A study of the reduced-order John Shaw SMA model and its extension for control applications

by

Shailaja Sajja

A thesis
presented to the University of Waterloo
in fulfillment of the
thesis requirement for the degree of
Master of Applied Science
in
Electrical and Computer Engineering

Waterloo, Ontario, Canada, 2012

AUTHOR'S DECLARATION

I hereby declare that I am the sole author of this thesis. This is a true copy of the thesis, including any required final revisions, as accepted by my examiners.

I understand that my thesis may be made electronically available to the public.

Abstract

SMA belongs to a class of so-called “smart materials” which possess properties that can be controlled by application of various types of stimuli – stress, temperature, electric field or magnetic field. In particular, SMA is a smart material which undergoes a temperature- or stress-dependent phase transformation giving it the property of remembering its original shape. Once deformed (up to a certain recoverable strain), SMA returns to its original shape upon heating.

In this thesis, a study of SMA models and techniques to improve the performance of SMA actuators was carried out. In general, an SMA model is required for 3 main purposes: simulation, analysis and for model-based hysteresis compensation.

In this work, the reduced-order form of John Shaw’s partial-differential equation model is chosen for implementation and simulation. The reduced-order form is used because its simpler structure makes it more useful for real-time control applications.

The parameters were estimated for the John Shaw model followed by its implementation in MATLAB. From the view of control applications, a limitation of the John Shaw model is the inability to reproduce the so-called ‘minor loop behavior’ which is observed when the material is subject to cycling resulting in incomplete phase transformations. Modeling minor loop behavior is particularly important in closed-loop strain (or position) control applications since achieving a specific target strain between the two (load-dependent) extremes requires partial phase transformation. Herein, the governing equations are modified to include minor loop behavior. This behavior was tested using damped signals which would be expected to trigger minor loops in the actual SMA and reasonable match is observed from the simulations.

The use of SMA actuators is limited by the relatively slow response time compared to other smart materials. The conventional current saturation (CS) scheme limits the maximum current into the wire at the manufacturer-specified safe current values in order to protect the wire from damage due to overheating. However, this is a conservative limit on the maximum current and hence, the response is artificially slowed. In order to improve the response time, a model-based temperature saturation (MBTS) scheme was developed, in which current is saturated based on model-predicted temperature. The MBTS scheme allows much higher currents to be applied to the wire, while ensuring that the wire is not damaged. Based on simulations using the reduced-order John Shaw model, it is observed that better tracking occurs using the MBTS scheme in the actuation scheme as compared to the CS scheme.

Acknowledgements

I would like to thank several people in helping my journey through this phase of graduate school. Firstly I would like to thank my supervisor Dr. Rob Gorbet for his sincere involvement towards my project and for having lots of patience in providing timely and helpful feedback towards creating better quality work. Thanks to Dr. Eric Kubica for his contribution to the ideas generated in this thesis and related work, for being a patient listener and above all ensuring that I learnt to make better figures. I would also like to thank Dr. Baris Fidan and Dr. Dana Kulic for agreeing to be readers for the thesis and providing timely feedback.

Thanks to General Motors Research Labs for their support in the form of financial assistance and opportunities to visit their facilities. I would also like to thank all the lab mates I have had over the years – Mohamed, Li, Brandon, Anita and Hunter for being helpful in their own special ways, as patient listeners, coffee friends, peaceful lab mates and comic relief(s).

Above all, I would like to thank my family: my dad, didi, Sue, my dog Pongo and my close friends for their continual love, support, life perspective and constant reminders that life is always wonderful and quite short to be taken too seriously.

Table of Contents

List of Figures

List of Tables

List of Symbols

Symbol	Description
L	Length
d	Diameter
β	Transformation Strain
E	Elastic Modulus
E_A	Elastic modulus of austenite
E_M	Elastic modulus of martensite
T_R	Reference Temperature
$T_R\Delta s$	Predicted enthalpy change
ρ	Mass density
μ_c	Critical driving force
v_o	Kinetic law stiffness
ρ_e	Electrical resistivity
h	Heat transfer co-efficient
P_e	Power
ε	Strain
ξ	Phase fraction of martensite
ξ_1	Tensile martensite
ξ_2	Compressive martensite
s_o	Specific entropy

c_I	Phase interaction parameter
T_a	Ambient temperature
γ	Strain gradient parameter
c_p	Specific heat

Chapter 1
Introduction

1.1 Introduction

The goal of this thesis is two-fold.

- To present a technique to improve the response time of shape memory alloy (SMA) actuators.
- To implement a suitable SMA model and increase its applicability to control applications by extending the model to show minor loop behaviour.

SMA belongs to a class of so-called "smart materials" which possess properties that can be controlled by application of various types of stimuli – stress, temperature, electric field or magnetic field. In particular, SMA is a smart material which has the property of remembering its original shape. Once deformed (up to a certain recoverable strain), SMA returns to its original shape upon heating [1]. SMAs typically have recoverable strains of 8%, which is higher than the strains achieved by the other commonly used smart materials – piezoelectrics and magnetostrictives [1].

In most SMA actuators, an SMA wire or spring has a bias load attached which is designed to stretch the SMA material when cold (reset the actuator) but allow some contraction when hot, so that mechanical work can be done. For convenience of control, these actuators are typically heated electrically by running current through them. Partly for their significantly higher electrical resistivity, NiTi alloys are preferred over Cu-based SMA for electrically-heated actuators. In addition, NiTi alloys, and particularly the Flexinol® alloy manufactured by Dynalloy [2], are more suited to repeat cycling at significant recoverable strains, and can achieve millions of cycles under appropriate design conditions.

SMA actuators have several applications in adaptive wing models, where the shape is curved by SMA wings, helicopter rotor blade control, in microgrippers and several others. It is worth noting here that usage of SMA actuators can be either in situations where the overall tracking performance is of relevance or the transient performance is important. In the simulations presented in this work in Chapter 5, the focus is more of achieving improvement with regard to the transient performance.

In this introductory chapter, the commonly employed SMA actuation scheme is discussed followed by a summary of the overall proposed scheme to improve the performance. The implementation of the overall scheme is beyond the scope of the thesis. However individual components have been explored. The summary of the overall scheme is followed by the specific goals of this project and discussion of the organization of the thesis.

1.2 Motivation

Besides high recoverable strains, shape memory alloys have the advantages of compact size and excellent cycling performance on repeated loading. However, shape memory alloys also have certain limitations – slow response time in comparison to other smart materials, hysteretic behaviour between phase transformation and temperature change, and low energy efficiency [1].

In this research, the goal is to improve the overall performance of the NiTi based SMA wire actuators. To achieve this goal, an actuation scheme is proposed with the following benefit:

- Improved response time by using a non-linear control strategy which allows for faster heating while still guarding against overheating.

The typical actuation scheme with some variations between authors [3, 4] for SMAs using strain feedback is illustrated in Figure 1-1. In this scheme, a linear control law (e.g., P/PI/PID) is used with the strain error as the input and the current as the output. In order to protect the SMA wire from overheating, the maximum current into the wire is limited at the manufacturer-specified safe current. This method of limiting the current is called "current saturation" (CS) in this thesis.

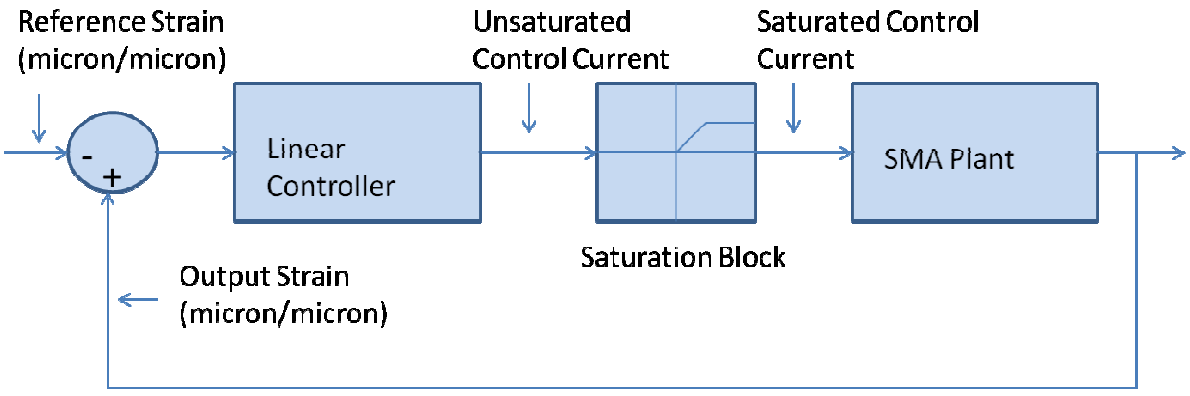

Figure 1-1: Strain-based feedback system for an SMA wire actuator, using so-called "current saturation" (CS)

This scheme has two drawbacks:

- First, the linear control law does not account for the nonlinear response of the SMA due to the hysteresis between the strain and the temperature (or current).

- Second, the manufacturer-specified safe current is conservative, and specified so as to limit the wire temperature to a safe region were the current to be applied indefinitely. Larger currents can safely be applied for short periods without reaching damaging temperatures, thus allowing for faster heating and better actuator time responses.

In this thesis, a methodology is provided to address the second drawback. A solution of the first drawback is discussed as future work to the thesis. In the next section, an overview of the methodology is provided to address the above-mentioned drawbacks.

1.3 Methodology

In this section, a high-level overview is provided of the methodology of the project. In the previous section, the two drawbacks of the CS actuation scheme were mentioned. The techniques proposed to deal with the drawbacks are described here. The issue of hysteretic nonlinearity has been addressed by several researchers [5, 6, 7, 8, 9, 10] by using model-based and non-model-based control techniques. One of the techniques of doing this is by employing the inverse of an SMA model to provide a feed forward component. The next step, as a future work to this thesis is to do model-based hysteresis compensation.

In this thesis, the John Shaw model [24, 25] was chosen for simulation purposes. The model itself is mature and has been tested to exhibit reasonable accuracy in modelling a variety of behaviours demonstrated by SMAs. However the complexity of the original John Shaw model precluded its applicability to control based applications. Recently John Shaw *et al.* [25] developed a reduced-order version of the model by introducing several simplifying assumptions. However a significant limitation of the model remains its inability to model the minor hysteresis loops which occur in SMA subject to partial phase transformation. Modelling minor loop behaviour is particularly important in closed-loop strain (or position) control applications since achieving a specific target strain between the two (load-dependent) extremes requires partial phase transformation. In this thesis, the applicability of the reduced-order John Shaw model to control applications is extended by modifying the model such that it can display minor loop behavior.

The manufacturer-specified safe current is a conservative limit and increases the time taken for the wire to attain the desired temperature, hence, slowing the response time. However, it is heating the wire above the safe temperature, and not the current itself, which can potentially destroy the wire. An improvement can be attained in the response time if the current is allowed to rise to substantially higher magnitudes because the corresponding joule heating will also be faster. However, the

magnitude of the current applied and the duration for which the current is applied must be carefully controlled to protect the wire from reaching temperatures which will damage it.

Thus, a hybrid controller could be designed which would perform strain tracking at one level, with a supervisory level ensuring the temperature does not exceed the safe value. However, the difficulty associated with direct wire temperature measurement is significant [37, 38]. This is primarily due to the challenges in attaching thermocouples to wires and the cost associated with non-contact temperature monitoring devices. This has been explained in further detail in Chapter 2, Section 2.7.

The technique proposed to deal with this problem is to use a heating model of SMAs which relates the applied current to temperature to predict the temperature of the wire. By using a model to predict the temperature, the need for additional temperature monitoring devices is eliminated.

Temperature instead of current is used as a deciding parameter regarding when to saturate the controller current. This is based on setting a threshold temperature, which is compared to the temperature value predicted by the SMA heating model. This value is used as an input to the controller. This method is referred to here as the "model-based temperature saturation" (MBTS) scheme. An illustration of the proposed scheme is presented in Figure 1-2. In Figure 1-2, the terms TMM and TEM stand for thermomechanical model and thermoelectric model respectively.

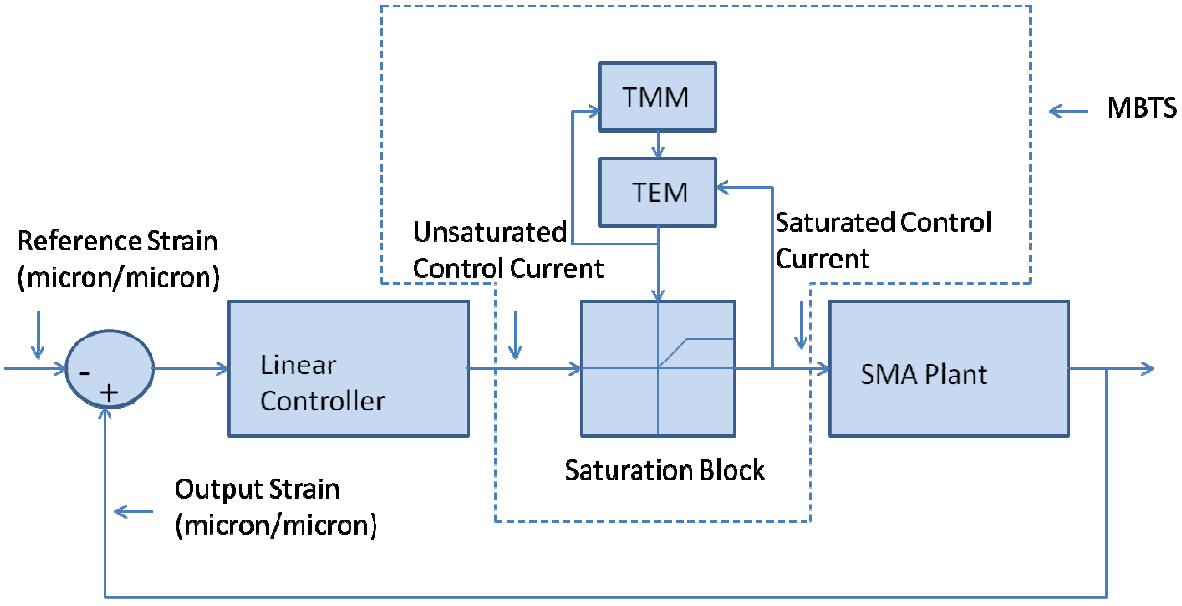

Figure 1-2: Improved strain-based feedback system for an SMA wire actuator, using model-based temperature saturation (MBTS)

The next section discusses the goals of the project. This is followed by a description of how the thesis comes together in providing the background required to understand and solve the problem, and the organization of the thesis.

1.4 Contributions

The overall contributions of this thesis are listed below:

- Extended the applicability of the John Shaw model to control applications by introducing minor loops.
- Implementation of the extended model in MATLAB and Simulink.
- Development of the concept of the "model-based temperature saturation" scheme. Comparison of the "current saturation" and "model-based temperature saturation" using the John Shaw model.

1.5 Thesis Organization

This thesis presents techniques to improve the performance of SMA-based actuators. The problem of hysteretic nonlinearity in SMAs has been addressed by several researchers and several techniques have been proposed to deal with the same. These techniques can be broadly classified into model-based and non-model-based control techniques. A review of different techniques is presented in Chapter 2.

To guide the choice of an appropriate model for the model-based controller used in this work, a detailed review of the existing models is also provided in Chapter 2. These models are developed from various perspectives: thermodynamics, thermomechanics, micromechanics and phenomenology. The reduced-order John Shaw model is chosen for use in the model-based controller scheme. A detailed overview of the John Shaw model – theory, parameter estimation and simulation results – is provided in Chapter 3**.** The MBTS control scheme is described and tested in Chapter 4. The conclusions and the future work are presented in Chapter 5 of the thesis.

Chapter 2
Background: Modeling and Control

2.1 Introduction

This chapter provides the background regarding the important SMA mechanics, the SMA modeling techniques and the significant models. The various behavioural aspects of SMAs are explained in the first section with emphasis on the phase transformations, shape memory effect and rate dependence. This is followed by the modeling section. SMA models typically consist of component modules with various degrees of coupling between them. This is explained through a block diagram which illustrates how the modules combine together. This is followed by a section which introduces the concept of the scales of modeling. The choice of the model used in this work is justified in this section. Finally, a review of the various control approaches used for SMA actuation is presented.

2.2 SMA Properties

SMA exists in two phases – low temperature martensite, which exhibits low lattice symmetry, and high temperature austenite, which exhibits high lattice symmetry. In particular, the material used in SMA is NiTi (typically consisting of 55-56% Nickel and 44-45% Titanium).

In some situations, a third intermediate phase is seen: the R-phase. However, the existence of R-phase is typically ignored in modeling and control applications since the R-phase related effects, even if present, are very minimal in the thermomechanical behaviour of SMAs. Hysteretic behaviour is observed in the transformation between the austenite and martensite phases. In Figure 2-1[11] the strain versus temperature response is illustrated for a typical SMA (NiTiCu) actuator initially at the strain of 4% under the load of 10 N. On increasing the temperature from austenite start (A_s) to austenite finish (A_f) temperatures it contracts. However, for the recovery process the elongation does not start until the martensite start (M_s) temperature is reached, producing the classical hysteresis loop. The various behaviours of SMAs are manifestations of the transformation which occurs between these two phases. These transformations can be brought about by the application of temperature or stress. Note that the transformation temperatures shown in Figure 2-1 [11] are for that specific material composition and applied stress. Transformation temperatures vary with composition and for a given material, increase approximately linearly with applied stress, as seen in Figure 2-2 [43].

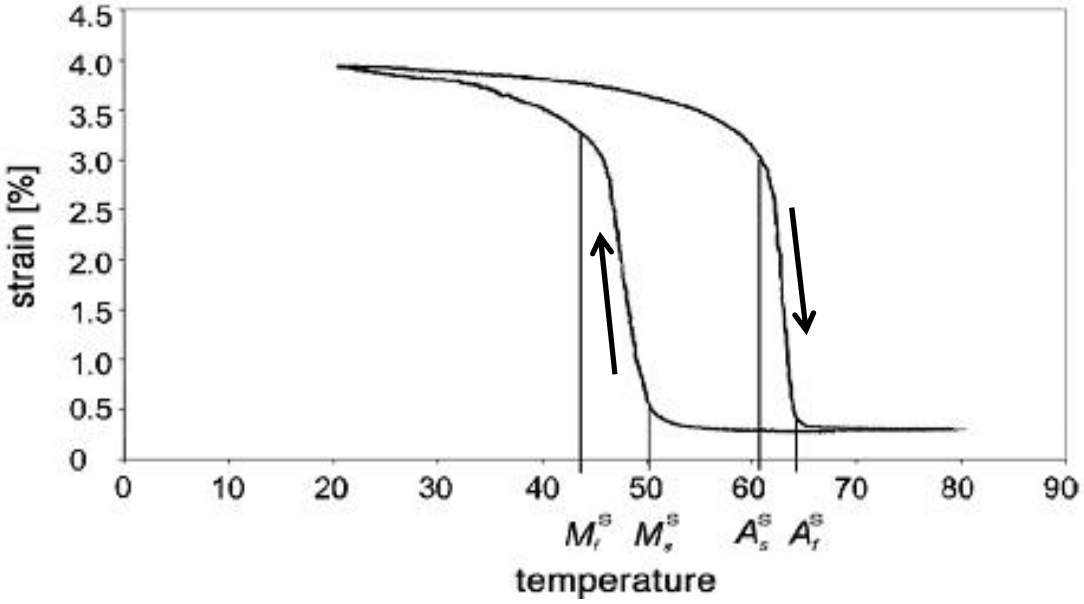

Figure 2-1: Deformation-temperature curve for a typical NiTiCu SMA wire loaded at 10N [11]

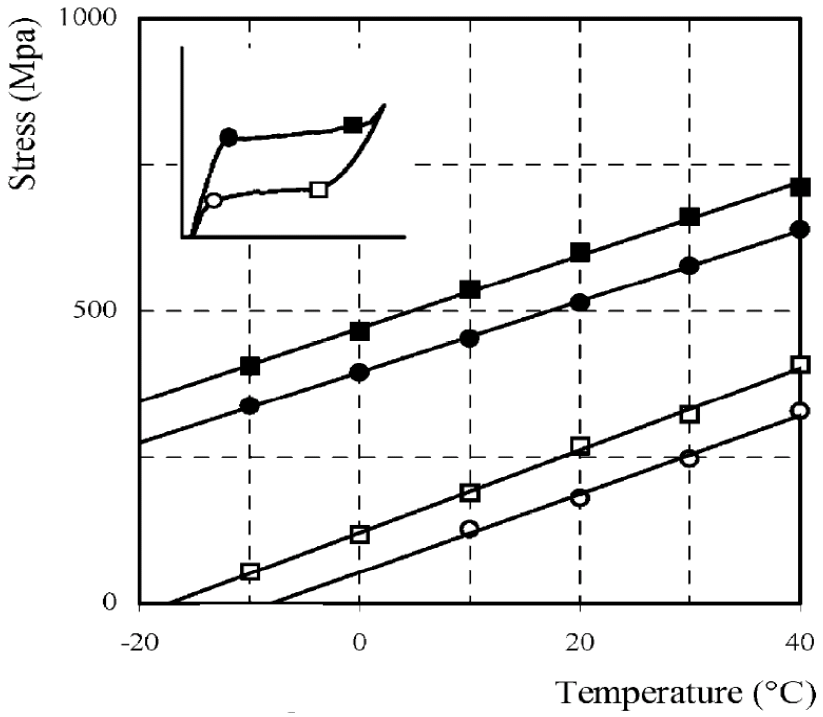

Figure 2-2: Increase in transformation temperatures with increasing stress (A_f and A_s are indicated by lines with open circles and open squares, whereas M_s and M_f are indicated by lines with black circles and black squares, respectively) [43]

Due to the low symmetry of the martensite, transformation from austenite to martensite can result in different variants of martensite. In the 3-D case, 24 possible variants of martensite variants can be formed. This is illustrated in Figure 2-3.

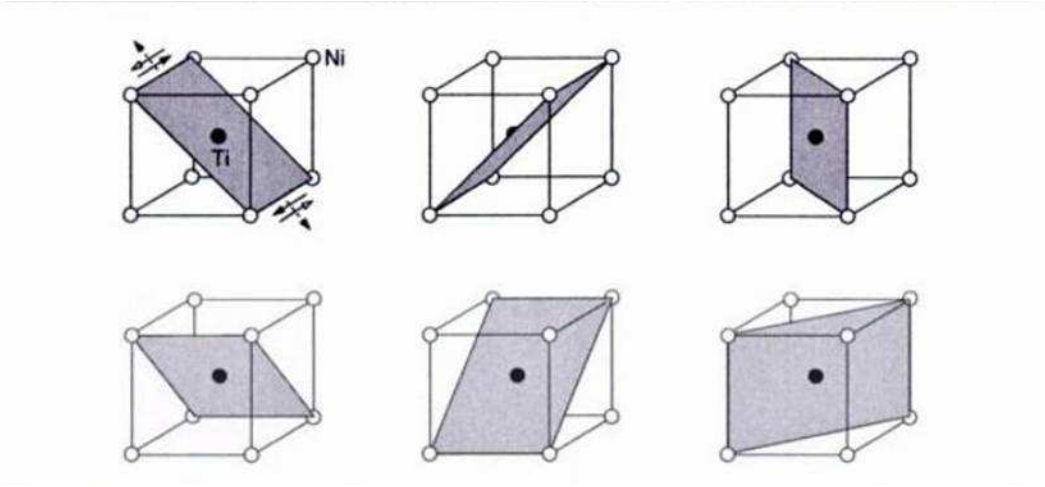

Figure 2-3: NiTi lattice cell illustrating the 6 face-diagonal planes. Each of the planes can shift in the direction indicated by the dashed arrows, and can shear in the direction of the open and closed arrows [14]

For this project, the scope is restricted to the 1-D or uniaxial case. This is justified since the stress is applied uniaxially along the length of the wire. The uniaxial case is considered while explaining the various physical properties and behaviour of SMAs. The transformation from austenite to martensite under zero load results in twinned martensite. This is illustrated in Figure 2-4. ε in this figure denotes the transformation strain. In Figure 2-4 (a), the austenite to martensite transformation is represented in 1-D. A single uniaxial lattice element of the austenitic phase and of the twinned martensitic phase is depicted in Figure 2-4 (b) and Figure 2-4(c) respectively. The two variants of detwinned martensite (M_+ and M_-) formed depending on the direction of the application of stress are depicted in Figure 2-4 (d). Twinned martensite refers to the martensite formed under stress-free conditions, and hence there are negligible macroscopic strains. If the formation of martensite occurs under the application of stress, then detwinned martensite is formed. In the uniaxial case, only two variants of detwinned martensite are possible, M_- and M_+, depending on the direction of the applied stress.

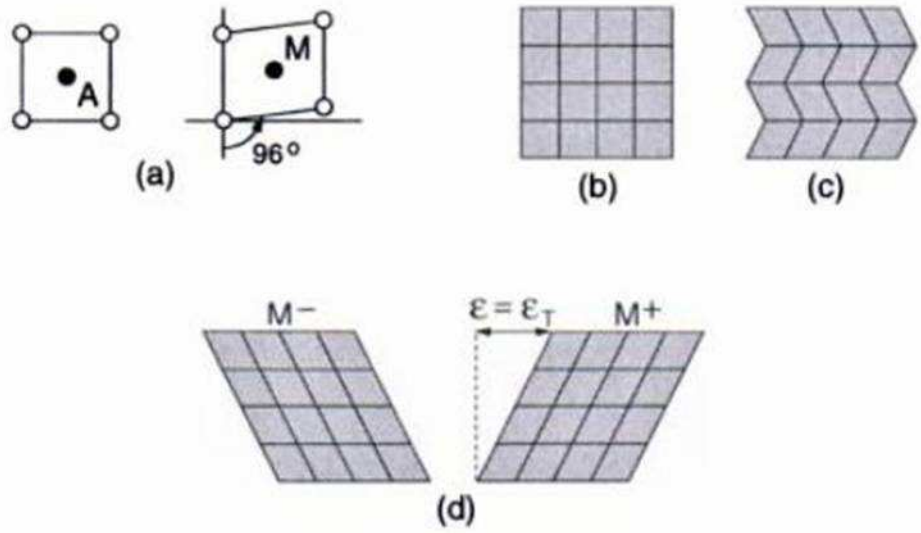

Figure 2-4: (a) Austenite to martensite transformation in 1-D (b) Uniaxial lattice element in the austenite phase (c) twinned martensite (d) detwinned martensite [14]

2.2.1 Temperature-induced phase transformation

Austenite is the stable phase at high temperatures and martensite is the stable phase at low temperatures. The temperatures at which the transformation from martensite to austenite starts and finishes are referred to as A_s and A_f respectively. Similarly the temperatures at which the transformation from austenite to martensite starts and finishes are referred to as M_s and M_f respectively.

Consider if a sample of shape memory alloy initially in the austenite state is cooled to temperatures below M_f in the absence of applied stress, twinned martensite is formed. Since the resulting macroscopic strain is minimal in the case of twinned martensite, this is naturally formed in the absence of external stress. If this sample is heated above A_f, the original crystallographic state is re-obtained. This is illustrated in Figure 2-5 (a). The hysteresis between phase fraction and temperature is illustrated in Figure 2-5 (b)

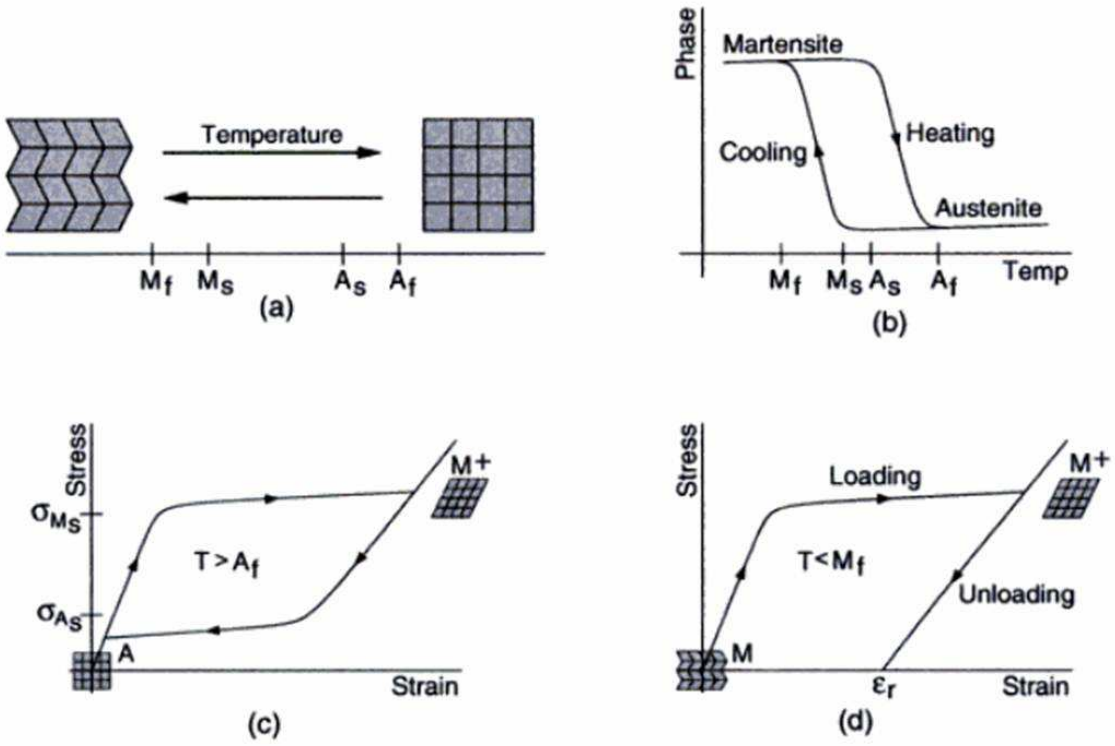

Figure 2-5: (a) Temperature-induced phase transformation between austenite and twinned martensite in the absence of an applied load (b) Hysteretic relation between austenite and martensite as a function of temperature (c) Stress-induced phase transformation and pseudoelastic behavior for $T>A_f$ (d) Quasiplastic behavior and residual strain ε_r generated when $T<M_f$ [14]

2.2.2 Stress-induced phase transformation

Stress can also be used to induce transformation between austenite and martensite. However, the response varies depending upon the value of the prescribed temperature (fixed).

Consider the case where the temperature $T>A_f$. The stress is gradually increased from 0 to values above σ_{Ms} (martensite starting stress) and then decreased to σ_{As} (austenite starting stress). The stress versus strain response under these conditions can be divided into three distinct regions.

- For $\sigma < \sigma_{Ms}$, the stress versus strain response is close to linear. Hooke's law can be used to model the response in this region.
- For $\sigma > \sigma_{Ms}$, transformation to detwinned martensite occurs and elastic behaviour is observed.
- After σ is increased to a value above σ_{Ms} and then reduced back to σ_{As}, reverse transformation occurs and the sample returns back to its original state with no residual strain.

Since the sample returns back to its original form on the removal of stress, this phenomenon is known as superelastic or pseudoelastic. This is because the elastic behaviour is only displayed in high temperature regimes. The behaviour described above is illustrated in Figure 2-5(c).

Consider now the case where the temperature $T<M_f$. On the application of stresses $\sigma > \sigma_{Ms}$, twinned martensite is transformed to detwinned martensite. This detwinning occurs since the variant aligned with the stress is energetically favoured. If the stress is decreased back to 0, some residual strain remains (as opposed to 0 strains in the high temperature case). This behaviour is termed quasiplastic behaviour, illustrated in Figure 2-5(d). The term quasiplastic distinguishes the behaviour from classical plastic deformation, since the residual strain or deformation is not permanent and can be removed on heating.

