POLYMER GEL ACTUATORS AND SENSORS

by

Ahmed Hamdi Mitwalli

S.B., Massachusetts Institute of Technology, 1991

S.M., Massachusetts Institute of Technology, 1993

Submitted to the Department of
Electrical Engineering and Computer Science
in partial fulfillment of the requirements
for the degree of

DOCTOR OF SCIENCE

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY

May 1998

Signature of Author _____
Department of Electrical Engineering and Computer Science
May 15, 1998

Certified by _____
Steven B. Leeb
Associate Professor of Electrical Engineering
Thesis Supervisor

Accepted by _____
Arthur C. Smith
Chairman, Department Committee on Graduate Students

POLYMER GEL ACTUATORS AND SENSORS

by

Ahmed Hamdi Mitwalli

Submitted to the Department of Electrical Engineering and Computer Science
on May 15, 1998 in partial fulfillment of the requirements for the degree of
Doctor of Science in Electrical Engineering

ABSTRACT

Gels consist of a cross-linked network of polymers suspended in a solvent. Under certain conditions, gels have been observed to undergo reversible changes in volume. These volume changes may be over 1000-fold and can be triggered by a variety of electrochemical conditions, including changes in temperature, solvent composition, or pH level. Gels could in principle be used as actuators in servomechanisms and sensors, which range in size from microscopic (silicon) mechanisms to mechanisms comparable in size and force density to biological systems. For example, polymer gels could act as synthetic muscles that provide direct linear motion with useful force densities. Several properties of polymer gels make them a potentially powerful engineering material that can be used in a wide variety of electromechanical devices. This thesis explores techniques for applying polymer gels as active components in engineering systems.

Chemical techniques and fabrication methods for optimizing the material for engineering applications are explored. For example, methods for increasing the tensile strength of polymer gels are discussed. Many polymer gels exhibit discontinuous equilibrium volume-phase transition curves. This makes them difficult to use, for example, as position regulators in servomechanisms. Fast switching of the gel trigger mechanism can be used to establish a dynamic equilibrium in cases where a static one is not possible. A prototype switched polymer gel actuator with pulse-width modulation (PWM) control is presented in this thesis. A dynamic model suitable for control design is derived and applied for the prototype system. The modeling and control results are generalized to a variety of switched systems. Two engineering applications of polymer gels are further addressed. First, a remotely-triggered magnetically-activated gel regulates the position of a suspended load. Second, a gel transition point sensor is used to detect the presence of metal ions in a solvent.

Thesis Supervisor: Steven B. Leeb
Thesis Supervisor Title: Associate Professor of Electrical Engineering

ACKNOWLEDGEMENTS

I thank my thesis supervisor Steven Leeb for his support and guidance. I have worked with him over the last seven years, on both my Master's and Doctoral degrees. During this time, he has been a source of constant support. Professor Leeb combines a strong understanding of theory and extensive practical experience. These attributes, as well as his enthusiasm and constant availability were critical to my successful completion of this work.

I also thank my thesis committee readers. This work would not have been possible without the guidance of Professor Toyoichi Tanaka. He was a constant source of help as the leading authority on polymer gels. Professor George Verghese made significant contributions to the mathematical modeling results in this thesis, and Professor James Kirtley provided good advice on the document structure.

I thank Professor Jacob White, my academic advisor, for his support and guidance. His kind, caring personality and his technical brilliance were always a source of inspiration.

Essential hardware and test equipment for this project was made available through generous donations from the Intel Corporation, Tektronix, and Hewlett-Packard. This research was also thankfully and generously supported by the NSF, CMSE, AT&T, and LFM.

I thank everyone in LEES lab. Deron Jackson, Tim Denison, and Steve Shaw are all exceptional engineers who provided a great deal of help throughout my project. Dave Perreault and Mark Thompson, my office mates for several years, endured my many questions. Steve Nagle, Craig Abler, Mary Jane Boyd, and Judy Cardell were always a source of good cheer. I would also like to thank John Ofori, Alex Stankovic, Kamakshi Srinivasan, Jeff Chapman, Haatchitaba Mweene, Mary Tolikas, Umair Khan, Aaron Schultz, and Vivian Mizuno.

Paolo Narvaez, Dahlene Fusco, and Dennis Evangelista assisted with parts of this work. I thank them for their contributions.

My great friends deserve thanks for their support throughout the last few years. Frank, Francis, Rajan, Sanjeev, Firdaus, Karim, Piero, Amr, Ghassan, Bassam, Riad, Dina, and Dana have always been there.

I would especially like to thank my loving parents, my sister Dalia, and my brother Mohamed. You have always been there and have given me everything. I love you all with all my heart. And to Nawal, who has been through the highs and lows with me the last couple of years: I thank you; I love you.

Contents

List of Figures

Chapter 1

Introduction and Background

For many electromechanical systems, the available actuator and sensor technology dictates significant performance limitations. For example, the absence of a conformable, flexible actuator with reasonable strength has hindered the development of dexterous multi-link manipulators. A polymer gel actuator may provide the solution for this and other applications. A gel consists of a tangled network of cross-linked polymers immersed in a solvent. Depending on how they are fabricated, polymer gels can exhibit abrupt, reversible volume changes in response to a variety of stimuli. These volume changes can be used as a sensor for a particular stimulus or as an actuator that converts some form of input energy to mechanical energy. Several properties of polymer gels make them a potentially powerful engineering material that can be used in a wide variety of electromechanical devices. While polymer-based materials serve as structural and insulating components in electrical devices, fabrics, and adhesives, few engineering systems use polymers as *active* components. This thesis explores the application of polymer gels as actuators and sensors. In this chapter, motivation for the thesis goals and background on polymer gels are presented, followed by an outline for the rest of the document.

1.1 Motivation and Goals

Mathematical modeling of complex mechanical systems, such as multi-degree-of-freedom robotic manipulators, is relatively well-developed [2]. With the advent of the micropro-

cessor and the rapid decrease in its price-to-performance ratio, it is possible to implement increasingly sophisticated control schemes. Performance limitations in servomechanical and robotic systems are often dictated by available sensor and actuator technologies.

A variety of actuator technologies are commonly used in motion control applications. Depending on the specific application requirements, different actuators may be more appropriate to use. For instance, the electric motor is very efficient for shaft powers over a few hundred watts and is therefore often used in high power applications. Electrostatic actuators generate forces as a result of the difference in electric potential between two conducting objects. Electrostatic forces become more significant when the distance between these objects is relatively small. Hence, these actuators are popular, for example, for implementing micromotors designed on silicon wafers [12]. Piezoelectric materials exhibit a change in strain when subjected to an electric field. Maximum strains on the order of 200-300 microstrains are possible[1], with research agressively pursuing a composition that will yield a 1% strain [86]. Frequency responses around 150 KHz have been observed [12]. Piezoelectric actuators are therefore used in high precision or microactuation applications. Research on novel materials for actuator design has produced magnetostrictive materials, which deform in response to applied magnetic fields. Strains on the order of 0.2% have been observed [87]. Other materials, such as shape memory alloys and electrorheological fluids, have also generated interest for actuator design. More on these materials can be found in [12],[7], [87], and [86].

Many of these actuators suffer from limitations that may be prohibitive in some applications. For example, a robotic, flexible multi-link manipulator that is capable of a range of motion and maneuverability similar to that of the human arm is still essentially unattainable. The reason is largely the absence of a soft, flexible actuator material that can be made into "synthetic muscles" to move the manipulator. Implementation of a manipulator with electromagnetic motors usually involves a multijoint system with a motor at each of several revolute joints. As more degrees of freedom are required, the number of joints increases, and so does the weight of the structure. The motors in the earlier stages of the kinematic

[1]More common strain values are 20-30 microstrains.

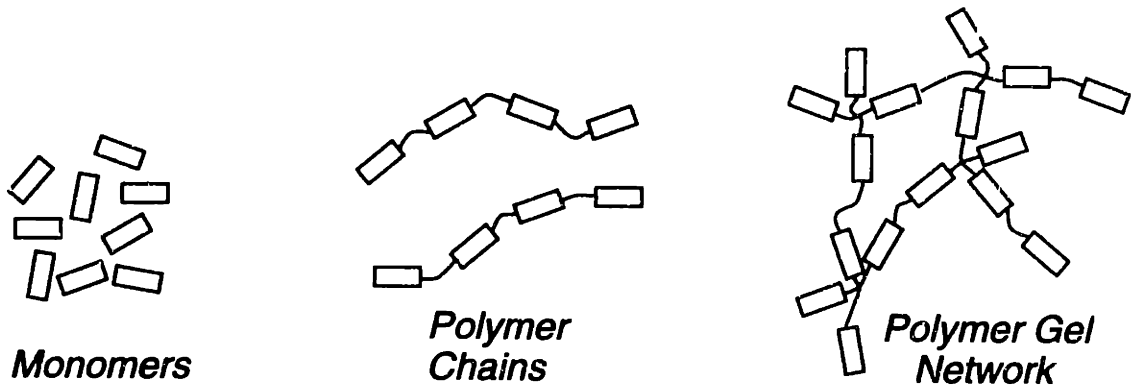


Figure 1.1: Polymer gel

chain need to have high torque ratings to support the various joints and the additional weight of the motors further down the manipulator. More than any of the conventional or novel actuators discussed above, polymer gel actuators would be ideal for this application. An overview of polymer gels will demonstrate various properties that make them good candidates for this and other servomechanical applications.

A polymer gel consists of a cross-linked network of polymers suspended in a solvent (See Figure 1.1). When properly constructed, polymer gels can be made to undergo abrupt, discontinuous changes in volume in response to a variety of environmental stimuli. These volume changes can be as large as 1000-fold or more and are caused by the gel network diffusing into and out of the solvent. Figure 1.2 illustrates this process. This volume change can either be used to perform mechanical work or observed to sense a change in some environmental variable. There are several advantages to using polymer gels as actuators and sensors :

- Polymer gels are flexible and conformable. A gel can be routed conformally with the underlying mechanical structure of a servomechanism to optimize weight distribution. The ability of the actuator to bend can also enhance maneuverability.

- The scalability of polymer gels also makes them potentially useful for many applica-

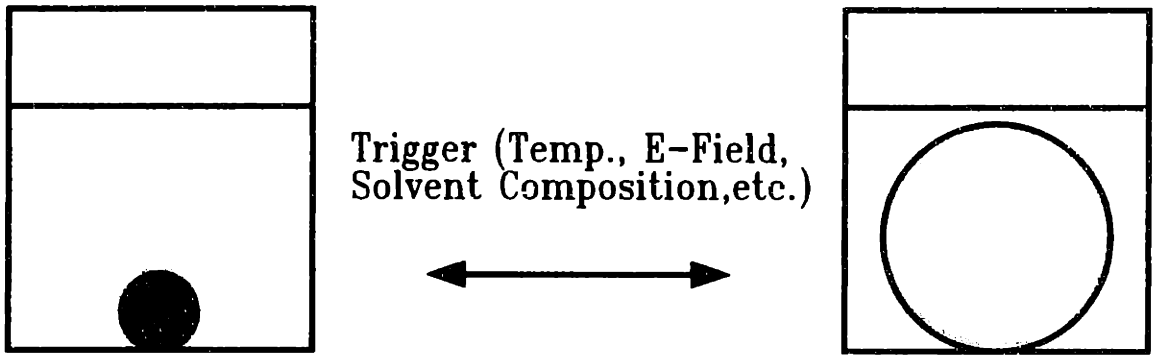


Figure 1.2: A polymer gel at two equilibrium points

tions. The fact that polymer gels retain their ability to change volume and generate force at any size or shape makes them attractive for microactuator applications. A microsized actuator can be constructed by forming the gel in a microscopic mold constructed by silicon machining, for instance.

- Different gels will undergo a volume change in response to a variety of stimuli or triggers. Volume phase-transitions in response to temperature, electric field, light, and solvent pH level have been observed [17]. This thesis will demonstrate the use of magnetically triggered gels for remotely-triggered devices. This property makes gels good candidates for use as actuators and sensors in a variety of applications and environments. Moreover, it may be possible to combine the sensor and actuator in a control application into one gel device. For instance, a valve that controls the flow of a temperature regulating fluid can be made using a thermo-sensitive polymer gel.

- Gels, in principle, have the ability to store large amounts of energy and release them with a small trigger or stimulus. A very large change in volume can occur in response to an infinitesimally small change in temperature, for instance.

The force to mass ratios and efficiency of polymer gels promise to be comparable to those of biological muscles, making them potentially practical actuators with significant advantages

over other actuation technologies for a variety of applications [8].

As a result, there has been an increasing interest in the use of polymer gels as sensor and actuators. Several innovative attempts at designing and building mechanical gel devices have been made. For example, a gel "looper" (an inchworm-like device that moves by repeatedly curling and straightening itself) was developed and demonstrated by Osada [82]. An artificial finger and an artificial flower constructed from thermo-sensitive polymer gels are presented in [47]. Similar examples are discussed in [94] and [93]. Other work has studied mechanical properties of certain polymer gels. Most research to date has been directed towards understanding static or equilibrium mechanical properties, such as tensile strength and force density [111] [90] [33] [34] [21]. The effect of mechanical load or tension on the equilibrium behavior of polymer gels has also been examined [44] [92].

Dynamic models for polymer gel actuators are also being researched. Most models, for example those developed in [88] and [23], are too complicated and not suitable for control design. A model for a pH-level activated polymer gel actuator is presented in [8] [52] [51]. The sixth-order model describes the dynamics of the irrigation system used to spray the gel with acid and base, the chemical reaction that produces the change in pH, and the dynamics of the volume transition of the gel. Ignoring the faster chemical processes, this model was reduced to a first-order approximation. Most gel research has observed two distinct rates of response for volume phase transitions [98], [17]. A single exponential governs the dominant behavior of the gel, but not the complete response. Moreover, the model in [52] and [51] only applies to gels that exhibit continuous equilibrium volume profiles.

The process of building a servomechanical system with polymer gel sensors and actuators involves several levels of design. The gel material itself needs to be designed. *Gel design* will be used to refer to the process of altering the chemical composition and shape of a polymer gel used in an application to best meet the application's requirements. *Actuator design* refers to higher level issues such as packaging, power delivery, and mechanical couplings. *System design* involves control, actuator and sensor locations, and the integrated design of polymer gel devices and underlying mechanical structure. Most of the focus of this thesis will be on exploring techniques for *gel design* and *modeling*. First, an understanding of how

to alter the material properties of a gel using its chemical composition and physical shape and size is developed. A simple yet predictive mathematical model of the gel dynamics is then derived. This model applies to a large class of gels, and is independent of the environmental trigger mechanism. Based on this model, control algorithms and techniques for effectively operating a polymer gel device are presented. *Actuator design* and *system design* are addressed in an application context in Chapters 8, 9, and Appendix A.

1.2 Background

The study of polymer network systems can be traced back to experiments on rubber by Gouth in 1805 [17]. His demonstration that rubber shrinks upon heating, unlike other materials, led to the understanding of the entropic nature of rubber elasticity. In the 1930's and 40's, Kuhn, Flory, and others provided quantitative treatments of the rubber network [17], [20], [59]. Much of this work became the basis of theories explaining polymer gel networks. Volume-phase transitions were theoretically predicted in 1967 and the first experimental observations were made ten years later. A good overview of the history and theory of these transitions is given in [17] and [59]. A brief introduction to the theory is presented here as background for work on modeling the gels for mechanical applications.

1.2.1 Physics of Volume-Phase Transition

Forces such as intermolecular interactions, electrical charges, and the pressure of moving particles contribute to the osmotic pressure of a polymer gel. At equilibrium, this osmotic pressure is zero as a balance of the various forces is maintained. At certain points of equilibrium, the balance of forces in the gel may be sufficiently altered by small (even infinitesimal) changes in the environmental variables that a volume-phase transition occurs, resulting in a large change in volume. A positive osmotic pressure tends to increase the volume occupied by the polymer network, and a negative one reduces the volume. The Flory-Huggins model [20] is an equation of state widely used to qualitatively model the balance of pressures on a polymer network. Further studies of the equilibrium behavior were carried out by Flory [20] [17], Dusek [25] [17], and Tanaka [97] [69]. Research on

the kinetics of polymer gels has lagged equilibrium modeling but has received increasing attention (Tanaka [98], Tanaka-Fillmore [99] [100], Li-Tanaka [59] [58], Candau [11]). The following sections discuss the different forces that contribute to osmotic pressure and briefly introduce equilibrium and kinetic gel models.

Osmotic Pressure and Forces Inducing Volume Changes

At least three different forces that resist or aid the volume change of a polymer gel have been identified [97]:

- **Rubber elasticity** is a mechanical property of the gel not unlike that of a rubber band. It arises from the elasticity of individual polymer strands. Rubber elasticity is a bidirectional force that opposes changes in the volume away from an equilibrium state. The freely jointed monomer units are in motion due to thermal energy. When the network is stretched, the motion of the monomer units tends to pull the boundary of the network inward. When compressed, the same motion resists the compression and pushes outward. The magnitude of the rubber elasticity increases with temperature, as thermal energy increases the motion of the monomer units. The sign of this force depends only on the volume with respect to the equilibrium volume of the gel.

- **Polymer-polymer affinity** results from the interactions between polymer strands with solvent molecules or with other polymer strands. This force can be negative (tending to shrink the gel) or positive (tending to expand the gel) depending on the electromechanical properties of the molecules in the solvent and the network. Four molecular interactions which account for the forces between network and solvent molecules are shown in Figure 1.3 (reproduced from [17]). A *van der Waals* force is the result of interactions between neighboring molecules in a non-polar solvent. Water molecules near a *hydrophobic* polymer chain form fixed structures as strong hydrogen bonding causes them to become highly ordered. The result is a reduction in the enthalpy and entropy of the system known as a hydrophobic interaction. *Hydrogen bonding* takes place when a hydrogen atom is located between two closely separated

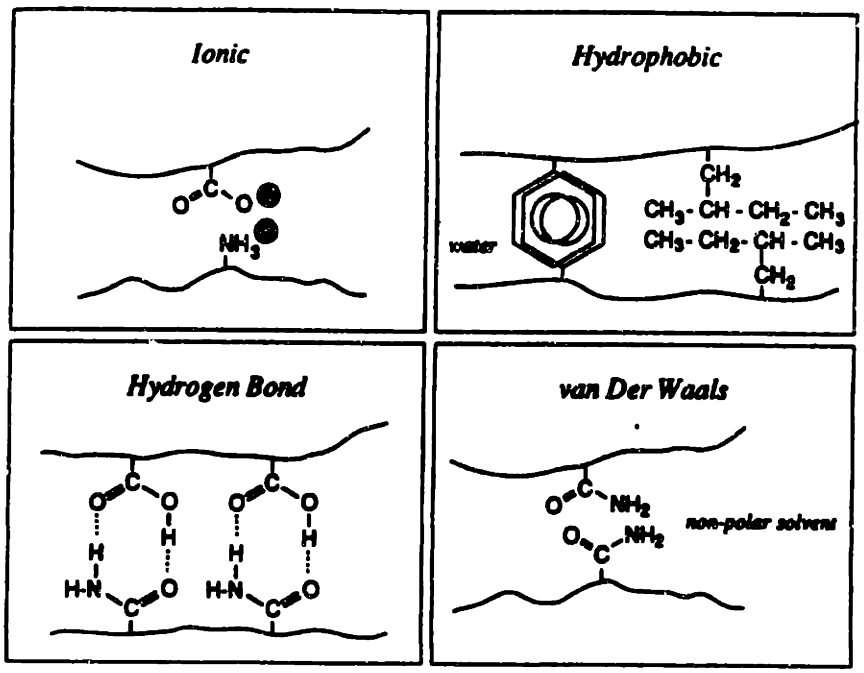


Figure 1.3: Four fundamental molecular interactions. Figure reproduced from [17]

atoms with a net negative charge. *Electrostatic* interactions are created by including positive or negative charges on the polymer chains. If one polarity is introduced and fixed on the network (a "polyelectrolyte"), counter ions position themselves near the chains to maintain electrical neutrality. The relative concentration of the counter ions inside and outside of the gel results in an osmotic pressure[2]. Another interesting electrostatic interaction is observed when both kinds of charges are bound onto the network (a "polyampholyte"). For more on the four molecular interactions refer to [17].

- **Hydrogen ion pressure** is a particular case of a polyelectrolyte interaction. The electrical charge of the hydrogen ions is screened by the background of negative charges in the polymer network. Nevertheless, as long as these ions move freely within the volume of the gel, they behave like a gas in a closed container, the gel network. The

[2]A positive osmotic pressure tends to increase the volume of a gel.

osmotic pressure due to the motion of these hydrogen ions is positive, tending to expand the gel.

Equilibrium Models

In order to understand the origins of the Flory-Huggins model, we review the pressure-volume relationship for an ideal gas and a van der Waals fluid [14]. For an ideal gas, the constitutive relation is

$$PV = mRT \tag{1.1}$$

where m is the mass of the gas, P, V, and T are the pressure, volume, and temperature, and R is the gas constant (which is a material property). For a van der Waals material, the constitutive relationship or equation of state becomes

$$P = \frac{RT}{V-b} - \frac{a}{V^2} \tag{1.2}$$

where b is the volume excluded by the dimensions of the molecules themselves and a accounts for an attractive force between molecules. While the ideal gas model describes a single phase system and does not predict a phase transition, the van der Waals equation can be interpreted as doing so. If it is rewritten as

$$PV^3 - (bP + RT)V^2 + aV - ab = 0 \tag{1.3}$$

the result is a third order polynomial in volume. For a constant temperature below some critical value, the relationship between pressure and volume is a curve similar to the "isotherm" in Figure 1.4. The middle section of the plot is known as a Maxwell loop. The van der Waals equation does not clearly predict a phase-transition. However, the Maxwell loop is a region of negative compressibility or instability that is commonly interpreted as a phase-transition by applying an "equal area" criterion. A line can be drawn across as shown in the figure to divide the regions near the local minimum and maximum to contain equal areas. This

line represents a discontinuity in volume at a particular pressure level. This discontinuous behavior is due to a change in the phase of the system. For more on this see [17] and [14].

Although the van der Waals equation does not represent the properties of any given substance with precision, it is presented here to illustrate the basis for the more complicated Flory-Huggins model. The gel equations of state are based on a pressure balance model similar to that in Equation 1.2. The classical theory for gels based on the Flory-Huggins model and developed further by Tanaka in [17] describes a system that exhibits a phase transition similar to a van der Waals system. The osmotic pressure of the gel Π arises from at least three components:

$$\Pi = \Pi_M + \Pi_{el} + \Pi_{ion} \tag{1.4}$$

The pressure Π_M is the product of the free energy of mixing (polymer-polymer affinity). Elastic forces in the polymer network are represented by Π_{el}, and Π_{ion} is a result of ionic forces that may exist among polymer constituents that carry electric charges. The mathematical assembly of the Flory-Huggins model is presented in [17]. Sample plots of pressure versus volume based on the Flory-Huggins model are shown in Figure 1.5. (The variable χ is the Flory interaction parameter; its interpretation depends on the particular gel.) The pressure-volume plots produced with the Flory-Huggins equation predict behavior that can be interpreted as a phase-transition in gels, just as the van der Waals equation predicts for a van der Waals fluid.

Kinetic Models

In 1979, it was demonstrated that the volume change of a polymer gel is determined by the collective diffusion of the polymer network in the solvent [98]. This process is the basis of recent kinetic modeling of gels. The collective diffusion coefficient D_c describes the rate at which the gel network moves relative to the solvent. The variable D_c can be defined as a function of the elastic modulus of the polymer network k and the frictional coefficient f between the network and the solvent [98], [68]:

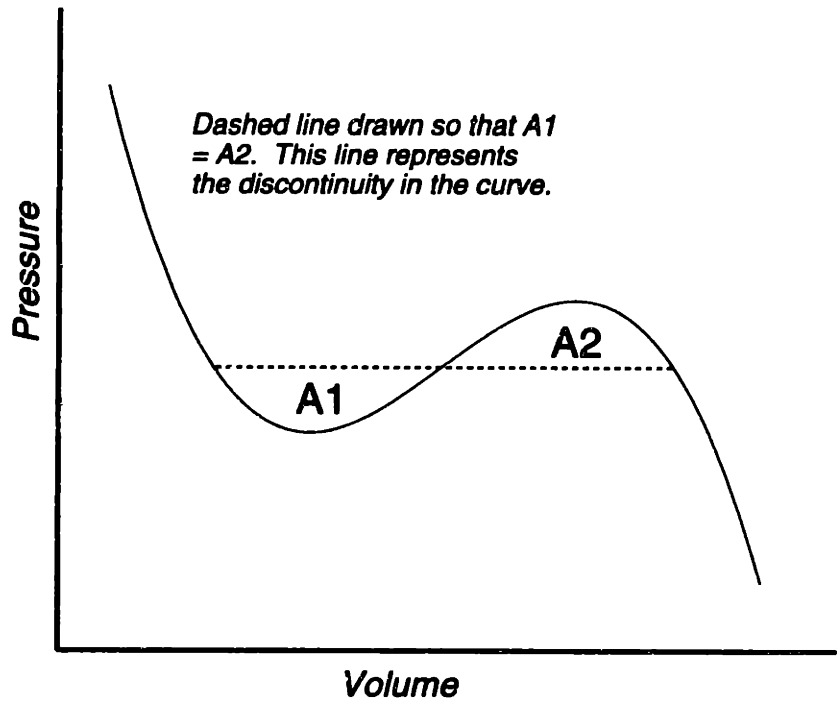


Figure 1.4: Discontinuous phase transition derived from Maxwell loop.

$$D_c = \frac{k}{f}. \tag{1.5}$$

These parameters and the diffusion coefficient have been shown to vary with the temperature, ionic content, and the volume of a gel [69], [98], [59]. In general, the kinetics of polymer gel volume-phase transitions are dominated by a single exponential that is a function of D_c and the size and shape of the polymer gel.

For certain symmetric shapes of gels, the kinetics of the volume change are only a function of a *critical dimension* of the gel. The critical dimension of a gel is generally its smallest dimension [59]. Since the volume phase transition is a diffusion process, it is fastest along this dimension. With a finite diffusion time and zero or negligible shear time [59], as soon as the gel changes in one dimension due to diffusion, the shearing due to the interconnection of the network causes the rest of the polymer gel to follow that

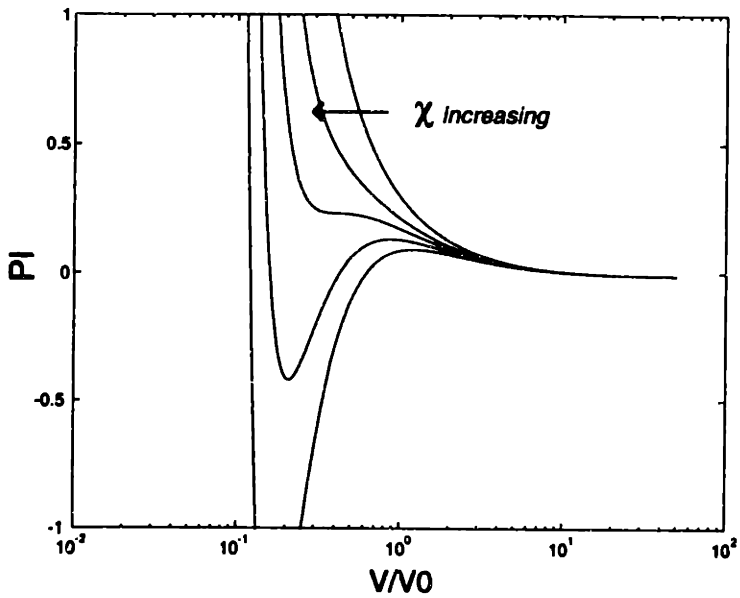


Figure 1.5: Plots of pressure versus volume based on the Flory-Huggins model

change. The result is an isotropic swelling or shrinking of the gel when unloaded and uniformly cross-linked. This is illustrated in Figure 1.6. The diffusion process in which the network undergoes a relative motion with respect to the solvent occurs mostly in the critical dimension. (This approximation is valid when there is a significant discrepancy between the shortest length scale and the rest of the dimensions of the gel.) In a sphere, this occurs in all three dimensions. It takes place in two and one dimensions in a cylinder and a disc, respectively. For the same material, with the same critical or "shortest" length from the center to the surface, a sphere will change its volume faster than a cylinder, and a cylinder faster than a disc. If the diffusion coefficient for a sphere, a cylinder, and a disk with the same critical length are defined as D_s, D_c, and D_d, respectively, the following approximate relationships hold [17]:

$$D_d = \frac{1}{2}D_c = \frac{1}{3}D_s. \tag{1.6}$$

Reference [98] discusses a nonlinear model for the dynamics of polymer gels. A set of partial

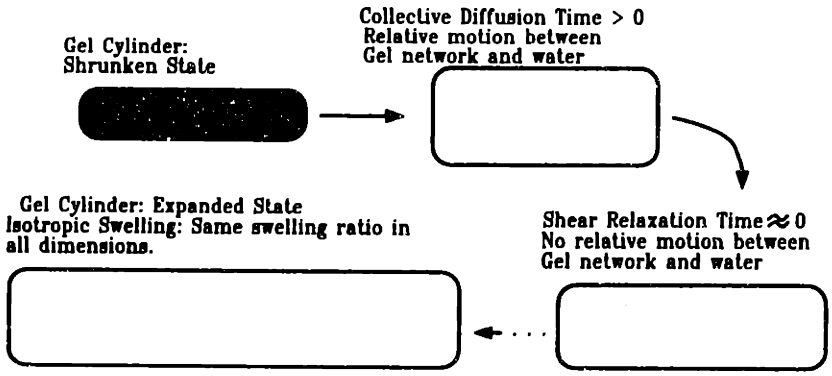


Figure 1.6: Isotropic volume changes due to short shear time constants

differential equations are used to describe the kinetics of a volume-phase transition. Based on these equations, an approximation for the dominant time constant of volume-phase transitions in polymer gels is presented in [17] and [98]:

$$\tau = L^2/D \tag{1.7}$$

Minimization of the critical dimension L can help achieve faster rates of response. When the volume change is small, the solution can be approximated by an infinite sum of exponentials [69]. This formulation is discussed further in Chapter 6.

1.2.2 NIPA

One gel formula that has been the subject of many experiments at MIT and in the polymer gel literature in general uses *N*-isopropylacrylamide (hereafter abbreviated as NIPA) as a base monomer [31]. Throughout this thesis, NIPA gel is used for analysis and experiments. NIPA is a thermo-sensitive gel that exhibits a phase-transition in response to a temperature change. This is especially convenient since a temperature change is readily and accurately achieved with a temperature-controlled fluid circulator. NIPA gels are also relatively easy to prepare, and their equilibirum physics and chemistry are well understood. In addition, with appropriate chemical modifications, they have been shown to respond to other stimuli such as light [95]. The results of the analysis and experiments in this thesis are general

and applicable to other polymer gels. Requirements on force density, tensile strength, and trigger mechanism, among other characteristics, would ultimately be used to choose the appropriate gel formula required for a particular application.

In a pure NIPA gel in an aqueous solvent, a balance between hydrophobic and hydrophillic interactions controls the gel state. Below the phase transition temperature, the polymer chains are relatively hydrophyllic, and the gel swells. Water forms a relatively ordered structure around the swollen chains. At higher temperatures, the ordered structure of the water is destroyed (above the transition temperature). The polymer chains are relatively hydrophobic, and the gel shrinks.

Reference [17] derives a relationship between gel volume and temperature from the Flory-Huggins equation for a NIPA gel. Figure 1.7 shows plots of curves generated with this model. The same Maxwell loop construction used for the van der Waals pressure plots can be used to predict the discontinuity point for these curves as well. Different curves are plotted for different concentrations of ionic groups in the gel. By incorporating ionic groups in the gel network, the equilibrium behavior can be altered as shown. A larger degree of ionization f results in a larger discontinuity in volume. Details regarding the fabrication of NIPA gels are discussed in the next chapter.

1.3 Thesis Organization

This chapter has presented the motivation for using polymer gels as sensors and mechanical actuators and provided relevant background. The following list outlines the purpose of the remaining chapters. Other details and supporting derivations are in the appendices.

Chapter 2 In order to effectively use polymer gels in engineering applications, the "best" gel given performance requirements should be fabricated or selected. The process of gel preparation is discussed in this chapter, with emphasis on the effect of each fabrication step or gel constituent on various gel properties. Manufacturing methods for different shapes and sizes of the gel are explained. Different gel form factors will be useful for different applications. The efficiency of a polymer gel actuator will vary depending on how it is used

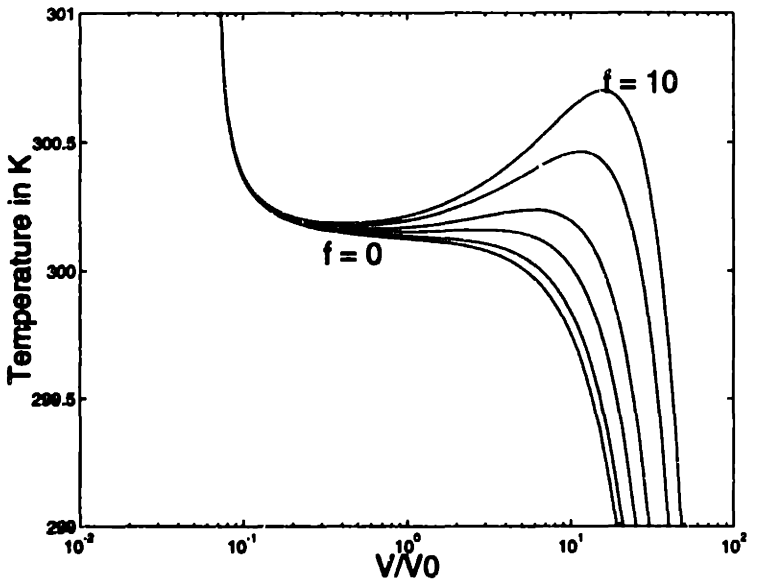


Figure 1.7: Plots of temperature versus volume based on the Flory-Huggins model

in a system. This is also discussed in Chapter 2.

Chapter 3 A gel with desirable properties may exhibit a discontinuous volume phase transition. One manner in which to obtain a continuum of force or position outputs from a discontinuous actuator is to employ switched control. Switched drives may be useful with other mechanical systems as well. In this chapter, different examples of switched systems are reviewed and the use of switching drives for a variety of actuators is motivated. A general model is suggested as a tool to explore the possibility of switching control and averaged model derivation for a large class of systems.

Chapter 4 Variations of the general model of Chapter 3 are examined in this chapter. The viability of switching drives for systems with different time-scale properties is addressed. Techniques that are primarily used in the modeling of switched power electronic ciruits are adapted to develop averaged models for a large class of switched actuators. Depending on the equilibrium properties and the underlying dynamics of the input, switch, and transducer

components of an actuator, different models are developed.

Chapter 5 An experimental setup was constructed to apply switching control to a polymer gel. By continuously switching the temperature of a NIPA gel, its volume was made to oscillate around an average point between its fully swollen and fully collapsed states. By using a switch period that is much shorter than the time constants associated with the gel volume-change, the volume ripple around the average point was minimized. Modulating the average temperature that the gel experiences during a switch cycle allowed the volume to be regulated to a dynamic equilibrium anywhere between its fully collapsed and fully swollen states.

Chapter 6 Using on the results of Chapter 4, an averaged model is developed for the prototype switched polymer gel actuator in Chapter 5. The structure of the model is based on the diffusion process that induces a volume phase transition and also the passive mechanical properties of the gel.

Chapter 7 Using temperature and gel length data recorded using the experimental setup described in Chapter 5, parameters for the model in Chapter 6 are identified. This chapter describes the identification algorithms and verifies the model under small-signal operation. In addition, the derived model is used for closed-loop control design to regulate the length of the gel and therefore the position of a suspended load.

Chapter 8 In this chapter, the application of polymer gels as actuators in remotely triggered devices is explored (for example, drug-delivery devices implanted under the skin). The development of magnetically activated gels for this purpose is discussed. Besides enabling remote activation, magnetic transfer of power is faster than the process used in the prototype in Chapter 5. This both improves the rate of response of the system and, as is shown in Chapter 4, makes the derivation of averaged models for switched systems easier. Experimental results of closed loop control on mechanically loaded magnetic gels will

demonstrate the effective optimization of performance through material design and the use the small-signal model for closed-loop control of different gels.

Chapter 9 Because of their ability to respond to a variety of environmental triggers, polymer gels can be used as sensors. Indeed, as mentioned earlier, sensing and actuating functions may be combined in a gel if it is fabricated appropriately. In Chapter 9, a gel sensor for metal ions is designed and implemented. The gel employed in this sensor has a transition temperature which is a function of metal ions in the gel solvent. A switching scheme is used to cycle the gel around this transition temperature which is then measured as an indicator for metal ions.

Chapter 10 The research work is summarized and the results of this thesis are reviewed. Future work based on these results is recommended and the potential of using polymer gels as sensors and actuators is evaluated.

Chapter 2

Fabrication and Properties of Polymer Gels

The chemical composition and physical shape of a gel influence its mechanical properties and dynamic response. For example, the chemical composition of the gel determines how it responds to different trigger mechanisms. By varying the quantities of different components, properties such as tensile strength and swelling ratio can be modified. Different shapes and sizes of polymer gels can be produced. Techniques have been developed to make cylindrical and spherical gels with the small dimensions required for short response times. Below, the basic formula for making a NIPA gel is introduced. The different components of NIPA gel and their influence are discussed. Methods of fabrication using variations of the basic formula to produce certain desirable shapes, sizes, and other gel properties are described. The tradeoffs involved in designing a polymer gel using these techniques for an engineering application are also discussed. Another important property of an actuator is its efficiency. By examining the thermodynamics of a gel work cycle some bounds can be placed on the efficiency of a polymer gel engine. These bounds and definitions of efficiency are also examined. Even though the discussion is in the context of NIPA gels, many of the chemical and physical priniciples apply to other gel compositions.

2.1 Chemistry

A polymer gel is a cross-linked network of long chains immersed in a solvent. In free-radical polymerization, the method by which a NIPA gel is commonly produced, an *initiator* compound is used to start the chemical reaction that causes monomers to combine into chains. This reaction can be aided by an *accelerator* compound or by heating the mixture, for instance. A *cross-linker* entangles or connects the chains together into a network. The network can then diffuse into and out of a solvent depending on different environmental parameters[1].

Other details of the chemical processes can vary depending on the particular gel. For instance, gel copolymers are made of more than one type of monomer. Including ionic compounds in the polymer network can result in different equilibrium and dynamic behavior [98] [17] [59]. To illustrate the effect of the main components of a polymer gel on various mechanical and dynamic properties, a standard NIPA gel formula is described below. Variations on the formula and fabrication technique are used to make gels with different characteristics and shapes.

2.1.1 Basic NIPA Formula

Pure NIPA gels are prepared by a free radical polymerization in water at room temperature. Tetramethyl ethylene diamine, or TEMED (the accelerator), is added to a solution of bisacrylamide, or BIS (the cross-linker), and the main constituent NIPA monomer. Ammonium persulfate, or APS (the initiator), starts the polymerization and gelation processes. The first step in polymerization is a reaction between APS and TEMED in which the TEMED molecule is left with an unpaired valence electron. The activated TEMED molecule can combine with a NIPA or BIS monomer, transferring the unpaired electron to the monomer unit. Another monomer can then be attached and activated in the same manner. Through this process, long polymer chains are formed as one monomer connects to another. The inclusion of BIS molecules turns these long chains into a web of inter-

[1]The cross-linking can take place simultaneously with the polymerization process; in other cases the polymerization is completed before the cross-linking begins.

connected chains. The BIS molecule has two functionalized sites and can be incorporated into two chains simultaneously. Figure 2.1 illustrates the chemistry of this process further. It is important to note that exposure to oxygen inhibits the polymerization process. The monomer solution is degassed before the addition of APS. Unless the container for the *pre-gel* solution is well sealed, a nitrogen atmosphere may also be necessary for the reaction to succeed. References [97], [82], [81], [95], [31] describe the chemical process in more detail.

2.1.2 Effect of Chemistry on Mechanical Properties

The behavior of a polymer gel is a result of its composition. For example, characteristics such as swelling ratio and tensile strength of the actuator are influenced by the concentration of cross-linker in the gel. These characteristics are important parameters in a servomechanical application. Every ingredient in the gel formula affects the gel properties in some manner. These effects are discussed below.

- The *base monomer*, in our case NIPA, is the main ingredient in the gel. (There are other possibilities, including copolymers, block copolymers, and interpenetrating polymer networks [17].) The monomer units are the links in the chains that form the gel network. The solubility of this monomer in the gel solvent under different environmental conditions determines in part what will cause a volume transition and to what degree.

- The *initiator* (APS) and the *accelerator* (TEMED) react to begin the formation of a polymer chain. By controlling the number of initiator molecules, we can control the number of polymer chains which begin to form during polymerization. For a fixed amount of monomer, including less initiator generally leads to the formation of longer chains. The longer polymer chains result in a less brittle gel, usually a desirable quality in a gel to be used in a servomechanical application. On the other hand, too little initiator may result in incomplete polymerization of the monomer. This is especially likely when it is difficult to control the environment and prevent oxygen from quenching the initiator, for instance. Increasing the amount of the accelerator

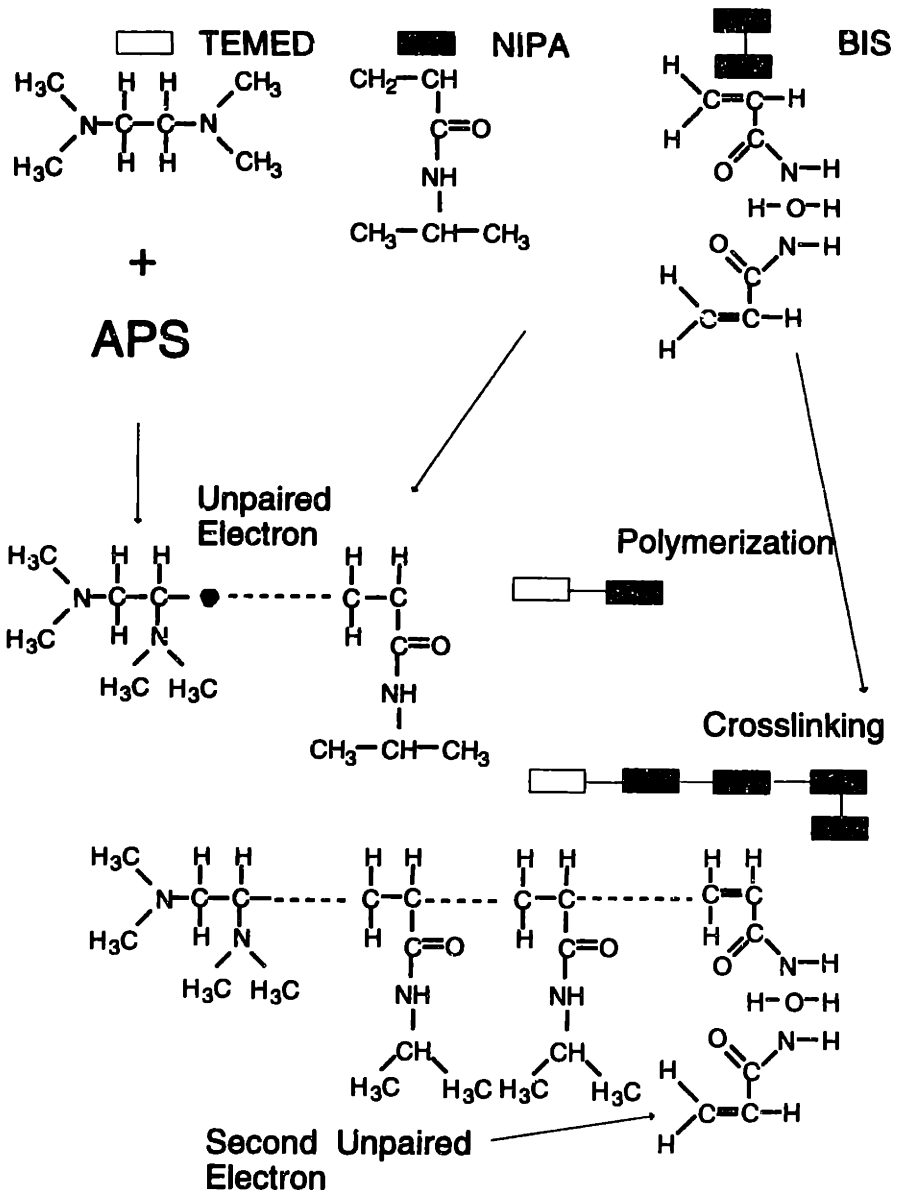


Figure 2.1: Chemical process of polymerization and gelation

and/or the initiator will also increase the rate of polymerization (and gelation in this case). This may also be an important parameter to adjust based on the process of fabricating the gel.

- The *cross-linker* (e.g. BIS) connects the different polymer chains into a tangled network. The higher the amount of cross-linker in the reaction, the more interconnections and the more intertwined the network. A denser network is stronger. However, it exhibits a smaller swelling ratio than would a more lightly cross-linked network. The cross-linker concentration in the gel formula can be adjusted to tradeoff strength and stroke for a particular application.

- Adding *ionic groups* to the monomer mix can result in a gel co-polymer with different mechanical properties than those of the pure NIPA gel. By adding ionic groups into the polymer chains, additional osmotic pressure can result in larger swelling ratios, different transition temperatures, and discontinuous equilibrium phase transition curves [31]. In this thesis, sodium acrylate or acrylic acid will be used to increase the gel stroke when necessary. A larger amount of these compounds produces larger swelling ratios. The transition temperature of the gel also increases with the concentration. A very high transition temperature may be a problem in some practical systems.

- *Inter-penetrating Polymer Networks* (IPN's) are formed by "interweaving" two polymer gel networks to give the gel different attributes. Preliminary results of our work show that incorporating poly-(vinyl alcohol) (PVA) polymer with the NIPA formula strengthens the resulting gel significantly, without considerably modifying the volume-phase transition characteristics. The PVA network appears to be cross-linked by a cooperative hydrogen interaction between the PVA molecules.

The polymer gel can be altered in other ways. For instance, the solvent itself can be changed. In [79], the salt content of the aqueous solvent is shown to have an effect on the volume-phase transition curve of a polymer gel. Heating the NIPA solution during gelation

can result in a micro-dispersed gel with an inhomogeneous network density. The result is a weaker gel with a faster rate of response [46]. More advanced techniques for synthesizing different polymer gels are discussed in [18], [101], [79], [36], [17], [111].

2.2 Form Factor

Different application requirements may dictate different shapes and specifications for the polymer gel material. Using the basic formula outlined above, different forms of NIPA polymer gel can be manufactured. As discussed in Section 1.2.1, in order to improve the response rate of a gel, the critical dimension has to be made very small. For the diffusion rates of NIPA gels, the dimensions necessary to achieve sub-second performance are on the order of tens of microns. A single gel at that size is too weak for most applications. One solution is to group many of the micron-sized gels in a manner that allows mechanical load-sharing while preserving the speed of response [2].

Depending on the application, different form factors or shapes of the gel units may be necessary. For a "synthetic muscle", a cylindrical shape or fiber is appropriate. A cylinder with a much larger length dimension than its radius will have a quick response time if the radius is small, and will deliver a significant stroke in the axial direction, ideal for a linear actuator. The cylindrical shape is also regular enough to simplify the gel modeling equations. Another useful and interesting shape is that of a sphere. Like a cylinder, a sphere's symmetry makes it easier to identify the critical dimension and to analyze the physics of the volume transition. Gel spheres or beads may be used to construct a "piston" that will push out when spheres packed inside a chamber are made to swell. Irregular shapes of gels will exhibit volume-phase transitions as well. Indeed, the ability to manufacture active gels in any shape or size makes them useful for many actuator or sensor applications. The fabrication of gel cylinders, fibers, and spheres with short critical lengths is discussed below. Variations on the NIPA formula from the previous section are used for each process. The symmetry of these shapes makes it easier to observe and analyze a volume change.

[2]The load-sharing approach also has the added advantages of any cellular system. For example, failure of one gel unit does not necessarily result in failure of the actuator.

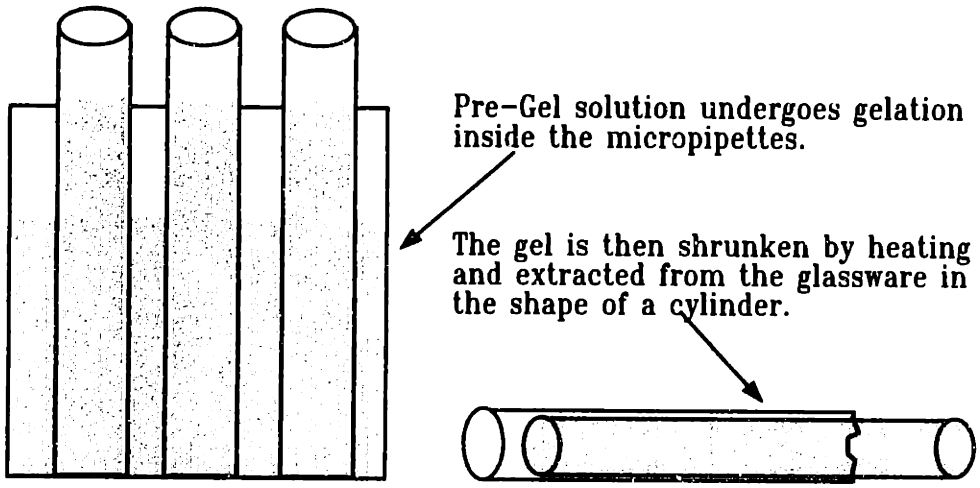


Figure 2.2: Making gel cylinders

2.2.1 Gel Cylinders

Thin cylinderical gels can be constructed by allowing the gel to form inside micropipettes. The gel is extracted from the micropipettes by immersion into warm water (with temperature above the transition point) and gently breaking the micropipette. The glass is broken but the more flexible gel is not. The cylinder of NIPA, shrunken due to the temperature of the surrounding water, is easily released from the micropipette. Figure 2.2 illustrates this procedure. The steps for fabricating a NIPA/PVA gel cylinder are:

1. Dissolve 115,000 molecular weight PVA (5 g) in 500mL of deionized water at 50°C for approximately 20 hours.

2. Dissolve APS (.2 g) in 5mL deionized water at room temperature. Degas the solution in a vacuum chamber.

3. Dissolve NIPA monomer (.78 g), and BIS (.013 g) in 5 ml of deionized water at room temperature. Mix this solution with 5 mL of the PVA solution from step 1, then degas the resulting 10 mL solution in a vacuum chamber. In this step, some of the NIPA monomer can be replaced with an equal amount of sodium acrylate. The result is a co-polymer with different volume transition properties. In our experiments, replacing

.016g of NIPA with sodium acrylate raised the transition temperature to around 38°C, compared to approximately 38°C for pure NIPA.

4. Mix TEMED (24 μL) and 40 μL of the APS solution from step 2 with the solution from step 3. Shake the mixture to insure good mixing. Polymerization and gelation begin with this step.

5. Place the cylindrical micropipettes of different diameters in the resulting solution immediately after step 4. Typical diameters range from 0.5 mm to 1 mm.

This formula is similar to the one used in [31]. The addition of the PVA to form an IPN and the lower level of APS to produce longer polymer chains were used to increase the tensile strength of the gel. This process of fabricating gel cylinders for eventual bundling to form an actuator is time-consuming. Also, the response times for the resulting gels are generally relatively long. The smallest gel that can be easily pulled out of its micropipette by hand is approximately 750μm in diameter (50μL micropipette), and its transition time is on the order of tens of seconds. However, these gels can be a good source of information about tensile strength and stroke length, as they are easy to construct and their properties are easily made homogeneous during fabrication. In particular, the degree of cross-linking and its spatial distribution, the size, and the shape of the final gel are all well controlled. The symmetric, homogeneous, long cylinders are useful in experiments where the quality and reproducibility of the data is important. They are used in this thesis to identify a mathematical model for the dynamics of gel under tension. Figure 2.3 illustrates the collapsed and expanded states of a gel cylinder with 0.16 g of sodium acrylate.

2.2.2 Gel Spheres

The same basic formula used in producing NIPA gel cylinders was used for making microscopic gel beads [3]. Initially, gel beads were made by filling tiny dimples machined into teflon blocks with pre-gel solutions. As the gel formed, it was to take the shape of the hole in which it lay. The gel beads formed using this method were not spherical and were

[3] I would like to thank Graham Fernandez for his help in starting this work.

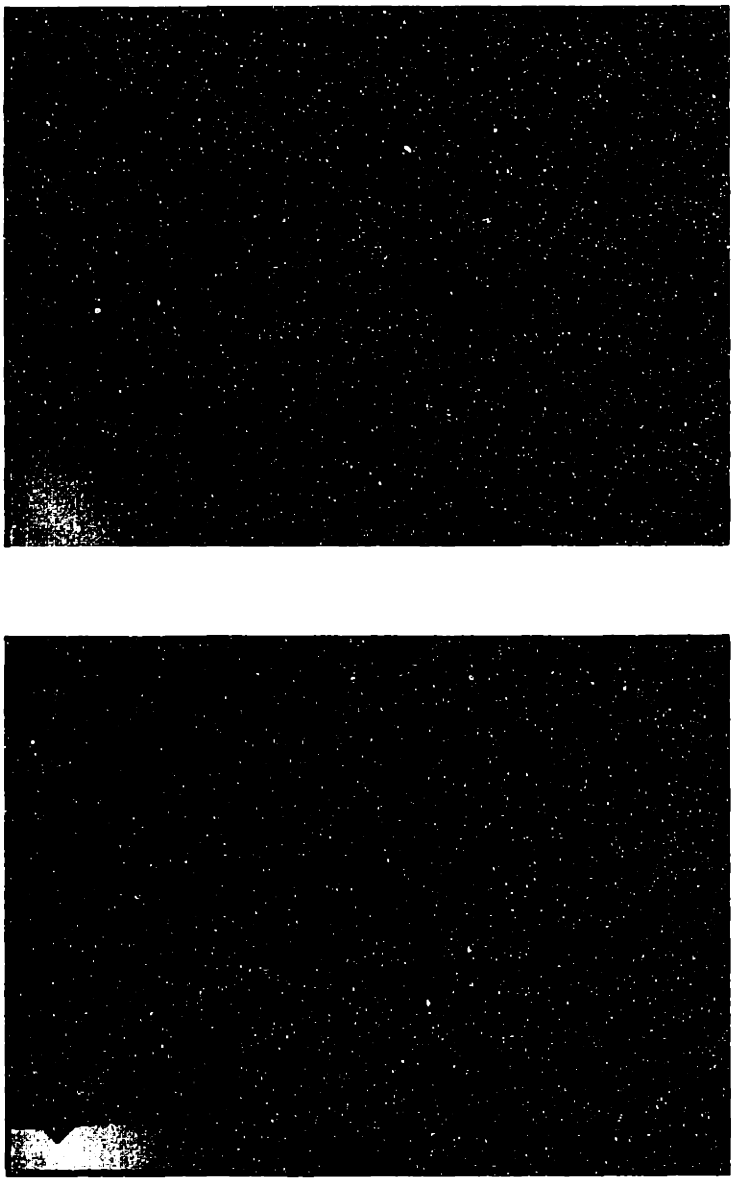


Figure 2.3: Volume transition of a gel cylinder as seen under the microscope

often damaged while being extracted from their mold. This procedure also makes mass production difficult.

Another method that is more commonly used in gel bead preparation is inverse suspension polymerization [69]. The gel constituents are dissolved in deionized water as they were previously. This solution is then injected into paraffin oil that is stirred with a magnetic stirrer or some other mechanism. Care is taken to saturate the paraffin oil and the pre-gel solution with nitrogen to ensure the absence of oxygen, which quenches the free-radicals generated by the initiator (APS). Figure 2.4 illustrates the method of inverse suspension polymerization. As the paraffin oil is stirred, the water is broken up into droplets that form into gel spheres. After the formation of the sphere, the mixture is removed and poured into a beaker containing deionized water. As the paraffin is agitated, the gel beads fall into the water layer (which is separated from the oil layer). In this manner, the beads are extracted from the paraffin. For a detailed description of this procedure, see [19]. Spherical gels with diameters in the hundreds of microns range were fabricated using the bead formation procedure outlined below.

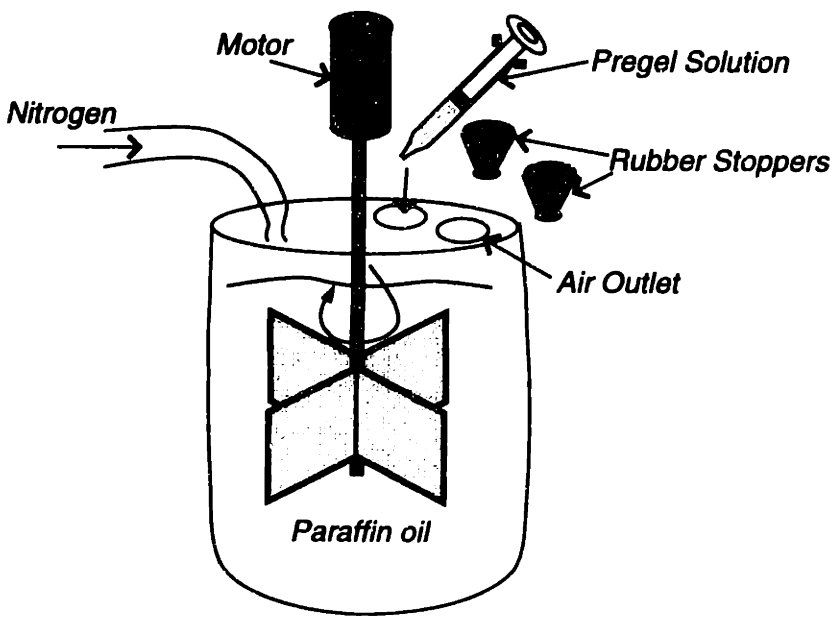


Figure 2.4: Gel bead formation by inverse suspension polymerization

1. Dissolve APS (.2 g) in 5mL deionized water at room temperature. Degas the solution in a vacuum chamber.

2. Dissolve NIPA monomer (.78 g), and BIS (.013 g) in 10 mL of deionized water at room temperature, then degas the resulting solution in a vacuum chamber. As with the cylinderical gels, some of the NIPA monomer can be replaced with an equal amount of sodium acrylate.

3. Pour parrafin oil into the container to immerse the paddle as shown in the figure.

4. Pump nitrogen into the oil as shown in the figure at 10 psi for approximately 30 minutes. At the same time, stir the oil to insure that the nitrogen flows into the paraffin oil and replaces the oxygen, which flows out of the open air outlet.

5. Turn off the nitrogen supply and seal the air flow outlet. A slightly positive nitrogen flow can also be maintained if a perfect seal of the jar is not possible to insure that the oil mixture remains oxygen free.

6. Mix TEMED (24 μL) and 100 μL of the APS solution from step 1 with the solution from step 2. Shake the mixture to insure good mixing and load into a syringe. Polymerization and gelation begin with this step.

7. With the paddle continuing to stir, inject the polymer solution into the oil. Allow to stir for 2 hours. The shaft velocity can be used to control the size of the formed gel beads. In our experiments, shaft speeds close to 300 rpm produced beads with diameters between 100 μm and 300 μm.

8. Stop the stirring and allow the water and formed gel beads to settle to the bottom of the container. The oil is then drained and the beads rinsed of any excess oil.

The protocol presented above can also be used with the addition of PVA as for the gel cylinders. However, the advantage of added tensile strength may not be significant enough to offset reduced swelling ratio in this case. Tensile strength is more significant in cylinders

and fibers. Note also that the amount of APS used is larger than before. The additional initiator speeds up the gelation process. The longer this process takes, the greater the possibility of oxygen permeating into the apparatus. Moreover, if the gelation is slow, some of the constituents may settle out of the pregel solution. Figure 2.5 shows a bead formed by this method in both its swollen and collapsed state. The bead shown is a gel co-polymer with 0.16 g of sodium acrylate substituted for an equal amount of NIPA in the formula above.

2.2.3 Gel Fibers

In order to achieve the fast performance necessary for many applications, gels that are only a few microns thick might need to be constructed. One way to create these fibers is by "pulling" or "spinning" a pre-gel polymer solution until it is the right thickness. In addition to providing very small volume gels, this method could potentially help to microscopically align the polymer chains in a network, which might yield higher performance in terms of stroke length and strength in an actuator.

A thick, viscous fluid can be spun or pulled to make thin fibers. A cross-linked gel cannot be spun. It has a specific equilibrium shape that is determined by its permanent cross-links. To make fibers, a solution of polymer chains that are not cross-linked is used. A formula similar to that of the gel cylinders or spheres can be used without the cross-linker to produce a NIPA polymer solution. The initiator and accelerator facilitate the polymerization process and produce polymer chains that are not cross-linked. The viscosity of this solution is a result of polymer chain length. We can therefore adjust the viscosity level by controlling the amount of APS in the pregel solution. After a polymer solution with an appropriate viscosity is prepared, diazido-stilbene (DAS) is dissolved in this pre-gel solution. DAS is an optical cross-linker which is activated through exposure to ultra-violet (UV) light. It is used to cross-link the polymer chains into a gel network after they have been spun into fibers.

After the polymer solution is prepared, it is loaded into a syringe and extruded through a needle onto spinning parallel bars as shown in Figure 2.6. The bars catch the viscous

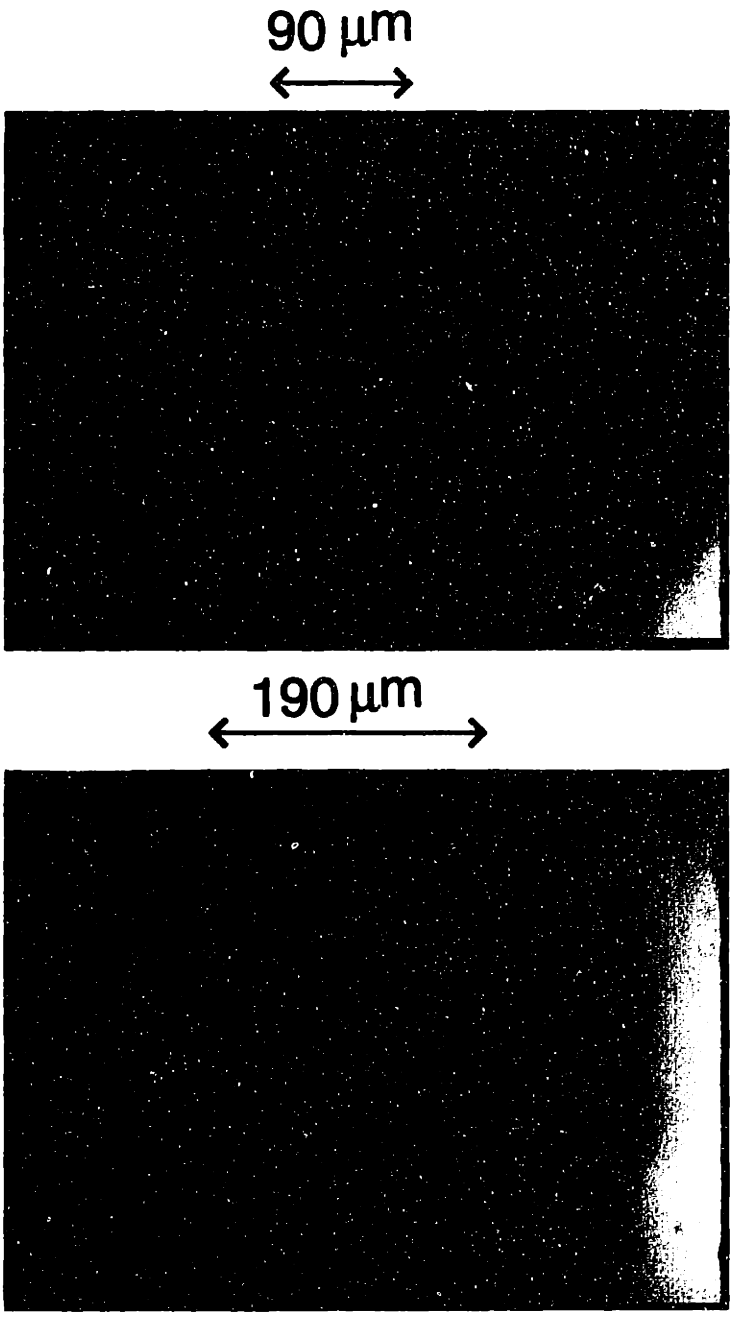


Figure 2.5: Volume transition of a gel sphere as seen under the microscope

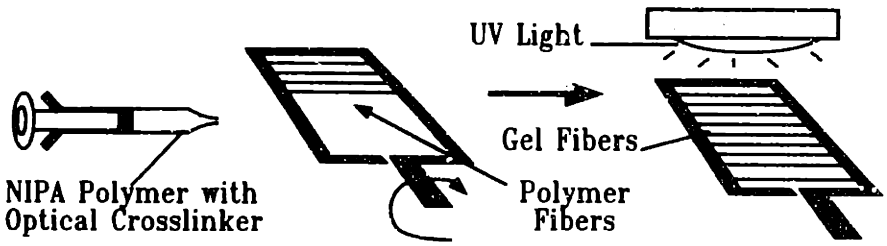


Figure 2.6: Spinning gel fibers

solution and pull it into a thin fiber. If the syringe is moved laterally (parallel to the bars) while this process takes place, many fibers will form between the bars. The thickness of these fibers can be controlled by the angular velocity of the bars and the extrusion rate. Once the fibers are formed, they are exposed to UV light to crosslink the polymer chains into a gel. The entire process is carried out in a nitrogen atmosphere. The steps for making gel fibers are:

1. Dissolve PVA (5 g) in 500mL of deionized water at 50°C for approximately 20 hours. The PVA used had molecular weight 115,000.

2. Dissolve APS (.2 g) in 5mL deionized water at room temperature.

3. Dissolve NIPA monomer (.78 g), in 7 mL of deionized water at room temperature. Mix this solution with 3 mL of the PVA solution from step 1 and 40 μL of the APS solution in step 2. Place the resulting solution in a vacuum chamber for approximately 1 hour. This will remove the oxygen which may inhibit polymerization. Allow 9 more hours for polymerization to complete.

4. Add DAS (.015 g) to the viscous polymer solution and vigorously shake on a Vortex Genie^{TM}. After approximately 4 minutes of continuous shaking, the DAS should be dissolved. Since DAS reacts in response to light, an opaque container should be used from this step onward. (Alternately, the glass vial can be wrapped in aluminum foil.)

5. The vigorous shaking of the viscous solution will trap air bubbles in the polymer/DAS

solution. Any trapped oxygen will hinder the crosslinking process. In addition, the bubbles will make it difficult to spin smooth fibers. The thick polymer solution should be degassed in a vacuum chamber again. This will take a longer time because of the viscosity of the solution. We have witnessed that 10 hours with the given formula was sufficient for releasing trapped air from the solution.

6. The degassed solution can be loaded into a syringe and spun as shown above. This should be done in an oxygen-free environment.

7. Place the polymer fibers under a UV light source for crosslinking. A reflective surface should be placed around the fibers to produce spatially homogeneous crosslinking. In our experiment, the UV source [4] was applied for 1 hour to complete cross-linking. Fibers with diameters greater than 100 μm were not successfully crosslinked with this UV source. A higher power UV source is required for penetrating thicker polymer fibers.

The cross-linked fibers are then grouped into a bundle by placing their ends in epoxy on aluminum sheets. The whole process is currently performed by hand, and assembling bundles with more than 40 or so fibers is very time-consuming. Moreover, for all the fibers to participate equally in an actuator, they need to be pulled to the same length at the same temperature, and they need to respond with similar transition rates. This requirement translates into nearly identical dimensions for all the fibers. To help produce bundles of a large number of homogeneous fibers, a machine was built to automate this process [78]. A gel fiber formed using the process described here is shown in Figure 2.7. The tiny spheres adjacent to the gel fiber are air bubbles in the solvent.

2.3 Thermodynamics and Efficiency

Polymer gel actuators will convert some form of energy to mechanical work. In the case of NIPA, for instance, the input energy is in the form of heat. An understanding of the energy

[4] The UV source used was a Spectroline PE-240T EPROM erasing UV lamp manufactured by Spectronics Corp., Westbury, NY.

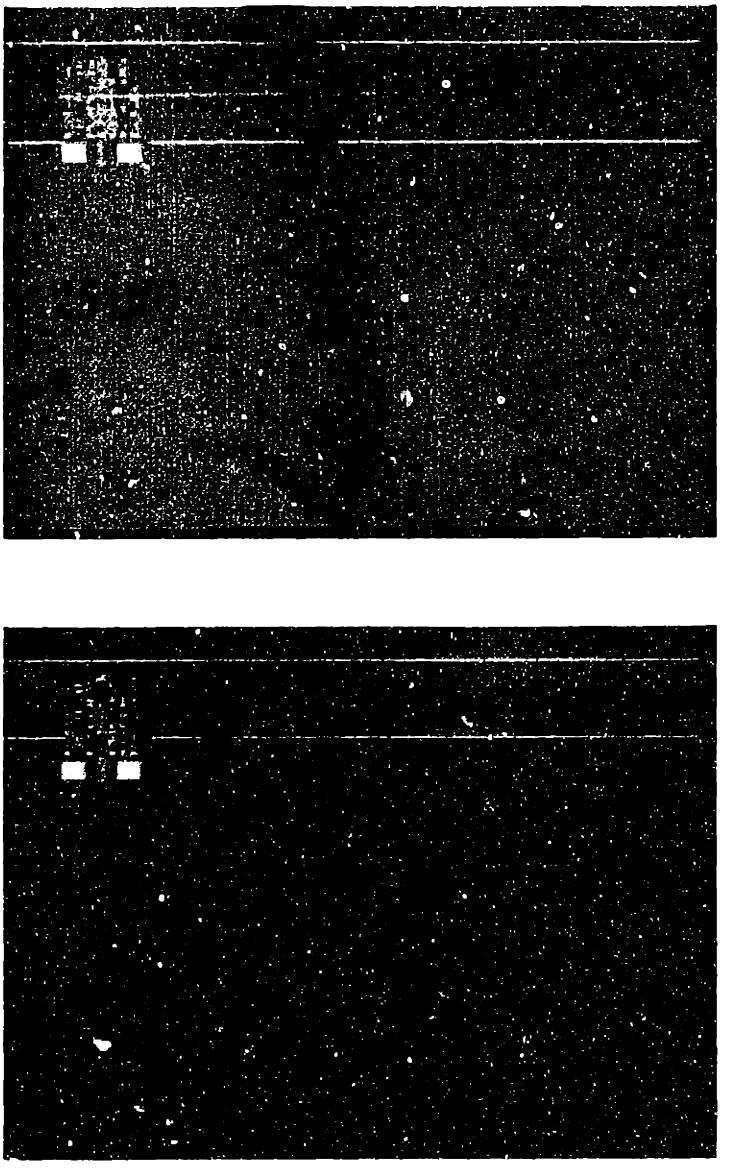


Figure 2.7: Volume transition of a gel fiber as seen under the microscope

flow in the gel actuator will help identify the input and output energy components of the system and determine its energy conversion efficiency. Free energy exchange in the network-solvent system is the basis for most of the theories describing phase transitions in polymer gels [17]. However, many of these theories, most notably the Flory-Huggins model, describe free-swelling gel systems and do not explore mechanical energy conversion or efficiency. Recent interest in polymer gel actuators has stimulated research on the mechanical work cycle of a volume-phase transition. In [33], it is shown that the efficiency of a gel engine increases with load. The gel in this study is triggered chemically and the difference in chemical potentials of the solvent is used to compute the input energy. No account for heat flow is made, and only the swelling transition is examined (pushing against external inward pressure). Solari in [90] uses a similar technique to demonstrate a change of efficiency with load, but provides no quantitative figures. The sections below discuss energy flow during a gel volume-phase transition and identify energy quantities that can be combined to examine the energy conversion properties of different polymer gel actuators.

2.3.1 The Ideal Gas Engine

As a starting point for the analysis of the gel thermodynamic cycle, consider the ideal gas engine. For a given mass of ideal gas, the product of the pressure and volume of the gas is constant at a constant temperature. We can use the expansion of the gas with its positive pressure to convert heat to useful work. For comparison and discussion, Figure 2.8 illustrates a mechanical work cycle, the Carnot cycle, for a gas engine. Two graphs are shown. A T-V (temperature versus volume) plot shows the equilibrium temperature as a function of the volume of the gas, and a P-V (pressure versus volume) plot shows the pressure/volume states of the ideal gas in a quasi-static equilibrium. The areas under the curves in the P-V plot indicate the work done on each part of the cycle [14].

In the following discussion, changes in volume occur slowly enough that the system is always in equilibrium, thus making the processes reversible by preventing dissipation. On the first part of the cycle, work is done on the gas engine to reduce the volume of the gas. This process is adiabatic, with no heat exchange across the boundaries of the system.

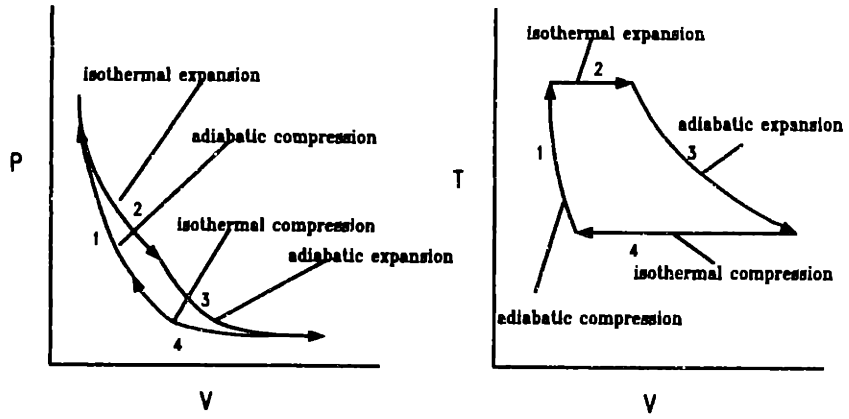


Figure 2.8: Work cycle for a gas engine

As a result, the temperature increases and so does the energy content of the system (by an amount equal to that of the work input into the system to reduce its volume). The curves labeled 2 in Figure 2.8 represent an isothermal process in which the temperature of the system is held constant by allowing it to communicate with a temperature reservoir at the temperature T_H. The gas expands, doing work on the environment as it moves the boundary of the system. The work output is equal to the heat that flows into the system, $W_H = Q_H$. The third component of the work cycle is an adiabatic expansion. The temperature drops to T_L and the energy content of the system is reduced as it does more work on the environment. On the fourth and final step of the closed cycle, the system is in communication with a temperature reservoir at T_L and contracts isothermally. The work input here is $W_L = Q_L$. It turns out that the work done during the positive and negative adiabatic cycles is equal in magnitude and opposite in sign. If the input into the system is defined to be Q_H, with T_L representing the temperature of the atmosphere for instance, then the efficiency is typically defined as:

$$\eta = \frac{W_H + W_L}{Q_H} = 1 - \frac{T_L}{T_H} \tag{2.1}$$

where W_L is a negative quantity. For more on this formulation, refer to [14].

2.3.2 NIPA Gel Volume-Phase Transition Cycle

Comparison to Ideal Gas Engine Cycle

The efficiency of the reversible cycle of a gas engine provides an upper bound on any heat engine operating between two thermal reservoirs [14]. It is therefore a good starting point to investigate the thermodynamics and efficiency of a thermosensitive polymer gel. For example, a NIPA gel can be made to go through a volume-phase transition cycle using a hot and a cold thermal reservoir to shrink and expand the network, respectively. Practical choices for the two temperatures, as will be seen in Chapter 5, are 44°C and 24°C, for example. Using the bound of Equation 2.1, the closed cycle efficiency is less than 10% (using absolute temperature values). For applications requiring high efficiency, therefore, gels that exhibit volume change in response to some other trigger may be more appropriate. While the ideal gas engine provides a good start for analyzing the gel cycle, we should note the following distinctions :

- The gel can do work on the environment on both parts of its cycle of operation, expansion and contraction. The sign of the gel pressure can follow that of the volume change, producing positive work flow during both transitions. The gel can push and pull!

- The adiabatic/isothermal work cycle used in the analysis of the gas engine facilitates the derivation of the efficiency limit. However, this is with the assumption that useful work is only done on one part of the cycle and work is negative as the gas volume decreases. Depending on how the gel is used and its configuration within a system, different efficiency definitions might be more appropriate.

- In the gas engine, there is no change of phase. Energy flow results in either mechanical work or a change in temperature. In the gel system, the change of phase involves the transfer of latent energy quantities that can be stored in the gel.

Despite these differences, an analysis similar to that of the gas engine can be developed to detail the work and heat components involved in different parts of a polymer gel volume-phase transition cycle. These quantities will be useful in computing efficiency bounds and efficiency-like metrics once specific parameters of the gel actuator are set.

Free-Swelling Gel Transition

NIPA and NIPA/PVA polymer gels are thermally activated. At equilibrium, the internal pressure of the gel is equal in magnitude and opposite in direction to the external pressure on the gel. A change in temperature disrupts this balance and the gel moves towards a new volume that will equate the pressures and restore the system to equilbrium. The cartoon in Figure 2.9 qualitatively illustrates the static relationships between the pressure and volume and the temperature and volume for a NIPA gel. Each of the traces in the T-V curve is an isobar with constant osmotic pressure acting on the gel network. For a free-swelling gel, the equilibrium volumes will all be at zero pressure, or, equivalently, on the curve Π_0 ($\Pi_0 = 0$ in this case). Consider a NIPA gel at equilibrium point 1 on the T-V curve. The gel is in the expanded phase at temperature T_1. The temperature is then raised to temperature T_2, moving the pressure of the gel system onto the curve labelled Π_{-1}. The change in temperature occurs much more rapidly than the change in volume. At point 2, the gel is at the same volume but is now experiencing a negative pressure due to the increase in temperature. The part of the cycle in which the gel moves from point 1 to 2 is approximately isochoric since the temperature change is much faster than a volume change [60]. In order to restore the balance of forces that produces zero pressure, the gel network begins to collapse. Keeping the temperature constant, the volume decreases along a horizontal line in the T-V plot until equilibrium is reached at the Π_0 curve ($\Pi_0 = 0$ for equilibrium). At the new equilibrium point 3, the gel is at a smaller volume and a higher temperature T_2. If the temperature is now dropped to T_1 again, the internal pressure increases to Π_2 at point 4 and the volume increases until the total pressure [5] drops back to 0.

[5]The total pressure in the case of a free-swelling non-ionic gel here is the same as Π_{gel}. As is shown in the next section, ionization or mechanical load adds another pressure component to the total osmotic pressure.

