THE DESIGN OF A MICROMACHINING INSTRUMENT: PHASE ONE

By

PAUL NEAL MOSS

Bachelor of Science

Oklahoma State University

Stillwater, Oklahoma

1982

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
December, 1993

OKLAHOMA STATE UNIVERSITY

THE DESIGN OF A MICROMACHINING INSTRUMENT: PHASE ONE

Thesis Approved:

Thesis Adviser

Dean of the Graduate College

ACKNOWLEDGMENTS

I wish to express my sincere thanks to all the members of my committee for their help and guidance. Sincere appreciation to Dr. Lucca for providing the opportunity to gain experience in a number of areas and for providing an environment which allowed me to grow personally. Thanks to Dr. Price for always being available to listen and to Dr. Misawa for help with the control system.

Charles Babcock and Yongwei Seo also deserve special mention for their help, both moral and technical, on those long dreary nights spent in the lab.

My deepest appreciation goes to my wife, Janelle. Without her companionship and unwavering support I would not have even taken up graduate school, much less taken up this project.

TABLE OF CONTENTS

LIST OF TABLES

Table Page

LIST OF FIGURES

LIST OF SYMBOLS

A	Area (cm^2 or in^2)
C	Capacitance (μF or nF)
d	Resolution (m)
d_c	Tool Post Cantilevering (nm)
d_{33}	Ratio of Strain to Electric Field parallel to poling direction (cm/V)
E	Modulus of Elasticity (MPa)
f	Frequency (Hz)
f_o	Resonant Frequency (Hz)
F_c	Cutting Force (N)
F_t	Thrust Force (N)
h	Sample Thickness (mm)
i	Electric Current (mA)
k	Stiffness of a structural Element (N/μm)
k_m	Material Constant = $4\pi E/3(1-\nu^2)$ (MPa)
k_o	Stiffness of PZ Stack (N/μm)
K_p	Proportional Gain
K_I	Integral Gain (1/s)
l_{bs}	Bolt Spacing (cm)
m	Mass (kg)
P_{cl}	Clamping Pressure (MPa)
P_{out}	Power Consumed by PZ Stack (W)

r_o Outer Radius of Circular Flexure (cm)

r_i Inner Radius of Circular Flexure (cm)

r_l Length Ratio

t Clamp Thickness (cm)

U Voltage Amplitude (V)

U_o Maximum Output Voltage of PZ Stack Driver (V)

δ Deflection (nm or μm)

ξ Aspect Ratio

ε Motional Error (μm)

λ Wavelength (μm)

κ Rotational Stiffness (μrad/Ncm)

Δ_t Total Off-Axis Tool Deflection (nm)

ν Poisson's Ratio

μ Coefficient of Friction

σ Stress (MPa)

σ_i Maximum Stress at Outer Radius of Circular Flexure (MPa)

σ_o Maximum Stress at Inner Radius of Circular Flexure (MPa)

σ_{ys} Uniaxial Yield Stress (MPa)

CHAPTER I

INTRODUCTION

1.1 Introduction

This chapter sets the stage for the design of the Micro-Machining Instrument (referred to as the MMI from this point forward) by reviewing some general technological trends and discussing some of the applications of ultraprecision machining. An outline of some of the current issues in specific areas will also be presented.

1.2 Ultraprecision Machining

The meaning of the term 'ultraprecision machining' tends to vary over time. It does generally refer to material removal techniques that are at the forefront of the current state of the art. However that definition encompasses a historically very broad range of processes and measurement precision. Indeed it has been claimed that in the last 200 years measurement precision in machining processes has improved by five orders of magnitude and that it quite likely will improve another three or four orders of magnitude [McKeown, 1987].

The processes are varied as well. In addition to processes which are on the surface similar to traditional machining, such as single point diamond turning, the term 'ultraprecision machining' also encompasses nontraditional processes such as laser machining, ion beam machining and molecular machining which makes use

of scanning tunneling processes to remove and place material on the atomic scale. Figure 1.1 shows the Taniguchi curve that illustrates some of the more salient points on the recent historical evolution of ultraprecision machining.

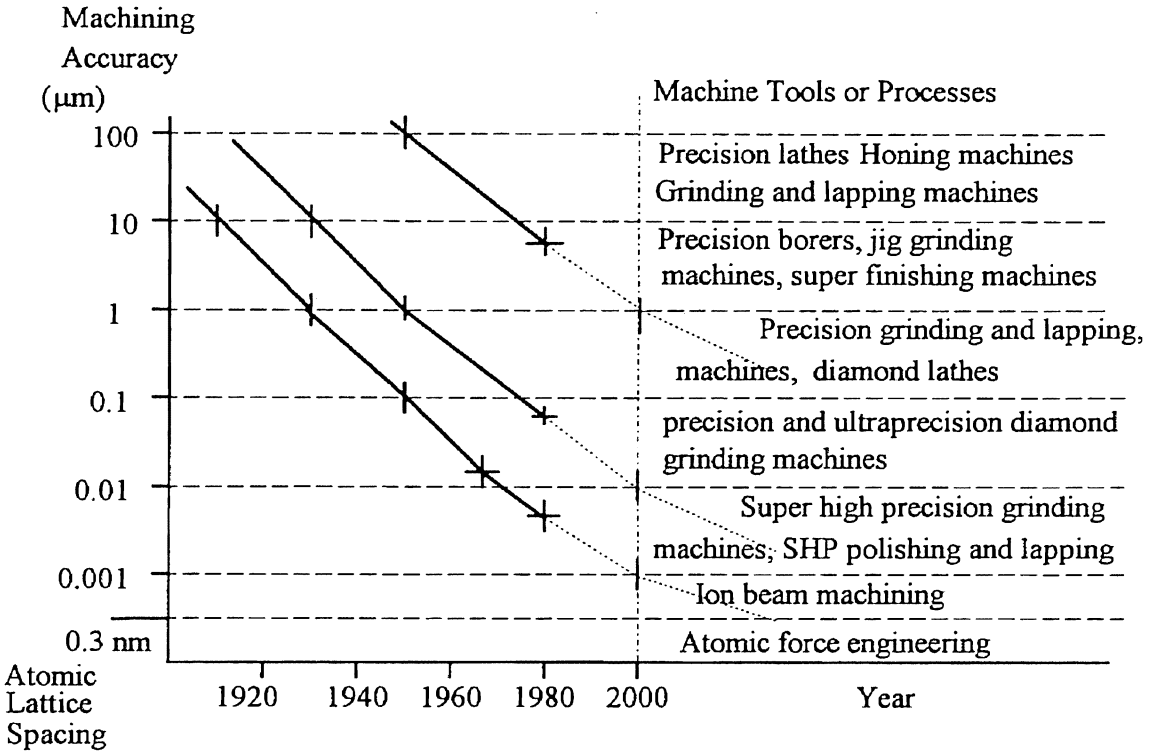


Figure 1.1. Development and Near Future of Machining Accuracy (After Tanaguchi and McKeown)

The curves show that what is considered ultraprecision machining varies with time and that current ultraprecision machining is the precision or normal machining of tomorrow. This trend leads to Taniguchi's definition of ultraprecision engineering (since ultraprecision machining requires ultraprecision engineering technology their respective distance scales coincide) as "the highest possible dimensional accuracy [that] is, or has been achieved at a given point in time" [Tanaguchi, 1983] (bracket added). Thus a quantitative definition can be set with the understanding that it is tentative and will almost certainly change in the

foreseeable future. Ultraprecision machining is then the achievement of overall dimensional tolerances (form accuracy) of less than one micron and average surface roughness on the order of a few nanometers. For the purposes of the present work the emphasis will be on those ultraprecision machining processes that bear a resemblance to traditional machining; that is with a mechanical tool (made of diamond) which 'cuts' the workpiece by straining to fracture a small region of the workpiece as in single point diamond turning. The word 'cuts' is emphasized because there is some debate as to exactly how to characterize the material removal process at these levels of precision and to highlight that this is definitely not traditional machining.

1.3 Application of Ultraprecision Machining

The mention of such small distances in the context of manufacturing can lead questions regarding the practical use for such high precision. Following are some examples of applications of diamond turning which range from specialized experimental equipment to technologies of more common devices.

Research in communications, non-linear optics and fusion power increasingly rely on high power lasers. Astronomical research also depends on high quality optics to gather small amounts of energy from distant objects. Diamond turning of optical surfaces for such advanced laboratory experiments and advanced weapons systems is routinely done. For example single crystals of Lithium Niobate, Calcium Fluoride and other compounds are essential to the operation of high power visible light lasers. The dimensions and surface finish influence the optical properties and the resistance to damage by the laser. Typical precision in machining these crystals is less than a 5μm depth of cut and surface roughness is less than 5 nm rms (as low as 0.75 nm rms for Potassium Dihydrogen Phosphate

crystals). Mirrors for X-ray telescopes use non-traditional optics (such as logarithmic surfaces) and may require form accuracy of less than 2 μm for a 1.5 m diameter mirror and 0.2 nm peak-to-valley surface finish [Ikawa, 1991]. More mundane optical systems require ultraprecision accuracy as well . The molds used to make contact lenses and high quality Fresnel lenses are diamond turned (Fresnel lenses are found in many traffic control lights).

The aluminum substrates of photocopier drums, compact video disks, VCR heads and computer hard disks also require surface finishes with peak-to-valley roughness of tens of nanometers and submicrometer form accuracy. One of the major obstacles to improving information packing density on computers has been the clearance between the read/write head and the hard disk. In the case of the computer hard disk the decrease in surface roughness contributed to the dramatic increase in information density, from about 2000 bit/in^2 in 1960 to over 10^9 bit/in^2 in 1990, by allowing the magnetic head to fly 300 nm above the disk [Ikawa, 1991; McKeown, 1987].

There are applications to more traditional workpieces as well. Over the last few years interest has developed in equipping aircraft turbofan engines with aft-mounted propfans. Rolls-Royce reports that by improving form accuracy to less than 10 μm and surface finish average roughness to less than 100 nm that compressor efficiency will approach 96% . In addition if the form accuracy and positioning accuracy of transmission gearing for these propfans can be improved to less than 1μm the power carrying capacity of the gear train will nearly double [McKeown, 1987].

In addition to assisting with the general evolution of technology there are other sound reasons for studying the machining process. Approximately 80% of the of the parts manufactured in this country require at least some machining [Dieter, 1986]. This contributes to the cost of those parts; in time required, energy spent

and material costs such as tools and equipment. This is one of the reasons the almost universal goal of investigations into cutting has been ultimately the determination of energies associated with the various cutting processes. The overall energy used is important in determining the economics of cutting on the large scale. Also important is the partition of the mechanical energy supplied by the tool; determining local temperature rises, the work done as plastic deformation and the wear rate of the tool. These variables in turn will also directly effect the economics (in terms of tool wear) and the quality of the finished workpiece. Substantial subsurface plastic deformation caused by worn tools is undesirable because it leaves residual stress and can render the surface more chemically active while high temperatures effect material surface properties.

The benefits of increasing the precision of machining processes is presented below [after McKeown, 1987].

1. Improve ease of assembly, especially for automation
2. Improve interchangeability of parts
3. Reduce scrap and time spent on inspection
4. Improve fatigue and wear life
5. Improve packing density and miniaturization

1.4 A Model of Traditional Machining

In the previous sections it was stated that traditional machining is to be distinguished from ultraprecision machining. The rational for this distinction is based on the lack of knowledge of the process physics in ultraprecision machining. Brief and simplified descriptions are presented to justify this distinction.

The standard model for traditional machining of ductile materials is based on an idealized two dimensional geometry in which the tool is considered to be a

single point tool characterized by its rake angle and clearance angle. The line formed by the intersection of the rake and flank faces is perpendicular to the cutting velocity (hence the term "orthogonal machining"). There are a number of mechanisms that dissipate the mechanical energy provided by the tool: (1) Heat is generated by the work done in the shear zone ahead of the tool. (2) Friction develops between the tool and the chip at the rake face. (3) Plowing occurs due to the fact that all tools are three dimensional objects and have a nonzero edge radius. (4) Friction also occurs due to elastic recovery at the flank face/workpiece interface. (5) Momen-tum changes as new material enters the shear zone. (6) New surface area is generated. (7) Energy is required to curl and break chips. The traditional model of orthogonal machining is shown in Figure 1.2.

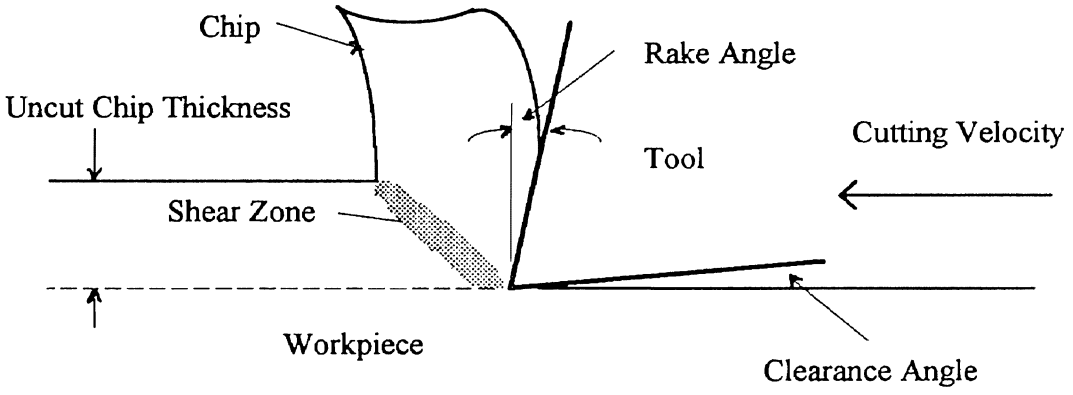


Figure 1.2. Traditional Geometry of Orthogonal Machining

The accepted analyses show that for traditional machining almost all the mechanical energy is dissipated by heat (95% to 99%) and approximately 80% is carried away by the chip. About 75% -80% of the heat energy is generated in the shear zone and 20% to 25% by rubbing at the chip/tool interface [Shaw, 1984]. In an unworn tool for large depths of cut the most important contact length is the contact between the newly formed chip and the rake face of the tool; since this is

where the energetically dominant processes take place this would be the characteristic length defining the scale of the process. More sophisticated models allow for some contact at the flank face due to tool wear. Approximately 5% of the heat energy may be generated due to rubbing between the flank face and the workpiece; but for an unworn tool the length of this sliding face is much smaller than the tool-chip sliding length and would be an even smaller fraction of the total energy. Negligible amounts of energy are partitioned by mechanisms (5)-(7). Another consequence of this energy partition is that the thermal properties of the tool do not have a strong effect on the distribution of energies. The size of the chip is sufficient that most of the energy is carried away by it; the rest passes into the tool and the workpiece.

1.5 A Model of Ultraprecision Machining

This approach is not adequate to model ultraprecision machining however. Even the sharpest diamond tools have an edge radius which can vary from 20 nm to 200 nm [Lucca and Seo, 1993]. This length scale is also in the range of uncut chip thickness found in ultraprecision machining, which varies from a few nanometers to hundreds of nanometers.

Figure 1.3 (on page 8) shows a proposed geometry that may be applicable to these depths of cut. Under this model there is a shift in geometry that in turn leads to a change in energy partition. The elastic recovery beneath the tool along the flank face (which is a small fraction of the depth of cut for traditional machining) may create a sliding length (which is dependent on the uncut chip thickness) that is now on the same order as, or perhaps two orders of magnitude greater than, the depth of cut [Lucca ae. al., 1992]. Flank wear would also contribute to this length.

This redefines the characteristic lengths of the problem. Now instead of the chip/rake face contact being a characteristic length it may be the case that the flank face/workpiece interface defines the characteristic length. The shear zone is perhaps now a large region (compared to the depth of cut) of plastic deformation both in front of and beneath the tool. Strong plowing could occur due to the effectively negative rake angle created by the tool edge radius. The flank face rubbing may also cause additional subsurface plastic deformation in the workpiece.

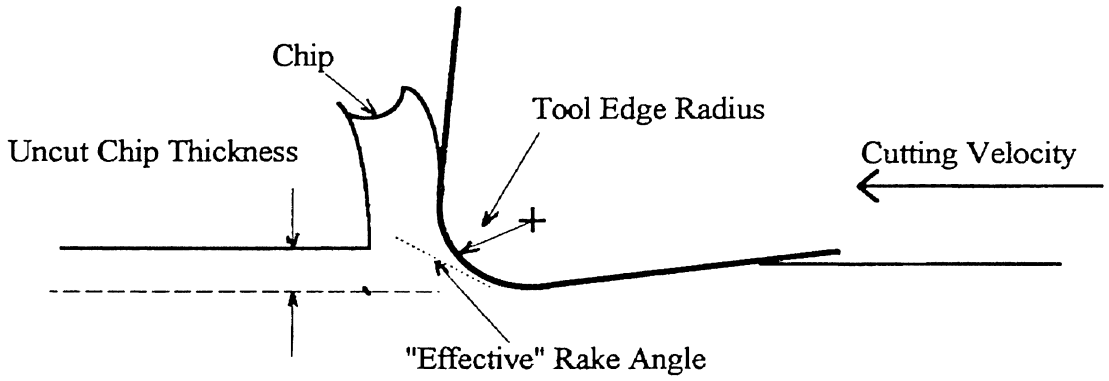


Figure 1.3. Ultraprecision Machining Geometry

These factors lead to a shift in the partition of the energy-the majority (60%) of the heat may now be generated at the tool edge and about 20% at the rake face and the shear zone (although mechanisms (5) -(7) remain negligible) [Lucca and Seo, 1989]. The contribution made by each of these sources to the energy partition depends on the depth of cut and the edge radius, which can change appreciably during the life of the tool [Lucca and Seo, 1993]. The chip is also very small and has a correspondingly small heat capacity; consequently the tool material can have profound effects on local temperatures and heat flow; as much as 90% of the heat generated may flow out through the tool [Lucca and Seo, 1989; Lucca et. al., 1992].

That the process of machining at small depths of cut is quite different from traditional machining is demonstrated by the so-called size effect. As the depth of cut decreases below a certain threshold the specific energy (or specific cutting force) is seen to increase. For cutting with single crystal diamond tools this effect begins to occur as the uncut chip thicknesses approaches a few microns [Nakayama and Tamura, 1968]. As the depth of cut is further decreased by two orders of magnitude the specific energy may increase by about the same amount [Lucca et. al., 1991a; Lucca et. al., 1992]. The ratio of the thrust to cut force reaches a minimum in this same region and increases dramatically as uncut chip thickness decreases. The increasing dominance of the thrust force is seen as evidence of increased plowing and sliding, and increased subsurface stress beneath the tool [Lucca et. al. 1992; Lucca et. al., 1993].

The role these mechanisms play in the partition of energy in ultraprecision machining is still the subject of much research. There are competing explanations for the size effect. It has been argued that it may be due an increase in strength caused by the smaller shear area; there are fewer dislocations in the smaller volume and so the strength would be more near the theoretical maximum [Shaw, 1950].

All the proceeding discussion is based on the assumption of ductile materials such as copper or aluminum alloys. With the advent of an advanced semiconductor industry there has also risen an interest in the machining of these materials. The II-VI group of compounds such as Cadmium Telluride, Cadmium Sulfide and other semiconductor materials such as Germanium are natural candidates for ultra-precision machining due to their common use in advanced optical and electronic devices. However as a class they are considered brittle materials. Interesting and poorly understood phenomena occur as the ultraprecision regime is approached: Germanium, which is normally considered

brittle undergoes a "transition" for critical force and depth of cut and appears to behave as a ductile material and exhibits a pitting phenomenon for which no mechanism has been clearly established [Lucca and Oquin, 1992; Blake et. al., 1988]. For uncut chip thickness down to 100 nm Germanium also does not undergo the transition from cutting force dominance to thrust force dominance. The increase in specific energy with decreasing uncut chip thickness is seen to occur however.

CHAPTER II

EXPERIMENTAL CONCEPT

2.1 Current Experimental Devices

The technology of single point diamond turning with ultraprecision accuracy brings together several different engineering disciplines. There is the mechanical design of the turning machine; but in addition there are now the technologies of displacement and measurement transducers capable of fine resolution. Good examples of measurement technologies are capacitive probes or linear variable differential transformers (LVDT's) which provide resolutions on the order of several nanometers (A leading manufacturer, Burliegh, claims resolutions of 0.1 nm are possible for LVDT's). The laser interferometer can attain resolutions of λ /500; for a laser operating at a wavelength of 500 nm the resolution is then 1 nm (As optics and electronics improve this resolution will improve to λ/4000.) [Slocum, 1993]. These measurement technologies can be coupled with displacement translators of the same resolution. One of the most effective is the piezoelectric transducer (PZT), discussed in Appendix A. The small distances involved also require the use of single or multiple feedback loop positioning systems. Much of this technology could not be integrated without the use of computers. In particular the rapidly expanding power/cost ratio of the desktop personal computer has helped by making it easier to provide high levels of computing power.

There are a number of features most experimental and production devices share. Since the forces and distances are small (on the order of a few newtons and nanometers, respectively) the mechanical structures must be stiff; on the order of hundreds of Newtons per micron [Patterson, 1985; Moriwaki, 1989]. For devices requiring travel larger than several millimeters aerostatic or hydrostatic slides can be used. For displacements on the order of microns down to nanometers, PZT's can be used [Patterson, 1985; Furukawa, 1988; Moriwaki, 1989]. In some cases to minimize thermal drift a feedback controlled thermal environment is provided [Furukawa, 1988; Hosler, 1990].

2.2 Measurement and Inferences

The temperature rise at the tool/workpiece contact is difficult to determine directly. The contact area may be on the order of hundreds of nanometers to a few microns; the wavelength of near infrared radiation is from (approximately) 0.9 μm to 5 μm. Since the region of interest is smaller than the wavelength of light emitted it is impossible for any diffraction limited device (such as an IR spot microscope) to detect. It might be possible to use near field technology to map thermal profiles of tools or workpieces. These instruments work on the principle that if the image is formed near the object (near meaning separation less than the wavelength of light used) then diffraction is no longer a limiting factor. In the case of infrared light this requires locating the aperture about 1 μm away from the object. Such positioning requires the use of the control comparable with scanning tunnelling microscopy, which is a related technology. Near field infra-red microscopes are still under development. The diamond turning machines generally used in this type of experiment have rotating pneumatic or hydraulic spindles which do not provide adequate space to mount small delicate instruments which

might be used for observation extremely near the workpiece or the tool. Since the temperatures occur over such small regions for even moderate speeds they are true flash temperatures and decay to ambient in very brief times [Gecim and Weiner, 1991]; by the time the temperature is registered by a detector it is too small to be considered the temperature of the contact area. Thus even if the detector could register the radiation it would probably not be an accurate measurement.

The approach this group has taken is to attempt to determine the regions of mechanical dissipation of energy and relate the mechanical energy to the heat and therefore the temperature. The types of data taken are measurements of the thrust and cutting forces from which the energy expended from work done by the cutting force can be calculated. As a part of this effort a technique for inferring the sliding length using wear of thin films deposited on the tool was developed [Lucca et. al 1992]. Also this group has measured the tool edge radius (as a profile) as the tool wears using an atomic force microscope [Lucca and Seo, 1993].

There are efforts to expand the kinds of information that can be used to infer energy dissipation. In particular the region of the flank face/workpiece contact is receiving additional attention. One approach has been to model this contact as a thermo-elastic or thermo-elastoplastic sliding indentor and calculate the distribution of equivalent stress in the workpiece [Lucca et. al., 1992; Lucca et. al., 1993]. In addition to the determination of sliding length by the methods previously mentioned attempts have been made to infer the stresses and directly measure the amount of subsurface plastic deformation [Nakayama and Tamura, 1968]. Investigations have been made to study the regions around the tool/work-piece interface directly while cutting is taking place using photoelastic analysis [Ramalingam, 1971; Bagchi and Wright, 1986]. However in these last two references the experimental scale was that of conventional machining. In particular Nakayama and Tamura had minimum uncut chip thicknesses of

approximately 2 μm and Bagchi and Wright used sapphire tools and uncut chip thicknesses of several hundred microns, well above that found in ultraprecision machining .

2.3 The Proposed Experiment

It has previously been noted that semiconductor materials are receiving more attention from the diamond turning community. These materials may offer an opportunity to study the machining process in a new way. The transmission spectra of some semiconductor ultraprecision machining candidates cover a wide range; some are opaque to visible light (400nm to 850nm in wavelength) but many are transparent in infrared light. The three semiconductor materials considered here are Cadmium Telluride, Cadmium Sulfide and Germanium; their transmission spectra are shown in Figure 2.1. The fact that they are transparent (for at least some range of the electromagnetic spectrum) invites the use of photoelastic techniques, in this case observing the workpiece instead of the tool. However, the phenomena to be observed occur on a much smaller scale than previous photo-elastic machining studies.

Some preliminary estimates can be made to determine the likelihood of detecting a fringe pattern set up in a photoelastic experiment. Two things are needed to make the estimate: (1) the optical properties of the material must be known and (2) an estimate of the stress field must be made. In contrast to the standard model materials used in photoelastic studies the single crystal semiconductor materials considered here are all anisotropic to some degree (See Table 2.1). The two practical consequences of this are that the photoelastic properties and behavior will be dependent on orientation and the crystals may

exhibit natural birefringence (as is the case for CdS). Naturally birefringent crystals become biaxial when subject to an arbitrary stress.

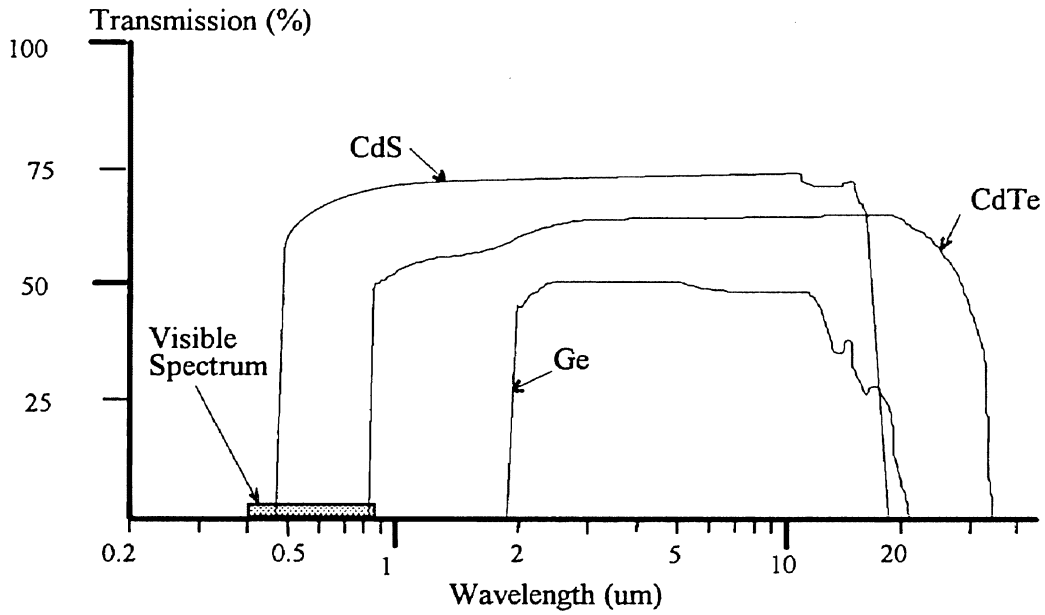


Figure 2.1. Transmission Spectra for CdTe, CdS and Ge

To obtain the optical material constants two different approaches must be related. The photoelastic constant most useful for an analysis of stress fields is the material fringe constant $f = \lambda/c$ where λ is the wavelength and c is the relative stress-optic constant. Generally however birefringence for these materials is measured as a change in index of refraction (δn) and non-uniform stress fields (and therefore fringe patterns) are deliberately avoided; the constants thus obtained are then related to physically more fundamental properties such as (electron) oscillator strength or valence band behavior. The two approaches can be related as follows. The standard photoelastic equation for the angular phase difference (Δ) between polarized light waves in a material subject to arbitrary stress depends on the three principle stresses. In the special case of light propagating parallel to one of these principle stresses (defined as σ_3 here) the equation reduces to the same

equation as for plane stress and for a sample thickness h is given by [Dally and Riley, 1991]

$$\Delta = \frac{2\pi hc}{\lambda}(\sigma_1 - \sigma_2) = \frac{2\pi h}{f}(\sigma_1 - \sigma_2) \tag{2.1}$$

In terms of the change in index of refraction (δn) the birefringence is

$$\delta n = Kh(\sigma_1 - \sigma_2) \tag{2.2}$$

where the form of the equation for the constant K depends on the particular physics under investigation. Since the angular phase difference is given by $\Delta = 2\pi h\delta n/\lambda$ the measured constants given by equation (2.2) can be related to the photoelastic properties of equation (2.1).

For diamond cubic symmetry (found in Germanium and Cadmium Telluride) there are a minimum of three constants necessary to completely characterize the optical properties. For hexagonal Cadmium Sulfide there are five constants. However for stress birefringence only one constant is required if the stresses are parallel to the <111> or <100> directions in the case of the cubic materials or if the stress is parallel to the c-axis for hexagonal materials (which then act as a uniaxial material) [Babonas, 1973]. For the values of the material constants see Table 2.1.

In order to predict the likelihood of seeing fringes an estimate of the stress field beneath the tool must be made. To this end a plain strain indenter of width 2L acting on an infinite half space is chosen as shown in Figure 2.2.

The stress fields for a contact pressure P_0 are given by [Y. Seo, 1993]

$$\sigma_{xx} = \frac{-2P_0}{\pi}\int_{-L}^{L}\frac{(x-\xi)^2 y}{[(x-\xi)^2 + y^2]^2}d\xi \tag{2.3}$$

$$\sigma_{yy} = \frac{-2P_o}{\pi}\int_{-L}^{L}\frac{y^3}{[(x-\xi)^2 + y^2]^2}d\xi \tag{2.4}$$

$$\sigma_{xy} = \frac{-2P_o}{\pi}\int_{-L}^{L}\frac{(x-\xi)y^2}{[(x-\xi)^2+y^2]^2}d\xi \qquad (2.5)$$

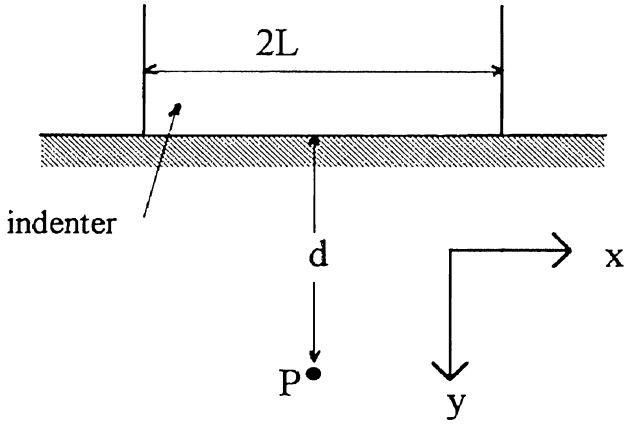


Figure 2.2. Indenter Geometry.

To avoid the singularity at the contact plane let $y \neq 0$. To evaluate the integrals make the substitution $\gamma = x\text{-}\xi$ and then $\gamma = y\tan\theta$. After carrying out the integration set $x = 0$ to obtain the stress directly beneath the indenter. Noting that under these conditions $\sigma_{xy} = 0$ and calling $\sigma_{xx} = \sigma_1$ and $\sigma_{yy} = \sigma_2$ the difference in principle stresses becomes

$$\sigma_1 - \sigma_2 = \frac{4P_0}{\pi}\frac{Ld}{L^2+d^2} \qquad (2.6)$$

If the indenter size L is small compared to the sample length and width then for points located under the indenter near the center of the sample a finite sample can approximate an infinite half space. Extinction occurs for $\Delta/2\pi = n$ (an integer); then in terms of the fringe order (n) for which extinction occurs equation (2.1) is given by

$$n = \frac{h}{f}(\sigma_1 - \sigma_2) \qquad (2.7)$$

By combining equation 2.6 and 2.7 the depth y_n below the indenter at which a particular fringe of order n will appear is then

$$y_n = \frac{1}{2} L(\beta_n \pm \sqrt{\beta_n^2 - 4}) \tag{2.8}$$

where $\beta_n = 4P_o h/\pi n f$ and h is the sample thickness.

The requirement that the distances be real imposes a condition on β_n and hence on the contact pressure. For the observation of fringes of order n for a given sample thickness the minimum contact pressure is

$$P_o^{min} = \pi n f/2h \tag{2.9}$$

A number of empirical constants are also required to carry out the estimate. They are the yield strength (just to get a rough idea of the stresses that can be applied), the material fringe value, the indenter length and the wavelength to be used. In each case a wavelength near the transmission edge close to visible light is chosen because the material fringe value depends on wavelength and it is assumed that the optics will perform better the closer the light is to the visible (See Appendix B). For the purposes of this estimate the contact stress is chosen as three times the maximum uniaxial yield stress. The results are summarized in Table 2.1.

The symmetry values for the cubic materials were calculated using the compliances (s_{ij}) with the formula $2(s_{11} - s_{12})/s_{44}$. Two values are needed to describe the symmetry of hexagonal materials. The elastic constants (c_{ij}) were used according to the formulas $(c_{11}-c_{22})/2c_{44}$ and c_{33}/c_{11}, and recorded in this order for CdS in Table 2.1 [Bolef, et. al., 1960].

The yield stresses were taken from the same sources that provided the optical constants. The samples used in these experiments were not generally compressed to gross yielding so that the stresses entered in the table are probably conservative. No yield stress for the <100> direction was listed for Germanium due the high

value of the fringe constant for this axis. According to equation (2.9) the yield stress would need to be over 300 MPa to allow the observation of fringes for the samples sizes considered. The yield stress reported for CdS is parallel to the c-axis. The material fringe values are for stress parallel to the <111> axis and <100> axis for the cubic materials; for CdS the stress is parallel to the c-axis. For the indenter geometry these two axes then would be parallel to the y-axis.

The entries y_1, y_2, y_1* and y_2* are the depths below the indenter for a given fringe order. The subscript on the y entries indicates the fringe order. The entries y_1 and y_2 are the a sample thickness of 1 mm, while y_1* and y_2* are for a sample 0.5 mm thick. For CdTe only the first order fringe for stress parallel to <100> direction in 1 mm thick sample was predicted to exist. The 0.5 mm sample is too thin for a first order extinction. For Ge for stress parallel to the <111> axis first order fringes exist for a 1 mm thick sample, but not for a 0.5 mm thick sample. The second order fringe for the 1 mm thick sample could probably not be resolved by diffraction limited optics because its spacing from the indenter is less than the wavelength of light used. For resolution equal to 0.6 wavelength (See Appendix B) the fringes for CdTe might be resolvable. For stress parallel to the c-axis first and second order fringes are resolvable in CdS for 1 mm and 0.5 mm thick samples.

One fact that must be kept in mind in evaluating this approximation is the natural limitation of the optics used to magnify the image. For example there may be a large number of closely spaced fringes near the corners of the indenter due to the singularity there. However the best resolution any traditional magnifying device can attain is in principle limited by diffraction to about half the wavelength of the light used. Fortunately virtually all research grade microscopes perform very close to this limit. However there would therefore be little chance of seeing fringes at the corners particularly for the long wavelength chosen for Germanium.

For each of the cases above, except where noted, the wavelengths chosen are sufficient to see the fringes at the predicted position and spacing. It appears that there is a reasonable chance of seeing fringes for at least Germanium and Cadmium Sulfide.

Property	CdTe	Ge	CdS
λ (μm)	1.2	2.2	0.6
Symmetry	2.12	1.66	1.08/1.11
σ_{ys} || <111> (MPa)	20	50	NA
σ_{ys} ||<100> (MPa)	20	—	NA
σ_{ys} || c-axis (MPa)	NA	NA	250
$f_{<111>}$ (10^4 N/m)	3.7	6.3	NA
$f_{<100>}$ (10^4 N/m)	2.6	59	NA
$f_{c\text{-}axis}$ (10^4 N/m)	NA	NA	2.9
y_1 (μm)	1 <100>	14 <111>	25
y_2 (μm)	—	1.5 <111>	13
y_1* (μm)	—	—	13
y_2* (μm)	—	—	7
Source (optical)	Wardzynski, 1971	Schmidt-Tiedemann, 1961	Babonas, 1973
Source (mechanical)	Wardzynski, 1971	Schmidt-Tiedemann, 1961	Babonas, 1973

Table 2.1. Summary of Photoelastic Model Results

2.4 Experimental Concept

The fundamental experimental concept is shown in Figure 2.3 (following page). Note that this figure establishes a directional convention to used consistently throughout.

The general specifications are that the MMI should be able to machine over approximately the same range as a diamond turning machine. The range of motion parameters are initially set then to provide a depth of cut (uncut chip thickness) of 5 μm to 0.020 μm. This means the tool should be capable of motion in the y-direction, in 0.020 μm (20 nm) increments, up to several microns. The actual accuracy and repeatability of the tool position will be determined by the gauging technology used and should clearly be less than the specified depth of cut. Capacitive gauging is to be used for the MMI and can attain resolutions of a few nanometers. The sample will need to move about 250 μm in the x-direction, at a range of velocities from 0.1 mm/s to 10 mm/s. During cutting the tool is fixed with respect to the lab frame. The workpiece is constrained to move in the x-direction only. Cutting is thus accomplished by pushing the workpiece across the stationary tool at the appropriate depth of cut. Since it will be impossible to tell when the tool first contacts the surface of the sample the tool will have to move in steps at the end of each stroke of the sample. This process will also help show consistency and repeatability.

An optical path will be provided for transmission or reflection microscopy utilizing filters and polarizers (visible and infra-red) and associated optics for studying stress birefringence. In addition, the tool/workpiece mechanism will have to allow for the positioning of high magnification microscope objectives near the tool/workpiece interface. To image properly such objectives may need to be as close as 200 μm away from the point where the tool touches the workpiece.

Hence there may be no structural features directly above or below the tool or workpiece.