2.2.3 Shape memory effect

The stress-induced phase transformations and the temperature-induced phase transformations render the shape memory alloys with unique memory capabilities. As seen in the case of temperature-induced phase transformations, when the stress applied to the detwinned martensite is reduced to zero, a certain residual strain still remains. However, if this sample is heated above A_f, the strain is recovered and the material returns to its original shape. This process of recovery of stress-induced strains through heating is called the shape memory effect.

2.2.4 Rate dependence

The hysteretic behaviour of shape memory alloys is affected by the rate of application of the stresses or the rate at which the strain is attained. It has been illustrated (Figure 2-6) in the works of Shaw and Kyriakides [12] that the stress-strain responses differ based on the stress-rate or strain-rate applied.

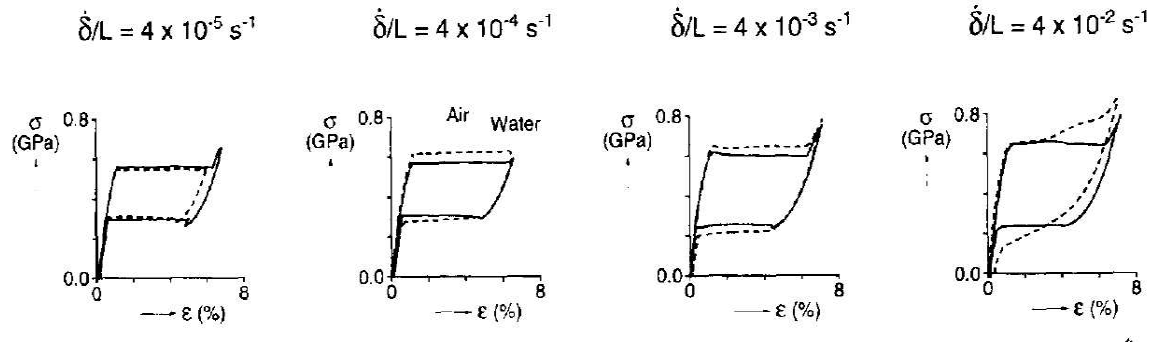

Figure 2-6: Comparison of the stress-strain responses (solid lines – water, dotted lines – air) resulting from different strain rates (denoted by $\dot{\delta}/L$) [12]

2.2.5 Minor Loops

Before defining the behaviour known as a 'minor loop', a major loop needs to be defined. A major loop is defined as the loop formed in the input-output plane when the SMA undergoes a complete transformation from 100% martensite to 100% austenite and back again. Looping, or hysteresis, behaviour can be seen notably in the relationships between temperature and phase fraction, temperature and strain, and stress and strain.

In cases when partial transformation takes place such that starting percentage of austenite is less than 100 or greater than 0, the loops formed on the transformation and back are labelled as minor loops, as illustrated in (Figure 2-7). Modelling minor loop behaviour is particularly important in closed-loop strain (or position) control applications since achieving a specific target strain between the two (load-dependent) extremes requires partial phase transformation.

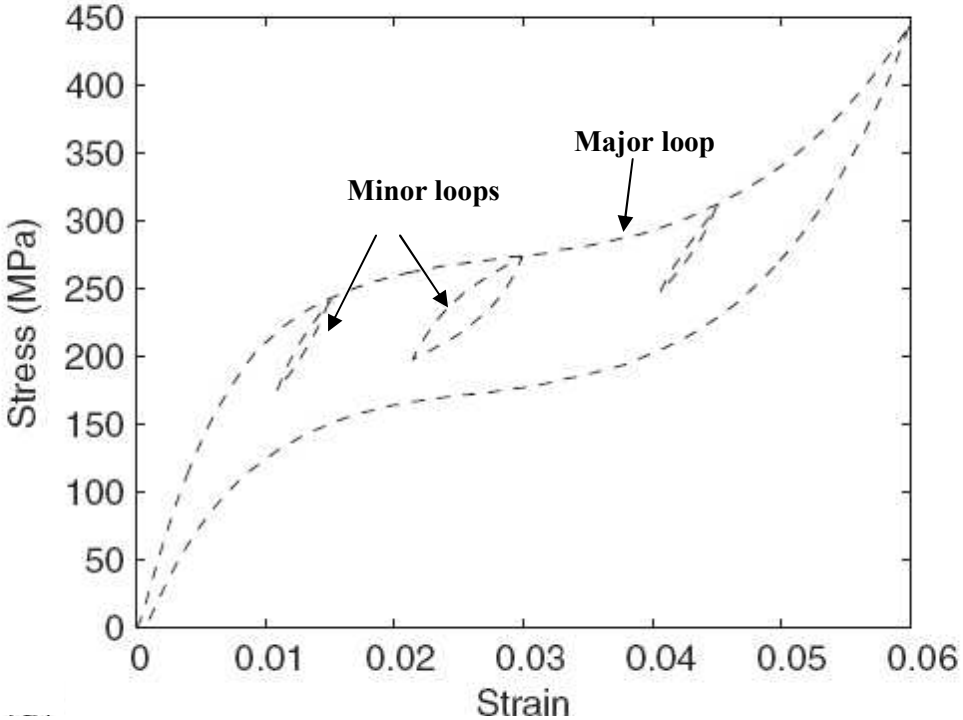

Figure 2-7: Stress-Strain curve illustrating minor loops and major loop [13]

2.3 Modeling Process

Modeling electrically-heated SMA actuators has several aspects to it: the thermoelectric relationship between the current and temperature, the phase kinetics relationship (or, "kinetic law") between temperature and phase fraction, and lastly, the mechanical relationship (or, "constitutive law") between the phase fraction and strain. Some of these relationships are coupled together in some of the models.

Besides modeling each of these relationships and coupling them together, there are other complexities which can also be modeled depending upon the form of the SMA actuator as well as the target application. The dimensionality of the model is a clear example of such a complexity as there are 1-D, 2-D and 3-D models, which determines the number of martensite variants obtained on transformation. There are 24 possible martensite variants in a 3-D situation which are characterized by changes from 90^o to 96^o in the lattice [12]. In a 1-D case, only two martensite variants, M_+ and M_- are formed. The variant formed depends on the direction of loading while converting from austenite to martensite. For a wire actuator due to its uniaxial nature, 1-D models are generally assumed.

Other examples of modeling complexity include whether the model accurately reproduces minor hysteresis loops which result in the phase kinetics from partial temperature cycling of the alloy, and whether the model predicts the effect of varying stress-rate on mechanical response (called "rate dependence"). Indeed, the importance of these effects (and hence, the selection of an appropriate model) depends on the specific application. In an on-off electrical actuator where partial thermal cycling does not occur, it may be desirable to neglect the impact of minor loops in favour of a model of reduced complexity. Similarly, there may be applications where the impact of varying stress-rate is less important.

The typical SMA model can be divided into three important sub-models, namely the heating model, phase kinetics model, and constitutive model, and an optional resistance model. The resistance model is included when the output of resistance is required. This resistance is calculated from the strains, temperature and phase fractions. Taking into account the impact of these variables on resistance can be useful, for example, in updating the heating model in real-time, since the input power is a function of the material resistance which changes with the transformation.

The heating module is also known as the thermoelectric model (TEM) and the phase kinetics and constitutive model together constitute the thermo-mechanical (TMM) model. This is discussed in greater detail in Section 2.3.4. The block diagram illustrating the smaller models constituting the SMA model is shown in Figure 2-8.

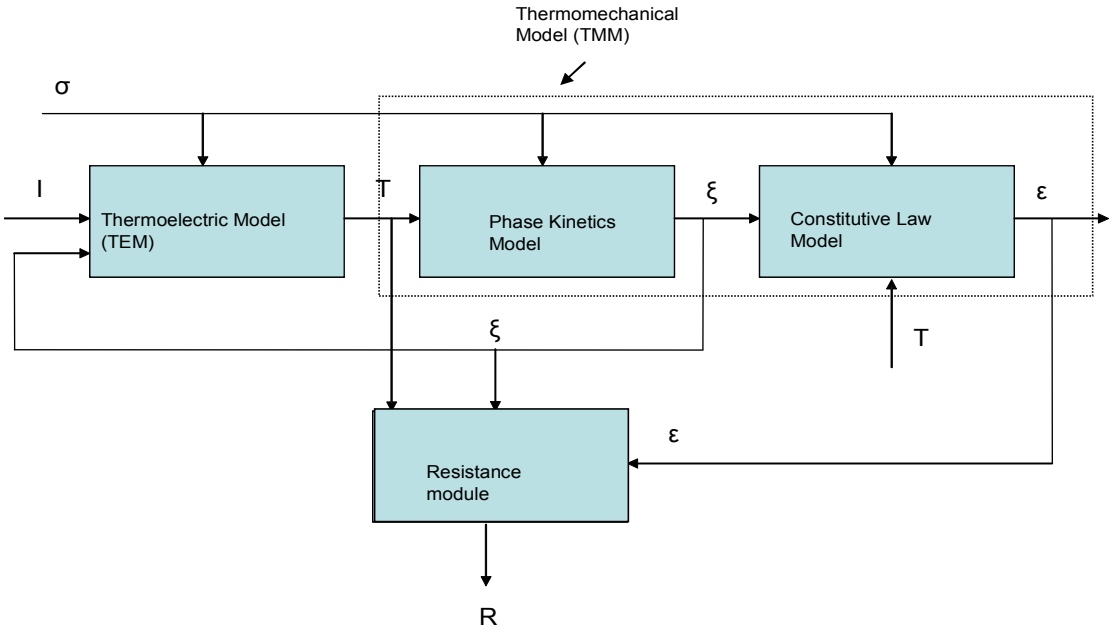

Figure 2-8: Block diagram of a typical SMA model

In this diagram, T denotes the temperature, σ denotes the stress, ε denotes the strain, ξ denotes the phase fractions, I denote the current and R denotes the resistance.

2.3.1 Heating Model (TEM)

In the heating model, the dynamic relationship between current, I, and temperature, T, is modelled. The heating equation is given in Equation 2.1 for natural convection of an electrically heated wire. In this equation, the parameters are as follows: ρ is the density of the wire material, V is the volume of the wire, A is the surface area of the wire, R is the electrical resistance of the wire, h is the heat transfer coefficient, and C_p is the specific heat of the material. Different models expand upon this basic equation by taking into account various factors such as additional heat losses due to radiation and conduction, and the temperature dependency of h and C_p. Some heating models also include an additional term on the right hand side of the equation to account for the latent heat released or absorbed by the system when phase transformation takes place.

$$\rho C_p V \frac{dT}{dt} = R\big(i(t)\big)^2 - hAT(t) \tag{2.1}$$

2.3.2 Phase Kinetics Model

The phase kinetics module models the relationship between the temperature and the phase fractions. The cyclic transformation between austenite and martensite is hysteretic and hence, the relationship between phase fraction and temperature is also hysteretic. Depending on the specific SMA model in consideration, the phase kinetics model can either be explicitly isolated as a separated model or is coupled with the remaining equations in the system where it is difficult to isolate. Stress is required as an input if the model accounts for stress-dependency of the transformation temperatures.

2.3.3 Constitutive Model

This module models the relationship between inputs namely, the temperature, the phase fractions, the stress and the output strain. It depends on the physical configuration of the actuator, the type of bias forces it is subject to, etc. It comes from the physical world and is defined by the physical constraints imposed on the material.

2.3.4 Coupling between TEM and TMM

As seen in Figure 2-8, the TEM and the TMM are coupled to each other through the phase fractions and stresses. The goal of coupling the TEM and the TMM is to have an accurate model description accounting for the effects of variable phase fractions and stress on the TEM. This also helps in accurate characterization of the temperature predicted by the model. As was briefly discussed in the introduction this temperature prediction is important in this work since it is used to protect the wire from overheating.

Phase transformations are accompanied by the release or absorption of latent heat as well as by the evolution of material properties such as thermal conductivity, heat capacity and electrical resistivity. Most of the existing modeling works take into account the release and absorption of latent heats. The evolution of material properties has received comparatively less attention in literature. There have however been some important works in modeling the evolution of the material properties in response to the phase fraction evolution [15, 16, 17]. It makes intuitive sense that these properties should evolve with phase fractions since their values are fairly different in the austenitic and martensitic phases.

The effects of the phase transformations from the TMM on the parameters of the TEM have been examined in terms of the work done previously by researchers [15, 16, 17]. The important parameters

in the heat equation which are affected by the phase fraction evolution are: K – thermal conductivity, C_p – specific heat capacity, and ρ_E – the electrical resistivity. In the initial work by Bhattacharya *et al.* [15], the authors look at the effect of varying each of the abovementioned parameters in response to phase evolution using the linear combinations of austenite and martensite values as shown in Equation 2.2. In Equation 2.2, the offsets (the first terms on the RHS of the equations) refer to the austenite values of the corresponding parameters and ξ refers to the martensite fractions which vary from 0 to 1.

$$\begin{aligned} K(\xi) &= K_A + \xi(K_M - K_A) \\ \rho_E(\xi) &= \rho_{E,A} + \xi(\rho_{E,M} - \rho_{E,A}) \\ C_p(\xi) &= C_{p,A} + \xi(C_{p,M} - C_{p,A}) \end{aligned} \tag{2.2}$$

The authors have shown that accounting for thermal conductivity and electrical resistivity helps predict a shorter time for transformation, however, accounting for heat capacity does not have much impact on the transformation [15]. In our current work, thermal conductivity is ignored so the above mentioned findings do not affect the modeling work in this thesis.

2.4 Modeling across Scales

The SMA models can be classified into microscopic, mesoscopic and macroscopic, based on the level at which the constitutive relations are developed.

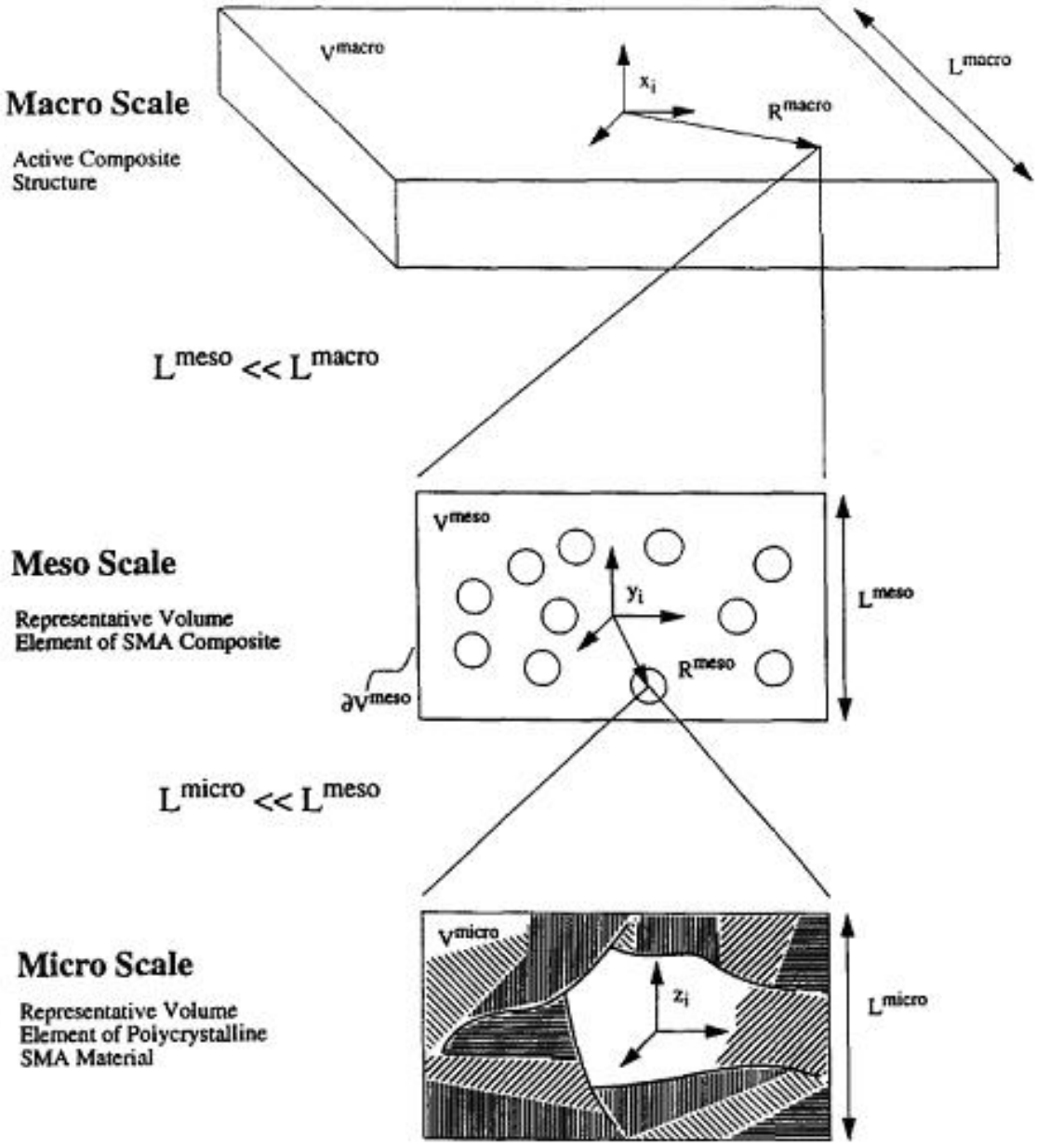

Figure 2-9: Depiction of modeling across scales [18]

The work of Boyd and Lagoudas [18] highlights modeling scales for SMA composite materials where it spans work done in the microscopic, mesoscopic and macroscopic realms. This concept of modeling across scales is illustrated in Figure 2-9. As such, a *point* on the *macroscopic* scale

corresponds to *lattice* volume elements which make up the *mesoscale* and this corresponds to individual *crystals* that make up the structure at the *microscopic* level. In the next three sections, the models belonging to each of the three scales are described. The microscopic and macroscopic models are described first, followed by the mesoscopic models. The mesoscopic models combine the favorable features of models belonging to the microscopic and macroscopic scales.

2.4.1 Microscopic Scale Modeling

This category of models concentrates on developing theories for the explanation of micro-scale behaviour [14]. The elastic, thermal and chemical free energies are constructed at the micro-scale level. These relations are used to quantify the phase transformations. These models are useful in explaining the fundamental behaviour of SMAs and also in understanding how the microscopic behaviour affects the macroscopic behaviour. However, the high degree of complexity of the models presents a severe limitation on their applicability to a wide array of engineering applications.

The significant work with regard to microscopic models has been done by Gao *et al.* [19] and Lu and Weng [20]. Gao *et al.* developed a 3-D multivariant model at the microscopic level. In the model by Lu and Weng, self-consistent relations are developed which connect the stress and phase transformation strains at the grain level with those at the polycrystal level.

2.4.2 Macroscopic Scale Modeling

The macroscopic models are designed based on the description of the material behaviour at the macroscopic level. These descriptions are based on phenomenological constructs, thermodynamics constructs or curve-fitting data [14]. Phenomenological construct refers to constructs derived purely from experimental data but without any physical basis. The macroscopic models are simple and accurate. However, some of the macroscopic models are developed based on phenomenological constructs. Hence, they lack strong physical significance. Some of the significant macroscopic models are the Tanaka-based models [21, 22, 23], the John Shaw model [24, 25] and the Preisach models [26, 27, 28, 29, 30, 31].

Three models are categorized together as the Tanaka-based models: the original model proposed by Tanaka [21], Liang and Roger's model [22] and Brinson's model [23]. In these models, one of the phase fractions is chosen as an internal variable whereas stress and temperature are the control variables. The functions governing the evolution of the phase fractions are exponential functions for the Tanaka models and cosine functions for the Liang and Rogers and Brinson's model.

Preisach modeling is an extensively worked upon area for modeling hysteresis in various types of materials – ferromagnets, piezoelectrics and shape memory alloys. Though developed to explain the physical aspects of magnetism, they are purely phenomenological models. The models have a strong mathematical framework and high degree of generality; hence, have been widely employed by several researchers. Some of the significant works with regard to Preisach modeling of SMAs have been done by: Huo [26], Ortin [27], Hughes and Wen [28], Gorbet, Wang and Morris [29], Ktena and Fotiadis [30], and Choi and Lee [31]. The basic kernel of the Preisach model is the hysteresis relay which is characterized by pair of switching values (α, β) such that $\alpha \geq \beta$ [32].

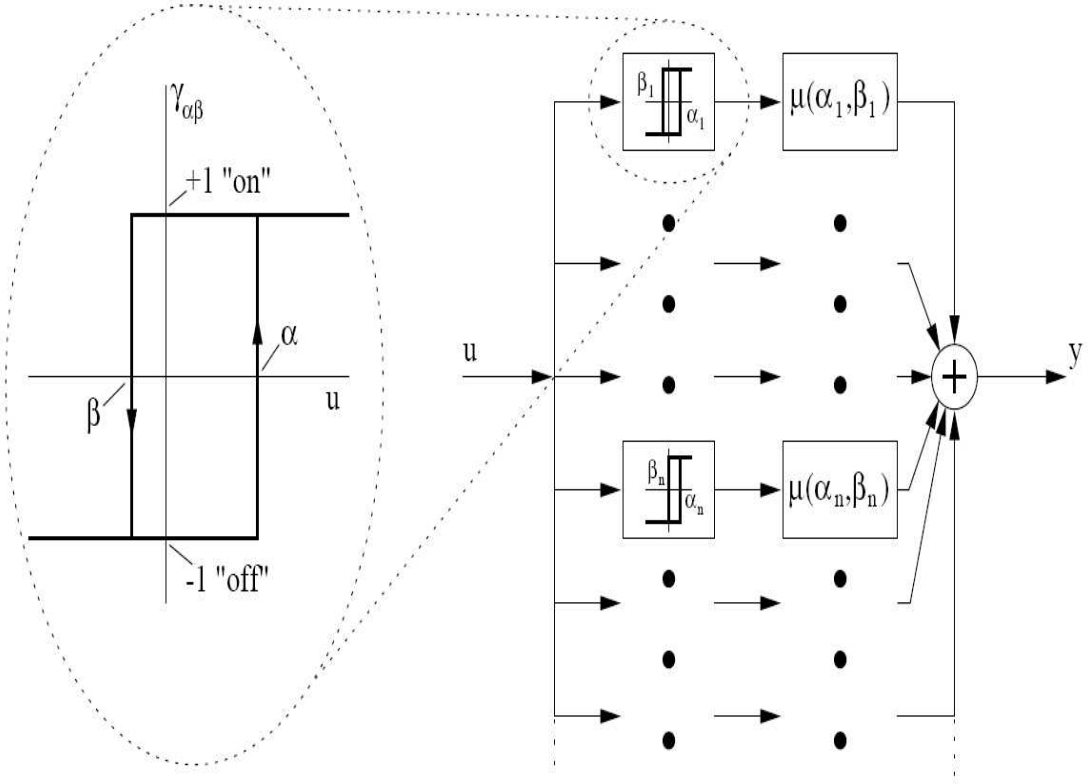

Figure 2-10: Schematic of the Preisach model [32]

The vertical portions of the relays are irreversible whereas the horizontal portions are reversible. Each of the weights $\mu(\alpha,\beta)$ describes the relative contribution of the relay to the overall hysteresis, as shown in Figure 2-10.

The overall output $(y(t))$ is written in terms of the input $u(t)$, the weights $\mu(\alpha,\beta)$ and the relays $\gamma_{\alpha\beta}$ as follows:

$$y(t) = \iint_P \mu(\alpha, \beta)[\gamma_{\alpha\beta} u](t) d\alpha d\beta \tag{2.3}$$

In the context of SMAs, the output *y(t)* usually refers to the strain and the input *u(t)* to the temperature. In discrete Preisach model, the number of relays is chosen based on the required accuracy. Greater number of relays results in more accurate representation. If the number of relays tends to infinity, the continuous Preisach model is obtained.

A sequence of works with regard to macroscopic modeling has also been done by John Shaw *et al* [24, 25]. This model combines a phenomenological approach for the constitutive models with the concepts of irreversible thermodynamics for the phase kinetics model. The initial work [24] is very comprehensive in its treatment – including the complexity, estimation and calibration of parameters and the evaluation of performance in various regimes. However, this model is complex in nature. In [25], a reduced-order model is presented, in which a reduced model complexity is achieved by making assumptions of uniform strain and temperature across the length of wire and equal values of the elastic moduli of austenite and martensite. The simpler structure presented by the reduced-order John Shaw model makes it more amenable for control applications.

2.4.3 Mesoscopic Scale Modeling

Mesoscopic models focus on the lattice level description of the material; hence, these models occupy the middle scale in the multiscale hierarchy. The energy relations in this category are constructed for a representative lattice or control volume. The energy relations then undergo averaging of some form to derive macroscopic constitutive relations. A series of significant mesoscopic models have been developed by Muller-Achenbach [33], followed by Seelecke [34]. The theories developed in these models have been adapted by Massad and Smith [35] with the primary goal of providing an energy formulation to the Preisach model. This led to the formation of the homogenized energy model (HEM). Compared to the Preisach model (Equation 2.3), in the HEM the basic kernel $\gamma_{\alpha\beta}$ is modified to make it energy based. If the output is strain, the energy-based kernel is defined for the single crystal strain and the weights of the Preisach model are modified to be distribution functions.

2.5 Choice of the model for the current work

In this research, the model chosen for study and implementation is the John Shaw model.

The favourable features of the John Shaw model are:

- The John Shaw model has demonstrated ability to model the properties of rate dependence, shape memory effect and superelastic effect with acceptable accuracy [24].
- Relatively few parameters and a simple parameter identification procedure.
- Also from the point of view of control, the reduced-order form of the model presents a simpler structure as compared to contemporary SMA models.

A detailed description of the John Shaw model – theory, parameter estimation and simulation results is presented in Chapter 3.

2.6 Control techniques for hysteresis compensation

In this section, a summary is provided of the various techniques used for the control of SMA actuators. Emphasis is placed on model-based control techniques, followed by a brief description of the other significant techniques. Also, a review is provided of the approaches employed by various researchers to improve the response time of SMA actuators.

2.6.1 Model-based control techniques

Adaptive control of the SMA wire actuator

This method of control has been utilized by Webb, Lagoudas and Kurdilla [5] for real-time control. This involves modifying the control law to accommodate the time varying nature of the system parameters. The use of an adaptive hysteresis model over a fixed hysteresis model is justified for two reasons: Firstly, a fixed hysteresis model can lead to significant errors in tracking control if there are regions in the input where the model is inadequate to capture the true response. Also, repeated cyclic deformation of SMAs under mechanical and thermal loads can lead to microscopic residual stresses that accumulate around the defects in the alloy. Hence, as these stresses accumulate, the envelope of the hysteretic behaviour gradually changes. This causes the identified model to gradually become mismatched and leads to an increase in the prediction error and eventual failure of the model. The adaptive hysteresis model developed in [5] was based on the Preisach model with the Krasnosel'skii and Pokrovskii (KP) operator as the kernel function. A comparison was carried out between closed-loop control using the adaptive hysteresis model and the open loop control using a fixed hysteresis model. A clear improvement is observed in the tracking performance on using the former scheme as compared to the latter.

Inverse model-based feed forward with model-based feedback

This method of control as developed by Majima, Kodama and Hasegawa [7], is composed of both a PID feedback loop and a feed forward loop. The SMA actuator for which the analysis is presented is a bias type SMA actuator consisting of an SMA coil spring, a bias spring, and a moving mass. A novel control system is proposed based on a static model of the SMA actuator. First, a Preisach-like model is developed for the actuator which relates temperature $T(t)$ and detwinned martensite fraction $\xi_s(t)$. Because the load of the bias spring is always above the detwinning stress, it is assumed that no twinned martensite is present in the material. The discretized Preisach model for the detwinned martensite fraction is then:

$$\xi_s(t) = \sum_{i=1}^{N} \frac{1}{N} \gamma_{\alpha_i \beta_i}[T(t)] \tag{2.4}$$

where $\gamma_{\alpha\beta}$ represents the elementary relay hysteresis operator whose output switches from 0 to 1 at the input α and from 1 to 0 at the input β. The cumbersome method of identifying parameters for a Preisach model is simplified in this work by setting each weighting function to $1/N$ and estimating the distribution of the switching points α and β instead of $\mu(\alpha, \beta)$. The identification procedure is described in [7] for an actuator using $N=500$ elements.

A dynamic system model is created by coupling the output of the above equation with relations for the SMA and bias spring force as a function of phase fraction, and finally with the standard equations for a moving mass subject to a force. At the input to this dynamic model, a first-order transfer function relates control signal duty ratio to temperature. For control, a static model of a form similar to (2.4) is identified from control duty cycle to position, by applying a slowly varying duty cycle and measuring displacement. This static model is then inverted in the feed forward block, and a comparison is made between position tracking performance using regular PID control and PID with the addition of the static model inverse in feed forward. An improved performance was achieved on using the model inverse based feed forward, and also the limit cycle oscillations observed previously were significantly reduced.

Neural network based control

Neural networks (NN) have also been employed for model-based control of SMA actuators in the works by Ma and Song [8, 9]. In the first work, open-loop inverse model-based feed forward control was used [8]. The steps involved are as follows: First, based on one representative hysteresis loop, the

forward neural network is designed. The inputs to the neural network are applied voltage and a tag signal (to denote whether the voltage signal is increasing or decreasing) and the output is displacement. The next step is the verification of the ability of the neural network to model the relationship between the voltage and the displacement. Finally, the neural network is used to model the inverse relationship based on the same data as the previous one. The network is tuned in order to minimize the error between predicted voltage and actual voltage. The maximum error is found to be 15% of the total input stroke. In the sequel by Song, Batur and Chaudary [9], the control proposed is of the form depicted in Equation 2.5:

$$i = i_{NN} + i_f - k_D r - \rho \tanh(ar) \qquad (2.5)$$

The purpose of each of the terms on the right hand side is as follows:

- i_{NN} is the neural network inverse controller, designed to perform the control action of significantly reducing the hysteresis (however, not exact cancellation due to the uncertainties not accounted for in the NN modeling)
- i_f is a feed forward current defined according to the first order model of an actuator with bias spring.
- $k_D r$ functions as a PD control with the r defined as $r = \dot{e} + \lambda e$ with λ a positive constant and e is the error.
- ρtanh(ar) is a sliding mode based robust compensator for hysteresis compensation as well as to increase the control accuracy and stability.

In their work, the error was found to be less than 2% on assessing the performance of the controller for tracking sinusoidal signals for a range of frequencies ranging from 1/60 Hz to 1/15 Hz.

2.6.2 Summary of model-based control

Based on the description provided of the various model-based control techniques, the following key aspects can be retained for future controller designs.

Adaptability

Including some intelligence and adaptation in the SMA model used for control allows the controller to compensate for eventual changes in material properties and/or operating conditions.

Model-based feed forward

All of the approaches reviewed use some information about the system being controlled in the derivation of control signals. Typically, an inverse plant model (possibly adaptive) is included in the feed forward path of the controller, combined with a linear feedback controller to compensate for modelling uncertainties. The effectiveness of this approach is demonstrated in the work of Majima *et al.* [33], which compares the response with and without the model-based feed forward term.

The neural network approach comes across as a simple alternative to the other models. However, it has certain hidden drawbacks:

- Inability to model the minor loop behaviour. Since for this property, the history of the system requires to be stored somewhere, a more complicated NN with memory elements will be required.
- There is no systematic procedure to determine the appropriate structure of the NN in terms of the number of layers and number of neurons. It is based on trial and error, which is a cumbersome procedure with no clear cut way to show stability.

We have done preliminary work on the use of the NN for modeling. The procedure followed and the results can be found in [36]. However the neural network based modeling was not pursued further due to the above mentioned drawbacks and also the large modeling errors observed.

2.6.3 Non-model based control techniques

There have also been some significant works done on control of SMA actuators using non-model based controller techniques. Two important techniques are briefly described here. In the work by Grant and Hayward [10], variable structure control has been proposed to control a novel shape memory alloy actuator consisting of several thin NiTi fibers. Variable structure control can be thought of as an algorithm that switches between a set of predefined controllers as the state of the system varies. This is equivalent to having several sets of controller gains which are indexed in a look-up table, based upon some measured system condition.

