Parallel-Plate Micro Servo for Probe-Based Data Storage

by

Michael Shiang-Cheng Lu

B.S. (National Tsing Hua University) 1991
M.S. (National Tsing Hua University) 1993

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

Doctor of Philosophy

in

Electrical and Computer Engineering

in the

Carnegie Institute of Technology

of the

Carnegie Mellon University

Committee:

Professor Gary K. Fedder, Advisor
Professor L. Richard Carley
Professor James A. Bain
Dr. Storrs Hoen (Agilent, Inc.)

2002
Parallel-Plate Micro Servo for Probe-Based Data Storage

© 2002

by

Michael Shiang-Cheng Lu
To my parents,
to Chun-Yen, Chun-Zho, and Chun-Yi,
and to my wife Ching-Yu.
Abstract

Parallel-Plate Micro Servo for Probe-Based Data Storage

by

Michael Shiang-Cheng Lu

Doctor of Philosophy in Electrical and Computer Engineering

Carnegie Mellon University

Professor Gary K. Fedder, Advisor

This thesis describes the use of closed-loop voltage control to extend the travel range of a parallel-plate electrostatic microactuator beyond the open-loop pull-in limit of one third of the gap. A general controller design procedure is presented to deal with the nonlinearities and unstable characteristic of the parallel-plate actuator. The resulting linear controller design is implemented easily for the desired application of a probe-based mass data storage device. In the envisioned data storage system, an array of parallel-plate tip actuators are employed to position the read/write probe tips about 10 nm away from the magnetic media for data access. Fabrication of tip actuators and servo and data channel circuits are conveniently integrated by use of the CMOS-MEMS micromachining process developed at Carnegie Mellon.

Controller design ensures system stability in the presence of a unstable pole beyond the pull-in limit. Desired transient response is achieved by a pre-filter added in front of the feedback loop to shape the step input command. The fabricated microactuator is characterized by static and dynamic measurements, with a spring constant of 0.17 N/m, mechanical resonant frequency of 12.4 kHz, and effective damping ratio from 0.55 to 0.35 for gaps between 2.3 to 2.65 µm. The minimum input-referred noise capacitance change is 0.5 aF/
√Hz measured at a gap of 5.7 µm, corresponding to a minimum input-referred noise displacement of 0.33 nm/√Hz. Measured closed-loop step response illustrates a maximum travel distance up to 60% of the initial gap with a rise time less than 5 ms.
# Contents

List of Figures \hspace{1cm} v

List of Tables \hspace{1cm} x

1 Introduction \hspace{1cm} 1

2 Electrostatic Gap Models \hspace{1cm} 9
   2.1 Introduction \hspace{1cm} 9
   2.2 Electrostatic Force \hspace{1cm} 10
   2.3 Gap Modeling \hspace{1cm} 14
   2.4 Simulation \hspace{1cm} 19
   2.5 Experiment \hspace{1cm} 21

3 Plant Design and Modeling \hspace{1cm} 23
   3.1 Introduction \hspace{1cm} 23
   3.2 CMOS-MEMS Fabrication \hspace{1cm} 23
   3.3 Material Property Characterization \hspace{1cm} 29
      3.3.1 Effective Young’s Modulus \hspace{1cm} 29
      3.3.2 Residual Stress and Vertical Stress Gradient \hspace{1cm} 31
      3.3.3 Cyclic Fatigue \hspace{1cm} 33
   3.4 Parallel-Plate Microactuator Design \hspace{1cm} 35
      3.4.1 Dynamic Equation \hspace{1cm} 35
      3.4.2 Design Description \hspace{1cm} 37
      3.4.3 Spring Constant \hspace{1cm} 41
      3.4.4 Modal Analysis \hspace{1cm} 47
      3.4.5 Squeeze-Film Damping \hspace{1cm} 54
   3.5 Capacitive Position Sensing \hspace{1cm} 60
      3.5.1 Pre-amp Circuit Design \hspace{1cm} 63
      3.5.2 Double-Balanced Demodulator \hspace{1cm} 70
   3.6 Noise Analysis and Minimum Detectable Signal \hspace{1cm} 72
      3.6.1 Pre-amp Noise \hspace{1cm} 72
      3.6.2 Brownian Noise \hspace{1cm} 77

4 Controller Design \hspace{1cm} 79
   4.1 Introduction \hspace{1cm} 79
   4.2 Classical Frequency-Domain Feedback Theory. \hspace{1cm} 79
      4.2.1 Quantification of Feedback Performance \hspace{1cm} 80
      4.2.2 Analysis of Unstable Plant \hspace{1cm} 82
         4.2.2.1 Stability Criterion \hspace{1cm} 82
         4.2.2.2 Bandwidth Limitations with a Real Unstable Pole \hspace{1cm} 83
List of Figures

1.1 Schematic representation of the envisioned micro disk drive. .................. 4
1.2 Top view and side view of the read/write probe head design. ............... 4
1.3 Side view of the bonded media-actuator die and the tip-actuator die, illustrating non-uniform gaps resulted from a tilted media and curl variations of tip actuators. .......... 6
1.4 Schematic representation of the feedback system using a frequency-multiplexed scheme to separate the actuation and sensing signals. Frequency responses in the plot illustrate the frequency components of signals at each node inside the loop. ...... 7
2.1 Schematic of two charged conductors. .................................................. 10
2.2 Path of integration in variable space. (a) For evaluating energy $W_e$. (b) For evaluating coenergy $W_e'$. ................................................................. 11
2.3 Schematic of parallel-plate actuator. .................................................... 13
2.4 Schematic of comb-finger actuator. (a) Top view. (b) Side view. .............. 13
2.5 Schematics of the electrode geometry for (a) two electrodes of the same width, (b) two electrodes in which one is much wider than the other, and (c) comb fingers. 14
2.6 Capacitance per unit length as a function of $g/h$ and $\theta$ for two electrodes of the same width. .............................................................. 15
2.7 Extracted $K_i$ values plotted as a function of the $w/h$ ratio and the sidewall angle $\theta$. 16
2.8 The distributed parallel-plate force is replaced by lumped forces applied at the ends of beam models. ......................................................... 18
2.9 Cantilever beam actuator. (a) Geometry from the top view, where $l = 100$ to 200 $\mu$m, $w = 2$ $\mu$m, $h = 2$ $\mu$m, $\theta = \pi/2$, and $g = 2$ $\mu$m. (b) Schematic representation of interconnected beams and gaps for NODAS simulation. .................................. 19
2.10 NODAS and electro-mechanical finite-element simulations of pull-in voltage for $l = 100$ to 200 $\mu$m, $w = 2$ $\mu$m, $h = 2$ $\mu$m, $\theta = \pi/2$, and $g = 2$ $\mu$m. .................. 20
2.11 Displacement-voltage characteristics from electro-mechanical finite-element and NODAS simulations for an 100$\mu$m-long beam actuator. .......................... 20
3.1 Cross section of the CMOS-MEMS process flow. (a) After CMOS processing. (b) After anisotropic dielectric reactive-ion etch for definition of structural sidewalls. (c) After anisotropic silicon etch. (d) After isotropic silicon etch for structural release. ................................................................. 25
3.2 (a) Fabricated crab-leg comb-drive resonator using the Agilent 0.5 $\mu$m three-metal n-well process. (b) Beam cross-section with three metal layers, inter-metal dielectric layers, and polysilicon. ...................................................... 26
3.3 Polymers stacked on beam sidewalls resulted from excessive polymerization during dielectric RIE. ........................................................................ 26
3.4 Failure mechanisms of electrical connections: (a) opened vias, and (b) lateral etch of refractory Ti/W layers deposited on top and bottom of metal layers. .......... 27
3.5 Resonant beam actuators for measuring effective Young’s modulus of composite beams. ......................................................................... 29
3.6 (a) SEM of a m3-m2-m1 bent-beam strain sensor. (b) Thermal-mechanical finite-element simulation by MEMCAD for residual stress analysis. ................. 32
3.7 Uniform curl of local metal-3 and metal-1-3 beams. ............................... 33
3.8  (a) SEM of a fan structure for durability test. (b) Measured frequency spectrum has a lateral mode at 13.1 kHz and a out-of-plane mode at 11.2 kHz. ............................................ 34
3.9  (a) Resonant frequency change of the fatigue structure with respect to the experienced cycles. (b) Side view of the notch before testing. (c) Broken notch after testing. .......................................................... 35
3.10 (a) Schematic of parallel-plate actuator represented by a lumped-parameter model. (b) Free-body diagram illustrating the forces acting on the plate, including the electrostatic force, $F_e$, the damping force, $F_d$, and the spring restoring force, $F_s$. .... 36
3.11 Normalized displacement-voltage characteristic of parallel-plate actuator. ......... 38
3.12 (a) Schematic of tip actuator design for probe-based data storage. (b) to (f): Side view at cross-section lines, illustrating electrical connections using three metal layers. 39
3.13 Side view schematic of actuator curl design. (a) Ends of springs are close to a straight line between anchors in design. (b) Ends of springs are away from anchors in design. ............................................................. 40
3.14 Schematic of tip actuator design. ................................................................. 42
3.15 Free-body diagram of one half of actuator used for deriving $z$-directional spring constant. .......................................................................................................................... 43
3.16 Comparison of analytic $z$-directional spring constants with finite-element analysis. ................................................................................................................................. 46
3.17 Twelve-degree-of-freedom beam finite element. ............................... 48
3.18 Transformation of a beam element from local to global ($x, y$) coordinates. .... 50
3.19 Schematic of one half of the actuator represented by beam elements for finite-element modal analysis. Nodes are numbered from 1 to $2n + 6$, with $n$ representing the number of connector beams. ....................................................................................................................... 51
3.20 Comparison of $z$-directional eigenmodes with numerical finite-element analysis. 54
3.21 Illustration of the first three modes of the actuator by numerical finite-element analysis. .......................................................................................................................... 55
3.22 Cross-section schematic illustrating a squeezed thin film of gas between two parallel plates separated by a gap, $g$, resulted from the vertical motion of the plate with velocity, $v$. The squeeze-film damping force, $F_D$, is produced opposing the direction of motion. ......................................................................................................................... 56
3.23 Equivalent circuit model using nonlinear inductors and resistors for squeeze-film damping analysis. Damping force and plate velocity are analogous to electrical current and voltage in the schematic. ................................................................. 57
3.24 Schematic of lumped-parameter actuator model driven by a sinusoidal force source for frequency-domain analysis. Squeeze-film damping forces of the three plates in design are modelled separately due to their different displacements and velocities... ................................................................................................................................. 59
3.25 Damping ratio of parallel-plate actuator versus gap separation at atmospheric pressure. ................................................................. 60
3.26 (a) Side-view schematic of parallel-plate actuator illustrating capacitive sensing scheme. (b) Equivalent capacitive sensing circuit model. ............ 61
3.27 (a) Schematic of sensing pre-amp using an operational amplifier in a non-inverting amplifier configuration. The bias transistor operates in the sub-threshold region to act as a large resistance. (b) Schematic of folded-cascode operational amplifier with a common-source amplifier as the output stage. ......... 64
3.28 Schematic of the constant-transconductance bias circuit for the folded-cascode amplifier. ....................................................... 67
3.29 Frequency response of the sensing pre-amp. Measured d.c. gain is 9.85, and the -3 dB frequency is 5.7 MHz. ................................. 69
3.30 Measured resistance of subthreshold biasing transistor versus gate-to-source voltage. .......................................................... 70
3.31 Gilbert multiplier with emitter degeneration applied to improve input voltage range. ............................................................. 71
3.32 Equivalent pre-amp noise model. ............................................. 73
3.33 (a) Minimum input-referred noise capacitance change versus gap. (b) Minimum input-referred noise displacement versus gap. .................. 76
3.34 Equivalent Brownian noise displacement of the micromechanical actuator and equivalent noise displacement derived from preamp noise plotted as a function of gap. 78
4.1 Block diagram of a linear time-invariant feedback system. Adequate controller design rejects input and output disturbances \(d_1, d_2\), and avoids noise \(n\) amplification. ...................................................... 80
4.2 (a) The notion of crossing with the ray between \((-\infty, -1 + j0]\) on the complex plane. (b) The notion of crossing on the Nichols chart, where the ray is defined as \(R_f = \{(\phi, r)|\phi = -180^\circ, r > 0\text{dB}\}. \) ........................................ 83
4.3 Graphical illustration of the sensitivity function magnitude for \(L(s)\) having at least two more poles than zeros. Sensitivity is larger than one when \(L(s)\) enters the unit circle centered at \(-1 + j0\). ........................................... 85
4.4 Minimum complementary sensitivity peak versus ratio of crossover frequency to unstable pole frequency. A rule of thumb for design requires that the crossover frequency at least twice as large as the unstable pole frequency with an one-pole rolloff to reduce resultant peak within 4 dB. ............................................. 88
4.5 Two-degree-of-freedom control system configuration. Sensitivity reduction and disturbance rejection are performed inside the loop by \(C(s)\), and tracking performance is achieved by design of \(F(s)\). .................................................. 90
4.6 Percentage ratio of the voltage difference between the transient and steady-state responses as a function of normalized displacement. (a) tracking bandwidth \(\omega_b = 0.05\omega_n\) to \(0.2\omega_n\), (b) tracking bandwidth \(\omega_b = 0.25\omega_n\) to \(\omega_n\) .................................................. 94
4.7 Ratio of maximum applied voltage over static pull-in voltage versus ratio of tracking bandwidth over actuator bandwidth. .................. 95
4.8 The d.c. actuator gain as a function of normalized displacement. At \(\alpha = 1/3\), the actuator acts like a “mechanical integrator” with an infinite d.c. gain. .......................... 97
4.9 The nonlinear plant is replaced by a set of linearized plants at different operating points. ......................................................... 98
4.10 With one unstable pole, \(L(s)\) must start from \(-180^\circ\) on the Nichols chart to result in a stable system, thereby producing a finite lower gain margin. ..................... 100
4.11 Flow chart for LTI proportional-gain controller design for a set of linearized unstable plants. Final controller design \(k_p\) must ensure adequate phase and gain margins for all the linearized plants. .................................................. 104
4.12 Maximum phase margin, and corresponding crossover frequency, lower gain margin, upper gain margin and controller gain are plotted from (a) to (e) as a function of normalized displacement. Selected LTI controller is \(k_p = 16\) which gives a minimum
phase margin around 60° at all displacements.

4.13 (a) Effective damping ratio as a function of normalized displacement. (b) Ratio of effective resonant frequency over mechanical resonant frequency as a function of normalized displacement.

4.14 (a) Open-loop transfer function of LTV design on the Nichols chart. (b) Open-loop transfer function of LTI design on the Nichols chart.

4.15 Maximum controller gain $k_p$ versus normalized plate displacement by stability requirement.

4.16 Ratio of disturbance force over first-order electrostatic force plotted as a function of normalized displacement for $\Delta z/Z_o = 0.025$ to 0.1.

4.17 High-order terms of the electrostatic force expanded by the Taylor’s series are formulated as an input disturbance force.

4.18 Perturbed displacement and voltage shown as an elliptical disc region around an operating point.

4.19 Concept of asymptotically stable.

4.20 The minimum controller gain plotted as a function of normalized displacement for input disturbance rejection. Also plotted are the optimal gain from 4.12(e) for an LTV design, and the maximum applied gain to maintain a minimum phase margin of 20°.

4.21 System block diagram at equilibrium point $z = Z_o$.

4.22 Percentage steady-state error as a function of normalized displacement. ($k_p = 16$)...

4.23 System block diagram at a new equilibrium point $z = Z_o + \Delta z_s$ due to the external shock force applied to the actuator.

4.24 Minimum required controller gain versus normalized displacement for suppressing the induced displacement from an 150 g shock force to 100 nm, 10 nm, and 1 nm...

4.25 Flow chart for computing QFT bounds [82].

4.26 Typical QFT bounds displayed on the Nichols chart for design of the loop transmission $L_o(j\omega)$. (type I: line. Type II: closed contour.)

4.27 (a) Templates of linearized unstable plants at various frequencies. (b) Robust stability bounds derived for the nominal unstable plant.

4.28 New stability lines on the Nyquist plot (left) and the Nichols chart (right), derived for unstable and/or non-minimum phase plants [84].

4.29 Robust stability bounds and stability lines derived for the new stable nominal plant.

4.30 Shaping of nominal open-loop function based on the derived bounds, such that $L_o'(j\omega)$ does not fall within the area as indicated in Figure 4.24. ($k_p = 30$)...

4.31 Input-disturbance bounds and loop-shaping result using $k_p = 106$.

4.32 Maximum phase margin, and corresponding crossover frequency, lower gain margin, upper gain margin and controller gain plotted as a function of normalized displacement from (a) to (e).

5.1 Photograph of custom assembly for setting initial gap and static interferometric measurement.

5.2 Schematic of custom interferometric setup for static-displacement measurement. Laser alignment is used for adjusting parallelism between the gold-coated glass and the
actuator die. ................................................................. 136
5.3 (a) SEM of released tip actuator fabricated by the AMS 0.5 \( \mu \)m CMOS process. (b) Structural curl measured by the WYCO white-light interferometric profiler. . . 137
5.4 Side view schematic of the actuator. ............................................. 138
5.5 The static displacement-voltage characteristic measures the actuator spring constant at 0.17 N/m. ................................................................. 139
5.6 Characteristic of pre-amp output versus plate displacement results in an extracted pre-amp input capacitance of 333 fF. ............................................ 140
5.7 Schematic of capacitive sensing circuit, illustrating a sensed voltage error introduced by capacitive feedthrough from actuation voltage, and reduction of feedthrough by high-pass filtering. ................................................... 141
5.8 Measured frequency response of the actuator driven at different d.c. bias, illustrating the resonant frequency shift due to the spring-softening effect. .......... 142
5.9 Measured resonant frequency versus d.c. actuation voltage. ............... 143
5.10 (a) Measured resonant peak and corresponding NODAS simulation results plotted as a function of d.c. actuation voltage in the frequency-response measurement. (b) Extracted effective damping ratio and corresponding NODAS simulation results plotted as a function of d.c. actuation voltage. ................................................... 144
5.11 Open-loop responses to 2 kHz square-wave actuation force varying in applied voltage amplitude from 1 V to 4.8 V (gap = 2.45 \( \mu \)m). ....................... 145
5.12 Open-loop responses to 2 kHz square-wave actuation force varying in applied voltage amplitude from 1 V to 7 V (gap = 3.2 \( \mu \)m). ....................... 146
5.13 (a) Pre-amp output spectrum measurement to a 2 kHz square-wave actuation voltage with an amplitude of 0.2 V. (b) Measured input-referred sensed voltage as a function of actuation gap. (c) Equivalent input-referred capacitance change as a function of actuation gap. ................................................................. 147
5.14 Optical photograph of the PCB with the tip actuator package and off-chip feedback control circuit. ................................................................. 148
5.15 (a) Measured controller output waveforms when displaced plate enters pull-in region and beyond. The inset illustrates its waveform when the input command turns from high to low. (b) Measured demodulated pre-amp output waveforms. (c) Plate displacement versus input command voltage. (gap = 3.3 \( \mu \)m) ....................... 149
5.16 Schematic representation of parallel-plate actuator design for closed-loop NODAS simulation. ................................................................. 151
5.17 Measured waveforms (dashed lines) and simulated waveforms (solid lines) of pre-filter output, controller output, demodulated pre-amp output, and plate displacement to a 50 Hz square-wave input-command voltage. ....................... 152
5.18 (a) Damping ratio derived from tilted plates and parallel plates as a function of normalized displacement. (b) Phase margin derived from tilted plates and parallel plates as a function of normalized displacement. ....................... 154
5.19 Noise gain from sensor output to controller output plotted for linearized plants at different displacements. ................................................................. 155
B.1 Schematic of feedback control board. Controller gain \( k_p = 24.3 \) is split into two parts \((24.3 = 10 \times 2.43)\). The gain of ten is placed after subtraction of the initial sensed voltage. ................................................................. 177
B.2 Schematic of demodulation circuit. ................................................................. 178
List of Tables

2.1 Measured pull-in voltage of polysilicon beam actuators compared with NODAS and electro-mechanical finite-element simulation results. ................................. 21
3.1 Measured effective Young’s modulus (in GPa) of composite beams from MOSIS N78H-AV run. ................................................................. 31
3.2 Measured effective Young’s modulus (in GPa) of composite beams from N6CJ-AV run and comparison with average values of N78H-AV run. ............... 31
3.3 Residual stress (in MPa) measured from MOSIS N78Q-AH, N79V-AQ, and N78H-AV runs. ................................................................. 32
3.4 Radius of curvatures (in mm) measured from dice of MOSIS N78Q-AH, N79V-AQ, and N78H-AV runs. ......................................................... 34
3.5 Transistor dimensions in the folded-cascode amplifier design. ............... 66
3.6 Transistor dimensions in the constant-transconductance bias circuit........ 68
5.1 Measured curl and plate tilt of the tip actuator. ................................. 138
5.2 Open-loop risetime and fall measurements (gap = 2.45 µm). ................. 143
5.3 Open-loop risetime and fall measurements (gap = 3.2 µm). ................. 146
Acknowledgements

I would like to express my sincere gratitude to my research advisor, Professor Gary K. Fedder, for his guidance, advice, and encouragement over the years of my research work at Carnegie Mellon. I’ll always cherish those great memories of working with him for the rest of my life! Many thanks to my other three thesis committee members, Professor L. Richard Carley, Professor James A. Bain, and Dr. Storrs Hoen for reviewing the manuscript and giving me invaluable advice. I am grateful to Dr. Tamal Mukherjee for his many research suggestions and help with simulation software, and to Professor Bill Messner for discussions on nonlinear control system.

I would like to thank many friends in the MEMS lab for their friendship, help, and many discussions that have pushed my research forward. I am very grateful to Xu Zhu for his help on chip-releases and taking those great-looking SEM photos, to Huikai for discussions on many optical questions and help on fabrication, to Hasnain Lakdawala for discussions on circuit design and feedback control, to Hao Luo for discussions on control circuit implementation and PCB design, to Qi Jing, Jan Vandemeer, and Sitaraman Iyer for their help on NODAS simulations, to Jiangfeng Wu and Gang Zhang for their many circuit design insights, to Bikram Baiyda for his help on layout extraction, and to Kai He for his Cadence support. I would like to thank many former and present officemates for the numerous things that they have helped me with over the years: Michael Stout, Steven Eagle, Dr. John Neumann, George Lopez, Lauren de Rosset, Dan Gaugel, Matthew Zeleznik, Brett Diamond, Kevin Frederick, Nilmoni Deb, John Ramsey, Keith Rebello, and Michael Kranz. Special thanks to Dr. Lars Erdmann for his help with the interferometric setup, which allowed me to complete the parallel-plate servo experiment.
It has been a wonderful experience to work on the MEMS data storage project, and to share expertise with many professors and fellow students (not mentioned elsewhere): Dr. Art Davidson, Professor Dave Greve, Dr. David Guillou, Suresh Santhanam, Fernando Alfaro, Seungook Min, Shan Liu, Xiaochun Wu, Li Zhang, and Rany El-Sayed. Many thanks to the cleanroom staff, Chris Bowman, Robert Harris, and Carsen Kline for their assistance which made our research work much easier.

Many thanks to Drew Danielson, and the MEMS lab administrator, Mary Moore, who has been a tremendous help for me in so many things, and truly made my research life a whole lot easier. I thank Lou Anschuetz, Jim McKinney, Matt Henson, and Dave Wilhelm of the ECE Computer facilities for the excellent computer facilities and services. Thanks to the ECE administrative staff, Elaine Lawrence and Elaine Zurcher, for their help in many things over the years.

I am grateful to the financial support from Defence Advanced Research Project Agency (DARPA). Otherwise this research work would not be made possible.

I would like to express my deepest gratitude to my parents for their consistent encouragement and support in so many ways during the long years of my graduate study. I truly thank them for their support of my graduate study in the U.S., and their forever love and belief in me. I sincerely thank my parents-in-law, who have always supported me and my family with their greatest love. I thank them very much for making great efforts to visit us every half a year from Taiwan because we were not able to come home very often.

Finally, I would like to express my deepest love and gratitude to my wife, Ching-Yu, for her love, support, and understanding during the course of my study. Her dedication to our family and children, Chun-Yen (Steven), Chun-Zho (Michelle), and Chun-Yi (Kath-
ryn), allowed me the greatest freedom to work towards my degree. I thank her very much for making me a better person and a better father, and I look forward to being a better husband and spending more time with her and our family!

So long, Pittsburgh. Thanks for the memory!
Chapter 1
Introduction

Electrostatic actuation is very attractive and widely used in microelectromechanical systems (MEMS) because of its low-power consumption, relative ease of fabrication, high energy density, and good scaling property with respect to inertial forces [1]. The two most commonly used actuation design concepts are based on the comb-finger and parallel-plate structures. Parallel-plate actuators generate larger forces than comb drives because intended displacements are in parallel with the major electric field lines. However the maximum displacement is limited by initial gap separation. Furthermore, a positive feedback mechanism exists in parallel-plate actuators such that the electrostatic force increases with respect to the decrease of gap, leading to the well-known pull-in instability [2][3] which occurs at one-third of the gap during open-loop operation.

Despite the relatively small displacement, parallel-plate actuation is preferred over comb drives in many applications [4-8] in which a small design form factor and/or large forces perpendicular/parallel to wafers are needed. In those applications, maximization of the travel range can be critical to system specifications, such as the maximum wavelength of light which can be effectively controlled in the optical transfer function of a polychromator reflective phase grating [11]. Proposed methods for extending travel range by open-loop operations include touch-mode actuators [9][12][13], curved-electrode actuators [10][14], the leverage bending method [15], nonlinear spring design [16], and the series capacitor approach [17][18]. Motion of touch-mode actuators using dielectric-coated electrodes are difficult to control due to stiction effects. For the others, large displacement is commonly achieved at the cost of a higher driving voltage. For the series capacitor approach, a capacitor is placed in series with the actuated capacitance. The actuation voltage and charge across the actuated capacitance is controlled by the voltage divider. Drawbacks of this approach include: (1) high driving voltage (up to five times of the pull-in voltage to close the entire gap), (2) maximum displacements are limited by parasitic capacitance, and (3) undesirable charge accumulated at the floating high-impedance
node between the actuated and series capacitances requires being reset by a switch. Switch turn-off leads to charge injection into the actuated capacitor, which can be destabilizing if the actuated capacitance is only several femtofarads. A stable travel distance up to 60% and 88% of the gap has been reported in [17] and [15], respectively.

Charge control using current drives [19-24] and a switched-capacitor circuit [25] has been reported to extend the travel range. For the current-drive implementation, it must have very small leakage current to control the injected charge into actuators. It can be very difficult to implement a current source with leakage current in the order of fA when actuated capacitance is only a few fF. A reset scheme for discharging the accumulated charge is required. Periodic reset can cause ringing around the final steady-state position [20]. For the switched-capacitor approach, it also must consider problems related to leakage current, charge injection, and parasitic capacitance.

Voltage control by closed-loop feedback is more attractive than open-loop methods because the desired actuation force can be fully generated using a lower driving voltage in the design phase, leading to a smaller design layout. The biggest challenge inherited by this approach is to design and implement the feedback system around a parallel-plate actuator, which is nonlinear, and exhibits unstable behavior beyond the pull-in limit. There have been only a few studies in the past regarding the closed-loop stability analysis and simulation [26][27]. To our best knowledge, no measurement on parallel-plate servo past the instability point has been reported.

