

Abstract

HATCH, ANDREW GRAYDON. Model Development and Control Design for Atomic Force Microscopy. (Committee Chair: Ralph C. Smith)

The development of energy-based models and model-based control designs necessary to achieve present and projected applications involving atomic force microscopy is investigated. Applications include real-time product diagnostics or monitoring of biological processes, nanoelectromechanical systems (NEMS) and employment of atomic force microscope (AFM) technology for spintronics. A crucial component in the AFM design is the piezoceramic (PZT)-based stage used to position the sample. Whereas PZT actuators provide the broadband and extremely high set point capabilities required by the AFM stages, they also exhibit frequency-dependent hysteresis and constitutive nonlinearities.

To characterize the field-polarization relation in PZT, low-order macroscopic models are constructed based on a combination of energy analysis at the mesoscopic level along with stochastic homogenization techniques. To account for nonuniformity and inhomogeneities in the material, local coercive field values are assumed to be distributed. Due to interactions among the dipoles, the effective field is also assumed to be distributed. Previous work has employed specific functions to describe these distributions. However, the fact that these choices are not based on energy considerations, motivates the use of general densities.

The dynamics of the actuator must be incorporated as well. A rod model is suitable for a stacked actuator whose cross-section is small compared to the length. The equation of motion for the rod can be derived using force balancing with boundary conditions determined from the fact that the rod is fixed at one end and pushes against the stage at the other.

At low frequencies, the hysteresis and constitutive nonlinearities inherent in PZT can be accommodated through PID or robust control designs. However, at the higher frequencies required by the previously outlined applications, increasing noise-to-data ratios and diminishing high-pass characteristics of control filters preclude a sole reliance on feedback laws to eliminate hysteresis. This motivates the development of control designs that incorporate and approximately compensate for hysteresis through model inverses employed as filters to linearize transducer responses for linear robust control design and PID control design. The inverse models are also tested in an open loop control experiment on a PZT stacked actuator.

MODEL DEVELOPMENT AND CONTROL DESIGN FOR ATOMIC FORCE MICROSCOPY

BY

ANDREW G. HATCH

A DISSERTATION SUBMITTED TO THE GRADUATE FACULTY OF
NORTH CAROLINA STATE UNIVERSITY
IN PARTIAL FULFILLMENT OF THE
REQUIREMENTS FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

APPLIED MATHEMATICS

RALEIGH, NORTH CAROLINA

SEPTEMBER 2004

APPROVED BY:

R.C. SMITH
CHAIR OF ADVISORY COMMITTEE

K. ITO

Z. LI

H.T. TRAN

Biography

The author was born in Lake Forest, Illinois and has lived most of his life in Raleigh, NC. He attended North Carolina State University, where he received the degrees of Bachelor of Science in Mathematics and Bachelor of Science in physics in 1996. He went on to receive the degree of Master of Science in Applied Mathematics from North Carolina State University in 2000. After spending nearly two years working as a tutor at Triangle Learning Consultants and a software consultant at Intelligent Information Systems, he returned to NCSU in 2002 to complete his doctoral work.

Acknowledgements

I would like to thank my advisor Dr. Ralph C. Smith as well as the members of the committee Dr. Kazufumi Ito, Dr. Zhilin Li and Dr. Hien T. Tran. I would also like to thank my late father Dr. George G. Hatch, my mother Susan, my brothers Timothy and Jonathan, my friend Mike Zager and my wife Elena.

Table of Contents

List of Figures	vi
1 Introduction	1
2 Model Development	9
2.1 Constitutive Relations	9
2.2 Well-Posedness Analysis	15
2.3 Parameter Estimation Problem	19
2.4 Coupled Electromechanical Constitutive Relations	20
2.5 Actuator Models	21
2.5.1 Stacked Actuator	21
2.5.2 Cylindrical Actuator	24
3 Approximation Techniques and Implementation Algorithms	29
3.1 Approximation Techniques for Polarization Model	30
3.2 Model Inversion	38
3.3 Parameter Estimation Techniques	41
3.3.1 Estimation of Densities by Constrained Optimization	41
3.3.2 Estimation of Joint Density	42
3.3.3 Estimation of Joint Density with Regularization	45
3.4 Approximation Techniques for Rod Model	46
3.4.1 Complete Rod Model	46
3.4.2 Lumped Rod Model	52
3.5 Approximation Techniques for Shell Model	54
4 Model Validation and Device Characterization	56
4.1 Field-Polarization	56

4.1.1	Lognormal/Normal Densities, Identification Using Major and Rayleigh Loops	58
4.1.2	General Densities, Identification Using Major and Rayleigh Loops	60
4.1.3	Lognormal/Normal Densities, Identification Using All Seven Loops	60
4.1.4	General Densities, Identification Using All Seven Loops	63
4.1.5	Joint Density	65
4.2	Field-Displacement	67
5	PI Control	71
5.1	PI Control Design	71
5.2	Numerical Examples	76
6	Robust Control	80
6.1	Robust Control Design	80
6.2	Numerical Examples	89
7	Open Loop Control	93
7.1	Experimental Design	93
7.2	Experimental Results	94
8	Conclusions and Future Directions	102
	List of References	104

List of Figures

1.1	Schematic of a prototypical atomic force microscope (AFM).	2
1.2	Fundamental components in single electron spin microscopy.	3
1.3	PZT-based AFM stage.	4
1.4	Frequency-dependent field-displacement data from an AFM. Sample rates of (a) 0.279 Hz, (b) 1.12 Hz, (c) 5.58 Hz, and (d) 27.9 Hz.	5
1.5	Nested minor loops in data collected at 0.1 Hz from an AFM.	6
2.1	Gibbs energy G and hysteron \overline{P} resulting from the necessary condition $\frac{\partial G}{\partial P} = 0$ for negligible thermal activation.	11
2.2	Cylindrical actuator.	25
4.1	PZT5H data collected at 0.2 Hz with a symmetric major loop, Rayleigh loop and five biased minor loops.	57
4.2	PZT5H data and model with lognormal and normal densities identified through a fit to major and Rayleigh loops.	59
4.3	Lognormal coercive and normal effective field densities identified through a fit to major and Rayleigh loops.	59
4.4	PZT5H data and model with general densities identified through a fit to major and Rayleigh loops.	61
4.5	General coercive and effective field densities identified through a fit to major and Rayleigh loops.	61
4.6	PZT5H data and model with lognormal and normal densities identified through a fit to all seven loops.	62
4.7	Lognormal coercive and normal effective field densities identified through a fit to all seven loops.	62
4.8	PZT5H data and model with general densities identified through a fit to all seven loops.	64
4.9	General coercive and effective field densities identified through a fit to all seven loops.	64

4.10	PZT5H data and model with joint density ν identified using the quadratic programming formulation (3.29) using data from all seven loops. (a) $N_i = N_j = 24$ and (b) $N_i = N_j = 48$	66
4.11	Joint density ν identified using the quadratic programming formulation (3.29) for (a) $N_i = N_j = 24$ and (b) $N_i = N_j = 48$	66
4.12	PZT5H data and model with joint density ν identified with Tikhonov regularization using data from all seven loops. (a) $N_i = N_j = 24$ and (b) $N_i = N_j = 48$	68
4.13	Joint density ν identified using Tikhonov regularization for (a) $N_i = N_j = 24$ and (b) $N_i = N_j = 48$	68
4.14	Characterization of AFM field-displacement behavior at 0.1 Hz.	69
4.15	Characterization of AFM field-displacement behavior with sample rates of (a) 0.279 Hz, (b) 1.12 Hz, (c) 5.58 Hz, and (d) 27.9 Hz.	70
5.1	On/off control.	72
5.2	Proportional control.	73
5.3	(a) Disturbance d due to scaled but unmodeled hysteresis and constitutive nonlinearities. (b) Disturbance d due to inverse filtering errors.	75
5.4	P design with no disturbance d and a frequency of 1 Hz. (a) Reference and simulated trajectory, and (b) tracking error.	77
5.5	PI design with no disturbance d and a frequency of 1 Hz. (a) Reference and simulated trajectory, and (b) tracking error.	77
5.6	PI design with disturbance d due to inversion error and a frequency of 1 Hz. (a) Reference and simulated trajectory, and (b) tracking error.	78
5.7	PI design with disturbance d due to uncompensated hysteresis and constitutive nonlinearities and a frequency of 1 Hz. (a) Reference and simulated trajectory, and (b) tracking error.	78
5.8	PI design with disturbance d due to inversion error and a frequency of 10 Hz. (a) Reference and simulated trajectory, and (b) tracking error.	79
5.9	PI design with disturbance d due to uncompensated hysteresis and constitutive nonlinearities and a frequency of 10 Hz. (a) Reference and simulated trajectory, and (b) tracking error.	79
6.1	System representation including input disturbance d and sensor noise s and n in the transducer.	81
6.2	Linear fractional transformation representation (LFT) of the transducer model.	82
6.3	(a) Frequency response of the passband filter W_s	83
6.4	Frequency response of the highpass filter W_n	84

6.5	Frequency response of W_d for the disturbance d due to scaled but uncompensated hysteresis and nonlinearities.	85
6.6	\mathcal{H}_2 design with sensor noise s and the disturbance d due to inversion error. (a) Reference and simulated trajectory, and (b) tracking error.	90
6.7	\mathcal{H}_2 design with sensor noise s and the disturbance d due to uncompensated hysteresis and constitutive nonlinearities. (a) Reference and simulated trajectory, and (b) tracking error.	90
6.8	\mathcal{H}_2 design with sensor noise s but no disturbance d . (a) Reference and simulated trajectory, and (b) tracking error.	91
6.9	\mathcal{H}_∞ design with sensor noise s and the disturbance d due to inversion error. (a) Reference and simulated trajectory, and (b) tracking error.	91
6.10	\mathcal{H}_∞ design with sensor noise s and the disturbance d due to uncompensated hysteresis and constitutive nonlinearities. (a) Reference and simulated trajectory, and (b) tracking error.	92
6.11	\mathcal{H}_∞ design with sensor noise s but no disturbance d . (a) Reference and simulated trajectory, and (b) tracking error.	92
7.1	(a) Desired displacement with frequency 0.279 Hz and amplitude 40.56 μm , displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.	95
7.2	(a) Desired displacement with frequency 0.279 Hz and amplitude 27.04 μm , displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.	96
7.3	(a) Desired displacement with frequency 2.79 Hz and amplitude 33.80 μm , displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.	97

7.4	(a) Desired displacement with frequency 2.79 Hz and amplitude 27.04 μm , displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.	98
7.5	(a) Desired displacement with frequency 27.9 Hz and amplitude 27.04 μm , displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.	99
7.6	(a) Desired displacement with frequency 27.9 Hz and amplitude 20.28 μm , displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.	100
7.7	Data collected before and after open loop experiments at (a) 0.279 Hz and (b) 27.9 Hz.	101

Chapter 1

Introduction

Developed in 1986, the atomic force microscope (AFM), depicted in Figure 1.1, relies on interatomic forces between a cantilever tip and the sample to obtain ultra-high resolution surface images [8]. The relatively low cost of the devices (less than \$100,000) and the fact that they require minimal sample preparation has made the AFM a standard diagnostic tool in research laboratories. However, several present and projected applications make requirements on the technology that present AFM designs are unable to consistently achieve. These limitations are primarily due to the high sample rates (in the MHz range) required for applications such as real-time product diagnostics or monitoring of biological processes, nanoelectromechanical (NEMS) applications, and employment of AFM technologies for spintronics. Real-time product diagnostics include analysis of contact lenses to detect defects or protein deposits and screening of semiconductor chips to maintain quality control. The real-time monitoring of biological processes has the potential for leading to treatment policies for ailments such as osteoporosis [31] as well as the potential for quantifying fundamental biological phenomena such as protein unfolding. Within the NEMS regime, the repulsive forces utilized in atomic force microscopy also lead to the potential for nanoconstruction using the cantilever as an actuator. Finally, the extreme accuracy provided

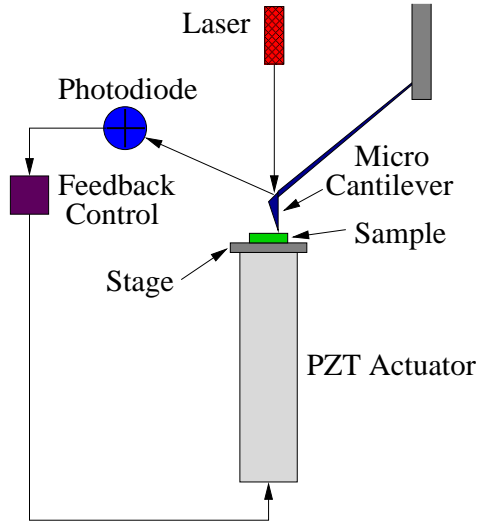


Figure 1.1: Schematic of a prototypical atomic force microscope (AFM).

by the AFM is presently being combined with nuclear magnetic resonance microscope (NMRM) technology to investigate the detection of single electron spins [4, 29, 46] with the proposed single electron spin microscope (SESM). In the operation of the SESM, shown schematically in Figure 1.2, an external magnetic field H_0 keeps the spins in alignment. Microwave pulses are then used to invert spins in the sample. The change in force of approximately 10^{-15} N caused by the inversion of a single spin is detected by the cantilever.

A crucial component in the AFM design for all of these applications is the piezoceramic (PZT)-based stage used to position the sample as depicted in Figure 1.3 [3, 8, 34, 39, 40]. Whereas PZT actuators provide the broadband and extremely high set point capabilities required by the AFM stages, they also exhibit frequency-dependent hysteresis as shown in Figure 1.4. Inertial effects in the rod are clearly visible in the higher frequency data as evidenced by the observation that the displacements continue to increase for a short time after field reversal. Nested minor loops in data collected at a fixed frequency of 0.1 Hz are shown in Figure 1.5.

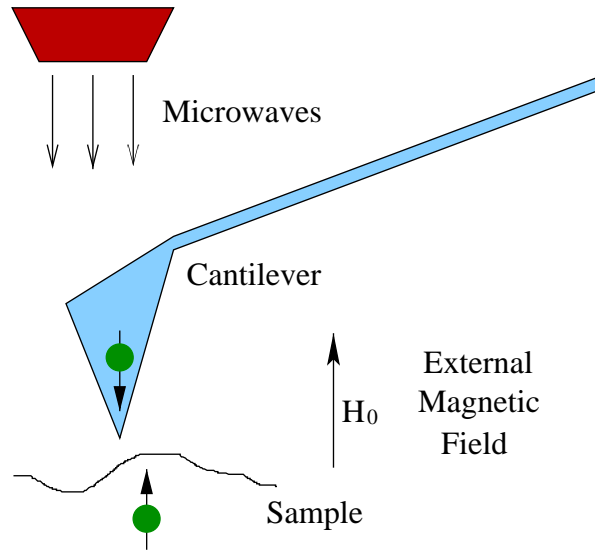


Figure 1.2: Fundamental components in single electron spin microscopy.

The effects of hysteresis in ferroelectric materials can be greatly reduced or even nearly eliminated by the use of current or charge controlled amplifiers in place of voltage controlled amplifiers [15, 16, 17, 18, 19]. However, this approach can be prohibitively expensive since it requires the construction of specialized amplifiers. Furthermore, if, as in many applications, there is a need to maintain DC offsets, current control will be ineffective. For example, this situation occurs when the x-actuator in an AFM must be kept fixed while the sample is scanned in the y-direction.

The two goals of this investigation are the development of models that characterize the hysteresis and constitutive nonlinearities inherent to PZT stages and model-based control design. We summarize first the state of the art and contributions in each case.

When PZT is placed in the poled state necessary to achieve bidirectional strains, low to moderate input fields will generate an approximately linear response. The initial development of linear models characterizing the behavior of piezoelectric materials is attributed to Voigt [45].

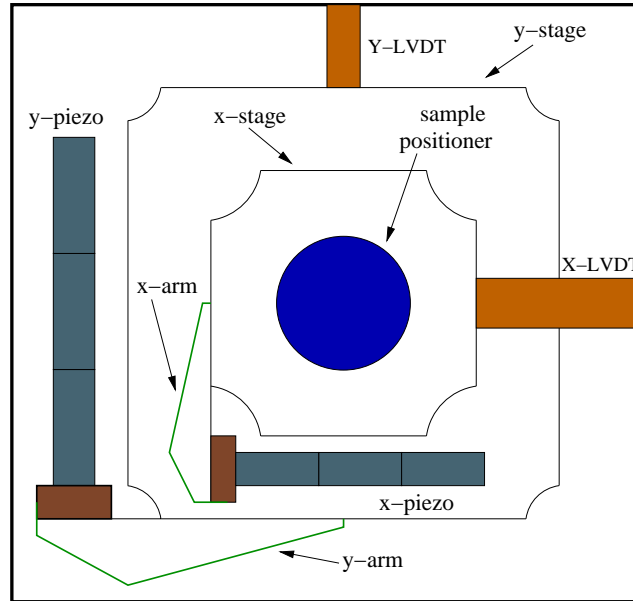


Figure 1.3: PZT-based AFM stage.

A second, more general, framework for characterizing hysteresis and constitutive nonlinearities is that of domain wall models. Domain wall models involve two basic mechanisms, domain wall bending, which is reversible, and domain wall translation, which is irreversible. The theory developed in [36, 37, 38] is based on these concepts. Additional information on the theory of domain processes can be found in [14] while a historical summary is provided in [36]. The theory is based on work originally proposed by Jiles and Atherton for ferromagnetic materials [12].

Preisach methods constitute another general class of models. For example, one of the first papers studying the application of Preisach models to ferroelectric compounds was written by Sreeran, Salvady and Naganathan [42] in 1993. An investigation of Preisach models for PZT, magnetic materials and shape memory alloys (SMAs) was conducted by Hughes and Wen [11] in 1997. Two examples of extensions of the Preisach methods for systems containing piezoceramic transducers are found in [6, 28].

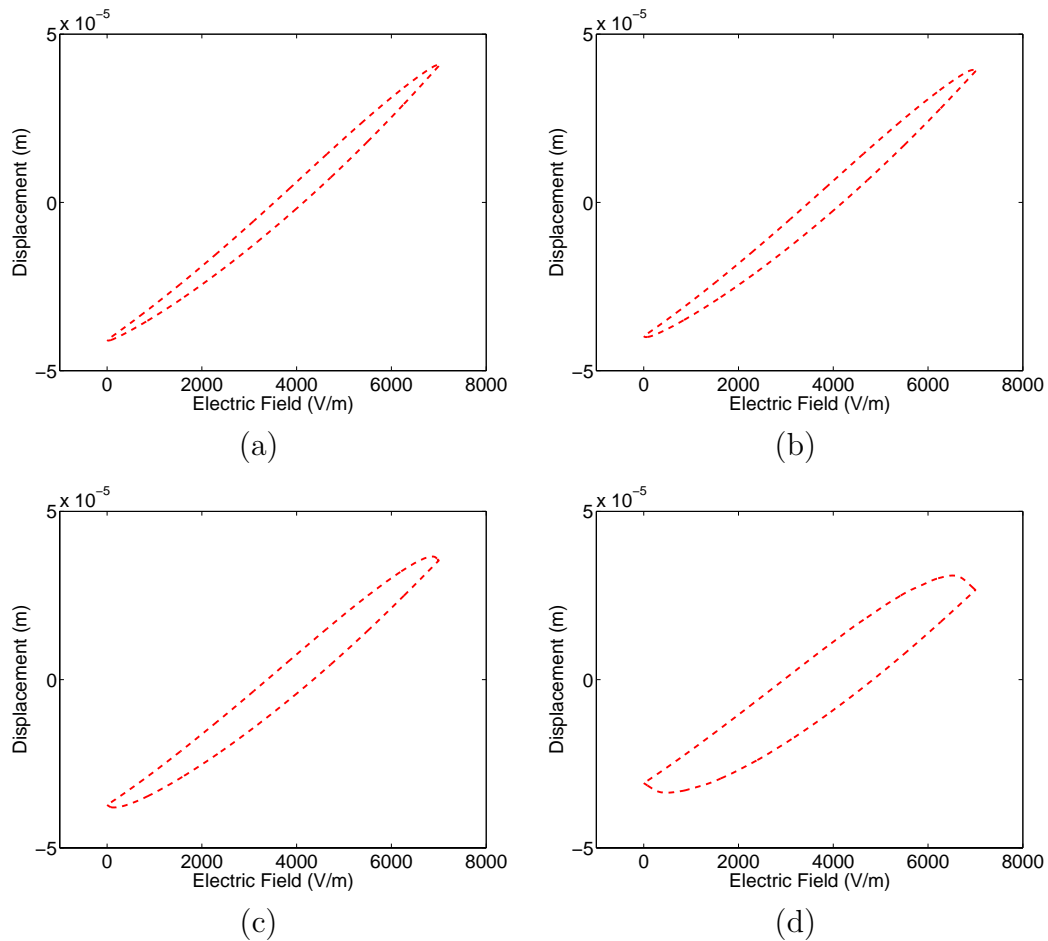


Figure 1.4: Frequency-dependent field-displacement data from an AFM. Sample rates of (a) 0.279 Hz, (b) 1.12 Hz, (c) 5.58 Hz, and (d) 27.9 Hz.

Classical Preisach methods are mathematical in structure and are thus very general as illustrated by Mayergoyz [20, 21], who established mathematical criteria to verify the applicability of these methods to a variety of magnetic hysteretic situations. Thus, Preisach models are suitable for characterizing hysteresis in a wide range of materials. However, the generality of Preisach techniques also creates drawbacks. Since the parameters in classical Preisach models are not physically based, it is difficult to account for changing operating conditions such as temperature, stress and frequency.

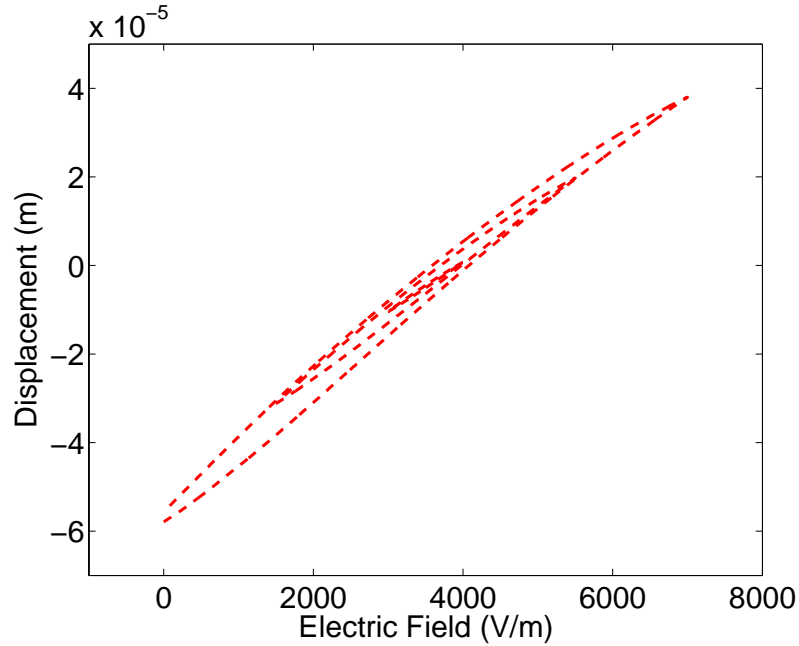


Figure 1.5: Nested minor loops in data collected at 0.1 Hz from an AFM.

Another limitation of the classical Preisach formulation is the closure of minor loops, a phenomenon not always observed in actual materials, where relaxation mechanisms can prevent this from occurring. For magnetic hysteresis models, extended Preisach formulations have been developed to address these limitations, although less has been done for ferroelectric materials in this regard. It should be noted that Preisach models were developed by Preisach in 1935 for use with magnetic materials [26] and most early research focused on ferromagnetic materials.

Statistically homogenized energy models provide yet another method for addressing hysteresis and these are the type of models that are developed in the present investigation. It is demonstrated in [39] that these models provide an energy basis for Preisach models. However, the limitations of the Preisach model can be addressed through the energy-based hysteresis kernel.

Chapter 2 contains the constitutive model development, based on theory previously developed in [41], and is constructed in three steps. In the first, a Helmholtz energy relation is derived using statistical mechanics principles under the assumption that dipoles are either aligned with the field or diametrically opposed to it. A quadratic approximation to this energy is then constructed to facilitate high speed implementation. In the case of low thermal activation, this leads to a piecewise linear relationship between the electric field E and the local polarization \bar{P} . If thermal activation is significant, the dipoles can change orientation at any time, with the probability of such a change determined from Boltzmann principles. The number of dipoles in each orientation evolves according to an ordinary differential equation.

To account for nonuniformity and inhomogeneities in the material, local coercive field values are assumed to be distributed. Due to interactions among the dipoles, the effective field is also assumed to be distributed. Previous work has employed specific functions to describe these distributions. For example, lognormal coercive field densities are common in the magnetics literature and one possible choice for the effective field density is a normal distribution. However, the fact that neither of these choices is based on energy considerations motivates the use of general densities. In this case, the function values of the densities are themselves parameters with the number of parameters varying according to the desired accuracy. These parameters must be identified along with the parameter η , which characterizes the slope of the relationship between E and \bar{P} .

In addition to the hysteresis, the dynamics of the actuator must be incorporated. For a stacked actuator, such as the one depicted in Figure 1.3, a rod model is suitable since the cross-section is small compared to the length. The equation of motion for the rod can be derived using force balancing with boundary conditions determined from the fact that the rod is fixed at one end and pushes against the stage at the other.

As discussed in Chapter 3 an approximate solution to the rod equation is found by first deriving a weak form and then discretizing in space with the finite element method and in time with finite difference techniques. Approximation techniques and algorithms for the field-polarization model are also discussed in this chapter.

Chapter 4 addresses model validation for both field-polarization model and the complete stacked actuator model. For the field-polarization model, a PZT5H data set collected at 0.2 Hz is used, whereas the AFM data shown in Figures 1.4 and 1.5 is used for the field-displacement model. The frequency-dependent model is employed for the data in Figure 1.4.

In addition to characterization, a second goal of the models in the context of the AFM is to develop model-based control algorithms. At low frequencies, the hysteresis inherent to smart materials can be accommodated through proportional-integral-derivative (PID) or robust control designs [2, 30]. However, at the higher frequencies required by the previously summarized AFM applications, increasing noise-to-data ratios and diminishing high-pass characteristics of control filters preclude a sole reliance on feedback laws to eliminate hysteresis. This motivates the development of control designs that incorporate and approximately compensate for hysteresis through model inverses employed either in feedback or feedforward loops. Chapters 5 and 6 discuss PID and robust control design, respectively. In Chapter 7 an open loop control experiment is described. An inverse model is used to predict the appropriate input field given a desired output field. The prediction is then tested on a stacked actuator. For comparison, a predicted input field derived by a linear scaling of the desired output is also tested. Chapter 8 contains conclusions and possible future directions.

Chapter 2

Model Development

In this chapter we develop models for the nonlinear and hysteretic relations between the electric field and polarization in PZT actuators. An abstract formulation that defines this model in terms of a compact operator is then described. This leads to an analysis of the well-posedness of the parameter estimation problem. After a discussion detailing the coupled electromechanical constitutive relations, rod and shell actuator models, which quantify the relation between field and strain, are developed.

2.1 Constitutive Relations

To model the constitutive behavior of the piezoceramic stacked actuator, the stress-strain relation is assumed to be linear. However, the relation between the applied field E , or the applied voltage V , and the polarization P exhibits nonlinearities and hysteresis. The actuator is also biased through poling so that the relation between P and the strain ε is approximately linear for the considered operating conditions. To characterize the hysteretic E - P behavior at the domain level, a Helmholtz energy relation was derived in [41] using statistical mechanics principles under the assumption that dipoles are either aligned with the field or diametrically opposed to it. This

model is appropriate for a single crystal with uniform effective fields. To construct a macroscopic model for a polycrystalline material with variable effective fields, the coercive and effective field values are then assumed to be distributed.

Under fixed temperature conditions with no applied stress σ , it is illustrated in [41] that a first order approximation to the statistical mechanics-based Helmholtz energy is the piecewise quadratic relation

$$\psi(P) = \begin{cases} \frac{1}{2}\eta(P + P_R)^2 & P \leq -P_I \\ \frac{1}{2}\eta(P - P_R)^2 & P \geq P_I \\ \frac{1}{2}\eta(P_I - P_R) \left(\frac{P^2}{P_I} - P_R \right) & |P| < P_I. \end{cases} \quad (2.1)$$

As shown in Figure 2.1, P_I is the positive inflection point and P_R is the polarization value at which the positive local minimum of ψ occurs. The parameter η is the reciprocal of the slope of the E - P relation after switching occurs. This fact can be used to establish an initial parameter value for η when modeling a specific data set.

