Pick and Place Silicon on Insulator Microassembly

by

Matthew Emanuel Last

B.S. (University of California, San Diego) 1998

A dissertation submitted in partial satisfaction of the requirements for the degree of

Doctor of Philosophy

in

Engineering – Electrical Engineering and Computer Sciences

In the

GRADUATE DIVISION

Of the

UNIVERSITY OF CALIFORNIA, BERKELEY

Committee in charge:

Professor Kristofer S.J. Pister, Chair

Professor Ming C. Wu

Professor Liwei Lin

Fall 2005

The dissertation of Matthew Emanuel Last is approved:

_____

Professor Kristofer S.J. Pister, Chair   Date

_____

Professor Ming C. Wu   Date

_____

Professor Liwei Lin   Date

University of California, Berkeley

Fall 2005

Pick and Place Silicon on Insulator Microassembly

Copyright 2005

by

Matthew Emanuel Last

Abstract

Pick and Place Silicon on Insulator Microassembly

by

Matthew Emanuel Last

Doctor of Philosophy in Engineering – Electrical Engineering and Computer Sciences

University of California, Berkeley

Professor Kristofer S.J. Pister, Chair

Deep reactive ion etching into Silicon on Insulator (SOI) wafers is a popular method of

fabricating high-performance MEMS devices. These include electrostatic and thermal

actuators, flexures that guide very precise motion, and ultra-flat and smooth reflectors to

make micro-mirrors. An important limitation of SOI-based MEMS is that it is not easy to

achieve a large range of out-of-plane motion without a complex fabrication process. This

is because of the difficulty in breaking symmetry in the vertical direction: the actuator

forces line up with the flexures that keep the parts moving in plane. Serial pick and place

microassembly provides a solution to this problem by allowing us to create complex

mechanisms in a simple process, then rotate these mechanisms ninety degrees out of

plane to convert the in-plane motion that SOI-based technology excels at to out-of-plane

motion. This approach improves processing yield and manufacturability while drastically

reducing design cycle time compared with other solutions to the out-of-plane motion

problem. This dissertation describes the novel micro-tooling used to pick up parts and automatically rotate them ninety degrees out of plane, a Zero Insertion Force (ZIF) connector technology that forms a strong permanent mechanical and electrical connection between the assembled parts and the rest of the chip, and several one- and two-axis out-of-plane rotation mechanisms that rival the performance of anything in the literature. All of these items were built in the same single-mask SOI process.

_____

Professor Kristofer S.J. Pister, Chair   Date

I dedicate this dissertation to my family and friends. Thank you for your patience

and support

Table of Contents

Acknowledgements

Thanks go to:

KSJ Pister: for always steering me in the right direction.

My dissertation committee: for helpful feedback and constructive criticism…

Subramaniam V: for becoming a full partner in the creation of the microassembly process. You not only helped refine good ideas, but even more importantly you helped me recognize my bad ones. Your help and original contributions throughout this project were instrumental to its success.

S Bergbreiter: for your insightful, detailed edits of this manuscript and interest in the progress of the microassembly process as it evolved. You were the first person I wound up telling about any breakthrough or failure, and after talking to you I'd feel better about either one.

V Milanovic: for teaching me pretty much everything I know about processing.

L Zhou: for great discussions about micromirror design.

B Cagdaser: for deepening my understanding of the dynamics of the mirrors and for showing us all how control circuits for micromirrors should be built.

C Keller: for being microassembly's strongest advocate. And for having forgotten more about fabrication than the rest of us are likely to ever know…

S Lanzisera, D Black, Mike Last, Mom and Dad: for letting me bounce ideas off you, for your nuggets of brilliance, and for your overall enthusiasms and interest over the years.

Microlab members and staff: For keeping the machines running.

Friends: for keeping me (somewhat) sane!

1. INTRODUCTION

Deep reactive ion etching into Silicon on Insulator (SOI) wafers has become a popular method for building high-performance MEMS structures. SOI MEMS advantages stem from the combination of high aspect-ratio flexures and actuators with the favorable material properties of single-crystal silicon. The resulting structures retain much of the design flexibility of surface micromachined MEMS, with the added benefits of high force-density actuators and smooth and flat surfaces suitable for optical applications. In contrast, surface micromachining incorporates multiple structural layers, offering three primary benefits over SOI micromachining: 1) structures can be folded up out of plane, enabling microsystems to move both in and out of the plane of the chip, 2) bushings can be designed to keep structures such as gears, motors, and other moving structures from moving in undesired directions, and 3) electrical wiring can be kept compact by using multiple routing layers. Finding a way to combine the positive features of surface micromachining with the benefits of SOI micromachining could improve the performance of many microsystems and enable the creation of others that have not yet been practical to build.

Both surface micromachining processes and SOI micromachining processes have been developed that try to incorporate the desirable features of the other. Sandia's SUMMiT V process is a surface micromachining process that creates high aspect-ratio actuators and flexures by stacking up to 5 thin polysilicon beams on top of one another, resulting in laminated structures up to 12μm thick. Since this process incorporates a chemical-

mechanical polishing (CMP) step after each layer is defined, the polysilicon surfaces are also smooth enough for optical applications. By controlling the deposition, doping, and annealing conditions, released polysilicon structures with a radius of curvature >25cm (Jensen, de Boer et al. 2001) can be created which also benefit optical applications.

The biggest hurdle preventing SOI processes from acquiring the desirable features of surface micromachining processes is the presence of the deep, high aspect ratio trenches that make SOI-based processes so compelling; while they are useful for making actuators and flexures, they make it very difficult to make additional layers that bridge those trenches. Once those trenches are etched, the topography makes spinning on a constant-thickness layer of photoresist very difficult since the photoresist is only ~1-2μm thick, while the trenches can be up to 50μm deep. Even if photoresist can be applied in an even layer, all optical lithography tools have an inherent tradeoff between resolution and depth of focus, and very few lithography tools have been designed to accommodate more than a couple of microns of topography, and none are available in the UC Berkeley microlab. The most straightforward method of planarizing the wafer after the deep trench is to refill the trenches with a sacrificial material, such as a low-pressure, chemical-vapor deposited (LPCVD) oxide. This process can deposit fairly conformal layers of glass up to 2-3μm thick without the layer's internal stress causing cracking to occur, placing an upper bound on the widths of the trenches that can be refilled. Any trenches wider than twice the thickness of the deposited layer will not refill completely. (Maharbiz, Cohn et al. 1999) and (Cohn, Bohringer et al. 1998) utilized this refill method to planarize the surface of their wafers, allowing them to create patterned gold bumps on their chips that were used

as rivets to transfer high aspect-ratio SOI structures to other substrates. These included both other SOI substrates and surface micromachined substrates. The drawback to this approach is the strong limitations on trench width imposed by the limited thickness of the refill layer.

Wider trenches have been bridged using the bump bonding techniques described in (Maharbiz, Howe et al. 1999). In this case, surface micromachined and molded polysilicon structures were transferred from the substrate on which they were manufactured to a target substrate. The target substrate had wide trenches that were successfully bridged by these transferred structures.

(Hollar, Flynn et al. 2003) planarized their SOI wafers using a spin-on glass followed by a chemical mechanical polishing step. After planarization, they did a complete 2-layer polysilicon surface micromachining process on top of their planarized SOI substrate. In this way, they were able to add hinged structures to the tops of their high aspect-ratio SOI actuators.

Finally, (Zhou, Last et al. 2003) used hydrophilic Silicon-silicon fusion bonding to stack two patterned SOI wafers on top of one another. By removing the handle wafer from one of the SOI wafers, the patterned device layer was transferred from one wafer to the other. Very small structures are able to be successfully transferred, and can bridge large gaps, limited only by the ability of the structure to withstand pull-down forces from the surface

tension of water encountered during various rinsing steps involved in fabrication. (Figure 1)

| Mag = 162 X | 100µm | EHT = 4.85 kV | Signal A = SE2 | Date :4 Dec 2000 |
| | | WD = 8 mm | Photo No. = 495 | Time :22:23 |
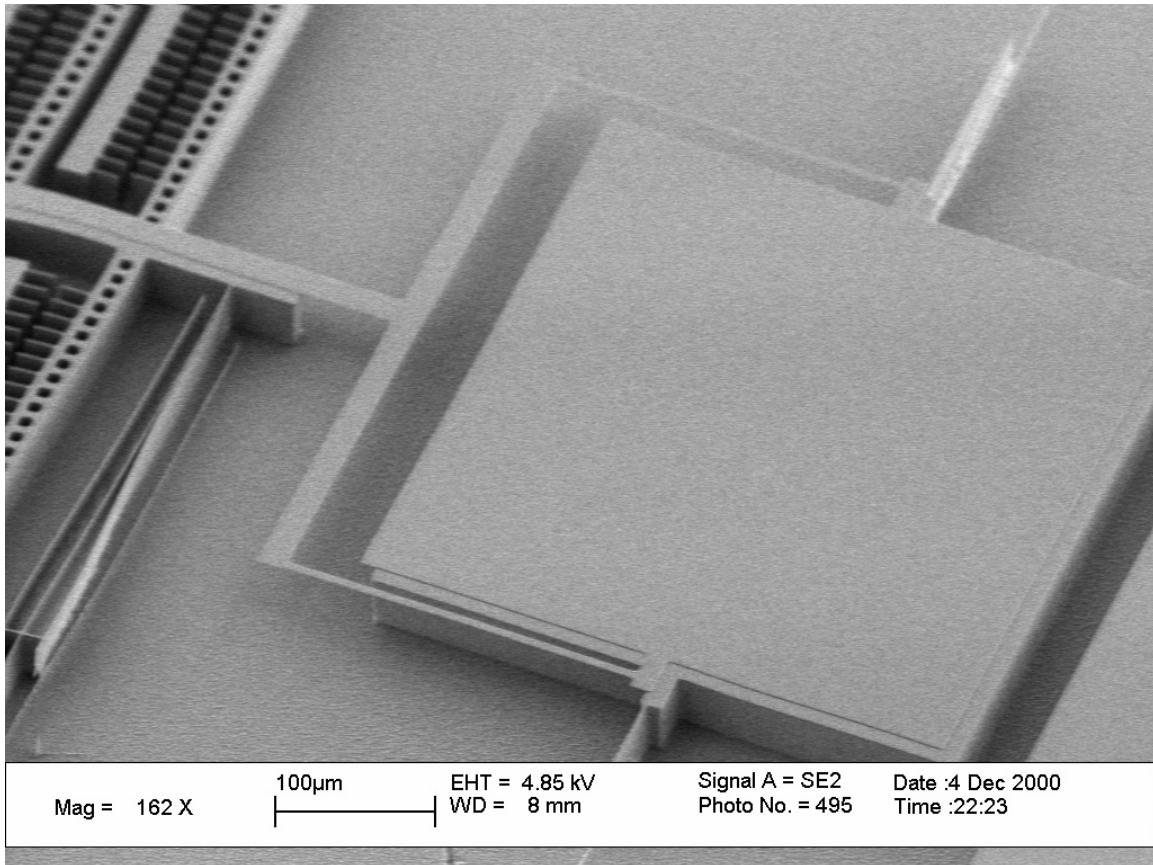
Figure 1. Single crystal silicon beams bridging deep trenches. The thin beams were transfered from a SOI wafer using hydrophilic silicon-silicon fusion bonding, after which the handle layer of the SOI wafer was removed.

Post-process assembly of SOI structures offers an alternative to previous attempts at complex processing that recovers the flexibility of surface micromachining processes. Complex structures can be fabricated on an SOI substrate, then assembled so that they form structures that bridge trenches and move out of plane. (Maharbiz et al. 1999) demonstrated a chip-level assembly process that utilized gold bump compression bonding

to transfer thin polysilicon beams on top of an already-etched SOI chip, bridging trenches and creating hinge structures to guide the manual assembly of SOI mirrors out of plane. Keller and Chang (Chang, Keller et al. 2005) used deep-reactive ion etched structures and manually assembled them into a space frame that was used as a micro-scale scalpel for slicing neural cells. Yet another assembly process was demonstrated by Skidmore et al (Skidmore, Ellis et al. 2004) at Zyvex, inc. In the Zyvex process, an SOI wafer was patterned, etched, and the parts fabricated on each chip were released. Afterwards, a macro-scale pick-and-place robot with a micro-scale end effector would grasp a part, break the tether holding it in place, rotate the part 90 degrees out of plane, and assemble it into a specially-designed snaplock located elsewhere on the chip. This approach promises to produce the whole feature set of a surface micromachining process – out-of-plane motion, bushings, and multilevel electrical wiring. Zyvex demonstrated multilevel electrical wiring (Huang, Nilsen et al. 2003) and out of plane motion (Zyvex Inc 2004), but never fully explored the mechanical capabilities of this process. No bushings or other interesting mechanical elements taking advantage of the out-of-plane nature of the assembly process were demonstrated. And the out-of-plane actuators that were built were never optimized or made to do anything useful. This dissertation addresses these limitations.

A successful microassembly process consists of four main elements: 1) a part-making process, 2) micro-tooling to grip, lift, rotate, and release parts at the appropriate time, 3) a macro-scale machine to maneuver the micro-tooling around the assembly workspace, and 4) some method of securely fastening micro-parts to one another.

There have been several examples of successful microassembly processes demonstrated in the literature. Dechev, Mills (Dechev, Cleghorn et al. 2002, 2003, 2004, 2004) use a surface micromachining process to build parts and micro-grippers to manipulate those parts, a 5 degree-of-freedom macro-scale robot with a soldering iron as an interface to the micro-world, and snaplocks to fasten parts to one another. This process has been used to make air-core inductors (Dechev et al. 2004). The biggest drawback to these inductors was the high contact resistance between the micro-parts and the sockets they were inserted into, yielding a low quality factor. No assembled moving micromechanical devices were built.

Guckel (Guckel, Skrobis et al. 1991; Christenson, Guckel et al. 1992; Guckel, Christenson et al. 1992, 1992; Guckel, Christenson et al. 1994) used the LIGA process to fabricate high aspect-ratio gears and hubs for a variable reluctance motor. In addition, a linear solenoid consisting of 5 stacked layers of LIGA-fabricated flexures was demonstrated (Fischer and Guckel 1998). While no specialized macro-tooling was described in the published work, high precision sub-micron alignment of the various parts was attained by using precision-fabricated alignment posts created again using LIGA. These alignment pegs were required in order to place the gears on their hubs and to mesh the various gears in the gear train. In addition, they were instrumental in aligning and holding the multiple layers of the solenoid immobile during operation. Both of these devices demonstrated solid performance comparable to or better than their monolithically-fabricated counterparts. No out-of-plane motion was ever demonstrated

with this technology however, due to the lack of out-of-plane rotation of the parts before assembly.

While primarily focusing on micro-tooling, Keller has also explored micro-assembly (Keller 2005) to build effective tools. His micro-grippers were the first demonstration of a strong mechanical connection formed by linking microfabricated silicon parts at 90 degrees to one another to constrain motion in all directions (Figure 2). He has extended this approach in his micro-scalpel tool (Chang et al. 2005), which used an assembled space frame to precisely position a microfabricated scalpel that was used to slice individual axons on neurons grown in-vitro.
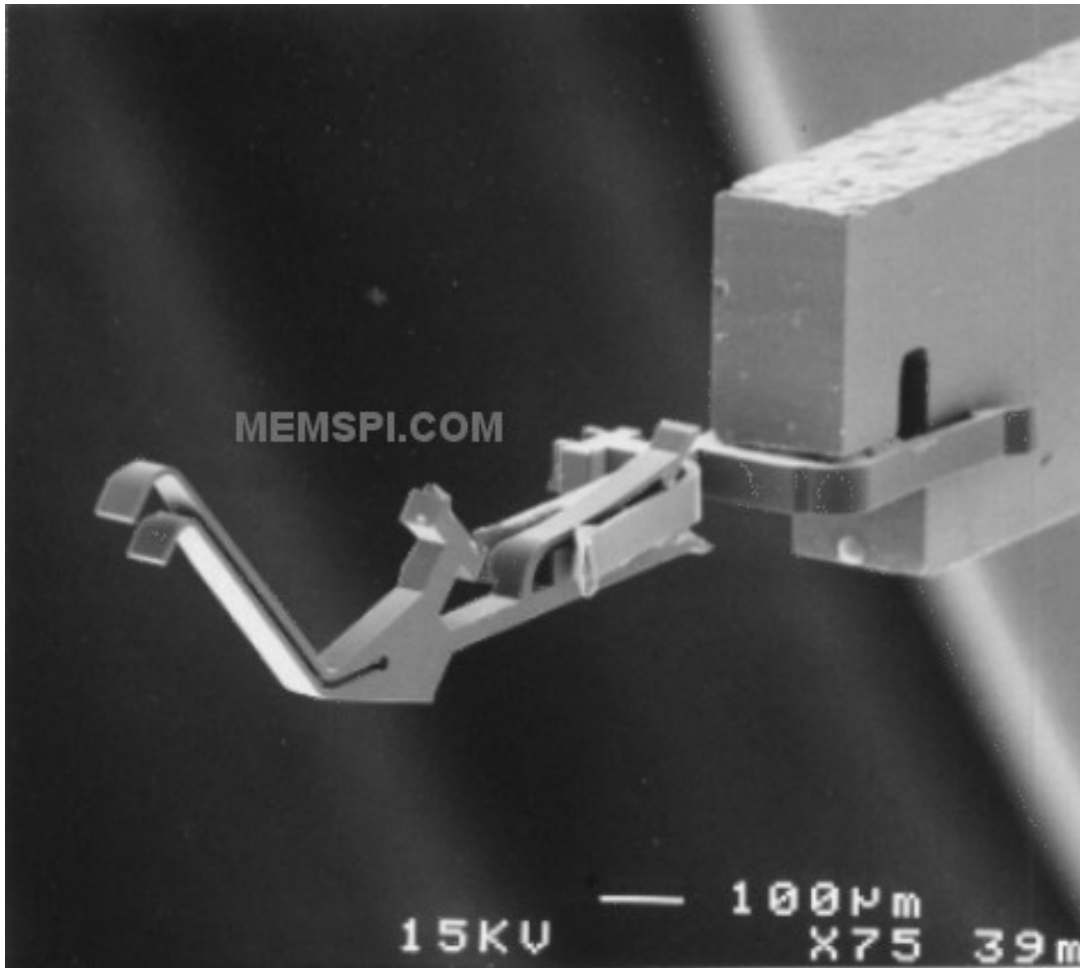
Figure 2. Electrostatic "picker"-type micro-gripper. This is a demonstration of the assembly of planar structures at ninety-degree angles to one another in order to form a 3D structure. From http://www.MEMSPI.com

There is a rich history of both monolithic and assembled approaches to building complicated MEMS structures. Both approaches have proven successful at combining the flexibility of surface micromachining with the performance of SOI/DRIE micromachining. The main lesson to be taken from this is that both approaches can create a wide variety of structures. Perhaps the choice of which approach to use depends on other factors than whether a particular device can be built or not.

1.1. Monolithic vs Assembled

Microassembly seems to have fallen out of favor in recent years in the MEMS world. Detractors note that microassembly is usually a serial process, with each assembly operation typically involving a skilled laborer manually assembling parts to build a finished device over the span of several minutes. This amount of labor will have adverse effects in terms of manufacturing throughput and cost, and runs counter to the drive towards parallel or batch-based manufacturing. The detractors have it wrong: it will be shown in this chapter that microassembly actually adds nominal cost to manufacturing and instead leverages the advantages of wafer-based manufacturing with the added benefit of a simpler manufacturing process.

Wafer-based manufacturing features two main improvements over traditional machining. The first benefit stems from the actual size of the devices: some mechanical problems are best solved using small machines. Due to steady progress in the technology of lithography, features smaller than 100 nanometers can now be reliably patterned. Examples of microsystems that benefit from downward scaling in the size of the devices are atomic force microscopy and inkjet printing. Scaling these devices downward in size improves their performance. The second advantage stems from economics; wafer-level manufacturing is a batch process where the cost of processing a single wafer remains roughly constant, while the number of devices per wafer scales upward with increased wafer size and decreased minimum feature size. These two effects are the same scaling phenomena that power Moore's law.

As with electronic microchips, all microsystems need to be handled individually at some point in the manufacturing process. This usually occurs at the very end of the manufacturing process, at the point where the chips are diced, tested, and packaged. The testing and packaging stages (called the post-process) are usually the most expensive part of the manufacturing process because the sudden shift from wafer-level batch processing to chip-level serial processing means that the cost to process each wafer increases with the number of chips, not the number of wafers. With microsystems, it is not unusual for the cost of packaging a finished device to dominate the cost of manufacturing the device. It therefore follows that any reduction in the cost of the post-process will couple strongly into the cost of the finished part. This desire to minimize cost in the post-process often leads designers to push towards wafer-level packaging solutions. However, automation offers a different approach to post-process cost reduction.

Robots have been developed that speed up and reduce the cost of the packaging process. Pick-and-place robots can move a tested chip from the diced wafer and bond it to a package in less than a second. Automated high-speed wirebonders can place over 30 bonds (15 wires) per second (Kulicke & Soffa inc 2005). Even though it is a serial process involving multiple operations per chip, the throughput of this operation can exceed the throughput of certain wafer-level operations such as etching. To illustrate this point, consider the capabilities of the UC Berkeley microlab. Fifty-two 1cmx1cm chips can be fit onto a 4" wafer. In the process described in the next section, etching through the device layer requires 18.3 minutes. If our lab's etcher and the die- and wirebonders cited above were running continuously, no bottleneck would form unless each chip

required more than 21 seconds to process; this is the equivalent of 316 wirebonds. This number drops to approximately 9 seconds for 6" wafers, and just under 5 seconds for 8" wafers. Since the cost of a deep-reactive ion etcher exceeds 10x the cost of a wirebonder, it is not unreasonable to expect the throughput of a DRIE-based process to be limited by the availability of etchers, not assembly machines.

Current state of the art of microassembly pick-and-place robots, such as the pick and place robot at Zyvex, can do roughly one assembly every 25 seconds (Zyvex Inc 2004). In the future, pick and place robots will be more similar to the high-speed wirebonders of today than the die-bonding or Zyvex-style robots. An assembly robot that uses the micromachined end-effector described later in this chapter would only need the same number of actuated degrees of freedom as the Kulicke & Soffa wirebonder cited above to perform microassembly tasks. Therefore, it is not unreasonable to expect to see sub-second microassembly times in the not-too-distant future.

2. FABRICATION: MAKING PARTS

The work in this thesis began as a thought experiment: what is the simplest manufacturing process possible that would result in good tip/tilt micromirrors? This came from a long period of time of pursuing monolithic processes that would create fully functional micromirrors by the wafer-full. None of these processes were successful in creating the high performance mirrors being pursued with anything approaching a useful yield. After a great deal of thought, it was determined that a single-mask process involving a single deep etch followed by dicing and a timed release was about as simple a process as possible. By pushing complexity out of the wafer-level fabrication and into the post-process, it was expected that yield and performance could be achieved at the same time.

The first step of this process was the fabrication of the Silicon on Insulator (SOI) wafers. SOI wafers consist of three layers: a handle wafer made of silicon, a buried insulating oxide, and a device layer also made of silicon. They are generally formed by bonding together two wafers, each of which has been oxidized. One side of the bonded wafer pair is then ground down and polished to form the device layer, leaving the un-ground wafer as the handle layer. The SOI wafers used in this process feature an extra-thick buried oxide. Although later generations of micro-grippers would render this issue moot, when this project was started it was determined that this thick layer would be necessary to provide clearance for the micro-grippers to grab the underside of the fabricated parts. It was determined that the largest practical buried oxide thickness was 5μm of wet oxide,

consisting of 2.5μm on each of the two bonded wafers. This maximum thickness was determined by constraints on available furnace time; at the furnace's maximum operating temperature of 1075°C, 14.75 hours were required to grow 2.5μm of oxide.

Fabrication of SOI wafers is a little bit of a black art. In order to achieve high-yield bonds, special care must be taken with the treatment of the bonding surfaces. In order to start with the most pristine surface possible, an initial oxidation of 5800Å at 1075°C was performed. This initial oxidation consumes silicon at the surface of the wafer. The interface between the edge of the oxidation front and the single-crystal silicon below it will be the bonding surface once the initial oxide is removed in dilute HF acid. This initial oxidation smoothes any irregularities in the silicon surface through two mechanisms: 1) the extra-high temperature of this oxidation step allows the oxide to flow, letting surface tension even out any irregularities in the oxide surface and 2) the increased oxidation rate at peaks in the silicon surface and decreased rate in valleys eventually smoothes out the etch front. This oxide is then stripped in 25:1 HF immediately before the oxidation step that creates the buried oxide, exposing the bonding surface to the environment for the least amount of time possible. It is important to use 25:1 HF, as Buffered HF or 10:1 HF can roughen the silicon surface and impair the ability of the wafers to bond (Ljungberg, Soderbarg et al. 1993). After the long oxidation that creates the buried oxide layer, the wafers are removed from the furnace, allowed to cool, and the surfaces are activated in a fresh batch of piranha (sulfuric acid and hydrogen peroxide). After rinsing in DI water until a resistivity of $10^{15}\,\Omega$-cm is achieved, the wafers are spun dry in a SRD (Spin-Rinse-Dryer). From this point on, exposure to ambient air is

minimized to reduce the chances of exposure to dust and other particulates that can destroy the yield of the bonding step. The cassette is removed from the SRD and placed in a covered box. A Teflon bonding chuck (Figure 3) consisting of two flat surfaces oriented orthogonally to one another provides coarse wafer alignment and a clean surface for bonding. Handling the wafers with a vacuum wand and only touching the non-bonding face of each wafer, the first wafer is placed face up on the bonding chuck and the second wafer is held face down above the first wafer. Releasing the wafer while it is still approximately 1cm above the first wafer traps a layer of air between the two wafers and allows the second wafer to float above the first wafer on this air bearing for several seconds. This is enough time for the wafer flats to align to one another before bonding occurs. Once the flats are aligned, the centers of the wafers are gently pressed together for approximately 2 seconds using a corner of the vacuum wand. This ensures that the bonding front has moved completely across the wafer, squeezing out any air from between the wafers. The pre-bonded pair are then placed back into a cassette and run through the SRD before they are placed into the furnace for an anneal at 1050°C for 10 hours in dry nitrogen. This second spindrying step puts the wafers under some mechanical stress. In this way, wafers that have not bonded well (ie have large voids) are identified before the anneal step, where they would have been permanently bonded together.

Figure 3. Making SOI wafers: Author shown bonding wafers using teflon bonding chuck.

A certain amount of paranoia about contamination results in improved bonding yield. It is found that in the UC Berkeley microlab, performing the bonding step late at night and/or on the weekend, wearing a fresh cleanroom gown, being careful to minimize time that the wafer is exposed to the environment (especially when the wafer is facing upwards and is at most risk of particulates landing on the surface to be bonded), and using a face mask and shield results in the highest bonding yield. This is most likely due to the lower amounts of foot traffic and the resultant decrease in the amount of airborne particulates shed by people, especially the person doing the bonding. Using the techniques described here, a very good yield is possible; the bonding run used to fabricate the devices in this dissertation created 28 bonded pairs of wafers and resulted in 22 perfect (zero defect) wafers and 6 wafers with either one or two voids less than 2mm

across. Two of those wafers with voids also contained a large void with a diameter of ~1cm.

$$\Delta t_{init} = 2*0.58\mu m = 1.16\mu m$$

Once the wafers have been bonded, they are sent to an outside vendor (in our case, Silicon Quest) for a high-precision grind/polish step. This vendor guarantees a final wafer thickness within +/-1μm. The device layer thickness variation is not determined by this step however; the tolerance of the starting wafers is +/-5μm from the wafer vendor we use (International Wafer Services). The wafer thicknesses, t_{handle}, were measured by averaging the measured thicknesses at three locations on the wafer surface after the initial oxidation (nominal value for t_{handle} is 400μm) . The wafers used as the handle wafer in the SOI wafers were cherry picked for their similar thicknesses. In this way, we were able to reduce the variation in wafer thickness from +/-5μm to +/-2μm.

The total thickness of the desired SOI wafer is a function of the measured thickness of the handle wafer and the various oxidations performed on the wafer during bonding. After the handle wafer thicknesses are measured, the initial oxide is stripped, reducing the thickness of the handle wafer. In our case, since 5800Å of oxide are present on each side of the wafer, the wafer thickness shrinks by . After this oxide is stripped, the handle wafer and the device layer wafer are each oxidized again by half the thickness of the SOI buried oxide and bonded together. Since silicon is being consumed to create silicon dioxide, every micron of silicon dioxide grown expands the wafer by only 540nm. Three of these oxide layers, the oxide on the backside of the handle wafer, the oxide on the frontside of the handle wafer, and the oxide contributed by the device wafer must be

included in the total thickness of the SOI wafer. In this case, the total increase in thickness is $\Delta t_{BOX} = 0.54*3*2.5\mu m = 4.1\mu m$. This yields a worst-case SOI wafer thickness of $t_{SOI} = t_{handle} - \Delta t_{init} + \Delta t_{BOX} + t_{device} = 423 \pm 3\mu m$, where the variation comes from the +/-2μm variation in the handle wafer thickness and the +/-1μm total thickness variation from the grind/polish step. Unfortunately, this variation in total wafer thickness shows up entirely in the device layer thickness since the total wafer thickness is controlled to within a micron. Therefore, all designs relying on the device layer thickness included a safety margin of +/-5μm to accommodate this uncertainty.

Once the wafers have been ground and polished, they are ready for lithography. 1.6μm of I-line photoresist is spun onto the wafer, followed by lithography in our lab's 4" GCA wafer stepper. Although this stepper is capable of resolving lines and spaces down to just below 1μm, no features smaller than 1.5μm were made. This removed much of the focus, exposure, and development time tweaking required to obtain features that small. A 60sec post-exposure bake at 120°C increases contrast and aids in resolving fine features. After one minute in I-line photoresist developer, the wafer is inspected, and the photoresist is stripped and lithography repeated if everything does not look right at this stage. Failure modes encountered included bubbles in the photoresist when the humidity in the lab is too high (Figure 4 and Figure 5), the wrong exposure settings on the stepper, design errors, and mask issues such as dirty masks and mask fabrication errors. After a 1-minute exposure to a 50-Watt O_2 plasma to remove photoresist scum in the trenches, a UV-assisted high-temperature hardbake step polymerizes the photoresist to increase the resistance of the photoresist to the deep etcher.
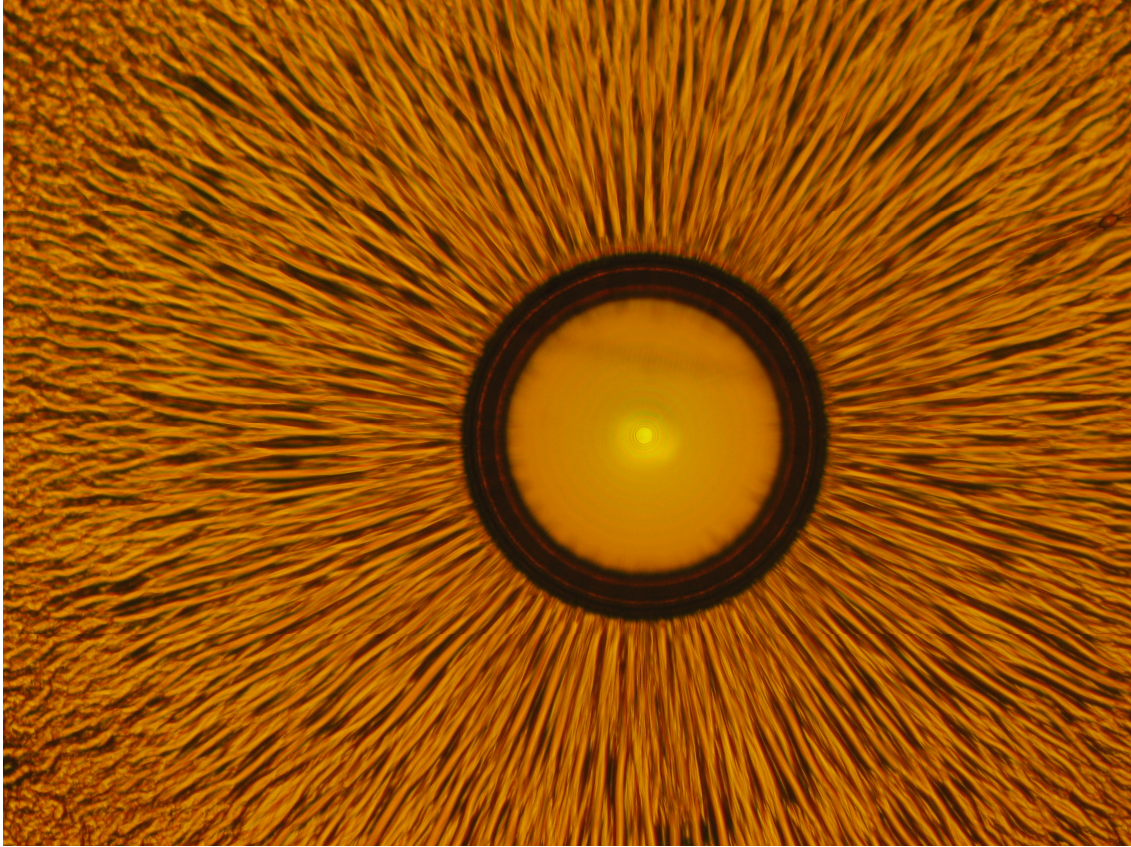
Figure 4. Photoresist bubble closeup
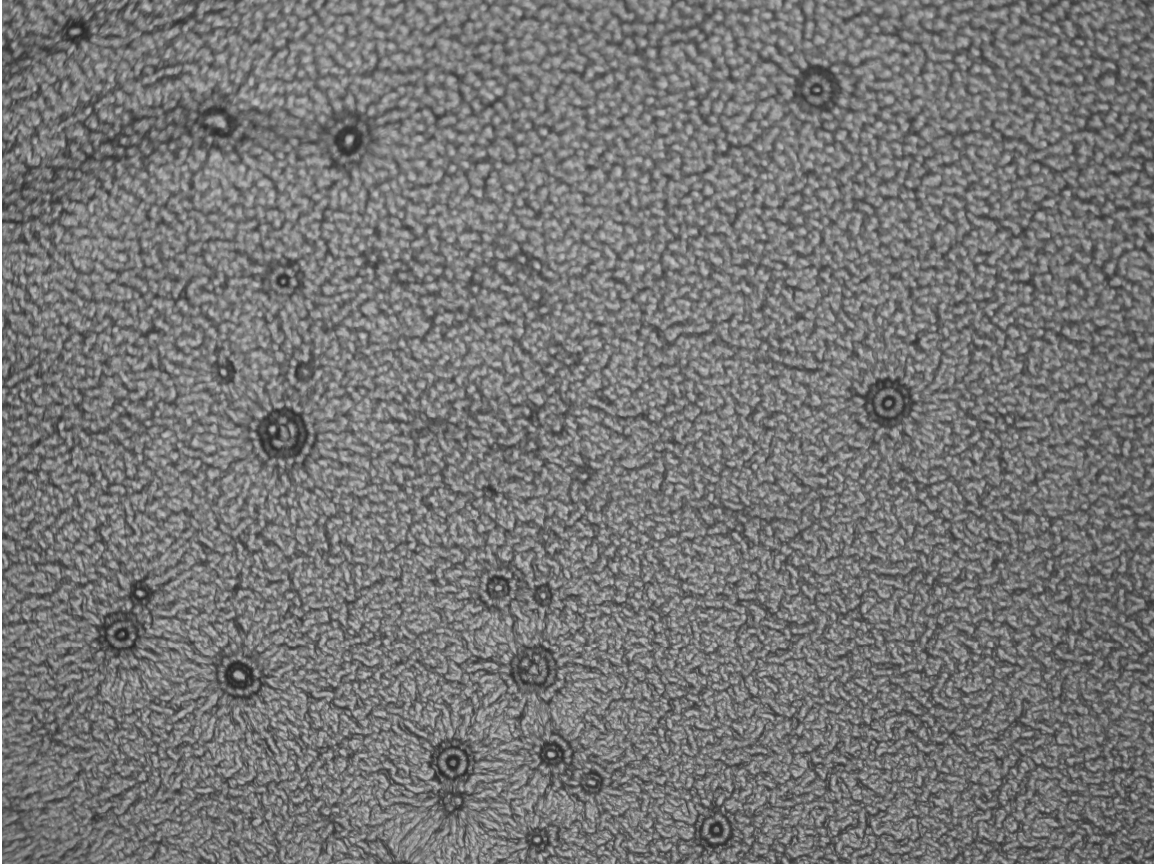
Figure 5. Small photoresist bubbles

Once hardbaked, the wafers are etched in the Surface Technology Systems (STS) deep
reactive-ion etcher (DRIE). This system utilizes the Bosch process, which alternates
short isotropic etches and short Teflon-like polymer deposition steps to achieve a large
trench etch aspect ratio ($AR = \dfrac{depth}{width}$). One drawback of this machine is that it suffers
two major etch rate non-uniformities: 1) cross-wafer, and 2) aspect-ratio dependent etch-
rate (ARDE, aka RIE lag). The cross-wafer etch-rate nonuniformity means that trenches
at the edges of the wafer etch approximately 10% faster than trenches at the center.
ARDE means that narrow trenches etch slower than wide trenches. This effect becomes

quite noticeable for trenches narrower than ~5µm. Together, these two effects result in a

lot of time at the end of the etch recipe where some trenches are etched all the way to the

buried oxide, while others are still being etched through silicon. Because it is necessary

to etch until all the trenches have reached the buried oxide, the areas of buried oxide that

are exposed first can collect a charge, reflecting reactive ions upward towards the bottom

of the trench sidewalls. This causes a "footing" effect (Figure 6), where the bottoms of

structures near the buried oxide are etched laterally. Footing has been observed to

completely undercut narrow structures. To counteract this effect, the bias voltage applied

to the chuck is pulsed on and off at 32kHz, allowing time for the trapped charge on the

oxide to be neutralized by the free electrons in the plasma and all but eliminating the

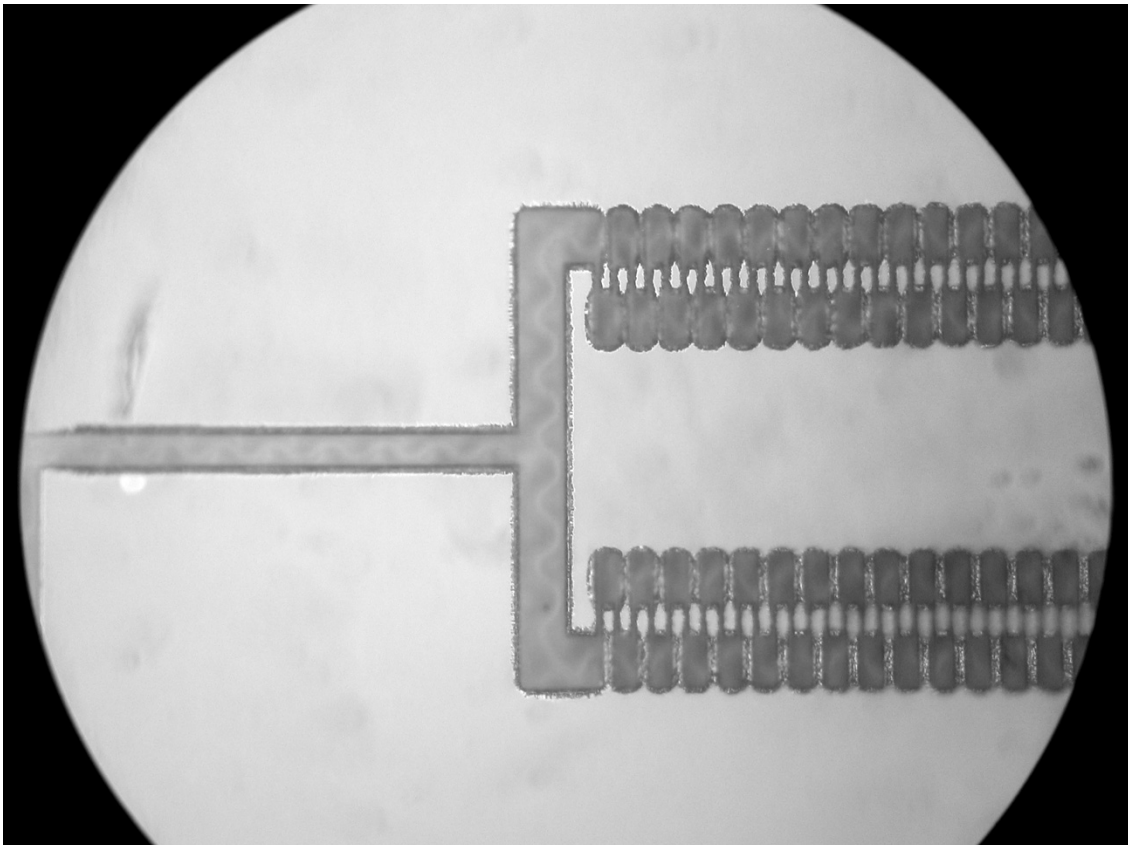footing effect observed during long over-etches.

