac{mmf_{mag} R_{sc}}{R_{mag}\left(R_{sc} + R_g\right) + R_{sc} R_g}. \tag{3.37}$$

Likewise, the flux density through the magnet is calculated in order to determine the operating point of the magnet,

$$B_{mag} = \frac{1}{A_{mag}} \cdot \frac{mmf_{mag}\left(R_{sc} + R_g\right)}{R_{mag}\left(R_{sc} + R_g\right) + R_{sc} R_g}. \tag{3.38}$$

3.4.7 Effective Path Length

The effective path length of the flux is not directly the distance between the surfaces for paths describing an arc between two parallel surfaces. Rather, the flux is able to travel a number of paths ranging from the direct distance d, all the way to a semi-circular path normal to both surfaces. The range of these paths is illustrated in Figure 3.15.

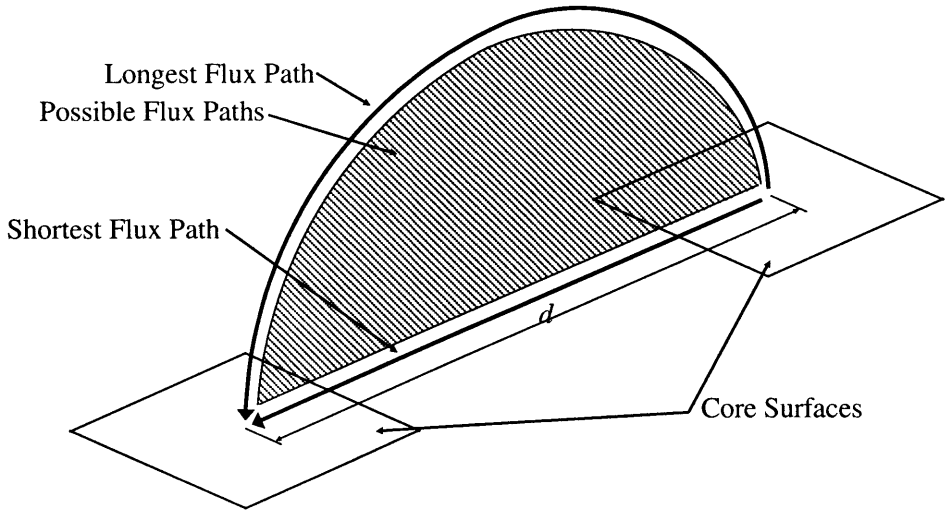

Figure 3.15: Schematic of possible flux paths between parallel plates.

The continuum of paths can be broken up into a large number of discrete paths to find the effective path length between such surfaces, each with a fraction of the cross sectional area, and each reaching some characteristic height from zero to $d/2$. The total resistance of the path is approximately 21% higher than the path resistance if calculated using full cross-sectional area and length d. This length scaling factor, L_{sf}, is used to scale all flux paths in which the start and end surfaces are not facing one another in the resistance definitions below.

3.4.8 Gap Resistance

The gap resistance of the magnetic circuit is composed of a number of path resistances in parallel, each defining a possible path across the gap between the cores composing the actuator. Only symmetric paths were chosen, as the actuator and also the magnetic field generated around it have three planes of symmetry (along each of the axes, X, Y, Z). Additionally, the vector nature of the flux means that overlapping paths will alter the field at that point to be the vector

sum. The net result is that each surface only has one path leaving it. The paths and associated resistances composing the gap resistance are shown in Figure 3.16.

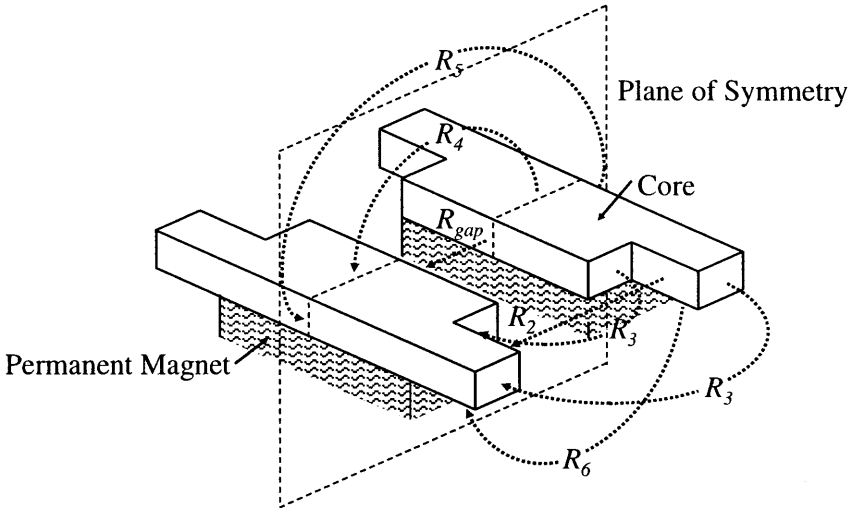

Figure 3.16: Flux paths corresponding to gap resistance.

Each gap resistance path can be defined in the form of Equation (3.35), using the actuator dimensions as laid out in Figure 3.10, to give:

$$R_g = \frac{h_{gap}}{\mu_0 \left(l_{mag} w_{gap} \right)} \qquad R_4 = \frac{L_{sf} \left(h_{gap} + w_{mag} \right)}{\mu_0 \left(l_{core} w_{mag} - \frac{l_{core} - l_{mag}}{2} w_{mag} \right)}$$

$$R_2 = \frac{h_{gap} + w_{mag}}{\mu_0 \left(\frac{l_{core} - l_{mag}}{2} w_{gap} \right)} \qquad R_5 = \frac{L_{sf} \left(h_{gap} + 2w_{mag} \right)}{\mu_0 \left(l_{core} w_{gap} \right)} \qquad (3.39)$$

$$R_3 = \frac{L_{sf} \left(h_{gap} + w_{mag} \right)}{\mu_0 \left(\frac{w_{mag}}{2} w_{gap} \right)} \qquad R_6 = \frac{L_{sf} \left(h_{gap} + 1.5w_{mag} \right)}{\mu_0 \left(\frac{l_{core} - l_{mag}}{2} \frac{w_{mag}}{2} \right)}$$

The total resistance of the gap is the summation of these path resistances in parallel, accounting for the multiple instances of certain paths:

$$R_g = \left(\frac{1}{R_{gap}} + \frac{2}{R_2} + \frac{4}{R_3} + \frac{1}{R_4} + \frac{1}{R_5} + \frac{1}{R_6} \right)^{-1} . \qquad (3.40)$$

3.4.9 Short Circuit Resistance

The short circuit resistance is calculated using the same methods as described for the gap resistance above. All possible symmetric paths composing the short circuit are represented in Figure 3.17.

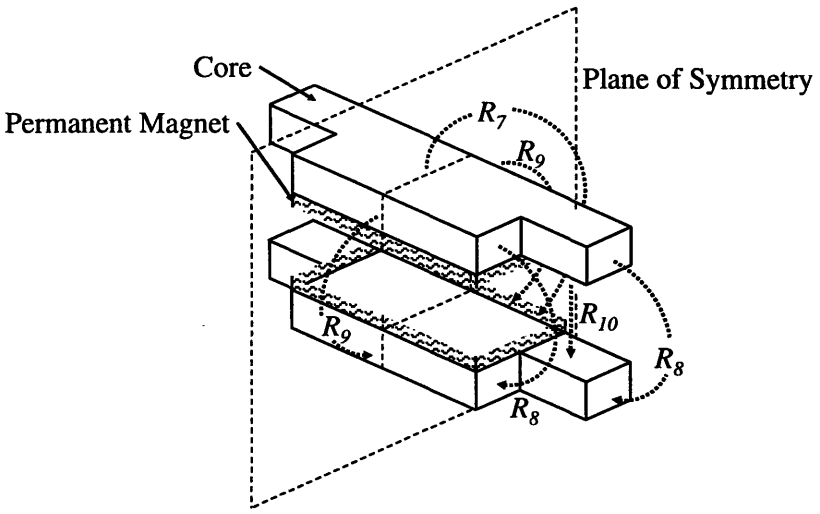

Figure 3.17: Flux paths corresponding to short circuit resistance.

Each short circuit resistance path can be defined in the form of Equation (3.35), using the actuator dimensions as laid out in Figure 3.10, to give:

$$R_7 = \frac{L_{sf}\left(h_{mag}+2w_{gap}\right)}{\mu_0\left(l_{core}w_{mag}-\frac{l_{core}-l_{mag}}{2}w_{mag}\right)} \qquad R_9 = \frac{L_{sf}\left(h_{mag}+w_{gap}\right)}{\mu_0\left(l_{core}w_{gap}\right)}$$

$$R_8 = \frac{L_{sf}\left(h_{mag}+w_{gap}\right)}{\mu_0\left(\frac{w_{mag}}{2}w_{gap}\right)} \qquad R_{10} = \frac{h_{mag}}{\mu_0\left(\frac{l_{core}-l_{mag}}{2}\frac{w_{mag}}{2}\right)} \tag{3.41}$$

The total resistance of the short circuit is the summation of these path resistances in parallel, accounting for the multiple instances of certain paths:

$$R_{sc} = \left(\frac{1}{R_7}+\frac{4}{R_8}+\frac{2}{R_9}+\frac{2}{R_{10}}\right)^{-1}. \tag{3.42}$$

3.4.10 Model Validation

The predicted gap flux density in the multi-flux path model (0.649 T) was compared to the measured flux density of the actuators in prototype C (0.645 T), and found to be in excellent agreement with 0.6% error. The single flux path model predicted a flux density of 1.5 T for the gap, indicating significant accuracy increase from 130% error to 0.6% error in the model predictions due to the addition of the extra flux paths.

3.4.11 Coil Model

The coil which carries current to create the Lorentz force must sit in the generated magnetic field and should drive the current perpendicular to both the magnetic flux and the desired force direction. The actuator coil is designed to sit in the gap between the cores in order to satisfy these conditions, with geometry as shown in Figure 3.18.

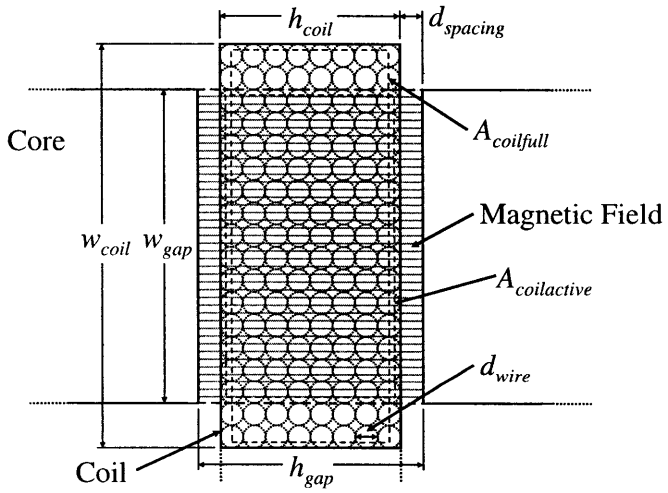

Figure 3.18: Cross section of Lorentz actuator coil.

A small spacing, $d_{spacing}$, is required between the coil and the cores to allow for motion of the coil in response to the generated force. The gap must be large enough to give the coil freedom to undergo rotations with the center stage in addition to vertical translations. The remainder of the area in the gap is assumed to be uniformly filled by the magnetic field, thus is ideal for the coil. This is calculated as the only part of the coil which interacts with the magnetic flux to generate force as a conservative estimate. The height of the coil, h_{coil}, parallel to the flux through this active area is

$$h_{coil} = h_{gap} - 2d_{spacing}, \tag{3.43}$$

and the area of this active region, $A_{coilact}$, is

$$A_{coilact} = h_{coil} w_{gap}. \tag{3.44}$$

The active area of the coil is assumed to be densely packed with magnet wire, a copper wire covered with a thin layer of lacquer insulation having a total diameter d_{wire}. The number of loops which will fit in the available area, $n_{coilact}$, is

$$n_{coilact} = \frac{2}{3} \frac{A_{coilact}}{d_{wire}^{\ 2}}, \tag{3.45}$$

where the packing density is estimated to be about 2/3 from early experiments and wires are assumed to stack in the square grid pattern shown in Figure 3.18. The total length of the wire in the active area, $l_{coilact}$, is calculated as

$$l_{wireact} = 2n_{coilact} l_{mag}. \tag{3.46}$$

To account for the actuator moving through the field, the actual coil width is increased beyond the width of the active area by the total displacement range possible for the coil. This added coil area is not considered part of the active area for force calculations, which ensures that the actuator performance when tested is slightly better than model predictions. The actual effective number of turns in the coil lies somewhere between $n_{coilact}$ and $n_{coilfull}$, indicating that the performance increase will be roughly 0-5%. The full geometry of the coil determines the amount of wire required for fabrication, as well as the power generation and other thermal properties. The full coil width w_{coil}, is defined as

$$w_{coil} = w_{gap} + \delta_{Zrange}. \tag{3.47}$$

The full area of the coil, $A_{coilfull}$, is:

$$A_{coilfull} = h_{coil} w_{coil}. \tag{3.48}$$

The full number of turns in the coil, $n_{coilact}$, is:

$$n_{coilfull} = \frac{2}{3} \frac{A_{coilfull}}{d_{wire}^{\ 2}}. \tag{3.49}$$

The full length of wire, $l_{wirefull}$, required for the actuator is calculated as the sum of two linear sections (active area), and two semi-circular end sections:

$$l_{wirefull} = \left[2l_{mag} + \pi \left(h_{mag} + w_{gap} \right) \right] n_{coil} \tag{3.50}$$

The resistance of the coil, R_{coil}, is calculated directly from the length and diameter of the wire, and is scaled by 25% due to early experiments with the wire:

$$R_{coil} = 1.25 \frac{l_{wirefull}}{\sigma_{cu} \pi \frac{d_{wire}^{\ 2}}{4}}. \tag{3.51}$$

This 25% error may be due to inaccuracies in the wire diameter.

The maximum current required for driving the actuator, I_{max}, is calculated from Equation (3.22) and the stiffness of the bearing determining the force at maximum displacement:

$$I_{max} = \frac{k_z \frac{\delta_{Zrange}}{2}}{3 B_{gap} l_{wireact}}. \tag{3.52}$$

The maximum voltage draw of the coil, V_{max}, is calculated using Ohm's Law as

$$V_{max} = I_{max} R_{coil}. \tag{3.53}$$

The maximum power generation of the coil, P_{max}, is calculated from the maximum current and resistance estimate as

$$P_{max} = I_{max}^{\ 2} R_{coil}. \tag{3.54}$$

The inductance of the coil, L_{coil}, is calculated using the definition of inductance as the ratio of the flux linkage over the current through the coil [37]. The average path length for the flux is assumed to be a circle of height, h_{mag}, resulting in the self-inductance expression:

$$L_{coil} = \frac{n_{coilfull}^{\ 2} \mu_0 h_{mag}}{2}. \tag{3.55}$$

3.4.12 Design Optimization Method

The actuator geometry is essentially dependent on 6 variables: the length, width, and height of the magnet, the width and height of the gap and the diameter of the wire. Several geometric variables such as the core length play a minor role in the actuator performance, but are mainly determined by the geometric layout of the components as described in Section 3.8. These variables, in conjunction with the magnetic material properties determine the coil properties, and thus the full actuator properties. The closed form equations are inverted as described in the bearing design optimization section, to solve for the actuator geometry as functions of the desired bearing performance. Six equations are required to fully define the six geometric variables.

The possible constraints on the actuator design are the maximum sustainable current through the wires, the maximum current output of the power electronics, I_{elec}, the maximum power generation in the coil and the maximum size of the entire actuator in terms of length, l_{max},

width, w_{max}, and height, h_{max}. It was initially assumed that the wire diameter would determine the power generation, however the analysis described for the coil showed that only the total cross-sectional area of conductor matters for the power generation. If the wire diameter is reduced, more turns are used with the net result of the exact same total area of conductor. The wire diameter does determine the number of turns required for fabrication, which suggests that a larger diameter wire will simply the fabrication of the coil. Additionally, it was found to set the maximum current draw of the actuator.

The set of design equations are the geometric constraints on the magnet height, width and length (3x equations), the equations setting the gap geometry as functions of the magnet dimensions (2x equations) as shown in Equations (3.31) and the equation setting the maximum current draw of the actuators to the maximum safe current output of the electronics (1x equation):

$$\begin{aligned} l_{mag} &= l_{max} \\ w_{mag} &= w_{max} \\ h_{mag} &= h_{max} \\ w_{gap} &= w_{gapideal} \\ h_{gap} &= h_{gapideal} \\ I_{max} &= I_{elec} \end{aligned} . \tag{3.56}$$

3.5 Sensor Design

3.5.1 Method

The sensor design is carried out by generating closed-formed models to predict sensor properties such as geometrical layout and strain to voltage transduction. These models are used to explore the available design space as bounded by the various constraints set on the HSFD, such as geometrical constraints and functional requirements. Little internal model analysis is required to determine the location of the ideal design in design space because of the few available design parameters and constraints. Most of the variables are independent and as much, do not require calculating performance tradeoff type optimization to reach the best design. The model is instead used for the multi-domain analysis described in Section (3.7), where it sets geometrical constraints on the bearing design and determines the type of strain gage needed to achieve optimal system resolution.

3.5.2 Strain Gages

Strain gages are used to transform surface strains into variations in an electrical property- resistance- which is easily measured and integrated into mechatronic systems. They are composed of resistors adhered to the surface of interest. The change in resistance is generated mainly through geometric effects, with slight enhancement from material dependent effects (piezoresistivity). A cross section of a strain gage resistor of length l, width w, height h, and area A, under strain loading along the X-axis of ε_X, and along the Y-axis of ε_Y. is shown in Figure 3.19.

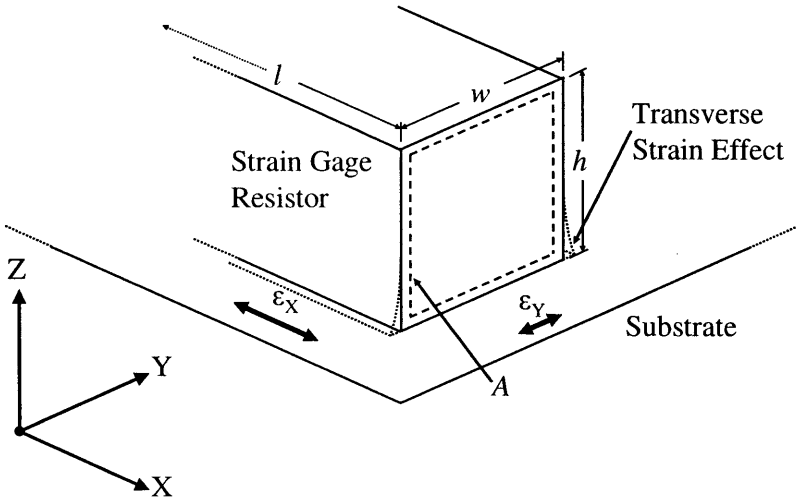

Figure 3.19: Strain gage under loading.

The resistance change in a strain gage, δR, is proportional to the gage factor, G_F, the unstrained resistance, R_0, and the strain ε, on the gage [39]:

$$\delta R = G_F R_0 \varepsilon. \tag{3.57}$$

The change in strain gage resistance may also be caused by temperature changes, which is modeled in the form:

$$\delta R = \alpha_R R_0 \Delta T, \tag{3.58}$$

where α_R is the strain gage resistance temperature coefficient, and ΔT is the change in temperature.

The strain gage response to strain in the X-axis can be examined through a simple analysis of strain based effects on the gage geometry and material properties. The general expression for the strain ε_i along the i^{th} axis in a material [40] can be written as:

$$\varepsilon_i = \frac{\delta l}{l} = \frac{1}{E}\left[\sigma_i - \nu\left(\sigma_j + \sigma_k\right)\right] + \alpha\Delta T, \tag{3.59}$$

Where δl is the differential length gain, l is the unstrained length, $\sigma_{i,j,k}$ is the stress along the i,j,kth orthogonal axes, and α is the temperature coefficient of the material. Even though two axes (X and Y) of the gage resistor are in mechanical contact with the substrate and thus able to transmit strain, the system can be approximated as only being exposed to X-axis strain. The length of the strain gage resistor is significantly longer than its height, meaning that the surface boundary condition of X-axis strain acts on what is essentially a 1-D wire. The strain relaxation that occurs up the height of the resistor is insignificant compared to the X-axis strain based displacement because $l >> h$. The width of the strain gage resistor is on the same order as its height, meaning that the surface boundary condition of Y-axis strain acts on a fully 2-D shape, as shown in Figure 3.19 as the cross sectional area of the resistor. The strain relaxation displacement up the height of the gage is able to counteract the applied Y-axis strain, resulting in a slightly deformed cross section, labeled as Transverse Strain Effect in Figure 3.19. Thus, the Y-axis strain is vastly attenuated, to the point of insignificance to the calculation. This results in a strain field ε_{Xsub} effectively only along the X-axis. Equation (3.60) can be used to find the strain, assuming no temperature changes, occurring along each of the axes of the strain gage, ε_{XSG}, ε_{YSG}, ε_{ZSG} as a result of the applied strain field,

$$\begin{aligned} \varepsilon_{XSG} &= \varepsilon_{Xsub} = \frac{\sigma_{XSG}}{E_{SG}} \\ \varepsilon_{YSG} &= -\nu_{SG}\varepsilon_{XSG} \\ \varepsilon_{ZSG} &= -\nu_{SG}\varepsilon_{XSG} \end{aligned}, \tag{3.60}$$

where σ_{XSG} is the stress in the strain gage along the X-axis, ν_{SG} is the Poisson's Ratio and E_{SG} is the Young's Modulus of the strain gage material. The X-axis stress is found from the X-axis strain, and used to calculate the strain in the other two axes. The definition of the strain gage resistance, R, is

$$R = \frac{l\rho}{A}, \tag{3.61}$$

where ρ is the electrical resistivity of the strain gage material. The differential variation in resistance can be written as

$$\delta R = \frac{\partial R}{\partial l}\delta l + \frac{\partial R}{\partial A}\delta A + \frac{\partial R}{\partial \rho}\delta \rho. \tag{3.62}$$

The differential change in resistivity, $\delta\rho$, is due to the piezoresistivity of the strain gage material, which is marginal in foil resistance gages [32]. It can be approximated as:

$$\delta\rho = \beta\varepsilon\rho, \tag{3.63}$$

where β is the piezoresistive coefficient. The geometrical changes in the resistor, δl and δA, can be written as functions of the internal strains and geometry of the strain gage. The differential change in resistance can be replaced with the strain gage equation (3.57), resulting in the expression:

$$G_F R_0 \varepsilon_{Xsub} = \frac{\rho}{A}\varepsilon_{XSG} l - \frac{l\rho}{A^2}(\varepsilon_{YSG}A + \varepsilon_{ZSG}A) + \frac{l}{A}\beta\varepsilon_{Xsub}\rho. \tag{3.64}$$

Equation (3.64) can be reduced to by using Equations (3.60) to rewrite the internal strains of the strain gage as functions of the applied strain field:

$$G_F = 1 + 2\nu + \beta. \tag{3.65}$$

This expression predicts a gage factor of nearly 2 for an incompressible material, because beta is generally small in foil gages (the kind commercially available). The predicted value is near the quoted value of 2.02 for the most commonly available metal foil gages.

The strain gage is predicted to be entirely insensitive to transverse (Y-axis) strains, as described previously in the discussion of the transverse strain effect illustrated in Figure 3.19. In reality, strain gage sensitivity to Y-axis strain is roughly 0-10% of the sensitivity to X-axis strain [39]. The incomplete attenuation of the Y-axis strain signal is likely due to the end of the zig-zag gage resistance pattern, where the resistor is briefly parallel to the Y-axis, as shown in Figure 2.14 above. This section is usually made significantly wider than the X-axis aligned sections to reduce its resistance, and thus any Y-axis strain effects.

Strain gages also have temperature dependent effects, where both strain gage and substrate heat up, resulting in a change in the strain gage resistance. This situation can be modeled using Equation (3.60), assuming: an external thermal strain field and the strain gage thermal strain. This results in the expressions for the strain gage internal strains:

$$\begin{aligned}
\varepsilon_{XSG} &= \alpha_{sub}\Delta T = \frac{\sigma_{XSG}}{E_{SG}} + \alpha_{SG}\Delta T \\
\varepsilon_{YSG} &= \left[(1+\nu_{SG})\alpha_{SG} - \nu_{SG}\alpha_{sub}\right]\Delta T, \\
\varepsilon_{ZSG} &= \left[(1+\nu_{SG})\alpha_{SG} - \nu_{SG}\alpha_{sub}\right]\Delta T
\end{aligned} \tag{3.66}$$

where α_{sub} and α_{SG} are the thermal expansion coefficients of the substrate and the strain gage respectively. Equations (3.58), (3.62), and (3.66) can be combined to generate an expression linking the strain gage resistance temperature coefficient to physical properties:

$$\alpha_R = (1+2\nu)\alpha_{sub} - (2+2\nu)\alpha_{SG}. \tag{3.67}$$

The net thermal effect is the difference between the two terms, by adjusting the value of the strain gage thermal expansion coefficient, the coefficient can be 'tuned' to be effectively zero for particular materials.

Both thermal and mechanical effects factors into the strain gage's response, thus the characteristic equation for the strain gage resistance, R_{SG}, must include the effects from both of these sources:

$$R_{SG} = R_0(1+G_F\varepsilon+\alpha_R\Delta T). \tag{3.68}$$

The signal of interest is the strain dependent resistance change, however this is only one of two small signals sitting on top of a large DC term. A Wheatstone bridge is commonly used to amplify the signal of interest.

3.5.3 Wheatstone Bridge

Wheatstone bridges are simple electrical circuits designed to allow the accurate measurement of small changes in resistance. The bridge is created from four resistors of roughly equal value $R_{1\text{-}4}$, arranged in two parallel paths of two resistors in series, as shown in Figure 3.20.

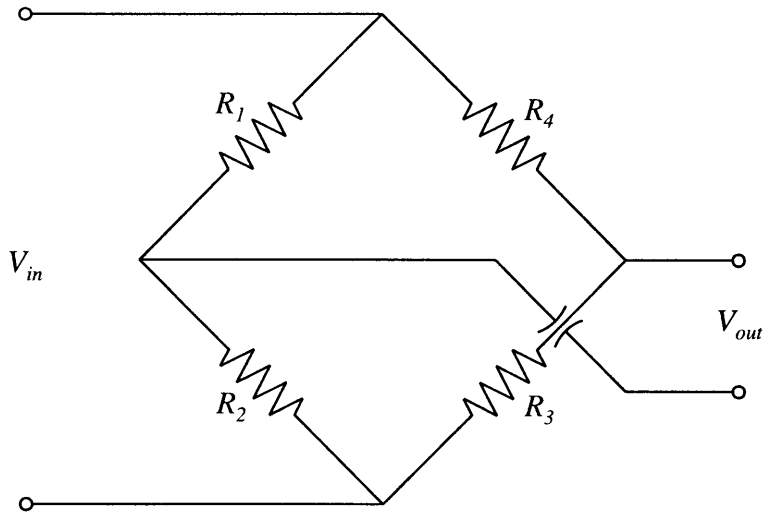

Figure 3.20: Wheatstone bridge circuit model.

The bridge is energized by a steady voltage V_{in}, which creates a voltage drop across the two arms of the bridge. The difference between the voltages at the midpoint of each arm, V_{out}, is measured and used to determine the change in resistance. The output voltage is zero when all of the resistors are of equal value, as the bridge removes the DC term through the differential voltage measurement. The Wheatstone bridge can be characterized by the expression [39]:

$$\frac{V_{out}}{V_{in}} = \frac{R_3}{R_3 + R_4} - \frac{R_2}{R_1 + R_2}. \tag{3.69}$$

Any change in the resistance of the resistors in the bridge acts to offset the voltage on the corresponding bridge arm, unbalancing the two arm measurements and causing the output voltage to shift, largely proportionally to the resistance change. This shift can be boosted and the response made more linear with respect to the resistance change depending on the number of resistors with actively varying resistance. The number of varying resistors determines the bridge type, which ranges from ¼ for only one active resistor, ½ for two and up to a full bridge for four active resistors.

A ¼ bridge contains the fewest active resistors, making it the easiest to place as only one sensor must be located on the strained surface. It also has the lowest strain sensitivity and the largest non-linearity. The ¼ bridge is unable to cancel out thermal effects on the reading assuming the active resistor is the only resistor exposed to strain and also temp changes. The Taylor expansion for the bridge is:

$$\left.\frac{V_{out}}{V_{in}}(\varepsilon,\Delta T)\right|_{1} = \left[\frac{1}{4}(G_F\varepsilon)-\frac{1}{8}(G_F\varepsilon)^2+\frac{1}{16}(G_F\varepsilon)^3+\ldots\right]+ \left[\frac{1}{4}(\alpha_R\Delta T)-\frac{1}{8}(\alpha_R\Delta T)^2+\frac{1}{16}(\alpha_R\Delta T)^3+\ldots\right]. \tag{3.70}$$

A ½ bridge requires twice as many active resistors and has twice the strain sensitivity of the ¼ bridge. The bridge can negate thermal effects on the reading depending on whether the two active resistors are placed on the same arm. The first type of ½ bridge offers the best performance, and involves making resistor 3 an active + response to strain and resistor 4 an active – response to strain. This can be done by placing the sensors on opposite faces of a flexure. The Taylor expansion for the bridge is:

$$\left.\frac{V_{out}}{V_{in}}(\varepsilon,\Delta T)\right|_{2a} = \left[\frac{1}{2}(G_F\varepsilon)\right]+[0]. \tag{3.71}$$

The second type of ½ bridge is still thermally insensitive, but contains higher-order non-linearities. This type is created by making resistor 3 an active + response to strain and resistor 2 an active – response to strain. The Taylor expansion for the bridge is:

$$\left.\frac{V_{out}}{V_{in}}(\varepsilon,\Delta T)\right|_{2b} = \left[\frac{1}{2}(G_F\varepsilon)+\frac{1}{8}(G_F\varepsilon)^3+\frac{1}{32}(G_F\varepsilon)^5+\ldots\right]+[0] \tag{3.72}$$

The third type of ½ bridge is thermally sensitive and contains higher-order non-linearities. It is the worst ½ bridge, but often the only kind possible as only + responses can be obtained from the strain gage sensors. This type is created by making resistor 3 an active + response to strain and resistor 1 an active + response to strain. The Taylor expansion for the bridge is:

$$\left.\frac{V_{out}}{V_{in}}(\varepsilon,\Delta T)\right|_{2c} = \left[\frac{1}{2}(G_F\varepsilon)-\frac{1}{4}(G_F\varepsilon)^2+\frac{1}{8}(G_F\varepsilon)^3+\ldots\right]+ \left[\frac{1}{2}(\alpha_R\Delta T)-\frac{1}{4}(\alpha_R\Delta T)^2+\frac{1}{8}(\alpha_R\Delta T)^3+\ldots\right]. \tag{3.73}$$

A full bridge requires four times as many active resistors and has four times the strain sensitivity of the ¼ bridge. It has no higher-order non-linearities and can negate thermal effects on the reading. This requires all the resistors to be active, with resistors 1 and 3 reading + changes, as well as resistors 2 and 4 reading – changes in response to strain. The Taylor expansion for the bridge is:

$$\left.\frac{V_{out}}{V_{in}}(\varepsilon,\Delta T)\right|_{4} = \left[(G_F\varepsilon)\right]+[0] \tag{3.74}$$

The bridge has two types, a strain type and a thermal type if there is any difference between the set of strain active resistors and thermally active resistors. The first part of the Taylor expansions above describes the sensitivities and non-linearities of that type of bridge in the strain domain, while the second part of each expansion above describes the sensitivities and non-linearities of that type of bridge in the thermal domain.

3.5.4 Sensor Layout

Each discrete flexure in the bearing can have gages on both sides, meaning that at minimum a ½-a or -b bridge can be used for sensing. A full bridge is desired for its higher sensitivity and complete thermal insensitivity. It was determined that the design should focus on a full bridge, with the ½-a bridge as the fall-back design. The full bridge can be laid out on the flexure as shown in Figure 3.21 to produce a positive output voltage when the flexure is deflected upwards.

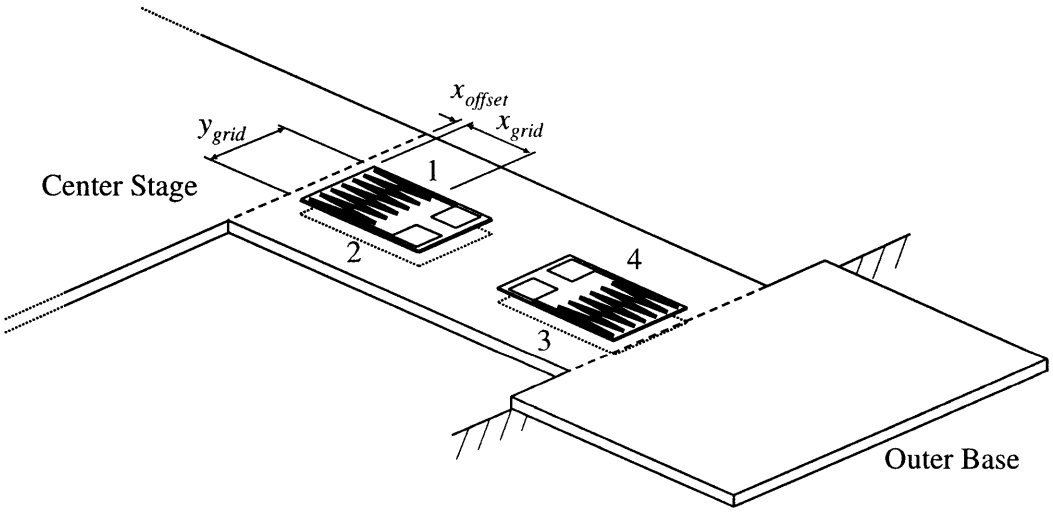

Figure 3.21: Strain gage geometric layout.

There is likely to be a temperature differential over the flexure owing to conductive heat transfer from the actuator into the center stage. Each arm of the Wheatstone bridge in the layout specified above is exposed to a unique temperature that is equally experienced by both resistors on the arm. This means the temperature effects are canceled out, even if the left arm of the bridge (resistors 1, 2) sees a different temperature than the right arm (resistors 3, 4). The bridge

could alternately have each arm on one face of the flexure, but then the temperature differential would cause a change in the common mode voltage of the output.

The strain gages are be placed at the location of highest strain in order to achieve the best displacement sensing. The ideal location is the base of the flexure, as in a clamped-guided beam, the strain increases linearly from center to reach its largest magnitude at either end. The finite length of the strain gage prevents it from being placed exactly at the location of highest strain, therefore the average strain of experienced by the gage must always be lower than the ideal max strain output of the flexure. The strain field is shown graphically in Figure 3.22 with the strain gage resistor grid overlaid to indicate the actual strain experienced by the sensor.

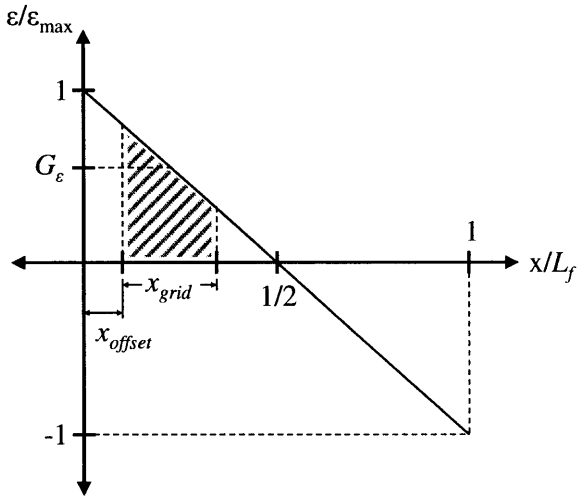

Figure 3.22: Diagram of strain distribution over strain gage.

The effect of the finite geometry of the strain gage can be described by a non-dimensional strain scaling factor G_ε, which is found by integrating over the resistor grid to find the average strain experienced by the gage due to its non-ideal location:

$$G_\varepsilon = \frac{\int_{x_{offset}}^{x_{offset}+x_{grid}} \left(1-\frac{2}{L_f}x\right)\delta x}{\int_{x_{offset}}^{x_{offset}+x_{grid}} 1\,\delta x} = 1-\frac{x_{grid}+2x_{offset}}{L_f}. \tag{3.75}$$

The max strain can be related through the coefficient $F_{\varepsilon Z}$ to the deflection of the flexure using Equations (3.5) and (3.6) as:

$$F_{\varepsilon z} = \frac{\delta\varepsilon}{\delta z} = \frac{3h_f}{L_f^{\,2}}. \tag{3.76}$$

The effective strain, ε_{out}, observed by the strain gage for a given strain gage geometric layout on the flexure can be found as a function of displacement by combining Equations (3.75) and (3.76) to give:

$$\varepsilon_{out}(z) = G_{\varepsilon} F_{\varepsilon z} z. \tag{3.77}$$

The three expected DOFs for the center stage require at least three sensors to measure these motions. More than three sensors would act to average out the readings through repeated measurement of the same system. This would increase resolution but would also increase the complexity of the wiring, fabrication and electronics. A simple model of the measurement as an average value with normal distribution suggests that repeat readings should reduce the standard deviation of the distribution by $1/\sqrt{(n)}$, resulting in relative resolutions as shown in Table 3.2.

Table 3.2: Number of sensors used in position measurement.

Number of Sensors	Normalized Resolution
3	1.0
4	0.87
5	0.77
6	0.71
7	0.65
8	0.61

The fabrication, wiring and electronics cost & complexity would increase proportionally to the number of sensors, but the resolution would only show minor and decreasing gains. For the sake of simplicity, it was decided to use only three sensors.

3.5.5 Zero Balance

The Wheatstone bridge will naturally output a slight voltage when under no temperature or strain loading, due to the slight mismatch in the resistors composing the bridge. This offset voltage is proportional to the energizing voltage, thus can be described by a dimensionless resistance factor, γ, as:

$$\frac{V_{out}}{V_{in}} = G_F \varepsilon + \gamma. \tag{3.78}$$

The bridge can be balanced to remove the offset voltage by adding a tunable resistor, R_{ZB}, to one arm as shown in Figure 3.23, which is then adjusted until it contributes the small resistance required to balance the bridge.

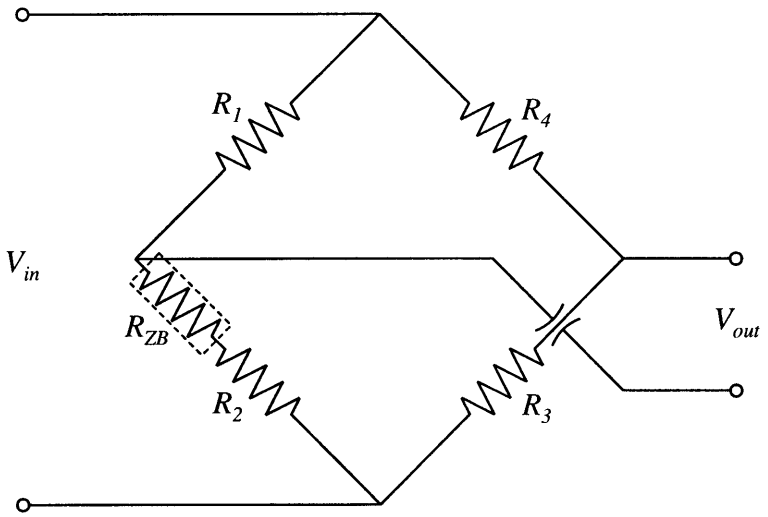

Figure 3.23: Wheatstone bridge with zero balance circuit model.