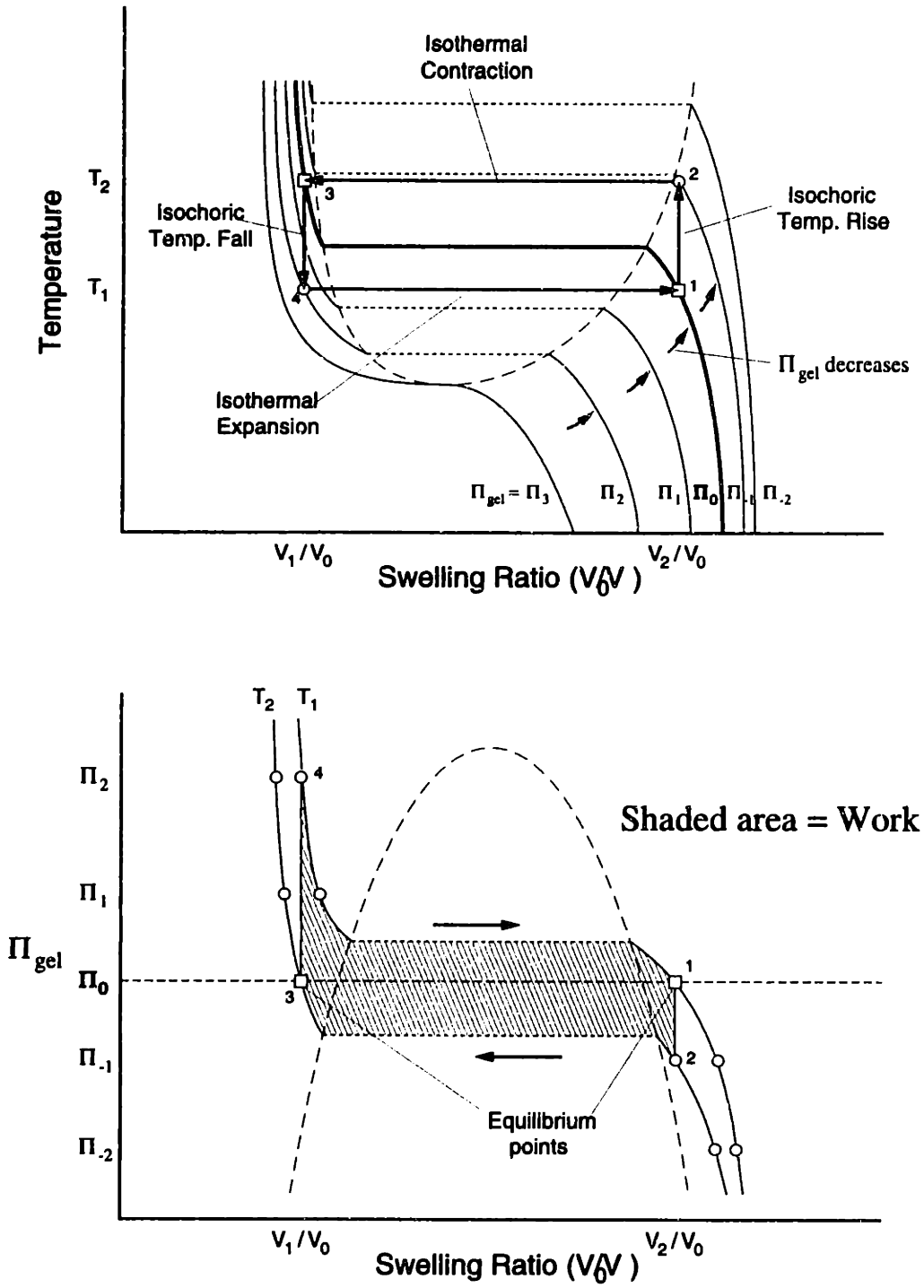


Figure 2.9: Volume transition cycle for a NIPA gel

As shown in the figure, the static Π-V curve can be obtained from the T-V curve. Each of the curves in the figure is an isotherm obtained by plotting points along a horizontal line in the T-V plot. The transition from one equilibrium point to another can be thought of as two consecutive processes. Initially, the temperature of the gel system jumps along an isochoric curve ($1-2$, and later $3-4$). The volume of the gel system is assumed constant during this transition. An isothermal expansion or contraction follows until the new equilibrium point is reached ($2-3$ or $4-1$). By analyzing the energy transfer during these transitions, an understanding of the work cycle and efficiency of a gel engine can be obtained. Note that the static plots have no information on the dynamics of the volume transition. Nevertheless, the area under the Π curves represents the work done in the different cycles of volume change and can be used to derive bounds on polymer gel efficiency.

Ionized and Loaded Gels

Ionic groups in a gel provide an additional source of internal pressure Π_{ion}. The condition for equilibrium is then:

$$\Pi_{tot} = \Pi_{ion} + \Pi_{gel} = 0 \tag{2.2}$$

where Π_{gel} is the gel pressure in the absence of any ionic forces. Ionic pressure generated by the inclusion of sodium acrylate in the NIPA gel is positive, favoring the expansion of the gel. In equilibrium, the total gel pressure Π_{tot} is zero by definition. As a result, Π_{gel} is negative, pulling the gel inward to offset the positive ionic pressure. The same analysis of a volume transition will occur around a negative isobar in the T-V curve (such as Π_{-1} or Π_{-2}). To maintain the same Π_{gel} and volume in the presence of ionic forces, a higher temperature is required. On the P-V curve, this translates to an upward shift of the isotherms. Figure 2.10 shows experimentally determined phase transition curves from [31]. The different curves in the top graph represent equilibrium points for NIPA gels with different concentrations of sodium acrylate. Although the total pressure $\Pi_{tot} = 0$ for all equilibrium points, the different plots can be thought of as isobars at different levels of $\Pi_{gel} = -\Pi_{ion}$, where Π_{ion} depends on the concentration of sodium acrylate. Figure 2.10 also illustrates the derivation

of the pressure-volume curves from the phase transition curves similar to Figure 2.9. The transition cycle shown is for a NIPA gel with a 50 mM concentration of sodium acrylate. In equilibrium, for this gel, $\Pi_{tot} = 0$ and $\Pi_{gel} = \Pi_{-3}$.

Effects of the ionic forces on the gel behavior are similar to mechanical effects. An external load that stretches the gel applies a positive pressure similar to Π_{ion}. In most cases, the pressure due to mechanical loading will not be uniform in all dimensions, unlike the pressure due to ionic forces. However, the behavior of a gel system subjected to ionic forces should approximate its behavior under mechanical load. For example, translation of isotherms in the P-V curves is consistent with the finding in [90] and [33] that the efficiency of a polymer gel actuator changes with load.

2.3.3 Energy Exchange in a Thermo-Sensitive Gel

In the ideal gas engine, the efficiency is calculated for a complete cycle. After every cycle, the piston is reset to its initial position and the heat transfer between temperature reservoirs is used to generate useful work. The total useful work output is divided by the heat input to compute the efficiency. Similar analysis can be carried out for gel engines. Specifics of the implementation of a polymer gel actuator need to be considered before a definition for the efficiency can be derived. For instance, the gel can do work in both the positive and negative directions (expanding and contracting), and depending on which part of the cycle is chosen to do the useful work, a different, non-conventional expression for the efficiency may be appropriate. Below, the energy transfer in the isochoric and isothermal processes during expansion and contraction is analyzed. Using expressions for energy quantities exchanged during these transitions, efficiency definitions for different kinds of gel engines are derived.

Isochoric Heating From equilibrium point 1 to point 2, the gel-solvent system is heated but the volume has no time to change. The flow of heat is

$$Q_{ICH} = m_s \int_{T_L}^{T_H} c_s(T)dT + m_n \int_{T_L}^{T_H} c_n(T)dT > 0 \tag{2.3}$$

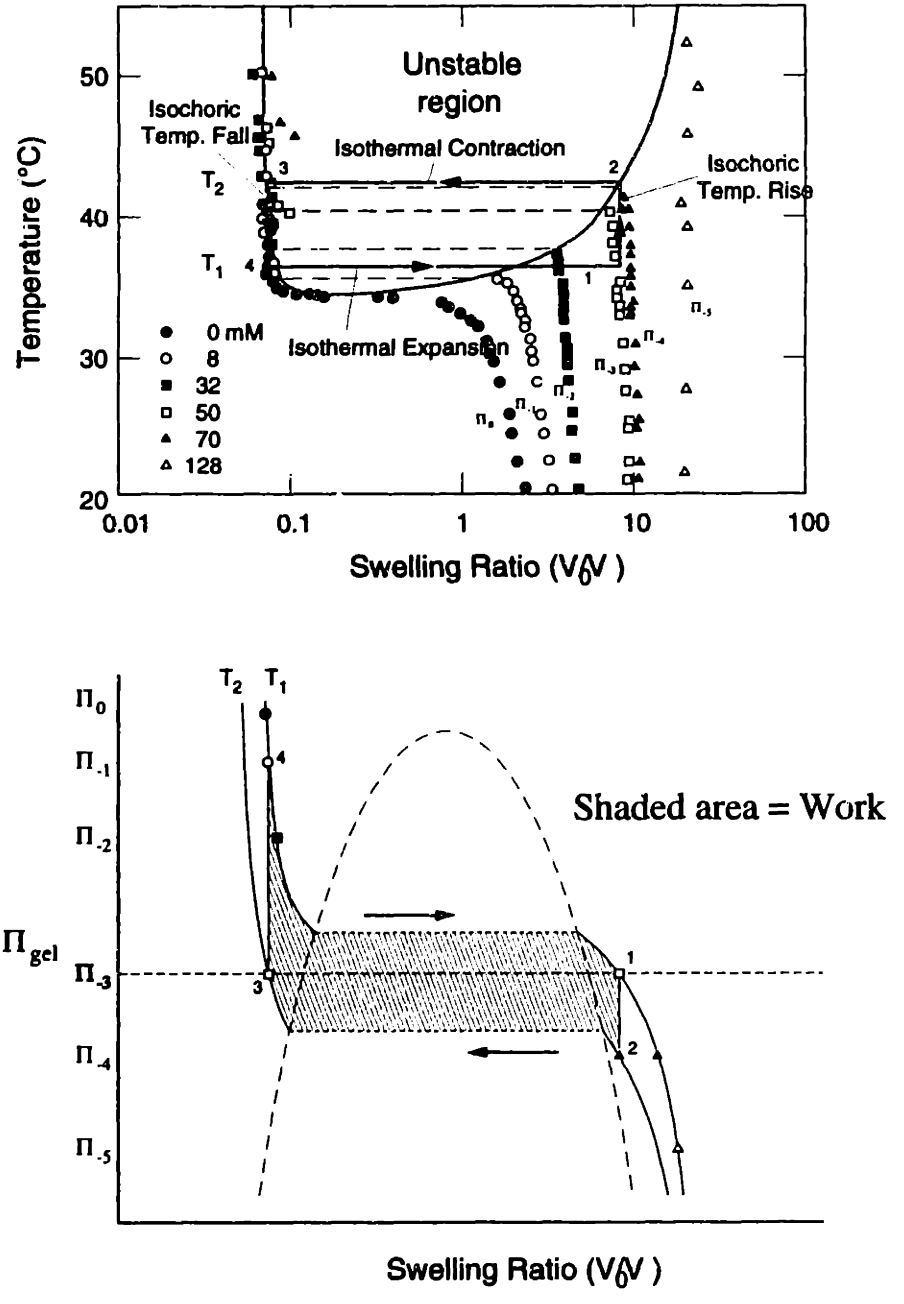


Figure 2.10: Volume transition cycle for a NIPA/sodium acrylate gel

where the variables m_s and c_s are the mass and specific heat of the solvent, and m_n and c_n are the mass and specific heat of the polymer component of the gel. We shall denote heat flowing into a system as positive. No work is done on this part of the cycle.

Isothermal Contraction This part of the volume transition takes place from point 2 to point 3 in Figures 2.9 and 2.10. The temperature is assumed to have stabilized at T_H and the gel undergoes this volume transition partly in a continuous regime and possibly partly through a phase transition. The total heat transfer in this part of the cycle, Q_{ITC}, is the latent heat transfer involved in the state transition as well as the heat causing the volume change :

$$Q_{ITC} = Q_{HL} + Q_{HW}, \tag{2.4}$$

where Q_{HL} is the latent heat of changing the state, and Q_{HW} is the energy that is converted to work through the volume change. The work done during this part of the cycle is :

$$W_{ITC} = \int_{V_2}^{V_1} \Pi(V) dV, \tag{2.5}$$

where the volume limits are the two equilibrium volumes of the gel at the high and low temperatures. The net heat transferred Q_{ITC} is positive (into the system) and for a practical system, this implies an increase in the disorder (entropy) in the system. A volume decrease is usually associated with an increased order in the system. For the gel system, however, the network becomes more ordered, but the solvent (water in this case) becomes less ordered, losing its structure and causing the decrease in volume.

Isochoric Cooling From equilibrium point 3 to point 4 in Figures 2.9 and 2.10, the gel-solvent system is cooled but the volume has no time to change. The flow of heat is :

$$Q_{ICC} = m_s \int_{T_H}^{T_L} c_s(T) dT + m_n \int_{T_H}^{T_L} c_n(T) dT < 0 \tag{2.6}$$

The negative heat flow indicates a flow out of the system. No work is done on this part of the cycle.

Isothermal Expansion This part of the volume transition takes place from point 4 to point 1. The temperature is assumed to have stabilized at T_L and the gel undergoes this volume transition partly in a continuous regime and partly through a phase transition. The total heat transfer in this part of the cycle, Q_{ITE}, is the latent heat involved in the state transition as well as the heat causing the volume change.

$$Q_{ITE} = Q_{CL} + Q_{CW} \tag{2.7}$$

where Q_{CL} is the latent heat of changing the state, and Q_{CW} is the energy that is converted to work through the volume change. The work done during this part of the cycle is :

$$W_{ITE} = \int_{V_1}^{V_2} \Pi(V) dV \tag{2.8}$$

The net heat transferred is negative (out of the system).

2.3.4 Possible Engines and Efficiency Definitions

Depending on the application of the gel in a specific motion system, quantities different from the conventional definition of efficiency may be useful as "efficiency" measurements. A wide range of metrics has been employed in the literature regarding gels and heat engines in general. The total cycle efficiency is an important measure of the performance of a gel that is to be used in a continuous, repetitive motion. This figure will change depending on whether useful work is done on the shrinking or expanding transition, or on both. The gel may also be used to store energy and then release it in a "one-shot" application[6]. For such an application, the ratio of the trigger energy to the stored energy may be more relevant.

For example, consider the application of polymer gels as muscle fibers. The gels will do work on the shrinking phase transition. Antagonistic pairs will be used to produce bi-directional motion. The work output in this configuration is essentially the work done by the shrinking transition. The input energy is the heat used to cycle the gel. A gel can be constructed with its transition point above that of the environment. In this case, the

[6]We indicate by "one-shot type application " any situation where the energy and time to reset the system is plentiful and inexpensive.

environment can be used as a low temperature reservoir at temperature T_L and no energy would be spent cooling the system. The "efficiency" of such an engine is

$$\eta = \frac{W_{ITC}}{Q_{ITC} + Q_{ICC}} \tag{2.9}$$

Another interesting case is the possibility of using the gel to produce useful work on both parts of its volume transition cycle. The ideal gas engine always has pressure in the positive (outward or upward) direction. Therefore, to move the system boundary in the negative direction, work has to be done on the system, subtracting from the total useful work done by the system. In the gel system, however, when the network is contracting, the gel pressure is negative, and when it is expanding, the gel pressure is positive. Useful mchanical work can conceivably be done on both parts of the cycle. With the environment cooling the gel, a useful metric for the "efficiency" of this system is :

$$\eta = \frac{W_{ITE} + W_{ITC}}{Q_{ITC} + Q_{ICC}} \tag{2.10}$$

Other definitions can be derived based on the specific application and the configuration of the gel in the motion mechanism. By computing the energy quantities presented in the previous section, bounds on the system efficiency can be derived.

2.4 Summary

The constituents and process of making a polymer gel determine the different properties of the gel. Many of these properties are important parameters when designing an actuator. It is therefore important to understand the fabrication and chemistry of a polymer gel. A system designer needs to be able to design the material or at least understand the limitations and tradeoffs involved in the design process. In this chapter, the manufacturing process of a polymer gel was discussed. The effects of the chemical components and physical shape and size of a gel on properties such as swelling ratio, tensile strength, and response rate were outlined. Using the formulae and processes for making gel spheres, cylinders, and fibers, different actuators can be built. Another important parameter in designing a transducer is

efficiency. In this chapter, the work cycle for a polymer gel was discussed. Depending on the actuator design, the gel can be made to do work on the expansion and/or contraction cycle. By understanding the energy exchange during the different parts of the transition, the efficiency of a particular gel engine can be derived. This chapter provides general guidelines for computing efficiency bounds on gel actuators. The dynamics of a volume phase transition are explored in the chapters that follow. A switching regulator is implemented and used to identify a mathematical model for the gel dynamics to be used for control design.

Chapter 3

Switched Actuators

An actuator produces a change in the mechanical state of a system by providing a change in stress, strain, shape, stiffness, position, natural frequency, damping, friction, fluid flow rate, or some other mechanical characteristic [86] [12]. In so doing, the actuator converts some form of input energy into mechanical energy. In practice, the choice of an actuator depends on application-specific requirements for force density, efficiency, weight, size, and cost. The ability to modulate input power to the actuator may also be an important factor for actuator selection. Controlling a servomechanical system will obviously be difficult if the inputs to the system actuators cannot be easily changed.

For example, a DC motor is an actuator that converts electrical energy to and from mechanical energy. Magnets or a field winding generate a magnetic field inside the motor. The armature coil current develops a force that rotates the shaft [7]. The transfer function from the electrical input (a voltage source or a current source) to angular velocity of the armature depends on the electrical and mechanical properties of the motor, such as coil resistance, inertia, and damping. A typical linear controller for a system that uses a DC motor to regulate the velocity or position of a shaft will command voltage or current as a function of the error between the desired trajectory and the measured or estimated trajectory. A simple controller with a proportional compensator will regulate the velocity of the shaft to a constant level by measuring the error between the desired velocity and the true velocity and increasing or decreasing the electrical input (and therefore the input power) by an amount proportional to that error.

A similar technique could be used to regulate the output voltage of a linear power supply. The error in voltage at the output due to a disturbance or change in load is used to vary the input power to restore the desired voltage. Ideally, the controller ultimately commands the amount of input current necessary to deliver enough power to maintain the output voltage at a desired level. Active controllers drive system outputs and states to desirable locations by modulating the power flowing through a system. For reasons that will become clear in the following sections, this method of control is not always suitable for a polymer gel actuator. A different technique described below is often more effective and useful for polymer gels as well as a variety of other systems.

3.1 Examples of Switched Systems

One control technique is to command the flow of power through the actuator or transducer. The input power of a linear power supply, for example, can be regulated by directly commanding a level of input current. Another technique for modulating power levels and other system state variables is by switching among discrete levels to produce effective "averaged" values that are a function of the switching scheme. This method is used frequently in power electronics [45] and somewhat less extensively in other systems. In the sections below, examples of switched control and systems that may lend themselves to this type of control are reviewed. The examples motivate the potential benefits of switching control and highlight the issues involved in implementing a switched system. As will be demonstrated, bounds on switching frequency and the characteristic time constants of different system components are vital considerations in system design. Because the high switch rates and fast power delivery of electronic circuits are not always easily achievable in other systems, the implementation and modeling of other types of switched systems is often more challenging. The rest of this section will illustrate these issues using a variety of examples. The next chapter will propose modeling techniques for switched actuators with different physical constraints. Using these results, a model for a switched polymer gel actuator will be developed.

3.1.1 Power Electronics

In switching power electronics, a DC-DC converter might use a switch to alternately connect and disconnect a power source to a linear circuit, resulting in a square wave of voltage at the input to a filter. The *Buck* converter shown in Figure 3.1 is one example of a switched DC-DC converter. By controlling the on-time of the switch (or the time the ideal voltage source is connected to the circuit) the average input power is modulated. If the RLC circuit is designed to be a low-pass filter with a cutoff frequency well below the switching frequency of the input, largely only the DC component of the voltage $v_m(t)$ will appear at the output. The resulting DC output voltage across the resistor can be controlled by the switching profile. Switches can be used to change the topology and dynamics of the circuit as well as the input in order to achieve desired performance[1]. The resulting time-varying systems pose challenging problems in modeling and control [45].

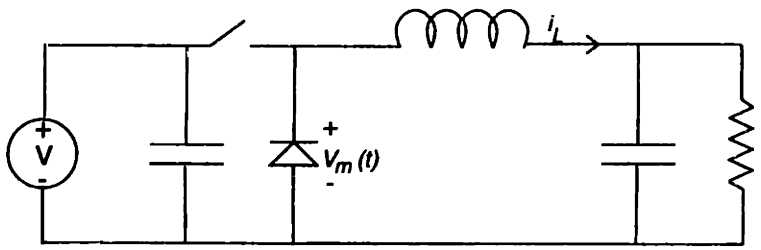


Figure 3.1: Buck converter

3.1.2 Temperature Control

Power supplies are not the only examples of *switched* systems. Consider the temperature control system shown in Figure 3.2. A typical water circulator regulates the temperature of the fluid in its reservoir using a compressor, heater, and a linear (PID) controller [106]. This fluid is pumped through the outer jacket of a double-jacketed glass container. The heat is then conducted to the contents of the glass jar through the glass wall in order to control the temperature in the inner chamber of the jar. This setup is a common labora-

[1]In the case of the buck converter the dynamics are the same regardless of the switch states.

tory arrangement for creating a temperature controlled environment for studying chemical reactions. If one programmable reservoir is used to control the temperature inside the jar, the dynamic response of the system will be limited by that of the reservoir temperature. The thermal inertia of the fluid in a large circulator is larger than that of the jar and its contents. The dominant time constant in the system is typically the time it takes to adjust the temperature of the fluid in the circulator. For a given step in input power, the time it takes the temperature of the jar contents to reach equilibrium is at least as long as the time it takes for the temperature of the circulator fluid to settle.

Using the two reservoir arrangement shown in Figure 3.3, a different control scheme for changing the temperature in the jar can produce faster response times. Each water circulator regulates its fluid supply to one of two temperature rails, T_H or T_L. The container temperature can be regulated by switching the flow through the outer jacket between the hot and cold reservoirs. Consider a periodic switching scheme where for m seconds the hot fluid is flowing in the outer jacket of the jar, then for n seconds the cold fluid is flowing. If the period, $m+n$ seconds, is short compared to the thermal time constant of the jar and its contents, the jar temperature will not change significantly over a switch period. The result is a steady-state jar temperature that can be modulated with the duty ratio $\frac{m}{m+n}$, with a rate of response limited primarily by the thermal inertia of the jar, not the circulators. If both temperature circulators and their pumps are perfectly matched, the heating and cooling processes are governed by the same dynamics, with different input levels. The thermal dynamics act as a filter for the square wave of temperature applied to jar and its contents. Much like the buck converter, the components of the input at the switching frequency and its harmonics are filtered by the system dynamics if the switching frequency is chosen to be much higher than the natural frequencies of the system. A reasonable model for this system is the thermal system driven by the DC component, or the average value, of the input temperature square wave. This DC component can be varied by changing the duty ratio of the switching waveform. The importance of *averaged* models will be discussed further in Chapter 4.

Model derivation is more challenging when the two switched systems are not identical.

The dynamics are time-varying because they depend on the switch state. To demonstrate this, the system in Figure 3.3 was assembled and tested[2]. The hot reservoir was set to a temperature of 44°C and the cold one to 24°C. The state of the valve can be controlled by a digital electronic signal to direct either the hot or the cold fluid to flow through the outer jacket of the glass jar. A thermocouple was placed in the water in the inner chamber of the jar to observe the temperature fluctuation as the flow of water in the outer jacket was switched. The pump speed on the hot water circulator provided a higher flow rate than that of the cold water circulator. This mismatch will illustrate how a time-varying system can result if the dynamics of the system change with the switch states.

In general, the heat transfer process between the circulating fluid and the jar contents is modeled by partial differential equations governing the thermal diffusion in the system. Under certain circumstances, however, the model can be simplified to an ordinary differential equation analogous to an RC-type circuit [56]. Under conditions of good mixing of the fluid and other assumptions discussed further in Section 4.3.1, a circuit analog composed primarily of linear components can be used to model the thermal system. The fluid circulator can be modeled as a voltage source (with temperature as voltage) that is delivering power to the thermal inertia of the jar and its contents (thermal mass or "capacitance")[3]. The electrical resistance in the RC circuit analog is a function of heat losses to the laboratory environment from the jar and tubing. The dynamics of heating and cooling the jar are governed by two different time constants because of the different flow rates. The switched system is therefore time-varying, with different dynamics depending on the switch state.

Figure 3.4 illustrates experimental results from a few switched cycles of the system in Figure 3.3. The solid trace is the experimental data. The lower trace is a simulation of a linear time-invariant (LTI) model with a single time constant of 28 seconds, which is the time constant of the heating process. In the simulation, this model is driven by the same switched temperature profile used in the experiment (i.e.. with a square wave alternating between 44°C and 28°C at the same times the valves were switched in the experiment). The

[2]The two water circulators used are VWR, model 1167.

[3]The reader may refer to Section 4.3.1 and Figure 4.4 for more detail on this circuit model.

results of the simulation do not reproduce the experimental results well, as shown in the figure. A more accurate model switches the time constant of the first-order system every time the valve state is changed. When the hot reservoir is in communication with the jar, the system is modeled by a 44°C input driving a first-order system with a time constant of 28 seconds. When the valve state changes and the cold water circulator is selected, an input of 24°C driving a first-order system with a time constant of 44 seconds is used as the model. This is a time varying model since the dynamics are a function of time. Simulation of this model shows better correspondence with observed results.

The *averaged* dynamics of this system describe the response of the average temperature of the jar contents to the average of the switching profile. The amplitude of the ripple around this average can be reduced in practice by using higher frequency switching. The temperature regulation specifications will determine the tolerable size of the ripple. In situations where the ripple is small and the average value of a state variable is always close to the instantaneous value, a model that predicts the average behavior of the systems may be a suitable basis for control design and performance estimation. One method for deriving a time-invariant model for the averaged dynamics of a time-varying switched system is *state-space averaging*, a technique used to model and predict the average behavior of state variables in switched circuits [45] [54] [55]. This technique will be used as the basis for developing averaged models for other systems, including a polymer gel actuator.

3.1.3 Mechanical Systems

Mechanical systems with switching drives might use sliding mode control or hysteretic control to ensure stability in the face of unmodeled dynamics [16] [110]. In sliding mode control, for example, a sliding surface is defined using a constraint on the states of the system [91]. Along this surface the distance between the desired trajectory of a system and its measured variables is guaranteed to decay to zero given an estimate of the system parameters. A discontinuous switching input is then used to force the system toward the sliding surface. The size of the discontinuity is larger with the degree of uncertainty in the system parameters. The result is *chattering* of the state variables around the sliding surface.

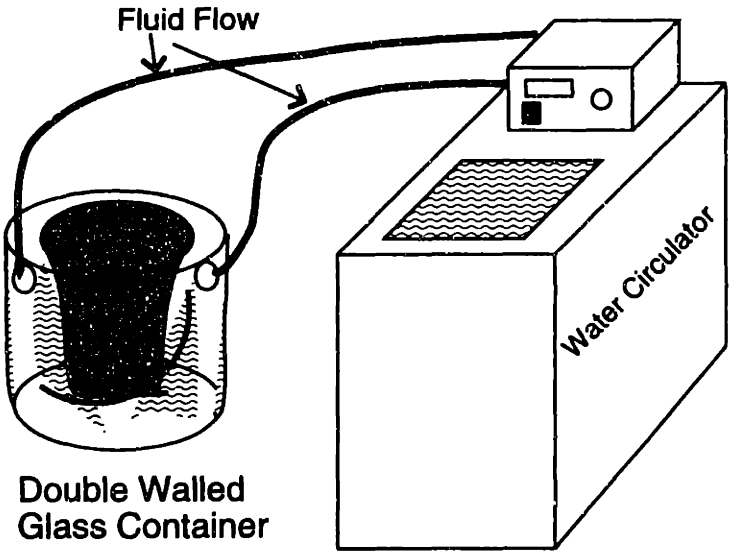


Figure 3.2: Temperature control system with one circulator

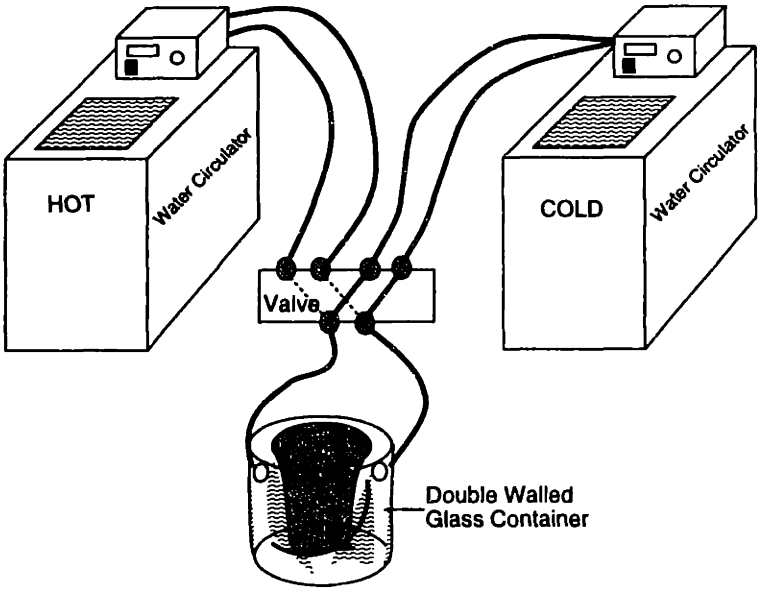


Figure 3.3: Switched temperature control system

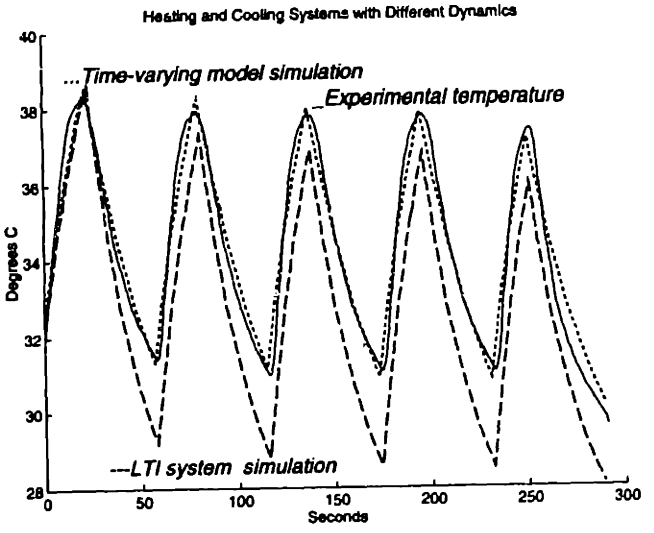


Figure 3.4: Switched temperature profile

This nonlinear control technique as well as the techniques based on state-space averaging that will be developed in the next chapter can be used with switched mechanical systems. Some examples of mechanical systems that may be suitable for this kind of operation are presented below.

Piezoelectric Actuators

One example of a switched, or pulsed, actuator is found in piezoelectric systems. Piezoelectric actuators can be broadly split into two categories [12]. *Traveling wave* actuators produce a wave of structural distortion in a stationary member (like a motor stator) that massages a moving element (like a rotor) and causes it to move in the direction of the wave. In *standing wave* devices, a vibrating element thrusts against a moving element periodically. Because the magnitude of the change in strain in piezoelectric material is extremely small (10E-7 mm per volt), it often cannot be used directly to produce the required motion. The standing wave devices provide a periodic mechanical pulse to a moving element, analogous to spinning a wheel, by periodically giving it a little kick [12]. The bandwidth of the material, as high as 150 KHz, allows high frequency pulsing. High spinning velocities can

be achieved. The resulting standing wave piezoelectric actuator is a switched mechanical system.

Human Muscular System

The human muscular system is also a switched biomechanical system [1]. A single stimulus from a nerve cell causes a muscle cell to contract briefly, followed by relaxation. This is known as a twitch. If a second shock is given to the muscle before the relaxation is complete, "summation" occurs. This is shown in Figure 3.5. With a high frequency of stimulation as shown, the muscle is in a contraction regime referred to as tetanus. Note also that the contracting and relaxing dynamics of the muscle are different. Contraction occurs at a much faster rate. The overall behavior of the muscle is a function of both dynamics, resulting in a time varying system crudely similar to the temperature control experiment described above.

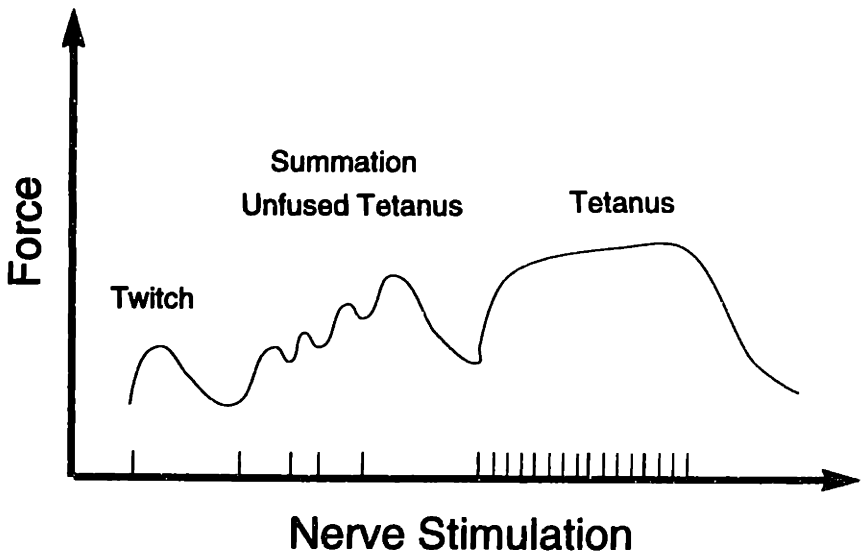


Figure 3.5: Muscular contraction in response to nerve stimulation adapted from [1]

Pneumatics

Pulse width modulation (PWM) has been used in mechanical systems to control pneumatic cylinders. Speed control of a pneumatic cylinder is often done open-loop using a flow control valve. Closed loop control becomes important when there are changes in friction or load force during operation. Especially at low speeds, speed regulation is difficult. One way to accomplish this is using electro-pneumatic servo or proportional valves, which are often cost-prohibitive. In [76], an inexpensive on-off solenoid valve is controlled using a pulsed waveform in order to regulate the speed of the pneumatic cylinder.

Robotics and Novel Actuator Material

Time-variance due to the dependence of dynamics on switch state can also be exhibited by switched mechanical transducers. Multi-jointed manipulators and tendon-driven robotic systems that use antagonistic actuation have been developed [39] [63]. One can envision a pulsing control technique that provides stiffness at a certain position by alternately pulling on each of the antagonistic actuators, which may be different. Polymer gels have also demonstrated this difference in dynamics between their shrinking and expanding volume transitions [90] [17]. Other novel actuator technology is based on shape-memory alloys (SMA), and magnetostrictive materials [87] [90] [86]. Depending on their activation and design, these actuators can exhibit disparate system responses on opposite sides of their work cycle. For example, SMA's are cycled by heating and cooling. The dynamics of water cooling (used in [12]) are likely to be different from those of the heating mechanism. As with the thermal system in Figure 3.4, switching these systems will likely result in a time-varying plant. Models and controllers for these systems are more challenging to develop than for LTI systems.

One actuator design discussed in [12] uses two antagonistic pieces of SMA. An SMA "memorizes" its form after being annealed. If the shape is deformed below a certain critical temperature, heating the alloy above this temperature will recover its shape. The antagonistic actuator design is shown in Figure 3.6. The elongated spring has been deformed while below the critical temperature. The contracted one has recovered its memorized shape after

being heated. As the heat is transferred from one SMA to the other, the hinged arm moves in an arc toward the heated SMA and pulls the other SMA out of shape. A similar actuator was built using polymer gels [8]. One of the SMA actuators is replaced with polymer gel while the other is replaced with a passive spring. When the gel is triggered (by changing the pH level of its solvent in this case) it contracts and moves the arm. When it is triggered to relax, the spring restoring force pulls the arm back in the opposite direction. Both the SMA and also the polymer gel described here exhibit a change in length that is *discontinuous* as a function of the input (temperature). Often in a servo-mechanical application, a continuous range of equilibrium mechanical states is desired. With a discontinuous actuator such as the gel or SMA, this cannot be achieved by varying the input continuously. Instead, by switching the direction of the states of the actuator around the desired trajectory, a dynamic equilibrium can be achieved where a static one was not attainable. This is discussed further below in the context of polymer gels.

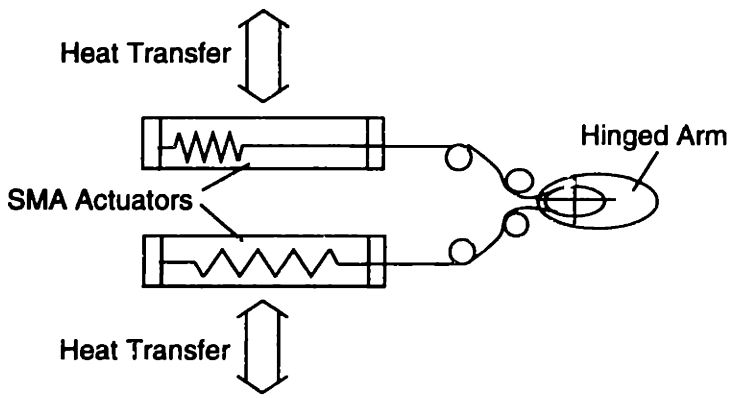


Figure 3.6: SMA actuator from [12]

3.2 Motivation for Switched Drives for Polymer Gels

Polymer gels can be fabricated to exhibit continuous volume change or discontinuous volume-phase transitions [98]. The stylized discontinuous volume-phase transition curve in Figure 3.7 is typical of many polymer gels [17] [97]. The use of discontinuous actuators as mo-

tion devices poses many challenges. The large discontinuity in volume represents mechanical states that are unattainable in equilibrium no matter what the steady-state temperature is. Consider a cylindrically shaped gel that is vertically suspended and attached to a weight at the bottom. The gel volume (and therefore length) can be controlled by temperature to regulate the position of the height of the weight. However, with the volume-phase transition curve shown in Figure 3.7, the range of attainable steady-state positions is very limited. In particular, any height that requires the volume of the gel to be in the unattainable region cannot be reached; a static equilibrium in this range is not possible. Even though the volume of the gel passes through this desired state during the course of its transition, it cannot settle there.

In a practical application, there is usually a specification of the error allowed between the desirable reference and the actual trajectory. If the reference dictates that the gel volume has to be in a state unattainable in equilibrium, then the gel can be made to continuously move in the vicinity of the reference. If the motion results in deviations that are within the error limits, the specification is met. For a discontinuous NIPA gel (NIPA is a thermo-responsive polymer gel that shrinks at higher temperatures. Fabrication of NIPA with ionic groups results in a discontinuous equilibrium volume-phase transition characteristic [17] [97] [109].), whenever the controller of the system senses that the variables of interest are moving away from their desired trajectory, it can change the temperature to induce motion that will reduce the error. If the sensory and correcting actions are done frequently enough that the error is never allowed to grow larger than some tolerance level, the output will ripple around the reference trajectory with a small ripple that is within the error tolerance level. In effect, the system is being switched around the desired trajectory. A dynamic equilibrium has been reached at a mechanical state where a static one does not exist.

The ability to employ and effectively control discontinuous phase transition polymer gels as mechanical actuators is important for several reasons:

- One of the advantages of polymer gels as actuators is their ability to respond to a variety of environmental changes. (This property can be used to eliminate a sensor

in some applications.) A gel chemically controlled to respond to some environmental trigger may unavoidably exhibit a discontinuous phase transition. It is therefore convenient to be able to control a discontinuous phase-transition polymer gel. Being able to use the largest class of gels, continuous and discontinuous, increases a designer's ability to optimize the other parameters in the system, such as cost or performance.

- Among the important design goals for an actuator are high force density, fast response, and efficient power consumption [12] [33] [90]. Discontinuous volume-phase transitions are characteristic of many of the polymer gels that give the best rates of response and potentially highest closed-cycle energy conversion efficiencies. For instance, it has been shown that the diffusion coefficient for a NIPA gel increases with the ionization content of the network [68], which results in a quicker gel. It has also been shown that ionized free-swelling NIPA gels exhibit discontinuous phase transitions while neutral ones do not. In the process of making a NIPA gel faster by ionizing its polymer chains, a discontinuous gel is created.

- The "latching" hysteretic behavior of discontinuous gels may be useful in many applications. As a gel goes through a phase change, it stores "latent" energy. This energy can be released with a small amount of "trigger" energy that alters the balance of osmotic forces in the gel, like releasing a coiled spring. This energy storage property can be useful in some applications.

- The largest volume change ratios have generally been observed in discontinuous phase-transition gels [97]. The mechanical stroke of any system with a gel actuator will be a function of this volume change.

- Even if the gel is fabricated to exhibit a continuous volume change in response to input change, its equilibrium phase-transition behavior may become discontinuous under certain loading or external tension conditions [92]. In a servomechanical application, the actuator is subjected to external tension that may change with the loading or operating conditions. It is therefore especially important to be able to control a

discontinuous transition gel in these applications.

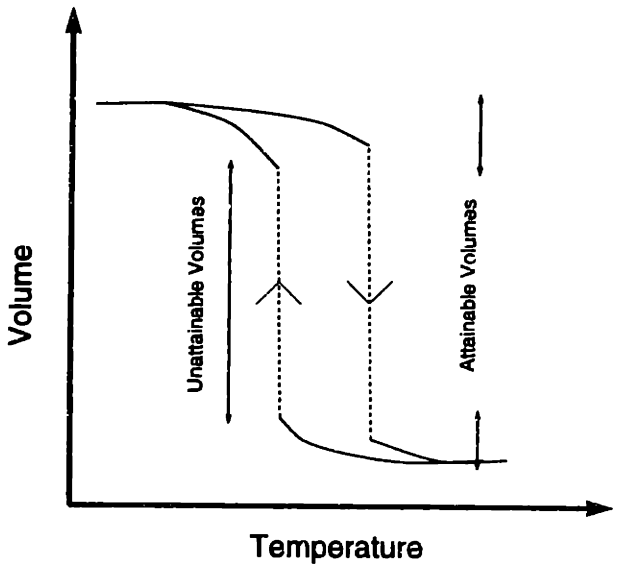


Figure 3.7: Equilibrium states attainable by a gel under constant external tension

3.3 Modeling Switched Systems

In most of the examples discussed above, the objective of the switching control is not to produce an oscillatory output. The switching of the input power or the interconnections of the system is used to modulate the average value of a state or output. The state trajectory is alternately switched in the direction of the desired value as it oscillates around it. The oscillations result in "switching" ripple around the target trajectory. If the switch frequency is chosen to be high compared to the bandwidth or natural frequencies associated with a state or output of interest, the ripple amplitude will be small compared to slower changes in the output [4]. In this manner, the output level can be controlled to within the ripple size. The actual value of the ripple is not important as long as it is maintained below a specified maximum. Given this time-scale separation, the average of the output over a switch period is a good estimate of the true value of the output. A model that describes the input/output

[4]Note that there is no assumption of constant frequency necessary for this statement to be true. As long as the switch period is maintained below a specified length, there will be a bound on the ripple amplitude.

relationship between the switching profile and this average can be used for control design and the analysis of the stability and performance properties of the closed-loop system.

Averaged models may be derived by applying a time-averaging operation to the full switched system model. Several properties of the underlying switched system will determine the degree of difficulty in obtaining an averaged model. In some cases, an averaged model can be derived only by imposing constraints on the operation of the switched system. In others, no analytical description can be derived for the averaged dynamics. When the systems that are being switched can be described by linear, time-invaritant (LTI) models, the overall switched system is piecewise LTI. Sometimes the system topology and dynamics are not a function of the switch states. The switches only modulate the amount of power delivered to the system every cycle. The temperature control system in Figure 3.3 with identical reservoirs and pumping systems is an example of this situation. The switched system is LTI and is unchanged by averaging the system equations. The input drive of the averaged system is the average of the switched input waveform. When the LTI systems in different switch states are different, the overall system is time-varying and the development of a model describing the average state behavior is more challenging. However, averaged systems can still be derived with the help of some simplifying assumptions. Much of the modeling and analysis of switched systems of this kind has been done in the context of power electronics [45], [54]. A power electronic circuit contains LTI components as well as switches. In effect, the state of the switches selects a certain LTI circuit topology and input to the system. Averaging the resulting linear, time-varying models is possible when the switching frequency is high compared to the rest of the dynamics in the system. This is discussed in detail in the Chapter 4.

When the dynamics of the underlying switched systems are nonlinear and the switching rates are not much higher than some of the natural frequencies in the system, the averaging process is more difficult. These conditions are more likely to occur in mechanical rather than electrical systems. Because of their difficulty to model and control, there are relatively few examples of switched mechanical systems. As mentioned above, one area where pulse-width modulation (PWM) has been introduced in mechanical systems is in the control of

pneumatic cylinders. Speed control of a pneumatic cylinder is often done open-loop using a flow control valve. Closed loop control becomes important when there are changes in friction or load force during operation. Especially at low speeds, speed regulation is difficult. One way to accomplish this is using electro-pneumatic servo or proportional valves, which are often cost-prohibitive. In [76], a cheap on-off solenoid valve is controlled using a pulsed waveform in order to regulate the speed of the pneumatic cylinder. No analytical model is developed for the system, however. Instead, measurements of system parameters at different duty ratios D (the fraction of the switch cycle time that a certain switch state is in effect) are measured and an operating range in which the parameters seem to behave linearly is chosen. The dynamics associated with transients in D are ignored. It is also shown that different switching frequencies produce different system response. However, a general model for using the valve at different operating frequencies is not given. Switching ripple is presented as a problem with this control scheme. Increasing the switching frequency to minimize the ripple shortens valve life. Hence, there is a tradeoff between the cost associated with replacing the valve and the cost of the switching frequency ripple in the cylinder velocity.

A similar system is proposed in [73]. A model for the whole system is built from the nonlinear components of a hydraulic actuator system controlled by two-way solenoid valves. The computer model is used to simulate the system and compare it to experimental results. A simple analytical model for the design of controllers to regulate average quantities is not presented, however. In [107], an instantaneous mass flow rate for simulating the dynamic behavior of pneumatic PWM control systems is presented. Furthermore, an equivalent mass flow rate model is developed. Like an averaged model, this model describes the dynamics of the "aggregate" behavior of the flow through the valve from one PWM cycle to the next. It is based on integrating the nonlinear dynamics over one cycle of operation. While this model is still nonlinear and discontinuous, it can be linearized in different regions and used for control design.

In the chapters that follow, the problem of deriving averaged models for mechanical systems is explored. First, in Chapter 4, a general structure of a switched actuator is defined and used to explore the possibility of deriving averaged models for systems with

different time-scale properties. The system components that are being switched are assumed to be LTI or at least modeled under some constraint by an LTI system. The averaging methods are borrowed from power electronics techniques and modified for the time-scale constraints that cannot be removed in some systems. The limitations of these methods are discussed along with the tradeoffs involved in their implementation. In Chapters 5 and 6, identification of a model for the averaged dynamics of a switched polymer gel actuator is presented. The dynamics of a volume transition in a polymer gel are nonlinear and vary with the state of the input switch. Direct identification of a small signal linear model for the average dynamics is motivated and the results presented. The small signal linear model describes a change in the length of a cylindrical gel near an operating point.

Chapter 4

Averaged Modeling of Switched Systems with Piece-wise LTI Dynamics

Averaged models for switched systems predict averages of the state variables and outputs[1]. If the switch period is much shorter than the characteristic time constants of the output response, the true output values are well approximated by these averages since the switch ripple is relatively small. An averaged model can therefore provide a simplified yet accurate approximation of the transfer characteristic of a switched system. A common technique for deriving averaged models for switched systems is direct evaluation of the averages of the differential equations governing the system dynamics. This technique is used extensively in power electronics. The fact that LTI models describe most power electronic circuits, for example, in a given switch state facilitates the derivation of averaged models. The basics of this approach are presented below[2]. More detail and more recent work is presented in [45],[10],[70], and [49]. The technique described below will serve as a starting point in analyzing switched mechanical systems and deriving models for their averaged dynamics. A general framework for modeling switched actuators will be developed and used to model the behavior of a polymer gel actuator.

Consider a system made of LTI components and switches, for now, and limit the possible

[1]The windowed averages are computed over a switch period.

[2]In the discussion below, two possible switch states are assumed for ease of illustration. Generalization to a larger number of switch states is straightforward.

switch states to two. In state 1 the system that relates the input $u(t)$ and the state vector $\boldsymbol{x}$ is:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}_1\boldsymbol{x}(t) + \boldsymbol{B}_1u(t). \tag{4.1}$$

In state 2, the equations become:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}_2\boldsymbol{x}(t) + \boldsymbol{B}_2u(t). \tag{4.2}$$

In order to develop a linear, time-invariant, ordinary differential system equation description, suitable for control design, begin by defining the waveform $q(t)$ as shown in Figure 4.1 [45]. The system is in state 1 when $q(t)$ is high. The waveform $q'(t)$ is the complement of $q(t)$. The system is in state 2 when $q'(t)$ is high (or when $q(t)$ is low). By multiplying each system equation by one of the switching functions $q(t)$ and $q'(t)$, the two sets of dynamics become mutually exclusive in time. An alternative description of the system is derived by adding these functions to combine them into one equation:

$$\dot{\boldsymbol{x}}(t) = q(t)\boldsymbol{A}_1\boldsymbol{x}(t) + q'(t)\boldsymbol{A}_2\boldsymbol{x}(t) + q(t)\boldsymbol{B}_1u(t) + q'(t)\boldsymbol{B}_2u(t) \tag{4.3}$$

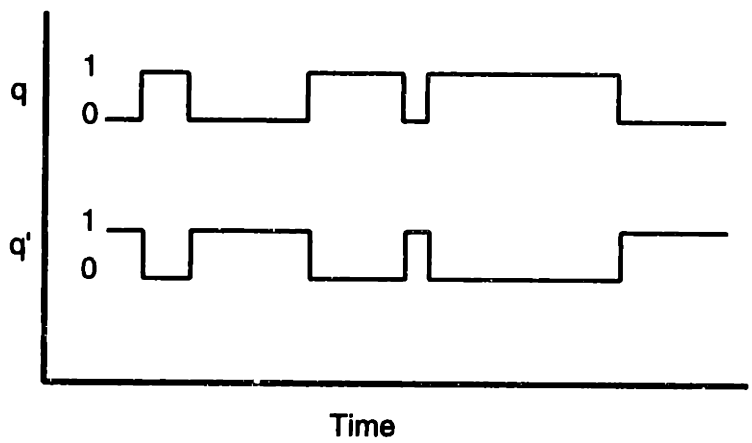


Figure 4.1: The switching functions,$q(t)$ and $q'(t)$

In general this is a time-varying system. If an attempt is made to control the system through the switch function $q(t)$, then the product of $q(t)$ (now an input) with $x(t)$ and $u(t)$ makes it nonlinear as well.

4.1 Time Averages of Switched Dynamics

One common technique for simplifying the analysis of switched systems and deriving useful models for control design is state-space averaging [45] [54]. This technique is used with the assumption that the switching is done at a rate much higher than the natural frequencies of the dynamics of interest in the system. As a result, the state and output trajectories of the switched system over a switch cycle do not deviate significantly from their one-cycle average. A model that describes the dynamics of the average is then a reasonable model for the true dynamics as well. In addition, it has the advantage of being time invariant, as will be shown below.

Consider a system operating with a fixed switch period. Specifically, fix the period of the switch functions $q(t)$ and $q'(t)$ to be T as shown in Figure 4.2. For each period, there is an associated duty ratio D that determines the time each switch state is selected. The function $q(t)$ is high for DT seconds, while $q'(t)$ is high for the remaining $(1-D)T$ seconds in the period.

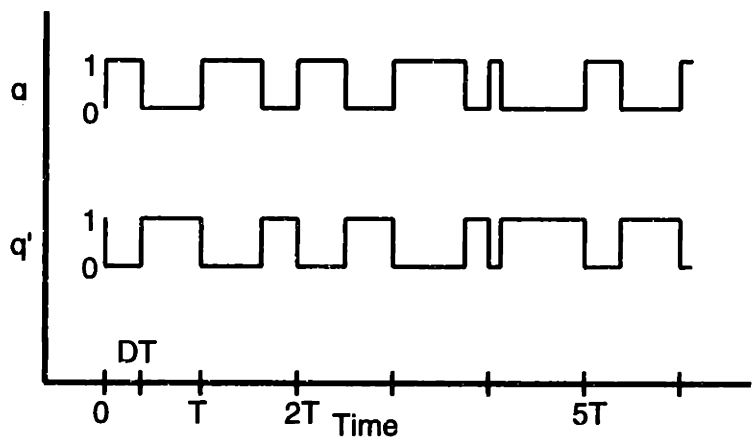


Figure 4.2: The switching functions,$q(t)$ and $q'(t)$ with a fixed switch period

Now define the windowed average over a switch period T of a variable $w(t)$ as:

$$\overline{w}(t) = \frac{1}{T}\int_{t-\frac{T}{2}}^{t+\frac{T}{2}} w(t)dt \tag{4.4}$$

Averaging the two sides of 4.3 over T yields:

$$\overline{\dot{\boldsymbol{x}}(t)} = \overline{q(t)\boldsymbol{A}_1\boldsymbol{x}(t)} + \overline{q'(t)\boldsymbol{A}_2\boldsymbol{x}(t)} + \overline{q(t)\boldsymbol{B}_1u(t)} + \overline{q'(t)\boldsymbol{B}_2u(t)} \tag{4.5}$$

Factoring out the constant system matrices describing the LTI systems:

$$\dot{\bar{\boldsymbol{x}}}(t) = \boldsymbol{A}_1\overline{q(t)\boldsymbol{x}(t)} + \boldsymbol{A}_2\overline{q'(t)\boldsymbol{x}(t)} + \boldsymbol{B}_1\overline{q(t)u(t)} + \boldsymbol{B}_2\overline{q'(t)u(t)} \tag{4.6}$$

In general, the averages above cannot be further simplified without additional assumptions.

4.1.1 time-scale Separation

As mentioned previously, the state-space averaging method for modeling switched systems is typically used when system quantities are well approximated by their average. For instance, it may be that a state variable to be controlled to a certain reference never deviates much from its average over a single switch period. A model for the average dynamics in this case is enough to design a controller to regulate the state variable by regulating its average. If it is feasible to implement the switching at a frequency T^{-1} that is much higher than all the natural frequencies of the system and the frequency content of the input, Equation 4.6 can be simplified to an LTI system description. The actual states and input values can be approximated by their averages over a switch period since they will not change significantly over this window of time:

$$\boldsymbol{x}(t) \approx \bar{\boldsymbol{x}}(t) \qquad u(t) \approx \bar{u}(t) \tag{4.7}$$

These values can then be assumed constant over a switch period and factored from the averaging integrals implicit in 4.5 and 4.6:

$$\dot{\bar{\boldsymbol{x}}}(t) = \bar{\boldsymbol{x}}(t)[\boldsymbol{A}_1\bar{q}(t) + \boldsymbol{A}_2\bar{q}'(t)] + \bar{u}(t)[\boldsymbol{B}_1\bar{q}(t) + \boldsymbol{B}_2\bar{q}'(t)] \tag{4.8}$$

The averages of the switch function over a period T are

$$\bar{q}(t) = D(t) \qquad \textit{and} \qquad \bar{q}'(t) = 1 - D(t) = D'(t), \tag{4.9}$$

where $D(t)$ is the duty ratio as in Figure 4.2. Equation 4.6 can be written as

$$\dot{\bar{x}}(t) = D(t)\boldsymbol{A}_1\bar{x}(t) + D'(t)\boldsymbol{A}_2\bar{x}(t) + D(t)\boldsymbol{B}_1\bar{u}(t) + D'(t)\boldsymbol{B}_2\bar{u}(t) \tag{4.10}$$

or, more compactly,

$$\dot{\bar{x}}(t) = \boldsymbol{A}\bar{x}(t) + \boldsymbol{B}\bar{u}(t), \tag{4.11}$$

where the *average system matrices* are defined as

$$\boldsymbol{A} = D(t)\boldsymbol{A}_1 + (1 - D(t))\boldsymbol{A}_2 \quad \textit{and} \quad \boldsymbol{B} = D(t)\boldsymbol{B}_1 + (1 - D(t))\boldsymbol{B}_2. \tag{4.12}$$

Equation 4.11, with a constant duty ratio $D(t) = D$, is a linear time-invariant description of the average dynamics of the switched system. When $D(t)$ is used to control the system and is a function of the states in feedback, the system becomes nonlinear, but is still time-invariant.

4.1.2 Special Cases

In some instances the requirement that the states and inputs be close to their average over T is not met. Nevertheless, it may still be desirable and possible to obtain a useful averaged model. Other constraints may allow the average dynamics of the system to be modeled by Equation 4.6. For instance, consider a case where the switch frequency T^{-1} is much higher than the natural frequencies of the system but the input value varies significantly during a switch period. The approximation of the input value by its average value over a switch period is no longer valid. In general, the second term on the right hand side of Equation 4.11 cannot be obtained because the average of the product of $q(t)$ and $u(t)$ is not the product of the averages in this case. However, if we constrain the $\boldsymbol{B}$ matrices so that

$$\boldsymbol{B}_1 = \boldsymbol{B}_2 = \boldsymbol{B}, \tag{4.13}$$

then the input terms in 4.6 can be simplified:

$$\overline{q(t)\boldsymbol{B}_1 u(t)} + \overline{q'(t)\boldsymbol{B}_2 u(t)} = \overline{(q(t) + q'(t))\boldsymbol{B}u(t)} \tag{4.14}$$

and since

$$q(t) + q'(t) = 1, \tag{4.15}$$

we can simplify Equation 4.14 by realizing that

$$\overline{(q(t) + q'(t))\boldsymbol{B}u(t)} = \boldsymbol{B}\bar{u}(t). \tag{4.16}$$

The average dynamics of the system will be governed by the average of the input in this case, allowing the use of the averaged model

$$\dot{\bar{\boldsymbol{x}}}(t) = \boldsymbol{A}\bar{\boldsymbol{x}}(t) + \boldsymbol{B}\bar{u}(t), \tag{4.17}$$

where $\boldsymbol{A}$ is the matrix from Equation 4.12.

A similar argument can be carried out for the $\boldsymbol{A}$ matrix. If $\boldsymbol{A}_1 = \boldsymbol{A}_2$ an averaged model can still be developed without the constraint of the switching frequency being much higher than the natural frequencies of the system.

A redundant case is one where the input is switched between two values one in each switch state. In this situation the value of the input for each state can be multiplied by the respective $\boldsymbol{B}$ matrix to produce a new input matrix with a constant input for all time. In this case, an averaged model can be derived as well. Assume the input can take on one of two constant values u_1 and u_2. In the first switch state

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}_1\boldsymbol{x}(t) + \boldsymbol{B}_1 u_1. \tag{4.18}$$

In the second switch state:

$$\dot{\boldsymbol{x}}(t) = \boldsymbol{A}_2\boldsymbol{x}(t) + \boldsymbol{B}_2 u_2. \tag{4.19}$$

If the new $\boldsymbol{B}$ matrices are defined:

$$B_1' = B_1 u_1, \qquad B_2' = B_2 u_2, \tag{4.20}$$

an equivalent system description in the first switch state is:

$$\dot{x}(t) = A_1 x(t) + B_1' U(t). \tag{4.21}$$

In the second switch state:

$$\dot{x}(t) = A_2 x(t) + B_2' U(t), \tag{4.22}$$

where $U(t)$ is the unit step function. For all $t > 0$ the input to this system is constant. An averaged input matrix B can be derived:

$$B' = D(t)B_1' + (1 - D(t))B_2' \tag{4.23}$$

4.2 General Structure of Switched Systems with Input and Transducer Dynamics

In power electronics, a switching frequency is usually chosen to be high compared to the dynamics of the systems being switched [45] [55]. Ripple is thus minimized and the averaged systems described above can be developed with the assumption of constant state and inputs over a switch period. The high bandwidth of electronic switches makes this possible. In other systems, however, this may not be the case. Physical constraints may limit the switch frequency, for instance. As an example, consider the SMA example shown in Figure 3.6. The bandwidth was limited to 1 Hz by the water cooling process [12]. The switching rate is limited by the rate of heat transfer into and out of the system. Suppose that the SMA is attached to a load that allows it to recover its shape, when heated, with a time constant close to a second. In this case the time-scale separation required to average a system modulated by the duty ratio of a fixed period switching function cannot be achieved.

In many physical applications there is a switch that changes the flow of power into the system through some input mechanism. The speed of this switch is the first limit on

the response rate of the system. The rate at which power can be delivered to the system component that accomplishes the energy conversion is also important. For example, in the temperature control application in Figure 3.3, the time it takes for the valve to switch states limits the switching frequency. Consider placing in the jar a polymer gel that converts the energy it receives in the form of heat to mechanical work. Once the switch state is set, the temperature-controlled fluid has to travel through the rubber tubing and heat or cool the jar before it can adjust the temperature of the gel. The transport delay through the pipes and the thermal inertia of the glass container can be modeled as an input stage that is delivering power to the gel. The speed of this "input" system places another limit on the switching speed of the system.

A reasonably general structure for actuators that have a switched input and input dynamics is shown in Figure 4.3. The different components are discussed below with the help of some examples.

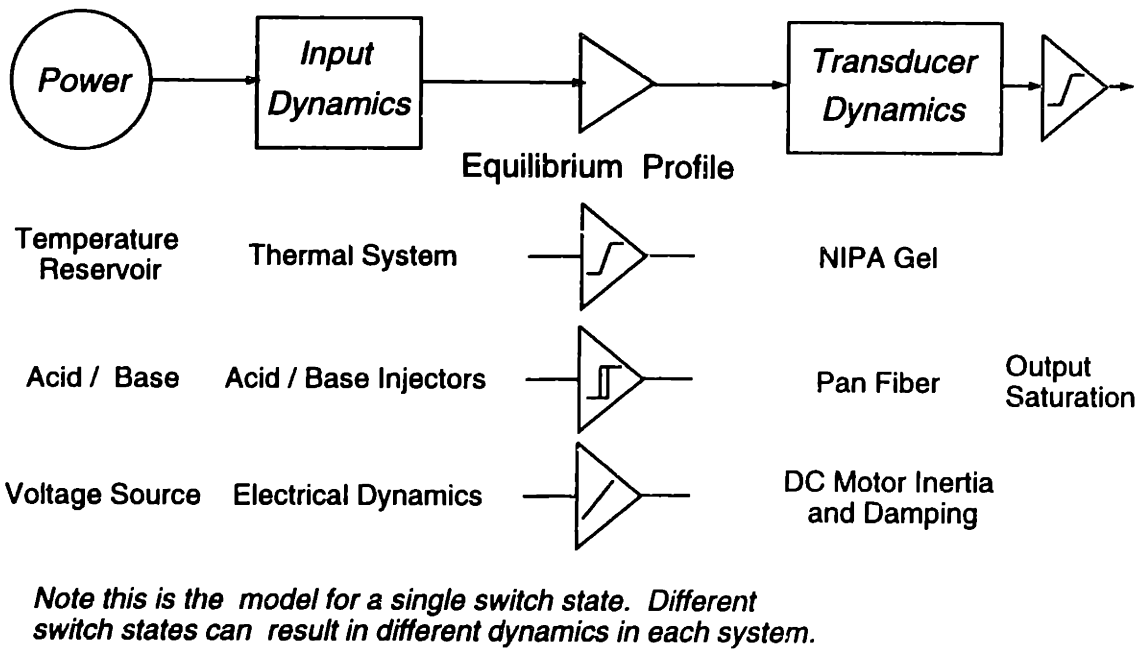


Figure 4.3: General structure for switching systems with input dynamics

- **Power Source and Input System:** These components of the system provide the input energy to be converted by the transducer. For example, the water circulators as mentioned above in Figure 3.3 hold the thermal energy to be delivered to the jar and converted by a polymer gel or some other engine. They are the source of input power in the system. The tubing and thermal inertia of the transporting system determine the dynamic response of the input delivery system. In the human muscular system, the energy is stored as chemical energy. This energy is converted through a chemical reaction to a form that biological muscles can use. Most actuator systems have a source of energy input (battery, hydraulic pressure) and a mechanism by which to deliver this energy to the actuator device or transducer that actually carries out the energy conversion (electrical circuit, pipe fixtures). The power source in the figure is the chemical energy of the muscle, the battery, or hydraulic pressure source.

- **Switches:** The figure illustrates only one switch state of the system. In a switched system, input power modulation is accomplished by changing the state of switches than control the communication between the power input and the rest of the system. Different switch states will select different input power levels to be connected to the system. By varying the switching profile, different amounts of power can be delivered to the system. The system may also exhibit a change in dynamics depending *directly* on the switch states (time-varying dynamics) or *indirectly* changing as a result of a change in the input or state of the system (nonlinear dynamics). Such a system is time-varying, as discussed previously[3]. Sometimes the state of the switch is a user input (controllable switches) and sometimes it is a function of the rest of the variables in the system (a diode switch or the temperature dependency in a gel, for instance).

- **Equilibrium Transfer Characteristic:** The equilibrium function stage describes the steady-state relationship between the input variable and the output variable. For linear systems, this is simply a gain block. This functional block can capture any

[3]Examples of this are changing circuit topologies in power electronics or different rates of response for shrinking and swelling transitions in gels.

static non-linearities in the system, such as saturation and hysteresis. As a result it is easier to capture the remaining behavior of the transducer using linear components and switches. Note the saturation block at the output. It may be more accurate to capture the saturation behavior of some systems at the output rather than the input.

- **Transducer Dynamics:** The transducer stage converts the input power through a process that can be described by some transfer characteristic. A DC motor will change electric power to mechanical power. A DC-DC converter will change the voltage and current levels of the electric power flowing through the system. An SMA converts heat to mechanical energy. This process has certain dynamic properties. The dynamics will be assumed to be linear in this chapter to facilitate time constant analysis and frequency domain arguments[4].

As alluded to before, the thermo-sensitive NIPA gel placed inside the glass container in Figure 3.3 with switched sources of temperature controlled fluid can be described by the model in Figure 4.3. The power sources are the water circulators and the input dynamics are those of the tubing and the transportation system dynamics. One of the switches in the system is the valve that selects the hot or cold water flow. The dynamics of the transducer are those of the gel volume transition in response to a temperature change. These dynamics may be different depending on whether the gel is shrinking or expanding, and therefore there is another "switch" that toggles at the points where the gel changes direction of volume change. The equilibrium function for the gel will approximately resemble a gain with saturation rails for a continuous volume transition NIPA gel, for instance.

Another gel example is a PAN hydrogel actuator [52]. Properly prepared PAN fibers exhibit a volume-phase transition in response to a change in the pH level of the gel solvent. A fluid spray system is used to deliver the acid and base doses necessary for pH control, and an irrigation system is used to remove the resulting salt. The energy is stored chemically in the solvent. The fluid spray system is the input system, with two inputs (acid and base)

[4]Nonlinear systems can often be modeled by linear dynamics. For instance, under small-signal operation, a linear approximation may be an accurate representation of a truly nonlinear system.

which can be alternately switched to cause the desired motion of the gel. The transducer is the PAN hydrogel, converting chemical energy to mechanical energy. The gel used in [52] is discontinuous, and an equilibrium block with two output states and hysteresis can be used to describe its steady-state behavior. A different example is a permanent-magnet DC motor. The electrical current is delivered to the coils through electrical dynamics governed by the inductive and resistive properties of the motor armature. This electrical subsystem is the input phase in Figure 4.3. The current in the coils is transformed to mechanical energy through slower mechanical dynamics (transducer component). If no linear power supply is available, the input current can be modulated by rapidly connecting and disconnecting a fixed input. The slower mechanical dynamics will generally filter out the ripple, and different velocities or output torques can be achieved.

The relative time constants of the components in the system in Figure 4.3 and the period of the switching function will interact to produce certain input and output profiles. The goal of the next few sections is to explore the effect of different switching frequencies on systems with different input and transducer time constants. Much of the work has been done with a thermo-sensitive NIPA gel in mind. Models based on a thermal input system and a mechanical actuator will be used for the analysis. These models are simple first-order models that serve to capture the dominant dynamics of the different system components. Without any loss of generality, this allows us to define input and output variables and to relate the results of the analysis to a physical system. The derivation of a mathematical model for this system using state-space averaging is explored in later sections. The model will be derived in terms of the system matrices of the underlying dynamic systems. Any physical system that has the same time-scale properties and structure as defined in Figure 4.3 can be modeled in the same manner. By varying the functional blocks and time-scale properties in this general structure, derivation of averaged models for a large class of physical systems can be explored. The discussion below will examine some interesting classes of actuator mechanisms.

4.3 Models for Switched Actuators with Discontinuous Equilibrium Behavior

First, the conceptual view of a switched system presented in Figure 4.3 is used below to formulate a general model for systems with *discontinuous* equilibrium properties[5]. Linear, first-order models for the input dynamics and the transducer dynamics are used, and the analysis is limited to two switch states. The system model assembled out of these reduced models and a discontinuous equilibrium profile can represent a large class of actuators. While it is important to maintain the generality of the model, it will be discussed in the context of a discontinuous volume phase-transition polymer gel that responds to temperature changes. This allows the discussion to refer to physical variables and should solidify the intuition behind the behavior of the system under different assumptions. The components of a polymer gel actuator system are briefly described below[6]. Simplified physical models are used to understand the origin of their dynamics. By exploring the behavior of switched systems in this framework with different time-scale properties, general conclusions can be drawn about whether or not an averaged model can be derived for the various cases and with what assumptions. The time constants associated with each system component in the reduced linear models used below represent the general scale of the dynamic response time.

4.3.1 Input Model

A model of a thermal input stage can be derived based on the system in Figure 3.3. Imagine that a polymer gel (or other thermally-activated actuator) is placed inside the glass container. The temperature of the gel, T_g, is the output of some thermal system delivering heat flow to the gel. An electrical analog for the thermal system can be developed [56]. The heat transfer process that raises the temperature of the gel solvent (water in the jar) is generally modeled by diffusion equations. With some additional assumptions discussed in Chapter 6, the diffusion effects can be lumped into a first-order dynamic model. The parameters of this model are a resistive path for the heat (R_{th}) and a thermal inertia (C_{th}).

[5] *Continuous* gels are discussed in Section 4.4

[6] Details of the gel model will be explained in Chapter 6.

A heat loss element, if the system is not well insulated for instance, could be added. The resistance R_{L1} represents dissipation from the power source into the surroundings and R_{L2} represents ambient dissipation by the jar and tubing. The result is shown in Figure 4.4. The equation governing the dynamics of the simplified system assuming perfect insulation ($R_{L1} = R_{L2} = \infty$) is:

$$\dot{T}_g = \frac{T_r}{R_{th}C_{th}} - \frac{T_g}{R_{th}C_{th}} \tag{4.24}$$

where T_r is the temperature of the temperature reservoir.

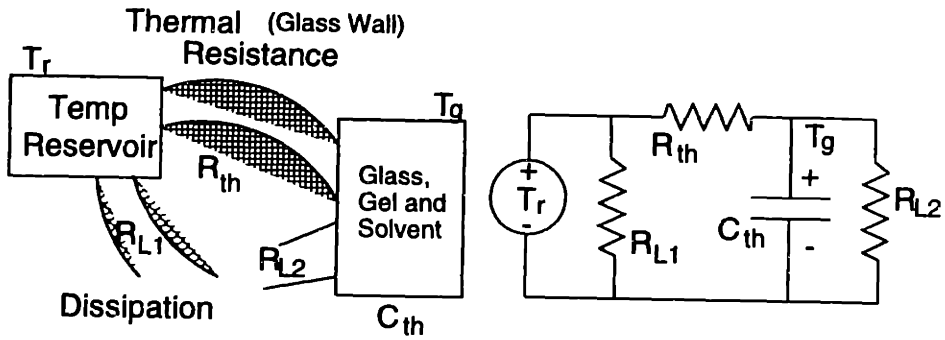


Figure 4.4: Temperature control and heat delivery system

The voltage source in the circuit is a constant temperature reservoir. Temperature is the analog of voltage in this model. Current in the model represents the heat flow (Watts). The ambient temperature is the ground reference in the circuit analog. The assumption of perfect insulation reduces the electrical analog of the thermal system to a series RC circuit that will be used later for simulations.

4.3.2 Transducer Model

A first-order model of a mechanical system is illustrated in Figure 4.5. The output variable $y(t)$ represents a mechanical *displacement*, which will be referred to as the transducer

output in the sections below. For a polymer gel, the process by which a volume transition occurs is the diffusion of the chains of a polymer network into and out of a solvent. The friction between the network and solvent is represented by the damper in the model. The disturbance in internal pressure for the gel produces a force input (U) that will result in a new equilibrium size of the network by stretching it or collapsing it through this diffusion process. The spring in the model represents network elasticity. The small mass of the network has been ignored, and the diffusion process has been oversimplified to a first-order response to facilitate the discussion, as mentioned earlier. By reducing the order of the system in this manner, an electrical RC circuit can be used as its analog. The first-order model does not necessarily provide an accurate description of the dynamics of the gel. It is meant only to assign a time constant or scale to the speed of the mechanical system or transducer in order to examine the effects of switching with different frequencies and different input dynamics.

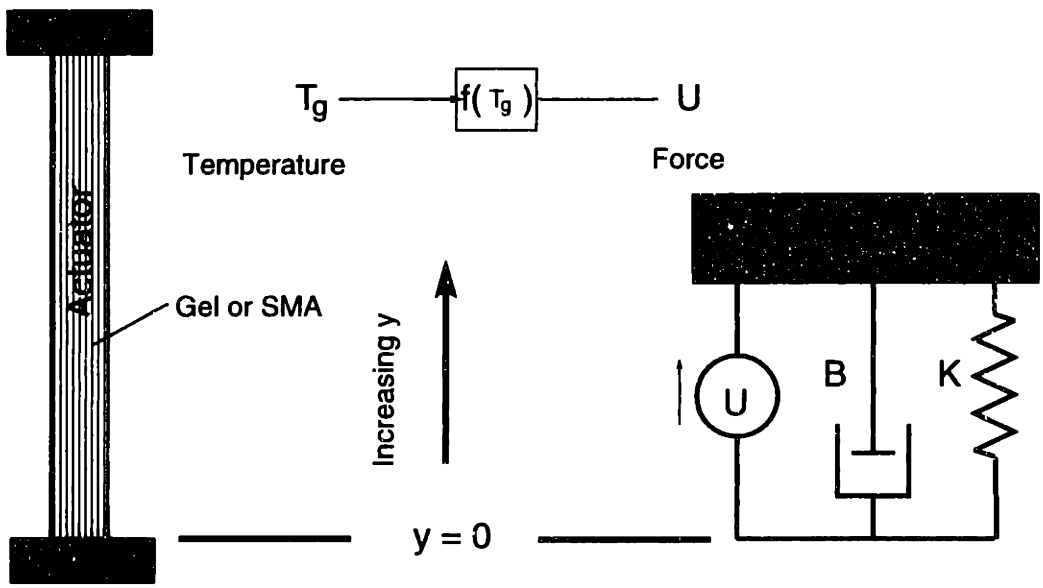


Figure 4.5: System variables as defined for gel control experiment

4.3.3 Modeling the Static Behavior of the Actuator

To model the discontinuous, hysteretic, *steady-state* or equilibrium behavior of a polymer gel, consider the transfer characteristic shown in Figure 4.6. A Schmitt trigger is an analogous electrical circuit that has this behavior. The output levels Y_h and Y_l, are the positions of the gel[7] in the model in Figure 4.5 corresponding to the fully collapsed and fully expanded gel equilibrium states, respectively. The temperatures T_1 and T_2 are the boundaries of the hysteresis band. A collapsed gel needs to be cooled to a temperature T_1 before it starts expanding, and a swollen gel needs to be heated to T_2 before it starts shrinking. This functional block can be used to connect the thermal and mechanical models discussed earlier. Assuming that the DC gain of the transducer system, which relates the gel temperature to the gel length, is equal to unity, the hysteretic block captures all of the equilibrium behavior of the mechanical system in response to temperature. (Any non-unity gain in the transducer model can be captured in this block by scaling the levels of the two rails in the hysteretic block appropriately.)

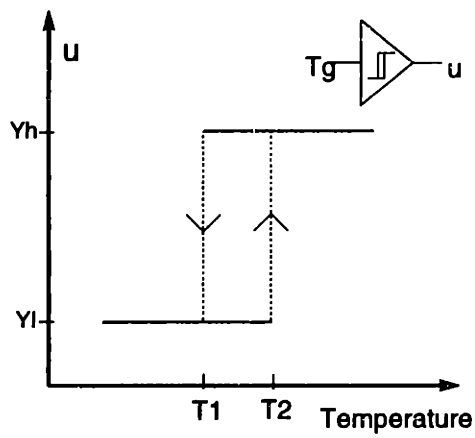


Figure 4.6: A Schmitt Trigger model for gel temperature induced discontinuous volume-phase transition

[7]Corresponding to a length of the gel, or a position of the suspended end in the gel.

4.3.4 Overall System Model

Connecting the input (thermal) system from Figure 4.4 to the mechanical system in Figure 4.5 with the Schmitt trigger in the previous section produces the model shown in Figure 4.7. The mechanical and thermal components are represented by electrical equivalents. The single-pole double-throw switch in the input system selects one of two inputs, or temperature levels, into the system. Note that the thermal system losses are ignored in this model. The diode switch in the transducer stage is turned on when the output of the Schmitt trigger is higher than the voltage on the capacitor. This results in different mechanical dynamics by combining R_d in parallel with R_M. This behavior is observed in several systems of interest. In a polymer gel, for instance, the dynamics of shrinking and expanding may be different. Note that the DC gain of the mechanical stage is unity. The equilibrium function describing the relationship between T_g and y is therefore the same as in Figure 4.6.

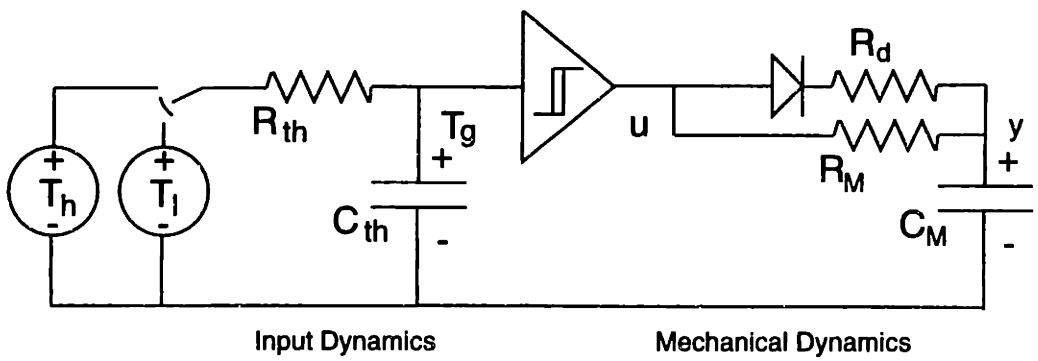


Figure 4.7: Complete switched mechanical system model equivalent

Assume that the input system has the same first-order dynamics regardless of the switch states. When T_H is connected to the system:

$$\dot{T}_g(t) = \boldsymbol{A}_I T_g(t) + \boldsymbol{B}_I T_h. \tag{4.25}$$

When T_L is connected:

$$\dot{T}_g(t) = \boldsymbol{A}_I T_g(t) + \boldsymbol{B}_I T_l. \tag{4.26}$$

The system matrices in this case are:

$$\boldsymbol{A}_i = -\frac{1}{R_{th}C_{th}} \qquad and \qquad \boldsymbol{B}_i = \frac{1}{R_{th}C_{th}}. \tag{4.27}$$

The mechanical system model depends on the gel temperature. When $u = Y_h$:

$$\dot{y}(t) = \boldsymbol{A}_1 y(t) + \boldsymbol{B}_1 u(t), \tag{4.28}$$

where

$$\boldsymbol{A}_1 = -\frac{1}{R_p C_M} \qquad and \qquad \boldsymbol{B}_1 = \frac{1}{R_p C_M} \tag{4.29}$$

The quantity R_p is the parallel combinations of resistors R_d and R_M :

$$R_p = \frac{R_M R_d}{R_M + R_d} \tag{4.30}$$

When $u = Y_l$:

$$\dot{y}(t) = \boldsymbol{A}_2 y(t) + \boldsymbol{B}_2 u(t) \tag{4.31}$$

where

$$\boldsymbol{A}_2 = -\frac{1}{R_M C_M} \quad and \quad \boldsymbol{B}_2 = \frac{1}{R_M C_M}. \tag{4.32}$$

Note that while most of the discussion will focus on first-order systems to simplify the development of intuition, the state-space models above can assume any order. In the sections that follow, assumptions about the time-scale separation between the time constants of different system components and the switch period will be used to derive averaged models. The results will hold for higher order models if the time-scale separation properties are the same for all of the natural frequencies of the component systems. For instance, we can assume that the switch frequency is much higher than the natural frequency of the RC circuit representing the input thermal system and derive some results based on this assumption. The same results will hold for a more complicated input model if it is true that all of its natural frequencies are much lower than the switch frequency.