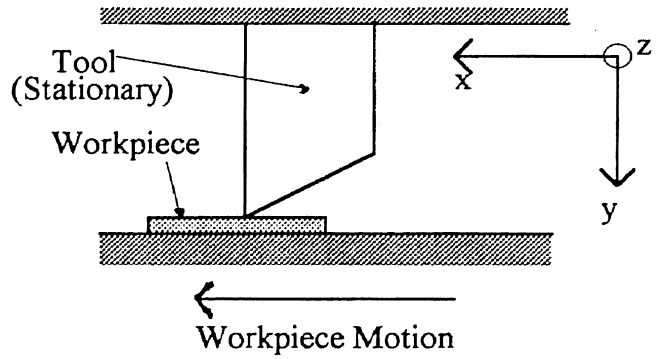


Figure 2.3. Experimental Concept Geometry

Comparison with fast tool servo designs indicate the static stiffness should be in the range of 50 N/μm to 100 N/μm [Patterson, 1990; Hosler, 1990]. To obtain the maximum possible stiffness the position of the tool should be controlled by a closed loop feedback system. Ideally, the system should sample and compensate at a length of cut equal to the depth of cut. However, for the smallest depth of cut at the highest speed of cut this is in the range of 200 kSamples/s. This will almost certainly turn out to be optimistically high. More realistic sampling rates will likely be several hundreds of Samples/s [Hosler, 1990]. Finally the forces generated in machining will need to be measured, using a dynamometer.

2.5 Inventory of Major Components

Following is a list of the major components used to put the overall experiment together. The list does not include items manufactured specially for this project.

Optical Components

1) Nikon MM-11U Measurescope

The microscope is fitted with a high capacity x-y stage that has

resolution of 1 μm. The objectives are of the metallurgical type with an additional super long working distance 40X objective. With the 100X metallurgical objective the maximum magnification through the 10X eyepieces is 1000X. The nosepiece is equipped with Wallaston prisms to provide Differential Interference Contrast (Nomarski) capabilities. For a discussion of optical specifications see Appendix B.

2) Hamamatsu C3500 ICCD Video Camera

This is a camera for which the CCD array is optically coupled to an intensifier. It make detailed images in very low light levels (even in starlight). In addition it has low lag time so that moving objects can be observed. Cameras of this type are also easily adapted to computer imaging and graphics enhancing equipment.

3) Nikon FX-35DX still camera

This is a 35 mm automatic camera with remote actuation capability and film sensing and compensating ability.

Positioning Components

1) Physik Instrumente P-243.20 Piezo Translator

The P-243.20 is a multiple disk translator with a stiffness of 420 N/μm, blocking force of 20,000 N and maximum expansion (under a -1000 V bias) of 20 μm. The nominal capacitance is 0.52 μF. The resonant frequency is 8 kHz and the coefficient of thermal expansion is 0.5 μm/°C.

2) ADE SC3204 Cylindrical Capacitance Gauge Probe and Module Board

This position detector uses changes in capacitance to sense displacement. This model has a resolution of 0.5 mV which corresponds to a displacement of 0.5 μin (12.5 nm). The maximum

range is ±0.005 in (125 μm), bandwidth is 40 kHz, linearity 0.1%.

3) Kepco BOP 1000M Power Supply

The high voltage bipolar power supply is for the P-243.20. It can be operated as a voltage supply (±1000 V) or as a current supply (±40 mA). The bandwidth is 1 kHz with a 6 V/μsec slew rate. Maximum power is 40 W. This power supply can also be operated remotely.

Computer and Data Acquisition Components

1) Compaq Deskpro 50M 486DX50 PC

This is an EISA bus computer with 16 Mbyte ram and 500 Mbyte hard drive.

2) Nicolet Pro 10 Digital Oscilloscope

This multichannel digital oscilloscope can be controlled either by a GPIB interface or a virtual instrument driver.

3) National Instruments AT-MIO-16F-5 Data Acquisition Board

This device contains a 12-bit ADC with a 200kS/s sustained sample rate with multiple analog and digital input/output ports. It also has 16-bit MDA capability and is integrated with the LabVIEW for Windows programming language.

4) National Instruments LabVIEW for Windows Graphical Programming Software

This graphical programming language and development system allows the writing of data acquisition programs, analysis and output routines by visual interface instead of traditional computer code.

2.6 Preliminaries to Design

Once the performance of the MMI is established the mechanical devices capable of giving the desired performance have to be settled upon. It is tempting to think that the control system can compensate for deflections and that therefore the stiffness is not critically important. However the control system acts only along one axis and it is necessary to minimize off-axis deflections as well. Therefore the standard design practice of maximizing the stiffness of the structure before any consideration is given to the control system was followed. As the MMI is to be essentially a stage for moving the tool and the sample; there are several devices which might suffice: 1) Linear stages; 2) Monolithic flexures; 3) Clamped flexures. Before a device is designed and machined the possibility of off-the-shelf components must be investigated; hence linear stages were investigated first.

2.6.1 Linear Stages. Linear stages come in a wide variety of configurations; there are mechanical stages which employ bearing surfaces or pneumatic or aerostatic stages. The simplest of the mechanical type are essentially two plates that slide relative to one another on a straight gib. For precision applications stages employing roller bearings are very common. The mechanical stages with the lowest noise and highest stiffness are the crossed roller bearing type. Since all mechanical stages have a clearance between moving surfaces and this clearance cannot be decreased without damaging the stage, the stage must have the requisite stiffness built into its mechanism. Stages using some type of fluid support (such as compressed air or hydraulic fluid) are capable of the highest stiffness.

Since x-y motion is necessary the stages would need to be mounted together (stacked) with the machining apparatus mounted on top of the stage assembly. High quality translation stages usually come with a set of stiffness specifications, given in two categories: linear stiffness and rotational stiffness (sometimes

compliance is used instead of stiffness). The geometry of such specifications is shown in Figure 2.4.

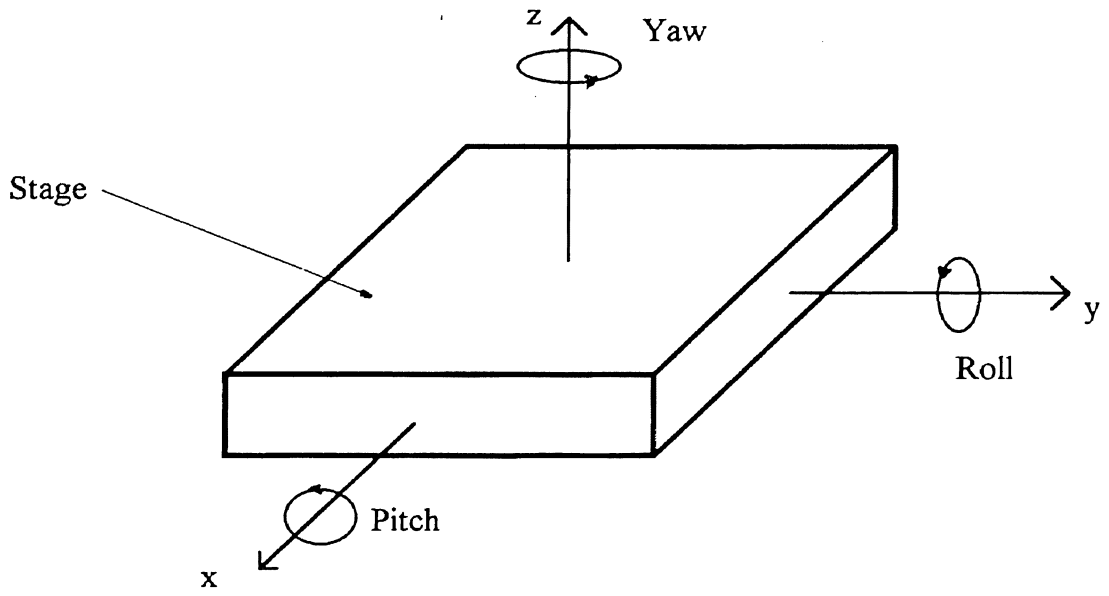


Figure 2.4. The Six Degrees of Freedom for Linear Stages

Two stacked stages can be modeled in two dimensions as shown on the following page in Figure 2.5. In this figure the equilibrium position of the stages with no force acting is shown as a dotted line. The equilibrium position of both stages when a force is applied (a distance l_1 above the center of gravity above stage 1) is show by dark lines. The compressed rectangle between the stages represents the linear stiffness of the two stages parallel to the applied force. In terms of the pitch and roll stiffness of the two stages (κ_{1p}, κ_{2p} and κ_{1r}, κ_{2r} respectively), the linear stiffness and the expected dimensions of the stage and tool holder the deflection can be shown to be approximately

$$D_y = [\kappa_{1p} l_1^2 + \kappa_{2r}(l_1 + l_2) l_2 + \frac{k_1 + k_2}{k_1 k_2}]F \qquad (2.8)$$

The approximation is the assumption that the angles of deflection are small enough that the linear displacements are proportional to the angle. Also it is assumed that the stages themselves do not deform.

High quality crossed roller bearing slides such as those made by Klinger-Newport typically have rotational stiffness on the order of 0.1 μrad/Ncm and linear stiffness around 50 N/μm. For a pair of stacked stages the model predicts deflections on the order of 150 nm for a force of 7 N.

These deflections are two orders of magnitude greater than the resolution necessary. This configuration also demonstrates one of the most serious pitfalls of precision instrument design: Abbe error. By allowing the motion to take place parallel to but not colinear with the force the deflections are amplified. Air bearing linear slides are probably the stiffest off the shelf linear slides available with stiffness in excess of 100 N/μm. However these slides are too large to fit under the microscope used for this experiment. Hydrostatic linear slides can be 10 times stiffer than aerostatic slides but generally must be custom designed [Slocum,1993].

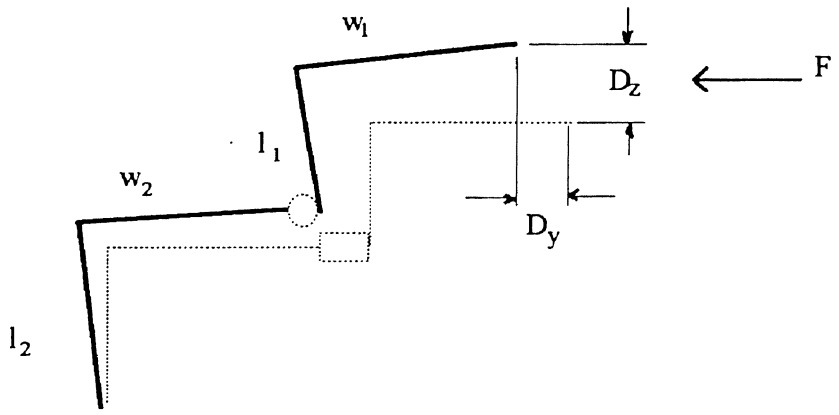


Figure 2.5. Deflection Model for Stacked Linear Stages

2.6.2 Flexures. The alternative is the use of flexures. Rather than depending on rolling elements or fluids to provide relative motion between two bodies, flexural

bearings have no sliding surfaces and depend on the elastic properties of the stage material itself. The advantages of flexures are absence of friction, high off-axis stiffness, low mechanical noise and relative ease of manufacture. There have been flexural bearing stages designed that are capable of subnanometer resolution [Scire, 1978]. There are two basic types of flexures: monolithic and clamped.

Monolithic flexures are made from a single piece of material. A monolithic flexure acting as a cantilever is shown in Figure 2.5. Although capable of high resolution their range of motion is limited to less than 1/100 their size (size here meaning the radius of the hinges). Consequently they are often used in complex arrangements that provide mechanical amplification [Scire, 1978]. In addition to provide the highest possible performance the required machining tolerances are quite close-requiring hole location tolerances of less than 10 μm [Scire, 1978]. Since these complex shapes are machined from a single block of material (usually by wire EDM) and cannot be modified after manufacture sophisticated design techniques such as FEM are employed. In addition, to control thermal errors the materials used are typically 'exotic' alloys such as Invar. Clearly these flexures are not even a remote candidate for the coarse positioning elements; given the difficulty of design and manufacture and the fact that the MMI is required to have some element of 'upgradeability', these flexures do not seem desirable.

Clamped flexures then seem the most reasonable. Although not as inherently stiff either on or off axis as monolithic flexures they are capable of motions one order of magnitude larger for a given size [Slocum, 1993]. The resulting device has a degree of modularity in that the performance can be changed by varying the thickness of the flexures, or adding tuning blocks to the flexures. It is even possible to cut patterns into the flexures themselves to vary properties for a given thickness [Trylinski, 1968]. An example of a rectangular clamped flexure acting as a cantilever with a tuning block is shown in Figure 2.6. The purpose of the

tuning block is to provide a degree of tunability to offset inaccuracies incurred during manufacture and assembly or variations in spring geometry or strength.

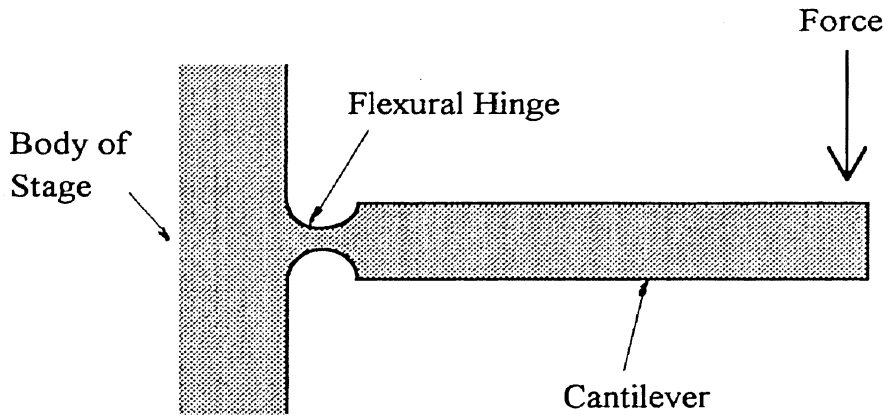


Figure 2.5. Monolithic Flexure

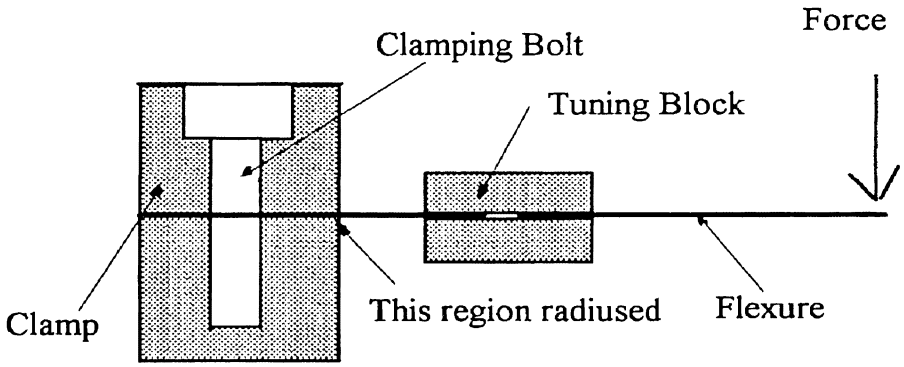


Figure 2.6. Clamped Flexure

CHAPTER III

ELEMENTS OF DESIGN

3.1 Introduction

The design strategy is to first devise an overall conceptual design and generate simple models based on the simple concepts. Next an iterative procedure is applied to verify the validity of the models (at least theoretically) and to derive the necessary dimensions. The objectives of this first phase are:

1) design and construction of an instrument consistent with the specifications on page 21,
2) creation of a control system,
3) testing of one axis (the thrust force axis) by indenting materials with well established photoelastic properties and
4) suggest improvements.

It should be noted that the resolution of the capacitance gauging system will not allow the verification of the first objective. Its resolution is at best 0.5 mV which is equivalent to 12.5 nm. This distance is over 60% of the minimum depth of cut and so could not be considered adequate for the final stage of development.

The design procedure for this type of instrument requires that absolutely no structural property be taken for granted; even the seemingly strongest and stiffest pieces must be carefully examined. One author has advised that the designer proceed as though the design is made of rubber and tabulate every possible error.

After considering several axes the sensitive directions are then determined [Slocum, 1993].

3.2 MMI Configuration

3.2.1 Tool Holder. The essential features of the tool holder assembly are shown in Figure 3.1 which is a side view. The tool holder consists of two concentric cylinders supported at each end by circular clamped flexures. The circular flexures act as flat washers which allow displacement in the axial direction but are relatively stiff along other axes. The flexures also hold the inner cylinder and keep it concentric with the outer cylinder. The outer cylinder is held fixed with respect to the lab frame during cutting and motion of the inner cylinder in the y direction is accomplished by a piezotranslator located at the back of the assembly. The tool is located at the other end of the inner cylinder and is placed so that the thrust force of machining is on the longitudinal (y) axis. The outer cylinder rests on an aluminum frame that allows coarse positioning in the x and y-directions. Both cylinders (and the flexure clamps) are made from 7075-T6 aluminum. The reason for choosing the high strength alloy will be discussed below. The flexures are made from half-hard stainless spring steel. Half-hard sheet stainless steel is half as strong as full-hard sheet (yield strengths of 690 MPa compared to 1380 MPa, respectively), however it is far easier to machine. The manufacture of the flexures require that holes for bolts be accurately spaced. By using the lower hardness material, the extra cost of the manufacture of special workholding jigs (made for this part only) was avoided. The outer rear clamp is integrated into the PZ stack housing. The flexures provide a preload for the stack which is adjustable through the bolt at the back plate. Alignment is provided by two ball bearings which also

prevent the transmission of off-axis forces to the stack. For a discussion of mounting considerations see Appendix A.

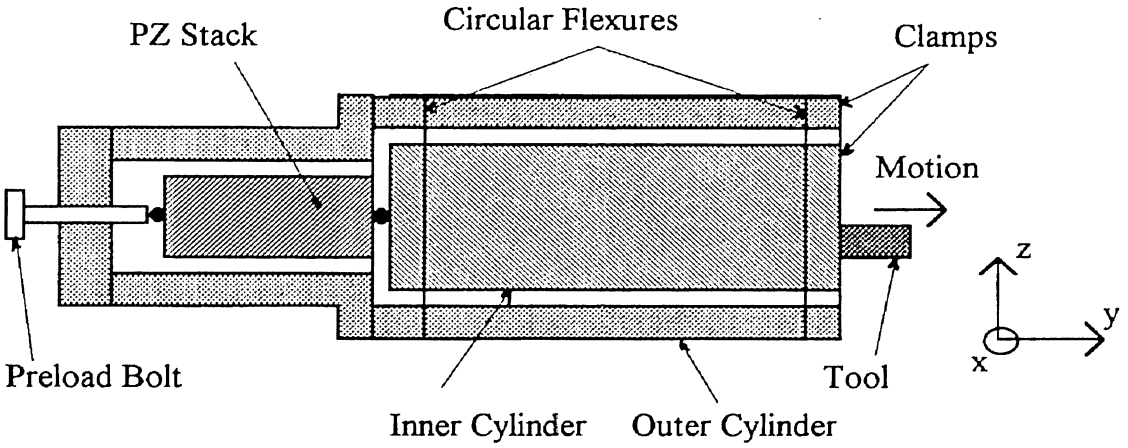


Figure 3.1 Tool Holder Schematic

3.2.2 Sample Holder. The sample holder consists of a small aluminum block which presses on the dynamometer. This assembly is carried on a larger aluminum block (the cross carriage) fixed by rectangular clamped flexures (All aluminum for the sample translating assembly is 6061-T6). A top view is shown in Figure 3.2. and a perspective view is shown in Figure 3.3. The sample is a small coupon approximately 12 mm × 15 mm × 0.5 mm. The tool is positioned against the 0.5 mm face to provide plain strain conditions (See Figures 3.4 and 3.5.). Sufficient clearance must be left for an objective to be positioned close to the sample, hence the stepped shape of the sample holder. Note that in Figure 3.2 an asterisk locates the optical axis which is perpendicular to the plane of the page.

The rectangular flexures are also made from half hard stainless steel spring steel. The sample holder assembly of Figure 3.2 also rests in a heavy aluminum frame that allows coarse movement in the z-direction. The x-direction motion of the cross carriage is to be accomplished by a PZ stack capable of an overall expansion of about 80 μm and fitted with a mechanical amplification system,

yielding the desired 250 μm motion (Note that the construction of this second axis is left to the next stage of the project.). For further notes on this aspect of the design see Appendix E. The mounting of the capacitance gauge is an issue that deserves separate treatment and will be addressed below.

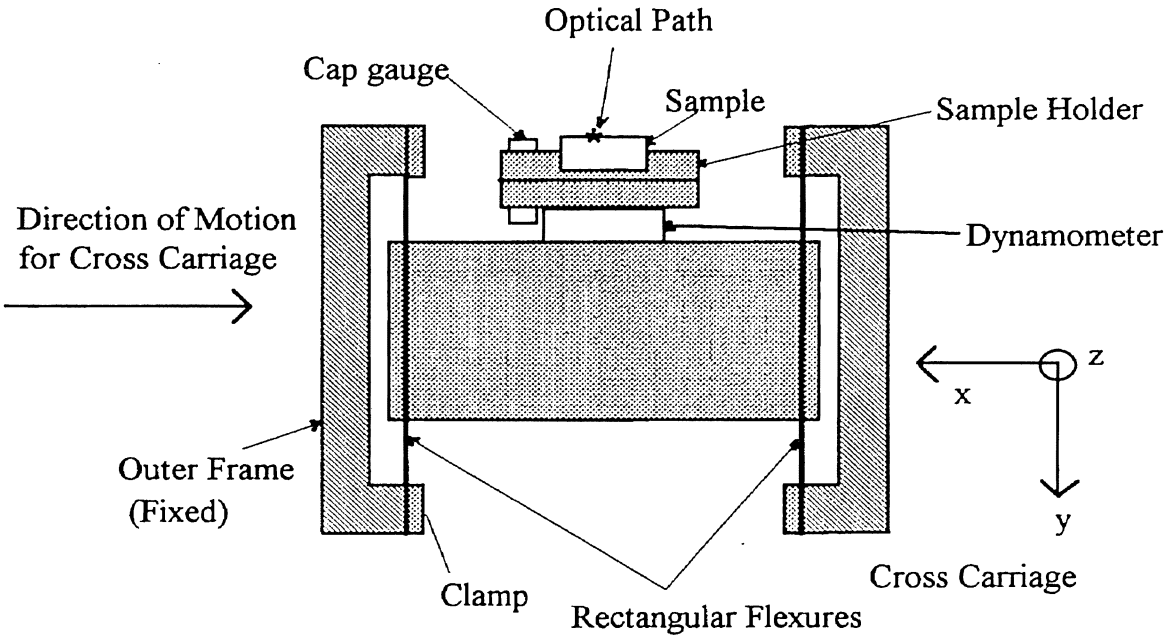


Figure 3.2. Schematic of Sample Holder

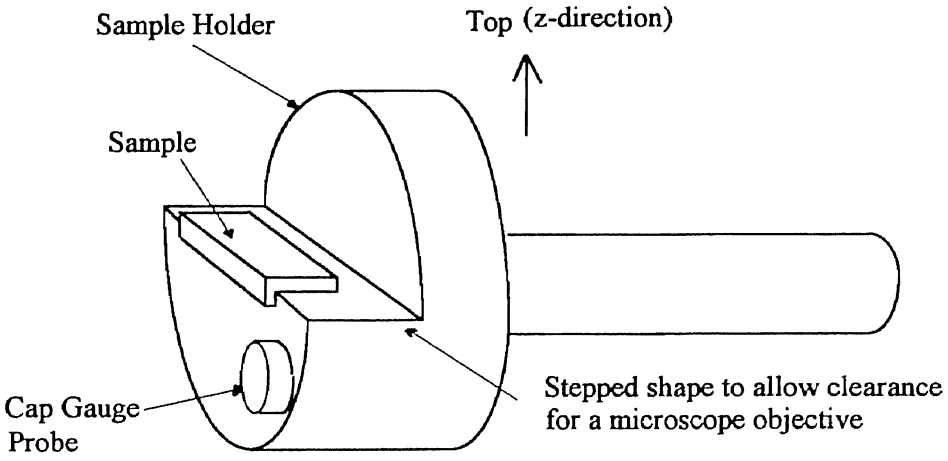


Figure 3.3. Perspective of Sample Holder.

The sample overhangs the holder slightly to allow light to pass through. The relative position of each part is shown in Figure 3.4. Again, the optical axis is perpendicular to the page and passes through the tool/sample contact point. Another view showing how the tool, sample and objective fit together is shown in Figure 3.5.

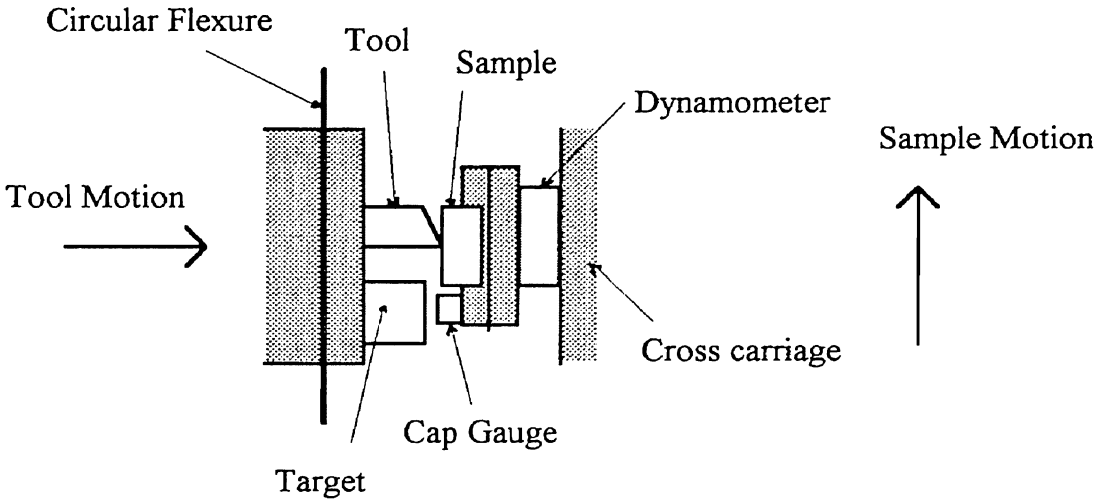


Figure 3.4. Position of Tool and Sample

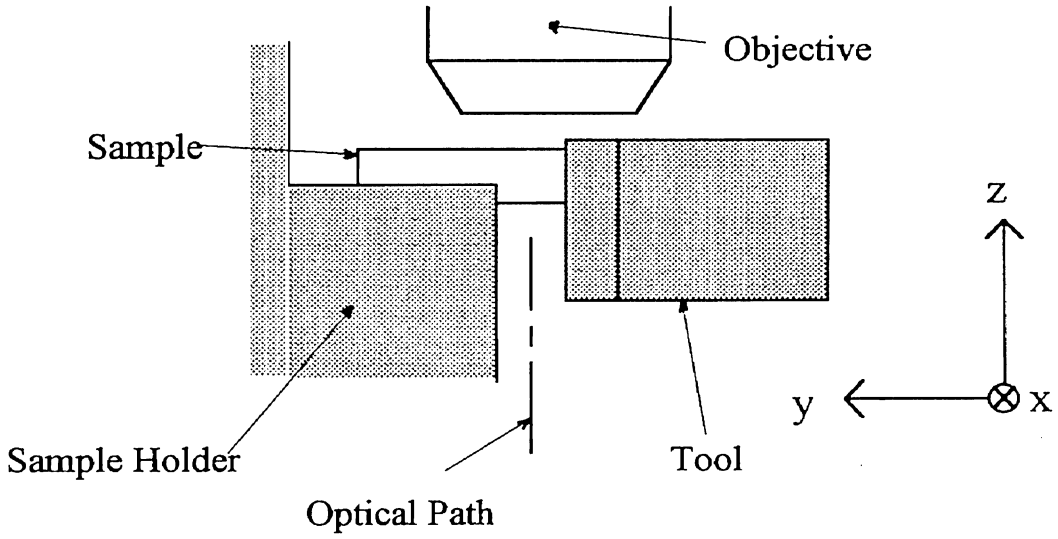


Figure 3.5. Objective in Relation to Sample

3.3 Tool Holder Design

There are a number of parameters that are fixed at the outset, particularly for the tool holder. Since the coarse mechanical positioning of the tool holder frame will allow positioning to within about 10 μm or so of the sample, the PZ stack will have to make up the remaining distance to provide the range of depths of cut. Thus the total expansion of the tool stack will have to be on the order of 20 μm to 30 μm to position the tool. An important constraint is the necessity of fitting the MMI under a microscope. If the outer dimensions are too large the MMI will obstruct the approach of the objective and the objective will not be able to come close enough to the sample/workpiece interface. This constraint limits the outer radius of the outer cylinder to a radius of 4.9 cm, naturally limiting the maximum outer radius of the flexures. The loads expected have been measured. The size of the precision stage which is to support the MMI will limit its overall dimensions as well.

The parameters to be determined are the stress, stiffness (axial and off-axis), off- axis deflections and the proper bolt spacing. Other properties to be determined are the bandwidth of the PZ stack when coupled to the tool holder and the power dissipated under the expected dynamic operation.

For the purposes of design the maximum load per unit width that might be measured during cutting are taken from Germanium data and is expected to be about 7 N/mm (Lucca and Oquin, 1992). Consequently all deflections are calculated based on a load bearing surface of 1 mm width for a force of 7 N (thus overestimating the force and leading to conservative estimates as the sample coupon is 0.5 mm wide).

Consider the tool flexures first. If the maximum performance is to be reached there can be no movement of the clamped part of the flexure. Even the slightest

microfretting or slip will cause mechanical noise and lead to a lack of repeatability and decrease in resolution. Thus care must be taken ensure the flexures are clamped properly. It is recommended that the product of the clamping pressure (P_{cl}) and the coefficient of friction (μ) be two or three times the maximum stress in the outermost fibers of the flexure. On the other hand the clamping pressure should be less than one half the yield strength of the weakest flexure material [Slocum, 1993]. If the stresses are going to approach this maximum limit it would be preferable for the clamping material to be as strong as the flexures. In order to maximize the thickness of the circular flexures (and the off-axis stiffness) the clamping material (for the tool holder only) is chosen to be 7075-T6 aluminum. The yield strength of 7075-T6 is 500 MPa compared to 270 MPa for 6061-T6.

For the maximum clamping force this leads to the conditions

$$\mu P_{cl} = \beta \sigma_{max}, \quad \beta = 2 \text{ to } 3 \tag{3.1}$$

and

$$P_{cl} = 0.5\sigma_{max} \tag{3.2}$$

or

$$\sigma_{max} = \mu\sigma_{ys}/2\beta \tag{3.3}$$

If a coefficient of friction is assumed to be 0.5 and the yield strength of 7075-T6 aluminum to be 500 MPa then the maximum stress range is 42 MPa to 63 MPa (6 ksi to 9.3 ksi), for β=2 and 3 respectively. This stress range also satisfies the general rule of flexure design that the stress of the flexure should be less than 10% to 15% of the yield stress [Slocum, 1993].

The circular flexures are modeled as shown in Figure 3.6. The maximum stress is the stress on the inner edge of the circular flexures and depends on the material, the thickness (t), the expansion along the axis (δ) and the inner and outer radii. In the development presented here this last factor is expressed in terms of the ratio of

the inner to outer radii (called the aspect ratio, ξ) and the outer radius. The boundary conditions are

$$\frac{dy}{dr}_{r=r_i} = \frac{dy}{dr}_{r=r_o} = 0 \tag{3.4}$$

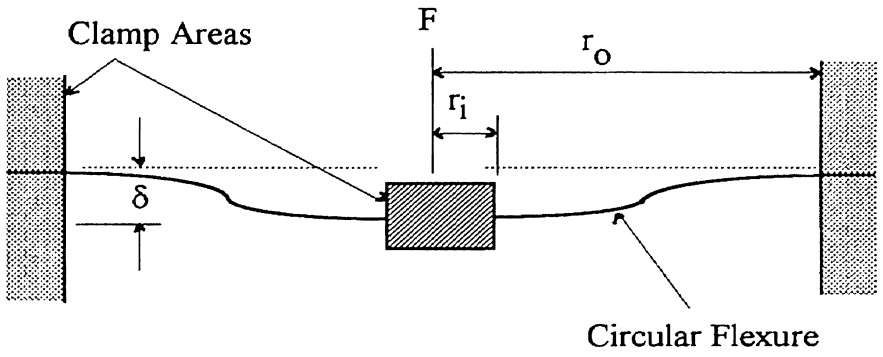


Figure 3.6. Deflection Model for Circular Flexures

From a handbook [Griffel, 1984] and after some rearranging the outer stress (σ_i) is given by

$$\sigma_i = \frac{3}{\pi} k_m f_1(\xi) f_4(\xi) \frac{t\delta}{r_0^2} \tag{3.5}$$

where $k_m = \frac{4\pi E}{3(1-\nu^2)} = 9.6\times10^{11}$ N/m^2 for steel

$$f_1(\xi) = [1 - \xi^2 - 4(\frac{\xi^2}{1-\xi^2})\ln^2(\xi)]^{-1} \tag{3.6}$$

$$f_4(\xi) = -[\frac{1}{2} - \frac{1}{1-\xi^2}\ln\xi]. \tag{3.7}$$

The outer stress (σ_o) is given by

$$\sigma_0 = \frac{f_2(\xi)}{f_4(\xi)}\sigma_i \tag{3.8}$$

where

$$f_2(\xi) = -[\frac{1}{2} + \frac{\xi^2}{1-\xi^2} \ln \xi]. \qquad (3.9)$$

To assure linearity it is assumed in all flexure calculations that the deflection is small compared to the thickness (deflection less than one-half the thickness).

The stress is directly proportional to the thickness and to the deflection. For all calculations the deflection was taken to be 30 μm. The resolution of the coarse adjustment in the y-direction is expected to be considerably better than this, so that this deflection (and a resulting higher stress) need not be obtained. In addition, the stress is inversely proportional to the square of the flexure thickness. In Figure 3.7 the effect of the aspect ratio on the stress is shown by plotting the product of the functions $f_1(\xi)$ and $f_4(\xi)$ against aspect ratio.

This is an important consideration. The size constraint and the stiffness requirement drive the design in opposite directions; in order that both constraints be satisfied as far as possible it will be necessary that the aspect ratio will be relatively large, greater than 0.5. However for large aspect ratios (those over about 0.6) small changes in aspect ratio give large changes in the stress. The aspect ratio is expected to be an important parameter in determining the final design. Whenever possible all design calculations are cast in terms of the aspect ratio. This also facilitates comparing one calculation with another. The final radii will have to be chosen carefully so that the stresses do not become too large.

Axial deflection will be largely determined by the stiffness of the PZ stack and the components along the y-axis. But the off-axis stiffness and deflection (which are considered errors) will be determined by the flexures themselves. The inner cylinder and flexures are modeled as an infinitely stiff beam supported by springs as shown in Figure 3.8. The cutting force (F_c) acts at the end of the beam (which is the very tip of the tool) in the x-direction.

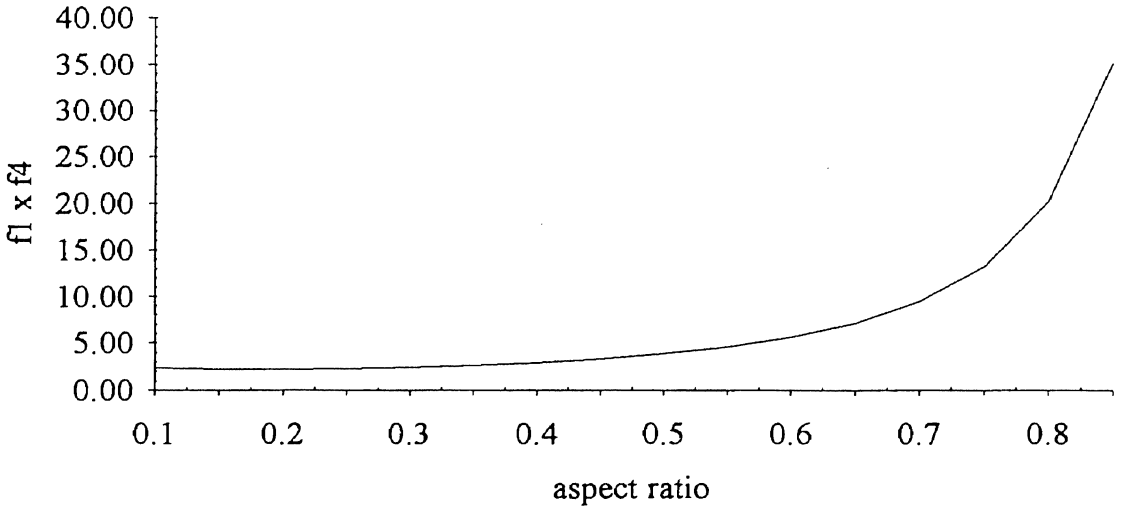


Figure 3.7. Influence of Aspect Ratio on Stress

The elastic forces applied by the flexures are represented by F_1 and F_2. As the cutting force is applied to the tool post the front and back flexures will be deformed by an amount δ_1 and δ_2, respectively. From the geometry of Figure 3.8 the tool deflection (Δ_t) is given by

$$\Delta_t = [r_l(1+\frac{\delta_2}{\delta_1})+1]\delta_1 + d_c \qquad (3.10)$$

Where r_l = length ratio = l_t/l_o and d_c is to account for tool post cantilevering. Note that the smaller r_l is the smaller the off-axis deflection. By assuming the tool post is a rectangular shape with square sides of 1 cm^2 and 2 cm long (typical dimensions for diamond tools), and applying the standard formula for an end loaded cantilever the term d_c is estimated to be about 5 nm.

