Electromagnetically-Driven Ultra-Fast Tool Servos for Diamond Turning

by

Xiaodong Lu

B.S., Automotive Engineering (1997)
B.S., Automation (1997)
S.M., Automotive Engineering (1999)
Tsinghua University

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

Doctor of Philosophy in Mechanical Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

September 2005

Author .....
Department of Mechanical Engineering
August 28, 2005

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

Accepted by .....
Lallit Anand
Chairman, Departmental Committee on Graduate Students

Electromagnetically-Driven Ultra-Fast Tool Servos for Diamond Turning

by

Xiaodong Lu

Submitted to the Department of Mechanical Engineering
on August 28, 2005, in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy in Mechanical Engineering

Abstract

This thesis presents the design, implementation, and control of a new class of fast tool servos (FTS). The primary thesis contributions include the design and experimental demonstration of: novel ultra-fast electromagnetic actuators, integration of these actuators in a new class of FTS, a novel real-time control computer with 1 million samples per second throughput, MIMO loop shaping techniques for parallel power amplifiers, and a novel configuration and controller tuning method for implementing adaptive feed-forward cancellation control. All of these elements have been successfully used for diamond turning of contoured surfaces.

Fast tool servos (FTS) are high bandwidth positioning devices, which, in conjunction with an ultra-precision lathe and diamond tooling, can produce free form surfaces with nanometer-scale resolution, such as required in micro-optical devices and light-enhancing films. The increasing complexity of such surfaces requires more components in shorter spatial wavelengths, and thus drives simultaneously the need for high bandwidth, high acceleration and high accuracy of the FTS. Conventional FTS solutions are based on piezoelectric stacks, which are typically limited to a few micron stroke at 1 kHz operation if not operated in a resonant mode.

As a promising alternative, this thesis demonstrates electromagnetically driven solutions for fast tool servos. The key new technology in these alternative designs is a new class of ultra fast electromagnetic drivers with thousands of G's acceleration capability in continuous operation. By separating the flux-biasing surfaces from the normal-flux working surfaces, this new driver design has a number of advantages: (1) actuating force linear to both excitation current and displacement, (2) modularity and parallel operating ability, (3) full magnetic stress utilization of normal surfaces, and (4) low heat dissipation. We analyze the operation characteristics, and also provide soft magnetic material selection criteria and motor design guidelines. Based on this ultra fast driver, we designed a linear fast tool servo with the theoretical capability for 1200 G acceleration in continuous operation.

To control such positioning devices at bandwidth of over 10 kHz, we developed a real-time computer architecture, utilizing three floating point digital signal proces-

sors (DSPs) in conjunction with a field-programmable gate array (FPGA) to significantly increase the processing rate. The real-time computer prototype experimentally demonstrated 1 million samples per second real-time control execution with a total latency of 1.9 microseconds when implementing a representative control algorithm of significant complexity. This processing system has capabilities far beyond what is commercially available for such real-time high-accuracy control tasks.

The power amplifier driving the FTS must supply 1 kVA (primarily reactive power) with over 100 kHz bandwidth. We present a solution of using 4 power operational amplifiers in parallel, each capable of supplying one fourth of the total power. To address the coupling issues among channels, a decoupling theory is developed to convert the associated MIMO plant into several SISO sub-plants, and thus ease the analog decentralized controller design of the power amplifier current feedback loops.

In order to enhance the FTS repetitive position trajectory tracking and disturbance rejection, adaptive feed-forward cancellation is embedded into a conventional motion control loop in our system. We provide a consistent loop shaping framework and intuitive parameter tuning and trade-off guidelines for this controller structure.

Experimental results with the first prototype FTS using powder iron cores demonstrate 23 kHz closed-loop bandwidth, as low as 1.7 nm RMS error, 30 micron stroke, 500 G peak acceleration at 10 kHz open-loop operation, and 2.1 nm (0.04%) error in tracking a 3 kHz sinusoid of 16 micron p-v. (The full 1200 G capability is expected to be experimentally demonstrated by the second prototype FTS using Ni-Fe tape cores). Using this FTS, we have diamond-turned two-dimensional sinusoidal surfaces in copper and aluminum with 0.5 degree azimuthal spatial period, 160 micron radial spatial period, and 2 micron peak-to-valley amplitude, at 500 RPM spindle speed.

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

Acknowledgments

I am most grateful to Professor Trumper for his guidance, encouragement, and help on my work. It is he who introduced me to the research field of electro-mechanical systems design and control, and helped me to build up the knowledge and skills. He is always patient and enthusiastic to explain and discuss the problems with me both on the theoretical part and experimental details. With his intelligence, perseverance, and enthusiasm, he has set a good example for me to learn from. He always encouraged me to pursue an academic career, brought me to conferences, and introduced me to experts in the fields whenever possible. Working with him taught me how to make creative ideas for design, how to get physical intuition on complex phenomena, how to make quick judgement on problems, and more important how to be an excellent researcher.

Professor Jeffrey Lang taught me electromechanics in 6.601, which established a solid foundation for my thesis work. I enjoyed very much his great explanation of the course content, which makes applying Maxwell equations to engineering problems a pleasure. In thesis committee meeting with him, he always pointed out key problems and suggested quick solutions. Professor Samir Nayfeh always gave me good ideas and sugestions on mechanical design, structural dynamics, and controls. I am amazed by his broad spectrum knowledge extending to almost every corner of mechanical engineering, and by his capability of both in-depth analysis and creative thinking. Professor David Perreault and Professor James Roberge provided valuable comments on my research work, and offered support and encouragement on my career development.

I am very grateful to Profess Minggao Yang and Professor Jianqiu Li at Tsinghua. They encouraged me to pursue overseas study at MIT, and also accommodated me at Tsinghua when I could not came back to MIT because of visa issues. During that time, I had the valuable opportunity to learn xPC Targets and graphical user interface programming. This unique experience later motivated many key ideas for the high-speed real-time computer in this thesis.

Lei Zuo is my classmate for more than 12 years, and is always a great sounding board to discuss various research topics. Our very interactive discussion often turned into very fruitful results. Having Lei around me at MIT made my life at MIT much more enjoyable. Richard Montesanti is a great engineer from LLNL with tremendous knowledge and experience in precision engineering. When I had questions on mechanical design, I often turned to him. His passion, hard-working, and well-orgnized style on research greatly influenced me. I enjoyed many discussion with him on circuits and controls issues, and he gave me valuable feedback on my power amplifiers. What's more, Rick constantly encouraged me to pursue an academic career. Marten Byl gave me a lot of help on using the diamond turning machines, and was always eager to show me his new experiment result. Katie Lilienkamp developed one of the most useful software: dynamic signal analyzer for dSPACE, which helped a lot for my dSPACE-based research. Ming-chih Weng gave me a lot of help during my early stage study at MIT. He walked me into his broomstick levitation projects and helped me established a solid fundation for my later work on moment/slope control of flexible structures. David Otten taught me about Gerber files for PCB board and provided suggestion on PCB design.

Over past many years, I was lucky to have many great colleagues and enjoyed very much working with them. They provided me useful advice, suggestion, or comments on my work. They have included Mike Liebman, Marty Vona, Amar Kendale, Andy Stein, Joseph Cattell, Justin Verdirame, Vijay Shilpiekandula, Aaron Mazzeo, Augusto Barton, Larry Hawe, David Cuff, Yi Xie. I thank David Rodriguera, Maggie Sullivan, Denise Moody, and Laura Zaganjori for their help ordering equipment. I thank Mark Belanger and Gerry Wentworth for their directing and help on the machine shop, and I always enjoyed their training in a relaxed style.

I am most indebted to my parents and Maggie. They are my source of support, love, and encouragement, and have made tremendous sacrifice for my study at MIT.

This project is supported by the US National Science Foundation under grant DMI-0322590.

For Mom, Dad, and Maggie

献给妈妈，爸爸，和光睿

Contents

List of Figures

List of Tables

Chapter 1

Introduction

There exist increasing needs in industry for free-form surfaces with optical quality surface finish. One of the most flexible ways to manufacture this kind of surface is single-point diamond turning enhanced with fast tool servo (FTS) technologies. Diamond turning operation can generate optical quality surfaces by using a single-crystal diamond tool on an ultra-precision lathe. The fast tool servo technology enables very fast tool motion at this diamond tool tip. As a result, the machined surface is not limited to rotationally symmetric surfaces as in a conventional turning process. The increasing complexity of surfaces requires more components in shorter spatial wavelengths, and thus drives simultaneously the need for high bandwidth, high acceleration and high accuracy of the FTS. The conventional solutions for FTS are primarily based on piezoelectric actuators, the performance of which is severely limited by the hysteresis energy loss at over 1 kHz operation. In order to advance the state-of-art technology of fast tool servos, this thesis explores an alternative solution, which is based on electromagnetic actuation. The initial experimental results have achieved much better bandwidth and higher acceleration than piezo-based fast tool servos designed for wide-band operation. In this chapter, we introduce the background of state-of-the-art fast tool servo technology. Next, we present an overview of this thesis and outline the whole thesis structure. Finally, we summarize the main contributions.

1.1 Background

Sculptured surfaces with nanometer-scale resolution requirements are used in a wide range of products [5], and these fine scale surfaces "offer designers additional freedom to create novel functions or combinations of functions" [15]. Examples of products with sculptured surfaces include films for brightness enhancement and controlled reflectivity, sinewave ring mirrors used in CO2-laser resonators, molds for contact lenses, fusion experiment targets, two-dimensional planar encoders, as well as in micro-optical devices such as Fresnel lenses, multi-focal lenses and micro-lens arrays.

A flexible way of manufacturing these surfaces is diamond turning in conjunction with fast tool servos. In the diamond turning process, a single-crystal diamond cutting tool is installed on an ultra-precision machine to cut surfaces with optical-quality surface finish. A conventional diamond turning machine can only cut rotationally symmetric surfaces, such as flat, sphere, and cone surfaces, because the diamond tool is maintained at an essentially constant position relative to the machined part within each revolution. When enhanced with a fast tool servo, the diamond tool tip can move rapidly within each revolution of the workpiece. As a result, free-form (or sculptured) surfaces can be machined. The types of surfaces that can be machined are highly dependent on the fast tool servo performance, mainly limited by the achievable acceleration and bandwidth. The increasing complexity of such surfaces requires more components in shorter spatial wavelengths, and thus drives simultaneously the need for high bandwidth, high acceleration and high accuracy of the FTS. The main focus of this thesis is to design new actuation mechanisms and associated enabling technologies to advance the state of the art in fast tool servos, in terms of bandwidth and acceleration.

1.2 Prior Art Fast Tool Servos

A significant amount of research effort has been devoted to FTS design. There are many FTSs reported in the literature using a variety of actuating mechanisms and

configurations, depending on the application needs. Ludwick's thesis has an extensive literature review on existing fast tool servos before 1999 [31]. Based on working principles, these FTSs can be categorized mainly as four types: piezoelectric actuators, magnetostrictive actuators, Lorentz force motors (including linear and rotary), and normal-stress electromagnetically-driven actuators. These devices can also be categorized according to their operating range, or stroke. We define short-stroke as less than 100 μm, intermediate as between 100 μm and 1mm, and long-stroke as above 1 mm. Other actuating mechanisms can also be used for FTSs, such as a hydraulic actuator with 180 μm stroke and 10 Hz bandwidth [81], but these categories are not the focus of this thesis because of their relatively low bandwidth.

1.2.1 Piezoelectric FTS

Most short-stroke FTSs and similar micro-positioning devices are based on piezoelectric stacks [74], which have the advantage of quick response, high acceleration, accurate micro positioning, and high stiffness (usually greater than 50 N/μm in the typical sizes used). The piezoelectric ceramic materials can produce up to 0.1% normal strain under electric excitation field, readily achieve bandwidth on the order of several kHz, and are capable of nanometer resolution positioning. What's more, piezo actuators are open-loop stable, and thus relatively easy to control.

However, piezoelectrically actuated FTSs have also disadvantages. Due to mechanical and electrical energy losses in the material when undergoing deformation with hysteresis, not only the tracking performance is degraded by the nonlinear behavior, but also a significant amount of energy is converted into heat. Especially in high-frequency operation, the effective stroke therefore reduces significantly. In some cases with both high frequency and high acceleration requirements, liquid cooling has to be applied. As the piezoelectric ceramic material can admit much smaller tensile stress than compressive stress, the piezo stacks must also be pre-loaded properly. As well, care must be taken in designing the motion guidance mechanism to couple the piezoelectric material to the moving payload in such a way as not to introduce parasitic strains in the actuator, which might lead to cracking. Usually, large and

expensive high-voltage amplifiers are required to drive these devices. Another problem is that the structural resonance modes of the Piezo stack itself limit the working frequency range, because operation near resonances can cause local tensile failure of the PZT ceramics.

In this class of actuators, we can cite the following contributions from the literature: Kouno published the design for a piezoelectric fast tool servo of 6.5 μm stroke, 10 nm resolution, 70 Hz bandwidth and 300N/μm stiffness [33]. The piezoelectric stack consists of 18 rings, each of which is 25 mm OD, 15 mm ID, and 1 mm thick.

Patterson and Magrab designed a fast tool servo for diamond turning with 2.5μm stroke, 660 Hz bandwidth (-3dB), and 1.3 nm dynamic repeatability at Lawrence Livermore National Laboratory [66]. A PZT stack of 6 mm diameter and 12.5 mm length provided stiffness sufficient to place the resonance above 1 kHz when supporting the maximum FTS body weight of 450 gram. The moving portion of the FTS consisted of a cylindrical shell supported by two diaphragm flexures fixed to a cylindrical support.

Faulter and Dow developed an FTS of 20 μm stroke and 2 kHz open-loop bandwidth using a hollow piezoelectric actuator (25 mm OD and 18mm long) [17, 13, 19]. The resonance frequency was approximately 10 kHz. However, at 1 kHz this FTS has a maximum stroke of 5 μm and could not work continuously because of internal heat generated by losses inside the piezoelectric actuators . Based on similar concepts, Jared developed a piezoelectric FTS of 11.4μm stroke and 200 Hz bandwidth for inertial confinement fusion targets [27].

Okazaki developed a fast tool servo with 5 μm stroke, 470 Hz bandwidth, and 5 nm resolution [62]. Later, Okazaki developed an improved fast tool servo with 15 μm stroke, 2.5 kHz bandwidth, 2 N maximum cutting force, and 2 nm resolution [63]. A stacked ring piezoelectric actuator of 25 mm OD, 14 mm ID and 19 mm length was fixed inside a stainless steel cylinder, with a size of 53 mm in diameter and 112 mm length. The primary resonance frequency was at 10 kHz. However, at 1 kHz continuous operation the effective stroke is only several micrometers.

Hara designed an FTS of 3.7 μm stroke and 2 kHz flat frequency response in open loop [24].

Rasmussen and Tsao described a system for asymmetric turning, by using a piezoelectric actuator to drive a tool through a lever assembly [70]. The FTS had 50 μm stroke and 200 Hz bandwidth.

Falter and Youden designed a device that is commercially available on a machine for contact lens generation. This fast tool servo has 400 μm stroke at DC and 600 Hz bandwidth. It uses a T lever to achieve 14:1 mechanical motion amplification [16]. A pair of PZT stacks against the horizontal part of the T lever provides push and pull motion.

Cuttino developed an FTS of 100 μm stroke and 100 Hz bandwidth [10]. In order to obtain the 100 μm stroke against the preloaded rod, the PZT stack was 13 cm long and had an open-loop stiffness of around 70 N/μm. As a result, the natural frequency of the piezoelectric system was 250 Hz.

Altintas presented a piezo tool actuator of 370 N/μm stiffness, 36 μm stroke, 50 Hz bandwidth for holding or 300 Hz bandwidth for dynamic contour machining, and 10 nm positioning during finish hard turning [1].

In summary, although piezoelectric actuators have significant applicability in practice, they appear to be limited to about 1 kHz operation and become more difficult to apply as the bandwidth and acceleration requirements are increased.

1.2.2 Magnetostrictive FTS

As a dual to piezoelectric materials that respond with strain to an applied electric field, magnetostrictive materials stain under an applied magnetic field, and this is another option for FTS. A common magnetostrictive material is called Terfenol, but it also has hysteresis problem [39]. It is claimed that magnetostrictive materials have higher maximum strain and less temperature dependent, in comparison to piezoelectric materials. However, their performance as reported in the literature appears to fall significantly behind that achieved by piezoelectric devices.

There are some research efforts on such devices in the literature. Eda developed a giant magnetostriction actuator with a stroke of 2 μm [14], but no frequency response of this device was reported. Liu *et al.* developed a magnetostrictively-actuated tool

holder of 50 μm stroke to investigate vibration abatement in turning processes [39]. The actuator had a resonance frequency at 1.5 kHz, but no closed loop bandwidth specification was reported for this device.

1.2.3 Lorentz Force FTS

Commonly used in, for example, voice coil motors, the Lorentz force can be applied to drive fast tool servo. This type of actuator can have much higher stroke than those of the piezoelectric FTSs. But the maximum acceleration is strictly limited by the heat generated in the coil and by the achievable magnetic flux density. The acceleration of these devices as reported in the literature tends to fall below 100 G's.

Douglass used a voice-coil-motor driven slide for a fast tool servo [12], which had a stroke of 500μm and 100 Hz bandwidth.

Wang described a voice-coil-based actuator to achieve 0.4 mm stroke and 0.1 μm resolution [83].

Weck *et al* developed a hybrid long stroke fast tool servo, which integrated a piezoelectric system and a linear motor [84]. The piezoelectric system has a stroke of 40 μm at 1 kHz bandwidth, and the linear motor has a stroke of 2 mm at 40 Hz. The piezoelectric translator had a maximum force of 2400 N, a stiffness of 50 N/μm and a resonance frequency with the loaded mass close to 2 kHz. The peak thrust force was 900 N with a time constant of 5 ms. The power loss of the linear motor and the piezoelectric translator was dissipated by a water-cooling system. Temperature stabilization was guaranteed by Peltier elements, which control the heat dissipation to the cooling media. A compensation of these forces to reduce excitation of the supporting machine was achieved by the use of a balance mass.

Greene and Shinstock developed a linear voice coil based fast tool servo [21]. This FTS had a 100 Hz bandwidth and 6 mm stroke. Because of the low mass of the actuator-tool assembly and the low bandwidth, the dynamic stiffness at 15 Hz was so low that it was considered not feasible for FTS application due to chatter instability.

Todd and Cuttino built a long-range traction-type FTS for diamond turning applications [79]. The tool shuttle was driven by a rotary motor through a steel ribbon.

The achieved performance was 1mm stroke at 20 Hz.

Ludwick and Trumper used a commercial brushless motor to develop a rotary fast tool servo with a bandwidth of 200 Hz, 50 G peak acceleration, micrometer resolution, and a stroke of 3 cm [45]. Because of its balanced rotary design, reaction forces are essentially cancelled, and only reaction torques are transmitted to the supporting machine base.

Montesanti designed a rotary FTS using a commercial moving magnet motor to achieve 50 μm stroke, 20 G's tool acceleration, and 2 kHz bandwidth [56].

1.2.4 Normal-Stress Electromagnetically-Driven FTS

As opposed to shear stress actuation in moving-coil or moving-magnet motors, normal-stress solenoid-type actuators can also be used for FTSs, because normal-stress motors potentially have higher force density and thus possibly higher acceleration than shear-stress motors. Very few people have looked at variable reluctance actuation for FTSs, probably because of the difficulty of controlling these devices in the presence of inherent nonlinearities.

Stancil, Gutierrez and Ro developed a normal-stress fast tool servo with a stroke of 800 μm [77]. This large stroke design did not take advantage of the high acceleration potential of the variable reluctance actuator, probably due to the low achievable flux density at such large air gaps. The achieved bandwidth is 100Hz, and the tracking performance at 30Hz is problematic [22]. Later a sliding mode control algorithm was applied, and the tracking performance of FTS at its resonance frequency 200Hz was greatly improved [23].

As a first design effort for this thesis, Lu and Trumper presented an electromagnetically actuated FTS with 50 μm stroke, 160 G acceleration to frequencies of 3 kHz, 10 kHz bandwidth, and 1.4 nm resolution [41, 43]. Figure 1-1 shows the cross-sectional representation of the configuration for the variable reluctance FTS. The two solenoids in yellow provide push and pull force to the armature in aqua blue, which is attached to a carbon fiber tube as the backbone of the moving system. The coil windings are implanted into the slots of the solenoids. The diamond tool holder is installed at

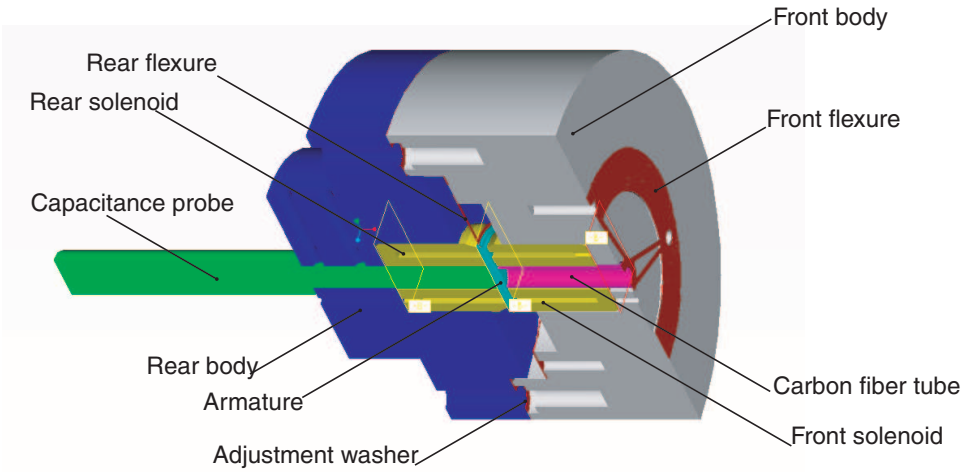


Figure 1-1: Configuration of a variable reluctance fast tool servo.

the other end of the carbon fiber tube. The whole moving system is suspended by a front flexure and a rear flexure. These flexures constrain five degrees of freedom of the moving system while allowing only motion in the axial direction of the carbon tube. The motion of the armature is measured by the capacitance probe which passes through the center of the rear solenoid. Figure 1-2 shows a part machined by this fast tool servo.

However, these solenoid type actuators are highly nonlinear, because the actuating force is proportional to current squared and inversely proportional to air gap squared. This nonlinear behavior will introduce significant errors in high frequency operation. Flux biasing can be used to eliminate nonlinearity [57, 42].

In our lab, Richard Montesanti designed a hybrid rotary/linear fast tool servo with a stroke of 70 μm and 10 kHz closed-loop bandwidth [57, 59]. This fast tool servo achieves 400 G acceleration at 5 kHz and 870 G acceleration at a mechanical resonance frequency of 10 kHz. This detailed electromagnetic design of this fast tool servo will be discussed further in Chapter 2.

This thesis describes a fast tool servo design of 30 μm stroke and 23 kHz closed-loop bandwidth [42, 44], with experimentally-demonstrated acceleration of 500 G.

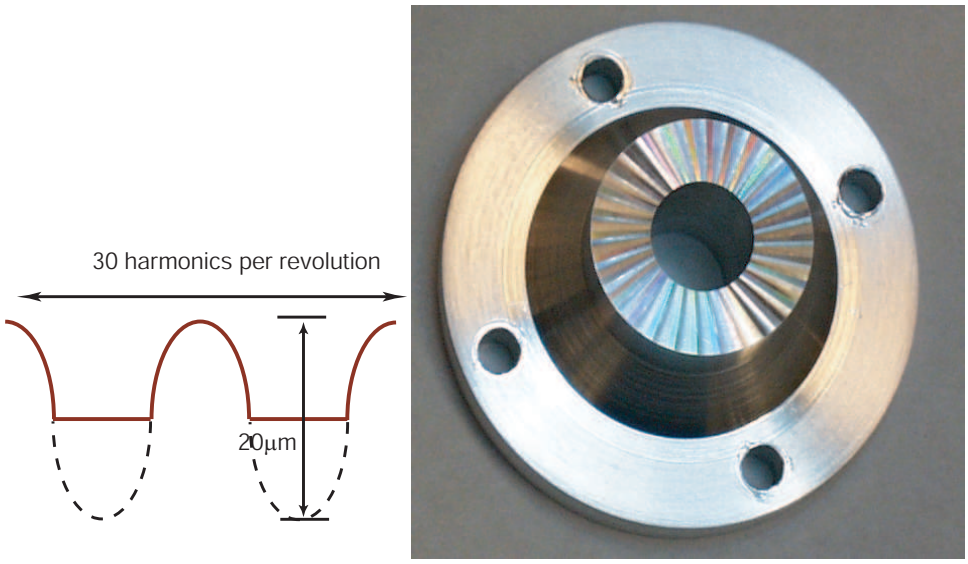


Figure 1-2: Diamond turned part by the electromagnetically driven prototype FTS. The surface is machined by face turning. The spindle speed is 1800rpm. The profile expanded in the circumference is half sinusoidal wave, as illustrated in the bottom left curve. There are 30 harmonics for each spindle revolution, and the peak to valley amplitude of the sine wave is 20 μm. The flat surface was machined first and then the sinusoidal surface was cut.

1.2.5 Performance Comparison

The DC stroke and small signal bandwidth performance of some typical FTSs discussed above are illustrated in the Figure 1-3. It can be observed that Lorentz FTSs have larger strokes but lower bandwidth than piezoelectric FTSs. However, one must be careful to interpret this data, because the DC stroke and the bandwidth are two separate performance parameters, they cannot be achieved simultaneously for most cases. For each FTS, the actual stroke is varying with frequencies when under pure-tone sinusoidal excitation, because the allowable stroke is determined by the power dissipating capability and has nothing related to the small signal frequency response. Unfortunately, most FTSs in the literature didn't report the actual travel as a function of frequency, so this figure gives the least available comparision metric.

1.3 Thesis Overview

Based on the literature review, we decided to design and build a new class of fast tool servos to advance the state of arts technology. By a theoretical analysis, we realize that a normal stress actuator can generate a pressure of about 9×10^5 Pa as shown in Figure 1-4, by assuming a flux density of 1.5 T, which is below the saturation of many soft magnetic materials. If we can apply this pressure to an armature iron disk of 3 mm thickness, then the resulting acceleration will be 4000 G.

Based on this analysis, we predicted that normal-stress actuators can be used to drive a FTS with unprecedented acceleration and bandwidth performance. The goal of this thesis is to fully release the high acceleration potential of normal stress actuators by novel electromagnetic configuration, and develop associated enabling technologies to demonstrate the achieved performance.

1.3.1 Ultra Fast Motor

A key contribution of this thesis is a new electromagnetic actuator configuration as shown in Figure 1-5, which we refer to as a canonical ultra fast motor. By intro-

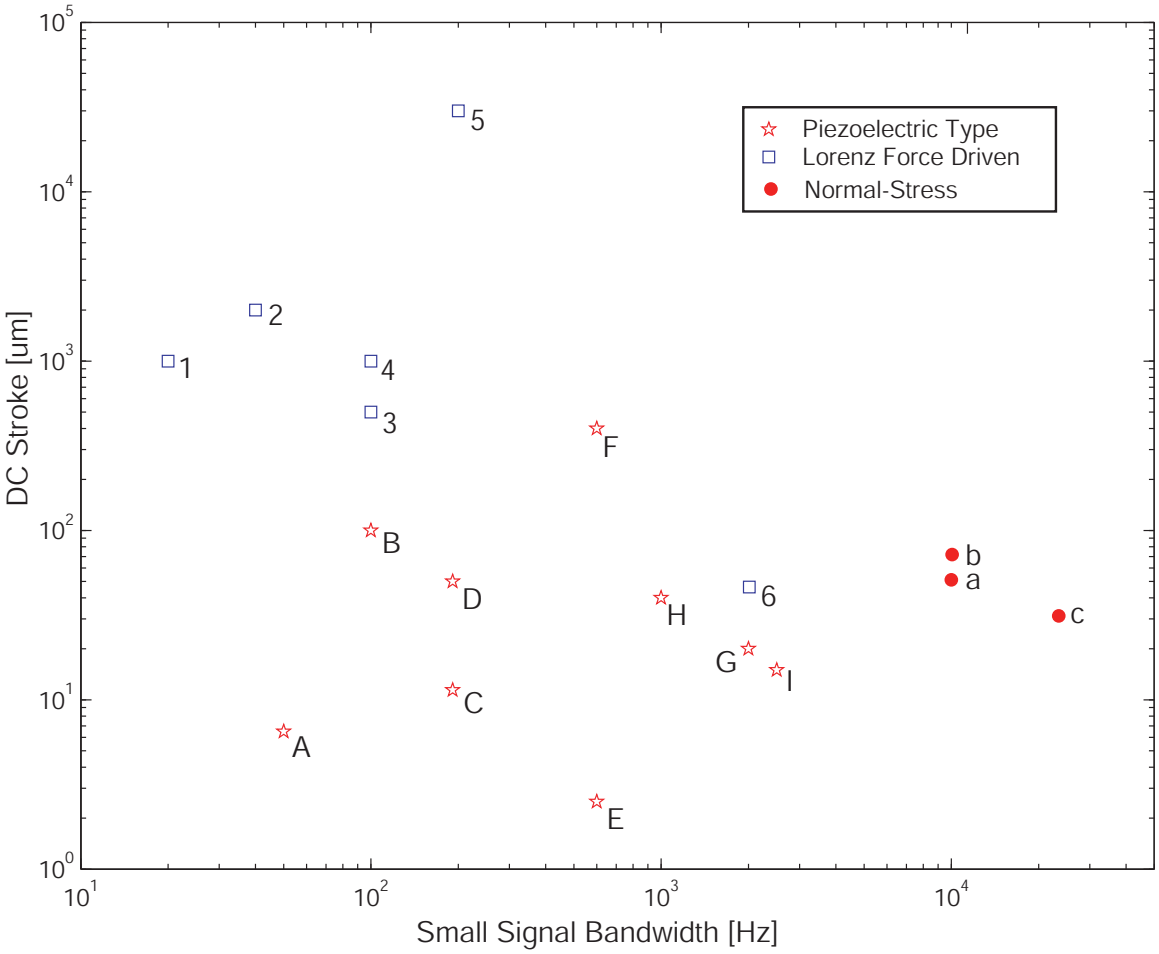


Figure 1-3: Performance comparison of main types of FTS, with reference listed in the following table.

Lorentz FTS		*Piezo FTS*		*Normal-stress FTS*	
1	Todd [79]	A	Kouno [33]	a	Lu [41]
2	Weck [84]	B	Cuttino [10]	b	Montesanti [59]
3	Douglass [12]	C	Jared and Dow [27]	c	Lu [44]
4	Greene [21]	D	Rasmussen [70]		
5	Ludwick [45]	E	Patterson [66]		
6	Montesanti [56]	F	Falter [16]		
		G	Dow [13]		
		H	Weck [84]		
		I	Okazaki [63]		

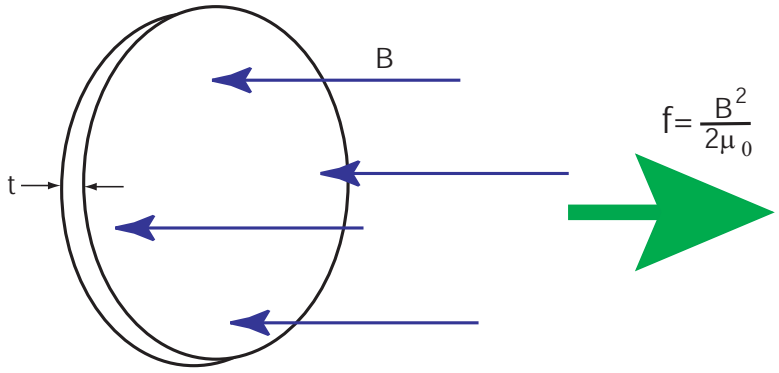


Figure 1-4: Force generation on an armature disk. Here f is the actuating force per unit area, B is the magnetic flux density, μ_o is the permeability of air, and t is the armature thickness.

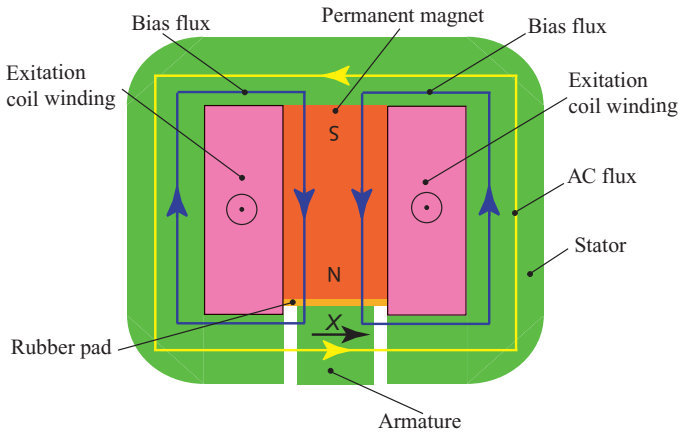


Figure 1-5: A cross-section of the canonical ultra fast motor.

ducing permanent-magnet-generated DC flux through a bias surface on the top, the full normal surface on the left and the right of the armature can be used for force generation. What's more, this actuator has linear characteristics and its actuating force is proportional to the excitation current. This linear behavior is very important for such precision application as FTS. At its maximum force operation, the heat generation of this actuator is only 25% that of a solenoid actuator with similar force and air gaps but much larger armature size. All these advantages combine to allow very high acceleration, high bandwidth, and high accuracy simultaneously. The detailed design and analysis of this actuator are presented in Chapter 2 and Chapter 3. Based on this building block, we also developed a series of electromagnetic actuators for various applications as presented in Chapter 2.

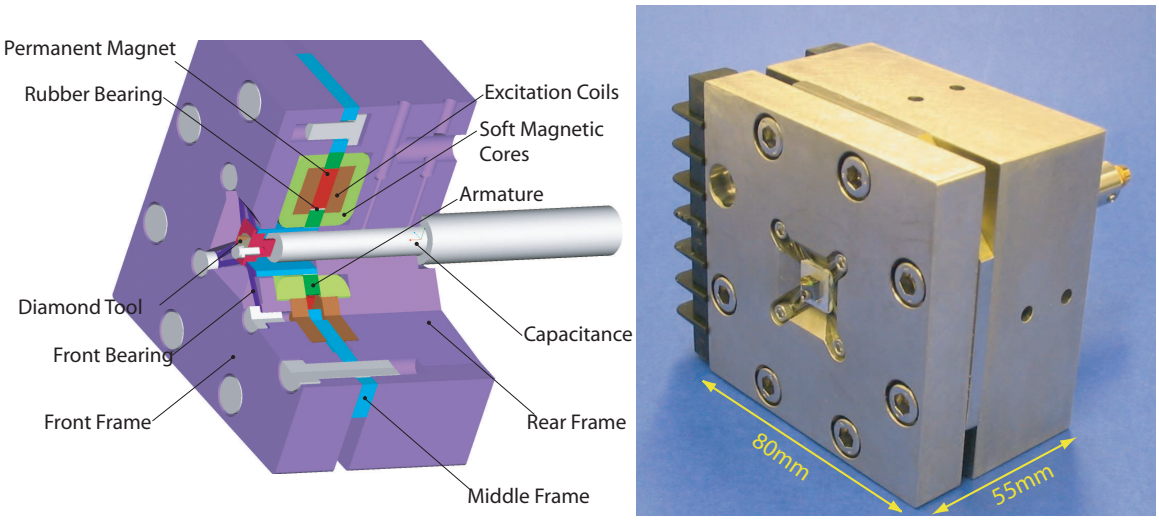


Figure 1-6: The ultra fast tool servo Pro-E model and photo.

1.3.2 Ultra Fast Tool Servo

Using the linear ultra fast motor as the core component, we created designs for ultra fast tool servos with two types of soft magnetic magnetic materials: SM2 powder-pressed iron and Ni-Fe wound-tape cores. Figure 1-6 shows the Pro-E model and a photo of the ultra fast tool servo (UFTS) using SM2 as the core material, the Ni-Fe device was not yet fabricated. The total moving mass is only 11.2 gram, and the first destabilizing structural resonance is at 45 kHz for SM2 based UFTS and 25 kHz for Ni-Fe based UFTS, because Ni-Fe is a laminated structure and thus not as rigid as the solid powder iron material. The design and fabrication details of the UFTS are in Chapter 4. In order to facilitate the structural design, sensing ability and controlling ability are proposed to evaluate structural dynamics, as described in Chapter 4. At the time of writing this thesis, only the first prototype using the SM2 powder-pressed iron as core material is fabricated and tested.

1.3.3 Power Amplifier

In order to drive the designed UFTS, a linear amplifier with 1000 VA reactive power capability and over 100 kHz bandwidth is required. As no commercial solution is available to meet these specification, we built a linear amplifier using four PA52

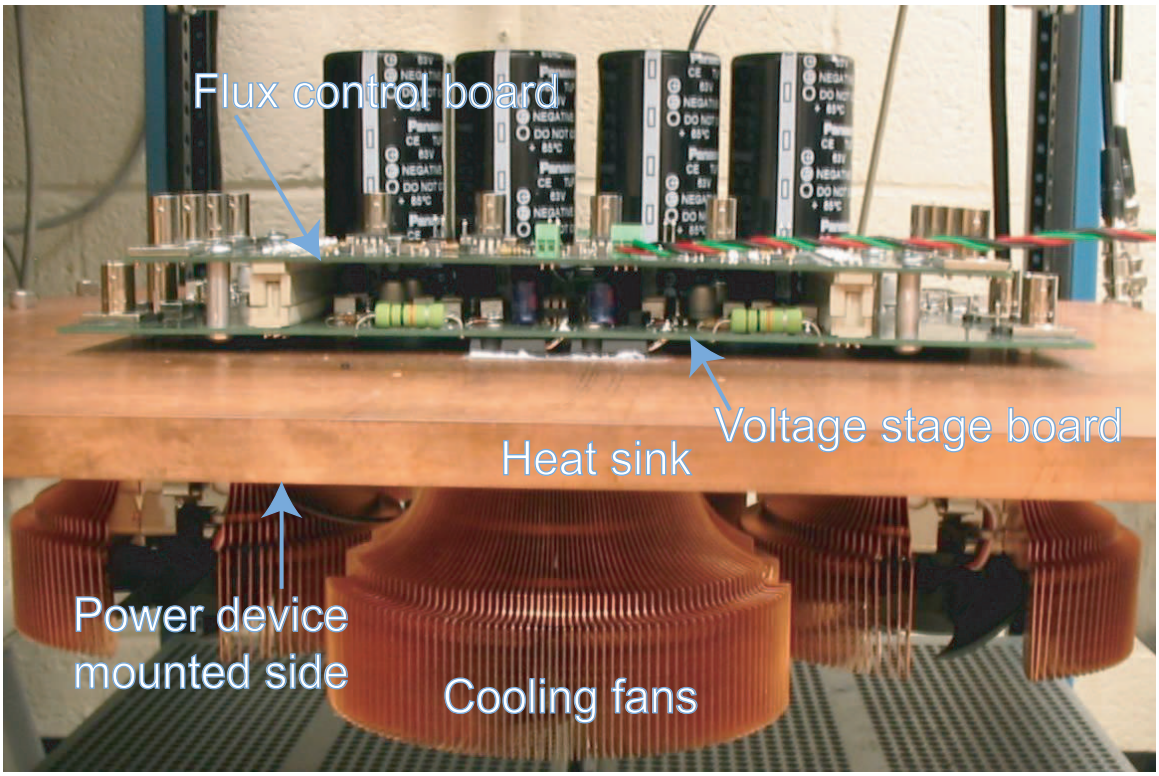


Figure 1-7: The prototype linear power current amplifier mounted on a 19-inch lab rack. A half-inch copper plate is used as both the mechanical backbone and the heat sink connecting between power devices and the cooling fans. Two separate PCB boards are stacked together for voltage control and current (or flux) control respectively.

power devices in parallel operation as shown in Figure 1-7, and presented in detail in Chapter 5. As there exists strong coupling among multiple power devices through the mutual inductance of the driven load, we developed a decoupling theory to greatly simplify the Multiple-Input-Multiple-Output (MIMO) controller design for high-power high-bandwidth current amplifiers.

1.3.4 Real-time Computer

The bandwidth specification on the UFTS is 10 kHz crossover frequency on the loop transmission. Accordingly, the required real-time computer should be capable of at least 200 kHz sampling rate for a control algorithm which includes significant complexity. As we couldn't identify any commercial solution meeting this specification,

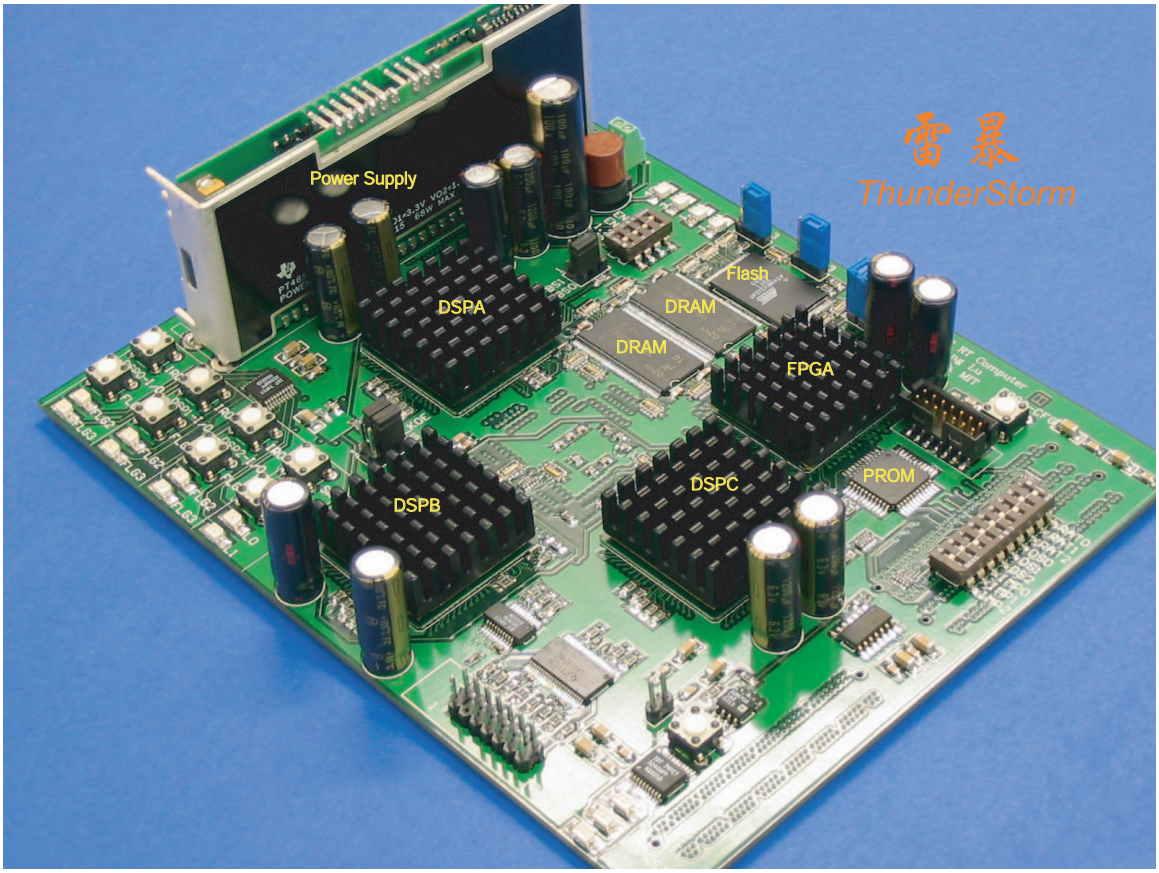


Figure 1-8: Picture of the real-time computer mother board. The dimension is 160×125 mm^2.

we designed a real-time computer using three floating-point digital signal processors as shown in Figure 1-8, and described in detail in Chapter 6. The experimental result shows that the total latency is 1.8 μs, and the achievable sampling rate is 1 MHz when running a representatively complex test algorithm.

1.3.5 Fast Tool Servo Control

Chapter 7 discusses the UFTS controller design. Figure 1-9 shows the controller structure. The compensated loop transmission without adaptive feedforward controllers (AFC) achieves a crossover frequency of 10 kHz with 30 degree phase margin as shown in Figure 1-10. For large signal response, this actuator has demonstrated 500 G acceleration at 10 kHz open-loop operation.

We present a new method for AFC controller implementation/design. By putting

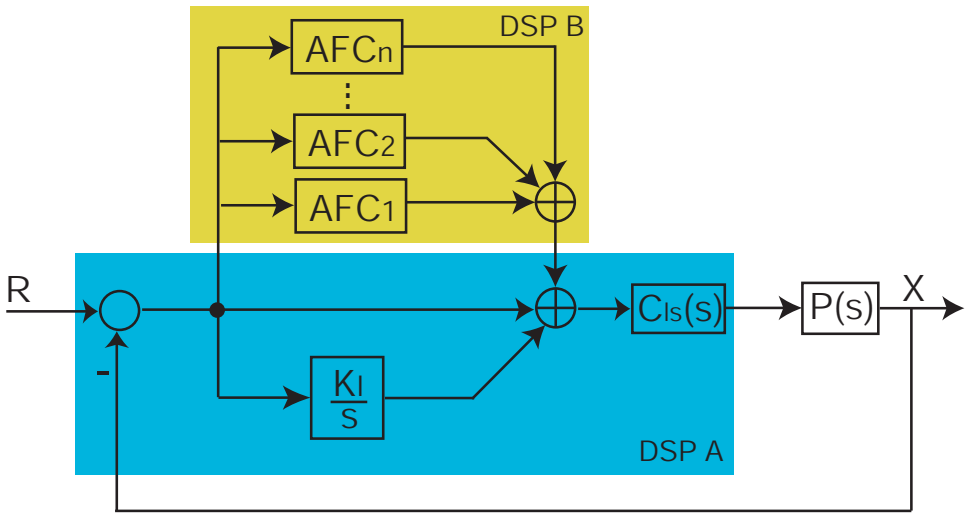


Figure 1-9: The controller structure for the SM2-based UFTS. The controller is composed of three parts: a loop-shaping controller $C_{ls}(s)$, an PI term $1 + K_I/s$, and adaptive feed-forward controller (AFCs).

the AFC controllers in parallel with the integrator, we can design the conventional motion controller ($C_{ls}(s)$ and K_I/s) and AFC controllers in one consistent loop. What's more, the parameter design for the AFC controllers can be greatly simplified with clear intuition. The inherent connection between integrator and the AFC controlllers are identified and the gain trade-off criteria are provided. Also a fundamental property for the AFC controller family, referred as equivalence of gain and bandwidth, are presented and proved for both sub-crossover AFCs and super-crossover AFCs. These properties provide important intuition for AFC controller design. The effectiveness of the designed AFC controllers has been demonstrated experimentally by tracking a 16 μm p-v sine wave at 3 kHz, and the resulting tracking error of only 2.1 nm RMS.

Other issues such as multi-processing and multi-sampling are also discussed. A synchronous filtering method is presented to attenuate aliasing noise and nonlinearity errors of the spindle encoder.

1.3.6 Experimental Results

Finally, we integrate the UFTS with a commercial diamond turning machine as described in Chapter 8. Both 1-D and 2-D sinusoidal surfaces have been cut on alu-

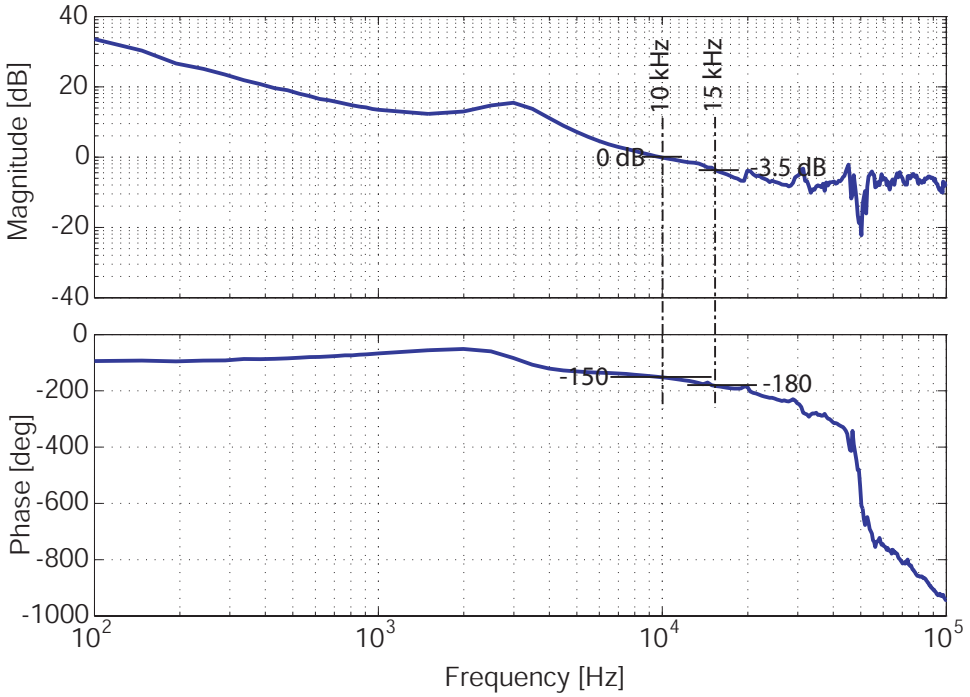


Figure 1-10: The measured negative loop transmission of the compensated UFTS without AFCs.

minum and copper parts with an azimuthal period of 0.5° at 500 RPM, and a corresponding UFTS working frequency of 6 kHz. No chatter was observed for cutting at depths from 0.5 μm to 30 μm.

Figure 1-11 shows an interferometric microscope image of the machined 2-D surface.

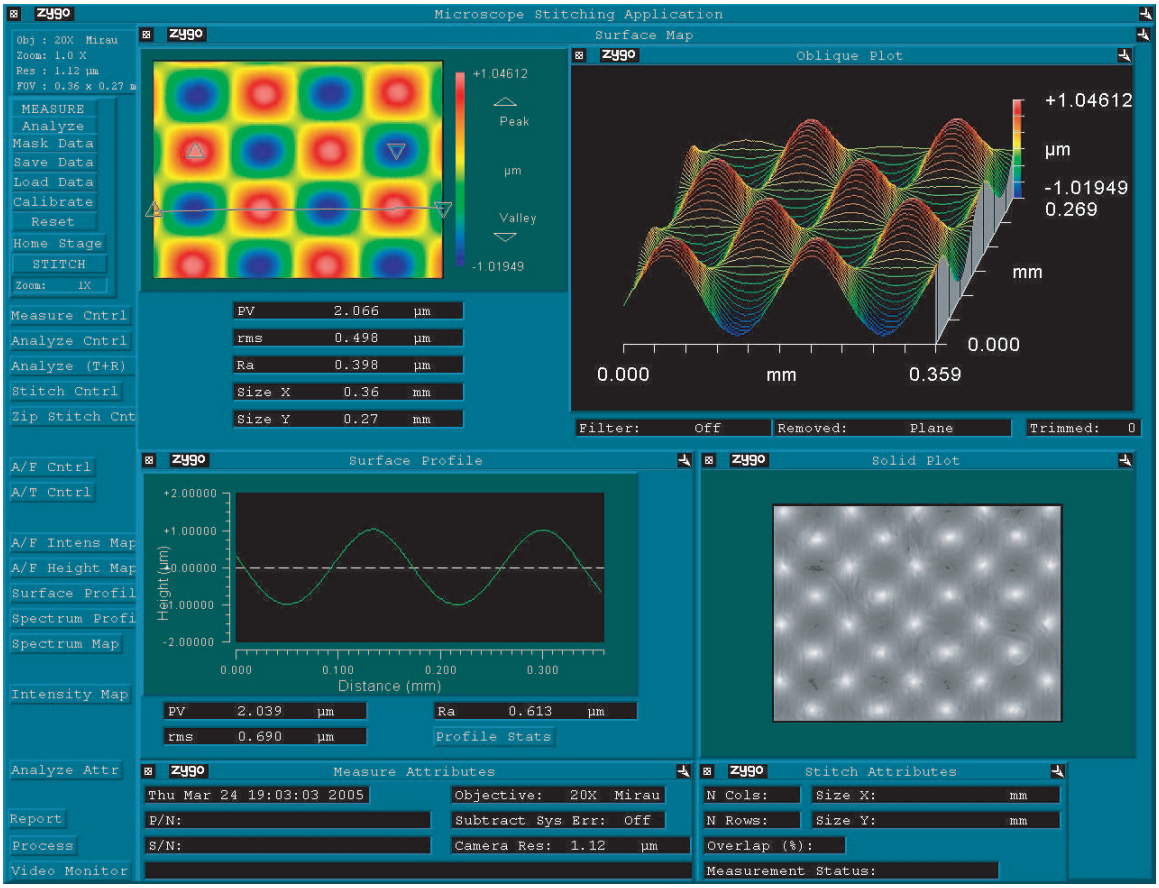


Figure 1-11: Surface metrology for diamond turned 2-D surface on an OFHC copper workpiece. The spatial period of the 2-D sinusoidal surface is 0.5 degree in the θ direction and 167 μm in the radial direction. The spindle speed is 500 rpm. The peak-to-valley amplitude of the sine wave is 2 μm.

Chapter 2

Ultra Fast Motor Design

In this chapter, we present the electromagnetic design of a new class of ultra fast motors [1] which have acceleration capability of over 1000 G's and a working stroke of tens of microns.

First, we examine two basic electromagnetic actuators configurations: a moving coil motor, and a normal stress motor. Moving coil actuators typically have a steady-state maximum acceleration of 100 G due to the thermal limit on current-carrying coils. Normal stress motors with an armature of 3 mm thickness can potentially achieve an acceleration of over 1000 G's.

Next, we review various normal-stress actuator designs in the literature. The most commonly-used normal stress actuators are solenoids. One key idea for achieving high acceleration without losing actuating force is to use a multi-polar design, which can be interpreted as multiple small solenoids working in parallel. A common problem in many normal-stress actuators is the non-linear force-current relationship, which can introduce significant tracking errors in high-performance servo applications. We examine several linearizing techniques, including current biasing and feedback linearization methods. However, there exist a special family of normal-stress actuators that have linear operation characteristics, and we refer to these as hard-linearized normal stress motors. The common design features of this family of actuators are

[1] In this thesis, we use the term of motor and actuator interchangeably. We also refer to **steady-state** acceleration capabilities; transient performance typically is higher.

summarized.

Finally, on the basis of the above review and examination, we present the design of a new class of normal-stress actuators, which belong to the family of hard-linearized normal-stress motors. This new class of normal-stress actuators are characterized by flux-biasing between armatures and stators, and thus are capable of using almost all armature face area for force generation. This feature is the fundamental difference from all the designs to be reviewed in the literature. Another advantage of this new class of actuators is that they are inherently modular and can easily work in parallel. Consequently, the actuator load capacity can be increased without losing high acceleration performance. All of this new class of normal-stress motors are centered on a canonical ultra fast motor design (CUFM), which is the building block for subsequently introduced ultra fast linear motors and ultra fast rotary motors. In this chapter, we present only the electromechanical design of the proposed motors. Their detailed analysis of force generation and material selection issues is presented in Chapter 3.

2.1 Electromagnetic Motor Fundamentals

In this section, we examine two kinds of electromagnetic actuators, which are candidates for high-acceleration positioning applications. These are moving coil motors, and normal-stress motors. After discussing their basic working principles, we illustrate fundamental differences in their acceleration capabilities.

2.1.1 Moving Coil Motors

Moving coil motors work on Lorentz forces generated by an external flux field on current carrying coils, and are also sometimes referred as a voice coil motor (VCM). Typical applications for VCM include hard disk drives and speakers. The actuating principle for such motors is shown in Figure 2-1. Here a conductor of mass density ρ and carrying current density J is immersed in an external field with flux density B. Ignoring additional inertial loads, the resulting acceleration of a moving coil can be

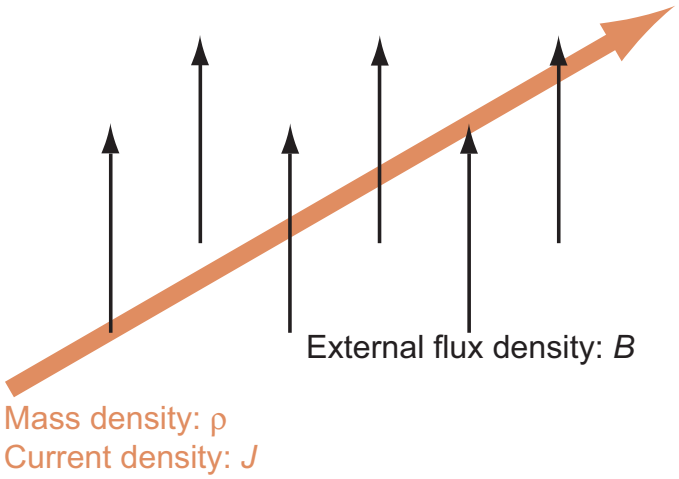


Figure 2-1: The force generation on a moving conductor. The current flows perpendicularly to the external field.

calculated as in [31]:

$$a = \frac{F}{M} = \frac{JBA_cL}{\rho A_cL} = \frac{JB}{\rho}, \tag{2.1}$$

where A_c is the cross-sectional area of the current carrying conductor and L is the conductor length.

Using this result, Table 2.1 shows the achievable acceleration for coils made of both copper and aluminum materials. Here, we assume a 3 A/mm^2 current carrying capacity of copper, which is an empirically-observed limit for the air-cooled case. For aluminum, by assuming the same allowable heat density, according to the respective electrical resistivity ρ_e, we can calculate that the aluminum current carrying capacity is 2.4 A/mm^2. Due to its lower mass density, a VCM using an aluminum coil has better acceleration performance, despite its lower allowable current density. If thermal design allows higher current densities, then the achievable acceleration will scale accordingly.

Table 2.1: Achievable acceleration of moving coil motors.

coil	ρ [gm/cc]	J [A/mm^2]	ρ_e [nΩ-m]	B [T]	a [G]
copper Cu	8.6	3	17.1	1.2	42.7
aluminum Al	2.7	2.4	26.5	1.2	109

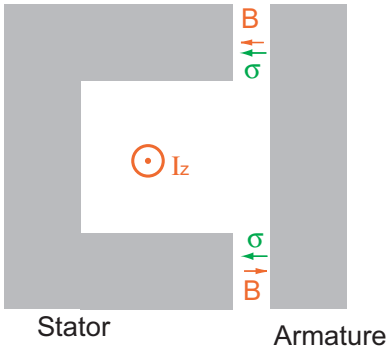


Figure 2-2: A schematic solenoid design.

2.1.2 Normal-Stress motors

Figure 2-2 shows a schematic configuration for a solenoid, which is a commonly used normal-stress motor. The excitation current I_z located inside the slot of the stator core generates magnetic flux density B in the air gap between the armature and stator. This magnetic flux density B generates a normal stress σ at the top and bottom pole face of the armature.

The achievable acceleration of normal stress motors can be illustrated by an armature shown in Figure 2-3. The armature disk has mass density ρ and thickness t. If we can apply flux density B in the normal direction of the armature face (this can be achieved by a proper electromagnetic configuration as discussed later), according to the Maxwell stress tensor theory [86], the generated force density f (or pressure) is calculated as

$$f = \frac{B^2}{2\mu_0}. \tag{2.2}$$

Consequently, the acceleration is

$$\begin{aligned} a &= \frac{F}{m} \\ &= \frac{f}{\rho t} \\ &= \frac{B^2}{2\mu_0 \rho t}. \end{aligned} \tag{2.3}$$

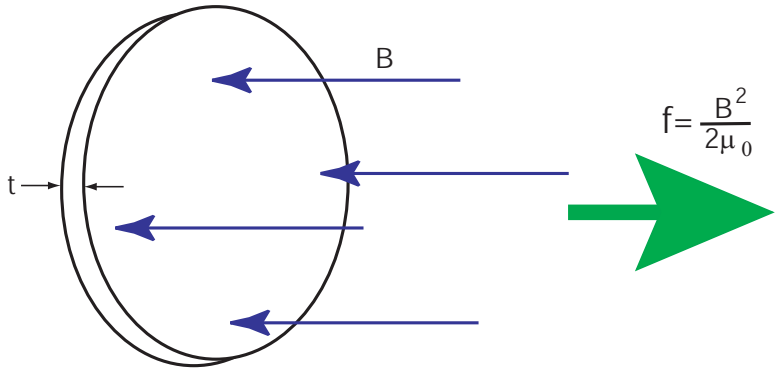


Figure 2-3: Force generation on an armature disk.

Note that the acceleration is inversely proportional to t. A small t is the key to achieving high acceleration. If we set B=1.5 T, iron mass density $\rho = 7.8$ gm/cc, and t=3 mm, the resulting acceleration will be 4000 G. This is much higher than that achievable with moving coil motors as calculated before.

Next, we examine prior-art designs for normal-stress motors and then present our ultra fast motor design.

2.2 Prior Art Normal-Stress Electromagnetic Actuators

In this section, we look at key design features of various normal-stress actuators in the literature. Typical normal-stress actuators include solenoids, magnetic bearings, and other short-stroke actuators for applications such as fast steering mirrors and electric engravers.

2.2.1 Solenoid

Figure 2-4 shows a solenoid design with a "U-I" configuration. An "I" shaped armature is located adjacent to the poles of a "U" shaped stator with an excitation coil in the slots. The current flowing through this coil can excite normal-flux on the armature and thereby generates normal stress on the armature surface. However, in this configuration only the ends of the armature contributes to force generation.

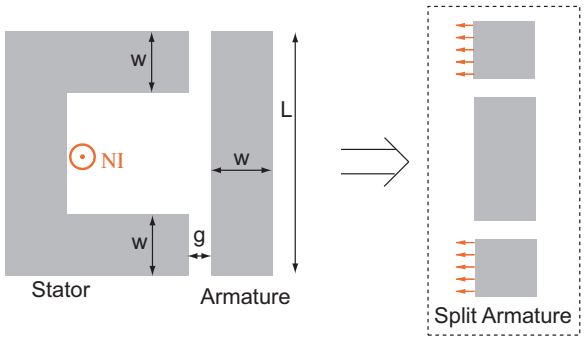


Figure 2-4: Force distribution on a "U-I" solenoid.

By splitting the armature into three parts as shown in the right of Figure 2-4, we can see that only the top and bottom parts of the armature have a force-generating surface and the middle part contributes no force but does add inertial load. Roughly speaking, for typical slot dimension only 50% of the armature surface area can be utilized in this solenoid design, assuming that $L = 4W$ as shown in the figure. With this observation, the maximum actuating force F_m and the resulting acceleration a_m can be calculated as

$$\begin{aligned} F_m &= \frac{B_{sat}^2}{2\mu_0} A \\ &= \frac{B_{sat}^2}{2\mu_0} 2WD; \end{aligned} \tag{2.4}$$

$$\begin{aligned} a_m &= \frac{F_m}{m} \\ &= \frac{B_{sat}^2}{2\mu_0} \frac{2WD}{4W^2 D\rho} \\ &= \frac{B_{sat}^2}{4W\mu_0\rho}, \end{aligned} \tag{2.5}$$

where B_{sat} is the saturation flux of the armature material, D is the depth into the paper, and ρ is the armature mass density. These equations show that the acceleration a_m is inversely proportional to armature width W while the force F_m is proportional to W. There exists a dilemma in that high acceleration is achieved at the sacrifice of actuating force. Higher acceleration requirement dictates a smaller W, and thus lighter armatures. However, this also results in smaller armature area and thus less actuating force. In reality, there also exist disturbance forces such as cutting forces in

fast tool servos, and additional inertial loads attached to the armature. Taking into account these disturbance and inertial forces, it is necessary to maintain a certain minimum of force capacity. In order to break this trade off relation in the "U-I" solenoid design, multipolar solenoids are presented in the next section.

2.2.2 Multi-pole Solenoids

Figure 2-5 shows a comparison between a "U-I" solenoid as discussed above to a "E-I" solenoid with three poles. The later can be taken as the most primitive multi-polar

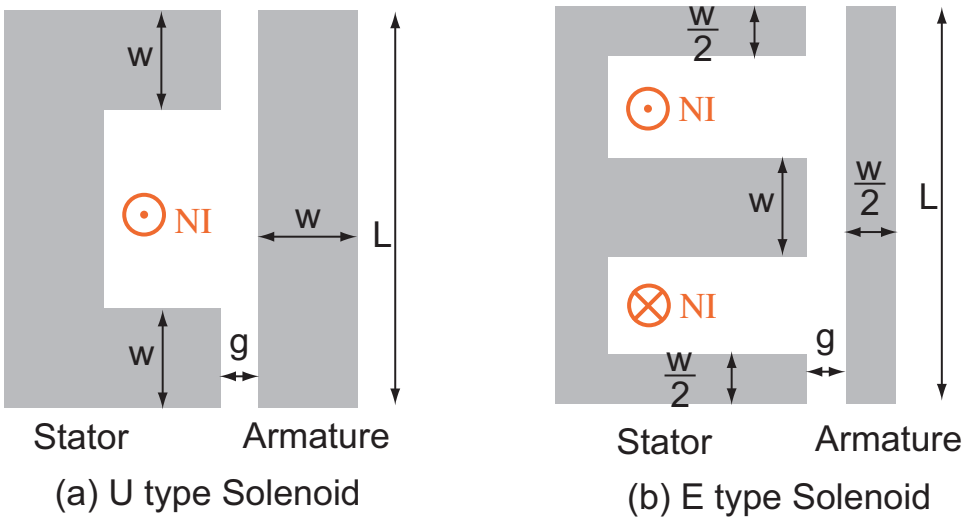


Figure 2-5: Comparison between "U-I" and "E-I" solenoids.

solenoid design. By changing the pole number of the stator core from 2 to 3, the armature width can be reduced without reducing the actuating force F_m. As a result, the armature mass is cut in half and the maximum acceleration is doubled. Another advantage of an "E-I" design is that one excitation coil winding can be used twice as opposed to once in a "U-I" design, but this feature is not our main concern in designing actuators for high acceleration applications.

Another way to interpret this difference in acceleration capability is to think about the "E-I" design in Figure 2-5(b) as two parallel-working "U-I" solenoids, the armature thickness of which is half that of the "U-I" solenoid in Figure 2-5(a). Each half-sized "U-I" solenoid achieves twice the acceleration of the "U-I" solenoid in fig-

ure (a) with half the force capacity. When two of these are working in parallel, the total force capacity of the actuator is doubled with the same acceleration as for each half-sized solenoid. This idea can be summarized as: **multiple small actuators working in parallel can increase acceleration without losing force.** As more slots are added, the armature thickness can be reduced correspondingly, thereby reducing the armature mass.

Using this idea more aggressively, Schechter designed a multi-pole "E" solenoid shown in Figure 2-6 [73], [87]. There exist other similar multipolar solenoid designs

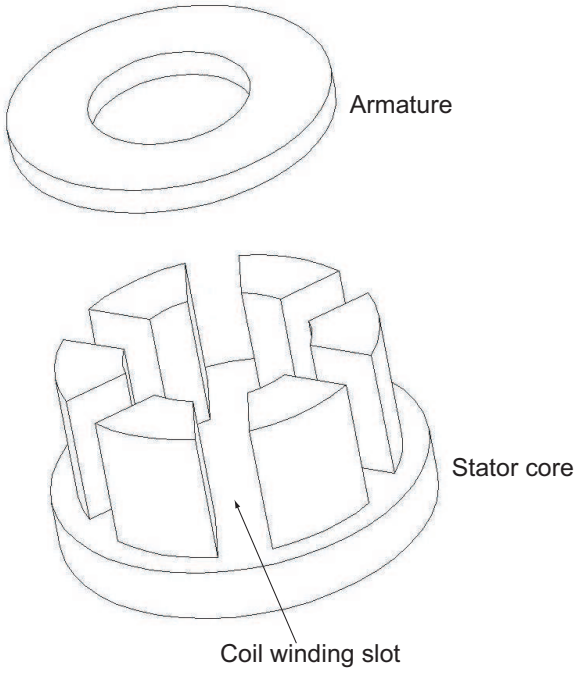


Figure 2-6: Multiple E-type solenoids.

[75, 76, 60]. Most of these actuators were designed for in-cylinder fuel injection control valves in automotive industry. All these designs can be explained as multiple small actuators working in parallel, and consequently acceleration can be greatly increased.

2.2.3 Nonlinear Magnetic Bearings

Solenoids as discussed above are generally used as binary actuators without position feedback. Many active magnetic bearings have similar electromagnetic configurations as the "U-I" or "E-I" solenoids, but the armatures in magnetic bearings are regulated

with a feedback position control system. Although most magnetic bearings are designed for holding armatures at a fixed position instead of high-acceleration trajectory tracking, magnetic bearings represent the majority of electromagnetic configurations using normal stress actuator in the literature. In particular, precision magnetic bearings have the same problems of nonlinearity and accurate-positioning as normal-stress actuators used for fast tool servos. Essentially, magnetic bearings are nonlinear actuators. The actuating force of an "U-I" or "E-I" design is proportional to the excitation current squared and inversely proportional to the air gap squared. In this section, we look at nonlinearities of magnetic bearings and corresponding linearizing methods, which are key issues in normal-stress actuator control.

Current-Biased Linearization

One widely-used solution to this nonlinearity issue is the current-biasing linearization method. As illustrated in Figure 2-7, a pair of "U-I" actuators are located on opposite faces of the armature.

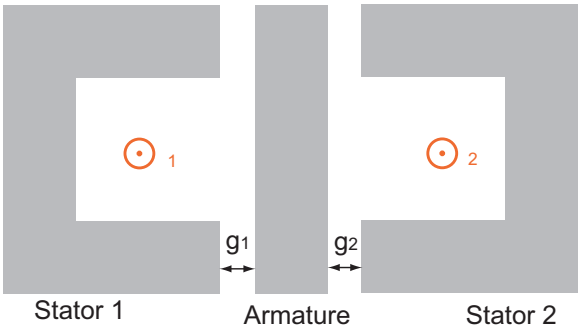


Figure 2-7: Current biased magnetic bearing.

The total actuating force is the difference between the two solenoids and can be expressed as

$$F = \frac{A\mu_0 I_1^2}{8g_1^2} - \frac{A\mu_0 I_2^2}{8g_2^2}, \tag{2.6}$$

where A is the armature force generating area, I_1 and I_2 are the excitation currents in stator 1 and stator 2 respectively, and g_1 and g_2 are the stator/armature air gaps.

When the armature is at the middle position, we have $g_1 = g_2 \equiv g_0$. The actuating force can be simplified at this position as

$$F = \frac{A\mu_0}{8g_0^2}(I_1 - I_2)(I_1 + I_2). \tag{2.7}$$

By defining $\overline{I} = (I_1 + I_2)/2$ and $\tilde{I} = (I_1 - I_2)/2$, we can express the excitation currents as $I_1 = \overline{I} + \tilde{I}$ and $I_2 = \overline{I} - \tilde{I}$. As a result, the force becomes

$$F = \frac{A\mu_0}{2g_0^2}\overline{I}\tilde{I}. \tag{2.8}$$

If we keep $\overline{I}$ constant and using $\tilde{I}$ as a control input, the force characteristic becomes linear with the virtual control input $\tilde{I}$ in this centered configuration. As only addition/subtraction operations are used in this transform, this current-biasing based control technique can be easily implemented with an analog circuit.

However, this linearization technique works well only when the armature is positioned near the center with $g_1 \simeq g_2$. Since excitation flux varies with both current and air gaps, the resulting force is not linear with either current or position when $g_1 \neq g_2$. In many magnetic bearing applications requiring only fixed position holding, current biasing is a good solution. When the armature is required to move over a large excursion, the uncancelled second order terms in force will introduce significant nonlinearity. For applications with precision control and/or high frequency tracking requirement, this nonlinearity is a problem. Next, we discuss the technique of linearization via position feedback, which is one way to solve this problem.

Position Feedback Linearization

To maintain good linearity over large positioning range, nonlinear compensation methods using a position feedback signal have been frequently used, for example in [28, 82].

This feedback linearization idea can be explained in many ways, for instance by a nonlinear coordinate transform [82]. Here we illustrate this idea by the block diagram

shown in Figure 2-8. A "U-I" solenoid is modelled as a nonlinear gain block $K(x)$

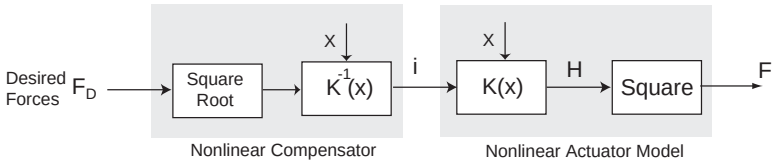


Figure 2-8: Feedback linearization of a solenoid-type magnetic bearing.

serially followed by a square operator block, where x is the air gap. The function $K(x)$ represents the relation of excitation current to magnetic field in the air gap, and the square block relates the magnetic field to the actuating force. If we neglect the leakage flux, $K(x)$ can be expressed as

$$K(x) \quad = \quad \frac{K_0}{2x}, \tag{2.9}$$

where K_0 is a constant related to actuator dimensions. In order to get a linear relation, a straightforward way is **to invert the actuator model block by block** as shown by the nonlinear compensator in Figure 2-8. A $K(x)^{-1}$ block and a "square-root" block in the compensator are used to invert the $K(x)$ block and the "square" block in the actuator model respectively. As a result, the relation of a desired force F_D to the actuating force F becomes linear. Note also that this result is independent of the dynamics from F to x in the mechanical system.

The shortcoming of this feedback linearization technique is also evident in this block diagram. The effectiveness of this technique is heavily dependent on how accurate our actuator model is. If there exist significant modelling errors in $K(x)$, then the resulting compensator block $K(x)^{-1}$ will not invert the actuator well. When applying this technique, an experimentally calibrated curve of $K(x)$ is usually used instead of the relation (2.9). Moreover, this nonlinear compensator can be easily implemented only with the aid of a digital control computer, and its sampling rate also greatly influences the performance as discussed in [82]. Today, with real-time computers capable of 1 MHz sampling rate, this linearizing idea can achieve pretty good results if eddy currents and hysteresis in the system are negligible.

2.2.4 Flux-Biased Magnetic Bearing

The current-biased linearization and the feedback linearization techniques discussed previously achieve a linear operation characteristic of magnetic bearing in which I will term "soft" ways: they are software-based control algorithms/transforms without changing the nonlinearity of magnetic bearings themselves. This sections discusses hard-linearized normal-stress actuators, which have essentially linear operation characteristics by proper electromagnetic configuration designs.

Studer's Design

Studer in 1977 presented a magnetic bearing design using a permanent magnet to generate biasing flux between two stators [78] as shown schematically in Figure 2-9. In the middle of this figure, a permanent magnet is sandwiched between two stators. A ring-shaped armature is located outside of the stators. Between the stators and the armature there are air gaps g_1 and g_2 on the left and the right, respectively. Between two stator slots are excitation coil windings, and the reference direction of currents is pointing out of the paper. Correspondingly, these excitation coils generate AC flux with the reference direction shown in the figure. The permanent magnet generates bias DC fluxes as shown in blue lines on both sides. According to the labelled flux directions with coil energized, the fluxes on the left pole surfaces of the armature are enhanced and those on the right are reduced. Consequently, a net force pointing to the right results on the armature. When the excitation current changes direction, so does the actuating force.

From this design, we can see the key features of such hard-linearized normal-stress actuators:

- **Invariant Air Gap**. The total length of all air gaps are invariant with the armature position. When the air gap g_1 decreases, the air gap g_2 increases, and vice versa. For any armature position, the sum of g_1 and g_2 is a constant. As a result, the current-generated excitation AC flux is independent of the armature position.

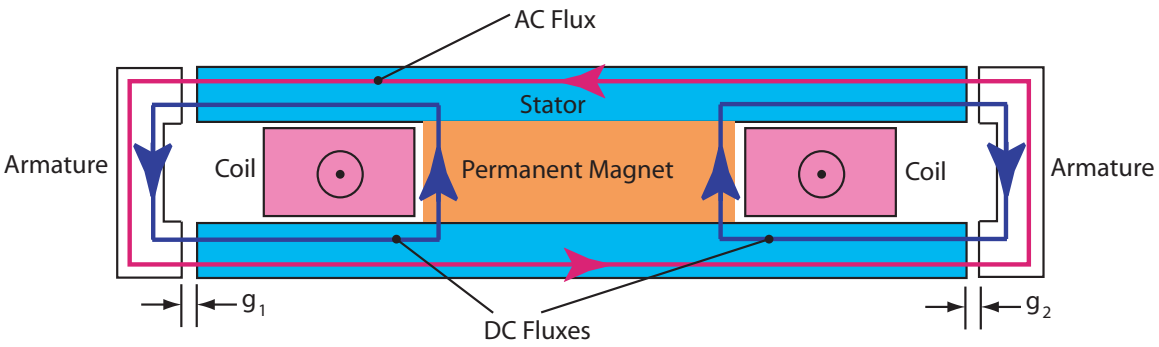


Figure 2-9: Flux-biased magnetic bearing using permanent magnet, of configuration introduced by Studer [78].

- **Invariant Bias Flux**. The permanent magnet generates DC fluxes on both sides. When the armature moves, these bias fluxes may change significantly. However, the total DC fluxes are invariant with the armature position, because the permanent magnet reluctance is much larger than reluctances in the air gaps.

- **Differential Excitation**. The excitation coil winding generates differential fluxes on both sides of the armature: one side of the AC flux sums with the DC flux and the other side the AC flux subtracts the DC flux.

- **Natural Subtraction**. The total force on the armature is the substraction between the normal stress on left pole surface and that on the right, as these two forces opposes to each other.

When a normal-stress actuator shares all of these key design features, we call it a hard-linearized normal-stress actuator. All these features combine to make the actuator generate an actuating force linear in the excitation current and in the armature position. Next, we will look at a couple of more designs in this actuator family.

Planar Magnetic Bearing

Molenaar developed a linear active magnetic bearing with both permanent magnets and electro-magnets [53, 54], as shown in an adapted drawing in Figure 2-10.

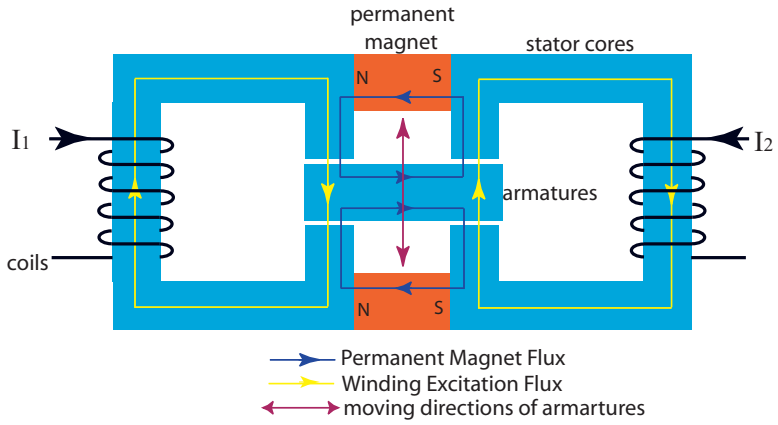


Figure 2-10: Principle of planer magnetic bearing, adapted from [53].

Two C-shaped stator cores surround an armature at the left and right sides, respectively. Inside the two stator cores are excitation coil windings. Two permanent magnets are placed between these two stator cores to generate bias fluxes.

Let's focus on the left arm of the planer magnetic bearing to examine the key features summarized previously. (1) The total air gap length is a constant. When the armature moves up, the top air gap decreases, and the bottom air gap increases by the same amount. (2) The total bias flux flowing from the permanent magnet into the armature bar is constant, as the permanent magnet has much larger reluctance than the air gaps. (3) The excitation current I_1 generates differential fluxes on the top normal surface and the bottom normal surface of the armature. (4) The net suspension force generated by the left arm stator is the subtraction between the normal force on the top surface of the armature and the normal force at the bottom. As a result, the actuating force is a linear function of the excitation current I_1 and the armature position. Similar analysis can be carried out with respect to the right arm stator.

When two coil current I_1 and I_2 have equal values, a levitation force can be generated to control the rotor in the translation motion in the vertical direction. Molenaar also pointed out another advantage of this non-coplanar configuration: the possibility of apply a torque to the rotor. When differential current are fed into two coils, differential forces will be generated on two ends of the rotor. This torque can

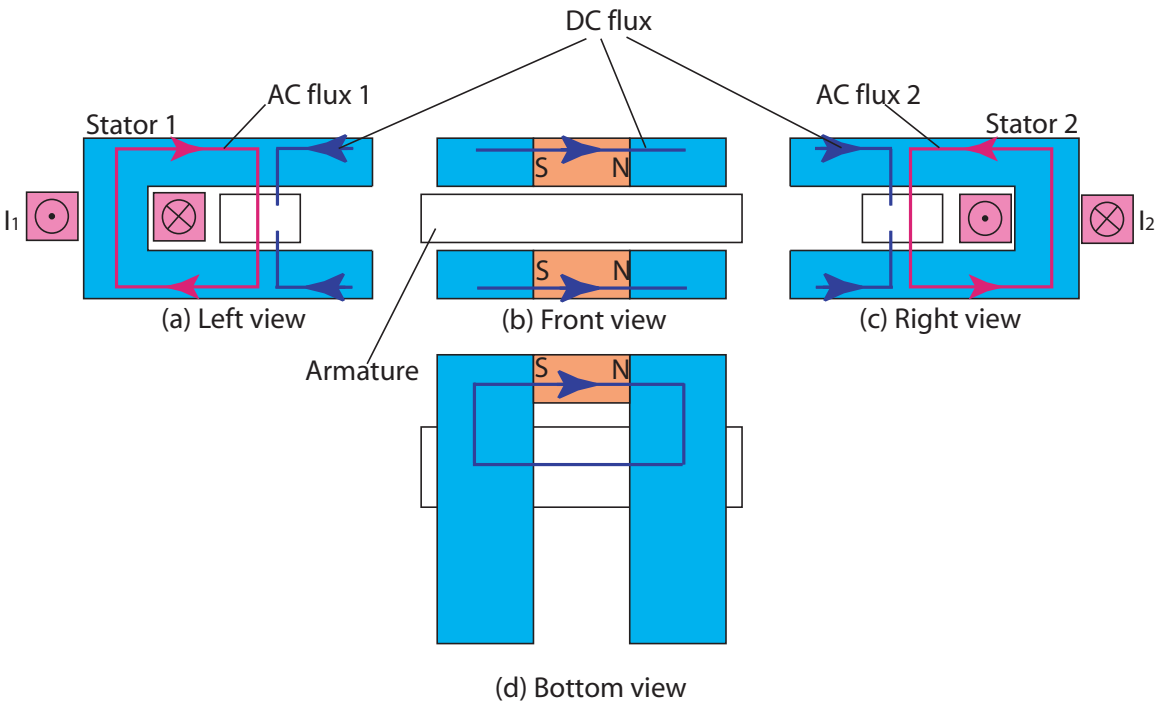


Figure 2-11: Flux-biased I-U suspension, adapted from [34]. Moving armature is outlined in black.

be used to control an additional degree of freedom of the rotor.

In Molenaar's planar magnetic bearing, another set of coils are used to translate the armature laterally (in propulsion direction), but not shown here. These propulsion coils can either be installed on armature (moving member) [54] or on the stator cores [26].

U-I Suspension

Similar to the planar magnetic bearing, another magnetic bearing design using permanent magnet to achieve linearity is the I-U suspension as shown in an adapted drawing in Figure 2-11 [34]. This I-U suspension is composed of U-shaped stator 1 and stator 2, and an I-shaped armature core. Permanent magnets are placed between these two stator cores. Inside each stator slot, there is an excitation coil winding, labelled as I_1 and I_2. The force generation can be analyzed in the same way as in the above planar magnetic bearing.

2.2.5 Flux-Biased Fast Motors

We have examined several flux-biased magnetic bearing designs, which are in the family of hard-linearized actuators. In this subsection, we will discuss some designs in this actuator family, which can be used for fast tool servo applications.

Flux-Biased Rotary Electric Engravers

Back in 1937, a rotary motor with flux-biasing was designed for printing plate engravers [40] as shown in Figure 2-12 [2].

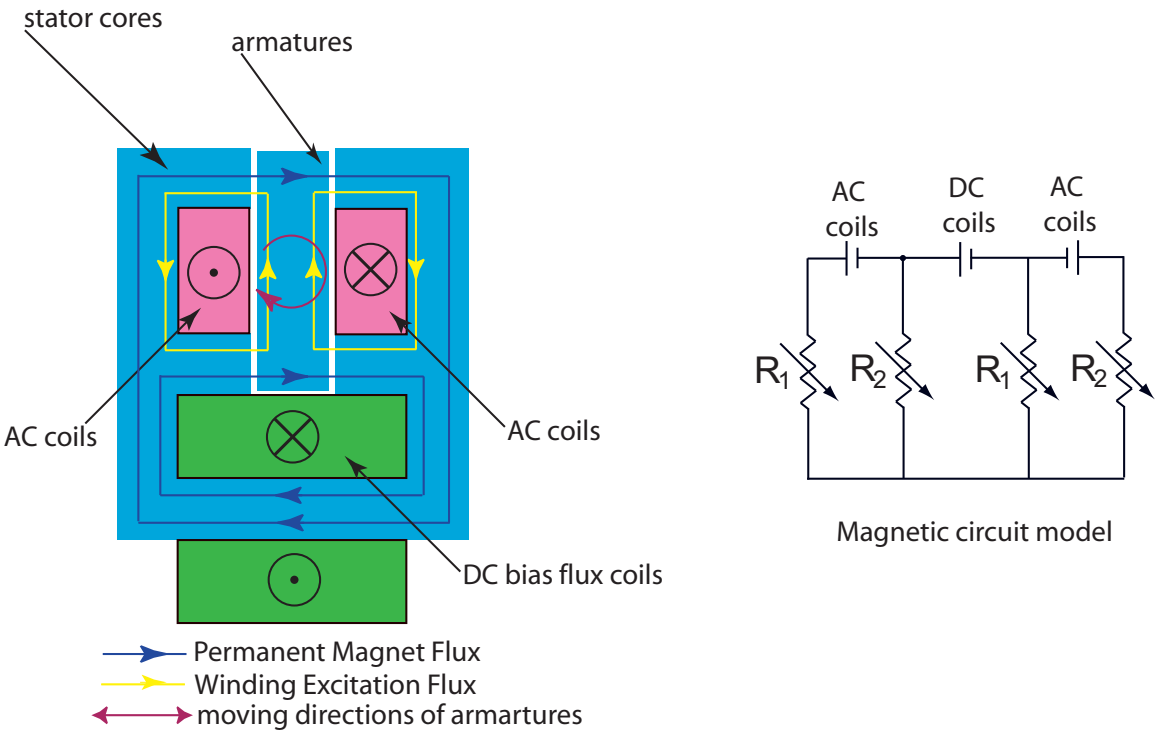


Figure 2-12: Flux biased fast rotary motor for engraver, as utilized in [40], along with magnetic circuit model. R_1 and R_2 are a differential pair of reluctances varying with armature rotation angle, with the constraint that $R_1 + R_2$ is constant.

In this design, there are a pair of C-shaped stator cores located on both sides of the armature. The armature rotates its geometric center, which is fixed by supporting bearings, typically flexure, not shown here. Inside the stator core slot is the excitation coil. These two stator cores are connected by a ferromagnetic core, around which is a

[2]Thanks to Rick Montesanti for digging up this reference and bringing it to our attention.

secondary DC coil to generate biasing fluxes. In this design the total air gap length is constant if the armature center is well supported and thus has no translation motion. However, the bias flux is not constant if the DC flux coil current is not regulated properly. According to the magnetic circuit model in Figure 2-12, the DC coil has almost zero reluctance, and the air gap reluctances R_1 and R_2 are strongly dependent on the armature angle position. If the magnetomotive force (the DC coil current) is constant, the generated DC bias flux will be strongly position-dependant. Hence, strictly speaking, this design is not an ideal hard-linearized normal-stress motor. In small-stroke applications where the armature excursion is much smaller than the air gap length, the resulting DC flux change can be ignored and the actuating toque can be approximated as a linear function of exciting AC current in the AC coils.

Montesanti and Trumper designed a fast tool servo with similar electromagnetic configuration [57], except that air gaps are intentionally added into the core connecting two C cores. As a result, the reluctance of the DC coil is increased and thus the DC flux change with position can be reduced. However, this increased reluctance of the DC coil is not comparable to the reluctance of the permanent magnet.

Flux Steering Motors

During the course of this research, Prof. Trumper and Richard Montesanti investigated a series of high-acceleration fast motors for fast tool servo applications. All these designs belong to the hard-linearized normal-stress motor family, as they have all the key features that we summarized in Section 2.2.4.

Figure 2-13(a) shows a representative design. Two C-shaped stator cores surround an armature at both the top and the bottom. Inside the two stator cores are excitation coil windings. Two permanent magnets are placed between these two stator cores to generate bias fluxes. More details of this design can be found in Richard Montesanti's thesis [59]. This configuration is very similar to the planar bearing in Figure 2-10, but these two design are targeted for different applications. If the top stator current flows out of the paper and the bottom stator current flows into the paper with equal amplitude as shown in the figure, an actuating torque can be generated in

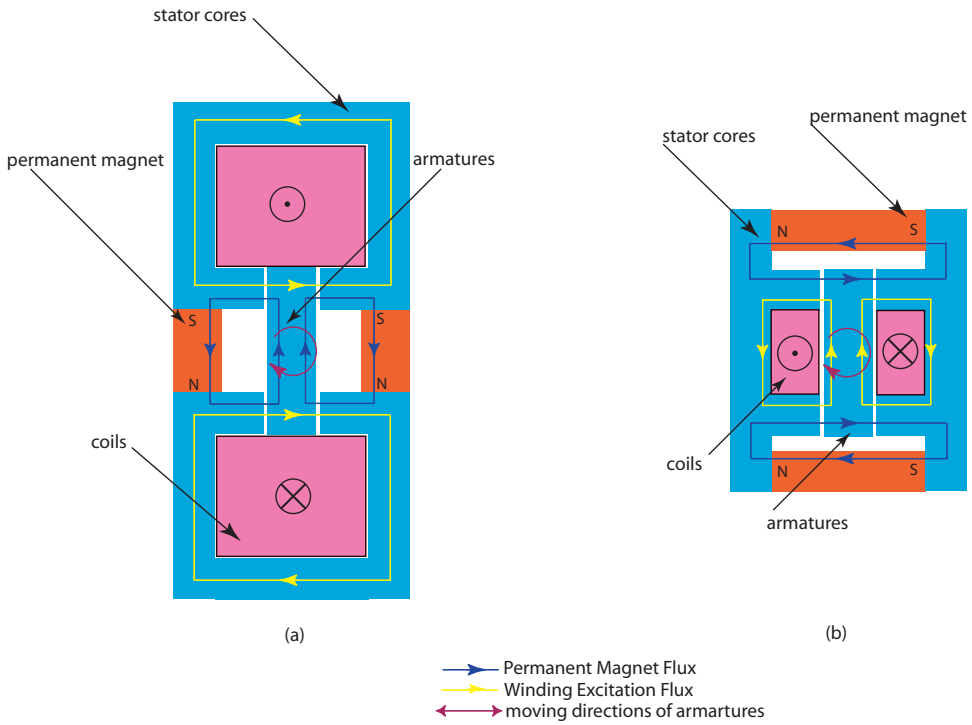


Figure 2-13: Flux biased fast rotary motor for fast tool servos.

the clockwise direction.

Generally, this motor design can generate both a rotary toque (proportional to differential mode current component, as shown in Figure 2-13(a)) and a linear force (proportional to common mode current component, as shown in Figure 2-14). Therefore, this motor can be operated in linear and/or rotary modes.

Trumper and Montesanti also developed a variation design shown in Figure 2-13(b). This design is similar to the flux-biased electric engraver shown in Figure 2-12, except that a permanent magnet is used to generate bias flux instead of a DC coil. However, this design does not have the linear operation characteristic in both the linear motion and the rotary motion as the design in Figure 2-13(a).

Armature-Armature Flux-Biased Fast Motors

All hard-linearized designs discussed until now are characterized by stator-stator biasing design, in which permanent magnets or DC coils are placed between two stators. If moving the permanent magnet from the stator into the armature as shown in Fig-

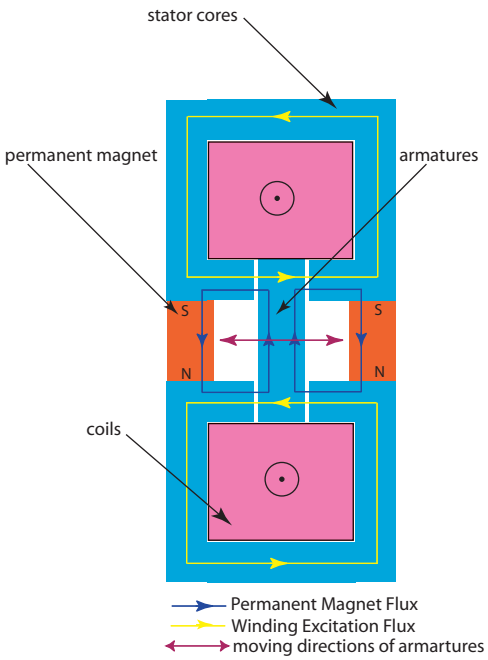


Figure 2-14: Flux biased fast linear motor for fast tool servos.

ure 2-15, we can get another design variation: armature-armature flux-biased design. The moving armature is the combination of a permanent magnet and soft magnetic cores. This design has all the key features summarized for a hard-linearized normal-stress actuator. With two excitation currents in the stator cores, both a linear force and a linear torque can be generated, similar to the design of Figure 2-14.

Actually, a rotary version of this armature-armature flux-biased motor has already existed in the literature [88] as shown in Figure 2-16. Multiple of this type of motors were used in a biaxial fast steering mirror for precision optical pointing systems.

2.2.6 Other Related Motor Designs

In the literature, there exist many actuators using permanent magnets to generate bias fluxes. However, if they lack any of the key features that we summarized for hard-linearized normal-stress actuators in section 2.2.4, most likely they don't have globally linear operation characteristics. In this subsection, we will look at several actuator designs operating in very different ways from hard-linearized normal-stress actuators, although they have both a permanent magnet and an excitation coil.

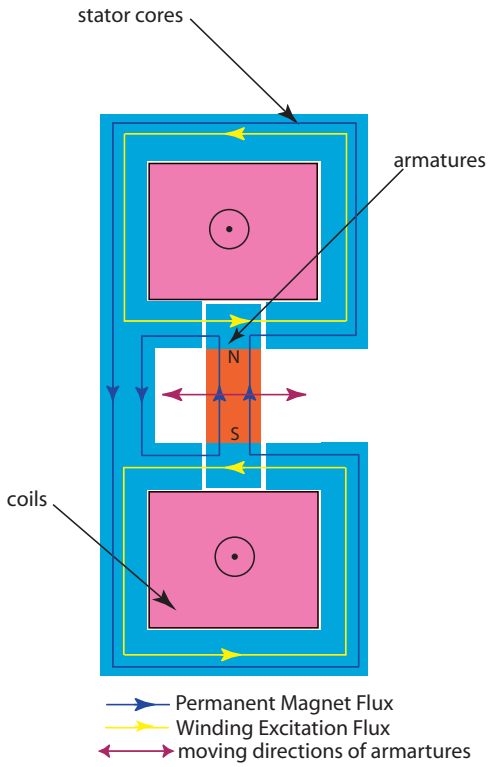


Figure 2-15: Armature-armature biased fast linear/rotary motor.

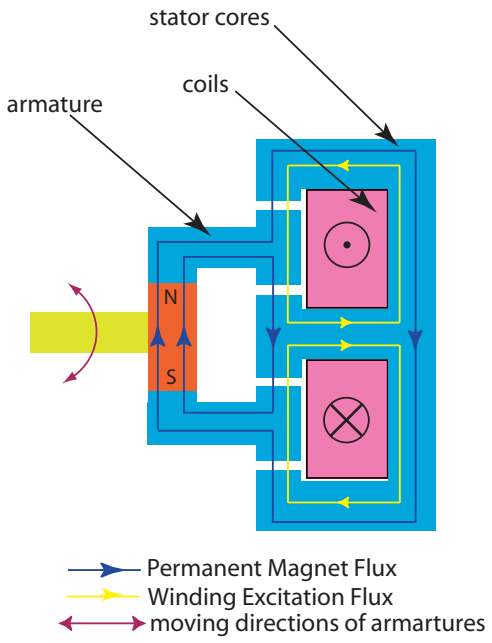


Figure 2-16: Armature-armature flux-biased fast rotary motor of a configuration similar to [88].

Flux-Biased Nonlinear Magnetic Bearing

An electromagnet design for DC operation is shown in Figure 2-17 [25, 11]. A perma-

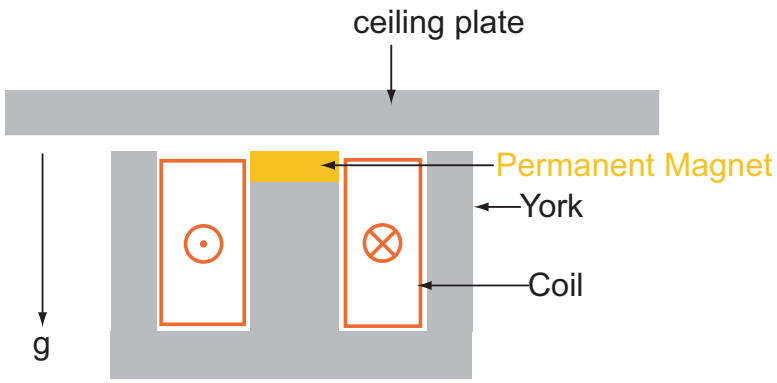


Figure 2-17: Flux biased magnetic bearing with weight balancing.

nent magnet is put into the middle leg of an E-shaped stator to serve two functions. First, a bias flux is generated and a levitating force is created to help balance the weight even when the excitation current is zero. As a result, most of the weight is balanced by the permanent magnet flux and the required excitation coil current can be reduced in comparison to a design without an embedded permanent magnet. Second, the nonlinearity of the operation characteristic is attenuated by the additional reluctance associated with the permanent magnet. However, these benefits are achieved at the sacrifice of dynamic force capability, as the response from the excitation current to the magnetic field is reduced due to the increased air gap. Obviously, this design does not have any feature summarized for hard-linearized normal-flux actuators. Therefore, this actuator does not have a linear relationship between the excitation current and the actuating force.

Flux-Biased Radial Magnetic Bearing

Figure 2-18 shows a flux-biased magnetic bearing design with permanent magnets between the stators [37]. This is proposed as a power-efficient replacement to current-based magnetic bearings. However, this design does not have properties such as invariant air gap and differential excitation. Therefore, this magnetic bearing is not in the family of hard-linearized normal-stress actuators.

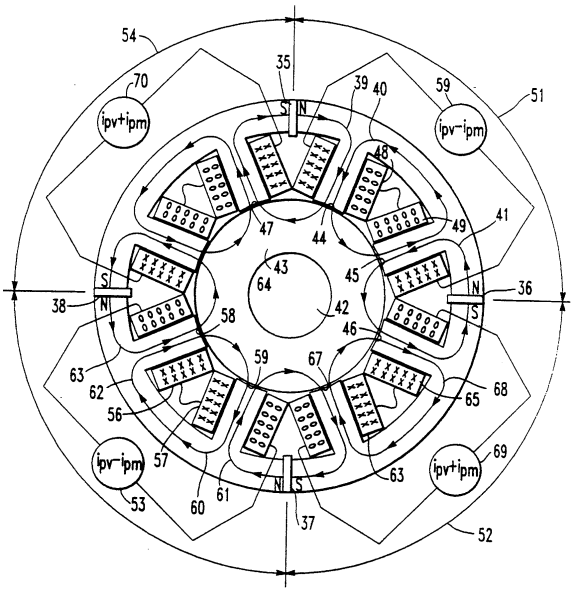


Figure 2-18: Flux-biased radial magnetic bearing with permanent magnets placed between stators. This figure is copied from *US Patent 5355042* [37].