In this thesis, we present closed-loop voltage control to extend travel range of a parallel-plate actuator beyond the pull-in limit. Research work on parallel-plate servo originates from development of a probe-based magnetic micro disk drive [28] in Carnegie Mellon University. This is a non-volatile, rewritable magnetic storage device featuring low power, high storage capacity, and miniaturized size. These micro disk drives can be used in mobile computing devices, and provide a quantum leap for the storage capacity of non-volatile memory which is now mainly served by flash-memory chips. Based on the progress of lithographical process for the next ten years (minimum feature size λ = 70 nm by 2009), it is not likely to get several gigabytes of storage capacity on a cm² size flash chip (each bit cell = 3 to 4 λ). Instead, probe-based data storage can be one of the best solutions when bit size is approaching 100 nm and below.
One of the best-developed probe-based data storage system is the IBM Millipede [29], which has demonstrated integration of a 32 × 32 tip array with a media of 200 Gb/in² storage density. The writing mechanism in this system uses tips to thermally melt pits in a polymer media. The readout signal is obtained by measuring the thermal conductance of the media. All the tips are moved up to the media for data access by actuation of the die, thereby no individual actuation and servo control of the tip is required. There are, however, a number of disadvantages of the read/write mechanism used in the Millipede. The data rate per tip is limited by the resonant frequency of the cantilevers and the contact mode for read/write raises questions about wear and durability of both probe tips and media. Also system power consumption can be negatively impacted by the thermal writing mechanism.

The target specifications of the Carnegie Mellon probed-based storage device include:

- Areal density = 10 GB/cm².
- Active power < 1 W.
- Standby power < 10 mW.
- Data transfer rate > 100 Mbits/s.
- Access time < 10 ms
- System volume = 2500 mm³.

In the envisioned system, a media-actuator die is flipped on top and bonded to a tip-actuator die at the bottom. For reading and writing a magnetic bit, the media actuator [30] with deposited media scans in a raster fashion in lateral directions (x and y), and an array of tip actuators are servoed in parallel to the top media (the z direction) to access data bits, as shown schematically in Figure 1.1. The read/write head is located on a square pallet carried by the tip actuator. A design of the magnetic probe head in plan view and side view is shown in Figure 1.2. This design is intended to allow a fine magnetic tip to be brought within 10 nm of the media surface for reliably reading and writing magnetic bits. Assuming an initial gap spacing of 3 μm between the pallet and the media, and a tip height of 500 nm, the displacement of each tip actuator has to be servoed up to 83% of the gap.
Figure 1.1 Schematic representation of the envisioned micro disk drive.

Figure 1.2 Top view and side view of the read/write probe head design.
The plan for the writing mechanism is to use the tip to generate local magnetic fields perpendicular to the media to reverse the magnetic domains. Magnetic fields are produced by a coil wrapped around the tip. If the field proves insufficient, it can be assisted by providing thermal energy to the media with a field emission current from the tip, in order to lower the required field for switching domains. A yoke wing with a large area of magnetic material is designed to efficiently close the magnetic circuit with the top media, as shown in the side view in Figure 1.2. Also shown is the reading process accomplished by conducting the magnetic flux from a magnetic bit through the tip and to a magneto-resistive (MR) sensor.

In designing the tip actuator, a key criterion is that the tip-actuator cell size is limited and must be equal to the media stroke in order to achieve 100% media coverage. Tip actuators are fabricated in a conventional CMOS process for convenient integration of control circuits and data channels. A total of six leads running through the actuator, including for reading, writing, and servo, can be accommodated by a two-anchored spring design, with each spring providing three leads. Minimum aluminum width, also the beam width, should be wide enough to provide sufficient capacity for the writing current (around 1 mA) in order to prevent electromigration.

In addition to stabilizing the parallel-plate actuator in the open-loop unstable regime, feedback control is desired to deal with the presence of plant uncertainties. For example, the electrostatic force changes due to the variation of actuation gaps, which may result from a tilted media after die-bonding, a curved media, and/or the variation of actuator curl, as depicted in Figure 1.3. Other uncertainties can include the variation of actuator spring constants across a die due to varying beam thicknesses, and the variation of damping coefficients due to varying actuation gaps. The implemented control circuit is preferred to be simple to fit into the actuator cell size. Thereby we select a linear time-invariant (LTI) controller design for its easier implementation by analog circuits than other alternatives, such as a linear time-varying (LTV) controller or a nonlinear controller.

A schematic representation of the feedback system is illustrated in Figure 1.4.
plex scheme in which the sensed signal with respect to the actuator displacement is obtained through modulation, amplification, demodulation, and low-pass filtering to produce the base-band signal as the input to the controller. The goal of this thesis will be focused on the control theory and experimental demonstration. Total circuit size and circuit power consumption will not be included as design constraints in the following presentation.

In parallel to our research efforts on parallel-plate servo using the CMOS-MEMS process, in Chapter 2, we have derived electrostatic gap models for use in a hierarchically structured MEMS design approach \cite{32}\cite{33} called NODAS, for Nodal Design of Actuators and Sensors. Physics-based parameterized gap models of comb fingers and parallel-plates are derived through a series of finite-element simulations and mathematical modeling. Derived models are aimed for low-aspect-ratio microstructures to account for the significant fringing capacitances, and expressed in terms of electrode dimensions, gap separations, and sidewall angles.

In Chapter 3, we review the CMOS-MEMS fabrication process. Material properties of the CMOS-MEMS microstructures are measured as the first attempt to establish the CMOS-MEMS material property database. Design and modeling aspects of the tip actua-
Chapter 4 details the controller design procedure for parallel-plate servo. The controller design for the highly nonlinear and unstable plant is based on linearization and disturbance-rejection techniques to ensure stability robustness. Design issues regarding steady-state error and shock rejection are also discussed. The Quantitative Feedback Theory (QFT) design technique is presented at last for robust controller design.

In Chapter 5, we present the experimental results on open-loop plant characterization and the closed-loop servo. The mechanical resonant frequency, damping ratio, spring constant and open-loop step responses of the tip actuator are measured for model validation. Measured closed-loop step response are compared with NODAS simulations in
which 3D elastic beam models, plate models, and squeeze-film damping models are used to fully represent the tip actuator dynamics.

In the final chapter, we form conclusions about the parallel-plate actuator servo, and outline some future research in this area.
Chapter 2
Electrostatic Gap Models

2.1 Introduction

A top-down hierarchical design using nodal simulation, called NODAS (for Nodal Design of Actuators and Sensors) has been developed at Carnegie Mellon [32][33] for design of complex microelectromechanical systems with a large number of multi-domain components. At the lowest level in the design hierarchy, MEMS components are represented as interconnected combinations of atomic behavioral models with geometric parameters. The set of elements, for example, includes beams, plates, and electrostatic gap models [31]. In this chapter, physics-based parameterized electrostatic gap models are derived for design of electrostatic actuators and capacitive sensors using NODAS.

We start with analytic derivation of electrostatic forces between two charged conductors, followed by the development of parameterized electrostatic gap models. A capacitance macromodel on a fixed-geometry accelerometer has been reported [34], in terms of in-plane and out-of-plane spatial coordinates but without geometric parameterization. Parameterized models of comb fingers with rectangular cross-section have been derived using the conformal mapping method [36][37]. In the low-level MEMS circuit schematic, electrostatic gap elements are interconnected with beam and plate elements to allow representation of most MEMS devices. By generating a set of gap models having geometric parameters, these MEMS circuit schematics can be simulated without resorting to intermediate modeling steps within the design loop

Derived models are based on low-aspect-ratio surface-micromachined microstructures, in which fringing capacitances become a significant portion of the total capacitance, and need to be characterized in order to better simulate the electromechanical behavior of devices. Parameterized electrode dimensions include the gap separation, the electrode thickness, the electrode width, and the sidewall angle. Similar research effort on parameterized capacitance modeling is found for estimation of interconnect delay [38] and microwave transmission line design [39], to replace time-consuming numerical simulations.
2.2 Electrostatic Force

We consider the electrostatic force between two charged conductors, as shown in Figure 2.1. The conservation of power of the system [40] can be written as,

\[
\frac{dW_e}{dt} = V \frac{dq}{dt} - F_e \frac{dz}{dt} \quad (2.1)
\]

where \( W_e \) is the electrical energy stored in the system, \( V \) is the voltage between the conductors, \( q \) is the amount of charge on each conductor, and \( z \) is the displacement of the top conductor. The term \( V(dq/dt) \) is the power input at the electrical terminals, and \( F_e(dz/dt) \) is the mechanical power output. As shown in Figure 2.1, both the electrostatic force and the displacement are defined in the same direction to satisfy the minus sign in (2.1), giving a positive \( F_e(dz/dt) \) product when the resultant force and velocity are in the same direction. Multiplication of (2.1) by \( dt \) gives the equation for conservation of energy,

\[
dW_e = Vdq - F_e dz \quad (2.2)
\]

The partial differentiation of (2.2) with respect to \( z \) yields,

\[
F_e = -\left. \frac{\partial W_e}{\partial z} \right|_q \quad (2.3)
\]

The evaluation of the change in \( W_e \) when \( q \) and/or \( z \) are varied requires a line integration of (2.2) through a two-dimensional variable space. For a conservative system, the stored energy is a function of its initial and final states only, and does not depend on what path through variable space is used to reach the final state.
The path of integration in the $q$-$z$ plane to be used in evaluating stored energy $W_e$ is shown in Figure 2.2(a). By integration of (2.2), the energy at point $(q, z)$ is,

$$W_e = \int_0^q -F_e(0, z')dz' + \int_0^q V(q', z) dq'$$

where the first term on the right of (2.4) is zero because $F_e$ is zero due to no charge ($q = 0$). Thus (2.4) can be written as,

$$W_e = \int_0^q V(q', z) dq'$$

The amount of charge on conductors is

$$q = C(z)V$$

where $C(z)$ is the capacitance expressed as a function of displacement $z$. Substituting (2.6) into (2.5) to eliminate $V$ gives,

$$W_e(q, z) = \frac{q^2}{2C(z)}$$

Substituting (2.7) into (2.3) yields,

$$F_e = \frac{q^2}{2C(z)^2} \frac{dC(z)}{dz}$$

Figure 2.2 Path of integration in variable space. (a) For evaluating energy $W_e$. (b) For evaluating coenergy $W_e'$. 
(2.8) represents $F_e$ with charge as the variable. It will be wrong to apply (2.3) to (2.7) with $q$ replaced by $CV$, in order to represent force with applied voltage. The magnitude of the resulting force will be correct, but the sign will be wrong. The magnitude will also be incorrect if $C$ is a function of $V$, i.e., if the $q$-$V$ relation is nonlinear. The correct force expression by voltage can be derived through coenergy, $W'_e$, given by,

$$W'_e = qV - W_e$$  \hspace{1cm} (2.9)

From $dW_e = d(qV) - dW_e$ and (2.2), we obtain,

$$dW'_e = qdV + F_e dz$$  \hspace{1cm} (2.10)

The partial differentiation of (2.10) with respect to $z$ yields,

$$F_e = \left. \frac{\partial W'_e}{\partial z} \right|_V$$  \hspace{1cm} (2.11)

By using the path of integration defined in Figure 2.2(b), the integration of (2.10) reduces to

$$W'_e = \int_0^V q(v', z) dv'$$  \hspace{1cm} (2.12)

Substituting (2.6) into (2.12) gives,

$$W'_e = \frac{1}{2} C(z)V^2$$  \hspace{1cm} (2.13)

and (2.3) becomes,

$$F_e = \frac{1}{2} \frac{dC(z)}{dz} V^2$$  \hspace{1cm} (2.14)

The parallel-plate actuator shown schematically in Figure 2.3 produces electrostatic force in the $z$ direction. Neglecting fringing fields, the parallel-plate capacitance is,

$$C = \frac{\varepsilon_o A}{g - z}$$  \hspace{1cm} (2.15)
where $\varepsilon_o$ is the permittivity of free space ($8.854 \times 10^{-12}$ F/m), $A$ is the plate area, $g$ is the initial air-gap spacing, and $z$ is the displacement of the movable plate at the bottom. Application of (2.14) gives the parallel-plate force,

$$F_e = \frac{\varepsilon_o A}{2(g-z)^2} V^2 \quad (2.16)$$

The comb-finger actuator shown schematically in Figure 2.4 was first illustrated by Tang [41]. The comb drive produces lateral displacement in the $x$ direction which is perpendicular to the major field lines. Neglecting fringing fields, the comb-finger capacitance is,

$$C = 2N \frac{\varepsilon_o (l+x)h}{g} \quad (2.17)$$
where $N$ is the number of rotor fingers, $l$ is the initial finger overlap, $x$ is the displacement of rotor fingers, $h$ is the finger thickness, and $g$ is the gap between comb fingers. Application of (2.14) gives the lateral comb-finger force,

$$F_e = \frac{N\varepsilon_0 h}{g} V^2$$

(2.18)

### 2.3 Gap Modeling

The three electrostatic gap models studied are two electrodes of the same width and of different widths (one is much wider than the other), and the comb-finger array, as illustrated in Figure 2.5(a), (b), and (c), respectively. The models are geometrically parameterized by gap separation ($g$), thickness ($h$), width ($w$) and sidewall angle ($\theta$). The geometries studied do not include an underlying ground plane. However, the modeling approach is extendable to such actuators.

No closed-form solution has been derived to take sloped sidewalls into account by the conventional conformal mapping method. To generate accurate gap models,

![Figure 2.5 Schematics of the electrode geometry for (a) two electrodes of the same width, (b) two electrodes in which one is much wider than the other, and (c) comb fingers.](image-url)
numerical data produced from finite-element simulations [42] are analyzed by least-squares fits. The capacitance equation used in fitting is:

\[
c = \frac{\varepsilon_0 h}{g} \left( K_1 \left( \frac{w}{h}, \theta \right) + \frac{g}{\pi h} K_2 \left( \frac{w}{h}, \theta \right) + K_3 \left( \frac{w}{h}, \theta \right) \ln \left( \frac{\pi h}{g} \right) \right)
\]  

(2.19)

where \( c \) is capacitance per unit length, and \( K_1, K_2, \) and \( K_3 \) are functions of cross-section geometry. (2.19) has a physical basis as the analytical equation derived for a zero-thickness plate on top of an infinite ground plane [43]. The capacitance expression depends on the ratios of electrode dimensions and gap separations.

Values of \( K_1, K_2, \) and \( K_3 \) are first obtained for every \( \left( \frac{w}{h}, \theta \right) \) pair by least-squares fitting of its corresponding capacitance vector with respect to the \( g/h \) ratio. Then a surface-fit on \( w/h \) and \( \theta \) gives the final coefficient functions. As an example, values of capacitance per unit length for the two electrodes of the same width are shown in Figure 2.6 for \( w/h \in [0.5, 5] \), \( g/h \in [0.1, 3] \), and \( \theta \in [1.33, \pi/2] \) (in radians. i.e., 76° to 90°).

Each \( \left( \frac{w}{h}, \theta \right) \) is corresponding a capacitance curve drawn with respect to the \( g/h \) ratio in the mesh. \( K_i \) values for each curve are obtained by curve-fitting using the capacitance expression in (2.19), as plotted in Figure 2.7. Non-physical basis functions are cho-
sen out of the set \((w/h)^{\alpha}, \theta^\beta, \ln(w/h), \ln(\theta), e^{(w/h)}, e^\theta,\) and out of products of these functions. The final coefficient functions listed below are obtained through another surface-fit:

\[
K_1 = -1.565 + 0.2818\left(\frac{w}{h}\right)^{-0.04348} - 2.986\theta^{5.9} + 0.4446e^{\theta^{6.002}} + 28.87\ln(\theta) \quad (2.20)
\]

\[
K_2 = 21.43 - 15.5\left(\frac{w}{h}\right)^{-0.02146} - 10.07e^{\theta\left(\frac{w}{h}\right)^{-0.03944}} - \theta^{-1.877} + 23.7\theta^{0.09913} + 0.0484e^{(w/h)}\theta^{0.8278\left(\frac{w}{h}\right)^{-2.244}} - 13.6\ln(\theta) \quad (2.21)
\]

\[
K_3 = 10.86 - 1.512\left(\frac{w}{h}\right)^{-0.2517} - 0.007964\theta^{15.31} + 0.005087\ln\left(\frac{w}{h}\right) - 62.71\ln(\theta) + 20.05\theta^{2.529}\ln(\theta)^{\left(\frac{w}{h}\right)^{-0.002656}} \quad (2.22)
\]

Using the same procedure for the case in which one electrode is much wider than the other, the coefficient functions are:

\[
K_1 = -1.353 - 2.827\theta^{5.988} + 0.4272e^{\theta^{6.076}} - 0.01112\ln\left(\frac{w}{h}\right) + 28.06\ln(\theta) \quad (2.23)
\]
Gap models of comb fingers are constructed for fingers at the center and the edge. The finger capacitance gradually increases as it moves away from the center. The coefficient functions for comb fingers close to the center are:

\[ K_2 = 2.058 - 0.1588 \ln \left( \frac{w}{h} \right) - 0.2527 \ln(\theta) + 30.9 \left( \frac{w}{h} \right)^{0.2687} \ln(\theta) \theta^{-3.177} \]  
\[ K_3 = 12.46 - 0.03697 \theta^{12.79} + 0.3212 \ln \left( \frac{w}{h} \right) - 109.4 \ln(\theta) \]  
\[ + 45.38 \theta^{1.989} \ln(\theta) \left( \frac{w}{h} \right)^{-0.001315} \]  

\[ K_1 = -1.441 - 0.0291 \left( \frac{w}{h} \right)^{-0.3762} - 2.2239 \theta^{6.254} + 0.3313 e^{\theta^{6.412}} + 25.28 \ln(\theta) \]  
\[ K_2 = -9.466 + 12 \left( \frac{w}{h} \right)^{-0.1935} - 1.137 \times 10^{-4} e^{\theta^{14.78}} + 2.998 \ln \left( \frac{w}{h} \right) \]  
\[ + 0.0895 \ln(\theta) \]  
\[ K_3 = 7.32 - 1.36 \left( \frac{w}{h} \right)^{-0.2075} + 0.63 \theta^{8.022} - 0.01269 e^{\theta^{12.24}} + 0.1648 \ln \left( \frac{w}{h} \right) \]  
\[ - 29.42 \ln(\theta) \]  

The coefficient functions for comb fingers at the edge are:

\[ K_1 = -1.799 + 0.5043 \left( \frac{w}{h} \right)^{-0.02533} - 2.97 \theta^{5.918} + 0.4446 \theta^{6.013} + 28.83 \ln(\theta) \]  
\[ K_2 = -0.4472 + 3.009 \left( \frac{w}{h} \right)^{0.2986} + 0.03332 e^{\theta^{1.226}} \left( \frac{w}{h} \right)^{-0.7201} - 3.385 \times 10^{-4} \theta^{16.02} \]  
\[ - 0.04904 \ln(\theta) \]  
\[ K_3 = 178.038 - 1.969 \left( \frac{w}{h} \right)^{0.207} + 347.593 e^{\theta \left( \frac{w}{h} \right)^{0.001652}} \theta^{14.847} - 395.697 \theta^{-4.252} \]  
\[ - 263.609 \ln(\theta) \]
By substituting the ratios of $w/h$ and $g/h$, and the sidewall angle $\theta$, values of capacitance per unit length are calculated using (2.19) to (2.31), and compared with the original data. For the case of two electrodes of the same width, it has an average deviation of 1.6% and a maximum deviation of 5.7%. The case with a single wider electrode has an average deviation of 1.4% and a maximum deviation of 4.9%. For comb fingers, center capacitance has an average deviation of 2.4% and a maximum deviation of 13.2%. The capacitance at the edge has an average deviation of 2.3% and a maximum deviation of 9.2%.

The gap model is connected to the nodes of two adjacent beam or plate elements at the ends. The through variable and across variable at a node are force and displacement, respectively. The capacitance and electrostatic forces are calculated by the separation, overlap, and beam dimensions. The comb-finger force, which is parallel to the beams, and denoted as $F_{comb}$, is applied directly to the nodes. By using (2.13) and (2.19), the comb-finger force magnitude is given by,

$$F_{comb} = \frac{1}{2} \frac{d((l + x)c)}{dx} V^2$$

$$= \frac{1}{2} c V^2$$

where $x$ is the displacement variable. The sum of parallel-plate force per unit length $\sigma$ of magnitude $\frac{1}{2} \frac{dc}{dz} V^2$ is lumped at the nodes as $F_1$ and $F_2$ by the principle of conservation of force and moment, as shown in Figure 2.8, where $F_1 = F_2 = \sigma l/2$. 

![Diagram](image)
2.4 Simulation

The geometry and the schematic representation of a cantilever beam actuator are shown in Figure 2.9(a) and (b), where $l = 100$ to $200 \, \mu m$, $w = 2 \, \mu m$, $h = 2 \, \mu m$, $\theta = \pi/2$, and $g = 2 \, \mu m$. A total of twenty beam instances (ten for the stator and ten for the movable beam) and ten gap instances are placed in the schematic. Behavioral models of MEMS elements are implemented in Analogy MAST® with simulation in Saber® [44]. Results of the NODAS and the electro-mechanical finite-element simulations [45] for pull-in voltage of the actuator as a function of beam length are plotted in Figure 2.10. A maximum of -1.7% deviation of pull-in voltage from the finite element simulation is obtained by NODAS simulations. Pull-in voltages in both simulations are inversely proportional to the beam length squared.

Beam displacement of an 100 $\mu m$ long beam with respect to the applied voltage is illustrated in Figure 2.11. Good agreement is shown except near the edge of pull-in point. For one data point, it takes NODAS simulation 1.7 seconds (CPU time) on the average to complete, compared to 280 seconds in electro-mechanical finite-element analysis.
Figure 2.10 NODAS and electro-mechanical finite-element simulations of pull-in voltage for $l = 100$ to $200\ \mu$m, $w = 2\ \mu$m, $h = 2\ \mu$m, $\theta = \pi/2$, and $g = 2\ \mu$m.

Figure 2.11 Displacement-voltage characteristics from electro-mechanical finite-element and NODAS simulations for an $100\mu$m-long beam actuator.
2.5 Experiment

Pull-in voltages of beam actuators fabricated by the polysilicon micromachining process (MUMPs [35]) are measured and compared with NODAS and finite-element simulations as listed in Table 2.1. Actual beam cross-section and gap separation are taken into account in both simulations \( (w = 1.7 \, \mu m, \, h = 2 \, \mu m, \, \theta = 1.496 \text{ radians, and } \, g = 2.3 \, \mu m) \); otherwise up to 25\% pull-in voltage deviation occurs with only nominal layout dimensions considered.

<table>
<thead>
<tr>
<th>( l(\mu m) )</th>
<th>Experiment (V)</th>
<th>NODAS (V)( (+%)</th>
<th>FEM (V)( (+%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>100</td>
<td>40.0</td>
<td>42.5( (+6.3%)</td>
<td>44.2( (+10.5%)</td>
</tr>
<tr>
<td>150</td>
<td>18.4</td>
<td>19.1( (+3.8%)</td>
<td>19.7( (+7.1%)</td>
</tr>
<tr>
<td>200</td>
<td>10.4</td>
<td>11.0( (+5.8%)</td>
<td>11.1( (+6.7%)</td>
</tr>
</tbody>
</table>

Table 2.1: Measured pull-in voltage of polysilicon beam actuators compared with NODAS and electro-mechanical finite-element simulation results.
Chapter 3

Plant Design and Modeling

3.1 Introduction

In this chapter, we will present the design and modeling of the controlled plant which consists of a micromechanical actuator driven by the parallel-plate electrostatic force, followed by a capacitive position sensor. Electrostatic actuation and capacitive sensing are used for their low power consumption which is ideal for the envisioned portable storage device. We begin with an introduction of device fabrication using CMOS-MEMS technology, and material property characterization of the CMOS-MEMS microstructures. The design and modeling aspects of the actuator and capacitive sensing circuit will be presented to establish a complete dynamic model of the plant for use in controller design, which is the topic of the next chapter.

3.2 CMOS-MEMS Fabrication

The conventional CMOS process has been used for MEMS fabrication in many ways, in which additional fabrication steps can be applied either before [46][47], in between [48], or after [49-51] the regular CMOS process. Our goal to achieve low-cost integration of micromechanical structures and high-performance electronics involves only a minimum of processing steps on CMOS layers after the completed CMOS process, thereby enabling a shift in research focus from processing details to the design of complex systems with multiple sensors and actuators on a single chip.
Microactuators described in this thesis are fabricated in a conventional CMOS processes followed by post-CMOS micromachining steps described in [52]. First, an anisotropic dielectric reactive-ion etch (RIE) with CHF$_3$/O$_2$ plasma defines the structural sidewalls using the top metal layer as an etch-resistant mask. Next, an anisotropic silicon deep reactive-ion etch (DRIE) is performed in an inductively-coupled-plasma etcher using SF$_6$ (etching) and C$_4$F$_8$ (passivation) plasmas. Finally microstructures are released from silicon substrate by an isotropic silicon etch with SF$_6$ plasma. Fabricated composite microstructures can consist of metal layers, dielectric layers, and polysilicon layers. A complete post-processing sequence is illustrated in Figure 3.1. A released crab-leg comb-drive resonator fabricated by the Agilent 0.5 µm three-metal n-well CMOS process is shown in Figure 3.2(a). A released beam cross-section with three metal layers (metal-3, metal-2, and metal-1) and one polysilicon layer is in Figure 3.2(b). When using dielectric RIE for post-CMOS processing, one must consider the following factors: (1) It is desirable to have high etch selectivity between the metal mask layer and dielectric layers (i.e., silicon dioxide and silicon nitride), and no selectivity between these dielectric layers in the vertical etch direction. (2) It is desirable to achieve directivity of sidewalls by control of passivation on sidewalls. However, too much polymerization on the surface will slow down dielectric etch and limit the smallest etched gaps and holes; too little polymerization does not provide enough sidewall protection, resulting in the loss of critical dimensions. A beam cross-section with excessive polymers stacked on sidewalls is illustrated in Figure 3.3. (3) It is essential to avoid electrical connection failure during etch. Failures can result from the removal of metal layers on vias connecting to the top metal layer, as shown in
Figure 3.1 Cross section of the CMOS-MEMS process flow. (a) After CMOS processing. (b) After anisotropic dielectric reactive-ion etch for definition of structural sidewalls. (c) After anisotropic silicon etch. (d) After isotropic silicon etch for structural release.
To summarize, the important etch effects include etch rate, loss of critical dimensions, survival of electrical connection, and polymer generated on the sidewall and in the field. The processing parameters in the plasma system which affect the RIE etch include gas flow rate, gas mixture, pressure, RF power, electrode spacing, electrode temperature,
electrode material, total wafer area (loading), and previous processing steps. Three critical processing parameters (pressure, power, and mixture of gases) are selected to quantify the process model in a Box-Behnken factorial experiment [53] reported by Zhu [54]. The CHF$_3$/O$_2$ gas mixture is chosen instead of the CF$_4$/O$_2$ mixture used previously in [52] because it can generate sufficient passivation with a reasonable etch rate and with fewer electrical connection failures. The dielectric RIE is performed in a Plasma-Therm 790 reactor. The processing region for achieving minimal lateral etch, minimum polymerization, and electrical continuity of vias is determined, resulting in the process parameters: 125 mTorr chamber pressure, 0.55 W/cm$^2$ RF power density, CHF$_3$ flow at 22.5 sccm, and O$_2$ flow at 16 sccm. Experimental results summarized in [54] indicate: (1) Etch rate increases with increasing RF power and chamber pressure, and is not significantly affected by O$_2$ concentration. (2) Increasing RF power and decreasing pressure reduce passivation. (3) High-power operation must be accompanied by increasing chamber pressure to reduce the loss of critical dimensions and opened vias resulting from the ion-mill-
ing effect. At an etch rate of 425 Å/min, it takes about two hours to etch through all the dielectric layers (~5 µm thick).