Figure 6. Footing effect. This wafer was prepared by bonding an SOI Wafer to another wafer, then removing the handle layer. The bottom side of the device layer is visible here through the buried oxide, which is seen as wavy lines where the buried oxide membrane is suspended over the trenches. The footing effect appears as the chewed-up silicon at the edges of the trenches where the bottoms of the trenches. Note that the thin beams constituting the fingers of the comb drive are undercut completely in places. Also visible is the RIE lag effect: where the comb fingers overlap the footing is less pronounced due to the shorter overetch time caused by the slower etching in the narrower trenches.

Once the wafers have been etched to the buried oxide, the wafer must be diced into individual chips using a diamond dicing saw. This saw generates a lot of silicon particles that can bond to exposed silicon or lodge inside trenches and interfere with the functioning of the fabricated parts. To prevent this, the wafers are coated with two layers of thick photoresist immediately before they are diced. First, the wafers are covered with a puddle of acetone to reduce the viscosity of the photoresist and aid it in wetting the deep, narrow trenches. As it is spun on, the photoresist displaces the acetone and leaves an even, thick (~8μm) coating over the entire surface of the wafer. Since the trenches are up to 20μm deep, the photoresist cannot fully planarize the wafer. In fact, after the application of the first layer of photoresist, the edges of the trenches closest to the center of the wafer are often left exposed. After softbaking the photoresist for 3min at 90°C, a second layer of thick photoresist is spun on, covering any exposed surfaces. The wafer is still not planarized, but it is covered completely. This is verified using an optical microscope.

Instead of cutting all the way through the wafer, the wafer is merely scribed using the dicing saw, leaving 200μm of silicon from the handle wafer to hold the chips together. This aids the cleaning steps since the photoresist removal (PRS-3000 at 80°C for 30min,

quick-dump rinse, SRD) and piranha cleaning (sink8 piranha 10min, quick-dump rinse, SRD) steps that follow dicing can be done to the entire wafer at once instead of on a chip-by-chip basis. Once cleaned, chips can be snapped off the wafer a row at a time and then snapped off the row to leave individual chips. Since the buried oxide is still attached to the device layer, the stress from the snapping-off step does not adversely affect yield. No particles from snapping have been observed either – if they are formed from the cleaving process they seem to either not stick to the chips or to be washed away during the release procedure.

The buried oxide is removed using a 10 minute wet etch consisting of 49% hydrofluoric (HF) acid with four drops of Triton-X100 (a surfactant) mixed in. The surfactant is used to keep the chip wet as it etches (bare silicon is hydrophobic) and to keep a bead of 49% HF on top of the chip during the transfer to the methanol rinse bath. As the buried oxide etches away, the chip can start to dewet. This reduces the ability of fresh fluoride ions to access deeply undercut structures, resulting in a slowdown in the etch rate as a function of structure undercut. Although it is unclear that the drying of the chip during transfer is an issue, it was determined that the Triton X-100 improved the uniformity of the timed etch and overall yield of the release process. The chips are soaked two times for approximately 1 minute each in two large Teflon dishes filled with methanol to rinse away the HF. Afterwards, they are dried using a Tousimis carbon dioxide critical-point dryer (CPD). The CO_2 flow rates of each step of the CPD process are adjusted before each use to ensure that the chips are not agitated during the drying process.

After drying the chip, a thin chrome-gold (60Å Cr, 300Å Au) layer is thermally evaporated onto the chips. A custom aluminum chuck was fabricated in the EECS machine shop that holds up to 8 chips upside-down above the chimney. This Cr/Au layer reduces the electrical resistance of the fabricated structures from the original 50-500Ω/square (corresponding to 1-10Ω-cm resistivity and a 20μm device layer thickness) to approximately 33Ω/square while eliminating stiction between the released parts and the handle wafer through some as-yet undetermined mechanism. Leading hypotheses are a) the stiction is mainly caused by electrostatic attraction between the device layer and handle wafer and the metal layer aids charge transfer when the two layers do come into contact, b) the silicon dioxide layer that is blamed for the permanent adhesion of parts suffering from stiction cannot form due to the presence of the metal, and c) the metal surface is sufficiently rough to prevent short-range adhesive forces from causing the two surfaces to bond. Which of these hypotheses are correct (if any) is an open question. The main downside to the use of the Cr/Au layer is that since the two metals are deposited using physically separated evaporation boats shadowing effects can cause the gold to be deposited directly onto the silicon. Without the diffusion layer to prevent the gold from migrating into the silicon, the gold can coalesce and wick into the silicon at high temperatures, forming a gold-silicon eutectic that liquefies at a temperature lower than either the gold or the silicon would individually. These high temperatures can be developed during device operation by running sufficient electrical current through the fabricated structures. (Figure 7) shows the effects of the formation of this gold-silicon eutectic under localized heating.
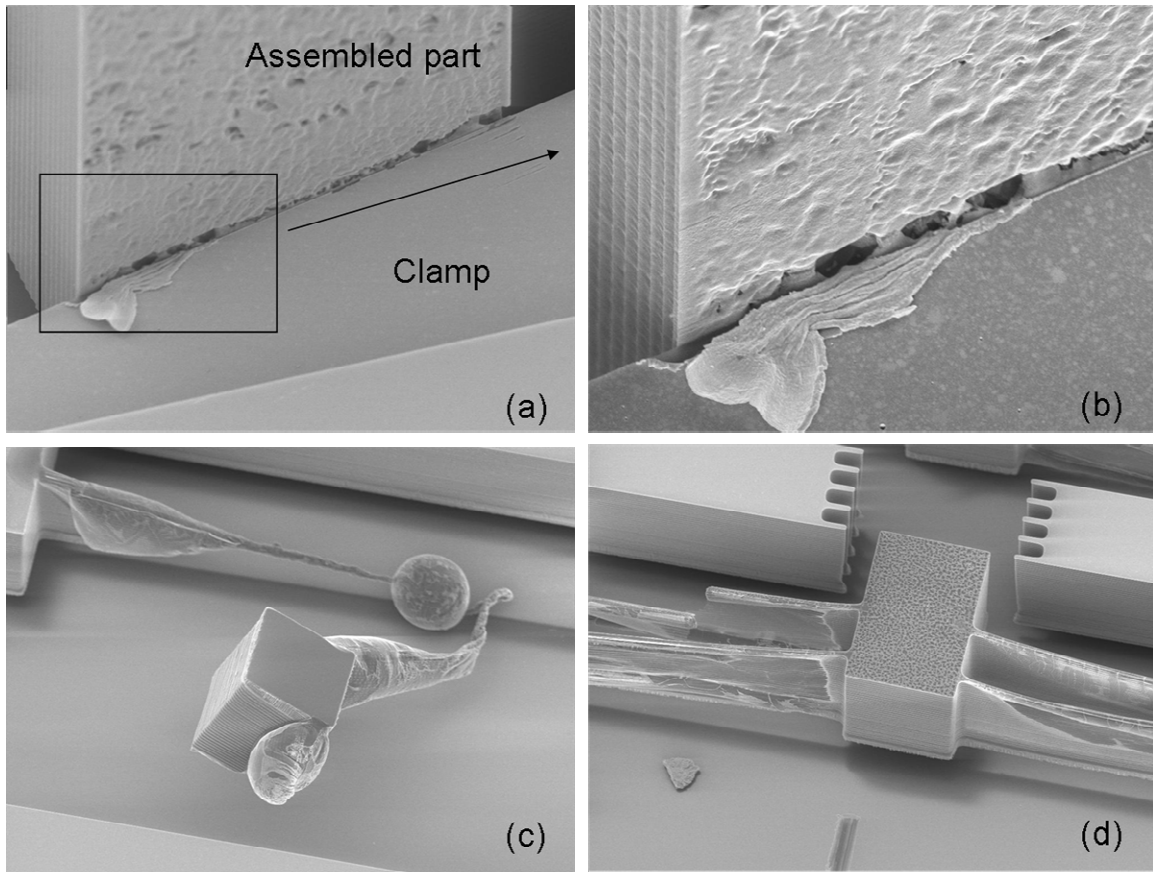
Figure 7. Formation of Gold-Silicon Eutectic. (a) Interface between assembled part and clamp after current is passed through. (b) Close-up of interface showing diffusion of gold away from interface. (c) An example of reflow of the gold-silicon eutectic. In this case the high aspect-ratio beams of the test structure completely reflowed, with surface tension causing the beams to take on a rounded appearance. (d) Thermal actuator test structure showing partial melting and flow of eutectic formed along thin suspension beams.

3. TOOLING FOR MICROASSEMBLY: THE ORTHO-GRIPPER

Once parts are fabricated as described in section 2, tools are required to pick them up and assemble them together. An ideal interface between the macro and micro world must grip the part firmly, be strong enough to easily break the tether holding the part to the chip, reliably rotate the part by exactly 90 degrees, and release the part once it has been placed in its socket. There must also be an easy way to position this tool within the assembly workspace.

The most essential element of the assembly process is the ability to reliably rotate the parts 90° out of plane. Zyvex (Skidmore, Ellis et al. 2003) and Dechev et al (Dechev et al. 2004) use a macro-scale rotation stage to perform this rotation. There are several drawbacks to this approach. The first is that it is difficult to calibrate the rotation such that the rotation axis is aligned with the part. If this is not accurately done, the rotation misalignment couples into both the horizontal and vertical axes of the robot. This can make it difficult to perform the assembly unless some sort of feedback is used to ensure accurate placement of the part. The second drawback is speed. The rotation actuator is placed in series with the Z-axis actuator in both cases cited. This means that the mass of the rotation stage must be moved by the Z-axis stage. For a macro-scale actuator, this mass will dwarf the mass of the part, dramatically reducing the speed of the robot. A collocated micromachined rotation stage suffers from neither of these problems since the rotation axis is automatically aligned with the part and the mass of the rotation stage is insignificant compared to the mass of the Z-axis actuator itself. The use of a low-mass rotation stage enables the use of high-speed, high-precision motors such as voice coil

actuators – these allow actuation over centimeter distances with nanometer precision and millisecond response times (BEI Kimco Inc. 2002).

Devices at the micro-scale often operate using different principles than their macro-scale counterparts. Macro-scale rotation stages generally employ either a ball bearing or bushing to allow the rotating part to rotate around the desired axis of motion without moving in any other direction. At the micro-scale, these types of pivots are notoriously difficult to engineer. Lack of lubrication, tolerances on the same scale of the dimensions of the rotating stage, and the need for multiple structural layers make simple downward scaling of macro-scale designs undesirable.

The ortho-tweezer (Thompson and Fearing 2001) is a device better suited to the micro-scale that can simultaneously grasp and rotate parts using only two single degree-of-freedom fingers (Figure 8). The ortho-tweezer consists of two fingers oriented at 90 degrees to one another (the "ortho" part of ortho-tweezer). Each of the fingers moves laterally. Forces on the order of milliNewtons can be applied, and with its integrated force gauges, the ortho-tweezer can sense forces in the tens of micronewtons. Fearing's group demonstrated manipulation primitives such as using the integrated force gauges to locate parts, grasping and ungrasping (not trivial at the micro-scale!) a part without rotating it, and rotating it both in-plane and out of plane utilizing an external fixture. They have used these manipulation primitives to demonstrate pick and place assembly of structures down to 50μm on a side. The ortho-tweezers they demonstrated were built from macro-scale components. However, they were designed with scaling to the micro-

scale in mind. They form the basis for an ideal micromanipulator for our assembly process.
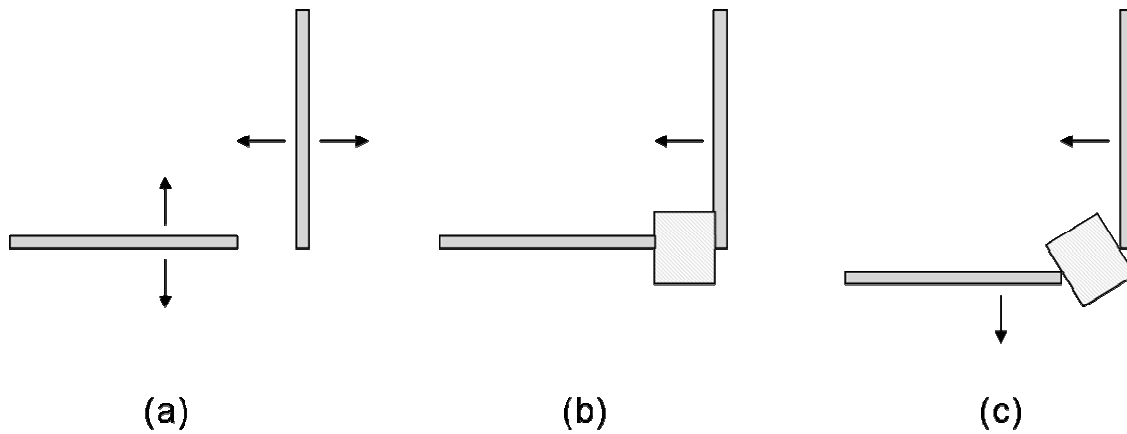
Figure 8. Ortho-tweezers consist of two fingers oriented orthogonally to one another that can each move with one degree of freedom (a). The ortho-tweezers are shown grasping an object (b), and rotating the object (c).

3.1. Micro-scale Orthogripper Design

The micro-tooling developed here is fabricated in the same single-mask SOI process described above and consists of a simplified version of the ortho-tweezers described above. This assembly process only requires the tool to perform one well-defined motion, so instead of the actively controlled ortho-tweezers we can use a purely passive compliant mechanism with no need for force feedback or active control. Because it is not as complex or capable as the general-purpose ortho-tweezer, the passive version was re-dubbed the "ortho-gripper".

The ortho-gripper has five phases of operation: 1) gripping the part, 2) breaking the tether holding the part to the substrate, 3) lifting and rotating the part out of the plane, 4)

27

carrying the part to its destination and placing it into a socket, and 5) un-gripping the part. In order to successfully perform each phase of operation, the forces acting on the part must be carefully controlled. In parts (1) and (2), the gripping forces must be sufficient to avoid slipping while the part is manipulated still attached to the substrate. In part (3), the springs must always generate a net torque in the correct direction while the energy stored in the springs is being released. In part (4), the gripping force must be enough to avoid slipping as the part is being slid into the socket. Finally, in part (5), the gripping force of the socket must be sufficiently greater than the gripper that the gripper can slide off the part without dislodging it from the socket.

The most interesting portion of the ortho-grippers to design is the rotation motion. At this point, the grippers have a secure grip on the part, one of the springs has been pre-loaded against one of the faces of the part (henceforth called the "preloaded finger"), and the tether has been broken (Figure 9). All that is left to do is pick up the part, rotate it, and carry it to the socket. In these phases, the grippers are only interacting with the part, not the tethers, not the sockets. The tethers and sockets can be designed later around the capabilities of the gripper.
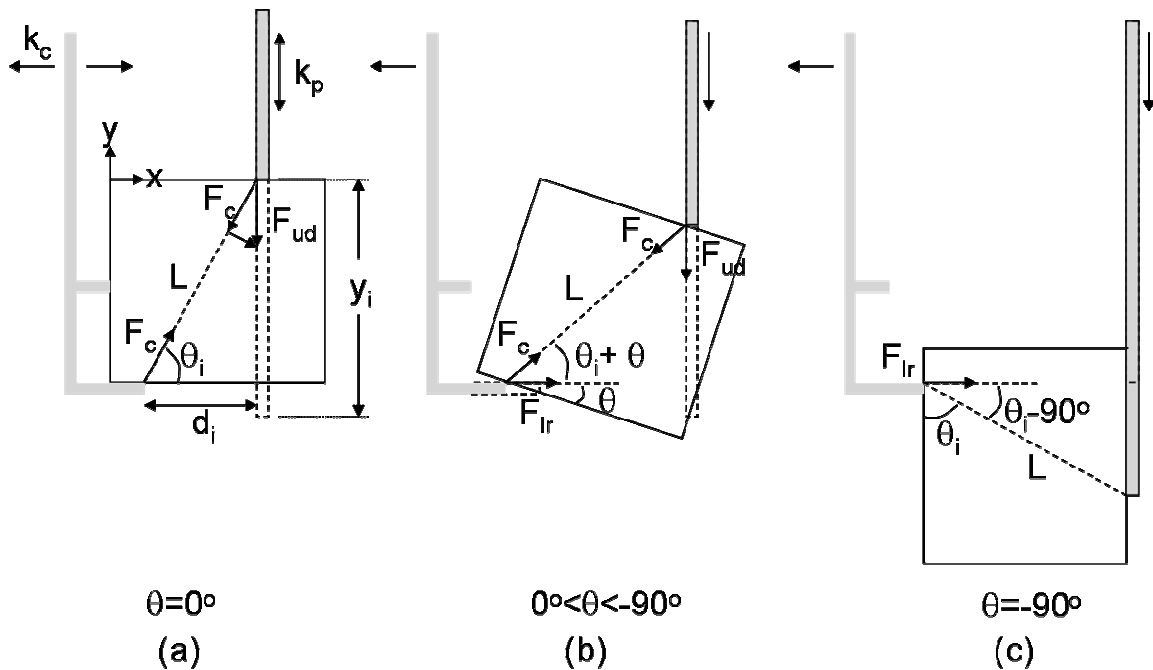
Figure 9. Ortho-gripper Free Body Diagram: Clockwise rotation means θ ranges from 0 -> -90 degrees. The contact angle θ_i the fingers make with the part must be greater than 90°- θ_{ns} = ~77 degrees for the non-slip condition to hold. Inversion points: when $\theta=-\theta_i$, the signs of the torques applied by each finger flip. Also, depending on the value of y_i, at some point in the rotation the preloaded (up/down) finger can pass through its rest position and from that point on requires an applied force to continue the rotation. This force is supplied by the energy stored in the clamping (left/right) spring.

There are two main design goals for the ortho-gripper's rotation motion. First, the grippers must not slip when gripping the part. Second, the parts must generate a net torque in the desired direction of rotation and stably hold the part in the fully rotated position. The non-slip criterion is met when the friction force causing the grippers to stick to the surface of the part exceeds the tangential force exerted by the grippers. This takes the form of a cone of acceptable angles (Figure 10) in which the force can be applied before the grippers start to slip. This is derived from the set of conditions where:
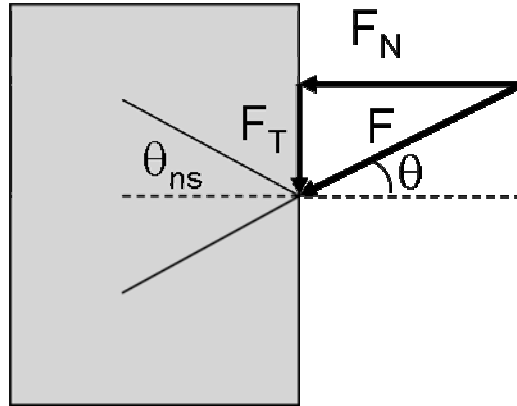
Figure 10. Application of a force where θ< θ_{ns} (the "friction cone") results in a fixed, non-slipping contact. θ>θ_{ns} results in a sliding contact. For contact between two silicon surfaces, qns is usually greater than 17 degrees.

$$F_T < F_{friction} = \mu F_N$$

$$\Rightarrow F \sin\left(\theta_{ns}\right) < \mu F \cos\left(\theta_{ns}\right) \quad\quad\quad (3.1)$$

$$\therefore \theta_{ns} < \arctan\left(\mu\right)$$

For silicon, whose friction coefficient μ ranges from ~0.3-1 (Lumbantobing and Komvopoulos 2005), the result is a worst case friction cone of roughly 17 degrees. This sets the minimum value of θ_i. Consequently, the maximum initial separation between the fingers is $d_i < \mu t_{part}$, where μ is the friction coefficient and t_{part} is the thickness of the part being grasped (for definition of d_i, see free body diagram in (Figure 9(a)).

One other aspect of the no-slip condition is that the nature of the contact between the fingers and the part changes after the full rotation has occurred. No longer are the fingers just two line contacts; now the entire side of the pre-loaded finger rests against the side of the part. The pre-loaded finger no longer serves as a pivot line, instead it provides a

plane constraint and prevents the part from rotating further. At this point, any force

applied by the pre-loaded spring is oriented tangentially to the surface being gripped.

Loading normal to the surface is provided by the other spring, and this is the only force

that prevents the pre-load finger from slipping. This brings the second no-slip criterion:

at the end of rotation, the gripper should not slip on the part. In order for this to occur,

the vector sum of the force provided by the pre-load finger and the force provided by the

clamping finger must lie within the friction cone derived above. This results in the

relationship:

$$F_T < F_{friction} = \mu F_N$$

$$\Rightarrow F_{preload} < \mu F_{clamp} \qquad (3.2)$$

$$\therefore \left. \frac{F_{preload}}{F_{clamp}} \right|_{\theta=-90^o} < \mu$$

This provides a second constraint on the ratio of the preload force to the clamping force,

only this time at the other end of the range of motion. This can be addressed

independently from the initial no-slip condition by adjusting the preload distance (or the

spring constant, though this can impact other design parameters discussed later), since the

amount of tangential force exerted by the preload finger at the end of travel is related to

the amount the preload spring is compressed or extended by after rotating the full 90°.

The second design goal is to ensure that the ortho-grippers generate a net torque in the

correct direction at all times. There are three parameters that can be used to ensure this:

the ratio of the spring constants of the two fingers' suspensions, the amount of

displacement of the pre-loaded finger, and the initial angle between the contact points of

the two fingers. Both the initial displacement of the preload finger and the initial angle of

the contact points were determined by different aspects of the no-slip criterion, leaving

just the ratio of the spring constants as a free variable. By plotting the torque as a

function of the rotation angle θ (Appendix B:), it can be ensured that the rotation will not

stall out before a full 90° has been achieved.

Each finger generates a force at an angle to the line connecting the two finger tips. This

generates a net torque on the part. Since the part is rigid, the torque generated by each

finger will be calculated by projecting the force vector onto a line extended out from the

opposite fingertip. The length of this moment arm varies continuously during the

rotation, and can be different for each finger since the forces applied by the fingers are

mutually orthogonal (again, see free body diagram, Figure 9). For the preloaded finger,

this moment arm length $\underline{L_{mp}}$ is:

$$L_{mp}(\theta) = L\cos(\theta_i + \theta) \tag{3.3}$$

And for the clamping finger, this moment arm length L_{mc} is:

$$L_{mc}(\theta) = L\sin(\theta_i + \theta) \tag{3.4}$$

Where L is the (constant) straight-line distance between the contact points of the two fingers, θ_i is the initial angle between the two fingers, and the variable θ is the amount the part has rotated.

This varying moment arm, combined with the changing amount of force applied by the two fingers as they transfer strain energy to one another, results in a system whose behavior is difficult to predict without careful analysis. For each angle, the separation of the gripper fingers is calculated and the force applied by each finger is determined based on the known spring constants of the suspensions supporting each finger. These separation distances are given by

$$F_p(\theta) = k_p y(\theta) = k_p \left(y_i + L\left(\sin(\theta_i + \theta) - \sin(\theta_i)\right)\right) \quad (3.5)$$

and

$$F_c(\theta) = k_c x(\theta) = k_c L\left[\cos(\theta_i) - \cos(\theta_i + \theta)\right] \quad (3.6)$$

Based on the geometry, the moment each arm applies around the pivot point supplied by the contact point of the opposite finger can be calculated. These two moments oppose one another, with the clamping finger providing a counterclockwise (positive) torque and the preload finger contributing a clockwise (negative) torque. Substituting in the forces and moment arms derived above, we get a net torque $\tau(\theta)$ of:

$$\tau(\theta) = F_c(\theta) L_{mc}(\theta) - F_p(\theta) L_{mp}(\theta) \tag{3.7}$$

If this net torque is negative (providing clockwise rotation) for all values of θ from 0° through -90°, the ortho-gripper will work as required to orient the part for placement.

The rotation requirement determines the ratio of the spring constants of the suspensions of the fingers. This ratio is determined numerically by plotting $\tau(\theta)$ and choosing a spring constant ratio that ensures a negative value for all θ. But the magnitude of the spring constants cannot be determined simply by the rotation requirement. To determine how strong these springs need to be, the required clamping force of the gripper after rotation can be evaluated. The gripper must not drop the part after rotation, and the stronger the grip, the less likely the grip will be broken by external disturbances such as gravity and acceleration of the gripping arm. The direction the grip is weakest is to that of a force tangential to the preload finger, where the part can slide out of the grippers without causing the fingers to separate. This can take the form of gravity or other acceleration pulling the part straight down out of the gripper, or a rotating torque shearing the part out of the gripper. Any force or torque other than the purely tangential slipping will pry the fingers apart, resulting in an increase in the clamping force holding the part. The friction from the clamping force must exceed the acceleration force that will exceed the no-slip condition derived above. The ability to resist the torque generated by an unbalanced part experiencing acceleration is less important since this torque just results in rotation of the part around the gripper. This rotation can be corrected by using an external jig to re-align the part before assembly or through additional design features on

34

the part that prevent it from rotating. Lateral motion however runs the risk of the part

sliding out from between the grippers and being dropped.

The no-slip condition under clamping can then be modified to add the following

constraint:

$$F_T \ll F_N$$

$$Nmg \ll \mu F_{clamp} \qquad\qquad (3.8)$$

$$F_{clamp} \gg \frac{Nmg}{\mu}$$

Where m is the mass of the part, g is the gravitational constant, F_{clamp} is the clamping

force normal to the direction of acceleration, μ is the friction coefficient of the part-

gripper interface, and N is the number of gravities of acceleration the part is experiencing.

To illustrate the magnitude of the required clamping force, a typical part might have a

volume of 500μm x 50μm x 20μm or 500e-15m^3. With silicon's density of 2330kg/m^3

and worst-case friction coefficient of 0.3, this gives a weight of 0.1nN and a minimum

clamping force under 1g of acceleration of 40nN. Therefore, accelerations of over 1000g

can be sustained without slipping with only a few tens of micronewtons of clamping

force.

Up until now, it has been assumed that the orthogripper's two fingers will initially grip

the top of the part with one finger and, using a hook-shaped finger, grip the bottom of the

part. This approach works just fine, but it suffers a couple of drawbacks. First, the tolerance on the thickness of the device layer (+/- 5μm) means that the orthogripper must be able to operate while accommodating parts with widely varying dimensions. In addition, because the hook finger needs to slide underneath the part, there needs to be either a very thick buried oxide to leave enough room between the device layer and handle wafer, or some sort of backside etch so that there is no handle wafer at all underneath the part.

These two problems can be eliminated completely if the orthogripper grips the side surfaces instead of the top and bottom surfaces. In this case, the part thickness variation becomes negligible (mask undercut from etching changes the width of the part by less than 1μm) and because the orthogripper never needs to slide underneath the part, the thickness of the buried oxide does not matter. An electron micrograph of an orthogripper design that grips the sidewalls of the part is shown in (Figure 11). The analysis of the orthogripper remains the same as above. The biggest difference between the sidewall-gripping orthogripper and the top- and bottom-gripping orthogripper is the coefficient of friction. Because the sidewall surfaces are rougher than the top/bottom surfaces, the coefficient of friction is expected to be lower. This is due to the decreased effect of adhesion associated with the reduction in real contact area. From rudimentary experiments conducted with structures coated with a self-assembled monolayer (SAM) the coefficient of friction has been found to be approximately 1. However, a worst-case coefficient of friction of 0.3 (Lumbantobing and Komvopoulos 2005) is used in all design

calculations. It is expected that the coefficient will be greater than this in most cases,

leading to improved resistance to slipping and a stronger grip overall.
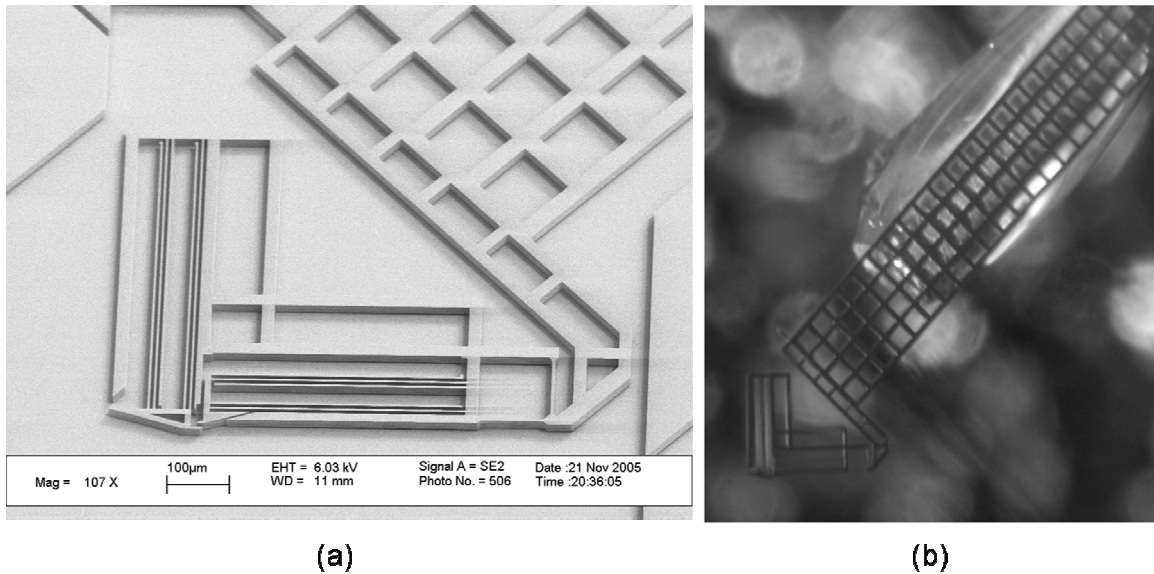

Figure 11. Sidewall-gripping Orthogripper. (a) SEM image of sidewall-gripping Orthogripper before being mounted on probe tip. (b) optical micrograph of mounted Orthogripper

Appendix B: Ortho-Gripper Design Worksheet for MathCAD contains a worksheet

showing the design of a candidate orthogripper. Based on this design, an orthogripper

was designed and built in the process described at the beginning of this chapter. This

orthogripper reliably grasps, rotates and ungrasps the part. The one area it has difficulties

is in constraining the part from rotating parallel to the substrate. Because the part is just

grasped on its sidewalls (the top and bottom surfaces after rotation), nothing prevents the

part from rotating between the two fingers. This does not pose a problem during

assembly since the part is easily rotated back. However, if automated assembly is to be

considered with this gripper, some method of guaranteeing orientation must be

developed. Possibilities consist of extra fingers on the gripper that prevent in-plane

rotation, outcroppings on the part that would interfere with the gripper if rotation occurred, or jigs on the chip that can be used to re-orient a rotated part before insertion.

Another alternative for automation is to get rid of the orthogripper entirely. If the part is fabricated on a different chip (the "source" chip) from the socket it is to be assembled into (the "target" chip), then the chips themselves can be mounted orthogonally to one another. In this case, the need to rotate the part prior to assembly is eliminated and a much simpler passive gripper can be used to grip the part. Although this was proposed early on in the development of this process (and some tooling was even made to support this assembly procedure), the idea was not pursued very far. With the benefit of hindsight, this assembly technique seems very appealing due to its simplicity, the increased robustness of the tooling that could be used, and the direct path to automated assembly that would be possible. Future work in this area is expected to greatly improve the reliability and speed of the assembly process.

3.2. Interfacing the Orthogripper with the Macro-scale World

One end of the orthogripper is used to interface with the micro-scale world, while the other end needs to interface with the macro-scale world. The orthogrippers are attached to a three-axis micromanipulator using a centimeter-scale submount. Currently, the orthogripper design includes a 5mm-long tang which is used to provide an attachment point for a tungsten probe tip. The tang is held in place using 6 tethers that attach it to the device layer of the chip. "Amazing Goop"-brand silicone adhesive is used to affix the probe tip to the tang.

The procedure for attaching the probe tip to the tang of the orthogripper is as follows. First, a freshly-released chip is mounted under the microscope and held down with either double-sided scotch tape or just a vacuum. Fresh silicone adhesive is dispensed onto a piece of smooth cleanroom tape. The probe tip is grasped with a pair of tweezers and dipped blunt-end first into the silicone and smoothly withdrawn, leaving a thin (~1mm-thick) coating over the last few millimeters of the probe tip. If a tail of adhesive is present, it is removed by wiping the end of the probe tip on the tape. After aligning the probe tip with the tang under the microscope (using the lowest-power objective to maximize depth of focus), the probe tip is lowered onto the tang by hand. Once contact is made, the probe tip is lifted straight upwards, breaking the tethers. The adhesive grips the tang strongly enough that the tethers are easily broken. The orthogrippers can now be used immediately, or the adhesive can be given time to cure. This curing time is essential if the grippers will be used in a vacuum chamber such as the electron microscope discussed in the next section, but is not necessary if the grippers will not be used in vacuum. When ready to use, the probe tip is inserted into the micromanipulator's probe tip receptacle, sharp end first. (Figure 12) shows an orthogripper mounted on a micromanipulator at a manual probe station.
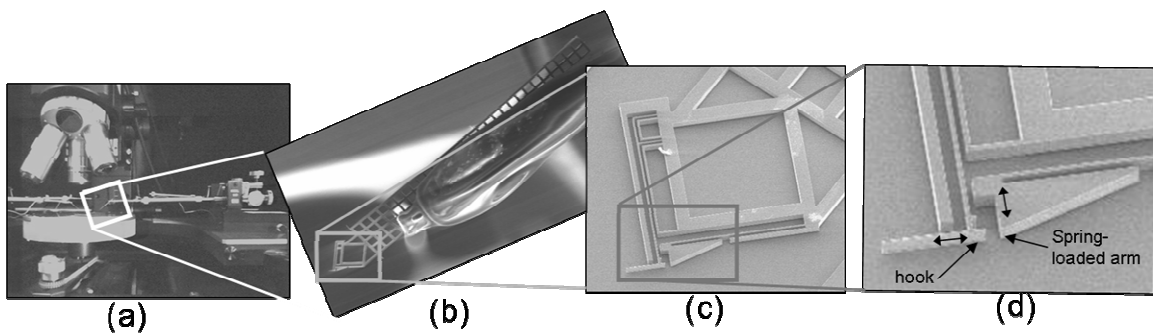

(a)   (b)   (c)   (d)

Figure 12. Orthogripper Mounted on Micromanipulator at Probe Station. (a) Probe station with 3-axis micromanipulator showing location of probe-mounted ortho-gripper (b) SEM of ortho-gripper mounted on end of probe tip (c) close-up of ortho-gripper end effector (d) section of ortho-gripper shown schematically in free-body diagram (Figure 9)

Mounted in this fashion, the orthogrippers have been tested on 3-axis micromanipulators at two different manual probe stations equipped with optical microscopes. Using the micromanipulators on the manual probe station, minimum steps of approximately 1 μm are possible in the X, Y, and Z axis. A 45 degree prisms with the hypotenuse coated with aluminum folds the optical path by 90 degrees and presents the user with a side view of the chip. By tilting the mirror slightly, a perspective view is obtained, greatly easing the task of microassembly (Figure 13).
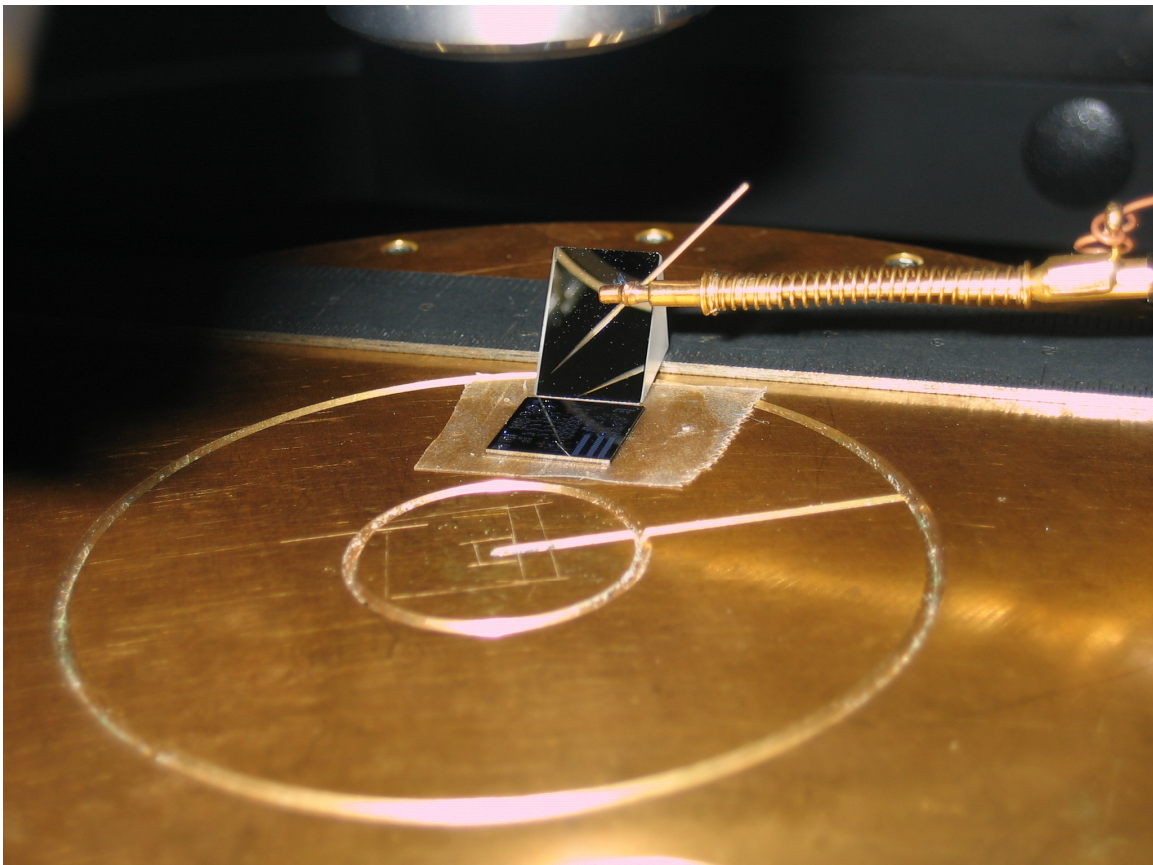
Figure 13. Use of 45 degree mirror to obtain side view of chip. A typical setup showing the objective of the microscope positioned over a chip being tested. A probe tip by itself or with an orthogripper mounted on it (not shown) is attached to a 3-axis micromanipulator . A 45 degree mirror provides a side view of the device under test when the user moves the microscope to a position above the mirror and changes focus appropriately. Photo by V Subramaniam

Figure 14 shows frame captures taken during a successful microassembly using a manual probe station. This video was captured using the 45 degree mirror to get a side view of the orthogripper as it picked up a part. To set up this video, a part was grasped and the tethers holding it in place were broken. The microscope was then moved to view the orthogripper from the side. The four captured frames show the rotation of the part as it is lifted from the substrate.