If there are no applied stresses σ , the Gibbs energy can be formulated as

$$G(E, P) = \psi(P) - EP, \quad (2.2)$$

where the second term represents the electrostatic energy due the applied field E . In the case of negligible thermal activation, the local average polarization \bar{P} is determined from the necessary conditions

$$\frac{\partial G}{\partial P} = 0 \quad , \quad \frac{\partial^2 G}{\partial P^2} > 0, \quad (2.3)$$

where the second derivative condition holds locally. Applying these conditions to

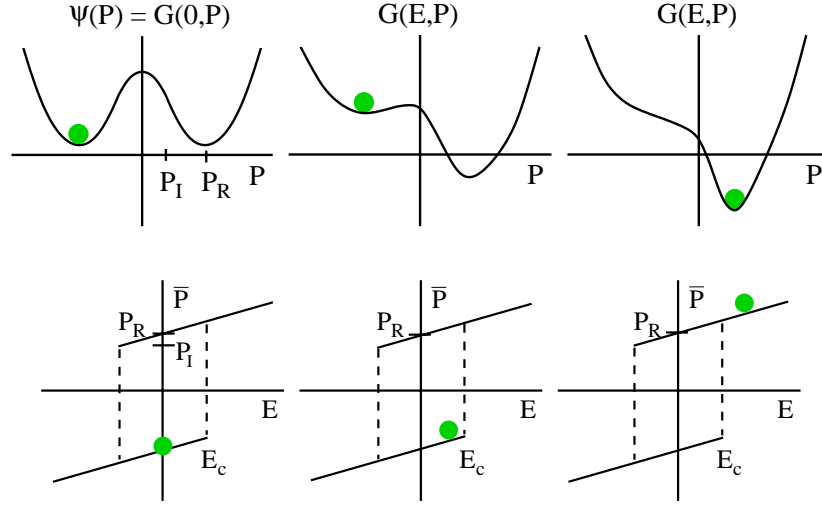


Figure 2.1: Gibbs energy G and hysteron \bar{P} resulting from the necessary condition $\frac{\partial G}{\partial P} = 0$ for negligible thermal activation.

(2.2) yields a piecewise linear E - P characterization

$$[\bar{P}(E; E_c, \xi)](t) = \begin{cases} [\bar{P}(E; E_c, \xi)](0), & \tau(t) = \emptyset \\ \frac{E}{\eta} - P_R, & \tau(t) \neq \emptyset \text{ and } E(\sup \tau(t)) = -E_c \\ \frac{E}{\eta} + P_R, & \tau(t) \neq \emptyset \text{ and } E(\sup \tau(t)) = E_c \end{cases} \quad (2.4)$$

with the initial dipole orientation given by

$$[\bar{P}(E; E_c, \xi)](0) = \begin{cases} \frac{E}{\eta} - P_R, & E(0) \leq -E_c \\ \xi, & -E_c < E(0) < E_c \\ \frac{E}{\eta} + P_R, & E(0) \geq E_c \end{cases} \quad (2.5)$$

Here $E_c = \eta(P_R - P_I)$ is the local coercive field and ξ is the set of initial dipole

configurations. The transition times τ are

$$\tau(t) = \{t \in (0, t_f] \mid E(t) = -E_c \text{ or } E(t) = E_c\}. \quad (2.6)$$

The local average polarization can also be written using more compact notation as

$$\bar{P} = \frac{1}{\eta} E + \delta P_R, \quad (2.7)$$

where $\delta = 1$ for positively oriented dipoles and $\delta = -1$ for negative orientations.

However, if thermal activation is significant, the dipoles can achieve the thermal energy required to switch in advance of the minimum Gibbs energy so the relative thermal and Gibbs energy must be balanced through Boltzmann principles. The probability density for achieving an energy level G is then given by

$$\mu(G) = C e^{-GV/kT}, \quad (2.8)$$

where k is Boltzmann's constant, V is a reference volume and C is a constant that is selected so that when $\mu(G)$ is integrated over all possible dipole orientations, a probability of 1 is achieved. If we let 2σ be the separation between possible polarization states around P_0 , the probabilities of reaching a polarization state having sufficient energy to switch orientations from positive to negative, and conversely, are

$$r_{+-} = \frac{\int_{P_0-\sigma}^{P_0+\sigma} e^{-G(E,P)V/kT} dP}{\int_{P_0-\sigma}^{\infty} e^{-G(E,P)V/kT} dP}, \quad r_{-+} = \frac{\int_{P_0-\sigma}^{P_0+\sigma} e^{-G(E,P)V/kT} dP}{\int_{-\infty}^{P_0+\sigma} e^{-G(E,P)V/kT} dP}. \quad (2.9)$$

The rates at which dipoles switch from a positive to a negative orientation and conversely are then

$$p_{+-} = \frac{1}{\tau} r_{+-}, \quad p_{-+} = \frac{1}{\tau} r_{-+}, \quad (2.10)$$

where τ is the relaxation time. The fractions of dipoles in each orientation evolve according to the ordinary differential equations

$$\frac{dx_+}{dt} = -p_{+-}x_+ + p_{-+}x_- \quad , \quad \frac{dx_-}{dt} = -p_{-+}x_- + p_{+-}x_+ . \quad (2.11)$$

The expected polarizations due to positively and negatively oriented dipoles are

$$\langle P_+ \rangle = \int_{P_0+\sigma}^{\infty} P\mu(G)dP \quad , \quad \langle P_- \rangle = \int_{-\infty}^{P_0-\sigma} P\mu(G)dP \quad (2.12)$$

so the evaluation of C yields

$$\langle P_+ \rangle = \frac{\int_{P_0+\sigma}^{\infty} P e^{-G(E,P,T)V/kT} dP}{\int_{P_0+\sigma}^{\infty} e^{-G(E,P,T)V/kT} dP} \quad , \quad \langle P_- \rangle = \frac{\int_{-\infty}^{P_0-\sigma} P e^{-G(E,P,T)V/kT} dP}{\int_{-\infty}^{P_0-\sigma} e^{-G(E,P,T)V/kT} dP} . \quad (2.13)$$

For a single crystal with uniform effective field, the local average polarization is subsequently

$$\bar{P} = x_+ \langle P_+ \rangle + x_- \langle P_- \rangle . \quad (2.14)$$

In the manner detailed in [41], the evaluation of the integrals in (2.9) and (2.13) can be simplified through approximations employing the inflection points $\pm P_I$ rather than the unstable equilibrium P_0 .

Both of the relations (2.4) and (2.14) are valid only for homogeneous single crystal materials with uniform effective fields. To account for nonuniformity and inhomogeneities in the materials, local coercive and effective fields are assumed to be manifestations of underlying distributions rather than constants. The macroscopic polarization model is then given by

$$P(E) = \int_0^{\infty} \int_{-\infty}^{\infty} \bar{P}(E + E_e, E_c, \xi) \nu_1(E_c) \nu_2(E_e) dE_e dE_c , \quad (2.15)$$

where ν_1 and ν_2 are appropriate densities.

Motivated by choices in the magnetics literature, ν_1 and ν_2 were respectively designated to be lognormal and normal densities in [41]; thus

$$\nu_1(E_c) = C_1 e^{-\ln[(E_c/\overline{E_c})/2c]^2} \quad (2.16)$$

$$\nu_2(E_e) = C_2 e^{-(E-E_e)^2/b}. \quad (2.17)$$

However, the fact that neither of these choices is based on energy considerations, motivates the consideration of general densities. Furthermore, as long as the densities are required to have positive arguments for the coercive field and exhibit certain decay properties, there is no physical reason why normal and lognormal functions are inherently preferable to general densities, which provide much greater flexibility for model construction. Based on physical principles, we assume that the general densities ν_1 and ν_2 satisfy the following conditions

$$(i) \quad \nu_1(E_c) \text{ is defined for } E_c > 0 \quad (2.18)$$

$$(ii) \quad \nu_2(E_e) = \nu_2(-E_e) \quad (2.19)$$

$$(iii) \quad |\nu_1(E_c)| \leq c_1 e^{-a_1 E_c}, |\nu_2(E_e)| \leq c_2 e^{-a_2 E_e}, \quad (2.20)$$

where a_1 , a_2 , c_1 and c_2 are nonnegative. The first condition is required since the coercive field E_c must be positive. The second condition reflects the assumption that the effective field E_e is symmetric. The third condition represents the physical observation that the coercive and effective fields decay as a function of distance. This also ensures finite polarization values when integrating the densities against the hysteresis kernel.

2.2 Well-Posedness Analysis

For the general densities defined by conditions (2.18), (2.19), and (2.20), we can also consider a single joint density function, where the general densities $\nu_1(E_c)$ and $\nu_2(E_e)$ are replaced with the joint density $\nu(E_c, E_e) = \nu_1(E_c)\nu_2(E_e)$ so that (2.15) has the form

$$P(E) = \int_0^\infty \int_{-\infty}^\infty \bar{P}(E + E_e, E_c, \xi) \nu(E_c, E_e) dE_e dE_c. \quad (2.21)$$

Making use of condition (2.20) allows (2.21) to be approximated to arbitrary accuracy by

$$P(E) = \iint_{\Omega_2} \bar{P}(E + E_e, E_c, \xi) \nu(E_c, E_e) dE_e dE_c, \quad (2.22)$$

where Ω_2 is the compact domain $\{(E_c, E_e) \in \mathbb{R}_+ \times \mathbb{R} | \nu(E_c, E_e) \geq \epsilon\}$. Now let E_{\min} and E_{\max} be the minimum and maximum allowable input fields and define Ω_1 to be the domain $[E_{\min}, E_{\max}]$. Next define the parameters $q = \nu$ in the parameter space

$$\mathcal{Q} = L^2(\Omega_2) \quad (2.23)$$

and let $k = \bar{P}$. Then, let

$$E \in C[\Omega_1] \subset L^2(\Omega_1) \quad (2.24)$$

and define the observation operator $\mathcal{C}P = P(E)$ on the observation space

$$\mathcal{Y} = L^2(P_{\min}, P_{\max}). \quad (2.25)$$

Finally, define the parameter-to-observation operator \mathcal{K} by

$$\mathcal{K}q = \mathcal{C} \iint_{\Omega_2} k(\cdot + E_e, E_c) q(E_c, E_e) dE_c dE_e. \quad (2.26)$$

The polarization model (2.22) can subsequently be rewritten as

$$y(E) = \mathcal{K}q(E). \quad (2.27)$$

Now, define $\Omega = \Omega_1 \times \Omega_2$ and note that since k has an affine construction, $k \in L^1(\Omega) \cap L^2(\Omega)$. The first property is typical for convolution operators. The second property allows easier construction of a generalized Fourier basis, which will be used in showing that the operator \mathcal{K} is compact. Note that \mathcal{K} is a special case of a Hilbert-Schmidt operator and has kernel in L^2 . The proof that follows is a modification of the proof in [10] for Hilbert-Schmidt operators with kernels in $L^2(\mathbb{R}^{2n})$. We first summarize Theorem 5.24.8 from [22]:

Theorem 1 *Let X and Y be Banach spaces and let $\mathcal{K}_N : X \rightarrow Y, N = 1, 2, \dots$, be a sequence of compact linear operators converging to a bounded linear operator $\mathcal{K} : X \rightarrow Y$; that is, $\|\mathcal{K}_N - \mathcal{K}\| \rightarrow 0$ as $N \rightarrow \infty$. Then \mathcal{K} is a compact linear operator.*

Thus, to show that \mathcal{K} is a compact operator, it must be demonstrated that \mathcal{K} is the limit of a sequence of finite operators. The first step is to construct an orthonormal basis $\{\phi_i\}$ for $L^2(\Omega)$. It is demonstrated in [22] that an orthonormal basis for $L^2(\Omega_1)$ is

$$\varphi_\ell(s) = \frac{1}{\sqrt{E_{\max} - E_{\min}}} \exp \left[2\pi i \ell \cdot \frac{s - E_{\min}}{E_{\max} - E_{\min}} \right], \ell = 0, \pm 1, \pm 2, \dots \quad (2.28)$$

An orthonormal basis for $L^2(\Omega_2)$ can be created in an analogous manner. Thus, an orthonormal basis for $L^2(\Omega)$ is

$$\phi_{\ell m}(s, t, v) = \varphi_\ell(s) \overline{\varphi_m(t, v)}. \quad (2.29)$$

For simplicity, this is reindexed as $\{\phi_i\}$. Using this basis, every $f \in L^2(\Omega)$ can be

written in the generalized Fourier series form

$$f = \sum_i \langle f, \phi_i \rangle \phi_i. \quad (2.30)$$

From Plancherel's theorem, it follows that

$$\|f\|^2 = \sum_i |\langle f, \phi_i \rangle|^2. \quad (2.31)$$

Letting $\psi_i = \mathcal{K}\phi_i$, allows \mathcal{K} and the approximating finite-rank operators \mathcal{K}_N to be represented by

$$\mathcal{K}f = \sum_i \langle f, \phi_i \rangle \psi_i \quad (2.32)$$

$$\mathcal{K}_N f = \sum_{i=1}^N \langle f, \phi_i \rangle \psi_i. \quad (2.33)$$

Then,

$$\begin{aligned} \sum_i \|\psi_i\|^2 &= \sum_i \left[\int_{\Omega_1} |\mathcal{K}\phi_i(E)|^2 dE \right] \\ &= \sum_i \int_{\Omega_1} \left| \iint_{\Omega_2} k(E + E_e, E_c) \phi_i(E_c, E_e) dE_e dE_c \right|^2 dE \\ &= \int_{\Omega_1} \left[\sum_i \left| \iint_{\Omega_2} k(E + E_e, E_c) \phi_i(E_c, E_e) dE_e dE_c \right|^2 \right] dE \\ &= \int_{\Omega_1} \left[\iint_{\Omega_2} |k(E + E_e, E_c)|^2 dE_e dE_c \right] dE < \infty, \end{aligned} \quad (2.34)$$

where the last step follows using Plancherel's theorem. There thus exists L such that

$$\sum_i \|\psi_i\|^2 \rightarrow L \quad (2.35)$$

and thus

$$\sum_{i \geq N+1} \|\psi_i\|^2 \rightarrow 0, \text{ as } N \rightarrow \infty. \quad (2.36)$$

Therefore, for $\varepsilon > 0$, there exists N_ε such that for $N > N_\varepsilon$,

$$\|\mathcal{K} - \mathcal{K}_N\| = \sup_{f \neq 0} \frac{\|\mathcal{K}f - \mathcal{K}_Nf\|}{\|f\|} < \varepsilon. \quad (2.37)$$

From this it can be concluded that

$$\lim_{N \rightarrow \infty} \|\mathcal{K} - \mathcal{K}_N\| = 0 \quad (2.38)$$

and since the range of \mathcal{K}_N is finite, \mathcal{K}_N is a compact operator.

Now we show the convergence $\mathcal{K} \rightarrow \mathcal{K}_N$. Using the Schwartz inequality, note that

$$\begin{aligned} \|\mathcal{K}f - \mathcal{K}_Nf\| &= \left\| \sum_{i \geq N+1} \langle f, \phi_i \rangle \psi_i \right\| \\ &\leq \sum_{i \geq N+1} |\langle f, \phi_i \rangle| \|\psi_i\| \\ &\leq \left[\sum_{i \geq N+1} |\langle f, \phi_i \rangle|^2 \right]^{\frac{1}{2}} \left[\sum_{i \geq N+1} \|\psi_i\|^2 \right]^{\frac{1}{2}} \\ &\leq \|f\| \left[\sum_{i \geq N+1} \|\psi_i\|^2 \right]^{\frac{1}{2}}. \end{aligned} \quad (2.39)$$

Thus by Theorem 1, \mathcal{K} is a compact operator since it is the norm limit of a sequence of compact operators.

2.3 Parameter Estimation Problem

Given data $(\widehat{E}, \widehat{P})$, where $\widehat{E} \in L^2(E_{\min}, E_{\max})$, the parameter estimation problem is to find $q \in \mathcal{Q}$ such that

$$\mathcal{K}q = \widehat{P}. \quad (2.40)$$

There is a classical solution to this problem if and only if $\widehat{P} \in \mathcal{R}(\mathcal{K})$, where $\mathcal{R}(\mathcal{K})$ is the range of \mathcal{K} . Since this will not be true in general, the least squares formulation

$$\min_{q \in \mathcal{Q}} T(q), \quad T(q) = \frac{1}{2} \|\mathcal{K}q - \widehat{P}\|^2 \quad (2.41)$$

is typically more appropriate. However, even this problem is ill-posed since the fact that \mathcal{K} is compact with infinite-dimensional range makes the Moore-Penrose inverse \mathcal{K}^\dagger discontinuous [7]. One way to deal with this difficulty is to solve the regularized problem

$$\min_{q \in \mathcal{Q}} T_\alpha(q), \quad T_\alpha(q) = \frac{1}{2} \|\mathcal{K}q - \widehat{P}\|^2 + \alpha \mathcal{J}(q), \quad (2.42)$$

where the additional term $\alpha \mathcal{J}(q)$ provides stability. As the regularization parameter $\alpha > 0$ is increased, stability increases as well, but at the cost of reducing the quality of the characterization of the data. One choice for \mathcal{J} is the Tikhonov functional

$$\mathcal{J}(q) = \frac{1}{2} \|q\|^2. \quad (2.43)$$

Implementation of (2.42) with Tikhonov regularization for data collected on PZT5H is discussed in Section 3.3.3.

2.4 Coupled Electromechanical Constitutive Relations

To include ferroelastic coupling, we utilize the extended Helmholtz relation

$$\psi_e(P, \varepsilon) = \psi(P) + \frac{1}{2}Y^P \varepsilon^2 - Y^P \gamma \varepsilon P, \quad (2.44)$$

where $\psi(P)$ is given by (2.1). The Gibbs energy is then

$$G(E, P, \varepsilon) = \psi(P) + \frac{1}{2}Y^P \varepsilon^2 - Y^P \gamma \varepsilon P - EP - \sigma \varepsilon, \quad (2.45)$$

where the term $\sigma \varepsilon$ incorporates the elastic energy. Note that Y^P is the Young's modulus for a constant polarization and γ is a ferroelastic coupling coefficient.

The equilibrium condition

$$\frac{\partial G}{\partial \varepsilon} = 0 \quad (2.46)$$

yields the elastic constitutive relation

$$\sigma = Y^P \varepsilon - Y^P \gamma P. \quad (2.47)$$

This relation along with the nonlinear polarization relation (2.15) quantifies the constitutive behavior for the piezoceramic materials employed in the AFM.

When a PZT shell is employed for actuation instead of a rod, longitudinal strains are generated by d_{31} rather than d_{33} mechanisms. Furthermore, as described in [1], longitudinal and circumferential stresses and strains are coupled due to the curvature.

The 2-D constitutive relations for the shell are

$$\begin{aligned}\sigma_x &= \frac{Y^P}{1-\nu^2}(e_x + \nu e_\theta) - \frac{Y^P\beta}{1-\nu}P(E) \\ \sigma_\theta &= \frac{Y^P}{1-\nu^2}(e_\theta + \nu e_x) - \frac{Y^P\beta}{1-\nu}P(E) \\ \sigma_{x\theta} &= \sigma_{\theta x} = \frac{Y^P}{2(1+\nu)}e_{x\theta}\end{aligned}\tag{2.48}$$

with the relation for $P(E)$ given by (2.15). Here, ε_x , σ_x , ε_θ and σ_θ represent the normal strains and stresses in the longitudinal and circumferential directions, while the shear strains and stresses are denoted by $e_{x\theta}$ and $\sigma_{x\theta}$. Also, ν is the Poisson ratio for the material and β is the electromechanical coupling coefficient.

2.5 Actuator Models

In addition to hysteresis, the dynamics of the actuator must be incorporated. Stacked actuators, which create motion in only one direction, are employed in pairs as in Figure 1.3 so that the stage can be moved in both the x and y directions. The cylindrical actuator is poled to generate movement in both directions by itself.

2.5.1 Stacked Actuator

For the stacked actuator, we assume that the rod has cross-sectional area A , length ℓ , density ρ and Young's modulus Y^P . Let c^P be the Kelvin-Voigt damping parameter and γ be the piezoelectric coupling coefficient. In the present stage design, depicted in Figure 1.3, one end of the actuator is fixed, while the attachment at the other end can be modeled as a damped spring-mass system. For this end, let M_ℓ , k_ℓ and

c_ℓ respectively denote the mass, stiffness and damping coefficients. Force balancing along the actuator then yields the relation

$$\rho A \frac{\partial^2 u}{\partial t^2} = \frac{\partial \mathcal{N}}{\partial x}, \quad (2.49)$$

where the resultant $\mathcal{N} = \int_A \sigma dA$ is given by

$$\mathcal{N} = Y^P A \frac{\partial u}{\partial x} + c^P A \frac{\partial^2 u}{\partial x \partial t} - Y^P A \gamma P(E). \quad (2.50)$$

Note that u denotes the displacement in the longitudinal x -direction and that the relation strain $\varepsilon = \frac{\partial u}{\partial x}$ is used in obtaining (2.50). The boundary conditions are

$$u(t, 0) = 0 \quad (2.51)$$

at the fixed end and

$$\mathcal{N}(t, \ell) = -k_\ell u(t, \ell) - c_\ell \frac{\partial u}{\partial t}(t, \ell) - M_\ell \frac{\partial^2 u}{\partial t^2}(t, \ell) \quad (2.52)$$

at the moving end. Initial conditions are given by

$$u(0, x) = \frac{\partial u}{\partial t}(0, x) = 0. \quad (2.53)$$

The polarization $P(E)$ is specified by (2.15).

To provide a setting that facilitates approximation, we also derive a weak formulation as described below in a manner similar to that in [39]. Let the states $z = (u(\cdot), u(\ell))$ be in the space $X = L^2(0, \ell) \times \mathbb{R}$, where the inner product is defined by

$$\langle \xi_1, \xi_2 \rangle_X = \int_0^\ell \rho A \xi_1 \xi_2 dx + M_\ell \xi_1(\ell) \xi_2(\ell) \quad (2.54)$$

and let the space of test functions be

$$V = \{\Psi = (\psi, \varphi) \in X \mid \psi \in H^1(0, \ell), \psi(0) = 0, \psi(\ell) = \varphi\}, \quad (2.55)$$

where the inner product is defined by

$$\langle \Psi_1, \Psi_2 \rangle_V = \int_0^\ell Y^P A \psi_1' \psi_2' dx + k_\ell \psi_1(\ell) \psi_2(\ell). \quad (2.56)$$

Multiplying (2.49) by $\psi \in H_0^1(0, \ell) = \{\psi \in H^1(0, \ell) \mid \psi(0) = 0\}$ and integrating gives

$$\int_0^\ell \rho A \frac{\partial^2 u}{\partial t^2} \psi dx = \int_0^\ell \frac{\partial \mathcal{N}}{\partial x} \psi dx. \quad (2.57)$$

Then, integrating the right side by parts,

$$\begin{aligned} \int_0^\ell \rho A \frac{\partial^2 u}{\partial t^2} \psi dx &= \psi \mathcal{N} \Big|_0^\ell - \int_0^\ell \mathcal{N} \frac{\partial \psi}{\partial x} dx \\ &= \psi(\ell) \mathcal{N}(t, \ell) - \psi(0) \mathcal{N}(t, 0) - \int_0^\ell \mathcal{N} \frac{\partial \psi}{\partial x} dx \\ &= \psi(\ell) \mathcal{N}(t, \ell) - \int_0^\ell \mathcal{N} \frac{\partial \psi}{\partial x} dx, \end{aligned}$$

where the fact that $\psi(0) = 0$ is used in the last step. Substituting for \mathcal{N} from (2.50)

and for $\mathcal{N}(t, \ell)$ from (2.52) yields

$$\begin{aligned} \int_0^\ell \rho A \frac{\partial^2 u}{\partial t^2} \psi dx &= \psi(\ell) \left[-k_\ell u(t, \ell) - c_\ell \frac{\partial u}{\partial t}(t, \ell) - M_\ell \frac{\partial^2 u}{\partial t^2}(t, \ell) \right] \\ &\quad - \int_0^\ell \left[Y^P A \frac{\partial u}{\partial x} + c^P A \frac{\partial^2 u}{\partial x \partial t} - Y^P A \beta P(E) \right] \frac{\partial \psi}{\partial x} dx. \end{aligned}$$

Rearranging terms gives the weak form

$$\begin{aligned} \int_0^\ell \rho A \frac{\partial^2 u}{\partial t^2} \psi dx + \psi(\ell) M_\ell \frac{\partial^2 u}{\partial t^2}(t, \ell) + \int_0^\ell \left[Y^P A \frac{\partial u}{\partial x} + c^P A \frac{\partial^2 u}{\partial x \partial t} \right] \frac{\partial \psi}{\partial x} dx \\ + \psi(\ell) \left[k_\ell u(t, \ell) + c_\ell \frac{\partial u}{\partial t}(t, \ell) \right] = \int_0^\ell Y^P A \beta P(E) \frac{\partial \psi}{\partial x} dx, \end{aligned} \quad (2.58)$$

which must hold for all $\psi \in V$. An approximation to the weak form (2.58) is found by discretizing in space with linear finite elements and in time with finite difference techniques as described in Section 3.4.

2.5.2 Cylindrical Actuator

Consider a cylindrical actuator composed of two shells stacked one on top of the other as depicted in Figure 2.2. The first shell creates displacements in the x - y direction while the second shell actuates in the z direction. Motions in the latter direction are used to maintain constant force between the sample and the micro-cantilever on the AFM. When analyzing the dynamics for this case, the mass of the x - y shell is combined with the mass of the sample to yield a total mass m , which acts as an inertial force at the free end of the z shell.

Assume that the z shell has length ℓ , thickness h , radius R , density ρ and Young's modulus Y^P . Let $\Gamma_0 = [0, \ell] \times [0, 2\pi]$ represent the reference or neutral surface of the z shell. The longitudinal, circumferential and transverse displacements are denoted by u , v and w . Consider the case where the bottom of the z shell ($x = 0$) is fixed, while the top ($x = \ell$) is free. Thus the motion of the top depends only on the inertial force of the combined mass m of the x - y shell and the sample. In the model development which follows, internal damping is assumed to be absent. However, descriptions of how to include Kelvin-Voigt damping by assuming that stress is a linear combination of strain, strain rate and polarization can be found in [1, 5, 39].

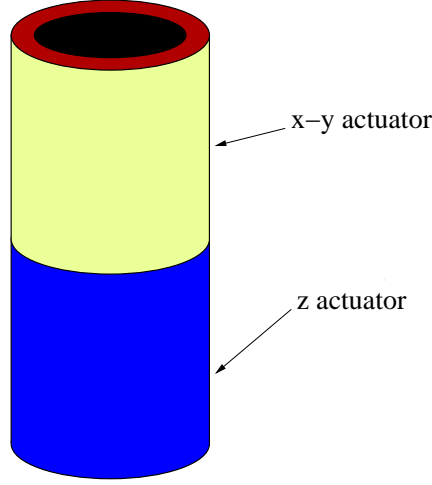


Figure 2.2: Cylindrical actuator.