The previous processing procedure described in [52] does not have the anisotropic silicon DRIE step inserted between the dielectric RIE and the isotropic release etch. Microstructural release is performed directly after dielectric RIE. The drawback is that the vertical etch and the lateral etch are coupled in an isotropic etch, hence the circuitry must be placed at least about as far as the silicon is etched down. The primary benefit of using the anisotropic silicon DRIE is that the vertical etch and the lateral etch are decoupled. Separation between the sensing device and silicon substrate is wider than achieved in [52], therefore parasitic capacitances can be reduced. The amount of the required lateral etch for structural release is also reduced depending on designed dimensions. The anisotropic silicon DRIE is performed in a Surface Technology Systems (STS) inductively-coupled-plasma (ICP) etcher. The process parameters for the etch part include 600 W coil power, 12 W platen power, SF₆ flow at 130 sccm, O₂ flow at 20 sccm, and 23 mTorr chamber pressure. The fluorocarbon polymer passivation cycle is performed using 600 W coil power, C₄H₈ flow at 85 sccm, 12 mTorr chamber pressure, and no platen power. The duration of the etching and passivation cycles are 12 seconds and 8 seconds, respectively, with achieved etch rate of 2.9 µm/min.

The isotropic silicon etch is also performed in the STS ICP etcher, because there is no noticeable decrement of aluminum layer thickness, as observed commonly in the parallel-plate system (e.g., Plasma-Therm 790 etcher). A Box-Behnken factorial experiment [53] has been performed with three major processing parameters: pressure, platen power, and SF₆ flow. The experimental results [55] show that the etch rate is primarily
determined by the pressure, followed by SF$_6$ flow, and not significantly affected by platen power especially under high pressure. The processing parameters are: 600 W coil power, 50 mTorr chamber pressure, SF$_6$ flow at 50 sccm, and no platen power.

3.3 Material Property Characterization

In this section, measured material properties of CMOS-MEMS microstructures are presented. Devices are fabricated by the Agilent 0.5 $\mu$m CMOS process available through the MOS Implementation Service (MOSIS). Reported material properties include effective Young’s modulus in the lateral direction, residual stress and vertical stress gradient [56] measured on dice from the same run, and dice from different runs as well. A durability test is performed on a resonating fatigue structure to demonstrate the effect of cyclic stress on composite micromechanical structures.

3.3.1 Effective Young’s Modulus

Effective Young’s modulus is determined by measuring the lateral resonant frequency of simple cantilever-beam actuators shown in Figure 3.5. Beams are excited elec-
trostatically by the actuators located near the tips of the cantilevers. The measured frequencies are substituted into the analytic equation for resonance frequency of a homogeneous cantilever beam [57],

$$f_r = 0.56 \sqrt{\frac{EI}{\rho AL^4}}$$ (3.1)

where \( L \) is the beam length, \( E \) is the effective Young’s modulus, \( I \) is the moment of inertia of the beam cross-section, \( \rho \) is the effective mass density, \( A \) is the beam cross-section area, and the unit of frequency in Hertz. The effective mass density of composite beams is the average density of constituent materials given by,

$$\rho = \frac{(\rho_{Al}A_{Al} + \rho_{ox}A_{ox} + \rho_{poly}A_{poly})}{A}$$ (3.2)

where \( \rho_{Al} \), \( \rho_{ox} \), and \( \rho_{poly} \) are the mass density of bulk aluminum, silicon dioxide, and polysilicon with values of 2700 kg/m\(^3\), 2500 kg/m\(^3\), and 2330 kg/m\(^3\), respectively. \( A_{Al} \), \( A_{ox} \), and \( A_{poly} \) are the cross-section areas occupied by aluminum, silicon dioxide and polysilicon.

Composite beams with fourteen combinations of metal and polysilicon layers are measured. Each cantilever is designed 2.1 \( \mu \)m wide and 90 \( \mu \)m long in layout. The actual dimensions of beam cross-sections are measured from scanning electron micrographs to calculate the moment of inertia and the effective mass density. Table 3.1 summarizes the effective Young’s modulus measured on three dice from the MOSIS N78H-AV run of Agilent 0.5 \( \mu \)m CMOS. The volume ratio of dielectric to metal in the right column illustrates that effective Young’s modulus tends to increase when structures contain more dielectric than metal. Another measurement of five commonly used composite beams on three dice
from the MOSIS N6CJ-AY run are tabulated in Table 3.2, with previous results from the
N78H-AV run for comparison. Both die-to-die and run-to-run variation of Young’s modu-
lus are within ±8%, and within ±3% for m3-m2-m1-poly beams.

### 3.3.2 Residual Stress and Vertical Stress Gradient

Residual stress is measured using the bent-beam strain sensor [58] shown in Fig-
ure 3.6(a). Relative apex displacement of the bent-beam rendered by the residual stress is
read out from a vernier. Residual stress of seven composite test structures was analyzed by
thermal-mechanical finite-element simulations in MEMCAD [45], in which the drawn model includes the bent-beam strain sensor and part of the anchor released from the substrate. Measured vernier displacement is matched in the simulation by introducing a stress, \( \sigma = E \cdot (\alpha \Delta T) \), with an equivalent temperature change, \( \Delta T \), and an equivalent thermal expansion coefficient, \( \alpha \), as illustrated in Figure 3.6(b). Table 3.3 summarizes the results of residual stress from dice of N78Q-AH, N79V-AQ, and N78H-AV runs. Residual stress
tends to increase when structures have more dielectric content. Beams containing all four conductors have relatively low compressive stress around 29.2 MPa with a 20% variation.

Curling of the structural material arises from the vertical stress gradient in the composite structures. Measured beam curl in a localized area as shown in Figure 3.7 displays much less variation than measured curl from different dice. Table 3.4 summarizes the measured out-of-plane radius of curvature of various 100 µm long, 2.1 µm wide composite beams from dice of N78Q-AH, N79V-AQ, and N78H-AV runs. Beams using metal-3 as the top layer display less curl than those using metal-2 or metal-1 as the top layer. Equations for predicting curl of multi-morph structures have been derived based on the parameter extraction technique [59].

### 3.3.3 Cyclic Fatigue

Cyclic fatigue of the CMOS microstructures is investigated using a modified fatigue structure [60] shown in Figure 3.8(a). Stress concentration induced by cyclic
motion near the notch at the base of the structure gives rise to the fatigue crack growth, and causes the failure of the structure.

Table 3.4  Radius of curvatures in mm measured from dice of MOSIS N78Q-AH, N79V-AQ, and N78H-AV runs.

<table>
<thead>
<tr>
<th>beam type</th>
<th>m3</th>
<th>m2</th>
<th>m1</th>
<th>poly</th>
</tr>
</thead>
<tbody>
<tr>
<td>N78Q-AH</td>
<td>die no.1</td>
<td>2.5</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>N79V-AQ</td>
<td>die no.1</td>
<td>2.7</td>
<td>1.7</td>
<td></td>
</tr>
<tr>
<td>N78H-AV</td>
<td>die no.1</td>
<td>0.9</td>
<td>1.7</td>
<td></td>
</tr>
</tbody>
</table>

Figure 3.8  (a) SEM of a fan structure for durability test. (b) Measured frequency spectrum has a lateral mode at 13.1 kHz and a out-of-plane mode at 11.2 kHz.

Frequency response of the fan structure shown in Figure 3.8(b) is measured from output of an off-chip transresistive amplifier by the Agilent 4395A spectrum/network analyzer. The lateral mode is found at 13.1 kHz, and the lower peak at 11.2 kHz is an out-of-plane mode excited by the pull-down force generated by the curled rotor fingers and stator
fingers. A durability test is performed with an a.c. driving voltage of 40 V amplitude biased at 55 V d.c., producing a maximum deflection of $10 \pm 1 \mu m$ measured at the free end of the structure. Failure of the notch is observed after 72 minutes for over 45 million experienced cycles. Figure 3.9(a) shows the change in resonance frequency with respect to the experienced cycles. Scanning electron micrographs of the notch taken before and after the test are shown in Figure 3.9(b) and Figure 3.9(c). The maximum stress value near the notch is $620 \pm 60$ MPa by finite element analysis.

### 3.4 Parallel-Plate Microactuator Design

#### 3.4.1 Dynamic Equation

A conventional parallel-plate electrostatic actuator shown schematically in Figure 3.10 is represented by a lumped-parameter model using mass, damper, and mechanical spring elements. Using the free-body diagram in Figure 3.10(b) and the Newton’s second law, we obtain the balanced force equation,
Figure 3.10 (a) Schematic of parallel-plate actuator represented by a lumped-parameter model. (b) Free-body diagram illustrating the forces acting on the plate, including the electrostatic force, $F_e$, the damping force, $F_D = b\ddot{z}$, and the spring restoring force, $F_S = kz$.

\[ m\dddot{z} = F_e - b\dot{z} - kz \quad (3.3) \]

where $m$ is the mass, $b$ is the damping coefficient between plates, $k$ is the spring constant, $F_e$ is the parallel-plate electrostatic force, and $z$ is the displacement of the movable plate at the bottom. By re-arranging (3.3) and replacing $F_e$ with (2.16), the dynamic equation of a parallel-plate actuator is given by,

\[ m\dddot{z} + b\dot{z} + kz = \frac{\varepsilon_o A}{2(g - z)^2}V^2 \quad (3.4) \]

By static analysis ($\ddot{z} = \dot{z} = 0$), (3.4) reduces to balanced spring restoring force $F_s$ and electrostatic force $F_e$. Plate displacement with respect to the applied voltage is the solution of a third-order algebraic equation,

\[ z^3 - 2gz^2 + g^2z - \frac{\varepsilon_o AV^2}{2k} = 0 \quad (3.5) \]
(3.5) suggests that parallel-plate actuators have multiple equilibrium points at a given voltage. To determine the convergence of these equilibrium points, we compare the force gradients of spring-restoring force and electrostatic force given by,

\[
\frac{dF_s}{dz} = k = \frac{\varepsilon_o A}{2z(g-z)^2} V^2 \tag{3.6}
\]

and

\[
\frac{\partial F_e}{\partial z} = \frac{\varepsilon_o A}{(g-z)^3} V^2 \tag{3.7}
\]

at an equilibrium point \(z = Z_o\). \(Z_o\) is a stable equilibrium point when the value of (3.6) is larger than that of (3.7), and is an unstable equilibrium point otherwise. The actuator is marginally stable at displacement of one-third of the gap known as the electrostatic pull-in limit. The pull-in voltage, \(V_{pi}\), corresponding to the pull-in point, is the maximum applied voltage for stable open-loop operation obtained by solving (3.5) with \(z = g/3\),

\[
V_{pi} = \sqrt{\frac{8kg^3}{27\varepsilon_o A}} \tag{3.8}
\]

A typical displacement-voltage curve of a parallel-plate actuator is in Figure 3.11, illustrating two equilibrium points at each voltage except at \(V = V_{pi}\).

### 3.4.2 Design Description

As introduced in Chapter 1, each magnetic probe tip actuator is servoed independently to complete the reading and writing actions. The actuator design for the envisioned system implementation must consider the following factors:
Two anchored mechanical suspensions to provide a total of six interconnects, including write current (around 1 mA), readback voltage, tunneling current for thermally-assisted writing, actuation voltage, capacitive sensing signal, and ground.

A 2 µm minimum suspension width is required for the 1 mA write current (1 mA/µm² in aluminum wires) to prevent electromigration.

The area of each actuator cell is 10,000 µm² with a dimension ratio of 1:1 or 4:1, in order to access 100% of the media with a given media-actuator stroke of ±50 µm.

The top-view schematic of the actuator design for probe-based data storage is illustrated in Figure 3.12(a). Its mechanical design has two anchored serpentine springs, followed by two plates for capacitive position sensing, and an actuation plate where the read/write probe head is to be located. The actuation plate is placed away from the anchors in order to increase the out-of-plane structural compliance for a maximum driving voltage less than 15 V. The ends of springs are close to an imaginary line between two anchors in

![Normalized displacement-voltage characteristic of parallel-plate actuator.](image)

Figure 3.11 Normalized displacement-voltage characteristic of parallel-plate actuator.
Figure 3.12 (a) Schematic of tip actuator design for probe-based data storage. (b) to (f): Side view at cross-section lines, illustrating electrical connections using three metal layers.
order to reduce the structural curl close to the anchor level as the springs fold back, thereby decreasing the eventual curl of actuation plate as shown in the side-view schematic of curl design in Figure 3.13(a). An alternate design with springs ending away from the anchors has more actuation-plate curl and out-of-plane initial displacement as shown in Figure 3.13(b). To ensure final structural release, the plates are separated to provide sufficient silicon undercut. From another perspective, the separation of the actuation and the sensing plates is also required to reduce signal feedthrough.

An external planar electrode will be placed on top of the actuator (not shown in Figure 3.12(a)) to form the actuated and sensing capacitances. A total of three electrical connections are required for parallel-plate servo: z actuation, capacitive sensing, and ground. The current design is for test of position servo only. Wires for read, write, and tunneling current can be considered as a.c. ground. Therefore they are d.c. grounded during the servo test. Actuation and sensing signals are carried by metal-1 through the left and right springs individually to the desired plates in order to reduce signal feedthrough, as shown in cross-section view in Figure 3.12(b), (c), (e), and (f). The metal-3 on top of the springs is d.c. grounded, producing nearly zero electrostatic force with respect to the top
electrode. A metal-3 cut is needed between the springs and the sensing plate, and between the sensing and actuation plates for separation of signals and ground, as shown in Figure 3.12(d), (e), and (f). A d.c.-ground shield provided by metal-2 is inserted between overlapped actuation and sensed wires to reduce signal feedthrough, as shown in Figure 3.12(d) and (e). Similarly a d.c.-ground mechanical bar between actuation and sensing plates, as shown in Figure 3.12(a) and (f), is designed to shunt field lines between plates for feedthrough reduction.

### 3.4.3 Spring Constant

In this section, the energy method (Castigliano’s second theorem) is used to derive an analytic formula for actuator spring constant in the out-of-plane direction. A comprehensive spring-constant analysis of four flexures that are commonly used in micro-mechanical designs is found in [61]. A schematic of the actuator design is re-drawn in Figure 3.14. Each meander in the spring consists of a connector beam of length \( a \) and width \( w_a \), and a span beam of length \( b \) and width \( w_b \). The last beam of the spring measures length \( c \) and width \( w_c \), and each plate is of size \( w \times l \). The distance between two meandering springs is \( d \). \( l_y \) is the distance defined from center of the actuation plate to the end of spring. \( l_0, l_1, \) and \( l_2 \) are the distances measured from the top of the spring to the center of the actuation plate, sensing plate 1, and sensing plate 2. The beams between plates are of length \( l_d \) and width \( w_d \). \( h \) is the beam thickness pointed into the \( z \) direction. The goal is to find the displacement resulting from a lumped electrostatic force applied at the center of the actuation plate. The spring constant is defined as the ratio of force over displacement.
By geometric symmetry of the actuator along the $y$ axis, we only need to analyze one half of the actuator with $n$ meanders using the free-body diagram in Figure 3.15. The connector beams are indexed from $i = 1$ to $n$ and the span beams are indexed from $j = 1$ to $n-1$. A $z$-directional force, $F_z$, and a moment around the $y$ axis, $T_y$, are applied to the center of the actuation plate. The rotational angle, $\phi_o$, at the center of the plate, is constrained by the guided-end boundary condition.

Because the plate is much wider than the beam, it is reasonable to assume the plate as a rigid body without structural deformation or rotation, thereby the stored strain energy is zero. The moment and torsion of each beam segment is deduced from the free-body diagram as the following,
Figure 3.15 Free-body diagram of one half of actuator used for deriving z-directional spring constant.
\[ M_c = F_z(l_y + x) \]  \hfill (3.9)

\[ T_c = T_o - F_z \frac{d}{2} \]  \hfill (3.10)

\[ M_{a,i} = T_o - F_z \left[ \frac{d}{2} + x + (i - 1)a \right] \quad i = 1 \text{ to } n \]  \hfill (3.11)

\[ T_{a,i} = F_z \left[ -l_y - c + \left( \frac{1 + (i - 1)k}{2} \right) b \right] \quad i = 1 \text{ to } n \]  \hfill (3.12)

\[ M_{b,j} = F_z(-L_y - c + x) \quad j = 1 \text{ to } n - 1 \]  \hfill (3.13)

\[ T_{b,j} = T_o - F_z \left( \frac{d}{2} + ja \right) \quad j = 1 \text{ to } n - 1 \]  \hfill (3.14)

where \( M_c \) and \( T_c \) are the moment and torsion of the last spring beam, \( M_{a,i} \) and \( T_{a,i} \) are the moment and torsion of the \( i \)-th connector beam, and \( M_{b,j} \) and \( T_{b,j} \) are the moment and torsion of the \( j \)-th span beam.

The total strain energy, defined as \( U \), of the actuator, is given by

\[
U = \sum_{i=1}^{n} \int_{0}^{a} \left( \frac{M_{a,i}^2}{2EI_{xa}} + \frac{T_{a,i}^2}{2GJ_{a}} \right) dx + \sum_{j=1}^{n-1} \int_{0}^{b} \left( \frac{M_{b,j}^2}{2EI_{xb}} + \frac{T_{b,j}^2}{2GJ_{b}} \right) dx + \int_{0}^{c} \left( \frac{M_c^2}{2EI_{xc}} + \frac{T_c^2}{2GJ_{c}} \right) dx
\]

\hfill (3.15)

where \( I_x \) is the moment of inertia around the \( x \) axis, \( J_a, J_b, \) and \( J_c \) are the polar moment of inertia, \( E \) is the Young’s modulus, and \( G \) is the torsion modulus of elasticity. The torsional modulus is related to Young’s modulus and Poisson’s ratio, \( \nu \), by

\[ G = \frac{E}{2(1 + \nu)} \]  \hfill (3.16)

The polar moment of inertia for a beam of rectangular cross-section \( h \times w \) is given by
\[ J = \frac{1}{3} h^3 w \left[ 1 - \frac{192 h}{\pi^5 w} \sum_{n = 1, \text{odd}}^{\infty} \frac{1}{n^3} \tanh \left( \frac{n \pi w}{2h} \right) \right] \tag{3.17} \]

where \( h < w \).

The actuator displacement \( z \) is found by solving the two simultaneous equations:

\[ \phi_o = \frac{\partial U}{\partial T_o} = 0 \tag{3.18} \]

\[ z = \frac{\partial U}{\partial F_z} \tag{3.19} \]

The \( z \)-directed actuator spring constant is \( k_z = 2F_z/z \). We assume that \( I_{xb} = I_{xc} \) and \( J_b = J_c \).

For \( n \) is always even in our design, the spring constant is given by,

\[ k_z = \frac{24S_a S_b G_a G_b (cS_a + b(n - 1)S_a + naG_b)}{4S_a^2 G_a G_b (c + b(n - 1)) (c^2 + 3cl_y + 3l_y^2) + b^3 (n - 1) + 3b^2 l_y (n - 1) + 3b^2 l_y^2 - 2aS_a G_b [-2c (c^2 G_a G_b + 3cG_a G_bl_y + 3(S_a S_b + G_a G_b)l_y^2) - 6bl_y (cS_a S_b + (n - 1)(S_a S_b + G_a G_b)l_y) - 3b^2 (n - 1)(cS_a S_b + 2(S_a S_b + G_a G_b)l_y)] - b^3 (3S_a S_b + 2G_a G_b) (n - 1) + n + a^4 S_b G_a G_b^2 n + 2a^3 S_a S_b G_a G_b n^2 (b(n - 1)^2 + 2cn) + a^2 S_a S_b n [12G_b^2 l_y^2 n + b^2 (S_a G_a (n - 2)(n - 1)^2 + 6G_b^2 n) + 2b (6G_b^2 l_y n + cS_a G_a (1 - 3n + 2n^2))]} \]

(3.20)

where \( S_a \equiv EI_{xa}, S_b \equiv EI_{xb}, G_a \equiv GJ_a \), and \( G_b \equiv GJ_b \).

(3.20) is further reduced assuming that all the beam cross-sections are the same \( (I_{xa} = I_{xb} = I_{xc} \text{ and } J_a = J_b = J_c) \), giving,
Analytic spring constants and finite-element simulations are compared using the following design parameters: $E = 70 \text{ GPa}$, $a = 4 \mu m$, $b = 108 \mu m$, $c = 113 \mu m$, $d = 15.5 \mu m$, $w_a = w_b = w_c = 2 \mu m$, $w_d = 5 \mu m$, $l_d = 10 \mu m$, $w = 20 \mu m$, $l = 80 \mu m$, $l_y = 70 \mu m$, and $h = 4.5 \mu m$. The number of connectors $n$ in each spring is even-numbered from 2 to 8 ($n = 6$ for the final design). The finite-element analysis and analytic calculations match to better than 0.33% for all cases as shown in Figure 3.16. The designed actuator has a spring constant $k_z = 0.167 \text{ N/m}$.

\[
k_z = \frac{24S_aG_a(bS_a + aG_a)}{4b^2S_aG_a(b^2 + 3bl_y + 3l_y^2) + 2ab[b^2(3S_a^2 + 2G_a^2) + 6b(S_a^2 + G_a^2)l_y^2 + n \left(6(S_a^2 + G_a^2)l_y^2 + a^4G_a^2n^2 + 2a^3bS_aG_a(1 + n^2) + a^2S_a[12bG_al_y + 12G_al_y^2 + b^2(6G_a + S_a(n^2 - 1))] \right) + 12G_al_y^2 + b^2(6G_a + S_a(n^2 - 1))} \right] \right)
\]

Figure 3.16 Comparison of analytic z-directional spring constants with finite-element analysis.
3.4.4 Modal Analysis

In this section, an analytical finite-element method [62][63] is utilized for the \( z \)-directional modal analysis of the continuous elastic microactuator. The result is expressed as an implicit solution of the formulated eigenvalue problem which can be conveniently solved by a computer program (MATLAB\textsuperscript{\textregistered} [64]), rather than an explicit analytical formula. Even so, the advantage of this method is that it takes much less computing time than the numerical finite-element method when many design iterations are needed, and gives satisfactory results comparable to the numerical finite-element method.

In the actuator design, each plate occupies a relatively small portion of the total area, resulting in high-order mechanical modes from the plate itself at very high frequencies. These modes can be neglected for elementary modal analysis. It is reasonable to represent plates with beam elements in analysis without resorting to a more complicated plate element, while having sufficient accuracy. By using interconnected beam elements with multiple degrees of freedom, the equation of motion of the actuator is formulated as,

\[
[m] \ddot{q} + [k]q = \{F\} \tag{3.22}
\]

where \([m]\) is the mass matrix, \([k]\) is the stiffness matrix, \(\{q\}\) is the vector of nodal displacements and rotations, \(\{\ddot{q}\}\) is the vector of accelerations, and \(\{F\}\) is the vector of external forces and moments. For the case of free vibration with natural frequency \(\omega\), (3.22) takes the following form:

\[
\{\theta\} = ([k] - \omega^2[m])q \tag{3.23}
\]

which is suited for eigenvalue solutions.
The beam element shown schematically in Figure 3.17 is used for the finite-element analysis, and is capable of resisting axial forces \((F_{xa} \text{ and } F_{xb})\), shearing forces \((F_{ya}, F_{yb}, F_{za}, \text{ and } F_{zb})\), bending moments \((M_{ya}, M_{za}, M_{yb}, \text{ and } M_{zb})\), and twisting torques \((M_{xa} \text{ and } M_{xb})\), allowing six degree-of-freedom motion \((x, y, \text{ and } z \text{ in translation. } \theta, \phi, \text{ and } \varphi \text{ in rotation}) at nodes \(a\) and \(b\). The stiffness and mass matrices of the beam element are obtained by substituting the beam displacement functions into the kinetic energy expression and the strain energy expression, and performing the differentiations in the Lagrange’s equation with respect to each degree-of-freedom [62]. For derivation of the \(z\)-directional mode, the vector \(\{q\}\) can be reduced to \(\{q\} = \{z_a, \theta_a, \phi_a, z_b, \theta_b, \phi_b\}\), leading to the stiffness and mass matrices given by,
where $\rho$ is the mass density, $l$ is the beam length, $A$ is the beam cross-section, and $J_p$ is the polar moment of inertia.

The force and moments acting on local coordinate axes of each beam element are transformed to the global coordinate axes used by the actuator as shown in Figure 3.18. Bars are added to all the symbols with reference to the local coordinate, which is oriented
at an angle $\Phi$ measured counterclockwise from the global $x$ axis. The resulting stiffness and mass matrices transformed from the local matrices are given by,

$$[k] = [T]^T[\bar{k}][T]$$  \hspace{1cm} (3.26)

$$[m] = [T]^T[\bar{m}][T]$$  \hspace{1cm} (3.27)

where $\{T\}$ is the transformation matrix given by,
For the \( z \)-directional modal analysis, only half of the actuator is considered because of its symmetry with respect to the \( y \) axis as shown in Figure 3.19. Starting from the connector at the left, a total of \((2n + 5)\) beam elements and \((2n + 6)\) nodes are numbered in a continuous sequence, with \( n \) as the total number of connectors in one half of the actuator. By tracing along the beam elements in ascending node number, the orientation of each beam element is found at 0° for all connectors, 90° for span beams pointed toward en.
the positive y direction, and -90° otherwise. The angle is -90° for the last spring beam, and the rest of beam elements used for representing plates. For the \( i \)-th beam elements \((i = 1\) to \(2n + 5)\), the displacement vector is \( \{q_i\} = [z_i, \theta_i, \phi_i, z_{i+1}, \theta_{i+1}, \phi_{i+1}]^T \), and the \( 6 \times 6 \) stiffness and mass matrices \([k_i], [m_i]\) are obtained by substituting the angle of beam elements into (3.26) through (3.28). The complete actuator stiffness and mass matrices \( \{K\}, \{M\} \) are assembled into \( 6(n + 3) \times 6(n + 3) \) matrices with \([k_i]\) and \([m_i]\) appearing in the diagonal. Entries of \([k_i]\), \([k_{i+1}]\), and \([m_i]\), \([m_{i+1}]\) are summed up at shared nodes of \( i \)-th and \((i + 1)\)-th beam elements, as indicated by the shaded areas (\( 3 \times 3 \) matrices) in (3.29) and (3.30).
The z-directional mode is constrained at the first node (anchor) with 
\( z_1 = \theta_1 = \phi_1 = 0 \), and the \((2n + 6)\)-th node with \( \phi_{2n+6} = 0 \) (no plate rotation). Rows and columns of \( \{K\} \) and \( \{M\} \) corresponding those zero displacements and rotations are removed, reducing the matrix size from \( 6(n + 3) \times 6(n + 3) \) to \( 6(n - 1) \times 6(n - 1) \). By substituting the new stiffness and mass matrices into (3.23), the z-directional mode is solved as the minimum of all eigenvalue solutions.