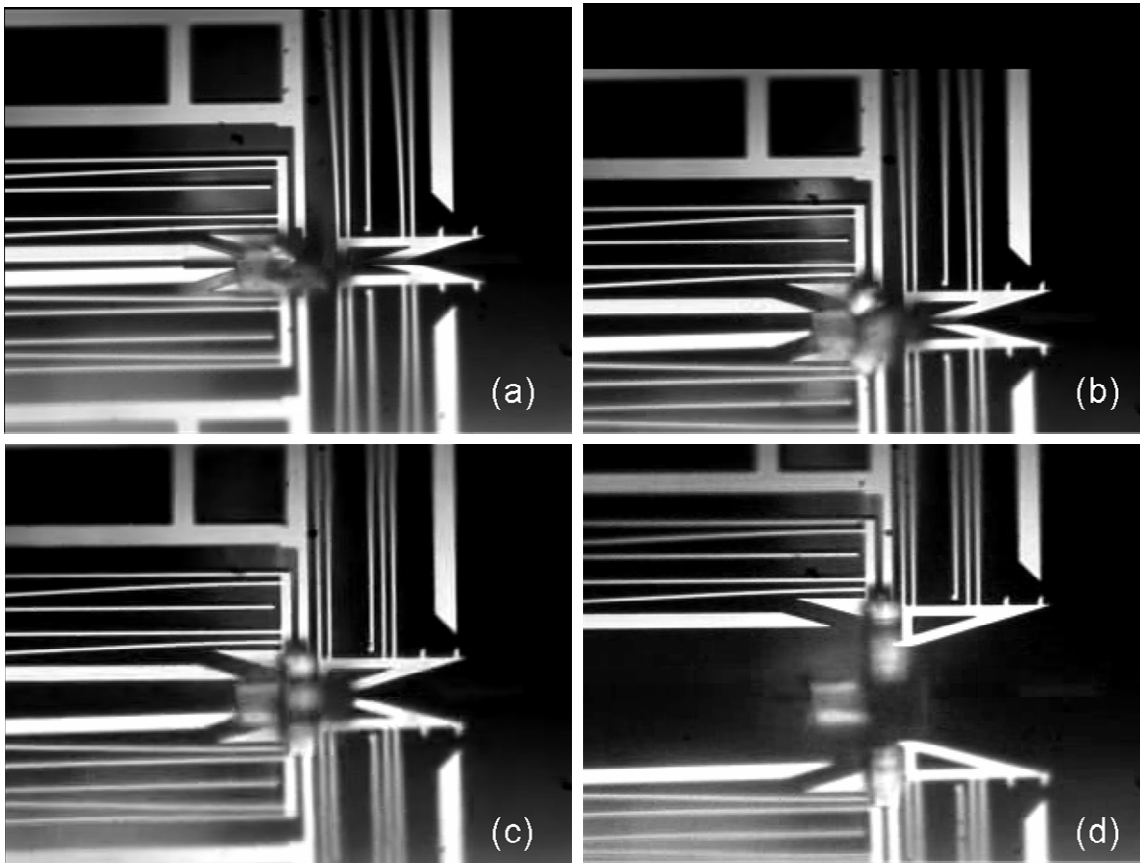**Figure 14. Side view of ortho-gripper rotating and lifting part. The ortho-gripper and its reflection from the mirror-like chip surface are visible. A sidewall-gripping orthogripper has grasped a part and the tether holding the part in place has been broken. Almost immediately, the part begins to rotate (a). As the orthogripper is lifted from the substrate (note the reflection from the mirror-like chip surface), the part continues to rotate (b) and (c). Once the part is no longer touching the chip, the rotation completes and the part is held in the rotated position as it is carried away by the orthogripper (d).**

3.3. Tethers

Tethers are used to connect parts to the chip until they are assembled. In addition to keeping the parts from sticking to the substrate during the release process, they also hold the parts in well-defined locations to aid the assembly process. But the most important job of a tether is to break reliably, with little applied force, and without generating particles.

Two different methods for breaking tethers were attempted. The first was to use the orthogripper itself to generate the force to break the tether. After gripping the part, the orthogripper would be moved sideways. The tethers were designed to break with just 200μN of applied force. Unfortunately, this force was also sufficient to substantially deflect the orthogripper. When the tether broke, the strain energy stored during the deflection of the orthogripper would often cause the part to be flung away, or at the very least, shift the grip of the orthogripper causing rotation to not work as well as it should.

The second method used a probe tip to break the tether. This can be problematic since the forces that can be applied by the probe tip can be so large at the micro-scale. By incorporating a microfabricated tether-breaking structure into the chip (Figure 15), this

probe tip force could be applied in a controlled manner. This prevents the part from

moving as the tether breaks by holding the part with a spring-loaded clip until the probe

tip can be removed. This approach eliminated all yield problems associated with tether

breaking and is incorporated in almost all new designs.
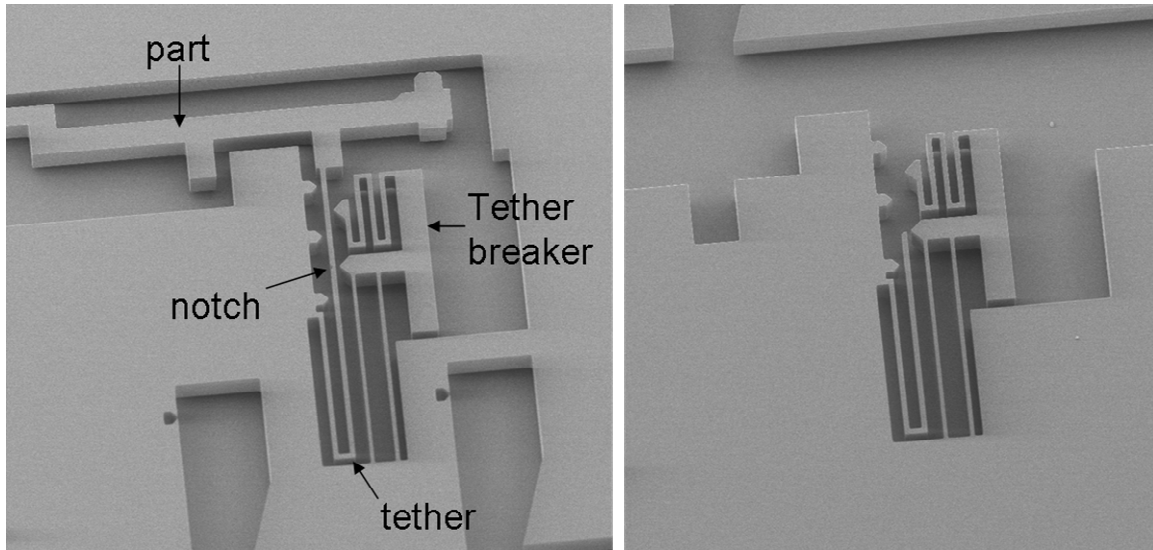
Figure 15. Tether with integrated tether-breaker. The tether breaker is pushed with a probe tip. The tip of the breaker and the tip at the end of the serpentine spring touch the tether at the same time. The spring-loaded tip holds the part in place while the tether is being broken and prevents the part from flying away until the breaker tip is disengaged from the tether.

Tethers are made to fracture at a certain point by incorporating a stress concentrating

notch. They are designed to have a stress of approximately 2.5GPa at the stress

concentrator (fracture stress of single crystal silicon is ~2GPa) when a pushing force of

100μN is applied to them.

4. MICRO-SCALE FASTENERS

Once the part has been removed from its initial location and rotated 90 degrees out of plane, it is ready to be carried to its final location and affixed to the substrate. An ideal connection will fuse the part and the substrate into the functional equivalent of a solid piece of silicon. The joined part should simultaneously provide a strong mechanical connection and a low electrical and thermal resistance. This connection is analogous to permanent macro-scale connectors like nails, screws, or welds, not temporary methods of attaching two parts like a latch.

There are two main principles by which a connector may operate: constraints and friction. An example of a constraint-based connector is a clevis pin, where the pin is constrained from sliding out of a hole by a cotter pin or snap ring. The pin cannot be removed from the hole unless the cotter pin/snap ring fails by shearing off. Friction does not play a part in the restraint of the pin. Instead, the strength of the material is what prevents the pin from sliding out. The downside of this type of connector is that there is usually a fair amount of play in the system unless some sort of preload mechanism is employed. This is especially true in MEMS devices like those described here where the layer thickness is not well controlled. In the case where there is a pre-load mechanism, the connector acts like a relatively weak linear spring with a threshold: above a certain displacement, the motional constraint prevents any further motion.

An example of a friction-based connector is a screw. The screw holds two parts together by generating a clamping force. This clamping action results in a frictional force that prevents the parts from sliding against one another. The screw cannot handle very much shear load by itself; instead it is the friction force of the parts rubbing together that supports the shear load. In addition, this clamping force results in the friction that prevents the screw from backing out of its hole.

Both types of connectors were built in this process. The constraint-based connector, dubbed the "snap lock" was developed first, and the friction-based connector, dubbed the "clamp" was developed next. This section discusses the design and testing results from these connectors. The connectors were evaluated based primarily on their ability to resist pull-out forces in various directions, though controlling the required insertion force is also very important due to its influence on assembly yield.

4.1. Snap Lock Connector

The snap lock uses two interlocking C-shaped pieces of silicon oriented at 90 degrees to one another that constrain rotation and translation in all directions except the axis in which they are slid together. A latch mechanism prevents the two pieces from sliding apart in that direction (Figure 16). This design is purely constraint-based: even if all pieces were free of friction, the connector would still operate.
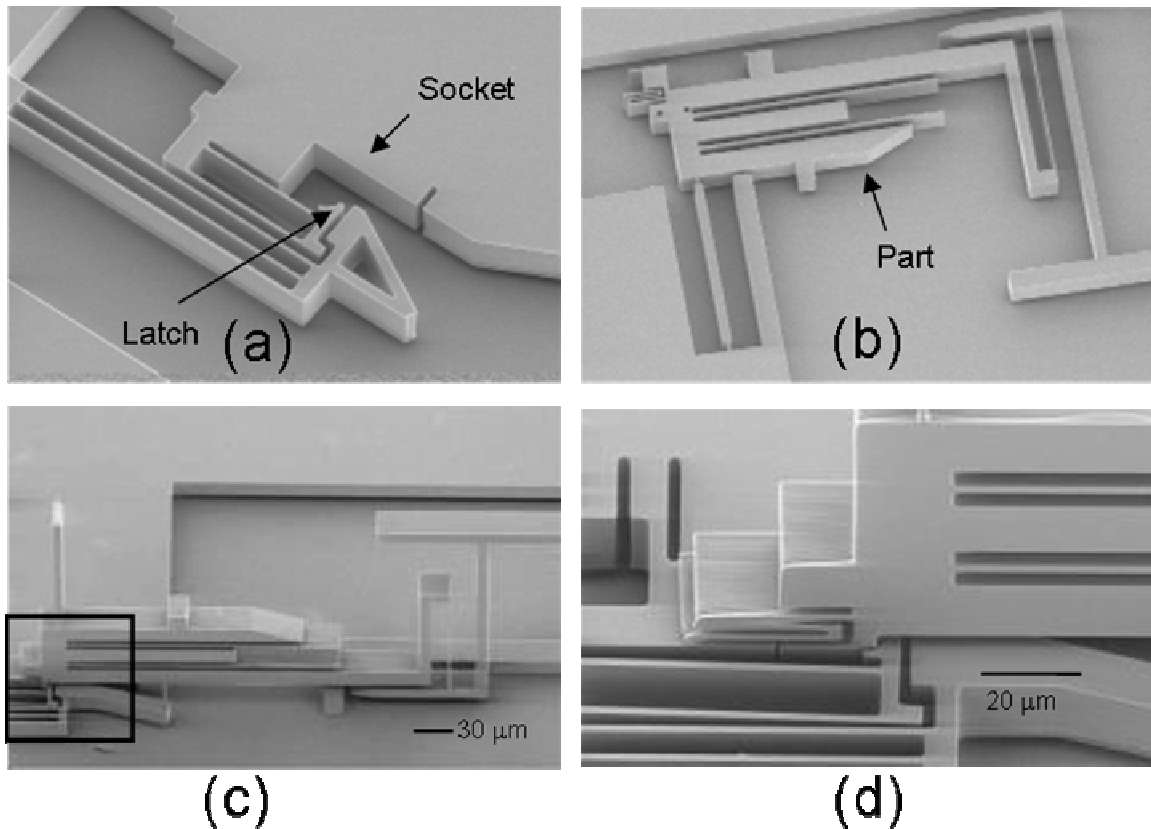
Figure 16. Snap Lock socket and mating part. (a) Snaplock socket, (b) part as fabricated (before assembly), (c) assembled part, (d) detail of assembled part

This design suffers from two primary problems. The first stems from its limited ability to handle parts of different thicknesses. The result is that there is a large amount of play between the two assembled parts. This play is accommodated using a preload spring that presses the parts together. The primary mechanism by which the parts are constrained is still the shape of the parts, however, the spring keeps the two parts in intimate contact when there is no external force applied to the connector. The second problem comes from incomplete constraints in the part as it was actually implemented: instead of being two interlocking "C"s, it consists of one "C" shape (the assembled part) and one "L" shape (the connector). Because of this, it does not resist simultaneous rotation and

46

translation as well as it would if it were actually shaped like a "C". With certain combinations of rotation and translation, it was possible to disengage the snap lock and pull the assembled part out of the connector.

In its implementation, this "snap lock" mechanism used a spring to latch behind the C-shaped part after it was inserted, preventing it from pulling back out of the connector. This spring was designed to accommodate variations in the thickness of the assembled part up to +/-5μm. This means that the spring should displace up to 10μm. Two methods were employed to achieve this displacement: a ramp that the part pushes against as it engages that opens the spring, and a hook that allows the spring to be pulled open using a probe tip on a micromanipulator stage. In the first case, opening the spring requires an increase in insertion force, up to several tens of micronewtons depending on the slope of the ramp. This force opposes the clamping force of the ortho-gripper, and can result in the part slipping or rotating during insertion. Once inserted, this snap lock takes up the gap in the connector that is required to accommodate the inevitable variations in part thickness by pushing the part against the corner of the "L". Because it is the same spring holding the part in place that also opposes insertion of the part, the total force clamping the part in place can only be a small amount larger than the insertion force. In the second case, where a probe tip opens the spring with an (for all practical purposes) unlimited external force, the clamping force can be much greater. But in both cases, the snap lock connector provides at most a couple of tens of micronewtons of preload to counteract play up to 1/2 of the width of the part. Although this preload is not the force that keeps

the part from being pulled out of the connector, it is partly responsible for helping to transmit forces through the connection.

As implemented, the snap lock does not provide much resistance to direct pull-out forces either. The promise was that in order for the part to pull out of the snap lock, the silicon itself would have to break. However, testing shows that the pullout force is roughly 65μN (value obtained using a spring-based force gauge) while the actuators can apply over 100μN. During operation of the electrostatically-actuated single degree-of-freedom rotation stages (described in later sections), the parts were seen to shift inside the socket. Clearly, this type of connector is only a very loose approximation of a solid chunk of silicon.

4.2. Clamp Connector

The second type of connector is the friction-based connector (Figure 17). As implemented, this consists of a fixed C-shaped piece. Inside the C-shaped piece, there is a triangular piece that clamps against the part. This triangular clamp is mounted on a parallel-bar flexure that moves to adjust for the range of thicknesses of the inserted parts. Next to the triangular clamp there is a wedge-shaped piece of silicon mounted on a parallel-bar flexure (moving orthogonally to the other parallel-bar flexure) that can slide between the clamp and the fixed C-shaped piece. This wedge is used to force the clamp against the inserted part and serves two purposes: to adjust for the tolerance in the part thickness and to provide mechanical advantage to increase the clamping force applied to

the part. This wedge is connected to an L-shaped handle that is pulled with a

micromanipulator-mounted probe tip. The clamping force exerted on the inserted part is

limited by the force used to engage the wedge and has been shown to be limited by the

strength of the silicon beams connecting the wedge to the handle. In the current design,

this is the fracture strength of the three 4μm-wide, 20μm (+/-5μm)-thick, 50μm-long

beams. The connector was intentionally designed to fail at these beams in order to avoid

accidental breakage of critical parts if too much force was applied while engaging the

clamp.
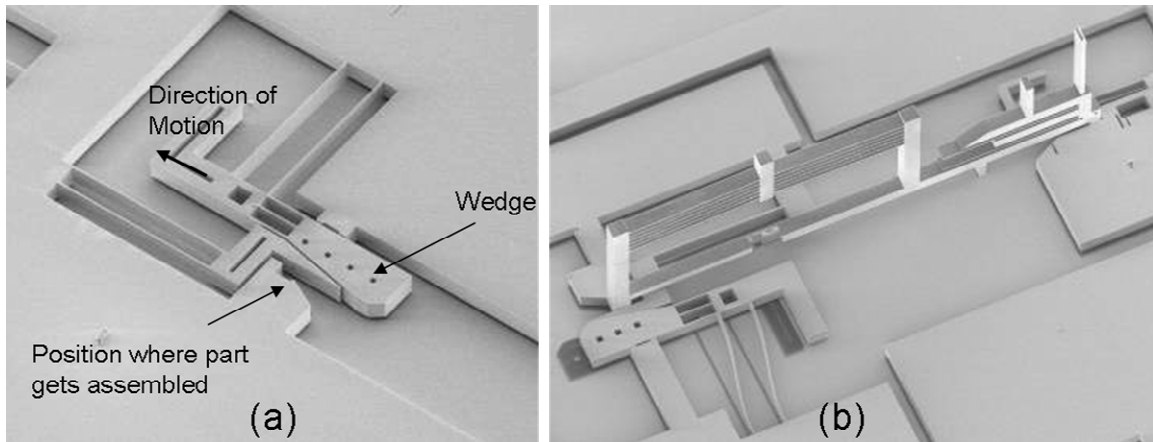
Figure 17. Clamp socket. (a) clamp as fabricated, (b) clamp after assembly. Shadow from gold deposition shows location of wedge before it is engaged.

Assembly of the part into the clamp starts the same way as the snaplock: the part is

removed from the substrate, rotated 90°, and carried to the clamp location. It is then

lowered to the substrate and slid into the connector. Once the part is roughly in position,

the wedge is slid into place with a probe tip, clamping the part firmly into the connector.

The ortho-gripper can now be easily removed from the part since the connector clamping

force is on the scale of several millinewtons while the orthogripper only grips with a force on the order of a few tens of micronewtons. By separating the insertion step from the clamping step, the forces required for each step remain independent of one another and can be separated by several orders of magnitude. In addition, the light pre-load provided by the clamp before the wedge is slid into place aids in alignment of the part relative to the connector before the part is permanently emplaced.

The wedge does not latch into place; instead, stiction forces hold the wedge in place. Normally an adverse effect in MEMS design, stiction increases the coefficient of friction of the wedge/clamp system and effectively welds it into the clamped position. The forces exerted by the part are well within the friction cone of the wedge/clamp system, so once in place, normal use will not pull the wedge out. In addition, nothing prevents the design from incorporating design aspects from the non-friction-based connector, namely the interlocking C-shaped constraints that block motion in undesired directions. In fact, the improved ability of the clamp to accommodate parts of different thicknesses and the high rigidity of the assembled clamp addresses all of the snaplock's weaknesses.

4.2.1. Clamping Force

The clamp is only as effective as the force it can apply to the part. The angled surface of the wedge and the clamp transform the force applied by the probe tip into a normal force on the part. Because the engagement force applied to the wedge is almost tangential to the interface (far outside the friction cone) we can assume that the clamp and wedge do not stick together while the wedge is being pulled into the clamped position (Figure 18).

To find the normal force F_N applied to the part, two sequential vector projections are performed: first the projection of the engagement force F_E onto the vector normal to the inclined surface of the wedge is performed; this force vector is then projected again onto the vector normal to the engagement force: $F_N = F_E \sin(\theta)\cos(\theta)$ where θ is the angle of the ramp of the wedge (Figure 19). This normal force allows us to predict the ability of the clamp to resist pull-out forces. The maximum tangential pull-out force the clamp can resist before the part begins to slip will be given by $F_{slip} = \mu F_N$. In addition to the part slipping, the wedge itself can also slip. The contribution of friction F_f to the sticking force holding the clamp in place is given by the sum of the friction forces acting on the two surfaces holding the wedge in place – one parallel to the direction of motion, and the other the angled face of the wedge: $F_f = \mu F_E (\sin\theta\cos\theta + \sin\theta) = \mu F_E \sin\theta(\cos\theta + 1)$.

The optimal angle is derived by maximization of the clamping force applied to the part under the constraints that the angle must remain small enough to avoid slipping but large enough to accommodate the part thickness variation with a reasonable wedge length. Decreasing the angle increases the chance of the wedge remaining stuck in place, but results in a smaller clamping force, a longer wedge, and a longer suspension in order for the wedge to travel this larger distance. A larger angle wedge increases the chances of the wedge slipping after assembly (the normal force vector gets closer to the edge of the friction cone), but can also result in increased clamping force. An optimal value will be close to the angle derived for the non-slip condition to hold ($\sim 16^\circ$ for a worst-case silicon-silicon friction coefficient of ~ 0.3), with enough margin-of-error to stick even in the case where an applied shock would dislodge the wedge. An angle of 12.7° was used in the first implementation of the clamp. This results in a clamping force on the part of

215µN for every millinewton of engagement force and a predicted minimum pull-out force of 65µN per millinewton of engagement force. The wedge is held in place even more firmly, with a predicted pull-out force F_f of 130µN per millinewton of engagement force.
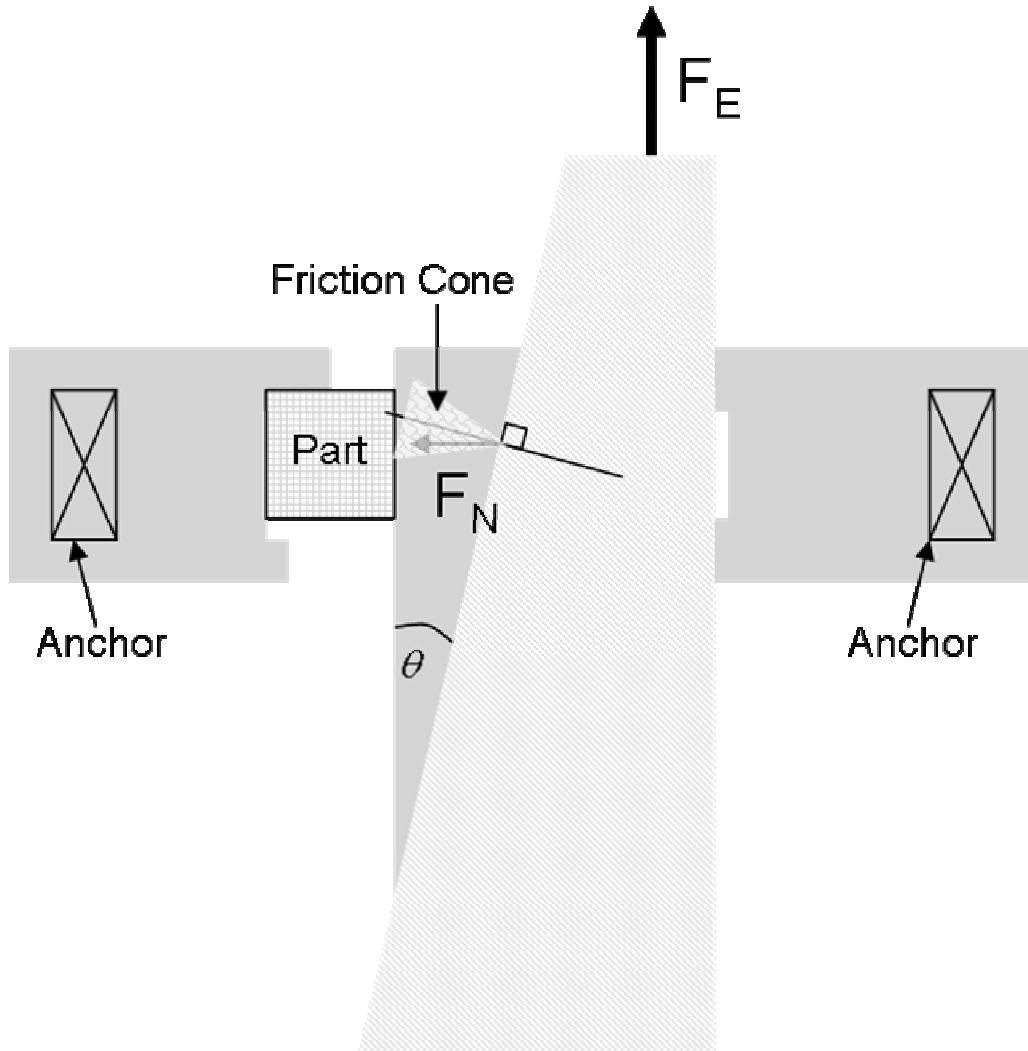
Figure 18. Clamp Socket Free Body Diagram
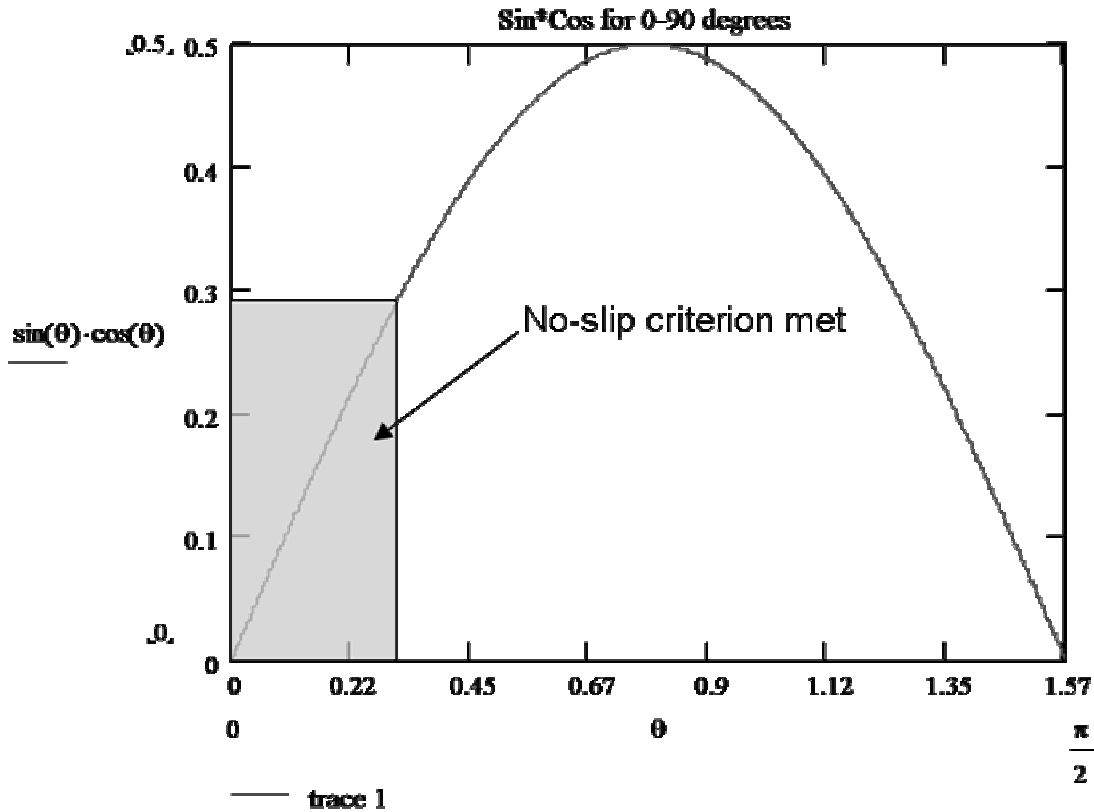
Figure 19. Clamp Wedge Design: engagement force multiplier vs theta. Clamping force is given by the engagement force multiplied by the engagement force multiplier. The range of acceptable angles that satisfy the non-slip criterion is indicated by the shaded region. This is for the worst-case situation where the friction coefficient equals 0.3. In most situations, this box will be larger.

The engagement force, limited by the breakage strength of the beams attaching the wedge to the handle the probe tip pulls on to engage the wedge, is estimated from the fracture strain of a beam in tension. The applied force at which fracture occurs is given by

$F_y = \sigma_f A$, where F_y is the applied force at which the beam yields through fracture, σ_f is the fracture stress, and A is the cross-sectional area of the beam. In (Tsuchiya 2005), the fracture stress of single crystal silicon was experimentally determined to be approximately 2GPa. With a cross-sectional area of 20μm x 5μm, this results in a yield force of approximately 200mN. If the wedge is pulled with the probe tip until this beam

breaks, the part must be pulled with a force greater than 13mN tangential to the gripping surface before it starts to slip. Although this is an overestimate on the strain the beam can accommodate (there are stress concentrators at the base of the beam and the applied load is not pure tension), the actual design as implemented uses three beams in parallel and is probably much stronger than this estimate. It was found that the part could easily resist more than 2mN of pull-out force, so the calculations above are at least the correct order of magnitude. A screen capture of an assembled part resisting a direct pull-out force of 2mN is seen in (Figure 20, Figure 21).
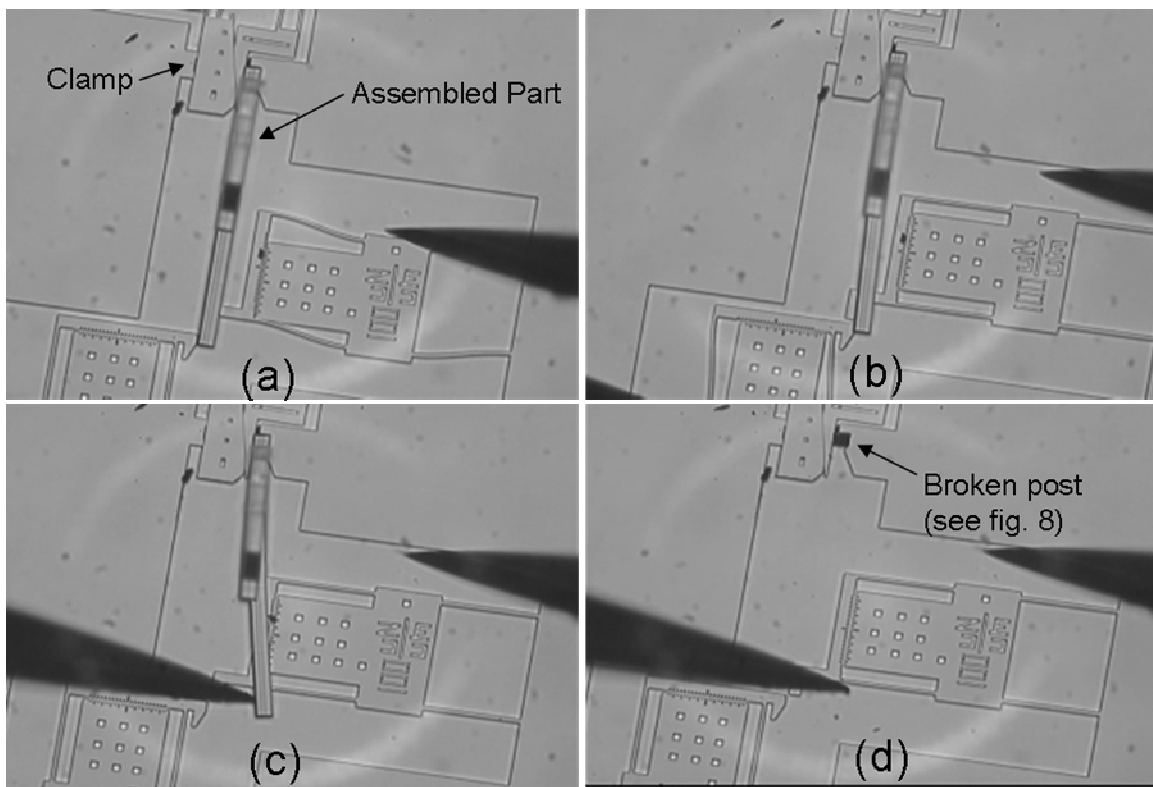

Figure 20. Pullout force test of clamp socket. (a) pulling straight out using a microfabricated force gauge. Maximum force the force gauge can apply is 2mN. (b) Pushing sideways using a force gauge, (c) pushing directly on the assembled part with a probe tip, (d) after part breaks

54
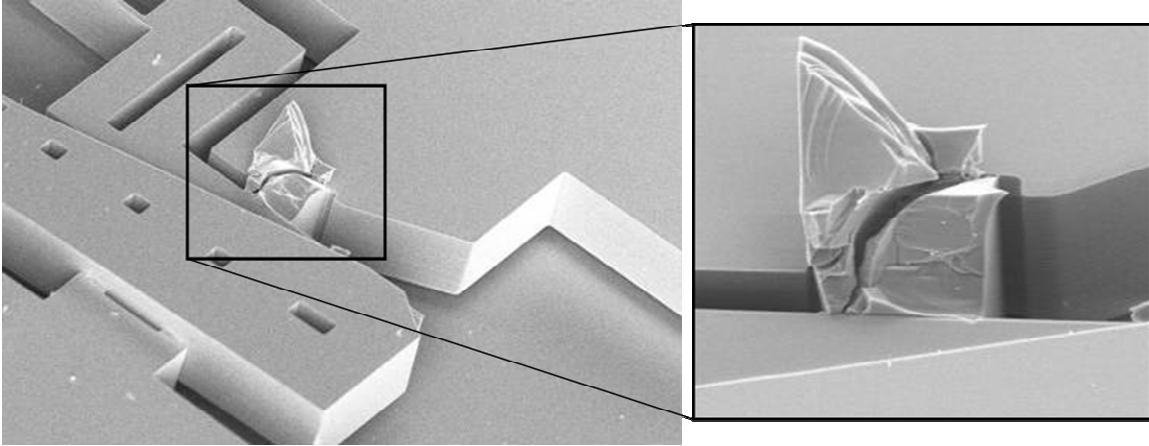
Figure 21. Post of assembled part in clamp after stress fracture

4.2.2. Moving Clamps

Clamps are useful to secure parts to anchored portions of the device layer to permanently affix them. However, it is essential to be able to make a rigid connection between parts and moving structures on the chip such as comb drives and other actuators. This requires clamps that are not anchored to the substrate. In order for the clamps to be able to move, two obstacles must be overcome. First, the clamp must be suspended in such a way that it can resist the forces involved during the engagement of the clamp's wedge. Second, the wedge must be constrained to move only in-plane; if it slides down out of plane too much, it can rub against the substrate and cause friction that interferes with the motion of the mechanism. In the current design, the first problem is addressed through the use of gap stops at the end of the range of motion of the actuator. During engagement of the wedge, the entire clamp translates until it runs into the gap stops, at which point it cannot move further. This provides the constraint necessary to keep the clamp in place while the wedge is pulled into position. The second problem is addressed by two methods: creating

a suspension for the wedge that is laterally flexible yet vertically and torsionally stiff, and ensuring the sidewalls are very close to vertical. The suspension design freely allows the wedge to move in-plane, but strongly resists any sort of vertical or twisting motion that would allow the wedge to touch the substrate. Meanwhile, the vertical sidewalls do not force the wedge out of plane like tilted sidewalls would, thereby limiting the forces directing the wedge out of plane in the first place. A surface profile scan using a white light interferometer (Figure 22) shows the worst-case out-of-plane motion of the wedge is less than 1.5μm. This is much less than the 5μm clearance between the device layer and the handle wafer. This value can be improved further by increasing the out-of-plane stiffness of the flexures supporting the wedge. This can be accomplished by adding additional flexures in parallel or increasing the distance between the existing flexures.
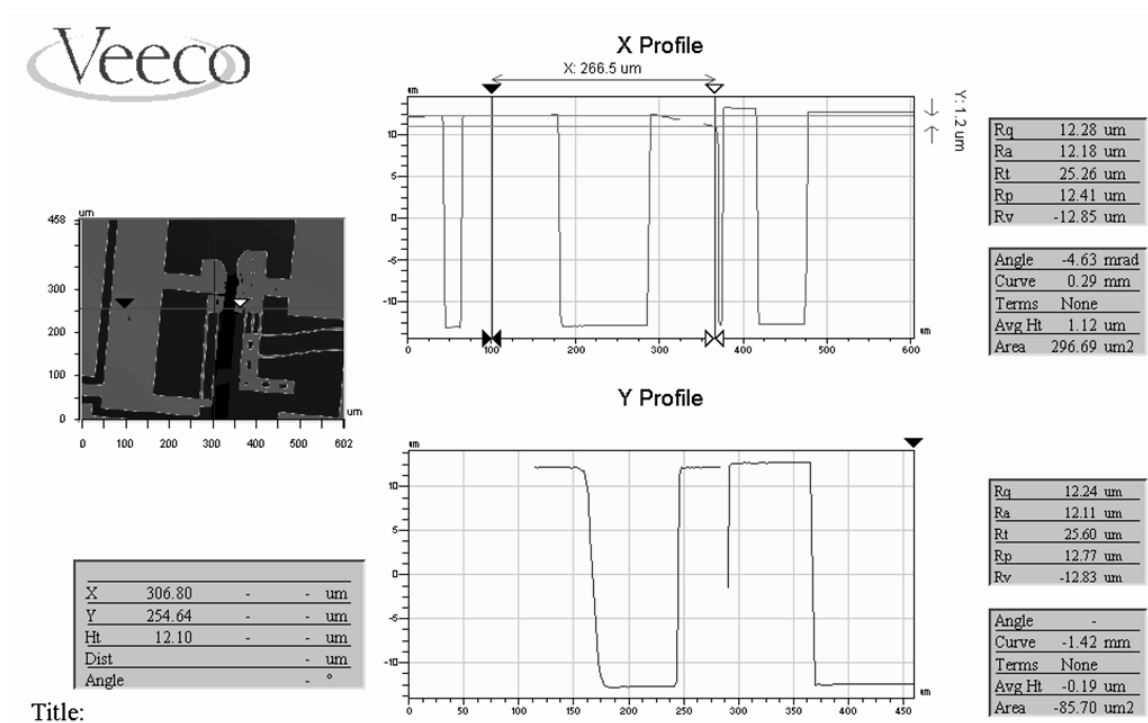
Figure 22. Surface profile of moving clamp. This was obtained using a Wyko white light interferometer

5. ASSEMBLED DEVICES

The microassembly concept is only as useful as the devices that can be made using this technology. To show that this approach is interesting from a practical as well as an academic standpoint, three types of devices were built, assembled, and characterized: 1) an assembled out-of-plane electrical interconnect, 2) a single-axis rotation mechanism, and 3) a dual-axis rotation mechanism.