This balancing is generally difficult or even impossible when using a full bridge, as all the resistors and bridge wiring are located on the strained surface, making it infeasible to add a tunable resistor to the circuit. There are several methods to take care of the zero balance that are described in Section 6.3.10, including attaching a flexible resistance wire to the circuit, removing the bias in the electronics, or jumping out from the bridge to a tunable resistor mounted nearby.

3.5.6 Zero Temperature Compensation

The Wheatstone bridge will generally have some temperature sensitivity when under no strain load, exempting ½-a, -b and full bridges. Zero temperature compensation (ZTC) describes the methods used to reduce or even remove this sensitivity. The bridge will ideally have no temperature sensitivity such that the output voltage will remain only a function of the offset voltage and the strain on the bridge, for highest sensor accuracy.

There are two ways to accomplish ZTC, the first is to use a thermally balanced bridge, such as ½-a, -b or full type. The equal magnitude and sign of the effect is canceled through the circuit design regardless of whether the strain gage sensors produce a thermal reading. Extra resistors can be added in near the strained surface to be exposed to only temperature, but not

strain, which can raise the thermal type of the bridge until it makes the bridge insensitive to thermal effects. The second method for accomplishing ZTC is to use sensors which are tuned to have zero strain gage resistance temperature coefficient, as described in Section 3.5.2 above. This removes any temperature based resistance change, and thus any bridge sensitivity to temperature regardless of the bridge type. Commercial strain gages are only tuned to a select set of materials- steel and aluminum- if the strained substrate is not one of these materials, then this method of ZTC is not available.

3.5.7 Span Temperature Compensation

The gage factor varies slightly when exposed to temperature variation. This generates a false change in the output signal based off a strain input to the sensor. The gage factor (and in the case of force sensing, the material Young's Modulus) ideally can be treated as a constant over all temperatures; however this is not acceptable for high precision analysis. Span temperature compensation (STC) is used to cancel out this thermal-strain cross term effect.

Temperature changes will be canceled through circuit symmetry for a full bridge, however the temperature variations in gage factor (and Young's Modulus for force sensing) will not be addressed. The bridge can be made nearly completely insensitive to temperatures if ZTC is used, depending on whether ZB is also used. This is explained below.

The bridge energizing voltage can be modulated to exactly counteract the material property temperature dependent changes which scale the strain reading,. This is done through the addition of resistors, R_{STC}, in series with the Wheatstone bridge, which have their own thermal sensitivities, as shown in Figure 3.24.

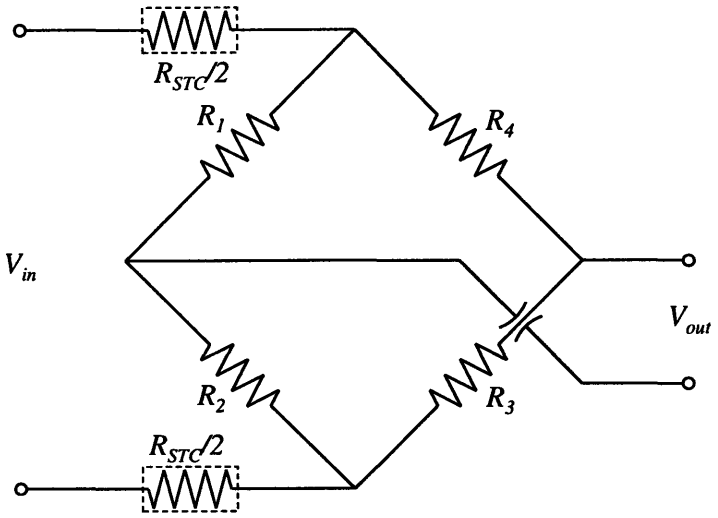

Figure 3.24: Wheatstone bridge with span temperature compensation circuit model.

The voltage divider effect created by the STC resistors results in the modification of the characteristic Wheatstone bridge expression, Equation (3.78), to be

$$\frac{V_{out}}{V_{in}} = (G_F \varepsilon + \gamma)\frac{R_{WB}}{R_{STC} + R_{WB}}, \tag{3.79}$$

where R_{WB}, is the resistance of the Wheatstone bridge. STC is achieved by determining the value of R_{STC} which precisely cancels the thermal-strain effects. The components of Equation (3.79) that do not directly affect the strain-to-voltage term, such as γ, will be ignored in this analysis, as STC is only intended to balance the strain-to-voltage term. The differential variation in the output voltage can be written as:

$$\delta V_{out} = \frac{\partial V_{out}}{\partial V_{in}}\delta V_{in} + \frac{\partial V_{out}}{\partial G_F}\delta G_F + \frac{\partial V_{out}}{\partial \varepsilon}\delta\varepsilon + \frac{\partial V_{out}}{\partial R_{STC}}\delta R_{STC} + \frac{\partial V_{out}}{\partial R_{WB}}\delta R_{WB}. \tag{3.80}$$

The differential variation is defined in terms of temperature variations and set to zero, which forces the bridge to be insensitive to temperature. Each partial differential can be rewritten in terms of temperature variation. The bridge energizing voltage is assumed to be insensitive, as is the actual device strain, because the strain is a result of the flexural geometry alone. Strain is written in terms of the beam stress and Young's Modulus of the flexure only if force, rather than displacement, is the value of interest. The Young's Modulus would bring an associated temperature variation coefficient into the analysis [39]. The gage factor and the span temperature resistor are linked to temperature variations through thermal coefficients α_{GF} and α_{STC}, respectively[39]. The strain gage resistance temperature coefficient is assumed to be

balanced to the aluminum flexure, resulting in effectively zero temperature sensitivity for al practical purposes in the Wheatstone bridge net resistance:

$$\begin{aligned} \delta V_{in} &= 0\,\Delta T \\ \delta G_F &= \alpha_{GF} G_F \Delta T \\ \delta \varepsilon &= 0\,\Delta T \\ \delta R_{STC} &= \alpha_{STC} R_{STC} \Delta T \\ \delta R_{WB} &= \alpha_R R_{WB} \Delta T = 0\,\Delta T \end{aligned} \quad . \tag{3.81}$$

The combination of Equations (3.80) and (3.81) can be used to find an expression for $\delta V_{out}/\Delta T$, which is then set to zero, and STC resistance is found to be:

$$R_{STC} = l_{STC} r_{STC} = \frac{\alpha_{GF} R_0}{\alpha_{STC} - \alpha_{GF}}, \tag{3.82}$$

where l_{STC} and r_{STC} are the length and resistance/length of the STC resistor, respectively. A full resistor of the desired material can be obtained commercially, or resistance wire of the STC material can be purchased, and the wire cut to the required length. The second option enables more accurate STC.

The thermal variation of the STC voltage bridge also affects the offset voltage, which exists only if ZB is not carried out (or carried out with a material of different resistance thermal coefficient than the strain gages), as in the case for the HSFD. The output variation through this mode can be calculated through rewriting Equation (3.80) to include γ, and combining the result with Equation (3.81) as done previously. The result is

$$\delta V_{out} = -V_{in} \frac{\alpha_{GF} (\alpha_{STC} - \alpha_{GF})}{\alpha_{STC}} \gamma \Delta T, \tag{3.83}$$

which shows the thermal sensitivity remaining in the Wheatstone bridge to be of a high order. The temperature coefficient is attenuated by the product of three 'small' terms (≈ 0.001) in the numerator: γ, α_{GF}, and the difference in thermal coefficients α_{STC} & α_{GF}. The numerator only has one small term, giving a net of two small terms in the temperature sensitivity. These terms are described as 'small' because they each act to reduce the input (temperature, voltage) by a factor of about 1000. The difference in thermal coefficients further increases this attenuation, as does the fact that the STC thermal coefficient is much larger than the gage factor thermal coefficient. The result is an extremely small temperature sensitivity for the bridge. The calibration accuracy (strain to voltage coefficient) will hold constant across a range of temperature, but the offset voltage from the sensors will vary by small amounts if the ambient temperature is altered. For

the best sensor accuracy, the sensors should be digitally zeroed whenever the device is brought into a new thermal equilibrium, and ideally the environment should be held constant during operation.

3.6 Electronics Design

3.6.1 Method

The electronics design is carried out by generating closed-formed models to predict electronics properties such as current output for the actuators and resolution for the sensors. These models are used to explore the available design space as bounded by the various constraints set on the HSFD, such as power draw and functional requirements. The models are used to adjust the design parameters in order to maximize the device performance while remaining in the design space bounded by the previously mentioned constraints.

3.6.2 Actuator Current Control

Two options are possible for driving the actuator, either a voltage or a current source. The actuator performance is solely dependent on the current, therefore a voltage source is of less utility. The voltage source will suffer from inaccuracies due to several reasons: the coil resistance changes with temperature, self-inductance of the coil will generate voltage drops reducing current and the back electro-magnetic force (EMF) created by moving the coil through the static magnetic field also generate voltage drops. The current controller (CC) will simply boost the applied voltage to cancel these effects and ensure the current flow through the actuator is maintained at the desired level if a current source is used. The current controller used for the HSFD was designed, built and kindly loaned to the author by Dariusz Golda [41]. The controller functions through an input voltage V_{in}, which produces a command that drives a power op-amp, PA75CX, to run the desired current through the Lorentz actuator. The power op-amp is held to the desired current through closed loop control using the LM741 to lock onto the command voltage, as shown in Figure 3.25.

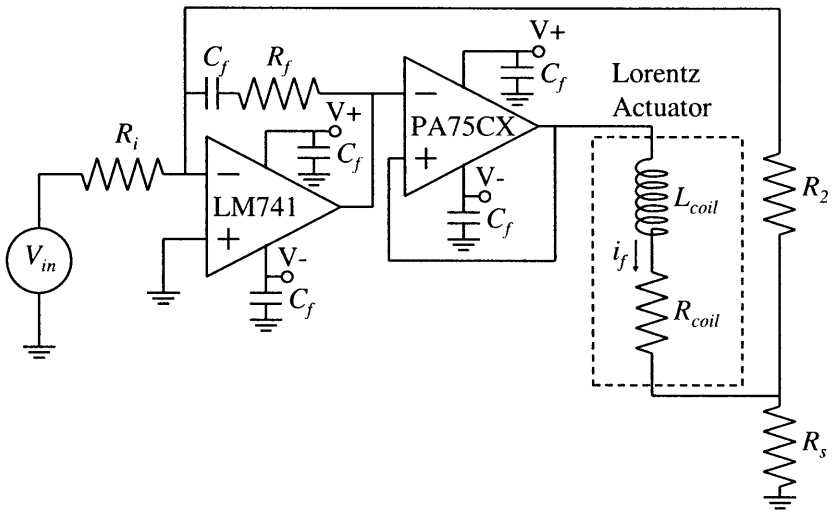

Figure 3.25: Lorentz actuator current controller circuit model.

The current controller can be described by the characteristic equation:

$$\frac{i_f}{V_{in}} = -\frac{R_2 + R_s}{R_i} \cdot \frac{C_f R_f s + 1}{C_f L_{coil}(R_2 + R_s)s^2 + C_f\left[R_{coil}(R_2 + R_s) + R_s(R_2 + R_f)\right]s + R_s}, \quad (3.84)$$

where i_f is the current through the actuator, C_f is the power supply low pass capacitor, and R_i, R_2, R_s, are the resistors used in the circuit. The steady state gain of the current controller is

$$\frac{i_f}{V_{in}} = -\frac{R_2 + R_s}{R_i R_s}. \quad (3.85)$$

The natural frequency, ω_n, of the current controller (in Hz) is

$$\omega_n = \frac{1}{2\pi}\sqrt{\frac{R_s}{L_{coil} C_f (R_2 + R_s)}}. \quad (3.86)$$

The damping ratio, ζ, of the current controller is

$$\zeta = \frac{1}{2}\sqrt{\frac{C_f}{L_{coil} R_s (R_2 + R_s)}}\left(R_2 R_{coil} + R_2 R_s + R_{coil} R_s + R_f R_s\right). \quad (3.87)$$

The system dynamics are analyzed through the pole and zero locations, which are calculated as

$$poles = -\zeta\omega_n \pm \omega_n\sqrt{1-\zeta^2}$$
$$zero = -\frac{1}{C_f R_f} \quad . \quad (3.88)$$

The system may be modeled as purely 2nd order depending on the relative distance of the zero from the origin in a pole-zero plot. The approximate 98% settling time, $t_{settling}$, of such a 2nd order system is

$$t_{settling} = \frac{4}{\zeta \omega_n}. \tag{3.89}$$

3.6.3 Sensor Signal Amplification

The signal leaving the Wheatstone bridge is small, on the scale of 10 mV maximum, yet the analog to digital converters (ADC) read this signal scale over ±10 V. The output voltage from the Wheatstone bridge sensor is boosted before being fed to the ADC to achieve the best resolution. The amplification system used to do this should have gains up to about 1000, as the estimates above show. The system must also have a fast response time of less than 1 ms to meet the function requirements laid out for both the HSFD overall in Table 1.1 and for the sensor in Table 2.7. It must have low non-linearity, cost, noise and a high common mode rejection ratio (CMRR). Typically an instrumentation amplifier (IA) is used for this operation, as it is specifically designed for low noise, high non-linearity and high CMRR. CMRR describes how the system handles the common mode component of the differential voltage input. The signal can be broken into two components- a common mode, V_c, and a differential voltage, V_d. The system ideally rejects the common mode and boosts only the differential voltage by gain G, to generate the Instrumentation Amplifier output, V_{IAout}, as

$$V_{IAout} = G \cdot V_d. \tag{3.90}$$

The circuitry cannot be made perfectly insensitive to the common mode of the signal in reality, therefore the output is described as

$$V_{IAout} = G \cdot V_d + CMRR \cdot V_c. \tag{3.91}$$

The sensor circuit is composed of only a few components- a power supply to drive the Wheatstone bridge strain gage sensor (SG WB) & IA, the WB sensor which produces the signal, the IA which boosts the signal, and the ADC which digitizes the signal. This is shown schematically for all three sensors in Figure 3.26.

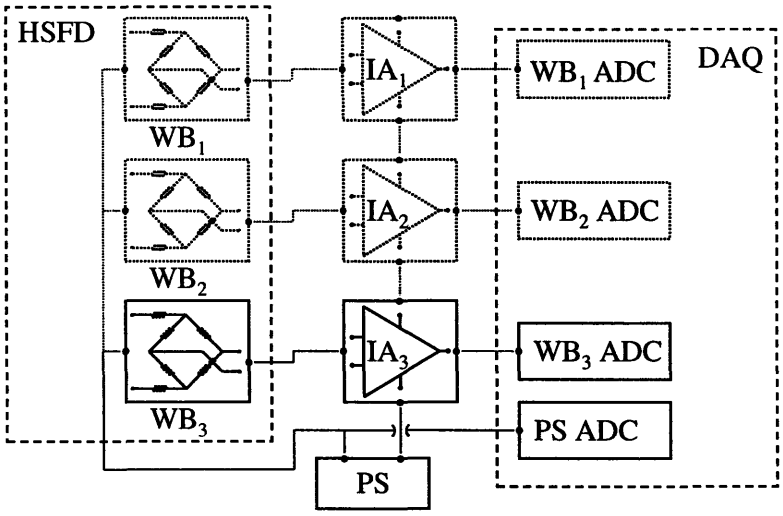

Figure 3.26: Schematic of sensor signal amplification circuitry.

The energizing voltage of the bridge is also measured to account for any variations- which would otherwise appear as strain readings to the system. The gain for the instrumentation amplifier must satisfy the expression:

$$\left|\delta_{Zmax} \cdot F_{\varepsilon z} \cdot G_{\varepsilon} \cdot G_F \cdot V_{in} \cdot G\right| = 10^V. \tag{3.92}$$

3.6.4 Layout

The component setup and wiring for a single sensor channel is laid out as shown in Figure 3.27. Each component of the sensor channel is described in detail below.

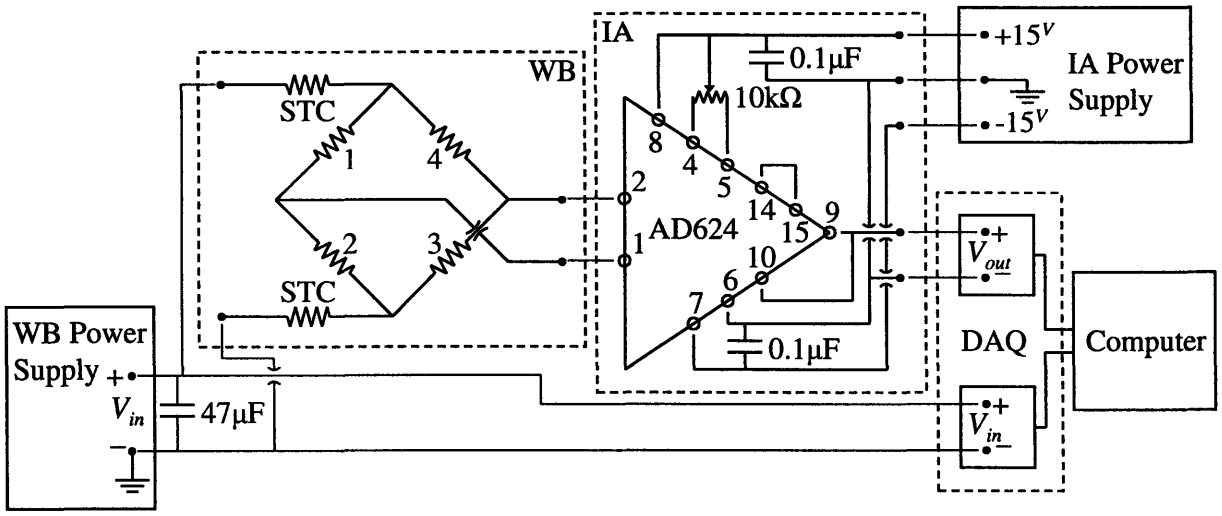

Figure 3.27: Wiring schematic for a single sensor channel.

3.6.5 DAQ board

All signals and commands are transferred between digital and analog by a DS1104a digital acquisition board (DAQ) from Dspace Inc. [42] with 8 input and 8 output channels. The board is linked to a computer through a PCI card. The ADC reads in voltages with 16-bit resolution and the digital to analog converter (DAC) outputs voltage commands with 16-bit resolution.

3.6.6 Wheatstone Bridge

The Wheatstone bridge is driven by a standalone power supply (WB PS) with a low pass filter placed to ground to reduce system noise. The filter is composed of a single 47 μF capacitor. The bridge is energized to a positive voltage and the negative terminal is linked to ground. The energizing voltage is fed to an ADC on the DAQ after the filter to capture the low frequency variation in the energizing voltage and mathematically remove it from the signal. All WB sensors are powered by a single Instek GW GPR-High series DC Power supply.

3.6.7 Instrumentation Amplifier

The output voltage from the SG WB sensor is fed to the IA, which modifies the signal as described by (3.91) to be both boosted and include an attenuated common mode component. The IA runs off a shared power supply (IA PS) which provides ±15 V. Input power supply filters are placed between the energized rails and the PS ground to reduce noise to the IA. The capacitors are placed before the IA such that the current flows (which the capacitor draws to hold the voltage steady) loop back to the PS without affecting the voltage down the line. The current sense input (9) on the IA is used to control the output voltage should the ADC impedance change. The output voltage offset (14, 15) is shorted to reduce added noise. The input voltage offset (4, 5) is tuned with a 10 kΩ potentiometer to zero out voltage offsets added by the electronics. All IA's are powered by a single Instek GW GPC-M series DC Power Supply.

3.6.8 Current Controller

The CC is given signals from the DAC on the DAQ, and the output from each channel is fed to the corresponding coil on the HSFD. The CCs are driven from a separate power supply, as they have hugely varying power draws, which may pull down the rail voltage momentarily,

disrupting any sensitive electronics also drawing power from the rail. All CCs are powered by a single Instek GW GPC-M series DC Power Supply.

3.6.9 Noise Sources

The electronic noise in the system is calculated following the assumption that Johnson noise is the dominant noise type, as confirmed by noise spectrum analysis on prototype C. Johnson noise is thermally generated and constant over all frequencies. Flicker noise is due to quantum variation in material band gaps, and is generally only dominant in microelectronics. Shot noise is related to the statistical nature of current and is also only dominant in microelectronics that utilize small currents on the order of nano-amps or below [43].

Johnson noise is generated by any dissipation mechanism which transforms mechanical or electrical energy into heat. These components generate noise due to small statistical fluctuations in their energy. The noise occurs equally across all frequencies with spectal density $S_n(f)$ [43],

$$S_n(f) = 4k_B TR, \tag{3.93}$$

where k_B is Boltzmann's Constant, T is the absolute temperature of the resistor and R is the resistance of the noise generating component. The voltage noise standard deviation, σ_n, generated from thermoelectric sources (Johnson) is found by integrating the noise spectral density over the full frequency range:

$$\sigma_n^2 = \int_0^\infty |H(2\pi jf)|^2 S_n(f)\,\delta f \tag{3.94}$$

The system in which the resistor is contained has a transfer function, $H(s)$, with an inherent bandwidth that limits the range of frequencies over which the noise is generated.

3.6.10 Electronic Filter

A software single pole low pass filter is used in the sensor channel to reduce noise in the system. The mechanical bandwidth of the system is at most 200 Hz, while the electronic noise has a much larger, theoretically unlimited bandwidth. The filter attenuates the higher frequency signals, which are disproportionately electronic noise. The filter is placed on the feedback loop for each axis' control loop, as will be shown in Figure 3.37. The effect of the addition of the filter is to provide a transfer function with finite bandwidth to the noise calculation in Equation

(3.94). The new expression for noise is calculated following the assumption of a bandwidth f_0, below the Nyquist sampling frequency of the controller:

$$\sigma_n^2 = \int_0^{\infty} \frac{1}{\left(\frac{f}{f_0}\right)^2 + 1} S_n(f)\,\delta f \approx S_n f_0 \frac{\pi}{2}. \tag{3.95}$$

3.6.11 Electronic Noise Model

A noise model is used to track the flow of the displacement signal and noise through a single sensor channel from device motion to sensor measurement result. This model outputs a prediction for the device resolution. The noise contribution of the separate components of the sensor may be analyzed to determine the dominant sources and the effect that design parameters have on these sources.

The standard deviation of the output, σ_Z, of a system characterized by output function Z = f(x1,x2,...), can be calculated through standard propagation of errors as:

$$\sigma_Z = \sqrt{\left(\frac{\partial Z}{\partial x1}\sigma_{x1}\right)^2 + \left(\frac{\partial Z}{\partial x2}\sigma_{x2}\right)^2 + \ldots}. \tag{3.96}$$

The standard deviation of the input variables be predicted using the Johnson noise model described by Equation (3.95) and experimentally validated for better model predictive accuracy. The model assembled to study the noise contributions of the components in the sensor circuit is shown in Figure 3.28.

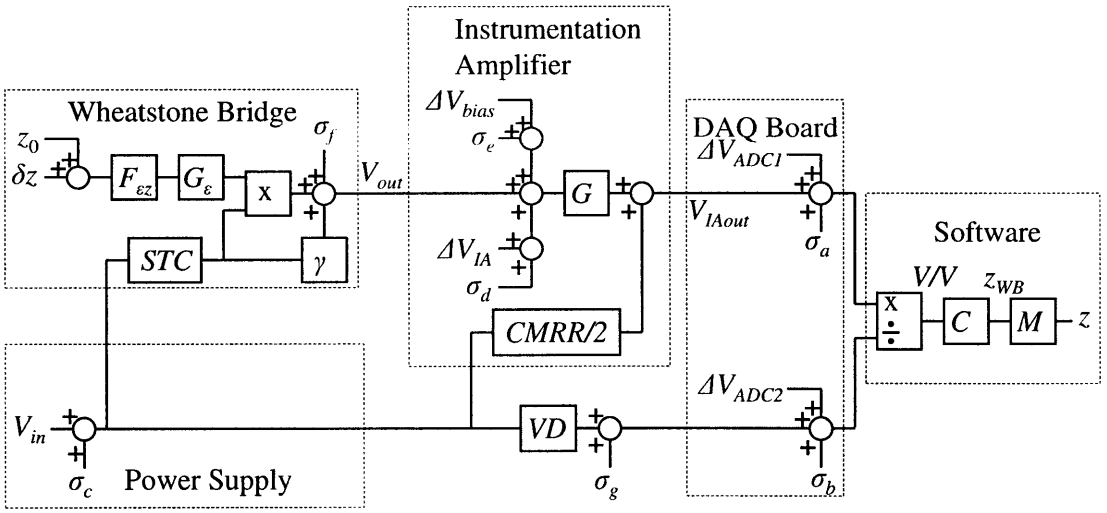

Figure 3.28: Electronic noise model for full system.

Five components are shown in the noise model, the ADC, PS, IA, WB, and software. Each contributes some noise (exempting software) and modulates the signal. The ADCs generate noise σ_a and σ_b, as well as adding voltage offsets ΔV_{ADC1} and ΔV_{ADC2} at their input. The PS generates the bridge excitation voltage which also contains noise σ_c, and is fed to the WB as well as the DAQ board. The PS voltage is attenuated by a voltage divider (VD) which scales the voltage down to a value within the range of the ADC before reaching the DAQ board. This adds noise σ_g to the signal. The IA adds in an inherent bias ΔV_{IA} and noise σ_d at its input. This offset can be adjusted by the tunable input voltage offset ΔV_{bias} which adds noise σ_e to the signal. The signal from the WB is boosted by the gain, and the common mode is added as described in Equation (3.91). The WB reads in an equilibrium offset, z_0, and a displacement δz, which are summed to give the deflection observed by the flexures. The deflection is scaled to strain as described in Equation (3.77). The strain is multiplied by the energizing voltage (which was attenuated by the STC voltage divider), and a voltage offset is added to the system as described in Equation (3.78). The WB also adds noise σ_f to generate the output voltage signal. The software inputs the digital signals from the DAQ board, divides out the measured bridge energizing voltage, and scales the voltage ratio by the calibration C to determine the displacement of the single sensor. The readings from the three sensors are then summed together through M to find the total stage displacement, z.

The model can be simplified if it is assumed that the inputs of the IA can be shorted, allowing a measurement of the inherent electronics offset, which is then negated through tuning the input voltage bias on the IA. Additionally, the effect of ΔV_{ADC2} is only important at resolution levels better than 1 part in 250,000. The resolution limit of the ADC is 16-bits, or 1 part in 16,000, meaning that ΔV_{ADC2} can be ignored without loss of model predictive accuracy. The model is then characterized by the resulting expression:

$$z = \frac{\left[\left[\left(\delta z + z_0\right) F_{\varepsilon z} G_\varepsilon G_F + \gamma\right] STC\left(V_{in} + \sigma_c\right) + \sigma_d + \sigma_e + \sigma_f\right] G + \dfrac{CMRR}{2}\sigma_c + \sigma_a}{\left(V_{in} + \sigma_c\right) VD + \sigma_b + \sigma_c} \cdot C \cdot M. \quad (3.97)$$

3.6.12 Model Constants

The constants used in the model are all defined as functions of the electronics design variable: the bridge energizing voltage. The voltage divider is adjusted to output a reading at the full range of the ADC:

$$VD = \frac{10}{V_{in}}. \tag{3.98}$$

The CMRR is found through the properties of the IA, which was chosen as described in Section 3.7.4 to be the AD624, to be at worst -80dB, more likely -120dB [44]. The adjustment factor to account for the effect of multiple sensors is calculated through Equation (3.96) and (5.8), to give a value of 0.682 (roughly $1/\sqrt{3}$). The same method can used to determine the noise in the rotational measurements simply by using a different row in Equation (5.8). The noise in θ_X can be found to be about 17.4x the noise in z and the noise in θ_Y can be found to be about 12.1x higher. The STC voltage divider is about 1000/1020, accounting for the small resistance of the STC scaling the voltage down over the 1 kΩ bridge. The calibration coefficient is calculated by setting the input displacement equal to the output displacement and assuming no noise or offsets. This requires the calibration to be:

$$C = \frac{VD}{F_{\varepsilon z} \cdot G_{\varepsilon} \cdot G_F \cdot G \cdot STC}. \tag{3.99}$$

The dimensionless resistance factor is estimated to be about 0.0005 based on measurements of prototype C. The IA gain is defined to ensure that the maximum displacement yields an output voltage which is within the voltage range of the ADC, in the form of Equation (3.92):

$$G = \frac{V_{max}}{\left|\delta_{Zmax} \cdot F_{\varepsilon z} \cdot G_{\varepsilon} \cdot G_F \cdot V_{in} \cdot STC\right|}. \tag{3.100}$$

The system constants are added into the model, resulting in an expression for the device resolution which is a function of the device parameters: (including noise), z, and V_{in}.

3.6.13 Model Calibration

The component noise quantities are measured through comparison of the noise in the system with and without the component of interest. The noise is recorded as components were sequentially added, enabling accurate measurements of each of the component noises. No filtering is used in order to read the largest noise possible. The model above is adjusted to reflect the system for each measurement, and the observed position noise is propagated back to the component. The assumption of constant spectral noise density makes it possible to estimate the Johnson noise from each component. These measured values were found to be within about 25%

of the estimated values indicating an adequate fit between theory and underlying physics of the noise.

3.7 Multi-Domain Optimization

The models of each of the components, bearing, actuator, sensor and electronics are integrated into a single overall model of the HSFD system. This multi-domain model contains the system constraints for all of the domains as well as the design equations used to optimize each of the components.

Previous to integration, each of the component models is examined using estimates of cross-domain variables to determine where in the design space the component was operating. The model geometry is set by inverting the closed form models to solve for component geometry as functions of the desired component performance. If any of constraints are encountered, they are substituted into the design equations (at the expense of equations setting geometry as functions of desired performance) to ensure the component is operating at the boundary of its design space for maximal performance. Once all relevant constraints are integrated into the design equations, the model operates along the boundary of maximum performance within its design space. The only remaining unknowns for determining the final parameters of each of the components are the cross-domain variables. The cross-domain variables can be linked to hard equations, rather than estimates by combining all of the component models. The full system performance is studied as a multi-dimensional function of the cross-domain variables. The volume of interest for this function is limited given the earlier estimates as the starting point. Only simple tests are needed to determine the general form of the performance function around this point. The component models, optimization process, and cross domain variables are shown schematically in Figure 3.29. The 'limiting factor' describes the variable of dominant importance, generally the variable which was studied while the converse variable was defined from the dominant variable.

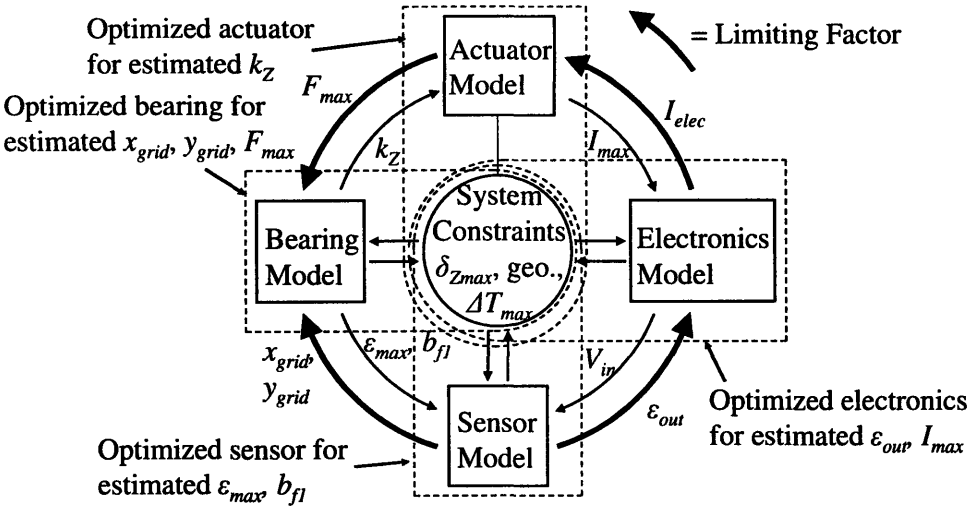

Figure 3.29: Schematic of multi-domain optimization process.

The cross-dimension variables fall into several categories. First, the variables linking the bearing and actuator model are the max force output of the actuator, and conversely the stiffness of the bearing. Setting either of these determines the other. If the max force is raised, the bearing is stiffer, has a higher natural frequency and improved ability to reject rotations, however the power generation grows rapidly (4^{th} power of the boost in natural frequency) setting a limited operating range for adjustment of this variable. Second, the variables linking the bearing and sensor model are the SG grid geometry and conversely the flexure geometry as well as the max strain on the flexure. Commercially available strain gages come in distinct geometries, which influence the size of the flexure and thus the max strain output of the bearing. These different types of gages also have different resistances and max energizing voltage. Third, the variables linking the sensor and electronics model are the effective strain out of the strain gage and the energizing voltage of the bridge. These variables determine the IA gain and system resolution, while strain gages have a limited energizing voltage which must not be exceeded. Fourth, the variable linking the actuator and electronics model is the max current output of the power electronics. This determines the wire gage of the actuator coil.

3.7.1 Bearing Optimization

The bearing optimization is conducted as a cross-domain study to find the optimal system performance over the range of strain gages available. Two strain gages were found with high resistance (low power generation) which were also thermally tuned to aluminum. Their grid

geometry was indicated previously in Figure 3.21, and the associated geometric limits placed on the flexures are shown in Table 3.3.

Table 3.3: Strain gage comparison.

Strain Gage Omega, Inc.	x_{grid} mm	y_{grid} mm	b_{min} mm	L_{min} mm	V_{RMS} V
SGD-7/1000-LY43	7.00	3.6	4.62	11.0	37
SGT-1A/1000-TY43	1.5	2.9	3.92	5.24	15

The transducer quality strain gages (SGT) are designed to be more compact, which generally makes them more difficult to place, but proves to be a large advantage for this application. The minimum width and length of the flexure is calculated assuming the gage excess plastic coating is trimmed and a gap of only 25 mils (0.635 mm) is used on the end and 20 mils (0.508 mm) on the sides.

The performance of the system is shown over a range of stock thicknesses and over the two different strain gages, as shown in Table 3.4. The stock thicknesses are determined from the materials presented in Table 3.1, working down from the top of the chart. Aluminum of type 2024-T4 has highest yield strain over the largest range of geometries and therefore is used in the optimization below. The length of the flexures, the width of the thinner flexure and the max output strain as calculated from Equation (3.77) are displayed as the system output. The 6th constraint used to complete the bearing design equations is listed in each case.

Table 3.4: Strain gage comparison.

Strain Gage Omega, Inc.	h_{stock} mm (mil-inch)	ε_{out} με	L_f mm	b_{fl} mm	6th Constraint
SGD-7/ 1000-LY43	0.406 (16)	390.5	22.1	15.0	$L_{fl} = 2L_{min}$
	0.508 (20)	488.2	22.1	7.64	$L_{fl} = 2L_{min}$
	0.635 (25)	561.9	23.4	4.62	$b_{fl} = b_{min}$
	0.762 (30)	510.8	28.1	4.62	$b_{fl} = b_{min}$
	0.813 (32)	491.4	30.0	4.62	$b_{fl} = b_{min}$
SGT-1A/ 1000-TY43	0.406 (16)	976.1	16.1	5.71	$SF = 0.25$
	0.508 (20)	998.2	18.0	4.07	$SF = 0.25$
	0.635 (25)	847.2	22.2	3.92	$b_{fl} = b_{min}$
	0.762 (30)	722.7	26.6	3.92	$b_{fl} = b_{min}$
	0.813 (32)	682.6	28.4	3.92	$b_{fl} = b_{min}$

The flexural design is always constrained by the size of the SGD strain gage, therefore it is never able to achieve ideal performance. For thin stock, the length limitation is dominant, while at about 0.635 mm (0.025 in.), the constraint becomes the width of the flexures. The system is able to put out at best 562 με. The much smaller size of the SGT strain gage enables the system to operate freely in design space for stock thicknesses under 0.635 mm (0.025 in.). The maximum strain output of 998 με occurs at 0.508 mm (0.020 in.). A stock thickness of 0.508 mm (0.020 in.) and the SGT-1A/1000-TY43 were chosen based on this analysis. The resulting bearing geometry and performance parameters are shown in Table 3.5.

Table 3.5: Flexural bearing design results.

Property	Value	Units
Material	2024-T4 Al	---
h_{f1}, h_{f2}	0.508 (20)	mm (mil-inch)
L_{f1}, L_{f2}	17.97	mm
b_{f1},	4.068	mm
b_{f2}	4.803	mm
SF	0.25	---
SF (max load)	0.36	---
COS (x,y)	(0, 11.36)	mm
COR (x,y)	(0, 11.36)	mm
k_X	80.4	MN/m
k_Y	30.8	MN/m
k_Z	64.3	kN/m
$k_{\theta X}$	3.71	kN/m
$k_{\theta Y}$	1.82	kN/m
ω_{nZ}	81.0	Hz

3.7.2 Actuator Optimization

The actuator optimization is carried out through an analysis of the power generation of the actuator as a function of the permanent magnet geometry. The CAD model of the system sets a length limit on the magnet of 7.62 cm (3 in.); however, the width and height of the magnet can be varied within a small range. A number of discrete geometries were tested to determine which generated the least power, resulting in Table 3.6. The dotted line shows the geometric constraints placed on the permanent magnet, the shading represents the power generation with

higher power darker. The test was carried out slightly beyond the geometric constraints in case the constraints were later altered.

Table 3.6: Lorentz actuator coil power output.

	3.18 (0.125)	6.35 (0.250)	9.53 (0.375)	12.7 (0.500)	15.9 (0.625)	19.1 (0.750)	w_{mag} mm (in.)
3.18 (0.125)	1300	230	84	42	25	16	
6.35 (0.250)	260	46	17	8.2	4.7	3.0	
9.53 (0.375)	120	22	7.9	3.9	2.2	1.4	
12.7 (0.500)	69	13	4.8	2.4	1.4	0.86	
15.9 (0.625)	46	8.8	3.3	1.6	0.93	0.60	
19.1 (0.750)	33	6.5	2.4	1.2	0.70	0.45	
h_{mag} mm (in.)							Power W

The power generation follows a clear pattern of decreasing towards the bottom right, suggesting a large square cross section is desired for the magnet. Based on this analysis, the size of the actuator is maximized to its geometrical constraints. The maximum height is determined by the thickness of the outer base of 12.5 mm (0.5 in.), while the maximum width is determined by the available space between aperture and envelope. The geometry of the permanent magnets (12.5*x*76.2*x*76.2 mm or 0.5*x*0.5*x*3 in. with transverse magnetization) requires custom magnets, which were ordered through Indigo Instruments [45].

The max current draw of the actuator is set by the power op-amp used in the current controller, which is limited to about 1.25 A output. For safety, a maximum value of 0.5 A was used. This sets the wire to be 26 AWG magnet wire, with a diameter of 0.406 mm (0.016 in.). The resulting bearing geometry and performance parameters are shown in Table 3.7.