In Figure 3.20, eigenvalue solutions with results from numerical finite-element analysis are compared using the design parameters described in Section 3.4.3, plus the effective mass density, \( \rho = 2650 \text{ kg/m}^3 \). The finite-element analysis and calculated eigenmodes match to better than 2\% for all cases. The designed actuator has a z-directional mode at 10.5 kHz. Using the spring constant \( k_z = 0.167 \text{ N/m} \), we obtain an effective lumped mass of \( 3.84 \times 10^{-11} \text{ kg} \). The first three modes of the actuator obtained by numerical finite-element analysis are shown in Figure 3.21. The actuator is designed such that the twisting mode is about three times of the z mode to prevent from being excited during closed-loop servo.
3.4.5 Squeeze-Film Damping

Viscous air damping is the dominant energy dissipation mechanism for microstructures operated in atmosphere. In the presence of a compressible gas (air) film between two parallel plates, as in the schematic in Figure 3.22, plate motion in the vertical direction squeezes the thin film of gas, and produces an opposing force due to the pressure in the film, known as the squeeze-film damping force. The characteristic of squeeze-film damping is frequency-dependent: at low operating frequencies, it behaves like a damper element as the gas film has enough time to flow out of the gap, and causes dissipation. At high frequencies, it behaves like a spring element as the gas film is trapped and squeezed between moving plates with low dissipation. Thereby an electrical equivalent circuit network consisting of nonlinear frequency-dependent components [67] is used for squeeze-film damping analysis.
Figure 3.21 Illustration of the first three modes of the actuator by numerical finite-element analysis.

Mode 1: 10.3 kHz, z

Mode 2: 11.7 kHz, x

Mode 3: 32.6 kHz, φ
Assume that the gas film undergoes an isothermal process during the entire motion, the distribution of the gas pressure between plates is governed by the compressible gas-film Reynolds equation [66]. In the analysis given in [65], the plate motion is assumed to be small, and perpendicular to the surface of two parallel plates, producing the linearized Reynolds equation:

\[
\frac{P_a g^2}{12 \eta_{\text{eff}}} \nabla^2 \left( \frac{p}{P_a} \right) - \frac{\partial}{\partial t} \left( \frac{p}{P_a} \right) = \frac{\partial \left( \frac{z}{g} \right)}{\partial t}
\]

(3.31)

where \( p \) is the pressure change of the static pressure, \( P_a \), \( g \) is the static gap spacing, \( z \) is the plate displacement, and \( \nabla^2 \) is the Laplacian operator. \( \eta_{\text{eff}} \) is the effective gas viscosity given by [67],

\[
\eta_{\text{eff}} = \frac{\eta}{1 + 9.638 K_n^{1.159}}
\]

(3.32)

where \( \eta \) is the viscosity coefficient of air. \( K_n = \frac{\lambda}{g} \) is the Knudsen number in which \( \lambda \) is defined as the mean free path of air. (3.32) is suitable for use in both continuum fluid mechanics when the air gap is much larger than the mean free path of air, and in molecular air-damping regime when the gap is narrow or/and the pressure is low. At air pressure of Figure 3.22 Cross-section schematic illustrating a squeezed thin film of gas between two parallel plates separated by a gap, \( g \), resulted from the vertical motion of the plate with velocity, \( v \). The squeeze-film damping force, \( F_D \), is produced opposing the direction of motion.
1 atm and room temperature of 25 °C, the mean free path is $\lambda = 70$ nm and the viscosity coefficient is $\eta = 22.6$ $\mu$N⋅s/m².

Squeeze-film damping force is simulated by a lumped parameter model consisting of parallel branches of series-connected nonlinear inductors and resistors [67], as shown schematically in Figure 3.23. By using the analogy between mechanical and electrical systems, the inductors and resistors are used to represent the spring behavior and the damping behavior of the gas film. The damping force and plate velocity are analogous to the through variable (electrical current) and the across variable (electrical voltage). The components $L_{mn}$ and $R_{mn}$ are given by,

$$L_{mn} = (mn)^2 \frac{\pi^4 (g-z)}{64lwP_a} \quad (3.33)$$

$$R_{mn} = (mn)^2 (m^2 l^2 + n^2 w^2) \frac{\pi^6 (g-z)^3}{768lw^3 \eta_{eff}} \quad (3.34)$$

Figure 3.23 Equivalent circuit model using nonlinear inductors and resistors for squeeze-film damping analysis. Damping force and plate velocity are analogous to electrical current and voltage in the schematic.
where \((m, n)\) is pair of odd integers, and \(l, w\) are dimensions of a rectangular plate. For plate size of tens of microns, non-zero pressure change at the edge is considered using [68],

\[
l_{\text{eff}} = l + g(0.8792 + 10000 \cdot l)
\]

and

\[
w_{\text{eff}} = w + g(0.8792 + 10000 \cdot w)
\]

to replace \(l\) and \(w\) in (3.33) and (3.34).

The damping coefficient, \(b\), with respect to the gap spacing, \(g\), is characterized by performing frequency-domain analysis of the parallel-plate actuator using NODAS simulation. The lumped-parameter actuator model shown in Figure 3.24 uses three squeeze-film damping elements to represent three parallel plates of different displacements and velocities. The displacements, \(z_{s1}\) and \(z_{s2}\), of the sensing plates 1 and 2, given by,

\[
z_{s1} = (l_1/l_0)z
\]

\[
z_{s2} = (l_2/l_0)z
\]

are less than the actuation-plate displacement, \(z\), due to the cantilever rotation. \(l_0, l_1,\) and \(l_2\) are 183 \(\mu\)m, 153 \(\mu\)m, and 123 \(\mu\)m in design, respectively. \(k\) and \(m\) are 0.167 N/m and 3.84\(\times\)10\(^{-11}\) kg, as calculated in Section 3.4.3 and 3.4.4. Parameters used in the squeeze-film damping model are: plate size = 20 \(\times\) 80 \(\mu\)m\(^2\), \(P_a = 10^5\) Pa, \(\eta = 22.5\) \(\mu\)N\(\cdot\)s/m\(^2\), and \(K_n = 7.052\cdot10^{-8}/g\) (\(g\) is in meters). The damping ratio, \(\xi = b/(2\sqrt{mk})\), is determined by matching the frequency response of the lumped-parameter model with that of a second-order transfer function, \(1/(ms^2 + bs + k)\). The damping ratio \(\xi\) is ranging from 240 to 0.2
Figure 3.24 Schematic of lumped-parameter actuator model driven by a sinusoidal force source in frequency-domain analysis. Squeeze-film damping of three plates in design is modelled separately due to their different displacements and velocities.
for gaps between 0.2 µm to 4 µm as shown in Figure 3.25. The designed actuator is over-
damped (ξ > 1) for gaps within 2.1 µm, and becomes underdamped (0 < ξ < 1) beyond
\( g = 2.1 \mu m \).

### 3.5 Capacitive Position Sensing

Position of the actuator is detected by the change of the parallel-plate sensing
 capacitances, \( C_{s1} \) and \( C_{s2} \), formed between the top electrode and the sensing plates 1 and 2, as shown in the actuator side-view schematic in Figure 3.26(a). The equivalent circuit schematic in Figure 3.26(b) illustrates that capacitance change is converted to a modulated sensed voltage using a single-ended capacitive bridge, followed by amplification, demod-
ulation, and low-pass filtering of the 1× and 2× carrier-frequency signals. Initial gap spacing, \( g \), is defined from center of the actuation plate to the top electrode, \( C_s \) is the sum of
\( C_{s1} \) and \( C_{s2} \), \( A_v \) is the voltage gain of the sensing pre-amp, \( V_m \) is the modulation voltage,
$V_c$ is the carrier voltage, $A_d$ is the voltage conversion gain of the demodulator, and $\omega_p$ is the corner (-3 dB) frequency of the low-pass filter with a d.c. gain of 1. The capacitance, $C_i$, includes the routing capacitance to the ground and the silicon substrate (an a.c. ground) throughout the meandering spring, and the pre-amp gate capacitance. Routing capacitance to the substrate can be neglected because the silicon is at least 30 $\mu$m below the actuator after the post CMOS-MEMS process. After low-pass filtering, a voltage subtracter is used to remove the d.c. component from the initial sensed signal.

Figure 3.26 (a) Side-view schematic of parallel-plate actuator illustrating capacitive sensing scheme. (b) Equivalent capacitive sensing circuit model.
The magnitude of the sensed voltage output after the low-pass filter is a function of actuation-plate displacement, \( z \), given by,

\[
[V_s] = A\,A_d\,|V_m|\left(\frac{C_{s1}(z) + C_{s2}(z)}{C_{s1}(z) + C_{s2}(z) + C_i} - \frac{C_{s10} + C_{s20}}{C_{s10} + C_{s20} + C_i}\right)
\]  

(3.39)

where \( C_{s10} \) and \( C_{s20} \) are the initial sensing capacitances, \( \varepsilon_oA/g \). The sensing capacitances, \( C_{s1}(z) \) and \( C_{s2}(z) \), are functions of the sensing-plate displacements in (3.37) and (3.38), given by,

\[
C_{s1}(z) = \frac{\varepsilon_oA}{g - (l_1/l_0)z}
\]  

(3.40)

\[
C_{s2}(z) = \frac{\varepsilon_oA}{g - (l_2/l_0)z}
\]  

(3.41)

The capacitive sensor gain at an operating point \( z = Z_o \) is defined as the differentiation of the sensed output with respect to the actuation-plate displacement. By including the dominant pole \( \omega_p \) from the low-pass filter, the capacitive sensor transfer function at \( z = Z_o \) is,

\[
H(s)|_{z = Z_o} = \frac{G_s(Z_o)}{1 + s/\omega_p}
\]  

(3.42)

where

\[
G_s(Z_o) = \frac{A\,A_d\,|V_m|\,C_i\varepsilon_oA[(q_1 + q_2)g^2 - 4q_1q_2Z_og + q_1q_2(q_1 + q_2)Z_o^2]}{[C_i(g - q_1z)(g - q_2z) + \varepsilon_oA(2g - (q_1 + q_2)Z_o)]^2}
\]  

(3.43)

Parameters used in (3.43) are: \( q_1 = l_1/l_0 = 0.8361, \, q_2 = l_2/l_0 = 0.6721, \, A_v = 10, \, A_d = 4.7, \, V_m = 1 \text{ V}, \, A = 1600 \mu m^2, \) and \( \omega_p = 120 \text{ kHz} \). The extracted capacitance \( C_i \) from layout is
408 fF, including the routing capacitance of 314 fF in the actuator, and the circuit input capacitance of 94 fF.

3.5.1 Pre-amp Circuit Design

The sensing pre-amp configuration in Figure 3.27(a) is a non-inverting amplifier implemented using a high-gain operational amplifier. The d.c. bias at the front end of the sensing pre-amp is provided by a bias transistor operated in the sub-threshold region. The operational amplifier shown in Figure 3.27(b) uses a differential input pair in the first stage, followed by a folded-cascode gain stage and a common-source amplifier as the output stage.

Two high-impedance nodes in the circuit are at the output of the folded-cascode stage (node 1) and the output stage (node 2). The dominant pole occurs at node 1 because of the high impedance provided by the cascode transistors, M6 and M7, and the wide-swing current mirror formed by the transistors M8, M9, M10, and M11. By using a small-signal analysis, the voltage gain at d.c. is given by,

\[
A_{vo} = \frac{v_o}{v_i} = A_1 A_2 = (g_{m1}R_{o1})(g_{m13}R_{o2})
\]  

(3.44)

where \(A_1\) and \(A_2\) are the gains of the first and second stage, \(g_{m1}\) and \(g_{m13}\) are the transconductances of the transistors M1 and M13, and \(R_{o1}\) and \(R_{o2}\) are the output impedances at node 1 and node 2 given by,

\[
R_{o1} = [g_{m7}r_{ds7}(r_{ds1} \parallel r_{ds4})] \parallel (g_{m9}r_{ds9}r_{ds11})
\]  

(3.45)

\[
R_{o2} = R_L \parallel r_{ds12} \parallel r_{ds13}
\]  

(3.46)
Figure 3.27 (a) Schematic of sensing pre-amp using an operational amplifier in a non-inverting amplifier configuration. The bias transistor operates in the sub-threshold region to act as a large resistance. (b) Schematic of folded-cascode operational amplifier with a common-source amplifier as the output stage.
where $r_{ds}$ is the output impedance across the drain and source of transistors, and $R_L$ is the load resistance. The transistor transconductance $g_m$ of a MOSFET transistor is given by,

$$g_m = \sqrt{\frac{2\mu C_{ox} W}{L} I_{DS}}$$  \hspace{1cm} (3.47)

where $\mu$ is the carrier mobility, $C_{ox}$ is the gate capacitance per unit area, and $I_{DS}$ is the drain-to-source current.

To provide sufficient phase margin for the circuit, the dominant pole at node 1 is moved to a lower frequency away from the second pole by pole-splitting compensation using the Miller capacitance, $C_c$. The resistance, $R_c$, combined with $C_c$, provides a phase-lead compensation. The resulting dominant pole after the compensation is given by,

$$\omega_{p1} \equiv \frac{1}{R_{o1}(A_2 C_c)} = \frac{1}{R_{o1}(g_{m13} R_{o2} C_c)}$$  \hspace{1cm} (3.48)

Multiplication of (3.44) and (3.48) yields the gain-bandwidth product,

$$GBW = \frac{g_{m1} C_c}{C_c}$$  \hspace{1cm} (3.49)

The circuit is fabricated in the AMS (Austria Micro System) 0.5-µm three-metal-two-poly CMOS process. The carrier mobility, threshold voltage, channel length modulation parameter, and gate capacitance per unit area for n-channel devices are:

$$\mu_n = 427 \text{ cm}^2/\text{V} \cdot \text{s}$$

$$V_{m} = 0.85 \text{ V}$$

$$\lambda_n = 0.131 \text{ V}^{-1}$$

$$C_{ox} = 0.0027 \text{ F/m}^2$$
and for p-channel devices:

\[
\mu_p = 112 \text{ cm}^2/\text{V} \cdot \text{s}
\]

\[
V_{tp} = -0.78 \text{ V}
\]

\[
\lambda_p = 0.036 \text{ V}^{-1}
\]

\[
C_{ox} = 0.0027 \text{ F/m}^2
\]

The transistor dimensions in design are tabulated in Table 3.5. The compensation resistor, \(R_c\), and capacitor, \(C_c\), are 720 \(\Omega\) and 3.54 pF. Each transistor is biased at 400 \(\mu A\) except for the transistor M3 which is biased at 800 \(\mu A\). Applying the bias current and the device parameters through (3.44) to (3.49) yields an open-loop gain of 99.3 dB, and a gain-bandwidth product of 64 MHz.

The bias voltages, \(V_{b1}, V_{b2}, V_{b3},\) and \(V_{b4}\), for the folded-cascode amplifier are provided by a constant-transconductance bias circuit [69] shown in Figure 3.28. The drain-source current, \(I_{DS3}\), can be determined by writing a Kirchhoff loop equation for the transistors M3 and M4,

\[
R_b I_{DS3} = V_{GS4} - V_{GS3}
\]

(3.50)

where the gate-to-source voltage, \(V_{GS}\), is given by,

<table>
<thead>
<tr>
<th>Transistor, W/L</th>
<th>Transistor, W/L</th>
<th>Transistor, W/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1, 20 (\mu m)/0.9 (\mu m)</td>
<td>M6, 450 (\mu m)/1.8 (\mu m)</td>
<td>M11, 150 (\mu m)/1.8 (\mu m)</td>
</tr>
<tr>
<td>M2, 20 (\mu m)/0.9 (\mu m)</td>
<td>M7, 450 (\mu m)/1.8 (\mu m)</td>
<td>M12, 1730 (\mu m)/1.8 (\mu m)</td>
</tr>
<tr>
<td>M3, 140 (\mu m)/1.8 (\mu m)</td>
<td>M8, 180 (\mu m)/1.8 (\mu m)</td>
<td>M13, 600 (\mu m)/1.8 (\mu m)</td>
</tr>
<tr>
<td>M4, 1040 (\mu m)/1.8 (\mu m)</td>
<td>M9, 180 (\mu m)/1.8 (\mu m)</td>
<td></td>
</tr>
<tr>
<td>M5, 1040 (\mu m)/1.8 (\mu m)</td>
<td>M10, 150 (\mu m)/1.8 (\mu m)</td>
<td></td>
</tr>
</tbody>
</table>

Table 3.5 Transistor dimensions in the folded-cascode amplifier design.
Substituting (3.51) into (3.50) gives,

\[ V_{GS} = V_{in} + \sqrt{\frac{2I_{DS}}{\mu n C_{ox}(W/L)}} \]  

(3.51)

Replacing \( I_{DS3} \) in (3.52) into (3.47), we obtain the transconductance of the transistor M3,

\[
I_{DS3} = \frac{2}{\mu n C_{ox} R_b^2} \left( \frac{1}{\sqrt{(W/L)_4}} - \frac{1}{\sqrt{(W/L)_3}} \right)^2
\]  

(3.52)

Replacing \( I_{DS3} \) in (3.52) into (3.47), we obtain the transconductance of the transistor M3,

\[
g_{m3} = \frac{2}{R_b} \left( \frac{W/L}{\sqrt{(W/L)_4}} - \frac{W/L}{\sqrt{(W/L)_3}} \right)
\]  

(3.53)

![Figure 3.28 Schematic of the constant-transconductance bias circuit for the folded-cascode amplifier.](image)
which implies that the transconductance of any n-channel transistor biased using this current reference is thereby independent of the mobility of the transistor, which varies with respect to the temperature [70].

The designed transistor dimensions are tabulated in Table 3.6. By applying the parameters of n-channel devices given in Section 3.5.1 and the transistor sizes in Table 3.6 to (3.53) for a bias current of 400 µA, the biasing resistor, $R_b$, is found at 306 Ω,

<table>
<thead>
<tr>
<th>Transistor, W/L</th>
<th>Transistor, W/L</th>
<th>Transistor, W/L</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1, 120 µm/1.8 µm</td>
<td>M7, 520 µm/1.8 µm</td>
<td>M13, 120 µm/1.8 µm</td>
</tr>
<tr>
<td>M2, 120 µm/1.8 µm</td>
<td>M8, 520 µm/1.8 µm</td>
<td>M14, 60 µm/1.8 µm</td>
</tr>
<tr>
<td>M3, 180 µm/0.9 µm</td>
<td>M9, 120 µm/1.8 µm</td>
<td>M15, 10 µm/0.6 µm</td>
</tr>
<tr>
<td>M4, 70 µm/0.9 µm</td>
<td>M10, 450 µm/1.8 µm</td>
<td>M16, 10 µm/0.6 µm</td>
</tr>
<tr>
<td>M5, 450 µm/1.8 µm</td>
<td>M11, 520 µm/1.8 µm</td>
<td>M17, 10 µm/0.6 µm</td>
</tr>
<tr>
<td>M6, 450 µm/1.8 µm</td>
<td>M12, 70 µm/1.8 µm</td>
<td>M18, 1.2 µm/10 µm</td>
</tr>
</tbody>
</table>

Table 3.6 Transistor dimensions in the constant-transconductance bias circuit.

Frequency response of the pre-amp, shown in Figure 3.29, is measured on a separately bonded test circuit in a dual in-line package (DIP) by the Agilent 4395A spectrum/network analyzer. The measuring active probe contacts directly with the output pin to eliminate the use of a coaxial cable, which induces additional capacitive loading to reduce the measured pre-amp bandwidth. Measured pre-amp gain is 9.85, which is 15% less than the projected gain of 10. The reason can be the imperfect matching of the on-chip polysilicon resistors which are used for setting the closed-loop gain. Measured bandwidth of pre-amp is 5.7 MHz, which is less than the simulated 8.8 MHz because of capacitive loading during the measurement.
For capacitive sensing, the d.c. path from the pre-amp input to the a.c. ground is provided by a large resistor realized by a n-channel transistor operated in the sub-threshold region. A pole frequency is formed by the resistor and the capacitive voltage divider in front of the pre-amp, and its magnitude is primarily determined by the pre-amp input capacitance and the resistor value. As the gate-to-source voltage of the d.c.-biasing transistor decreases, the resistance increases to result in the shifting of the pole frequency to a lower frequency. By applying a sweeping sinusoidal voltage to the top electrode which is placed about 3 μm away from the microactuator, frequency response is measured at pre-amp output using the network analyzer. The biasing resistance is extracted from the measured pole frequency of frequency responses, and plotted as a function of gate-to-source voltage in Figure 3.30.

Figure 3.29 Frequency response of the sensing pre-amp. Measured d.c. gain is 9.85, and the -3 dB frequency is 5.7 MHz.
3.5.2 Double-Balanced Demodulator

Demodulation of the amplitude-modulated sensed signal after the pre-amp is realized off the microactuator chip by a BJT (Bipolar Junction Transistor) double-balanced demodulator, the so-called Gilbert cell [71] shown schematically in Figure 3.31. The advantage of using a double-balanced demodulator over a single-balanced demodulator is that the differential pairs of Q3-Q4, and Q5-Q6 add the amplified carrier-signal feedthrough with opposite phases, thereby providing a first-order cancellation which leads to a less demanding low-pass filtering following the demodulation. To increase the linear input range for operation, emitter degeneration is utilized by adding the resistor, $R_E$, in the lower emitter-coupled pair, Q1 and Q2. By applying a sufficiently large carrier signal to the cross-coupled transistors, Q3-Q6 and Q4-Q5, each pair conducts current alternatively to generate output across the load resistance, $R_L$.

![Figure 3.30](image)  
Figure 3.30 Measured resistance of subthreshold biasing transistor versus gate-to-source voltage.
The collector currents, $I_{c1}$ and $I_{c2}$, of the transistors, Q1 and Q2, can be determined by numerically solving the following equations from a large-signal analysis,

$$\frac{I_{c1}}{I_{c2}} = \exp \left[ \frac{V_i - (R_E/\alpha)(I_{c1} - I_{c2})}{V_T} \right]$$  \hspace{1cm} (3.54)

$$I_{c1} + I_{c2} = \alpha I_E$$  \hspace{1cm} (3.55)

where $V_i$ is the amplitude of the input signal voltage, $\alpha$ is the base transport factor (ideally $\alpha \rightarrow 1$), and $V_T$ is the thermal voltage ($V_T = 26$ mV at 300°K). The voltage conversion gain, $A_d$, is the product of the difference of $I_{c1}$ and $I_{c2}$, and the load resistance, $R_L$, divided by the input signal, $V_{in}$, and a factor of 2 due to the multiplication of $V_c(t)$ by a sinusoidal carrier signal $\sin(\omega_c t)$:

![Gilbert multiplier with emitter degeneration applied to improve input voltage range.](image-url)
By applying the parameters: $\alpha = 0.99$, $I_E = 10.6 \, mA$, $R_E = 75 \, \Omega$, and $R_L = 750 \, \Omega$ from (3.54) to (3.56), the voltage conversion gain is calculated from 4.62 to 4.7 for an input voltage range between 400 to 800 mV. Measured voltage conversion gain is between 4.9 to 5, which is about 6% higher than calculations.

### 3.6 Noise Analysis and Minimum Detectable Signal

#### 3.6.1 Pre-amp Noise

The primary noise sources in the folded-cascode opamp shown in Figure 3.27(b) are the transistor pairs M1-M2 and M4-M5 in the differential-input circuit, and M10-M11 in the cascode stage. The transistors M1-M2, M4-M5, and M10-M11 are assumed to be matched in the following derivations. By adding a voltage noise source at the input of these transistors, the contributed output noise voltage from the folded-cascode operational amplifier is given by,

$$ V_{n_o}^2 = 2(g_{m1} R_{o1} g_{m13} R_{o2})^2 V_{n1}^2 + 2(g_{m4} R_{o1} g_{m13} R_{o2})^2 V_{n4}^2 + 2(g_{m11} R_{o1} g_{m13} R_{o2})^2 V_{n11}^2 $$

(3.57)

where $V_{n1}$, $V_{n4}$, and $V_{n11}$ are the input noise voltage source of the transistors M1, M4, and M11. The output noise voltage is related back to an equivalent input noise voltage, $V_{n_eq}$, by dividing it by the gain squared given in (3.44):

$$ V_{n_{eq}}^2 = 2V_{n1}^2 + 2\left(\frac{g_{m4}}{g_{m1}}\right)^2 V_{n4}^2 + 2\left(\frac{g_{m11}}{g_{m1}}\right)^2 V_{n11}^2 $$

(3.58)
The input-referred voltage noise source for a MOSFET transistor are represented primarily as the sum of the thermal noise (white noise) and the flicker (1/f) noise at the low frequencies. By using a sufficiently high modulation frequency (> 1 MHz) [72] for capacitive sensing, the 1/f noise is assumed to be out of the signal frequency range and can be neglected, leading to,

\[
V_{ni}^2 = 4k_B T \left( \frac{2}{3} \right) \left( \frac{1}{g_{mi}} \right)
\]  

(3.59)

where \(i\) is the transistor index, \(k_B\) is the Boltzmann’s constant \((1.38 \times 10^{-23} \text{ J/K})\), and \(T\) is the temperature in Kelvins \((T = 300 \text{ K at room temperature})\). Applying the device parameters given in Section 3.5.1 to (3.57), (3.58), and (3.59), we obtain the equivalent input-referred noise voltage \(V_{neq} = 9.23 \text{nV/\sqrt{Hz}}\).

To calculate the equivalent input-referred noise voltage of the entire pre-amp, an equivalent noise model shown in Figure 3.32 is used with thermal noise of the resistors \(R_1\) and \(R_2\) being represented by current noise sources, \(I_{n1}\) and \(I_{n2}\), and thermal noise of the

![Figure 3.32 Equivalent pre-amp noise model.](image_url)
d.c.-bias transistor is represented by a voltage noise source, $V_{n3}$, where the transistor is modelled as an equivalent resistor, $R_3$. They are given by,

$$I_{n1}^2 = \frac{4k_BT}{R_1}$$

(3.60)

$$I_{n2}^2 = \frac{4k_BT}{R_2}$$

(3.61)

$$V_{n3}^2 = 4k_BGR_3$$

(3.62)

The equivalent input-referred noise voltage of the pre-amp is given by,

$$V_{neq, total}^2 = V_{neq}^2 + V_{n3}^2 + (I_{n1}^2 + I_{n2}^2) \cdot \frac{R_1^2 R_2^2}{(R_1 + R_2)^2}$$

(3.63)

Substituting $R_1 = 1.517 \, k\Omega$, $R_2 = 13.65 \, k\Omega$, and $R_3 = 100 \, M\Omega$ through (3.60) to (3.63), we obtain $V_{neq, total} = 1.29 \, \mu V/\sqrt{Hz}$, which is primarily contributed from the d.c.-bias resistor $R_3$.