5.1. Multi-level Electrical Wiring

One of the major limitations of single-mask SOI processes is the inability to make wires that cross one another without being shorted together. This can make wiring up certain structures very complicated. A point of concern when using an assembled piece of silicon to bridge two silicon traces is the electrical resistance of the contact between the part and the clamp. Because the assembled silicon is not fused into a single piece, but consists instead of two separate pieces that are pushed together by the clamp, the electrical connection might be less than ideal. Precisely how much less is an area that can be explored using a test structure.

To test this clamp resistance, an assembled interconnect was built and tested. Adjacent to it, a reference structure was built that was identical in every way, except the assembled silicon part was instead replaced with an in-plane, solid piece of silicon of the same length as the assembled part. If done correctly, this reference can be used to subtract off the resistance caused by the interface between the silicon, wirebonds, and the resistance

of the silicon itself from the total resistance of the assembled structure, giving just the resistance of the two assembled contacts. Using this approach, a contact resistance of 8.4Ω was measured.

A second generation of the contact resistance measurement structure utilizes a part with three clampable posts and a mating structure that integrates three clamps (Figure 23). The central clamp has two bond pads attached to it. The resistance of the single contact at the center of the assembled piece is what is measured. Current is run from a wirebond, through an anchored pad, into the assembled part, through the contact to another anchored pad, and out through another wirebond. Voltage is measured from the across the other two pads, ie from a wirebond to a pad, across the contact, through the other portion of the assembled part, and out a second pad and wirebond. Because no current flows through this portion of the structure, its resistances and that of the second assembled contact do not cause a drop in voltage. The only voltage drop measured is the resistance of the part of the structure where both current is flowing and voltage is being measured, ie the region of the central contact. In all these contact resistance experiments, the parts were coated with a thin seed layer of chromium (~10Å) and a thicker layer of gold (~500Å). This dramatically reduces the resistivity of all parts, the friction coefficient, and the tendency for parts to stick to the substrate. It also substantially reduces the resistance of the contact between the part and the clamp.
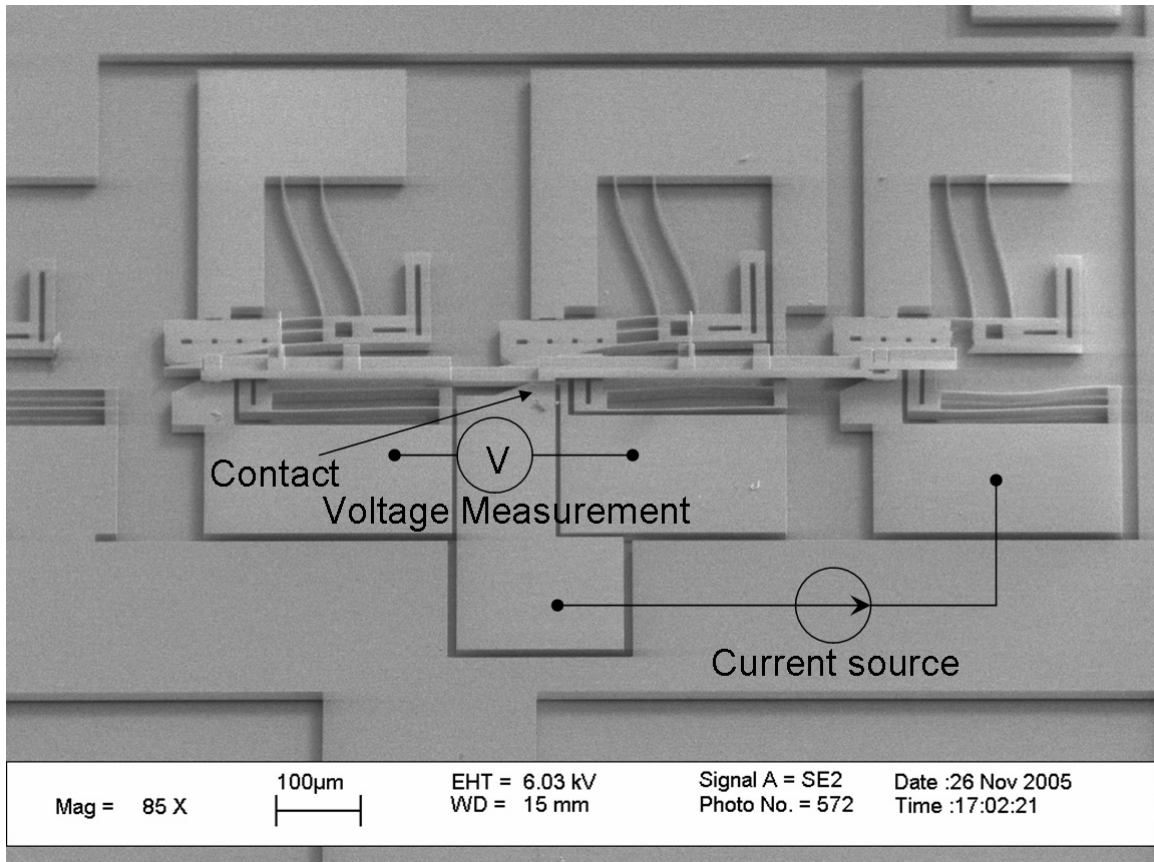
Figure 23. Contact resistance test structure. Inspired by the Van der Pauws test structure, this device cancels out all resistances except for the contact resistance between the assembled part and the device layer.

Using this structure, contact resistance values ranging between 20Ω and 50Ω were measured (after the device had been broken in). One interesting effect that was observed was a marked decrease in the contact resistance after a large current was flowed through the part. This break-in effect was observed as both a time-dependent drop in resistance and a current-dependent drop with a substantial threshold effect between 0.1mA and 1mA (this value varied substantially from run-to-run). This is attributed to localized resistive heating and the subsequent formation of a eutectic at the gold/silicon interface where the assembled part meets the sidewall of the clamp. This gold/silicon eutectic has

60

a much lower resistance than the bulk silicon and forms a solid bridge between the

assembled part and the clamp, effectively forming a weld joint between the two parts.

Electron micrographs of parts where the gold/silicon eutectic has formed are shown in

Figure 7. A plot of contact resistance vs current for a typical device is shown (Figure

24).

**Figure 24. Contact resistance vs current for a typical device. As the current through the part
increases, the contact heats up. At some threshold current, the contact resistance drops
dramatically.**

5.2. Single-axis Rotation Stage

The multi-level wiring device demonstrated the ability to electrically couple assembled parts to the substrate. Equally important is mechanical coupling between the parts and the substrate. This allows actuators still anchored to the sub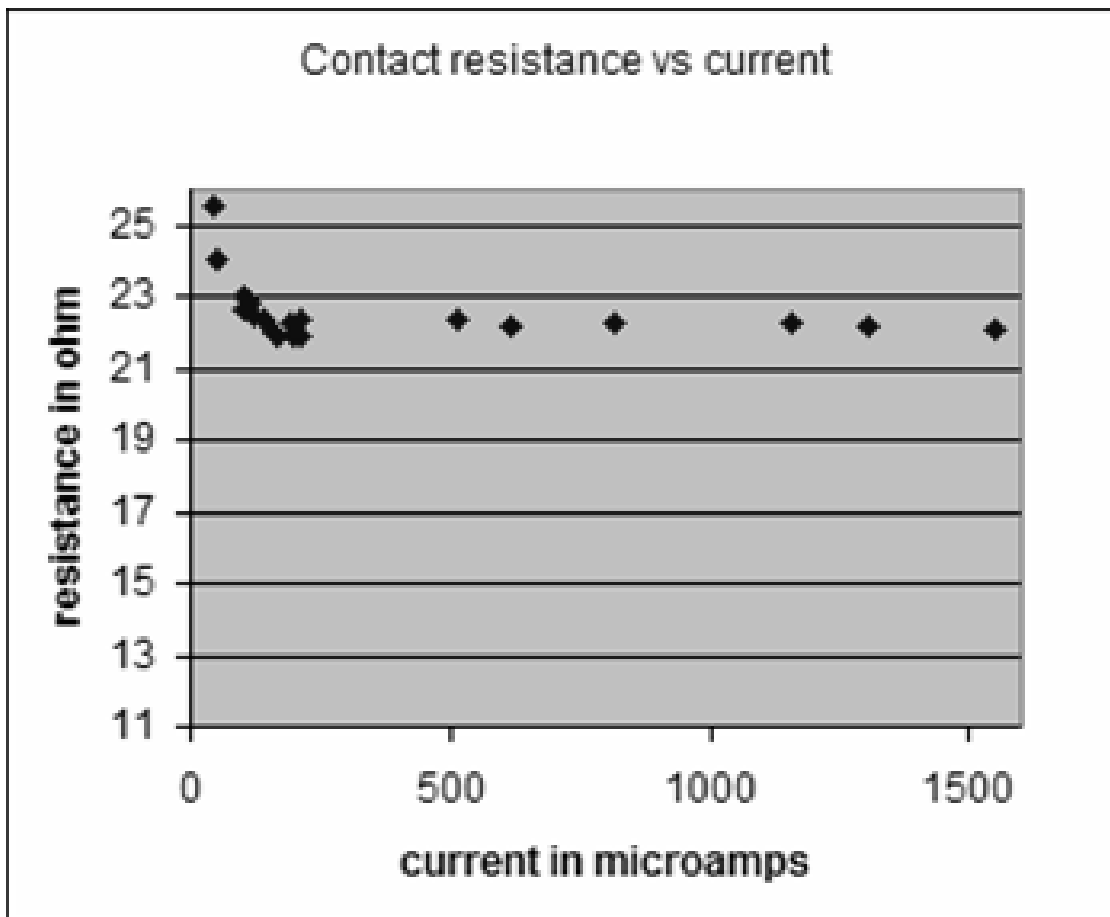strate to interact with assembled mechanisms, which is important because very few actuators can be made in a single-mask SOI process that can still function after being assembled. This limitation stems from the inability to fabricate parts that consist of mechanically connected yet electrically isolated components with just the single mask. It must be shown that the actuators don't have to undergo assembly, but can instead remain on the substrate while assembled mechanisms couple the in-plane motion of the actuators into out-of-plane directions.

The simplest useful device that can demonstrate out of plane motion using in-plane actuators is a single axis rotation stage. This type of rotating platform can be used as a mount for a reflecting surface, forming a single-axis micromirror. It could also be used in a microrobot as a very simple leg or as an aerodynamic control surface.

In its most basic form, the rotation stage consists of an anchor, a flexure, and a moment arm (Figure 25). The anchor, either a clamp or a snaplock, holds the front of the assembled part in place. The flexure is the pivot about which the stage rotates. The moment arm rests against a hook attached to the actuator and converts the force provided by the actuator into a moment applied about the end of the flexure. Different variations on this design have been built – versions with a clamp or snaplock, versions with a straight or serpentine flexure, and different lengths and designs of the moment arm.
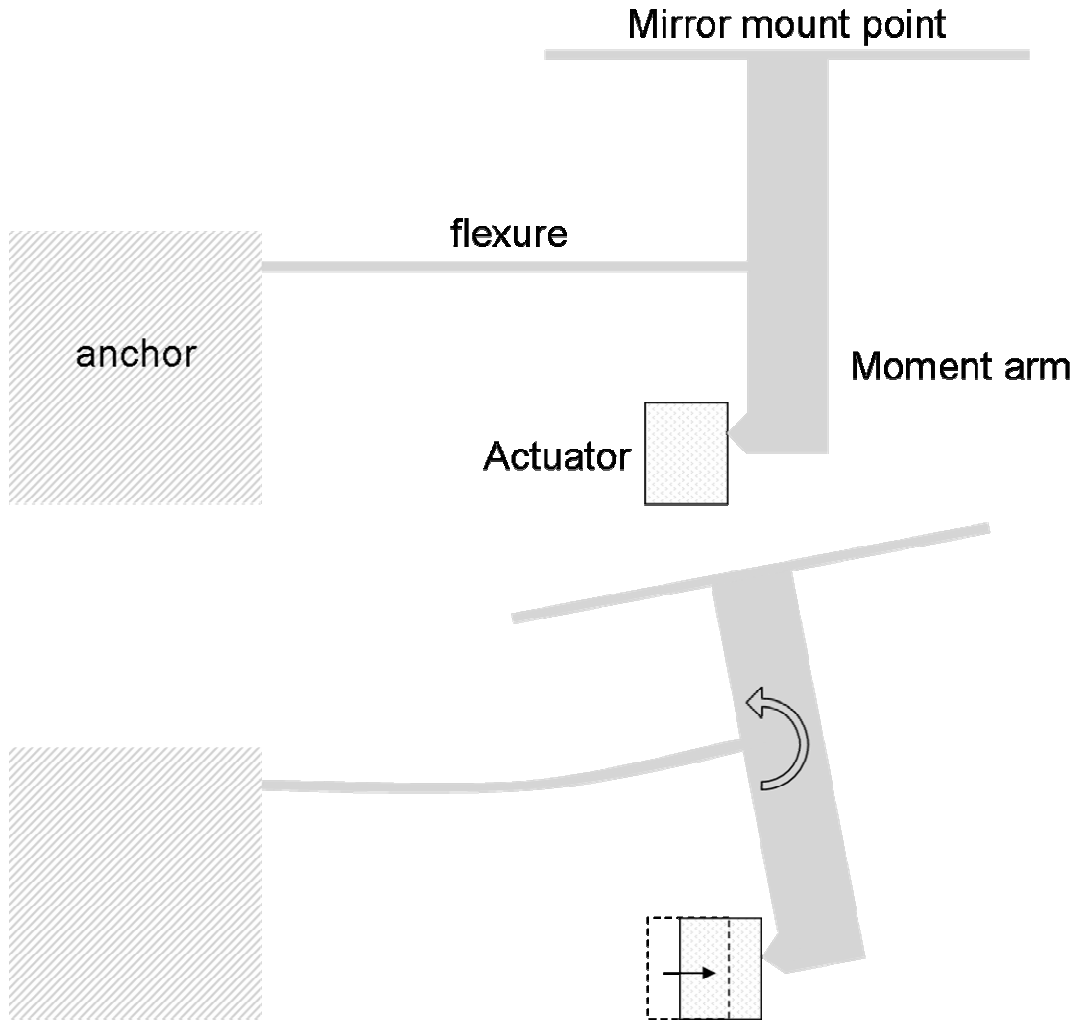
Figure 25. Single-axis Rotation Stage Concept. Pictured is a schematic of the side view of a single-axis rotation stage. In-plane motion of the actuator causes the mirror mount point to rotate.

Assuming a perfect anchor and small deflections (up to ~15°), a straight flexure can be

modeled as a cantilevered beam with a pure moment applied to it. The effects of tension

on the beam are neglected here. The angle at the tip of the beam is given by $\theta = \dfrac{LM}{EI}$,

where L is the length of the beam, the moment applied to the tip $M = Fd$ is equal to the

actuator force F times the length of the moment arm d, E is the Young's modulus of

silicon, and the second moment of area of a flexure with rectangular cross-section

$I = \dfrac{a^3 b}{12}$ where a is the dimension of the flexure in the direction that is getting flexed, and

b is the width of the flexure. Under tension, the angle reduces by a small factor.

The interface between the moment arm and the actuator is a knife-edge pivot (Figure 26). The advantage of this type of pivot over a flexural pivot is zero stiffness. Another advantage is that assembly of this type of structure is easier than for one with a flexural pivot since the end of the pivot does not need to be attached rigidly to the actuator. The main shortcoming of this device is that it is subject to slippage at large deflections. This causes problems with repeatability. However, because of the presence of the flexure, the rotation stage always returns to the same starting position. The result is a simple single degree-of-freedom transmission that converts the linear in-plane motion of the actuator to a rotational motion.
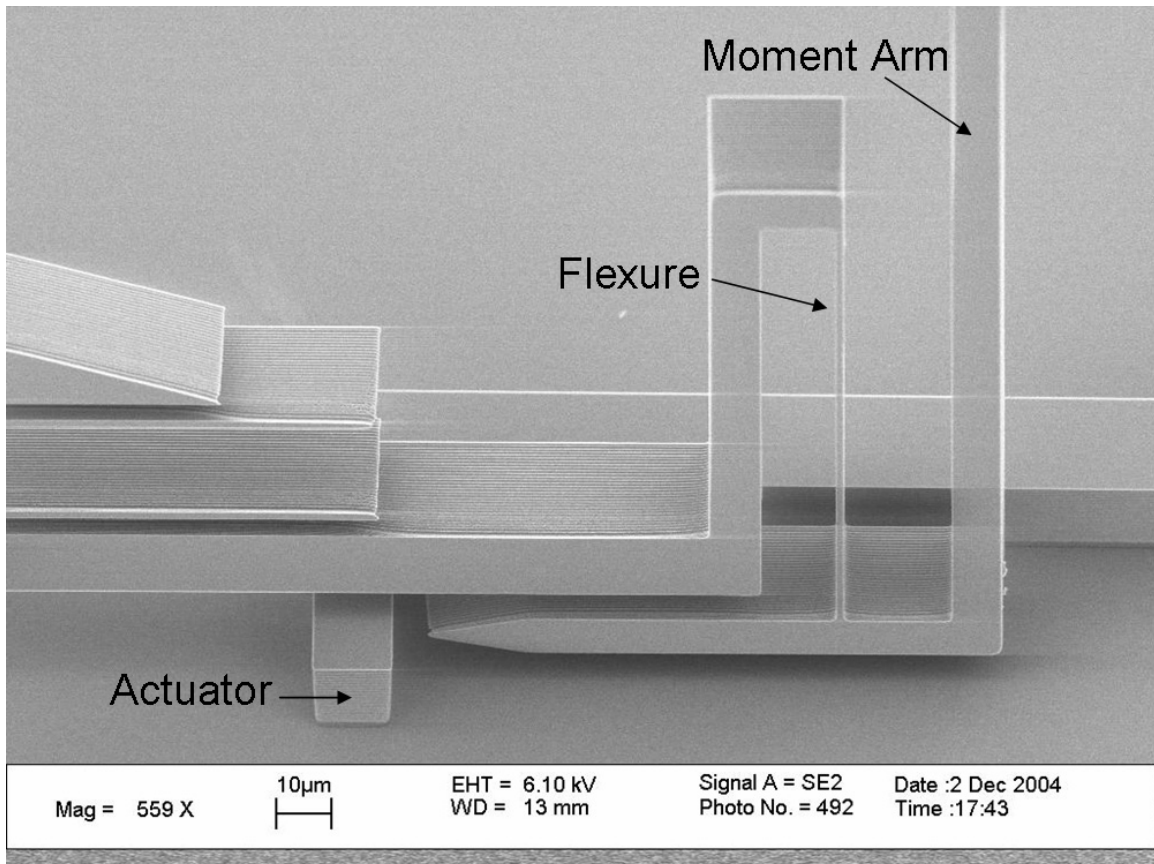
Figure 26. Detail of knife-edge pivot on single-axis rotation stage. The knife-edge pivot does not require force to rotate, but suffers from slippage and hysteresis.

Several designs have been built and tested to date (Figure 27). The largest deflection angle was achieved using the serpentine flexure design shown in Figure 27(b) which delivered 17° of static mechanical rotation. Greater rotation was prevented by the pull-in of the electrostatic actuators, and we believe that the maximum rotation angle can be increased significantly by minor design modifications such as reducing the length of the moment arm and by changing the knife edge design. A different rotation stage achieved 4.2^o of static mechanical rotation when a force of 7.9µN was applied using in-plane electrostatic actuators. The system was found to resonate at 3.75 kHz and attained a maximum
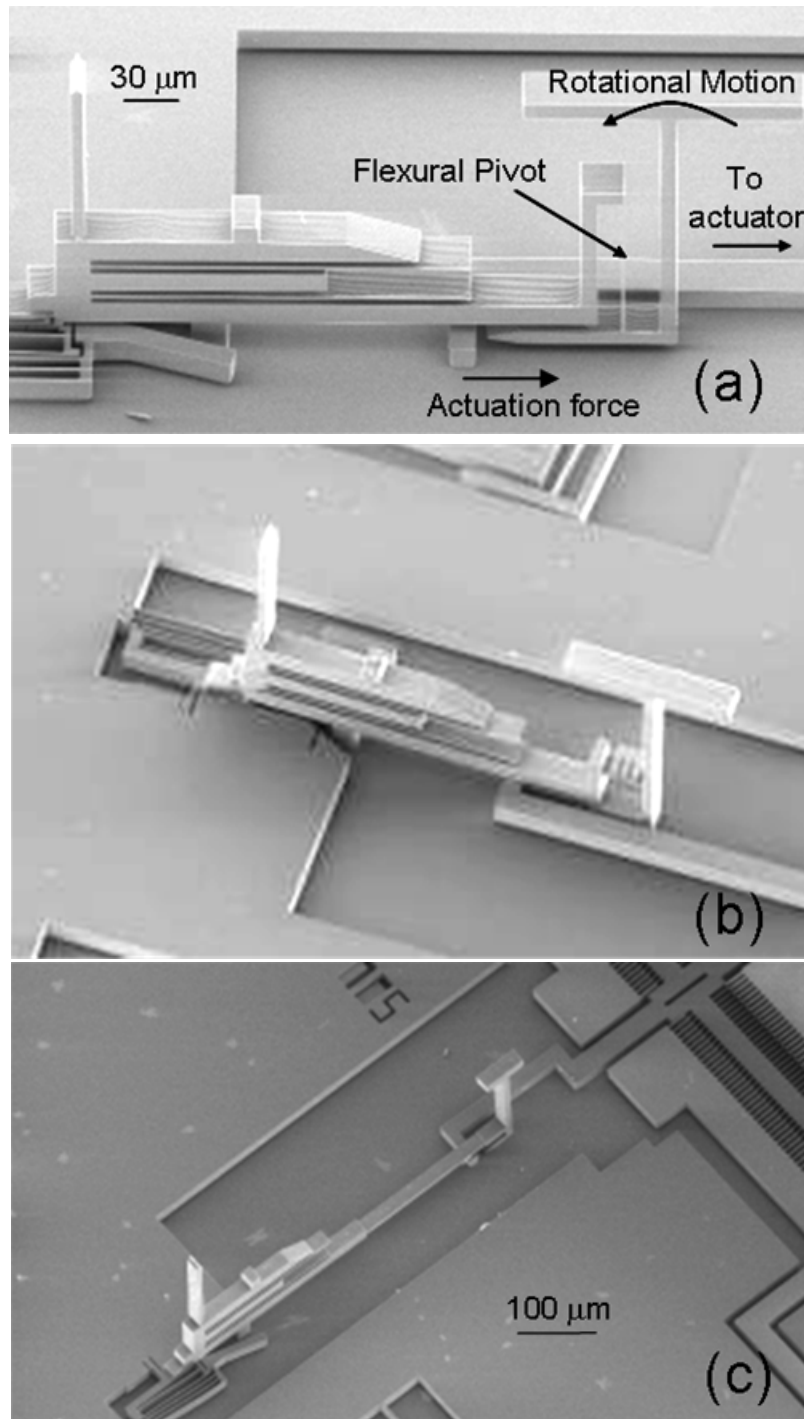
Figure 27. Electron Micrographs of Assembled Single-axis Rotation Stages. (a) Principle of operation of rotation stage: In-plane force provided by actuators is transformed into rotational motion through a flexural pivot. Actuated rotation stage in (b) achieves 17o static mechanical rotation. In (c), the interface between the assembled part and an electrostatic comb drive array is visible.

rotation of 5.7^o at resonance. Since the actuator can exert up to 100μN, these devices are not operating in an actuator force-limited regime. Therefore, with more careful design, it should be possible to extend the range of static deflection and/or resonant frequency.

5.3. Two-axis Rotation Stage

A two-axis rotation stage has also been built. This is discussed in depth in chapter 8.

6. EXTENSIONS TO THE MANUFACTURING PROCESS

Research into microassembly is motivated by the desire to build tip/tilt micromirrors and micro-robots that can walk, hop, climb, and assemble other structures. To build walking robots, it may be possible to build multi-degree-of-freedom legs in the process as described in this chapter. But there are other areas in which some change in the process could greatly improve the performance or capabilities of a robot. For instance, higher force density actuators or high-density energy storage could be integrated into the process. Or if active structures such as actuators or electronic sensors/amplifiers could be assembled as parts, the capabilities of the assembly process would be extended dramatically.

Just a few relatively simple process modifications give us many of these additional capabilities. A method for providing electrical isolation between islands of mechanically connected pieces of silicon would allow us to pick up entire actuators and assemble them as parts. It would also allow us to route multiple electrical signals through a single assembled part. A second modification might involve patterning the handle wafer. This would decrease the mass of any robot built in this process, ease the challenge of gripping and assembling parts, and also greatly reduce stiction of released parts. Another avenue for investigation involves moving from 100mm to 150mm wafers. This opens up UC Berkeley's most modern lithography tools, and allows us to scale the parts down by approximately a factor of 5 in all dimensions. Finally, if CMOS could be assembled as parts, active devices such as microcontrollers, sensors, power converters, or radios could

be integrated directly into the MEMS devices without the need to build the CMOS and MEMS on the same substrate.

6.1. Electrical Isolation with Mechanical Connection: Nitride Trenches

In (Brosnihan, Bustillo et al. 1997), (Bellew, Hollar et al. 2003), (Hollar et al. 2003), (Sarajlic, Berenschot et al. 2005), structures built in the device layer of Silicon on Insulator wafers were mechanically connected yet electrically isolated by way of trenches that were refilled with silicon nitride and undoped polysilicon. This was used in (Warneke, Scott et al. 2002; Bellew et al. 2003; Hollar et al. 2003) to create electrically isolated solar cells and high-voltage electronics that could be combined in series to make high-voltage solar cell arrays. In addition, they demonstrated the use of these silicon nitride isolation trenches in moving MEMS structures to route electrical signals to electrostatic actuators embedded in the fully released device. Finally, Sarajlic and Elwenspoek (Sarajlic, Berenschot et al. 2003, 2003; Sarajlic, de Boer et al. 2004; Sarajlic et al. 2005) used silicon nitride isolation trenches that extended slightly below the device layer to create inchworm motors with several orders of magnitude greater force density (up to $14,427\frac{N}{m^2}$) than the best inchworm motors made in our lab (up to $87\frac{N}{m^2}$)(Yeh 2001). In these devices, the silicon nitride performs two functions: it is used to prevent the pull-in instability at extremely small final gap sizes without requiring extra structural silicon to stiffen the structures, and it allows both electrodes of the electrostatic actuators to be mechanically connected.

Incorporating this extension to the single-mask SOI process would allow integration of the same types of high-force electrostatic actuators, but more importantly, it would create more interesting "active" parts that could be assembled into the connectors described in the previous sections. These active parts could incorporate complete electrostatic actuators and multiple electrical traces. Potential applications include spring clips that could provide electrical connection to multiple electrical pads on an assembled chip or to provide all the electrical connections to a complex device such as an inchworm motor. Assembled parts could integrate many actuators for multi-degree-of-freedom legs. These actuators might include vertical and rotational comb drives and inchworm motors. Insulated "feet" at the ends of assembled robot legs could be used to grip conducting surfaces with electrostatic clamps. Finally, since anchors to the handle wafer would no longer be required to provide electrical isolation between different portions of the device, fully released devices could dispense entirely with the handle wafer. In this way, micro-robots might be fabricated and assembled on a chip that incorporates a handle wafer, but after assembly might break the tethers that hold them to the chip and just walk away.

6.2. Patterned Backside Etch

Patterning the handle wafer with a deep reactive ion etch also expands the capability of the process. There are three main advantages: tailoring the mass or mass distribution of the chip, providing access holes for the buried oxide etchant from the backside of the wafer to eliminate them from the device layer, and the ability to create structures that extend down from the bottom of the device layer as well as up from the top either because of assembly or in the course of their normal operation.

The patterning of the handle wafer has been used by Hollar et al (Hollar et al. 2003) to tailor the mass profile of the chip. This was used for weight reduction of his walking micro-robot. By minimizing the mass, the force requirement of the actuators was reduced, reducing the area (and therefore further reducing the mass) of the actuators. Bellew took this idea even further, tailoring the mass distribution to control the mode of gliding of silicon chips. His variable density honeycomb structures would autorotate like a maple seed by lumping mass at one end of a wing-like structure. This autorotating wing fell at a slower rate than similar wings that did not have engineered weight distributions. Finally, (Lindsay, Teasdale et al. 2001) was able to minimize the heat transfer through the handle wafer of his micro-rockets by reducing the area of the silicon in contact with the combustion chamber. By incorporation of air gaps and increase of the distance for heat to escape by creation of a meandering path through the sidewalls of his rocket chamber, the thermal gradient was decreased, reducing thermal losses. Microassembly could be used in a process similar to this to add aerodynamic control surfaces to the rocket to allow it to steer itself in flight.

Certain optical devices such as micro-mirrors and solar cells are adversely affected by the presence of etch holes in the device layer. But without etch holes, the buried oxide beneath the device layer is inaccessible. By instead etching access ports into the handle wafer, the hydrofluoric acid used to etch away the buried oxide can gain access from the backside of the wafer. This allows large, continuous sheets of device layer to be released, picked up, and assembled.

Finally, many of these large area devices (especially micromirrors) can benefit from being able to tilt or move vertically. Currently, this motion is limited by the presence of the handle wafer. This is also useful for devices that move in-plane, but suffer from too much damping (eg. Couette flow underneath comb drives), stiction to the substrate, or electrostatic levitation forces ruining the symmetry of the actuator (eg. some vibratory rate gyroscopes). Additionally, some assembled parts benefit from free access to the underside of the device layer. For instance, assembled thrust bearings designed to keep the shuttles of inchworm motors traveling in a straight line would benefit from the ability to constrain motion both upwards and downwards from the plane of the device layer.

6.3. Device Scaling

Another useful process modification is switching from 4" (100mm) diameter wafers to 6" (150mm) diameter wafers. Although switching wafer sizes does not have any inherent differences, in the UC Berkeley Microlab it means that different lithography and etching equipment will be used. While the etching tools have very similar capabilities at both 4" and 6", the primary difference comes in the lithography tools. 6" wafers can be processed using one of two step-and-repeat lithography tools available: the ASML DUV stepper model 5500/90, and the GCA stepper model 8500. Both tools use 5x reduction optics; the difference is that the ASML is capable of resolving features as small as 0.35µm with a field size up to 21x21mm in a production environment, whereas the GCA is only capable of resolving features as small as 0.7µm with a field size up to 15x15mm in production. The primary reason for the difference in performance between the two

steppers is the wavelength of light used. This determines the types of photoresist available for the process. The GCA uses the same I-line photoresist as is used for 4" wafers, while the ASML uses deep ultraviolet (DUV) photoresist only. Process development, mostly in the area of determining etch selectivity, is required before making the switch to DUV resist. Because each of these steppers utilizes 5x reduction optics, the photomasks must be made with smaller features as well. This is not a problem when using the GCA stepper, but unfortunately the minimum feature that can be produced by our in-house pattern generator is 2μm, yielding a minimum feature size of 0.4μm after going through the 5x reduction optics. In order to take advantage of the full abilities of the ASML stepper, an outside vendor is required that can produce quartz photomasks with minimum features ~1.5μm. Luckily, several such vendors operate in the bay area. With educational discounts, such masks can be made for less than $500/mask.

Because the etch aspect ratio does not change between the 4" and 6" deep reactive ion etchers, reducing the minimum feature size with the new lithography tools means that the device layer must be made thinner as well. The current minimum feature size of 1.5μm (on the wafer, not the mask) with a device layer thickness of 20μm yields an aspect ratio (AR) of 13. To scale down to 0.35μm features, a device layer slightly less than 5μm thick is required. However, in (Marty, Saadany et al. 2005), it was shown that the aspect ratio of deep reactive ion etching improves as the trenches get smaller than 1μm. This aspect ratio improvement is yet to be demonstrated in our lab.

There are many implications for the downward scaling of devices, some negative and some positive. The first is that it becomes more difficult to handle parts. As the minimum feature sizes drop into the submicron range, it becomes more and more difficult to use optical microscopes to observe and direct the assembly procedure. This is partly ameliorated by the use of the Zyvex S-100 nanomanipulator in an electron microscope to maneuver ortho-grippers around the workspace. Creating end-effectors is not a stumbling block because the ortho-grippers are manufactured in the same process as the parts that they need to grip. Therefore, they can scale down at the same rate as the rest of the devices being built. However, the parts themselves become more difficult to handle since spring constants scale down linearly, while electrostatic forces are independent of geometric scaling (Fearing 1995). This means that the parts will tend to stick to surfaces more easily, making part manipulation (especially releasing parts) more difficult. One benefit is that electrostatic actuator force density will increase as the minimum feature size decreases (see Appendix A). Another benefit is that the downward scaling of spring constants makes designing flexural hinges easier.

It is currently not known what the lower limit of scaling might be for pick and place assembly. Drexler (Drexler 1986), thinks that pick and place assembly of individual atoms is possible. However, Smalley (Smalley 2003) disagrees, saying that sufficiently small parts will be impossible to structure in such a way that they can be grasped and released reliably. At any rate, pick and place assembly is probably not the right approach for fashioning complicated nanostructures. Instead, the approach taken in this work utilizes one or just a few assembly operations of already-patterned parts to make more

complex devices. In this way, we remove ourselves from the argument over assembly's

lower limit, since at a certain point downward scaling will reduce the capabilities of each

assembled part, possibly at a greater rate than assembly increases the capability of the

device.

Although parts may be more difficult to handle when fabricated at this smaller scale, it is

not always necessary to build parts at the lower limit of what is possible. Hybrid

structures can take advantage of different size scales. For instance, a comb drive actuator

could take advantage of fine-line lithography, while a micromirror does not require fine

features and would instead benefit from tall supporting ribs to help stiffen the mirror

without adding very much mass. Or a leg for a robot would benefit from the extra out-of-

plane stiffness that the greater width of being built in a thicker SOI process would give.

One of the greatest benefits of assembly in this context is that each component of a

device can be built in its own process, using whatever SOI thicknesses and line widths

are most optimal for the particular component.

6.4. CMOS Integration

Integration of MEMS devices with CMOS electronics has been an elusive goal for years,

with compromises required both in the capabilities of the MEMS process and the CMOS

process in order to combine both of them on the same chip. The successful approaches

so far can be categorized as MEMS-first or CMOS-first, referring to which was built first

on the substrate. Bosch and Analog Devices both chose a MEMS-first approach for their

inertial sensors. The advantage of this approach is that it can accommodate the high-

temperature deposition and oxidation steps that are part of the MEMS process. Franke and Howe (Franke, Heck John et al. 2003) developed CMOS-first approaches using low-temperature Silicon-Germanium surface micromachining. In general, MEMS-first approaches tend to place the MEMS structures in a separate section of the chip from the CMOS circuitry. In addition, they tend not to utilize the most advanced CMOS processes. In CMOS-first, the main restriction involves the severe temperature constraints imposed by the risk of melting the aluminum traces. This severely restricts the materials available for the subsequent MEMS processing. In both cases, by integrating both the MEMS and the CMOS monolithically on the same chip, the process complexity goes up dramatically and therefore yield can suffer. The key is to minimize this yield hit by using as simple a process as possible to combine the CMOS and MEMS parts of the microsystem.

Microassembly provides a different path to CMOS integration, where both the CMOS and MEMS are made concurrently on separate chips and assembled at the end. The advantage of this approach is that a defect in either the CMOS or MEMS chip does not doom the entire microsystem to failure. Instead, the defective chip is discarded and a fully functional copy is used instead. The trick to this approach lies in finding a simple process for connecting the two chips together.

Flip chip assembly utilizing gold bumps or solder balls is one approach that has shown success for connecting MEMS and CMOS chips. Developed to increase the density of off-chip electrical interconnects and to decrease the size and cost of packaging CMOS

chips, this approach has also shown promise in integrating CMOS and MEMS chips. Unfortunately, there are integration issues that stem from where the release step appears in the MEMS process. From the MEMS chip's perspective, assembly is best performed before the sacrificial layer has been removed when the chip is most able to withstand the stresses of bonding, while from the CMOS chip's perspective assembly is best performed after release since the inter-layer dielectric of the CMOS chip or the solder bumps used to connect the chips will not withstand the acid used to release the MEMS chip.

Making the chips compatible with the assembly process described here would eliminate the processing issues involved with flip-chip bonding, since assembly could be performed after the MEMS chip was released. One method to do this would be to perform a deep etch into the frontside of a CMOS wafer, then bond the wafer face-down to a handle wafer and grind/polish the wafer until the trenches have been exposed. This would create patterned parts that could be picked up and assembled into a clamp fabricated on a separate MEMS chip.

But this single-mask post-processing step has more power than just making the CMOS chips compatible with MEMS chips. This process has the same capabilities as the single-mask SOI process described in this chapter, meaning that assembly tooling, clamps, parts, and actuators could be made just using CMOS chips, with no separate MEMS chip required. In this case, the SOI MEMS chip would be useful as an inexpensive micro-scale backplane to plug these CMOS MEMS chips into. Working as a sort of MEMS

circuit board, the SOI MEMS process would be an inexpensive method to tie the active circuit-MEMS hybrids together.