Table 3.7: Lorentz actuator design results.

Property	Value		Units
	Channel 1	Channel 2&3	
l_{mag}	76.2 (3.00)		mm (in.)
h_{mag}	12.7 (0.500)		mm (in.)
w_{mag}	12.7 (0.500)		mm (in.)
h_{gap}	5.08 (0.200)		mm (in.)
w_{gap}	5.28		mm
B_{gap}	0.729		T
h_{coil}	4.57 (0.180)		mm (in.)
w_{coil}	5.79		mm
l_{core}	108 (4.25)		mm (in.)
$d_{spacing}$	0.254 (0.010)		mm (in.)
d_{wire} (26 AWG)	0.406		mm
Magnetic Material	N52		---
$l_{wirefull}$	20.3		m
$l_{wireact}$	14.9		m
$n_{coilfull}$	107		turns
R_{coil}	3.25		Ω
L_{coil}	90.9		μH
I_{max}	0.173	0.490	A
V_{max}	0.561	1.59	V
P_{max}	0.096	0.782	W

The multi-flux path model was later used to further refine the actuator optimization process as described in Appendix A. The refined actuator model was not fabricated due to time constraints on the project and the difficulty of fabricating coils. A CAD model was generated for an updated HSFD prototype (E) utilizing the optimized actuators.

3.7.3 Sensor Optimization

The sensor optimization is carried out in conjunction with the bearing design as described earlier. To maximize the sensor readout, a full bridge is used on the bearing flexures. The gages are placed 0.508 mm (0.025 in.) from the ends of the flexures to produce the highest strain reading. The geometry of the SGT-1A/1000-TY43 gage combined with the offset enables the strain factor to be calculated at 0.846. The STC resistance was calculated to be 20.4 Ω of SGB-36 Balco resistance wire (19.7 Ω/ft [39]), purchased from Omega, Inc. The Balco wire is separated into two sections of 6.25 inches for each of the two STC resistors. The thermal coefficients of the gage factor ($9.0x10^{5}$) and the Balco wire ($4.5x10^{-3}$) were obtained from the

manufacturer [39], Omega Inc. The resulting sensor geometry and performance parameters are shown in Table 3.8.

Table 3.8: Strain gage Wheatstone bridge sensor design results.

Property	Value	Units
Sensor	SGT-1A/1000-TY43	Omega Inc.
x_{grid}	1.5	mm
y_{grid}	2.9	mm
x_{offset}	0.635 (0.025)	mm (in.)
G_F	2.02	---
R_0	1.0	kΩ
G_ε	0.846	---
ε_{out}	998.2	με
Span. Temp. Comp.	SGB-36	Omega Inc.
α_{STC}	0.45	%/°C
α_{GF}	0.0090	%/°C
r_{STC}	64.6	Ω/m
l_{STC}	2x159 (6.25)	in. (mm)
R_{STC}	20.4	Ω

3.7.4 Electronics Optimization

The electronics optimization is carried out in two relatively independent steps. The current controller design is adjusted to produce the desired actuator performance, and the sensor circuitry is adjusted to produce maximum resolution.

The current controller output is scaled to drive the stage to maximum displacement at 75% of the maximum current output. The extra space in the command range is left to enable dynamic commands to not hit saturation even at max displacement. The required currents are determined from prototype C, which contained the same actuators as prototype D. Only R_i is varied between the channels to produce the different gains. The components used in the current controller are described in Table 3.9.

Table 3.9: Current controller component values.

Component	Value		Units
	Channel 1	Channel 2&3	
R_2	100	100	kΩ
R_s	10	10	Ω
R_{coil}	3.25	3.25	Ω
R_i	432	150	kΩ
R_f	100	100	Ω
C_f	100	100	pF
L_{coil}	91	91	μH

The resulting properties of the current controller are shown in Table 3.10.

Table 3.10: Current controller properties.

Property	Value		Units
	Channel 1	Channel 2&3	
Gain	-0.023	-0.067	---
ω_n	16.7	16.7	kHz
ζ	0.695	0.695	---
poles	-72.9±75.4	-72.9±75.4	krad/s
zero	-100	-100	Mrad/s
$t_{settling}$	54.9	54.9	μs

The system dynamics can accurately be characterized as a second order system because the zero is much greater than 5x further out than the poles. The settling time for the system is significantly (≈100x) faster than the 5 ms ideal settling time of the stage, suggesting that the current controller dynamics will play an insignificant role in determining the system response.

The bandwidth of the low pass filter is set to be 1 kHz, 10x higher than the stepping speed of the mechanical system (100 Hz) to minimize the attenuation of mechanical signals. The frequency is as low as possible to reduce noise, as described earlier.

Two main choices were found for the instrumentation amplifier in the sensor circuits, AD624 [44] and AD8221 [46]. The properties of the two instrumentation amplifiers are shown in Table 3.11.

Table 3.11: Instrumentation amplifier comparison.

Property	AD624	AD8221	Units
Noise	4	8	nV/rt(Hz)
Bandwidth	20	10	kHz
Cost	50	10	$

The cost of the AD8221 is significantly less due to sacrifices in the noise floor of the device and a slight decrease in the operating bandwidth. Cost estimates for the HSFD carried out after fabrication are shown in Appendix B, and indicate that the added cost for the AD624 does not significantly alter the device cost. This negated the main benefit of the AD8221, resulting in the choice of the AD624.

The system noise model was used to study the effect of varying the bridge energizing voltage between 10 and 20 V. The results of the study are shown in Figure 3.30.

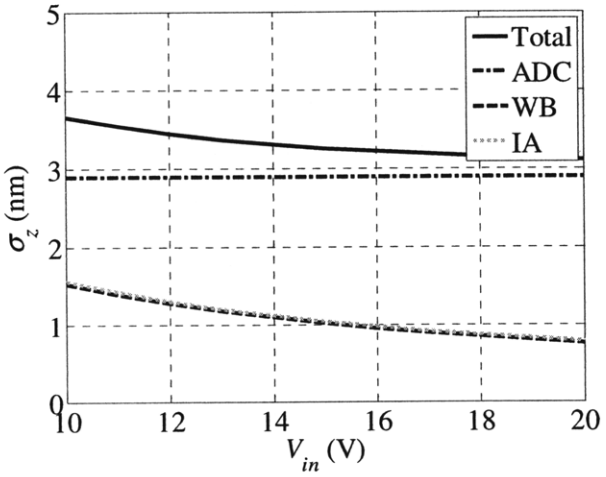

Figure 3.30: Effect of energizing voltage on predicted electronic noise.

The total noise is estimated to be about 3-4 nm, mainly generated by the noise at the ADC. The two other main noise sources, WB and IA, are reduced by boosting the energizing voltage, however they play a secondary role to the ADC and little benefit is gained from boosting the energizing voltage. If the ADC noise is reduced through purchasing a higher resolution (18 or 20 bit) converter, large gains may be accrued in the system resolution, and the energizing voltage would significantly attenuate noise. The thermal effects of boosting the energizing voltage were considered to not be significant, therefore despite the low return, the

energizing voltage is set to be about 20 V. The resulting electronic components and performance parameters are shown in Table 3.12.

Table 3.12: Electronics design results.

Property	Value	Units
V_{in}	20	V
Ins. Amp	AD624	Analog Devices Inc.
G	186.5	---
Filter Bandwidth	1.0	kHz

3.8 Geometric Design

The HSFD components are laid out according to Figure 3.31, using the designed parameters for the bearing and actuators.

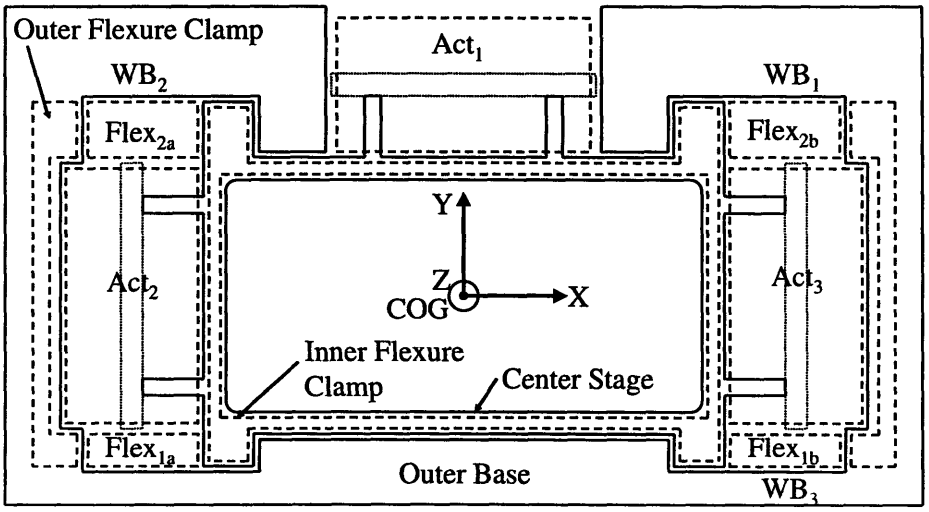

Figure 3.31: Component geometric layout.

The center stage and outer based are restricted by the geometric envelop of the aperture and device. The three actuators are fit in around the stage, nested between the discrete flexures of the bearing. The actuator cores and permanent magnets are anchored to the outer base, while the actuator coils are anchored to the center stage. The cores contain flanges, as shown in Figure 3.10 previously, through which mounting screws are run to anchor the cores to the outer base. The compressive force generated by anchoring the cores on both sides holds the permanent

magnet at the desired location. A 0.254 mm (0.010 in.) steel shim is placed between one of the cores and the magnet to ensure that the cores are clamped down properly on the permanent magnet. The clamping method is illustrated in Figure 3.32.

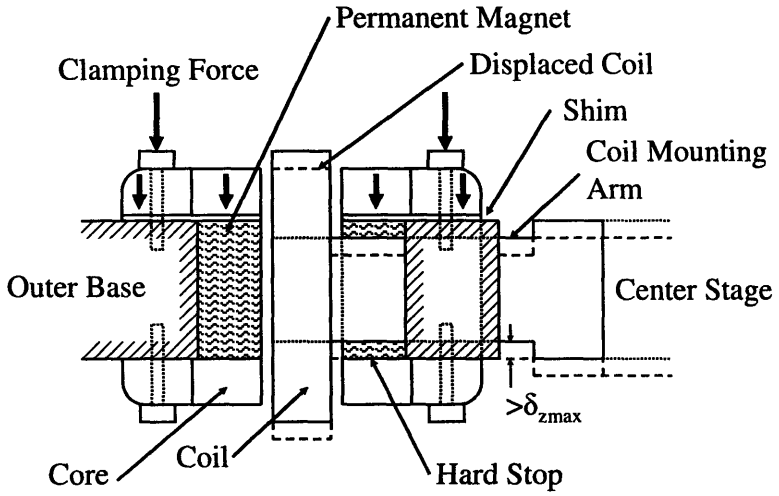

Figure 3.32: Actuator mounting and hard stop.

Clearance holes for the clamping screws enable the actuator to be adjusted in-plane to align to the actuator during fabrication. The core flanges serve a second purpose as hard stops for the center stage. The coil mounting arms are machined down by 0.254 mm (0.010 in.) on the same side as the shim, and 0.508 mm (0.020 in.) on the opposite side. When the center stage is displaced by about 500 μm in either direction, the coil mounting arms come into contact with the cores, limiting any further displacement which might damage the flexures. The range of 500 μm is chosen as the sum of the maximum travel (250 μm), maximum weight settling (130 μm) and a roughly 25% dynamic overshoot for the largest possible step (125 μm).

The flexure sheets are clamped over most of their area to immobilize the sheet except for the small regions where a discrete flexure is desired. The center stage end of the flexure is anchored with a much thicker 'inner flexure clamp' which is intended to spread the pressure cone of the flexure clamping screw such that the entire end of the flexure is tightly clamped. The outer ends of the flexures are anchored with the 'outer flexure clamp'. The four possible outer flexure clamp locations are integrated into two pieces to simplify assembly and alignment.

3.9 Thermal Design

3.9.1 Method

A thermal model of the HSFD is used to determine the thermal profile of the device based on the geometric design and component parameters as described above. A closed form model is used identify the dominant thermal flows within the device, and a FEA simulation is used to confirm the model. The FEA model was tuned to match the measured thermal profile of prototype C based on the dominant parameters identified in the closed form model.

3.9.2 Closed Form Model

The layout of the thermal flow is assumed to mainly be from the coil to the center stage and from the center stage to the outer base. Losses occur at each step to the environment. The strain gage sensors also generate heat, which flows out of the flexures into both the center stage and the outer base. The full closed form model can be visualized schematically as in Figure 3.33.

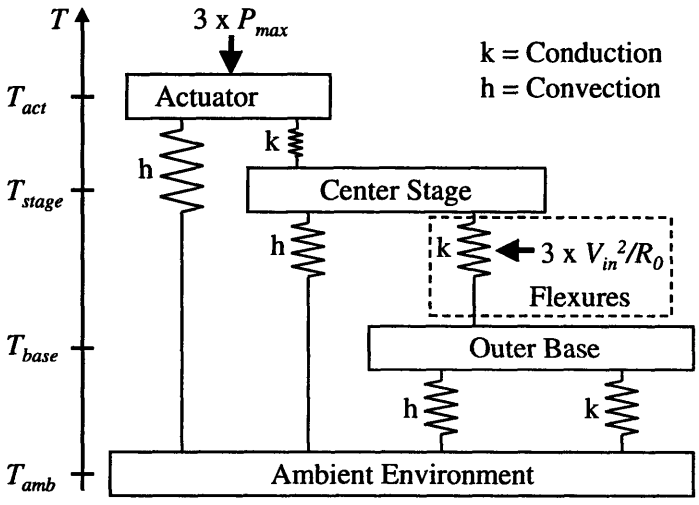

Figure 3.33: Schematic of thermal model.

The temperature of the coils must be known to determine the heat flux from the coils. This temperature is dependent on several parameters, including the convective heat flow from the coil. The convective heat transfer coefficient of the coil surface can be predicted from the Nusselt number of the coil, which is determined by several parameters [47]. The characteristic

dimension of the coil surface, d_{surf}, is its length in the direction of the convective heat transfer (upwards), and is calculated as:

$$d_{surf} = h_{mag} + 2w_{coil}. \tag{3.101}$$

The surface area of the coil, A_{surf}, can be approximated as:

$$A_{surf} = l_{mag}\,2\left(d_{surf} + h_{gap} - 2d_{spacing}\right). \tag{3.102}$$

The Prandl, Pr, Grashoff, Gr, and Rayleigh, Ra, number of the fluid flow are calculated for a surface temperature difference of ΔT, as [47]:

$$\begin{aligned} \mathrm{Pr} &= \frac{\mu_{air} c_{pair}}{k_{air}} \\ \mathrm{Gr} &= \frac{\frac{1}{T_{air}} \Delta T g d_{surf}^{\,3}}{\nu_{air}^{\,2}}, \\ \mathrm{Ra} &= \mathrm{Pr}\cdot\mathrm{Gr} \end{aligned} \tag{3.103}$$

where the properties of air at standard temperature and pressure are used for calculation [48]- specific heat capacity c_{pair}, of 1010 J/K, dynamic viscosity μ_{air}, of $1.83x10^{-5}$ Pas, kinematic viscosity ν_{air}, of $1.54x10^{-5}$ m^2/s, thermal conductivity k_{air}, of 0.0257 W/(mK), ambient temperature viscosity μ_{air}, of 25 °C, and gravity g, of 9.81 m/s^2. A rough estimate of the convection coefficient can be calculated from these dimensionless numbers using the Nusselt number, Nu, definition for natural convection on a vertical surface [47]:

$$\begin{aligned} \mathrm{Nu} &= 0.68 + 0.670\left(\mathrm{Ra}\cdot\Psi\right)^{\frac{1}{4}} \\ \Psi &= \left[1 + \left(\frac{0.492}{\mathrm{Pr}}\right)^{\frac{9}{16}}\right]^{\frac{-16}{9}} \end{aligned} \tag{3.104}$$

The convective heat transfer coefficient, h_{eff}, of the coil surface can be found from the definition of the Nusselt number [47]:

$$h_{eff} = \mathrm{Nu}\frac{k_{air}}{d_{surf}}. \tag{3.105}$$

Radiative heat transfer, q_{rad}, must also be included. This occurs in the non-linear form below, where ε is the surface emissivity (about 1 for lacquered coils), σ is the Stephan-Boltzmann constant ($5.67*10^{-8}$ $W/(m^2K^4)$), T is the emitting surface temperature, and T_{amb} is the ambient air temperature. The expression can be linearized and treated as an effective convection coefficient, h_{rad}, for small variations around a steady surface temperature:

$$h_{rad} \approx \frac{\delta q_{rad}}{\delta T} = \varepsilon\sigma 4T^3. \tag{3.106}$$

The total convective heat transfer coefficient for the coil surface, h_{conv}, is the sum of these two heat transfer modes:

$$h_{conv} = h_{rad} + h_{eff}. \tag{3.107}$$

The natural convection and radiation coefficients for the range of operation of the device are of roughly equal magnitude: 4.6 and 6.8 W/m^{2}K, respectively.

The desired heat transfer occurring from the coil to the central stage may be visualized as a small leakage of the generated heat through the plastic mandril and aluminum coil mounting arms. The majority of the heat should be convectively dissipated. The coil mounting arms are made thin, long to ensure this is the case in the real device, and the mandril is made of material with low thermal conductivity. The two modes of heat transfer for the coil are compared using the values in Table 3.13, where $k_{mandril}$ is the thermal conductivity of the plastic mandril [48], A_{cs} is the cross sectional area of the mandril, and $l_{mandril}$ is the length of the thermal path through the mandril.

Table 3.13: Estimated thermal parameters for coil.

Property	Value	Units
$k_{mandril}$	0.15	W/mK
A_{cs}	40	mm^2
$l_{mandril}$	3	mm
h_{conv}	11.2	W/m^{2}K

The thermal properties are inserted into the expression below, which is a ratio of convection, Q_{conv}, to conduction, Q_{cond}, out of the coil:

$$\left(\frac{Q_{conv}}{Q_{cond}}\right)_{coil} = \frac{\left(h_{conv}A_{surf}\right)}{\left(\frac{2k_{mandril}A_{cs}}{l_{mandril}}\right)}. \tag{3.108}$$

The ratio of the heat transfer modes suggests convection should be about 13x larger than conduction, meaning it effectively sets the temperature of the coil. This simplifies the model, as it suggests the coil temperature is largely independent of the component temperatures down the thermal flow path.

The heat flow from the coil to center stage can be visualized as three discrete small flows from the coils and from the strain gages as described below, to a large lumped body mass owing to the low thermal resistance of the aluminum center stage. The center stage has no heat generation sources, and thus simply acts as a thermal mass elastically held between the heated coils and ground. The samples on the center stage will likely be thermally insulated, owing to the fact that sample holders are commonly made of plastic.

The heat flow from the center stage can be visualized as 8 flows through the bearing flexures in parallel with convection to the ambient environment. As with the coil analysis, the relevant geometry and thermal parameters are shown in Table 3.14, where A_{surfcs} is the surface area of the center stage, k_{al} is the thermal conductivity of aluminum [48], $A_{flexure}$ is the sum cross sectional area of a flexure from set 1 & from set 2, and $l_{flexure}$ is the length of the flexures.

Table 3.14: Estimated thermal parameters for center stage.

Property	Value	Units
A_{surfcs}	17500	mm^2
k_{al}	240	W/mK
$A_{flexure}$	11.6	mm^2
$l_{flexure}$	28.6	mm

The thermal properties are inserted into the expression below:

$$\left(\frac{Q_{conv}}{Q_{cond}}\right)_{cs} = \frac{\left(h_{conv}A_{surfcs}\right)}{\left(\frac{4k_{al}A_{flexure}}{l_{flexure}}\right)}. \tag{3.109}$$

The ratio of the heat transfer modes suggests conduction should be roughly 2x convection, assuming the temperature drop to air and outer base are equivalent. This means both must be considered in setting the center stage temperature. The center stage temperature is set largely by the actuator and ambient environment temperature in the case that the outer base is in thermal equilibrium with the center stage. The heat flow across the flexures is significantly less than the heat generated by the strain gages on the flexure. This heat is expected to flow in both directions, to the center stage and outer base.

The outer base is assumed to be thermally grounded in the best case scenario. This scenario involves the device being clamped down to the much larger thermal mass of the

Confocal microscopy stage, which is always effectively at the ambient temperature. In this case, conduction is the dominant mode of heat transfer controlling the center stage temperature. In the worst case scenario, the outer base is not in thermal contact with any other surface and must convectively transfer out the heat. This second case is similar to the test setup used for the thermal experiments, where the device was supported on three thin metal stands. In both cases, the outer base is thermally linked directly to the ambient environment; the only difference is the strength of the link depending on which heat transfer mode dominates.

3.9.3 Finite Element Model

A finite element analysis model (FEA) was generated from the CAD design of the HSFD using CosmosWorks Solver to determine the thermal profile and ensure the system thermal constraints are met over the whole device. The mesh is finely discretized on the flexures and mandril to obtain better special resolution at these points of interest. The device is split along its plane of symmetry (normal to the X-axis) for computational speed, and the conduction set to zero at these exposed surfaces. Material properties are imported for the aluminum and permanent magnets. Three convection parameters are used to adjust the thermal profile: the coil surface, h_{coil}, center stage, h_{stage}, and outer base, h_{base}, convection coefficient. The conductivity of the plastic mandril, $k_{mandril}$, is the fourth parameter which is adjusted to fit the model to observed data.

3.9.4 Model Calibration

The four adjustable parameters in the FEA model were altered until the results of the simulation largely matched the observed temperature profile in prototype C. The adjustment process is carried out sequentially, working down the thermal flow path shown in Figure 3.33 above. The coil temperature is fit through adjustment of h_{coil} until the simulation profile matches the measured coil surface temperature. The center stage temperature is fit through adjustment of $k_{mandril}$ and h_{stage} until the simulation profile matches the measured stage surface temperature. Finally, the outer base temperature is fit through adjustment of h_{base} until the profile matches the measured stage surface temperature. The estimated heat transfer coefficient in Equation (3.107) is used as the starting location for the adjustment. The resulting values for the parameters are shown in Table 3.15.

Table 3.15: Thermal model calibration parameters.

Property	Value	Units
h_{coil}	25	W/m^2K
h_{stage}	6	W/m^2K
h_{base}	5	W/m^2K
$k_{mandril}$	1.8	W/mK

The center stage and outer base convection coefficients are slightly lower than expected from predictions, which may be due to low emissivity for aluminum which is roughly 0.1. This low emissivity heavily attenuates the radiation term, resulting in a convection coefficient of around 5.5 W/m^2K, exactly as observed.

The heat transfer coefficient for the coil is significantly higher than expected, about twice the predicted value. This may be due to a semi-forced convection occurring around the coil as shown in Figure 3.34.

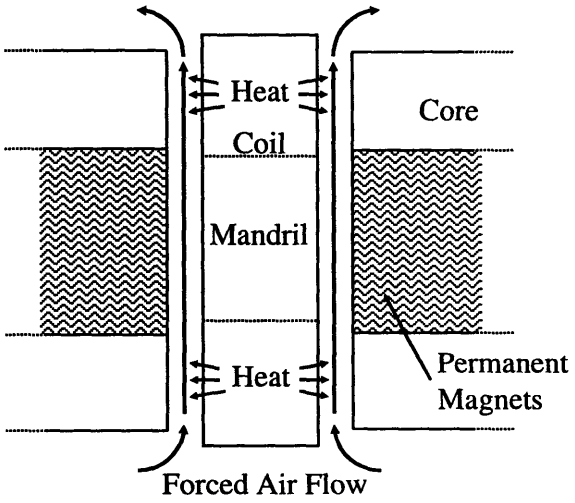

Figure 3.34: Mechanism for forced air flow increasing convection.

The mechanism for this effect is believed to be the heating of the air in the thin column on either side of the coil generating a buoyancy force, driving the air out of the gap. The unique situation around the coil is the extreme dimensions of the gap. The rising, heated air draws a flow of cold air in from below to fill its spot, driving a flow over the surface and boosting the heat transfer coefficient.

3.9.5 Results

The FEA model thermal profile of prototype D at steady state maximum power output is shown in Figure 3.35. The power generation terms correspond to the measured currents required to displace the center stage 250 μm, and are 0.02 W for actuator 1, 0.83 W for actuators 2 &3, and 0.08 W for each of the strain gages.

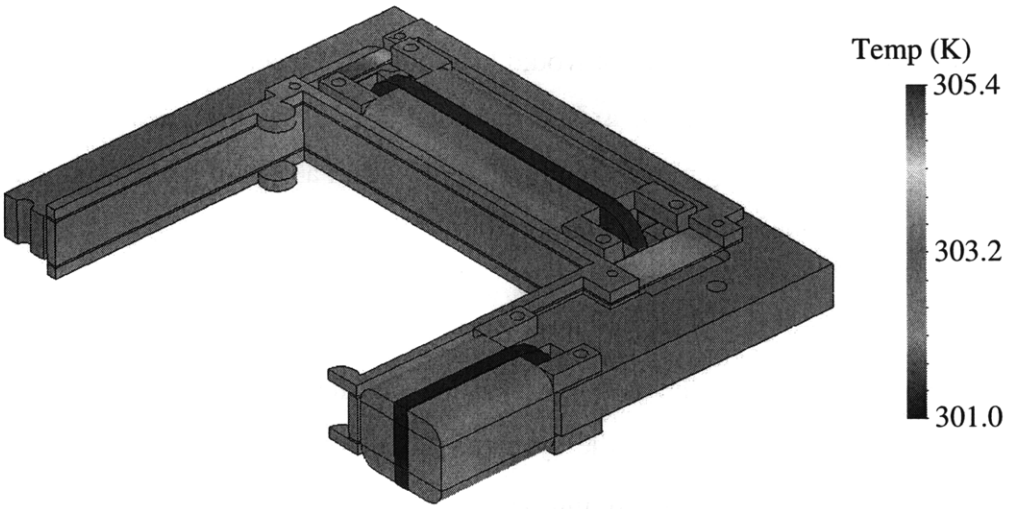

Figure 3.35: FEA Thermal Simulation of Prototype D.

The temperatures of the components, measured at the same locations as in the thermal experiment described later in Section 5.11.1, are shown in Table 3.16.

Table 3.16: Thermal model results for Prototype D.

Component	Temperature (°C)
Coil	5.4
Center Stage	3.4
Outer Base	3.3

The high voltage strain gage sensors add noticeably to the thermal profile, but not significantly. The effect of the gages is to bring both the center stage and the outer base to nearly the same temperature, indicating that the center stage must be transferring heat out mainly through convection. Actuator 1 is cooler than 2 and 3, because the COS is near the COG. Plastic screws are used to anchor the mandril to the coil mounting arm due to the need for thermal insulation

and the low precision/force required for the connection. The thermal model indicates that the system will operate safely within its thermal constraints.

3.10 System Operation

3.10.1 Method

A transfer function model of the system is used to explore the dynamic properties of the device and determine what sort of control would be required to meet the performance constraints on the system operation. Only the Z axis is considered as it has the most difficult operation parameters. The rotational axes are only required to be within a certain range around zero, while the Z-axis must have an operating bandwidth of 100-200 Hz.

3.10.2 Plant

The system plant is modeled as three cascaded transfer functions: the current controller, actuator and bearing. The bearing of the system is modeled as a 2nd order system for initial analysis. The bearing design results give the estimated mass, m_{est} as 0.25 kg, and the stiffness as 58.6 kN/m, with an estimated natural frequency of about 77 Hz. The damping is assumed to be negligible due to the lack of friction or direct contact with any damping mechanism besides air. The estimated damping ratio, ζ_{est}, is set as 0.1, as a safely high estimate. The estimated damping coefficient, b_{est}, can be calculated from the damping ratio as:

$$b_{est} = \frac{2\zeta_{est} k_Z}{\omega_n}. \tag{3.110}$$

The actuators are modeled as transfer functions whose dynamics are contained in the current controller. The current controller is modeled simply as a unity gain transfer function because the gain is only important to calibration, and the dynamics were previously described in Table 3.10 to be much greater than 5x the bearing dynamics. The current controller dynamics are not included for simplicity in the estimated system model, shown in Figure 3.36.

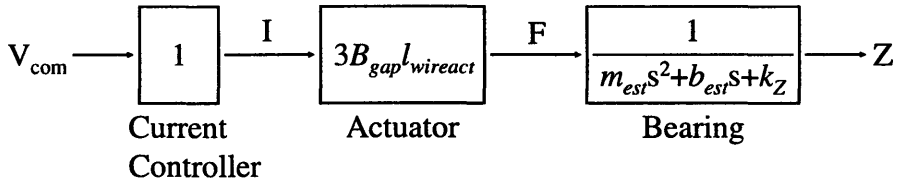

Figure 3.36: Estimated open loop Z-axis transfer function model.

Quick analysis using Equation (3.89) shows that the settling time of the system is about 83 ms, far slower than the 10 ms required to reach 100 Hz bandwidth. Open-loop operation also has difficulty rejecting thermal errors, accounting for loading variations and minimizing parasitic rotations. The bandwidth of the device can be adjusted through the addition of a compensator but the error rejection cannot be accomplished with open-loop control.

3.10.3 Closed Loop Control

The problems identified in the open loop system analysis may be solved through the use of closed loop control. The axial dynamic decoupling through COS/COR alignment indicates that the three axes can be controlled through three independent loops in parallel. Standard negative feedback loop is used for the axes, owing to its simplicity. The control loop for the Z-axis is shown in Figure 3.37, with compensator $C_Z(s)$ and sensor $H(s)$ transfer functions added to the loop.

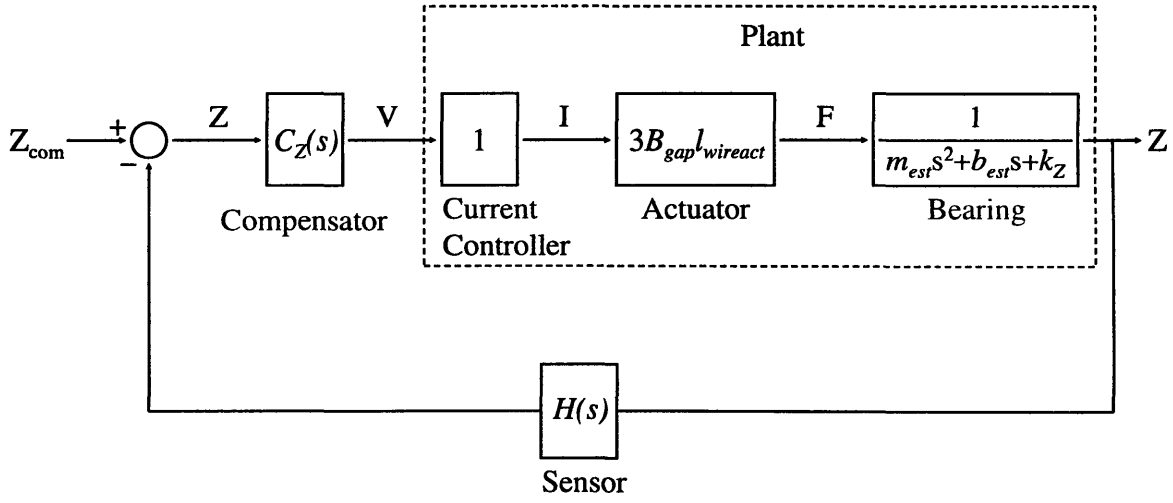

Figure 3.37: Estimated closed loop Z-axis transfer function model.

The compensator design relies heavily on the actual plant dynamics, particularly in the case of the HSFD where the desired operation bandwidth lies well into the dynamic range of the plant (above natural frequency). Thermal expansion and other slow changes such as loading variations occur at lower frequency and should be adequately rejected by any closed-loop control.

CHAPTER

4

FABRICATION AND ASSEMBLY

4.1 Introduction

In this chapter, the fabrication and assembly process for the HSFD is described with a focus on issues of particular importance to be observed during fabrication and assembly. Fabrication methods and operations of importance to device performance are discussed. The optimal assembly process is described for ensuring component alignment and minimizing the difficulty of assembling magnetic components. Specific dimensions of the components for prototype D are shown in Appendix C.

4.2 Fabrication

The HSFD is designed to utilize a small range of repeated fabrication operations for as much of the process as possible to simplify fabrication and reduce costs. All parts of the HSFD, excluding the coils, are fabricated through a combination of water jet cutting and milling. All other operations used in the HSFD fabrication, such as tapping and deburring, do not require large machine shop equipment. The water jet cutting is carried out on sheet aluminum of thickness from 0.508 mm (0.020 in.) up to 12.5 mm (0.500 in.), and on steel of thickness 6.35 mm (0.250 in.). Laser cutting may also be used in place of the water jet. The use of the laser cutter may slightly decrease the yield strength and fatigue life of the flexures due to the localized thermal stresses generated on the edges of the flexure during fabrication.

All parts in the HSFD are made from aluminum with the exception of the coil mandrils which are made from plastic, and the cores, which are made from 1020 steel. The flexures are

composed of 2024-T4 aluminum, while the rest of the structure is composed of 6061-T651 aluminum.

The assemblies are held together with 4-40 screws due to the need for small diameters when anchoring the center stage side of the flexures. If the center stage is made thicker through the relaxation of the system geometric constraints such as aperature size, then 6-32 screws would be recommended because of their larger pressure cones and higher clamping forces.

The flexure blades require the highest fabrication accuracy, about 25 μm (0.001 in.), in order to ensure the COS/COR remain aligned within 1 mm. For all of the other components, dimensional accuracies of about 125 μm (0.005 in.) are sufficient. The design is intended utilize alignment during assembly to account for fabrication errors. The center stage and outer base should be cut from the same sheet of material to ensure equivalent thicknesses within a few mils. The matching thicknesses will minimize the strain in the flexures once assembled.

The coil fabrication process consists of anchoring the mandrils between two parallel plates separated by a gap of h_{coil} as described by Equation (3.43). The magnet wire is wrapped around the mandril in successive layers and anchored with epoxy. This replicates the geometry desired for the coil. For the prototype coils, 'Krazy Glue' standard formula epoxy is used on the coils because it dries rapidly and Teflon plates are used to hold the coil geometry during fabrication.

4.3 Assembly

4.3.1 Wheatstone Bridge Assembly

The Wheatstone bridge is assembled by gluing strain gages to the top and bottom surface of the flexures using thermal set glue TT300 from Omega Inc. The strain gages are located on the flexure as shown in Figure 3.21 previously, with the wiring carried out according to Figure 4.1. The wires used to link the strain gages are 36 AWG stranded flexible wire, chosen to be as compliant as possible during device operation. The STC resistance wire is adhered to the outer base near the anchored end of the flexure using type 384 Loctite brand thermally conductive adhesive to minimize the temperature differences between the span temperature compensation and the strain gages on the Wheatstone bridge.

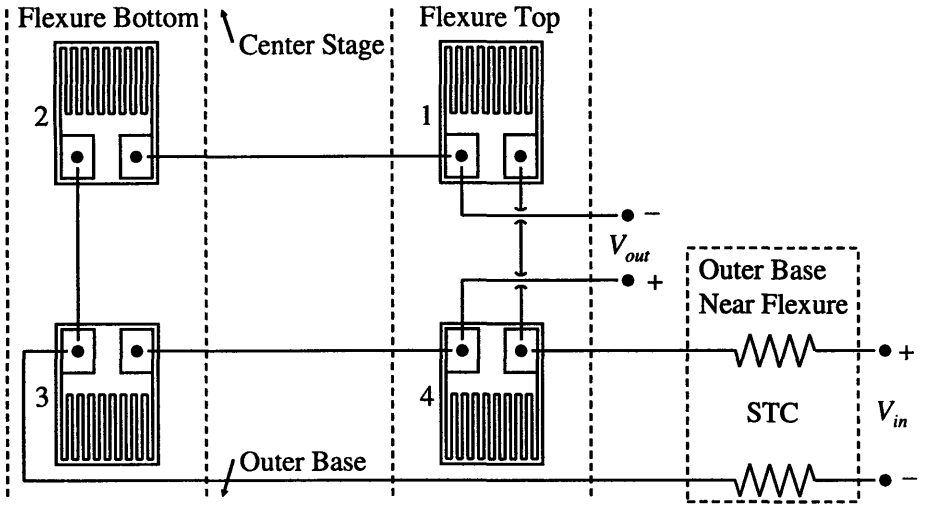

Figure 4.1: Strain gage wiring schematic.

The interface between the circuitry and the sensors is a parallel connector, which enables the WB wires to be disconnected if the flexure sheet needs to be removed. Insulated, foil wrapped, grounded, and twisted wires are used to electrically link the parallel connector to the electronics. The insulation protects the signals from making electrical contact elsewhere, the foil protects the wires from varying electric fields, the grounding prevents static buildup on the foil and the twisting cancels out the voltages generated by varying magnetic fields.

4.3.2 Coil and Magnet Assembly

The first step in the HSFD mechanical assembly is to place the permanent magnets flush against the sides of the coils, with a plastic shim between the magnet and coil, as shown in Figure 4.2. Plastic is used for the shim to avoid scratching the coils, particularly during the final shim removal step. This shim determines the gap between the magnet and coil, therefore is ideally 0.254 mm (0.010 in.), to set the spacing to the design parameter. Any warpage in the coil decrease the gap when the shim is removed. For safety, the shim size set at 0.380 mm (0.015 in.). If the coil warpage is reduced, as discussed in Section 6.3.2, then the 0.254 mm shims should be used. The magnets are placed with anti-parallel magnetization directions. The attractive force between the magnets, even with no cores, is on the scale of 100 N, therefore caution is necessary during assembly.

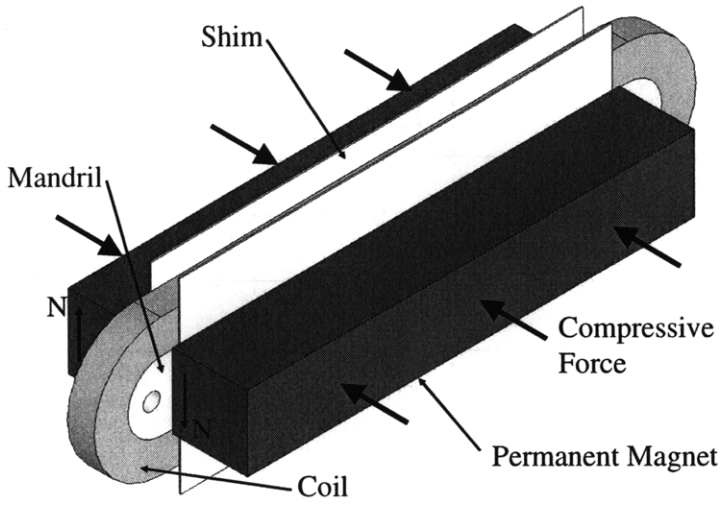

Figure 4.2: Coil and magnet assembly.

4.3.3 Addition of Top Cores

The second step in the assembly is to insert the top cores on each of the actuators, with the metal spacing shim placed between the magnet and cores, all as shown in Figure 4.3. The surface of the magnet and steel is cleaned of any debris prior to assembly, such that the components mesh correctly with no air gap. The steel shims are placed to ensure that the cores clamp the permanent magnets rather than the outer base, as described in Section 3.8 earlier. The first cores are added at this step as it becomes progressively more difficult to add them later without damaging nearby parts, including the flexures. The addition of each core boosts the magnetic field, increasing the attractive pull of the magnets. As many cores as possible are placed at this step, the remainder are inserted once the actuators are seated in the outer base.