Figure 2-19 shows another flux-biased radial magnetic bearing design, in which permanent magnets are placed between a stator and a rotor [32]. In comparison with a current-biased magnetic bearing, this design is more power efficient. Moreover, this design introduces two current-free stator poles made of permanent magnets on the top of the rotor. As a result, the rotor weight can be partially balanced by the normal forces on these poles. However, this design does not have the features of invariant air gaps and differential flux excitation. As a result, this magnetic bearing is fundamentally different from the hard-linearized normal-stress actuators.

Flux-Biased Shear-Stress Motor

Figure 2-20 shows a rotary actuator design developed by Montagu [55] at General Scanning Inc[3]. This design has a pair of stators located on the top and bottom side of a salient-pole armature. Two permanent magnets are placed between these two stators to generate bias flux. This design appears to have some common features with the flux-biased fast rotary motor in Figure 2-13(a). However, these two actuators operate

[3]The company name has since been changed to GSI Lumonics Inc.

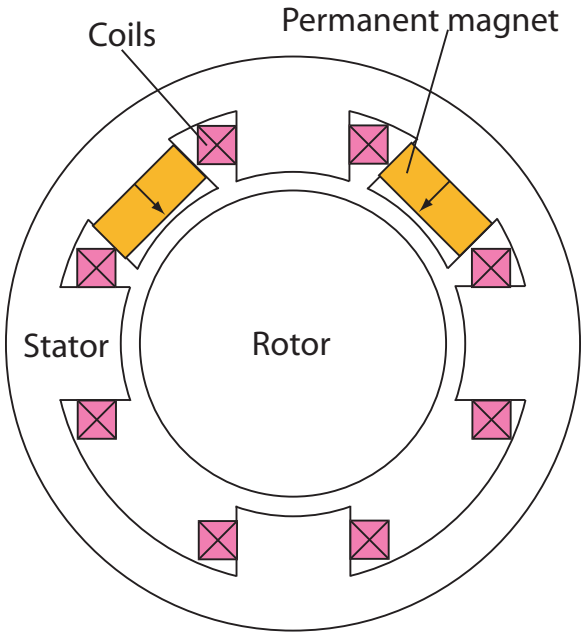


Figure 2-19: Flux-biased radial magnetic bearing with permanent magnets placed between a stator and a rotor.

in completely different ways. This design is a shear-stress motor with relatively large stroke. The actuating torque is a linear function of excitation current, but is a nonlinear function of the rotation angle.

2.2.7 Long-Stroke Solenoids

This subsection discusses two long-stroke solenoid designs. Although long-stroke actuators are not used for fast tool servo applications, there exist some good design ideas that can be adapted to short stroke applications. These long-stroke actuators are mainly designed for applications such as electronic valve controls for camless engines. In contrast with the fuel injection application that needs a stroke of about 100 μm, intake-valve control requires a stroke around 5−−8 mm. For such an air-gap, the current for saturation-level normal flux driving is not feasible, because the current winding window area should be at least 50 times larger than that of a solenoid with 100 μm storke if an conventional solenoid design is used.

To solve this problem, Pischinger presented a bistable actuator design, in which a double-spring preloaded armature is located between a push-pull solenoid pair as

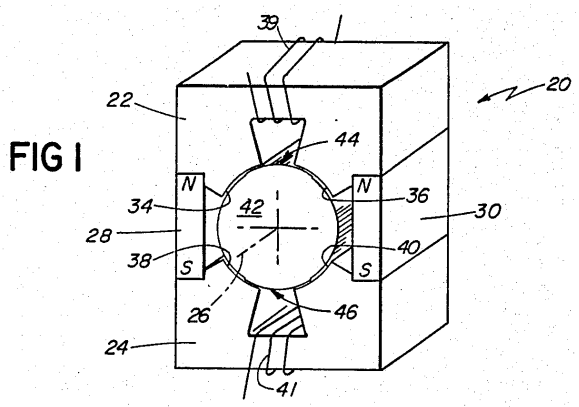


Figure 2-20: Flux-biased shear-stress motor. This figure is copied From *US Patent 4528533* [55].

shown in Figure 2-21 [68].

As shown in the figure, initially the top spring 16 is unloaded and the armature is pushed to its top end position by the bottom spring 17. In order to pre-load the top spring 16 and thus move the neutral point to the center of two solenoid poles, both the top solenoid coil 12 and a compressing coil 14 are energized to compress the top spring 16. After this start-up process, the armature is preloaded by a pair of springs and behaves like a pendulum. In order for the armature to move to the bottom-end position, the top solenoid is de-energized first, and then the armature swings from the top end to the bottom end in half of its pendulum natural period. Once the armature is near the bottom end, the bottom solenoid is energized and further holds the armature at this position. When the armature needs to go to the top end position, the bottom solenoid is de-energized and thus the armature swings back to the top end, where it is held by the energized top solenoid. During the whole process, the spring provides the actuating force for the armature to travel from one position to the other, and the solenoids work only when the armature is in the vicinity to its pole face. This design illustrates the fundamentals of many later electric valve control systems.

In Pischinger's design, holding the armature needs excitation current in the solenoids and consumes energy. To further remove this energy consumption associated with position holding, Lequesne described a bistable actuator design shown in Figure 2-22

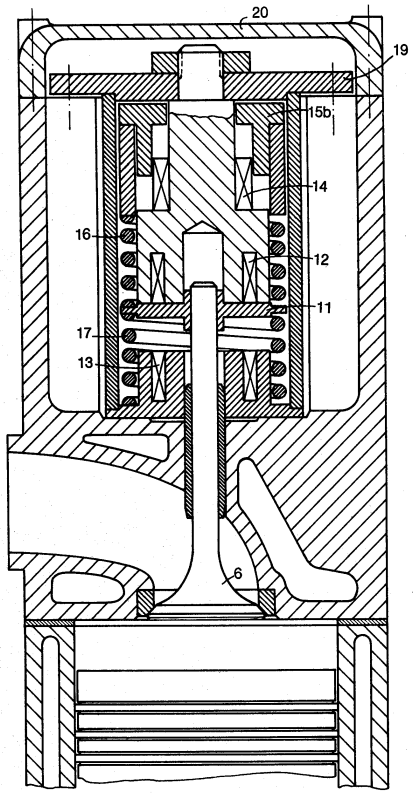


Figure 2-21: Bistable long-stroke solenoid. This figure is copied from *US Patent 4 455 543* [68].

[35]. By putting a permanent magnet between an armature and an stator, a magnetic latch is created. "Holding the moving mechanism in either extreme position is achieved by the permanent magnet instead of the holding current. Motion is initiated by reducing the magnetic force to a level that is below the level of the spring force. This is achieved by inducing a current in the coil located next to the closed air gap to produce a flux opposing the permanent magnet" [36]. Even though this design is for long-stroke applications, it is one of the closest prior-art designs to the ultra fast motor design that I present in next sections.

2.3 Canonical Ultra Fast Motor

All the designs (except for Lequesne's bistable long-stroke actuator in Figure 2-22) reviewed previously have a common problem: only a portion of the armature area is for

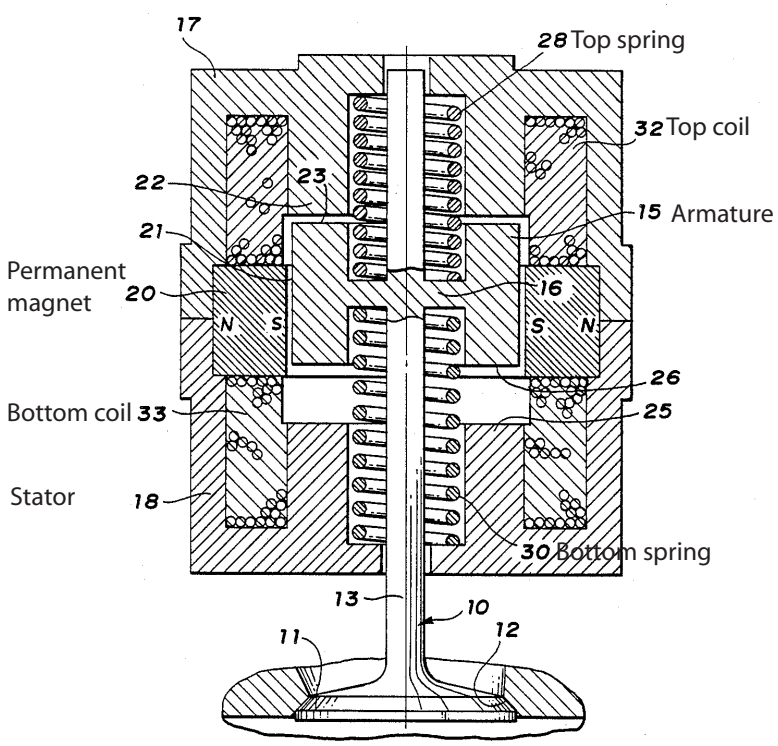


Figure 2-22: Bistable long-stroke actuator using permanent magnet. This figure is adapted from *US Patent 4,779,582* [35].

force generation, and thus these designs cannot realize the full potential acceleration in the normal-stress actuator model shown in Fig 2-3. To achieve this goal, here we present a canonical ultra fast motor design (CUFM) shown in Figure 2-23, which is the core electromagnetic design in this thesis and is also the building block of several possible designs to be described later. By "ultra fast", we mean that almost all the armature area contributes to force generation, and this design feature differentiates our design from other hard-linearized normal-stress motors reviewed previously.

The CUFM in Figure 2-23 comprises an armature, a stator core, a permanent magnet, and excitation coil windings. Although not part of the magnetic design, in our typical use, between the permanent magnet and the armature there is a rubber pad to allow the shear motion of the armature (the only moving part) in the X-direction relative to the stationary permanent magnet. Both the armature and the stator core are made of soft magnetic material. The bias flux generated by the permanent magnet enters the armature from its top surface, then splits inside the

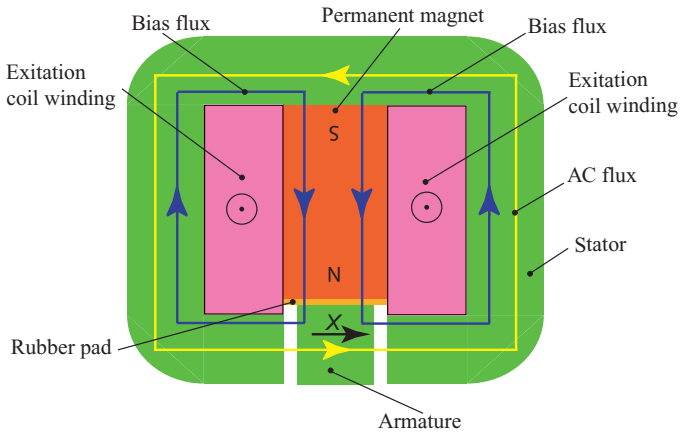


Figure 2-23: A cross section of the canonical ultra fast motor.

armature, and then exits the armature on both the left surface and the right surfaces horizontally.

In this design, all the key features of hard-linearized normal-stress motors are represented. There are two armature/stator air gaps on opposite faces of the armature, and thus no matter where the armature is, the total air gap length is fixed. The total flux from the permanent magnet is essentially constant, as the permanent magnet reluctance is much larger than that of the air gaps. The excitation coil generates differential flux on the left and right surfaces of the armature. The net actuating force is the subtraction between the normal force on the left surface of the armature and that on the right surface. Therefore, this design is a hard-linearized normal stress motor.

In comparison with other hard-linearized normal-stress actuators reviewed previously, there are several distinct features. First, the CUFM design is characterized by placing a permanent magnet between an armature and a stator, in contrast to those previously reviewed designs in which permanent magnets are placed between either two stators or between two armatures. Second, the CUFM is a self-contained modular design. In the design with permanent magnets placed between stators, only one pair of stators can work together. The CUFM needs only one stator and one armature. As a result, this CUFM can be used as an module so that multiple of them can work together in parallel to increase load capacity without losing acceleration.

Finally, the most important feature in the CUFM is that almost all of the armature surface area can be distributed with normal flux and thus be used for normal-stress generation. Consequently, the achievable acceleration can be close to the limit shown in Figure 2-3.

Lequesne's long-stroke actuator (LLSA) design shown in Figure 2-22 is the closest prior-art to the CUFM. However, there are fundamental differences between the LLSA and the CUFM. The air gap in the LLSA (5∼8 mm) is much larger than the CUFM (less than 100 μm). As a result, in the LLSA the actuating force generated by the coil mainly works on leakage fluxes. With the aid of two springs, two coils are energized separately. Most importantly, the LLSA works in a binary on/off mode, and thus both the linearity issue and continuous positioning issue were not addressed in [36].

Based on the CUFM, we have developed constructed several other designs for various application requirements, as discussed below.

2.4 Linear Ultra Fast Motors

Using four CUFMs surrounding a motion backbone, a linear ultra fast motor with only one degree of freedom motion is constructed as shown in Figure 2-24. The motion backbone is made of high specific stiffness ceramic materials such as silicon carbide and alumina. Based on this motor, we designed and built an ultra fast tool servo for diamond turning, which is discussed in more detail in Chapter 4.

2.5 Rotary Ultra Fast Motor

Figure 2-25 shows two rotary ultra fast motor designs, which have slight differences in stator construction. In Figure 2-25(a), the stator is laminated along the plane perpendicular to the rotor axis. The stator in Figure 2-25(b) can be composed of four wound tape cores installed inside a solid housing. In each design, four CUFMs work in parallel to drive the rotor. As actuating forces can be generated in the X, Y, and rotation (around rotor axis) directions, this has three degrees of freedom (DOF)

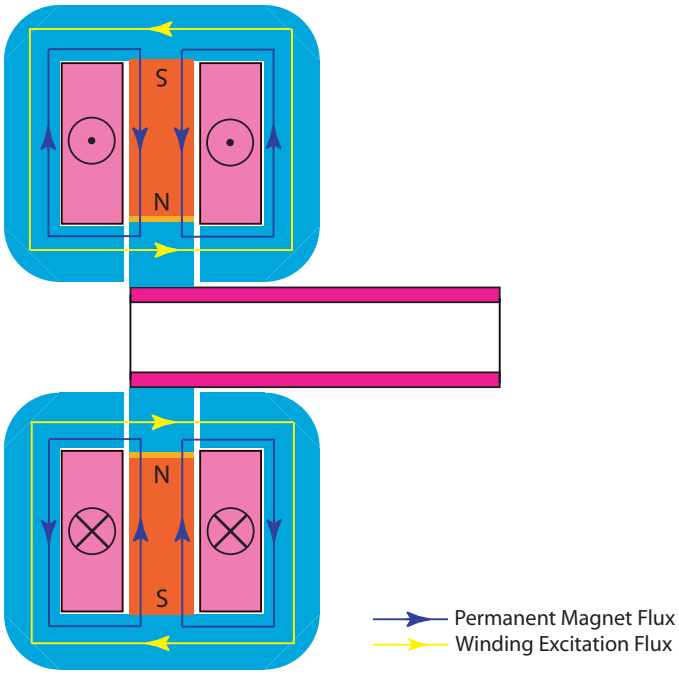


Figure 2-24: Linear ultra fast motor design.

actuation. These designs well exemplify the idea of multiple small actuators working in parallel to achieve high acceleration and high force capacity.

As there are many interesting control issues in actuating the moving member of these rotary ultra fast motors in 3 DOF, we discuss the decoupling drive and control methods of these devices with more details in Appendix C.

2.6 5 DOF Ultra Fast Motor

As one rotary ultra fast rotary motor presented in the last section can have three DOF actuation, by connecting two of them together as shown in Figure 2-26, the rotor can have 5 DOF active actuation. The sixth DOF is the axial direction of the rotor, and in this direction the rotor is passively restrained by permanent magnets. The main advantages of this 5 DOF ultra fast motor is that no mechanical bearing is needed. Therefore, a lot of problems resulted from either flexural bearings or rubber bearings can be avoided. The rotor potentially can achieve linear acceleration at thousands of G's level in two orthogonal directions, and differential accelerations at

the thousands of G's level in two rotary axis, and thousands of G's rotary acceleration at the armature around the rotor axis. Possible applications include high performance laser steering systems. It can also be used as fast tool servos and printing plate electric engravers. When used for cutting tools, if the axial stiffness of the passive magnetic bearing is not sufficient, either a linear ultra fast motor as presented in previous section can be added to achieve 6 DOF active control or a uniaxial flexural bearing can be attached to the rotor.

2.7 Rotary Fast Motor for Scanning Application

Figure 2-27 shows another variation of the CUFM, except that the armature is driven by a torque instead of a force. This design is very simple and easy to build. However, the achievable acceleration of this design is not as high as the proposed ultra fast rotary motors mainly for two reasons. First, the inertia of the rotor is larger because the middle part of the armature has no force contribution. Second, only one side of the armature is used for force generation. As the motion backbone of this motor is made of soft magnetic material, its structural dynamics is not as good as those designs with a ceramic backbone. Another C-core can be placed in the other side of the armature to achieve higher acceleration. Even so the structural problem and the inertial problem will not be solved. The bearing supporting the rotor could be a flexural hinge bearing made of steel or a rubber bearing between two cylindrical surfaces as proposed by Rivin [71].

2.8 Summary

This chapter first reviews normal stress actuator designs in the literature, including solenoids, magnetic bearings with current biasing, magnetic bearings with position feedback linearization, hard-linearized magnetic bearings and fast motors. With this review in mind, we present a canonical ultra fast motor design, which belongs to the family of hard-linearized normal-stress actuators. The ultra fast actuator is self-

contained, and easy to work in parallel. More important, almost the full armature pole surface area can be used for force generation. Further, using the CUFM as a building block, we discusse the design of a linear ultra fast motor, a 3 DOF rotary ultra fast motor, and a 5-DOF ultra fast motor. All these actuators can potentially used in high acceleration applications.

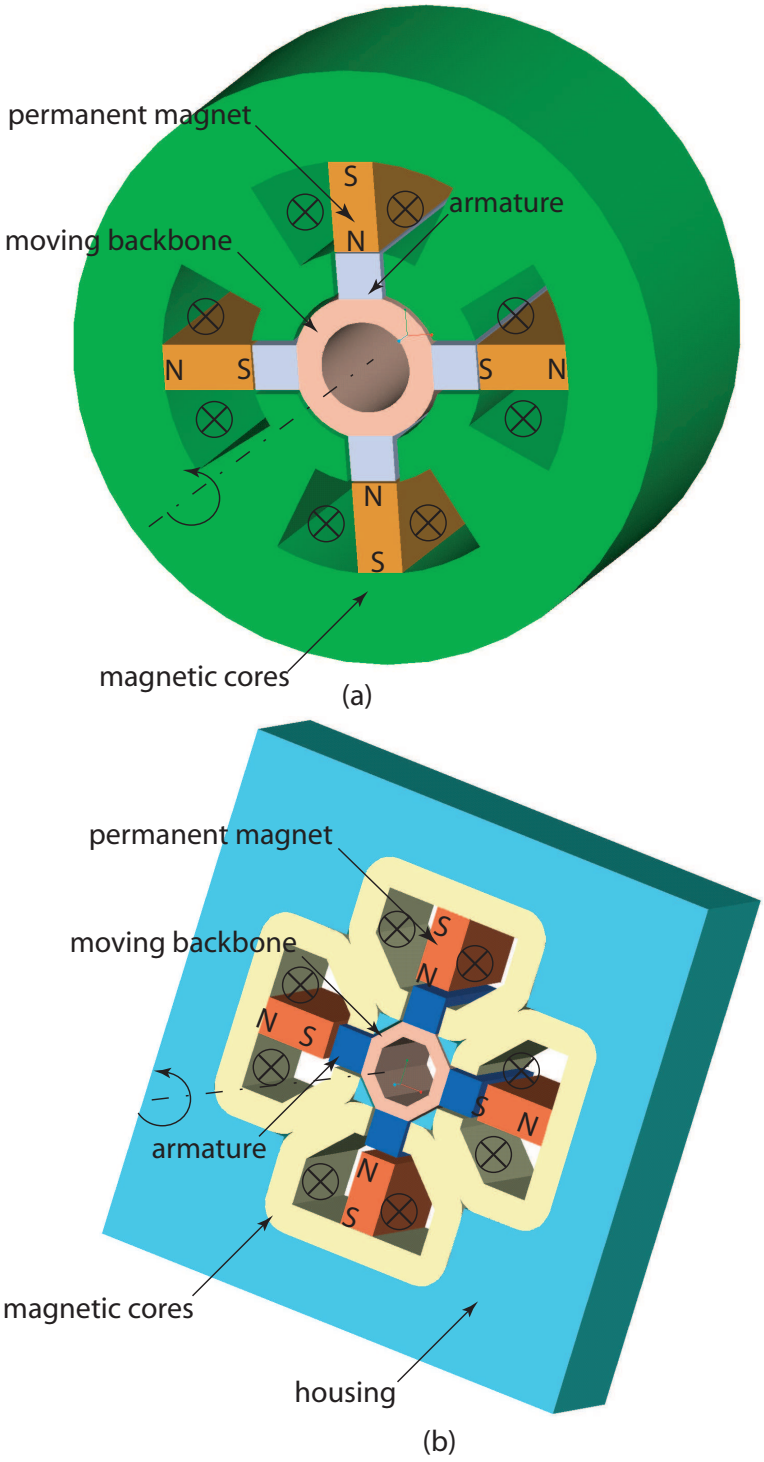


Figure 2-25: Ultra fast rotary motor design.

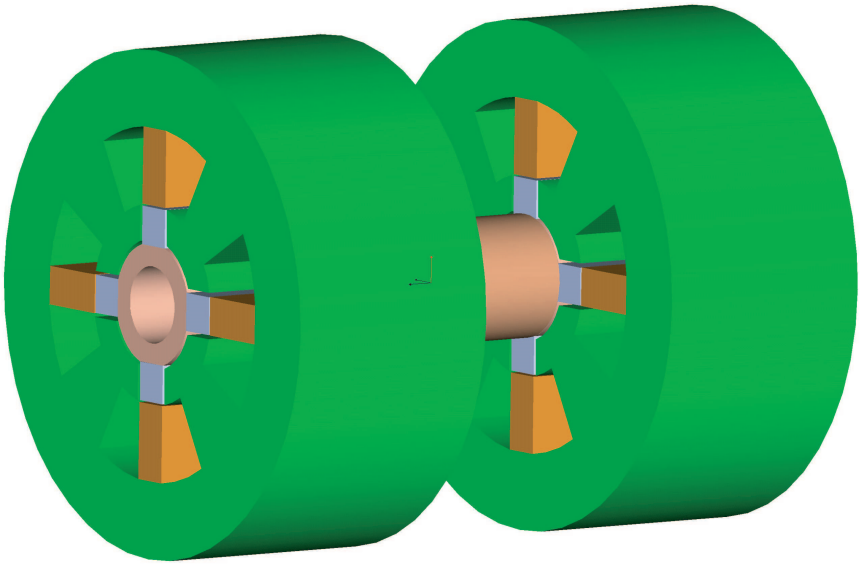


Figure 2-26: Magnetically levitated 5 DOF ultra fast motor.

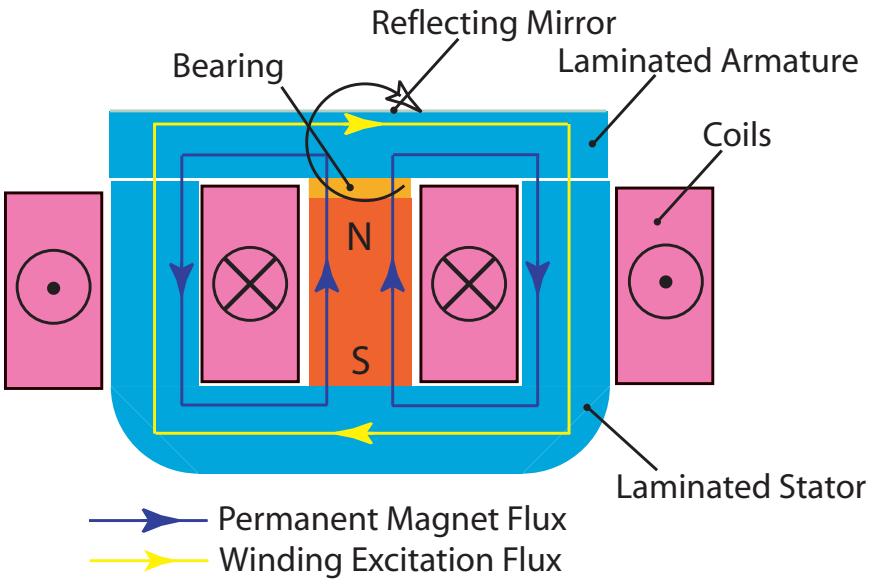


Figure 2-27: A simplified fast rotary motor.

Chapter 3

Ultra Fast Motor Analysis

This chapter analyzes the working principles of the canonical ultra fast motor (CUFM) depicted in Figure 3-1, which is the building block for many possible designs presented in Chapter 2. Most of the resulting analysis can be applied directly to other hard-linearized normal stress actuators of the types shown in Chapter 2.

Figure 3-1 shows a cross section of CUFM, which is comprised of the moving armature, a stator core, a permanent magnet, and excitation coil windings. While not part of the magnetics, the figure shows the rubber pad bearing located between the permanent magnet and the armature, which is designed to allow shear motion of the armature in X-direction relative to the stationary magnet. Both the armature and the stator core are made of soft magnetic material. The air gaps on the left and right side of the armature are denoted as $X_0 + X$ and $X_0 - X$ respectively, where X_0 is the air gap when the armature is centered and X is the armature displacement taken as positive for movement to the right. The permanent magnet plus the excitation coils generate total flux B_1 on the right side of the armature and total flux B_2 on the left. The reference direction of both total fluxes is the same as that of the displacement X. These fluxes contain both DC components $\overline{B}_1$ and $\overline{B}_2$ generated by the permanent magnet and AC components $\tilde{B}_1$ and $\tilde{B}_2$ generated by the excitation coil windings. Note that the reference directions for $\overline{B}_1$ and $\overline{B}_2$ are outwordly directed fro mthe armature to give the bias paths shown in Figure 3-1.

We next analyze these flux components in Sections 3.1 and 3.2. The supposition

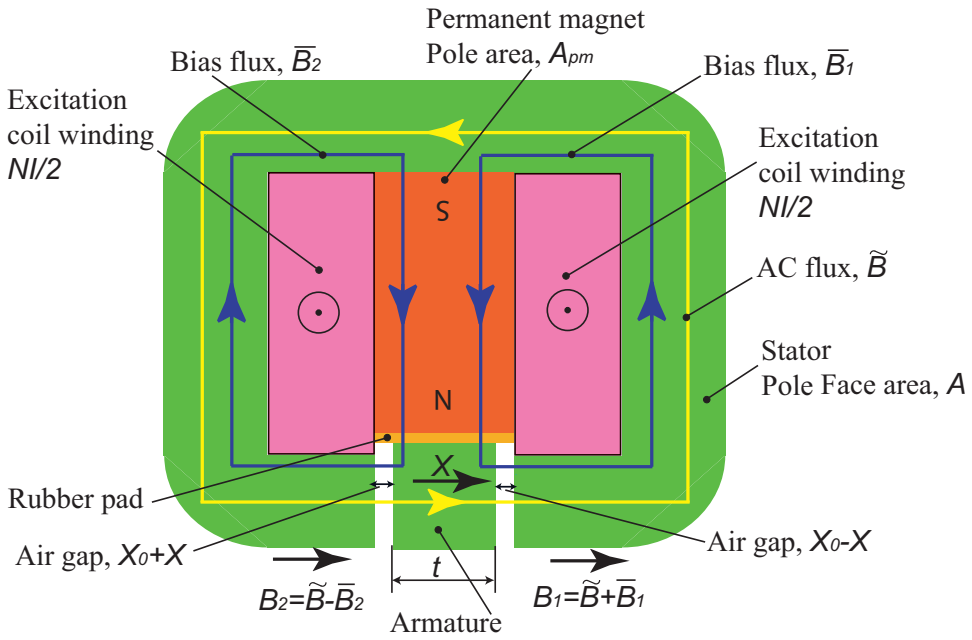


Figure 3-1: Cross section of canonical ultra fast motor. The total excitation coil windings NI are split into two halves in the left side and the right side. The armature thickness is labelled as t.

of them is shown in Section 3.3. Based on this flux analysis, actuating forces are calculated using Maxwell's tensor method, energy method, and co-energy method, in Sections 3.4, 3.5, and 3.6 respectively. The force expression is further simplified into a normalized format in Section 3.7 for more general applicability. Consequently, achievable acceleration is derived for the CUFM using typical soft magnetic materials in Section 3.8. The required power and voltage issues are discussed in Section 3.10. We conclude this chapter in Section 3.11.

3.1 DC flux

Figure 3-2 shows a simplified schematic cross-section of CFUM, which contains only components related to DC flux generation. The permanent magnet of length L with north pole at the armature generates bias flux $\overline{B}_1$ and $\overline{B}_2$ as shown in Figure 3-2. Note that $\overline{B}_1$ and $\overline{B}_2$ have different reference directions; they are positive when outwordly directed at the armature.

The total flux flowing from the permanent magnet into the armature is denoted

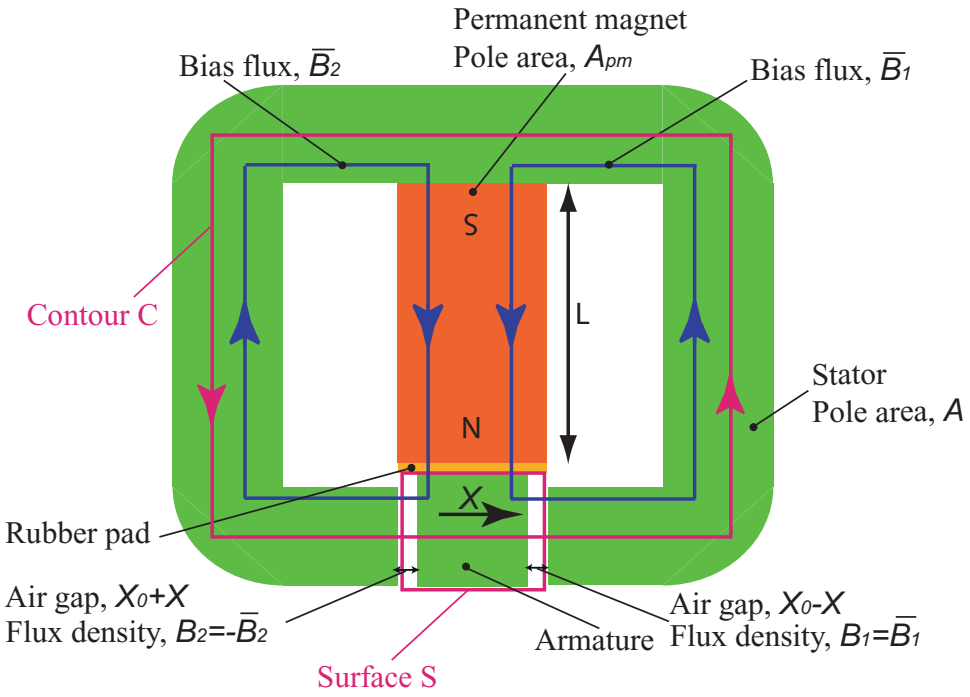


Figure 3-2: DC flux path in canonical ultra fast motor.

as $\overline{\Phi}$. Applying Gauss's Law on the surface S in Figure 3-2, we find

$$(\overline{B_1} + \overline{B_2})A - \overline{\Phi} = 0, \tag{3.1}$$

where A is the stator pole face area at each working gap.

Assuming that the permeability μ is infinite inside both the armature and the stator and with the coil current set at zero, we can apply Ampere's Law to the contour C in Figure 3-2 to yield

$$\begin{aligned} \oint_C \overrightarrow{H} \cdot \overrightarrow{dl} &= 0, \\ \overline{B_1}(X_0 - X) - \overline{B_2}(X_0 + X) &= 0. \end{aligned} \tag{3.2}$$

Solving (3.1) and (3.2) yields

$$\begin{aligned} \overline{B}_1 &= \frac{X_0 + X}{X_0}\overline{B}, \\ \overline{B}_2 &= \frac{X_0 - X}{X_0}\overline{B}, \end{aligned} \tag{3.3}$$

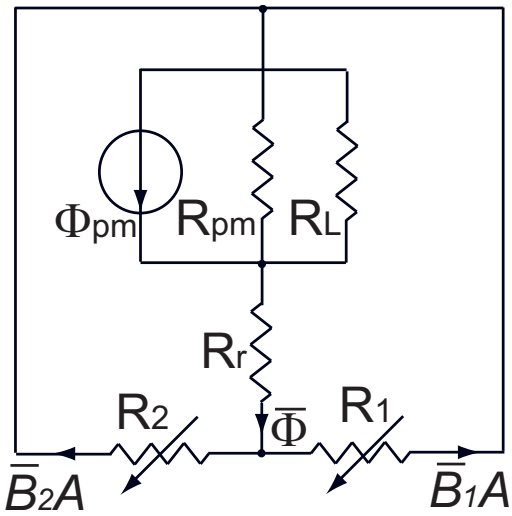


Figure 3-3: DC flux magnetic circuit model of the canonical ultra fast motor.

where $\overline{B} = \overline{\Phi}/2A$.

Next, we will solve for $\overline{\Phi}$ using a magnetic circuit model shown in Figure 3-3. By rearranging the topology, the model in Figure 3-3 can be equivalently transformed into that in Figure 3-4. The permanent magnet is modelled as a constant magnetic flux source Φ_{pm} with an internal-reluctance R_{pm}, where Φ_{pm} is the total flux through the permanent magnet when its internal magnetic field $H = 0$. According to the permanent magnet model in Appendix A, $\Phi_{pm} = A_{pm}B_r$, where B_r is the remanence of the permanent magnet and A_{pm} is the pole face area of the permanent magnet. The leakage path of the permanent magnet is modelled as a reluctance R_L. The rubber pad ($\mu = \mu_0$) is modelled as a reluctance R_r. Air gaps in the right side and the left side are modelled as variable reluctances R_1 and R_2 respectively. According to electromagnetic theory, these reluctances are expressed as

$$R_1 = \frac{X_0 - X}{\mu_0 A}, \tag{3.4}$$

$$R_2 = \frac{X_0 + X}{\mu_0 A}, \tag{3.5}$$

$$R_{pm} = \frac{L}{\mu_0 A_{pm}}, \tag{3.6}$$

$$R_r = \frac{t_R}{\mu_0 A_{pm}}, \tag{3.7}$$

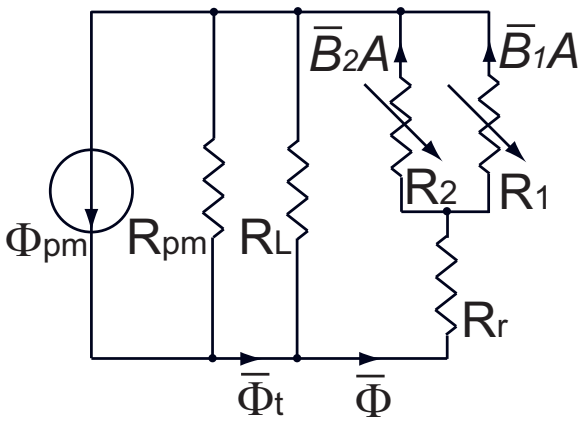


Figure 3-4: Equivalent magnetic circuit model for DC flux distribution in CUFM. $\overline{\Phi}$ is the flux from from permanent magnet into the armature. The total flux flowing out of the permanent magnet is labelled as $\overline{\Phi}_t$. The difference between $\overline{\Phi}_t$ and $\overline{\Phi}$ goes through the leakage path modelled as R_L.

where μ_0 is vacuum permeability and t_R is the rubber pad thickness. The leakage reluctance R_L is difficult to model using lumped parameters, but it can be derived from an FEA result.

Figure 3-5 shows a field plot using the finite element analysis package ANSYS for a design of typical parameters: X_0=0.1 mm, t_R=0.4 mm; the armature cross section is 4 mm by 4 mm. By counting the flux lines we can see that $\overline{\Phi}_t$ corresponds to 45 flux lines and $\overline{\Phi}$ 37 flux lines.

From this result, we can calculate the reluctance ratio as

$$\frac{\overline{\Phi}}{\overline{\Phi}_t} = \frac{R_L}{R_1//R_2 + R_r + R_L} = \frac{37}{45} = 0.82, \tag{3.8}$$

$$\Rightarrow R_L = 5.6(R_1//R_2 + R_r). \tag{3.9}$$

Further, we can show that $R_1//R_2$ is much less than R_r. $R_1//R_2$ has the maximum value when the armature is centered, and in this position we calculate

$$\frac{R_1//R_2}{R_r} = \frac{X_0}{2t_R}\frac{A_{pm}}{A} = \frac{X_0}{2t_R} = 0.125, \tag{3.10}$$

Combining (3.9) and (3.10) yields $R_L = 6.3R_r$. As a result, $R_1//R_2$ and its variation with position X are negligible when compared to R_r and R_L. This calculation

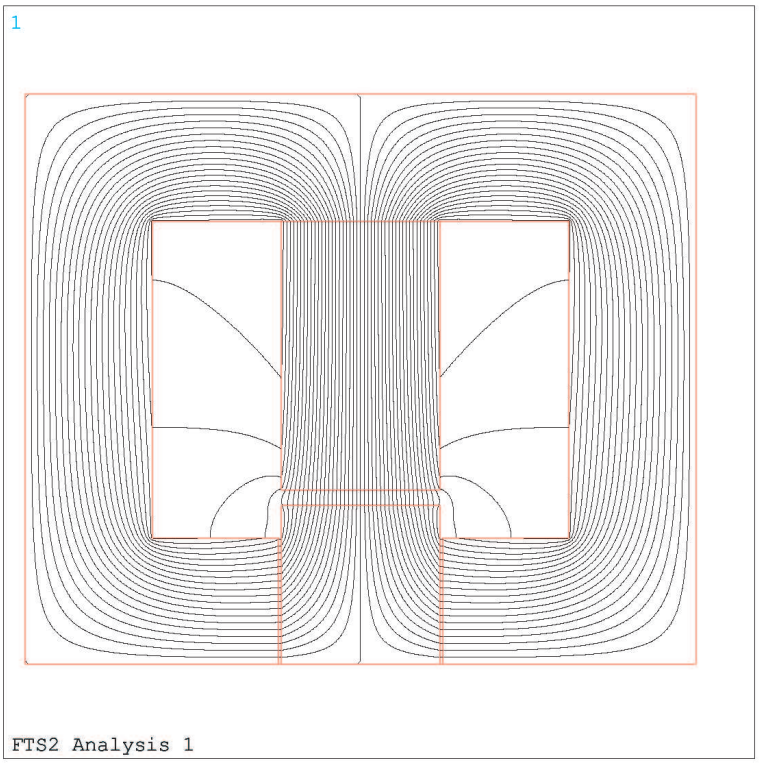


Figure 3-5: FEA solution of DC field in CUFM using typical design parameters. The permeability of the soft magnetic materials is assumed as $5000\mu_0$.

for a typical CUFM design is not intended to give exact value of R_L in general, but can be used to illustrate the qualitative relation of reluctances in the magnetic circuit model for typical designs of the CUFM.

Now, we can apply circuit theories to the model in Figure 3-4, and express $\overline{\Phi}$ as

$$\overline{\Phi} = \alpha_L \Phi_{pm}, \tag{3.11}$$

where α_L is defined as the leakage factor with an expression

$$\alpha_L = \frac{R_{pm}//R_L}{R_{pm}//R_L + R_r + R_1//R_2}. \tag{3.12}$$

Setting the permanent magnet length L=14 mm for a typical design into (3.6),

we can get

$$R_{pm} = \frac{L}{t_R} R_r = 35 R_r. \tag{3.13}$$

As shown before, $R_1 // R_2$ is much less than R_r. Therefore, α_L can be well approximated as a constant in our typical designs, despite the fact that $R_1 // R_2$ varies with the armature position X. According to (3.12), for this typical design we can calculate the leakage factor α_L as

$$\alpha_L = \frac{35R_r // 6.3R_r}{35R_r // 6.3R_r + R_r + 0.125R_r} = 0.826. \tag{3.14}$$

This result is very close to the ratio of $\overline{\Phi}/\overline{\Phi}_t = 0.82$. The underlying intuition is that R_{pm} can be approximated as infinity. Hence, for typical designs we can use the ratio of $\overline{\Phi}/\overline{\Phi}_t$ calculated from FEA as the leakage factor α_L.

As a result, the variables $\overline{\Phi}$ and $\overline{B}$ that we introduce earlier are constants independent of X. Therefore with $\overline{B}$ assumed constant, the bias flux $\overline{B}_1$ and $\overline{B}_2$ are linear functions of X according to (3.3).

3.2 AC flux

Figure 3-6 shows a cross-section of the CUFM including only the components components related to AC flux generation.

The total coil winding NI Ampere-turns (pointing out of the paper) generates flux densities $\tilde{B}_1$ and $\tilde{B}_2$ at the right air gap and the left air gap of the armature, respectively. Allying Gauss's Law to the surface S in Figure 3-6, we get

$$\tilde{B}_1 A - \tilde{B}_2 A = 0, \tag{3.15}$$

As a result, we can use $\tilde{B} = \tilde{B}_1 = \tilde{B}_2$ to represent fluxes in both air gaps.

Further, by applying Ampere's Law to the contour C in Figure 3-6 and assuming infinite permeability in the soft magnetic materials for both the armature and the

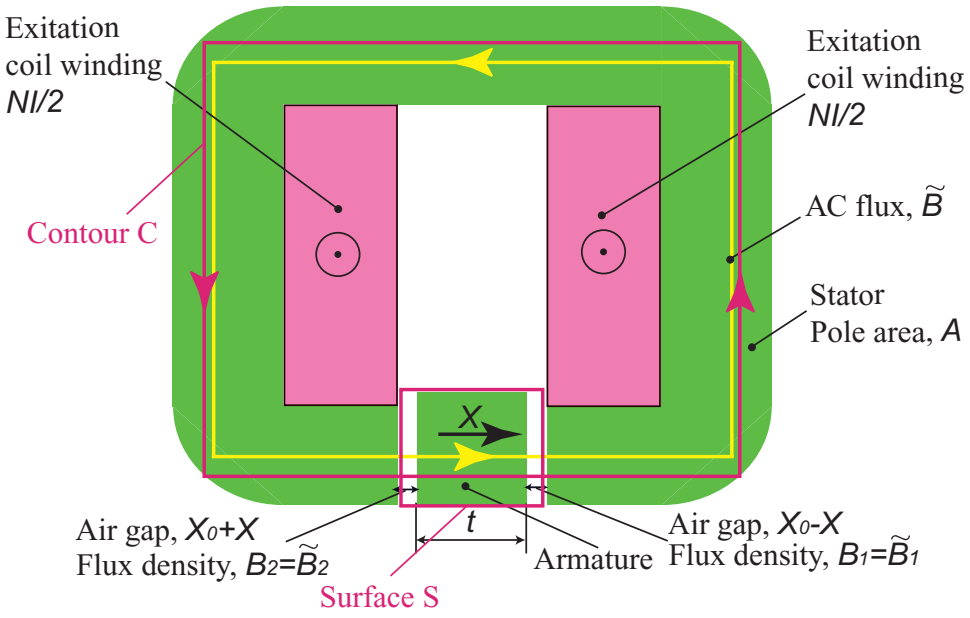


Figure 3-6: AC flux path in the canonical ultra fast motor.

stator, we can get

$$\begin{aligned} \oint_C \overrightarrow{H} \cdot \overrightarrow{dl} &= NI, \\ \frac{\tilde{B}_2}{\mu_0}(X_0 + X) + \frac{\tilde{B}_1}{\mu_0}(X_0 - X) &= NI, \\ \tilde{B} &= \frac{NI\mu_0}{2X_0}. \end{aligned} \tag{3.16}$$

Figure 3-7 shows a magnetic circuit model to illustrate this system. As R_1 and R_2 are much less than $R_{pm}//R_L + R_r$, we can ignore the reluctances associated with the permanent magnet and the rubber pad. Although the reluctances R_1 and R_2 vary with the armature position X, the sum of them $R_1 + R_2$ is a constant independent of X. Therefore, the AC flux $\tilde{B}$ is a linear function of excitation current I, independent of changes in X.

3.3 Total flux

Figure 3-8 shows a magnetic circuit model for the whole CUFM, by superposition of both the DC flux and AC flux. The flux B_1 in the right air gap and the flux B_2 in

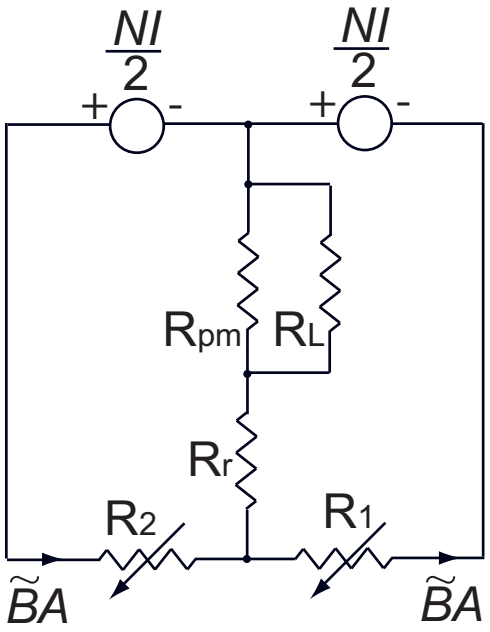


Figure 3-7: Magnetic circuit model of AC flux in CUFM.

the left air gap are

$$B_1 = \tilde{B} + \overline{B_1}, \tag{3.17}$$

$$B_2 = \tilde{B} - \overline{B_2}. \tag{3.18}$$

Collecting the results from previous sections thus gives

$$B_1 = \frac{NI\mu_0}{2X_0} + \frac{X_0 + X}{X_0}\overline{B}, \tag{3.19}$$

$$B_2 = \frac{NI\mu_0}{2X_0} - \frac{X_0 - X}{X_0}\overline{B}. \tag{3.20}$$

The reference direction for B_1 and B_2 is taken as positive in the X-direction.

Using circuit-transform theories, we can further obtain a simplified magnetic circuit model for CFUM as shown in Figure 3-9, which is equivalent to the model in Figure 3-8.

In this model, $\overline{\Phi}_{pm}$ and $\overline{R}_{pm}$ are constants defined as

$$\overline{\Phi}_{pm} = \frac{R_L // R_{pm}}{R_L // R_{pm} + R_r}\Phi_{pm}, \tag{3.21}$$

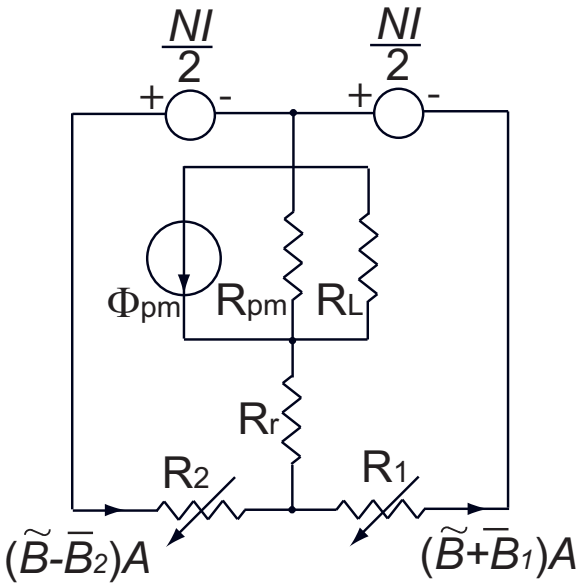


Figure 3-8: Full circuit model of the canonical ultra fast motor.

$$\overline{R}_{pm} = R_r + R_L//R_{pm}. \tag{3.22}$$

Later, this model will be used for force calculation via energy and co-energy methods.

3.4 Tensor Point of View on Force Generation

In this section, we use the Maxwell stress tensor theory to calculate the net forces acting on the armature. Figure 3-10 shows fluxes on the armature in both X-direction and Y-direction. $B_3 = \overline{\Phi}/A_{pm}$ is the flux on the top surface of the armature. The magnetic force in the i-direction on the armature F_i can be calculated according to the Maxwell stress tensor theory in [86] as

$$F_i = \oint_S T_{ij} n_j da, \tag{3.23}$$

where i and j are the coordinate directions and T_{ij} is the magnetic stress tensor,

$$T_{ij} \;=\; \mu_0 H_i H_j - \frac{\mu_0}{2}\delta_{ij} H_k H_k, \tag{3.24}$$

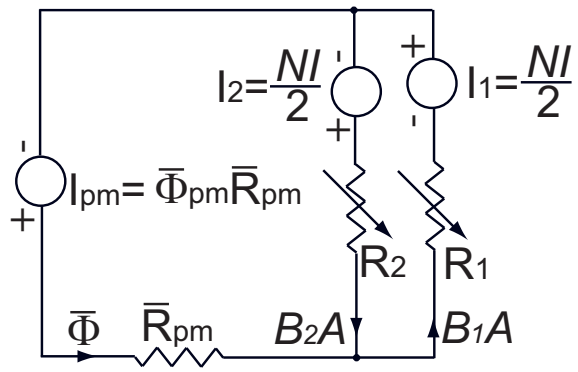


Figure 3-9: Simplified full circuit model of CUFM. The flux source (Φ_{pm}) and the reluctances (R_pm, R_L, and R_r) in Figure 3-8 are equivalently replaced with the excitation current source I_{pm} and the reluctance $\overline{R}_{pm}$.

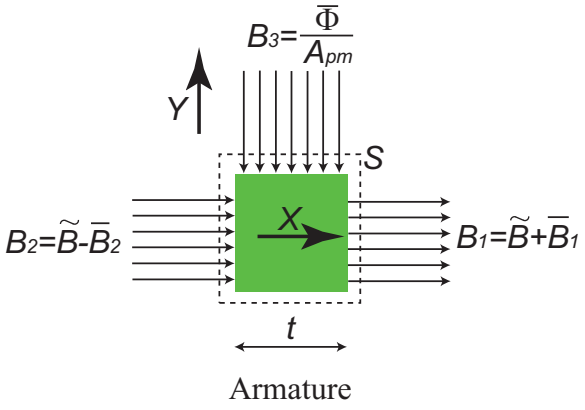


Figure 3-10: Flux distribution on the armature.

$$\delta_{ij} = \begin{cases} 1, & \text{when i} = \text{j} \\ 0, & \text{when i} \neq \text{j} \end{cases}, \tag{3.25}$$

$$H_k H_k = H_x H_x + H_y H_y + H_z H_z. \tag{3.26}$$

By integrating over the surface S in Figure 3-10, the actuating force on the armature in the X-direction F_X is thus calculated as

$$\begin{aligned} F_X &= \frac{B_1^2 - B_2^2}{2\mu_0} A \\ &= \frac{A}{2\mu_0}(B_1 + B_2)(B_1 - B_2). \end{aligned} \tag{3.27}$$

Combining (3.19) and (3.20) with this equation yields

$$F_X = \frac{A}{\mu_0}\overline{B}\left(\frac{\mu_0 NI}{X_0} + \frac{2X}{X_0}\overline{B}\right). \tag{3.28}$$

Similarly, the actuating force on the armature in the Y-direction F_Y is

$$F_Y = \frac{B_3^2}{2\mu_0}A_{pm} = \frac{A_{pm}}{2\mu_0}\left(\frac{2A\overline{B}}{A_{pm}}\right)^2 = \frac{2A^2\overline{B}^2}{\mu_0 A_{pm}}. \tag{3.29}$$

3.5 Energy Point of View on Force Generation

In this section, the actuating force is derived again using the energy method based on the magnetic circuit model in Figure 3-9.

According to [86], the actuating force can be calculated from the system stored energy function $W(\Lambda_e, X)$, where Λ_e is the flux vector of all excitation sources, and X is the displacement. The actuating force F_x in the X-direction is calculated as

$$F_X = -\frac{\partial}{\partial X}W(\Lambda_e, X)|_{\Lambda_e}. \tag{3.30}$$

When assuming the permeability of the soft magnetic material as infinity, energy can only be stored in air gaps, i.e., reluctances. Therefore, the total stored energy in a lumped-parameter system can be calculated as

$$W(\Lambda_e, X) = \sum_{i=1}^{n}\frac{\Phi_i^2(\Lambda_e, X)R_i(X)}{2}, \tag{3.31}$$

where n is the number of reluctances in the system, and Φ_i is the flux flowing through the reluctance R_i.

In the lumped parameter model shown in Figure 3-9, the excitation flux vector Λ_e

includes two coil winding fluxes, and the flux associated with the permanent magnet,

$$\Lambda_e = \begin{bmatrix} \overline{\Phi} \\ B_1 A \\ B_2 A \end{bmatrix}. \tag{3.32}$$

The total stored energy is

$$W(\Lambda_e, X) = \frac{1}{2} R_1 (B_1 A)^2 + \frac{1}{2} R_2 (B_2 A)^2 + \frac{1}{2} \overline{R}_{pm} \overline{\Phi}^2. \tag{3.33}$$

According to (3.30), the actuating force in the X-direction is calculated as

$$\begin{aligned} F_X &= -\frac{\partial}{\partial X} W(\Lambda_e, X)|_{\Lambda_e} \\ &= -\frac{\partial}{\partial X} W(\Lambda_e, X)|_{B_1,\ B_2,\ \overline{\Phi}} \\ &= -\frac{1}{2}(B_1 A)^2 \frac{\partial}{\partial X} R_1 - \frac{1}{2}(B_2 A)^2 \frac{\partial}{\partial X} R_2 - \frac{1}{2}\overline{\Phi}^2 \frac{\partial}{\partial X} \overline{R}_{pm}. \end{aligned} \tag{3.34}$$

The third term in the right side of the above expression is zero, because $\overline{R}_{pm}$ is constant. The force expression is thus transformed into

$$\begin{aligned} F_X &= -\frac{1}{2}(B_1 A)^2 \frac{\partial}{\partial X} \left(\frac{X_0 - X}{\mu_0 A} \right) - \frac{1}{2}(B_2 A)^2 \frac{\partial}{\partial X} \left(\frac{X_0 + X}{\mu_0 A} \right) \\ &= \frac{A}{2\mu_0}(B_1^2 - B_2^2). \end{aligned} \tag{3.35}$$

This result is the same as that derived using the Maxwell stress tensor method in Section 3.4.

3.6 Co-Energy Point of View on Force Generation

In this section, the actuating force is derived again by applying co-energy method to the model in Figure 3-9. According to [86], the actuating force can be calculated from the system co-energy function $W'(I_e, X)$, where I_e is the magnetomotive force vector,

and X is the displacement. The actuating force F_x in the X-direction is expressed as

$$F_X = \frac{\partial}{\partial X} W'(I_e, X)|_{I_e}. \tag{3.36}$$

For the model of Figure 3-9, the magnetomotive force vector I_e is

$$I_e = \begin{bmatrix} I_{pm} \\ I_1 \\ I_2 \end{bmatrix}. \tag{3.37}$$

The co-energy is calculated as

$$\begin{aligned}
W'(I_e,\ X) &= \int_{I_{pm}=0,\ I_1=0,\ I_2=0,\ X}^{I_{pm}=\overline{\Phi}_{pm}\overline{R}_{pm},\ I_1=\frac{NI}{2},\ I_2=\frac{NI}{2},\ X} \overline{\Phi} dI_{pm} + B_1 A dI_1 + B_2 A dI_2 \\
&= \int_{I_{pm}=0,\ I_1=I_2=0,\ X}^{I_{pm}=\overline{\Phi}_{pm}\overline{R}_{pm},\ I_1=I_2=0,\ X} \overline{\Phi} dI_{pm} \\
&\quad + \int_{I_{pm}=\overline{\Phi}_{pm}\overline{R}_{pm},\ I_1=I_2=0,\ X}^{I_{pm}=\overline{\Phi}_{pm}\overline{R}_{pm},\ I_1=I_2=\frac{NI}{2},\ X} (B_1 A + B_2 A)\, dI_1 \\
&= \int_{I_{pm}=0,\ I_1=I_2=0,\ X}^{I_{pm}=\overline{\Phi}_{pm}\overline{R}_{pm},\ I_1=I_2=0,\ X} \frac{I_{pm}}{\overline{R}_{pm} + R_1//R_2} dI_{pm} \\
&\quad + \int_{I_{pm}=\overline{\Phi}_{pm}\overline{R}_{pm},\ I_1=I_2=0,\ X}^{I_{pm}=\overline{\Phi}_{pm}\overline{R}_{pm},\ I_1=I_2=\frac{NI}{2},\ X} \left(\overline{B}_1 - \overline{B}_2 + 2\tilde{B}\right) A dI_1 \\
&= \frac{\overline{\Phi}_{pm}^2 \overline{R}_{pm}^2}{2\left(\overline{R}_{pm} + R_1//R_2\right)} + \overline{\Phi}\frac{R_2 - R_1}{R_1 + R_2}\frac{NI}{2} \\
&\quad + \frac{N^2 I^2}{2\left(R_1 + R_2\right)}.
\end{aligned} \tag{3.38}$$

The actuating force in the X-direction is calculated as

$$\begin{aligned}
F_X &= \frac{\partial}{\partial X} W'(I_{pm},\ I_1,\ I_2,\ X)|_{I_1,\ I_2,\ I_{pm}} \qquad (3.39) \\
&= -\frac{\overline{\Phi}_{pm}^2 \overline{R}_{pm}^2}{2\left(\overline{R}_{pm} + R_1//R_2\right)^2} \frac{d}{dX}\left(R_1//R_2\right) \\
&\quad + \overline{\Phi}\frac{1}{R_1 + R_2}\frac{NI}{2}\frac{d}{dX}\left(R_2 - R_1\right)
\end{aligned}$$

$$= \frac{\overline{\Phi}_{pm}^2 \overline{R}_{pm}^2}{2\left(\overline{R}_{pm} + R_1//R_2\right)^2} \frac{X}{\mu_0 A X_0} + \overline{\Phi}\frac{NI}{2X_0}$$
$$= \frac{\overline{\Phi}^2}{2}\frac{X}{\mu_0 A X_0} + \overline{\Phi}\frac{NI}{2X_0}. \quad (3.40)$$

Plugging $\overline{\Phi} = 2A\overline{B}$ into (3.40) yields

$$F_X = \frac{2A\overline{B}^2}{\mu_0}\frac{X}{X_0} + A\overline{B}\frac{NI}{X_0}, \quad (3.41)$$

which is the same as (3.28).

3.7 Normalized Expression

In last three sections, we obtained the same force expression using three different methods. In this section, we will simplify the results derived previously, by introducing the following characterizing variables,

$$B_0 = \frac{B_{sat}}{2}, \quad (3.42)$$
$$I_0 = \frac{X_0 B_{sat}}{\mu_0 N}, \quad (3.43)$$
$$F_0 = \frac{B_{sat}^2 A}{2\mu_0}, \quad (3.44)$$

where B_{sat} is the saturation flux of the soft magnetic material used for both the armature and the stator, I_0 is the excitation current which will drive the AC flux $\tilde{B}$ to B_0, and F_0 is the actuating force at saturation induction. Using B_0, I_0, and F_0, we can define normalized variables as

$$i = \frac{I}{I_0}, \quad (3.45)$$
$$x = \frac{X}{X_0}, \quad (3.46)$$
$$\eta = \frac{\overline{B}}{B_0}, \quad (3.47)$$

$$
\begin{aligned}
f_x &= \frac{F_X}{F_0}, && (3.48)\\
f_y &= \frac{F_Y}{F_0}, && (3.49)\\
\tilde{b} &= \frac{\tilde{B}}{B_0}, && (3.50)\\
b_1 &= \frac{B_1}{B_0}, && (3.51)\\
b_2 &= \frac{B_2}{B_0}, && (3.52)\\
b_3 &= \frac{B_3}{B_0}, && (3.53)
\end{aligned}
$$

where η is defined as the normalized biasing strength representing how strong the bias flux is in a CUFM design, relative to half the magnetic material saturation.

Consequently, previous results can be reformulated as

$$
\begin{aligned}
\tilde{b} &= i; && (3.54)\\
b_1 &= (1+x)\eta + i = \lambda + \eta; && (3.55)\\
b_2 &= (1-x)\eta + i = \lambda - \eta; && (3.56)\\
b_3 &= \frac{2A\overline{B}}{A_{pm}B_0} = 2\beta\eta; && (3.57)\\
f_x &= \eta(i + x\eta) = \eta\lambda; && (3.58)\\
f_y &= \frac{F_Y}{F_0} = \frac{A}{A_{pm}}\eta^2 = \beta\eta^2, && (3.59)
\end{aligned}
$$

where $\beta = A/A_{pm}$ is a design parameter representing the area ratio between stator pole face area A and permanent magnet pole face area A_{pm}, and $\lambda = i + x\eta$ represents the normalized flux linkage. As the total flux linkage of excitation coil windings Λ is calculated as

$$
\begin{aligned}
\Lambda &= \frac{NA}{2}B_2 + \frac{NA}{2}B_1,\\
&= \frac{NAB_0}{2}(b_1 + b_2),\\
&= NAB_0\lambda,\\
&= \Lambda_0\lambda, && (3.60)
\end{aligned}
$$

where $\Lambda_0 = NAB_0$ is a constant used to normalize Λ.

In order not to saturate the soft magnetic material, we need to ensure that the fluxes in both air gaps are less than the B_{sat}

$$|B_1| < B_{sat}, \tag{3.61}$$

$$|B_2| < B_{sat}, \tag{3.62}$$

$$|B_3| < B_{sat}. \tag{3.63}$$

Correspondingly, the normalized format of these constraints is

$$|b_1| = |\lambda + \eta| < 2; \tag{3.64}$$

$$|b_2| = |\lambda - \eta| < 2; \tag{3.65}$$

$$|b_3| = |2\beta\eta| < 2 \Longrightarrow \beta\eta < 1. \tag{3.66}$$

Actually, $\beta\eta = \alpha_L B_r / B_{\text{sat}}$. Hence, for typical designs (B_{sat} = 1.4 T, B_r =1.2 T, and $\alpha_L < 1$), the condition $\beta\eta < 1$ can be easily satisfied. Therefore, the concerned constraints become

$$\begin{aligned} \eta - 2 < \lambda < 2 - \eta, \\ \text{or } (1-x)\eta - 2 < i < 2 - (1+x)\eta. \end{aligned} \tag{3.67}$$

Equation (3.67) sets the proper operation zone for the excitation flux linkage λ or the excitation current i, which are illustrated with shaded areas in Figure 3-11.

According to this constraint, the maximum achievable actuating force f_x is calculated as

$$|f_x| = |\eta\lambda| < |(2-\eta)\eta|, \tag{3.68}$$

which shows that the maximum achievable force is a quadratic function of the design parameter η. Figure 3-12 shows the relation between f_x and λ at various biasing strength levels η. It can be seen that the maximum normalized force $f_x = 1$ is

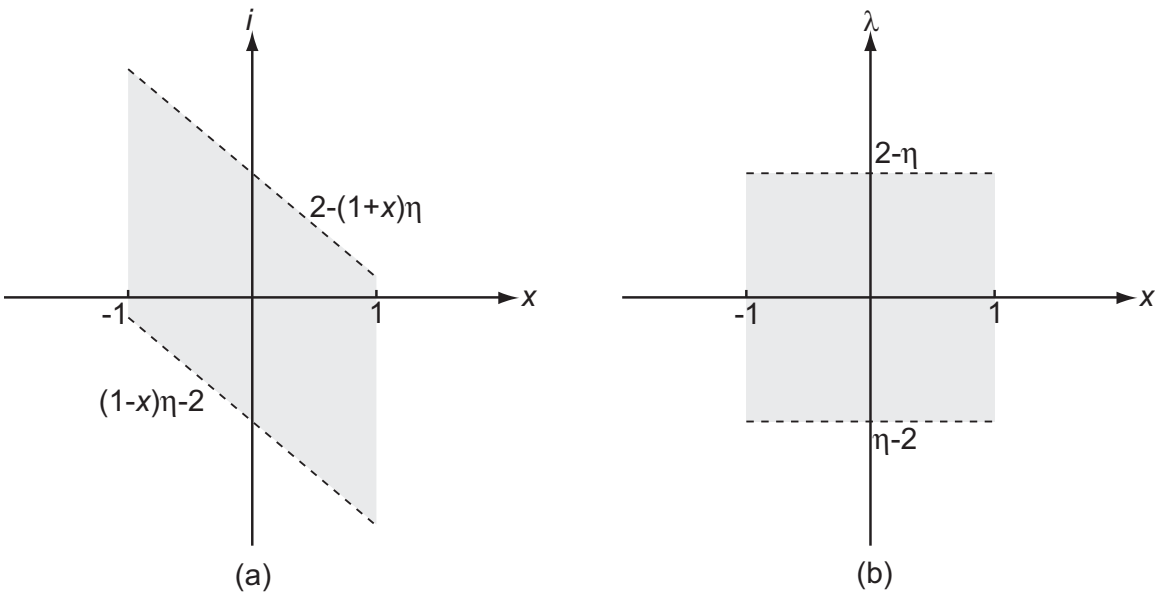


Figure 3-11: Permissible excitation zone. (a)Permissible region for the normalized excitation current i. (b) Permissible region for the normalized flux linkage λ.

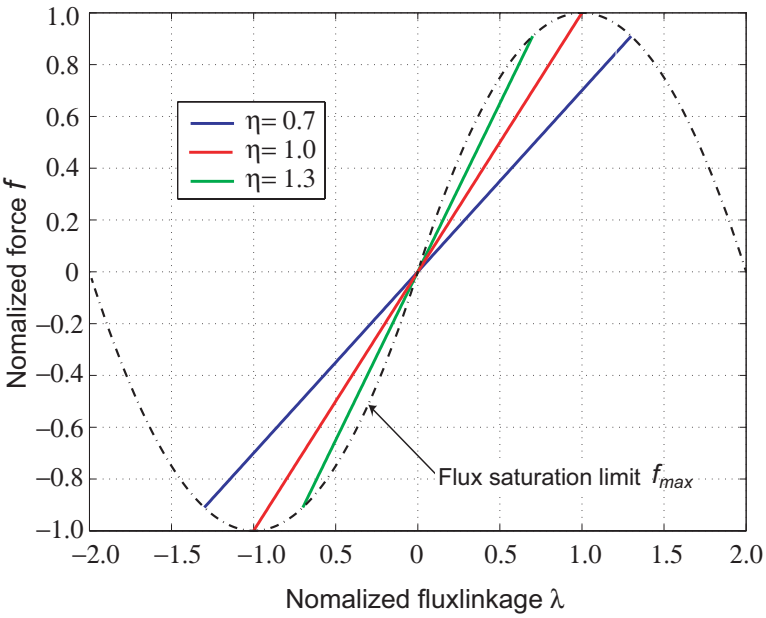


Figure 3-12: Operation curves between the normalized actuating force f_x and the normalized excitation flux linkage λ.

achieved only in a critically biased design η=1. However, for $0.7 < \eta < 1.3$, the maximum achievable normalized force is still more than 90%. An advantage for a over-biased design ($\eta > 1$) is that a higher gain from current to force can be achieved, and thus the actuator is more energy efficient. For example, if η is designed at 1.3, then the achievable maximum force is 91% of that for η=1.0, but the required current is 70% of that for η=1.0, and thus the ohmic heat generated in the coil is only 50% of the nominal η=1 design . Therefore, a biasing strength η between 1 and 1.3 is desired in order to reduce coil heating.

This force relation is further illustrated in Figure 3-13. In a critically biased case ($\eta = 1$) as shown in the middle column, the full excitation level is $\lambda = \pm 1$, with the flux saturated on one side and zero flux on the other side. In an under-biased case ($\eta < 1$) as shown in the left column, the full excitation level is $\lambda = \pm(2 - \lambda)$, with the flux saturated on one side and the flux reversed on the other side. In an over-biased case ($\eta > 1$) as shown in the right column, the full excitation level is $\lambda = \pm(2 - \lambda)$, with the flux saturated on one side and still some residual flux on the other side.

3.8 Acceleration Performance Curves

In the last section, we showed that the maximum excitation is limited by material saturation $|\lambda| < (2 - \eta)$. Here we discuss another constraint limiting the flux excitation level, which is the power loss of the magnetic materials. In high frequency AC operation, the heat generated by eddy current and hysteresis power loss inside soft magnetic materials may exceed the thermally sustainable limit even though the excitation level is well below the magnetic saturation.

Appendix B discusses in detail the power loss inside magnetic materials. Briefly speaking, the volumetric power loss p can be expressed as a function of frequency ω and excitation flux B

$$p = E(\omega, B), \tag{3.69}$$

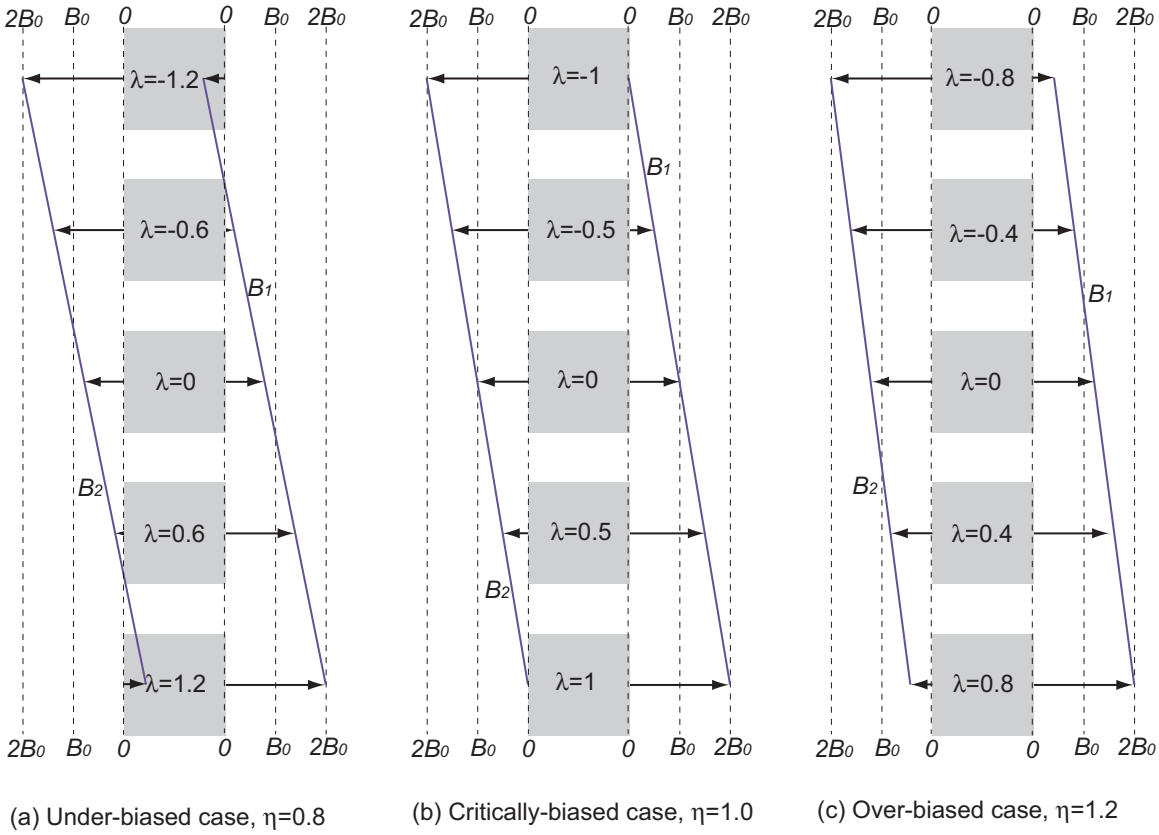


Figure 3-13: Flux densities B_1 and B_2 illustrated at various biasing η and excitation λ levels.

where E is the function representing the power loss. Appendix B gives some analytical results for the function E, but they are only used to illustrate the qualitative intuition. In reality, most soft magnetic material manufactures provide experimentally measured volumetric power loss either in curves or tables. Consequently, for a permissible volumetric power loss p_e (set by the heat transfer capability), we can conversely derive the maximum permissible excitation flux density as

$$B_{\max}(\omega) \quad = \quad E^{-1}(\omega, p_e), \tag{3.70}$$

and we call this $B_{\max}(\omega)$ the heat-limited maximum allowable flux excitation at the frequency ω. In normalized format, the maximum normalized excitation flux is

$$\lambda_{\max}(\omega) \quad = \quad \min\left(2-\eta, \frac{B_{\max}(\omega)}{B_0}\right). \tag{3.71}$$

As $B_{\max}(\omega)$ decreases with frequency, there exists a break frequency point ω_B, at which $B_{\max}(\omega_B)/B_0 = 2-\eta$. At frequencies lower than ω_B, $\lambda_{\max} = (2-\eta)$ is a constant limiting value. For $\omega > \omega_B$, we go into the limiting value $\lambda_{\max} = B_{\max}(\omega)/B_0$.

Using $\lambda_{\max}$, we can calculate the maximum achievable acceleration at various frequencies. Assuming that the armature is the only inertial load and that the full armature face area is used for force generation, the resulting acceleration a on the armature is calculated as

$$\begin{aligned} a_{\max} &= \frac{F_{\max}}{M} \\ &= \frac{F_0 f_{\max}}{A\rho t} \\ &= \frac{F_0 \eta \lambda_{\max}}{At\rho} \\ &= \frac{B_{sat}^2 \eta \lambda_{\max}}{2\mu_0 t\rho}, \end{aligned} \tag{3.72}$$

where t is the armature thickness and ρ is the armature mass density.

(3.72) also shows a further trade off in designing the biasing strength η. At high frequencies where $\lambda_{\max}$ is determined by $B_{\max}(\omega)/B_0$, the maximum acceleration $a_{\max}$ increases with biasing strength η. However, at low frequencies where $\lambda_{\max}$ is set by $2-\eta$, when $\eta > 1$ an increase in biasing strength will reduce the maximum acceleration.

Figure 3-14 shows the calculated achievable acceleration based on (3.72) and the experimental power loss data from soft magnetic material suppliers, under the assumption : (1) the biasing strength is set at $\eta = 1$; (2) the armature thickness $t = 3$ mm; and (3) the permissible volumetric power loss is $p_a = 0.25\ W/\text{cm}^3$. For non-homogenous material such as wound tape cores, the effective mass density ρ_{e} [1] and the effective flux saturation $B_{\text{sat,e}}$ [2] are used instead. We call these curves shown in Figure 3-14 as acceleration performance curves, which represent the maximum

[1] $\rho_e = \chi\rho_{\text{sm}} + (1-\chi)\rho_{\text{ep}}$, where ρ_{sm} is the mass density of the soft magnetic material, ρ_{ep} is the mass density of expoy used to bond tape cores together, and χ is the volumetric ratio between soft magnetic tapes and the overall tape-wound cores.

[2] $B_{sat,e} = \chi B_{\text{sat}}$, where B_{sat} is the saturation flux density of the tape material itself.

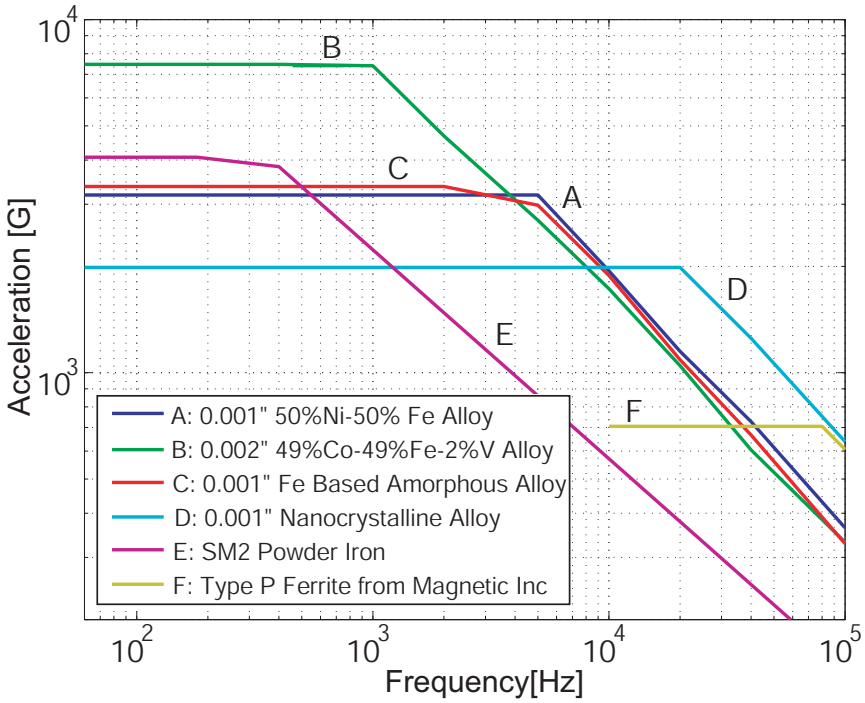


Figure 3-14: Achievable acceleration of CUFM using typical soft magnetic materials.

achievable acceleration of CUFM using typical magnetic materials.

For this specific design, these curves show that the nanocrystaline alloy core developed by Hitachi metal is good for high frequency operation, while the cobalt-iron material has the potential for over 7000 G acceleration for drive frequencies below about 1 kHz. Practical acceleration values will be reduced from these curves by additional mass loads attached to the armature. From this figure, we can see that the break frequency ω_B is 400 Hz for a CUFM using SM2 powder pressed iron, 5 kHz for Ni-Fe (50%-50%), and 20 kHz for nano-crystalline alloy. For other thickness t and permissible power loss p_e, the acceleration performance curves can be scaled properly according to (3.72).

3.9 Vertical Force Analysis

In previous sections, we mainly focused on the actuating force F_x, since this is the axis of controlled motion. This section discusses the magnetic force in the vertical direction F_y. In previous models, we assumed that the rubber pad was a constant

thickness t_R. In reality, the vertical force F_y can deform the rubber and thus change the rubber pad thickness, which again may change the vertical force. This negative-spring effect in F_y resulting from the permanent magnet should be evaluated before designing the rubber-pad bearing.

The intuition is that the force F_y is almost a constant force because the permanent magnet can be approximated as a constant flux source in our CFUM. To verify this intuition, the rubber pad thickness is expressed as $t_R - Y$ and correspondingly its associated reluctance is expressed as

$$R_r(Y) = \frac{t_R - Y}{\mu_0 A_{pm}}, \tag{3.73}$$

where Y is the displacement of the armature in the Y-direction as shown in Figure 3-10. The vertical force stiffness K_Y is defined as

$$K_Y = \frac{dF_Y(Y)}{dY}. \tag{3.74}$$

According to (3.29),

$$\begin{aligned} K_Y &= \frac{d}{dY}\left(\frac{\alpha_L^2 \Phi_{pm}^2}{2\mu_0 A_{pm}}\right) \\ &= \frac{\alpha_L \Phi_{pm}^2}{\mu_0 A_{pm}} \frac{d}{dY}\alpha_L. \end{aligned} \tag{3.75}$$

From (3.11), the derivative of α_L is

$$\begin{aligned} \frac{d}{dY}\alpha_L &= \frac{d}{dY}\left(\frac{R_{pm}//R_L}{R_{pm}//R_L + R_r + R_1//R_2}\right) \\ &= \frac{\alpha_L R_r(0)}{t_R(R_{pm}//R_L + R_r + R_1//R_2)} \\ &= \frac{\alpha_L(1-\alpha_L)}{t_R}. \end{aligned} \tag{3.76}$$

Consequently,

$$K_Y = \frac{\alpha_L^2 \Phi_{pm}^2 (1-\alpha_L)}{\mu_0 t_R A_{pm}}$$

$$
\begin{aligned}
&= \frac{\overline{\Phi}^2(1-\alpha_L)}{\mu_0 t_R A_{pm}} \\
&= \frac{4A^2B_0^2\eta^2(1-\alpha_L)}{\mu_0 t_R A_{pm}}.
\end{aligned} \tag{3.77}
$$

To evaluate the significance of this vertical force stiffness K_Y, we need to calculate the force stiffness in the X-direction K_X and then compare them. The force stiffness K_X in the X-direction is

$$
\begin{aligned}
K_X &= \frac{d}{dX}F_X \\
&= \frac{F_0}{X_0}\frac{d}{dx}f_x \\
&= \frac{F_0}{X_0}\frac{d}{dx}\eta(i+\eta x) \\
&= \frac{\eta^2 F_0}{X_0} \\
&= \frac{2B_0^2 A\eta^2}{\mu_0 X_0}.
\end{aligned} \tag{3.78}
$$

Further, the force stiffness ratio K_X/K_Y is calculated as

$$
\begin{aligned}
\frac{K_X}{K_Y} &= \frac{2B_0^2A\eta^2}{\mu_0 X_0}\frac{\mu_0 t_R A_{pm}}{4A^2B_0^2\eta^2(1-\alpha_L)} \\
&= \frac{t_R}{X_0}\frac{1}{2\beta(1-\alpha_L)}.
\end{aligned} \tag{3.79}
$$

Plugging $\alpha = 0.82$ for a typical design into (3.79) and assuming $\beta = 1$, we have

$$
\frac{K_X}{K_Y} = \frac{0.4}{0.1} \times \frac{1}{2 \times 1 \times (1-0.82)} = 11.1, \tag{3.80}
$$

which shows that K_X is about ten times higher than K_Y.

Based on this result, in a design with $F_0 = 100$ N and $X_0 = 100$ μm, K_X is expressed as $F_0/X_0 = 1$ N/μm by assuming η=1. Therefore, K_Y is approximately estimated as $K_X/10$=0.1N/μm, which is much less than rubber bearing stiffness in typical designs. Hence, the vertical force position stability will not be a problem in CUFM. In the 5 D.O.F. ultra fast rotary motor design described in Chapter 2, the

forces in vertical direction acting on the armatures are negligible, because K_Y is much less than K_X. The negative stiffness K_X does however affect the open-loop dynamics and needs to be considered in our control loop design.

3.10 Power Calculation

This section analyzes the power required to drive the CUFM. As the CUFM is a inductive load for the power amplifier, the instantaneous electrical power flowing into CUFM is dominated by reactive power, which can be calculated as

$$\begin{aligned} P(t) &= VI \\ &= \frac{d\Lambda}{dt} I \\ &= \Lambda_0 I_0 \frac{d\lambda}{dt} i \\ &= \Lambda_0 I_0 \frac{d\lambda}{dt} (\lambda - x\eta) \\ &= \Lambda_0 I_0 \frac{d\lambda}{dt} \lambda (1 - \frac{x}{\lambda}\eta). \end{aligned} \tag{3.81}$$

In this power analysis, the maximum power is needed at the high frequency operation, where the gain from flux λ to displacement x is negligible. Consequently, we can approximate (3.81) as

$$P(t) \simeq \Lambda_0 I_0 \frac{d\lambda}{dt} \lambda. \tag{3.82}$$

For steady state sinusoidal operation at frequency ω, assuming $\lambda = \lambda_{\max} \sin(\omega t)$, we can get

$$\begin{aligned} P(\omega) &\simeq \Lambda_0 I_0 \frac{d\lambda}{dt} \lambda \\ &\simeq \omega \Lambda_0 I_0 \lambda_{\max}^2(\omega) \sin(\omega t) \cos(\omega t) \\ &\simeq \frac{\omega \Lambda_0 I_0 \lambda_{\max}^2(\omega)}{2} \sin(2\omega t) \end{aligned} \tag{3.83}$$

Based on the power loss analysis for soft magnetic materials in Appendix B, the

maximum permissible excitation $\lambda_{\max}$ remains as constant $2-\eta$ for frequency region $\omega < \omega_B$. For frequency region $\omega > \omega_B$, $\lambda_{\max}$ decays as a function of frequency at either -20dB/decade or -15dB/decade, depending on whether the eddy current is in full penetration or not. As a result, $\omega\lambda^2_{\max}(\omega)$ achieves its maximum value at ω_B. Therefore, according to (3.83), the maximum required power is calculated as

$$P_{\max} \simeq \frac{\Lambda_0 I_0 (2-\eta)^2 \omega_B}{2}. \tag{3.84}$$

The maximum required driving voltage can be calculated as

$$\begin{aligned} V) &= \frac{d}{dt}(\Lambda_0 \lambda) \\ &= \frac{d}{dt}(\Lambda_0 \lambda_{\max} \sin(\omega t)) \\ &= \omega \Lambda_0 \lambda_{\max} \cos(\omega t). \end{aligned} \tag{3.85}$$

Consequently, the maximum driving voltage $V_{\max}$ is at the maximum operation frequency $\omega_{\max}$.

$$V_{\max} = (2-\eta)\omega_{\max}\Lambda_0. \tag{3.86}$$

These results will be used to guide the power amplifier design discussed in later chapters.

3.11 Summary

In this chapter, we analyzed in detail the flux generation and distribution in the CUFM. We introduced a lumped-parameter magnetic circuit model to demonstrate that the biasing flux $\overline{\Phi}$ can be approximated as a constant regardless of the armature displacement X for a CUFM design with typical parameters. As a result, DC fluxes in both air gaps are linear functions of X. The AC flux is shown to be identical in both air gaps, and is a linear function of excitation current. Based on the flux analysis, we calculated the actuating force using three different methods which lead to the

same key result: the actuating force F_x is a linear function of the excitation current I and the armature displacement X. Further, we derived the maximum achievable acceleration for common soft magnetic materials, based on the heat-limited excitation flux. The magnetic force in the vertical direction of the armature was also analyzed, and it was shown that this force is essentially constant and will not be a concern in CUFM design. Finally, we calculated the maximum required power and driving voltage for the power amplifier design.

In the next chapter, we will present ultra fast tool servo designs based on this chapter's analysis.

Chapter 4

Ultra Fast Tool Servo

This chapter presents the design and fabrication of an ultra fast tool servo (UFTS), which is driven by the linear ultra fast motor described in Chapter 2. The main design decisions and the associated performance predictions are based on the CUFM analysis in Chapter 3.

4.1 Design Objectives

The goal of this design is to demonstrate the feasibility of an ultra fast tool servo design with the following performance:

- 1000 G acceleration up to 20kHz operation;
- 1 nm RMS positioning error;
- 20 kHz closed-loop bandwidth;
- 50 μm stroke;
- no less than 20 N/μm static lateral stiffness;
- no less than 40 N/μm axial stiffness for all frequencies.

This target UFTS performance exceeds any wide-band fast tool servos so far reported in the literature, and so advances the state of art in fast tool servo technology.

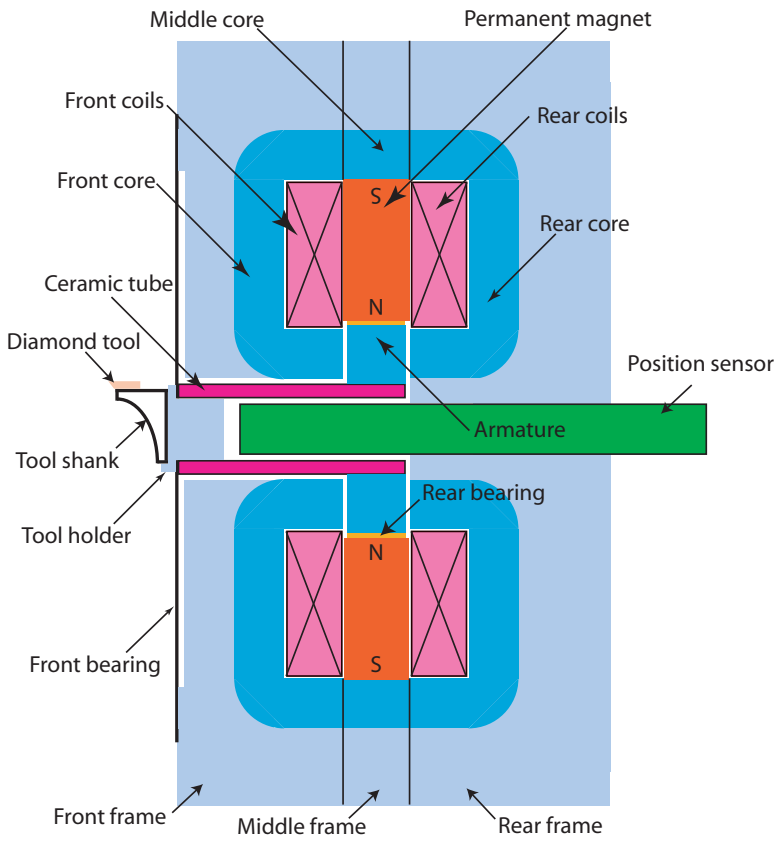


Figure 4-1: Schematic design of the ultra fast tool servo.

4.2 Schematic Design

The schematic design of our UFTS is shown in Figure 4-1, which builds upon the ultra fast linear motor of Chapter 2 by including metrology, supporting bearings and a diamond tool holder.

As shown in the figure, the motion of the armature is transferred to the tool holder by a segment of ceramic tube. The moving assembly is suspended by a spring steel flexure bearing in the front and a rubber pad bearing in the back. Consequently, the diamond tool moves primarily along the axial direction of the ceramic tube and the other five degrees of freedom are fully constrained. The air gap X_0 is set at 100 μm to easily ensure 50 μm stroke. In this schematic design, the stator cores are split into three parts: the front core, the middle core, and the rear core to simplify manufacturing and fabrication. All frames (including the front frame, the middle frame, and the rear frame) are made of Type 304 non-magnetic stainless steel to provide good

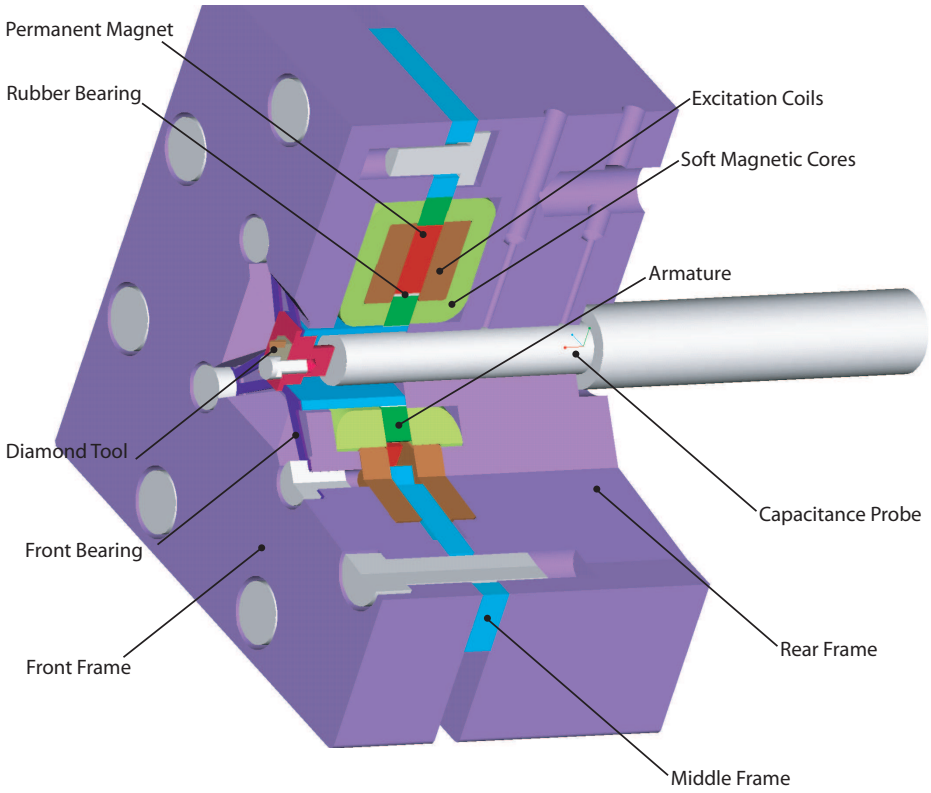


Figure 4-2: ProE model of the ultra fast tool servo.

structural stiffness. The permanent magnets in the UFTS are Ne-Fe-B permanent magnets of 1.2T remanence, which are rated up to 120°C operation. A capacitance probe (ADE 5501 active probe) of 100 kHz bandwidth directly measures the tool holder motion from behind. Therefore, the tool can be controlled accurately, in despite of structural length variations induced by thermal expansion or cutting/inertial forces. Figure 4-2 shows a detailed cutaway view of the UFTS.

Figure 4-3 gives an overall view of all flux-carrying components in the Ni-Fe based design: the armature core, the front core, the middle cores, and the rear core.

In the next sections, we will describe in parallel two ultra fast tool servo designs using two magnetic materials: 25 μm Ni-Fe tape wound core and SM2 powder pressed iron as characterized in Chapter 3. Ni-Fe tape wound cores have much lower eddy current and thus allow higher flux density, especially in high-frequency operation. In contrast, magnetic cores made of SM2 powdered iron have much higher eddy current loss, but they are very easy to manufacture so as to allow quick concept proof. We set

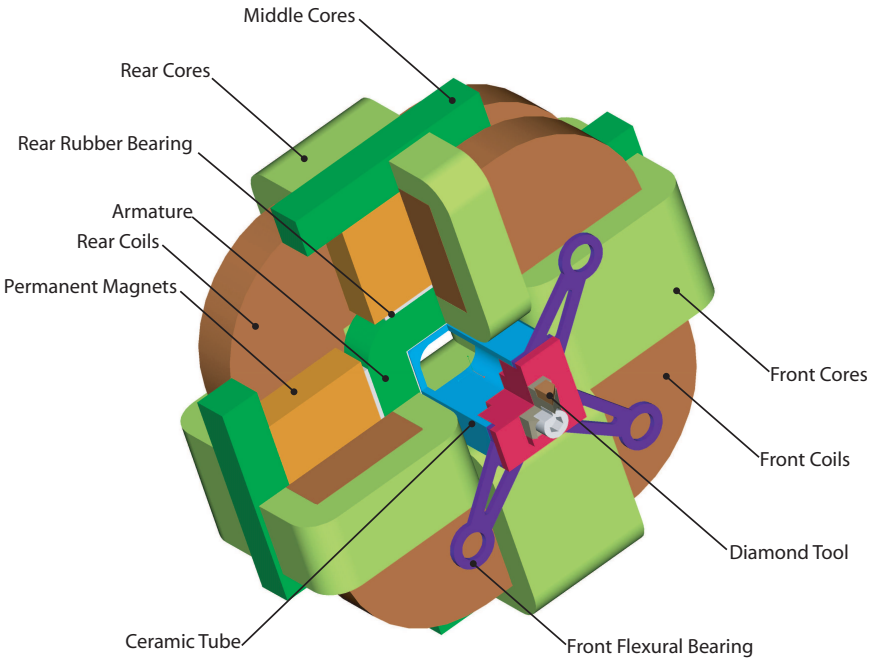


Figure 4-3: 3-D Magnetic Path of the Ni-Fe based UFTS.

main design parameters to optimize the performance of the Ni-Fe based UFTS, as the Ni-Fe material can yield better performance than the SM2 material. To maximize the number of interchangeable parts between the two designs, we kept the same design parameters for the SM2-based UFTS, although these parameters are not optimized for SM2-based UFTS. We built a SM2-based UFTS to allow quick concept proof, and not for maximum performance demonstration.

4.3 Moving Assembly

The key assembly of the UFTS is the moving assembly as shown in Figure 4-4, which includes an armature made of soft magnetic materials, a structural backbone made of ceramic materials, a flexural front bearing, and a rear bearing made of four rubber pads interposed between the armature and the permanent magnets. In order to simultaneously achieve high acceleration, high control bandwidth, and chatter free cutting, we need to ensure light weight, good structural dynamics, and high dynamic stiffness in the moving assembly.

In the moving assembly, the most important part is the armature, which will be

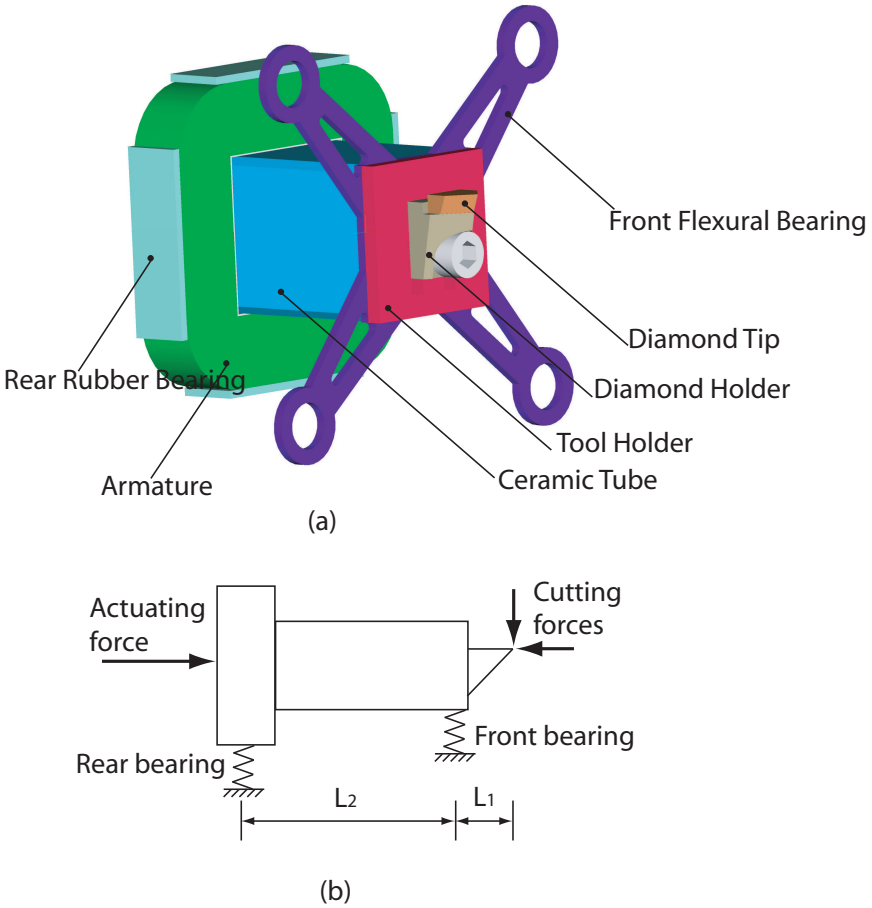


Figure 4-4: Moving Assembly. (a) ProE model of the moving assembly. (b) Simplified dynamical model.

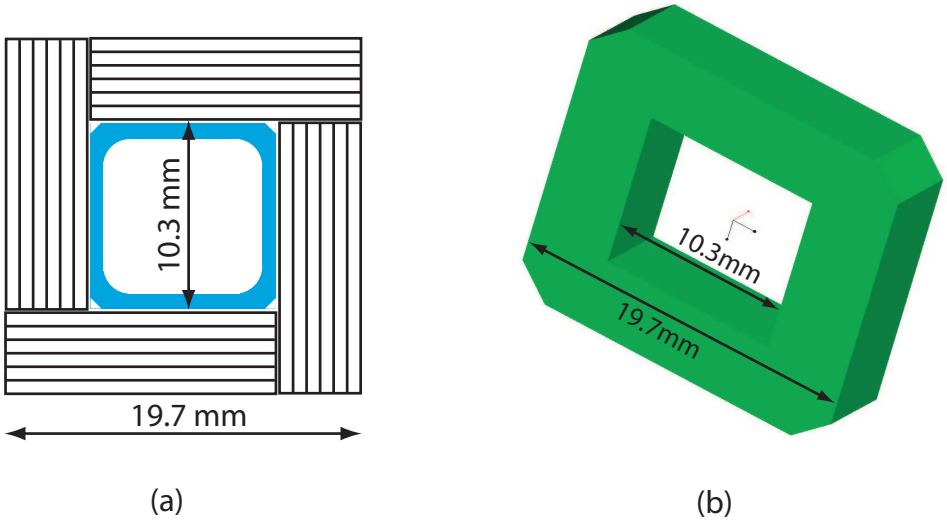


Figure 4-5: Two armature designs. (a) Ni-Fe tape laminated armature. (b) SM2 powder-pressed iron armature.

discussed first. Then we will go through the remaining parts of the moving assembly.

