A PHENOMENOLOGICAL MODEL OF SHAPE MEMORY ALLOYS INCLUDING TIME-VARYING STRESS

by

Arati Pai

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, 2007

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 this thesis may be made electronically available to the public.

Arati Pai

Abstract

Shape memory alloys (SMAs) are metallic materials, which have two main stable crystalline phases: austenite, a high temperature phase and martensite, a low temperature phase. Austenite and martensite each have unique physical and mechanical properties, and transformation between these phases enables two effects known as the shape memory effect (SME) and superelasticity. When a material that displays the SME is plastically deformed at low temperature, a heat input will cause the SMA to return to its original shape before the deformation. At higher temperatures, the material displays an effect called superelasticity, where strains of up to 10% are recoverable. These characteristics of SMA allow for significant amounts of strain recovery, and enable the design of SMA actuators. The temperature in an SMA actuator is generally controlled by resistive heating, also know as joule heating, and the strain recovery capabilities are used to do work on a load, thereby creating an electro-mechanical actuator. SMA actuators have attractive properties such as high energy density, smooth and silent actuation, reduced part counts compared to traditional alternatives, and scalability down to the micromechanical level.

The phase transformation in SMA actuators, however, is highly non-linear. Therefore, the use of SMA as actuators, for example in positioning systems, benefits from the development of good models to predict and control the materials. The goals of this work are to develop a model suitable for real-time implementation, and that reproduces the observed behaviour of SMA actuators. The model is then inverted and used to develop a model-based controller, used in conjunction with traditional PID control to improve the precision and robustness of SMA actuators.

The modelling portion of this work consists of the development of a phenomenological SMA model. The forward model is split into three blocks: a heating block, a phase kinetics block and a mechanical block. Since joule heating is commonly used in SMA actuators to bring about an

increase in temperature, the heating block presents equations to convert a current input into the temperature of the wire. The phase kinetics block equations convert the calculated temperature and applied stress to the fraction of martensite present in the SMA. Finally, the mechanical model calculates the strain in the material from the martensite fraction and the applied stress. Once the model equations are presented, experimental verification tests are shown to compare physical SMA behaviour with that predicted by the model. Each of the blocks of the forward model are then inverted in order to be used as a feedforward linearizing controller.

The control section of this thesis deals with the response of two common types of SMA actuators: a constant force SMA actuator and a spring-biased SMA actuator. The response of the system to step and sinusoidal signals with period of 5 seconds is investigated using two types of controllers: a traditional PI controller and the inverse-model controller in feedforward with a PI controller in feedback. Additionally, the robustness of the system is investigated through the response of the system to transient and sinusoidal stress disturbances. The disturbance rejection is investigated on a constant force actuator both with and without the presence of a force sensor.

Acknowledgments

I would first and foremost like to thank the Lord my God without whom I am truly nothing. I am most grateful of the incredible support and guidance of my supervisor, Professor Robert B. Gorbet, who not only gave me invaluable insight on my research but also on how to be a better engineer (and always think about units!).

I would like to thank the Canadian Space Agency for allowing us to use their equipment for some of our experiments. I am grateful to Professors David Wang and Amir Khajepour for reviewing this thesis.

I would like to thank the Ontario government and the University of Waterloo for their financial support through the Ontario graduate scholarship, the president's scholarship and the Faculty of Engineering scholarships.

My appreciation also goes to my fellow group members and classmates who helped me keep my sanity during the past two years.

Last but by no means the least, I am extremely grateful for the support and advice extended to me by my family.

Contents

List of Tables

List of Figures

Chapter 1

Introduction

In recent years there has been a trend towards the development of intelligent actuators to be used in various applications such as aerospace, automotive and medical to mention a few. Shape memory alloys (SMAs) are metallic materials displaying superior characteristics in their application as actuators. SMA actuators have attractive properties such as high energy density, smooth and silent actuation, reduced part counts compared to traditional alternatives, and scalability down to several microns. These properties have been exploited in various complex applications in the past ten years.

1.1 Motivation

The actuation mechanism in SMA actuators is enabled by a thermally-induced phase transition inherent to the material. The phase transition, however, is hysteretic in nature. The complexity of the transition and the coupling between transformation temperatures, stress, strain and mechani-

cal properties, for example, add to the non-linearity in SMAs. The use of SMA as actuators, for example in positioning systems, benefits from the development of good models to predict and control the highly non-linear behaviour inherent in the materials. The purpose of this work is to develop a model that reproduces the salient features of SMA behaviour, and can be implemented in real-time for control purposes. The proposed model is subsequently used in control loops for precise and robust control of the SMA.

1.2 Goals

The goals of this research are as follows:

1. Develop a model which is suitable for real-time implementation, reproduces the observed behaviour of SMA actuators including minor loop behaviour, and can account for time varying stress.

2. Derivation of inverse model.

3. Compare inverse model-based control with traditional PID control on the following control problems:

    - Step and sinusoidal strain tracking on common SMA actuator configurations, namely: a constant force SMA actuator, a spring-biased SMA actuator and a variable force SMA actuator where the variable force is different from a spring force.
    - Rejection of transient and sinusoidal stress disturbances on a constant force SMA actuator using two control configurations, both with and without the presence of a force sensor.

1.3 Overview

This thesis is organised as follows:

- Chapter 2 provides the background information required to understand shape memory alloys and their properties.
- Chapter 3 presents the equations defining the SMA model and its inverse. The plant model is broken down into a heating model, a phase kinetics model and a mechanical model. Each model is presented together with its inverse. Additionally, the equations for a spring-biased SMA actuator, are presented.
- Chapter 4 presents the quantitative verification of the model via simulation results. The simulations show that the model is able to reproduce key phase kinetic and mechanical SMA behaviour.
- Chapter 5 presents the experimental verification of the model. Details of the experimental set up and procedure are defined followed by a presentation of simulation and experimental results.
- Chapter 6 presents experimental results for the controls investigation into strain tracking and disturbance rejection.

Chapter 2

Background

This chapter provides fundamental theories of shape memory alloys required to understand their physical behaviour. Additionally, SMA modelling history is also presented.

2.1 Shape Memory Alloys

Shape memory alloys are metallic materials that, under particular external conditions, display two effects known as the shape memory effect (SME) and superelasticity [1]. When a material that displays the SME is plastically deformed at low temperature, a heat input will cause the SMA to return to its original shape before the deformation [2]. At higher temperatures, the material can be 'reversibly deformed up to approximately 10% of its original length' [2] displaying an effect called superelasticity. These characteristics of SMA allow for significant amounts of strain recovery. Based on the temperature of the system, the strain is recovered either through the SME by application of heat or through superelasticity upon unloading [3].

The most commonly used SMA for actuator applications is nickel-titanium (NiTi), traditionally called NiTiNOL, after the Naval Ordinance Laboratory at which it was discovered by W.J. Buehler ([4], [2]). Since then, a number of other alloys that display the SME have been discovered including copper based alloys such as copper zinc aluminium (CuZnAl), copper aluminium nickel (CuAlNi), as well as alloys of metals such as gold, silver, indium and platinum. NiTi alloys and the copper based alloys are the two most commercially viable options. The nickel-titanium alloy in particular has a resistivity that is about 80 times higher than copper making it attractive for actuator applications, which typically use joule heating to bring about an increase in temperature. Additionally, NiTi has high tensile strength upto 1000 MPa, ability to recover strains in the order of 8%, elongation to failure reaching 50% and significant internal damping [5, 1, 6]. Flexinol® is a type of NiTi wire, manufactured by Dynalloy Inc, that can be used for high cyclic applications such as in door latches and locking actuators [7]. Flexinol was used in all experiments in this research work.

The following sections presents SME, microscopic and macroscopic properties of SMA, hysteresis characteristics, stress-strain characteristics and superelasticity.

2.2 Shape Memory Effect

The shape memory effect (SME) refers to the ability of a material to remember a predefined shape after deformation [8] with an increase in temperature. In order for a SMA to display the SME, it has to go through a process known as annealing. Once constrained in the desired "remembered" shape, the alloy is heated to a temperature in the range of 450^oC - 550^oC for about fifteen minutes and then air cooled or water quenched [9].

2.3 Microscopic Properties

Shape memory alloys display three main phases: austenite, martensite and R-phase. In this work, the R-phase will not be considered because its presence is negligible in materials that exhibit strong SME [10] such as Flexinol. Austenite is the phase that exists at higher temperatures and is the stronger of the two. Martensite is the weaker, lower temperature phase. It is the characteristics of the austenite and martensite phases that give SMAs the ability to perform shape recovery after deformation. The section below discusses these phases in greater detail.

2.3.1 Austenite

Austenite has a very stable crystalline structure that comprises a body-centered cubic (BCC) lattice [6] shown in Figure 2.1[1]. In the body-centered structure, there are Ti atoms at each vertex of a cube and Ni atoms at the center. It is this stable structure that gives austenite its strength properties.

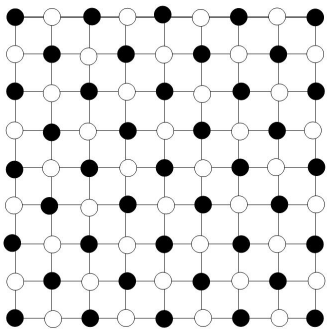

Figure 2.1: Body-centered cubic structure of austenite

[1]Note that the figure shows the 2D schematic of the austenite lattice

2.3.2 Martensite

Martensite has either a monoclinic crystal lattice [6] or rhombic crystal lattice [8] depending on the applied stress. The monoclinic lattice occurs at low/no stress and is formed when the shape memory alloy cools to low temperature[1] from high temperature. Therefore, microscopically, the low temperature monoclinic structure is formed by conversion from the high temperature body-centered cubic structure of austenite. When compared to austenite, the monoclinic lattice has a low degree of symmetry and therefore several energetically equivalent configurations, called variants, can be formed [8, 6]. One such variant is formed by *twinning*. Figure 2.2 shows a twinned variant. The twin boundary in the figure is called as such because an atom situated on a boundary sees mirror views on either side of the boundary. Twinning is a method of accommodating the transformation from the parent phase, in this case austenite, to a different phase while preserving the external macroscopic dimensions [8, 6]. Due to the presence of twin boundaries, the martensite in this state is called *twinned martensite*. The twin boundaries are low energy and move easily with the application of stress. When stress is applied to twinned martensite, the martensite is said to *detwin* because the twin boundaries move to accommodate the applied stress and the lattice structure converts to a rhombic structure. The martensite in this state is called *detwinned martensite* [8, 6]. Figure 2.3 shows the transformation from twinned to detwinned martensite.

[1] For the type of martensite used in this work, the room tempertature is sufficiently low for martensite formation.

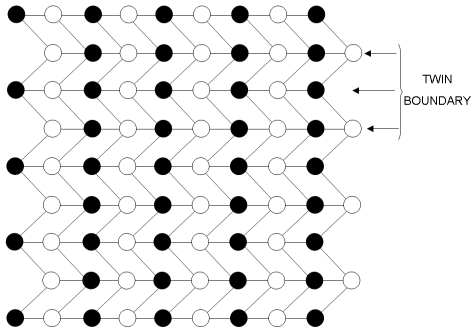

Figure 2.2: Twinned martensite

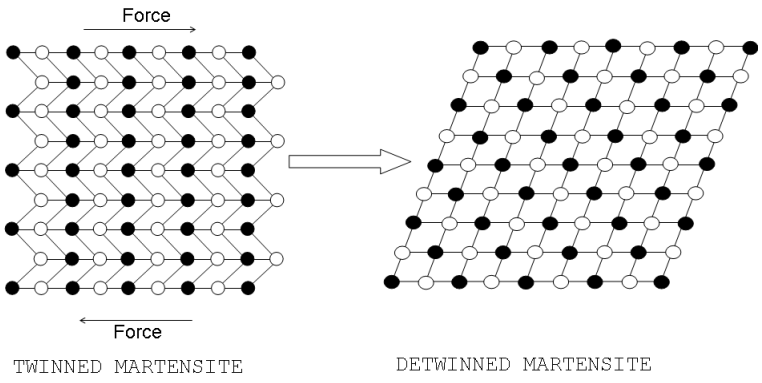

Figure 2.3: Transformation from twinned to detwinned martensite

2.4 Macroscopic Properties

An important characteristic of shape memory alloys is that although austenite and twinned martensite have different microscopic characteristics, they have identical macroscopic shapes. Figure 2.4 shows the microscopic and macroscopic views of the SMA. The foundation for the shape memory effect lies in the fact that although martensite can form in several different variants, austenite has only one crystal variant, BCC. Recall that the formation of twin boundaries is a result of stress accommodation in order to maintain local shape. Therefore, austenite and twinned martensite have the same macroscopic shape. When twinned martensite is loaded, it converts to detwinned martensite with a rhombic structure. The orientation of the rhombus depends on the direction of the applied stress. Macroscopically, loading is accompanied with deformation. On heating, the conversion has only one path, that is, to the body centered cubic structure of austenite and the conversion to austenite is accompanied by a macroscopic recovery of the original shape. The key to this recovery is the congruence between the macroscopic dimensions of austenite and twinned martensite [8, 6].