Another significant work was done by Selden, Cho and Asada [6]. In this work, a new approach is presented both for the design and control of SMA actuators. The SMA wire is divided into smaller segments. Each segment is controlled individually in the binary mode with two temperature thresholds, one hot and one cold. Peltier elements in contact with each wire segment are used for heating. The approach used here is of digital control as opposed to the analog control methods used previously. The proposed scheme demonstrates an improvement in the response time and a reduction

in the power consumption. This approach is interesting in some applications but the additional complexity of adding individual Peltier heating devices reduces its applicability to a range of applications.

2.7 Techniques to improve the response time

In the previous sections, a review was provided of the control approaches used by various researchers to control SMA actuators. These approaches have been geared towards addressing the problem of hysteretic behaviour of SMAs. In this section, a review is provided of the various approaches employed by researchers to improve the response time of SMA actuators.

The response time (or speed) of SMA actuators can be decreased by using techniques for faster heating or faster cooling. Examples of some techniques for faster cooling are: forced air and forced water cooling. In order to increase the speed of heating, higher currents can be applied. However, care should be taken to ensure that the wire is not overheated.

There are manufacturer specified safe currents (MSSC) for SMA wire, which are currents that can be supplied to the wire for indefinite periods of time without damaging the wire. Most present control techniques are designed such that the maximum current into the wire is limited at the MSSC. This limit on the current is conservative. Since it is actually the temperature which damages the wire, it should be possible to supply much higher currents intelligently, such that the current is shut off or decreased before the temperature goes above the safe temperature.

There have been some approaches proposed in the past by researchers [37, 38, 39, 40] to potentially supply much higher currents without damaging the wire. The first category of these approaches have been designed such that temperature is directly monitored using a thermocouple or other temperature sensor. In the latter category, other more easily monitorable quantities like resistance have been used to switch between currents.

In the work of Kuribayashi [37], originally a push-pull type of NiTi SMA actuator is controlled using PWM driver. The height of the pulse (corresponding to the applied voltage) using PWM is maintained at a conservatively lower value to avoid the SMA wire from overheating. Hence, a technique is proposed to limit the temperature of the wire instead of the voltage by directly monitoring it using a Cu-Constantan thermocouple. A clear improvement was seen in the performance using this technique. Also, on combining the thermocouple with ventilation techniques, an improvement was achieved in the overall response for heating as well as cooling.

Thermocouples are attached to the SMA wire. This is problematic for a few reasons – Firstly; it is difficult to attach them to thin SMA wire hence, larger diameter SMA wires are required. Also, achieving electrical isolation between the SMA wire and thermocouple is difficult, and the need for electrical isolation and thermal conductivity counter each other. Lastly, the constant flexing movement of the SMA wire makes the attachment of thermocouple more challenging. Hence, Russell and Gorbet [38] in their work proposed the use of an infrared temperature sensor to monitor the temperature of the SMA wire actuator. Also in this work, a novel actuator design is proposed such that the mobile heat sink is attached to the actuator for faster cooling. The most effective cooling techniques employed for cooling in SMA wires are water immersion and fixed heat sink. These techniques do help in faster cooling but also the power consumption on their usage is increased substantially due to the heat lost to the cooling medium. The idea of the mobile heat sink is proposed as a solution to this problem.

Although the temperature sensor is non-contact, there is the additional hardware cost of having these temperature monitoring devices as a part of the system. This motivated the development of techniques where a quantity which can be easily monitored, such as resistance, is used to switch between the currents. The major works using this technique have been by Featherstone and Teh [39, 40] where resistance feedback is used to switch between the currents. This scheme does away with the requirement for the temperature sensors. However, the accuracy is compromised since the exact temperature is not determined based on the resistance and instead a threshold resistance is proposed for switching between two values of currents as explained below.

The schemes used in [39] have been briefly described below and illustrated in Figure 2-11 and Figure 2-12. The scheme utilizes the property that the resistance is different in the austenite and martensite phases. There is hysteresis between the resistance and temperature as illustrated in Figure 2-11, and the exact temperature cannot be uniquely determined with a single resistance measurement. It is possible though to determine the temperature if the history of the resistance measurements exists. However, R_{thres} is proposed as the boundary between the 'safe state' and 'possibly unsafe state'. Using R_{thres}, a scheme is proposed to switch between the I_{safe} (MSSC) and I_{high} (the maximum value of the rapid heating current, possibly limited by the power supply ratings), also rather than having a discontinuity between I_{safe} and I_{high}, a smooth transition is proposed by using a ramp. This is depicted mathematically in Equation 2.6.

$$I_{\max}(R) = \begin{cases} I_{high} & R \geq R_{ramp} \\ I_{safe} + \left(\dfrac{I_{high} - I_{safe}}{R_{ramp} - R_{thres}} \right)(R - R_{thres}) & R_{thres} < R \leq R_{ramp} \\ I_{safe} & R < R_{thres} \end{cases} \quad (2.6)$$

Also once this value is calculated, the appropriate heating current (I_h) is calculated using the Equation 2.7, where I_d is the current from the controller.

$$I_h = \min(I_d, I_{\max}(R_{meas})) \quad (2.7)$$

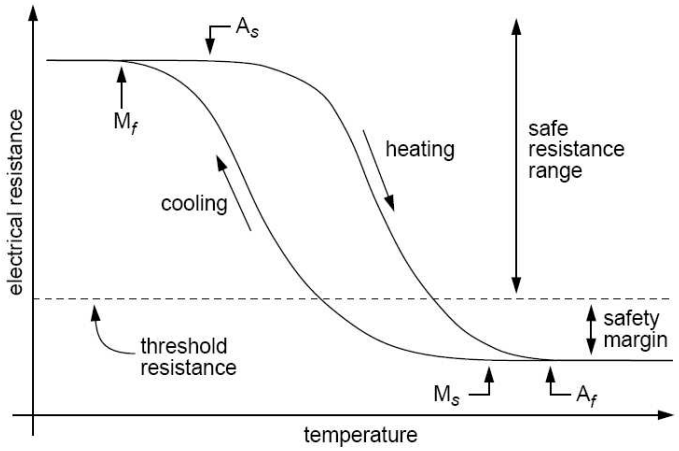

Figure 2-11: Plot of resistance versus temperature for Nitinol [39]

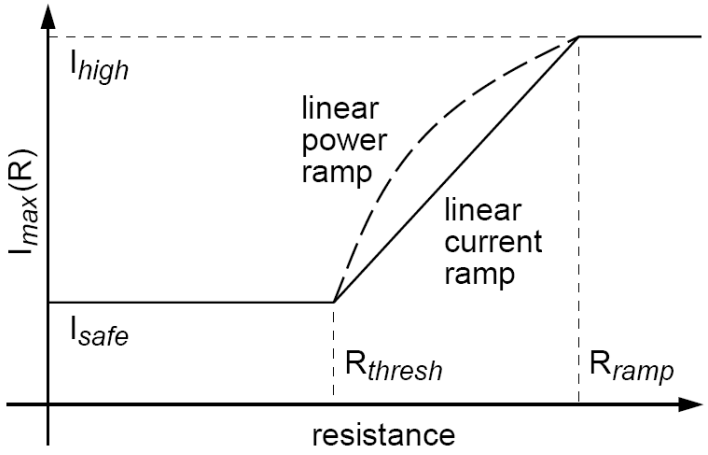

Figure 2-12: Scheme for current control based on resistance in [39]

In the later work [40], instead of current, power is saturated based on the scheme. In these works, there is no modeling involved; a value of R_{thres} is identified experimentally. This value of R_{thres} is subject to changes due to variations in the applied stresses. Although a safety margin is added to the R_{thres} value, the accuracy is compromised. Hence, our goal is to further improve upon these proposed schemes by using an accurate TEM to predict the temperature directly based on the current and the phase fractions (and stress) and accordingly ensure the temperature is not high enough to overheat the wire.

2.8 Summary of the Chapter

This chapter provides the background to the thesis. The chapter starts with a description regarding the behavioural aspects of SMAs. This is followed by background regarding SMA models. It is concluded that the John Shaw model is a suitable choice of model from the perspective of control applications. Subsequently, the various control techniques for SMA actuation are reviewed. Also, based on the previously employed schemes to improve the response time of SMA actuators, the proposed scheme is introduced briefly.

Chapter 3
John Shaw Model: Theory and Parameter Estimation

Over the past few decades, several SMA constitutive models of varying complexity have been developed. The present goals of SMA researchers have been towards developing models which can capture various SMA behaviors. In order to achieve the goals of higher accuracy and rigor, more complex models have been developed. The term "rigor" is used in this context to denote the thoroughness of the model in demonstrating actual SMA behavior. For instance: a model which demonstrates rate dependence or minor loop behavior is more rigorous than a model that does not. However, this requirement for greater accuracy leads to the development of fairly complicated models, typically involving one or more partial differential equation forms [24].

The original John Shaw model was in a computationally intensive form [24, 25]. The governing equations were in the form of partial differential equations. The complexity of this form precluded its applicability to control applications, particularly real-time. By making some simplifications, the model was converted to an ODE form. This renders it more amenable for real-time control applications.

This chapter is organized as follows: The theory of the John Shaw model is provided in the first section. The original full-order PDE model is briefly introduced, followed by a description of the reduced-order ODE model. This sections ends with a discussion of the limitations of the John Shaw model.

The algorithm proposed for extending the ability of the John Shaw model to display minor loops is described in Section 3.3. The experimental set-up used for the parameter estimation is described in Section 3.4. This is followed by a description of the parameter estimation procedure.

The model is developed for a thin NiTi wire under uniaxial tensile loading conditions. The wire set-up is considered to be immersed in a convective thermal environment. The scope is restricted to

macroscopic evolution during phase transformations as the model does not include the microstructure evolution during the phase transformations. The state of each point (x) along the SMA wire is determined by the strain field $\varepsilon(x,t)$, the temperature field $T(x,t)$ and the internal phase field vector $\xi(x,t)=\{\xi_1, \xi_2\}$ for tensile and compressive variants of martensite respectively. The constitutive equations of the John Shaw model were derived from the specific Helmholtz free energy, described in greater detail in [24]. In the next section, the assumptions made by John Shaw *et al.* to arrive at the ODE version of the John Shaw model [25] are presented, followed by a discussion of the individual equations of the model.

3.1.1 Assumptions and Reduced-order governing equations

The assumptions made to simplify the system of partial differential equations to ordinary differential equations are presented below:

- The values of the strain, temperature and phase fractions are considered to be uniform along the length of the wire. This assumption is reasonable if the SMA wire has already been conditioned to display repeatable cyclic behavior and also if the wire is thermally insulated at its ends.
- The SMA element is always under sufficient tension, so as to avoid thermal martensite. This assumption is valid if the applied stress is always greater than minimum load required for detwinning martensite. In position control applications, this is a reasonable condition to satisfy since there will always be an applied stress which "resets" the actuator so that it can be cycled.
- The interaction energy (also called mixing energy) between austenite and martensite is neglected. Very limited explanation has been provided regarding the interaction energy by John Shaw *et al.* and other authors in the context of SMAs. By generalizing based on the general definition of interaction energy, it is the difference between the energies of the phases in isolation and their combined energy. The authors state that during superelastic transformation, any hardening and softening is neglected.
- The elastic moduli of austenite and martensite are the same ($\Delta E=0$). The elastic moduli of austenite and martensite are denoted as E, used in place of E_A and E_M. This assumption is certainly not justified for the values of E_A and E_M identified for our set-up. However this assumption was made on the premise of restricting the operation of the actuator in certain regimes. It is stated that as long the martensite is not unloaded during operation; this assumption does not pose significant restrictions [25]. Since it is expected that a certain stress, greater in value than the detwinning stress (172 MPa) is always applied, this should not cause major issues.

However in order to get a sense of the implications of this assumption, some analysis is done on the impact of the equations when ΔE is non-zero. This is presented in Appendix A, where it is demonstrated that the maximum error induced by assuming $\Delta E=0$ is approximately 2%. For simulation purposes, it has been assumed that $\Delta E=0$.

- The electrical power P_e and the ambient temperature T_a are chosen as piecewise constant functions of time.
- The characteristic speed of phase transformation is fast compared to the heat transfer rate and mechanical loading rate. This is a reasonable assumption for slow to moderate loading rates (for Nitinol, this is valid when the elongation rate (denoted by $\dot{\delta}$) $<10^2$ m s^{-1}). Basically this is valid as long the inherent velocities of the martensitic transformations are large compared to the displacement rates.

Table 1: List of variables used in the John Shaw model

Symbol	**Description**
L	Length
d	Diameter
β	Transformation Strain
E	Elastic Modulus
T_R	Reference Temperature
$T_R\Delta s$	Predicted enthalpy change
ρ	Mass density
μ_c	Critical driving force
v_o	Kinetic law stiffness
ρ_e	Electrical resistivity
h	Heat transfer co-efficient
P_e	Power

ε	Strain
ξ	Phase fraction of martensite
s_o	Specific entropy
c_I	Phase interaction parameter
T_a	Ambient temperature
γ	Strain gradient parameter
c_p	Specific heat

Simplified governing equations

The state of the wire in the reduced-order form is defined by three time-dependent variables, the strain $\varepsilon(t)$, the temperature $T(t)$ and the single phase fraction $\xi(t)$ which represents tensile martensite.

In this section, the governing equations of the simplified model are presented and explained. The governing equations for the actuator's operation are algebraic equations for equilibrium and phase transformation, and an ordinary differential equation in time for the heat transfer.

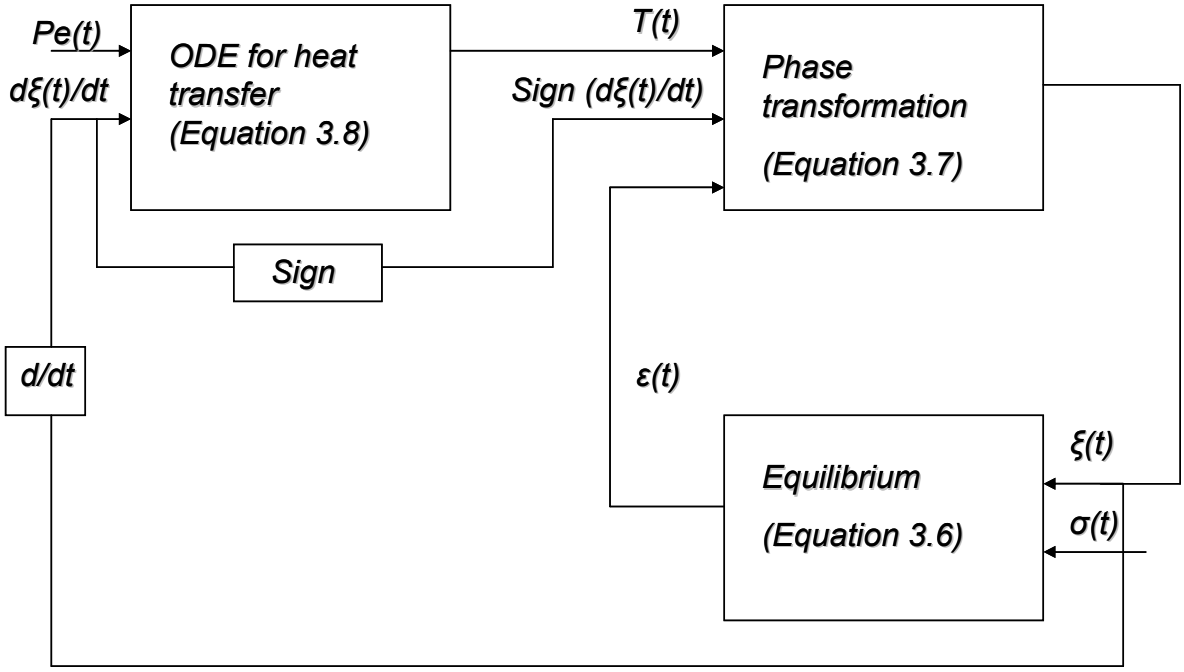

Figure 3-1: Depiction of the John Shaw model

The equilibrium equation is developed on the premise that the inertial effects are ignored. For the axial force (*F(t))* in the SMA wire to be uniform, it is balanced by the externally applied force *(F_{app}(t))* (Equation (3.6)).

$$F(t) = F_{app}(t) = EA(\varepsilon - \beta\xi)$$
$$F(t) = \sigma(t)A \tag{3.6}$$

In order for the phase transformation from austenite to martensite to occur, two conditions need to be satisfied.

- Thermodynamic driving force has reached its critical value μ_c.
- The amount of tensile martensite has not yet reached saturation *($0 \leq \xi < 1$)*

Conversely, for the reverse transformation from martensite to austensite to occur, two conditions must be satisfied.

- Thermodynamic driving force has reached its critical value $-\mu_c$.
- The amount of austenite has not yet reached saturation *($0 < \xi \leq 1$)*

If neither of these sets of conditions is satisfied, no phase transformation occurs. During phase transformation, the stress and temperature are coupled through Equation 3.7.

$$\frac{\beta\sigma}{\rho} + (T - T_R)\Delta s = \mu_c \operatorname{sgn}(\dot{\xi})$$
$$\frac{\beta \mathrm{E}(\varepsilon - \beta\xi)}{\rho} + (T - T_R)\Delta s = \mu_c \operatorname{sgn}(\dot{\xi}) \tag{3.7}$$

The heat equation is represented by Equation 3.8. The term on the left hand side is the rate of the energy associated with sensible heat, whereas the first term on the right hand side is the latent heat source, the second term is the heat loss rate to the environment and the last term is the applied power P_e.

$$\rho A L c_0 \dot{T} = \rho A L[\operatorname{sgn}(\dot{\xi})\mu_c - T\Delta s]\dot{\xi} - h\pi d L(T - T_a) + P_e \tag{3.8}$$

The John Shaw model is depicted diagrammatically in the Figure 3-1 . The heat equation inputs are the applied power and the rate of evolution of phase fraction, and the output is temperature. The heat equation module corresponds to Equation 3.8.

The input to the phase fraction module is temperature, strain and the rate of change of phase fraction and the output is phase fraction. The phase transformation module corresponds to Equation 3.7. The input to the equilibrium module is the phase fraction of the martensite and the applied stress. The equilibrium module corresponds to Equation 3.6.

In this model, since constant stress is applied throughout, the influence of stress on the heating module and phase kinetics is eliminated. Hence one would observe that in Figure 2-8) in Chapter 2, stress is depicted as an input to each of the modules of the SMA model. However for the representation of the John Shaw model (Figure 3-1), stress is only depicted as input to the equilibrium module.

From a high-level, the inputs to the John Shaw model are applied power and stress and the output is strain. Since in the typical SMA model usage as a plant, the input is usually current, an additional block for current to power conversion is needed. In this block the relationship *(Power* $=I^2*R$*)* is used. Also the resistance of SMAs varies between the austenite and martensite phases. In order to accurately determine the current, a phase fraction feedback is needed from the model to determine the resistance as function of the phase fraction.

3.1.2 Limitations of the reduced-order John Shaw model

There are certain limitations to the reduced-order form of the model which limit its widespread applicability to control applications. Firstly, the transformation kinetics equation does not account for the fact that phase fractions only attain values between 0 and 1 and saturate beyond those values. Hence when the model is implemented, the saturation has to be explicitly programmed. Another significant limitation of the model is the inability to model minor loops. Modelling minor loop behaviour is particularly important in closed-loop strain (or position) control applications since achieving a specific target strain between the two (load-dependent) extremes requires partial phase transformation.

3.2 Minor loops extension

As briefly mentioned above, the John Shaw model lacks the ability to model minor loops. In this section, two possible approaches to extend the ability of the model to display minor loop behavior are

discussed. The first approach is based on the work developed by Madill *et al.* [41] in which the authors extend the Ikuta model [42] to display minor loop behavior. This work is the only one (to the best of my knowledge) where the original equations of the SMA model were modified to account for minor loop behavior. Hence the procedure followed by Madill *et al.* is described here.

The main steps of the procedure developed by Madill *et al.* are as follows:

- The phase kinetics equation is converted from a function of form $\xi(T)$ (phase fraction as a function of temperature) to a function of the form $\xi(t,T)$ (phase fraction as a function of time and temperature), so that the history dependence is captured as well.
- Hence the constants in the previous form of the original phase fraction equations in Madill *et al.*'s model are now piecewise constant functions of time. Their values change every time the directions of the temperature signal changes.
- In order to determine these constants when switching takes place, certain conditions are utilized. For switching from heating to cooling, the continuity condition and the common origin conditions are used.
- *Continuity condition:* According to this condition, there cannot be an instantaneous change in the martensite fraction. Hence the value of the phase fraction on the heating curves at the switching time t_s equals the value of the phase fraction on the cooling curves at the same time (Equation 3.9). In this equation, $\xi^{C \text{ or } H}(t,T)$ refer to cooling and heating equations respectively.

$$\xi^{H}(t_s, T_s) = \xi^{C}(t_s, T_s) \tag{3.9}$$

- *Common origin condition:* According to this condition, the strain is fully recovered after heating the wire and cooling it down again to ambient temperature. The name arises since this condition enforces a closed major hysteresis loop. However, this condition only applies to situations where the stress is maintained at a constant level.
- Similarly when the switching takes place from cooling to heating, the continuity equation and the common limit equations are used.

In order to apply this procedure to the John Shaw model, there is a limitation with the model itself, which needs to be resolved. The phase kinetics equation does not account for upper and lower saturation limits at 1 and 0 respectively. For implementation purposes, the saturation can be explicitly

programmed. However in order to convert it to a form which can handle minor loops, a closed form of the phase kinetics equation is required which automatically limits at 0 and 1 respectively.

Given the above-mentioned issue, an alternative approach was proposed to model minor loop behavior.

Once programmed, the functioning is similar to that of the backlash operator. The procedure is composed of the following steps:

- Assume that at the start of the simulation/trial/experiment, we start on the major hysteresis loop.
- The most likely start temperatures are either the ambient temperatures (below M_f for our wire) or temperatures above A_f(austenite finish temperature).
- Depending on whether the phase fraction of martensite is increasing or decreasing, choose the cooling equation or the heating equation respectively (Equation 3.7).
- If switching takes place from cooling to heating, hold the value of the phase fraction at the value ($\xi^C(t_s)$)calculated according to the cooling equation. Meanwhile, as simulation time progresses, calculate the values of the phase fractions according to the heating equation ($\xi^H(t)$). At each time instant, compare $\xi^C(t_s)$ with $\xi^H(t)$. If the values are equal, the heating equation is valid and active. Until switching occurs again, ξ is calculated according to the heating equation.
- Similarly, when switching takes place from heating to cooling, hold the value of the phase fraction at the value ($\xi^H(t_s)$) calculated according to the heating equation. Meanwhile, as simulation time progresses, calculate the values of the phase fractions according to the cooling equation ($\xi^C(t)$). At each time instant, compare $\xi^H(t_s)$ with $\xi^C(t)$. If the values are equal, the cooling equation is valid and active. Until switching occurs again, ξ is calculated according to the cooling equation. Since an exact match does not always occur, a certain tolerance was used while detecting whether the phase fraction values match.
- Also at each time step, if the value of ξ is >1 or <0, it is saturated at 1 and 0 respectively.

3.3 Experimental Set-up

In the previous section, the theory and limitations of the John Shaw model were presented, along with an algorithm for extending the existing reduced-order John Shaw model to include minor loop behavior. Prior to software implementation of the extended model, the parameters required to build the model has to be estimated. In this section the experimental set-up is presented, followed by the

parameter estimation procedure. The parameters required to build the model alongside their physical meaning is provided in Section 3.4.

The experimental set-up used for the purposes of estimating the parameters is described in this section. The set-up is illustrated in Figure 3-2. At a high level the functioning of the set-up can be understood as follows: The SMA wire is connected along the XY-positioning stage. The elongation/contraction of the wire is monitored by the position encoder. Force is applied to the wire via the force actuator, which is powered by the current amplifier. The applied force is monitored via the force sensor. Current is supplied to the SMA wire via a current-controlled voltage supply. Software-based control of the setup (via real-time workshop in MATLAB) is done by the computer and connected to the voltage power supply and the current amplifier. In this set-up, current and force are controlled and position, force and voltage are measured.

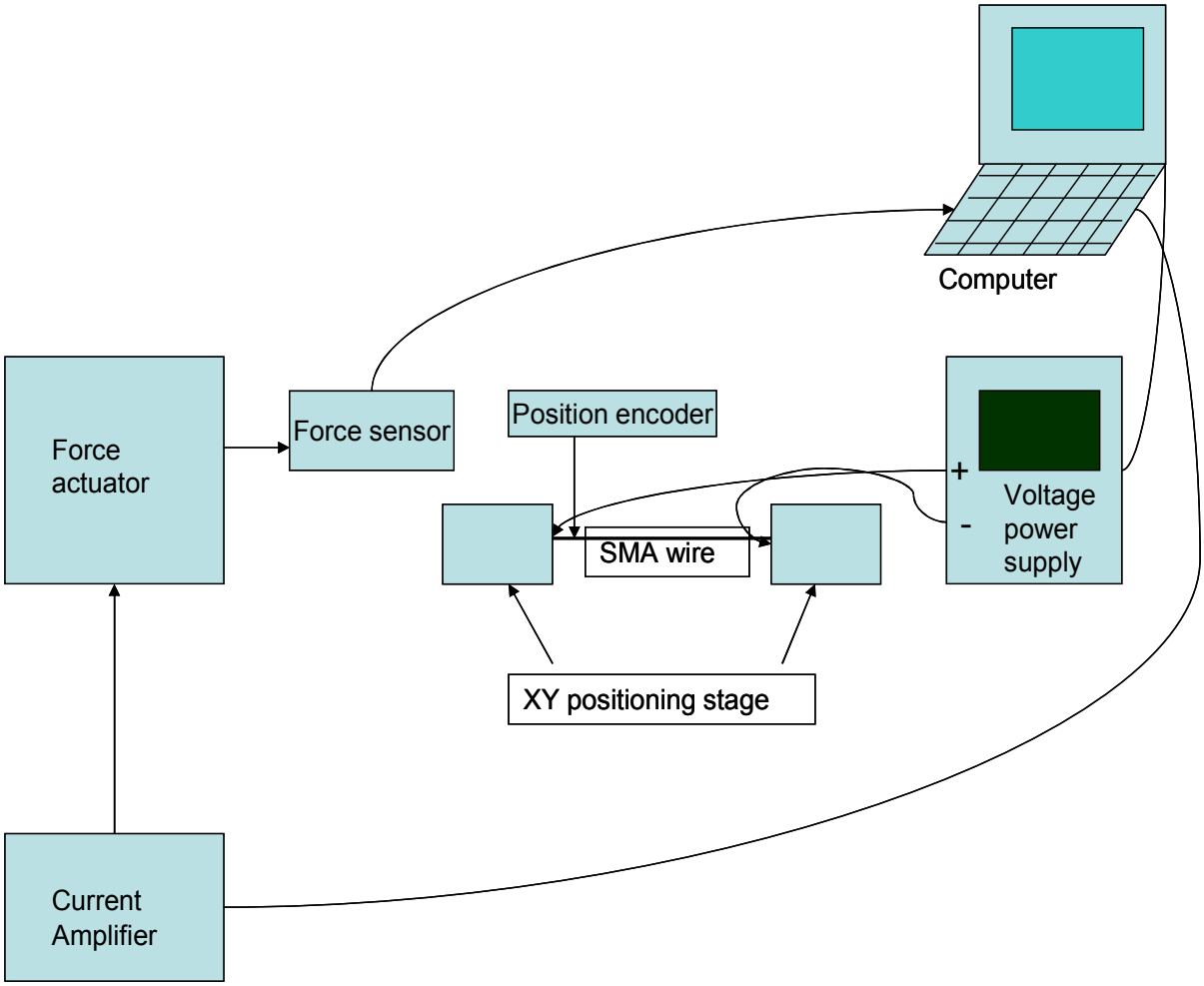

Figure 3-2: A high-level illustration of the experimental set-up in our lab

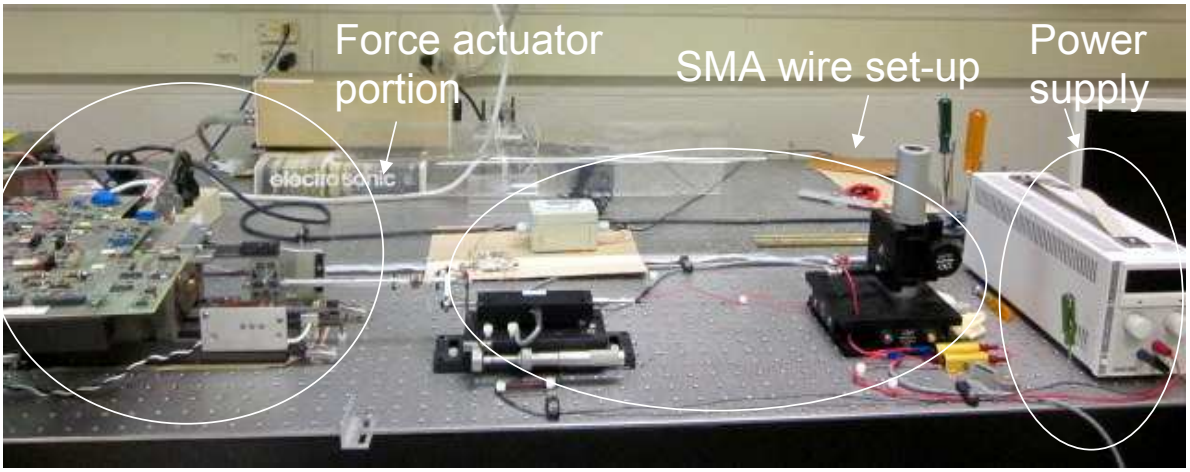

Figure 3-3: Overview of the actual set-up

More description of the various portions of the set-up along with their actual photographs is provided here. The photograph of the complete set-up is found in Figure 3-3.

a) Force actuator portion (includes current amplifier, force actuator, force sensor)

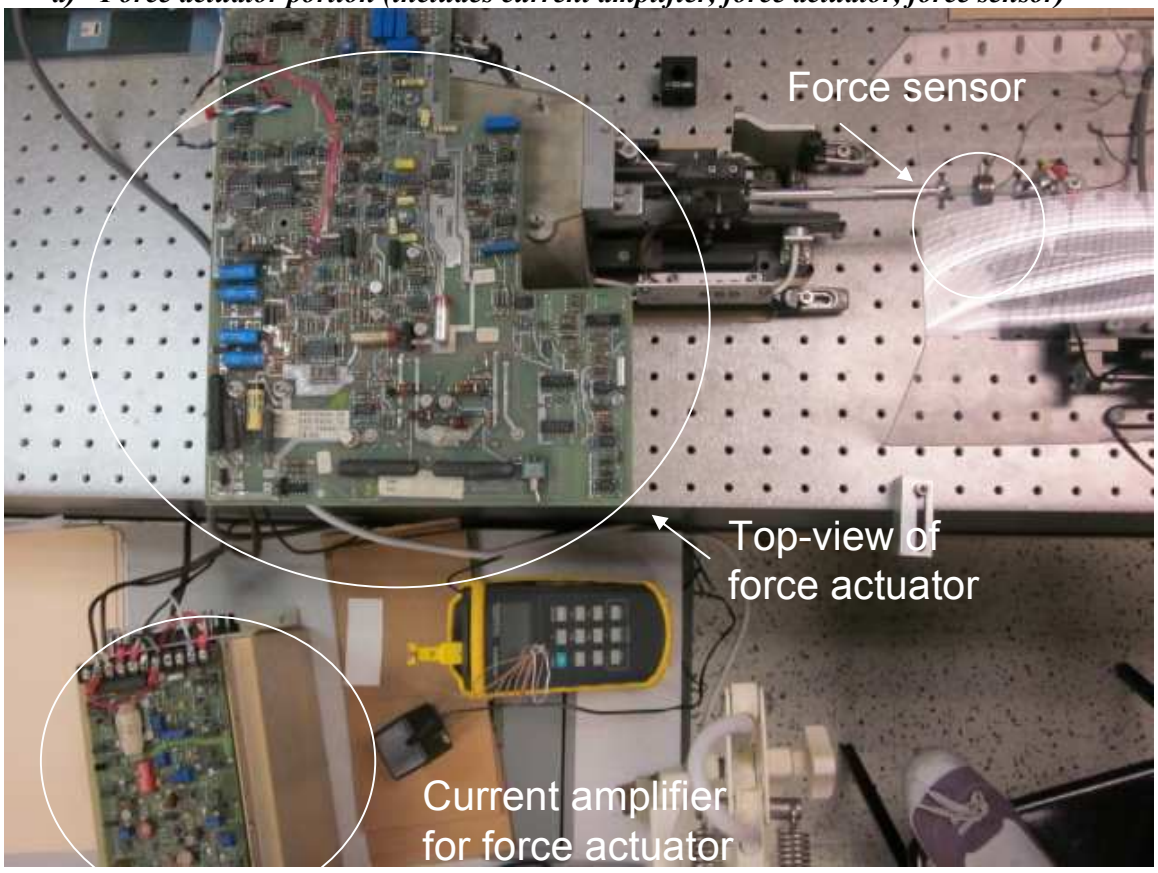

Figure 3-4: Depiction of the force actuator portion

The force actuator applies the force to the wire. The force actuator is powered by the current amplifier labeled in the figure (EG&G torque systems model C0502-001). The applied force from the force actuator is measured by the force sensor. A feedback controller is designed around the force supplied by the force actuator as illustrated in Figure 3-5. Firstly, the error is calculated between the force measured by the force sensor and the software-commanded force. This error is fed into the PI controller. The output is multiplied by the force actuator calibration and current amplifier gains respectively. Also the maximum input into the force actuator is set at 4 A. This is considering that the maximum output from the DAQ card is 5V and the current amplifier gain is 0.8. The output voltage

from the computer (supplied to the current amplifier) is again multiplied by the load cell (or force sensor) calibration gain to calculate the actual force supplied to the SMA set-up by the force actuator.