Furthermore, the compression/extension of the individual flexures will be different since different forces act on each. These are given by

$$F_1 = (1+r_l)F_c \qquad (3.11a)$$

$$F_2 = r_l F_c \qquad (3.11b)$$

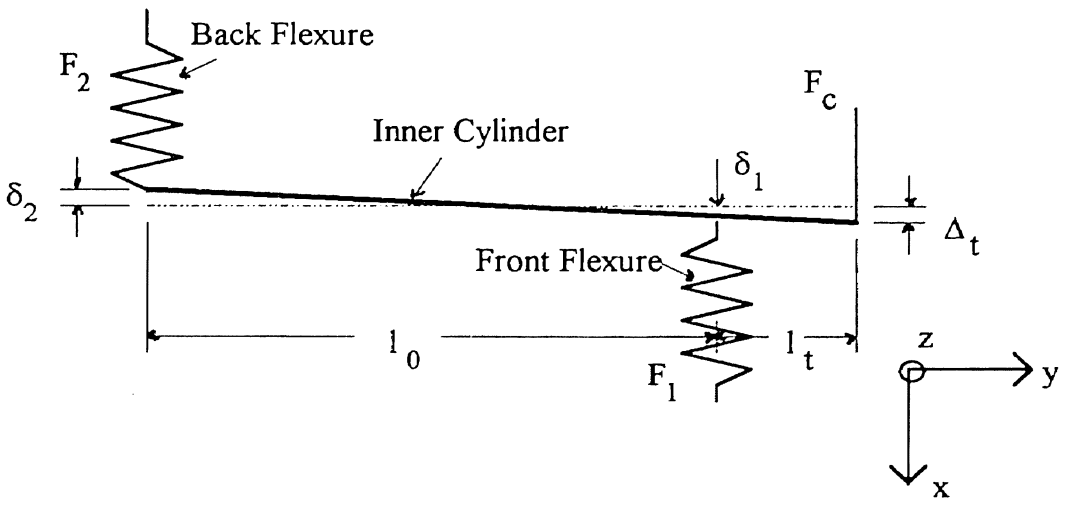


Figure 3.8. Deflection Geometry for Tool Holder

In order to calculate Δ_t an estimate of the off-axis stiffness of the flexure must be obtained. The deflection of each flexure (δ_i, i=1,2) is defined by $\delta_i = F_i/k_f$ but the off-axis (transverse) stiffness of the flexure (k_f) is not known. Two different 'models' were used, both of which are extremely simple and overestimate the deflection of the flexures as the upper half of the flexure provide some constraint. Instead of trying to calculate the deflection of a circular shape set on edge, a semicircle is used. The structure is then assumed to be a column of thickness t (the flexure thickness) and width *L*. The difference between the two model is in the way the length *L* (and therefore the area over which the force acted) was chosen. The stiffness is then written $k_f = EA/a$. It is understood that this can provide only the crudest estimate of deflection, since part of the structure (namely the other half of the flexure) is ignored. The unstressed length (a) in each case was chosen to be r_o-r_i. In this case the tool deflection (3.10) can be rewritten, using equations (3.11a) and (3.11b) to obtain

$$\Delta_t = f(r_l)\frac{r_0 - r_i}{AE}F_c + d_c \qquad (3.12)$$

where

$$f(r_1) = 2r_1^2 + 2r_1 + 1. \tag{3.13}$$

Note that the smaller r_1 the smaller the off-axis deflection.

In the first model the area is calculated based on the area of a column of width r_i as shown in Figure 3.9.

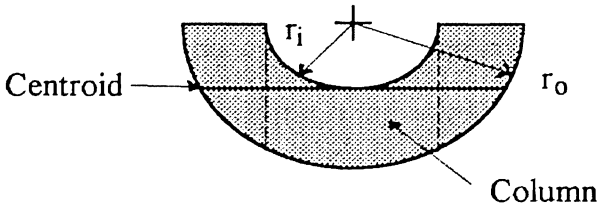


Figure 3.9. Off-Axis Stiffness Model.

The area is then $2r_i t$ and the deflection is given by

$$\Delta_t = \frac{F_c}{E} f(r_1) \frac{1-\xi}{2\xi} \cdot \frac{1}{t} + d_c \tag{3.13}$$

For the second model the area was based on the centroid of the region shown in Figure 3.9. In this case the centroid $\bar{r}$ is

$$\bar{r} = \frac{4r_0}{3\pi}\left(\frac{1-\xi^3}{1-\xi^2}\right). \tag{3.14}$$

The tool deflection is then

$$\Delta_t = \frac{F_c}{E} f(r_1) f_3(\xi) t + d_c \tag{3.15}$$

where

$$f_3(\xi) = \frac{1}{2}(1-\xi)\left[1 - \frac{16}{9\pi^2}\left(\frac{1-\xi^3}{1-\xi^2}\right)^2\right]^{-\frac{1}{2}}. \tag{3.16}$$

Figure 3.10 shows the relative effect of the aspect ratio on the deflection for aspect ratios greater than 0.5. This influence is demonstrated by graphing the aspect ratio dependent part of equations (3.14) and (3.16) versus the aspect ratio.

The two different estimates give widely divergent deflections for small aspect ratios, but become similar for aspect ratios greater than 0.8. A number of different schemes for choosing the area all exhibited this behavior.

Another consideration for calculating the dimensions of the flexures is their axial stiffness. Although the static axial stiffness of the tool holding assembly will be determined mainly by the PZ stack and frame the stiffness of the flexures themselves should not be neglected. First, the higher the axial stiffness the higher the off-axis stiffness (and vice versa) and the smaller the off axis deflections will be. Second, the flexure stiffness will have a direct bearing on the resonant frequency of the assembly. The inner cylinder is not rigidly attached to the PZ stack; if the flexures have a too low a stiffness it will greatly reduce the range of dynamic performance. Ideally, the flexures should have the same dynamic range as the PZ stack. However due to the mass of the inner cylinder and naturally lower stiffness and smaller section size of the flexure materials (steel as compared to ceramics) it is expected that the flexures will not be nearly as stiff as the stack. The stiffness of the PI stack selected for the MMI is 420 N/μm. A reasonable stiffness for the flexures will be 1 or 2 N/μm.

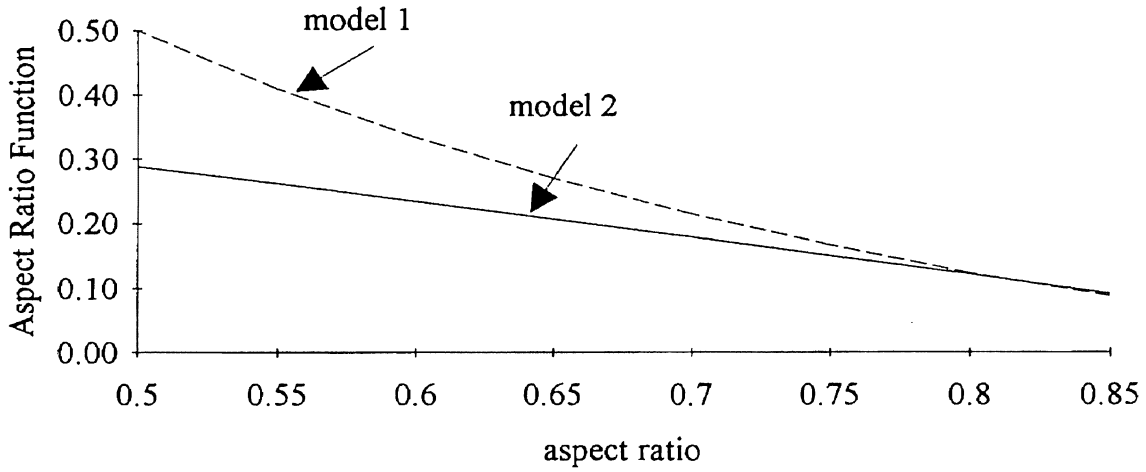


Figure 3.10. Influence of Aspect ratio on Off-Axis Deflection.

The stiffness (k) of the flexures is given by

$$k = k_m f_1(\xi)\frac{t^3}{r_0^2} \tag{3.18}$$

where $f_1(\xi)$ and k_m are the same as before. The thickness has a strong effect on stiffness, as does the aspect ratio. The effect of flexure dimensions on dynamic properties is discussed in Chapter V.

The suggested bolt spacing for the clamping arrangement shown in Figure 3.11 is derived from the optimal stress cone overlap. The bolt spacing (l_{bs}) should provide approximately a 30% overlap in the 45° pressure cones formed by the bolt head [Slocum, 1993]. In addition to avoid any unclamped areas the width (w) of the clamping zone should be about 60% of the total diameter of the 45° pressure cones [Slocum, 1993].

These conditions lead to the formulas

$$w = D_w + \frac{4}{3}t \quad \text{and} \quad l_{bs} = D_w + \frac{5}{3}t \tag{3.19}$$

where D_w is the diameter of the bolt head. It should also be noted that the final width of the flexure clamp and base are slightly larger than those calculated based on bolt spacing. The clamp and base are given a small radius at the inner edge of the flexure to avoid stress concentrations [Slocum, 1993].

To choose dimensions for the tool holder assembly an iterative procedure is used. First a spreadsheet is made to calculate the inner stress for a variety of aspect ratios, thicknesses and outer radii (all using an expansion of 30 μm). From this spreadsheet all configurations which fall into the allowable stress range are extracted. These configurations are then used to set up a second spreadsheet to calculate off-axis deflections (the results of the two models is averaged). From this spreadsheet those configurations with the least deflection are selected and the bolt spacing necessary is calculated. Many configurations which fell into the

allowable stress range had either large off-axis deflections, were not large enough to support the expected clamping pressure or would not fit under the microscope. Some parameters had to be fixed, such as the length ratio r_l. From a measurement of the microscope it was judged that this parameter could not be much smaller than 0.2 and still allow the MMI to fit on the microscope. The final dimensions for the flexures with some of the specifications is given in Table 3.1

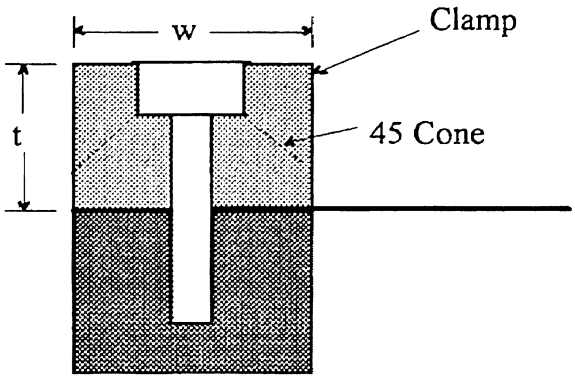


Figure 3.11. Bolt Spacing Parameters

Now, Δ_t is the total tool deflection in the x-direction. As the tool cuts and is deflected in the x-direction it may back out slightly (similar to the motion of a pendulum) giving an error (e) in the (uncorrected) depth of cut. This situation is illustrated in Figure 3.12. If this is a non-sensitive direction Δ_t will not induce a large error along the y-axis.

The length l_t is chosen because it gives an upper bound estimate and it makes the calculation easier. Neglecting terms of order $(e)^2$, the deflection is

$$e = \frac{\Delta_t^2}{2l_t} \tag{3.20}$$

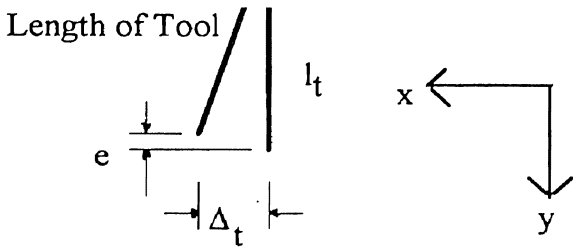


Figure 3.12. Error Induced by Off-axis Deflection

For $l_t = 2$ cm and Δ_t in the range of 50nm to 150 nm this distance is in the range 0.0003 nm to 0.002nm. It appears that the x-axis (and by symmetry the z-axis) is not a sensitive axis.

A summary of the tool holder dimensions is given in Table 3.1

Dimensions	
Inner radius (cm)	2.1
Outer radius (cm)	3.4
Aspect ratio	0.6
Tool length (cm)	2.5
Flexure separation (cm)	15
Length ratio	0.17
Flexure thickness (mm)	0.30
Clamp thickness (cm)	1.1
Clamp width (cm)	1.6
Specifications	
Axial stiffness (N/μm)	1
Off-axis stiffness (N/μm)	100
Resonant frequency* (Hz)	150
Total expansion (μm)	30
Maximum force (N)	20,000
Maximum stress** (MPa)	65

* Undamped

** For a 30 μm maximum expansion and 0.50 mm thick flexure

Table 3.1. Dimensions for Tool Holder

3.4 Rectangular Flexure Design

The sample carriage translates the sample across the tool. It is based on a clamped flexure in the shape of a flat, wide cantilever. The beam is modeled as shown in Figure 3.13.

The deflection and stresses are determined by solving the standard beam equations

$$\frac{d^4y}{dx^4} = q(x); \qquad \frac{d^2y}{dx^2} = \frac{M}{EI} \tag{3.21}$$

where q(x) = 0; subject to the boundary conditions

$$y\,(x = 0) = 0; \tag{3.22}$$

and

$$\frac{dy}{dx}(x = 0) = \frac{dy}{dx}(x = L) = 0. \tag{3.23}$$

The deflection y_{max} is equal to the length of cut which is 250 μm. The deflection of the beam is given by

$$y_{max} = \frac{1}{6}\frac{FL^3}{EI}(1 - \nu^2) \tag{3.23}$$

Where the $1\text{-}\upsilon^2$ term accounts for the fact that the width of the cantilever is much greater than its thickness [Griffel, 1984].

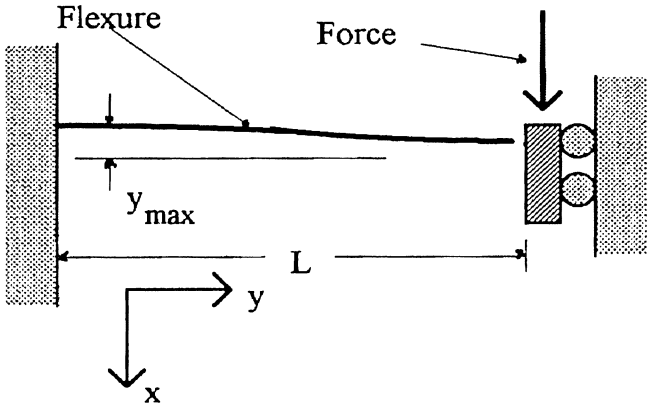


Figure 3.13. Model of Rectangular Flexure

From here the stress in the outer fibers of the flexure can be found to be

$$\sigma_{max} = \frac{3Ety_{max}}{L^2(1-\nu^2)}. \tag{3.25}$$

The stiffness in the direction of deflection is

$$k = \frac{Ew}{2(1-\nu^2)}\left(\frac{t}{L}\right)^3. \tag{3.26}$$

As with the tool holder flexures there are a number of parameters and dimensions that are fixed by the materials or the physical constraint of the microscope. Using the criterion for stress developed in section 3.3 the allowable stresses are from 23 to 35 MPa with the frame and clamping material made from 6061-T6. The lower strength alloy is used for this assembly because the orientation of the flexures will allow them to be longer and stresses to be lower. To avoid obstructing the microscope the flexures cannot be more than 5 cm wide (width here referring to the dimension parallel to the z-axis). When high precision is to be attained the relative thickness of the flexure must be more than one-half the total travel of the flexure to insure linearity. In the case of the tool holder flexures this requirement was not a concern because the total travel was small (30 μm) and the thickness of the flexures (as dictated by other concerns) guaranteed that the thickness would be sufficient on this count. However the sample holder travel, at 250 μm, is to be 10 times that of the tool holder; thus the flexures should not be less than 0.5 mm thick. In addition the y-axis stiffness (a sensitive direction) will need to be as high as possible; the larger the section size of the rectangular flexures the higher the stiffness in this direction. Indeed this consideration lead to the rejection of a sample holder geometry based on circular flexures. Circular flexures (if small enough to mount under the microscope) would have to be too thin to maintain reasonable stresses thus giving away stiffness and

requiring a range of motion well beyond one-half the thickness. A range of thicknesses is then established from 0.5 mm (the minimum necessary to assure linearity) to several millimeters for design calculations.

The thrust force acts parallel to the length of the flexures and will tend to extend and compress them. To calculate the deflection along the y-axis the flexure is modeled as a rectangular column with stiffness equal to EA/L where L is the length (parallel to the y-axis) of the flexures and A is the product of the width times the thickness.

To determine the final dimensions a series of spreadsheets is created for a variety of lengths, widths and thicknesses. Only those configurations that have stresses within the accepted range and with the lowest off-axis deflections are retained and the bolt spacing is calculated for each configuration. For a thickness of 1.5 mm and length of 6 cm, the predicted deflection is less than 10 nm.

The dimensions and specifications for the tool holder are given in Table 3.2

Dimensions	
Length (cm)	6
Width (cm)	5
Thickness (mm)	1.0
Cross carriage width (cm)	11
Clamp width (cm)	1.7
Specifications	
Travel (μm)	250
Axial stiffness (N/μm)	0.1
Off-axis stiffness (N/μm)	300
Resonant frequency* (Hz)	80

*Undamped

Table 3.2. Dimensions for the Sample Holder Assembly

Of concern, however is the extent to which these flexures will act as slender column with respect to the y-axis forces. This tendency is somewhat mitigated by the flexures on the tool post side of the cross carriage which act as supports (because they are in tension). The critical load required to cause buckling in a free standing beam is directly proportional to the ratio E/L^2. If it is assumed the compressive flexures are free standing columns, for the design loads the ratio of E/L^2 (where L is the length in the y-direction) to cause buckling are on the order of 10^{-12} m^2. For the expected order of magnitude dimensions of the cross carriage flexures the ratio of E/L^2 is about 10^{-10} m^2 , so the forces expected are to be about two orders of magnitude below those necessary to cause buckling.

The sample holder assembly is more sensitive to assembly errors than for the circular flexures. If the lengths are not the same on both sides there will be a runout error as the cross carriage moves; this situation is made somewhat worse by using multiple sheets (to permit layering with an RTV silicon which will provide damping) on each side.

Melles Griot NanoFlex™ Series of stages use single and differential rectangular flexures to obtain nanometer resolution. NanoFlex stages would not allow mounting of the dynamomter, however.

3.5 Capacitance Gauge Probe Placement

Ideally the capacitance gauge probe should be rigidly mounted parallel and colinear with the sensitive axis to reduce Abbe errors. This would entail placing it along the y-axis of the MMI inside the inner cylinder.

The probe was place outside the inner cylinder for several reasons. The requirement of using the microscope requires that sometimes both the tool holder and the sample holder will have to be adjusted (to position both the sample and

tool under the objective). There is then no way to insure then that the proper gap would always be maintained for the probe; it would need to be adjustable. The manufacture of an adjusting mechanism inside the cylinder would be complicated and expensive. The current mount has the probe riding in two O-rings and held with a Teflon tipped set screw. Also placing the probe behind the tool would leave several centimeters of uncompensated length; this problem is generally handled by making all fixtures which are outside of the feedback loop out of low expansion alloys. By placing the probe directly next to the tool and sample the uncompensated length of the tool assembly is kept to a minimum. The in-cylinder placement would also require removing a significant amount of material hence reducing stiffness. The concern is that since the microscope severely restricts the size, internal sections need to be left as large as possible.

A major drawback of the conventional mounting scheme is that sample holder errors cannot be compensated for since the reference surfaces are the tool holder and the frame. By mounting the probe on the sample holder y-axis errors for both tool and sample holder can be compensated.

Runout (motion of the tool holder deviating from the y-axis) or a misalignment of the probe and target surface will result in a cosine offset error. For surfaces separated by a distance y_0 the error (ε) in absolute position is given by [Slocum, 1993]

$$\varepsilon = \frac{y_0 \theta^2}{2} \tag{3.27}$$

where θ is the angle the sensor makes with the axis of motion. The probe sensing width is 0.125". Thus every 0.001" of misalignment (y) results in an absolute position error of 10 nm. This is the maximum misalignment that is still within the noise of the capacitance gauge module board. See Figure 3.14 for the definition of the variables.

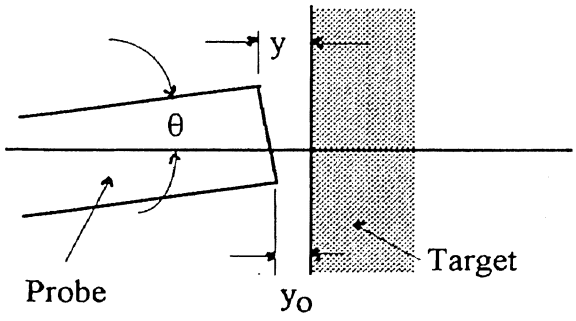


Figure 3.14. Capacitance Gauge Misalignment

3.6. Assembly

Figures 3.15 shows the wrap around frame and the overall position of the tool holder and cross carriage frame. Not shown is the slide that allows the tool holder to move relative to the sample. Figures 3.16a and 3.16b show the entire assembly as it rests on the Nikon Measurescope.

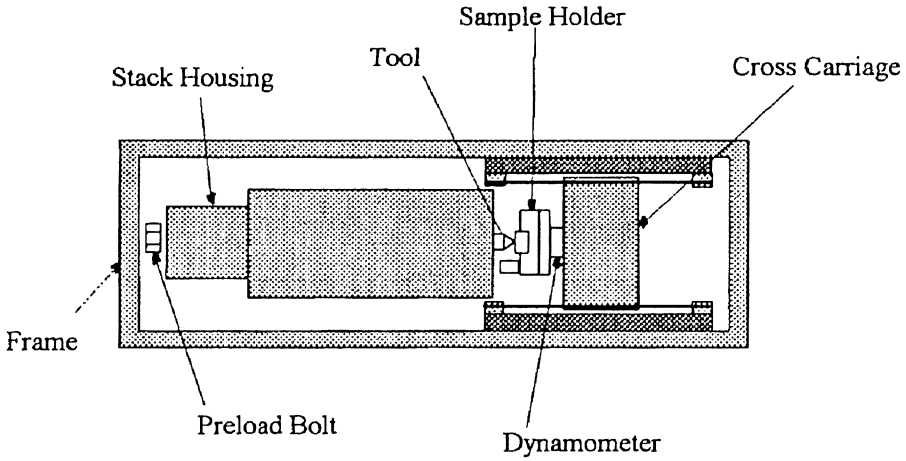


Figure 3.15. MMI Assembly

Mechanical contact between the stack and the inner cylinder and the stack and the preload bolt is provided by two 5/32" steel ball bearings which fit into conical seats at each end of the stack (See Figure 3.1, page 32). There are 18 10-24 bolts

clamping the flexures on the outer clamps and four 10-24 bolts in the inner clamps.

The clearance between the inner and outer cylinders is about 0.01". The inner cylinder has locating pins machined at its center which align the flexures to the inner cylinder. To center the flexures and the inner cylinder each end of the inner cylinder (inside the flexures) has a layer of cellophane tape; this may also provide some frictional damping. The fit of the flexures with the outer clamp is loose enough that without the tape the inner and outer cylinder can make contact which effectively locks the inner cylinder. With the tape the inner cylinder is free to move but the alignment (checked at the front and back by inserting a wire gauge at four points) cannot be better than 0.005".

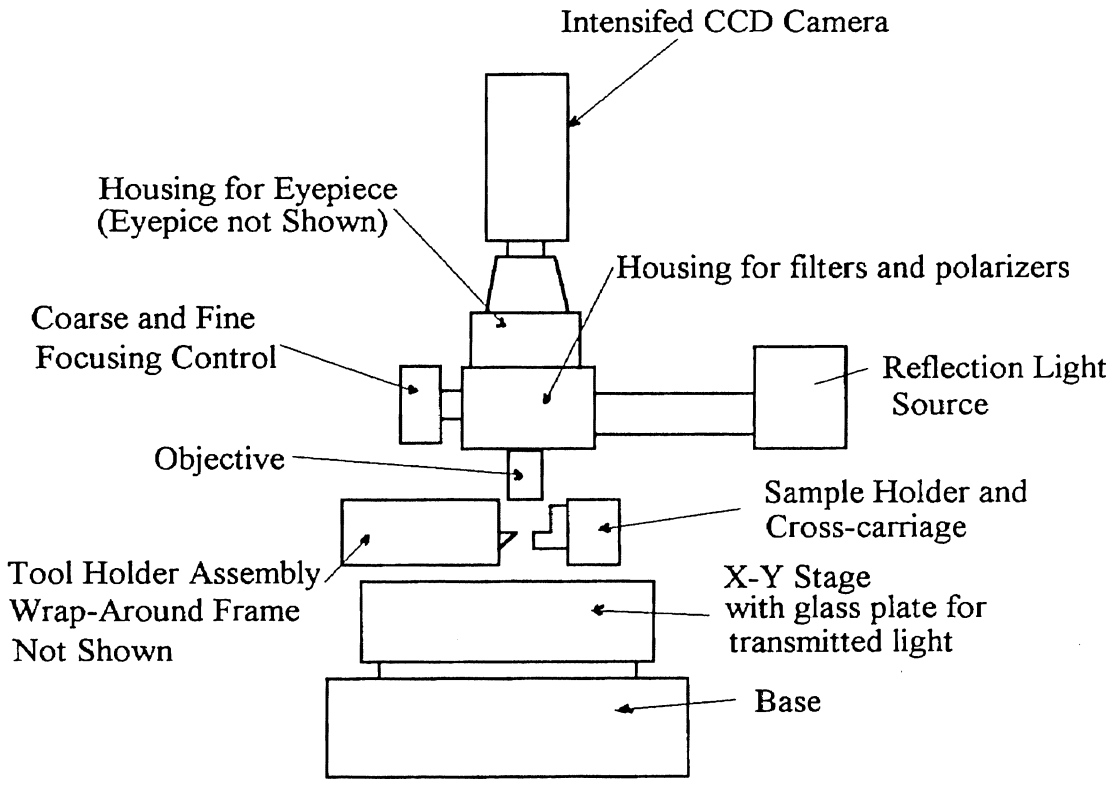


Figure 3.16a. Front View of Microscope and MMI

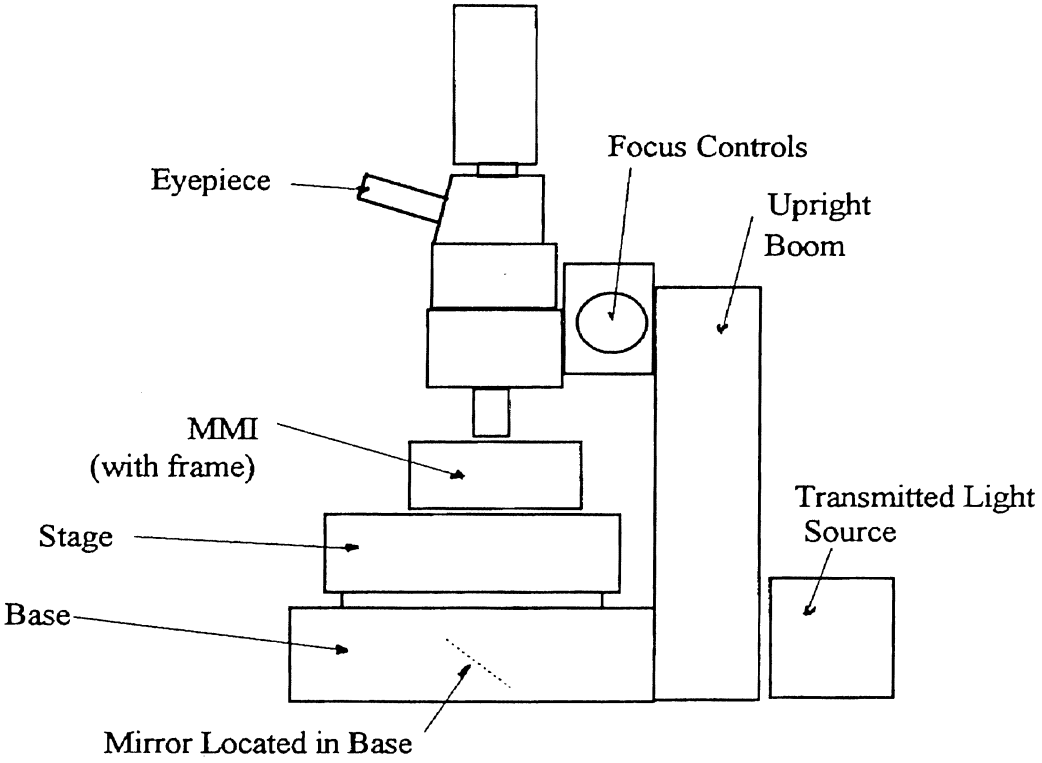


Figure 3.16b. Side View of Microscope and MMI

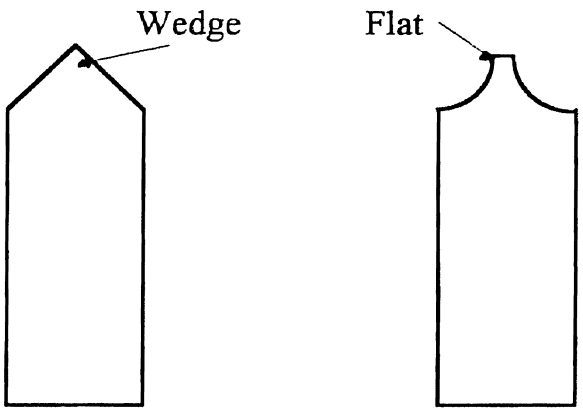


Figure 3.17. Indenters

Shown above in Figure 3.17 are the two indenters manufactured for this phase of the MMI. One has a square profile approximately 500 μm across and the wedge has a nose radius of approximately 20 μm.

3.7 Control System Design

A closed loop feedback control system provides several advantages. It will increase stiffness by compensating for small errors in motion, correcting for compliance, correcting for thermal drift and allow virtually all operation of the MMI to be integrated into a single overall computer control.

A diagram of the physical model used is shown in Figure 3.18. The inner cylinder is represented by the mass M and the flexures by the springs S. As shown previously the capacitance gauge probe is mounted on the sample holder across from the target mounted on the tool holder. This arrangement affords a small air space of several thousands of an inch between the gauge and the target, thus forming a parallel plate air capacitor. Changes in the position of the target causes a change in the capacitance which is processed by the board. The output of the board is a calibrated voltage (the calibration being provided by ADE) between -5V and +5V with each millivolt change representing a displacement of 1 μin (25 nm), and a reading of zero volts when the stand-off distance is 0.007500 in. This output is then read by the data acquisition board and an error based on a reference voltage is calculated. The computer supplies a voltage to the amplifier to correct the error. The position of the probe is adjustable to accommodate tools and samples of different length. The amplifier used is capable of high speed (high slew rate) and sensitive to the level of inductive or capacitive loading. The amplifier is itself a closed loop feedback instrument and if the load is not sufficiently resistive the amplifier will become unstable. The piezoelectric transducer is an almost purely capacitive load and so the resistor is used to prevent the amplifier from becoming unstable.

The A/D board has a resolution of 12 bits and software adjustable gains and voltage ranges. To determine the resolution (r) for a given voltage range (ΔV) and gain (G) use the formula

$$r = \frac{\Delta V}{G \times 2^{12}} \tag{3.28}$$

The gains available are 0.5, 1, 2, 5, 10, 20, 50 and 100. The voltage ranges are ±5 V, 0V to 5 V and ±10 V.

The resolution of the module board supplied with the capacitance gauge probe is 0.5 mV. The board output can be as high as 10 V but in practice it is a simple matter to position the probe so the output for zero expansion is on the order of 10 mV. A full expansion of the stack would then result in an 800 mV output from the module board. Thus with a gain of 5 on the 0 V to 5 V range the resolution is 0.25 mV which is 0.25 mV better than the capacitance gauge.

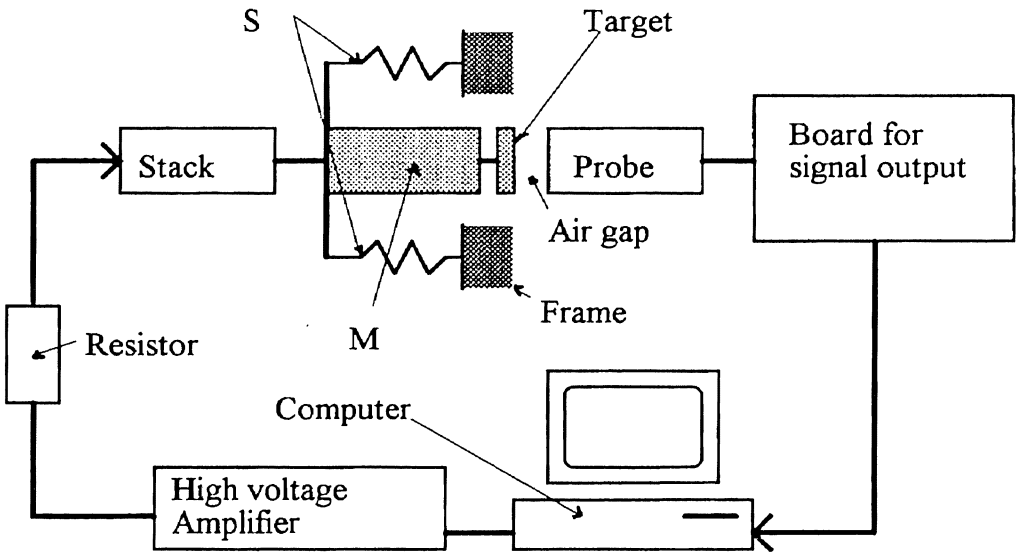


Figure 3.18. Schematic of Control System.

For the control system a proportional-integral (PI) controller is assumed and the two unknown constants, K_p and K_I are to be determined from frequency

response data (as with the Ziegler-Nichols method with $K_D = 0$, for example) without prior knowledge of the plant or modeling of system dynamics [Ellis, 1991]. The arrangement of the controller is shown in Figure 3.19.

Once the controller is designed and equations are written for the time domain the frequency domain equations must be determined and furthermore since a digital computer is to be used the z-domain equations must be determined. The z-transform is used to map the analog algorithm into a digital algorithm for use on a computer.

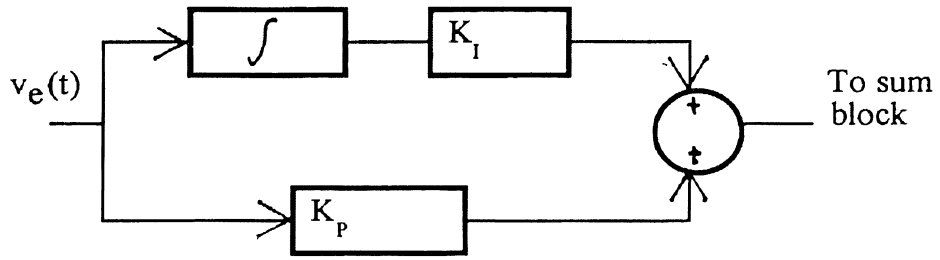


Figure 3.19. Block Diagram of a PI Controller

The digital computer cannot continuously sample the input. The delay caused by the sample and hold period can be represented as a transform of the form $z = e^{-sT}$ where T is the sample rate. By using a Taylor series expansion the following approximation of this transform can be derived:

$$s = \frac{2}{T}\frac{z-1}{z+1}. \tag{3.29}$$

In terms of the frequency domain the PI controller, for the input $V_e(s)$and the computer output V(s), can be written as

$$V(s) = K_I \frac{V_e(s)}{s} + K_p V_e(s) \tag{3.30}$$

For the PI controller the resulting algorithm after the transform is applied is

$$v_n = v_{n-1} + \alpha[K_p(e_n - e_{n-1}) + \frac{K_I T}{2}(e_n + e_{n-1})] \tag{3.31}$$

where T is the sample time, e is the measured error and v is the voltage output from the computer. The proportional constant isK_p and K_I is the integral constant. The subscript n denotes the iteration number for each sample; hence the nth output depends on the nth error and the previous output and error. The constant α is not a result of the transform; it is used to mach the scale of the error terms to the output. The sampling frequency should be at least twice as fast as the reciprocal of the characteristic time constants of the plant. These quantities can be determined from dynamic response data to be discussed in the next chapter.

The algorithms are programmed using the LabVIEW for Windows graphical programming language with a 486DX 50MHz computer as the platform. By using this language the entire operation can be automated from the computer and operated by an individual with no knowledge of computer programming. The graphical interface generated by the program creates a virtual instrument that controls the MMI. An example of the control algorithm of equation (3.31) programmed as a LabVIEW VI is shown in Figure 3.20. See Appendix C for more information about LabVIEW.

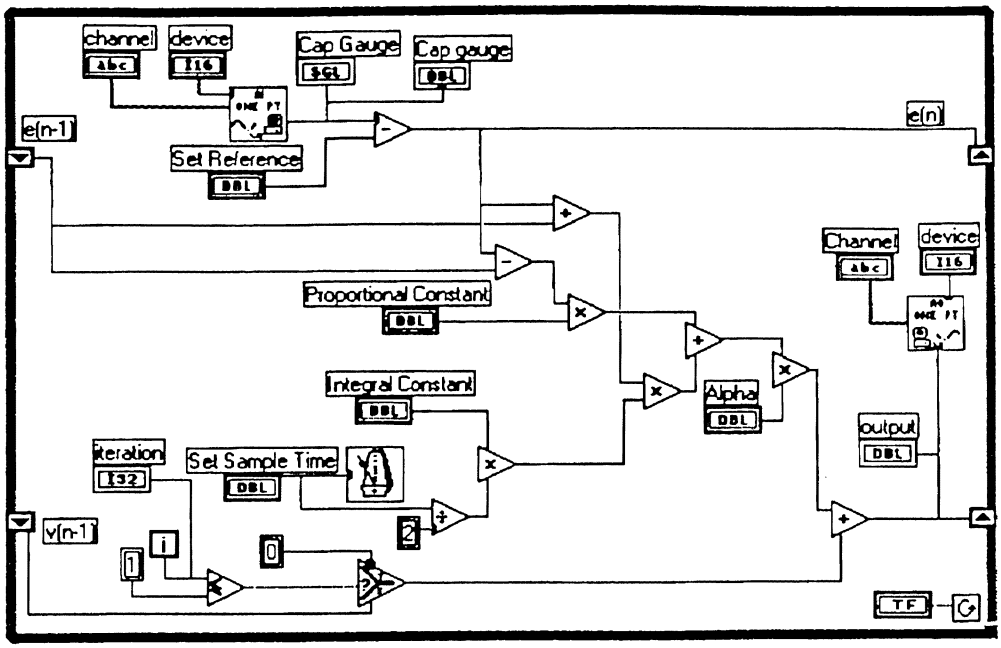


Figure 3.20. Block Diagram of Control Algorithm

CHAPTER IV

OVERALL INTEGRATION

4.1 Introduction

The experimental device consists of two independent subsystems: The microscope with its optics and the mechanical structure (the MMI). The MMI has two interdependent subsystems in that the control system and the mechanical structure interact. Ideally performance should be calculated and then the performance measured and compared to the calculation; whenever such a comparison cannot be made due to lack of data a method for making the measurement is suggested. Evaluation of the optical performance of the microscope is somewhat qualitative. Also, most of the distance measurements are taken with the capacitance gauge; when its performance influences a measurement an appropriate note will be made.