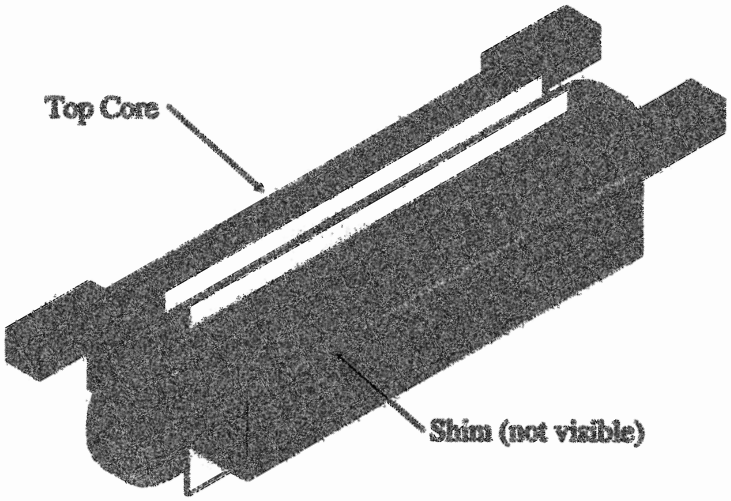

Figure 4.3: Addition of top cores.

4.3.4 Attachment to Center Stage

The third step in the assembly is to anchor all of the actuators to the center stage as shown in Figure 4.4. The actuators can be attached to the center stage at this step with relatively little force interaction. The wiring leading from the actuators are on one side, for later electrical connections. The coils are attached and clamped down to the coil mounting arms when the compressive force of the magnets is forcing them into a completely flat shape. The coils will relax back to their natural shape to a degree depending on how tightly they are anchored once the shims are released.

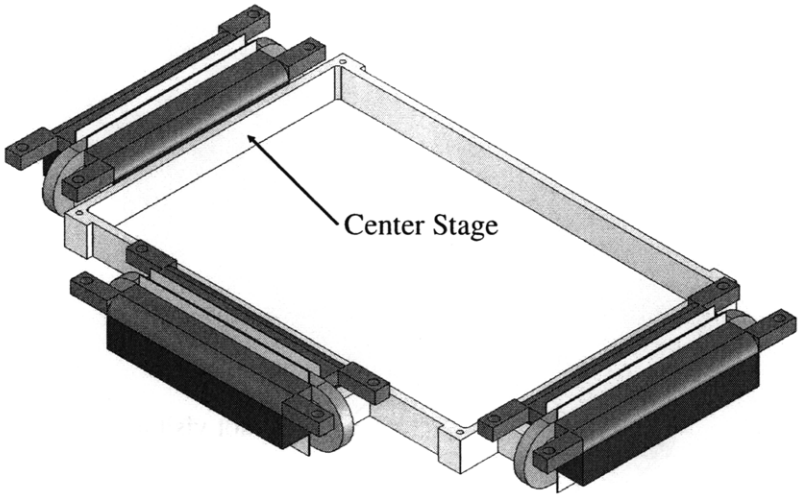

Figure 4.4: Attachment to center stage.

4.3.5 Addition of Top Flexure Sheet

The fourth step in the assembly is to settle the center stage into the outer base and clamp it in place using the top flexure sheet, as shown in Figure 4.5. The center stage is placed in the outer base and both are seated on a flat surface, to ensure their surfaces are flush. After cleaning all of the clamping surfaces, the top flexure sheet is laid over the center stage and tightly clamped in place with both inner and outer flexure clamps. Type 316 stainless steel washers are used with each 4-40 screw to evenly distribute the clamping pressure and reduce damage on the parts. This step ensures that the flexures are placed and anchored with minimal stress. The actuators are aligned with their mounting holes only once the flexure alignment is complete.

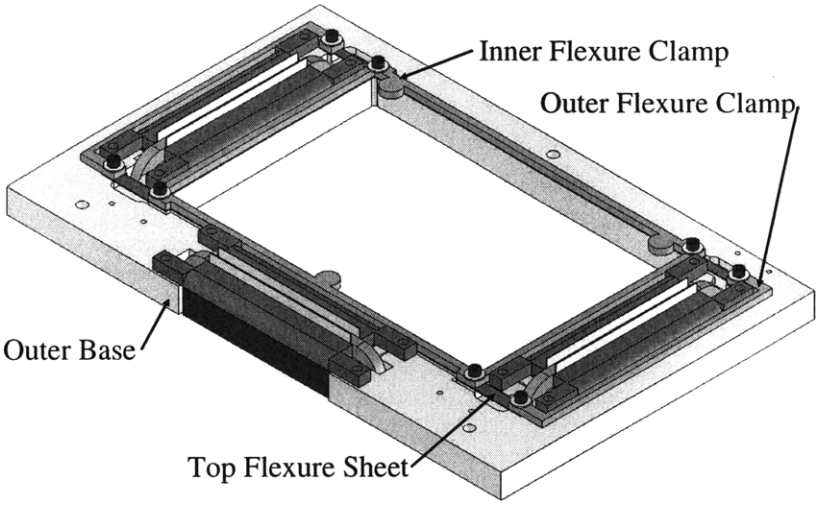

Figure 4.5: Addition of top flexure sheet.

4.3.6 Alignment of Actuators

The fifth step in the assembly is to adjust the magnets and actuator cores into correct alignment with the outer base, and screw the actuators in place without tightening the screws, as shown in Figure 4.6. Washers are used with all of the clamping screws. The cores are intended to not make contact with the outer base on the top, to ensure that the permanent magnets are tightly clamped. The bottom cores make contact, but the top cores do not. The alignment of the actuators at this point ensures that the flexures are located with minimal internal stresses and that the actuators are located relative to the coil locations.

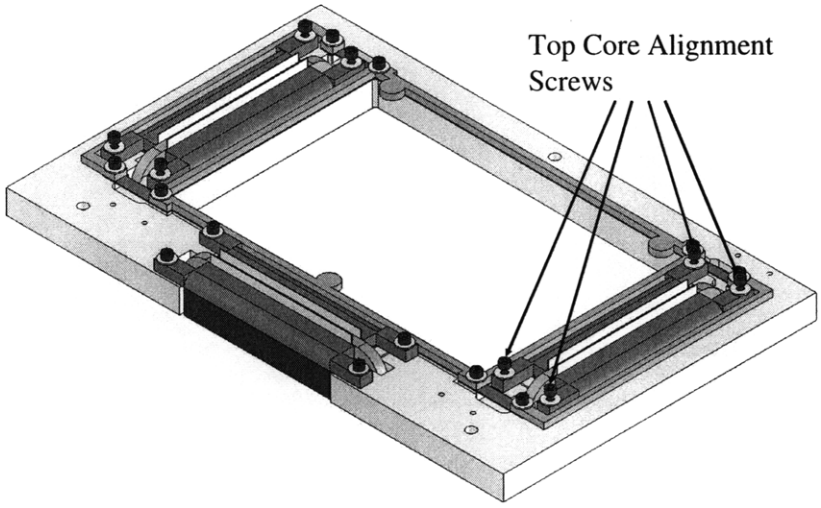

Figure 4.6: Alignment of actuators.

4.3.7 Addition of Bottom Components- Cores and Flexures

The sixth step in the assembly is to flip the device and add in all of the cores and flexure sheet on the bottom, as shown in Figure 4.7. The cores are added first, and are slid in from the side to minimize the pulling force of the magnet. The bottom flexure sheet is placed and firmly clamped down once all of the cores are seated and the magnets centered in their gaps. The bottom cores are then firmly clamped down to the outer base, setting the actuator location.

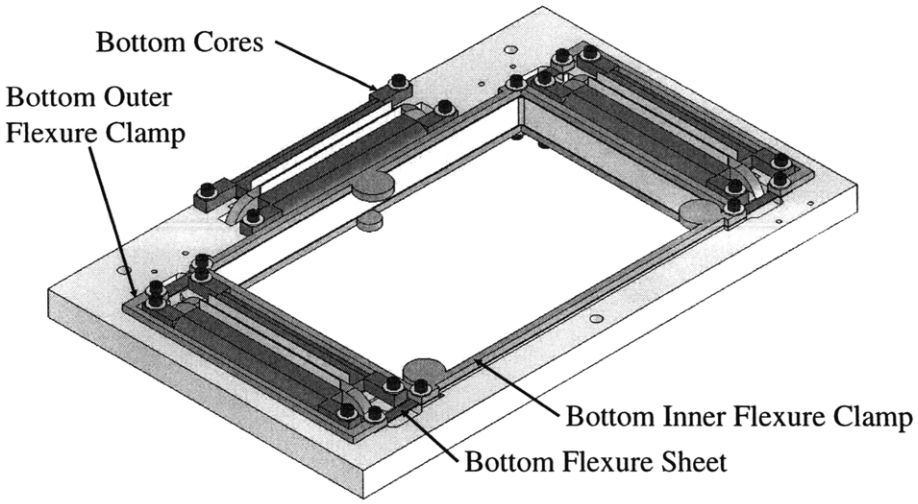

Figure 4.7: Addition of bottom components- cores and flexures.

4.3.8 Anchoring of Actuators

The seventh step in the assembly is to flip the device right side up, tighten the actuator screws and remove the plastic shims, resulting in the completed device as shown in Figure 4.8. The magnets will snap in to contact the coil when the shims are removed and the assembly process must be restarted if the actuators are not tightly clamped. Pressurized air and thinner plastic shims are used to ensure that no debris is located in the gap between coil and magnet/core.

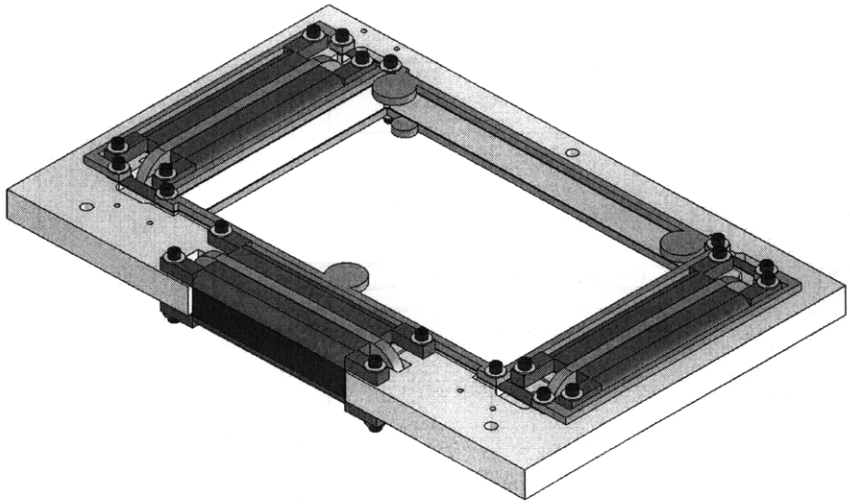

Figure 4.8: Anchoring of actuators.

The circular extrusions on the inner flexure clamp are for characterization purposes only, and are not needed for the final device. The capacitance probe position measurement system described in Section 5.2.3 requires 9.53 mm (0.375 in.) diameter conducting targets on the bottom of the moving stage. The top extrusions are used as plates for the placement of weights when applying static moments to the device to characterize the bearing.

The completed HSFD is shown in Figure 1.1, with all components added, shims removed and anchored down to the testing stage. The labels are used to identify the wires which connect the actuators and sensors to their associated electronics.

4.4 Electronics

4.4.1 Layout

The system electronics are laid out according to Figure 4.9, with the actuation channels running across the top, and the sensing channels running across the bottom of the schematic. Shielding is important for a precision machine such as the HSFD, particularly on the components carrying the un-amplified position signal.

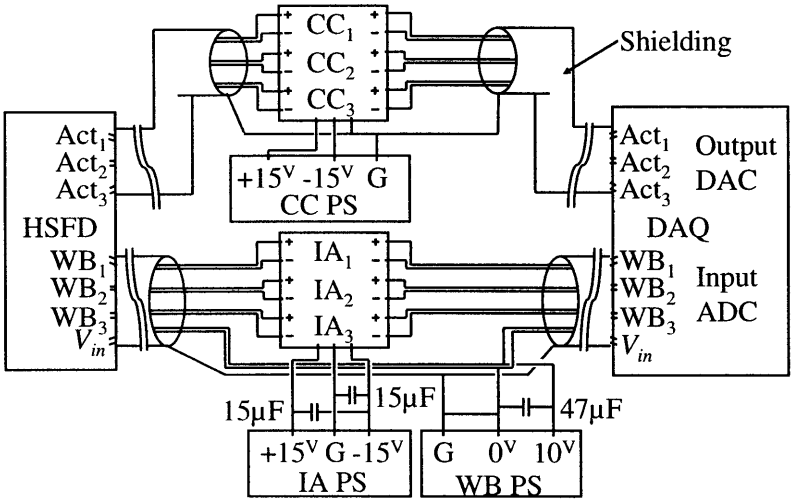

Figure 4.9: Wiring schematic for HSFD electronics.

The signal is boosted by approximately 200*x* after the IAs, reducing the effect of further noise. The wires are shielded wherever possible to reduce EM noise and cross coupling in the device. The actual electronics layout is shown in Figure 4.10, with the main electronics sections boxed for clarity. The setup is carried out on an optical table and all wires are strain relieved to the table when possible.

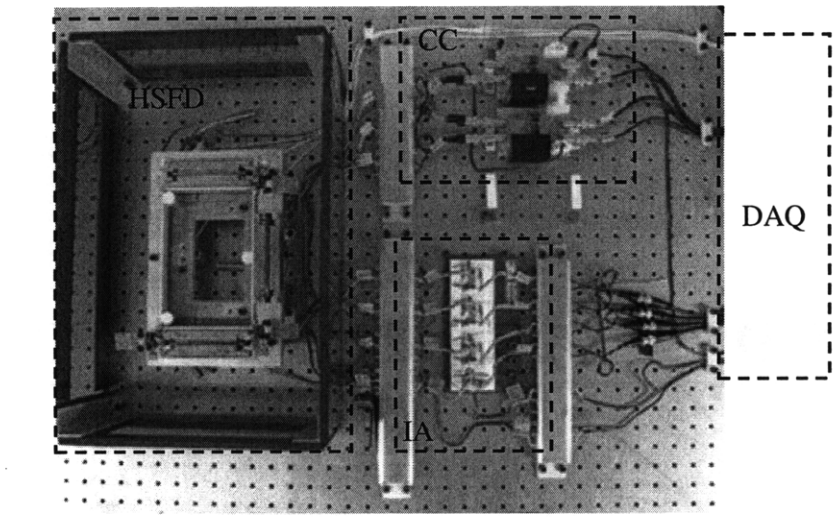

Figure 4.10: Layout of HSFD electronics.

Terminal blocks are used between each section of electronics to (i) act as strain relief and (ii) to allow for the wire type to be changed, such as the transfer from shielded, grounded and twisted wires to BNC cables leading to the DAQ board. The terminal blocks are shielded to reduce EM noise. The IA circuitry is laid out on a breadboard to ease adjustment and debugging of the circuitry. This exposed section of the electronics is likely picking up significant noise. The HSFD is placed in an environmental enclosure (orange) to prevent stray air currents and radiative heating from the external sources disturbing the system during characterization. The top of the enclosure has been removed for image clarity.

CHAPTER

5

CHARACTERIZATION AND TUNING

5.1 Introduction

In this chapter, the components of the HSFD are characterized and compared to the design performance predictions. The methods and model used for the characterization are described along with the results obtained. The full system dynamics are measured and a compensator design is implemented to enable the required device bandwidth. The full system performance is characterized and found to meet the performance requirements.

5.2 Position Measurement

5.2.1 Method

The position sensing on the HSFD is done at three discrete locations, however it is not these three displacements which are desired for device operation, but rather the height, tip and tilt of the center stage. The information must be mathematically transformed into a position reading of the stage starting from three sensor measurements, which are read into the controlling software.

5.2.2 Coordinate Transform

The coordinate transform inputs three displacements along the Z axis, corresponding to the transverse displacement of the sensed flexures. These measurements are taken at a known in-plane location defined relative to the COG. The exact in-plane locations are found from the CAD model of the HSFD. The result of the three sensor measurements is the definition of three

sensor location points in 3-D space as shown in Figure 5.1. The flexures are anchored to the stage on their guided end, enforcing the requirement that the central stage plane must pass through each of the three sensor locations.

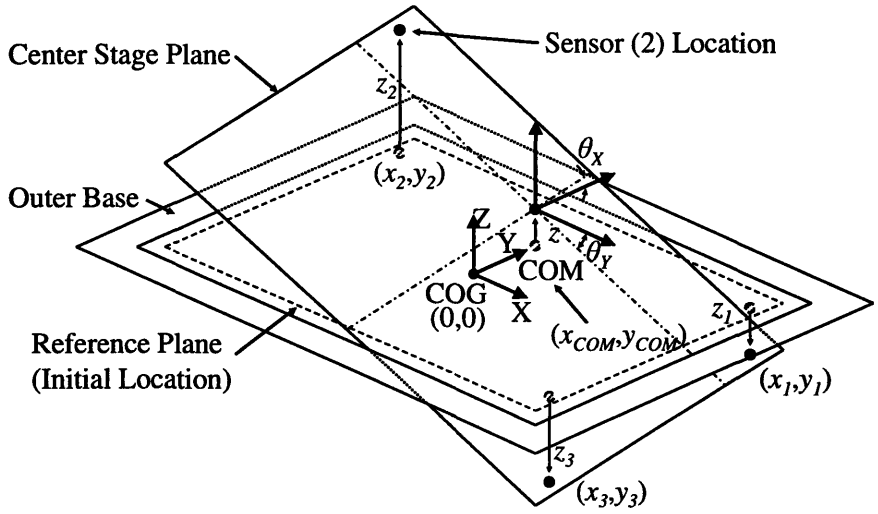

Figure 5.1: Coordinate transform geometric layout.

The position measurement is only relative to the initial location (Reference Plane) of the center stage plane, as it is not known in absolute terms how the stage sits relative to external devices such as the Confocal microscope. The sensor locations are transformed into a set of coordinates locating the center stage in three degrees of freedom. While three points perfectly define a plain, more than three points may be used for this sort of measurement. The effect is to average out the measurement error as described in Section 3.5.4 previously, making the position sensing more accurate. The coordinate transform of the sensor locations must be done in real-time to enable the control loop to control the center stage position. Only three bits need to be fed to the coordinate transform operation because six of the nine bits of information determining the sensor locations are constant.

The coordinate transform needs to be written as a transfer function utilizing only linear algebra operations on the sensor locations as shown in Figure 5.2, such that it can be easily integrated into the controlling software. The sensor vertical displacements are fed into the transform as a vector, $\boldsymbol{s}$, which is operated on by the transform matrix to result in another vector, $\boldsymbol{p}$, describing the height, tip and tilt of the center stage.

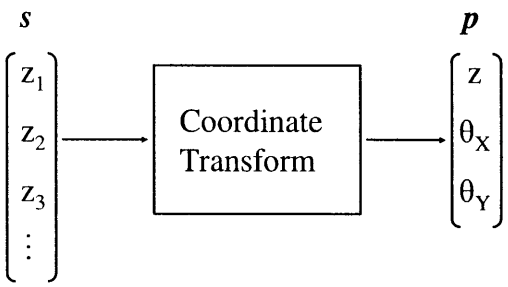

Figure 5.2: Schematic of coordinate transform.

The coordinate transform process can be decomposed into two main steps: finding the plane defined by the sensor locations (center stage plane) and using the plane parameters to find the height, tip & tilt of the plane.

The center stage plane can be described by an equation defining the height of the plane, $Z(x, y)$, at any given in-plane location (x, y),

$$Z(x,y) = C_0 + C_1 x + C_2 y, \tag{5.1}$$

where $C_{0\text{-}2}$ are constants. Equation (5.1) must be satisfied at every point along the center stage plane, including the points measured by the sensors. The particular instantiations of Equation (5.1) at each sensor in-plane location can be stacked into a matrix, $\boldsymbol{A}$, relating the vector of plane constants, $\boldsymbol{c}$, to the sensor measurements:

$$\boldsymbol{s} = \mathbf{A} \cdot \boldsymbol{c}$$

$$\begin{bmatrix} z_1 \\ z_2 \\ z_3 \\ \vdots \end{bmatrix} = \begin{bmatrix} 1 & x_1 & y_1 \\ 1 & x_2 & y_2 \\ 1 & x_3 & y_3 \\ & \vdots & \end{bmatrix} \cdot \begin{bmatrix} C_0 \\ C_1 \\ C_2 \end{bmatrix}. \tag{5.2}$$

The plane defining matrix $\boldsymbol{A}$ will have n rows corresponding to the n sensor measurements read into the control loop and three columns corresponding to the three plane constants. The sensor position measurements are input to the transform, while the plane constants are the output. To find the plane constants, the least squares method of fitting [49] is used, giving:

$$\boldsymbol{c} = \left[\left(\mathbf{A}^{\mathrm{T}} \mathbf{A} \right)^{-1} \mathbf{A}^{\mathrm{T}} \right] \cdot \boldsymbol{s}. \tag{5.3}$$

The plane constants vector defines the least squares plane fit to the sensor locations.

The plane constants are transformed into a height, tip, and tilt of the plane to complete the coordinate transform. The height of the plane, z, is defined by Equation (5.1) at the center of

measurement (COM), a particular point on the center stage plane with in-plane coordinates (x_{COM}, y_{COM}):

$$z = Z(x_{COM}, y_{COM}) = C_0 + C_1 x_{COM} + C_2 y_{COM}. \tag{5.4}$$

The rotation around the X-axis, θ_X, is defined for small deflections as the slope of the plane equation:

$$\theta_X = \frac{\partial Z(x,y)}{\partial y} = C_2. \tag{5.5}$$

The rotation around the Y-axis, θ_Y, is likewise defined for small deflections as:

$$\theta_Y = -\frac{\partial Z(x,y)}{\partial x} = -C_1. \tag{5.6}$$

The definitions for the plane coordinates, Equations (5.4), (5.5), and (5.6) are stacked into a matrix, $\boldsymbol{B}$, which transforms the plane constants into the plane position in terms of height, tip and tilt:

$$\boldsymbol{p} = \mathbf{B} \cdot \boldsymbol{c}$$

$$\begin{bmatrix} z \\ \theta_X \\ \theta_Y \end{bmatrix} = \begin{bmatrix} 1 & x_{COM} & y_{COM} \\ 0 & 0 & 1 \\ 0 & -1 & 0 \end{bmatrix} \cdot \begin{bmatrix} C_0 \\ C_1 \\ C_2 \end{bmatrix}. \tag{5.7}$$

The full coordinate transform matrix, $\boldsymbol{T}$, defining the center stage plane position directly from the sensor readings, is found through combining Equations (5.3) and (5.7) to give:

$$\boldsymbol{p} = \mathbf{T} \cdot \boldsymbol{s}$$

$$\boldsymbol{p} = \left[\mathbf{B}\left(\mathbf{A}^{\mathrm{T}}\mathbf{A}\right)^{-1}\mathbf{A}^{\mathrm{T}} \right] \cdot \boldsymbol{s}. \tag{5.8}$$

Equation (5.8) must be used for calculating the coordinate transform matrix when $n > 3$, but the plane is exactly defined when $n = 3$, thus least squares fitting is un-necessary and Equation (5.8) reduces to:

$$\boldsymbol{p} = \mathbf{B}\mathbf{A}^{-1} \cdot \boldsymbol{s}$$

$$\begin{bmatrix} z \\ \theta_X \\ \theta_Y \end{bmatrix} = \frac{1}{x_1(y_2 - y_3) + x_2(y_3 - y_1) + x_3(y_1 - y_2)} \begin{bmatrix} \chi_1 & \chi_2 & \chi_3 \\ x_3 - x_2 & x_1 - x_3 & x_2 - x_1 \\ y_3 - y_2 & y_1 - y_3 & y_2 - y_1 \end{bmatrix} \cdot \begin{bmatrix} z_1 \\ z_2 \\ z_3 \end{bmatrix}. \tag{5.9}$$

$$\chi_1 = x_{COM}(y_2 - y_3) + y_{COM}(x_3 - x_2) + x_2 y_3 - x_3 y_2$$

$$\chi_2 = x_{COM}(y_3 - y_1) + y_{COM}(x_1 - x_3) + x_3 y_1 - x_1 y_3$$

$$\chi_3 = x_{COM}(y_1 - y_2) + y_{COM}(x_2 - x_1) + x_1 y_2 - x_2 y_1$$

5.2.3 Capacitance Probes

Capacitance displacement measurement probes (CAP) are used to facilitate the characterization of the HSFD before strain gage sensors are attached. Additionally, the CAP probes are used to calibrate the strain gage sensors. Three Lion Precision DMT12 CAP probes were placed below the stage and adjusted to read the location of the targets built into the center stage, at the locations shown in Figure 5.3.

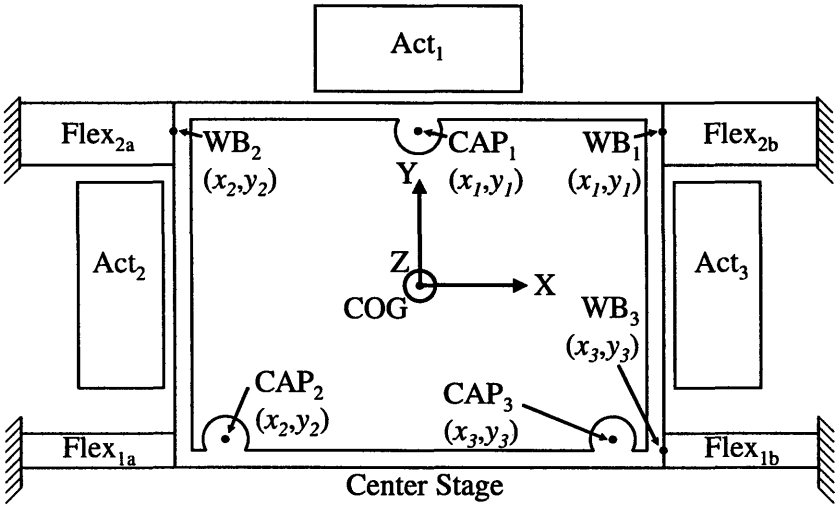

Figure 5.3: General sensor geometric layout.

5.2.4 Coordinate Transform Constants

The location of the CAP and WB strain gage sensors as shown in Figure 5.3 above are calculated from the CAD model of the HSFD and listed in Table 5.1.

Table 5.1: Sensor coordinates.

Sensor	X (mm)	Y (mm)
CAP_1	0.00	52.14
CAP_2	-70.78	-52.14
CAP_3	71.20	-52.14
WB_1	85.68	57.79
WB_2	-85.68	57.79
WB_3	85.68	-57.80

5.3 Mass Characterization

The masses of each of the components in the moving center stage are listed in Table 5.2.

Table 5.2: Masses of center stage components.

Component	Quantity	Mass ±1 (g)
Center stage	1	77.0
Coil	3	35.5
Flexure sheet	2	5.5
Flexure clamp	2	18.5
Screws	14	0.46
Total	---	240.0

The center stage should be as light as possible to raise the natural frequency and simplify the control scheme. The component masses show that while it may be possible to remove material from the center stage, it is not the dominant mass and would show minor gains for large structure stiffness losses. The coils are the dominant mass in the system because of the large volume of copper wiring used.

5.4 Stiffness and COS Characterization

5.4.1 Method

The stiffness and COS of the bearing is determined through the application of loads to the bearing while measuring displacement. CAP probes were used to measure displacements at the locations shown in Figure 5.3, with 1000-point moving averages of the readings, sampled at 1 kHz. The stage is loaded with non-magnetic masses on the three plates built into the top inner flexure clamp. These masses generate known forces at known locations, allowing for the loading forces and moments to be accurately calculated.

For the measurement of the stiffness and COS, it is desired that the COM lie on top of the COS such that rotations do not also appear as vertical motion. The COM is iteratively adjusted along one axis at a time after the data is collected, altering the coordinate transform matrix. The adjusted position data is graphed against applied torques (assuming at each point that the COM is actually on the COS) until the apparent rotational stiffness around the perpendicular axis is the

same over the range of rotations. As the COM (and thus the assumed COS location) is shifted, the lever arm calculated for the masses is altered, changing the calculated moments. The measured rotations are not affected by changing the COM, only the applied moments. At the true COS, the rotational stiffness for rotations in either direction (+ or -) is the same. The value of y_{COS} is determined by aligning the rotational stiffness around the X-axis and vice versa. The vertical stiffness is measured through applying roughly equal weights at all three plates and measuring the displacement at the now known COS once the stiffnesses are matched for both rotational axes.

5.4.2 Results

The error in the linear stiffness fit curve shows no trends for either rotational stiffness measurement, but does show a noticeable trend for the vertical stiffness, as shown in Figure 5.4.

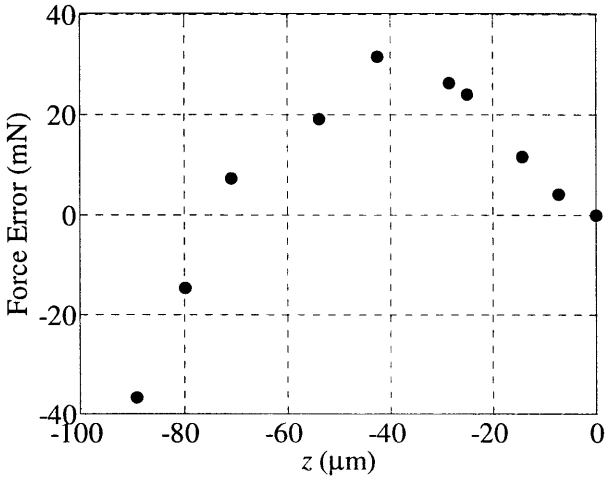

Figure 5.4: Linear stiffness error in Z-axis.

The clear trend in the Z-stiffness error appears sinusoidal, with a proportional error out to about 40 μm, at which point the fit begins to successively under-predict the required force. This appears to be due to a slight strain stiffening effect in the flexures, which is expected as the flexures are designed to operate at high strains.

Table 5.3: Measured stiffness and COS properties.

Parameter	Value	Predicted	Units
k_Z	49.86 ± 0.31	64.25	kN/m
$k_{\theta X}$	0.293 ± 0.0011	3.714	kN/rad
$k_{\theta Y}$	1.051 ± 0.005	1.815	kN/rad
x_{COS}	4.85 ± 0.05	0	mm
y_{COS}	11.93 ± 0.05	11.36	mm

The measured axial stiffnesses and COS location is shown in Table 5.3, along with the measurement errors and predicted values. The linear stiffness of the bearing is significantly below the predicted value. This is consistent with the previous prototypes, measured during the COR alignment described in Section 3.3.8, which also showed proportionally low stiffness as shown in Table 5.4.

Table 5.4: Error in stiffness predictions.

Prototype	Iteration #	Predicted kN/m	Measured kN/m	k_z Error (%)	$k_{\theta z}$ Error (%)
C	1	58.6	56.3	-3.9	-91.1
D	2	48.6	41.2	-29.7	-92.7
D	3	62.6	44.9	-28.3	-92.5
D	4	64.3	49.9	-22.4	-92.1

The error is low for prototype C and but increases significantly for prototype D. This may be due to a consistent clamping issue with the new prototype as described in the section below, or may be due to material property variations between the aluminum used for the flexure sheets. The latter is unlikely. The high strains of the prototype D flexures may require more significant clamping force to ensure proper rigid anchoring.

The rotational stiffnesses are also extremely far off from predictions which suggests that the assumption of center stage rigidity may not be valid. The error in one of the rotational stiffnesses over the various prototypes is shown in Table 5.4 above. The stiffness error is steady, which is consistent with the existence of another equivalent spring in series with the flexural bearing. Adding rigidity to the center stage should significantly reduce this issue. The Y-axis rotational stiffness is much closer to the predicted value, suggesting that the axial mode of

flexural loading, which set the rotational stiffness around the Y-axis is a more accurate model of the bearings behavior than the shear loading which set the rotational stiffness around the X-axis.

The large error in x_{COS} is significantly beyond the scale of the fabrication error observed in y_{COS} (about 0.5 mm). This is likely due to a clamping issue as described in Section 5.4.3.

5.4.3 Clamping

Several issues occur in relation to the clamping of the flexures. The COS of prototype D shows significant position error despite symmetric flexure dimensions. Additionally, the clamping force affects the COS, as shown in Section 3.3.8, where brass screws were experimentally found to provide insufficient force to rigidly anchor the flexures. The contact surfaces are cleaned and deburred before each time the flexure is clamped, ruling out grit as an issue. The end surfaces of the flexure clamps may not be flush, thus when the flexure is engaged in one direction its effective length is less than in the other direction, as shown in Figure 5.5.

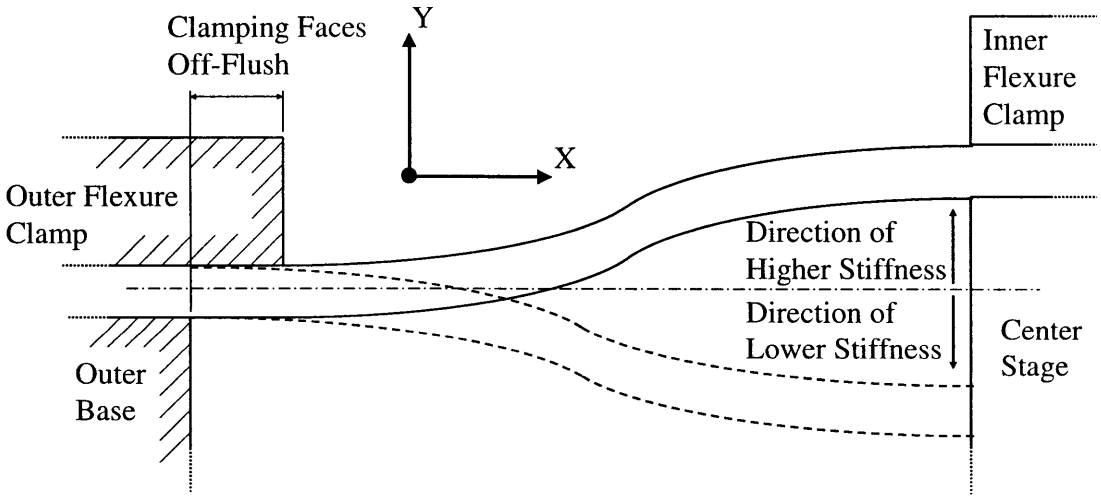

Figure 5.5: Stiffness variation due to clamping error.

The result of such a mis-alignment would be a slightly different stiffness in one direction as compared to the other direction. This will throw off the COS measurement, possibly producing the error observed in prototype D. If both the flexures on one side of the device are clamped off-flush in the same direction, then the there would be no effect on y_{COS}, but a noticeable effect on x_{COS}. Several solutions to this issue are presented in Section 6.3.3 later.

5.5 COR Characterization

5.5.1 Method

The COR is determined through driving the device at the first resonance for each of the rotational axes sequentially while scanning the COM over the in-plane axis perpendicular to the one undergoing resonance. For instance, y_{COR} is found by driving the system at the primary θ_X resonance, and scanning the along the Y-axis. The actuators must be calibrated sufficiently to generate a stage rotation. The coordinate transform matrix is changed when the COM is altered. To do this in real time, the coordinate transform matrix is split to have a constant component and a component dependent on the COM coordinates, which is scaled during device operation. The sum of these components produces an adjustable coordinate transform matrix. The amplitude of the measured vertical displacement is divided by the amplitude of the command signal fed into the rotational axis. This ratio is referred to as the gain of the system.

5.5.2 Results

The gain of the system is graphed versus the commanded COM (translated along the Y-axis) in Figure 5.6.

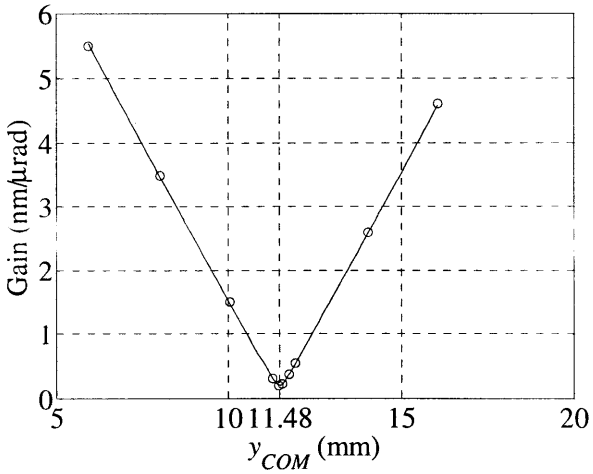

Figure 5.6: Measurement of COR.

The system gain shows a slight hyperbolic form around the COR, which rapidly settles into symmetric linearly increasing asymptotes at large distances from the COR. The system output, *gain(y)* follows the hyperbolic equation:

$$gain(y) = \sqrt{a^2 + b^2 (y - y_{COR})^2} + g_{min} - a, \tag{5.10}$$

where a controls the scale of curvature in the transition regime, b sets the asymptotic slope, g_{min} sets the minimum gain and y_{COR} sets the Y-axis coordinate of the minimum gain. The results of the curve fits for measurements along both axes are shown in Table 5.5.

Table 5.5: Measured COR properties.

Parameter	Value	Predicted	Units
x_{COR}	-0.47 ± 0.1	0	mm
y_{COR}	11.48 ± 0.1	11.36	mm

It is believed that the non-zero value for x_{COR} is due to the anomalously large location error in x_{COS}. The empirical relation determined between the COS and COR suggests that the large positive value of the COS should draw the COR negative by a fraction of the distance, as is observed in this case.

5.6 Actuator Characterization

5.6.1 Method

The actuator performance is measured by commanding each actuator individually over a range of voltages and recording the center stage position in all three axes as a function of these command voltages. CAP probes were used to measure displacement at 1 kHz sampling with 1000-point averages on the data, as done with the earlier characterizations. The voltage command given to the actuator passes through several systems before producing the measured stage position. The open loop transfer function model of this signal propagation is shown in Figure 5.7.

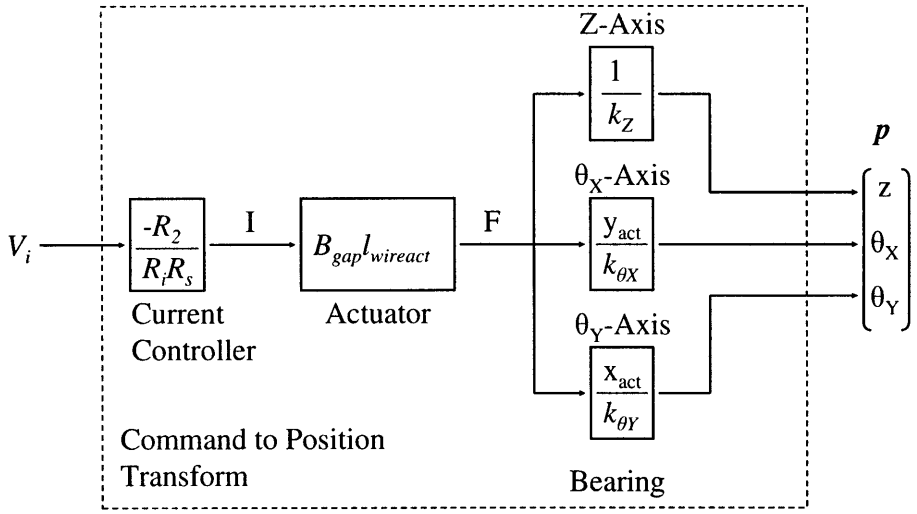

Figure 5.7: Coil voltage command to position transfer function model.