To determine the minimum input-referred noise capacitance change and the minimum input-referred noise displacement from the input-referred noise voltage, $V_{neq, total}$, the relations between the capacitance change and the sensed output voltage, and between the displacement and the sensed output voltage are derived. Substitution of a small capacitance change, $\Delta C = (C_{s1} + C_{s2}) - (C_{s10} + C_{s20})$, into (3.39) leads to the linearized relation between the capacitance change and the sensed output voltage amplitude,

$$\frac{\Delta C}{V_s} = \frac{1}{A_v A_d |V_m|} (C_{10} + C_{20} + C_i)$$

(3.64)

Equations (3.40) and (3.41) can be re-written with a small actuation-plate displacement, $z$, around the initial gap, $g$,
\[ C_{s1} = C_{10} \left(1 + q_1 \frac{z}{g} \right) \]  
(3.65)

\[ C_{s2} = C_{20} \left(1 + q_2 \frac{z}{g} \right) \]  
(3.66)

Substitution of (3.65) and (3.66) into (3.39) leads to the linearized relation between the displacement and the sensed output voltage amplitude,

\[ \frac{z}{|V_s|} = \frac{1}{A_v A_d |V_m|} \frac{(C_{10} + C_{20} + C_i)g}{q_1 C_{10} + q_2 C_{20}} \]  
(3.67)

where the d.c.-bias resistor can be neglected because its magnitude is \(10^2\) to \(10^3\) times higher than the impedance value of pre-amp input capacitance at selected modulation frequency. Replacing the sensed voltage in (3.70) and (3.71) with the output noise voltage, \(A_v A_d V_{neq,total}\), we find the input-referred noise capacitance change,

\[ \Delta C_n^2 = \frac{1}{|V_m|^2} (C_{10} + C_{20} + C_i)^2 V_{neq,total}^2 \]  
(3.68)

and the input-referred noise displacement,

\[ z_n^2 = \frac{1}{|V_m|^2} \frac{(C_{10} + C_{20} + C_i)^2 g^2}{(q_1 C_{10} + q_2 C_{20})^2} V_{neq,total}^2 \]  
(3.69)

By substituting the parameters: \(|V_m| = 1\), \(C_{10} = C_{20} = 8.854 \times 10^{-12} \times (1600 \mu \text{m}^2)/g\), \(C_i = 408 \text{ fF}\), and \(V_{neq,total} = 1.29 \mu \text{V/\sqrt{Hz}}\), in (3.68) and (3.69), the minimum input-referred noise capacitance change and the minimum input-referred noise displacement are plotted as a function of gap in Figure 3.33(a) and (b). The minimum input-referred noise capacitance change is larger at smaller gaps, illustrating the difficulty to sense small capacitance change out of large initial sensing capacitances. However, the minimum
Figure 3.33 (a) Minimum input-referred noise capacitance change versus gap. (b) Minimum input-referred noise displacement versus gap.
input-referred noise displacement reduces with the decrease of gap because of the large capacitive sensor gain at small gaps.

### 3.6.2 Brownian Noise

Air molecules impinging on the micromechanical plates gives rise to a Brownian noise force, \( f_n^2 \), which is related to the damping coefficient, \( b \), of the plate as given by [73],

\[
f_n^2 = 4k_BTb
\]  

(3.70)

The damping coefficient can be expressed in terms of the damping ratio, \( \xi_z \), the actuator resonant frequency, \( \omega_z \), and the mass, \( m \), by

\[
b = 2\xi_z\omega_zm
\]  

(3.71)

Substitution of (3.71) into (3.70) gives,

\[
f_n^2 = 8k_BT\xi_z\omega_zm
\]  

(3.72)

Multiplication of the Brownian noise force with the frequency response of the actuator leads to the frequency response of the noise plate displacement given by,

\[
z_b^2 = \frac{8k_BT\xi_z}{m\omega_z^3}\left(\frac{1}{(1 - \omega^2/\omega_z^2)^2 + (2\xi_z\omega/\omega_z)^2}\right)
\]  

(3.73)

Substituting the actuator parameters: \( m = 3.84 \times 10^{-11} \text{ kg} \), \( \omega_z = 10.5 \text{ kHz} \), and \( \xi_z \) in Figure 3.25, we obtain the equivalent d.c. Brownian noise displacement ranging from \( 2.69 \times 10^{-2} \) nm/\( \sqrt{\text{Hz}} \) to \( 7.75 \times 10^{-4} \) nm/\( \sqrt{\text{Hz}} \) for gaps between 0.2 \( \mu \text{m} \) to 4 \( \mu \text{m} \) as shown in Figure 3.34. Also shown in Figure 3.34 is the equivalent pre-amp noise displacement plotted in Figure
Figure 3.34 Equivalent Brownian noise displacement of the micromechanical actuator and equivalent noise displacement derived from pre-amp noise plotted as a function of gap.

3.33(b). The comparison shows that for open-loop system, the minimum detectable signal is limited by the Brownian noise for gaps below around 0.6 µm at the atmospheric pressure. For the closed-loop system, the Brownian noise force acting on the plates is analogous to a plant-input disturbance. The resultant noise displacement can be lowered by applying a large open-loop gain.
Chapter 4

Controller Design

4.1 Introduction

The controller design for parallel-plate servo uses linearization and input-disturbance rejection techniques to cope with the nonlinearity and open-loop instability of the plant. In this chapter, we will first review the classical frequency-domain design methodology with focus on quantification of essential feedback properties and their trade-offs imposed by a given unstable plant. It is beneficial to understand the cost of feedback for linear unstable plants, such as bandwidth requirement and noise amplification, before designing for the nonlinear unstable plant. Phase margin of the linearized open-loop transfer function is optimized along the plate motion trajectory by use of a proportional-gain controller. Effects of initial conditions (i.e., plate acceleration and velocity) and displacement perturbation are modelled as an input-disturbance force into the linearized system. Disturbance rejection is performed for sensitivity reduction with respect to the disturbance force. Those techniques are applied in the Quantitative Feedback Theory (QFT), in which lots of plant uncertainties can be considered for analysis of stability robustness.

4.2 Classical Frequency-Domain Feedback Theory

Feedback control is used to stabilize unstable plants, to reduce the sensitivity with respect to external disturbance, and to reduce the effect from variations of plant parameters. On the other hand, potential disadvantages also exist. Improper controller design can destabilize systems when plants are stable. Large loop bandwidth can induce
excessive sensor noise at the plant input to deteriorate stability robustness as well. All these properties have to be quantified to understand design limitations and trade-offs, especially for design of unstable plants.

### 4.2.1 Quantification of Feedback Performance

Consider the linear time-invariant closed-loop feedback system shown in Figure 4.1, where $P(s)$ is the plant, $C(s)$ is the controller, $r(s)$ is the command input, $y(s)$ is the system output, $d_1(s)$ is the disturbance at plant input, $d_2(s)$ is the disturbance at plant output, and $n(s)$ is sensor noise.

Define the open-loop transfer function,

$$L(s) = C(s)P(s) \quad (4.1)$$

the sensitivity function,

$$S(s) = \frac{1}{1 + L(s)} \quad (4.2)$$

and the complementary sensitivity function,

$$T(s) = \frac{L(s)}{1 + L(s)} \quad (4.3)$$
By linear superposition, the outputs of the system contributed from \( r, d_1, d_2, \) and \( n \) are summed up and represented by the sensitivity and complementary sensitivity functions,

\[
y(s) = y_r(s) + y_{d1}(s) + y_{d2}(s) + y_n(s) \quad (4.4)
\]

where

\[
y_r(s) = T(s)r(s) \quad (4.5)
\]

\[
y_{d1}(s) = P(s)S(s)d_1(s) \quad (4.6)
\]

\[
y_{d2}(s) = S(s)d_2(s) \quad (4.7)
\]

and

\[
y_n(s) = -T(s)n(s) \quad (4.8)
\]

Similarly, the input to the plant is

\[
u(s) = u_r(s) + u_{d1}(s) + u_{d2}(s) + u_n(s) \quad (4.9)
\]

where

\[
u_r(s) = C(s)S(s)r(s) \quad (4.10)
\]

\[
u_{d1}(s) = S(s)d_1(s) \quad (4.11)
\]

\[
u_{d2}(s) = -C(s)S(s)d_2(s) \quad (4.12)
\]

and

\[
u_n(s) = -C(s)S(s)n(s) \quad (4.13)
\]

Classical design by shaping of open-loop transfer function indirectly alters the feedback properties (sensitivity and complementary sensitivity functions) of the system. Rules of thumb for shaping the open-loop transfer function use a large loop gain at low frequencies for sensitivity reduction and disturbance rejection, and a small loop gain at
high frequencies to achieve small noise response. Phase margin is determined by the roll-off rate of loop gain at the vicinity of the crossover frequency. Large phase margin avoids amplification of disturbances and noise, given that,

\[ |L(j\omega)| = 1, \angle L(j\omega) \equiv -180^\circ \iff |S(j\omega)| \gg 1, |T(j\omega)| \gg 1 \quad (4.14) \]

It is not desirable to have a large loop gain at frequencies for which the plant gain is small. The gains of (4.12) and (4.13) reduce to \(|P^{-1}(s)|\) at those frequencies, thereby inducing large responses at the plant input due to output disturbance and noise.

### 4.2.2 Analysis of Unstable Plant

#### 4.2.2.1 Stability Criterion

Given a feedback system whose open-loop transfer function, \( L(s) \), is strictly proper with possible poles and/or zeros on the \( j\omega \) axis, and \( n \) poles in the right-half \( s \) plane, the Nyquist stability criterion is given as the follows:

**Theorem 4.1** Define the standard Nyquist contour, \( \Gamma \), as the semi-circle of radius \( R \) in the clockwise direction with necessary \( j\omega \)-axis indentations to account for the imaginary poles of \( L(s) \). Let \( R_o \) be the ray connecting between \((-\infty, -1 + j0]\). The crossing between \( R_o \) and the \( L(s) \) locus is said to be positive if the direction is counter-clockwise, and negative otherwise, as shown in Figure 4.2(a). The feedback system is stable, if and only if the Nyquist plot of \( L(s) \) does not intersect \(-1 + j0\), and the net sum of its crossings of \( R_o \) is equal to \( n \) [74].
For stability analysis on the Nichols chart, the criterion is modified as the following [75]:

**Theorem 4.2** Let $R_1$ be the ray defined as $R_1 = \{(\phi, r) | \phi = -180^\circ, r > 0 \text{ dB}\}$. The crossing between $R_1$ and the locus of $L(s)$ on the Nichols chart is said to be positive if the direction is counter-clockwise, and negative otherwise as shown in Figure 4.2(b). The feedback system is stable, if and only if $L(s)$ does not intersect the point $(-180^\circ, 0 \text{ dB})$, and the net sum of its crossings of $R_1$ is equal to $n$.

### 4.2.2.2 Bandwidth Limitations with a Real Unstable Pole

Stabilization of a linear unstable plant requires pulling of unstable poles from the right-half complex plane into the left half plane, thereby implying a lower-bounded open-loop gain. The Bode gain-phase relation [76] states that loop gain must not decrease faster than 40 dB/decade, indicating that the resultant crossover frequency is also lower-bounded.
for stabilizing an unstable plant. To illustrate the increase of crossover frequency, let the open-loop transfer function containing one unstable pole be given by,

\[
L(s) = P(s)C(s)
\]

\[
= \frac{P_s(s)}{s-a}C(s)
\]

\[
= \frac{P_s(s)C(s)}{s+a}, \frac{s+a}{s-a}
\]

\[= L_s(s)A(s)
\]

(4.15)

where \(L_s(s)\) is the stable part of the unstable open-loop function \(L(s)\), and \(A(s)\) is an unstable all-pass function with gain of one throughout all frequencies. The phases of \(L_s(j\omega)\) and \(L(j\omega)\) at the crossover frequency \(\omega_c\) are given by,

\[
\angle L_s(j\omega_c) = -180 + \theta_m
\]

(4.16)

\[
\angle L(j\omega) = \angle L_s(j\omega_c) - 180 + 2\theta|_{\omega = \omega_c}, \tan \theta = \frac{\omega_c}{a}
\]

(4.17)

where \(\theta_m\) is the phase margin of open-loop function \(L_s(s)\). To obtain the same amount of phase margin for \(L_s(s)\) and \(L(s)\), comparing (4.16) and (4.17) gives,

\[
\angle L_s(j\omega_c) = -2\theta|_{\omega = \omega_c} + \theta_m
\]

(4.18)

\(\angle L_s(j\omega_c)\) is the allowed phase lag at the crossover frequency by \(L_s\) for loop-gain roll-off in the stable case. By applying (4.18) to obtain the same amount of phase lag for loop-gain decrease in the unstable case, the crossover frequency, \(\omega_c\), should be at least twice as large as the unstable pole frequency [77].
4.2.2.3 Unstable Poles and the Sensitivity Function

For stable open-loop $L(s)$ with at least two more poles than zeros, there will always exist a frequency range over which sensitivity function in (4.2) is greater than one, as $L(s)$ will have a phase lag of at least $-180^\circ$ for $\omega \to \infty$. A graphical illustration of the sensitivity function magnitude on the complex plane is in Figure 4.3. According to the Bode sensitivity integral, the sensitivity function must satisfy [78],

$$\int_0^\infty \log |S(j\omega)| d\omega = 0 \quad (4.19)$$

which implies that a trade-off exists between sensitivity properties in different frequency ranges, since the area of sensitivity reduction must be equal to the area of sensitivity increase in units of decibels $\times$ (rad/sec).

When $L(s)$ has poles in the open right half plane, denoted by $\{p_i; i = 1, \ldots, n\}$, the sensitivity function will instead satisfy [78],

Figure 4.3  Graphical illustration of the sensitivity function magnitude for $L(s)$ having at least two more poles than zeros. Sensitivity is larger than one when $L(s)$ enters the unit circle centered at $-1 + j0$. 

85
which implies that the area of sensitivity increase is larger than the area of sensitivity reduction by an amount proportional to the distance of unstable poles to the origin. The amount of difference can also be interpreted as the required open-loop gain to pull the unstable poles into the left half complex plane. This open-loop gain can otherwise be used for sensitivity reduction if all poles are stable. However, (4.20) does not give any design implication regarding design trade-offs and limitations. By combining (4.20) and constraints on crossover frequency, loop-gain roll-off rate, and required sensitivity reduction, the peak value of the sensitivity function was derived by Freudenberg [78]. Let \( L(s) \) have a frequency-dependent bound after the crossover frequency given by,

\[
|L(j\omega)| \leq \varepsilon \left[ \frac{\omega_c}{\omega} \right]^{1+k}, \quad \forall \omega \geq \omega_c
\]  

(4.21)

where \( \varepsilon < 1/2 \) and the roll-off rate \( k > 0 \). The sensitivity function is required to satisfy an upper bound for \( \omega \leq \omega_o \),

\[
|S(j\omega)| \leq M_s < 1, \quad \forall \omega \leq \omega_o \leq \omega_c
\]  

(4.22)

(4.20) is manipulated to satisfy both constraints from (4.21) and (4.22), giving

\[
\sup \log |S(j\omega)| \geq \frac{1}{\omega_c - \omega_o} \left\{ \pi \sum_{i=1}^{n} Re[p_i] + \omega_o \log \left( \frac{1}{M_s} \right) - \frac{3 \varepsilon \omega_c}{2k} \right\}, \quad \omega \in (\omega_o, \omega_c)
\]  

(4.23)

(4.23) quantifies the largest peak value of the sensitivity function occurs at frequencies between \( \omega_o \) and \( \omega_c \) due to either of the following reasons: 1) these two frequencies are
very close, 2) the distance of unstable poles to the origin is large, 3) a small $M_s$ is required for sensitivity reduction, and 4) the roll-off rate is fast. Hence a design trade-off is imposed by the bandwidth constraint from (4.21) and the peak sensitivity constraint from (4.23).

### 4.2.2.4 Unstable Poles and the Complementary Sensitivity Function

Two types of frequency-dependent bounds are used to limit the gain of the complementary sensitivity function at high frequencies for noise reduction. They are,

$$ |T(j\omega)| \leq \left( \frac{\omega_c}{\omega} \right)^k, \forall \omega \geq \omega_c \quad (4.24) $$

and

$$ |T(j\omega)| \leq M_t \leq 1, \forall \omega \geq \omega_c \quad (4.25) $$

Denote a real unstable pole of $L(s)$, $p_I$, and $\phi$ is the angle from the crossover frequency on the imaginary axis to the pole.

$$ \phi = \tan^{-1}\left( \frac{\omega_c}{p_1} \right) \quad (4.26) $$

Based on the roll-off rate constraint in (4.24), Looze [79] derived the peak value of the complementary function,

$$ \sup \log |T(j\omega)| \geq k \cdot \frac{-\log(\tan(\phi))\left(\frac{\pi}{2} - \phi\right) + \frac{1}{2} Cl_2(2\phi) + \frac{1}{2} Cl_2(\pi - 2\phi)}{\phi}, \forall \omega > 0 \quad (4.27) $$

where $Cl_2(\cdot)$ is the Clausen integral,

$$ Cl_2(\theta) = -\int_0^\theta \log\left(2\sin\left(\frac{x}{2}\right)\right) dx \quad (4.28) $$
Based on the high-frequency constraint in (4.25), the peak value of the complementary function is constrained by,

\[
\sup_{\omega} \log |T(j\omega)| \geq \frac{\log \left( \frac{1}{M_t} \right) \left( \frac{\pi - \phi}{2} \right)}{\phi}, \quad \forall \omega > 0
\]  

(4.29)

A design trade-off is imposed by the roll-off rate, \( k \), from (4.24) and the complementary sensitivity peak from (4.27), as shown graphically in Figure 4.4. (4.29) is used with \( M_t = 0.5 \) to plot for the curve of \( k = 0 \). As discussed in Section 4.2.2.2, a rule of thumb states that the crossover frequency should be at least twice as large as the unstable pole frequency, which results in a minimum peak value of \(|T(j\omega)|\) less than 4 dB with a one-pole roll-off as shown in Figure 4.4. It is desirable to minimize the peak value in order to maximize phase margin as indicated by (4.14).

Figure 4.4 Minimum complementary sensitivity peak versus ratio of crossover frequency to unstable pole frequency. A rule of thumb for design requires that the crossover frequency at least twice as large as the unstable pole frequency with an one-pole rolloff to reduce the resultant peak within 4 dB.
4.3 Controller Design by Linearization

For given small signals, a nonlinear system behaves similarly to its linearized approximation at each operating point. Local stability is analyzed using the frequency-domain method with phase margin and gain margin, as opposed to the Lyapunov’s linearization method [80] for the state-space approach in which stability is ensured by placing all the closed-loop poles in the left-half complex plane.

Input-command shaping and input-disturbance rejection are utilized to extend local stability around an operating point to global stability as needed in a position-servo problem. Input command shaping ensures that the nonlinear plant behaves mostly like the linearized plants around a time-varying operating point. Input-disturbance rejection reduces the effects of initial conditions and neglected high-order electrostatic force terms during linearization of the nonlinear plant.

A two-degree-of-freedom control system configuration is used for realizing the design procedures stated above. The nonlinear plant is linearized in the stable and unstable regimes, and the maximum phase margin is analyzed using a proportional-gain controller. Controller design regarding the input-disturbance rejection, steady-state error, and shock-force rejection are presented following the stability analysis.

4.3.1 Two-Degree-of-Freedom Control Systems

Two types of trade-off exist in controller design between the sensitivity and complementary sensitivity functions. One is the algebraic trade-off at each frequency, as evidenced by the relation,

\[ S(j\omega) + T(j\omega) = 1 \]  

(4.30)
and the other is the analytical trade-off performed at different frequency ranges as quantified by the integral equation such as (4.20). In other words, the stability requirement and tracking performance can not always be achieved simultaneously. A two-degree-of-freedom system configuration depicted in Figure 4.5 gives another degree of freedom in complementary sensitivity function such that,

$$T(s) = F(s) \frac{L(s)}{1 + L(s)}$$  \hspace{1cm} (4.31)$$

and thus decouples the mutual dependency between $S(j\omega)$ and $T(j\omega)$.

By design of the controller, $C(s)$, sensitivity reduction and disturbance rejection are performed within the loop to achieve stability robustness. The pre-filter design, $F(s)$, shapes the input command to achieve robust performance. The controller and pre-filter are both linear as suggested in the block diagram.

**4.3.2 Closed-Loop Step Response**

The closed-loop step response and the system tracking bandwidth are determined through the pre-filter design. The intended pre-filter design has a relatively smaller band-
width than that of the actuator, thereby the actuator is operated “quasi-statically” during closed-loop servo. Since the linearization technique is utilized for controller design, in this section, the difference of actuation voltage between the quasi-static operation and the static operation is calculated along the planned motion trajectories. This voltage difference represents the initial-condition effect, which is included as part of the plant input-disturbance force in Section 4.3.5.

The resultant closed-loop response of the displacement, $z(s)$, over the step-input command, $r(s)$, can be described by a third-order transfer function,

$$ T(s) = \frac{z(s)}{r(s)} = \frac{1}{(s/\omega_b + 1)(s/\omega_h + 1)^2} $$

(4.32)

where $T(s)$ has a dominant pole $\omega = \omega_b$ which is the pre-filter bandwidth, and two far-away poles at $\omega = \omega_h$. $\omega_h$ is close to the closed-loop bandwidth, namely, the actuator bandwidth $\omega_n = \sqrt{k/m}$. Substituting pre-filter bandwidth $\omega_b = r_1\omega_n$ and $\omega_h = \omega_n$ into (4.32) gives,

$$ T(s) = \frac{1}{\left(\frac{s}{r_1\omega_n} + 1\right)\left(\frac{s}{\omega_n} + 1\right)^2} $$

(4.33)

$T(s)$ is at least a third-order transfer function in order to satisfy the initial conditions, $z(0) = \dot{z}(0) = \ddot{z}(0) = 0$, as suggested by the initial value theorem; otherwise the derived output response has an initial velocity and/or an initial acceleration at $t = 0$, which is not practical for physical systems.
The closed-loop transient response is determined by the selection of pre-filter, and can be approximated by applying a displacement step input command to this third-order closed-loop transfer function. Substituting \( r(s) = \frac{Z_{o,max}}{s} \) into (4.33) and applying the inverse Laplace transform gives,

\[
z(t) = Z_{o,max} \left(1 - \left(\frac{1}{1 - r_1}\right)^2 e^{-r_1\omega_n t} - \left(\frac{(r_1 - 2)r_1}{(1 - r_1)^2} + \frac{r_1\omega_n t}{r_1 - 1}\right) e^{-\omega_n t}\right)
\]  

(4.34)

\( \dot{z}(t) \) and \( \ddot{z}(t) \) can be obtained accordingly by differentiations of (4.34). Dividing (4.34) by the initial gap, \( g \), and using the normalized time axis \( \tau = \omega_n t \), we obtain the normalized displacement,

\[
\tilde{z}(\tau) = \alpha_{max} \left(1 - \left(\frac{r_2}{r_2 - r_1}\right)^2 e^{-r_1\tau} - \left(\frac{(r_1 - 2)r_2}{(r_2 - r_1)^2} + \frac{r_1r_2\tau}{r_1 - r_2}\right) e^{-r_2\tau}\right)
\]  

(4.35)

where \( \tilde{z}(\tau) = z(\tau)/g \), \( \alpha_{max} = Z_{o,max}/g \), and \( \alpha = Z_o/g \) is the normalized displacement. Rearranging the parallel-plate dynamic equation in (3.4) with the normalized displacement \( \tilde{z}(\tau) \) yields the applied voltage,

\[
V(\tau) = (1 - \tilde{z}(\tau))g \cdot \sqrt{\frac{2m g}{\varepsilon_o A}} \left(\frac{d^2\tilde{z}(\tau)}{d\tau^2} + 2\xi \omega_n \frac{d\tilde{z}(\tau)}{d\tau} + \omega_n^2 \tilde{z}(\tau)\right)
\]  

(4.36)

where the damping ratio \( \xi = b/(2\sqrt{mk}) \). To calculate the steady-state actuation voltage for displacement up to \( \alpha_{max}g \), we remove the dynamic terms in (4.36). The steady-state actuation voltage for normalized displacements between \([0, Z_{o,max}/g]\) is,

\[
V_{ss} = (1 - \alpha)\omega_n g \cdot \sqrt{\frac{2m g}{\varepsilon_o A}} \alpha
\]  

(4.37)
From (4.36) and (4.37), the ratio of the voltage difference between the transient and steady-state voltages over the steady-state voltage at each normalized position $\beta$ is given by,

\[ \frac{V(\tau) - V_{ss}}{V_{ss}} = \frac{(1 - \ddot{z}(\tau)) \left[ \frac{1}{\omega_n^2} \frac{d^2}{d\tau^2} \ddot{z}(\tau) + \frac{2\zeta}{\omega_n} \dot{z}(\tau) + \ddot{z}(\tau) \right] - (1 - \alpha) \sqrt{\alpha}}{(1 - \alpha) \sqrt{\alpha}} \]  

(4.38)

The ratio of the maximum transient voltage, $V_{max}$, over static actuator pull-in voltage $V_{pi}$ in (3.8) is,

\[ \frac{V_{max}}{V_{pi}} = \frac{3\sqrt{3}}{2} \cdot \max \left( (1 - \ddot{z}(\tau)) \left[ \frac{1}{\omega_n^2} \frac{d^2}{d\tau^2} \ddot{z}(\tau) + \frac{2\zeta}{\omega_n} \dot{z}(\tau) + \ddot{z}(\tau) \right] \right), \forall \tau, \tau \leq 1 \]  

(4.39)

The plant parameters given in Chapter 3 and a target displacement of 90% of the initial gap ($g = 3 \mu m$) are substituted in (4.35) and (4.38) for slow closed-loop responses ($r_1 = 0.05$ to 0.2) and fast closed-loop responses ($r_1 = 0.25$ to 1). The voltage differences of quasi-static and static operations are plotted in Figure 4.6(a) and (b) as a function of normalized displacement. Both graphs show that the difference increases as the closed-loop response becomes faster. In addition, the responses exhibit mostly transient behavior at small displacements to provide initial acceleration for the plate, and the difference increases again at large displacements due to the increasing damping ratio. We will perform the position servo for $\omega_b$ less than 0.05 $\omega_n$ to meet the quasi-static conditions used by the linearization technique. The deviation from a linear operation shown in Figure 4.6, is an initial condition which is considered in Section 4.3.5 as an input disturbance to the linearized plant.
Figure 4.6 Percentage ratio of the voltage difference between the transient and steady-state responses as a function of normalized displacement. (a) tracking bandwidth $\omega_b/\omega_n = 0.05$ to $0.20\omega_n$. (b) tracking bandwidth $\omega_b/\omega_n = 0.25\omega_n$ to $\omega_n$. 

94
We also apply the plant parameters into (4.39) and plot in Figure 4.7 the ratio of the maximum transient voltage over the static pull-in voltage as a function of the tracking bandwidth. A maximum supplied voltage about five times of the static pull-in voltage is required to servo the actuator at its mechanical bandwidth.