The derivation of the Donnell-Mushtari shell equations

$$\begin{aligned}
 R\rho h \frac{\partial^2 u}{\partial t^2} - R \frac{\partial \mathcal{N}_x}{\partial x} - \frac{\partial \mathcal{N}_{x\theta}}{\partial \theta} &= 0 \\
 R\rho h \frac{\partial^2 v}{\partial t^2} - \frac{\partial \mathcal{N}_\theta}{\partial \theta} - R \frac{\partial \mathcal{N}_{x\theta}}{\partial x} &= 0
 \end{aligned} \tag{2.59}$$

$$R\rho h \frac{\partial^2 u}{\partial t^2} - R \frac{\partial \mathcal{M}_x}{\partial x^2} - \frac{1}{R} \frac{\partial^2 \mathcal{M}_\theta}{\partial \theta^2} - 2 \frac{\mathcal{M}_{x\theta}}{\partial x \partial \theta} + \mathcal{N}_\theta = 0$$

from force and moment balancing is described in [1]. Here, \mathcal{N}_x , \mathcal{N}_θ and $\mathcal{N}_{x\theta}$ are general force resultants and \mathcal{M}_x , \mathcal{M}_θ and $\mathcal{M}_{x\theta}$ are moment resultants. Boundary conditions are enforced at $x = 0$ and $x = \ell$. For $x = 0$

$$u = v = w = \frac{\partial w}{\partial x} = 0 \tag{2.60}$$

and for $x = \ell$

$$\mathcal{N}_x = -m \frac{\partial^2 u}{\partial t^2} \quad \mathcal{N}_{x\theta} + \frac{1}{R} \mathcal{M}_{x\theta} = 0$$

$$\mathcal{Q}_x + \frac{1}{R} \frac{\partial \mathcal{M}_{x\theta}}{\partial \theta} = 0 \quad \mathcal{M}_x = 0,$$

where the first condition at $x = \ell$ incorporates the inertial force from the combined mass m . To determine the force and moment resultants when there are no shear stresses, the stress relations (2.48) are integrated through the thickness of the shell to give

$$\mathcal{N}_x = \frac{Y^P h}{1 - \nu^2} (e_x + \nu e_\theta) - \frac{Y^P h \beta}{1 - \nu} P(E) \quad (2.61)$$

$$\mathcal{N}_\theta = \frac{Y^P h}{1 - \nu^2} (e_\theta + \nu e_x) - \frac{Y^P h \beta}{1 - \nu} P(E) \quad (2.62)$$

$$\mathcal{N}_{x\theta} = \frac{Y^P h}{2(1 + \nu)} e_{x\theta} \quad (2.63)$$

and

$$\mathcal{M}_x = \frac{Y^P h^3}{12(1 - \nu^2)} (\kappa_x + \nu \kappa_\theta) \quad (2.64)$$

$$\mathcal{M}_\theta = \frac{Y^P h^3}{12(1 - \nu^2)} (\kappa_\theta + \nu \kappa_x) \quad (2.65)$$

$$\mathcal{M}_{x\theta} = \frac{Y^P h^3}{24(1 + \nu)} \tau. \quad (2.66)$$

In addition, the midsurface strains and changes in curvature are

$$\begin{aligned}
 e_x &= \frac{\partial u}{\partial x} & \kappa_x &= -\frac{\partial^2 w}{\partial x^2} \\
 e_\theta &= \frac{1}{R} \left(\frac{\partial v}{\partial \theta} + w \right) & \kappa_\theta &= -\frac{1}{R^2} \frac{\partial^2 w}{\partial \theta^2} \\
 e_{x\theta} &= \frac{\partial v}{\partial x} + \frac{1}{R} \frac{\partial u}{\partial \theta} & \tau &= -\frac{2}{R} \frac{\partial^2 w}{\partial x \partial \theta}.
 \end{aligned}$$

An approximate solution to (2.59) is found by first deriving a weak form as described in [39] and summarized here. First, let the state space be

$$X = L^2(\Gamma_0) \times L^2(\Gamma_0) \times L^2(\Gamma_0). \quad (2.67)$$

Then let $\phi = (u, v, w)$ and $\psi = (\eta_1, \eta_2, \eta_3)$ and define the inner product

$$\langle \phi, \psi \rangle_X = \int_{\Gamma_0} \rho h u \overline{\eta_1} d\gamma + \int_{\Gamma_0} \rho h v \overline{\eta_2} d\gamma + \int_{\Gamma_0} \rho h w \overline{\eta_3} d\gamma + m u \overline{\eta_1}(\ell). \quad (2.68)$$

Next, define

$$\begin{aligned}
 H_\ell^1(\Gamma_0) &= \{\eta \in H^1(\Gamma_0) | \eta(0, \theta) = 0\} \\
 H_\ell^2(\Gamma_0) &= \{\eta \in H^2(\Gamma_0) | \eta(0, \theta) = \eta_x(0, \theta) = 0\}
 \end{aligned}$$

and let the space of test functions be

$$V = H_\ell^1(\Gamma_0) \times H_\ell^1(\Gamma_0) \times H_\ell^2(\Gamma_0). \quad (2.69)$$

Define the inner product on this space as

$$\begin{aligned}
\langle (Y^P)\phi, \psi \rangle_V &= \int_{\Gamma_0} \frac{Y^P h}{1 - \nu^2} \left[(e_x + \nu e_\theta) \frac{\overline{\partial \eta_1}}{\partial x} + \frac{1}{2R} (1 - \nu) e_{x\theta} \frac{\overline{\partial \eta_1}}{\partial \theta} \right] d\gamma \\
&+ \int_{\Gamma_0} \frac{Y^P h}{1 - \nu^2} \left[(e_\theta + \nu e_x) \frac{\overline{\partial \eta_2}}{\partial \theta} + \frac{1}{2R} (1 - \nu) e_{x\theta} \frac{\overline{\partial \eta_2}}{\partial x} \right] d\gamma \\
&+ \int_{\Gamma_0} \frac{Y^P h}{1 - \nu^2} \left[\frac{1}{R} (e_\theta + \nu e_x) \overline{\eta_3} - \frac{h^2}{12} (\kappa_x + \nu \kappa_\theta) \frac{\overline{\partial^2 \eta_3}}{\partial x^2} \right. \\
&\quad \left. - \frac{h^2}{12R^2} (\kappa_\theta + \nu \kappa_x) \frac{\overline{\partial^2 \eta_3}}{\partial \theta^2} - \frac{h^2}{12R} (1 - \nu) \tau \frac{\overline{\partial^2 \eta_3}}{\partial x \partial \theta} \right] d\gamma.
\end{aligned} \tag{2.70}$$

The weak form of the model (2.59) can then be written as

$$\begin{aligned}
&\int_{\Gamma_0} \left[R\rho h \frac{\partial^2 u}{\partial t^2} \overline{\eta_1} + RN_x \frac{\overline{\partial \eta_1}}{\partial x} + N_{x\theta} \frac{\overline{\partial \eta_1}}{\partial \theta} - R \frac{Y^P h \beta}{1 - \nu} P(E) \frac{\overline{\partial \eta_1}}{\partial x} \right] d\gamma \\
&\quad - Rm \frac{\partial^2 u}{\partial t^2} \overline{\eta_1}(\ell) = 0 \\
&\int_{\Gamma_0} \left[R\rho h \frac{\partial^2 v}{\partial t^2} \overline{\eta_2} + N_\theta \frac{\overline{\partial \eta_2}}{\partial \theta} + RN_{x\theta} \frac{\overline{\partial \eta_2}}{\partial x} - \frac{Y^P h \beta}{1 - \nu} P(E) \frac{\overline{\partial \eta_2}}{\partial \theta} \right] d\gamma = 0 \\
&\int_{\Gamma_0} \left[R\rho h \frac{\partial^2 w}{\partial t^2} \overline{\eta_3} + N_\theta \overline{\eta_3} - RM_x \frac{\overline{\partial^2 \eta_3}}{\partial x^2} - \frac{1}{R} M_\theta \frac{\overline{\partial^2 \eta_3}}{\partial \theta^2} - 2M_{x\theta} \frac{\overline{\partial^2 \eta_3}}{\partial x \partial \theta} \right] d\gamma
\end{aligned} \tag{2.71}$$

for all $\psi = (\eta_1, \eta_2, \eta_3) \in V$. An approximation to the weak form (2.71) is found by discretizing in space with a tensored basis comprised of cubic B splines and Fourier elements and in time with finite difference techniques as described in [39] and summarized in Section 3.5.

Chapter 3

Approximation Techniques and Implementation Algorithms

In this chapter, we begin by discussing approximation techniques and implementation algorithms for the field-polarization models described in Chapter 2. The first approach uses constrained optimization with a least squares cost function to determine the general density parameters and η . In the second method, the general densities are combined into a single joint density and the resulting linear system is solved using the techniques employed in [43] for identifying density functions in Preisach models. In the final approach, the constrained minimization problem for the joint density is modified to include Tikhonov regularization to accommodate the ill-posedness of the original formulation. All three methods are validated with a PZT5H data set collected at 0.2 Hz. Then an inverse algorithm quantifying the polarization-field relation is developed for use in control design. Section 3.3 addresses parameter estimation techniques for the field-polarization algorithms from Section 3.1. Finally, approximation techniques and implementation algorithms for the rod and shell actuator models are outlined.

3.1 Approximation Techniques for Polarization

Model

To implement the nonlinear polarization model (2.15), the integrals must be approximated. One approach is to use Gaussian quadrature routines constructed for infinite or semi-infinite domains. Alternatively, the decay of the densities ν_1 and ν_2 allows the intervals to be truncated so that Gauss-Legendre rules can be employed. Both cases are described in detail in [41]. In either case, the discrete model will be of the form

$$P(E) = \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \bar{P}(E + E_{e_j}, E_{c_i}, \xi_i) \nu_1(E_{e_j}) \nu_2(E_{c_i}) v_i w_j, \quad (3.1)$$

where E_{e_j}, E_{c_i} are the abscissas and v_i, w_j are the weights. For the characterization examples in the sections which follow, we use a 4 point composite Gauss-Legendre quadrature rule for truncated domains. For the integral over E_e let the truncated domain be $[-L, L]$ and let N_{q_j} be the number of subintervals for the composite Gaussian quadrature. The subintervals are then $[h_{q_j-1}, h_{q_j}]$, where $h_{q_j} = -L + q_j h$ with $q_j = 1 \dots N_{q_j}$ and $h = 2L/N_{q_j}$. If we use a four point rule, $N_j = 4N_{q_j}$ and the points and weights on each subinterval are

$$\begin{aligned} \varepsilon_{q_j 1} &= h_{q_j-1} + h \left[\frac{1}{2} - \frac{\sqrt{15 + 2\sqrt{30}}}{2\sqrt{35}} \right], & w_{q_j 1} &= \frac{49h}{12(18 + \sqrt{30})} \\ \varepsilon_{q_j 2} &= h_{q_j-1} + h \left[\frac{1}{2} - \frac{\sqrt{15 - 2\sqrt{30}}}{2\sqrt{35}} \right], & w_{q_j 2} &= \frac{49h}{12(18 - \sqrt{30})} \\ \varepsilon_{q_j 3} &= h_{q_j-1} + h \left[\frac{1}{2} + \frac{\sqrt{15 - 2\sqrt{30}}}{2\sqrt{35}} \right], & w_{q_j 3} &= \frac{49h}{12(18 - \sqrt{30})} \\ \varepsilon_{q_j 4} &= h_{q_j-1} + h \left[\frac{1}{2} + \frac{\sqrt{15 + 2\sqrt{30}}}{2\sqrt{35}} \right], & w_{q_j 4} &= \frac{49h}{12(18 + \sqrt{30})}. \end{aligned}$$

For the integral over E_c let the truncated domain be $[0, K]$ and let N_{q_i} be the number of subintervals. Here, the subintervals are $[g_{q_i-1}, g_{q_i}]$, where $g_{q_i} = q_i g$ with $q_i = 1 \dots N_{q_i}$ and $g = K/N_{q_i}$. Again using a four point rule, we have $N_i = 4N_{q_i}$ and analogous nodes and weights. If the joint density formulation (2.21) is used, the discretized version becomes

$$P(E) = \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \bar{P}(E + E_{e_j}, E_{c_i}, \xi_i) \nu(E_{e_j}, E_{c_i}) v_i w_j. \quad (3.2)$$

The number of parameters required for the construction of (3.1) depends on the choice of densities. For the lognormal and normal densities there are five parameters; η , C , $\overline{E_c}$, c and b , where $C = C_1 C_2$. Note that for any values of η , C and P_R , rescaling allows P_R to be set to 1. Thus, P_R does not need to be identified. For the general densities the $N_i + N_j + 1$ parameters $[\nu_1(E_{c_1}), \dots, \nu_1(E_{c_{N_i}})]$, $[\nu_2(E_{e_1}), \dots, \nu_2(E_{e_{N_j}})]$ and η need to be identified. The joint density ν requires the identification of $N_i N_j + 1$ parameters. The larger number of parameters in the general density and joint density formulations allows the construction of models which are highly accurate for a wide range of drive regimes. The trade off is the increased time and effort required to identify the parameters. Numerical simulations have indicated that N_i and N_j must be on the order of 40 to 80 to achieve convergence of (3.1). Therefore, highly efficient optimization and parameter estimation techniques are required. However, implementation of the model in control design is equally efficient for both the lognormal and normal densities and general densities since each involves the multiplication of $N_i \times 1$ and $N_j \times 1$ vectors.

An implementation of the forward model (3.1) based on the description in [41] is now summarized. Formulate the local polarization (2.4) as

$$\bar{P} = \frac{E}{\eta} + P_R \Delta, \quad (3.3)$$

where Δ is an $N_i \times N_j$ matrix whose elements are either 1 or -1 . The ij th entry in Δ indicates whether the j th effective field value E_{e_j} has crossed the i th coercive field value E_{c_i} and thus whether the associated polarization value is on the upper or lower branch of the hysteron. As the field is changed, the entries in Δ can be updated either by using *if-then* statements or by employing the following matrix multiplication algorithm. First define the matrices,

$$\begin{aligned} \Delta_{\text{init}} &= \begin{bmatrix} -1 & \cdots & -1 & 1 & \cdots & 1 \\ \vdots & & \vdots & \vdots & & \vdots \\ -1 & \cdots & -1 & 1 & \cdots & 1 \end{bmatrix}_{N_i \times N_j} & \varepsilon_c &= \begin{bmatrix} E_{c_1} & \cdots & E_{c_1} \\ \vdots & & \vdots \\ E_{c_{N_i}} & \cdots & E_{c_{N_i}} \end{bmatrix}_{N_i \times N_j} \\ \varepsilon_k &= \begin{bmatrix} E_k + E_{e_1} & \cdots & E_k + E_{e_{N_j}} \\ \vdots & & \vdots \\ E_k + E_{e_1} & \cdots & E_k + E_{e_{N_j}} \end{bmatrix}_{N_i \times N_j} & \mathcal{O} &= \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{N_i \times N_j} \end{aligned} \quad (3.4)$$

and the weight vectors

$$V^T = \left[v_1 \nu_1 (E_{c_1}), \dots, v_{N_i} \nu_1 (E_{c_{N_i}}) \right]_{1 \times N_i} \quad (3.5)$$

$$W^T = \left[w_1 \nu_2 (E_{e_1}), \dots, w_{N_j} \nu_2 (E_{e_{N_j}}) \right]_{1 \times N_j}, \quad (3.6)$$

where E_k is the k th value of the input field. The calculation of P_k is summarized in Algorithm 1, where \cdot^* denotes componentwise matrix multiplication and sgn is the signum function.

Algorithm 1. *Forward Algorithm (matrix multiplication version).*

```

 $\Delta = \Delta_{\text{init}}$ 
for  $k = 2 : N_k$ 
   $dE = E_k - E_{k-1}$ 
   $\varepsilon_k = \varepsilon_k + dE$ 
  if  $dE > 0$ 
     $\Delta = \left(\frac{\Delta+1}{2}\right) - \left(\frac{\Delta-1}{2}\right) \cdot * \text{sgn}(\varepsilon_k - \varepsilon_c)$ 
  else
     $\Delta = \left(\frac{\Delta+1}{2}\right) \cdot * \text{sgn}(\varepsilon_k - \varepsilon_c) - \left(\frac{\Delta-1}{2}\right)$ 
  end
   $\bar{P} = \frac{1}{\eta} \varepsilon_k + P_R \Delta$ 
   $P_k = V^T \bar{P} W$ 
end

```

Depending on the programming language, Algorithm 1 can be significantly faster than using if-then statements and has the advantage of being easy to code. However, both of these methods update the state of all of the hysteresis kernels for every change in dE . Also, there is no clear way to directly invert (3.3) to give a relation $E(P)$.

To construct a model inverse to be employed as filter in the robust control design, we first describe an alternative implementation of the forward model (3.1). Under reasonable assumptions about the data, the states of most of the hysteresis kernels will not change for a given step of dE . First, assume that there are not large changes in dE for a given step. Second, assume that there are not frequent changes in the sign of dE relative to the number of data points. In the case when these assumptions hold, we propose a modified algorithm to take advantage of the fact that only a small number of the entries in Δ change at each step. Updating only those entries in Δ that need to be updated allows for even faster computational times. Note however that

even if the above assumptions do not hold, the algorithm will still work, but may not be faster than Algorithm 1. The point at which this occurs is difficult to determine *a priori* since it depends on the data and this is the reason for the impreciseness of the assumptions.

To begin, define the vectors

$$\Upsilon_n = -\text{vec}(\varepsilon_k - \varepsilon_c) \quad (3.7)$$

$$\Upsilon_p = -\text{vec}(\varepsilon_k + \varepsilon_c), \quad (3.8)$$

where ‘vec’ is the column vector concatenation of a matrix and thus Υ_n and Υ_p are $N_i N_j \times 1$ vectors. Next, define

$$\Omega_W = \begin{bmatrix} W_1 & \cdots & W_1 \\ \vdots & & \vdots \\ W_{N_i} & \cdots & W_{N_i} \end{bmatrix}_{N_i \times N_j}, \quad \Omega_V = \begin{bmatrix} V_1 & \cdots & V_{N_j} \\ \vdots & & \vdots \\ V_1 & \cdots & V_{N_j} \end{bmatrix}_{N_i \times N_j} \quad (3.9)$$

and

$$P_{vals} = \text{vec}(P_R \cdot * \Omega_W \cdot * \Omega_V). \quad (3.10)$$

Note that the entries in P_{vals} are the contributions to the polarization from each hysteresis kernel. Now define

$$\varepsilon_n = \#sort \begin{bmatrix} \Upsilon_n & P_{vals} & \Delta_{row} & \Delta_{col} \end{bmatrix}_{N_i N_j \times 4} \quad (3.11)$$

$$\varepsilon_p = \#sort \begin{bmatrix} \Upsilon_p & P_{vals} & \Delta_{row} & \Delta_{col} \end{bmatrix}_{N_i N_j \times 4}, \quad (3.12)$$

where $\#sort$ sorts the rows of its argument matrix from smallest to largest based on the first entry in each row and Δ_{row} and Δ_{col} are the row and columns indices of Δ

that correspond to the entries of P_{vals} . Finally, let

$$\Delta_0 = \begin{bmatrix} 1 & \cdots & 1 \\ \vdots & & \vdots \\ 1 & \cdots & 1 \end{bmatrix}_{N_i \times N_j} \quad (3.13)$$

and define

$$\varphi = \frac{W\Delta_0V^T}{\eta}. \quad (3.14)$$

The first column of ε_n contains the values of E that must be surpassed in order for each hysteresis kernel to switch from its lower branch to its upper branch. The corresponding change in P is contained in the second column. The situation is analogous for ε_p with the switches being from the upper branch to the lower branch. In order to illustrate the general idea of the algorithm, consider the case when $dE > 0$ and assume that the value of P_k is known. Then find each value in the first column of ε_n which is greater than E_k and less than E_{k+1} . Next update the corresponding entries of Δ if necessary. Finally, add the appropriate increase in polarization to calculate P_{k+1} . There are two parts to this: the increase from the switches and the linear increase. In the practical implementation, we only keep track of the values of E and P at which the switch occurs and then solve a linear equation to get P_{k+1} once the correct point has been reached. The situation is analogous for the case when $dE < 0$. The next question is what to do when dE changes sign. To illustrate, consider the case when dE changes from positive to negative. Then, we must search through ε_p to find the first value that is less than E_k . This tells us the first hysteresis kernel that will switch from the upper branch to the lower branch (assuming it is not already on the lower branch). Once a linear adjustment to the current value of P is made, we can continue with the general case for $dE < 0$. There are two special situations to consider. First, how do we find the value of P_2 ? This is really the same

situation as when dE changes sign. The type of sign change depends on whether E_2 is positive or negative. Second, suppose $dE > 0$ and E is large enough that all of the hysteresis kernels are on the upper branch. Then, once the last switching value has been reached, P_{k+1} can be found by solving a linear equation. An outline of this process is summarized in Algorithm 2.

Algorithm 2. *General Outline of New Forward Algorithm.*

$$\bar{P} = \varepsilon_k / \eta + P_R \Delta$$

$$P_1 = CW^T \bar{P} V$$

if $dE > 0$

Set $E_{prev} = 0$ and determine initial values of E_{cur} and P_{cur} using ε_n

else

Set $E_{prev} = 0$ and determine initial values of E_{cur} and P_{cur} using ε_p

end

for $k = 2 : N_k$

if $dE > 0$

Update E_{prev} , E_{cur} , P_{cur} and Δ based on E_k and ε_n .

Calculate P_k

If dE changes sign, update E_{prev} , E_{cur} and P_{cur} using ε_p

else

Update E_{prev} , E_{cur} , P_{cur} and Δ based on E_k and ε_p .

Calculate P_k

If dE changes sign, update E_{prev} , E_{cur} and P_{cur} using ε_n

end

end

Algorithm 3 summarizes details for the case $dE > 0$ before the main loop. The situation for $dE < 0$ is analogous.

Algorithm 3. $dE > 0$ initially

$$\begin{aligned} E_{prev} &= 0 \\ E_{cur} &= \min(\varepsilon_n(:, 1)) > E_{prev} \\ P_{cur} &= P_1 + \varphi(E_{cur} - E_{prev}) \end{aligned}$$

Algorithm 4 provides details for the case $dE > 0$ inside the main loop. The situation for $dE < 0$ is analogous. The notation $\Delta[E_{cur}]$ refers to the value of Δ for the hysteresis kernel whose switching value is E_{cur} . Similarly, $P_{val}[E_{cur}]$ is the change in polarization that occurs when E becomes greater than E_{cur} .

Algorithm 4. $dE > 0$ inside main loop

$$\begin{aligned} &\text{while } E_{cur} < E_k \text{ and } E_{ind} < N_q \\ &\quad \text{if } \Delta[E_{cur}] = -1 \\ &\quad\quad P_{cur} = P_{cur} + 2P_{val}[E_{cur}] \\ &\quad\quad \Delta[E_{cur}] = 1 \\ &\quad \text{end} \\ &\quad E_{prev} = E_{cur} \\ &\quad E_{ind} = E_{ind} + 1 \\ &\quad E_{cur} = \varepsilon_n(E_{ind}, 1) \\ &\quad P_{cur} = P_{cur} + \varphi(E_{cur} - E_{prev}) \\ &\text{end} \\ &\text{if } E_{cur} < E_k \text{ and } E_{ind} = N_q \\ &\quad \text{if } \Delta[E_{cur}] = -1 \\ &\quad\quad P_{cur} = P_{cur} + 2P_{val}[E_{cur}] \\ &\quad\quad \Delta[E_{cur}] = 1 \end{aligned}$$

```

    end
end

$$P_k = \varphi(E_k - E_{cur}) + P_{cur}$$

if  $k \neq N_k$  and  $E_{k+1} < E_k$ 
     $dE = -$ 
     $E_{prev} = E_k$ 
     $P_{cur} = P_{cur} - \varphi(E_{cur} - E_{prev})$ 
     $E_{cur} = \max(\varepsilon_p(:, 1)) < E_{prev}$ 
     $P_{cur} = P_{cur} + \varphi(E_{cur} - E_{prev})$ 
end

```

In initial testing this algorithm was approximately five to ten times faster than the original matrix algorithm. Since the speed of this algorithm varies depending on the data, more testing needs to be done to ensure its efficacy.

3.2 Model Inversion

The forward algorithm can easily be modified to create an inverse algorithm. That is, given a desired value of the polarization P , determine the value of E that will give that value of P when input to the forward algorithm. There are two main modifications. First, the while loop is terminated based on P_k rather than E_k . Second, the desired value of P_k may occur at one of the switching values. In this case, the value of E at the switching value is selected for E_k . Although putting this value of E into the forward algorithm will not return the desired value of P , this is the best estimate for E_k that can be obtained from the model. Algorithm 5 outlines the process used to construct the inverse model. The details for the case $dP > 0$ before the main loop are provided separately in Algorithm 6. The situation for $dP < 0$ is analogous.

Algorithm 5. *Outline of Inverse Algorithm.*

$$E_0 = 0$$

$$\bar{P} = \varepsilon_k / \eta + P_R \Delta$$

$$P_{init} = CW^T \bar{P} V$$

if $P_2 > P_{init}$

Set $E_{prev} = 0$ and determine initial values of E_{cur} and P_{cur} using ε_n

else

Set $E_{prev} = 0$ and determine initial values of E_{cur} and P_{cur} using ε_p

end

for $k = 2 : N_k$

if $dE > 0$

Update E_{prev} , E_{cur} , P_{cur} and Δ based on P_k and ε_n .

Calculate E_k

If dE changes sign, update E_{prev} , E_{cur} and P_{cur} using ε_p

else

Update E_{prev} , E_{cur} , P_{cur} and Δ based on P_k and ε_p .

Calculate E_k

If dE changes sign, update E_{prev} , E_{cur} and P_{cur} using ε_n

end

end

Algorithm 6. $dP > 0$ initially

$$dE = +$$

$$E_{prev} = 0$$

$$E_{cur} = \min(\varepsilon_n(:, 1)) > E_{prev}$$

$$P_{cur} = P_{init} + \varphi(E_{cur} - E_{prev})$$

Algorithm 7 summarizes the details for the case $dP > 0$ inside the main loop. The situation for $dP < 0$ is analogous.

Algorithm 7. $dP > 0$ inside main loop

```

while  $P_{cur} < P_k$  and  $E_{ind} < N_q$ 
  if  $\Delta[E_{cur}] = -1$ 
     $P_{cur} = P_{cur} + 2P_{val}[E_{cur}]$ 
     $\Delta[E_{cur}] = 1$ 
  end
   $E_{prev} = E_{cur}$ 
   $E_{ind} = E_{ind} + 1$ 
   $E_{cur} = \varepsilon_n(E_{ind}, 1)$ 
   $P_{cur} = P_{cur} + \varphi(E_{cur} - E_{prev})$ 
end
if  $P_{cur} < P_k$  and  $E_{ind} = N_q$ 
  if  $\Delta[E_{cur}] = -1$ 
     $P_{cur} = P_{cur} + 2P_{val}[E_{cur}]$ 
     $\Delta[E_{cur}] = 1$ 
  end
   $E_k = (P_k - P_{cur})/\varphi + E_{cur}$ 
else
   $P_{mid} = P_{cur} - \varphi(E_{cur} - E_{prev})$ 
  if  $P_{mid} > P_k$ 
     $E_k = E_{prev}$ 
  else
     $E_k = (P_k - P_{cur})/\varphi + E_{cur}$ 
  end
end

```

```

end
if  $k \neq N_k$  and  $P_{k+1} < P_k$ 
   $dE = -$ 
   $E_{prev} = E_k$ 
   $P_{cur} = P_{cur} - \varphi(E_{cur} - E_{prev})$ 
   $E_{cur} = \max(\varepsilon_p(\cdot, 1)) < E_{prev}$ 
   $P_{cur} = P_{cur} + \varphi(E_{cur} - E_{prev})$ 
end

```

The inverse algorithm is only slightly slower than the forward algorithm. The reason for this is the need to calculate P_{mid} . The time needed to do this is insignificant compared to the total time required to execute the algorithm.

3.3 Parameter Estimation Techniques

3.3.1 Estimation of Densities by Constrained Optimization

Consider data $(\widehat{E}_k, \widehat{P}_k), k = 1, \dots, N_d$. A least squares fit to this data may be employed to estimate the $N_i + N_j + 1$ parameters

$$[\nu_1(E_{c_1}), \dots, \nu_1(E_{c_i})], [\nu_2(E_{c_1}), \dots, \nu_2(E_{c_j})] \text{ and } \eta.$$

Let $P(\widehat{E}_k; \theta)$ represent the modeled parameter-dependent polarization values of (3.1) for the N_d input field values \widehat{E}_k . Then define

$$\theta = [\nu_1(E_{c_1}), \dots, \nu_1(E_{c_i}), \nu_2(E_{c_1}), \dots, \nu_2(E_{c_j}), \eta] \in \mathbb{R}^{N_i + N_j + 1}.$$

The constrained optimization problem is to find the θ that minimizes

$$f(\theta) = \frac{1}{2} \|P(\widehat{E}_k; \theta) - \widehat{P}_k\|^2, \quad (3.15)$$

$$\theta \geq 0, \quad (3.16)$$

where $\|\cdot\|$ is the Euclidean norm in \mathcal{R}^{N_d} . The MATLAB routine *fmincon.m* or *fminsearch.m* with the positivity constraint enforced can be employed to carry out the optimization. The accuracy of the resulting model will depend to a large extent on the amount and variability of the data used to identify the parameters. For example if only a major loop is used in the optimization process, the ability of the model to predict minor loops will be hampered. If a larger data set that includes biased minor loops is available, the accuracy of the model in these regimes will be greatly improved. Finally, note that solving (3.15) is analogous to the method used in [43] to estimate parameters in discretized Preisach models.

3.3.2 Estimation of Joint Density

Consider the more general formulation described by (2.21). The discretized equation (3.2) has the advantage of exhibiting a linear dependence on the parameters, although there are now $N_i N_j$ parameters to be identified as compared to the $N_i + N_j + 1$ for the nonlinear constrained optimization problem described in Section 3.3.1. This is also seen to be a quadratic programming problem when solved in the least squares sense. Note that η is not included with the other parameters for this method, but is estimated directly by the slope $\frac{\partial P}{\partial E}$ of major loop after all switching has occurred.