4.4 Performance and time-scales in Discontinuous Systems

The circuit in Figure 4.7 is controlled by changing the state of the switches. The switches undergo a full cycle (changing states twice to get back to the original state) once every switch period T. This switching frequency and the dynamics of the mechanical and input systems produce the behavior of the switched actuator. Specifically, the behavior of the switched actuator depends on the frequency of switching f_S relative to the natural frequencies of the mechanical system f_M and the input or thermal system f_I[8]. The sections below examine the behavior of the switched, discontinuous actuator under different arrangements of these frequencies. In particular, the following three[9] cases are considered:

- First, systems with a **low switch frequency** compared to the natural frequencies of the mechanical system are considered. The output ripple in this case is large and the actuator cannot be used to maintain a load at a specified reference.

- A **slow input system** that cannot produce a temperature change on the order of the hysteresis gap in the gel phase-transition curve will also result in large ripple in the output of the discontinuous actuator. For a discontinuous actuator, the input system response time as well as the switch period must be short compared to the time constants of the transducer.

- With a **slow mechanical system**, the gel position can be regulated to an average value with a small amount of switch ripple. However, different techniques are needed to derive averaged models for the system depending on the relative magnitudes of the input and switch frequencies.

[8]Throughout this discussion, these variables (f_S, f_I, and f_M) will be used to compare the rate of response of the different subsystems relative to the switching frequency. A large value for either f_I or f_M compared to f_S indicates that the dynamics of the input or mechanical system, respectively, will completely settle within the time window of a switch period. A small valued f_I or f_M indicates insignificant response for either system during the time window of a switch period. In the s-plane, the dominant pole of the stable, mechanical system is $-f_M$, and the pole of the input system is $-f_I$. We will compare these pole locations to the quantity $-f_S$ to explore the time-scale properties of the overall system.

[9]The three cases are further divided to more specific instances. For example, the case where the switching frequency is higher than both the mechanical and the input frequencies is covered partly in the second item and partly in the third. All possible relevant cases are therefore covered in the three sections.

4.4.1 Low Switch Frequency

As discussed in Section 4.1.1, time-scale separation between the switch frequency and the switched system dynamics facilitates the averaging of system matrices. With this assumption, the state variables and inputs can be approximated by constant values and factored out of the averaging integral. Another practical motivation for this time-scale separation requirement is ripple reduction. Without significant separation between the time constants and the switch period, the ripple in the mechanical system variables will be large, and regulation close to a reference may not be possible.

To illustrate the difficulty of developing a useful average model in the case of low switching frequency, consider the system pole locations in Figure 4.8. Since the input dynamics

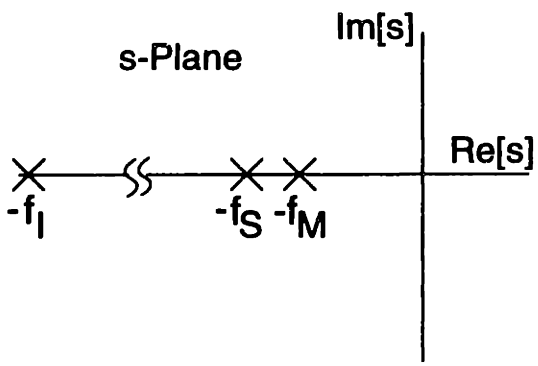


Figure 4.8: Switching and natural frequencies of system with time-scale separation between mechanical and input frequencies. The switching frequency is selected to be much higher than the input frequency but much lower than the mechanical frequency.

are fast compared to the switch period, the output of the input system, T_g, will settle to its final value before the switch states are changed. On the time-scale of the switch frequency, the waveform $T_g(t)$ will be approximately a square wave that is synchronized with the input switching function. This is shown in Figure 4.9. Notice in particular that the duty ratio D is approximately the same for both the driving switch function and the output of the first stage T_g, which is also the input to the mechanical system[10]. The fast dynamics of

[10]Unless it is required to make a specific point, the time index will be suppressed from the function $D(t)$

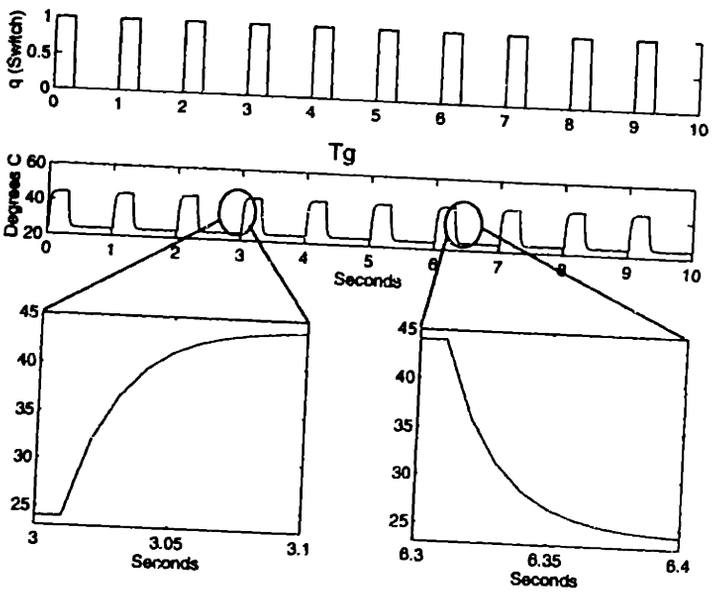


Figure 4.9: Switching and T_g waveforms when switch frequency is lower than input natural frequency. Switch frequency is 1 Hertz and input natural frequency is 50 Hertz.

the input system can therefore be ignored, and the overall switched system will behave as if directly driven by the input waveform directly. In every switch period, for DT seconds, the overall system looks like:

$$\dot{y}(t) = \boldsymbol{A}_1 y(t) + \boldsymbol{B}_1 Y_h, \tag{4.33}$$

and for $(1-D)T$ seconds:

$$\dot{y}(t) = \boldsymbol{A}_2 y(t) + \boldsymbol{B}_2 Y_l, \tag{4.34}$$

where the system matrices and input values have been defined earlier.

A simulation of a switched system with a switch frequency close to the natural frequency of the mechanical system is shown in Figure 4.10. In this figure, the time constants of the two mechanical systems being switched are 0.1 and 0.2 seconds, with $Y_h = 10$ and $Y_l = 2$, respectively. In the lower graph, the full, switched response of the output is shown along

to simplify the notation. A constant value D will be used instead.

with two "averaged" responses. The dashed response represents the true, computed average of the output response. The solid line is the average response as predicted by an averaged model derived according to Equation 4.11 using the duty ratio D of the input switching function. The averaged model is not accurate since there is insufficient time-scale separation between the state variable dynamics and the switch period. Simulation of the average output based on this model is clearly not a good match for the average of the actual output.

If the input system has slower dynamics it may filter the switch ripple before it reaches the mechanical system. However, with the discontinuous equilibrium characteristic, the result is still a square wave of input to the transducer system at a frequency equal to or lower than the switch frequency. (In the continuous system discussed in the next section the slower input dynamics can help). The resulting large magnitude of ripple in the output makes this system useless for regulating the output close to a reference, as illustrated in the next section.

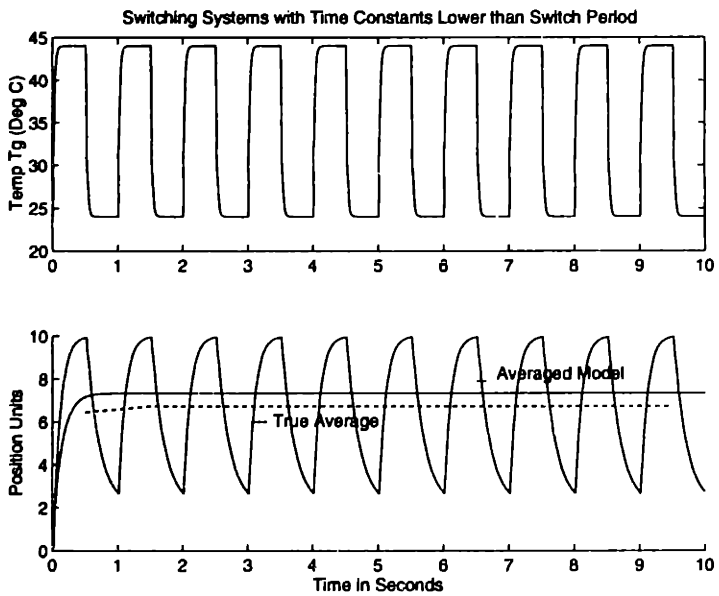


Figure 4.10: Simulation of System with a Low Switching Frequency

4.4.2 Slow Input System

When the input system response time is close to that of the mechanical system or longer, the ripple in the output may be large even when the switch frequency is relatively high. A slow thermal system will take a long time to cross the hysteresis gap in order to change the state trajectories of the mechanical system which are deviating from their reference. By the time the thermal system reaches the opposite side of the hysteresis gap, the faster mechanical system will have traveled far from the reference. If we are primarily concerned with controlling the mechanical system, it will be extremely difficult to achieve good regulation with an input system that changes more slowly than the rest of the system. As an example of this behavior, the simulation in Figure 4.11 shows a discontinuous transducer driven with a switching profile of constant frequency and duty ratio of 0.7 and 0.3 alternating every 20 switch cycles. The output of the Schmitt trigger is either 10 or 2 and the hysteresis gap is between 32°C and 36°C. The temperature drive rails are 44°C and 24°C, and the time constant of the temperature dynamics is 15 seconds. The top trace displays the input switching function, and the middle trace shows the slow input dynamics filtering out the ripple. This in general will limit the switching frequency seen by the transducer to around the natural frequency of the input system. If this natural frequency is close to that of the transducer natural frequencies, the ripple in the output will be large. The case where the input system is slow compared to the switch period and the output ripple is still small because the mechanical dynamics are even slower is addressed as part of the next section.

4.4.3 Slow Mechanical System

Finally, consider the case where the input system is significantly faster than the mechanical system. This is often a requirement for good system performance. It is therefore a reasonable physical assumption in many cases. For instance, as shown above, when the input system is at the same speed or slower than the mechanical system, the mechanical system ripple will be large. Assume also that the switching frequency is much higher than the natural frequencies of the mechanical system. In this section, the input system dynamics are assumed to be significantly faster than the transducer system dynamics.

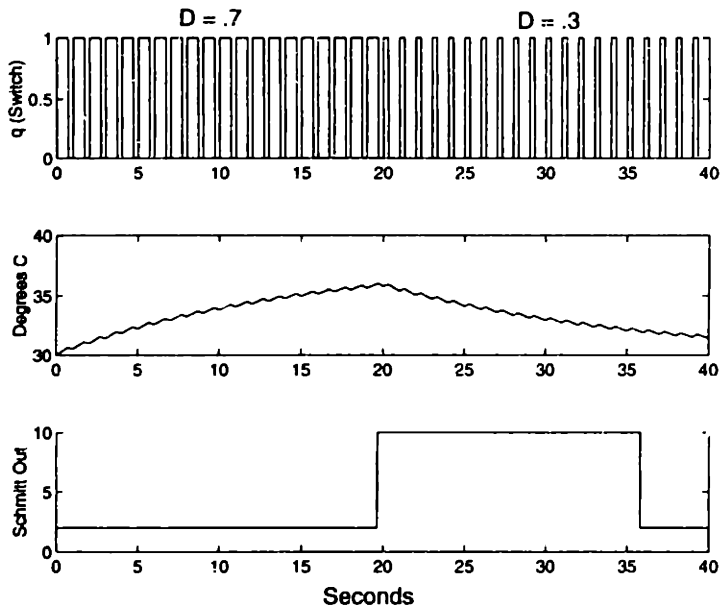


Figure 4.11: Simulation of discontinuous system with slow input dynamics

Switching Frequency Lower than Input Natural Frequency

Figure 4.12 shows widely separated natural frequencies for the input and mechanical system components. With wide enough separation, it is possible to pick a switching frequency that is much higher than the mechanical system frequency and much lower than the input system frequency, as shown in the figure. As discussed in the previous section and illustrated in Figure 4.9, the input dynamics can be ignored in this case. However, unlike in the previous section, the resulting reduced system is a slow mechanical system that exhibits small ripple for the given switching frequency. An averaged model can therefore be derived for this system since the actual output is well approximated by its average over one switch cycle. The switched system is still modeled by Equations 4.33 and 4.34. The averaged system is:

$$\dot{\bar{y}}(t) = \boldsymbol{A}\bar{y}(t) + \boldsymbol{B}, \tag{4.35}$$

where

$$\boldsymbol{A} = D\boldsymbol{A}_1 + (1-D)\boldsymbol{A}_2 \text{ and } \boldsymbol{B} = D\boldsymbol{B}_1 Y_h + (1-D)\boldsymbol{B}_2 Y_l. \tag{4.36}$$

Figure 4.13 shows a simulation of the full switched system and its averaged model. The top

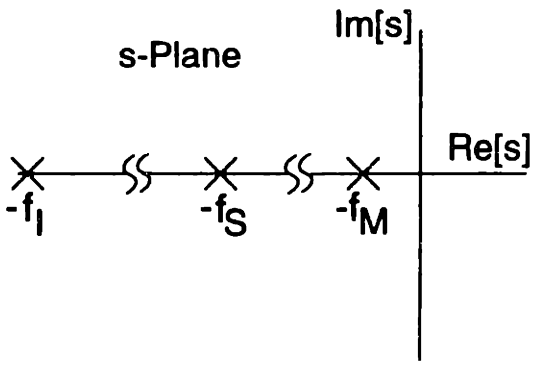


Figure 4.12: Switching and natural frequencies of system without time-scale separation between mechanical and switching frequencies

trace illustrates the full switched simulation of the output and the prediction of the average using the averaged model. The temperature profile is a square wave on the time-scale of the transducer dynamics, as shown in the bottom left plot in Figure 4.13. A closer look at the output waveform shows good prediction of the true, computed average of the average position using the averaged model.

Switching Frequency close to Input Natural Frequency

In some cases, the input and mechanical systems may not exhibit sufficient time-scale separation to allow a switch frequency that both effectively eliminates the dynamics of the input system and achieves the desired ripple constraints for the mechanical variables. Figure 4.14 illustrates the case where the input natural frequency and the switching frequency are close. When this happens, the quantity T_g is not allowed to settle before the switch state is changed. The temperature dynamics cannot be ignored in this case, since T_g does *not* look like a square wave on the switch period time-scale. However, because of the Schmitt trigger that follows the input stage, the mechanical system still sees a square wave. This square wave is not the same as the switch signal, with duty ratio D. It is generated by the exponential waveform T_g as it passes through the hysteresis band of the Schmitt trigger.

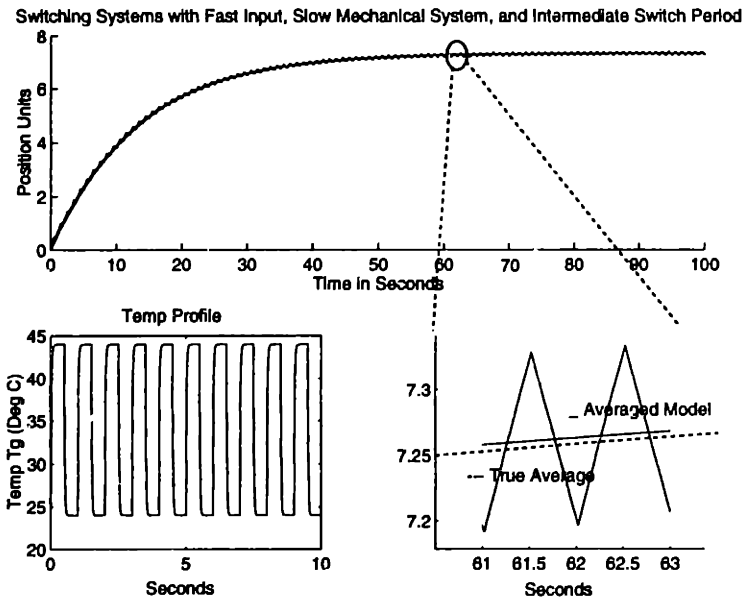


Figure 4.13: Simulation of system with switch period between the mechanical and the input time constants

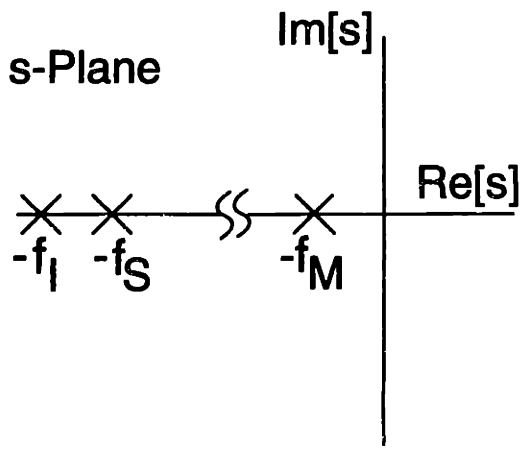


Figure 4.14: Switching and natural frequency of input system not significantly separated

Figure 4.15 shows how the input to the mechanical system is produced.

Note that the dynamics of the switched system are still described by Equations 4.33 and 4.34. Now, there is still sufficient time-scale separation between the switching period and the mechanical time constant to ensure low output ripple amplitude. An averaged model can be derived for this system. However, the averaging will need to be done using D_1, the duty ratio of the square wave driving the mechanical system, and not D. The averaged system in this case is:

$$\dot{\bar{y}}(t) = \boldsymbol{A}\bar{y}(t) + \boldsymbol{B} \tag{4.37}$$

where

$$\boldsymbol{A} = D_1\boldsymbol{A}_1 + (1 - D_1)\boldsymbol{A}_2 \ \textit{and} \ \boldsymbol{B} = D_1\boldsymbol{B}_1Y_h + (1 - D_1)\boldsymbol{B}_2Y_l \tag{4.38}$$

Figures 4.16 and 4.17 show simulations of a system with a switch frequency on the order of the input dynamics. In Figure 4.16 shows the error between the true average dynamics and those predicted by an averaged model based on the duty ratio D. The temperature profile is *not* a square wave and therefore does not reproduce the same duty ratio at the input to the transducer. Figure 4.17 illustrates an improved averaged model based on D_1, the duty ratio of the switching function (square wave) seen by the transducer. The effective duty ratio D_1 is a nonlinear function of D in the same cycle, the value of T_g at the beginning of the cycle, and the Schmitt trigger parameters. Appendices C and D describe control implementation using a microprocessor that exploits knowledge of this nonlinear mapping.

High Switching Frequency

When very high switch frequencies are possible, the ripple in the input system state (the temperature T_g in this case) can be made very small. This case is illustrated in Figure 4.18. The input natural frequency is still assumed much higher than the mechanical system natural frequency. A fixed duty ratio will always result in a temperature that is rippling around an average value. With the switch period close in duration to the time constant of the temperature dynamics, the ripple within a period is large enough to intersect the ends of the hysteresis gap of the Schmitt trigger at the input of the mechanical system

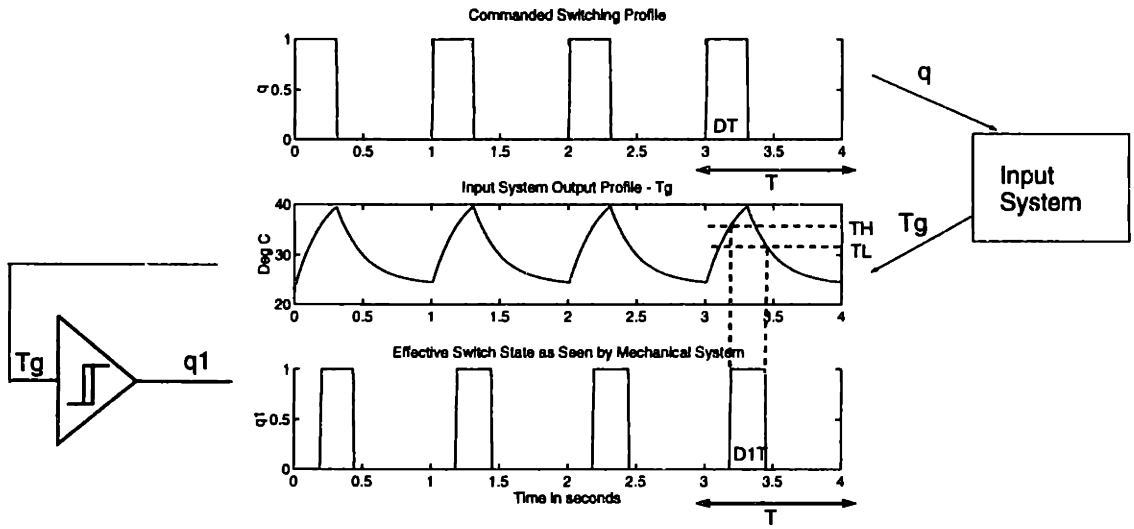


Figure 4.15: Converting the exponential profile of T_g into a square wave

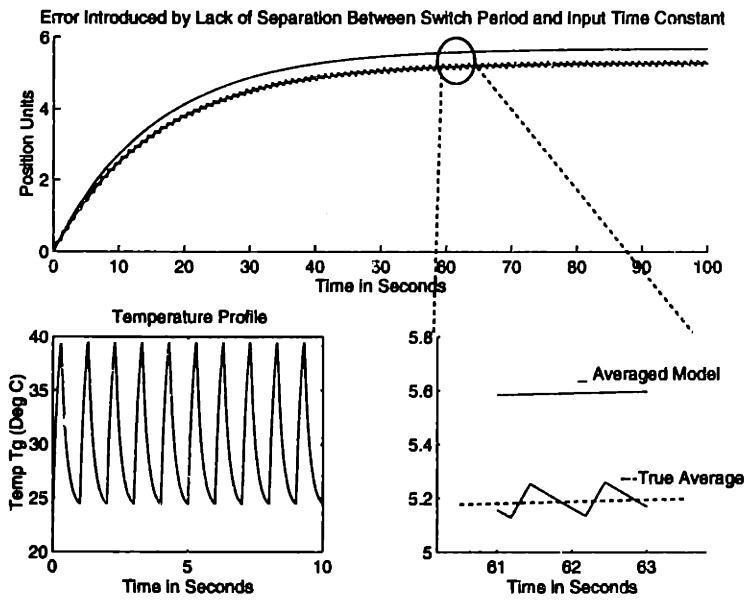


Figure 4.16: Simulation of system with switch period close to input time constant. Average system derived using D.

within one switch cycle. A shorter switch cycle reduces the ripple to a size smaller than the hysteresis gap. A constant duty ratio will keep the temperature, and the mechanical system, on one side of the switch. Figure 4.19 shows a simulation with a switch frequency of 1Hz, a temperature natural frequency of (1/3)Hz, and mechanical natural frequencies of (1/30)Hz and (1/60)Hz. The duty ratio is switched from $D = 0.7$ to $D = 0.3$ every ten switch cycles. The mechanical system will not be switched between its two states in one switch cycle. However, as shown in the figure, the transducer system state is switched with another switch function $q_1(t)$. Average models can be developed based on the conversion from the commanded duty ratio to the effective duty ratio D_1 as discussed above. (The effective duty ratio for the mechanical system within a switch cycle will either be 0 or 1 most of the time. The duty ratio for the input system is still D. This duty ratio D can be used to estimate the average T_g fed into the Schmitt trigger.) However, the large swings in the duty ratio will produce a system not well modeled by an LTI model over a switch cycle if the the systems alternately selected have different dynamics. Since the switch frequency is usually user-specified, one way to arrive at an averaged model is to reduce this rate to a range where the results of the previous two sections can be applied. Usually, there are upper limits on frequency but not lower limits. When this is not possible, more elaborate schemes need to be developed. The effective switching frequency of the mechanical system, which is limited by the input system dynamics, can be estimated along with its duty ratio (using some form of observer or real-time estimator) from the input switch profile and the various system parameters. Averaged models for the transducer systems can then be obtained by averaging over this new switch period using the observed duty ratios. (See Appendix C.)

4.4.4 Identical Matrices A_1 and A_2

The discussion above has assumed that the matrices A_1 and A_2 are different. Sometimes changing the state of the switch changes only the input power level and not the dynamics of the system. For the system in Figure 4.7, this is equivalent to replacing the resistor R_d with an open circuit. The average of the input terms (i.e. $B_1 Y_h$ and $B_2 Y_l$) still needs to be computed as weighted by D or D_1. The averages of the states and outputs of the system

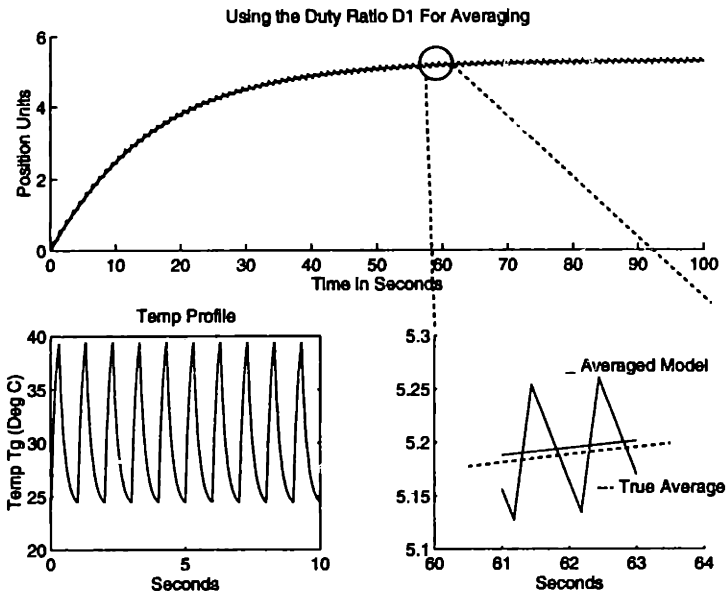


Figure 4.17: Simulation of system with switch period close to input time constant. Average system derived using D_1.

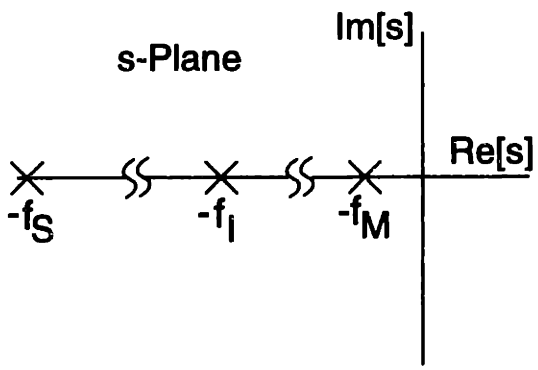


Figure 4.18: Switching frequency high above both input and mechanical system frequencies

are described by the response of the system to this averaged input. The average matrix ($\boldsymbol{A}$) can be computed similar to the computation of $\boldsymbol{A}$ in Section 4.1.2. The following equation describes this situation:

$$\boldsymbol{A}_1 = \boldsymbol{A}_2 = \boldsymbol{A}. \tag{4.39}$$

We can therefore simplify the averaged expression :

$$\overline{q(t)\boldsymbol{A}_1 x(t)} + \overline{q'(t)\boldsymbol{A}_2 x(t)} = \overline{(q(t) + q'(t))\boldsymbol{A}x(t)}, \tag{4.40}$$

and since

$$q(t) + q'(t) = 1, \tag{4.41}$$

we can simplify further

$$\overline{(q(t) + q'(t))\boldsymbol{A}x(t)} = \boldsymbol{A}\bar{x}(t). \tag{4.42}$$

The overall system can then be described by

$$\dot{\bar{\boldsymbol{x}}}(t) = \boldsymbol{A}\bar{\boldsymbol{x}}(t) + \boldsymbol{B} \tag{4.43}$$

and

$$\boldsymbol{B} = \boldsymbol{B}_1 Y_h D + \boldsymbol{B}_2 Y_l (1 - D). \tag{4.44}$$

4.5 Performance and Time-Scales in Continuous Systems

The previous section examined the properties of switched systems where the transducer equilibrium output is a discontinuous function of the input variable. This model describes the behavior of many physical systems, including a variety of polymer gel actuators. Other systems demonstrate a continuous equilibrium response. Some polymer gels have volumes that are continuous functions of temperature [90] [98]. A DC motor, for instance, produces a torque linearly related to its input current [86] [12]. Continuous systems can be analyzed similarly by changing the equilibrium function in the general model. As in the case with the discontinuous system, the dynamics of the mechanical system are captured in the first-order model. Only the static behavior is captured in the shape of the equilibrium block.

The functional block illustrated in Figure 4.20 is used to model the equilibrium behavior of a continuous transition gel or other mechanical system. In the linear region, an increase in the input level will result in a proportional increase in the output level in steady state. (This linear function can represent a linear system or the approximation of a nonlinear system near an operating point.) The saturation levels are characteristic of many physical systems. The volume of a polymer gel will decrease with temperature only until it reaches the totally collapsed state. Saturation of the magnetic core of a motor will result in similar limiting behavior. Note also that no hysteresis is featured in this model for equilibrium states.

Another system characteristic that could change whenever the switch states are altered is the DC gain of the system. The block in Figure 4.20 does not capture this behavior. The DC gain is the slope of the equilibrium transfer function and it does not change with the state of the diode. Different DC gains can be assigned to the different states of the system without any loss of generality. The gain is a function of the system parameters in the system matrices $\boldsymbol{A}$ and $\boldsymbol{B}$. If an averaged model can be derived using the mathematical techniques discussed below, the overall DC gain will be the one associated with this averaged model. The assumption of a constant DC gain is often a good one and is physically motivated. Without hysteresis or other non-decaying memory elements (such as an integrator or unstable mode) a system is expected to settle to the same equilibrium point with a given excitation, regardless of its direction of approach to that point.

Combining the function in Figure 4.20 with the electrical analogs of the mechanical system and the input system in Figures 4.4 and 4.5, the model in Figure 4.21 is derived. The input equations are the same as in the discontinuous case, and are reproduced here for convenience. This structure may, for instance, represent a simplified model for a continuous volume transition gel. When the input is at T_H:

$$\dot{T}_g(t) = \boldsymbol{A}_I T_g(t) + \boldsymbol{B}_I T_H. \tag{4.45}$$

When T_L is the driving source:

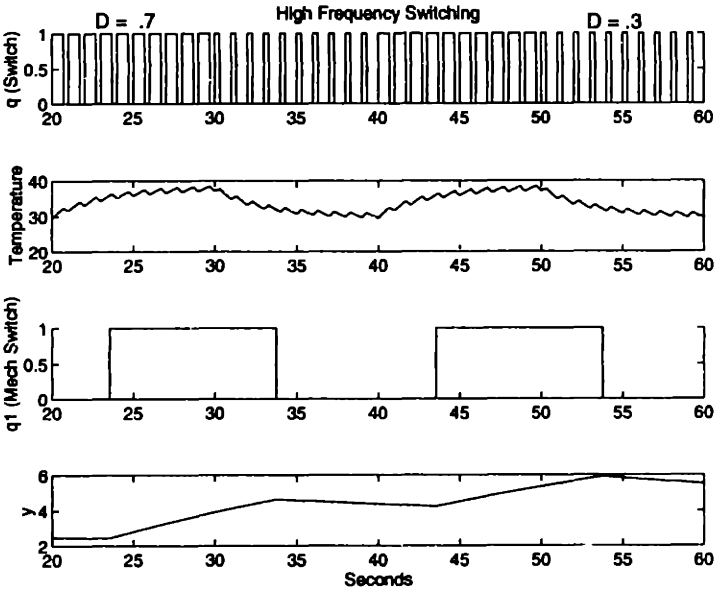


Figure 4.19: Simulation of system with alternating duty ratio. Switching frequency high above both input and mechanical system frequencies

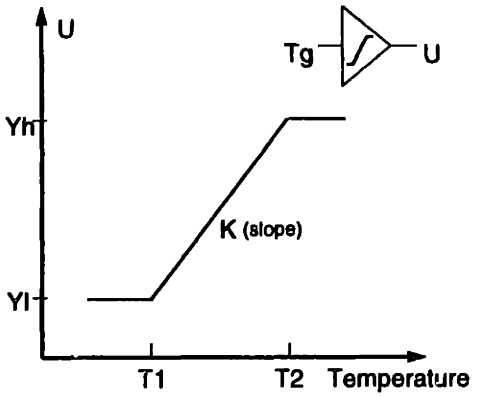


Figure 4.20: Continuous linear relationship with saturation

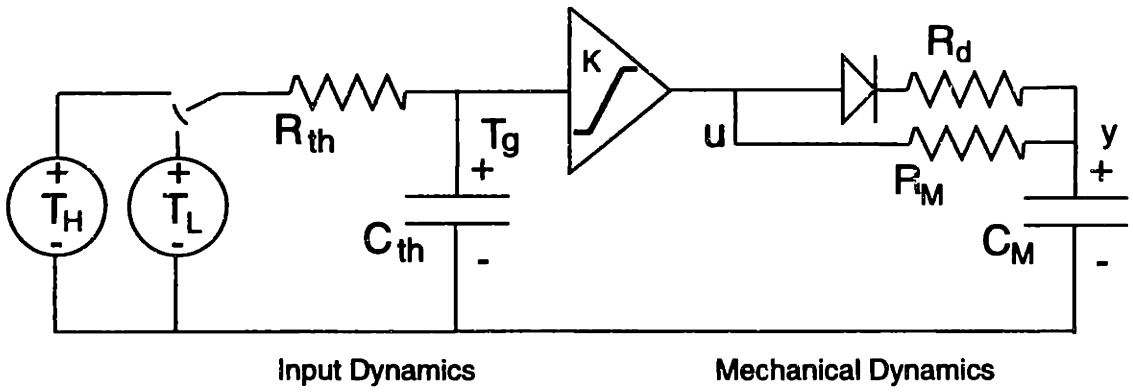


Figure 4.21: Overall continuous system

$$\dot{T}_g(t) = \boldsymbol{A}_I T_g(t) + \boldsymbol{B}_I T_L. \tag{4.46}$$

The system matrices in this case are:

$$\boldsymbol{A}_I = -\frac{1}{R_{th}C_{th}} \quad \textit{and} \quad \boldsymbol{B}_I = \frac{1}{R_{th}C_{th}} \tag{4.47}$$

The mechanical system has the same time constants as before. However, the switching here is different. While the state of the controllable switch determines which power source is connected to the input system, the mechanical system dynamics are determined by the diode state. In the earlier section with the discontinuous system, the diode was always on if the Schmitt trigger output was high, and always off when it was low. The ideal diode in the circuit in Figure 4.21 is on when

$$u(t) > y(t) \tag{4.48}$$

In this state, the mechanical system dynamics are described by:

$$\dot{y}(t) = \boldsymbol{A}_1 y(t) + \boldsymbol{B}_1 u(t) \tag{4.49}$$

where

$$A_1 = -\frac{1}{R_p C_M}, \qquad B_1 = \frac{1}{R_p C_M}, \tag{4.50}$$

and

$$R_p = \frac{R_M R_d}{R_M + R_d}. \tag{4.51}$$

When the diode is off:

$$u(t) \leq y(t), \tag{4.52}$$

and the mechanical system is described by:

$$\dot{y}(t) = A_2 y(t) + B_2 u(t) \tag{4.53}$$

where

$$A_2 = -\frac{1}{R_M C_M} \quad \textit{and} \quad B_2 = \frac{1}{R_M C_M} \tag{4.54}$$

The input to the mechanical system $u(t)$ is the output of the gain block. In the linear region of the equilibrium block:

$$u(t) = K T_g(t) \tag{4.55}$$

For the following sections, we will ignore the saturation bands to simplify the discussion. Saturation can be avoided in implementation by insuring the temperature drives are in the range between T_1 and T_2. Different time-scale assignments among the system components are explored. Specifically, the following interesting cases are considered:

- The system can be controlled using a **high switching frequency** compared to the natural frequencies of both the input and mechanical systems[11]. Recall that in the case of the discontinuous actuator, the transducer still had a square wave at its input at a frequency largely dependent on the speed of the input system. However, the transducer in the continuous system has a smooth waveform at its input. Regardless of the speed of the input system relative to the mechanical dynamics, the ripple in the output variable will be small. However, different simplifying assumptions can be made

[11] Any reference to the speed of the transducer or input systems or the switch frequency is in relation to the other two. A high transducer natural frequency indicates that it is much higher than the input natural frequency and the switch frequency, for example.

to facilitate the derivation of averaged models depending on the relative magnitudes of the thermal and mechanical time constants.

- With a **fast input system**, the transducer will have large switch ripple at its input. The output will have small ripple if the dynamics of the mechanical system are much slower than the switch frequency. However, the averaging problem is different depending on the separation between the switch frequency and the thermal system natural frequency.

- Unlike a discontinuous actuator, a continuous system with **fast transducer dynamics** can still produce small ripple in the output variable. A slow input system could filter out the switch ripple. The derivation of averaged models for an actuator with a fast transducer stage and a slow input stage is discussed.

Before the specific cases are discussed, an augmented state-space model is derived for the switched continuous actuator.

4.5.1 Augmented State-Space Model for Switched System

Ignoring the saturation effects in the equilibrium block, the overall system can be modeled as a linear one realized by the series connection of the input system, a gain block, and the transducer system. The controllable switch selects the input system configuration and the diode state determines the topology of the transducer. There are four possible switch states.

- With the input switch connecting the system to T_H and the diode turned on:

$$\begin{bmatrix} \dot{y} \\ \dot{T}_g \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{B}_1 \\ 0 & \boldsymbol{A}_I \end{bmatrix} \begin{bmatrix} y \\ T_g \end{bmatrix} + \begin{bmatrix} 0 \\ \boldsymbol{B}_I \end{bmatrix} T_H \tag{4.56}$$

- With the input switch connecting the system to T_H and the diode turned off:

$$\begin{bmatrix} \dot{y} \\ \dot{T}_g \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_2 & \boldsymbol{B}_2 \\ 0 & \boldsymbol{A}_I \end{bmatrix} \begin{bmatrix} y \\ T_g \end{bmatrix} + \begin{bmatrix} 0 \\ \boldsymbol{B}_I \end{bmatrix} T_H \tag{4.57}$$

- With the input switch connecting the system to T_L and the diode on:

$$\begin{bmatrix} \dot{y} \\ \dot{T}_g \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{B}_1 \\ 0 & \boldsymbol{A}_I \end{bmatrix} \begin{bmatrix} y \\ T_g \end{bmatrix} + \begin{bmatrix} 0 \\ \boldsymbol{B}_I \end{bmatrix} T_L \tag{4.58}$$

- Finally, with the input switch connecting the system to T_L and the diode off:

$$\begin{bmatrix} \dot{y} \\ \dot{T}_g \end{bmatrix} = \begin{bmatrix} \boldsymbol{A}_2 & \boldsymbol{B}_2 \\ 0 & \boldsymbol{A}_I \end{bmatrix} \begin{bmatrix} y \\ T_g \end{bmatrix} + \begin{bmatrix} 0 \\ \boldsymbol{B}_I \end{bmatrix} T_L \tag{4.59}$$

4.5.2 Switching Frequency Higher than Both Input and Mechanical Natural Frequencies

The case where the switching frequency is much higher than both the mechanical and input natural frequencies is illustrated in Figure 4.22. The qualitative behavior of the switched system in this case does not vary with the relative positions of the mechanical and the input time constants. However, the ripple size and rate of response will be dominated by the slower of the two. Note that in the four different state-space descriptions in the previous sections, the input is a constant term (T_L or T_H). This is similar to the special case in Section 4.1.2. Even though the input to the overall system is changing significantly over a switch period, it is held at a single value during each switch state. It can therefore be considered a constant multiplier of the input ($\boldsymbol{B}$) matrix for that particular state. This, along with the assumption of small ripple in the state variables (T_g and y in this augmented formulation), allows the use of the averaging result in that section. The averages of the switch functions can be used to compute a weighted average of the different state matrices. The result is an averaged model of the system.

The two switches in the system above result in two different switch functions and two duty ratios. Define the switch profile of the controllable switch as $q(t)$ and the switch profile of the diode switch as $q_1(t)$. In the model above, the input system (represented by the lower rows of the matrices) is in the state where T_H is connected to the system when $q(t)$ is high and is driven by T_L when $q(t)$ is low. The transducer system (upper row of system matrices) is switched on $q_1(t)$. Defining averaged functions $D(t)$ and $D_1(t)$ for $q(t)$

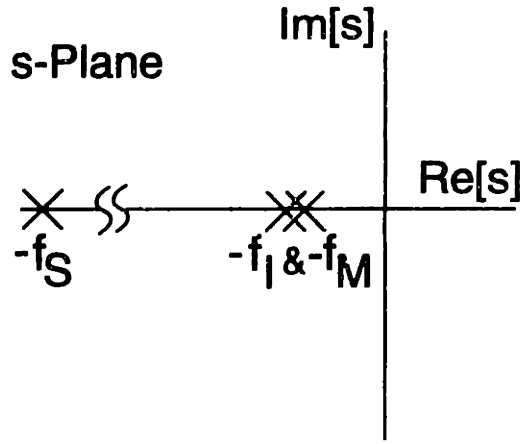


Figure 4.22: Pole locations when switching frequency is much higher than input and transducer frequencies

and $q_1(t)$, respectively, and suppressing the time index, the following averaged model can be derived:

$$\boldsymbol{A} = \begin{bmatrix} D_1\boldsymbol{I} & 0 \\ 0 & D\boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{B}_1 \\ 0 & \boldsymbol{A}_I \end{bmatrix} + \begin{bmatrix} (1-D_1)\boldsymbol{I} & 0 \\ 0 & (1-D)\boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_2 & \boldsymbol{B}_2 \\ 0 & \boldsymbol{A}_I \end{bmatrix} \tag{4.60}$$

The matrix $\boldsymbol{I}$ is the identity matrix. In the first-order example it is simply 1. For the average $\boldsymbol{B}$ matrix:

$$\boldsymbol{B} = D \begin{bmatrix} 0 \\ \boldsymbol{B_I} T_H \end{bmatrix} + (1-D) \begin{bmatrix} 0 \\ \boldsymbol{B_I} T_L \end{bmatrix}. \tag{4.61}$$

The average model is therefore:

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\bar{T}}_g \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} \bar{y} \\ \bar{T}_g \end{bmatrix} + \boldsymbol{B}. \tag{4.62}$$

The simulation in Figure 4.23 illustrates an application of this averaged model. The top trace is a switching function $q(t)$ with a constant $D = 0.5$. The variable T_g is the temperature waveform. Notice the low amplitude ripple in the temperature since the switching frequency is higher than the natural frequency of the input dynamics. The input dynamics

are first-order with time constant $\tau = 5$, and the switch period is 1. The diode state is displayed in the third curve, as the function $q_1(t)$ with duty ratio $D_1(t)$. In equilibrium, the diode is switching with the same period of 1 second with a duty ratio that is a function of the two switched transducer systems as well as D. The two transducer systems here have characteristic time constants of 4 and 8 seconds. Using $D(t)$ and $D_1(t)$, an averaged model is derived and used to predict the response of the average of $y(t)$. The results are shown in the bottom graph in Figure 4.23, contrasted with the actual (true average of the full switched simulation) response. The averaged model is accurate because the input and state ripple is small.

4.5.3 Fast Input System

Frequently, the input dynamics are not as slow as the transducer dynamics. This situation is often desirable as it allows for the elimination of the input dynamics and the reduction of the system model order. One example of a fast input stage is the electrical dynamics of a DC motor. The mechanical dynamics of the moving parts and load may be slower than the electrical dynamics [105]. In the previous section we explored the case where the switching frequency was much higher than the input and transducer natural frequencies. In this section, the cases where the input time constants are as short as or shorter than the switch period are explored.

Switch Frequency Lower than Input Natural Frequency

The situation when the switch frequency is lower than the input natural frequency is the same one illustrated earlier in Figure 4.12. With enough of a gap between the mechanical and input time constants, a switch period can be selected to keep the ripple in the mechanical system low while at the same time allowing the input system to settle every time the switch state is altered. Direct application of the result in Equations 4.60 - 4.62 will not produce an accurate model. It is not valid to average the input system since its state variable has significant harmonic content at the switch frequency. To illustrate this point further, the model in Equations 4.60 - 4.62 is derived for a continuous system with the time-scale

properties in 4.12. The full switched system is simulated, as is the averaged model. The results are shown in Figure 4.24. Note that the mechanical switch function q_1 is the same as the input switch function q. However, as the bottom trace demonstrates, an averaged model based on D (which is equivalent to D_1 in this case) predicts the wrong average for the mechanical variable $y(t)$.

One solution in this particular case exploits the fact that the variable T_g is essentially a square wave on the time-scale of the switch period. This square wave has the same shape and phase as that of the switch. We can reduce the order of the model in Equations 4.60 - 4.62 by ignoring the dynamics of the input. Since the input states settle much faster than the mechanical states and before the switch state is changed, we can replace the input system by a wire that conducts either T_H or T_L to the gain block. Denote the output of the gain block corresponding to an input T_H by U_h and the output due to T_L by U_l. The output of the gain block on the time-scale of the mechanical system will be a square wave of duty ratio D (that of the controllable switch) oscillating between U_h and U_l. By using the simplification introduced in Section 4.1.2 and averaging the mechanical dynamics over a switch period with duty ratio D, the following averaged model can be derived for the total system:

$$\dot{\bar{y}}(t) = \boldsymbol{A}\bar{y}(t) + \boldsymbol{B} \tag{4.63}$$

where

$$\boldsymbol{A} = D\boldsymbol{A}_1 + (1-D)\boldsymbol{A}_2 \text{ and } \boldsymbol{B} = D\boldsymbol{B}_1 U_h + (1-D)\boldsymbol{B}_2 U_l \tag{4.64}$$

Figure 4.25 simulates this averaged model.

Switch Frequency Comparable to Input Natural Frequency

When the switch frequency is close to the natural frequency of the input system (See Figure 4.14), the same problem of averaging the input dynamics persists. As Figure 4.26 shows, the profile of the signal T_g is not that of a square wave. Ignoring saturation, $u(t)$ will be a scaled version of this waveform T_g. The harmonic content of this waveform at the

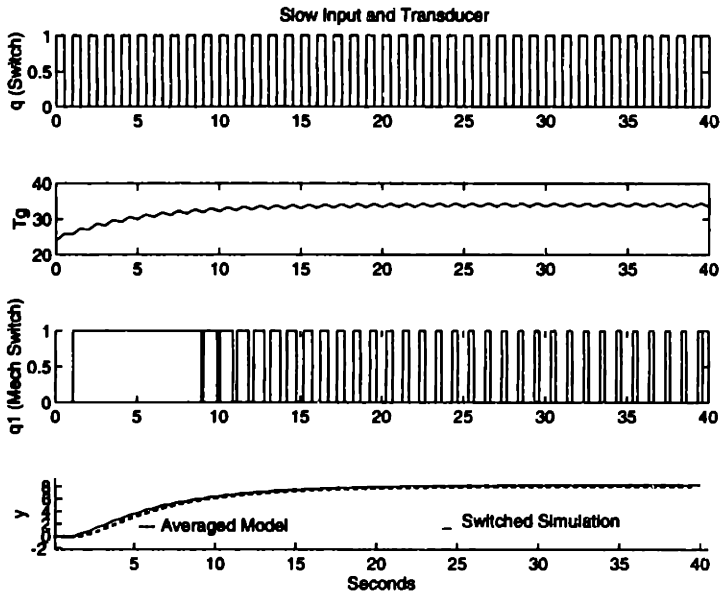


Figure 4.23: Simulation of continuous system with high switch frequency

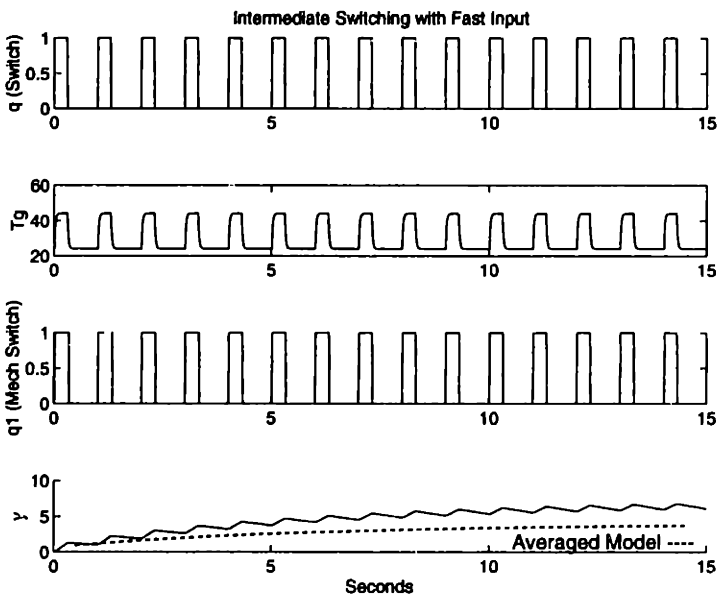


Figure 4.24: Simulation of continuous system with intermediate switching frequency

switch frequency is still significant. Moreover, this results in a different switching function $q_1(t)$ that is seen by the transducer system (as a result of diode switching). Even assuming that the duty ratio of this function, D_1, can be computed or observed, it is still difficult to derive an averaged model for this system with any of the techniques explored earlier. Since the input into the transducer is not a square wave, it is not possible to combine it with the $\boldsymbol{B}$ matrices to produce new input matrices and a constant input. Instead, this temperature T_g has to be treated as a state of the system. The large ripple in this waveform voids the assumption that the states and inputs can be approximated by their averages over a switch period. However, as mentioned in Section 4.1.2, there are special cases where averaging will work even if the ripple in a particular variable is large. It was shown that, for instance, if the input to a system was significantly varying over a switch period, the system could still be averaged if the input matrices $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$ were identical. In this case, the rippling variable is a state multiplied by the input dynamics state matrix $\boldsymbol{A}_I$ and the transducer dynamics input matrices $\boldsymbol{B}_1$ and $\boldsymbol{B}_2$. The input system in the problem formulation is always constant. If the transducer input matrices are also constant:

$$\boldsymbol{B}_1 = \boldsymbol{B}_2 = \boldsymbol{B} \tag{4.65}$$

then the averaging will be valid. The resulting averaged model is :

$$\begin{bmatrix} \dot{\bar{y}} \\ \dot{\bar{T}}_g \end{bmatrix} = \boldsymbol{A} \begin{bmatrix} \bar{y} \\ \bar{T}_g \end{bmatrix} + \boldsymbol{B} \tag{4.66}$$

where the matrix $\boldsymbol{A}$ is:

$$\boldsymbol{A} = \begin{bmatrix} D_1\boldsymbol{I} & 0 \\ 0 & D\boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_1 & \boldsymbol{B} \\ 0 & \boldsymbol{A}_I \end{bmatrix} + \begin{bmatrix} (1-D_1)\boldsymbol{I} & 0 \\ 0 & (1-D)\boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{A}_2 & \boldsymbol{B} \\ 0 & \boldsymbol{A}_I \end{bmatrix}, \tag{4.67}$$

and the matrix $\boldsymbol{B}$ is:

$$\boldsymbol{B} = D \begin{bmatrix} 0 \\ \boldsymbol{B}_I T_H \end{bmatrix} + (1-D) \begin{bmatrix} 0 \\ \boldsymbol{B}_I T_L \end{bmatrix}. \tag{4.68}$$

The constraint that the input matrices are the same under different switch states is a reasonable one. Although some of the dynamics of a system may change depending on

its state, the mechanism of input delivery may remain the same. Most importantly, this result illustrates the potential for the existence of averaged models for switched systems even if practical constraints require an internal state to contain significant switch ripple. Figure 4.27 simulates a system with the two identical transducer $\boldsymbol{B}$ matrices[12].

4.5.4 Fast Transducer Dynamics

Recall that with the discontinuous equilibrium function it was not practical to switch at a frequency close to the natural frequency of the mechanical dynamics. The result would be large ripple in the output. However, there are cases with the continuous system in which a fast transducer (See Figure 4.28) may be switched in this manner. Specifically, if the input dynamics are slow enough to filter the ripple in the switched input, this ripple would not appear in the mechanical output $y(t)$. This is only possible due to the series connection of the input dynamics and the transducer in the system configuration. Figure 4.29 illustrates this result. The mechanical time constants are 0.4 and 0.8 seconds, with a switch period of 1 second and input characteristic time constant of 5 seconds. The duty ratio of the switch function $q(t)$ is stepped to 0.4 and then 0.2. The temperature T_g responds to the step with little ripple. Again, the resulting diode states are described by $q_1(t)$.

Because the dynamics of the transducer are much faster than the temperature, the output profile is the same as that of T_g. The fast dynamics can be ignored and replaced by a simple gain to reduce the order of the model. The gain will be a function of the DC gains of the transducer in its two states weighted by the duty ratio $D_1(t)$ (the average of $q_1(t)$). By setting the derivative $\dot{\bar{y}}$ to zero (as a consequence of its fast settling on the time-scale of the input system dynamics):

$$[D_1\boldsymbol{IA}_1 + (1 - D_1)\boldsymbol{IA}_2]\bar{y} = [D_1\boldsymbol{IB}_1 + (1 - D_1)\boldsymbol{IB}_2]\overline{T}_g. \tag{4.69}$$

This matrix equation can be used to find the static gains that approximate the relationship between the averaged temperature and output. Combined with the input dynamics, this

[12]The two switched mechanical systems in this simulation have different DC gains. In the first-order models used, constraining the input matrices as well as the DC gains to be equal results in identical systems.

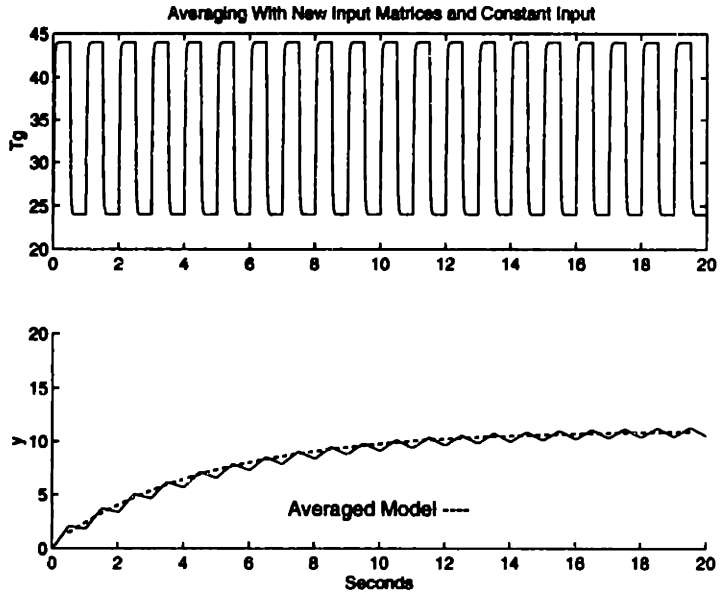


Figure 4.25: Averaged model simulation For intermediate frequency switching

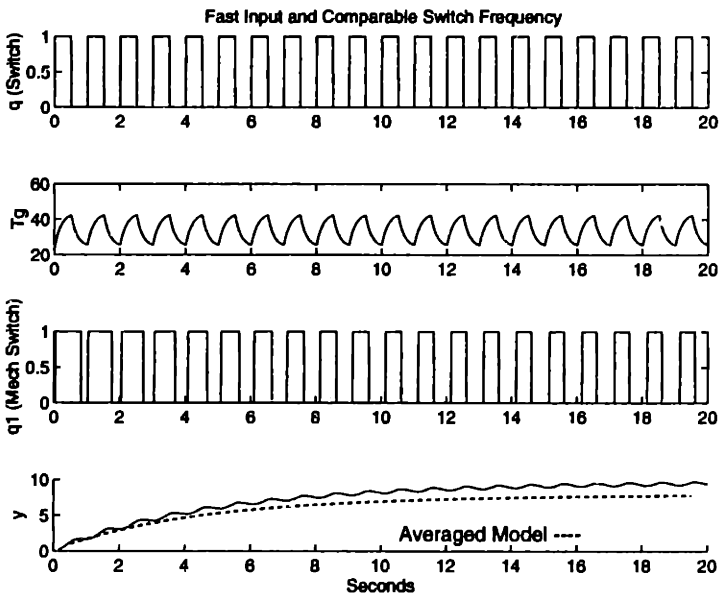


Figure 4.26: Input and switch frequencies comparable

should produce a good model for the overall system. The averaged model prediction shown in the bottom trace in Figure 4.29 is accurate despite the lack of time-scale separation between the switch period and mechanical time constants.

Another significant implication of this result concerns the ability to respond to disturbances in the mechanical variable and correct for deviations from the desired state trajectories. Because the system is controlled through the slower input dynamics, correcting action may be limited by this speed. Particularly when some external disturbance can excite the much faster dynamics of the mechanical system, we may not be able to respond before a significant error has occurred. One example of this is the series elastic component (SEC) of biological muscles [1]. This passive component of the muscle exhibits much faster response than the active motor units. For instance, consider a sudden unexpected strike to the hand when the elbow is bent. Significant deflection in the elbow angle will occur before the active muscle tissue can respond and restore the original position. This issue will be revisited when a dynamic model is developed for polymer gels in the following chapters.

4.6 Summary

The techniques presented above for deriving averaged models are borrowed or modified versions of techniques often used in power electronics. The results that are mostly discussed in the literature rely on the assumption that both input and state variables are well approximated by their average values over a switch cycle. This is a good assumption in power electronics because of the high bandwidth of switching devices and the ability to deliver power quickly and efficiently using almost ideal voltage and current sources. In other systems, such as mechanical actuators, these assumptions are not always as easy to make. Often, it is infeasible or cost prohibitive to achieve certain time-scale properties for a system. As a result it is more challenging to use switched systems and to derive averaged models for them. In [10], averaging based on higher-order harmonics as well as the DC component of the state and input variables over a switch cycle provides more accurate models when the system variables have significant ripple. Sliding mode control [91] and

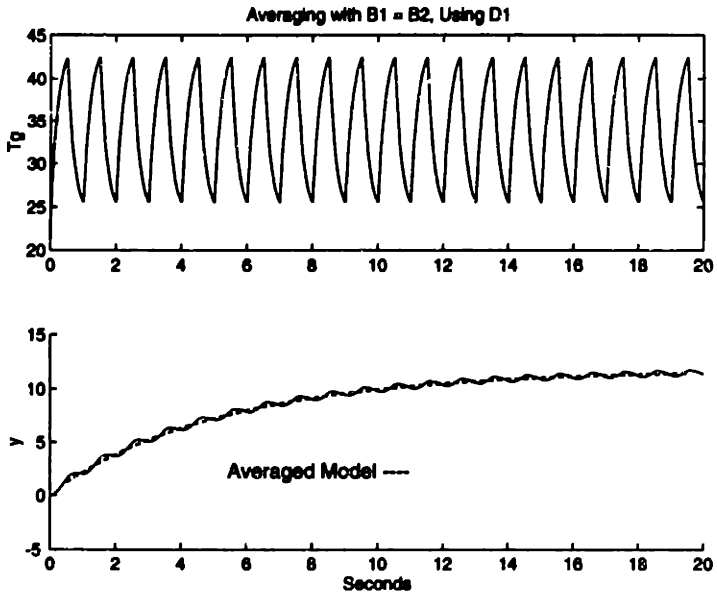


Figure 4.27: Averaged model for comparable input and switched frequencies for systems with identical input matrices

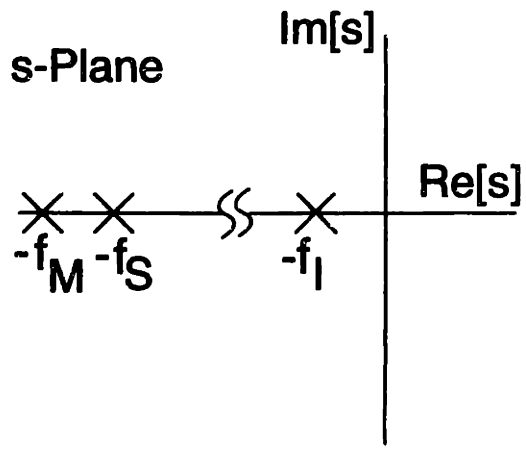


Figure 4.28: Pole arrangement with a fast mechanical system

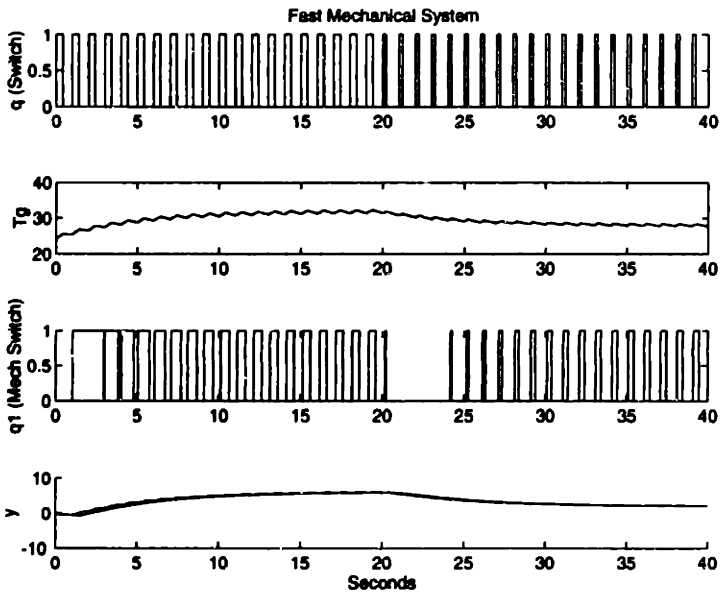


Figure 4.29: Simulation of system with fast mechanical mode

other examples of switched systems ([73] [76] [107]) illustrate different examples of using and modeling mechanical switched systems. In the sections above, a large class of potential switched actuators was explored. Using general system structures with power sources, input dynamics, continuous and discontinuous transducers, and switches to control the power flow and system topology, averaged models for systems with different time-scale properties and component models were explored. When possible, analytical models for the averaged dynamics were developed. Linear dynamic models were used, coupled with nonlinear switches and equilibrium profiles. Many physical systems are either linear or can be approximated by linear dynamics in some regime of operation.

In the structures explored above, a forced switching function $q(t)$ often produced another function $q_1(t)$. The resulting averaged model required knowledge of the duty ratio D_1 associated with $q_1(t)$. In general, this duty ratio will be a nonlinear function of the system states, inputs, and D. Examples of methods for approximating D_1 are discussed in Appendix C. A significant result of this chapter is the demonstration that, under some

system constraints, averaged models can exist even if the usual averaging requirements, such as time-scale separation, are not fully met. Using this result, an averaged model for a switched polymer gel actuator is motivated and derived in the next three chapters. Also, recognizing that certain system properties can simplify the modeling task can guide system design. For instance, a fast input system allows us to drive the transducer directly with a square wave, which reduces the overall order of the model and facilitates system averaging. This was one of several motivations for developing magnetically triggered gels[13]. As will be discussed later, the magnetic activation of polymer gels provides faster input dynamics than other heat transfer systems that were explored.

[13]See Chapter 8.

Chapter 5

Experimental Setup

An experimental setup was constructed to investigate the possibility of utilizing switching control techniques with polymer gel actuators. Figure 5.1 shows an illustration of this setup. Two regulated temperature reservoirs serve as the power sources for exciting the thermosensitive NIPA gel suspended in the glass container. The rubber tubing, the glass walls of the container, and the thermal inertia of the gel solvent and network determine the dynamics of the thermal input system delivering power to the polymer gel. The NIPA gel cylinder is the transducer that converts heat to mechanical work and vertically displaces a suspended load. In order to switch the temperature of the gel, a DC motor with a pinching arm is used to block the flow from one of the fluid circulators as shown. Using this switch to alternate the thermal source connected to the gel container, the polymer gel can be controlled to a cyclic steady state around an equilibrium point. A microcontroller outputs a switching profile to power electronic switches that in turn control the state of the mechanical switch. This switching profile can be stored in memory, or varied based on the sensed position signal to implement a closed-loop controller. The position sensor is implemented with a linear variable differential transformer (LVDT). Both the position signal and the temperature of the gel are fed into a data acquisition system[1]. They are the input (temperature) and output (position) that will be used to derive a mathematical model for the gel transducer. In order to explore the response of the polymer gel actuator

[1]The temperature signal can also be used by the microcontroller in closed-loop control. In this thesis only the position signal was used.

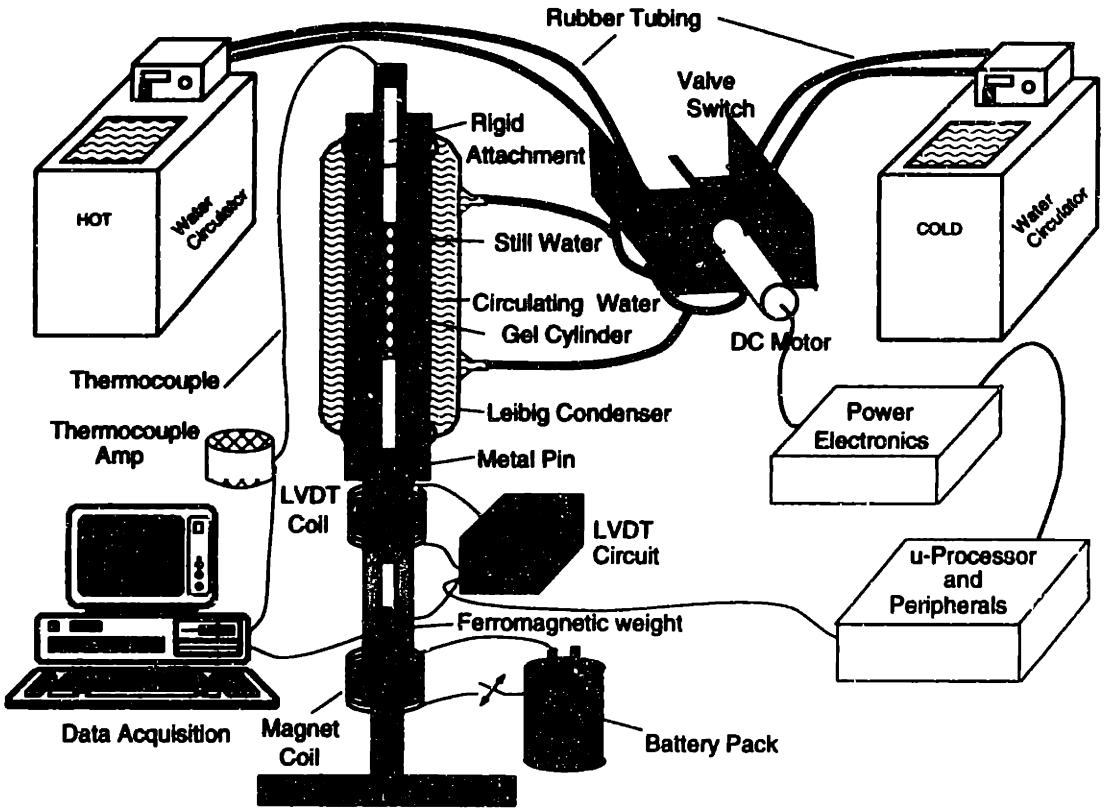


Figure 5.1: Overall Experimental Setup for Switching Control of Polymer Gel

to mechanical disturbances, force can be applied with an electromagnet or a permanent magnet to pull on a ferromagnetic weight attached to the gel. The components of this experimental setup are discussed in more detail in the following sections. Their design and limitations are outlined and used to classify the overall system into one of the general cases of switched actuators analyzed in the previous chapter.

5.1 Polymer Gel and Attachments

The polymer gel transducer is a cylindrical NIPA/PVA gel formed inside a glass tube using the method for making gel cylinders described in Chapter 2. The gel was formed inside a tube with an inner diameter of 0.51 cm. The gel was 8 cm long inside the tube at room temperature (27 °C). The amounts of APS, TEMED, PVA, and NIPA monomer were not varied and no ionic groups were incorporated into the gel network. The amount of BIS was reduced from the formula in Chapter 2 to 0.0075 g. (The effect of this reduction is explained shortly). The gel was attached to a rigid, hollow plastic tube with inner diameter 8 mm. The fixed dimensions of the system shown in Figure 5.2 and the variable volume of the gel are important in predicting the passive dynamic behavior as will be shown in the next two chapters. For the operating point of most of the experiments, the length of the gel was approximately 13 cm and the radius 0.33 cm. Short lengths at the ends of the polymer gel cylinder were covered with Epoxy Patch[2] and glued into the hollow plastic attachements. The top attachement was grabbed with a clamp to suspend the gel into the Liebig condenser. At the bottom end, a steel pin was used as the core for the LVDT position sensor as explained below. Another plastic spacer connects the steel pin to a ferromagnetic weight that can be grabbed by a magnet to produce an external mechanical disturbance to the system. Several aspects of this experimental design should be noted:

- The large 0.51 cm diameter of the tube was necessary to achieve acceptable ripple amplitude with the limited switching frequency of this demonstration system. Ripple on the order of 0.1 cm was achieved with a switching period of 58 seconds under

[2]Product of Hysiol Engineering Adhesives, The Dexter Corporations, Seabrook, NH 03874.

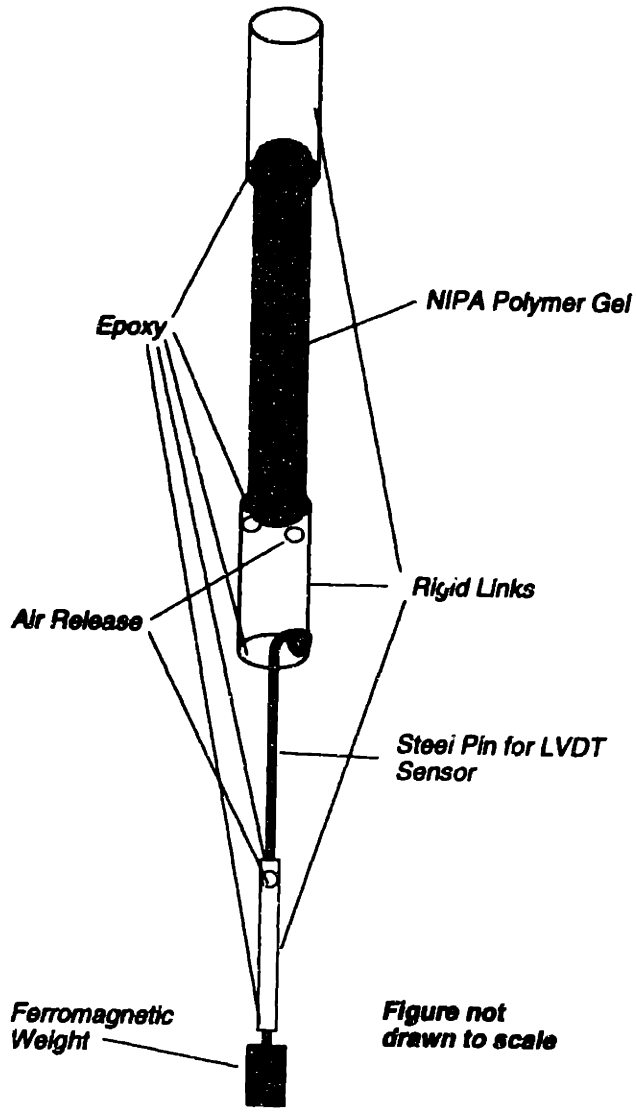


Figure 5.2: Polymer gel and attachements

the conditions of the experiment described below. A larger size gel would be difficult to maneuver inside the Liebig condenser and would result in impractically long experiments.

- The reduction in BIS was used to increase the swelling ratio. A larger swelling ratio helps increase the working range of the actuator while still avoiding the nonlinear effects of operating near the fully compressed or expanded volumes. This is important in developing the small-signal model as discussed later.

- Only rigid attachements were used in the experiment. This helped ensure that the observed dynamics were due to the elastic properties of the gel and not other components in the system.

- Extreme care was exercised to ensure perfect alignment of the gel and its attachments. Use of a narrow glass containter was important to minimize the thermal inertia of the system. As a result, the slightest misalignment between the attachments can result in friction with the glass walls that would alter the dynamics of the gel response.

- Holes were used at the top of the different hollow plastic spacers to ensure that they were always filled with water. Different amounts of trapped air in the fixture would change the bouyancy forces and therefore the operating point of the experiment.

- The separation between the ferromagnetic weight and the steel pin was used to prevent any coupling that may occur between the magnetic forces pulling on the ferromagnetic weight and the LVDT.

5.2 Thermal System

The thermal system consists of two VWR Scientific fluid circulators model 1197, the fluid transport system, the Liebig condenser, and the polymer gel and surrounding solvent. Similar to the system described earlier in Figure 3.3, the switching of two circulators to control the temperature allows us to ignore the slow dynamics associated with changing the temperature of the water reservoirs. A mechanical switch blocks the flow of either the hot fluid

or the cold fluid. The temperature of the cold fluid reservoir is regulated to 32°C and the hot to 36°C[3]. The thermal dynamics in each switch state are those of the fluid transportation and the Liebig condenser and its contents. The following characteristics of the thermal system should be noted:

- The experiment was designed to minimize the transport delay of the circulating fluid from the reservoir to the Liebig condenser. The length of the rubber tubing was minimized and the flow rate on the fluid circulator pumps increased until the delay was less than 2 seconds. As will be shown later, the switching period used was much longer than this delay. The short delay did not significantly influence the modeling and control design for the gel actuator.

- The double-walled container was used to ensure that the gel solvent is unperturbed by the flow of the temperature-controlled fluid. The gel cylinder resides in the inner chamber in stationary water. Any recorded dynamics of the gel length are therefore the product of gel dynamics and not circulator fluid dynamics. While direct application of the temperature-controlled fluid to the gel would have produced faster response, the identification of a dynamic model for the gel would have been more difficult.

- When the switch state is changed from hot to cold, for instance, hot water inside the tubing and the outer jacket of the Liebig condenser is mixed with the water in the cold reservoir. A similar amount of cold water is mixed into the hot water reservoir when the switch is thrown in the other direction. Care has to be taken to ensure that the temperature level in the water reservoir is not significantly disturbed. The temperature rails, the volume of the fluid reservoirs relative to the amount of water in the tubing and the Liebig condenser, and the internal compensation of the water circulator are all parameters that determine the degree of this disturbance. The observed change in temperature due to this effect was less than 0.2°C and was

[3]Moving the temperature rails further apart beyond the 32°C and 36°C resulted in too much disturbance to the water bath temperatures following a change in the switch state. This is explained in the third item below.

corrected within 5 seconds for our specific experiments. This did not affect the results significantly but is one potential source of error in the closed-loop control experiments. For the purpose of model identification, as will be discussed later, this effect can be ignored. The temperature of the gel solvent was used as the input to the gel model. Any change in the level of the reservoir temperatures would have been captured in these measurements.

- The inner diameter of the Liebig condenser is 2 cm. This narrow container was used in order to minimize the amount of solvent and therefore the thermal inertia of the system. A slow input system (thermal in this case) complicates the switching control implementation as discussed earlier. The height of the Liebig condenser is 30 cm. This results in a large ratio of surface area of the inner tube to its volume. As will be shown later, the symmetric distribution of this large surface area around the volume of solvent and gel allows us to make simplifying assumptions for the purpose of modeling the thermal system.

- The mechanical switch response time is a fraction of a second, and it will be approximated as an ideal switch.

The thermal system dynamics are the input dynamics in the switched actuator model. The water circulators are the power sources and the mechanical switch connects the power sources to the input dynamics.

5.3 Control and Data Acquisition

Switching of the mechanical switch is achieved by changing the polarity of current into a DC motor. A mechanical arm connected to the motor shaft as shown in Figure 5.1 swings to one side until it reaches a mechanical stopper and pinches one set of rubber tubing, blocking the flow from that particular reservoir. Limit switches disconnect power from the motor right before the mechanical arm hits the stopper. The power electronics that drive the DC motor are controlled by the output of a digital microcontroller. The 16-bit Intel

80C196KC microprocessor [37] and EV80C196KC [38] evaluation board were used to implement the control and switching sequence generation. One 10-bit A/D channel is used to sample the position of the gel indicated by the output of the LVDT circuitry. Microprocessor code generates a pre-programmed output switching profile (open-loop control) or uses this position signal to adjust the duty ratio of the output switch command (closed-loop control). Special function registers on the 80C196KC generate interrupts that initiate the A/D conversion and sampling, and set the timing to enable pulse-width modulation under constant switching frequency operation. The switch period used in the experiments was 58 seconds. The LVDT signal was conditioned to map 1.67 cm to the 5 volt range of the A/D converter. This provided maximum resolution for the range used in system identification.

A separate system is used for data acquisition. The PCL-818L data acquisition card and the Genie Lite 2.0c data acquisition software[4] were used to sample the position signal from the LVDT and the temperature signal. Using the separate data acquisition system had several advantages:

- The graphical user interface facilitated monitoring of the experiment and easy variation of data-logging parameters, such as sample rate and A/D dynamic range.

- Using a separate system for data acquisition provided added flexibility. For instance, during a long experimental run, data records of parts of the experiment can be examined without interrupting the run.

- The 16 bit A/D converter of the Genie card was essential in recording the temperature signal. For the gel volume changes and the switch period used in the experiments, the change in average temperature was often a very small fraction of the temperature ripple amplitude. In order to resolve the change in the average and capture the ripple, more than 10 bits were needed.

- The 80C196KC was programmed in C. This provided flexibility in implementing the controller code and therefore the Genie software was only used for data-acquisition.

[4]The PCL-818L card and Genie software are products of American Advantech Corporation.

5.4 Sensors

A linear variable differential transformer (LVDT) made of three inductor coils wrapped around an extension of the inner glass tube is illustrated in Figure 5.1. (It is represented in the figure as one coil.) A primary coil is excited by an alternating current, generating magnetic flux in two secondary coils. The secondary coils are placed above and below the primary coil. If the length of the metal pin is not evenly divided between the two secondary coils, the effective core permeability of each secondary is different, and a different amount of flux is linked onto each one. This difference in flux and therefore winding voltage is a measure of the vertical displacement of the metal pin, and therefore of the length of the polymer gel. The LVDT circuitry in the figure conditions this differential signal to provide a position readout to be fed into the data acquisition system and the microprocessor board. The LVDT allows the inner glass container to remain completely sealed. Other sensors (such as a linear potentiometer for instance) would require penetration of the glass container and would probably affect the mechanical dynamics of the system due to extraneous damping or tugging on the gel.