The actuator channel characterization captures the gains of all of the components along the command signals path, including the current controller, actuator and bearing. The complete black box form of the measurement enables extremely accurate calibration, as the outputs of the system (position) are directly fit to the digital inputs to the system (V_i, sent from the DAC).

The position response to the voltage command is expected to be linear and is only measured as a difference in displacement, thus has no offset term. This enables the system response to be written as a product of a calibration coefficient times the command voltage. The nine coefficients linking the coil command voltages, $\boldsymbol{v}$, to the center stage position are gathered into a calibration matrix, $\boldsymbol{C}$, to give the expression:

$$\boldsymbol{p} = \mathbf{C} \cdot \boldsymbol{v}$$

$$\begin{bmatrix} z \\ \theta_X \\ \theta_Y \end{bmatrix} = \begin{bmatrix} a_{1Z} & a_{2Z} & a_{3Z} \\ a_{1X} & a_{2X} & a_{3X} \\ a_{1Y} & a_{2Y} & a_{3Y} \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix}. \tag{5.11}$$

The position vector of the stage can be used as a command, which is then translated into command voltages to each of the coils by inverting the calibration matrix.

5.6.2 Results

The actuator calibration resulted in the actuation matrix, $\boldsymbol{C^{-1}}$, written in units of V/μm for the first row and V/μrad for the second and third rows:

$$\mathbf{C}^{-1} = \begin{bmatrix} -2.637 \cdot 10^{-2} & -1.584 \cdot 10^{-2} & 1.847 \cdot 10^{-3} \\ -2.564 \cdot 10^{-2} & 2.860 \cdot 10^{-3} & -1.477 \cdot 10^{-2} \\ -3.013 \cdot 10^{-2} & 2.636 \cdot 10^{-3} & 1.374 \cdot 10^{-2} \end{bmatrix}. \tag{5.12}$$

The first column of the actuation matrix describes the relative voltage commands given to the actuators for commands of purely vertical motion. All of the actuators are commanded at roughly the same voltage. This is due to the adjustment of the current controller channel gains such that all of the actuators would operate over the full scale of command voltages for the full device displacement. The second column of the actuation matrix describes the relative voltage commands for rotation around the X-axis. Actuator 1 is the dominant generator of this motion, while actuators 2 and 3 are roughly 1/10 as utilized. The third column of the actuation matrix describes the relative voltage commands for rotation around the Y-axis. Actuator 1 plays a minor role in this case and actuators 2 and 3 play the dominant role with their commands opposite one another, precisely as expected.

The actuation matrix was used to drive the center stage open loop, and the position response is compared to the position command in Figure 5.8.

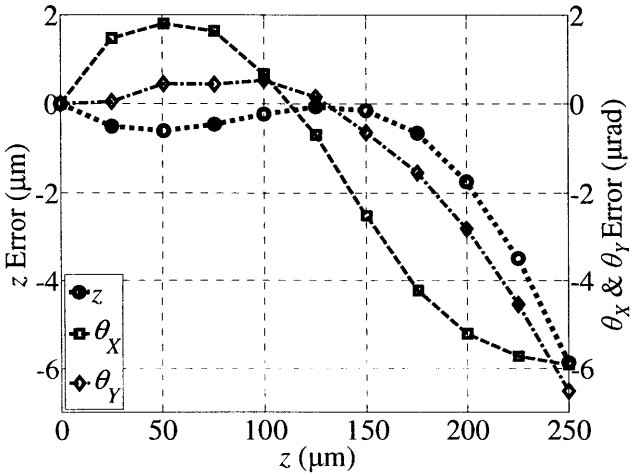

Figure 5.8: Actuator calibration error in all 3 axes.

The error in the actuator calibration is within 1% for vertical displacements up to about 150 μm, beyond which the actuators non-linearity increases significantly in all axes. This is likely due to the actuators moving out of the magnetic field, lowering the average value of the magnetic flux density observed by the coil. The coil force output drops as the flux density decreases, resulting

in growing error between position command and measurement. This is particularly noticeable in the Z axis, where the error remains nearly constant around zero until a 150 μm, at which point the error begins growing exponentially relative to further displacement. The field non-linearity is addressed in the fully optimized actuator design presented in Appendix A.

5.6.3 Range Restriction

The current controller interaction with the sensor electronics is illustrated through the SG WB sensor based actuator calibration data. The interaction is likely due to a linked ground between the current controller circuitry and the sensor circuitry. Only the negative current commands on the actuator show problems, as shown in Figure 5.9. The entire negative range of the device is essentially un-usable as the coupling between the electronics causes significant variations in the sensor output which do not correspond to actual stage displacements.

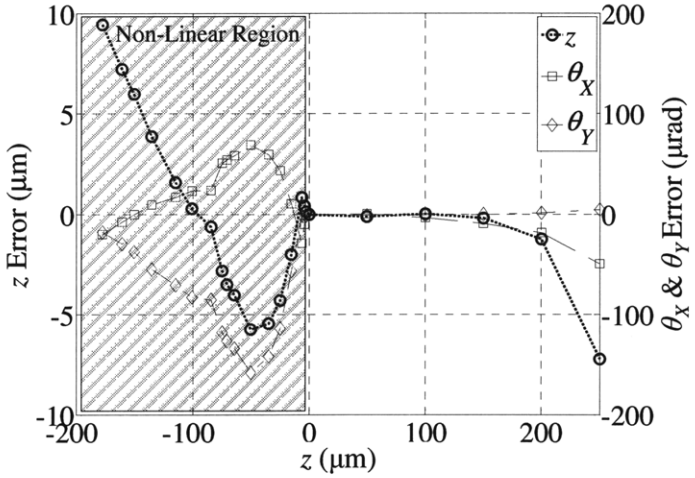

Figure 5.9: Actuator calibration error over full range.

It appears that when the current controller is commanded in the negative direction, the switch in the direction of current flow somehow raises or lowers the ground voltage to which the sensor is linked. All direct ground links have been removed and the current controller is run on an independent power supply, but the issue has not been resolved. The conductive optical table on which the device is anchored may be acting as the ground link. Simultaneous measurements of the device position with CAP probes confirm that the WB sensor readings are not capturing actual device motion. The CAP probes additionally show a tripling of noise when the current

controller is operating in the non-linear region. The HSFD is only operated over positive displacements in lieu of rebuilding all of the electronics.

5.6.4 Coil Characterization

The coils properties are shown in Table 5.6.

Table 5.6: Measured coil properties.

Coil	Turns	R ±0.01 (Ω)
1	102	3.24
2	105	3.25
3	105	3.36

The number of turns in coil 1 is less than the other two coils because damage to the coil during the device assembly required that several turns be removed to regain a usable electrical connection. The resistance of coil 3 is significantly higher than the other two coils, possibly because of an error in the counted number of turns.

5.6.5 Flux Density Measurement

The average gap flux in the actuators can be measured by comparing the full actuator channel characterizations described in Section 5.6.1 to models of the channel containing the bearing, current controller and actuator properties, as shown in Figure 5.7 previously.

The current controller is operating in steady state, therefore only the channel gain is of importance. The current controller is modeled as in Equation (3.85) for each channel. The equations can be stacked to form a diagonal matrix, $\boldsymbol{D}$, defining the current output vector, $\boldsymbol{i}$, in relation to the voltage command vector input to the current controller channels:

$$\boldsymbol{i} = \mathbf{D} \cdot \boldsymbol{v}$$

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} = \begin{bmatrix} \left[-\frac{R_2}{R_i R_s}\right]_{ch1} & 0 & 0 \\ 0 & \left[-\frac{R_2}{R_i R_s}\right]_{ch2} & 0 \\ 0 & 0 & \left[-\frac{R_2}{R_i R_s}\right]_{ch3} \end{bmatrix} \cdot \begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} \tag{5.13}$$

The current controller gain matrix can be used to relate current to position through the matrix $\boldsymbol{E}$, which is generated from Equations (5.11) and (5.13):

$$\begin{aligned} \boldsymbol{p} &= \mathbf{E} \cdot \boldsymbol{i} \\ \boldsymbol{p} &= \left(\mathbf{C}\,\mathbf{D}^{-1}\right) \cdot \boldsymbol{i} \end{aligned} \tag{5.14}$$

The first row of the current to position matrix gives a measured value for the ratio of z/I for each coil. These ratios combined with the known bearing stiffness and average coil wire length, L_{active} enable Equation (3.52) to be used to calculate the actual gap flux density as

$$B_{gap} = \frac{k_Z \sum_{i=1}^{3} \mathbf{E}_{1,i}}{3\bar{L}_{active}}. \tag{5.15}$$

The average wire length is calculated from Equation (3.46) using the measured number of turns in each coil as shown in Table 5.6 previously, assuming every turn is exposed to magnetic flux. The measured value of the gap flux density is 0.785 T, about 7% higher than the predicted value of 0.729 T made by the multi-flux path model. The error is likely caused by differences in the actual gap width from the design dimension and variation in the permanent magnet properties, as the quoted remnant flux and coercivity of NE52 magnets varies by about 5% between manufacturers.

5.7 Sensor Characterization

5.7.1 Method

The sensor properties are determined by commanding the stage over a range of displacements while simultaneously measuring all of the strain gage WB sensor voltage outputs, the WB PS voltage and the actual stage position. Initial offsets of the WB sensors limit the energizing voltage of the bridges to be 15.43 V, rather than the planned 20 V. CAP probes were used to measure displacement at 1 kHz sampling with 1000-point averages on the data, as done with the earlier characterizations.

The calibrated actuators are given open loop commands during operation over a set of displacements covering the device operating range. Immediately before each actuator command is given, the position measurement from the CAP probes and sensor voltages are zeroed to only record the change in both position and voltage when the stage moves. This is intended to remove slow drifts in the readings, such as thermal growth. The initial offset value of the SG voltage for

each sensor (measured at the beginning of the calibration) is added back to the differential voltage measurements to regenerate the full voltage signals associated with a given differential displacement. The HSFD can operate in the 3-D space defined by its three DOFs, but the desired operation is largely 1-D, only along the Z-axis over the range of ±250 μm, with minimal rotations of less than 10 μrad. The calibration is carried out along the center (rotations less than 1 μrad) of this line of common operation as shown in Figure 5.10.

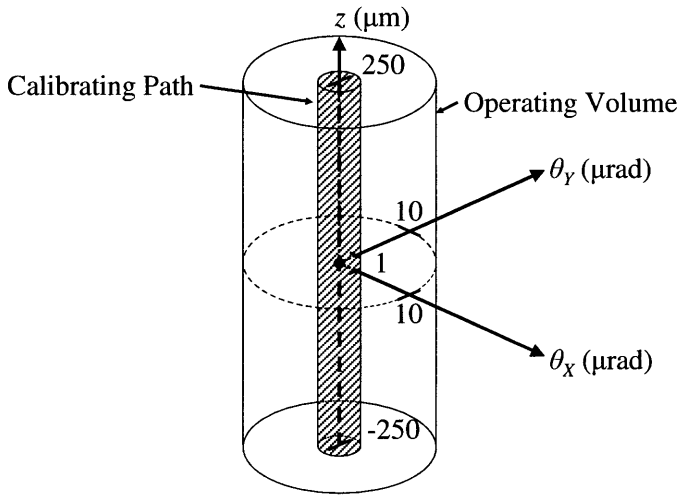

Figure 5.10: Strain gage Wheatstone bridge calibration path.

The calibration essentially captures only the level translation of the entire stage, and ensures that this form of motion is held to the highest accuracy. The non-linear region of the device operation, as described in Section 5.6.3 above, is not included in the calibration.

Sensor calibration is carried out by calculating the vertical displacement of each of the SG WB and comparing this to the measured voltage at the given displacement. A calibration curve can be drawn from multiple measurements of these two variables, which is then used to translate the sensor voltage into vertical displacements at the sensor location. The vertical displacement of the i^{th} SG WB sensor, z_{WBi}, in vector form, $\boldsymbol{s}_{\boldsymbol{WB}}$, can be determined from the measured stage position vector through application of the small angle approximation. The equations calculating the vertical displacement of each sensor can be rewritten into the transfer matrix $\boldsymbol{F}$ as:

$$\mathbf{s}_{WB} = \mathbf{F} \cdot \boldsymbol{p}$$

$$\begin{bmatrix} z_{WB1} \\ z_{WB2} \\ z_{WB3} \end{bmatrix} = \begin{bmatrix} 1 & y_1 - y_{COM} & x_{COM} - x_1 \\ 1 & y_2 - y_{COM} & x_{COM} - x_2 \\ 1 & y_3 - y_{COM} & x_{COM} - x_3 \end{bmatrix} \cdot \begin{bmatrix} z \\ \theta_X \\ \theta_Y \end{bmatrix}. \tag{5.16}$$

The voltage from the SG WB sensor, V_{WBi}, is normalized by the WB PS voltage, V_{PS}, to cancel out voltage fluctuations in the energizing voltage. This creates a non-dimensional voltage term which is nearly perfectly insensitive to power supply variation and noise. A third order least squares fit is applied to the displacement versus non-dimensional voltage curve for each sensor. The best fit is found to be third order, likely because of the slight 2nd order non-linearity of the flexural bearing under high strain as well as the repeatable non-linearity of the instrumentation amplifier:

$$z_{WBi} = a_3 \left(\frac{V_{WBi}}{V_{PS}} \right)^3 + a_2 \left(\frac{V_{WBi}}{V_{PS}} \right)^2 + a_1 \left(\frac{V_{WBi}}{V_{PS}} \right) + a_0, \tag{5.17}$$

where $a_{0\text{-}3}$ are the calibration constants for the sensor.

5.7.2 Results

The calibration coefficients for the SG WB sensors are listed in Table 5.7.

Table 5.7: Calibration coefficients for strain gage Wheatstone bridge sensors.

Sensor	a_3 (μm)	a_2 (μm)	a_1 (μm)	a_0 (μm)
WB_1	-14.2684	-25.5181	374.470	435.396
WB_2	-14.4897	7.11184	388.230	111.5951
WB_3	-7.80499	16.4165	345.987	2.86226

The calibration coefficients show that the second and third order non-linearities of the gages are roughly two orders of magnitude smaller than their first order, linear, term. The system is largely linear, as expected from the sensor and bearing models. The 0th order term corresponds largely to the amplified bridge imbalance (if zero-balance is not carried out on the bridges).

Sensor output voltages at the stage displacement extremes can be determined using Equation (5.17) for each channel. These values are shown in Table 5.8.

Table 5.8: Voltage range of strain gage Wheatstone bridge sensors.

Sensor	Voltage (V) at Δz (μm)		
	-250	0	250
WB_1	-14.5	-9.0	-4.0
WB_2	-7.6	-2.3	2.7
WB_3	-6.0	-0.1	5.4

The offset of each bridge is a significant component of the entire sensor signal and varies over a large range between the channels. The full voltage range of one of the channels exceeds the ±10 V range of the ADC, meaning that the signal would not get picked up over some of the range. The offset in the sensor signals can be removed by biasing the output of the IAs, as described in Section 6.3.10 later.

The error in the SG WB calibration is shown in Figure 5.11. No general form is discernable in the error shown below as the 3rd order calibration has accounted for all of the trends which have an effect significantly greater than the measurement noise.

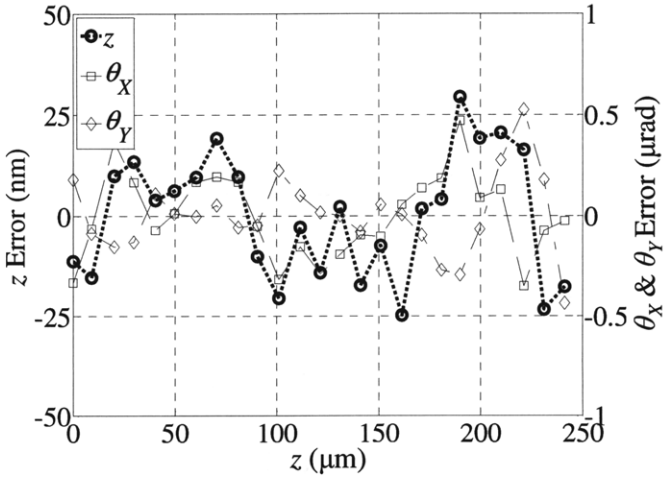

Figure 5.11: Sensor calibration error in all axes.

The error in the Z-axis is on the scale of ≤0.005% of the total device operating range, about 25 nm over 500 μm, significantly exceeding the device requirements for accuracy. The error for rotations is about ≤0.5 μrad. A rapid first order decay ($\tau \approx 5$ s) of magnitude roughly 2% of the step is observed in the in the CAP probe derived vertical displacement reading as well as the SG WB reading upon making a step. This effect is largely gone after 3 time constants (15 s) at which point the data point is taken for the calibration. The error in the SG WB sensors once this

effect is removed is extremely small, indicating that beyond the issue of the first order decay, discussed below, the SG WB sensors are extraordinarily accurate and repeatable, down to the electronic resolution of the system (20 nm). The accuracy of the system can be separated into two components, the high speed, dynamic accuracy in which the first order decay error dominates and the steady state, static accuracy in which the first order decay has dissipated and the inherent strain gage error dominates. The bandwidth of the first order decay error is roughly 0.2 Hz, corresponding to a 5 s time constant. Stepping operations for this prototype at speeds significantly above 0.2 Hz will observe the dynamic accuracy while operations at speeds significantly below 0.2 Hz will observe the static accuracy.

The calibration method described above, is intended to be done in such a short time that thermal growth is removed from the differential measurement. The zeroing of the position and voltage immediately before each measurement result in the case that only the change caused by the stage motion (high speed) and not the thermal expansion (low speed) are measured, especially if the measurement is made rapidly. A rapid first order decay ($\tau \approx 5$ s) is observed in the in the CAP probe derived vertical displacement reading as well as the SG WB sensors, however the two sensor types read the decay in opposite directions- the CAP probe reading shows a drop, while the SG WB reading shows a rise. The decay time is about three orders of magnitude faster than the thermal time constants of the device components, as measured in Section 5.11. For both the reason of the opposite direction of decay and the speed of the decay, this error does not appear to be thermally induced. Simple thermal tests of locally heating the outer base around the coil location did not replicate this effect, further suggesting that thermal growth is not the cause.

Mechanical cause is ruled out by the opposite magnitudes of the decay in the two sensors, leaving the main culprit as CAP probe electronic coupling with the current controllers, perhaps through ground. Possible solutions to this issue are discussed in Section 6.3.11 later.

The actual strain scaling factor of the flexures can be calculated using the voltage/position slope $\delta V_{WB}/\delta z$, from the curve fits, in the expression:

$$G_\varepsilon = \left(\frac{\delta V_{WB}}{\delta z}\right)\frac{1}{G \cdot V_{in} \cdot STC \cdot G_F}\frac{1}{F_{\varepsilon z}}. \tag{5.18}$$

The measured strain scaling factor of 0.744 is about 14% lower than the predicted value of 0.846, likely because the assembly process for the Wheatstone bridge is not accurate enough

when done by hand, to place the strain gages within the desired 0.254 mm (0.010 in.) of the base of the flexure.

5.8 Open Loop Characterization

5.8.1 Method

Characterizations of prototype C showed that the higher resonances in the device played a significant role in determining the compensator design. These higher resonances are not described by the simple system model laid out in Section 3.10 earlier. FEA simulation of the device dynamics should indicate the frequency and form of the modes, however the unknown damping ratio of the resonances and the inaccuracies in the simulations leave direct frequency response measurement (Bode plots) as the only effective means for full system identification. Bode plots contain sufficient information in the magnitude and phase diagram to generally identify all of the axis's poles and zeros over the measured frequency range. Plots are taken for each axis when self-excited and for each permutation of driving and measured axis. This results in nine plots, of which the three charts describing the axes self-excitation are most important. The transfer function models for each axis's dynamics are derived from the self-excitation Bode plots. These transfer function models are used to determine a compensation scheme whose closed loop performance meets the required system specifications, as described in Section 5.9 later.

The frequency response of the axes of prototype D are shown overlaid on the measurements of prototype C to demonstrate the effect of the COS/COR alignment which composed the main difference between the two prototypes.

5.8.2 Bode Plots

Bode plots are generated by driving the system of interest with a pure sinusoidal signal and measuring system response to this command. The ratio of the amplitudes of the response over command is calculated for each frequency input, and graphed as the system gain. The phase of the response relative to the signal is graphed as the phase lag. Gain and phase lag of the device are charted over the measured frequency range to create a Bode plot.

The system frequency response is captured through a Matlab Simulink Dynamic Signal Analyzer (DSA) developed by Katherine Lilienkamp for the Trumper Lab [50]. This program outputs a command sinusoid of adjustable amplitude and reads in the system response. Gain and phase lag are automatically calculated and the process restarted for the next frequency in the input vector of frequencies to measure. The system is run at the maximum stable speed possible for the control loop, which is 33 kHz for the prototype D control loop. A frequency cutoff point of roughly 2 kHz is used for the measurements of the prototype D axes, as it gives an extra decade of information beyond the operating range of the device. The gains of the axes are generally found to be on the level of the measurement noise at 2 kHz. The generalized system measured by the Bode plot characterization is shown in Figure 5.12.

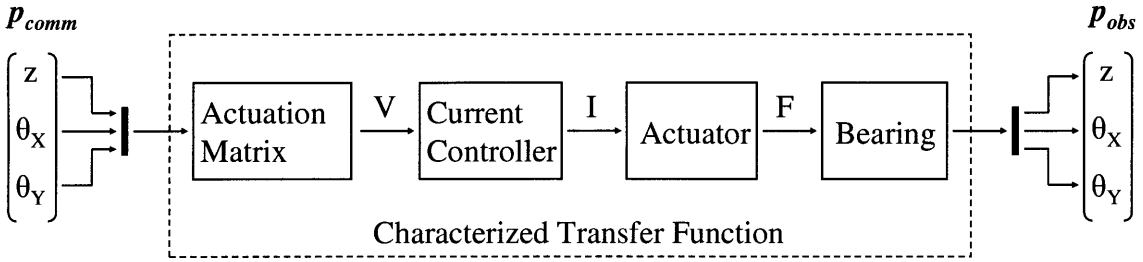

Figure 5.12: Transfer function model of system measured by Bode plots.

The DSA command, p_{comm}, only contains a sinusoidal command on one axis at a time and is fed through the actuation matrix to generate voltage commands for the actuators. This operation is carried out in software and thus does not add any dynamics to the signal. The voltage commands are sent through a DAC to the current controller, which drives a current proportional to the voltage. The current controller adds high frequency dynamics (16 kHz) to the signal, as described in Section 3.7.4 previously, which will not appear in the 2 kHz frequency measurement range. The current controller will mainly contribute a gain to the signal, like the actuation matrix. The actuator describes the transfer from current to force occurring in the actuators. Coil dynamics are not included as they are part of the current controller dynamics. The bearing is the main dynamic system of interest and transforms the force to a position, p_{obs}, of which only one axis at a time is measured by the DSA. All possible permutations of command and measurement are carried out, resulting in nine bode plots. Cross-axis bode plots indicate the degree of axial coupling in the system.

Z-axis motions are measured in μm and rotations around the X- & Y-axis are measured in μrad. The cross-axis gains are not unitless as is normally found in Bode plots, but are rather ratios of the units described above.

5.8.3 Transfer Function Modeling

Transfer function models are fit to each of the self-excitation Bode plots up to 1 kHz. The model for each axis includes the predicted current controller transfer function, *CC(s)*, calculated from the component values described in Table 3.9 previously. The gain of the device forward path transfer function, shown in Figure 5.12, has been normalized by the actuation matrix, thus only the dynamics of the current controller are of interest for model fitting:

$$CC(s) = \frac{1 \cdot 10^{-8} s + 1}{9.1 \cdot 10^{-11} s + 1.33 \cdot 10^{-5} s + 1}. \tag{5.19}$$

A time delay expression, *D(s)*, is also added to the transfer function model to account for the delay added by the discrete nature of the digital controller [51]. This delay is an exponential function of the controller time step, *T*, and the frequency of the signal:

$$D(s) = e^{-sT}. \tag{5.20}$$

The rest of the transfer function model is empirically fit to the measured Bode plot data through a simple process: the phase chart is used to determine the location and number of singularities (poles or zeroes), while the gain chart confirms their location and is used to determine the damping ratio if the poles/zeroes occur in pairs. A 90° phase drop is associated with a pole (180° for pole pair) while a 90° phase increase is associated with a zero (180° for zero pair). The frequency at the midpoint of this drop, 45°, provides an excellent estimate of the singularity characteristic. First order singularities, either single poles or zeroes, $T_1(s)$, are written with unity gain in the form [52]:

$$T_1(s) = \frac{1}{\omega} s + 1, \tag{5.21}$$

where ω is the frequency of the first order singularity. Second order singularities, either pairs of poles or zeroes, $T_2(s)$, are written with unity gain in the form [52]:

$$T_2(s) = \frac{1}{\omega_n^2} s^2 + \frac{2\varsigma}{\omega_n} s + 1 \tag{5.22}$$

where ω_n is the resonant frequency of the pair, and ζ is the damping ratio. These singularities are stacked into a transfer function model with zeroes in the numerator, poles in the denominator, starting from the low frequencies and working up.

5.8.4 Z-Axis Driving Z-Axis

The Z-axis self-excitation Bode plot is compared to the earlier prototype C dynamics, as shown in Figure 5.13. The transfer function model is also overlaid on the measurements to illustrate the degree of alignment between the model and real axis dynamics. The data shows an excellent match to the simple 2nd order system described in Section 3.10.2 earlier, however this match is only up to about 250 Hz, at which point higher order dynamics begin to occur.

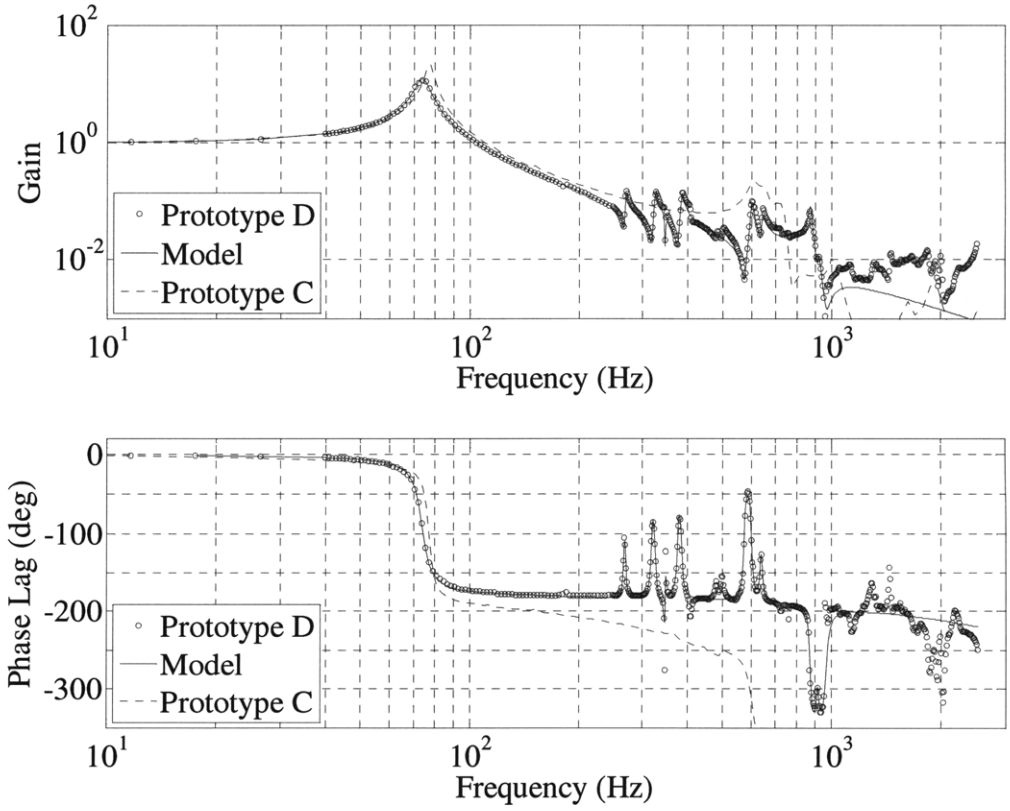

Figure 5.13: Bode plot for Z-axis driving Z-axis.

The singularities in the transfer function model empirically fit to the Z-axis are listed in Table 5.9. The high frequency poles occur as collocated singularities, with both a pole pair and a zero pair. The pairs are highlighted in the chart below.

Table 5.9: Poles and zeros for model of Z-axis driving Z-axis.

Type	Number	ω or ω_n (Hz)	ζ
P	2	74	0.040
Z	2	267	0.007
P	2	270	0.007
Z	2	318	0.010
P	2	326	0.010
P	2	346.8	0.0041
Z	2	347	0.001
Z	2	376	0.010
P	2	386	0.010
Z	2	574.7	0.010
P	2	603.3	0.010
Z	2	637.7	0.009
P	2	644.3	0.009
P	2	874	0.010
Z	2	912	0.010
P	2	920	0.010
Z	2	967	0.017

The most common reason for collocated poles is when the main structure in the system is flexible, and can undergo harmonic modes. These modes generate both a pole and zero pair, which causes only local distortion in the frequency response [51]. The structural proximity of the sensor relative to the actuator plays a large role in determining the location of the zeroes relative to the poles- with closer location resulting in a greater degree of pole-zero overlap [53]. The flexible structure in the case of the HSFD is the center stage, which has relatively low rigidity owing to the geometric constraints on the structure. The assumption of center stage rigidity appears to hold until about 300 Hz, at which point collocated pole/zero pair sets begin to occur with great frequency. These resonances occur across the three main axes in some cases, indicating modes which the coordinate transform mathematics translates into motion in multiple axes. Multi-axis resonances driving both the Z-axis and θ_X-axis occur at roughly 270, 320, 500, 600 and 750 Hz. A multi-axis resonance driving both the Z-axis and θ_Y-axis occurs at roughly 350 Hz. All three axes are driven to resonance at 400 and 900 Hz. These multi-axis resonances are the largest impediment for Z-axis control, as explained in Section 5.9 below. These dynamics can be attenuated by adding greater reinforcement to the center stage structure as is

described in Section 6.3.4 later. Comparison of the Bode plots between the two prototypes (C & D) indicates that the COS/COR alignment did not have any significant effect on the Z-axis or the multi-axis resonances occurring at >300 Hz. This is likely because the COS/COR alignment plays no role in the rigidity of the stage, and only alters the lower frequency dynamics where the center stage rigidity assumption still holds.

5.8.5 θ_X-Axis driving θ_X-Axis

The θ_X-axis self-excitation Bode plot is compared to the earlier prototype dynamics, as shown in Figure 5.14. The transfer function model is also overlaid on the measurements to illustrate the degree of alignment between the model and real axis dynamics.

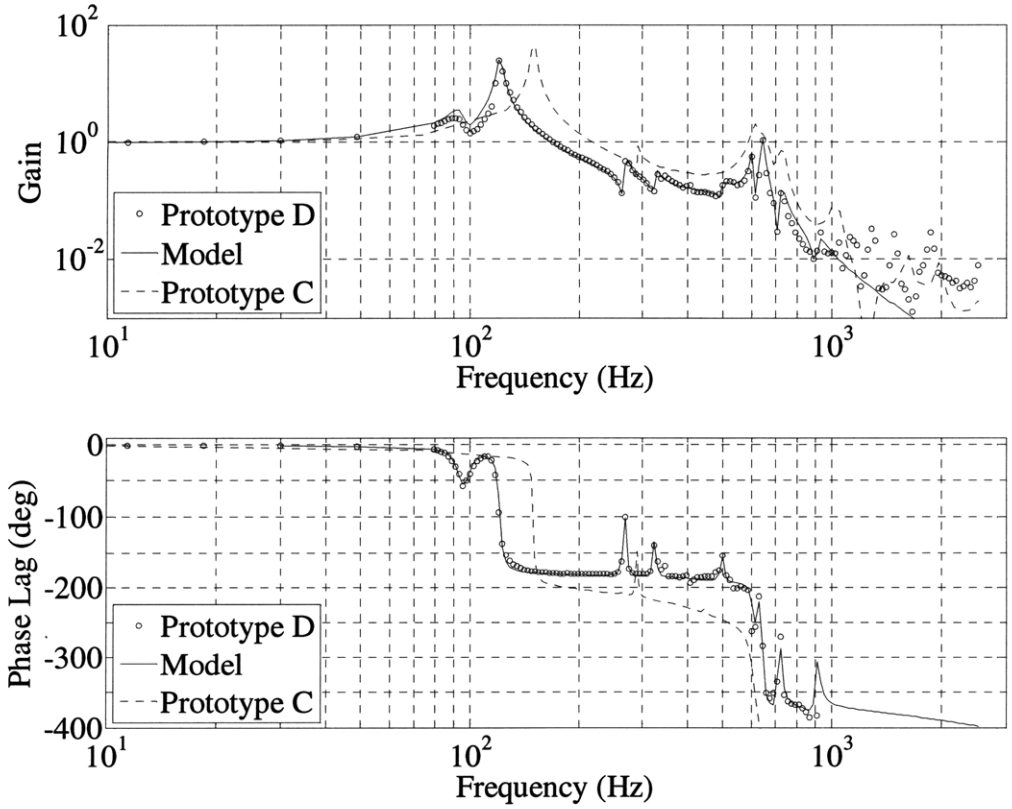

Figure 5.14: Bode plot for θ_X-axis driving θ_X-axis.

The singularities in the transfer function model empirically fit to the θ_X-axis are listed in Table 5.10. The collocated pole and zero pairs are highlighted in the chart below.

Table 5.10: Poles and zeros for model of θ_X-axis driving θ_X-axis.

Type	Number	ω or ω_n (Hz)	ζ
P	2	94	0.045
Z	2	98	0.045
P	2	121	0.015
Z	2	265.6	0.009
P	2	270.4	0.009
Z	2	323.2	0.010
P	2	326.8	0.010
Z	2	497.1	0.010
P	2	500.9	0.010
P	2	607	0.005
Z	2	615	0.005
P	2	646	0.010
Z	2	710	0.006
P	2	726	0.014
Z	2	903	0.010
Z	2	919	0.021

The main resonance frequency of the θ_X-axis decreased between prototypes C and D. This is likely caused by the reduction in flexure width between prototypes, which lowered the effective shear stiffness of the flexures (as well as the torsional stiffness of the flexure anchors), the dominant stiffness for the θ_X-axis. The Z-θ_X resonances described previously are apparent in this bode plot at 270, 320, 500, 600, 750 Hz and create a significant peak in the gain plot which even slightly exceeds unity around 600-750 Hz. Almost every high frequency pole in the θ_X-axis is a multi-axis resonance. The full Z-θ_X-θ_Y resonances at 400 and 900 Hz do not significantly affect the frequency response in the θ_X-axis.

5.8.6 θ_Y-Axis Driving θ_Y-Axis

The θ_Y-axis self-excitation Bode plot is compared to the earlier prototype dynamics, as shown in Figure 5.15. The transfer function model is also overlaid on the measurements to illustrate the degree of alignment between the model and real axis dynamics.

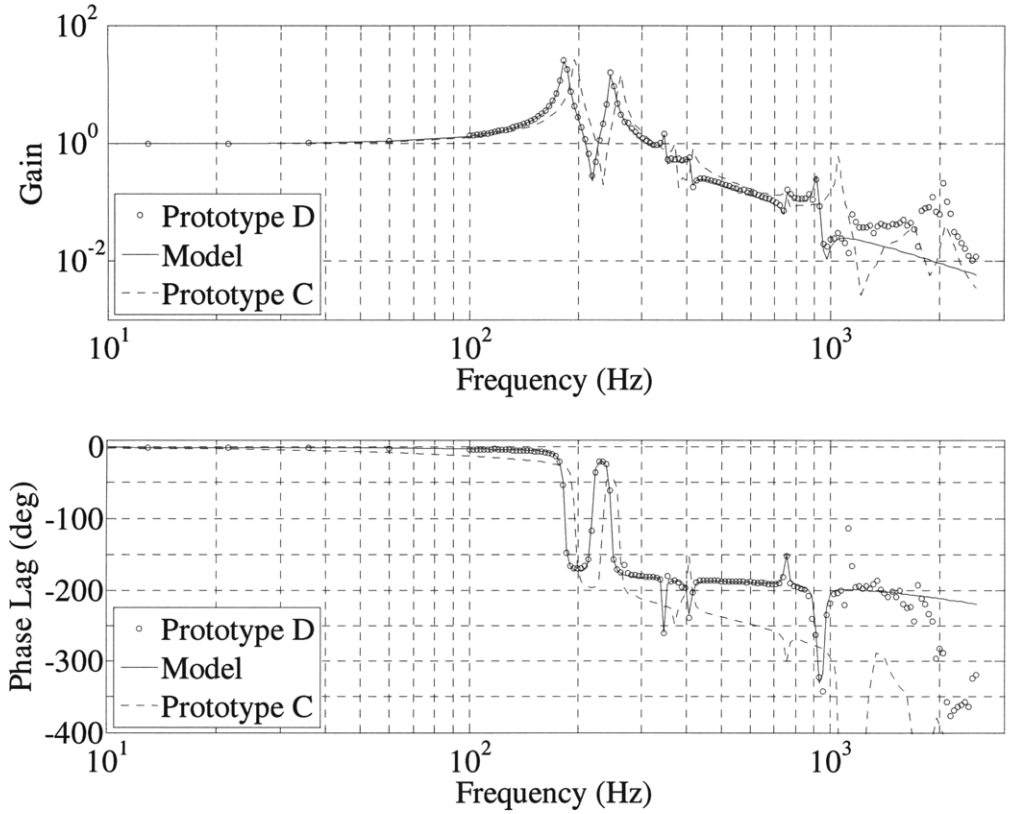

Figure 5.15: Bode plot for θ_Y-axis driving θ_Y-axis.

The singularities in the transfer function model empirically fit to the θ_X-axis are listed in Table 5.11. The collocated pole and zero pairs are highlighted in the chart below.

Table 5.11: Poles and zeros for model of θ_Y-axis driving θ_Y-axis.

Type	Number	ω or ω_n (Hz)	ζ
P	2	183.7	0.010
Z	2	219	0.010
P	2	247	0.010
P	2	346.3	0.0041
Z	2	348.5	0.0027
P	2	407.6	0.010
Z	2	412.4	0.010
Z	2	752.1	0.010
P	2	755.9	0.010
P	2	915	0.010
Z	2	965	0.010

The same reduction in resonances as shown in the θ_X-axis previously can be seen in the θ_Y-axis. The reduction is not as large, likely due to the reason of the θ_Y-axis stiffness being dependent on the tensional mode of the flexures, which was not greatly altered between prototypes. Several large resonances occur below 300 Hz in the θ_Y-axis, in particular at 220 and 250 Hz. Few high frequency resonances are apparent in this axis, however the ones that do appear are multi-axis resonances. The Z-θ_X resonance at 350 Hz is clearly apparent, as are the full Z-θ_X-θ_Y resonances at 400 and 900 Hz yet it is significantly lower than the peak observed in prototype C.