For the case of no thermal activation, evaluating (2.7) at the input field data

values \widehat{E}_k gives

$$P(E) = \sum_{i=1}^{N_i} \sum_{j=1}^{N_j} \left[\frac{1}{\eta} (\widehat{E}_k + E_{e_j}) + P_R \delta(\widehat{E}_k; E_{e_j}, E_{c_i}) \right] \nu(E_{e_j}, (E_{c_i}) v_i w_j \quad (3.17)$$

for $k = 1, \dots, N_d$. Note that δ depends on both the coercive and effective field quadrature points. Now define A_k and Φ to be $N_i \times N_j$ matrices with components

$$[A_k]_{ij} = \left[\frac{1}{\eta} (\widehat{E}_k + E_{e_j}) + P_R \delta(\widehat{E}_k; E_{e_j}, E_{c_i}) \right] v_i w_j \quad (3.18)$$

$$[\Phi]_{ij} = \nu(E_{e_j}, E_{c_i}). \quad (3.19)$$

Let $N = N_i N_j$ and define the $N \times 1$ vector θ and $1 \times N$ vector a_k by

$$\theta = \text{vec}(\Phi) \quad , \quad a_k = [\text{vec}(A_k)]^T, \quad (3.20)$$

where ‘vec’ is the column vector concatenation of a matrix. Next, define \mathcal{P} and $\widehat{\mathcal{P}}$ to be $N_d \times 1$ vectors with components

$$[\mathcal{P}]_k = P(\widehat{E}_k) \quad , \quad [\widehat{\mathcal{P}}]_k = \widehat{P}_k. \quad (3.21)$$

Finally, define A to be an $N_d \times N$ matrix with rows

$$[A]_k = a_k. \quad (3.22)$$

As a result of these definitions, (3.17) can be treated as the linear system

$$A\theta = \mathcal{P}. \quad (3.23)$$

In order to identify θ for a particular set of data $\{\widehat{E}_k, \widehat{P}_k\}$, $k = 1, \dots, N_d$, we minimize

$$f(\theta) = \frac{1}{2} \|A\theta - \widehat{\mathcal{P}}\|^2 = \frac{1}{2} \theta^T A^T A \theta - \widehat{\mathcal{P}}^T A \theta + \widehat{\mathcal{P}}^2 \quad (3.24)$$

subject to

$$\theta_i \geq 0, \quad i = 1, \dots, N.$$

Note that the term $\widehat{\mathcal{P}}^2$ is a constant and thus can be dropped for implementation purposes.

The solution to (3.24) is found using the techniques employed in [32] for identifying density functions in Preisach models. First, since $A^T A$ is symmetric, its singular value decomposition can be written as

$$A^T A = U S U^T, \quad (3.25)$$

where the $N \times N$ diagonal matrix S contains the singular values of $A^T A$. Since the columns u_i of U are orthogonal,

$$u_i^T u_j = \delta_{ij}, \quad (3.26)$$

where δ_{ij} is the Kronecker delta. Now suppose $\text{rank}(A) = \text{rank}(S) = q < \min N, N_d$. This will be the situation when there number of data points is not great and large values of N_i and N_j are used to allow fine details of the hysteresis loops to be characterized. In this case, the rows and columns of S and U corresponding to zero singular values are eliminated to form the $q \times q$ matrix \widehat{S} and $N \times q$ matrix \widehat{U} . It may be necessary to remove small singular values as well if they are causing instability. Then

$$A^T A = \widehat{U} \widehat{S} \widehat{U}^T \quad (3.27)$$

and

$$\widehat{U}^T \widehat{U} = 1. \quad (3.28)$$

Letting $x = \widehat{U}^T \theta$ or equivalently $\theta = \widehat{U} x$, allows (3.24) to be rewritten as

$$f(x) = \frac{1}{2} x^T \widehat{U}^T \widehat{U} \widehat{S} \widehat{U}^T \widehat{U} x - \widehat{\mathcal{P}}^T A \widehat{U} x = \frac{1}{2} x^T \widehat{S} x - \widehat{\mathcal{P}}^T A \widehat{U} x, \quad (3.29)$$

where the unneeded $\widehat{\mathcal{P}}^2$ has been dropped. The minimization constraint is now

$$g(x) = \widehat{U} x \geq 0. \quad (3.30)$$

The MATLAB routine *quadprog.m* may then be used to solve the quadratic programming problem (3.29) for x^* . The solution θ^* is then obtained from $\theta^* = \widehat{U} x^*$.

3.3.3 Estimation of Joint Density with Regularization

In Section (2.2), it was shown that the polarization model (2.21) is an integral equation with a compact operator having infinite dimensional range. Thus (3.24) is ill-posed for determining a unique solution as well as for the continuous dependence of θ on the measured data \widehat{P} . So, as described in Section (2.2), Tikhonov regularization may be added so that the constrained minimization problem becomes

$$f_\alpha(\theta) = \frac{1}{2} \theta^T A^T A \theta - \widehat{\mathcal{P}}^T A \theta + \widehat{\mathcal{P}}^2 + \frac{\alpha}{2} \|\theta\|^2 \quad (3.31)$$

subject to $\theta_i \geq 0, \quad i = 1 \dots, N.$

A method of solution for (3.31) and ways to optimize α to balance the need for stability with the desire for accurate characterization of the data is described in [44].

3.4 Approximation Techniques for Rod Model

3.4.1 Complete Rod Model

To implement the rod model (2.58), it is necessary to approximate the infinite dimensional spatial and temporal behavior. We first derive a semidiscrete system using finite elements, construct a uniform partition of $[0, \ell]$ into N subintervals of length $h = \ell/N$. The subinterval points are then $x_i = ih$, $i = 0, 1, \dots, N$. The spatial basis $\{\psi_i\}_{i=1}^N$ is comprised of linear splines

$$\psi_i(x) = \frac{1}{h} \begin{cases} (x - x_{i-1}), & x_{i-1} \leq x \leq x_i \\ (x_{i+1} - x), & x_i \leq x \leq x_{i+1} \\ 0, & \text{otherwise} \end{cases}, \quad i = 1, \dots, N-1 \quad (3.32)$$

$$\psi_N(x) = \frac{1}{h} \begin{cases} (x - x_{N-1}), & x_{N-1} \leq x \leq x_N \\ 0, & \text{otherwise} \end{cases}. \quad (3.33)$$

The approximate solution $u^N(t, x)$ to the solution $u(t, x)$ of (2.58) is given by the expansion

$$u^N(t, x) = \sum_{j=1}^N u_j(t) \psi_j(x). \quad (3.34)$$

Note that the essential boundary condition $u^N(t, 0) = 0$ is satisfied since

$$V^N = \text{span}\{\psi_i, \varphi_i\}_{i=1}^N \subset V = \{(\psi, \varphi) \in X \mid \psi \in H^1(0, \ell), \psi(0) = 0, \psi(\ell) = \varphi\}. \quad (3.35)$$

Now consider test functions ψ_i and substitute $u^N(t, x)$ for $u(t, x)$ in (2.58) to get

$$\begin{aligned}
& \int_0^\ell \rho A \frac{\partial^2}{\partial t^2} \sum_{j=1}^N u_j(t) \psi_j(x) \psi_i(x) dx + \psi_i(\ell) M_L \frac{\partial^2}{\partial t^2} \sum_{j=1}^N u_j(t) \psi_j(\ell) \\
& + \int_0^\ell \left[Y^P A \frac{\partial}{\partial x} \sum_{j=1}^N u_j(t) \psi_j(x) + c^P A \frac{\partial^2}{\partial x \partial t} \sum_{j=1}^N u_j(t) \psi_j(x) \right] \frac{\partial \psi_i}{\partial x} dx \\
& + \psi_i(\ell) \left[k_L \sum_{j=1}^N u_j(t) \psi_j(\ell) + c_L \frac{\partial}{\partial t} \sum_{j=1}^N u_j(t) \psi_j(\ell) \right] \\
& = \int_0^\ell Y^P A \gamma P(E) \frac{\partial \psi_i}{\partial x} dx,
\end{aligned} \tag{3.36}$$

which holds for $i = 1 \cdots, N$. Note that the terms involving $\psi_i(\ell)$ in (3.36) are zero unless $i = j = N$, in which case $\psi_N(\ell) = 1$. Define the $N \times N$ matrices Q , K and C and the $N \times 1$ vector $F(t)$ by

$$[Q]_{ij} = \int_0^\ell \rho A \psi_i \psi_j dx, \quad i \neq N \text{ or } j \neq N \quad , \quad [Q]_{NN} = \int_0^\ell \rho A \psi_i \psi_j dx + M_\ell \tag{3.37}$$

$$[K]_{ij} = \int_0^\ell Y^P A \psi'_i \psi'_j dx, \quad i \neq N \text{ or } j \neq N \quad , \quad [K]_{ij} = \int_0^\ell Y^P A \psi'_i \psi'_j dx + k_\ell \tag{3.38}$$

$$[C]_{ij} = \int_0^\ell c^P A \psi'_i \psi'_j dx, \quad i \neq N \text{ or } j \neq N \quad , \quad [C]_{ij} = \int_0^\ell c^P A \psi'_i \psi'_j dx + c_\ell \tag{3.39}$$

$$[F]_i = \int_0^\ell Y^P A \gamma \psi'_i P(E) dx, \quad i = 1, \cdots, N \tag{3.40}$$

and let $v = [u_1, \dots, u_N]^T$. Then (3.36) can be written as

$$Q\ddot{v} + C\dot{v} + Kv = F(t). \quad (3.41)$$

Multiplying by Q^{-1} and rearranging terms gives

$$\ddot{v} = -Q^{-1}C\dot{v} - Q^{-1}Kv + Q^{-1}F(t). \quad (3.42)$$

Letting $y = [v, \dot{v}]^T$, this can be reformulated as the first order system

$$\dot{y}(t) = Ay(t) + B(t), \quad y(0) = 0, \quad (3.43)$$

where

$$A = \begin{bmatrix} 0 & I \\ -Q^{-1}K & -Q^{-1}C \end{bmatrix} \quad (3.44)$$

and

$$B = \begin{bmatrix} 0 \\ Q^{-1}F \end{bmatrix}. \quad (3.45)$$

To determine the values of the entries in the matrices Q , K , C and the vector F , several integrals involving the basis functions $\psi_i(x)$ must be evaluated. To simplify the calculations, the property

$$\int_a^b f(x)dx = \int_{a-c}^{b-c} f(x+c)dx \quad (3.46)$$

is used throughout. Thus,

$$\int_0^\ell \psi_i \psi_j dx = 0, \quad |i - j| > 1 \quad (3.47)$$

$$\int_0^\ell \psi_i^2 dx = \frac{1}{h^2} \int_0^h x^2 dx + \frac{1}{h^2} \int_0^h (h - x)^2 dx = \frac{2}{3}h, \quad i = 1, \dots, N - 1 \quad (3.48)$$

$$\int_0^\ell \psi_i \psi_{i+1} dx = \frac{1}{h^2} \int_0^h x(h - x) dx = \frac{1}{6}h, \quad i = 1, \dots, N - 1 \quad (3.49)$$

$$\int_0^\ell \psi_N^2 dx = \frac{1}{h^2} \int_0^h x^2 dx = \frac{1}{3}h \quad (3.50)$$

$$\int_0^\ell (\psi'_i)^2 dx = \int_0^h \left(\frac{1}{h}\right)^2 dx + \int_0^h \left(\frac{-1}{h}\right)^2 dx = \frac{2}{h}, \quad i = 1, \dots, N - 1 \quad (3.51)$$

$$\int_0^\ell \psi'_i \psi'_{i+1} dx = \int_0^h \left(\frac{1}{h}\right) \left(\frac{-1}{h}\right) dx = \frac{-1}{h}, \quad i = 1, \dots, N - 1 \quad (3.52)$$

$$\int_0^\ell (\psi'_N)^2 dx = \int_0^h \left(\frac{1}{h}\right)^2 dx = \frac{1}{h} \quad (3.53)$$

$$\int_0^\ell (\psi'_i) dx = \int_0^h \left(\frac{1}{h}\right) dx + \int_0^h \left(\frac{-1}{h}\right) dx = 0, \quad i = 1, \dots, N - 1 \quad (3.54)$$

$$\int_0^\ell (\psi'_N) dx = \int_0^h \left(\frac{1}{h}\right) dx = 1 \quad (3.55)$$

Hence, Q , K and C are the tridiagonal matrices

$$Q = \rho A \begin{bmatrix} \frac{2}{3}h & \frac{1}{6}h & & 0 \\ \frac{1}{6}h & \frac{2}{3}h & \ddots & \\ & \ddots & \ddots & \\ 0 & & \frac{2}{3}h & \frac{1}{6}h \\ & & \frac{1}{6}h & \frac{1}{3}h + M_L \end{bmatrix},$$

$$K = Y^P A \begin{bmatrix} \frac{2}{h} & \frac{-1}{h} & & 0 \\ \frac{-1}{h} & \frac{2}{h} & \ddots & \\ & \ddots & \ddots & \\ 0 & & \frac{2}{h} & \frac{-1}{h} \\ & & \frac{-1}{h} & \frac{1}{h} + k_L \end{bmatrix},$$

and

$$C = c^P A \begin{bmatrix} \frac{2}{h} & \frac{-1}{h} & & 0 \\ \frac{-1}{h} & \frac{2}{h} & \ddots & \\ & \ddots & \ddots & \\ 0 & & \frac{2}{h} & \frac{-1}{h} \\ & & \frac{-1}{h} & \frac{1}{h} + c_L \end{bmatrix}.$$

The vector F is

$$F = Y^P A \gamma \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

For simulation purposes, the system (3.43), is now discretized in time using a modified trapezoid rule. First, let $t_j = j\Delta t$ and integrate from t_j to t_{j+1} . Thus,

$$\int_{t_j}^{t_{j+1}} \dot{y} dt = \int_{t_j}^{t_{j+1}} [Ay + B(t)] dt \quad (3.56)$$

Using the trapezoid rule to approximate the second integral yields

$$y(t_{j+1}) - y(t_j) = \left[A \left(\frac{y(t_{j+1}) + y(t_j)}{2} \right) + \left(\frac{B(t_{j+1}) + B(t_j)}{2} \right) \right] \Delta t. \quad (3.57)$$

Next, let y_j approximate $y(t_j)$ so that

$$y_{j+1} - y_j = \left[A \left(\frac{y_{j+1} + y_j}{2} \right) + \left(\frac{B(t_{j+1}) + B(t_j)}{2} \right) \right] \Delta t. \quad (3.58)$$

Then, solving for y_{j+1} gives

$$\begin{aligned} y_{j+1} - y_j &= \Delta t A \left(\frac{y_{j+1} + y_j}{2} \right) + \Delta t \left(\frac{B(t_{j+1}) + B(t_j)}{2} \right) \\ \left[I - \frac{\Delta t}{2} A \right] y_{j+1} &= \left[I + \frac{\Delta t}{2} A \right] y_j + \Delta t \left(\frac{B(t_{j+1}) + B(t_j)}{2} \right) \\ y_{j+1} &= \left[I - \frac{\Delta t}{2} A \right]^{-1} \left[I + \frac{\Delta t}{2} A \right] y_j + \Delta t \left[I - \frac{\Delta t}{2} A \right]^{-1} \left(\frac{B(t_{j+1}) + B(t_j)}{2} \right). \end{aligned} \quad (3.59)$$

However, since $B(t_{j+1})$ is not known in advance, $\frac{B(t_{j+1}) + B(t_j)}{2}$ is replaced with $B(t_j)$. The difference equations can then be written as

$$y_{j+1} = \mathcal{A}y_j + \mathcal{B}(t_j), \quad y_0 = 0, \quad (3.60)$$

where

$$\mathcal{A} = \left[I - \frac{\Delta t}{2} A \right]^{-1} \left[I + \frac{\Delta t}{2} A \right] \quad (3.61)$$

and

$$\mathcal{B} = \Delta t \left[I - \frac{\Delta t}{2} A \right]^{-1} B(t_j). \quad (3.62)$$

This choice of numerical method is advantageous for both simulation purposes and subsequent control.

3.4.2 Lumped Rod Model

To determine the displacement of the rod at the end $x = \ell$, the rod model can be approximated by a lumped spring-mass model since the cross-section is small compared to the length and the electric field E is approximately uniform throughout the actuator. The resulting ODE model is

$$m_r \frac{d^2 u}{dt^2} + c_r \frac{du}{dt} + k_r u = \gamma_r P(E) - m_\ell \frac{d^2 u}{dt^2} - c_\ell \frac{du}{dt} - k_\ell u. \quad (3.63)$$

This can be simplified to

$$m \frac{d^2 u}{dt^2} + c \frac{du}{dt} + k u = \gamma_r P(E), \quad (3.64)$$

where $m = m_r + m_\ell$, $c = c_r + c_\ell$ and $k = k_r + k_\ell$. Rearranging terms and switching notation for the derivatives then gives

$$\ddot{u} = -\frac{k}{m} u - \frac{c}{m} \dot{u} + \frac{\gamma_r}{m} P(E). \quad (3.65)$$

Then, as was done for the rod model, letting $y = [u \dot{u}]^T$, the second order ODE may be written as the first order system

$$\dot{y}(t) = Ay(t) + B(t), \quad y(0) = 0, \quad (3.66)$$

where

$$A = \begin{bmatrix} 0 & 1 \\ -\frac{k}{m} & -\frac{c}{m} \end{bmatrix}$$

and

$$B = \begin{bmatrix} 0 \\ \frac{\gamma_x}{m} P(E) \end{bmatrix}.$$

To approximate the solution to (3.66), the modified trapezoid rule may be used, or to increase speed with some loss in accuracy, the backward Euler method may be employed. In this case, (3.66) is discretized as

$$\frac{y_{j+1} - y_j}{\Delta t} = Ay_{j+1} + \Delta t B(t_j). \quad (3.67)$$

Solving for y_{j+1} yields

$$\begin{aligned} y_{j+1} - y_j &= \Delta t Ay_{j+1} + \Delta t B(t_j) \\ [I - \Delta t A] y_{j+1} &= y_j + \Delta t B(t_j) \\ y_{j+1} &= [I - \Delta t A]^{-1} y_j + [I - \Delta t A]^{-1} \Delta t B(t_j) \end{aligned} \quad (3.68)$$

Numerical simulations and physical experiments demonstrate that the ODE model is a good approximation to the PDE model for specifying tip displacement. The ODE model has the advantage of faster computational times compared to the discretized PDE.

3.5 Approximation Techniques for Shell Model

Following the discussion in [39], we summarize approximation techniques for the weak form (2.71) of the shell model using a spline-Fourier based Galerkin method, which was developed for thin shells in [5]. Define

$$\mathcal{B}_{u_k}(\theta, x) = e^{im\theta} B_{u_j}(x)$$

$$\mathcal{B}_{v_k}(\theta, x) = e^{im\theta} B_{v_j}(x)$$

$$\mathcal{B}_{w_k}(\theta, x) = e^{im\theta} B_{w_j}(x)$$

to be the bases for the u , v and w displacements. Here B_{u_j} , B_{v_j} , and B_{w_j} are cubic B splines that are modified to satisfy the boundary conditions. Methods for these modifications can be found in [27]. Now let the approximating subspace be

$$V^{\mathcal{N}} = \text{span}\{\mathcal{B}_{u_k}\} \times \text{span}\{\mathcal{B}_{v_k}\} \times \text{span}\{\mathcal{B}_{w_k}\}. \quad (3.69)$$

Approximations to the displacements u , v , and w are the given by the expansions

$$\begin{aligned} u^{\mathcal{N}}(t, \theta, x) &= \sum_{k=1}^{\mathcal{N}_u} u_k(t) \mathcal{B}_{u_k}(\theta, x) \\ v^{\mathcal{N}}(t, \theta, x) &= \sum_{k=1}^{\mathcal{N}_v} v_k(t) \mathcal{B}_{v_k}(\theta, x) \\ w^{\mathcal{N}}(t, \theta, x) &= \sum_{k=1}^{\mathcal{N}_w} w_k(t) \mathcal{B}_{w_k}(\theta, x). \end{aligned}$$

By letting

$$\vartheta(t) = [u_1(t), \dots, u_{\mathcal{N}_u}(t), v_1(t), \dots, v_{\mathcal{N}_v}(t), w_1(t), \dots, w_{\mathcal{N}_w}(t)]^T \quad (3.70)$$

and defining

$$y(t) = [\vartheta(t), \dot{\vartheta}(t)]^T, \quad (3.71)$$

the weak form (2.71) can then be approximated by the semidiscrete system

$$\dot{y}(t) = Ay(t) + B(t) + G(t), \quad y(0) = y_0, \quad (3.72)$$

where A and $B(t)$ have same general form as (3.44) and (3.45). Details regarding the construction of the matrices Q , K , and C can be found in [5]. The vector

$$G(t) = \begin{bmatrix} 0 \\ Q^{-1}g(t) \end{bmatrix}$$

represents the force at the end of the cylinder.

Chapter 4

Model Validation and Device Characterization

In this chapter, we provide model validation for the field-polarization model in the case of no thermal activation. The methods described in Section 3.3 are all employed with the same data set. Then the polarization models which include and neglect thermal activation are combined with the rod actuator model to yield a field-displacement model. The two models are then validated with corresponding sets of varying frequency data and single frequency data. In both cases, the constrained optimization method from Section 3.3 is employed.

4.1 Field-Polarization

Consider the material PZT5H which for data collected at 0.2 Hz exhibits the hysteretic and nonlinear behavior shown in Figure 4.1 (note that this data was not collected on an AFM). We will demonstrate the ability of the model to accurately characterize the biased minor loop behavior. In order to show the accuracy of the model using the general densities as compared to the model using the lognormal/normal densities, four cases will be considered.

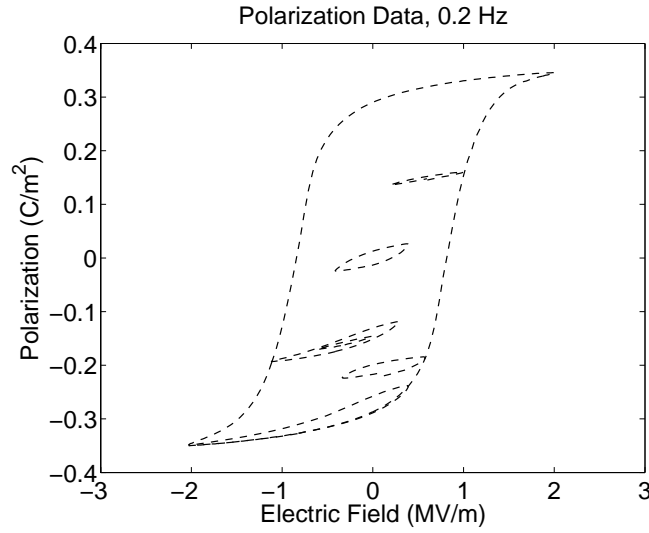


Figure 4.1: PZT5H data collected at 0.2 Hz with a symmetric major loop, Rayleigh loop and five biased minor loops.

1. Lognormal/normal densities, identification by fit to major loop and Rayleigh loop.
2. General densities, identification by fit to major loop and Rayleigh loop.
3. Lognormal/normal densities, identification by fit to all seven loops.
4. General densities, identification by fit to all seven loops.

In all four cases $N_i = N_j = 80$ was used and was sufficient for convergence. The accuracy of the model is quantified by the residual

$$\mathcal{R} = \left[\frac{1}{N_d} \sum_{k=1}^{N_d} |P(\widehat{E}_k) - \widehat{P}_k|^2 \right]^{\frac{1}{2}} \quad (4.1)$$

4.1.1 Lognormal/Normal Densities, Identification Using Major and Rayleigh Loops

The first case employed the discretized model (3.1) with lognormal and normal densities. The parameters $\eta = 1.03 \times 10^7$ MVm/C, $\overline{E_c} = 7.58 \times 10^5$ V/m, $c = 2.39 \times 10^{-1}$, $b = 2.98 \times 10^{10}$ V²/m² and $C = 1.42 \times 10^{-12}$ were estimated through a least squares fit to the major loop and the Rayleigh loop following the method outlined in Section 3.3.1 for the nonlinear constrained optimization problem (3.15). The resulting model fit is plotted in Figure 4.2 and the lognormal and normal densities are shown in Figure 4.3. The fit exhibits moderate accuracy for the major loop and Rayleigh loop and is inaccurate for the nested minor loops. The residual for the major loop and the Rayleigh loop is $\mathcal{R} = 9.37 \times 10^{-3}$ and that for all seven loops is $\mathcal{R} = 1.36 \times 10^{-2}$ as summarized in Table 4.1.

	(a)	(b)	(c)	(d)
Lognormal/Normal Densities ($N_i = N_j = 80$)	0.0094	0.0136	0.0098	0.0132
General Densities ν_1, ν_2 ($N_i = N_j = 80$)	0.0050	0.0115	0.0065	0.0057
Joint Density ν ($N_i = N_j = 24$)	0.0028	–	–	–
Joint Density ν ($N_i = N_j = 48$)	0.0106	–	–	–
Joint Density ν , Regularized ($N_i = N_j = 24$)	0.0037	–	–	–
Joint Density ν , Regularized ($N_i = N_j = 48$)	0.0024	–	–	–

Table 4.1: Residuals \mathcal{R} given by (4.1) with ν_1, ν_2 identified using the nonlinear constrained optimization method (3.15) and ν identified with the quadratic programming algorithm (3.24) and regularized method (3.31). (a) Identification and residual using major loop and Rayleigh loop. (b) Identification using major and Rayleigh loops, residual over all seven loops. (c) Identification over all seven loops, residual for major loop and Rayleigh loop. (d) Identification and residual over all seven loops.

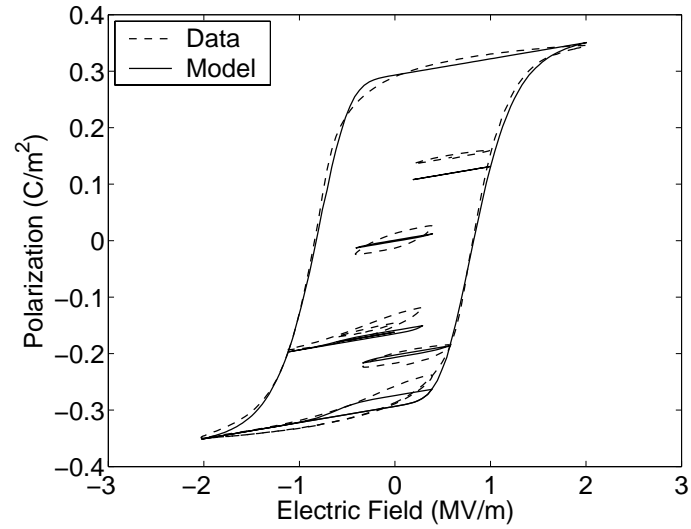


Figure 4.2: PZT5H data and model with lognormal and normal densities identified through a fit to major and Rayleigh loops.

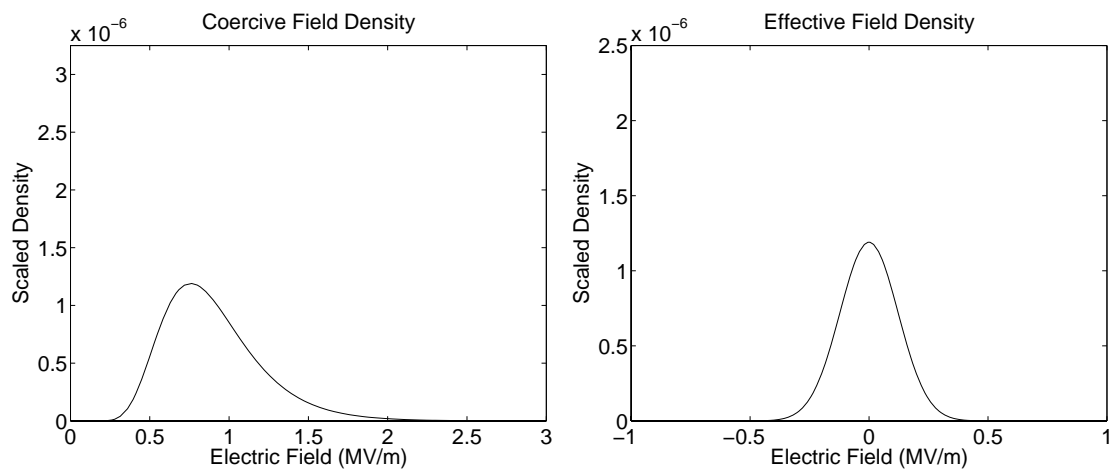


Figure 4.3: Lognormal coercive and normal effective field densities identified through a fit to major and Rayleigh loops.