2.5 Hysteresis Characteristics

Consider an SMA wire with no stress applied to it. At room temperature, the SMA is at its low temperature phase, martensite. Moreover, since there is no stress, the SMA is in the twinned martensite state. An increase of temperature will cause it to transform to austenite. The temperature at which austenite starts forming is called the austenite start temperature, A_s. A further increase of temperature will cause more austenite to form and at a certain temperature called the austenite finish temperature, A_f, the material is entirely austenite. Similarly, if the SMA starts

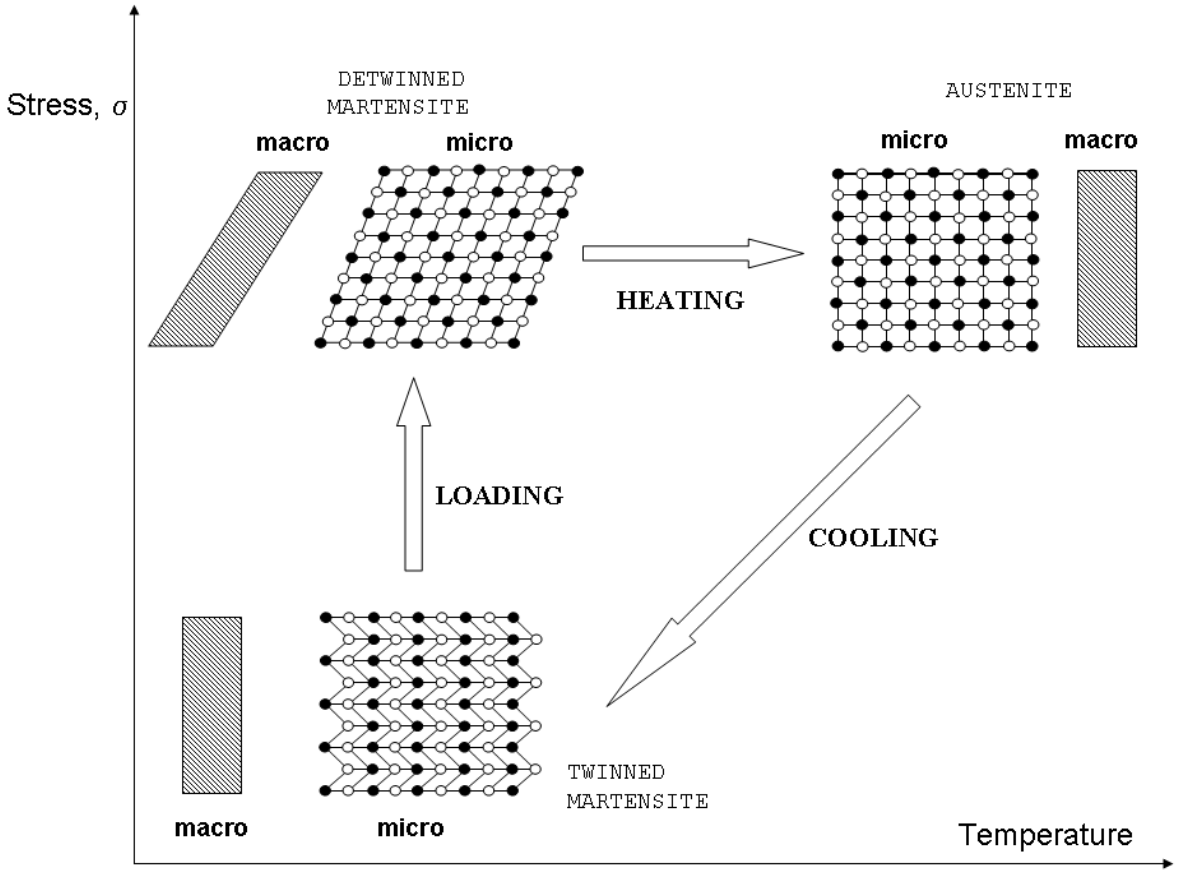

Figure 2.4: Macroscopic properties of SMA

at a fully austenite state and the temperature is decreased, then the material will transform to martensite. Martensite variants form at the martensite start temperature, M_s, and the material is fully martensite by the martensite finish temperature, M_f. The transformation from martensite to austenite via heating and from austenite to martensite via cooling follow two different paths, thereby forming a hysteresis loop as shown in Figure 2.5. The y-axis of Figure 2.5 is the martensite fraction, R_m. When the material is fully martensite, $R_m = 1$ and when the material is fully austenite, $R_m = 0$. For R_m between 1 and 0, the material is a combination of martensite and austenite. The transformation temperatures, M_s, M_f, A_s, A_f are difficult to measure because it is difficult to know when exactly the martensite or austenite molecules form in the SMA. Traditionally, to quantify the transformation temperatures, the hysteresis loop is fitted with a quadilateral and M_f, M_s, A_s, A_f are denoted as the temperatures at each of the vertices of the quadilateral as shown in Figure 2.5.

Temperature reversals in SMA hysteresis cause immediate R_m reversals. Consider an SMA material in martensite phase. If the temperature is increased so as to convert the SMA completely into austenite and then the material cooled so as to transform completely back to martensite, the ensuing hysteresis loop is called the major loop. If the input temperature is partially cycled between the transformation temperatures, internal branching behaviour results. Consider an increasing temperature profile such that the SMA, initially martensite, starts to convert to austenite, but before the material is fully austenite, the temperature profile switches to cooling. In this case, the SMA will follow a minor loop branch. A similar minor loop branch will result when the temperature profile is switched from cooling to heating before the material is fully martenstite, given that it was originally austenite. Additionally, it is assumed that the major and minor loops are closed. The major and minor loop behaviour is shown in Figure 2.6.

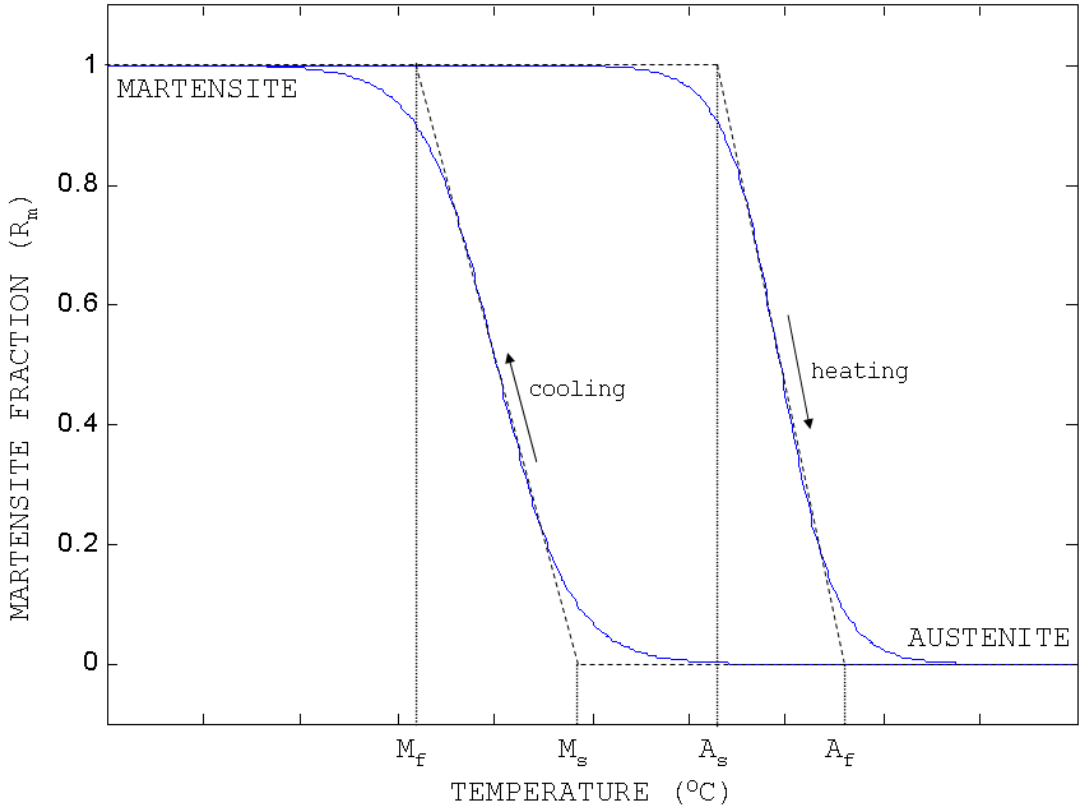

Figure 2.5: Hysteresis loop in SMA showing transformation temperatures

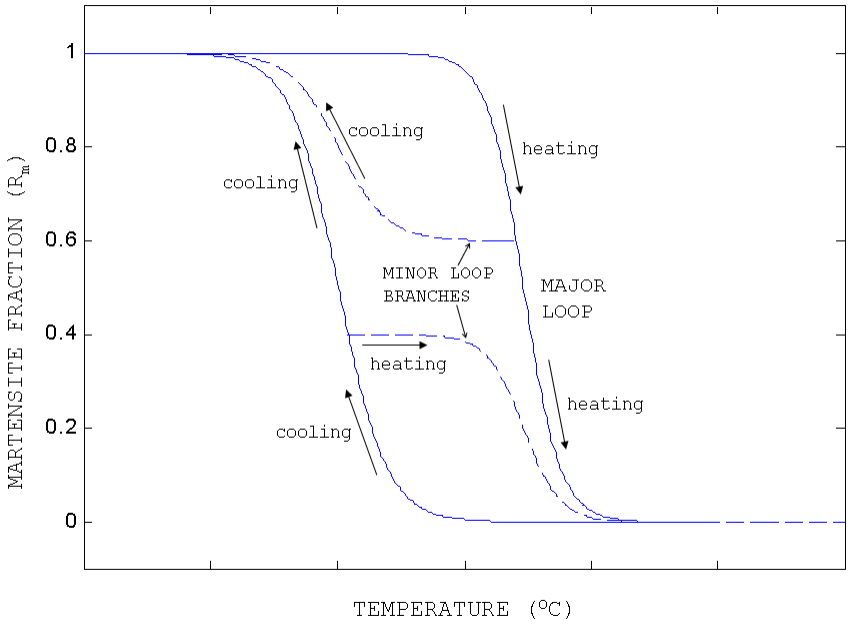

Figure 2.6: Hysteresis loop in SMA showing minor loop branching

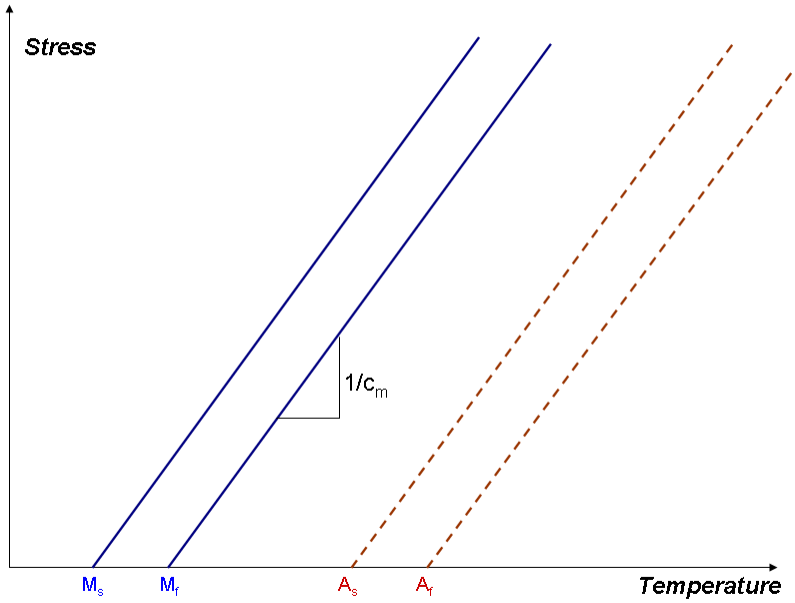

Figure 2.7: Effect of stress on SMA transformation temperatures

2.5.1 Effect of Stress on Hysteresis Curves

The application of stress on the SMA wire causes the transformation temperatures to shift to higher values [11] as shown in Figure 2.7, where $1/c_m$ is the slope of the stress - temperature curves. This shifting of the transformation temperatures causes the hysteresis loops to translate to higher temperatures with increasing stress as shown in Figure 2.8.

The martensite fraction, R_m is a difficult quantity to measure. Therefore, the hysteresis curves are traditionally plotted with strain, ϵ, on the y-axis. The shape of the resulting hysteresis curve is similar to using R_m on the y-axis. When stress is applied to the SMA, however, the total strain in the material also increases with increasing stress. Therefore, in addition to the $\epsilon - T$

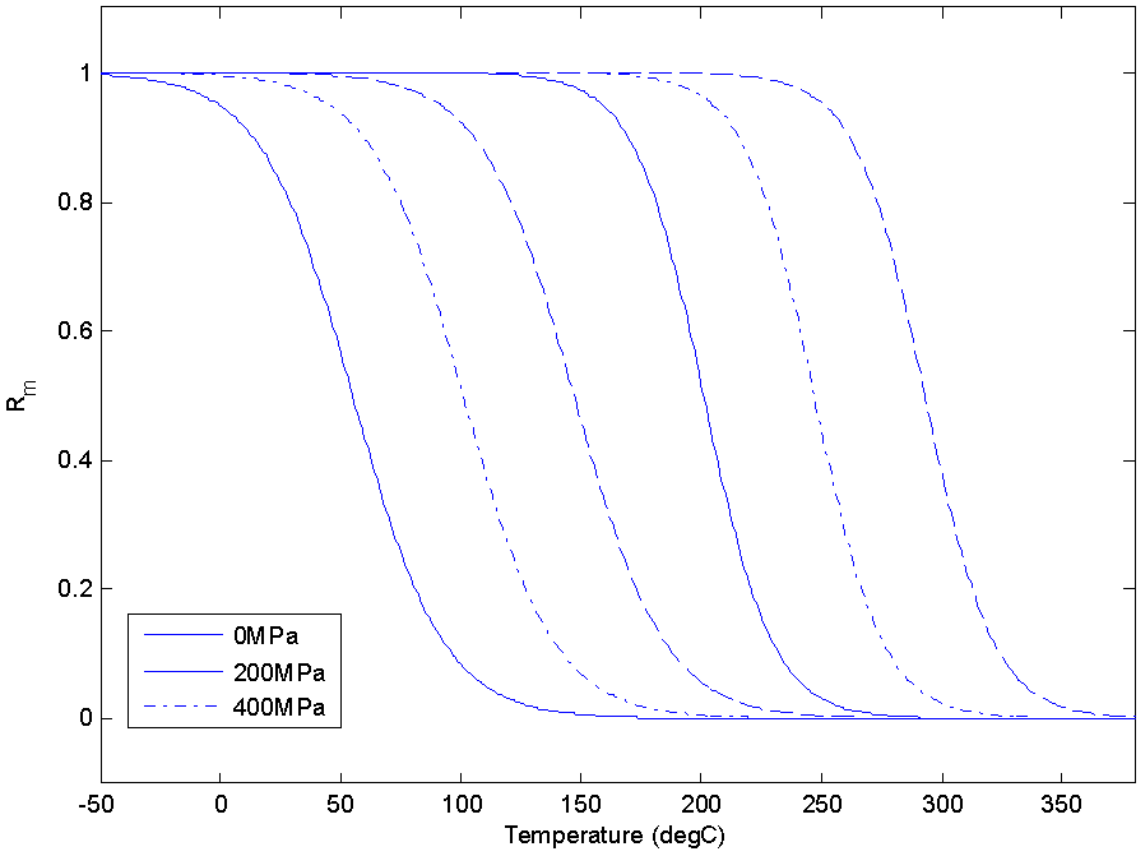

Figure 2.8: Hysteresis loops shifting with increasing stress

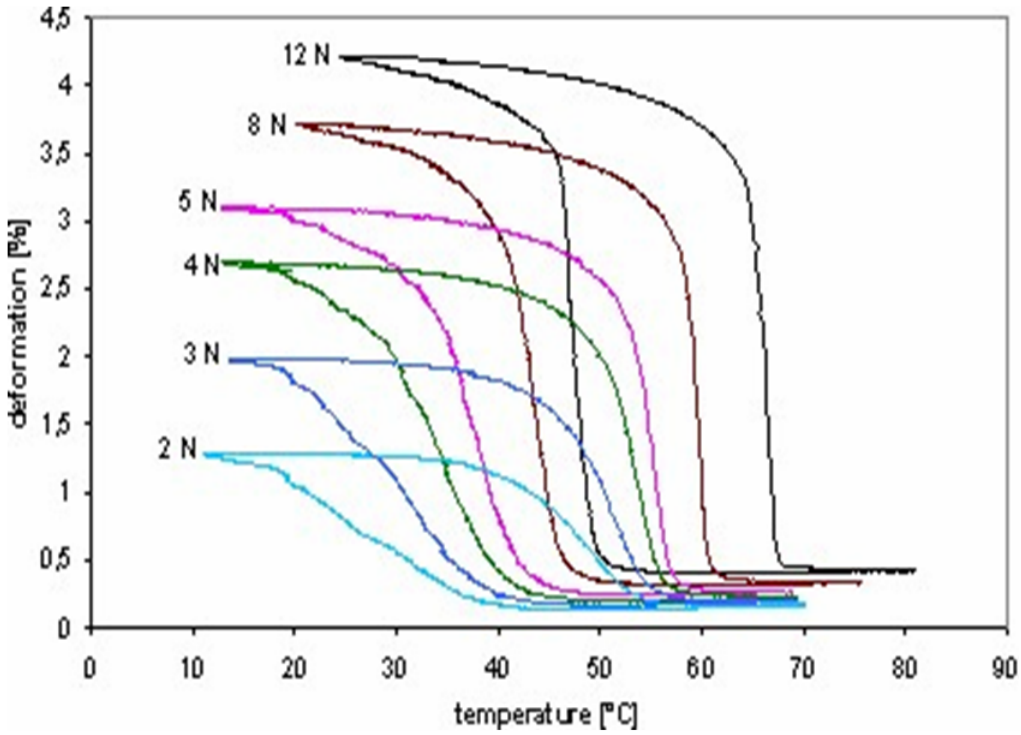

Figure 2.9: Experimental hysteresis loops shifting with increasing stress [12]

curves shifting to higher temperatures, they also increase in height. Figure 2.9 shows experimental hysteresis curves taken from [12], plotted as strain (or deformation) versus temperature with increasing stress.