The temperature of the gel is sensed using a type K thermocouple and Iomega TX91-K2 temperature transmitter. Since the gel is extremely sensitive to small changes in temperature, a very accurate measurement is required. Furthermore, this accuracy needs to be sustained over a range covering the full ripple in the temperature waveform. The current output of the transmitter provides a clean and accurate measurement of temperature over the desirable range. As mentioned earlier, the use of the Genie card 16 bit A/D converter also helped resolve small changes in average temperature.

5.5 Mechanical Input

In order to effect disturbances in mechanical load, a ferromagnetic weight was mounted below the metal pin. By pulling on this weight using a magnetic field, a mechanical disturbance can be applied without having to penetrate the sealed Liebig condenser. Two methods were used to apply this external force.

1. Polymer gels exhibit damped mechanical oscillations in response to applied external force, not unlike a rubber band for instance. An automotive solenoid coil with the sliding bar removed was used to excite these modes. The electromagnet provided a clean pull and release of the ferromagnetic weight. Applying a clean step in magnetic field manually using a permanent magnet would be difficult since the observed period of oscillation was a fraction of a second. A 12-Volt battery was used to supply the necessary 4 Amps required to pull the load into the coil, and then release it to produce mechanical oscillations due to a step in tension. A larger load would have required less current, but would have damaged the polymer gel. The current was only applied for a few seconds due to the short time constants associated with the purely mechanical modes of the gel.

2. As discussed in Chapter 2, a change in external mechanical pressure produced a change in osmotic forces similar to a change in ionic content of the gel. The osmotic forces result in a volume and length change that can be observed by the LVDT position sensor. Because of the long time constants associated with a volume change, the mechanical disturbance has to be applied for a long time (several hours for the gel used) to observe a response that exposes the different osmotic modes. This was not feasible with the large amount of current flowing through the electromagnet. Instead, a permanent magnet was used to pull the weight for a long time and then release it to produce the step in tension.

As will be more rigorously shown later, there is significant time-scale separation between the mechanical and osmotic modes. On the short time-scale of the electromagnet experiment, only the mechanical modes were observed and used to identify these modes. On the longer time-scale of the permanent magnet experiments, only the osmotic modes are observed. This is discussed further in Chapter 7.

The permanent magnet was also used to fix the gel at the bottom while the gel was stretched by applying force through a pulley system like the one presented later in Chapter 8. By applying small measurable perturbations in the operating point tension and measuring

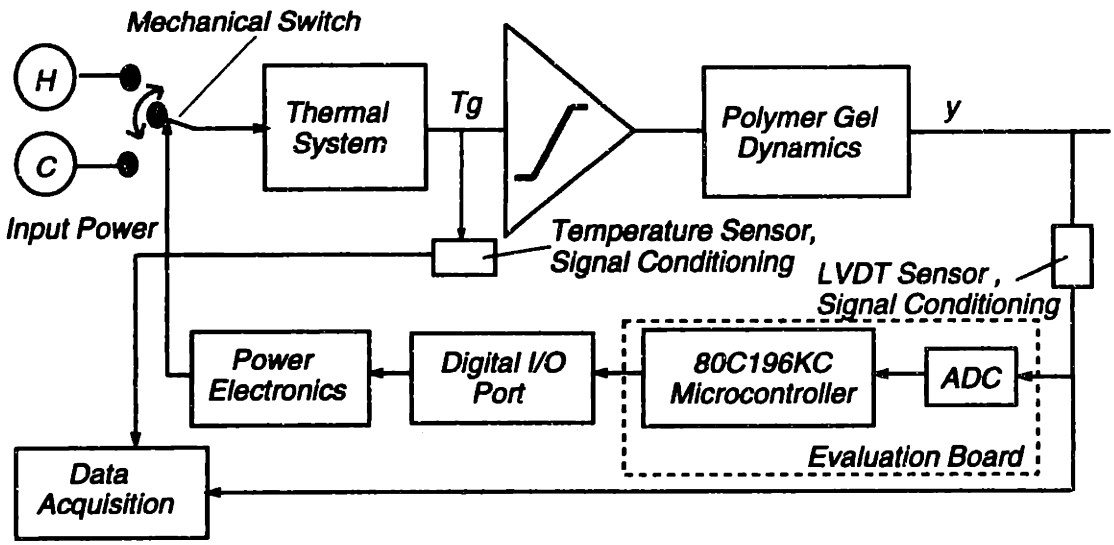


Figure 5.3: Block diagram description of experimental setup

the resulting change in length, the incremental spring constant for the gel was approximated. This value was used to verify the results of the model for the purely mechanical components of the system. This is also discussed further in Chapter 7.

5.6 Overall Layout of Experiment

The components discussed above are used to implement an experimental test-bed for a switched polymer gel actuator system. Figure 5.3 is a block diagram description of the system and its interconnections. Note that the top portion of this system is a special case of the generic switched system description in Figure 4.3. The power sources are the water reservoirs. The switch in the figure is the mechanical switch implemented with the DC motor. The dynamics of the input stage are those of the transport delay of water through the tubing[5] and the thermal inertia and resistance of the Liebig condenser and its contents. The polymer gel dynamics are to be determined by observation of the gel temperature (input) and gel position (output). While a continuous volume-phase transition

[5]This delay is hereafter ignored since it is short compared to the switch period.

gel is represented in the block diagram, the gel may be discontinuous depending on the load and other environmental conditions.

The goal of the experimental design was to achieve time-scale separation between the thermal system and the gel transducer. If the input system is much faster than the transducer system, the switch period can be chosen so that the input to the transducer is a square wave of the same duty ratio as the switching profile. This simplifies the modeling problem for linear systems as shown in Section 4.4.3 or Section 4.5.3. The large cylindrical polymer gel was used because it would have longer time constants than the thermal system. The size of the gel was limited in two ways. It had to fit comfortably inside the inner chamber of the Liebig condenser so that friction with the side walls would be avoided. In addition, an extremely slow gel would require long experiments. Control of the experimental parameters would be difficult beyond a certain number of hours. Electronic drift and other disturbances would affect the quality of the data collected for system identification. The experiments in Chapter 7 for identifying the gel described above were completed in two weeks. Single experiment times ranged from 3 or 4 hours to a full day.

With the constraints on the polymer gel size and the thermal system rate of response, arbitrarily large time-scale separation could not be achieved between the two systems. The switching period of 58 seconds was short enough to produce small ripple in the polymer gel length and load position, but not long enough for the solvent temperature to settle before the switch is thrown. As a result, the time-scales of the system were distributed as in Figure 4.14 and shown again here in Figure 5.4[6]. An example of the temperature and position profiles under these operating conditions is also shown in Figure 5.4. As was shown in the previous discussion on deriving averaged models for switched linear systems, there are special cases where LTI averaged models exist for this set of time-scales. First, models for the different sub-systems (gel transducer, thermal system) are needed. Linear models, or models that can be linearized under certain operating conditions, will lend themselves to averaging. The next chapter develops small-signal linear models for the thermal and

[6]Recall that the single poles in these figures are used to represent the lumped behavior of the dynamics of the subsystem and do not imply first-order dynamics.

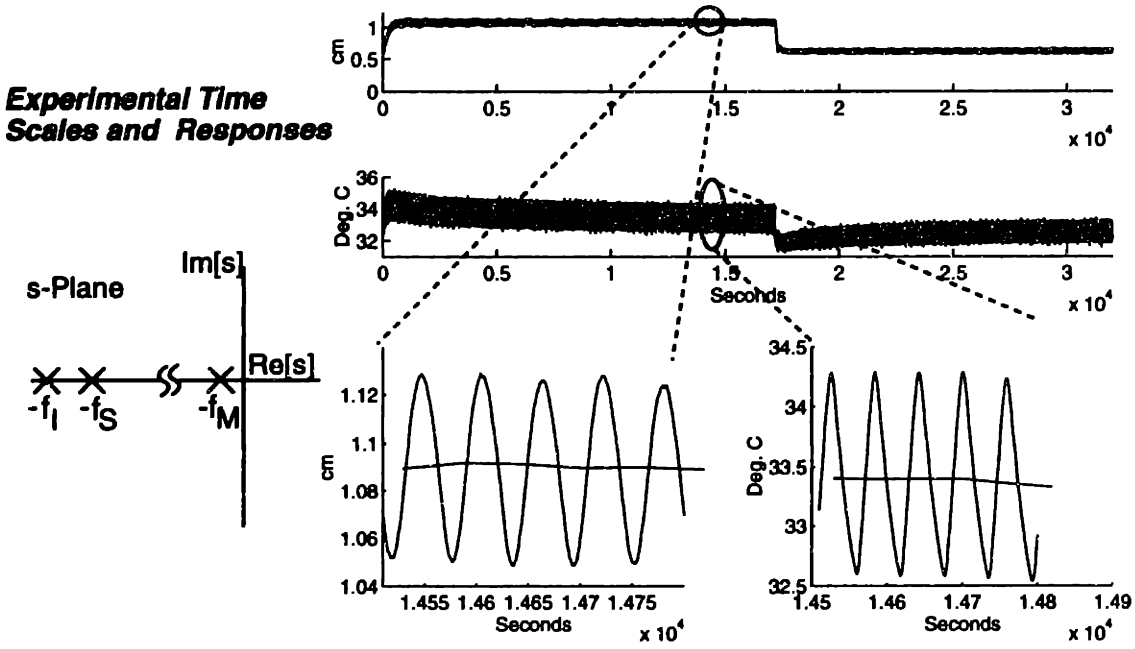


Figure 5.4: Sample response and time-scale properties from the experimental setup

polymer gel systems that are used in Chapter 7 for identifying averaged models for the switched polymer gel actuator.

Chapter 6

A Dynamic Model of the Switched Polymer Gel Actuator

In the following sections, a model for a polymer gel actuator under switching control is proposed. The actuator system to be modeled is the experimental setup discussed in Chapter 5. Analytical descriptions of the input (thermal) and transducer (polymer gel) systems are derived and motivated. The input system model is based on lumped parameter modeling of thermal diffusion systems under simplifying assumptions of uniform mixing and good insulation [56]. A linear, small-signal model is defined for the gel transducer. Based on prior work on the kinetics of gel volume-phase transition and our own observations of the physical behavior of the gel, a parameterized fourth-order model that captures the response of the gel to mechanical disturbance and temperature change is formulated. In the next chapter, identification of the parameters of the gel model is discussed, and an averaged model for the switched polymer gel actuator is derived. Eperimental results are presented for various inputs and input profiles, and the model is verified.

6.1 Thermal System Model

The thermal input system is that described in Section 5.2. Depending on the mechanical switch setting, one of the fluid circulators is in communication with the Liebig condenser. The fluid flows in a closed circuit through the tubing and the outer jacket of the container. Heat transfer due to the temperature gradient between the gel and solvent and the circulat-

ing fluid raises or lowers the temperature of the gel. A model for this system is developed below based on a similar setup in [56].

The energy needed to make a required temperature change ΔT in the gel and solvent is:

$$\Delta E = m_w c_w \Delta T + m_g c_g \Delta T, \tag{6.1}$$

where m_w, c_w and m_g, c_g are masses and specific heats of the water and gel network, respectively. Since typically the mass of the polymer network is much smaller than the mass of the solvent, and the specific heats are comparable [60], we will ignore the second term in the equation and simplify the expression:

$$\Delta E = mc\Delta T, \tag{6.2}$$

where the mass m and the specific heat coefficient c are those of the water in the gel (m_w, c_w). The temperature T is in degrees C. The dynamics by which this temperature change occurs are governed by complicated convection and diffusion processes. Partial differential equations are generally used to model these dynamics, but a more tractable model can be developed with some simplifying assumptions. For instance, if the solvent is continuously mixed, its temperature can be assumed homogeneous and diffusion can be ignored. However, as mentioned earlier, mixing would result in undesirable fluid dynamics. Nonetheless, the same assumption can be made in our experimental setup because of the geometry of the glass container. The Liebig condenser used in the experiment has a height much greater than its radius. At a high flow rate, the temperature-controlled fluid is uniformly distributed around the outside of the internal chamber. Heat will be delivered through the large surface area of the cylinder to the relatively small volume of solvent inside. The result is an approximately uniform temperature in the gel and solvent. Assuming further that the specific heat of water is constant, lumped-parameter LTI models for the thermal system can be developed. Instantaneously, the rate of change of the temperature of the mass m can be related to the energy input using Equation 6.2. An expression for the

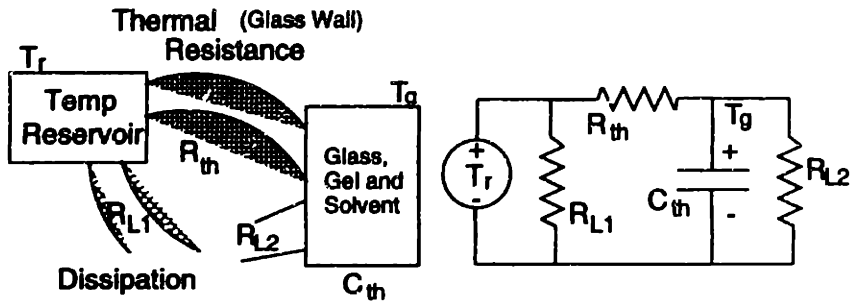


Figure 6.1: Temperature control and heat delivery system - same as Figure 4.4.

heat energy flow W measured in watts is:

$$W = \frac{dE}{dt} = mc\frac{dT}{dt} = C\frac{dT_g}{dt} \tag{6.3}$$

where T_g is used to represent the gel temperature and to avoid notational confusion later with the switch period T. Based on this constitutive law, the contents of the Liebig condenser are analogous to a capacitor C with voltage T_g and current W. The temperature reservoirs are regulated to maintain a programmed temperature and can be modeled as a voltage source in this analogy. The heat in the circulating fluid can flow in several directions. It can be lost to the ambient environment and it can travel through the inner glass wall to the gel and solvent. The path of heat flow can be modeled as a resistor. A good thermal insulator will have a high thermal resistance value, for example. It will create a large temperature (voltage) gradient or drop across its terminals. The ambient temperature is the "ground" potential in our electrical analogy.

Figure 4.4, repeated below for convenience as Figure 6.1, shows an electrical analog of the thermal system. Heat can be lost through the tubing or outer glass wall, represented in the figure by R_{L1}. Heat can be delivered to the internal chamber of the Liebig condenser through the resistance of the glass wall denoted R_{th}. This heat can be used to raise the

temperature of the gel and solvent thermal inertia, modeled by C_{th}[1].

The model in Figure 4.4 can be reduced further. Heat loss can be minimized by good insulation, therefore minimizing the effect of the loss resistors R_{L1} and R_{L2} in the circuit. The thermal system analog is still first-order and can be respresented by the transfer function:

$$\frac{T_g}{T_r} = \frac{1}{\tau_{th}s + 1}, \tag{6.4}$$

where T_r is the temperature of the circulated fluid and $\tau_{th} = R_{th}C_{th}$. A state-space description of this system is:

$$\dot{T}_g = -\frac{T_g}{\tau_{th}} + \frac{T_r}{\tau_{th}}. \tag{6.5}$$

The equations above describe the thermal system in one switch state[2]. Changing the switch state changes the driving temperature T_r. We will assume that the heat flow path and other thermal system parameters are independent of switch position. That is, the dynamics of the temperature model will not be time-varying under switched control. Figure 6.2 shows the response of water temperature inside a Liebig condenser switched between temperature drives of 44°C and 24°C. Simulation of a first-order system model in Equation 6.5 with $\tau_{th} = 28.9$ seconds driven by temperature input switched between 44°C and 24°C reproduces the experimental temperature. The thermal (input) stage of the system can therefore be modeled by first-order linear dynamics.

6.2 Polymer Gel Model

Polymer gel equilibrium behavior is often nonlinear and discontinuous as discussed in Chapter 1. The dynamics generally vary with the gel volume and its surrounding environment. Much of the work on modeling gel behavior has resulted in complicated nonlinear differential equations, with many of the models more qualitatively insightful than accurate and

[1]Note that after the switch is thrown, it takes a finite amount of time for the fluid from the selected circulator to reach the glass container. This can be modeled as a transport delay. However, in the experimental setup, this delay was much shorter than the time constants of the system and the switch period, and was therefore ignored.

[2]If the system is not well insulated, a first-order description can still be used, but the DC gain will be less than unity.

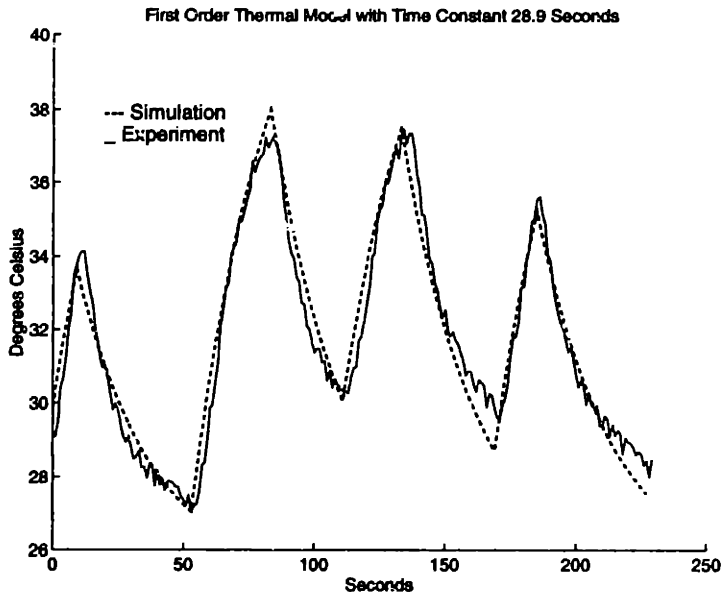


Figure 6.2: first-order model simulation of temperature dynamics

predictive [17] [98]. The goal of this work is to develop a polymer gel actuator model that is reasonably predictive of the mechanical behavior of gels yet easy enough to analyze and solve. A linear model of the gel behavior would be ideal for stability and performance analysis as well as control design. Because of the general nonlinearity of gel volume transition, however, linear models will not be valid except under certain constraints. This difficulty is compounded in the case of switched polymer gels. As shown previously, the switching operation can produce time-varying or nonlinear effects in a system. The sections that follow propose a linear model that is valid in small-signal operation, or for small changes in the gel volume and other variables.

6.2.1 Background

As discussed in Chapter 1, the equilibrium behavior of polymer gel volume transitions is well understood, at least qualitatively. The fundamental forces producing an imbalance in the gel osmotic pressure that ultimately results in a volume change are ionic, hydrophobic, hydrogen bond, and va Der Waals forces. During a volume change, the swelling ratio of a

gel is the same in all dimensions due to the short shear relaxation time of the network. The slower dynamics of isotropic swelling are governed by the diffusion process of the network into the solvent and depend on the gel composition as well as its shape and size. These dynamics are discussed in more detail below.

Kinetics of Gel Transitions

The understanding of kinetics of polymer gel volume phase transitions has lagged behind understanding of the equilibrium behavior. While accurate quantitative modeling for the steady state behavior of gels is still unavailable, a basic understanding is provided by qualitative models [20] [25] [97] [17]. Kinetic modeling has been studied in [98], [99], [100], [59], [58], and [11]. In 1979, it was demonstrated that the volume change processes of a gel are determined by the collective diffusion of the polymer network in the solvent [98]. Previously, the diffusion of individual water molecules into the polymer network was thought to govern the dynamics. Research on the factors determining this collective diffusion coefficient defines it as a function of the elastic modulus of the polymer network k and the frictional coefficient f between the network and the solvent [98] [68]:

$$D_c = \frac{k}{f}. \tag{6.6}$$

Friction of the network with the solvent and elastic properties of polymer gels have been explored in [112] [22] [103] [104]. The kinetics of a polymer gel as it goes through a full volume-phase transition are generally nonlinear. The factors determining the diffusion coefficient, and therefore the dynamics, depend on the actual volume of the gel as well as its final equilibrium state. The response of the gel is therefore described by nonlinear partial differential equations that may require numerical solutions. When the volume change is small, however, the equation of motion for a spherical NIPA gel with final equilibrium radius a and total radius change Δa_o can be reduced to a diffusion equation [69]:

$$u(a,t) = \frac{6\Delta a_o}{\pi^2} \sum_{n=1}^{\infty} \frac{e^{\frac{-n^2 t}{\tau}}}{n^2}, \tag{6.7}$$

where $u(a,t)$ is the displacement of a fixed point on the gel network from its equilibrium location a at time t. The time constant τ is defined as:

$$\tau = \frac{a^2}{\pi^2 D_c}. \tag{6.8}$$

Similar results have been obtained for cylindrical gels. For further detail on this formulation and its origin, see [69]. Note that according to this mathematical description, the radius of the spherical gel changes rapidly for times smaller than τ and follows a single exponential profile for the rest of the time as it approaches its final value. This dominant single exponential $e^{\frac{-t}{\tau}}$ is a function of the dimensions of the gel as well as its diffusion coefficient. This two-time-scale property of the sum of exponentials solution will be the basis for a second-order linear approximation for the diffusion dynamics under small volume change.

Polymer gels also exhibit disparate rates of response between their shrinking and expanding transitions [97] [98] [17]. We have generally witnessed the NIPA gel to collapse faster than is swells. Equations 6.7 and 6.8 above would indicate, however, that if the transitions are kept small enough, the time constants for small volume changes around a nominal point depend only on the nominal volume and not the small deviations.

Gels as Mechanical Devices

The understanding of polymer gel kinetics and the dynamics of volume-phase transitions is essential to their application as mechanical actuators. Polymer gels are appealing as mechanical devices because of their conformability, low cost, and variety of trigger mechanisms. Hence, the mechanical properties and the potential use of polymer gels as actuators have been the subject of active research for over a decade. However, previous work has mostly focused either on static mechanical properties of polymer gels [44] [90] [33] [92] [34] [21], or attempts to devise clever ways of using the gels as actuation devices without the use of an analytical model. Examples of this are applications in [47] [94] [93] [81] or the gel looper discussed in [82] and described in Chapter 1. Others have explored the dynamic modeling of polymer gels for mechanical design and control [8] [52] [88] [23]. Invariably, the complete volume phase transition of a gel was characterized. In general, the results have been com-

plicated nonlinear models. The development of a dynamic model describing the behavior of a switched or pulsed polymer gel poses additional challenges. The interesting variables in a pulsed or PWM application are the averaged variables. Further analysis may be required to arrive at the behavior of these quantities. In order to be able to use a polymer gel effectively as an actuation device, a switching approach may be necessary. A simple yet predictive dynamic model for a switched gel is important for these applications. The model should apply to a large class of gels, and should be independent of the actuation mechanism.

6.2.2 Polymer Gel Mechanical Model

Figure 6.3 shows a polymer gel suspended in a temperature-controlled solvent. We will assume that the temperature is uniform within the gel structure. This is a good assumption since the thermal diffusion in a gel network is much faster than the network/solvent diffusion process [60]. A change in the temperature of the gel can result in an imbalance of forces that causes the gel to move toward a new equilibrium point. For instance, if the temperature change causes a positive osmotic pressure pushing the gel outward, the gel volume will increase. The new equilibrium is not attained instantaneously, but rather at a rate governed by the network solvent diffusion process. In effect, the internal process produced by a change in temperature produces a force at the gel boundary through the dynamics of gel network diffusion. This process is represented in the figure as the active gel unit (AGU). This unit converts heat into mechanical energy by taking a temperature input $T(t)$ and producing a force output $F_1(t)$.

The passive gel unit (PGU) represents the mechanical linkage of the gel and its load. The force $F_1(t)$ causes a change in the volume of the gel by moving the network strands and any attachments. In the one dimensional actuator in Figure 6.3, $F_1(t)$ can move the mass of the polymer gel and the mechanical load. The gel network and solvent, as well as the load, comprise the PGU. The dynamics of this unit depend on the mechanical properties of the gel and the load. Note that the mechanical output $y(t)$, or the position of the load, can be altered by the application of an external force $F_2(t)$ that stretches the gel. As a result, the PGU has two inputs as shown. A polymer gel can undergo a mechanical

deformation in one of two ways. A change in the internal balance of forces results in a diffusion process that produces a volume change. The friction between the network strands and the solvent as they move past each other makes this a slow process. On the other hand, an external tugging or compressing force can be applied to the gel. This can result in a mechanical deformation without any volume change. As shown in Figure 6.4, the gel can be stretched and made thinner and longer. This process can be much faster than the diffusion process since it does not involve relative motion between the gel network and the solvent[3]. Models for the dynamics of both the diffusion process and the mechanical linkage are developed below. It is important at this point, however, to reiterate the distinction in the model between the two components. The AGU represents the process of diffusion that can convert the input energy to mechanical energy by causing a volume change. The PGU represents the passive mechanical structure of the gel. The dynamics of this component are generally faster because of the short shear time constants of the polymer gels. A conceptual overview of this system structure is illustrated in Figure 6.5.

The AGU input $T(t)$ is the temperature that can produce a volume change in a thermo-sensitive NIPA gel. Another input to the AGU shown in Figure 6.5 is the mechanical tension $F_2(t)$. Often, many variables will influence the balance of forces in a gel that can produce a volume transition. Most of these effects, such as chemical changes in the solvent, can be eliminated and ignored in this system. However, the effects of mechanical disturbances need to be considered. In [92], at least the equilibrium characteristics of polymer gels are shown to change with external load. Mechanical loading provides tension that induces osmotic pressure imbalances much like ionic groups in a gel network [17]. In a servo-mechanical application, a situation where the load changes is very likely to occur.

In the sections that follow, mathematical models for the two components of the gel mechanical transducer will be developed. These models are based on the discussion of the processes above, observed and documented gel behavior, and the basic physics governing

[3]Recall from Chapter 1 that in the case of a cylindrical gel, most of the diffusion occurs in the radial direction. The shear relaxation process, which is much faster than the diffusion process, keeps the ratio of the dimensions of the gel constant as it changes its volume. This shear relaxation process witnesses no relative motion between solvent and network and is a constant volume process. In the model in 6.3, it is represented by the PGU. The diffusion process is the AGU.

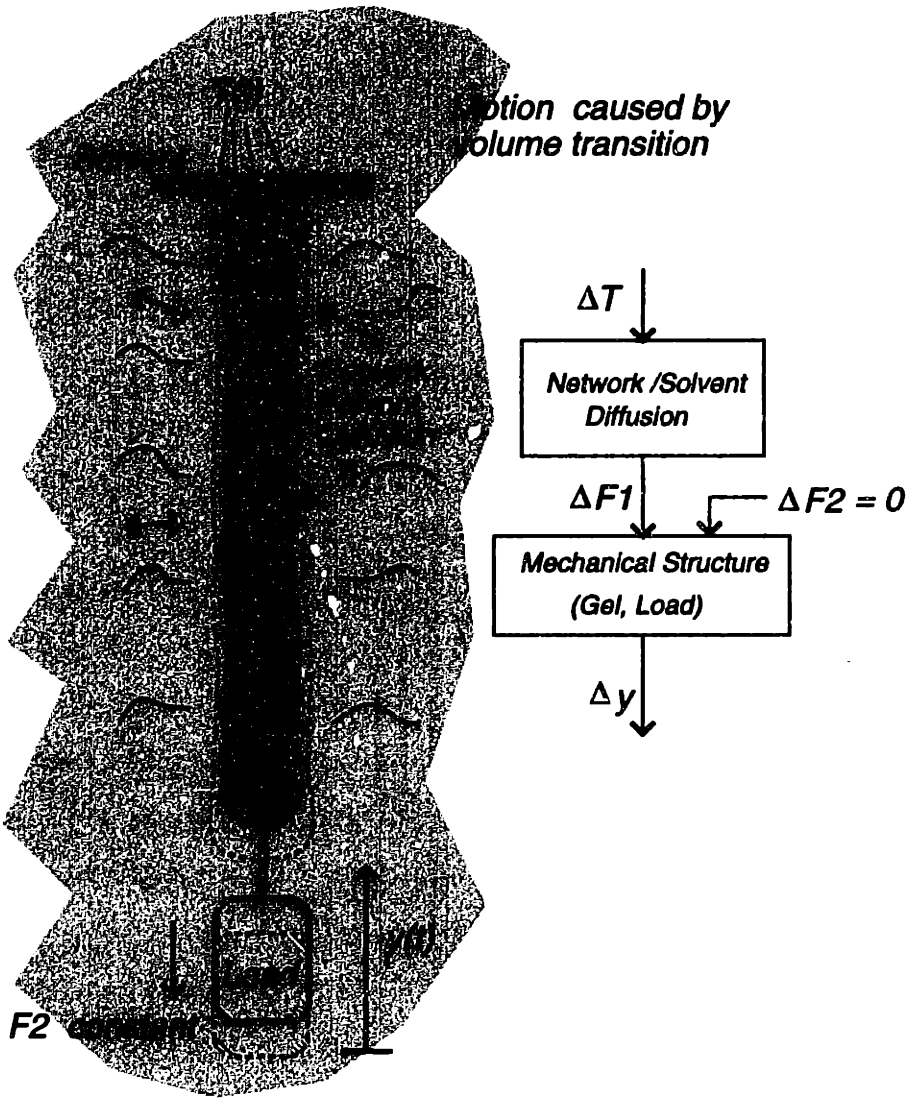


Figure 6.3: Change in mechanical states produced by a volume transition

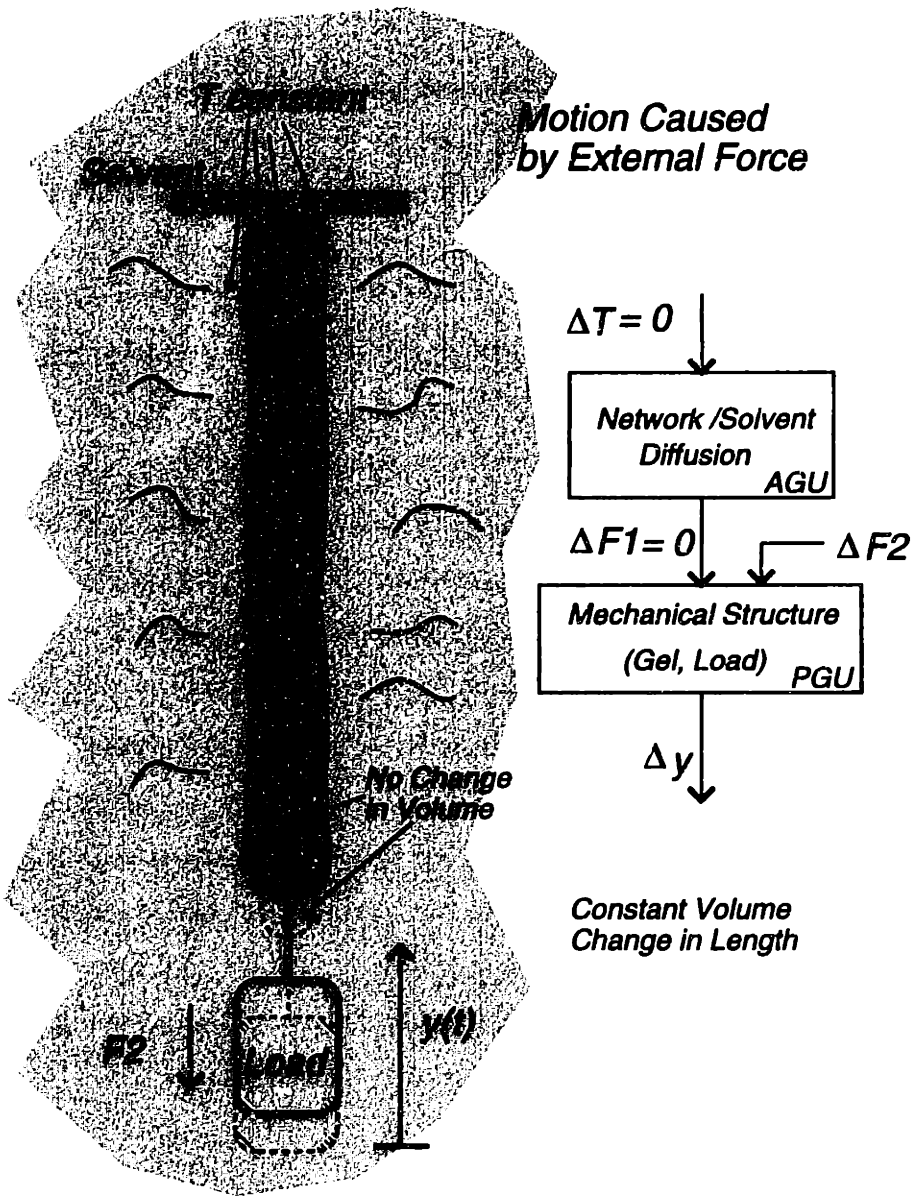


Figure 6.4: Change in mechanical states produced by external force at constant gel volume

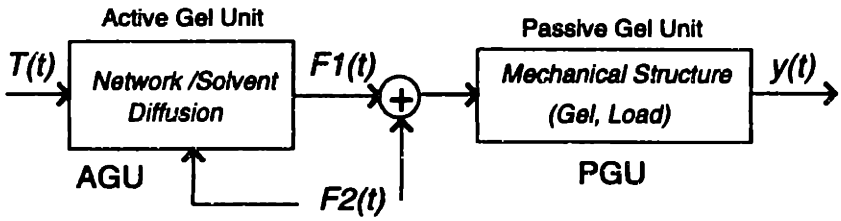


Figure 6.5: Conceptual model of gel system and input output relationships

volume phase transitions and the shear deformation of the gel. We will assume that motion is one dimensional and that any environmental variables other than mechanical tension and temperature can be controlled or kept constant. A characterization of the response of the output position $y(t)$ to inputs $T'(t)$ and $F_2(t)$ will capture the behavior of the gel actuator under these conditions.

It is important to note here that the model developed below is a parameterized LTI model that is an accurate description of the gel system with specific parameter values under small-signal operation. The same parameterization will hold at different equilibrium point, but with new parameter values. In order to use polymer gels in trajectory tracking robotic applications, for instance, gain scheduling or other adaptive control techniques should be used in conjunction with parameter estimation algorithms. While future work should address the issues of large signal model development and understanding how the parameters of the function vary with the operating point, the small-signal model is a good starting point for understanding the gel behavior and establishing guidelines for model identification and control design for gel-powered systems.

6.2.3 Active Gel Unit Model

The network/solvent diffusion process can be modeled by nonlinear partial differential equations that are difficult to solve analytically. Under the assumption of small volume change, the sum of exponentials in Equations 6.7 and 6.8 is a good approximation of the solution. This solution demonstrates two distinct rates of response. The slower rate of response is

dominated by the slowest exponential term. The faster rate is the lumped behavior of the rest of the exponentials. In [69], this dual rate response was observed in NIPA/sodium acrylate gel spheres subjected to steps in temperature and is illustrated in Figure 6.6. Similar behavior occured in gel cylinders and in experiments with the polymer gel used in our experimental setup. The curves in Figure 6.7 illustrate changes in the length of a cylindrical NIPA/PVA polymer gel due to a step in temperature. In this experiment, the mechanical switch in the experimental setup in Figure 5.1 is thrown once only to step the temperature of the polymer gel as shown. The same gel system was also subjected to a switching temperature drive as described in the previous chapter. The mechanical switch is cycled between the hot and cold reservoirs every 58 seconds. The result is a change in the length of the gel over a switch period that is a function of the duty ratio of the switching function. The averaged response of the gel length to a step in the duty ratio of the switching sequence that alternately heats and cools the gel is shown in Figure 6.8. Figures 6.9 and 6.10 show the response of NIPA gels seeded with nickel flakes and immersed in ferrofluid, respectively, to a step in magnetic field[4]. All of the responses shown exhibit two distinct rates of change. This observation is used below to develop a second order linear model for the polymer gel diffusion process.

It is important to note here that the sum of exponentials as stated in [69] describes the dynamics of the change in the *radius* of a sphere, which is a *size* variable and not a *force*. The dynamic responses shown also illustrate the *lengths* of gels as they change with a temperature induced volume transition. However, when the gel mechanical state is altered only due to a change in volume, the response is dominated by the dynamics of the diffusion process. The observed dynamics of length or radius are therefore those of the force F_1. Specifically, in the cases of free swelling beads or the responses of cylindrical gels under constant mechanical load, the diffusion process is too slow to excite the other sources of mechanical dynamics in the system. This will become clear later as time-scale separation between the different model components is discussed. For now, it suffices to say that the

[4]Magnetically triggered gels were developed as a result of this research effort and are discussed in Chapter 8

dynamic response of a volume or length variable under free swelling or constant tension conditions reflects the dynamics of the diffusion process, which will be modeled by the AGU with output F_1. Since these conditions are true of the responses illustrated here, the dynamics of the length or radius variables shown are those of F_1.

Model Equations

The AGU has inputs $T(t)$ and $F_2(t)$ and output $F_1(t)$. We will first develop a parameterized linear model for the transfer characteristic from temperature to output force. While the static and dynamic properties of volume changes due to temperature have been widely observed and recorded, work on the effects of mechanical tension on volume transition has been limited[5]. In either case, however, the volume change occurs as a result of network/solvent diffusion. Since the volume change process is the same regardless of which input produces it, it should demonstrate the same dynamics in either case. This will result in the same eigenvalues or pole locations in the transfer function from $F_2(t)$ to $F_1(t)$ as in the function from $T(t)$ to $F_1(t)$.

The simplest second-order description for an over-damped linear system is the two-pole model:

$$\frac{F_1(s)}{T(s)} = \frac{a}{s^2 + bs + c} = \frac{a}{(\tau_1 s + 1)(\tau_2 s + 1)}. \tag{6.9}$$

The solution to a step response of this system is a weighted sum of two exponentials, with τ_1 and τ_2 governing the decay times of the exponentials. A first attempt at modeling the gel dynamics matches these time constants with the two response times witnessed in a gel response. It turns out that there are not enough degrees of freedom in this model to produce step responses similar to the ones in Figures 6.6 - 6.10. This is illustrated below. Consider the Laplace transform of the second-order system above when multiplied by a frequency domain step $\frac{1}{s}$ to yield a step response:

[5]This can partly be attributed to the fact that it is more difficult to impose uniform three dimensional tension than it is to subject the whole volume of a gel to a temperature change. Studies of the effects of uniaxial tension have been carried out in [92] [32], for instance, but only the static behavior was explored

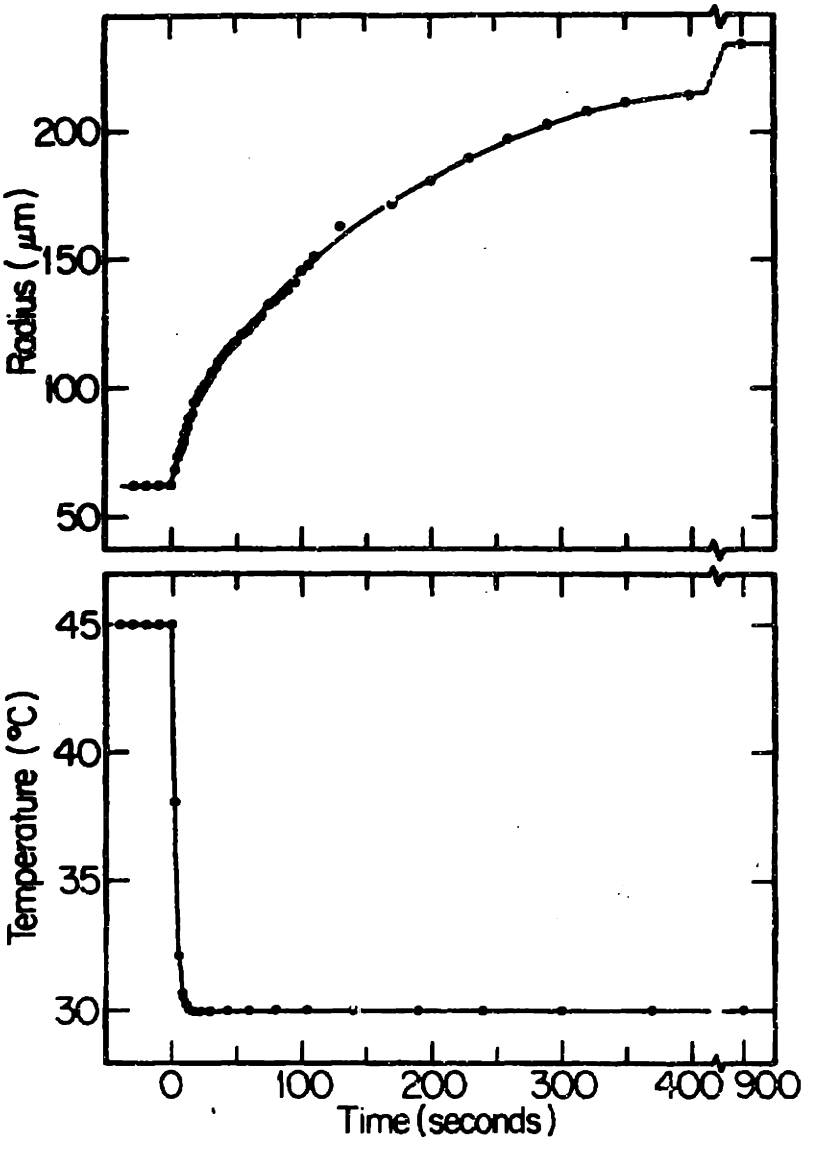


Figure 6.6: Dynamic response of NIPA/sodium acrylate gel bead from [69]

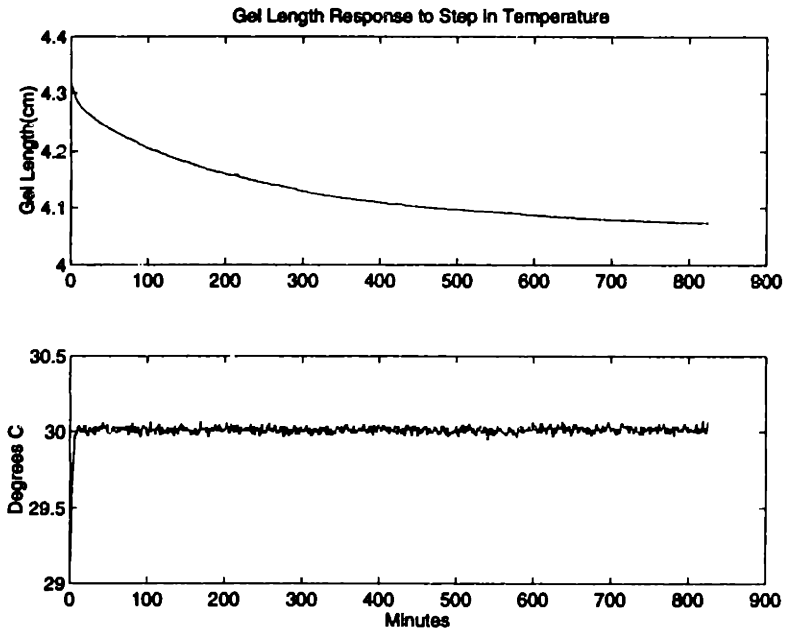


Figure 6.7: Response of NIPA/PVA cylindrical gel to a step in temperature

$$F_1(s) = \frac{a}{s^2 + bs + c}\frac{1}{s} = \frac{a}{(s - s_1)(s - s_2)s}, \tag{6.10}$$

where s_1 and s_2 are the solutions of the quadratic equation

$$s = \frac{-b \pm \sqrt{b^2 - 4a}}{2a}. \tag{6.11}$$

The partial fraction decomposition of this signal is:

$$F_1(s) = \frac{A}{s} + \frac{B}{s - s_1} + \frac{C}{s - s_2}. \tag{6.12}$$

The coefficient multiplying the step component (the DC gain of the system) is

$$A = \frac{a}{s_1 s_2}. \tag{6.13}$$

This coefficient A multiplies a constant part of the response to a step in the input. The other two parts of the response are the two exponentials that are the natural frequencies of

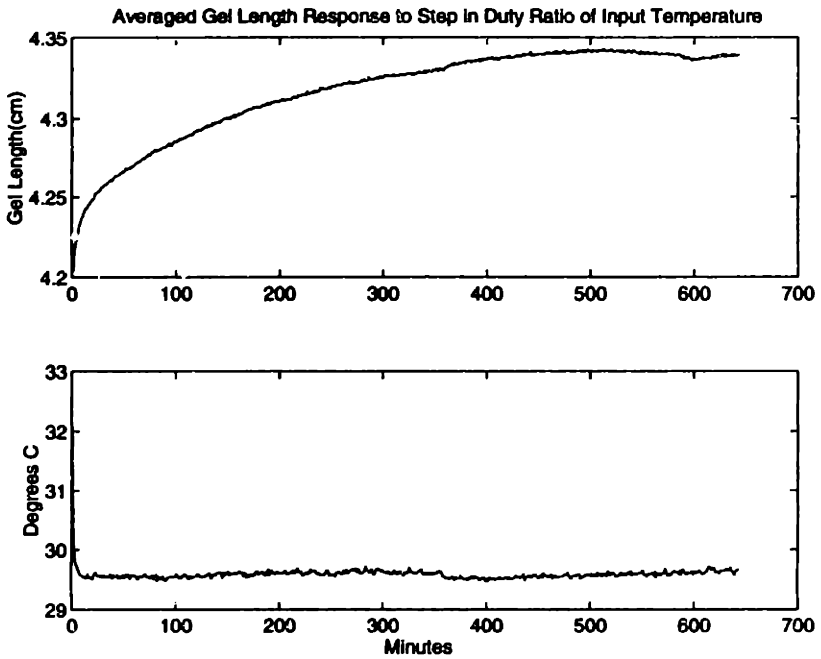


Figure 6.8: Response of NIPA/PVA cylindrical gel to a step in input duty ratio. The input switch selects either a hot or cold temperature drive for the gel solvent. The average temperature and length of the gel are shown.

the system. These two exponentials are added to the constant value of the drive multiplied by the DC gain with the following weights:

$$B = \frac{a}{s_1(s_1 - s_2)} \qquad and \qquad C = \frac{a}{s_2(s_2 - s_1)}. \tag{6.14}$$

Cross-multiplying and solving for a in terms of B and C in the above equation:

$$a = (s_1^2 - s_1 s_2)B = (s_2^2 - s_1 s_2)C, \tag{6.15}$$

and

$$\frac{B}{C} = \frac{(s_2^2 - s_1 s_2)}{(s_1^2 - s_1 s_2)}. \tag{6.16}$$

For a ratio of pole magnitudes

$$\frac{s_1}{s_2} = R, \tag{6.17}$$

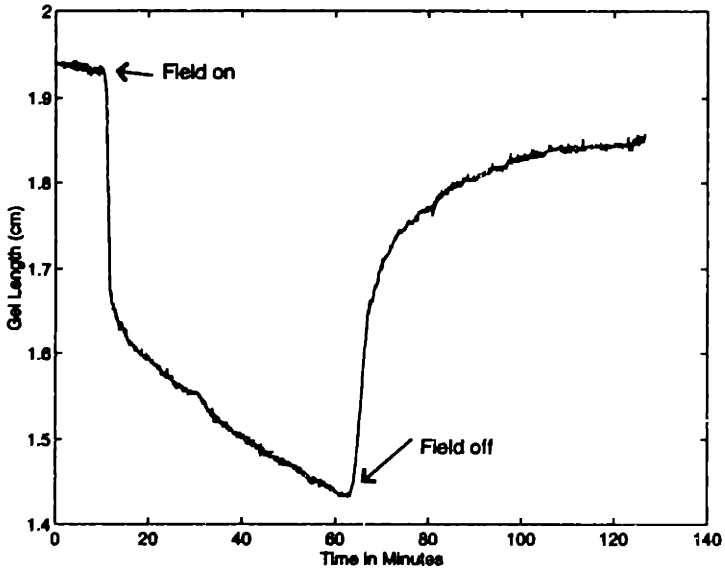


Figure 6.9: Dynamic response of nickel seeded magnetically triggered gels

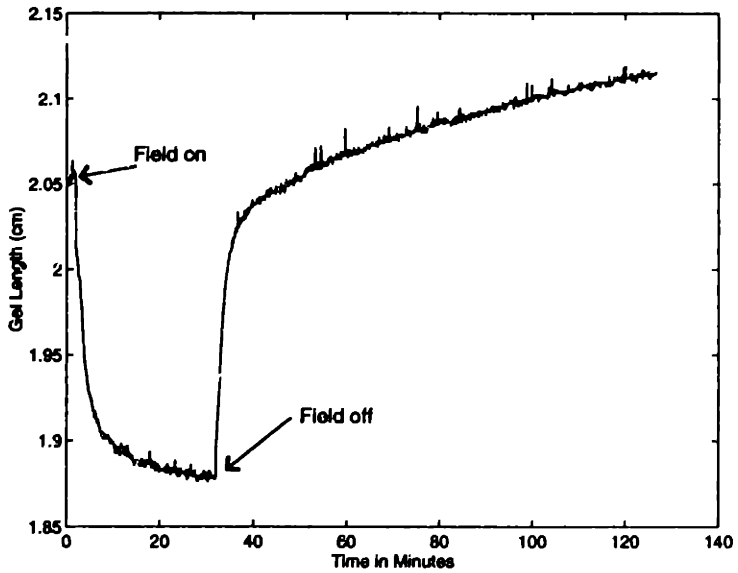


Figure 6.10: Dynamic response of ferrofluid solvent magnetically triggered gels

we can rewrite the ratio:

$$\frac{B}{C} = \frac{(s_2^2 - Rs_2^2)}{(R^2 s_2^2 - Rs_2^2)} = \frac{1-R}{R(R-1)} = \frac{-1}{R}. \tag{6.18}$$

The slower exponential is weighted in the step response by a weight that is R times larger than the weight of the fast exponential, where R is the ratio of the magnitude of the faster pole to the slower one. For any significant separation between these poles, such as in a heavily damped system, the fast pole has negligible effect on the response. The step response of such a system will resemble that of a first-order system. In the case of the gel experiments shown above, the two distinct time constants are at least an order of magnitude apart. Yet they both are significant contributors to the step response profile. Another degree of freedom to remove the constraint on the weights of the two exponentials is required. The minimum order system that can exhibit this behavior is a second order system with one finite zero. The transfer function for this system is:

$$\frac{F_1(s)}{T(s)} = \frac{ds + a}{s^2 + bs + c}. \tag{6.19}$$

The form of the partial fraction decomposition for this system is the same as the one in Equation 6.12. The partial fraction coefficient of the DC component is still the same ($A = \frac{a}{s_1 s_2}$). However, the weights of the two exponentials are now:

$$B = \frac{ds_1 + a}{s_1(s_1 - s_2)} \quad \text{and} \quad C = \frac{ds_2 + a}{s_2(s_2 - s_1)}. \tag{6.20}$$

By choosing d and a independently, the two weights can be arbitrarily adjusted. Equation 6.19 is chosen as a linear model for the transfer characteristic from the gel temperature to its volume. Recall that, since the diffusion process producing a volume change is the same whether it is triggered by $T(t)$ or by the external tension on the gel through the mechanical linkage, $F_2(t)$, the poles of the transfer function to the output from both inputs should be the same. The effect of external tension on volume change can therefore be described by:

$$\frac{F_1(s)}{F_2(s)} = \frac{ms + n}{s^2 + bs + c} \tag{6.21}$$

This multi-input single output system can be realized in state-space form:

$$\begin{bmatrix} \dot{x} \\ \dot{F}_1 \end{bmatrix} = \begin{bmatrix} 0 & -c \\ 1 & -b \end{bmatrix} \begin{bmatrix} x \\ F_1 \end{bmatrix} + \begin{bmatrix} a & m \\ d & n \end{bmatrix} \begin{bmatrix} T \\ F_2 \end{bmatrix} \tag{6.22}$$

Note that using the observable form to realize the transfer functions eliminates the need for different output matrices. The effects of the zeros in the transfer functions are captured in the input matrix. As a result, the output is one of the states. This will be important in developing a compact state-space model for the total gel system.

small-signal Linearity

The general behavior of the diffusion process, as mentioned earlier, is described by nonlinear equations. The linear model formulated above will only be a valid description of the process under specific constraints. The solution in Equations 6.7 and 6.8 is only valid when the change in gel volume is small compared to the maximum swelling ratio of the gel. The linear model for the AGU will be specific to a particular equilibrium gel volume, becoming less accurate as the volume of the gel deviates from that value. Another source of nonlinearity is the diffusion coefficient D_c. It is one of the factors determining time constants in Equation 6.8, and it varies with temperature, load tension, ionic content, and volume [17] [97]. In addition to volume, the load needs to be held near an operating point value[6]. Ionic content and other chemical effects are not significant since the actuation mechanism is temperature.

Temperature may not be easy to constrain. In a switched system, even in equilibrium, the temperature value may vary significantly over a switch period. This may affect the exponential solution. The time constant in Equation 6.8 depends on the final radius a of the spherical gel. The gel spheres studied in [69] were driven by steps in temperature. The size of the step implied a final gel volume. If the temperature profile is varied during the volume transition, a change in a would result. The variation in D_c and a during a switch period may be a source of nonlinear behavior. Recall that we were successful in deriving an

[6]The tension due to load, and not necessarily the mass of the load, influences the diffusion coefficient. In most cases, specifically the one explored here, gravitational force pulls the load, and therefore its mass is significant as a source of tension.

averaged model when switching among two or more linear systems. If the dynamics of the temperature system are fast enough compared to the gel dynamics, a switching frequency can be selected so that the temperature settles to the rail values in every switch state. The constant temperature during a switch state enables the derivation of an LTI model for that state. The different LTI models can then be averaged using the techniques discussed earlier. If, however, the temperature profile is more complicated, the dynamics of the gel in one switch state will depend on the input. The result is a nonlinear system that is more difficult to average. Averaged models may still exist, however, as will be shown in the next chapter.

6.2.4 Passive Gel Unit

The PGU represents the passive mechanical structure of the gel system. The polymer gel is elastic, with the network strands compressing or relaxing in response to pressure. In the one dimensional mechanical actuator, the external force $F_2(t)$ can cause the length of the gel, and the mechanical output $y(t)$, to change. A step in downward force, for instance, will first cause the gel to stretch, increasing in length while remaining at a constant volume. Eventually, if the force input persists, a volume change will occur at the rate of the AGU dynamic response. The dynamics of the initial stretch in $y(t)$ are the dynamics of the PGU. They depend on mechanical properties such as the mass of the system, friction with surroundings, and the stiffness of the gel network. A model for these dynamics is developed below.

Model Equations

A linear model for the mechanical structure of the PGU is shown in Figure 6.11. The spring coefficient depends on the rubber elasticity of the gel, as well as its network density. The rubber elasticity is a property of the polymer gel strands, and is a function of the chemical makeup and the temperature of the gel. The network density varies with the volume of the gel. The more collapsed the gel, the denser and more entangled the network, and the more resistant to external force. The damping coefficient models the friction of the moving elements with their surroundings. Under constant volume conditions, the gel can stretch

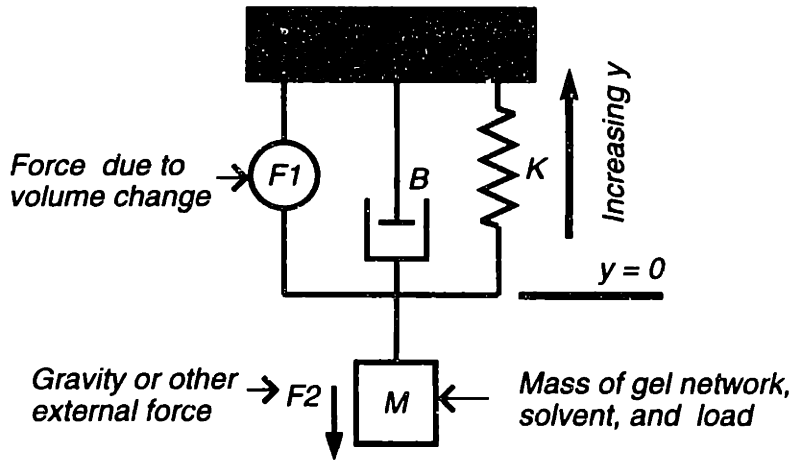


Figure 6.11: Linear model of the polymer gel mechanical structure

and become longer. The gel and its load move through the water. Recall, however, that there is not relative motion between the network and solvent, as in the case of the AGU diffusion process. The only friction is between the outer surface of the gel or load and the solvent. This is represented by the damper in Figure 6.11. This system is less damped than the diffusion system, and may exhibit oscillatory behavior depending on what mass is attached to it. The mass in the model is that of the network, the load, and the solvent inside the gel. On the time-scale of the dynamics of this system, the solvent does not travel across the boundaries of the network because the process of diffusion is too slow. Both the network and solvent will move together. This motion can be produced by applying an external force $F_2(t)$ or an internal force $F_1(t)$ due to a temperature change. When the temperature changes, the volume of the gel changes. However, on a short time-scale, the assumption that there is no relative motion between network and solvent is still a good one. This is important for the validity and linearity of the PGU model, as will be discussed below. The transfer function describing the relationship between internal force $F_1(t)$ and output $y(t)$ is:

$$\frac{Y(s)}{F_1(s)} = \frac{1}{Ms^2 + Bs + K}. \tag{6.23}$$

The same transfer function relates $F_2(t)$ and $y(t)$:

$$\frac{Y(s)}{F_2(s)} = \frac{1}{Ms^2 + Bs + K}. \tag{6.24}$$

A state-space description of this system is:

$$\begin{bmatrix} \dot{y} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\frac{k}{M} & -\frac{B}{M} \end{bmatrix} \begin{bmatrix} y \\ v \end{bmatrix} + \begin{bmatrix} 0 & 0 \\ -\frac{1}{M} & -\frac{1}{M} \end{bmatrix} \begin{bmatrix} F_1 \\ F_2 \end{bmatrix}. \tag{6.25}$$

Although the two transfer functions are identical, separating the input in Equation 6.25 makes it possible to combine this model with that of the AGU to obtain the complete gel model.

small-signal Linearity

The linear models of the components in the passive gel unit description ignore dependencies of the component constitutive laws and parameters on changes in the states or inputs of the system. The damping effect is due to the friction between the outer surface of the gel and load attachments with the surrounding solvent. Under small-signal conditions for the volume, the surface area and the texture of the gel cylinder will not vary significantly. Moreover, the viscosity of water should remain relatively unchanged for temperatures between the two rails. As a result, the linear damper should be a good model for friction under small-signal operation. The mass in the model is that of the gel network, the solvent in the gel, and any attached load. The only component of the mass that can change is the solvent in the gel. Again, under small-signal operation, the amount of water in the gel will not vary significantly and the mass will be approximately constant. The spring constant depends on the elasticity of the polymer strands as well as the density of the network. Network density is a function of the volume of the gel, and should remain relatively unchanged under small volume changes. The elasticity of the network depends on the temperature. The elasticity

goes to zero at 0 Kelvin and is directly proportional to the temperature. Around the transition temperature of NIPA of approximately 34°C, even a temperature swing of 20°C will result in less than a 10% change in the elasticity. It is reasonable to assume that the linear spring in the model is a good model of the elastic behavior of the PGU [112] [22].

6.2.5 Complete Polymer Gel Model

The mathematical models developed in the previous two sections for the two parts of the gel system can be combined using the interconnections in the overview in Figure 6.3. The overall model is fourth order, with each of the gel model components, the AGU and the PGU, accounting for two poles. External inputs to the model are the temperature $T(t)$ and the external force $F_2(t)$. The internal interaction between the forces creates coupling between the systems. The output of the AGU, $F_1(t)$. is an input to the PGU. The effect of external force on the diffusion system is the external input $F_2(t)$ exciting the slow states $x(t)$ and and $F_1(t)$. This model can be summarized in the equations:

$$\begin{bmatrix} \dot{x} \\ \dot{F}_1 \\ \dot{y} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & -c & 0 & 0 \\ 1 & -b & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & \frac{1}{M} & -\frac{k}{M} & -\frac{B}{M} \end{bmatrix} \begin{bmatrix} x \\ F_1 \\ y \\ v \end{bmatrix} + \begin{bmatrix} a & m \\ d & n \\ 0 & 0 \\ 0 & -\frac{1}{M} \end{bmatrix} \begin{bmatrix} T \\ F_2 \end{bmatrix} \tag{6.26}$$

The nonzero term in the fourth row, second column of the state dynamics matrix represents the series connection by which the force due to the diffusion process is transferred to the mechanical linkage.

Using this mathematical model, the response of the length of the gel $y(t)$ to mechanical and temperature excitation can be predicted. By changing the temperature of the gel, $T(t)$, the diffusion process is initiated, inducing a change in the force $F_1(t)$. The dynamics of this process are governed by the eigenvalues of the block matrix $\boldsymbol{A}_D$:

$$\boldsymbol{A}_D = \begin{bmatrix} 0 & -c \\ 1 & -b \end{bmatrix}. \tag{6.27}$$

These dynamics are slow relative to the dynamics of the mechanical system. The force $F_1(t)$ is fed into the mechanical system. As a result, the position $y(t)$ is excited and a change in

the length of the gel is produced by changing the temperature control variable. Another manner in which volume change can occur is through the input $F_2(t)$. This effect is captured by the input matrix terms m and n. The same slow eigenvalues of $\boldsymbol{A}_M$ the diffusion process can be excited by $F_2(t)$ through these input terms. The mechanical structure of the gel has much faster modes of response. The natural frequencies of this system component are the eigenvalues of the matrix $\boldsymbol{A}_M$:

$$\boldsymbol{A}_M = \begin{bmatrix} 0 & 1 \\ -\frac{k}{M} & -\frac{B}{M} \end{bmatrix} \tag{6.28}$$

This system's dynamics are very fast compared to the dynamics of the diffusion system. These modes may be excited by F_1 or the external force F_2. This force is a load or disturbance input that can be applied directly to the mechanical linkage of the gel system, just as it can directly affect the diffusion system.

The overall system with both components is illustrated in Figure 6.12. Simulations of step responses of the system are shown in Figures 6.13 and 6.14. The simulation parameters do not describe a particular physical system. Parameters of the diffusion system were chosen to display two different exponentials in the step response. Parameters of the mechanical system were chosen to produces damped oscillations. Gels composed of NIPA and PVA behave in this manner when pulled and released. The mechanical time constants were chosen to be much shorter than the diffusion time constants. Note that the response of the model is qualitatively similar to the gel behavior described above and observed in the literature. A step in temperature produces a slow change in $F_1(t)$ due to the volume change by network solvent diffusion. The fast mechanical modes are not observed in this response. The length of the gel $y(t)$ just tracks the slowly changing $F_1(t)$. The step in F_2 excites all modes of response. On a short time-scale, the mechanical system oscillations are observed and the gel seems to settle to a new length. The gel has been pulled by the external force, changing $y(t)$ without changing the volume of the gel. On a much longer time-scale, if the application of F_2 persists, a volume transition causing F_1 to change will result in further change in $y(t)$.

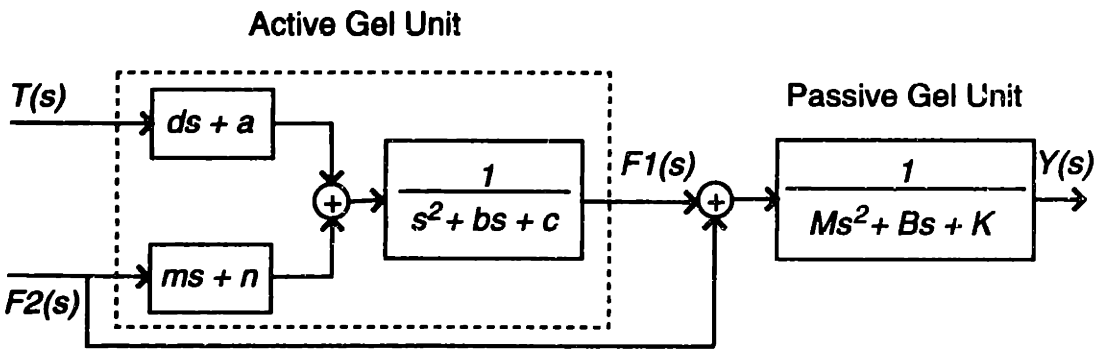


Figure 6.12: Complete model of polymer gel actuator

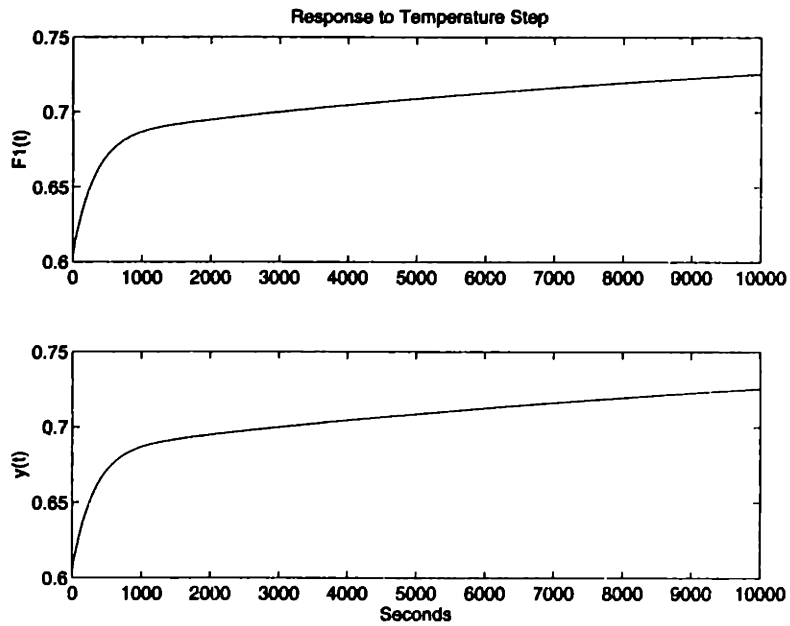


Figure 6.13: Simulation of complete gel model respone to a step in temperature

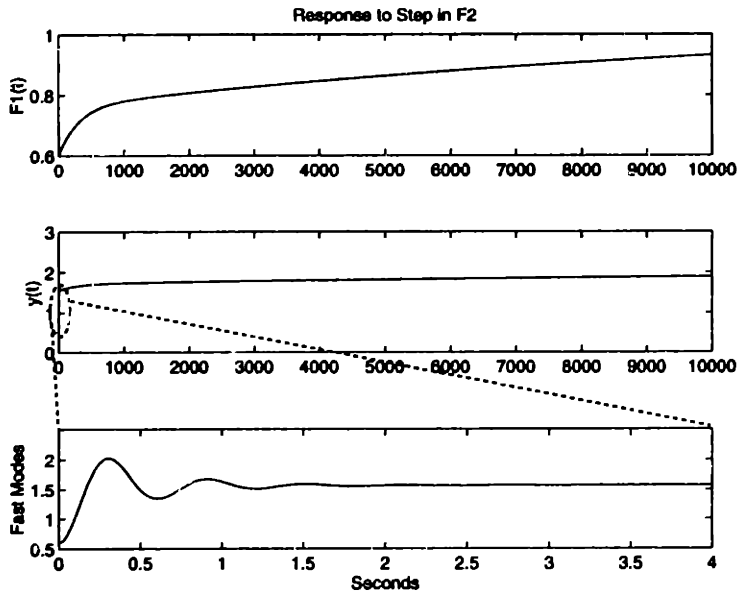


Figure 6.14: Simulation of complete gel model respone to a step in F_2

6.3 Model Summary

The total actuator system consists of the power sources, thermal input system, and polymer gel mechanical transducer. Figure 6.15 illustrates the total system dynamics. Both the thermal dynamics and the gel dynamics may be different depending on the state of the mechanical switch or some other state of the gel or environment. For instance, the gel may exhibit different dynamics depending on the direction of volume transition. This would be reflected in a change in the coefficients of the system polynomials in Figure 6.15. If an averaged system description exists, its state matrices would have the same structure as those describing the shrinking and expanding transitions. The coefficients can be derived using one of the methods in Chapter 4. In the next chapter, this result will be exploited to simplify the derivation of the actuator model. When linear models of the switched systems are known, they can be combined using averaging techniques to derive the averaged model. In the case of the switched polymer gel, no such models exist. However, the averaged model may be derived directly if we assume it has the same form of the system functions presented

above. Parameters of the model can be estimated from *averaged* input and output data. The next chapter identifies an averaged model in this manner and validates it on different experiments.

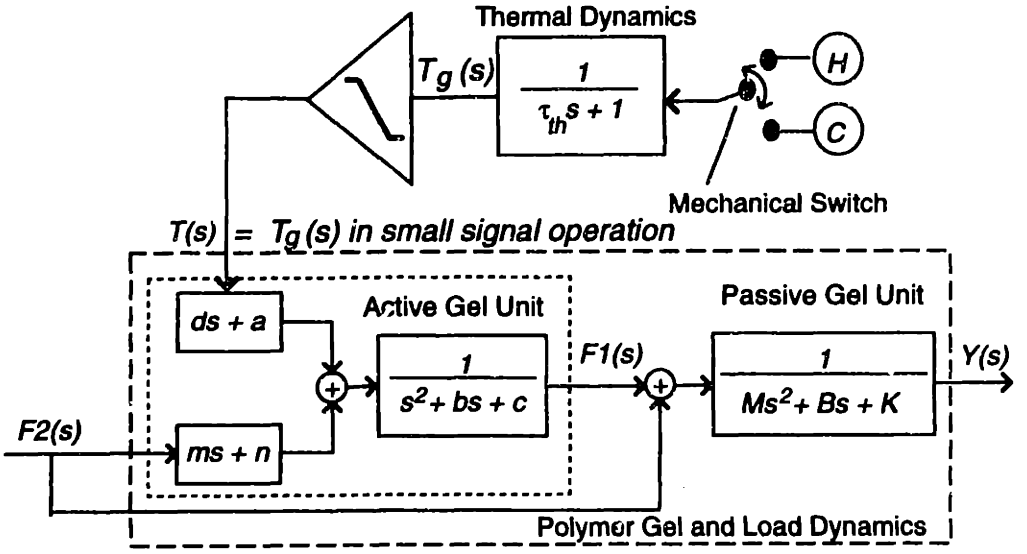


Figure 6.15: Full parameterized gel actuator model

Chapter 7

System Identification and Experimental Results

Using the experimental setup and the parameterized model presented in the previous two chapters, a small-signal dynamic model is developed for the linear-motion polymer gel actuator. With input and output data from the different sensors in the experimental setup, nonlinear parameter estimation algorithms were used to identify the model constants. The precision required in assembling and maintaining the experiment and the assumptions of the model that need to be satisfied experimentally were emphasized in Chapters 5 and 6. The following sections outline the identification process and explain the techniques used in solving the estimation problem. The model is verified through cross-validation and its use for the design and implementation of closed-loop controllers for the polymer gel actuator.

7.1 Response to Thermal Excitation

We will first identify a model for the dynamics of the network/solvent diffusion process as a result of a change in the control variable, the switch command. The input to the actuator system is the switching function from the microprocessor. The thermal system converts this into a temperature profile that is then input to the gel transducer. A change in the temperature produces osmotic forces that cause a change in volume by providing the output F_1 in the gel model. It is the relationship between the output F_1 and the input switching profile that we seek to identify.

7.1.1 Using Gel Temperature as Input - modeling the AGU

In order to achieve the best dynamic response and to avoid the nonlinear effects of total network collapse or swelling, the operating point for the experiments was chosen near the half-way point between the two volume extremes. The temperature rails were centered around the transition temperature of NIPA in order to achieve a wide range of volume states. Under these conditions, the gel mechanical state is very sensitive to temperature changes. This complicated the identification process. Small fluctuations in the thermal system (discussed earlier) made it difficult to predict the temperature with an accuracy greater than 0.5°C based on the switching function. This produced a significant error in the gel volume that made it difficult to identify the system with the switching function as input[1].

In order to eliminate the uncertainty of the thermal model, the directly measured gel temperature was used as the input to the gel model in the parameter estimation problem. Any unmodeled fluctuations in temperature will be captured in the input measurement. In addition, the model developed will primarily describe the dynamics of the network/solvent diffusion process regardless of the input mechanism. Such a general model can then be easily adapted to represent other gels with different input mechanisms. The input model will vary depending on the implementation of the actuator. The object of the following sections is therefore to identify a model for the Active Gel Unit, and not the whole system from the switching function as input to output force F_1. Parameters of the thermal system are later estimated for the purpose of implementing closed-loop control. This will complete the description of the system with the switching function as input and the force F_1 as output.

[1]For control design and practical uses of this actuator, the uncertainty in the thermal system is not as significant. First, a more compact and speedy power delivery system will need to be constructed for a practical application. Second, smart control of the actuator (high gain closed-loop control or adaptive control) can be used to reject the error due to this uncertainty.

7.1.2 Direct Identification of Averaged Model

When a time-varying, switched system can be described at any instant in time by an LTI model, state-space averaging can generally be used to develop an averaged model of the system that is time-invariant. Even if the underlying switched systems are not LTI, linearized models can be developed and averaged, or averages of the nonlinear dynamics can be developed [107] and used to model the *aggregate* behavior of the system. In the case of the polymer gel, the complicated partial differential equation models of a gel are not convenient for averaging. Development of simpler models for both the shrinking and swelling transitions was considered. However, this is difficult for at least two reasons:

- In order to model the shrinking transition, for instance, a positive step in temperature may be applied and the shrinking dynamics observed. However, to stay within the small-signal limits of operation, this step in temperature would have to be extremely small. Because of the non-linearity of the gel behavior with temperature, the shrinking response due to a small step in temperature is different from the response to the larger temperature variation during switched-mode operation. The average behavior of the gel is a result of the latter, and not the former.

- Identification of the shrinking and swelling dynamics from the switching ripple was also considered. The size of the ripple produced poor signal to noise ratio, however, since a relatively high switching frequency was chosen to minimize the ripple in the position variable.

A different method was used to determine a model for the averaged dynamics of the polymer gel network/solvent diffusion process. Assume that the AGU is well-modeled under small-signal operating conditions by Equation 6.22 in both shrinking and swelling volume transitions. (The parameterized description is the same for both processes but the model parameters may be different.) The shrinking dynamics of the AGU due to a temperature input are:

$$\begin{bmatrix} \dot{x} \\ \dot{F}_1 \end{bmatrix} = \begin{bmatrix} 0 & -c_1 \\ 1 & -b_1 \end{bmatrix} \begin{bmatrix} x \\ F_1 \end{bmatrix} + \begin{bmatrix} a_1 \\ d_1 \end{bmatrix} T. \tag{7.1}$$

The swelling dynamics are:

$$\begin{bmatrix} \dot{x} \\ \dot{F}_1 \end{bmatrix} = \begin{bmatrix} 0 & -c_2 \\ 1 & -b_2 \end{bmatrix} \begin{bmatrix} x \\ F_1 \end{bmatrix} + \begin{bmatrix} a_2 \\ d_2 \end{bmatrix} T. \tag{7.2}$$

Now denote the matrices:

$$\boldsymbol{A}_1 = \begin{bmatrix} 0 & -c_1 \\ 1 & -b_1 \end{bmatrix} \quad \boldsymbol{B}_1 = \begin{bmatrix} a_1 \\ d_1 \end{bmatrix}, \tag{7.3}$$

and for the swelling system:

$$\boldsymbol{A}_2 = \begin{bmatrix} 0 & -c_2 \\ 1 & -b_2 \end{bmatrix} \quad \boldsymbol{B}_2 = \begin{bmatrix} a_2 \\ d_2 \end{bmatrix}. \tag{7.4}$$

We further denote the input T by u. This input waveform is the temperature of the polymer gel. It is a periodic waveform with significant deviation from its average value over the course of a switch cycle. The assumption that the input frequency content is much lower than the switch frequency is not valid. However, as discussed earlier in Chapter 4, an averaged model may still exist in this case. Specifically, refer to Figure 4.14 and Sections 4.4.3 and 4.5.3 for discontinuous and continuous volume-phase transition gels, respectively. In general, the averaged model of the gel will be:

$$\dot{\bar{\boldsymbol{x}}}(t) = (D\boldsymbol{A}_1 + (1-D)\boldsymbol{A}_2)\bar{\boldsymbol{x}}(t) + (D\boldsymbol{B}_1 + (1-D)\boldsymbol{B}_2)\bar{u}(t), \tag{7.5}$$

where D is the duty ratio for the switch function which is high when the gel is shrinking (in state 1 with system matrices $\boldsymbol{A}_1$ and $\boldsymbol{A}_2$) and low when it is swelling (in state 2)[2]. The system in Equation 7.5 is time-invariant. With a constant duty ratio, it is also a linear system. However, when $D(t)$ changes as a function of the states and inputs, in feedback control for instance, the system is nonlinear:

$$\dot{\bar{\boldsymbol{x}}}(t) = F(\bar{\boldsymbol{x}}(t), \bar{u}(t), D(t)). \tag{7.6}$$

[2]In the case of continuous gels, where the input matrix has to be the same for both systems in order for an averaged model to exist, $\boldsymbol{B}_1 = \boldsymbol{B}_2 = \boldsymbol{B}$.

Linearization of this system around a nominal operating point will result in a linear system description of its small-signal dynamics in response to incremental changes in input. For clarity, the overbar indicating an averaged value and the time argument t will be suppressed in the following analysis. Thus, $\boldsymbol{x}$ and u will denote the *averaged* states and input levels. Denote the equilibrium or nominal quantities of state, duty ratio, and input as $\boldsymbol{X}$, D_o, and U, respectively. Deviations from these nominal values are then defined as:

$$\tilde{\boldsymbol{x}} = \boldsymbol{x} - \boldsymbol{X}, \qquad \tilde{u} = u - U, \qquad \tilde{D} = D - D_o. \tag{7.7}$$

At the nominal point $(\boldsymbol{X}, U, D_o)$, a linearized model is derived using the partial derivatives of the nonlinear function with respect to the different inputs and states:

$$\dot{\tilde{\boldsymbol{x}}} \approx \tilde{\boldsymbol{x}} \left[\frac{\partial F}{\partial \boldsymbol{x}}\right]_{(\boldsymbol{X},U,D_o)} + \tilde{u} \left[\frac{\partial F}{\partial u}\right]_{(\boldsymbol{X},U,D_o)} + \tilde{D} \left[\frac{\partial F}{\partial D}\right]_{(\boldsymbol{X},U,D_o)}, \tag{7.8}$$

where

$$\left[\frac{\partial F}{\partial \boldsymbol{x}}\right]_{(\boldsymbol{X},U,D_o)} = \boldsymbol{A}_1 D_o + \boldsymbol{A}_2(1 - D_o) = \boldsymbol{A}_{av}, \tag{7.9}$$

$$\left[\frac{\partial F}{\partial u}\right]_{(\boldsymbol{X},U,D_o)} = \boldsymbol{B}_1 D_o + \boldsymbol{B}_2(1 - D_o) = \boldsymbol{B}_{av}, \tag{7.10}$$

and

$$\left[\frac{\partial F}{\partial D}\right]_{(\boldsymbol{X},U,D_o)} = (\boldsymbol{A}_1 - \boldsymbol{A}_2)\boldsymbol{X} + (\boldsymbol{B}_1 - \boldsymbol{B}_2)U = \boldsymbol{B}_d. \tag{7.11}$$

This result is a general one for closed-loop switched systems controlled by PWM (changing the duty ratio of a fixed period pulse). It has been derived and used in modeling power electronic circuits [45]. The state matrices of the averaged model above, $\boldsymbol{A}_{av}$ and $\boldsymbol{B}_{av}$ are weighted averages of the state matrices of the shrinking and swelling volume transitions. Furthermore, it was determined that, for reasons discussed in Appendix D, the two inputs $\tilde{u}$ and the duty ratio $\tilde{D}$ can be lumped into one. The new small-signal averaged model for the gel has the form:

$$\begin{bmatrix} \dot{\tilde{x}} \\ \dot{\tilde{F}}_1 \end{bmatrix} = \begin{bmatrix} 0 & -c \\ 1 & -b \end{bmatrix} \begin{bmatrix} \tilde{x} \\ \tilde{F}_1 \end{bmatrix} + \begin{bmatrix} a \\ d \end{bmatrix} \tilde{u}, \tag{7.12}$$

where the overbar denoting average value has been suppressed and the input matrix represents the combined effects of the duty ratio and the temperature input on the dynamics. The state and input variables in the equation above represent averaged values. By averaging the input temperature and the output F_1 measured in the experiment, the model parameters can be identified. Direct estimation of the averaged model parameters is used below to identify the polymer gel AGU averaged model.