This solution is interesting, but suffers from the same problem as all other CMOS-first approaches, namely, a yield hit on the MEMS post-process dooms CMOS that would have been functional. A better approach would take CMOS dice and assemble them without the requirement for additional post-processing. Many methods could be envisioned to do this, from sliding a CMOS chip edge-on into a connector fabricated in a MEMS chip, to modifying the gold bump bond mask so that the bumps would be the right shape for the MEMS chip to clamp onto (Cohn 1997; Cohn et al. 1998), to assembling microfabricated spring clips that would take the place of wirebonds.

7. FUTURE OF MICROASSEMBLY

There are 3 main reasons to use a microassembly-based process: 1) to rapidly make a functional MEMS prototype, 2) to reduce the cost of making the part by reducing fabrication time and complexity, or 3) when it results in the best-performing device. In many cases, the simple microassembly-based process presented here will be too simple to meet all specifications for a given device. The power of this process however stems from its speed and flexibility: many different designs are possible in rudimentary form without altering the process, and unaltered, this manufacturing process takes only days to cycle through. This process was designed to be as simple as possible; a thought experiment in what could be done with just a single mask, a single etch, and a release step. All of the complexity was pushed into the post-process, i.e. after the wafer was divided into chips and released. Additions to the process, such as additional masks or materials must be carefully weighed against the increase in process complexity and increase in design cycle time to avoid eliminating the main benefits of this process. Assembly is not yet routine or easy, and as of the writing of this paper involves a substantial yield hit (up to ~50% for the two-axis rotator described in Section 7 below, less for easier parts to assemble). But compared to adding process complexity, it still comes out ahead for rapid prototyping. Its value in a mass-production environment is yet to be seen.

Microassembly also shows the potential to be a powerful approach to making hybrid microsystems that combine chips made in different processes. Any process that can create and pattern thin structures (~50μm or less) will be compatible with the

microassembly approach as described here. These could be electronic components, optoelectronic, or mechanical components made in a different process. The MEMS device would act like a micro- or nano-circuit board, allowing multiple components to be linked together through mechanical clamps and/or gold-silicon eutectic bonding. The potential to perform snap-together integration of specialized components promises to make complicated integrated micro- and nano-systems a reality, rather than just micro-devices.

8. TIP/TILT MICROMIRROR

The design, numerical optimization, and experimental verification of a 2-axis micromirror are presented in this chapter. First, the design goals of this micromirror are listed and the reasons for choosing this manufacturing process explained. Next, the expected mirror performance is derived using the equation of motion and simple beam-bending theory. The design is then refined using finite element analysis and the performance of the device is optimized with regards to DC deflection angle, bandwidth, and manufacturability. Finally, experimental results are presented and compared to the theoretical and numerical predictions.

8.1. Figure of Merit

The micromirror design is driven by its optical performance. One possible figure of merit (FOM) that can be used to encapsulate most of the important features of a micromirror consists of the number optical spots it can aim at per second. This touches on the size and quality of the mirror (these affect the size of the spot), the maximum angle the mirror can move to (the number of spots), and its bandwidth (how fast it can switch between spots).

8.1.1. Spot Size

The first part of this figure of merit, how big the spot is, is based on the diffraction limit

of light bouncing off the mirror. For a flat, circular mirror normal to an incident beam,

this is given by:

$$\theta_{div} = \frac{1.22\lambda}{D_m} \tag{8.1}$$

Where θ_{div} is the full-width angle between the first null-points of the intensity profile of

light in the far field, λ is the wavelength of the light, and D_m is the diameter of the mirror.

When the mirror is tilted, this diffraction-limited beam divergence increases, leaving a

less tightly-collimated beam. This effect, where the apparent size of the aperture

decreases, is just caused by the reduction in the projected area of the aperture and is

accounted for by the addition of a cosine term:

$$D_a = D_m \cos(\theta_m) \tag{8.2}$$

Where D_a is the apparent size of the aperture and θ_m is the angle the mirror has been tilted

by. Taking this into account, the new diffraction-limited beam divergence becomes a

function of the mirror tilt angle and is given by:

$$\theta_{div}(\theta_m) = \frac{1.22\lambda}{D_a} = \frac{1.22\lambda}{D_m \cos(\theta_m)} \tag{8.3}$$

This beam divergence angle assumes a perfectly flat mirror. Any curvature of the mirror will increase this angle.

8.1.2. Number of Spots

The second portion of the figure of merit is the maximum angle θ_{max} the mirror can move to. Combining this angle with the diffraction-limited divergence angle of the mirror, we get the number of spots the mirror can resolve:

$$N_{spots} = \frac{\theta_{\max}}{\theta_{div}(\theta_m)} \tag{8.4}$$

Although the exact number of non-overlapping spots is possible to derive, the formulation is complex and reduces the benefit of the theory in developing intuition about the problem. Instead, we replace the exact number of spots for two bounds on the number of spots: the lower bound is when $\theta_{div}(\theta_m)$ is replaced with $\theta_{div}(\theta_{\max})$, and the upper bound is when $\theta_{div}(\theta_m)$ is replaced with $\theta_{div}(0) = \theta_{div}$ (the original formulation for the diffraction angle of a circular aperture).

8.1.3. Speed

The third aspect of the figure of merit is the mechanical bandwidth of the mirror. This determines the number of spots per second the mirror can illuminate. For micromirrors with a flexure-based suspension, the speed and the maximum angle are tied together through the equation of motion (in angular form):

$$I\ddot{\theta} + B\dot{\theta} + J\theta = \tau \qquad\qquad (8.5)$$

Where I is the moment of inertia of the moving parts, B is the damping in the system, and J is the spring constant of the suspension. τ is the torque exerted by the actuator on the mirror. In the case of a static torque applied to the mirror, $\dot{\theta}$ and $\ddot{\theta}$ are both zero, giving $\theta = \dfrac{\tau}{J}$. So $\theta_{\max} = \dfrac{\tau_m}{J}$ where τ_m is the maximum torque that can be applied to the micromirror. The natural frequency of the system is $\omega_n = \sqrt{\dfrac{J}{I}}$. This provides an upper limit on the angle traversed per second, since any increase in the frequency of the mirror oscillation past the natural frequency results in a decrease in the angle traversed. This shrinkage of the traversal angle is proportional to frequency squared, so the tradeoff between frequency and angle is not a good one at high frequencies. We can define an angle-bandwidth product $P_{ABW} = \theta_{\max}\omega_n^2 = \dfrac{\tau_m}{I}$ that is independent of the stiffness of the suspension. To trade between angle and bandwidth squared, we can vary the suspension stiffness, but the product of the two remains invariant. This provides a convenient way to combine maximum angle and speed while taking into account the mechanics of the system. The final term of the equation of motion, B tells us about the damping present in the system. The quality factor Q describes the ratio of energy storage in the system to energy dissipation; $Q < 1$ is overdamped, $Q > 1$ is underdamped, and $Q = 1$ is critically damped. Most MEMS micromirrors, especially large ones (~1mm

diameter), are underdamped. This adversely affects step response time, given by

$$t_s = \frac{8Q}{\omega_n} \text{ for } Q > 1.$$

8.1.4. Summary

So we can now formulate an actual figure of merit, using the mechanical angle-

bandwidth invariant:

$$FOM = N_{spots}\left(\frac{spots}{\sec}\right)^2 = N_{spots}\left(\frac{1}{t_s}\right)^2 = \frac{\theta_{\max}}{\theta_{div}}\frac{\omega_n^2}{64Q^2} = \frac{P_{ABW}}{64\theta_{div}Q^2} \qquad (8.6)$$

This figure of merit is given by the total number of spots the mirror can aim at (dynamic

range) times the square of the number of spots per second the mirror can aim at (speed).

Both P_{ABW} and θ_{div} are related to the radius of the mirror. This formulation tells us

several things about what makes an ideal micromirror-based beamsteering system. First,

large, lightly damped mirrors are slow to settle into their new location, so the mirror

should be kept as small and light as possible. However, external constraints on beam

divergence often limit the minimum diameter of the mirror or what wavelengths of light

are to be used.

8.2. Micromirror Design

The formulation derived in section 8.1.4 gives us some insight into what makes a good micromirror system, but we now need some way to relate this to the physical design parameters of the system. By expanding these parameters, we get to the actual design variables.

8.2.1. Angle-Bandwidth Product

The angle-bandwidth product P_{ABW} is related to physical parameters through the moment of inertia I of the mirror, which for a flat circular plate is given by:

$$I = \frac{\pi}{4}\rho_{Si}t_m r_m^4 \qquad (8.7)$$

where ρ_{Si} is the density of silicon, t_m is the thickness of the mirror plate, and r_m is the radius of the mirror. Substituting this into the equation for P_{ABW} we get:

$$P_{ABW} = \frac{\tau_m}{I} = \frac{4\tau_m}{\pi\rho_{Si}t_m r_m^4} \qquad (8.8)$$

8.2.2. Quality Factor

The quality factor Q is given by:

$$Q = \omega_n \frac{I}{B} \qquad (8.9)$$

Where the damping term B is the total damping of the micromirror system. The damping comes from just a few sources: viscous damping of air moving around the mirror as it tilts, squeeze-film damping as the air gets compressed or expanded between the mirror and the substrate, and in the case where the mirror is actuated using an electrostatic comb drive, there is damping due to couette flow between the fingers of the actuator.

Viscous drag is proportional to the square of the velocity of an object moving through a fluid. Since the velocity of a given point on the mirror is a function of its distance from the axis of rotation, an integration is required to figure out the total torque required to counteract the effects of viscous drag. For a circular plate undergoing rotation about its y-axis (or by symmetry, any other axis through the mirror's midpoint and along its surface), this integral is:

$$\tau_{vd} = -\int_{-R}^{R}\int_{-\sqrt{R^2-x^2}}^{\sqrt{R^2-x^2}} C_d \rho_{air}\left(y\dot{\theta}\right)^2 dydx = -\frac{\pi}{4}C_d\rho_{air}R^4\dot{\theta}^2 \qquad (8.10)$$

Where C_d is approximately 1.5 for typical Reynolds numbers and speeds a 1mm-diameter mirror might encounter (Re $\approx$ 1-10), $\rho_{air} = 1.29\frac{kg}{m^3}$, R is the radius of the mirror, and $\dot{\theta}$ is the angular velocity of the mirror.

Under normal circumstances, the layer of a fluid in contact with a surface will not move relative to the surface. This effect is called the non-slip condition. If this surface moves, the air immediately next to it moves at the same velocity as the surface, while air further away moves at a velocity in between that of the surface and the bulk fluid. The energy

88

required to shear the fluid creates a drag force. When no other surfaces are nearby, the slope of this gradient is determined by the viscosity μ of the fluid. If a second motionless surface is placed close by, the non-slip condition of that surface causes the fluid velocity gradient between the two surfaces to increase. The closer together these two surfaces get, the larger the gradient. This shearing motion of fluid between two surfaces is called couette flow and the force drag force imparted to the surfaces by the fluid is given by:

$$F_c = -\frac{4\mu A}{g} \dot{x} \qquad (8.11)$$

Where A is the area of the smaller surface, g is the gap between the surfaces, and $\dot{x}$ is the velocity of the moving surface.

The region between moving comb fingers and stationary fingers experiences this type of drag. Transforming this linear force into a torque allows one to compare the magnitude of this damping term with the others described here. To go from linear motion to rotation, the relation $\dot{x} = \frac{dx}{dt} = \frac{d\left(L_m \sin(\theta)\right)}{dt} = L_m \cos\left(\dot{\theta}\right)$ is used. To transform between linear force and torque, the relation $\tau = F_{couette} L_m \cos(\theta)$ is used. Finally, since there are two gaps per comb finger and we use bi-directional drive (there are twice the number of fingers, therefore twice the damping), we have a total area of $A = 2 N_f L_f (AR)$ where N_f is the number of comb fingers, L_f is the overlap distance of the comb fingers at zero displacement (under the assumption that the comb fingers never fully disengage from one

another at any time), and $AR = \dfrac{h}{g}$ is the aspect ratio of the trench between the fingers.

Combining these equations, the mirror damping from couette flow between the comb fingers is given by:

$$\tau_{cf} = -4\mu N_f L_f (AR) L_m^2 \cos^2(\theta)\dot{\theta} \qquad (8.12)$$

The squeeze film damping between a micromirror and the substrate beneath it becomes significant as the gap between mirror and substrate gets to be small compared to the radius of the mirror. An expression for the squeeze film damping between a torsionally suspended micromirror and the substrate beneath it was derived in (Pan, Kubby et al. 1998). Their linearized (small deflection) model for the the damping torque on a rectangular mirror of length L and width W is given by:

$$\tau_{sf} \cong -\frac{48\mu L W^5}{\pi^6\left(\left(\dfrac{L}{W}\right)^2 + 4\right)g_0^3}\dot{\theta} \qquad (8.13)$$

Where μ is the viscosity of the air (given above) and g_0 is the initial gap between the mirror and the substrate.

In the case of a large mirror (~1mm diameter) well removed from the substrate, the viscous drag of the mirror will be the dominant damping force, followed by squeeze film damping between the comb fingers (and to a more limited extent, between the moving

90

structures and the substrate), followed by squeeze film damping between the mirror and the substrate.

The damping coefficient B is determined by plugging these equations into the equation of motion ($I\ddot{\theta} + B\dot{\theta} + J\theta = \tau$) and setting θ and $\ddot{\theta}$ to zero. B is obtained by summing the torques from the different sources of damping and dividing through by $\dot{\theta}$, giving:

$$B = \frac{\tau_{vd} + \tau_{cf} + \tau_{sf}}{\dot{\theta}} \qquad (8.14)$$

The figure of merit can now be expressed in terms of physical parameters of the beamsteering system:

$$FOM = \frac{\left(\dfrac{\tau_m}{I}\right)}{64\left(\dfrac{1.22\lambda}{2r_m}\right)\left(\omega_n \dfrac{I}{B}\right)^2} = \frac{\tau_m B^2 r_m}{32(1.22)\lambda\left(\dfrac{J}{I}\right)I^3} = \frac{\tau_m B^2}{2(1.22)\lambda J\pi^2 \rho_{Si}^2 t_m^2 r_m^7} \qquad (8.15)$$

8.2.3. Reflector Design

Looking at this equation, the importance of the mirror radius on the *FOM* is clear. The first step in designing any micromirror system should be to determine the minimum mirror radius acceptable for the application. Even small changes in this parameter will greatly affect the beamsteering system performance. Once the mirror radius is minimized, its volume and density should be minimized. This can be done by replacing a

solid mirror with a thin mirrored face sheet stiffened with supporting ribs. Material

should be removed until unacceptable amounts of curvature from thermal mismatch

between the mirror's structural and reflective materials are introduced, or dynamic

deformations during operation become excessive (Conant, Nee et al. 2000).

An example of this type of mirror is shown (Figure 28(a)). This mirror was constructed

during an earlier attempt at creating micromirrors. It consists of a thin (2µm-thick)

single-crystal silicon face sheet attached to 20µm-thick high aspect-ratio single-crystal

silicon ribs using polysilicon staples. Other ribbed mirrors have been made using

hydrophilic silicon-silicon fusion bonding with two patterned SOI wafers (Figure 28(b)).

These micromirrors are better because they do not require holes in the face sheet

wherever the ribs connect to it.
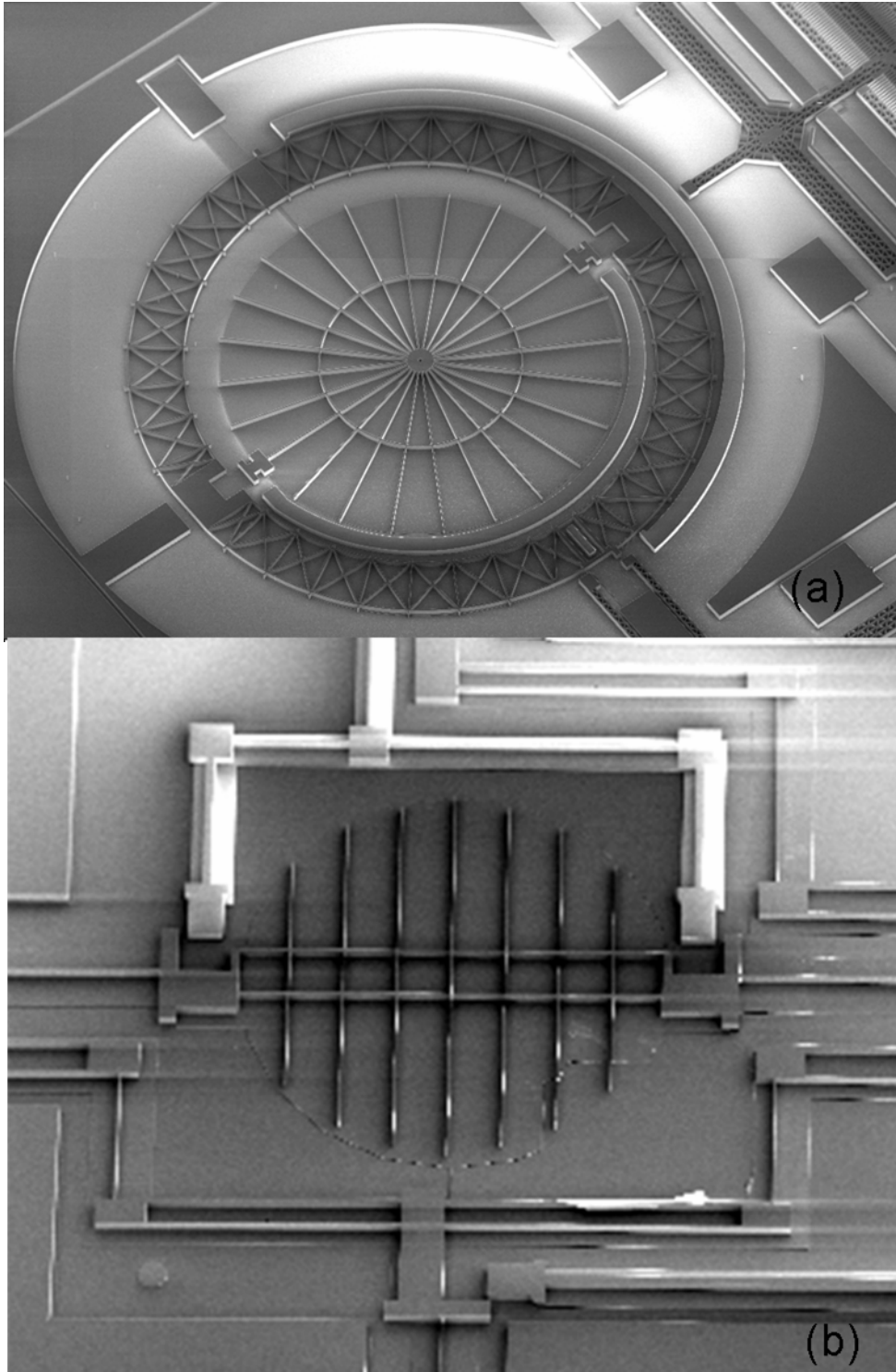
Figure 28. Ribbed Micromirrors with thin face sheets. These are 1mm diameter mirrors fabricated using two different technologies. Mirror (a) uses a conformal polysilicon deposition to connect the ribs to the mirror face sheet (circular disk below ribs), while mirror (b) is constructed by directly bonding the ribs to the mirror face sheet. In both cases, the reflective mirror surface is the far side of the face sheet. In these processes, the mirror is designed to be exposed using an etch through the handle wafer that the face sheet is bonded to.

93

The next target for optimization is mirror damping. Using the equations above, it is possible to show that for millimeter-scale mirrors, the damping is dominated by viscous drag on the mirror and couette flow in the actuators. To increase these damping effects, dampers can be added to the actuators (or the actuator area could be increased), or drag-enhancing hair-like structures can be added to the edge of the mirror. The hairs on the mirror have the benefit of also damping undesired modes of the mirror, especially the uncontrolled vertical bounce mode of the mirror. Another approach might be to place the micromirror in a package filled with a heavy, inert gas such as krypton or xenon. The increased density of the gas should result in more damping and faster overall step response time.

8.2.4. Moment Arm

Once all efforts to optimize the size, composition, and packaging of the mirror are complete, it is time to examine the other aspects of the beamsteering system. The first thing to choose is the amount of torque applied by the actuators. When linear electrostatic actuators are used to supply the driving force, a mechanical linkage is required to convert the linear motion of the actuators into the rotational motion required for the mirror. Consisting of a beam of length L_m and a flexural pivot, this linkage acts as a moment arm that converts the force generated by the actuator into a torque. This torque is given by $\tau_a = F_a L_m \cos(\theta)$, where F_a is the force generated by the actuator, L_m is the length of the moment arm, and θ is the angle of the mirror, given by:

$$\theta = \sin^{-1}\left(\frac{x_a}{L_m}\right) \qquad\qquad (8.16)$$

Where x_a is the linear displacement of the actuator. From these two equations it can be seen that, all else being equal, there is a tradeoff involved in choosing L_m: if it is very long, the torque will be large but the rotation angle will be small, while a short L_m yields a low torque but large rotation angle. Additionally, many actuator technologies face an inherent tradeoff between force generated (F_a) and distance traveled (x_a). So again we face a tradeoff between torque and angle.

Unfortunately, our intent to use the angle-bandwidth product to eliminate the spring constant term from the figure of merit is thwarted by the presence of the natural frequency ω_n in the equation for the quality factor Q of the mirror. This reintroduces the spring constant J into the denominator of the *FOM* equation.

8.2.5. Suspension

The suspension of the micromirror is based on silicon flexures to eliminate friction and hysteresis in the system. It is designed to be built in the microassembly process described in this thesis. We would like to use the fewest number of assembled parts possible to maximize the assembly yield. If we use both the bending and twisting modes of a beam,

we can achieve two-axis rotation with a single assembled part. The rotational spring constant of a flat rectangular beam in torsion is given by:

$$J_t = \frac{G_{Si} t^3 w}{3L} \tag{8.17}$$

While the rotational spring constant of a flat rectangular beam in flexion (with a pure moment applied at the tip of the beam) is given by:

$$J_f = \frac{EI}{L} = \frac{E_{Si} t^3 w}{12L} \tag{8.18}$$

where t is the thickness of the beam (the short side of the cross-section of the beam), w is the width of the beam (the long side of the cross-section), L is the length of the beam, and E_{Si} and G_{Si} are the young's and shear moduli (respectively) of silicon.

8.2.5.1. Design Goal: Maximum Range of Motion

Many design goals are possible, including maximizing the range of motion of the mirror, maximizing the resonant frequency of the mirror, or matching the natural frequencies of the two modes. Because the FOM weights the resonant frequency more heavily than range of motion, that would be a natural quantity to maximize. However, many micromirror applications care more for the number of spots the mirror can hit than the speed with which the micromirror can hit them. Applications where this is the case include point-to-point communication between stationary or moving transceivers, all-

96

optical fiberoptic switches, and many brightness-limited optical systems where the mirror must dwell in a certain location for long enough to allow sufficient photons to pass. In all these cases, the micromirror is sufficiently fast to meet the application's needs and further improvement is wasted effort. In many applications, this point of diminishing returns occurs in the high hundreds of hertz or low kilohertz range.

Because of this, we will temporarily ignore our quest to maximize the figure of merit and instead attempt to maximize the range of motion. To do this, we need to choose design parameters that keep the spring constant low. The easiest to start with is w, since it is defined by the thickness of the device layer of the SOI wafer. In this case, the device layer is 20μm thick. The thickness of the flexure t is the larger of the two lower limits on beam width: minimum lithographic feature size, or minimum line width as determined by the etching aspect ratio. The current process has a minimum lithographic feature size of 1.5μm, and an etching aspect ratio of ~15. So t is 1.5μm. The length of the beam L is a free variable which we will use to ensure that the maximum stress in the beam is kept well below the yield stress.

8.2.5.2. Maximum Stress

The maximum stress in a beam in flexion is encountered at the base, where the beam meets the anchor. It is given by (Young 1989) (pg 94 eq 1):

$$\sigma_f(y) = \frac{My}{I} \tag{8.19}$$

where M is the bending moment at the base of the beam, y is the distance from the neutral axis of the beam, and I is the second moment of area of the beam, given by (Young 1989) (pg 62):

$$I = \frac{t^3 w}{12}$$

(8.20)

for a rectangular beam of thickness t and width w (Figure 29). The maximum stress occurs along the top and bottom corners of the beam where it attaches to the anchor, at $y = \pm \frac{t}{2}$. This gives a maximum stress of $\sigma_{f \max} = \dfrac{M\left(\dfrac{t}{2}\right)}{\left(\dfrac{t^3 w}{12}\right)} = \dfrac{6M}{t^2 w}$.
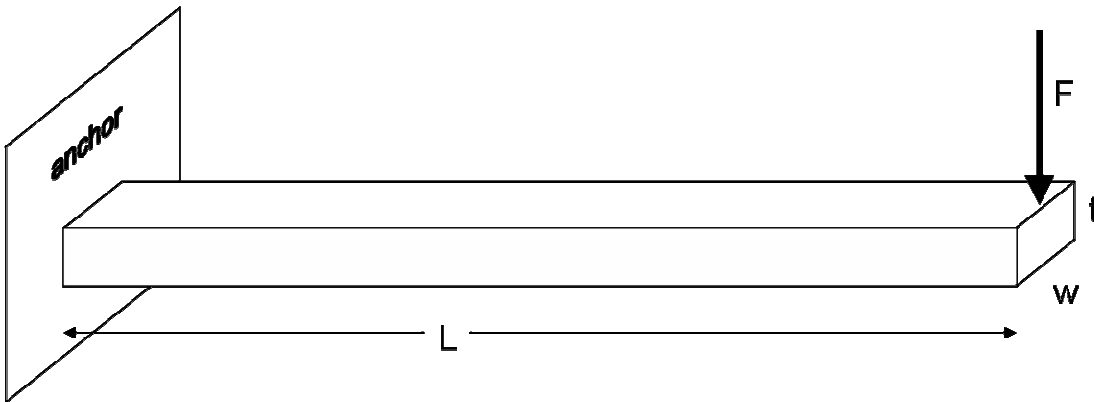

Figure 29. Rectangular cross-section beam of length L, width w, and thickness t. The thickness t is defined as the direction aligned with the applied bending force.

A beam in torsion sees a maximum shear stress along a line down the middle of the thin side of the beam. This is given by:

$$\sigma_t = \frac{Mr}{J_a} \qquad (8.21)$$

Where M is the torque on the beam, r is the distance from the neutral axis of the beam, and J_a is the polar moment of inertia of the beam. For a beam with rectangular cross-section, J_a is given by the sum of the 2nd moments of area about the width and thickness axes:

$$J_a = I_x + I_y = \frac{t^3 w}{12} + \frac{w^3 t}{12} = \frac{tw\left(t^2 + w^2\right)}{12} \qquad (8.22)$$

Giving a maximum shear stress of $\sigma_{t\,max} = \dfrac{M\left(\dfrac{w}{2}\right)}{\left(\dfrac{tw\left(t^2 + w^2\right)}{12}\right)} = \dfrac{6M}{t\left(t^2 + w^2\right)}$ for $w>t$.

The largest angle that a micromirror might ever need to achieve is +/-45 degrees in each direction. Since the angle of incidence is the same as the angle of reflection, a light beam normal to the mirror will make a right-angle turn for a mirror tilted by 45 degrees. If the mirror can rotate to any angle within a 45 degree cone in both tip and tilt axes, a laser beam can be steered anywhere in the hemisphere. Design of the suspension flexures is complicated by the fact that these equations assume small angles. To figure out the

deflection and stress of the suspension for larger angles, we will switch to numerical methods.

8.3. Implementation

This section describes the design, simulation, and experimental verification of the assembled two-axis rotation stage.

8.3.1. Modeling

We have previously developed an actuator capable of +/-15μm linear travel (x_a) with an output force up to 100μN (F_a). We will use this actuator for the micromirror. Our design goal will be +/-45° rotation in both axes. Step one is to choose the moment arm L_m that transforms the linear motion of the actuator to a rotational motion. Rearranging terms on equation (8.16) above, we find that $L_m = \dfrac{x_a}{\sin\theta} = \dfrac{2(15\mu m)}{\sqrt{2}} = 21\mu m$. The moment that this generates once coupled with the actuator

is $M = F_a L_m = (100\mu N)(21\mu m) = 2.1e-9Nm$. Since two springs in parallel are required (one attaching the mirror to the fixed anchor, and one attaching the mirror to the actuator – see Figure 30) this requires a spring constant of less than

$$J \le \frac{M}{2\theta} = \frac{(2.1e-9Nm)}{2(0.785rad)} = 1.3e-9Nm$$ in order to achieve 45 degrees of rotation. In the

microassembly technology with a 20μm-thick device layer and a 1.5μm lithographic limit, this requires a spring length of at least

$$L_t \geq \frac{G_{Si}t^3 w}{3J} = \frac{(80GPa)(1.5\mu m)^3(20\mu m)}{3(1.3e-9Nm)} = 1.4mm \text{ for the torsion beam or}$$

$$L_f \geq \frac{E_{Si}t^3 w}{12J} = \frac{(160GPa)(1.5\mu m)^3(20\mu m)}{12(1.3e-9Nm)} = 692\mu m \text{ for the flexion beam. These beams}$$

are fairly long, and it is unclear if a 1.4mm-long beam could be assembled. Even if they

could be assembled, it is not likely that these equations will apply to angles as large as 45
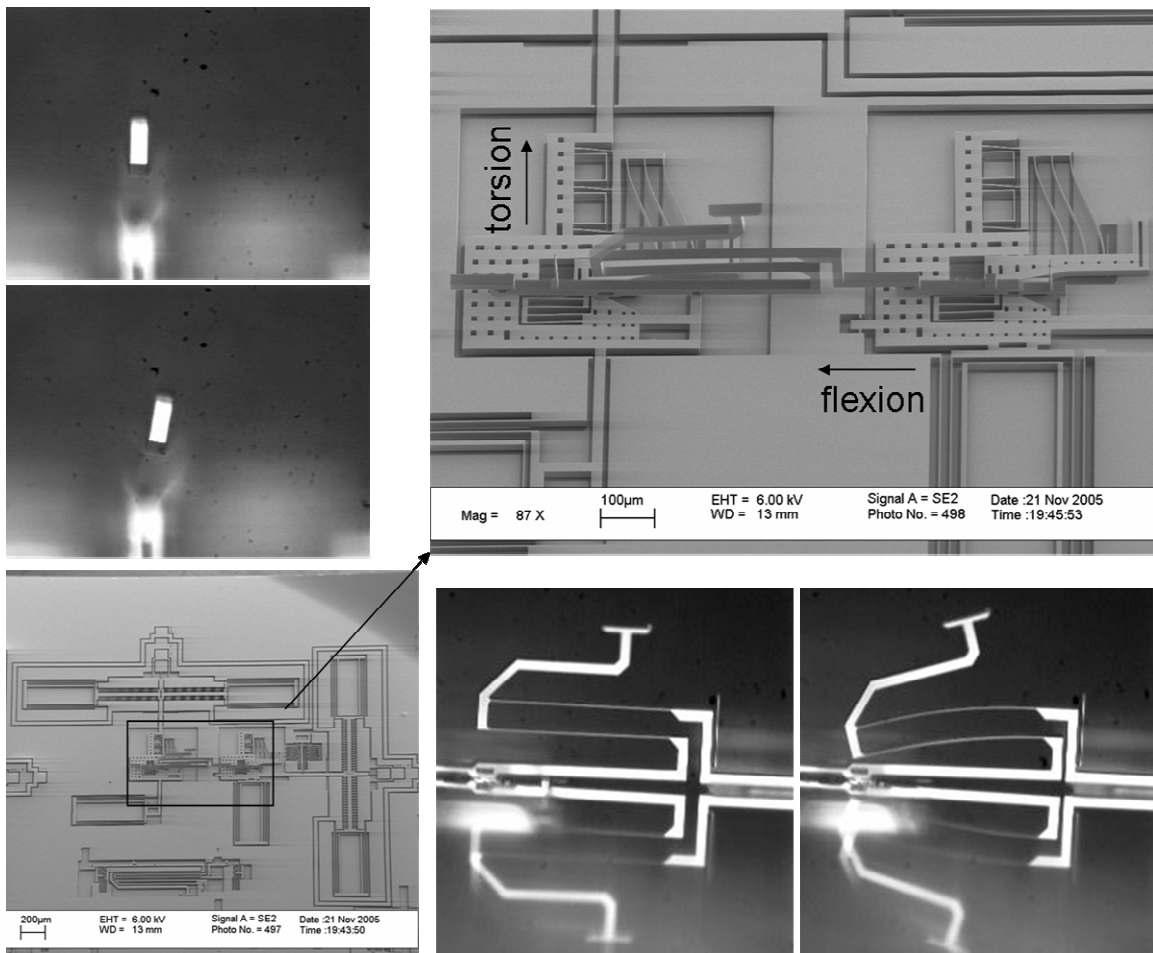
degrees.

Figure 30. 2-axis rotation stage. Bottom left electron micrograph shows birds-eye view of comb drive actuators, moving clamps, and assembled rotation stage. Larger electron micrograph shows a zoomed perspective of the assembled rotation stage and the clamps that hold each end. Photomicrographs show an end-on view of motion about the torsion axis (left) and a side view of

motion about the flex axis (bottom). The photomicrographs were taken using an optical microscope looking through a 45-degree mirror placed next to the chip.

Backing off on the design goal is the most straightforward way to address these problems. A design goal of $+/-15^o$ falls within the small angle approximation that the derivations of these equations use. With this new design goal, $L_m = 58.0 \mu m$, $M = 5.8e - 9Nm$, and $J \le 11.1e - 9Nm$, yielding $L_t \ge 163 \mu m$ and $L_f \ge 81.1 \mu m$. Since the suspension flexure operates in both bending and torsion modes, we choose the longer of the two required lengths for the flexure. This implies that if force-limited, the rotation stage will rotate two times further in the flexion mode than the torsion mode for the same input force, or $+/-15^o$ in the torsion mode and $+/-30^o$ in the flexion mode.

The maximum stress on the beam can be derived from the equations given in the previous section: $\sigma_{f\max} = \dfrac{6M}{t^2 w} = \dfrac{6(5.8e - 9Nm)}{(1.5e - 6m)^2 (20e - 6m)} = 770MPa$ for the beam's flexion

mode, and $\sigma_{t\max} = \dfrac{6M}{t(t^2 + w^2)} = \dfrac{6(5.8e - 9Nm)}{(1.5\mu m)\left[(1.5\mu m)^2 + (20\mu m)^2\right]} = 58MPa$ for the

beam's torsion mode. Unfortunately, 770MPa is above the stress limit for silicon; if that load were placed on the beam, there is a good chance that the beam will break. Luckily, our actuators cannot apply the full moment to the beam due to limitations in travel range of the actuator. Because the actuator can only travel 15µm in either direction, the rotation angle of the flexion axis will be limited to 15° because of the length of the moment arm. In other words, the flexural mode of the rotation stage operates at the

actuator displacement limit, not the actuator force limit. This implies that either the voltage used to drive the actuator can be reduced or the actuator size can be reduced, and the maximum stress experienced by the beam in flexion will be well within the fracture stress limit.

The final design prediction we can make is the natural frequency of the rotation stage with a mirror mounted on top. The moment of inertia of a circular plate pivoting about a line through its central axis is $I_m = \dfrac{\pi}{4}\rho_{Si}tR^4$. This would ideally be a ribbed mirror with a very thin reflective face sheet. However, in this experiment, we will consider just a solid circular piece of single-crystal silicon that has not been optimized with regards to moment of inertia. This will provide a lower bound on the natural frequency of the micromirror. The unoptimized mirror face sheet will be 20μm thick and 1mm in diameter, giving $I_m = \dfrac{\pi}{4}\left(2330\dfrac{kg}{m^3}\right)(20\mu m)(500\mu m)^4 = 2.29e-15 kg \cdot m^2$. For the spring constant derived above, this yields a natural frequency of :

$$\omega_n = \sqrt{\frac{J}{I_m}} \tag{8.23}$$

or $\omega_{nf} = \sqrt{\dfrac{\frac{1}{2}(11.1e-9Nm)}{2.29e-15 kgm^2}} = 1.56e3\dfrac{rad}{\sec} = 248Hz$ for the flexural mode, and

$\omega_{nt} = \sqrt{\dfrac{11.1e-9Nm}{2.29e-15 kgm^2}} = 2.20e3\dfrac{rad}{\sec} = 350Hz$ for the torsional mode.

103

An optimally designed reflector would have an I_m dominated by the mirror face sheet. This could be very thin; a lower bound is not known, but is expected to be in the range of a micron or so. A 1µm-thick, 1mm diameter face sheet has a moment of inertia 20x less than that derived above. This yields a 4.5x improvement in resonant frequency.

8.3.2. Simulation

The two-axis rotation stage is simulated in ANSYS using its nonlinear modeling capabilities. First, a reduced-complexity solid model is generated using the built-in modeling tool. The dimensions for each of the flexures and the moment arm are taken from the design generated in the previous section. The ANSYS model consists of two long (350µm), thin (1.5µm) parallel beams that act as the flexural elements in the rotation stage and a solid beam connecting the ends of the flexures that functions as the moment arm (Figure 31). These are longer than the beams derived in the previous section in order to provide a safety margin in terms of stress and required force. The faces of the beams at the opposite end from the moment arm have constraints applied to them: one beam is fixed, preventing it from moving in any direction, while the other beam has a displacement constraint applied to it. This constraint forces the face to move a specified distance along the z-axis (flexion axis) or along the x-axis (torsion axis).
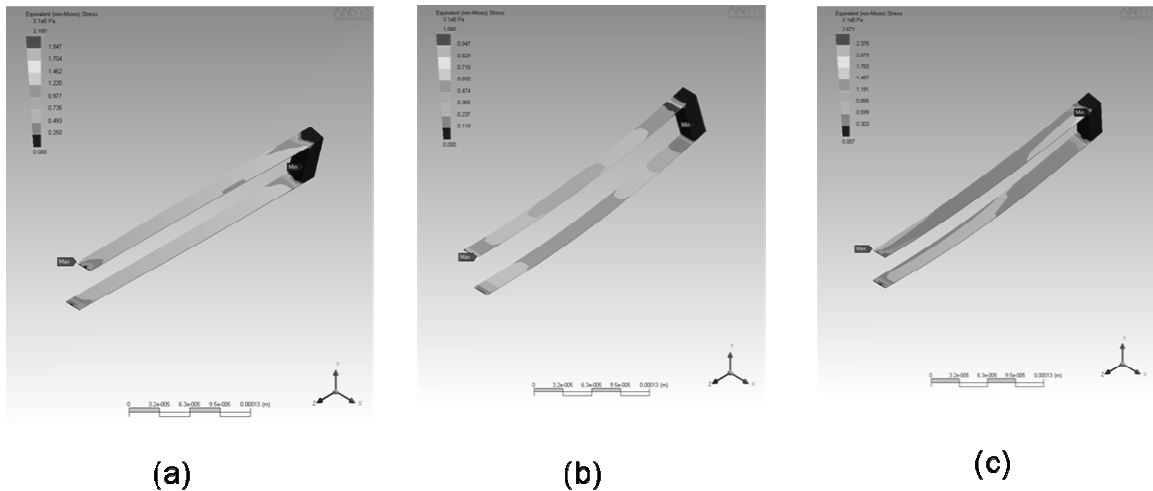
Figure 31. Simplified ANSYS models of 2-axis rotation stage. Von Mises stress obtained from FEM simulations at maximum deflection in (a) torsion mode, (b) flexion mode, and (c) combined torsion and flexion.