4.3.1 Armature

The armature designs using SM2 and Ni-Fe are different. In the armature made of Ni-Fe wound tape cores, four pieces of laminated cores will be epoxy-glued together around the ceramic tube as shown in Figure 4-5(a). In the SM2 armature design, the armature is a monolithic piece as shown in Figure 4-5(b). The inner square hole of both armatures is 0.2 mm wider than the ceramic tube so that an 0.1 mm epoxy bonding joint is allowed on each side.

In the armature design, we mainly consider the acceleration performance. Another consideration is the dynamic stiffness requirement. The minimum axial dynamics stiffness seen at the tool-tip happens at the loop transmission unity-gain crossover frequency ω_c of the motion servo loop. The unity-gain frequency stiffness K_c can be approximated as $K_c = m\omega_c^2$ based on the total moving mass m. Even though the capacitance probe position sensor is claimed to have 100 kHz bandwidth, it is very difficult to push ω_c to more than 10 kHz mainly due to the significant phase drop from the Butterworth filter inside the ADE capacitance-probe position sensor, because this

phase drop cannot be recovered without amplifying the high frequency noise. As a result, maintaining a certain minimum amount of total inertial mass on the moving assembly is the only way to ensure high dynamic stiffness.

The armature design is an iteration process. We start with an armature cross-section design of t=4 mm thick on the basic of the acceleration results in Chapter 3. To keep reasonable stiffness, the armature is made square in cross-section with a height of 4 mm. We then work around this initial design to check the maximum acceleration, total moving mass, structural dynamics. We end up with the pole face design shown in Figure 4-5. This design has an armature pole face area of $A_a = 295$ mm^2. In order to reduce the leakage flux (flowing directly from permanent magnet into stator cores), the stator pole face area A is set at $A = 200$ mm^2. Correspondingly, we set the permanent magnet width at $W_{pm} = 15$ mm, slightly less than the armature width of 19.7 mm. Next, we will describe the armature thickness optimization based on these design parameters.

In addition to the acceleration performance curves for CUFM given in Chapter 3, we need to consider another two facts in evaluation of the UFTS acceleration: (1) the force generation armature pole face area A (equal to the stator pole face area) is slightly less than the total armature face area A_a; (2) there are additional mass loads m_l attached to the armature. As a result, the maximum achievable acceleration on the moving assembly is

$$\begin{aligned} a_{\max} &= \frac{F_{\max}}{m} \\ &= \frac{B_{\text{sat}}^2 A\eta(2-\eta)}{2\mu_0}\frac{1}{A_a\rho_a t + m_l}, \end{aligned} \tag{4.1}$$

where B_{sat} is the material saturation flux, ρ_a is the armature mass density, m_l =3.3 gm is the mass of additional inertial loads (moving mass other than the armature itself) as listed in Table 4.1, t is the armature thickness and also the permanent magnet thickness. The biasing strength η can be related to the armature thickness t as

$$\eta = \frac{\overline{B}}{B_0}$$

Table 4.1: Additional mass loads on the moving assembly.

Part	mass density [gm/cc]	volume [mm^3]	mass [gm]
Diamond tip	3.5	4.6	0.016
Tool shank	17.7	30	0.531
Tool shank screw	7.8	22.5	0.176
Tool holder	2.7	286	0.772
Ceramic tube	3.9	459	1.79
Total			3.3

$$\begin{aligned} &= \frac{\frac{A_{pm}B_r\alpha_L}{2A}}{\frac{B_{\text{sat}}}{2}} \\ &= \frac{A_{pm}B_r\alpha_L}{AB_{\text{sat}}} \\ &= \frac{4W_{pm}tB_r\alpha_L}{AB_{\text{sat}}}, \end{aligned} \tag{4.2}$$

where W_{pm} is the permanent magnet width, B_r is the permanent magnet remanence, and α_L is leakage factor.

Making use of (4.1), Figure 4-6 plots the maximum achievable acceleration versus the armature thickness t for the Ni-Fe based UFTS. The over-mass factor α_M is defined as the ratio of the total moving mass over the force-generating mass:

$$\alpha_{\text{M}} = \frac{A_a\rho_a t + m_l}{A\rho_a t}. \tag{4.3}$$

Using Ni-Fe lamination with an armature thickness $t = 4$ mm, as indicated by red dots in the figure, we can achieve an over-mass factor α_M =2.1, a biasing strength η=0.95, a total moving mass of 11.2 gm, and a maximum acceleration $a_{\max}$ =1200 G. Although a $t = 2.5$ mm design can result in a slightly higher maximum acceleration of 1300 G, the resonance frequencies of the thinner moving assembly will be significantly reduced and the dynamic stiffness will also decrease because of the reduction in the total mass. What's more, considering the uncertainty of the void factor of the wound tape material, the armature thickness is set at $t = 4$ mm to achieve high mechanical reliability.

For the SM2-based UFTS design, similar performance curves are shown in Fig-

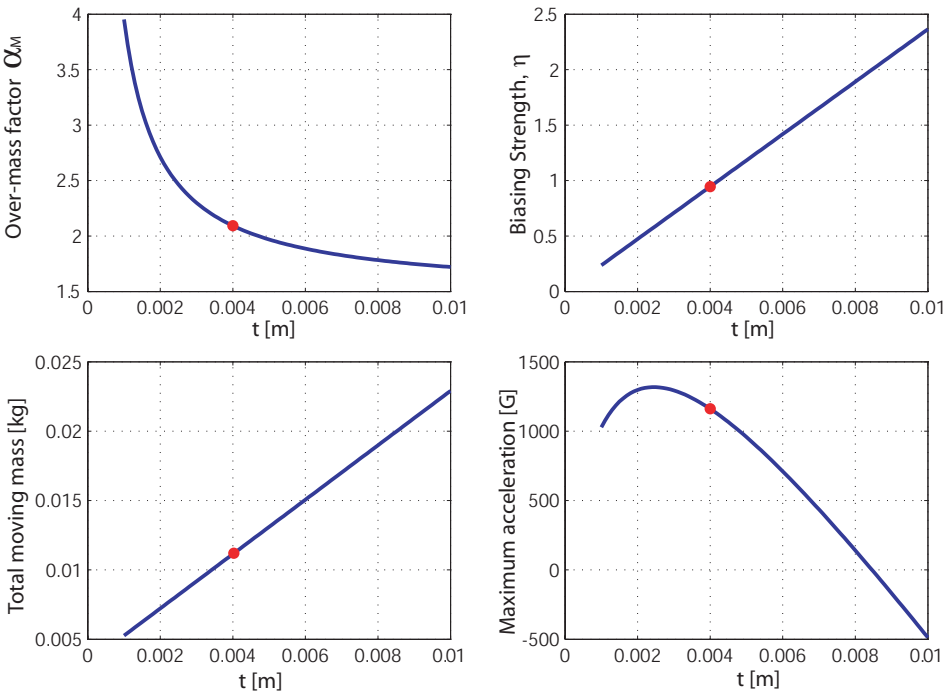


Figure 4-6: The armature thickness design trade off for the Ni-Fe based UFTS. The red dots shows the final design decision of $t = 4$ mm.

ure 4-7. By setting $t = 4$ mm, we can achieve a over-mass factor α_M =1.9, a biasing strength η=0.8 , a total moving mass of 11.5 gm, and a maximum acceleration $a_{\max}$ =1500 G. This value is higher than the Ni-Fe design with the same thickness, because the Ni-Fe has a lower equivalent saturation level than SM2 material.

For these two designs, according to the analysis in Chapter 3 the basic characteristics of the UFTS' are calculated as listed in Table 4.2.

Table 4.2: Performance of the ultra fast tool servo design

Material	X_0 [μm]	B_0[T]	N [turn]	I_0[A]	F_0[N]	λ_0[V-ms]	$V_{\max}$(10kHz)[V]
Ni-Fe	100	0.63	6×4	4.2	127	3	47×4
SM2	100	0.75	6×4	4.9	179	3.6	56×4

4.3.2 Bearings

As shown in Figure 4-4, the motion system is built upon a ceramic tube structural backbone. The ceramic tube is supported by two bearings: a flexure bearing in the front and a rubber bearing in the back. The ratio of L_2 (the distance between the

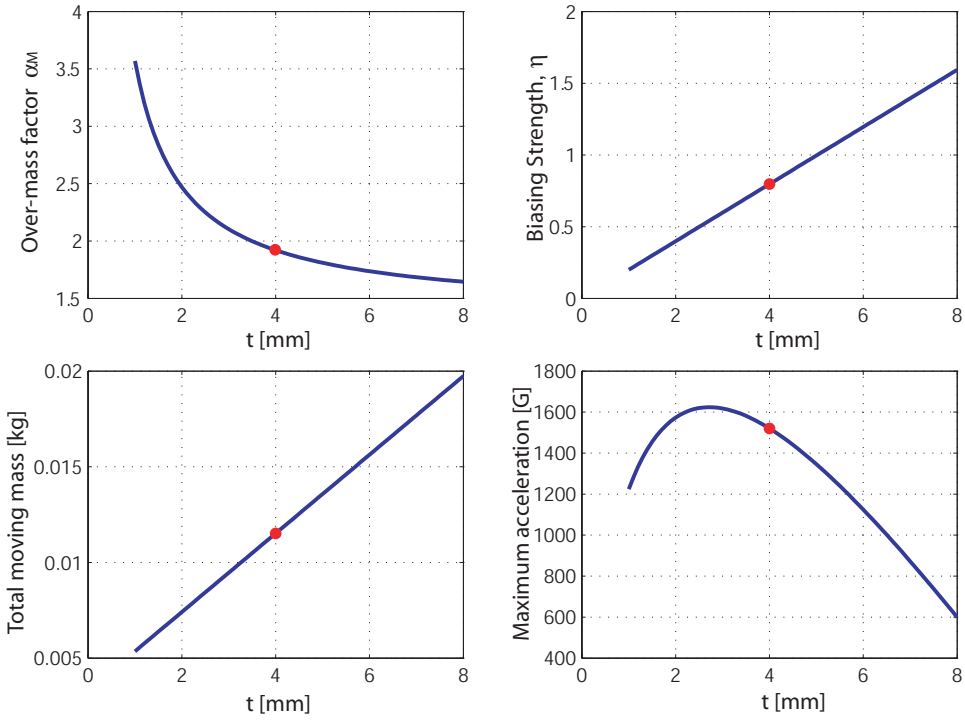


Figure 4-7: The armature thickness design trade off for the SM2-based UFTS. The red dot shows the final design decision of $t = 4$ mm.

tool tip and the front bearing) over L_1 (the distance between the rear bearing and the front bearing) is 2.7.

The lateral stiffness at the tool tip K_{tl} is related to the front bearing lateral stiffness K_{fl} and the rear bearing lateral stiffness K_{rl} as

$$\begin{aligned} K_{tl} &= (K_{rl}\frac{L_2^2}{L_1^2})//(K_{fl}\frac{L_2^2}{(L_1+L_2)^2}) \\ &= 7.3K_{rl}//0.53K_{fl}. \end{aligned} \tag{4.4}$$

Here we have used the parallel sysmbol // to indicate $a//b = ab/(a+b)$; this operation is appropriate since the two stiffness are mechanically in series. The key point here is that the stiffness of the rear bearing is about 15 times less important than that of the front bearing, and thus can be designed to be much less than that of the front bearing.

In axial direction, the axial stiffness at the tool tip K_{ta} is related to the front

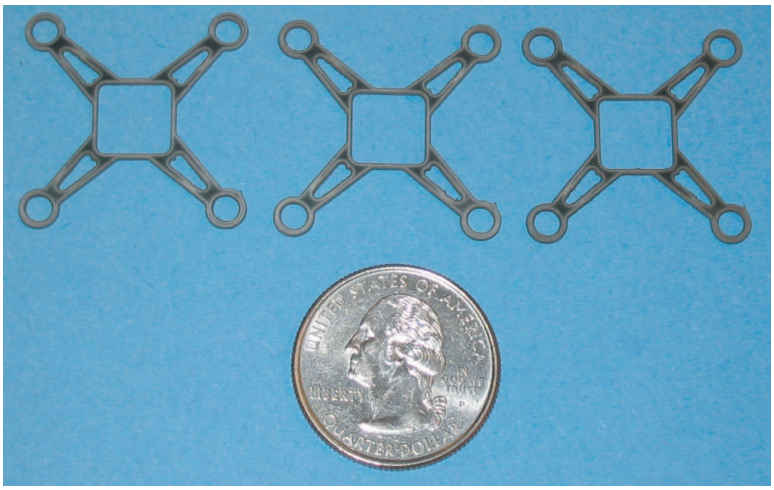


Figure 4-8: Front flexure bearings made of spring steel sheets.

bearing axial stiffness K_{fa} and the rear bearing axial stiffness K_{ra} as

$$K_{ta} = K_{ra} + K_{fa}. \tag{4.5}$$

The two bearings contribute equally to the axial stiffness. As we ensure the axial stiffness of our UFTS mainly by the active control of the moving assembly, this passive axial stiffness K_{ta} should be design as little as possible in order to minimize the nonlinear force from the rubber pad bearing.

Front Flexure Bearing

Figure 4-8 shows the front flexure bearing design with 8 legs in the shape of a letter "X", which provides in-plane lateral stiffness and rotational stiffness. This flexure is cut from blue-tempered spring steel sheets using a water-jet machine. Assuming a clamped-clamped boundary condition, the lateral stiffness K_{fl} and the axial stiffness K_{fa} (in the axial direction of the ceramic tube) of the front bearing is calculated as

$$K_{fl} = 8\frac{\sqrt{2}}{2}\frac{Et_f w_f}{L_f} \tag{4.6}$$

$$\text{and } K_{fa} = 8\frac{Et_f^3 w_f}{L_f^3}, \tag{4.7}$$

where t_f is the flexure bearing thickness, $E = 200$ GPa is the Young's modulus of steel, and each leg has a width of w_f=1.2 mm and a length of L_f =4.5 mm.

In the first assembled prototype, a steel sheet of thickness t_f=0.64 mm is used and the resulting stiffness is $K_{fl} = 193$ N/μm and $K_{fa} = 5.5$ N/μm. This design yields unnecessarily high stiffness in lateral direction. In future prototypes, 0.3 mm thick steel should be used instead to reduce the axial stiffness, and correspondingly the stiffness would be 90 N/μm in lateral direction and 0.56 N/μm in axial direction.

Rear Rubber Bearing

Rubber sheets have been proposed to be used as bearings for more than 30 years [20, 71, 38]. A rubber sheet bearing has the property of very high normal stiffness K_z (in the normal direction) and very low shear stiffness $K_{x,y}$ (along the shear direction in the rubber plane). The shear stiffness K_z and the normal stiffness $K_{x,y}$ under a uni-axial strain assumption[1] are calculated according to Hook's Law as

$$K_z = \frac{E(1-\gamma)A_r}{(1-2\gamma)(1+\gamma)T_r}, \tag{4.8}$$

$$K_{x,\,y} = \frac{EA_r}{2(1+\gamma)T_r}$$

$$\frac{K_z}{K_{x,\,y}} = \frac{2(1-\gamma)}{(1-2\gamma)}, \tag{4.9}$$

where E is the Young's modulus, γ is the Poisson ratio, A_r is the rubber pad area, and T_r is the rubber pad thickness. As γ of rubber material is very close to 0.5, this stiffness ratio can be further simplified as

$$\frac{K_z}{K_{x,\,y}} = \frac{1}{1-2\gamma}, \tag{4.10}$$

This high-stiffness ratio result shows that the rubber pad bearing can be very stiff in the normal direction and very soft in the shear direction, which is a desirable

[1]When the X and Y dimensions of the rubber pad are much larger than the Z dimension, such that the rubber has no way to escape in X and Y direction under stress in Z direction. Hence, the strain in X and Y will be $\varepsilon_x = \varepsilon_x = 0$

characteristic for flexural bearing. For instance, assuming γ=0.499, the stiffness ratio $K_z/K_{x,y}$ is 500. In reality, the uni-axial strain assumption does not always hold. To take into account the deviation of the actual rubber pad from the ideal uni-axial strain model, the shape factor $S.F.$ is used in [38, 20]

$$S.F. = \frac{\text{load area } A_L}{\text{bulge area } A_B}. \tag{4.11}$$

When $S.F. = \infty$, the rubber will satisfy the uni-axis strain condition. For rubber bearings with finite shape factors, the rubber stiffness can be obtained by FEA, analytical solution, or experimental measurement. For simplification of calculation, the compression modulus E_c is defined in [38] as

$$E_c = \frac{K_z T_r}{A_r}. \tag{4.12}$$

As a result, the relation between E_c and $S.F.$ can be established for rubber materials with various hardness as shown in Figure 4-9, adapted from [38].

In our UFTS designs, we construct the rear bearing with four rubber pads as shown in Figure 4-4. Figure 4-10 shows the dimension of one rubber pad used in the rear bearing. The shape factor of each rubber pad is calculated as

$$S.F. = \frac{W_r L_r}{2(W_r + L_r)T_r} = 3.2. \tag{4.13}$$

By using the curves in Figure 4-9, the compression modulus and shear modulus can be determined as E_c=40 Mpa and G=0.83 Mpa, respectively, for rubber materials of 55A durometer[2]. Consequently, the total lateral stiffness and the total axial stiffness of the four rear rubber bearings are calculated as

$$K_{ra} = 4G\frac{W_r L_r}{T_r} = 0.4 \text{ N}/\mu\text{m}, \tag{4.14}$$

$$K_{rl} = 2G\frac{W_r L_r}{T_r} + 2E_c\frac{W_r L_r}{T_r} = 10 \text{ N}/\mu\text{m}. \tag{4.15}$$

[2]This hardness of 55A durometer is between the hardness of a pencil eraser and that of a vehicle tire.

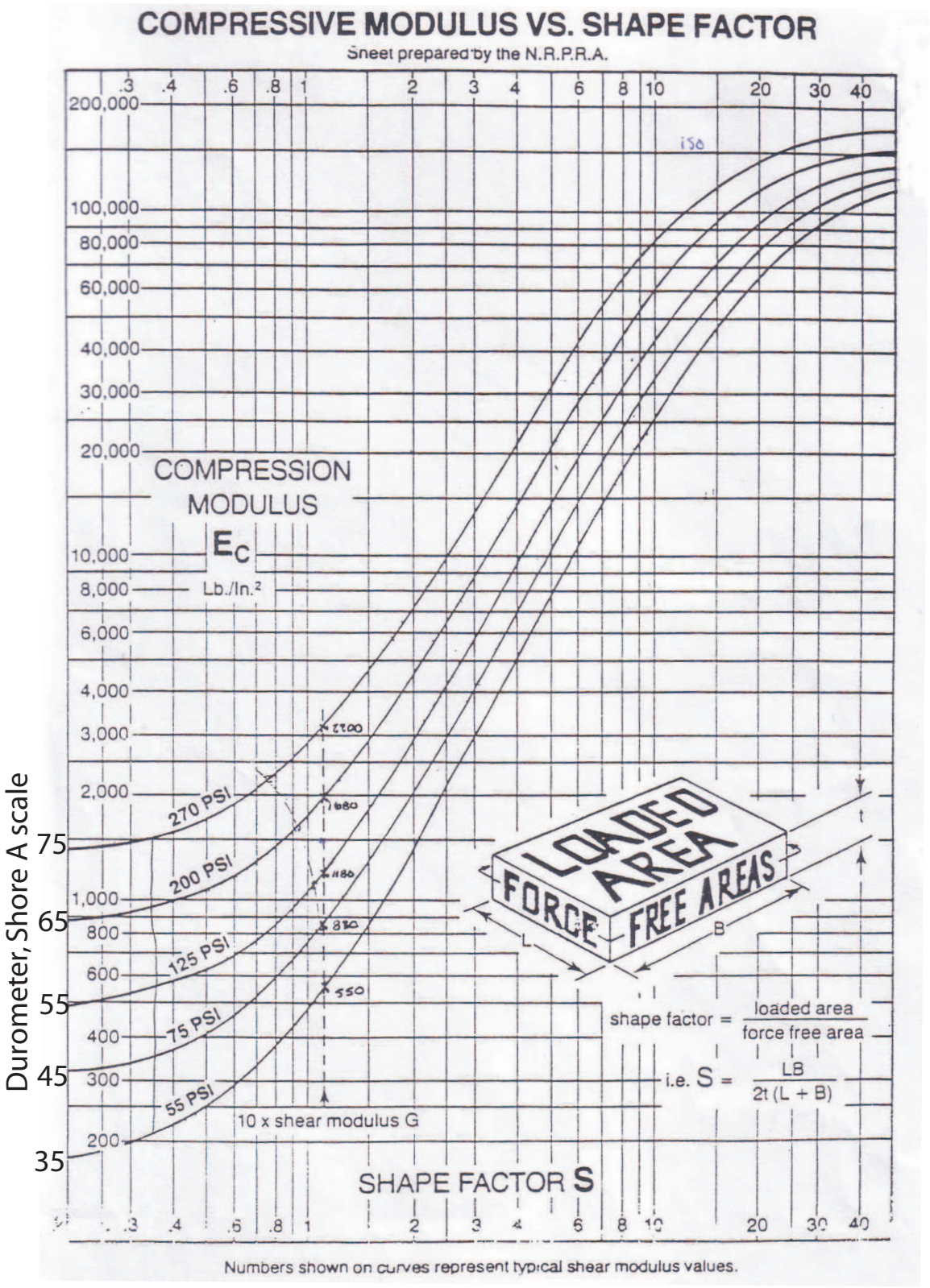


Figure 4-9: Compression stiffness curves (assuming $\gamma = 0.4995$). Adapted from "Engineering Design with Natural Rubber" [38].

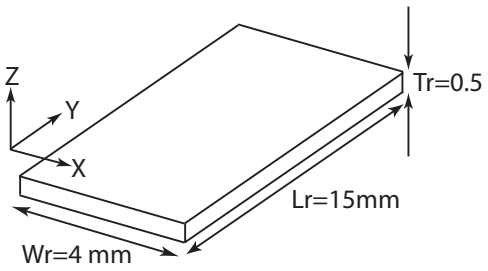


Figure 4-10: Rubber pad used in the rear bearing.

Using(4.4) and (4.5), we can calculate the lateral stiffness and the axial stiffness at the tool tip as

$$K_{tl} = \begin{cases} 42 \text{ N}/\mu\text{m}, & t_f = 0.64 \text{ mm} \\ 28 \text{ N}/\mu\text{m}, & t_f = 0.3 \text{ mm} \end{cases}, \tag{4.16}$$

$$K_{ta} = \begin{cases} 5.9 \text{ N}/\mu\text{m}, & t_f = 0.64 \text{ mm} \\ 0.96 \text{ N}/\mu\text{m}, & t_f = 0.3 \text{ mm} \end{cases}. \tag{4.17}$$

The front bearing designs of both t_f=0.64 mm and t_f=0.3 mm meet the lateral stiffness requirement of 20 N/μm listed at the beginning of this chapter. The soft axial stiffness from the bearings is preferred, because a softer passive stiffness in axial direction means less nonlinear force from the rubber bearings to the UFTS control loop.

All the above analyses of rubber bearings are based on static operation at room temperature. The compression modulus E_c and G will increase when the operation frequency is increased or when the temperature is reduced. Fortunately, there exists the inverse relationship between temperature and frequency effects[61]. As a result, the stiffness G can be plotted as a function of single variable of $f\alpha_T$. In this thesis, I haven't considered further this dynamic property of rubber as the rubber materials commercially available usually don't come with this temperature and frequency dependent data. We kept in mind that there would be stiffness increase at high frequencies, and this increase can be countered at high temperature. According to Dr. Peter Nelson formerly of TMC Corp., for highly damped materials, the stiffness increase with frequency could be a factor of 3. Augusto Barton of our lab found

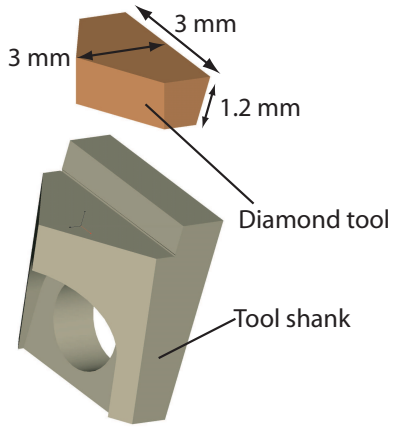


Figure 4-11: Diamond cutting tool and tool shank.

in his test experiment on Silicon rubber of durometer 50A that the stiffness at high frequencies (1-10 kHz) is 100 times higher than that at DC. In our designed UFTS, we select Neoprene with a working temperature of -30°C $\sim$ 104°C and durometer of 55A as the rear bearing material. There exist other materials suitable for this operation temperature range (such as Buna-N, Neoprene, Vinyl, and Viton), but I didn't explore these alternative solutions due to time limit.

4.3.3 Diamond Tool

Figure 4-11 shows the diamond tool and its tool shank. The diamond tip is 3 mm long, 1.2 mm tall, and 3 mm wide. The diamond tip width is intentionally designed larger than our cutting requirement, in order to increase the stiffness of the tool shank and the tool tip sub-assembly, because diamond has a very high specific stiffness and its cost is insensitive to the diamond size. The diamond cutting edge geometry design as shown in Table 4.3 is adapted from a tool tip design by Montesanti [59].

The diamond shank is made of Super Chatter Free[3] Tungsten-based alloy in order to ensure high supporting stiffness. A #1-72 screw is used to install the tool shank on the tool holder, which is made of aluminum 4104 to ensure light weight and high strength on the thread.

Figure 4-12 shows the manufacturable surfaces as constrained by the tool geom-

[3]A trade name of Mi-Tech Metals.

Table 4.3: Diamond tool geometry parameters.

No.	Geometry	Dimension
1	Nose radius	0.03±0.003 mm, conical
2	Primary clearance angle	15°
3	Waviness	<0.2 μm
4	Tool nose sweep angle	>50°
5	Cutting edge radius	<50 nm
6	Tool edge roughness	<1 nm RMS
7	Rake angle	0°
8	Included angle	90°

etry. On the top view, the sweep angle θ_s determine the maximum derivative of the part surface profile $Z(X)$ as

$$\left|\frac{dZ}{dX}\right|_{\max} < \tan\left(\frac{\theta_s}{2}\right). \tag{4.18}$$

The tool nose radius R_n set the maximum curvature of $Z(X)$ on the top view as

$$\frac{Z''}{(1+Z'^2)^{\frac{3}{2}}} < \frac{1}{R_n}. \tag{4.19}$$

On the side view, the clearance angle θ_c sets the maximum derivative of the part surface profile $Y(X)$ as

$$\frac{dZ}{dY} < \tan(\theta_c). \tag{4.20}$$

According to these constraints, the tool with geometric parameters listed in Table 4.3 can cut 2 μm peak-to-valley sinusoidal surface with 40 μm pitch on the top view and 30 μm pitch in the side view.

4.3.4 Ceramic Tube

Figure 4-13 shows the ceramic tube, which is the backbone of the moving assembly. Two candidate materials (silicon carbide and alumina, both from Coorstek[4]) are

[4]http://www.coorstek.com/.

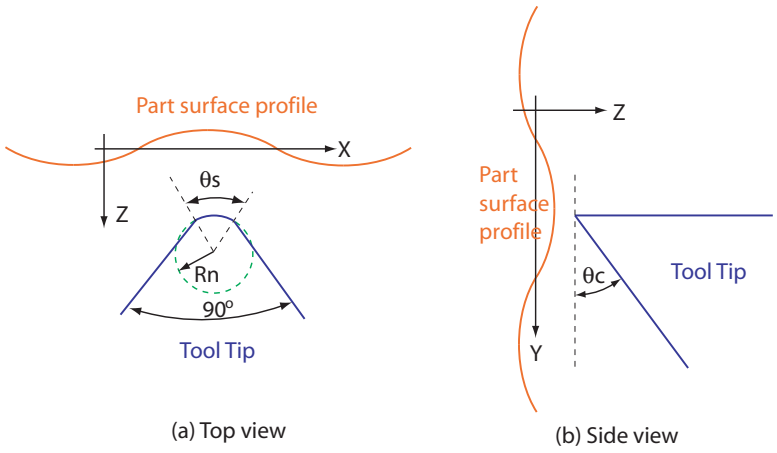


Figure 4-12: Diamond tool geometries and cutting surfaces.

considered. Although silicon carbide (SC-DS of Coorstek, mass density 3.15 g/cc, and Young's modulus 410 Gpa) has 40% higher specific stiffness and lower mass density than alumina (AD-99.5 of Coorstek, mass density 3.9 g/cc, and Young's modulus 370 GPa), alumina is selected in the final design, mainly due to the difficulty of manufacturing such a small part from silicon-carbide. Coorstek provided the materials and manufactured the part to our specifications. The center hole is drilled off first, and then all four surfaces are ground flat to dimensional tolerances. The tube length and wall thickness are picked based on FEA structural dynamics analysis to ensure that the tube itself does not contribute flexible modes within 50kHz. The wall thickness of the tube is chosen to be at 0.75 mm, which is the minimum value acceptable to Coorstek for manufacturing purposes. The length is set at 14.8 mm to allow sufficient space to accommodate stator the cores and coil windings. In the next section, we discuss the structural dynamics and the sound transmission delay in the ceramic tube.

4.3.5 Structural Dynamics Analysis

Structural analysis plays an important role for the design of any mechatronic devices. It is important to design both mode frequencies and mode shapes to be compatible wit the desired controlled performance. In our device, the critical structural component is the moving assembly.

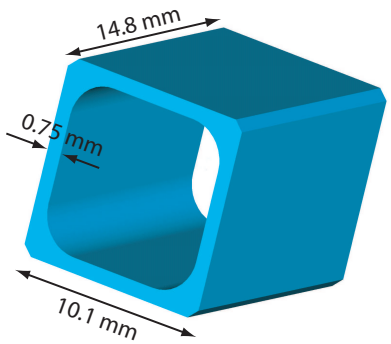


Figure 4-13: Ceramic tube design.

There are some related prior art on structural control in the literature. For concentrated force actuating or concentrated actuator sensing, Rankers has investigated approaches that minimize or even eliminate the destabilizing effect of arbitrary modes [69], by locating driving force at a node of the mode and/or shifting sensor towards a node of mode. He also mentioned the idea of applying additional forces or additional sensors to make the specific mode not to be excited or measured. In a more systematic way, Ming-chih Weng [85] developed actuating averaging/sensor averaging methods and investigated how to design the weighting function W_a and W_s to create spatial filtering effects to attenuate undesired resonance modes.

In order to represent the distributed actuating/sensing on such flexible structures, we present the concept of Actuating Ability and Sensing Ability, which are general evaluation criterion on how excitable a specific mode is. These Acuating Ability and Sensing Ability concepts can be generally applied to distributed actuating/sensing on structures. Further, we apply these concepts to analyze the moving assembly structural dynamics.

Actuating Ability and Sensing Ability

When the frequency region of interest is high enough, mechanical systems need to be modelled as flexible structures rather than as rigid bodies. What's more, for flexible structures with short wavelengthes of high-frequency modes, modelling of force-actuating and position-sensing should also take into account the spatial actuating/sensing distribution effect rather than modelling these as single-point position

sensing.

We express the distributed actuating force as the product of a temporal function and a spatial function, under the assumption that the flexible structure motion does not change the surface force distribution. Consequently, the actuating force density can be modelled as a separable expression

$$f(t,x,y,z) \;=\; F(t)W_a(x,y,z), \tag{4.21}$$

where $f(t,x,y,z)$ is the actuating force surface density, $F(t)$ is the total actuating force and $W_a(x,y,z)$ is the force weighting function representing force distribution. Based on this definition, we have

$$F(t) \;=\; \iint_{S_a} f(t,x,y,z)dS_a, \tag{4.22}$$

where S_a is the force-distributed area. Plugging (4.21) into (4.22) yields the constraint condition

$$\iint_{S_a} W_a(x,y,z)dS_a \;=\; 1, \tag{4.23}$$

Similarly for distributed position sensing, under the assumption that the flexible structure motion does not change the displacement sensing, we can model the displacement output as

$$X(t) \;=\; \iint_{S_s} X(t,x,y,z)W_s(x,y,z)dS_s, \tag{4.24}$$

where $X(t)$ is the measured total position feedback, $X(t,x,y,z)$ is the motion of the structure, S_s is the sensed surface, and $W_s(x,y,z)$ is the weighting function representing sensing distribution with the constraint

$$\iint_{S_s} W_s(x,y,z)dS_s \;=\; 1. \tag{4.25}$$

Consequently, for a flexible structure with distributed actuating and sensing as

in the above analysis, according to classical structural dynamics [49] the transfer function from $F(t)$ to $X(t)$ can be expressed as:

$$\begin{aligned} G(s) &= \frac{X(s)}{F(s)} \\ &= \sum_{n=1}^{\infty} \frac{\iint_{S_a} \Phi_n(x_a,\ y_a,\ z_a) W_a(x_a,\ y_a,\ z_a) dS_a \iint_{S_s} \Phi_n(x_s,\ y_s,\ z_s) W_s(x_s,\ y_s,\ z_s) dS_s}{M_n(s^2 + 2\zeta_n \omega_n s + \omega_n^2)} \end{aligned} \tag{4.26}$$

where $G(s)$ is the transfer function, $X(s)$ and $F(s)$ are the Laplace Transforms of $X(t)$ and $F(t)$ respectively, Φ_n is the n-th modal shape, ω_n is the n-th modal frequency, ζ_n is the n-th modal damping ratio, and M_n is the n-th modal mass. The modal mass M_n is related to the modal shape Φ_n and the structural mass density ρ as

$$M_n = \iiint_V \Phi_n^T \Phi_n \rho dv, \tag{4.27}$$

where V is the whole structural volume.

Here, we introduce to the notation of the Actuating Ability A_n and the Sensing Ability S_n of the n-th mode as

$$A_n = \frac{\iint_{S_a} \Phi_n(x_a,\ y_a,\ z_a) W_a(x_a,\ y_a,\ z_a) dS_a}{\sqrt{M_n}}, \tag{4.28}$$

$$S_n = \frac{\iint_{S_s} \Phi_n(x_s,\ y_s,\ z_s) W_s(x_s,\ y_s,\ z_s) dS_s}{\sqrt{M_n}}. \tag{4.29}$$

With this definition, both A_n and S_n have physical units of $\mathrm{Kg}^{-0.5}$, but they are independent of the mode shape scalings. For instance, using Φ_n or $2\Phi_n$ as the n-th mode shape yields the same values of A_n and S_n.

Based on A_n and S_n, the transfer function in (4.26) can be simplified as

$$\begin{aligned} G(s) &= \frac{X(s)}{F(s)} \\ &= \sum_{n=1}^{\infty} \frac{A_n S_n}{s^2 + 2\zeta_n \omega_n s + \omega_n^2}. \end{aligned} \tag{4.30}$$

We can see that the actuating ability A_n and the sensing ability S_n of the n-

th mode determine how much this mode will show up in the system overall transfer function $G(s)$. If a mode has small values of A_n or S_n, then that mode will not appear strongly in the system frequency response, and will not be a serious problem with respect to stabilizing controller design. If $A_n S_n$ is significant, then we call mode-n as an excitable mode. The actuating ability and sensing ability definitions are related by are very different from the standard controllability/observability as defined in control theories, which correspond to whether a matrix is full rank or singular, and thus cannot indicate the continuous nature represented by the Actuating/Sensing Abilities as defined above.

We refer to the product of Actuating Ability and Sensing Ability as the excitability of the associated mode. In structural design for a mechatronic system, we can evaluate the excitability of vibration modes to identify the key destabilizing modes by finite element analysis. Based on these results, structure design can be modified to reduce the excitability of key destabilizing modes if necessary. Next we use the Sensing Ability and Actuating Ability to evaluate how much a specific mode influences the transfer function in the moving assembly design.

Moving Assembly Structural Analysis

With these analytical fundamentals in hand, we come back to the structural design of the moving assembly in the UFTS. The electromagnetic actuating force can be modelled as a uniform pressure acting on the armature pole face opposite to a given stator pole face, and thus W_a is a constant $W_a = 1/S_a$. The position sensing can be modelled as the displacement output from a concentrated point (W_s is a singular function), because the capacitance probe active area is only 1 mm in diameter, which is much less than the wave-lengthes of the concerned modes of interest. As a result, the sensing ability expression for our UFTS can be simplified to

$$S_n = \frac{\Phi_n(x_s,\ y_s,\ z_s)}{\sqrt{M_n}}, \tag{4.31}$$

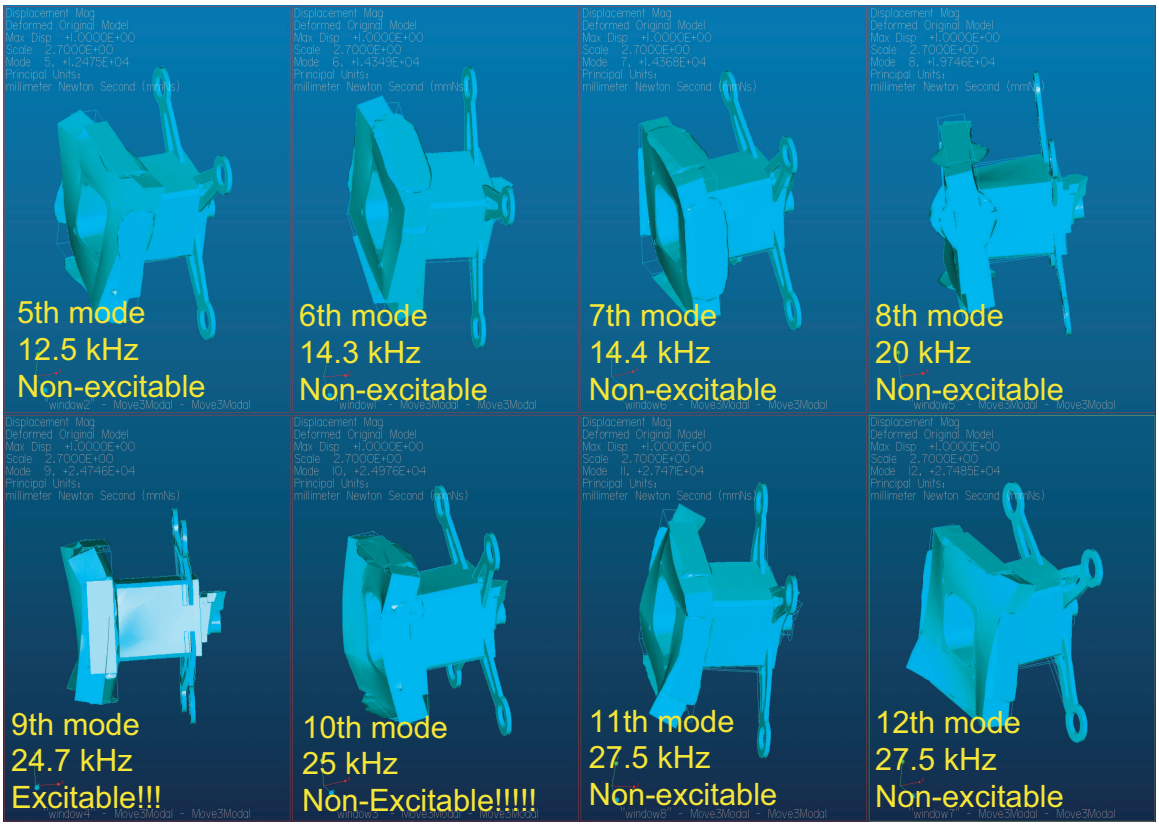


Figure 4-14: Flexible modes of the moving assembly of Ni-Fe wound-tape cores. This armature is composed of four pieces as shown in Figure 4-5, each piece is stacked up with 25 μm-thick Ni-Fe layers and 5-μm-thick epoxy layers alternatively. As suggested by Rick Montesanti and Professor Nayfeh, an orthotropic homogenous model is used to represent each segment of orthotropic non-homogenous core by mechanical property equivalence at the overall size. The joints between two segment of cores or between the armature and the ceramic tube are modelled by an epoxy layer of 0.1 mm thickness.

where $(x_s,\ y_s,\ z_s)$ is the point on the moving assembly that the capacitance position sensor is measuring.

For the moving assembly using Ni-Fe wound-tape cores, Figure 4-14 shows a finite element analysis of the moving assembly. By simple inspection, we can see that all flexible modes except the 9th mode exhibit either an anti-symmetric mode shape on the actuating area or no displacement along the X-direction at the measured point. As the Actuating Ability of an anti-symmetric mode shape ($A_n = W_a \iint_{S_a} \Phi_n(x_s,\ y_s,\ z_s) dS_a/\sqrt{M_n}$) is zero and the Sensing Ability of a nodal point ($S_n = \Phi_n(x_s,\ y_s,\ z_s)/\sqrt{M_n}$) is zero, these modes have very low or zero values of

AnS_n. The first flexible mode showing up significantly on the system transfer function will be the 9th-mode with a resonance frequency of 25 kHz.

Similarly, Figure 4-15 shows the finite element analysis of the moving assembly using the SM2 core. The first excitable mode is the 12th mode with a resonance frequency of 45 kHz. This frequency is much higher than the first excitable mode in the moving assembly of Ni-Fe cores, mainly because the SM2-based armature is a monolithic sintered piece, in comparison to the four pieces of laminated composite cores in the Ni-Fe armature design.

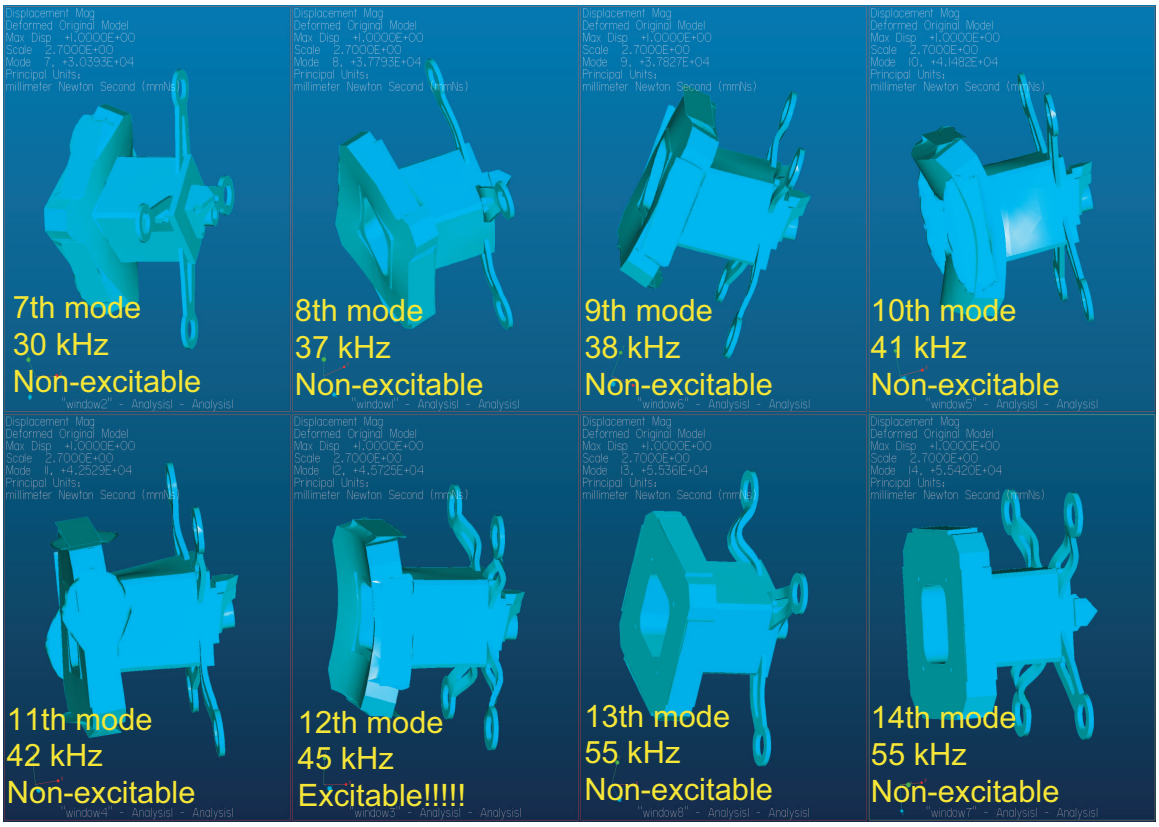


Figure 4-15: Flexible modes of the moving assembly of the SM2 armature core.

The existence of the 9th mode on the Ni-Fe moving assembly (or the 12th mode on the SM2 moving assembly) in the system open loop dynamics imposes a limit on the achievable unity-gain cross-over frequency of the compensated loop transmission.

One additional structural dynamic property that we need to take into account is the transmission time delay from the actuated area to the sensed area, which is the time required for the pressure wave to travel inside the flexible structures. Specifically

for the moving assembly of our UFTS, the actuated area is on the armature end of the ceramic tube, and the position sensing is at the other end. As a result, the delay τ_d is the time required for the sound to transmit axially through the whole ceramic tube, which is calculated as

$$\begin{aligned} \tau_d &= \frac{L}{V} \\ &= L\sqrt{\frac{\rho}{E}} \\ &= 14.6 \times 10^{-3}\sqrt{\frac{3.9 \times 10^3}{370 \times 10^9}} \\ &= 1.5\ \mu\text{s}, \end{aligned} \tag{4.32}$$

where L is the ceramic tube length, V is the sound speed in the alumina tube, E is the Young's Modulus of alumina, and ρ is the mass density of alumina.

In summary, for the purpose of feedback controller design, the system dynamics of the moving assembly can be approximated by a delay unit in series connection to the sum of the rigid body mode(a concentrated mass spring system) and the first excitable flexible mode as

$$G(s) \simeq e^{-\tau_d s}\left(\frac{1}{M(s^2 + 2\zeta_1\omega_1 s + \omega_1^2)} + \frac{A_k S_k}{s^2 + 2\zeta_k\omega_k s + \omega_k^2}\right), \tag{4.33}$$

where M is the total moving mass and k-th mode is the first excitable flexible mode: the 9th mode of the Ni-Fe armature and the 12th mode of the SM2 armature.

4.4 Front Assembly

Figure 4-16 shows the front assembly design in the Ni-Fe based UFTS. The whole front assembly is comprised of a front coil winding, a Ni-Fe based front core, and a stainless steel front frame. The front core is made up of four pieces of "C" core as shown in Figure 4-17. The lamination direction of each "C" core is shown in the bottom right corner. The winding slot area of the front core is 70 mm^2 (14 mm wide and 5 mm deep). Inside this slot, there are two coils and each coil is wound

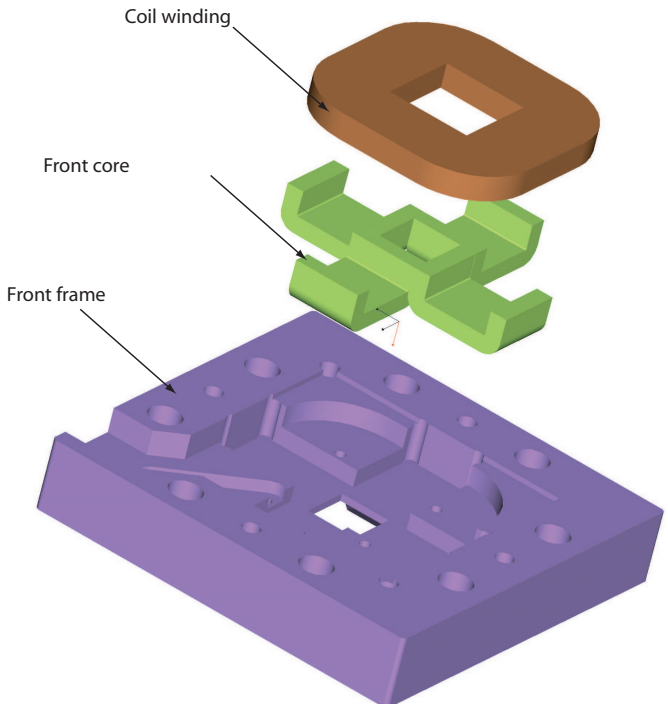


Figure 4-16: Front assembly of the Ni-Fe based UFTS.

with 6 turns. Each coil is made up of 4 parallel SPNSN served litz wire of the type 20/34 [5] to minimize the eddy current effect in the copper wire at frequency region up to 20 kHz. As a result, each coil has a copper cross section area of 3175.2 circular mils [6], which can continuously conduct 4 A DC current or an AC current of 5.7 A based on the rule of thumb 800 circular mils per Ampere (2.5 A/mm^2) as suggested by wire manufactures. According to Table 4.2, we have I_0=4.2 A and thus the coil has maximum normalized AC current i_{max}=1.35 A, which is beyond the required $2 - \eta = 1$. Therefore, this coil design is sufficient to conduct the current required for the maximum AC acceleration of 1200 G. The selection of gauge 34 wire is based on the fast tool servo working frequency of 20kHz, at which the AC impedance of the wire has negligible increase induced by eddy current.

Figure 4-18 shows the corresponding front assembly design using SM2 cores. The

[5] 20/34 SPNSN means that this litz wire is a bundle of 20 strands of AWG-34 size copper wires. Each strand is coated with 9 μm thick Polyurethane-Nylon, and the whole bundle is served with nylon sleeve.

[6] The circular mil is a unit of area standard in coil manufacturing literature. It is equivalent to the area of a circle whose diameter is 0.001 inch.

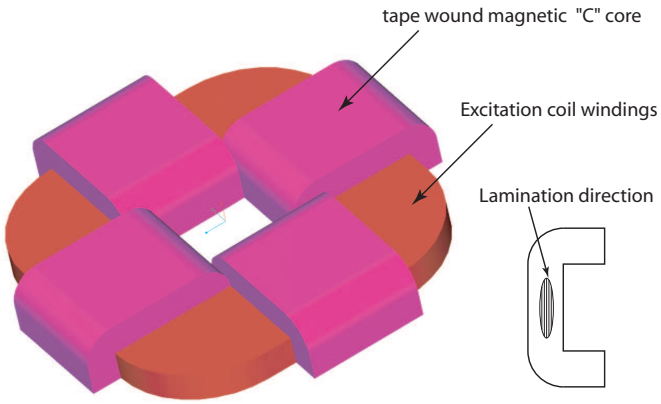


Figure 4-17: Front core design using tape wound material.

front core is a monolithic piece, and its winding slot is the same as that of the Ni-Fe front core, in order to maintain interchangeability between two designs.

4.5 Middle Assembly

Figure 4-19 shows the middle assembly ProE model, which is the same for both design except the middle core material. The middle assembly is made in two halves, which can squeeze together to further pre-load the rubber pads, in addition to the magnetic force between the armature and the permanent magnet. Each half is made of two permanent magnets, two middle cores, and a stainless steel middle frame. The permanent magnet is Ne-Fe-B of 1.2 T remanence, which is rated up to 120^{o}C operation. All permanent magnets are oriented in the way that all north poles point to the inside. The thickness of the middle assembly is 4.2 mm, which is 200 μm thicker than the armature thickness to ensure a 100 μm air gap on each side of the armature. For the Ni-Fe based middle core, the lamination direction is parallel to its attached permanent magnet pole face.

4.6 Rear Assembly

The rear assembly design is almost a mirror to that of the front assembly, except that there exists a round hole for clamping the cap probe and the rear assembly can be

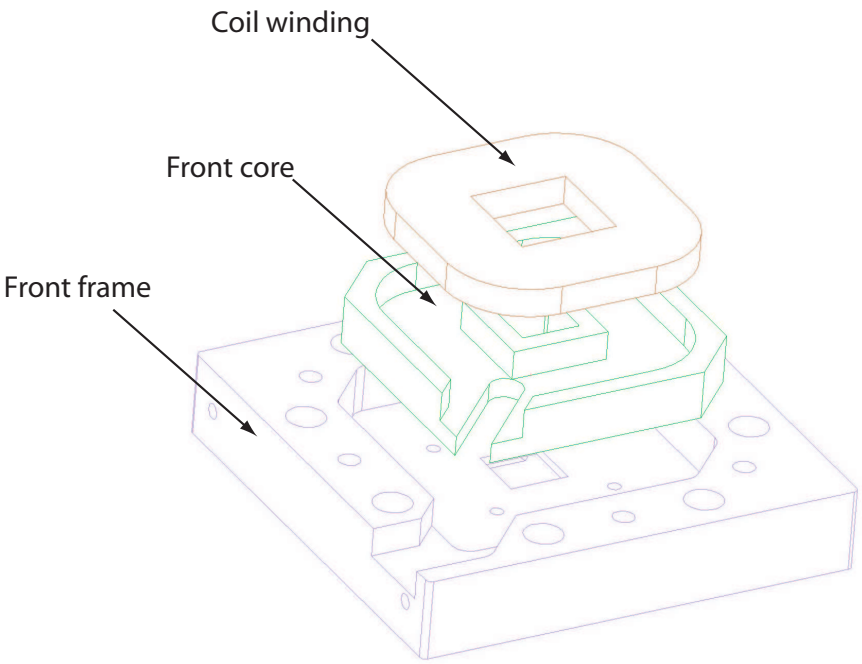


Figure 4-18: Front assembly of the SM2-based UFTS.

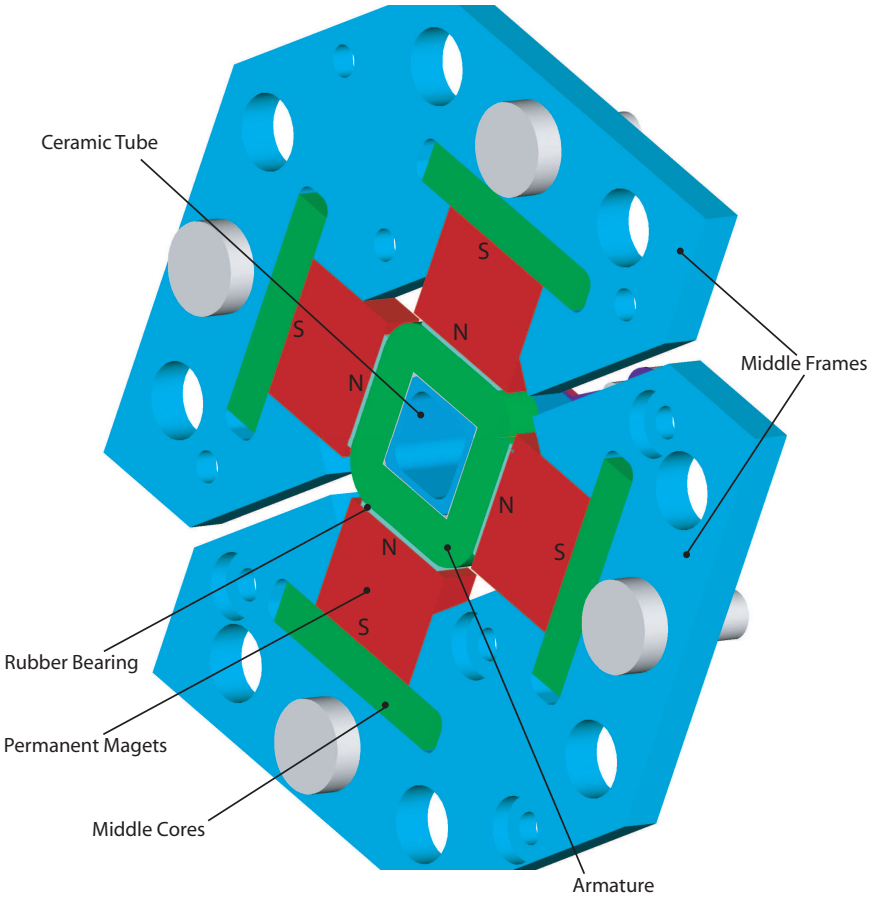


Figure 4-19: Middle Assembly design.

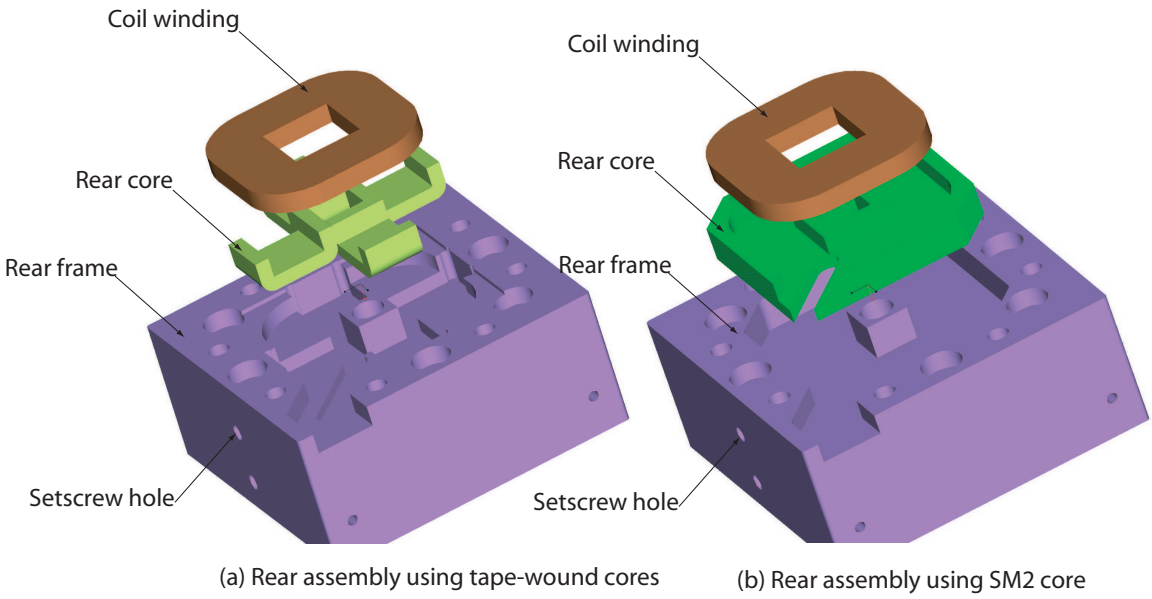


Figure 4-20: Rear assembly designs.

much thicker. Figure 4-20 shows the rear assembly designs of both materials. The rear assembly is composed of a rear coil winding, a rear core, and a rear frame. The rear coil winding and the rear core are identical to those in the front assembly. The capacitance probe is clamped to the rear frame by four soft-tip setscrews, with two in the vertical direction and two in the horizontal direction.

4.7 Fabrication

At the time of writing this thesis, I have assembled and tested only the SM2-based UFTS. All parts of the Ni-Fe based UFTS are ready, but have not been put together; this task remains for later studies.

This section thus mainly describes the fabrication of the SM2-based UFTS, as shown in Figure 4-21. The overall size is $80\times80\times80$ mm^3 excluding the tail of the capacitance probe. In the center of this figure, we can see the diamond tool installed on the tool holder, which is supported by the steel front bearing. Figure 4-22 shows all component parts for this UFTS.

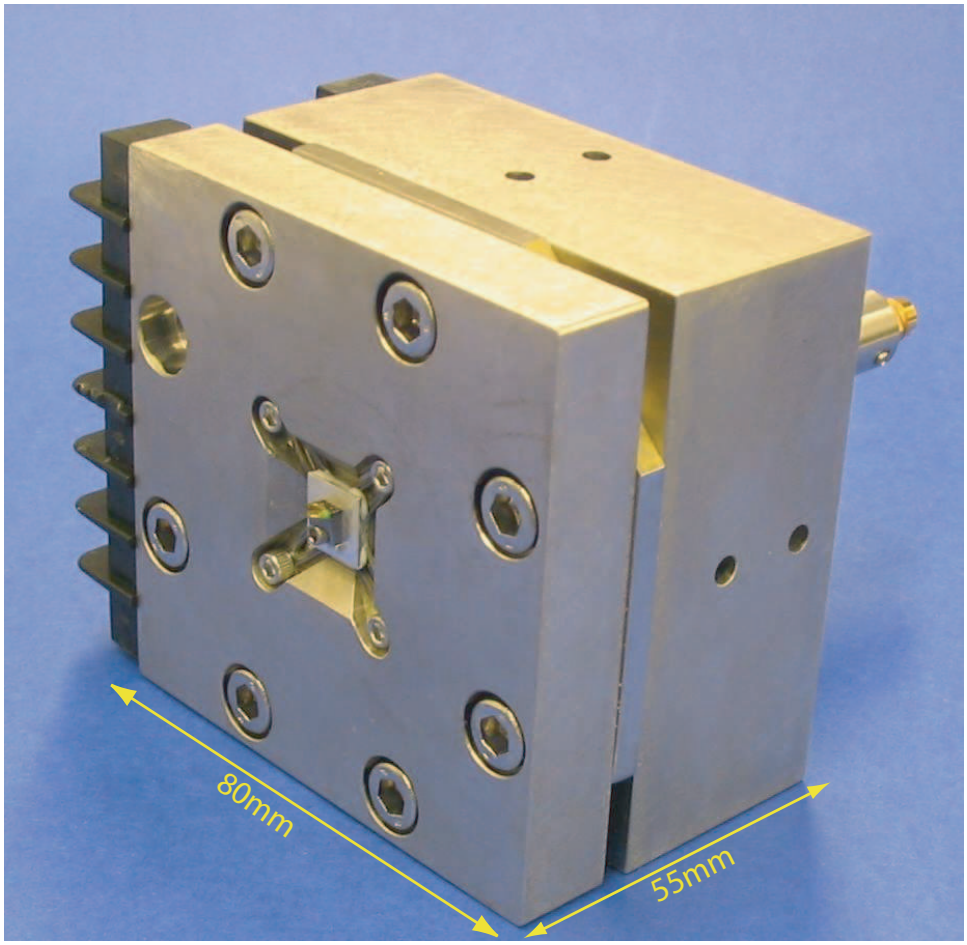


Figure 4-21: Pictures of the assembled SM2-based UFTS prototype.

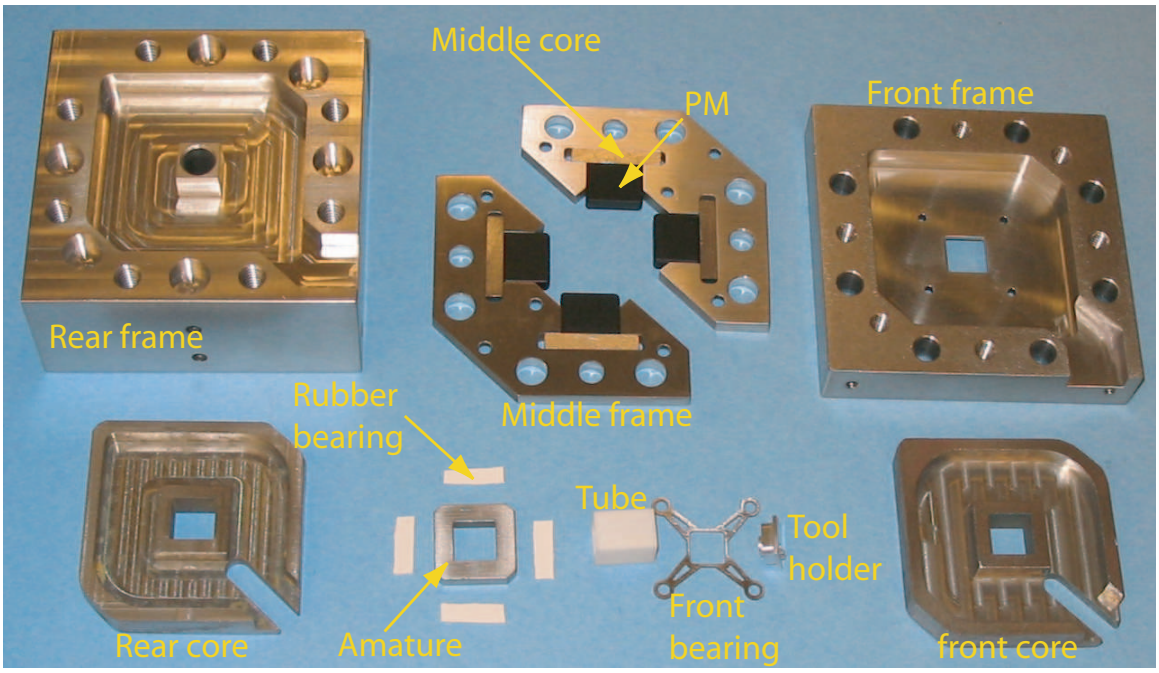


Figure 4-22: Pictures of all component parts for the SM2-based UFTS.

Moving Assembly Fabrication

The moving assembly is shown Figure 4-23, in which all parts are put together but are not glued yet. The first step in armature assembly is to clean the armature and the ceramic tube by using an ultrasonic cleaner. The next step is to glue the armature to the ceramic tube assembly using high strength epoxy EA9360[7] as shown in Figure 4-24. Plastic shims are used between the bonded surface for two purposes: (1) to center the tube relative to the armature; (2) to ensure a epoxy bonding layer of 100 μm thickness and thus to avoid adhesive starvation [67]. Both the armature and the ceramic tube are aligned against a flat surface. The residual epoxy can be easily cleaned off with a knife after the epoxy is fully cured.

Next, the ceramic tube along with the armature is inserted into the square hole in the middle of the front assembly (which should be assembled well before this step), and the tool holder and the front bearing are glued to the ceramic tube with epoxy at the other end of the alumina tube. After this step, the moving assembly will not

[7]A product of Loctite, www.loctite.com. Its shear strength is 34.5 MPa at 25°C, and 8.3 MPa at 93°C. The curing time is 5 days.

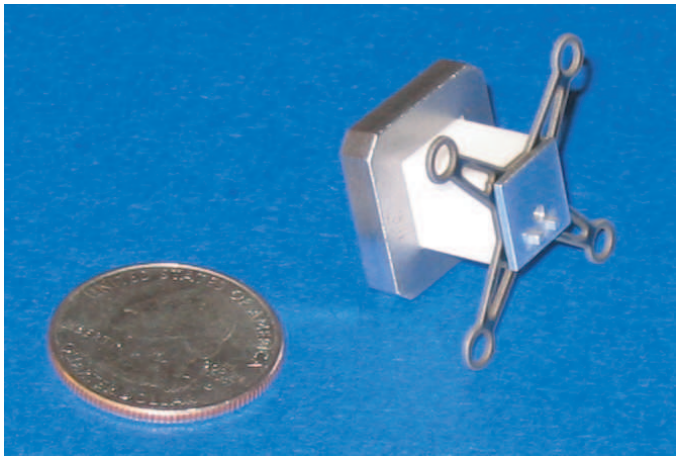


Figure 4-23: Picture of the moving assembly.

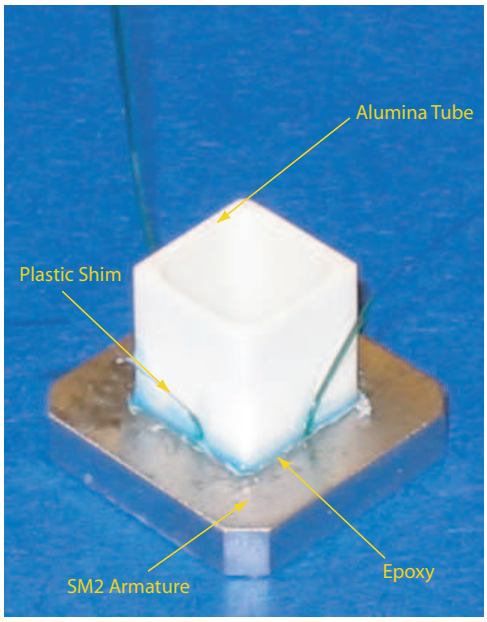


Figure 4-24: Pictures of the middle assembly fabrication.

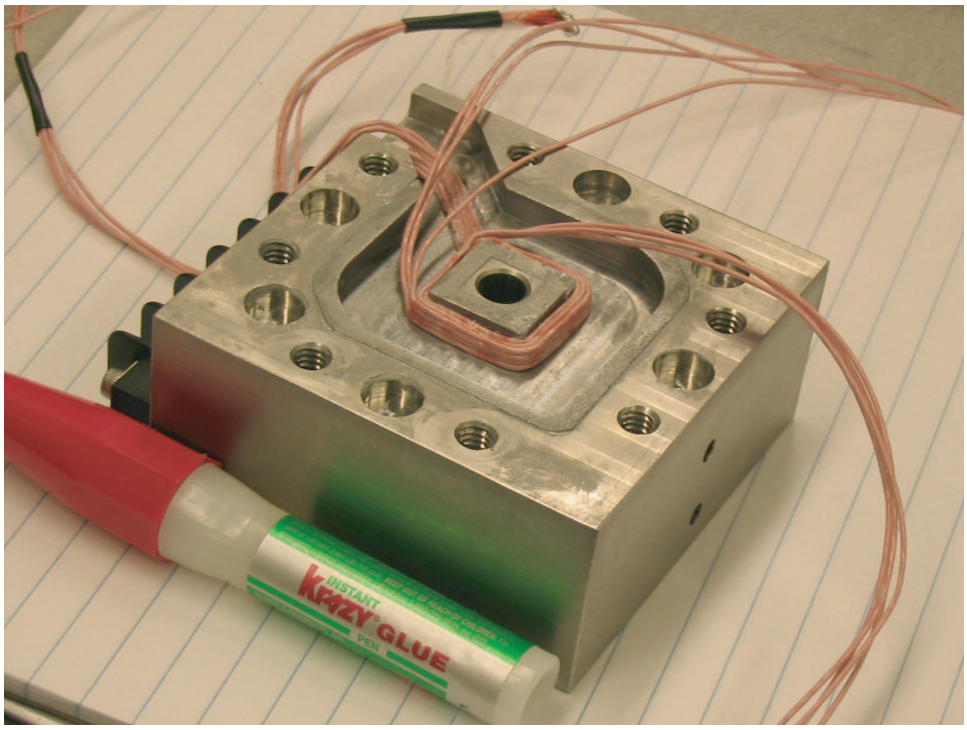


Figure 4-25: Pictures of the coil winding fabrication.

be able to be separated from the front assembly.

Fabrication of the Front and Rear Assemblies

The front core is glued to the front frame with epoxy EN-20S [8]. Next, the Litz wires can be wound into the front core slot as shown in Figure 4-25. In order to tack down the flexible wires during the fabrication process, a cyanoacrylate glue is used to spot attach the winding layers. Figure 4-26 shows the fabricated front assembly with two front coils inside.

The rear assembly is fabricated similarly. After two rear coils are wound inside the slot, epoxy EN-20S is applied on top of the coil to fill up to the frame surface. After the epoxy is cured, both the front assembly and rear assembly are ground flat with a grinding wheel. Figure 4-27 shows the finished rear assembly.

[8]EN-20S is a product of Loctite, www.loctite.com. The shear strength of EN-20S is 10 MPa at 22°, and 5 MPa at 93°C. It needs 24 hours curing time.

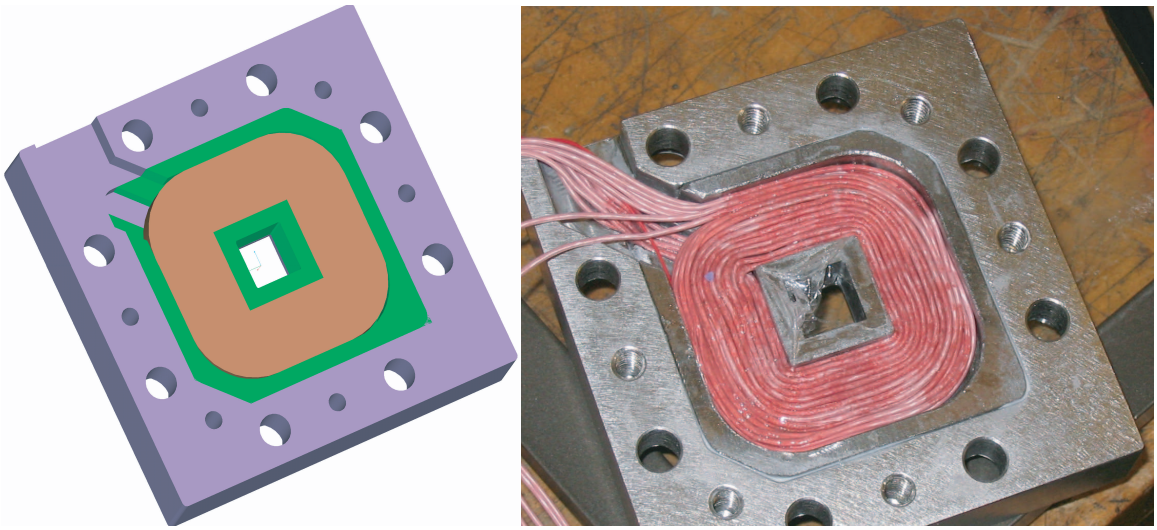


Figure 4-26: The SM2-based front assembly.

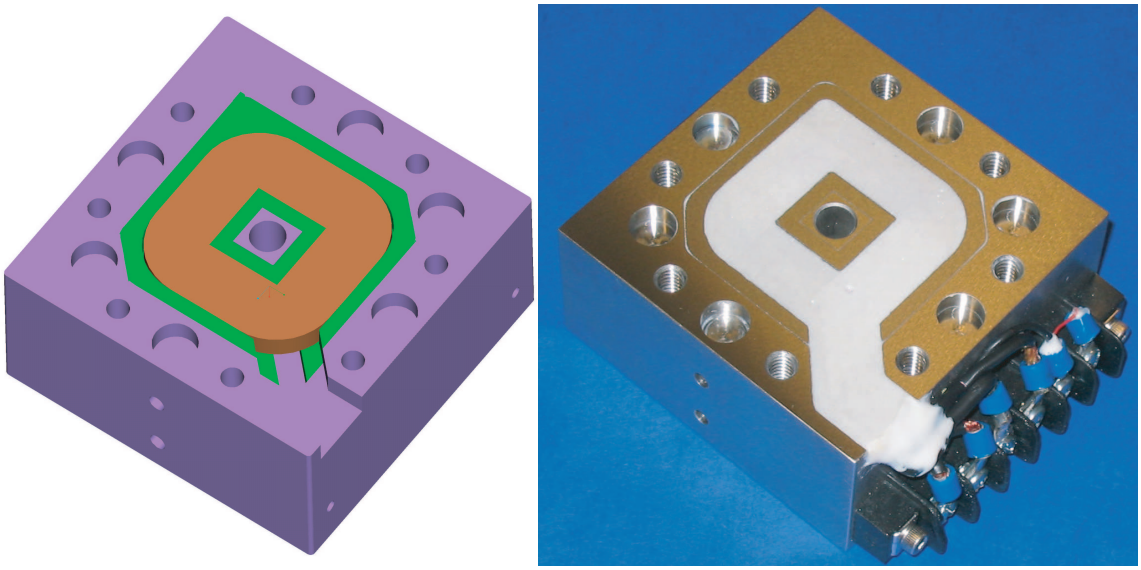


Figure 4-27: The SM2-based rear assembly.

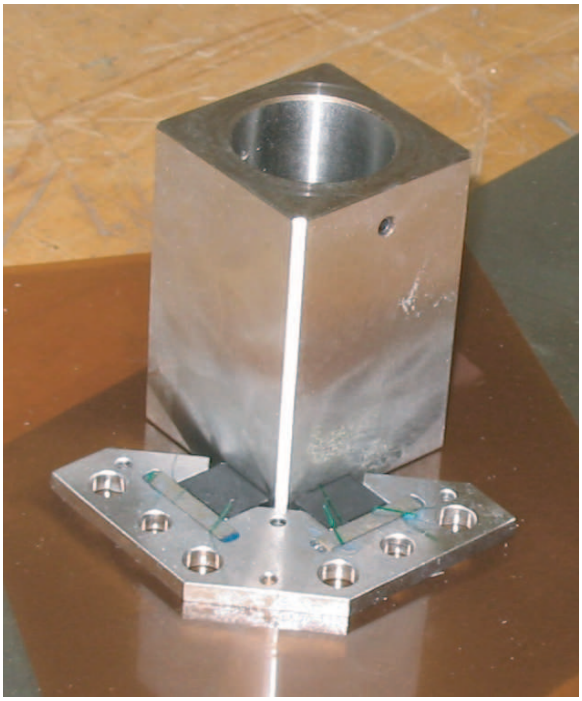


Figure 4-28: Picture of the middle frame fabrication.

Middle Assembly Fabrication

As shown in Figure 4-28, a square steel block is used as a fixture for the middle assembly. The pole surfaces of two permanent magnets are attached to two orthogonal surfaces on the block. Note that all magnetic poles facing the square block must be the same (we use North poles in the design). In order to maintain a 100 μm thick epoxy bonding layer[9], plastic shims are inserted between the middle cores and the middle frame, and also between the middle core and the permanent magnets. We use high strength epoxy EA9360 for all bonding in the middle assembly.

After two halves of the middle assembly have been fully cured, they are ground together on the grinding machine to a thickness of 4.2 mm. Both sides of the middle assembly need to be ground. As shown in Figure 4-29, six counter-bore holes labelled by a letter "F" are designed to fix the middle assembly on a flat steel surface for part setup on the grinding machine. Therefore, all mating surfaces in three assemblies (front, middle, and rear) are ground, and thus the parallelism and dimensions of the

[9]Usually a 50$\sim$150 μm epoxy bonding layer is preferred according to the manufacture guide.

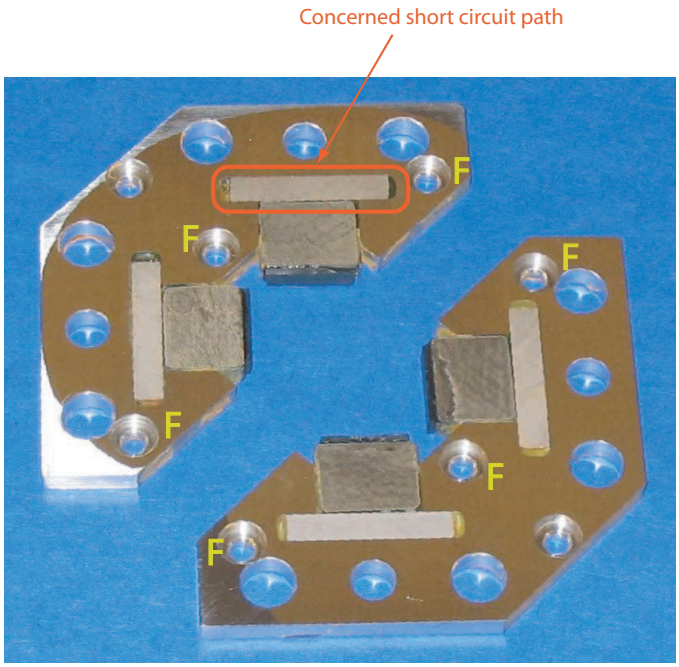


Figure 4-29: Middle assembly. The sixe holes labelled by letter "F" are used to fix the middle assembly on the grinding machine bed.

air gaps between the armature and the stators can be easily guaranteed. The Ne-Fe-B permanent magnets are epoxy coated when ordered from the manufacturer, mainly to prevent eddy currents occurred in the possible short circuit path shown in Figure 4-29. Because the permanent magnet is a good conductor, the AC flux flowing through the middle cores would generate significant eddy currents if there existed electrical contacts between both ends of the permanent magnet and the steel frames.

Figure 4-30 shows the middle assembly installed on the front assembly. Note that 50 μm thick 3M tapes[10] are used to cover the armature surface to prevent metal-to-metal contacts between the stator surface and the armature surface, and thereby limit contact forces should the system be overdriven.

4.8 Summary

This chapter showed two ultra fast tool servo designs based on Ni-Fe (50%-50%) wound tape cores and SM2 powder pressed cores, respectively. The UFTS using the

[10]The part number is 3M 5419, and the Digikey order number is 3m591934.

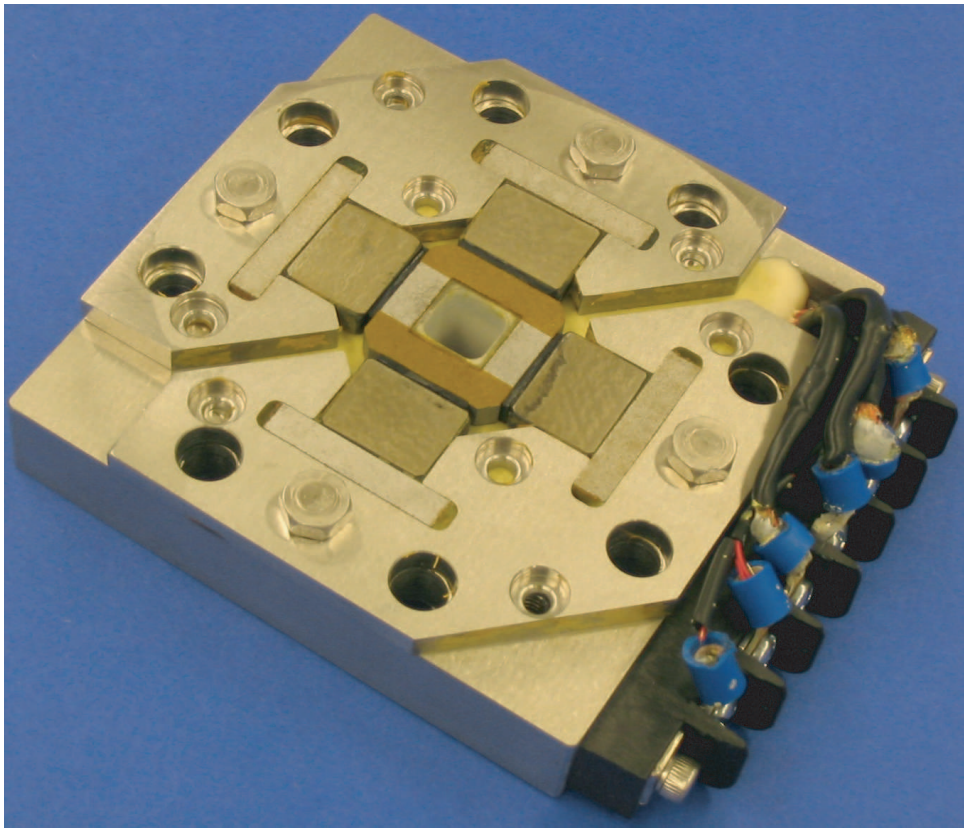


Figure 4-30: Pictures of the middle assembly installed in the front assembly.

Ni-Fe material can achieve a maximum acceleration of 1200 G, while the UFTS using the SM2 material is for conceptual proof only. The armature thickness was set at 4 mm, as a tradeoff between the achievable acceleration and the structural stiffness. The front flexure bearing was designed in "X" shape with a thickness of 0.3 mm, and the rear bearing was composed of four rubber pads. The resulting lateral static stiffness at the tool tip is 28 N/μm.

We presented the definitions of Actuating Ability and Sensing Ability for flexible modes. Their multiplication indicates how excitable a mode is as seen by a controller. When applying this criteria to analyzing the moving assembly structural dynamics, we can generally ignore modes of little displacement at the position sensing point (little Sensing Ability) and/or anti-symmetric armature mode shapes (little Actuating Ability). What's more, the wave transmission delay effect was discussed from the motion control point of view.

Finally, we showed the fabrication details of the SM2-based UFTS. With the system design and fabrication in hand, in the next chapter we turn our attention to the issues of power amplifiers driving the actuator coils.

Chapter 5

Power Amplifier

This chapter presents the design, analysis, and control of a linear high-power high-bandwidth power current amplifier as shown in Figure 5-1, which is used to drive the ultra fast tool servo (UFTS). The specifications of this power amplifier are 1000 W power and 100 kHz closed-loop bandwidth.

The power amplifier consists of four parallel-working channels. Each channel consists two cascaded loops: an internal voltage control loop and an outer current (or flux) control loop. In the voltage control loop, a power operational amplifier PA52 forms a power voltage amplifier with output voltage feedback. Further, the output current of this power voltage amplifier is compared with a reference current command, and their difference is fed into a current controller. The output from this current controller drives the input to the power voltage amplifier. As a result, the reference current command dictates the output current of the overall closed-loop system.

In this chapter, we first present an architecture solution to meet the 1000 W requirement. Then we discuss the thermal design. Next, we go into the detailed design and compensation techniques for both the voltage control loops and the current control loops. In the current control loop, four parallel channels are strongly coupled together through mutual inductance of the driven load, and thus the system represents a multiple-input-multiple-output (MIMO) problem. To address this problem we present a decoupling theory that can transform the MIMO control problem into com-

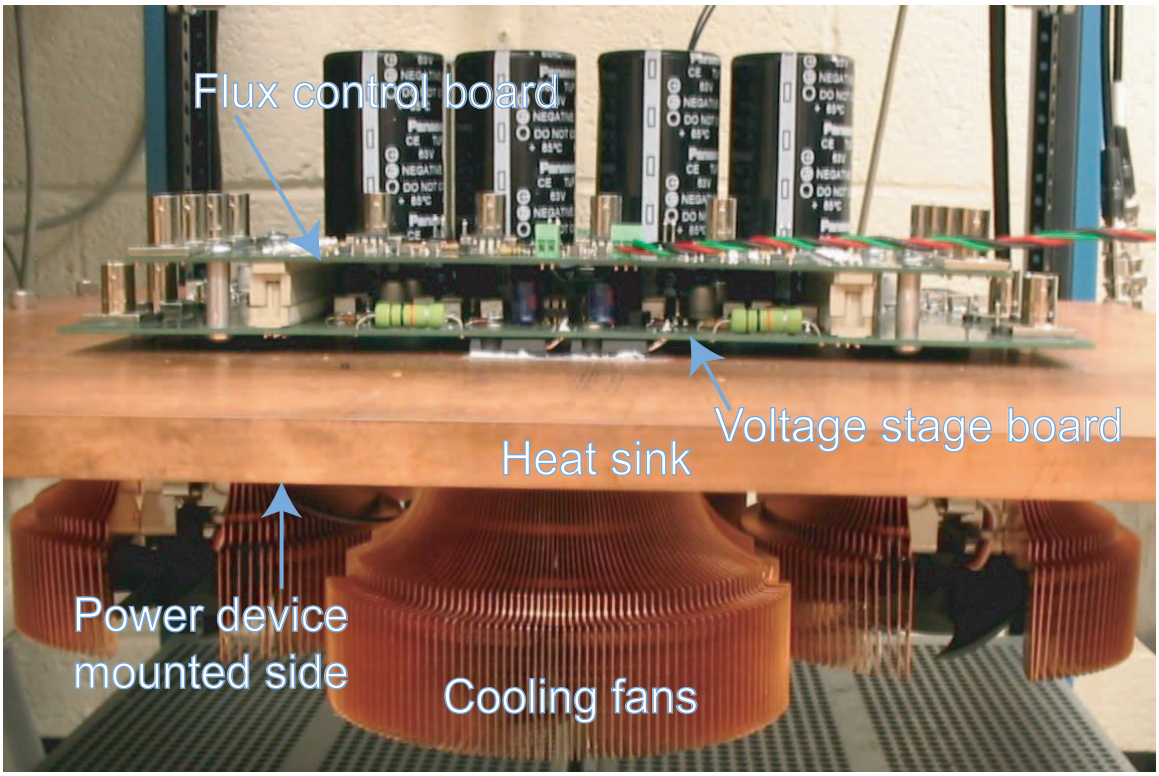


Figure 5-1: The prototype linear power current amplifier mounted on a 19-inch lab rack. A half-inch copper plate is used as both the mechanical backbone and the heat sink connecting between power devices and the cooling fans. Two separate PCB boards are stacked together for voltage control and current (or flux) control respectively.

pensation of several SISO loops. At the end of this chapter, we introduce future ideas for flux sensing and control, although they have not been investigated experimentally.

5.1 Design Requirement

The power current amplifier supplies excitation current to the actuator coils in the UFTS, and thus to drive the UFTS to the desired motion. According to Table 4.2, for the Ni-Fe based UFTS, the maximum output current is expressed as

$$\begin{aligned} I_{max} &= I_0(2-\eta) \\ &= 4.2 \times (2 - 0.95) \\ &= 4.5 \text{ [A]}. \end{aligned} \tag{5.1}$$

The maximum output voltage is V_{max}=47 V for each coil at 10 kHz operation. To meet this we set the power supply voltage at V_s=60 V. Recognizing the 90 degree phase shift between the output voltage and the output current at this frequency, the maximum mean dissipated power P_d (for linear amplifier only) and the maximum instantaneous output power P_o can be calculated as

$$\begin{aligned} P_d &= 4\frac{1}{2\pi}\left(\int_0^{\pi} I_{max}sin(\theta)(V_s - V_{max}cos(\theta))d\theta + \int_{\pi}^{2\pi} -I_{max}sin(\theta)(V_{max}cos(\theta) + V_s)d\theta\right) \\ &= 4\frac{2}{\pi}I_{max}V_s \\ &= 4 \times 4.5 \times 60 \times \frac{2}{\pi} \\ &= 700 \text{ [W]}, \end{aligned} \tag{5.2}$$

$$\begin{aligned} P_o &= 4 \times \frac{1}{2}V_{max}I_{max} \\ &= 4 \times \frac{1}{2} \times 4.5 \times 47 \\ &= 423 \text{ [W]}. \end{aligned} \tag{5.3}$$

These numbers are calculated for 10 kHz operation without any DC component of output current. As a DC current component I_{dc} requires additional dissipated power capability of $I_{dc}V_s$, we set the design requirement of the power amplifier at 1000 W power dissipating capability to allow I_{dc}=1.25 A in additional to the 10 kHz AC dissipated power on each channel.

The overall motion control system bandwidth of 20 kHz requires that the power current amplifier bandwidth (from the current reference command to the actual current output) should be a least 100 kHz, to ensure that the power amplifier does not introduce significant phase lag to the motion control loop. What's more, a higher bandwidth design means that higher loop gain can attenuate more nonlinearity in the amplifier system. At the same time, the power current amplifier should be capable of four-quadrant operation.

5.2 Architecture Design

Generally, there are two power amplifier architectures: switching and linear. Although a switching architecture is more energy-efficient, it is not be considered as an appropriate solution for our UFTS mainly because of its low bandwidth and high EMI noise. The most typical switching amplifier mode is to used pulse-width modulation (PWM). The industry standard frequency for PWM is 20 kHz. Even if a switching amplifier can be designed with a switching frequency at 100 kHz, the resulting bandwidth will be around 10 kHz or even less. Most importantly, the significant switching noise introduced by the switching amplifier will give a ripple force on our high-bandwidth actuator, and will also pose a serious problem to the high-accuracy capacitance position sensor. Therefore, we have selected a linear solution, even though it presents a significant design challenge.

So far, a 1kW and 100 kHz bandwidth linear power amplifier is not commercially available, and only a home-designed linear power amplifier can meet our requirements. One way to design a power amplifier is to start with discrete components. Using this approach can possibly achieve the best performance, but it brings the most complexity. To avoid this time-consuming process, commercial power voltage amplifiers PA52A[1] are selected as the voltage stage components, mainly due to their high power dissipating ability provided by their low thermal resistance MO-127 package.

The overall power amplifier architecture is shown in Figure 5-2. There are four channels operating in parallel. In each channel, there is a PA52A driving an individual actuator coil. A dual-loop structure is implemented. A voltage controller $C_v(s)$ controls the PA52A voltage output. Outside this voltage loop, there is a dedicated current controller $C_c(s)$ for each channel. This sequential compensation architecture has been widely used in a number of theses of our lab. However, the multiple channel architecture has not been implemented before and the problem of stabilizing the strongly coupled MIMO loads has not been addressed yet.

[1]PA52A is made by Apex Microtechnology. www.apexmicrotech.com

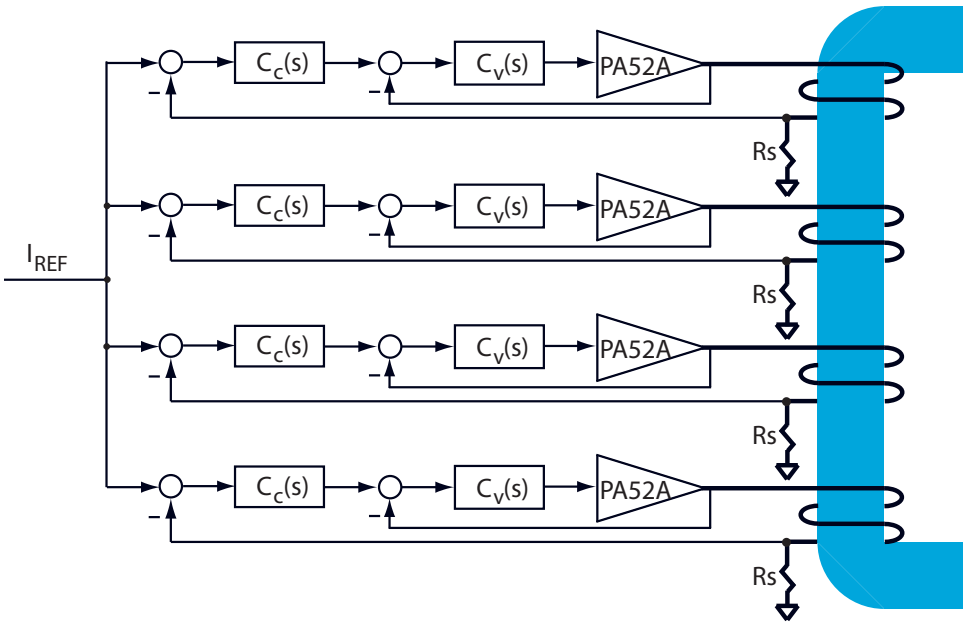


Figure 5-2: Power amplifier architecture using four PA52A devices in parallel.

5.3 Thermal Design

The start of designing this linear high bandwidth power amplifier is the thermal analysis. Each PA52A is rated to dissipate 400W. However, this is under the condition that sufficient cooling exists to remove all the generated heat and thus keep the case temperature at ambient. In reality, the dissipated power is further limited by the heat sink design. For 1 kW total power dissipation, 4 power devices are used so that each power amplifier dissipates up to 250 W. This dissipated 1 kW power must be transferred by a heat sink from the power devices to ambient without overheating the semiconductor junction. A tour to the local computer shop lead to the choice of a CPU cooler CNPS7000A-Cu[2] with thermal resistance of $0.2^{o}/W$ with its fan in operation. This cooler design is unique in that copper plates are stacked together and then are spread out to make fins as shown in Figure 5-1. In comparison, extruded heat sinks usually have very coarse fin density due to limits of extrusion processing, and thus are not very thermally efficient. Heat sinks with epoxy bonded fins can achieve higher fin density than extruded heat sinks, but these specialty devices are quite expensive, and epoxy itself does not conduct heat very well. The selected cooler

[2]It is made by Zalman, www.zalman.co.kr

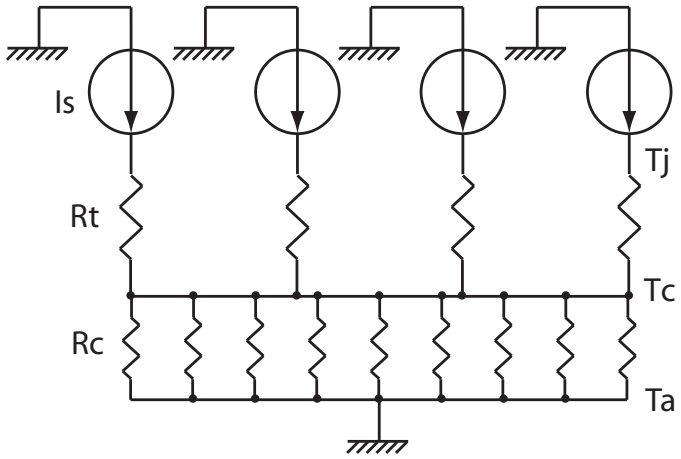


Figure 5-3: Thermal model of the power amplifier.

obviously has the advantage of both high conductivity and high fin density, as well as relatively low unit cost.

A total of 9 CPU coolers are sufficient for the power amplifier as illustrated by the thermal model in Figure 5-3. Based on the number of coolers and the cooler size, a copper plate is selected as the mechanical backbone and the thermal conduction path between the power devices and the coolers. Each power device is modelled at full power as a heat source of $Is = 250$ W, and its junction-to-package thermal resistance is $R_t = 0.31$°C/W. The copper plate is modelled as an isothermal point of negligible resistance. Each of the 9 CPU coolers is modelled as a thermal resistor of $Rc = 0.2$°C/W. The junction temperature T_j and the PA52 package case temperature T_c are calculated as

$$\begin{aligned} T_j &= 4 \times Is \times (\frac{R_t}{4} + \frac{R_c}{9}) + T_a \\ &= 100 + T_a \ [^{\circ}\mathrm{C}], \end{aligned} \tag{5.4}$$

$$\begin{aligned} T_c &= 4 \times Is \times (\frac{R_c}{9}) + T_a \\ &= 22 + T_a \ [^{\circ}\mathrm{C}], \end{aligned} \tag{5.5}$$

where T_c is the ambient temperature. Considering the maximum rated junction temperature is 150°C for PA52A, the maximum allowable ambient temperature is 50°C for 1000 W dissipation, and the corresponding case temperature is T_c=22°C

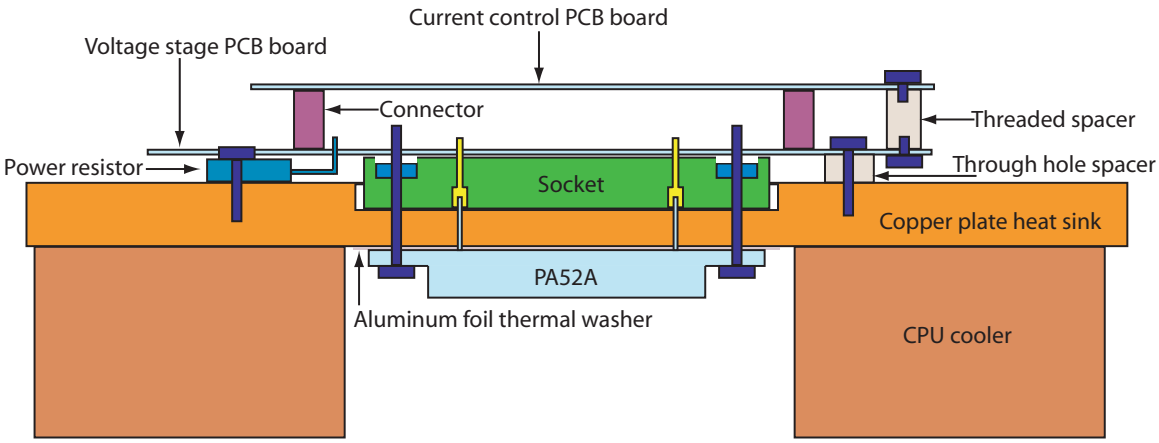


Figure 5-4: Power Amplifier Assembly Structure Design

above ambient. According to the power derating curve of PA52 [52], the internal dissipating power at T_c=22°C is 250 W, which just meets our required power for each device. Therefore, the heat sink design can support 1000 W dissipation at up to 50°C ambient temperature.

The detailed heat sink assembly thermal-mechanical design is shown in Figure 5-4. Both the power devices and the CPU coolers are mounted on the copper plate. The PA52As are mounted on the bottom side of the plate with an aluminum thermal washer[3] in between to compensate the surface unevenness. Pins of PA52A are wrapped with heat-shrink tube to avoid contact with the copper plate and the thermal washer, and go through the clearance holes on the copper plate. These pins insert into the sockets on the other side of the copper plate. To ensure that at least 3 mm pin length is inserted into socket, pockets are machined out on the socket side of the copper plate. Also the current sensing power resistors and power devices used in the boost circuit for PA52A are mounted on the top side of the copper plate. In Figure 5-5, CPU coolers are shown mounted on the copper plates. Current sensing resistors are mounted on the other side of the copper plate as shown in Fig 5-6.