4.3.3 Plant Linearization

The electrostatic actuator dynamics (3.4) is re-written here by,

\[ m \ddot{z} + b \dot{z} + k z = F_e \]

\[ = \frac{\varepsilon_o A V^2}{2(g - z)^2} \]  

(4.40)

Given a small variation of \( \Delta z \) and \( \Delta v \) around the operating point \((Z_o, V_o)\), the linearized dynamic equation expanded by the Taylor’s series is,
\[ m\ddot{z} + b\dot{z} + k(Z_o + \Delta z) = F_e|_{Z_o, V_o} + \Delta F_e \]

\[ \equiv F_e|_{Z_o, V_o} + \frac{\partial F_e}{\partial z}|_{Z_o, V_o} \Delta z + \frac{\partial F_e}{\partial V}|_{Z_o, V_o} \Delta v \]  \hspace{1cm} (4.41)

Removing the balanced spring and electrostatic forces in (4.41), and substituting the balanced-force equation at d.c. equilibrium,

\[ kZ_o = \frac{\varepsilon_o A}{2(g-Z_o)^2} V_o^2 \]  \hspace{1cm} (4.42)

into (4.41), we have,

\[ m\ddot{z} + b\dot{z} + (k + k_e)\Delta z = \frac{2kZ_o}{V_o} \Delta v \]  \hspace{1cm} (4.43)

where,

\[ k_e = -\frac{\partial F_e}{\partial z}|_{Z_o, V_o} = -\left(\frac{2\alpha}{1-\alpha}\right)k \]  \hspace{1cm} (4.44)

\( \alpha \) in (4.44) is the normalized displacement, \( Z_o/g \). The electrostatic force gradient, \( k_e \), induces a “negative” spring constant which is a function of the plate displacement. As the ratio \( \alpha \) becomes 1/3, electrical spring constant completely negates the mechanical spring constant, thereby creating a pole frequency at zero. The actuator behaves like a “mechanical integrator” at this point, and becomes unstable as \( \alpha \) increases beyond 1/3.

By using (4.43) and the pull-in voltage expression in (3.8), the d.c. actuator gain is given by,

\[ G_{act} = \frac{4\sqrt{3}\alpha}{9|1-3\alpha|} \frac{g}{V_{pi}} \]  \hspace{1cm} (4.45)
The normalized d.c. actuator gain as a function of normalized displacement is plotted in Figure 4.8. A large gain variation shown in the neighborhood close to $\alpha = 1/3$ raises the stability concern regarding designed gain margins. The low d.c. gain at the small-displacement region implies that the actuator transient response can be sluggish initially.

By combining the transfer functions of the actuator in (4.43) and the capacitive sensor in (3.42), a complete plant transfer function is given by,

$$ P(s) = \frac{G_s(Z_o) \cdot \left(\frac{2kZ_o}{V_o}\right)}{\left(\frac{s}{\omega_p} + 1\right)(ms^2 + bs + (k + k_e))} \quad (4.46) $$

In the system block diagram in Figure 4.9, the nonlinear plant is represented by a family of linearized plants at different operating points.

We define the effective resonant frequency and effective damping ratio affected by the electrostatic force gradient as,
Figure 4.9  The nonlinear plant is replaced by a set of linearized plants at different operating points.

\[
\omega_e = \frac{1 - 3\alpha\omega_n}{\sqrt{1 - \alpha}} \quad (4.47)
\]

\[
\xi_e = \sqrt{\frac{1 - \alpha}{1 - 3\alpha}} \xi \quad (4.48)
\]

in which \( \omega_n = \sqrt{k/m} \) and \( \xi = b/(2\sqrt{mk}) \). Substituting (4.47) and (4.48) into (4.46) yields the plant model in the stable and unstable regimes (\( \alpha \leq 1/3 \) and \( \alpha > 1/3 \)),

\[
P(s) = \frac{G_s(Z_0)\left(\frac{2Z_o}{V_o}\right)\omega_n^2}{\left(\frac{s}{\omega_p} + 1\right)\left(s^2 + 2\xi_e\omega_e s + \omega_e^2\right)} , \text{ stable} \quad (4.49)
\]

\[
P(s) = \frac{G_s(Z_0)\left(\frac{2Z_o}{V_o}\right)\omega_n^2}{\left(\frac{s}{\omega_p} + 1\right)\left(s^2 + 2\xi_e\omega_e s - \omega_e^2\right)} , \text{ unstable} \quad (4.50)
\]

### 4.3.4 Phase-Margin Optimization using a Proportional-Gain Controller

The Nyquist stability criterion indicates that an unstable open-loop with \( n \) unstable poles must have \( n \) crossings of the ray \((-1 + j0, -\infty)\) on the Nyquist plot, or the ray:

\[\{(\phi, r) | \phi = -180^\circ, r > 0 \text{ dB}\}\] on the Nichols chart. With one unstable pole, and the num-
ber of poles is at least two more than the number of zeros, the \( L(j\omega) \) trajectory must start from \(-180^\circ\) on the Nichols chart (or the Bode plots) to result in a stable system, as shown in Figure 4.10(a) and (b). The lower gain-margin, \( G_L \), and the upper gain-margins, \( G_H \), are defined at frequencies when \( \angle L(j\omega) = -180^\circ \), and the phase margin, \( \phi \), is defined at the loop unity-gain frequency when \(|L(j\omega)| = 1\). \( L(j\omega) \) has a finite lower gain margin as a result of having an unstable pole, as compared to an infinite lower gain-margin for a stable open-loop transfer function.

Given a set of linearized unstable plants, we will analyze the maximum attainable phase margin using a proportional-gain controller, and find the resultant lower and upper gain-margins, and the minimum crossover frequency accordingly.

### 4.3.4.1 Plant in the Unstable Regime

By multiplying a proportional-gain controller, \( C(s) = k_p \), with the unstable linearized plant in (4.50), the resultant open-loop function is,

\[
L(s) = C(s)P(s) = \frac{K}{\left(\frac{s}{\omega_p} + 1\right)\left(s^2 + 2\xi\omega_n s - \omega_n^2\right)}
\]

where

\[
K = G_s(Z_o)\left(\frac{2Z_o}{V_o}\right)\omega_n^2 k_p
\]

By using standard complex number arithmetic, magnitude and phase of \( L(j\omega) \) are expressed by,
Figure 4.10 With one unstable pole, $\angle L(j\omega)$ must start from $-180^\circ$ to result in a stable system, as shown in (a) the Nichols chart, and (b) the Bode plots.
\[ |L(j\omega)| = \frac{K}{\sqrt{1 + \omega^2/\omega_p^2}(\omega^2 + \omega_e^2)^2 + (2\xi_e \omega_e \omega)^2} \] (4.53)

and

\[ \angle L(j\omega) = -\pi + \tan^{-1}\left(\frac{2\xi_e \omega_e \omega^2}{\omega^2 + \omega_e^2}\right) - \tan^{-1}\left(\frac{\omega}{\omega_p}\right) \] (4.54)

At zero frequency \((L(j0) = G_L)\), and (4.53) becomes,

\[ G_L = \frac{K}{\omega_e^2} \] (4.55)

Substituting (4.55) into (4.53), and equating (4.53) to unity at the crossover frequency, \(\omega_e\), yields

\[ G_L = \frac{\sqrt{\omega_p^2 + \omega_e^2}}{\omega_p \omega_e^2} \sqrt{(\omega_e^2 + \omega_e^2)^2 + (2\xi_e \omega_e \omega_e)^2} \] (4.56)

From (4.55) and (4.56), we have,

\[ K = \left(\frac{r_1^2 + r_2^2}{r_2^2} \sqrt{r_1^2 + 1} + (2\xi_e r_1)^2\right) \cdot \omega_e^2 \] (4.57)

in which \(r_1 = \omega_e/\omega_p\) and \(r_2 = \omega_p/\omega_e\). For stability consideration, \(G_L\) must be larger than one (0 dB), thus the following inequality must hold,

\[ (\omega_p^2 + \omega_e^2)((\omega_e^2 + \omega_e^2)^2 + (2\xi_e \omega_e \omega_e)^2) > \omega_p^2 \omega_e^4 \] (4.58)

Equating \(\angle L(j\omega)\) to \(-\pi\) at the upper gain-margin frequency, \(\omega_g\), gives,

\[ \omega_g = \sqrt{\omega_e (2\xi_e \omega_p - \omega_e)} \] (4.59)

Substituting (4.59) into (4.53) gives the upper gain margin,
\[ G_H = \frac{2\xi_e \omega_p (1 + 2\xi_e \omega_e / \omega_p - \omega_e^2 / \omega_p^2)}{G_L \omega_e} \]  \hspace{1cm} (4.60)

\( G_H \) must be larger than unity for stability consideration. At \( \omega = \omega_c \), the phase angle of \( L(j\omega) \) in (4.54) gives,

\[-\pi + \phi = -\pi + \tan^{-1}\left(\frac{2\xi_e \omega_e \omega_c}{\omega_c^2 + \omega_e^2}\right) - \tan^{-1}\left(\frac{\omega_c}{\omega_p}\right) \hspace{1cm} (4.61)\]

\[ \phi = \tan^{-1}\left(\frac{\frac{2\xi_e \omega_e \omega_c}{\omega_c^2 + \omega_e^2} - \omega_p}{1 + \frac{2\xi_e \omega_e \omega_c}{\omega_c^2 + \omega_e^2} \cdot \omega_p}\right) \hspace{1cm} (4.62)\]

To get the maximum phase margin, (4.54) is differentiated with respect to \( \omega \) at \( \omega = \omega_c \).

The resultant crossover frequency is the solution of the following 4th-order equation:

\[-\omega_p - 2\xi_e \omega_e \omega_c + (- 2\xi_e \omega_p^2 \omega_e - 2\omega_p \omega_e^2 - 4\xi_e^2 \omega_p \omega_e^2 + 2\xi_e \omega_e^3) \omega_c^2 + (2\xi_e \omega_p^2 \omega_e^3 - \omega_p \omega_e^3) = 0 \hspace{1cm} (4.63)\]

At each operating point, \( \omega_e \) and \( \xi_e \) are first computed according to (4.47) and (4.48). Given the \( \omega_p \)/\( \omega_e \) ratio and \( \xi_e \), \( \omega_c \) is found from (4.63). Then the maximum phase margin, the upper and lower gain-margins and the controller gain are solved from (4.62), (4.60), (4.56), and (4.52), respectively, at each operating point. The optimal controller gain which gives the maximum phase margin varies with the displacement, implying that the optimal controller must be linear time-varying (LTV) in order to secure the maximum phase margins at all displacements. We implement a linear time-invariant (LTI) controller whose gain is selected among the \( k_p \)'s obtained at all operating points while giving reason-
able phase and gain margins for all displacements. A complete LTI controller design flow is depicted in Figure 4.11.

Using the plant parameters given in Chapter 3 with an initial gap of 3 µm, we plot the maximum phase margin, the corresponding lower and upper gain margins, the crossover frequency, and the controller gain from Figure 4.12(a) to Figure 4.12(e). The phase margin at large displacements is over 60° due to the large damping ratios at small gaps. The selected LTI controller is $k_p = 16$ which gives all the operating points a minimum phase margin of 60°. The effective damping ratio and the effective resonant frequency are plotted as a function of normalized displacement in Figure 4.13(a) and (b), respectively. The resultant open-loop transfer functions of LTV and LTI designs are plotted on the Nichols chart in Figure 4.14(a) and (b) for selected normalized displacements.

### 4.3.4.2 Plant in the Stable Regime

By multiplying a proportional-gain controller, $C(s) = k_p$, with the stable linearized plant in (4.49), the open-loop function below the pull-in limit is,

$$L(s) = C(s)P(s) = \frac{K}{\left(\frac{s}{\omega_p} + 1\right)(s^2 + 2\xi_e \omega_e s + \omega_e^2)}$$

Without unstable poles, the phase of $L(j\omega)$ starts from $0^\circ$ for $\alpha < 1/3$ or $-90^\circ$ for $\alpha = 1/3$, thereby the lower gain margin does not exist.

Equating $\angle L(j\omega)$ of (4.64) to $-\pi$ at the gain-margin frequency, $\omega_g$, gives,

$$\omega_g = \sqrt{\omega_e (2\xi_e \omega_p + \omega_e)}$$

(4.65)
Figure 4.11 Flow chart for LTI proportional-gain controller design for a set of linearized unstable plants. Final controller design $k_p$ must ensure adequate phase and gain margins for all the linearized plants.
Figure 4.12 Maximum phase margin, and corresponding crossover frequency, lower gain margin, upper gain margin and controller gain plotted as a function of normalized displacement from (a) to (e). Selected LTI controller is $k_p = 16$ which gives a minimum phase margin around $60^\circ$ at all displacements.
Figure 4.13 (a) Effective damping ratio as a function of normalized displacement. (b) Ratio of effective resonant frequency over mechanical resonant frequency as a function of normalized displacement.
Figure 4.14 (a) Open-loop transfer function of LTV design on the Nichols chart. (b) Open-loop transfer function of LTI design on the Nichols chart.
Substituting (4.65) into (4.64) gives,

\[ G_H = \frac{2 \xi_e \omega_c (\omega_p^2 + 2 \xi_e \omega_c \omega_p + \omega_e^2)}{K \omega_p} \]  \hspace{1cm} (4.66)

For stability consideration, \( G_H \) must be larger than unity (0 dB). Thus the following inequality sets the upper controller gain limit:

\[ k_p < \frac{V_o}{G(Z_o)Z_o} \cdot \frac{\xi_e \omega_c (\omega_p^2 + 2 \xi_e \omega_c \omega_p + \omega_e^2)}{\omega_p \omega_n^2} \]  \hspace{1cm} (4.67)

Substituting the plant parameters given in Chapter 3 for plate displacement up to one-third of the gap \( (g = 3 \mu m) \), we plot the maximum stable controller gain as a function of normalized displacement in Figure 4.15. The \( k_p \) values are at least an order of magnitude larger than \( k_p \) values obtained in the unstable regime as shown in Figure 4.12(e). Hence there is no conflict in the designed controller gain in the stable and unstable regimes.

Figure 4.15 Maximum controller gain \( k_p \) versus normalized plate displacement by stability requirement.
4.3.5 Plant Input Disturbance Rejection

Using linearization, we neglect the initial conditions of plate and the high-order terms of electrostatic force which are important to the nonlinear system stability. These effects are modelled as plant input disturbances in this section. We perform disturbance rejection to reduce the induced plate displacement due to disturbances, thereby increasing the stability robustness of the system.

4.3.5.1 Problem Formulation

By linearization, only the first-order terms of the expanded nonlinear electrostatic force in the Taylor’s series are considered as in (4.41). The sum of the remaining high-order terms, denoted as $\Delta F_d$, at operation point $(Z_o, V_o)$ is given by,

$$
\Delta F_d = \sum_{i=2}^{\infty} \frac{\partial F_e}{\partial z^i} \Delta z^i + \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} \left( \frac{\partial F_e}{\partial z^i} \frac{\partial F_e}{\partial v^j} \right) \Delta z^i \Delta v^j + \sum_{i=2}^{\infty} \frac{\partial F_e}{\partial v^i} \Delta v^i
$$

(4.68)

where $\Delta F_{d,z}$, $\Delta F_{d,zv}$ and $\Delta F_{d,v}$ are the small-signal disturbance forces with respect to $\Delta z$, $\Delta z$ and $\Delta v$, respectively. Substituting the electrostatic force expression of (2.16) into (4.68), we can represent $\Delta F_{d,z}$, $\Delta F_{d,zv}$ and $\Delta F_{d,v}$ as the multiple of the spring restoring force at $z = Z_o$ as the following:

$$
\Delta F_{d,z} = \left( \sum_{n=2}^{\infty} \frac{(n+1)(\alpha r_z)^n}{(1-\alpha)^n} \right) \cdot kZ_o
$$

(4.69)

$$
\Delta F_{d,zv} = \left( \sum_{n=2}^{\infty} \frac{2r_v(\alpha r_z)^{n-1}}{(1-\alpha)^{n-1}} + \sum_{n=2}^{\infty} \frac{2r_v^2(\alpha r_z)^{n-1}}{(n+1)(1-\alpha)^{n-1}} \right) \cdot kZ_o
$$

(4.70)
\[ \Delta F_{d,z} = r_v^2 kZ_o \]

where \( r_z = \Delta z/Z_o \) and \( r_v = \Delta v/V_o \).

\( \Delta F_{d,z} \) is the dominant disturbance force at small gaps. The ratio of \( \Delta F_{d,z} \) over the first-order electrostatic force, \( \Delta F_{e,1} \), as a function of normalized displacement is shown in Figure 4.16, which illustrates that \( \Delta F_{d,z} \) dominates over the first-order term for \( Z_o/g > 0.8 \) and \( \Delta z/Z_o > 0.075 \). Therefore performing disturbance rejection in this region is especially important.

The block diagram with the input disturbance force is shown in Figure 4.17, in which the voltage-to-force gain of \( 2kZ_o/V_o \) in the original plant is included within the redefined controller, \( C'(s) = C(s) \cdot 2kZ_o/V_o \), and the plant is redefined as \( P'(s) = P(s)/(2kZ_o/V_o) \).

The aforementioned linear stability analysis assumes that the motion trajectory exactly follows the steady-state displacement and applied voltage, yet deviations exist due
to the quasi-static operation, as shown by the elliptical disc with radii $\Delta z$ and $\Delta v$ on the static displacement-voltage plot in Figure 4.18. The implication of performing disturbance rejection with respect to the perturbed displacement and voltage and their induced disturbance force, $\Delta F_d$, is to suppress the final perturbed displacement $\Delta \tilde{z}$ as shown in Figure 4.17. Therefore the quasi-static trajectory is brought closer to the neighborhood of linearized operating point through disturbance rejection.

### 4.3.5.2 Controller Design

Graphically speaking, as shown in Figure 4.19, the operating point $z = Z_o$ is said to be asymptotically stable if for $\Delta z > 0$ such that for $\|z(t_o)\| < \Delta z$, then $\|z(t_o + t)\| \rightarrow Z_o$ for $t \rightarrow \infty$. The final destination would be biased from $Z_o$ by $\Delta \tilde{z}$ under the plant input disturbance, shown as the point $Z'_o$ on the same plot. By performing disturbance rejection with a large loop transmission gain at low frequencies, asymptotically stable operation can be achieved with $Z'_o \equiv Z_o$. 

![Diagram](image.png)
From (4.6) and Figure 4.17, the input disturbance force and the induced sensor output are related by,

\[
\left| \frac{\Delta V_s}{\Delta F_d} \right| = \left| \frac{P'(s)}{1 + C'(s)P'(s)} \right| \tag{4.72}
\]

where,
\[
C'(s) = \frac{2kk_pZ_o}{V_o} \tag{4.73}
\]
\[
P'(s) = \frac{G_s(Z_o)}{(ms^2 + bs + (k + k_c))(1 + s/\omega_p)} \tag{4.74}
\]

Assume that \(|C'P'| > 1\) at low frequencies, we substitute \(\Delta v_s = \Delta z G_s(Z_o)\), (4.73) and (4.74) into (4.72), producing the steady-state displacement, \(\Delta \tilde{z}\), expressed by,

\[
\left| \frac{\Delta \tilde{z}}{\Delta F_d} \right| = \frac{|P'(0)|}{|C'(0)||P'(0)|}
= \frac{V_o}{2kk_pZ_oG_s(Z_o)}
\tag{4.75}
\]

Hence, required \(k_p\) is given by,

\[
k_p = \frac{V_o \Delta F_d}{2kZ_o G_s(Z_o) \Delta \tilde{z}} \tag{4.76}
\]

\(\Delta F_d\) is calculated at each normalized displacement by substituting the perturbed displacement ratio, \(\Delta z/Z_o = 0.075\), and \(\Delta v/V_o\) from Figure 4.6(a) \((\omega / \omega_n = 0.05)\) through (4.69) to (4.71). By substituting the plant parameters given in Chapter 3, the \(\Delta F_d\) values, and the resultant displacement ratio, \(\Delta \tilde{z}/Z_o = 0.01\), into (4.76), the required controller gain for disturbance rejection is plotted as a function of normalized displacement in Figure 4.20. Also shown in the plot are the optimal LTV design from Figure 4.12(e), and the maximum LTI controller gain \(k_p = 220\) to maintain marginal system stability (phase margin = 0°). The controller gain increases below \(Z_d/g < 0.2\) and beyond \(Z_d/g > 0.8\) due to the large \(\Delta F_d\) values resulting from large \(\Delta v/V_o\) and \(\Delta z/Z_o\), respectively. The required con-
controller gain in the middle region is close to our LTI controller design \( (k_p = 16) \). System stability in the small-displacement region is not affected because the actuator is in the stable regime, and required controller gain for disturbance rejection meets our LTI design before reaching into the unstable regime. However for \( Z_o/g > 0.8 \), the large controller gain can be in conflict with the controller gain for stability requirement as shown in Figure 4.20, which illustrates the previous LTV controller design for comparison.

### 4.3.6 Steady-State Error

The use of a proportional-gain controller results in finite open-loop gain at d.c., thereby producing a steady-state error when the system is at equilibrium. The system diagram at steady state is shown in Figure 4.21, in which \( e_o, V_o, F_{eo}, Z_o, \) and \( V_{so} \) are the steady-state values of error, controller output voltage, electrostatic force, plate displace-
ment, and sensed voltage. The voltage-to-force gain, \( f(Z_o) \), is a function of displacement given by,

\[
f(Z_o) = \frac{\sqrt{2\varepsilon_oAkZ_o}}{2(g - Z_o)}
\]  

(4.77)

and the large-signal sensor gain at steady state is,

\[
G_{so}(Z_o) = \frac{V_{so}}{Z_o}
\]

(4.78)

Multiplying all the gains within the loop gives the total d.c. loop gain at \( z = Z_o \) as,

\[
|L(j0)|_{z = Z_o} = \frac{k_p f(Z_o) G_o(Z_o)}{k}
\]

\[
= \frac{k_p V_{so}}{2k(g - Z_o)\sqrt{\frac{2\varepsilon_oAk}{Z_o}}}
\]

(4.79)

The steady-state error is related to the input-command magnitude, \( r \), by,

\[
\frac{e_o}{r} = \frac{1}{1 + |L(j0)|_{z = Z_o}}
\]

(4.80)

By substituting (4.79) into (4.80), and assuming that \( |L(j0)|_{z = Z_o} \approx 1 \), (4.80) becomes,

\[
\frac{e_o}{r} \approx \frac{2k(g - Z_o)}{k_p V_{so} \sqrt{\frac{Z_o}{2\varepsilon_oAk}}}
\]

(4.81)
Substituting the plant parameters given in Chapter 3 and \( k_p = 16 \) into (4.81), we plot the percentage steady-state error as a function of normalized displacement in Figure 4.22. Both \( f(Z_o) \) and \( G_{so}(Z_o) \) increase with respect to the increasing plate displacement, thereby resulting in the decrease of steady-state error.

### 4.3.7 Shock Rejection

In this section, we apply an external shock force to the actuator and use steady-state analysis to compute the maximum shock-induced displacement with respect to the controller gain, \( k_p \). We assume that the actuator is initially at equilibrium when a shock force is applied, which temporarily changes the steady state of the actuator until being completely removed.

The system block diagram after the application of external shock force, \( F_s \), is illustrated in Figure 4.23, in which \( e_o \) and \( Z_o \) are the initial steady-state values of error and
plate displacement, $\Delta e_s$ and $\Delta z_s$ are the changes of steady-state error and steady-state displacement, and $V_{o'}$, $F_{eo'}$, and $V_{so'}$ are the new steady-state values of controller output voltage, electrostatic force, and sensed voltage. By solving the balanced electrostatic force, spring restoring force, and shock force, the voltage-to-force function, $f'$, is given by,

$$f' = \frac{\sqrt{2}\varepsilon_o A(k(Z_o + \Delta z_s) - F_s)}{2(g - (Z_o + \Delta z_s))}$$

(4.82)

The open-loop functions before and after the application of shock force are defined as $L$ and $L'$, respectively. From Figure 4.22 and Figure 4.23, the change of the sensor output is given by,

$$V_{so'} - V_{so} = \left( \frac{L' r}{1 + L'} + \frac{P' F_s}{1 + L'} \right) - \frac{L r}{1 + L}$$

$$= \left( \frac{L'}{1 + L'} - \frac{L}{1 + L} \right)r + \frac{P' F_s}{1 + L'}$$

(4.83)

Define

$$\psi \equiv \frac{f \cdot G_{so}(Z_o)}{k}$$

(4.84)

and
Substituting (4.84), (4.85) and $r$ into (4.83) gives,

$$
\psi' = \frac{f' \cdot G_{so}(Z_o + \Delta z_s)}{k}
$$

(4.85)

Substituting (4.84), (4.85) and $r = (1 + L)V_{so}/L$ into (4.83) gives,

$$
k_p = \frac{V_{so}(\psi' - \psi) + \psi P'F_s}{(V_{so} - V_{so})\psi \psi'} - \frac{1}{\psi'}
$$

(4.86)

We apply the plant parameters given in Chapter 3 and a 150 g shock force into (4.86). The minimum controller gain for suppressing the shock-induced displacement, $\Delta z_s$, under 100 nm, 10 nm and 1 nm is plotted in Figure 4.24 as a function of normalized displacement. Both $f'$ and $G_{so}'$ increase with respect to the increasing plate displacement, thereby reducing the controller gain for shock rejection. For the current design of $k_p = 16$, $\Delta z_s$ varies from a few hundred nanometers to several nanometers depending on the steady-state position at which the shock force is applied.

Figure 4.24 Minimum required controller gain versus normalized displacement for suppressing the induced displacement from an 150 g shock force to 100 nm, 10 nm, and 1 nm.
4.4 Controller Design by the QFT (Quantitative Feedback Theory) Method

The frequency-domain design procedure discussed in Section 4.3 only deals with a nominal nonlinear plant, whose parameters can be uncertain within a range instead of being a fixed value. For example, the initial gap can vary between 2 to 3 µm due to the variation of actuator curl. To design efficiently with plant uncertainties, the previous approach is inappropriate for its lack of robustness consideration, which is why the QFT method is presented in this section.

The QFT is a frequency-domain design method developed by Horowitz [81] based on two important observation of feedback systems: (1) the use of feedback is to achieved a desired plant output response in the presence of plant uncertainty, and/or unknown external disturbances; and (2) controller design with the least required bandwidth is preferred. The canonical QFT system configuration is the two degree-of-freedom diagram shown in Figure 4.5. We modify the design procedure given in [81] for use in our design problem. The design steps are:

1. Design of the pre-filter based on the desired tracking bandwidth as discussed in Section 4.3.2.
2. Derivation of templates of the uncertain linearized plants on the complex plane (the Nichols chart) at frequencies \( \omega_i \) \((i = 1, 2, ..., n)\).
3. Selection of a nominal plant, \( P_o(s) \), among all linearized plants, and derivation of bounds on the nominal open-loop function, \( L_o(j\omega) \), at \( \omega = \omega_i \) based on stability and disturbance-rejection specifications.
4. Design \( L_o(j\omega) \) based on bounds at \( \omega = \omega_i \).
4.4.1 Generation of Bounds

The essence of the QFT method is the transformation of robust stability and robust performance specifications into domains on the complex plane (the Nichols chart), referred to as bounds, which a nominal loop transmission should obey in design. The specifications considered in our design include:

1. Gain and phase margins:
   \[ |T(j\omega)| \leq \lambda \]  
   \[ (4.87) \]

2. Plant input disturbance rejection:
   \[ |P(j\omega)S(j\omega)| \leq w_1(\omega) \]  
   \[ (4.88) \]

The lower gain margin and phase margin are related to \( \lambda \) by [82],

\[ G_L = 1 + \frac{1}{\lambda}. \]  
\[ (4.89) \]

and

\[ \phi = 180^{\circ} - \cos^{-1}\left(\frac{1}{2\lambda^2} - 1\right) \]  
\[ (4.90) \]

For example, \( \lambda = 1.2 \) implies at least 50\(^{\circ}\) phase margin and at least 1.66 (4.4 dB) upper gain margin.