4.2 Stiffness

For calculating the stiffness of the sensitive axis (the y-axis) the MMI can be visualized as a collection of springs connected in series. The equivalent stiffness of springs in series is given by $k_{eq} = \sum 1/k_i$ where k_i is the stiffness of each element. The individual elements corresponding to each spring are shown in Table 4.1 with the estimated stiffness of each component. The stiffness is calculated as $k_i = EA_i/L_i$ where A_i is the gross section area and L_i is the unstressed length. The

total estimated stiffness is 65 N/μm. This estimate is within the range established as a goal on page 22. However this is just an estimate; the stiffness of the ball bearings was estimated as one half the stiffness of a cylinder 5/32" in diameter and 5/32" high.

MMI component	Estimated y-axis Stiffness (N/μm)
Inner Cylinder	1.2×10^3
Tool and Tool Holder	350
PZ Stack	420
Ball Bearings and Preload Bolt	180
Frame	5×10^3
Sample Holder	3×10^3
Cross-Carriage	7×10^3
Rectangular Flexures	300

Table 4.1. MMI Elements and Their Stiffness.

Verification of the stiffness can be made by mounting a square 5/16" piece of tool stock in the tool holder and expanding the cylinder/flexure assembly with either the PZ stack or by mechanical means. The dynamometer will give force data and the capacitance gauge will read deflection.

4.3 Static Performance of Stack/Inner Cylinder

The amplifier, PZ stack and cylinder assembly together constitute an electromechanical device and electrical and mechanical behavior should be kept separate as far as possible. For example the expansion of the stack is proportional to the charge on the disk surfaces and at low frequencies the stack can be thought of as a purely capacitive load. Mechanically however there is some hysteresis that this simple electrical model cannot account for, and at higher frequencies the rate of mechanical expansion is not proportional to dq/dt of the driver.

The Kepco amplifier has insufficient phase margin to drive a purely capacitive load. When connected to the PZ stack a positive feedback develops and the stack is driven into a 60 Hz (frequency measured with a Fluke 88 multimeter) low amplitude oscillation. A decade resistor placed in series with the amp and the stack adjusted to a minimum of 200 Ω was usually sufficient to prevent this oscillation; however variations in the background noise or small bits of dust in the connections sometimes raised the minimum resistance to 300 Ω or 400 Ω. It should be noted that this resistance guarantees static stability only.

To verify the accuracy of the capacitance gauge measurements a series of voltages were applied and the voltage output of the cap gauge read and converted into a displacement. The expansion of the target was simultaneously measured by a Rank-Pneumo Model 1045 gauge head which uses a very sensitive LVDT (linear variable differential transformer) to obtain resolution of approximately 10 nm. To the author's knowledge the R-P gauge head had not been recently calibrated. It was assumed that if the two different instruments (the gauge head and the cap gauge) gave identical displacements, read from the same point, for the same input then it was likely that both read correctly. By comparison the resolution of the capacitance gauge (with the board provided) is 12.5 nm. A small amount of drift is present (a standard behavior for PZ stacks) and shown graphically for a 2 μm expansion (from 0 μm to 2 μm in 0.5 s) in Figure 4.1. These measurements were taken as voltage readings by the capacitance gauge read by the computer using a LabVIEW for Windows visual interface (VI) and from the R-P gauge head. The voltage range was 0 V to 5 V and the gain was 5. The noise was measured with the Nicolet Pro 10 oscilloscope, with the time/division set high enough that there was no aliasing. The major component of the noise was a 3 mV 60 Hz signal. This signal made the drift difficult to see; the curve was generated by recording the voltage on the LabVIEW VI. The noise was very consistent so a curve that

was the average of the maximum and the minimum of each point is the graph of Figure 4.1. The output of the capacitance gauge was consistent with the R-P gauge head; the linearity of the gauge head is 0.1%.

Another common performance aspect of PZ stacks is observed as well: hysteresis. The hysteresis curve of the stack/cylinder assembly is shown in Figure 4.2. This curve is generated by raising the voltage gradually (over a period of several seconds) in 100 V steps and then reducing the voltage zero at the same rate. The output of the capacitance gauge is read from a Fluke 88 multimeter and converted into a distance. These curves demonstrate a non-linear character and the assumption that the expansion of the stack is directly proportional to the voltage applied by the amplifier is false.

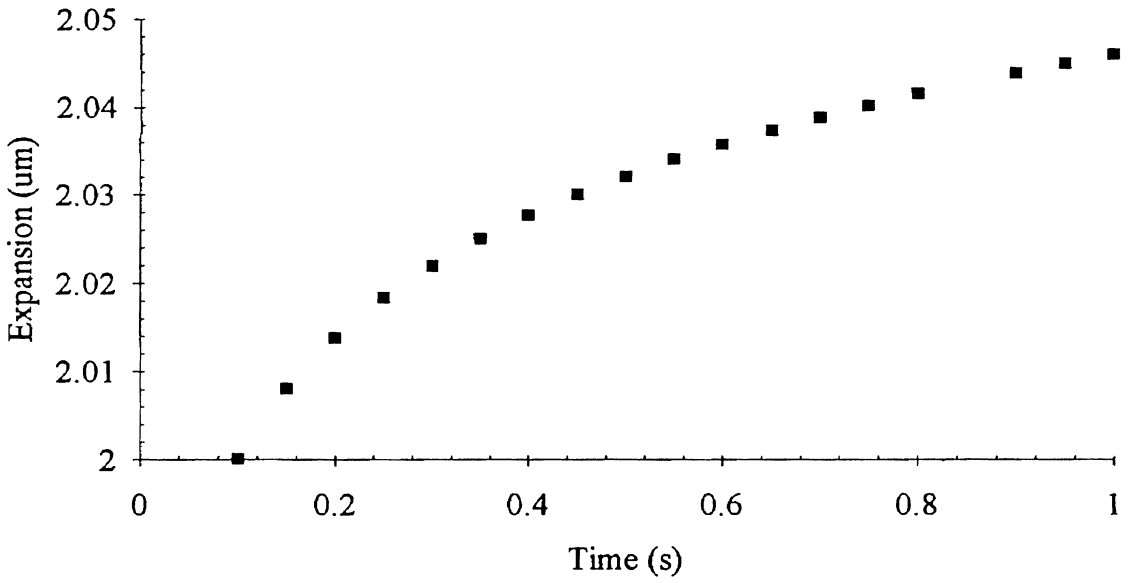


Figure 4.1. Drift of Stack and Cylinder

The effect of drift on open loop system response depends on the slope of the ramp as the voltage changes. For step responses it is a maximum and for slow increases taking several seconds it is negligible (immeasurable). In fact careful

examination of the step response of Figure 4.3 also displays a small amount of drift. The manufacturer claims that under closed loop operation (up to 1 kHz) drift and hysteresis can both reduced to immeasurable levels. The hysteresis is about 14% of the total expansion which is consistent with the manufacturers specifications. The manufacturer also claims that the relative hysteresis (represented here as a percentage) does not depend on the magnitude of the expansion.

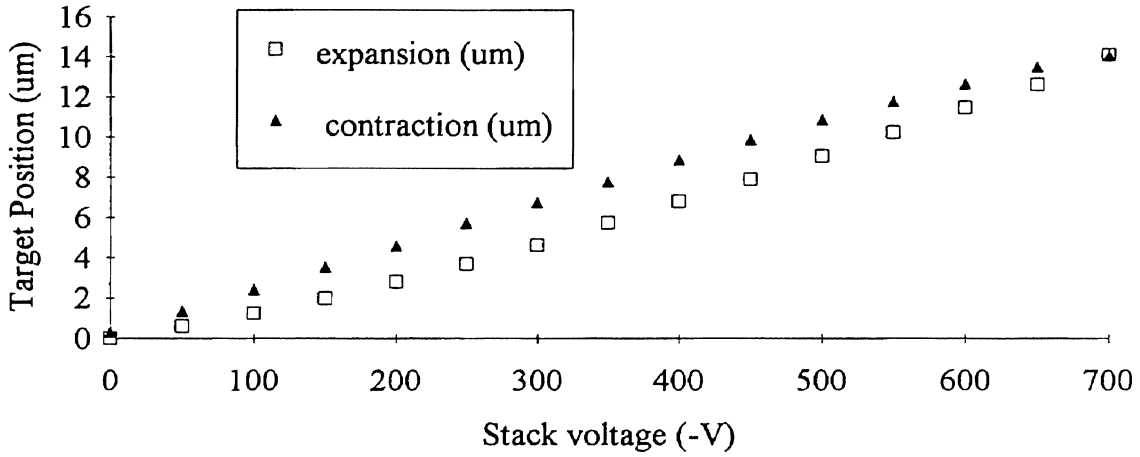


Figure 4.2. Hysteresis of Stack and Cylinder Assembly.

Note that in Figure 4.2 the stack does not return to the original zero position. There are four possible explanations for this observation: 1) the drift phase of the motion is not complete, 2) the power dissipated in the stack has caused some thermal expansion, 3) fretting of the flexures has not allowed the target to return to the initial zero (even though the stack is at zero)and 4) there is a small residual polarization even though the applied field is zero. Possibility 4) seems the most likely (even though the others may influence the total amount of residual expansion) because in open loop operation there is always a residual polarization; that is why the expansion depends on past voltage history (hysteresis).

It should be noted that the decade resistor has negligible effect on these static tests as the stack has nominal capacitance of 0.52 μF which gives a time constant for the stack/resistor combination of about 0.15 ms. In addition at quasistatic operation the effective leakage resistance is on the order of 10 MΩ. However the capacitance of the stack is not constant and is a function of the applied voltage. The specific variation of capacitance with voltage is not provided by the manufacturer. As a general statement though PI claims that the capacitance at full voltage is approximately 30% higher than nominal.

Finally, the noise of the capacitance gauge signal varied considerably; from a minimum of about 1 mV to as high at times as 7 mV (The noise component was always verified with the oscilloscope set at high enougn speed to avoid aliasing.). The noise was usually a 60 Hz signal and could sometimes be diminished by changing the grounding scheme from differential to pseudodifferential or vice versa. The connections themselves seemed to contribute to the amount of noise as well. Another source of noise was probably the target of the capacitance gauge probe which was grounded with only an alligator clip; the best arrangement would have the probe target drilled for a terminal screw.

4.4 Dynamic Performance

In addition to obtaining open loop performance for the tuning and initial adjustment of the controller the general dynamic character of the plant is of interest; the closed loop and open loop frequency response performance can be compared to evaluate the degree of improvement. To this end two types of input signals are selected from the LabVIEW analysis libraries: step and sinusoidal. The step inputs range from 1 V to 8 V in one volt increments. There are 10^4 points in each step output to the amplifier. The pulses are two seconds long so the

computer updates the signal every 200 μs. This input should yield an output from the amplifier of 100 V to 800 V in 100 V steps. This was not verified by measurement.

Both the board input (from this point forward "board" will mean the data acquisition board) and the target response (as measured by the Nicolet Pro 10 digital oscilloscope for a sample rate of 100 Samples/s and voltage resolution of 3.2 mV for the response and 0.8 mV for the board output) to the 1 V, 5 V and 8 V inputs is shown in Figures 4.3, 4.4 and 4.5. See Appendix D for a more complete catalog of waveforms. The cross carriage is clamped in place by two machine screws to prevent motion and the capacitance gauge probe is held as described in section 3.5. The flexures have a 5 μm preload.

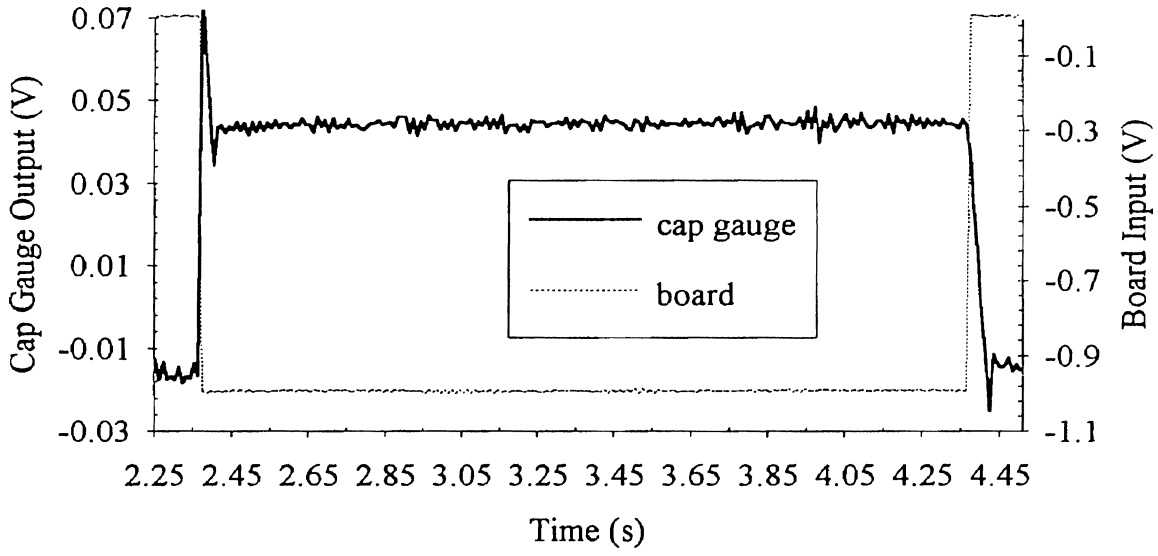


Figure 4.3. 1 V Step Response

There are a number of important features in these graphs. Figure 4.3 shows a 40% overshoot which is characteristic of an underdamped system. However the per cent overshoot in a linear system depends only on the system properties and

not on the initial step height. The relative amount of overshoot decreases as step height increases showing another non-linearity in the system (in addition to the hysteresis). This pattern was present for several step inputs of the same voltage taken on different days. The largest expansion is approximately 16 μm which is within the linear range of the flexures. It is known that when switched rapidly PZ stacks due exhibit a ringing behavior. However the amount of overshoot for an (unloaded) maximum expansion with a series resistance comparable to the present one is typically that shown for the 1 V response (30% to 40%). The reduction in overshoot for the higher voltage responses may be due to non-linear friction between the cylinders and the alignment tape. Another reason for the reduced overshoot may be inertial forces generated by the mass of the inner cylinder. The reader is referred to Appendix A.

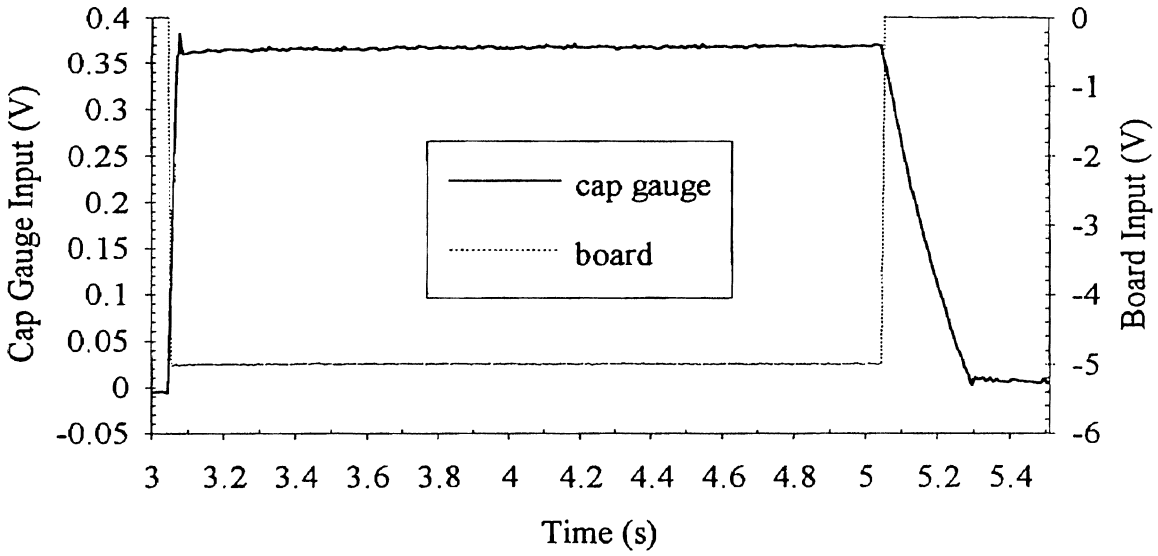


Figure 4.4. 5 V Step Response

Although the stack does not overshoot as the voltage increases there can be a 'ringing down' phenomenon caused by the generation of tensile internal forces as the voltage decreases suddenly [O'Niell et. al., 1980]. The stack used here has an

internal mechanical preload which prevents damage from such forces. Figure 4.3 shows that ringing down occurred for the 1 V input.

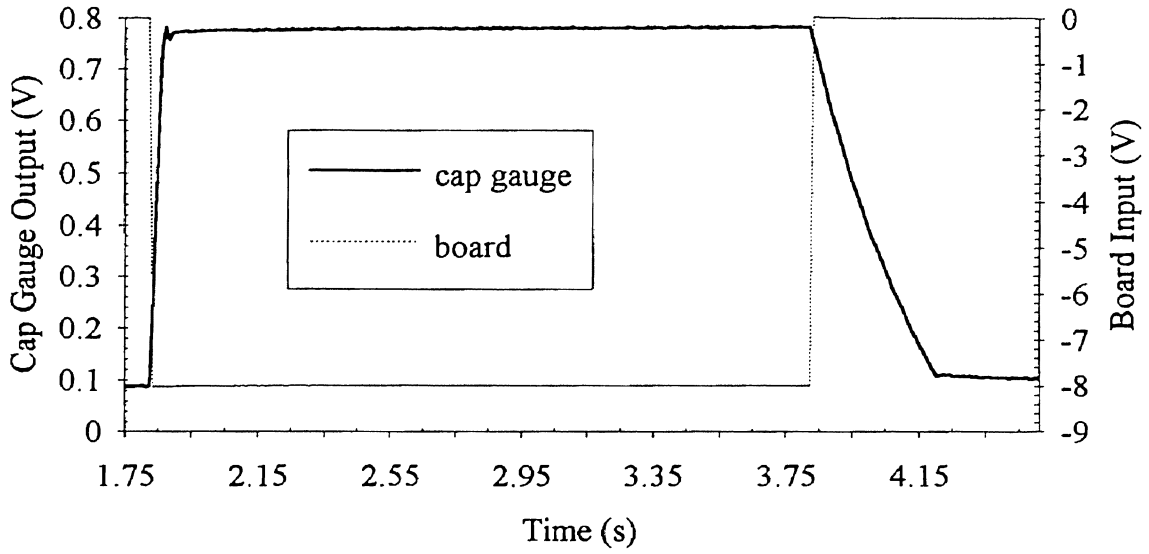


Figure 4.5. 8 V Step Response

In the waveforms for voltages greater than 1 V a rapid decrease in voltage does not show an equally rapid decrease in target position. The magnitude of the slope for the settling time increases steadily from 1 V/s for the 1 V step input to 1.9 V/s for the 8 V step input. While this increase may be due to the friction previously mentioned the most likely explanation lies with the set up of the amplifier.

It is essential that there be no positive voltages applied to the stack. When the amplifier was set up the voltage limiter was set so that the maximum voltage that could be applied was approximately -0.5 V. At the same time the current limiter was set to limit the maximum positive current to approximately 1 mA (which was the best resolution that could be got from the amplifiers front panel analog display). The sign of the current as the stack expands is the same sign as the voltage (negative). However step function contraction requires a positive current (even though the voltage remains negative).

The slopes of the voltage decreases may then be the leaking of charge off the stack as determined by the preset current limit. The different slope for each step is due to the fact that the capacitance of the stack is a function of the applied voltage. The actual form of the increase in capacitance is not known. If it is assumed the capacitance increases linearly by 30% from 0 V to -1000 V and the charge is calculated by $q = CV$ then for each slope $i = dq/dt$ ranges from 1 mA to 1.2 mA. The smaller steps have the smaller currents; if the increase in capacitance is assumed to be logarithmic the currents are even more consistent. But generally this is compatible with the preset of approximately 1 mA. The plot of Figure 4.6 shows further evidence that this sloping settling time was not a property of the stack and cylinder. It was generated with the positive current limit increased. Whereas in Figure 4.4 the settling time was approximately 0.3 s in Figure 4.6 it is at most 0.01 s (sample rates and voltage resolution are the same as previously stated). Note also that the amount of overshoot is repeated.

To prevent aliasing the sample rate should be at least twice the characteristic time constant (defined as the reciprocal of the resonant frequency) of the system. The undamped resonant frequency of the stack and cylinder is approximately 150 Hz (See Table 3.1 on page 45.); if there is a substantial amount of friction present it is probably lower than this, indicating a time constant of about 0.01 s (consistent with Figure 4.6). Thus the sampling rate was too low to allow the reliable calculation of system properties from the step response graphs. The resonant frequency of the stack is 8 kHz. However the stack is not connected to the inner cylinder and so could not drive it at this frequency even if such a noise component was present.

To obtain an open loop frequency response plot a series of sine waves was generated by the computer using a LabVIEW VI. See Appendix C for examples of the VI's used. The computer generated waveform has five cycles consisting of

1000 points so that each cycle consists of 200 points. The amplitude of each cycle is 1 V. To avoid applying a positive voltage to the stack a -1 V dc offset is superimposed on the sinusoidal pattern; the voltage then varies from -2 V to 0 V. The waveforms are shown in a normalized form which does not show the offset. The output of the board and the response of the capacitance gauge was read by the Nicolet Pro 10. The sample rate for the 3 Hz and 10 Hz signals was 200 Samples/s; for the 50 Hz signal the sample rate was 1000 Samples/s (for all sinusoidal plots with frequency greater than 15 Hz the sampling rate was 1000 Samples/s). The voltage resolution was 0.8 mV for the board input and 3.2 mV for the response for all plots. The initial preload was light; approximately 2 μm to 5 μm. As with the step responses there is a catalog of sine wave inputs and responses in Appendix D.

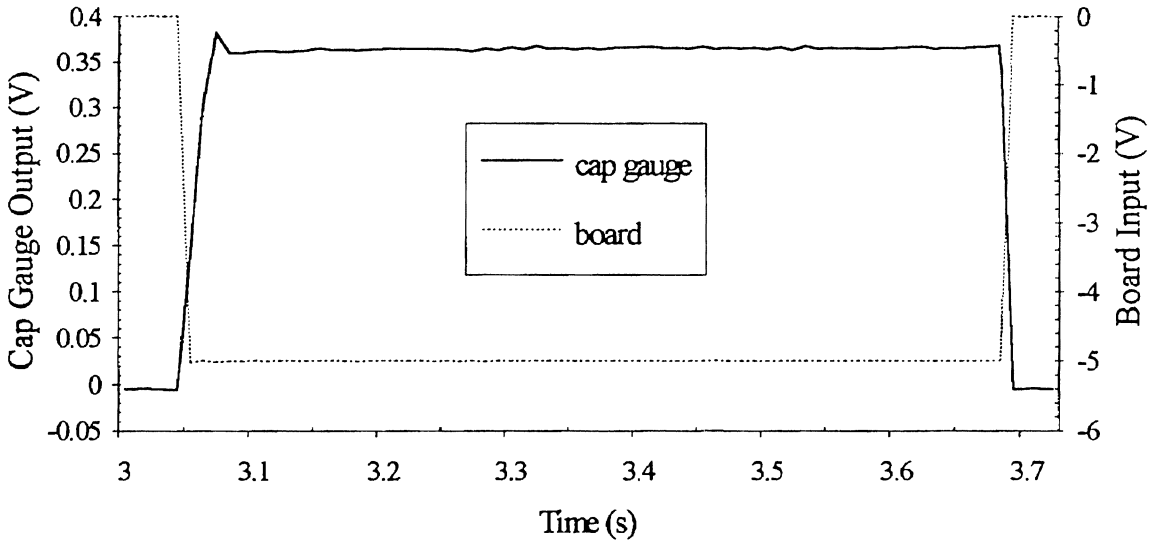


Figure 4.6. 5 V Step Response with Higher Current Limit

The lowest frequency (3 Hz) response of the plant (consisting of amplifier, stack, flexures and inner cylinder) is shown in Figure 4.7. Note that the apparent difference in phase is due to a sign difference. As the board output becomes more

negative this causes an expansion of the stack which registers as in increasingly positive output from the capacitance gauge. The voltage of the stack is in phase with the board input. The series resistor is too small (at these frequencies) to affect the phase between the amplifier and the stack.

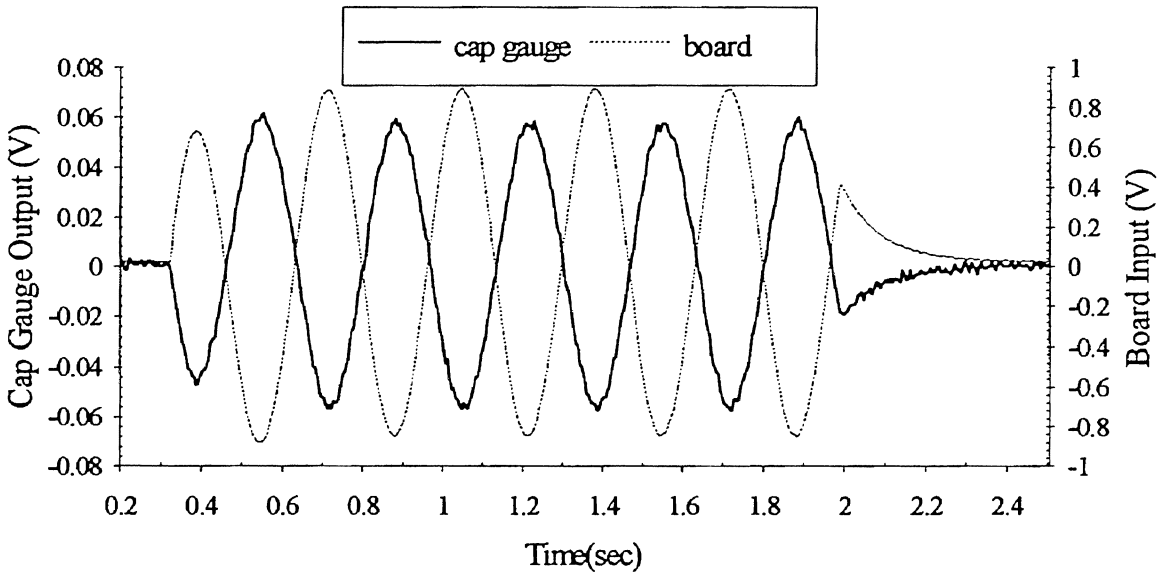


Figure 4.7. 3 Hz Response

The response of Figure 4.7 is sinusoidal to within the resolution obtained. The maximum expansion is 0.06 V which is 1.5 μm. The maximum expansion of the stack under static conditions is 20 μm for a -1000 V bias. Given the hysteretic behavior of the stack it is not surprising this expansion is not 1/10 of the maximum.

At 10 Hz the motion of the target shown in Figure 4.8 is not sinusoidal. The magnitude of the response is a maximum 0.06 V for the first peak then 0.05 V for the remaining peaks. In addition the only part of the wave that appears even slightly sinusoidal is the expansion portion (cap gauge response greater than zero). Note that when the cap gauge response is less than zero the current required is positive.

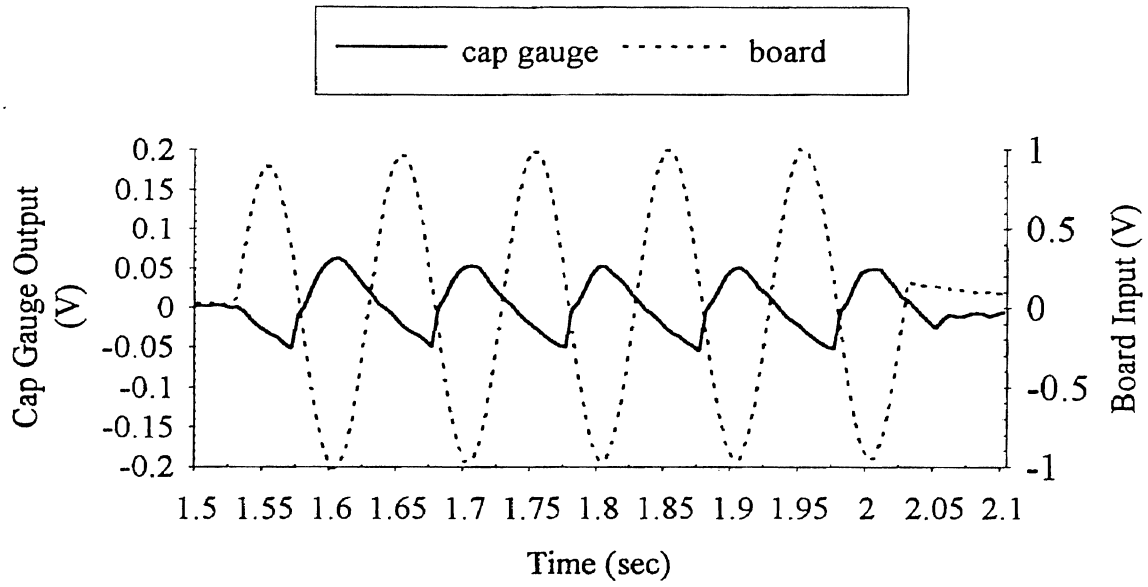


Figure 4.8. 10 Hz Response

The deterioration of the response continued as waveforms were recorded at 5 Hz intervals up to 50 Hz (See Appendix D). The 50 Hz response of Figure 4.9 shows the degeneration of the response. As with the other waveforms there is little motion for contraction as compared to expansion. The expansion portion however is not sinusoidal.

There are a number of possibilities for this response. The lack of contraction may be due to the current limiting setting previously mentioned. The current required in sinusoidal operation is dependent on frequency and at low frequencies the current needed is small; for a 3 Hz input less than 1 mA is needed (See Appendix A). The response of the stack directly (not acting on the cylinder) is not known (See Chapter V); thus it is not a certainty that the waveforms represent a mechanical deviation from the truly sinusoidal input. If the problem is that the stack does not follow the board input this does not necessarily imply that the stack is defective (although that is a possibility). As previously stated the BOP amplifier is very sensitive to reactive loads and the manual states that to obtain maximum

mechanical deviation from the truly sinusoidal input. If the problem is that the stack does not follow the board input this does not necessarily imply that the stack is defective (although that is a possibility). As previously stated the BOP amplifier is very sensitive to reactive loads and the manual states that to obtain maximum performance the load should be *mostly resistive*. The series resistor for this series of plots was 300 Ω and the reactance at 50 Hz is 6 kΩ. Thus the load is mainly reactive. It may be that under these circumstances the 300 Ω resistor is satisfactory for low speed (low frequency) input but the amplifier clips at higher frequencies because it cannot drive the load at higher speed (higher frequency).

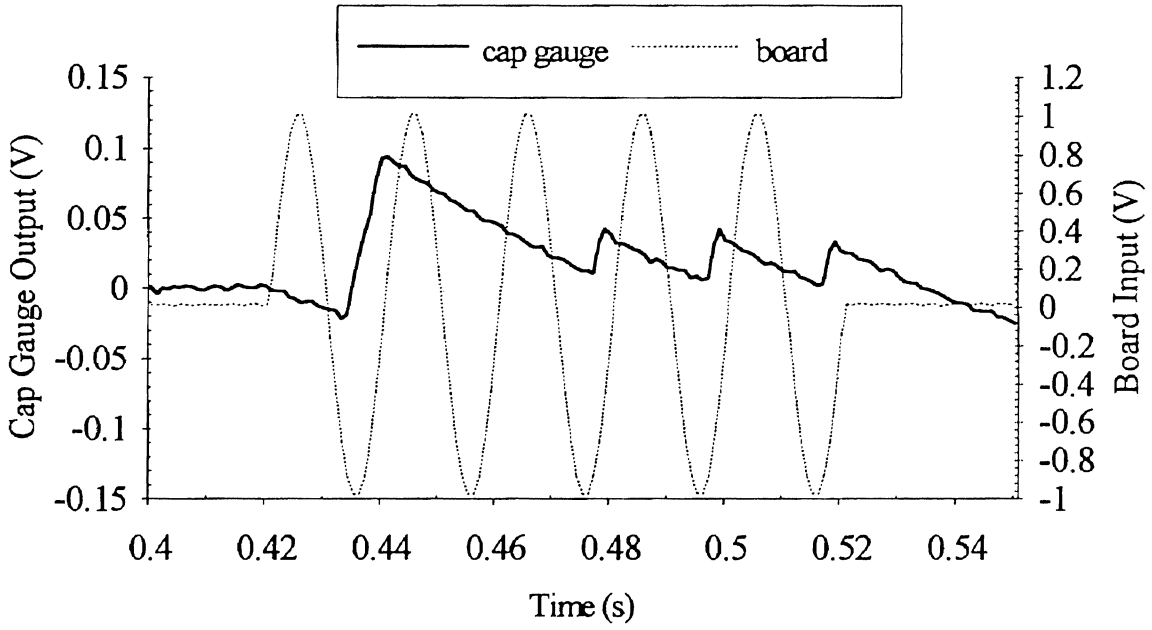


Figure 4.9. 50 Hz Response

Initially it was thought that at least part of the plot was due to backlash between the stack, ball bearings and inner cylinder. The preload was fairly small and so it seemed reasonable that the three components might become separated briefly as the stack oscillated, particularly if there was an appreciable degree of misalignment between these components. As the stack had been cycled at least hundreds of times the balls and the conical cups they rest in were examined under

the microscope for signs of unusual wear. The only signs of wear were in the aluminum cup at the rear of the inner cylinder and the wear land was well centered and uniform. Also numerous attempts to relate the time from the initiation of the waveform to the time of the first expansion to some characteristic length of the system (such as a possible separation of the ball and cylinder) did not yield consistent results. The same procedure applied to the maximum contraction interval gave the same inconclusive results.

4.5 Control System

Since the dynamic character of the amplifier/stack/cylinder/flexure combination is nonlinear and difficult to model it is treated as a "plant" that is to be controlled. A common method in this situation is to use PID (proportional-integral-differential) control with the initial value of the three gains set by measurements of responses made to known (simple) inputs. One of the most used is the Zeigler-Nichols method of which there are several variations [Ellis, 1991]. Since the controller here is a two term controller (PI only) the third gain is zero.

Recall from Chapter III that the algorithm is given by

$$v_n = v_{n-1} + \alpha[K_p(e_n - e_{n-1}) + \frac{K_I T}{2}(e_n + e_{n-1})] \qquad (3.31)$$

v_n is the output of the nth iteration and e_n is the error of the nth iteration. The constant α is arbitrary in the sense that it does not result from calculation on the control system; it merely scales the correction terms to the voltage scale of the output to the amplifier which can vary over several orders of magnitude depending on the desired position of the target.

To obtain the initial settings for the two gains the procedure recommended by Ellis was followed. The sample period (T_0) was set at 3 ms; this is sufficient as the resonant period of the inner cylinder is about 1 ms. First the integral gain was

set to zero and the proportional gain gradually increased until the system became marginally stable (sustained large amplitude oscillations). The value of the proportional gain at this point was 13 and is called K_{max}. Then the following formulas were used to calculate the initial gains for the PI controller [Ellis, 1991]

$$K_p = 0.45\ K_{max};\quad K_I = 1.2K_p/T_0. \tag{4.1}$$

The values recommended are then $K_p = 5.85$ and $K_I = 2000$.

Three tests were performed on the control system. Because of the failure to obtain good open loop dynamic response and the uncertainty as to its cause the tests were limited to step commands, hysteresis control and an indenting of polycarbonate.

To observe the response to a step command the amplifier was first set to 150 V and the cap gauge output to 200 mV. Then using the LabVIEW VI the command (reference value) was also set to 200 mV. The constant α was set at 0.0008 (A setting arrived at by trial and error as a value that kept the system stable.). Then the command value was discontinuously changed to 300 mV. The new command was recieved 3 ms after it was sent (one cycle if the VI). This procedure was repeated for a number values of α over a period of several hours. Some representative results are shown in Figure 4.10.

The response curves were recorded with the VI itself and the individual points measured from those curves and placed on a spreadsheet for display. The offset was measured by the Fluke 88 multimeter which has a max/min setting which allows the filtering of noise. For $\alpha = 0.0008$ the response held steadily at 300.0 mV, although it takes over 0.5 s to reach this value. For $\alpha = 0.0025$ the response had a DC offset of 2 to 3 mV and for $\alpha = 0.0035$ the DC offset was 5 mV. This offset may have been due to a low setting for the integral gain which was accidentally set at 1000. The response time for the high value of α was faster as the new reference was reached in less than 10 ms. When the new value had been

reached the target was then given a small push by hand (an external disturbance) to see if it would come back to the 300mV setting; it did seek the command for all settings.

The next test was for hysteresis control. The reference value is slowly changed (over approximately 3 s) as the VI is running and the amplifier voltage is read. This procedure is repeated for increasing and decreasing displacements several times. The amplifier always returned to the same voltage (for a given displacement) to within 0.5 V. The constant α was set at 0.0008. The sampling period remained at 3 ms. The results are shown in Figure 4.11.