2.6 Stress - Strain Characteristics

2.6.1 Martensite

Figure 2.10 shows that the martensite stress strain curve has three distinct portions. For strains from 0% to about 1%, the material is twinned martensite. The stress strain relationship is linear and elastic up to 1% strain, where the first yield point occurs. After this yield point, the material starts to detwin and there is a mixture of twinned and detwinned martensite in the sample. The second yield point occurs at about 5% strain. At this point, the material is fully detwinned and atomic bonds within the rhombic structure can be stretched [9], giving a second linear elastic region. If more stress is applied, beyond the final yield point, atomic bonds break and irrecoverable deformation occurs.

The unloading behavior of martensite is as follows: when the material is fully twinned martensite ($\epsilon < 1\%$), the unloading is elastic. When the material is fully detwinned martensite ($\epsilon > 5\%$), such as at point ***b*** in Figure 2.11, the unloading follows the curve labelled **1** in the figure which may have an elastic modulus different from the loading modulus for detwinned martensite. When the material is a combination of twinned and detwinned martensite ($1\% < \epsilon < 5\%$), such as at point ***a*** in Figure 2.11, unloading follows the curve labelled ***2*** which has a modulus that is proportional to the fraction of twinned and detwinned martensite present in the material at the point of unloading. Both of the latter types of unloading result in residual strain in the material upon complete unloading. This residual strain can be recovered by heating the material.

Reloading occurs when the stress changes direction from decreasing to increasing. Reloading behaviour occurs during cyclic loading and unloading. When the stress begins to increase after a

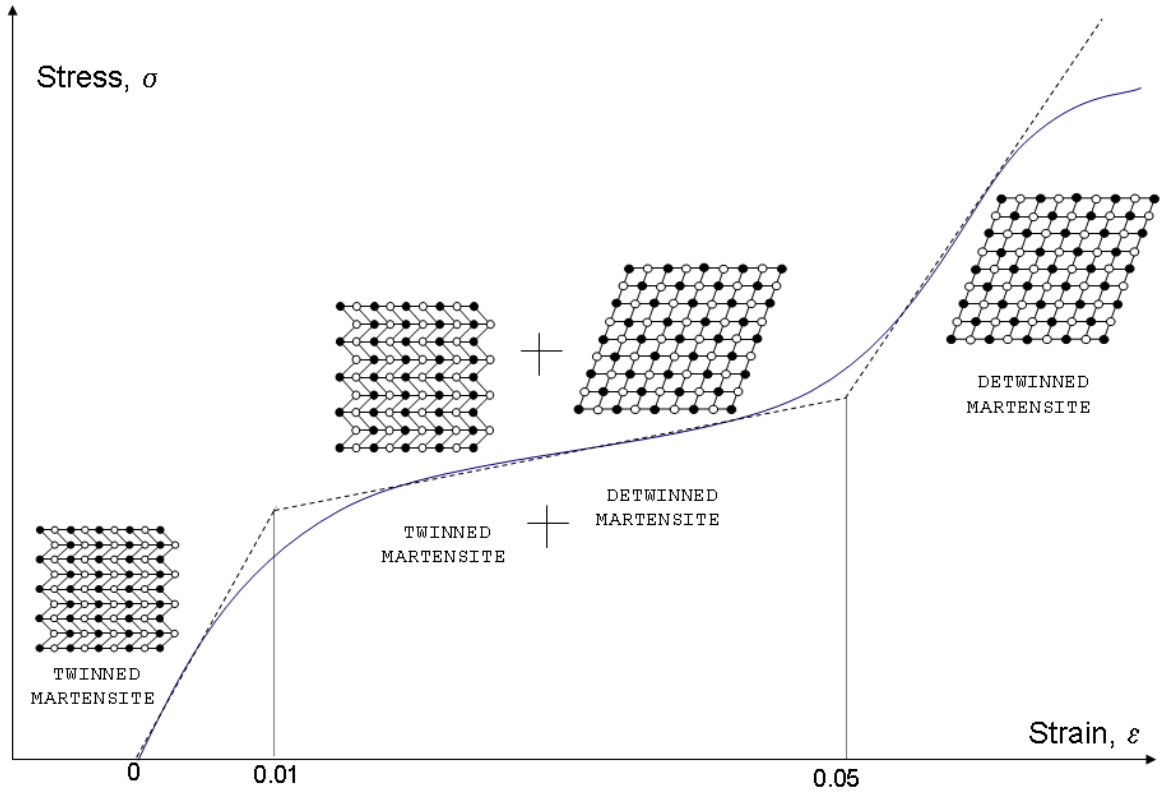

Figure 2.10: Martensite stress - strain curve showing loading behaviour

period of unloading, the reloading follows the same curve as the unloading until it joins the original loading curve. If reloading were to occur when the martensite is fully twinned, reloading is identical to normal loading since the unloading and loading curves for twinned martensite overlap. If reloading were to occur after initial unloading from either the martensite plateau or when martensite is fully detwinned, then the reloading behaviour follows the unloading curve until position ***a*** or ***b*** in Figure 2.11 respectively, after which reloading proceeds as normal loading.

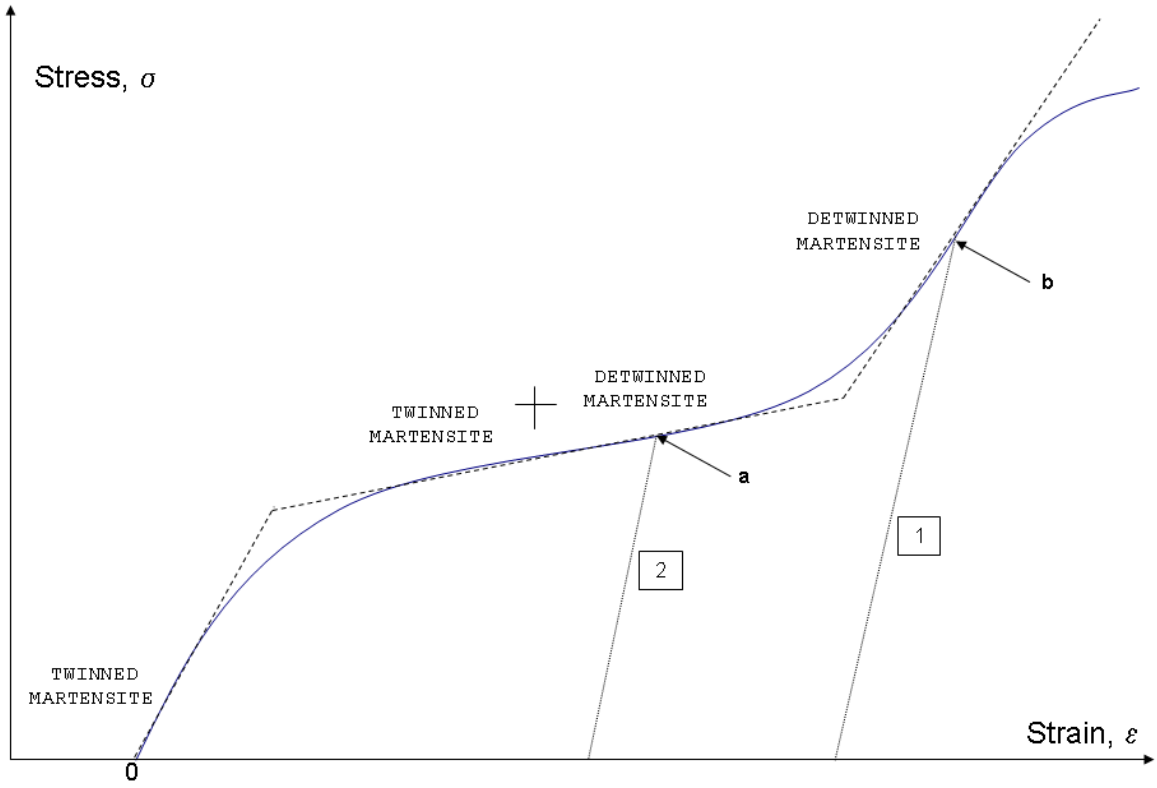

Figure 2.11: Martensite stress - strain curve showing unloading and reloading behaviour

2.6.2 Austenite

Figure 2.12 shows that the austenite stress strain curve, similar to the martensite curve, has three distinct portions. For strains from 0% to approximately 0.8%, the material is purely austenite. The stress strain relationship is linear and elastic until the first yield point. After this yield point, the austenite begins to convert to stress induced martensite (SIM). Stress induced martensite has the same chacteristics as detwinned martensite. The name stress induced martensite is used to differentiate it from thermally induced martensite that is formed by cooling the austenite below the martensite finish temperature. The material is a combination of austenite and SIM, and said to be on the 'SIM plateau', until the second yield point after which a further increase in stress causes the material to be entirely SIM. Unloading in austenite before the first yield point is elastic. When austenite is unloaded after the first yield point, it displays superelasticity (also called pseudoelasticity), discussed below.

2.6.3 Superelasticity

When the SMA is held at a temperature greater than the austenite finish temperature, A_f, such that the material is completely austenite, and is loaded beyond the second yield point and then subsequently unloaded, it follows the curve labeled ***1*** in Figure 2.13. It is common to model this recovery as complete ([13], [14] e.g.), although in practice some small residual strain is sometimes observed. The SMA at high temperature therefore displays full strain recovery and macroscopically behaves as if it were a material with high elasticity like twinned martensite. This behaviour is known as the superelastic (or pseudoelastic) effect.

If the material is unloaded from the SIM plateau (at point **b**) in Figure 2.13, it still behaves

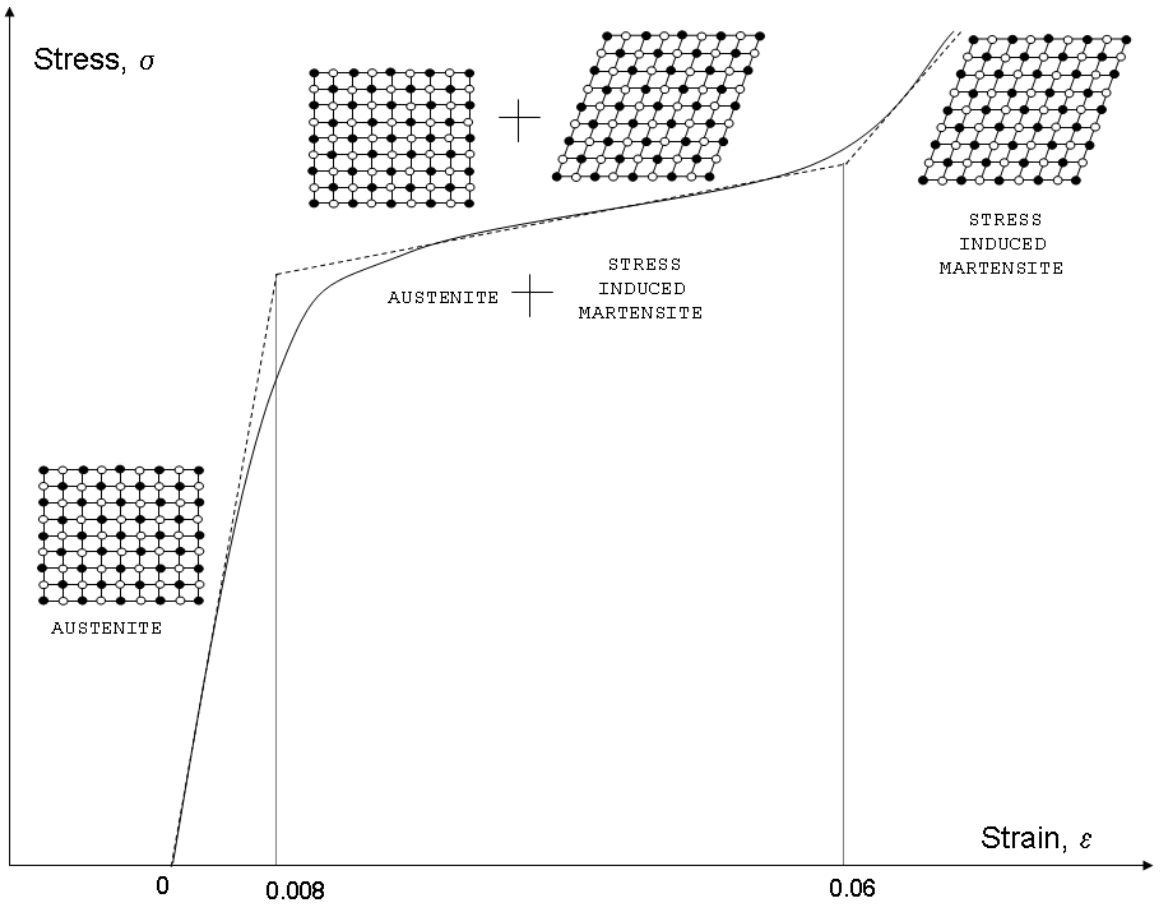

Figure 2.12: Austenite stress - strain curve showing loading behaviour

according to the superelastic effect by displaying an unloading modulus (curve ***2***) that is proportional to the amount of austenite and SIM present at the time of unloading, until it intersects with the pseudoelastic loop at point **a**, after which unloading proceeds as before.

Reloading behaviour in austenite is similar to that in martensite. If reloading were to occur when the material is pure austenite, it proceeds as normal loading. If reloading were to occur when the material is on some portion of the pseudoelastic loop, such as point **c** in Figure 2.13, after unloading either from the SIM plateau or when fully SIM, then the reloading proceeds between points **c** and **d** as is shown in the figure with a reloading modulus that is proportional to the amount of austenite and SIM present at the reloading instant. After **d**, the reloading proceeds as normal loading. Several loading and unloading cycles within the major pseudoelastic loop cause the formation of minor pseudoelastic loops as shown in Figure 2.14 from [15].