To derive bounds for stability margins, we represent the controller and plant transfer functions in their polar form: \( C = ce^{j\phi} \) and \( P = pe^{j\theta} \), and substitute them into (4.87),

\[ \frac{|cpe^{j(\phi + \theta)}|}{|1 + cpe^{j(\phi + \theta)}|} \leq \lambda \]  
\[ (4.91) \]

Evaluating the magnitude and squaring both sides yields a second-order inequality.
Similarly, the input-disturbance rejection specification in (4.88) gives the inequality,

\[

c^2 + [2p\cos(\phi + \theta)]c + 1 \geq 0
\]  \hspace{1cm} (4.92)

Following the procedure described in [82], as illustrated in the flow chart in Figure 4.25, the bounds for the nominal loop transmission \(L_o(j\omega)\) are obtained. The final controller design is not affected by the selected nominal plant. Graphically speaking, the bound computed at each frequency is categorized into two types as depicted on the Nichols chart in Figure 4.26. The area indicated by arrows is where the nominal loop transmission must fall within if given one of them, to satisfy the inequalities (4.92) and (4.93).

The uncertain plant that we consider has a gap variation between 2 to 3 \(\mu m\). The gap is discretized by \(n\) points as,

\[
g = \{g_i, i = 1, \ldots, n\}, g_i \in [g_{min}, g_{max}]
\]  \hspace{1cm} (4.94)

We set \(n = 3\), \(g_{min} = 2\ \mu m\), and \(g_{max} = 3\ \mu m\). At each \(g_i\), the nonlinear plant is linearized at a total of \(m\) positions, \(a_{ij}\ (j = 1 \text{ to } m)\), equally spaced by \(\Delta\alpha_i = (\alpha_{i, max} - \alpha_{i, min})/(m - 1)\), where \(\alpha\) is the normalized displacement. We set \(m = 5\).

The total number of linearized plants used for bound generation is \(n \times m = 15\). The three discrete gaps are \(g_1 = 2\ \mu m\), \(g_2 = 2.5\ \mu m\), and \(g_3 = 3\ \mu m\), with the minimum and maximum \(\alpha_i\)'s given by: \([\alpha_{1, min}, \alpha_{1, max}] = [0.34, 0.85]\), \([\alpha_{2, min}, \alpha_{2, max}] = [0.34, 0.88]\), and \([\alpha_{3, min}, \alpha_{3, max}] = [0.34, 0.90]\). The nominal plant is chosen with \(g = g_3\), and \(\alpha = \alpha_{31} = 0.34\). We substitute \(\lambda = 1.3\) into (4.89) and (4.90) to get a minimum phase
1. Discretize frequency $\omega$ into a finite set: $\Omega = \{\omega_i | i = 1, \ldots, n\}$.

2. Select a nominal plant $P_o(s) \in \mathcal{G} = \{P_1(s), \ldots, P_m(s)\}$, with $P_o(j\omega) = p_o e^{j\theta_o}$.

3. Define a phase range $\Phi$, and discretize it for the controller phase $\varphi_i$.

4. Choose a single frequency $\omega \in \Omega$.

5. Choose a single controller’s phase $\varphi_i \in \Phi$.

6. Choose a single plant $P_k(s) \in \mathcal{G}$.

7. Compute $g_{\text{max}}(P_k)$ and $g_{\text{min}}(P_k)$ using (4.92) and (4.93).

8. Set $g_{\text{max}}(\varphi) = \max\{g_{\text{max}}(P_k)\}$ and $g_{\text{min}}(\varphi) = \min\{g_{\text{min}}(P_k)\}$.

9. Set $L_{o,\text{max}}(\psi_o) = p_o g_{\text{max}}(\varphi)$ and $L_{o,\text{min}}(\psi_o) = p_o g_{\text{min}}(\varphi)$, with $\psi_o = \varphi + \theta_o$.

Figure 4.25 Flow chart for computing QFT bounds [82].
margin and an upper gain margin of $45^\circ$ and 4.95 dB. At selected frequencies $\omega = 10, 1000, 10000, 50000, 100000$, and $500000$, the magnitude and phase of the linearized plant templates are plotted on the Nichols chart shown in Figure 4.27(a). The generated stability bounds and the stability line for the nominal unstable plant is illustrated on the Nichols chart in Figure 4.27(b).

4.4.2 Modification of Bounds and Loop Shaping

For loop-shaping with an unstable and/or non-minimum phase nominal plant, the bounds must be shifted to convert the loop-shaping problem to that for a stable minimum phase nominal plant [83][84]. The reason is that it is more convenient to work with a minimum phase function in numerical design because the Bode integral can be used, and the optimal loop shaping can be derived [81].

Given a uncertain unstable/non-minimum-phase plant,
where \( \hat{N}(s) \) and \( \hat{D}(s) \) denote the parts with right half-plane poles and zeros. An arbitrary selected nominal plant from the plant family is given by,

\[
P(s) = \frac{\hat{N}(s)N(-s)}{\hat{D}(s)D(-s)}
\]
We define an all-pass filter containing the right-half plane poles and zeros as,

$$P_o(s) = \frac{N_o(s)N_o(-s)}{D_o(s)D_o(-s)}$$  \hspace{1cm} (4.96)

which relates the new nominal stable/minimum-phase plant to the original nominal plant by,

$$A(s) = \frac{N_o(-s)D_o(s)}{N_o(s)D_o(-s)}$$  \hspace{1cm} (4.97)

Thus the loop transmission $L(s)$ is related to the new loop transmission $L'(s)$ by,

$$L(s) = C(s)P(s) = C(s)P'(s)A(s) = L'(s)A(s)$$  \hspace{1cm} (4.99)

The bounds imposing on $L'(s)$ are the same in magnitude as the bounds imposing on $L(s)$ yet with a horizontal phase shift of $-\angle A(j\omega)$ \[83\]. In addition, the stability line defined on the Nyquist plot and the Nichols chart in Section 4.2.2.1 becomes,

$$R'_o(\omega) = \{(x, y) | x= -(k + 1) \cos(\angle(-A(j\omega))), \ y= -(k + 1) \sin(\angle(-A(j\omega))), \ k > 0 \}$$  \hspace{1cm} (4.100)

and

$$R'_1(\omega) = \{ (\phi, r) | \phi= -(2q + 1)180^\circ + \angle A^{-1}(j\omega), \ r > 0, \ q= \pm 1, \pm 2, \ldots \}$$  \hspace{1cm} (4.101)

respectively, as shown in Figure 4.28. Due to the shift of those stability lines, the stability criterion is modified as given in Theorem 4.3 [84].

**Theorem 4.3** For the plant with \( n \) right half plane poles, the closed-loop system is stable if and only if \( L'(j\omega) \) as defined in (4.99) does not intersect the point \(-A(j\omega)\) at the same frequency \( \omega \), and the net number of positive and negative crossings on the stability line \( R'_1(\omega) \) at the same frequency is \( n \).

The nominal unstable plant is modified to a nominal stable plant using (4.98).

The final bounds, including the stability lines, after being shifted by \( \angle A^{-1}(j\omega) \), are shown in Figure 4.29.

The loop shaping of \( L_o'(j\omega) \) is in Figure 4.30, which shows that derived bounds are not violated by \( L_o'(j\omega) \) at \( \omega = \omega_p \), indicating that the desired phase margin and upper gain margin are satisfied by design. The designed proportional-gain controller is \( k_p = 30 \),
which is larger than previous design ($k_p = 16$) for a single nonlinear plant in Section 4.3, because additional gain-bandwidth must be paid for sensitivity reduction of plant uncertainties ($g = 2$ to $3 \mu m$ instead of $3 \mu m$ only).

Figure 4.29 Robust stability bounds and stability lines derived for the new stable nominal plant.

Figure 4.30 Shaping of nominal open-loop function based on the derived bounds, such that $L_o'(j\omega)$ does not fall within the area as indicated in Figure 4.26. ($k_p = 30$)
4.4.3 Plant Input Disturbance Rejection

For plant input-disturbance rejection, the plant templates and bound generation are constructed point-by-point for a total of \( n \times m \) times. The final bounds, \( B_D \), are the intersection of the total \( n \times m \) bounds. Each set of bounds \( B_{d,ij} \) is scaled by the ratio, \( 2kZ_{o,ij}/V_{o,ij} \), over that of the selected nominal plant, \( 2kZ_{o,i,j_o}/V_{o,i,j_o} \), so all the bounds are referred as for the nominal plant. \( Z_{o,i,j_o} \) and \( V_{o,i,j_o} \) are the operating displacement and the operating voltage of the nominal plant selected at \( g = g_3 \), and \( \alpha = \alpha_{35} = 0.9 \). \( Z_{o,ij} \) and \( V_{o,ij} \) are the operating displacement and the operating voltage of the other plants. The final bounds are given by,

\[
B_D = \bigcap_{i = 1}^{n} \bigcap_{j = 1}^{m} \left( \frac{B_{d,ij}}{2kZ_{o,ij}/V_{o,ij}} \right) = \bigcap_{i = 1}^{n} \bigcap_{j = 1}^{m} \left( \frac{B_{d,ij}V_{o,ij}Z_{o,i,j_o}}{2kZ_{o,i,j_o}/V_{o,i,j_o}} \right)
\]

(4.102)

We substitute \( r_v = 0.1 \) and \( r_z = 0.075 \) into (4.69) to (4.71), and obtain the disturbance force, \( F_{d,ij} \), at all displacements, \( \alpha_{ij} \), as multiples of the spring restoring force, \( D_{ij}kZ_{o,ij} \).

We set the ratio, \( \Delta z_{ij}/Z_{o,ij} \), of the final perturbed displacement over the steady-state displacement to 0.01. Substituting the capacitive sensor gain given by (3.43) into (4.75) yields the sensitivity-reduction specification,

\[
\frac{\Delta v_{s,ij}}{F_{d,ij}} = \frac{G(Z_{o,ij})\Delta z_{ij}}{D_{ij}kZ_{o,ij}}
\]

(4.103)
where $\Delta v_{s, ij}$ is the final perturbed sensed voltage. The final bounds on the Nichols chart are illustrated in Figure 4.31 for frequencies up to 100,000 rad/sec. Also shown is the nominal loop-shaping using a minimum controller gain, $k_p = 106$ to satisfy the bounds. This large $k_p$ is similar to previous results in Figure 4.20.

### 4.5 Scaling of Plant Parameters

Control design procedures for a given fixed plant and a given plant with initial gap variation have been introduced in Section 4.3 and 4.4. As design guidance for both plant and controller, it will be beneficial to understand how the scaling of plant parameters, such as mass, spring, and plate area, affects the controller design, in order to provide directions for achieving a good plant design. A good plant design for parallel-plate servo should help to reduce the cost of feedback, namely, the open-loop bandwidth, in dealing with the unstable pole after the pull-in limit. Implemented gain-bandwidth product of the servo and sensing circuits can therefore be smaller to reduce power consumption.
The rule of thumb discussed in Section 4.2.2.2 and 4.2.2.4 states that the open-loop unity-gain frequency should be at least twice the unstable pole frequency in order to achieve a 60° phase margin. By making the unstable pole close to the origin on the complex plane, the resulting open-loop unity-gain frequency is reduced, and so is the circuit gain-bandwidth product required to be implemented. Implemented circuits within the feedback loop include the controller, pre-amp, demodulator, and subtracter. The linearized actuator d.c. gain ranges from 10^{-6} to 10^{-7} (unit = m/V), and the capacitive sensor gain is in the range of 10^{-4} (unit = V/m). Implemented d.c. voltage gain of control and sensing circuits can be in the range of a few hundreds to several thousands. If the open-loop unity-gain frequency is 20 kHz and the d.c. circuit gain is 1000, for example, then the total circuit gain-bandwidth product should be at least 20 MHz to ensure that the dominant pole from circuits does not affect phase margin at the open-loop unity-gain frequency.

From (4.47), (4.48), and (4.50), the right-half-plane unstable pole is given by,

\[
\omega_{rhp} = \left(\frac{\xi^2}{\frac{3\alpha - 1}{1 - \alpha} - \xi}\right) \omega_n
\]

\[\quad (4.104)\]

To reduce the unstable pole frequency, the damping ratio \(\xi\) can be increased and/or the actuator resonant frequency can be decreased (i.e., decrease of spring constant and/or increase of mass). Rise time of the output response will increase due to the use of either approach. Increase of damping ratio is preferred at larger displacements when the spring-softening effect is more significant. However the rise time can be negatively impacted with a large damping ratio at small displacements. Assume that the increase of plate area, \(A\), is proportional to the increase of mass, total linearized d.c. gain of the parallel-plate actuator and the capacitive sensor is proportional to:
according to (4.45), (3.8), and (3.43). The implemented circuit gain-bandwidth product can therefore be reduced by $A^{3/2}$ if the open-loop gain remains the same. An easy way to reduce the gain-bandwidth product without involving actuator design is to use a large modulation voltage $V_m$ to improve the capacitive sensitivity. The drawback is that $V_m$ contributes an effective d.c. actuation voltage which limits the maximum controlled displacement. Also, there is a capacitive feedthrough issue in which carrier-frequency signals sitting on the actuation voltage pass to the sensing plates, and are demodulated back to a d.c. sensed voltage which eventually adds to the actuation voltage. More discussion of capacitive feedthrough will be presented in Chapter 5.

Phase margin is negatively impacted by the reduction of damping ratio because of the increase of unstable pole frequency. By using the plant parameters given in Chapter 3 except a fixed damping ratio ($\xi = 0.3$ to 1.5) is used for all displacements, an optimal time-varying proportional-gain controller is designed with the resulting phase margin, the corresponding lower and upper gain margins, the crossover frequency, and the controller gain shown from Figure 4.32(a) to (e). Phase margin decreases monotonically with increasing displacement due to the use of a fixed damping ratio for all displacements. Damping in air is not fixed, however these plots show the trend for decreasing damping. The system can not be stabilized over the entire gap at underdamped and slightly overdamped situations unless phase-lead compensation is employed. Although phase margin can be attained with phase-lead compensation, the resultant large open-loop bandwidth

$$G_{act}G_s \propto \frac{1}{V_{pi}} \cdot A$$

$$= A^{3/2}$$
Figure 4.32 Maximum phase margin, and corresponding crossover frequency, lower gain margin, upper gain margin and controller gain plotted as a function of normalized displacement from (a) to (e).
includes more sensor noise, and the disturbance in the expanded bandwidth is likely to excite the higher mechanical modes.

To summarize, the scaling of plant parameter should be directed towards the reduction of cost of feedback, and thereby the implemented gain-bandwidth product of control and sensing circuits. Gain of the circuits can be reduced with increasing actuator gain and capacitive sensor gain resulting from a larger actuation/sensing plate area. Bandwidth of the circuits can be reduced by decreasing the unstable pole frequency resulting from a larger damping ratio and/or a smaller mechanical bandwidth.
Chapter 5
Experimental Results

5.1 Introduction

In this section, we present the experimental results on plant characterization and closed-loop feedback control. We will first discuss the interferometric setup for static measurement. Then curl measurement of the microactuator is presented, followed by static displacement-voltage measurement using capacitive sensing to extract the spring constant of the actuator and the total circuit input capacitance. We discuss capacitive feedthrough and its removal by high-pass filtering. Measured frequency response of the actuator is verified with NODAS simulation to extract the mechanical resonant frequency and the damping ratios at different gaps. Measured transient response is presented to characterize the risetime and falltime of the actuator. We have measured the minimum input-referred noise voltage to determine the minimum input-referred noise displacement at 0.33 nm/√Hz. Last, experimental results of closed-loop step response are presented, which illustrate a maximum plate displacement up to 60% of the gap.

5.2 Experimental Setup

We have built a custom micropositioning assembly with interferometric measuring capability on an air table, as shown by the photograph in Figure 5.1. The complete servo system is realized on a printed-circuit board (PCB), which consists of a 40-pin dual inline package (DIP) with the attached tip-actuator die, and external control and demodulation circuits. The PCB is bolted to a six-degree-freedom micropositioner with a resolution of 1 µm for coarse control of initial gap spacing. The electrode placed on top of the actuator is made of a thin glass piece coated with a thin gold layer about a few hundred nanometers. The glass piece is attached to a piezo-actuator (pz 38 from piezosystem jena [85]) with a resolution of 0.2375 µm/V for fine control of initial gap spacing. A schematic
representation of the interferometric setup in Figure 5.2 illustrates that fringe patterns are formed by light beams reflected from the gold layer and the actuator, and the image is collected using a charge-coupled-device (CCD) camera. An orange/red light-emitting diode (LED) ($\lambda = 635$ nm) is used as the light source. Each fringe occurs at multiples of a half wavelength of the LED. It is necessary to ensure the parallelism between the actuator and the top electrode before placing them closely. We shine a He-Ne laser beam on both of
them and perform alignment of reflected beam spots on a white wall 1.1 meters away using the six-degree-of-freedom micropositioner. After alignment, we use the six-degree-of-freedom micropositioner for coarse-positioning by approaching the die to the top electrode until about 10 µm apart where fringe patterns begin to show up. Then we can combine the use of the micropositioner and piezo-actuator for fine control of gap. The piezo-actuator is only used for setting up the initial gap and adjusting gap spacing with its sub-µm accuracy. No a.c. driving of the piezo-actuator is required in our static/dynamic measurements. The alignment process is integrated conveniently with the interferometric setup as shown in Figure 5.2, in which the CCD camera and the objective lens have to be dismounted to pass the laser beam onto the glass piece and the test die.

5.3 Curl Measurement

The microactuator fabricated by the AMS (Austria Micro System) 0.5 µm CMOS process after post-CMOS process is shown in Figure 5.3(a). Structural curl of the actuator is measured by the WYCO NT3300 white-light interferometric profiler, as shown in Figure 5.3(b), where the spikes on the surrounding metal-3 surface is a measuring artifact due to the slotted holes required by design rules.

A comprehensive curl measurement over ten dies has been performed to provide information on curl variation for building an array of tip actuators. A side view schematic of the actuator is depicted in Figure 5.4, in which the actuation-plate curl height is defined $h_{at}$, height from center of the actuation plate to center of the sensing plate 1 is $h_{asl}$, height
from center of the actuation plate to center of the sensing plate 2 is $h_{as2}$, the actuation-plate tilting angle is $\theta_a$, the sensing-plate-1 tilting angle is $\theta_{s1}$, and the sensing-plate-2 tilting angle is $\theta_{s2}$. The angles are defined from the curl heights of plates at edges with respect to plate centers. Measured results are tabulated in Table 5.1, in which a maximum

<table>
<thead>
<tr>
<th>Die</th>
<th>$h_a$ (µm)</th>
<th>$h_{as1}$ (µm)</th>
<th>$h_{as2}$ (µm)</th>
<th>$\theta_a$ (deg.)</th>
<th>$\theta_{s1}$ (deg.)</th>
<th>$\theta_{s2}$ (deg.)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.9</td>
<td>0.8</td>
<td>0.7</td>
<td>2.58</td>
<td>1.43</td>
<td>-2.86</td>
</tr>
<tr>
<td>2</td>
<td>1.3</td>
<td>1.1</td>
<td>1.1</td>
<td>2.29</td>
<td>1.43</td>
<td>-2.58</td>
</tr>
<tr>
<td>3</td>
<td>1.2</td>
<td>1.0</td>
<td>1.0</td>
<td>2.86</td>
<td>1.43</td>
<td>-2.86</td>
</tr>
<tr>
<td>4</td>
<td>1.2</td>
<td>0.9</td>
<td>0.9</td>
<td>2.58</td>
<td>1.72</td>
<td>-2.86</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>0.7</td>
<td>0.5</td>
<td>2.15</td>
<td>1.43</td>
<td>-2.86</td>
</tr>
<tr>
<td>6</td>
<td>0.8</td>
<td>1.0</td>
<td>0.6</td>
<td>2.29</td>
<td>1.43</td>
<td>-2.86</td>
</tr>
<tr>
<td>7</td>
<td>1.3</td>
<td>1.0</td>
<td>1.0</td>
<td>2.29</td>
<td>1.72</td>
<td>-2.86</td>
</tr>
<tr>
<td>8</td>
<td>1.6</td>
<td>1.1</td>
<td>1.3</td>
<td>2.86</td>
<td>2.00</td>
<td>-2.86</td>
</tr>
<tr>
<td>9</td>
<td>1.8</td>
<td>1.1</td>
<td>1.0</td>
<td>2.86</td>
<td>1.72</td>
<td>-3.43</td>
</tr>
<tr>
<td>10</td>
<td>1.8</td>
<td>1.2</td>
<td>1.0</td>
<td>3.72</td>
<td>1.72</td>
<td>-3.58</td>
</tr>
<tr>
<td></td>
<td>Average (µm, deg)</td>
<td>1.19</td>
<td>0.99</td>
<td>0.91</td>
<td>2.65</td>
<td>1.60</td>
</tr>
<tr>
<td></td>
<td>Std. Dev. (µm, deg.)</td>
<td>0.54</td>
<td>0.15</td>
<td>0.24</td>
<td>0.46</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Table 5.1: Measured curl and plate tilt of the tip actuator.

difference of 0.9 µm in $h_a$ is illustrated. Large $\theta_{s1}$, $\theta_{s2}$, $h_{as1}$, $h_{as2}$ can reduce damping coefficient compared to a perfect parallel-plate actuator, resulting in phase-margin reduction. The AMS process does not perform Chemical Mechanical Polishing (CMP) after thin-film deposition, which can be a factor affecting the curl variation.

Figure 5.4  Side view schematic of the actuator.
5.4 Static Characterization

To measure the z-directional spring constant and the pre-amp input capacitance, we have performed a series of static displacement-voltage tests at various initial gaps using capacitive sensing. Plate displacement is determined by the fringe displacement with a measuring error around ±10 nm due to the image contrast of fringe patterns. A d.c. voltage is incrementally applied to the actuator, and an a.c. modulation voltage of 1 V amplitude at 4.5 MHz is applied to the top electrode for capacitive sensing. Measurement is performed within the pull-in limit to prevent the actuator from snapping onto the top electrode.

Actuation and sensing capacitances are modelled using the Maxwell™ simulation to account for the plate tilt and the fringing field due to the finite metal thickness. Capacitance polynomials using least-squares fit are used to replace the simple parallel-plate capacitance model in our analysis. The root-mean-square voltage contributed from the a.c. modulation voltage is included in the d.c. driving voltage. The measured and simulated displacement-voltage characteristics shown in Figure 5.5 extract the spring constant of the actuator at 0.17 N/m, as compared to the analytical value of 0.167 N/m. The pre-
amp input capacitance, $C_i$, is 333 fF extracted from the measured pre-amp output amplitude with respect to the actuator displacement in Figure 5.6. Measured value is smaller than the extracted capacitance, 408 fF, from layout.

### 5.5 Capacitive Feedthrough

The capacitance feedthrough from the electrostatic actuator and the capacitive position sensor gives rise to incorrect sensor outputs after demodulation. As depicted in Figure 5.7, the baseband actuation voltage, $V_a$, is coupled to the pre-amp output via a feedthrough capacitance, $C_f$, which consists of the fringing capacitance, $C_{f1}$, between the crossed actuation and sensing interconnects (a ground shield is inserted in between), and another fringing capacitance, $C_{f2}$, between the top metal layers of the actuation and sensing plates. The total of $C_f$ is 1.38 fF, with $C_{f1} = 0.42$ fF and $C_{f2} = 0.96$ fF obtained by Maxwell™ simulation. After the pre-amp, the baseband feedthrough subsequently feeds through the demodulator with a gain, $A_{df}$, to produce a sensed output error, $V_f$. 

Figure 5.6 Characteristic of pre-amp output versus plate displacement results in an extracted pre-amp input capacitance of 333 fF.
We measure the feedthrough transfer function, $V_f(s)/V_a(s)$, using the Agilent 4395A spectrum/network analyzer by applying a sinusoidal voltage to the actuation plate without a top electrode. The measured feedthrough gain is -35.9 dB, with -26.4 dB measured after the pre-amp, and -9.5 dB from the pre-amp to the demodulator. Substituting measured values and $C_i = 333 \text{ fF}$ into (5.1) gives $C_f = 1.6 \text{ fF}$. For a 10 V actuation voltage, it produces a significant feedthrough output of 160 mV, which is equivalent to a sensed voltage produced by a 0.7-µm plate displacement at a 3-µm initial gap.

To suppress the capacitive feedthrough, we perform successive high-pass filtering in the pre-amp and right after the pre-amp, as shown in Figure 5.7, to reduce the feedthrough gain to -81 dB. The resistors and capacitors used for high-pass filtering are: $C_1 = 1 \text{ nF}$, $C_2 = 390 \text{ pF}$, $R_1 = 1.52 \text{ k} \Omega$, $R_2 = 13.65 \text{ k} \Omega$, and $R_3 = 3 \text{ k} \Omega$.

5.6 Open-Loop Frequency Response

The actuator frequency response of the actuator in Figure 5.8 is measured electronically by the network analyzer, which supplies an a.c. voltage amplitude of 700 mV.
with incrementally increased d.c. bias from 0 to 4 V to the actuator, and gathers readout signal after demodulation. At a frequency around 30 kHz, the measured gain magnitude starts to increase because of the capacitive feedthrough. Measured resonant frequency decreases with respect to the increasing d.c. bias due to the spring-softening effect. The root-mean-square voltage contributed from the a.c. modulation voltage of 1 V amplitude is included in the d.c. driving voltage for analysis. Measured resonant-frequency shift and NODAS simulation results are compared in Figure 5.9 for initial gaps of 2.3 µm and 2.65 µm, in which more frequency drop is observed for the smaller initial gap at the same d.c. actuation voltage. The original actuator resonant frequency, \( \omega_n \), is found at 12.4 kHz from the plot.

The effective damping ratio, \( \xi_e \), given in (4.48), is obtained by comparing the resonant peak magnitude of the measured frequency response in Figure 5.10(a) with that of a second-order system transfer function. \( \xi_e \) is between 0.55 to 0.35 for gaps between 2.3 to 2.65 µm as shown in Figure 5.10(b). Also shown in Figure 5.10(a) and (b) are results from the NODAS simulation, in which the squeeze-film damping model and the electrostatic gap model are used for frequency-response simulation. We have considered structural curl

Figure 5.8  Measured frequency response of the actuator driven at different d.c. bias, illustrating the resonant frequency shift due to the spring-softening effect.
by using different gaps for the actuation and sensing plates, and a ±5% gap variation is used to account for the plate tilt.

5.7 Open-Loop Step Response

We apply a 2 kHz square-wave voltage with amplitudes ranging from 1 V to 4.8 V (minimum = 0 V) to the actuator with an initial gap spacing of 2.45 µm. The sensed signal is gathered after demodulation and amplification by a gain of ten, and displayed on an Agilent 54624A digital oscilloscope averaged over 128 sweeps to remove most of the 1× and 2× carrier-frequency signals. Measured demodulator output voltage waveforms are shown in Figure 5.11 from which risetime (0 to 100%) and falltime (100 to 0%) are measured and tabulated in Table 5.2. The risetime increases with the closing of gap due to the increasing damping coefficient, and spring-softening-induced bandwidth reduction. The actuator is more lightly damped when moving back to the initial position as evidenced by the measured larger overshoot, shorter falltime, and longer settling time in Figure 5.11.