4.1.2 General Densities, Identification Using Major and Rayleigh Loops

We now consider the model performance when the general densities ν_1 and ν_2 are employed. There are $N_i + N_j + 1$ parameters $[\nu_1(E_{c_1}), \dots, \nu_1(E_{c_i}), \nu_2(E_{c_1}), \dots, \nu_2(E_{c_j})]$ and η to be estimated. Note that all $N_i + N_j + 1$ parameters are required to be positive and ν_2 must be symmetric. The model fit is shown in Figure 4.4 with the corresponding general densities in Figure 4.5. The fit for the major loop and Rayleigh loop is greatly improved as confirmed by the residual of $\mathcal{R} = 5.04 \times 10^{-3}$ for these loops. However the fit to the other loops is only slightly better than for lognormal and normal densities and the residual for all seven loops is $\mathcal{R} = 1.15 \times 10^{-2}$.

Comparing the lognormal and normal densities of Figure 4.3 and the general densities of Figure 4.5 reveals that the qualitative behavior is similar in both cases. However the specific behavior of the general densities is much less regular.

4.1.3 Lognormal/Normal Densities, Identification Using All Seven Loops

Using the data from all seven loops, a least squares fit yielded the parameters $\eta = 8.94 \times 10^6$ MVm/C, $\overline{E_c} = 7.58 \times 10^5$ V/m, $c = 2.37 \times 10^{-1}$, $b = 3.24 \times 10^{10}$ V²/m² and $C = 1.35 \times 10^{-12}$. The model behavior, illustrated in Figure 4.6, shows a slight increase in accuracy in the biased minor loops as compared to first case when only the major and Rayleigh loops were used to identify the parameters. This is verified by residual of $\mathcal{R} = 1.32 \times 10^{-2}$ for all seven loops. There is a small decrease in accuracy in the residual of $\mathcal{R} = 9.75 \times 10^{-3}$ for the major loop and Rayleigh loop, but this is compensated for by the overall increase in the accuracy of the fit to all seven loops. The lognormal and normal densities for this case are shown in Figure 4.7.

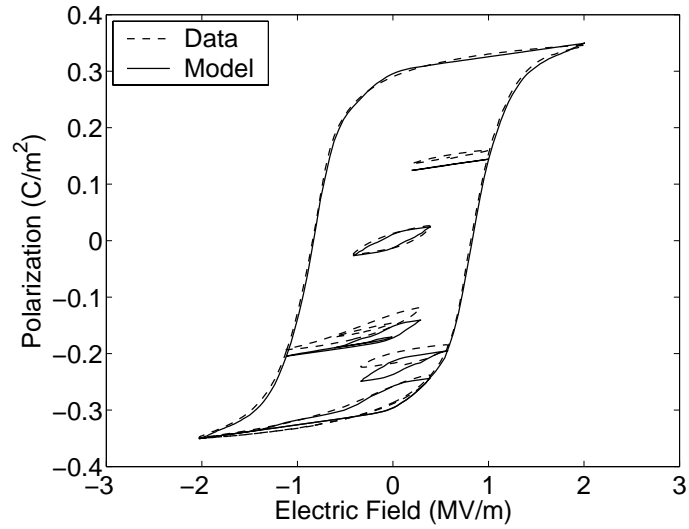


Figure 4.4: PZT5H data and model with general densities identified through a fit to major and Rayleigh loops.

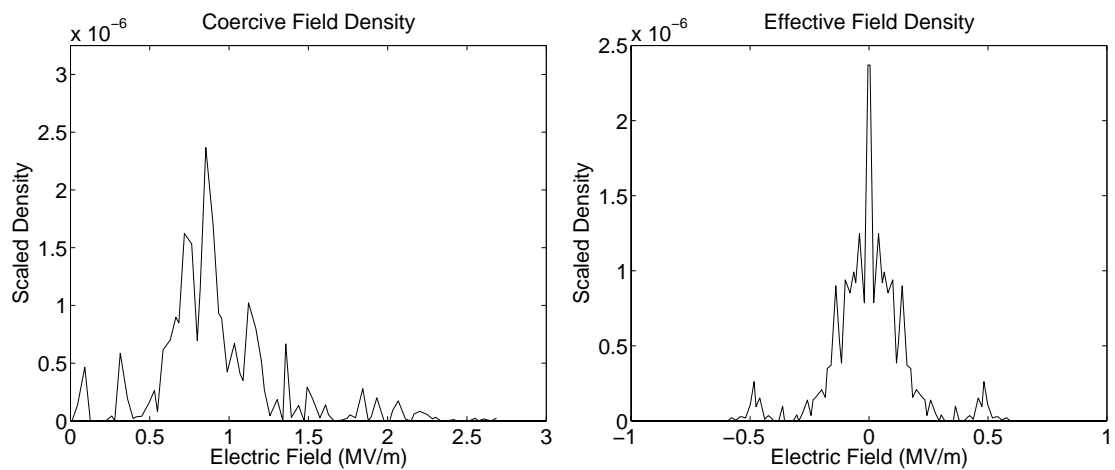


Figure 4.5: General coercive and effective field densities identified through a fit to major and Rayleigh loops.

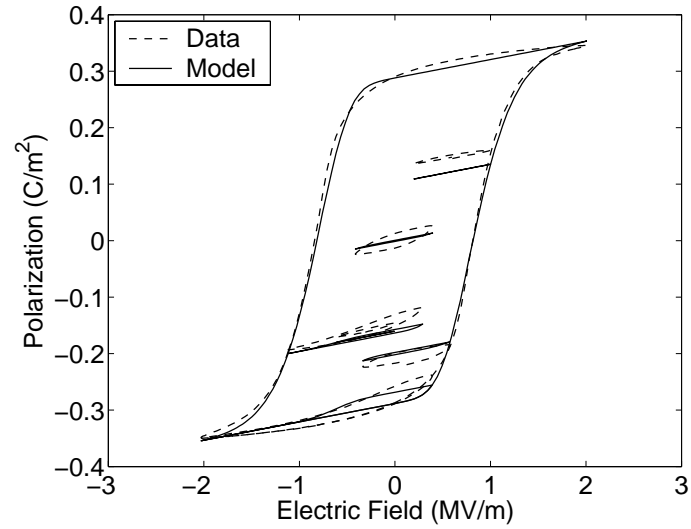


Figure 4.6: PZT5H data and model with lognormal and normal densities identified through a fit to all seven loops.

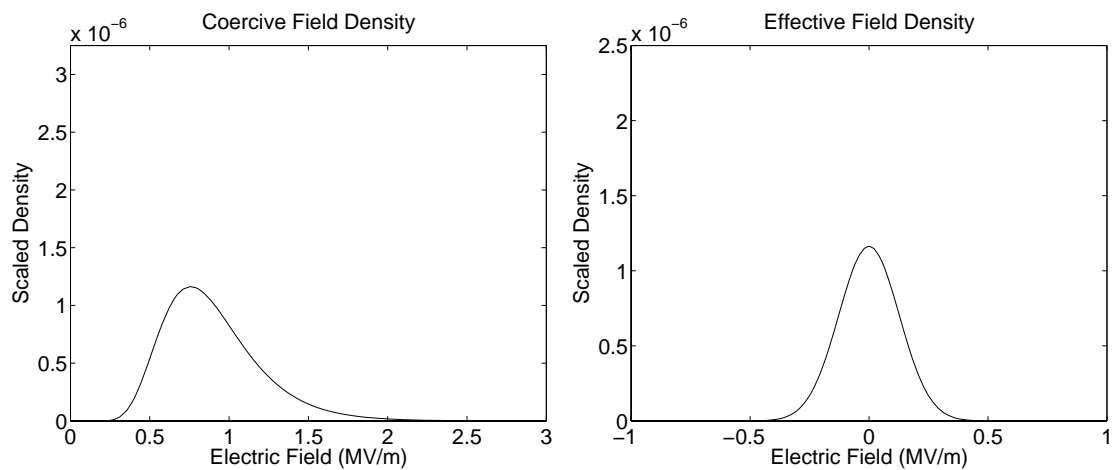


Figure 4.7: Lognormal coercive and normal effective field densities identified through a fit to all seven loops.

4.1.4 General Densities, Identification Using All Seven Loops

Finally, a model fit was constructed with both the general densities and the data from all seven loops. As demonstrated by Figure 4.8 the model fit achieved is very accurate for both the major loop and the Rayleigh loop as well as the biased minor loops. The residual of $\mathcal{R} = 5.73 \times 10^{-3}$ for all seven loops is less than half that when only the major loop and Rayleigh loop were used for identification, whereas the residual of $\mathcal{R} = 6.52 \times 10^{-3}$ for the major loop and Rayleigh loop is only slightly worse. This illustrates the limitations of the predictive capabilities of the model when only a small amount of data is available. This fit compares even more favorably to the third case when lognormal and normal densities were used to construct a fit with all seven loops. Thus the general densities, shown in Figure 4.9 are better able to capture the fine detail of the behavior of the loops.

Note that the accuracy of the fit to the biased minor loops is extremely dependent on the accuracy of the excursion point. Small errors in the fit to the major loop, where the gradient $\frac{\partial P}{\partial E}$ is large, can lead to large errors in the fit to the biased minor loops. This also explains the slight oscillations in the model fit before saturation is reached. They are caused by the optimization of the densities to accommodate minor loop behavior and not by any numerical instability related to the stepsize of the input field.

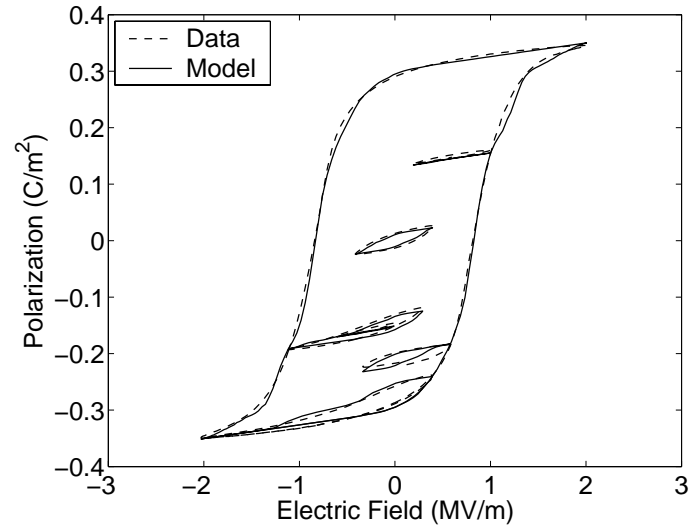


Figure 4.8: PZT5H data and model with general densities identified through a fit to all seven loops.

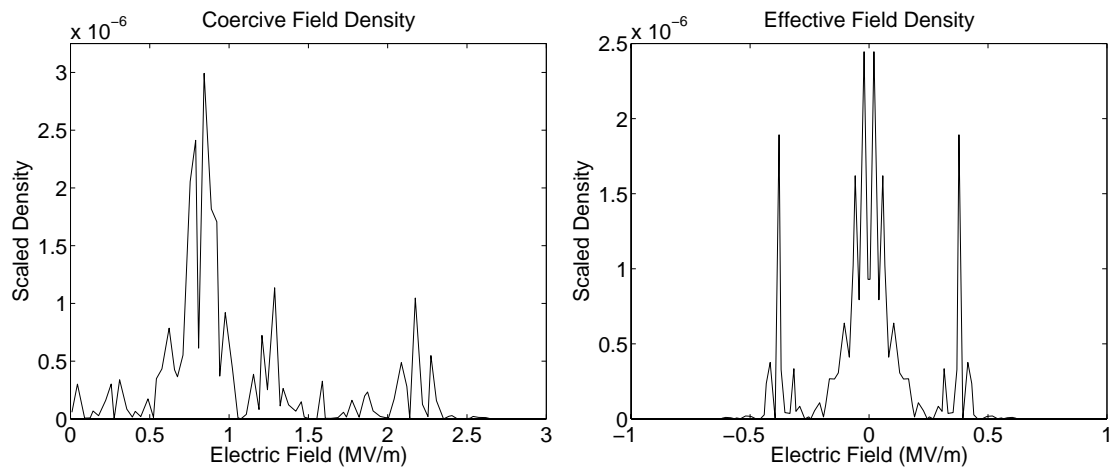


Figure 4.9: General coercive and effective field densities identified through a fit to all seven loops.

4.1.5 Joint Density

A method for identifying the joint density ν was discussed in Section (3.3.2). This technique takes advantage of a linear parameterization which allows the constrained optimization problem (3.24) to be solved instead as the quadratic programming problem (3.31). The main disadvantage of this approach is the large number of parameters, $N_i N_j$ as opposed to $N_i + N_j + 1$. However solution methods for this type of problem have been employed for a number of applications.

For the examples in this section, we use all seven loops of the PZT5H data from Figure 4.1. The effects of ill-posedness will be illustrated by the two levels of discretization $N_i = N_j = 24$ and $N_i = N_j = 48$. The values are lower than those used in Section 4.1 because of the memory limitations of MATLAB when doing singular value decomposition (SVD) computations. The model fits and joint density identifications are shown in Figures 4.10 and 4.11. The residuals for the two cases are respectively $\mathcal{R} = 0.0028$ and $\mathcal{R} = 0.0106$. It is noted that the residual is larger for the higher discretization level as is easily visible from inspection of the model fits. The explanation is the ill-posedness that occurs as the level of discretization is increased. At the lower discretization level, the fit is comparable to that obtained for the nonlinear constrained optimization with general densities using all seven loops. However, for the joint density formulation to obtain improved results, the loss in accuracy at higher discretization levels must be addressed. This leads to the examples using the formulation (3.31).

First note that the densities of Figure 4.11 validate the decay assumption (2.20) as well as the restriction of the operator \mathcal{K} to the compact domains. The discretization levels $N_i = N_j = 24$ and $N_i = N_j = 48$ were again used for the regularized least squares formulation and α was chosen to be $\alpha = 5 \times 10^{20}$ after repeated testing.

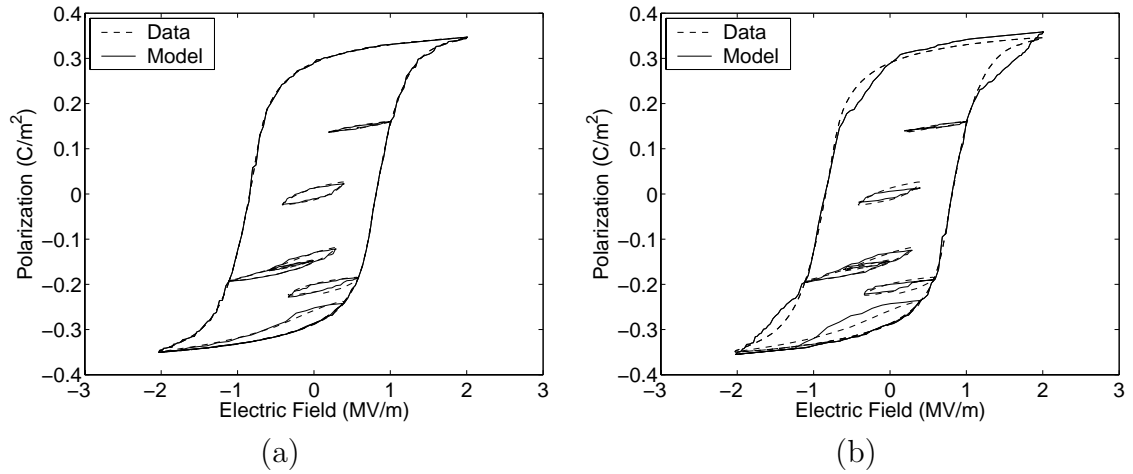


Figure 4.10: PZT5H data and model with joint density ν identified using the quadratic programming formulation (3.29) using data from all seven loops. (a) $N_i = N_j = 24$ and (b) $N_i = N_j = 48$.

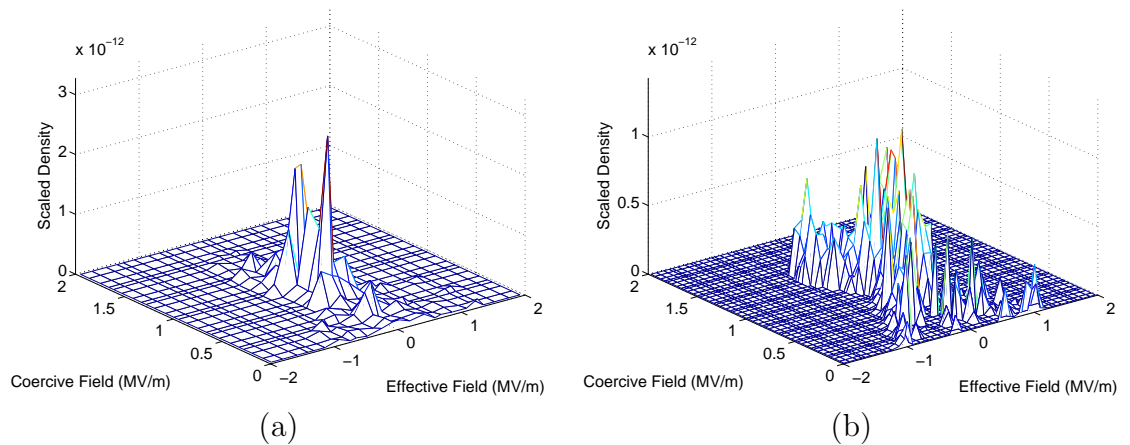


Figure 4.11: Joint density ν identified using the quadratic programming formulation (3.29) for (a) $N_i = N_j = 24$ and (b) $N_i = N_j = 48$.

The model fits are illustrated in Figure 4.12 while the corresponding joint densities are shown in Figure 4.13. The residuals of $\mathcal{R} = 0.0037$ and $\mathcal{R} = 0.0024$ illustrate the improvement in accuracy for higher discretization levels when regularization is employed. This is clearly visible by comparing Figures 4.10b and 4.12b. It is also apparent that the regularized joint density in Figure 4.13b is much cleaner than the joint density without regularization in Figure 4.11b. Finally, note that this is regularization by coarsening. In practice the lower discretization level without regularization can be used. However, the regularization was useful in verifying the reason for the problems that occurred as the discretization level was increased.

4.2 Field-Displacement

To demonstrate the accuracy and efficiency of the hysteresis model, we consider the characterization of nested minor loops collected at 0.1 Hz as well as E - P behavior at frequencies ranging from 0.279 Hz to 27.9 Hz. In both cases, displacements were computed using the ODE model (3.66) constructed using the stress relation (2.47) and hysteresis model (2.15) with general densities ν_1 and ν_2 identified using the techniques detailed in [33, 35].

Figure 4.14 illustrates the capability of the model to characterize nested, biased minor loop behavior and Figure 4.15 demonstrates the characterization of frequency-dependent dynamics. The latter involves the quantification of both thermal relaxation and inertial effects as illustrated by the change in sign of the slope $\frac{\partial P}{\partial E}$ following field reversal. Further details demonstrating properties of the model for characterizing hysteresis in various PZT compounds can be found in [33, 34, 35, 41].

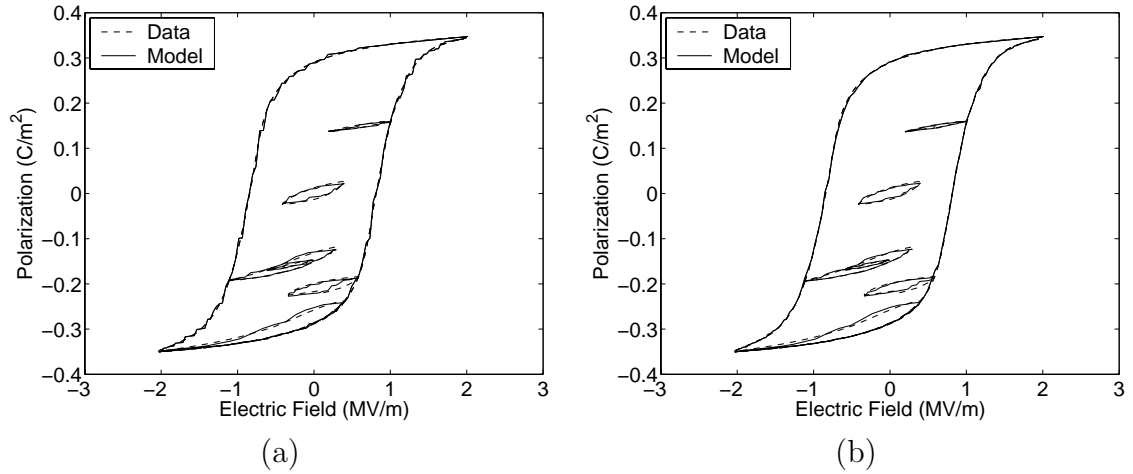


Figure 4.12: PZT5H data and model with joint density ν identified with Tikhonov regularization using data from all seven loops. (a) $N_i = N_j = 24$ and (b) $N_i = N_j = 48$.

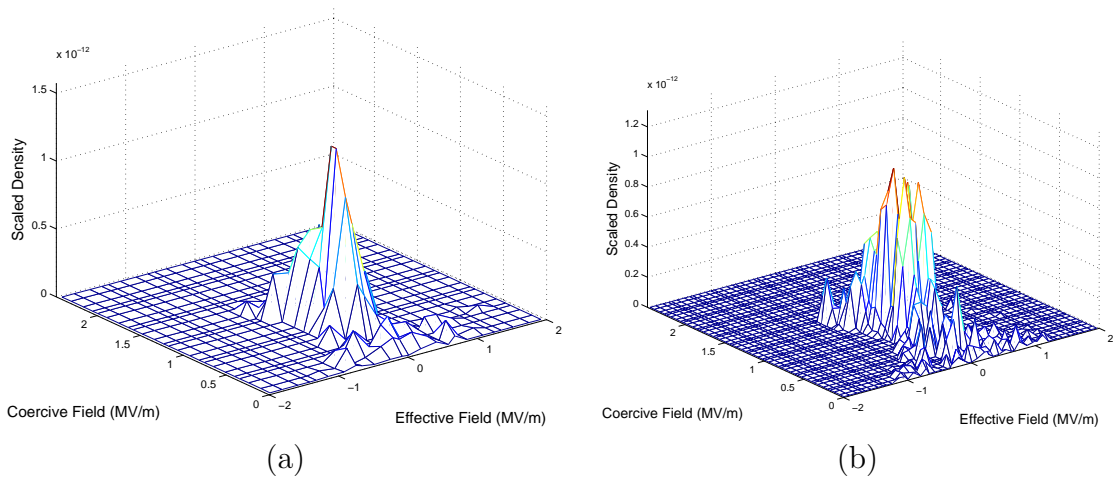


Figure 4.13: Joint density ν identified using Tikhonov regularization for (a) $N_i = N_j = 24$ and (b) $N_i = N_j = 48$.

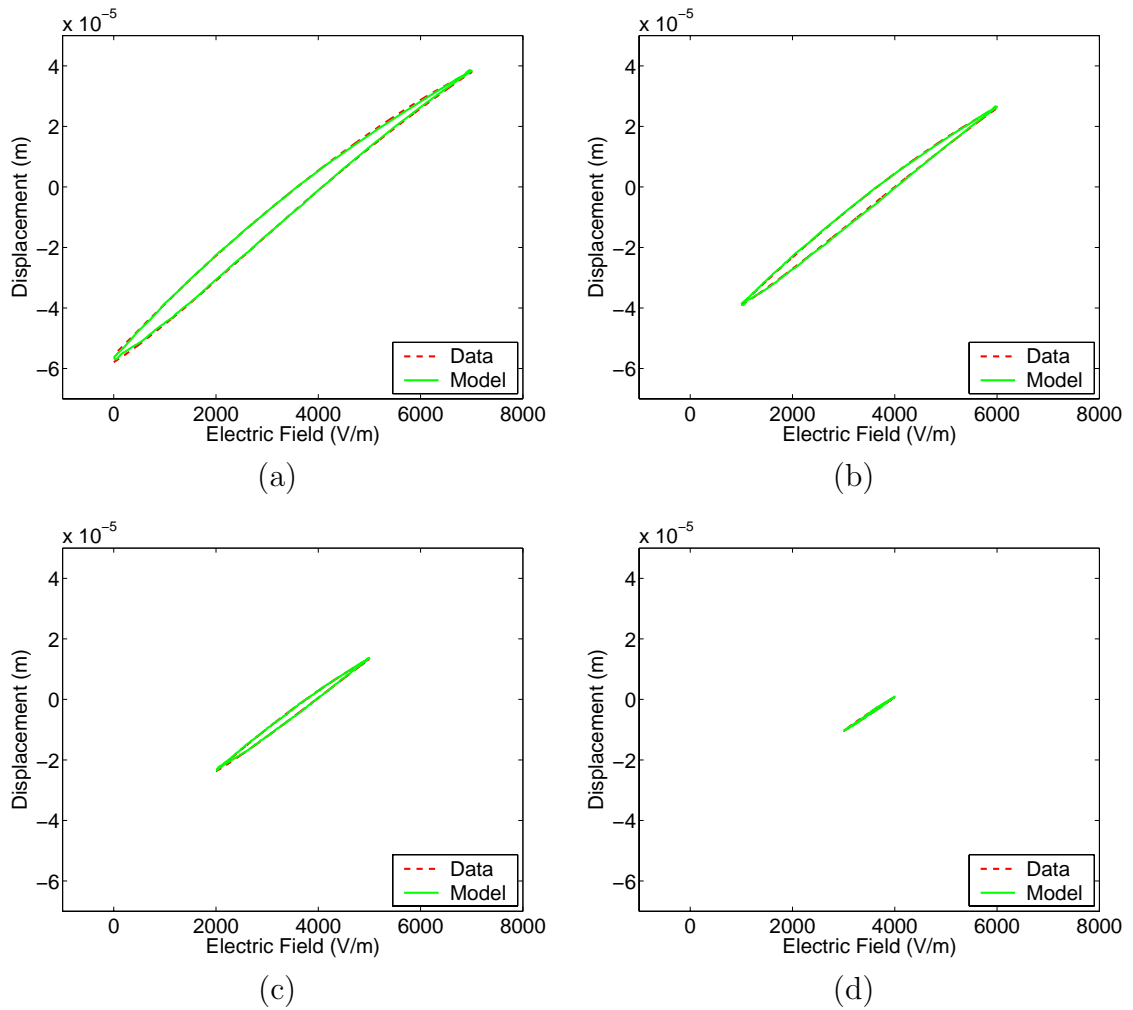


Figure 4.14: Characterization of AFM field-displacement behavior at 0.1 Hz.

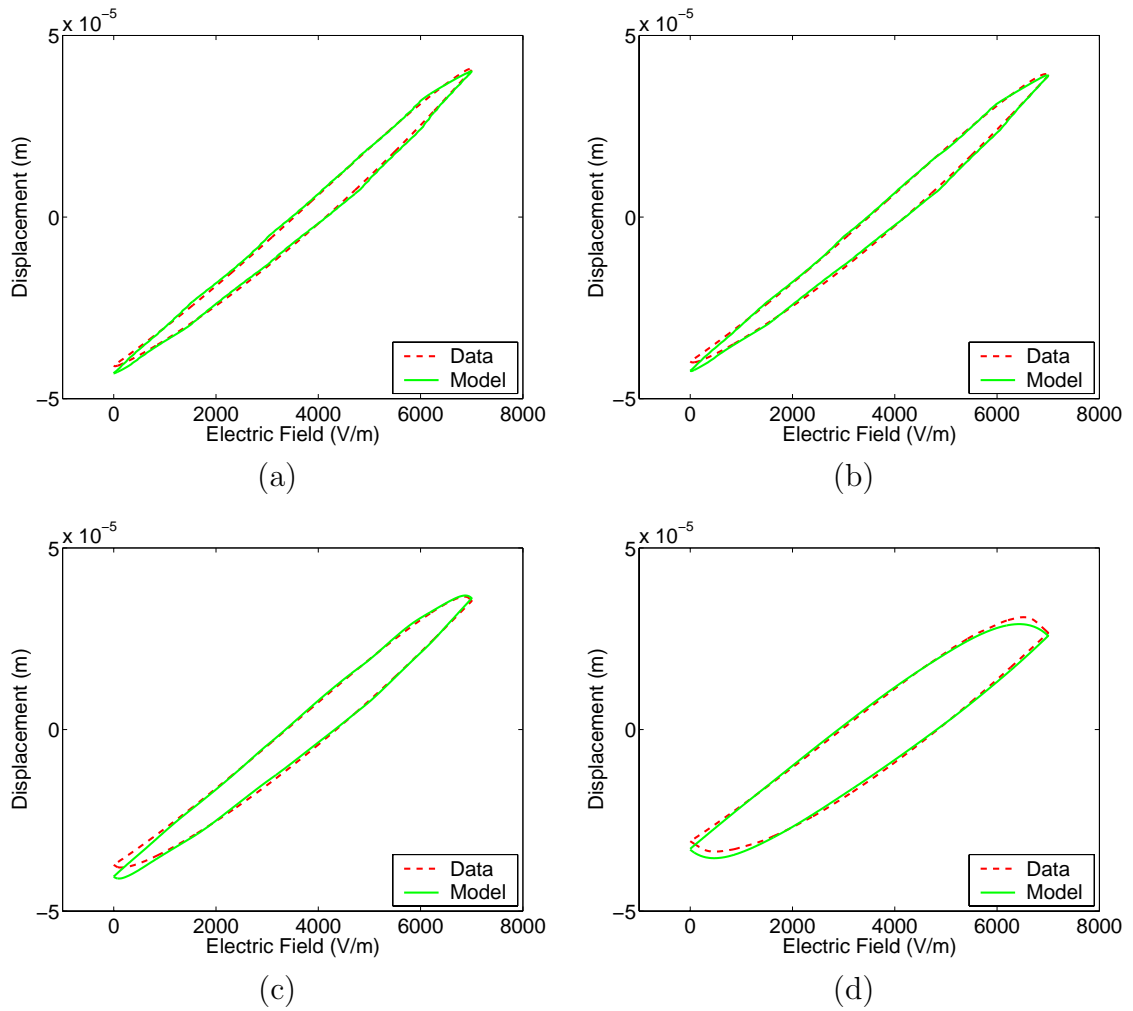


Figure 4.15: Characterization of AFM field-displacement behavior with sample rates of (a) 0.279 Hz, (b) 1.12 Hz, (c) 5.58 Hz, and (d) 27.9 Hz.