When the temperature is increased above the austenite finish temperature, A_f, the superelastic loops shift to higher stresses. Figure 2.15 from [13] shows experimental and simulated results of the temperature sensitivity of the superelastic loops.

2.6.4 Stress, Strain, Temperature Relationship in SMAs

Section 2.5 discussed the martensite fraction, R_m, or strain, ϵ, versus the temperature, T, of the SMA. Section 2.6 discussed the stress, σ, versus the strain, ϵ, in the SMA. This section describes the link between the stress, strain and temperature of the SMA. Figure 2.16 shows a three-dimensional SMA characteristic curve taken from [16]. Starting at twinned martensite at 0 stress and strain, the material is loaded up to position *1* and unloaded to position *2*. The SMA displays some residual strain at point *2*. The material is then heated and it follows the heating hysteresis loop from *2* to *5*. The SMA experiences complete strain recovery through the heating

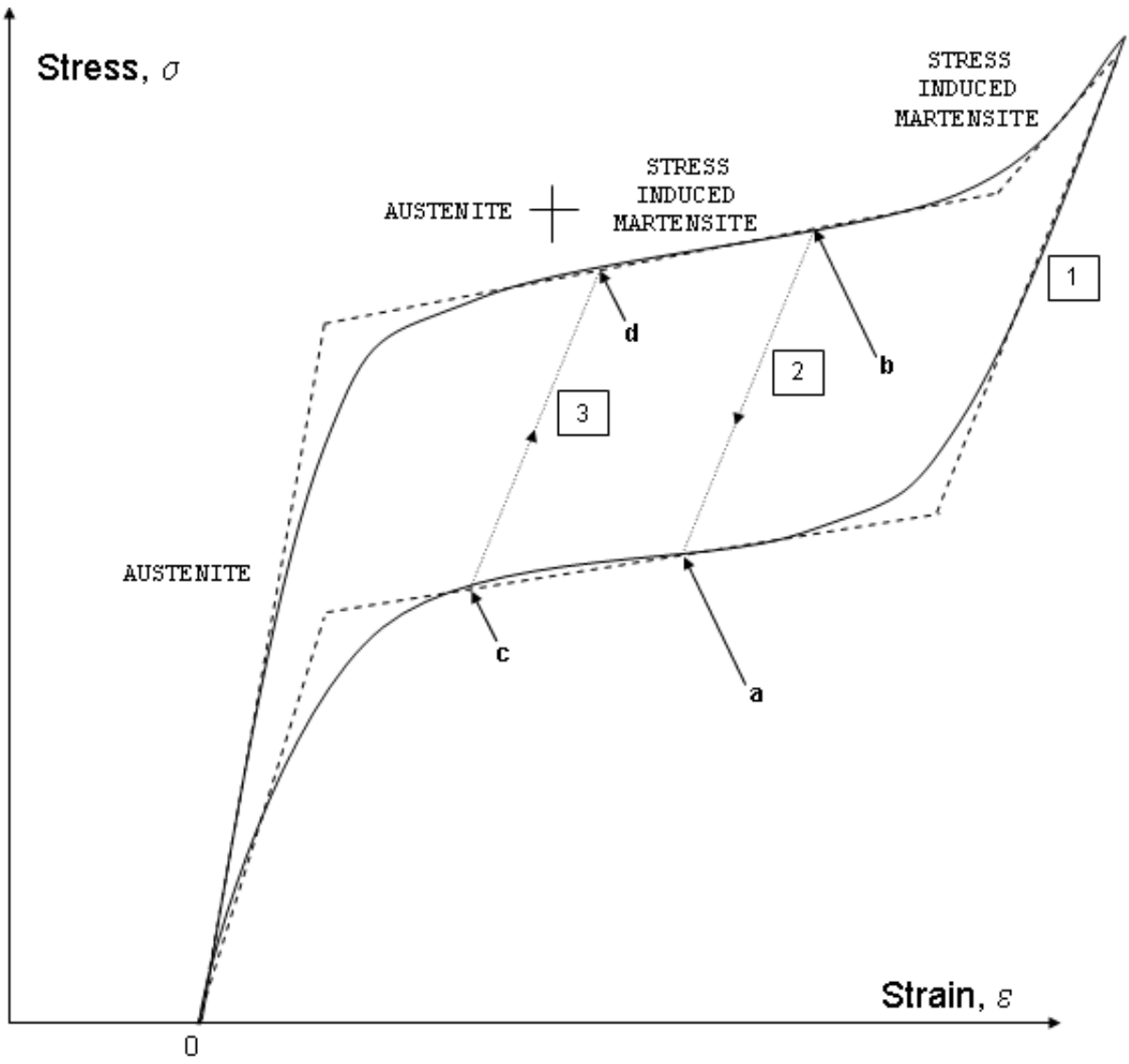

Figure 2.13: Austenite stress - strain curve showing unloading and reloading behaviour

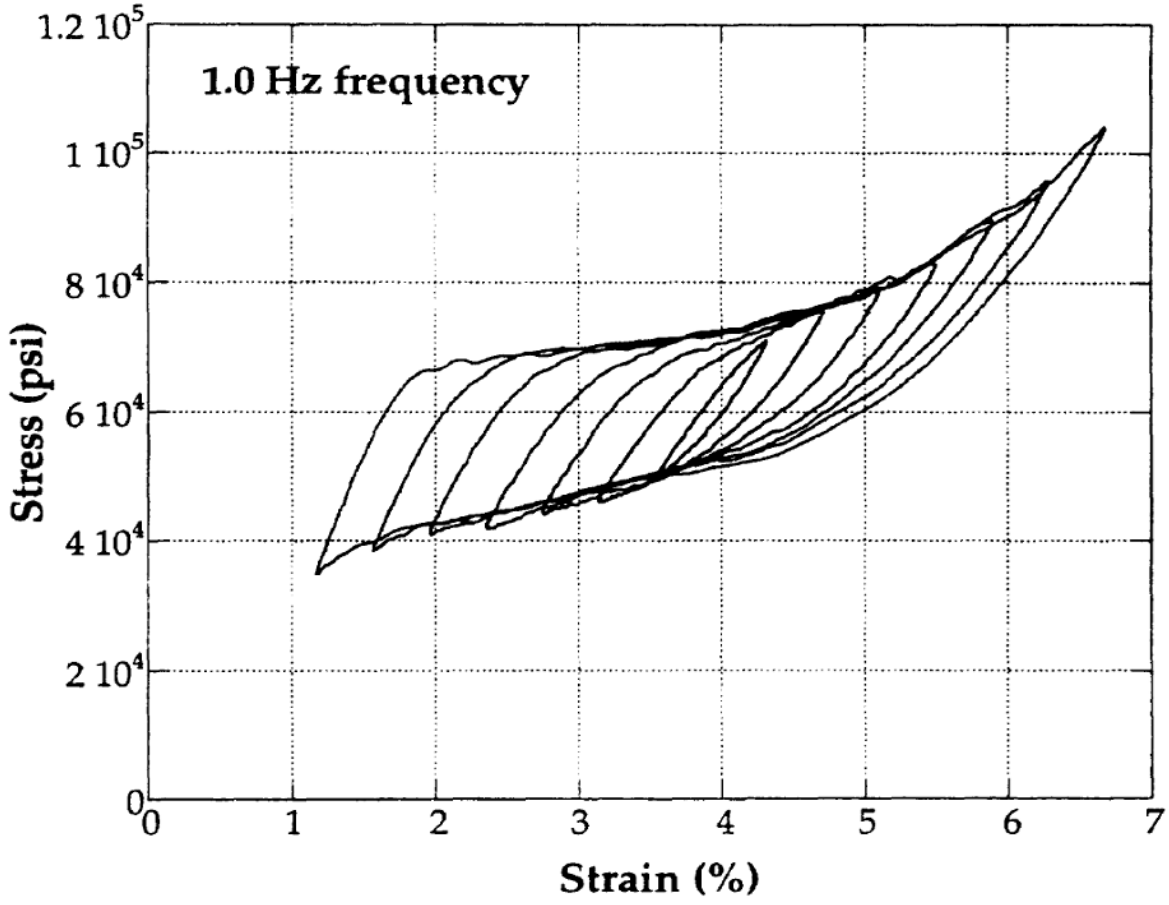

Figure 2.14: Experimental cyclic loading and unloading loops of austenite [15]

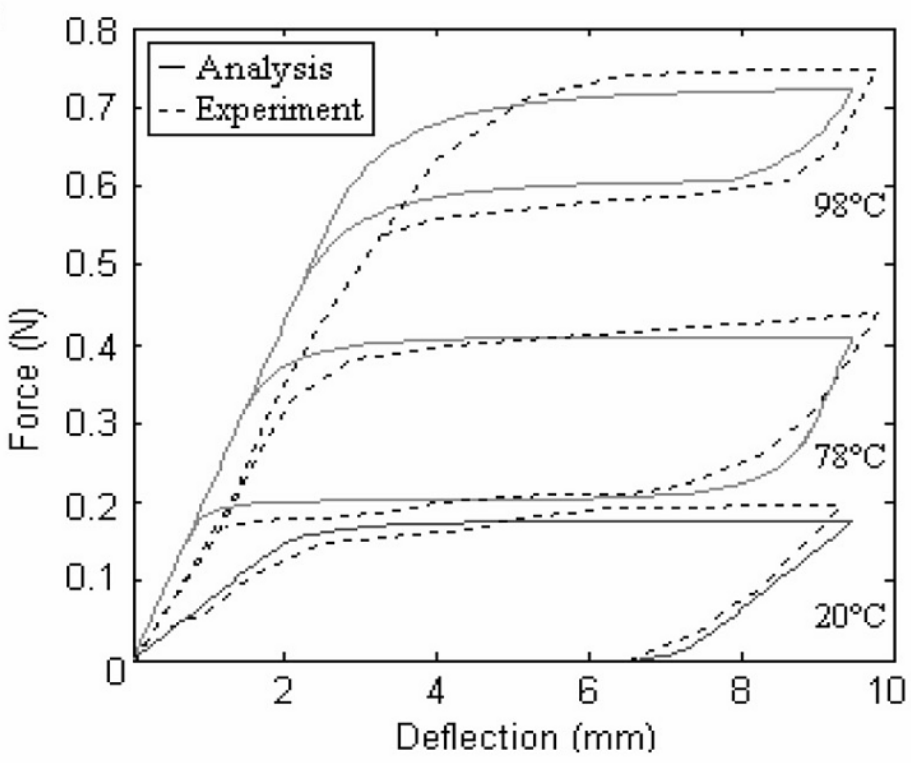

Figure 2.15: Temperature sensitivity of superelastic loops [13]

and is converted to austenite. The austenite is then loaded until the formation of SIM at *7* and then unloaded from *7* to *10*, displaying the superelastic effect.

2.7 SMA Actuators

There are two common SMA actuator configurations: the first is an SMA wire with constant force and the second is a spring-biased SMA actuator, both of which are presented in the following sections.

2.7.1 Constant Force SMA Actuator

Figures 2.17 shows a simple SMA actuator where a weight is attached to an SMA wire which stretches by a certain amount based on the mass of the weight and the diameter of the wire. Figure 2.18 shows the corresponding 'path' followed by the actuator on a stress-strain plot. At point **1** in Figure 2.18, the SMA wire is unstressed and in its twinned martensite state. When the weight is attached to the SMA wire when it is still at low temperature, the SMA will follow the loading curve between positions **1** and **2** as shown in Figure 2.18 corresponding to a stretch in the wire. Now consider that the temperature is increased by passing a current into the wire. The stress is constant due to the weight and the SMA will transition from the martensite curve to the austenite curve (**2** to **3**), thereby causing the wire to contract towards its austenite length. During this phase transition, the SMA wire does work by lifting the weight (Figure 2.17). On cooling, the material transforms back to the low-stiffness martensite and the weight elongates the wire. The SMA in this example acts as a simple actuator that raises and lowers a weight by heating and cooling the wire. Detaching the weight at high temperature will cause unloading in

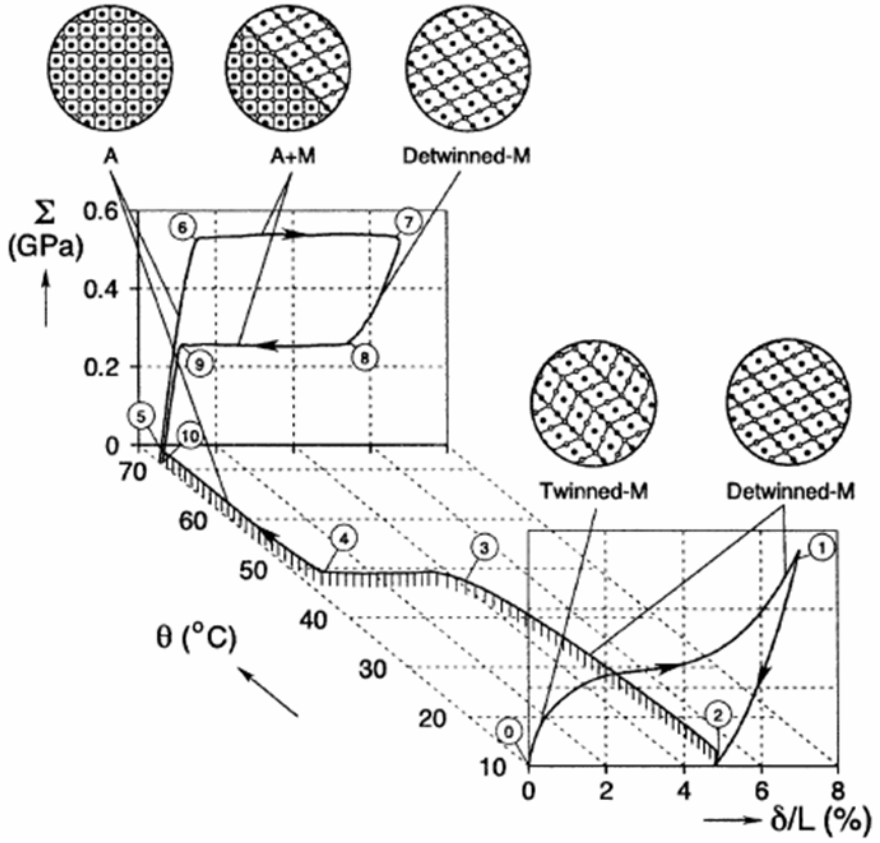

Figure 2.16: Three-dimensional SMA characteristic curve [16]

the material and the strain will reduce to 0% (Figure 2.18, **3** to **1**). Note that between positions **2** and **3** the temperature in the SMA wire is increasing and strain decreases according to the heating hysteresis curve described in section 2.5.