7.1.3 Experiments and Identification

The input/output relationship for the AGU describes the osmotic force $F_1(t)$ (output) produced by temperature $T(t)$ (input). Using a force sensor to measure the output profile would require additional hardware. Instead, the position sensor was used to determine the osmotic force. This was possible due to the significant time-scale separation between the mechanical modes (PGU) and the diffusion modes (AGU). If the suspended load is constant, and the gel is only excited by slow changes in temperature, the mechanical linkage behaves as a rigid structure on the time-scale of the temperature-induced response. As a result, the position of the load $y(t)$ will track the profile of the force due to the volume change $F_1(t)$. Indeed, since the PGU is essentially always in equilibrium, the following approximate relationship between the osmotic force and the position of the load holds:

$$y(t) \approx \frac{F_1(t)}{k}. \tag{7.13}$$

The linearity of the spring constant k was addressed earlier in the modeling sections. The constant-load experiments were thus used to identify the second order model for the diffusion process with temperature input. By separately identifying the diffusion process and the mechanical modes, the problem is simplified to two second-order systems instead of one fourth-order system.

To identify the second-order diffusion system, several experiments were carried out. Sinusoidal excitation at different frequencies to identify the frequency response of the system was considered. However, with the long response time of the gel, sinusoidal steady-state would require very long experiments. The same problem exists with a step excitation.

Theoretically, a step input is rich enough in harmonic content to expose all the natural frequencies of the system. Again, the time it would take for the response to settle and expose all the dynamics of the system would be approximately one day. The quality of the experiment and control of the different variables that can influence the behavior of the gel, the sensors, and other components of the setup may be compromised over the length of time it would take to carry out several of these experiments. A different scheme for the identification is therefore used.

DC gain Identification

First, the DC gain of the gel around the operating point is determined using closed loop control. In order to achieve shorter settling times for the load position, closed-loop control was used. With knowledge of the general structure of the model for the gel diffusion process, a proportional-accumulator (PA)[3] controller was experimentally designed to yield reasonable response times and to regulate the average volume of the gel under constant load to small deviations around the nominal operating point. Details of this process are described in Appendix D. The microprocessor was programmed with PA gains to control the sampled position signal from the LVDT to a digital reference. The digital reference was chosen to represent a point near the middle of the working range of the gel, as described earlier. By observing the steady-state gel position and temperature as shown in Figure 7.1, one point on the small-signal equilibrium curve can be obtained. The PA controller was used to ensure perfect tracking of the reference, thereby enabling regulation of the gel to the chosen nominal point. Several reference points can then be used to determine the small-signal DC gain of the gel system as shown in Figure 7.2. By acquiring the DC gain in this manner, it can be used as an additional constraint to guide the system identification algorithm in determining the coefficients of the system transfer function as described below. This constraint is:

[3]With the digital controller implemented using the microprocessor, proportional accumulator algorithms were used, as opposed to proportional integral controllers for continuous-time or analog implementations.

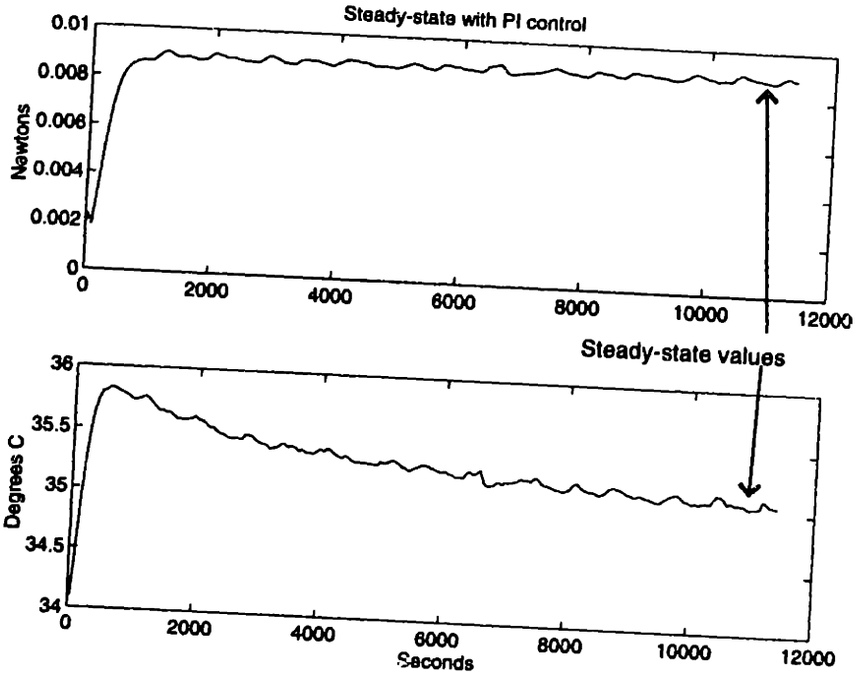


Figure 7.1: Example of PA control response used for equilibrium behavior measurement.

$$\frac{a}{c} \approx .0083. \tag{7.14}$$

System Dynamics

To determine the coefficients of the transfer function of the thermally induced diffusion process (Equation 6.22), several different types of experiments were used in addition to the steady-state experiments:

- **Partial Step Responses:** Partial step responses can expose some of the dynamics of the gel diffusion process without demanding the time it would take to complete a full step response. By stepping the duty ratio of the square-wave command into the mechanical switch, the average temperature profile is sent to a new equilibrium. The dynamics of the thermal system are faster than those of the polymer gel volume

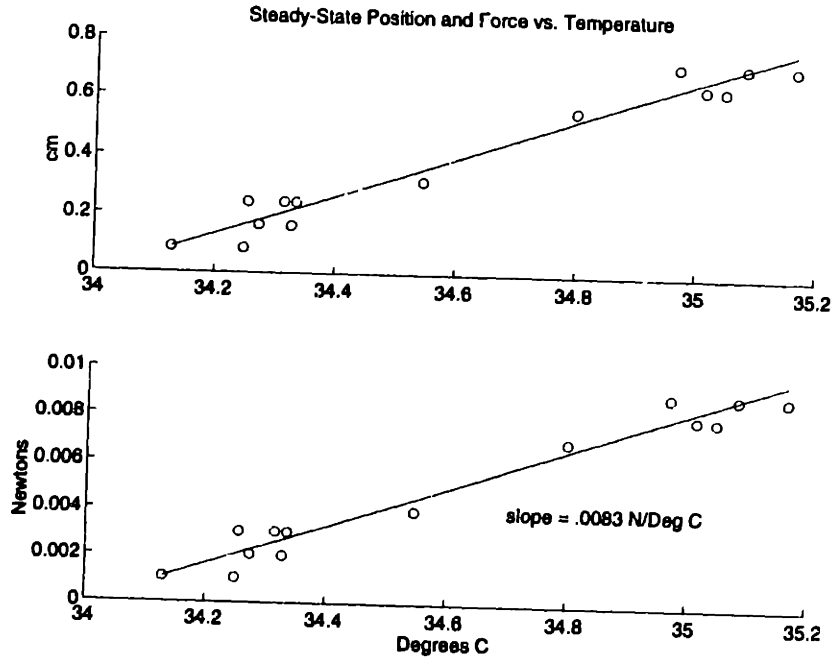


Figure 7.2: Identification of DC gain using various steady-state measurements.

change. Over the time-scale of the response of the gel, the average input is a step.

- **Closed Loop Responses:** By varying the gains in the closed-loop controller used to obtain the steady-state data described earlier, different temperature profiles can be generated. Steps in the digital reference then produced dynamics in the input that were determined by the location of the closed-loop poles. In this manner, different frequencies of excitation can be generated. The input and output data from these experiments was used in both the identification process and the model verification.

- **Periodic Signals:** Again using the closed-loop setting and periodically changing the reference, the response of the gel to inputs with varying frequency content can be observed. Use of the closed loop controllers allowed the system to reach periodic steady-state more quickly than the open-loop natural response of the gel would permit. These input/output profiles were used for verification of the model as described in the following section.

Using the input (temperature) profiles from these experiments, a simulation of the system output can be produced with a set of parameters for the transfer function. Define the output of the simulation F_{1s} using the transfer function:

$$\frac{F_{1s}(s)}{T_e(s)} = \frac{ds + a}{s^2 + bs + c} \tag{7.15}$$

for the model described by the set of parameters (a, b, c, d). The measured experimental temperature $T_e(t)$ is used as the input in the simulation. The set of models defined by different parameters (a, b, c, d) can then be searched for the model that best predicts the experimental output F_{1e}. The implementation of the search algorithm is described below.

Parameter Estimation

Using the information from the dynamic responses and the DC gain measurements, the parameters of the system function were identified. The responses in Figure 7.3 were used along with the steady-state measurements for this purpose. The output of a linear system is a nonlinear function of the coefficients of the transfer function polynomials [108]. To solve the nonlinear parameter estimation problem, the MATLAB ^{TM} function *fmins* [4] was used to minimize the following cost function over the space of the parameters (a, b, c, d). The solution is the parameter set defining the transfer function description of the small-signal polymer gel actuator. The first component of the cost function reflects the DC-gain information in Equation 7.14:

$$C_1 = (\frac{\hat{a}}{\hat{c}} - .0083)^2. \tag{7.16}$$

The force F_{1e} measured with the LVDT position sensor is compared to the force simulated using the model with parameters (a, b, c, d):

$$C_2 = \sum_{i=1}^{i=3} (F_{1ei} - F_{1si})^T (F_{1ei} - F_{1si}), \tag{7.17}$$

[4] *fmins* uses the simplex method.

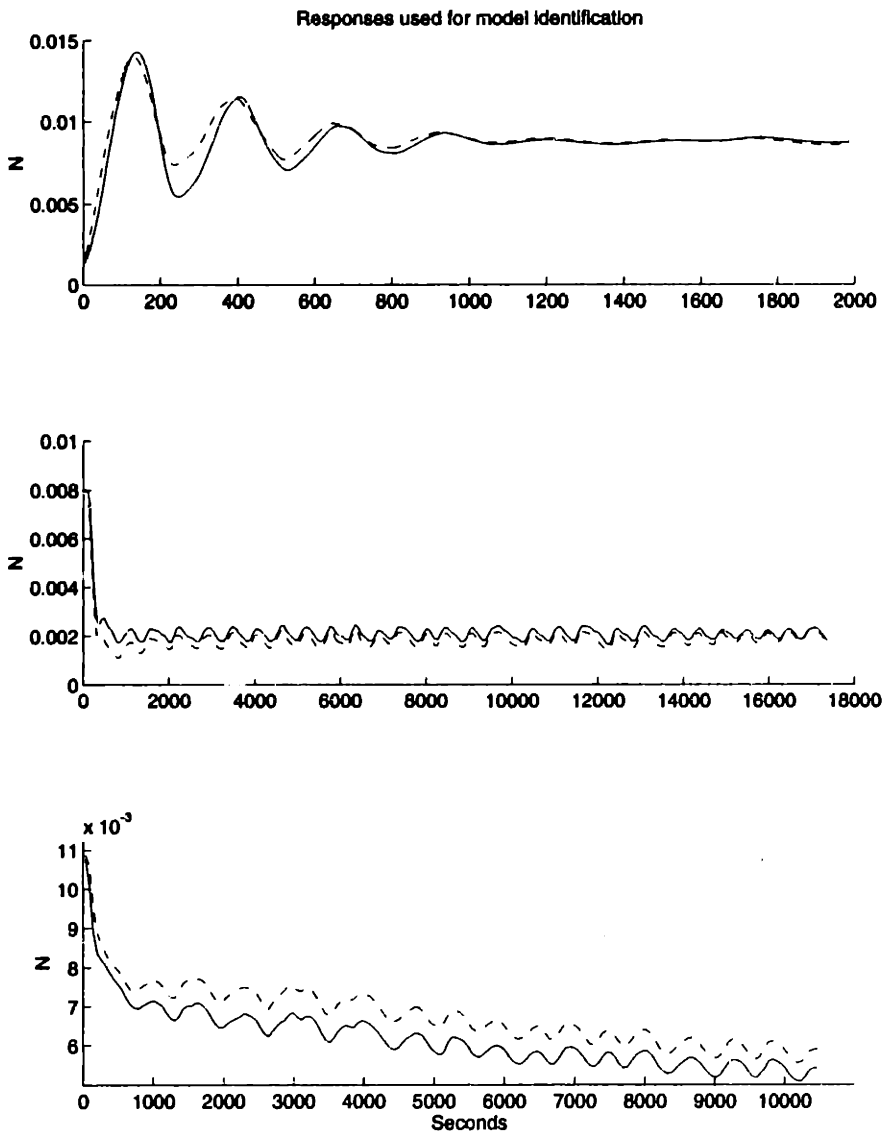


Figure 7.3: Dynamic responses used for system identification. Solid lines are the experimental output. Dashed responses are generated by simulation of the identified model with experimental temperature as input.

where F_{1ei} is the experimental, or measured force for the ith experiment depicted in Figure 7.3 and F_{1si} is the simulated output force using the measured temperature for the ith experiment as input. The following combination of c_1 and c_2 resulted in a gel model that predicted the three responses above and was verified further as discussed in the next section.

$$C = 10C_1 + C2. \tag{7.18}$$

Minimizing this cost over the space of the system parameters:

$$(\hat{a}, \hat{b}, \hat{c}, \hat{d}) = min(C(a, b, c, d)) \tag{7.19}$$

produced the following numbers:

$$\hat{a} = 3.5166e-7, \hat{b} = 2.3497e-2, \hat{c} = 1.43525495e-6, \hat{d} = 2.811e-3. \tag{7.20}$$

Details of the parameter estimation and MATLAB code implementation are presented in Appendix G.

7.1.4 Cross-Validation

In order to verify the validity of the model, additional experiments with different input profiles were conducted. Figure 7.4 illustrates the AGU model's prediction of the gel response to steps in temperture. Figures 7.6 and 7.5 illustrate responses due to input generated using different PA and proportional gains in a closed-loop setting, respectively. Using a proportional gain controller, periodic temperature input was generated by providing a triangular wave reference of different frequencies. This is illustrated in Figure 7.7. As the figures illustrate, the model accurately predicts the behavior of the gel diffusion process well within approximately a 0.012 N region of operation (about 1 cm of change in length with an operating point $k = 1.25$ N/m). Away fom the operating point, a model with the same structure but different parameter values will be needed.

The frequency response of the AGU model is illustrated in Figure 7.8. The frequency content of the input profiles used to generate the different force profiles shown in Figures 7.4 - 7.7 is shown in Figure 7.9. The model predicts the gel response to input frequencies up to

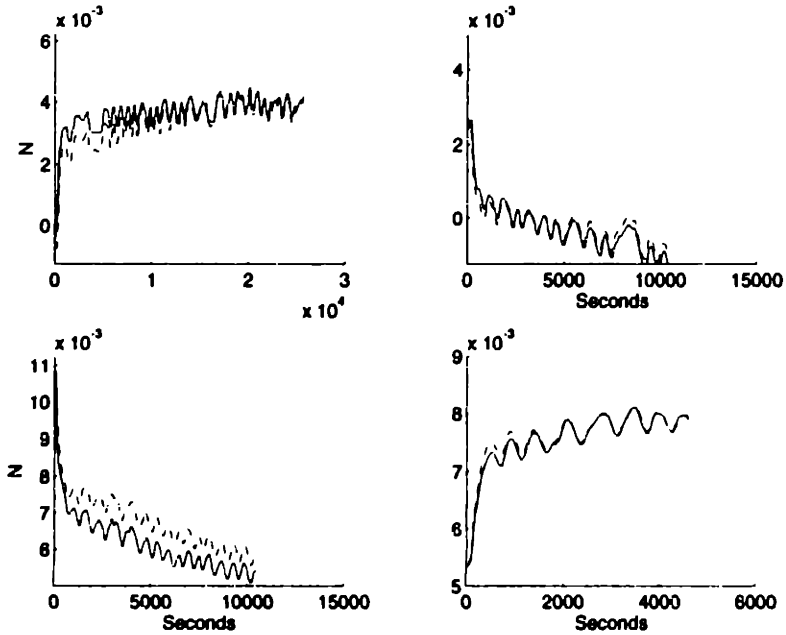


Figure 7.4: Responses due to temperature steps. Solid lines are the experimental output. Dashed responses are generated by simulation of the identified model with experimental temperature as input.

half of the switch frequency, which is expected since high frequency variations were ignored in the averaged model [45] [70] [49].

7.2 Mechanical Excitation

To identify the components of the gel model that are unexposed by thermal excitation, a controlled mechanical disturbance was built into the experimental setup as discussed earlier. Refer again to Figure 5.1. The metallic load attached to the bottom of the cylindrical gel actuator can be attracted with a magnetic field. Two methods for applying this magnetic force were used in the experiments. An automotive solenoid was used to provide a clean step in magnetic field that can be used to expose the fast mechanical modes of the gel. A permanent magnet was used to expose the effect of external mechanical force on the diffusion process or the osmotic forces. The permanent magnet was necessary to avoid pulling 4 Amperes of current into the solenoid for the long time required to observe the

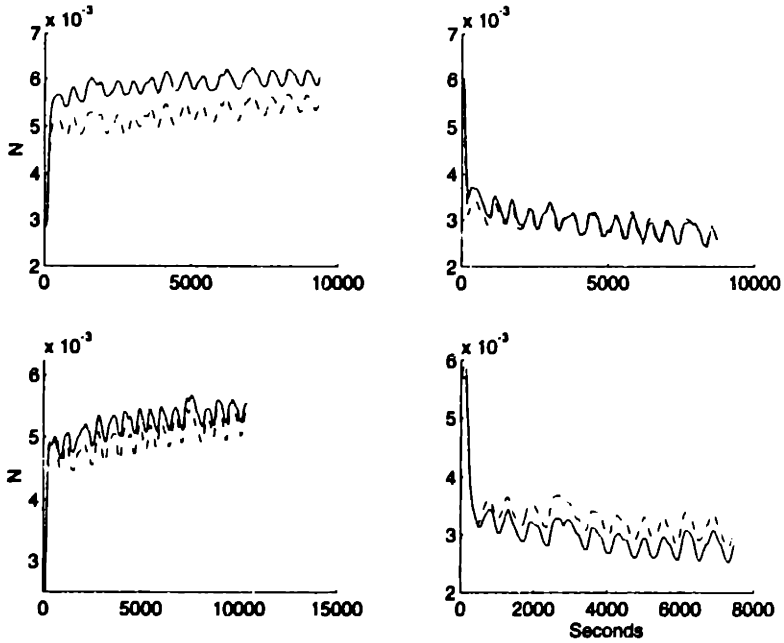


Figure 7.5: Responses under proportional control and steps in reference. Solid lines are the experimental output. Dashed responses are generated by simulation of the identified model with experimental temperature as input.

dynamics of volume change.

7.2.1 Passive Gel Unit

The temperature drive was used to regulate the gel volume and length to a steady-state position around the operating point under constant mechanical load. At that point, the solenoid was turned on to pull in the metal weight and stretch the polymer gel beyond its equilibrium length. A few seconds later, the solenoid was turned off, releasing the gel to respond to initial condition on its length (or position $y(t)$). The gel experienced damped oscillations as it settled at the equilibrium position. During the few seconds that are involved in applying and releasing the external disturbance and the time it takes for the damped oscillations to settle, the volume of the gel can be assumed constant, and consequently also the internal force $F_1(t)$. As a result, only the mechanical modes of the gel participate in the observed output.

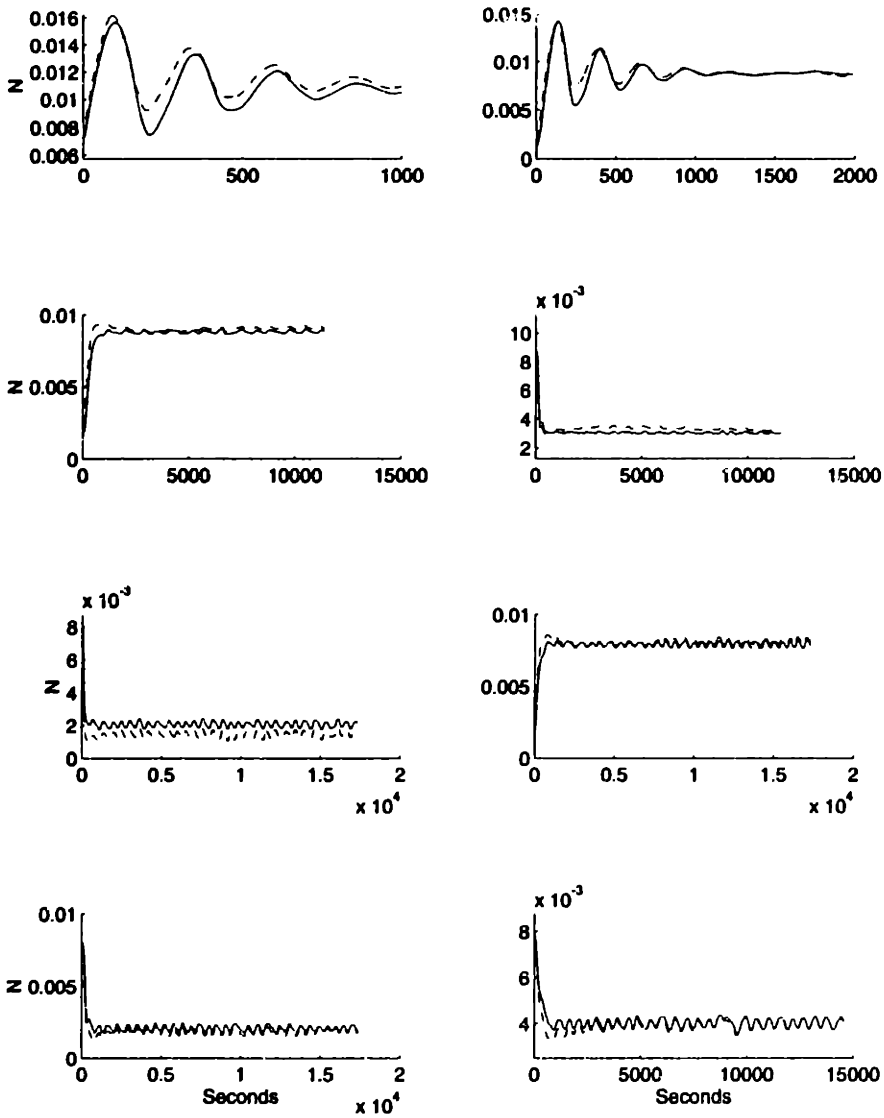


Figure 7.6: Responses Under PA control. Solid lines are the experimental output. Dashed responses are generated by simulation of the identified model with experimental temperature as input.

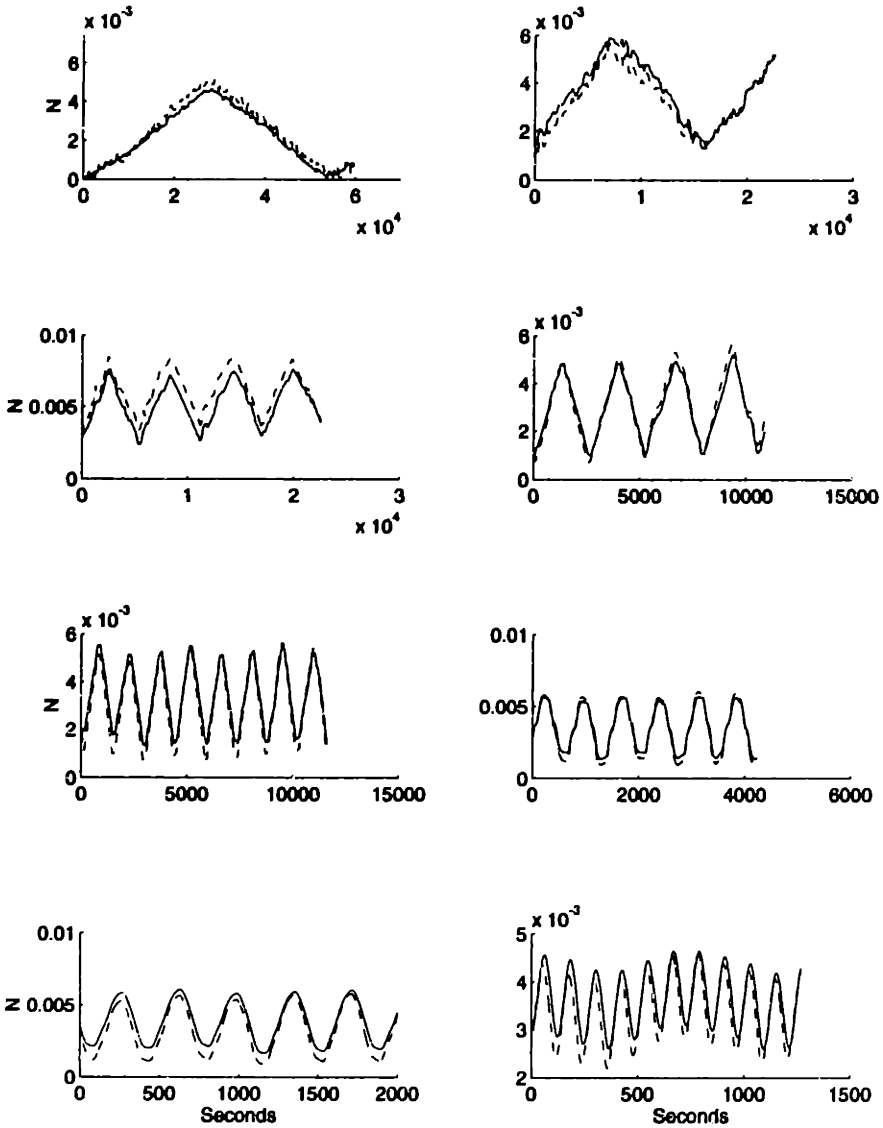


Figure 7.7: Responses due to periodic temperature inputs. Solid lines are the experimental output. Dashed responses are generated by simulation of the identified model with experimental temperature as input.

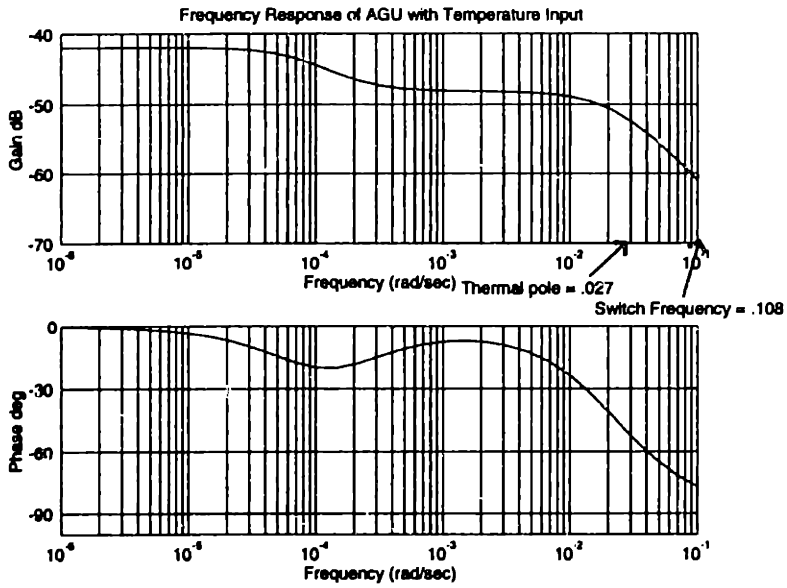


Figure 7.8: Frequency response of the AGU.

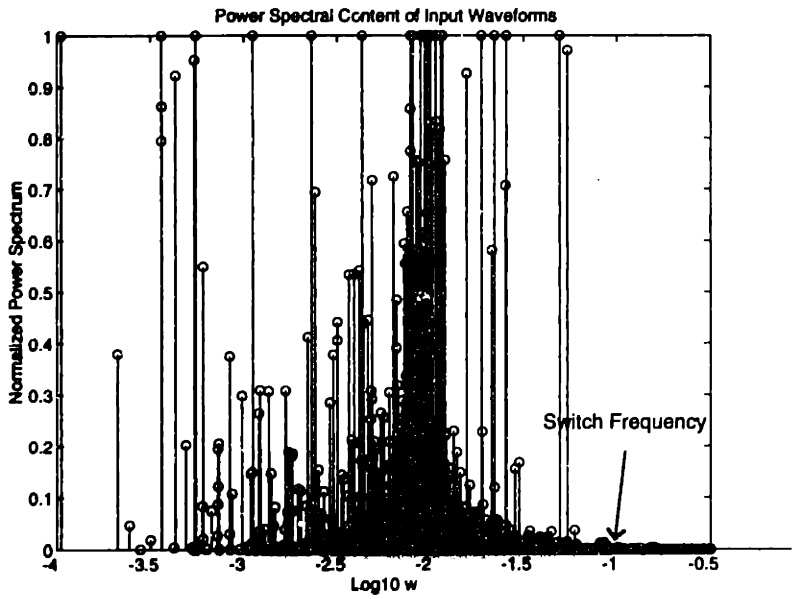


Figure 7.9: Power spectrum of inputs used in model verification.

The spring constant and mass of the gel and moving attachments were measured to verify the physical basis for the model of the PGU. To measure the spring constant, a permanent magnet was used to fix the bottom end of the gel structure. The top end of the gel was then connected to a load using a pulley system as shown in Figure 7.10. A nominal weight was used to stretch the gel to its equilibrium length under nominal temperature conditions. Perturbations in weight were then used to stretch or relax the gel to varying degrees within 0.5 cm of the nominal point. By measuring the deflections in length and the perturbations in weight, an approximate, small-signal linear spring constant can be calculated. The measured spring constant was:

$$k = 1.25\text{N/m} = .0125\text{N/cm}. \tag{7.21}$$

The mass of the moving parts was estimated by weighing the metal and plastic attachments outside water, as well as a dry gel of the same dimensions and chemical composition as the gel used in the experiment. It is important to note here that the water trapped in the gel network should be part of the moving mass. Over the few seconds of the experiment and the response of the mechanical modes, the solvent trapped in the gel does not diffuse into and out of the network structure. Instead, the solvent oscillates along with the rest of the moving parts, and should therefore be considered part of the mass in the mechanical linkage model. Details of this measurement are in Appendix E. The measured mass was approximately:

$$M = 9.745\text{g} = .009745\text{Kg}. \tag{7.22}$$

Figure 7.11 illustrates the observed output when the solenoid is deactivated and the gel is released. A simulation of the PGU model with the measured mass and spring constant is also shown. This simulation was carried out in MATLAB^{TM} with a step in force driving the PGU model. The slow dynamics of the AGU are ignored and the force F_1 is constant throughout the simulation. The value of the damping coefficient identified from the experiment is:

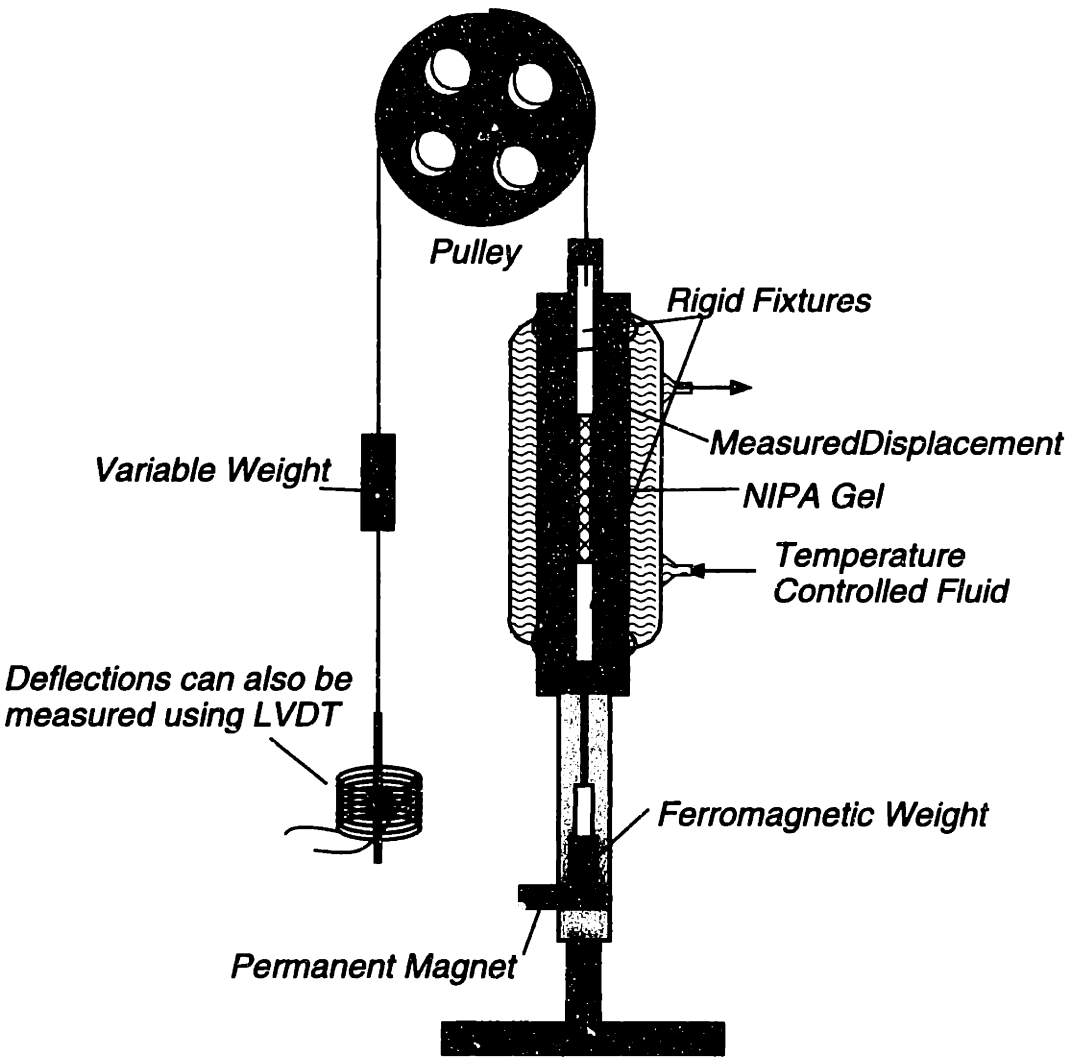


Figure 7.10: Measuring the spring constant of the gel.

$$B = .053\text{Ns/m}. \tag{7.23}$$

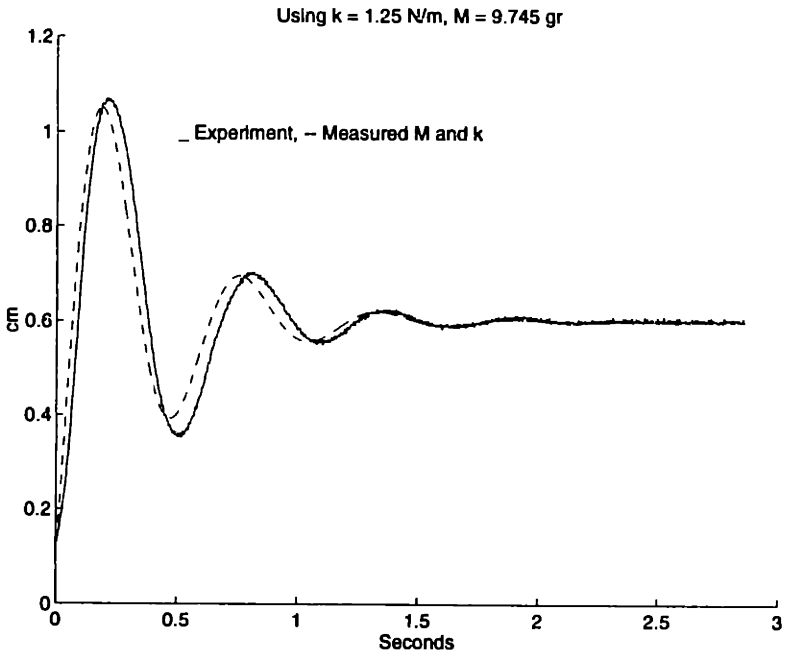


Figure 7.11: Short time response to step in mechanical disturbance.

In order to verify the linearity of the mass and spring constant further the same experiment was carried out with small perturbations in the steady-state temperature and volume. The observed natural frequency and damping of the gel mechanical structure was approximately constant over the range shown in Figure 7.12.

7.2.2 Active Gel Unit

As mentioned earlier, volume change in a gel results from network/solvent diffusion it is driven by a temperature change or external force. The dynamics should exhibit the same natural modes in either case. We have already identified the parameters that determine these modes of response for the diffusion process (b, c). In order to identify the parameters (m, n) to complete the description of the transfer function from external force F_2 to osmotic force F_1, however, the reponse of the diffusion process to mechanical excitation needs to be

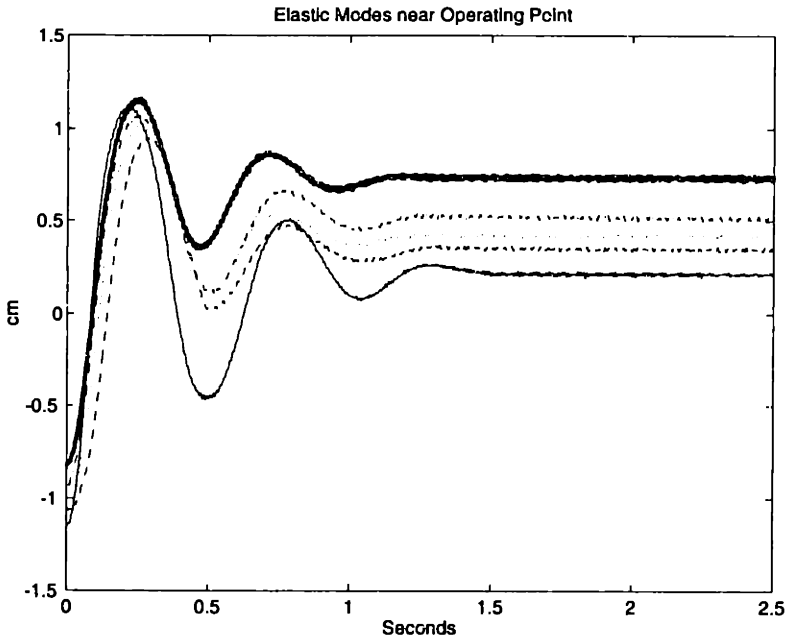


Figure 7.12: PGU response exhibits approximately constant natural frequencies over the small-signal range.

observed. Ideally, different input force profiles $F_2(t)$ would be applied and the output force $F_1(t)$ observed. However, experimental limitations prohibited the application of controlled force levels. With the pole locations of the AGU model identified earlier, the unknown parameters (m, n) of the transfer function:

$$\frac{ns + m}{s^2 + bs + c} \tag{7.24}$$

were determined using the experiment described below.

By stretching the gel to the position of the magnet, mechanical forces can be applied to the system. As the equilibrium volume changes due to this applied force, however, the stretching force varies. We therefore do not know the force profile $F_2(t)$ applied to the gel. However, an approximate value for the coefficients of the transfer function can still be found:

1. Using closed-loop PA control, the gel was made to settle at the nominal equilibrium

point. The duty ratio command was then fixed at its steady-state value and the controller disconnected. Denote the position of the load at this point by y_1.

2. A permanent magnet was used to stretch the gel by attracting and holding the metal weight in place (stretched approximately 1.6 cm). Because the closed-loop controller was disconnected, the temperature command did not change in response to this position error. Under constant temperature, any change in volume that the gel undergoes will be due to the tension exerted by the pulling force of the permanent magnet.

3. The gel length is fixed in this manner for over 36 hours[5], allowing the diffusion modes to complete their response and the gel to attain a new equilibrium. At that point, the gel is released and the new position of the gel is recorded, y_2.

4. The change in gel length $(y_1 - y_2)$ measured was 0.33 cm. At equilibrium, then, the gel was stretched due to an additional applied F_2 by 1.27cm (1.6-0.33). Using the spring constant $k = 1.25$N/m, the final applied external force is 0.015875N[6]. In equilibrium, this resulted in a change in $F_1 = k(y_1 - y_2)$. The DC gain of the transfer function from F_2 to F_1 is therefore:

$$\frac{m}{c} = \frac{0.33}{1.6 - 0.33} \approx 0.26. \tag{7.25}$$

5. Using this DC gain and the fact that the release of the metal weight when the gel is in a state of equilibrium can be modeled as a step response (where the step in F_2 is the difference between the equilibrium force applied at the end of 36 hours and the nominal force applied by the weight of the attachments), the dynamic response can be used to identify a value for the parameter n. As Figure 7.13 illustrates, $n = 0$ provided the best fit for the experimental results. Several sources of error account

[5]The temperature of the gel was recorded during this time to ensure that it maintained a constant average. As mentioned earlier, however, longer experiments are more susceptible to drift effects. This is a potential source of error in the model components identified here.

[6]The extension by the magnet brought the gel outside the region of validity for the model. The external force is probably larger than is calculated using the linear spring approximation, making the DC gain found here larger than the true DC gain.

for the mismatch between the observed experimental output and the simulation. In order to produce enough volume change, the 1.6 cm stretch was needed. This caused the gel to be outside the range for which the model is most accurate for a long time. In addition, the long duration of the experiment makes the possibility of error due to component drift or noise more likely. The difference between the simulation and observed output was on the same order as fluctuations in the LVDT sensor output over the course of 24 hours.

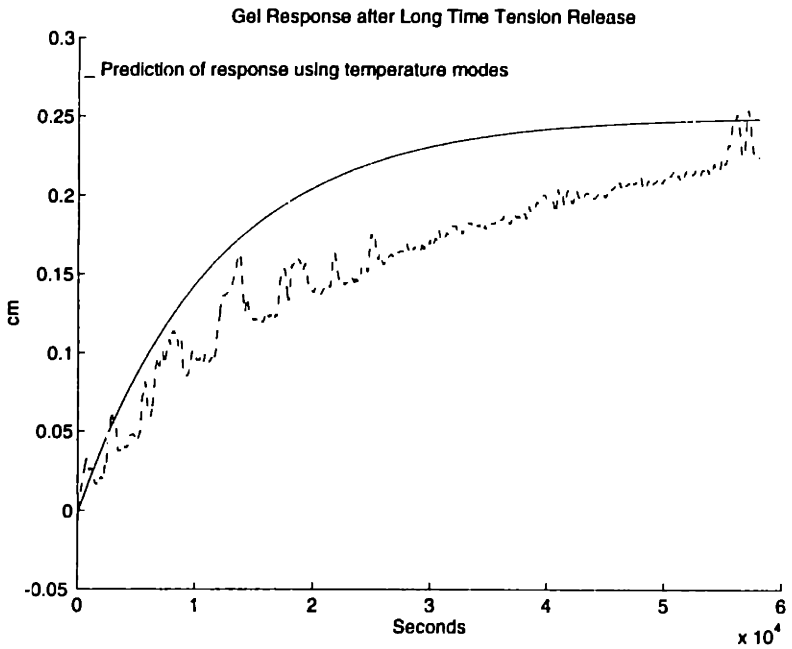


Figure 7.13: Response of slow diffusion modes to mechanical excitation.

The numerical values for the parameters (m, n) are approximate. The primary of this part of the model identification is to demonstrate the existence of this effect of mechanical disturbance on the osmotic pressure in a gel. The resulting dynamics are important to note particularly in a servo-mechanical application where the mechanical load is likely to change. A mechanism for applying arbitrary force profiles and observing the diffusion process through some means of accurate force measurement is necessary to characterize the relationship between external pressure and the osmotic forces of the gel.

7.3 Closed-Loop Control

With the parameter values obtained in the previous sections, the overall model for the polymer gel is:

$$\begin{bmatrix} \dot{x} \\ \dot{F}_1 \\ \dot{y} \\ \dot{v} \end{bmatrix} = \begin{bmatrix} 0 & -1.4353e-6 & 0 & 0 \\ 1 & -2.3497e-2 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1.02617e2 & -1.2827 & -5.4387 \end{bmatrix} \begin{bmatrix} x \\ F_1 \\ y \\ v \end{bmatrix} + \begin{bmatrix} 3.5166e-7 & 3.732e-7 \\ 2.811e-3 & 0 \\ 0 & 0 \\ 0 & -1.02617e2 \end{bmatrix} \begin{bmatrix} T \\ F_2 \end{bmatrix} \quad (7.26)$$

Using this model, closed-loop controllers can be designed to regulate the position of a load. Using a first-order model for the thermal system presented above, a MATLAB ^{TM} simulation of the closed loop system was constructed[7]. A controller provided a duty ratio command that was used to drive the thermal model with a square-wave. The average temperature of the thermal system was used to drive the average model of the gel derived here. Using different controllers, simulations of the system responding to steps in position reference were carried out. The control algorithms were also implemented using the microprocessor in the experimental setup. The output position in the simulation is compared to the output of the experiments in Figures 7.14 and 7.15. Note the zero steady-state error with PA control. Also note the low frequency disturbance on the experimental waveforms. Fluid exchange between the two water reservoirs resulted in low frequency temperature noise. With higher proportional gain and PA control, the disturbance rejection in the system was improved as illustrated in the figure.

7.4 Summary and Discussion of Results

The prototype linear gel actuator system was used as a position regulator using the switching scheme described above. Using several assumptions, a linear time-invariant model for

[7]Details of the control design are provided in Appendices H and D.

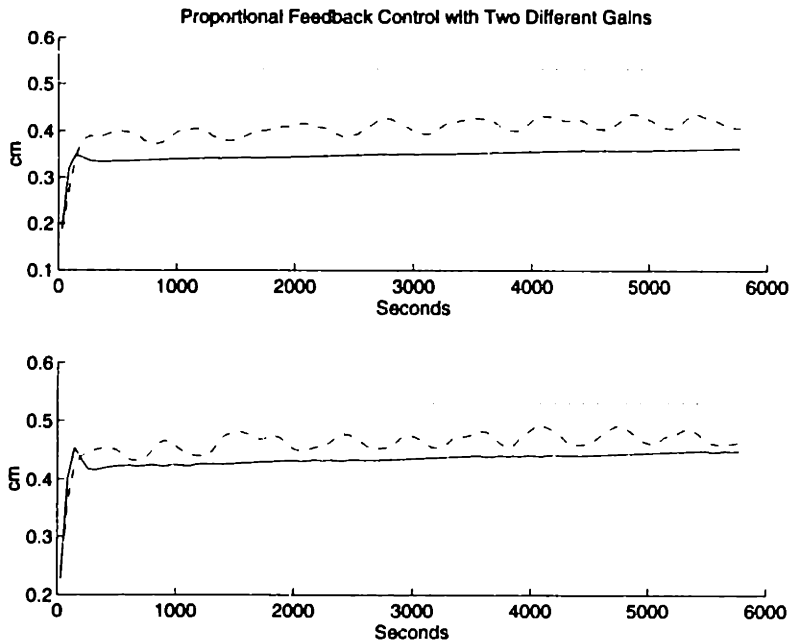


Figure 7.14: Closed-loop control (proportional) based on the polymer gel actuator odel. Short dashes represent the reference, longer dashes the experimental output, and solid lines the simulation.

the average dynamics of the system was developed. First, an LTI description of the diffusion process was assumed for both the shrinking and swelling transitions under small-signal operation. The two transitions have different dynamics, but both exhibit the same qualitative two-rate response observed in the literature and our experiments and modeled here by a second-order LTI model. Second. although large separation between the dynamics of the input and the switching frequency did not exist in the experiment, an averaged model derived from the underlying shrinking and swelling dynamics was assumed to exist. This assumption was possible given the analysis in Chapter 4. Under some special constraints, it was shown that the averaged dynamics for a system with fast input can be modeled by LTI system matrices that are weighted averages of the underlying system matrices. The same model structure assigned to the shrinking and swelling transitions was therefore used to directly identify an *averaged model* for the polymer gel system. The model developed described the gel behavior under small-signal and constant load conditions. The model

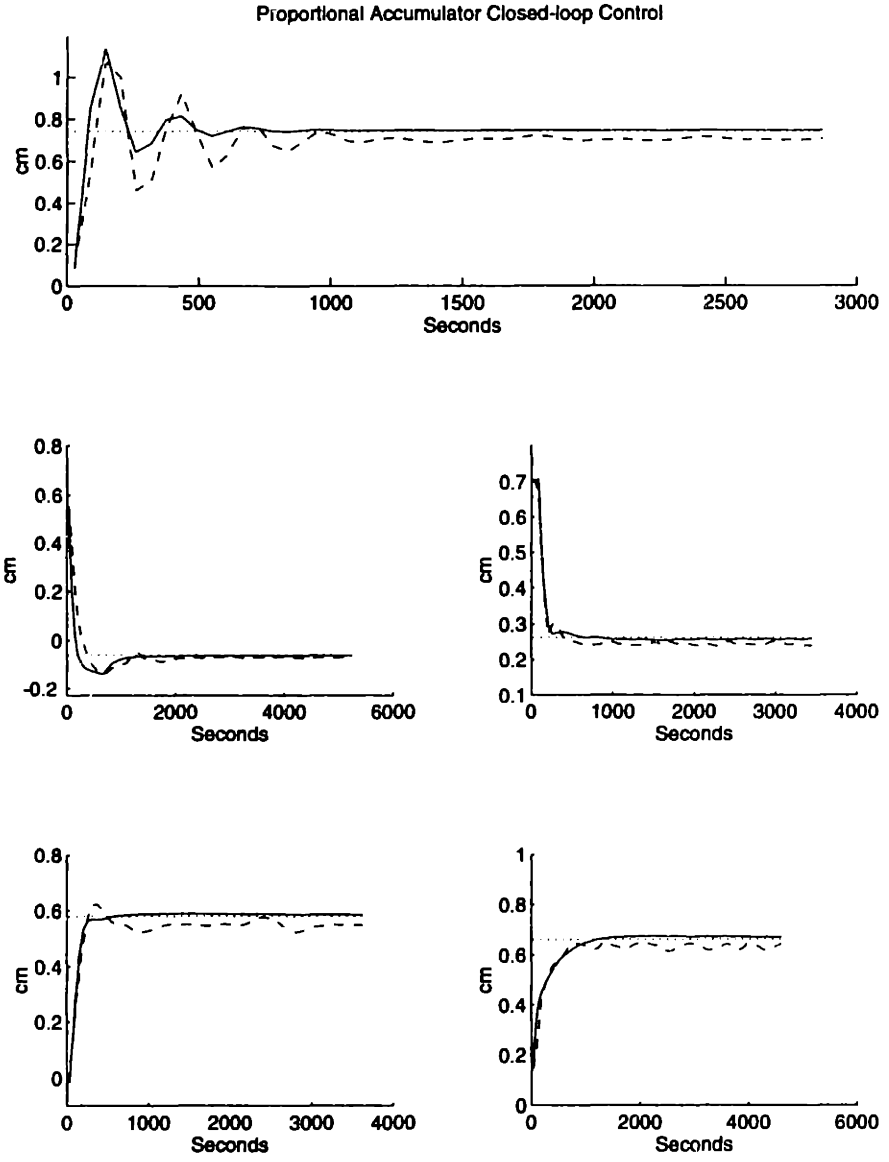


Figure 7.15: Closed-loop control (with accumulator) based on the polymer gel actuator model. Short dashes represent the reference, longer dashes the experimental output, and solid lines the simulation.

validity within approximately 0.5 cm of the nominal averaged gel length and 1.5°C of the nominal averaged temperature was verified using a variety of input profiles. Closed-loop control was further implemented using an approximate LTI model for the input (thermal) system. The experimental results were well-predicted by the simulation. It is important to note that the AGU averaged model was derived and tested using a maximum frequency of excitation of $f = \frac{1}{116}$Hz, or half of the switching frequency. The model is not guaranteed to be accurate for higher frequencies. For practical purposes, higher frequencies can be ignored. The switching frequency was chosen so that switching ripple is significantly attenuated by the system dynamics. Any higher frequencies will be attenuated to a larger degree.

To further validate the model, the settling time of its step response was compared to similar responses observed by other gel researchers. For example, in [69], a NIPA/sodium acrylate gel sphere with radius 0.234mm completes a swelling transition in approximately 500 seconds. As mentioned earlier, the response time of a gel volume transition has been predicted and observed to vary approximately with the square of the critical length, the radius in this case [17]. According to this approximation, a sphere of the same gel studied in [69] with a radius of 3.3mm would have a settling time of:

$$\tau_s \approx \frac{3.3^2}{0.234^2} 500 = 99441\text{s} \approx 27\text{hrs}. \tag{7.27}$$

Compare this to the settling time associated with the gel model, or:

$$\tau_c = \frac{5}{s_d} \approx 17\text{hrs} \tag{7.28}$$

where s_d is the dominant pole in the gel model[8].

While the parameters of the gel model will change depending on the size and chemistry, as well as the operating point of the system, the model structure should stay the same within the vicinity of a given equilibrium. The validity of a model with specific parameters is

[8]The difference in geometry has been ignored in this comparison. Furthermore, the comparison was purposely made with a free-swelling ionic gel. As mentioned earlier, the effect of the mechanical loading can be similar to that of the ionic pressure component in an ionic gel. However, no effort was made to quantitatively relate the two effects.

limited to small deviations from the equilibrium average position (gel length), temperature, and load. For robotic applications, large deviations that are part of an overall trajectory reference are common. Gain scheduling and adaptive control based on real-time model parameter estimation can be used in such a setting. Determination of model parameters at different operating points may aid the development of a large-signal nonlinear model. Such a model can be linearized at different operating points to update parameter values and facilitate adaptive control. It may also lend itself to well-known nonlinear control techniques and should be considered in future work.

The PWM technique and the small-signal averaged model present an engineer with valuable design tools and control methods. By employing switching control, continuous and discontinuous gels can be used equally effectively in servo-mechanical applications. The small-signal model allows the design of stable controllers with good prediction of performance in the vicinity of an equilibrium point. Prior existing models have been too complicated ([88], [23], [52]) because they attempted to model the large-signal behavior and/or included the power delivery mechanism (i.e. thermal or chemical delivery system) in the model. In [52], for example, after the dynamics for the acid/base delivery system were ignored because they are much faster than the diffusion process dynamics, a first-order system was used to describe the network/solvent diffusion. We have found that the second order model has significantly improved predictive power over a first-order one. In addition, none of the prior art seems to address the difficulty of controlling a discontinuous polymer gel.

The time-scale separation between the two components of the model (the diffusion process and the purely mechanical natural response) was helpful in solving the system identification problem. However, it also reflects a property of the gel that will be significant in the design of mechanical systems that use polymer gels. The control variable, or temperature, in the linear actuator system excites the mechanical modes indirectly through the diffusion process. This, coupled with saturation effects of the input command and/or the gel transducer may greatly limit the ability to correct for fast mechanical disturbances. An analogous situation occurs in the human muscular system. For instance, if the elbow joint

is at a right angle and a heavy load unexpectedly impacts the hand, a significant deflection will occur before the muscular system can correct for the angular position. Similarly, because of the elastic nature of the gel, a step in load will produce a large error due to the inability of the gel diffusion process to respond fast enough[9].

One way to avoid the large disturbances due to mechanical load change is to employ compliance control with antagonistic pairs. In the human arm, for instance, the biceps and triceps muscle groups are used to rotate the elbow joint in opposite directions. Moreover, the stiffness of the joint in any angular position can be controlled by applying a "common mode" command to both muscle groups. A system with antagonistic gel actuators is under development and should be the focus of future research. An added benefit to using gels in an antagonistic configuration is one of speed of response. In general, the gel collapses faster than it swells. With antagonistic pairs, the faster shrinking response can be used to produce motion in both directions for one joint[10]. In [1], a muscle model with a fast series elastic component and an active component, similar to the model developed here, is presented. Indeed, the apparent similarities between polymer gels and muscle tissue suggest borrowing techniques for power delivery and control from biological systems. For instance, muscles are bundles of fibers. Recall that polymer gel actuators will need to be constructed out of groups or bundles of smaller units in order to achieve both speed and strength. Such a system presents its own unique challenges of modeling and control, and the polymer gel model developed above should provide a good starting point for tackling these problems.

[9]Theoretically, closed-loop control can increase the bandwidth of the response due to temperature and produce a much more stiff system that responds well to even small deflections. However, saturation limits on the input and the gel transducer may prevent this in a practical setting.

[10]This is only possible due to the elastic nature of the gel and the fast response of the mechanical modes. So while the fast mechanical modes present a problem in a single gel configuration, they appear to facilitate the implementation of an antagonistic pair device.

Chapter 8

Closed-Loop Control of Magnetically-Activated Gels

The previous chapters demonstrated the effective application of switching techniques for regulating the volume of a polymer gel. Using switching drives, even a discontinuous volume-phase transition gel can be controlled to any point between its totally collapsed and totally swollen states. Using the fourth-order model consisting of the combined AGU and PGU, the small-signal behavior of a gel linear actuator can be predicted accurately. Closed-loop control was implemented using this model to achieve stable regulators with different performance parameters. However, this demonstration system has response times and construction tradeoffs that are extremely prohibitive in a practical setting.

- The natural response time of the polymer gel was approximately 27 hours. In order to create a strong, homogeneously cross-linked polymer gel, the NIPA/PVA gel was fabricated inside a micropipette. Spinning gel fibers would have produced thinner, faster responding gels[1]. However, the non-uniform cross-linking and gel dimensions would result in weak points along the gel structure that could cause the fibers to break during experimentation. In addition, bundling of such fine gel fibers would be necessary for providing enough strength to carry the mechanical load and sensor components described earlier. Without the ability to accurately control the properties

[1]Recall that, because the volume-change process is due to network/solvent diffusion, a smaller gel is faster to respond.

of the different components of the bundle, effective load-sharing among the different fibers would be difficult.

- Another limitation on the achievable open-loop speed of the gel actuator was imposed by the size of the load to be carried. The smallest available radius micro-pipette gel could not be used without bundling. The sensor and magnetic load used were too heavy for a single gel at that size. Bundling would be required again. A relatively large gel cylinder strong enough to carry this load was needed.

- In order to achieve low ripple levels, the polymer gel natural time constants had to be much longer than the switch period. However, as discussed earlier, the switching frequency was limited by fluid transport delays and the thermal system dynamics. The slow thermal system in the experiment required an otherwise unnecessarily large gel. Of all the constraints on the gel size, this one was the most limiting, resulting in the extremely slow gel in the experiment.

Even with slow open-loop dynamics, closed-loop control can generally be used to achieve faster settling times. In our experimental system, however, this was limited by the thermal input mechanism[2]. The slow fluid-circulating thermal system is also heavy, bulky, and inefficient. In order to heat the polymer gel network, all of the solvent in the inner chamber of the Liebig condenser has to be heated through the thermal resistance of the glass.

In this chapter, magnetic triggering of polymer gels is introduced to implement a faster, more efficient PWM linear actuator. Higher switching frequencies are achieved since the magnetic field can be switched by controlling an electrical variable (current). Moreover, because the magnetic field produces thermal losses in seed material *inside* the polymer gel, the transport delay in the system is minimized. Since the energy transfer to the gel system occurs through a magnetic field, the system is contactless. It could be used to trigger a gel under the skin or in any remote location where an electromagnetic field can penetrate[3]. A

[2]With the long switch period, the fluid transport delay, and the slow thermal dynamics, the closed-loop controller gain was limited. Recall also that water exchange between the two circulators caused an input disturbance that was minimized by keeping the gap between the two source temperatures small.

[3]I would like to acknowledge Deron Jackson and Timothy Denison for their effort in this work. Deron Jackson designed and built the power electronics to deliver the alternating magnetic field to the polymer

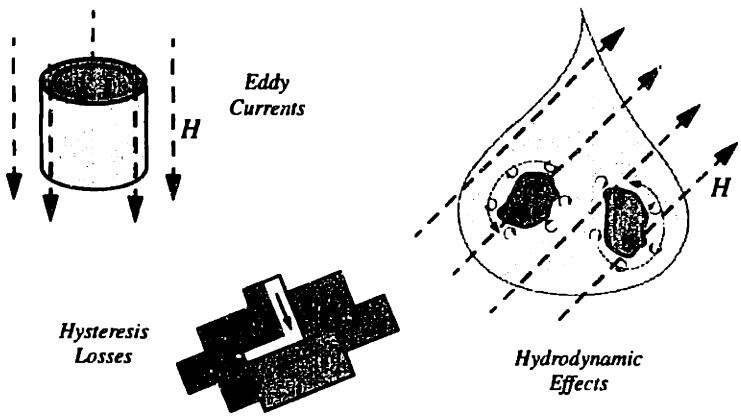


Figure 8.1: Loss mechanisms.

more detailed discussion of this work is provided in [71], [42], [43], and [40].

8.1 Magnetically-Activated Gels

Losses in a ferromagnetic seed material can be generated with the application of a quasi-static [26] magnetic field [24]. By including such ferromagnetic seed material in a polymer gel network, magnetically activated gels were developed in [42] and [53]. By applying an external alternating magnetic field, the thermal losses in the seed material can induce heating of the gel that initiates a volume-phase transition. The dominant loss mechanisms that serve to couple the magnetic field to the gel temperature are generally a function of field frequency, seed material, and geometry. These heating mechanisms, illustrated in Figure 8.1, include eddy current losses, hysteresis losses, and hydrodynamic friction. An overview of the impact of the choice of seed material on gel design is given in [42]. A brief summary is presented below.

gel samples. Timothy Denison helped implement the closed-loop control and data-acquisition for the active control experiments.

8.1.1 Lumped-Seed/Powdered-Seed Gels

A macroscopic seed material such as a steel pin can be used to develop magnetically triggered polymer gels [42]. The seed material could be inserted into the gel during or following gel fabrication. Large-scale seed materials are excellent choices for developing magnetically-activated gels that respond to relatively low field strengths and frequencies, and are easy to model accurately. They are used in [42] to develop good experimental design and analytical models for the energy transfer. However, because lumped seed materials can restrict the swelling ratio of a gel, this seeding approach is not preferred for developing gel actuators with significant stroke. The distribution of a finely powdered seed material within the gel yields a more flexible gel structure with larger stroke. In addition, the distribution of the seed material throughout the gel network results in relatively homogeneous heating. In the experiments presented here, the physical size and shape of the metal particles resulted in significant loss contributions from both eddy currents and hysteresis losses [42]. A gel was fabricated with 5%-by-weight nickel "leafing-grade" flakes provided by Novamet corporation[4]. These flakes have a typical thickness of 0.4 μm and a typical diameter of 30 μm. The gel was formed in a cylinder 0.51 cm in diameter and 7.37 cm in length. Because the relative dimensions of the flakes are small in comparison to those of the gel, the linear stroke of the seeded gel, between its collapsed and swollen states, is comparable to that of a gel without seed particles.

8.1.2 Ferrofluid Solvent

Ferrofluid solvents can also be used to couple the magnetic field to the temperature of the polymer matrix is a ferrofluid solvent. A typical ferrofluid is composed of ferromagnetic particles on the order of 100 Å (0.01 μm) suspended in an aqueous solvent by a surfactant[5]. Ferrofluids can yield significant heating when immersed in an alternating magnetic field [42]. Physical models and experiments suggest that hydrodynamic friction between rotating magnetic domains and the ferrofluid solvent is responsible for the thermal losses [102]. This is

[4]Novamet Specialty Products Corporation, 10 Lawlins Park, Wyckoff, NJ 07481 USA.
[5]Solvents other than water may be used.

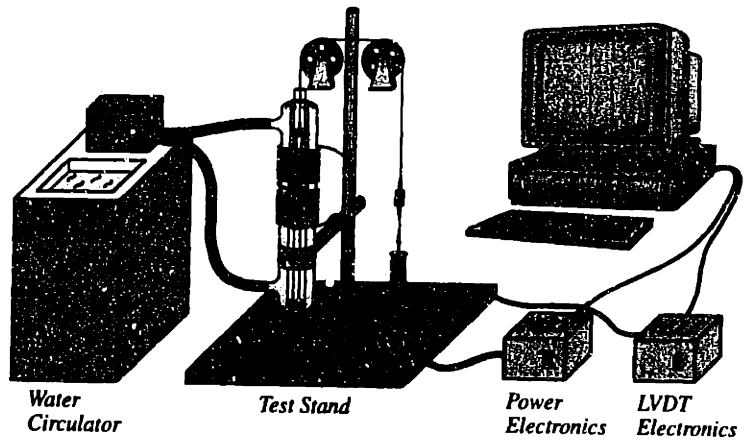


Figure 8.2: Experimental apparatus for magnetic gel experiments.

supported by the finding in [42] that even at field strengths and frequencies (around 3 MHz) where eddy current dissipation and hysteresis losses should be negligible, very substantial heating is observed. A second gel was synthesized using an aqueous ferrofluid as its solvent and used in the active control experiments below.

8.2 Experimental Overview

Experiments were conducted using NIPA/PVA gels seeded with induction heating targets to investigate the dynamic properties and the suitability of these gels as actuators. As mentioned earlier, pure NIPA gels exhibit a nearly discontinuous phase-transition curve around a phase-transition temperature of approximately 34 °C [31]. Flake and ferrofluid seeded NIPA gels exhibit similar phase-transition curves [42].

An apparatus was constructed to regulate the temperature, magnetic field, and mechanical load applied to the gel (Figure 8.2). A linear variable differential transformer (LVDT) was used to measure the length of the gel [61]. The first set of experiments was conducted on a NIPA gel fabricated with Novamet nickel flakes implanted in the polymer matrix. The second set of experiments used a standard NIPA gel immersed in a ferrofluid solvent.

8.2.1 Gel Preparation

The main constituents of the gel were the same as those introduced in the cylinder fabrication technique discussed in Chapter 2. The technique was modified to incorporate the magnetic seeds into the polymer network. The modified process is detailed below. Gels made from recrystallized *N*-isopropylacrylamide (NIPA, Kodak) monomer were prepared by free radical polymerization in water at room temperature. Ammonium persulfate (APS, Mallinckrodt) and tetramethyl ethylene diamine (TEMED, Bio-Rad) were used as the initiator and accelerator, respectively. The crosslinker used was *N,N'*-methylenebisacrylamide (BIS, Bio-Rad). Poly(vinyl alcohol) (PVA, Aldrich) polymer with 115,000 molecular weight was used to form a semi-interpenetrating polymer network (IPN) in order to increase the tensile strength of the gel [72]. In the case of the powdered seed gel, PVA also helped suspend the seed material until gelation occurred.

Powdered Seed Gel

The powdered seed material is composed of Novamet ferromagnetic nickel "leafing grade" flakes. In the powdered seed gel, the ferromagnetic particles are entrained in the matrix during polymerization. This is achieved by suspending the seed material in solution as gelation occurs, effectively trapping the ferromagnetic material in the gel network structure. The polymer gel formula described below is a variation on the NIPA gel formula described by [72]. The preparation steps for this gel are:

1. PVA (5 g) was dissolved in 500 ml of deionized water at 50 °C for approximately 20 hours.

2. APS (0.2 g) was dissolved in 5 ml deionized water at room temperature. This solution was degassed in a vacuum chamber.

3. Nickel flakes (0.5 g) were mixed in 5 ml of the PVA solution from step 1.

4. NIPA monomer (0.78 g), and BIS (0.013 g) were dissolved in 5 ml of deionized water at room temperature. This solution was mixed with the mixture of PVA and nickel

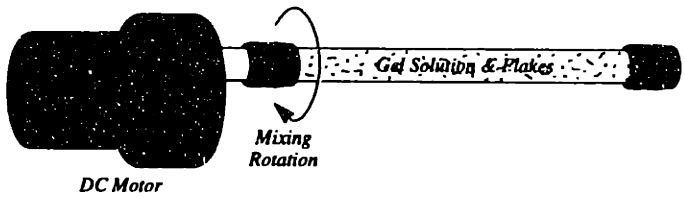


Figure 8.3: Spinning solution to maintain the nickel flakes in suspension.

flakes from step 3. The resulting 10 ml solution was degassed.

5. TEMED (24 μL) and 100 μL of the APS solution from step 2 were mixed with the solution from step 4. The new mixture was then vigorously shaken with a vortex mixer.

6. The solution in 5 was immediately poured into a glass cylinder 0.51 cm in diameter. It is important not to let the nickel flakes settle before transferring the solution.

7. The cylinder was sealed at both ends and mounted onto the shaft of a DC motor. The motor serves to spin the cylinder at approximately 100 rpm horizontally (with the glass cylinder parallel to the ground) to keep the nickel flakes suspended in the solution as gelation occurred. This is illustrated in Figure 8.3. When gelation was complete, the gel was extracted from the glass tube and immersed in deionized water.

PVA serves several roles in the preparation of this gel. In addition to enhancing the tensile strength of the formed gel, PVA helps to suspend and coat the nickel particles in solution during the transfer into the glass cylinder and while gelation occurs. Coating the nickel flakes with PVA prior to gelation also appears to enhance the efficiency of the polymerization initiator, APS.

Ferrofluid Gel

In the ferrofluid gel, the ferromagnetic material is suspended in the solvent around the gel network. The preparation protocol for the ferrofluid solvent gel is essentially the same

as that given for the nickel flake gel in the previous section. However, for the case of the ferrofluid gel, the addition of nickel flakes in the third step is eliminated[6]. The gel was formed in a cylinder 0.51 cm in diameter and 6.67 cm in length. When gelation was complete, the gel was extracted from the glass tube and immersed into the ferrofluid[7].

8.2.2 Electromechanical System

Figure 8.4 illustrates the electromechanical system used to excite and observe the dynamics of the magnetic gels. The NIPA gel was mounted vertically in a tube within the Liebig condenser (See Figure 8.5for more detail). The base of the gel was fixed with a mechanical support. The water temperature within the outer jacket of the condenser was regulated with a Lauda RC 6 water circulator, which prevented any direct heat transfer from the excitation coil to the gel. An air barrier was left as a buffer to help further stabilize the temperature of the inner glass cylinder in the experiments with the nickel seeded gel. The circulating Lauda water was set to 30 °C, which is near to, but below the gel transition temperature (34 °C). In the second series of experiments with the ferrofluid gel, the air barrier was filled with water to increase the coupling with the thermal bath. In addition, the circulating water bath was adjusted to 15 °C to increase the cooling rate of the gel temperature.

A silk fiber was attached to the top of the NIPA gel with epoxy. The fiber passed over a low-friction pulley system to an adjustable weight and light steel rod. The rod was suspended within a LVDT that sensed the relative position of the gel. The output of the LVDT circuitry was sampled and incorporated directly into the MATLAB processing environment (Mathworks,1997). MATLAB also provided the command signal to the power electronics (described in [42]) that drive the excitation coil surrounding the Liebig condenser. The use of MATLAB as a control and computation environment allowed for relatively easy synthesis of different feedback compensation schemes.

Significant electromagnetic interference between the magnetic field applied by the power electronic inverter and the LVDT prevented an accurate position measurement while the

[6]The gel used is essentially the same as the cylindrical gels described in Chapter 2, immersed in ferrofluid solvent.

[7]Lignosite FML made by Georgia Pacific

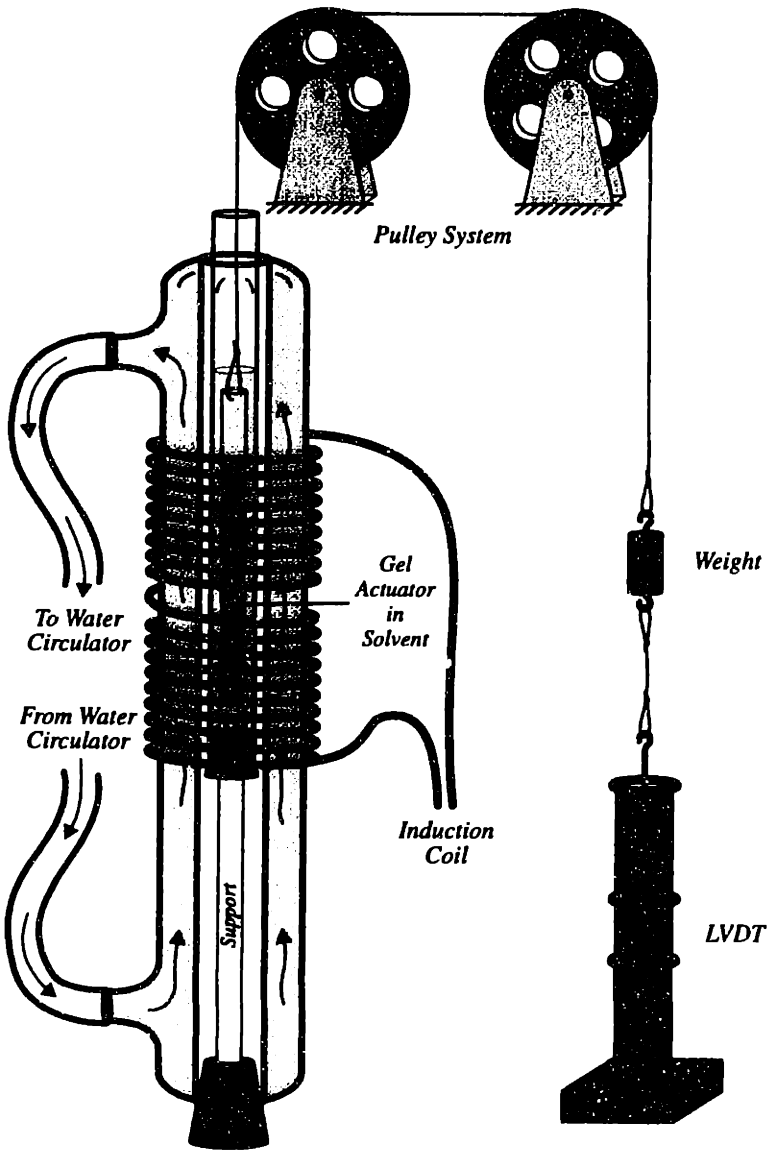
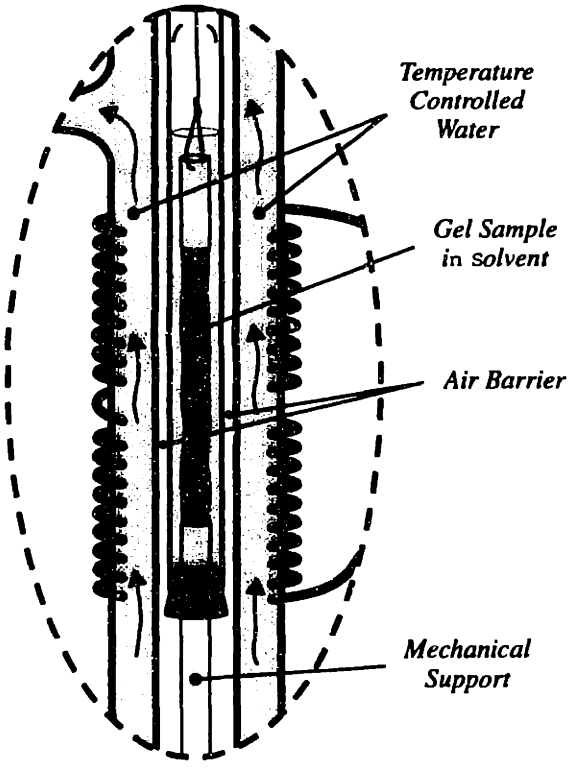


Figure 8.4: Liebig condenser and suspended gel.



(The induction coil has been removed for clarity.)

Figure 8.5: Closeup of Liebig condenser.

inverter was active. To avoid interference, the inverter was disabled during the sampling of position. This operating scheme is illustrated in Figure 8.6. A magnetic field is applied to the gel for two seconds by the inverter. Then, the inverter is disabled for two seconds to allow the synchronous detector following the LVDT to stabilize. The position voltage was then sampled, and the cycle began again. This four second switching period was substantially shorter than the dominant dynamic time constants of the gel.

8.2.3 Control Implementation

In order to close a feedback loop on the gel's position, a means of modulating the energy (induced thermal dissipation) delivered to the gel was required. The simplest control scheme given the experimental apparatus was to adjust the duty cycle of the applied field. As mentioned previously, an experimental period in the open-loop experiments consisted of a two second window of applied magnetic field, followed by a two second window to allow the synchronous detector to stabilize before sampling. The MATLAB interface facilitates adjustment of the duration of the magnetic excitation during the first two second window. Varying the duty cycle D of the applied field interval varies the average heat flow into the gel. An experimental period is always four seconds long, with at least two seconds at the end of the period when the field is off in order to facilitate accurate position measurement. This constrains the maximum duty cycle to a value of 0.5 in our experiments.

8.3 Results

Using the experimental setup and data acquisition scheme described above, both the nickel-seeded NIPA gel and the ferrofluid NIPA gel were placed in an alternating magnetic field. The field frequency in all experiments was 2.46 MHz, and the field strength was 2640 A·T/m. The dynamic response of the gel length in each experiment was recorded by the data acquisition system.

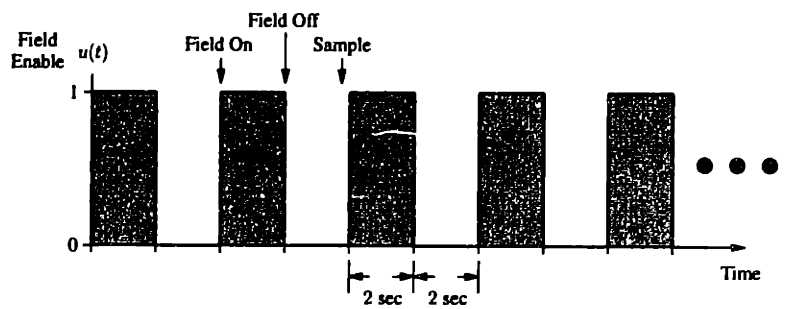


Figure 8.6: Timing sequence for heating and data acquisition.

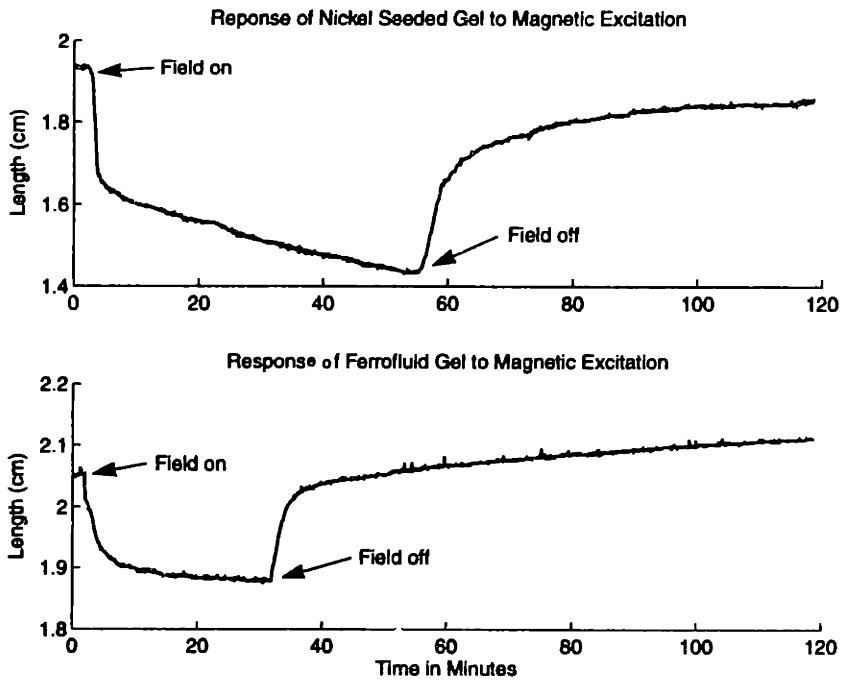


Figure 8.7: Gel open-loop responses.