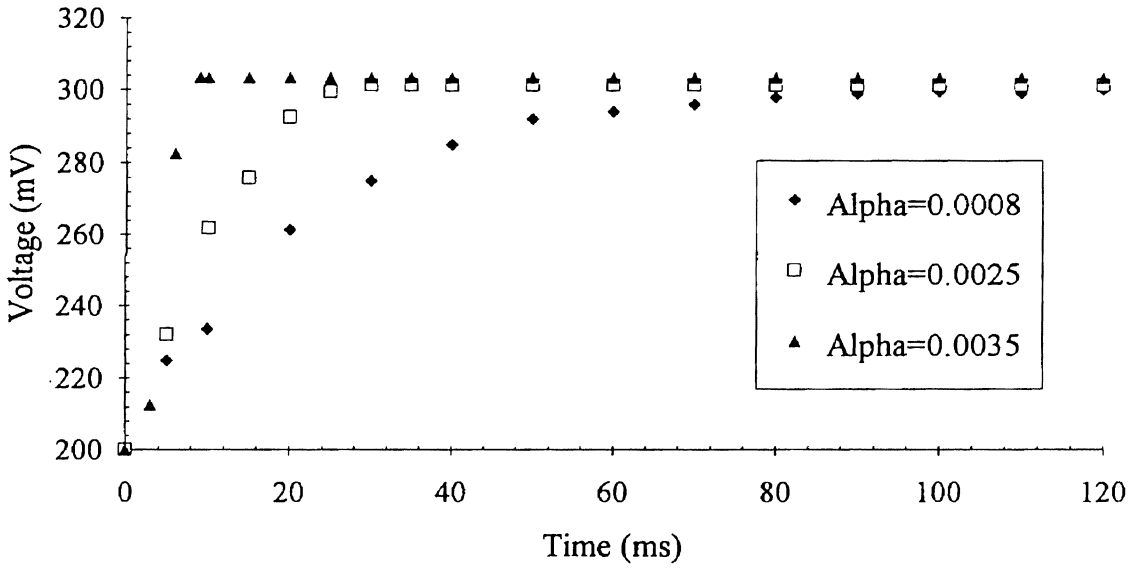


Figure 4.10. Control System Response to a Step Command

For a final test a polycarbonate sample was mounted and the wedge indenter (Shown on page 53) was used to indent the sample at a reference depth of 5 μm. Again the small value of α was used and the sample period was 3 ms. Contact with the surface was assumed to occur at the first visible sign of deformation when viewed under the microscope. At this point the 5 μm command is given (200 mV) and the indenter pressed into the sample. A photo is shown in Figure 4.12. The

scale was obtained by photographing an etched gold grid with a known spacing of 5 μm between grids under the same objective immediately after the indent. The load could not be calculated from this stress field for two reasons. First the fringe order cannot be determined from the photo; the widely spaced fringes at the bottom may be of order ten to twenty. Second the photo was taken in polarized light without any sort of waveplate. Thus the fringes include both the isoclinic pattern (for principle stress direction) and the so-called isochromatic pattern (for principle stress difference).

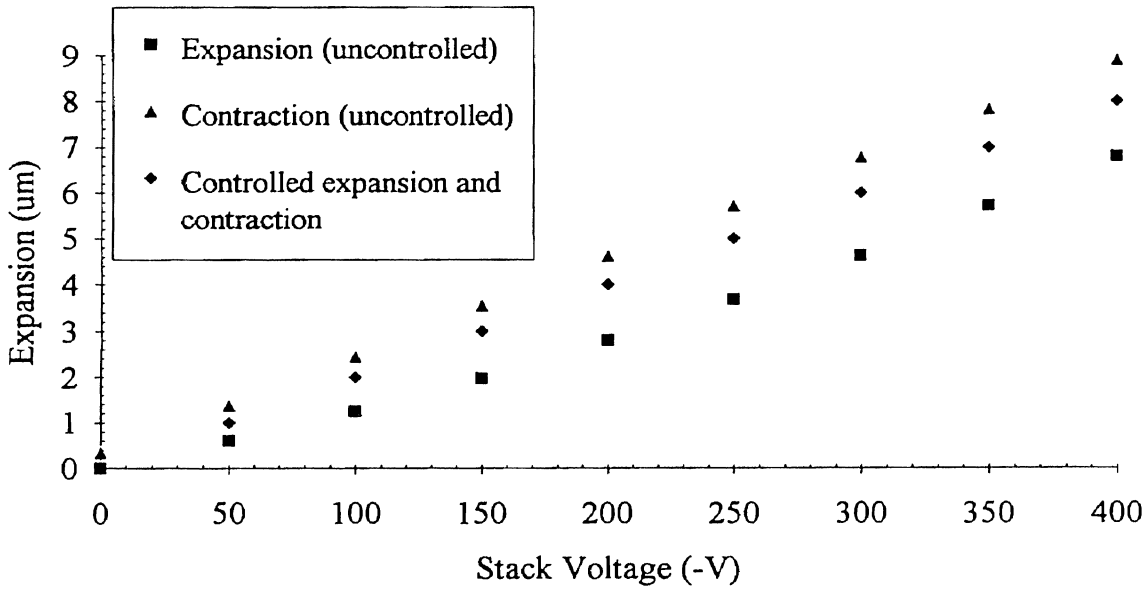


Figure 4.11. Hysteresis Control.

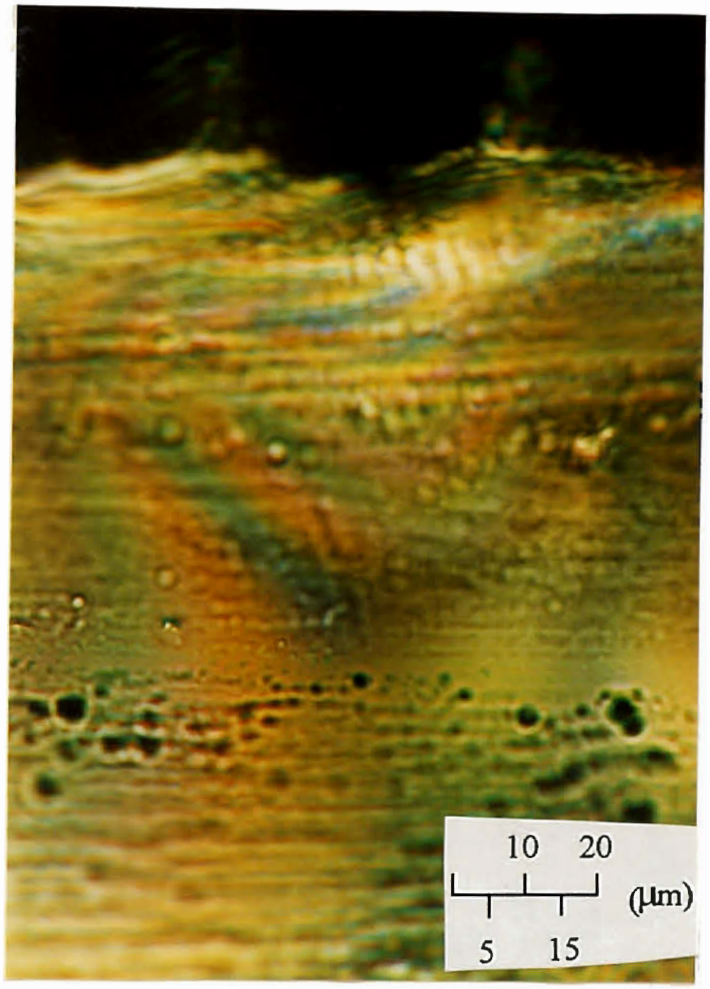


Figure 4.12. Stess Field in Polycarbonate

4.6 Optical Performance

The microscope optics can resolve images very nearly at the theoretical limits set by diffraction. For visible light the nominal wavelength can be taken as 550 nm and the numerical aperture of the 100x objective is 0.95; giving a resolution of 750 nm (See Appendix B). Under the 100x objective 1 μm diameter spheres were clearly resolvable. However 0.5 μm spheres were not resolvable as spherical shapes. Under high magnification the depth of view is small-less than 20 μm.

With the MMI under the microscope a number of qualitative tests were done to establish general operating procedures. Using the wedge indenter several pieces of Plexiglas and polycarbonate were indented. Due its smaller fringe value polycarbonate exhibits observable birefringence for smaller indents (as small as 2 μm). It is not surprising that polycarbonate is a standard model material used in stress studies. However, by the time a fringe is seen for the Plexiglas significant plastic deformation has occurred. It is also found that sample preparation is important. The polymer sample should be polished on both sides to maximize transmission of light and to minimize stray birefringence. The sample of Figure 4.12 had been polished but over time had become scratched. The surface of the sample holder and the flat portion of the tool holder were machined so that they are perpendicular to the objective (the frame is used as a reference surface); this provides a focused image across the entire field of view.

The quality of the edge facing the indenter is also important. It must be as square as possible; here square meaning that the normal and tangent of the edge surface must be perpendicular to the optical axis. In the case of Plexiglas this can be obtained by careful machining with an end mill. However polycarbonate is too soft to give a satisfactory edge when prepared this way. Machining on a diamond

turning machine at relatively small depths of cut (200 nm) produced far better surfaces than with traditional milling.

For birefringence the Nomarski prisms are not necessary and so were switched out of the optical path. In order to maximize the light transmitted the neutral density filter is also switched out. The best images are obtained for transmitted light; there are too many stray reflections when reflected light is used. The photo of Figure 4.12 is for transmitted light. In addition the image quality is enhanced when polarizers are placed directly above and below the sample/tool interface. The polarizers supplied with the microscope are dichroic sheet polarizers; again see Appendix B for details. The small working distance of the standard metallurgical objectives precludes this for magnifications higher than 10X, so the polarizers used were the ones provided with the microscope and are in the optical path several centimeters away from the sample.

4.7 General Positioning Performance

The overall movement of the MMD assemblies was also observed under the microscope using a microfilar attachment. The microfilar is an eyepiece with an adjustable vernier scale which can be used to measure lengths under the microscope. Under the 40x objective the resolution of the microfilar is 1 μm. The zero distance line of the microfilar is a single line that runs across the entire field of view. By positioning this line parallel with the motion and watching a single point on the object to be observed (in this case the indenter) a measurement of displacement perpendicular to the axis of motion can be obtained. To establish the reference axis the zero line of the microfilar was aligned with the axis of the x-y stage. The procedure was repeated several times to measure consistency. Coarse positioning in the x and y directions produced no noticeable runout; however

coarse positioning of the sample holder (the z-direction) did result a deviation in the negative y direction (toward the tool) of 20 μm, but only at one particular point on the vertical travel. This may due to a bad spot on the coarse adjustment screw thread. This might require repositioning the sample under the objective if the height of the sample relative to the tool is changed. However when indenting there was no noticeable motion of the sample holder (which is fixed by four large bolts).

Extending and releasing the cylinder/flexure assembly with the preload bolt found that the tool returns to the same position to within about 100 nm (This observation made by use of the capacitance gauge and the Rank-Pnuemo gauge head with the MMI not under the microscope.). This may be due to stiction caused by the tape layer used for alignment, a small amount of fretting which could probably be alleviated by the application of adhesive to the contact surfaces of the flexure and clamp [Slocum, 1993] or by a small temperature difference between the inner and outer cylinder. The inner cylinder is in contact with the external environment through the inner and outer clamps only. Any changes in temperature must be transmitted through those points.

When extending and indenting using a voltage applied to the stack there is no observable deviation from the y-direction. This last observation must be made with the microfilar as a special jig would be required to observe this with the capacitance gauge. This indicates there is negligible Abbe offset error (due to the motion of the cylinder) at the capacitance gauge target (See page 51 for an estimate of Abbe error).

4.8 Thermal Errors

Flexures tend to be prone to thermal errors due to their large ratio of surface area to volume. The use of aluminum is an advantage here as its high thermal

diffusivity (0.3 m^2/hr) allows aluminum to respond faster and makes it less likely to develop thermal gradients [Slocum, 1993].

An estimate of the thermal sensitivity of the flexures can be made using a simple model [Slocum, 1993]. In this model the expansion of one side of a rectangular flexure is assumed to expand linearly by an amount Δw and the thermal sensitivity is then defined as the angle φ (shown in Figure 4.13) divided by the length parallel to the temperature gradient. The error introduced by a thermal gradient parallel to the z-axis is shown in Figure 4.13. The thermal sensitivity (s_t) is given by

$$s_t = \frac{\phi}{\Delta T} = \frac{w_0 \alpha}{L} \tag{4.3}$$

where α is the linear coefficient of thermal expansion.

For our flexures the thermal sensitivity is about 12 μrad/°C. If the cutting surface of the sample is to be kept parallel to the optical axis (to within the minimum depth of cut) the maximum angle of the sample surface makes with the optical axis is 1×10^{-4} rad (for a 0.5 mm thick sample). Then the maximum temperature difference across the flexures is a very acceptable 80 °C.

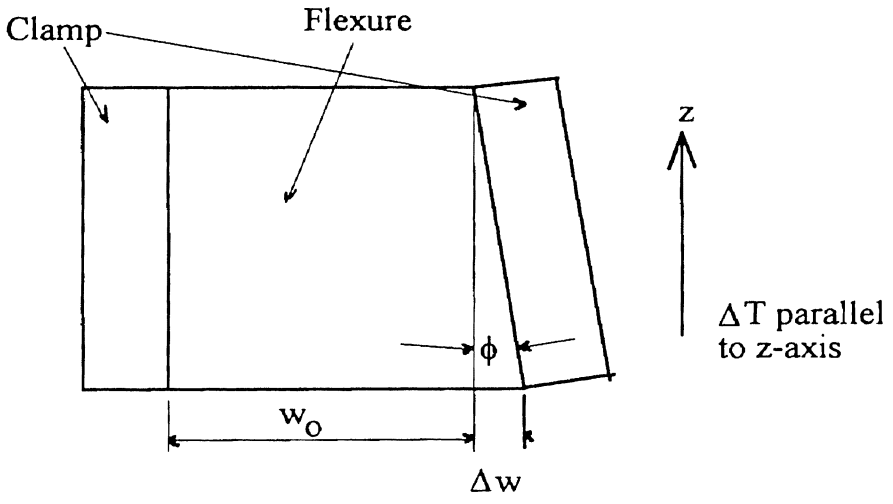


Figure 4.10. Rectangular Flexure Expansion

If the thermal gradient is across the cross-carriage in the x-direction the tolerance is not so large. For the same sensitivity to hold the motion to within the minimum depth of cut (for a sample 15 mm wide) the maximum temperature gradient is 0.3 °C.

Because they are smaller than the rectangular flexures and because they have no free edges the circular flexures are not as sensitive to thermal gradients [Slocum, 1993]. They are not affected by thermal gradients in the y-direction. To hold tolerance in the x- and z-directions the allowable thermal gradients are even larger than for the rectangular flexure z-direction.

The coefficient of linear thermal expansion for the stack is 0.5 μm/°C. A temperature variation of 2°C will cause the stack length to vary by more than the minimum depth of cut.

The thermal diffusivity of steel (0.043 m^2/hr) is smaller than that of aluminum by almost an order of magnitude. This implies that for a given change in temperature it will take the steel about 10 times longer to come to equilibrium. But the mass of the flexures is less than 1/100 that of the aluminum so that gradual changes in temperature over time should not be a problem. However rapid changes such as those caused by traffic in and out of the lab may pose a problem.

There are several sources of thermal energy. The closest are the illumination sources of the microscope; in particular the reflected light source is directly above the sample holder section of the MMI. These are halogen light sources that get warm to the touch even though they are enclosed. Another source is thermal energy generated by persons in the micromachining lab itself. The presence of two or three people can cause expansions of 50 nm to 150 nm; the variation becomes worse if there is substantial traffic into and out of the room as this causes significant changes in the ambient temperature. The only way to deal with the people is to keep them out of the lab for extended periods or to surround the MMI

with a temperature controlled environment. The halogen lights might be dealt with (as a first attempt) by putting a shield of aluminum foil between the MMI and the light. This would reduce the IR radiation falling on the MMI but could not affect the rise in ambient temperature due to convection. The control system may also be able to reduce the effects of such disturbances.

4.9 Life of the Flexures

Since the flexures will be operating dynamically the question arises as to how long they will last. Although the exact type of stainless steel is unknown the flexures are magnetic thus indicating a body-centered cubic stainless. Assuming the flexure material exhibits a knee in the S-N curve; a generalized S-N curve for the flexures would then have a knee at about 490 MPA. Since the maximum stress should stay below 20% of the yield strength of the flexure material the life of the flexures should be infinite.

CHAPTER V

SUGGESTIONS

5.1 Introduction

The development of this instrument has required techniques from several engineering backgrounds: mechanical design, control/electronics and optics. Each area has certain problems that need to be addressed.

5.2 Mechanical Design

Overall the mechanical design seems basically sound. There are several important modifications which need to be made and would not be difficult to execute. The two major areas of concern for the tool holder are the stack cylinder and connection, and the dynamic response of the tool holder itself.

5.2.1 Tool Holder Assembly First there is the matter of the stack and inner cylinder interaction. Currently there are two 5/32" ball bearings at each end of the stack; one bearing pushes directly on the inner cylinder while the other pushes on a preload bolt. It is advisable to leave at least one (the rear bearing acting on the preload bolt) in place although it ought to be increased in diameter or replaced with one made of Tungsten Carbide to increase stiffness [Slocum, 1993]. Also the bearings should be clean. There is currently a thin film of grease on them to aid in assembly. A small annular ring on the end of the stack or cylinder could aid in assembly and do away with the need for the grease.

At the other end of the stack/cylinder assembly there are four alternatives. The first would be to leave the design as it is now and replace the front ball with a stiffer one. This has the advantage of requiring little or no modification of the design and will compensate for any misalignment. The second option is to attach the inner cylinder directly to the stack. Masses are commonly attached to stacks which operate dynamically. Indeed PI stacks are all available with an optional mounting system for just such an application; the stack currently in the MMI is outfitted with this option. However in the present case the mass is constrained to move in only one direction; any off- axis forces applied will risk damaging the stack. It may be that one bearing at the rear of the stack is sufficient to prevent this; at any rate before this type of mount is implemented the manufacturer should be consulted. The author knows of no tool servo applications where direct mounting is used. The advantage of such a mount is that the resonant frequency will be determined by the stack resonant frequency and the cylinder mass, which is much higher than the cylinder/flexure resonant frequency (See Appendix A). The third option is to grind the mating surfaces at the front and back of the stack parallel. The advantage of this arrangement is that there is no compliant ball bearing between the stack and the inner cylinder. This mount is used in at least two fast tool servos [Patterson, 1988: Moriwaki, 1985]. It appears to be used also in Hosler's grinding servo but his discussion is not clear on this point. The fourth alternative is to provide some external preload such as a Belleville spring. This alternative is the least desirable due to the expense of implementing and it may introduce more noise. A discussion of this final possibility is included in Appendix E.

The most important step to be taken next is the resolution of the dynamic properties of the stack. The proposals made in the preceding paragraph may not even need to be considered if there is indeed no mechanical backlash. It was not

established that the poor responses shown in Chapter IV were due to mechanical backlash. Each component should be individually tested to answer the following questions: 1) Can the amplifier follow the computer output with the current value of series resistance?, 2) Does the mechanical output of the stack follow the amplifier? and 3) If a larger resistance is needed, what value is needed to get the amplifier to properly drive the stack?.

Such a test could be carried out as shown in Figure 5.1. The stack can be mounted to the vibration isolation table and a washer with a lapped surface glued to the stack facing the capacitance gauge probe. A ground wire should be soldered to the washer. For a given input the output of the capacitance gauge can be read thus revealing its response characteristics and the output waveform of the amplifier can be determined by measuring the voltage across the resistor. The voltage across the resistor will depend on the relative values of the resistance and the impedance of the stack. Three channels need to be observed and compared: the DAQ board output, the voltage across the resistor (or the output direct from the amplifier, if it is set low enough to avoid damaging the oscilloscope) and the capacitance gauge output. Since the oscilloscope has only two channels the tests will need to repeated and compared or a LabVIEW VI could be written to separately observe the gauge output (A VI for this purpose is in the LabVIEW VI library directory under the filename SIMPLE.VI).

Another point to consider when evaluating the dynamic performance of the stack is the voltage amplitude. Since the dynamic performance is needed to correct for (ostensibly) small errors, the amplitude of the stack needs to be tested for small amplitudes. Limitations at large amplitude do not imply the same limitations at small amplitude. The tests of Chapter IV were for a 1 V amplitude which is an expansion of approximately 2 μm. This is very large compared to the

minimum uncut chip thickness specified. A series of tests as discussed above should be done for voltages an order of magnitude smaller as well.

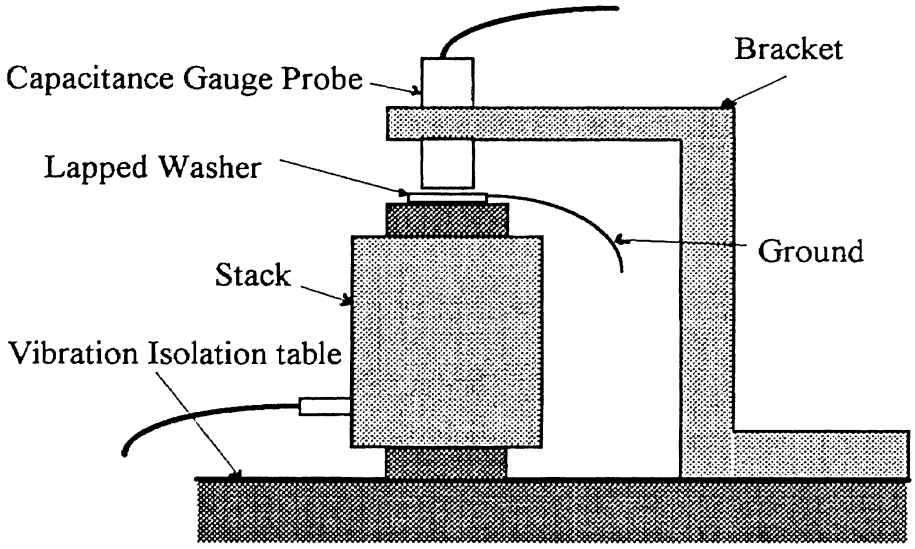


Figure 5.1. Test of Stack Response

Increasing the value of the series resistance will have at least some detrimental effects on the MMI. While it may improve amplifier response it will decrease the rise time of the stack to step inputs. The resistance will also introduce a phase difference between the amplifier voltage and the stack voltage according to the formula $\phi = \tan^{-1}(R/X_C)$. This will bear on the response of the control system.

Damping may still remain a problem as well. It appears for the current configuration the damping may not be linear; hence it may be too little at low speed and too much at high speed. One solution may be to use multiple sheets for each flexure and coat the inside of each layer with an RTV silicone adhesive. Although this has been shown to work [Slocum, 1993] it might prove difficult to implement. The aspect ratio of the flexures is relatively high and since the adhesive can only be applied to those areas which are not in contact with a clamping surface (this criterion itself might be hard to verify) there will be very

little adhesive under the flexures. Also, the precision of the assembly of clamped flexures is extremely important [Slocum, 1993] and the presence of multiple layers makes manipulation of the flexures difficult; the alignment of bolt holes for the inner and outer clamps would have to be maintained. If the alignment tape is removed the inner cylinder will be more difficult to center in the outer cylinder. Proper alignment may call for locating pins in the inner and outer cylinder.

One other possible solution to the damping problem would be to expand the current damping mechanism. Wrapped around inner cylinder at the front and back are two thin layers of cellophane tape. These were very carefully applied, using nearly constant tension and preventing layers from overlapping. The original motivation was to provide a method of centering the inner cylinder as the clamping bolts are tightened. But the tape also provides some frictional damping. This might be carried further in that the entire cylinder might be wrapped with nylon shim stock to center the cylinder and provide a friction surface. As the contact would be light it might be necessary to alter the surface texture of the inner cylinder since at small loads the surface texture would strongly influence the frictional characteristics.

Another important characteristic to consider when considering changes is the resonant frequency of the cylinder/flexure assembly. With a stiffness of about 1 N/μm and a mass of 1.1 kg the undamped resonant frequency is about 150 Hz. For the range of cutting speeds set at the beginning (0.1mm/s to 10 mm/s) the control system could at best sample from 375 times to 3.75 times, respectively.

The response of the system can be improved in three ways. One has already been mentioned; attach the stack to the cylinder. The second method is to lower the mass of the inner cylinder by changing its dimensions. The cylinder can be made thinner; but this will also lower its stiffness. The cylinder can also be made shorter, raising its stiffness in the process; although this would require shortening

the outer cylinder as well. The original length of the cylinders was determined by the desire to minimize off-axis angular errors. The longer the cylinder the smaller the errors (See equation 3.16). It may be that a shorter cylinder will still keep angular errors within acceptable limits. The influence of changes in length and diameter on resonant frequency are shown in Figures 5.2 and 5.3.

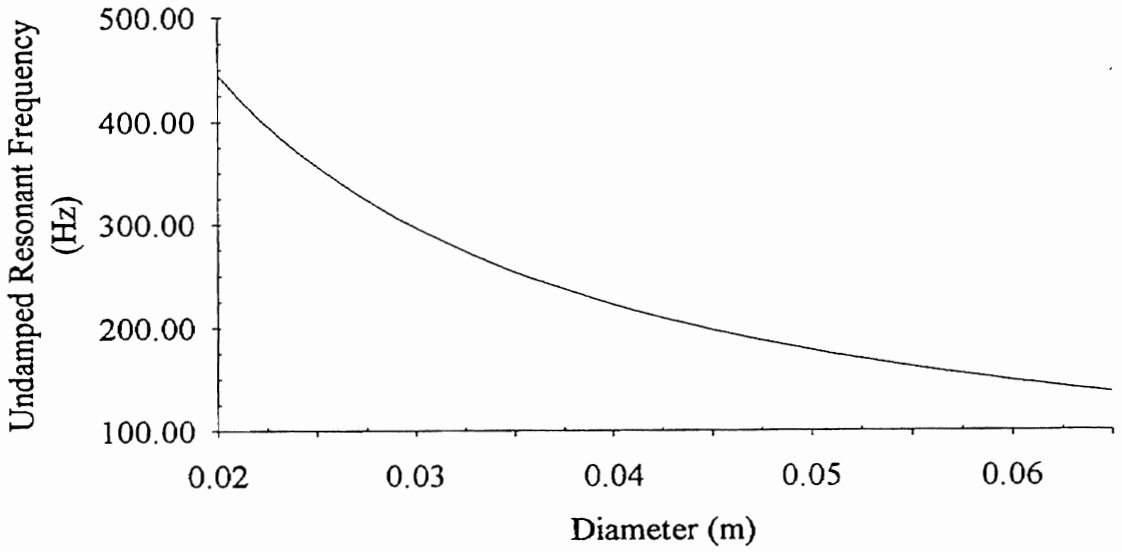


Figure 5.2. Effect of Diameter Change on Resonant Frequency

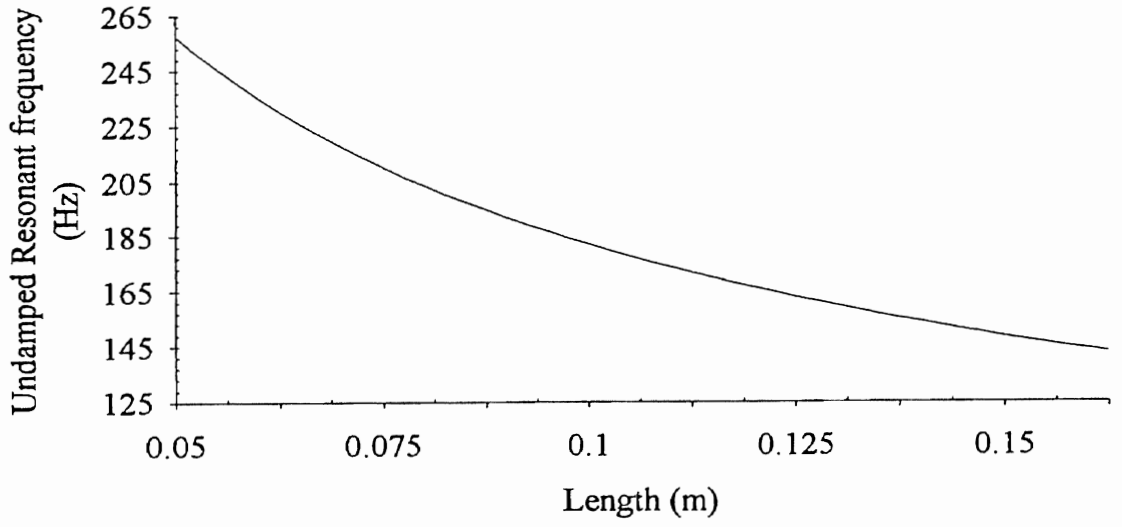
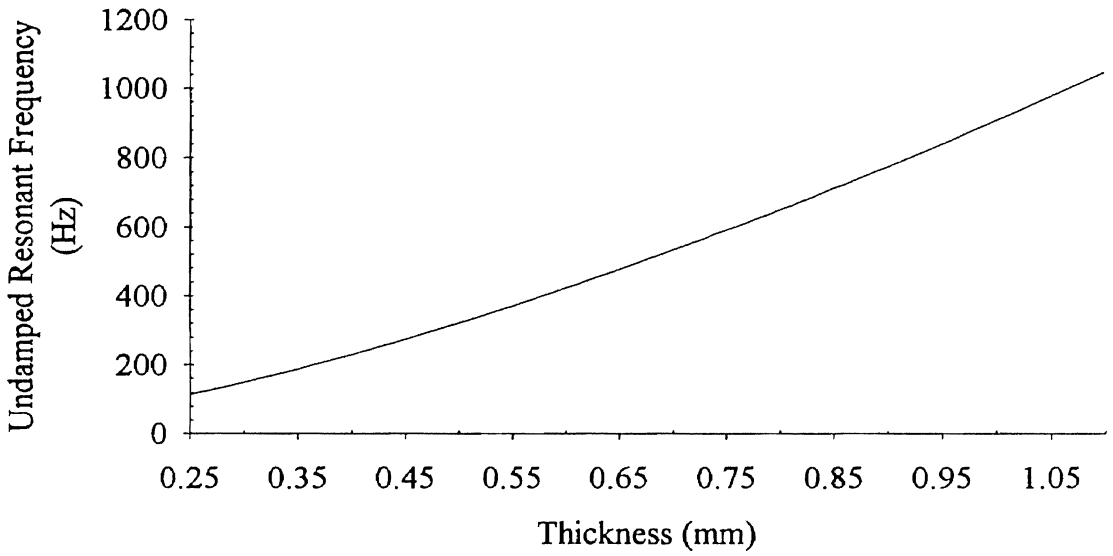


Figure 5.3. Effect of Length Change on Resonant Frequency.

The third way to improve system response is to increase the stiffness of the flexures. The flexures currently installed are thinner and hence less stiff than the maximum allowable for a 30 μm expansion. The maximum thickness for a 30 μm expansion is 0.55 mm. Thicker flexures can be used for smaller expansions. The maximum stress is directly proportional to the thickness. Thus a 20 μm maximum expansion implies a 0.825 maximum thickness. The effect of flexure thickness on resonant frequency is shown in Figure 5.4, using equation 3.18.



5.4. Effect of Flexure Thickness on Resonant Frequency

5.2.2 Sample Holder Assembly. The sample cross carriage needs some modification as well. In the current design the flexures are pinned at each end and each flexure is made of several thin sheets (in anticipation of the application of an RTV silicone as mentioned earlier). To obtain smooth motion from each flexure with this configuration each sheet on each side and the two opposite sides must be very nearly the same length (probably to within 1 μm or so) [Slocum, 1993]. This level of precision was not obtained in the original (and it would be impractical to

expect it from the facilities presently available). As a result the cross carriage flexures exhibit a "break-over" phenomenon as they travel through the x-axis motion. There is resistance to motion as the flexures move along the x-direction; but when a certain point is reached the carriage suddenly accelerates to the stop at the other end of the axis. This break-over feature occurs over a travel ten times longer than the planned length of cut; however it indicates an instability which should be eliminated. A single thickness of spring steel on each side would be better than the multiple thickness now used. In addition the flexures should be put in pure tension to avoid this behavior.

Figure 5.5 illustrates a new design for the cross carriage frame which is very similar to the present frame (viewed from the top). This new flexure scheme puts the front part (the part facing the tool) of the flexures in tension. This particular design has been shown to provide one-dimension motion with an off-axis error of less than 100 nm over a 2 cm travel [Biernson, 1988].

In addition, this design exhibits virtually infinite adjustability by varying the thickness of the rear tension springs. Since the tension springs act as cantilevers under a moment load the deflection for a given moment is inversely proportional to the third power of the thickness. It would even be possible to vary the response of each flexure individually by using a different thickness for each spring and using a shim under the clamp. Another advantage is that by shortening and increasing the thickness of the rectangular flexures the stiffness in the control direction is increased.

As another alternative tuning blocks as shown in Figure 2.6 on page 29 might be considered. Tuning blocks can yield order of magnitude improvements in resolution [Slocum, 1993].

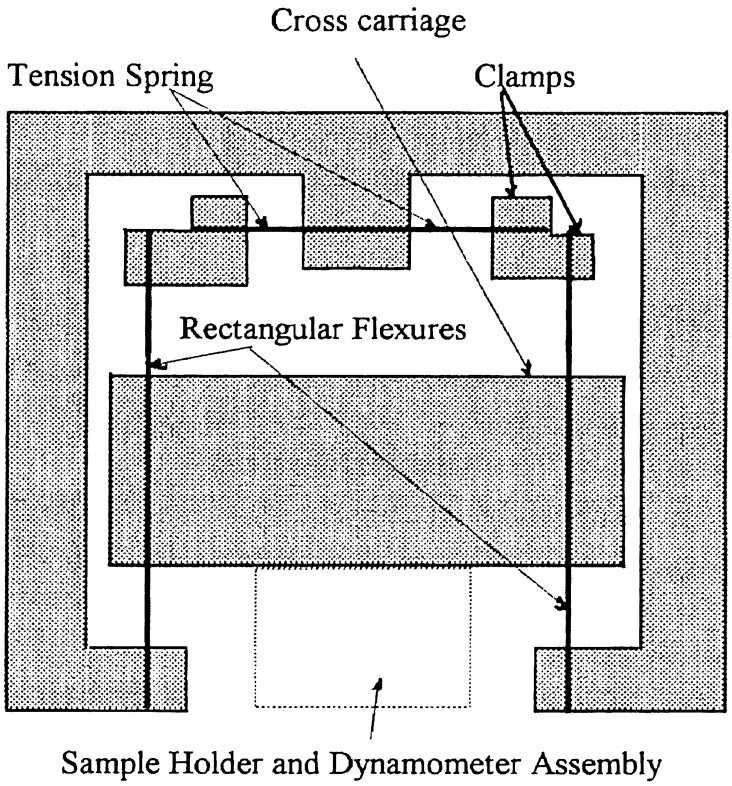


Figure 5.5. Alternative Flexure Design.

5.3 Optics

Since there was no design for the microscope and it works extremely well there is little to suggest for its improvement. There are a number of peripheral items worth discussing.

Observation under transmitted light is the best way to observe birefringence for this geometry. In fact transmitted light is used for birefringence only when stress or strain is to be observed at the interface between two materials, one of which is transparent. Most transmitted light observations are made with condensers in the optical path (See Appendix B.). The base and stage of the Nikon MM-11U do not provide for a condenser; there is a plane mirror to direct the light vertically through the sample. This is not unusual for this type of installation. The transmitted light is used for illumination of a sample only so that a distance

measurement can be made. The application here calls for filtering of the light using narrow bandpass filters. The narrower the bandpass region the lower the overall intensity of the light. Under these circumstances it would be quite helpful to be able to focus the light directly on the sample. There is not room in the stage or in the MMI itself for a condenser; even it there were the mounting of a condenser with its complex lens system requires expert optical knowledge.

A possible compromise would be the use of a Fresnel lens (shown in Figure 5.6). These lenses are very thin (1mm to 3 mm) and of very high optical quality. They are available in virtually any focal length and are commercially available made from plastics designed for infra-red transmittance. There is also a considerable cost advantage as a condenser can easily cost $1000 and Fresnel lenses cost $5 to $60. Needless to say a Fresnel lens does not have the overall optical performance of a condenser; but in the present context where the broadband performance is not essential it would make a reasonable alternative.

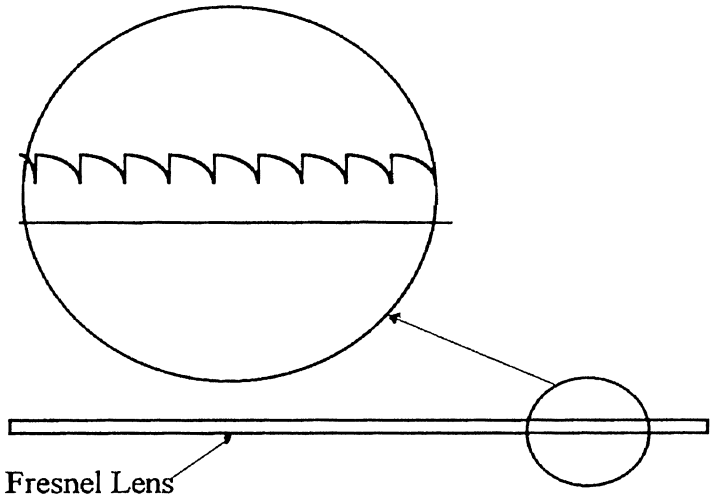


Figure 5.6. A Fresnel Lens

To use a Fresnel lens the side with the raised ridges is located away from the point upon which light is to be focused. The Fresnel lens then behaves as a thick lens focusing the light. See Figure 5.7.