<table>
<thead>
<tr>
<th>$V_{\text{act}}$ (V)</th>
<th>Risetime ($\mu$s)</th>
<th>Falltime ($\mu$s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>29.0</td>
<td>28.0</td>
</tr>
<tr>
<td>2.0</td>
<td>29.2</td>
<td>28.5</td>
</tr>
<tr>
<td>2.8</td>
<td>30.0</td>
<td>28.7</td>
</tr>
<tr>
<td>3.6</td>
<td>30.6</td>
<td>28.8</td>
</tr>
<tr>
<td>4.2</td>
<td>32.0</td>
<td>28.8</td>
</tr>
<tr>
<td>4.8</td>
<td>33.3</td>
<td>28.8</td>
</tr>
</tbody>
</table>

Table 5.2: Open-loop risetime and fall measurements (gap = 2.45 µm).

![Figure 5.9 Measured resonant frequency versus d.c. actuation voltage.](image)
Figure 5.10 (a) Measured resonant peak and corresponding NODAS simulation results plotted as a function of d.c. actuation voltage in the frequency-response measurement. (b) Extracted effective damping ratio and corresponding NODAS simulation results plotted as a function of d.c. actuation voltage.
Another measurement of step-response waveforms performed at an initial gap of 3.2 \( \mu \text{m} \) with actuation voltage ranging from 1 V to 7 V is shown in Figure 5.12, and the measured risetime and falltime are tabulated in Table 5.3. From Table 5.2 and Table 5.3, we observe that the risetime does display an increasing trend with respect to the actuation

Figure 5.11 Open-loop responses to 2 kHz square-wave actuation force varying in applied voltage amplitude from 1 V to 4.8 V (gap = 2.45 \( \mu \text{m} \)).

Figure 5.12 Open-loop responses to 2 kHz square-wave actuation force varying in applied voltage amplitude from 1 V to 7 V (gap = 3.2 \( \mu \text{m} \)).
voltage, yet not a very definite trend for the falltime. Also, both of the risetime and fall-
time decrease as the initial gap is widened due to the decreasing damping coefficient and spring-softening effect.

<table>
<thead>
<tr>
<th>$V_{\text{act}}$ (V)</th>
<th>Risetime (µs)</th>
<th>Falltime (µs)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>26.3</td>
<td>25.2</td>
</tr>
<tr>
<td>2</td>
<td>26.5</td>
<td>25.5</td>
</tr>
<tr>
<td>3</td>
<td>26.8</td>
<td>26.0</td>
</tr>
<tr>
<td>4</td>
<td>27.0</td>
<td>25.5</td>
</tr>
<tr>
<td>5</td>
<td>27.5</td>
<td>25.8</td>
</tr>
<tr>
<td>6</td>
<td>27.6</td>
<td>25.5</td>
</tr>
<tr>
<td>7</td>
<td>28.6</td>
<td>25.8</td>
</tr>
</tbody>
</table>

Table 5.3: Open-loop risetime and fall measurements (gap = 3.2 µm).

5.8 Input-Referenced Noise Displacement

We apply a 2 kHz square-wave voltage with a 0.2 V amplitude (minimum = 0 V) to the actuator separated by 2.65 µm to the top electrode. The modulation voltage on the top electrode for capacitive sensing has an amplitude of 1 V at 4 MHz. The gap spacing is incrementally increased by pulling the top electrode away from the actuator using the piezo-actuator. At each displacement, we measure the sensed signal from the pre-amp output at 4 MHz ± 2 kHz using the spectrum analyzer. The measured spectrum at a 1 Hz resolution bandwidth is in Figure 5.13(a), in which the frequency component at 4 MHz is the initial sensed voltage magnitude of 490 mV, and the spikes on the spectrum are contributed from the 60 Hz feedthrough.

Measured output voltage at 1 Hz resolution bandwidth with respect to gap spacing is scaled back to the pre-amp input by the pre-amp gain, $A_v = 7.76$ at 4 MHz, as shown in Figure 5.13(b). The minimum input-referred noise voltage is measured around 1.4 µV/√Hz at a gap of 5.7 µm. The equivalent capacitance change at 1 Hz bandwidth corresponding to the measured voltage output is plotted in Figure 5.13(c), which illustrates a minimum input-referred noise capacitance change around 0.5 aF/√Hz, equivalent to a minimum input-referred noise displacement of 0.33 nm/√Hz. Measured input-referred
Figure 5.13  (a) Pre-amp output spectrum measurement to a 2 kHz square-wave actuation voltage with an amplitude of 0.2 V. (b) Measured input-referred sensed voltage as a function of actuation gap. (c) Equivalent input-referred capacitance change as a function of actuation gap.
noise capacitance change and input-referred noise displacement at \( g = 5.7 \mu m \) are primarily contributed from the thermal noise of d.c.-bias resistor in front of the pre-amp, as depicted in Figure 3.33(a) and (b).

5.9 Closed-Loop Step Response

The photograph in Figure 5.14 illustrates the PCB with the tip-actuator package and the off-chip feedback control circuit. The final controller design is \( C(s) = 24.3 \), and the pre-filter design is \( F(s) = 1/(s/1800 + 1) \). The controller gain is higher than the previously designed \( k_p = 16 \) primarily because of the pre-amp gain \( A_v = 7.76 \) at the modulation frequency 4.5 MHz, instead of \( A_v = 10 \). We apply a 100 Hz square-wave voltage as the input command for the closed-loop response measurement. Before the command is applied, an initial d.c. sensed voltage output (103 mV at \( g = 3.3 \mu m \)) is nulled to zero, and the amplitude of the \( 1 \times \) and \( 2 \times \) carrier-frequency signals \( (f_c = 4.5 \text{ MHz}) \) at the controller output is reduced to within 1 V by low-pass filtering \( (\omega_p = 140 \text{ kHz}) \).

Measured controller output waveforms and demodulated pre-amp output waveforms as the actuated plate moves across and beyond the pull-in instability are averaged over 128 sweeps on the oscilloscope, as shown in Figure 5.15(a) and Figure 5.15(b). The relation of plate displacement versus input command voltage amplitude is depicted in Figure 5.15(c). The decrease of controller output after the pull-in voltage \( (V_{pi} = 10.43 \text{ V}) \)
Figure 5.15 (a) Measured controller output waveforms when displaced plate enters pull-in region and beyond. The inset illustrates its waveform when the input command turns from high to low. (b) Measured demodulated pre-amp output waveforms. (c) Plate displacement versus input command voltage. (gap = 3.3 µm)
reduces the stored charge on the plate in order to stabilize the actuator in the open-loop unstable regime. As the input command changes from high to low, the controller output first decreases, and then increases as the actuator retracts back to the rest position to maintain stability, as shown by the inset in Figure 5.15(a). The peak magnitude shown in the inset increases as the damping coefficient decreases. The rise time of the plate displacement increases with increased input command magnitude from 2 ms to 5 ms as illustrated by the demodulated pre-amp output waveforms in Figure 5.15(b). Therefore we have to decrease the input-command frequency from 100 Hz to 50 Hz for large input commands.

Measured controller output waveforms and demodulated pre-amp output waveforms are compared with closed-loop NODAS simulations, in which we model the parallel-plate actuator using 3D elastic beam elements, 3D elastic plate elements, squeeze-film damping elements, and electrostatic gap models as shown in the schematic in Figure 5.16. The actual gap values used in the squeeze-film damping models are increased by 1.2 µm to account for the reduced damping due to plate tilt. The measured and simulated pre-filter output, controller output, demodulated pre-amp output, and plate displacement to a 50 Hz square-wave input command are plotted in Figure 5.17 with very good agreement. A maximum displacement of 2 µm, equivalent to 60% of the gap (g = 3.3 µm) is illustrated. Measured falltime of the demodulated pre-amp output waveform is significantly shorter than its risetime because the actuator dynamics is different in the falling period, in which the spring restoring force, not the electrostatic force, provides acceleration for the actuator when it retracts back to the rest position. Also the experienced damping force reduces as the actuator moves away from the top electrode.

5.10 Discussion

Maximum travel range of 60% of the gap is experimentally demonstrated by closed-loop servo. Critical factors which contribute to the failure of the closed-loop servo beyond 60% of the gap are:

1. Reduced phase margins resulted from an reduced damping coefficient due to plate tilt.

The measured controller output waveform during the input-command transition is related the damping coefficient at that instant. We match the waveform in NODAS
Figure 5.16 Schematic representation of parallel-plate actuator design for closed-loop NODAS simulation.
Figure 5.17 Measured waveforms (dashed lines) and simulated waveforms (solid lines) of pre-filter output, controller output, demodulated pre-amp output, and plate displacement to a 50 Hz square-wave input-command voltage.
simulation by increasing the gap spacing of squeeze-film damping elements to accommodate for the plate tilt. The damping ratio as a function of normalized displacement is plotted in Figure 5.18(a) for both tilted and parallel plates. The resulting phase margin versus normalized displacement is shown in Figure 5.18(b), which illustrates a significant phase margin loss due to plate tilt as compared to the parallel-plate case. The phase margin at 60% of the gap is only 11°.

2) **Sensor noise amplification at the controller output.** As noted in Section 4.2.1, a large loop gain at frequencies for which the plant gain is small results in large noise amplification at the controller output, and probable system instability. The situation of noise amplification is worsened when the d.c gain of the actuator reduces as the plate displacement increases as shown in Figure 4.8, leading to noise amplification at low frequencies. We substitute the linearized plant at 60% of the gap (2-μm displacement) and the controller into (4.13), and plot the noise gain from the capacitive sensor output to the controller output in Figure 5.19. The plot includes noise gains for displacements of 1.8 μm and 1.6 μm, demonstrating that as the placement increases, both the noise gain at low frequencies and its peak value increase. A noise spectral density of 70 μV/√Hz from the capacitive sensor output is amplified by 44 to 51 dB for frequencies up to 100 kHz at the controller output. Therefore system stability can be seriously affected.

3) **Changing sensor output due to thermally-induced plate displacement, charging effect, and piezo-actuator drift.** The actuator is constantly illuminated by a LED, which produces a 4 °F difference measured between the actuator and the ambient temperatures. We have observed up to 300 nm plate displacement due to the heating effect. Charging may occur at the high-impedance sensing node of the pre-amp due to the charge in the surrounding dielectric layer and air. Due to the fact that there is no feedback loop around the piezo-actuator, drift up to 100 nm can happen to affect the preset gap spacing. All those effects result in a d.c. offset at the capacitive sensor output, leading to a d.c. offset about 1 to 2 V at the controller output during the experiment. This offset cannot be nullled immediately when it happens, therefore system stability may be affected.
Figure 5.18 (a) Damping ratio derived from tilted plates and parallel plates as a function of normalized displacement. (b) Phase margin derived tilted plates and parallel plates as a function of normalized displacement.
Figure 5.19 Noise gain from sensor output to controller output plotted for linearized plants at different displacements.
Chapter 6

Conclusions

In this thesis, we have presented a controller design procedure which effectively resolves the nonlinear and unstable characteristics of the parallel-plate actuator, and leads to a linear controller design suitable for implementation. We are the first to report successful experimental results of closed-loop voltage control that extend the parallel-plate displacement to 60% of the initial gap, almost doubling the maximum displacement limited by the pull-in instability. For the proposed probe-based micro disk drives, we prefer to achieve a displacement up to 75 to 83% of the initial gap ($g = 2$ to $3\,\mu m$) to accommodate a tip height of 500 nm. Given the achieved 60% travel range, an increase of tip height to $1.2\,\mu m$ is suggested by stacking up the thin films underneath the tip.

The CMOS-MEMS process developed at Carnegie Mellon is the best to date enabling technology for implementation of the proposed micro disk drive. Fabrication of the tip actuator array can benefit from the convenient circuit integration to get high-speed electronics, and the use of existing multiple interconnects within mechanical structures for signal routing. To provide material properties for mechanical design, we have characterized Young’s modulus, stress gradient, and residual stress of various CMOS-MEMS structures fabricated from different Agilent 0.5 $\mu m$ CMOS runs. The commonly used m3-m2-m1 beam has an average Young’s modulus of 62 GPa, an average radius curvature of 1.37 mm, and an average residual stress of 42.8 MPa.

Controller design is through linearization of the nonlinear plant. Local stability of the system is analyzed using gain margin and phase margin. To extend from local stability to global stability in a nonlinear control problem, we perform input-disturbance rejection to reduce the effects from initial conditions and high-order electrostatic force terms neglected during plant linearization. There is a design trade-off at large displacements where a large controller gain is desired for disturbance rejection, yet the large gain decreases the phase margin at small and median displacements, and induces more noise amplification at the controller output. Both of them can lead to system instability.
We have shown that selection of the two degree-of-freedom control system configuration (controller and pre-filter) is crucial as it decouples the design processes for closed-loop stability and output tracking. The controller ensures system stability in the presence of the unstable pole, and the pre-filter shapes the input command for the feedback loop to follow. The selected pre-filter design is based on a desired output risetime of 2 to 4 ms. Without the pre-filter, the resultant output may exhibit a large overshoot according to our design experience.

As mentioned in Chapter 4, to maintain a 60° phase margin at an unstable operating point, the minimum open-loop bandwidth should be at least twice of the unstable pole frequency. Frequency of the unstable pole increases with increasing plate displacement and decreasing damping ratio. If the system were to be operated at a low pressure for application of thermally-assisted writing using tunneling current, the reduced damping ratio will require a large open-loop bandwidth for system stabilization, subsequently resulting in excessive noise amplification at the controller output, and/or excitation of higher mechanical modes to destabilize the system. Furthermore, the implemented open-loop bandwidth is limited by the use of low-pass filtering after the demodulator, thereby the achieved phase margin is restricted at large displacements and low operating pressures, and design of a stable controller is even more difficult.

Steady-state error resulting from the use of a proportional-gain controller can be eliminated by adding an integral controller. However, the 90° phase lag introduced by an integrator has to be compensated by adding a zero at a very low frequency (tens of Hertz); otherwise the resultant phase margin can be significantly reduced. Special circuit design techniques are required to realize the low-frequency zero on chip with die area taken into consideration. The implementation of an integrator requires a reset scheme for cancelling the integral of circuit non-idealities, such as input offset voltage.

Experimental closed-loop responses have very good agreement with the NODAS simulations in which 3D plates, 3D beams, squeeze-film damping elements, and electrostatic gap models are used to represent the actuator. Reduced phase margin and increase in loop noise bandwidth are the key factors which are believed to contribute to the failure of the closed-loop servo beyond 60% of the gap. The d.c. offset voltage introduced by charging effect, thermally-induced plate displacement, and piezo-actuator drift can also affect
stability if it is not completely nulled during the experiment. The possibility that instability at atmospheric pressure is caused by excitation of the high-order mechanical modes is ruled out based on simulation results from NODAS.

6.1 Future Work

This thesis provides a general design procedure for parallel-plate servo using a linear time-invariant controller, which is less versatile than a nonlinear controller or a linear time-varying controller in terms of bandwidth efficiency and attainable gain and phase margins. These types of controller design are promising to further extend the travel range to the entire gap. The control problem can be of interest to the control system community as well. One of the possible options for design of a nonlinear controller can use a nonlinear circuit network to first cancel the plant nonlinearity, followed by conventional design of a linear controller.

Control of the actuator as it returns back to the rest position is another research topic equally important to the problem of extending its travel range. We have shown in Chapter 5 that the closed-loop step response exhibits different risetime and falltime due to the dissimilar actuator dynamics in the rising and falling periods. Therefore system stability is not automatically guaranteed in the falling period by the current controller design. We have found that a large damping coefficient at small gaps can destabilize the system during the input-command transition, contrary to its stabilizing effect in the rising period. The stability issue with respect to the squeeze-film damping coefficient in the returning pull-out path has to be further analyzed to complete the position-servo cycle. The results can be utilized in design of the actuator to obtain desired value of damping coefficient.

There is much work to do regarding the fabrication and testing of the tip-actuator array. Size of the on-chip sensing and control electronics has to be minimized to fit into a tip-actuator cell, and the circuit power consumption has to be reduced to increase the total number of active tip actuators for data read and write. The trade-off is that the circuit gain-bandwidth and noise performance will be negatively impacted by decreased transistor size. It is reasonable to say that each of the circuit blocks, including controller, pre-amp, demodulator, subtracter, and pre-filter, can be implemented with a size of at least 30 to
50 µm on the side. Therefore, fitting them and the tip actuator all into an 10,000 µm² tip actuator cell will be a challenge. The media stroke may have to increase from the current ±50 µm by 50% more to increase the tip actuator cell size by 125%. An even larger increase of the actuator cell size may be necessary if operational amplifiers are used to implement pre-amp, controller, and pre-filter for precise gains, which are highly desirable to better ensure stability robustness. In addition, the passive components used for filtering out capacitive feedthrough can also take up large areas. A better demodulation circuit with less feedthrough needs to be built to prevent undesired signals from being modulated into the signal band. High-voltage MOSFET transistors providing up to ±15 V need to be fabricated. This voltage value becomes larger if the desired closed-loop response has to be much faster than milli-seconds. From the supplied-voltage standpoint, it is not necessary to use a state-of-the-art sub-micron CMOS process for fabrication of the tip-actuator array. The amount of current in the control circuit is not limited by slew rate since the actuated capacitance is only 10 to 20 fF, and the controller output changes in milli-seconds. Required current is more related to providing large transconductances for building high-speed electronics (gain-bandwidth product > 20 MHz). Finally, stability robustness against external shock forces should also be analyzed and experimentally tested as the commercial portable disk drives. Currently there is not a pull-down electrode in the actuator design to actively prevent the snap-in. This part of the research can be further explored in the future.
Bibliography


Appendix A

Electrostatic Gap Model

Electrostatic gap models presented in Chapter 2 are programmed by the Analogy MAST® language for simulation in SABER. The following file (gap_parallel.sin) is the implemented electrostatic gap model for parallel plates.

```plaintext
#*******************************************************************#
#       MAST Code for an electrostatic gap in the x-direction
# #
# Jan E. Vandemeer and Michael Kranz
# Carnegie Mellon University
# Michael Lu   8/6/98
# Qi Jing     8/10/98
#*******************************************************************#
element template gap_parallel x_a_t x_b_t x_a_b x_b_b \ 
  y_a_t y_a_b y_b_t y_b_b \ 
  phi_a_t phi_a_b phi_b_t phi_b_b \ 
  v_a_t v_a_b v_b_t v_b_b = finger_w_t, finger_w_b, finger_length, gap, overlap, topology

translational_pos x_a_t, #x-displacement and force at port "x_a_t"
  y_a_t, #y-displacement and force at port "y_a_t"
  x_a_b, #x-displacement and force at port "x_a_b"
  y_a_b, #y-displacement and force at port "y_a_b"
  x_b_t, #x-displacement and force at port "x_b_t"
  y_b_t, #y-displacement and force at port "y_b_t"
  x_b_b, #x-displacement and force at port "x_b_b"
  y_b_b, #y-displacement and force at port "y_b_b"

rotational_ang phi_a_t, #Angle displacement about z angle and torque
  phi_a_b, #Angle displacement about z angle and torque
  phi_b_t, #Angle displacement about z angle and torque
  phi_b_b #Angle displacement about z angle and torque

electrical         v_a_t, v_a_b, v_b_t, v_b_b #input & output voltage

number       finger_w_t = undef    #width of the top finger
number       finger_w_b = undef    #width of the bottom finger
number       finger_length = undef #width of the top and bottom fingers
number       gap = undef           #Gap between the two beams
number       overlap = undef       #X-overlap between the beams
number       topology = undef      #Situation of top beam relative to bottom beam
```
{174
<br/ Số x, th, k1, k2, k3
<br/ Số ntx = 20n # Native oxide thickness around the fingers
<br/ val frc_N fy1x, fy2x
<br/ val frc_N fx1x, fx2x
<br/ val pos_m ovr1x # overlap of beams
<br/ val pos_m pt_x, pb_x, mt_x, mb_x
<br/ val pos_m pt_y, pb_y, mt_y, mb_y
<br/ val pos_m tmp
<br/ val pos_m dy # The displacements in y
<br/ number ydc # The DC value of y position.
<br/ val ang_rad dphi
<br/ val v vlt
<br/ val c cap # capacitance between comb-fingers
<br/ val v vltg # v^2
<br/ val nu const, # state, digital on or off - alternates force function
<br/ const2 # the spring constant (after crashing)
<br/ val nu length # comb-finger length
<br/ val g q1, q2
<br/ val pos_m g_eff
<br/ val nu w # force per unit length
<br/ val nu slope
<br/ # whenever you have the two beams crash, shorten time steps to help simulator
<br/><br/ # Parameters section
<br/ parameters{
<br/ ydc = gap + finger_w_t/2 + finger_w_b/2
<br/ if (overlap == undef){
<br/ error("Overlap is undefined, please give a numerical value.")
<br/ }
<br/ if (finger_w_b == undef){
<br/ error("Bottom fingers width is undefined, please give a numerical value.")
<br/ }
<br/ if (finger_w_t == undef){
<br/ error("Top fingers width is undefined, please give a numerical value.")
<br/ }
<br/ if (finger_length == undef){
<br/ error("Top and bottom finger length is undefined, please give a numerical value.")
<br/ }
<br/ if (gap == undef){
<br/ error("Gap is undefined, please give a numerical value.")
<br/ }
<br/ if (topology < 0 | topology > 1){
<br/ error("Invalid topology, must be either '0' or '1', see 'help' for descriptions.")
<br/>}
<br/>174}
\[
\text{coeff. in capacitance and electrostatic-force equations}
\]

\[
x = \frac{\text{finger}_w_t}{\text{poly}_1_t} \quad \# \text{ratio of finger width to finger thickness}
\]

\[
\text{th} = 1.496 \quad \# \text{sidewall angle}
\]

\[
k_1 = -1.565 + 0.2818/(x^{0.04348}) - 2.986\times{(\text{th}^{5.900})} + 0.4446\times\exp(\text{th})\times(\text{th}^{6.002}) + 28.87\times\ln(\text{th})
\]

\[
k_2 = 21.43 - 15.50/(x^{0.02146}) - 10.07\times\exp(\text{th})/(x^{0.03944})/(\text{th}^{1.877}) + 23.70\times(\text{th}^{0.09913}) + 0.04840\times\exp(x)\times(\text{th}^{0.8278})/(x^{2.244}) - 13.60\times\ln(\text{th})
\]

\[
k_3 = 10.86 - 1.512/(x^{0.2517}) - 0.007964\times(\text{th}^{15.31}) + 0.005087\times\ln(x) - 62.71\times\ln(\text{th}) + 20.05\times(\text{th}^{2.529})\times\ln(\text{th})/(x^{0.002656})
\]

\# values section

\[
\text{pt}_x = \text{pos}_m(x_b_t)
\]

\[
\text{pb}_x = \text{pos}_m(x_b_b)
\]

\[
\text{mt}_x = \text{pos}_m(x_a_t)
\]

\[
\text{mb}_x = \text{pos}_m(x_a_b)
\]

\[
\text{pt}_y = \text{pos}_m(y_b_t)
\]

\[
\text{pb}_y = \text{pos}_m(y_b_b)
\]

\[
\text{mt}_y = \text{pos}_m(y_a_t)
\]

\[
\text{mb}_y = \text{pos}_m(y_a_b)
\]

\# topology = 0 means that the top finger is located to the right of the bottom finger, and
\# topology = 1 means the opposite

\[
\text{if} (\text{topology} == 0) \{
    \text{ovrlp} = \text{overlap} + (\text{pb}_x - \text{mt}_x)
    \text{dy} = (\text{pos}_m(y_b_b, y_b_t) + \text{pos}_m(y_a_b, y_a_t))/2
    \text{vlt} = v(v_b_b, v_a_t)
\}
\]

\[
\text{else if} (\text{topology} == 1)\{
    \text{ovrlp} = \text{overlap} + (\text{pt}_x - \text{mb}_x)
    \text{dy} = (\text{pos}_m(y_b_b, y_b_t) + \text{pos}_m(y_a_b, y_a_t))/2
    \text{vlt} = v(v_a_b, v_b_t)
\}
\]

\[
\text{slope} = E\times\text{ovrlp}\times\text{poly}_1_t/(\text{finger}_w_t/2 + \text{finger}_w_b/2)
\]

\[
\text{if}((\text{gap} - \text{dy}) > \text{ntv}_ox_t)\{
    \text{const} = 1
    \text{const2} = 0
\}
\]

\[
\text{else}\{
    \text{const} = 0
    \text{const2} = \text{slope}
\}
\]
vltg = (vlt)**2

\[ g_{eff} = (1-const)\times ntv_{ox_t} + const\times (gap - dy) \]

\[ cap = \varepsilon_0 \times poly1_t \times ovrlp \times (k1 + g_{eff} / \pi / poly1_t \times (k2 + k3 \times \ln(\pi / poly1_t / g_{eff}))) / g_{eff} \]

if (topology == 1) {
    phi = ang_rad(phi_b_t, phi_a_b)
    fxd1 = -vltg/2*\varepsilon_0*poly1_t*(k1+g_{eff}/\pi/poly1_t*(k2+k3*ln(\pi*poly1_t/g_{eff})])/g_{eff}
}

w = (vltg/2*(k1*\varepsilon_0*poly1_t*ovrlp/((g_{eff})**2)+k3*\varepsilon_0*ovrlp/\pi/g_{eff}) - const2*(dy-gap+ntv_{ox_t})/(ovrlp+1p))

fxd2 = 0

\[ fx = \frac{w}{2}\times (ovrlp+(ovrlp**2)/(finger_length+1p)) \]

\[ fy = \frac{w}{2}\times (ovrlp-(ovrlp**2)/(finger_length+1p)) \]

q1 = abs(v(v_b_t, v_a_b))*cap
q2 = 0

else if (topology == 0) {
    phi = ang_rad(phi_a_t, phi_b_b)
    fxd1 = 0
    w = (vltg/2*(k1*\varepsilon_0*poly1_t*ovrlp/((g_{eff})**2)+k3*\varepsilon_0*ovrlp/\pi/g_{eff}) - const2*(dy-gap+ntv_{ox_t})/(ovrlp+1p))
}

fzd2 = vltg/2*\varepsilon_0*poly1_t*(k1+g_{eff}/\pi/poly1_t*(k2+k3*ln(\pi*poly1_t/g_{eff})])/g_{eff}

\[ fx = \frac{w}{2}\times (ovrlp+(ovrlp**2)/(finger_length+1p)) \]

\[ fy = \frac{w}{2}\times (ovrlp**2)/(finger_length+1p) \]

q1 = 0
q2 = abs(v(v_a_t, v_b_b))*cap

} equations {
    \[ i(v_a_b->v_b_t) += d_{by\_dt}(q1) \]
    \[ i(v_a_t->v_b_b) += d_{by\_dt}(q2) \]
    \[ frc_N(x_b_t->x_a_b) -= fxd1 \]
    \[ frc_N(x_a_t->x_b_b) -= fxd2 \]
    \[ frc_N(y_b_t->y_a_b) -= fyd1 \]
    \[ frc_N(y_a_t->y_b_b) -= fyd2 \]
    
}
Figure B.1  Schematic of feedback control board. Controller gain $k_p = 24.3$ is split into two parts ($24.3 = 10 \times 2.43$). The gain of ten is placed after subtraction of the initial sensed voltage.
Figure B.2  Schematic of demodulation circuit.