2.7.2 Spring-biased SMA Actuator

Figure 2.19 shows the configuration for a spring-biased SMA actuator. The spring, also known as a return spring, acts as a reset mechanism for the SMA. Unlike the constant load configuration, the force exerted is a function of spring (and hence SMA) elongation, and is therefore time-varying. Figure 2.20 shows the path followed by a spring-biased SMA actuator. The spring constant of the return spring is k_{spr}. Strain ϵ_1 is the total strain the system, set by varying the length, L, between the posts in Figure 2.19. Length, L, is fixed during the duration of the experiment, therefore the total strain in the system is constant during the the experiment. The individual strains in the spring and SMA change based upon the temperature of the SMA. Assuming actuation does not include dynamics of motion for example inertia in the spring, the forces in the spring and the SMA are equal. When the spring is attached to the wire at low temperature, and the total strain in the system is set, the SMA is stretched and follows the martensite stress-strain curve in Figure 2.20 until **2**. When the SMA wire is heated, it contracts stretching the spring and increasing the force in both the SMA and the spring. This corresponds to line **2** - **3** where the slope of the path of travel is the spring constant, k_{spr}. When the SMA wire cools, it transforms to the low stiffness martensite and the higher force in the reset spring stretches the SMA. The spring force is thereby reduced and a new equilibrium position is reached with the wire stretched once again, at **2**.

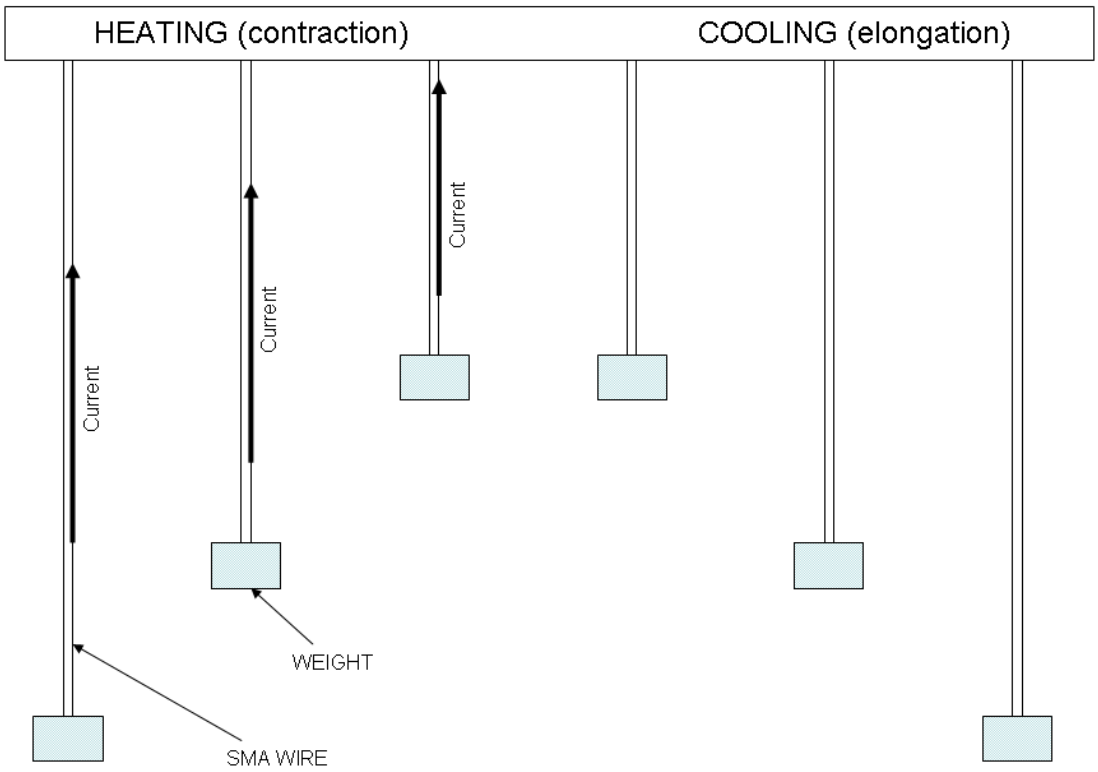

Figure 2.17: Constant force SMA actuator

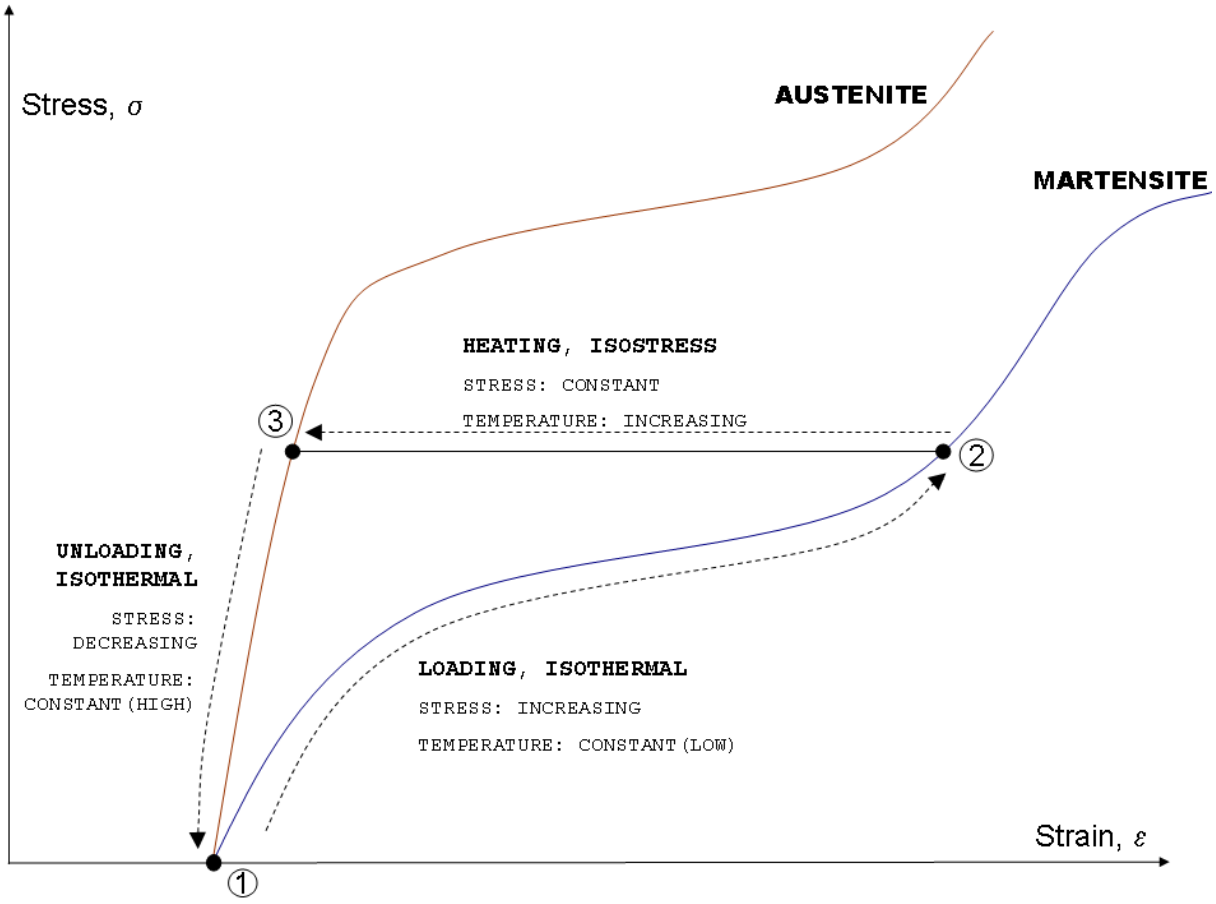

Figure 2.18: Stress-strain plot for constant force SMA actuator

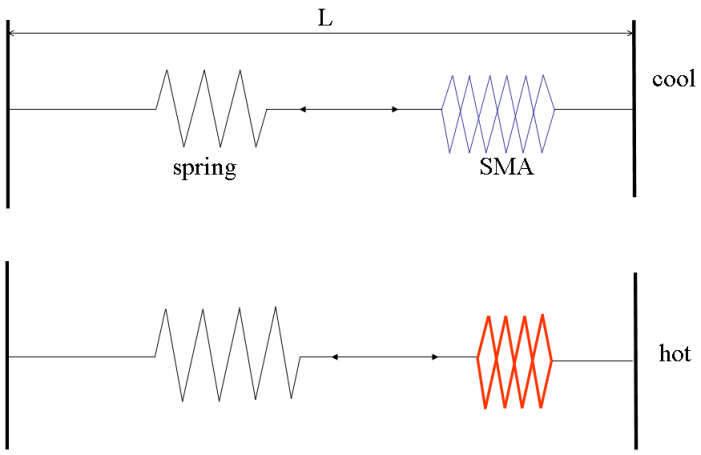

Figure 2.19: Spring-biased SMA actuator configuration

2.8 Literature Review on SMA Modelling

In the past 15 years, models to describe SMA behaviour have been developed from various perspectives such as thermodynamics, micromechanics, thermomechanical and phenomenological and finite element analysis models.

Thermodynamic based models have been developed by Achenbach [17] whose model breaks down the SMA into lattice particles arranged in layers whose shape and potential energy depends upon applied stress and temperature. Seelecke and Mueller [2] extended Achenbach's model in order to make quantitative predictions through the use of the Helmholtz free energy and in order to simulate SMA mechanical behaviour such as elastic and residual deformation, yielding through flipping of layers and shape recovery by conversion of martensite layers to austenite. Boyd and Lagoudas [18] developed a thermodynamic-based model to describe SME

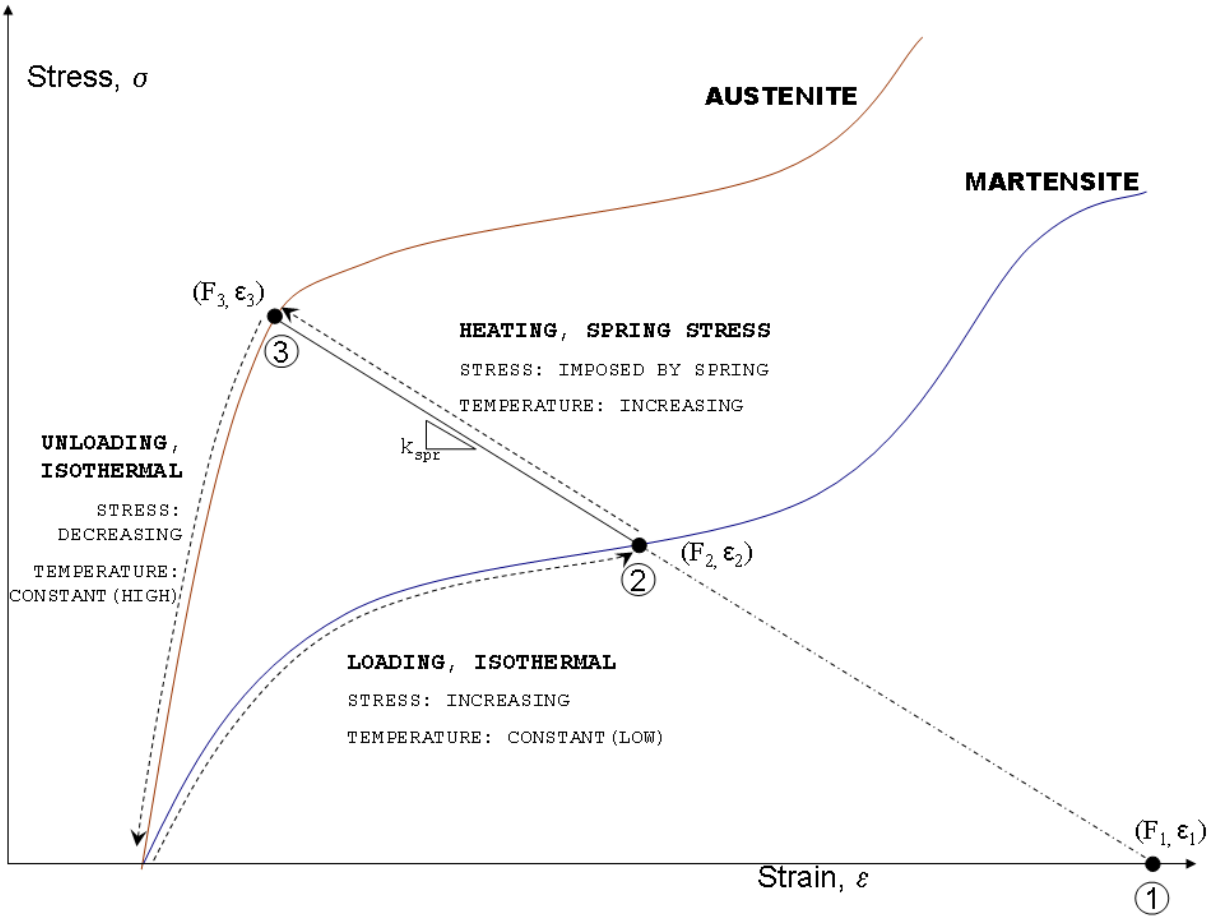

Figure 2.20: Stress-strain plot for spring biased SMA actuator

and pseudoelasticity using a free energy function and a dissipation potential. Ivshin and Pence [19] define hysteretic transformation based on temperature changes of an internal variable for the phase fraction based on a thermodynamical consideration. The hysteresis equations are governed by the differential equations for the Duhem-Madelung hysteresis model. Goo et al. [20] derived a micromechanics model based on the Helmholtz free energy and then developed equations describing kinetic relationships in the SMA, martensite nucleation and the reorientation of martensite variants. The model is able to describe pseudoelastic behaviour. Finite element models have been developed by Brinson and Lammerung [21] and Amalraj [22].

SMA models have also been developed from a phenomenological view point where equations describing experimental behaviour are developed. Phenomenological models are attractive choices in use for controls experiments due to the fact that they directly describe observed behaviour, are generally simpler, computationally faster and therefore more conducive for use in control loops. Phenomenological models have been developed by Preisach [23] whose model to describes hysteresis as a parallel summation of weighted hysteresis relays. The weighting of each of the relay modules defines its contribution to the final hysteresis loop. While the Preisach model is commonly used, it has a large number of variables that need to be identified. Tanaka et al. [24] developed a phenomenological model whose thermomechanical constitutive equations and kinetics of transformation are derived though the Claussius-Duhem inequality. Additionally, an exponential function is used quantify the relationship of the martensite fraction to the stress and the temperature in order to reproduce observed behaviour in the stress-temperature and stress-strain planes to model pseudoelasticity. Liang and Rogers [25] use a similar approach to Tanaka but with cosine functions. Brinson [14] later extended the work of Tanaka and Liang and Rogers by the separation of the martensite fraction into temperature-induced and stress-induced

parts and the introduction of twinned and detwinned martensite in order to model low temperature SME in addition to pseudoelasticity. Ikuta [11] also developed a phenomenological model similar to Tanaka's that used the martensite fraction and an exponential function to describe shape memory alloy behaviour. The exponential function developed by Ikuta was equipped with variables to define minor loops but no extention to minor loop behaviour was presented. This extension was developed by Madill and Wang [10] who used so called 'switching conditions' in order to describe the behaviour of SMAs under partial temperature cyling. Ikuta's model and Madill's extension, however, deal with constant stress. Additionally, Madill developed equations to describe the loading behaviour of martensite and austenite, although austenite was modelled only until its first yield point. A superiority of Madill's model is that there are relatively few parameters to identify when compared to other phenomenological models such as the Preisach model. Additionally, the relative simplicity of the model allows for it to be implemented in real-time for use in control experiments. In this work, Madill's phase kinetics model is further extended to include time-varying stresses. A mechanical model is also developed that describes the complete loading, unloading and reloading behaviour of SMAs. The model is then inverted and used as a feed-forward controller to linearize the SMA wire behaviour to improve the control of SMA actuators.