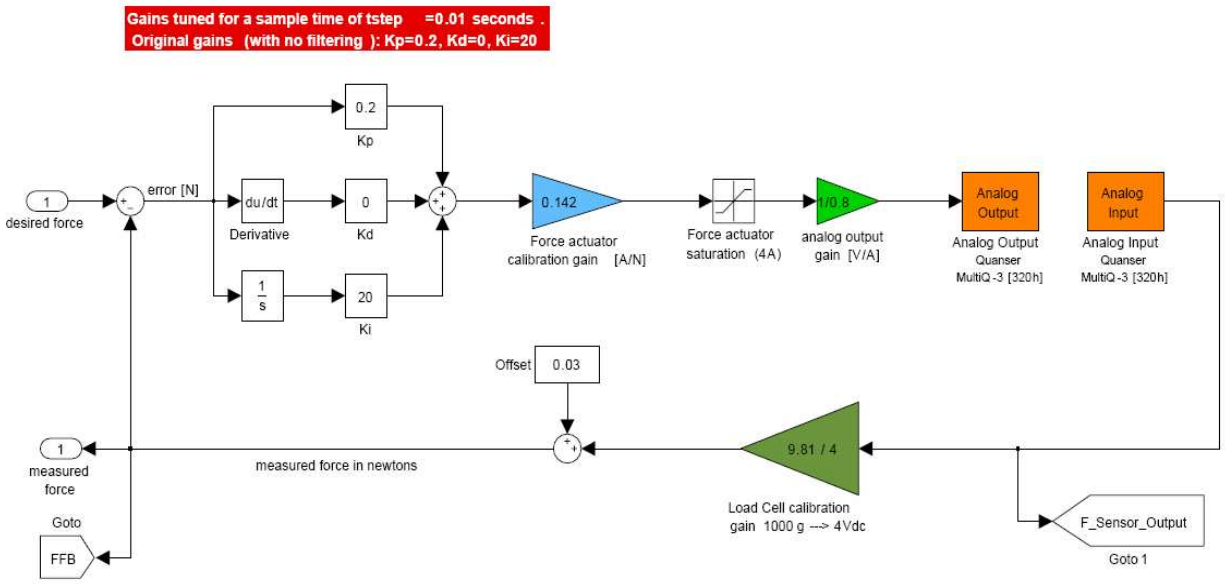

Figure 3-5: SIMULINK set-up of the force controller

b) SMA wire set-up

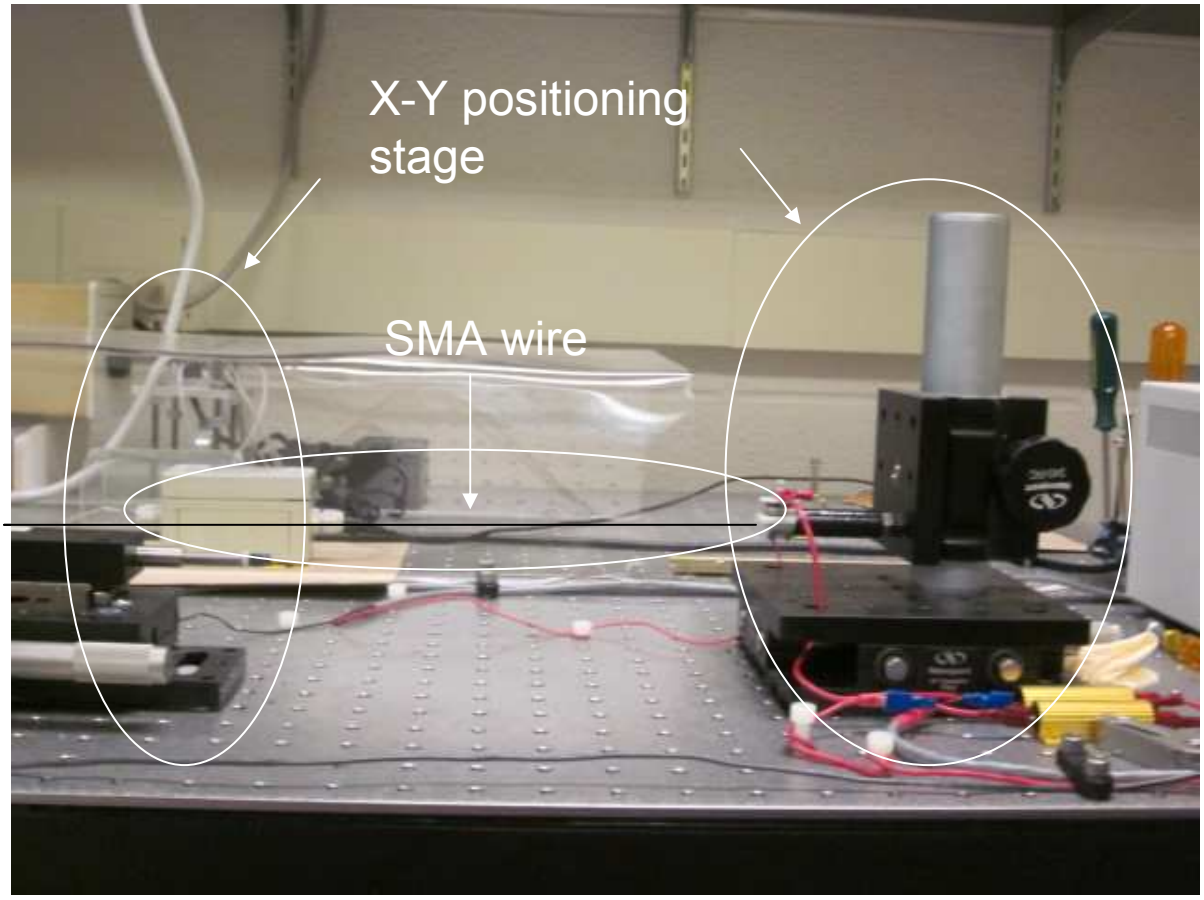

Figure 3-6: SMA wire set-up

The SMA wire set-up portion (illustrated in Figure 3-6) includes the SMA wire (darkened in the figure to emphasize location) held in position by means of an XY-positioning stage. The US digital encoder monitors the contraction/expansion of the SMA wire. In order to limit air currents to make *h* consistent since it is assumed to be constant, the setup is protected by a Plexiglass® covering.

c) Power supply and current controller

The current required to heat the wire is supplied by the Xantrex power supply (Figure 3-7). The resistance circuit is used as a voltage divider, to scale the SMA voltages to values that better match the ADC input range.

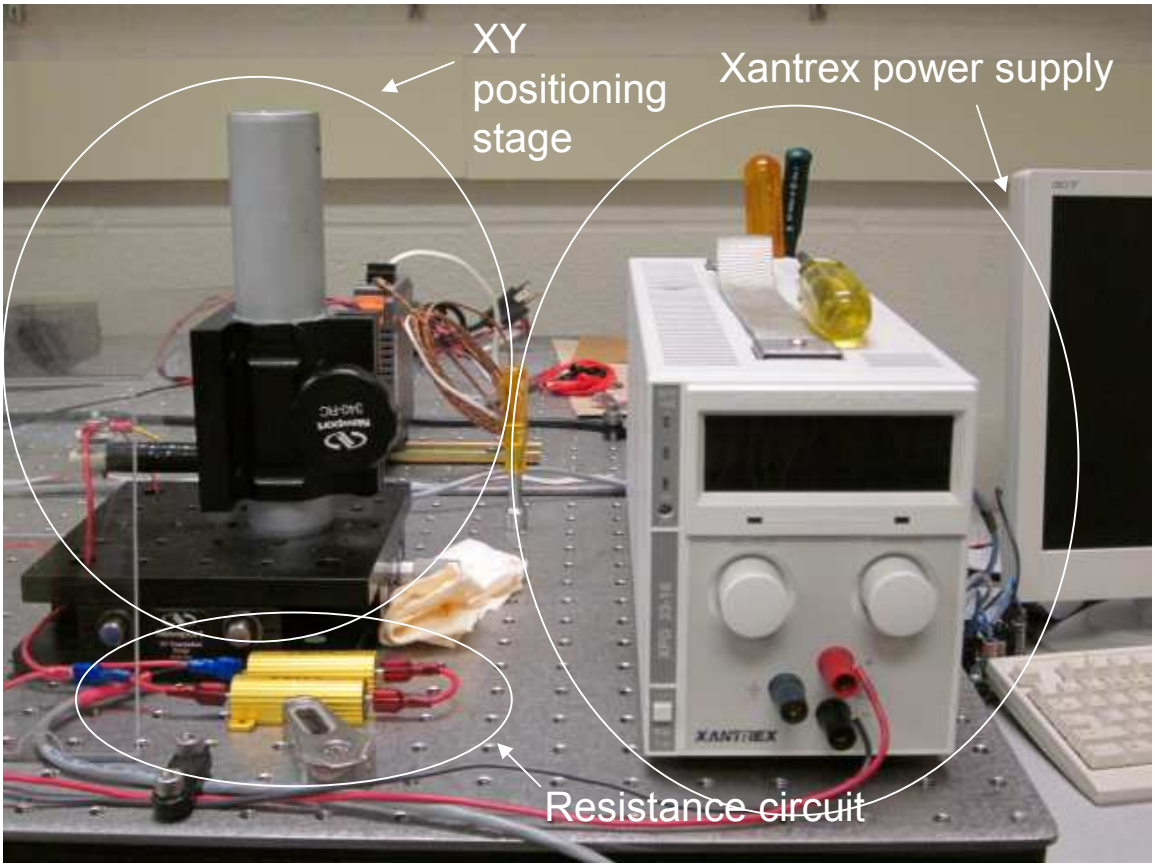

Figure 3-7: Xantrex power supply

c) DAQ break out board

The function of the DAQ breakout board is to make connections from the experimental set-up to the computer and vice versa.

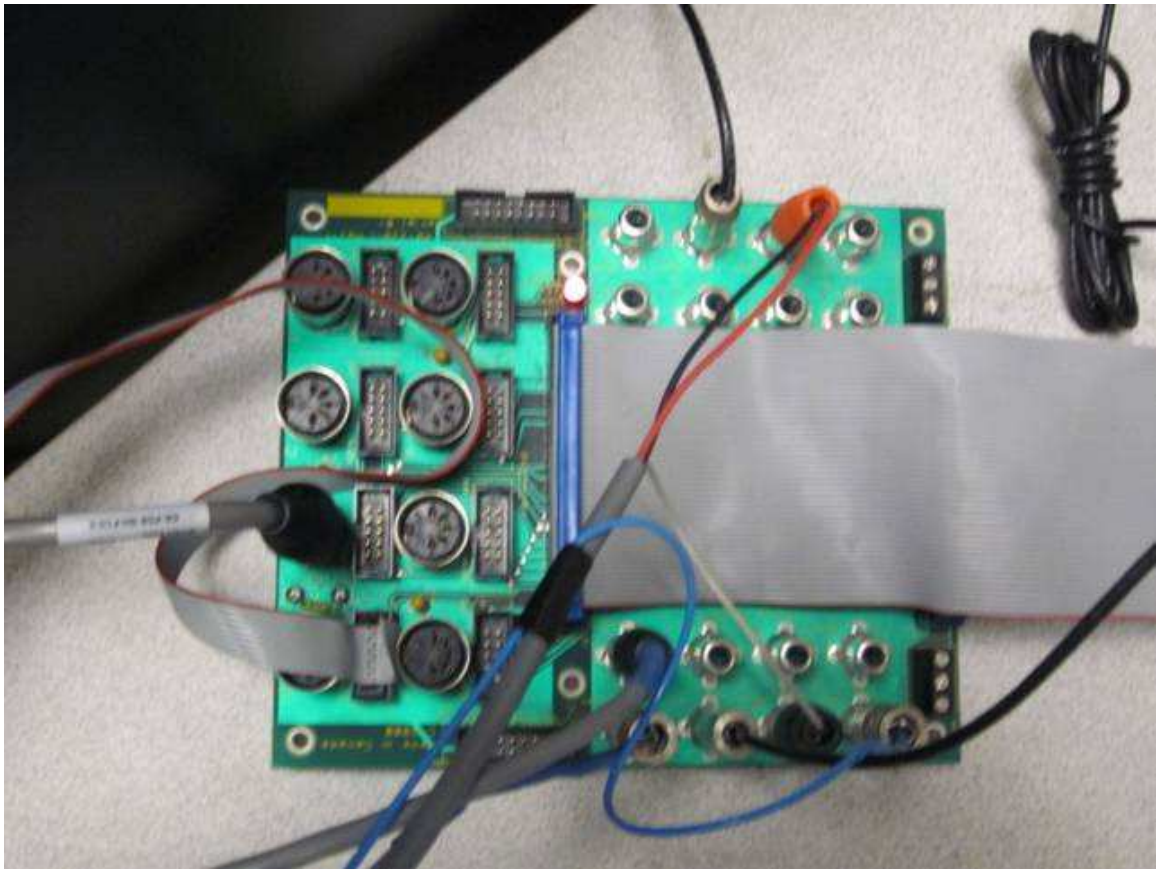

Figure 3-8: DAQ break-out board

3.4 Parameter Identification

In this section, the procedure followed for the estimation of parameters is described. The section is organized as follows: the parameters are listed, alongside their description. Also, important notes regarding the estimation of parameters are provided. This is followed by a description of the two procedures required for the estimation of parameters.

Table 2: List of parameters

Parameter Symbol	**Description**	**Value**	**Source**
L	Length	410 mm	*
d	Diameter	127 micron	*
β	Transformation Strain	0.0535	**

E	Elastic Modulus	48.23 GPa	*
T_R	Reference Temperature	314.3 K	*
$T_R \Delta s$	Predicted enthalpy change	-21.11 J/g	[25]
c_o	Specific Heat	0.5 J/gK	[25]
ρ	Mass density	6455000 g/m^3	[25]
μ_c	Critical driving force	1.01 J/g	[25]
ρ_e	Electrical resistivity	10^{-6} ohm-m	[25]
T_a	Ambient temperature	300 K	*
h	Convection co-efficient	90 W/m^{2}K	[25]
Pe	Power	User given	*

* measured empirically

** determined experimentally, procedure described in Section 3.4.2

Important notes about the estimation of parameters

- The length L, the diameter d, the mass density ρ, the electrical resistivity ρ_e', and h are pre-determined parameters and do not require any experimental validation.
- The ambient temperature T_a and applied input power P_e are user-defined inputs.
- T_R is the average of the stress-free transformation temperatures. In John Shaw *et al.* [24, 25], T_R is defined as the average of the austenite start and martensite start temperatures. It can be calculated using manufacturer specified data. Alternatively, it can be calculated from the differential scanning calorimetry (DSC) data [25]. DSC is a technique in which the difference in the amount of heat required to increase the temperature of a sample and reference is measured as a function of temperature. In this work, it is calculated as follows: From the Figure 3-9, A_s (austenite start temperature) and M_s (martensite start temperature) at 172 MPa can be determined. Given the stress-temperature coefficient is 0.23C/MPa, A_s and M_s at 0 MPa can be calculated. A_s was found to be 90C from the graph and M_s was approximately 72C.

$$A_s(0) = A_s(\sigma = 172) - 0.23\sigma$$
$$M_s(0) = M_s(\sigma = 172) - 0.23\sigma$$
$$T_R = \frac{A_s(0) + M_s(0)}{2} \tag{3.10}$$

- The stress-free transformation strain β and elastic moduli for austenite and martensite, respectively E_A and E_M can be determined from the stress-strain hysteresis curves. The procedure to do so is briefly described in section 3.5.1.
- The predicted enthalpy change $T_R\Delta s$ and the critical driving force μ_c are taken from John Shaw *et al.* [23]. Using the value of T_R calculated for 90C flexinol, the corresponding value of $T_R\Delta s$ is accordingly calculated.

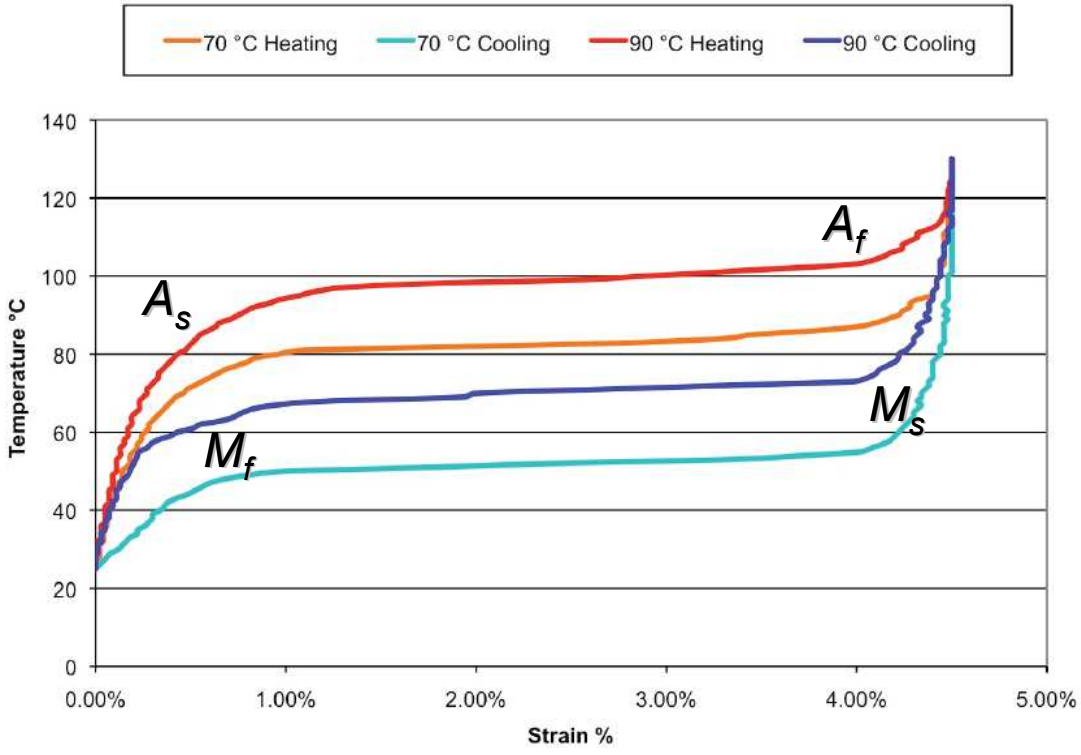

Figure 3-9: Strain temperature curves (Dynalloy) - 70C and 90C flexinol at 172 MPa [2]

3.4.1 Stress-strain hysteresis curves

As mentioned above, the elastic moduli and the transformation strain can be derived from the stress-strain hysteresis curves [12]. The procedure for the identification of parameters is illustrated in Figure 3-10.

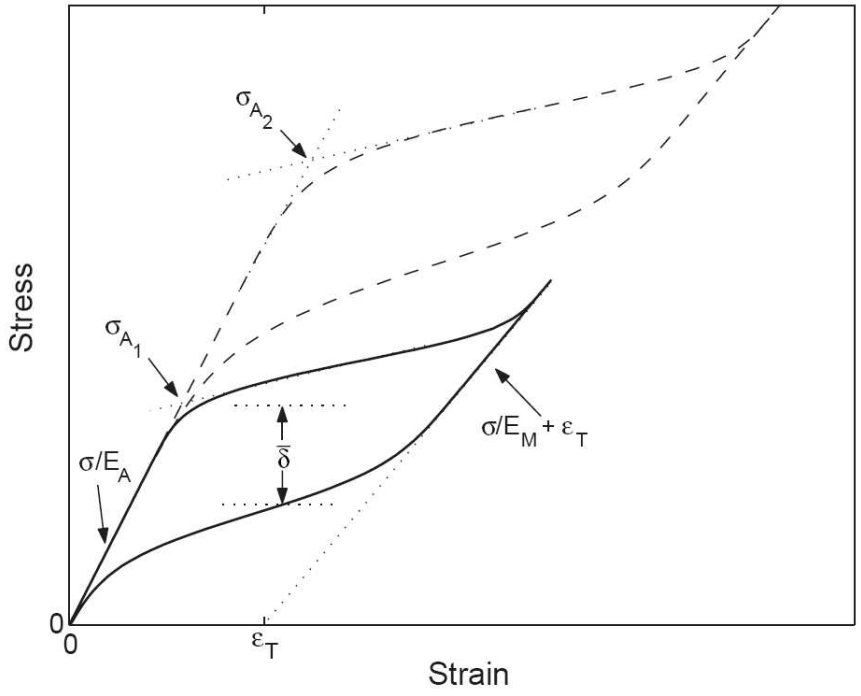

Figure 3-10: Identification of material parameters using stress-strain data (dashed hysteresis curve corresponds to higher ambient temperature) [12]

The linear portions of the loading and unloading curves can be used to derive the elastic moduli. In order to derive E_A, the slopes of the linear portions of the loading curves at sufficiently high temperatures (typically above A_f) have to be determined. E_M can be calculated either from the slopes of the loading superelastic curves or from the slopes of the unloading curves at room temperatures. The stress-free transformation strain β (depicted by ε_T in Figure 3-10) is the offset on the strain-axis. This is obtained by extrapolating the unloading portion of the stress-strain curves at low temperatures.

3.4.2 Procedure for deriving the stress-strain curves using the lab set-up

For the parameter identification, we require several isothermal stress-strain curves at various temperatures. We apply a triangular stress profile between 0 newtons and 10 newtons, and measure stress and strain. Since we cannot control or measure temperature directly in our setup due to the absence of temperature monitoring devices, we rely on the application of constant currents and assume that the steady-state temperature can be predicted by zeroing out the dynamics in the TEM (Equation 3.11). Thus, the current required to achieve a specific steady-state temperature T_{ss} is given by:

$$I_{ss}^{\ 2} = T_{ss}\left(\frac{R}{hA}\right) \tag{3.11}$$

The specific test protocol for a given T_{ss} is as follows:

- For 0-10 seconds, the Flexinol 127 μm 90C wire is heated at 0.3 ampere and 0 stress to attain its zero-stress austenite length.
- From 10-15 seconds, the current is ramped down to I_{ss}
- From 15-215 seconds, the force is ramped from 0 newtons to 10 newtons at 0.05 N/s maintaining a constant current at I_{ss}.
- From 215-415 seconds, the force is ramped down from 10 newtons to 0 newtons with no change in the current.
- The stress-strain pulls are carried out at 10 different equidistant temperatures corresponding to the following current values: 0.3 A, 0.283 A, 0.265 A, 0.245 A, 0.224 A, 0.2 A, 0.173 A, 0.1414 A and 0.1 A. Though currents are specified at this precision, the accuracy of the current supply is around 50 mA.

The stress-strain pulls are displayed in Figure 3-11.

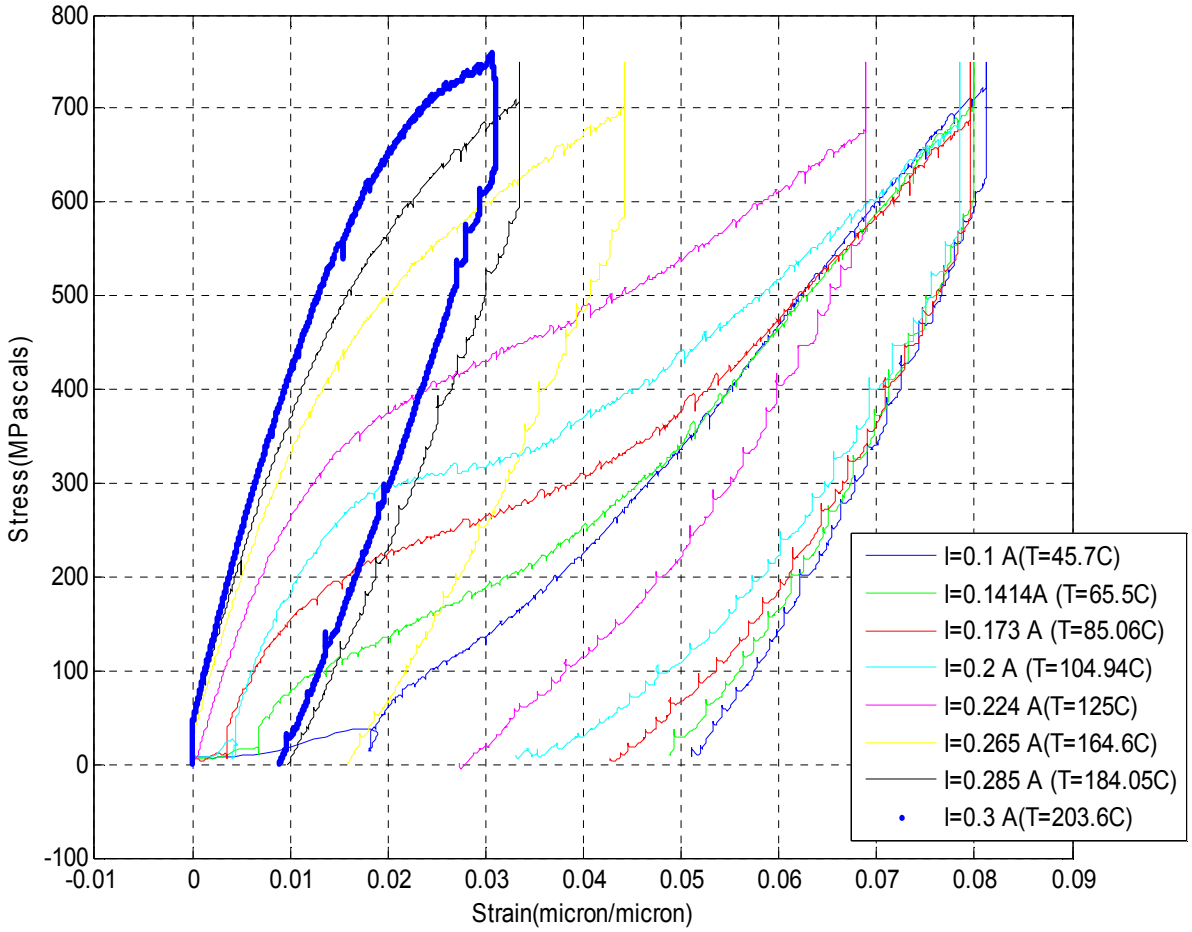

Figure 3-11: Experimental stress-strain curves for parameter estimation.

3.5 Summary of the Chapter

In this chapter, the theory regarding the John Shaw model is presented. Also the experimental set-up is presented, followed by the estimation of the parameters.

Chapter 4 John Shaw Model: Results and Discussion

4.1 Introduction

In the previous chapter, the John Shaw model was introduced and the estimation of parameters was discussed. In this chapter, the simulation results of the John Shaw model are presented along with a discussion of the simulation results. The minor loops algorithm which is a novel contribution of this work is tested thoroughly and the results are discussed. On the whole, a good match is obtained between the simulated behavior and what is expected experimentally.

4.2 Simulations

The techniques followed in order to test the John Shaw model are discussed in this section. Each of the governing equations or blocks of the John Shaw model (Figure 4-1) was implemented as a separate block for testing purposes (Figure 4-2, Figure 4-7 and Figure 4-12). The heat equation was implemented as a separate block with the applied power and the rate of evolution of phase fractions as inputs. The output of the block is temperature. The inputs to the transformation kinetics block are applied power, the rate of evolution of phase fractions and strain. The output is the phase fraction in the current step. The input to the equilibrium block is phase fraction and the output is strain. After the individual modules have been tested, the ability of the model to display minor loop behaviour is tested. All the simulations were carried out at a constant stress (σ) of 172 MPa.

Table 3: List of main simulation parameters

Simulation Parameter	**Value**
Stress *(σ)*	172 MPa
A_s (172 MPa)	90C (363 K)
E	48.23 MPa
β	0.0535

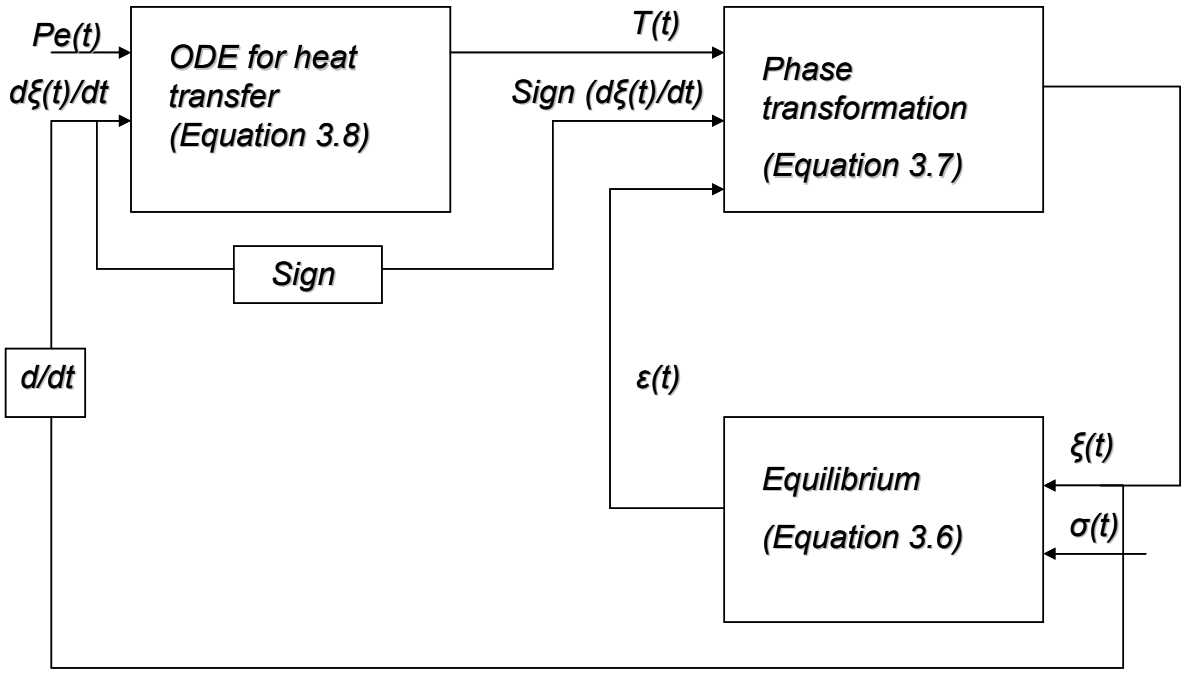

Figure 4-1: Block diagram of the John Shaw Model

4.2.1 Heat equation block tests

This block was tested using two tests cases – constant phase fraction and decreasing phase fraction. These two input signals were chosen to test:

- Whether the heat equation block works in term of accurately changing the temperature based on the power applied. It should be noted that the actual value of temperature cannot be determined since there is no technique to measure the temperature. However the actual equation is used to determine what the corresponding the steady state temperature value should be and it is compared with the values noted from the simulation.
- Whether the latent heat of transformation term is functional.