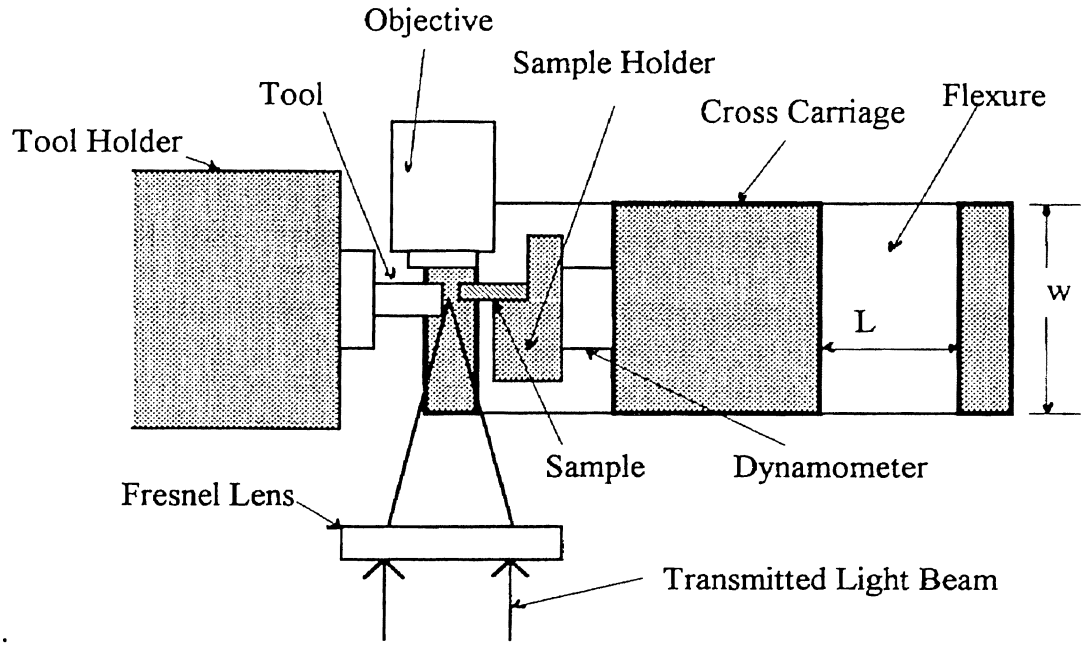


Figure 5.7. Application of a Fresnel Lens to the MMI.

One possible candidate is Oriel Optics part no. 43000. This Fresnel lens has a focal length of 22 mm and a diameter of 2" and costs $15. Thus it would fit between the sample holder and the base plate. Its response begins to fall off beyond about 1.5 μm. Companies such as Poly IR or Fresnel Technologies provide infrared Fresnel lenses for about the same cost.

5.4 Control System

The major problem with the control system was the constant presence of noise. The majority of the time the major component was a 60 Hz signal that varied from 1 mV to 7 mV. The claimed 0.5 mV resolution of the module board was never approached. There are two things that can be done to control this noise or limit its influence. The capacitance gauge probe target needs to be properly grounded (this point has been made earlier). The target also needs to be lapped. The LabVIEW system has a large number of filter sub-VI's (For a description of a VI see

Appendix C.) A notch filter placed in the capacitance gauge input to the computer centered on 60 Hz may help significantly.

The next modification of the algorithm itself may be to include a differentiating term to the controller. This type of controller (a PID controller) was used by Hosler on a similar design to obtain infinite stiffness up to 300 Hz [Hosler, 1990].

CHAPTER VI

SUMMARY AND CONCLUSION

6.1 Mechanical design

This thesis has presented the effort so far to design and build an instrument capable of emulating the cutting performance necessary to do single point diamond turning while under a research microscope and using a wide spectrum of illumination. Due to the complexity of the project it is to be accomplished in stages with the first stage being the design of the mechanical elements and testing of the electromechanical performance and the optical performance.

The design parameters set at the beginning push the chosen design (using clamped flexures) to the limits of what this type of flexure will do. The circular flexures supporting the inner cylinder give every indication so far that the design is quite sound. A full test (meaning to the limits of resolution of the instrument) of the off-axis deflections was not possible with the available equipment.

Static tests were performed on the tool holder assembly. The displacements registered by the capacitance gauge were compared to and are consistent with the displacements indicated by the Rank Pnuemo Model 1045 Gauge Head. Hysteresis of 14% was observed which is consistent with the manufacturer's claims. On a 2 μm step expansion a drift (continuing the expansion) of approximately 50 nm over a period of several seconds was seen as well. Observations of the motion of the tool as the stack was expanded over a period of

several seconds showed no noticeable deviation from the y-axis. The resolution of these observations was approximately 1 μm.

A series of dynamic tests were performed on the tool holder as well. The response to step inputs from 1 V to 8 V showed that the overshoot was a function of the step input, varying from a high of 40% for the 1 V input to 5% for the 8 V input. This is also a non-linear characteristic. Furthermore it appeared that the positive current limit setting on the amplifier influenced the settling time for the step responses. Frequency response characteristics were especially poor. The target yielded sinusoidal motions only for very low frequencies of less than 10 Hz. The cause of this is not clear; however Chapter V contains several suggestions to answer some of the questions raised by this behavior.

A number of improvements for the tool holder are suggested in Chapter V; mostly focusing on improving the frequency response of the tool holder. As it stands now the bandwidth of the tool holder even if the stack did drive it perfectly is not more than 100 Hz. The mechanical structure should be capable of attaining a bandwidth of at least 300 Hz (under feedback control) [Hosler, 1990].

The cross carriage flexures are more problematic. Certainly their performance could be improved as well with relatively simple modifications. However the performance required of them is high. Ideally there should not be more runout than the minimum depth of cut. However at 20 nm this small an amount of runout would have the flexure performance competing with the very best air bearing spindles, which typically have a runout (radial or axially) on this same order [Lucca, 1993]. Just in comparison to very good crossed-roller bearing stages, which have a runout of about 50 nm/cm, the performance is high. Hopefully the closed loop control will bridge the gap and provide the needed performance. At the very least the attainment of the highest level of performance from the flexures

will require continuous refinement, and suggestions on this point are offered in Chapter V as well.

6.2 Control System and Electronics

The non-contact gauging system works extremely well and needs only the most trivial of improvements; mainly the target needs to be lapped and fitted with a terminal screw for a ground. The labVIEW development system is also very remarkable. It is easy to use and can perform an array of functions that would consume huge amounts of programming effort for a conventional language.

The control system is an entirely different matter. Again the performance goal is high enough that refinement will be necessary. A proportional-integral controller was programmed and tested for step response, hysteresis control and indenting. The controller will follow step responses to within 1 mV (allowing for noise averaging) which could give a resolution of 25 nm. The highest resolution required the longest time to stabilize. Hysteresis control was very good as the stack could be commanded to return the same displacement for an amplifier voltage that varied by approximately 0.5 V. The control system was used to perform a 5 μm deep indent on a 1 mm wide polycarbonate sample. Noise was a major source of trouble with the control system and it is suggested that dealing with this problem take high priority.

6.3 Optical System

The optics are excellent as well. There is no reason to believe that the quality of the equipment will in any way determine the ultimate success of the MMI. There are many manufacturers that make filters, lenses and polarizers that fit Nikon products so that the performance can be tailored to the needs at hand.

6.4 Overall Appraisal

Consider the objectives on page 30. With regard to 1) it is clear that the design does not meet the original specifications. It is not clear that the design cannot meet those specifications. The main shortcomings are the bandwidth of the tool holder, the frequency response of the stack/cylinder assembly and the sample holder flexures. An addition shortcoming that was recognized at the start is that the current gauging device, even if noise is controlled, cannot meet the original specifications. With regard to 2) a control system was developed which can follow a command and compensate for disturbances. However given the dynamic response problems its main use was for indenting. An indent was done and that was objective 3). Finally many of the issues raised in 1) were addressed and a series of proposals and solutions was presented in chapter V.

6.5 Conclusion

The hope is that by the use of deterministic thinking the performance of the MMI can be brought to specification. Deterministic thinking in precision machine design is characterised by the belief that there are no random occurances in the behavior of machines. The development of instruments using this philosophy requires that the designer be patient and take nothing for granted [Bryan, 1984]. Continuity of effort is important to assure that workers that start later can benifit from earlier experience. Although a number of things were taken for grantes, such as the performance of the stack, it is now clear that those assumptions must be carefully re-examined. This project is the first attempt by this research group at the design and development of an ultraprecision instrument. As such it represents

the first phase of a continuing process. The specific contributions of this work are 1) the design and manufacture of a precision cutting device that is compact enough to fit under a microscope, 2) the testing of the instrument and working out of some of the details and problems incurred when trying to integrate several different systems and 3) hopefully pass on the experience gained. Not every test was a success, but procedures were suggested to help resolve questions.

REFERENCES

Andreeva, L. E. (1966). Elastic Elements of Instruments. Jerusalem: Israel Program for Scientific Translations.

Bradbury, S. (1984). An Introduction to the Optical Microscope. New York, NY: Oxford University Press

Babonas, A. et. al. (1973). Birefringence of CdS Under Uniaxial Stress. Physica Statica Solidus (b). Vol. 55, 321-327.

Bagchi, A. and Wright, P.K. (1987). Stress Analysis in Machining with the use of Sapphire Tools. Proceedings of the Royal Society of London, A. Vol. 409, 99-113.

Biernson, J. (1988). Advanced Control Systems. New York, N. Y: Academic Press.

Blake, P. et. al. (1988). Precision Machining of Ceramic Materials. Ceramic Bulletin, Vol. 67, no.67, 1038 to 1044

Bolef, D. I., et. al. (1960). Elastic Constants of Hexagonal Cadmium Sulfide. J. Phys. Chem. Solids (GB), Vol. 17, No. 1, 143-148.

Bryan, J. B. (1984). The Power of Deterministic Thinking in Machine Tool Accuracy. First International Machine Tool Engineers Conference. Nov. 7-8, Tokyo, Japan. 39-53.

Chapman, G. F. (1988). Recent Developments in the Generation of Glass Aspherical Surfaces. Proceedings of the International Society for Optical Engineering, Vol. 1015, 36 - 44.

Davidson, A. (1972). Handbook of Precision Engineering, Vol. 6. New York, NY: McGraw-Hill Book Company.

Dorf, D. C. (1989). Modern Control Systems. New York, NY: Addison-Wesley Publishing Company.

Dally, E. and Riley, R. (1991) Experimental Stress Analysis. New York, N. Y: McGraw Hill Publishing Co.

Dieter, G. E. (1986). Mechanical Metallurgy. New York, N.Y: McGraw Hill Publishing Co.

Epstein, S. (1970). Scientific Instruments. Englewood, NJ: Franklin Publishing Company.

Ellis, G. (1991). Control System Design Guide. New York, NY: Academic Press

Furukawa, Y. and Moronuki, N. (1988). Effect of Material Properties on Ulta-Precise Cutting Processes. Annals of the CIRP., Vol. 37 No. 1, 113-116.

Gutmanas, E. Y. et. al. (1979). The Mechanical Behavior of CdTe. Scripta Metalurgica., Vol. 13, 293-297.

Gecim, B. and Weiner, W. O. (1991). Transient Temperatures in the Vicinity of an Asperity Contact. Journal of Tribology . Vol. 107, 333-342.

Griffel, A. (1984). Handbook of Formulas for Stress and Strain. New York, N.Y: F. Ungar Publishing Co.

Hartshorne, N. H. and Stuart, A. (1970). Crystals and the Polarizing Microscope. New York, NY: American Elsevier Publishing Co.

Hosler, J. B. (1990). Precision Grinding of Ulta Thin Quartz Wafers. Masters Thesis submitted at Boston University.

Ikawa, N. et. al. (1991). Ultraprecision Metal Cutting-The Past Present and Future. Annals of the CIRP, Vol. 40 No. 2, 587 to 594.

Kingslake, R., Ed. (1965). Applied Optics and Optical Engineering, Vol. I. New York, NY: Academic Press

Kingslake, R., Ed. (1965). Applied Optics and Optical Engineering. Vol. III New York, NY: Academic Press.

Lucca, D. A. and Seo, Y. W. (1989). Prediction of the Partition of Energies and Tempertaures in Orthogonal Ultraprecision Machining. Presented at the 5th International Precision Engineering Seminar of the ASPE.

Lucca, D. A. et. al. (1991a). Energy Dissipation in Ultra-Precision Machining. Proceedings of the NSF Design and Manufacturing Systems Grantees Conference. 263-267.

Lucca, D. A. et. al. (1991b). Energy Dissipation in the Ultra-Precision Machining of Copper. Annals of the CIRP. Vol. 40, No. 1, 69-72.

Lucca, D. A. and Oquin, T. R. (1992). Orthogonal Ultra-Precision Machining of Single Crystal Germanium. Proccedings of the Seventh Annual Meeting of The American Society for Precision Engineering. 211-213.

Lucca,D. A. et. al. (1992). Energies in the Ultra-Precision Machining of Ductile Materials. Proceedings of the 1992 NSF Design and Manufacturing Systems Conference. 123-129.

Lucca D. A. et. al. (1993). Energy Dissipation in Ultra-Precision Machining. Proceedings of the 1993 NSF Design and Manufacturing Systems Conference. 225-267.

Lucca, D. A. and Seo, Y. W. (1993). Effect of Tool Edge Geomerty on Energy Dissipation in Ultraprecision Machining. Annals of the CIRP. Vol. 42, No.1, 225-229.

Lucca, D. A. (1993). Private Communication based on data taken at Los Alomos National Laboratory with Dr. R. Rhorer.

McDonald, A. C. and Lowe, H. (1981). Feedback and Control Systems. Reston VA: Prentice Hall.

McKeown, P. A. (1987). The Role of Precision Engineering in Manufacuring of the Future. Annals of the CIRP, Vol no. , 495 - 501.

Moriwaki, T. et. al. (1989). Development of a Micro Tool Positioning System for Precision Machining. International Conference on Advanced Mechatronics, May 21-24. Tokyo, Japan. 147-152.

Nakayama, K. and Tamura, K. (1968). Size Effect in Metal-Cutting Force. Journal of Engineering for Industry. Feb. 119 - 126.

O,Niell, G. (1990) Piezomotors-Electromechanical Muscle. Transactions of the Society of Automotive Engineers. Vol. 21. No. 2, 510-520.

Patterson, S. R. and Magrab, E. B. (1985). Design and Testing of a Fast Tool Servo for Diamond Turning. Precision Engineering. Vol. 7, No. 3, 123-128.

Ramalingam, S. (1971). A Photoelastic Study of Stress Distribution During Orthogonal Cutting Part 1: Workpiece Stress Distribution. Transactions of the ASME: Journal of Engineering for Industry. May 1971. 527-537.

Shannon, R. S. and Wyant, W. C., Ed. (1987). Applied Optics and Optical Engineering, Vol. X. San Diego, CA: Academic Press

Scire, F. E. and Teague, E. C. (1978). Piezodriven 50 μm Range Rtage with Subnanometer Resolution. Review of Scientific Instruments., Vol.49, No. 12, 1735-1740.

Scmidt-Tiedemann, K. J. (1961). Stress Optical Constants of Germnium. Journal of Applied Physics, Vol. 32, No. 10, 2058-9.

Seo, Y. W. (1993). Energy Dissipation in Orthogonal Ultra-Precision Machining of Ductile Materials. Doctoral Dissertation submitted at Oklahoma State University.

Shaw, M. C. (1984). Metal Cutting Principles. London: Oxford University Press.

Shaw, M. C. (1950). A Quantized Theory of Strain Hardening as Applied to the Cutting of Metals. Journal of Applied Physics, Vol. 21, No 6, 599.

Slocum, Alex (1993). Precision Machine Design. New York, N. Y: Prentice Hall

Spencer, M. (1982). Fundamentals of Light Microscopy. New York, NY: Cambridge University Press.

Taniguchi, N. (1983). Current Status in, and Future Trends of, Ultraprecision Machining and Ultrafine Materials Processing. Annals of the CIRP, Vol. 32 no. 2, 573 - 582.

Trylinski, W. (1971). Fine Mechanisms and Precision Instruments. London: Oxford University Press.

Wahl, R. (1944). Mechanical Spring Design. 2nd. Ed. New York, N. Y. John Wiley and Sons.

Wardzynski, W. (1970). Birefringence Induced by Uniaxial Stress in Cadmium Telluride Single Crystals. Journal of Physics (C): Solid State Physics. Vol. 3, 1251-1263.

APPENDIX A

PIEZOELECTRIC TRANSDUCERS

A.1 Introduction

The sources for this appendix are the catalog literature of Physik Instrumente, Burliegh Instruments, Inc and Morgan Matroc.

Piezoelectric materials develop a surface charge when physical pressure is applied to the surface of the material. Conversely, when a charge is placed on the surface, the material undergoes a change in dimension. There are a number of naturally occurring piezoelectric materials, quartz and Rochelle salts being common examples. For most transducer applications man-made ceramics are used. The original material for application as a translator was PZT (Lead Zirconium Titanate); over time the acronym PZT has come to be commonly applied to all piezoelectric transducers regardless of their material makeup. To manufacture a PZT element a fine powder mixed with polymer powder is sintered at over 1000°C. The resulting polycrystalline solid has poorly defined piezoelectric properties. However by orienting the ceramic in a strong electric field at sufficiently high temperature (a process called 'poling') and keeping the field constant as the ceramic cools the piezoelectric properties of the individual crystals are aligned to the electric field (The direction of the field applied during poling defines the directional properties of the element). In this appendix discussion is limited to those applications where the element is used as a position

translator; that is an applied electric field parallel to the poling direction causes an expansion.

These devices have a number of advantages over traditional transducers. The displacement resolution is limited only by the precision of the electric field control. It is common for integrated stacks with control systems to have resolutions on the order of 1 nm. Since the piezoelectric effect is a bulk effect the response times are limited only by the speed of sound in the ceramic and can be on the order of microseconds. Electrically at low frequencies they are very close to ideal capacitors; consequently they have very low leakage currents and can store charge for very long times (Although the impedance of the PZT driving circuitry affects the leakage also.). In addition the power dissipation can be negligible (but is also strongly dependent on frequency). Ceramic PZT elements can also apply very high forces: up to 5 tons (with a positional accuracy of 100 μm).

There are a number of disadvantages. Despite being sintered with polymers to improve toughness the elements are fragile. There are severe limits on the pulling forces that the elements can apply; in addition whenever the element expands it contract in the other directions. Great care must be exercised in mounting to prevent even the smallest shear loads or misalignment which causes edge stresses as this may destroy the stack. Typically the stacks come in a preloaded enclosure to provide protection. Some typical mounting schemes are shown in Figure A.1. The dual ball-in-cup design was used for the tool holder actuator.

Each element is capable of only small dimensional changes (about 0.001 in/in). Thus to get longer expansions the elements must be stacked which decreases stiffness and increases capacitance (the elements are wired in parallel). In addition the stiffness of a stack is dramatically dependent on the bonding agent used to join the individual elements in the stacks; indeed these bonding materials and techniques are very closely guarded secrets in the translator industry. Finally most

PZT stacks require high voltages (1000 V to 1500 V) to obtain full expansion. In fact the magnitude of the applied voltage is the limiting factor for expansion (Above 1500 V dielectric breakdown occurs destroying the elements.). For high performance this requires expensive high power MOSFET driven amplifiers (besides being dangerous); although recently this drawback has motivated the development of "soft" PZT materials which while more compliant can achieve the same resolution (for smaller loads) at one-tenth the voltage. Perhaps the most notorious drawback is hysteresis which can be up to 15% of the total displacement covered; although some manufacturers use clever construction to limit this to (a claimed) less than 1% or 2% .

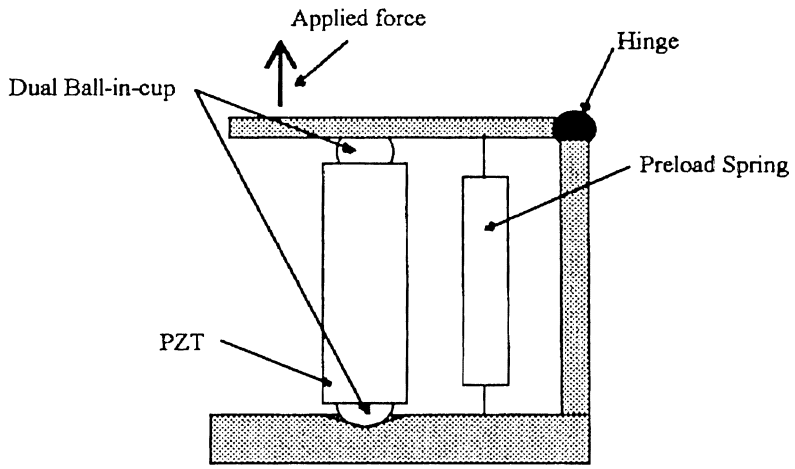


Figure A.1. A Mounting Scheme for a PZ Stack

PZT translators made of stacks of elements are available in wide variety of geometries from stacks and segmented tubes to unimorphs, bimorphs and hybrids. Segmented tubes are used as positioning elements in scanning probe microscopes which resolve surface details on the atomic level. The stack used in the MMI is constructed from thin disks. Some manufacturers, such as Morgan Matroc, supply raw disks or other shapes to be assembled by the buyer.

A.2. Properties of Piezoelectric Translators

The maximum expansion (δ) possible for a PZT depends on the material as expressed by a constant d_{33}, the applied electric field E, the zero field length L_O, the applied force F and the stiffness of the translator, k_O . The constants d_{ij} are components of a tensor relating expansion along an axis to the applied electric field. Standard notation denotes the expansion axis by i=3; j=3 and the contraction axes as i=3; j=1,2. Then δ is given by

$$\delta = d_{33}EL_0 + \frac{F}{k_0}. \tag{A.1}$$

Note that if the PZT is subject to a compressive force that F is negative. If the constant force F is high enough to prevent any expansion ($\delta = 0$) then F is the maximum force specified for the PZT and is called the blocking force. On the other hand if F is path dependent, such as for a spring with a spring constant k_s, then

$$\delta = d_{33}EL_0\left(\frac{k_0}{k_0 + k_s}\right) \tag{A.2}$$

For the stack used in the tool holder the stiffness is greater than 400 N/μm while the flexure stiffness is about 1 N/μm. Consequently the flexures have a negligible effect on the total expansion. Even the maximum constant cutting force of about 7 N has a negligible effect as the blocking force for this PZT is about 20,000 N. The conclusion is that this PZT will still be capable of its no-load expansion.

The dynamic properties of the PZT are usually expressed through its resonant frequency and minimum expansion time. The resonant frequency (f_O) for the preloaded stack is given as a standard specification. However the additional mass of the inner cylinder and the force applied by the flexures will alter the resonant

frequency of the system. The mass of the inner cylinder is approximately 1 kilogram and for an additional mass the resonant frequency f_0' is

$$f_0' = f_0\sqrt{\frac{m}{m+2M}} \quad . \tag{A.3}$$

Note that this applies only when the mass is a constant force. For a spring loaded mass attached to the stack the mass acts as a displacement driven oscillator (Thus low stiffness springs do not necessarily dramatically lower the resonant frequenct of the combination). The resonant frequency for the stack and cylinder if they are rigidly attached is approximately 2 kHz. Since the cylinder and stack are not attached the resonant frequency of the system will be determined by the flexure/inner cylinder combination. This undamped resonant frequency of this combination is on the order of 150 Hz.

All PZ stacks exhibit hysteresis. One factor that contributes to the shape of the hysteresis curve is the phenomenon of drift, found in all PZTs. If a new voltage command is sent to the stack it will continue to expand slowly, after the initial expansion, in the same direction as the original command. The amount of drift is generally small and up to an additive constant is proportional to the natural log of the time (the logarithm is taken as a ratio of time to a standard time unit; usually 100 ms). Drift, like hysteresis, is eliminated by the use of a position control system. A schematic of drift is shown in Figure A.2. In the figure, drift = $\Delta x/\Delta t$.

The capabilities of the amplifier must be known to estimate the performance of the stack. In fact the stack and the amplifier need to be integrated so closely that some instrumentation texts consider them a single unit. The current needed to drive the stack in quasi-static operation is

$$i = \frac{dQ}{dt} = C\frac{dV}{dt} \tag{A.4}$$

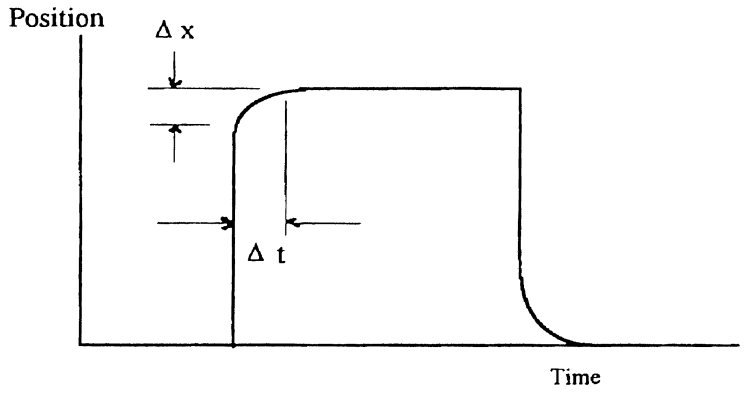


Figure A.2. Drift

The current and power output needed to drive a translator at a given frequency (f) is given by

$$i = 2CUf \tag{A.5}$$

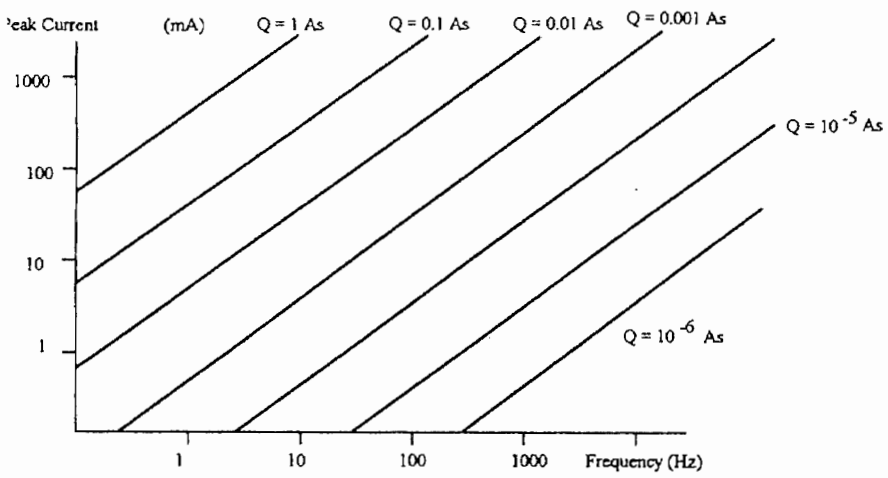
$$P_{out} = CUU_0f \tag{A.6}$$

where C is the nominal capacitance (generally the capacitance is itself a function of applied voltage) and U is the operating voltage (the amplitude of the varying voltage, not a dc component that may be initially applied). In the second equation U_0 is the maximum output voltage of the amplifier. For sinusoidal operation the current requirements are determined by the product of the stack capacitance and the voltage amplitude. A graph for determining the current for sinusoidal operation is shown in Figure A.3 on the following page.

The issue of current needed also brings up the question of power dissipation. Since the stack will be completely enclosed it must generate very little heat. The thermal heat generated by a PI stack of capacitance C operating at a frequency f under an operating voltage U is given by

$$P = fCU^2 \tan \delta. \tag{A.7}$$

The tanδ factor is called the angle of loss and is approximately 0.05.



A.3. Current Requirement for Sinusoidal Operation

APPENDIX B

BASIC OPTICAL CONSIDERATIONS

B.1 The Microscope

The microscope is a collection of a variety of optical instruments. The most important are the objectives, the filters and polarizers and the camera systems.

B.1.1. Microscope Objectives. The heart of the microscope is the objective. The objective the collection of lenses closest to the sample, and has several jobs; it must 1) collect light from the sample and form the primary image to be presented to the eyepiece lens; 2) provide high resolution; 3) provide (usually) the greatest part of the magnification.

Microscope objectives are generally classified according to their optical corrections or according to some special properties they might possess. For all types of microscopy there are 20 or so types of objective. Fortunately for general lab use and for the MMI only a few need to be considered. These are

A. Achromatic
B. Semiapochromatic
C. Apochromatic
D. Metallurgical
E. Differential Interference Contrast (Nomarski or DIC)

Far and away the most important and common are the first three which are classifications based on corrections made to chromatic aberration present in the lens system. It is well known that the refraction of light is wavelength dependent. Consequently if not corrected different colors will focus at different points on the optical axis of the objective. Achromats focus two colors (red and blue) at a single point on the optical axis and then are corrected for one intermediate wavelength. This is relatively inexpensive to do and so achromatic objectives are widely employed. Semiapochromats are essentially achromats which have additional elements made of fluorite or fluorite glasses which provide more correction of the residual color effects of the achromats. In addition these fluorite based glasses greatly enhance light gathering ability and contrast. Sometimes these types of objectives are called fluorites or neofluorites. Apochromats are the most highly corrected of all. Red, green and blue are all brought to focus at a single point and additional elements correct for two intermediate colors. Since an apochromat may contain twice as many optical elements as an achromat this type of objective is far more expensive. The extent of correction is an important issue for this project. The more highly corrected the objective for the visible part of the spectrum the less likely it is to perform well at bandwidths outside the visible spectrum.

A special note about flatness of image. At one time it was standard to include objectives further corrected to provide a flat field of view as its own category of objective. However modern objectives are almost always corrected for flatness of image, although manufacturers still draw attention to the fact by calling their objectives 'Plan-achromats', etc.

Standard microscope objectives are designed to work in air with specimens using a cover glass on a glass slide, reflecting the widespread use of the microscope in the biological sciences. This requires that the optical effects of the cover glass (refraction) be incorporated into the objective design. The

metallurgical objective is one corrected for use without a cover glass and usually has additional antireflection coatings since it is commonly used to view opaque objects in incident, or reflected, light (These are intended for fairly general use although the term metallurgical falsely suggests some particular application to metals). Metallurgical objectives are available in the different categorize of color correction.

Differential interference contrast (DIC) is a technique for obtaining qualitative information about variations in height in a microscope image. It was traditionally developed by G. Nomarski as a technique for examining the surface characteristics of metals. DIC objectives are metallurgical objectives which are strain free because the technique requires the use of polarized light. Indeed when the possibility exists that polarized light is to be used the objective (regardless of type) should be strain free to avoid distortion of the image color and quality by birefringence present in the objective elements.

A final class of objective to be considered is the super long working distance objective. These objectives use mirror systems to greatly increase the light gathering ability of the objective. The result is the formation of an image with the objective further away from the sample. For a 100X objective with a working distance of 300 μm the SLWD version may have a working distance of 4.5 mm.

B.1.2 Objective performance. The most obvious specification of an objective is its magnification, which is defined as the ratio of the image size to the object size. Typical ranges are from 2X to 200X. Useful magnification is the maximum magnification that allows details to be resolved by the observer. Empty magnification is any magnification beyond useful magnification and can be used to make viewing more comfortable. The total magnification of the image is obtained by multiplying the objective magnification by the eyepiece magnification.

Working distance refers to the distance away from the first objective lens to the sample surface when the image is focused. Generally the higher the magnification the smaller the working distance. For high magnification objectives the working distance can be as small as 200 μm.

Another important specification is optical tube length. This is the length from the back focal plane of the objective to the image plane. In order for an objective to be compatible with the microscope body it must have the correct tube length. Different manufacturers tend to use different tube lengths making it generally impossible to interchange different manufacturers objectives on the same body. For the Nikon MM-11U used for the MMI a 210 mm tube length is used.

One of the most important measures of microscope performance is resolving power, most of which is done by the objective. Resolving power (d) is the minimum distance between two objects that can be detected. In general it is given by

$$d = \frac{\alpha\lambda}{n\sin\theta} \tag{B.1}$$

where α is a constant that depends on the specific (arbitrary) criterion used to determine define resolvability and n is the index of refraction of the medium between the sample and the objective. The most common criterion is the Raleigh criterion for which $\alpha = 0.61$. The term $n\sin\theta$ is called the numerical aperture (N) and has a geometric interpretation as shown in Figure B.1.

If the objective is to be used in air (for which n = 1) then $N = \sin\theta$ and the numerical aperture is the sine of the half-angle defining the light cone of the light entering the objective (called the half-angle of acceptance). This is a standard number used to compare objectives; it has a maximum value of about 0.95. The importance of N cannot be overstressed. It determines the amount of illumination needed (higher N, higher illumination), and the resolving power of the objective.

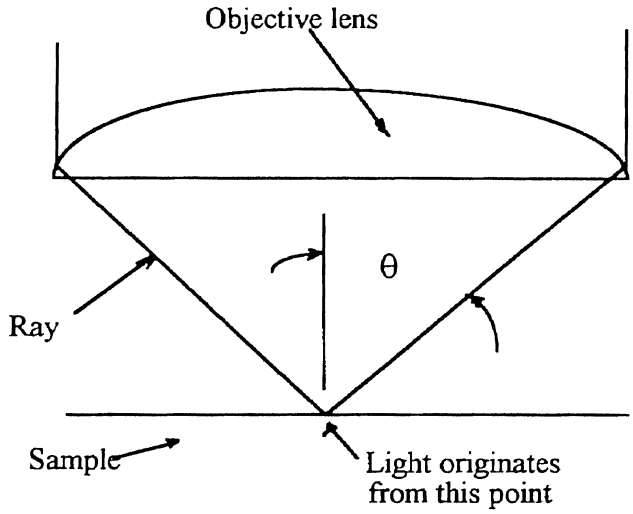


Figure B.1. Light Cone Defining Numerical Aperture.

This development assumes the light is reflected from the sample. For the case of transmitted light the transmission path can affect the resolving power also. Microscopes generally have an optical device called a condenser in the transmitted light path. The condenser is itself a carefully designed and corrected collection of optics. In this case, if the numerical aperture of the condenser is less than that of the objective, the resolving power is

$$d = \frac{\alpha\lambda}{N_a + N_c} . \qquad (B.2)$$

It can be seen that the condenser can aid resolving power. For practical calculations standard practice is followed here and the constant α is chosen to be unity.

Depth of focus is the depth through which the objective can be moved without any noticeable change in the image. It is strongly dependent on numerical aperture or magnification with high numerical aperture (and high magnification) objectives having small depths of focus (perhaps only tens of microns). The depth of field will also be different for photographic and naked-eye observations.

The materials used to construct the objectives are of concern for this project. Modern objectives are corrected for use in the visible part of the spectrum and so the materials transmit well in the visible. There performance in the infrared is an unknown quantity and is material dependent. Another concern is overall bandwidth. If objectives are found that work well in the visible and the infrared it will certainly be the case that they will not work well in the ultra-violet. Lens materials are of course closely guarded proprietary secrets. As with the uncertainty with regard to corrections and IR performance the best way to evaluate this factor is by direct observation and comparison. Before the Nikon MM-11U was selected a series of on-sight demonstrations was undertaken to evaluate the Nikon and microscopes from other manufacturers to evaluate their infra-red performance.

B.2. Filters

The clearest images and the greatest information is obtained from stress birefringence when monochromatic light is used. The illumination sources of microscopes generate light over the near ultra-violet to visible to the infrared end of the spectrum. Consequently filters will be needed to narrow the wavelength of light incident on the sample. Filters are available in a bewildering range of types and performance. There are two types which will be of concern here: Neutral density filters and narrow bandpass filters.

B.2.1. Neutral density filters. The spectral density (intensity vs. wavelength) of the illumination source is a function of the transformer voltage supplied to it. For example the Nikon units will approximate the spectral density of the Sun at approximately 9 V. The purpose of a neutral density filter is to reduce the intensity of the output by the same fraction for each wavelength. This is most

important for photographic considerations but it can also help improve image quality as it allows the reduction of light intensity (which often improves contrast) without simultaneously changing the wavelengths being returned to the eyepiece. Neutral density filters for microscopes are usually made by evaporating a thin film Inconel onto a substrate or by some combination of thin sheet polarizing elements.

B.2.2. Narrow bandpass filters. A narrow bandpass filter transmits only a narrow range of wavelengths. There are many ways to construct these filters but generally microscopes (and in particular the MMI) need small filters that fit into small spaces. The only candidate here is the interference filter. Interference is the combining of two waves to generate a new wave by superposition. There are two kinds of interference: wavefront splitting as occurs with thin slits and amplitude splitting as occurs with thin films.

Interference filters operate on the principle of thin-film interference (amplitude splitting). When light is incident on a surface which is coated with a thin dielectric film it is partially transmitted and partially reflected. The transmitted portion is then reflected off the substrate and recombines with the original reflected ray. Since the index of refraction for the coating is greater than that of air the first reflected ray undergoes a phase shift of 180° (if the incidence is normal to the surface). By combining a number of layers to repeat this effect over and over a filter can be made which transmits a number of well defined wavelengths (due to constructive interference) and no others (which are canceled by destructive interference). The undesired spikes can then be eliminated (called blocking by filter makers), leaving only one spike. Some of the important characteristics are shown in Figure B.2.

As with any device increases in performance in one area have to be paid for with decreases in another area. It is possible to obtain off the shelf interference

filters with full width at half maximum transmittance (FWHM) of 1 nm. This is a very narrow bandwidth and provides an excellent approximation to monochromatic light. However due to the large number of layers and the blocking coatings the peak transmittance is only about 40% to 50% (compared to unfiltered light). In fact these filters are generally used with intense sources such as lasers. To obtain high transmission the bandwidth will be wider. Transmission of up to 90% can be obtained with bandwidths as narrow as 10 nm off the shelf. When selecting filters blocking should be kept to the minimum to avoid unnecessarily degrading performance.