[3]This thermal washer is made by Apex Microtechnology, www.apexmicrotech.com.

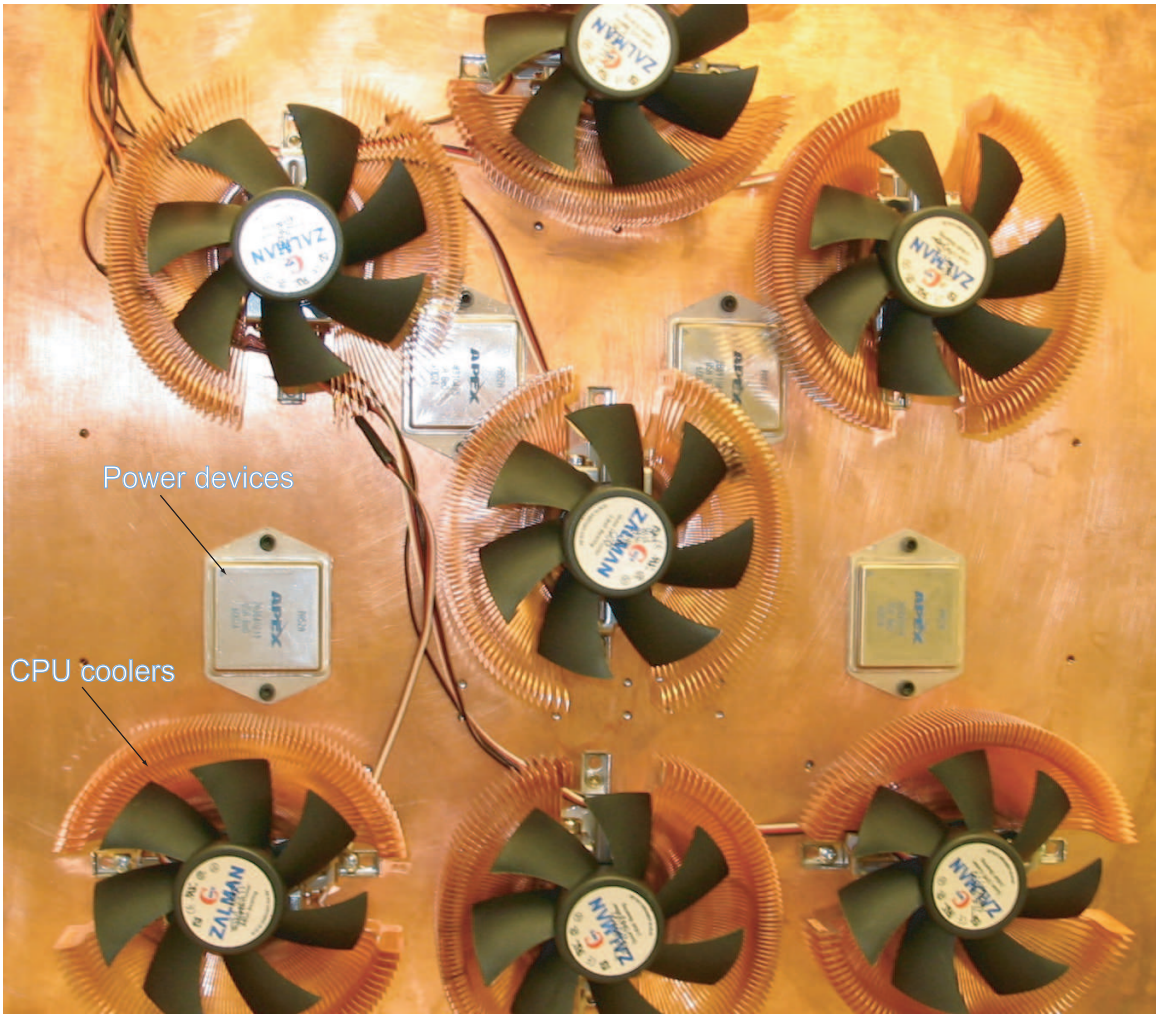


Figure 5-5: The heat sink of the power amplifier

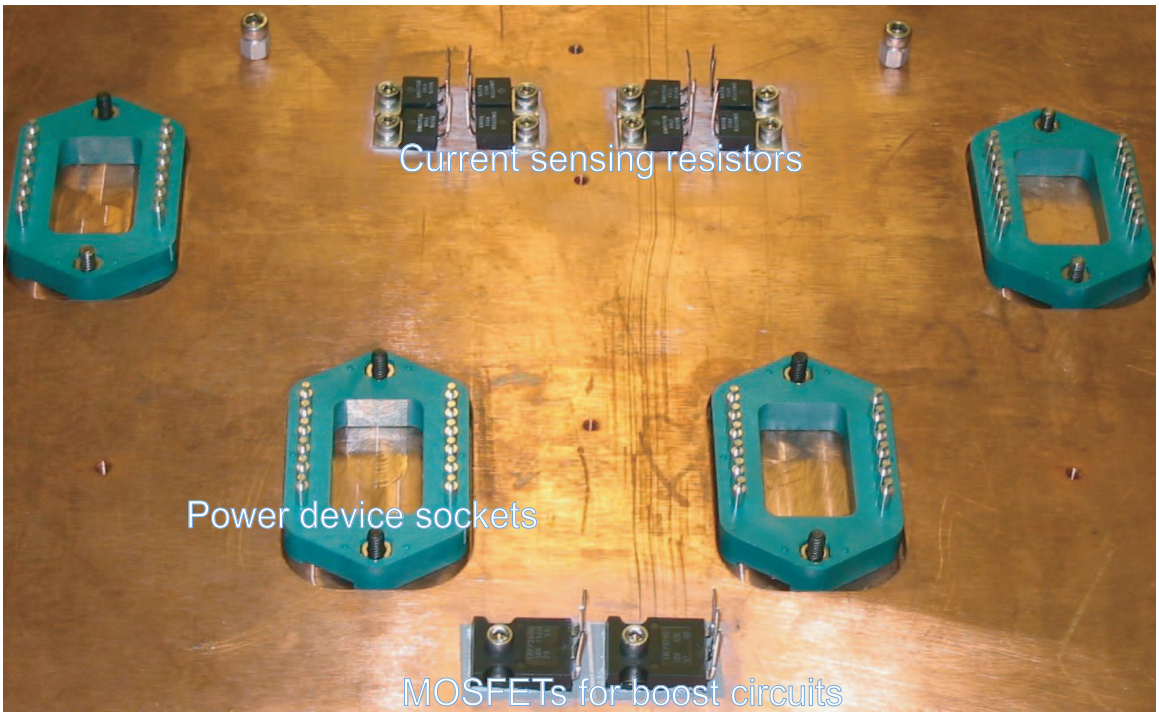


Figure 5-6: Current sensing resistors and MOSFET mounting on component side of the copper plate. Power device sockets are also shown.

5.4 Voltage Control Stage

The voltage stage board of the power amplifier is shown in Figure 5-7. The boost circuit is located at the bottom middle part to power up the signal circuits for all PA52s.

5.4.1 Voltage Loop Compensation

Adapted from textbook[72], the principle of the non-inverting voltage control stage is shown in Figure 5-8(a), and the blocks in shadow are protection circuits. The PA52 has both the small signal power supply $\pm V_b$ and the output power supply $\pm V_s$. Resistors R_a and R_b set the closed-loop DC gain of the voltage stage. This gain is designed based on the input and output dynamic range. In the design with $V_s = 60$ V, the output V_o swings between ± 60 V, and the input to the voltage stage is from the current-control op-amp powered by ± 15 V. Hence, the DC gain is set at 5 with R_a=4k and R_b=16k. The input resistor R_c is set at $R_a//R_b = 3.2$k to balance

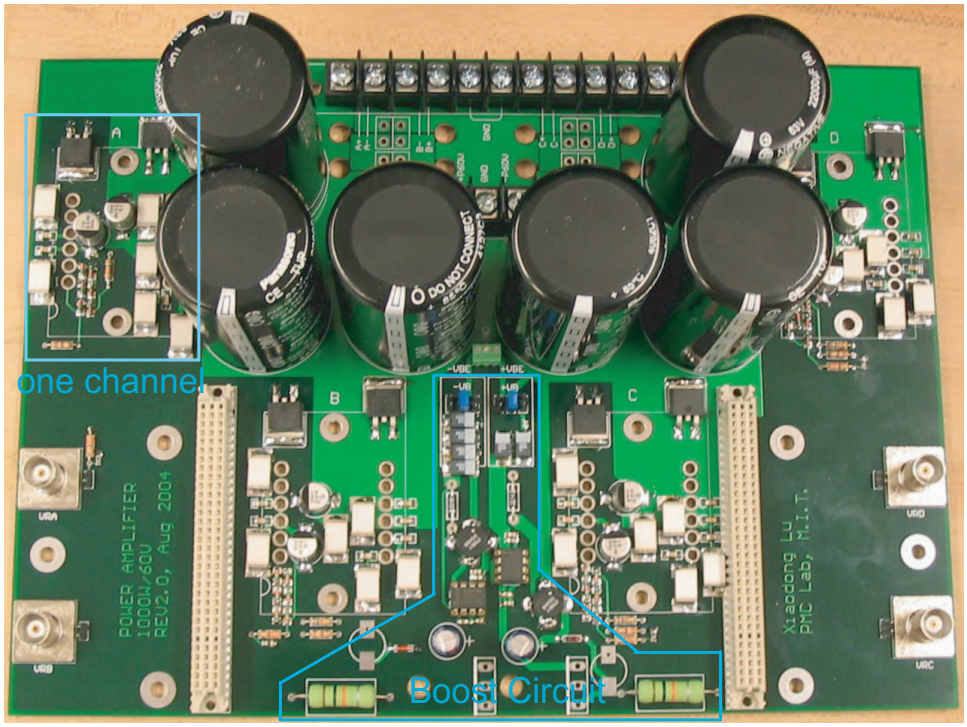


Figure 5-7: Voltage stage print circuit board

the input impedance on both input terminals of the PA52 to ensure good linearity.

The input network $Z_i = R_i//C_i$ is used to reduce the loop transmission gain at the desired crossover frequency. From the circuit equivalence in Figure 5-8(b) and the block diagram in Figure 5-8(c), we can see that the closed-loop DC gain is determined solely by R_a and R_b but the loop transmission can be modified by the input network Z_i.

Figure 5-9 shows the PA52 small signal response. According to the phase response, the desired loop-transmission crossover frequency is set at 1 MHz with 65 degree phase margin. As this frequency response is based on small signal measurement and a 4 Ω nominal resistive load, the actual response with parasitic capacitance at large signal could possibly be worse and thus more aggressive crossover frequency may sacrifice the system robustness. In Figure 5-9, the red curves shows the frequency response of the feedback network to achieve 1MHz cross over frequency. As the open-loop gain of PA52A at 1MHz is 10, the compensation circuit should have a total gain of 0.1 at the crossover frequency. The feedback path has only a gain of 0.2,

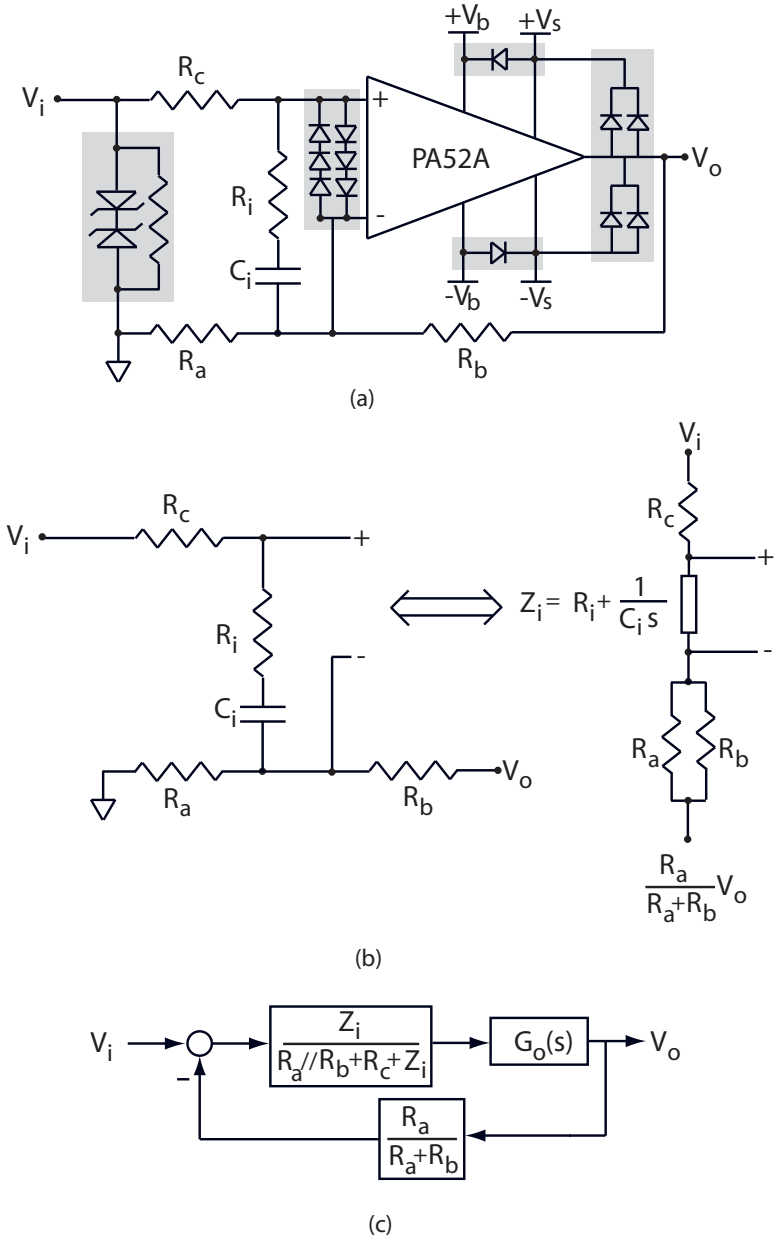


Figure 5-8: Voltage Control Principle. (a) the schematic circuit of the voltage stage. (b) the feedback network and its equivalence. (c) the block diagram of the voltage stage.

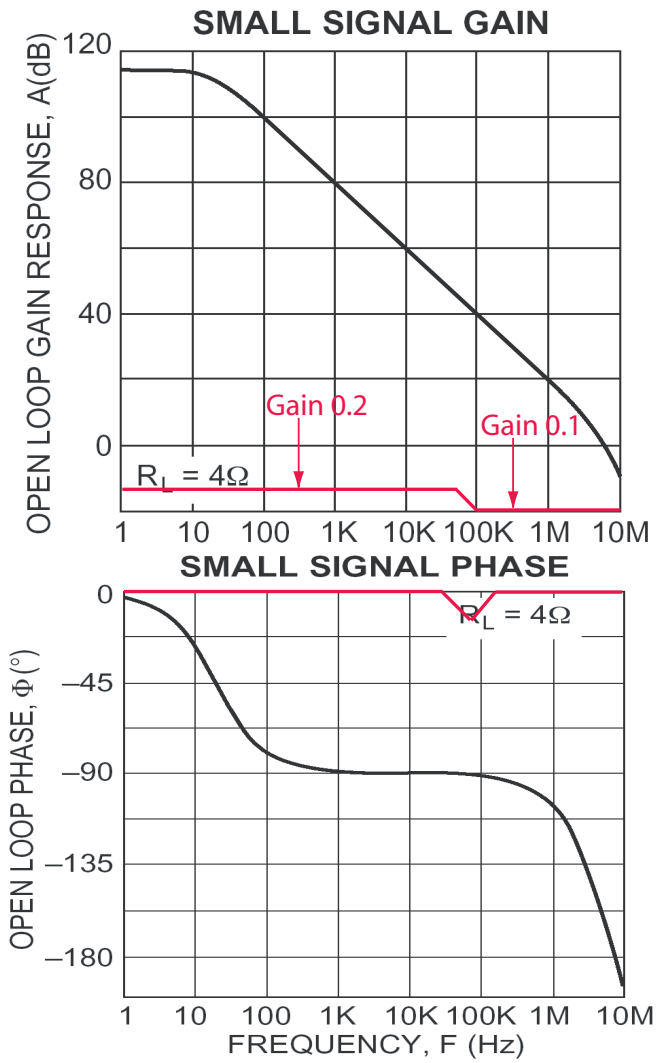


Figure 5-9: Frequency response of a PA52 and its feedback compensation network. Adapted from APEX PA52 data sheet [52].

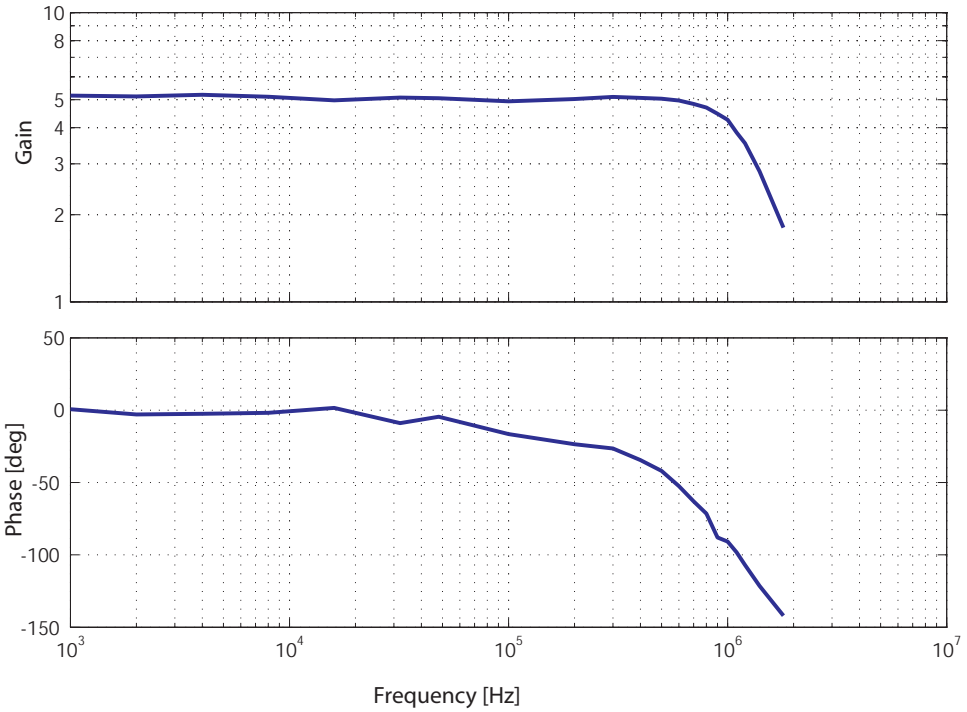


Figure 5-10: Frequency response of the closed-loop voltage stage.

so the input network should contribute a gain of 0.5 to the loop transmission. As a result, R_i=R_c+R_a//R_b=6.4k. The capacitor C_i should be near a short circuit at the crossover frequency 1 MHz, and is used to increase the low-frequency loop transmission reduced by the input resistor R_i. Hence, the R-C network corner frequency is set at 100 kHz, which corresponds to

$$C_i = \frac{1}{2\pi \times 100k \times 6.4k} = 250\text{pF}. \tag{5.6}$$

The actual value of C_i is designed at 330 pF.

Figure 5-10 shows the experimentally measured closed-loop response of the voltage stage with the above parameters, which shows a bandwidth of about 1 MHz.

5.4.2 Boost Circuit

The PA52 has two pairs of power supplies: $\pm V_s$ the main power supply for power output stage, and $\pm V_b$ the power supply for small signal circuit. In order for the

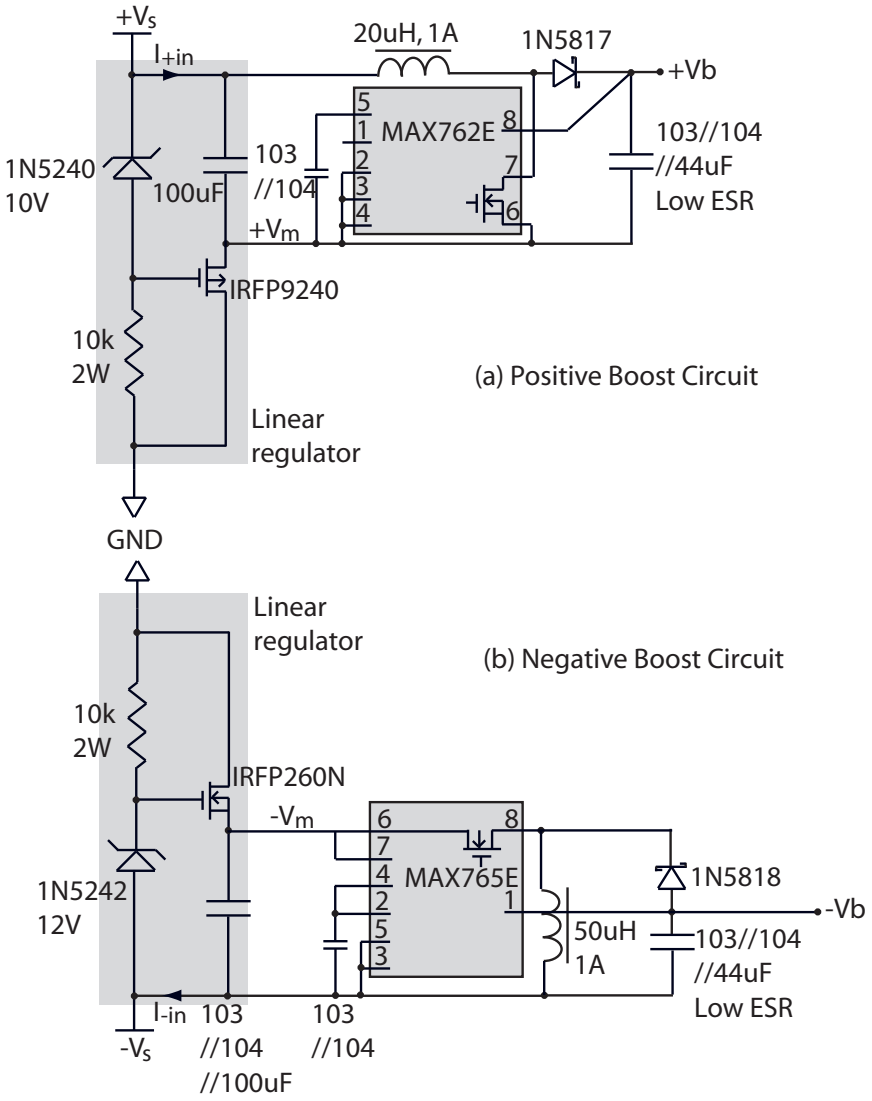


Figure 5-11: The boost circuit block diagram.

output to swing fully rail-to-rail and thereby reduce output stage power dissipation, $\pm V_b$ should be ± 10 Volts beyond $\pm V_s$. Here we describe a boost circuit for the PA52A as shown in Figure 5-11, which is a modified version of the circuit in APEX application note [51].

On $+V_b$ side, a P-channel MOSFET is used as a linear regulator to generate an intermediate potential $+V_m$, which is about 6 Volts below $+V_s$. Further, $+V_s$ and $+V_m$ are connected as inputs to a DC-to-DC booster converter, and its output $+V_b$ is +15 Volts higher than $+V_m$. As a result, $+V_b$ is +9 Volts above $+V_s$. Each PA52A consumes 30 mA from $+V_b$. The total output current of the $+V_b$ boost circuit is

120 mA. Then the input current of the $+V_b$ boost circuit can be calculated as

$$\begin{aligned} I_{+in} &= 120 \times \frac{+V_b - +V_m}{(+V_s - +V_m)\eta_b} \\ &= 120\frac{15}{6 \times 0.9} \\ &= 333 \text{ mA}, \end{aligned} \tag{5.7}$$

where η_b=0.9 is the efficiency of the boost converter. When $+V_s$=60 V, the dissipated power on the P-MOSFET will be $V_s \times I_{+in} = 0.333 \times 60 = 20$ W.

Similarly on the $-V_b$ side, a N-channel MOSEFET is used as a linear regulator to generate an intermediate potential $-V_m$, which is about 8 volts above $-V_s$. Further, $-V_s$ and $-V_m$ are connected as inputs to a DC-to-DC flyback convertor, and its output $-V_b$ is 12 Volts less than $-V_s$. The total output current of $-V_b$ is 120 mA. Then the input current of the $-V_b$ boost circuit can be calculated as

$$\begin{aligned} I_{-in} &= 120 \times \frac{-V_s - -V_b}{(-V_m - -V_s)\eta_f} \\ &= 120\frac{12}{8 \times 0.7} \\ &= 260 \text{ mA}, \end{aligned} \tag{5.8}$$

where η_b=0.7 is the efficiency of the flyback converter. When $-V_s$=-60 V, the dissipated power on the N-MOSFET will be $V_s \times I_{-in} = 0.26 \times 60 = 15$ W.

According to the above power calculation for the MOSFETs , they need to be mounted on the heat sink as shown in Figure 5-6. Note that the 30 mA current for $\pm V_b$ is labelled as quiescent current, excluding the dynamic load situation. Test results show that there is no noticeable current increase for operation below 50 kHz. For operation at over 50 kHz frequency, the currents to $\pm V_b$ will increase significantly because the MOSFET gate capacitors become heavy loads. Hence, the boost circuit design sets a limit that the power amplifier can output full power only below 50 kHz, which is sufficient for driving our UFTS to a power bandwidth of 10–20 kHz.

5.4.3 Power Device Protection

There are several ways that power devices can be damaged if not protected properly. The protection circuits are shadowed in Figure 5-8(a).

The first important thing is to prevent the PA52A output voltage from saturation. For the voltage stage, the input voltage V_{in} should be limited as follows:

$$|V_{in}| < \frac{V_s}{G_{DC}}, \tag{5.9}$$

where G_{DC} is the DC gain of the closed-loop voltage stage. When V_{in} is beyond the above limit, due to some internal PA52 details the output may saturate far from $\pm V_s$ as captured by a test result shown in Figure 5-12. In this figure, the PA52A is powered by $\pm V_s = \pm 40$ Volts and $\pm V_b = \pm 50$ Volts. This result shows that the PA52A output stays somewhere in the middle instead of going to the rails when it is overdriven. The interaction between this abnormal behavior of the PA52 and the current controller may result in sustained oscillation and further damage to the device. As shown in Figure 5-8(a), V_{in} is clamped by a pair of back-to-back zener diodes. The resistor in parallel with the zener diodes is used to bring V_{in} to ground when the input is floating.

Secondly, during the power up process, the driven load need to be cut off from the PA52. Before the boost rails $\pm V_b$ reach the minimum working voltage $\pm$20V, the PA52 is in an abnormal state, where either the N-MOSFETs or the P-MOSFETs in the PA52 output stage is conducting. If the PA52 output is connected to a low impedance load, the power supply $\pm V_s$ will be discharged directly and thus $\pm V_b$ will be never rise high enough to bring the PA52 out of this abnormal state.

Thirdly, the input differential voltage need to be protected within $\pm$20V . As shown in Figure 5-8(a), several diodes stacked together are sufficient. These diodes need to be either Shottky diodes or fast recovery diodes. The number of the didoes should be set such that the total forward voltage is less than the rated input differential offset but higher than the input differential voltage required for dynamic output.

Finally, the output of the PA52 needs to be protected using Shottky diodes to pre-

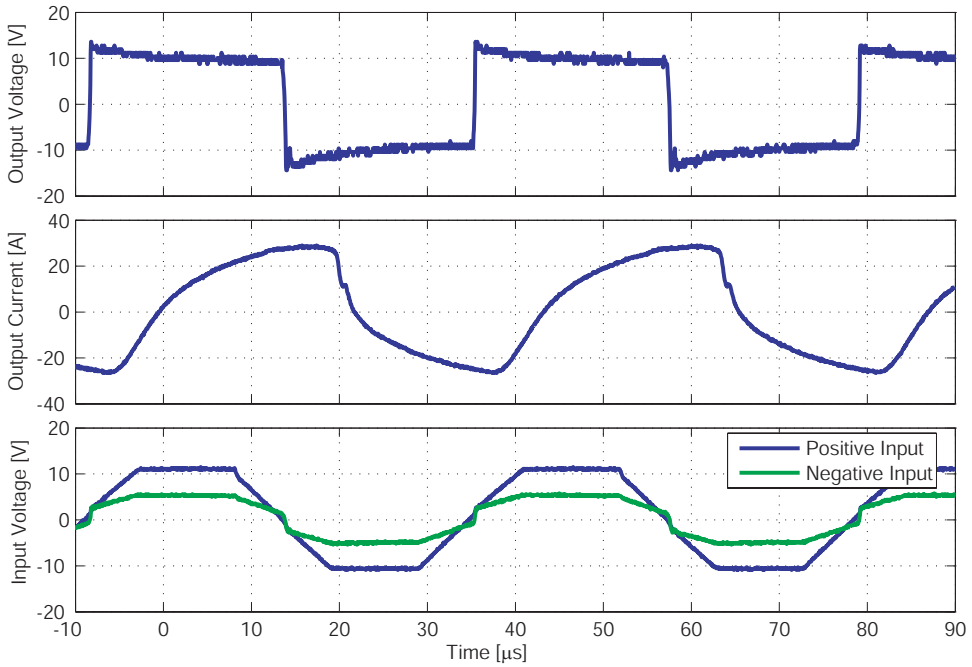


Figure 5-12: Output waveform of the over-driven PA52A. $\pm V_s = \pm 40$ Volts and $\pm V_b = \pm 50$ Volts.

vent large reverse currents going through the in-body parasitic diodes of MOSFETs. Also Shottky diodes need to be placed between $\pm V_b$ and $\pm V_s$ to protect the PA52 in case the boost circuit fails.

5.5 Current Control Loops

This section describes the loop-compensation method for the current control loops. As there are multiple starts in the coils of our designed UFTS, it can be operated in three modes as shown in 5-13: the uni-channel mode, the dual-channel mode, and the quad-channel mode.

Section 5.5.1 shows the current controller circuit design. Then the compensation methods are discussed for all three modes. For the dual-channel mode and the quad-channel mode operation, we present a MIMO decoupling control theory to guide the compensation of these multiple loops.

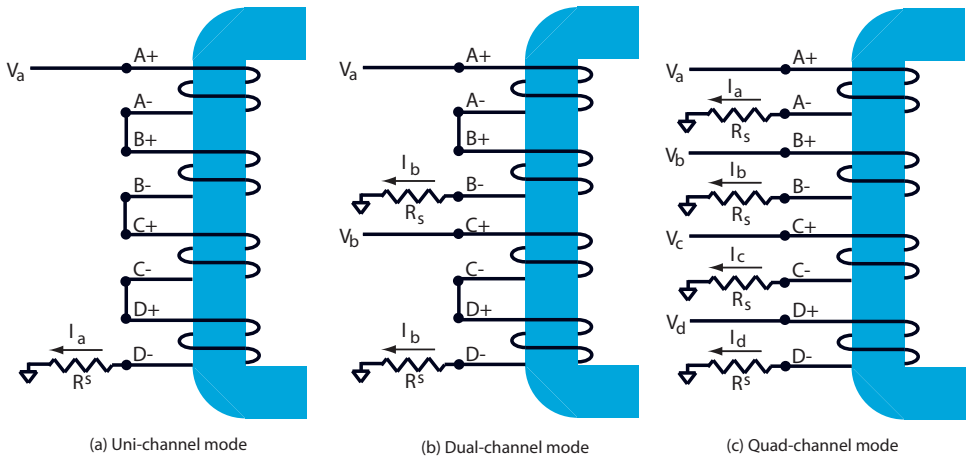


Figure 5-13: The UFTS terminal connection.

5.5.1 Current control circuit

Figure 5-14 shows the current control circuit for one channel. For multiple channel modes, each channel has an identical control circuit. The whole current control circuit can be divided into four parts: the voltage stage, the current sensing, the flux summing, and the PI controller.

The current sensing circuit is used to amplify the voltage difference across the current sensing resistor by a gain of 5. An OP37 is used in this circuit due to its high gain bandwidth product so that the current sensing has a bandwidth over 2 MHz. The current sensing resistor R_s is 0.2 Ω. As a result, the current sensing stage has a total gain of -1 V/A from the sensed current to the current sensing output. The resistors of 5k and 1k needs to be matched so as to have good common mode rejection.

The flux summing stage can be treated as a pure gain of -2 with R_{57} = 10k and R_{59} = 5.1k. It can also be used as a summing point for flux feedback signal, which will be discussed later in section 5.7.

To compensate the whole current loop, a PI controller is used with high frequency rolling-off, and can be implemented in three ways as shown in Figure 5-15. In our amplifier, we used the Figure 5-15(b), which I learned from Rick [58].

Figure 5-14(b) shows the block diagram of the whole circuit including the inner loop of the PI controller. When designing high-bandwidth high-power current amplifiers to drive heavy inductance loads, it is important to make sure that the inner loop

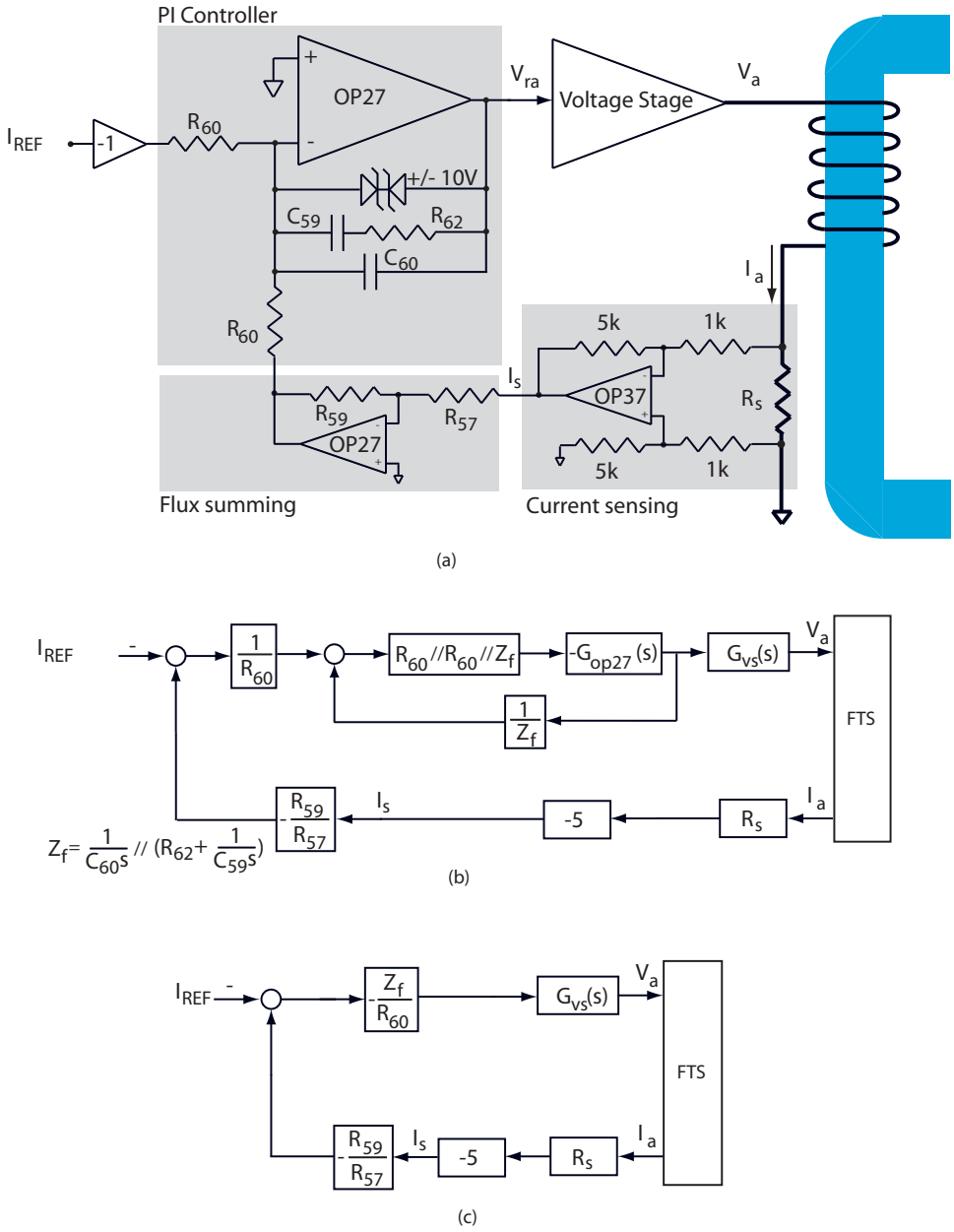


Figure 5-14: Current control circuit. (a) Current feedback and control circuit. (b) The block diagram of current control circuit, G_{op27} is the open loop transfer function of OP27, and G_{vs} is the transfer function of the closed-loop voltage stage. (c) A simplified block diagram of (b).

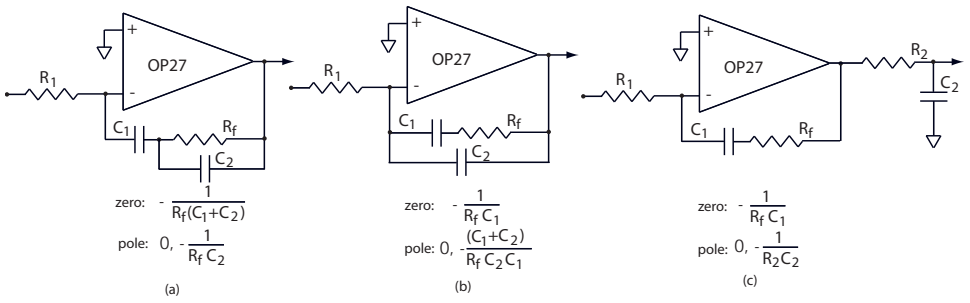


Figure 5-15: Three equivalent implementations of PI controllers.

bandwidth of current control circuit is one decade above the intended overall current loop bandwidth. If the loop transmission cross-over frequency of this internal loop was around 100 kHz, it would be impossible to achieve 100 kHz bandwidth on the overall current loop. By incorporating three amplifier stages (current sensing, flux summing, and current control) into the current control feedback loop, the required gain bandwidth product is distributed among three op-amps. Only after verifying that each stage has a much higher bandwidth than the intended overall system bandwidth, the system block diagram can further be simplified into Figure 5-14(c).

5.5.2 Current Control for Uni-Channel Mode Operation

This subsection describes controller tuning for single channel operation as shown in Figure 5-13(a). All excitation coils are connected in series. The advantage of this setup is that the MIMO control issues can be avoided, but only a quarter of the whole power amplifier capability is utilized, since only one of the four PA52s is active.

The controller design starts with the coil impedance measurement. For such a non-ideal magnetic material as SM2 (low permeability and high eddy current), Figure 5-16 shows the experimentally measured frequency response from the voltage stage input $V_i(s)$ to the current output $I_o(s)$ measured by swept sinusoidal excitation.

Based on the measured frequency response, the compensation parameters are designed at $C_{60} = 150\text{pF}$, C_{59} and R_{62} remain unpopulated, and $R_{60} = 2k\Omega$. The resulting loop transmission is shown in Figure 5-17 with a crossover frequency of 200 kHz and a phase margin of 50 degrees. A 1 A step response is shown in Figure 5-

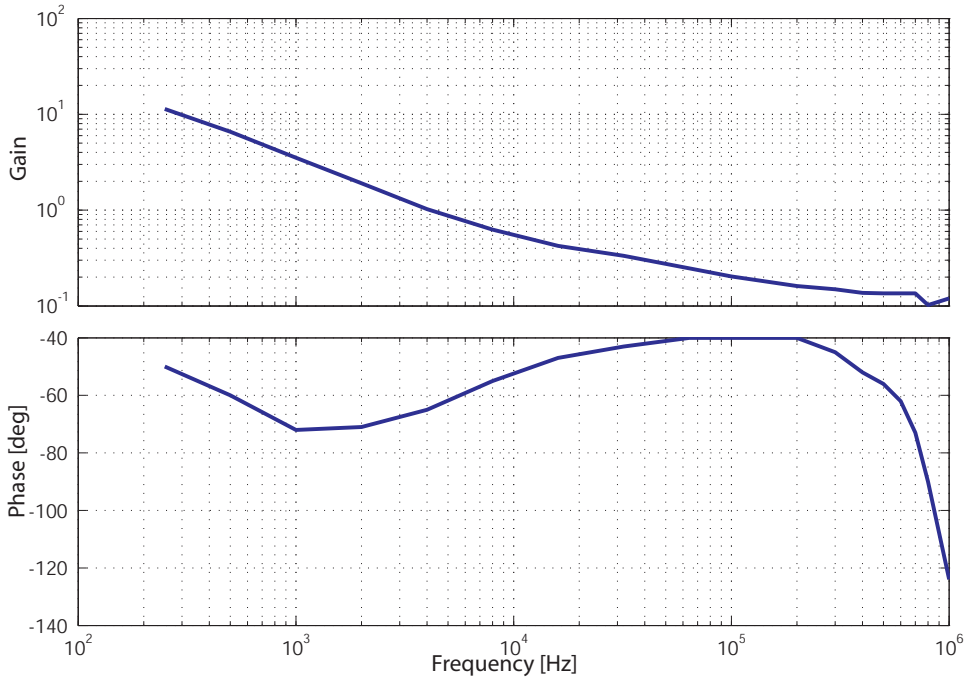


Figure 5-16: Experimentally measured voltage to current frequency response of the power amplifier driving an SM2-based UFTS in single channel mode

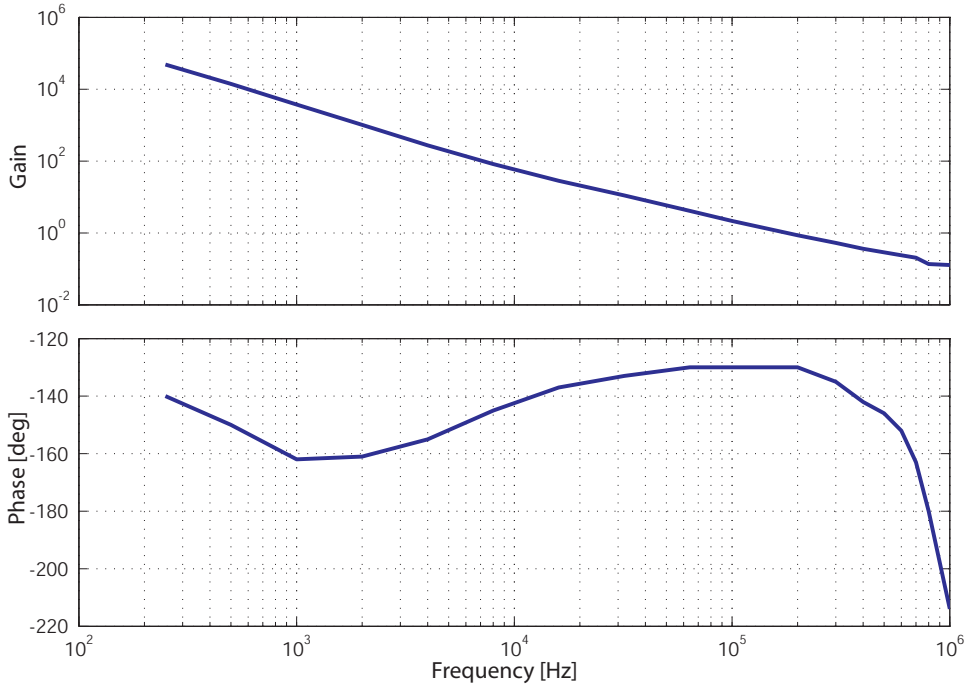


Figure 5-17: Compensated negative of the loop transmission for uni-channel mode SM2-based UFTS

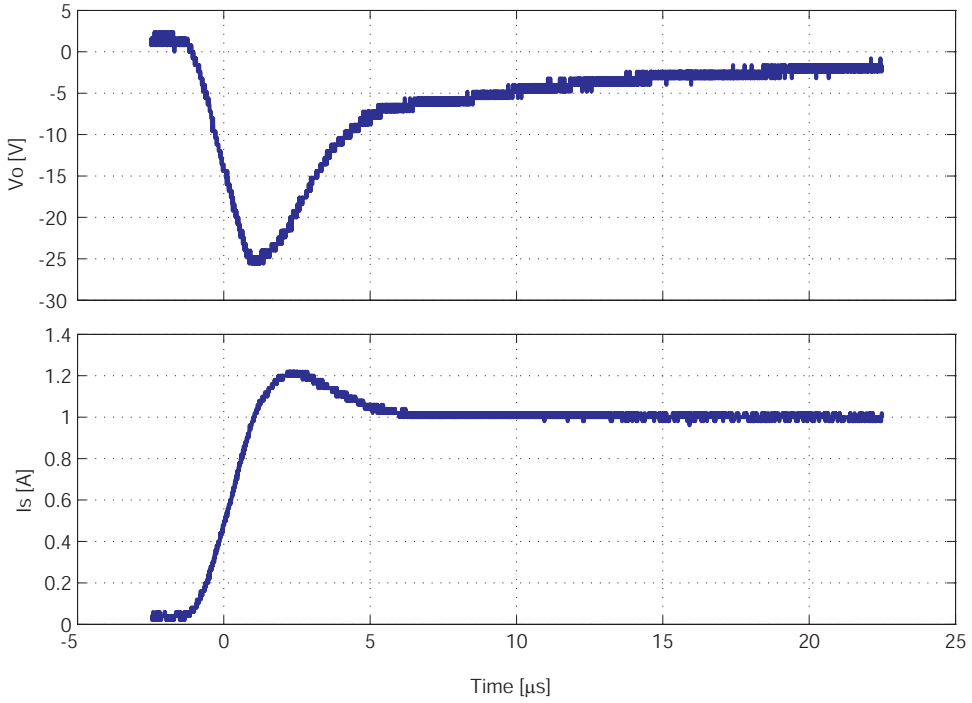


Figure 5-18: Step response of the current amplifier driving an SM2-based UFTS in uni-channel mode.

18, where the current sensing output I_s and PA52A output V_o are plotted. The rise time is about 2 μs.

5.5.3 MIMO Control of A Fully-Coupled Multiple-Start Actuator

In this subsection, we present a MIMO compensation method for fully-coupled coils, in which the leakage flux is negligible. With no leakage, the mutual-inductance between two coils and the self-inductance of each coil have the same values.

The system can be modelled as a multiple-input (V_a, V_b, V_c, and V_d) and multiple-output (I_a, I_b, I_c, and I_d) system, and they are related by $P(s)$ as

$$\begin{bmatrix} I_a \\ I_b \\ I_c \\ I_d \end{bmatrix} = P(s) \begin{bmatrix} V_a \\ V_b \\ V_c \\ V_d \end{bmatrix}, \tag{5.10}$$

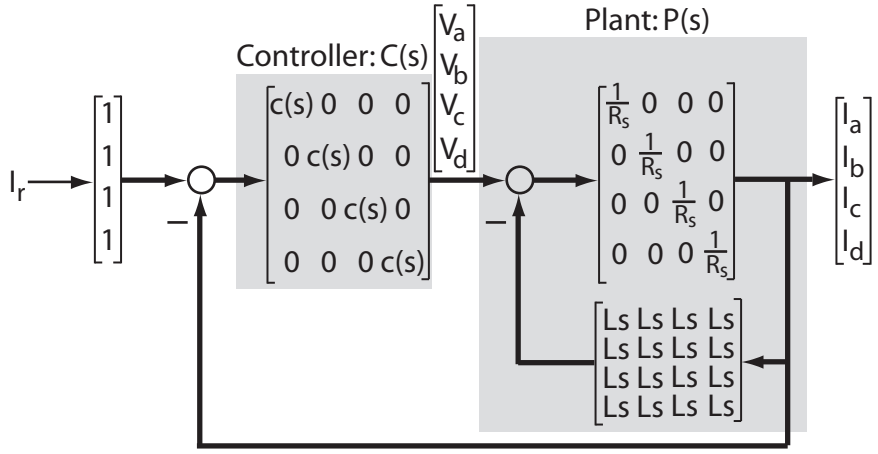
where

$$P(s)^{-1} = \begin{bmatrix} R_s + Ls & Ls & Ls & Ls \\ Ls & R_s + Ls & Ls & Ls \\ Ls & Ls & R_s + Ls & Ls \\ Ls & Ls & Ls & R_s + Ls \end{bmatrix}, \tag{5.11}$$

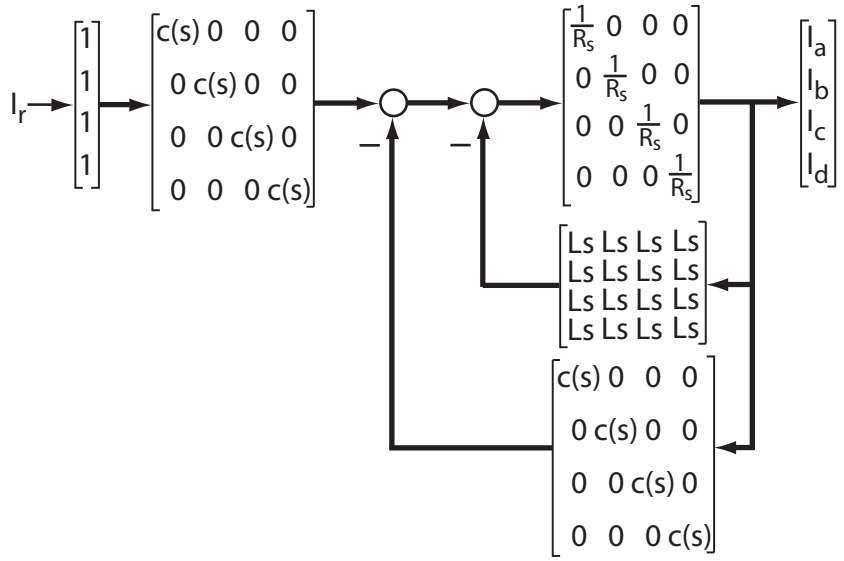
where the inverse of $P(s)$ is expressed analytically due to its simplicity.

Next, through a series of straightforward block diagram transformations we show that the compensation of the MIMO plant in (5.10) can be converted to the compensation of two SISO plants. Figure 5-19(a) shows the system block diagram when using decentralized controller $C(s)$ with diagonal elements $c(s)$ to control the plant $P(s)$, which is represented in a feedback format. The reference current command for each channel is I_r. By moving the matrix $C(s)$ into the feedback path, we can get the block diagram in Figure 5-19(b). Further, the two diagonal matrices R_s and $C(s)$ can be compressed into one block by Black's formula as shown in Figure 5-19(c). According to the rank-1 property of the inductance matrix in the feedback path of Figure 5-19(c), we can decompose it into the multiplication of three rank-1 matrices as shown in Figure 5-20(a). By introducing a virtual variable I_m, which is the mean value of the current in all channels, the block diagram from I_r to I_m can be simplified into a SISO system as in Figure 5-20(b), which can be further transformed from (c) through (f). As a result, original MIMO system in Figure 5-19(a) is converted to a SISO system in Figure 5-20(e). Note that the SISO system in Figure 5-20(e) does not capture all the dynamics of the original MIMO system. As only the mean current of all channels I_m is controlled, the difference among all channels remain uncontrollable and unobservable in Figure 5-20(e). To ensure the stability of the original MIMO system in Figure 5-19(a), we also need $\frac{1}{c(s)+R_s}$ in Figure 5-20(b) to be stable. In conclusion, the compensation of the MIMO plant is converted to design one controller $c(s)$ to stabilize two SISO plants $\frac{1}{R_s}$ and $\frac{1}{R_s+4Ls}$.

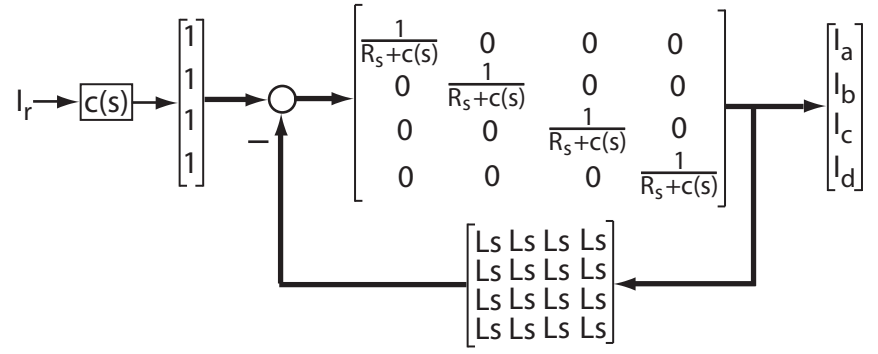
This result can be generalized to arbitrary number of fully-coupled multiple-start coils. **For a MIMO system with n sets of multiple-start fully coupled coils**



(a)



(b)



(c)

Figure 5-19: Block diagram reduction **I** for MIMO control of fully coupled coils.

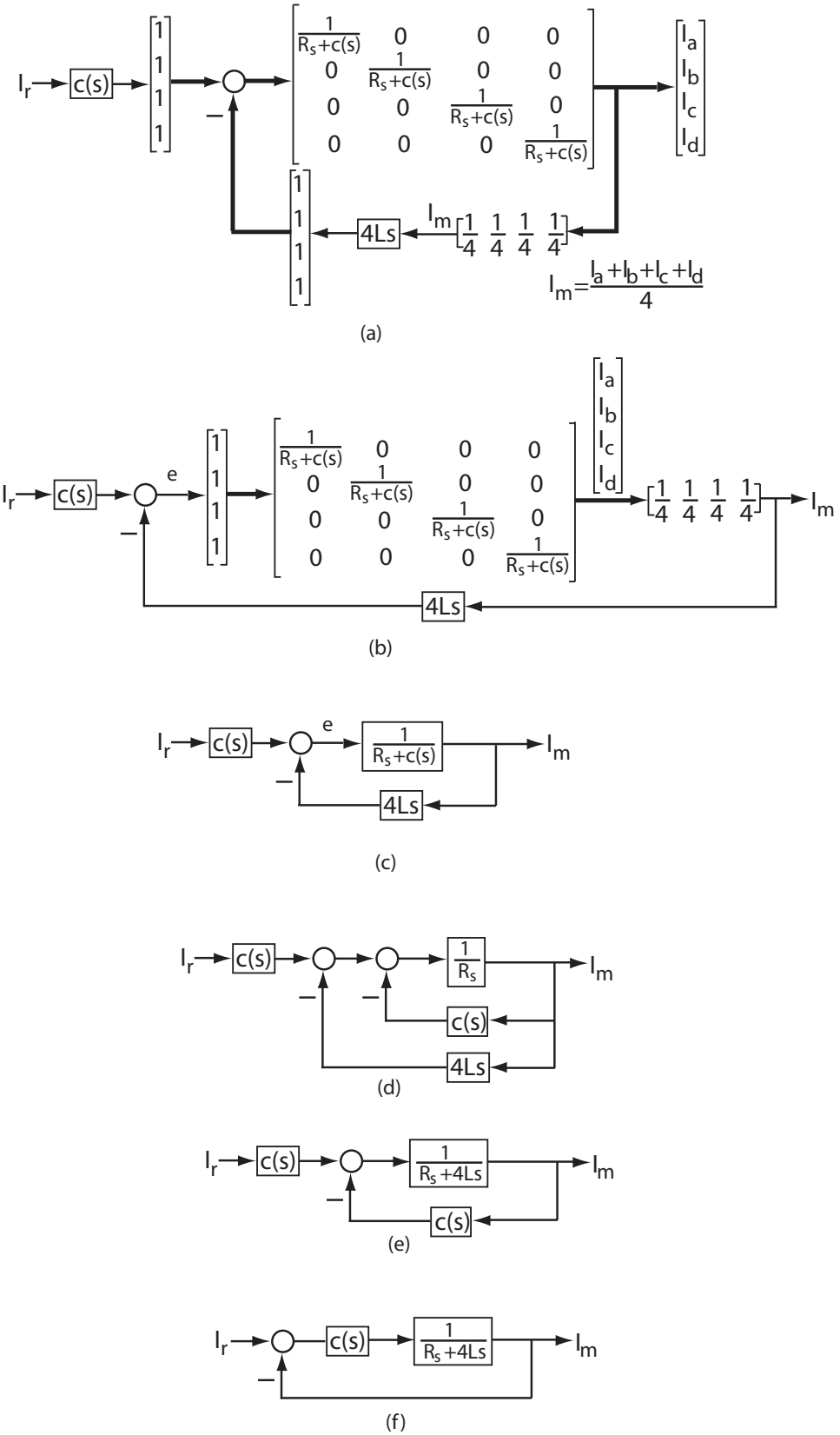


Figure 5-20: Block diagram reduction **II** for MIMO control of fully coupled coils.

with identical inductance (both the self inductance and the mutual inductance) L and resistance R_s, a decentralized controller with diagonal element $c(s)$ can stabilize the whole MIMO system if and only if $c(s)$ can stabilize $\frac{1}{R_s+nLs}$ and $\frac{1}{R_s}$.

However, this compensation method cannot be applied to our SM2-based UFTS, because the fully coupled assumption does not hold due to the low permeability (less than $500\mu_0$) and high eddy currents in SM2 material. It is expected that for the more ideal Ni-Fe-based UFTS this theory can guide the current controller design. In the rest of this section, we will discuss how to compensate the SM2-based UFTS in multiple-channel modes, by taking into account of the leakage flux. We start with the dual-channel mode, and go to the quad-channel mode.

5.5.4 Current Control in the Dual-Channel Mode with Leakage Flux

We now present a tuning method for the SM2-based UFTS in the dual-channel mode as shown in Figure 5-13(b).

The impedance matrix from the current vector to the voltage vector is modelled as

$$\begin{bmatrix} V_a \\ V_b \end{bmatrix} = \begin{bmatrix} Z_{aa} & Z_{ab} \\ Z_{ba} & Z_{bb} \end{bmatrix} \begin{bmatrix} I_a \\ I_b \end{bmatrix}. \tag{5.12}$$

Because of symmetry, we have the following relation

$$Z_{aa} = Z_{bb} \equiv Z_0 \tag{5.13}$$

$$Z_{ab} = Z_{ba} \equiv Z_1. \tag{5.14}$$

With this relation, the transfer matrix can be simplified as

$$\begin{bmatrix} V_a \\ V_b \end{bmatrix} = \begin{bmatrix} Z_0 & Z_1 \\ Z_1 & Z_0 \end{bmatrix} \begin{bmatrix} I_a \\ I_b \end{bmatrix}. \tag{5.15}$$

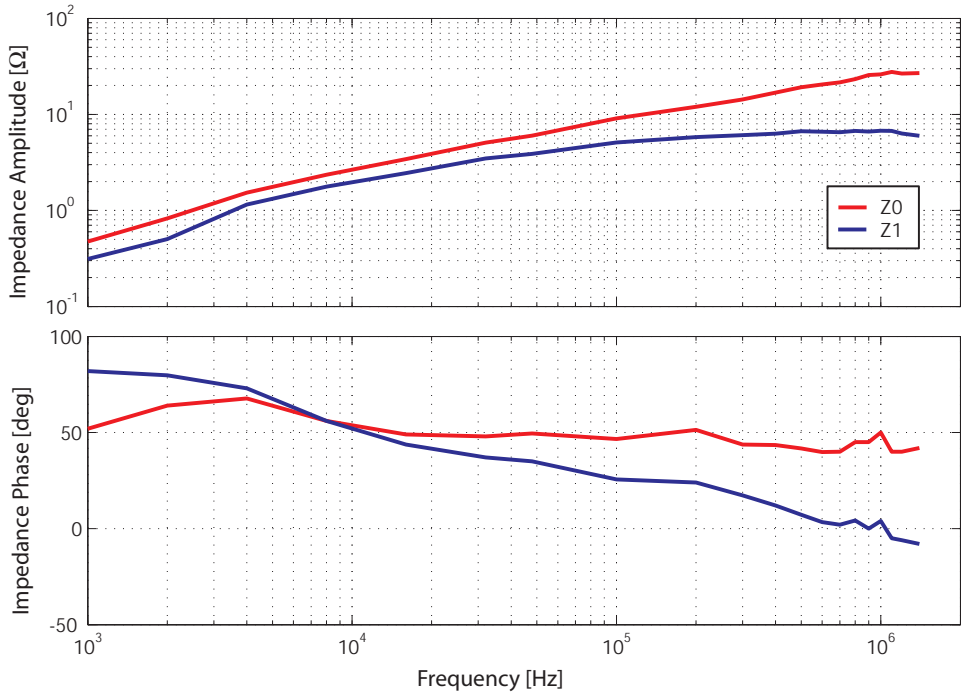


Figure 5-21: Measured impedance frequency response of the SM2-based UFTS in the dual-channel mode.

Note $Z_0 \neq Z_1$ due to the leakage flux.

Figure 5-21 shows the experimentally measured impedance frequency response for Z_0 and Z_1. Their difference shows that this actuator coils are far from fully-coupled.

In order to compensate for this dual-input-dual-output (DIDO) system, we reformulate the relation from the input voltage vector to the output current vector as

$$\begin{bmatrix} I_a \\ I_b \end{bmatrix} = \begin{bmatrix} Z_0^{-1} & 0 \\ 0 & Z_0^{-1} \end{bmatrix} \left(\begin{bmatrix} V_a \\ V_b \end{bmatrix} - \begin{bmatrix} 0 & Z_1 \\ Z_1 & 0 \end{bmatrix} \begin{bmatrix} I_a \\ I_b \end{bmatrix} \right). \tag{5.16}$$

Correspondingly, the block diagram is as shown in Figure 5-22(a). Using the same block diagram transformation method in Section 5.5.3, the DIDO system can be converted into a SISO system through a series of transforms as shown in Figure 5-23(e). Again, we require the element $1/(Z_0 + c(s) - Z_1)$ in Figure 5-23(b) to be stable to ensure that the differential component between I_a and I_b is also stable.

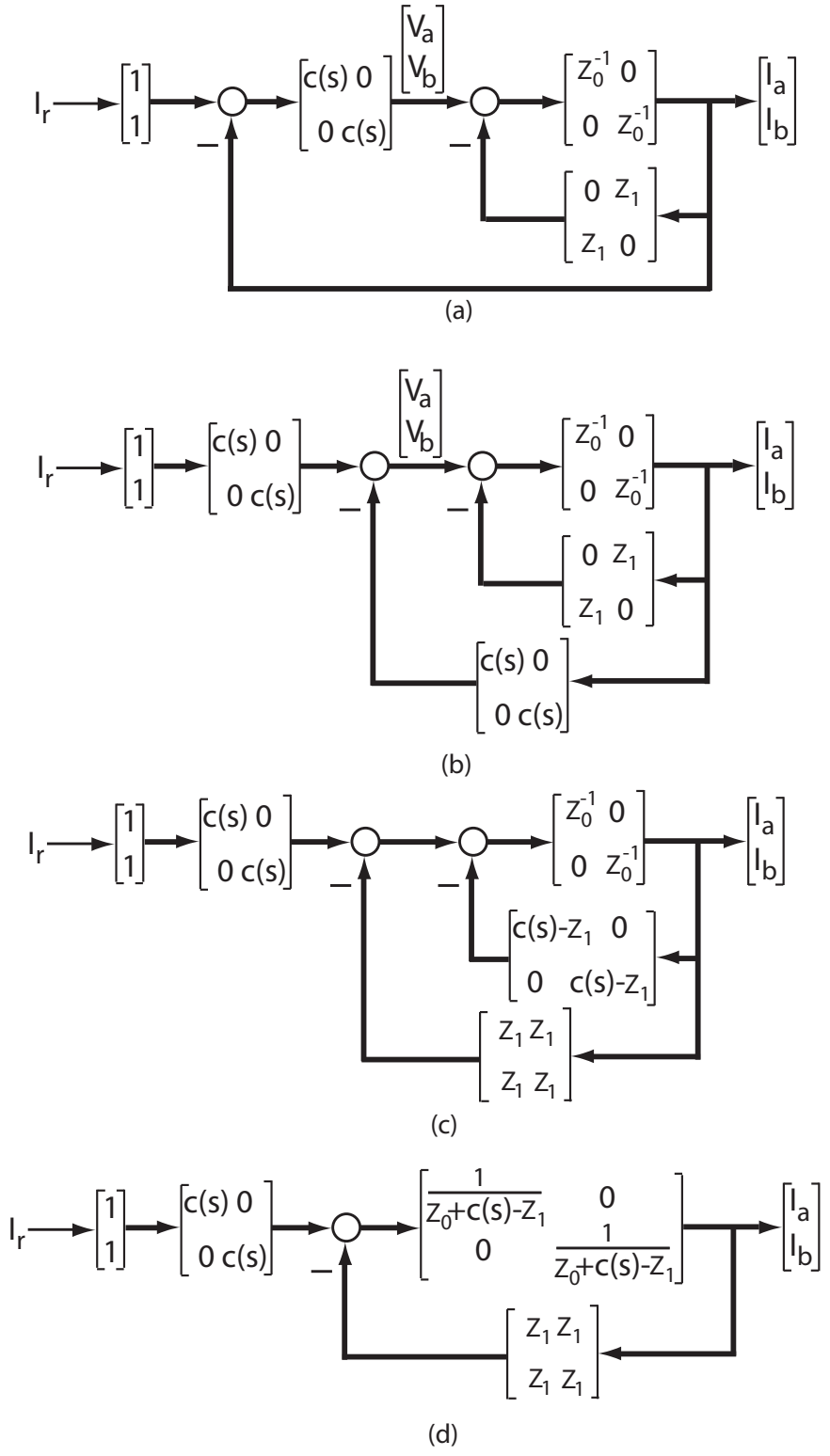


Figure 5-22: Block diagram reduction **I** for DIDO control of the SM2-based UFTS.

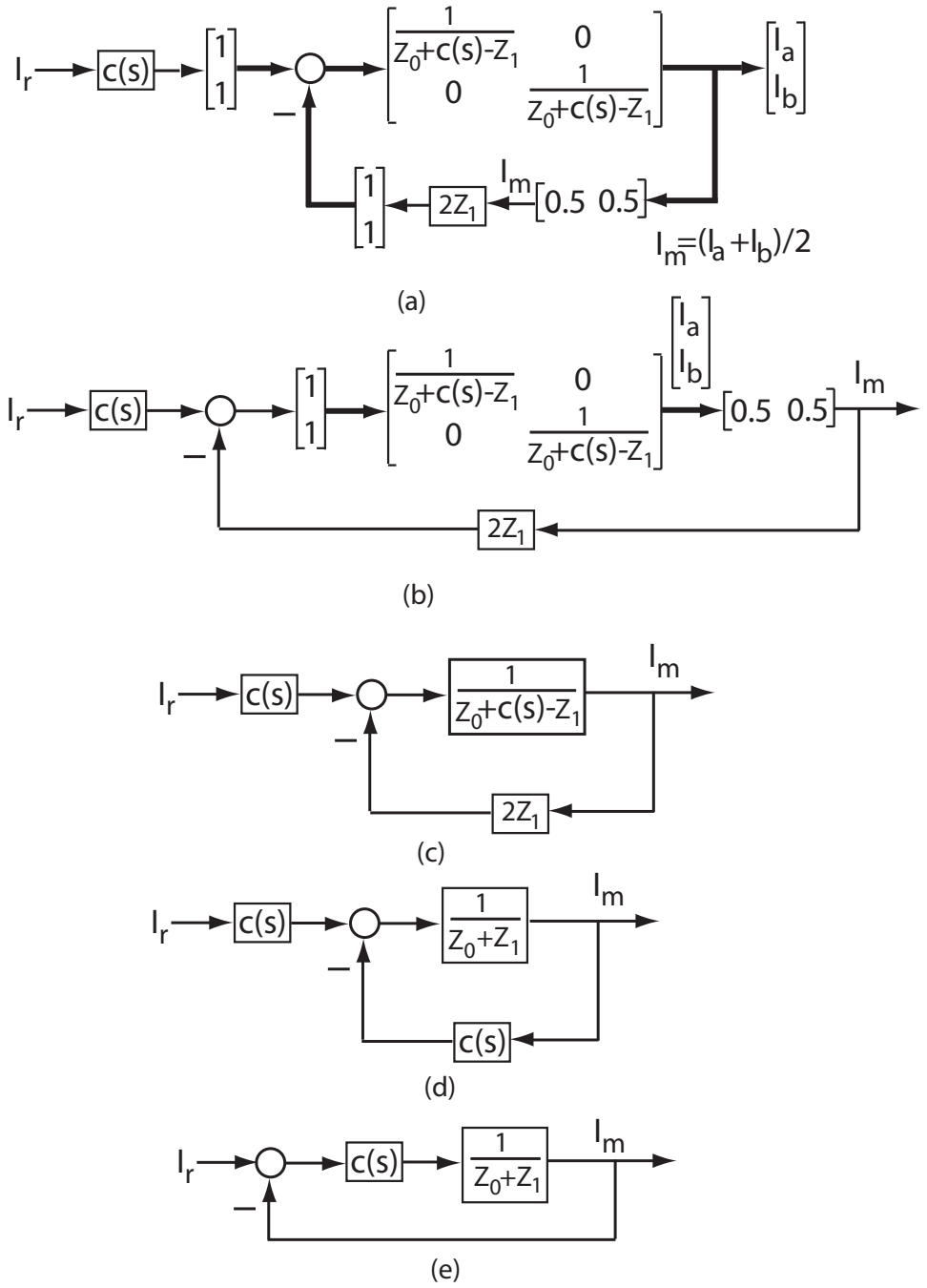


Figure 5-23: Block diagram reduction **II** for DIDO control of the SM2-based UFTS.

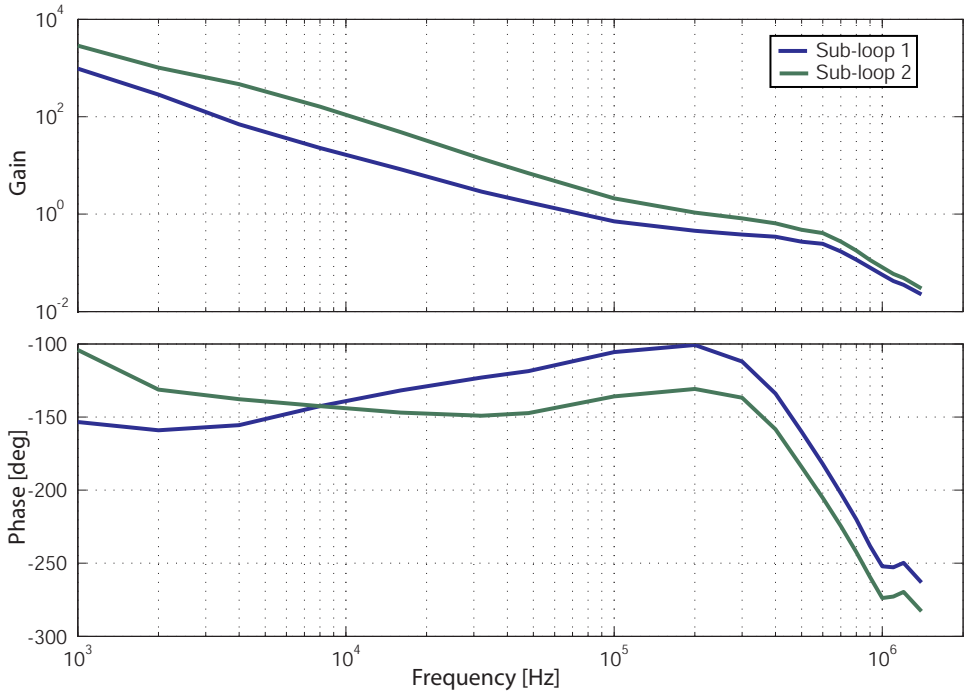


Figure 5-24: Compensated negative sub-loop transmission of the SM2-based UFTS in the DIDO mode. Sub-loop 1 is $c(s)/(Z_0 + Z_1)$, and sub-loop 2 is $c(s)/(Z_0 - Z_1)$.

The above result can be summarized as: **For the DIDO plant as (5.15), a decentralized controller with diagonal element $c(s)$ can stabilize the whole DIDO system if and only if $c(s)$ can stabilize two sub-plants: $\frac{1}{Z_0+Z_1}$ and $\frac{1}{Z_0-Z_1}$, where Z_0 is the self impedance and Z_1 is the mutual impedance.**

Applying this theory to the measured impedance in Figure 5-21, we design the control parameters as follows: $C_{60} = 100$pF, $R_{62} = 1.33$k, $R_{60} = 2$k, and $C_{59} = 1000$pF. With this compensation, Figure 5-24 shows the loop transmission of two sub-loops. The crossover frequency of the sub-loop 1 is 80 kHz with 70 degree phase margin. Sub-loop 2 has a crossover frequency of 200 kHz and 45 degree phase margin. It can be seen that sub-loop 2 is more prone to instability. Based on the previous block diagram transformation, the closed-loop dynamics (from I_r to I_m) is only determined by sub-loop 1. Figure 5-25 shows the measured closed-loop frequency response at two excitation levels of 0.1A and 1A peak-to-valley.

Figure 5-26 shows the 1-A step response of the closed-loop system with a rise time of about 3.5 μs.

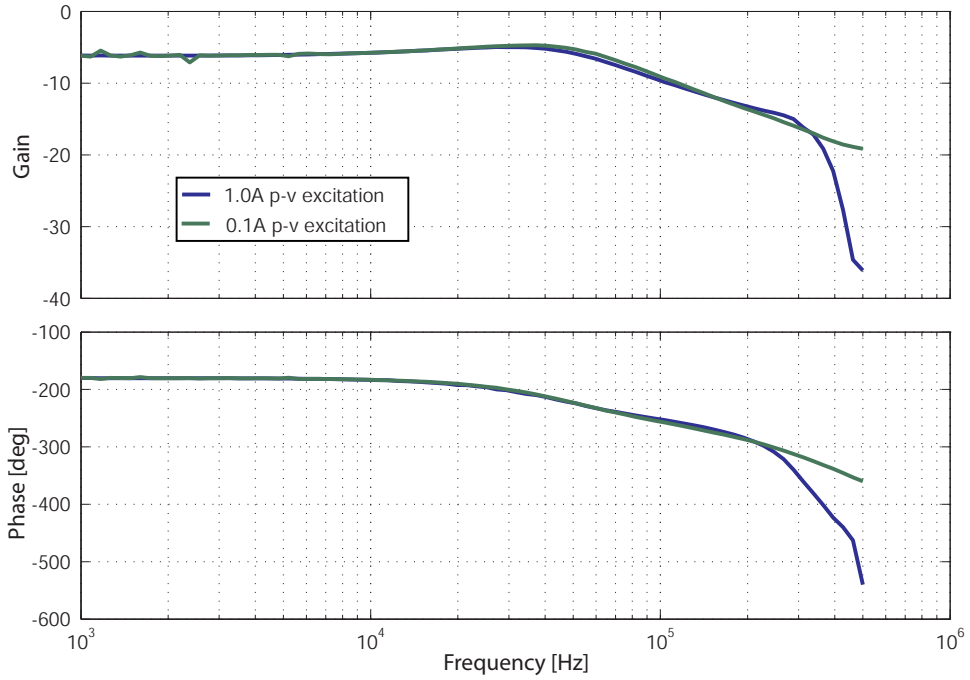


Figure 5-25: Closed-loop frequency response of the SM2-based UFTS in the dual-channel mode. The deviation at high frequencies comes from the slew-rate saturation of the current control OP27 op-amps.

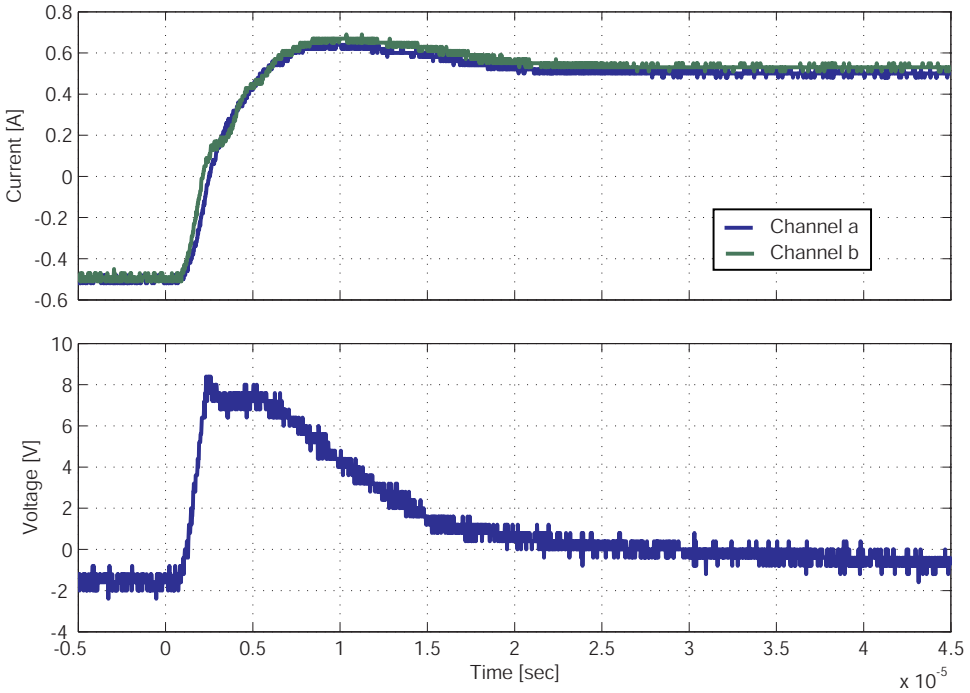


Figure 5-26: Measured 1-A step response of the SM2-based UFTS in the dual-channel mode.

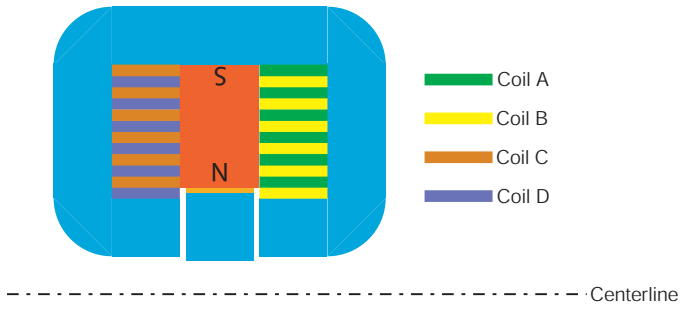


Figure 5-27: Coil winding construction.

5.5.5 Current Control in the Quad-Channel Mode with Leakage Flux

The best way to take full advantage of the designed power amplifier is the quad-channel mode as shown in Figure 5-13(c). First, we present a method to compensate the quad-input-quad-output (QIQO) plant, in which significant leakage fluxes exist. Then this method is applied to the SM2-based UFTS in the quad-channel mode and experimental results are gave.

The 4-by-4 transfer matrix $P(s)$ from the voltage vector to the current vector can be represented as

$$\begin{bmatrix} I_a \\ I_b \\ I_c \\ I_d \end{bmatrix} = P(s) \begin{bmatrix} V_a \\ V_b \\ V_c \\ V_d \end{bmatrix}, \tag{5.17}$$

$$\text{where } P(s)^{-1} = \begin{bmatrix} Z_{aa} & Z_{ab} & Z_{ac} & Z_{ad} \\ Z_{ba} & Z_{bb} & Z_{bc} & Z_{bd} \\ Z_{ca} & Z_{cb} & Z_{cc} & Z_{cd} \\ Z_{da} & Z_{db} & Z_{dc} & Z_{dd} \end{bmatrix}. \tag{5.18}$$

Matrix $P(s)^{-1}$ is the impedance matrix.

The four coils are constructed as shown in Figure 5-27. Coil A and coil B are closely coupled together on the left, and coil C and coil D are closely coupled on the

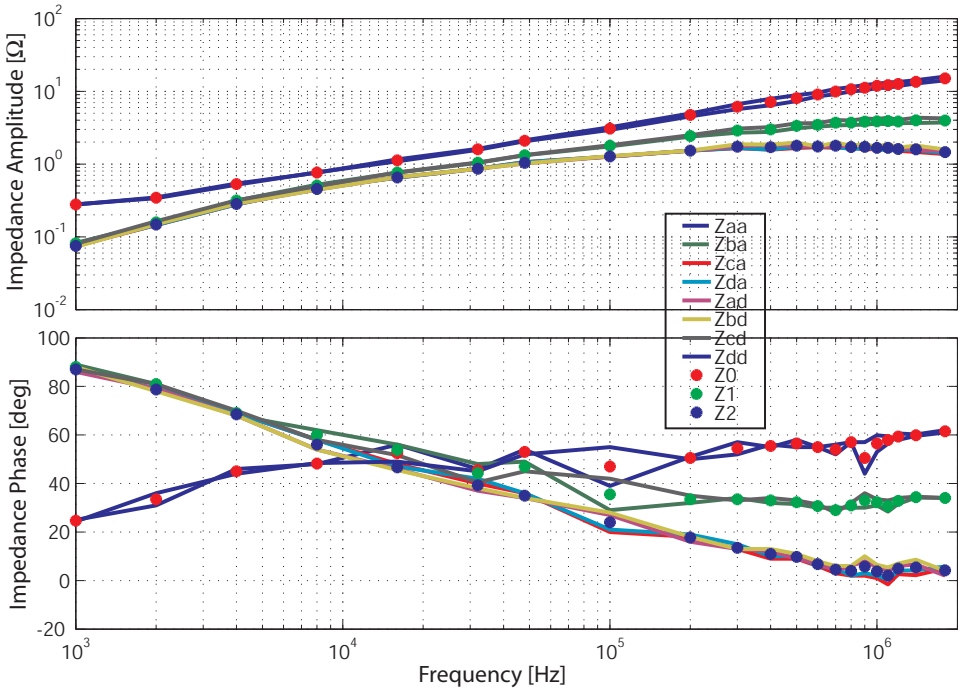


Figure 5-28: Measured impedance frequency responses of the SM2-based UFTS in the quad-channel mode.

other side. Given this assumption, the coupling between A or B on the left and C or D on the right is not so strong as that between two coils on the same side.

Using this symmetrical property, we can classify the elements in the impedance matrix into three categories: self impedance Z_0, closely-coupled impedance Z_1, and loosely-coupled impedance Z_2 as below

$$Z_0 = Z_{aa} = Z_{bb} = Z_{cc} = Z_{dd}; \tag{5.19}$$

$$Z_1 = Z_{ab} = Z_{ba} = Z_{cd} = Z_{dc}; \tag{5.20}$$

$$Z_2 = Z_{ac} = Z_{bc} = Z_{ad} = Z_{bd} = Z_{ca} = Z_{cb} = Z_{da} = Z_{db}. \tag{5.21}$$

This classification is verified by the experimentally measured impedance curves as shown in Figure 5-28. In this figure, Z_0 is the geometric average of Z_{aa} and Z_{dd}, Z_1 is the geometric average of Z_{ab} and Z_{cd}, and Z_2 is the geometric average of Z_{ad}, Z_{bd},Z_{ca}, and Z_{da}.

Consequently, the MIMO plant transfer matrix $P(s)$ can be simplified as

$$P = \begin{bmatrix} Z_0 & Z_1 & Z_2 & Z_2 \\ Z_1 & Z_0 & Z_2 & Z_2 \\ Z_2 & Z_2 & Z_0 & Z_1 \\ Z_2 & Z_2 & Z_1 & Z_0 \end{bmatrix}^{-1}. \tag{5.22}$$

Further, we can relate the input voltage and the output current as

$$\begin{bmatrix} I_a \\ I_b \\ I_c \\ I_d \end{bmatrix} = \begin{bmatrix} Z_0^{-1} & 0 & 0 & 0 \\ 0 & Z_0^{-1} & 0 & 0 \\ 0 & 0 & Z_0^{-1} & 0 \\ 0 & 0 & 0 & Z_0^{-1} \end{bmatrix} \left(\begin{bmatrix} V_a \\ V_b \\ V_c \\ V_d \end{bmatrix} - \begin{bmatrix} 0 & Z_1 & Z_2 & Z_2 \\ Z_1 & 0 & Z_2 & Z_2 \\ Z_2 & Z_2 & 0 & Z_1 \\ Z_2 & Z_2 & Z_1 & 0 \end{bmatrix} \begin{bmatrix} I_a \\ I_b \\ I_c \\ I_d \end{bmatrix} \right). \tag{5.23}$$

This relation corresponds to the block diagram in Figure 5-29(a), which can be transformed into (b) and (c), and further into Figure 5-30(a), (b), and (c). In Figure 5-30(c), we define two new virtual variables as

$$I_{m1} = \frac{I_a + I_b}{2}, \tag{5.24}$$

$$I_{m2} = \frac{I_c + I_d}{2}. \tag{5.25}$$

As the original four outputs I_a, I_b, I_c, and I_d are reduced into I_{m1} and I_{m2}, the differential components $I_a - I_b$ and $I_c - I_d$ become uncontrollable and unobservable. To ensure the stability of the overall system, each diagonal element in Figure 5-30(c) $1/(c(s) + Z_0 - Z_1)$ should be stable. Equivalently, the controller $c(s)$ must be able to stabilize the sub-plant $P_2(s) = 1/(Z_0 - Z_1)$.

Continuing with Figure 5-31, the two-input-two-output system in (a) can be simplified into (c), in which the diagonal element $1/(c(s) + Z_0 + Z_1 - 2Z_2)$ should be stable. As a result, the controller $c(s)$ must be able to stabilize sub-plant $P_3(s) = 1/(Z_0 + Z_1 - 2Z_2)$.

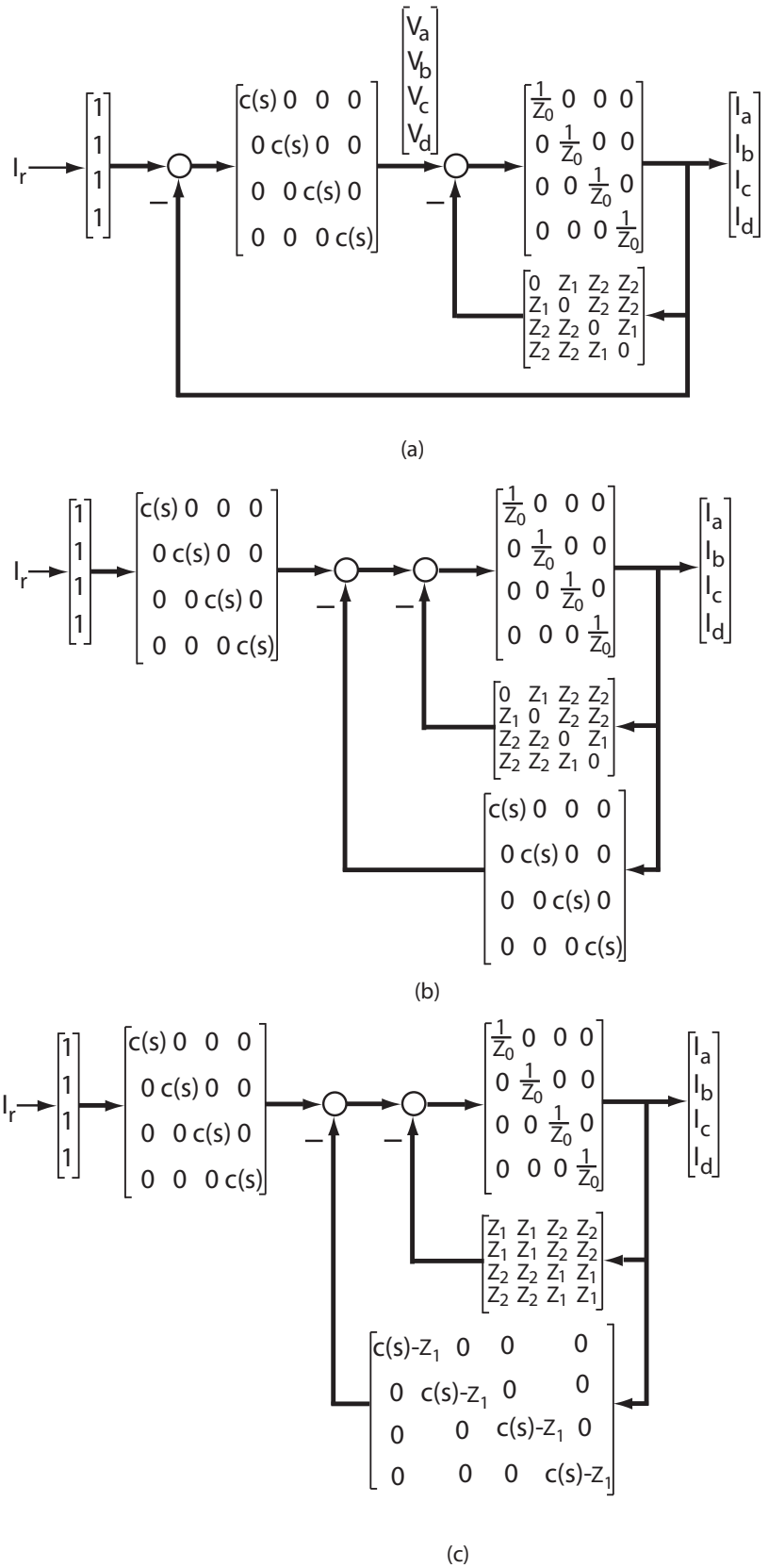


Figure 5-29: Block diagram reduction **I** for QIQO control of the SM2-based UFTS.

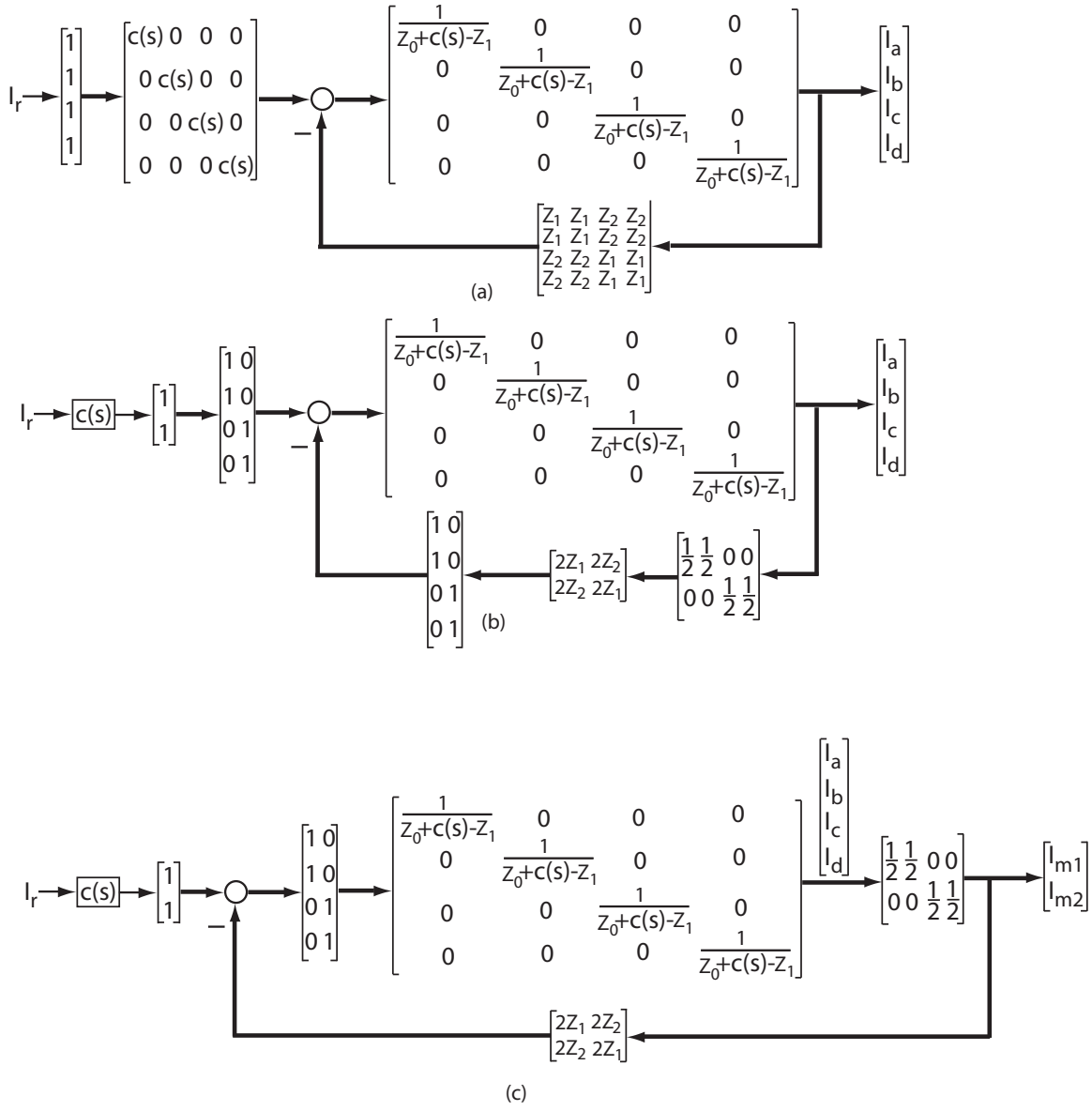


Figure 5-30: Block diagram reduction **II** for QIQO control of the SM2-based UFTS.

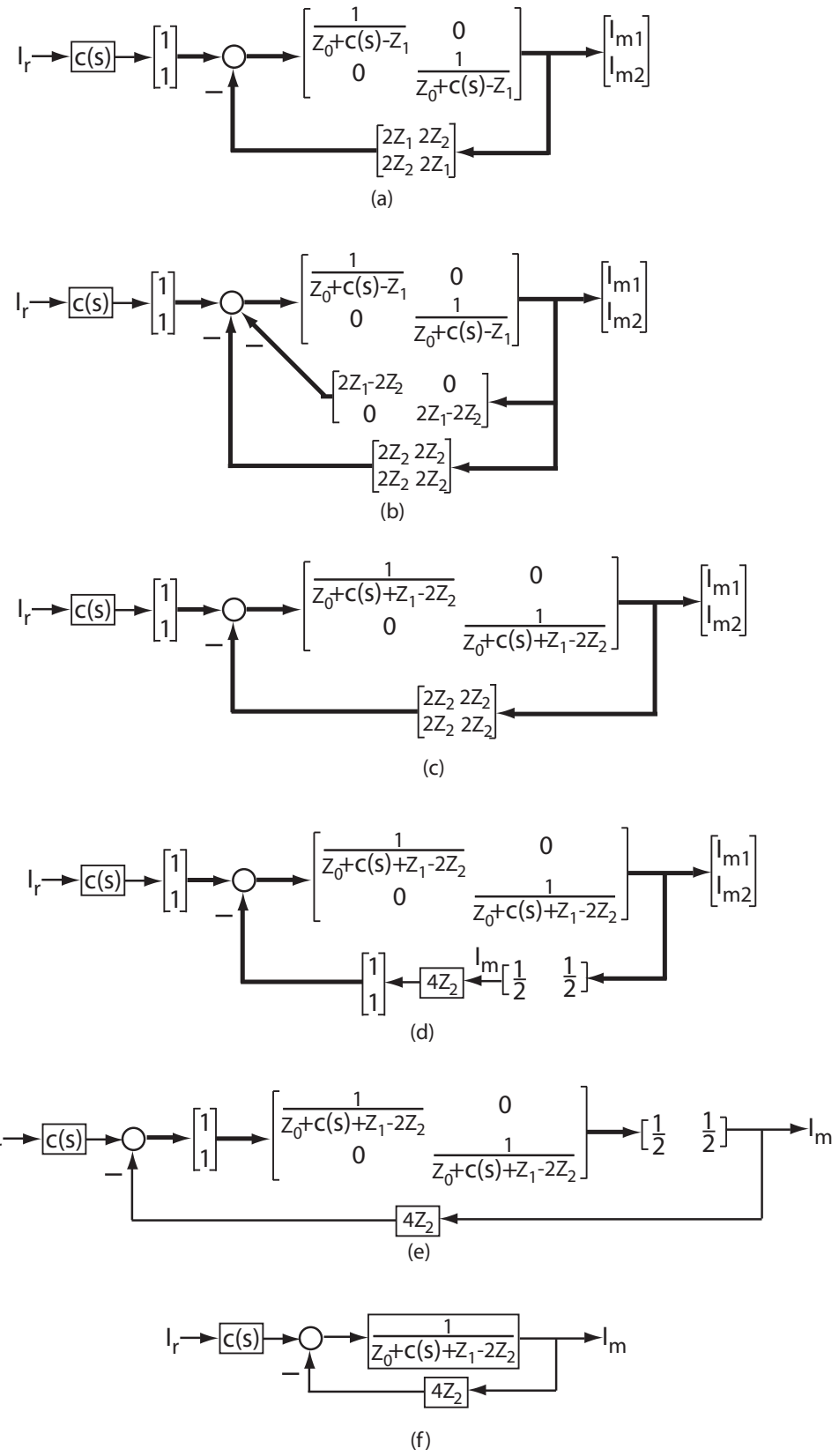


Figure 5-31: Block diagram reduction **III** for QIQO control of the SM2-based UFTS.

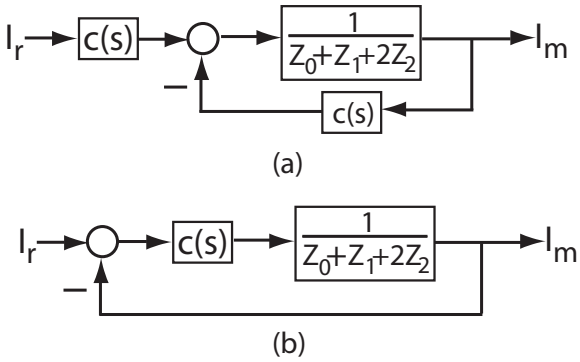


Figure 5-32: Block diagram reduction **IV** for QIQO control of the SM2-based UFTS.

In Figure 5-32, the system is transformed into a new SISO system with output I_m, which is the mean value of all channel currents. To ensure the stability of this SISO system, the controller $c(s)$ must be able to stabilize $P_1(s) = 1/(Z_0 + Z_1 + 2Z_2)$.

This leads to the following result. **For the QIQO multi-start actuator coils in Equation 5.22, a decentralized controller with diagonal element $c(s)$ can stabilize the whole MIMO system if and only if $c(s)$ can stabilize three sub-plants $P_1(s)$, $P_2(s)$, and $P_3(s)$, as defined above.**

Applying this result to the measured impedances in Figure 5-28, we design the current control circuit parameters as $C_{60} = 150$pF, $R_{62} = 1$k, $R_{60} = 1$k, and $C_{59} = 1200$pF for each channel. The loop transmissions of the three compensated sub-plants are shown in Figure 5-33. The crossover frequency of sub-loop 1 $c(s)P_1(s)$ is 100 kHz with 65 degrees phase margin. Sub-loop 2 $c(s)P_2(s)$ has a crossover frequency of 400 kHz and 40 degrees phase margin. In sub-loop 3 $c(s)P_3(s)$, the crossover frequency is 200 kHz and the phase margin is 50 degrees. The overall average-current closed-loop dynamics is only determined by sub-loop 1.

Figure 5-34 shows the closed-loop frequency response at various excitation levels. At large signal excitation, the frequency response significantly deviates from the small signal response. The deviation is caused by the fact that the current control op-amp OP27 has a measured slew rate saturation limit of 0.8V/μs, which is less than the claimed value of 2.7V/μs. As our design is only expected to be used up to 20 kHz at full power level, this slew rate saturation around 100 kHz will not be a problem.