TEST CASE 1: Constant phase fraction (Phase fraction of martensite =1)

The phase fraction and temperature behaviours are illustrated in Figure 4-3 and Figure 4-4. The temperature behaviour is noted in response to fixed phase fractions (no transformation) and constant applied power. Qualitatively a good match is obtained, since there is a 0 change in temperature when

the applied power is 0, and also the temperatures reached are higher as the input powers are increased. The change in temperature is given by $P_e/h\pi dL$. This is derived in Equations (4.1 and 4.2).

$$\rho ALc_0 \dot{T} = \rho AL[\text{sgn}(\dot{\xi})\mu_c - T\Delta s]\dot{\xi} - h\pi dL(T - T_a) + P_e \tag{4.1}$$

At steady state, $\dot{T}$ and $\dot{\xi}$ terms are 0, hence the resulting steady state change in temperature is given by Equation 4.12.

$$\Delta T = T - T_a = \frac{P_e}{h\pi dL} \tag{4.2}$$

The expected values of ΔT calculated from Equation 4.2 are the same as those seen in Figure 4-4. For 0.25 watts, ΔT =17 Kelvin, for 0.5 watt, ΔT =34 Kelvin, for 0.75 watts, ΔT =51 Kelvin and for 1 watt, ΔT =68 Kelvin.

TEST CASE 2: Decreasing phase fraction (transforming from full martensite to full austenite)

The phase fraction and temperature behaviours are observed in Figure 4-5 and Figure 4-6. The temperature behaviour is noted in response to decreasing phase fractions and constant applied power.

In comparison to results in Figure 4-4, it is observed that the temperatures rises up to the same steady state values as seen for constant phase fraction case. However, after a certain time the temperatures start to decrease due to the latent heat absorbed for transformation from martensite to austenite. This is as expected, since the heat equation of the John Shaw model has a term to account for latent heat of transformation.

At this point, it is however worth a comment that in reality, changes in temperature causes changes in phase fraction, and not vice versa. Hence the testing done for this module is purely for point of view of understanding mathematically whether or not it works. It is not physical.

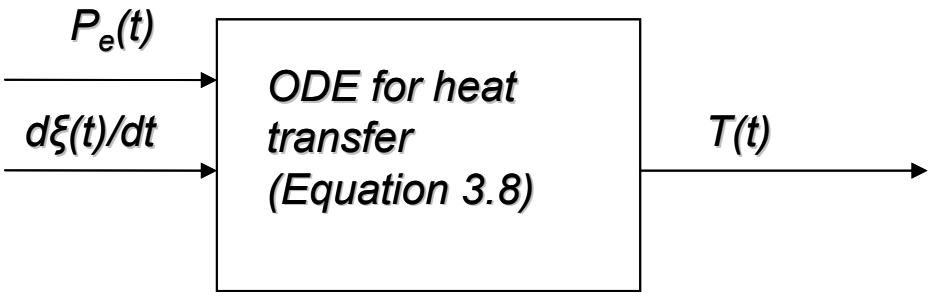

Figure 4-2: Test block diagram for the heat equation

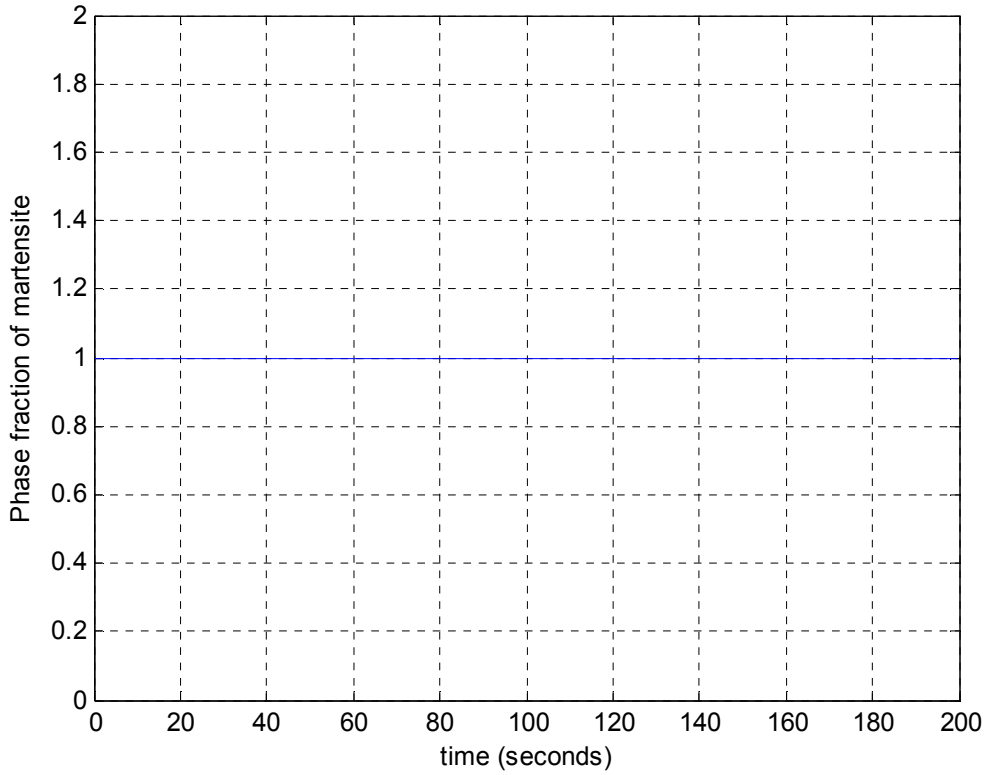

Figure 4-3: Input phase fraction (martensite)

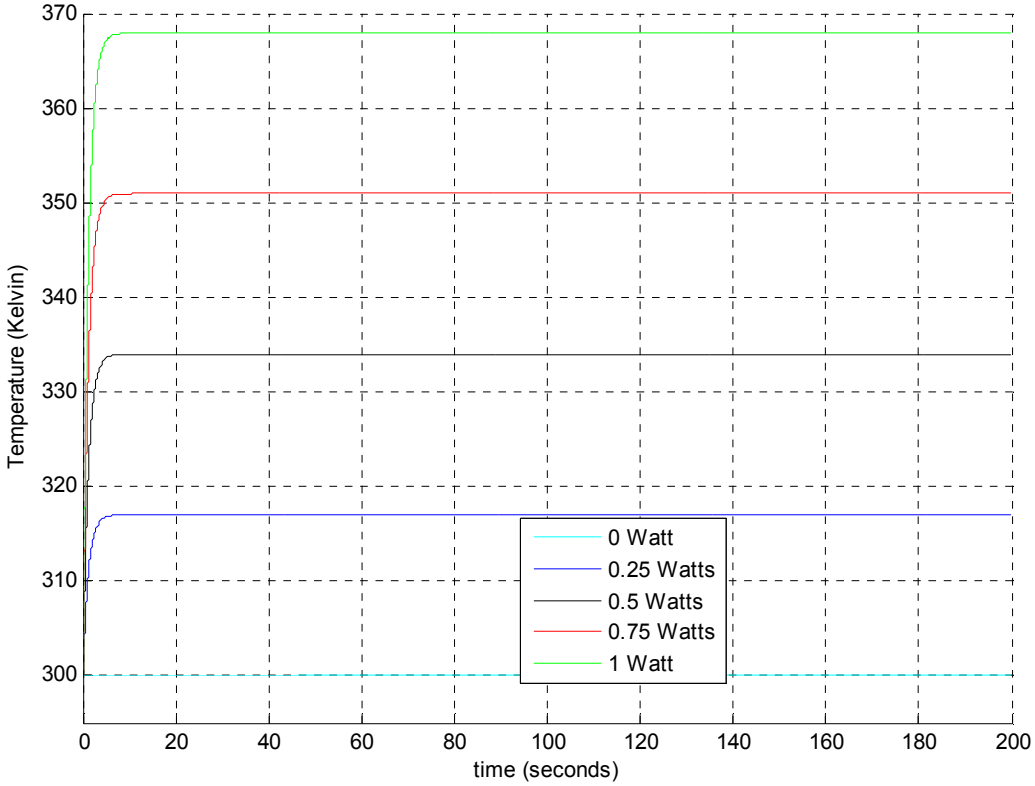

Figure 4-4: The observed temperatures corresponding to the different powers, in response to the phase fraction illustrated in **Figure 4-3**

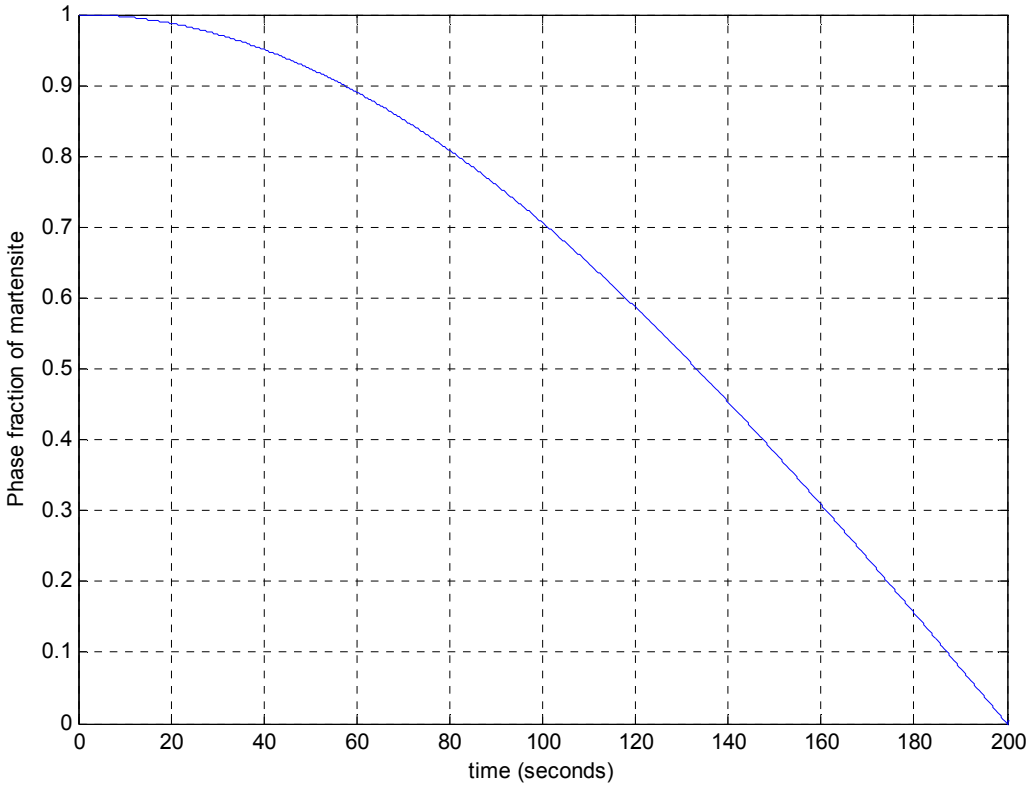

Figure 4-5: Input phase fraction (martensite)

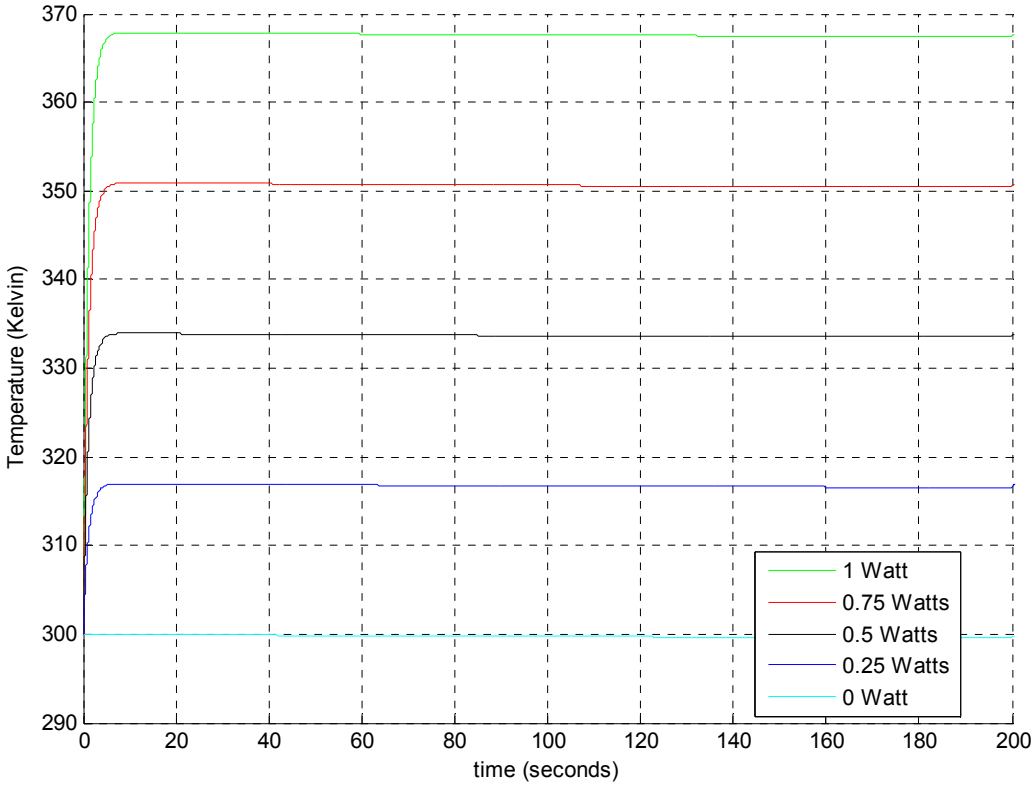

Figure 4-6: The observed temperatures corresponding to the different powers, in response to the phase fraction illustrated in **Figure 4-5**

4.2.2 Equilibrium block tests

The equilibrium block (Figure 4-7) was tested by applying the following input signals:

- **Phase fraction input:** A triangular input phase fraction signal (varying from 0 to 1, back to 0, in order to complete the full cycle of transformation) (Figure 4-8)
- **Stress inputs:** 1) A set of constant stresses (0, 100, 200 and 300 MPa) 2) Time varying stress input (varying from 0 to 150 MPa).

$$\sigma(t)A = F(t) = F_{app}(t) = EA(\varepsilon - \beta\xi) \tag{4.3}$$

The strain output in response to different constant stresses is illustrated in Figure 4-9. It is observed that the difference between the maximum and minimum strain is equal to the stress-free transformation strain β (~0.0535), and also the strain is shifted by ~0.002 for every 100 MPa of stress applied. Both these values are close to what is expected, as is justified mathematically using Equation 4.4. Using Equation 4.14, it is noted that given a constant stress σ, difference between the strains

corresponding to ξ=1 and ξ=0 is given by β. Also it is clearly seen from Equation 4.4 (in which Equation 4.3 is rearranged) that impact of stress σ on strain ε is given by $\frac{\sigma(t)}{E}$, which given the current values is equal to ~0.002.

$$\varepsilon(t) = \frac{\sigma(t)}{E} + \beta\xi \tag{4.4}$$

In Figure 4-11, the output strain signal in response to the constant input stress and varying input stress signals (Figure 4-10) are illustrated. It is seen that in response to the varying stress, although the initial strain is 0, the difference between the strains, corresponding to varying stress (blue) and constant 150 MPa stress (red) becomes smaller as the varying stress input increases in magnitude. This is as expected.

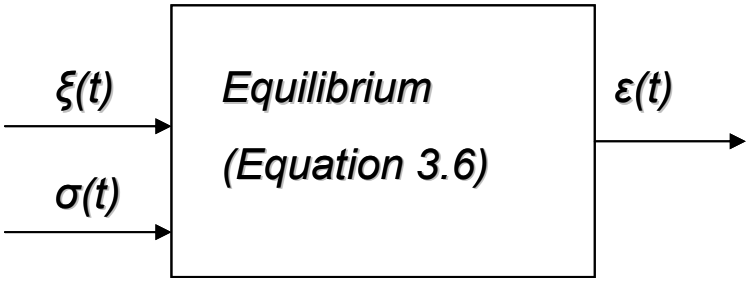

Figure 4-7: Test block diagram for the equilibrium equation

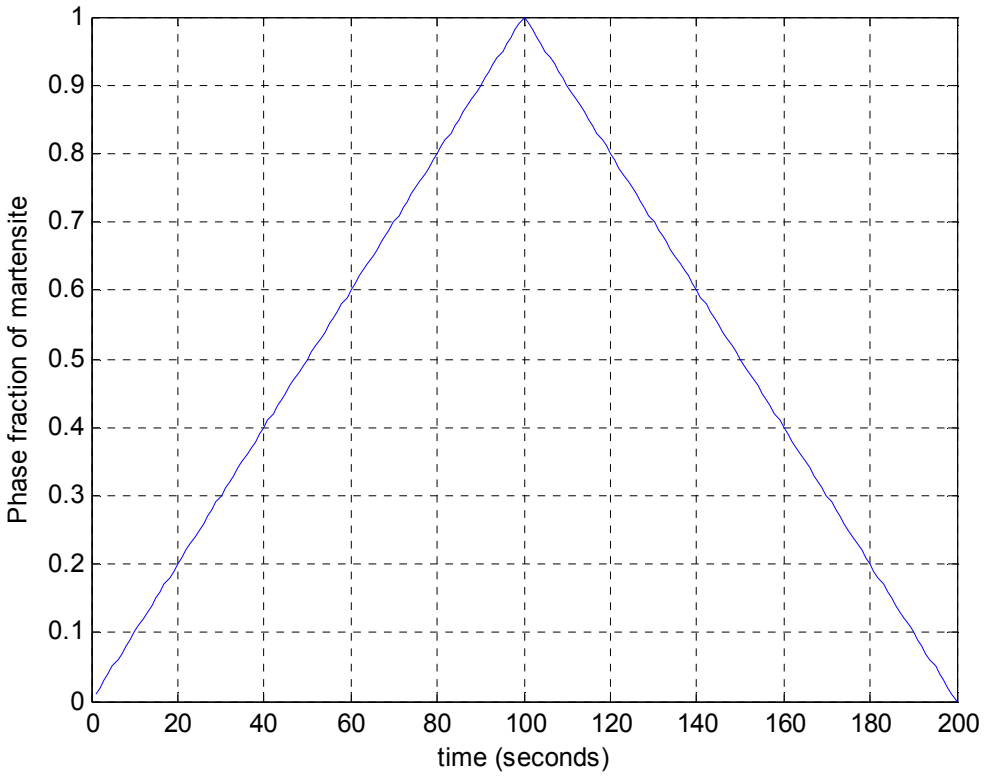

Figure 4-8: Test input phase fraction signal

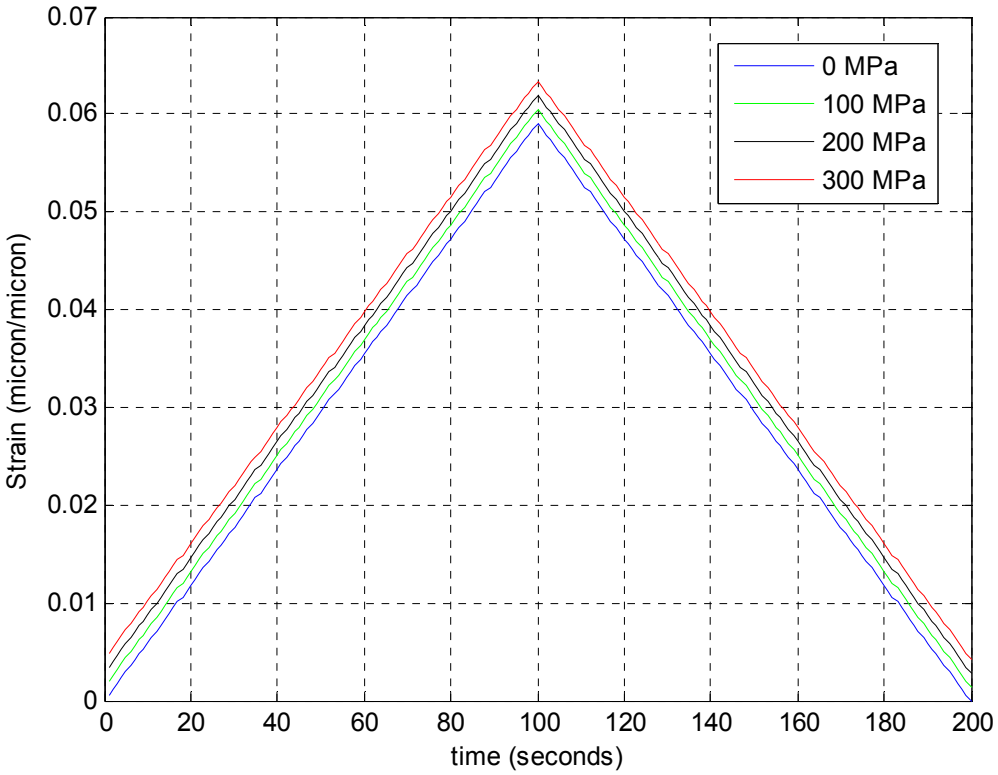

Figure 4-9: Output strain signal

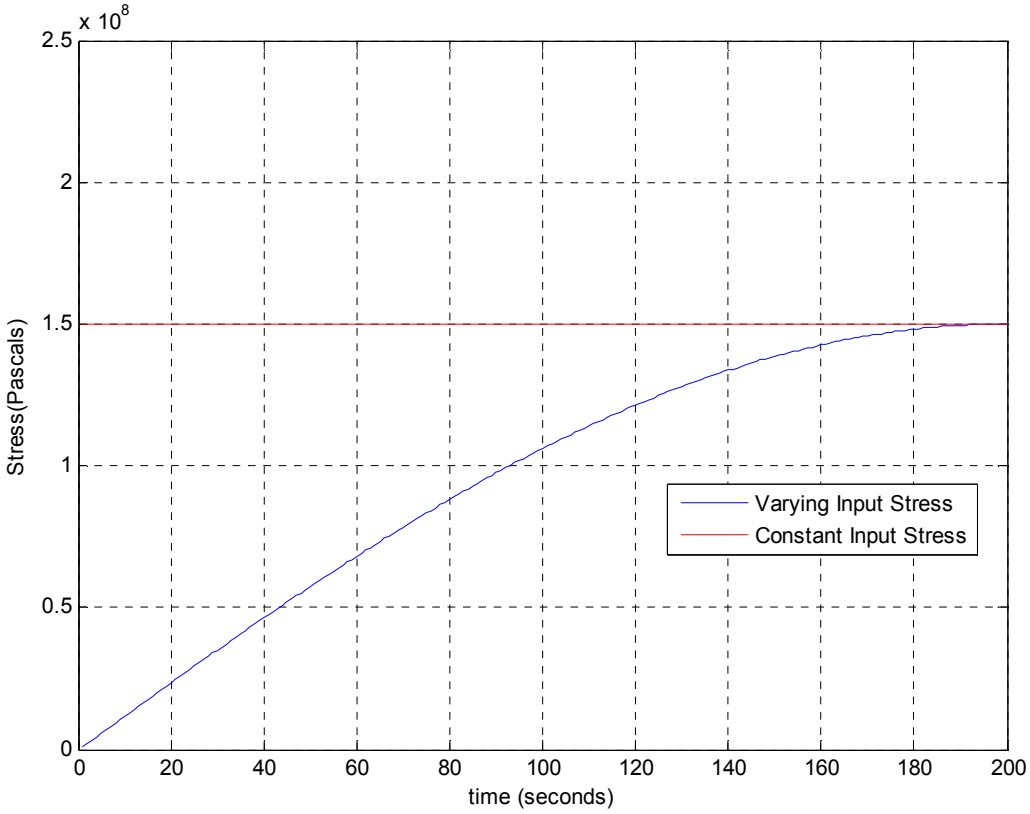

Figure 4-10: Input stress signals for testing the equilibrium block

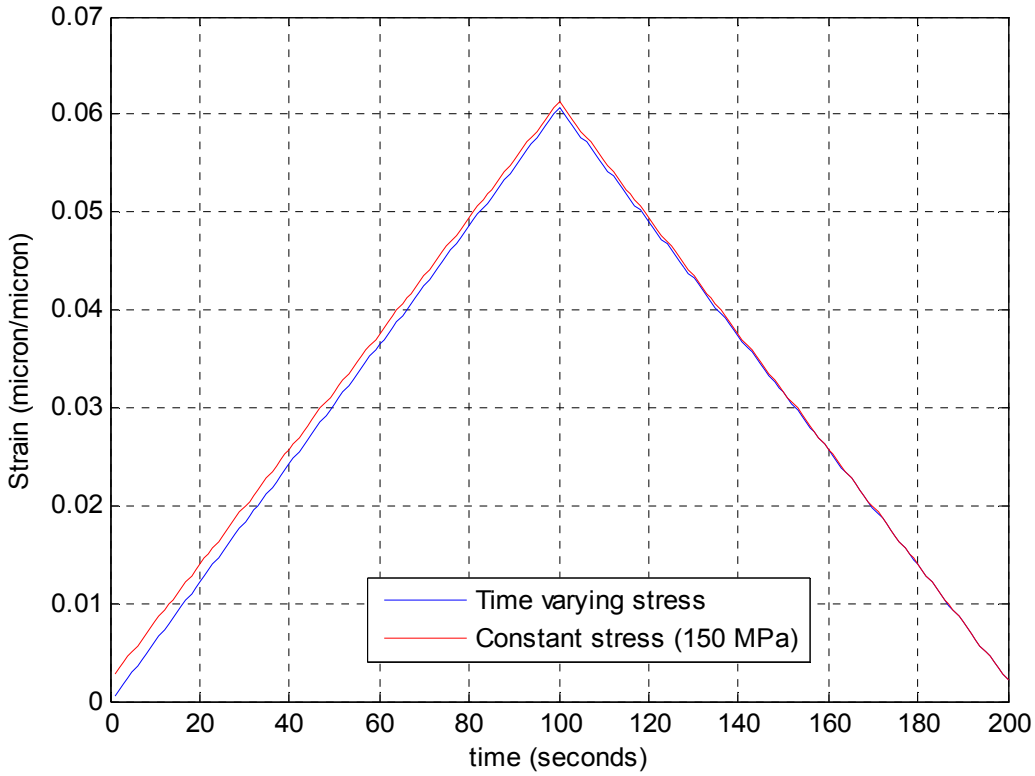

Figure 4-11: Output strain signals in response to the stress signals (**Figure 4-10**) and Phase fraction signal (**Figure 4-8**)

4.2.3 Phase transformation equation block tests

In this section, the testing of the phase transformation equation is described. As illustrated in Figure 4-12, this module takes in 3 inputs: temperature, rate of change of phase fraction, and the strain signals. Instead of providing all three signals as arbitrary inputs, the equilibrium module (tested) provides the strain input. The temperature is provided as an external input. The rate of phase fraction is calculated from the output phase fraction and fed back as an input into the system. In this test, the input temperature signal varies along a rectified sinusoidal signal between 275 Kelvin and 375 Kelvin. The input temperature (heating), the output strain and phase fractions are shown in Figure 4-13. From the strain and phase fraction subplots, it is observed that the transformation to austenite starts at around 350 K and finishes at 362 K. This is reflected in both the phase fraction and strain plots. Also, it is observed that the transformation to martensite starts at around 320 K and finishes at 308 K. The strain versus temperature hysteresis is shown in Figure 4-14. The strains versus temperature plots for different values of stresses (172 MPa – 700 MPa) are illustrated in Figure 4-15.

Qualitatively, a good match is obtained, in terms of the hysteretic behavior observed. In comparison with austenite start and martensite start temperatures according to the Dynalloy website [2] (illustrated in Figure 3-9), A_s at 172 MPa applied stress is 90C. This value is quite close to what is observed in the simulation results depicted in Figure 4-14 and Figure 4-15. Also it is observed in Figure 4-15 that both the residual strain and the transformation temperatures increase linearly with stress. This is as expected physically. Also on cross checking with Figure 4-9, it is found that the austenite and martensite strains corresponding to 300 MPa match with Figure 4-15.