5.8.7 Z-Axis Driving θ_X-Axis

The Z-axis driving θ_X-axis Bode plot is compared to the earlier prototype dynamics, as shown in Figure 5.16.

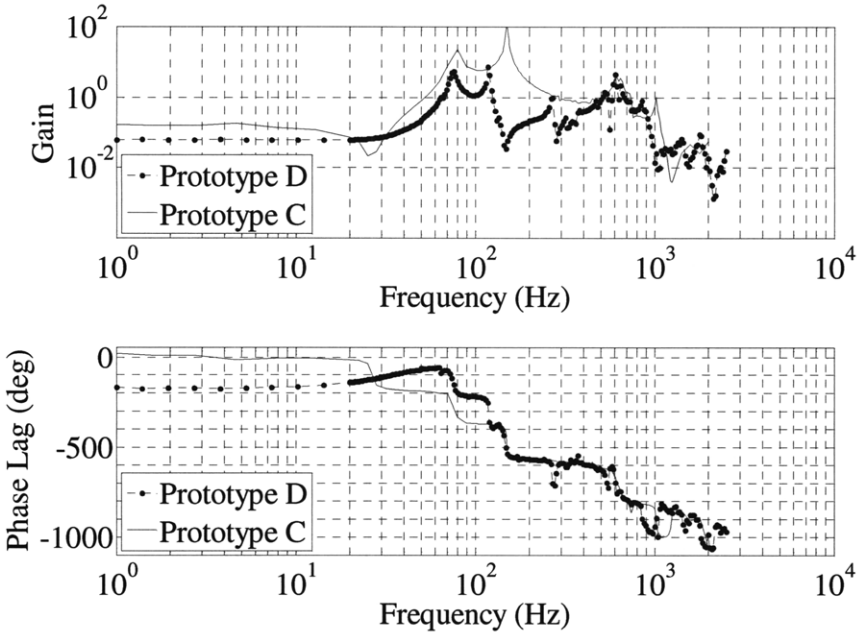

Figure 5.16: Bode plot for Z-axis driving θ_X-axis.

The gain is significantly less than one, due to the actuator calibration. At low frequencies, the gain sits around 0.05. Beyond the Z-axis primary resonance, the dynamics of the two axes alter the plot. The gain at the θ_X-axis primary resonance is greatly attenuated for prototype D over prototype C, this is the result of the COS/COR alignment. There is no

significant attenuation of the dynamics at frequencies above 400 Hz, suggesting that the rigid body assumption of the center stage holds below 400 Hz, but not above. The Z-θ_X resonance shows in this plot as a significant and wide peak, as both axes contribute. The Z-θ_X-θ_Y resonance is barely visible in the peak of the early resonances.

5.8.8 Z-Axis Driving θ_Y-Axis

The Z-axis driving θ_Y-axis Bode plot is compared to the earlier prototype dynamics, as shown in Figure 5.17.

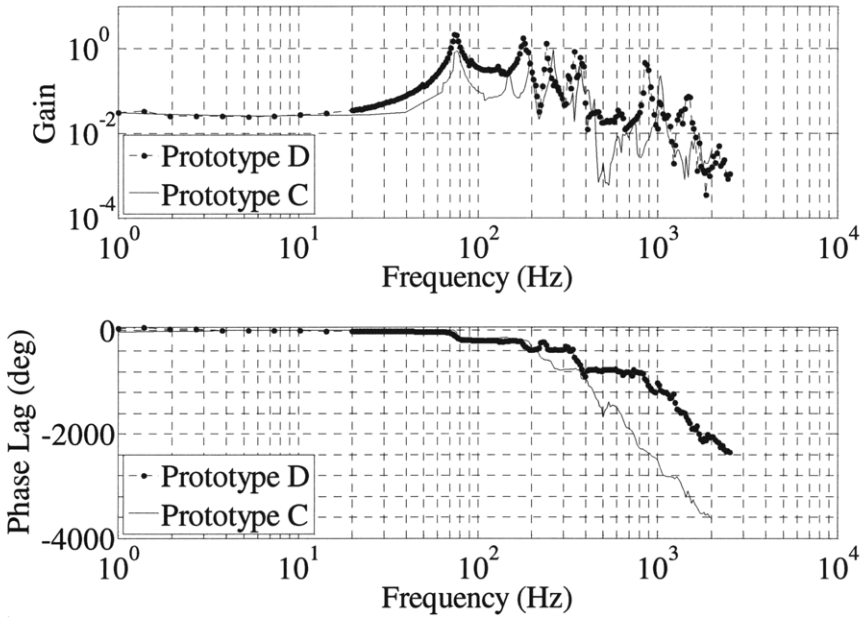

Figure 5.17: Bode plot for Z-axis driving θ_Y-axis.

The gain sits around 0.01 at low frequencies, like the cross-axis plot shown previously. Beyond the Z-axis primary resonance, the dynamics of the two axes alter the plot. The gain in the region from 75 to 200 Hz is significantly larger than in prototype C. It is believed that this is due to the large value of x_{COR}, which causes the COR to interchange energy with the θ_X-axis, more than in prototype C. Beyond 400 Hz, the error is likely due to measurement resolution limits in the prototype C data. The Z-θ_X-θ_Y resonance appears in this cross-talk plot as a sharp peak with roughly unity gain.

5.8.9 θ_X-Axis Driving Z-Axis

The θ_X-axis driving Z-axis Bode plot is compared to the earlier prototype dynamics, as shown in Figure 5.18.

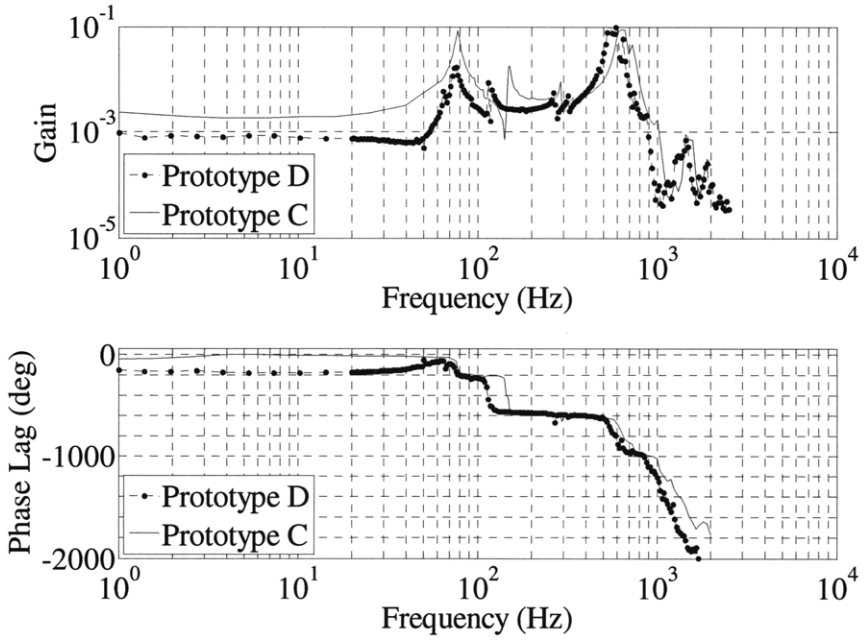

Figure 5.18: Bode plot for θ_X-axis driving Z-axis.

The low frequency gain is significantly less than observed in prototype C, likely because of better calibration. The dynamics from the Z-axis main resonance up to about 400 Hz are also attenuated, while the center stage compliance resonances are not altered beyond being shifted down in frequency. The Z-θ_X resonance shows the highest gain of about 0.1. Over the entire frequency range, the cross-axis gain is extremely small.

5.8.10 θ_X-Axis Driving θ_Y-Axis

The θ_X-axis driving θ_Y-axis Bode plot is compared to the earlier prototype dynamics, as shown in Figure 5.19.

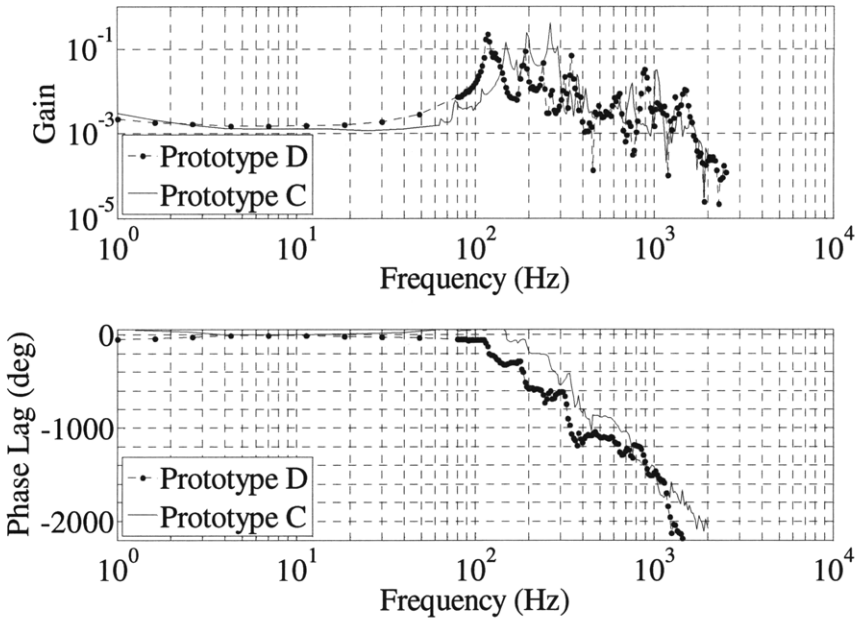

Figure 5.19: Bode plot for θ_X-axis driving θ_Y-axis.

The large numbers of resonances from the θ_X- and θ_Y-axes generate the complex frequency response between roughly 100 and 1000 Hz. The gain between 100-400 Hz is lower for prototype D, but no variation is observed above that range. The exact gain between these axes is difficult to measure as the thousand-fold reduction in input generates an output on the same range as the measurement resolution. Over the entire frequency range, the cross-axis gain is extremely small.

5.8.11 θ_Y-Axis Driving Z-Axis

The θ_Y-axis driving Z-axis Bode plot is compared to the earlier prototype dynamics, as shown in Figure 5.20.

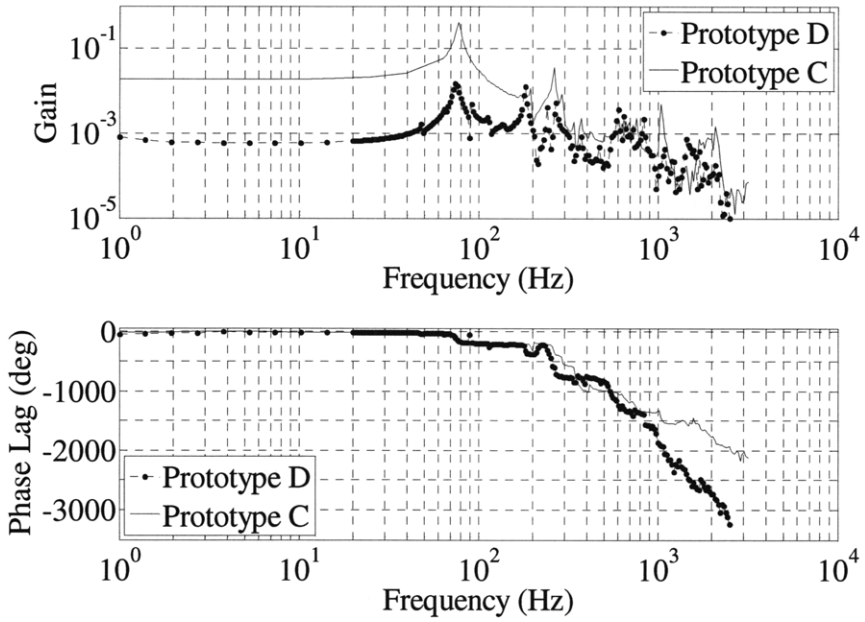

Figure 5.20: Bode plot for θ_Y-axis driving Z-axis.

The low frequency gain for this cross-axis response is significantly reduced in prototype D, it is believed this is a result of better calibration. The attenuation in the system dynamics from the Z-axis main resonance out to 400 Hz suggests that the effect is also due to the COS/COR alignment. Low gain is found over the entire frequency range.

5.8.12 θ_Y-Axis Driving θ_X-Axis

The θ_Y-axis driving θ_X-axis Bode plot is compared to the earlier prototype dynamics, as shown in Figure 5.21.

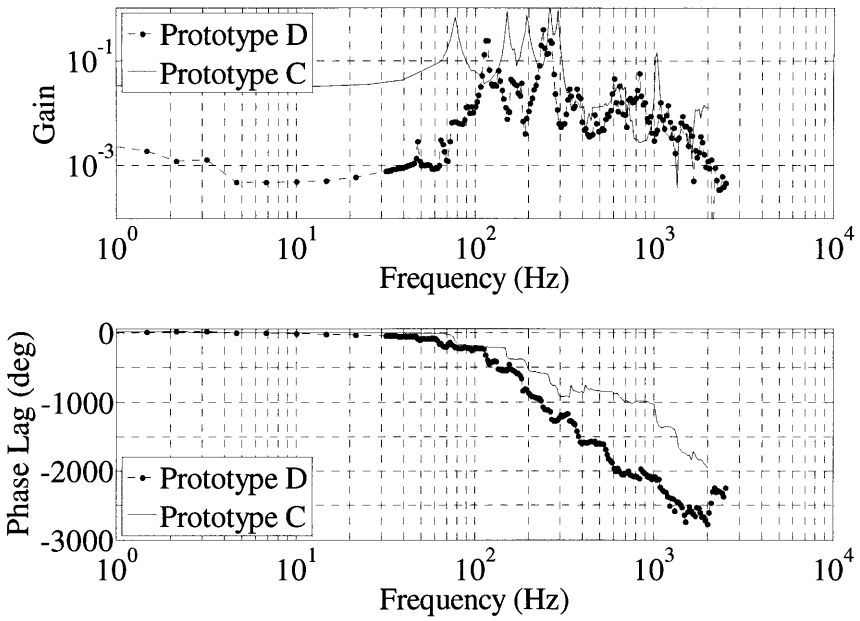

Figure 5.21: Bode plot for θ_Y-axis driving θ_X-axis.

The gain of the cross-axis plot is again low for low frequency signals, but rises almost to unity at one of the resonances of the θ_Y-axis. The dynamic range from the Z-axis primary resonance out to 400 Hz is nearly completely attenuated, beyond which little variation is observed between the two prototypes. The COS/COR alignment appears to have removed the Z-axis resonance from the frequency response, likely because the misalignment was allowing energy to sneak between the rotations through the Z-axis.

5.9 Control Loop Tuning

5.9.1 Method

The system control scheme can be determined once the open loop response is accurately known. The empirically fit transfer function models provide an excellent starting point for this design process. The design of the compensation is carried out by examining the location of the singularities in the transfer function model and adding in poles and zeroes to the open loop response to modify the device dynamics. The performance of the control scheme is confirmed using simulation of the axis closed loop response to a unit step command.

Three axis control is required to ensure proper (flat) stepping at the desired speed. The ideal case is three independent control loops, each of which is designed specifically from the self-excitation Bode plot of the corresponding axis. This method would enable the most rapid implementation and simplest tuning process.

5.9.2 Cross Talk

The feasibility of the parallel control loop scheme is dependent on the degree of cross-talk which occurs between the three axes of motion. The cross talk can be quantified using the information gained in the Bode plots shown above. The axial gain coefficients for each of the permutations of driving and reading are listed in Table 5.12. These coefficients describe how well signals are able to propagate through the system axes. Prototype C results are shown below the Prototype D results to show the effect of the COS/COR alignment.

Table 5.12: Axial cross talk coefficients.

	Z	θ_X	θ_Y	Z	θ_X	θ_Y	Measured axis
Z	1	0.06	0.03	20	8	2	
	(1)	(0.2)	(0.03)	(20)	(100)	(1)	
θ_X	0.001	1	0.001	0.01	50	0.1	
	(0.002)	(1)	(0.001)	(0.9)	(50)	(0.3)	
θ_Y	0.0007	0.001	1	0.02	0.4	30	
	(0.02)	(0.03)	(1)	(0.4)	(1)	(30)	
Driving axis		Static Coupling		Prototype D (Prototype C)	Dynamic Coupling		

The static coupling matrix describes the zero frequency response of each of the system axes to each possible command. Steady state gains are mainly a function of the degree of orthogonalization of the actuator calibration. The better this calibration is carried out, the less that motion occurs in non-commanded axes. Each column describes the particular axis response to the commands from all the axes. The ideal case for independent operation is an identity matrix for the static coupling. The measured static coupling is nearly (<5% error) ideal, with most of the cross-axis coefficients being roughly three orders of magnitude less than the self-actuation command coefficient. This means that the command signal for low frequencies can be approximated as being read only by its own axis. The Z-axis driving the θ_X-axis is the largest

off-axis coefficient, and corresponds to a θ_X-axis response that is 94.2% controlled by commands set to the θ_X-axis (desired), 5.6% controlled by commands sent to the Z-axis (undesired) and 0.1% controlled by commands sent to the θ_Y-axis, (undesired). The prototype C static coupling is worse, particularly in the θ_X-axis.

The dynamic coupling describes the axes' response to commands given at the measured axes' primary resonance frequency: 75 Hz for the Z-axis, 119 Hz for the θ_X-axis, and 182 Hz for the θ_Y-axis. The dynamic coupling gains give a measure of the energy leaking between dynamically active axis through geometric effects like the COS/COR coupling. Each column describes the particular axis response to the commands (at its primary resonance) from all the axes. The ideal case is an identity matrix, as with the static coupling. The measured dynamic coupling approaches this ideal case for the Z-axis, is slightly worse for the θ_Y-axis, and shows the most significant deviation in the θ_X-axis. The coupling between the Z- and θ_X-axis is again the dominant problem. For axial commands of equal amplitude, the θ_X-axis response is 85.6% controlled by the θ_X-axis commands, 13.7% controlled by the Z-axis commands, and 0.7% controlled by the θ_Y-axis commands. The self-actuation command still dominates the signal, if not quite as completely as in the static coupling. The prototype C dynamic coupling is worse in nearly every cross-axis case, especially for the Z-axis driving θ_X-axis coefficient, which is the most important to reduce.

In both static and dynamic cases, the self-actuation commands are the dominant controlling signals on the axes, indicating that independent control should be possible. The Z-axis driving θ_X-axis case is the main concern when operating under the assumption of independent control. COS/COR has reduced this cross-axis coefficient by over 10-fold, as well as nearly every other dynamic cross-axis coefficient. Based on this improved decoupling, a control scheme of three independent control loops for the three axes is used for closed loop control of the HSFD.

5.9.3 Z-Axis Control Loop

The initial Z-axis control loop design of a proportional-integral-derivative (PID) controller, explained in Section 5.9.4 below, was found to be unsuitable for the high bandwidth required on the axis. The cause of this unsuitability is the large high-frequency multi-axis resonances in the Z-axis. Several design requirements were placed on the compensation scheme-

90% settling time of 10 ms, closed loop bandwidth >100 Hz, zero steady state error to a step command, fine positioning error decay constant <10 ms. Practical limitations on the controller DAQ hardware set the sampling rate to be 25 kHz and required that the compensation scheme have minimal complexity to ensure real-time performance. Finally, the steady state behavior of the system should be as stable as possible, without the controller introducing added oscillations (mechanical noise) to degrade the resolution. A compensation design incorporating lead, and ideal integral compensation is investigated based on the helpful suggestions of Aaron Mazzeo and Kevin Miu, graduate students working in the Precision Motion Control Laboratory under Prof. David Trumper. Notch filters are included in the compensation scheme to reduce the gain of the high frequency poles.

Ideal integral (II) compensation is alternately referred to as Proportional-Integral compensation, and acts like a lag compensator with an integrator- a pole at the origin. The integrator raises the system by a type and ensures zero steady state error to a step command. The combination of an integrator and a low frequency zero acts to nearly cancel out high frequency dynamics from the II compensation while allowing the compensator to drag the system steady state error down to zero. The net result is a compensation which adjusts the steady state behavior of the system response, such as steady state error and fine settling time. The fine settling time of the system is caused by the single closed loop pole which moves along the root locus from the integrator to the zero as gain is increased. The higher this pole is, the faster the system settles to the command value. When designing the Ideal Integral compensator, the free variable is the location of the zero, raising it drags the closed loop pole further from the origin, speeding up the fine settling. An ideal integral compensation shifts down the phase at low frequencies by 90° due to the integrator. The higher frequency zero returns the phase to 0°, however this takes up to roughly $10x$ the frequency of the zero to achieve within 5°. Generally the zero is placed at least a decade lower frequency than the crossover frequency to ensure no phase delay. The final zero location is determined by a compromise between these two design goals [51]. The HSFD compensation scheme places the zero within a decade of the crossover frequency and the reduction in phase margin is offset with increased phase from the lead compensation, allowing for faster fine settling. The ideal integral compensator *II(s)*, with unity gain, is given as:

$$II(s) = \frac{\frac{1}{450}s + 1}{s}. \tag{5.23}$$

Lead compensation acts as a PD compensator (a zero) with the addition of a pole placed at a higher frequency (generally about 5x-10x) than the zero. The lead and PD enable adjustment of the system transient behavior such as setting time and overshoot. The zero compensation adds phase to the system at the zero frequency and above, increasing the system stability and therefore improving the dynamic response. The zero also causes an upward slope in the gain with respect to frequency, which would result in infinite gain at infinite frequency. This will cause instability due to white noise. The high frequency response is mitigated by placing a pole higher than the zero, such that the gain increase is not infinite. Lead compensation is generally designed around phase addition to the frequency response, with the location and amount of phase setting the various parameters in the compensator. Lead compensators, *L(s)*, can be written as:

$$L(s) = \frac{\alpha\tau s + 1}{\tau s + 1} \quad \alpha > 1, \tag{5.24}$$

where α is a dimensionless ratio corresponding to the separation between the pole and the zero frequencies and τ is the frequency of the higher singularity- the pole. The maximum phase gain generated by the lead compensation, φ_{max}, is calculated by [51]:

$$\varphi_{max} = \sin^{-1}\left(\frac{\alpha - 1}{\alpha + 1}\right). \tag{5.25}$$

This phase gain occurs at ω_{max}, the geometric middle of the pole and zero in the frequency plot, which is calculated by [51]:

$$\omega_{max} = \frac{1}{\tau\sqrt{\alpha}}. \tag{5.26}$$

Generally, the maximum phase gain is placed at the crossover frequency to boost the system dynamic response and stability. The lead compensation for the HSFD is designed to add about 65 degrees of phase at ≈200 Hz, (1300 rad/s), resulting in a value of α of about 20 and τ of about $1.71x10^{-4}$. This is about the maximum accepted compromise between phase margin and high frequency gain increase. The higher order complexity of the axis requires that the phase margin be shifted up to stabilize the higher frequency resonances, resulting in a lead compensator, *Lead(s)*, with unity gain, of:

$$Lead(s) = \frac{\frac{1}{288.2}s + 1}{\frac{1}{5864}s + 1}. \tag{5.27}$$

A notch filter is used in the compensation to remove the high-frequency multi-axis resonances from the open loop Z-axis response. The filter creates a drop in gain over a tight frequency range as well as an 'S' curve in the phase lag through the use of the resonance peak of a zero pair which is then canceled with a critically damped pole pair. The notch compensation, N(s), can be written as:

$$N(s)=\frac{\frac{1}{\omega_n^2}s^2+\frac{2\zeta}{\omega_n}s+1}{\left(\frac{1}{\omega_n}s+1\right)^2}, \tag{5.28}$$

where ω_n and ζ are the frequency of the notch and the damping ratio of the zero pair. The under-damped zero pair creates a notch at the natural frequency, about which the gain rises at a slope of 2. The critically damped poles do not affect the notch, but cancel out the slope, leaving only the dip in gain at the notch filter frequency. The phase drops due to the critically damped poles, then shifts up by 90 at the filter frequency from the under-damped zeroes and continues dropping back down to 0 at significantly higher frequencies. The phase loss at lower frequencies reduces the phase margin for a system like the HSFD where the crossover frequency is below the notch filter frequency. This is solved through the addition of a second lead compensator to add phase to the system. Two notch filters are used, and are aligned with the main sets of collocated poles in the Z-axis, at 318 Hz (2000 rad/s) and 875 Hz (5500 rad/s), each having a damping ratio of 0.1 for the zeroes. This damping ratio is chosen to attenuate the gain as much as possible (lower damping ratio) while reducing the negative phase added by the notch (raising damping ratio). The equation describing the full notch filter compensator *Notch(s)*, with unity gain, is:

$$Notch(s)=\frac{\frac{1}{2000^2}s^2+\frac{2\cdot 0.1}{2000}s+1}{\left(\frac{1}{2000}s+1\right)^2}\cdot\frac{\frac{1}{5500^2}s^2+\frac{2\cdot 0.1}{5500}s+1}{\left(\frac{1}{5500}s+1\right)^2}. \tag{5.29}$$

The system open loop gain, *K*, is adjusted to 125 to set the loop transmission bode plot to have a crossover frequency of about 125 Hz (780 rad/s), which sets the closed loop bandwidth to be approximately the same frequency. The system gain is set low enough to ensure that the high frequency resonances are gain stabilized (<0 dB). These modes are observed to cause steady state oscillation in the controller when their gain exceeds unity, likely due to cross axis excitation. The full Z-axis compensation $C_Z(s)$, is composed of the compensators described above and shown in cascaded form as:

$$C_Z(s) = K \cdot Lead^2(s) \cdot II(s) \cdot Notch(s)$$

$$C_Z(s) = 125 \cdot \left[\frac{3.470 \cdot 10^{-3} s + 1}{1.705 \cdot 10^{-4} s + 1}\right]^2 \cdot \left[\frac{2.222 \cdot 10^{-3} s + 1}{s}\right] \cdot \ldots \quad (5.30)$$

$$\left[\frac{2.5 \cdot 10^{-7} s^2 + 1 \cdot 10^{-4} s + 1}{2.5 \cdot 10^{-7} s^2 + 1 \cdot 10^{-3} s + 1} \cdot \frac{3.306 \cdot 10^{-7} s^2 + 3.636 \cdot 10^{-5} s + 1}{3.306 \cdot 10^{-7} s^2 + 3.636 \cdot 10^{-4} s + 1}\right]$$

5.9.4 θ_X-Axis Control Loop

PID compensation is used for the rotational axes. PID compensation can be easily tuned to alter both the system transients and steady state error. A large range of rotations are allowed (10 μm) for the rotational axes, therefore they do not need to be as precisely controlled as the vertical axis during the prototyping phase. The controller hardware also places significant limitations on the complexity of the controlling scheme. The system multi-axis resonances are observed to cause instabilities in the controlled system at high gains. The Z-axis is implemented first in order to meet the system requirements for stability, then the rotational compensation schemes are added in while the Z-axis control is running. This enables direct observation of cross-axis instabilities, should the control loops begin to fight one another. This effect can not be modeled using single-input single-output methods and is observed to be a significant factor when attempting to drive the axial controllers to their design gain values, if multiple loops are running. This problem is surmounted by experimental adjustment of the θ_X- and θ_Y-axis control loops, for which PID compensation is ideally suited.

The PID compensator, $C_{PID}(s)$, can be broken into an integrator and two zeroes. The integrator sets the steady state error to 0 and one of the zeroes is used to negate the transient behavior of the integrator following the design rules described previously for the ideal integral compensation. The other zero is used to raise the frequency of the system poles, driving down the settling time and setting the overshoot. The standard form in which the compensator is written has coefficients for the proportional gain, K_P, integral gain, K_I, and derivative gain, K_D:

$$C_{PID}(s) = \frac{K_D s^2 + K_P s + K_I}{s}. \quad (5.31)$$

The PID compensator for the θ_X-axis is experimentally tuned by adjusting the three coefficients during device operation. The resulting compensator, $C_{\theta X}(s)$, is

$$C_{\theta X}(s) = \frac{2 \cdot 10^{-4} s^2 + 0.3 s + 20}{s}. \quad (5.32)$$

5.9.5 θ_Y-Axis Control Loop

The PID compensator for θ_Y-axis, $C_{\theta Y}(s)$, is found through the same process as described for the θ_X-axis compensator:

$$C_{\theta Y}(s) = \frac{5 \cdot 10^{-5} s^2 + 0.11 s + 20}{s}. \tag{5.33}$$

5.9.6 Input Modulation

A real time filter, *IM(s)*, is used to modulate the input to the control loop, as shown in Figure 5.22, to remove the high frequency components of the command signal. The filter is tuned to modulate a step command to output a smooth ramp-up command, which significantly reduces overshoot and device ringing.

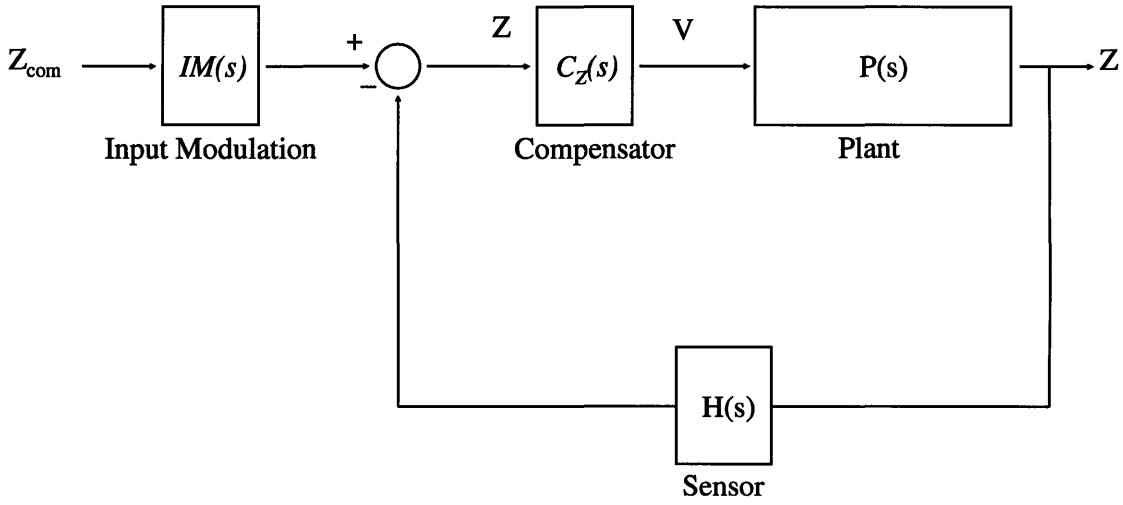

Figure 5.22: Transfer function model of closed loop control with input modulation.

The first component of the filter is a pole pair which modulates the step into the carrier command, a smooth critically damped 2nd order track. The control loop will generate the least excitation of higher modes by tracking this path, however if the response is too slow, then the bandwidth and response time of the system is slowed. The carrier command, *Carrier(s)*, is be generated from a pole pair with natural frequency 350 rad/s (55 Hz) and damping ratio 1,

$$Carrier(s) = \frac{1}{\left(\frac{1}{350} s + 1\right)^2}. \tag{5.34}$$

The second component of the filter, *FSE(s)*, is added to cancel out the observed closed loop fine settling error. This error is generated by a single pole on the real axis at roughly -100, the same pole discussed previously in the lag design. This pole settles much slower than the dominant pole pair, thus even once the main dynamics are completed, the system continues to settle in at a first order exponential decay. This is canceled by a placing a pole-zero pair in the input modulation. The zero of the pair is set on top of the closed loop pole responsible for the fine settling error. The pole is placed at a higher frequency, essentially shifting the fine positioning error pole out to the new pole location. The closed loop fine settling error pole is generated by the compensation dynamics, meaning that its location is relatively well known. The fine settling error for the HSFD is cancelled through a zero at 13 Hz (81 rad/s) and a pole at 38 Hz (238 rad/s),

$$FSE(s) = \frac{\frac{1}{81}s+1}{\frac{1}{238}s+1}. \tag{5.35}$$

The addition of these transfer functions was modeled to ensure that the closed loop bandwidth remains sufficiently high to meet the design goals- above 100 Hz. The total input modulation, IM(s) for the HSFD is:

$$IM(s) = Carrier(s) \cdot FSE(s)$$

$$IM(s) = \left[\frac{1}{8.163\cdot 10^{-6}s^2 + 5.714\cdot 10^{-3}s+1}\right]\cdot\left[\frac{1.235\cdot 10^{-2}s+1}{4.200\cdot 10^{-3}s+1}\right]\cdot \tag{5.36}$$

Further cancellation of other system dynamics is possible, for instance the dominant 2nd order pole oscillations could be largely removed by placing a pair of complex zeroes near these poles. A new set of poles would have to be added in to retain the same general system response, if these are given the same frequency as the zero pair, but with critical damping, then a notch filter is effectively used at the input. This will remove the exitation of the main dynamics, replacing it with a critically damped response. This method is slightly more dangerous as it alters the main system response, which may result in the higher frequency poles becoming more noticeable.

5.10 Closed Loop Response

5.10.1 Step Response

The step response of the closed loop system is measured using all three parallel control loops running at 33 kHz sampling and reading from the SG WB sensors. The resulting response of all axes is shown along with the predicted system response in Figure 5.23.

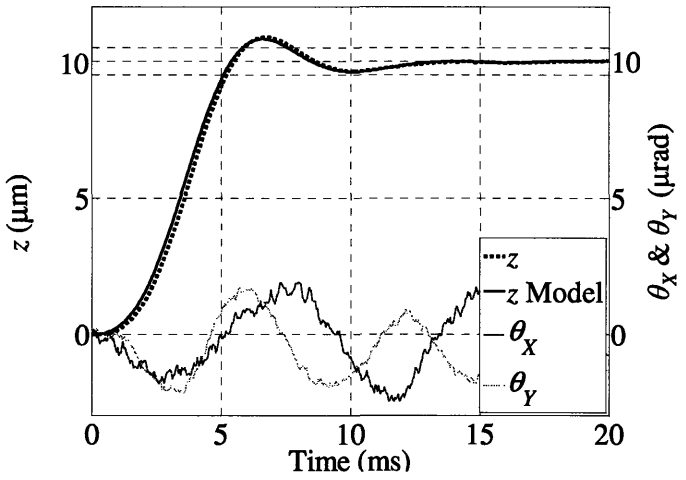

Figure 5.23: Step response for all 3 axes.

The 95% settling line is shown on the figure as the lines bounding the 10 μm mark to indicate the approximate cutoff used for judging the settling time of the device, roughly 8 ms. After this point the stage remains within the 5% bounds indicated on the chart. A step size of 10 μm is used to minimize the effect of measurement error, enabling the dynamic response to be clearly observed. The θ_X-axis receives the largest cross-axis driving signal, but not by a significant amount. Both rotations continue past the 95% settling time of the Z-axis. The amplitude of the oscillations are small, but the lack of settling indicates the need for a better compensation design in these two axes. The frequencies of the axes' rotations are almost exactly double that of the Z-axis oscillations, which may be due to the Z-axis driving the rotations through cross-axis energy transfer.

5.10.2 Continuous Stepping Response

A second type of input command was tested on the HSFD to simulate the desired stepping motion expected during Confocal imaging. This input command is composed of a string of step commands spaced out by 10 ms intervals. The steps are carried out with 500 nm amplitudes to mimic the HSFD making small, quick steps during imaging. The result of the stepping command on all axes for the first three steps is shown in Figure 5.24.

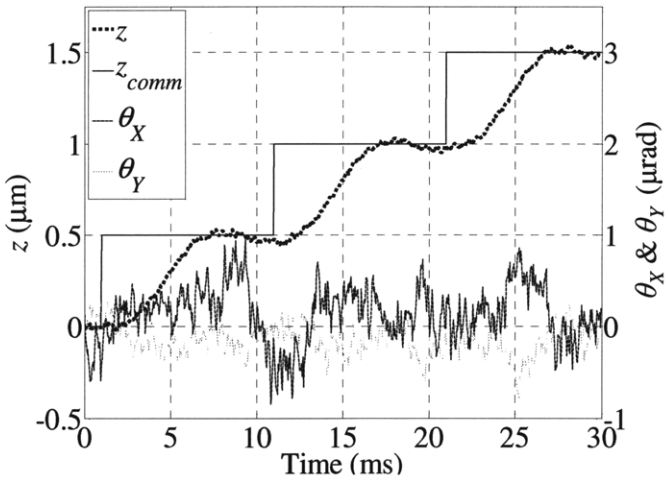

Figure 5.24: Continuous stepping response for all 3 axes.

The transient associated with each step command have nearly completely settled out before the next step begins. The system response would continue to oscillate around the command value for roughly another 2 ms, however the next step command now obscures this slower behavior. The stage rotations are maintained at a constant small state of oscillation that is driven by the stepping command. Better design for the compensation of the rotational axes could significantly reduce this undesired motion, as with the Z-axis. The system is well controlled even though rapid motion is occurring.

5.10.3 Resolution

The system resolution is calculated by observing the system under several different conditions. The electrical noise of the sensors is measured through recording stage position while the device is mechanically restrained via a high damping material. The standard deviation of the position measurement is defined as the root-mean-square of the noise. The open-loop

mechanical noise of the stage is measured through recording the stage position while the device is free to move. The geometric sum of the open-loop mechanical noise and the electronic noise compose the measured noise. The closed-loop noise is measured in the same was as the open loop noise except the system is run closed loop. The system is run at 33 kHz sampling, but the data is stored at 1 kHz. Down-sampling reduces the processing load on the controller and does not alter the standard deviation of the noise. A Fast Fourier transform (FFT) of the closed-loop full system noise is shown in Figure 5.25.

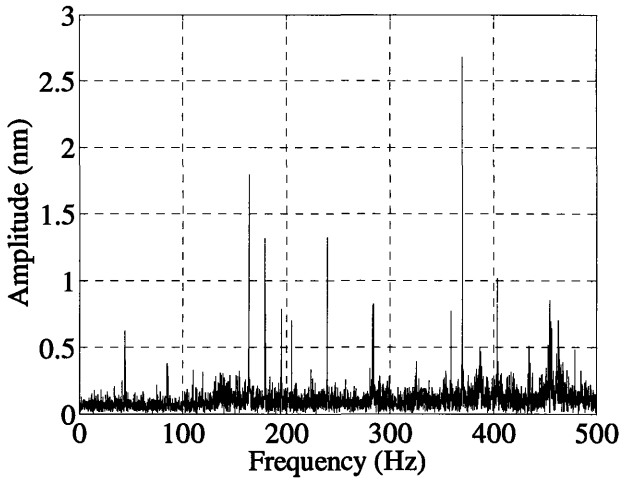

Figure 5.25: Frequency distribution of system noise.