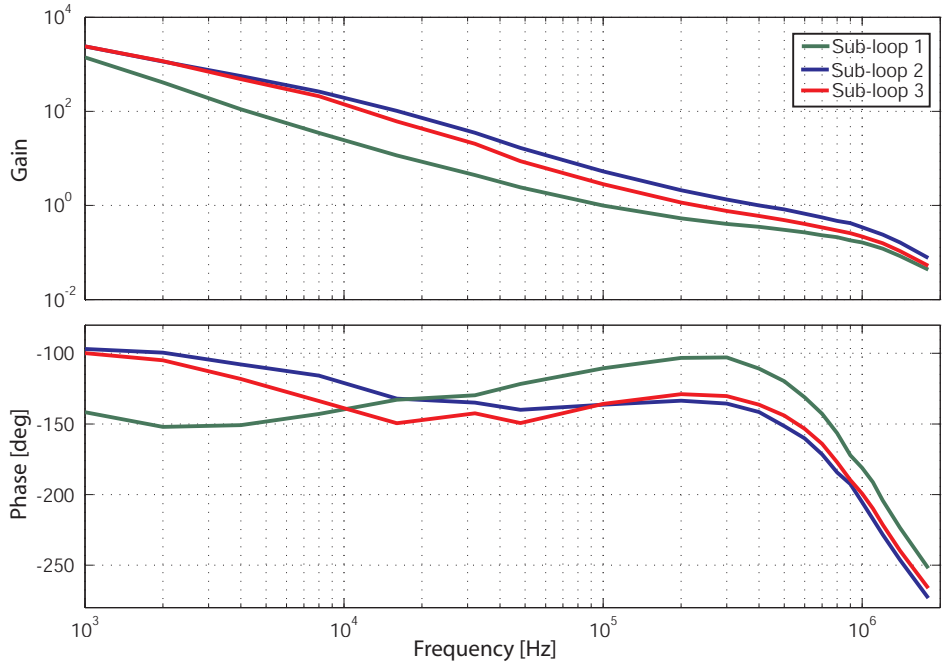


Figure 5-33: Loop transmissions of the compensated sub-plants of the SM2-based UFTS in the quad-channel mode.

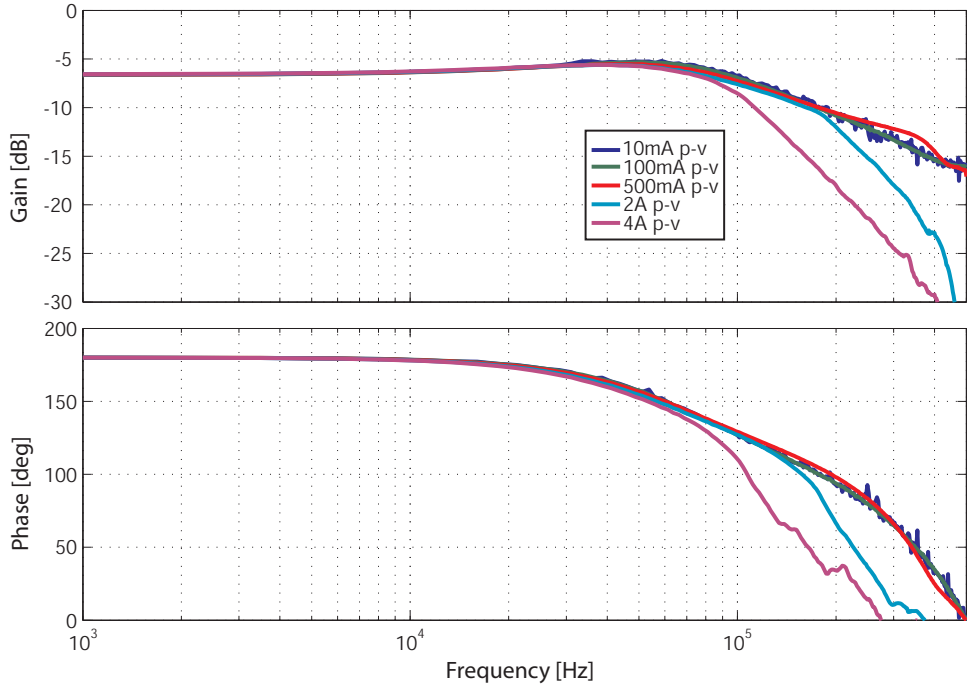


Figure 5-34: Measured frequency response of the closed-loop current amplifier driving the SM2-based UFTS in the quad-channel mode.

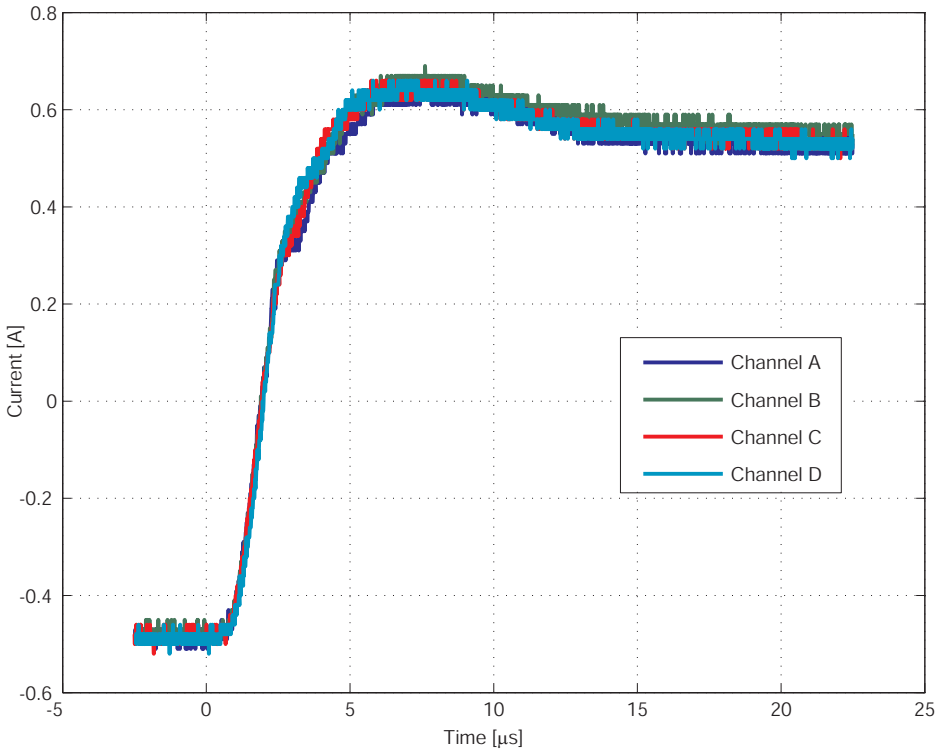


Figure 5-35: Measured step response of the closed-loop current amplifier in the quad-channel mode.

When designing a power amplifier for operation beyond 100 kHz, an op-amp with higher slew rate should be used.

Figure 5-35 shows a measured 1-A step response of the closed-loop system.

5.6 Generalized Theory of the Decentralized Current Controller for MIMO Electrical Loads

In this section, we extend the previous results to more generalized results for the current control of electrical loads. If a n-input and n-output plant $P(s)$ can be

expressed as

$$\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} = P(s) \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix}, \tag{5.26}$$

and its impedance matrix P^{-1} can be factorized as

$$P^{-1}(s) \;=\; VD^{-1}(s)V^{-1}, \tag{5.27}$$

where $D(s)$ is a diagonal matrix with diagonal elements: $D_1(s)$, $D_2(s)$, ..., $D_n(s)$ and V is an invertible minimum phase system, then a decentralized controller $C(s)$ with identical diagonal elements $c(s)$ can stabilize the original plant $P(s)$, if and only if $c(s)$ can stabilize all SISO sub-plants $D_i(s)$, where $i = 1, ..., n$.

Proof: When the MIMO system $P(s)$ is controlled by the decentralized controller C(s), the negative loop transmission matrix is

$$-L.T. \;=\; P(s)C(s). \tag{5.28}$$

The eigenvalues of matrix $P(j\omega)C(j\omega)$ can be plotted as closed loci on the complex plane when ω changes from $-\infty$ to ∞. These loci are referred as characteristic loci. According to the generalized Nyquist stability theory [46], the closed-loop system of $P(s)$ controlled by $C(s)$ is stable if and only if the characteristic loci $\lambda\left(-L.T.(j\omega)\right)$ has U times anti-clockwise encirclement around "-1" point, where U is the number of unstable poles in $P(s)C(s)$.

Applying this theory to our system, the characteristic loci are expressed as

$$\begin{aligned} \lambda\left[P(j\omega)C(j\omega)\right] &= \lambda\left[VD(j\omega)V^{-1}C(j\omega)\right] \\ &= \lambda\left[VD(j\omega)C(j\omega)V^{-1}\right] \\ &= \lambda\left[D(j\omega)C(j\omega)\right] \end{aligned}$$

$$
\begin{aligned}
&= \lambda\left(\begin{bmatrix} D_1(j\omega)c(j\omega) & 0 & \cdots & 0 \\ 0 & D_2(j\omega)c(j\omega) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & D_n(j\omega)c(j\omega) \end{bmatrix}\right) \\
&= \{D_1(j\omega)c(j\omega),\ D_2(j\omega)c(j\omega),\ ...,\ D_n(j\omega)c(j\omega)\}. \qquad (5.29)
\end{aligned}
$$

As a result, the stability of the original MIMO system is equivalent to that each of $D_i(s)c(s)$ satisfy the SISO Nyquist stability criteria for $i = 1, ..., n$ or say that $c(s)$ can stabilize all $D_i(s)$. This completes the proof.

As a special case, we apply this generalized theory to the QIQO system in Eqation 5.22,

$$
\begin{aligned}
P^{-1}(s) &= \begin{bmatrix} Z_0 & Z_1 & Z_2 & Z_2 \\ Z_1 & Z_0 & Z_2 & Z_2 \\ Z_2 & Z_2 & Z_0 & Z_1 \\ Z_2 & Z_2 & Z_1 & Z_0 \end{bmatrix} \\
&= \begin{bmatrix} -1 & 0 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{Z_0-Z_1} & 0 & 0 & 0 \\ 0 & \frac{1}{Z_0-Z_1} & 0 & 0 \\ 0 & 0 & \frac{1}{Z_0+2Z_2+Z_1} & 0 \\ 0 & 0 & 0 & \frac{1}{Z_0-2Z_2+Z_1} \end{bmatrix}^{-1} \\
&\quad \begin{bmatrix} -1 & 0 & 1 & -1 \\ 1 & 0 & 1 & -1 \\ 0 & -1 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{bmatrix}^{-1} . \qquad (5.30)
\end{aligned}
$$

Based on this factorization, we see that $D_1(s) = 1/(Z_0 - Z_1)$, $D_2(s) = 1/(Z_0 - Z_1)$, $D_3(s) = 1/(Z_0 + Z_1 + 2Z_2)$, and $D_4(s) = 1/(Z_0 + Z_1 - 2Z_2)$. According to the generalized theory, the stability of the closed-loop QIQO system is equivalent to that $c(s)$ can simultaneously stabilize D_1, D_2, D_3, D_4. This is the same as the result in the previous section. Hence, this generalized theory illustrates the mathematical

foundation underlying the block transformation discussed before.

This generalized theory can be applied to a wide range of MIMO current controller designs. For example, by applying this theory we have derived the current control method for the ultra fast rotary motors of various winding modes and coil connections in Appendix 2.

5.7 Flux sensing and control

This section introduces a flux control method, which may be integrated into the power amplifier to improve the linearity of the UFTS. Due to time limitations, we have not tested this method experimentally.

As discussed in Chapter 3, the normalized actuating force f is a linear function of the normalized excitation flux λ

$$f = \eta\lambda, \tag{5.31}$$

$$\lambda = i + \eta x. \tag{5.32}$$

If the amplifier can feed back the flux and thus control the excitation flux directly, better linearity may be achieved. Next, we will describe a possible flux feedback and control method.

Figure 5-36 shows the proposed flux control circuit and the flux control block diagram. The output of a secondary flux sensing coil is fed into an instrument amplifier INA111 with a gain of 1 and a closed-loop bandwidth of 5MHz. The 100k resistor at the positive input terminal to INA111 is used as a bias current path. Then the output of INA111 is summed together with the armature displacement X and the feedback current I_s with proper gain scaling at the flux summing stage.

The detailed block diagram of the whole flux control circuit is shown in Figure 5-36(b). The synthesized signal S_1 is the weighted sum of feedback current I and the armature position X to reconstruct signal $\lambda = i + \eta x$. In parallel, S_2 represents the direct flux feedback signal. According to (5.32), S_1 and S_2 are redundant signals

of the same physical meaning. However, there exist some differences between them. The signal to noise ratio (SNR) of S_2 is very poor at low frequencies but is very good at high frequencies. The SNR of S_1 is relatively independent of frequency, but is not an accurate indication of flux for all frequencies because of the leakage flux, magnetic material nonlinearity, and eddy currents. A synthesized signal S takes the low-frequency component of S_1 and the high-frequency component of S_2 by a complementary filter at the flux summing stage. This signal S then has advantages from both S_1 and S_2. At low-frequency operation S_1 dominates S_2, and thus the amplifier operates in current mode. At high frequencies, the current signal is rolled off and the signal from the secondary coil is used. In the complimentary filter, no matter what time constant τ we set, we always have: $S = S_1 = S_2$. As a result, we can use the same current controller parameters as designed before.

5.8 Summary

This section presents a linear current power amplifier design with 1 KW power dissipating ability and 100 kHz bandwidth, which is an essential component of our UFTS system.

The overall power amplifier has four channels, and each channel has two cascaded loops for voltage control and current control respectively. We set the voltage stage at a gain of 5 with 1 MHz bandwidth. In this stage, the input and output of the power device PA52A should be protected from input over driven, boost circuit failure, and so on.

The current control is more complicated when multiple channels work in parallel, due to the strong coupling among coils of the UFTS. We first present a decoupling control method for driving coils without leakage flux. Unfortunately, the SM2-based UFTS has significant flux leakage. To address this issue, we discuss in detail the decoupling control issues for the SM2-based UFTS operated in both the dual-channel mode and the quad-channel mode. Finally, these results are extended to a generalized decoupling control theory for power amplifiers to drive multi-start loads in parallel.

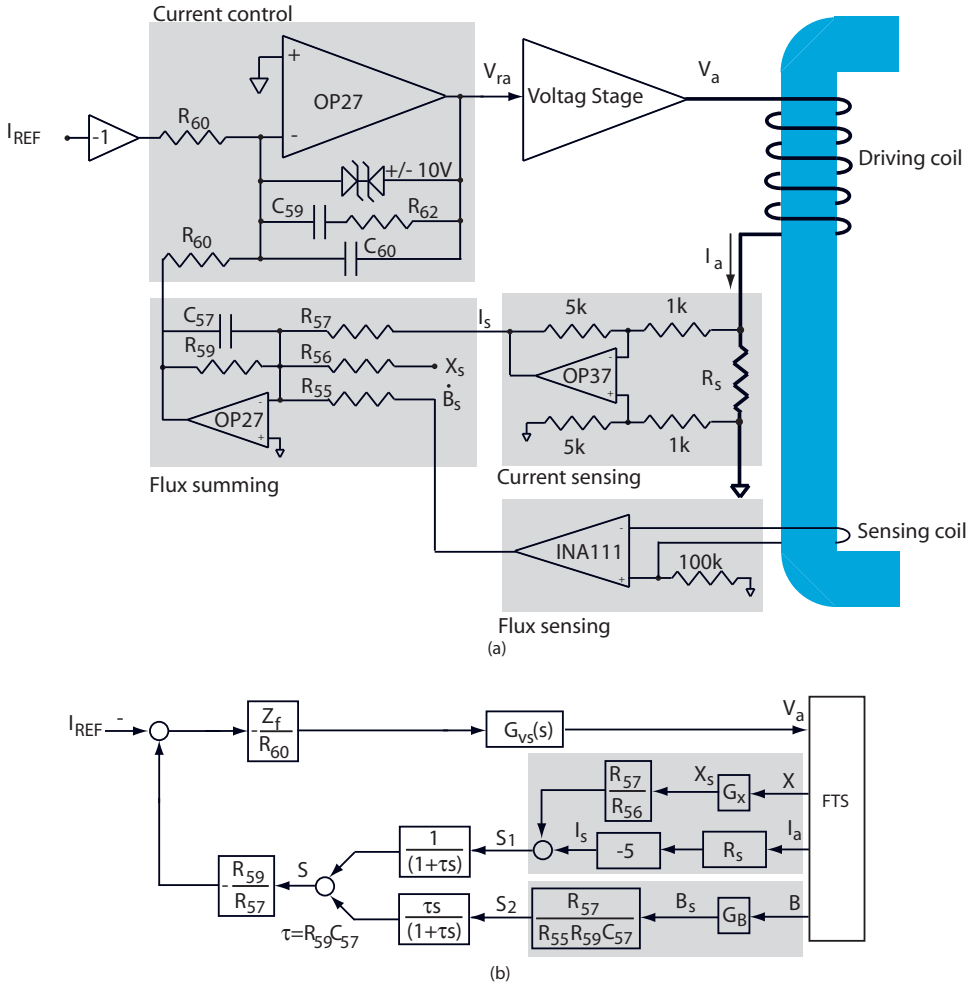


Figure 5-36: Flux sensing and control circuit implementation and block diagram.

Finally, we briefly introduced a flux feedback and control method, which may be used to improve the linearity of the UFTS, but has not yet been experimentally verified.

Chapter 6

Real Time Computer

To control the UFTS with a closed-loop bandwidth of over 20 kHz, the controller implementation includes both digital and analog solutions. Obviously, an op-amp based analog controller has the advantages of high bandwidth, low latency, low cost, and fast prototyping. However, there exist some problems in the analog implementation of controllers. First, an analog controller is of low flexibility, because the implementable algorithms are basically limited to gain, summing junction, and transfer functions. Algorithms such as angle-based adaptive feed-forward cancellation and/or 2-D surface trajectory generation are very difficult to implement using analog circuits. Secondly, the dynamic range and noise of an op-amp become very difficult to deal with when the controllers become complicated. Thirdly, the gain-bandwidth product is limited to about 10 MHz for common op-amps. Last, the interface between a analog controller and digital sensors, such as quadrature digital output, becomes a problem. All these problems can be handled very easily in a digital implementation. Once the sensor outputs are digitized, they can be manipulated with much more flexibility, especially in floating-point digital processors.

This chapter discusses the real-time computer design, which is a digital implementation of our controllers. Section 6.1 presents digital control fundamentals, which reveals that various delays caused in digital controls are key limiting factors on the achievable bandwidth of closed-loop systems. In Section 6.2 we examine the state of art architecture design of real-time control computers and then we present a novel

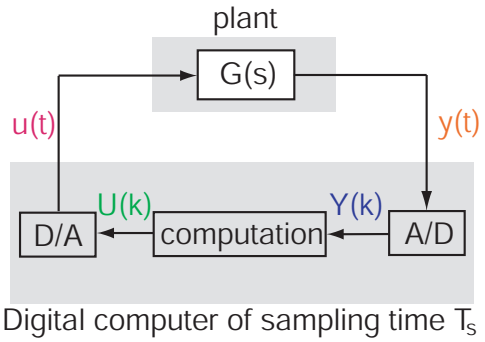


Figure 6-1: Block diagram of general digital control systems.

architecture design which can potentially achieve much smaller delays. Section 6.3 describes the function block diagram design for a computer that implements this novel architecture. Its unique encoder interface and serial interface design is presented in Section 6.4 and Section 6.5. We describes the hardware PCB design in Section 6.6 and Section 6.7. The graphical user interface is described in Section 6.8. We introduce the test results of the designed computer in Section 6.9.

6.1 Digital Control Fundamentals

A typical digital control system can be modelled by the block diagram shown in Figure 6-1. The output $y(t)$ from an analog plant $G(s)$, including anti-aliasing filters as part of the plant, is first sampled by an A/D convertor to generate a discreet variable $Y(k)$. These signals are related as

$$Y(k) \quad = \quad y(kT_s) + e_{AD}, \tag{6.1}$$

where e_{AD} is the A/D conversion error. But this relation only reflects their magnitude relationship. To model the relation in time axis, Figure 6-2 shows the time history of all variables, no matter whether they are continuous or discreet. In the time axis, there exists a conversion time delay T_{AD} between $y(t)$ and $Y(k)$. According to the A/D result $Y(k)$, the control output $U(k)$ is calculated as

$$U(Z^{-1}) \quad = \quad C(Z^{-1})Y(Z^{-1}), \tag{6.2}$$

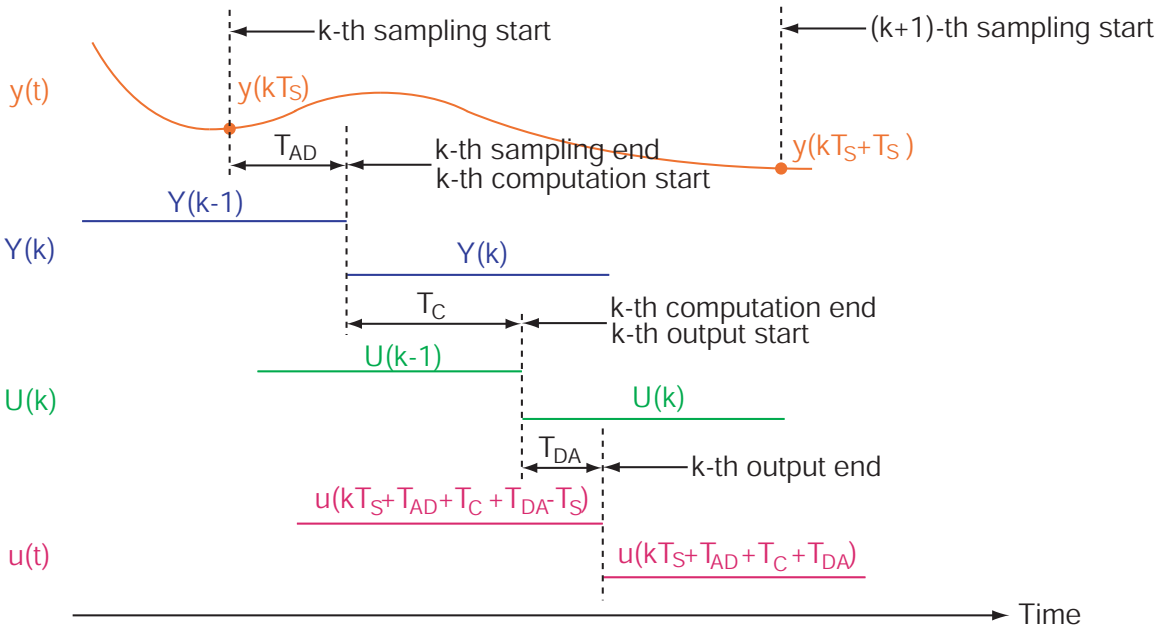


Figure 6-2: Time history of variables in the digital control system.

where $C(Z^{-1})$ represents the control algorithm implemented in the computation block. This computation process takes a time duration of T_c. Next, U(k) is converted to a continuous variable $u(t)$, with an associated delay T_{DA}

$$u(kT_s + T_{AD} + T_C + T_{DA}) \quad = \quad U(k) + e_{DA}, \tag{6.3}$$

where e_{DA} is the D/A conversion error. Finally, the D/A result $u(t)$ goes back to the plant input.

Based on this process, we can establish a more precise block diagram for the sampled-data system as shown in Figure 6-3. A delay element is introduced with $T_d = T_{AD} + T_C + T_{DA}$ to represent the delay resulting from the actual processes. In this way, the discrete-time transfer function, the sampler, and the ZOH become ideal time-free blocks.

According to the sampled-data theory in [65], the Z-domain equivalent transfer function from input U(k) to output Y(k) becomes

$$G_z(Z) \quad = \quad \frac{Y(Z)}{U(Z)}$$

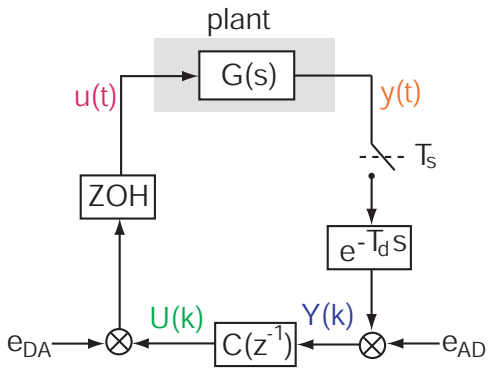


Figure 6-3: Sample-data system model of digital control systems.

$$= \mathbf{Z}\{\frac{e^{-T_d s}(1-e^{-T_s s})G(s)}{s}\}, \tag{6.4}$$

where T_s is the sampling time of the digital control system and $\mathbf{Z}(\cdot)$ is the Z-transform of samples of the impulse response. Further, its frequency response can be expressed as

$$G_z(e^{j\omega T_s}) = \sum_{n=-\infty}^{\infty} \frac{1}{T_s} P\left(j(\omega - n\omega_s)\right), \tag{6.5}$$

$$\text{where } P(j\omega) = \frac{e^{-j\omega T_d}(1-e^{-j\omega T_s})G(j\omega)}{j\omega}, \tag{6.6}$$

$$\text{and } \omega_s = \frac{2\pi}{T_s}. \tag{6.7}$$

When $P(j\omega)$ is negligible in the frequency range of $\omega > \omega_s/2$ by proper anti-aliasing design and/or the rolling-off of G(s) itself, (6.5) can be simplified as

$$G_z(e^{-j\omega T_s}) = \frac{1}{T_s} P\left(j\omega\right). \tag{6.8}$$

Further, for the frequency region of $\omega < \omega_s/4$ where the feedback control is mainly concerned, the zero-order-hold element $(1-e^{-T_s s})/s$ can be approximated by a delay element $T_s e^{-\frac{T_s s}{2}}$ with less than 10% gain error and no phase error. As a result, in the region of $\omega < \omega_s/4$ we can represent the discrete domain plant transfer function as

$$G_z(e^{-j\omega T_s}) \simeq e^{-j\omega(T_d + \frac{T_s}{2})} G(j\omega), \text{ for } \omega < \frac{\omega_s}{4} \tag{6.9}$$

Note that (6.9) is true only when the aliasing component is negligible. If this assumption doesn't hold, then both sides of (6.9) can be very different. For example, when connecting the D/A output to the A/D input of a digital computer directly by a wire (loop-back setup), the measured time delay by a sweeping sine-wave identification from $U(e^{j\omega})$ to $Y(e^{j\omega})$ is not the computation and sampling delay, because in this setup $G(s)$ equals to 1 and thus significant aliasing exists. However, according to (6.4), we can get $G_z(e^{-j\omega T_s}) = e^{-j\omega k T_s}$ for this loop back set up, where k is the minimum integer greater than $T_d + T_s/2$. This analytical result matches the experimental measurement very well.

From (6.9), we can see that the time-delay element (non-minimum phase element) $e^{(T_d+\frac{T_s}{2})s}$ is one of the main limiting factors for the achievable bandwidth. From this point of view, a 20 MHz $\Sigma - \Delta$ A/D convertor with 10 μs pipeline delay is almost as bad as a 100k Hz SAR A/D convertor. Of course any non-minimum phase element in $G(s)$ (such as the force transmission delay in the UFTS backbone ceramic tube discussed in Chapter 4 also further limits the achievable bandwidth.

A rule of thumb is that the sampling frequency of a digital control system needs to be at least 20 times the intended cross-over frequency of the loop transmission, if the dynamics of the plant $G(s)$ is clean without non-minimum phase elements. This rule works quite well for a number of computer systems that I used. When I used a dSPACE DS1003 to control the dual-solenoid FTS that I designed, the DS1003 system maximum sampling frequency was about 20 kHz due to its hardware limits, and the achievable loop-transmission crossover frequency got stuck at 1 kHz. When I used a dSPACE DS1103 to control the same FTS, the maximum sampling frequency was increased to 100 kHz, and the corresponding achievable crossover frequency was 5 kHz. In these experiments, the achievable crossover frequency was achieved with at least a 30 degree phase margin and 3 dB gain margin.

For our UFTS, an intended 20 kHz crossover frequency of loop transmission, the minimum sampling frequency is at least 400kHz based on the above rule of thumb. Considering that there exist mechanical resonance, the capacitance probe filters, and the force transmission delay, we need a digital computer capable of about 1 MHz

real-time throughput. This is the design goal of the real-time computer discussed in this chapter.

6.2 Architecture Design

The time delay T_d discussed in the previous section is significantly related to the architecture of a real-time computer. We first review the state of art real-time computer design, and then present a new architecture, which can potentially reduce the time delay and the sample time significantly.

6.2.1 Review of Existing Real-Time Computer Architecture

Real-time computers oriented to feedback control applications need to accomplish at least two distinctive tasks: (1) a periodic real-time service (RTS) and (2) a graphical user interface (GUI). Each execution of a RTS includes feedback signal acquisition (such as A/D conversion), control algorithm computation, and control variable output (such as D/A conversion), typically at a fixed frequency. This execution process usually does not need a lot of CPU load but requires low latency. The GUI provides an interface through which a user can input supervising actions, including tuning parameters in the RTS, monitoring and logging RTS variables for display or later analysis. The GUI service usually needs a fair amount of CPU load, but the response time can be much longer than that of the RTS, because human beings can only respond in the order of seconds in any case.

Two system architectures are commonly used to accomplish these two tasks. The earliest computer control system used a foreground-background architecture within one computer, for instance, a PC. The RTS was executed in an interrupt as a background task, while the GUI was a task executed in foreground. This architecture was widely used before the Windows operating system came to the market. We refer this type of system as Uni-Body Architecture as shown in Figure 6-4. The latency marked in orange is mainly caused by the interrupt response time, such as the time used to push all the current CPU register values into a stack before entering the inter-

rupt and the time to initialize the CPU instruction pipelines. However, this latency should not be counted into T_d in (6.9), because this delay occurs before the sampling start (the starting point of the A/D conversion). The GUI service basically fill the gaps among background execution time slots. According to this figure, the sampling time T_s should be the worst case of RTS execution time (in green) plus latency (in orange) and a marginal time. If the reserved marginal time is set too small, then the foreground will be slowed down significantly, because there are not many gaps to fill in. The data exchange between GUI and RTS is implemented via a shared memory (such as global variables that tasks in both grounds can access).

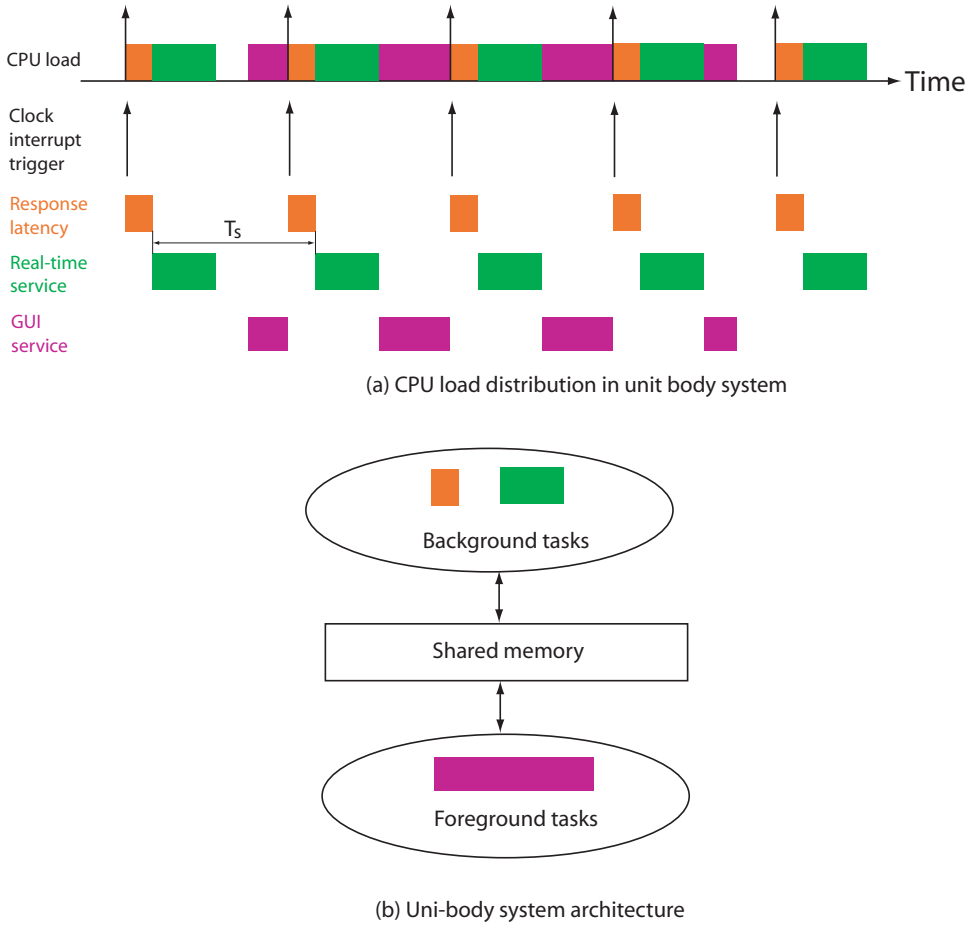


Figure 6-4: Uni-body architecture implementation of tasks.

When Windows operating systems became widely used, the uni-body architecture met some serious problems. The interrupt latency was increased with significant jitter introduced by the Windows operating system. The same RTS under Windows

requires much slower sampling frequency than that under DOS. However, Windows make the GUI design much easier. As a result, there emerged a dual-body architecture implementation of real-time computers: the host-target configuration. This dual-body architecture combines advantages of both low latency on DOS-like machines and a nice-looking GUI on a Windows-based PC, as shown in Figure 6-5. In this architecture, a target computer is dedicated for RTS, and a host computer is used for GUI implementation. On the target computer, foreground-background architecture is still used, but the foreground is only a specially tailored real-time kernel (instead of a complete operation system) only for host-target communication. On the host computer, the GUI is running under Windows. The physical connection between the host and the target can be any media such as ISA, PCI, RS-232, Ethernet, USB, and so on. Correspondingly, the GUI program has communication routines talking with the kernel service on the target through the physical media connection.

So far, almost all the existing commercial real-time computer systems fall into one of these two architectures. Another important feature of today's commercial control-computer systems is to support Real-Time Workshop of Mathworks or Autocode of MATRIXx, and thus a control algorithm created using graphical diagram languages such as Simulink or SystemBuild can be automatically converted into C codes and further be compiled into machine language. Next, we will review some current leading-edge commercial products.

Real-Time Windows Target

Real-Time Windows Target is a solution provided by Mathworks[1]. The code generated from Simulink runs on the same machine where GUI is running. Obviously, this is a uni-body architecture under Windows. It is reported in [47] that a 10 kHz sampling rate is achieved for most models.

[1] www.mathworks.com

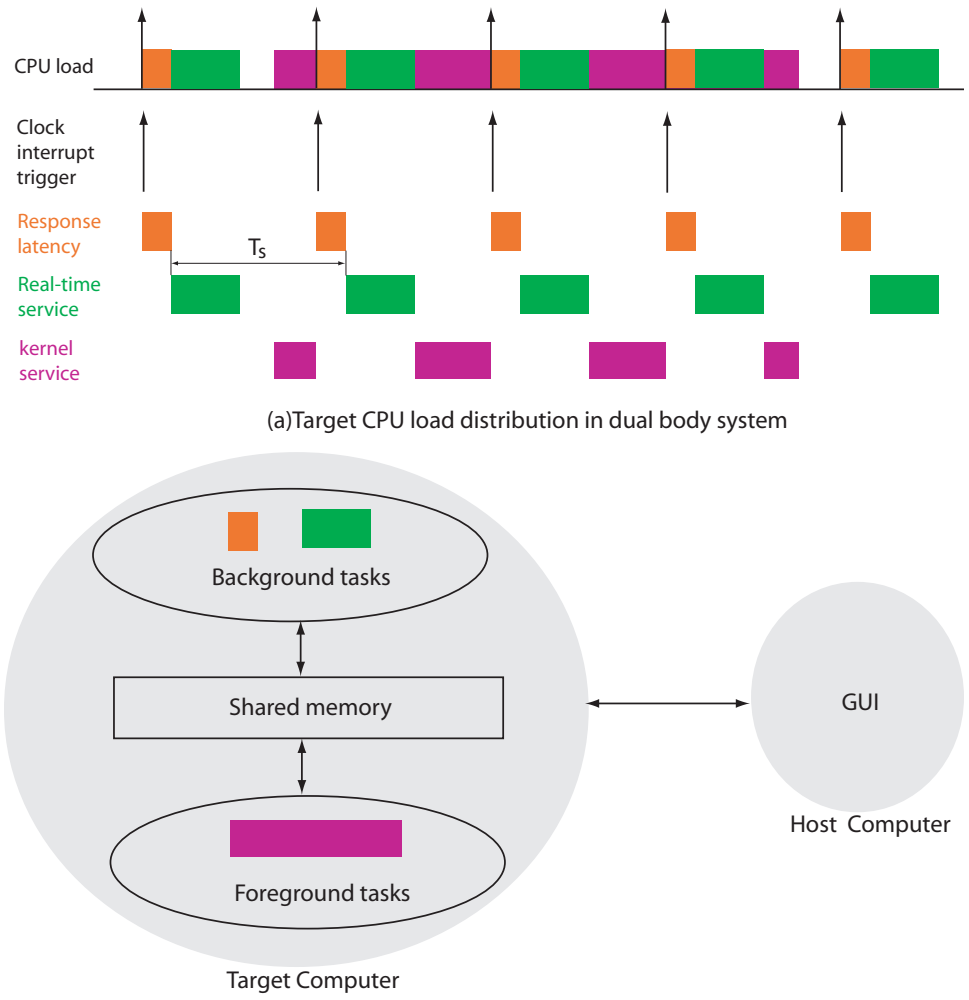


Figure 6-5: Dual-body architecture implementation of real time computers.

Quanser

Based on Ardence's RTX[2], Quanser[3] developed a real-time computer solution, which can achieve a mean latency of 4.5 μs (the worst case is 12.21 μs) and a maximum jitter of 7.7 μs on a high-performance Pentium 4 PC [2]. Although this solution is an uni-body architecture, it is better than other uni-body systems. As a real-time extension to Windows, RTX is a kernel at a level lower than the Windows operating system, and thus is not controlled by Windows. Actually, the RTX is booted up before Windows. As a result, RTX takes over the authority of the timer interrupt,

[2]Previously it was VenturCom's RTX.
[3]www.quanser.com

and make it invisible to Windows. Hence, the problem of jitter and interrupt latency caused by the contention of Windows is alleviated. Softsevo[4] products are also based on Ardence's RTX for industrial control of machine tools using one PC.

dSPACE

One of the most successful dual-body architecture real-time computers is the real-time control solution developed by dSPACE[5]. The target is based on either a PowerPC or an AMD Opteron processor, and the GUI software called Controldesk runs on a host PC under Windows. The host-target communication is a PCI bus (for DS1104), or an ISA bus (for DS1103 and other modular products). Based on my experience on a Year 2002 Version DS1103, a single-task RTS can run at 100 kHz. When multiple interrupts are used, the DS1103 can support two services simultaneously: a fast tool servo control routine runs at 83 kHz and a diamond turning machine control routine runs at 4 kHz. Both routines have fair amount of complexity including adaptive feed-forward cancellation, stateflow, and trajectory generation and so on.

xPC Target

Another dual-body architecture real-time computer called xPC Target [6] receives increasing popularity. Both the target and the host are PCs. On the target PC runs a specially tailored real-time kernel, and the RTS is executed in a timer interrupt. On the host PC, a GUI is developed and runs on Windows. According to xPC Target 2.6 data sheet [48], the maximum achievable sampling rate is 100 kHz. Actually, the xPC Target software does not allow the sampling time to be set at less than 8 μs. Table 6.1 shows the test results that I collected on various machine platforms. These results include the minimum achievable sampling time for a model containing only a gain element without any I/O access, the interrupt context-switching time, and the time required for a data log operation.

[4] www.softservo.com
[5] www.dspaceinc.com
[6] xPC Target is a trade mark of Mathworks. www.mathworks.com

Table 6.1: Test results of xPC Target, xPC Target V1.2 and Matlab V6.1.

Target PC	Minimum sampling [μs]	Interrupt latency [μs]	Data log [μs]
Pentium II, 400MHz	13	4.25	5.25
Pentium III, 933MHz	8	1.7	3.0
Pentium 4, 1.7GHz	10	2.9	3.5
AMD Athlon, 1.8GHz	11	4.22	4.77

RT-Lab

Opal-RT[7] has a product called RT-Lab, which is a PC-based dual-body architecture real-time computer solution. The target PC runs on QNX operation system, and the RTS is a task under QNX. For RT-Lab Version 6, the system overhead and jitter is around 20 μs on a Pentium 3 machine of 600 MHz [64]. As a result, a reasonable sampling time is about 30 kHz on this machine.

One main problem on all the computer systems described above is the interrupt-latency. To reduce this latency, xPC provide an operation mode called polling mode. In this mode, the target PC polls the system timer to get the starting point first, and then executes RTS in a round-bobbin fashion. As a result, all the system overhead is slashed and the interrupt latency is avoided. However, the target PC loses the connection to the host machine, because there is no time reserved for host-target communication. Hence, all the CPU resource of the target PC is dedicated to the RTS, and on the host machine the user cannot monitor and control the target any more. This vital shortcoming is not tolerated for most control application. Similarly, the RT-Lab also allows the target PC to be configured in XHP (eXtrem High Performance) mode. The underlying idea is the same to the polling mode of xPC Target, and thus has the same problem.

In order to improve the computation power, both dSPACE and RT-Lab support multi-processor system, which include one host machine and more than one target machines. Each target machine in a multi-processor system works almost the same ways as a target in a dual body system, except that inter-processor communication is also included in the RTS of each target. As a result, the interrupt-latency problem

[7] www.opal-rt.com

still exist in these multi-processor systems. A multiple processor system handle more computation load than a conventional dual-body architecture system with the same sampling time, but the minimum sampling time will not be reduced. In this sense, these multi-processor systems are still categorized into dual-body architecture.

Next, we will present a new architecture, which can eliminate the interrupt overhead without losing the host-target communication.

6.2.2 ThunderStorm Architecture

Here we describe a triple-body architecture, which I developed for this project: Target-Servant-Host, as shown in Figure 6-6. On the basis of the host-target con-

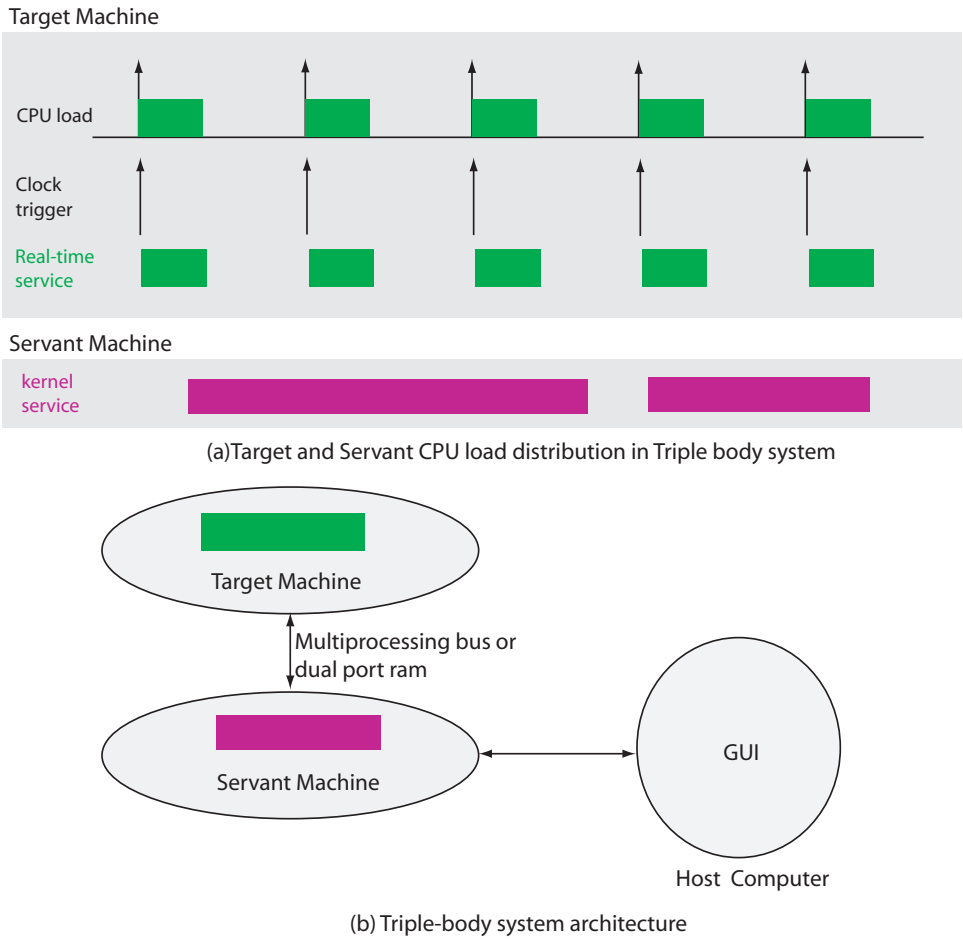


Figure 6-6: Triple-body architecture of real-time computers.

figuration, we relocate the real-time kernel that is dedicated for host-target commu-

nication from the target machine to a third machine, which is named as Servant. The servant-host communication is the same as the host-target communication in the dual-body architecture. The RTS is running on the target machine in polling mode in order to eliminate interrupt latency. To realize the servant-target communication without any load on the target machine, processors with multiprocessing capability are selected for both the target and the servant processor, so that the target and the servant are connected to the same multiprocessing bus. In comparison with the host-target configuration in dual-body architecture, the triple-body architecture realizes the parallel processing of the RTS and the kernel service of communication handling, and thus eliminates the task-switching associated latency/overhead.

To implement this triple-body architecture, I selected TigerSharc digital signal processors[8] (DSPs) for both the target and the servant. In order to further enhance the target processing capability, we use two TigerSharc DSPs as Targets, and one TigerSharc DSP as the servant. The block diagram of this multiple-target triple-body architecture is shown in Figure 6-7, which I have named the ThunderStorm architecture.

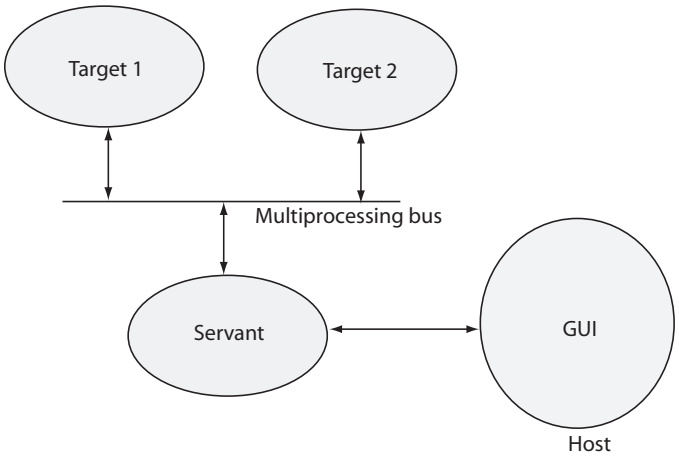


Figure 6-7: Triple-body system implementation of a real-time computer.

One more feature in the ThunderStorm architecture is to allow the parallel processing of both A/D conversion and control computation. The A/D conversion is performed in A/D convertors and the control computation is executed in the target

[8]TigerSharc is a trade mark from Analog Device Inc.

processors. Hence, they can work in parallel. Figure 6-8 shows a comparison between the RTS load distribution in a conventional computer system and that in the ThunderStorm architecture.

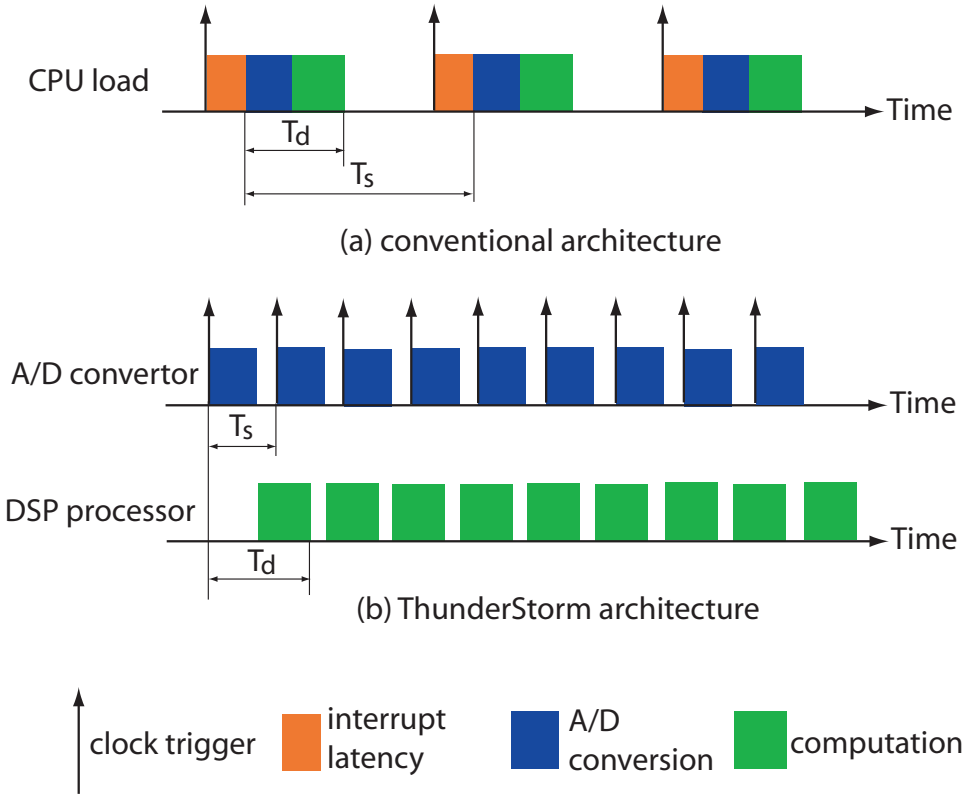


Figure 6-8: Parallel operation in ThunderStorm Architecture

Figure 6-8(a) represents a conventional system such as from dSPACE or the xPC Target. The A/D conversion is initiated after entering the interrupt (a latency after the clock trigger). Next, the CPU polls the A/D conversion status until the end of A/D conversion. Finally, based on the A/D conversion result the control algorithm is calculated on the CPU. We can see that the sum of these three time blocks sets a lower bond on the sampling time.

Figure 6-8(b) shows the ThunderStorm architecture. Not only is the latency eliminated, but also the A/D conversion is not part of the processor execution time. The A/D conversion is automatically initiated by the clock trigger, and the CPU only needs to poll the end of the A/D conversion as the starting point of control calculation. As a result, only the greater value of the A/D conversion time and the

control computation time sets the lower bond of the sampling time. Although the time delays T_d in both systems (a) and (b) are the same, the sampling frequency in ThunderStorm can be greatly increased. This is an important advantages for real-time control.

In the rest of this chapter, we will present the detailed design of a computer system implementing this ThunderStorm architecture.

6.3 Functional Block Diagram

The block diagram design of the Thunderstorm real-time computer is shown in Figure 6-9. DSP A, DSP B, and DSP C are symmetrically interconnected via a 100-MHz

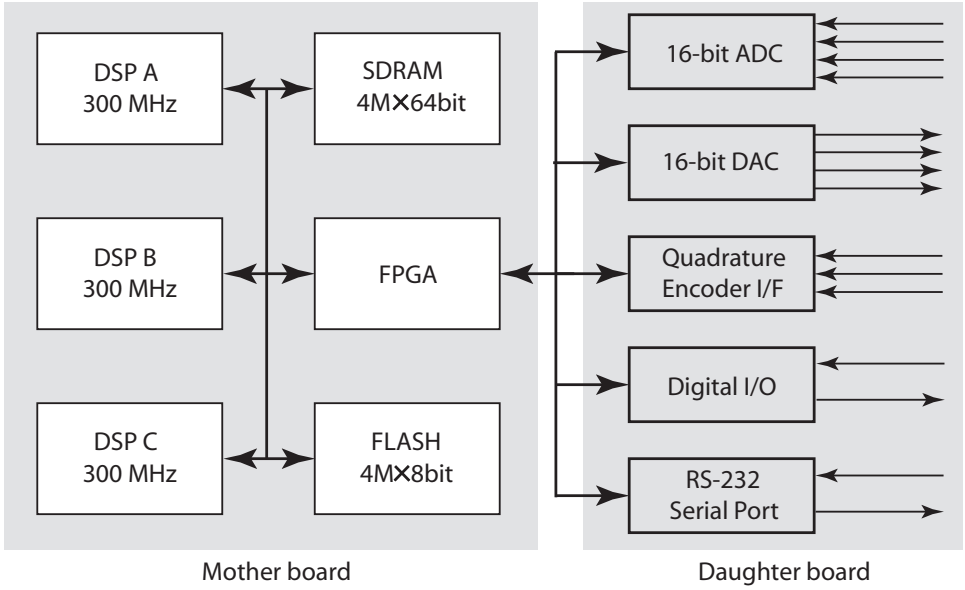


Figure 6-9: Parallel operation in ThunderStorm Architecture

64-bit multiprocessing bus, which is referred as a cluster bus. The model number of this processors is ADSP-TS101. On this bus, there are a 4M×64 bit synchronous dynamic RAM (SDRAM), a 4MB (bytes) non-volatile NOR FLASH memory, and a field programmable gate array (FPGA). The SDRAM is used to store logged data. The flash memory can be used to store an application program so that the computer can execute this program after power-up. All the I/O resources are connected to the cluster bus via the FPGA. The I/O resources include 4 channels of 16-bit SAR A/D

convertors with 0.4 μs conversion time, 4 channels of 16-bit D/A converters capable of 30 MHz maximum sampling frequency, digital I/Os, 3 quadrature encoder interfaces (I/F), and a serial port.

The serial port is a RS-232 UART (universal asynchronous receiver and transmitter) used for host-servant communication. Actually, the UART engine is implemented inside the FPGA, and thus the serial transmit and receive signals directly connects to the FPGA. Similarly, the quadrature counters are also implemented in the FPGA, and the quadrature input signals are connected directly to the FPGA. The advantage of having an FPGA inside the system is that the hardware resources can be reconfigured easily. All the digital inputs/outputs (including UART, digital in/out, and quadrature inputs) are optical isolated from the analog part of the computer so as to reduce noise problems. The details of the UART and the quadrature encoders will be discussed in next sections.

6.4 2-D Quadrature Encoder Counter

Quadrature encoders are widely used as motion feedback sensors. Output signals of a quadrature encoder is two square waves with 90^o spatial phase difference. In our diamond turning machine, quadrature encoder outputs are used in both spindle rotation and X/Z stage position measurement. [9] The working principle of a conventional quadrature encoder counter is illustrated by signals A, B and S in Figure 6-10. Signals A and B (in blue color) are quadrature outputs from a quadrature encoder. A spatial counter S in 4$\times$ mode increases or decreases the counter register value by 1 whenever there is change in A or B. Otherwise, the counter register holds its value. Inherently, a quadrature encoder is a spatial-domain sampler, which updates the measured position of a motion axis at constant spatial intervals.

In order to measure the momentary speed from two consecutive spatial samples

[9] Reflective laser scales are mounted on the machine base, and laser read heads are on the machine stages. Outputs from a laser scale read head are two sinusoidal waves with 90 degree spatial phase difference. An signal-processing card first interpolates these analog sinusoidal signals by a factor of 16, and then convert the interpolation result into quadrature encoder signals as output.

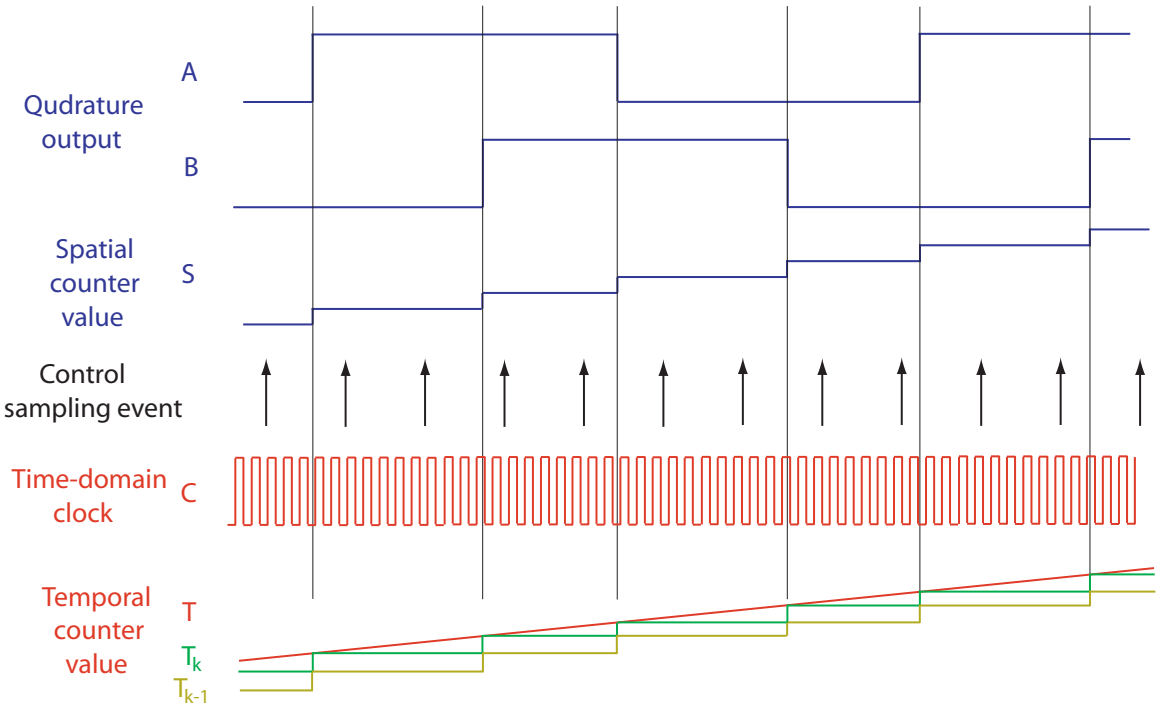


Figure 6-10: Signals in the spatial-temporal quadrature encoder counter. Spatial signals are shown in blue, and temporal signals in red.

and thus allow for high-frequency spindle/encoder dynamic performance analysis, we designed a temporal-spatial two dimensional counter. Besides the spatial counter S to count spatial reference signal A and B, we introduce a temporal reference signal C, which has much higher temporal spatial frequency than signals A and B. As shown in Figure 6-10, a temporal counter T is incorporated to count a fixed frequency clock signal C (labelled in red color) so as to generate a time reference. At every moment that the spatial filter change its value (equivalently the signal A or B flips), the temporal counter value T is latched as T_k. From the current latched timer value T_k and its previous value T_{k-1}, the momentary speed during the latest two adjacent spatial counts can be calculated as

$$\begin{aligned} n &= \frac{S_k - S_{k-1}}{T_k - T_{k-1}} \\ &= \frac{1}{T_k - T_{k-1}} [\frac{\text{spatial count}}{\text{timer period}}] \end{aligned} \tag{6.10}$$

This 2-D quadrature encoder counter is implemented in the FPGA using a 100 MHz

clock as signal C. During momentary spindle speed analysis, we logged both the spindle spatial positions and the latched temporal counter value at each spatial sample. In Chapter 8, we introduce the result of applying this 2-D counter to measure the spindle momentary speed on our diamond turning machine.

For a motion axis at near a constant speed such as a spindle, another benefit from this 2-D encoder interface is that an interpolated position $\hat{S}$ with resolution higher than the encoder can be derived from the temporal reference signal T as

$$\hat{S} = \frac{T - T_k}{T_k - T_{k-1}} + S, \tag{6.11}$$

where T and S are the current temporal and spatial counter values at the time a control sampling event, T_{k-1} and T_k are the latched temporal counter values at the last two spatial sampling event of the encoder.

6.5 Serial Interface

The serial interface includes both transmitting and receiving components.

As shown in Figure 6-11, when transmitting a 64-bit long word out of the computer, a DSP only need to write the data to the UART-specific address, and the data will be latched into an outgoing register inside the FPGA. At the same time, a transmission-ready flag inside the FPGA will be cleared automatically, indicating that the transmitting interface is busy with current data and is not ready to send the next data. The unpacking engine then converts the latched 64-bit data into eight 8-bit bytes, and consequently each 8-bit byte is serially sent out as a bit stream according to standard RS-232 protocol in the transmitter. When all eight bytes are sent out, the transmission-ready flag will be set to indicate that the serial interface is ready to transmit another 64-bit long word.

In the receiving interface, after eight continuous bytes of serial data are received according to RS-232 protocol, they are compacted into a 64-bit long word in the packing engine. Further, this long word is moved into a receiving register, and cor-

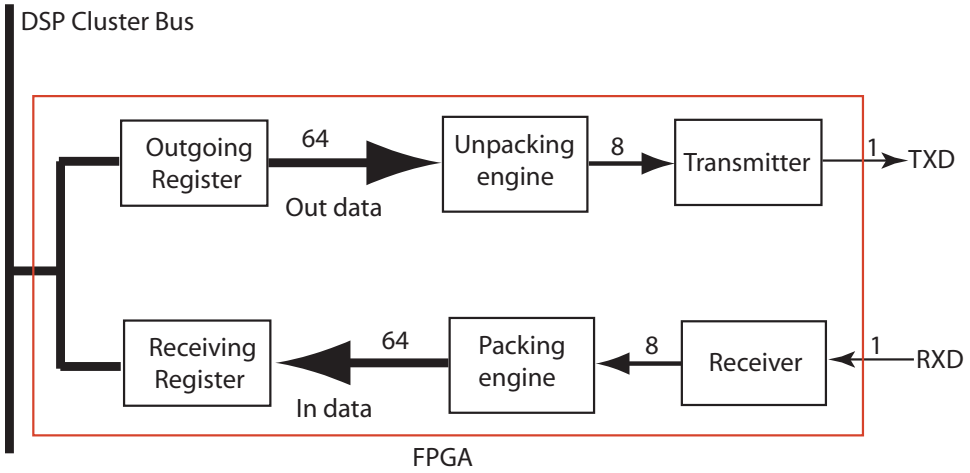


Figure 6-11: Simplified block diagram of the serial interface .

respondingly a receiver-full flag will change from 0 to 1, indicating to DSPs that the data is ready to be read. Once a DSP reads the receiving register, the receiver-full flag will be cleared automatically.

Both the packing/unpacking engine and transmitter/receiver are programmed in VHDL by using finite-state machines and implemented in the FPGA. The maximum supported data rate is a 1 MHz baud rate, much higher than the 115.2 kHz supported by a PC serial port.

With this serial interface, the DSPs see only a 64-bit UART on the cluster bus, and the bus load is significantly reduced when compared to a conventional 8-bit UART. For each transmitting/receiving of a 64-bit, only a write/read transaction is required on a DSP, with all the rest operation completed in the FPGA. As a result, this design allows DSP/FPGA parallel operation and thus minimizes the DSP involvement with the serial communication.

6.6 Motherboard Design

Figure 6-12 shows the mother board design of the ThunderStorm computer, which I designed. There are 12 layers in the print circuit board made of FR-4 [10], the stack-up

[10]FR-4 is the abbreviation of flame resistant 4. It is an epoxy material reinforced with a woven fiber glass material. Its dielectric constant (relative permittivity) is 4.6 at 1 MHz, and its dielectric strength is 500 V/mil.

design of which is shown in Figure 6-13. There are eight signal layers of 0.5 oz copper[11] and four power layers of 1 oz copper. The signal layers are mainly used for signal routing, and the power layers are used for ground plane, +3.3V, +1.8V, and +1.2V power planes respectively. The dielectric thickness between each two adjacent layers is designed to ensure 50 Ω characteristic impedance on all signal tracks to control signal edge integrity. The track width of controlled-impedance signals on layer 3, 4, 6, 7, 9, and 10 is designed to be 5 mil, and that of controlled-impedance signals on layer 1 and 12 is designed to be 9 mil, because layer 1 and 12 have one side directly exposed to air. In this design, all digital signals are single-ended. The stack-up design is based on FEA electromagnetic solver from Polar Instrument[12]. The PCB board layout was designed with guidelines on signal routing in [29].

Figure 6-14 shows the microscope image of a cross section of a coupon with stack-up design identical to the mother board. The PCB board manufacture did a TDR (time-domain-reflectometry)[13] test on the coupon, indicating that the impedance of all controlled-impedance tracks is within 10% of the required 50 Ω.

6.7 Daughter Board Design

Figure 6-15 shows the designed daughter board, which handles many I/O functions. The mother board and the daughter board are assembled together back-to-back via four 140-pin connectors. Figure 6-16 shows that two boards are assembled together.

Each analog input channel has a pair of differential signals $\pm A_{IN}$ with a differential input range of ±10 V. Figure 6-17 shows the simplified circuit of an analog input. The differential inputs to the A/D convertor require a input range of $0 \sim 4.096$ V and a common mode voltage of 2.048 V. A full-differential amplifier is used to convert the differential analog inputs $\pm A_{IN}$ to the required differential inputs to the A/D

[11] 0.5 oz copper means an electroplated copper foil with a thickness of 0.5 oz per square feet. Similarly, the thickness of 1 oz copper is 1.4 mil or 35.5 μm.

[12] www.polarinstrument.com.

[13] Time domain Reflectometry is a widely used method to measure the impedance characteristic of cables and PCB board traces. The working principle is like a radar: an energy pulse is sent out and the reflected pulsed is recorded to derive the impedance characteristics of the transmission path.

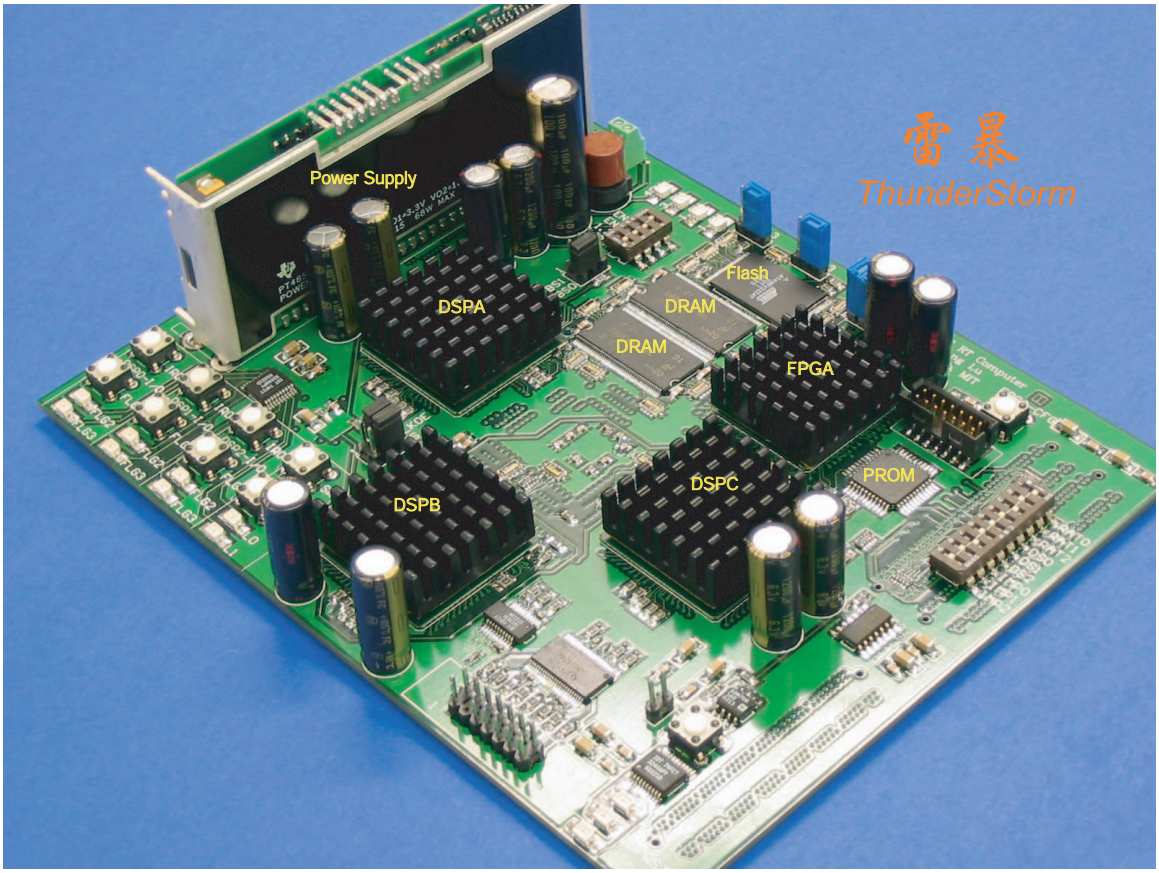


Figure 6-12: Picture of the ThunderStorm computer mother board. The dimension is 160×125 mm^2.

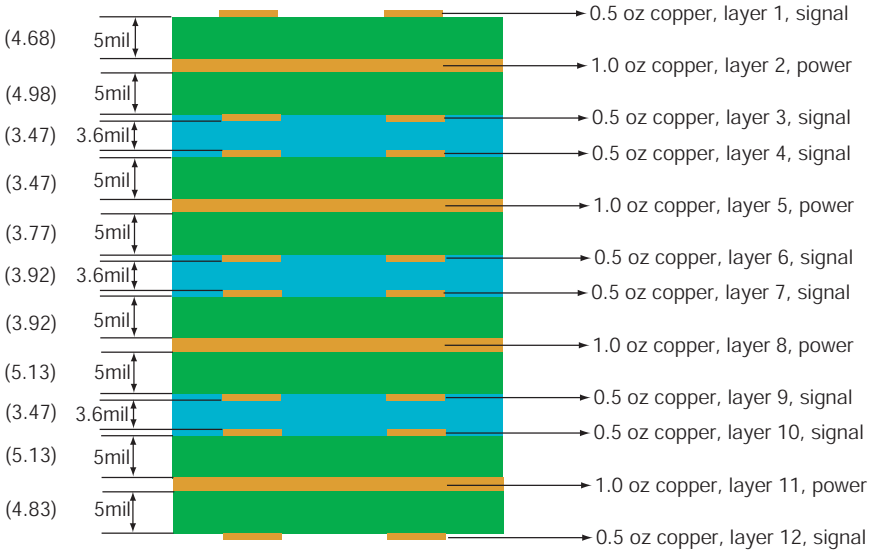


Figure 6-13: Stack-up design of the mother board. The PCB board manufacture www.2justforyou.com measured the layer thickness of the actual finished board, and the results are shown inside brackets in unit of mils.

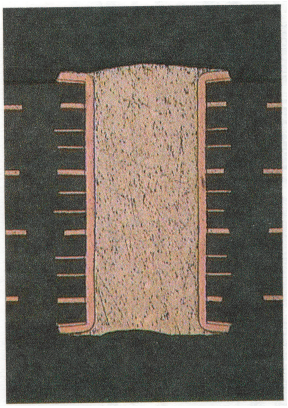


Figure 6-14: Cross section of a via in the sample PCB of the designed stack-up.

convertor. In order not to exceed the common mode input voltage range of ± 3.4 V on the differential amplifier, the common mode voltage of $\pm A_{IN}$ must be limited to less than 6.9 V.

Each analog output channel has also a pair of differential signals with zero common mode voltage with respect to board analog common and a differential output range of ±20 V.

6.8 GUI Design

Figure 6-18 shows the fast tool servo control graphical user interface (GUI) using Matlab GUIDE tool. Using this GUI, all the variables inside the DSPs can be modified, monitored, and logged.

6.9 Test Results

Figure 6-19 shows an experimental test result of the total latency of the Thunderstorm computer executing a sample program. This sample program includes four simultaneous A/D conversions, a 20th-order filter, a sine computation, a cosine computation, a square root computation, and finally four simultaneous D/A conversions. The value sent to the channel 1 D/A convertor is the channel 1 A/D conversion result. The whole program is executed at a rate of 1 million times per second. A square

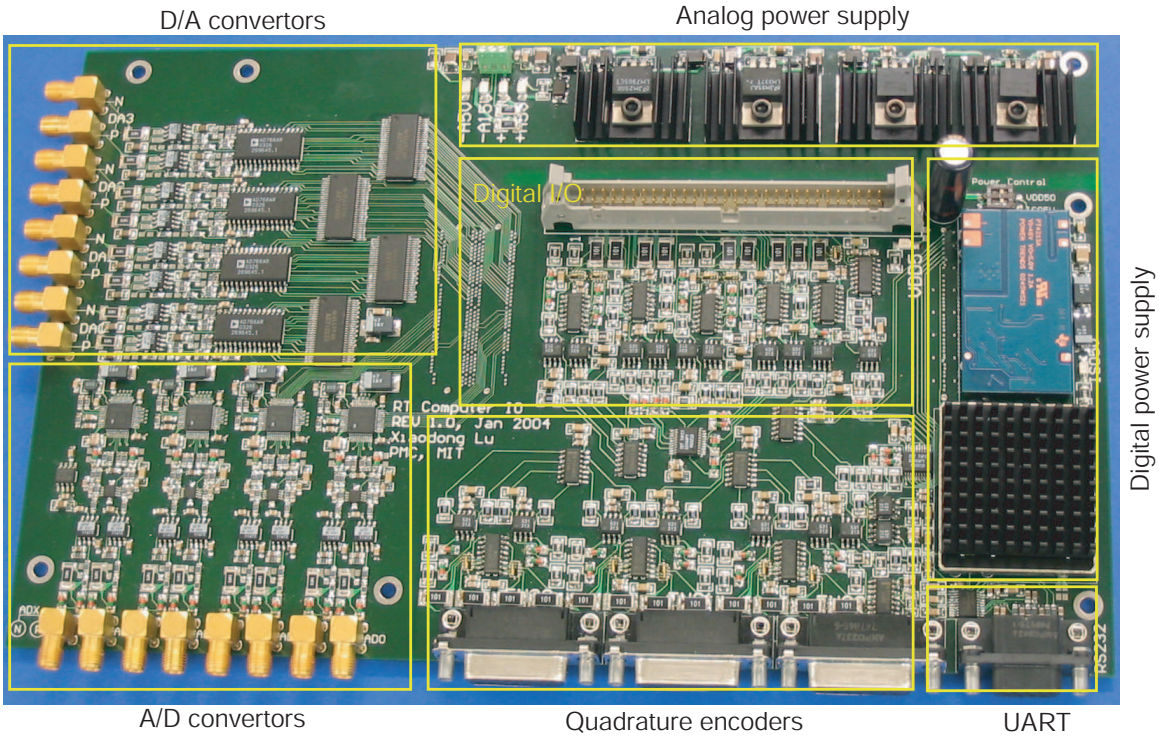


Figure 6-15: Picture of the ThunderStorm computer daughter board. The dimension is 270×165 mm^2.

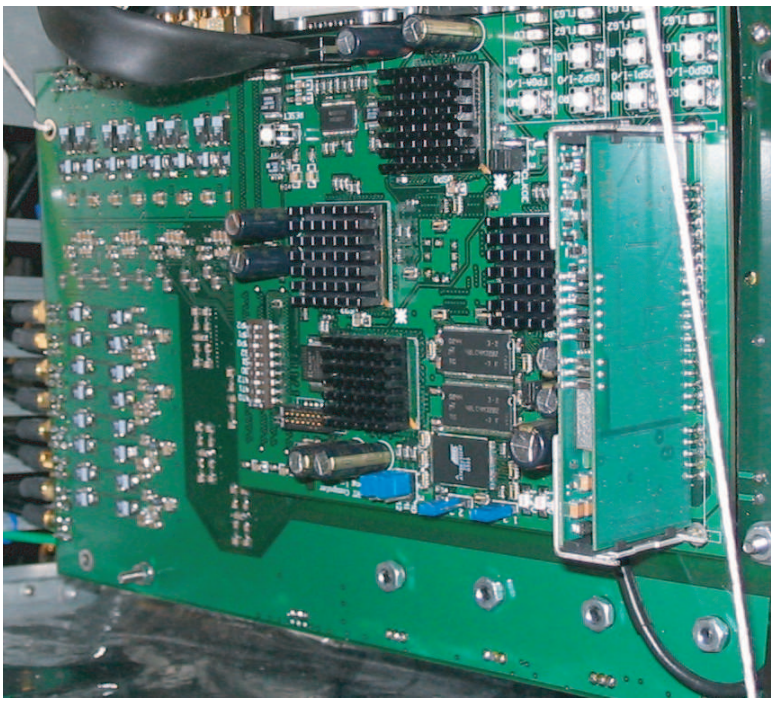


Figure 6-16: Daughter board assembled with the mother board.

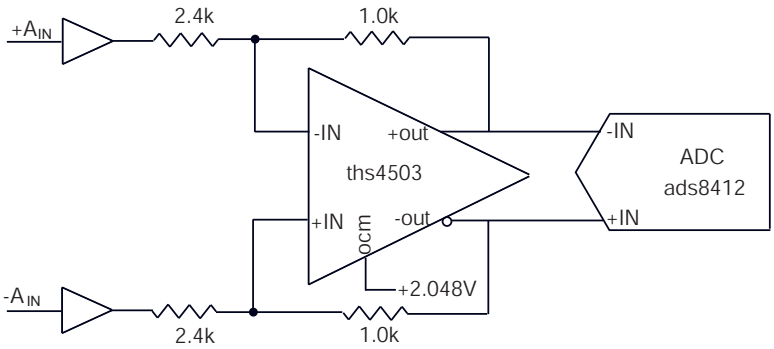


Figure 6-17: Simplified circuit for analog inputs.

wave from a signal generator is connected to the channel 1 analog input in the test. As shown in the figure, the output of channel 1 D/A relative to the signal generator signal has a variable delay ranging from 1.4 μs to 2.4 μs, because of the asynchronism between the signal generator and the sampling events of the real-time computer. The minimum delay reflects the actual delay T_d=1.4 μs, and the total variation reflects the sampling time T_s=1.0 μs. As a result, the total delay element which will be seen in a control loop in this real time computer is $T_d + \frac{T_s}{2}$=1.9 μs.

Figure 6-20 shows the A/D conversion result when the analog input is grounded by a 50 Ω resistor. This result indicates that the RMS noise of the A/D channel floor is 0.7 least significant bit (LSB) of the A/D convertor. Its spectrum shows that the ground floor is nearly a white noise. This performance is significantly better than most commercially available A/D boards at this sampling rate.

Figure 6-21 shows the total nonlinearity performance of the A/D convertor and D/A convertor. The system is in a direct loop-back setup by connecting the D/A output directly to the A/D input with a BNC cable. A sine wave of 50 kHz was computed inside the digital signal processor, then this signal is sent to the D/A convertor, finally the A/D conversion result is logged and uploaded to the host computer. Therefore, the resulting data contains nonlinearities of both the D/A convertor and the A/D convertor. The total harmonic distortion (THD, the root mean square of all the super harmonic components divided by the amplitude of the base harmonic component) is -88 dB. In practice, the nonlinearity of the analog input is more important than the analog output. To test the nonlinearity of the A/D convertor alone, a very

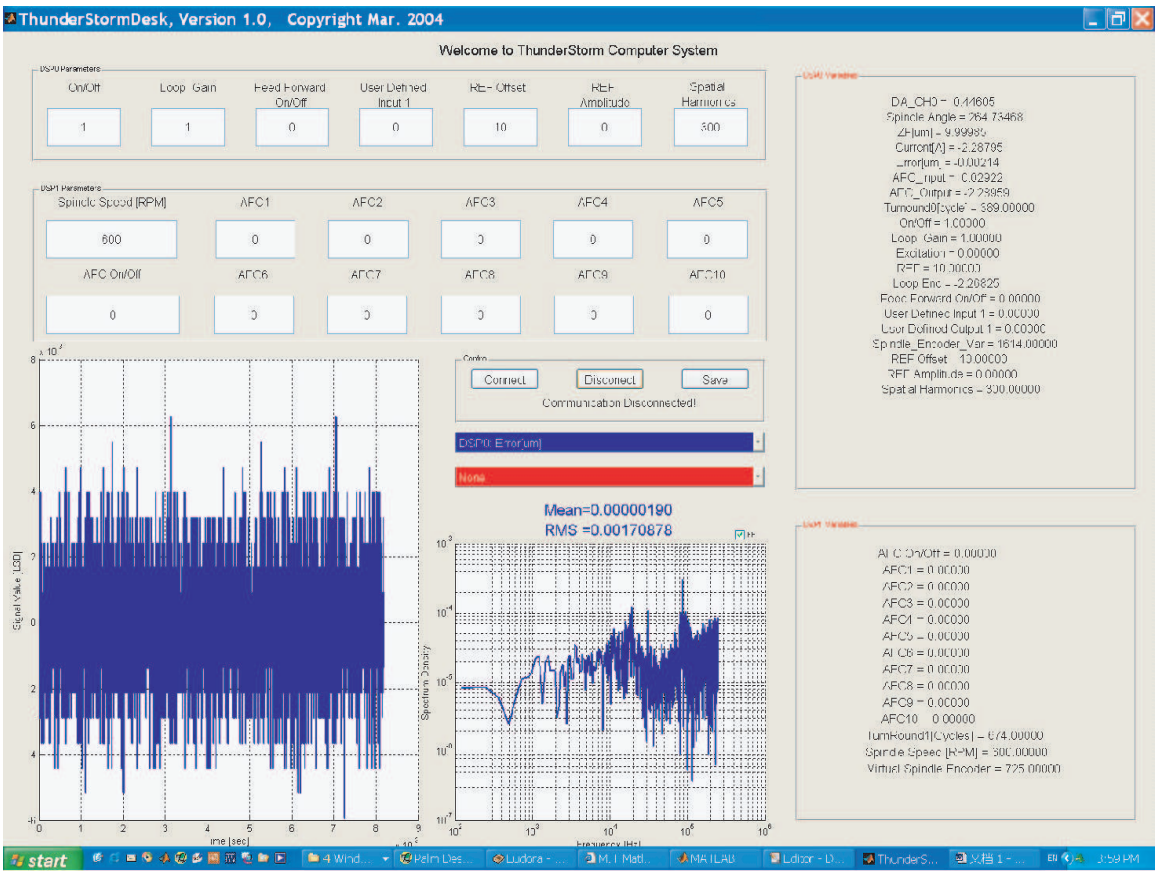


Figure 6-18: Graphical user interface design.

high quality sinusoidal signal should be used. As we do not have a signal generator with better quality than our D/A analog outputs, this test has not been performed yet. It is expected that the A/D convertor has a THD of near -90 dB.

6.10 Summary

The achievable bandwidth of digital control systems are limited by the total delay caused by the A/D conversion, computation, and sampling time. For the ultra-fast tool servo control application, we designed a real-time computer capable of 1 MHz sampling rate.

After examining existing control computer architectures, we designed a thunderstorm architecture, incorporating a servant machine between the target and host. As a result, the target machine can work in fully polling mode and the interrupt

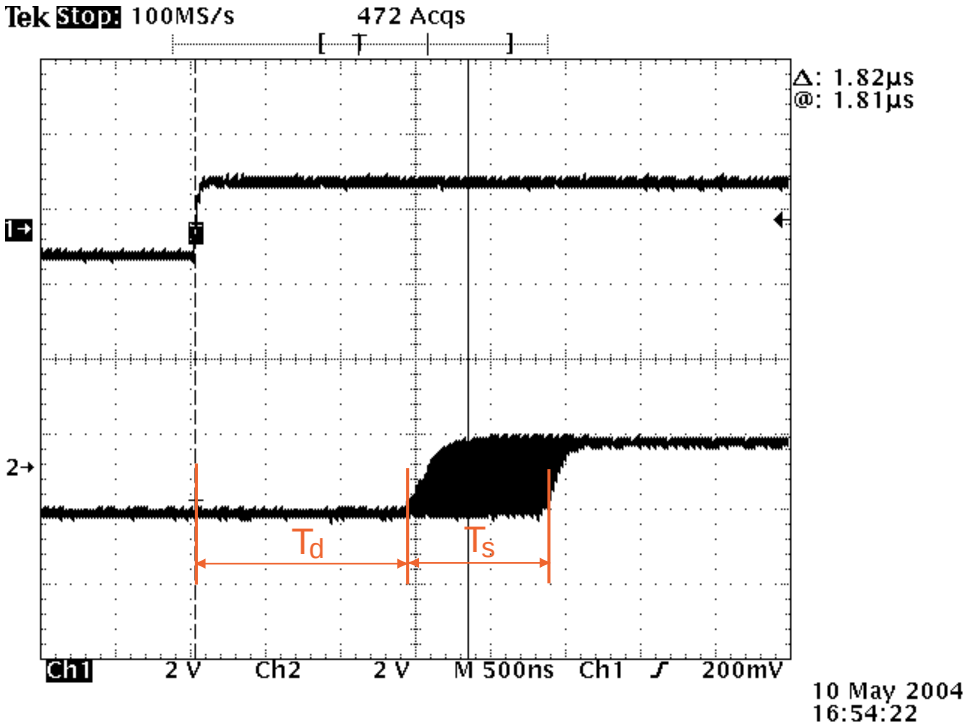


Figure 6-19: Total latency test. The top waveform is the input to the channel 1 analog input. The bottom waveform is the output from the channel 1 analog output.

associated latency can be eliminated.

The designed real-time computer are based on three floating-point digital signal processors and a FPGA for I/O interface. The mother board is a 12-layer PCB with controlled impedance on all high-frequency signal traces. Inside the FPGA, a 64-bit RS232 transmiter/receiver and quadrature encoder counters are implemented. The quadrature encoder interfaces have two-dimensional counters, not only counting the spatial incremental but also latching a 100 MHz timer value at the spatial sampling point. Therefore momentary speed of the motion axis can be calculated for every spatial sample interval. Both analog inputs and outputs are in differential pairs, with 16-bit resolution and over 2 MHz sampling rate capability. The digital I/Os, serial communication ports, and encoder interfaces are optically isolated from the rest of the computer.

Test results demonstrate that the total latency is 1.8 microseconds at 1 MHz sampling rate. The floor noise of the analog inputs is 0.7 LSB, and the total harmonic distortion is -88 dB.

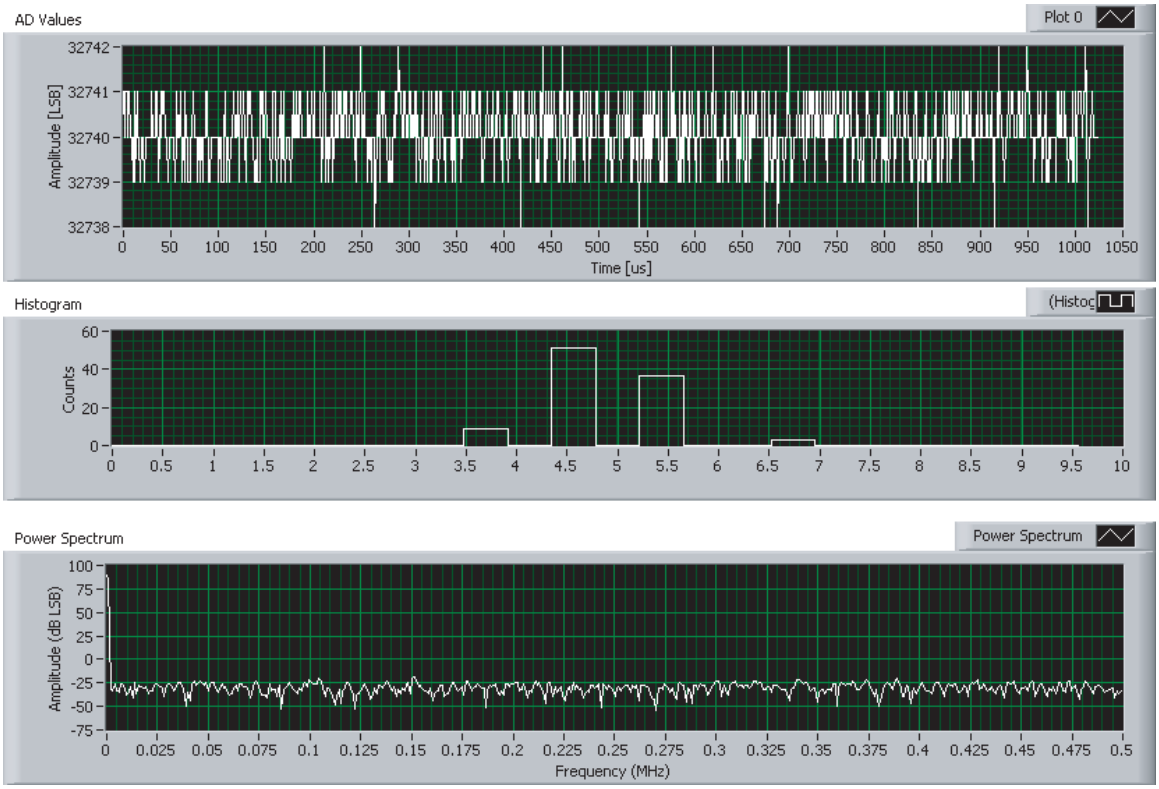


Figure 6-20: Analog input noise floor test, with input connected to analog common with a 50 Ohm resistor. The top graph is the raw sampled data by the A/D convertor at 1 MSPS, and the unit is LSB of the A/D convertor. The middle graph shows the statistics of the A/D results. The bottom graph is the power spectrum (FFT) of the raw data using a Hanning window.

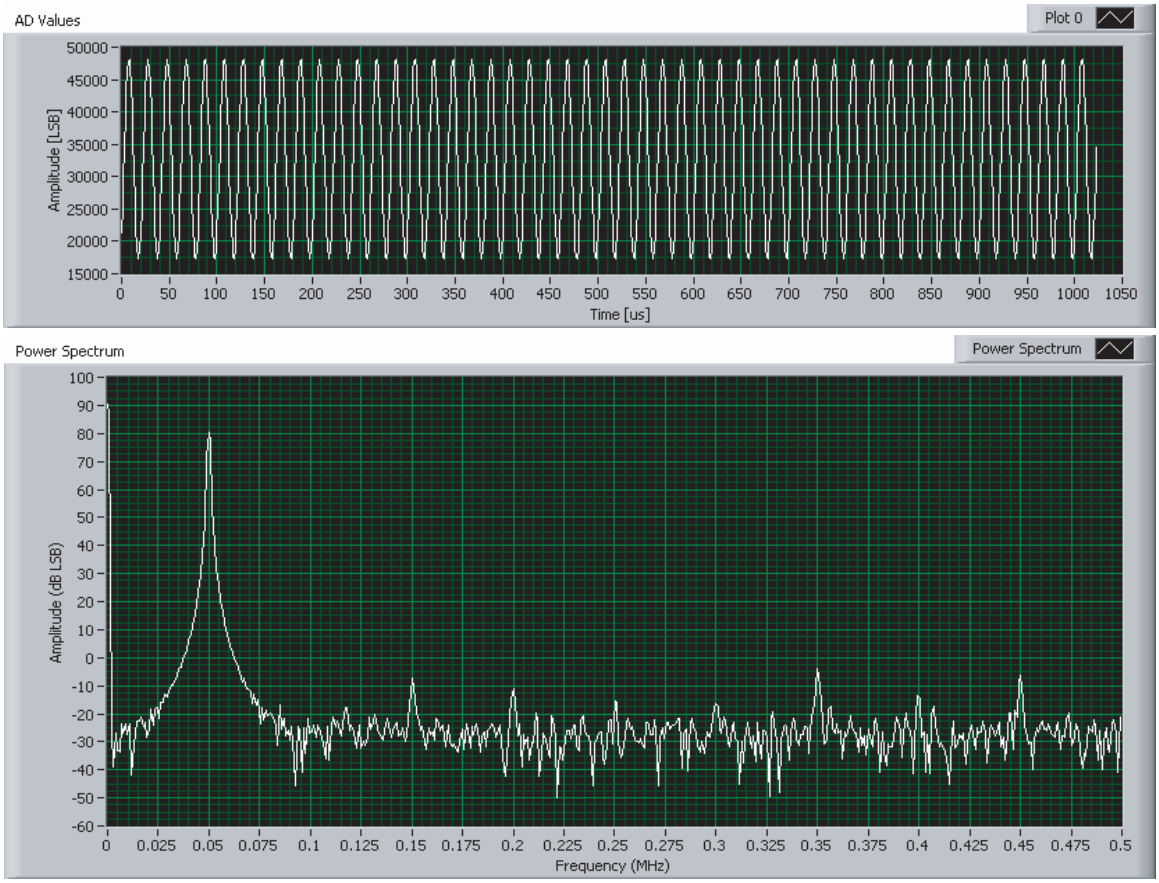


Figure 6-21: Total harmonic distortion test, using a 50 kHz sine wave generated by the D/A output. The upper graph is the A/D sampling result in unit LSB of the convertor. The lower graph is the spectrum of the A/D sampling result.

Next Chapter, we will use this real-time computer to control our designed ultra-fast tool servos.

Chapter 7

Fast Tool Servo Control

This chapter presents control of the SM2-based UFTS. First, we discuss the plant identification in Section 7.1. Based on the experimentally measured frequency response, we compensate the loop transmission to achieve a 10 kHz crossover frequency in Section 7.2. To further enhance the repetitive trajectory tracking and disturbance rejection, we introduce the adaptive feed-forward controller design in Section 7.3. At the end, trajectory generation issues are discussed.

7.1 Plant Identification

We experimentally measured the frequency response using swept sine waves at both small signal and large signal levels.

7.1.1 Small-Signal Response

Figure 7-1 is a small-signal measurement result of the frequency response from the excitation current I [A] to the displacement output X [μm] with data acquired at a sampling frequency of 500 kHz. This result is measured while under closed-loop control, and the armature is commanded to move away from stator pole faces. At low-frequency region up to 3 kHz, the frequency response is a flat gain resulting from the stiffness of bearings. There is a resonance frequency at 3 kHz, which is caused

by the resonance of the bearing axial stiffness and the moving assembly mass. In the frequency region from 3 kHz to 45 kHz, the system exhibits a mass-dominated characteristics. Beyond 45 kHz, the measured frequency response most likely results from electrical cross-talk between the actuator coils and the capacitance probe sensor, and does not reflect the mechanical dynamics. We now discuss the response for each region in detail.

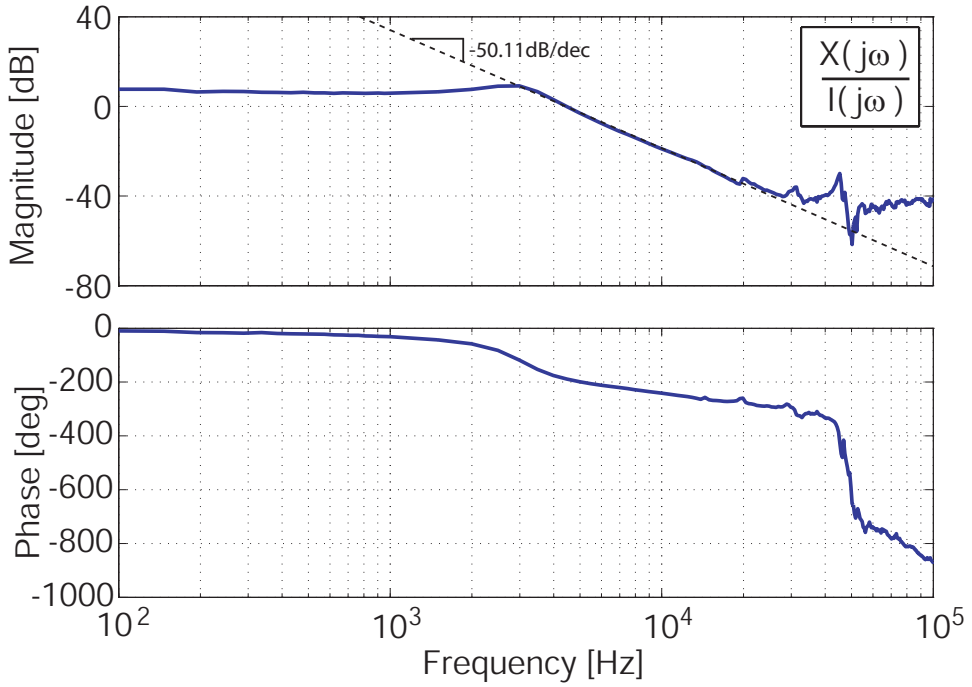


Figure 7-1: Experimentally-measured frequency response of the SM2-based UFTS, from the excitation current in Amperes to the displacement output in micrometers. The sampling frequency is 500 kHz.

The amplitude of the plant frequency response in the region from 0 to 3 kHz is about 8 dB or equivalently 2.5 μm/A. According to the total moving mass of m =11.2 gm, we can also estimate the total bearing stiffness K_{DC} from the resonance frequency ω_N at 3 kHz

$$\begin{aligned} K_{DC} &= m\omega_N^2 \\ &= 11.2 \times 10^{-3} \times (3 \times 10^3 \times 2\pi)^2 \\ &= 4\ N/\mu\text{m}. \end{aligned} \tag{7.1}$$

Further, we can derive the actuator force coefficient K_f (the gain from current to actuating force) $K_f = 4 \times 2.5 = 10$ N/A, which is consistent with earlier result from the static tests. In the static test, we first commanded the UFTS in closed-loop (including integrator) with a constant reference position. Then we used a force gauge to push against the tool holder. By recording the power supply current and the force gauge outputs, we can calculate the static force coefficient. However, in both the dynamic test and the static test, the resulting force coefficient is much smaller than the theoretical value $K_f = F_0/X_0 = 35.5$ N/A for the SM2-based UFTS calculated in Table 4.2. The reason for this discrepancy is that the permeability of the SM2 material has a maximum permeability of $500\mu_0$, and the theoretical calculation is based on infinite permeability assumption. As a result, in the SM2-based UFTS significant magnetomotive force is dropped inside the magnetic cores. In contrast, the Ni-Fe material has a minimum permeability of $10^4\mu_0$ for induction level between 0 to 1.5 T, and thus should give a higher force coefficient.

The frequency response between 3 kHz and 30 kHz can be approximated by a straight line with a slope of -50.11 dB/decade, as indicated in Figure 7-1. Because the eddy current effect inside SM2 material contributes another -10 dB/decade as discussed in Appendix B, this identified frequency response deviates from a typical mass-dominated response of -40 dB/decade.

At 45 kHz, there is a strong non-collocated mode, which surprisingly matches very well with the 12-th mode of the FEA result shown in Figure 4-15. In the frequency region between 20 kHz and 30 kHz, there also exist two small bumps, which come from the weakly-excitable modes. According to their phase response, we can tell that they are collocated modes with little sensing ability and/or actuating ability.

7.1.2 Large-Signal Response

Figure 7-2 shows measured frequency response in open-loop operation at various AC levels and a DC bias current of -2A. This DC bias current was used to move the armature away from the stator poles surfaces, because the steel flexural bearing is preloaded. In this result, we can see that the low-frequency gain changes with

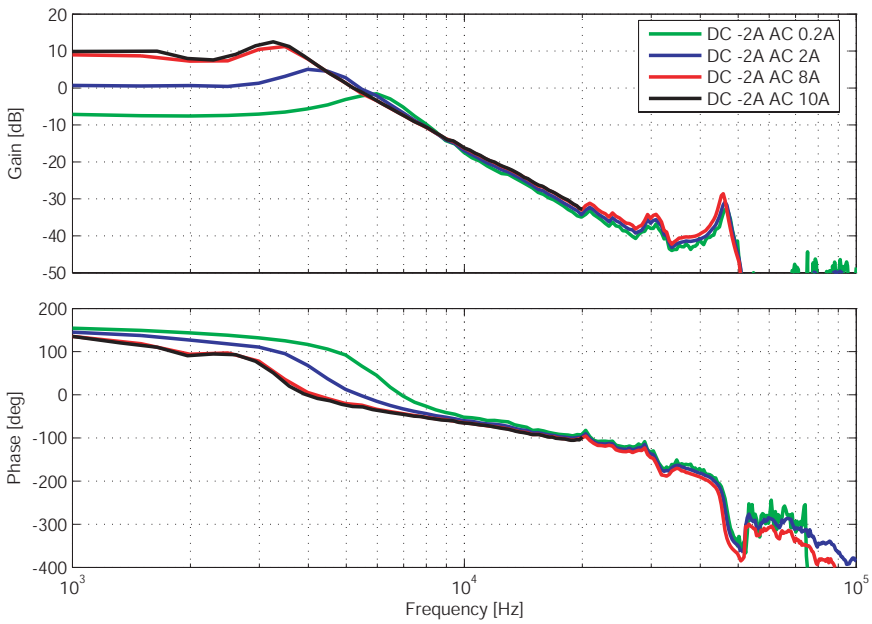


Figure 7-2: Open-loop frequency response of the UFTS at various AC level. The DC current is -2 A. The AC level in the figure is peak-to-valley value.

excitation amplitudes from -8 dB at a small signal to 10 dB at a large signal. This change mainly comes from the contact stiffness of the tape on the armature pole face. At small-amplitude excitation current, the generated force is not strong enough to move the armature completely away from the stator pole face, and thus the measured stiffness comes from the bearing stiffness in parallel with the plastic tape stiffness. At large signal excitation, the armature is moved away from the stator pole face and thus only the bearing stiffness is measured.