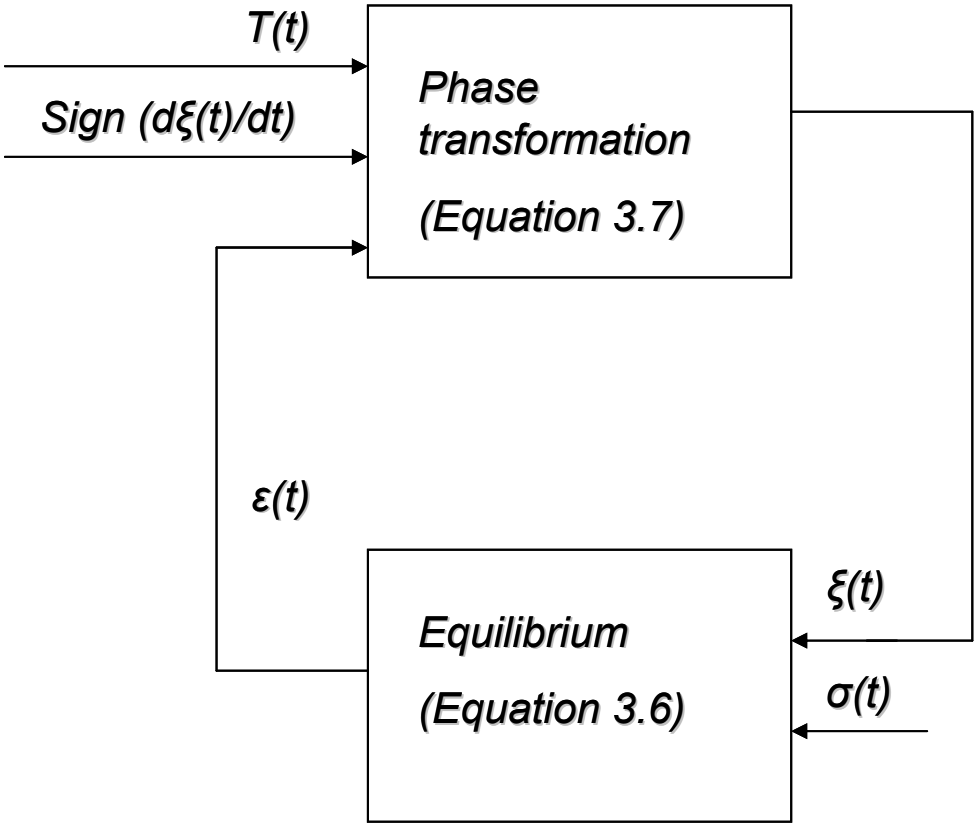

Figure 4-12: Test block diagram for the phase transformation block

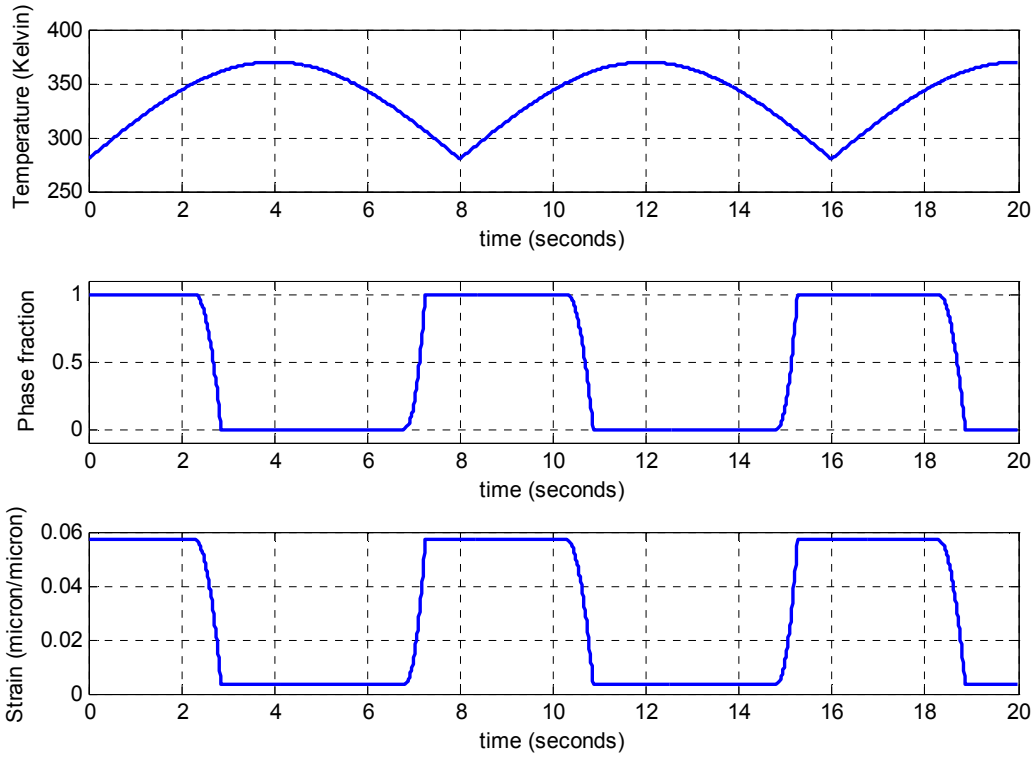

Figure 4-13: The temperature, strain and phase fraction plots for rectified sinusoidal temperature signals

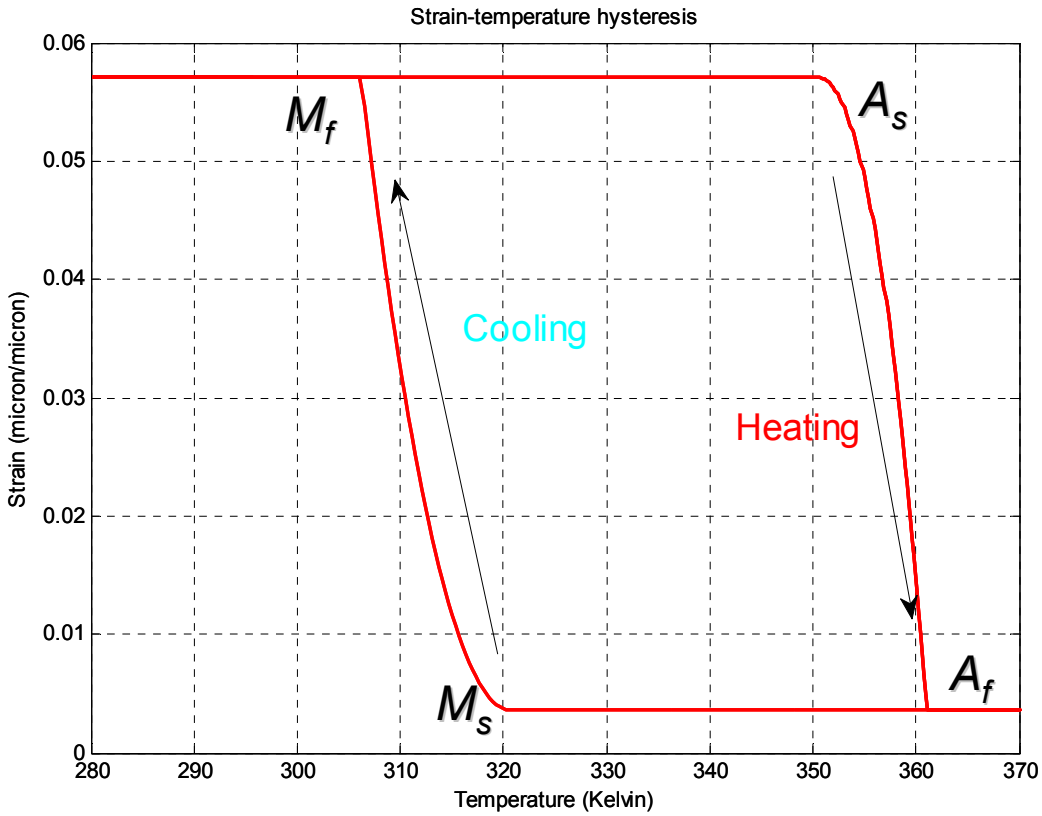

Figure 4-14: Strain versus temperature plot displaying hysteresis

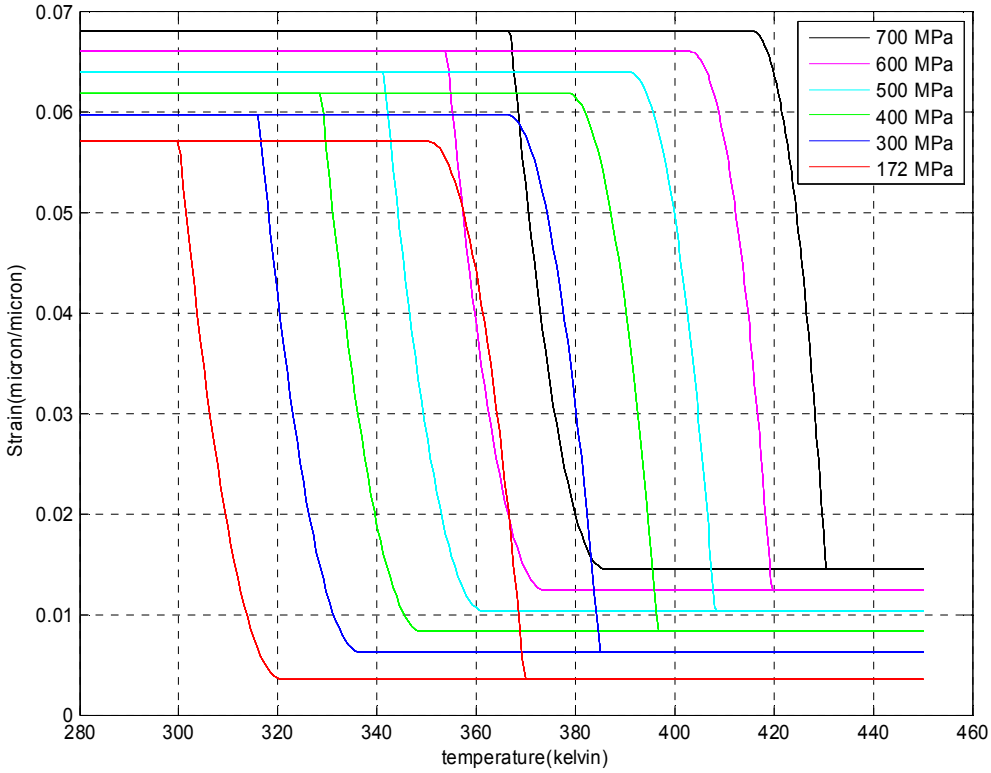

Figure 4-15: Strain versus temperature curves for different values of stresses (172 MPa-700 MPa)

4.2.4 Minor loop simulations

In this section, the minor loop algorithm is tested in simulation by using different test scenarios. The simulation set-up, the different test scenarios and the responses are discussed.

Simulation set-up

The simulation set-up is illustrated in Figure 4-16. The phase kinetics module with the minor loops code is created as a subsystem. The input temperature signal is provided as an input to the block. The delay blocks are required to store values from the previous time steps.

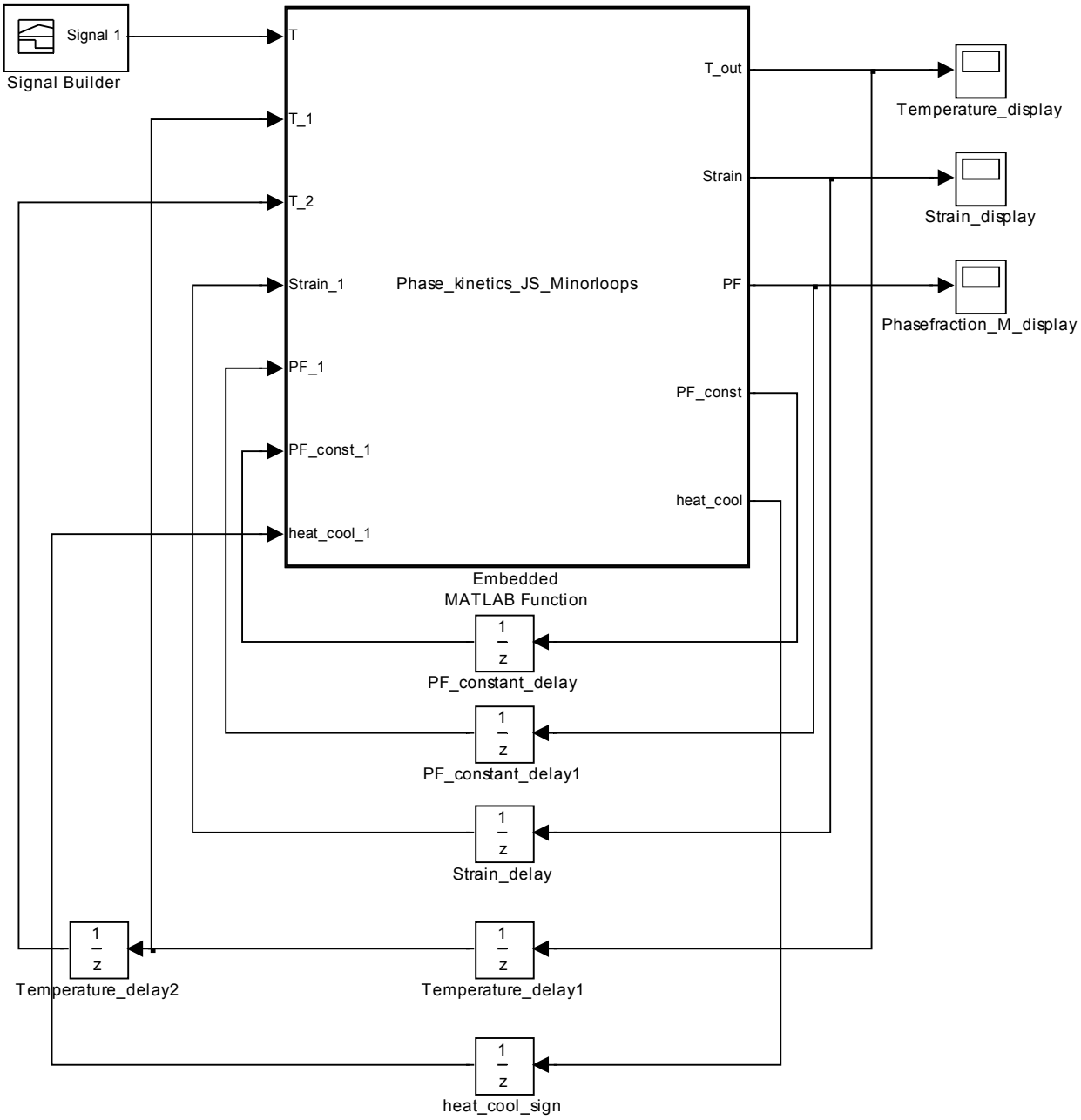

Figure 4-16: SIMULINK block diagram for testing the phase kinetics module with minor loops

The algorithm was first tested using a damped sinusoidal signal. The input signal is illustrated in Figure 4-17. The corresponding phase fraction signal is illustrated in Figure 4-18. The phase fraction versus temperature hysteresis plot is illustrated in Figure 4-19. In the hysteresis plot, a gradual drift is

observed in the major loops as the amplitude of the temperature signal changes. The nature of the temperature signal is such that with the change in the amplitude of signal, the rate of temperature change (*dT/dt*) also changes. This is because the period remains constant, although the amplitude changes.

This behavior is unexpected: major loops resulting from repeated cycling between temperatures below M_f and above A_f should overlap. The reason for this unexpected drift with varying *dT/dt* is numerical rather than due to the model itself. This drift can also be minimized by using smaller time – steps.

In order to see the expected behavior, with decreasing temperature signals (before the minor loops are triggered), $d\xi/dT$ should be rate-independent (essentially independent of the value of *dT/dt*). With sinusoidal temperature inputs of decreasing amplitude, *dT/dt* changes in each cycle. Hence if $d\xi/dt$ is proportional to *dT/dt*, then there would be no drift. In the model, using the phase kinetics and equilibrium equations (Equations 4.6 and 4.8), the constant terms are replaced with constants K_1, K_2, K_3 and K_4. This gives rises to the set of Equations (4.5 -4.9). Equation 3.5 represents the phase kinetics equations, with the constant terms replaced with K_1, K_2, and K_3. It is differentiated with respect to time and represented in Equation 4.6. Similarly the equilibrium equation is represented in Equation 4.7 and differentiated with respect to time in Equation 4.8. Equation 4.9 is arrived at by combining Equations 4.6 and 4.8. It can be seen that effectively $d\xi/dt$ is proportional to *dT/dt* as per the model.

However in the implementation, when the phase fraction is calculated in the current time step, the strain from the previous time step is used (since the strain from the current time step is not yet available). This strain is based on the phase fractions of the previous time steps. The resulting equation of $d\xi/dt$ is not straightforward, since previous ξ comes into the equation hence resulting in the drift.

$$\xi = K_1\varepsilon + K_2T \pm K_3 \tag{4.5}$$

$$\frac{d\xi}{dt} = K_1\frac{d\varepsilon}{dt} + K_2\frac{dT}{dt} \tag{4.6}$$

$$\varepsilon = \frac{\sigma}{E} + K_4\xi \tag{4.7}$$

$$\frac{d\varepsilon}{dt} = K_4 \frac{d\xi}{dt} \tag{4.8}$$

$$\frac{d\xi}{dt} = \left(\frac{K_2}{1 - K_1 K_4} \right) \frac{dT}{dt} \tag{4.9}$$

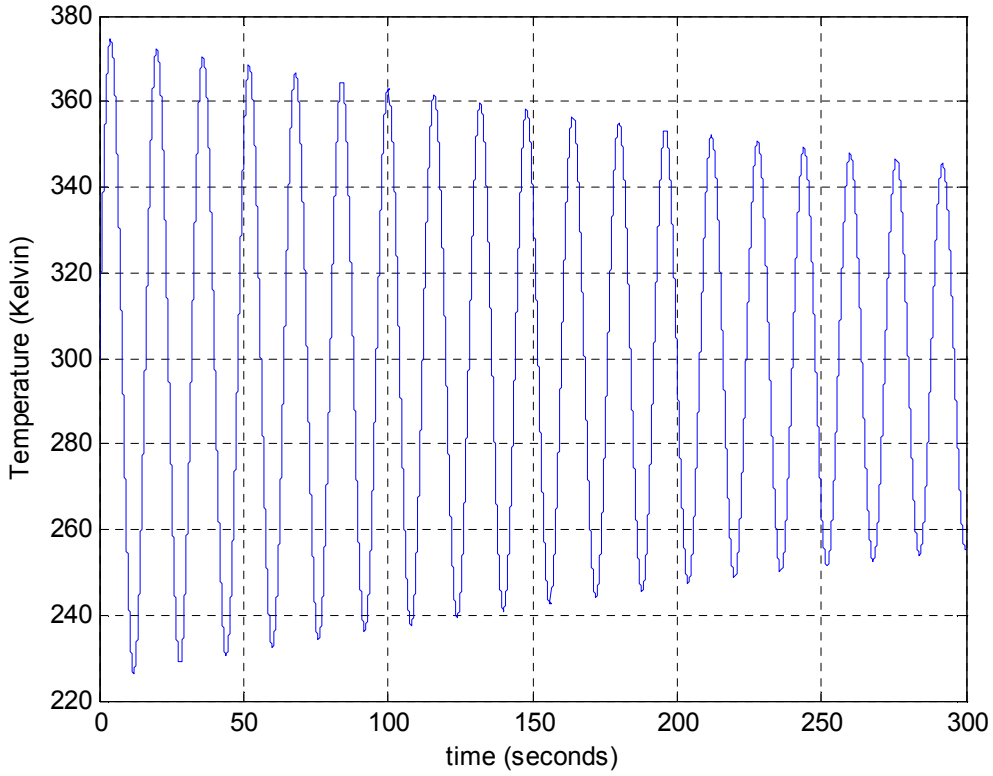

Figure 4-17: Input temperature signal as a damped sinusoidal signal

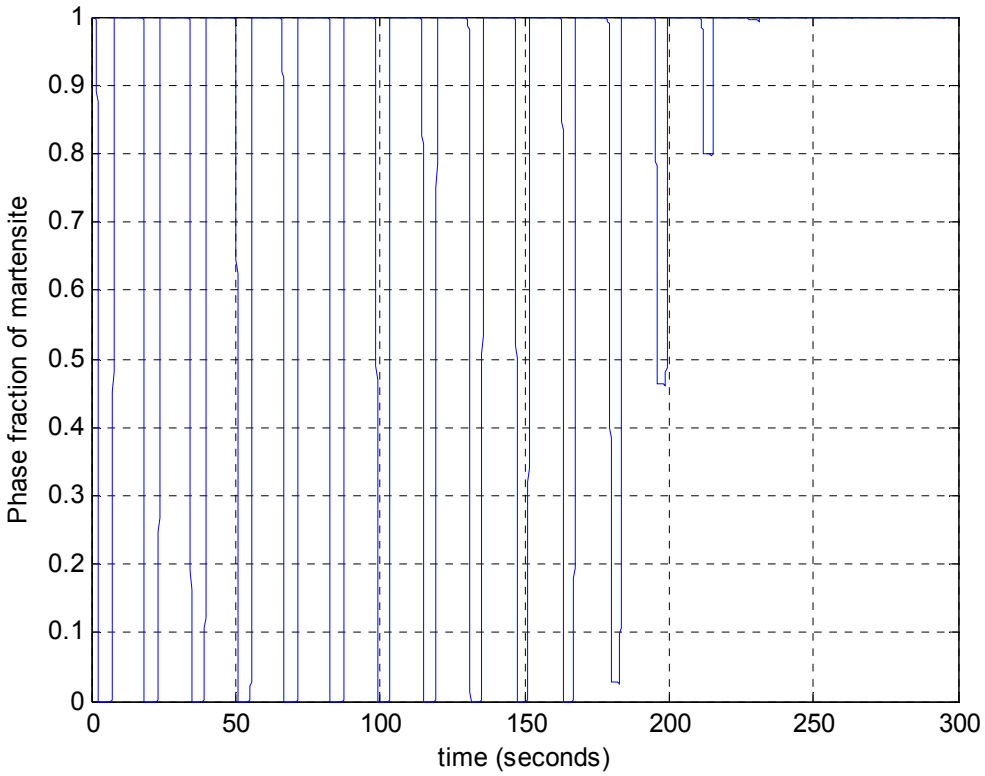

Figure 4-18: Output phase fraction in response to input signal given in **Figure 4-17**

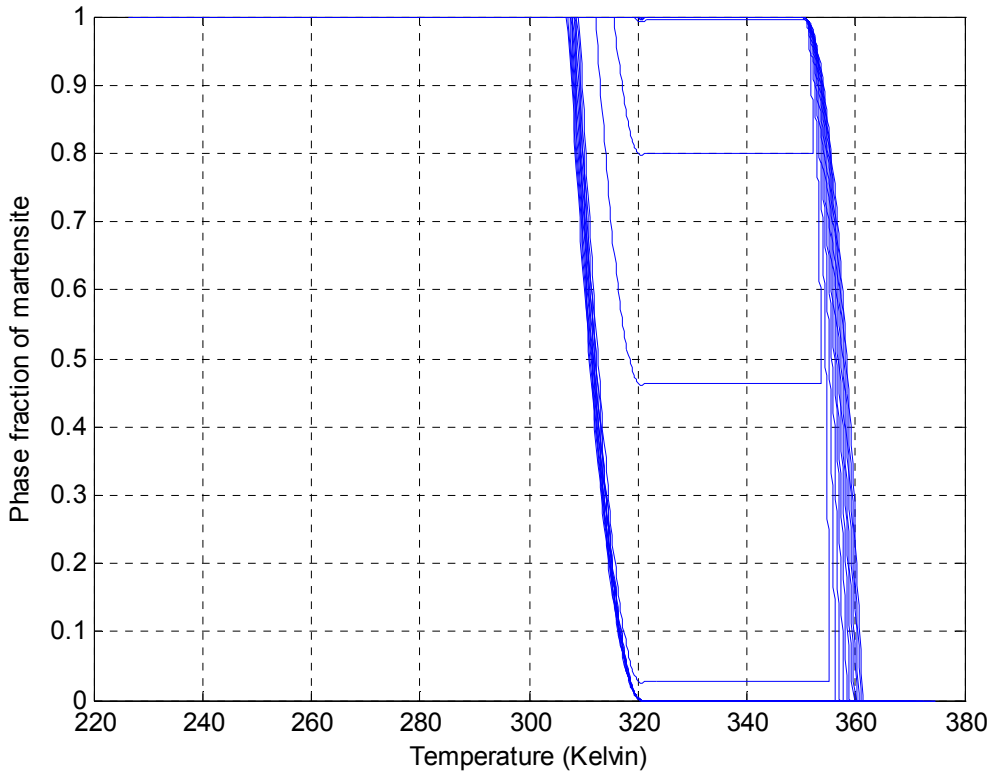

Figure 4-19: Phase fraction versus temperature hysteresis curve

Another observation with regard to the modeling of minor loops is the existence of "bumps" at the points where the minor loop starts to catch the major loop. This is illustrated in Figure 4-20, where the plot observed in Figure 4-19 is further zoomed around the "bumpy" regions. The reasoning for the existence of the "bumps" is as follows: There is a difference between the phase fraction according to the major loop equation and the phase fraction value on the minor loop equation, when the match occurs (given the tolerance). This causes the "bump" in the behavior. Reducing the tolerance does reduce the magnitude of the bump and has been tested in simulations.

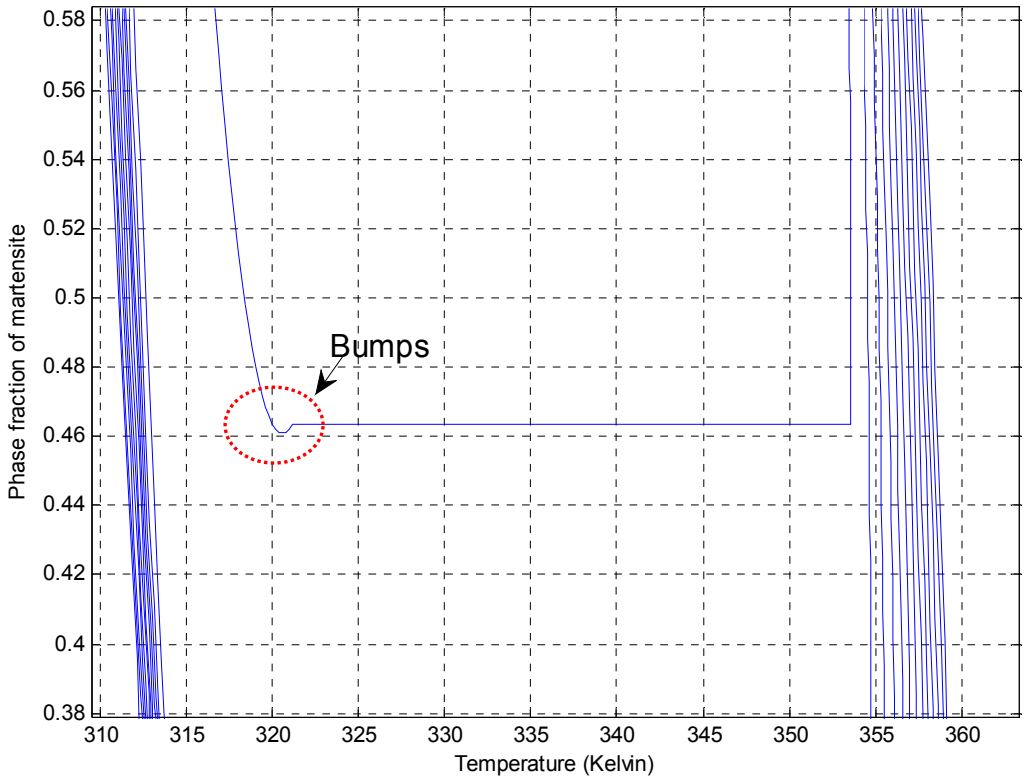

Figure 4-20: Zoomed version of **Figure 4-19** to illustrate the "bumps"

Damped triangle wave with constant dT/dt

In order to verify that the drift was definitely being caused by the varying *dT/dt*, another simulation was done using a triangle signal with constant *dT/dt* (Figure 4-21). A triangle wave was chosen over the sinusoidal signal, solely because of the ease of construction of a triangle wave with constant *dT/dt* as compared to a sinusoidal signal with constant *dT/dt*. It is observed in Figure 4-23 that the drift is completely minimized. It is worth mentioning at this point that the drift can also be reduced by using smaller time steps, however this increases the computational time. This is an important compromise to consider with view of closed-loop control, since the temperature signal is an intermediate quantity in cases where strain is the feedback variable hence controlling *dT/dt* is not straightforward.

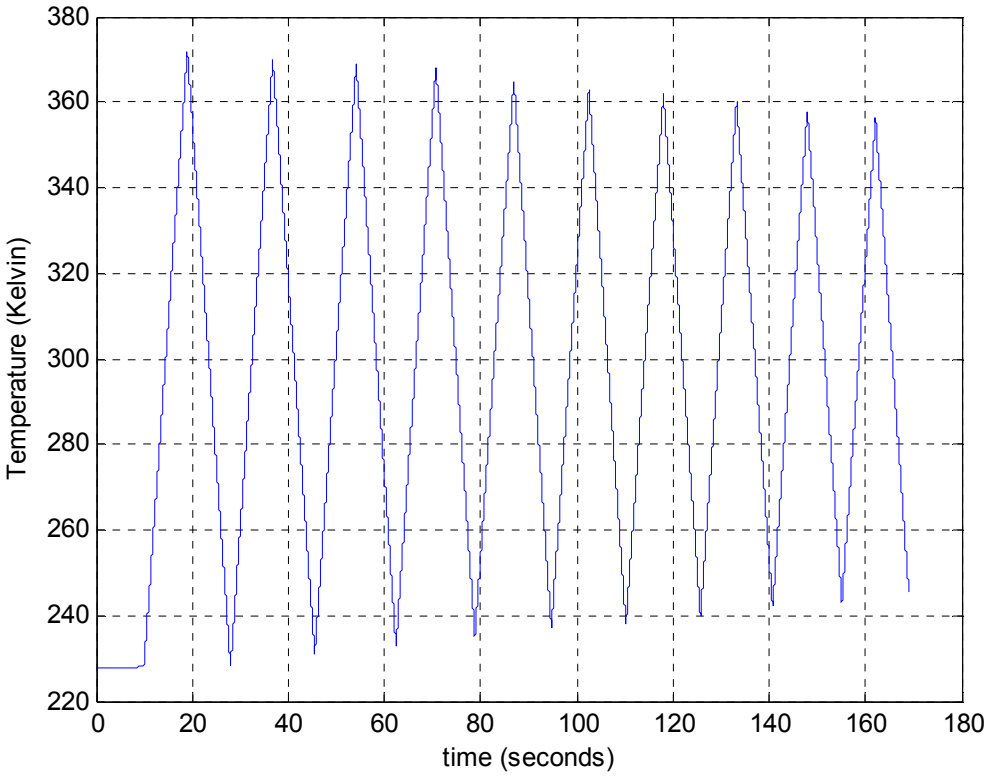

Figure 4-21: Damped input triangle wave with constant *dT/dt*

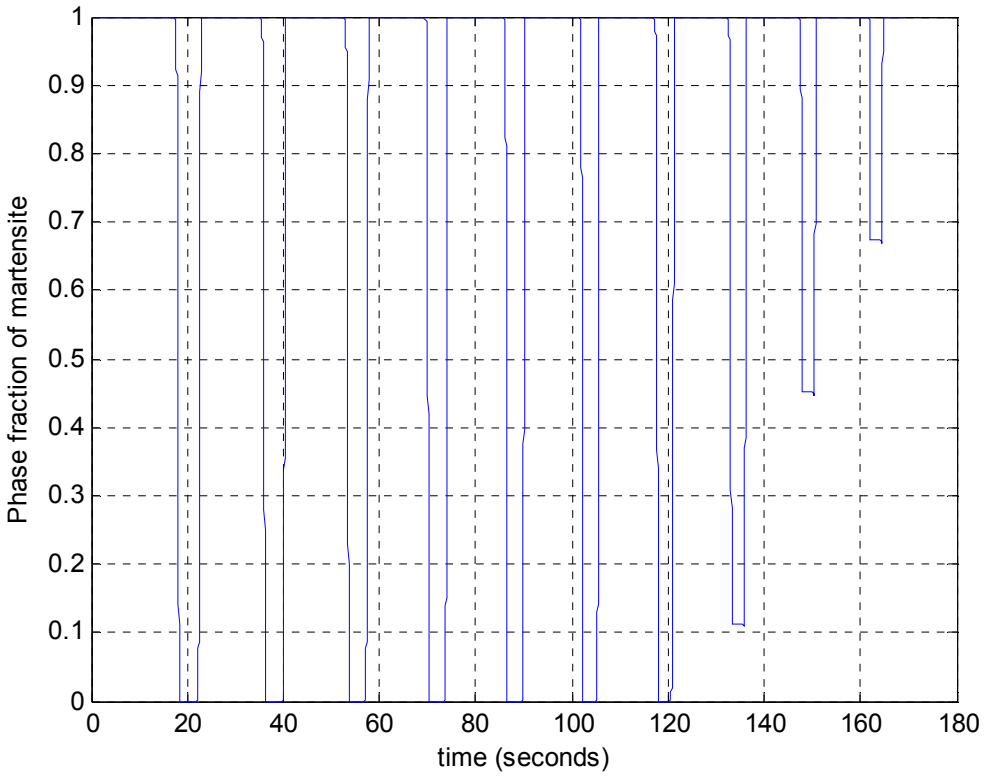

Figure 4-22: Output phase fraction in response to input signal given in **Figure 4-21**

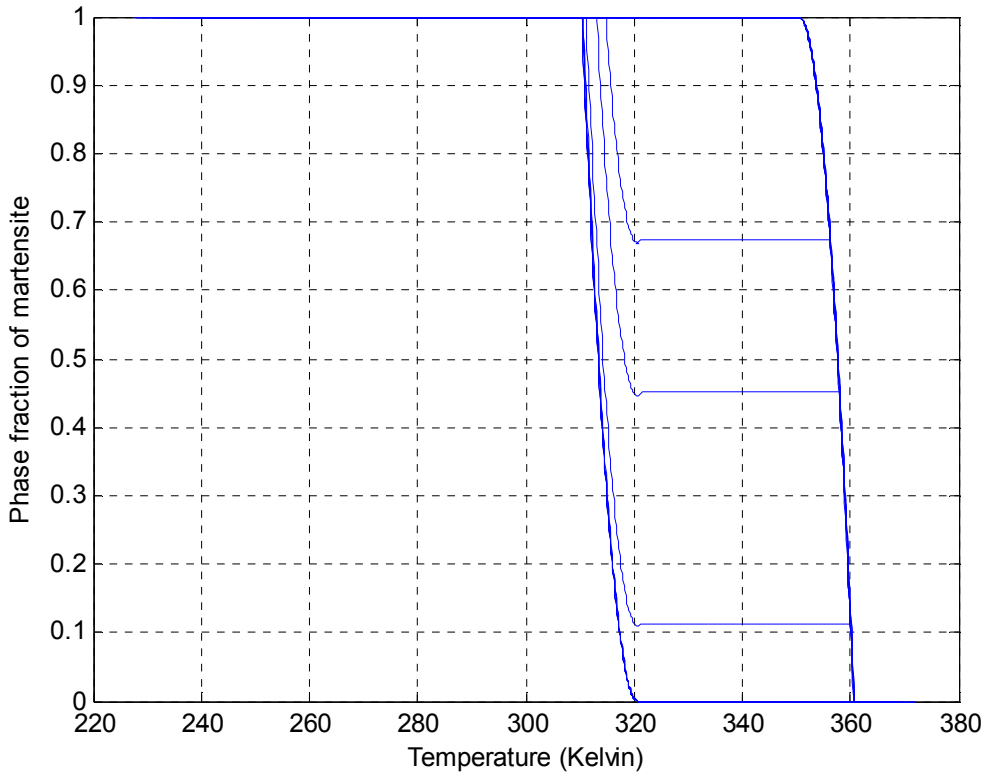

Figure 4-23: Phase fraction versus temperature hysteresis plot

Composite signal

For completeness and in order to ensure that the algorithm developed deals with the minor loops/partial transformation behavior under different conditions, different cases need to be tested:

- Major loop hysteresis
- Minor loop (when temperature changes direction from cooling to heating and vice versa) on the major loop.
- When a change in the direction of the temperature occurs on the constant portion of a minor loop.
- When a change in the direction of the temperature occurs when the transformation is occurring (but on the minor loop).