This model makes the assumption that the rest of the structure surrounding the flexures is rigid and is constrained to move in straight lines while not rotating at all. In order to ensure that this assumption is warranted, careful design of the supporting structures and flexures suspending the actuators and clamps is required.

Once these constraints are applied to the flexures, various displacements are applied to the end of the beam and the resulting reaction forces and beam tip displacements are recorded in a file. From these data points, the angle and vertical displacement of the moment arm can be calculated. This provides a matrix of angular displacement values and the applied force required to achieve those values. Due to the symmetry of the simulated structure, this only needs to be done for one quadrant of the range of motion of the actuator. In the simulations done here, the positive x direction and positive z direction are exercised (Figure 32).
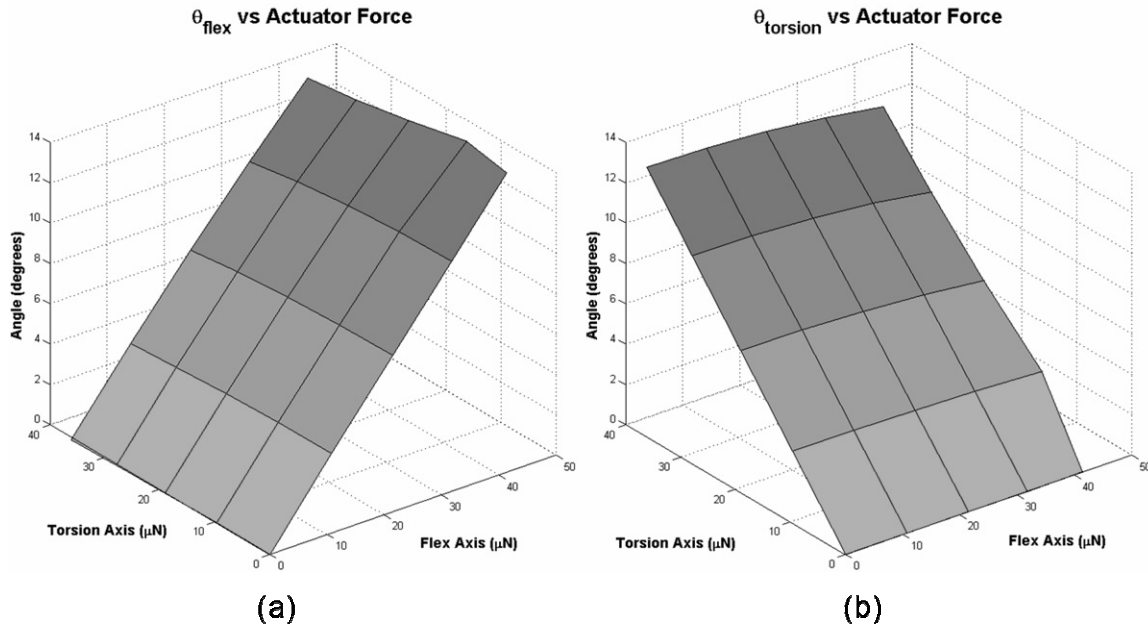
Figure 32. Curves of Angle vs Input Force. (a) Flexural angle vs input forces on each axis. Note minimal cross-coupling from torsion axis actuator even at large deflection angles. (b) Torsion angle vs input forces on each axis. The torsion axis does not reject cross-talk from the flex-axis actuator as well as (a) does. Total worst-case cross-coupling is less than 1 degree.

The simulation results show that due to the extra length of the flexures (350μm instead of the 163μm requirement derived above), the force required to achieve roughly 15 degrees deflection is well within the force that the actuators can deliver (100μN). In addition, these results show very good decoupling between the axes. Figure 32(a) shows the rotation about the flexural axis as a function of input force. As can be seen from the graph, the stage rotates fairly linearly with increasing force applied to the appropriate (flex-axis) actuator. In addition, it can be seen that it rejects the effects of the torsion axis actuator, even at large deflections. Similarly, Figure 32(b) shows the rotation about the torsion axis. Again, the rotation is roughly linear with force. However the torsion axis is not quite as decoupled from the flexural axis inputs. Although the flexures reject flex-axis input at no torsion angle, at maximum rotation about the torsion axis, almost a full

106

degree of rotation can come from the flexural axis actuator. This can be seen more easily

when the rotation angle versus flexure displacement is plotted (Figure 33).
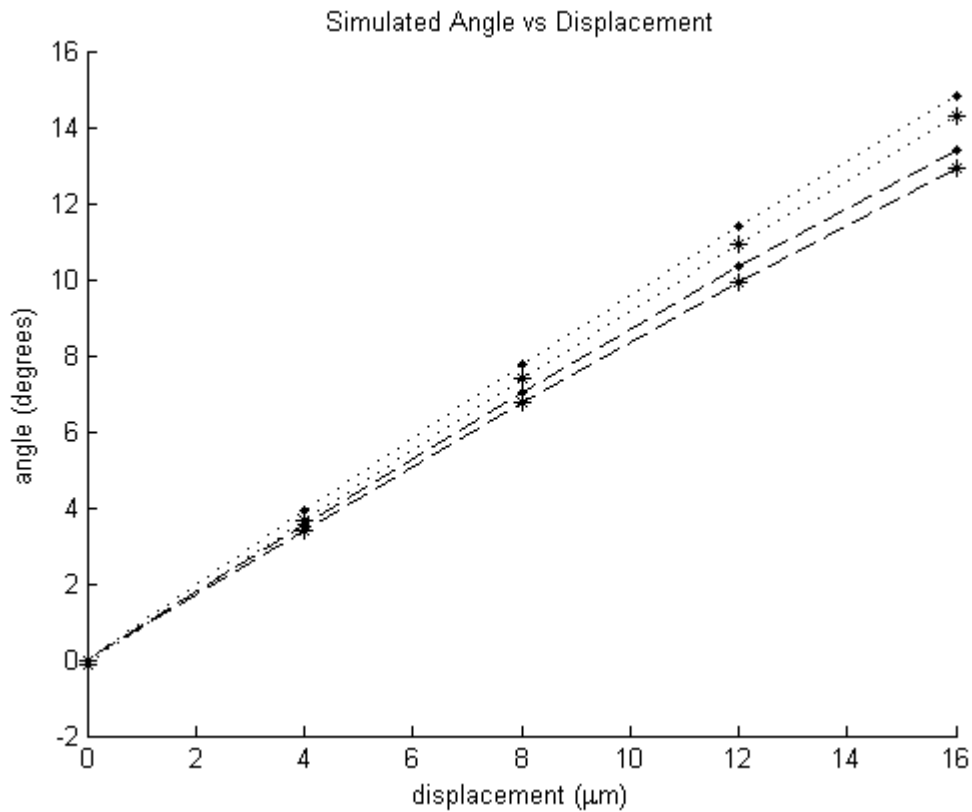
Figure 33. Angle vs endpoint displacement from finite element simulation. Dashed lines represent linear displacement and corresponding rotation angle for the torsion axis; dotted lines represent the flexural axis. For each data series, the worst-case cross-coupling is shown with stars, and no cross-coupling is shown with large dots.

Another drawback to this suspension stems from its vertical bounce mode. Because the

flexure is essentially a parallel-bar linkage, it has a natural inclination to move in the

vertical direction. In fact, the vertical bounce mode is the lowest-frequency mode of

operation, and if the micromirror needs to control the phase of the reflected light, this

suspension topology will not be the right one. However, if all that is important is the

107

ability to change the angle of the outgoing beam, the important metric is the coupling

between the bounce mode and the rotation modes (Figure 34). To simulate this, the

flexure endpoint is deflected by 15μm in each axis. This is done because the parallel-bar

linkage automatically decouples the bounce mode from other modes at zero deflection.

In fact, this decoupling property is one of the reasons why parallel-bar linkages are used

so often in mechanical design. Once rotated to the maximum angle, a vertical load is

placed on the moment arm. This vertical load is varied, and the change in angle in each

axis is plotted. As can be seen from the graph, the suspension is still quite resistant to

cross-coupling, with close to 9 gravities of acceleration required to cause the mirror

pointing angle to deviate by as much as one optical spot (assuming a mirror diameter of

1mm). Additionally, viscous and squeeze-film damping should cause this bounce mode

to have a reasonably low quality factor, meaning that resonance during operation will

most likely be kept to manageable levels. Unless the mirror is expected to operate in an

extremely high vibration environment or while experiencing large shocks, this suspension

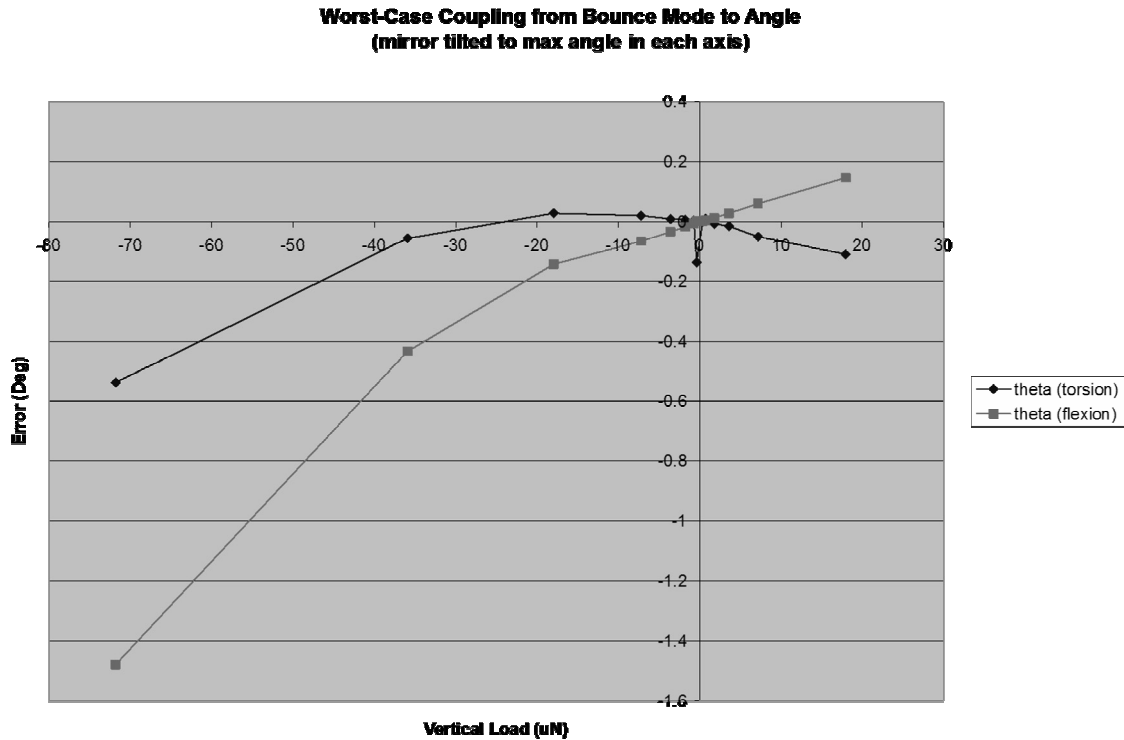adequately rejects bounce mode cross-coupling.

Figure 34. ANSYS simulation of worst-case coupling between bounce mode motion and rotation. The design of the flexure decouples the bounce and rotation modes when the flexure is at zero deflection, and only starts to couple into the rotation modes at large angles. The weight of the micromirror (approximately 0.36uN for a 20um-thick, 1mm diameter solid mirror) means that the mirror will rotate in the flexion axis by one diffraction-limited spot-width at approximately 9 gravities of acceleration.

This simulation of the coupling between the vertical bounce mode and the angular deflection of the mirror does not capture the whole picture however. There are two shortcomings to this model, both of which stem from the offset between the axes of rotation and the center of mass of the mirror. The first reason for this separation is the placement of the mirror mounting point far above the flexures that form the suspension. This is required in order to provide clearance between the mirror face sheet and the substrate. The second reason is that the axes of rotation do not intersect. The torsional axis of rotation runs through the stationary flexure (the flexure connected to the flex-axis actuators that is designed not to move in the same direction as the torsion-axis actuators).

109

Meanwhile, the flexural axis of rotation is a "virtual pivot" point located halfway between the two flexures, roughly a third of the distance down the length of the beam. The mirror mounting point is located as close to directly above this point as possible to minimize the vertical translation of the mirror face sheet as a response to rotation input. The design of the rotation stage is asymmetrical. This alone virtually guarantees that the coupling between rotation and bounce-mode motion will not be completely eliminated.

In order to operate properly, the actuator and clamp suspensions must only move in the desired directions. This is made more difficult by the design of the rotation stage. Because the moment arm travels a large distance in the vertical direction when it rotates around the flexural axis, the bottom flexure is located almost 150μm above the top of the device layer after assembly. This offset means that as the actuators move, the reaction force from the flexures generates a substantial moment that is felt by the actuator and clamp suspensions. Without careful attention to the design of these suspensions, this moment can cause the clamp and other in-plane structures to tip downwards until they rub against the handle wafer. This would cause undesirable amounts of friction, and depending on the exact conditions of the contact, might actually cause the moving structures to become immobilized.

Based on this concern, the double-folded flexures suspending the actuators and moving clamps were simulated using the moment calculated from the rotation stage geometry and expected reaction loads. Since the spring constant of the suspension in the vertical direction can decrease as the suspension deflects in-plane, the vertical spring constant of

the suspension was simulated both with no deflection and at the maximum deflection that the actuators can move before they run into their end-of-travel gap stops (15μm). Less than 0.5% difference in spring constant is visible in the simulation, so spring softening is not a concern at the small deflections simulated here.

Larger deflections are encountered when the actuator suspension has a moment applied to it, or even worse, when the suspension is deflected and then has a moment applied to it. The out-of-plane spring softening is clearly visible in this case (Figure 35).
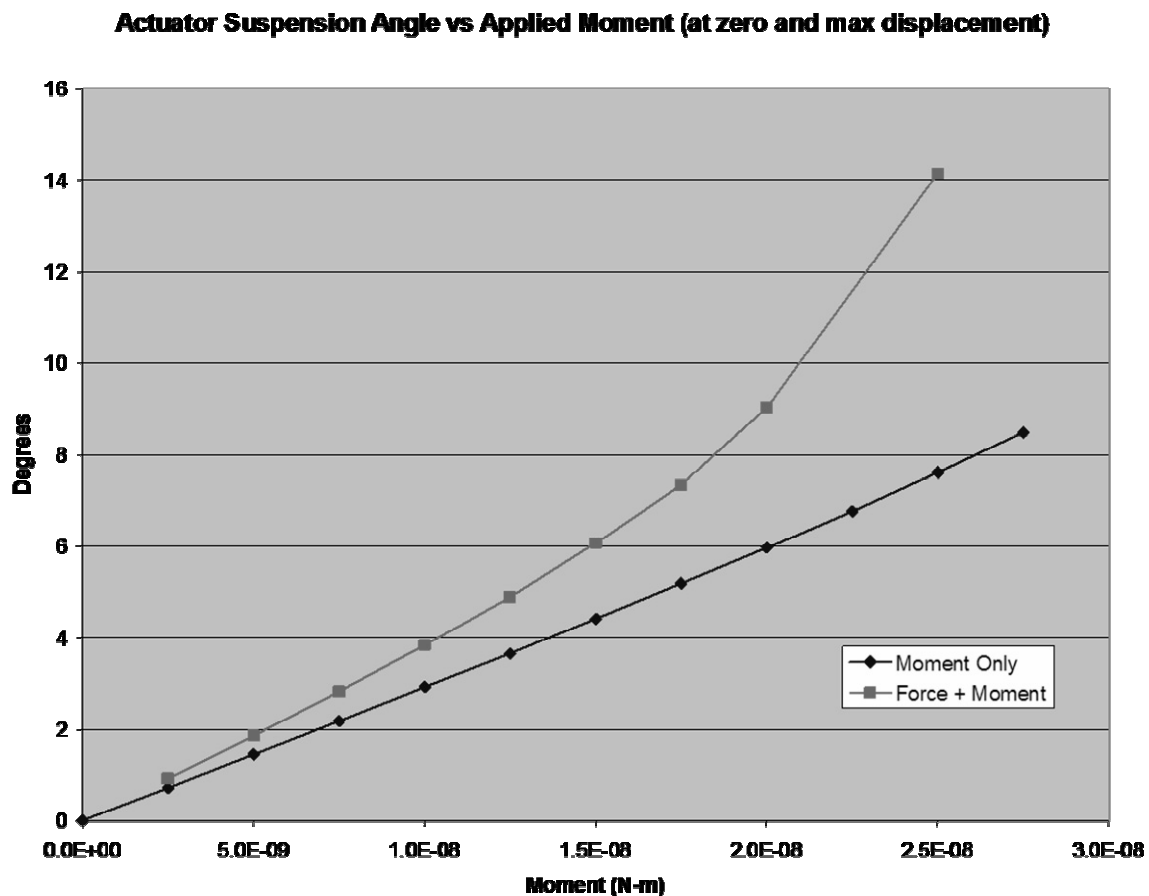

Figure 35. ANSYS Simulation of the actuator suspension. Smaller slope is better. Both a pure moment (straight line) and a combination of pre-deflecting the suspension to the maximum actuator throw in addition to the moment (curved line) are shown applied to the suspension. A reaction moment of approximately 5e-9Nm is expected as a result of actuating the rotation stage. This results

in a maximum vertical displacement of the actuator of approximately 1.5 microns at the point most distant from the axis of rotation.

This simulation is of a single double-folded suspension as a moment is applied to it. In the actual layout, the situation is much better. There are three separate copies of this suspension (one attached to the clamp and two attached to the actuator), leading to three times the vertical spring constant. In addition, the large distance between these suspensions causes the resistance to applied moments to be much better than what one might guess from just seeing the results of this simulation. It is expected based on this data that the clamps and actuators will not move enough in the vertical direction to touch the substrate.

8.3.3. Fabrication and Assembly

The two-axis rotation stage is fabricated as three separate parts: a bidirectional actuator for the flexural axis, a bidirectional actuator for the torsional axis, and an out-of-plane mechanism that converts the x- or y-axis motion of the actuators into out-of-plane rotation. Two separate single-axis actuators are used instead of a combined 2-axis actuator in order to conserve die area and to help decouple the axes from one another. Each of these actuators is connected to a fully-suspended clamp and an auxiliary suspension that helps resist the reaction moment generated by the rotation stage. The assembled rotation stage mechanism (shown as fabricated in Figure 36) inserts into the clamps using two square pegs with rounded ends. Tethers with integrated "tether-breakers" (see Figure 15 for details) hold the part in place until it is ready for assembly.
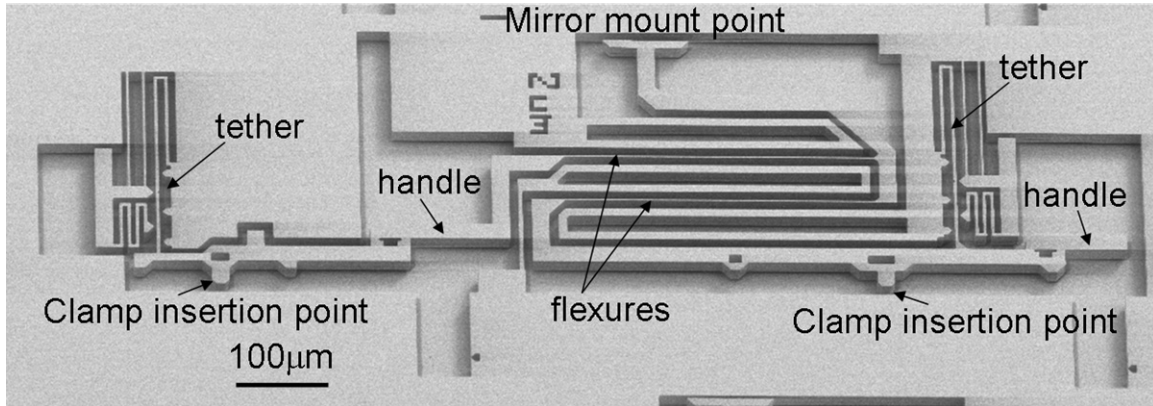
Figure 36. Electron micrograph of rotation stage (as fabricated). Two tethers hold the rotation stage in place. The rotation stage has two "handles" that can be gripped by an ortho-gripper, one at the right-hand end of the part, and one near the middle. Two pegs insert into clamps after assembly and provide the mechanical connection between the rotation stage and the actuators. Flexures are surrounded by dummy structures that ensure a uniform trench width. This minimizes the effect of RIE lag.

The assembled part has many interesting design features. Tethers hold the part in place until it is ready to be assembled. Integrated tether-breakers gently release the part from these restraints once the part is ready to be assembled. There are two different places on the part that the ortho-gripper can grab: one in the middle of the part and one at the very front of the part (right-hand side in Figure 36). In the results presented here, the handle in the middle of the part is used exclusively. These handles can interface with either the sidewall-gripping ortho-grippers or the original style of top/bottom-surface-gripping ortho-grippers. The interfaces to the clamps consist of pegs that are clamped in place and two additional "feet" that rest on the top surface of the device layer (Figure 37). Once assembled, these pegs extend downwards below the surface of the device layer a sufficient distance to remain well-clamped (15μm), but not so far as to extend past the bottom of the device layer and risk rubbing on the substrate. The pegs and feet are attached to 30μm-wide solid beams of silicon that are rigid in comparison to the flexures.

These rigid beams form the anchors for the flexures. The flexures are placed approximately 50μm above the rigid beam below them in order to provide clearance for the full range of motion of the flexural axis. Being very narrow, the flexures are quite susceptible to the effects of RIE lag and footing. In order to partially mitigate these effects, the trench width next to the flexures is kept exactly 10μm wide on each side using dummy structures that are attached to the device layer. Even with the dummy structures, mask undercut causes the flexures to become narrower. After fabrication, flexures that are drawn 2μm wide end up approximately 1.5μm wide.
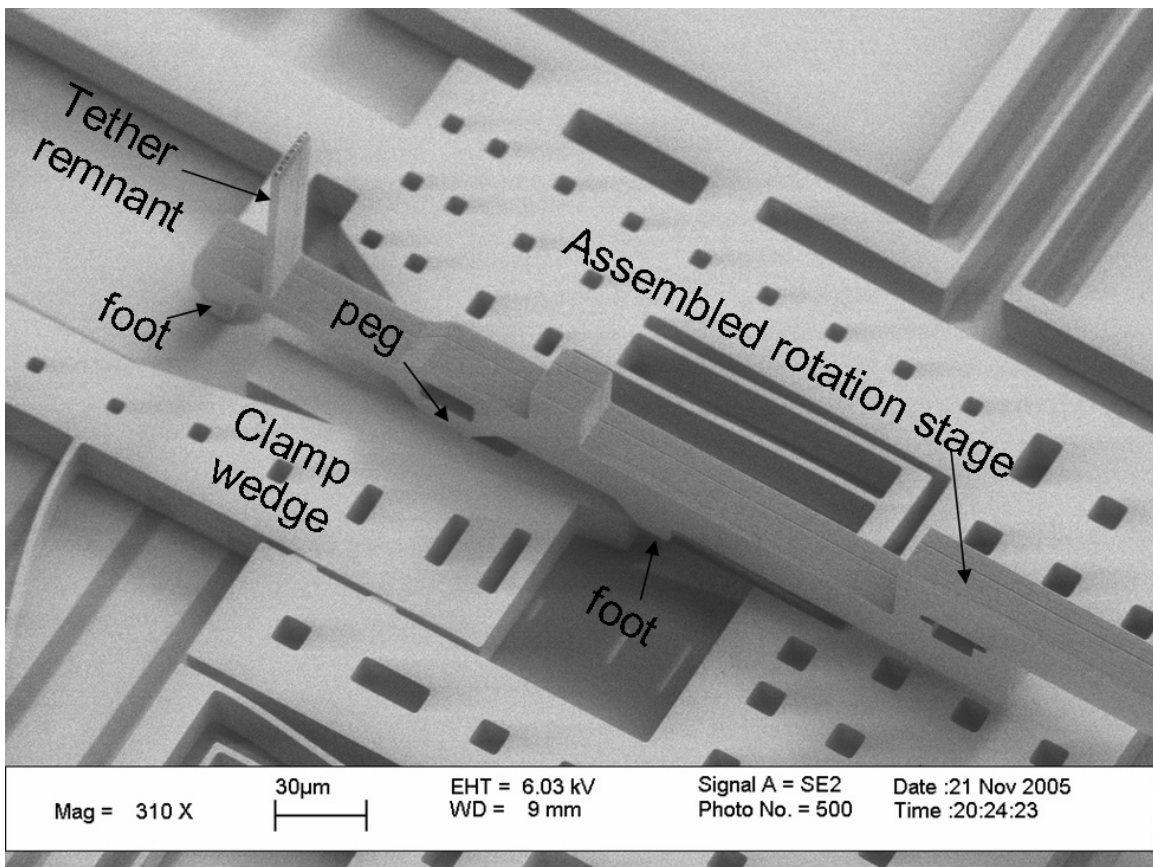

Figure 37. Close-up electron micrograph of interface between assembled part and moving clamp. Visible is the back end of the rotation stage mechanism and details of how it inserts into the fully-suspended clamp.

114

Design features of the moving clamps are described in detail in section 4.2.2 above (pg. 55). Additional problems are encountered when assembling the two-axis rotation stage. Because the two ends of the structure are only connected through the two long, flexible flexures in series, the two ends of the part are difficult to maneuver as a unit. As soon as one end of the part makes contact with the clamp, the other end can continue to move under the influence of the ortho-gripper. This effect makes it difficult to assemble both ends of the part into their respective clamps simultaneously with just the ortho-gripper. A second probe tip can help to hold the peg on the distal end of the part in its clamp and flush against the device layer of the chip. This involves a painstaking alignment process, switching back and forth between the overhead view and the side view seen through the 45-degree mirror. Future investigators should spend some effort to find an easier, more reliable method for performing this assembly step.

8.3.4. Testing

DC deflection angles are obtained using two testing methods. The first involves using the 45-degree mirror (Figure 13) to image the side of the assembled rotation stage. A video camera is used to record still images and video of the structure. Quantitative data is obtained using image processing software to measure angles between substructures of the rotation stage in the undeflected and deflected states. This approach is hampered by the limited resolution of the camera, resulting in low precision measurements (on the order of ½ degree) of the device deflection. In addition, only one axis of rotation can be measured at a time, so cross-axis coupling at large deflections cannot be observed easily.

The second approach to measuring DC deflection is to bounce a laser beam off of the mirror mounting platform. The area of this reflector is very small, resulting in very little light available to be measured. Consequently, the measurements of beam deflection had to be performed in the dark. This made recording the data difficult as glow-in-the-dark ink was not utilized. In addition, the small size of the reflector results in a large amount of diffraction of the laser beam, stretching out the spot in the torsion direction. Despite these limitations, the sensitivity of this approach far exceeded that of the camera-based approach, and in addition, motion about both axes could be measured simultaneously.

Results obtained using both techniques are shown in (Figure 38). Maximum DC deflection angles are less than 2% away from the predicted actuator displacement-limited maximum in the flexural axis, and approximately 10% away from the predicted rotation value in the torsional axis. For comparison, the best tip/tilt mirror designed for large deflection angles using a large reflector area (Milanovic, McCormick et al. 2004) is also shown on the graph. They achieve a smaller deflection angle at much higher voltages. On the other hand, their dynamic performance is well characterized and very good out to several kilohertz, while the dynamic performance of this rotation stage with a comparable reflecting surface mounted on it is currently unknown. That said, this new rotation stage sets the new "high water" mark in terms of DC optical deflection.
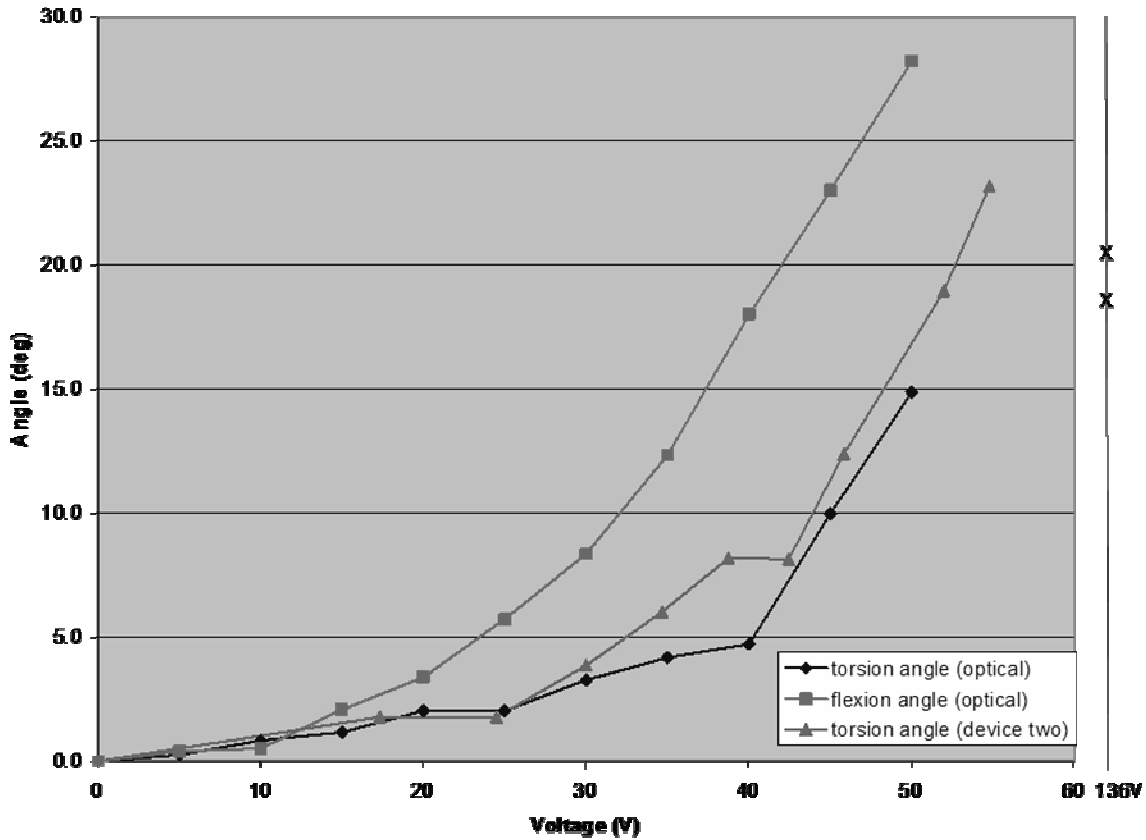
Figure 38. Optical deflection angle of a laser beam bounced off of the top of the rotation stage. Mechanical rotation angle is one half the optical angle shown here. Two devices were tested: Device one had two functional actuators (pull-mode about torsion and flex axes) and was tested by bouncing the laser beam off of the top of the device. The second device had only one functional axis (torsion) and was tested using the side-view mirror and video camera (no laser). For comparison, the best quadrant of operation of the highest-performing large tip/tilt mirror in the literature is shown (Milanovic, McCormick et al. 2004). When fully functional, this rotation stage design will achieve larger static deflection angles at lower voltages.

Future work includes obtaining measurement data from all four quadrants (push and pull modes for both axes), gluing a reflector onto the rotation stage, and measuring the dynamics of the rotation stage/mirror combination using a laser Doppler vibrometer.

8.4. Discussion

117

This two-axis rotation stage can achieve more DC deflection angle than any other flexure-based 2-axis rotation stage presently reported in the literature. Because it performs this rotation at the displacement limit of the actuator instead of the force limit, this approach promises to be able to rotate even further in the future. All that is required is a slight reduction in the length of the moment arm separating the two flexures. The limit of DC deflection angle required for any optical application is +/-45 degrees of mechanical rotation. At that point, a laser beam can be steered to any point in the hemisphere. The rotation stage demonstrated in this section demonstrated just under 14 degrees of mechanical rotation about the flexural axis and just under 12 degrees about the torsional axis. If the opposing actuators can be made to work, it is expected that this structure will achieve the same deflection in the other direction as well, doubling the total angular deflection. Extending the angular deflection by the factor of three required to solve the DC pointing problem once and for all is not a clear path however. The equations derived here depend on the deflection angles remaining small (~15 degrees). At larger angles, it is expected that the behavior will deviate quite significantly from a simple linear scaling of forces and spring constants.

The static performance of this rotation stage is quite good. However, the dynamic performance is currently unknown and is likely to not be very good due to asymmetries in the flexure design. It is expected that this will manifest through a strong coupling between the rotation (both flexion and torsion) and bounce modes. Because the bounce mode is expected at a lower frequency than either of the rotational modes, the dynamics

of the bounce mode will not be able to be neglected. A different design might be required.

Future work should focus on several key areas: first and foremost, the dynamics of the rotation stage with a mirror face sheet attached to it is an area that should be investigated. The optimization of that mirror face sheet for flatness and moment of inertia is another topic that should be pursued. Finally, in order to place this device into a system, a closed loop control system should be implemented.

9. CONCLUSION

This dissertation describes a serial pick-and-place microassembly technique based on parts manufactured in a single-mask, Silicon On Insulator (SOI) process. A simple fabrication process that starts with standard double-sided polished wafers and ends with anti-stiction-coated micro-parts is presented. This process is used to create micro-mechanisms that are rotated ninety degrees out of plane and affixed elsewhere on the chip, a novel tool (the "ortho-gripper") for handling the parts and rotating them reliably, and a new zero insertion-force socket that forms a good mechanical and electrical connection between the chip and the assembled part. Together, these form a novel manufacturing process that extends the capabilities of single-mask SOI processing and allows it to be used in applications that require large amounts of out-of-plane motion.

This manufacturing process is more than just a novelty. It is used to make devices that would not otherwise be possible in a single-mask process. Multi-level electrical wiring is demonstrated. One- and Two-axis rotation stages are demonstrated, with the two-axis rotation stage setting the new high water mark for the largest static pointing angle of any two-axis rotation stage currently in the literature.

In general, SOI-based processes are exceptionally good at making in-plane micromechanisms. These consist of actuators, sensors, and rigid and flexible beam elements. However, SOI-based micromechanisms have difficulty developing large out-of-plane motions. This is due to the alignment of the forces generated by in-plane

actuators and the flexures that guide their motion. A method of making mechanical elements that skew the vertical symmetry of the SOI process is required in order for out-of-plane moments and forces to develop.

This approach solves the symmetry problem. It is not unique in doing so – other processes have added beams above and/or below the centerline of the device layer, or thinned down regions of the device layer from the front or backside of the wafer, or tilted portions of the device layer up or down. All these approaches break the vertical symmetry at the expense of additional wafer-level or die-level process complexity.   The manufacturing process described here takes a different approach, pushing all the complexity from the wafer/die-level processing to the post-process. This results in a manufacturing process that can be completed in a weekend. At the same time, it provides additional flexibility by allowing the designer to create libraries of parts that can be assembled into a custom device whose performance characteristics can be determined during assembly.

Bibliography

BEI Kimco Inc. (2002) "Voice Coil Actuators Application Guide." **Volume**, 12 DOI:

Bellew, C. L., S. Hollar and K. S. J. Pister (2003). <u>An SOI process for fabrication of solar cells, transistors and electrostatic actuators</u>. IEEE International Solid-State Sensors and Actuators Conference. Boston, MA.

Brosnihan, T. J., J. M. Bustillo, A. P. Pisano and R. T. Howe (1997). <u>Embedded interconnect and electrical isolation for high-aspect-ratio, SOI inertial instruments</u>. Proceedings of International Solid State Sensors and Actuators Conference (Transducers '97). Chicago, IL.

Cagdaser, B. and B. E. Boser (2005). <u>Resonant drive for stabilizing parallel-plate actuators beyond the pull-in point</u>. TRANSDUCERS '05. The 13th International Conference on Solid-State Sensors, Actuators and Microsystems. Digest of Technical Papers. Seoul, South Korea. Korean Sensors Soc. 5-9 June 2005.

Cagdaser, B. and B. E. Boser (2005). <u>Resonant drive: sense and high voltage electrostatic drive using single MEMS capacitor and low voltage electronics</u>. 18th IEEE International Conference on Micro Electro Mechanical Systems. Miami Beach, FL.

Cagdaser, B., A. Jog, M. Last, B. S. Leibowitz, L. Zhou, E. Shelton, K. S. J. Pister and B. E. Boser (2004). Capacitive Sense Feedback Control for MEMS Beam Steering Mirrors. Hilton Head 2004. Hilton Head, S.C.

Chang, W. C., C. G. Keller, E. A. Hawkes and D. W. Sretavan (2005). Microdevice components for a cellular microsurgery suite. Transducers 2005. Seoul, Korea, IEEE. **1:** 209-212 Vol. 1.

Che-Heung, K., J. Hee-Moon, J. Jong-Up and K. Yong-Kweon (2003). "Silicon micro XY-stage with a large area shuttle and no-etching holes for SPM-based data storage." Journal of Microelectromechanical Systems **12**(4): 470-8.

Christenson, T. R., H. Guckel, K. J. Skrobis and T. S. Jung (1992). Preliminary results for a planar microdynamometer. Hilton Head Island, SC.

Clark, W. A., R. T. Howe and R. Horowitz (1996). Surface micromachined Z-axis vibratory rate gyroscope. Technical Digest Solid-State Sensor and Actuator Workshop. Hilton Head Island, SC.

Cohn, M. B. (1997). Assembly Techniques for Microelectromechanical Systems. Engineering -- Electrical Engineering and Computer Sciences. Berkeley, University of California, Berkeley. **PhD:** 150.

Cohn, M. B., K. F. Bohringer, J. M. Noworolski, A. Singh, C. G. Keller, K. Y. Goldberg and R. T. Howe (1998). Microassembly technologies for MEMS. Micromachining and Microfabrication Process Technology IV. Santa Clara, CA.

Conant, R. A., J. T. Nee, K. Y. Lau and R. S. Muller (2000). Dynamic deformation of scanning mirrors. 2000 IEEE/LEOS International Conference on Optical MEMS. Kauai, HI.