Figure 7-3 shows the time-domain results at 10 kHz open loop operation. The phase shift is 225 degrees between current input to the displacement output. The tool motion is 2.6 μm peak-to-valley, which corresponds to 523 G acceleration at 10 kHz. When the operation frequency is around the resonance (at 3 kHz and 45 kHz), the acceleration can be significantly amplified to over 1000 G.

7.2 Loop Compensation

In order to control this SM2-based UFTS, we design the controller structure as shown in Figure 7-4. The whole controller is composed of three parts: an integrator K_I/s

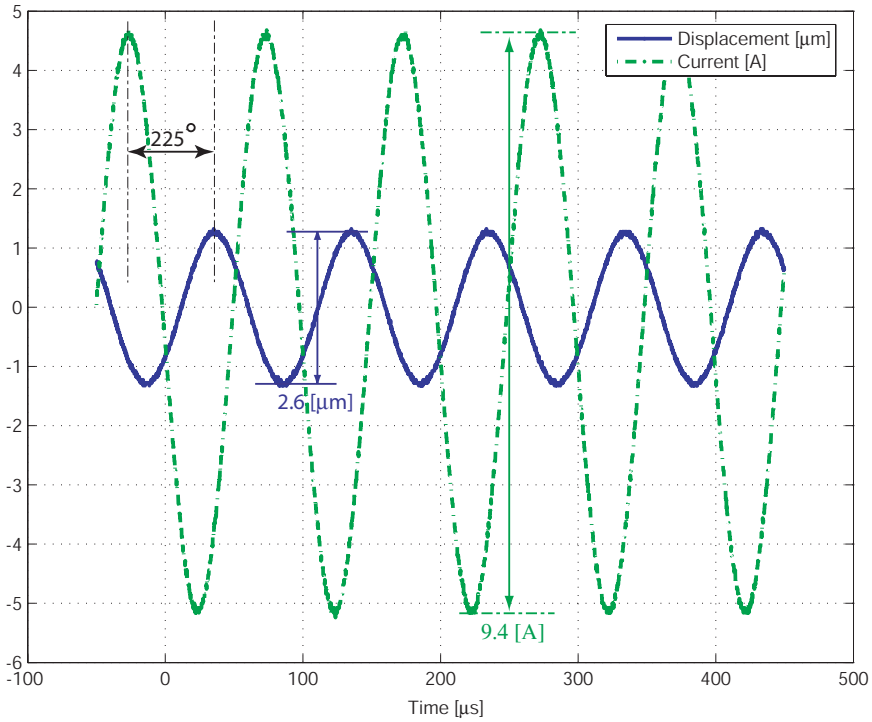


Figure 7-3: Acceleration capability test at 10 kHz operation. The moving assembly displacement is shown in blue with units of μm. The current is shown in green with units A.

implemented in DSP A, a loop-shaping controller $C_{ls}(s)$ implemented in DSP A, and adaptive feed-forward cancellation (AFC) controllers $AFC_n(s)$ implemented in DSP B. This section introduces the integrator and the loop shaping controller design; the AFC controllers design will be discussed in the next section.

The integrator K_I/s gain is designed at $K_I = 2\pi f_c/5$ to eliminate steady-state errors, where f_c in units of Hz is the intended unity-gain crossover frequency f_c of the loop transmission. Here we intentionally use a relatively higher integrator gain (in comparison with a typical integrator gain of $K_I = 2\pi f_c/10$) in order to reserve enough gains for the later AFC design. The loop-shaping controller $C_{ls}(s)$ is designed such that the feedback loop has at least 30 [deg] phase margin at the crossover frequency f_c and the loop transmission rolls off for frequencies above f_c. The loop shaping controller also notches out the resonance peak at 45 kHz. After a couple of iterations,

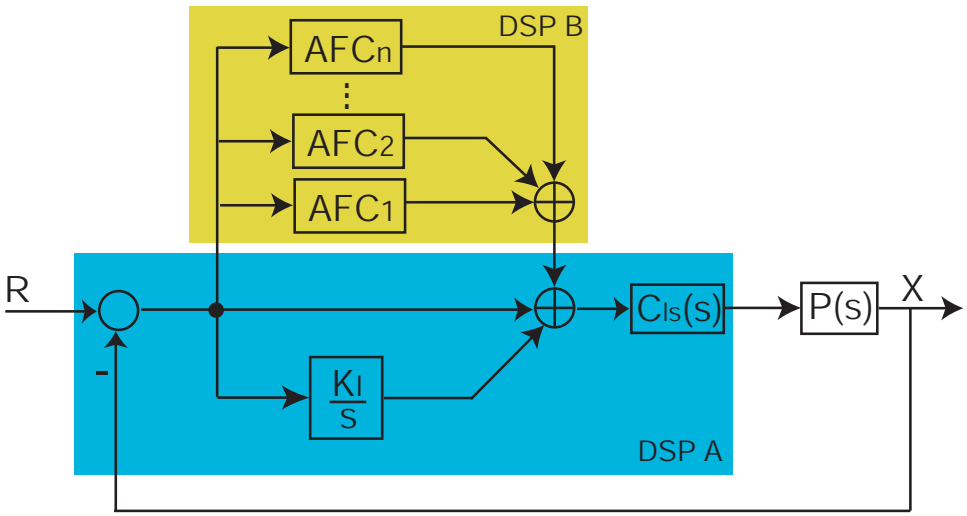


Figure 7-4: The controller structure for the SM2-based UFTS.

we design the loop-shaping controller $C_{ls}(s)$ as

$$C_{ls}(s) = 4.15 \times 10^{13} \left(\frac{s+25130}{s+301680}\right)\left(\frac{s+18850}{s+439800}\right)\left(\frac{s^2+20610s+8.72\times 10^{10}}{s^2+43830s+9.87\times 10^{10}}\right), \tag{7.2}$$

which is composed of a double-lead compensator and a notch filter. The crossover frequency is at 10 kHz, and correspondingly K_I is set at 12566. This continuous controller is converted into a discrete controller by matched pole-zero method [18] with a sampling frequency of 500 kHz. Figure 7-5 shows the computed frequency response of the whole compensator including both the integrator and the loop shaping controller $C_{ls}(s)$. The resulting measured loop transmission is shown in Figure 7-6. The unity gain crossover frequency of the negative loop transmission is 10 kHz with 30 degree phase margin. The -180° phase crossover frequency is at 15kHz with 3.5 dB gain margin as shown in the figure.

Figure 7-7 shows the closed-loop response of the SM2-based UFTS from the reference command to the actual UFTS motion. The -3dB bandwidth of the closed-loop response is 23 kHz. The closed-loop magnitude response has a peak at 13.35 kHz, where the negative loop transmission has the minimum vector margin.

Figure 7-8 shows a 500 nm step response of the closed-loop system. The $10\%-90\%$

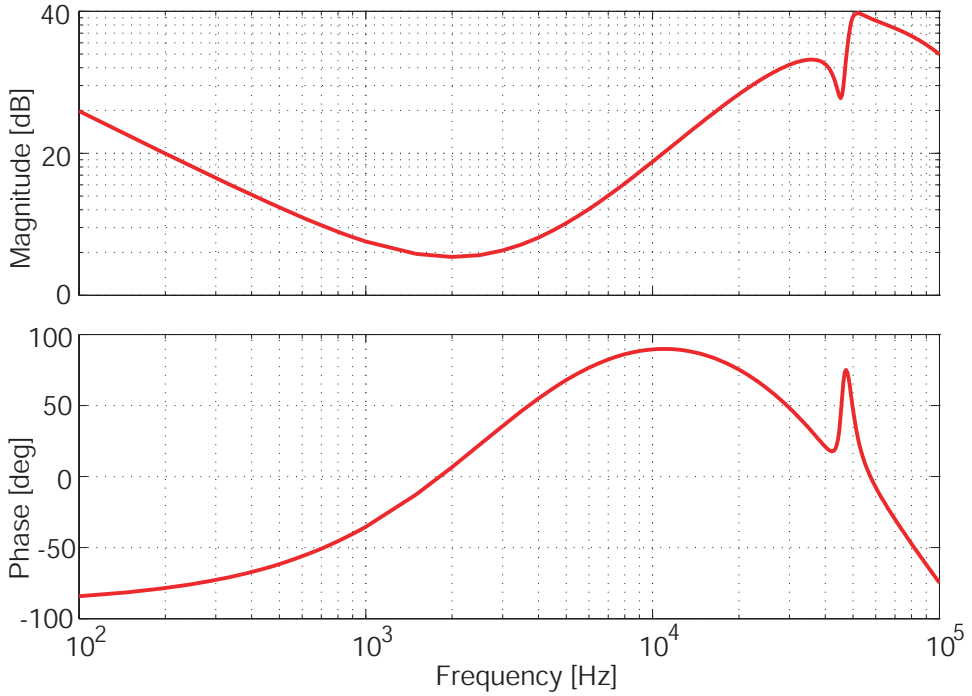


Figure 7-5: Frequency response of the loop shaping controller and the integrator.

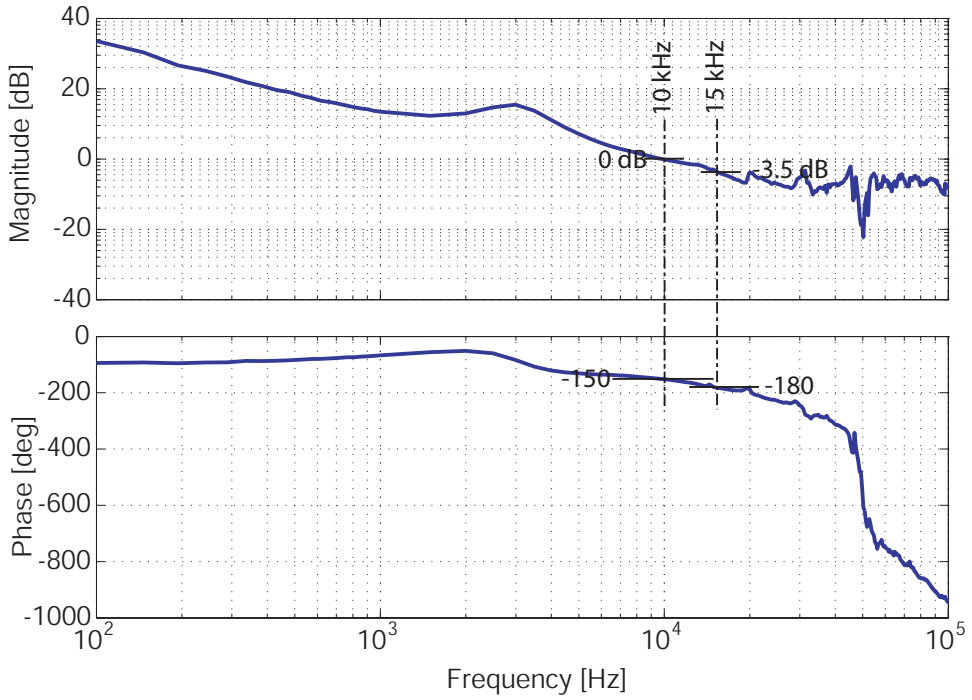


Figure 7-6: The measured negative loop transmission of the compensated UFTS.

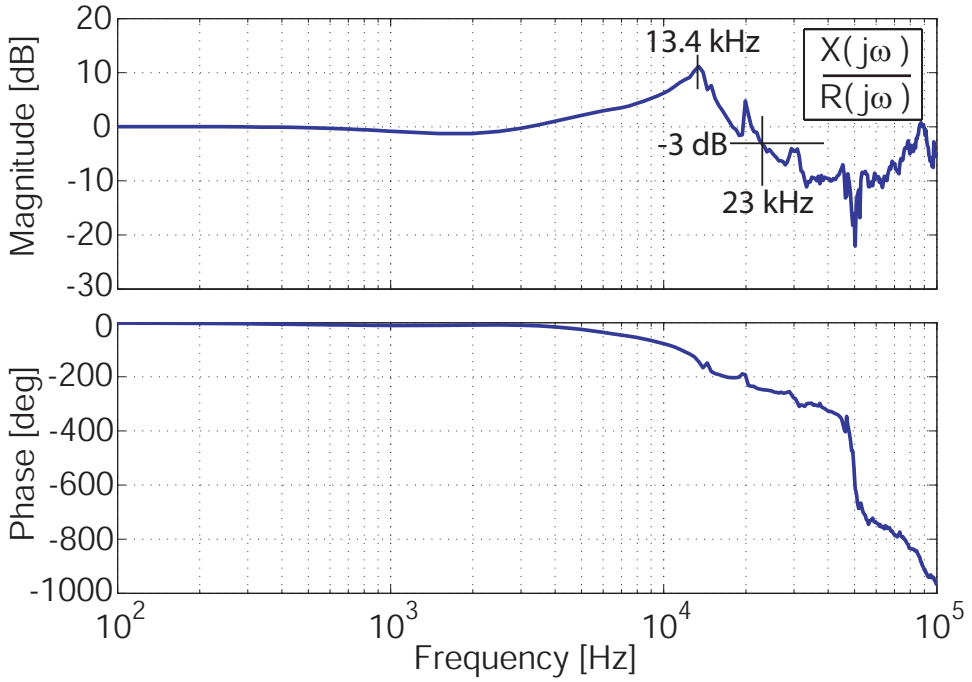


Figure 7-7: The measured closed-loop frequency response .

rise time is 25 μs, and the dominant pole shows an oscillation of 75 μs, corresponding to the peak-amplitude frequency of 13.35 kHz in Figure 7-7.

For position regulation, the steady-state positioning error is 1.7 nm RMS.

7.3 Adaptive Feed-forward Cancellation

In the diamond-turning application with a UFTS, the tool path has a quasi-periodic pattern at harmonics of spindle rotation frequency, as the tool path in two consecutive revolutions are almost identical, given the typical a small feed-rate. Moreover, there also exist periodic disturbance forces from both the spindle unbalance and the cutting process. To enhance the reference-tracking/disturbance-rejection performance at harmonic frequencies of spindle rotation, an adaptive feed-forward cancellation (AFC) controller has been studied and applied.

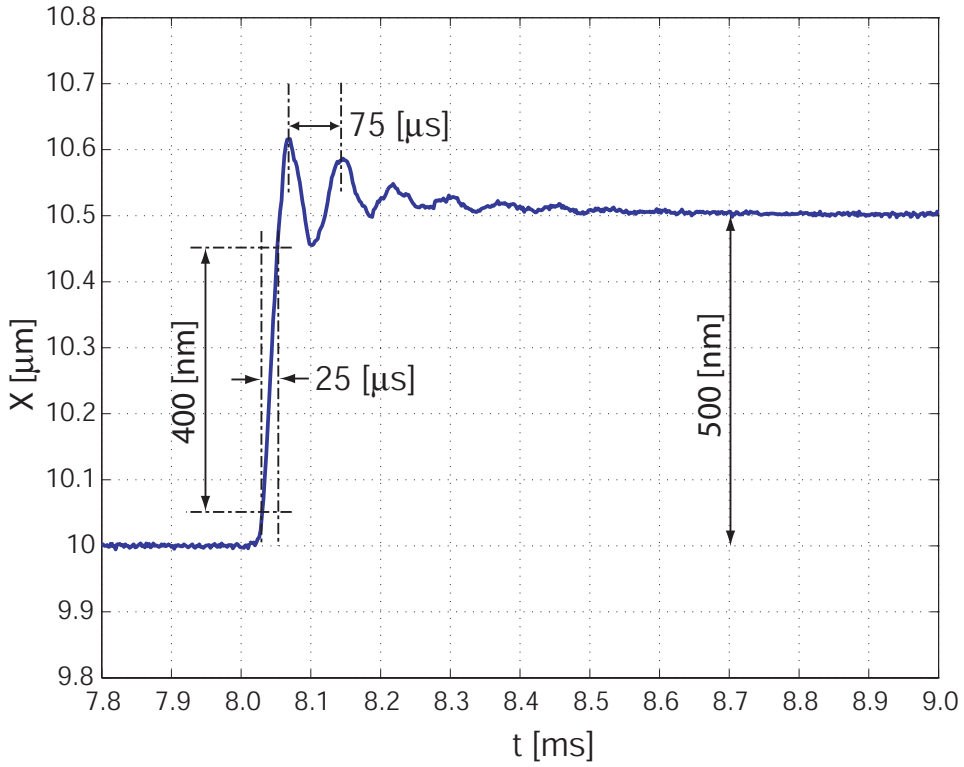


Figure 7-8: 500 nm step response of the closed-loop system.

7.3.1 Prior-Art AFC control

There have been a number of thesis in our lab on the adaptive feed-forward control theory and its application to fast tool servos [31, 8, 7]. These thesis have included a rather complete literature review of AFC control. The Adaptive Feed-forward Cancellation controller was developed to reject periodic harmonic disturbances as shown in Figure 7-9 [4, 50]. A plant $G(s)$ is disturbed by a periodic signal D of a known frequency ω_n. In order to maintain zero output from this plant, according to linear system theory a control signal U_n of the same frequency can be input to the plant to cancel the disturbance, by setting the amplitude and phase of U_n properly. The adaptive feed-forward cancellation is a method used to automatically calculate this required amplitude and phase for U_n. The AFC controller shown in the shaded region is composed of three parts: (1) a demodulator used to estimate the sine and cosine components of the error signal; (2) a parameter adaptation law; and (3) output signal generation. This nonlinear-looking algorithm in fact is surprisingly equivalent

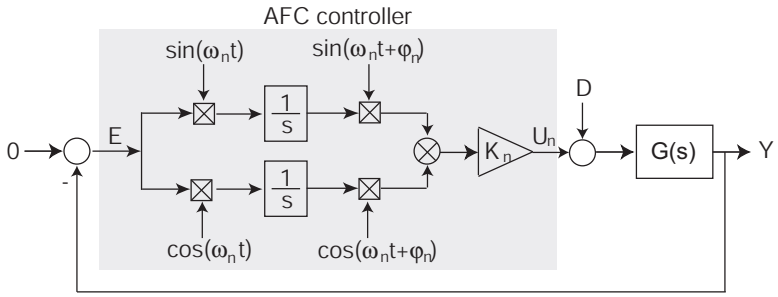


Figure 7-9: General principle of Adaptive Feed-forward Cancellation controller . The square block with cross inside represents a multiplier. Signal E is the tracking error, signal U_n is the AFC controller output, ω_n is the frequency of a periodic disturbance signal D, φ_n is the phase advance angle and K_n is the gain of the AFC controller.

to the linear-time-invariant system [30, 4, 50] with transfer function

$$A_n(s) \quad = \quad K_n \frac{s\cos\varphi_n - \omega_n \sin\varphi_n}{s^2 + \omega_n^2}. \tag{7.3}$$

Based on this result, the AFC controller analysis and design can be greatly simplified. As the AFC controller is equivalent to this resonator with the resonance frequency ω_n, we can understand its effectiveness with linear analysis tools. Because the loop transmission gain is infinity at frequency ω, the transfer function from the disturbance to the plant output is zero at the frequency ω_n. In the discrete-time domain, there exist similar results [50].

The AFC controller design reduces to the selection of the controller gain K_n and the phase advance angle φ_n. The phase advance angle was suggested to be chosen equal to $\varphi_n = -\angle G(\omega_n j)$ in [50] to ensure that the root locus of the closed loop system depart from the imaginary axis at right angles. In [31], it was shown that by setting the phase advance angle this way the closed loop transfer function from reference command to the plant output has a zero derivative. Root locus plot is suggested to set the AFC controller gains in [50, 31], and the Nyquist plot of $A_n(s)G(s)$ was used for the multiple AFC case in [6, 7].

AFC control can completely reject the periodic disturbance and ideally track the command signal at the frequency ω_n. However, in most engineering control problems, these requirements are not sufficient, because an AFC control almost leave the plant

$G(s)$ almost at open loop at frequencies other than ω_n. Therefore, a conventional controller, such as lead-lag compensator/PID, is still required. The question is where the AFC controller should be placed in the signal flow if a conventional controller is still used.

A natural and simple way to integrate the AFC controller into a conventional control loop is shown in Figure 7-10(a). The plant P(s) is first controlled in a closed loop by a conventional controller $C(s)$. This closed-loop system is viewed as a new plant $G(s)$, for which the previously discussed AFC controller design methods can be applied in a secondary loop. In this implementation, the tracking performance at frequencies other than resonating frequencies of AFCs suffers a lot because the secondary outer loop transmission gain is very low at non-resonating frequencies. This problem is not serious if the trajectory contains only sinusoidal components at resonating frequencies of AFC. For wide-spectrum trajectory tracking, this implementation is problematic. However, this shortcoming can be easily fixed by a plug-in type implementation as shown in Figure 7-10(b), which was used by Tomizuka [80] to implement repetitive controllers on a hard disk drive. From disturbance rejection point of view, two implementations in Figure 7-10(a) and (b) are exactly the same, but certainly (b) has better tracking performance than (a) at frequencies other than resonance. Figure 7-10(c) shows a third implementation by placing the AFC in parallel with a conventional controller C(s) in [89]. As the AFC is part of the loop transmission, there is no significant reduce on tracking perforce at frequencies other than AFC resonance frequencies.

So far, in the literature the AFC design methods associated with all these three implementations use a two-step procedure. First, a conventional controller $C(s)$ is designed to form a closed-loop system, and then this closed-loop system is treated as a new plant, upon which AFC controller design is executed. This design procedure does not directly consider the interaction between the conventional controller $C(s)$ and the AFC controller. The AFC controller is put in an inferior status, as the AFC has to accept the result caused by the conventional controller. Actually, the inner-loop controller design strongly influences the AFC design and this influence is not

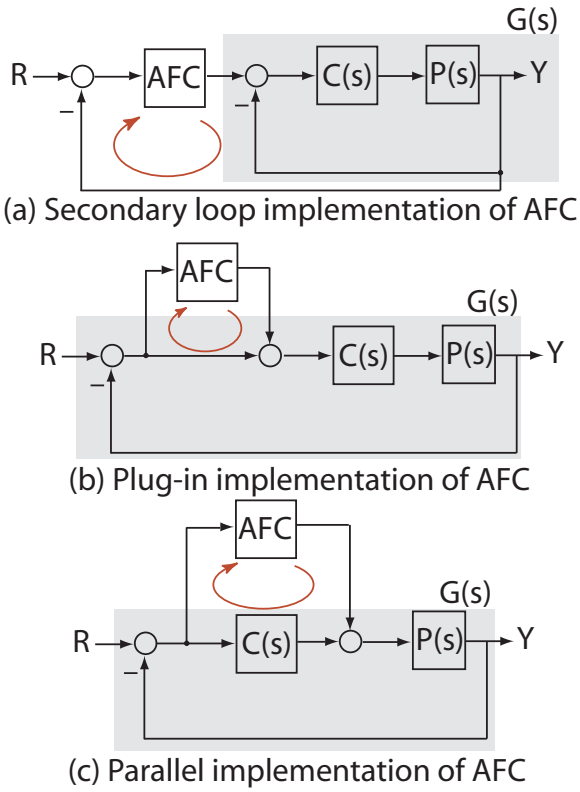


Figure 7-10: Three ways of incorporating AFC into the conventional control loop.

considered in the inner-loop design. Another problem in this sequential loop closure design of AFC controllers is that the phase/gain margin in the outer secondary loop is not very meaningful when using Nyquist-plot based methods. In fact, the so called new plant $G(s)$ is not a true plant, but a combination of a true physical plant P(s) to be controlled and a controller (quite often digital). From robustness point of view, only the uncertainty of $P(s)$ should be of concern and be accommodated with phase/gain margins, as there is almost no uncertainty in the digital controller, which can be implemented much more accurately than the physical plant $P(s)$. In Figure 7-10(a), (b) or (c), when designing the AFC controller using the loop transmission frequency response of the outer loop as indicated by the brown arrow, the phase/gain margins of the negative loop transmission of $AFC(j\omega)G(j\omega)$ are misleading, because these margins are for the uncertainty of $G(s)$. In reality, the uncertainty can happen only in $P(s)$, and all the other blocks are digital controllers with almost no uncertainty.

To illustrate the possibly distorted margins from the outer loop point of view,

Figure 7-11 shows an example system from two views. The plant $P(s) = 1/S^2$, and the inner loop controller is $C(s) = 2 \times 10^6(s + 400)/(s + 2000)$. The resulting inner loop system $P(s)C(s)$ has a phase margin of 42° and a crossover frequency of 1000 rad/sec. Further an outer loop controller $C_2(s) = 300/s$ is added into the feedback loop. If from the outer loop point of view as shown in Figure 7-11(a), the loop transmission of $C_2(s)G(s) = C_2(s)P(s)C(s)/\left(1 + P(s)C(s)\right)$ has a phase margin of 82°, which appears to be more than sufficient. However, if looking the same plant with the same control from the loop transmission $\left(1 + C_2(s)\right)P(s)C(s)$ in which $P(s)$ is a serial element as shown in Figure 7-11(b), the phase margin reduces to 25°, which is the phase change of the true plant $P(s)$ that can be accommodated. In this sense, the outer loop point of view is very dangerous, and could result in a very aggressive controller without being noticed. In short, **classical Nyquist-like loop shaping with digital controllers should be applied in the loop where the controlled plant $P(s)$ is a serial element, because only in this loop the claimed phase/gain margin can be enjoyed by the $P(s)$, which is the only element that should be accommodated with uncertainty.** In other loops where the plant $P(s)$ is not a serial element, the resulting margins are questionable and sometimes misleading, no matter their values.

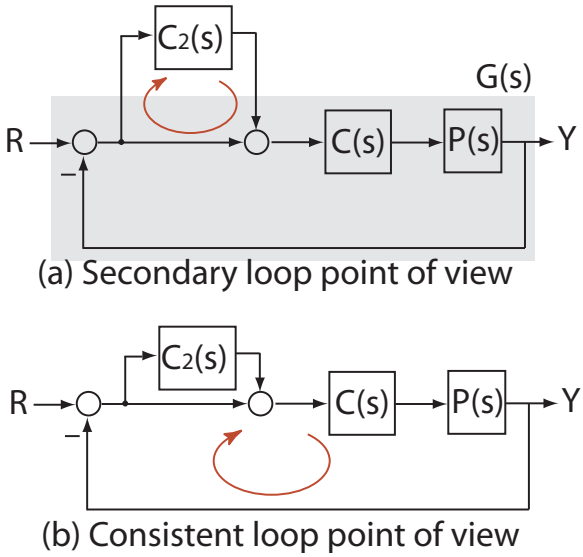


Figure 7-11: Two point of views on loop shaping.

Next, we discuss a controller design framework utilizing the above insight, in which the conventional controller and the AFC controller design are analyzed and designed in one consistent loop, and the phase/gain margins are always considered for plant $P(s)$. More importantly, the design trade-off between the conventional controller and the AFC controller can be easily performed with clear intuition.

7.3.2 Embedded AFC Controller Implementation

Adaptive feed-forward controllers are resonators, which have infinite gains at resonance frequencies. Their effectiveness in disturbance rejection and trajectory tracking comes from this infinite gain property. Actually, this philosophy is not new to controls. An integrator is applied in the same way at zero frequency, to completely reject a DC disturbance or to achieve zero steady state tracking error under a step input. Mathematically, an integrator has the same format as an AFC controller,

$$A_0(s) \quad = \quad K_0 \frac{s}{s^2 + 0^2} = K_0 \frac{1}{s}. \tag{7.4}$$

As a result, an integrator can be considered a member of the AFC controller family. Since all non-zero-resonating AFCs can be implemented in a parallel format, why should an integrator be treated differently? In fact, almost all conventional motion controllers already have an integrator inside. As all AFC controllers belongs to the same family, the place where an integrator is located could also be considered as the place for other AFC controllers. Following this logic, we put the AFC controllers in parallel with the integrator of the conventional controller as shown in Figure 7-12(b). In Figure 7-12(a), a plant $P(s)$ is controlled by a conventional controller, which is composed of an integrator of gain K_I in parallel format and a loop shaping controller $C_{ls}(s)$ (usually a lead-lag like controller). On this basis, AFC controllers $A_n(s)$ $(n = 1, 2, ..., N)$ are embedded into the loop transmission in parallel with the integrator, the gain of which is adjusted to K_0 to ensure the stability. In next sections, we will discuss the AFC controller parameter design.

In this embedded implementation, the negative loop transmission is changed from

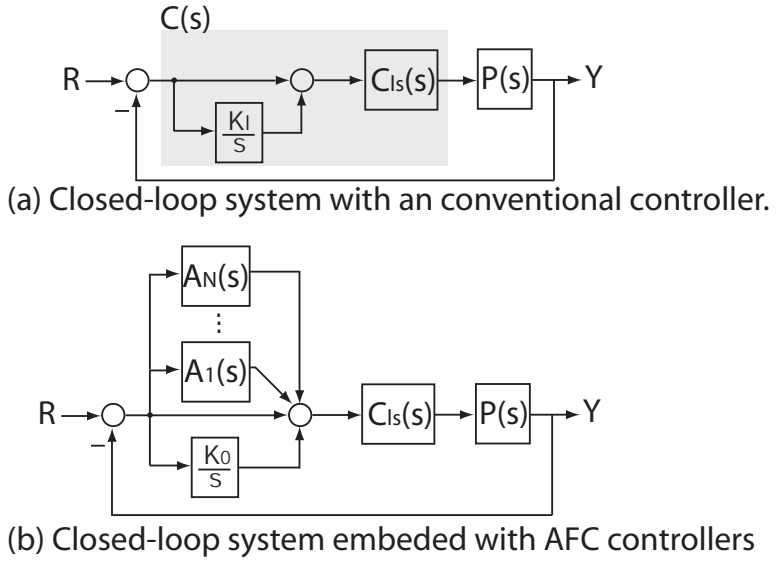


Figure 7-12: Embedded implementation of AFC controllers.

$(1 + K_I/s)C_{ls}(s)P(s)$ to $(1 + K_0/s + A_1(s) + ... + A_N(s))C_{ls}(s)P(s)$. The basic shape of the loop transmission remain unchanged, except that in the AFC active regions the loop transmission is greatly enhanced. Therefore, there is no tracking performance loss issue as in Figure 7-10(a). Roughly speaking, this embedded implementation is similar to Figure 7-10(b) and (c), but the fundamental difference is that in this embedded implementation of AFC's the AFC controllers and the conventional controller can be designed in a unified framework (a consistent loop transmission). Note that there is no need to worry about the complex zeros in this embedded framework, because $\left|1 + K_0/s + A_1(s) + ... + A_N(s)|_{s=j\omega}\right| > 1$. As a result, there will be no multiple crossover frequencies when investigating the loop transmission of $(1 + K_0/s + A_1(s) + ... + A_N(s))C_{ls}(s)P(s)$. This is very different from the outer loop point of view, in which there are multiple crossover frequencies in the loop transmission of $(A_1(s) + ... + A_N(s))\,G(s)$ in Figure 7-10.

Next, we will present the controller design intuition according to the generic relation between an integrator and the AFC controller family.

7.3.3 Sub-crossover AFC Design

In this subsection, we focus on the parameter design for sub-crossover AFCs, which have resonance frequencies less than half of the unity gain crossover frequency of the negative loop transmission $(1 + K_I/s)C_{ls}(s)P(s)$. When we design the conventional controller (including an integrator), typically the integrator gain is designed at 0.1--0.2 times of the crossover frequency of loop transmission to make sure that the phase lag from the integrator element $1 + K_I/s$ can be compensated properly, for instance $K_I = 0.2\omega_c$, where ω_c is the crossover frequency of the loop transmission. Actually, only the frequency response around the frequency ω_c will influence the stability of the closed-loop system. When the AFC controllers are plugged in, they will bring in additional phase lag at the crossover frequency ω_c and thus impair the phase margin. For an AFC $A_n(s)$ with zero phase advance $\varphi_n = 0$, its response at the crossover frequency is

$$\begin{aligned} A_n(j\omega_c) &= K_n \frac{j\omega_c}{(j\omega_c)^2 + \omega_n^2} \\ &= \frac{K_n\omega_c^2}{\omega_c^2 - \omega_n^2}\frac{1}{\omega_c j}. \end{aligned} \tag{7.5}$$

This value is exactly a pure imaginary number. In comparison with an integrator, the AFC controller $A_n(s)$ with a gain K_n has the same frequency response at frequency ω_c as an integrator with a gain $K_n\omega_c^2/(\omega_c^2 - \omega_n^2)$. In this sense, an AFC controller can be equivalently treated as an integrator with proper gain adjustment. We refer this adjustment as resonance gain weighting function $W(\omega_n)$, which is defined as

$$W(\omega_n) = \frac{\omega_c^2}{\omega_c^2 - \omega_n^2}. \tag{7.6}$$

Figure 7-13 shows this gain weighting function. The horizontal axis is the normalized resonance frequency and the vertical axis is the gain weighting coefficient. From this curve, when $\omega_n < \omega_c$, the associated AFC controller $A_n(s)$ can be directly treated as an integrator. The higher the resonance frequency ω_n, the more weighting. This relation is highly nonlinear. For $\omega_n < 0.5\omega_c$, the gain weighting coefficient W do not

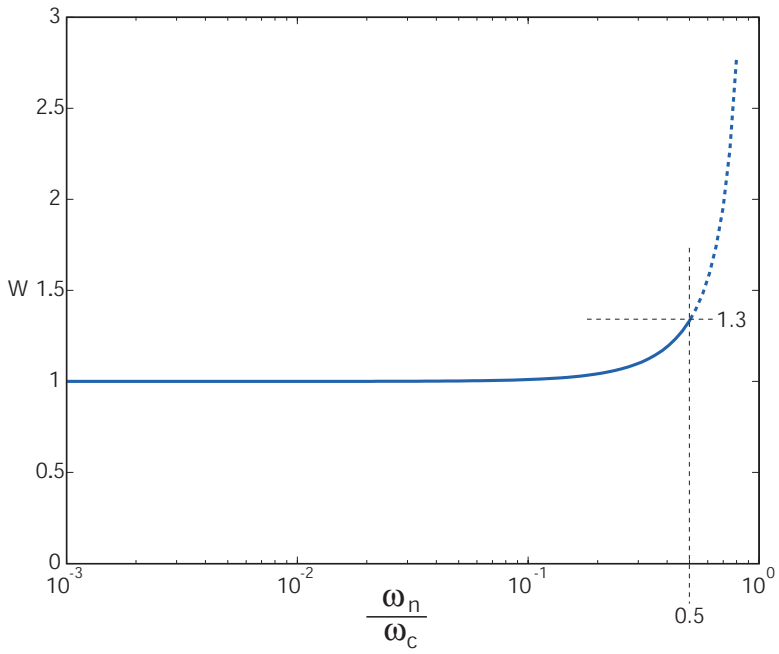


Figure 7-13: Resonator gain weighting function. ω_c is the crossover frequency, and the ω_n is the resonance frequency of a AFC controller $A_n(s)$.

change much with the highest value of 1.3 at $\omega_n = 0.5\omega_c$. Based on this result, for $\omega_n < \omega_c/2$, we can approximate $A_n(s) = K_n s/(s^2 + \omega_n^2)$ as K_n/s, from the stability point of view.

This relation greatly simplifies the AFC controller design, especially for AFC's with multiple resonance frequencies. The controller in Section 7.2 is

$$C(s) = \left(1 + \frac{K_I}{s}\right) C_{ls}(s). \tag{7.7}$$

After embedding with AFC controllers, the new controller is

$$C_E(s) = \left(1 + \frac{K_0}{s} + \sum_{n=1}^{N} \frac{K_n s}{s^2 + \omega_n^2}\right) C_{ls}(s), \tag{7.8}$$

where N is the total number of AFCs with resonating frequencies less than ω_c, and K_0 is the adjusted integrator gain after embedding AFCs.

To ensure controller $C(s)$ and controller $C_E(s)$ have the same complex gain at the crossover frequency ω_c of the compensated loop in Section 7.2, we need match their

response at $s = j\omega_c$,

$$\begin{aligned}
\left(1 + \frac{K_I}{j\omega_c}\right) C_{ls}(j\omega_c) &= \left(1 + \frac{K_0}{j\omega_c} + \sum_{n=1}^{N} K_n \frac{j\omega_c}{(-\omega_c^2 + \omega_n^2)}\right) C_{ls}(j\omega_c), \\
\Longrightarrow -\frac{K_I}{\omega_c} &= -\frac{K_0}{\omega_c} + \sum_{n=1}^{N} \frac{K_n \omega_c}{(-\omega_c^2 + \omega_n^2)}, \\
\Longrightarrow K_I &= K_0 + \sum_{n=1}^{N} \frac{K_n \omega_c^2}{(\omega_c^2 - \omega_n^2)} \\
\Longrightarrow K_I &= \sum_{n=0}^{N} K_n W(\omega_n). \qquad (7.9)
\end{aligned}$$

When the phase balancing equation (7.9) is satisfied, the controller with AFCs embedded in will provide the same phase margin and the same crossover frequency as the the controller in Section 7.2.

This result can be summarized as follows: **If a plant $P(s)$ is stabilized by $\left(1 + \frac{K_I}{s}\right) C_{ls}(s)$ with a crossover frequency ω_c, then the plant $P(s)$ can be stabilized by $\left(1 + \frac{K_0}{s} + \sum_{n=1}^{N} \frac{K_n s}{s^2 + \omega_n^2}\right) C_{ls}(s)$ with the same crossover frequency, the same phase margin, and improved gain margin, as long as $K_I = \sum_{n=0}^{N} K_n W(\omega_n)$ and $\omega_n < \omega_c$ for all $0 \le n \le N$.**

As for setting each AFC gain, the tradeoff picture is very clear. The total allowable gain sum is K_I, and is distributed among all K_n's. As for exactly how to set the gain at each resonance frequency, we will not give a rule, because this depends completely on the specific case. For instance, if the actual plant has very strong disturbance at the third harmonics ω_3 and a small disturbance at the second harmonics ω_2, then we can make K_3 much bigger than K_2. If the disturbance and trajectory spectrum is not clear, we suggest the following design as a starting point,

$$K_0 = \frac{K_I}{2}, \qquad (7.10)$$

$$K_n = \frac{K_I}{2NW(\omega_n)}, \text{ for } 1 \le n \le N. \qquad (7.11)$$

This rule is based on the argument that the integrator is more important, as we always need to follow step response properly and reject some DC disturbance. Next

we explore further some generic properties of the AFC controller family.

Gain and Bandwidth Equivalence in AFC controllers

We define the active region of an AFC controller as the frequency region in which its frequency response has an amplitude greater than or equal to 1. This definition implies that inside the active region the AFC controller can help enhance the loop transmission, and outside this region the integrator is negligible. The width of this region is defined as its bandwidth. Obviously, the active region of an integrator K_0/s is from $\omega = 0$ to $\omega = K_0$, and correspondingly its bandwidth is K_0. Here, we get an very interesting relation that the gain term of an integrator is equal to its bandwidth. In this sense, we can say that the integrator gain is equivalent to its bandwidth. Actually, this is a common feature of all AFC's without phase advance. For an AFC controller $A_n = K_n s/(s^2+\omega_n^2)$, it is active in a band, and the active region is between a lower bond frequency ω_{nL} and a higher bond frequency ω_{nH}. These two frequencies are the solutions to the following equation

$$\|K_n \frac{\omega j}{\omega_n^2 - \omega^2}\| = 1. \tag{7.12}$$

Solving this equation, we have

$$\begin{aligned} K_n^2\omega^2 &= (\omega_n^2 - \omega^2)^2 \\ (\omega_n^2 - K_n\omega - \omega^2)(\omega_n^2 + K_n\omega - \omega^2) &= 0 \end{aligned} \tag{7.13}$$

As a result,

$$\begin{aligned} \omega_{nH} - \omega_{nL} &= K_n, \\ \text{and } \omega_{nH}\omega_{nL} &= \omega_n^2. \end{aligned} \tag{7.14}$$

These relations show that the bandwidth the AFC controller $A_n(s)$ is the same as its gain K_n, and ω_n is the geometric measure of these two frequencies. Therefore, by setting the gain of an AFC controller, we know the width of its effective region.

From this point of view, the AFC controllers design can be interpreted as the total bandwidth of all AFCs (including the integrator) is restricted to $\omega_c/5$ in the loop shaping controller design, and then we redistribute the total bandwidth of $\omega_c/5$ among the integrator and other AFCs based on the weighting function. Hence, we have a rough picture of over how wide a frequency range an AFC controller can alter the frequency response of $C_{ls}(s)P(s)$. Once we know the crossover frequency limit of a system by playing with $C_{ls}(s)$, we can get a rough idea on the allowable bandwidth (gain) of an AFC controller.

7.3.4 Super-crossover Frequency AFC design

From the previous discussion, in Figure 7-13 we have seen that the resonance gain weighting function $W(\omega_n)$ defined in equation (7.6) will increase rapidly when the resonance frequency ω_n of $A_n(s)$ is greater than $\omega_c/2$. As an extreme case, the weight goes to infinity when $\omega_n = \omega_c$, and thus it is impossible to design a AFC of resonance frequency at ω_c without phase advance. Therefore, a phase advance parameter φ_n should be introduced to help the AFC controller design when $\omega_n > \omega_c/2$. We refer such an AFC controller

$$A_n(s) \quad = \quad K_n \frac{s\cos\varphi_n - \omega_n \sin\varphi_n}{s^2 + \omega_n^2} \tag{7.15}$$

as super-crossover AFC controller. Next, we investigate the bandwidth and phase advance of super-crossover AFC controllers.

Let's denote ω_{nL} and ω_{nH} as the frequencies where the AFC controller $A_n(s)$ has a unity gain frequency response, and the distance $\omega_{nH} - \omega_{nL}$ is the bandwidth of $A_n(s)$. So ω_{nH} and ω_{nL} are the solution to

$$|A_n(j\omega)| \quad = \quad 1. \tag{7.16}$$

Solving this equation yields

$$\begin{aligned} |K_n \frac{j\omega \cos\varphi_n - \omega_n \sin\varphi_n}{\omega_n^2 - \omega^2}| &= 1 \\ \frac{K_n}{\omega_n^2 - \omega^2} \sqrt{\omega^2 \cos^2\varphi_n + \omega_n^2 \sin^2\varphi_n} &= \pm 1 \\ \frac{K_n \omega_n}{\omega_n^2 - \omega^2} \sqrt{\frac{\omega^2 \cos^2\varphi_n + \omega_n^2 \sin^2\varphi_n}{\omega_n^2}} &= \pm 1 \\ \sqrt{1 + \frac{(\omega^2 - \omega_n^2)\cos^2\varphi_n}{\omega_n^2}} &= \pm \frac{\omega_n^2 - \omega^2}{K_n \omega_n} \end{aligned} \tag{7.17}$$

By defining $\delta = (\omega - \omega_n)/\omega_n$ (or equivalently $\omega = \omega_n(1+\delta)$), we can simplify the above equation as

$$\begin{aligned} \sqrt{1 + \delta(\delta+2)\cos^2\varphi_n} &= \pm\delta \frac{\omega_n}{K_n}(2+\delta) \\ \delta &= \pm \frac{K_n}{\omega_n} \frac{\sqrt{1 + \delta(\delta+2)\cos^2\varphi_n}}{2+\delta}. \end{aligned} \tag{7.18}$$

In super-crossover AFC controllers, the generated sine/cos waves amplitude parameters in Figure 7-9 are usually adapted at a time constant $(1/K_n)$ much longer than one sinusoidal period $2\pi/\omega_n$. Hence we can assume that $K_n << \omega_n$. Under this assumption, we have $\delta << 1$ according to equation (7.18). As a result, we can get approximate solutions to equation (7.18) as

$$\delta = \pm \frac{K_n}{2\omega_n}. \tag{7.19}$$

According to the definition of δ, we get solutions to equation (7.16) $\omega_{nH} = \omega_n + K_n/2$ and $\omega_{nL} = \omega_n - K_n/2$. Therefore the bandwidth of the super-crossover AFC controller $A_n(s)$ is $\omega_{nH} - \omega_{nL} = K_n$. We arrive at the same conclusion of gain and bandwidth equivalence as that for the AFC controllers without phase advance, but under the assumption of $K_n << \omega_n$.

Another property of super-crossover AFC controller is the phase advance φ_n. The

frequency response of $A_n(s)$ can be calculated as

$$\begin{aligned} A_n(j\omega) &= K_n \frac{s\cos\varphi_n - \omega_n \sin\varphi_n}{s^2+\omega_n^2}\bigg|_{s=j\omega} \\ &= K_n \frac{s}{s^2+\omega_n^2} \frac{s\cos\varphi_n - \omega_n \sin\varphi_n}{s}\bigg|_{s=j\omega} \\ &= K_n \frac{s}{s^2+\omega_n^2}\bigg|_{s=j\omega} \left(\cos\varphi_n + j\frac{\omega_n}{\omega}\sin\varphi_n\right) \\ &= K_n \frac{s}{s^2+\omega_n^2}\bigg|_{s=j\omega} \left(\cos\varphi_n + j\frac{1}{1+\delta}\sin\varphi_n\right), \end{aligned} \tag{7.20}$$

where $\delta = (\omega - \omega_n)/\omega$ as defined before. In the active region $\omega_{nL} < \omega < \omega_{nH}$ of the controller $A_n(s)$, we have $\delta << 1$ as discussed above. Therefore equation (7.20) can be approximated as:

$$A_n(j\omega) \simeq K_n \frac{s}{s^2+\omega_n^2}\bigg|_{s=j\omega} (\cos\varphi_n + j\sin\varphi_n). \tag{7.21}$$

This result shows that inside the active region of the super-crossover AFC controller, the frequency response is shifted by an angle of φ_n relative to that of an AFC controller without phase advance.

To illustrate this result, Figure 7-14 (a) shows the Nyquist plot of $1 + A_n(j\omega)$ with a zero phase advance in $A_n(s)$ for ω from 0 to infinity. It starts at the point 1 in the complex plane, goes to infinity vertically, then surrounds the right half plane, and returns to 1. If a phase advance angle φ_n is incorporated into the controller $A_n(s)$, the corresponding Nyquist plot of $1 + A_n(j\omega)$ will rotate anti-clockwise by an angle of φ_n in the active region of $A_n(j\omega)$ as shown in Figure 7-14 (b). The Nyquist plot $C_{ls}(j\omega)P(j\omega)$ of the conventional loop without AFC controllers is shown in Figure 7-14(c). On this basis, further the Nyquist plot of the whole loop transmission $(1 + A_n(j\omega))\, C_{ls}(j\omega)P(j\omega)$ with an AFC controller $A_n(s)$ is shown in Figure 7-14(d). In order for the loop transmission not to encircle the "-1" point in the complex plane, the big half circle generated by the AFC controller $A_n(s)$ should be rotated such that its diameter should directly face the "-1" point as shown in Figure 7-14 (d). According

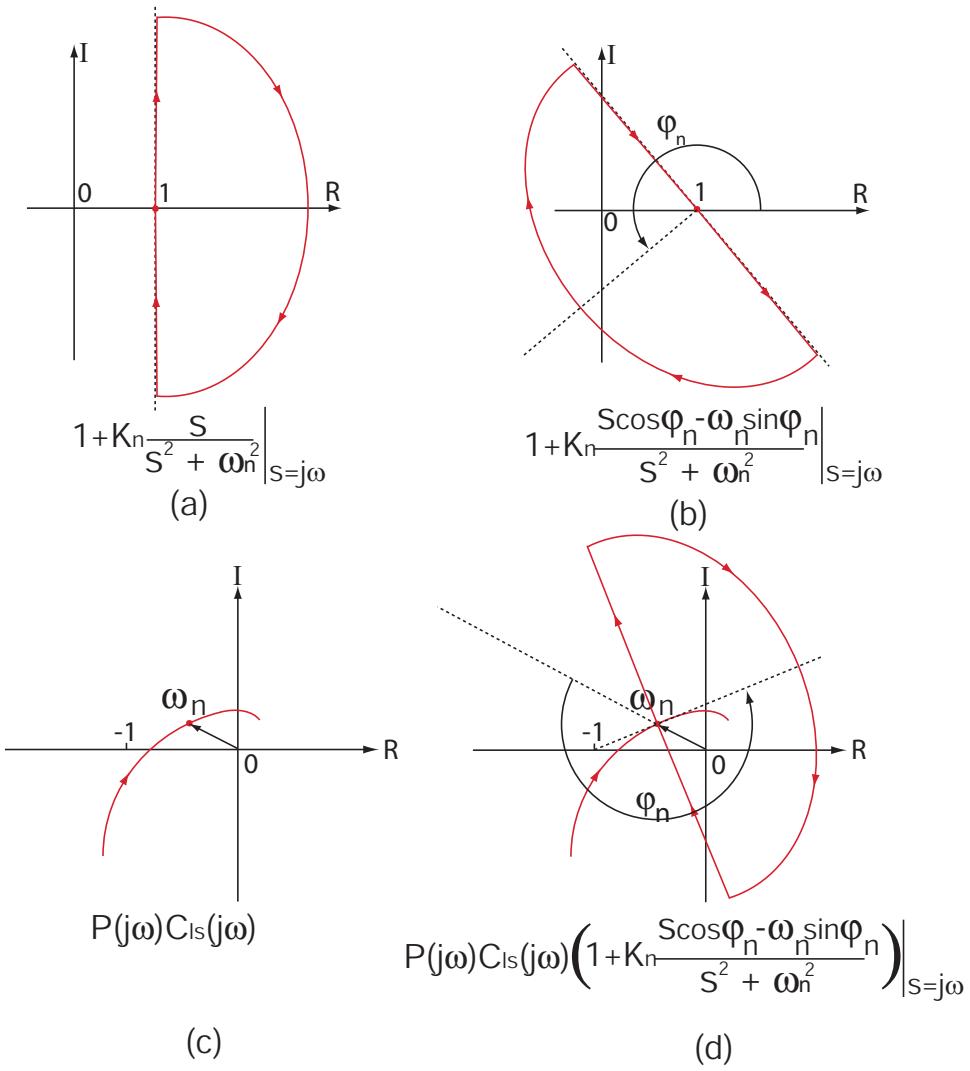


Figure 7-14: The Nyquist plot for the loop transmission with a super-crossover AFC controller embedded in. (a) The Nyquist plot of AFC controller $A_n(s)$ with a zero phase advance angle. (b) The Nyquist plot of AFC controller $A_n(s)$ with a phase advance angle φ_n. (c) The Nyquist plot of the conventional loop transmission $C_{ls}(s)P(s)$. (d) The Nyquist plot of loop transmission $(1+A_n(s))\,C_{ls}(s)P(s)$ with a phase advance angle φ_n.

to this rule, the phase advance angle is the difference between the phase of complex vector $P(j\omega_n)C_{ls}(j\omega_n) - (-1)$ and that of $P(j\omega_n)C_{ls}(j\omega_n)$, which is expressed as

$$\varphi_n = \angle \frac{1 + P(j\omega_n)C_{ls}(j\omega_n)}{P(j\omega_n)C_{ls}(j\omega_n)}. \tag{7.22}$$

This is the opposite value of the phase angle of $P(s)C_{ls}(s)/(1 + P(s)C_{ls}(s))$, and this result is the same as the prior art in [50, 31].

As shown in the Figure 7-14(d), in the vicinity of ω_n, the loop transmission is not a problem because of the properly set phase advance. Based on the above analysis, the active region (where the Nyquist plot blows a big "balloon") is between $\omega_n - K_n/2$ and $\omega_n + K_n/2$. Beyond this region, the loop transmission is almost identical to the conventional control loop transmission $P(j\omega)C_{ls}(j\omega)$. If we follow the same rule of designing K_n for sub-crossover AFC controllers, then we can have $K_n = K_I/(2N)$. Considering that $K_I = \omega_c/5$, we have $K_n = \omega_c/(10N)$. Since $\omega_n > \omega_c/2$, we get $K_n < \omega_n/(5N)$, which satisfy our assumption of $K_n << \omega_n$. For example, our conventional control loop crossover frequency is 10 kHz, and correspondingly we set the total AFC bandwidth as $K_I = \omega_c/5 = 2$ kHz. If we want to design an AFC controller with resonance frequency at 12 kHz, we can set its gain at 200 Hz, which means that the whole loop transmission is only altered in the region from 11.9 kHz to 12.1 kHz. For such a narrow band, we even do not need to check the Nyquist plot. If we want to design an super-crossover AFC controller with a more aggressive gain, we need to plot the Nyquist plot carefully to ensure the proper phase/gain margin.

7.3.5 AFC Experimental Result

Based on our presented embedded implementation of AFC and the associated controller design methods, we experimentally test its effectiveness on the UFTS. The command trajectory for the fast tool servo to track is a 3 kHz sine wave of 16 μm peak-to-valley amplitude, the peak acceleration of which is $\pm$300 G. The reason for selecting this frequency is that we can easily test both the sub-crossover AFC controller design and the super-crossover AFC controller design, given that the crossover

frequency of the compensated loop is 10 kHz as discussed before.

We design and implement 6 AFC controllers at $\omega_1 = 3$ kHz, $\omega_2 = 6$ kHz, $\omega_3 = 9$ kHz, $\omega_4 = 12$ kHz, $\omega_5 = 15$ kHz, and $\omega_6 = 18$ kHz. The controller parameters are listed in Table 7.1. In this design, we set the bandwidth of each AFC controller

Table 7.1: AFC control parameters for 3 kHz trajectory tracking.

Harmonic index	Harmonic frequency [Hz]	Gain	Phase [deg]
1	3000	200	0
2	6000	200	0
3	9000	200	100
4	12000	200	200
5	15000	200	200
6	18000	200	200

at 200 rad/sec, which corresponds to 33 Hz. As the gain sum of all AFC controllers is only 200 Hz bandwidth, which is much less than the total available 2 kHz bandwidth, we keep the integrator gain the same as before $K_0 = K_I = 2\pi \times 2000 = 12560$. In this case, we didn't lose any integrator gain at all after incorporating the AFC controllers in the embedded implementation. For the super-crossover frequency, we set the phase advance angles only based on a rough estimation of the phase of the conventional closed-loop system, without carefully reading the actual phase from the closed-loop response curve, because the big balloon in Figure 7-14(d) can accommodate the phase rotation error up to ± 90 degrees. As the super-cross AFC controllers have a bandwidth of only 30 Hz, the Nyquist plot of the loop transmission is altered at a fairly narrow band. With these intuition in mind, we didn't bother to check the resulting Nyquist plot. In fact, the designed AFC controller always work well, so I haven't get a chance to plot any Nyquist plot with "balloons".

The experimental results are shown in Figure 7-15, with temporal-domain results on the left side and frequency-domain results on the right. When no AFC is implemented, the tracking error is 0.77 μm RMS as shown in the top row, and is dominated by the base harmonic component at 3 kHz. After the first AFC controller $A_1(s)$ with $\omega_1 = 3$ kHz is activated by turning its gain from 0 to 200, the tracking error is reduced from 770 nm to 42 nm RMS, which is dominated by higher order harmonics. From

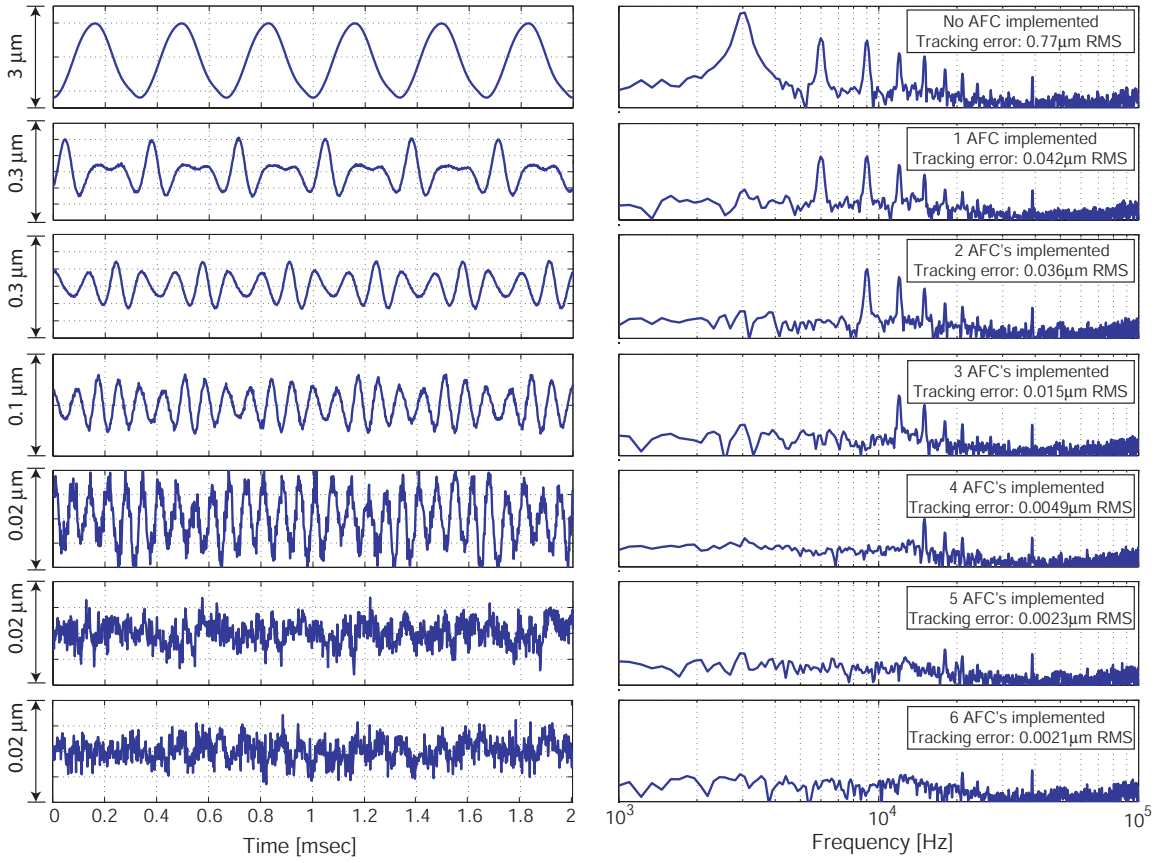


Figure 7-15: Adaptive feed-forward cancellation control tracking errors when the trajectory is 16 μm peak-to-valley 3 kHz sinusoidal trajectory with peak acceleration of $\pm$300 G . The left side shows the time domain tracking errors, and the right side shows the FFT of the tracking error. From the top to the bottom, we sequentially shows the experimental result of no AFC, 1 AFC, ..., 6 AFCs. For example, the nomenclature 3 AFCs mean that three AFCs are implemented at 3 kHz, 6 kHz, and 9 kHz.

the frequency plot, we can see that the error component at 3 kHz is completely eliminated. When both $A_1(s)$ and $A_2(s)$ are activated, the error reduces to 36 nm RMS and the error component at 6 kHz is also eliminated. After all 6 AFC controllers are activated, the tracking error reduces to 2.1 nm RMS!! Remember that the regulation error of the UFTS under the conventional control loop is 1.9 nm RMS for constant position regulation. For such a high acceleration trajectory, the AFC controllers greatly improve the tracking performance, almost as good as constant position holding. In the test experiment, we did not see benefits in reducing the steady state error when higher AFC gains were used.

7.3.6 AFC Controller Design Summary

We have presented a new framework for AFC controller implementation and design. We put the AFC controller in parallel with the already implemented integrator, which is shown to be a member of the AFC controller family. This embedded format enables us to focus on only one consistent loop transmission and thus to use an unified point of view on examining both the conventional controller design and the AFC controller design. The resulting benefits include

- after incorporating the AFC as in Figure 7-12(b), the tracking performance has almost no change, except gaining some enhancement in the vicinity of AFC resonant frequencies;

- we can consistently design both the integrator and AFC controllers on only one negative loop transmission, which is $(1 + K_I/s + AFC(s))\, C_{ls}(s)P(s)$. As a result, the trade-off among the loop shaping controller design, the integrator, and AFC controllers can be clearly evaluated in an unified framework;

- based on the similarity between AFCs and an integrator, the controller tuning process can be greatly simplified by redistributing the integrator gain.

We further explored the equivalence property of gain and bandwidth for AFC controllers. Based on this property, we design sub-crossover AFCs according to a phase balance condition in equation (7.9). With this intuition in mind, we clearly know how much gain can be set and how to trade off gains among AFC controllers. We can easily set the gain of 1000 sub-crossover AFC controllers in one second, thanks to the presented generic property inside the AFC controller family.

For super-crossover AFC controllers, we presented the equivalence property of gain and bandwidth for super-crossover AFC controllers with arbitrary phase advance, by assuming that the parameter adaptation time is much longer than one signal period, which is acceptable for most cases. Consequently, we can set the super-crossover AFC controller gains by controlling their bandwidth. We illustrated how to set the phase advance angle. For more aggressive super-crossover AFC controller gain tuning, a

Nyquist plot of $(1 + A_n(s))\, C_{ls}(s)P(s)$ should be carefully drawn, but in the loop where $P(s)$ is a serial element.

7.4 Multi-rate Sampling Implementation

The whole digital controller is implemented as shown in Figure 7-4. The conventional controller is calculated in DSP A, and AFC controllers are in a second processor DSP B. This allows for multi-rate sampling. The direct benefit from multi-rate sampling is that more computational resource can be available for controller calculations. For example, the conventional control is calculated at 500 kHz in DSP A, and AFC controllers are calculated at 250 kHz. In this way, we can get twice the computation time for the AFC controllers. The experimental results show that up to 10 AFC controllers can be implemented in DSP B with a sampling rate of 250 kHz. The reason to implement AFC controllers at a slower sampling frequency is that AFC controllers are narrow-band controllers in parallel with a direct feed-through path and thus the additional delay on the AFC controllers have no significant effects on the overall loop transmission. If the conventional loop controllers are implemented at a slower rate, then the increased delay will directly reflect on the loop transmission.

One important thing in multi-rate implementation with multi-processing is that the sampling rates of two processors must exactly have an integer multiplicative relation, otherwise the system performance will significantly degrade. For example, if DSP A samples at 500 kHz and DSP B samples at 251 kHZ, then a 3 kHZ signal generated by DSP B will be accompanied by an aliased component at 5kHz ($=3 + 251 \times 2 - 500$), which seriously disturbs the system. However, if DSP B samples at 250 kHz instead, then the aliased component at 253 kHz will has negligible impact on the system performance, because the plant has almost no output at that high frequency.

7.5 Trajectory Generation

Accurate tool motion depends not only on good tracking performance of fast tool servo control, but also on the trajectory generation, which is a function of spindle rotation angle. In reality, the trajectory for the fast tool servo is generated based on a spindle position sensor, usually a quadrature encoder. When such a spatially-sampled sensor[1] is used for trajectory generation, two sampling events are involved: the spatial sampling by the encoder and the temporal sampling by the digital controller. Usually, the digital controller samples at a fixed temporal interval with a frequency f_{TS}. When the motion axis moves at constant speed, the spatial sampling by the encoder will correspond to another sampling frequency f_{SS}. As these two sampling events are not synchronized with each other, significant aliasing can be resulted. As shown in Figure 7-16(a), the position θ of a rotation axis is sampled at fixed spatial interval at frequency $f_{SS} =$. This value will be kept in the quadrature encoder spatial counter until the next spatial sample, and correspondingly this process is modelled as a spatial-domain zero-order hold (ZOH). The output from this spatial ZOH is sampled again by control sampling events at frequency f_{TS}. To better understand the aliasing process, Figure 7-16(b) shows an equivalent block diagram. As the trajectory generation is an algebraic function, it can be swapped with other elements. In Figure 7-16(b), the ideal trajectory Z^* is first sampled at f_{SS}. When passing through the ZOH, low-frequency components in Z^* will be duplicated at infinite number of higher frequencies. When sampled at f_{TS} again, the duplicated components will be fold back to low frequencies and contribute to aliasing.

For example, when an encoder with a resolution of 10000 counts/rev is mounted on a spindle rotating at 500 RPM, its output update rate is $10000 \times 500/60 = 83.3$ kHz. When we implement our control algorithm at a frequency of 500 kHz, this asynchronous nature between the computer sampling and the sensor update will introduce significant aliasing errors as illustrated in Figure 7-17. Figure 7-17(a) is the encoder output, which is a ramp signal and rolls back to zero every 10000 counts. When

[1]See more information in Chapter 6.4 about quadrature encoder.

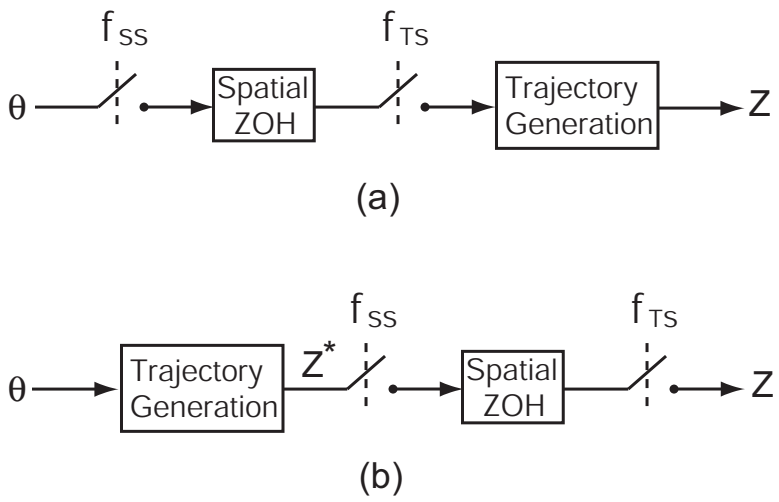


Figure 7-16: Spatial-temporal multiple sampling process. (a) The multiple sampling process in trajectory generation. (b) An equivalent block diagram for trajectory generation.

zoomed into a small square box on this ramp, the detailed waveform is a stair-case signal as shown by the blue line in Figure 7-17(b). If we generate a sine trajectory $Z = \sin(300\theta)$ directly based on the encoder output at 500 kHz sampling frequency, the resulting trajectory will be the blue curve in Figure 7-17(c), and correspondingly its FFT is represented by the blue curve in Figure 7-17(d), which contains a lot of aliased components. This up-sampling is very different from exact integer multiplicative up-sampling in conventional digital signal processing. This non-integer and asynchronous up-sampling can be understood as two steps: (1) up-sampling from 83.3kHz to the continuous domain by a zero-order-hold, and then (2) down-sampling this wide spectrum signal to 500 kHz.

Another problem in the spindle encoder is its nonlinear property caused by the misalignment in the manufacturing and/or the installation. This nonlinearity will result in harmonic error components in the generated trajectory.

To deal with this problem, a band-pass filter has been used to filter the trajectory generated directly on the encoder output [59] as shown in Figure 7-18(a). For single frequency trajectory generation, this method is very effective. But the filter design is highly dependent on the trajectory and they are strongly coupled together. When the trajectory frequency is changed, the filter needs to be changed accordingly. If

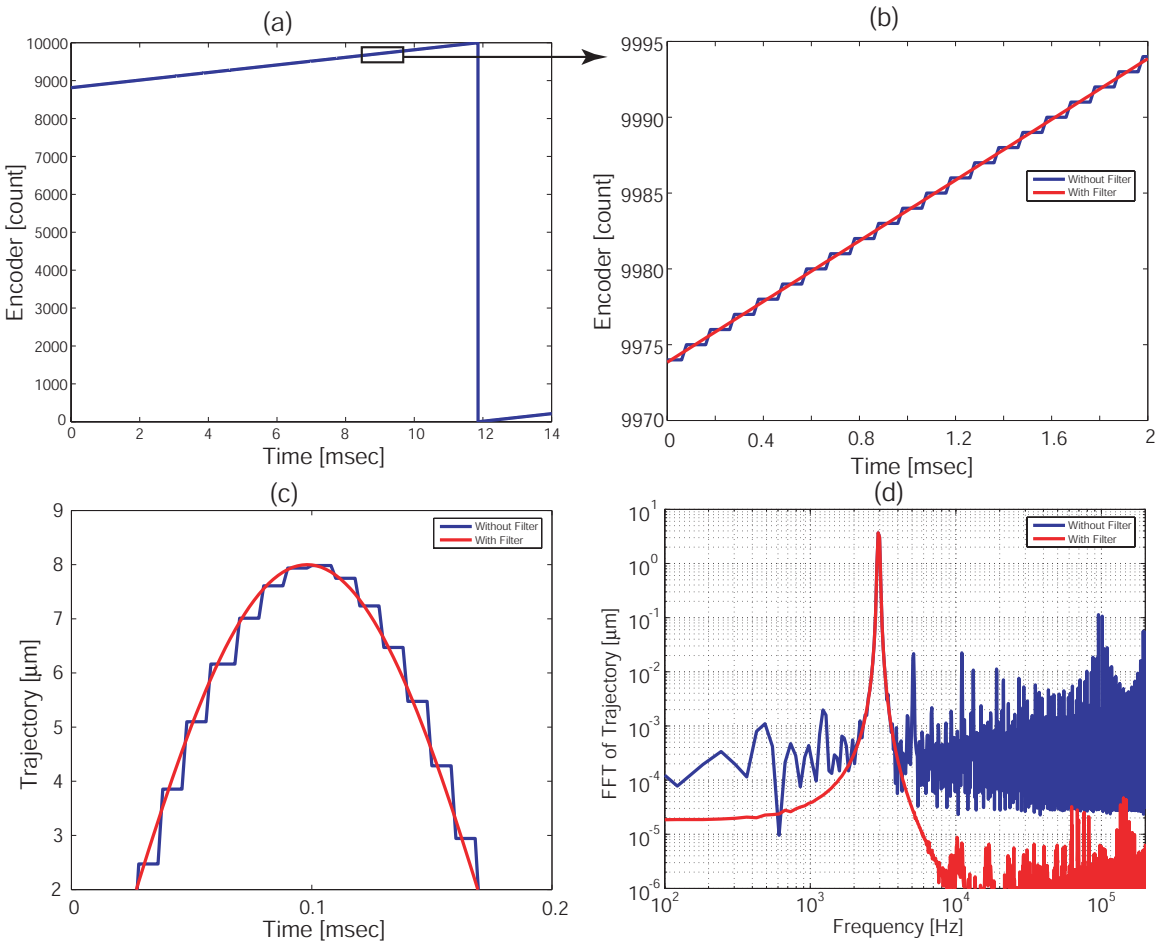


Figure 7-17: Simulated trajectory generation process. The spindle speed is at 590 RPM. The spindle encoder has 10000 counts per revolution. The sampling frequency for the spindle encoder is 500 kHz. The trajectory to be generated is $Z = 8\sin(300 \times 2\pi \times \frac{\theta}{10000})[\mu\text{m}]$, where θ is the spindle position in counts and Z is the generated trajectory in micron.

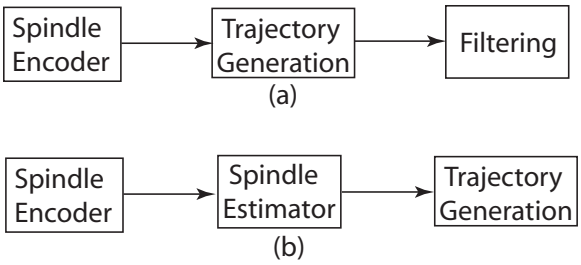


Figure 7-18: Two trajectory generation schemes.

the spindle is fluctuating, then it is difficult to maintain the filter gain at unity over a frequency band without phase distortion. If the desired trajectory has a wide spectrum, it will be very difficult to filter the trajectory, because both the generated trajectory and the error components have high frequency components and they are mixed together.

In order to solve all these problems, we present a solution, referred as synchronous filtering. We decouple the filtering and the trajectory generation. Instead of filtering the generated trajectory to remove the error components, we filter directly the spindle encoder output, and then generate a trajectory based on the filtered spindle position output as shown in Figure 7-18(b). As the main problems lie in the encoder and its output does not represent the spindle rotation accurately enough, the synchronous filtering is equivalent to a spindle position estimator, which takes the encoder output as its input and reconstructs the true spindle motion. As long as the spindle position estimator is designed properly, it can be used for a wide range of trajectories.

The difficulty in filtering the encoder output is its abrupt change of 10000 counts in each revolution as shown in Figure 7-17(a). If using a conventional filter on the encoder, significant error will be introduced at the discontinuity. We need to synchronously shift the spindle position estimator state. For this reason, we refer to our method as synchronous filtering. The principle of synchronous filtering is shown in Figure 7-19. The spindle rotation process is modelled as an integrator with spindle speed estimation n^* as its input, and the integrator output is the spindle position estimation. In the spindle model, we need to accommodate the encoder output discontinuity by introducing a signal d. When a discontinuity is detected from the

encoder output, an impulse with an amplitude of -10000 is inserted into the estimated spindle position. The position estimation error e is the difference between the encoder output and the spindle position. By properly designing filter $C_f(s)$, good tracking performance can be achieved at low frequencies. At high frequencies, where there exist significant aliasing and nonlinearity errors in the spindle encoder output, the filter $C_f(s)$ need to reduce its gain and does not respond to these errors. The overall effect of synchronous filtering is equivalent to a low pass filter, except that the spindle encoder discontinuity can be built into the spindle model and all the variables have a clear physical meaning.

In order to maintain zero steady-state error in tracking a ramp signal θ, the $C_f(s)$ is required to contain at least one integrator. For instance, $C_f(s)$ can be designed as

$$C_f(s) = \left(1 + \frac{\omega_f}{3s}\right) \frac{\omega_f}{1 + \frac{s}{3\omega_f}}, \tag{7.23}$$

where ω_f is the desired loop transmission cross-over frequency and the closed loop bandwidth of the spindle estimator. The bandwidth ω_f design is based on the actual spindle behavior, such that the synchronous filter can track the spindle speed variation (usually at low frequencies) and reject the aliasing/nonlinearity errors (usually at high frequencies). Therefore, ω_f should be located between them.

Applying this synchronous filter with 10 Hz bandwidth to the encoder output shown in Figure 7-17(a), the synchronous filter output result is plotted as the red curve in Figure 7-17(b), and the trajectory generated with this estimated spindle position are red curves of Figure 7-17(c) and (d). We can see that the aliasing errors have been greatly reduced.

7.6 Summary

The ultra-fast tool servo achieves a unity-gain crossover frequency of 10 kHz on the loop transmission with 30 degree phase margin by a loop shaping controller design. The limiting factor for the achievable crossover frequency is the phase delay caused

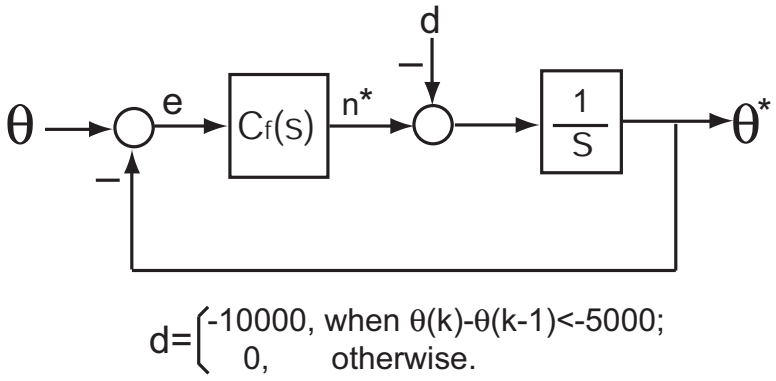


Figure 7-19: Synchronous filter structure.

by the filter inside the capacitance probe. In a 500 nm step response, the rise time is shown as 25 μs. The achievable acceleration is 500 G at 10 kHz. In the vicinity of resonance, the UFTS can achieve 1000 G acceleration at 3 kHz and 45 kHz. However, these high acceleration experiments are conducted in open-loop.

Adaptive feed-forward controllers are designed in a new framework, in which the AFC controller design can be greatly simplified and the trade off with the conventional motion controller can be evaluated in a consistent loop transmission. What's more, the equivalence between gain and bandwidth property are analyzed, and has been used to guide design of AFC controllers. The effectiveness of designed AFC controllers are experimentally demonstrated.

The whole controller was implemented in multi-rate sampling with multi-processors. The sampling rate ratio between two processors must be integer.

For accurate trajectory generation, a synchronous filtering method is presented to remove both aliasing error, encoder nonlinearities, and at the same time significantly increase the encoder resolution.

Chapter 8

Diamond Turning Experiments and Spindle Position Estimation

This chapter presents diamond turning experiment results using our SM2-based ultra fast tool servo (UFTS). Figure 8-1 shows the diamond turning machine setup for our face turning experiment. The workpiece is installed on the spindle with more details shown in Figure 8-2. The surface to be turned is the front ring-shaped surface. The dimension of this surface is 40 mm in outer diameter and 12.5 mm in inner diameter. The workpiece has a hole in the center to reduce the need for critical tool-height adjustments. Three surface patterns are cut in our experiments, including flat surfaces, one-dimensional(1-D) sine waves in the azimuthal direction, and two-dimensional(2-D) sine waves in both the azimuthal and the radial directions. Two materials are used in our test cut, which are aluminum 6061-T6 and oxygen free high conductivity(OFHC) copper. In the cutting experiments, cutting depths from 30 μm to 0.5 μm are tested without any symptom of chattering.

Figure 8-3 shows an interferometric microscope image of a 1-D surface turned on an aluminum workpiece. Figure 8-4 shows an interferometric microscope image of a 2-D surface on an OFHC copper workpiece.

In this chapter, we first look at the diamond turning machine and its controls for running the cutting test. Next, the flat surface cutting result is described in Section 8.3. Then we discuss the estimation and filtering of the spindle position signal,

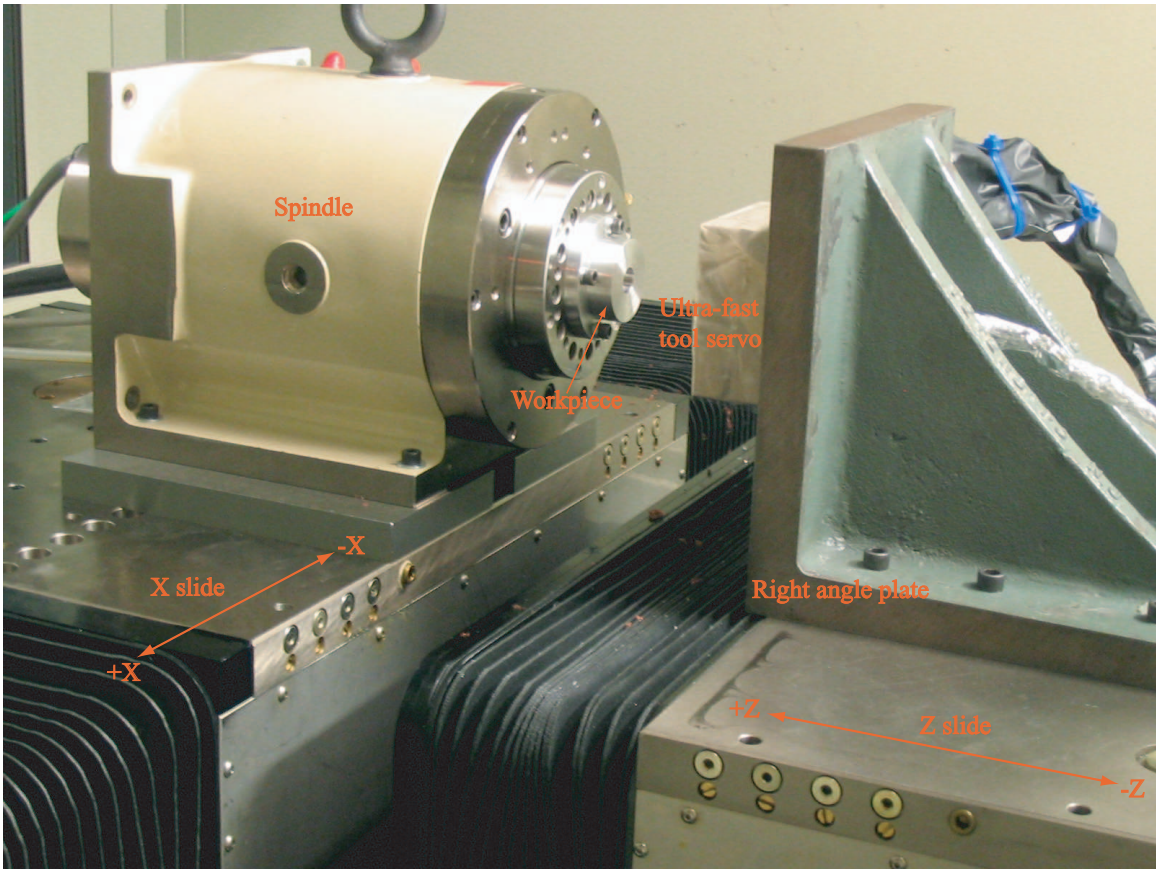


Figure 8-1: Diamond turning machine with the SM2-based UFTS installed. The front side of the UFTS is covered by a piece of latex rubber to prevent cutting fluid from entering the UFTS.

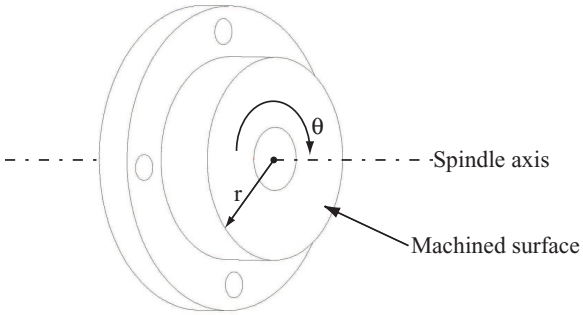


Figure 8-2: Workpiece used in the face-turning experiment.

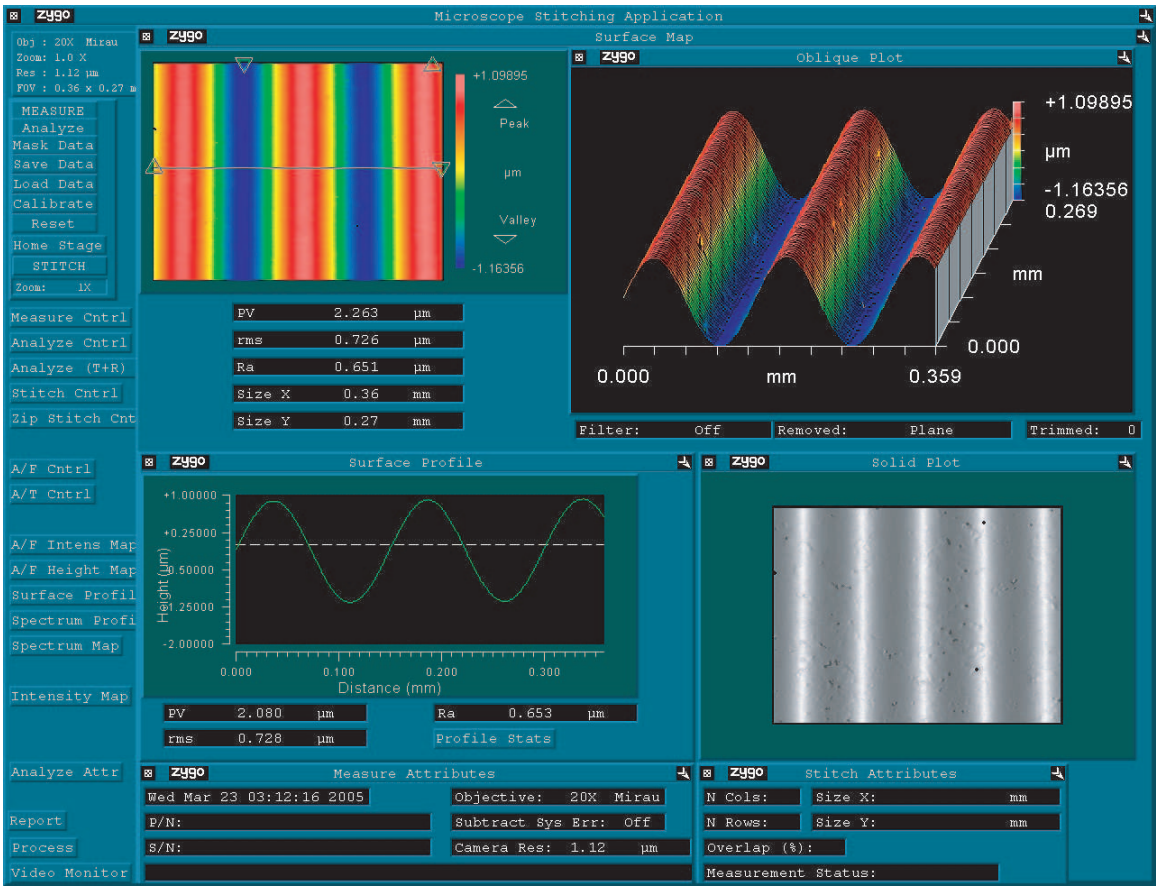


Figure 8-3: Surface metrology for diamond turned 1-D surface on an aluminum 6061-T6 workpiece. The sine-wave is in the θ direction with a spatial period of 0.5 degree, which corresponds to a 174 μm pitch on the outer edge and a 55 μm pitch on the inner edge of the machined surface. The spindle speed is 500 rpm. The peak-to-valley amplitude of the sine wave is 2 μm.

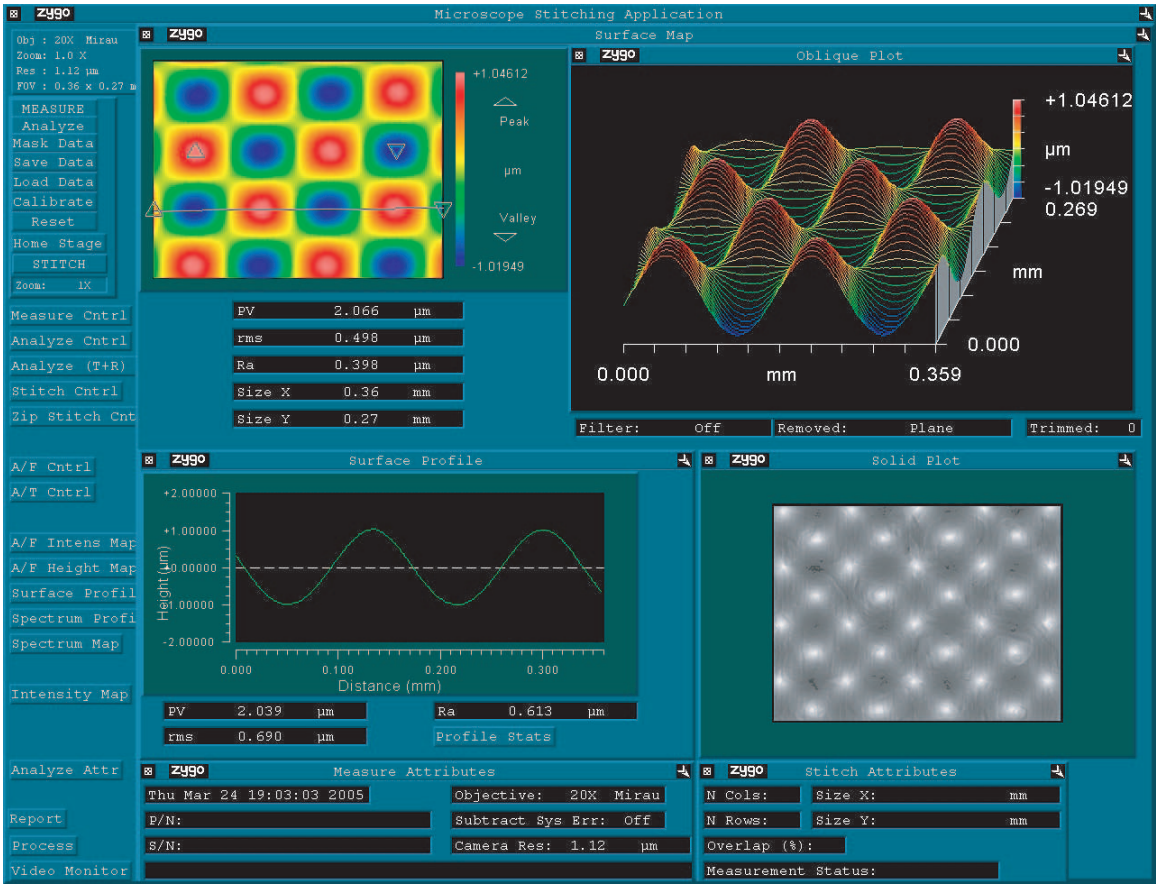


Figure 8-4: Surface metrology for diamond turned 2-D surface on an OFHC copper workpiece. The spatial period of the 2-D sinusoidal surface is 0.5 degree in the θ direction and 167 μm in the radial direction. The spindle speed is 500 rpm. The peak-to-valley amplitude of the sine wave is 2 μm.

which is a very important variable for trajectory generation for a non-rotationally symmetric surfaces. Based on this technique, we introduce the final 1-D and 2-D turning experiments.

8.1 Diamond Turning Machine and Its Control

8.1.1 Diamond Turning Machine

As shown in Figure 8-1, the SM2-based UFTS is installed on a commercial diamond turning machine[1] This diamond turning machine has a typical "T" construction: a Z slide travelling in the ±Z direction and a cross X slide travelling in ±X direction as shown in the figure. Both slides are suspended by hydrostatic bearings (with an oil pressure of 2 MPa), and are driven by iron-less three-phase linear motors. For each slide, there is a laser scale unit[2] for the slide position feedback. The laser scale unit outputs quadrature sinusoidal signals with a wavelength of 0.1379 μm, which is further interpolated by a laser scale detector board[3] with a factor of 16. As a result, the digital quadrature outputs have a resolution of about 8.6 nm. Each linear motor is driven by a commercial 3-phase PWM amplifier with a built-in current controller.

On top of the X slide, a motorized air-bearing spindle[4] is installed. At 150 psi (1 MPa) air pressure, this spindle achieves less than 25 nanometer error motion in the radial and axial directions and less than 0.1 micro radian tilt error motion. During our turning experiment, we supply compressed air of only 80 psi to the air-bearing spindle, well below the rated 150 psi. As pointed out by Richard Montesanti, our 140-psi boost pump generates a lot of floor vibration and cyclic pressure variation, which disturbs the spindle motion and so it was not used in these experiments. As the cutting force in our turning experiments are relatively small, operation at 80 psi is acceptable for our experiment.

[1]This diamond turning machine is manufactured by Moore Nanotechnology Systems, LLC, Keene, NH. The model number is Nanotech 350UPL.

[2]Manufactured by Sony Precision Technology, Inc. The model number is BS75A.

[3]Manufactured by Sony Precision Technology, Inc. The model number is BD15.

[4]This spindle is made by Professional Instruments Company, Hopkins, MN

The spindle motor PWM power amplifier has an integrated speed controller, which controls the spindle at a constant speed using a spindle-mounted resolver feedback signal. A second position feedback sensor (Heidenhain optical encoder) is also mounted on the spindle with digital quadrature outputs. This spindle position sensor is used for the tool trajectory generation. The resolution of this spindle encoder is 10,000 counts per revolution.

On top of the Z slide, the ultra-fast tool servo is mounted to a right angle plate, which is bolted to the Z slide.

8.1.2 Diamond Turning Machine Control

One graduate student in our lab, Marty Byl did the initial set up of the diamond turning machine, including the mechanical, pneumatic, and hydraulic systems. Marty also wired the electrical systems together and interfaced all the sensors and actuators to the dSPACE DS1103 control system. Based on his work, I constructed a diamond turning machine controller using Simulink/Stateflow. This controller runs in the DS1103 board.

Figure 8-5 shows two independent computer control systems: (a) the dSPACE computer, which controls the X and Z slide motion, and sends out the spindle command signal; (b) the ThunderStorm computer, which controls the UFTS servo motion. Both of them share the position signals of the X-Stage, Z-Stage, and Spindle optical encoders. Each controller system has its own host computer running its own graphical user interface (GUI).

All the servo loops of machine stages and the stage path trajectory generation are calculated by a Simulink program with a sampling frequency of 20kHz in dSPACE. The trajectory planing/switching/coordination and the system diagnosis are performed by a Stateflow program in dSPACE. The Stateflow program is calculated at a frequency of 1 kHz or whenever there is a new command from the Controldesk or whenever alarm signal is triggered. Basically, Stateflow is a finite state machine programming language and is good at event-based decision making. For example, a complete cutting process includes a sequence of trajectories: (1) the Z stage moves

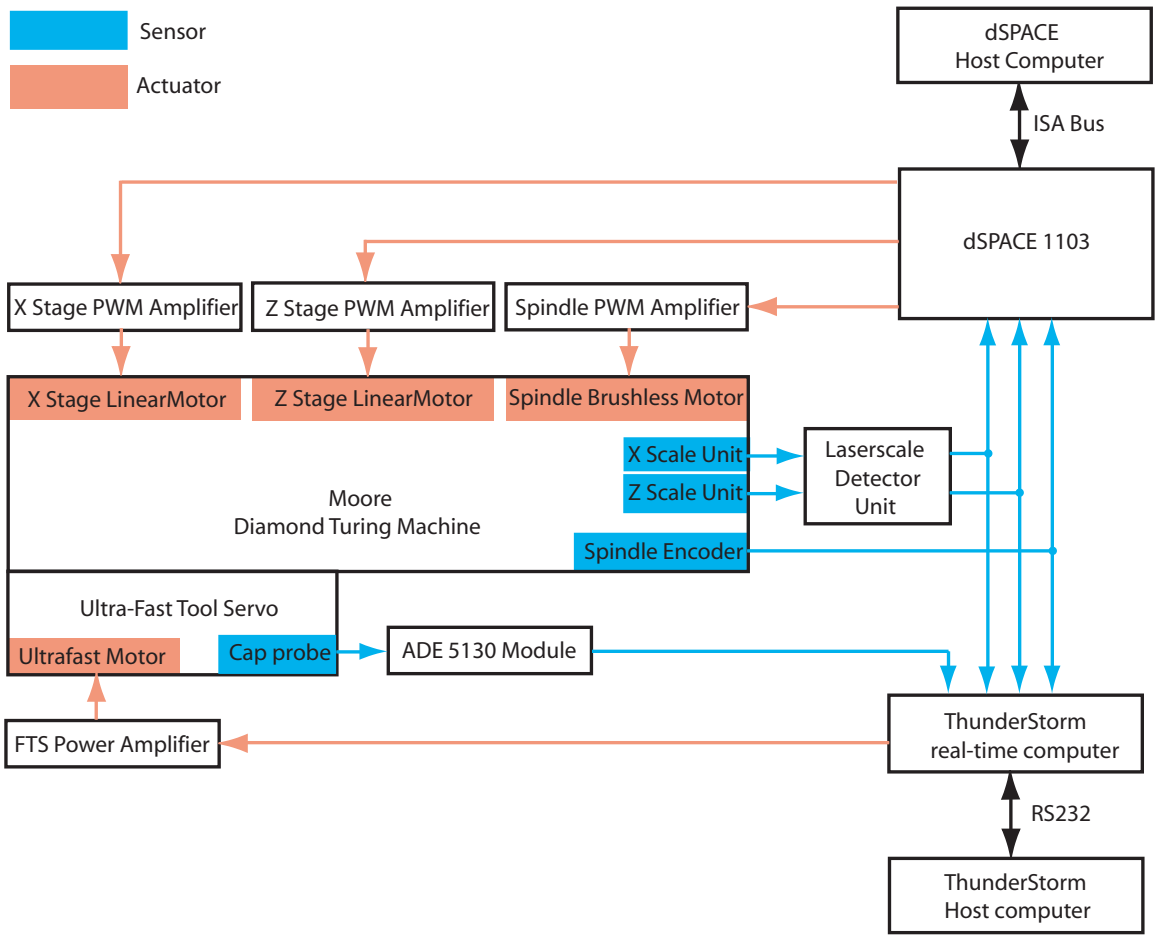


Figure 8-5: Diamond turning machine control system architecture.

to the desired cutting depth in the +Z direction; (2) the X stage moves along the X direction to cut the workpiece; (3) finally the Z stage retrieves back along the -Z direction. The switching between these trajectories can be easily implemented in Stateflow.

Each servo loop of the X-stage and Z-stage is compensated with a combination of a lead-lag compensator and a notch filter. At low frequencies, the loop transmission is enhanced with an integrator in parallel to the lead-lag controller. The crossover frequency of each servo loop is set at 100 Hz. Figure 8-6 and Figure 8-7 show the positioning errors of the X-stage and the Z-stage respectively. These errors result mainly from the 8.6 nm resolution of the linear position sensors.

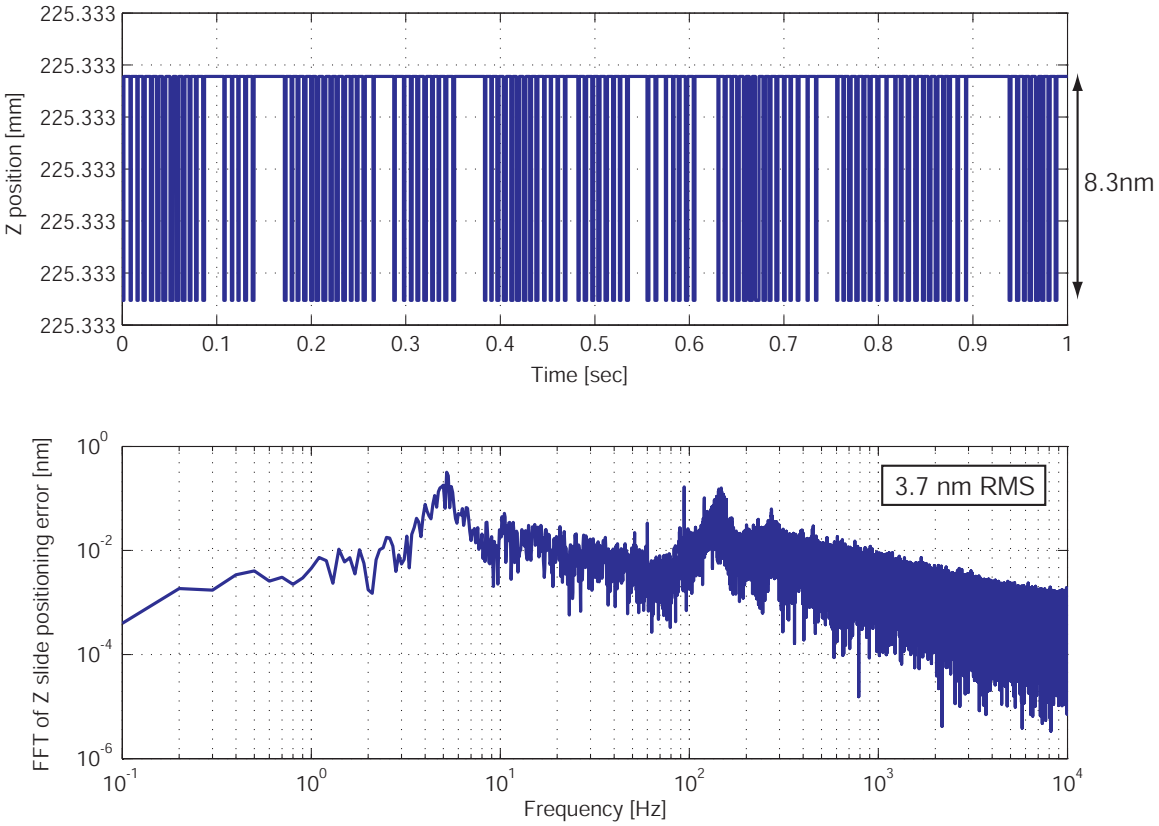


Figure 8-6: Z slide regulation error during air cutting. The RMS noise is 3.7nm

8.2 Ultra-Fast Tool Servo Integrated with Diamond Turning Machine

When the UFTS is installed on the DTM, the UFTS position regulation error degrades to 2.7 nm RMS (Sometimes, this value may further degrade to 3.7 nm RMS because of the interference from the power lines.), mainly because the capacitance position sensor picks up interference from the spindle PWM amplifier. With no excitation current in the UFTS, Figure 8-8 shows the position sensor output results with spindle-on and spindle-off. This comparison indicates that the spindle amplifier interferes with the capacitance probe output when the spindle is turned on. The closed-loop positioning error is shown in Figure 8-9 with the spindle turned on. However, the ThunderStorm computer doesn't pick up interference from the spindle PWM amplifier at all, as the A/D convertor ground floor noise is still 0.5∼0.6 nm, whether the spindle is on or off.