Chapter 3

SMA Wire Forward and Inverse Models

The section presents the equations for the proposed forward and inverse model of the SMA wire as well as the equations for a spring-biased SMA actuator.

3.1 Block Diagrams

The forward model of the SMA wire is divided into three blocks:

- Heating model: relates the electrical current, I, into the SMA wire to the temperature difference from ambient, θ, of the wire.
- Phase Kinetics model: relates the wire stress, σ, and temperature difference from ambient, θ, to the martensite fraction, R_m.
- Mechanical model: relates the wire stress, σ, and martensite fraction, R_m, to strain, ϵ, of the wire.

The relationships between each of these blocks is shown in Figure 3.1.

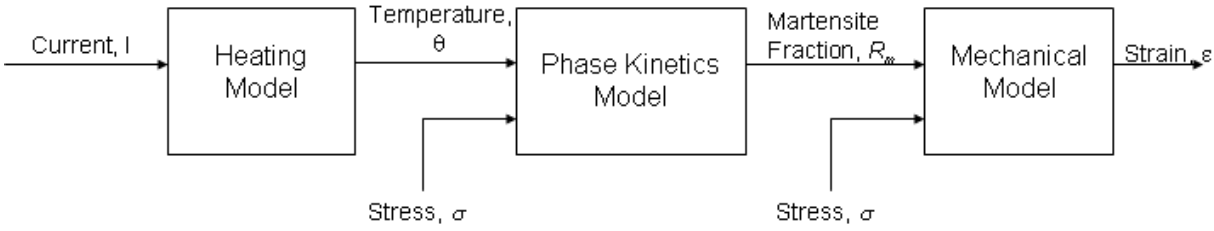

Figure 3.1: SMA forward model components

The inverse model consists of inverses of each of the blocks in the forward model. The inverse model is used for control purposes to linearize the SMA wire behaviour (Figure 3.2) in order to improve tracking performance and robustness of a control architecture that uses the SMA wire as the plant.

Section 3.2 describes the heating model with its inverse in Section 3.3, Section 3.4 describes

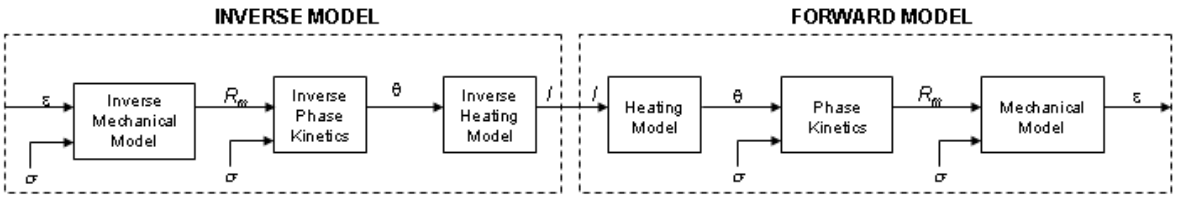

Figure 3.2: SMA inverse and forward model components

the phase kinetics model with its inverse in Section 3.5[1], and Section 3.6 describes the mechanical model with it inverse in Section 3.7. Additionally, Section 3.8 presents the equations and

[1]Note that in all control experiments, the wire is initially heated until completely austenite and then cooled to completely martensite. This ensures that at the start of the experiment, the SMA wire is on the major loop, therefore, initial conditions do not need to be included as inputs in the block diagram.

simulation results for a spring-biased SMA actuator.

3.2 Heating Model

In this model, resistive heating is used to raise the temperature of the SMA material. The heating model (3.1) relates input current, I, to wire temperature above ambient, θ. It is a general first order heating model as is commonly used in SMA modelling [10, e.g.]. The variables in (3.1) are as follows:

ρ = material density [kg/m^3]

C = specific heat [J/(kg^{o}C)]

A = wire surface area [m^2]

V = wire volume [m^3]

h = convection heat transfer coefficient [W/(m^{2o}C)]

t = time [s]

T = wire temperature [^{o}C]

I = wire current [A]

R = wire resistance [Ω]

T_∞ = ambient temperature [^{o}C]

$$CV\rho\frac{dT}{dt} = I(t)^2R - hA(T(t) - T_\infty) \tag{3.1}$$

The heat equation in (3.1) can be explained intuitively as follows: The rate of change of temperature depends upon the electrical energy input and the energy lost from the system through convection which depends on the surface area of the wire. Note that (3.1) includes only the ef-

fects of convection via the heat transfer coefficient, h. Conduction is ignored because the wire is short and thin and therefore the temperature is assumed to be constant in the material [9]. Heat transfer by radiation is negligible compared to convection at the operating temperatures of the SMA [9]. In SMAs, h and the heat capacity, C, are temperature dependant [26, 27]. Some researchers have developed expressions for this dependance [28], however, since the focus of this work is on the phase transformation and mechanical behaviour the SMA, h and C are taken as constant lumped parameters. Additionally, although the volume and the surface area will change during operation, the effects are negligible and volume and surface area are taken as constant parameters [4]. The resistance is assumed to be constant and is calculated by multiplying the length of the wire with the linear resistance listed in the SMA data sheet [7]. This assumption is limiting because resistance has been shown to change considerably with changes in phase [29].

Defining $\theta(t) = T(t) - T_\infty$ and taking the Laplace transform of (3.1) after substituting for power, $P(t) = I(t)^2 R$, gives the transfer function of the heating model (3.2). Note that the wire is assumed to start at ambient temperature and therefore $T(0) = T_\infty$.

$$CV\rho(s\theta(s) - \theta(0)) = P(s) - hA\theta(s)$$

$$\theta(0) = T(0) - T_\infty = 0$$

$$\frac{\theta(s)}{P(s)} = \frac{\frac{1}{CV\rho}}{s + \frac{hA}{CV\rho}} \tag{3.2}$$

3.3 Inverse Heating Model

The heating model inverse is constructed by inverting the transfer function (3.2) for the linear heating equation and adding a fast pole to make the result proper. The resulting transfer function is:

$$\frac{P(s)}{\theta(s)} = \frac{100(s + \frac{hA_{sma}}{C\rho V_{sma}})}{\frac{1}{C\rho V_{sma}}(s + 100)}. \tag{3.3}$$

3.4 Phase Kinetics Model

The phase kinetics model describes the hysteretic phase transformation in the SMA. The model equations presented in this section reproduce the following SMA hysteresis characteristics initially described in Section 2.5:

- The relationship between the transformation temperatures, M_s, M_f, A_s, A_f and the hysteresis loop.
- The formation of major loops by complete temperature reversal.
- The formation of minor loop branches by partial temperature reversal.
- The shifting of the hysteresis loops due to changes in stress.
- The formation of closed hysteresis loops.
- Immediate output reversal on input reversal.

The model takes stress and temperature as its inputs and computes the martensite fraction, R_m. It is assumed that temperature determines only two phases in the material, austenite and martensite,

so R_m is sufficient to fully describe the material state. The distinction between twinned and detwinned martensite, and the presence of stress-induced martensite, is determined later as a function of stress in the mechanical model. The modeling approach used for the phase kinetics behaviour is a phenomenological model that is based on the initial work of Ikuta [11] and Madill [10]. Ikuta's model equations describe the major hysteresis loop that defines the phase kinetics in SMA and Madill's model elaborated on Ikuta's model to describe minor loop behaviour. Madill's model, though, is valid for only constant stresses. The phase kinetics model presented in this section is an extension of Madill's model to include time varying stresses. The sections below present an overview of Ikuta and Madill's work and the extension to Madill's model. Note that the variable names in Ikuta's and Madill's models have been changed for simplicity.

3.4.1 Ikuta's Model

The basic approach, first developed by Ikuta [11], models the observed behaviour of the major hysteresis curve with exponential functions. The parameters of the functions are identified from experimental major loop data, hence, the classification of this model as phenomenological. Ikuta defined the hysteresis loop using the exponential function:[1]

$$R_m(\theta) = \frac{1}{1 + e^{k_m(\theta - \beta)}}. \tag{3.4}$$

Figure 3.3 shows the relationship of the variables of (3.4) with the general shape of the hysteresis loop. The slope of the curve at the inflection, S, is related to k_m, which can take on either of two values, k_m^C or k_m^H, depending on whether R_m is increasing (i.e., the wire is cooling

[1]Note that in (3.4), the independent variable T, temperature has been changed to θ, where θ is the temperature difference from ambient temperature, for consistency between Ikuta's and Madill's model

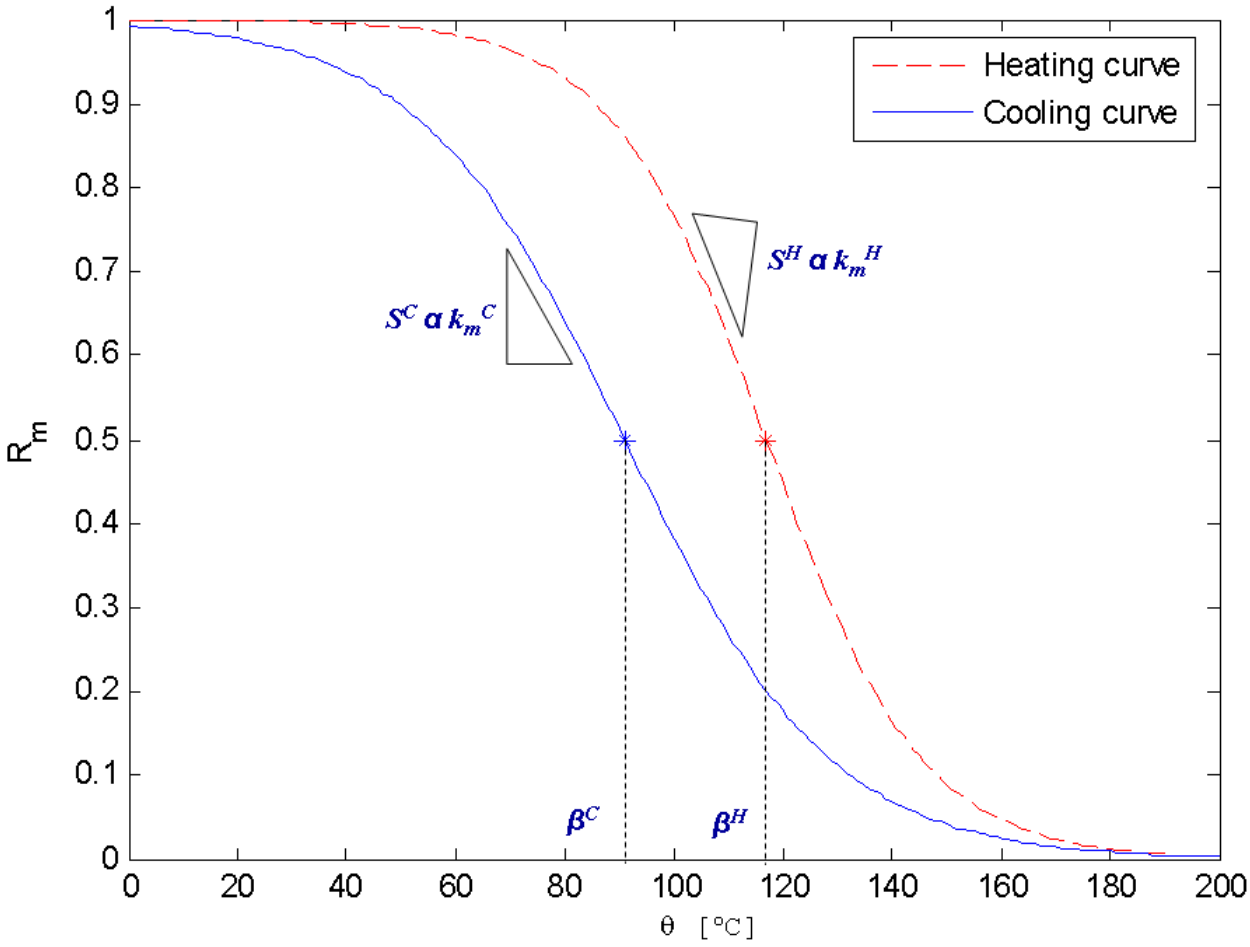

Figure 3.3: SMA hysteresis loop

and slope is S^C) or decreasing (i.e., the wire is heating and slope is S^H), respectively. The temperature of the inflection point in the $R_m - \theta$ curve is denoted by β. Similar to k_m, different values of β are associated with each of the cooling and heating transformations. These are also distinguished by the superscripts "C" and "H". The parameters of (3.4) can be extracted from the major hysteresis loop (Figure 3.3) as follows: The values of k_m^C and k_m^H are found by computing the derivative at the point of inflection (3.5) of the cooling and heating curve, respectively.

$$\begin{aligned} \left.\frac{dR_m}{d\theta}\right|_{\theta=\beta} &= \left.\frac{d}{d\theta}\left[\frac{1}{1+e^{k_m(\theta(t)-\beta)}}\right]\right|_{\theta=\beta} \\ S &= \left.\frac{-k_m e^{k_m(\theta-\beta)}}{(1+e^{k_m(\theta(t)-\beta)})^2}\right] |_{\theta=\beta} \\ &= \frac{-k_m}{4} \\ k_m &= -4S \end{aligned} \tag{3.5}$$

Parameter β^H is related to the transformation temperatures A_s and A_f, and β^C to M_s and M_f, the martensite start and finish temperatures. Additionally, β includes the effect of stress on the hysteresis curves. Changes in stress cause shifting of the transformation temperatures, as shown in Figure 3.4. The constant, c_m[^{o}C/Pa], represents the effect of stress on the transformation temperatures in that $1/c_m$ is the slope of the $\sigma - T$ curve (see Section 2.5.1). Therefore:

$$\frac{1}{c_m} = \frac{d\sigma}{dT} \Longrightarrow dT = c_m d\sigma. \tag{3.6}$$

The equations defining β are in (3.7), where M_s, M_f, A_s and A_f are the transformation temperatures at stress σ_o (often, $\sigma_o = 0$ Pa). The relationship between stress and the hysteresis curve is shown in Figure 3.5 where the heating/cooling curve is observed to have a horizontal shift due