A composite signal is constructed (depicted in Figure 4-24, green) in order to cover the above-mentioned cases. The phase fraction response to the signal is also illustrated in Figure 4-24 (blue).

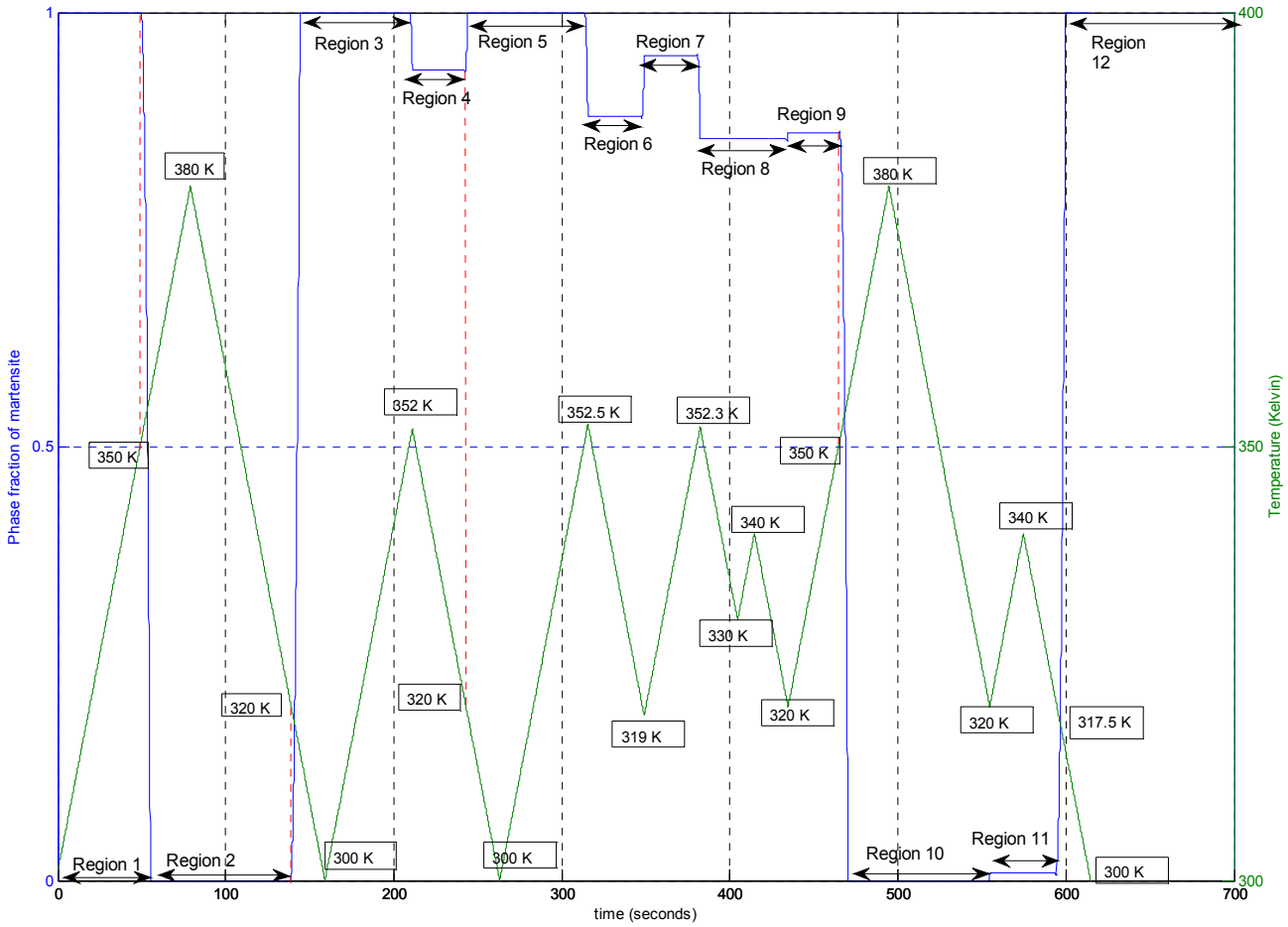

Figure 4-24: Input temperature signal (green) and the output phase fraction signal (blue)

The initial conditions at the start of the simulation are as follows: Phase fraction of martensite (ξ) = 1, Temperature (T) = 300 K, Strain (ε) = 0.0535, chosen equal to the stress-free transformation strain, applied stress (σ) = 172 MPa. An explanation is provided of the observed behavior. For better understanding, the time period of 700 seconds is divided into 12 regions. The demarcation between regions is made at points when a change in the direction of the phase fraction occurs.

Region 1: The temperature increases from 300 K to 350K. At 350K, the martensite to austenite transformation starts to occur. This is as expected since the A_s (Austenite start temperature) in these simulations is 350 K.

Region 2: The temperature continues to increases to 380 K and then decreases. The phase fraction remains constant till T= 320 K. At this point, the transformation from austenite to martensite starts. This is also as expected since M_s (Martensite state temperature) in these simulations is 320 K.

Region 3 and Region 4: The temperature decreases to 300 K and then increases to 352 K. During this time, the phase fraction remains constant at 1 till 350 K, when the transformation from martensite to austenite starts. However since the temperature changes direction at 352 K, the transformation is not completed and the minor loop is triggered. The phase fraction remains constant until temperature equals 320 K. At this point, the transformation back to the martensite starts to occur.

Region 5: In this region, the temperature decreases to 300 K and then increases up to 352.5 K. During this time, the phase fraction remains constant at 1, till 350 K, when the transformation from martensite to austenite starts. However since the temperature changes direction at 352 K, the transformation is not completed and the minor loop is triggered.

Region 6: The phase fraction remains constant until the decreasing temperature equals 320 K. At this point, the transformation back to the martensite starts to occur. However at 319 K, the temperature changes direction and starts to increase again.

Region 7: The temperature continues to increase from 319 K. Around 350 K, the transformation to austenite starts. However the temperature changes direction at 352.5 K, and the minor loop is triggered and the phase fraction remains constant.

Region 8 and Region 9: The phase fraction continues to remains constant. The temperature changes direction at 330 K, hence double reversal occurs i.e., the temperature changes direction when the constant portion of the minor loops is being traversed. The temperature increases up to 340K and then again changes direction falls down to 320 K. At this point M_s is reached and hence the minor loops starts following along the major loop. When temperature is equal to 350 K (A_s), phase fraction starts to transform back to austenite.

Region 10, Region 11 and Region 12: The temperature rises to 380 K and decreases to 320 K. During this time, the phase fraction stays at 0 (Austenite phase). At 320 K, it starts to transform to

martensite. However at 340 K, the temperature direction changes again and a minor loop is formed. Around 318 K, the major loop is met and the phase fraction starts to convert back to martensite and stays there.

The phase transformation versus temperature hysteresis plot is illustrated in Figure 4-25. A zoomed version of this figure is illustrated in Figure 4-26. The direction of movement along the curve is marked by arrows and each cycle is marked by a different colored set of arrows (green -> lilac ->black ->brown).

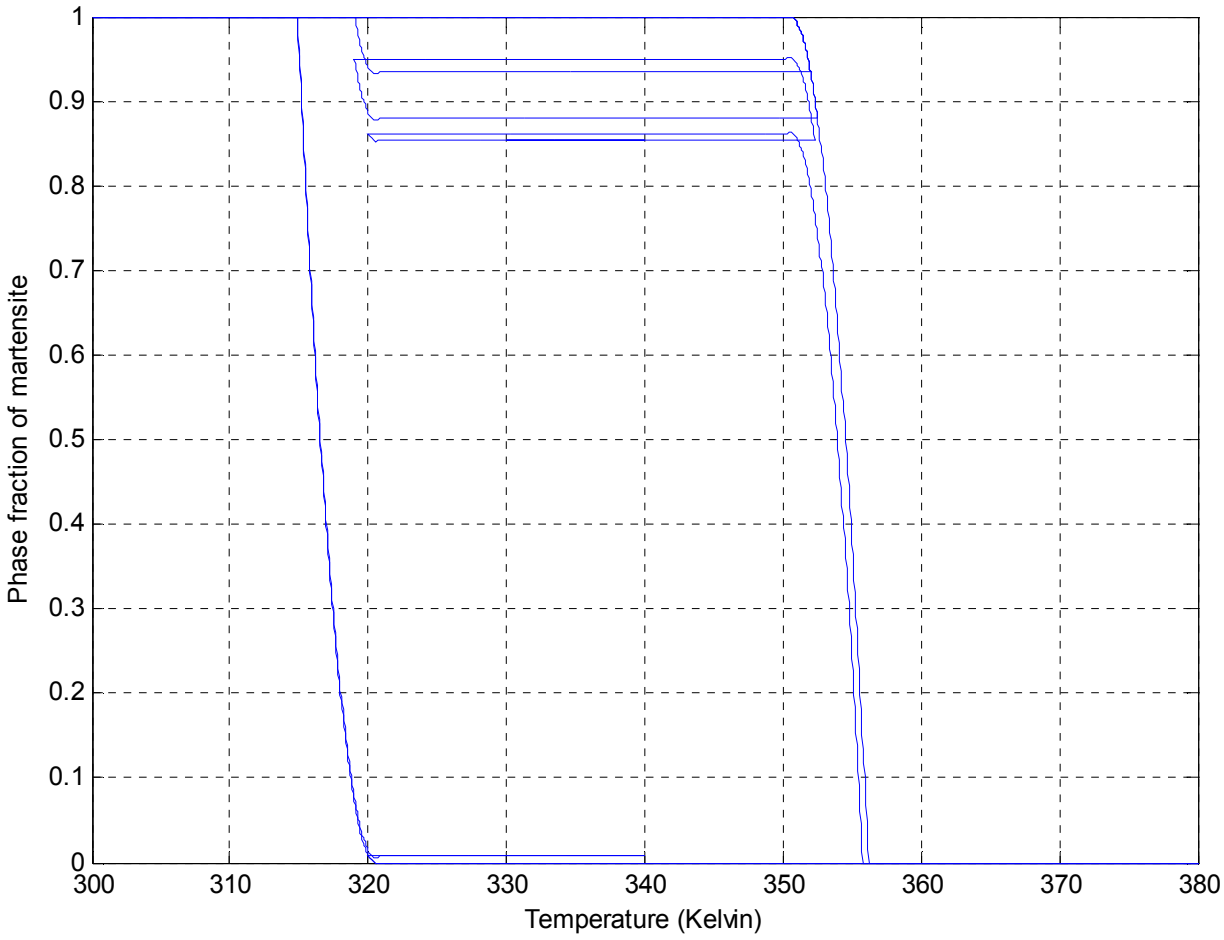

Figure 4-25: Phase fraction versus temperature hysteresis plot

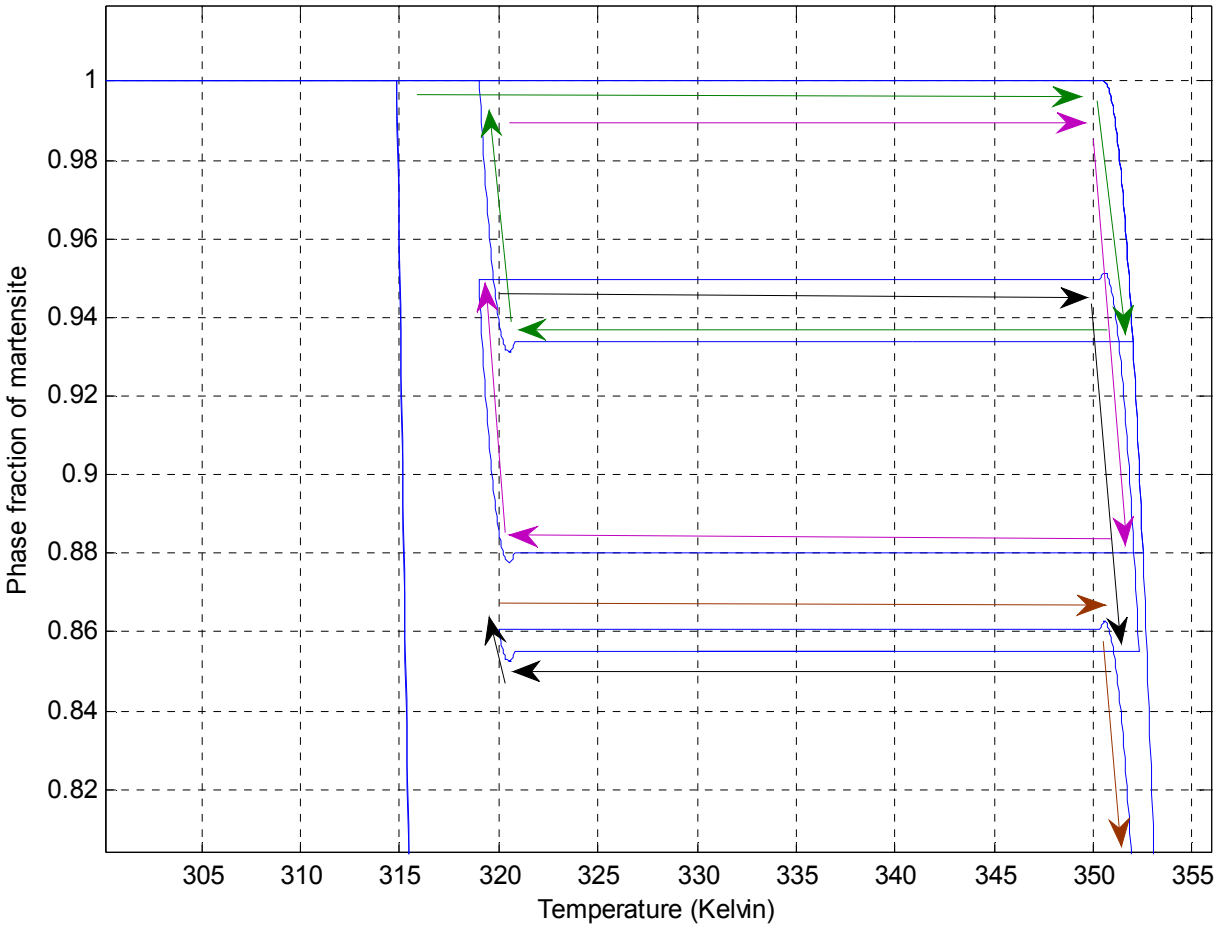

Figure 4-26: Zoomed image of the phase fraction versus temperature hysteresis

4.3 Summary of the chapter

In this chapter, the individual modules of the John Shaw model are tested in simulation and qualitatively analyzed. Also the ability of the model to display minor loop behavior is tested. The observed behavior is explained and it is concluded that the minor loop behavior is suitably displayed by the improved model. This renders the John Shaw model with a greater applicability to control applications.

Chapter 5 Improving the response time of SMA actuators

In this chapter, the concept of model-based temperature saturation (MBTS) scheme is discussed. This is followed by closed-loop simulations comparing the performance of the MBTS scheme with the conventional CS scheme. The plant (SMA wire) in the actuator system is represented by using the first order heating equation combined with the thermomechanical part of the John Shaw model (phase kinetics and the equilibrium portion of the John Shaw model). The first order heating equation was used instead of the John Shaw heating equation in order to keep consistency between the TEM model used for MBTS and in the plant.

An improvement is seen in the initial portion of the cycle, when a significant change is needed in the strain. For situations requiring smaller strain changes, the performance of the MBTS scheme and CS scheme are similar.

5.1 Concept of Model-based Temperature Saturation

A commonly used actuation scheme for SMA actuators with strain feedback is illustrated in Figure 5-1 and has been briefly introduced in Chapter 1. In this scheme, the error signal is fed into a linear controller (e.g. P/PI/PID) where the output is current (A). This current is saturated at the manufacturer specified safe current (MSSC). The current is fed into the SMA plant and accordingly the output strain is fed back and subtracted from the desired strain to get the error.

This actuation scheme presents two limitations:

- Firstly, limiting the current at the manufacturer specified safe current presents a very conservative limit on the maximum current that can potentially be applied to the wire without damaging the wire.
- The control portion includes only a linear control part. However, SMA exhibits nonlinear behaviour due to the hysteresis. The controller should be augmented to account for this nonlinear behaviour, resulting in improved performance of the SMA actuator over a wider range of operating conditions.

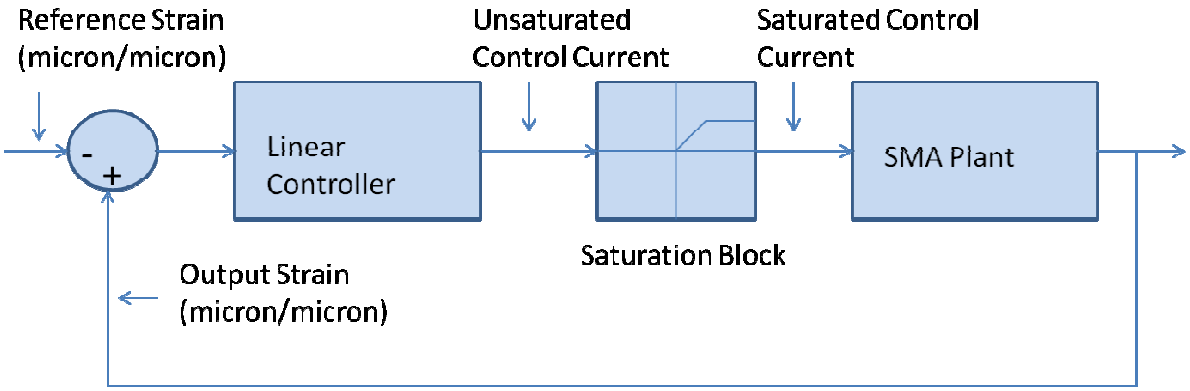

Figure 5-1: Actuator control using conventional current saturation

In this chapter, a technique is presented to deal with the first limitation mentioned above. The MSSC is specified on the premise that it could be applied indefinitely. Hence, it is undoubtedly a conservative limit on the maximum current. Note that it is actually high temperature and not high current that damages the wire. Thus, an improved technique is to use the wire temperature to saturate the current only when the temperature has risen above a particular critical value (chosen as a value below the annealing temperature and above the austenite finish temperature, preferably closer to the latter). The annealing temperature is the temperature at which the SMA reaches a state where its properties can be altered; hence it is important to ensure that the chosen critical value of the temperature is lower than the annealing temperature.

The two approaches employed previously by researchers to achieve the goal of saturating the current based on the temperature are as follows:

- Direct temperature measurements [37, 38] using non-contact temperature sensors or thermocouples.
- Indirect measurements [39, 40], in which some other measurable quantities, such as resistance, have been used to estimate the temperature of the SMA.

A summary of the work done in this regard has been presented in Section 2.7. In this project, the goal is to further improve upon the proposed schemes in [39, 40] by using an accurate TEM and calculating the temperature accordingly. The TEM can be used to predict the temperature based on the current, phase fractions, and stress, (as depicted in Figure 5-2) and accordingly ensure the temperature is not high enough to overheat the wire. Hence, a threshold temperature (T_{Thres}) is chosen whereby the temperature of the wire is not allowed to rise above T_{Thres}. A switching condition

(Equation 5.1) is set for the current into the SMA plant such that when the temperature of the plant is below the threshold value, the controller current is fed into the plant. Otherwise, the current is shut off in order to passively cool the wire as fast as possible.

$$I_{plant} = \begin{matrix} I_{control} & T \le T_{Thres} \\ 0 & T > T_{Thres} \end{matrix} \qquad (5.1)$$

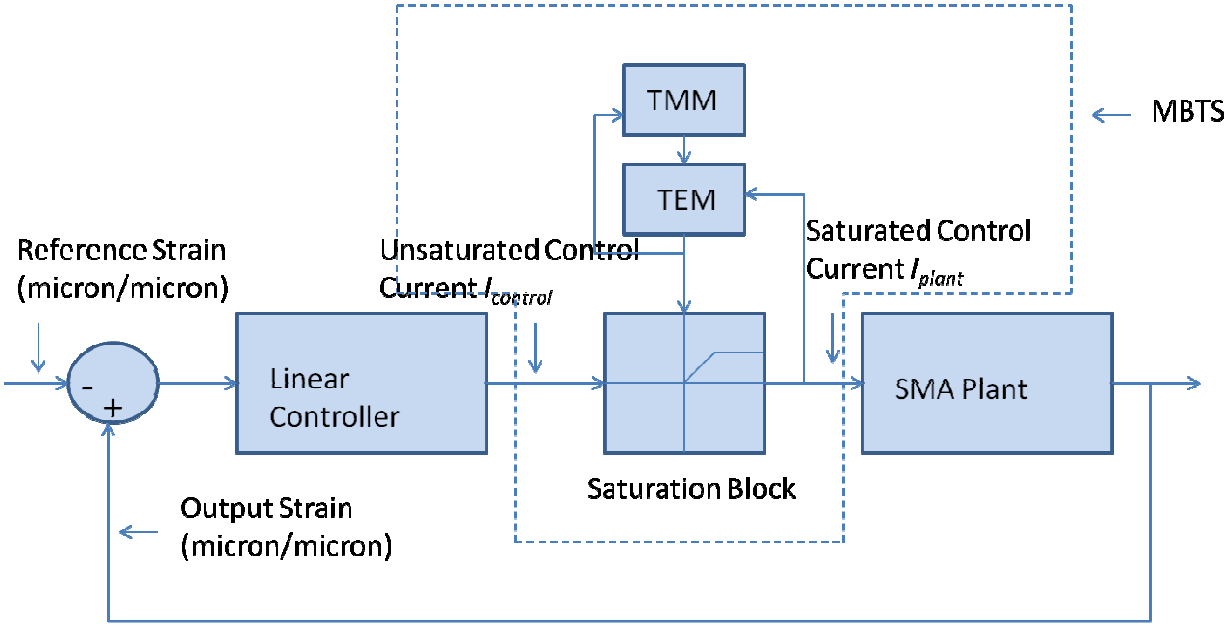

Figure 5-2: Actuator control using temperature saturation via a thermoelectric model (TEM)

However, as stated in [15, 16, 17] and also summarized in Chapter 2, there exists coupling between the TEM and TMM. In particular, the material properties used in the TEM, such as resistance and heat transfer coefficient, change with the evolution of the phase fractions [15, 16, 17]. The scheme proposed in Figure 5-2 can be improved upon by coupling the TEM with the TMM. The degree to which the coupling between the TEM and the TMM needs to be taken into consideration depends upon:

- The difference occurring in the model-predicted temperature values using only the TEM, versus taking into account the coupling between the TEM and the TMM.
- How sensitive is the actuation time to the switching temperature? This implies impact on the actuation time if the switching temperature is a certain number of degrees higher or lower than the chosen value.

These two criteria together determine the degree to which coupling is essential. In the works of [15, 16, 17], it is stated that the thermal and electrical parameters of the heating model are different in the martensite and austenite phases. Rather than assuming a discrete change in the value of the parameters on occurrence of full transformation from austenite to martensite or vice versa, it makes more physical sense for the values to evolve as the phase fractions evolve. The authors make an assumption of linear dependence on the phase fractions [17]. Based on the experimental data presented in [44], it is seen that the behaviour is not exactly linear, but shows peak-like occurrence(s) both for electrical resistivity and specific heat capacities. The authors note, however, that this is highly dependent on the composition of the alloy. Due to the absence of corresponding data for Nitinol, for simplicity the variation was taken to be linear by the authors [44]. In this chapter, the simulation results are presented of the MBTS based only on TEM. However, the coupling of the TEM and TMM is an interesting and important future direction to pursue.

5.1.1 Comparison of MBTS versus CS simulations

In order to illustrate the improvement observed on using the MBTS scheme over the CS scheme, two sets of simulations were carried out using a SIMULINK version of the actuator scheme illustrated in Figure 5-2. First was a preliminary validation of the idea of MBTS using a simplified SMA model consisting of a backlash operator coupled with an equilibrium equation to convert phase fraction to strain. The second set of simulations was done with the extended John Shaw model as the plant. For purposes of brevity, only the second set of simulations is included in the thesis. The results of the former can be found in [45], a deliverable submitted to General Motors.

The controller used is a P-controller. In order to simulate the plant, the TEM and TMM are modelled separately. The output (temperature) of the TEM block feeds as an input to the TMM block. The coupling between the TEM and TMM is not modeled. The TEM is modelled using the first-order heating equation (Equation 5.2). The temperature from the TEM is an input to the hysteretic temperature to phase fraction block. For the temperature to phase kinetics portion, the phase kinetics module of the John Shaw model (with minor loops) is used. This was described in greater detail in Chapter 3.

$$\rho C_p V \frac{dT}{dt} = R\left(i(t)\right)^2 - hAT(t) \tag{5.2}$$

In the current set of simulations, the SMA plant is represented by the John Shaw model. However certain limits are placed on K_p (the controller proportional gain) and nature of the input signals into the system. These limits are posed due to a limitation in the implementation of the minor loops algorithm. It should be noted that if this limitation is suitably eliminated, then a greater improvement can be achieved using the MBTS over the CS scheme as compared to the improvement noted in the simulations in this chapter.

The limitation is as follows (illustrated in Figure 5-3): the performance of the model is dependent on the rate of change of temperature. This is specifically with regard to detecting the point where minor loops meet the major loop. The algorithm for minor loops uses a tolerance value while detecting when the value of the phase fraction (held constant) on the minor loop matches with the value of the phase fraction according to the minor loop (heating or cooling equation).From the phase kinetics equation, it is seen that the phase fraction is dependent on the temperature hence the rate of change of phase fraction is dependent on the rate of change of temperature. If the temperature changes very rapidly between time steps then the phase fraction also changes very rapidly between time steps. If the difference between the phase fraction values in two consecutive steps is greater than the tolerance, then the point where the match occurs between the minor loop and the major loop is missed. Consider that the temperature signal changes at a rate such that the temperature values at the consequent time steps are T_1 and T_2 and the corresponding phase fractions are ξ_1 and ξ_2 respectively. If the difference between ξ_1 and ξ_2 is greater than the tolerance specified for the simulations, then it is clearly seen that point of match is not detected and the minor loop continues along its path.

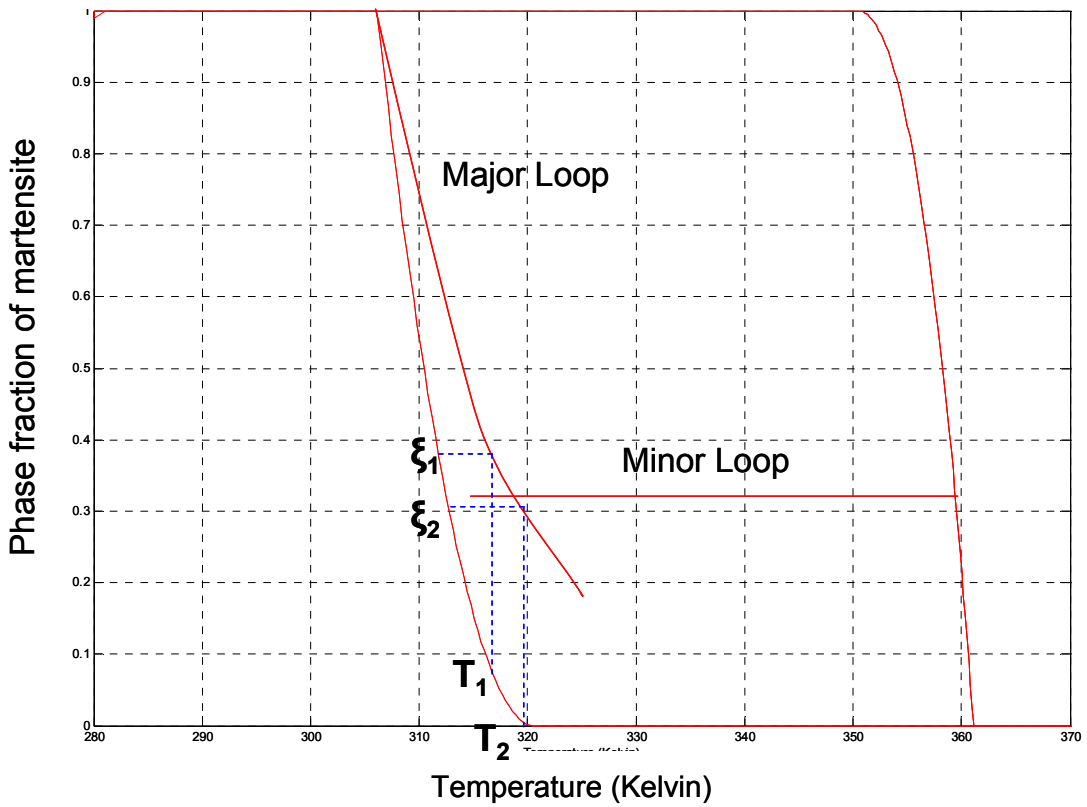

Figure 5-3: Figure illustrating the tolerance related issue

To some extent, this issue can be dealt with by increasing the tolerance and decreasing the step size. However increasing the tolerance decreases the accuracy and the "bumpy" behavior becomes more prominent. Also decreasing the step size increases the computational time. For the simulations in this chapter, the phase fraction tolerance was set to 0.005 and time step, to 0.001 seconds.

While testing the model components with open-loop signals in Chapter 3, the input temperature signals were chosen such that the rate of change of temperature was not high enough to trigger the issues mentioned above. However, in the strain-based feedback system such as the one used for the MBTS versus CS simulations, choosing the strain such that the resulting temperature signal does not change faster than a certain rate is not as straightforward. Hence the gain/input signals were chosen after some trial-and-error.