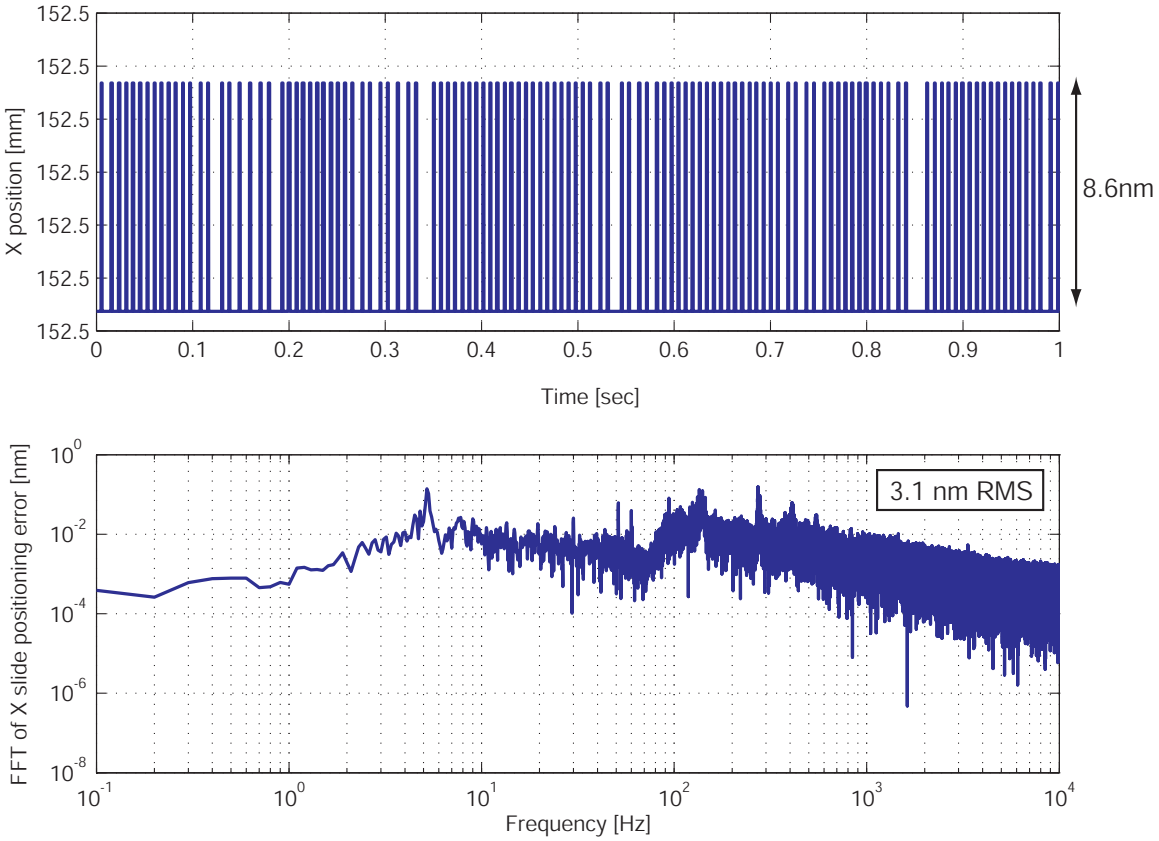


Figure 8-7: X slide regulation error during air cutting. The RMS noise is 3.1nm

This is the same as measured on the bench away from the DTM. This result is due to the analog differential inputs and careful regional management on the ThunderStorm board.

During the diamond turning operation, we apply a thin layer of mineral oil on both the diamond tip and the workpiece surface for cooling and lubrication, as suggested by Richard Montesanti in our lab.

One lesson that we learned from the diamond turning test with our UFTS is that the mineral oil will enter the inside of the UFTS via capillary effect if the UFTS is not sealed. Initially, we ran the cutting test without any preventive measures, and the UFTS lost its accelerating ability after four hours of continuous operation. The measured loop transmission indicated that the loop gain had been significantly reduced for frequency range from 100 Hz to 30 kHz. After disassembling the UFTS, it was found that the applied mineral oil entered every surface inside the UFTS as

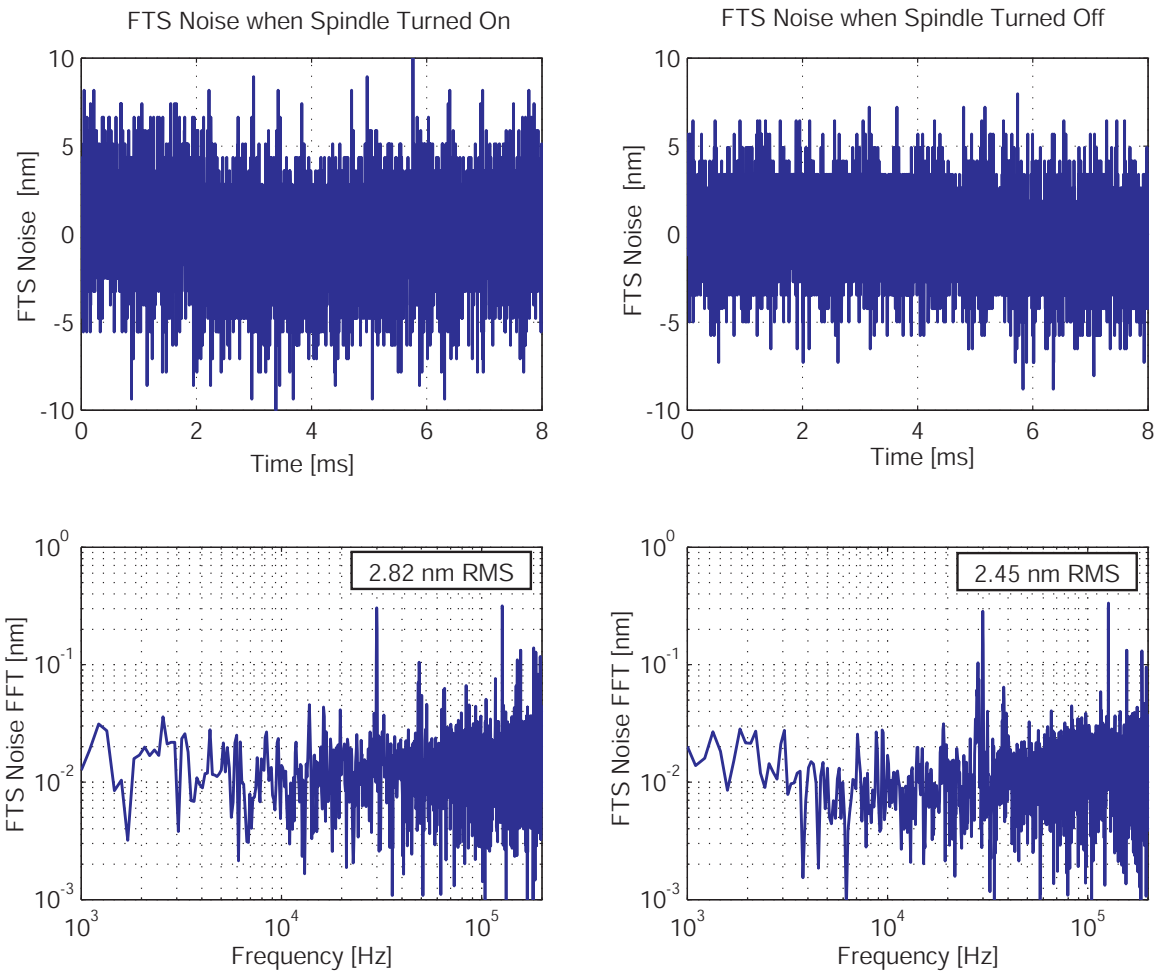


Figure 8-8: UFTS capacitance position sensor noise measurement under two conditions: spindle-on and spindle-off.

shown in Figure 8-10. Hence, the air gap between the armature and the stator core was completely filled with mineral oil. This viscous and incompressible fluid generates a huge amount of damping on the armature motion. To solve this problem, we use a piece of latex rubber to cover the whole front side of the UFTS as shown in Figure 8-1, and then install the tool shank on the tool holder with the latex rubber clamped in-between.

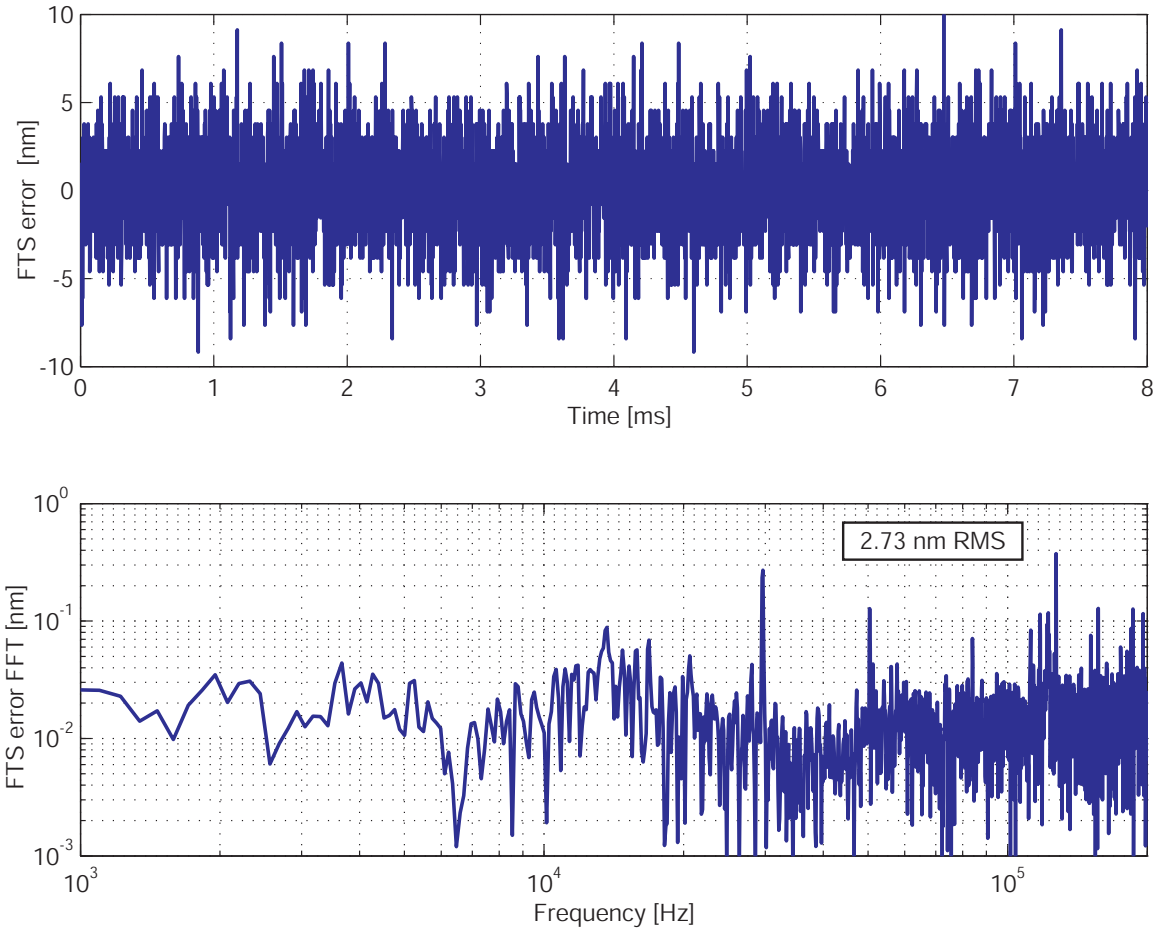


Figure 8-9: UFTS position regulation error when spindle is turned on.

8.3 Flat Surface Turning

During flat-surface cutting of either an aluminum part or a copper part, the regulation error of the UFTS is not noticeably different from that under air cutting. Hence, we can conclude that the cutting force disturbance is negligible when using the combination of our UFTS and the designed diamond tool of a nose radius of 30 μm.

Figure 8-11 shows an interferometric microscope image of a diamond turned aluminum surface. The flatness of this surface is 7 nm RMS over an area of 643×540 μm^2. The sequence of cutting depth is 5 μm, 5 μm, 4 μm, 3 μm, 2 μm, 1 μm, 1 μm, 1 μm, 1 μm, 0.5 μm, and 0.5 μm. For the rough cuts, the feed rate varies from 12 μm/rev to 2.4 μm/rev. The last cut condition is 500 rpm spindle speed with a feed rate

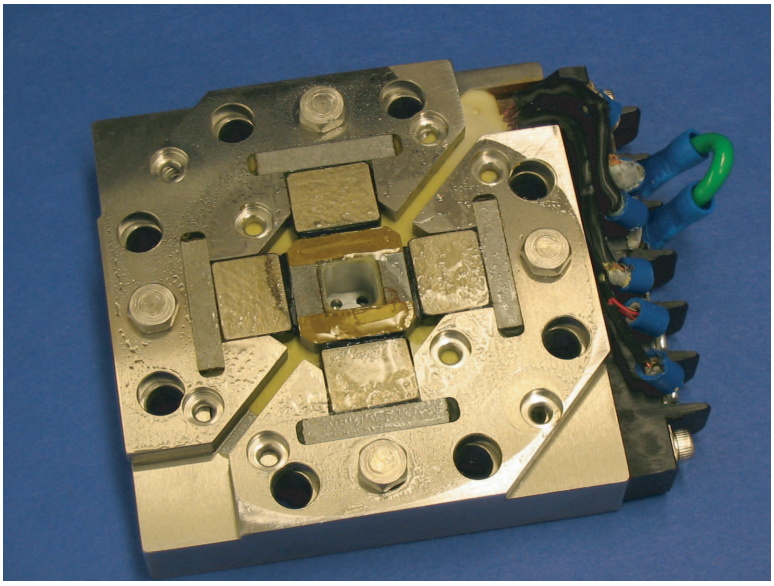


Figure 8-10: Disassembled UFTS shows that the applied mineral oil has entered the UFTS by capillary effect.

of 1.2 μm/rev. The UFTS positioning error Z_f is 3.7 nm RMS, and the Z slide positioning error is 4$\sim$5 nm RMS. The combined error of the diamond tip relative to workpiece is at least 6.4 nm RMS, which is quite close to the measured surface roughness of 7 nm RMS.

Figure 8-12 shows a flat surface turned on an OFHC copper part. The cutting condition is similar to that of the previous aluminum workpiece. This copper surface has a roughness of 8 nm RMS, a little bit higher than that of the aluminum workpiece.

8.4 Spindle Position Estimation

As opposed to flat surface turning, turning a non-rotationally symmetric surface (such as a 1-D or 2-D sine wave surface) requires the spindle position signal for trajectory generation, because the tool motion needs to be synchronized with the spindle rotation. Although a spindle position encoder is installed on the spindle with a resolution of 10,000 counts/rev, there still exist significant position errors for two main reasons

- The encoder updates its output 10,000 times per revolution, which corresponds

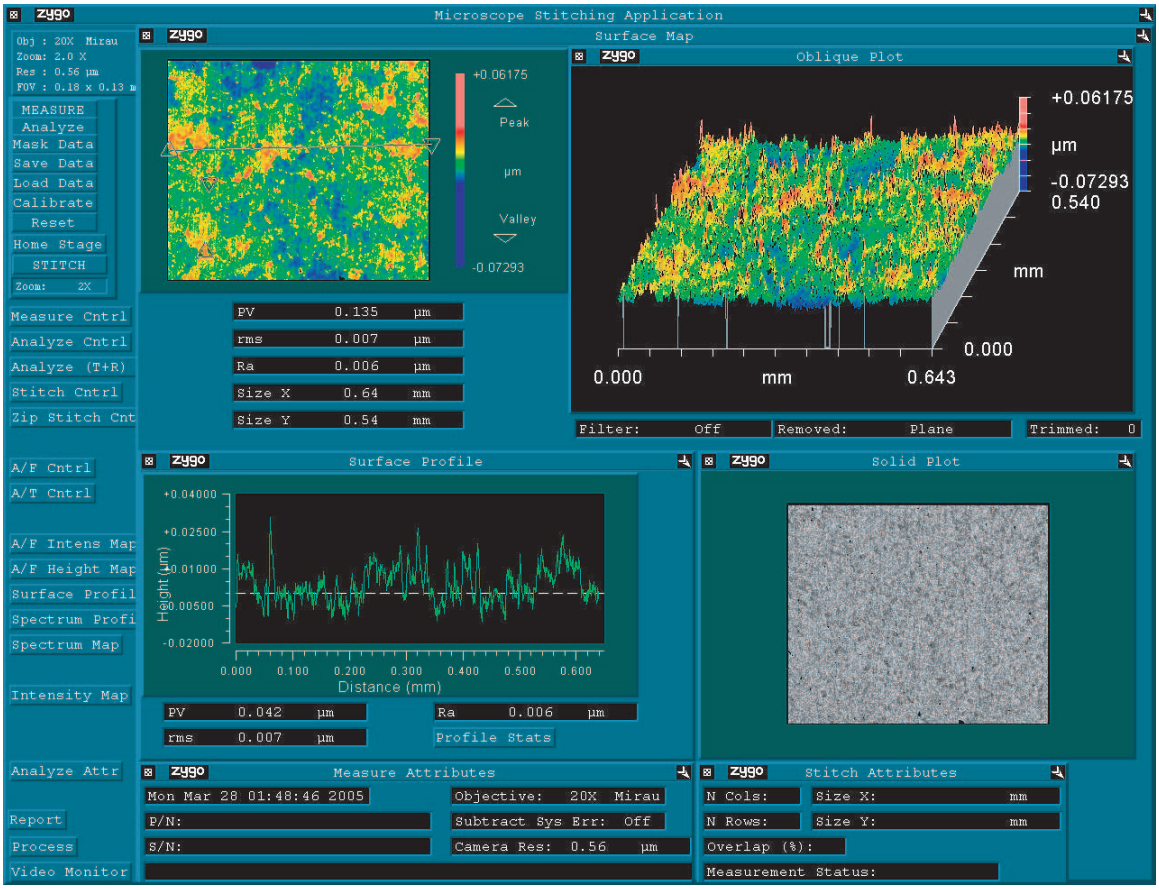


Figure 8-11: UFTS turned flat surface of an aluminum(6061-T6) part. The image is stitched together from many small captures, and the overall area is 643×540 μm^2. The surface roughness is 7 nm RMS. The final cut condition is 500 rpm with 0.5 μm cutting depth and 1.2 μm/rev feed rate.

83.3kHz when the spindle speed is 500 RPM. This frequency is much lower than our computer sampling frequency of 500kHz or higher. As a result, the current value in the encoder quadrature counter doesn't reflect the current spindle position.

- The encoder itself has nonlinearity error, mainly due to misalignment in manufacturing and installation.

As mentioned earlier in Chapter 7, a synchronous filter is a solution to deal with these two problems. However, the synchronous filter bandwidth design is highly dependent on the spindle speed characteristics. This section discusses the spindle motion characterization and its effect on the turning operation of the UFTS.

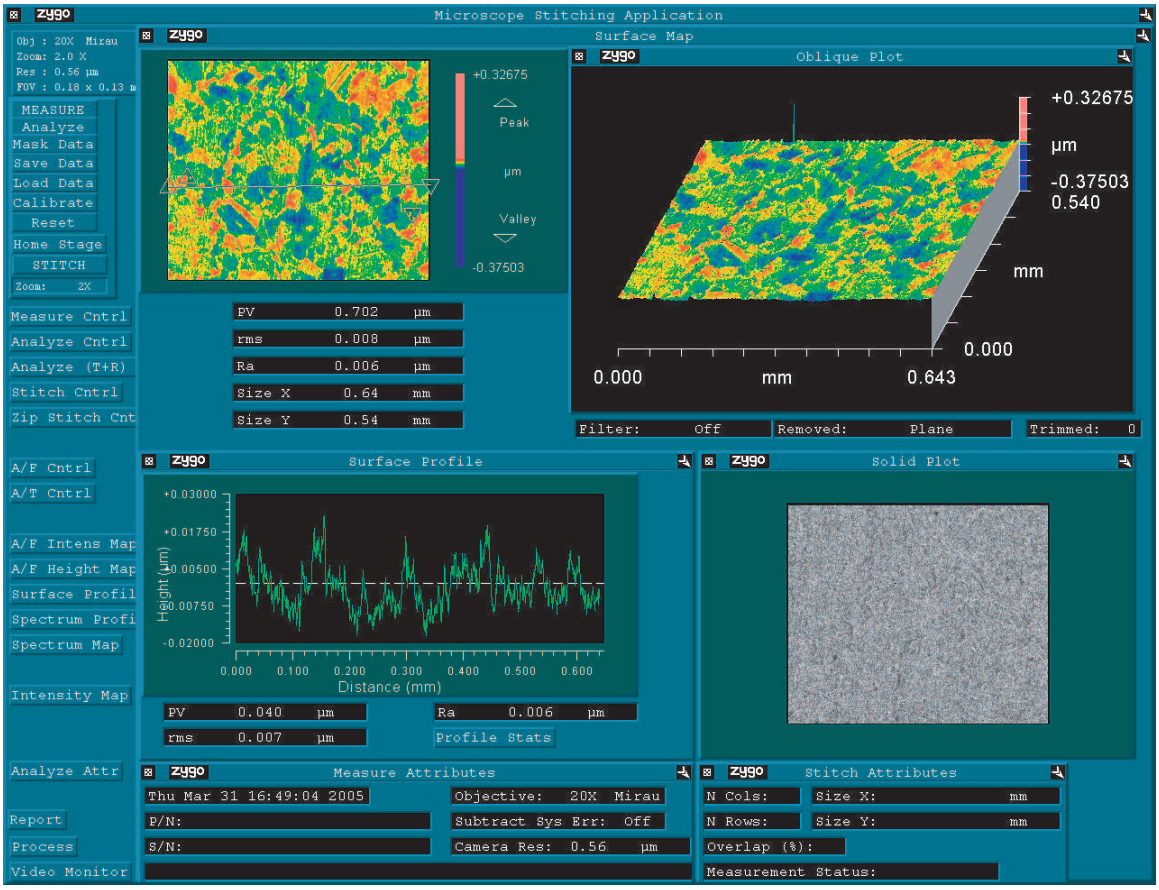
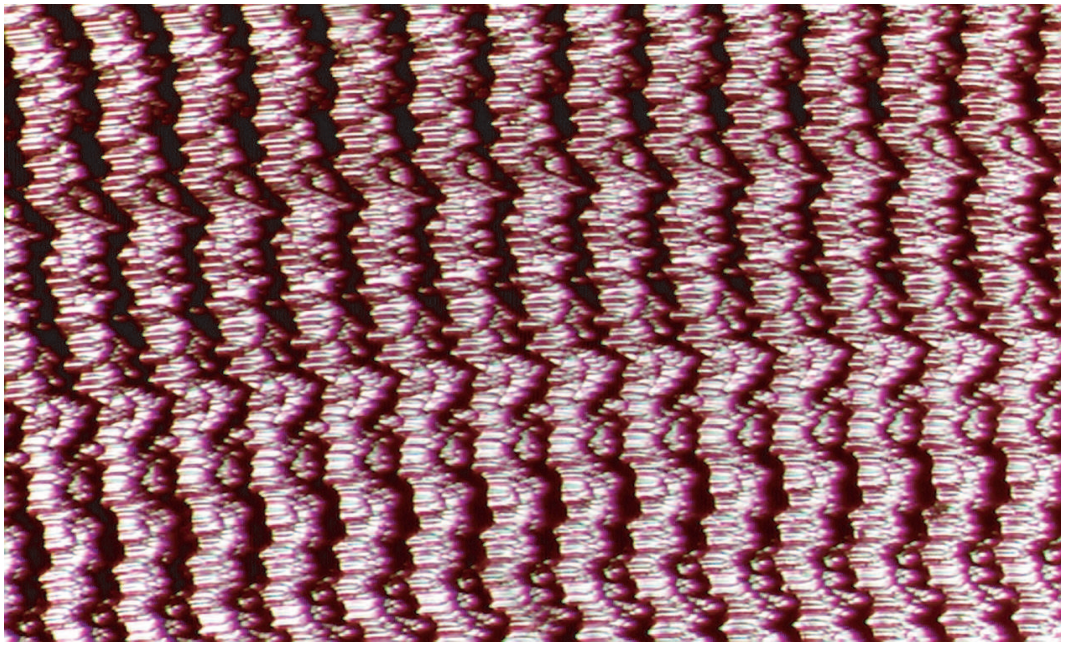
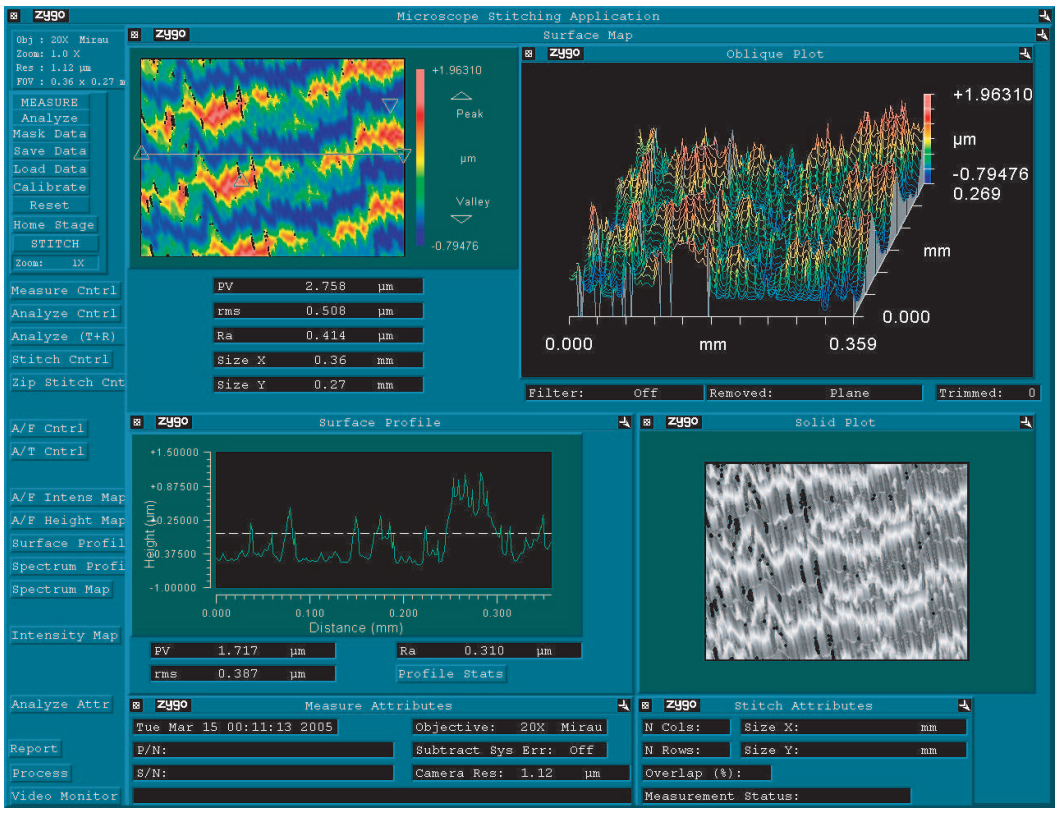


Figure 8-12: UFTS turned flat surface on an OFHC copper part. The image is stitched together from many small captures, and the overall area is 643×540 μm^2. The roughness is 8 nm RMS over this area. The final-cut condition is 500 rpm, 0.5 μm cutting depth, and 1.2 μm/rev feed rate.

First, we assume that the spindle speed is constant under a fixed spindle speed command. Based on this assumption, we design a synchronous filter with a bandwidth of 0.1 Hz. Using a trajectory generated with this synchronous filter at 500 rpm spindle speed, we cut a 1-D copper surface with an expression $Z = \sin(720\theta)$ μm, where θ is the spindle position angle. The resulting surface is shown in Figure 8-13. Obviously, this result is far from what we intend to cut. It shows significant inconsistency (phase error) from revolution to revolution, but the surface profile from one sinusoidal cycle to another in the same revolution track is much better. Based on this observation, we analyzed the speed signal variation in two time scales: (1) revolutionary mean speed representing slow dynamic change from revolution to revolution, (2) momentary speed



(a)



(b)

Figure 8-13: Images of a diamond turned 1-D sinusoidal surface on OFHC copper, using a 0.1 Hz synchronous filter for spindle position estimation. (a) is an image under an optical microscope; (b) is an image under an interferometric microscope.

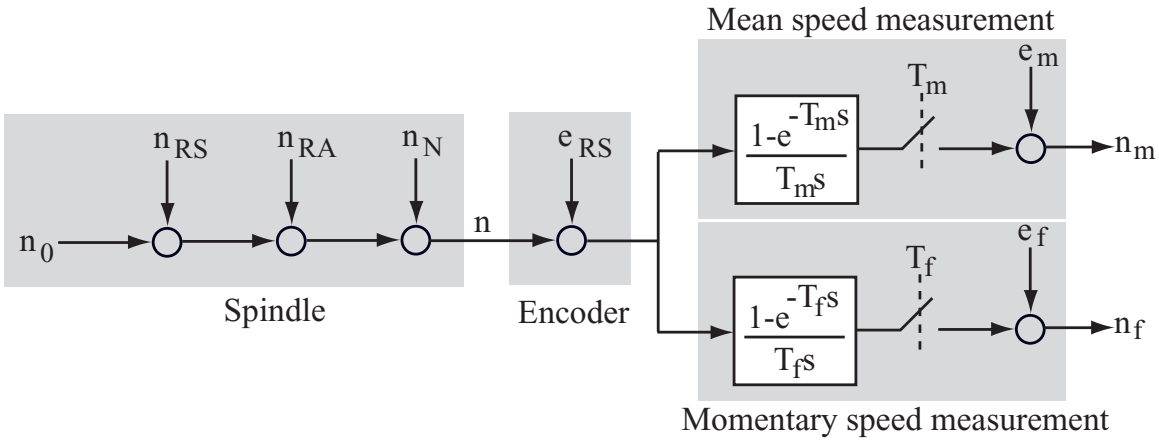


Figure 8-14: Block diagram for spindle speed measurement modelling.

representing dynamics much faster than once revolution.

Figure 8-14 shows a block diagram modeling the spindle speed measurement process. The true spindle speed n [rpm]can be modelled as a constant component n_0 with addition of three AC components: n_{RS} repeatable synchronous speed, n_{RA} repeatable asynchronous speed, and n_N non-repeatable (random) speed. Repeatable synchronous speed is the component that is periodic in spindle angle domain. This component n_{RS} can result from the repeatable synchronous error in the resolver used for spindle speed control, and can also be resulted from synchronous torque ripples. Repeatable asynchronous speed is the component that is periodic in time domain, such as periodic disturbances from PWM amplifiers and AC lines. The signal n_N represents all non-repeatable components, such as the speed variation caused by ground vibration and/or air pressure fluctuation. The spindle encoder is modelled as a summing junction with additional repeatable synchronous speed error e_{RS} caused by the misalignment and/or encoder manufacturing errors, such as non-uniform grating or phase error between gratings A and B.

We performed experiments to quantify two speeds: the mean speed n_m [5] and the momentary speed n_f [6]. The mean speed n_m is defined as the average speed in each

[5]The subscript "m" indicates that this speed measurement only represent the slow dynamics of the spindle.

[6]The subscript "f" indicates that this speed measurement represent the fast dynamics of the spindle.

revolution in order to capture the slow varying spindle speed dynamics

$$n_m = \frac{\int_{t-T_m}^{t} n dt}{T_m} = \frac{1 - e^{T_m s}}{T_m s} n, \tag{8.1}$$

where t is the current time, and T_m is the current spindle revolution period. As shown in Figure 8-14, if captured once per interval t_m the mean speed measurement can be modeled as a zero-order-hold followed by a sampling element with period T_m and measurement counting error e_m. In the mean speed measurement, the zero-order-hold element filters out all the repeatable synchronous elements. As a result, n_m is only related to n_N and n_{RA}.

Similarly, the momentary speed n_f is the average speed between two adjacent spindle encoder counts, and can be expressed as

$$n_f = \frac{\int_{t-T_f}^{t} n dt}{T_f} = \frac{1 - e^{T_f s}}{T_f s} n, \tag{8.2}$$

where T_f is the time duration between two consecutive encoder counts. In the block diagram, if updated red at each encoder count the momentary speed can be modelled as the output of spindle speed n going through a zero-order-hold and a sampling element of a period T_f.

It is nice to sample the spindle speed at each encoder count over a very long time to capture both n_f and n_s in one set of data. However, it is impractical to do so, in large part due to the huge amount of data to be handled. Therefore, we performed two separate experiments to get n_m and n_f respectively.

8.4.1 Mean Spindle Speed Measurement

We use a 500 kHz "soft" clock [7] to count the clock numbers N_m between two adjacent index pulses [8] from the spindle encoder. The spindle rotation period T_m is calculated

[7] By "soft", we mean that a software program running at 500kHz is used instead of a hardware true clock.

[8] The index signal from the spindle encoder output a single pulse every spindle revolution.

as

$$T_m = \frac{N_m}{500 \times 10^3} \text{ sec}, \tag{8.3}$$

and then the average speed n_m over this revolution is

$$n_m = \frac{60}{T_m} = \frac{60 \times 500 \times 10^3}{N_m} = \frac{3 \times 10^7}{N_m} \; rpm. \tag{8.4}$$

Figure 8-15 shows the mean spindle speed measurement result at a nominal speed of 500 rpm and its time domain FFT. At this speed, the measurement error e_m

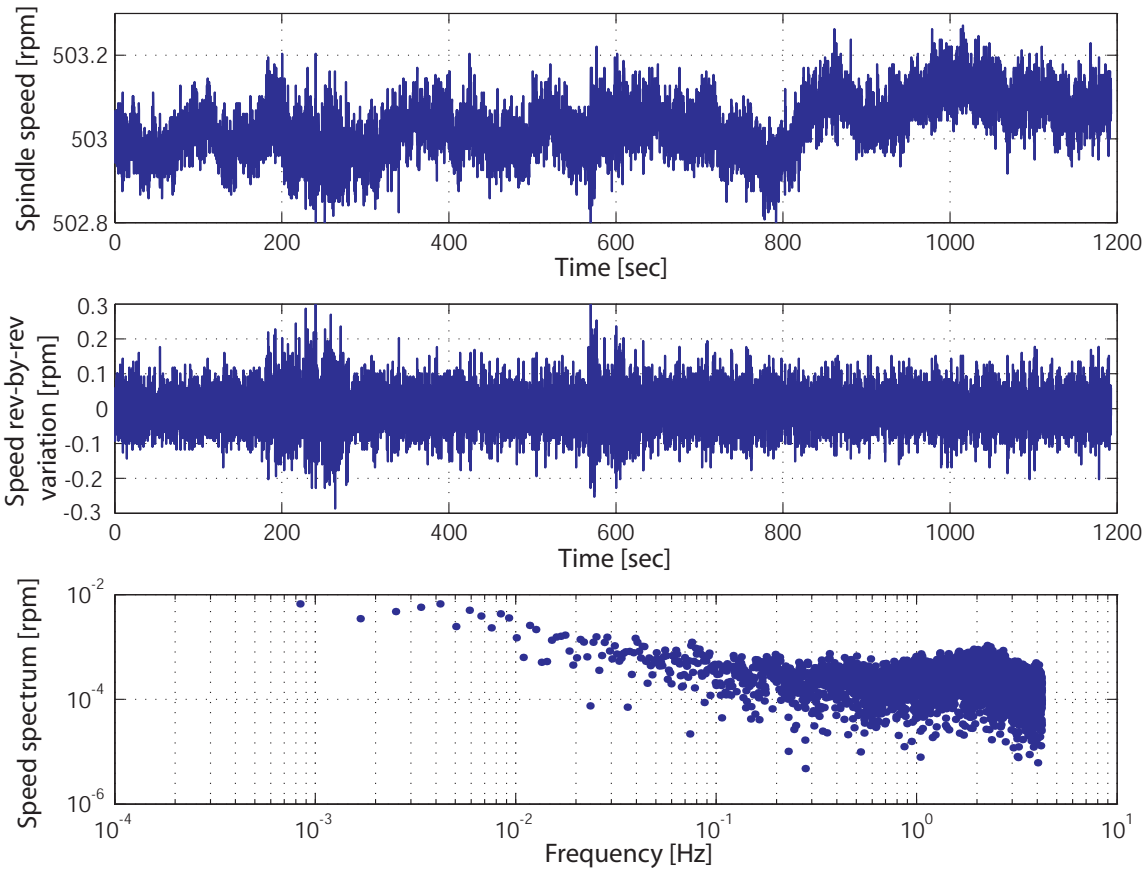


Figure 8-15: Mean spindle speed measurement results at a nominal speed of 500 rpm.

resulting from the clock counting error $\delta N_m (\leq 1)$ can be estimated as

$$\begin{aligned} |e_m| &\leq \left| \delta N_m \frac{d}{dN_m} n_m \right| \\ &\leq \left| \frac{3 \times 10^7}{N_m^2} \right| \end{aligned}$$

$$
\begin{aligned}
&= \left| \frac{n_m^2}{3 \times 10^7} \right| \\
&= \frac{500^2}{3 \times 10^7} \\
&= 0.01 \quad [rpm]. \qquad (8.5)
\end{aligned}
$$

Hence, the error resulting from a 500kHz soft clock is negligible in comparison with the speed fluctuation in Figure 8-15, and thus this measurement reflects the true speed.

In this result, the mean spindle speed has 0.5 rpm peak-to-valley variation over 20 minutes duration, which is consistent with the 0.1% speed regulation specification for the spindle controller. The speed spectrum shows that the spindle speed is nearly a white noise for frequencies less than 4 Hz and thus the low frequency dynamics are dominated by non-repeatable speed fluctuation n_N.

Also the revolution-to-revolution variation (the difference between the mean speed in the the current revolution and that in the previous revolution) ranges from +0.3 to -0.3 rpm. Using a spindle speed estimator of 0.1 Hz bandwidth, only mean speed over 80 revolutions can be estimated. As a result, the revolution-to-revolution spindle speed variation will contribute significantly to the spindle speed estimation error. Assuming that there is a Δn_m=0.2 rpm mean speed estimation error, for the 1-D trajectory of $Z = \sin(720\theta)$ μm the accumulative phase error over one revolution can be calculated as

$$
\begin{aligned}
|\Delta\phi| &= 720 \times \frac{\Delta n_m}{n_m} \times 360 \\
&= \frac{720 \times 0.2 \times 360}{500} \\
&= 103.7 \quad \text{deg}, \qquad (8.6)
\end{aligned}
$$

which sufficiently explains the inconsistency between two consecutive tracks in Figure 8-13.

8.4.2 Momentary Spindle Speed Measurement

A 100 MHz clock in the ThunderStorm computer is used to count the clock numbers N_f between two consecutive output counts (10,000 per revolution) from the spindle encoder. The time duration T_f between two consecutive encoder pulses is then given by

$$T_f = \frac{N_f}{100 \times 10^6} \text{ [sec]}. \tag{8.7}$$

Thus the momentary speed n_f for the current clock count between encoder counts is

$$n_f = \frac{60}{T_f \times 10^4} = \frac{6 \times 10^5}{N_f} \text{ [rpm]}, \tag{8.8}$$

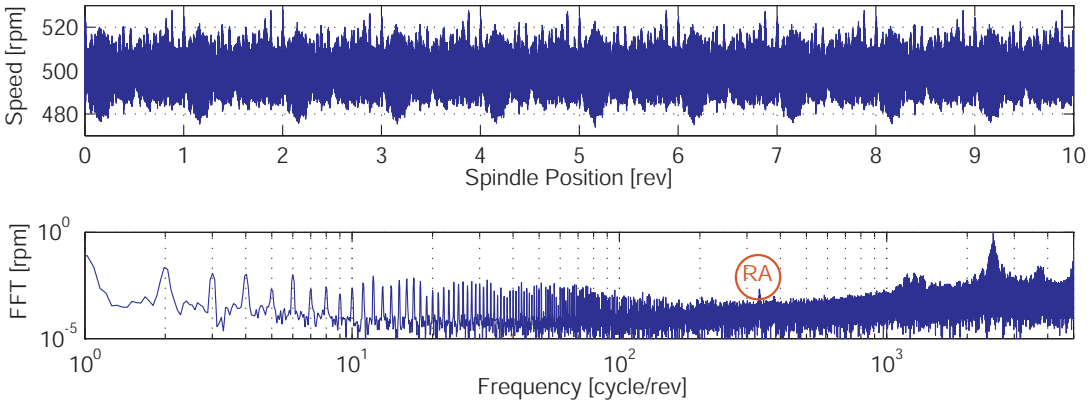
where 10^4 is the spindle encoder count number per revolution.

Figure 8-16 shows experimental momentary speed measurement results based on (8.8) at nominal speeds of 500 rpm and 400 rpm. The speed measurement error e_f at 500 rpm resulting from the clock discrete counting error $\Delta N_f (<1)$ can be estimated as

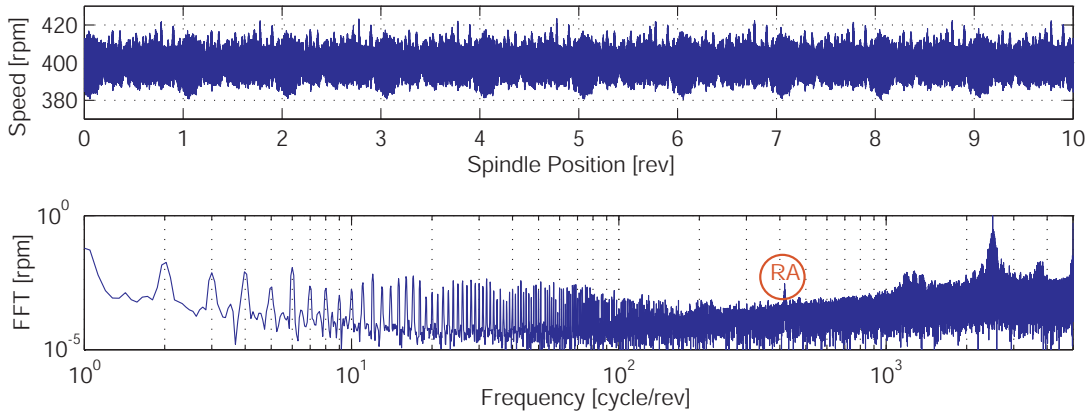
$$\begin{aligned} |e_f| &\leq \left| \Delta N_f \frac{d}{dN_f} n_f \right| \\ &\leq \left| \frac{6 \times 10^5}{N_f^2} \right| \\ &\leq \frac{n_f^2}{6 \times 10^5} \\ &= \frac{500^2}{6 \times 10^5} \\ &= 0.4 \text{ [rpm]}. \end{aligned} \tag{8.9}$$

Similarly, the error e_f is even less at nominal speed of 400 rpm. Because this error is far below the speed fluctuation in Figure 8-16, this clock-related measurement measurement error is negligible.

Figure 8-16 also shows the associated spindle angle domain FFT, which clearly indicates that there are many peaks at harmonic spindle rotation frequencies. The FFT of course can be calculated in time domain. However, it is found that the high



(a) Spindle momentary speed waveforms at norminal speed of 500 [rpm]



(b) Spindle momentary speed waveforms at norminal speed of 400 [rpm]

Figure 8-16: Angle-domain momentary speed measurement in time and FFT-domains. Data captured at each encoder count edge, over the course of 10 spindle revolutions.

frequency harmonics (higher than spindle rotation frequency) at these two speeds are consistent in the angle-domain FFT rather than in the time-domain FFT. Therefore these peaks are either n_{RS} or e_{RS}. There also exist peaks at 2.780 kHz, which are not coincident in the angle-domain but coincident in the time-domain (333.657 cycle/rev at 500.062 rpm and 416.965 cycle/rev at 400.167 rpm) marked by red circles and labels of "RA" in Figure 8-16. This frequency is exactly one-sixth of the spindle PWM amplifier frequency (16.67 kHz). Clearly these components are related to the spindle PWM amplifier. But I am not clear about the details of how this speed component is generated. As these repeatable asynchronous components are relatively small (0.002∼0.003 rpm), in the following analysis and spindle estimator design we assume $n_{RA} = 0$.

For the repeatable synchronous components, we cannot tell which parts are from n_{RS} or e_{RS} without installing a second spindle position measurement of higher accuracy to act as a reference. Since the encoder has 10,000 counts per revolution, the 2500 and 5000 cycle/rev components are related to the spindle encoder grating errors or read-head alignment errors.

8.4.3 Dynamical Correction Mapping

The effects of position estimation errors can be understood with reference to Figure 8-17. The shaded area is the turned surface, and the part is both spinning and feeding in the indicated directions. The tool tip is moving in the direction pointing into the paper. Let's look at two cutting tracks labelled as A and B. If there is non-repeatable position estimation error, radial inconsistency will be generated as exemplified in Figure 8-13. For instance, there will exist phase shifts between A1 and B1, and between A2 and B2. If there are repeatable synchronous errors in the spindle position estimation, azimuthal inconsistency will result. For instance, the distance between A1 and A2 will be stretched or compressed relative to the desired shape.

Therefore, the spindle position estimator design requires trade-off between two requirements: (1) remove the encoder error component e_{RS} to maintain azimuthal consistency, (2) track the spindle speed n low frequency variation to maintain radial

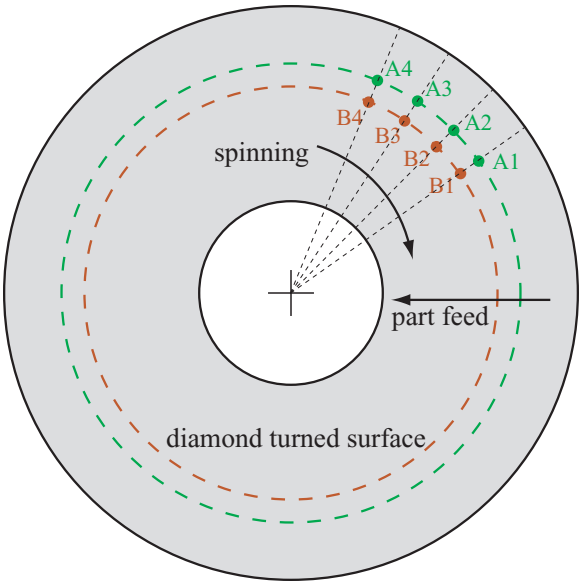


Figure 8-17: Tool trajectory on the machined surface.

consistency and n high frequency variation to maintain azimuthal consistency.

One way to remove the encoder error e_{RS} is to calibrate the spindle encoder by another encoder with higher resolution as a reference. Within the scope of this thesis, we lacked time and equipments to explore this method, so we cannot tell e_{RS} from n_{RS}. In order to move forward, we assume that $n_{RS} = 0$, i.e., the spindle speed is near constant within one revolution. This is a reasonable assumption for such a high precision air-bearing spindle. Under this assumption, we present a method called dynamical correction mapping to remove e_{RS} in the position estimation.

The idea of dynamical correction mapping is to first collect the momentary speed measurement data synchronously in M continuous revolutions $n_{i,j}$, which is the momentary speed in the i-th count after encoder index pulse of j-th revolution, and then based on this data to derive the encoder nonlinearity characteristics.

As data in multiple revolution are collected, we can remove the non-repeatable and repeatable asynchronous components by averaging the data synchronously

$$\overline{n}_i = \sum_{j=1}^{M} n_{i,j}. \tag{8.10}$$

This new variable $\overline{n}_i$ contains only the encoder repeatable synchronous error. Based

on this result, when the encoder outputs the i-th count after index pulse, we can calculate the true count as

$$i^* = \frac{\sum_{j=1}^{i} \frac{1}{\overline{n}_j}}{\sum_{j=1}^{10000} \frac{1}{\overline{n}_j}}. \tag{8.11}$$

Figure 8-18 shows both the synchronously captured momentary speed in the first revolution $n_{i,1}$ for $1 \leq i \leq 10^4$ and the synchronously averaged speed $\overline{n}_i$ for $1 \leq i \leq 10^4$ over $M = 83$ revolutions. These two plots share very similar pattern, indicating that the repeatable synchronous components dominate repeatable asynchronous components and non-repeatable components. Also in the figure, the calculated correction map i^* and the correction value $i^* - i$ are plotted versus the spindle encoder output i.

8.4.4 Position Estimator Design

The overall spindle position estimator design is shown in Figure 8-19. First the encoder nonlinearity error is removed by correction mapping (either calibrated according to a higher-accuracy sensor or using dynamic correction mapping as described above), then a synchronous filter is designed to remove up-sampling (from T_f to T_c) error, where T_c is the sampling time of ThunderStorm computer control algorithm. The synchronous filter bandwidth should be not only high enough to cover most spectrum component in n but also should be as low as possible to filter off the aliasing-induced position noise. From the mean speed measurement in Figure 8-15, we can see that the non-repeatable spindle fluctuation can be up to 4Hz at a nominal speed of 500 rpm. So we design the bandwidth of the synchronous filter at 70 Hz, which is more than 10 times higher than the spindle speed fluctuation.

In the turning experiment described with more details in the next sections, we do both the cutting test with dynamic mapping turned on and the test with dynamic mapping turned off. When the dynamic mapping is turned on, the UFTS motion for a spatial harmonics can be heard as a pure tone, in contrast to a trembling tone when the dynamic mapping is turned off. This demonstrates that the dynamic mapping method can effectively remove the synchronous repeatable component. However, we

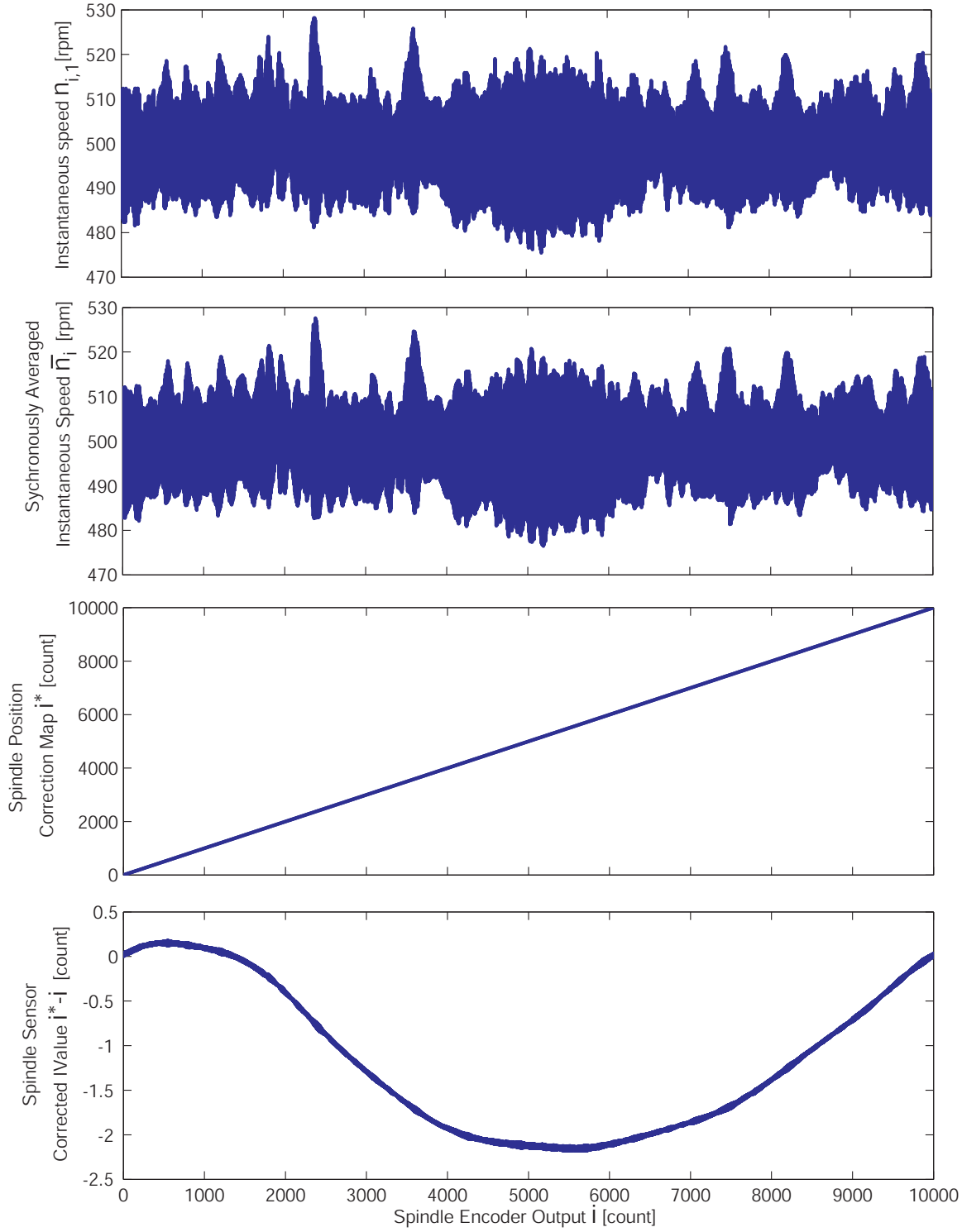


Figure 8-18: Spindle encoder dynamic correction mapping result.

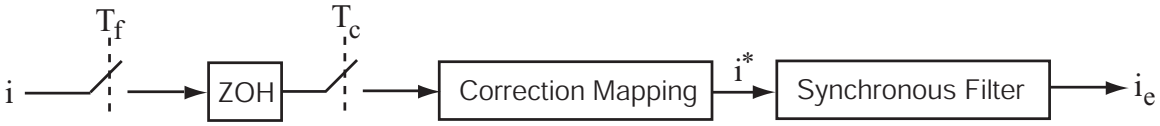


Figure 8-19: Spindle position estimator design.

do not yet know whether the synchronous repeatable components are from the speed fluctuation or from the encoder error. If the former case is true, then the dynamic mapping will introduce significant synchronous position error and make the turned pattern stretched or compressed in the azimuthal direction. If the encoder error contributes mainly to the synchronous repeatable components in the encoder output, then the dynamic mapping can effectively remove them and generate surfaces with more consistency in the azimuthal direction. Whether the mapping is turned on or off, there is no noticeable difference in the UFTS tracking error.

8.5 1-D Sinusoidal Surface Turning

We turned a 1-D sinusoidal surface $z = \sin(720\theta)$ at 500 rpm spindle speed, which corresponds to a 6kHz sine wave with a peak acceleration of ± 150 G. The cutting sequence is 1 μm depth at 12 μm/rev, 1 μm depth at 12 μm/rev, 1 μm depth at 12 μm/rev, 1 μm depth at 6 μm/rev, 0.5 μm depth at 2.4 μm/rev, and 0.5 μm depth at 1.2 μm/rev. The AFC control parameters Gain/Phase/Frequency for this 1-D cutting experiment, are listed in Table 8.1.

Table 8.1: AFC control parameters in 1-D turning.

Harmonic index	Harmonic frequency [Hz]	Gain	Phase [deg]
1	6000	200	0
2	12000	200	100
3	18000	200	200

In the cutting experiment, the tracking error on the UFTS servo loop is varying from 4.8 nm RMS to 5.7 nm RMS as shown in Figure 8-20. The tracking error on the Z slide is 9 nm RMS as shown in Figure 8-21, and the tracking error on the X slide is 3 nm RMS as shown in Figure 8-22.

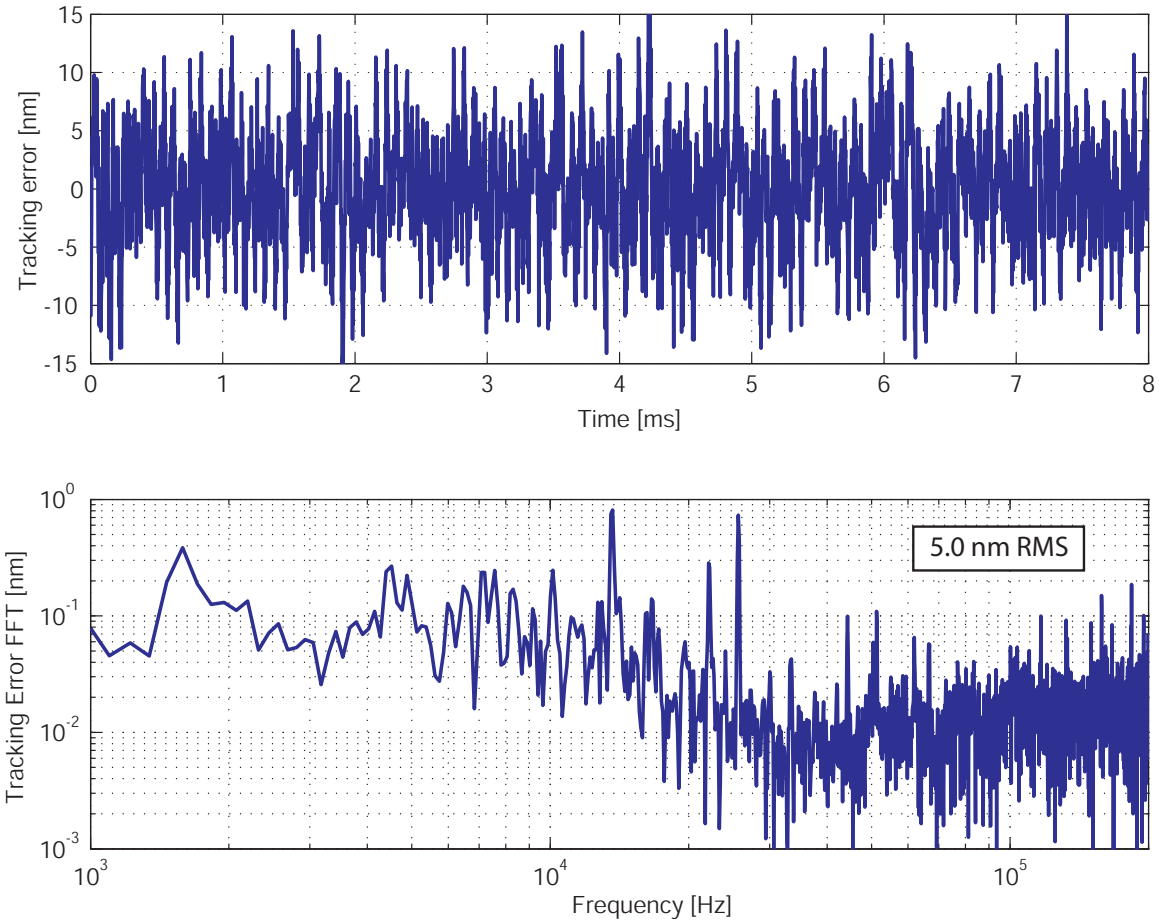


Figure 8-20: Tracking error of the UFTS control loop in the 1-D surface diamond turning experiment. In comparison with Figure 8-9, the tracking error has degraded from 2.73 nm RMS holding error to 5.0 nm RMS tracking error in turning a 1-D sine wave.

Figure 8-23 shows a Zeiss microscope image of a turned 1-D aluminum surface. By increasing the synchronous filter bandwidth from 0.1Hz to 70Hz, the spindle speed estimator can track the spindle speed fluctuation. As a result, the problem of asynchronous tool path exhibited in Figure 8-13 has been eliminated. Figure 8-24 shows an interferometric microscope image of this 1-D surface on aluminum. The peak-to-valley value shown in the figure is 2.22 μm, which the difference between the highest point and lowest point and not the true peak-to-valley value of the sine wave. The tile of the sine wave can be seen on the surface profile. Because of this tilt, the peak-to-valley value of 2.22 μ does not match well with the intended 2 μm. Figure 8-25 indicates that the roughness of this surface is 10 nm RMS by measuring along a peak

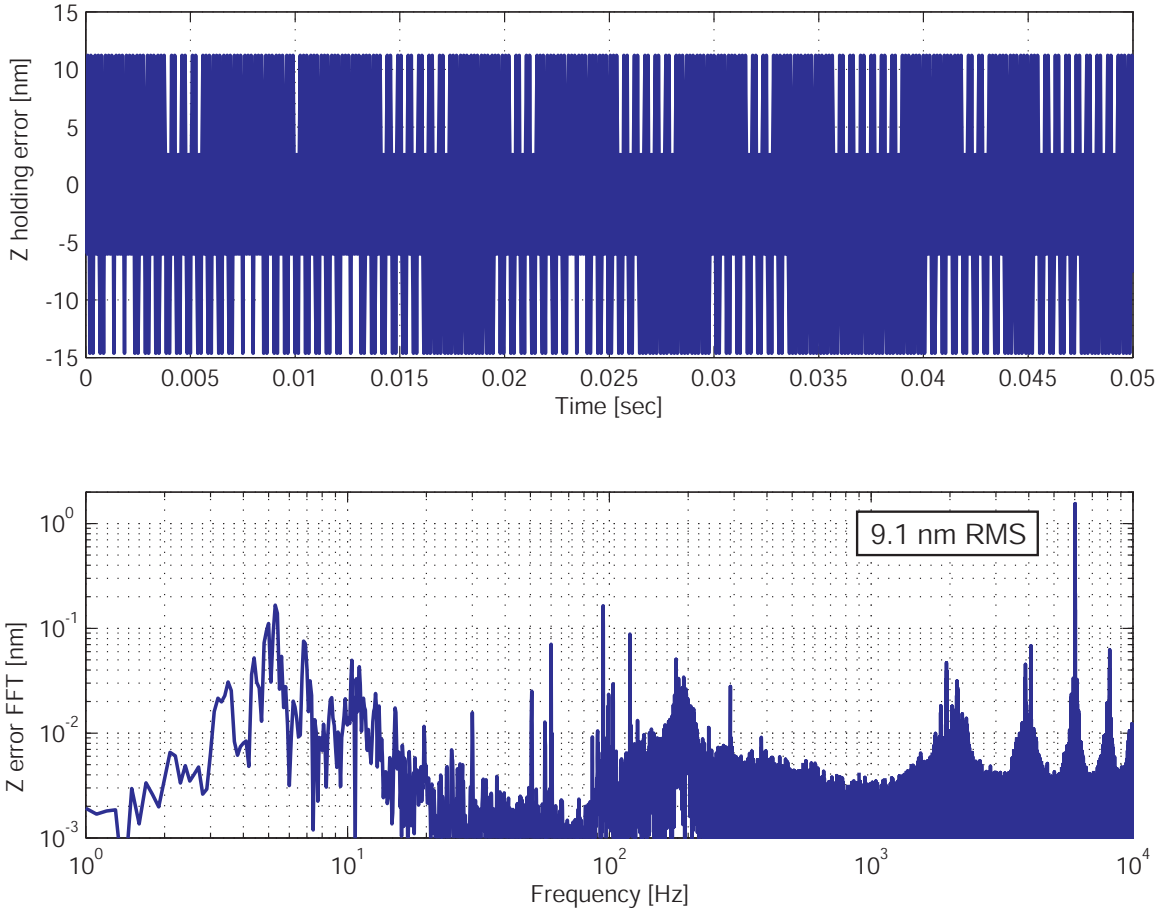


Figure 8-21: Regulation errors of the Z slide control loop in the 1-D surface diamond turning experiment. In comparison with Figure 8-6, the regulation error has increased from 3.7nm RMS when no FTS motion to 9.1 nm RMS error when turning 1-D sine wave, mainly due to the 6 kHz component generated by the reaction force of the UFTS motion. In this figure, we also see significant components at 2kHz, 4kHz, and 8kHz. These error components are most likely aliased components from the motion at the 18 kHz, 24 kHz, and 12 kHz respectively, as this Z slide position signal is sampled at 20kHz in the DS1103 board.

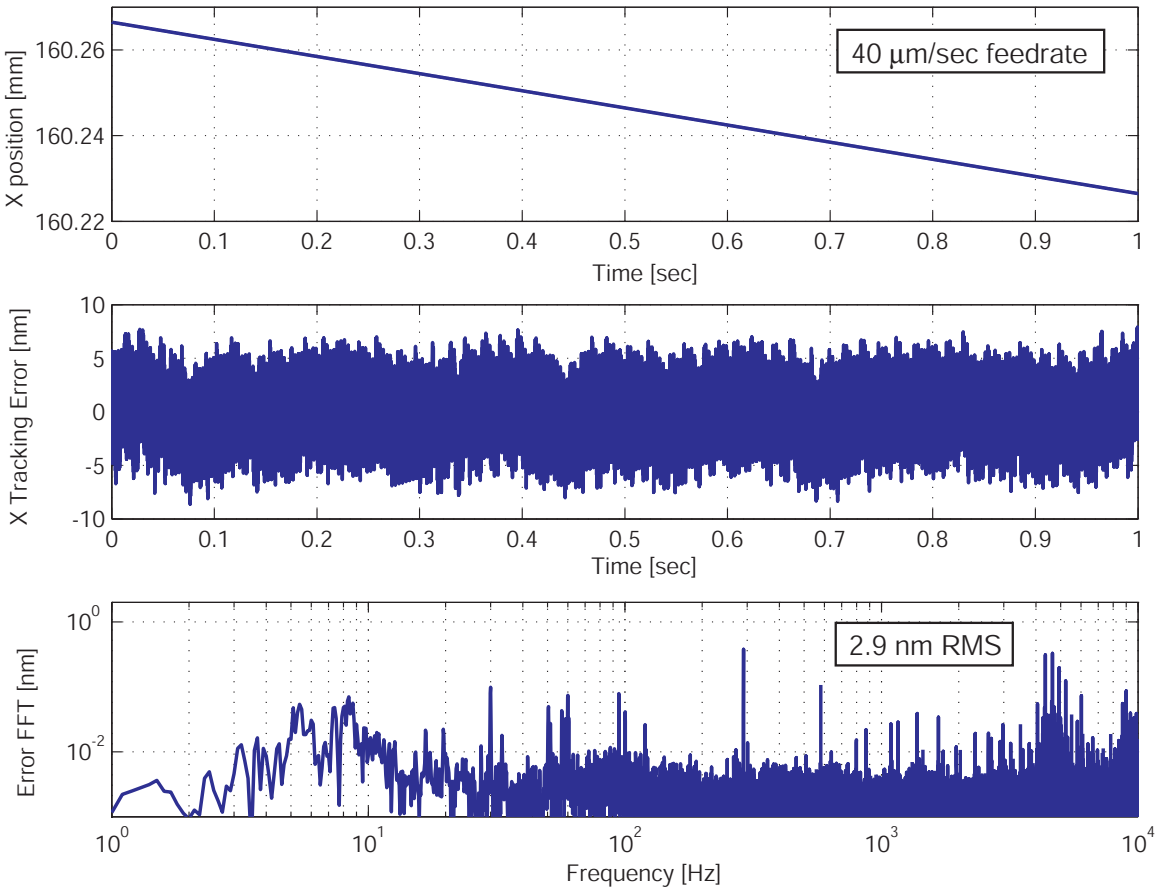


Figure 8-22: Tracking errors of the X slide control loop in the 1-D surface diamond turning experiment. The X slide motion is a ramp with a slope of 40 μm/sec. By extracting an ideal ramp from the captured position signal, we can get the tracking error as shown in the middle. In comparison with Figure 8-7, which is under no cutting condition, there is no noticeable difference, except that the error pattern is not limited to two discrete values. Therefore, the UFTS motion has no disturbance on the X slide.

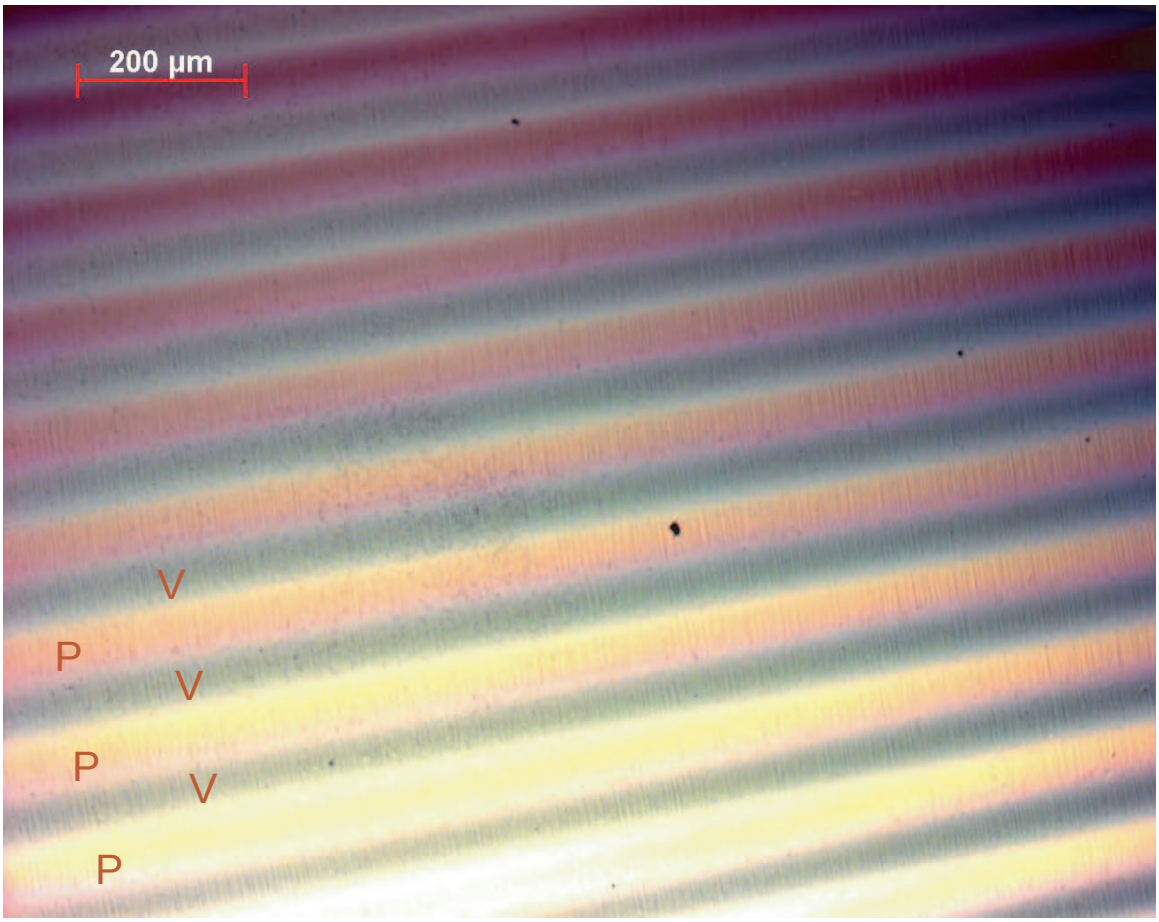


Figure 8-23: Zeiss microscope image of a turned 1D aluminum surface. The bright parts are peaks indicated by letter P, and the dark parts indicated by letter V correspond to the valleys of the sine wave.

line of length 520 μ. This result matches the Z slide holding error of 9.1 nm RMS in Figure 8-21.

For copper parts, we obtained similar results as for aluminum parts. Figure 8-26 shows a Zeiss microscope image of a turned 1-D copper surface. Figure 8-27 shows an interferometric microscope image of a 1-D surface, which indicates that the peak-to-valley value is 2.246 μm. Again, the peak-to-valley value is the difference between the highest point and the lowest point in the profile and does not reflect the true peak-to-valley value of the sine wave. The difference between this value and the intended 2 μm mainly comes from the tilt of this part set on the microscope. Figure 8-28 shows that the roughness of this surface is 9 nm by measuring along a valley line, which matches the Z slide holding error of 9.1 nm RMS in Figure 8-21.

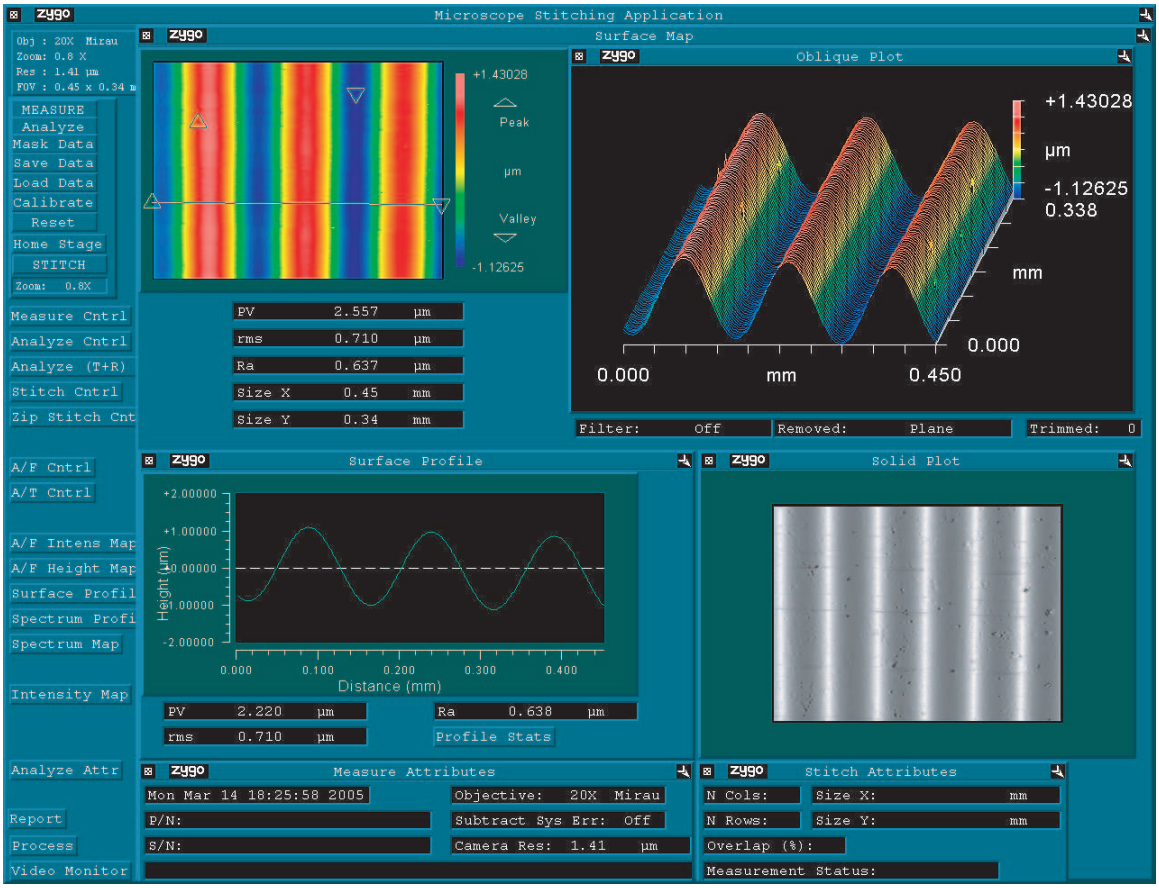


Figure 8-24: Zygo interferometric image of a turned 1-D aluminum surface. The image size is 450 μm long and 338 μm wide. The image lateral resolution is 1.41 μm.

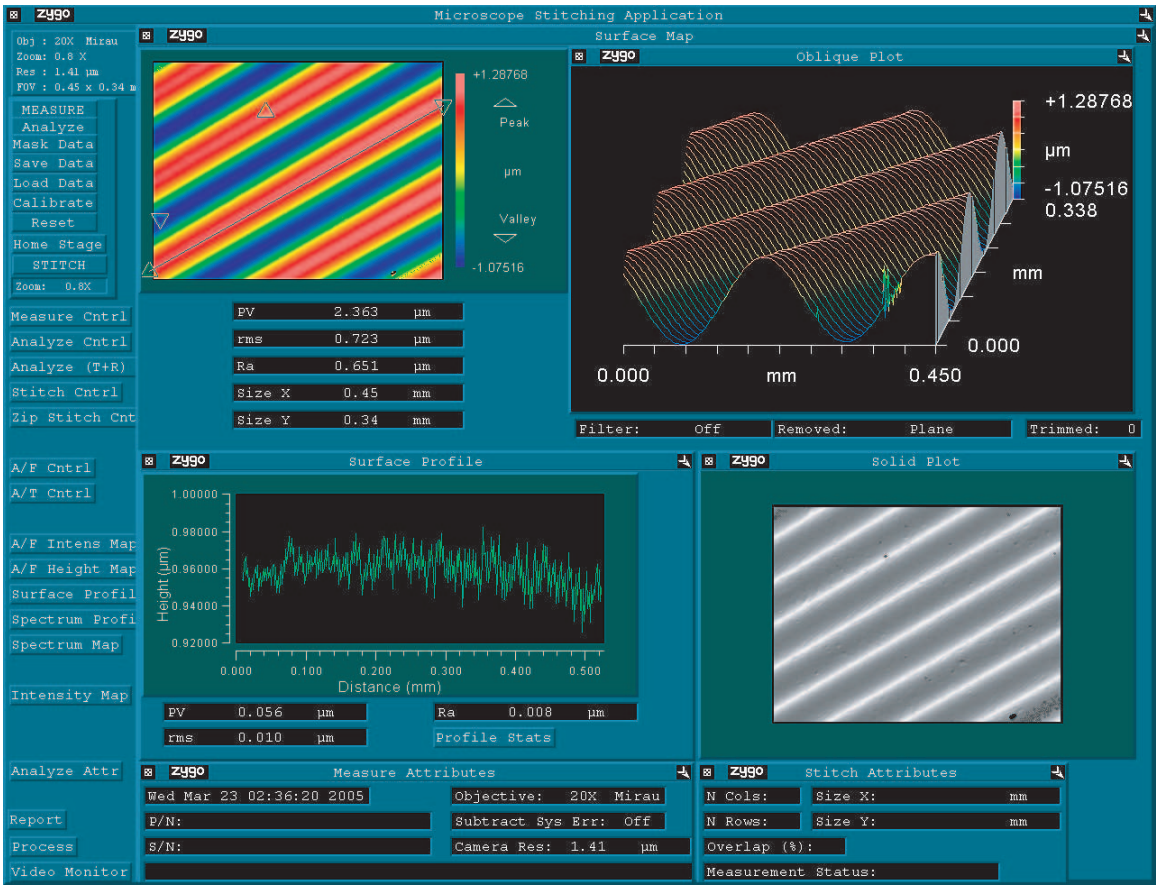


Figure 8-25: Zygo interferometric image of a turned 1-D aluminum surface. The image size is 450 μm long and 338 μm wide. The image lateral resolution is 1.41 μm. This figure shows that the roughness is 10 nm RMS when measured along a peak line.

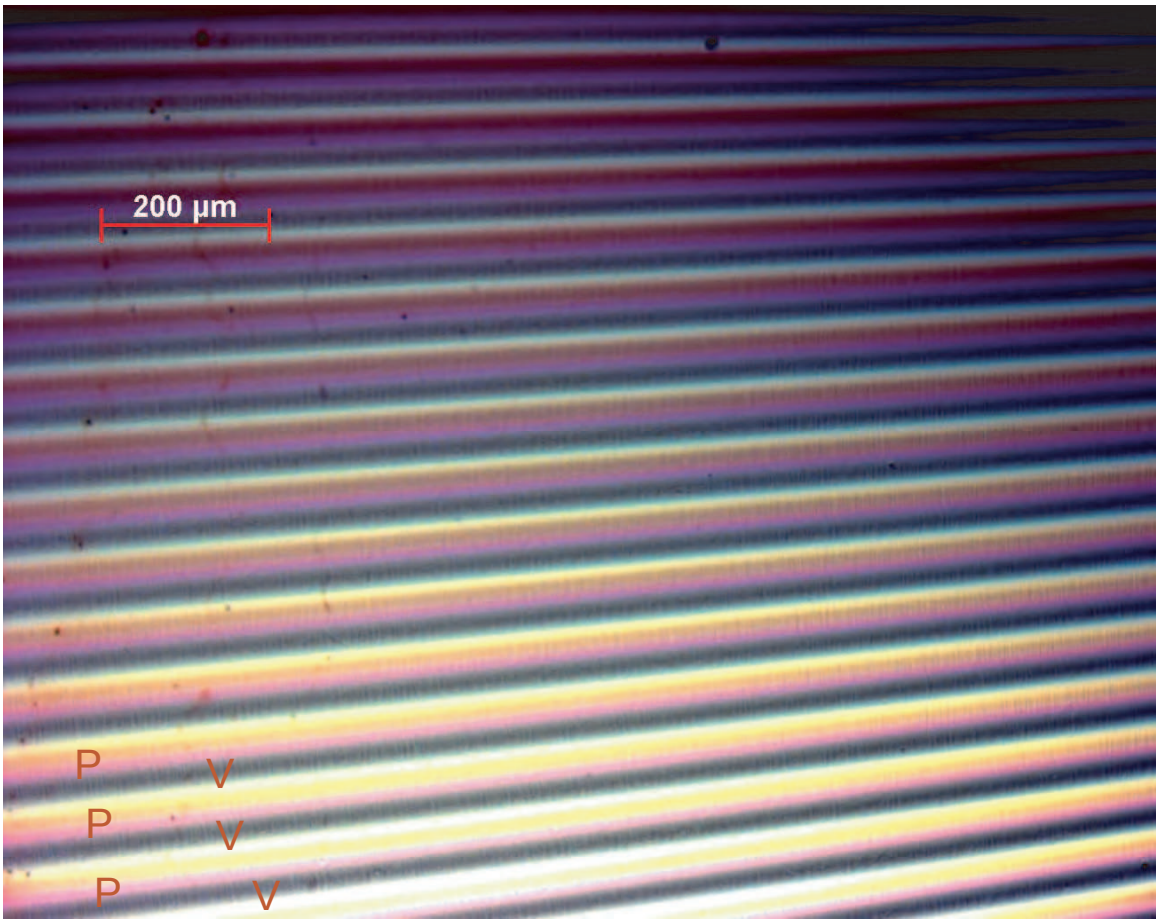


Figure 8-26: Zeiss microscope image of a diamond turned 1-D copper surface. Peaks are indicated by the letter P, and valleys are indicated by the letter V.

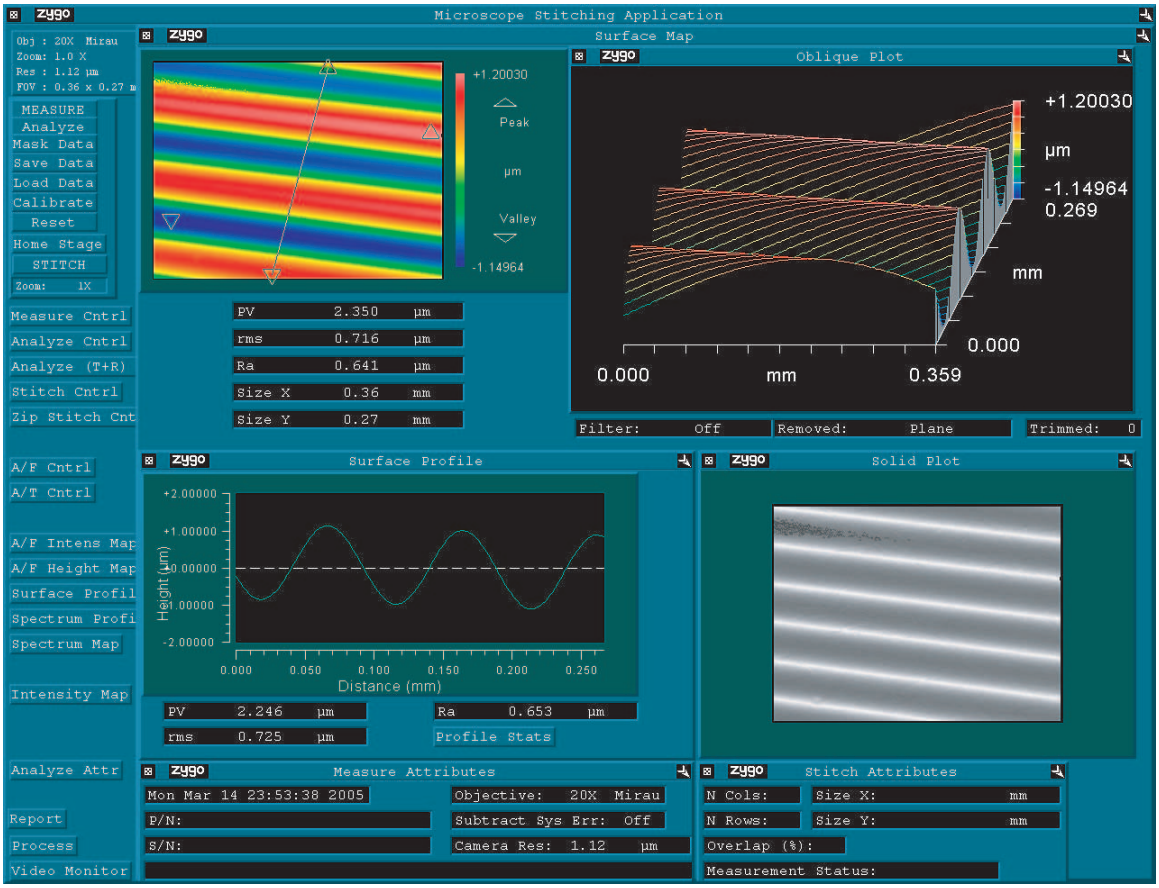


Figure 8-27: Zygo image of a diamond turned 1-D copper surface. The image size is 360 μm long and 270 μm wide. The image lateral resolution is 1.12 μm.

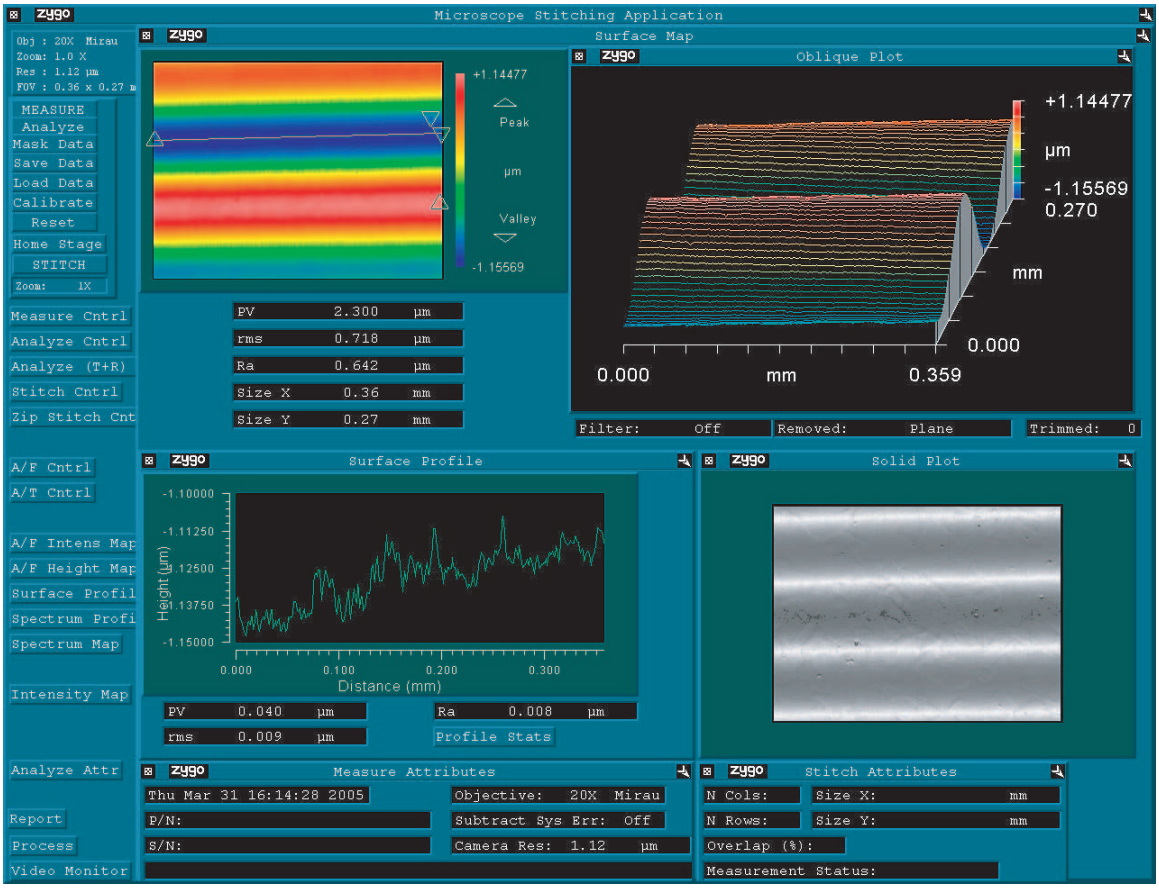


Figure 8-28: Zygo image of a diamond turned 1-D surface. The image size is 360 μm long and 270 μm wide. The image lateral resolution is 1.12 μm. This figure shows that the surface roughness along a groove line is 9 nm RMS.

8.6 2-D Sinusoidal Surface Turning

The 2-D surface that we cut in the turning experiment is expressed as $z = \cos(720\theta) \times \cos(2\pi \times 6 \times x)$, where θ is the spindle position angle and x is the X slide position in units of mm. This surface has a radial period of 166.67 μm and an azimuthal period of 0.5°. We cut this 2-D surface at 500 rpm spindle speed, which corresponds to a 6 kHz sinusoidal wave with a peak acceleration of ±150 G acceleration. At a feed rate of 20 μm/sec (2.4 μm/rev), the tool trajectory is a 6 kHz sine wave modulated by a 0.12 Hz sine wave.

We first turned the part to make a flat surface, and then did a sequence of 2-D cutting: 1 μm depth at feed rate 6 μm/rev, 1 μm depth at 2.4 μm/rev, 1 μm depth at 2.4 μm/rev, 1 μm depth 2.4 μm/rev, 1 μm depth at 2.4 μm/rev, and 0.5 μm depth at 1.2 μm/rev. The AFC control parameters Gain/Phase/Frequency for this 2-D cutting experiment are listed in Table 8.2:

Table 8.2: AFC control parameters in 2-D turning.

Harmonic index	Harmonic frequency [Hz]	Gain	Phase [deg]
1	6000	400	0
2	12000	200	100
3	18000	200	200

During the cutting experiment, the tracking error on the UFTS servo loop varies from 4.0 nm RMS to 5.4 nm RMS as shown in Figure 8-29. The tracking error on the Z slide is 8.3 nm RMS as shown in Figure 8-30, and the tracking error on X slide is 3 nm RMS as shown Figure 8-31.

Figure 8-32, Figure 8-33, and Figure 8-34 show the interferometric microscope images of a turned 2-D copper surface. It can bee seen that the peak-to-valley value is 2.041 μm. The finished surface is very close to the intended 2 μm peak-to-valley vallue. These figures also show that the radial pitch is about 165 μm.

Figure 8-35 and Figure 8-36 are the interferometric microscope images of a 2-D aluminum surface. It can bee seen that the the peak-to-valley amplitude is 2.012 μm and the radial pitch is 165 μm.

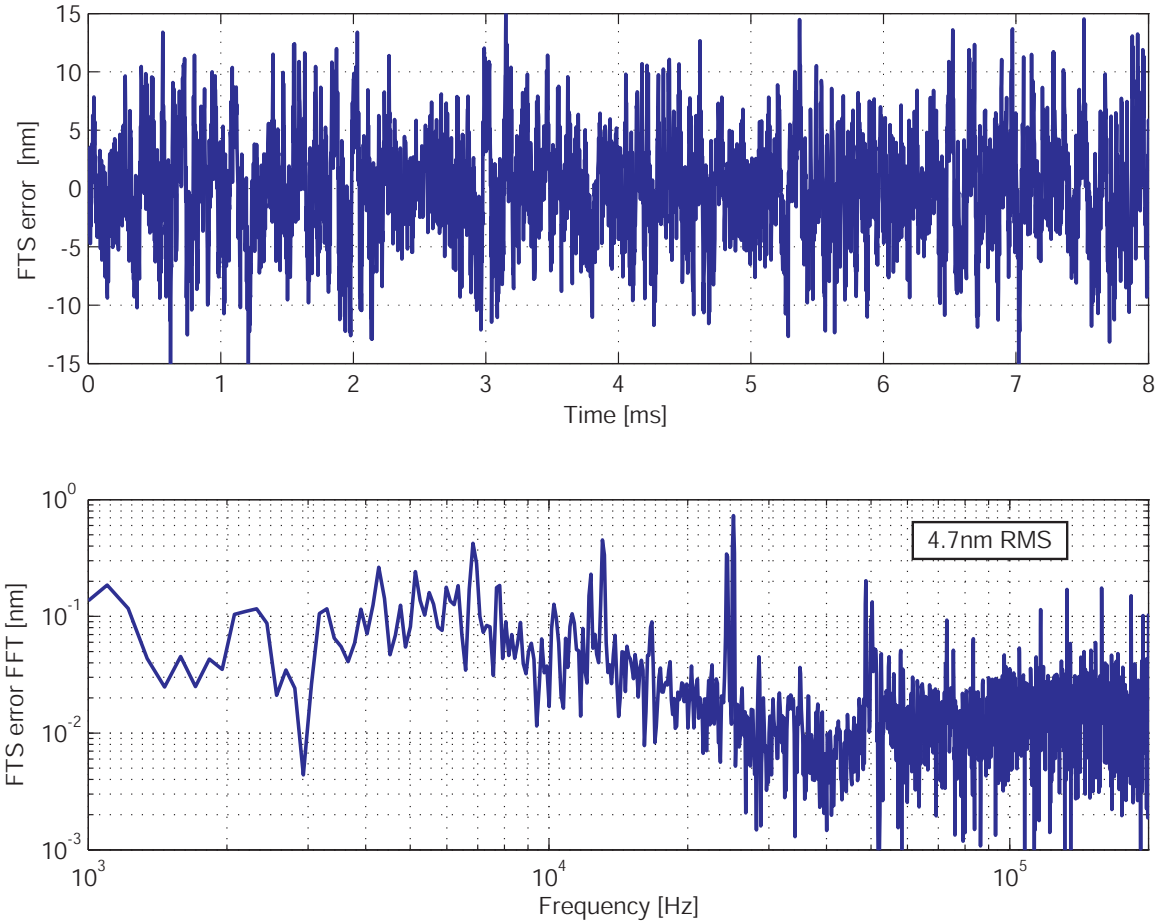


Figure 8-29: Tracking error of the UFTS control loop in a 2-D surface diamond turning experiment. In comparison with the 1-D cutting in Figure 8-20, there is no noticeable degradation on the UFTS tracking performance.

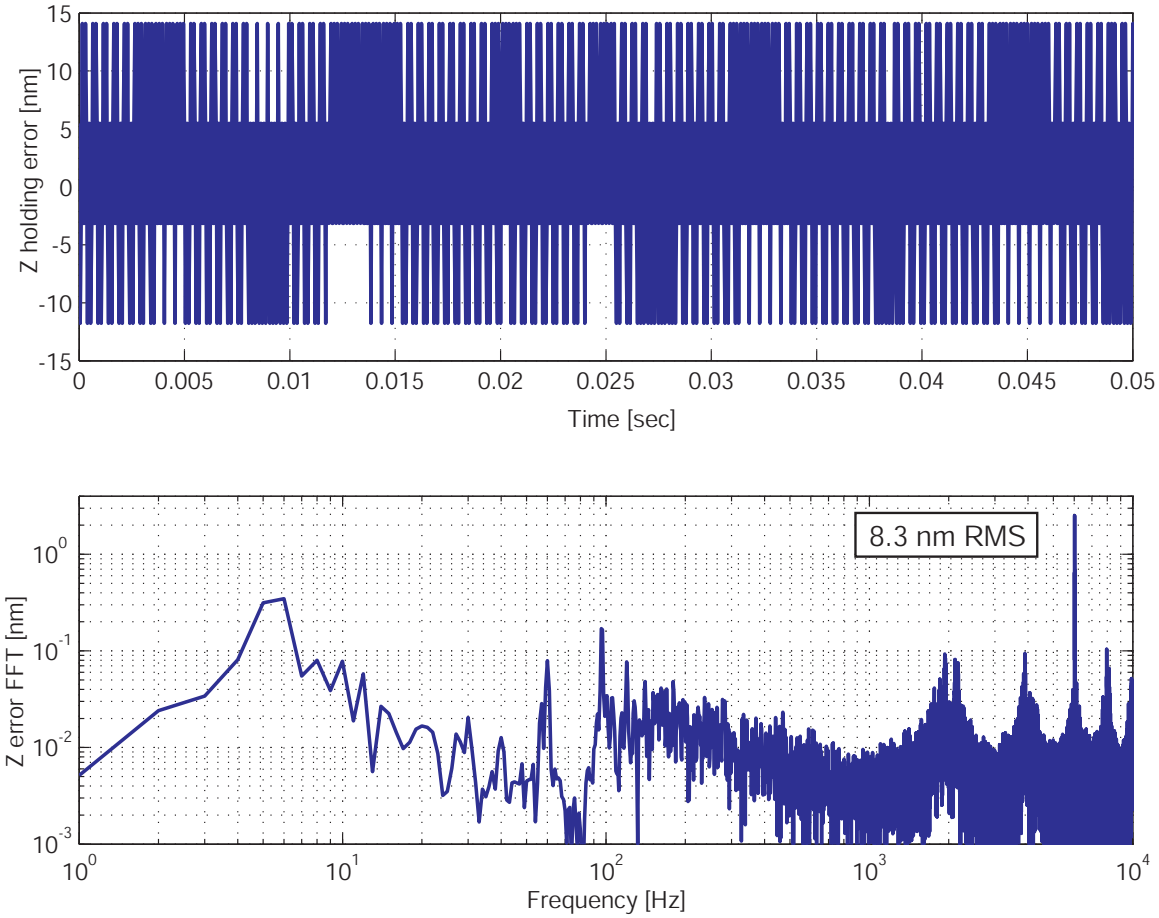


Figure 8-30: Regulation error of the Z slide control loop in the 2-D surface diamond turning experiment. In comparison with Figure 8-6, the regulation accuracy has degraded from 3.7nm RMS without FTS motion to 8.3 nm RMS error when turning a 2-D sine wave. The main cause of this increase is the 6 kHz component generated by the reaction force of the UFTS motion. In this figure, we also see significant components at 2kHz, 4kHz, and 8kHz. These error components are most likely aliased components from the motion at the 18 kHz, 24 kHz, and 12 kHz respectively, as this Z slide position signal is sampled at 20kHz in the DS1103 board.