Dechev, N., W. L. Cleghorn and J. K. Mills (2002). "Micro-assembly of microelectromechanical components into 3-D MEMS." Canadian Journal of Electrical & Computer Engineering 27(1): 7-15.

Dechev, N., W. L. Cleghorn and J. K. Mills (2003). Microassembly of 3D MEMS structures utilizing a MEMS microgripper with a robotic manipulator. IEEE International Conference on Robotics and Automation. IEEE ICRA 2003 Conference Proceedings. Taipei, Taiwan. IEEE Robotics & Autom. Soc. Nat. Sci. Council, Taiwan,. Ministr. Educ., Taiwan. 14-19 Sept. 2003.

Dechev, N., W. L. Cleghorn and J. K. Mills (2004). "Microassembly of 3-D microstructures using a compliant, passive microgripper." Journal of Microelectromechanical Systems 13(2): 176-89.

Dechev, N., W. L. Cleghorn and J. K. Mills (2004). <u>Tether and joint design for microcomponents used in microassembly of 3D microstructures</u>. MEMS/MOEMS Components and their Applications. San Jose, CA,USA. 26-27 Jan. 2004.

Drexler, E. K. (1986). <u>Engines of Creation: The Coming Era of Nanotechnology</u>. New York, Anchor Books, Doubleday.

Dutto, D. J., D. F. Hoyt, H. M. Clayton, E. A. Cogger and W. S. J (2004). "Moments and power generated by the horse (Equus caballus) hind limb during jumping." <u>Journal of Experimental Biology</u>(207(pt 4)): 667-74.

Fearing, R. S. (1995). <u>Survey of sticking effects for micro parts handling</u>. Proceedings 1995 IEEE/RSJ International Conference on Intelligent Robots and Systems. Human Robot Interaction and Cooperative Robots. Pittsburgh, PA.

Fischer, K. and H. Guckel (1998). "Long throw linear magnetic actuators stackable to one millimeter of structural height." <u>Microsystem Technologies</u> **4**(4): 180-3.

Franke, A. E., M. Heck John, K. Tsu-Jae and R. T. Howe (2003). "Polycrystalline silicon-germanium films for integrated microsystems." <u>Journal of Microelectromechanical Systems</u> **12**(2): 160-71.

Grade, J. D., H. Jerman and T. W. Kenny (2000). A large-deflection electrostatic actuator for optical switching applications. Technical Digest. Solid-State Sensor and Actuator Workshop. Hilton Head Island, SC.

Grade, J. D., H. Jerman and T. W. Kenny (2003). "Design of large deflection electrostatic actuators." Journal of Microelectromechanical Systems **12**(3): 335-43.

Guckel, H., T. Christenson and K. Skrobis (1992). Fabrication and testing of metal micromechanisms with rotational and translational motion. Sensors and Sensor Systems for Guidance and Navigation II. Orlando, FL.

Guckel, H., T. Christenson and K. Skrobis (1992). "Metal micromechanisms via deep X-ray lithography, electroplating and assembly." Journal of Micromechanics & Microengineering **2**(4): 225-8.

Guckel, H., T. R. Christenson, K. J. Skrobis and J. Klein (1994). "Planar rotational magnetic micromotors." International Journal of Applied Electromagnetics in Materials **4**(4): 377-82.

Guckel, H., K. J. Skrobis, T. R. Christenson, J. Klein, S. Han, B. Choi, E. G. Lovell and T. W. Chapman (1991). "Fabrication and testing of the planar magnetic micromotor." Journal of Micromechanics & Microengineering **1**(3): 135-8.

Hirano, T., T. Furuhata, K. J. Gabriel and H. Fujita (1991). Operation of sub-micron gap electrostatic comb-drive actuators. San Francisco, CA.

Hirano, T., T. Furuhata, K. J. Gabriel and H. Fujita (1992). "Design, fabrication, and operation of submicron gap comb-drive microactuators." Journal of Microelectromechanical Systems **1**(1): 52-9.

Hollar, S., A. Flynn, C. Bellew and K. S. J. Pister (2003). Solar powered 10 mg silicon robot. Proceedings IEEE Sixteenth Annual International Conference on Micro Electro Mechanical Systems. Kyoto, Japan. 19-23 Jan. 2003.

Huang, T., E. Nilsen, M. Ellis, K. Kabseog, K. Tsui, G. Skidmore, C. Goldsmith, A. Nallani and J. B. Lee (2003). 3-D, self-aligned, micro-assembled, electrical interconnects for heterogeneous integration [MEMS example]. MEMS Components and Applications for Industry, Automobiles, Aerospace, and Communication II. San Jose, CA.

Jensen, B. D., M. P. de Boer, N. D. Masters, F. Bitsie and D. A. LaVan (2001). "Interferometry of Actuated Microcantilevers to Determine Material Properties and Test Structure Nonidealities in MEMS." Journal of Microelectromechanical Systems **10**(3): 336-346.

Keller, C. G. (2005). "MEMS Precision Instruments Website." from http://www.memspi.com.

Kulicke & Soffa inc (2005). Max◯mPlus Datasheet.

Last, M. and K. S. J. Pister (1999). <u>2-DOF actuated micromirror designed for large DC deflection</u>. Proceedings of MOEMS 99. Mainz, Germany. Controlware. CSEM. Haas Laser. IOT Integrierte Optik Technol. Schott. Siemens Electromech. Components. et al. 30 Aug.-1 Sept. 1999.

Legtenberg, R., A. W. Groeneveld and M. Elwenspoek (1996). "Comb-drive actuators for large displacements." <u>Journal of Micromechanics & Microengineering</u> **6**(3): 320-9.

Lindsay, W., D. Teasdale, V. Milanovic, K. Pister and C. Fernandez-Pello (2001). <u>Thrust and electrical power from solid propellant microrockets.</u> . Technical Digest. MEMS 2001. 14th IEEE International Conference on Micro Electro Mechanical Systems. Interlaken, Switzerland. IEEE Robotics & Autom. Soc. 21-25 Jan. 2001.

Ljungberg, K., A. Soderbarg and Y. Backlund (1993). "Spontaneous Bonding of Hydrophobic Silicon Surfaces." <u>Applied Physics Letters</u> **62**(12): 1362-1364.

Lumbantobing, A. and K. Komvopoulos (2005). "Static friction in polysilicon surface micromachines." <u>Journal of Microelectromechanical Systems</u> **14**(4): 651-63.

Maharbiz, M. M., M. B. Cohn, R. T. Howe, R. Horowitz and A. P. Pisano (1999). <u>Batch micropackaging by compression-bonded wafer-wafer transfer</u>. Proceedings of 12th International Workshop on Micro Electro Mechanical Systems - MEMS. Orlando, FL.

Maharbiz, M. M., R. T. Howe and K. S. J. Pister (1999). Batch Transfer Assembly of Micro-Components Onto Surface and SOI MEMS. Transducers. Sendai, Japan.

Marty, F., B. Saadany, T. Bourouina, Y. Mita and T. Shibata (2005). High Aspect Ratio Nano-structures (HARNS) for Photonic MEMS Based on Vertical DBR Architecture. Transducers 2005. Seoul, Korea.

Milanovic, V., M. Last and K. S. J. Pister (2001). Torsional micromirrors with lateral actuators. Proceedings of 11th International Conference on Solid State Sensors and Actuators Transducers '01/Eurosensors XV. Munich, Germany. 10-14 June 2001.

Milanovic, V., M. Last and K. S. J. Pister (2003). "Laterally actuated torsional micromirrors for large static deflection." IEEE Photonics Technology Letters **15**(2): 245-7.

Milanovic, V., D. T. McCormick and G. A. Matus (2004). "Gimbal-less Monolithic Silicon Actuators for Tip-Tilt-Piston Micromirror Applications." IEEE Journal of Select Topics in Quantum Electronics **10**(3): 462-471.

Nguyen, C. (1994). Micromechanical Signal Processors. EECS. Berkeley, University of California, Berkeley. **PhD**.

Ono, T., D. Y. Sim and M. Esashi (2000). "Micro-discharge and Electric Breakdown in a Micro-Gap." Journal of Micromechanics & Microengineering **10**: 445-451.

Pan, F., J. Kubby, E. Peeters, A. T. Tran and S. Mukherjee (1998). Squeeze Film Damping Effect on the Dynamic Response of a MEMS Torsion Mirror. International Conference on Modeling and Simulation of Microsystems, Nano Science and Technology Institute.

Sarajlic, E., E. Berenschot, G. Krijnen and M. Elwenspoek (2003). Low volume, large force (>1mN) and nanometer resolution, electrostatic microactuator for low displacement applications. NanoTech 2003. San Francisco, CA.

Sarajlic, E., E. Berenschot, G. Krijnen and M. Elwenspoek (2003). Versatile trench isolation technology for the fabrication of microactuators. Micro- and Nano-Engineering 2002. 28th International Conference on Micro- and Nano-Engineering. Lugano, Switzerland. 16-19 Sept. 2002.

Sarajlic, E., E. Berenschot, N. Tas, H. Fujita, G. Krijnen and M. Elwenspoek (2005). High performance bidirectional electrostatic inchworm motor fabricated by trench isolation technology. TRANSDUCERS '05. The 13th International Conference on Solid-State Sensors, Actuators and Microsystems. Digest of Technical Papers. Seoul, South Korea. Korean Sensors Soc. 5-9 June 2005.

Sarajlic, E., M. J. de Boer, H. V. Jansen, N. Arnal, M. Puech, G. Krijnen and M. Elwenspoek (2004). "Advanced plasma processing combined with trench isolation technology for fabrication and fast prototyping of high aspect ratio MEMS in standard silicon wafers." Journal of Micromechanics & Microengineering 14(9): S70-5.

130

Skidmore, G., M. Ellis, A. Geisberger, K. Tsui, R. Saini, T. Huang and J. Randall (2003). Parallel assembly of microsystems using Si micro electro mechanical systems. Micro- and Nano-Engineering 2002. 28th International Conference on Micro- and Nano-Engineering. Lugano, Switzerland. 16-19 Sept. 2002.

Skidmore, G., M. Ellis, A. Geisberger, K. Tsui, K. Tuck, R. Saini, T. Udeshi, M. Nolan, R. Stallcup and J. Von Ehr, II (2004). Assembly technology across multiple length scales from the micro-scale to the nano-scale. 17th IEEE International Conference on Micro Electro Mechanical Systems. Maastricht MEMS 2004 Technical Digest. Maastricht, Netherlands. IEEE. Robotics and Automation Soc. 25-29 Jan. 2004.

Smalley, R. E. (2003). Nanofallacies: Of Chemistry, Love, and Nanobots. Scientific American. **285:** 76-78.

Tang, W. C., C. Tu, H. Nguyen and R. T. Howe (1989). "Laterally driven polysilicon resonant microstructures." Sensors & Actuators **20**(1-2): 25-32.

Thompson, J. A. and R. S. Fearing (2001). Automating microassembly with ortho-tweezers and force sensing. Proceedings of RSJ/IEEE International Conference on Intelligent Robots and Systems. Maui, HI.

Tsuchiya, T. (2005). "Tensile testing of silicon thin films." Fatigue & Fracture of Engineering Materials & Structures **28**(8): 665-74.

Warneke, B. A., M. D. Scott, B. S. Leibowitz, Z. Lixia, C. L. Bellew, J. A. Chediak, J. M. Kahn, B. E. Boser and K. S. J. Pister (2002). An autonomous 16 mm/sup 3/ solar-powered node for distributed wireless sensor networks. Proceedings of IEEE SENSORS 2002. Orlando, FL.

Yeh, R. (2001). Articulated Mechanisms and Electrostatic Actuators for Autonomous Microrobots. Engineering - Electrical Engineering and Computer Science. Berkeley, University of California, Berkeley. **PhD:** 130.

Young, W. C. (1989). Roark's Formulas for Stress & Strain. New York, McGraw-Hill, Inc.

Zhou, L., M. Last, V. Milanovic, J. M. Kahn and K. S. J. Pister (2003). Two-axis scanning mirror for free-space optical communication between UAVs. 2003 IEEE/LEOS International Conference on Optical MEMS. Waikoloa, HI.

Zyvex Inc (2004). Zyvex Microassembly**:** Microassembly capabilities at Zyvex. Videos and SEMs of assembled structures and assembly robot.

Appendix A: High Voltage Electrostatic Comb Drive Actuators

A.1 Introduction

Electrostatic comb drive actuators are useful in microsystems due to their ease of fabrication, high precision motion, and favorable scaling characteristics. These scaling characteristics were first pointed out in (Tang, Tu et al. 1989; Hirano, Furuhata et al. 1991, 1992) and indicate that the energy density of the actuator increases as the comb fingers and gaps are scaled downward in size. Because of these characteristics, comb drives have found use in micro-optical systems (Last and Pister 1999; Maharbiz et al. 1999; Milanovic, Last et al. 2001, 2003; Zhou et al. 2003), inertial sensors (Clark, Howe et al. 1996), motion stages for scanning probe microscopes (Che-Heung, Hee-Moon et al. 2003), and numerous other applications. In these cases, the actuator performance was limited by actuator instability caused by inadequate stiffness of the actuator suspension. This actuator instability is addressed in (Grade, Jerman et al. 2003), and design rules that ensure stability over large actuator strokes are presented. However, (Grade et al. 2003) did not explicitly deal with the effects of instability introduced by sideways flexing of the comb fingers themselves. When combined with the breakdown voltage of air, this presents an upper bound on the energy density of the current design of comb drive actuators. All constraints necessary for generating comb drive actuators that are optimized for energy density are presented.

In many applications today, this maximum energy density is never approached due to the requirement of an extremely high drive voltage. This high voltage, in the range of 200-300V, has been seen as a problem due to the incompatibility with CMOS electronics and

the difficulty of sensing the position of the comb drive without overwhelming the sense signal with feedthrough from the drive electronics. Two main approaches are discussed in this appendix: Electromechanical Amplitude Modulation (EAM) and Resonant Drive. Either approach can be applied to mitigate these problems.

Actuators can be designed to be uni- or bi-directional. Each approach has tradeoffs in terms of suspension, comb finger, and control system design. These tradeoffs are enumerated in this appendix.

A.2 Electrostatic Actuation

Electrostatic actuators convert potential energy stored in a battery to physical motion by minimizing the potential energy of a capacitor with moving electrodes. The work done by an electrostatic actuator as it moves from one position to another under voltage control is simply the change in potential energy of the capacitor:

$$\Delta E = \frac{1}{2}\Delta C V^2 \tag{A.1}$$

For a comb drive actuator moving along the y-axis, ΔC is related to the change in overlap area:

$$\Delta C = N_f \varepsilon_0 \frac{h\left(y_f - y_i\right)}{g} \tag{A.2}$$

Where N_f is the number of fingers, ε_0 is the permittivity of the gap material (8.852e-12

F/m for vacuum or air), h is the height of the comb fingers, g is the gap between comb

fingers, and y_f and y_i are the final and initial comb drive positions respectively (Figure

39). Typically, the number of fingers in a comb drive is in the dozens to hundreds.
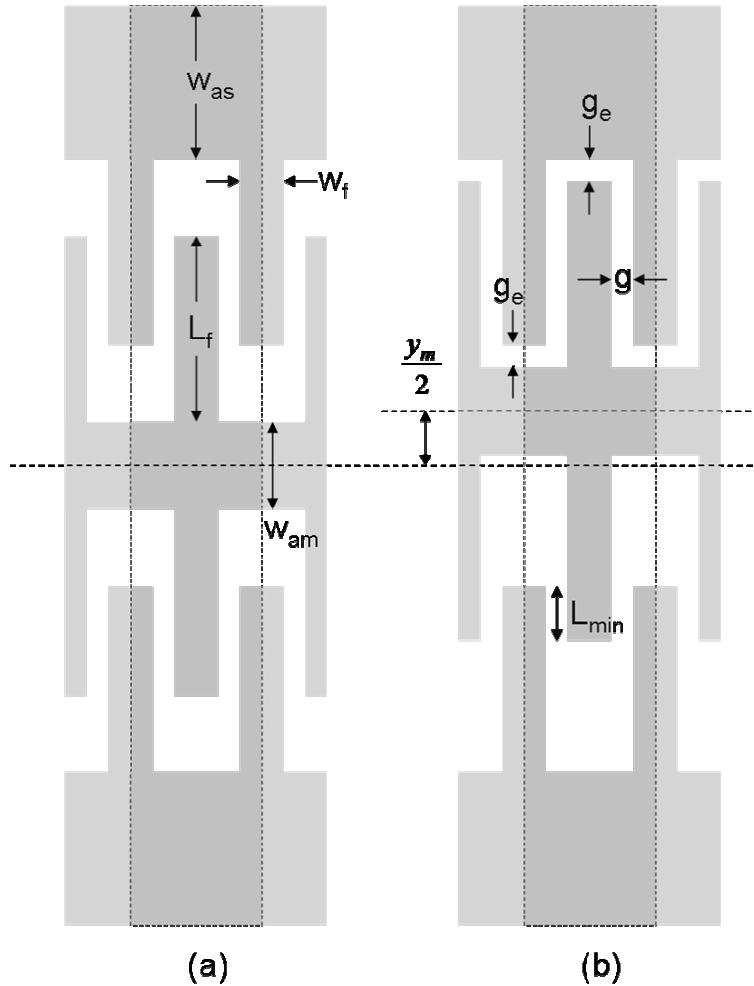

Figure 39. Comb drive actuator unit cell used in energy density calculations. Comb drive shown (a) at rest (zero applied voltage), and (b) maximally displaced. Unit cell is shaded.

For gap-closing actuators moving along the y-axis, ΔC is related to the change in the gap

between fingers:

$$\Delta C = N_f \varepsilon_0 \frac{h L_f}{\left(y_f - y_i\right)} \qquad \text{(A.3)}$$

Where N_f is the number of fingers, ε_0 is the permittivity of the gap material (8.852e-12 F/m for vacuum or air), h is the height of the comb fingers, L_f is the distance that the comb fingers overlap, and y_f and y_i are the final and initial comb drive positions respectively. The number of fingers can range from a single large plate, to an array of comb fingers numbering in the dozens to hundreds.

In both the case of the comb drive and the gap-closing actuator, we will define longitudinal (y-axis) motion as the desired direction of motion of the actuator, while lateral (x-axis) motion will be defined as the axis orthogonal to that desired direction of motion (but still in-plane).

The energy stored in the capacitor increases quadratically with voltage. By increasing the maximum voltage at which the actuator operates, the energy stored in the capacitor can be increased quite dramatically. The only physical limit to increasing this voltage comes from the breakdown voltage of the gap. In air, this is related to both the pressure and the gap size by the Paschen curve (Figure 40)(Ono, Sim et al. 2000). Although the Paschen curve seems to imply that at submicron gaps one could continue scaling the voltage upwards, in practice this will not be the case. This is because even if the gap between the comb fingers is smaller than the gap at where the breakdown voltage is at its minimum

(implying a larger breakdown voltage), there will almost always be a larger gap for which

the breakdown voltage is lower. Examples of such larger gaps are the one between non-

adjacent comb fingers (by arcing above or underneath the fingers) and the gap between

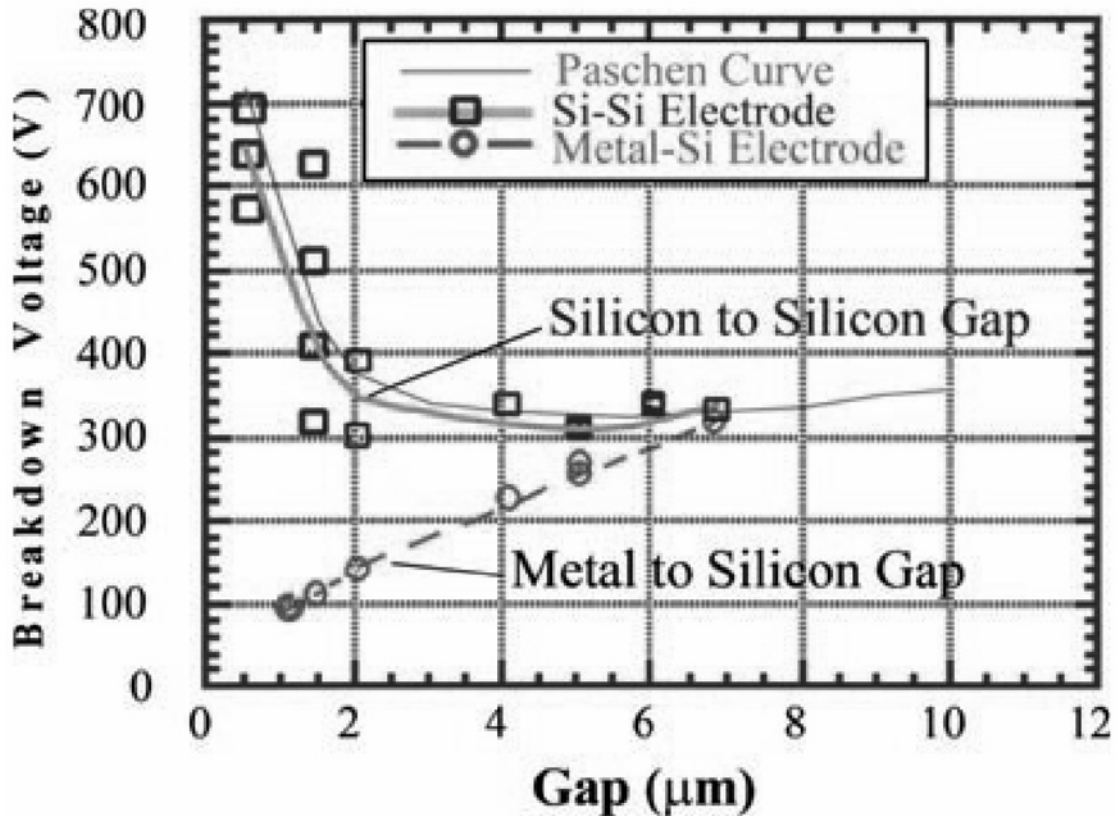the comb finger tip and the sidewall further down the gap.

Figure 40. Breakdown voltage as a function of gap distance for air at standard temperature and pressure. Gaps between silicon electrodes follow the paschen curve very closely and reach a minimum at approximately 300V. This provides an upper bound on voltage scaling for high energy-density electrostatic actuators.

A.3 Stability

If the control system driving these actuators is a simple voltage controller, both comb

drive and gap-closing actuators are limited in their range of motion by pull-in

instabilities. The gap-closing actuator has two main instabilities: 1) longitudinal snap-through, where the force of the electrostatic actuator overwhelms the actuator suspension's ability to resist motion, and 2) finger pull-in, where the stiffness of the individual comb fingers are not sufficient to keep opposing fingers from touching.

There are four main instabilities that affect comb drive actuators. All of these involve the parasitic gap-closing actuator formed between the sides of the comb fingers. Under normal conditions, the forces on each side of the comb fingers balance each other; however this is not always the case. These instabilities are: 1) longitudinal pull-in, where the gap-closing actuator formed between the end of a comb finger and the base of the opposing comb becomes non-negligible, 2) lateral pull-in, where the stiffness of the actuator suspension against x-axis motion is less than required to keep the actuator centered, 3) rotational pull-in, where the stiffness of the suspension against in-plane rotation is less than required to keep the actuator centered, and 4) lateral finger pull-in, where the stiffness of the fingers themselves are not sufficient to remain stable.

Instability occurs when the total system energy as a function of displacement is concave down, or $\dfrac{\partial^2 E}{\partial \vec{r}^2} < 0$ where r is the direction of motion. In this case the energy of the system is the combination of the potential energy of the actuator summed with the energy stored in the actuator suspension or comb finger, or $E = -\dfrac{1}{2}CV_m^2 + \dfrac{1}{2}k_r r^2$, where k_r is the spring constant in the $\vec{r}$ direction and r is the magnitude of the displacement in the $\vec{r}$ direction. Since all of the instabilities of interest are the gap-closing variety, we use
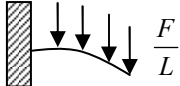
$C = N_f \varepsilon_0 h\, L_{mo} \left[\left(g_1 + r \right)^{-1} + \left(g_2 - r \right)^{-1} \right]$, where L_{mo} is the maximum length the fingers
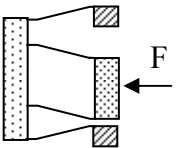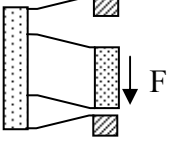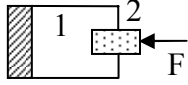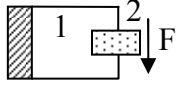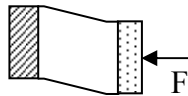
overlap by ($L_{mo} = L_f - g_e$), g_2 is the gap being closed, and g_1 is the back-gap. This

creates the stability requirement:

$$k_r > \frac{1}{2} N_f \varepsilon_0 h L_{mo} V_m^2 \left(\frac{2}{\left(g_1 + r \right)^3} + \frac{2}{\left(g_2 - r \right)^3} \right) \qquad \text{(A.4)}$$

In each of the instabilities mentioned, the main customization required is to choose the

appropriate value for k_r based on the calculation being done. Some commonly used

equations for spring constants found in common electrostatic actuators are listed in the

table below:

Actuator Component	Beam end conditions	Diagram	Formula	Reference
Comb drive finger, lateral stiffness	Fixed-free, distributed load		$\dfrac{k_l}{L} = \dfrac{2Ehb^3}{3L^4}$	[Roark's] Table 3, eqn 2(a)

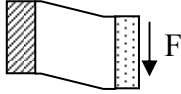| Double-folded flexure, axial stiffness	Fixed-guided, point load		$$k_a = \frac{50Ehb^3}{L\left(25b^2 + 9y^2\right)}$$	Elwenspoek, eq (11, 23)
Double-folded flexure, lateral stiffness	Fixed-guided, point load		$$k_l = \frac{Ehb^3}{8L^3}$$	Derived from [Roark's] Table 3, eqn 1(a)
Crab-leg flexure, axial stiffness	Fixed-guided, point load		$$k_a = \frac{12EI_2}{L_2^3}\left\{\frac{L_1I_2 + 2L_2I_1}{2L_1I_2 + L_2I_1}\right\}$$	Elwenspoek, eq (8)
Crab-leg flexure, lateral stiffness	Fixed-guided, point load		$$k_l = \frac{24EI_1}{L_1^3}\left\{\frac{L_1I_2 + L_2I_1}{L_1I_2 + 4L_2I_1}\right\}$$	Elwenspoek, eq (9)
Parallel bar flexure, axial stiffness	Fixed-guided, point load		$$k_a = \frac{2Ebh}{L}, \text{ for zero y-axis deflections}$$	Derived from stress/strain equation

Parallel bar flexure, lateral stiffness	Fixed-guided, point load			$k_l = \dfrac{Ehb^3}{4L^3}$	[Roark's] Table 3, eqn 1(a)

Table 1: Commonly used spring formulas. In each beam, $I = \dfrac{hb^3}{12}$, where h is the height of the beam and b is the width of the beam in the direction of bending. L is the length of the beam, y is the displacement of the actuator in the longitudinal direction, and E is the Young's modulus of the spring material (~160GPa for silicon).

In order to ensure stability over the desired range of travel, the designer must ensure that the instability condition in equation (A.4) above is not met for values of r that might be encountered during operation of the actuator. When designing a comb drive, this calculation would typically need to be repeated for the case of lateral pull-in of the actuator, longitudinal snap-through of the comb fingers at the end of their travel, the lateral pull-in of the comb fingers, and potentially rotational stability of the actuator. The stability of an actuator against lateral pull-in is derived in (Hirano et al. 1991; Grade et al. 2003). The actuator is stable against side pull-in if:

$$\frac{2k_a y^2}{k_l g^2} \frac{1 + 3\left(\dfrac{x}{g}\right)^2}{\left(1 - \left(\dfrac{x}{g}\right)^2\right)^2} \le 1 \qquad \text{(A.5)}$$

Where k_a is the axial spring constant of the suspension, k_l is the longitudinal spring constant of the suspension, y is the longitudinal displacement of the actuator, x is the lateral displacement of the actuator (corresponding to an axial displacement of the suspension), and g is the gap between comb fingers.

Rotational pull-in is related to lateral pull-in, except instead of the actuator pulling directly to one side or another, the top half of the actuator will pull to the side in one direction and the bottom half will pull in the opposite direction. If an actuator is designed such that it is stable against lateral pull-in, it will be stable against rotational pull-in as well if the suspension is attached to the actuator above and below the combs (Figure 41). This is because the torque generated by the actuator is bounded by the lateral force times the distance along the y-axis of the most distant comb finger. The restoring torque provided by the suspension is exactly the lateral force times the distance along the y-axis of the attachment point of the suspension. Therefore, if the actuator is stable against lateral pull-in and the distance of the attachment point is greater than the distance of the most remote comb finger, the actuator will be stable against rotational pull-in as well.
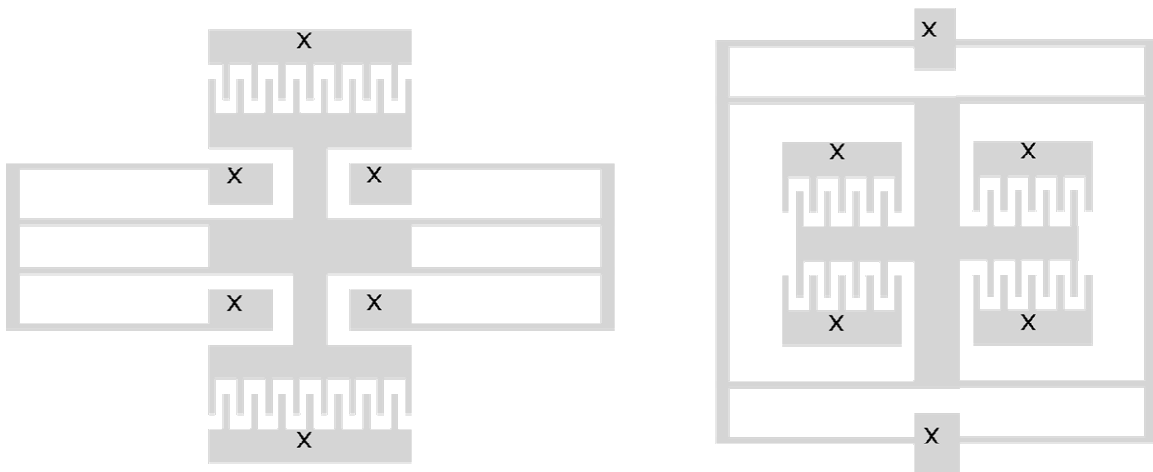
Figure 41. If a comb drive is stable against lateral pull-in, it will be stable against rotational pull-in as well if the suspension is attached above and below the active region. Comb drive on left may or may not be rotationally stable even if laterally stable, whereas comb drive on right will be stable against rotational pull-in if laterally stable. "x"s denote anchored parts.

The stability of a comb finger against lateral pull-in is derived by inserting the spring constant formula for a cantilevered beam under a distributed transverse load from the table above into the stability equation (A.4). Assuming a uniform load over the length of the comb finger, the energy equation is modified to fit the distributed nature of the electrostatic force, yielding the instability equation:

$$\frac{3L_{mo}^4 \varepsilon_0 h V_m^2}{2Eb^3}\left(\frac{1}{(g_1+r)^3}+\frac{1}{(g_2-r)^3}\right) \geq 1 \tag{A.6}$$

Where in this case $g_1 = g_2$, r will be either the positive or negative x direction, b is the width of the comb finger W_f and $L_{mo} = L_f - g_e$. This results in the design constraint on the dimension W_f:

$$W_f \geq \sqrt[3]{\frac{3}{2}\left(\frac{\varepsilon_0}{E}\right)V_m^2\left(L_f - g_e\right)L_f^3\left[\frac{1}{(g+x)^3}+\frac{1}{(g-x)^3}\right]} \tag{A.7}$$

The parameter g_e is chosen to ensure the stability of the actuator against longitudinal pull-in at the end of travel. In this case, the longitudinal stiffness of the actuator suspension k_y is what balances the longitudinal actuator force. Since the back-gap is so large

143

($g_1 = y_m + g_e$), the back-gap capacitance can be neglected, leading to the instability relation:

$$\frac{2N_f\varepsilon_0 hW_f V_m^2}{k_y(g_e - y)^3} \geq 1 \tag{A.8}$$

This yields the design constraint on the dimension g_e:

$$g_e \geq \left(\frac{2N_f\varepsilon_0 hW_f V_m^2}{k_y}\right)^{\frac{1}{3}} + y \tag{A.9}$$

Which is solved simultaneously with the equation for W_f. This equation actually underestimates the stability of the actuator against longitudinal snap-through because it does not include the stiffness of the load the actuator is driving. This means that a smaller value of g_e can be used in practice, decreasing the length of the comb finger and increasing the energy density of the array.

A.4 Optimization: Volumetric Energy Density

The volumetric energy density of an actuator will be defined as:

$$D_E = \frac{\Delta E}{V} \tag{A.10}$$

Where ΔE is defined above, and V is the actuator volume not including suspension and anchor volumes. The geometry of a bidirectional lateral comb drive actuator is shown in Figure 39. As shown in the figure, the volume is:

$$V = 4 N_f L_{cell} W_{cell} h = 4 N_f \left(W_{as} + \frac{1}{2} W_{am} + L_f + y_m + g_e + L_0 \right) \left(W_f + g \right) h \quad (A.11)$$

Where the variables and their constraints are summarized in Table 2. The derivation of these constraints are shown next. Optimizing the actuator array relies upon finding an optimal value for the energy density of the actuator subject to the constraints placed on each of these parameters. These constraints come from several sources: 1) external requirements, 2) process limitations, 3) prevention of instability (see analysis above), and 4) optimization of resonant frequency.

Parameter	Description	Constraint Equation	Constrained by
N_f	# of unit cells in the array	$N_f = \dfrac{F}{\varepsilon_0 V_m^2} \left(\dfrac{g}{h} \right)$	External constraint
W_{as}	Width of stationary arm	$W_{as} = \dfrac{2 f_n \pi L^2 \left(W_f + g \right) - \dfrac{W_f L_f}{2} \sqrt{\dfrac{E}{\rho}}}{\left(W_f + g \right) \sqrt{\dfrac{E}{\rho}}}$	Resonant Frequency

W_{am}	Width of moving arm	$$W_{as} = \frac{2f_n\pi L^2\left(W_f + g\right) - W_f L_f \sqrt{\dfrac{E}{\rho}}}{\left(W_f + g\right)\sqrt{\dfrac{E}{\rho}}}$$	Resonant Frequency
L_f	Length of comb finger	$$L_f = L_0 + y_m + g_e$$	Dependent variable
y_m	Maximum actuator displacement (sum of both directions)	Free variable	Energy Density (global optimization)
g_e	End gap between tip of comb finger and opposing arm when $y = \dfrac{y_m}{2}$	$$g_e \geq \left(\frac{2N_f\varepsilon_0 hW_f V_m^2}{k_y}\right)^{\frac{1}{3}} + y$$	Stability
L_0	Initial finger overlap distance	$$L_0 = 0.1L_f$$	Linearity (Grade et al. 2003)
W_f	Width of comb finger	$$W_f \geq \sqrt[3]{\frac{3}{2}\left(\frac{\varepsilon_0}{E}\right)V_m^2\left(L_f - g_e\right)L_f^3\left[\frac{1}{\left(g+x\right)^3} + \frac{}{\left(g\right.}\right.}$$	Stability

g	Gap between comb fingers	$g = \max\left(\dfrac{h}{AR}, \lambda\right)$	Process
h	Thickness of actuator	$h = \lambda \cdot AR$	Process

Table 2: Comb drive unit cell dimensions and their constraints. See text for variable definitions.

A.4.1 External Requirements

External requirements are defined as the design goals for the actuator. They can include the maximum force of the actuator, maximum displacement, linearity, or any number of other requirements. This paper deals with the optimization of energy density. Because energy density is a combination of force, displacement, and volume, the designer should keep in mind that the most energy dense actuator may not meet all their design requirements. For example, an actuator may have an extremely high force over an extremely short distance, making it very energetically dense. However, that actuator might not meet the needs of a designer who requires a more moderate force that is linear over a large displacement. For additional discussion on this topic, see (Yeh 2001). In this discussion, the only design parameter is the total force of the array, as controlled by the number of comb fingers.

A.4.2 Manufacturing Process Limitations

The fabrication process influences the design of the actuator through two parameters, the thickness of the actuator h and the minimum gap between comb fingers g. These two parameters are linked through the minimum lithographic feature size λ and the maximum

trench aspect ratio $AR = \dfrac{depth}{width}$ that the etcher is capable of. When the thickness h is

chosen to match the lithographic capabilities to the etching capabilities of the fabrication

process, the actuator thickness is defined as $h = \lambda \cdot AR$, and the gap $g = \lambda$. When the

thickness is not defined by the processing capabilities, the gap is limited either by

lithography (if the thickness is less than the matched value) or by etcher aspect ratio (if

the thickness is greater than the matched value). In this case, $g = \max\left(\dfrac{h}{AR}, \lambda\right)$.

A.4.3 Stability

There is often a tradeoff between travel distance and stability. This can result in a clear

maximum value for energy density for each stability constraint. An example of this is

shown in (Figure 42). In this case overly-long travel distances result in wide comb

fingers to combat the side pull-in problem -- this increases the size of the unit cell and

reduces the energy density. At the same time, continued reduction in the comb finger

width and consequent reduction in the size of the unit cell does not compensate for the

reduction in travel distance of the actuator.