Chapter 5

PI Control

The second goal of the models developed in Chapter 2 is to develop model-based control design algorithms for the AFM. We begin with a discussion of proportional integral (PI) control, which is the control design presently implemented on the AFM.

5.1 PI Control Design

PI control is one of the oldest and simplest feedback control methods. Because of its relative ease of implementation and ability to achieve satisfactory results in a variety of situations, it is still commonly used for a number of applications. We will summarize the principles of PI control as described in detail in [13]. The PI controller is composed of two parts, the proportional part and the integral part. Each will be described in turn, but as a prelude, the operation of an on/off controller should first be explained.

The on/off controller has two possible values for the control variable u with the larger value usually associated with the on value. If r is the set point and y is the process output, the control law may be formulated as

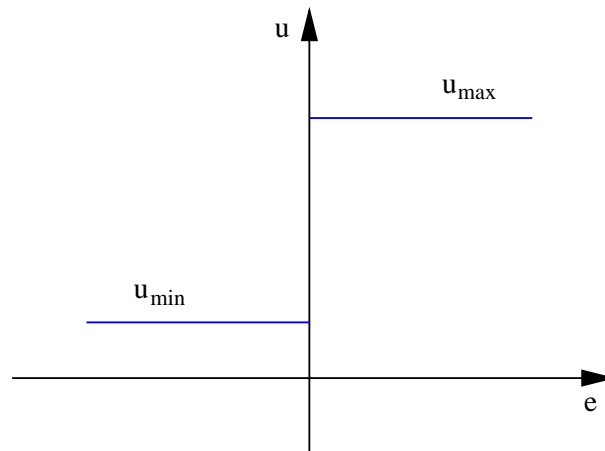


Figure 5.1: On/off control.

$$u = \begin{cases} u_{\min}, & e < 0 \\ u_{\max}, & e > 0, \end{cases} \quad (5.1)$$

where $e = r - y$ is the control error. The control law for an on/off controller is illustrated in Figure 5.1. Note that depending on the particular process, the conditions in (5.1) may be reversed. One example of the use of on/off control occurs in the thermostat for the furnace heating a home. A major disadvantage of on/off control is oscillation about a constant set point. In the thermostat example, the unit will continually switch on and off as the temperature falls below and then heats above the set point. One solution to this problem is to prescribe a range around the set point, where no control action is allowed. This greatly reduces the frequency of on/off oscillations and is in fact the solution that is implemented for the thermostat.

A second method to reduce the oscillations is to use a proportional controller. In this case, the controller is no longer on/off but is continuous in the range where previously no controller action was allowed. The control law for the proportional controller is

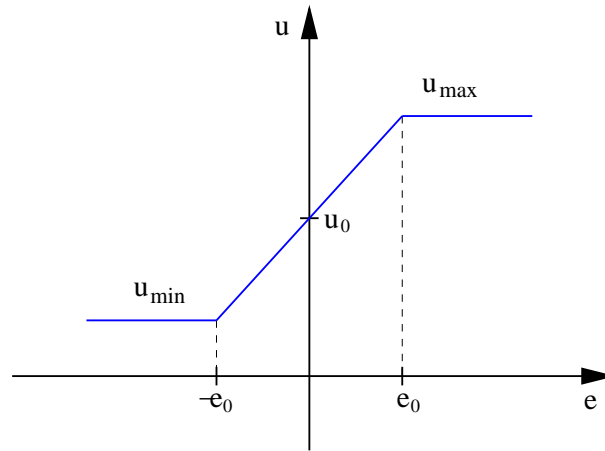


Figure 5.2: Proportional control.

$$u = \begin{cases} u_{\min}, & e < -e_0 \\ u_0 + K_c e, & -e_0 < e < e_0 \\ u_{\max}, & e > e_0, \end{cases} \quad (5.2)$$

where u_0 is the control value when the error is zero and K_c is the proportional gain of the controller. The control law for a proportional controller is illustrated in Figure 5.2. Thus the control value is proportional to the error in the region around the set point. It is assumed that u in (5.2) is continuous so we require that

$$u_{\min} = u_0 - K_c e_0 \quad (5.3)$$

$$u_{\max} = u_0 + K_c e_0. \quad (5.4)$$

Whereas the use of proportional control eliminates the problem of controller oscillation, the possibility of steady state error now arises. Thus, even after transients have been attenuated, a constant difference between the set point and the process output may persist. This can be seen by noting that when the error is small, it will be given

by

$$e = \frac{u - u_0}{K_c}. \quad (5.5)$$

Therefore, the error will be zero if and only if either $K_c \rightarrow \infty$ or $u_{ss} = u_0$. The first condition effectively means that the proportional controller has reduced back to an on/off controller. The second condition will only be true if $r = u_0$. In practice, the gain K_c is increased as large as possible in order to minimize the steady state error e_{ss} . However, if K_c is made too large, the closed-loop system will be unstable. Typically, u_0 is chosen to be some average of the expected values of r , thus minimizing the error for the greatest possible range of operating conditions.

A solution to the problem of steady state error is achieved by introducing the integral or I part of the controller. The integral portion essentially continually adjusts u_0 to the appropriate value for the current set point. The combination of of integral control with proportional control is called PI control and the control law for a PI controller is given by

$$u = \begin{cases} u_{\min}, & e < -e_0 \\ K_c \left[\frac{1}{T_i} \int e dt + e \right], & -e_0 < e < e_0 \\ u_{\max}, & e > e_0 \end{cases} \quad (5.6)$$

Note that the value of u_0 from the P controller has been replaced by

$$u_0 = \frac{K_c}{T_i} \int e dt. \quad (5.7)$$

The I part of the controller can be interpreted as the area under the error curve over time. Therefore, if an error persists over time, the integral portion of the controller will add up the error and adjust the control u so as to eliminate the error.

To construct a PI control design that accommodates the hysteresis and constitutive nonlinearities, the model can be inverted in the manner outlined in Section 3.2 and employed as a filter. Discretization errors combined with modeling errors will prevent complete elimination of error by the filter, but the disturbance d associated with the actuator can be significantly reduced in this manner. To provide a basis for comparison in the numerical examples which follow, we also consider a linear filter. In either case, the incoming signal, which has units of polarization, was fed through the filter. The result was subsequently input to the forward model. This process is depicted in Figure 5.3 for the two choices of disturbance. Finally, we treat the case of no disturbance d . In this situation, both the inverse filter and the forward polarization model are removed from the simulation.

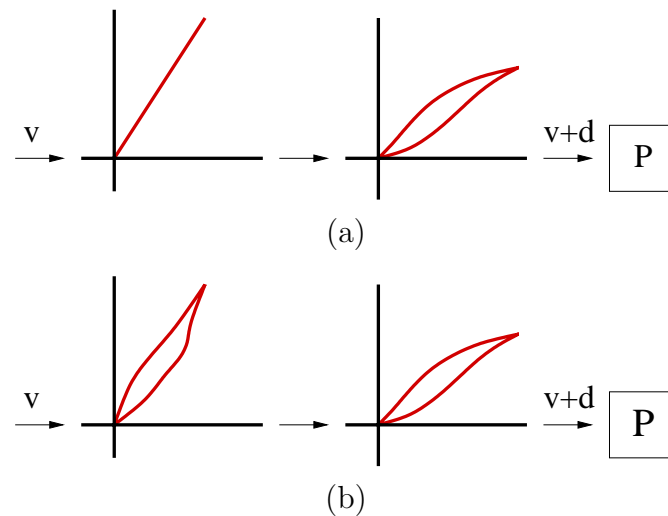


Figure 5.3: (a) Disturbance d due to scaled but unmodeled hysteresis and constitutive nonlinearities. (b) Disturbance d due to inverse filtering errors.

5.2 Numerical Examples

We now illustrate the performance of PI control with some numerical examples. Consider first the case where there is no disturbance d . The desired signal consists of a constant portion followed by a ramp, then another constant portion followed by a sine wave of frequency 1 Hz.

The tracking results for a P controller are shown in Figure 5.4. The values of $e_0 = 1 \times 10^{-5}$, $K_c = 1 \times 10^5$ and $u_0 = 0.5$ were used. The problem of steady state error is observed during the first constant portion of the desired trajectory.

The results for the same case but with a PI controller are shown in Figure 5.5. The values of e_0 and K_c are the same and $T_i = 5 \times 10^{-4}$. Note that the steady state error problem has been eliminated and the tracking is significantly more accurate.

Next, the PI controller is applied to the cases where the disturbance d is due to scaled but uncompensated hysteresis and where the inverse compensation procedure is used to approximately linearize the transducer response as shown in Figure 5.3. The tracking capabilities and errors for these cases are shown in Figure 5.6 and Figure 5.7. Both of these cases were considered again when the sine portion of the desired signal was increased to 10 Hz. The results are shown in Figure 5.8 and Figure 5.9. The tracking results are good for both cases at the lower frequency, but at the higher frequency a phase delay in the tracking is clearly evident. Thus, since PI control by itself has no predictive capability, it does not perform as well for rapidly changing desired trajectories.

One possible solution to this problem is the use of robust control. The development of model-based robust control designs is the focus of Chapter 6.

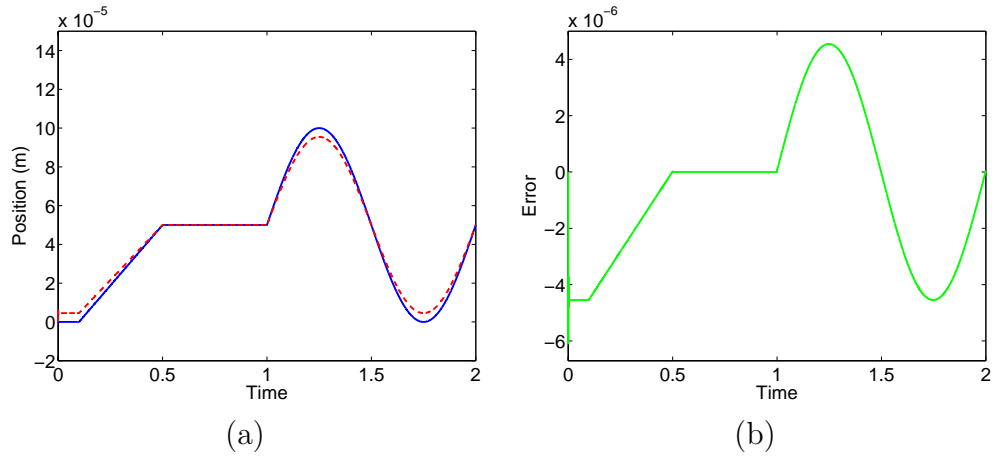


Figure 5.4: P design with no disturbance d and a frequency of 1 Hz. (a) Reference and simulated trajectory, and (b) tracking error.

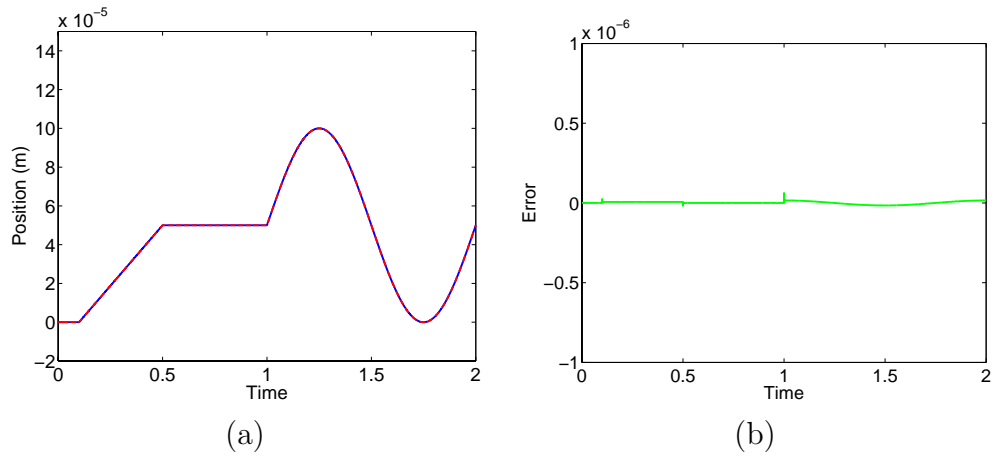


Figure 5.5: PI design with no disturbance d and a frequency of 1 Hz. (a) Reference and simulated trajectory, and (b) tracking error.

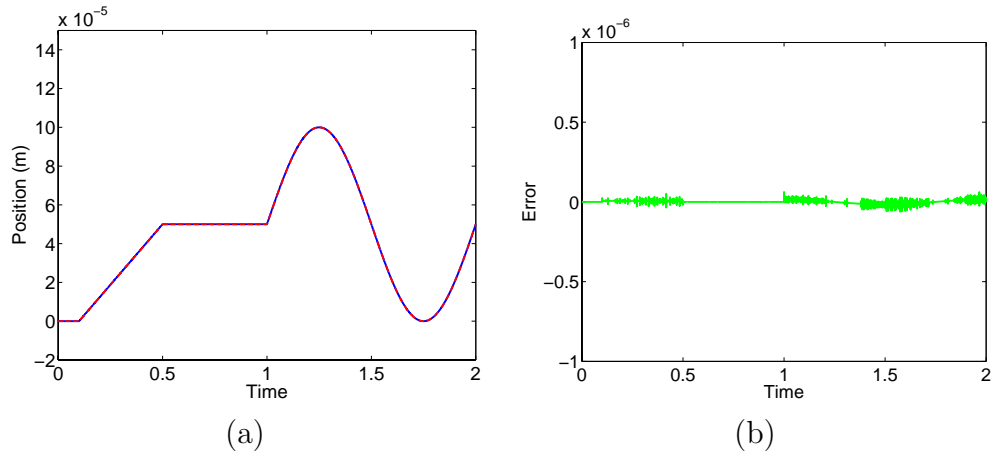


Figure 5.6: PI design with disturbance d due to inversion error and a frequency of 1 Hz. (a) Reference and simulated trajectory, and (b) tracking error.

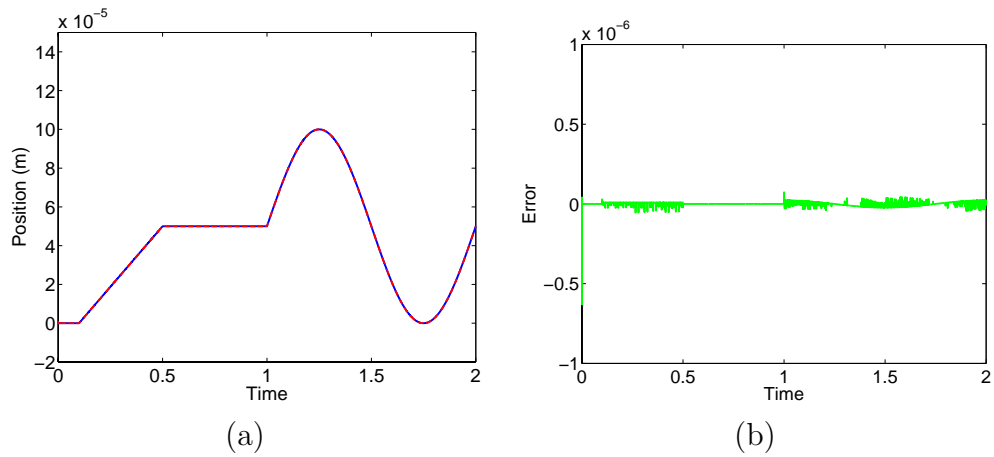


Figure 5.7: PI design with disturbance d due to uncompensated hysteresis and constitutive nonlinearities and a frequency of 1 Hz. (a) Reference and simulated trajectory, and (b) tracking error.

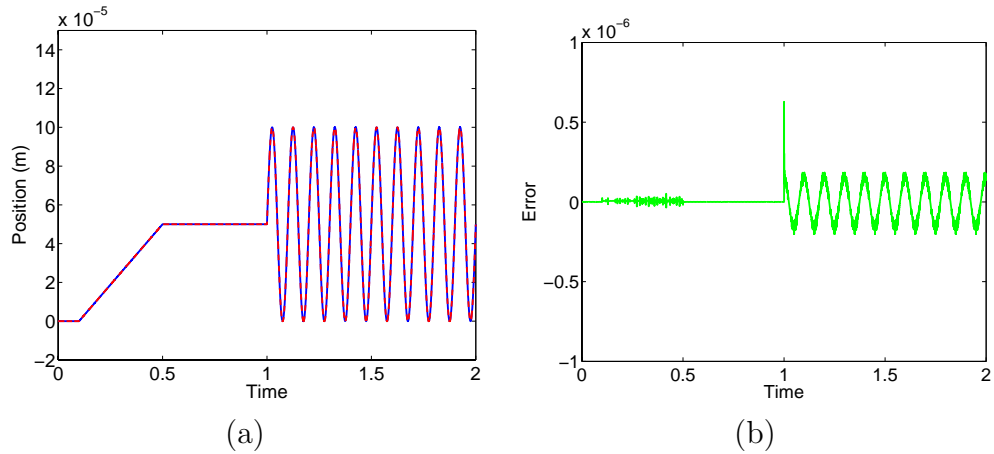


Figure 5.8: PI design with disturbance d due to inversion error and a frequency of 10 Hz. (a) Reference and simulated trajectory, and (b) tracking error.

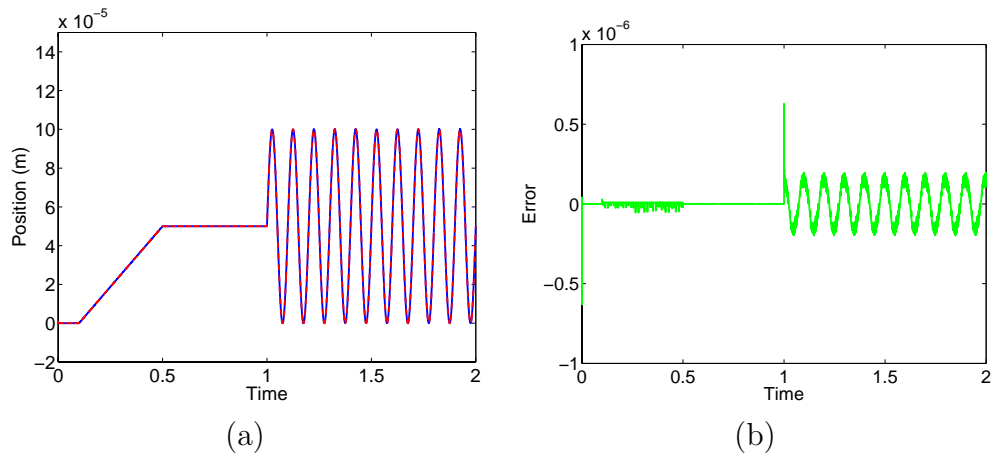


Figure 5.9: PI design with disturbance d due to uncompensated hysteresis and constitutive nonlinearities and a frequency of 10 Hz. (a) Reference and simulated trajectory, and (b) tracking error.

Chapter 6

Robust Control

6.1 Robust Control Design

To construct a robust control design that accommodates the hysteresis and constitutive nonlinearities, we again employ a model inverse as a filter as was done in Chapter 5. We will construct \mathcal{H}_2 and \mathcal{H}_∞ algorithms that utilize the model inverse. After a discussion of the system representation, numerical examples are given to demonstrate the performance of the control designs.

The control design is analogous to that described in [23, 24, 25] and is based on theory in [47]. The physical control system is the AFM and the control objective is to track a reference trajectory r , which represents the displacement of the actuator. The system representation is shown in Figure 6.1. The plant P represents the ODE model for the PZT actuator. The sensor noise in the measurement of y is separated into two components, s , taken to be 60 Hz noise to simulate electromagnetic disturbances, and n , taken to be high frequency noise, which can be attributed to the sensing device or external disturbances. The disturbance d represents errors in the plant input caused by either unattenuated hysteresis and constitutive nonlinearities or by discretization errors from the inverse filter. The output signal \hat{e} represents the weighted tracking

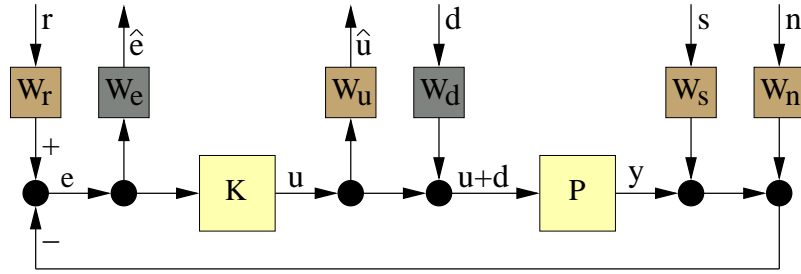


Figure 6.1: System representation including input disturbance d and sensor noise s and n in the transducer.

error and \hat{u} is the weighted output of the controller K . The weighting functions W_r , W_e , W_u , W_d , W_s , and W_n are selected to achieve the best performance from the controller. They are chosen based on known information about the corresponding signals.

To describe the open loop system, maps from the inputs r , d , s , and n to the outputs e , \hat{e} and \hat{u} are designated as

$$\begin{aligned} e &= W_r[r] - (P[W_d[d] + u] + W_n[n] + W_s[s]) \\ &= W_r[r] - P[W_d[d]] - W_n[n] - W_s[s] - P[u] \\ \hat{e} &= W_e[e] \\ \hat{u} &= W_u[u]. \end{aligned}$$

The transfer function matrix G from the inputs r , d , n , s and u to the outputs e , \hat{e} and \hat{u} is then specified by

$$G = \begin{bmatrix} W_e W_r & -W_e P W_d & -W_e W_n & -W_e W_s & -W_e P \\ 0 & 0 & 0 & 0 & W_u \\ W_r & -P W_d & -W_n & -W_s & -P \end{bmatrix}. \quad (6.1)$$

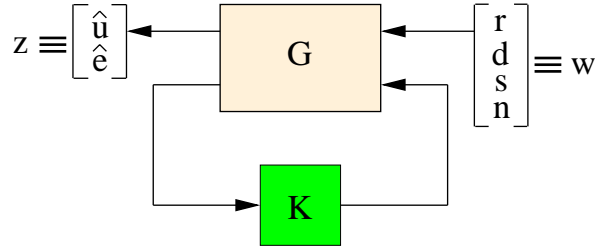


Figure 6.2: Linear fractional transformation representation (LFT) of the transducer model.

As shown in [25, 47], the linear fractional transformation system can be represented in the manner depicted in Figure 6.2. If $G(s)$ is partitioned as

$$G(s) = \begin{bmatrix} G_{11} & G_{12} \\ G_{21} & G_{22} \end{bmatrix}, \quad (6.2)$$

where G_{11} , G_{12} , G_{21} and G_{22} are the transfer functions from w to z , u to z , w to e and u to e and are defined by

$$\begin{aligned} G_{11} &= \left[\begin{array}{c|c} A & B_1 \\ \hline C_1 & 0 \end{array} \right], & G_{12} &= \left[\begin{array}{c|c} A & B_2 \\ \hline C_1 & D_{12} \end{array} \right] \\ G_{21} &= \left[\begin{array}{c|c} A & B_1 \\ \hline C_2 & D_{21} \end{array} \right], & G_{22} &= \left[\begin{array}{c|c} A & B_2 \\ \hline C_2 & 0 \end{array} \right] \end{aligned} \quad (6.3)$$

then $G(s)$ can be written as

$$G(s) = \left[\begin{array}{c|cc} A & B_1 & B_2 \\ \hline C_1 & 0 & D_{12} \\ C_2 & D_{21} & 0 \end{array} \right]. \quad (6.4)$$

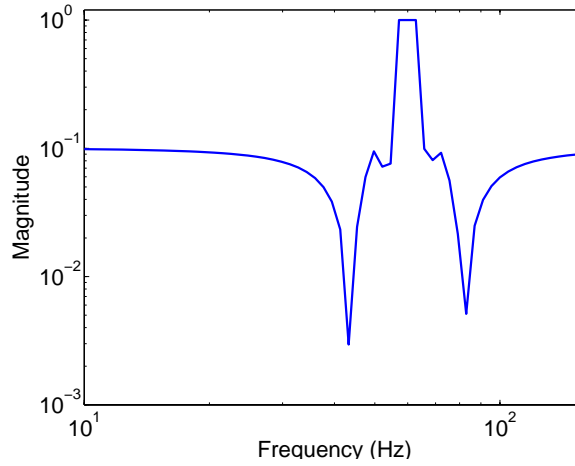


Figure 6.3: (a) Frequency response of the passband filter W_s .

The choice of weighting functions is an important part of the control design. To filter the 60 Hz sensor noise s , a sixth-order passband Chebyshev filter with a range of 55 Hz to 65 Hz is used for W_s . The frequency response of W_s is shown in Figure 6.3. An n th-order Chebyshev filter is a system whose frequency response function satisfies

$$|H_{\text{cheby}}(\omega)| = \frac{1}{1 + \epsilon_p C_n^2(\omega/\omega_s)}. \quad (6.5)$$

Here ω_s is the sampling frequency, ϵ_p is a parameter that determines the rolloff rate, and C_n are the n th-order Chebyshev polynomials, which are given by the recursion formulas

$$C_0(\omega) = 1 \quad (6.6)$$

$$C_1(\omega) = \omega \quad (6.7)$$

$$C_{n+1}(\omega) = 2\omega C_n(\omega) - C_{n-1}(\omega). \quad (6.8)$$

The choice for W_n is a second-order highpass Butterworth filter with a cutoff of

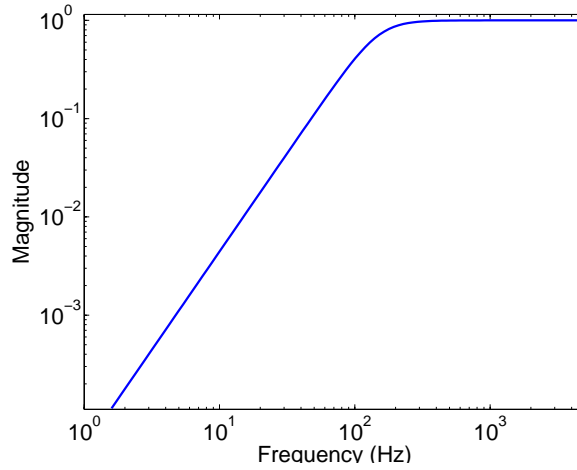


Figure 6.4: Frequency response of the highpass filter W_n .

150 Hz, which accommodates the high frequency sensor noise n . The frequency response of W_n is shown in Figure 6.4. The n th order Butterworth filter has the frequency response

$$|H_{\text{butter}}(\omega)| = \frac{1}{\sqrt{1 + (\omega/\omega_s)^{2n}}}. \quad (6.9)$$

The weight W_r is based on which components of the reference signal are most important during tracking. For example, considering the 0.25 Hz sinusoidal part of the reference signal to be most important, a sixth-order passband Chebyshev filter with a range of 0.125 Hz to 0.375 Hz was employed. To determine the weighting filter W_d , the reference signal was scaled to have the same order of magnitude as the polarization. This signal was then run through the inverse and linear filtering processes that were described in Section 5.1 and depicted in Figure 5.3. In both cases, the power spectrum of the output d had only low frequency components. Thus a second-order lowpass Butterworth filter with a cutoff of 10 Hz was used for W_d . The frequency response of W_d in the case when the linear filter was used is shown in Figure 6.5. The

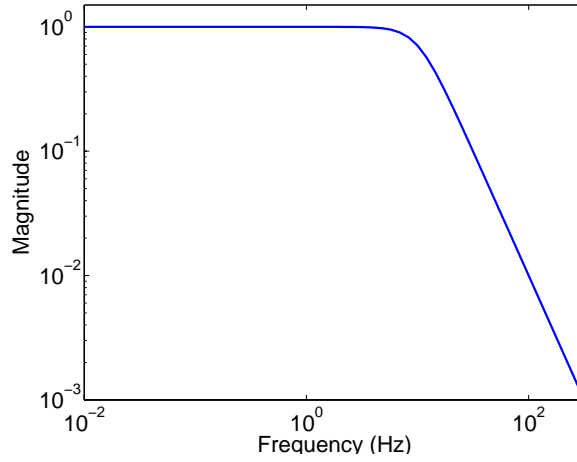


Figure 6.5: Frequency response of W_d for the disturbance d due to scaled but uncompensated hysteresis and nonlinearities.

weighting function on the error signal e was chosen to be

$$W_e = \frac{\gamma_e}{s + \epsilon_e} \quad (6.10)$$

with $\gamma_e = 1 \times 10^1$ and $\epsilon_e = 1 \times 10^{-8}$. An integrator was chosen so that the error would not achieve steady state at a nonzero value. Also, the pole was shifted slightly off zero to ensure that the controller was realizable. Finally, the weighting function on the controller output was chosen to be $W_u = 5 \times 10^{-6}$.