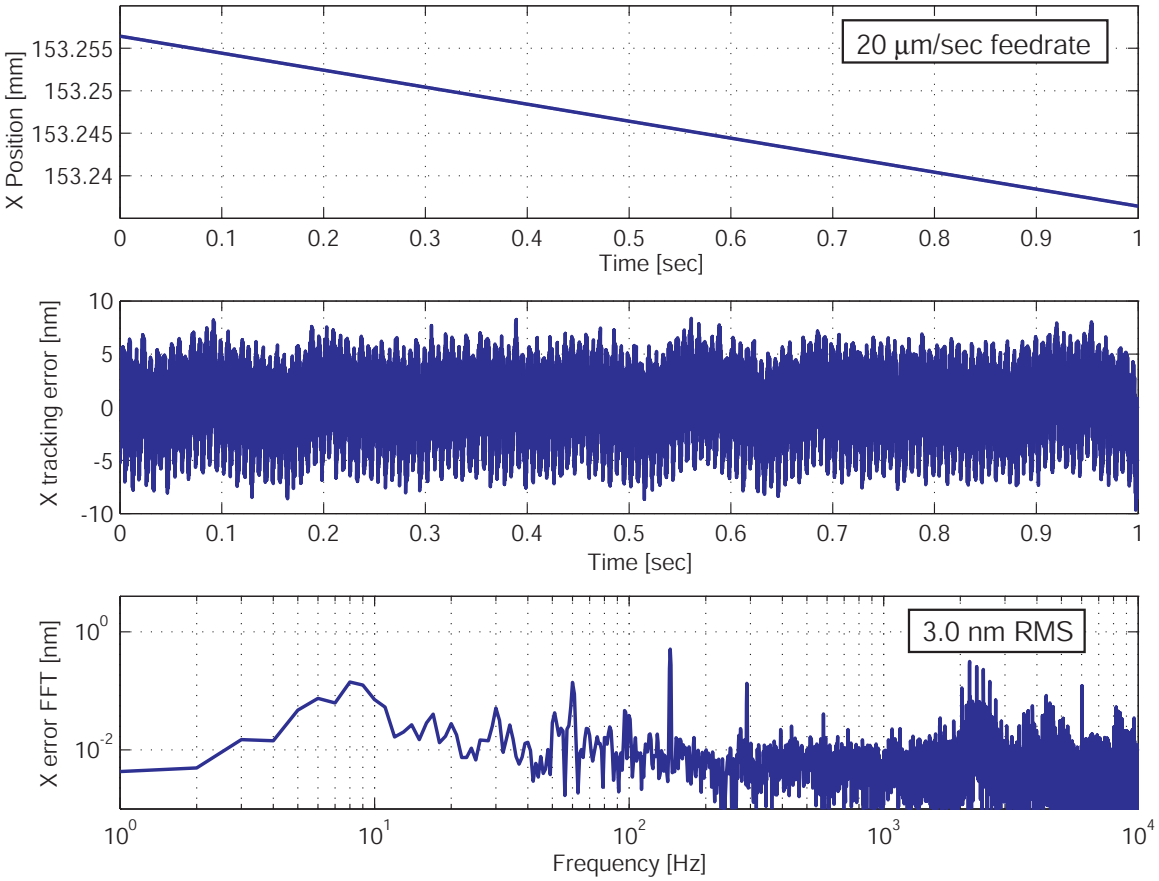


Figure 8-31: Tracking error of the X slide control loop in the 2-D surface diamond turning experiment. As the X slide is the feeding slide, its position signal is a ramp with a slope of 20 μm per second. By extracting an ideal ramp from this captured position signal, we can get the tracking error shown in the middle graph. In comparison with Figure 8-7 under no cutting condition, there is no noticeable difference, except that the error pattern is no longer limited to two discrete values. Therefore, we can conclude that the UFTS motion has no disturbance on the X slide.

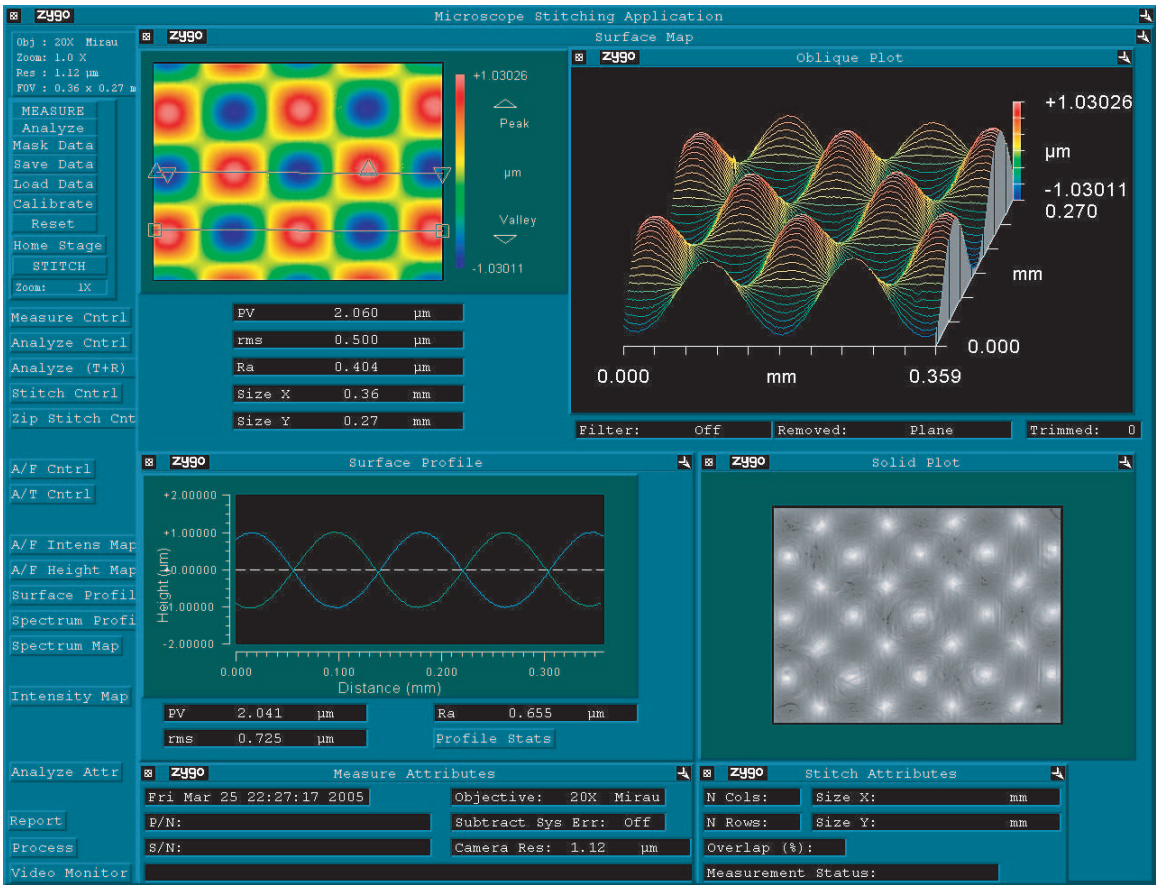


Figure 8-32: Zygo image of a diamond-turned 2-D copper surface. The imaged area is 360 μm long and 270 μm wide, with a lateral resolution of 1.12 μm. The horizontal axis is the radial direction, and the vertical axis is the azimuthal direction. The surface profile shows that along the radial direction the pitch is about 165 μm and the peak-to-valley amplitude is 2.041 μm.

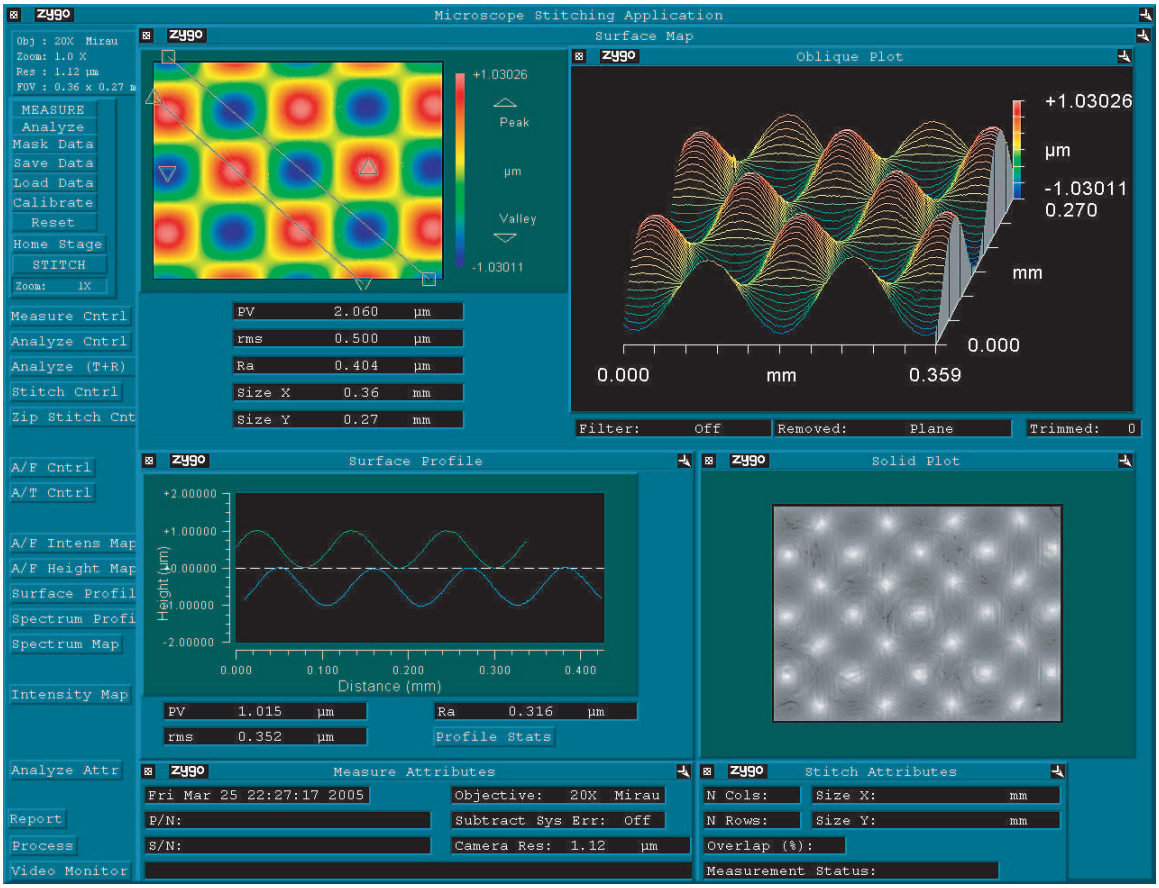


Figure 8-33: Zygo image of a diamond-turned 2-D copper surface. The surface profile shows that along the diagonal direction the plots are sine waves with a peak-to-valley amplitude of 1.015 μm.

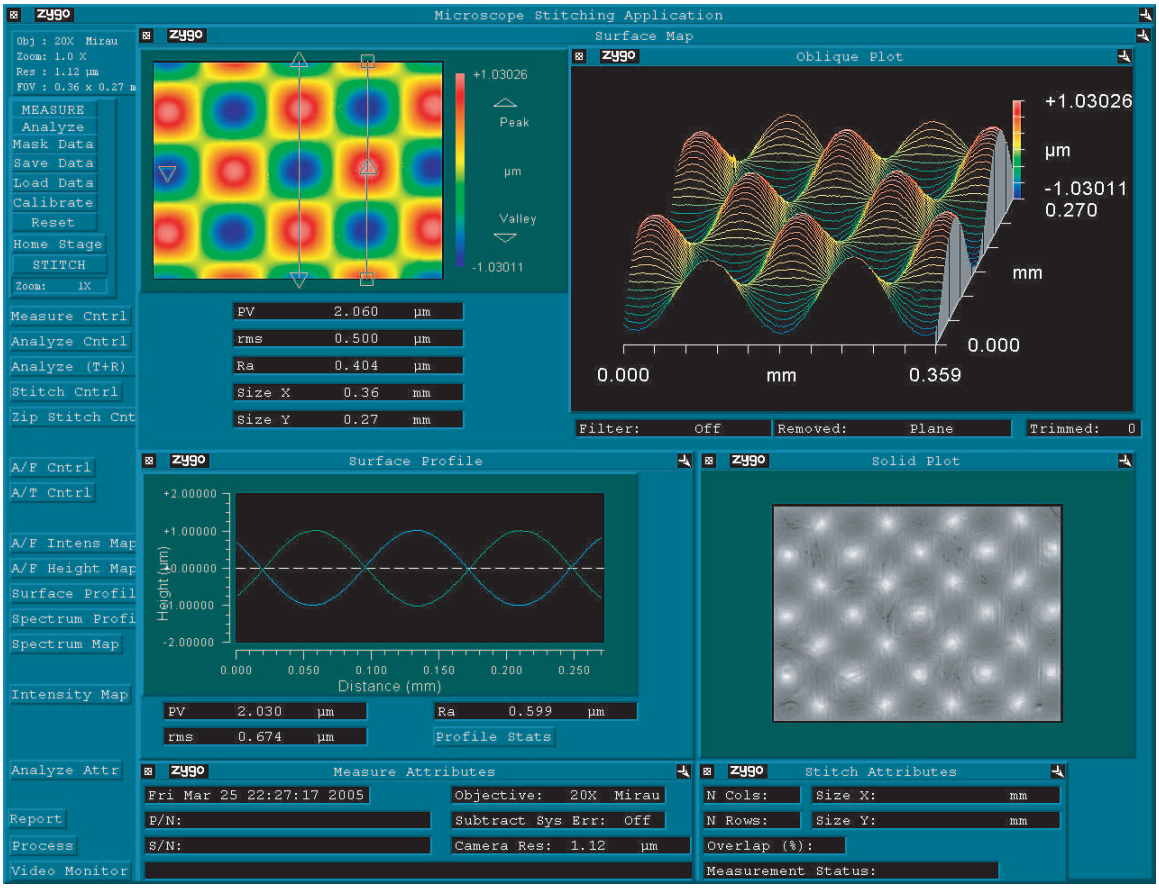


Figure 8-34: Zygo image of a diamond-turned 2-D copper surface. The surface profile shows that along the azimuthal direction the pitch is about 135 μm and the peak-to-valley amplitude is 2.030 μm.

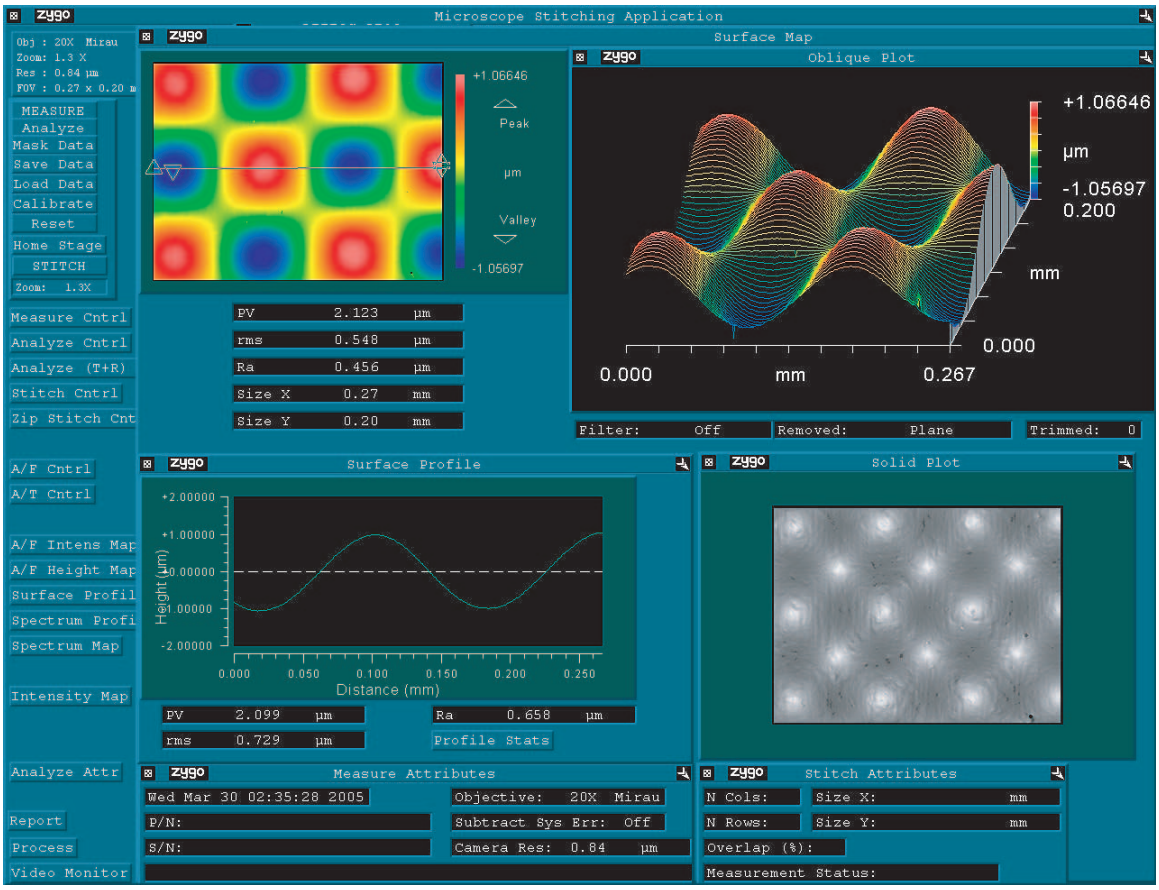


Figure 8-35: Zygo image of a diamond-turned 2-D aluminum surface. The imaged area is 267 μm long and 200 μm wide, with a lateral resolution of 0.84 μm. The horizontal axis is the radial direction, and the vertical axis is the azimuthal direction. The surface profile shows that the radial pitch is about 165 μm and the peak-to-valley amplitude is 2.1 μm.

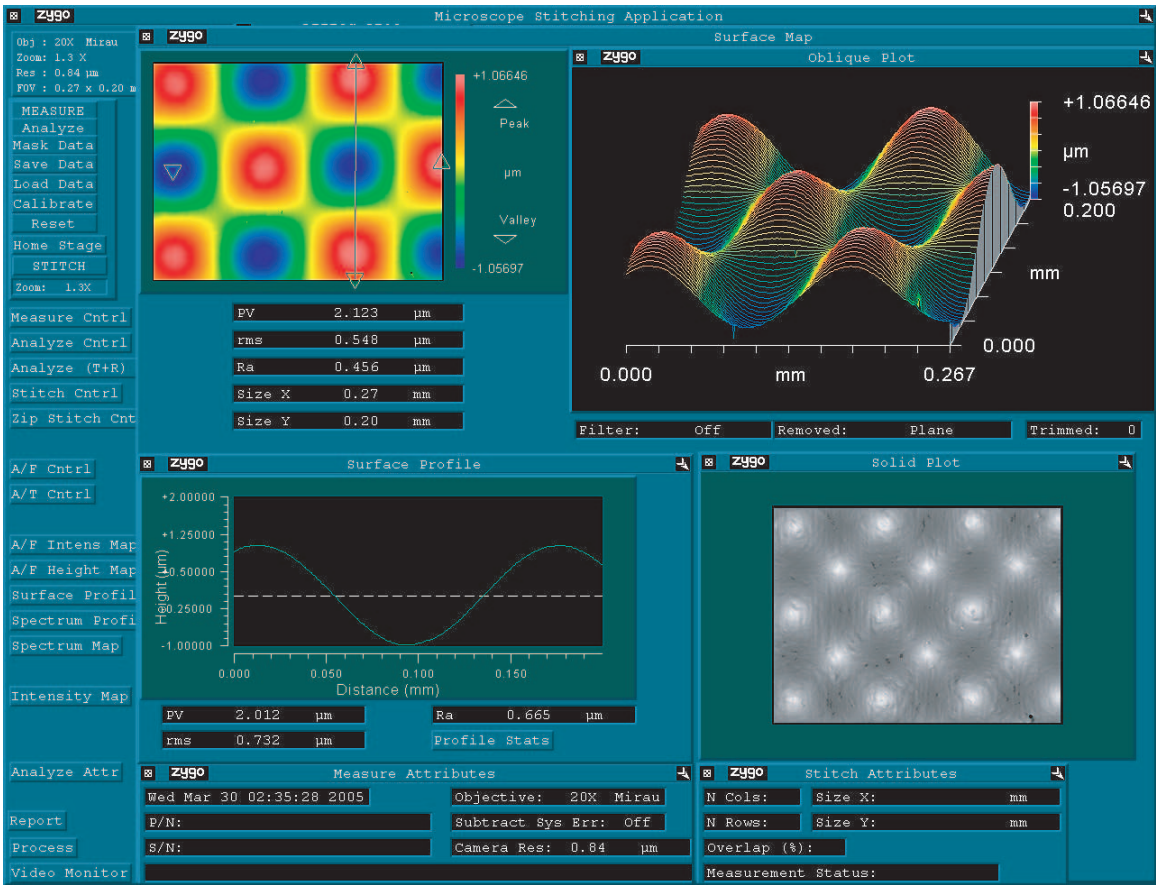


Figure 8-36: Zygo image of a diamond-turned 2-D aluminum surface. The imaged area is 267 μm long and 200 μm wide, with a lateral resolution of 0.84 μm. The horizontal axis is the radial direction, and the vertical axis is the azimuthal direction. The surface profile shows that the azimuthal pitch is about 170 μm and the peak-to-valley value is 2.012 μm.

8.7 Summary

This chapter described the diamond turning results for a flat surface, a 1-D surface, and a 2-D surface. Both aluminum and copper materials were tested. In all the experiments, no chatter was observed for cutting depth ranging from 0.5 nm to 30 nm.

In the flat surface turning, the UFTS has a holding error of 3.7 nm RMS, and the Z-stage has a holding error of 4−5 nm RMS. The resulting surface has a roughness of 7−8 nm RMS over an area of 643×540 μm^2.

In the 1-D surface turning experiment, the UFTS servo control loop has a tracking error of 4.8−5.7 nm RMS, the Z-stage has a holding error of 9 nm RMS, and the X-stage has a tracking error of 3 nm RMS. The roughness of the resulting surface is 9−10 nm along a peak line or a grove line.

In the 2-D surface turning experiment, the tracking error results are very similar to that in the 1-D surface turning. The tracking error is 4.0−5.4 nm RMS in the UFTS servo loop, 8.3 nm RMS in the Z-stage, and 3.0 nm RMS in the X-stage. The resulting surface has a peak-to-valley amplitude of 2.041 nm, only 2% deviation from the command value of 2 nm. This deviation is mostly likely caused by the tilt in the sample set-up under the microscope or the mismatch in the calibration between the capacitance probe sensor and the microscope.

A key issue in non-rotational surface turning is the trajectory generation. As the spindle encoder outputs less than 14 samples in one cycle of the cut sine wave, the spindle speed estimation is very important to surface quality. We analyzed and measured the spindle speed characteristics of both slow dynamics (revolution mean speed) and fast dynamics (mean speed between two consecutive counts). Based on this result, we designed the spindle filter bandwidth to be 70 Hz. For the encoder non-linearity, we developed a dynamic correctional mapping method, which can eliminate synchronous harmonic components.

Chapter 9

Conclusion and Future Work

9.1 Conclusion

In summary, this thesis provides a foundation for significant advancements in fast tool servo technology. The primary contribution of this thesis include:

1. Invented a new class of ultra fast motors with thousands of Gs acceleration capability (Section 2.3-2.7). Presented analysis and design guidelines for this new class of ultra fast motor (Chapter 3).

2. Designed and prototyped a new control-oriented real-time computer design with 1 million samples per second processing capability. (Section 6.2), and demonstrated its high throughput capability (Section 6.3-6.9).

3. Presented a decentralized analog current controller design method for multiple-input-multiple-output linear power current amplifiers. Designed and prototyped a 1 kW 100 kHz power current amplifier and demonstrated the MIMO controller design techniques (Chapter 5).

4. Designed 1200 G acceleration linear ultra fast tool servos based on the presented ultra fast motor concept (Chapter 4). The first prototype using powder pressed iron material has been assembled, and has demonstrated 500 G capability up to 10 kHz operation (Section 7.1) and the usability in diamond turning

experiment(Chapter 8).

5. Presented an embedded structure implementation of adaptive feed-forward cancellation and the corresponding adaptive feed-forward cancellation controller tuning technique and demonstrated effectiveness in attenuating trajectory tracking errors (Section 7.3).

6. Proposed a definition of sensing ability and actuating ability for structural design and control. (Section 4.3.5).

7. Presented a synchronous filter concept and the design method for filtering aliasing and nonlinearity errors of spindle encoder (Section 7.5), and this method has been applied to diamond turning operations to verify its effectiveness (Section 8.4).

9.2 Future Works

Looking into the future, we outline here some suggested future work to fully demonstrate the performance of the UFTS.

9.2.1 Ni-Fe Based UFTS

In this thesis, we have designed two UFTS. One uses 25 μm Ni-Fe as magnetic cores, and the other uses the SM2 powder-pressed iron. Due to time limits, only the SM2-based UFTS is fabricated. In Section 3.8, we have shown by analysis that Ni-Fe can achieve much higher acceleration at high frequency operation because of much lower eddy current loss. Also due to the much higher permeability of Ni-Fe material than SM2, the actuating force of Ni-Fe based UFTS is expected to be five times that of SM2-based FTS for the same air gap and coil winding design. One problem associated with the Ni-Fe based material is its poor structural properties resulting from laminations, and thus the Ni-Fe based UFTS has a destabilizing resonance at 25 kHz, in comparison to 45 kHz of SM2-based UFTS. Even so, the achieved crossover

frequency is still expected to be around 10 kHz by proper controller design. Another potential issue is the mechanical strength of the wound-tape core material. The epoxy-glued lamination might fail in operation, as cores are typically wound by transformer manufactures and we thus had no control on the bonding quality. Eventually, it is ideal to build magnetic cores from Ni-Fe tapes, and the tape bonding and cutting process need to be investigated ensure sufficient bonding strength. It is expected that the assembled Ni-Fe UFTS will demonstrate 1200 G acceleration at non-resonating frequencies, and over 2000 G at resonance.

9.2.2 Flux Sensing and Control

We have presented a flux sensing and control scheme in Section 5.7. Better linearity is expected to be achievable by using flux sensing. Consequently, the requirement on AFC control can be reduced and a simpler controller can also achieve very good performance. One potential problem in flux sensing is the sensing coil shielding, which is very important to isolate electrostatic coupling between the driving coil and sensing coil.

Another control scheme is to use the driving coil voltage as a flux sensing signal at high frequencies and thus the secondary sensing coil can be eliminated. At high frequencies, the driving coil voltage is dominated by the inductive voltage and the voltage drop caused by the coil resistance can be negligible. Therefore, the amplifier works at current mode in low frequencies and at voltage mode in high frequencies. The associated multiple-channel amplifier compensation method for this hybrid working mode needs to be further investigated. This method most likely works well for high permeability materials such as Ni-Fe cores.

9.2.3 Laser Interferometric Position Sensing

The bandwidth of the capacitance probe for position measurement imposes a bottle-neck on the achievable bandwidth of the fast tool servo. The state of art technology of capacitance probes is 100 kHz bandwidth. A laser interferometric position sensor has

the advantage of both high bandwidth and high resolution with low latency. In order to use a laser interferometer as position feedback, a low-inertia mirror surface needs to be installed on the back of the tool holder. In the current design using capacitance position sensor, the backside of the tool holder is a flat aluminum surface. Possible ways to make a mirror surface on the back side tool holder include diamond turning, chemical electrical polishing and/or thin film coating, as opposed to mounting a glass mirror, which will increase the moving inertia significantly.

Anther advantage of using laser interferometer is that the actuator design is not limited by the capacitance probe length and the probe clamping space. Therefore, it is possible to replace the rear rubber bearing with a spring-steel flexural bearing for the motion assembly by extending the ceramic tube length at the back, and the nonlinear and potential high-frequency stiffening problems associated with the rubber bearing can be avoided.

A possible issue is the speed and acceleration limit imposed by the laser interferometer. When the UFTS speed exceeds this limit, the position sensing will lose accuracy. Also proper shielding of the laser path should be considered to ensure position sensing accuracy.

9.2.4 Ultrafast Rotary Motors

In Chapter 2, we have proposed several other ultra rotary motors, which are of significant interest for mirror steering application. In particular, the fully-levitated ultra fast motor with 5 D.O.F. has no limit imposed by mechanical bearings, and thus is expected to demonstrate very high performance. The envisioned challenges include

- Amplifier designs capable of driving eight coils, each of which may require 1 kW reactive power. The decoupling control method proposed in Appendix C needs to be verified experimentally;
- Real-time computer with much higher computation power to control 5 or 6 loops simultaneously;
- 6 D.O.F position sensing without increasing the moving assembly inertia;

- Multiple D.O.F control of the levitated moving assembly without any mechanical damping.

Appendix A

Permanent Magnets

This appendix describes a permanent magnet model, which is used for the flux analysis in Chapter 3. Properties of commonly used rare-earth permanent magnets are also introduced. Based on these properties, we select NeFeB permanent magnets in the ultra fast tool servo design presented in Chapter 4.

A.1 Permanent Magnet Model

Inside a permanent magnet, based on the classical electromagnetic theories [86] the magnetic flux density B is related to the magnetic field strength H as

$$\begin{aligned} B(H) &= \mu_0 (H + M) \\ &= \mu_0 H + \mu_0 M, \end{aligned} \tag{A.1}$$

where M is the magnetization of the permanent magnet. The term $\mu_0 H$ is referred as the excitation flux density, M is the magnetization, and the summation of them is flux density B. As shown in Figure A-1, $\mu_0 M$ is a function of H, which is modelled as a constant $\mu_0 M(0)$ if $\mu_0 H$ is to the right side of point "a", the starting point of demagnetization. When $\mu_0 H$ is reduced from 0 to a negative value and then back to zero, the remanence can be fully recovered if the value of $\mu_0 H$ is never less than point "a". For more negative values of $\mu_0 H$, at point "b", the magnet is

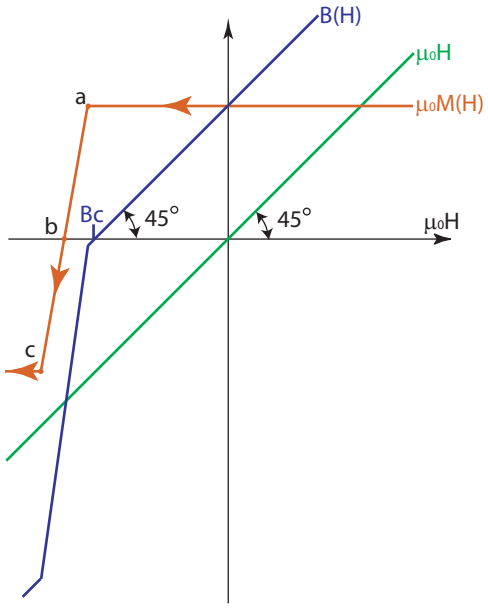


Figure A-1: Typical rare earth permanent magnet characteristics. B_c is referred as coercive flux, the excitation flux that cancels out the remanence flux.

completely demagnetized. When further decreasing $\mu_0 H$ to point "c", the magnet will be magnetized in the reverse direction. The term $\mu_0 H$ is usually magnetized to its maximum strength when magnets are shipped out of factory. However, $\mu_0 H$ is also dependent on temperature. With an increase of temperature, not only is $\mu_0 M(0)$ reduced, but also the whole $\mu_0 M$ curve will move towards the right side, and thus is more readily demagnetized. When choosing a permanent magnet, detailed specifications from the magnet manufactures should be consulted.

A.2 Commonly Used Permanent Magnetic Materials

The most commonly used rare earth permanent magnets include SmCo (Samarium Cobalt) and NdFeB (Neo). The main differences between them are listed as follows: [9]

SmCo: (1) Relatively lower strength, $\mu_0 M(0)$=1.1 T remanence flux. (2) Higher working temperature, greater magnetic output at temperature above 150°C. High

Curie point, up to 750°C or 825°C depending on the material content percentage. (3) Corrosion resistant, no protective coating required. (4) Relatively more expensive. (5) Higher brittleness, very easily to chip and crack, and thus requires round corners of radius of at least 0.1 mm.

Neo: (1) Relatively higher strength, up to $\mu_o M(0)$=1.3 T remanence. (2) Relatively lower working temperature and lower temperature stability. Curie point is about 310°C. (3) Less resistant to oxidation and corrosion, so it needs epoxy coating and/or nickel plating. (4) Very low cost. (5) Relatively less brittle.

Appendix B

Soft Magnetic Materials

This chapter describes power loss models for soft magnetic materials. The simplest model for soft magnetic materials is to assume the permeability μ as infinity. In reality, the effective permeability is dependent not only on magnetic field strength (a nonlinear effect referred as hysteresis), but also on frequency (referred as the eddy current effect). These two effects contribute significantly to the power loss inside soft magnetic materials. These power losses are converted to heat and increase the temperature of soft magnetic materials. As a result, the thermal condition imposes a limit on allowable excitation field strength (especially at high frequencies), which again sets the achievable acceleration limits of our designed ultra fast tool servos.

B.1 Hysteresis Loss

Based on an experimental B-H curve, the volumetric hysteresis power loss p_h can be roughly treated as a linear function of ω and is approximated as:

$$\begin{aligned} p_h(\omega) &= \alpha_h \omega B H, \\ &= \alpha_h \frac{\omega B^2}{\mu}, \end{aligned} \tag{B.1}$$

where α_h is a non-dimensional constant coefficient representing the area encircled by the B-H curves in one cycle, B represents the AC excitation flux amplitude, H is the

AC magnetic field magnitude, and μ is their ratio. p_h is independent of the size of the material, and is the dominant loss at low frequencies.

A B-H curve is a characteristic measured under quasi-static conditions (the magnetic field scans slowly). At high frequencies where eddy current plays a dominant role, the significance of B-H curves is vague, because the field H and B are both position dependent and the curve shape depends on the size and shape of the test sample. As a result, B-H curves measured with the conventional method are not "meaningful" at high frequencies. Instead, we use frequency responses to characterize material properties at high frequency. In next sections, we use a half-order lumped-parameter model to illustrate the frequency response of soft magnetic materials.

B.2 Half-Order Frequency Response of Soft Magnetic Materials

Eddy current loss is a significant loss at high frequencies, and is highly dependent on the size of magnetic materials. In order to reduce the eddy current loss, soft magnetic materials are in shapes of laminations, wound tapes, sintered powder, or pressed powder. Figure B-1 shows the flux distribution inside a one-dimensional lamination of width $2d$ under steady state sinusoidal excitation. The boundary condition is that the magnetic field strength in both ends have the same value

$$H(x = \pm d,\ j\omega) = H_0 e^{(j\omega t)}, \tag{B.2}$$

where H_0 is the magnetic field generated by the excitation current[1] , regardless of the eddy current inside the magnetic material.

According to Maxwell's Equations, the diffusion equation [86] for this one-dimensional

[1]H field can be thought as the superposition of two components: the magnetic field generated by eddy current and the magnetic field generated by external excitation current, although eddy current is generated by the external current.

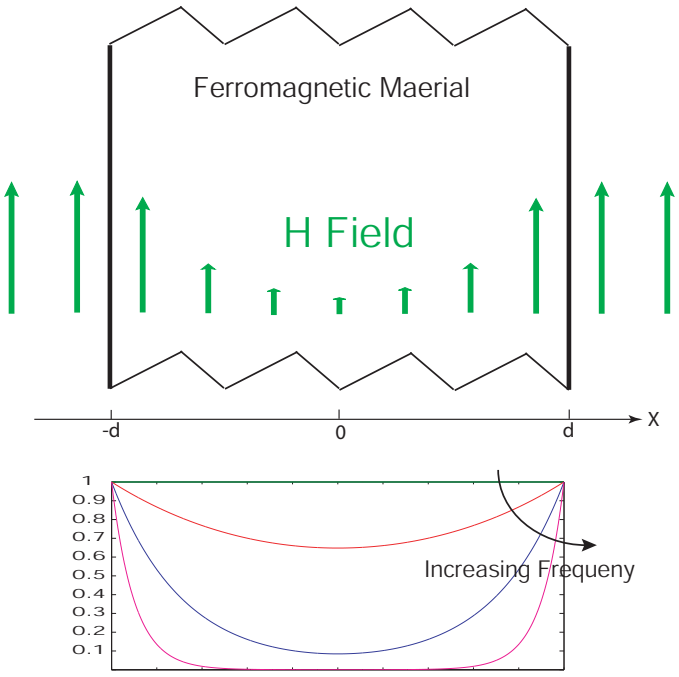


Figure B-1: Magnetic field distribution inside one dimensional lamination immersed in a uniform AC field.

lamination is

$$\frac{1}{\mu\sigma}\frac{\partial^2 H}{\partial x^2} = \frac{\partial H}{\partial t}. \tag{B.3}$$

Combining (B.2) and (B.3), the magnetic field solution is expressed as

$$H(x,t) = H_0 \frac{\cosh\left(\frac{1+j}{\delta}x\right)}{\cosh\left(\frac{1+j}{\delta}d\right)} e^{j\omega t}, \tag{B.4}$$

where $\delta = \sqrt{\frac{2}{\mu\sigma\omega}}$ is the skin depth.

The total flux flow through the material is calculated as

$$\begin{aligned}
\Phi(j\omega) &= \mu \int_{-d}^{d} H(x,t)dx \\
&= \frac{H_0 \mu e^{j\omega t}}{\cosh\left(\frac{1+j}{\delta}d\right)} \int_{-d}^{d} \cosh\left(\frac{1+j}{\delta}x\right) dx \\
&= H_0 \mu e^{j\omega t} \tanh\left(\frac{1+j}{\delta}d\right) \frac{2\delta}{1+j}.
\end{aligned} \tag{B.5}$$

In the case that there is no eddy current inside the magnetic material (by assuming zero conductivity), the total flux $\Phi_0(j\omega)$ generated only by the same excitation field H_0 is

$$\Phi_0(j\omega) = 2d\mu H_0 e^{j\omega t}. \tag{B.6}$$

By comparing $\Phi(j\omega)$ and $\Phi_0(j\omega)$ we can define the AC relative permeability μ_r as

$$\begin{aligned} \mu_r(j\omega) &= \frac{\Phi(j\omega)}{\Phi_0(j\omega)} \\ &= \frac{H_0\mu e^{j\omega t}\tanh\left(\frac{1+j}{\delta}d\right)\frac{2\delta}{1+j}}{2dH_0\mu e^{j\omega t}} \\ &= \frac{\tanh(X)}{X}|_{X=(1+j)\sqrt{\frac{\omega}{\omega_s}}}, \end{aligned} \tag{B.7}$$

where $\omega_s = \frac{2}{\mu\sigma d^2}$ is the skin frequency, at which the material thickness is twice the skin depth. Consequently, $\Phi(j\omega) = 2d\mu\mu_r H_0 e^{j\omega t}$. In this way, we can represent the original distributed parameter system with a lumped parameter model μ_r, which is a half-order system with frequency response shown in Figure B-2. According to Figure B-2, the magnetic field almost fully penetrates the lamination at excitation frequency ω below ω_s, and the total flux reduces at a slope of -10dB/decade for $\omega > \omega_s$. Similar frequency responses can be derived for materials of other shapes. For cylindrical materials with radius "d", the relative AC permeability can be derived as

$$\mu_r(j\omega) = \frac{2}{XI_0(X)}\frac{dI_0(X)}{dX}|_{X=(1+j)\sqrt{\frac{\omega}{\omega_s}}}, \tag{B.8}$$

where $I_0(\cdot)$ is the zero-order Bessel function of the first kind. This frequency response is very similar to that (B.7).

The relative AC permeability can be tested by the setup shown in Figure B-3. Using the experimental measured current $I(j\omega)$ and voltage $V(j\omega)$, The relative AC

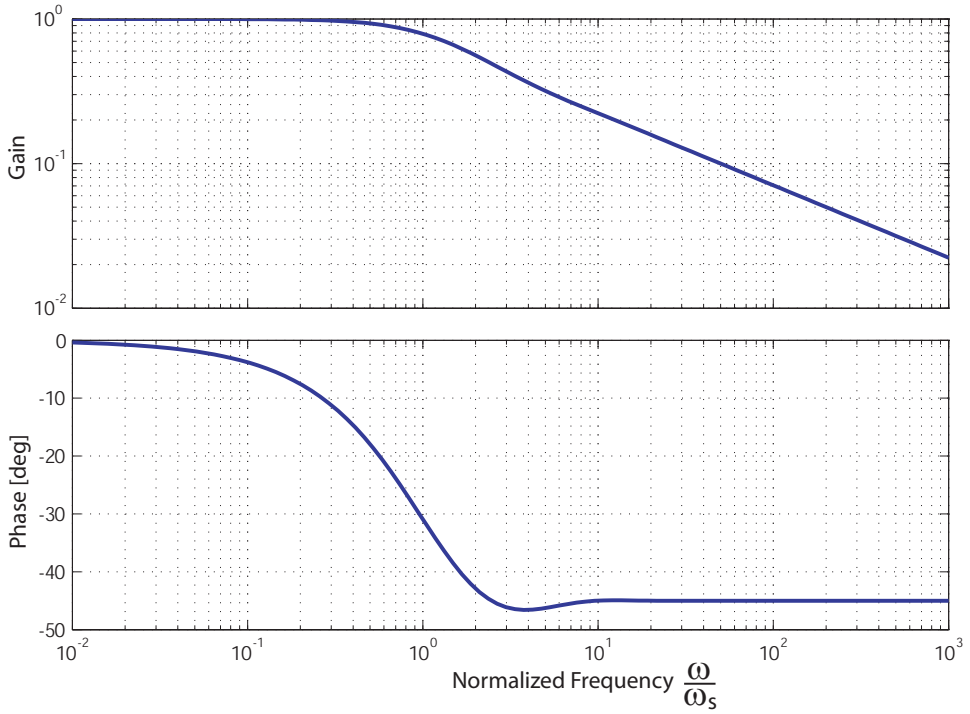


Figure B-2: Simulated frequency response of the AC permeability μ_r for one dimensional lamination.

permeability can be expressed as

$$\begin{aligned} \mu_r(j\omega) &= \frac{\Phi(j\omega)}{\Phi_0(j\omega)} \\ &= \frac{\frac{V(j\omega)}{j\omega n}}{A\mu n\frac{I(j\omega)}{L}} \\ &= \frac{M_{12}(j\omega)}{M(0)}, \end{aligned} \tag{B.9}$$

where A is the core cross section area, n is the coil turn number, L is the core length, and $M_{12}(j\omega) = V(j\omega)/I(j\omega)/j\omega$ is the AC mutual inductance between two coils. For non-conducting magnetic cores, $M_{12}(j\omega)$ is a real constant referred as mutual inductance, and is independent of frequency. However, when there exist eddy currents in magnetic cores, $M_{12}(j\omega)$ will become frequency-dependent complex valued function.

I used an experimental setup similar to Figure B-3 to measure the properties of three materials. Figure B-4 shows experimental μ_r measurement results: sintered

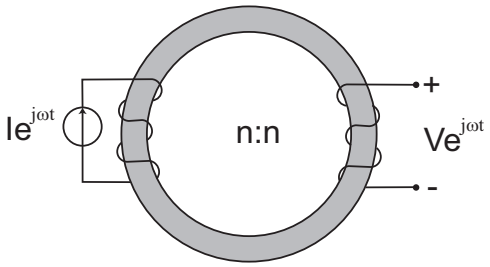


Figure B-3: The test setup for measuring the relative AC permeability μ_r.

powder material (SMS), pressed powder material (SM2), and 175 μm laminated silicon steel (Si-Fe). The SMS result matches the theoretical prediction very well, and the skin frequency is at about 250 Hz. The $\mu_r(j\omega)$ shows a -10 dB/decade slope and -45 deg phase lag above the skin frequency. The SM2 μ_r measurement shows that the skin frequency is 100 kHz. In the frequency region above 1 MHz, the SM2 μ_r frequency response doesn't match the theory well because the input impedance of the oscilloscope is close to the coil output impedance and a L-C resonance resulted. In the Si-Fe μ_r measurement result, the slope is more than -10 dBdecade at frequencies above the 1 kHz skin frequency, mainly because of the electrical resonance which occurred at 20 kHz. A mechanical resonance is also shown at 5kHz in Si-Fe μ_r curve, which is believed to result from the contact stiffness of two "E" cores (instead of a full core) used in the test setup.

B.3 Distributed Air Gaps in Powdered Irons

For magnetic paths including air gaps, the magnetic circuit model can be modified according to the AC relative magnetic permeability as shown in Figure B-5. The AC reluctance of the soft magnetic core $R_m(j\omega)$ can be expressed as

$$R_m(j\omega) = \frac{R_{m0}}{\mu_r(j\omega)}, \tag{B.10}$$

where R_{m0} is the DC reluctance (the length divided by permeability and cross section area). The air gap reluctance R_a is a constant independent of frequency. Figure B-5(c) shows the frequency response of $R_m(j\omega)$, R_a, and their sum. It can be seen that

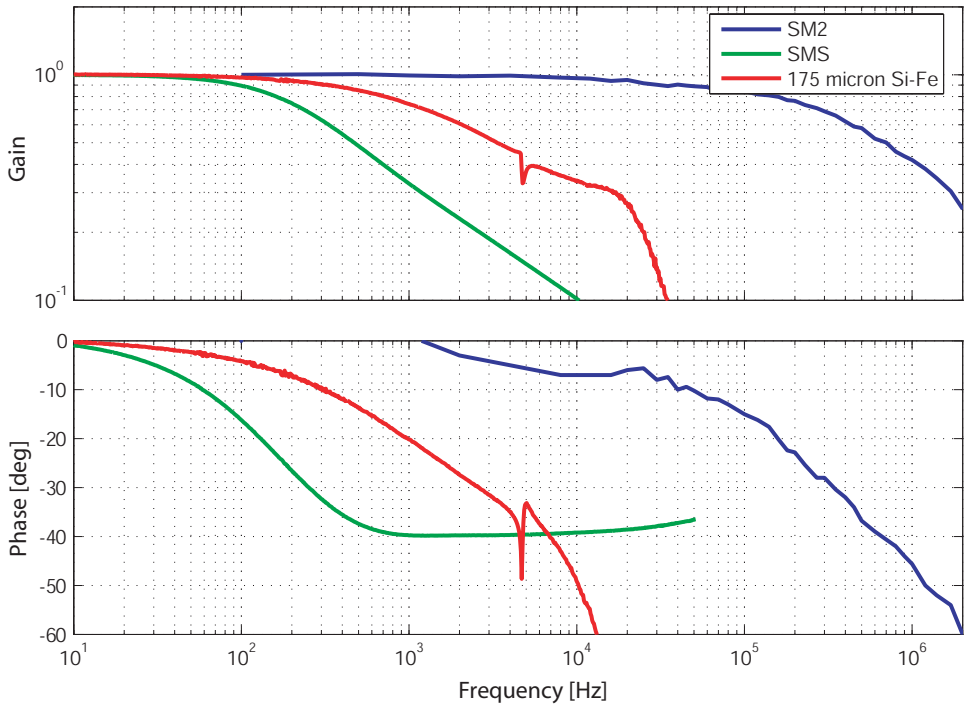


Figure B-4: Experimentally measured frequency responses for three soft magnetic materials.

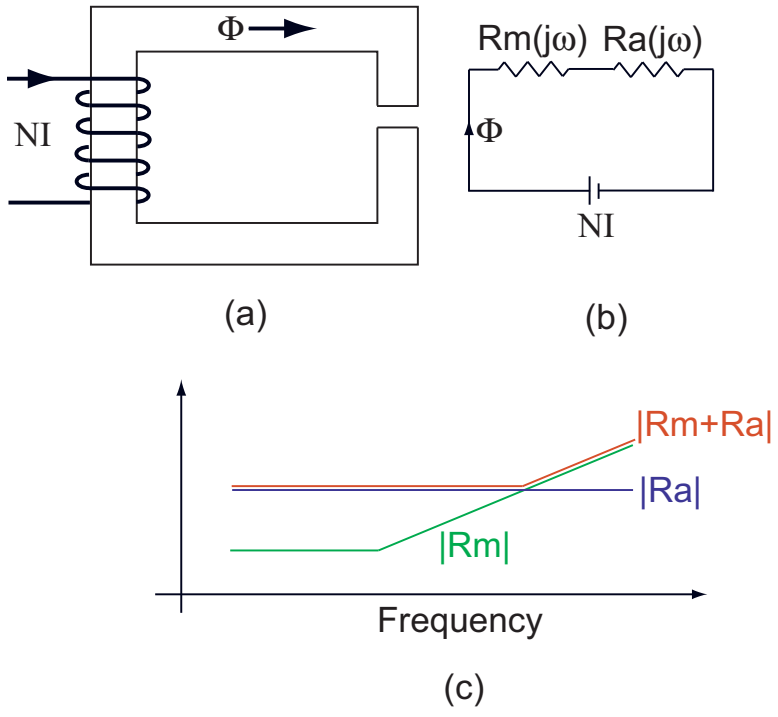


Figure B-5: AC magnetic circuit model of a magnetic path including both a soft iron core and an air gap. Taking into account the eddy current effect, the reluctance of the soft magnetic core is modelled as $R(j\omega)$ instead of a constant.

the break frequency of $R_m(j\omega) + R_a$ has been pushed higher than that of $R_m(j\omega)$ because of the inclusion of the air gap in the magnetic path. If $R_{m0} \ll R_a$ and thus the break frequency of $R_m + R_a$ is well above the frequency range of interest, then in the magnetic circuit analysis we can ignore $R_m(j\omega)$. Otherwise, the eddy current effect should be taken into account by using the AC reluctance $R_m(j\omega)$.

The power pressed iron core SM2 can be modelled as a magnetic path including distributed air gaps. Because of the polymer coating on the iron particles, the break frequency of the reluctance has been significantly increased. This is the main reason that the SM2 material has a very high break frequency in the frequency response in Figure B-4. The penalty for this higher break frequencies is a large DC reluctance in the core.

B.4 Eddy Current Loss

Using the experiment as shown in Figure B-3, in sinusoidal steady state operation we can calculate the eddy current loss as

$$P(\omega) = \frac{1}{2}\mathrm{Re}\{V^*(j\omega)I(j\omega)\}, \tag{B.11}$$

where $V^*(j\omega)$ represents the complex conjugate of $V(j\omega)$. The voltage $V(j\omega)$ and the current $I(j\omega)$ can be related to the total flux Φ flowing through the magnetic core as:

$$V(j\omega) = j\omega N\Phi(j\omega), \tag{B.12}$$

$$I(\omega) = \frac{L\Phi(j\omega)}{\mu AN\mu_r(j\omega)}, \tag{B.13}$$

where N is the turn number of coil windings, L is magnetic path length, and A is the cross section area of the magnetic core. Consequently,

$$P(\omega) = \frac{1}{2}\mathrm{Re}\{V^*(j\omega)I(j\omega)\}$$

$$= -\frac{\omega L|\Phi|^2}{2|\mu_r(j\omega)|A\mu}\sin(\angle\mu_r(j\omega)), \tag{B.14}$$

which is always a positive number for $\omega > 0$.

Because of the eddy currents, the flux distribution inside the magnetic core is no longer uniform. We can introduce a mean flux density $\overline{B} = \frac{\Phi}{A}$, and then relate the volumetric power loss to $\overline{B}$ (instead of to B in hysteresis loss case) as

$$\begin{aligned} p(j\omega) &= \frac{P(\omega)}{AL} \\ &= -\frac{\overline{B}^2\omega}{2|\mu_r(j\omega)|\mu}\sin(\angle\mu_r(j\omega)) \\ &= \mathrm{Re}\{\frac{-j}{2\mu_r(j\omega)}\}\frac{\overline{B}^2\omega}{\mu}. \end{aligned} \tag{B.15}$$

We can see that volumetric power loss is proportional to the mean flux squared $\overline{B}^2$. At frequencies higher than skin frequency ω_s, the volumetric power loss increases with frequency at 30 dB/decade, because $|\mu_r(j\omega)|$ decreases at a slope of -10dB/decade. For frequencies below skin frequency ω_s, (B.15) is convenient to use. In this case, we can use a simple full penetration model to estimate the power loss as in [3].

Figure B-6 shows the magnetic flux $B_Y(j\omega)$ and its induced electric field $E_Z(j\omega)$. They are related as

$$\begin{aligned} \nabla \times \overrightarrow{E} &= \frac{\partial}{\partial t}\overrightarrow{B} \\ -\frac{\partial}{\partial X}E_Z(j\omega, X) &= -\frac{\partial}{\partial t}B_Y(j\omega, X), \\ \frac{\partial}{\partial X}E_Z(j\omega, X) &= j\omega B_Y(j\omega, X), \\ E_Z(j\omega, X) &= j\omega B_Y(j\omega)X. \end{aligned} \tag{B.16}$$

Consequently, the volumetric power loss is

$$p(j\omega) = \frac{\int_{-d}^{d}\frac{1}{2}|E_Z(j\omega)|^2\sigma dX}{2d}$$

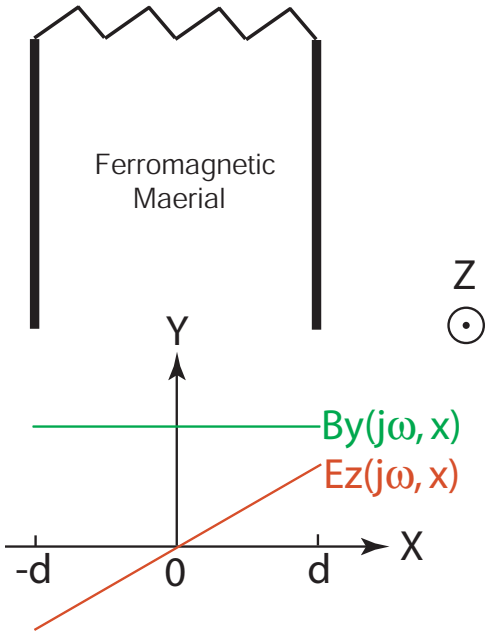


Figure B-6: Eddy current model in full penetration case.

$$= \frac{\sigma d^2 \omega^2}{6} B_Y^2$$
$$= \frac{\omega}{3\omega_s} \frac{B^2 \omega}{\mu}, \tag{B.17}$$

where we remove the sub-script in B_Y because B_Y is the only component of B.

B.5 Power Loss Summary

In conclusion, the power loss can be expressed as

$$p(j\omega) = \begin{cases} \alpha_h \frac{B^2 \omega}{\mu}, & \text{hysteresis dominant region} \\ \frac{\omega}{3\omega_s} \frac{B^2 \omega}{\mu}, & \text{eddy current dominant, but } \omega < \omega_s \\ \mathrm{Re}\{\frac{-j}{2\mu_r(j\omega)}\} \frac{\overline{B}^2 \omega}{\mu}, & \omega > \omega_s \end{cases} . \tag{B.18}$$

At low frequencies where hysteresis loss is dominant, the power loss is proportional to ω. At middle frequencies where the eddy current loss is dominant but the flux can fully penetrate into the material, the power loss is proportional to ω^2 according to (B.15). At high frequencies where the flux cannot evenly distribute within the

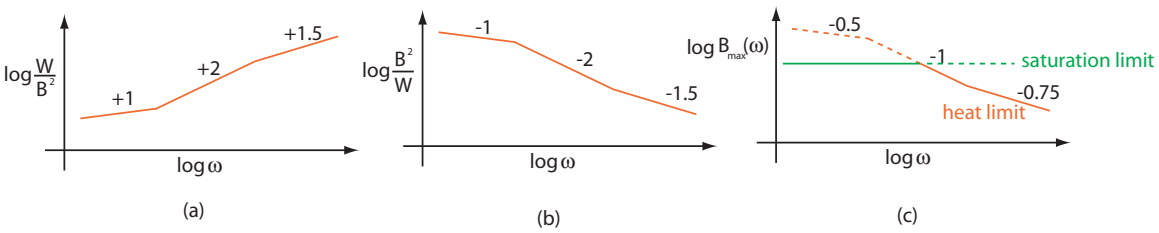


Figure B-7: AC operation energy loss in soft magnetic materials. (a) Volumetric power loss versus frequency. (b) The reciprocal of volumetric power loss versus frequency . (c) Heat limited excitation flux limit versus frequency.

material, the power loss is proportional to $\omega^{1.5}$ according to (B.17). This relation can be expressed by the curves in Figure B-7(a) and its reciprocal is shown in Figure B-7(b). When the electromagnetic actuators are in AC operation, the heat dissipating capability of the soft magnetic material will impose a limit of maximum allowable dissipated energy density, which again sets the maximum allowable flux excitation $B_{\max}$ as shown in Figure B-7(c).

Although this simplified model is not expected to explain the exact power loss mechanism inside soft magnetic materials, the manufactures' power loss data of many materials exhibit the trends illustrated by this simple model.

When calculating the power loss of a specific material, the analytical equations derived here are only used to illustrate intuition and the experimentally measured data from material suppliers should be consulted.

Appendix C

Electrical Drives for the Rotary Ultra Fast Motor

This chapter discusses the driving and control issues for the 3 DOF ultra fast rotary motors presented in Chapter 2.5. We introduce two possible coil winding patterns: 1-2-3-4 winding and A-B-C-D winding. Although 1-2-3-4 winding is not recommend in practice, we use it to illustrate the actuator operation and control principles. Based on the analysis of 1-2-3-4 winding, we discuss the driving and control issues of the motor with A-B-C-D winding. Finally, we present alternative permanent magnet orientations for the rotary ultra fast motor designs.

C.1 1-2-3-4 Winding Pattern

Figure C-1 shows a cross-section of the rotary ultra fast motor. Four coils labelled as 1, 2, 3, and 4 are wound around four separate cores respectively. This winding pattern will be called 1-2-3-4 winding pattern to differentiate from the A-B-C-D pattern discussed subsequently. Four actuating forces F_1, F_2, F_3, and F_4 are shown

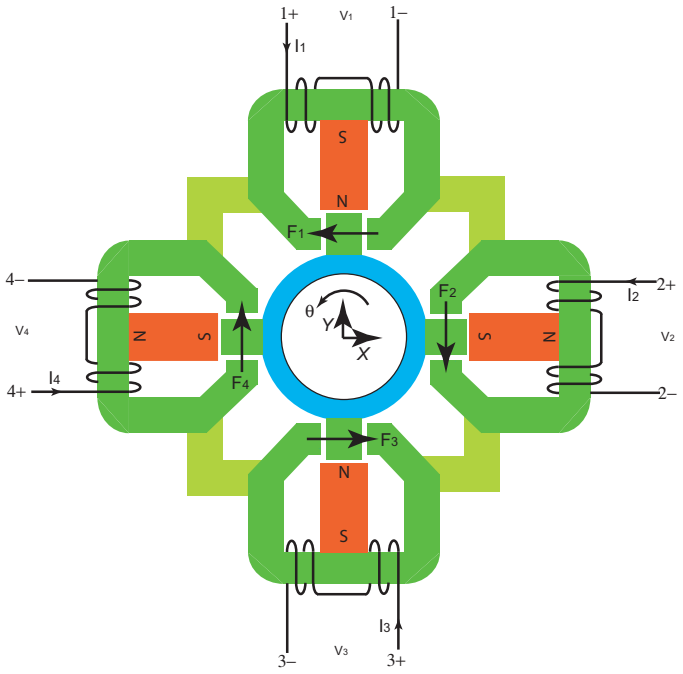


Figure C-1: 1-2-3-4 winding pattern of the rotary ultra fast motor.

in the figure; they are related to excitation currents as

$$\begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{bmatrix} = \begin{bmatrix} K & 0 & 0 & 0 \\ 0 & K & 0 & 0 \\ 0 & 0 & K & 0 \\ 0 & 0 & 0 & K \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix}, \tag{C.1}$$

where K is a constant independent excitation current and armature position, and has been calculated[1] in Chapter 3.

To control this rotary ultra fast motor, we prefer to express the actuating forces in the coordinate of X, Y and θ as shown in Figure C-1. This transform can be

[1]In each armature, a negative spring force is also generated, but is only a function of displacement as analyzed in Chapter 3. To simplify analysis, these forces are ignored here.

calculated as

$$\begin{bmatrix} F_x \\ F_y \\ T \end{bmatrix} = \begin{bmatrix} -1 & 0 & 1 & 0 \\ 0 & -1 & 0 & 1 \\ L & -L & L & -L \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \\ F_3 \\ F_4 \end{bmatrix}, \tag{C.2}$$

where F_x is the X-direction actuating force, F_y is the Y-direction actuating force, T is the actuating torque in θ-direction, L is the distance from the armature to the rotor axis.

Combining (C.1) and (C.3) yields

$$\begin{bmatrix} F_x \\ F_y \\ T \end{bmatrix} = \begin{bmatrix} -K & 0 & K & 0 \\ 0 & -K & 0 & K \\ KL & -KL & KL & -KL \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix}. \tag{C.3}$$

By simple observation, we can establish a decoupling control transform as,

$$\begin{bmatrix} I_{1r} \\ I_{2r} \\ I_{3r} \\ I_{4r} \end{bmatrix} = \begin{bmatrix} -\frac{1}{2K} & 0 & \frac{1}{2KL} \\ 0 & -\frac{1}{2K} & -\frac{1}{2KL} \\ \frac{1}{2K} & 0 & \frac{1}{2KL} \\ 0 & \frac{1}{2K} & -\frac{1}{2KL} \end{bmatrix} \begin{bmatrix} F_{xr} \\ F_{yr} \\ T_r \end{bmatrix}, \tag{C.4}$$

where F_{xr}, F_{yr}, T_r are force reference commands, and I_{1r}, I_{2r}, I_{3r},I_{4r} are command currents. The block diagram for this transform is shown in Figure C-2.

If the current amplifier output current I_i tracks the current command I_{ir} for $i = 1, 2, 3, 4$, then the actuating forces F_x, F_y and T will track their command respectively. The transform in (C.4) not only inverts the actuator characteristics, but also ensures that there is no common-mode current in the four coils, i.e., all currents always sum to zero. This avoids common-mode flux paths in the actuator housing.

Figure C-3 shows one way to connect the four coils, which we refer to as the

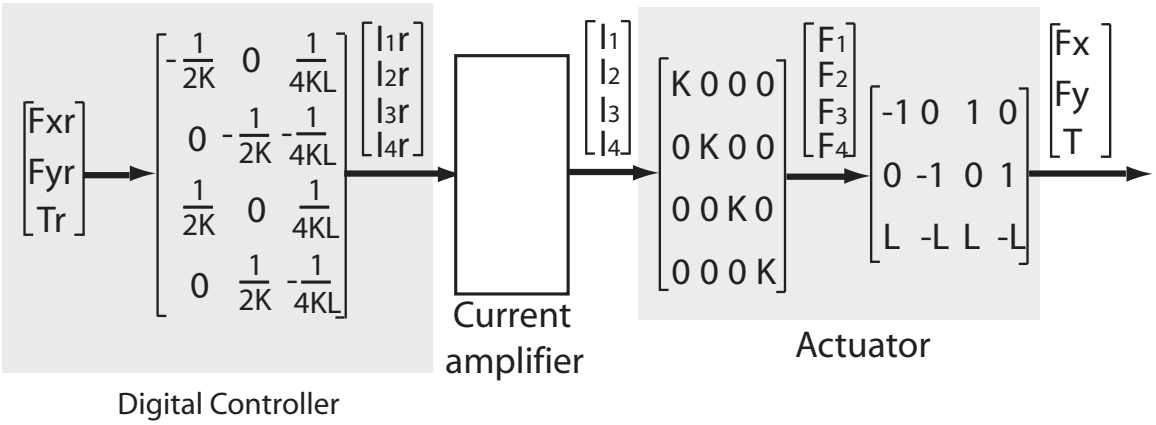


Figure C-2: Reference current generated from designed actuating forces and torque.

X-connection. The neutral point in the X-connection center should be left floating. Otherwise, it is very difficult to ensure no common mode current flowing into the system and there exist very strong common-mode magnetic coupling among the four channels. Connecting the coils in this way, only four terminals need to be connected to the power amplifier. When designing power amplifiers for this system, three currents should be sensed and controlled (I_1, I_2, and I_3), with the fourth channel running in voltage mode ($V_{4r} = -V_{1r} - V_{2r} - V_{3r}$), where V_{ir} is the voltage stage reference signal in the i-th coil. As a result, if the actuators are assumed identical, zero reference potential is created at the neutral point. In this way, there is no coupling among currents in different coils. Therefore, the current controller design is the compensation of three single-input-single-output(SISO) independent loops. This simplicity of electronic configuration is the main advantage of using the X-connection in the 1-2-3-4 winding pattern.

C.2 A-B-C-D Winding Pattern

This section present a winding pattern which is more compact than the 1-2-3-4 winding pattern. As shown in Figure C-4, each coil links two cores; this pattern is referred to as the A-B-C-D winding pattern. Very different from the 1-2-3-4 winding pattern (coils are completely isolated from each other), there exist strong coupling between coil windings in the A-B-C-D pattern. One main advantage of A-B-C-D winding

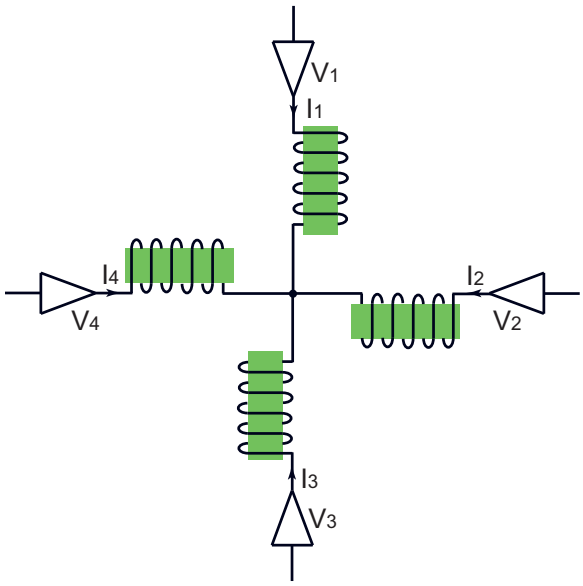


Figure C-3: X-connection of the rotary ultra fast motor coil .

over 1-2-3-4 winding is that the coils in the A-B-C-D winding pattern are twice as compact, have about half the resistance, and thus generate half the Ohmic heat.

Apparently, there exists strong coupling among coils in this pattern. As a result, the current control becomes a Multiple-Input-Multiple-Output(MIMO) problem. In order to investigate actuator characteristic and further design current controllers, we need to relate the currents in the A-B-C-D pattern to those in 1-2-3-4 pattern as

$$\begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} = W_1 \begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix}, \tag{C.5}$$

$$\text{where } W_1 = \begin{bmatrix} -1 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & 1 & -1 \end{bmatrix}, \tag{C.6}$$

and I_A, I_B, I_C, and I_D are the coil currents in coils A, B, C and D respectively.

Although matrix W_1 is singular, we can invert (C.5) to solve for $I_{A,\ B,\ C,\ D}$ by

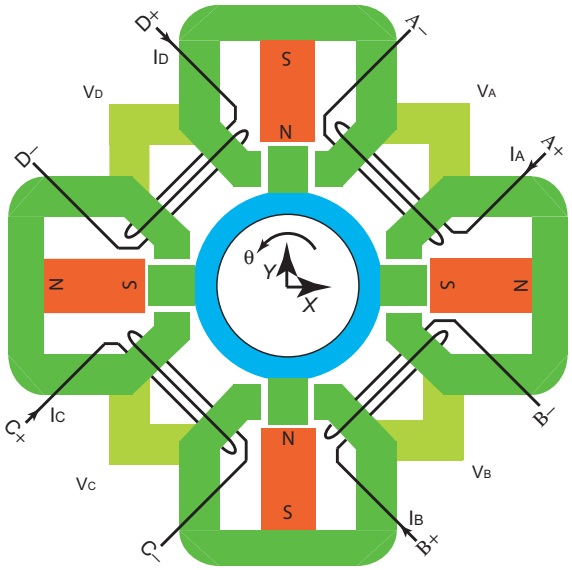


Figure C-4: A-B-C-D winding pattern for the rotary ultra fast motor.

introducing variable I_m

$$\begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix} = W_2 \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} I_m, \tag{C.7}$$

$$\text{where } I_m = \begin{bmatrix} \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} \end{bmatrix} \begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix}, \tag{C.8}$$

$$\text{and } W_2 = \frac{1}{4} \begin{bmatrix} -1 & 2 & 1 & 0 \\ -1 & -2 & 1 & 0 \\ -1 & -2 & -3 & 0 \\ 3 & 2 & 1 & 0 \end{bmatrix}. \tag{C.9}$$

By defining $Z(s)$ as the mutual impedance between two adjacent coils and R as the resistance of a single coil in A-B-C-D winding pattern, the plant transfer function from coil currents to coil terminal voltage drops V_A, V_B, V_C , and V_D can be expressed

as

$$\begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix} = P \begin{bmatrix} V_A \\ V_B \\ V_C \\ V_D \end{bmatrix}, \tag{C.10}$$

$$\text{where } P = \begin{bmatrix} 2Z(s)+R & -Z(s) & 0 & -Z(s) \\ -Z(s) & 2Z(s)+R & -Z(s) & 0 \\ 0 & -Z(s) & 2Z(s)+R & -Z(s) \\ -Z(s) & 0 & -Z(s) & 2Z(s)+R \end{bmatrix}^{-1}. \tag{C.11}$$

Figure C-5 shows three ways to connect four coils in A-B-C-D winding pattern: X-connection, Xg-connection, and ⋄-connection. The center point in the X-connection is left floating, while the center point in Xg-connection is grounded. As systems in these three connections behave very differently, we present the drive and control for each connection separately.

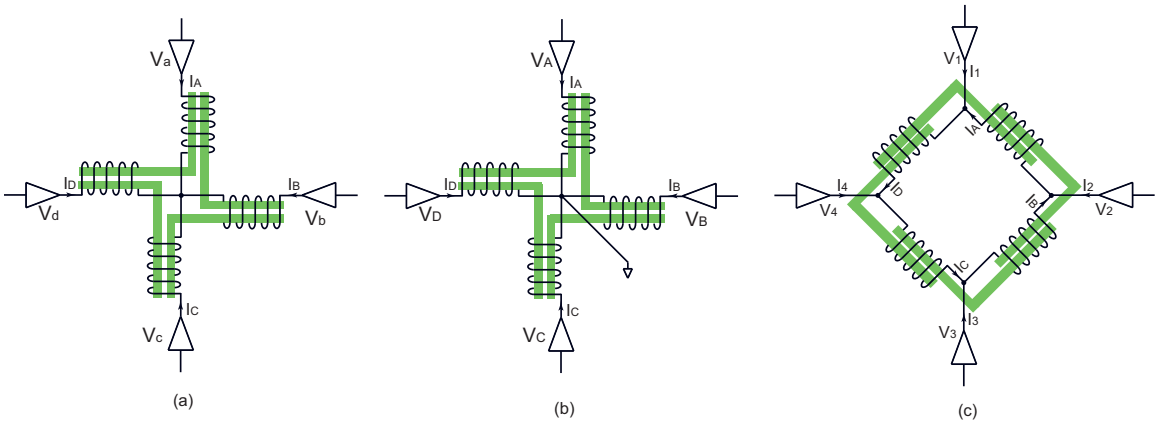


Figure C-5: Three connection methods for the Rotary motor coil in the A-B-C-D winding pattern. (a) is X-connection, (b) is Xg-connection, and (c) is ⋄-connection.

C.2.1 Driving in X-connection of A-B-C-D Winding Pattern

In X-connection of the A-B-C-D winding pattern, we need to control three coils in current mode (coil A, B , and C) and control the fourth coil in voltage mode (coil

D) by keeping the common mode component of all driving reference voltages at zero $V_{dr} = -V_{ar} - V_{br} - V_{cr}$. As a result, the common mode component of all driving currents are absolutely zero, $I_m = 0$. Assuming the center point potential is V_g, according to (C.11), we have

$$\begin{bmatrix} V_A \\ V_B \\ V_C \\ V_D \end{bmatrix} = \begin{bmatrix} V_a - V_g \\ V_b - V_g \\ V_c - V_g \\ V_d - V_g \end{bmatrix} = P^{-1} \begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix}, \tag{C.12}$$

where V_a, V_b, V_c, V_d is the driving terminal potential of coil a, b, c, and d respectively. By multiplying a constant matrix on both sides we get

$$\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} V_a - V_g \\ V_b - V_g \\ V_c - V_g \\ V_d - V_g \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} P^{-1} \begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix}, \tag{C.13}$$

$$V_a + V_b + V_c + V_d - 4V_g = R(I_A + I_B + I_C + I_D) = 0, \tag{C.14}$$

$$V_g = \frac{V_a + V_b + V_c + V_d}{4}. \tag{C.15}$$

As we command zero common mode voltage on V_a, V_b, V_c, and V_d, V_g is the mean value of the four voltage stage tracking errors.

With the constraint $I_m = 0$, the plant transfer matrix can be reduced as:

$$\begin{bmatrix} V_A \\ V_B \\ V_C \\ V_D \end{bmatrix} = P^{-1} \begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix}$$

$$= P^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -1 & -1 \end{bmatrix} \begin{bmatrix} I_A \\ I_B \\ I_C \end{bmatrix}$$

$$= \begin{bmatrix} 3Z(s)+R & 0 & Z(s) \\ -Z(s) & 2Z(s)+R & -Z(s) \\ Z(s) & 0 & 3Z(s)+R \\ -3Z(s)-R & -2Z(s)-R & -3Z(s)-R \end{bmatrix} \begin{bmatrix} I_A \\ I_B \\ I_C \end{bmatrix}. \quad \text{(C.16)}$$

$$\begin{bmatrix} I_A \\ I_B \\ I_C \end{bmatrix} = P_{ABC} \begin{bmatrix} V_A \\ V_B \\ V_C \end{bmatrix}, \quad \text{(C.17)}$$

$$\text{where } P_{ABC} = \begin{bmatrix} 3Z(s)+R & 0 & Z(s) \\ -Z(s) & 2Z(s)+R & -Z(s) \\ Z(s) & 0 & 3Z(s)+R \end{bmatrix}^{-1}. \quad \text{(C.18)}$$

The system control block diagram can be represented as in Figure C-6. Here a decentralized current controller with identical diagonal element $c(s)$ is used to control plant P_{ABC}. Applying the theory derived in Chapter 5, the system is stable if and only if $c(s)$ can simultaneously stabilize the following two SISO (Single-Input-Single-Output) sub-plants: $\frac{1}{4Z(s)+R}$ and $\frac{1}{2Z(s)+R}$.

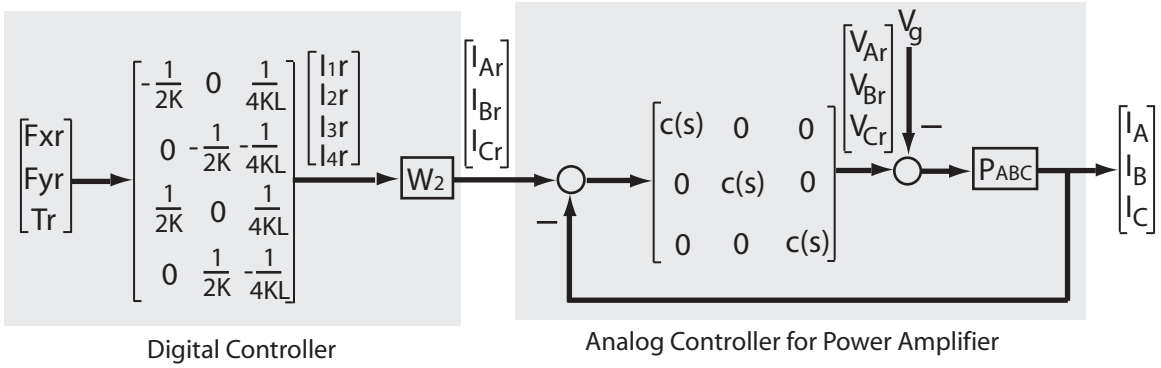


Figure C-6: Block diagram for current control of the rotary ultra fast motor in X-connection of A-B-C-D winding pattern.

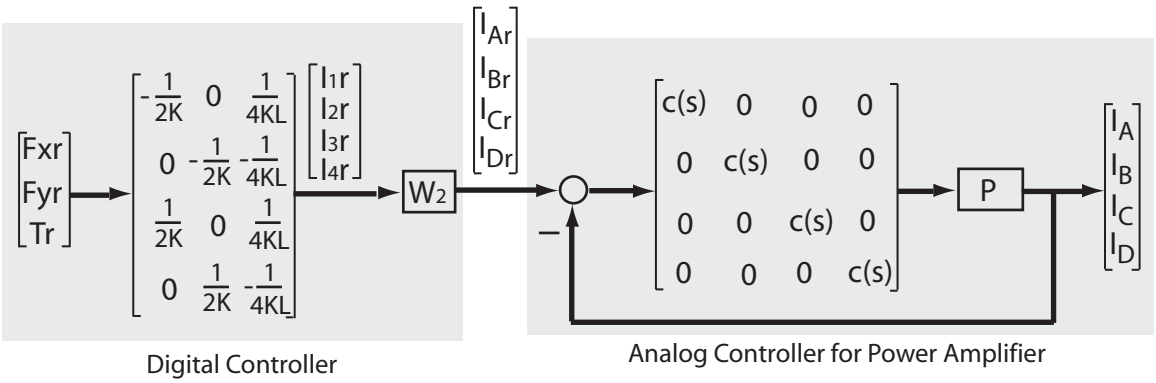


Figure C-7: Block diagram for current control of the rotary ultra fast motor in Xg-connection of A-B-C-D winding pattern.

C.2.2 Driving in Xg-connection of A-B-C-D Winding Pattern

In the Xg-connection of A-B-C-D winding pattern, all four coils should be controlled in current mode, with the center ground connection acting as a return path. One main advantage is that the current sensing can be easily implemented by single-ended power resistors on the ground side of the each coil. In other connections, differential voltage sensing of floating resistor must be implemented. The system control block diagram is shown in Figure C-7. Here a decentralized current controller $c(s)$ is used again to control plant P of (C.11). According to the theory derived in Chapter 5, the system is stable if and only if $c(s)$ can simultaneously stabilize the following three SISO (Single-Input-Single-Output) sub-plants: $\frac{1}{R}$, $\frac{1}{4Z(s)+R}$, and $\frac{1}{2Z(s)+R}$.

C.2.3 Driving in $\diamond$-connection of A-B-C-D Winding pattern

In the $\diamond$-connection of the A-B-C-D winding pattern of Figure C-5(c), the winding currents I_A, I_B, I_C, and I_D become implicit to power amplifiers. These currents are replaced by I_1, I_2, I_3, and I_4, and they have the consistent physical meaning as those in (C.5). We need to control three terminals in current mode (coil 1, 2 , and 3) and control the fourth terminal in voltage mode (terminal 4) by ensuring the common mode component of all driving voltages at zero $V_{4r} = -V_{3r} - V_{2r} - V_{1r}$. The system

is constrained by

$$I_1 + I_2 + I_3 + I_4 = 0, \tag{C.19}$$

$$V_A + V_B + V_C + V_D = 0. \tag{C.20}$$

Now, we need to analyze the common mode current among all coil windings (current circulating inside four coils, but not observable to I_1, I_2, I_3, and I_4). According to (C.11), we have:

$$\begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} V_A \\ V_B \\ V_C \\ V_D \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 \end{bmatrix} P^{-1} \begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix}, \tag{C.21}$$

$$V_A + V_B + V_C + V_D = 4RI_m, \tag{C.22}$$

$$I_m = \frac{V_A + V_B + V_C + V_D}{4R} = 0. \tag{C.23}$$

This result shows that the symmetric electrical loads ensure that there is no circulating current.

The physical connection also imposes the following relation between V_1, V_2, V_3, V_4 and V_A, V_B, V_C, V_D

$$\begin{bmatrix} V_A \\ V_B \\ V_C \\ V_D \end{bmatrix} = W_3 \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix}, \tag{C.24}$$

$$\text{where } W_3 = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ 1 & 0 & 0 & -1 \end{bmatrix}. \tag{C.25}$$

Although matrix W_3 is singular, we can get the required inverse by introducing

variable V_e, which is the common mode component of V_1, V_2, V_3, V_4

$$\begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix} = W_4 \begin{bmatrix} V_A \\ V_B \\ V_C \\ V_D \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} V_e, \tag{C.26}$$

$$\text{where } W_4 = \frac{1}{4} \begin{bmatrix} -3 & -2 & -1 & 0 \\ 1 & -2 & -1 & 0 \\ 1 & 2 & -1 & 0 \\ 1 & 2 & 3 & 0 \end{bmatrix}. \tag{C.27}$$

The plant transfer matrix can be expressed as :

$$\begin{aligned} \begin{bmatrix} V_1 \\ V_2 \\ V_3 \\ V_4 \end{bmatrix} &= W_4 \begin{bmatrix} V_A \\ V_B \\ V_C \\ V_D \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} V_e \\ &= W_4 P^{-1} \begin{bmatrix} I_A \\ I_B \\ I_C \\ I_D \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} V_e \\ &= W_4 P^{-1} W_2 \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} V_e \\ &= P_{1234}^{-1} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \\ I_4 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} V_e, \end{aligned} \tag{C.28}$$

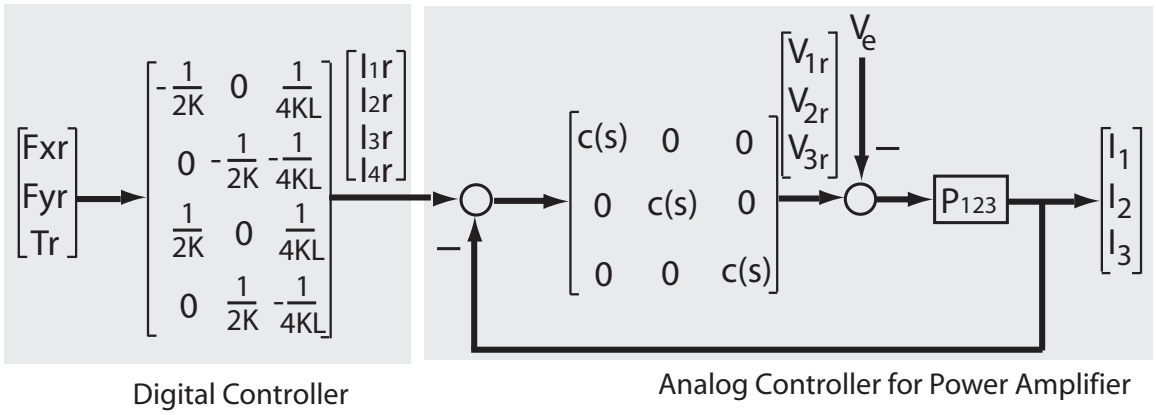


Figure C-8: Block diagram for current control of the rotary ultra fast motor in $\diamond$-connection of A-B-C-D winding pattern.

$$\text{where } P_{1234} = \begin{bmatrix} Z(s)+\frac{3}{8}R & 0 & -\frac{1}{8}R & 0 \\ \frac{1}{8}R & Z(s)+\frac{1}{2}R & \frac{1}{8}R & 0 \\ -\frac{1}{8}R & 0 & Z(s)+\frac{3}{8}R & 0 \\ -\frac{3}{8}R-Z(s) & -Z(s)-\frac{1}{2}R & -\frac{3}{8}R-Z(s) & 0 \end{bmatrix}^{-1}. \quad \text{(C.29)}$$

Further, the plant can be simplified as:

$$\begin{bmatrix} V_1 \\ V_2 \\ V_3 \end{bmatrix} = P_{123}^{-1} \begin{bmatrix} I_1 \\ I_2 \\ I_3 \end{bmatrix} + \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} V_e, \quad \text{(C.30)}$$

$$\text{where } P_{123} = \begin{bmatrix} Z(s)+\frac{3}{8}R & 0 & -\frac{1}{8}R \\ \frac{1}{8}R & Z(s)+\frac{1}{2}R & \frac{1}{8}R \\ -\frac{1}{8}R & 0 & Z(s)+\frac{3}{8}R \end{bmatrix}^{-1}. \quad \text{(C.31)}$$

The system control block diagram is shown in Figure C-8. Here we use a decentralized current controller $c(s)$ to control plant P_{123} . Based on the theory derived in Chapter 5, the system is stable if and only if $c(s)$ can simultaneously stabilize the following two SISO (Single-Input-Single-Output) sub-plants: $\frac{4}{4Z(s)+R}$ and $\frac{2}{2Z(s)+R}$.

C.3 Design Variation of Permanent Magnet Orientation

The design in Figure 2-25 illustrates only one possible way of orienting the permanent magnets. Although the permanent magnets cannot be oriented arbitrarily, there exist other possibilities. An acceptable orientation requires that there exists a decoupling matrix to transform the actuating characteristic in a decoupled format. It can be verified that only three orientation designs as shown in Figure C-9 are acceptable.

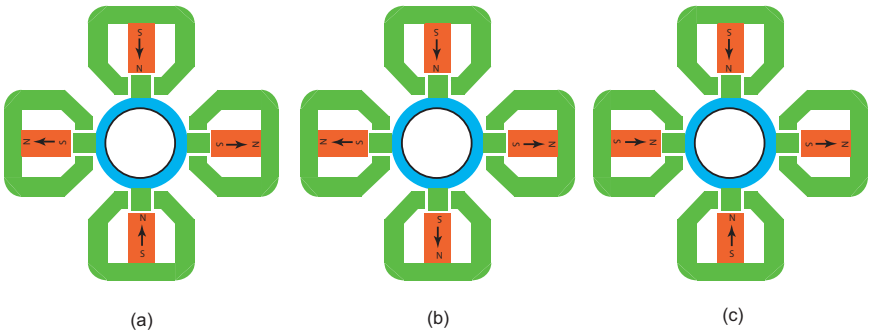


Figure C-9: Permanent magnet orientations for ultra fast rotary motor designs.

Bibliography

[1] Y. Altintas and A. Woronko. A piezo tool actuator for precision turning of hardened shafts. *Annals of the CIRP*, 51(1), 2002.

[2] Jacob Apkarian. Data sheet of qstudiorp power i. *www.quanser.com*, 2005.

[3] H. Wayne Beaty and James L. Kirtley Jr. *Electric motor handbook*. McGraw-Hill, 1998.

[4] M. Bodson, A. Sacks, and P. Khosla. Harmonic generation in adaptive feedforward cancellation schemes. *IEEE Tran. on Automatic Control*, 39(9):1939–1944, 1994.

[5] E. Brinksmeier and W. Preuss. Complex surfaces - applications and generation by diamond machining. *Proceedings of the ASPE 1997 Annual Meeting*, 1997.

[6] M. Byl, S. Lucwick, and D. Trumper. A loop shaping perspective for tuning controllers with adaptive feedforward cancellation. *Precision Engineering*, 29(1):27–40, 2005.

[7] Marten F. Byl. *Design and Control of a Long Stroke Fast Tool Servo*. PhD dissertation, M.I.T., Department of Mechanical Engineering, 2005.

[8] Joseph Harry Cattel. *Adaptive Feedforward Cancellation Viewed from an Oscillator Amplitude Control Perspective*. MS thesis, M.I.T., Department of Mechanical Engineering, 2003.

[9] Arnold Magnetic Technologies Corporation. How to select the appropriate permanent magnet material. *http://www.arnoldmagnetics.com, TN 0205,*, 2002.

[10] J. F. Cutttino. Performance optimization of a fast tool servo for single-point diamond. *IEEE/ASME Trans. on Mechatronics*, 4(2), 1999.

[11] Angelo C.P. de Klerky, Georgo Z. Angelisy, and Jan van Eijk. Design of a next generation 6 dof stage for scanning application in vacuum with nanometer accuracy and mgauss magnetic stray field. *Proceedings from ASPE 2004 Annual Meeting*, pages 60–63, 2004.

[12] S. Douglass. *A machining system for turning non-axis-symmetric surfaces*. PhD dissertation, the University of Tennessee, Knoxville, 1983.

[13] T. A. Dow, M. H. Miller, and P.J. Falter. Application of a fast tool servo for diamond turning of nonrotationally symmetric surfaces. *Precision Engineering*, 13(4), 1991.

[14] H. Eda. Ultra-precise machine tool equipped with a giant magnetostriction actuator - development of new materials and their application. *CIRP annals*, 41(1), 1992.

[15] Chris J. Evans and James B. Bryan. "structured", "textured" or "engineered" surfaces. *CIRP annals*, 48(2), 1999.

[16] K. J. Falter and David H. Youden. The characterization and testing of a long stroke fast tool servo. *International Progress in Precision Engineering: Proceedings of 8th International Precision Engineering Seminar*, 1995.

[17] P. Falter and T. Dow. A diamond-turning apparatus for fabrication of non-rotationally symmetric surface generation. *Proceedings of the International Congress for Ultraprecision Technology, Aachen*, pages 187–201, 1988.

[18] G.F. Franklin, J.D. Powell, and M.L. Workman. *Digital Control of Dynamic Systems, Third Edition*. Addison-Wesley, 1997.

[19] II G. M. Moorefield. Generation of rotationally asymmetric optical surfaces using a fast tool servo. *Proc. of ASPE 1995 annual meeting*.

[20] Alan N. Gent. *Engineering with Rubber: How to Design Rubber Components.* Hanser Publishers, 1995.

[21] W. Greene and D. Shinstock. Design of a linear voice coil actuator for fast tool servo applications. *Proc. of ASPE 1997 annual meeting.*

[22] H. M. Gutierrez and P. I. Ro. Parametric modeling and control of a long-range actuator using magnetic servo-levitation. *IEEE Tran. on Magnetics*, 34(5), 1998.

[23] H. M. Gutierrez and P. I. Ro. Sliding-mode control of a nonlinear-input system: application to a magnetically levitated fast tool servo. *IEEE Tran. on Industrial Electronics*, 45(6), 1998.

[24] Y. Hara, S. Motonishi, K. Yoshida, and N. Ikawa. A new micro-cutting device with high stiffness and resolution. *Annals of the CIRP*, 39(1), 1990.

[25] Yasushi Horiuchi. Development of magnetically suspended sliders for deployable antena test facility. *Proceedings of the Fifth International Symposium on Magnetic Bearings*, pages 363–368, 1996.

[26] L. Jabben, P.M. Overschie, A., and H.F. van Beek. Loretz motor with stationary magnets and coils applied in a 6-dof motion state. *Proceedings of the ASPE 2002 Spring Topical Meeting*, page 3942, 2001.

[27] B. Jared. Fabrication of surface perturbation on inertial confinement fusion targets. *Proc. of ASPE 1996 annual meeting.*

[28] B.V. Jayawant, R.L. Hodkinson, A.R. Wheeler, and R.J. Whorlow. Transducers and their influence in the design of magnetically suspended vehicles. *I.E.E. Conf. on Control Aspect of New Forms of Guided Land Transport, I.E.E. Publication No. 117*, pages 200–206, August 1974.

[29] Howard W. Johnson and Martin Graham. *High-speed Digital Design: A Handbook of Black Magic.* Prentice Hall, 1993.

[30] John R. Glover Jr. Adaptive noise cancelling applied to sinusoidal interferences. *IEEE Tran. on Acoustics, Speech, and Signal Processing*, December 1977.

[31] Stephen Joseph Ludwick Jr. *A Rotary Fast Tool Servo for Diamond Turning of Asymmetric Optics*. PhD dissertation, M.I.T., Department of Mechanical Engineering, 1999.

[32] Andrew Kenny and Alan B. Palazzolo. Single plane radial, magnetic bearings biased with poles containing permanent magnets. *Tran. of ASME, Journal of Mechanical Design*, 125:178–185, 2003.

[33] E. Kouno. A fast reponse piezoelectric actuator for servo correction of systematic errors in precision machining. *Annals of the CIRP*, 33(1):369–372, 1984.

[34] Dick A.H. Laro, Jan van Eijk, and Anton Lebedev. Two degree of freedom actuator for suspension and propulsion. *Proceedings of the ASPE 2002 Spring Topical Meeting*, pages 50+, 2004.

[35] Bruno Lequesne. Bistable electromechanical valve actuator. *US Patent 4,779,582*, 1988.

[36] Bruno Lequesne. Fast-acting, long-stroke solenoids with two springs. *IEEE Tran. on Industry Applications*, 26(5):848+, 1990.

[37] D. W. Lewis, R. R. Humphris, E. H. Maslen, and R. D. Williams. Magnetic bearing for pumps, compressors, and other rotating machinery. *US Patent 5355042*, 1994.

[38] P. B. Lindley. *Engineering Design with Natural Rubber*. NR Technical Bulletin, 1978.

[39] D. Liu. Surface texture improvement in the turning process via application of a magnetostrictively actuatored tool holder. *ASME Journal of DSMC*, 120, 1998.

[40] George E. Losier. Electric machine. *US Patent 2092765*, 1937.

[41] Xiaodong Lu and David L. Trumper. Electromagnetically driven fast tool servo. *Proceedings from ASPE 2003 Annual Meeting*, pages 103–106, 2003.

[42] Xiaodong Lu and David L. Trumper. Electromagnetically driven ultrafast tool servo. *Proceedings from ASPE 2004 Annual Meeting*, pages 269–272, 2004.

[43] Xiaodong Lu and David L. Trumper. High bandwidth fast tool servo control. *American Control Conference 2004*, 2004.

[44] Xiaodong Lu and David L. Trumper. Ultra fast tool servos for diamond turning. *Annals of the CIRP*, 54(1), 2005.

[45] S. Ludwick and D. L. Trumper. Design of a rotary fast tool servo for ophthalmic lens fabrication. *Precision Engineering*, 23(4):253–259, 1999.

[46] Jan Marian Maciejowski. *Multivariable feedback design*. Electronic systems engineering series. Addison-Wesley, 1989.

[47] Mathworks. Data sheet, real-time windows target 2.5. *www.mathworks.com*, 2004.

[48] Mathworks. xpc target 2.6, perform real-time rapid prototyping and hardware-in-the-loop simulation using pc hardware. *www.mathworks.com*, 2004.

[49] Leonard Meirovitch. *Elements of Vibraton Analysis*. McGraw-Hill, 1986.

[50] W. Messner and M. Bodson. Design of adaptive feedforward algorithms using internal modal equivalence. *International Journal of Adaptive Control and Signal Processing*, 1995.

[51] Apex Microtechnology. *Evaluation Kit for PA04 Pin-Out*.

[52] Apex Microtechnology. *PA52 Datasheet*.

[53] A. Molenaar, H.F. van Beek, and M.J.L. Sanders. A new linear magnetic bearing configuration for high accuracy positioning. *Proceedings of MAG97: Magnetic Bearings, Magnetic Drives and Dry Gas Seals Conference & Exhibition*, 1997.

[54] A. Molenaar, H.F. van Beek, and M.J.L. Sanders. A novel low dissipation long stroke planar magnetic suspension and propulsion stage. *Proceedings of the Sixth International Symposium on Magnetic Bearings*, pages 650–659, 1998.

[55] Jean I. Montagu. Actuator with compensating flux path. *US Patent 4528533*, 1985.

[56] R. C. Montesanti and D. L. Trumper. High bandwidth short stroke rotary fast tool servo. *Proceedings from ASPE 2003 Annual Meeting*, pages 115–118, 2003.

[57] R. C. Montesanti and D. L. Trumper. A 10 khz short-stroke rotary fast tool servo. *Proceedings from ASPE 2004 Annual Meeting*, October 2004.

[58] R. C. Montesanti and D. L. Trumper. Design and implementation of the control system for a 2 khz rotary fast tool servo. *Proceedings of the ASPE 2002 Spring Topical Meeting*, pages 28–33, 2004.

[59] Richard Montesanti. *High Bandwidth Rotary Fast Tool Servos and a Hybrid Rotary/Linear Electromagnetic Actuator*. PhD dissertation, M.I.T., Department of Mechanical Engineering, 2005.

[60] H. Mutai and K. Yamsawa. Fundamental operations of a multipolar disk-solenoid. *IEEE Tran. on Magnetics*, 31(4), July 1995.

[61] Ahid D. Nashif, David I. G. Jones, and John P. Henderson. *Vibration Damping-ing*. John Wiley & Sons, 1985.

[62] Y. Okazaki. A micro-positioning tool post using a piezoelectric actuator for diamond turning machines. *Precision Engineering*, 12(3), 1990.

[63] Yuichi Okazaki. Fast tool servo system and its application to three dimensional fine surface figures. *Proc. of ASPE 1998 annual meeting*.

[64] OPAL-RT. Rt-lab: Feature highlights. *www.opal-rt.com*, 2003.

[65] Alan V. Oppenheim and Ronald W. Schafer. *Discrete-Time Signal Processing*. Prentice-Hall, 1998.

[66] S. R. Patterson and E. B. Magrab. The design and testing of a fast tool servo for diamond turning. *Precision Engineering*, 7(3):123–128, 1985.

[67] Edward M. Petrie. *Handbook of adhesives and sealants*. McGraw-Hill, 2000.

[68] F. Pischinger and P. Kreuter. Electromagnetically operating actuator. *US Patent 4 455 543*, 1984.

[69] A.M. Rankers. *Machine Dynamics in Mechatronic Systems - An Engineering Approch*. PhD dissertation enschede, The Netherlands: Univ. Twente, 1997.

[70] J. D. Rasmussen. Dynamic variable depth of cut machining using piezoelectric actuators. *International Journal of Machine Tools and Manufacture*, 34(3):379–392, 1994.

[71] E. Rivin. Properties and prospective applications of ultra thin layered rubber-metal laminates for limited travel bearings. *Tribology Int.*, 16(1):17–25, 1983.

[72] James K. Roberge. *Operational Amplifiers, Theory and Practice*. John Willey & Sons, 1975.

[73] Michael M. Schechter. Fast response multipole solenoid. *SAE paper n0. 820203*, 1982.

[74] P. Schellekens. Design for precision: Current status and trends. *Annals of the CIRP*, 47(2):557–586, 1998.

[75] A. H. Seilly. Helenoid actuations - a new concept in extremely fast actiing solenoids. *SAE paper n0. 790119*, 1979.

[76] A. H. Seilly. Colenoid actuators - further development in extremely fast acting solenoids. *SAE paper n0. 810462*, 1981.

[77] B. Stancil, H. Gutierrez, and P. Ro. Design of a long range fast tool servo system using magnetic servo levitation. *Proceedings from ASPE 1995 Annual Meeting*, pages 301–304, 1995.

[78] Philip A. Studer. A practical magnetic bearing. *IEEE Tran. on Magnetics*, 13(5):1155+, 1977.

[79] M. W. Todd and J. F. Cuttino. Development of a long range, traction drive fast tool servo for diamond turning applications. *Proc. of ASPE 1997 annual meeting*.

[80] M. Tomizuka and C. Kempf. Design of discrete time repetitive controllers with applications to mechanical systems. *Proceedings of the 11th IFAC World Congress*, 3:243–248, 1987.

[81] H. Tran and D. DeBra. Design of a fast short-stroke hydraulic actuator. *Annals of the CIRP*, 43(1):469–472, 1994.

[82] David L. Trumper. *Magnetic Suspension Techniques for Precision Motion Control*. PhD dissertation, M.I.T., Department of Electrical Engineering and Computer Science, 1990.

[83] X. Wang. Experimental research on the linear motor micro-feed devices with high frequency response, long travel and high accuracy. *Annals of the CIRP*, 40(1):379–382, 1991.

[84] M. H. Weck. A new hybrid concept for a long stroke fast tool servo system. *Proc. of ASPE 1995 annual meeting*, 1995.

[85] Mingchih Weng. *Magnetic Suspension and Vibration Control of Flexible Structures for Non-contact Processing*. PhD dissertation, M.I.T., Department of Mechanical Engineering, February 2000.

[86] Herbert H. Woodson and James R. Melcher. *Electromechanical Dynamics*. John Willey & Sons, 1968.

[87] Minggao Yang. Study of the injection control valve in a new electronic diesel fuel system. *SAE paper n0. 980813*, 1998.

[88] J. G. Zaremba. A biaxial fast steering mirror for precision optical pointing systems. *AIAA Guidance, Navigation and Control Conference, Minneapolis, MN, Aug. 15-17, 1988*, pages 471–478, 1988.

[89] Yu Zhou, Maarten Steinbuch, and George Leenknegt. A cost-effective scheme to improve radial tracking performance for higher-speed optical disk drives. *Journal of Vibration and Control*, pages 795–810, 2004.