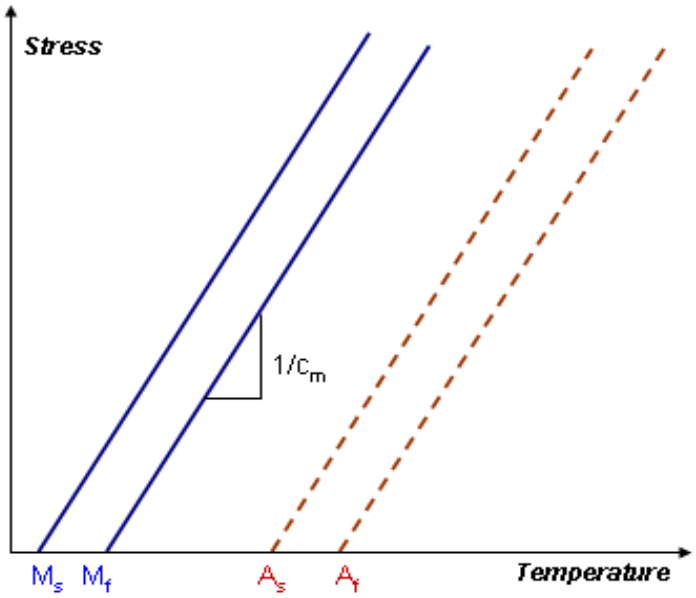

Figure 3.4: Shifting of transformation temperatures due to stress

to an increase in stress.

$$\begin{aligned} \beta^C &= \frac{1}{2}(M_s + M_f) - T_\infty + c_m(\sigma - \sigma_o) \\ \beta^H &= \frac{1}{2}(A_s + A_f) - T_\infty + c_m(\sigma - \sigma_o) \end{aligned} \tag{3.7}$$

3.4.2 Madill's Model

In the later extension by Madill [10], two new parameters $\alpha(t)$ and $\gamma(t)$ are introduced to (3.4) to incorporate minor loop behaviour. Additionally, to ensure tht the wire starts on the major loop, α is split into α_p and an exponential function. Parameters α_p and γ are set to 1 and 0 to describe the major loop [10]. This modification ensures that at room temperature ($\theta = 0$), $R_m = 1$ and as

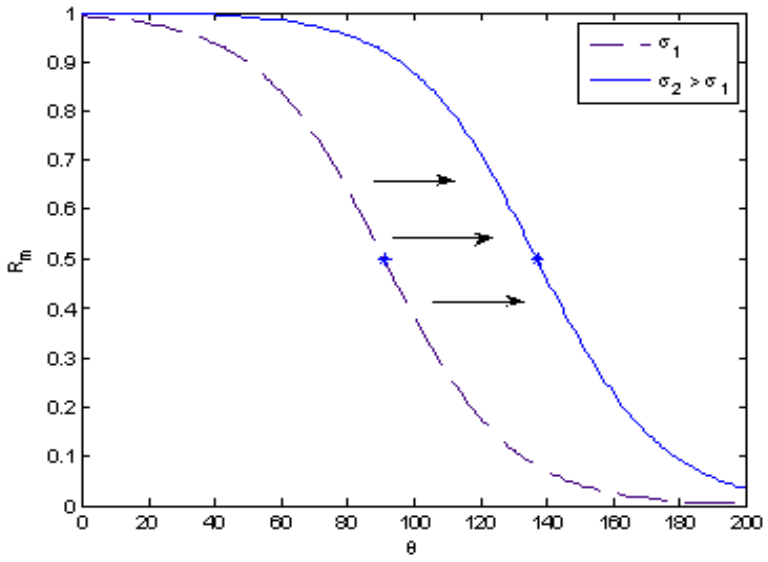

Figure 3.5: Shifting of hysteresis loop due to stress

$\theta \to \infty$, $R_m = 0$.

$$\begin{aligned} R_m(\theta, \sigma) &= \frac{\alpha(t)}{1 + e^{k_m(\theta(t) - \beta(\sigma(t)))}} + \gamma(t) \\ &= \frac{\alpha_p(t)(1 + e^{-k_m \beta(\sigma)})}{1 + e^{k_m(\theta(t) - \beta(\sigma))}} + \gamma(t). \end{aligned} \tag{3.8}$$

Parameter α_p determines the scaling of the curve while γ determines the offset. By adjusting α_p and γ, scaled and shifted curves can be generated. Portions of these modified curves are used to model the internal branching behaviour seen in SMAs when the transformation is incomplete due to input reversal (cf. Section 2.5). Figure 3.6 shows the major loop and two scaled loops with their respective α_p and γ values. The portions of the scaled loops that are enclosed in the

major loop are traversed during partial cycling to form minor loop branches.

In Madill's model, α and γ are piecewise constant functions that are recalculated on switching between heating and cooling to ensure continuity of the martensite fraction. These switching conditions are re-derived for the case of time varying stress, in the following section. Superscripts "C" and "H" are added onto α and γ to differentiate between heating and cooling curves.

3.4.3 Extension to Time-varying Stress

A novel contribution to the model is in the incorporation of time-varying stresses. The time-varying stress affects the calculation of β and the calculation of α and γ. Therefore, in the extended model, α and γ are piecewise continuous (but not necessarily constant) functions.

Initial Conditions

To initialize the model, the wire is assumed at $t = 0$ to be fully martensite ($R_m = 1$) and at room temperature. Therefore the following variables are set in (3.8):

$$\begin{aligned} \alpha_p(0) &= 1 \\ \alpha(0) &= \alpha_p(0)(1 + e^{-k_m\beta(\sigma(0))}) \\ \gamma(0) &= 0 \end{aligned} \tag{3.9}$$

with $\beta(\sigma(0))$ computed from system constants and initial stress according to either β^H or β^C in (3.7) depending upon whether the SMA is initially heating or cooling. Setting the initial parameters as in (3.9) ensures that $R_m(0, \sigma) = 1$ for any value of σ.

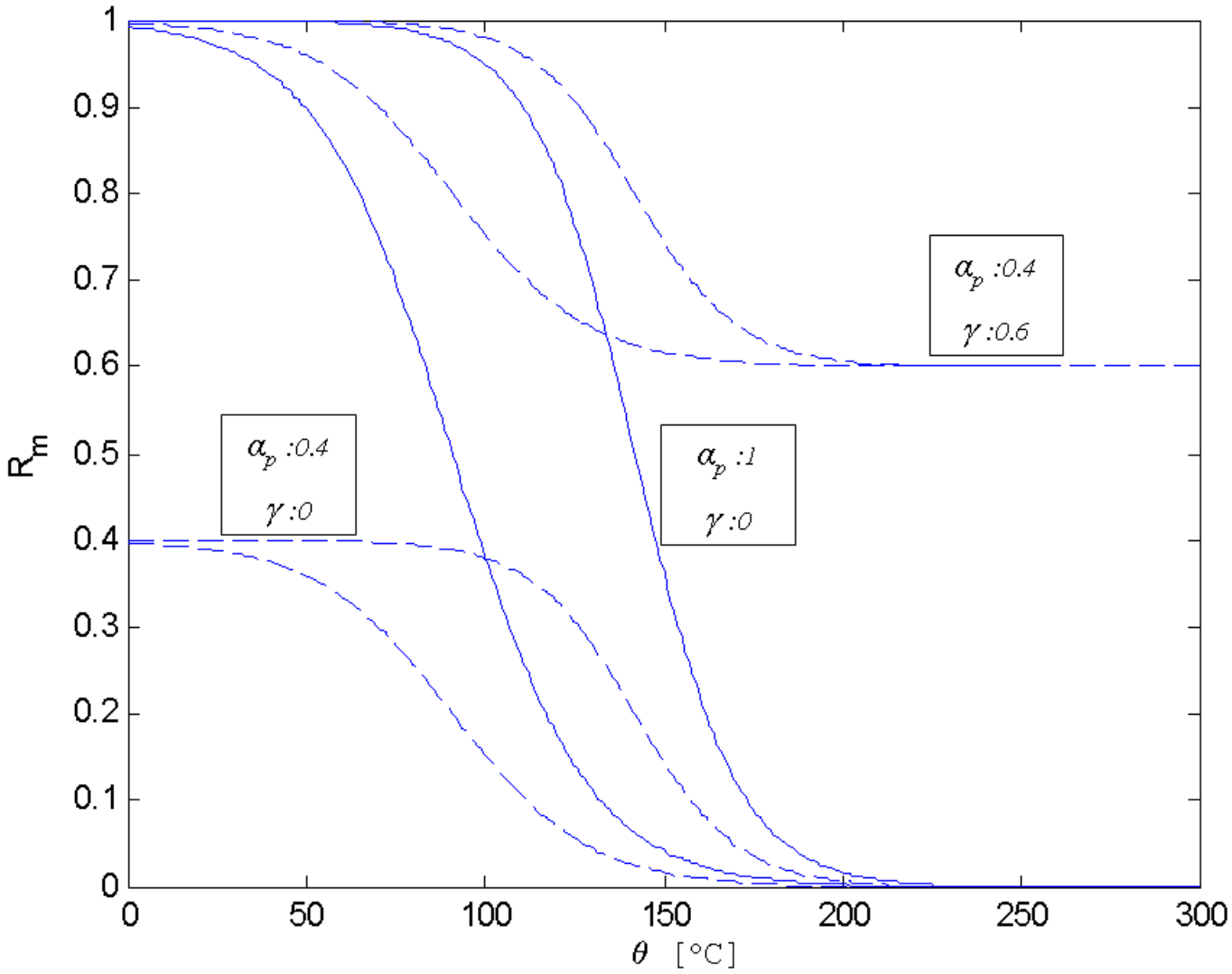

Figure 3.6: Major and scaled hysteresis loops at constant stress

The temperature input is separated into segments of monotonic heating and cooling. 'Switching' is said to occur at the transition between the monotonic heating or cooling segments and at these switching instants, the parameters defining the hysteresis loop need to be recalculated based on matching boundary conditions. The sections below present the equations for the monotonic heating and cooling segments as well at the equations at the switching instants.

Monotonic Heating or Cooling

When either monotonically heating or cooling, α_p and γ retain their previous values. At $t = 0$, α_p and γ values are those set by the initial conditions (3.9). The remaining variables are calculated with (3.10) and (3.11).

Heating:

$$\begin{aligned} \alpha^H(t) &= \alpha_p(1 + e^{(-k_m^H \beta^H(\sigma(t)))}) \\ \gamma^H &= \gamma \\ R_m^H(\theta, \sigma) &= \frac{\alpha^H(t)}{1 + e^{k_m^H(\theta(t) - \beta^H(\sigma(t)))}} + \gamma^H. \end{aligned} \tag{3.10}$$

Cooling:

$$\begin{aligned} \alpha^C(t) &= \alpha_p(1 + e^{(-k_m^C \beta^C(\sigma(t)))}) \\ \gamma^C &= \gamma \\ R_m^C(\theta, \sigma) &= \frac{\alpha^C(t)}{1 + e^{k_m^C(\theta(t) - \beta^C(\sigma(t)))}} + \gamma^C. \end{aligned} \tag{3.11}$$

Switching Conditions

When the temperature changes direction, the R_m value changes direction as well since input (θ) reversal causes immediate output (R_m) reversal. If the switching occurs when R_m is neither 1 nor 0, minor loops are formed. Maintaining the continuity of R_m during switching establishes one of the boundary conditions required to compute the new α_p and γ. The next two sub-sections present the model equations for temperature switch from heating to cooling and from cooling to heating. These equations have been updated from those derived by Madill [10] in order to include the effects of time-varying stress.

Heating to Cooling:

Consider a monotonically increasing temperature profile that starts at room temperature ($R_m = 1$). The material follows the heating curve as shown by the bold curve in Figure 3.7a. Consider a certain temperature, θ_s, before R_m is 0, at which the temperature profile changes from increasing to decreasing. Let the stress at this 'switching point' be σ_s. The material now has to follow a minor cooling curve shown as the dotted curve in Figure 3.7b. In order to define this minor cooling curve, its α and γ values need to be determined. Switching conditions are used to provide two equations to solve for these two unknowns. Point 1 in Figure 3.7b occurs right when $\theta = \theta_s$. At this point, the value of the original heating curve, R_m^H, and the value of the subsequent cooling curve, R_m^C, must be equal to ensure continuity in R_m. This is called the *continuity condition*. Further examination of Figure 3.7b shows that the major loop heating curve, the major loop cooling curve and the minor loop cooling curve all originate at the same point, labeled 2. This condition called the *common origin condition*, must be satisfied since the SMA major loop is closed. The switching conditions are presented in (3.12) with $t = t_s$ as the switching time:

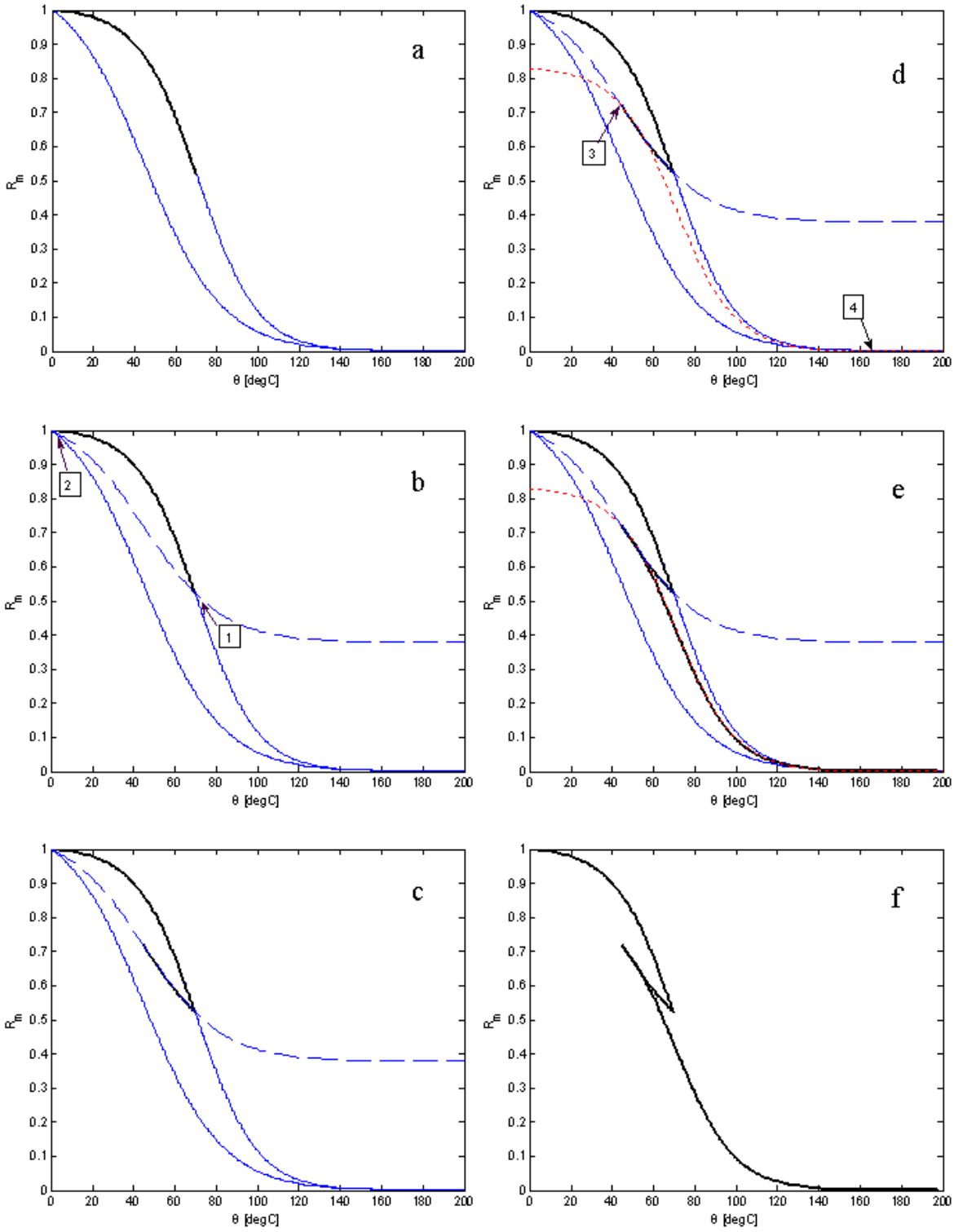

Figure 3.7: Evolution of minor loop branching

$$Continuity\ Condition: R_m^C(\theta_s, \sigma_s) = R_m^H(\theta_s, \sigma_s)$$

$$Common\ Origin: R_m^C(\theta = 0, \sigma_s)\big|_{\alpha_p=\alpha_p(t_s),\gamma=\gamma(t_s)} = R_m^H(\theta = 0, \sigma_s)\big|_{\alpha_p=\alpha_p(0),\gamma=\gamma(0)} \quad (3.12)$$