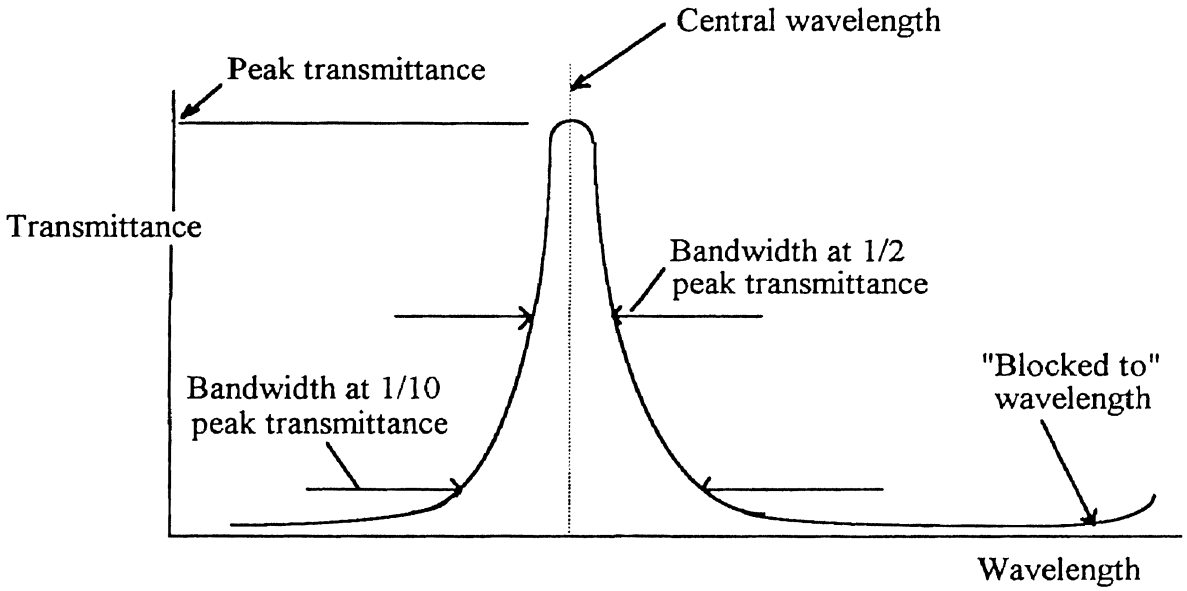


Figure B.2. Spectral Characteristics of an Interference Filter.

It should be stressed that these coatings are indeed very thin and fragile. In order to generate interference there has to be a certain amount of coherence in the reflected and transmitted light (there has to be some nonrandom phase relation). The coherence length (which is the length over which this phase relation can be maintained) for white light is about 1μm. Consequently the coatings on these filters must be thinner than this.

B.3. Polarizers

Light propagates as a transverse electromagnetic wave. Typically the direction of the electric fields is random. Polarization is the selecting of one component of the electromagnetic field while obstructing all others. The direction of polarization is referenced parallel to the transmitted electric field and is called the transmission axis. There are a number of ways to specify the performance of polarizers.

B.3.1. Specifications. As with filters the transmittance of a polarizer is an important consideration. However the situation is more complex with polarizers. Consider a linearly polarized beam incident on a polarizer. If the transmission axis is parallel to the incident polarized beam the ratio of incident to transmitted intensities is called the major principle transmittance, k_1. If the transmission axis is perpendicular to the electric field this ratio is called the minor principle transmittance, k_2. The extinction ratio is defined as k_2/k_1. This extinction ratio is the primary specification used to evaluate and compare polarizers. The lower the extinction the better the polarizer. This property is also very dependent on wavelength. In fact it is good policy not to purchase a polarizer until a graph of extinction vs. wavelength has been seen. When unpolarized light is incident on two polarizers the maximum transmittance (both transmission axes parallel) is $\frac{1}{2}(k_1^2 + k_2^2)$. The minimum (transmission axes crossed) is $k_1 k_2$.

B.3.2. Types of polarizers. Dichroism is a standard way to polarize light. The most common are "Polaroids" made from stretched polyvinyl alcohol treated with iodine. An easily made error is to assume the electric field is polarized parallel to

the stretched polymer. However the iodine provides free electrons which are excited to motion and thereby absorb the energy of the field in this direction. As a result the field is polarized *perpendicular* to the stretch polymer. These polarizers can have good extinction characteristics (as low as 0.001) in the ultraviolet, visible and near infrared.

Polarization in the infrared is also accomplished by wire grid polarizers. A series of fine parallel metal wires can act as a polarizer by the same mechanism as used in dichroic polarizers. Etched gold and silver films are commonly used for this purpose. Since the spacing of the wires must be less than the wavelength it is not usual to see these used for wavelengths shorter than the infrared. Like dichroic sheet polarizers these can be made small and are suitable for use in a microscope.

The last category is the birefringence polarizer. Due to optical anisotropy, when light passes through a birefringent crystal the wavefront splits into two separate orthogonally polarized beams. The angle between the beams depends on the crystal orientation. If the crystal is cut and rejoined using a glue with an index of refraction less than the crystal one of the beams can be rejected (by total internal reflection) while the other passes through. These are not as compact as other polarizers so mounting has to be carefully thought out, and they are far more expensive. However their extinction ratios are on the order of 100 times better than dichroic sheet polarizers over very wide bandwidths (from 400 to 2500 nm).

For further information, consult these references (found in the bibliography) used for this Appendix.

Bradbury, 1984
Hartshorne, 1970
Kingslake, 1965a
Kingslake, 1965b
Shannon and Wyant, 1987
Spencer, 1982

APPENDIX C

LabVIEW FOR WINDOWS

C.1 Introduction

Computer controlled systems are a very important tool due to their speed and ability to gather and combine data from a number of different sources. The computer programs are generally written in high level languages such as FORTRAN or C which the computer translates via a compiler into machine language. Many of these high level languages are text based and mimic many of the structures of English in an attempt to make them easier to learn and use. The language used for this project is a graphical user interface or visual interface (VI). The program is not written in lines of code; rather a picture is drawn representing the logical relationships between various programming structures and commands. There are a number of these programs commercially available; the one used by this lab is LabVIEW™ for Windows developed by National Instruments. A compiled LabVIEW visual instrument program (VI) runs approximately as fast as a compiled C program.

C.2 Structure of LabVIEW

Programs written with visual interfaces are called virtual instruments because they emulate a real instrument. A virtual instrument in the LabVIEW system consists of two parts: the front panel and the block diagram.

The front panel contains buttons, knobs and readouts which look like the panel of an instrument. The individual control and display elements are chosen by the programmer depending the set of instructions he/she wishes the computer to execute. There are virtually an infinite number of front panel configurations for a given program. Knobs and buttons are controlled by the cursor and mouse keys.

The block diagram is the actual source code for the computer. It contains all the instructions and logical relations that are compiled to object code. Objects in the block diagram fall into several different categories. Most are either nodes, terminals or wires. Nodes are the elements that execute program instructions and serve the same purpose as statements or subroutines in standard languages. An example of a node for adding two numbers is shown in Figure C.1 LabVIEW has very extensive libraries of nodes for simple functions up to Fast Fourier Analysis, filtering, signal generation and statistical signal processing.

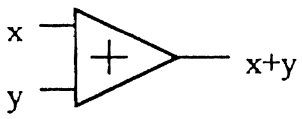


Figure C.1. Node for Addition.

Terminals are the block diagram elements that correspond to a control or display on the front panel. With them information such as data type or precision is set, or the results of a calculation or measurement are displayed. Nodes and terminals are connected by wires which "carry" the data from one part of the program to another. The color or style of the wire indicates the type of data that it is carrying.

The example shown in C.1 is an arithmetic function (of which the LabVIEW library has 19). In addition there are logical functions, trigonometric functions,

logarithmic and exponential functions or functions for operating on data such as comparison, conversion or array/cluster functions.

An important kind of node is the structure, which acts as a control structure in standard languages. Programs are written inside these nodes which then provide an organization for data flow. There are four structures in LabVIEW: the for loop, the while loop, the case structure and the sequence structure. The for loop executes a program n times (n is set by the programmer). The while loop executes a program until a Boolean variable has a false value. The case structure is composed of a layer of two diagrams (like a deck of cards). Each layer is labeled true or false and different subprograms are placed in each structure. The program that runs at a particular time is determined by the Boolean variable and can be controlled by the front panel. Like the case structure the sequence structure is layered; but with as many layers as are needed. The programs are executed in the sequence they appear in the layers and the data from one layer is passed on to the next if necessary.

C.3 Examples of LabVIEW Visual Interfaces

Probably the best way to get a feel for how the system works is to give some examples. Below is a front panel (Figure C.2) and block diagram (Figure C.3) for a simple counting program. The broad border around the block diagram is a while loop. The small square in the lower right corner is the terminal which controls the loop; if the value wired to it is true the loop runs. In the figure there are two TF terminals wired to the control terminal-one for the front panel control (a toggle switch) and the other for the front panel indicator. The arrows on the sides of the loop are called shift registers and the data entering on the right shift register is transferred to the shift register on the left. The sum node takes the value from the

previous cycle (stored by the left shift register) and adds one and then sends that value to the right shift register (to be picked up for the next cycle by the left shift register) and to the display (which is a dial with a digital indicator below for this VI). Since program is placed in a while loop it executes as long as the switch of the front panel is 'on' (true). Thus a continuous counting loop is established. The timing of the loop is controlled by the timing node that looks like a metronome. The interval is wired to it in units of milliseconds; one count per second is thus 1000 ms.

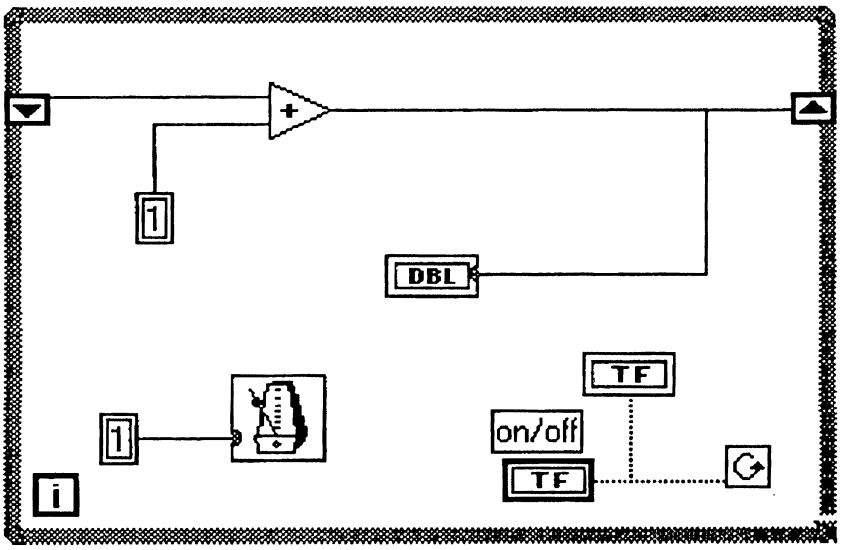


Figure C.2. Counting Program Block Diagram

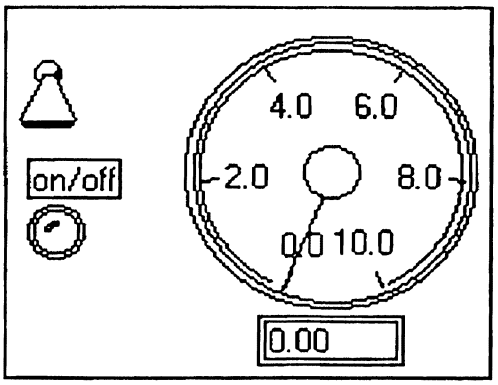


Figure C.3. Counting Program Front Panel.

An important use of VI's in this project is the waveform generation and signal processing VI. There are a number of waveforms that can be generated: sine and sinc, square, ramp impulse, gaussian noise just to name a few. A typical waveform generated is the sine wave; the VI is shown in Figure C.4.

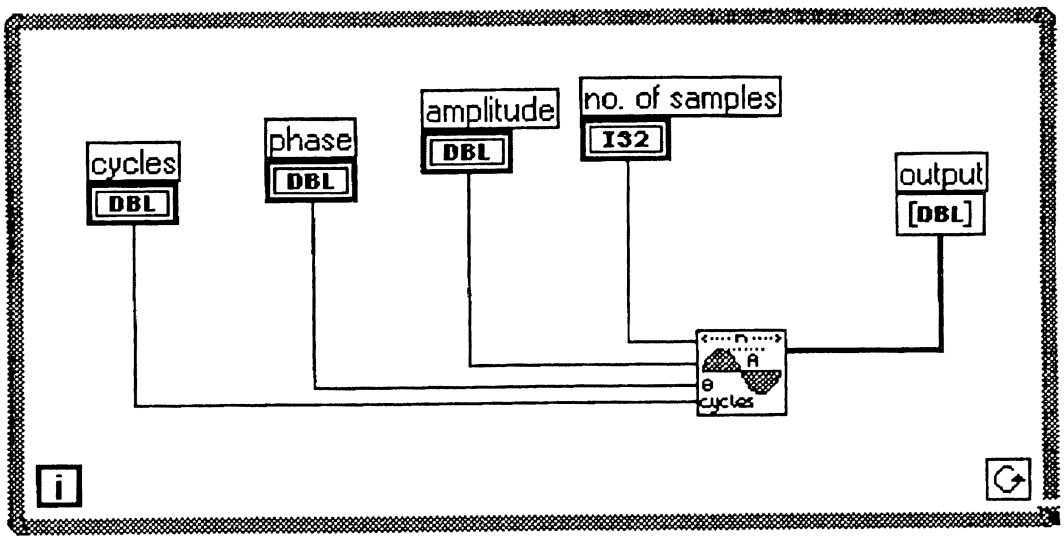


Figure C.4. Sine Generation VI

The number of samples is essentially the number of points in the total wave form. Amplitude determines the output voltage; by choosing the appropriate scales on the board output it can be up to ±10 V. Phase determines the delay and cycles is the number of cycles in the waveform and can be given in fractional numbers. Note that the number of samples covers the entire waveform; for example if 1000 samples is chosen for 4 cycles this means each cycle will contain 250 samples (points). The relationship between the number of samples and the number of cycles must be closely watched if aliasing is to be avoided. These nodes can be wired to output VIs so that the signal can be applied to preamplifiers. Note: Although wave generation symbols and other symbols which perform complex data operations conform to the definition of a node, due to their complexity they are sometimes called sub-VI's.

An example of an output VI block diagram is shown in Figure C.5. In this case the waveform generated by the computer is a square pulse. The controls corresponding to the appropriate nodes are shown in Figure C.6.

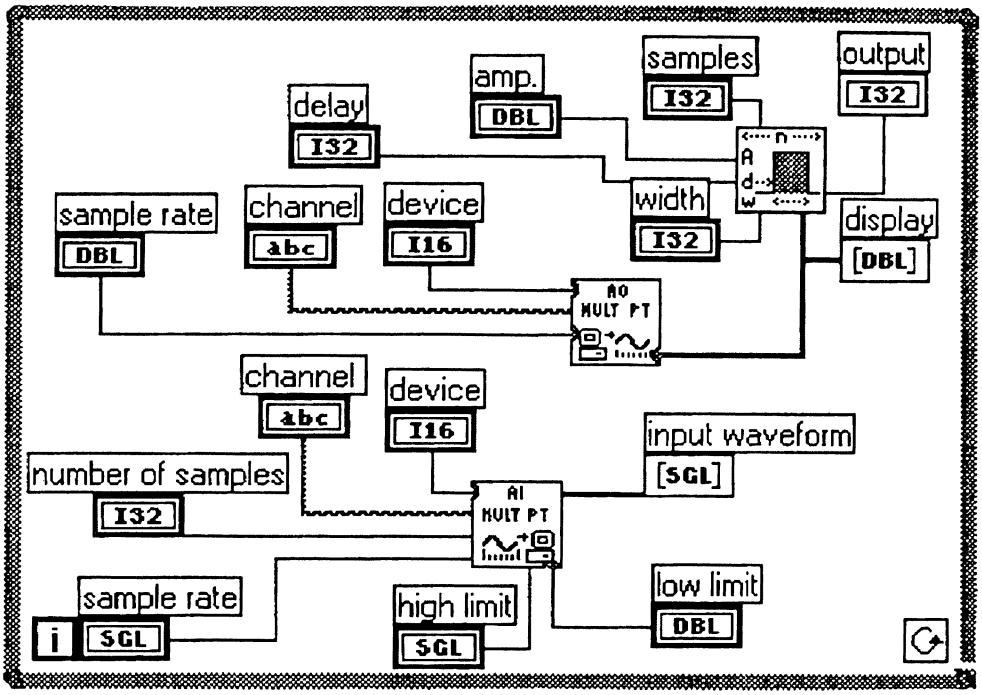


Figure C.5. Waveform Generation/ Reading
VI Block Diagram.

The samples, amplitude and delay parameters have the same meaning as for the sine generation VI; the width selects the number of samples used for the actual square wave (so width must be less than the number of samples). All these controls are grouped together on the front panel under "configure pulse pattern". Compare the labels on the block diagram with the labels on the front panel.

The generated waveform must be fed to the DAQ board to become output. The AO (analog output) node handles this function. The device (the slot on the mother board that has the DAQ board) is selected as is the channel to be used on the board. The rate that the waveform is to be output is also selected. These controls are grouped together under "configure output" on the front panel.

Finally this program is set to read a waveform (in this case a response to the square wave signal output). The same parameters for the board have to be selected in addition to the total number of samples to read and the high and low voltage limits. This node is also wired to display the waveform acquired.

The reader will note the different data types shown in Figure C.6. The abc terminal denotes a string (also note it has a distinctive wire). No other data type may be wired to this terminal and the node. Incorrect data type is a common error in this type of programming; in fact LabVIEW has a special debugging program which can, among other things, detect this error. DBL and SGL denote single and double precision floating point numbers while I8 and I16 indicate byte and word integers, respectively.

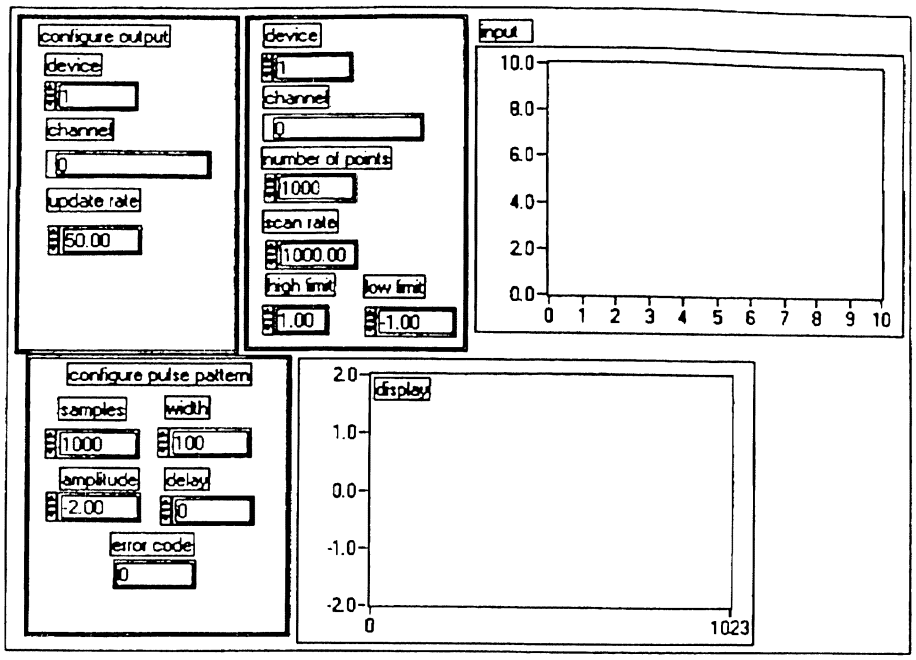


Figure C.6. Waveform Generation/Reading VI front Panel.

All information in this Appendix can be found in the LabVIEW user manuals.

APPENDIX D

WAVEFORM CATALOG

D.1 Step Function

This appendix is a more complete collection of waveform plots. The step function plots are shown first followed by the sinusoidal plots. Not every graph has been presented; however the file name of every plot is listed at the end of this appendix. For further information see Dr. D. A. Lucca.