8.3.1 Open-Loop Response

The duty ratio of applied magnetic field was stepped between 0 and 0.5 to produce a change in volume for the nickel-seeded and ferrofluid gels. In each case, the gels were loaded with a 5.1 g mass. Gel open-loop responses are shown in Figure 8.7. Note that the lengths indicated in the plots are the lengths of the segment of the steel rod in the bore of the LVDT, not the absolute lengths of the gels. For the time between the "Field on" and "Field off" markers in each subplot, the magnetic field was applied to the gels according to the scheme illustrated in Figure 8.6. For the time after the "Field off" marker in each plot, the magnetic field was disabled and the gels expanded. The dynamic profiles of both gels exhibit two distinct time constants, which is consistent with the observations and modeling results of the previous chapters.

8.3.2 Dependence on Load

Figure 8.8 plots the open-loop dynamic responses of the nickel-seeded gel under different loading conditions. Loads of 5.1 g and 10.19 g were used, and in each case the magnetic excitation was the same as the open-loop experiments in the previous section. The static equilibrium points of the gel volume or length will vary with load due to the gel's elasticity. The gel in Figure 8.8 is longer in both the shrunken and expanded states when it is more heavily loaded. Other properties of polymer gels are known to change under mechanical load. As mentioned earlier, mechanical tension has been shown to have an effect on the transition temperature of thermo-sensitive gels [92], [31]. The experiments in Figure8.8 demonstrate that the shape and rate of the dynamic response of the gel can also change with the variation of mechanical load. This is consistent with observations made in Chapter 7, where the LTI model description was only valid under small-signal variations of operating conditions including the mechanical tension.

8.3.3 Active Control

A closed-loop controller can be implemented by developing a compensator to modify the duty cycle as a function of the error between the measured gel length and the desired gel length. A well-designed compensator commands a corrective excitation profile for the magnetic field to minimize the position error. Figure 8.9 illustrates this scheme.

The length of the gel $y[n]$ is sampled every $T = 4$ seconds by the MATLAB-based digital data acquisition system and controller. The discrete-time index n increments once each period T. The measured error $e[n]$ is the difference between this measurement and the reference input $r[n]$:

$$e[n] = r[n] - y[n]. \tag{8.1}$$

To adjust the input to the gel system based on this error signal, a pulse-width-modulation (PWM) scheme was used [45]. This technique is illustrated in Figure 8.10. At the beginning of every n-th four second period, the magnetic field is turned on. The field will remain on for a fraction $D[n]$ of the period T, but not longer than two seconds, and then switch off.

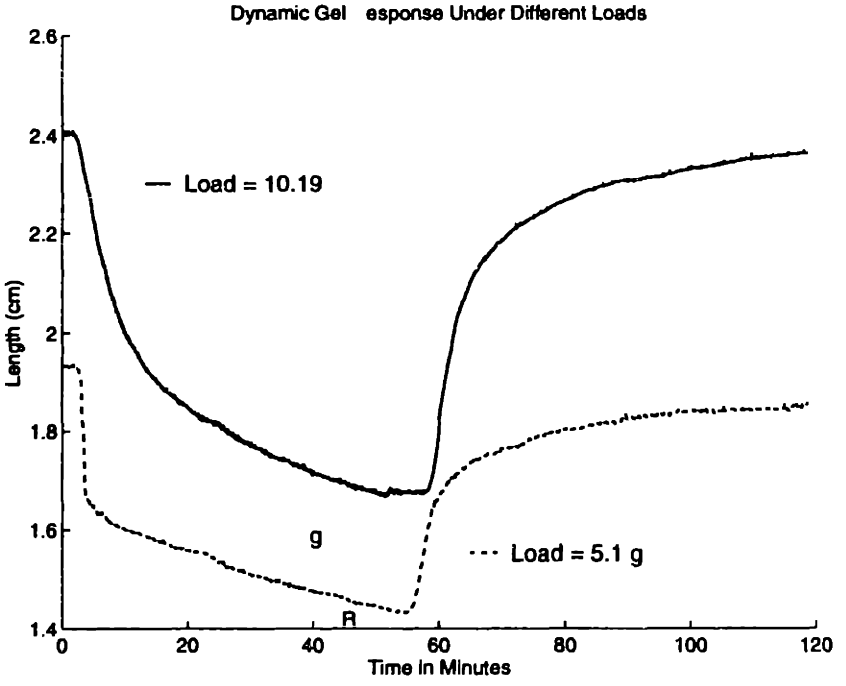


Figure 8.8: Magnetic gel responses under different loading conditions.

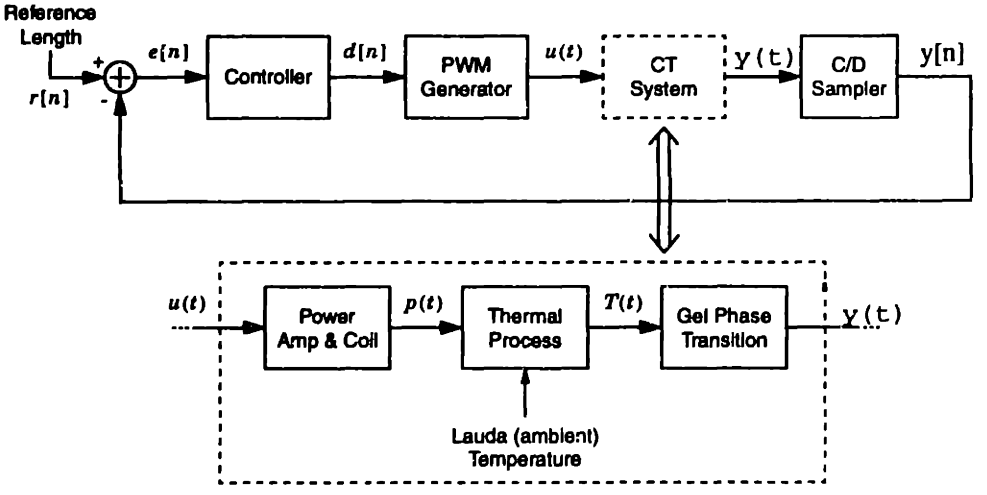


Figure 8.9: Block diagram.

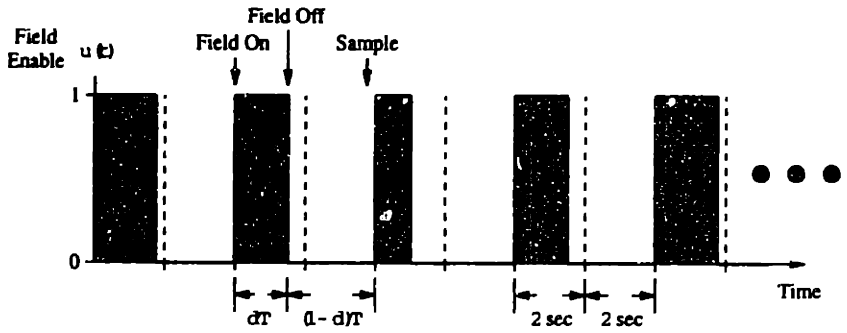


Figure 8.10: Pulse-width-modulation.

That is, the maximum duty cycle is again limited to $D[n] = 0.5$. By increasing $D[n]$, more heat is delivered to the gel over one switch period and its temperature rises. Decreasing $D[n]$, on the other hand, permits the ambient cooling from the circulating water in the outer jacket to lower the temperature of the gel. The closed-loop controller can therefore modulate the value of $D[n]$ to drive the gel position in the right direction to minimize the position error. The compensator was implemented as a digital controller in MATLAB.

To design a stable compensator with adequate dynamic performance, a model for the continuous-time system in Figure 8.9 is needed. The same form of the gel model developed earlier was used. Even though the gel is being heated by a magnetic losses in the seed material instead of through communication with a temperature-controlled circulating fluid, the dynamics of volume change and the mechanical response should be similar. The network/solvent diffusion process by which the gel changes volume is qualitatively the same. The presence of magnetic seed material in the gel network and the surfactants and magnetic material in the ferrofluid solvent will alter the specific rates of response and parameters of both the AGU and PGU models. However, the following assumptions should hold:

- The purely mechanical response, or the dynamics of the PGU, should still be much faster than the dynamics of the diffusion process. Moreover, the switching period $T = 4s$ is still much longer than the settling time of the purely mechanical modes. We

will therefore conduct the closed-loop experiments under constant load. The PGU will always be assumed in equilibrium and its dynamics can be ignored. It can therefore be replaced by a static gain that translates force to position through a spring constant

$$y(t) = \frac{F_1}{k}, \tag{8.2}$$

where k is the spring constant in the PGU model and F_1 is the output of the AGU, force applied to the load by the volume change process.

- The dynamics of both the magnetic field power amplifier and coil and also of the temperature diffusion from the seed targets to the gel were ignored in the control design, because these dynamics are both faster than the dominant gel kinetics. This approximation is valid only because of the large size and slow response of the polymer gel used in the experiments presented here.

- The open-loop profiles illustrated above have the same shape that was used to argue the *averaged*[8] model structure for the AGU earlier:

$$\frac{F_1(s)}{T(s)} = \frac{ds + a}{s^2 + bs + c}, \tag{8.3}$$

with parameters that are different from the ones identified earlier and that depend on the formula of the gel as well as the operating point in the experiment.

Using these assumptions, an overall *averaged* model for the continuous-time system in Figure 8.9 has the form:

$$\frac{Y(s)}{U(s)} = G\frac{ds + a}{s^2 + bs + c}, \tag{8.4}$$

where the gain $\frac{Ga}{c}$ captures the static relationship between the average of the square-wave input $u(t)$ and the average output position $y(t)$. Identification of the specific parameters in this model can be carried out as described in Chapter 7 and was not repeated for these

[8]Note that the temperature and force quantities in the transfer function are averaged over a switch period T. We will ignore the overbar sign and assume that all states and inputs are averaged quantities.

experiments. This model can then be used along with a zero-order hold mapping [4] to design discrete-time controllers. Even without specific values for the parameters, the general structure of the model can be used to guide control design. As shown earlier, a proportional feedback controller can be used to shorten the response time of the gel and to regulate its volume or length to a specified reference. With this particular model structure, however, a proportional controller results in a steady-state error between the reference and the output position. A proportional accumulator algorithm can be implemented in digital control to achieve perfect tracking or zero steady-state error. Details of the control design and software implementation are discussed in Appendices D and G.

The performance of the closed-loop controllers is illustrated in Figures 8.11 and 8.12. Rising and falling step references were provided as trajectories to the closed-loop gel system. Figure8.11 plots the closed-loop responses of the nickel-seeded gel. In the top portion of the figure, the higher gain proportional controller forces the gel to settle at a length that is closer to the reference signal. The proportional-accumulator controller eliminates the steady-state error as the gel length settles to the value specified by the reference trajectory.

Similar experiments were carried out on the ferrofluid gels. In Figure 8.12, the proportional controller yields a non-zero steady-state error while the proportional-accumulator controller exhibits perfect tracking in equilibrium. Note also that the closed-loop dynamics are all different from the open-loop dynamics in Figure 8.7.

8.4 Summary and Conclusions

Two types of magnetically-triggered polymer gels were investigated. The powdered-seed gel was fabricated with nickel flakes embedded in the polymer matrix. The ferromagnetic particles in the ferrofluid gel were suspended in the gel solvent. Successful remote triggering of the volume-phase transition of both gels by application of an external alternating magnetic field was demonstrated.

Magnetic activation of polymer gels has several advantages over other methods of triggering. As mentioned earlier, thermal activation was preferred to chemical methods of

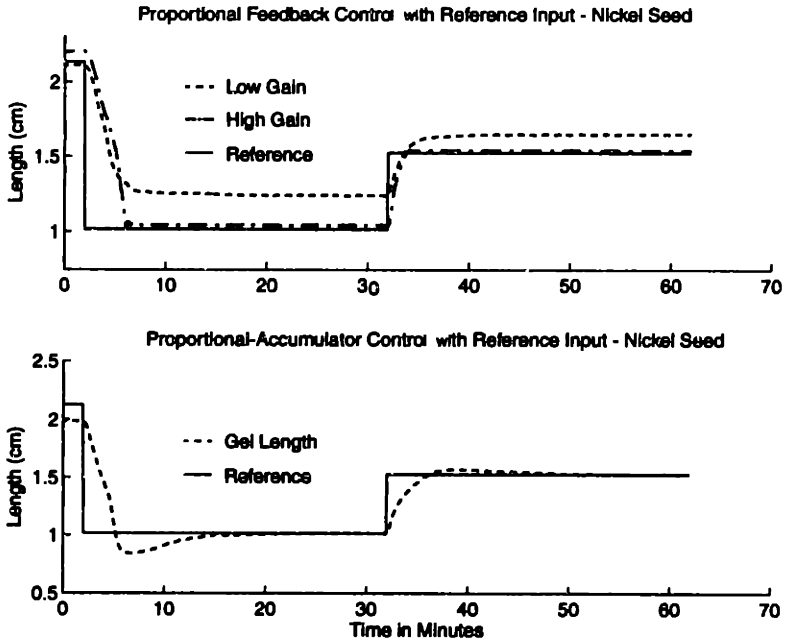


Figure 8.11: Closed-loop control – nickel-seeded gel.

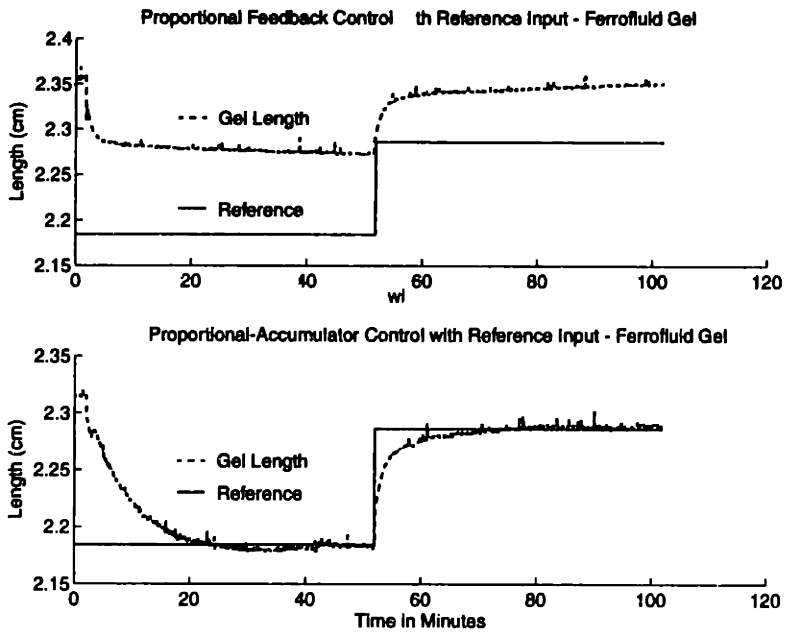


Figure 8.12: Closed-loop control – ferrofluid gel.

activation because it did not require the manipulation of large chemical reservoirs. It is also "cleaner", not resulting in chemical by-products that need to be disposed of through an elaborate irrigation mechanism [8] [52]. While the prototype system described in Chapters 5 - 7 demonstrated the potential for using polymer gels as actuators and motion regulators, it may be impractical to use fluid circulators to control the temperature of the gel in some motion applications. By using a magnetic field to induce thermal losses in the gel system, the fluid circulation system can be replaced by a power electronic system. The electromagnetic system has several potential advantages. First, faster thermal dynamics are more easily achieved since the gel network is heated directly from within. The faster dynamics as well as the higher bandwidth of electrical switches allowed us to switch at a much higher frequency. This is important both in terms of simplifying model development and allowing faster gels to be used under PWM control. An electronic system is also more easily reduced in size than a purely thermal, fluid-circulating system. In addition, the ability to trigger the gel contactlessly through a magnetic field makes it easier to power gels remotely through physical boundaries.

Several potential applications for magnetically-activated gels are currently under consideration. The ability to provide a remote trigger makes magnetic gels a good candidate for controlling *in vivo* drug delivery systems across the physical boundary of skin. We are also considering the use of magnetically-seeded gels in synthetic muscles, and multipart industrial chemical delivery systems [42]. Using ferrofluid-based magnetic gels, we are exploring the creation of fluids that exhibit a variable viscosity in response to an applied magnetic field. For example, ferrofluids can be loaded with millimeter- to micrometer-diameter gel beads. When the beads are small (having collapsed in response to an applied magnetic field heating the solvent), they occupy a small fraction of the solution volume. The viscosity of the solution is that of the sorrounding ferrofluid. If the magnetic field is then turned off, the gel beads swell and and occupy a substantial part of the solution volume. The overall solution exhibits a higher viscosity.

Still, several challenges need to be met before a practical, cost-effective, robotic or servomechanical system can be powered by polymer gel actuators. Expensive custom elec-

tronics were required to produce the high frequencies and power levels required to produce the current system performance. Improved seeding methods can help produce higher concentrations of metal seed. The overall efficiency of the system may also be improved by reducing the amount of solvent through which heat is dissipated. If a gel is fabricated to have a transition temperature well above that of the ambient temperature, the active cooling system may be unnecessary. With the help of such an integrated solution and design approach, magnetically triggered polymer gels can be incorporated successfully as active components of electromechanical systems.

Chapter 9

Polymer Gel Transition Point Sensor

While Chapter 1 highlighted the potential for using polymer gels as actuators *and* sensors, most of the discussion thereafter has focused on the cylindrical linear actuator. Often, however, a polymer gel actuator can also act as a sensor. Recall that one of the potential advantages of using polymer gels as active components or actuators is the ability to eliminate the need for a separate sensor and actuator. Because polymer gels can be fabricated to respond to a variety of stimuli, they may be constructed to respond to a variable in the environment that they will be used to control. For instance, a drug-delivery system can be implemented using a polymer gel capsule. If the gel responds to a certain level of drug p by swelling and increasing its pour size, it can release drug q to help bring down the level of p. Once the level of p drops below the "transition" level of the gel, the capsule shrinks and no more drug q is released. The gel is both the sensor *and* the actuator in this setting.

A more explicit application of a polymer gel as a sensing device is described below. A polymer gel transition point sensor was designed and tested[1]. One of the parameters of a gel system that can change due to a change in the environment is the transition point (e.g., the temperature at which a gel exhibits a large, discontinuous volume change). For instance, as mentioned earlier, the transition temperature of a NIPA gel will change in response to external load [92]. If the transition temperature can be measured, the NIPA gel

[1]This work was the result of a collaborative effort among Deron Jackson, Dhalene Fusco, Changnan Wang, and the author

can be used for tension measurements. In the previous chapters, we were able to hold the gel at a dynamic equilibrium point by switching the input at a high rate so that the volume will oscillate around the desired point. In this chapter, a similar technique is employed to develop a sensor for the presence of metal ions. A gel was fabricated whose transition temperature was a function of the presence or absence of divalent metal ions. The phase-transition was sensed by observing the optical properties of the gel (the gel turns opaque when it collapses, and clear when it swells.). Using a switching technique, the polymer gel average temperature was modulated based on its transition point temperature. Changes in the transition temperature were then used to detect the presence of polyvalent metal ions in the gel solvent. The use of the thermal, optical, and chemical properties of a polymer gel, and the switching technique to hold the gel at its transition point to implement this sensor, are described below and in more detail in [41] and [74].

9.1 Gel Design

The design of NIPA-based gels that exhibit transition temperatures that vary with the existence of certain metal ions is discussed below. First, the fabrication process is outlined. Second, the dependence of these gels on the presence of various metal ions is demonstrated by illustrating the empirical phase-transition plots for the polymer gels after being washed in solutions with different metal-chloride baths. Finally, a brief explanation of the chemical process by which the gel volume-change process recognizes the existence of these metal ions is provided.

9.1.1 Gel Fabrication

Interpenetrating polymer networks (IPN) containing *N*-isopropylacrylamide (NIPA), acrylic acid, and poly(vinyl alcohol) (PVA) were formed by free-radical copolymerization. The gels formed had the same basic components as the gel used in the linear actuator experiments. However, the inclusion of the acrylic acid added an ionic group onto the NIPA polymer chains, producing a co-polymer gel of NIPA and acrylic acid. The NIPA monomers were responsible for the volume response to temperature, while the acrylic acid was essential in

forming a complex with the metal ions to be detected. The gel was formed as follows:

1. 2 g of 115,000 molecular weight poly(vinyl alcohol) (Aldrich), 7.24 g of recrystallized NIPA (Kodak), 340 μL of acrylic acid (Aldrich), 200 mg of sodium hydroxide (Aldrich), 80 mg of *N,N'*-methylenebisacrylamide (cross-linker, Bio-Rad), and 240 μL of tetramethylethylenediamine (accelerator, Bio-Rad) were dissolved in 100mL of degassed, deionized water at 20.0 °C.

2. To initiate the polymerization, 16 mg of ammonium persulfate (Mallinckrodt) were added to the mixture in stcp 1.

3. Glass micropipettes of 1.30 mm inner diameter were inserted into the solution in step 2 prior to gelation (as soon as APS was added)[2].

4. After gelation, the cylindrical gels were removed from the micropipettes and washed for two weeks in distilled, deionized water.

5. Three gel samples were then immersed in 0.05 M baths of copper chloride ($CuCl_2$), zinc chloride ($ZnCl_2$), and manganese chloride ($MnCl_2$), respectively. These samples were washed in the metal-chloride baths for several weeks.

9.1.2 Dependence of Transition Temperature on Metal Ions

Gels consisting purely of NIPA and a non-ionic cross-linker are well known to exhibit a nearly discontinuous volume-phase transition in response to changes in temperature [31]. The gels described above in the fabrication section will be referred to as IPN gels to distinguish them from NIPA gels with no PVA polymers. IPN gels exhibit a similar, albeit more continuous, volume-phase transition. Figure 9.1 shows the equilibrium diameter of the three IPN gel samples versus temperature. Diameter measurements for each gel were collected sequentially from low to high temperature, allowing sufficient time at each temperature for the gel diameter to equilibrate. In each case, the gels shrank continuously over the entire

[2]It may be necessary to leave the micropipettes and the pregel solution under a vacuum while gelation occurs to ensure no oxygen contaminates the solution.

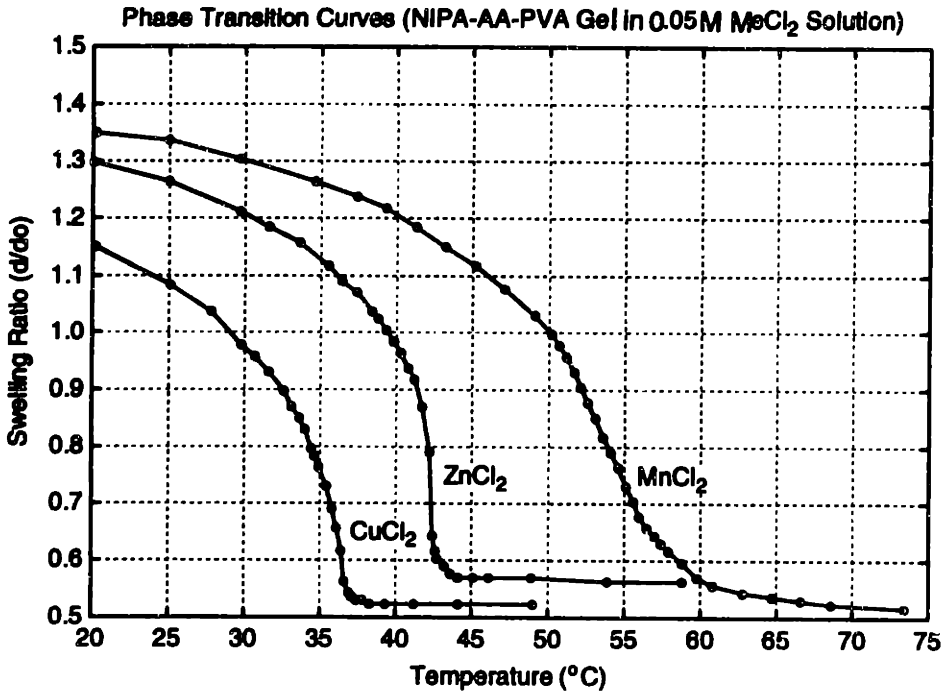


Figure 9.1: Equilibrium phase-transition curves.

temperature range with the largest effects generally occurring in a single, relatively steep transition region. This region is centered around a "transition temperature." The presence of acrylic acid in the gel causes the transition temperature to vary with the presence of metal ions as shown by Figure 9.1. Similar results were observed in the lab with metal-chloride bath concentrations as low as 10 μM. Gels made without PVA still exhibited the same metal ion dependence and had a more clearly defined tansition temperature in a discontinuous volume-phase transition curve. However, these gels were found to be too brittle for use in our preliminary experimentation. Hence, PVA was added to substantially increase the tensile strength of the gel and thereby ease sensor construction.

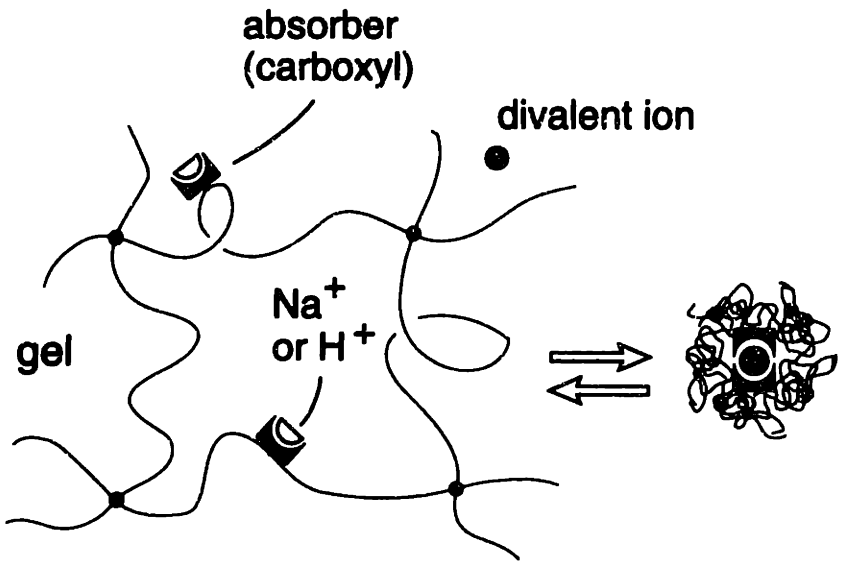


Figure 9.2: Chelation of divalent metal ions.

9.1.3 Theoretical Overview

The essence of designing a gel that responds to metal ions is the creation of a polymer consisting of two groups of monomers. One group of monomers forms a complex with the metal ions, and the other allows the gel to undergo volume transitions in response to temperature changes. The carboxyl groups of two acrylic acid monomers (CH_2=CHCOOH) in the NIPA/acrylic acid copolymer gel can form a complex with one divalent ion. In this gel system, proximity can be controlled through the reversible phase transition of the gel. At a fixed temperature, the copolymer gel collapses when the metal ion concentration exceeds a certain threshold. The absorption site can be destroyed when the gel is forced to expand through a temperature induced volume-phase transition. When the groups come closer, the absorption site is restored and can capture divalent metal ions [41]. This process is illustrated in Figure 9.2.

9.2 Experimental Setup

Figure 9.3 illustrates the experimental setup used for our prototype gel sensor. In order to detect the presence of metal ions in the gel system, it is sufficient to measure the transition temperature of the gel. This was accomplished by exploiting the optical properties of the gel. When the gel is cooled below its transition temperature, it swells and becomes optically clear. When the gel is heated above its transition temperature, it shrinks and becomes opaque. By detecting how much light passes through the gel sample, the optical sensor in the figure can determine whether the gel is above or below the transition temperature. The stacked peltier junction plates can heat or cool the gel sample based on the direction of current through them. Control electronics provide feedback from the photosensor output to the peltier junction plates. If the gel is swollen and optically clear, feedback causes the peltier plates to heat the gel. The gel subsequently shrinks and becomes opaque as it crosses the transition temperature. The peltier plates respond to the change in opacity by cooling the gel, inducing a phase change in the opposite direction. This feedback arrangement leads to a continuous oscillation of the gel sample between the swollen and shrunken states.

The gel sensor imposes and measures the non-equilibrium cyclic temperature around an average point of optical clarity. A good correspondence between the optical clarity and the physical diameter, or swelling ratio, of IPN gels throughout their phase-transition curve was observed in the lab. The oscillations are therefore occuring around the "transition" temperature of the gel, or the midpoint of the "steep" region as discussed above. By filtering the oscillating temperature profile, an average temperature that roughly corresponds to the temperature of the "steep" region can then be measured.

To estimate the temperature of the gel sample, a thermocouple was placed on the surface of the peltier junctions, close to the gel sample. A running average of the thermocouple temperature is used as the gel-sensor output. This, average gel temperature provides a good estimate of the gel transition temperature. Possible sources of error are presented in the next section. Because the sample is kept in an oscillating state, the sensor outputs a continuous, real-time estimate that allows dynamic changes in temperature as well as the

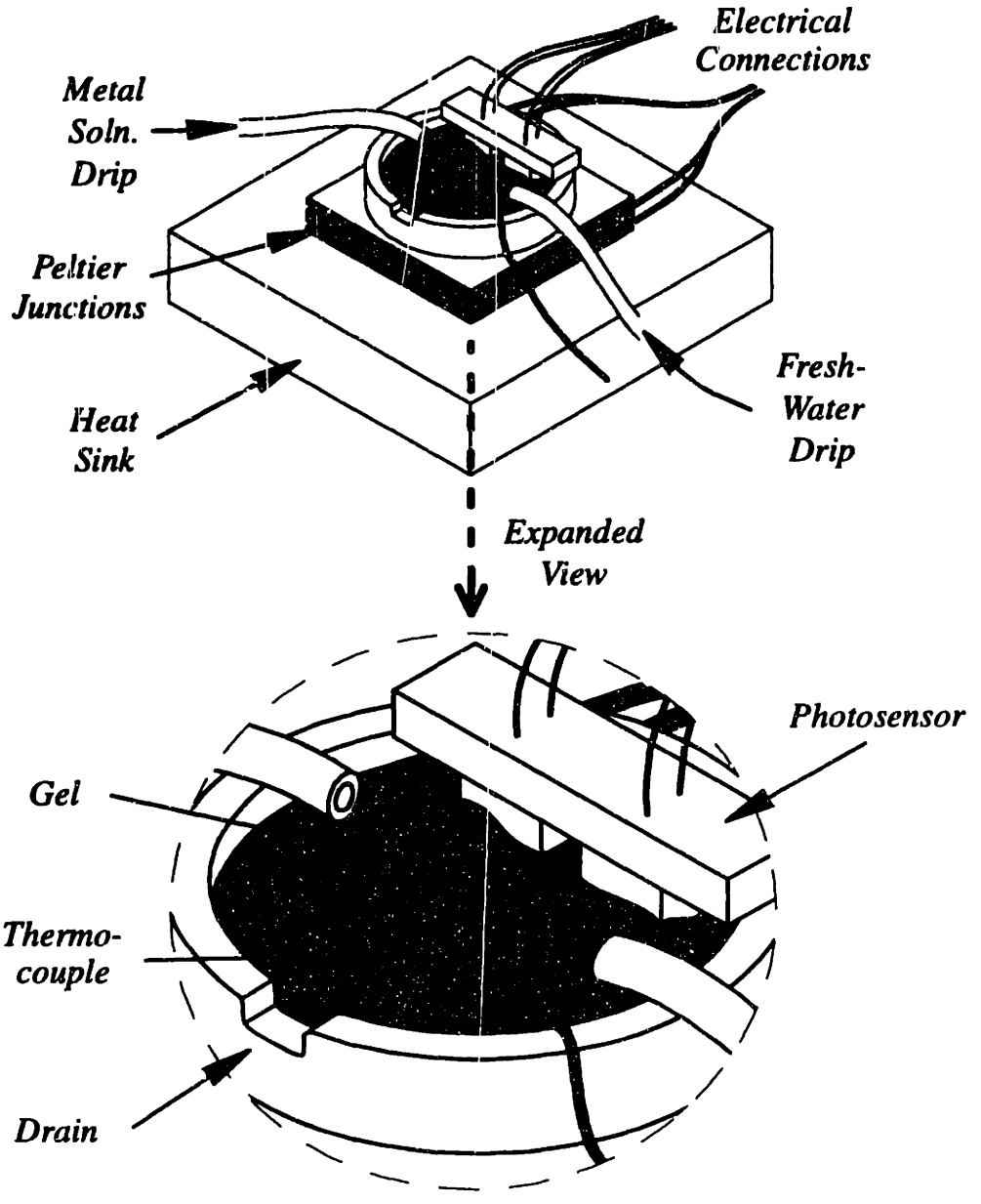


Figure 9.3: Experimental gel sensor.

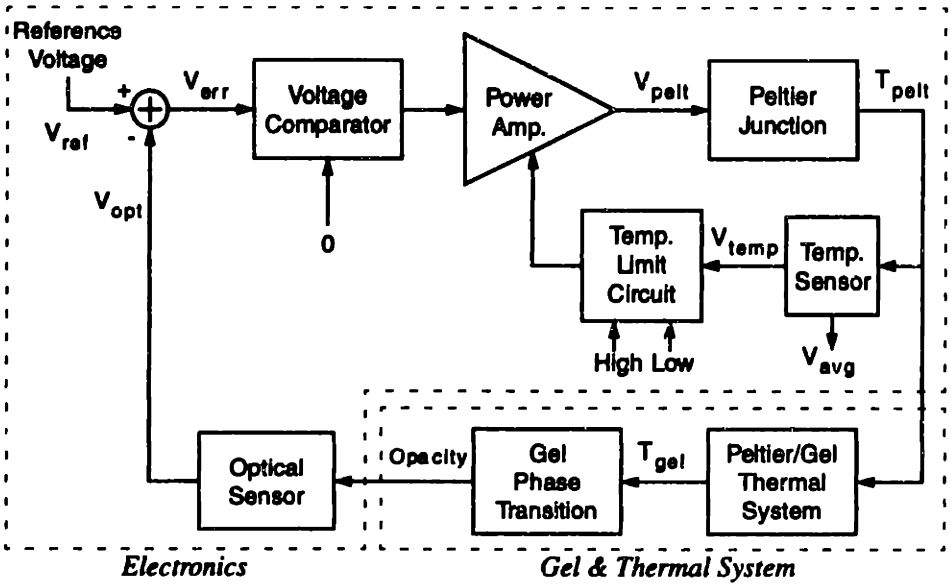


Figure 9.4: Gel-sensor block diagram.

equilibrium phase-transition temperature to be observed.

Analog control circuitry was used to implement control of the gel transition point sensor. As shown in the block diagram in Figure 9.4, the controller consisted of two loops. The inner loop was used to limit the maximum and minimum temperatures of the peltier junctions [74]. The outer loop determines whether the sampled output of the photosensor is above or below a set reference V_{ref}. Based on the sign of the output of the comparator in the loop, a power amplifier heats or cools the gel-side surface of the peltier junction plates, thereby heating or cooling the gel sample. The temperature change causes a gel transition, causing the photosensor output to change and, in time, the output of the comparator. In this manner, a cyclic steady state around the transition temperature is achieved.

Sample experimental waveforms measured from the prototype system are shown in Figure 9.5. The data was recorded while the system was in steady-state operation with a gel sample soaked in 0.05 M $CuCl_2$ solvent. Figure 9.5 (a) is the voltage applied across the peltier junctions. A positive voltage indicates heating, while a negative voltage indicates

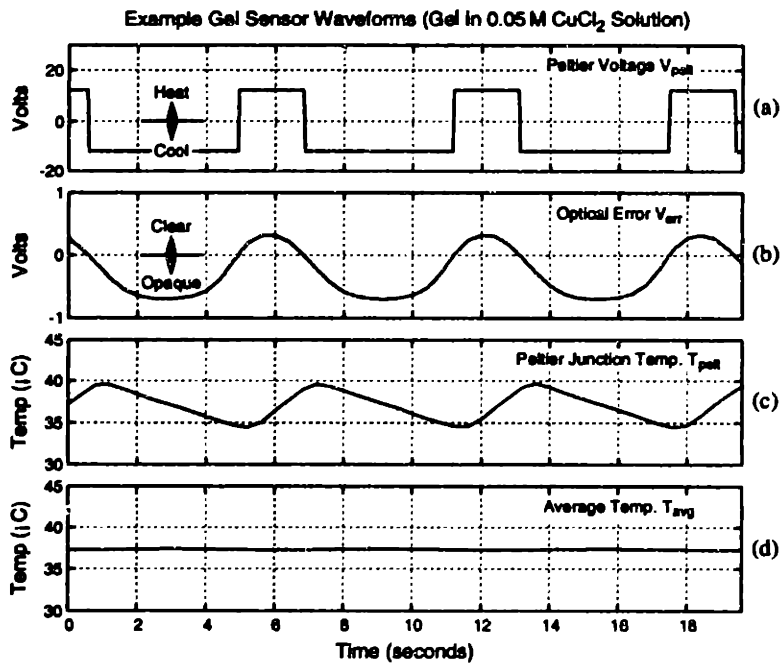


Figure 9.5: Experimental gel-sensor waveforms.

cooling. Figure 9.5 (c) is the resulting temperature as measured from the thermocouple mounted on the surface of the peltier junctions. Figure 9.5 (b) is the photosensor output subtracted from the reference voltage V_{ref}. Thus, positive and negative values of the voltage V_{err} indicate opaque and clear conditions, respectively. Figure 9.5 (d) is the averaged (low-pass filtered in the analog implementation), peltier junction temperature T_{avg}. The figure indicates a steady transition temperature of about 37.3 °C.

The continuous estimate of the gel transition temperature can be used to monitor dynamic changes in the phase-transition temperature of a gel. These changes could result, for example, from a change in the gel solvent composition. In the next section, dynamic changes in the transition temperature due to the introduction of metal ions are observed. They are used to demonstrate that it is possible to use this device to detect the sudden introduction of polyvalent metal ions into the gel-sensor solvent bath.

9.3 Experimental Results and Discussion

The sensor described in the previous section was loaded with a gel that had been washed in deionized water, but never exposed to metal ions. Two drip lines, shown in Figure 9.3, are used to introduce fluids into the solvent bath, which surrounds the gel sample. The first line was used to provide a continuous drip of deionized water into the pool at a rate of approximately 0.35 ml per minute. Contaminating 0.05 M metal-chloride solution flows into the bath through the second drip line at a rate of approximately 0.14 ml per minute. Excess bath fluid drains through a drain port as indicated in Figure 9.3. The bath fluid volume is thus maintained at a nearly constant 2.5 ml.

Using the second drip line, contaminating metal-chloride solutions were introduced into the pool of gel solvent. Figure 9.6 shows the output of the gel sensor shortly before and after commencement of the contaminant flow. The plots in Figure 9.6 show the averaged peltier junction temperature T_{avg} versus time for three different contaminants. Each experiment was started with a fresh IPN gel in a pure, deionized water bath. Since the fresh IPN gel does not exhibit a phase transition below 66 °C, the control circuitry pins the temperature of the peltier junction at a maximum allowed temperature of 66 °C prior to the introduction of the contaminant drip.

The metal-chloride drip was initiated three minutes after the start of each experiment. Almost immediately afterward, the outer surface of the IPN gel absorbed divalent metal ions. The contaminated gel was above its transition temperature (66°C), and therefore began to turn opaque. The sensor control circuitry responded by cooling the gel with the peltier junctions, thus initiating the astable temperature oscillation. The gel sensor eventually reached a steady-state temperature that corresponded well with the equilibrium phase-transition temperature. The three curves shown in Figure 9.6 correspond, from top to bottom, to experiments with manganese chloride, zinc chloride, and copper chloride. The final steady-state sensor outputs T_{avg} for each experiment are 50.6, 43.5, and 37.3 °C, respectively.

Another experiment, shown in Figure 9.7, displays the steady-state response of the

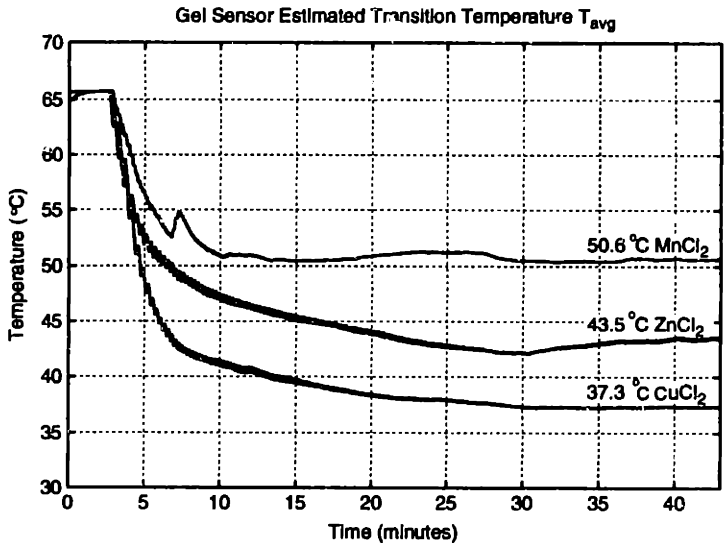


Figure 9.6: Experimental dynamic gel-sensor output.

sensor after exposure to a mixture of contaminating solutions. In this experiment the contaminating solution was a 50/50 mix of 0.05 M copper chloride and manganese chloride. The steady-state output temperature T_{avg} from the sensor was 37.5 °C. The significance of this result is discussed in the next section.

In order to reuse a gel once it has been contaminated, the IPN gel had to be cleared of chelated metal ions by being washed in the *swollen* state in a 1mM solution of sodium chloride for 60 minutes. Once the IPN gel sensor has been exposed to a metal-chloride solution and a steady-state output temperature has been reached, the sensor will continue to oscillate about that temperature indefinitely (over 6 hours in our experiments) even after the metal-chloride drip is removed. Apparently, the metal ions remain chelated while the gel is in an oscillating state.

The steady-state temperatures from Figure 9.6 are consistent with the equilibrium phase-transition curves shown in Figure 9.1. The repeatability of the sensor measurements are affected by a variety of error sources. The thermocouple temperature sensor detects the temperature of the peltier junction, which is slightly different from the temperature of the gel. This was necessary since it was difficult to place the thermocouple inside the gel

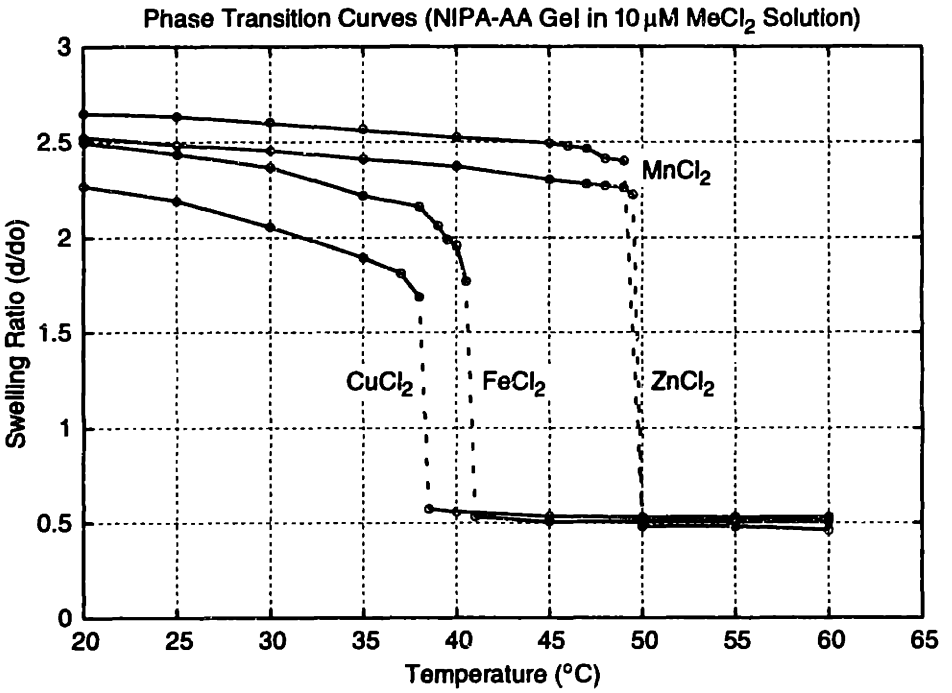


Figure 9.7: Experimental gel-sensor output when exposed to a 50/50 solution of $MnCl_2$ and $CuCl_2$.

without interfering with the optical measurements. This problem may be solved by placing the thermocouple inside an inert gel placed in close proximity to the active gel. The inert gel would not interfere with the operation of the sensor, and its temperature should mimic that within the active gel.

A second source of error stems from difficulties with the optical measurement. Bubbles are generated in the water solvent, especially at higher temperatures. As a result, the optical path may be distorted, and the gel position may be disturbed. The small hump that appears approximately 7 minutes into the $MnCl_2$ data in Figure 9.6 is the result of gross movement in the gel position. Gross motions of the gel may be eliminated in the future by using a mechanical fixture, which might incorporate gel-bond film [104] to firmly hold the gel in place.

9.4 Conclusion and Future Work

A successful prototype of a gel transition point sensor was designed and implemented. A switching control scheme was used to identify the gel's transition temperature. By sensing this temperature, the presence of specific metal ion molecules that alter the transition temperature of the gel was detected. As was the case with the linear actuator, proper gel material design as well as system design was important. Ionic groups had to be incorporated into the polymer gel network in order for the gel to have the ability to abosorb the metal ions. Moreover, by exploiting the optical properties of the gel and their fast change during a volume transition, an inexpensive, practical mechanism for detecting volume transitions was possible.

The mechanism used by this sensor can be applied more generally to sense a variety of other environmental changes. Any thermosensitive gel that exhibits an optical change during its phase transition, and whose transition characteristics are affected by an environmental parameter of interest, could be used in the sensor[3]. Even more generally, the switching scheme developed here can be used to sense the transition point of any gel. By sensing volume, opacity, pressure, or some other parameter that changes during a volume-phase transition the direction of the transition can be detected. If the gel is then made to change directions using a particular trigger (acid or base for a pH level triggered gel for instance), the same oscillatory behavior can result around the transition point. Sensors for a variety of chemicals or environmental parameters that change the transition point of a particular gel can be implemented in this manner.

The gel transition point sensor cannot always identify every individual constituent in a metal ion mixture. Figure 9.7 shows the steady-state output from the gel sensor after exposure to a contaminant drip of copper chloride and manganese chloride in equal parts. The sensor outputs a temperature of 37.5 °C, which is very near the expected 37.3 °C output for a solution containing copper chloride only. One suggestion for future work is to

[3]For instance, work is underway to develop gels whose phase-transition characteristics are strongly affected by the presence or absence of specific target molecules. Professor Toyoichi Tanaka is a Principal Investigato in one such effort.

exploit potentially different mobilities of the metal contaminants in the gel. As a mixture of contaminants separates in the gel due to this disparity in mobilities, multiple optical sensors could be placed along the length of the gel to scan the different phase-transition temperatures in different regions of the gel, thus identifying the components of the mix.

Chapter 10

Conclusions

10.1 Summary

This thesis has investigated the use of polymer gels as active components in engineering applications. First, general properties of polymer gels were introduced to motivate their potential as actuators and sensors. Specifically, the volume-phase transitions of polymer gels can be used to sense a particular stimulus and to convert some form of input energy to mechanical energy. Flexibility, conformability, scalibility, and the ability to combine the sensing and actuation functions into one device are potential advantages of polymer gels over other devices[1]. Second, as a starting point for an efficient, integrated design of a polymer gel actuator, the fabrication process and material properties of gels were discussed. By understanding the components that make up polymer gels, variations on a formula can be used to tradeoff different properties in order to optimize the system design. Techniques for making different form factors (spheres, cylinders, and fibers) in small sizes were presented. The ability to make small, quick gels is critical to achieve the high bandwidths required by many applications.

Often, a polymer gel with the desirable properties for a specific application (swelling ratio, tensile strength) will exhibit a discontinuous volume-phase transition. In order to produce a continuum of force or position profiles that may be required by a servomechanical

[1]Appendix A explores the potential of using a conformable, mass-distributable actuator in a flexible manipulator. A survey of other actuator technologies is used to highlight the strengths and weaknesses of polymer gels.

application using both discontinuous and also continuous volume transition gels, a switching drive was proposed. By switching the gel input at a high frequency[2] the volume of the gel can be controlled to a dynamic equilibrium with low amplitude ripple around an average equilibrium point where a static one was not achievable. An experimental setup was constructed to investigate a prototype linear polymer gel actuator using switching control, and the volume of a gel was successfully controlled using a switching function.

In order to effectively design controllers for a gel-based device, a good dynamic model is required. Previously existing models were either too complicated for control design or not meant to model discontinuous volume-phase transition gels. With the goal of the implemented switching drive to minimize the amount of ripple around an average state or output, an *averaged* model would be sufficient to capture the relevant behavior of the system. One area where averaged models have been frequently developed and used for switched systems is power electronics. Using *state-space averaging*, LTI models for the dynamics of the average quantities in a switched power supply, for example, can be derived, given certain assumptions about the time-scale properties in the system. Chapter 4 extended the results of state-space averaging (through additional constraining assumptions or through a nonlinear mapping of the input, for example) to a larger class of switched systems with different time-scale properties. For example, using these new techniques, an LTI averaged model for a system without significant time-scale separation between the switching frequency and input natural frequency (such as the prototype switched polymer gel actuator) may be developed.

The results in Chapter 4 motivated the existence of an LTI averaged model for the polymer gel switched actuator. A small-signal, LTI model structure was developed for the polymer gel system based on physical observations and empirical data. Using averaged, dynamic measurements from the experimental setup, the LTI model parameters were identified. The model was verified using cross-validation on different sets of data and control designs. The small-signal model describes the network/solvent diffusion process and the mechanical properties of the NIPA/PVA gel, and should be similar in structure for other kinds of gels. The work in Chapters 5 - 7 provides effective tools for regulating the volume or

[2]Relative to the dynamics of the gel.

force output of a polymer gel and identifying a model for performance analysis and control design.

The final two chapters of the thesis demonstrated the application of knowledge of the gel material and system design issues to specific problems. The slow thermal input system in the experimental setup in Chapter 5 limited the performance and speed of the overall actuator. In Chapter 8, remote triggering of a NIPA/PVA gel using a magnetic field was demonstrated. Higher switching frequencies and shorter activation delays are possible with this system. Higher closed-loop bandwidths can be achieved with this system, which may be critical in many motion applications. In addition, the electromagnetic power delivery system is potentially more easily miniaturized.

A gel sensing application was investigated in Chapter 9. A device was implemented that can detect changes in the transition temperature of a gel. By designing the gel to change its transition temperature in response to certain metal ion contaminants, the presence of these metals could be detected. Both the magnetic gel and the transition point sensor are good examples of combining system level design with the chemical or fabrication details of the material in order to achieve the desired performance in an application. Because of the many choices of shapes, sizes, and trigger mechanisms, an integrated design approach at the material, actuator, and system level can improve the final product. The PWM technique and the dynamic model developed above provide the basis for designing controllers for many polymer gel-powered systems.

10.2 Future Work

The design of systems with polymer gel actuators and sensors will involve at least three strongly coupled steps : *gel design*, *actuator design*, and *system design*. This thesis has investigated the gel design component and presented a behavioral dynamic model for interfacing the material with the rest of the system. Development of the PWM control technique, the magnetically activated gels, and the gel transition point sensor exposed some higher level design issues and potential solutions. However, a wide range of issues should be addressed

before a polymer gel can be used in a practical application. Often, these issues will be application specific. Some of the challenges of future work include:

- Systems powered by antagonistic pairs of gels should be explored. Very often, the shrinking transition of the gel appears to be faster and more powerful than the expanding one. Moreover, it is frequently easier to transfer power in one direction (heating a magnetic gel for instance). The human muscular system operates in this manner. In general, because of their soft texture and flexible structure, gels seem more suited to "pulling" than "pushing". Using an antagonistic pair as opposed to a gel and a passive spring, for instance, also provides the opportunity to control the compliance of the system at a given position by applying a "common mode" input to the two gels.

- The use of polymer gels in micro-actuation should be investigated. Because of their scalability, gels may be ideal for such applications. In addition, a pregel solution can be injected into a small mold of any shape. In this manner, micro-sized actuators of any shape and size can be fabricated. Combined with modern micro-machining techniques, very intriguing systems and applications may be developed.

- Polymer gels operate in a wet environment. They change volume by exchanging solvent with the surrounding space. A flexible "skin" should be developed that will contain the solvent and survive a rugged environment (depending on the application) without constraining the motion or flexibility of the polymer gel. This skin will also have to allow power transfer to the gel actuator. Magnetic activation is nonintrusive and may make it easier to solve this packaging problem.

- Effective bundling of small units of gels should be achieved in order to create fast, strong gels. Accurate, repeatable, and well-controlled techniques for making the units as well as grouping them will be essential in order for the load to be well distributed across a bundle. In addition, models for gel bundles will be more challenging to derive. Theoretically, they can be derived from the single gel model developed in this thesis.

For instance, multiple gel cylinder models can be connected in parallel to represent the mechanical model for a cylinder or fiber bundle. However, effects of the neighboring gels on the diffusion process of one gel strand will add unmodeled effects that may significantly influence the aggregate behavior.

- The transition-point sensor in Chapter 9 was not capable of identifying mopre than one divalent metal ion in solution. One suggestion for future work is to exploit potentially different mobilities of the metal contaminants in the gel. As a mixture of contaminants separates in the gel due to this disparity in mobilities, multiple optical sensors could be placed along the length of the gel to scan the different phase-transition temperatures in different regions of the gel, thus identifying the components of the mix.

Additional work on enhancing the gel material itself should also be addressed. The gels used here exhibit relatively low tensile strength. PAN fibers used in [51], for example, can be much stronger. However, they have the undesirable property of using pH level as a trigger and requiring an irrigation system to dispose of the salt resulting from acid/base reactions. In addition, the development of homogeneously crosslinked, uniform gel fibers that are a few microns in diameter is currently under research. The speed of polymer gels that can be fabricated using other methods may not be satisfactory in all cases.

The small-signal model should be verified for other types of gels and at several operating points. We believe the structure of the model should be the same given our observation of dynamic responses of a variety of gels in the lab and in the literature. Similar experiments on other kinds of gels should help solidify this claim. In addition, identification of model parameters at different operating points in the gel is essential if the gel is to be controlled over a large range of motion. A table of parameters for different operating conditions or an analytical description for the parameters in terms of the gel states can be used to implement adaptive control algorithms that will make the small-signal model useful over a large-signal range.

Appendix A

Case Study: Actuator Mass Distribution in a Flexible Manipulator

Gel actuators provide flexibility and the ability to distribute actuator mass on a manipulator. This may potentially be useful in improving the dynamic response of the mechanical system or providing increased maneuverability. This appendix explores the potential value of polymer gel actuators in this regard. First, other approaches to enhancing the performance of flexible, multi-link manipulators will be presented. For example, several efforts to vary the mechanical design of the manipulator structure and place the actuators at more favorable positions are discussed. The goal of this work was to alleviate some the force requirements on the drive system and therefore reduce its mass. Second, the challenges encountered in designing flexible manipulators illustrated by these designs are used to motivate the need for a flexible, conformable actuator that can be layered on a given skeletal structure. A survey of linear actuators than can potentially be used to distribute the actuator mass in a manipulator is used to compare gel actuators to the available technology. Third, mass-distribution of a linear actuator is used to analytically motivate the potential advantages of using a mass-distributed polymer gel actuator.

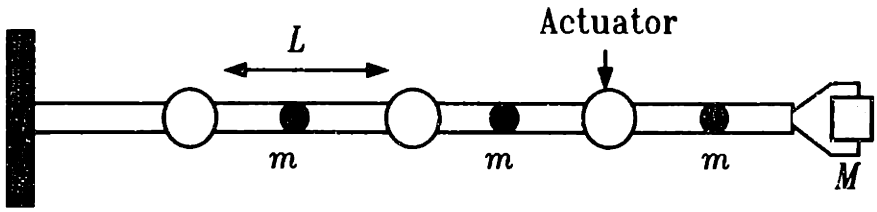


Figure A.1: Conventional multijoint manipulator with actuators at the joints.

A.1 Traditional Multijoint Manipulator Designs

Several attempts have been proposed to alleviate the weight restriction of actuator mass on multijoint manipulators. Some designs employ a powerful slider at the base of the arm that reduces the load off the other actuators by bearing much of the driving force [30]. Others use light weight designs for joints. Special light weight rotary joints are discussed in [3] and [77]. A more common mechanism of concentrating the actuators at the base of the robot and transmitting power to each joint through tendons or transmission cables is explored in [29], [27], and [39]. Reference [63] proposes another design which increases the payload capability of a multijoint manipulator by using the traction forces of the power-transmitting tendons positively, as will be shown below.

Consider the multijoint manipulator shown in Figure A.1. In this configuration, there is an actuator at every joint. It is assumed that this arm is in the horizontal configuration shown as this is the worst case in terms of gravitational support. For simplicity, consider the case with uniform link lengths L and uniform link mass m. (Note that we can consider the actuator mass lumped into the link mass.) Also denote the load mass M, the mass to be supported by the end-effector. For a structure with n actuators indexed with i from the base to the tip, it can be shown that torque τ_i to be generated by actuator i is [63]:

$$\tau_i = (n+1-i)MgL + \frac{1}{2}(n+1-i)^2 mgL \tag{A.1}$$

This equation shows that very powerful actuators are required close to the base. The problem is compounded as more flexibility and more degrees of freedom are needed.

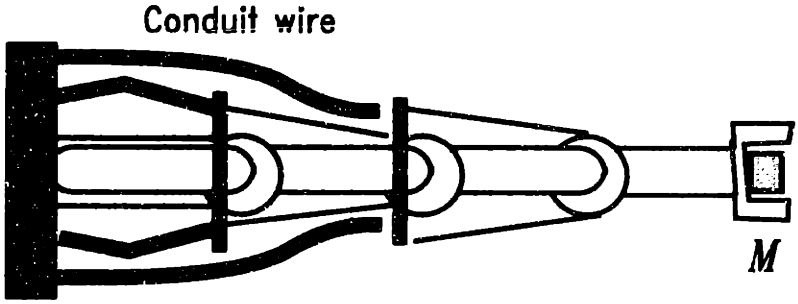


Figure A.2: Design of multijoint manipulator with joints driven remotely by base actuators and conduit wires.

Figure A.2 shows one way to reduce the weight of the structure and thus the torque required by the actuators near the base. In this example, conduits are used to carry wires on cables that transmit mechanical power. Power is generated by actuators at the base, and the joints are remotely driven. By removing the actuators from the links and mounting them at the base, much of the weight of actuators that the base actuator has to carry is eliminated, and the base actuator torque requirement is reduced. In this design (with all the motors at the base), however, large tendon traction forces are still required from the base side actuators. Moreover, when large forces are exerted, friction in the conduit wires becomes significant.

A different design reduces friction by getting rid of the conduit wires. Instead, tendons of former joints are passed through the center portion of the rear joints. Tendon traction forces for driving the base side actuators are still large, however. As mentioned earlier, [63] proposes a design to improve on these cable driven designs. Figure A.3 is borrowed from the reference and shows the proposed manipulator. The cables are wound in such a way that when torque T_i is generated at joint i, the same torque is generated at joints 1 to $i-1$. In this manner, the effect of the coupled drive is positively used, and the result is a reduction in tendon traction forces. The mechanism is described in detail in [63].

It is also noteworthy that flexibility and maneuverability could potentially be achieved

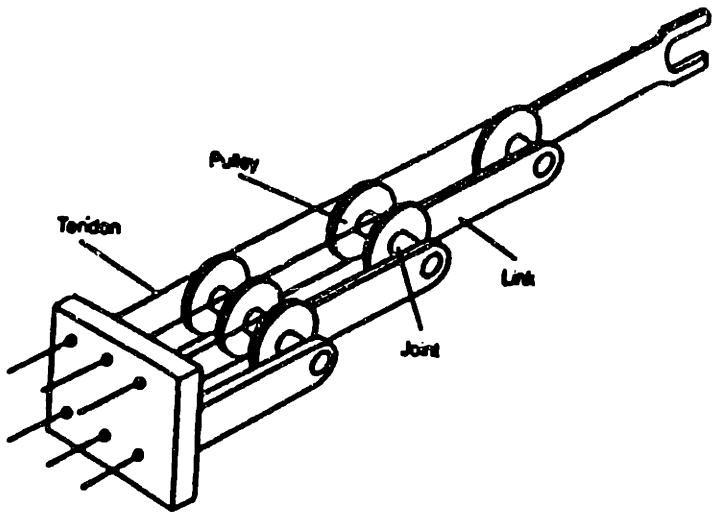


Figure A.3: Coupled-tendon driven manipulator from [63].

by using flexible components as well as kinematic redundancy. An example of this approach is given in [5]. Elastic components usually excite undesirable harmonics that degrade performance and pose serious challenges to the control system. The next section argues that perhaps a clever choice of joint and actuator types, coupled with a rigid "skeleton" could be the answer to this need for flexible, kinematically redundant robots.

A.2 Mass-Distributed Linear Actuators

The designs mentioned above either moved the actuators to the base joint or sought lighter material for building them. A re-examination of the goals of these efforts leads to another approach. The primary goals are flexibility and maneuverability. Adding flexible or elastic parts introduces undesirable dynamic performance due to excess harmonics. Increasing the number of actuators increases the force requirements on the actuators, especially those that have to bear much of the structures weight. It is not the weight of the extra parts that matters, however, but the manner in which they couple with the earlier actuators in the kinematic chain. A rigid series of links with the appropriate joints and a well-chosen

distribution of the mass of the actuators can lead to an improvement in performance.

A good example is the structure of the human body, or that of many other animals. The human arm combines rigid skeletal support and a layered muscular structure to produce a wide range of motion with great maneuverability. The rigid bone structure gives it strength and support. Actuation of the structure is accomplished through a system of distributed mass linear actuators, the muscular system. Ball-and-socket joints with properly placed actuators allow for multi degree-of-freedom movement. Obviously, the human system controller is very advanced. However, the problem of controlling this structure if we are able to build it should not be too difficult to solve. Research on the analysis and modeling of the human motion control system exists and should provide the know-how for controlling such structures soon [57]. To be able to mimic this design, linear flexible actuators are necessary. These actuators need to have the necessary force densities, use an easily controlled form of energy for actuation, and meet other feasibily requirements. A brief description of actuators that are currently available or are being researched and may hold promise for use in this application is given below.

A.2.1 Lead Screw

The simplest "linear" actuator to be used in a manipulator system would be a system that employs a rotary motion motor that uses a lead screw to convert the rotational motion to linear motion. A popular approach is to use stepping motors to drive the screw. High accuracy and high levels of force can be achieved. A lead screw can also exert force in both directions, eliminating the need for using a pair of actuators to control one joint (as in the human body). An example given in [7] can provide 40 lbs of linear force and increments of on mil. (You can increase the size of the driving motor for more force, but if the actuator is further along in the kinematic chain than the second link, then the large driving motor has to be carried by the structure.) Speed and weight could be of concern in lead screw systems. Moreover, the mechanical drive components result in backlash as well as wear and tear of the components. A more subtle drawback is that the screw has to be as long as the maximum length required by the range of motion, and has to stay that way. It does not

shrink, but protrudes out of the structure, and that may have undesirable effects on the dynamic characteristics of the system under certain configurations.

A.2.2 Linear Motors

Linear motors, unlike systems that use lead screws, eliminate the need for mechanical conversion by directly converting electrical current into linear force. This simplifies the motion system's mechanical design and reduce the wear with time mentioned in the previous section. Position can be very high, depending on the sensing mechanism. A linear motor generally consists of a wound coil and a permanent magnet. Force is produced by a magnetic field acting on current in the coil. Controlling the current flow controls the output force. A linear motor example in [7] is 58 in. long has a linear stoke up to 18 in. and up to 500 lbs of force. A primary concern with linear motors in multi-joint manipulators would be weight. Flexibility and maleability, necessary for flexible multi-degree of freedom motion, are lacking. The lead screw problem of a fixed length equal to the maximum required is also a difficulty here. Like the lead screw, only one actuator per joint is required.

A.2.3 Voice Coil Actuators

Voice coil actuators are direct drive, limited motion devices that use a permanent magnet field and a conductive coil winding to produce a force proportional to the current applied to the coil. The electromechanical conversion mechanism of a voice coil uses the Lorentz Force Priniciple, which states that a force acts on a conductor carrying current and placed in a magnetic field. They are used in applications requiring high acceleration or high frequency actuation [6].

As an actuator for multijoint manipulator, voice coil actuators would be limited to use in joints where high force but limited angular or linear displacement is required. In a multijoint manipulator where flexibility and maneuverability are essential, most joints are required to effect large displacements (angular or linear) to achieve redundancy and thus be able to reach places that are not easily accessible by attaining a certain configuration.

A.2.4 Hydraulics and Neumatics

Hydraulic and neumatic actuators are also used as linear actuators. A hydraulic fluid (oil) is pressurized using a pump that is driven by an AC motor. A hydraulic valve regulates the fluid into the actuator (both flow rate and pressure). For a linear actuator, a piston-cylinder is used [15].

A flexible actuator has been built using hydraulic actuators [96]. The actuator is made of fiber-reinforced rubber. Its deformation is controlled by an electro-hydraulic pressure control system. The actuator is made of three sections that are reinforced with fibers in the radial direction, so they deform in the axial direction. When all three sections are pressurized equally, the actuator stretches axially. Otherwise, it bends away from the direction of the section with the most pressure. By controlling the pressure in the three chambers, the actuator can be bent in any direction.

The hydraulic flexible actuator exhibits flexibility and a range of motion that are both superior to other linear actuators. However, as presented in [96], it comes in the shape of a cylinder and does not seem maleable enough to shape into a variety of configurations. Hydraylic actuators are superior to electric motors in terms of the torque/mass ratio and they are not limited by the saturation of electric motors at some level of magnetic flux density. Large amounts of force are not necessarily an important goal of flexible, multi-jointed actuators, however. Moreover, hydraulic pumps are usually inefficient due to dissipation, leakage, and compressibility effects. The compressors which power the hydraulic actuators are themselves driven by heavy electric motors.

A.2.5 Other Actuators

Other actuators are also used in applications that require linear force or displacement. **Piezoelectric** actuators use reverse piezoelectric action to generate very precise motion (Piezoelectric action is the generation of electrical charge and associated potential differences when subjected to mechanical stress or strain.) [15]. They are used in miniature stepper motors, accurate positioning stages, and delicate test equipment. They are, however, too rigid and are too limited in their range of motion for the flexible manipulator

application. **Shape memory alloys**, such as Nitinol, have been used in robotics, but the thermal to mechanical energy conversion has proven difficult in practical applications. **Magnetostrictive** material, such as Terfenol, has been used to build the "Elastic Wave Motor" [87]. A Terfenol rod alternately contracts, elongates, and reclamps itself when subjected to a magnetic field induced by current running through a wire. This results in worm-like motion than can be controlled very precisely. The main application has been in the paper manufacturing industry, and more applications are being explored. However, the actuator as constructed is encased in a rigid structure and would not be suitable for flexible manipulators. Other novel actuators convert rotary motion to linear motion [28] for high resolution for instance.

Despite the abundance of research in linear actuator technology, none of the available actuators seem ready to be used as artificial muscles in multijoint manipulators. Polymer gel actuators are closer than the above technologies in texture, flexibility, maleability, and range of motion to the human muscle. Actuators of that kind are still in the development stages. More research on using this and other materials for flexible manipulator designs is needed to develop the actuators to where they can be used as "artificial muscles" .

A.3 A Design Example

This section analyzes a simple, planar, two-link manipulator with a lumped actuator and a distributed-mass actuator. In the example used here, the mass has been very simply stretched out into a line that connects the two links. The issue of complimentary "muscle" pairs is not addressed again for simplicity. Some of the actuators mentioned above can exert force in both directions and can therefore be used in the simplified model presented here. Others would need to work in antagonistic pairs, with each of two actuators causing a rotation in one direction. It is also important to note that in a three dimensional setup with more links, the improvements can be significantly increased. Moreover, the mass can be distributed much more elaborately and computer analysis tools can be used to determine the ideal shape of the actuators for a given system requirements. However, the simple shape

used here is one that can be implemented with all the actuators surveyed above (More elaborate shapes may not be realizabe with some.) and makes the analysis straight forward yet informative and illustrative of the potential improvements. The following analysis will focus on the forces exerted on the base actuator due to the mass of the actuator controlling the second link, as well as the force requirements on the actuators to generate a desirable endpoint force.

Figures A.4 and A.5 illustrate the two systems to be compared in most of this analysis, with the first being the traditional direct drive system and the second being the proposed distributed linear actuator system. An understanding of the force mapping from the actuators to the endpoint, as well as static forces needed to support the structures, is essential for building controllers for the system. It also serves as a guide to the force requirements on the actuator used. It is important to remeber that polymer gel actuators are not being explored for use as mechanical actuators for their high force densities but for their flexibility and conformability. Moreover, any deficiencies in power or efficiency that may be concluded are likely to improve as the material side of the research progresses. While it will turn out that the force requirements on the linear actuator are larger than the torque requirements on the direct-drive rotary one (an obvious result since the joint is revolute), there are several reasons to explore this system. The weight distribution of the linear actuator will result in reduced power requirement on the base joint actuator. As the gel and fiber spinning technologies advance, more forces will be available from the actuators. Even without large forces, a light-weight, small, easily maneuverable system has many applications. Although this system may not be complex enough to demonstrate all the potential of polymer gel actuated systems, it serves as a good starting point. It is also the same technique used in the human arm, which is a superior multi-degree-of-freedom manupilator.

A.4 Base Actuator Torque

A.4.1 Static Analysis

The relationship between an endpoint force $\boldsymbol{F}$ and joint force vector $\boldsymbol{\tau}$ is given in [2] as:

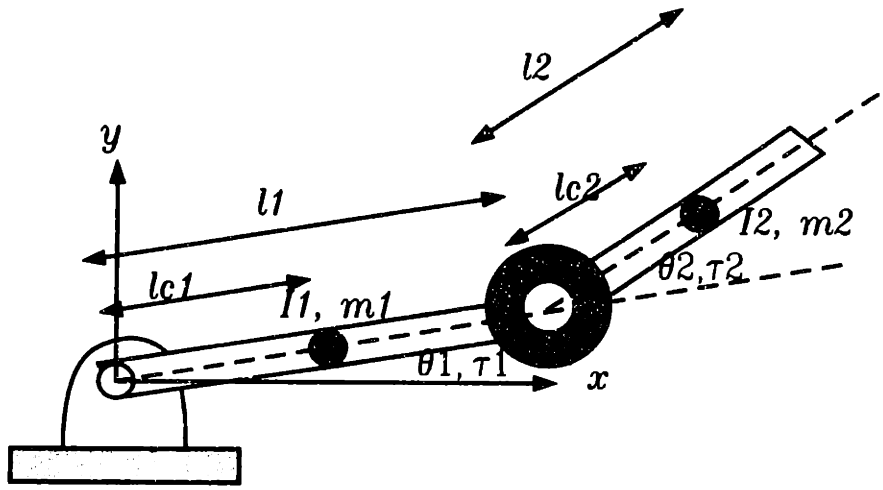


Figure A.4: Two-link planar manipulator with a lumped-mass actuator.

$$\tau = \boldsymbol{J}^T \boldsymbol{F}, \tag{A.2}$$

where $\boldsymbol{J}$ is the Jacobian relating endpoint velocities to joint velocities. This expression for τ ignores the link weights and actuator weights. To analyze the effect of a direct drive motor at the base of the second link on the torque exerted by the first actuator, the Jacobian relating the actuator velocity to the base joint angular velocity should be computed. The position of the actuator, assuming a lumped mass at one point and using Figure A.4, is:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} l_1 \cos(\theta_1) \\ l_1 \sin(\theta_1) \end{bmatrix} \tag{A.3}$$

Taking partial derivatives with respect to θ_1 and θ_2 :

$$\boldsymbol{J_A} = \begin{bmatrix} -l_1 \sin(\theta_1) & 0 \\ l_1 \cos(\theta_1) & 0 \end{bmatrix} \tag{A.4}$$

The force vector due to a point mass actuator with mass M is then used to compute the excess torque required from the actuators:

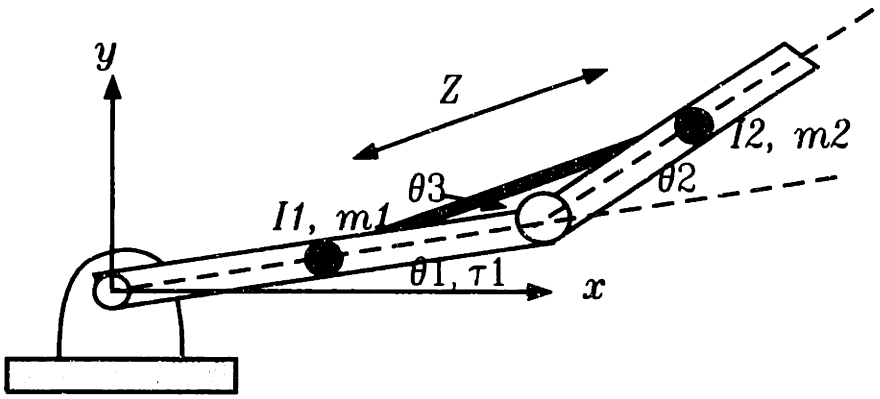


Figure A.5: Two-link planar manipulator with distributed-mass actuator.

$$\begin{bmatrix} \tau_{A1} \\ \tau_{A2} \end{bmatrix} = \begin{bmatrix} -l_1 \sin(\theta_1) & l_1 \cos(\theta_1) \\ 0 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ Mg \end{bmatrix} = \begin{bmatrix} Mgl_1 \cos(\theta_1) \\ 0 \end{bmatrix} \tag{A.5}$$

As expected from the model of a point mass at the connection between the two links, the extra torque requirement affects only the base actuator. This is the relevant actuator for the analysis here.

Figure A.5 illustrates a different manner of actuating the second link. The mass of the actuator is distributed over a line and the actuator acts linearly. The joint converts this linear motion to a rotary motion. Assuming the mass of this actuator is equal to M, the mass of the point actuator in the previous example, define the mass density ρ such that $\rho z = M$, where z is the length of the actuator. (There are other considerations, such as that it may not be feasible to maintain the same mass yet provide the same amount of force for the second link, but the concern here is the first actuator force requirements and this simple model serves as a good starting point.) Other parameters of interest here are the angle θ_3 and the points of contact of the actuator and the links. Once the points of contact are known, the law of sines or cosines can be used, with the angles $\theta_{1,2}$ to determine θ_3 and z. Given these parameters, the position of a point on the actuator is:

$$\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} \alpha l_1 \cos(\theta_1) + \zeta \cos(\theta_1 + \theta_3) \\ \alpha l_1 \sin(\theta_1) + \zeta \sin(\theta_1 + \theta_3) \end{bmatrix} \tag{A.6}$$

where α is the fraction of l_1 before the contact with the actuator and ζ is a point on the actuator and ranges from 0 to z. The "incremental" force exerted on the base actuator by a "point" on the linear actuator is ρg in the vertical direction. Differentiating with respect to θ_1 gives the first row of ${J_A}^T$ and differentiating with respect to ζ gives the second row. Only the first row is relevant in this analysis as mentioned earlier (The second actuator also needs force to carry itself in this setup, but that is ignored in this analysis.), and so the expression for the extra required torque from the base actuator due to one point on the linear actuator is:

$$\tau_{a1} = (\alpha l_1 \cos(\theta_1) + \zeta \cos(\theta_1 + \theta_3))\rho g. \tag{A.7}$$

Integrating this expression as ζ goes from 0 to z and substituting $M = \rho z$, the total extra torque required from the base actuator is:

$$\tau_{A1} = Mg(\alpha l_1 \cos(\theta_1) + \frac{z}{2} \cos(\theta_1 + \theta_3)). \tag{A.8}$$

If the $\theta_3 = 0$, $\alpha = 1/2$, and $z = l_1$, the solution simplifies to that of the point mass seen earlier, which is to be expected. Note that this expression can be reduced by decreasing z or α or increasing θ_3 when compared with the point mass expression. In essence, moving the actuator closer to the base reduces the extra torque. In a three dimensional example, these effects should be more noticeable, as the mass distribution of the actuator is allowed to be more elaborate and different actuator shapes are permitted.

It should be noted that for a given workspace and attach points of the actuator, a certain stroke is required. With some actuators, this may allow us to use an actuator for only part of the length z and complement it with a taught string, or tendon. If the tendon is assumed massless, then the result is lower torque requirements on the base actuator, even if the total mass of the actuator is the same in both cases. If physically, this change is thought of as using only part of the actuator, then the mass of the actuator is reduced as well. However,

it may be that the same energy and force requirements are needed from the linear actuator, and that may roughly translate to the mass remaining constant. It is this second case that will be explored.

In order to compute the required length of actuator in a certain configuration for a certain workspace (The workspace here is defined mainly by θ_2, as that is the only joint coordinate that varies with the length of the linear actuator.), two ends of the workspace need to be defined, θ_{2I} (initial) and θ_{2F} (final). The total length of the string (tendon) and the actuator at both angles and the contraction factor k_{cf} (a number between zero and 1) can be used to compute the length of actuator required. If we restrict both θ_{2I} and θ_{2F} to lie between 0 and 2π, the actuator length is:

$$z_{part} = \|z_{\theta_{2I}} - z_{\theta_{2F}}\|/k_{cf}. \tag{A.9}$$

Of course, for certain parameters and contraction factors, some workspaces are not achievable.

The excess actuator torque due to the same mass M but a different length z_{part} is simply given by the same expression for the mass-distributed actuator but with z_{part} replacing z :

$$\tau_{A1} = Mg(\alpha l_1 \cos(\theta_1) + \frac{z_{part}}{2}\cos(\theta_1 + \theta_3)). \tag{A.10}$$

A.5 Dynamic Analysis

To describe the forces required by the joints to move the manipulator, [48] uses the Lagrangian L:

$$L(q_i, \dot{q}_i) = T - U \tag{A.11}$$

Where q_i is a joint coordinate (i.e. θ_1), T is the kinetic energy, and U is the potential energy. This expression is differentiated in the following manner to compute the required torques for achieving certain generalized coordinates and velocities:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = \tau_i \tag{A.12}$$

The right-hand-side torques are the generalized forces corresponding to the generalized coordinates q_i. Since differentiation is a linear operation, the contribution of the kinetic and potential energies can be computed independently using A.12 (by substituting T or $-U$ for L). As in the static analysis, only the extra terms due to the actuator mass need to be computed.