The transfer function matrix (6.1) gives a representation for the transducer system with weighting filters included. In order to formulate the control laws for computing the gains K , \mathcal{H}_2 and \mathcal{H}_∞ norms of the closed loop system representation T are minimized. These are

$$\|T\|_2^2 = \frac{1}{2\pi} \int_{-\infty}^{\infty} \text{trace}[T^*(j\omega)T(j\omega)] d\omega \quad , \quad \|T\|_\infty = \sup_{\omega \in \mathbf{R}} \bar{\sigma}[T(j\omega)], \quad (6.11)$$

where $\bar{\sigma}[T(j\omega)]$ represent the maximal singular values of the closed loop map T . As

described in [47], to formulate the \mathcal{H}_2 control algorithm, first define the matrices

$$H_2 = \begin{bmatrix} A - B_2 R_1^{-1} D_{12}^* C_1 & -B_2 R_1^{-1} B_2^* \\ -C_1^* (I - D_{12} R_1^{-1} D_{12}^*) C_1 & -(A - B_2 R_1^{-1} D_{12}^* C_1)^* \end{bmatrix}$$

and

$$J_2 = \begin{bmatrix} (A - B_2 R_1^{-1} D_{12}^* C_1)^* & -C_2^* R_2^{-1} C_2 \\ -B_1 (I - D_{21}^* R_2^{-1} D_{21}) B_1^* & -(A - B_2 R_1^{-1} D_{12}^* C_1) \end{bmatrix}$$

and the Ricatti equations

$$\begin{aligned} (A - B_2 R_1^{-1} D_{12}^* C_1)^* X_2 + X_2 (A - B_2 R_1^{-1} D_{12}^* C_1) \\ + X_2 (-B_2 R_1^{-1} B_2^*) X_2 - C_1^* (I - D_{12} R_1^{-1} D_{12}^*) C_1 = 0 \end{aligned} \quad (6.14)$$

$$\begin{aligned} (A - B_2 R_1^{-1} D_{12}^* C_1) Y_2 + Y_2 (A - B_2 R_1^{-1} D_{12}^* C_1)^* \\ + Y_2 (-C_2^* R_2^{-1} C_2) Y_2 - B_1 (I - D_{21}^* R_2^{-1} D_{21}) B_1^* = 0. \end{aligned} \quad (6.15)$$

Note that $R_1 \equiv D_{12}^* D_{12} > 0$ and $R_2 \equiv D_{21} D_{21}^* > 0$. To guarantee the existence of a unique \mathcal{H}_2 feedback gain K , we use the following theorem from [47].

Theorem 2 *There exists a unique controller which minimizes the \mathcal{H}_2 norm of the closed loop system if*

$$H_2 \in \text{dom}(\text{Ric}) \text{ and } X_2 \equiv \text{Ric}(H_2) > 0 \quad (6.16)$$

and

$$J_2 \in \text{dom}(\text{Ric}) \text{ and } Y_2 \equiv \text{Ric}(J_2) > 0 \quad (6.17)$$

The optimal \mathcal{H}_2 is then

$$K = \left[\begin{array}{c|c} A_2 & -L_2 \\ \hline F_2 & 0 \end{array} \right], \quad (6.18)$$

where

$$A_2 = A + B_2 F_2 + L_2 C_2 \quad (6.19)$$

$$L_2 = -(Y_2 C_2^* + B_1 D_{21}^*) R_2^{-1} \quad (6.20)$$

$$F_2 = -R_1^{-1} (B_2^* X_2 + D_{12}^* C_1). \quad (6.21)$$

If the \mathcal{H}_∞ norm is used, we will design a control law such that $\|T\|_\infty < \gamma$, where $\gamma > 0$. In this case, define the matrices

$$H_\infty = \begin{bmatrix} A & \gamma^{-2} B_1 B_1^* - B_2 B_2^* \\ -C_1^* C_1 & -A^* \end{bmatrix}, \quad J_\infty = \begin{bmatrix} A^* & \gamma^{-2} C_1^* C_1 - C_2^* C_2 \\ -B_1 B_1^* & -A \end{bmatrix}$$

and the Ricatti equations

$$A^* X_\infty + X_\infty A + X_\infty (\gamma^{-2} B_1 B_1^* - B_2 B_2^*) X_\infty + C_1^* C_1 = 0 \quad (6.22)$$

$$A Y_\infty + Y_\infty A^* + Y_\infty (\gamma^{-2} C_1^* C_1 - C_2^* C_2) Y_\infty + B_1 B_1^* = 0. \quad (6.23)$$

To guarantee the existence of an \mathcal{H}_∞ feedback gain K , we use the following theorem from [47].

Theorem 3 *There exists an admissible controller such that $\|T\|_\infty < \gamma$ if and only if*

$$H_\infty \in \text{dom}(\text{Ric}) \text{ and } X_\infty \equiv \text{Ric}(H_\infty) > 0, \quad (6.24)$$

$$J_\infty \in \text{dom}(\text{Ric}) \text{ and } Y_\infty \equiv \text{Ric}(J_\infty) > 0 \quad (6.25)$$

$$\rho(X_\infty, Y_\infty) < \gamma^{-2}. \quad (6.26)$$

The \mathcal{H}_∞ suboptimal control gain is given by

$$K = \left[\begin{array}{c|c} A_\infty & -Z_\infty L_\infty \\ \hline F_\infty & 0 \end{array} \right], \quad (6.27)$$

where

$$A_\infty = A + \gamma^{-2} B_1 B_1^* X_\infty + B_2 F_\infty + Z_\infty L_\infty C_2 \quad (6.28)$$

$$L_\infty = -Y_\infty C_2^* \quad (6.29)$$

$$F_\infty = -B_2^* X_\infty. \quad (6.30)$$

As described in [25], some assumptions must be made for the system representation (6.4) to assure the existence of unique control inputs for these \mathcal{H}_2 and \mathcal{H}_∞ control algorithms:

- (i) (A, B_1) is controllable and (C_1, A) is observable.
- (ii) (A, B_2) is stabilizable and (C_2, A) is detectable.
- (iii) $D_{12} = [0, I]^T$ and $D_{21} = [0, I]$.
- (iv) $\begin{bmatrix} A - j\omega I & B_2 \\ C_1 & D_{12} \end{bmatrix}$ has full column rank for all ω .
- (v) $\begin{bmatrix} A - j\omega I & B_1 \\ C_2 & D_{21} \end{bmatrix}$ has full row rank for all ω .

Assumption (i) guarantees the existence of a stabilizing controller, (ii), (iv) and (v) guarantee the existence of solutions to the associated Riccati equations and (iii) ensures that the \mathcal{H}_2 and \mathcal{H}_∞ problems are nonsingular.

6.2 Numerical Examples

We now illustrate the performance of robust control with some numerical examples similar to those discussed for the PI control design. Consider the two possibilities for the disturbance depicted in Figure 5.3. The first, which is due to scaled but uncompensated hysteresis, yields the tracking results and errors shown in Figure 6.6. In the second case, the disturbance is due to errors existing in the inverse compensation procedure used to approximately linearize the transducer response. The tracking capabilities and errors for this case are shown in Figure 6.7. A baseline case with no disturbance was also included. Figure 6.8 illustrates the tracking performance and errors for this case. In all three cases, the sinusoidal part of the tracking signal is set at 0.25 Hz.

A comparison between Figures 6.6 and 6.7 illustrates that highly accurate tracking is obtained in both cases, with errors less than $2 \mu\text{m}$ maintained after the commencement of the periodic cycle. The equivalence in accuracy for the uncompensated and compensated designs is attributed to the low level of hysteresis present in this low frequency drive regime as demonstrated by the 0.2 Hz data in Figure 1.4. Future investigations will focus on the extension of these control designs to higher frequencies, where increasing hysteresis levels necessitate inverse compensation as demonstrated for analogous magnetic model based controllers in [25].

Very similar results, shown in Figures 6.9, 6.10 and 6.11, are obtained using the \mathcal{H}_∞ design, indicating that for this application, the construction of models and control filters may play a more important role than the choice of robust control laws for high accuracy tracking.

In Chapter 7 we discuss the results of the experimental implementation of a model-based open loop control algorithm. This is a step towards the eventual implementation of a closed loop model-based PI or robust control algorithm.

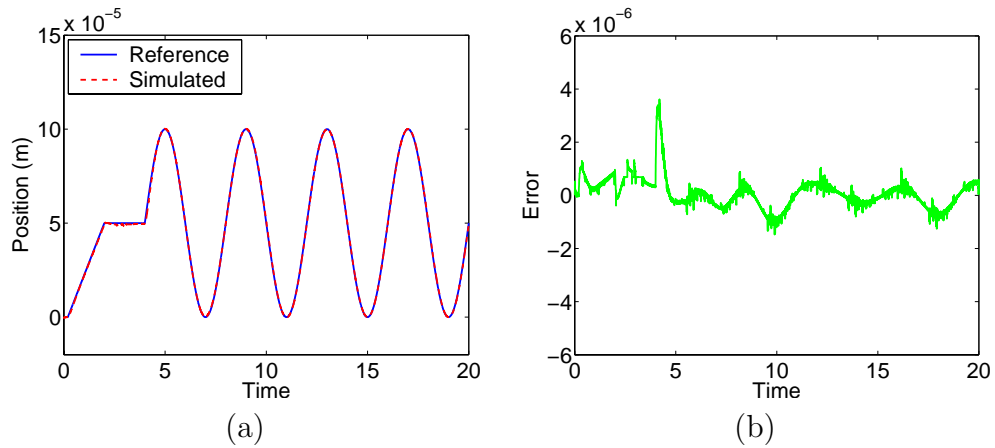


Figure 6.6: \mathcal{H}_2 design with sensor noise s and the disturbance d due to inversion error. (a) Reference and simulated trajectory, and (b) tracking error.

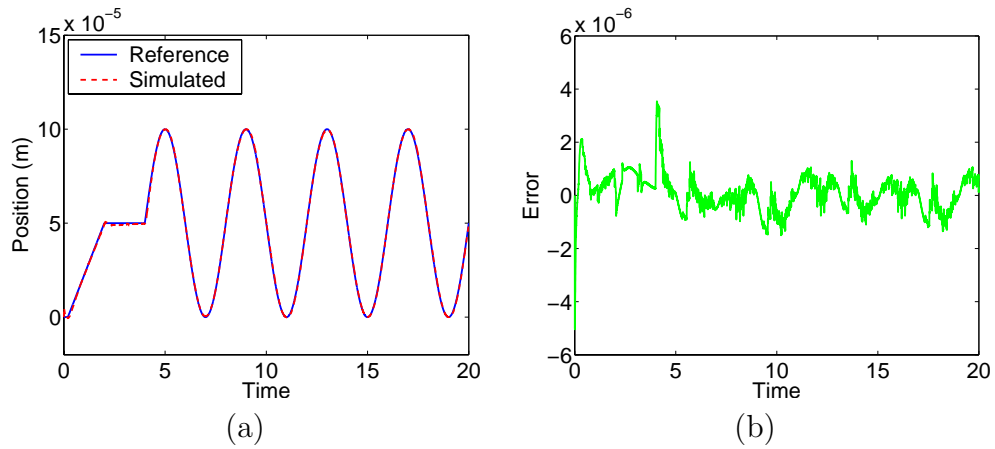


Figure 6.7: \mathcal{H}_2 design with sensor noise s and the disturbance d due to uncompensated hysteresis and constitutive nonlinearities. (a) Reference and simulated trajectory, and (b) tracking error.

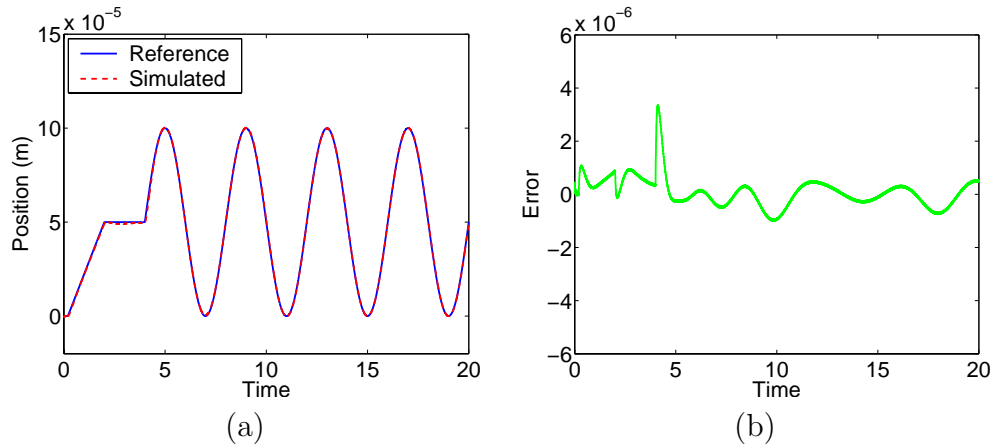


Figure 6.8: \mathcal{H}_2 design with sensor noise s but no disturbance d . (a) Reference and simulated trajectory, and (b) tracking error.

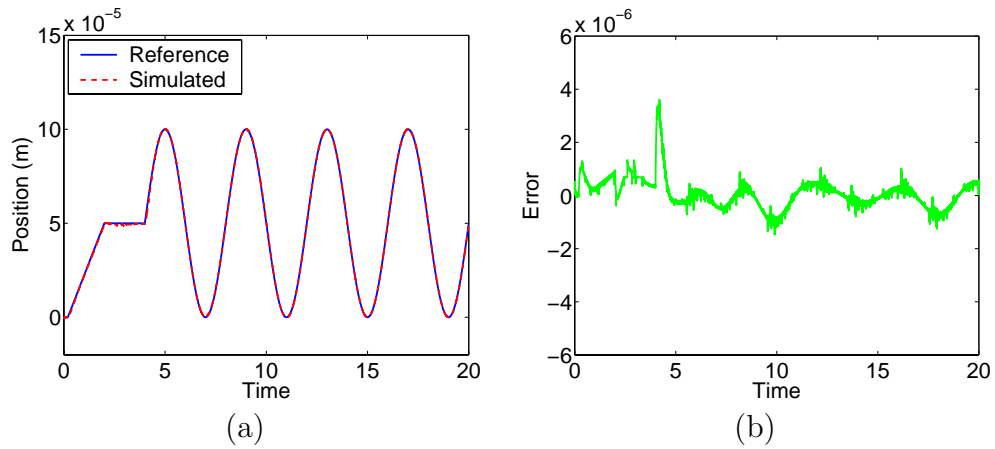


Figure 6.9: \mathcal{H}_∞ design with sensor noise s and the disturbance d due to inversion error. (a) Reference and simulated trajectory, and (b) tracking error.

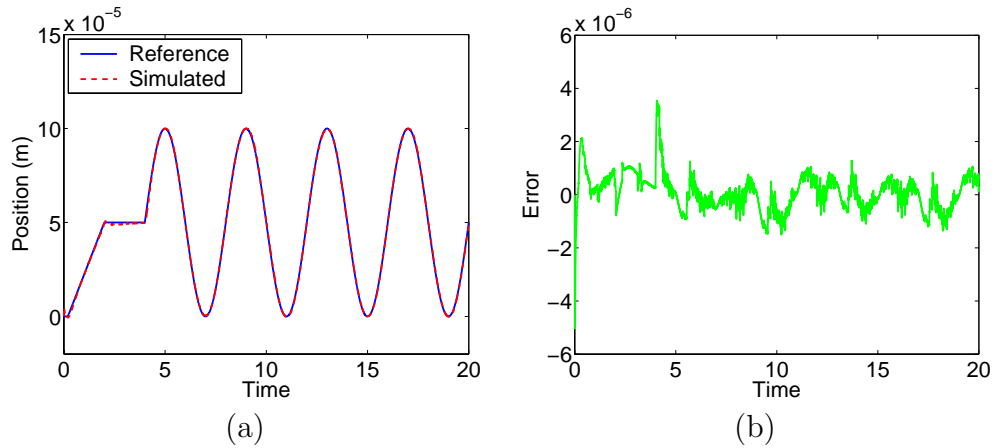


Figure 6.10: \mathcal{H}_∞ design with sensor noise s and the disturbance d due to uncompensated hysteresis and constitutive nonlinearities. (a) Reference and simulated trajectory, and (b) tracking error.

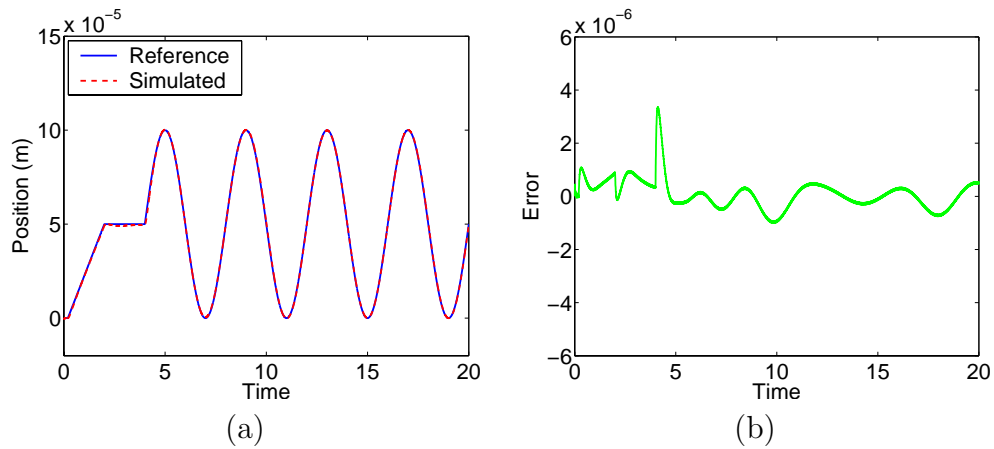


Figure 6.11: \mathcal{H}_∞ design with sensor noise s but no disturbance d . (a) Reference and simulated trajectory, and (b) tracking error.

Chapter 7

Open Loop Control

7.1 Experimental Design

We report here experiments run to ascertain the capability of the model to predict a control for an open loop control implementation. In the initial experiment, data was first collected at the three frequencies 0.279 Hz, 2.79 Hz and 27.9 Hz, with the input field chosen to be a triangle wave with a range of 0 to 7000 V/m. The data was then used to identify parameters for the model in the manner described in Chapter 4. A desired displacement profile for the stacked actuator was specified and the model was inverted and used to estimate the necessary input field to achieve the desired displacement. The predicted input field was then fed to the actuator and the actual displacement was compared to the desired displacement.

For any particular desired output displacement, a second predicted input field was determined by a linear scaling of the desired output displacement. This input field was also fed to the actuator and its ability to reproduce the desired output displacement was compared to that of the model determined input field. The scaling factor was derived from the manufacturer specifications for actuator travel range when the field is increased from the minimum to the maximum allowed values. This is of course a

linear approximation and does not account for the hysteresis. The actuator used for the results given in Section 7.2 was determined to have a scaling factor of 6.86×10^{-1} .

The process used for finding the model inverse is simple. The input field to the forward model was incremented by some fixed stepsize until the output displacement crossed the desired output displacement. The predicted input field was then calculated by interpolation. Smaller stepsizes will result in a more accurate prediction but at the cost of increased computation time.

7.2 Experimental Results

The desired output displacement for the open loop control experiments is chosen to be a triangle wave. For each of the three frequencies, two choices of amplitude are considered. Figure 7.1(a) compares the actual and desired output displacements, Figure 7.1(b) depicts the predicted input fields and Figures 7.1(c) and (d) show the tracking errors. Figures 7.2-7.6 are analogous for each of the other five cases. Table 7.1 shows the residuals for each of these cases when both the model inverse and the linear scaling are used to predict the input field. To avoid having the predicted input fields be outside of the allowable range, lower amplitudes must be chosen as the frequency increases. It is observed that the predicted input field from the model inverse performs up to a factor of ten better than the predicted input field from the linear scaling.

Frequency (Hz)	0.279	0.279	2.79	2.79	27.9	27.9
Amplitude (μm)	40.56	27.04	33.80	27.04	27.04	20.28
Model	0.7934	0.4836	0.8145	0.9283	0.9083	0.9954
Linear	3.1425	1.8070	3.1123	2.5665	6.2826	4.8981

Table 7.1: Residuals for open loop control for desired output triangle waves of different frequencies and amplitudes.

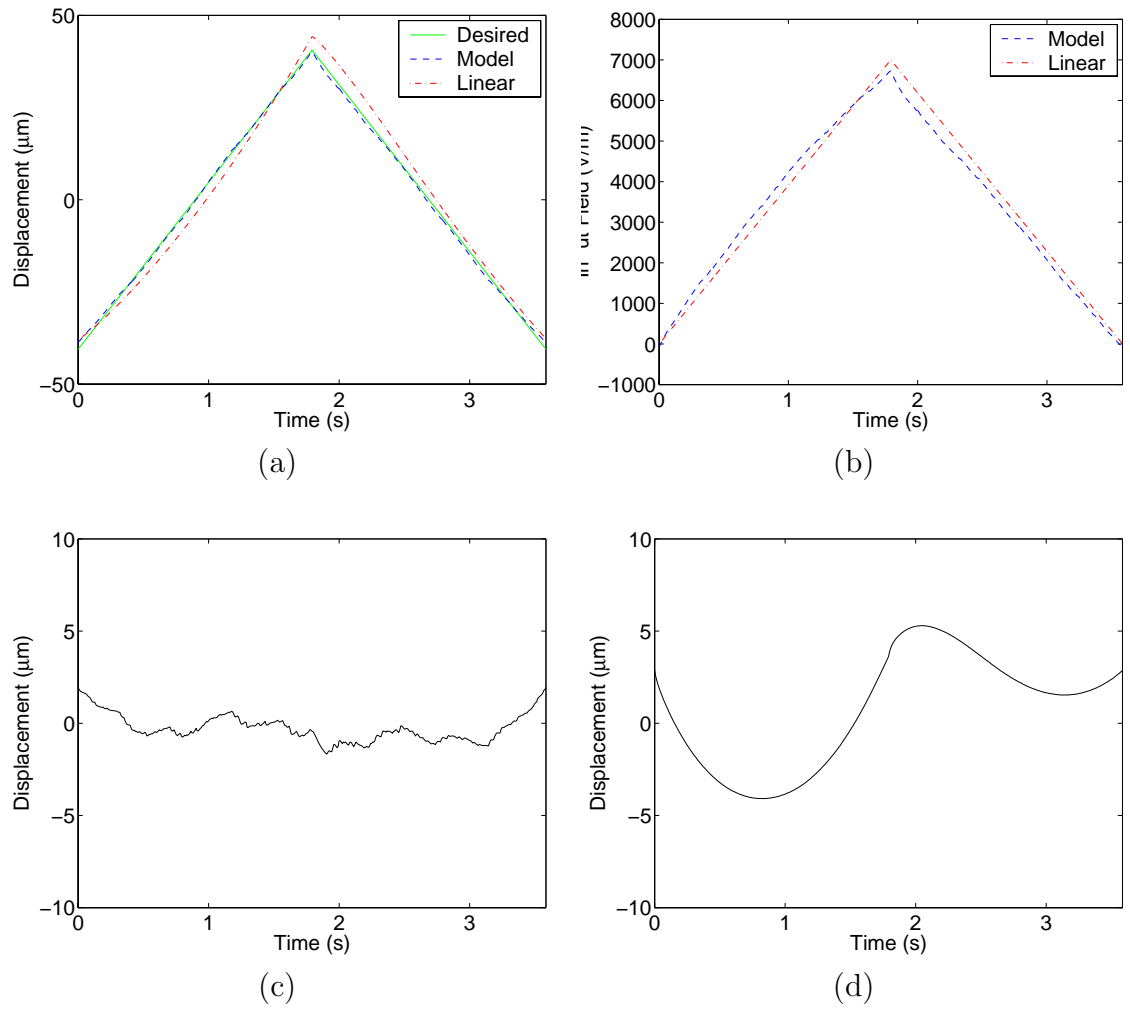


Figure 7.1: (a) Desired displacement with frequency 0.279 Hz and amplitude $40.56 \mu\text{m}$, displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.

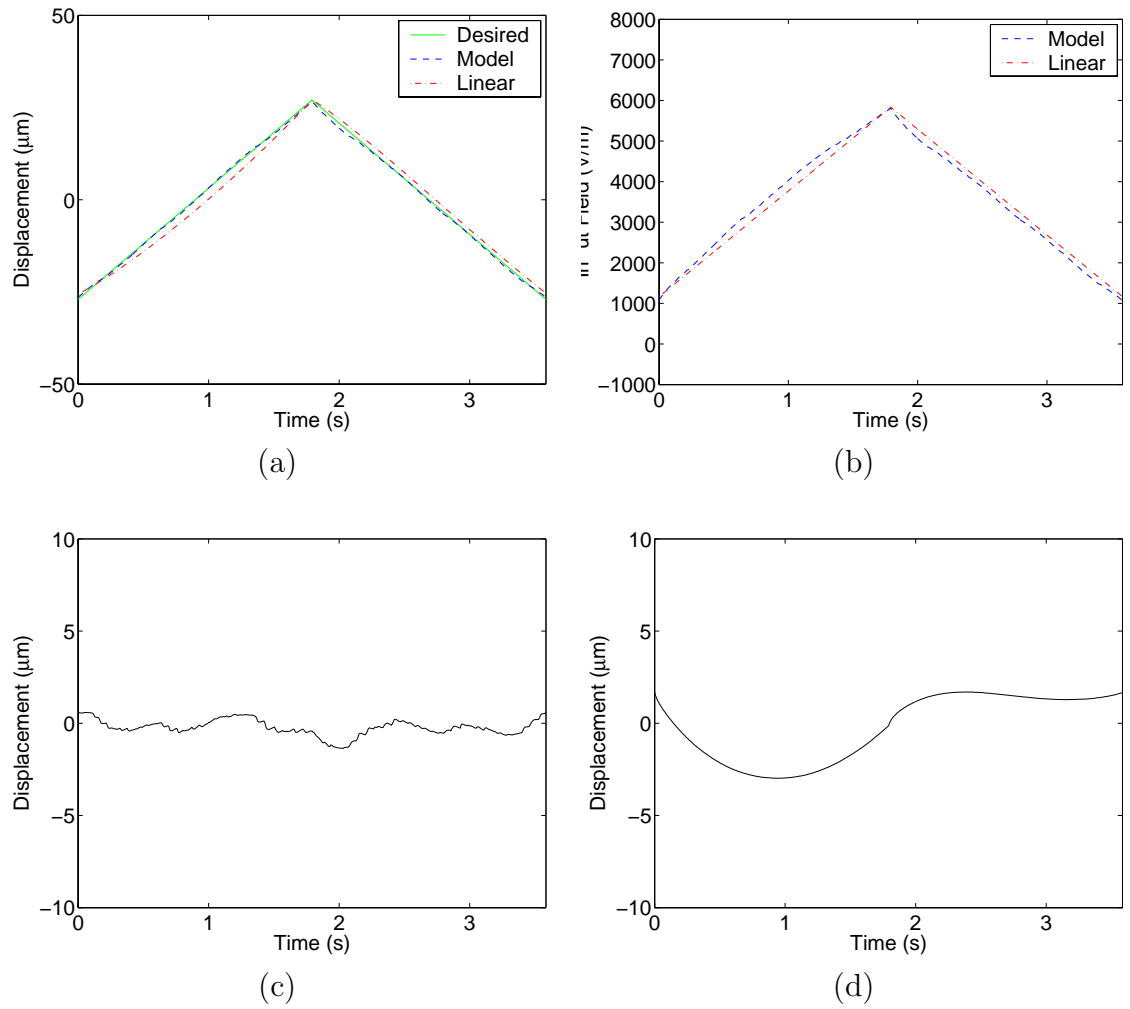


Figure 7.2: (a) Desired displacement with frequency 0.279 Hz and amplitude $27.04 \mu\text{m}$, displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.