Energy density vs actuator travel

Legend (left side):
- 1.532×10^5
- $E_v(y, 0.1\ \mu m)$ ——
- $E_v(y, 0.5\ \mu m)$ - - - - -
- $E_v(y, 1\ \mu m)$ — · —
- $E_v(y, 2\ \mu m)$ — —
- $E_v(y, 5\ \mu m)$ — · · —
- 307.798

Y-axis: $2 \cdot 10^5$, $1.5 \cdot 10^5$, $1 \cdot 10^5$, $5 \cdot 10^4$, 0

X-axis: $2.5 \cdot 10^{-5}$, $5 \cdot 10^{-5}$, $7.5 \cdot 10^{-5}$, $1 \cdot 10^{-4}$; y; $0.1\ \mu m$ (left), $100\ \mu m$ (right)
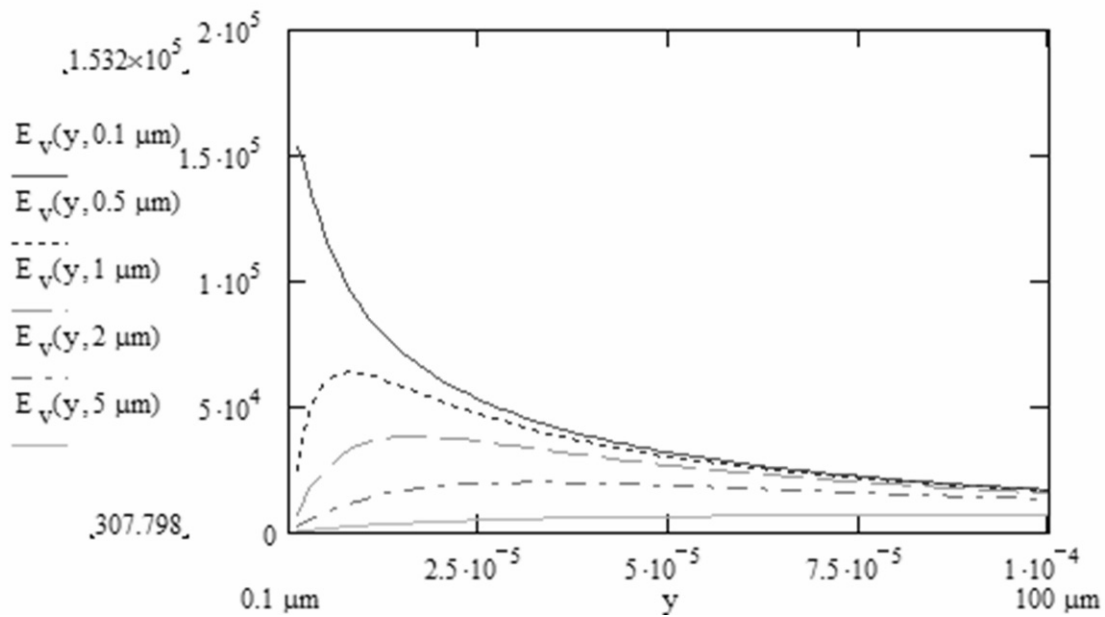
Figure 42. Energy density (in Joules per square meter) of a unit cell as a function of actuator travel distance y. This plot shows the effects of the finger stability requirement on the actuator energy density. Different curves represent different gap sizes g.

The stability constraints in the table are all derived using the energy method described in the previous section. One issue that arises when using these equations is that often the position variable x or y appears in the design equation. These values refer to the position of the actuator relative to its initial position, which in all cases derived here, is defined as the end of the range of travel of the actuator. So the local coordinate system keeps $x = 0$ at the same location, but shifts the origin of y to $\frac{y_m}{2}$. The easiest thing for the designer to do is assume that the comb drive does not move from the local origin (set x and y to zero) and to calculate the values for these parameters accordingly. However, it would be prudent to factor in a margin to accommodate overshoot caused by underdamping,

149

resonance, and imperfections in the direction of travel of the actuator (eg. if the load

driven by the actuator has a lateral component of force or if the suspension travels in an

arc instead of a straight line). These safety factors are very design-specific and are

probably best chosen with the aid of FEA tools.

A.4.4 Resonant Frequency

There are two main situations where the resonant frequency constraints become active:

during the optimization of power density, and control system design. The power output

of an actuator is defined as the energy output of a single cycle times the number of cycles

per unit time. Above the resonant frequency, any increase in the number of cycles per

unit time will result in reduced energy output per cycle. Since mechanical spring-mass

systems have at least two poles, the energy output per cycle will decrease faster than the

frequency increases. This means that the resonant frequency is also the frequency at

which the power output of the actuator peaks. In order to maximize this power output, it

is important to ensure that the first resonant mode of the actuator is as high in frequency

as possible. This means that no part of the actuator should unintentionally start

resonating before the entire actuator starts resonating. This sets a constraint on the

dimensions of some of the larger actuator substructures, such as the actuator spine and

arms. However, this constraint may not be tight enough. When a controller is built to

close the loop on the position of the actuator, it is important from the standpoint of

stability to have no poles nearby the two dominant poles. So, depending on the design of

the controller, it may be necessary to push the resonant frequencies of the actuator

substructures higher by several orders of magnitude than the resonant frequency of the actuator itself.

A.4.5 Energy Density vs Power Density

This paper addresses energy density of actuators instead of power density because in many cases of interest, the structure being driven by the actuator sets the dominant poles of the system rather than the actuator itself. In this case, the peak power output of the actuator by itself is less important than the amount of energy per cycle it can deliver since the combined system will never move fast enough to utilize the full power output of the actuator. This is analogous to a game of tug-of-war between a motor scooter and a horse. The scooter puts out about 20 peak horsepower vs the horse's 9 peak horsepower (Dutto, Hoyt et al. 2004), but the scooter's lack of low-end torque prevents it from getting to a speed where it can put it to good use.

That said, it is still important to have a good separation between the lowest-frequency pole and higher-order modes in order to ensure that the structure does not go unstable once it is placed in a closed-loop controller. To ensure this, we calculate the resonant modes for some of the substructures of the actuator. The arms supporting the comb fingers and the actuator spine are the substructures most likely to have low-frequency vibrations. The 1st resonant mode of a cantilevered beam with a uniformly distributed mass (such as the stationary or moving arm of a comb drive) is given by [Roark's, Table 36, eq 3(b)]:

$$f_n = \frac{3.52}{2\pi}\sqrt{\frac{EI}{\left(\frac{m}{L}\right)L^4}} = \frac{3.52}{2\pi L^2}\sqrt{\frac{Ehb_{eff}^3}{12\left(\frac{\rho b_{eff}hL}{L}\right)}} = \frac{b_{eff}}{2\pi L^2}\sqrt{\frac{E}{\rho}} \qquad (A.12)$$

Where ρ is the density of the actuator material, E is the Young's modulus of the actuator material, L is the length of the beam, and b_{eff} is the effective width of the beam. The effective width b is approximated by the average over a unit cell of the width of the beam. In the case of the actuator shown in Figure 39, this average is given by:

$$\text{Stationary arm: } b_{eff} = \frac{W_f\left(L_f + W_{as}\right) + \left(W_f + 2g\right)W_{as}}{2\left(W_f + g\right)} \qquad (A.13)$$

$$\text{Moving arm: } b_{eff} = \frac{W_f\left(2L_f + W_{am}\right) + \left(W_f + 2g\right)W_{am}}{2\left(W_f + g\right)} \qquad (A.14)$$

Plugging this into the equation for the resonant frequency and solving for $W_{a[s,m]}$ gives the constraints in the table:

$$\text{Stationary arm: } W_{as} = \frac{2f_n\pi L^2\left(W_f + g\right) - \dfrac{W_f L_f}{2}\sqrt{\dfrac{E}{\rho}}}{\left(W_f + g\right)\sqrt{\dfrac{E}{\rho}}} \qquad (A.15)$$

$$\text{Moving arm: } W_{as} = \frac{2f_n\pi L^2\left(W_f + g\right) - W_f L_f\sqrt{\dfrac{E}{\rho}}}{\left(W_f + g\right)\sqrt{\dfrac{E}{\rho}}} \qquad (A.16)$$

152

A.5 Actuator Suspension

So far this analysis has focused on the optimization of the unit cell of the actuator itself and has ignored the design of the actuator suspension except where it influences the design of the unit cell. This is because the suspension is generally not desired to be the dominant stiffness in the actuator-load system. Its sole purpose is to constrain the motion to the desired linear path while providing the least impediment to longitudinal motion. Three main suspension topologies have dominated the literature: the parallel-bar flexure, the crab-leg flexure and the double-folded flexure.

The first of these, the parallel-bar flexure, is not as desirable as the other two topologies because the actuator does not follow a straight line; instead it follows an arc roughly radius L_s, the length of the suspension beam. Attaching a second parallel-bar flexure creates a pair of clamped-clamped beams, returning the path of the actuator to a straight line. However, the beams must now be stretched axially in order to move in the longitudinal direction. This causes a nonlinear increase in the longitudinal spring constant of the suspension as a function of displacement. Additionally, thermal expansion or residual stress in the actuator material can easily cause the suspension beam to buckle, causing the actuator to move without applied voltage.

The crab-leg topology addresses the problem of axial extension of the spring by adding short beams ("hips", beam 2 in figure in Table 1) orthogonal to the parallel bars ("legs", beam 1 in figure in Table 1) to relieve the strain in the axial members. Instead of the legs

stretching axially, the hips bend laterally. This allows the longitudinal spring constant to remain constant instead of changing substantially with displacement. In addition, thermal expansion or contraction and residual stress are also accommodated by the hips. The main problem with this topology is that the hips are indiscriminate about what lateral motion they allow. This can have an adverse effect on the actuator stability against side pull-in (Table 1).

The double-folded flexure is an improvement over both these approaches since it is a) more compact than either actuator for the same longitudinal spring constant, b) it automatically relieves strain from residual stress or thermal expansion through symmetry, and c) it follows a straight line without losing axial stiffness.

In each of these topologies, the key figure of merit is the stiffness ratio $\dfrac{k_a}{k_l}$, the ratio of axial (x-axis) stiffness to lateral (y-axis) stiffness. This stiffness ratio is derived for the clamped-clamped beam, the crab-leg and double-folded flexures in (Legtenberg, Groeneveld et al. 1996), and summarized in the table below:

Suspension type	Stiffness ratio
Clamped-clamped beam suspension	$\dfrac{k_a}{k_l} = \left(\dfrac{L}{b}\right)^2$
Crab-leg suspension	$\dfrac{k_a}{k_l} = \dfrac{L^3}{L_{hip}^3}\left[\dfrac{L^2 + 6LL_{hip} + 8L_{hip}^2}{2L^2 + 3LL_{hip} + L_{hip}^2}\right]$

Double-folded flexure	$$\frac{k_a}{k_l} = \left(\frac{L}{b}\right)^2$$

Table 3: Stiffness ratio for three different suspension topologies. The long beam has length *L* and width *b*. In the crab-leg suspension, L_{hip} refers to the length of the short strain relieving beam.

The stiffness ratios above apply to small longitudinal displacements of ~10% of the suspension length *L*. For large displacements, the suspension starts to become less stiff. This softening occurs in both the in-plane and out-of-plane directions. This is because for large deflections, the tip of the beam is quite a distance from where it was lined up with the anchor end of the beam. Any deflection from this point creates a moment arm that transforms lateral forces or out-of-plane forces into moments. Out-of-plane forces result in torsion of the beam, while lateral forces become bending moments.

In the case of unidirectional actuator motion, one can fabricate the beam in a pre-bent shape such that it straightens out as it moves (Grade, Jerman et al. 2000). This prevents this behavior since the smallest moment arm coincides with the largest applied voltage. This reduces the undesired moment since the reduction in moment arm offsets the increase in lateral force.

With bidirectional actuator motion, the spring softening problem can be solved in a different way. One possibility is to increase the length of the suspension until the longitudinal displacement falls within the range for which the stiffness ratio is valid. The downside to this approach is that the long beams are very susceptible to buckling since the force required to buckle a beam goes down with the square of its length. By using

multiple copies of the beam right next to one another, the force required to buckle the beam goes up. Of course, the multiple copies of the beam cause the longitudinal spring constant to increase as well. But the spring constant of the beam decreases with the cube of the length while it only increases linearly with the number of copies, so this increase in spring constant can be easily compensated for by the original lengthening of the beam.

A.6 High Voltage Drive and Sense

A.6.1 Electromechanical Amplitude Modulation (EAM)

Electromechanical Amplitude Modulation (Nguyen 1994; Cagdaser, Jog et al. 2004) uses a high-voltage linear amplifier to generate the drive signal. This requires a DC voltage supply above the maximum desired drive voltage, and a high-voltage amplifier. By utilizing separate drive and sense comb fingers, the high voltage drive signal is decoupled from the low-voltage sense signal. Additionally, the sense signal is a high-frequency signal, allowing any remaining feedthrough from the drive to the sense signals to be filtered out in the electronic domain (Figure 43).
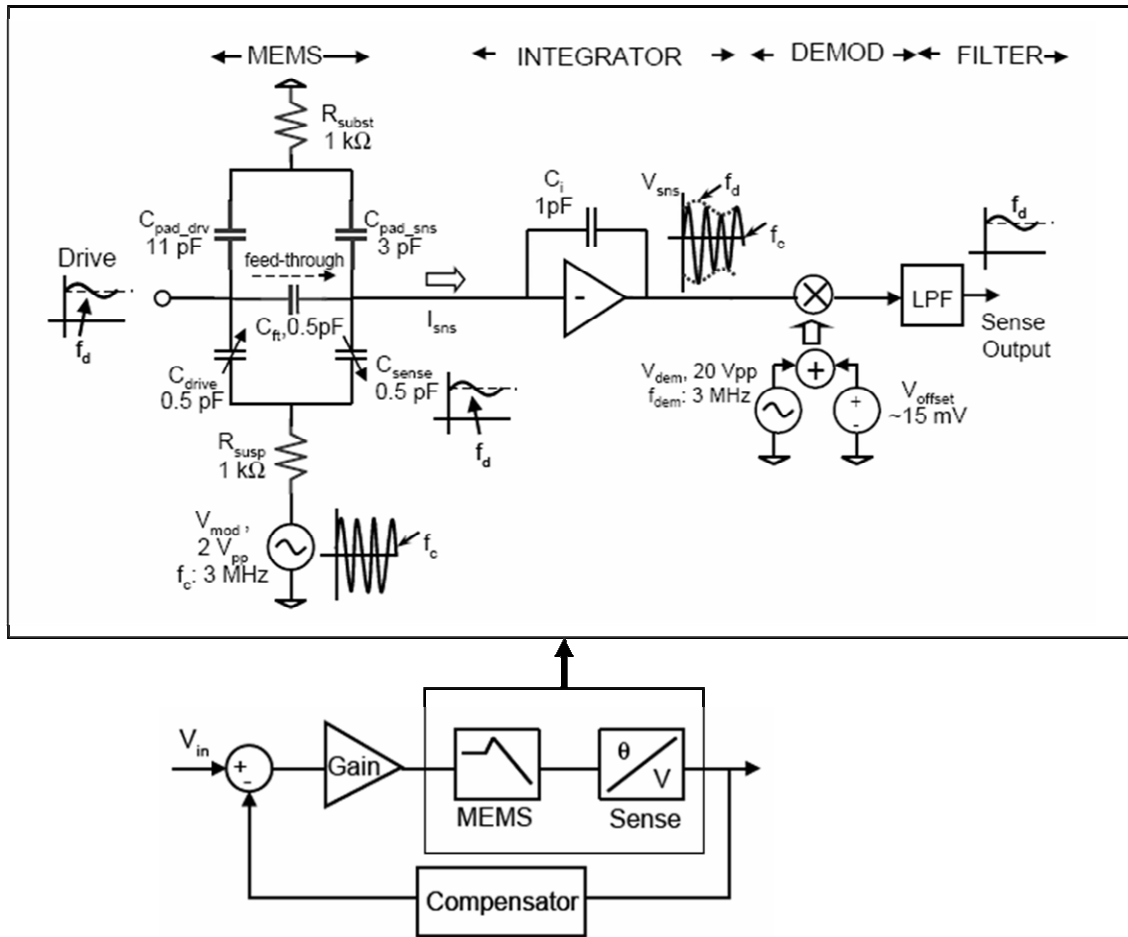
Figure 43. The electromechanical amplitude modulation technique. A closed-loop control system consists of a gain stage driving a MEMS actuator. A high-frequency signal is applied to a separate set of differential position sensing capacitors. The high-frequency signal is demodulated and filtered, resulting in a position sense signal. This is fed back into a compensator and subtracted from the control signal. From (Cagdaser, Last et al, 2004)

A.6.2 Resonant Drive

Resonant drive (Cagdaser and Boser 2005, 2005) utilizes a high-Q series LC tank circuit whose resonant frequency is much higher than the mechanical resonance of the MEMS system being controlled (Figure 44). The inductor is typically an off-chip inductor, while the capacitor is the comb drive itself. A variable-gain amplifier causes the tank to start resonating. At resonance, the LC tank provides a voltage boost across the comb drive

equal to its quality factor Q. By changing the amplifier gain, the amplitude of the

resonant signal, and therefore the voltage the comb drive sees, is controlled. Since

$F \propto V^2 = \left(QA\sin\left(\omega_{tank}t\right)\right)^2 = \dfrac{\left(QA\right)^2}{2}\left(1+\cos\left(2\omega_{tank}t\right)\right)$, the comb drive converts the AC

drive signal into a force with two components – a DC force, and an AC force at twice the

original frequency. The high-frequency oscillations are low-pass filtered by the

mechanical structure, reducing their effect to negligible amounts if ω_{tank}>>$\omega_{mechanical}$.

With this system, CMOS electronics operating at low voltages can drive MEMS

structures requiring voltages a factor of 10 or more above what the electronics could

survive. Additionally, because the comb drive is a variable capacitor, the resonant

frequency of the LC tank is a function of the displacement of the comb drive. This

resonant frequency can be measured through a capacitive voltage divider by the low-

voltage CMOS. The two properties of a sine wave, its amplitude and frequency, can

therefore be used for a feedback control system where the amplitude is used for the feed-

forward path, and the frequency is used to provide position feedback. And because the

circuit is always oscillating at its resonant frequency, Q is always maximized for the

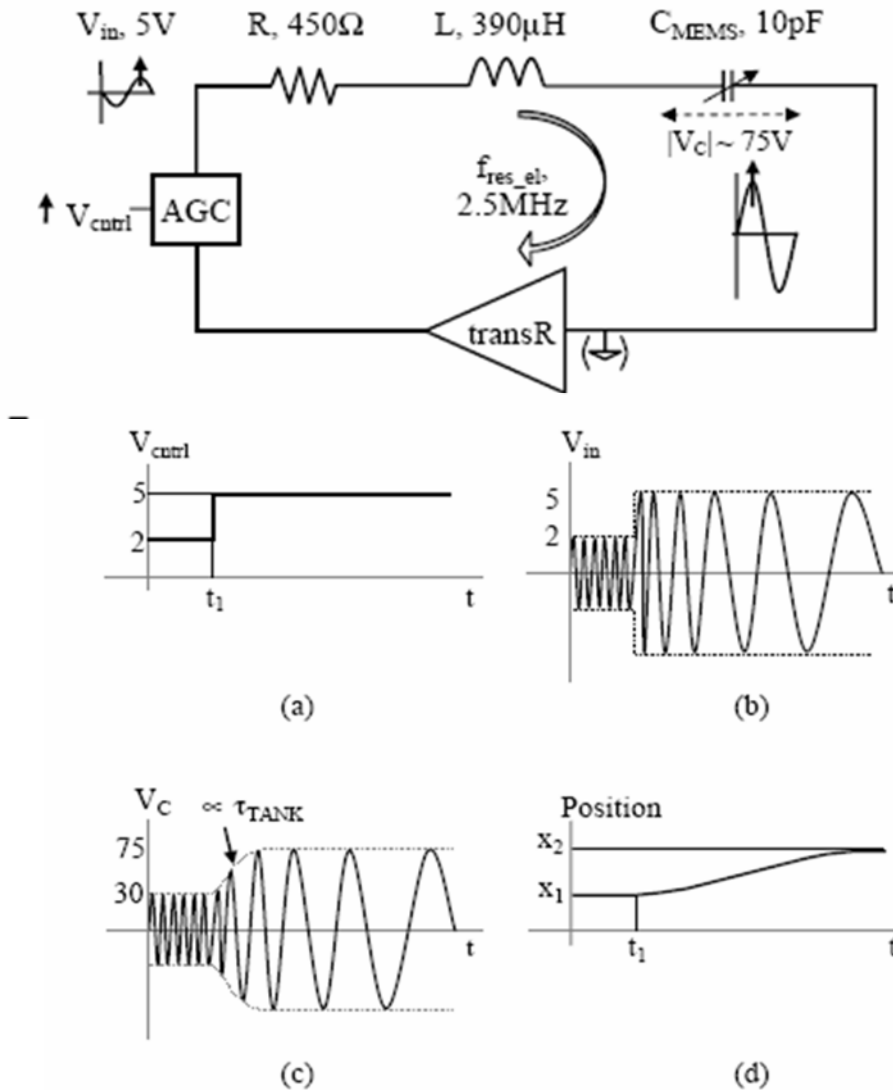circuit, causing the maximum available voltage boost to be achieved.

Figure 44. The resonant drive technique. An external inductor forms a series LC tank circuit with the electrostatic MEMS actuator as "C" part of the LC tank. An automatic gain control (AGC) circuit starts the LC tank resonating at its natural frequency. The series LC passively boosts the voltage across the MEMS capacitor by the quality factor of the circuit (usually up to a factor of ~10). A phase-locked loop (PLL) (not shown) is used to sense the frequency, which is directly related to the value of the MEMS capacitor and therefore the position of the actuator. This is fed back into a compensator and subtracted from the control input. Typical operational waveforms are shown below: To change the position of the actuator, the gain of the AGC is adjusted (a). This increases the amplitude of the oscillations across the LC tank (b). As the MEMS actuator shifts position, the frequency of oscillation changes (c). This is read out as a change in position using the PLL (d). (From Cagdaser and Boser, 2005)

A.6.3 EAM vs Resonant Drive

The biggest advantages of resonant drive are the ability to use the same comb fingers for drive and sense with minimal feedthrough, the minimization of external components (the only components are the controller chip, a variable-gain amplifier, an inductor, and the MEMS chip), and the simplicity of control. Its biggest limitations stem from the limited Q of the inductor: power is required even when the comb drive is not moving due to the need to keep the oscillator ringing, and the voltage boost is usually limited by the Q of the inductor. However, even if the Quality factor of the inductor was infinite, there would be other problems with this approach. Since the time to ring up or ring down an inductor is proportional to the Q of the resonator, any increase in Q requires an increase in the LC tank's resonant frequency in order to avoid slowing the response time of the actuator. Unfortunately, as inductors increase in self-resonant frequency, usually their Quality factors decrease as well. This property limits the resonant drive approach to low-frequency applications or applications only needing voltage boosts in the range of 10-15.

EAM does not suffer from these limitations. A high-efficiency DC-DC converter is used to drive a high-speed high-voltage linear amplifier. The use of two separate modules decouples the design of the two components and can result in a higher performance system. Although there is still power dissipated when the comb drive is not moving (the bias current in the high-voltage amplifier), it can be made to be less than the power dissipated in the LC tank of the resonant drive circuit with the right choice of high-voltage amplifier. However, EAM does require a fairly complex electronics module consisting of multiple chips and passive components (see figure [from Cagdaser]). Additionally, the overhead of separate comb fingers required to sense position can be

considered a decrease in actuator volumetric energy density since the sense combs do contribute to actuator volume but do not contribute to force generation.

A.6.4 Differential Drive and Sense

In many cases involving bidirectional actuation or where unidirectional actuation with increased linearity is desired, differential driving and sensing are employed. Differential sensing is useful for reducing the influence of common mode noise and parasitics on the sense signal. In addition, differential drive is useful for linearizing the drive signal. If the drive comb fingers are always engaged (even at either extreme of the actuator's range of motion) then differential drive can be employed. Applying this system to an actuator moving up and down in the y-axis, the moving combs are grounded and separate voltages are applied on each of the opposing combs. In the undeflected position, both voltages are the same: $V_{up} = V_{down} = \dfrac{V_m}{2}$. This results in zero net force on the actuator. In order to move the actuator, the two voltages are adjusted by an equal and opposite amount. This creates a net force on the actuator of

$$F_{net} = N_f \varepsilon_0 \left(\frac{h}{g} \right) \left(V_{up}^2 - V_{down}^2 \right) = N_f \varepsilon_0 \left(\frac{h}{g} \right) \left[\left(\frac{V_m}{2} + V \right)^2 - \left(\frac{V_m}{2} - V \right)^2 \right] = 2 N_f \varepsilon_0 V_m \left(\frac{h}{g} \right) V$$

$$(A.17)$$

Where all variables are the same as defined above, and V is the differential adjustment to the voltage applied to each side of the comb drive. Note how the net force has gone from a quadratic to a linear function of the applied voltage, yet the maximum and minimum forces applied by the actuator are still the same (when $V = \pm \dfrac{V_m}{2}$) as with the single-ended drive.

A.7 Conclusion

Electrostatic comb drives are the actuators of choice in applications where high accuracy analog positioning at minimal power levels are desired. However, their largest drawback has traditionally been seen as their lack of force. Many applications have not taken full advantage of the energy density achievable using these actuators due to actuator instabilities. These instabilities limit the maximum voltage at which the actuators can be operated. By designing the actuator such that it does not suffer these instabilities, the voltage can be raised. This appendix provides all the information necessary to design actuators that are stable up to whatever voltage the designer chooses. The only limit to actuator voltage scaling is electrical breakdown of the gaps; this occurs in air around 300V for micron-scale gaps. Circuits able to drive 300V are not common; usually special processes are required to fabricate such circuits, leading circuit designers to place voltage restrictions on the MEMS portion of the system. If high-voltage components are acceptable in the system, Electromechanical Amplitude Modulation provides a high-performance method of controlling the comb drive actuator. If high-voltage components are not acceptable, resonant drive provides a method to reduce the voltage requirement by

a factor of 10 or more while simultaneously providing a method to sense the position of

the actuator without requiring a separate set of sense comb fingers.

Appendix B: Ortho-Gripper Design Worksheet for MathCAD

Orthogripper calculations

Constants:

$$\rho_{Si} := 2330 \frac{kg}{m^3} \qquad g = 9.807 \frac{m}{s^2} \qquad \mu m := 10^{-6} m \qquad \mu N := 10^{-6} N$$

Parameters:

$$k_p := 3 \frac{\mu N}{\mu m} \qquad \text{Preload spring constant}$$

$$k_c := 6 \frac{\mu N}{\mu m} \qquad \text{Clamping spring constant}$$

$$t_{part} := 25 \mu m \qquad \text{Thickness of part}$$

$$d_i := 5 \mu m \qquad \text{Initial separation between fingers}$$

$$y_i := t_{part} + 10 \mu m \qquad \text{Preload distance on preload spring}$$

$$\mu := 0.3 \qquad \text{Worst-case (lowest) friction coefficient of Si-Si interface}$$

$$V_{part} := 20 \mu m \cdot 100 \mu m \cdot 500 \mu m \qquad V_{part} = 1 \times 10^{-12} m^3 \qquad \text{Volume of part to be manipulated}$$

Derived Parameters:

$$\theta_i := atan\left(\frac{t_{part}}{d_i}\right) \qquad \theta_i = 78.69 deg \qquad \text{Initial angle between contact points}$$

$$L := \sqrt{t_{part}^2 + d_i^2} \qquad L = 25.495 \mu m \qquad \text{Straight-line distance between contact points}$$

$$\theta_{ns} := atan(\mu) \qquad \theta_{ns} = 16.699 deg \qquad \text{Angle of friction cone}$$

$$m_{part} := \rho_{Si} \cdot V_{part} \qquad m_{part} = 2.33 \times 10^{-9} kg \qquad \text{Mass of manipulated part}$$

$$w_{part} := m_{part} \cdot g \qquad w_{part} = 2.285 \times 10^{-8} N \qquad \text{weight of manipulated part in 1 gravity}$$

Calculations:

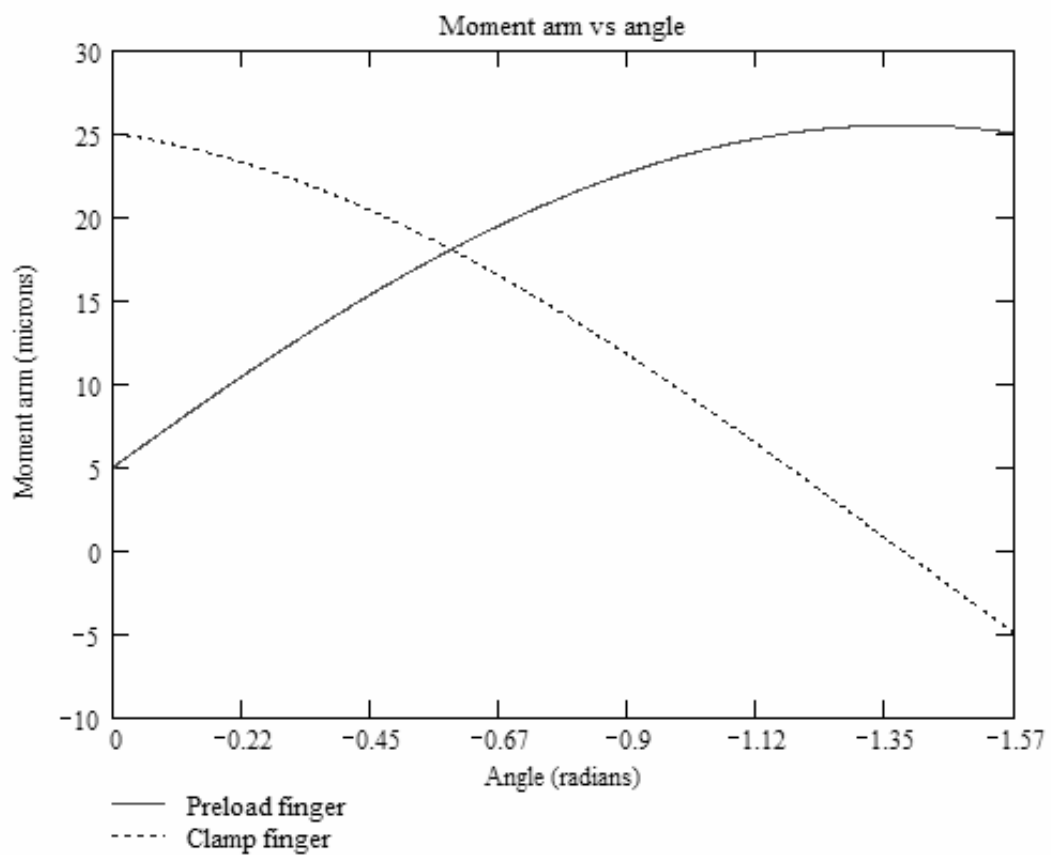$$d_{imax} := \mu \cdot t_{part} \qquad d_{imax} = 7.5\,\mu m \qquad \text{Non-slip condition only met if d.i <= d.imax}$$

$$d_i \le d_{imax} = 1 \qquad \text{0=false, 1= true}$$

Moment arm calculations:

$$L_{mp}(\theta) := L \cdot \cos(\theta_i + \theta)$$

$$L_{mc}(\theta) := L \cdot \sin(\theta_i + \theta)$$

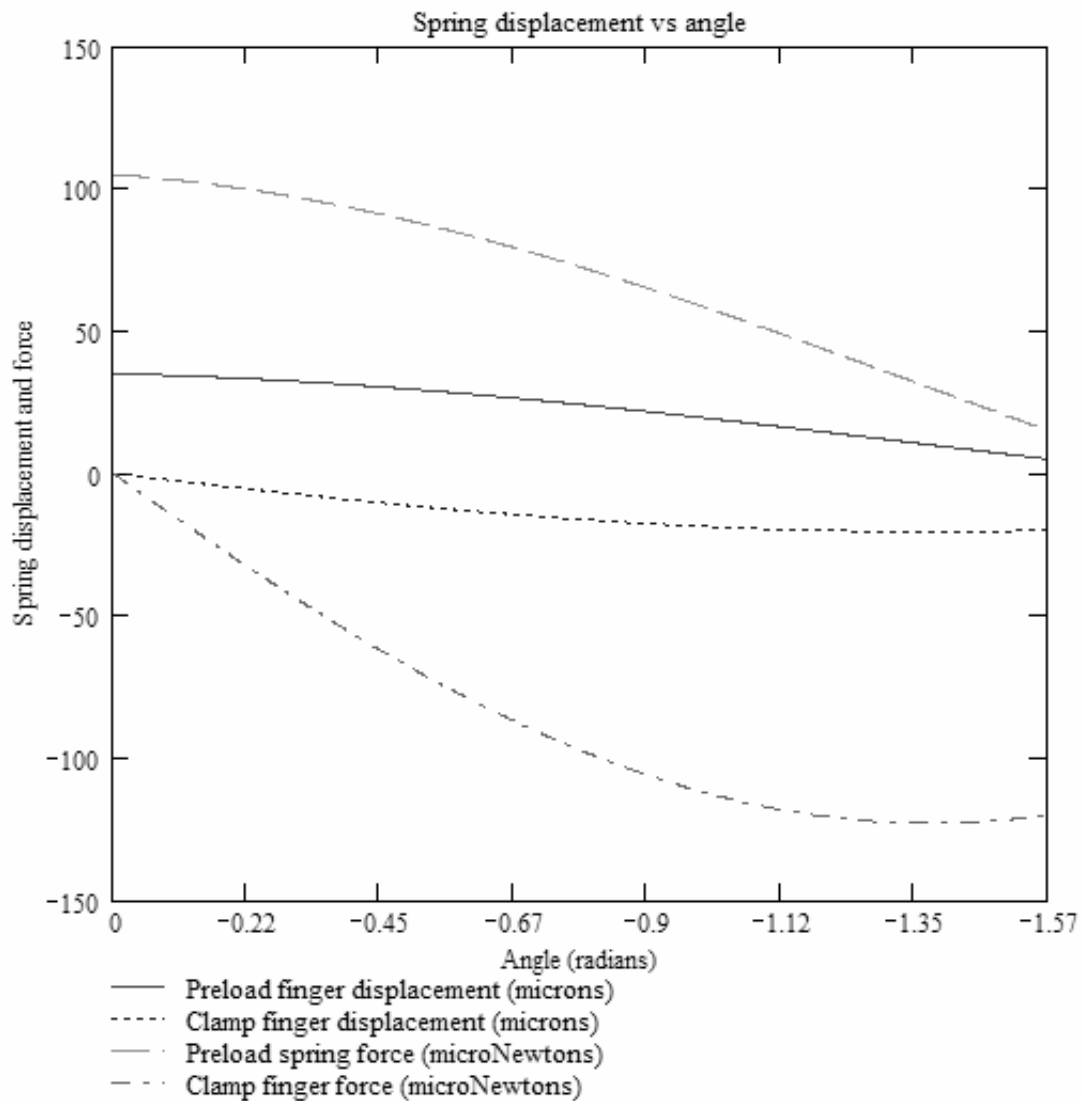Moment arm vs angle

Preload finger
---- Clamp finger

Spring displacement calculations:

$$y_p(\theta) := y_i - L \cdot \left(\sin(\theta_i) - \sin(\theta_i + \theta) \right)$$

$$F_p(\theta) := k_p \cdot y_p(\theta)$$

$$x_c(\theta) := L \cdot \left(\cos(\theta_i) - \cos(\theta_i + \theta) \right)$$

$$F_c(\theta) := k_c \cdot x_c(\theta)$$

Spring displacement vs angle

——— Preload finger displacement (microns)
- - - - Clamp finger displacement (microns)
— · Preload spring force (microNewtons)
— · · Clamp finger force (microNewtons)
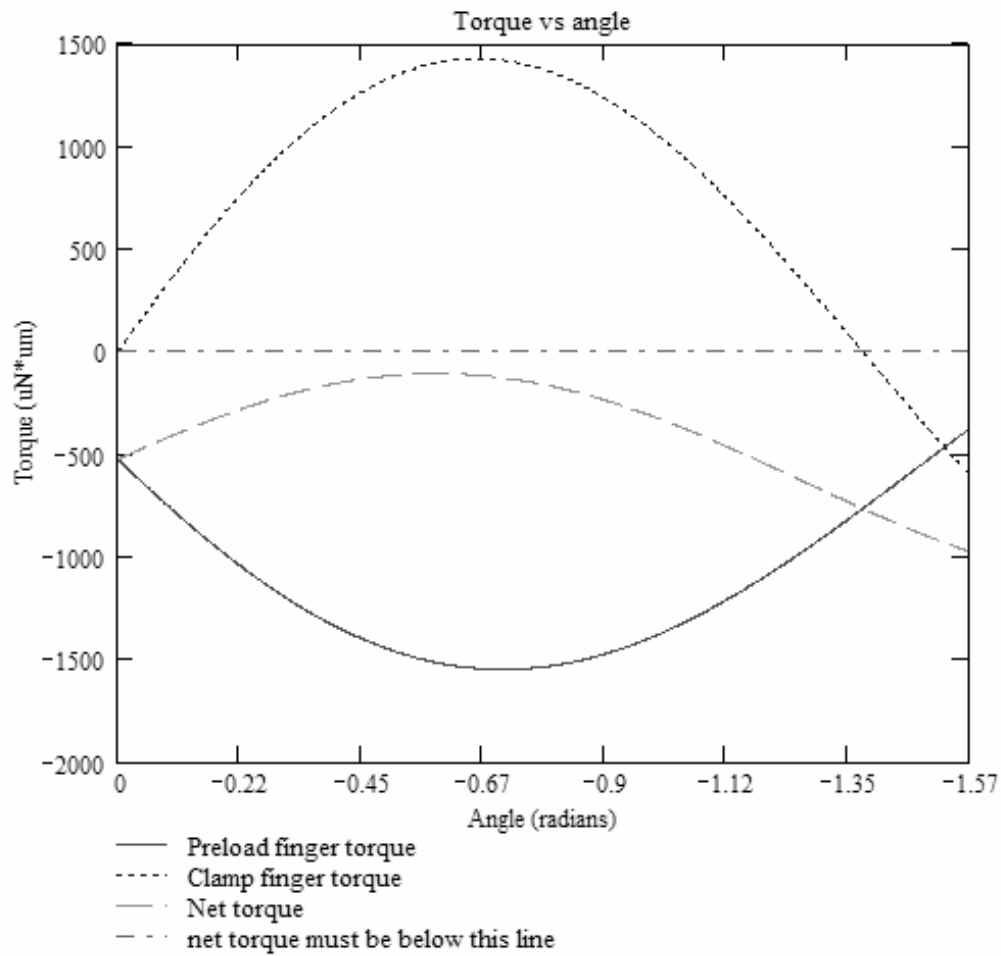
Torque calculations:

$$\tau_p(\theta) := -L_{mp}(\theta) \cdot F_p(\theta)$$

Need negative sign to orient torque in correct direction (comes from geometry)

$$\tau_c(\theta) := L_{mc}(\theta) \cdot -F_c(\theta)$$

Need negative sign to orient torque in correct direction (comes from needing to negate the negative sign in the force from expanding the spring)

$$\tau(\theta) := \tau_c(\theta) + \tau_p(\theta)$$

Torque vs angle

— Preload finger torque
---- Clamp finger torque
— · Net torque
— ·· net torque must be below this line

At θ=-90 degrees:

Clamping force:

$$\left| F_c(-90 \cdot deg) \right| = 120 \, \mu N$$

$$N_g := \frac{\mu \cdot \left| F_c(-90 \cdot deg) \right|}{w_{part}} \qquad N_g = 1.576 \times 10^3 \qquad \text{Number of gravities of acceleration before part begins to slip out of grippers}$$

Tangential force from preload finger:

$$\frac{\left| F_p(-90 deg) \right|}{\left| F_c(-90 deg) \right|} \leq \mu = 1 \qquad \frac{\left| F_p(-90 deg) \right|}{\left| F_c(-90 deg) \right|} = 0.125 \quad \text{No-slip requirement: 0 = slip, 1 = no slip}$$