A.5.1 Point Mass

The extra potential energy due to the lumped actuator in Figure A.4 is:

$$U_A = Mgl_1 \sin(\theta_1) \tag{A.13}$$

The additional torque to be generated by the base actuator due to this increase in potential energy is:

$$\tau_{A1U} = \frac{d}{d\theta_1} U_A = Mgl_1 \cos(\theta_1) \tag{A.14}$$

The excess kinetic energy is:

$$T = \frac{1}{2} M \|v\|^2 = \frac{1}{2} I \omega^2 \tag{A.15}$$

where v is the linear velocity, I is the inertia of the point mass about the base joint axis, and ω is the angular velocity of the actuator.

$$\|v\|^2 = l_1^2 \dot{\theta}_1^2 \tag{A.16}$$

$$I = Ml_1^2 \tag{A.17}$$

$$\omega = \dot{\theta}_1 \tag{A.18}$$

Using these substitutions, the kinetic energy of the actuator is:

$$T_A = \frac{1}{2} M l_1^2 \dot{\theta}_1^2 \tag{A.19}$$

and the extra torque required from the base actuator is:

$$\tau_{A1T} = \frac{d}{dt}\left(\frac{d}{d\dot{\theta}_1} T_A\right) = M l_1^2 \ddot{\theta}_1 \tag{A.20}$$

Adding the terms due to potential and kinetic energy, the total torque required from the base actuator for moving the point actuator is:

$$\tau_{A1} = \tau_{A1U} + \tau_{A1T} = M l_1^2 \ddot{\theta}_1 + M g l_1 \cos(\theta_1) \tag{A.21}$$

A.5.2 Distributed Mass Actuator

The same model for mass distribution is used as in the static analysis. For the manipulator shown in Figure A.5, again define the mass density of the actuator as ρ where $M = \rho z$. The total potential energy of the actuator is simply the product of the total mass, the gravitational acceleration, and the height of the center of mass, which is the midpoint of the actuator in the case of uniform mass distribution:

$$U_A = M g(\alpha l_1 \sin(\theta_1) + \frac{z}{2} \sin(\theta_1 + \theta 3) \tag{A.22}$$

The Torque due to this energy is:

$$\tau_{A1U} = \frac{d}{d\theta_1}(U_A) = M g(\alpha l_1 \cos(\theta_1) + \frac{z}{2} \cos(\theta_1 + \theta_3)) \tag{A.23}$$

As expected, this is equal to the torque needed in the static case. The compute the kinetic energy of a single "point" on the actuator, the linear (v) and ($\omega = \dot{\theta}_1$) angular velocities as well as the inertia (i) of that point at a radius r from the base are required:

$$T_a = \frac{1}{2} \rho \|v\|^2 = \frac{1}{2} i \omega^2 \tag{A.24}$$

The linear velocity is given by the time derivative of the position given in A.6:

$$v = \frac{d}{dt}\begin{bmatrix} x \\ y \end{bmatrix} = \frac{d}{d\theta_1}\begin{bmatrix} x \\ y \end{bmatrix}\dot{\theta}_1 \tag{A.25}$$

Note that the variation in ζ is ignored for simplicity. This is a reasonable assumption as the actuator may be connected close enough to the base of the second link that a small change in the actuator length is all that is needed to provide the roational motion needed. This expression simplifies to:

$$\begin{bmatrix} v_x \\ v_y \end{bmatrix} = \begin{bmatrix} -\alpha l_1 \sin(\theta_1) - \zeta \sin(\theta_1 + \theta_3) \\ \alpha l_1 \cos(\theta_1) + \zeta \cos(\theta_1 + \theta_3) \end{bmatrix}\dot{\theta}_1 \tag{A.26}$$

$$\|v\|^2 = v^T v = (\alpha^2 l_1^2 + \zeta^2 + \alpha l_1 \zeta \cos(\theta_3))\dot{\theta}_1^2 \tag{A.27}$$

The inertia term has:

$$i = \rho r^2 \tag{A.28}$$

$$r^2 = x^2 + y^2 = \alpha^2 l_1^2 + \zeta^2 + \alpha l_1 \zeta \cos(\theta_3) \tag{A.29}$$

Substituting in the expression for kinetic energy:

$$T_a = \frac{\rho}{2}\dot{\theta}_1^2(\alpha^2 l_1^2 + \zeta^2 + \alpha l_1 \zeta \cos(\theta_3)) \tag{A.30}$$

The total actuator kinetic energy is:

$$T_A = \int_0^z T_a d\zeta = \frac{\rho}{2}\dot{\theta}_1^2(\alpha^2 l_1^2 z + \frac{1}{3}z^3 + \frac{1}{2}\alpha l_1 z^2 \cos(\theta_3)) \tag{A.31}$$

Note that this expression does not depend on θ_1 (θ_3 is independent of θ_1 as defined above). Only the first term on the left of A.12 contributes to the generalized force for the kinetic energy. The torque due to this component of the energy is:

$$\tau_{A1T} = M\ddot{\theta}_1(\alpha^2 l_1^2 + \frac{1}{3}z^2 + \frac{1}{2}\alpha l_1 z \cos(\theta_3)) \tag{A.32}$$

The total required torque for moving the linear actuator is:

$$\tau_{A1} = Mg(\alpha l_1 \cos(\theta_1) + \frac{z}{2}\cos(\theta_1 + \theta 3) + M\ddot{\theta}_1(\alpha^2 l_1^2 + \frac{1}{3}z^2 + \frac{1}{2}\alpha l_1 z \cos(\theta_3)) \tag{A.33}$$

It is straight forward to see that this expression evaluates to a smaller number than the expression for the point mass actuator for some ranges of values of z, α and θ_3.

Part Actuator and Part String As mentioned at the end of the last section on static analysis, the given workspace and parameters of the actuators may allow us to combine a massless string with the linear actuator to form the total distributed mass actuator. Again, if we maintain the mass of the actuator, effectively increasing the mass density ρ for the shorter actuator, the only difference in the excess torque expressions is the substitution of z by z_{part}:

$$\tau_{A1} = Mg(\alpha l_1 \cos(\theta_1) + \frac{z_{part}}{2}\cos(\theta_1 + \theta 3) + M\ddot{\theta}_1(\alpha^2 l_1^2 + \frac{1}{3}z_{part}^2 + \frac{1}{2}\alpha l_1 z_{part}\cos(\theta_3)) \quad \text{(A.34)}$$

A.6 Endpoint Force

Assuming the same mass for the revolute and linear mass-distributed actuators produces torque requirements on the base actuators that favor the linear actuator. However, to fairly compare the two setups, the torques or forces required by the actuators to perform a task, ie. to produce an endpoint force, have to be examined. There are many ways of computing these relationships, and some are presented below. Although the results may not seem identical, computation of the forces using all results in Matlab confirms that they are different forms of the same result.

A.6.1 Revolute Actuator

As mentioned earlier, the relationship between endpoint forces F and joint torques τ is given by $\boldsymbol{\tau} = \boldsymbol{J}^T\boldsymbol{F}$. $\boldsymbol{J}$ is the Jacobian relating the endpoint velocity components to joint velocities. If the endpoint velocity is a vector $\boldsymbol{v}$ with the x and y components and the vector $\dot{\boldsymbol{q}}$ contains $dot\theta_1$ and $\dot{\theta}_2$, the two are related by

$$\dot{\boldsymbol{v}} = \boldsymbol{J}\dot{\boldsymbol{q}} \quad \text{(A.35)}$$

$$\begin{bmatrix} \dot{x} \\ \dot{y} \end{bmatrix} = \begin{bmatrix} -l_1\sin(\theta_1) - l_2\sin(\theta_1+\theta_2) & -l_2\sin(\theta_1+\theta_2) \\ l_1\cos(\theta_1) + l_2\cos(\theta_1+\theta_2) & l_2\cos(\theta_1+\theta_2) \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} \quad \text{(A.36)}$$

The forces are related by

$$\begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} -l_1 \sin(\theta_1) - l_2 \sin(\theta_1 + \theta_2) & l_1 \cos(\theta_1) + l_2 \cos(\theta_1 + \theta_2) \\ -l_2 \sin(\theta_1 + \theta_2) & l_2 \cos(\theta_1 + \theta_2) \end{bmatrix} \begin{bmatrix} F_x \\ F_y \end{bmatrix} \tag{A.37}$$

A.6.2 Linear Actuator System: Coordinate Transformation

If the velocities of the coordinates θ_1 and z, the actuator coordinates of the linear system, are related to the velocities of θ_1 and θ_2, the actuator coordinates of the revolute joint system, by

$$\dot{\boldsymbol{q}} = \begin{bmatrix} \dot{\theta_1} \\ \dot{\theta_2} \end{bmatrix} = \boldsymbol{J_c} \begin{bmatrix} \dot{\theta_1} \\ \dot{z} \end{bmatrix} = \boldsymbol{J_c} \dot{\boldsymbol{q}}_{\boldsymbol{l}} \tag{A.38}$$

then the endpoint velocity is related to the linear actuator system coordinate velocities by

$$\boldsymbol{v} = \boldsymbol{J} \boldsymbol{J_c} \dot{\boldsymbol{q}}_{\boldsymbol{l}} \tag{A.39}$$

and the actuator and endpoint forces are then related by

$$\boldsymbol{\tau_l} = \begin{bmatrix} \tau_1 \\ F_z \end{bmatrix} = \boldsymbol{J_c}^T \boldsymbol{J}^T \boldsymbol{F}. \tag{A.40}$$

An inspection of the partial derivatives that populate $\boldsymbol{J_c}$ leads to this form of the matrix:

$$\boldsymbol{J_c} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{d\theta_2}{dz} \end{bmatrix} \tag{A.41}$$

The zeroes in the off-diagonal positions are intuitive. The rate of change of θ_1 does not depend on how the length of the linear actuator is varying, and the rate of change of θ_2 does not depend on $\dot{\theta_2}$, as in the case of the revolute actuator. To compute $\boldsymbol{J_{c22}}$, a function relating θ_2 and z is needed. Using the law of cosines and substituting for $\cos(\pi - \theta_2)$ with $-\cos(\theta_2)$:

$$z^2 = (\alpha_2 l_2)^2 + ((1 - \alpha_1) l_1)^2 + 2\alpha_2 l_2 (1 - \alpha_1) l_1 \cos(\theta_2) \tag{A.42}$$

$$2z \frac{dz}{d\theta_2} = -2\alpha_2 l_2 (1 - \alpha_1) l_1 \sin(\theta_2) \tag{A.43}$$

$$\frac{d\theta_2}{dz} = \frac{-z}{\alpha_2 l_2 (1-\alpha_1) l_1 \sin(\theta_2)} \tag{A.44}$$

A.6.3 Linear Actuator System: Direct Differentiation

Another method for computing the Jacobian for the linear actuator system is to directly differentiate the position of the endpoint with respect to the generalized coordinates. This is the method used to compute for instance the Jacobian for the revolute actuator system above. Define θ_4 as the angle between the line on which the linear actuator lies and the link l_2. The endpoint coordinates are then:

$$x = \alpha_1 l_1 \cos(\theta_1) + z\cos(\theta_1+\theta_3) + (1-\alpha_2) l_2 \cos(\theta_1+\theta_3+\theta_4) \tag{A.45}$$

$$y = \alpha_1 l_1 \sin(\theta_1) + z\sin(\theta_1+\theta_3) + (1-\alpha_2) l_2 \sin(\theta_1+\theta_3+\theta_4) \tag{A.46}$$

The jacobian is then found by partial differentiation with respect to θ_1 and z:

$$J_l = \begin{bmatrix} \frac{dx}{d\theta_1} & \frac{dx}{dz} \\ \frac{dy}{d\theta_1} & \frac{dy}{dz} \end{bmatrix} \tag{A.47}$$

These derivatives are messy, as we need the derivatives of $\theta_{3,4}$ with respect to z to apply the chain rule. However, a numerical approximation of the derivative using these functions can be used to check the expression in the first section.

A.6.4 Linear Actuator System: Endpoint Torque Analysis

A more physical approach to computing the force F_z required by the linear actuator to transmit an endpoint force is to examine the moments around the second joint due to different forces. The torque τ_1 required by the base actuator is the same as in the revolute actuator case. Note that F_z can only produce endpoint force that is orthogonal to the second link l_2, call it $F_\perp$. The moment generated by F_z around the second joint and the moment generated by F (or equivalently $F_\perp$) around the same joint must be equal. The moment arms are $\alpha_2 l_2$ and l_2, respectively. Equating the magnitudes of the two moments

$$F_z \alpha_2 l_2 \sin(\theta_4) = \sqrt{F_x^2 + F_y^2} l_2 \sin(\theta_1 + \theta_2 - \theta_e) \tag{A.48}$$

$$F_z = \frac{\sqrt{F_x^2 + F_y^2} \sin(\theta_1 + \theta_2 - \theta_e)}{\alpha_2 \sin(\theta_4)} \tag{A.49}$$

where θ_e is the angle of the endpoint force vector.

$$\theta_e = \arccos\left(\frac{F_x}{\sqrt{F_x^2 + F_y^2}}\right) \quad 0 \leq \theta_e \leq \pi \tag{A.50}$$

or

$$\theta_e = \arcsin\left(\frac{F_y}{\sqrt{F_x^2 + F_y^2}}\right) \quad \frac{-\pi}{2} \leq \theta_e \leq \frac{\pi}{2} \tag{A.51}$$

From this expression for the linear actuator force, it is clear that when the arm is almost flat, $\theta_4 \approx 0$, or when the contact point α_2 is small (short moment arm for the linear actuator), the required F_z becomes very large.

A.6.5 Two Different Loads

The analysis that was developed above represents the endpoint force in cartesian coordinates. This representation can be used to describe a variety of loads, such as carrying a weight or pressing against a flat surface. However, it is perhaps more instructive in the analysis of the forces to be exerted by the second actuator, revolute or linear, to look at a force that is orthogonal to the second link. This force, or $F_\perp$ as referred to earlier, couples directly and completely with the force exerted by the actuator moving the second joint. The other component of the force is exerted on the stiff link and will be compensated for either by the stiffness of the structure or the actuator at the base joint. To incorporate this load into the framework described above, $F_\perp$ is transformed into its x and y components as follows:

$$\boldsymbol{F} = \begin{bmatrix} F_x \\ F_y \end{bmatrix} = \begin{bmatrix} \sin(\theta_1 + \theta_2) \\ -\cos(\theta_1 + \theta_2) \end{bmatrix} F_\perp \tag{A.52}$$

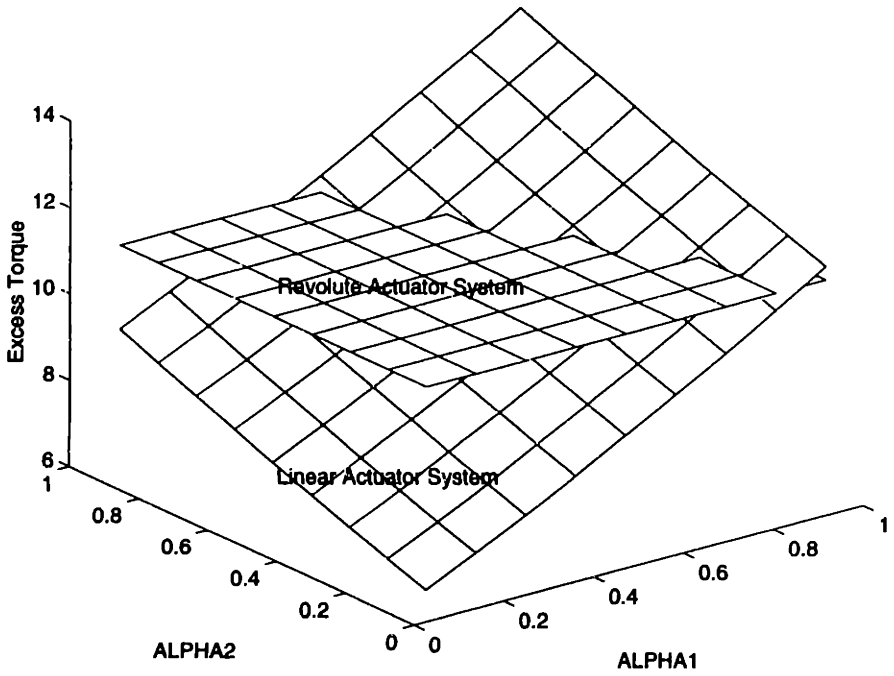


Figure A.6: Excess Torque Requirements on Base Actuator Due to Second Actuator Plotted Against Contact Points. Graph is for a given position in space and unit acceleration and mass.

A.7 Optimizing the Basic Design

Using the equations presented above, the basic linear actuator with contacts on the two links can be explored and optimized. By varying the contact points on the arm, different torque requirements on the base actuator can be achieved. The transfer function from linear actuator to endpoint force also changes with different contact points. Once the application specifications and the material properties of the actuators, such as mass and force density, are known, these tools can be used to design the arm. Figures A.6, A.7, A.8, and A.9 are plots of some of the functions presented above. Of course, a model of the dynamic system based on these forces should be used to simulate motion of the arm following the desired trajectories to further understand and optimize the design and control strategies to be used with this system. This is the subject of the next appendix.

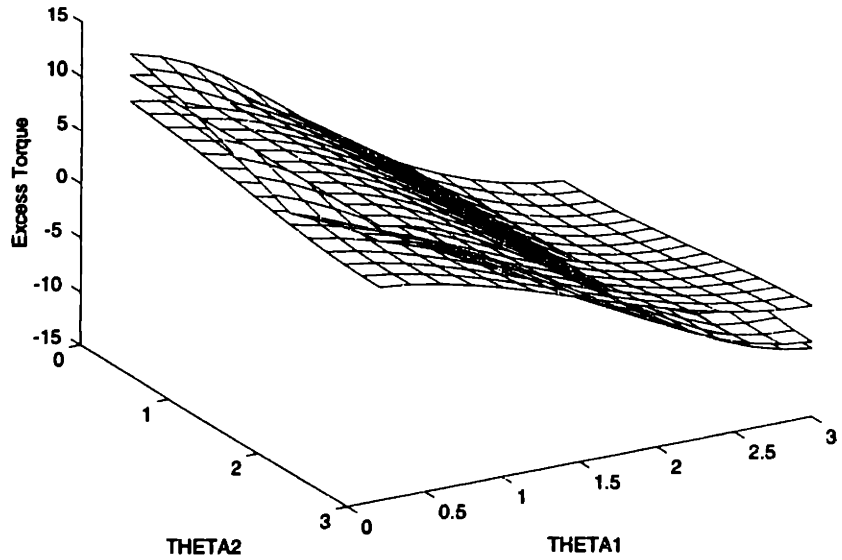


Figure A.7: Excess Torque Requirements on Base Actuator Due to Second Actuator Plotted Against Workspace Position. The outermost curve is for the revolute actuator system and the rest of the curves are for different contact points with the distributed mass actautor.

A.8 Future Designs

More complex mechanical designs and greater freedom in the distribution of the mass of the polymer gels will result in further improved dynamic response and range of motion for the manipulator system. A three-dimensional model with ball-and-socket joints as well as polymer gel actuators with their mass distributed in three-dimensions and according to numerical optimization will emphasize the advantages of a conformable actuator much more. In the model presented above the linear actuator needs to generate infinite force to move the second link when the arm is fully extended. The addition of a pulley the increases the angle of contact of the gel and the link may be the solution, with the pulley friction a new variable to be dealt with. Similar mechanical system design issues need to be explored for a practical system. Design optimization tools and creative mechanical designs should be the subject of further research.

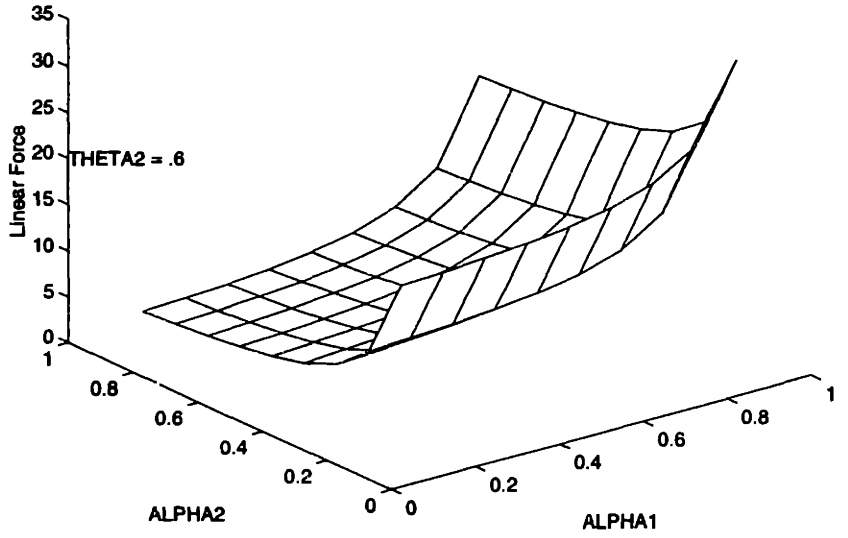


Figure A.8: Transfer Function from Endpoint Force Orthogonal to Second Link to Linear Actuator Force Plotted Against Contact Points. Graph is for a specific workspace position

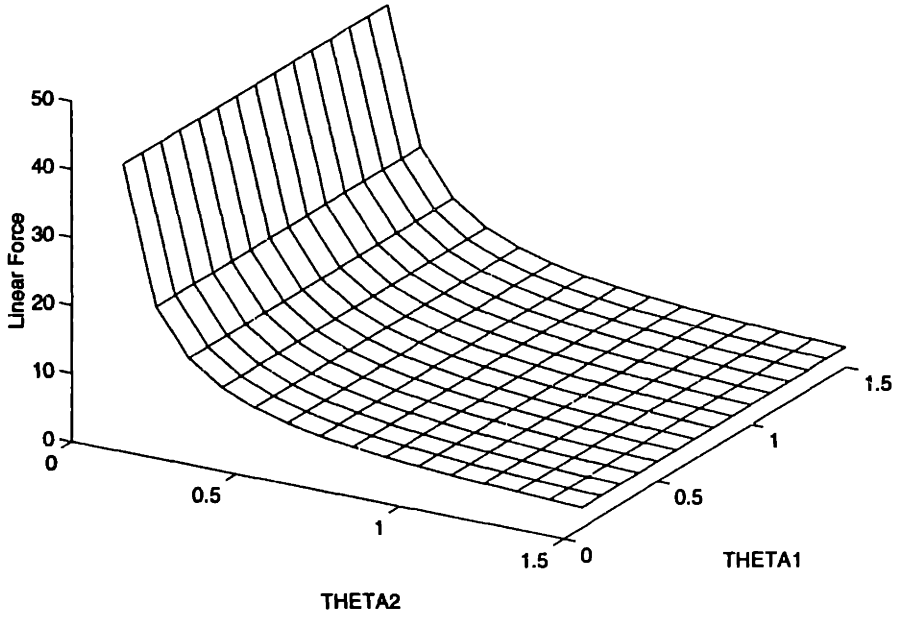


Figure A.9: Transfer Function from Endpoint Force Orthogonal to Second Link to Linear Actuator Force Plotted Against Joint Angles. Graph is for a specific pair of contact points.

Appendix B

Flexible Manipulator : Modeling and Control

In order to simulate the dynamics of the system and design controllers for it, we need a model of the arm. In the following sections, a basic model for the lumped actuator arm is developed and then modified to include the effects of actuator mass. The model is further altered to describe the distributed actuator system presented above. Possible methods of controlling the arm are introduced as well.

B.1 Modeling

This section describes the modeling of the manipulator system assuming the availability of a torque or force input from the actuators.

B.1.1 Basic Dynamic Model

The two revolute joint, direct-drive rotary motor system shown in Figure A.4 above can be modeled using its generalized coordinates, the joint angles, and generalized forces, the joint torques. The equations describing the system are shown below. They were found using the Lagrangian formulation of manipulator dynamics [2], and are written in the form shown with parameters a_i to aid in the adaptive control described later. The Lagrangian L is defined by:

$$L(q_i, \dot{q}_i) = T - U \tag{B.1}$$

With the dynamic equations given by:

$$\frac{d}{dt}\frac{\partial L}{\partial \dot{q}_i} - \frac{\partial L}{\partial q_i} = Q_i \tag{B.2}$$

T is the kinetic energy, U the potential energy, q_i the generalized coordinate (θ_i in the case of the revolute joint manipulator discussed below), and Q_i the generalized force (τ_i in the manipulator below). Working the equations out and assuming massless actuators:

$$\boldsymbol{\tau} = \boldsymbol{H}\ddot{\boldsymbol{\theta}} + \boldsymbol{C}\dot{\boldsymbol{\theta}} + \boldsymbol{G} \tag{B.3}$$

$$\begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} H11 & H12 \\ H21 & H22 \end{bmatrix} \begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} + \begin{bmatrix} -h\dot{\theta}_2 & -h(\dot{\theta}_1 + \dot{\theta}_2) \\ h\dot{\theta}_1 & 0 \end{bmatrix} \begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix} \begin{bmatrix} G_1 \\ G_2 \end{bmatrix} \tag{B.4}$$

where

$$H11 = a_1 + 2a_3\cos(\theta_2) \tag{B.5}$$

$$H12 = H21 = a_2 + a_3\cos(\theta_2) \tag{B.6}$$

$$H22 = a_2 \tag{B.7}$$

$$h = a_3\sin(\theta_2) \tag{B.8}$$

$$G_1 = a_4\cos(\theta_1) + a_3\frac{g}{l_1}\cos(\theta_1 + \theta_2) \tag{B.9}$$

$$G_2 = a_3\frac{g}{l_1}\cos(\theta_1 + \theta_2) \tag{B.10}$$

and

$$a_1 = I_1 + m_1 l_{c1}^2 + I_2 + m_2 l_{c2}^2 + m_2 l_1^2 \tag{B.11}$$

$$a_2 = I_2 + m_2 l_{c2}^2 \tag{B.12}$$

$$a_3 = m_2 l_1 l_{c2} \tag{B.13}$$

$$a_4 = (m_1 l_{c1} + m_2 l_1)g \tag{B.14}$$

I_i is the inertia of link i, m_i the mass, l_i the length of the link, l_{ci} the distance from a joint to the center of mass of a link, and g is the gravitational constant . There is no load in this setup for simplicity. Also, the mass of the links is not assumed to include the actuator masses. These will be added later in the development of the model, and endpoint forces can be included relatively easily to simulate an endpoint load.

B.1.2 Distributed Mass Actuator - Linear Gel

The dynamic equations for the distributed mass actuator are derived from those of the point mass actuator. The mass of the actuator is again left out in this analysis and added in the following sections. Since the joints of this system constrain the motion of the links in the same way as the system with the rotary actuator, they basically convert the motion of the linear actuator to that which was produced by the rotary one corresponding to the generalized force defined above. This allows us to capture the dynamics of the physical system by a coordinate transformation. Specifically, [2] relates two sets of generalized forces and velocities in the followin manner. If

$$\boldsymbol{v_1} = \boldsymbol{J v_2} \tag{B.15}$$

then

$$\boldsymbol{\tau_2} = \boldsymbol{J^T \tau_1} \tag{B.16}$$

Using this relationship, the input torque of the first system (with two direct drive rotary actuators) is transformed to that of the second through:

$$\boldsymbol{\tau_z} = \boldsymbol{J_c^T \tau} \qquad \boldsymbol{\tau} = \boldsymbol{J_c^{-T} \tau_z} \tag{B.17}$$

and

$$\boldsymbol{J_c} = \begin{bmatrix} 1 & 0 \\ 0 & \frac{d\theta_2}{dz} \end{bmatrix} \tag{B.18}$$

The new coordinates, for the linear actuator system, are θ_1 like before and z, the length of the linear actuator. The new forces are τ_1 as before and F_z, the linear force along the length of the linear actuator. In Figure A.5, $\alpha_{1,2}$ are the contact points of the linear actuator on the links 1 and 2, respectively. The zeroes in the off-diagonal positions are intuitive. The rate of change of θ_1 does not depend on how the length of the linear actuator is varying, and the rate of change of θ_2 does not depend on $\dot{\theta}_2$, as in the case of the revolute actuator. To compute $\boldsymbol{J_{c22}}$, a function relating θ_2 and z is needed. Using the law of cosines and substituting for $\cos(\pi - \theta_2)$ with $-\cos(\theta_2)$, we find the expression found in the previous appendix:

$$\frac{d\theta_2}{dz} = \frac{-z}{\alpha_2 l_2 (1-\alpha_1) l_1 \sin(\theta_2)} \tag{B.19}$$

With this new matrix, the dynamic equations of the new system can be kept the same for the purpose of simulations and the command input transformed to the rotary torque command through this transformation. For purposes of control, the required rotary torque can be computed from the measured outputs and states of the system and then transformed into the required linear force using the linear transformation. It is noteworthy that the new dynamic equations can also be aquired through this transformation, which amounts to a similarity transformation for the H and C matrices and a pre-multiplication by $\boldsymbol{J^T}$ of G.

B.1.3 Additional Dynamics due to Actuator Mass

The previous section introduced the models for the system with two rotary actuators and the system with a rotary base joint actuator and a linear actuator for the second joint. The models ignored the actuator masses, however, and the only difference was in the force transfer from the linear actuator to the revolute joint. Incorporating the effects of the actuator masses involves adding the torque terms computed in Appendix A to the massless actuator models. The torque terms (total additional torque on the base actuator due to the second joint actuator) are repeated here for clarity.For the direct-drive, two rotary actuator system:

$$\tau_{A1} = \tau_{A1U} + \tau_{A1T} = Ml_1^2\ddot{\theta}_1 + Mgl_1\cos(\theta_1) \tag{B.20}$$

For the linear, distributed-mass actuator:

$$\tau_{A1} = Mg(\alpha l_1\cos(\theta_1) + \frac{z}{2}\cos(\theta_1 + \theta 3) + M\ddot{\theta}_1(\alpha^2 l_1^2 + \frac{1}{3}z^2 + \frac{1}{2}\alpha l_1 z\cos(\theta_3)) \tag{B.21}$$

In order to incorporate these terms into the dynamic equations, note that the additional torque terms are the superposition of two terms, one the result of potential energy and one the result of kinetic energy.

$$\tau_{A1} = \tau_{A1U} + \tau_{A1T} = \tau_{A1U} + H11_A\ddot{\theta}_1 \tag{B.22}$$

The new dynamic system is the same as the old one with τ_{A1U} added to G_1 and $H11_A$ added to $H11$ in Equation B.4. The new lumped actuator system is:

$$\begin{bmatrix} \tau_1 \\ \tau_2 \end{bmatrix} = \begin{bmatrix} H11 + H11_A & H12 \\ H21 & H22 \end{bmatrix}\begin{bmatrix} \ddot{\theta}_1 \\ \ddot{\theta}_2 \end{bmatrix} + \begin{bmatrix} -h\dot{\theta}_2 & -h(\dot{\theta}_1 + \dot{\theta}_2) \\ h\dot{\theta}_1 & 0 \end{bmatrix}\begin{bmatrix} \dot{\theta}_1 \\ \dot{\theta}_2 \end{bmatrix}\begin{bmatrix} G_1 + \tau_{A1U} \\ G_2 \end{bmatrix} \tag{B.23}$$

The equation for the distributed actuator system is achieved through the same transformation of the linear force to rotary torque at the second joint. τ_1 is the same as the lumped actuator system since the transformation does not affect the base torque. τ_2 is unaffected by the extra torque due to the actuator mass since this load is only seen by the base actuator.

B.2 Control

In order fully control the gel actuated system, a model for the polymer gel actuator as well as the mechanical system is required. A model for the actuator will depend on the mechanical model describing the the material's behavior when actuated (by changing temperature for instance) as well as how the overall artificial muscle or linear actuator is constructed. For instance, if the discontinous phase transition is modeled by a step input of force to a mass-spring-damper system, that model is at the heart of the actuator description. If many

strands of actuator are then used to somehow additively deliver force to the manipulator arm, the aggregate model of these gel units is the desired actuator model. The result may be a quantized force delivery unit, for instance, and the effects of the quantization have to be understood and dealt with. The control designs introduced below assume a continuous, unlimited force input from the actuators. While the effects of any limitations on the actuators can be examined from simulations with various controllers, these limitations have not been analytically examined in the control design due to the lack of a satisfactory actuator model at the time being. The additional mass of the actuators has been included in the control design.

B.2.1 Position Control

For position control, a PD (Proportional Derivative) controller is enough to stabilize the system around the desired position as an equilibrium point. Choosing positive definite matrices$\boldsymbol{K_P}$ and $\boldsymbol{K_D}$ such that

$$\boldsymbol{\tau} = -\boldsymbol{K_P}\tilde{\boldsymbol{\theta}} - \boldsymbol{K_D}\dot{\boldsymbol{\theta}} \tag{B.24}$$

and choosing the lyapunov function

$$V = \frac{1}{2}[\dot{\boldsymbol{\theta}}^T \boldsymbol{H}\dot{\boldsymbol{\theta}} + \tilde{\boldsymbol{\theta}}^T \boldsymbol{K_P}\tilde{\boldsymbol{\theta}}] \tag{B.25}$$

then

$$\dot{V} = -\dot{\boldsymbol{\theta}}^T \boldsymbol{K_D}\dot{\boldsymbol{\theta}} \tag{B.26}$$

which is negative semidefinite guarantee the stability of the system.

B.2.2 Feedback Linearization

Feedback linearization control aims to cancel out the nonlinear dynamics of a system through feedback control to result in a linear plant which is easier to understand and control. If the matrices defined in Equations B.4 and B.23 are used to construct the control law given that the controller can command a desired torque, choose a function of the form:

$$\boldsymbol{\tau} = \boldsymbol{H\nu} + \boldsymbol{C\dot{\theta}} + \boldsymbol{G} \tag{B.27}$$

If the desired angles, $\boldsymbol{\theta_d}$ are computed from the desired endpoint trajectories as well as their first and second derivatives, $\boldsymbol{\nu}$ can be chosen to be:

$$\boldsymbol{\nu} = \boldsymbol{\ddot{\theta}_d} - 2\lambda\dot{\tilde{\boldsymbol{\theta}}} - \lambda^2\tilde{\boldsymbol{\theta}} \tag{B.28}$$

where the error in a variable $x - x_d$ is denoted by $\tilde{x}$. When this expression is used for the torque on the left side of the dynamic equations above, the equations simplify to:

$$\ddot{\tilde{\boldsymbol{\theta}}} + 2\lambda\dot{\tilde{\boldsymbol{\theta}}} + \lambda^2\tilde{\boldsymbol{\theta}} = 0 \tag{B.29}$$

This equation describes a second order system where the two poles are at $-\lambda$ on the real axis. By choosing λ, the speed of the system, or the decay of the error, can be determined. For different trajectories, different convergence rates may be desired. For more on this control strategy, see [91].

For the lumped actuator system, the matrices with the additional terms for the excess inertia of the actuator mass should be used in the formulation above. For the distributed actuator system, the appropriate matrices should be used as well. Moreover, once the controller computes the torque required in the θ_2 coordinate, or the second element of $\boldsymbol{\tau}$, that element should be transformed into the required linear force in the linear actuator system. It can then command that force from the polymer gel actuator, within its physical limitations.

B.2.3 Robust Control

In many situations, uncertainty in the properties or environment of the motion system prevent the designer from capturing the full dynamic behavior of the system. The errors in the mathematical model may be large enough to cause instability or other loss of performance. One nonlinear control technique used to insure that the system remains stable with such modeling errors is robust, or switching mode, control. Given limits on the errors in the

model of the system, a robust controller provides stability for all systems with parameters within these bounds. As a common controller for nonlinear systems, robust control will be examined in the context of the flexible manipulator.

Define the following variables:

$$\dot{\boldsymbol{\theta}}_r = \dot{\boldsymbol{\theta}}_d - \boldsymbol{\Lambda}\tilde{\boldsymbol{\theta}} \tag{B.30}$$

$$\boldsymbol{s} = \dot{\boldsymbol{\theta}} - \dot{\boldsymbol{\theta}}_r = \dot{\tilde{\boldsymbol{\theta}}} + \boldsymbol{\Lambda}\tilde{\boldsymbol{\theta}} \tag{B.31}$$

According to Lyapunov theory, if a positive linear function is found with a time derivative that is negative definite along the trajectories of a system, then the system is stable. This is discussed in detail along with more on the mathematical formulation in [91]. If the positive definite function V is defined:

$$V = \frac{1}{2}\boldsymbol{s}^T\boldsymbol{H}\boldsymbol{s} \tag{B.32}$$

and $\dot{V}$ is found using the chain rule and the dynamic equations of the system to be:

$$\dot{V} = \boldsymbol{s}^T(\boldsymbol{\tau} - \boldsymbol{H}\ddot{\boldsymbol{\theta}}_r - \boldsymbol{C}\dot{\boldsymbol{\theta}}_r - \boldsymbol{G}) \tag{B.33}$$

The following control torque is defined using the estimates of the parameters of the dynamic system:

$$\hat{\boldsymbol{\tau}} = \hat{\boldsymbol{H}}\ddot{\boldsymbol{\theta}}_r + \hat{\boldsymbol{C}}\dot{\boldsymbol{\theta}}_r + \hat{\boldsymbol{G}} \tag{B.34}$$

$$\boldsymbol{\tau} = \hat{\boldsymbol{\tau}} - \boldsymbol{k}sgn(\boldsymbol{s}) \tag{B.35}$$

In order to guarantee an upper bound on $\dot{V}$ the following condition on $\boldsymbol{k}$ is needed:

$$\boldsymbol{k}_i \geq \|[\tilde{\boldsymbol{\tau}} = \tilde{\boldsymbol{H}}\ddot{\boldsymbol{\theta}}_r + \tilde{\boldsymbol{C}}\dot{\boldsymbol{\theta}}_r + \tilde{\boldsymbol{G}}]_i\| + \eta_i \tag{B.36}$$

The notation $\tilde{x} = \hat{x} - x$ where $\hat{x}$ is the estimate for the parameter x. This is satisfied if $\boldsymbol{k}$ is computed using upper limits on the deviations from the dynamic system matrices:

$$\|\tilde{\boldsymbol{H}}\| \leq \boldsymbol{L_H} \quad \|\tilde{\boldsymbol{C}}\| \leq \boldsymbol{L_C} \quad \|\tilde{\boldsymbol{G}}\| \leq \boldsymbol{L_G} \tag{B.37}$$

$$\boldsymbol{k} = \boldsymbol{\eta} + \boldsymbol{L_H}\|\ddot{\boldsymbol{\theta}}_r\| + \boldsymbol{L_C}\|\dot{\boldsymbol{\theta}}_r\| + \boldsymbol{L_G} \tag{B.38}$$

where the inequality signs are element-wise. Increasing $\boldsymbol{\eta}$ increases the bandwidth of the system and results in higher frequency chattering but smaller error around the desired trajectory. $\boldsymbol{\Lambda}$ governs the rate of convergence of the error.

B.2.4 Adaptive Control

Robust control methods require some knowledge of the bounds on the uncertain terms in the dynamic system model. If such knowledge is not available, but the designer knows the structure of the system and where the uncertainty enters the system, then adaptive control allows for the on-line estimation of these parameters and modification of the controller based on the updated model of the system. In particular, [91] outlines a method for the implementation of this technique that would be applied to the flexible manipulator system in the following manner. Using the system matrices defined above, factor the following expression into a vector with the structure of the mechanical system multiplied by a vector of unknown parameters.

$$\boldsymbol{\tau} = \boldsymbol{H}\ddot{\boldsymbol{\theta}}_r + \boldsymbol{C}\dot{\boldsymbol{\theta}}_r + \boldsymbol{G} = \boldsymbol{Y}(\boldsymbol{\theta}, \dot{\boldsymbol{\theta}}, \dot{\boldsymbol{\theta}}_r, \ddot{\boldsymbol{\theta}}_r)\boldsymbol{a} \tag{B.39}$$

The choice of lyapunov function for adaptive control in this case will be:

$$V = \frac{1}{2}(\boldsymbol{s}^T\boldsymbol{H}\boldsymbol{s} + \tilde{\boldsymbol{a}}\boldsymbol{\Gamma}^{-1}\tilde{\boldsymbol{a}}) \tag{B.40}$$

Choosing an expression for the control torque based on initial estimates for the parameters $\hat{\boldsymbol{a}}$, the derivative of the lyapunov function simplifies to:

$$\boldsymbol{\tau} = \boldsymbol{Y}\hat{\boldsymbol{a}} - \boldsymbol{K_D}\boldsymbol{s} \tag{B.41}$$

$$\dot{V} = s^T Y \tilde{a} - s^T K_D s + \dot{\hat{a}}^T \Gamma^{-1} \tilde{a} \tag{B.42}$$

where K_D is a positive definite matrix. If the adaptation law for the unknown parameters is chosen to be:

$$\dot{\hat{a}} = -\Gamma Y^T s \tag{B.43}$$

$$\dot{V} = -s^T K s \leq 0 \tag{B.44}$$

Changing Γ controls the rate at which the parameters converge to values that allow the system to track the desired trajectory. (These are not necessarily the correct values. If the excitation of the system is not rich enough there may be more than one set of parameters that allow it to track the commanded trajectory.) K_D is related to the rate of convergence of the error and its bounds.

Appendix C

Computing the Effective Duty Ratio D_1 for Control

In Chapter 4, averaged models for systems with different time-scale properties were developed. For some cases, the averaged model was derived using a new "effective" duty ratio seen by the mechanical system. For example, for a discontinuous volume-phase transition gel, with subsystem time scales as shown in Figure 4.14 and with the circuit analogy in Figure 4.7, using the input duty ratio D to directly define an averaged model fails because of violation of time-scale separation requirements for averaging. The input natural frequency is not much higher than the switching frequency, and we cannot therefore approximate the input over a switch cycle by its average. Figure 4.16 illustrated the failure of this technique to describe the average dynamics in the system. However, as shown in Figures 4.15 and 4.17, using the duty ratio D_1 of the square wave at the output of the Schmitt Trigger due to the original switching function, we can derive accurate averaged models. This Appendix describes how to obtain this duty ratio D_1 given the thermal dynamics and characteristics of the nonlinear equilibrium block. By inverting this mapping, the user specified duty ratio D can be used for control.

C.1 D_1 as a Function of D

We assume no hysteresis for this example. This simplification will reduce the number of constants and allow a more succinct mathematical description of the problem. Generalization

to hysteresis equilibrium blocks is straight forward. The relationship between the functions $D(t)$ and $D_1(t)$ can be summarized as shown in Figure C.1. At the input to the complete system, we have the switching function $q(t)$ driving a thermal or input subsystem. At the output of the thermal system, the gel temperature $T_g(t)$ is a switched, exponential waveform as shown[1]. This temperature waveform drives the gel dynamics through the nonlinear equilibrium block. If $T_g(t)$ is higher than the transition temperature T_t, the gel shrinks in response to an input of U_h, and $y(t)$ increases in the circuit analogy. When $T_g(t) < T_t$, the gel expands and $y(t)$ decreases in response to the input U_l. Whether the gel temperature is above or below transition is determined by the switching function $q_1(t)$. When $q_1(t)$ is high, the input U_h drives the circuit. When it is low, or 0, U_l is the drive. With these constant drives and the duty ratio $D_1(t)$ associated with $q_1(t)$, an averaged model can be derived as was illustrated in Figure 4.17. We need to be able to estimate the function $D_1(t)$ in order to derive this model, however.

Figure C.2 will aid in the development of analytical expressions relating $D_1(t)$ and $D(t)$. The temperature *rails*, or the two temperatures of the water circulators, are T_h and T_l, a high and a low temperature, repsectively (on either side of the transition temperature T_t). Assume that the temperature $T_g(t)$ is below the transition temperature when the switch is flipped to connect the hot water reservoir (at a temperature T_h) to the thermal system. The thermal system is represented by a first order linear RC circuit model. For a fraction of the switch period, namely for DT seconds, the temperature $T_g(t)$ is described by the equation:

$$T_g(t) = T_h(1 - e^{\frac{-t}{\tau}}) + T_o e^{\frac{-t}{\tau}}, \tag{C.1}$$

where T_o is the initial temperature of the gel, $\tau = RC$, and we have directly applied the solution to a first order LTI system driven by a step input and an initial condition. When the gel temperature is above transition and the switch is flipped to the cold reservoir, the the temperature T_g is:

[1]Recall that when the thermal natural frequency was much higher than the switching frequency, the temperature was a square-wave function.

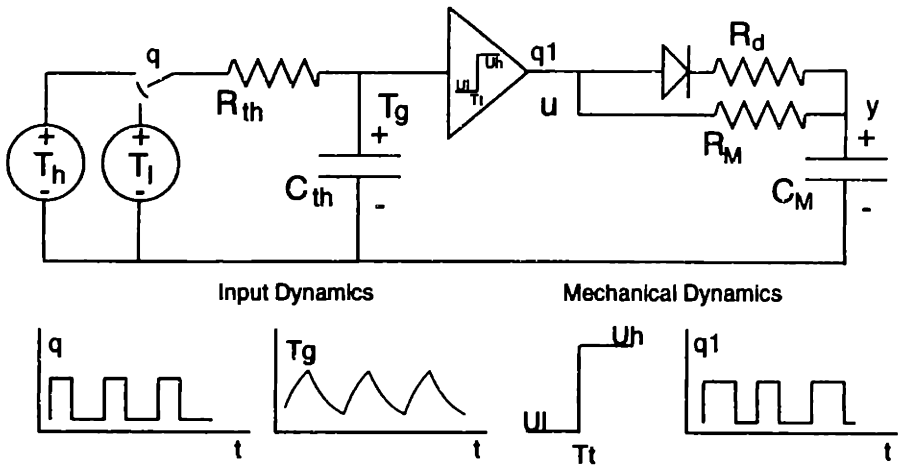


Figure C.1: Switching functions in an actuator system with slow input dynamics.

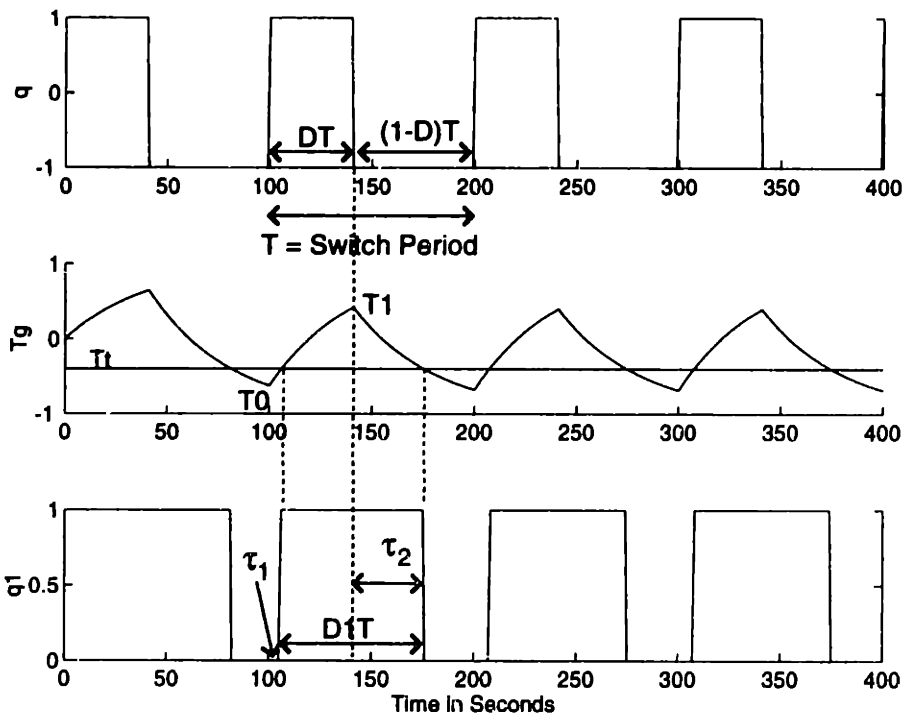


Figure C.2: Output switching function q_1 as a result of input switch q and the thermal dynamics.

$$T_g(t) = T_l(1 - e^{\frac{-t}{\tau}}) + T_1 e^{\frac{-t}{\tau}} \tag{C.2}$$

where T_1 is the initial value of the temperature when the switch is flipped from hot to cold. This temperature T_1 is also the final value before the switch is turned, expressed as :

$$T_1 = T_h(1 - e^{\frac{-DT}{\tau}}) + T_o e^{\frac{-DT}{\tau}} \tag{C.3}$$

The temperature $T_g(T)$ crosses the transition temperature at times τ_1 and $\tau = \tau_2 + DT$. The transition temperature is a function of τ_1 :

$$T_t = T_h(1 - e^{\frac{-\tau_1}{\tau}}) + T_o e^{\frac{-\tau_1}{\tau}} \tag{C.4}$$

and τ_2 :

$$T_t = T_l(1 - e^{\frac{-\tau_2}{\tau}}) + T_1 e^{\frac{-\tau_2}{\tau}} \tag{C.5}$$

Rearranging these equations, we find expressions for the exponentials that are a function of the times τ_1 and τ_2. First, for τ_1 :

$$e^{\frac{-\tau_1}{\tau}} = \frac{T_t - T_h}{T_o - T_h} \tag{C.6}$$

For τ_2 :

$$e^{\frac{-\tau_2}{\tau}} = \frac{T_t - T_l}{T_1 - T_l} \tag{C.7}$$

Solving the two equations above for τ_1 :

$$\tau_1 = \tau \ln \frac{T_t - T_h}{T_o - T_h} \tag{C.8}$$

and for τ_2

$$\tau_2 = \tau \ln \frac{T_t - T_l}{T_1 - T_l} \tag{C.9}$$

Using these time quantities, the amount of time that $T_g(t)$ is above the transition temperature during this cycle can be determined:

$$D_1 T = DT - \tau_1 + \tau_2 \tag{C.10}$$

Therefore, dividing by the period T, the new duty ratio is :

$$D_1 = D - \frac{\tau_1}{T} + \frac{\tau_2}{T} \tag{C.11}$$

Using this relationship, and the set of equations described above, we can find the duty ratio D_1 as a function of the initial temperature T_o at the beginning of a period, the transition temperature T_t, and the switch period T. The switch period is constant in our operation scheme, and D is user-defined at the beginning of each cycle. Using a thermal sensor, $T_o[n+1]$ can be identified at the beginning of each period, or predicted using an observer and a sampled data model that describes the dynamics of T_o given a switching profile. Given the temperature $T_1[n]$ and $D[n]$ for switcht cycle n, the initial temperature for the next cycle $T_o[n+1]$ can be approximated:

$$T_o[n+1] = T_1[n]e^{\frac{-(1-D[n])T}{RC}} + T_l(1 - e^{\frac{-(1-D[n])T)}{RC}}) \tag{C.12}$$

C.2 3-D Matlab Plot of the Nonlinear Surface

The equations of the last section provide a nonlinear description of D_1 as a function of D and T_o given a fixed period T. The following Matlab code was used to investigate properties of this function. The file arbitrarily uses values of $T_h = 50$, $T_l = 18$, $R = 2.5e5$, $C = 6e-5$, $T_t = 34$. The value of D_1 saturates at 1 and at 0. This file also implements an observer for the initial value of the temperature at the beginning of every cycle to eliminiate the need for a thermal sensor.

```
%This file generates different values for D1 given
%a specific value for TO and D.

%%Thermal System Parameters
rup = 60;
rdown = 20;
R = 1e6;
C = 6e-5;
tau = R*C;
```

```
%Range for duty ratios D as input to function
dnl = .05:.05:.95;
%Range of initial temperatures T0 as input to function
t0nl = 22:.5:59;
%Transition Temp
tt = 37;
%Switch Period
T = 174;
%Outer loop for cycling through the init temperature
for i = 1:length(t0nl)
%Inner loop for cycling through D values
    for j = 1:length(dnl)
tau1 = -R*C*log((tt - rup)/(t0nl(i)-rup));
        t1(i,j) = exp((-dnl(j)*T)/(R*C))*t0nl(i) +
rup*(1-exp((-dnl(j)*T)/(R*C)));
        tau2 = dnl(j)*T - R*C*log((tt -
rdown)/(t1(i,j)-rdown));
%getting the next t0 observed in terms of d and t0
t0test(i,j) = exp((-(1-dnl(j))*T)/(R*C))*t1(i,j) +
rdown*(1-exp((-(1-dnl(j))*T)/(R*C)));
d1nl(i,j) = (tau2-tau1)/T;
    end
end
figure(1)
%plot 3D plot
figure(1)
mesh(dnl,t0nl,d1nl);
xlabel('D')
ylabel('T0')
zlabel('D1')
figure(2)
mesh(dnl,t0nl,t0test);
xlabel('D0[n]')
ylabel('T0[n-1]')
zlabel('T0[n]')
```

Using this script, three-dimensional plots of D_1 as a function of D and T_o are generated. Figure C.3 illustrates an example for the parameters in the code above. Figure C.4 shows the initial temperature at the beginning of the n*th* cycle as a function of the initial temperature of the previous (n-1) cycle and the duty ratio D. Given the user-defined D and the switch cycle initial temperature T_o, D_1 may be identified and used to generate the averaged model for the system.

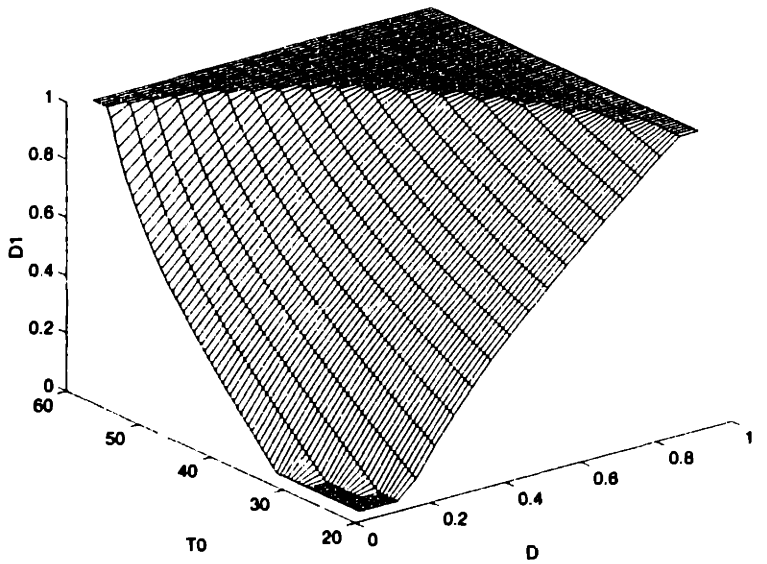


Figure C.3: Three-dimensional plot of D_1 as a function of D and T_o.

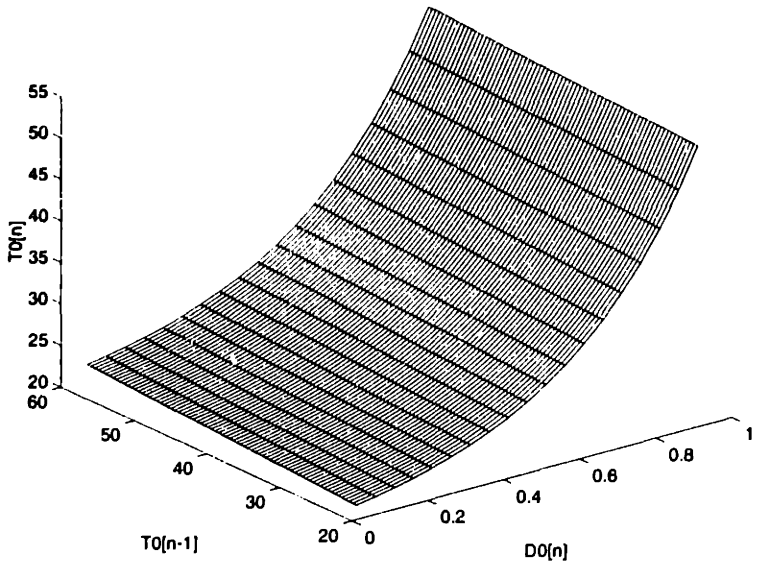


Figure C.4: Three-dimensional plot of $T_o[n]$ as a function of D[n] and $T_o[n-1]$.

C.3 Inverting the Mapping from D to D_1

In the case where the duty ratio seen by the mechanical system is different from the user generated duty ratio, another approach to control design can be used. Define the mapping between D, T_o, and D_1 as :

$$D_1 = f(D, T_o) \tag{C.13}$$

If this function is inverted in the controller, the overall effect will be a cancellation of this mapping. This permits direct application of the duty ratio of the input switch function to the mechanical system. However, in general, this inverse function must be solved using nonlinear minimization routines which are not suitable for real-time control. A sampled-data system requires strict constraints on the time of execution of the control algorithm, which may not be guaranteed with the minimization routine. However, using some approximations, an invertible function may be used to represent the relationship between D and D_1. Figure C.5 shows the plots of D_1 versus D for three different initial temperatures. For small-signal operation, where D and T_o do not vary significantly around an operating point, an approximation can be used to describe this relationship. For instance, for $T_o \approx 23°C$ and $0.4 < D < 0.8$, the linear approximation

$$\tilde{D}_1 \approx 1.2\tilde{D} - 0.07 \tag{C.14}$$

can be used, as shown in Figure C.5. We can invert this function by using the mapping :

$$D_2 = \frac{\tilde{D} + 0.07}{1.2} \tag{C.15}$$

Including this function in series with a controller cancels out the effect of the system mapping on the duty ratio and allows the user to design controllers with the variable D as the output of a linear controller directly commanding the duty ratio seen by the gel or mechanical system. This is illustrated in Figure C.6. More accurate approximations can be derived using higher order descriptions or by including the effect of T_o on the duty ratio.

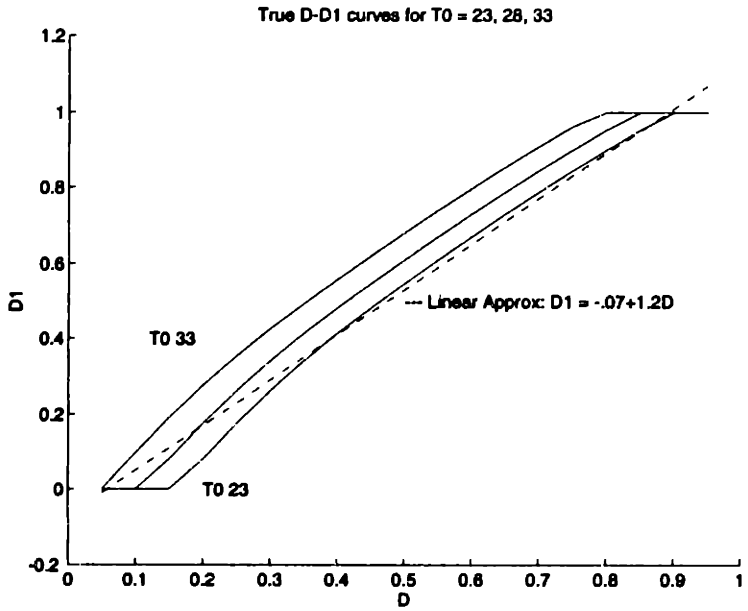


Figure C.5: D_1 vs. D near at different, fixed initial temperatures.

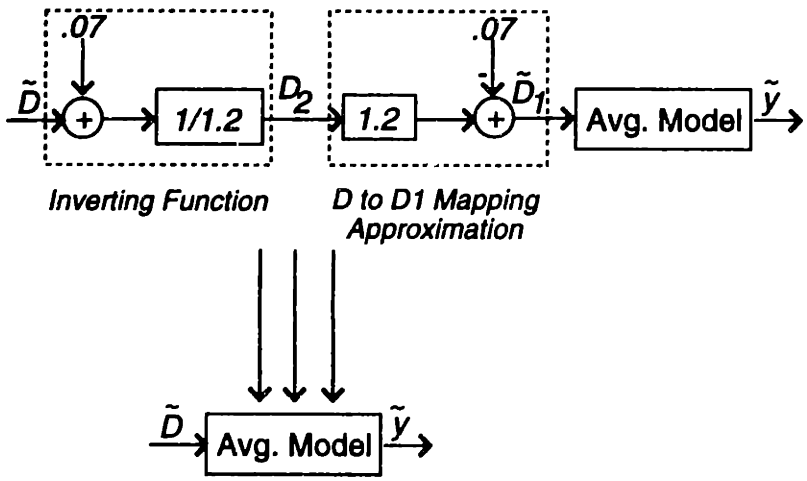


Figure C.6: Inverting the mapping of D to D_1.

Appendix D

PWM Closed-Loop Control Using the Small-Signal Gel Model

In Chapter 7, closed-loop controllers were designed and implemented for the prototype polymer gel linear actuator. This appendix describes the process of control design using the small-signal LTI gel model.

D.1 Developing LTI Averaged Models for Thermally Driven Gel Actuators

The block diagram in Figure D.1 represents the different components of the prototype gel actuator and the experimental setup described in Chapter 5. The thermal reservoirs, T_h and T_l, have been lumped into the thermal system. By modulating the duty ratio D of the fixed frequency switch function $q(t, D)$, the temperature of the gel $T_g(t)$ can be altered, thus producing a change in the gel length or mechanical load position $y(t)$. In general, both the thermal and gel systems might not be LTI. For example, the dynamics of the thermal system may depend on the switch state, causing it to be time-varying, or the dynamics of the gel may depend on the input $T_g(t)$, resulting in a nonlinear system. Figure D.1 illustrates the general case where the input switching function affects the dynamics of the gel and thermal systems.

As mentioned in Chapter 4, by selecting an appropriate switching frequency $\frac{1}{T}$ in this system configuration, the switch ripple in the output variable may be minimized, thus reg-

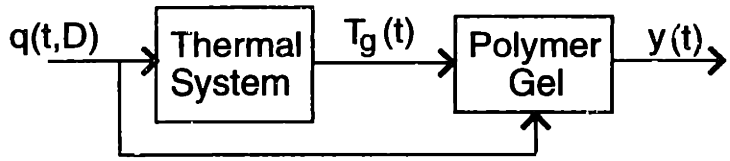


Figure D.1: Full switched system.

ulating $y(t)$ to an average value with small, cyclic deviations in its vicinity. A compensator can be used in closed-loop around this system to modulate the pulse-width of the square wave DT in order to control this average value. State-space averaging and small-signal linearization were used in Chapter 7 to illustrate that, in general, an LTI averaged model for the *gel* system in Figure D.1 will take on the following structure:

$$\dot{\tilde{\boldsymbol{x}}}(t) \approx \boldsymbol{A}_{av}\tilde{\boldsymbol{x}}(t) + \boldsymbol{B}_{av}\tilde{u}(t) + \boldsymbol{B}_{d}\tilde{D}(t), \tag{D.1}$$

where $\boldsymbol{x}$ is the *averaged* state vector for the gel system, and u is the *averaged* temperature of the gel ($\overline{T}_g = u$). The tilde superscripts indicate small deviations from some nominal value, emphasizing that this is a small-signal model. The matrices $\boldsymbol{A}_{av}$, $\boldsymbol{B}_{av}$, and $\boldsymbol{B}_{d}$ are only valid around a nominal operating point. The output variable $\tilde{y}(t)$ is a weighted combination of the states, $\tilde{\boldsymbol{x}}(t)$, and the inputs $\tilde{u}(t)$ and $\tilde{D}(t)$. For more on this formulation refer to Equations 7.6 - 7.11.

An averaged model for the thermal system may also be developed to relate the duty ratio $\tilde{D}(t)$ to the average temperature $\tilde{u}(t)$. For example, if the thermal system natural frequency is much higher than the switching frequency, then the average temperature (averaged over a switch cycle) is:

$$\tilde{u}(t) = \tilde{D}(t)T_h + (1 - \tilde{D}(t))T_l = T_l + \tilde{D}(t)(T_h - T_l) \tag{D.2}$$

where we have assumed that the fluid at temperature T_h circulates for $\tilde{D}(t)T$ seconds in a switch period and the fluid at temperature T_l circulates for $(1 - \tilde{D}(t))T$ seconds. The average model reduces to a gain block with a disturbance input T_l. Even if the thermal

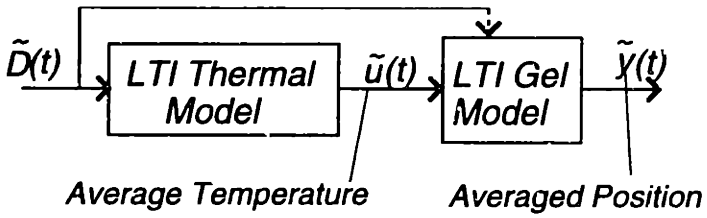


Figure D.2: Cascading the thermal and gel models for control design.

system dynamics are *not* that fast, LTI averaged models may still be derived. The thermal system, as described in Chapter 6, is well-modeled by a first-order LTI system in each switch state:

$$\dot{T}_g(t) = -\frac{T_g(t)}{\tau_{th}} + \frac{T_r(t)}{\tau_{th}} \tag{D.3}$$

where $T_r(t)$ is the input to the system, or the temperature of the circulated fluid from the water reservoirs (T_h or T_l), and $T_g(t)$ is the temperature of the gel. The time constant τ_{th} is determined by the thermal inertia and resistance of the glass container and the solvent. In general, the time constant will change as a function of the switch state. As a result, the averaged system matrices will be a function of the duty ratio. In this case, the resulting average thermal system is nonlinear (if $\tilde{D}(t)$ is a modulated control input). Under small-signal assumptions, however, an LTI averaged model may be derived as shown in Chapter 7 and in [45]. The overall system, shown in Figure D.1, may now be modeled by LTI system components as shown in Figure D.2. Using this LTI model description for the thermal and gel systems, controllers may be designed for a gel actuator.

D.2 Averaged Modeling and Control for Prototype Gel Actuator

D.2.1 Thermal System Model

For our experimental setup, it was generally possible to tune the two system states, the heating and cooling, so that the first-order model is approximately the same for both.

Assume in the thermal model above (Equation D.3) that for $D(t)T$ seconds in a switch period T, the circulating fluid has temperature $T_r = T_h$. For the remaining $(1 - D(t))T$ seconds in a switch period, it has temperature $T_r = T_l$. If the time constant τ_{th} is the same regardless of the switch state, then an averaged model for the thermal system is:

$$\dot{u}(t) = -\frac{u(t)}{\tau_{th}} + \frac{D(t)(T_h - T_l)}{\tau_{th}} + \frac{T_l}{\tau_{th}} \tag{D.4}$$

where u is again the average gel temperature. This equation describes a first-order LTI model with input $D(t)$ and disturbance input T_l.

D.2.2 Polymer Gel Model

In Chapter 7, an averaged model for the gel dynamics was identified. The model relates the gel states and output to the average temperature $\tilde{u}(t)$ and an external force $F_2(t)$. The external force input $F_2(t)$ is regarded as a disturbance input in our model. It is interesting to note that the identified average gel model was not a function of the input duty ratio as is the general case shown in Figure D.2. This model simplification can occur in one of the following cases:

- For slow enough variations in the duty ratio, deviations in the average temperature $\tilde{u}(t)$ will track changes in $\tilde{D}(t)$, so that $\tilde{D}(t) = \alpha\tilde{u}(t)$. The effect is that the two inputs, $\tilde{u}(t)$ and $\tilde{D}(t)$ can be lumped:

$$\boldsymbol{B}_{av}\tilde{u}(t) + \boldsymbol{B}_d\tilde{D}(t) = (\boldsymbol{B}_{av} + \alpha\boldsymbol{B}_d)\tilde{u}(t) \tag{D.5}$$

- When the thermal system natural frequency is much higher than the switching frequency, the temperature profile is a square wave on the time-scale of the switch period. The average temperature is linearly related to the duty ratio. If the fluid at temperature T_h is circulated for $\tilde{D}(t)T$ seconds and the fluid at T_l for $(1 - \tilde{D}(t))T$, then the averaged temperature over a cylce is:

$$\tilde{u}(t) = \tilde{D}(t)T_h + (1 - \tilde{D}(t))T_l = T_l + \tilde{D}(t)(T_h - T_l) \tag{D.6}$$

Again, a change in the duty ratio $\tilde{D}(t)$ is proportionally related to a change in the average temperature u :

$$\Delta\tilde{D}(t) = \frac{1}{T_h - T_l}\Delta\tilde{u}(t) = \alpha\Delta\tilde{u}(t) \tag{D.7}$$

and the two inputs can be lumped into one as above.

- If the system matrices describing the gel dynamics in either switch state are identical[1], then the analysis in Equations 7.6 - 7.11 suggests that the input matrix $\boldsymbol{B}_d = 0$, eliminating the duty ratio as an input to the system.

The small-signal LTI averaged model relating the inputs, averaged temperature $\tilde{u}(t)$ and force $(F_2(t))$ and the output gel length $(\tilde{y}(t))$ was formulated in Chapter 6 and identified in Chapter 7. By combining this model in series with the thermal averaged model derived above in Equation D.4, an overall averaged model can be constructed. This model can be used for control design as illustrated in Figure D.3. The variables shown are Laplace transforms of the small signal inputs, states, and outputs of the system. The gel model in the figure was derived in Chapter 6 and the thermal model is a block diagram representation of Equation D.4.

D.3 Gel Model Includes D_1 Mapping

If the gel model is a result of averaging the response of a discontinuous volume-phase transition gel, the transducer dynamics are not necessarily switched using the controller-specified or input switching function. A different switching function, $q_1(t)$, with duty ratio $D_1(t)$, was the output of a discontinuous equilibrium block and the input driving the transducer dynamics as shown in Figure D.4. This idea was introduced in Chapter 4 and analyzed further in Appendix C. In general, to effectively develop averaged models using the underlying dynamics of switched systems, this duty ratio mapping should be explicitly incorporated

[1]For example, if the system matrices $\boldsymbol{A}_1$ and $\boldsymbol{B}_1$ describe the shrinking volume change and the matrices $\boldsymbol{A}_2$ and $\boldsymbol{B}_2$ describe the shrinking volume change, then dynamics that are independent of the gel state (shrinking or expanding) translate to $\boldsymbol{A}_1 = \boldsymbol{A}_2$ and $\boldsymbol{B}_1 = \boldsymbol{B}_2$.

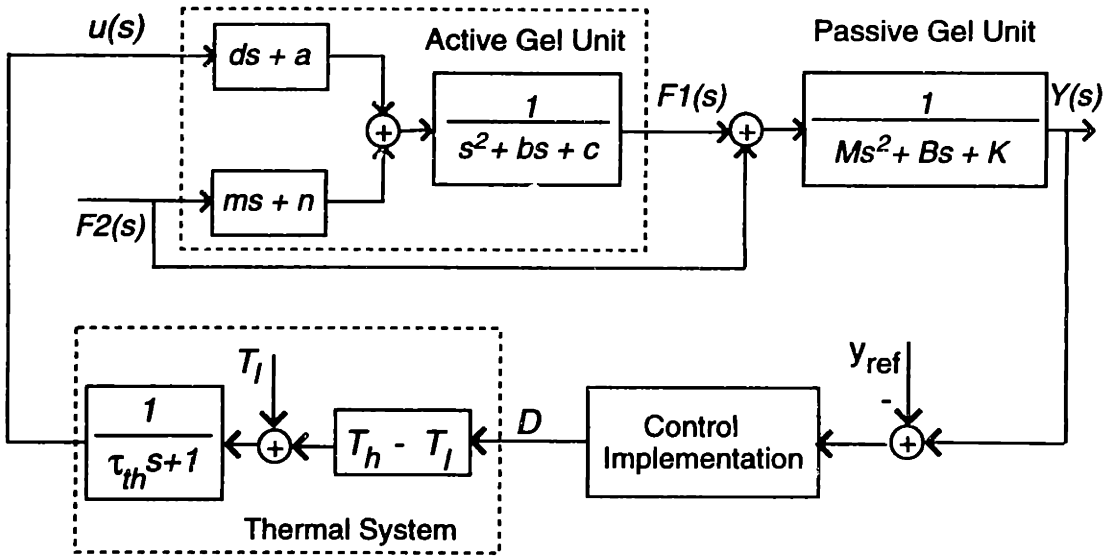


Figure D.3: Closed-loop control design for gel system.

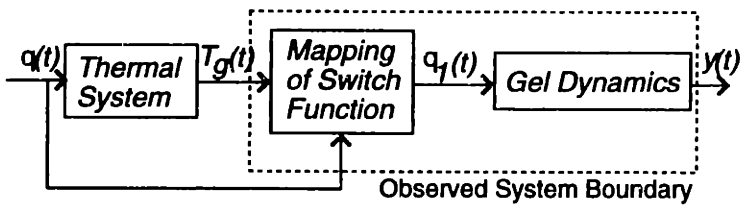


Figure D.4: Switch function mapping

into the averaged model. For example, if the exact transition temperature and hysteresis gap of a discontinuous gel under given loading conditions are known, this mapping can be derived with knowledge of the thermal system dynamics as demonstrated in Appendix C.

The technique used for deriving an averaged model for switched gel actuators in Chapters 5 - 7 identified the gel averaged model *directly* from the averaged waveforms. Averages of the inputs and outputs of the system outlined by the dashed lines in Figure D.4 can be calculated and used to identify a transfer function for the average dynamics. In general, as mentioned earlier, the identified average model will have both $\tilde{D}(t)$ (the small-signal average of $q(t)$) and $\tilde{u}$ (the small-signal average of T_g) as input, and the average position

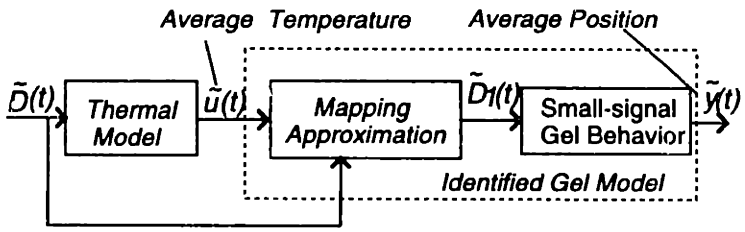


Figure D.5: Averaged model includes switch mappingas a function of D and u

as output. This model will include the effect of the switch function mapping. While in general this mapping is nonlinear, we may be able to identify dynamic models that are LTI. Under small-signal operation, the nonlinear mapping can be approximated by a linear function as was shown in Appendix C. An approximate, averaged, LTI model is illustrated in Figure D.5. The average temperature $\tilde{u}(t)$ and the duty ratio $\tilde{D}(t)$ are inputs to the identified, averaged gel model, which is the series connection of the mapping approximation and the averaged gel dynamics as a function of the duty ratio $D_1(t)$. The direct identification technique avoids the separate identification of the switch mapping function.

The gel model derived in this thesis did not have the duty ratio D as an input. As discussed in the previous section, there are several situations in which an averaged model may be derived with only the averaged temperature $\tilde{u}(t)$ as input. In this case, the direct effect of the duty ratio on the gel system (apart from its effect on the thermal system) may be lumped with the effect of the averaged temperature or completely ignored. In the case of a discontinuous gel where a mapping of the switch function exists, an approximation of the mapping function may be derived as a function of the averaged temperature $\tilde{u}(t)$ in the vicinity of an operating point duty ratio D_o[2]. As shown in Figure D.6, this mapping is also part of the identified gel model from Chapter 7. This LTI gel model can be used directly in a closed-loop configuration as described in the previous section to design controllers for the gel actuator, without explicitly inverting the mapping function. The mapping function approximation may appear as a gain in the gel model. The eventual control design will

[2]This is similar to linear approximations of $\tilde{D}_1(t)$ as a function of $\tilde{D}(t)$ alone in the vicinity of an operating point temperature as discussed in Appendix C.

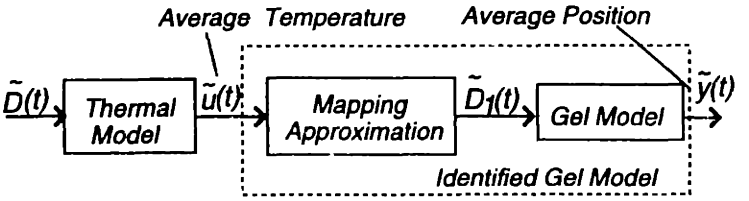


Figure D.6: Averaged model includes switch mapping as a function of u only

directly include the mapping function inversion discussed in Appendix C.

Appendix E

Measurement of Mechanical Properties of the Gel

Measuring the mass M and spring constant k of the polymer gel was carried out to verify the physical basis of the model for the PGU. For measuring k, the setup in Figrue 7.10 was used. Weights were added on the outside of the Leibig condenser across the pulley as shown in Figure 7.10 to provide an extension force. As the gel stretched, the rigid attachment at the top of the gel moved upward. The volume of the rigid cylinder moving out of the water (solvent) in the inner Leibig condenser then displaced less of the solvent. This resulted in a *negative* increment in force due to the added load. Adding a weight W on the pulley produced a change in applied force to the gel:

$$\Delta F = W - F_b, \tag{E.1}$$

where F_b is the bouyancy force applied by the solvent to the volume of the gel support that was moved higher than the level of the solvent. Since the radius of the cylindrical support is 4 mm = 0.4 cm, the volume change is approximately:

$$\Delta V_b = \pi(0.4)^2 y \, \text{cm}^3, \tag{E.2}$$

where y is the vertical displacement of the cylinder outside the solvent in centimeters. We approximate this displacement by the stretch in the gel. While the true displacement will be slightly higher (since the water level drops slightly as the cylinder moves out), this is

a good approximation. The cross-sectional area of the inner chamber of Leibig condenser is at least 4 times larger than the cross-sectional area of the cylinder attachment[1]. The resulting change in force, since the density of water is 1 g/cm^3, is therefore:

$$\Delta F = W - \pi(0.4)^2 y \, \mathrm{g}, \tag{E.3}$$

and the incremental spring constant k would then be:

$$k = \Delta F y. \tag{E.4}$$

The incremental load applied, $W = 0.28$g, produced a stretch of $y = 0.157$cm. The resulting spring constant is:

$$k = \frac{(0.28 - .0789)\mathrm{g}}{.157\mathrm{cm}} = \frac{(0.201)(9.8)(10^{-3})\mathrm{N}}{(1.57)(10^{-3})\mathrm{m}} = 1.25\mathrm{N/m} \tag{E.5}$$

For measuring the mass M of the gel structure, refer back to Figure 5.2. We ignore the mass of the water in the cylindrical attachment just below the gel. The holes at the top of this plastic fixture and the fact that the bottom is not sealed prevent water from being trapped and moving with the oscillations of the gel. However, the mass of the attachment itself is important and was measured to be $0.8g$ (including a small metal hook at its base). The mass of the dry gel network was estimated based on a dry sample of the same formula to be approximately 1 g. The thin plastic attachment at the bottom, along with the LVDT sensor pin and the ferromagnetic weight are approximately 3 g. Since this attachement is sealed and contains water, we need to account for the mass of the water as the gel moves. With the volume of the thin cylinder :

$$V_c = \pi(0.15)^2(7)\mathrm{cm}^3 \tag{E.6}$$

the mass of the water is approximately $0.495g$. The mass of the water solvent in the gel network should also be included in this measurement. Since the time scale of the mechanical response is much faster than the time scales of the diffusion process, this solvent is effectively

[1]This results in a 25% error in the bouyancy force, which is only about 25% of the applied force. The resulting error margin is within the limits of measurement error. Measurement of the actual displacement of the water level inside the Leibig condenser was difficult.

trapped inside the network and moves with it through the oscillations. The volume of the gel around the operating point is :

$$V_g = \pi(0.33)^2(13)\text{cm}^3 \tag{E.7}$$

which results in a mass of water approximately $4.45g$. Adding up all the contributions, the total moving mass is :

$$M = 9.745\text{g} = (9.745)(10)^{-3}\text{Kg} \tag{E.8}$$

Appendix F

Polymer Gel Formulas and Fabrication

The different formulas and gel fabrication techniques are repeated below with more details of the chemicals used and their sources.