The FFT runs from the frequency associated with the length of the test (9 sec, ≈ 0.1 Hz) up to the Nyquist frequency for the 1 kHz sampling of 500 Hz. A roughly even background noise level can be seen over the whole spectrum, indicating that the Johnson noise is not only present, but is also the dominant form of noise in the system. Low frequency variations are attenuated slightly because the control loop is able to reject the motion. The FFT shows no noticeable mechanical resonance at 200 Hz, but instead what appears to be a resonance at around 450 Hz. This may be a secondary loop resonance. There are a large number of additional frequency spikes in the data, with the largest appearing at roughly 160, 180, 240, 280, and 360 Hz. Three of these frequencies are associated with the harmonics of three-phase power. The sharpness of the peaks suggests an electrical noise source as compared to a mechanical frequency source. All of these peaks are believed to be external noise, as they did not appear in the FFTs of the components noise. The

likely cause of this external noise is insufficient shielding over the instrumentation amplifies and unwanted electronic coupling of the sensor circuitry to the actuator circuitry.

The noise shows a normal distribution, suggesting that while several peaks are observable in the FFT, in majority the noise is occurring as random variations that are characteristic of Johnson noise. The standard deviations of the noise from the electronics, open-loop and closed-loop operation are shown in Table 5.13.

Table 5.13: Noise source characterization.

Type	St. Dev. of Noise (nm)
Electronic	10.5
OL mech.	6.5
CL mech.	8.7
Total CL	13.7

The electronic noise can be considered 'false' information fed into the system, the lower this value is, the better the device can resolve position changes. Actual location changes which occur on the scale of this level or below cannot be accurately measured, setting an effective spatial resolution limit. The system would ideally read in no electronic noise.

The mechanical noise corresponds to actual device motion, and should be measured as well as possible. Only this 'noise' would ideally be captured, and the control loop would fight these random disturbances to hold the device at the command location. The small difference between open- and closed-loop operation shows that the control system is not adding significant additional jitters to the device position. The scale of the open-loop mechanical noise suggests that the open loop noise is caused by the resolution limit of the DAC generating the voltage that drives the current controllers. The DAC is 16-bit, which translates to 7.7 nm steps over the 500 μm range, roughly the noise level observed. This value might increase significantly if worse resolution DACs are used.

The closed-loop noise is less a function of the actuator noise, but rather the sensors. Stage motion occurring on the same scale as the noise in the sensors is indistinguishable from the noise-based random apparent motion, therefore the controller cannot reject the actual motions. The result of this limit is that the closed-loop mechanical noise is at, or slightly less than, the electronic noise. High resolution sensors allow the system to observe that the stage is moving

off of position, and modulate the actuator command over the range of command values to return the stage to the desired location. As long as the condition of a higher electrical bandwidth than mechanical bandwidth is met, then the actuator's effect is smeared out, averaged, by the mechanical low frequency response. This averaging makes it possible to achieve better resolution than the discrete minimum step of the DAC, simply by actively stepping the actuator between the two values at a ratio corresponding to the desired signal. The situation described above is not generally the case, because high resolution, low noise DACs are less expensive and more common than high resolution ADCs. The ADC is usually the lower resolution component, meaning the closed loop noise is slightly higher. Table 5.13 shows this is occurring in the HSFD, as the sensor noise is greater than the actuator noise.

System measurement resolution is found from the electronic noise. The actual location of the HSFD is within ±20.6 nm of any sensor reading, with 95% confidence. The 95% boundary corresponds to ±1.96 standard deviations. Holding resolution is the actual device position (not measured position) compared to the commanded value and is found from the closed-loop noise. The location of the device is within ±17 nm of the commanded locations 95% of the time. The total resolution is the geometric sum of the electronic and closed-loop mechanical noise. This total resolution relates commanded position to the measured (not actual) position, with 95% confidence this separation is within ±26.9 nm. These different resolutions serve different purposes. The electronic resolution is useful for determining how well the device can resolve position changes. The holding resolution defines how stable the stage is, how well it can hold a particular location. The total resolution defines what will be seen as the measurement noise during device operation. Stable control should, in general, hold the closed-loop mechanical noise below the electronic noise. This means that the electronic resolution can be considered a conservative estimate of the holding resolution. A rough estimate of the electronic noise can be obtained from the total resolution of a closed loop system by dividing the total resolution by $\sqrt{2}$, (and vice-versa) following the worst case assumption of electrical and mechanical noise being equal in magnitude.

The predicted electronic noise of about 4 nm is significantly below the observed value of 10.7 nm. The same error in predictions occurred for prototype C measurements, where both errors correspond to a roughly 12 nm unaccounted source. It is believed that this may be due to the insufficient shielding in the sensor electronics.

5.11 Thermal Characterization

5.11.1 Method

The thermal response of the system is measured by driving the system to maximum displacement and measuring the temperatures of the coil, center stage and outer base over time. The maximum displacement command drives the largest current through the actuators and thus produces the greatest heat generation in the system. The thermal profile collected by this measurement corresponds to the worst case thermal scenario for the HSFD.

The HSFD is anchored to a floated metal optical table using three long metal stands and is left exposed to free air currents. This setup mimics the worst case scenario for the device where only convection is available for cooling. The device is driven up to a position of (250 μm, 0 μrad, 0 μrad) in open loop, and held at those coil currents for the length of the test of about 5 hours. The current running through each actuator is calculated using Equation (5.14), and this information can be scaled into power inputs for thermal models, as described previously in Section 3.9.3. Every few minutes during the test, the temperatures of the HSFD components are measured with an OS65 series non-contact IR sensor from Omega Inc. The emissivity of the measured surface is set to roughly 1 by covering the surface with masking tape, which acts as a black-body emitter. The measurements are carried out at the top of the coil, the inner face of the center stage and near the base of one of the flexures on the outer base. An additional ambient temperature reading is made on the surface of the optical table nearby with each measurement. The voltage output from the IR sensor is run through the DAQ board and read in at 1 kHz sampling, with a moving 1000-point average to boost resolution to 0.1 °C.

Experimental results shows that the temperature of the coil, center stage and outer base are dependent on the ambient temperature, as expected from the closed form model described previously in Section 3.9.2. This effect is accounted for by subtracting off the ambient temperature at the time from each from each set of measurements. This reduces noise in the measurements, as the room temperature variations scale both temperatures, leaving the difference unaffected. A first order decay fit function is used to characterize the thermal response of each of the components. The fit function, $\Delta T(t)$, is written in the form:

$$\Delta T(t) = \Delta T_{SS}\left(1 - e^{-\frac{t}{\tau_T}}\right) \tag{5.37}$$

where ΔT_{ss} is the steady state temperature of the component above ambient, and τ_T is the thermal time constant of the component.

5.11.2 Results

The thermal response of the HSFD, measured as described above and driven by actuators that held the stage at maximum displacement is shown in Figure 5.26. The figure contains the measurements made of the component temperatures over the 6 hr run of the experiment along with the fit curves for each component.

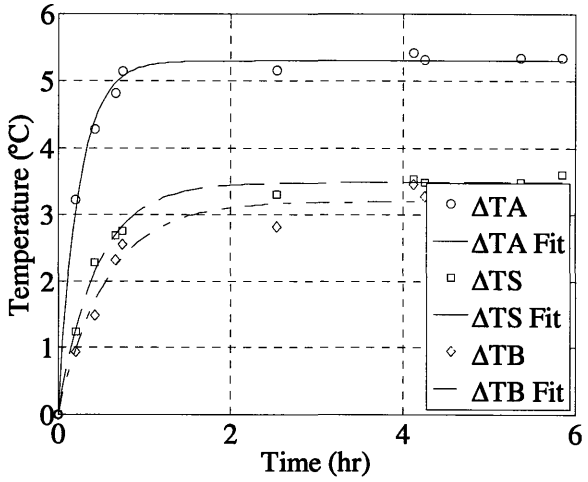

Figure 5.26: Temperature dynamics at maximum power draw.

The data shows a clear first order trend with both decreasing amplitude from coil to center stage to outer base. The time constant is also increasing from coil to outer base in the same order. The agreement between the data and first order curves supports the prediction of the closed form model that each component is strongly linked to the ambient temperature through convection, particularly in the case of the center stage, where the temperature difference to the outer base is minimal. Conduction is thus also minimal. The plastic mandril appears to be largely successful in thermally shielding the device. The values used for the fit curves for each component are shown in Table 5.14.

Table 5.14: Temperature model fit coefficients.

Component	τ_T ±3 (min)	ΔT_{SS} ±0.1 (°C)	Predicted ΔT_{SS} (°C)
Actuator	14.3	5.3	5.4
Center Stage	26.9	3.5	3.4
Outer Base	33.8	3.2	3.3

The fit coefficients are close to the predicted values as described in Table 3.16 previously. The predicted values are derived from a thermal FEA model which is calibrated based on thermal data from prototype C. The strong agreement in the values suggests that the FEA model is capturing the dominant thermal physics with its four adjustable parameters. For this prototype, both FEA model and experiment agree that the system is operating within the thermal constraints and has no significant thermal issues.

CHAPTER 6
CONCLUSION

6.1 Introduction

The intent of this research was to generate the knowledge required to design, fabricate and operate a device capable of high speed nano-scale vertical positioning of microscopy samples. A high speed focusing device (HSFD) was created which used a new combination of technologies for the purpose of imaging: Lorentz coil actuation, flexural bearings and strain gage sensing. The application of these technologies with precision design principles, yielded a device that was capable of achieving a combination of performance and cost gains over competing nano-positioning imaging stages. The HSFD is shown to perform steps with 8 ms 95% settling time, 2% dynamic accuracy, and 0.005% static accuracy while operating with a resolution of 10.5 nm (1σ) over a range of 500 μm at a cost of about $1400. This performance is $3x$ faster stepping, $2x$ better dynamic accuracy, $\approx 100x$ better static accuracy, with equivalent resolution and range to the top of the line commercial devices at less than half of the cost. The reduced cost is envisioned to enable greater distribution and use of nano-positioning imaging stages, while the increased performance is envisioned to enable faster, more benign (in the case of biological sciences) and more precise imaging. The increased use and data gathering ability of the new device is envisioned to enable fields of research such as biology and materials science to extend their bounds further into the micro and nano-scale as well as further along the time scale for both high speed and low speed processes.

6.2 Summary

The work presented in this thesis is a first step towards the development of a fully functional, ready-for-production HSFD. The main results of this work are presented below.

A novel combination of compatible technologies was chosen for the HSFD. Flexural bearings were chosen for guiding the stage position. These bearings have a compact design, fine spatial resolution, high repeatability, and require only a simple fabrication & assembly process. Lorentz actuators were chosen to drive the stage to the desired location over three axes of control. These actuators have high speed, fine resolution, high repeatability, and are low cost. Strain gage sensors were chosen for position sensing of the stage in all three axes. These sensors combine compact design, high bandwidth, minimal drift, fine resolution and low cost. The decision process leading to these choices is described in Chapter 2. The high repeatability and linearity of each of these technologies made it possible to generate reliable predictive models for all aspects of the device.

Models for all of the device components, bearing, actuator, sensor, and electronics were generated and combined into a single general model. The full HSFD model encompasses multiple domains, including mechanical, electrical, thermal, electro-magnetic and electronic. Cross-domain variations were identified and linked between component models. The performance was measured across all domains and optimized through adjustment of the cross-domain variables. Results from intermediate prototypes were used to adjust the models for greater accuracy. The theory underlying these models as well as the chosen device parameters are described in Chapter 3.

The device fabrication and assembly details were described in Chapter 4, with all processes of critical performance laid out in detail. The magnetic assembly of the device was found to require a particular assembly process to ensure the best performance and least likelihood of damage to the components.

The components of the HSFD were characterized and compared to their predicted performance to validate the design models. Calibration was carried out on the actuation and sensing such that the device could be operated as intended. The open-loop dynamics of the device axes were measured and related back to the physical structure of the HSFD. A compensation scheme to enable closed-loop control was developed, and the compensators for

each axis designed to produce the best dynamic performance. The full system operating under closed-loop control was characterized and compared against predictions to ascertain the accuracy of the control models. The methods and results of the characterization for both components and the full system are described in Chapter 5.

This work resulted in the creation of a prototype HSFD whose properties are compared to the initial requirements in Table 6.1. The required and desired values are drawn from Table 1.1 earlier.

Table 6.1: HSFD properties.

Functional Requirement	Measured	Required	Desired	Units
Range	500	100	500	μm
Repeatability (1σ)	10.5	40	1	nm
Resolution (1σ)	10.5	100	10	nm
Bandwidth	125	100	200	Hz
Accuracy	0.005	0.5	0.05	% of travel
Maximum Load	500	500	500	g
Natural Frequency	0.0742	>1	>1	kHz
In-plane Tilt	<1	<10	<10	μrad
Temperature range	5 to 50	5 to 50	5 to 50	°C
Target system cost	1430 (est.)	<2500	<2500	$
Material	Al	Any	Al	---
Power Draw	<2	<30	<10	W
Stage Temp	<3.5	<5	<2	°C (ΔT)
Aperture (x,y)	165*x*110	≥85*x*65	165*x*112	mm
Envelope (x,y,z)	260*x*165*x*23	≤260*x*165*x*26	<260*x*165x26	mm

The HSFD prototype meets or exceeds nearly every requirement described above. The largest underperformance is the center stage natural frequency. The decision to ignore this functional requirement was based on an extreme mismatch in stiffness requirements between actuation power limitations and the natural frequency requirement laid out above, as explained in Section 3.3.9 above. A high natural frequency is required for rapid open-loop response, but if the system is operated in closed loop, the response can be adjusted to be faster than the inherent open loop performance. The range of the device is designed to be ±250 μm, as listed in the HSFD properties above; but at present, prototype D is limited to only +250 μm due to un-diagnosed error in the current controller electronics, as described in Section 5.6.3 above. The cost of the device was estimated from the component costs of the prototype in Appendix B.

6.3 Future Work

6.3.1 Purpose

Several technical issues remain for the HSFD before the design can be finalized into a production ready device. These issues are a combination of problems observed in prototype D, general device limitations which the author lacked the experience to fix, such as controls and electronics, and final packaging design required to ensure the device is not damaged during regular operation.

6.3.2 Mandril

A small amount of warp was observed in the coils used in the prototypes. This warp acts to reduce the gap width on either side of the coil. If the warp is too large, then the coil may make contact with the magnets and cores composing the anchored part of the actuator. This will seriously degrade device performance through stiction at the contact point and the possible wear and shorting of the coil wires. The mandril plastic should be as rigid as possible to avoid this and should not have residual stresses which will cause bending when the mandril is machined. The coil should also be held in a mold of significantly higher rigidity than the mandril itself. For the HSFD prototype, the mold was made of Teflon, which were found to not meet this criterion.

Metal mandrils could be substituted in for the plastic presently used if coil warpage cannot be solved by the methods described above. Plastic inserts placed between the mandril and the coil mounting arms could provide thermal insulation for the rest of the system.

6.3.3 Flexure Clamps

Several issues were observed with the prototype flexures, including the anomalously low bearing stiffness and variation in the COS location. It is believed that these issues are due to the clamping problems at the base of the flexures. A small amount of epoxy binder could be placed between the flexure sheet and the clamping surfaces to increase the effectiveness of the flexure anchoring method. If the binder is placed over all of the surfaces in contact with the flexure sheet, then the result should be the enforcement of the no-slip contact condition between the separate pieces. The epoxy will ensure excellent surface contact between the clamping surfaces and the flexure sheet. The thinner the layer of epoxy, the stiffer it is and the closer the interface

is to perfectly rigid. The adhered surfaces will distribute the clamping pressure more evenly over the flexure ends and better resist flexing under the high loads at the ends of the flexure. The device will be more difficult to disassemble for repair, therefore a thermal based epoxy may be used. This should enable the components to be separated once the center stage is adequately heated.

The issue of clamping error as described in Section 5.4.3 earlier can be minimized through several methods. The surfaces defining the edge of the clamp on the center stage should be milled to ensure that they are at the same location on the top and bottom of the piece. This would be easily accomplished if the center stage were precut to slightly larger dimensions through water jet or laser cutting, then all of the surfaces finished through a final single milling operation. Reducing the screw hole size to have less clearance should ensure the location of clamps with better accuracy. Finally, the center stage clamp (the likely cause of the error in x_{COS}) can be located by pushing its edge flush against a reference plane which is also in contact with the side of the center stage. This should ensure both surfaces are flush within roughly 25 μm. If it is difficult to get high accuracy when cutting the inner flexure clamp, it may be beneficial to separate it into two pieces, such that flexures on either side of the Y-axis each have a separate clamp that can be aligned to their particular clamping edges using the reference plane method. Large clearance in the screw holes would be possible if the reference plane method is used.

Larger screws should be used for clamping the flexures, somewhere in the range of 6-32. The present prototype does not use these due to the thin center stage. It is recommended in Section 6.3.4 that the stage be thickened slightly or sample holding flanges extending yet further inwards be included onto the center stage to increase rigidity. This will require an analysis of the aperture size to determine how to best add structural support to the center stage without compromising the ability to hold samples, such as 96 well plates. The screw hole sizes at the corners can safely be increased for the larger screws if this is done. The higher clamping force generated by these larger screws should also aid the flexure clamps.

6.3.4 Center Stage Rigidity

It is believed that at frequencies above 400 Hz the center stage cannot be treated as a rigid structure and its associated dynamics are causing the range of multi-axis resonances

described in Section 5.8. The stage compliance appears to also be reducing the rotational stiffness of the bearing.

The rigidity of the stage can be increased through several means, any of which should both raise the bearing stiffnesses and push the multi-axis resonances to higher frequencies. The center stage cannot be easily extended outwards to increase thickness, instead the gain must come from a reduction in the aperture. A detailed analysis must be carried out on the sample holders that are to be placed in the aperture. This will determine what space is available to add in reinforcement. It is expected that filleting the interior corners of the center stage should significantly help the rotational stiffnesses, and should not interfere with sample holders. Another possible method is to include a flange in the center stage that extends inwards. This extra geometry will serve two functions, to seat samples and even if extending only about 2.54 mm (0.1 in.) would have a large impact on the rigidity of the stage. The extension of the flange inwards would significantly increase the second moment of area of the center stage sides undergoing in-plane bending. It is recommended that the flange be added by increasing the width of the bottom inner flexure clamp, such that it extends inwards to cover some of the aperture, as shown in Figure 6.1.

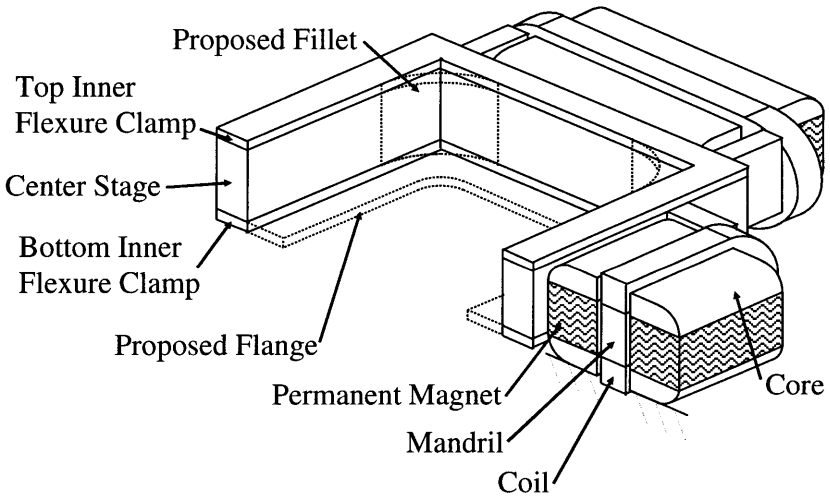

Figure 6.1: Proposed improvements in center stage rigidity.

The four screw mounting points of the center stage to the flexure clamp, as well as the proposed epoxy bonding over the top surface of the center stage will ensure that the two pieces- center stage and flexure clamp- may be modeled as a single rigid component. The benefit of the

flange would scale roughly with the cube of its width. The geometric changes would be simple to include into the present design, as neither requires adding additional fabrication steps to the device- the planar geometry of both components are maintained.

6.3.5 Mechanical Shielding

Shielding plates should be placed over the surface of the device to protect the strain gage wiring from EM noise and mechanical damage. The flexures should also be covered over to prevent damage. The steel cores do not need to be covered, but doing this would ensure they could not be hit and moved. The shielding plate does not need to be very thick, as it can be supported over nearly the entire surface of the outer base. The vertical gaps on either side of the coil should not be covered, as the high convection coefficient over the coil surface depends on the driven air flow as shown in Figure 3.34 previously. Small gaps could be cut in the mechanical shield to allow the air through these gaps.

6.3.6 HSFD Wiring

The hand soldered wiring for the strain gage bridges is infeasible for general production of the HSFD. It also has a large geometric profile, which would be difficult to safely cover for a final product. A flex-circuit could be used to replace these wires. The circuit would have a section which would bend around under the flexure to make electrical contact with the bottom flexures. It would need to be cut to be flexible, thereby avoiding interference with the bearing performance. Bending the flex circuit slightly to stand off from the flexure surface would significantly aid in increasing its compliance The strain gages could either be incorporated into the flex circuit, or bond pads on the flex circuit could be used to make connection with strain gages on the flexure. Additionally, the flex circuit could either have the STC resistance wire integrated into it, or have bond pads which would attach to the STC resistor. The circuit could then be routed over the surface of the outer base to the electrical output port for the device. This would vastly simplify the wiring design, as well as the fabrication and assembly for the electronics on the device.

6.3.7 Noise Reduction

Any noise reduction in the system will result in better electronic resolution as a general rule. This will improve the system performance without requiring any significant sacrifices in other aspects of the device. The system electronics should be thoroughly shielded in a faraday cage container. The actuator power circuitry generates significant power draw dynamics and large time varying signals. Every effort should be made to keep the power channels as far from the sensor channels as possible. A third separate power channel should be used for driving the bridges, as the sensor Wheatstone bridges do not vary in power draw and therefore will not introduce unwanted signals to one another. The power supply channels for each of these electronics: IAs, WBs, and CCs should be kept separate if possible. This would likely mean that each have their own voltage regulator. The three voltage regulators can be powered by a single transformer or voltage regulator that transforms the general socket voltage down to the voltage required to drive the regulators. Precision voltage regulator chips, such as the μA723 made by Analog Devices, are inexpensive ($5) and will attenuate variations in the power supply voltage on the order of 10-100x, better insulating the sensor electronics from both actuator and wall power noise.

The wires linking the HSFD internal wiring to the electronics box should be shielded, twisted, and grounded if at all possible. The sensor return wiring should be shielded separately from the actuator power wiring to reduce cross capacitance. It is strongly recommended that the command signals sent to the electronics box be transmitted in digital form rather than analog. The resolution of the system is presently obtained by straining the system to the limit of the electronics; in particular, the resolution of the system is presently dominated by the noise generated by the ADC reading in the sensor signals. Adding another DAC to output the computer's command and an ADC to read the command into the electronics box will impair both the total system resolution and the holding resolution. If digital command inputs and digital position outputs are used between the electronics box and controlling computer then no extra noise will be added to the system, and single nanometer resolution is possible. It is believed that a USB connection could handle the task of transferring the 3-vector command to the device and reading the 3-vector position back out in real-time.

The 16-bit ADCs reading the sensor outputs are presently the dominant noise source. If the electronics is properly shielded and the present 16-bit ADCs are replaced with 18- or 20-bit

ADCs, the resolution model predicts the resolution should be reduced by 2-4*x*. This is recommended as the cost difference between 16-bit ($10) and 18-bit ($20) or even 20-bit ($30) is not significant when compared to the costs of the other components in the system, as shown in Appendix B.

6.3.8 Bridge Energizing Voltage

The high bridge energizing voltage presently used in prototype D is not effectively boosting the resolution due to the fact the ADCs are the dominant noise source. The bridge energizing voltage should be reduced to 10 V to reduce the thermal load on the system if 16-bit ADCs are used. A bridge energizing voltage of about 15 volts would be a good compromise between power generation and noise if proper shielding and 18- or 20- bit ADCs are used. The higher resolution electronics should drop the ADC noise to roughly 1 nm, 25% of its present value of about 4 nm, while the bridge energizing voltage would have reduced the bridge and IA noise to roughly the same 1 nm, based on Figure 3.30 above. This will balance all the noise sources, resulting in the best resolution for the system. The resolution model should be used to study the benefit of increasing the energizing voltage once the ADC choice is made and the noise in the properly shielded electronics is measured. If the voltage is reduced to 10-15 V, this should lower the center stage temperature by roughly 30-50%. A precision voltage regulator such as the μA723 should be sufficiently accurate to remove the need for direct measurement of the bridge energizing voltage. The removal of the power supply measurement would result in an immediate halving of the system resolution.

6.3.9 Current Controller

The actuator power electronics should be fixed to avoid generating the observed non-linear region shown in Figure 5.9 above. This could be done through replacing the op-amps in the circuit with higher quality components, or redesigning the bridge. It is expected that the problem is not inherent to the current controller design, as it is a common circuit, but rather to the type of components used or to damage in the circuits through laboratory miss-use.

6.3.10 Sensor Voltage Offset

The SG WB sensors show a significant voltage offset due to mismatched strain gage resistors in the bridge. This offset is large enough to push the sensor voltage range out of the measurable range of the ADC. The sensor offset also enables thermal variations to affect the sensor reading, as described in Section 3.5.7. The present voltage offset in the sensors restricts the operating range of the HSFD prototype to only a fraction of the designed operating range, as described in Section 5.7.2. A full 10 V offset only requires about a 12 Ω difference in resistance on one of the arms.

The voltage offset can be minimized or removed through three different methods. First, a zero balance resistance wire can be used to link the strain gages to form the bridge. The resistance wire would have to be cut and placed uniquely for each bridge, and it would have to have high flexibility, high resistivity, and a low resistance temperature coefficient. Second, a potentiometer could be placed on the outer base near each flexure, and used to manually zero-balance the bridge during fabrication. The potentiometer can be placed between resistors 1 and 2 as shown in Figure 6.2.

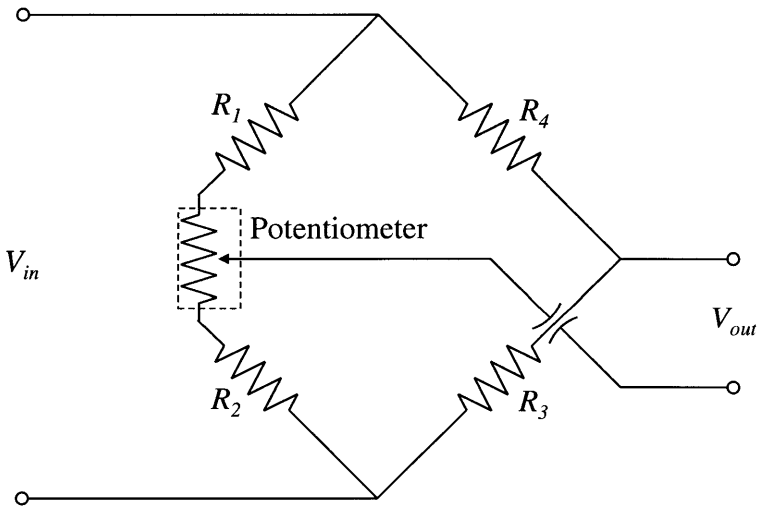

Figure 6.2: Proposed improvements to sensor bridge.

The wiring linking the potentiometer into the bridge would be relatively simple to add if flex circuitry is used for the HSFD. The closer the resistance temperature coefficient of the potentiometer is matched to the strain gage, the greater the attenuation in the thermal sensitivity of the bridge. Third, the voltage output of the sensor IA can be biased to remove the offset

before it is fed to the ADC. This can be implemented by sending a bias voltage using a unity gain op-amp (for the low impedance) to the reference terminal of the AD624, as explained in the IA product description [44]. The voltage for each channel would need to be tuned after each sensor is fabricated. This will remove the offset, but not the thermal sensitivity of the offset.

6.3.11 Strain Gage Sensor Error

The error in the SG WB sensors has two sources. The first is the error in ability of the sensors to read the correct displacement and the second is the motion of the sample that is not captured by the sensors. The sensor errors are captured through the calibration process, while the un-sensed motions are generally much more difficult to quantify.

The calibration error for the HSFD is dominated by the first order decay that is observed in the CAP sensors during calibration, as described in Section 5.7.2. The SG WB sensors did not share this signal, thus it is difficult to determine if the first order decay corresponds to actual motion. The time constant of the system is too slow to be caused by the CAP bandwidth and too fast to be caused by a thermal response. It is believed that the decay is instead a function of the ground plane (which includes the optical table top) used as part of an electrical circuit by the current controller. The steady state value output by the CAP probes is likely the true displacement value. This means that the SG calibration should be carried out once the first order decay is settled, which can be done by waiting about 15 seconds after the position command is given. The accuracy of the calibration should be improved by this change.

Thermal errors between the measured position and the actual sample position can be minimized by reducing the thermally active section of the ground path linking the sample to the microscope stage. An accurate measurement of the sample location must link the sample to the sensors and the sensors to the absolute ground. The path defined by this linking is called the 'ground path' and can be traced through Figure 6.3 by following the mechanical linkage from the ground plane through the outer base to the present sensor plane, then back down the center stage to the sample plane. Ideally, the only motion that occurs over the ground path is on the flexures, which is captured by the sensors. The present sensor plane is distant from both the ground plane and the sample plane with the result that thermally induced expansion occurring over either the height of the outer base or the height of the center stage will cause an error in the measurement

of the sample location. The effect would cancel out if these errors were of the same magnitude; however neither the lengths nor the expected temperatures match.

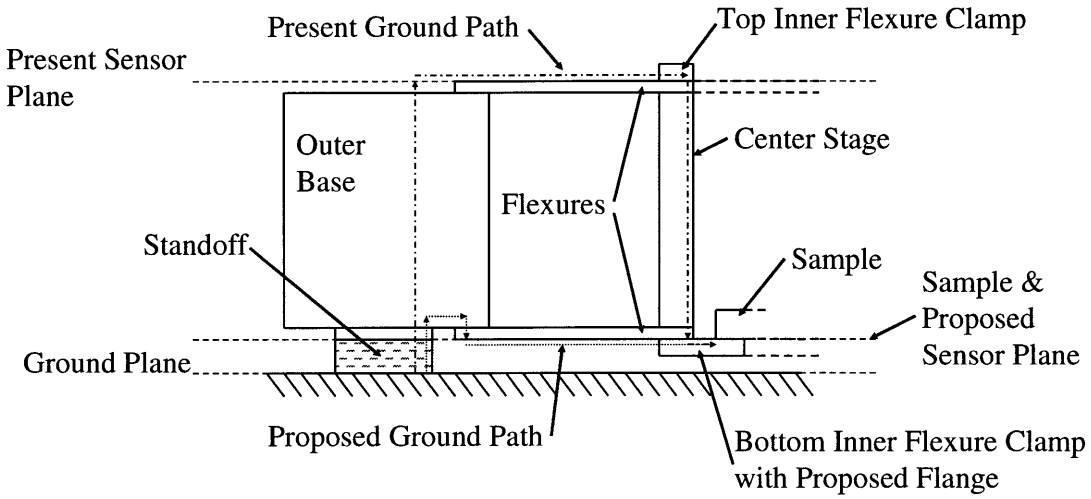

Figure 6.3: Proposed improvements to sensing accuracy.

The simple solution to this ground path thermal sensitivity it to reduce the length of the thermally sensitive parts of the path. This can be accomplished by shifting the sensor plane to the bottom of the device, that is by measuring the bottom flexure sheet. The distance from the sample to the sensor plane is thus reduced from the thickness of the center stage to zero because the sensors are now on the exact same plane as the sample. The distance from the sensor plane to the ground is also reduced to no more than the thickness of the standoffs holding the outer base off from the ground. Thermally insensitive materials such as Invar or Zerodur can be used to make the standoff, with the result that the ground path passing through these stands has virtually no thermal growth. The standoff in this case would have to be anchored to the outer base at the sensor plane and would have to pass through the mechanical shielding to make direct contact with the ground. An aluminum shim of the same thickness as the flexure sheet should be placed between the stand and the outer base to bring the end of the stand onto the same plane as the strain gages.

This proposed ground path will generate at worst 20 nm of thermal error for Invar or 3 nm for Zerodur, again assuming 6.35 mm (0.25 in.) stands and 3 °C stage temperature rise. The pseudo-kinematic contact of the stands will also ensure that the device does not rock on the microscope stage, which would cause position errors. It is recommended that stands be used and

made of Invar, owing to its easier machinability. Further reduction in the thermal error can be found through adjusting the thickness of the aluminum shim under each stand. If the aluminum shim is made thinner, the ground path through the aluminum outer base and flexure sheet is not completely negated, therefore the path is made sensitive to thermal errors again. This can be seen in Figure 6.3 above. The slight upwards growth of the stands can be counteracted by the downwards expansion of the flexure sheet thickness by properly designing the shim thickness. A rough estimate of the system using an Invar stand height of roughly 6.35 mm (0.25 in.) suggests that a shim thickness of 0.127 mm (0.005 in.) should remove the thermal sensitivity of the proposed ground path if both shim and stand are heated equally. Calibration should be done on the bottom surface of the center stage to measure as close to the sensor plane as possible. This will further reduce any thermal error in the calibration.

Further error reduction, if desired, can be gained through thermal control of the electronics box. This can be directly implemented with a thermoelectric cooler driven by an ADN8831 chip from Analog Devices ($10). The Peltier thermoelectric cooling device ($15) must be thermally grounded to the board on which the electronics are attached and also to a heat sink, which is preferably exposed to free air outside of the electronics box. Full thermal control of the electronics will prevent errors in the measurement which can grow up to 0.03% (assuming 10 °C possible rise during operation), mainly through the thermal sensitivity of the IA and the voltage regulator driving the Wheatstone bridge.

6.3.12 Control Hardware

The control loop should be run on a dedicated microcontroller or FPGA. The sampling rate at which the control loop is run precludes any ability to have the computer running the control loop. It is estimated that a single processor should be capable of running the three loops in parallel, and should cost somewhere in the range of $100. An alternate scheme is to implement the control loop in hardware [52], which is feasible given that the compensation design remains simple. The noise added by the hardware implementation would likely be large compared to the present device noise level.

6.3.13 Control Loop

The proper design of the control loop should be able to roughly halve the system settling time, especially if it is carried out for all three axes. A state-space observer-based controller should be capable of handling the multiple input, multiple output nature of the device. The present control loop design of parallel independent controllers has some difficulty with controlling the higher frequency center stage modes, which are presently the limiting factor in driving the device to settling times of 5 ms or less. Improving the compensation schemes for these axes should reduce the cross axis disturbances, enabling a significant decrease in the stage settling time. An increase in the center stage rigidity will shift the center stage modes further up in frequency, which would simplify the compensation scheme. The parallel independent controller scheme may then be capable of 5 ms settling times or better, without requiring the use of state space control.

REFERENCES

[1] G. D. Rondeau, Applied Scientific Instrumentation Inc., "Stage Assembly and Method For Optical Microscope Including Z-axis Stage and Piezoelectric Actuator For Rectilinear Translation of Z Stage," U.S. Patent No. 7,180,662 B2, Feb. 2007.

[2] D. S. George, Image Analysing Computers Ltd., "Precision Microscope Stage," U.S. Patent No. 3,652,146, Mar. 1972.

[3] J. P. Stewart, Applied Precision, Inc., "Orthogonal High Accuracy Microscope Stage," U.S. Patent No. 5,812,310, Sep. 1998.

[4] H. J. B. Marsman, R. Stricker, R. W. Wijnaendts van Resandt, G. J. Brakenhoff, and P. Blom, "Mechanical Scan System for Microscopic Applications," *Review of Scientific Instruments*, vol. 54, pp. 1047-1052, 1983.

[5] V. Bansal, S. Patel, and P Saggau, "A High-Speed Confocal Laser-Scanning Microscope Based on Acousto-Optic Deflectors and a Digital Micromirror Device," *Proceedings of the 25th Annual International conference of the IEEE EMBS*, Cancun, Mexico, Sept 17-21, 2003.

[6] S. A. Jewell and J. R. Sambles, "Enhanced Confocal Microscopy Imaging of the In-Plane Switching of Cholesteric Liquid Crystal Cells," *Proceedings of SPIE*, vol. 7050, pp. 70500G, 2008.

[7] H. Liu, G. D. Bachand, H. Kim, C. C. Hayden, E. A. Abate, and D. Y. Sasaki, "Lipid Nanotube Formation from Streptavidin-Membrane Binding," *Langmuir*, vol. 24, pp. 3686-3689, 2008.

[8] L. L. Kang, Y. X. Huang, W. J. Liu, X. J. Zheng, Z. J. Wu, and M. Luo, "Confocal Raman Microscopy on Single Living Young and Old Erythrocytes," *Biopolymers*, vol. 89, pp. 951-959, 2008.

[9] H. Fang, L. Qiu, E. Vitkin, M. M. Zaman, C. Andersson, S. Salahuddin, L. M. Kimerer, P. B. Cipolloni, M. D. Modell, B. S. Turner, S. E. Keates, I. Bigio, S. D. Freedman, R. Bansil, E. B. Hanlon, and L. T. Perelman, "Confocal Light Absorption and Scattering Scpectroscopic Microscopy," *Applied Optics*, vol. 46, pp. 1760-1769, 2007.

[10] J. Hubbuch, M. R. Kula, "Confocal Laser Scanning Microscopy as an Analytical Tool in Chromatographic Research," *Bioprocess and Biosystems Engineering*, vol. 31, pp. 241-259, 2008.

[11] M. Minsky, "Microscopy Apparatus," U.S. Patent No. 3,013,467, Dec. 1961.

[12] D. Yan, K. Li, "Measurement of Wet Fiber Flexibility by Confocal Laser Scanning Microscopy," *Journal of Materials Science*, vol. 43, pp. 2869-2878, 2008.

[13] A. V. Failla, S. Jäger, T. Züchner, M. Steiner, and A. J. Meixner, "Topology Measurements of Metal Nanoparticles with 1 nm Accuracy by Confocal Interference Scattering Microscopy," *Optics Express*, vol. 15, pp. 8532-8542, 2007.

[14] M. Petráň and M. Hadravský, "Tandem-Scanning Reflected-Light Microscope," *Journal of the Optical Society of America*, vol. 58, pp 661-664, 1968.

[15] Image reprinted with permission from the Optical Society of America: M. Petráň and M. Hadravský, "Tandem-Scanning Reflected-Light Microscope," *Journal of the Optical Society of America*, vol. 58, pp 661-664, 1968. Copyright 1968 by the OSA.