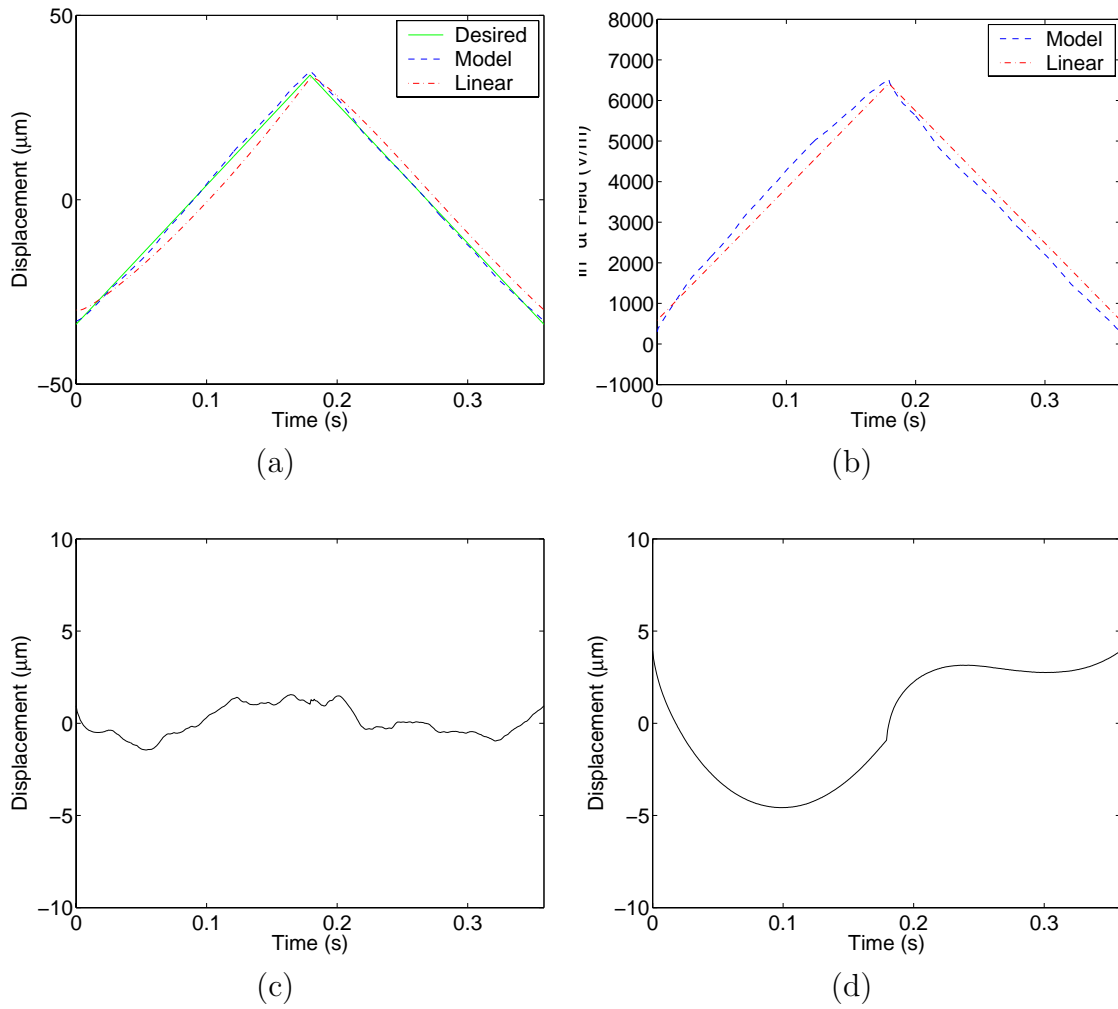


Figure 7.3: (a) Desired displacement with frequency 2.79 Hz and amplitude $33.80 \mu\text{m}$, displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.

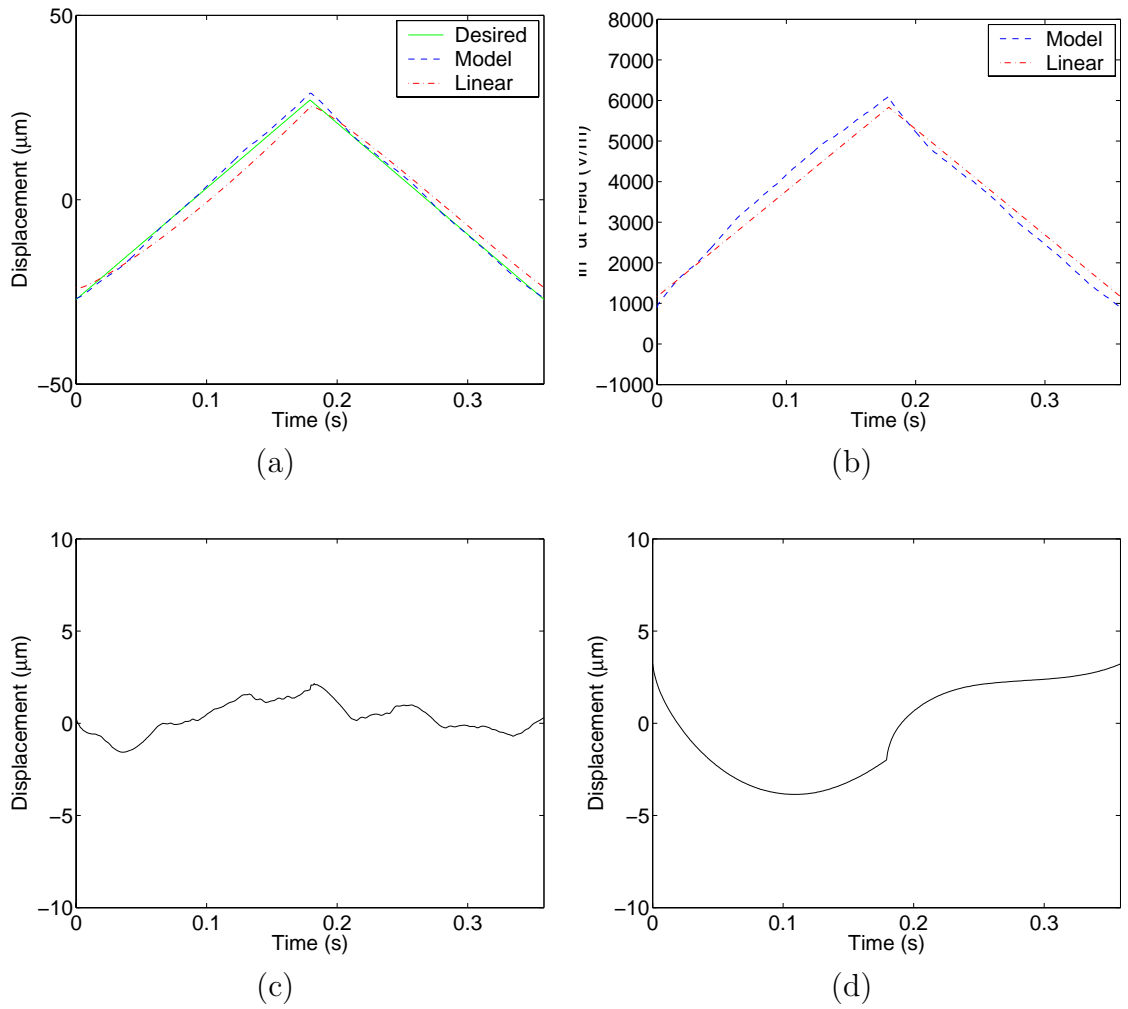


Figure 7.4: (a) Desired displacement with frequency 2.79 Hz and amplitude $27.04 \mu\text{m}$, displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.

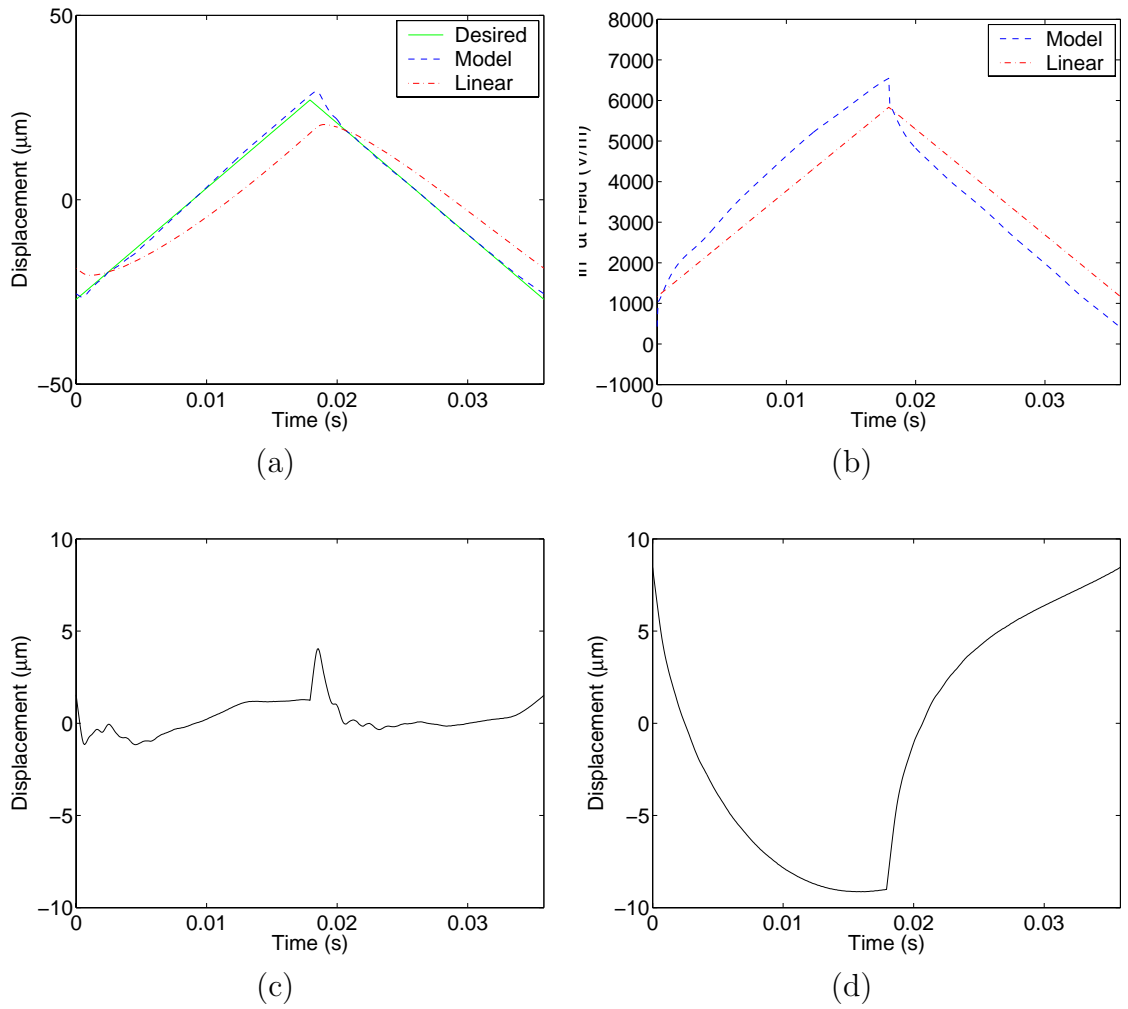


Figure 7.5: (a) Desired displacement with frequency 27.9 Hz and amplitude 27.04 μm , displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.

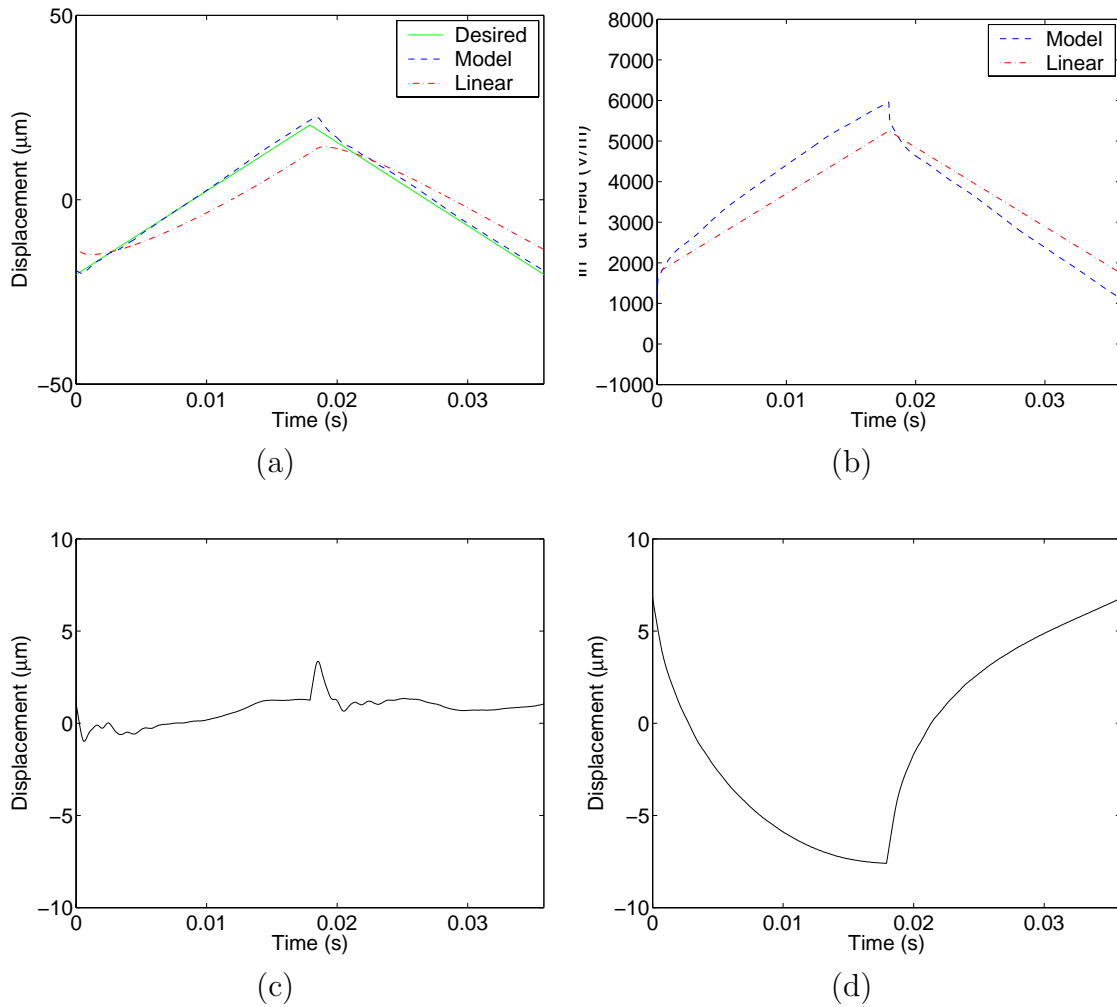


Figure 7.6: (a) Desired displacement with frequency 27.9 Hz and amplitude $20.28 \mu\text{m}$, displacement achieved with input field determined by inverse model and displacement achieved using input field determined by linear scaling. (b) Input field determined by inverse model and input field determined by linear scaling. (c) Tracking error for input field determined from model inverse. (d) Tracking error for input field determined from linear scaling.

The major sources of error in the experiment are error in the model fit, error in the inversion process and non-repeatability of the data. Errors of less than $1 \mu\text{m}$ in the model fit are achievable using the general densities ν_1 and ν_2 . The error in the inversion process depends directly on the stepsize. However, every decrease in the stepsize causes an increase in computation time. This error is much smaller than that from the other two sources. Non-repeatability of the data is seen when the data is collected again after the open loop experiment is conducted. The data does not match the data collected initially for the model fit. Figures 7.7(a) and (b) show the data collected before and after the experiment for the frequencies 0.279 Hz and 27.9 Hz. In both cases, the second set differs by a significant amount with a residual of $\mathcal{R} = 1.5102$ for the low frequency set and a residual of $\mathcal{R} = 1.9749$ for the high frequency set. Note that this is larger than the typical model error. One possible explanation is that heating of the actuator occurs when the data is collected and when the experiment is run.

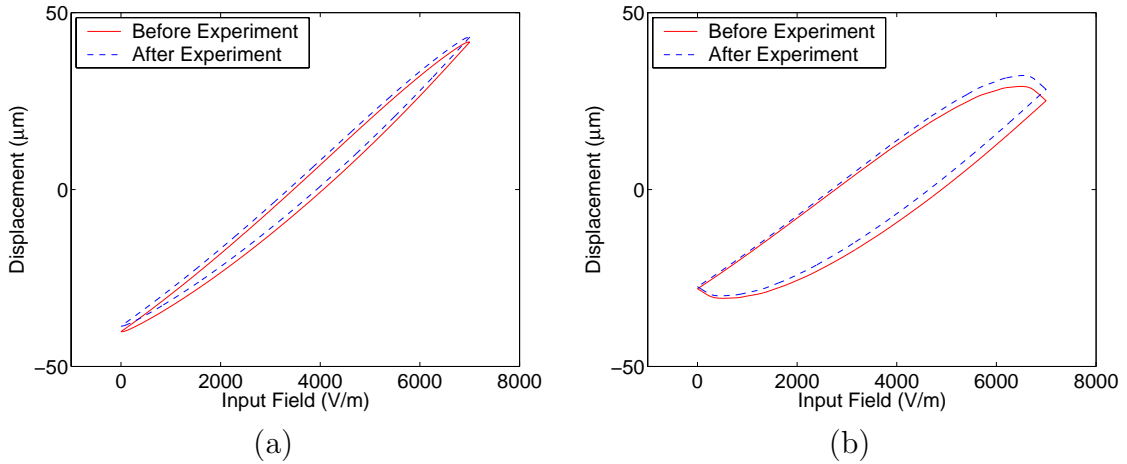


Figure 7.7: Data collected before and after open loop experiments at (a) 0.279 Hz and (b) 27.9 Hz.

Chapter 8

Conclusions and Future Directions

We have developed macroscopic constitutive relations and models to characterize the hysteretic drive dynamics of a PZT-based AFM stage. These models extend previous formulations through the use of general density representations, which provide high accuracy while maintaining implementational efficiency. In particular, the ability of the models to characterize both PZT data and frequency dependent AFM data validates the modeling approach. The modeling framework also facilitates inversion for linear control design.

The performance of initial PID and robust control designs for high accuracy AFM drive regimes was illustrated through numerical examples. The model-based control designs were able to maintain a tracking error of less than $2 \mu\text{m}$ during the tracking phases. Results from an open loop control experiment, employing the model inverse, on a PZT stacked actuator were given as well. In this experiment, the objective was to demonstrate that a model-based inverse could be used to predict the correct input field required to achieve a specified output displacement trajectory. The tracking errors achieved with this filter were reduced by a factor of approximately ten when compared with the unfiltered case, which used only a linear scaling of the output displacement to predict the input field.

The use of adaptive control algorithms is one possible option to account for the non-repeatability of data in the open loop experiment. This approach is feasible because of the relatively small changes in the data for an experimental run with equivalent input conditions. To implement the model in a closed loop control, the speed of the inverse algorithm must be significantly increased. For example, updating the stepsize adaptively instead of using a fixed stepsize could lead to significantly faster runtimes. Future research will focus on these issues.

List of References

- [1] H.T. Banks, R.C. Smith and Y. Wang, *Smart Material Structures: Modeling, Estimation and Control*, Mason/John Wiley, Paris/Chichester, 1996.
- [2] D. Croft, G. Shed and S. Devasia, “Creep, hysteresis, and vibration compensation for piezoactuators: Atomic force microscopy application,” *Journal of Dynamic Systems, Measurement, and Control*, 23, pp. 35-43, 2001.
- [3] A. Daniele, S. Salapaka, M.V. Salapaka and M. Dahleh, “Piezoelectric scanners for atomic force microscopes: Design of lateral sensors, identification and control,” Proceedings of the America Control Conference, San Diego, CA, pp. 253–257, 1999.
- [4] DARPA Program on Molecular Observation, Spectroscopy and Imaging Using Cantilevers (MOSAIC), <http://www.darpa.mil/dso/thrust/biosci/mosaic.htm>.
- [5] R.C.H. del Rosario and R.C. Smith, “Spline approximation of thin shell dynamics”, *International Journal for Numerical Methods in Engineering*, 40, pp. 2807-2840, 1997.
- [6] P. Ge and M. Jouaneh, “Modeling hysteresis in piezoceramic actuators,” *Precision Engineering*, 17, pp. 211-221, 1995.

- [7] C.W. Groetsch, *The Theory of Tikhonov Regularization for Fredholm Equations of the First Kind*, Pitman, Boston, 1984.
- [8] P.K. Hansma, V.B. Elings, O. Marti and C.E. Bracker, "Scanning tunneling microscopy and atomic force microscopy: Application to biology and technology," *Science*, 242, pp. 209-242, 1988.
- [9] A.G. Hatch, R.C. Smith and T. De, "Model development and control design for high speed atomic force microscopy," Proceedings of the SPIE, Smart Structures and Materials 2004, San Diego, CA, Vol. 5383, pp. 457-468, 2004.
- [10] P.D. Hislop and I.M. Sigal, *Introduction to Spectral Theory with Applications to Schrödinger Operators*, Springer-Verlag, New York, 1996.
- [11] D. Hughes and J.T. Wen, "Preisach modeling of piezoceramic and shape memory alloy hysteresis," *Smart Materials and Structures*, 6, pp. 287-300, 1997.
- [12] D.C. Jiles and D.L. Atherton, "Theory of ferromagnetic hysteresis," *Magnetism and Magnetic Materials*, 134, pp. 143-160, 1994.
- [13] T.K. Kiong, W. Quing-Guo and H.C. Chieh, T.J. Hägglund. *Advances in PID Control*, Springer Verlag, London, 1999.
- [14] M.E. Lines and A.M. Glass, *Principles and Applications of Ferroelectrics and Related Materials*, Oxford University Press, Oxford, UK, 1977, Oxford Classics Series, 2001.
- [15] J. Luan and F.C. Lee, "Design of a high frequency switching amplifier for smart material actuators with improved current mode control," PESC 1998 Record, 29th Annual Power Electronics Specialists Conference, Vol. 1, pp. 59-64, 1998.

- [16] J.A. Main and E. Garcia, "Design impact of piezoelectric actuator nonlinearities," *Journal of Guidance, Control, and Dynamics*, pp. 327-332, 1997.
- [17] J.A. Main and E. Garcia, "Piezoelectric stack actuators and control system design: strategies and pitfalls," *Journal of Guidance, Control, and Dynamics*, 20(3), pp. 479-485, 1997.
- [18] J.A. Main, E. Garcia and D.V. Newton, "Precision position control of piezoelectric actuators using charge feedback," *Journal of Guidance, Control, and Dynamics*, 18(5), pp. 1068-1073, 1995.
- [19] J.A. Main, D. Newton, L. Massengil and E. Garcia, "Efficient power amplifiers for piezoelectric applications," *Smart Materials and Structures*, 5(6), pp. 766-775, 1996.
- [20] I.D. Mayergoyz, *Mathematical Models of Hysteresis*, Springer-Verlag, New York, 1991.
- [21] I.D. Mayergoyz, *Nonlinear Diffusion of Electromagnetic Fields with Applications to Eddy Currents and Superconductivity*, Academic Press, New York, 1998.
- [22] A.W. Naylor and G.R. Sell, *Linear Operator Theory in Engineering and Science*, Springer-Verlag, New York, 2000.
- [23] J.M. Nealis and R.C. Smith, " \mathcal{H}_∞ Control Design for a Magnetostrictive Transducer," Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, HA, Vol. 2, pp. 1801-1806, 2003.
- [24] J.M. Nealis and R.C. Smith, "Robust Control of a Magnetostrictive Actuator," Proceedings of the SPIE, Smart Structures and Materials 2003, Vol. 5049, pp. 221-232, 2003.

- [25] J. Nealis and R.C. Smith “Model-based robust control design for magnetostrictive transducers operating in hysteretic and nonlinear regimes,” CRSC Technical Report CRSC-TR03-25; *IEEE Transactions on Automatic Control*, submitted.
- [26] F. Preisach, “Über die magnetische nachwirkung,” *Zeitschrift für Physik*, 94, pp. 277-302, 1935.
- [27] P.M. Prenter, *Splines and Variational Methods*, Wiley, New York, 1975.
- [28] G. Robert, D. Damjanovic and N. Setter, “Preisach modeling of piezoelectric nonlinearity in ferroelectric ceramics,” *Journal of Applied Physics*, 89(9), pp. 5067-5074, 2001.
- [29] D. Rugar, O. Züger, S.T. Hoen, C.S. Yannoni, H.-M. Vieth and R.D. Kendrick, “Force detection of nuclear magnetic resonance,” *Science*, 264, pp. 1560-1563, 1994.
- [30] S. Salapaka, A. Sebastian, J.P. Cleveland and M.V. Salapaka, “High bandwidth nano-positioner: A robust control approach,” *Review of Scientific Instruments*, 73(9), pp. 3232-3241, 2002.
- [31] Skeletal Tissues Research Groups, Leeds Dental Institute, University of Leeds, <http://www.leeds.ac.uk/dental/research/skeletal/skelprojs.htm>
- [32] M.E. Shirley and R. Venkataraman, “On the identification of Preisach measures,” Proceedings of the SPIE, Smart Structures and Materials, 2003, Modeling, Signal Processing, and Control, Vol. 5049, pp. 326-336, 2003.
- [33] R.C. Smith and A. Hatch, “Parameter estimation techniques for nonlinear hysteresis models,” Proceedings of the SPIE, Smart Structures and Materials, 2004, San Diego, CA, Vol. 5383, pp. 155-163, 2004.

- [34] R.C. Smith, A. Hatch and T. De, "Model development for piezoceramic nanopositioners," Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, HA, Vol. 3, pp. 2638-2643, 2003.
- [35] R.C. Smith, A. Hatch, B. Mukherjee and S. Liu, "A homogenized energy model for hysteresis in ferroelectric materials: General density formulation," CRSC Technical Report CRSC-TRO4-23; *Journal of Intelligent Material Systems and Structures*, accepted.
- [36] R.C. Smith and C.L. Hom, "Domain wall theory for ferroelectric hysteresis," *Journal of Intelligent Material Systems and Structures*, 10(3), pp. 195-213, 1999.
- [37] R.C. Smith and Z. Ounaies, "A domain wall model for hysteresis in piezoelectric materials," *Journal of Intelligent Material Systems and Structures*, 11(1), pp. 62-79, 2000.
- [38] R.C. Smith and Z. Ounaies, "Model development for high performance piezoelectric polymers," Material Research Society Symposium Proceedings Vol. 698, pp. 217-222, 2002.
- [39] R.C. Smith and M. Salapaka, "Model development for the positioning mechanisms in an atomic force microscope," *International Series of Numerical Mathematics*, Vol. 143, pp. 249-269, 2002.
- [40] R.C. Smith, M.V. Salapaka, A. Hatch, J. Smith and T. De, "Model development and inverse compensator design for high speed nanopositioning," Proceedings of the 41st IEEE Conference on Decision and Control, Las Vegas, NV, Vol. 4, pp 3652-3657, 2002.
- [41] R.C. Smith, S. Seelecke, Z. Ounaies and J. Smith, "A free energy model for

- hysteresis in ferroelectric materials,” *Journal of Intelligent Material Systems and Structures*, 14(11), pp. 719–739, 2003.
- [42] P.N. Sreeran, G. Salvady and N.G. Naganathan, “Hysteresis prediction for a piezoceramic material system,” Proceedings of the ASME Winter Annual Meeting, New Orleans, LA, Volume AD-VOL 35, pp. 35-42, 1993.
- [43] X. Tan, R. Venkataraman and P.S. Krishnaprasad, “Control of hysteresis: Theory and experimental results,” Proceedings of the SPIE, Smart Structures and Materials, 2001, Modeling Signal Processing and Control in Smart Structures, Vol. 4326, pp. 101-112, 2001.
- [44] C.R. Vogel, *Computational Methods for Inverse Problems*, SIAM, Philadelphia, 2002.
- [45] W. Voigt, *Lehrbuch der Kristallphysik*, B.G. Teubner, Leipzig, Berlin, 1928.
- [46] S.A. Wolf, D.D. Awschalom, R.A. Buhrman, J.M. Daughton, S. von Molnár, M.L. Chtchelkanova and D.M. Teger, “Spintronics: A spin-based electronics vision for the future,” *Science*, 294, pp. 1488-1495, 2001.
- [47] K. Zhou and J.C. Doyle, *Essentials of Robust Control*, Prentice Hall, New Jersey, 1998.