F.1 Chemicals Used

The following references in the text of the chemical procedures will be used to indicate the chemicals used :

- NIPA
  - *N*-isopropylacrylamide - base monomer responsible for the thermosensitive behavior of the gel.
  - Formula shown in 2.1
  - Kodak Co.
- APS
  - Ammonium Persulfate - initiator
  - Formula : $(NH_4)_2S_2O_8$
  - Mallinkckrodt

- TEMED
  - *N,N,N'N'*-Tetramethyl-ethylenediamine - accelerator
  - Formula : $C_6H_{16}N_2Sm$ F.W. 116.21
  - EM Science, A divison of EM Industries, Inc., 480 S. Democrat Rd., Gibbstown, NJ 08027.
- BIS
  - *N,N'*-Methylene-bis-acrylamide - cross-linker
  - Formula shown in Figure 2.1
  - Bio-Rad Laboratories
- DAS
  - 4,4'-Diazido-Stilbene-2,2'-Disodium-Sulfonate - optical cross-linker
  - Polysciences, Inc., Warrington PA 18976-2590.
- PVA
  - Poly(vinyl alcohol), 100% hydrolyzed, M.W. 115,000
  - Formula : $[\text{-}CH_2CH(OH)\text{-}]_n$
  - Aldrich Chemical Company, Inc.
- Sodium acrylate Ionic group incorporated in NIPA copolymer to alter equilibrium behavior.
- Acrylic acid
  - Poly-(Acrylic acid) - M.W. 250,000, ionic group
  - Formula : $[\text{-}CH_2CH(COH)\text{-}]_n$
  - Aldrich Chemical Company, Inc.
- Nickel flakes

  - Microscopic nickel "leafing-grade" particles
  - Typical thickness - 0.4μm, typical diameter - 30μm
  - Novamet Specialty Products Corporation, 10 Lawlins Park, Wyckoff, NJ 07481.

- Ferrofluid solvent
  - Fluid with microscopic (order of .01μm in diameter) ferromagnetic particles suspended in some solvent by a surfactant.
  - Lignosite FML was used below
  - Georgia Pacific

F.2 Gel Cylinders in Micropipettes

1. Dissolve 115,000 molecular weight PVA (5 g) in 500mL of deionized water at 50°C for approximately 20 hours.

2. Dissolve APS (.2 g) in 5mL deionized water at room temperature. Degas the solution in a vacuum chamber.

3. Dissolve NIPA monomer (.78 g), and BIS (.013 g) in 5 ml of deionized water at room temperature. Mix this solution with 5 mL of the PVA solution from step 1, then degas the resulting 10 mL solution in a vacuum chamber. In this step, some of the NIPA monomer can be replaced with an equal amount of sodium acrylate. The result is a co-polymer with different volume transition properties. In our experiments, replacing .016g of NIPA with sodium acrylate raised the transition temperature to around 38°C, compared to approximately 38°C for pure NIPA.

4. Mix TEMED (24 μL) and 40 μL of the APS solution from step 2 with the solution from step 3. Shake the mixture to insure good mixing. Polymerization and gelation begin with this step.

5. Place the cylindrical micropipettes of different diameters in the resulting solution immediately after step 4. Typical diameters range from .5 mm to 1 mm.

F.3 Microscopic Gel Spheres using Inverse Suspension Polymerization

1. Dissolve APS (.2 g) in 5mL deionized water at room temperature. Degas the solution in a vacuum chamber.

2. Dissolve NIPA monomer (.78 g), and BIS (.013 g) in 10 mL of deionized water at room temperature, then degas the resulting solution in a vacuum chamber. As with the cylinderical gels, some of the NIPA monomer can be replaced with an equal amount of sodium acrylate.

3. Pour parrafin oil into the container to immerse the paddle as shown in the figure.

4. Pump nitrogen into the oil as shown in the figure at 10 psi for approximately 30 minutes. At the same time, stir the oil to insure that the nitrogen flows into the paraffin oil and replaces the oxygen, which flows out of the open air outlet.

5. Turn off the nitrogen supply and seal the air flow outlet. A slightly positive nitrogen flow can also be maintained if a perfect seal of the jar is not possible to insure that the oil mixture remains oxygen free.

6. Mix TEMED (24 μL) and 100 μL of the APS solution from step 1 with the solution from step 2. Shake the mixture to insure good mixing and load into a syringe. Polymerization and gelation begin with this step.

7. With the paddle continuing to stir, inject the polymer solution into the oil. Allow to stir for 2 hours. The shaft velocity can be used to control the size of the formed gel beads. In our experiments, shaft speeds close to 300 rpm produced beads with diameters between 100 μm and 300 μm.

8. Stop the stirring and allow the water and formed gel beads to settle to the bottom of the container. The oil is then drained and the beads rinsed of any excess oil.

F.4 Polymer Gel Fiber Spinning

1. Dissolve PVA (5 g) in 500mL of deionized water at 50°C for approximately 20 hours. The PVA used had molecular weight 115,000.

2. Dissolve APS (.2 g) in 5mL deionized water at room temperature.

3. Dissolve NIPA monomer (.78 g), in 7 mL of deionized water at room temperature. Mix this solution with 3 mL of the PVA solution from step 1 and 40 μL of the APS solution in step 2. Place the resulting solution in a vacuum chamber for approximately 1 hour. This will remove the oxygen which may inhibit polymerization. Allow 9 more hours for polymerization to complete.

4. Add DAS (.015 g) to the viscous polymer solution and vigorously shake on a Vortex Genie^{TM}. After approximately 4 minutes of continuous shaking, the DAS should be dissolved. Since DAS reacts in response to light, an opaque container should be used from this step onward. (Alternately, the glass vial can be wrapped in aluminum foil.)

5. The vigorous shaking of the viscous solution will trap air bubbles in the polymer/DAS solution. Any trapped oxygen will hinder the crosslinking process. In addition, the bubbles will make it difficult to spin smooth fibers. The thick polymer solution should be degassed in a vacuum chamber again. This will take a longer time because of the viscosity of the solution. We have witnessed that 10 hours with the given formula was sufficient for releasing trapped air from the solution.

6. The degassed solution can be loaded into a syringe and spun as shown above. This should be done in an oxygen-free environment.

7. Place the polymer fibers under a UV light source for crosslinking. A reflective surface should be placed around the fibers to produce spatially homogeneous crosslinking. In our experiment, the UV source [1] was applied for 1 hour to complete cross-linking.

[1]The UV source used was a Spectroline PE-240T EPROM erasing UV lamp manufactured by Spectronics Corp., Westbury, NY.

Fibers with diameters greater than 100 μm were not successfully crosslinked with this UV source. A higher power UV source is required for penetrating thicker polymer fibers.

F.5 Fabrication of Gel Cylinder Used in Prototype Gel Actuator

1. Dissolve 115,000 molecular weight PVA (5 g) in 500mL of deionized water at 50°C for approximately 20 hours.

2. Dissolve APS (.2 g) in 5 mL deionized water at room temperature. Degas the solution in a vacuum chamber.

3. Dissolve NIPA monomer (.78 g), and BIS (.0075 g) in 7 ml of deionized water at room temperature. Mix this solution with 3 mL of the PVA solution from step 1, then degas the resulting 10 mL solution in a vacuum chamber.

4. Mix TEMED (24 μL) and 40 μL of the APS solution from step 2 with the solution from step 3. Shake the mixture to insure good mixing. Polymerization and gelation begin with this step.

5. Place a glass tube with inner diameter 0.51 cm and length 10cm in the resulting solution immediately after step 4.

F.6 Nickel-Seeded Polymer Gel Cylinder

1. PVA (5 g) was dissolved in 500 ml of deionized water at 50 °C for approximately 20 hours.

2. APS (0.2 g) was dissolved in 5 ml deionized water at room temperature. This solution was degassed in a vacuum chamber.

3. Nickel flakes (0.5 g) were mixed in 5 ml of the PVA solution from step 1.

4. NIPA monomer (0.78 g), and BIS (0.013 g) were dissolved in 5 ml of deionized water at room temperature. This solution was mixed with the mixture of PVA and nickel flakes from step 3. The resulting 10 ml solution was degassed.

5. TEMED (24 μL) and 100 μL of the APS solution from step 2 were mixed with the solution from step 4. The new mixture was then vigorously shaken with a vortex mixer.

6. The solution in 5 was immediately poured into a glass cylinder 0.51 cm in diameter. It is important not to let the nickel flakes settle before transferring the solution.

7. The cylinder was sealed at both ends and mounted onto the shaft of a DC motor. The motor serves to spin the cylinder at approximately 100 rpm horizontally (with the glass cylinder parallel to the ground) to keep the nickel flakes suspended in the solution as gelation occurred. This is illustrated in Figure 8.3. When gelation was complete, the gel was extracted from the glass tube and immersed in deionized water.

F.7 Polymer Gel in Ferrofluid Solvent

1. PVA (5 g) was dissolved in 500 ml of deionized water at 50 °C for approximately 20 hours.

2. APS (0.2 g) was dissolved in 5 ml deionized water at room temperature. This solution was degassed in a vacuum chamber.

3. NIPA monomer (0.78 g), and BIS (0.013 g) were dissolved in 5 ml of deionized water at room temperature. This solution was mixed with the PVA solution from step 1. The resulting 10 ml solution was degassed.

4. TEMED (24 μL) and 40 μL of the APS solution from step 2 were mixed with the solution from step 3. The new mixture was then vigorously shaken with a vortex mixer.

5. The solution in 4 was immediately poured into a glass cylinder 0.51 cm in diameter. Gelation was completed in few hours.

6. After gelation was complete, the resulting cylinder of gel was immersed in ferrofluid.

Appendix G

Matlab Code

G.1 Data Filtering and Averaging

G.1.1 lp1.m

In order to minimize sensor noise in the measurements, a first order low pass filter was used. This filter was applied to both the input and output data as shown later.

```
function xfilt = lp1(x,step)
%xf = lp1(x, step)
%This function filters the input x to get rid of sensor noise.  'step' is
%the sampling period, or time between samples. xf is the filtered output

num = 1;

%Set filter bandwidth
tau = 10;
den = [tau 1];

%Convert to state-space form to be able to specify initial condition
[A,B,C,D] = tf2ss(num,den);
t = 0:step:length(x)-1;
[xfilt,st] = lsim(A,B,C,D,x,t,x(1)/C);
```

G.1.2 cave.m

The averaged model was identified based on averaged responses. These responses were frequently derived by averaging *and* sampling the averaged waveform once every period. The data before averaging was sampled at $1Hz$, with a switch period of 58 seconds. Frequently, it was possible to compute one average point every 58 seconds. With most of the profiles, the harmonic content was significantly lower in frequency than $f = \frac{1}{58}Hz$. As a result, this sampling frequency was sufficient, and made the averaging procedure computationally efficient. For profiles that are more rapidly varying, another function was used as shown in the next section.

```
function [ave,t] = cave(step, T, dum)
%[ave, t] = cave(step, T, indata)
%indata is an nx2 array with the first column the sensed temperature (voltage)
%and the second column the sensed position (voltage). This file assumes the
%sampling time is 'step' and produces a two column array 'ave' that is the averaged
%temperature and position signals over the period 'T'. The output is also sampled
% every 'T'. This file also produces a time vector 't' that is as long as the output
%data and starts at T/2, with spacing 'T' between consecutive time samples.


x = dum(:,2);
u = dum(:,1);

%filter the temperature and position data

xf = lp1(x,1);
uf = lp1(u,1);

%initialize 'pointer' to position in array

j1 = 0;

%initialize output vectors of position (xav) and temperature (uav)

xav = zeros([1 ((length(xf))/(T/step)-1)]);
uav = zeros([1 ((length(uf))/(T/step)-1)]);

for i = 1:round(length(xf)/(T/step) -1)
```

```
    t(i) = (i-1)*T + T/2;

    for j = (i-1)*T + step :step: T*i
        j1 = j1+1;
        xav(i) = xav(i) + xf(j1);
uav(i) = uav(i) + uf(j1);
    end
    xav(i) = xav(i)/(T/step);
    uav(i) = uav(i)/(T/step);
end

ave = [uav' xav'];
```

G.1.3 cavec.m

For data that exhibited significant harmonic content that can be lost or aliased if sampled with a period of T, the following function mimiced *cave.m* but used a moving window without sampling.

```
function [ave,t] = cavec(step, T, dum)
%[ave, t] = cave(step, T, indata)
%indata is an nx2 array with the first column the sensed temperature (voltage)
%and the second column the sensed position (voltage). This file assumes the
%sampling time is 'step' and produces a two column array 'ave' that is the averaged
%temperature and position signals over the period 'T'.  This file also produces a time
%vector 't' that is as long as the output data.

x = dum(:,2);
u = dum(:,1);

xf = lp1(x,1);
uf = lp1(u,1);

j1 = 0;
xav = zeros([1 length(uf)-T]);
uav = zeros([1 length(uf)-T]);
size(xav)
for i = 1:length(uf)-T
```

```
    t(i) = i-1;

    for j = i:step:i+T-1
        xav(i) = xav(i) + xf(j);
uav(i) = uav(i) + uf(j);
    end
    xav(i) = xav(i)/(T/step);
    uav(i) = uav(i)/(T/step);
end

ave = [uav' xav'];
```

G.1.4 convert.m

The sensors provided a voltage signal that represented the temperature and position of the signal. This function was used to convert the data array to the appropriate units, of degrees celcius and length in centimeters.

```
function degcm = convert(rdata)
%degcm = convert(rdata)
%The raw data rdata in the same nx2 array format is converted
% to a similar array with the numbers in the first column scaled to
%degrees and the ones in the second to cm

uav = rdata(:,1); %Temperature
uav = 10*uav+34;
xav = rdata(:,2); %Position
xav = 2.54*xav./7.77;

degcm = [uav xav];
```

G.2 Parameter Identification Routines

Once the data is conditioned and averaged, it is used by the identification and parameter estimation routines below to determine a good set of parameters for the system.

G.2.1 test1.m

Using the Matlab function *fmins*, the cost function below was minimized. This function simulates an output given a set of parameters and a measured input temperature. The squared error between the simulated vector and the true (measured) position vector is combined with a weighted error between the previously determined DC gain to generate a cost to be associated with these parameters. This function is called by *fmins* in the functions *fitav1.m* and *tallav.m* described later.

```
function cost = test1(param,pnew,out,ts)
%cost = test1(param,pnew,out,ts)
%This function generates a cost associated with a set of parameters 'param' using the
%measured position 'pnew', the temperature 'out' and the time base 'ts'

global count; %Counter for keeping track of fmins cycle on screen
count = count+1

step = ts(2)-ts(1); %Determine time step for simulation
n = 1;
yout(1) = pnew(1);
x(1) = pnew(1)/param(1); %Get value of initial state
xdot(1) = 0; %Assign initial derivative state to be zero

%%%System parameters
a1 = param(1);
b1 = param(2);
c1 = param(3);
d1 = param(4);
s11 = (-b1+(sqrt(b1^2 - 4*c1)))/2; %System Poles
s12 = (-b1-(sqrt(b1^2 - 4*c1)))/2;

%Simulation based on solution of second order system
for n = 2:length(ts)
            bu1 = out(n);
            Bg = (xdot(n-1)-s11*x(n-1)+s11*(bu1/c1))/(s12-s11);
            Ag = x(n-1)-Bg-bu1/c1;
```

```
            x(n) = Ag*exp(step*s11)+Bg*exp(step*s12)+bu1/c1;
            xdot(n) = s11*Ag*exp(step*s11)+s12*Bg*exp(step*s12);
            yout(n) = a1*x(n)+d1*xdot(n);
end;

x2 = yout;

%Error between simulation and experiment
cost = pnew(1:length(x2))'-x2;

%Use this to weigh different parts of waveform differently
%cost = [1*cost(1:88) 2*cost(89:length(cost))];

%Squared Error
cost = cost*cost';

%Bias estimate to previously measured DC gain
cost1 = (.2-a1/c1)^2;

%Total cost function
cost = cost + cost1*10;

%Weigh function against unstable systems
if (s11 > 0) cost = 10000000;end;
if (s12 > 0) cost = 10000000;end;

%Weigh function against unreasonable DC gains
%if ((a1/c1)>.2) cost = 10000000000;end;
%if ((a1/c1)<.03) cost = 10000000;end;

%Weigh function against oscillating, complex poles
pol = roots([1 b1 c1]);
pol1 = pol(1);
if (abs(imag(pol1))>0) cost = 10000000000;end;
```

G.2.2 fitav.m

```
%This file calls on 'fitav1' through the function 'fmins' to find a minimum
%for the cost function in 'fitav1' and a "good" set of parameters in a model
% based on ONE set of data -- one temperature and the corresponding position profile
```

```
clear;

global count;
count = 0;

%Offsets in raw data to set operating point at zero position and zero temperature

xoff = 1.1525;
toff = -.1817;

%load averaged data matrix

load newdat3\pi4a
xfull = ave(:,2);
ufull = ave(:,1);

%specify starting and ending points of data vectors to be used

c11 = 45;
c12 = 100;
c6 = c12-c11+1; %%length of data vectors
x = xfull(c11:c12) - xoff;
u = ufull(c11:c12) - toff;

x6 = x;u6 = u;
load datnew;

x = x22-.4;
u = u22;

%%specify initial values of parameter a,b,c,d in estimation using
%% time constants and DC gain

den = poly([-1/10 -1/1000]);
num1 = 20*den(3);
num2 = num1*1000;
initv = [num1 den(2) den(3) num2];

T = 58; %Switching period

ts = T/2:T:(length(u)-1)*T + T/2; %Time base for averaged and sampled waveform

%Option vector for fmins
options = [0,1e-8,1e-8,0, 0, 0, 0, 0, 0, 0, 0, 0, 2000, 0, 0, 0, 0];
[param,out2] = fmins('tav1new',initv,options,[],x,u,ts);
param
```

G.2.3 tallav.m

```
function cost = tavall(param,pnew,tmpnew,c)
%cost = tavall(param,pnew,tmpnew,c)
%This function calls on test1 multiple times to find a minimum
%cost set of parameters for more than one data set.
%'param' is the parameter set for the specific simulation and cost
%'pnew and 'tmpnew' are concatenated vectors of several experimental positions
% and temperatures, respectively, and 'c' is a vector that contains the
%lengths of the different sets of data

global count; %Counter to show fmins cycle on screen
count = count+1

%initialize costs
cost1 = 0;
cost = 0;

T = 58;
cc = 0; %Pointer to beginning of 'current' data set

for i = 1:length(c);
   pnew1 = pnew(cc+1:cc+c(i));
   u1 = tmpnew(cc+1:cc+c(i));
   t1 = T/2:T:(length(u1)-1)*T + T/2;
   cc = cc+c(i);
   cost1 = tav1new(param,pnew1,u1,t1);
   cost = cost + cost1;
end;
```

G.2.4 fallav.m

This function calls upon *tallav.m* using fmins to determine a minimum cost set of parameters for multiple data sets. This function concatenates the different position and temperature profiles into arrays and creates an array with the length of the different data sets to pass to *tallav.m*.

```
clear;

global count; %counter for keeping track of fmins cycles

T = 58;

%Offsets for setting equilibrium position and temperature at zero.
xoff = 1.1525;
toff = -.1817;

%Load data environment
load datnew;
load x24;
load u24;
count = 0;

%%Construct xall and uall as an array of different data sets
xall = [x2 x19]; %xi, ui, ci are in data file datnew
uall = [u2 u19];
c = [c2 c19];

%pick initial conditions using time constants and DC gain
den = poly([-1/200 -1/10000]);
num1 = 20*den(3);
num2 = num1*2000;
initv = [num1 den(2) den(3) num2];

%Can also initialize parameter values to previously determined set
%load p6_10_20
%load newdatp\ol1_2
%load paramn;
%param = paramn;
%initv = param;
options = [0,1e-8,1e-8,0, 0, 0, 0, 0, 0, 0, 0, 0, 1600, 0, 0, 0, 0];
[param,out2] = fmins('tallav',initv,options,[],xall,uall,c);
param
```

G.3 Simulation of Closed-Loop System

In order to test the identified model in designing closed-loop controllers, simulations of the total system, including the thermal input system, were needed. The following files were used to simulate the AGU (under constant load, the PGU is always in equilibrium) and the thermal system and combine them with a controller.

G.3.1 ctr.m

```
function [D,intout,err,err1] = control(pos,ref,gain,integ)
%[D,intout,err,err1] = control(pos,ref,gain,integ)
%Controller generating D.. pos is the position sensed
%gain is 1 number for proportional control, 2 for proportional/integral
%and the argument integ is value for integral content at current simulation
%step.  err and err1 are small-signal deviations in D and the bias, or
%nominal D in the controller.


%ks in microprocessor control is used to get higher resolution with fixed point arithmetic
%The resulting D is represented in Digital by a number between zero and 100000,
%represents the fraction of the 58 second period. In the simulation, this scale up and down
%has been ignored except where it affects some parameters in the control.

ks = 100000; %Digital gain in microprocessor for better fixed point accuracy
cinv = 20;
ce = 1400/ks; %in controller, ce is premult by ks
xv = pos;   %voltage sensed from LVDT
x = xv*1023/5;  %digital position, 1023/5 from AD

if (length(gain) == 1)
%If 1 gain element, assume proportional control
err = cinv*(ref - x)/gain;
%we divide by gain to mimic real controller
%which does that to avoid truncation errors
err1 = cinv*ce;
D = err+err1;
intout = 0;
else
%If 2 gain elements, assume proportional/accumulator control
k = gain(1);
hi = gain(2);
err = cinv*(ref - x)/k;
err1 = cinv*ce;
intout = integ+(ref-x);
```

```
interr = cinv*intout/hi;
D = err+err1+interr;
end;

if (D>.95) D=.95;end;
if (D<.05) D = .05;end;
```

G.3.2 tempsim.m

```
function [fullt,ave,t] = tempsim(temp0,D,taus,taue,rup,rdn,step,T)
%[full,ave] = tempsim(temp0,D,tau,rup,rdn,step,T)
%Returns a temperature profile (1 period) with average
%or just a point simulated given an initial temp
%and duty ratio.. Step is simulation step and
%taue, taus are RC time constants from temperature model
%(allows for simulation of time-varying thermal system)
%rdn and rup are the two temperature rails
%step is the simulation step
%T is the period of the switching
%D is the duty ratio, and temp0 is initial condition on temp
%outputs fullt: one period temp profile, ave :temp average, t:time vector


At1 = -1/taus;
Bt1 = 1/taus;
Ct1 = 1;
Dt1 = 0;

%%Allow for simulation of temperature system that has different dynamics
%%depending on switch state
At2 = -1/taue;
Bt2 = 1/taue;
Ct2 = 1;
Dt2 = 0;

ts = step:step:D*T; %time vector for heating
te = ts(length(ts))+step:step:T; %time vector for cooling
t = [ts te];

in1 = ones(size(ts))*rup;
[full1,x1] = lsim(At1,Bt1,Ct1,Dt1,in1,ts,temp0);
```

```
in2 = ones(size(te))*rdn;
temp0 = full1(length(full1));
[full2,x2] = lsim(At2,Bt2,Ct2,Dt2,in2,te,temp0);

fullt = [full1; full2];
ave = mean(fullt);
```

G.3.3 gelsim2.m

```
function [y,D1,yav,x,xdot] = gelsim(init,temp,ts,par1,par2)
%[y,D1,yav,x,xdot= gelsim(init,temp,ts,par1,par2)
%Simulates ONE POINT or ONE PERIOD gel position given the input temp
%ts is the time base of the simulation.  par1 and par2 are the
%system params for the shrinking and expanding, and T is period

a1 = par1(1);
b1 = par1(2);
c1 = par1(3);
d1 = par1(4);
s11 = (-b1+(sqrt(b1^2 - 4*c1)))/2;
s12 = (-b1-(sqrt(b1^2 - 4*c1)))/2;

a2 = par2(1);
b2 = par2(2);
c2 = par2(3);
d2 = par2(4);
s21 = (-b2+(sqrt(b2^2 - 4*c2)))/2;
s22 = (-b2-(sqrt(b2^2 - 4*c2)))/2;
dcount1 = 0;
dcount2 = 0;

x0 = init(1);
xd0 = init(2);
y0 = x0*a1+xd0*d1;

%x(1) = x0;
%xdot(1) = xd0;
%y(1) = y0;
if (length(ts) == 1) step = ts;
else step = ts(2)-ts(1);end;
```

```
for n = 1:length(ts)
gap = temp(n);

        if ((a1/c1)*gap > y0)
            dcount1 = dcount1+1;
            Bg = (xd0-s11*x0+s11*(gap/c1))/(s12-s11);
            Ag = x0-Bg-gap/c1;
            x(n) = Ag*exp(step*s11)+Bg*exp(step*s12)+gap/c1;
            xdot(n) = s11*Ag*exp(step*s11)+s12*Bg*exp(step*s12);
            y(n) = a1*x(n)+d1*xdot(n);
        else
            dcount2 = dcount2+1;
            Bg = (xd0-s21*x0+s21*(gap/c2))/(s22-s21);
            Ag = x0-Bg-gap/c2;
            x(n) = Ag*exp(step*s21)+Bg*exp(step*s22)+gap/c2;
            xdot(n) = s21*Ag*exp(step*s21)+s22*Bg*exp(step*s22);
            y(n) = a2*x(n)+d2*xdot(n);
        end;
x0 = x(n);
xd0 = xdot(n);
y0 = y(n);
end;

if (length(ts) == 1)
D1 = -1;
yav = y;
else
D1 = dcount1/(dcount1+dcount2);
yav = mean(y);
end;
```

G.3.4 simx.m

This file uses the different components of the closed-loop (and assumes an ideal switch in the system) to carry out a simulation of the total prototype.

```
%This file is for simulating total system and
%variations of that (ie averaging temp then feeding, etc.)
%By changing the gains on the controller and the thermal system
%parameters (according to observed thermal response at time
```

```
%of experiment), the gel model can be tested. Also using this
%software, response of the system with different controller gains
%can be predicted.

clear;

%%%%Load System Parameters
load ..\..\sec\paramn
param = paramn;

%Switching Period
T = 58;

y0 = 1.4-xoff;
ya(1) = y0;
x0 = y0/param(1);
xd0 = 0;

%%%Parameters

%%%%%%%Temp parameters and init conds
%%Initialize full temperature array
temptot = [];
%%Initialize average temperature array
tmpava = zeros(1,60);

%Set rdown and rup (units of volts before conversion to degrees
rdown = -.34-toff;
rup = .29-toff;
rsep = (rup-rdown)/2;

%%Thermal time constants
tautmp1 = 52;
tautmp2 = 58;
%%%%%%%%%%%%%%%%%%%%%%%%
y = x6(1:length(x6));

%%%%%controller
ref = 700; %%Reference to closed-loop
gain = [18000 35000];  %%Controller Gains
integ = 451; %%Initial value in integrator
%%%%%%%%%%%%%%%%%%%%%%%%

ts = T/2:T:50*T-T/2;
for i = 2:50;
refa(i) = ref;
[D,intout,err,err1] = ctr(y0+xoff,ref,gain,integ);
erra(i) = err;
```

```
err1a(i) = err1;
intouta(i) = intout;
integ = intout;
Da(i) = D;
[fullt,tmpav,t] = tempsim(u0,D,tautmp1,tautmp2,rup,rdown,.1,58);
u0 = fullt(length(fullt));
temptot = [temptot;fullt];
tmpava(i) = tmpav;
[yout,D1,yav,xout,xdout] = gelsim2([x0 xd0],tmpav,T,param,param);
x0 = xout;
xd0 = xdout;
D1a(i) = D1;
ya(i) = yout;
y0 = yout;
end;
```

G.4 Simulating the PGU

G.4.1 Verifying the PGU Model

In order to simulate the response of the PGU, the measured mechnical parameters, k and M, were used to build a second order mechanical model along with the damping coefficient that best reproduces the observed response. The simulated output was compared to the observed response using the following script.

```
load t335osc; %%Loads observed response
k = 1.25; %%Measured Spring constant
b = .053; %%damping that best fits response
Ms = .011; %%Mass that best fits response
M = 9.745/1000; %%Measured mass

%radius .33, height 13 - mass of water is 4.4476

%%%%System Matrices using measured mass
A2 = [0 1;-k/M -b/M];
B2 = -[0 0;1/M -1/M];
B2 = [0; 1/M];
C2 = [1 0];
```

```
D2 = 0;

%%Using best estimate for mass
A2s = [0 1;-k/Ms -b/Ms];
B2s = -[0 0;1/Ms -1/Ms];
B2s = [0; 1/Ms];
C2s = [1 0];
D2s = [0 0];
D2s = 0;
c = 1;
off = .6;
%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
t = .001:.001:2.861;

[y,x] = lsim(A2,B2,C2,D2,(k*.6)*ones(size(t)),t,[-.38 0]);
[ys,xs] = lsim(A2s,B2s,C2s,D2s,(k*.6)*ones(size(t)),t,[-.38 0]);
```

G.4.2 Simulations of Overall Model Response to Temperature and Force Inputs

The following file combines the two components of the model to illustrate the dominance of certain modes over different time-scales of response.

```
%%Load AGU Parameters
load paramn;
param = paramn;

a = param(1);
b = param(2);
c = param(3);
d = param(4);

%%State Space Description of AGU with Temperature input
A1 = [0 -c; 1 -b];
B1 = [a;d];
C1 = [0 1];
D1 = 0;

temp = 27.79;  %temp input for operating point, goes with init conds - no charge;
t0 = 20; %temp offset;
```

```
tempin = temp-t0;

%%For efficient simulation, first part of time base is finely sampled, rest has larger step
t1 = 0:.001:4;
t2 = 6:4:66000;
t = [t1 t2];

%%%%Define the PGU parameters
k = 1.25;
b = .053;
M = .011;

A2 = [0 1;-k/M -b/M];
B2 = [0 0;1/M -1/M];
C2 = [1 0];
D2 = [0 0];

%%%%Steps in force
f21 = .4*ones(size(t1));
f22 = .4*ones(size(t2));
f2 = .4*ones(size(t));
Adm = [0 0; 0 0];
Adm = [0 0; 0 0];
Amd = [0 0; 0 k/M]; %note that the force is multiplied by k!!

%%Total, 4th order ssytem
Atot = [A1 Adm;Amd A2];
Btot = [B1 -B1*5;[0 0]' [0 -1/M]'];
Ctot = [0 1 1 0];
Dtot = [0 0];

%%Use different initial conditions, and steps in temperature
%%and/or external force to observe output
%[ytot,xtot] = lsim(Atot,Btot,Ctot,Dtot,[tempin*ones(size(t))'
%%0*f2'],t,[0.002048   0.77155   0.77155  0] );
[y1tot,x1tot] = lsim(Atot,Btot,Ctot,Dtot,[tempin*ones(size(t1))'
f21'],t1,[0.00161062606923   0.60553071147685 (.38-.38+0.60553071147685)
 0.00000000010426] );
[y2tot,x2tot] = lsim(Atot,Btot,Ctot,Dtot,[tempin*ones(size(t2))'
f22'],t2,x1tot(length(x1tot),:));
ytot = [y1tot;y2tot];
xtot = [x1tot;x2tot];
```

Appendix H

Microcontroller Code

H.1 adsample.c

```
/*************************************************************/
/**This program generates samples off of the AD converter **/
/**and sends them to the output port. Top 8 bits of the   **/
/**digital representation of the sample can be observed on**/
/**LED's and are used for calibrating the sensor, etc.    **/
/**                                                       **/
/*************************************************************/

/*Reserve locations for use by the evaluation board*/
register char apple[9];
#pragma locate (apple = 0x30)
#pragma model(kc)

/*Declare interrupt routines*/
#pragma interrupt (analog_conversion_done = 1)

#include<80C196.h>                /*Microcontroller model*/

/*Register Variables*/
register unsigned int v1;        /*Used to hold lower byte of AD register*/
register unsigned int v2;        /*Used to hold bits 9,10 of AD register*/
register unsigned int vo;        /*holds 10 bit AD value*/
register unsigned char dum;      /*Samples AD register*/

void software_timer(void)
{
        ad_command = 8;
}
```

```
/*Interrupt Service Routine (ISR) entered when the AD conversion is done*/
void analog_conversion_done(void)
{
        /*Sample the 10 bits of the AD register */
        dum = ad_result_hi;
        v1 = (unsigned int) ad_result_hi;
        ioport1 = (unsigned char) dum;
        v1 = v1 << 2;
        v2 = (unsigned int) ad_result_lo;
        v2 = v2 >> 6;
        vo = v1 + v2;

        ad_command = 8;
}

main()
{
        /*Initialize output port*/
        ioport1 = 0;
        int_mask = 0x22;
        int_pending = 0;
        hso_command = 0x18;
        hso_time = timer1 + 10000;

        enable();
        while(1);
}
```

H.2 hyst.c

```
/*****************************************************************************/
/**This program implements Hysterisis closed-loop control.  By sampling***/
/*the position and comparing it to a band around the reference signal, it*/
/*decide whether to place the switch in the 'cold' or 'hot' position.  ***/
/**                                                                       ***/
/*****************************************************************************/

/*Reserve locations for use by the evaluation board*/
register char apple[9];
#pragma locate (apple = 0x30)
#pragma model(kc)

/*Declare interrupt routines*/
#pragma interrupt (software_timer = 5)
#pragma interrupt (analog_conversion_done = 1)

#include<80C196.h>                /*Microcontroller model*/

/*Register Variables*/
register unsigned int v1;         /*Used to hold lower byte of AD register*/
register unsigned int v2;         /*Used to hold bits 9,10 of AD register*/
register unsigned int vo;         /*holds 10 bit AD value*/
register long x;                  /*Here, simply vo. Can be used for diff Control*/
register unsigned char dum;       /*Samples AD register*/
register unsigned char state;     /*Switch State*/

void software_timer(void)
{
        /*Here the reference is 290 with +- 5 hyst band.*/
        if ((x < (unsigned int) 295) && (state ==  1))
                {
                ioport1 = (unsigned char) 1;
                state = 0;
        }
        else if ((x > (unsigned int) 305) && (state == 0))
                {
                ioport1 = (unsigned char) 2;
                state = 1;
        }

        ad_command = 8;
}

/*Interrupt Service Routine (ISR) entered when the AD conversion is done*/
```

```
/*Sample the 10 bits of the AD register */
void analog_conversion_done(void)
{
        /*Sample the 10 bits of the AD register */
        dum = ad_result_hi;
        v1 = (unsigned int) ad_result_hi;
        v1 = v1 << 2;
        v2 = (unsigned int) ad_result_lo;
        v2 = v2 >> 6;
        vo = v1 + v2;

        x = (long) vo;

        /*Start HSO timer for next cycle*/
        hso_command = 0x18;
        hso_time = timer1 + 10000;
}

main()
{

        /*Initialize output*/
        ioport1 = 2;
        /*Set interrupt mask .. Here, bits 2 and 6*/
        int_mask = 0x22;
        int_pending =1;
        hso_command = 0x18;
        hso_time = timer1 + 10000;
        /*Initial State -  Gel expanded and Switch on Cold*/
        state = 0;
        enable();
        while(1);
}
```

H.3 dconst.c

```
/*****************************************************************/
/**This program impresses a constant duty ratio square        **/
/** wave on the output port lines controlling the valve.  It **/
/** also steps this D value, for open-loop responses.         **/
/**                                                           **/
/*****************************************************************/

/*Reserve locations for use by the evaluation board*/
register char apple[9];
#pragma locate (apple = 0x30)
#pragma model(kc)

/*Declare interrupt routines*/
#pragma interrupt (software_timer = 5)
#pragma interrupt (analog_conversion_done = 1)

#include<80C196.h>                  /*Microcontroller model*/

/*Register Variables*/
register long refcnt;               /*Counts number of cycles elapsed*/
register unsigned char dum;         /*Samples AD register*/
register unsigned int v1;           /*Used to hold lower byte of AD register*/
register unsigned int v2;           /*Used to hold bits 9,10 of AD register*/
register unsigned int vo;           /*holds 10 bit AD value*/
register long p;                    /*period over ks.. define in main*/
register long count;                /*Counts timer ticks, indicates switch time*/
register long ton;                  /*DT in timer interrupts*/
register long toff;                 /*(1-D)T in timer interrupts*/
register long d;                    /*Duty ratio, 0 to 100000 (ks)*/
register long dbar;                 /*ks - d*/
register long ks;                   /*Digital Scale for fixed point*/
register long x;                    /*Sampled AD voltage*/

/*Interrupt Service Routine (ISR) entered when the Software Timer triggers*/
void software_timer(void)
{

        count = count + 1;
        if (count < ton)
        /*Cycle through with switch on HOT for DT*/
                {
                ioport1 = 1;
                hso_command = 0x18;
                /*Adding 570 ticks to timer 1, period approx 58 seconds*/
```

```
                hso_time = timer1 + 570;
        }
        else if (count < ton+toff)
        /*Cycle through with switch on COLD for (1-D)T*/
                {
                ioport1 = 2;
                hso_command = 0x18;
                hso_time = timer1 + 570;
        }

        else ad_command = 8;

}

/*Interrupt Service Routine (ISR) entered when the AD conversion is done*/
void analog_conversion_done(void)
{
        /*Sample the 10 bits of the AD register */
        dum = ad_result_hi;
        v1 = (unsigned int) ad_result_hi;
        ioport1 = (unsigned char) dum;
        v1 = v1 << 2;
        v2 = (unsigned int) ad_result_lo;
        v2 = v2 >> 6;
        vo = v1 + v2;

        x = (long) vo;

        refcnt = refcnt + 1;                    /*Increment Cycle counter*/
        if (refcnt < (long) 60)
        {
        d = 27000;
        }

        else if (refcnt < (long) 1200)
                {
                d = 37000;

        }
                else if (refcnt < (long)1600)
                {
                d = 23000;
        }
        else
                {
                d = 27000;
        }
```

```
dbar = ks - d;
ton = p * d;
toff = p * dbar;
count = 0;

/*Start HSO timer for next cycle*/
hso_command = 0x18;
hso_time = timer1 + 100;
}

main()
{
        ioport1 = 0;
        int_mask = 0x22;
        int_pending =1;
        refcnt = 0;
        ks = 100000;
        d = 40000;
        p = 100000/ks;
        ad_command = 8;

        enable();
        while(1);
}
```

H.4 dtraj1.c

```
/*************************************************************/
/**This program carries out proportional control but      **/
/**generates periodic triangular wave references.  This   **/
/**is used to test the gel response at diff frequencies.  **/
/**                                                       **/
/*************************************************************/

/*Reserve locations for use by the evaluation board*/
register char apple[9];
#pragma locate (apple = 0x30)
#pragma model(kc)

/*Declare interrupt routines*/
#pragma interrupt (software_timer = 5)
#pragma interrupt (analog_conversion_done = 1)

#include<80C196.h>                /*Microcontroller model*/

/*Register Variables*/
register unsigned char dum;        /*Samples AD register*/
register unsigned int v1;          /*Used to hold lower byte of AD register*/
register unsigned int v2;          /*Used to hold bits 9,10 of AD register*/
register unsigned int vo;          /*holds 10 bit AD value*/
register long p;                   /*period over ks.. define in main*/
register long count;               /*Counts timer ticks, indicates switch time*/
register long ton;                 /*DT in timer interrupts*/
register long toff;                /*(1-D)T in timer interrupts*/
register long d;                   /*Duty ratio, 0 to 100000 (ks)*/
register long dbar;                /*ks - d*/
register long err;                 /*Error due to small-signal deviation*/
register long err1;                /*Error used to bias gel*/
register long ks;                  /*Digital Scale for fixed point*/
register long k;                   /*Proportional gain (inverted)*/
register long cinv;                /*Normalizing gain for CL control*/
register long ce;                  /*Input to bias gel around operating point*/
register long ref;                 /*Controller reference*/
register long x;                   /*Sensed Position*/
register char state2;              /*State indicating if ramping up or down*/
register long count1;              /*counter for number of cycles not ramping*/
register long count2;              /*counter for number of ref increments*/

/*Interrupt Service Routine (ISR) entered when the Software Timer triggers*/
void software_timer(void)
{
```

```
/*********************************************************/
/*Create a periodic wavefrom by combining up-ramps,      */
/*down-ramps, and straight sections, using different,    */
/*states for each, and then periodically cycling through*/
/*By changing ramp increments and timer counts, the      */
/*period length can be changed for each ref waveform     */
/*                                                       */
/*********************************************************/

if ((ref > (long) 695) && (state2 == 0))
        {
        state2 = 1;
        count1 = 0;
}
if ((count1 == 1) && (state2 == 1)) state2 = 2;
if ((count1 == 1) && (state2 == 3)) state2 = 0;

if ((count2 > (char) 10) && (state2 == 0))
{
        ref = ref + 10;
        count2 = 0;
}
if ((count2 == (char) 2) && (state2 == 1))
        {

        ref = 700;
        count1 = count1 + 1;
        count2 = 0;
}
if ((count2 == (char) 2) && (state2 == 3))
        {

        ref = 300;
        count1 = count1 + 1;
        count2 = 0;
}
if ((count2 > (char) 10) &&  (state2 == 2))
        {
        ref = ref - 10;
        count2 = 0;
}

if ((ref < (long) 305) && (state2 == 2))
        {
        state2 = 3;
        count1 = 0;
}
count = count + 1;
```

```
        if (count < ton)
        /*Cycle through with switch on HOT for DT*/
                {
                ioport1 = 1;
                hso_command = 0x18;
                hso_time = timer1 + 570;
        }
        else if (count < ton+toff)
        /*Cycle through with switch on COLD for (1-D)T*/
                {
                ioport1 = 2;
                hso_command = 0x18;
                hso_time = timer1 + 570;
        }

        else
                {
                ad_command = 8;

                count2 = count2+1;
                /*count2 counts periods for stepping ref*/
        }

}

/*Interrupt Service Routine (ISR) entered when the AD conversion is done*/
void analog_conversion_done(void)
{
        /*Sample the 10 bits of the AD register */
        dum = ad_result_hi;
        v1 = (unsigned int) ad_result_hi;
        ioport1 = (unsigned char) dum;
        v1 = v1 << 2;
        v2 = (unsigned int) ad_result_lo;
        v2 = v2 >> 6;
        vo = v1 + v2;


        x = (long) vo;
        err = ref-x;
        err = ks * err;
        err = cinv * err;

        err = err/k;

        err1 = cinv * ce;
        d = err + err1;
```

```
        /*make sure d is between 0 and 1*/
        if (d < 5000) d = 5000;
        if (d > ks) d = ks;

        dbar = ks - d;
        ton = p * d;
        toff = p * dbar;
        count = 0;

        /*Start HSO timer for next cycle*/
        hso_command = 0x18;
        hso_time = timer1 + 1000;
}

main()
{
        ioport1 = 0;
        int_mask = 0x22;
        int_pending =1;
        /*hso_command = 0x18;
        hso_time = timer1 + 100;*/
        ks = 100000;
        count2 = 0;
        count1 = 0;
        state2 = 0;
        k = 25000;
        ref = 305;
        ce = 1400;
        /*note this is ced x10 since i made cinv cinvby 10*/
        /*in order to make sure there is no overflow in err*/
        cinv = 20;
        p = 100000/ks;
        ad_command = 8;

        enable();
        while(1);
}
```

H.5 dmult.c

```
/*****************************************************************/
/**This program applies Closed-Loop proportional control     **/
/**It also allows different references and steps to be pre-  **/
/**programmed to avoid interruption in operation.  Using     **/
/**the flag cloop, it can disable the controller in steady-  **/
/**state to observe the gel settled in open-loop.            **/
/**                                                          **/
/*****************************************************************/

/*Reserve locations for use by the evaluation board*/
register char apple[9];
#pragma locate (apple = 0x30)
#pragma model(kc)

/*Declare interrupt routines*/
#pragma interrupt (software_timer = 5)
#pragma interrupt (analog_conversion_done = 1)

#include<80C196.h>                /*Microcontroller model*/

/*Register Variables*/
register long dol;                /*Open-loop duty ratio.*/
register unsigned char cloop;     /*Flag indicating if controller on*/
register long refcnt;             /*Counts number of cycles elapsed*/
register unsigned char dum;       /*Samples AD register*/
register unsigned int v1;         /*Used to hold lower byte of AD register*/
register unsigned int v2;         /*Used to hold bits 9,10 of AD register*/
register unsigned int vo;         /*holds 10 bit AD value*/
register long p;                  /*period over ks.. define in main*/
register long count;              /*Counts timer ticks, indicates switch time*/
register long ton;                /*DT in timer interrupts*/
register long toff;               /*(1-D)T in timer interrupts*/
register long d;                  /*Duty ratio, 0 to 100000 (ks)*/
register long dbar;               /*ks - d*/
register long err;                /*Error due to small-signal deviation*/
register long err1;               /*Error used to bias gel*/
register long ks;                 /*Digital Scale for fixed point*/
register long k;                  /*Proportional gain (inverted)*/
register long cinv;               /*Normalizing gain for CL control*/
register long ce;                 /*Input to bias gel around operating point*/
register long ref;                /*Controller reference*/
register long x;                  /*Sensed Position*/

/*Interrupt Service Routine (ISR) entered when the Software Timer triggers*/
void software_timer(void)
```

```
{
        count = count + 1;
        if (count < ton)
        /*Cycle through with switch on HOT for DT*/
                {
                ioport1 = 1;
                hso_command = 0x18;
                hso_time = timer1 + 570;/*570;*/
        }
        else if (count < ton+toff)
        /*Cycle through with switch on COLD for (1-D)T*/
                {
                ioport1 = 2;
                hso_command = 0x18;
                hso_time = timer1 + 570;/*570 - 250*/
        }

        else ad_command = 8;

}

/*Interrupt Service Routine (ISR) entered when the AD conversion is done*/
void analog_conversion_done(void)
{
        /*Sample the 10 bits of the AD register */
        dum = ad_result_hi;
        v1 = (unsigned int) ad_result_hi;
        ioport1 = (unsigned char) dum;
        v1 = v1 << 2;
        v2 = (unsigned int) ad_result_lo;
        v2 = v2 >> 6;
        vo = v1 + v2;
        x = (long) vo;
        refcnt = refcnt + 1;                /*Increment cycle counter*/

        /*For different parts of the experiments, provide different (constant)*/
        /*references. Also, at end of each part, presumably when st. state */
        /*is reached, disable loop and let drift (optional)*/
        if (refcnt < (long) 20)
        {
        ref = 300;
        cloop = 1;
        k = 30000;
        }
        /*else if (refcnt < (long) 105)
                {
                dol = d;
                cloop = 0;
```

```
        }*/
        else if (refcnt < (long) 220)
                {
                ref = 700;
                cloop = 1;
                k = 30000;

        }
        /*else if (refcnt < (long) 215)
                {
                 cloop = 0;
                 dol = d;
         }*/
        else if (refcnt < (long) 420)
                {
                ref = 300;
                cloop = 1;
                k = 100000;

        }
        else if (refcnt < (long) 620)
                {
                ref = 700;
                cloop = 1;
                k = 50000;
        }
        else if (refcnt < (long) 820)
                {
                ref = 300;
                cloop = 1;
                k = 50000;
        }
        else
                {
                cloop = 0;
                dol = d;
        }

/*If we are in closed loop, carry out duty ratio command computation*/
if (cloop == (unsigned char) 1)
        {

        err = ref - x;
        err = ks * err;
        err = cinv * err;
        err = err/k;
        err1 = cinv * ce;
        d = err + err1;
```

```
        }
/*If in open loop, just use open loop D, which is last D used in cl. loop*/
        else d = dol;
        /*make sure d is between 0 and 1*/
        if (d < 0) d = 0;
        if (d > ks) d = ks;
        dbar = ks - d;
        ton = p * d;
        toff = p * dbar;
        count = 0;

        /*Start HSO timer for next cycle*/
        hso_command = 0x18;
        hso_time = timer1 + 100;
}

main()
{
        ioport1 = 0;
        int_mask = 0x22;
        int_pending =1;
        ks = 100000;
        k = 15000;
        ref = 480;
        cloop = (unsigned char) 1;
        ce = 1400;
        refcnt = 0;
        /*note this is ced x10 since i made cinv cinvby 10*/
        /*in order to make sure there is no overflow in err*/
        cinv = 20;
        p = 100000/ks;
        ad_command = 8;

        enable();
        while(1);
}
```

H.6 pidtynew.c

```
/*************************************************************/
/**This program applies closed-loop PI (PA) control        **/
/**It also allows different references and steps to be pre-**/
/**programmed to avoid interruption in operation.  Using   **/
/**the flag cloop, it can disable the controller in steady-**/
/**state to observe the gel settled in open-loop.          **/
/**                                                        **/
/*************************************************************/

/*Reserve locations for use by the evaluation board*/
register char apple[9];
#pragma locate (apple = 0x30)
#pragma model(kc)

/*Declare interrupt routines*/
#pragma interrupt (software_timer = 5)
#pragma interrupt (analog_conversion_done = 1)

#include<80C196.h>                /*Microcontroller model*/

/*Register Variables*/
register long dol;                /*Open-loop duty ratio.*/
register unsigned char cloop;     /*Flag indicating if controller on*/
register long refcnt;             /*Counts number of cycles elapsed*/
register unsigned char dum;       /*Samples AD register*/
register unsigned int v1;         /*Used to hold lower byte of AD register*/
register unsigned int v2;         /*Used to hold bits 9,10 of AD register*/
register unsigned int vo;         /*holds 10 bit AD value*/
register long p;                  /*period over ks.. define in main*/
register long count;              /*Counts timer ticks, indicates switch time*/
register long ton;                /*DT in timer interrupts*/
register long toff;               /*(1-D)T in timer interrupts*/
register long d;                  /*Duty ratio, 0 to 100000 (ks)*/
register long dbar;               /*ks - d*/
register long err;                /*Error due to small-signal deviation*/
register long err1;               /*Error used to bias gel*/
register long ks;                 /*Digital Scale for fixed point*/
register long k;                  /*Proportional gain (inverted)*/
register long cinv;               /*Normalizing gain for CL control*/
register long ce;                 /*Input to bias gel around operating point*/
register long ref;                /*Controller reference*/
register long x;                  /*Sensed Position*/
register long integ;              /*Accumulator Value*/
register long interr;             /*Command comp due to integration*/
```

```
/*Interrupt Service Routine (ISR) entered when the Software Timer triggers*/
void software_timer(void)
{
        count = count + 1;
        if (count < ton)
        /*Cycle through with switch on HOT for DT*/
                {
                ioport1 = 1;
                hso_command = 0x18;
                hso_time = timer1 + 570;/*570;*/
                }
        else if (count < ton+toff)
        /*Cycle through with switch on COLD for (1-D)T*/
                {
                ioport1 = 2;
                hso_command = 0x18;
                hso_time = timer1 + 570;
        }

        else ad_command = 8;

}

/*Interrupt Service Routine (ISR) entered when the AD conversion is done*/
void analog_conversion_done(void)
{
/*Sample the 10 bits of the AD register */
dum = ad_result_hi;
v1 = (unsigned int) ad_result_hi;
ioport1 = (unsigned char) dum;
v1 = v1 << 2;
v2 = (unsigned int) ad_result_lo;
v2 = v2 >> 6;
vo = v1 + v2;
x = (long) vo;
refcnt = refcnt + 1;                /*Increment cycle counter*/

/*For different parts of the experiments, provide different (constant)*/
/*references. Also, at end of each part, presumably when st. state */
/*is reached, disable loop and let drift (optional)*/
if (refcnt < (long) 50)
        {
        cloop = 1;
        k = 10000;
        hi = 30000;

        }
        else if (refcnt < (long) 250)
```

```
        {
        /*Can reset integrator after step*/
        if (refcnt == (long) 50) integ = 0;
        ref = 700;
        cloop = 1;
        k = 45000;
        hi = 84000;
}
/*else if (refcnt < (long) 333)
        {
        dol = d;
        cloop = 0;
}*/
else if (refcnt < (long) 450)
        {

        if (refcnt == (long) 250) integ = 0;
        ref = 300;
        cloop = 1;
        k = 80000;
        hi = 154000;
}
/*else if (refcnt < (long) 215)
        {
         cloop = 0;
         dol = d;
 }*/
else if (refcnt < (long) 650)
        {

        if (refcnt == (long) 450) integ = 0;
        ref = 700;
        cloop = 1;
        k = 1000000;
        hi = 100000;
}
else if (refcnt < (long) 800)
        {

        if (refcnt == (long) 650) integ = 0;
        ref = 300;
        cloop = 1;
        k = 1000000;
        hi = 100000;
}
else
        {
        cloop = 0;
```

```
                dol = d;
        }

/*If we are in closed loop, carry out duty ratio command computation*/
if (cloop == (unsigned char) 1)
        {
        err = ref - x;
        integ = integ+err;
        err = ks * err;
        err = cinv * err;
        interr = ks * integ;
        interr = cinv * interr;
        interr = interr / hi;
        err1 = cinv * ce;
        d = err + err1;
        d = d + interr;
        }
/*If in open loop, just use open loop D, which is last D used in cl. loop*/
        else d = dol;
        /*make sure d is between 0 and 1*/
        if (d < (long) 5000) d = 5000;
        if (d > (long) 95000) d = 95000;
        dbar = ks - d;
        ton = p * d;
        toff = p * dbar;
        count = 0;

        /*Start HSO timer for next cycle*/
        hso_command = 0x18;
        hso_time = timer1 + 100;
}

main()
{

        ioport1 = 0;
        int_mask = 0x22;
        int_pending =1;
        ks = 100000;
        k = 60000;
        interr = 0;
        integ = 0;
        hi = 100000;
        ref = 350;
        cloop = (unsigned char) 1;
        ce = 1400;
        refcnt = 0;
        /*note this is ced x10 since i made cinv cinvby 10*/
```

```
        /*in order to make sure there is no overflow in err*/
        cinv = 20;
        p = 100000/ks;
        ad_command = 8;

        enable();
        while(1);
}
```

Appendix I

Signal Conditioning for Data-Acquisition

As mentioned in Chapter 5, the temperature and position signals were recorded and used for system identification and model verification. The volume of the gel around the operating point is very sensitive to changes in the average temperature, since the gel is operated near its transition point. One of the instrumentation challenges was to be able to resolve the changes in average temperature while recording the full range of temperature ripple. In addition, to obtain the best results in the identification process, the position signal had to exhibit fine quantization and low noise levels. In order to ensure the accuracy and precision required for these measurements, the signal conditioning circuits in Figure I.1 were used.

The op-amp circuit is used to adjust the measured position signal so that the gel length change of interest is mapped to the full dynamic range of the A/D converter. The LVDT signal is offset and scaled by the circuit. In this particular case, the LVDT output (for the range of motion of interest) ranged from -2.8V to 0V. The circuitry shown mapped that range to the range from 0V to approximately 5V, the range of the A/D converter on the 80C196 microcontroller. An RC filter is used at the output of this mapping circuit to provide a clean signal to the data acquisition board and the microcontroller. The Omega temperature transmitter (TX91-K2) provides an output current of 4mA to 20 mA for the temperature from 0°F to 200°F. This current is converted to a voltage using the 1.1kΩ resistor at the input to analog input channel 1 on the Genie card. At the operating

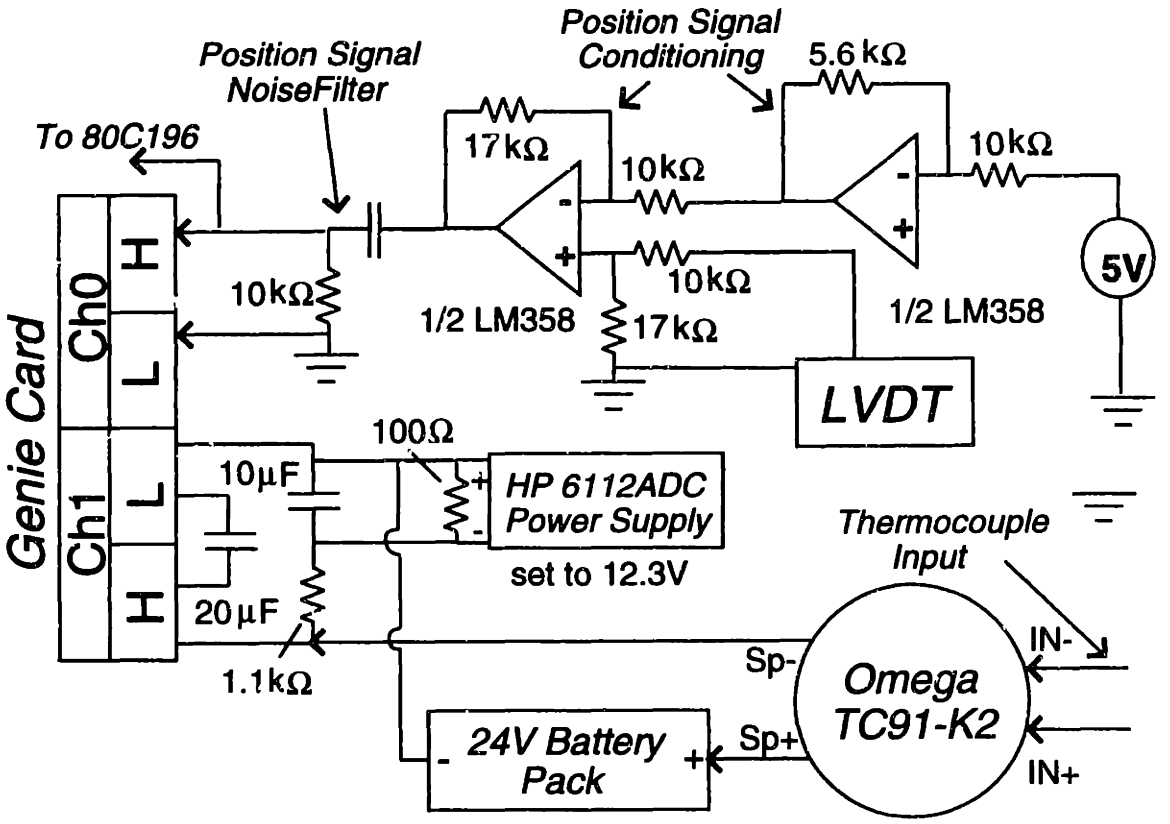


Figure I.1: Output switching function q_1 as a result of input switch q and the thermal dynamics.

point of the gel actuator, around 34°C, this voltage is approximately 11V. The HP power supply provides an offset that centers this voltage around 0V. The Genie card input is maximally rated at 5 volts. The Genie software was then adjusted to measure voltages in the range from -0.625V to 0.625V on channel 1 to maximize the resolution for the temperature measurement. The input to channel 1 is also filtered as shown.

Bibliography

[1] K. Akazawa, H. Kusumoto, "Cybernetic Actuators - Muscle Mechanics and Hardware Model," International Biomedical Engineering Days, pp. 48-57, 1992.

[2] H. Asada and J.E. Slotine, "*Robot Analysis and Control*", John Wiley & Sons, 1986.

[3] K. Asano *et al.*, "Multijoint inspection robot," *IEEE Trans. Ind. Electron.*, vol.IE-30, no.3, pp. 277-281, 1983.

[4] K. Astrom, B. Wittenmark, *Computer Controlled Systems*, Englewood Cliffs, New Jersey: Prentice-Hall, 1984.

[5] J. Baillieul, "Kinematically redundant robots with flexible components," *IEEE Control Systems*, vol. 13 # 1, Feb. 93.

[6] B. Black, M. Lopez, A. Morcos, "Basics of Voice Coil Actuators," *PCIM*, July 93.

[7] K. Bosecker, "Linear motors provide smooth, precise positioning," *PCIM*, Aug. 92.

[8] D.L. Brock, "Dynamic Model and Control of an Artificial Muscle based on Contractile Polymers," Aritificial Intelligence Laboratory memo. No. 1331, Nov. 1991.

[9] K. G. Budinski, "*Engineering Materials - Properties and Selection*," Prentice Hall, 1989.

[10] V.A. Caliskan, G.C. Verghese, A.M. Stankovic, "Multi-Frequency Averaging of DC-Dc Converters," *IEEE Workshop on Computers in Power Electronics*, pp. 113-119, 1996.

[11] S. Candau, C. Young, T. Tanaka, P. Lemarechal, J. Bastide, "Intensity of Light Scattered from Polymeric Gels : Influence of the Structure of the Networks," *Journal of Chemical Physics*, vol. 70, no. 10, pp. 4694-4699, 1979.

[12] P. Carter, F. Naghdy, N. Laszlo, C.D. Cook, P. Wong, "Study of Actuator Technologies for a Miniature Distributed Manipulation Environment," *IEEE International Workshop on Emerging Technologies and Factory Automation - Technology for the Intelligent Factory*, Proc. pp. 661-665, 1992.

[13] K. Choe, H. Baruh, "Actuator Placement in Structural Control,"Journal of Guidance, Control, and Dynamics, vol. 15, no. 1, pp. 40-48, 1994.

[14] E. Cravalho, J. Smith, Jr., "*Engineering Thermodynamics*," Pitman Publishing, 1981.

[15] C. W. deSilva, "*Control Sensors and Actuators*," Prentice Hall, 1989.

[16] M. Dunnigan,"Sliding Mode Control Approach to Reduce the Influence of Manipulator Disturbances on an Underwater Vehicle's Position", *Transactions of Institute of Measurement and Control*, vol. 18, no. 3, pp. 125-34, 1995.

[17] K. Dusek, "*Responsive Gels*," Springer-Verlag Berlin Heidelberg 1993.

[18] D. Fanning,"Light Induced Phase Transitions of Gels," S.B. Thesis, MIT, May 1993.

[19] G. Fernandez and S.B. Leeb, "Preparation of N-isopropylacrylamide Bead Gels of Submillimeter Diameter," Internal document - research project report.

[20] P. J. Flory,*Principles of Polymer Chemistry.*, Ithaca, NY, Cornell University Press, 1953.

[21] H. Fujiwara, M. Shibayama, J. Chen, S. Nomura, "Preparation of High-Strength Poly(vinyl Alcohol) Fibers by Crosslinking Wet Spinning," *Journal of Applied Polymer Science*, vol. 37, pp. 1403-1414, 1989.

[22] E. Geissler, F. Horkay, A. Hecht, M. Zrinyi, "Elastic Free Energy in Swollen Polymer Networks," *Journal of Chemical Physics*, vol. 90, no. 3, Feb. 1989.

[23] J.P. Gong, T. Nitta, Y. Osada, "electrokinetic Modeling of the Contraction of Polyelectrolyte Gels," *Proceedings of the Second International Conference on Intelligent Materials, ICIM* '94pp. 556-564, 1994.

[24] S. Haider, T. Cetas, J. Wait, J. Chen, "Power Absorption in Ferromagnetic Implants from Radio Frequency Magnetic Fields and the Problem of Optimization", *IEEE Trans. on Microwave Theory and Techniques*, 39(11):1817–1827, 1991.

[25] J. Hasa, M. Ilavsky, K. Dusek, "Deformational, Swelling, and Potentiometric Behavior of Ionized poly(methacrylic acid) gels. I. Theory," *Journal of Polymer Science, Polymer Physics Edition*, vol. 13, no.2, pp. 253-262, 1975.

[26] H. Haus, J. Melcher, *Electromagnetic Fields and Energy.* Englewood Cliffs, New Jersey: Prentice-Hall, 1989.

[27] A. Hemami,"Studies on a Lightweight and Flexible Robot Manipulator", *Robotica*, 1:27–36, 1985.

[28] "A high-resolution rotarty-to-linear motion converter," NASA Tech Briefs, Nov. 92.

[29] S. Hirose, T. Kado, Y. Umetani, "Tenser Actuated Elastic Manipulator", *Proceedings 6th IFToMM World Congress*, 2:978–981, 1983.

[30] S. Hirose, S. Ma, "Moray Drive for Multijoint Manipulator", *Fifth International Conference on Advanced Robotics*, 1:521–526, 1991.

[31] S. Hirotsu, Y. Hirokawa, T. Tanaka, "Volume-Phase Transitions of Ionized N-isopropylacrylamide Gels", *Journal of Chemistry and Physics*, 87(2):1392–1395, 1987.

[32] S. Hirotsu, A. Onuki,"Volume Phase Transition of Gels under Uniaxial Tension," *Journal of the Physics Society of Japan*, vol. 58, no. 5, pp. 1508-1511, 1989.

[33] F. Horkay, M. Zrinyi, "Mechano-chemical Energy Conversion of Neutral Polymer Gels," *Makromol. Chem., Macromol. Symp.* 30, pp. 133-143, 1989.

[34] F. Horkay, M. Zrinyi, "Studies on Mechanical and Swelling Behavior of Polymer Networks on the Basis of the Scaling Concept. 6. Gels Immersed in Polymer Solutions," *Journal of Macromolecular Science*, PHYSICS, B 25(3), pp. 307-334, 1986.

[35] I.W. Hunter, S. Lafontaine, P.M.F. Nielsen, P.J. Hunter, J. Hollerbach, "Manipulation and Dynamic Mechanical Testing of Microscopic Objects Using a Tele-Micro-Robot System," *IEEE Control Systems Magazine*, vol. 10 no. 2, pp. 3-8, Feb. 90.

[36] F. Ilmain, T. Tanaka, E. Kukufuta,"Volume Transition in a Gel Driven by Hydrogen Bonding," *Nature*, vol. 349, no. 6308, pp. 400-401, Jan. 1991.

[37] Intel Corp., *80C196KB User's Guide*, Oct. 1990.

[38] Intel Corp., *EV80C196KB Microcontroller Evaluation Board User's Manual*, Release 001, Feb. 20 1989.

[39] S.C. Jacobsen *et al.*, "Design of the Utah/MIT dexterous hand," in *Proc. 1986 IEEE Int. Conf. Robotics and Automation*, vol. 3, pp. 21-50, 1986.

[40] D. Jackson, "Inductively-Coupled Power Transfer for Electromechanical Systems," *Ph.D. Thesis*, MIT, June 1998.

[41] D. Jackson, S. Leeb, A. Mitwalli, D. Fusco, C. Wang, T. Tanaka, "A Sensor for Measuring Gel Phase-Transition Temperature, with Potential as a Metal Ion Detector," *Journal of Intelligent Material Systems and Structures*, vol. 8, no. 2, pp. 184-190, 1997.

[42] D. Jackson, S. Leeb, A. Mitwalli, P. Narvaez, D. Fusco, and E.C. Lupton, "Power Electronic Drives for Magnetically Triggered Phase Transition Gels", *IEEE Power Electronics Specialists Conference*, June 23-27, Baveno, Italy, pp. 302–309, 1996.

[43] D. Jackson, S. Leeb, A. Mitwalli, P. Narvaez, D. Fusco, and E.C. Lupton, "Power Electronic Drives for Magnetically Triggered Phase Transition Gels", *IEEE Transactions on Industrial Electronics*, 44(2):217–225, 1997.

[44] B. Jang, Z. Zhang, "Thermally- and Phase Transformation-Induced Volume Changes of Polymers for Actuator applications," *Proceedings of the Second International Conference on Intelligent Materials, ICIM* '94pp., pp. 654-663, 1994.

[45] J.G. Kassakian, M.F. Schlect, G.C. Verghese, "*Principles of Power Electronics,*" Addison-Wesley, 1991.

[46] B. Kabra, S. Gehrke, "Synthesis of Fast Response, Temperature-Sensitive poly(N-isopropylacrylamide) gel," *Polymer Communications*, vol. 32, no. 11, p. 322, 1991.

[47] R. Kishi, H. Ichijo, and O. Hirasa, "Thermo-Responsive Devices Using Poly(vinyl methyl ether) Hydrogels," *Journal of Intelligent Material Systems and Structures*, Vol. 4, pp. 533-537, October 1993.

[48] S. Krandall, D. Karnopp, E. Kurtz, D. Pridmore-Brown, "*Dynamics of Mechanical and Electromechanical Systems,*" Robert E. Krieger, 1968.

[49] P. Krein, J. Bentsman, R. Bass, B. Lesieutre, "On the Use of Averaging for the Analysis of Power Electronic systems," *IEEE Transactions on Power Electronics*, vol. 5, pp. 182-190, 1990.

[50] R. Ledesma, S. Devasia, E. Bayo, "Inverse Dynamics of Spatial Open-Chain Flexible Manipulators with Lumped and Distributed Actuators," *Journal of Robotic Systems*, vol. 11, no. 4, pp 327-338, 1993.

[51] W. Lee, "Polymer Gel Based Actuator: Dynamic Model of Gel for Real Time Control," *Ph.D. Thesis*, MIT 1996.

[52] W. Lee, D. Brock, "Lumped Parameter Modeling of Polymer Gel for Real Time Control," *Proceedings SPIE - International Society for Optical Engineering*, vol. 2779, pp.353-8, 1996.

[53] S. Leeb *et al.* "Electromagnetically Triggered, Responsive Gel Based Drug Delivery Device", U.S. Patent #5,643,246, July 1, 1997.

[54] S.B. Leeb, "Recognition of Dynamic Patterns in High Frequency DC-DC Switching Converters," *S.M. Thesis*, MIT, Feb. 1989.

[55] S. Leeb, J. Kirtley, G. Verghese, "Recognition of Dynamic Patterns in DC-DC Switching Converters," *IEEE Transactions on Power Electronics*, vol. 6, no. 2, 1991.

[56] S.B. Leeb, R. Surydevara, "Digitally Controlled Thermoelectric Temperature Regulator," Report for Advanced Mechatronics Project Laboratory, MIT, 1995.

[57] W.S. Levine and G.E. Loeb, "The neural control of limb movement," *EEE Control Systems*, vol. 9 # 6, Dec. 92.

[58] Y. Li, T. Tanaka, "Kinetics of Swelling and Shrinking of Gels," *Journal of Chemical Physics*, vol. 92, no. 2, p. 1365-71, 1990.

[59] Y. Li, T. Tanaka, "Phase Transitions of Gels", *Annual Review of Material Science*, 22:243–277, 1992.

[60] Y. Li, T. Tanaka, "Study of the Universality Class of the Gel Network System," *The Journal of Chemical Physics*, vol. 90, on. 9, 1989.

[61] Linear Technology. *LTC1043 Data Sheet*, Linear Databook, pp. 11-15 – 11-30, 1990.

[62] L. Ljung, S. Soderstrom, *Theory and Practice of Recursive Identification*, Massachusetts Institute of Technology, 1983.

[63] S. Ma, S. Hirose, and H. Yoshinada, "Design and experiments for a coupled tendon-driven manipulator," *IEEE Control Systems*, vol. 13 # 1, Feb. 93.

[64] D. Madill, W. Wang, "The Modelling and L_2-Stability of a Shape Memory Alloy Position control System," *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, San Diego, CA, USA, May 8-13, 1994.

[65] "Magnetostrictive roller-drive stepping motor," NASA Tech Briefs, May 93.

[66] A. Mamada, T. Tanaka, D. Kungwatchakun, M. Irie,"Photoinduced Phase Transition of Gels," *Macromolecules*, vol. 23, pp. 1517-1519.

[67] Mathworks. *MATLAB Student Version 5.0*, Englewood Cliffs, New Jersey: Prentice-Hall, 1997.

[68] E. Matsuo, T. Tanaka, "Collective Diffusion Coefficients of N-Isopropylacrylamide Gels Near Phase Transition," *Phase Transitions*, vol. 46, pp. 217-228, 1994.

[69] E. Matsuo, T. Tanaka, "Kinetics of Discontinuous Volume-phase Transition of Gels," *Journal of Chemical Physics*, vol. 89, no. 3, pp. 1695-1703, 1 Aug. 88.

[70] R. Middlebrook, S. Cuk, "A General Unified Approach to Modeling Switching Converter Power Stages," *Proceedings of IEEE Power Electronics Specialists Conference (PESC)*, pp. 18-34, 1976.

[71] A.H. Mitwalli, T.A. Denison, D.K. Jackson, S.B. Leeb, T. Tanaka, "Closed-Loop Feedback Control of Magnetically-Activated Gels," *Journal of Intelligent Material Systems and Structures*,vol. 8, pp. 596-604, 1997.

[72] A. Mitwalli, S. Leeb, T. Tanaka, U. Sinha, "Polymer Gel Actuators - Status Report", *Universities Power Engineering Conference*, Galway, Ireland, pp. 871–875, 1994.

[73] T. Muto, H. Yamada, Y. Suematsu, "PWM-Digital Control of a Hydraulic Actuator Utilizing Two-Way Solenoid Valves'" *The Journal of Fluid Control*, vol. 20, no. 2, pp. 24-41, 1990.

[74] P. Narvaez, "A Sensor for Measuring Gel Phase-Transition Temperature" *S.B. Thesis*, MIT, 1996.

[75] K. Neuffer, W. Brehm,"Fast Switching PWM-Solenoid for Automatic Transmissions," *Proceedings of the Institution of Mechanical Engineers Seventh International Conference. Automotive Electronics*, pp. 99-104, London, UK, Oct 9-13, 1989.

[76] T. Noritsugo, "Development of PWM Mode Electro-Neumatic Servomechanism, Part I:Speed Control of a Pneumatic Cylinder," *The Journal of Fluid Control*, vol. 17, no. 1, pp. 65-80, 1987.

[77] M. Obama *et al.*, "A locomotive inspection of robot for turbine building interior inspection in nuclear power plants," in *Proc. 1985 Int. Conf. Advanced Robotics*, pp. 355-362, 1985.

[78] C. Odero, "Polymer Gel Spinning Machine," *S.M. Thesis*, MIT, 1994.

[79] I. Ohmine, T. Tanaka,"Salt Effects on the Phase Transition of Ionic Gels," *Journal of Chemical Physics*, vol. 77, on. 11, 1982.

[80] Omega Technologies, The Temperature Handbook, vol. 28, 1992.

[81] T. Osada and M. Hasebe, "Electrically activated mechanochemical devices using polyelectrolyte gels," *Chemistry Letters*, 1985.

[82] Y. Osada and S.B. Ross-Murphy, "Intelligent Gels," *Scientific American*, May 93.

[83] M. Rahman, N. Cheung, K. Lim,"Position Estimation in Solenoid Actuators," *IEEE Transactions on Industry Applications*, vol. 32, r.o. 3, 1996.

[84] H. Robertshaw, "Control Approaches for Intelligent Material Systems-What Can We Learn from Nature," *Proceedings of the Second International Conference on Intelligent Materials, ICIM* '94pp., pp. 63-69, 1994.

[85] "*Robot Control: Dynamics, Motion Planning, and Analysis,*" Book Review, IEEE Control Systems Magazine, vol. 13, no. 1, Feb. 93.

[86] C.A. Rogers, "Intelligent Material Systems - The Dawn of a New Materials Age," *Journal of Intelligent Material Systems and Structures*, vol. 4, Jan. 1993, pp. 4 - 12.

[87] R. Roth, "Elastic wave motor produces linear motion with precise, high force," *PCIM*, Aug. 92.

[88] M. Shahinpoor, G. Wang, M. Mojjarad, "Electro-Thermo-Mechanics of Spring-Loaded Contractile Fiber Bundles with Applications to Ionic Polymeric Gel and SMA Actuators," *Proceedings of the Second International Conference on Intelligent Materials, ICIM* '94pp., pp. 1105-1116, 1994

[89] I. Shimoyama, H. Miura, K. Suzuki, Y. Ezura, "Insect-like Microrobots with External Skeletons," *IEEE Control Systems Magazine*, vol. 13 no. 1, pp. 37-41, Feb 93.

[90] M. Solari, "Evaluation of the Mechanical Properties of a Hydrogel Fiber in the Development of a Polymeric Actuator," *Journal of Intelligent Material Systems and Structures*, Vol. 5, pp.295-304, May 1994.

[91] J.J. Slotine, W. Li, "*Applied Nonlinear Control,*" Prentice Hall, 1991.

[92] S.G. Starodubtsev, A.R. Khokhlov, V.V. Vasilevskaya, "Collapse of Polyacrylamide Gels: Effect of Mechanical Deformation of the Sample and the Type of Solvent," *Akademia Nauk SSSR, Proceedings of Physical Chemistry Section*, 280-282, pp. 484-486, 1985.

[93] I.Z. Steinberg, A. Oplatka, A. Katchalsky, "Mechanochemical Engines," *Nature*, vol. 210, pp. 568-571, May 7, 1966.

[94] M.V. Sussman, A. Katchalsky, "Mechanochemical Turbine: A New Power Cycle," *Science*, vol. 1, pp 45-47, 1969.

[95] A. Suzuki, T. Tanaka, "Phase Transition in Polymer Gels Induced by Visible Light", *Nature*, 346(6282):345–347, 1990.

[96] K. Suzumori, S. Likura, H. Tanaka, "Applying a flexible microactuator to robotic Mechanisms," *IEEE Control Systems*, Feb. 92.

[97] T. Tanaka, "Gels," *Scientific American*, vol. 244, no. 1, Jan. 81.

[98] T. Tanaka, "Kinetics of Phase Transition in Polymer Gels," *Physica* 140A, pp 261-268, 1986.

[99] T. Tanaka, D. Filmore, "Kinetics of Swelling Gels," *Journal of Chemical Physics*, vol. 70, no. 3, pp. 1214-1218, 1979.

[100] T. Tanaka, D. Fillmore, S. Shao-Tang, I. Nishio, G. Swislow, A. Shah, "Phase Transitions in Ionic Gels," *Physical Review Letters*, vol. 45, no.20, pp. 1636-1639, 1980.

[101] T. Tanaka, I. Nishio, S. Sun, S. Ueno-Nishio,"Collapse of Gels in an Electric Field," *Science*, vol. 218, Oct. 29, 1982.

[102] M. Thompson, "Magnetically-Induced Heating Effects in Ferromagnetic Fluids", EECS Department Area Examination Report, Massachusetts Institute of Technology, 1996.

[103] M. Tokita, T. Tanaka, "Friction Coefficient of Polymer Networks of Gels," *Journal of Chemical Physics*, vol. 95, no. 6, Sept, 1991.

[104] M. Tokita, T. Tanaka, "Reversible Decrease of Gel-Solvent Friction", *Science*, 253(6):1121–1123, 1991.

[105] M. Velez-Reyes, K. Minami, G. Verghese, "Recursive Speed and Parameter Estimation for Induction Machines," *IEEE Industry Applications Society Annual Meeting*, vol. 1, pp. 607-611, 1990.

[106] VWR Scientific Catalog, 1996.

[107] N. Ye, S. Scavarda, M. Betemps, A. Jutard, "Models of a Pneumatic PWM Solenoid Valve for Engineering Applications," *Transactions of the ASME*, vol. 114, pp. 680-688.

[108] P. Young,"Parameter Estimation for Continuous-Time Models, a Survey," *5th IFAC Symposium on Identification and System Parameter Estimation*, pp. 23-29, 1979.

[109] X. Yu, A. Tanaka, K. Tanaka, T. Tanaka, "Phase Transition of a poly(acrylic acid) Gel Induced by Polymer Complexation," *Journal of Chemical Physics*, 97(10), Nov. 15, 1992.

[110] M. Zhihong, D. Habibi,"Robust Adaptive Sliding-Mode Control for Rigid Robotics Manipulators with Arbitrary Bounded Input Disturbances", *Journal of Intelligent and Robotics Systems, Theory and Applications*, vol. 17, no. 4, pp. 371-86, 1996.

[111] M. Zrinyi, "Magnetic-Field-Sensitive Polymer Gels," *TRIP*, vol. 5, no. 9, pp. 280-285, Sept. 1997.

[112] M. Zrinyi, F. Horkay, "On the Elastic Modulus of Swollen Gels," *Polymer*, vol. 28, pp. 1139-1143, June 1987.