[16] E. M. Slayter and H. S. Slayter, "Light and Electron Microscopy," New York NY: Cambridge University Press, 1994.
[17] Y. Shao, D. L. Dickensheets, "MOEMs 3D Scanning Mirror for Single-Point Control of Beam Deflection and Focus," *Journal of Microlithography, Microfabrication, and Microsystems*, vol. 4, pp. 041502, 2005.
[18] Olympus Corporation, "LEXT OLS3100," accessed Feb. 2009, <http://www.olympusmicroimaging.com>.
[19] A. Diaspro et al., "Confocal and Two-photon Microscopy: Foundations, Applications, and Advances," New York, NY: Wiley-Liss, Inc., 2002.
[20] Image reprinted with permission from the Society of Photo-Optical Instrumentation Engineers: S. A. Jewell and J. R. Sambles, "Enhanced Confocal Microscopy Imaging of the In-Plane Switching of Cholesteric Liquid Crystal Cells," *Proceedings of SPIE*, vol. 7050, pp. 70500G, 2008. Copyright 2008 by the SPIE.
[21] Image reprinted with permission from the American Chemical Society: H. Liu, G. D. Bachand, H. Kim, C. C. Hayden, E. A. Abate, and D. Y. Sasaki, "Lipid Nanotube Formation from Streptavidin-Membrane Binding," *Langmuir*, vol. 24, pp. 3686-3689, 2008. Copyright 2008 by the ACS.
[22] Y. Saga, Y. Shibata, S. S. Itoh, H. Tamiaki, "Direct Counting of Submicrometer-Sized Photosynthetic Apparatus Dispersed in Medium at Cryogenic Temperature by Confocal Laser Fluorescence Microscopy: Estimation of the Number of Bacteriochlorophyll c in Single Light-Harvesting Antenna Complexes Chlorosomes of Green Photosynthetic Bacteria," *The Journal of Physical Chemistry B*, vol. 111, pp. 12605-12609, 2007.
[23] Image reprinted with permission from the American Chemical Society: Y. Saga, Y. Shibata, S. S. Itoh, H. Tamiaki, "Direct Counting of Submicrometer-Sized Photosynthetic Apparatus Dispersed in Medium at Cryogenic Temperature by Confocal Laser Fluorescence Microscopy: Estimation of the Number of Bacteriochlorophyll c in Single Light-Harvesting Antenna Complexes Chlorosomes of Green Photosynthetic Bacteria," *The Journal of Physical Chemistry B*, vol. 111, pp. 12605-12609, 2007. Copyright 2007 by the ACS.
[24] Image reprinted with permission from Springer: D. Yan, K. Li, "Measurement of Wet Fiber Flexibility by Confocal Laser Scanning Microscopy," *Journal of Materials Science*, vol. 43, pp. 2869-2878, 2008. Copyright 2008 by Springer.
[25] D. S. Moore, Vickers PLC, "Microscope Stage," U.S. Patent No. 4,688,908, Aug. 1987.
[26] G. G. Gerhard, Applied Precision, Inc., "Orthogonal Motion Microscope Stage," U.S. Patent No. 5,684,628, Nov. 1997.
[27] A. Lewis and K. Lieberman, "Flat Scanning Stage for Scanned Probe Microscopy," U.S. Patent No. 5,705,878, Jan. 1998.
[28] P. E. West, TopoMetrix Corp., "High Resolution Electromechanical Translation Device," U.S. Patent No. 5,260,622, Nov. 1993.
[29] Image reprinted with permission from the American Institute of Physics: H. J. B. Marsman, R. Stricker, R. W. Wijnaendts van Resandt, G. J. Brakenhoff, and P. Blom, "Mechanical Scan System for Microscopic Applications," *Review of Scientific Instruments*, vol. 54, pp. 1047-1052, 1983. Copyright 1983 by the AIP.
[30] I. Ilev, R. Waynant, I. Gannot, and A. Gandjbakhche, "Simple Fiber-Optic Confocal Microscopy with Nanoscale Depth Resolution Beyond the Diffraction Barrier," *Review of Scientific Instruments*, vol. 78, pp. 093703, 2007.

[31] Physik Instruments, "P541.Z Piezo Z stage," accessed Feb. 2009, <www.physikinstruments.com>.
[32] A. H. Slocum, "Precision Machine Design," Eaglewood Cliffs, NJ: Prentice-Hall, Inc., 1992.
[33] J. Hopkins, "Design of Parallel Flexure Systems via Freedom and Constraint Topologies (FACT)," MIT Masters Thesis, Cambridge, MA, 2007.
[34] Matweb, "Materials Property Data," accessed Sep. 2007, <http://www.matweb.com>
[35] R. J. Roark, W. C. Young, and R. G. Budynas, "Roark's Formulas for Stress & Strain," 7th ed., New York, NY: McGraw Hill Companies Inc., 2002.
[36] E. Oberg, F. D. Jones, H. L. Horton, and H. H. Ryffel, "Machinery's Handbook," 25th ed., New York, NY: Industrial Press, Inc., 1996.
[37] A. E. Fitzgerald, C. Kingsley Jr., S. D. Umans, "Electric Machinery," 6th ed., Singapore: McGraw-Hill Companies Inc., 2003.
[38] H. A. Haus, J. R. Melcher, "Electromagnetic Fields and Energy," Eaglewood Cliffs, NJ: Prentice Hall, 1989.
[39] Omega Engineering, Inc., "The Pressure Strain and Force Handbook," 8th ed., Stamford, CT: Omega Engineering Inc., 2006.
[40] S. H. Crandall, N. C. Dahl, T. J. Lardner, "An Introduction to the Mechanics of Solids," 2nd ed., New York, NY: McGraw-Hill Companies, Inc., 1999.
[41] D. S. Golda, "Design of High-Speed, Meso-Scale Nanopositioners Driven by Electromagnetic Actuators," MIT Doctoral Thesis, Cambridge, MA, 2008.
[42] Dspace Inc., "DS1104 R&D Controller Board," accessed Feb. 2009, <http://www.dspaceinc.com>.
[43] S. D. Senturia, "Microsystem Design," Boston, MA: Kluwer Academic Publishers, 2001.
[44] Analog Devices Inc., "AD624 Precision Instrumentation Amplifier," accessed Feb. 2009, <http://analog.com>.
[45] Indigo Instruments, "Rare Earth Magnets," accessed Dec. 2007 <http://www.indigo.com>
[46] Analog Devices Inc., "AD8221 Precision Instrumentation Amplifier," accessed Feb. 2009, <http://analog.com>.
[47] E. Cravalho, J. L. Smith Jr., J. Brisson II, G. McKinley, "2.006 Thermal-Fluids Engineering," unpublished class notes for MIT class 2.006, Cambridge, MA, 2006.
[48] E. Cravalho, J. L. Smith Jr., J. Brisson II, G. McKinley, "Thermal-Fluid Property Data," unpublished class reference for MIT class 2.006, Cambridge, MA, 2006.
[49] G. Strang, "Introduction to Linear Algebra," 3rd ed., Wellesley, MA: Wellesley – Cambridge Press, 2003.
[50] K. A. Lilienkamp, "A Simulink Driven Dynamic Signal Analyzer," MIT Undergraduate Thesis, Cambridge, MA, 1999.
[51] G. F. Franklin, J. D. Powell, and A. Emami-Naeini, "Feedback Control of Dynamic Systems," 5th ed., Upper Saddle River, NJ: Pearson Prentice Hall, 2006.
[52] N. S. Nyse, "Control Systems Engineering," 5th ed., Hoboken, NJ: Wiley, 2008.
[53] A. Preumont, "Vibration Control of Active Structures," 2nd ed., Dordrecht, The Netherlands: Kluwer Academic Publishers, 2002.

APPENDIX

A

ACTUATOR OPTIMIZATION

A.1 Motivation

The actuator used in prototype D is based on the single flux path model and was initially designed for prototype C. Data generated from analysis of prototype C led to the generation of the multi-flux path model, from which a better actuator could be designed. The actuator was not rebuilt when prototype D was made due to time constraints. This appendix describes the analytical design of a optimized actuator which should outperform the present actuator in both heat generation (about 20% reduction) and linearity (out to 250 μm, twice as far).

A.2 Force Non-Linearity

The actuator in prototypes C and D was observed to output a non-linear force over is range in steady state. Simple models of the system suggest that the linear stiffness and the constant magnetic flux should enforce a perfect linearity between current through the actuator coil and steady state displacement, as described by Equation (3.52) previously. Static measurements of the bearing confirm that the stiffness variation described in Figure 5.4 previously is insufficient to explain the non-linear actuation of the stage as shown in Figure 5.8 previously.

It is believed that this non-linearity is caused dominantly by a reduction in the average gap flux density observed by the coil as it moves away from the center of the gap. The reduction in the effective flux density value for the coil drags the force output of the coil down, resulting in un-intended center stage motions. The force output of each actuator is not reduced by the exact same value, resulting in the stage slowly rotating at large displacements. This is exactly what is observed in the calibration error shown in Figure 5.8 previously. The gap flux density variation

also explains the asymptotic increase in the un-intended motions beyond a certain range of coil displacement, as the field begins to fall off to zero. The linear model relies on the assumption that the coil is always completely occupying the region of full flux density, shown as states 'a' and 'b' in Figure A.1. Extra turns of the coil extend it above and below this region of constant flux such that the coil motions have no effect on the number of turns in contact with the magnetic flux. The end of the coil passes into the region of significant force producing flux at high displacements, shown in the figure as state 'c', thus reducing the full integral of flux density and current described in Equation (3.22) previously.

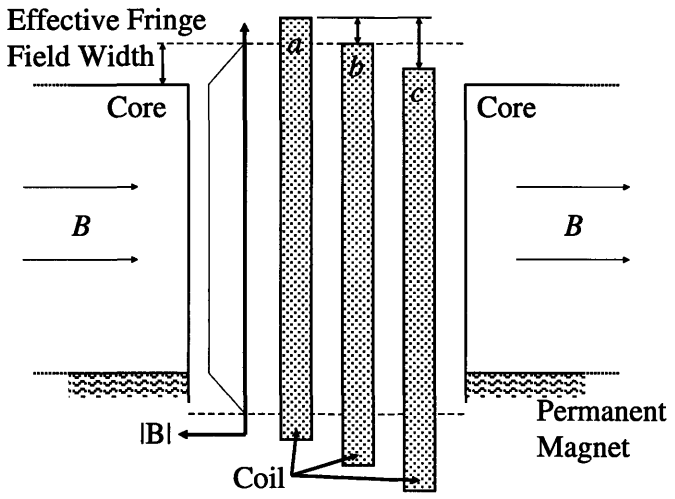

Figure A.1: Schematic of actuator non-linearity.

State 'a' illustrates the system when the assumption of constant flux density may be made. State 'b' illustrates the threshold where the assumption of constant flux density fails. State 'c' illustrates the system when the effective flux density observed by the coil is a function of position. The threshold state 'b' corresponds to a center stage displacement of roughly 150 μm, as the actuator calibration error shown in Figure 5.8 previously is nearly steady up to this displacement (state 'a'), then begins to grow significantly afterwards (state 'c'). The effective edge of the field can thus be defined as being at the height of the bottom surface of the coil at 150 μm upwards displacement. This surface begins 250 μm below the actuator, sinks down 40 μm due to center stage weight, and is driven up 150 μm by the actuators to reach the edge of the magnetic flux at roughly 140 μm below the surface of the core. At a maximum loading of 0.5 kg, the stage sinks 90 μm, bringing the top of the coil closer to the flux edge. The coil must start

at about 500 μm above the surface of the core in order to remain in state 'a' throughout the entire ±250 μm operating range. The increase in the coil height is an extra 250 μm on either side of the coil gap for both top and bottom gap resulting in a coil that is an extra 0.5mm taller, but should show a significantly more linear response over its full range.

A.3 Actuator Optimization

The actuator design is altered in two ways: increased width to boost linearity and geometry adjustment to boost efficiency using the multi-flux path model. Linearity is designed in though the coil model Equation (3.47), which is modified to have an extra term of 250 μm on the right hand side. The geometry adjustment occurs in several steps. The various dimensions of the actuator are each examined for their effect on the actuator performance according to the multi-flux path model, and these dimensions are adjusted to maximize the performance within the constraint range. The anchoring arms emit wasted flux both across the gap and short circuiting the coil. Their length is reduced by setting the length of the core to 104 mm (4.100 in.). The shims, placed between the core and magnet to ensure the magnet is clamped as described in Section 3.8 above, are removed in favor of simply increasing the height of the magnet to 1.37 mm (0.540 in.). This ensures 500 μm (0.020 in.) of travel on either side of the center stage, reduces the number of required pieces (no shim), simplifies fabrication by removing the surfacing operation on the coil mounting arms, and boosts the actuator efficiency. The width of the gap is increased to 5.59 mm (0.220 in.), which is the maximum width for the coil at furthest extension to travel no more than 6.35 mm (0.250 in.) beyond the surface of the outer base. This cutoff value is based on the operating envelope of the device as laid out in Table 1.1 previously.

The gap height, h_{gap}, is the only variable which does not have a performance maximum at a constraint boundary. The maximum is instead found through searching over the range of heights from 1 to 5 mm, as shown in Figure A.2. The tradeoff represented by this curve is between the space for extra turns and the gap flux density. The net heat generation of all of the coils is graphed as the variable to minimize. The minimum heat generation is found at 2.67 mm (0.105 in.). For design purposes, any value from 2 to 4 mm should achieve a significant power reduction over the present version.

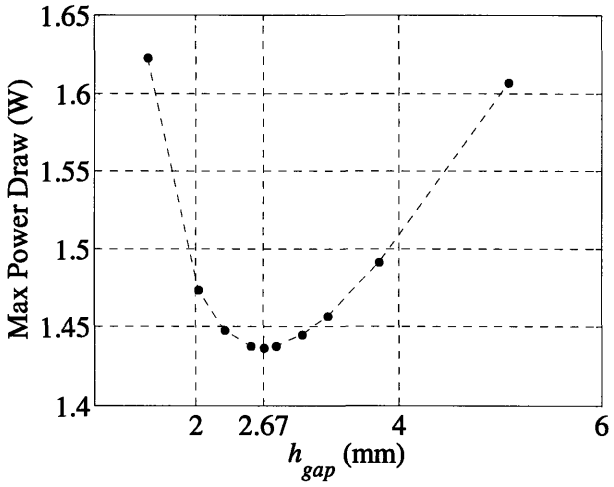

Figure A.2: Lorentz actuator efficiency optimization through gap adjustment.

The resulting optimized actuator geometry and properties are listed in Table A.1.

Table A.1: Lorentz actuator optimization design results.

Property	Value		Units
	Channel 1	Channel 2&3	
l_{mag}	76.2 (3.00)	76.2 (3.00)	mm (in.)
h_{mag}	13.7 (0.540)	13.7 (0.540)	mm (in.)
w_{mag}	12.7 (0.500)	12.7 (0.500)	mm (in.)
h_{gap}	2.67 (0.105)	2.67 (0.105)	mm (in.)
w_{gap}	5.59 (0.220)	5.59 (0.220)	mm (in.)
B_{gap}	1.16	1.16	T
h_{coil}	2.16 (0.085)	2.16 (0.085)	mm (in.)
w_{coil}	6.58	6.58	mm
l_{core}	104 (4.10)	104 (4.10)	mm (in.)
$d_{spacing}$	0.254 (0.010)	0.254 (0.010)	mm (in.)
d_{wire} (26 AWG)	0.406	0.406	mm
Magnetic Material	N52	N52	---
$l_{wirefull}$	11.05	11.05	m
$l_{wireact}$	7.42	7.42	m
$n_{coilfull}$	57	57	turns
R_{coil}	1.79	1.79	Ω
L_{coil}	28.4	28.4	μH
I_{max}	0.217	0.615	A
V_{max}	0.387	1.1	V
P_{max}	0.084	0.676	W

APPENDIX B

COST ESTIMATE

This appendix presents the components resulting in the cost estimate quoted in the conclusion chapter. The cost is broken into three main sections: Electronics, Hardware, and Assembly.

The electronics cost is further decomposed into the expected cost of the current controller, the instrumentation amplifiers, the Analog-to-Digital converters to read in the sensor outputs, the Digital-to-Analog converters to output commands to the actuators, the voltage regulator circuit which drive the Wheatstone bridge, the strain gages on the flexure and the microcontroller/FPGA used to command the device. The current controller circuitry is described in Section 3.6.8, and the necessary components are listed in Section 3.7.4. It is assumed that the AD8221 instrumentation amplifier is used due its lower cost and the dominant noise source lies elsewhere in the system. It is also assumed that the ADC and DAC are 16 bits, and these are common, off-the-shelf components, as is the voltage regulator. The strain gages can be obtained from Omega Engineering, and the microcontroller is also an off-the-shelf component.

The hardware cost is further decomposed into the expected cost of the coils, magnets, steel and aluminum. The coil cost is likely dominated by the cost of having them fabricated. The magnet cost is predicted from estimates provided by Indigo Instruments, the supplier of the magnets used in the prototype. The steel and aluminum costs are largely dominated by the 12.7 mm (0.500 in) aluminum block used as the outer base.

The assembly cost is conservatively estimated to be about 10 hours based on prototype assembly experience, at $50/hr. This depends on whether the strain gages are pre-attached to the flexures, on a flex circuit or must be glued on during assembly. Additionally, the level of electronics integration affects this value. The total cost is conservatively estimated at $1430.

Table B.1: Cost estimates for HSFD.

Property	Number		Cost ($)
Electronics	---		580
Current Controller	3	90	
Ins. Amp	3	40	
ADC	3	100	
DAC	3	100	
Voltage Reg.	1	10	
Strain Gages	12	150	
Microcontroller	1	100	
Hardware	---		350
Coils	3	150	
Magnets	6	100	
Steel & Al.	---	100	
Assembly	---		500
Total	---		1430

APPENDIX

C

PROTOTYPE D PART DRAWINGS

All dimensions are in inches. The hole diameters are given entirely for 4-40 screws, except for holes J and L on the Outer Base.

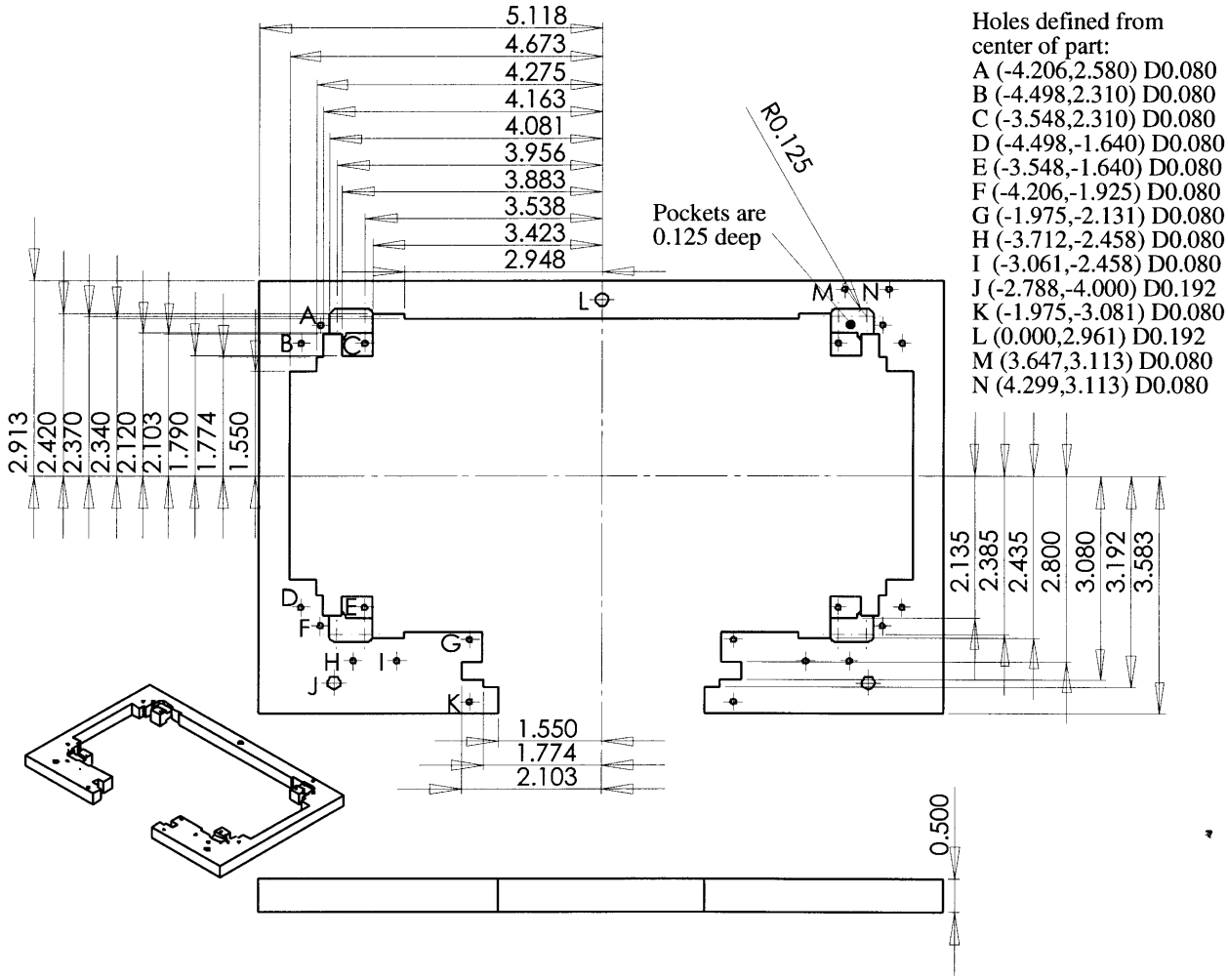

Figure C.1: Prototype D Outer Base.

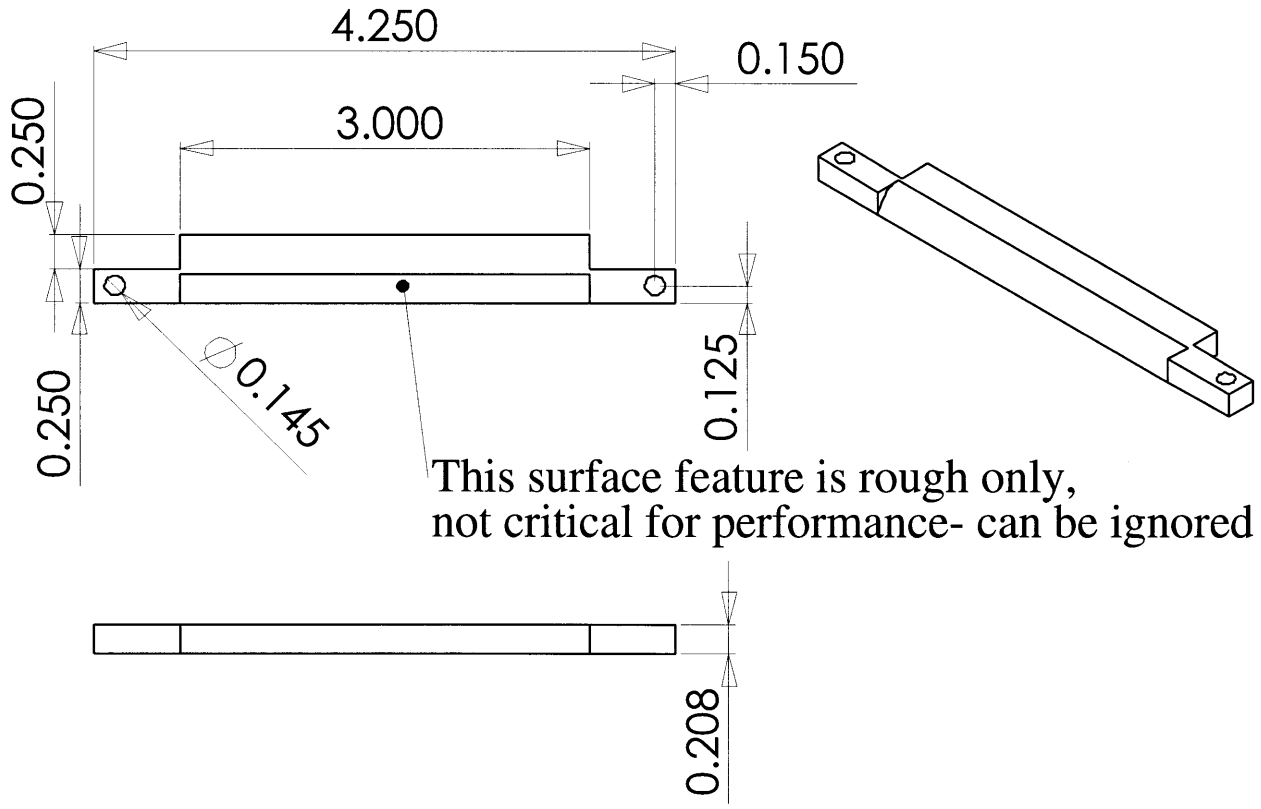

Figure C.2: Prototype D Core.

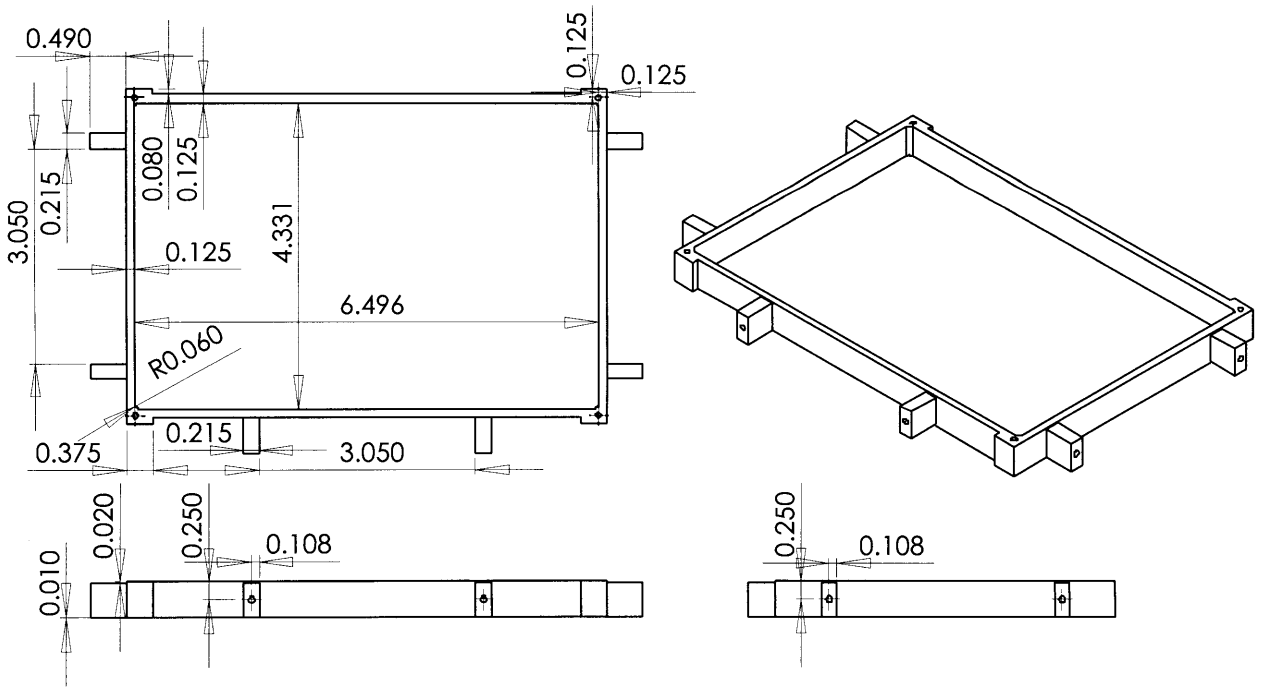

Figure C.3: Prototype D Center Stage.

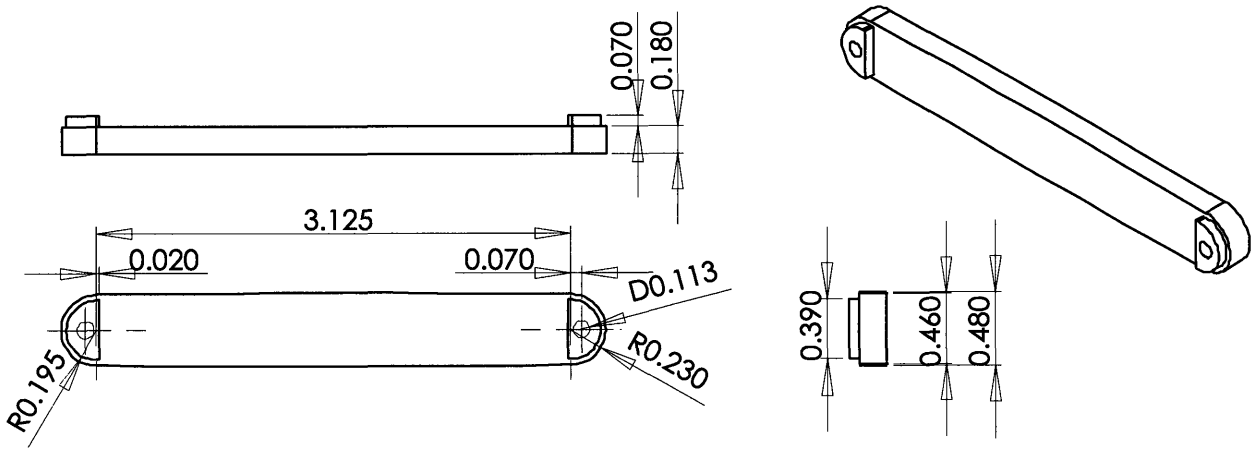

Figure C.4: Prototype D Mandril.

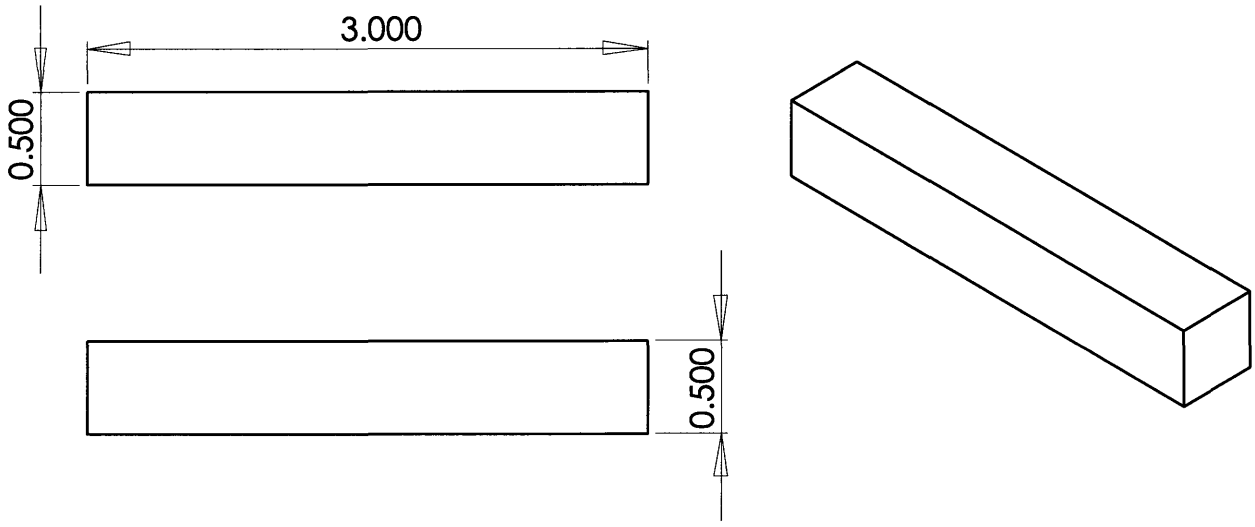

Figure C.5: Prototype D Permanent Magnet.

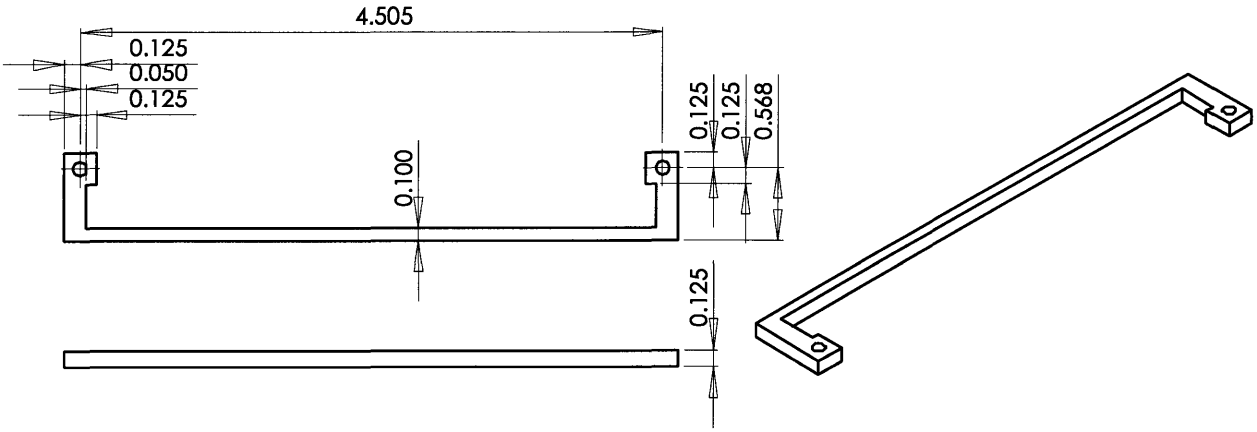

Figure C.6: Prototype D Outer Flexure Clamp.

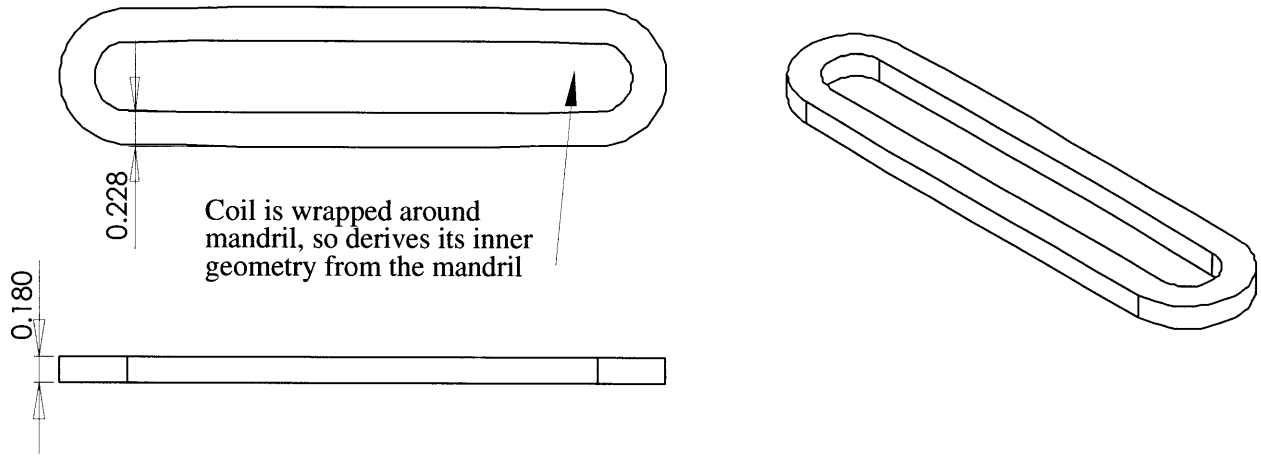

Figure C.7: Prototype D Coil.

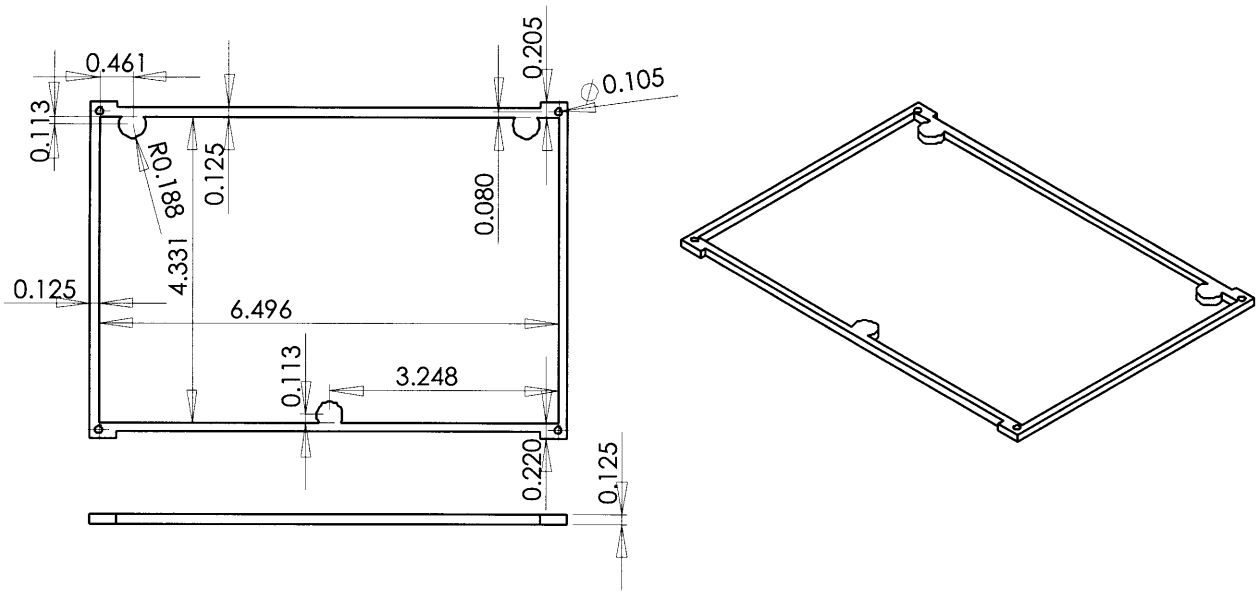

Figure C.8: Prototype D Inner Flexure Clamp.

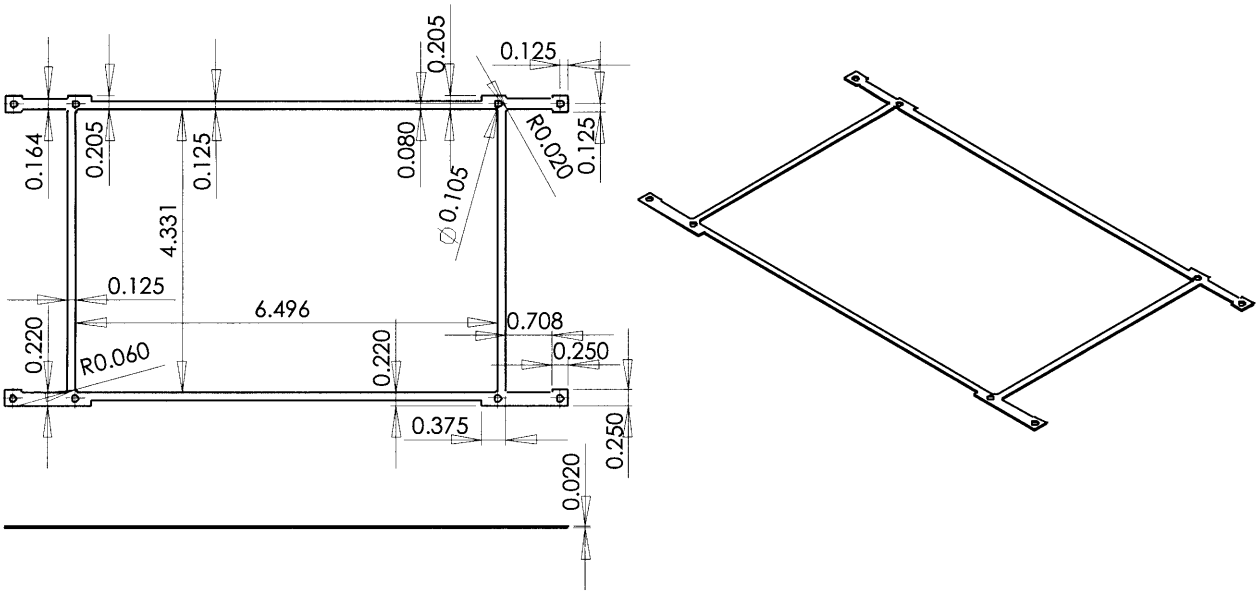

Figure C.9: Prototype D Flexure Sheet.