5.1.2 Results

A comparison is done of the tracking performance, using a reference signal of the form (Equation 5.3) using MBTS scheme versus the CS scheme. This is further illustrated in Figure 5-4, Figure 5-5 and Figure 5-6.

$$r(t) = 0.04\left|\sin(kt)\right| + 0.01 \tag{5.3}$$

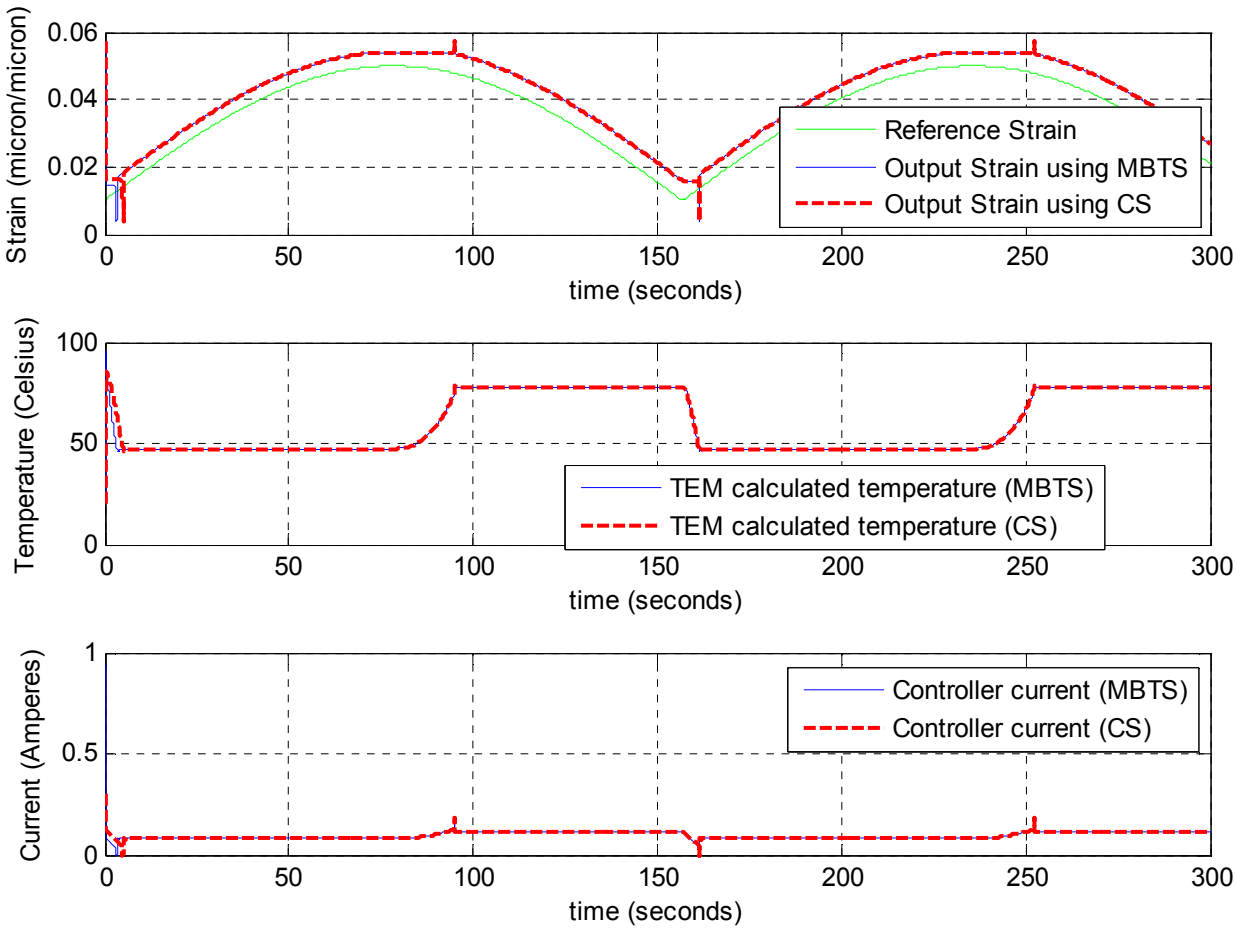

Figure 5-4: Comparison of the strain profiles and corresponding controller current and temperature signals for a rectified sinusoidal signal with k=0.02 radians/second using the CS scheme versus MBTS scheme.

At the start of the simulation, the output strains corresponding to both the MBTS as well as the CS schemes are at 0.0571 (micron/micron). For the MBTS scheme, the current initially attains 0.94 amperes and the temperature rises rapidly. As the signal increases, the error (output strain – reference

strain) effectively reduces. This causes the current to decrease, and after a delay, the temperature follows and starts to decrease too. In response to the decrease in temperature, the strain first falls down rapidly to a value of 0.014 (micron/micron). At this point, the temperature changes direction and the minor loop is triggered, this causes the strain to remain constant along the minor loop till the time of 3.45 seconds. At this point, after an initial "bump" the output strain starts tracking the reference strain with a constant steady state error.

For the CS scheme, the current initially attains 0.3 amperes since this is the maximum limit of the current (MSSC) for wire of diameter 127 micron.

As the signal increases, the error (output strain – reference strain) effectively reduces. This causes the current to decrease, and after a delay, the temperature follows and starts to decrease too. In response to this, the strain first falls down rapidly, but slower compared to the MBTS scheme, to a value of 0.013 (micron/micron). At this point, the temperature changes direction and the minor loop is triggered, this causes the strain to remain constant along the minor loop till the time of 5.15 seconds. At this point, after an initial "bump" the output strain starts tracking the reference strain with a constant steady state error. This "bump" is similar to ones observed while testing the minor loops in Chapter 3. There is a difference between the phase fraction according to the major loop equation and the phase fraction value on the minor loop equation, when the match occurs (given the tolerance). This causes the "bump" in the behavior. Reducing the tolerance does reduce the magnitude of the bump and has been tested in simulations.

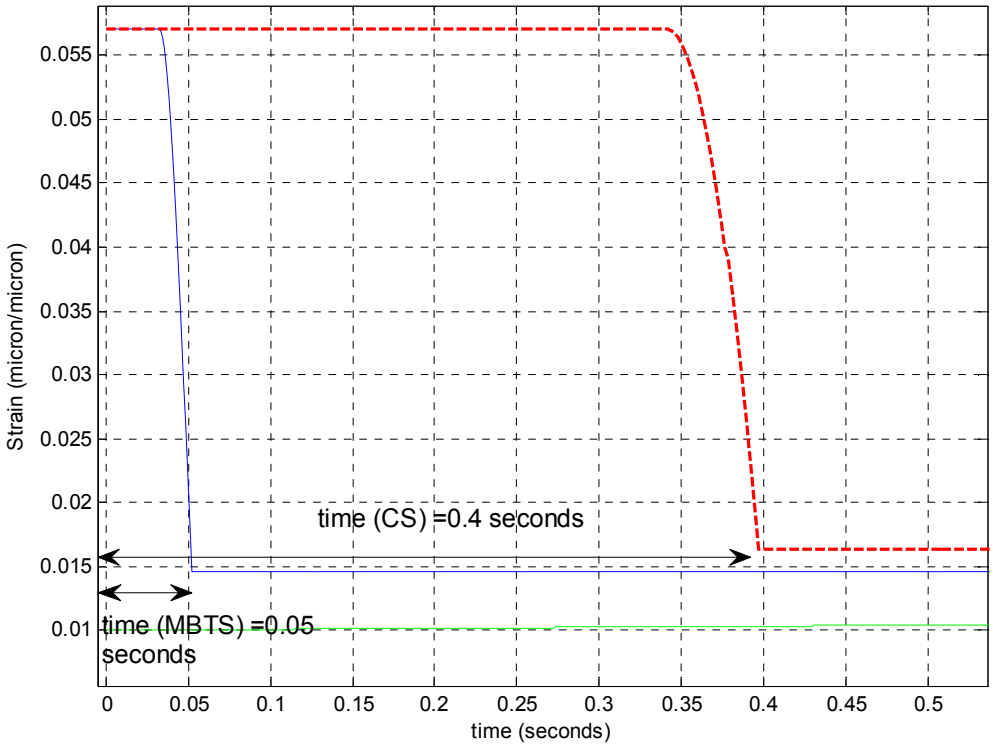

Figure 5-5: Comparison of time taken using MBTS versus CS schemes (green – reference strain, dashed red – output strain using current saturation, blue – output strain using temperature based current saturation)

In Figure 5-5 and Figure 5-6, the initial few seconds of the cycle have been zoomed to clearly observe the improvement in time observed by using MBTS scheme over the CS scheme. We observe in Figure 5-5, that the time taken by the strain corresponding to the MBTS scheme is 0.05 seconds, whereas the corresponding time for the CS scheme is 0.4 seconds. This illustrates an improvement in time by a factor of 8. However this is followed by a portion of time, during which the strain is held constant at its value, so the overall time to start tracking is 3.45 seconds (for MBTS scheme) and 5.15 seconds (for the CS scheme).

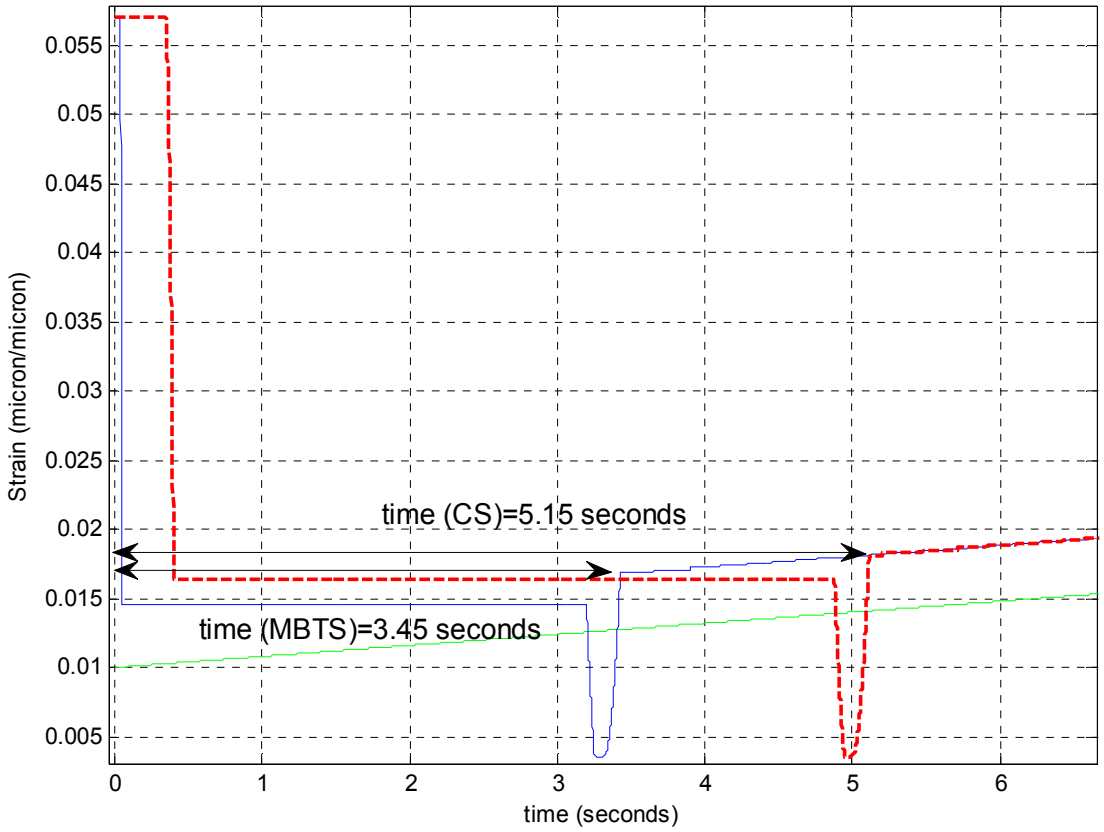

Figure 5-6: Comparison of the time taken to start tracking the trajectory using the MBTS and CS schemes (green – reference strain, dashed red – output strain using current saturation, blue – output strain using temperature based current saturation)

5.2 Summary

In this chapter, the concept of model-based temperature saturation is discussed. The simulations are provided and discussed. By monitoring temperature rather than imposing a strict saturation limit on wire current, the MBTS scheme was able to safely apply currents over three times the magnitude of the MSSC. This produced a significant improvement in the response time using the MBTS scheme, compared with the CS scheme.

Also, the above simulations were done using the P-controller. Introducing the integral (I) and derivative terms in the controller is expected to bring some different in the simulation results. Having a PD controller, the D-term would provide a phase lead, which would have an impact on speeding up the response time. Based on the values of D-terms chosen, this would impact the improvement seen on using the MBTS over the CS scheme.

Chapter 6 Conclusions and Future Work

In this thesis, a study of SMA models and control techniques was carried out. The reduced-order John Shaw model was chosen for implementation and its capability was extended to display minor loop behavior. A technique for improving the response time of SMAs was also proposed and implemented.

The reduced-order John Shaw model is suitable for control applications as compared to the original full-order John Shaw model due to its simpler structure. However the John Shaw model (both the full-order and reduced-order versions) lacks the ability to display minor loop behavior, which is important in closed-loop strain (or position) control applications since achieving a specific target strain between the two (load-dependent) extremes, requires partial phase transformation. Hence, in this work, an algorithm for modeling minor loops was proposed, implemented and tested in simulation.

The parameters required for building this model were either chosen from manufacturer data sheets or estimated from the experimental set-up in our lab. The individual modules of the model and the ability to display minor loop were tested by providing suitable input signals. The results were qualitatively analyzed. In general the following conclusions were made regarding the model simulations

- The simulations closely match what is expected experimentally. However further experimental validation is required. The value of Δs is a potential source of error that causes the stress-strain curves to shift rightwards from the expected curves.
- The tolerance used in minor loop-major loop detection has to be chosen carefully. Higher tolerance increases the magnitude of the "bump" whereas with lower tolerance, we run into the risk of not detecting the points where minor loop meets the major loop.
- The performance of the MBTS scheme was compared with the conventional CS scheme in strain tracking applications. It was found that an improvement occurs when using the MBTS scheme. More specifically the performance was tested using a reference signal given by Equation 6.1. An improvement by a factor of 8 was observed during the start of the cycle on using the MBTS scheme as compared to the CS scheme.

$$r(t) = 0.04\left|\sin(kt)\right| + 0.01 \tag{6.1}$$

6.1 Future Directions

The interesting directions to pursue further research with regard to this project are listed below.

- Perform experimental validation of the model.
- Devise techniques to do model-based hysteresis compensation. One of the techniques to do this is to invert the John Shaw model to provide a feed forward component for hysteresis compensation. This is illustrated in Figure 6-1 more clearly, whereby the proposed actuation scheme is further expanded to include an inverse model-based feed forward term. This is to compensate for the hysteretic nonlinearity for which the linear feedback law would not suffice.
- Improve the accuracy of MBTS scheme, by combining the TEM (thermoelectric model) with the TMM (thermomechanical model). This was discussed in greater detail in Chapter 2 and Chapter 4. The parameters of the TEM change in value depending on the value of the phase fraction. In order to accurately determine the temperature, the phase fraction value (from the TMM) is required to adapt the parameters of the TEM in real-time.
- Establish a state-space representation of the John Shaw model and perform a passivity study. The preliminary work on the passivity of the John Shaw model is listed in Appendix A.2

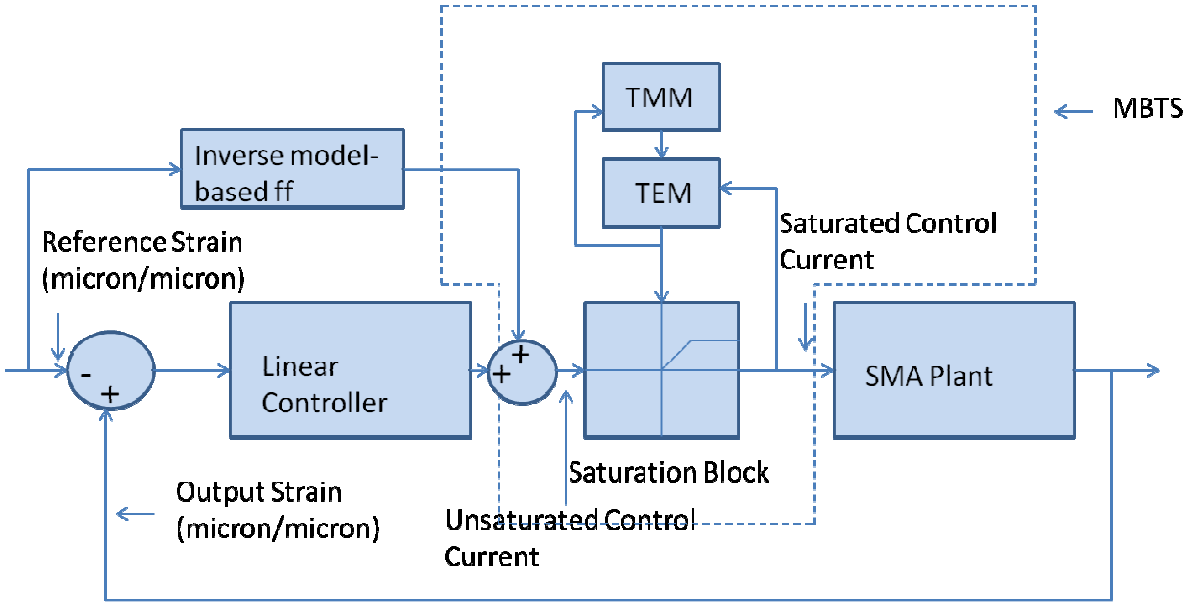

Figure 6-1: Actuator scheme with the inverse model-based feed forward component

References

1. B. Culshaw, *Smart Structures and Materials*, Artech House, Norwood, MA, 1996.
2. Dynalloy ,Inc., "Flexinol® Technical Data," http://www.dynalloy.com/TechnicalData.html (current 26 March 2012).
3. N. Ma and G. Song, "Control of shape memory alloy actuator using pulse width modulation," *Institute of Physics Publishing*, Vol. 12, 2003, pp.712-719.
4. G. Song, B. Kelly, B.N. Agrawal, P.C. Lam and T.S. Srivatsan, "Application of Shape Memory Alloy Wire Actuator for Precision Position Control of a Composite Beam," *Journal of Materials Engineering and Performance*, Vol. 9, No. 3, June 2000, pp. 300-333.
5. Glenn V. Webb, Dimitris C. Lagoudas and Andrew J. Kurdila**, "**Hysteresis Modeling of SMA Actuators for Control Applications," *Journal of Intelligent Material Systems and Structures*, Vol. 9, No. 6, 1998,pp. 432-448.
6. B. Selden, K. Cho and H. Asada, "Segmented shape memory alloy actuators using hysteresis loop control," *Smart Material Structures*, Vol. 15, 2006, pp. 642-652.
7. S. Majima, K. Kodama and T. Hasegawa, "Modeling of shape memory alloy actuator and tracking control system with the model," *IEEE Transactions on. Control Systems Technology,* Vol. 9, Jan 2001, pp. 54-59.
8. G. Song, V. Chaudhry and C. Batur, "A neural network inverse model for a shape memory alloy actuator," *Journal of Intelligent Material Systems and Structures*, Vol. 14, No. 6, 2003, 371-377.
9. G. Song, V. Chaudhry and C. Batur, "Precision tracking control of shape memory alloy actuators using neural networks and sliding-mode based robust controller," *Smart Material Structures*, Vol. 12, 2003, pp. 223-231.
10. D Grant and V Hayward, "Variable Structure Control of shape memory alloy actuators," *IEEE Control Systems Magazine*, Vol. 17, June 1997, pp. 80-88.
11. S. Seelecke and I. Muller, "Shape memory alloy actuators in smart structures – Modeling and simulation," *ASME Applied Mechanics Reviews*, Vol. 57, No. 1, 2004, pp. 23-46.
12. J.A. Shaw and S. Kyriakides, "Thermomechanical aspects of NiTi," *Journal of the Mechanics and Physics of Solids*, Vol. 43, No. 1, 1995, pp. 1243-1281.
13. R.C. Smith, S. Seelecke, M.J. Dapino, "A Unified Framework for modeling hysteresis in ferroic materials", *Journal of Mechanics and Physics of Solids*, Vol. 54, 2006, pp.46-85.

14. Ralph Smith, *Smart Materials Systems: Model development*, SIAM, Philadelphia, 2005. Copyright © 2005, Society for Industrial and Applied Mathematics, Reprinted with Permission, All rights reserved.
15. A. Bhattacharya, J.J. Amalraj and M.G. Faulkner, "Finite element modeling of phase transformation in shape memory alloy wires with variable material properties", *Smart Materials and Structures.*, Vol. 9, 2000, pp. 622-631.
16. J.J. Amalraj, A. Bhattacharya and M.G. Faulkner, "Finite-element modeling of cyclic thermal response of shape memory alloy wires with variable material properties", *Computational. Materials Science*, Vol. 17, 2000, pp. 93-104.
17. K. Sadek, A. Bhattacharya and W. Moussa, "Effect of variable material properties and environmental conditions on thermomechanical phase transformations in shape memory alloy wires", *Computational Materials Science*, Vol. 27, No. 4, June 2003, 493-506.
18. Boyd J.G. and Lagoudas D.C., "A thermodynamical constitutive model for shape memory materials. Part 1. The monolithic shape memory alloy," *International Journal of Plasticity*, Vol. 12, No. 6, 1996 , pp. 805-842.
19. X. Gao, M. Huang and L.C. Brinson, "A multivariant micromechanical model for SMAs, Part 1: crystallographic issues for single crystal model," *International Journal of Plasticity*, Vol. 16, 2000, pp. 1345-1369.
20. Z.K. Lu and G. J. Weng, "A self-consistent model for the stress-strain behavior of shape-memory alloy polycrystals," *Acta Materialia*, Vol. 46, 1998, pp. 5423-5433.
21. K. Tanaka, "A thermomechanical sketch of Shape memory effect: One-dimensional tensile behavior," *Res. Mechanica*, Vol. 18, 1986, pp. 251-263.
22. C. Liang and C. A. Rogers, "One-dimensional thermomechanical constitutive relations for shape memory materials, *Journal of Intelligent Material Systems and Structures*, Vol. 1, 1990, pp. 207-234.
23. L. C. Brinson, "One-dimensional constitutive behavior of shape memory alloys: Thermomechanical derivation with non-constant material functions and refined martensite internal variable," *Journal of Intelligent Material Systems and Structures*, Vol. 4, 1993, pp. 229-242.
24. J.A. Shaw,M.A. Iadicola and B.C.Chang, "Thermodynamics of shape memory alloy wire: modeling, experiments, and application," *Continuum Mechanics and. Thermodynamics*, Vol. 18, ,2006, pp. 83-118.

25. J.A. Shaw, C.B. Churchill, "A Reduced-Order Thermomechanical Model and Analytical Solution for Uniaxial SMA wire actuators", *Smart Materials and Structures*, Vol. 18, 2009, pp.1-21.
26. Y. Huo, "A mathematical model for the hysteresis in shape memory alloys." *Continuum Mechanics and Thermodynamics*, Vol. 1, 1989, 283-303.
27. J. Ortin, "Preisach modeling for hysteresis for a pseudoelastic Cu-Zn-Al single crystal." *Journal of Applied Physics*, Vol. 71, No. 3, February 1992, pp. 1454-1461.
28. D. C. Hughes. Piezoceramic and SMA hysteresis modeling and passivity analysis. *Doctoral Research Proposal*, Rensselear Polytechnic Institute, Troy, New York, 1994.
29. R.B. Gorbet, D.W.L. Wang and K.A. Morris, "Preisach model identification of a two-wire SMA actuator," *IEEE International Conference on Robotics and Automation*, Vol. 3, 1998, pp. 2161-2167.
30. Ktena, D.I. Fotiadis, P.D. Spanos, C.V. Massalas, "A Preisach model identification procedure and simulation of hysteresis in ferromagnets and shape memory alloys", *Physica B: Condensed Matter,* Vol. 306 , December 2001, 84-90.
31. B.J. Choi, Y. J. Lee and B. Y. Choi, "Fast Preisach Modeling method for shape memory alloy actuators using major hysteresis loops," *Institute of Physics Publishing*, Vol. 13, 2004, pp. 1069-1080.
32. R.B. Gorbet, "Control of Hysteretic Systems with Preisach Representations," *Ph.D. thesis,* University of Waterloo, Waterloo, 1997.
33. M. Achenbach and I. Muller," Simulation of material behavior of alloys with shape memory," *Archives of Mechanics*, Vol. 37, No.6, 1985, pp.573-585.
34. S. Seelecke, "Modeling of dynamic behavior of shape memory alloys," *International Journal of Non-linear Mechanics*, Vol. 37, 2002, pp. 1363-1374
35. J.E. Massad and R.C. Smith, "A homogenized free energy model for hysteresis in thin-film shape memory alloys," *Thin Solid Films*, Vol. 489, 2005, pp. 266-290.
36. S. Sajja, R.B. Gorbet, "Linear Controller benchmarks and neural network control of a Flexinol wire actuator", Technical Report TR-GM-0803, October 1, 2008.
37. K. Kuribayashi, "Improvement of the Response of an SMA Actuator Using a Temperature Sensor", *International Journal of Robotics Research*, Vol. 10, No. 1, 1991, pp.13-20.
38. R.A. Russell and R.B. Gorbet, "Improving the Response of SMA Actuators, *Proc. of IEEE International Conference on Robotics and Automation*, Vol. 3, May 1995 pp. 2299-2304,.

39. R. Featherstone and Y.H. Teh, “Improving the Speed of Shape Memory Alloy Actuators by Faster Electrical Heating”, *Experimental Robotics IX, Springer Tracts in Advanced Robotics*, Vol. 21, 2006, pp. 67-76.
40. Y.H. Teh and R. Featherstone, “An Architecture for Fast and Accurate Control of Shape Memory Alloy Actuators,” *International Journal of Robotics Research,* pp.595-611, 2008.
41. D.R. Madill and D. Wang, “Modeling and L2-Stability of a shape memory alloy control system, *IEEE Transactions on Control Systems Technology*, Vol. 6, No. 4, July 1998, pp. 473-481.
42. K. Ikuta, M. Tsukamoto, and S. Hirose, “Shape memory alloy actuator system with electrical resistance feedback and application foractive endoscope,” *in Proc. 1988 IEEE International. Conference of. Robotics Automation*, 1988, pp. 427–430
43. Mauro Dolce, Donatello Cardone, “Mechanical behaviour of shape memory alloys for seismic applications 2. Austenite NiTi wires subjected to tension”, *International Journal of Mechanical Sciences*, Vol. 43, Issue 11, November 2001, pp. 2657-2677.
44. H. Funakubo, *Shape Memory Alloys*, pp. 74-85, University of Tokyo, Gordon and Breach Science Publishers, 1984.
45. S. Sajja, R.B. Gorbet and E.G. Kubica, “Model based Techniques to improve the performance of SMA actuators”, Technical Report TR-GM-1001, May 31, 2010.
46. S. Valadkhan, K.A. Morris and A. Khajepour, “Passivity of magnetostrictive materials*,” SIAM Journal of applied mathematics,* Vol. 3, 2007, pp. 667-686, 2007.
47. R.B. Gorbet, K.A. Morris and D.W.L. Wang, “Passivity based stability and control of hysteresis in smart actuators,” *IEEE Transactions on Control Systems Technology*, Vol.9, 2001, pp. 5-16.
48. H. Khalil, *Nonlinear Systems*, Prentice Hall, 2002.

Appendix A

A.1 Analysis of the case of elastic moduli of austenite and martensite not being equal

As discussed previously, the assumption that the elastic moduli of austenite and martensite are equal is a very limiting assumption. Further from the values calculated from our set-up, they were found to be significantly different.

In order to investigate the impact of this assumption on the model, the reduced-order model equations are modified to account for the ΔE term, based on the original full-order model equations discussed in [22]. Since it is independent of the material modulus, the heat equation (Equation A.1) remains unchanged. The phase kinetics equation becomes a quadratic in ξ (Equation A.2). Though not verified at this point, it is expected that one solution will be eliminated since the phase fractions are always between 0 and 1.

$$\rho A L c_0 \dot{T} = \rho A L \left(\operatorname{sgn}\left(\dot{\xi}\right)\mu_c - T\Delta s \right)\left(\dot{\xi}\right) - h\pi d L\left(T - T_a\right) + P_e \tag{A.1}$$

$$\mu_c = \frac{\beta F(t)}{\rho A} - \frac{\Delta E}{2\rho}\left(\frac{\sigma}{E_A + \xi\Delta E}\right)^2 + \left(T - T_R\right)\Delta s \tag{A.2}$$

$$F(t) = \left(E_A + \xi\Delta E\right)\left(\varepsilon - \xi\Delta E\right) \tag{A.3}$$

In order to understand the impact of $\Delta E=0$ on Equation A.2, it is re-ordered and parameter values are substituted from Table 2 at the boundary conditions of $\xi=1$ and $\xi=0$. For ease of analysis, ΔE is taken as $0.5E$, a value which is relatively close to what is from the identified parameters. Stress, σ, is assumed to be 200MPa. The results are shown in Equations A.4 and A.5, respectively. It is found that the first term is order of 100 (100 for $\xi=1$ and 50 for $\xi=0$) times larger than the second term (with the ΔE term in it) and hence neglecting this term is justified.

For ξ=1:

$$\mu_c = \frac{\sigma}{\rho}\left(\beta - \frac{0.5\sigma}{(1.5)^2 2E_A}\right) + (T - T_R)\Delta s$$
$$\mu_c = \frac{\sigma}{\rho}\left(\frac{5.35}{100} - \frac{4.6057}{10000}\right) + (T - T_R)\Delta s \quad \text{(A.4)}$$

For ξ=0:

$$\mu_c = \frac{\sigma}{\rho}\left(\beta - \frac{0.5\sigma}{2E_A}\right) + (T - T_R)\Delta s$$
$$\mu_c = \frac{\sigma}{\rho}\left(\frac{5.35}{100} - \frac{1}{1000}\right) + (T - T_R)\Delta s \quad \text{(A.5)}$$

A.2 Passivity of the John Shaw model

Establishing the stability of closed loop systems employing hysteretic systems is often a challenge. This is because of the dependence of hysteresis on external conditions and the nonlinear nature of such systems [46]. Passivity is one of most useful techniques in demonstrating stability of nonlinear systems [46].

At the core, passivity can be visualized as an input-output property. It states that the sum of the energy taken out of the system and the energy stored in the system is no greater than the energy supplied to the system.

Based on a brief literature survey, two significant works [46, 47] were found in this regard. The first one [46] established the passivity of the Preisach model (a significant model used for modeling smart materials) using both input-output stability and state-space approach. The second one [47] established the passivity of magnetostrictive (a different smart material) actuators using a model-independent approach. The Helmholtz energy was chosen to be a suitable storage function. Based on the above mentioned works and the theory presented in Khalil [48], the process of establishing passivity of the John Shaw model can be seen as composed of the following steps:

- Choosing the input-output pairs. In this case, the choice is fairly straightforward with the input being the applied power and output is strain. The corresponding product corresponds to power consumption, since strain itself is dimensionless.
- Develop the state-space representation of the model. This is the challenging part of the whole problem, since the equations are not readily available in the state model form. Two possible sets of states are as follows: Input $u=P_e$, output $y= \varepsilon (t)$ and the states $\{x_1,x_2\} =\{\xi,T\}$. Input $u=P_e$, output $y=\varepsilon(t)$ and the states $\{x_1,x_2,x_3\}=\{ \xi,d\xi/dt,T\}$.
- Choose an appropriate Lyapunov function $V(x)$ (storage function in this case) which is continuously differentiable positive semi-definite. A good starting function is the Helmholtz energy itself, since it is equation from which the constitutive equations of the model are developed. However it is essential to establish and state the external conditions under which it is continuously differentiable and positive semi-definite.
- Establish the condition for passivity.

$$u^T y \geq \frac{\delta V}{\delta x}\left(f\left(x,u\right)\right) \tag{A.6}$$

Currently the challenge of this analysis is in finding a state-space representation for the John Shaw model. Based on some preliminary work done in this regard, it was found to be a non-trivial task and remains a part of the future work.