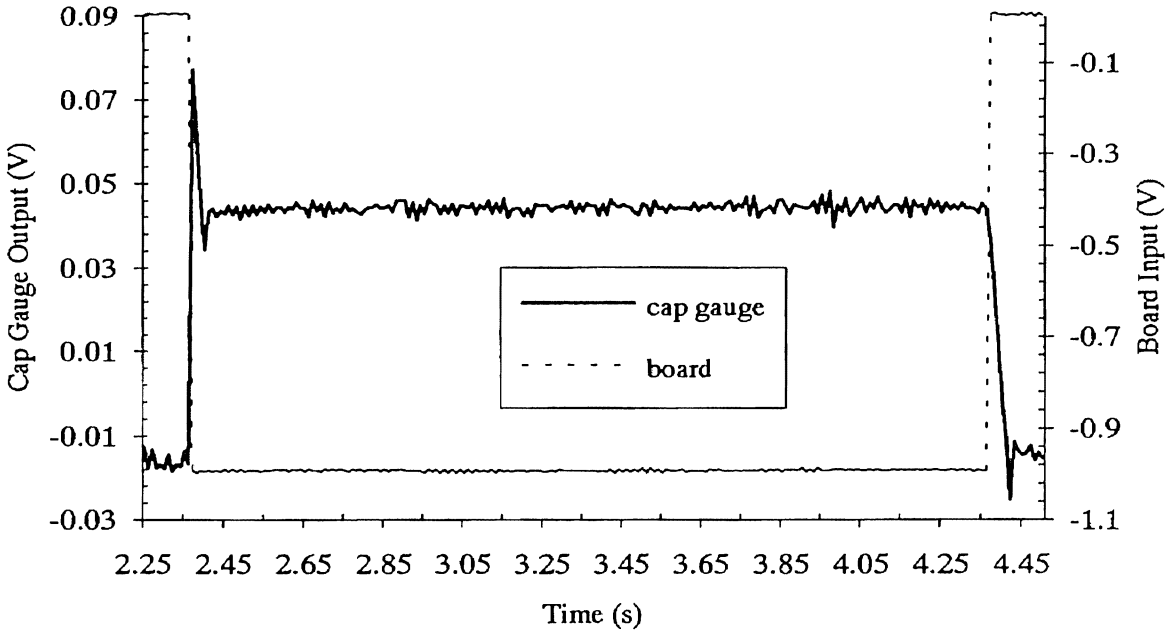


Figure D.1. 1 V Step Response

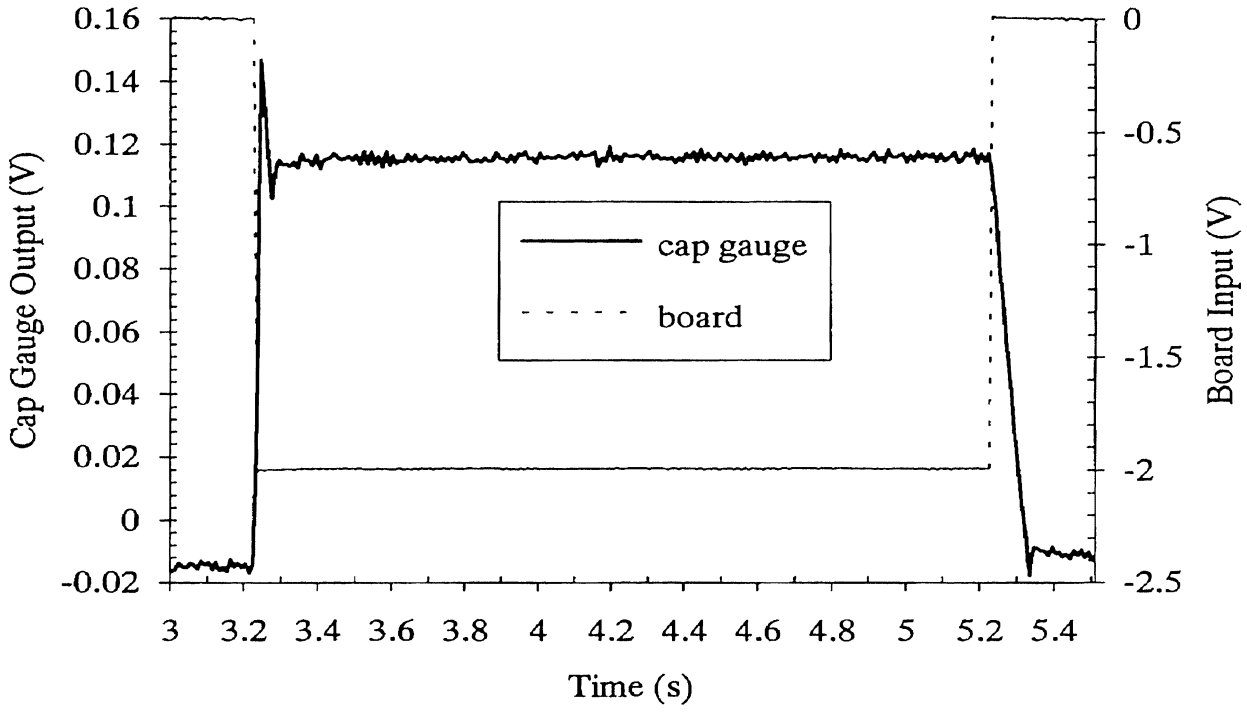


Figure D.2. 2 V Step Response

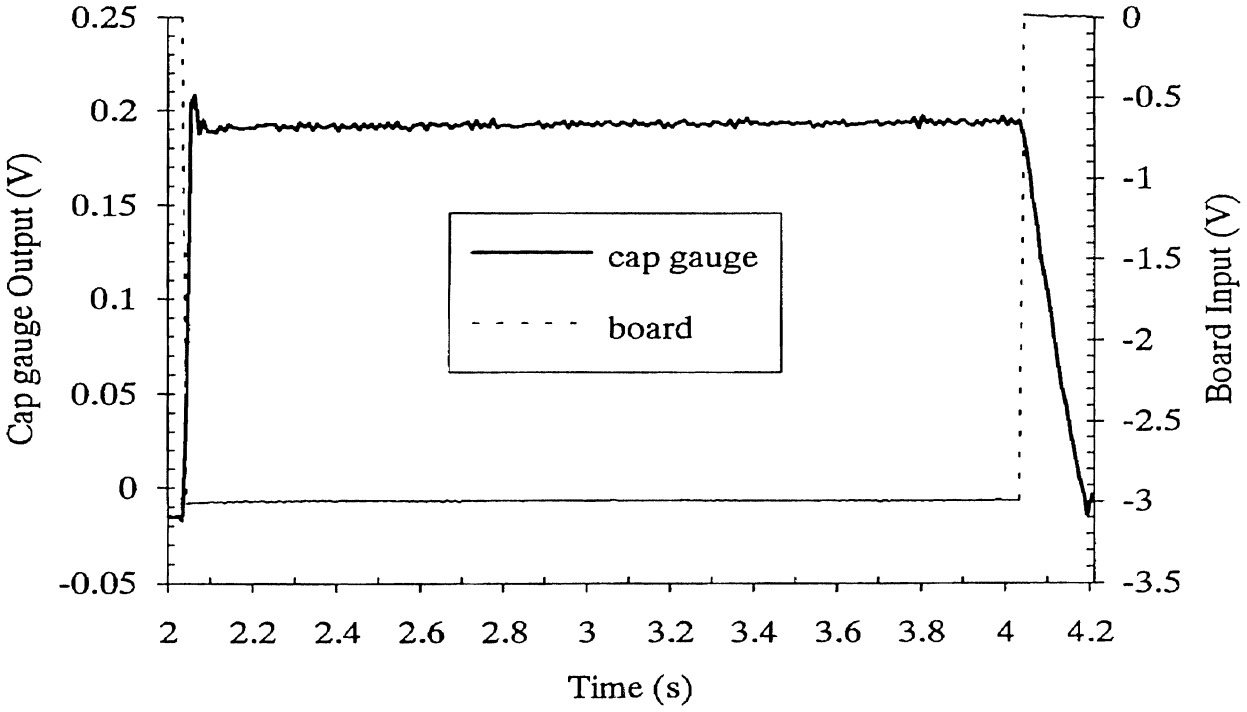


Figure D.3. 3 V Step Response

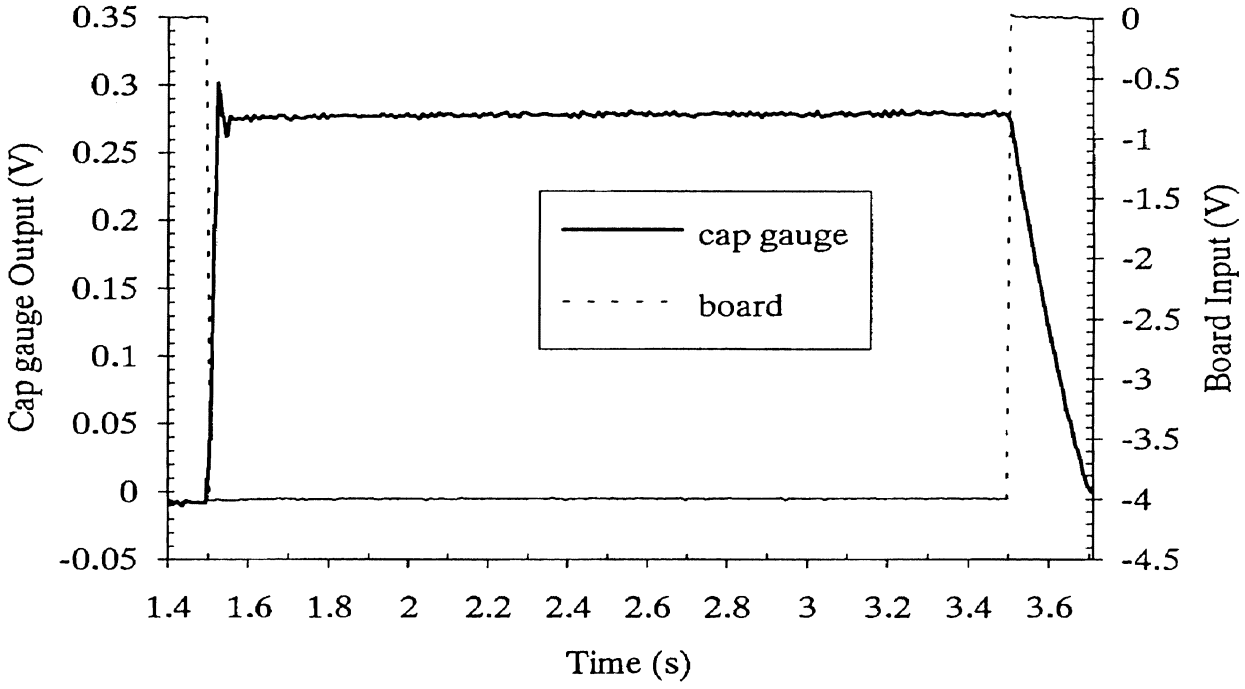


Figure D.4. 4 V Step Response

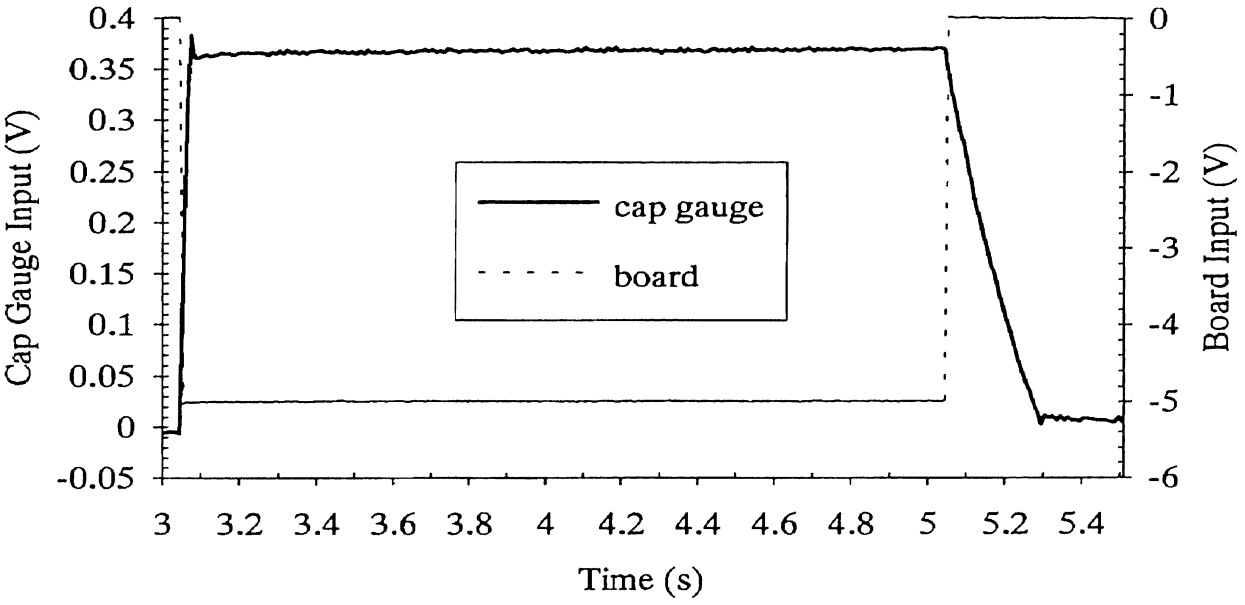


Figure D.5. 5 V Step Response

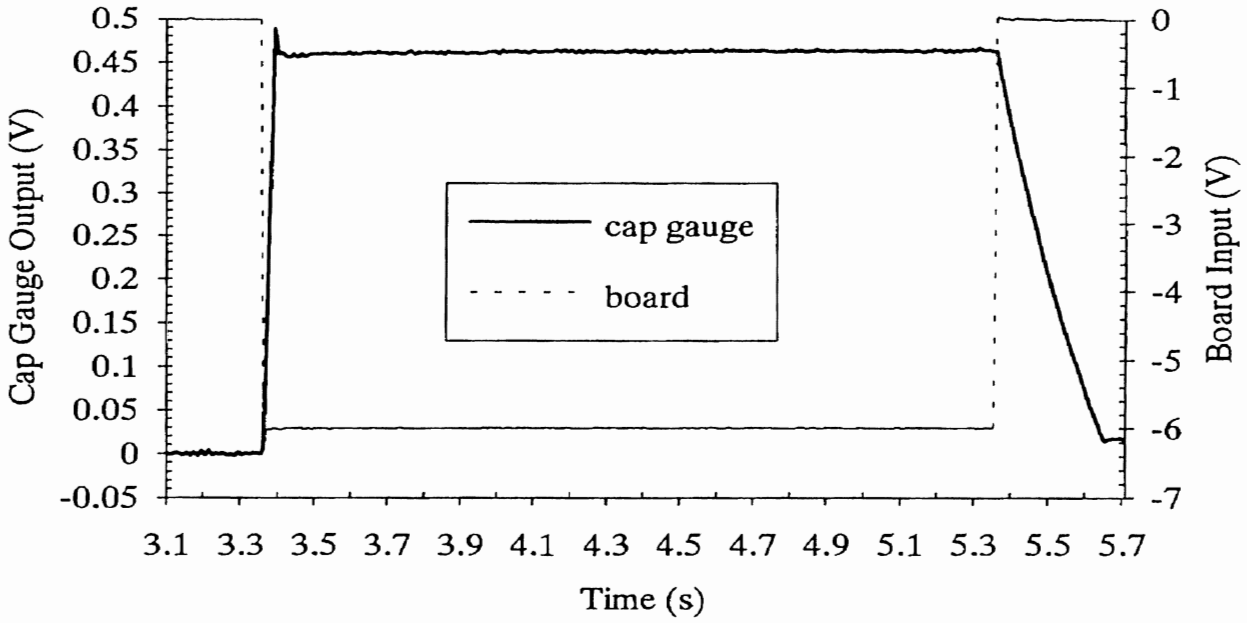


Figure D.6. 6 V Step Response

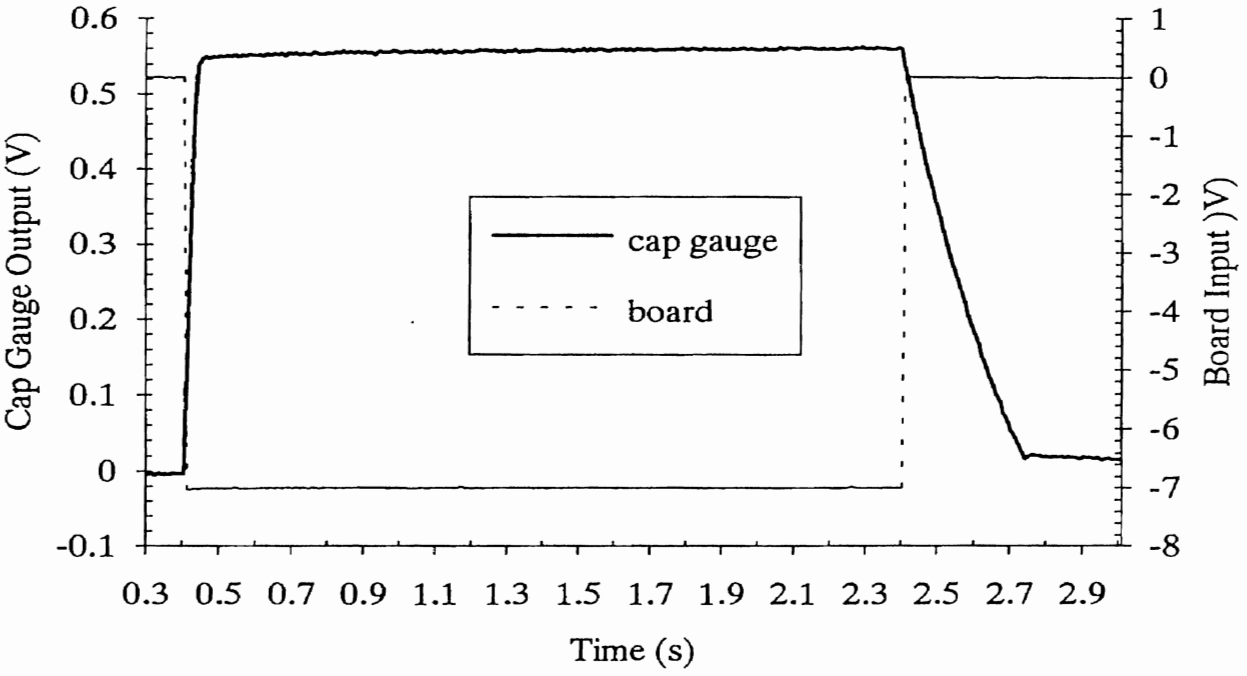


Figure D.7. 7 V Step Response

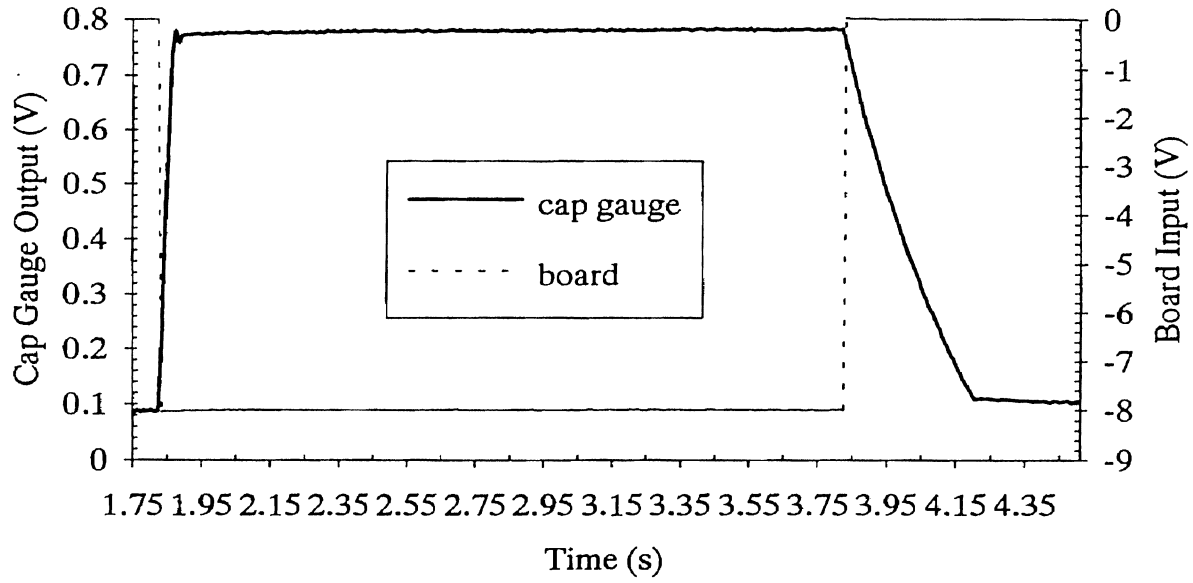


Figure D.8. 8 V Step Response

D.2. Sinusoidal Plots

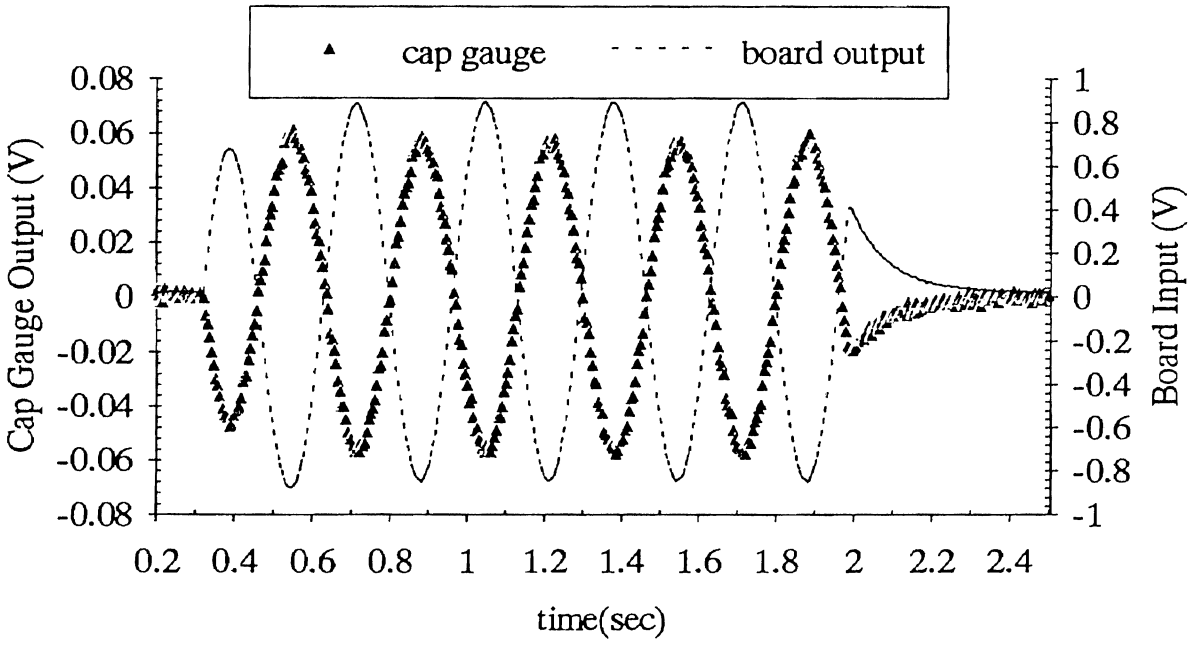


Figure D.9. 3 Hz Response

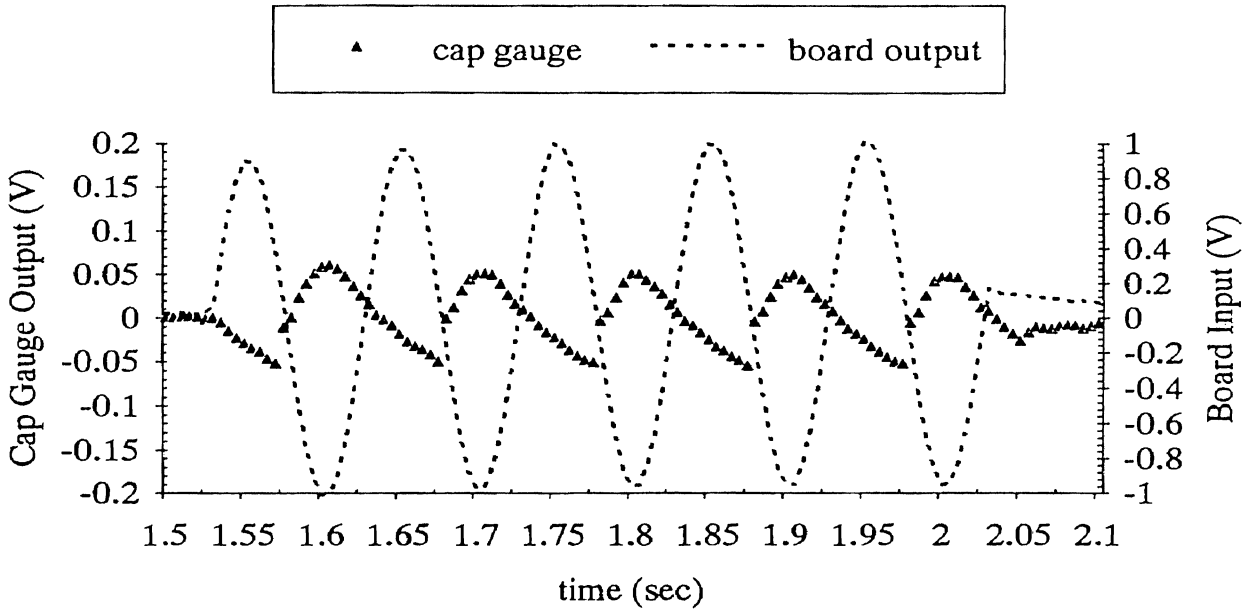


Figure D.9. 10 Hz Response

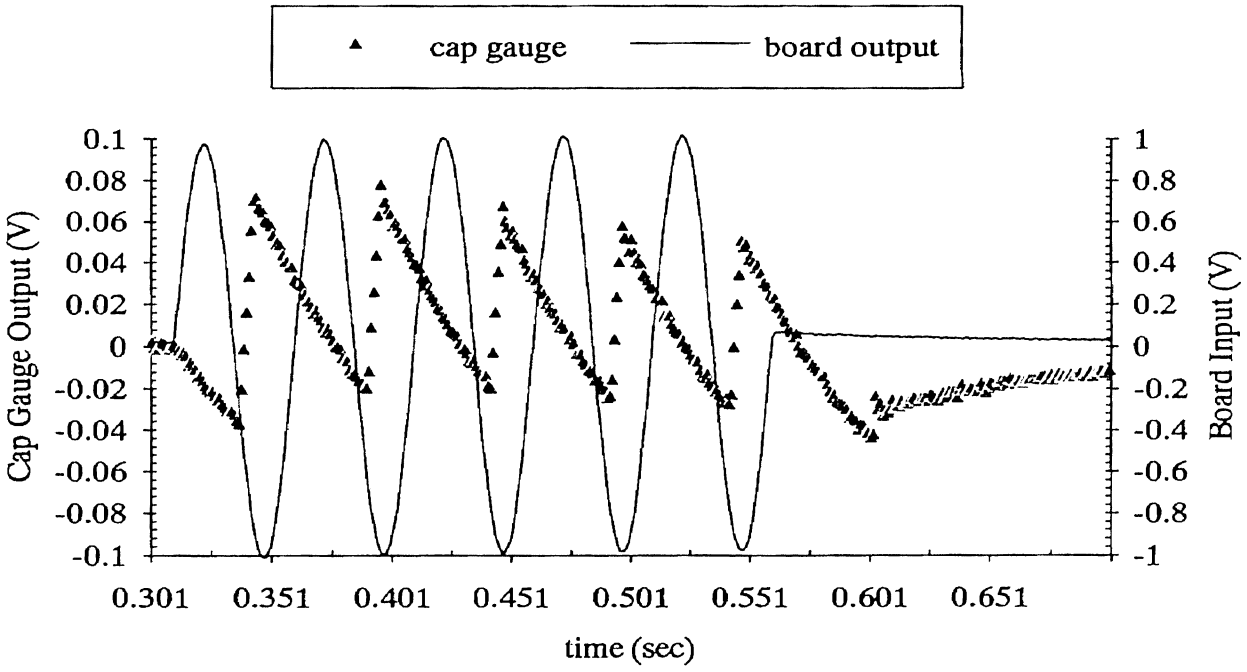


Figure D. 10. 20 Hz Response

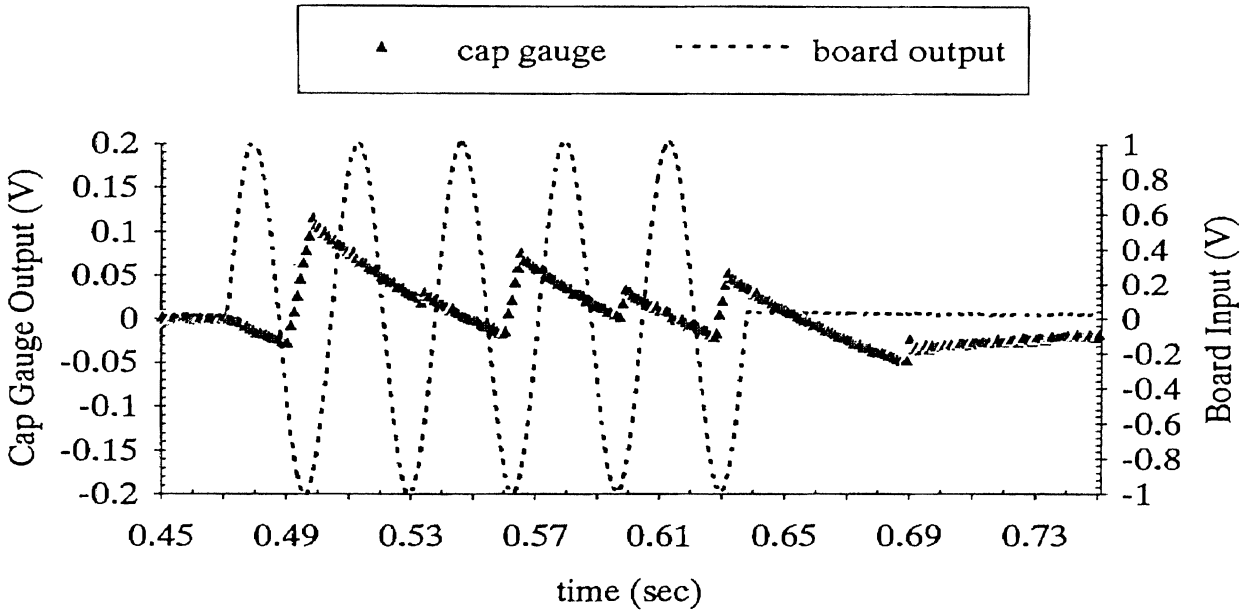


Figure D.11. 30 Hz Response

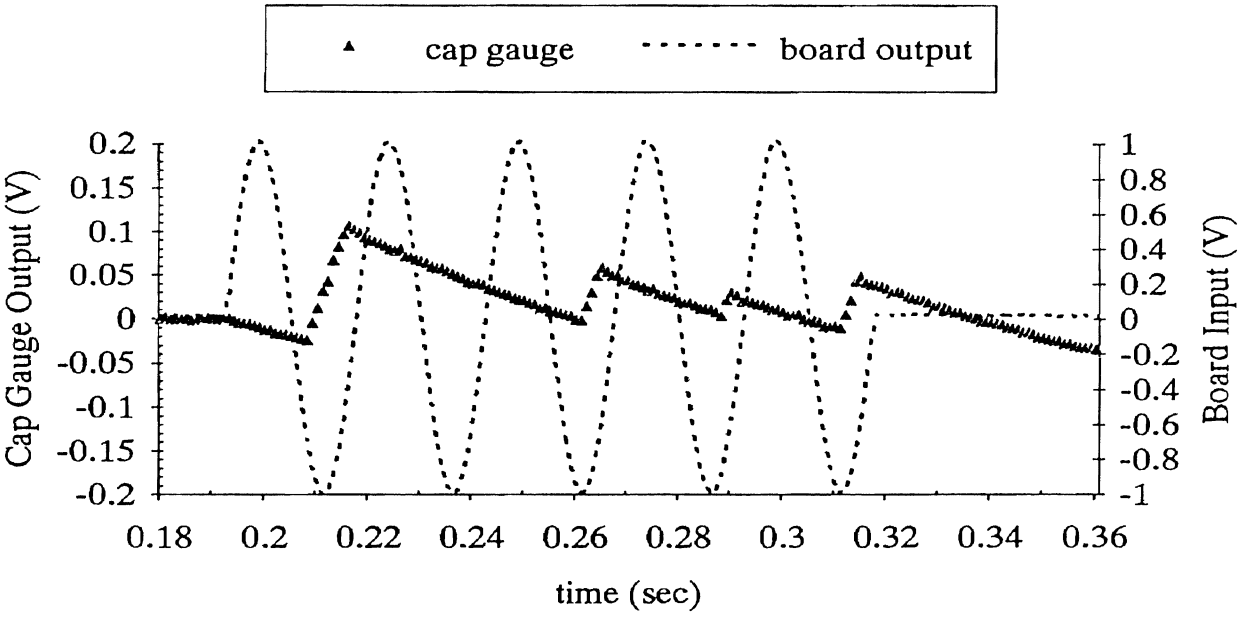


Figure D.12. 40 Hz Response

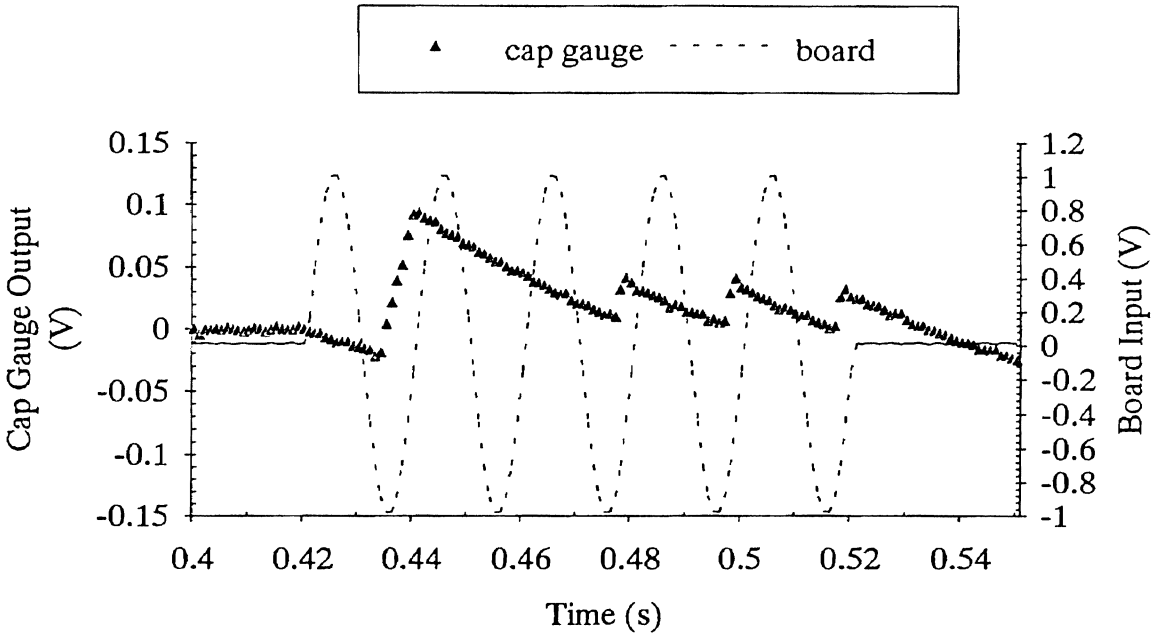


Figure D. 13. 50 Hz Response

D.3. List of Waveform Files

Following is an inventory of computer files that contain waveform data stored as ASCII floating point files. The voltage and time data are normalized and arranged into two columns with the voltage data in the first column. The first row is a header and the second contains the axis labels, select delimiter and CR-LF terminator. The actual data starts in the third row. Originally this data was imported into Microsoft Excel for Windows and the delimiter selected was a comma.

Also included in the inventory is the original data file as it was recorded by the Nicolet Pro 10 Oscilloscope. The Nicolet waveform utilities disk will perform a number of useful conversions on this raw data, including converting to other ASCII formats. These files have the same name as the ASCII flies but with the extension *.wft. Finally there are also a number of Fast Fourier Transform plots for the step inputs; however the conversion program needed for this data is

wavconv.exe which is not on the copy of utilities disk (Version 1.4) in the possession of the lab . The raw data for the FFT plots have the file extension *.wff.

Step Input Value	**Board Input Filename**	**Response Filename**
1 Volt	1vbop.flt	1vcapgag.flt
2 V	2vboard.flt	2vcapgag.flt
3 V	3vboard.flt	3vcapgag.flt
4 V	4vboard.flt	4vcapgag.flt
5 V	5vboard.flt	5vcapgag.flt
6 V	6vboard.flt	6vcapgag.flt
7 V	7vboard.flt	7vcapgag.flt
8 V	8vboard.flt	8vcapgag.flt
Sinusoidal Input	**Board Input Filename**	**Response Filename**
3 Hz	wave0039.flt	wave0040.flt
4 Hz	wave0037.flt	wave0038.flt
5 Hz	wave0033.flt	wave0034.flt
10 Hz	wave0035.flt	wave0036.flt
15 Hz	wave0041.flt	wave0042.flt
20 Hz	wave0043.flt	wave0044.flt
25 Hz	wave0045.flt	wave0046.flt
30 Hz	wave0047.flt	wave0048.flt
35 Hz	wave0049.flt	wave0050.flt
40 Hz	wave0051.flt	wave0052.flt
45 Hz	wave0053.flt	wave0054.flt
50 Hz	wave0055.flt	wave0056.flt

Table D.1. List of Waveform Files

APPENDIX E

DESIGN NOTES

E.1 Tool Holder Notes

Regardless of what measures are taken to improve the stiffness and frequency response of the tool holder assembly a preload mechanism may need to be mounted to improve the mechanical aspect of the dynamic performance. This will increase the resonant frequency of the tool holder assembly and provide the preload necessary to keep the stack/cylinder system in contact with its support. One possible configuration using a Belleville washer is shown in Figure E.1; this arrangement has the advantage of requiring only a new outer ring clamp for the outer cylinder. There is currently an excess clearance of several millimeters on the tool holder side of the objective which gives ample room for a slightly thicker circular clamp. Belleville springs are widely available in a variety of sizes and stiffnesses. The strength of the preload should be carefully considered because the force it applies will diminish the maximum expansion of the stack. The extent of this influence can be calculated or determined from a load stress line. A graphical representation of the effect of preload on maximum expansion is shown in Figure E.2.

As with the flexures the stiffness of a Belleville spring depends on the thickness t, the initial cone height h, and the aspect ratio ξ. In general the load is not a linear function of the deflection [Wahl, 1944] and therefore the stiffness is

not linear; however if the deflection is much less than the thickness (as is the case here) then the load is directly proportional to the deflection and the stiffness k is given by

$$k = \frac{E}{1-\nu^2} C_1 C_2 \frac{t^3}{r_0^2} \tag{E.1}$$

where

$$C_1 = \left(\frac{h}{t}\right)^2 + 1 \tag{E.2}$$

and

$$C_2 = \pi\left(\frac{1+\xi}{1-\xi} + \frac{2}{\ln \xi}\right)\left(\frac{1}{1-\xi}\right)^2 \tag{E.3}$$

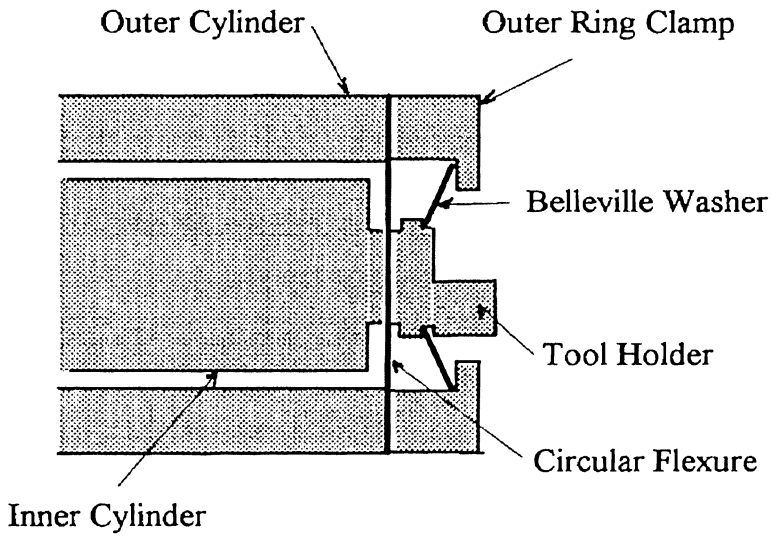


Figure E.1. A Possible Preload Mechanism

The aspect ratio ξ is fixed at about 0.6 and although it could be made somewhat larger it could not be made large enough to significantly affect the stiffness. The stiffness of a Belleville spring as a function of stiffness for two different cone heights is shown in Figure E.4.

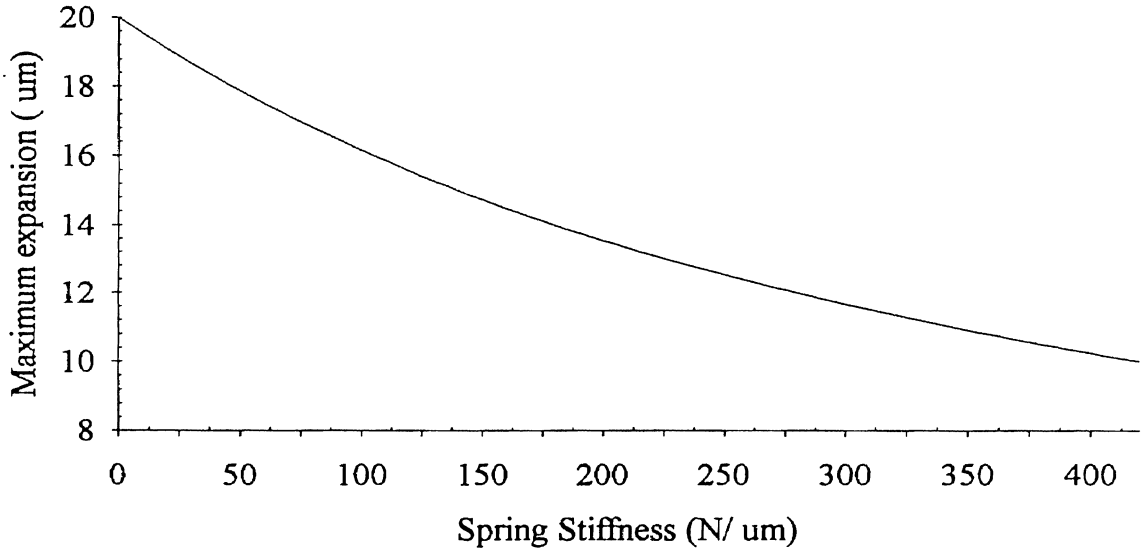


Figure E.2. Influence of Preload on Expansion of PZ Stack

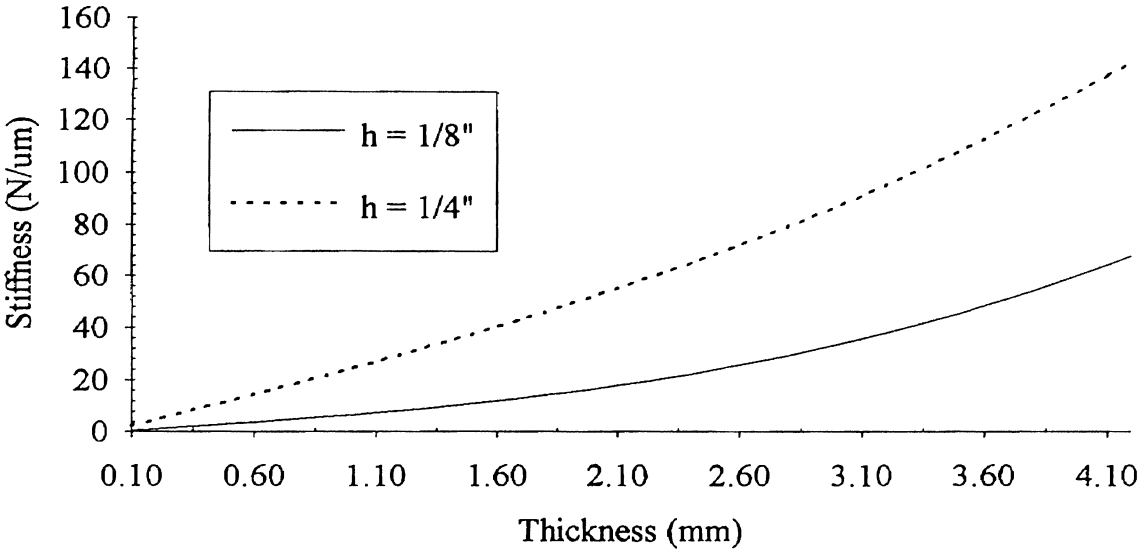


Figure E.3. Stiffness of Belleville Springs

It can be seen that a thickness of 3 mm and cone height of 6 mm will yield a stiffness of approximately 80 N/μm which should significantly improve preload and dynamic response. To control mechanical noise Teflon shim stock might be

placed under the areas where the Belleville spring contacts the inner and outer clamps. Under compression the Teflon would fill the asperities of the metal washer and provide a uniform contact area.

E.2 Cross Carriage Actuation Mechanism.

At some point a second device for driving the cross carriage will be installed. It is the purpose of this section to provide suggestions for this device.

The overall travel of the cross carriage is to be 250 μm. There are no PZ stacks with this travel that are not excessively long and compliant (more than 12" long and stiffnesses less than 0.1 N/μm). There are various bimorph shapes which are very compact but they are still too compliant. It should be noted that stiffness is not nearly as important an issue for this device as compared to the tool handling stack. The primary function of this driver is to provide a constant velocity motion for the sample. Structural stiffness is to be provided by the rectangular flexures. Stiffness is not irrelevant in that it influences the minimum expansion time for the assembly. However the speeds needed should not require extremely high stiffness. In fact rapid expansion should be avoided as it will apply forces proportional to the acceleration. Considering this force with the cutting force there will be a couple applied to the cross carriage which will tend to twist the carriage about an axis parallel to the z-axis. Figure E.5 shows a possible drive mechanism for the cross carriage. The stack pushes on a piston which contacts a flat circular spring near the outer edge. The flat washer in turn pushes on a ball bearing and pushrod located at the center of the spring. The mechanical amplification is proportional to r_0/d where r_0 is the outer radius and d is the distance from the outer edge to the contact point of the piston.

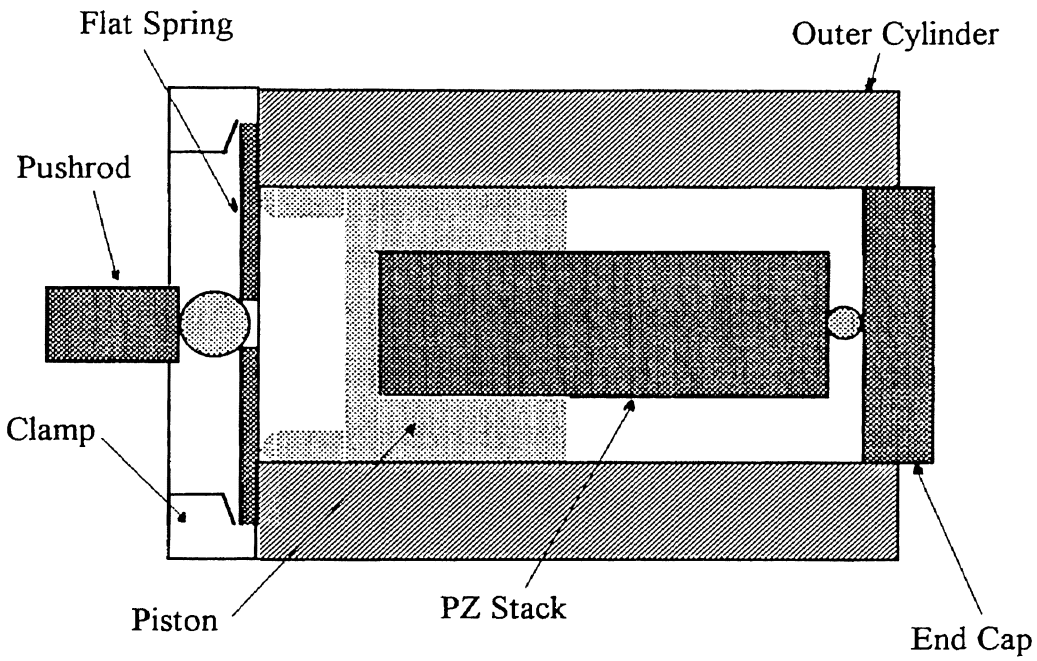


Figure E.5. Proposed Drive Mechanism for Sample Carriage

The most recently contemplated PZ stack has a nominal expansion of 80 μm. To obtain the desired expansion for this stack the amplification would have to be about 3:1. The stiffness of a flat washer for small expansions is [Wahl,1944]

$$k = \frac{Et^3}{nCr_0^2} \tag{E.4}$$

where n is the mechanical amplification factor; in this case 3, and

$$C = 0.551\left(1+\xi^2\right) + 1.614\left(\frac{\xi^2}{1-\xi^2}\ln^2\xi\right). \tag{E.5}$$

The stiffness of a flat washer of outer diameter 5 cm is shown as a function of aspect ratio for thicknesses of 2 mm, 3 mm and 4 mm. To fit a ball bearing the inner diameter will likely be from several millimeters to 1 cm. Thus the aspect ratio will be between 0.7 and 0.2. The lower the stiffness the better in this application so the minimum thickness should be chosen consistent with small

deflection or bowing of the disk itself. Stiffness as a function of aspect ratio for several thicknesses is shown in Figure E.6.

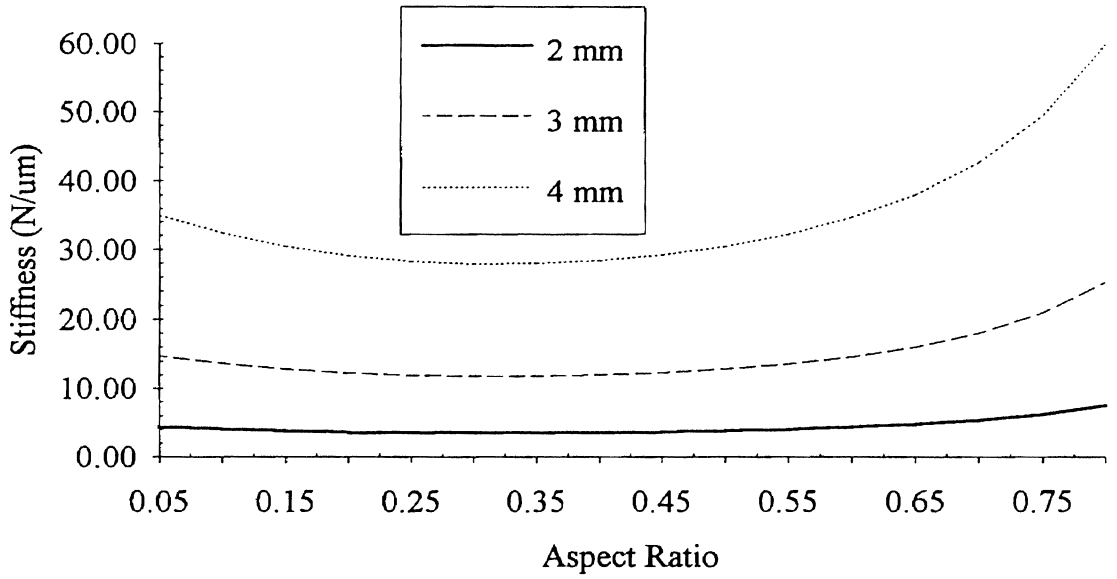


Figure E.6. Stiffness of a Flat Washer for Several Thicknesses

The last point to be made about the second axis is with regard to electrical properties. The two stacks considered are the PI P-245.50 and the P-245.70. Their relevant specifications are shown in Table E.1. The expansion should proceed as shown in Figure E.7, with t_1 being the time for expansion to the maximum displacement and maximum voltage. The slew rate required for a full 250 μm expansion of the cross carriage (80 μm and 120 μm for the two stacks) for a 1000 V PZT at 0.1 mm/s and 10 mm/s is 400 V/s and 4×10^5 V/s, respectively. These slew rate are within the capability of many power supplies, including the BOP used on the first axis. The maximum current required is given by equation (A.4) in Appendix A. The currents for the 10 mm/s expansion are well beyond the abilities of most off-the-shelf power supplies. The solution is to find lower

capacitance PZT's ("soft" PZT's may help in this regard), use a different type of driver such as a stepper motor or lower the maximum cutting speed.

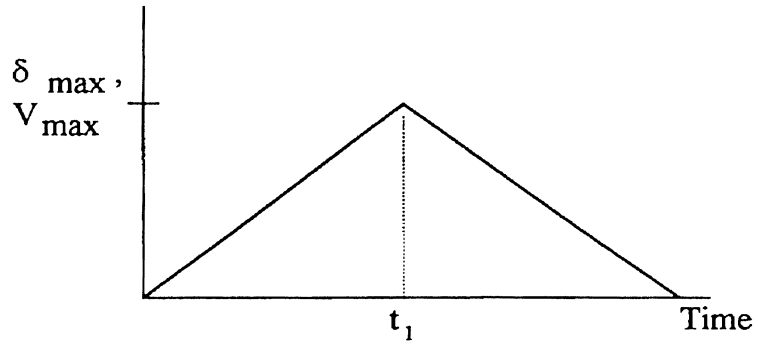


Figure E.7. Voltage Profile for Second Axis

Specification	P-245.50	P-245.70
Expansion @ 1000V (μm)	80	120
Stiffness (N/μm)	15	8
Capacitance @ 0 V (nF)	310	450
Resonant Frequency (kHz)	8	5
Coefficient of Thermal Expansion (μm/°C)	0.7	0.9
Current Required for 0.1 mm/s Expansion	0.12 mA	0.18 mA
Current Required for 10 mm/s Expansion	127 mA	180 mA

Table E.1. Specifications of Two PI Stacks

VITA

Paul N. Moss

Candidate for the Degree of

Master of Science

Thesis: THE DESIGN OF A MICROMACHINING DEVICE: PHASE ONE

Major Field: Mechanical Engineering

Biographical:

Personal Data: Born in Tulsa, Oklahoma on July 28, 1958, the son of Paul A. and Roberta Moss.

Education: Graduated from Tulsa Central High School in May 1976; received Bachelor of Science Degree in Physics from Oklahoma State University in December 1982. Master of Science in Mechanical Engineering received December 1993.

Professional Experience: Teaching Assistant for the Department of Physics January 1991 to August 1992; Research Assistant for the Department of Mechanical Engineering September 1992 to July 1993. Physics Department Head at Spartan School of Aeronautics November 1988 to December 1990.

VITA 2

Paul N. Moss

Candidate for the Degree of

Master of Science

Thesis: THE DESIGN OF A MICROMACHINING INSTRUMENT: PHASE ONE.

Major Field: Mechanical Engineering

Biographical:

Personal Data: Born in Tulsa, Oklahoma on July 28, 1958, the son of Paul A. and Roberta Moss.

Education: Graduated from Tulsa Central High School in May 1976; received Bachelor of Science Degree in Physics from Oklahoma State University in December 1982. Master of Science in Mechanical Engineering received December, 1993.

Professional Experience: Teaching Assistant for the Department of Physics January 1991 to August 1992; Research Assistant for the Department of Mechanical Engineering September 1992 to July 1993. Physics Department Head at Spartan School of Aeronautics November 1988 to December 1990.