Using t_s^- and t_s^+ to specify instants before and after switching, defining θ and σ at t_s^+ and t_s^- as: $\theta_s^+ = \theta(t_s^+), \theta_s^- = \theta(t_s^-)$, $\sigma_s^+ = \sigma(t_s^+), \sigma_s^- = \sigma(t_s^-)$ and substituting into (3.10) and (3.11) gives:

$$\begin{aligned} Before\ Switching: \quad R_m^H(\theta_s^-, \sigma_s^-) &= \frac{\alpha_p^H(t_s^-)(1 + e^{(-k_m^H \beta^H(\sigma_s^-))})}{1 + e^{k_m^H(\theta_s^- - \beta^H(\sigma_s^-))}} + \gamma^H(t_s^-) \\ After\ Switching: \quad R_m^C(\theta_s^+, \sigma_s^+) &= \frac{\alpha_p^C(t_s^+)(1 + e^{(-k_m^C \beta^C(\sigma_s^+))})}{1 + e^{k_m^C(\theta_s^+ - \beta^C(\sigma_s^+))}} + \gamma^C(t_s^+) \end{aligned} \quad (3.13)$$

From the common origin condition:

$$\begin{aligned} \alpha_p^C(t_s^+) + \gamma^C(t_s^+) &= \alpha_p^H(0) + \gamma^H(0) \\ \gamma^C(t_s^+) &= \alpha_p^H(0) + \gamma^H(0) - \alpha_p^C(t_s^+) \end{aligned} \quad (3.14)$$

From the continuity condition:

$$\begin{aligned} R_m^C(\theta_s^+, \sigma_s^+) &= R_m^H(\theta_s^-, \sigma_s^-) \\ \frac{\alpha_p^C(t_s^+)(1+e^{(-k_m^C\beta^C(\sigma_s^+))})}{1+e^{k_m^C(\theta_s^+-\beta^C(\sigma_s^+))}} + \gamma^C(t_s^+) &= R_m^H(\theta_s^-, \sigma_s^-) \\ \gamma^C(t_s^+) &= R_m^H(\theta_s^-, \sigma_s^-) - \frac{\alpha_p^C(t_s^+)(1+e^{(-k_m^C\beta^C(\sigma_s^+))})}{1+e^{k_m^C(\theta(t_s^+)-\beta^C(\sigma_s^+))}} \end{aligned} \tag{3.15}$$

Equating (3.14) and (3.15) and rearranging gives $\alpha_p^C(t_s^+)$ and subsequently $\alpha^C(t_s^+)$:

$$\alpha_p^C(t_s^+) = (R_m^H(\theta_s^-, \sigma_s^-) - \alpha_p^H(0) - \gamma^H(0)) \left[\frac{1+e^{k_m^C(\theta_s^+-\beta^C(\sigma_s^+))}}{e^{-k_m^C\beta^C(\sigma_s^+)} - e^{(\theta(t_s^+)-k_m^C\beta^C(\sigma_s^+))}}\right] \tag{3.16}$$

$$\alpha^C(t_s^+) = \alpha_p^C(t_s^+)(1+e^{-k_m^C\beta^C(\sigma_s^+)}). \tag{3.17}$$

Substituting $\alpha_p^C(t_s^+)$ into (3.14) gives $\gamma^C(t_s^+)$ to give an R_m curve according to Figure 3.7c

Cooling to Heating:

Consider now that before the temperature reduces down to ambient, the temperature profile switches from cooling to heating at certain temperature, θ_s. The material will now follow a minor heating curve as shown by the fine dotted curve in Figure 3.7d. Similarly to the heating to cooling case, the α and γ values need to be determined to define this minor heating curve. The continuity condition still holds at point 3. At point 4, it is observed that the major loop heating and cooling curve and the minor loop heating curve all terminate at the same point at $\theta \to \infty$.

This is called the *common limit condition* written mathematically as:

$$Common\ Limit: \quad \lim_{\theta\to\infty} R_m^C(\theta,\sigma_s)\big|_{\alpha_p=\alpha_p(t_s),\gamma=\gamma(t_s)} = \lim_{\theta\to\infty} R_m^H(\theta,\sigma)\big|_{\alpha_p=\alpha_p(0),\gamma=\gamma(0)} \tag{3.18}$$

In a similar derivation as for the transition to cooling, (3.18) and the continuity condition in (3.12) give unknowns for the minor loop heating curve in Figure 3.7e as (3.19).

$$\begin{aligned} \gamma^H(t_s^+) &= \gamma(0) \\ \alpha^H(t_s^+) &= (R_m^C(\theta_s^-,\sigma_s^-)) - \gamma^H(t_s^+))(1+e^{k_m^H(\theta_s^+ - \beta^H(\sigma_s^+))}) \end{aligned} \tag{3.19}$$

The final curve for the entire temperature profile is shown in Figure 3.7f

3.5 Inverse Phase Kinetics Model

The inverse phase kinetics model converts the martensite fraction and stress to temperature. The equations defining this block are found by solving for temperature in the phase kinetics equations in Section 3.4.

3.5.1 General Equation

The general equation of inverse phase kinetics model, (3.20), calculates the value θ, the temperature difference from ambient, from information about the applied stress profile and the variation

in the martensite fraction, R_m.

$$\theta = \frac{1}{k_m} ln \left(\frac{\overbrace{\alpha^p_{inv}(t)(1 + e^{(-k_m \beta(t))})}^{\alpha_{inv}(t)}}{R_m(t) - \gamma_{inv}(t)} - 1 \right) + \beta(t) \tag{3.20}$$

Variables $\alpha_{inv}(t)$ and $\gamma_{inv}(t)$ are time-varying piecewise continuous functions that are used to define the major and minor loops. They are the unknowns in the equation. Specifically, $\alpha_{inv}(t)$ defines the 'length' of the loop and $\gamma_{inv}(t)$ defines the offset of the loop. The major loop is defined by $\alpha^p_{inv} = 1$ and $\gamma_{inv} = 0$. Additionally, α^p_{inv} and γ_{inv} lie between 0 and 1 inclusive. Variables $\alpha_{inv}(t)$, and $\gamma_{inv}(t)$ are also given superscripts C and H to differentiate between heating and cooling.

Initial Conditions

At $t = 0$, $\alpha_{inv}(t)$, $\gamma_{inv}(t)$ and $\beta(t)$ need to be initialised. α^p_{inv} and γ_{inv} are set at 1 and 0 to signify the major loop with no offsets. These values can be changed depending on experimental observations.

$$\begin{aligned} \alpha^p_{inv}(0) &= 1; \\ \gamma_{inv}(0) &= 0; \\ \alpha_{inv}(0) &= alpha^p_{inv}(0)(1 + e^{(-k_m \beta(\sigma(0)))}) \end{aligned} \tag{3.21}$$

Monotonic Heating or Cooling

When either monotonically heating or cooling, $\alpha_{inv}(t)$, $\gamma_{inv}(t)$ are as follows

Heating:

$$\begin{aligned} \alpha_{inv}^{H} &= \alpha_{inv}^{p}(1+e^{(-k_m^H \beta^H(\sigma(t)))}) \\ \gamma_{inv}^{H} &= \gamma_{inv} \\ \theta^H(t) &= \tfrac{1}{k_m^H} ln\left(\tfrac{\alpha_{inv}^H}{R_m-\gamma_{inv}^H}-1\right)+\beta^H(t). \end{aligned} \tag{3.22}$$

Parameters α_{inv}^{p} and γ_{inv} retain their previous values. For the first time step, if the material starts heating then α_{inv}^{p} and γ_{inv} are the same values as those set by the initial conditions (3.21).

cooling:

$$\begin{aligned} \alpha_{inv}^{C} &= \alpha_{inv}^{p}(1+e^{(-k_m^C \beta^C(\sigma(t)))}) \\ \gamma_{inv}^{C} &= \gamma_{inv} \\ \theta^C(t) &= \tfrac{1}{k_m^C} ln\left(\tfrac{\alpha_{inv}^C}{R_m-\gamma_{inv}^C}-1\right)+\beta^C(t). \end{aligned} \tag{3.23}$$

Switching Condtions

When the R_m value changes direction, the temperature too has to switch directions. If the switching occurs when R_m is neither 1 nor 0, minor loops are formed. Consider the switching from heating to cooling: At the switching instant, in order to maintain continuity, the temperature of the original heating curve, θ^H, and the temperature of the new cooling curve, θ^C, have to be identical. This is called the continuity condition. Additionally, the new cooling curve and the original heating curve have to originate at the same point (where $R_m = 1$). This condition is called the common origin condition. On the switching from cooling to heating, the continuity condition still holds. Additionally, the

new heating curve and the original cooling curve have to terminate at the same point as R_m tends to 0. These switching conditions are derived by solving the forward phase kinetics model switching equations presented in Section (3.4) for θ with t_s as the switching time. Variables $R_m^{t_s}$ is $R_m(t_s)$, $\theta(R_m^{t_s}, t_s)$ is the θ value at t_s. Using t_s^- and t_s^+ to specify instants before and after switching, the equations for θ before and after switching are in (3.26) and (3.27).

The continuity and common origin conditions in mathematical form are:

Continuity condition:

$$\theta(R_m^{t_s}, t_s)^H = \theta(R_m(t_s), t_s)^C \tag{3.24}$$

Common origin condition:

$$\theta(0, t_s)^H = \theta^C(0, 0) = \theta(0, 0)^H \tag{3.25}$$

Before Switching:

$$\theta^H(R_m^{t_s^-}, t_s^-) = \frac{1}{k_m^H} ln\left(\frac{\alpha_{inv}^H(t_s^-)}{R_m(t_s^-)\gamma_{inv}^H(t_s^-)} - 1\right) + \beta^H(t_s^-) \tag{3.26}$$

After Switching:

$$\theta^C(R_m^{t_s^+}, t_s^+) = \frac{1}{k_m^C} ln\left(\frac{\alpha_{inv}^C(t_s^+)}{R_m(t_s^+) - \gamma_{inv}^C(t_s^+)} - 1\right) + \beta^C(t_s^+) \tag{3.27}$$

Heating to Cooling switching equation:

$$\begin{aligned}
\alpha^p_{inv}(t_s^+) &= (R_m^{t_s^-} - \alpha^p_{inv}(0) - \gamma_{inv}(0)) \left(\frac{1+e^{k_m^C(\theta(t_s^-)-\beta^C)}}{e^{-k_m^C\beta^C} - e^{k_m^C(\theta(t_s^-)-\beta^C)}} \right) \\
\alpha^C_{inv}(t_s^+) &= \alpha^p_{inv}(t_s^+)(1+e^{(-k_m^C\beta^C)}) \\
\gamma^C_{inv}(t_s^+) &= \alpha^p_{inv}(0) + \gamma_{inv}(0) - \alpha^C_{inv}(t_s^+)
\end{aligned} \tag{3.28}$$

Cooling to Heating switching equation:

$$\begin{aligned}
\alpha^p_{inv}(t_s^+) &= (R_m^{t_s^-} - \gamma_{inv}(0)) \left(\frac{1+e^{k_m^H(\theta(t_s^-)-\beta^H)}}{1+e^{-k_m^H\beta^H}} \right) \\
\alpha^H_{inv}(t_s^+) &= \alpha^p_{inv}(t_s^+)(1+e^{(-k_m^H\beta^H)}) \\
\gamma^H_{inv}(t_s^+) &= \gamma_{inv}(0)
\end{aligned} \tag{3.29}$$

3.6 Mechanical Model

The Mechanical Model computes strain in the SMA as a function of martensite fraction, R_m, and applied stress, σ. The model equations presented in this section reproduce the following SMA mechanical behaviours initially described in Section 2.6:

- The loading behaviour of martensite and austenite based upon percentages of martensite and austenite present in the material.
- The unloading and reloading behaviour of martensite based upon location of unloading in the stress-strain plane.
- The unloading and reloading behaviour of pure austenite.
- The unloading and reloading behaviour when the SMA is a combination of austenite and SIM. This includes superelasticty and the formation of minor superelastic loops on cyclic loading and unloading.
- The shifting of superelastic loops on increase of temperature above the austenite finish temperature, A_f.
- The behaviour of a spring-biased SMA actuator.

In modelling the mechanical behaviour, some basic assumptions about the configuration of martensite and austenite layers in the SMA are made. Literature shows the use of two material configurations: a parallel configuration [11], [10] and a series configuration [19]. The parallel configuration assumes that the martensite and austenite layers are parallel to each other, therefore experience the same strain for any given stress input. The series model, on the other hand,

assumes the martensite and austenite layers are in series with each other therefore each experiencing the same stress and each contributing to the final strain. Each of these configurations is independently implemented, allowing comparison of parallel and series models with experimentally observed behaviour. While the parallel and series configurations have different results for a mixture of austenite and martensite in the SMA ($0 < R_m < 1$), the results are, naturally, identical when the material is either fully martensite ($R_m = 1$) or fully austenite ($R_m = 0$). Additionally the wire is modelled as a 1D problem with a uniaxial force. The organisation of this section is as follows: the stress - strain behaviour for pure martensite and austenite phases will be presented. This will be followed by a presentation of loading, unloading and reloading behaviour, for each of the parallel and series configurations.

3.6.1 Stress-Strain Behaviour of Martensite and Austenite

Martensite can exist either as fully twinned, fully detwinned or a combination of twinned and detwinned martensite. Similarly the austenite can be pure austenite, stress-induced-martensite (SIM) or a combination of the two. Figure 3.8 shows the characteristic martensite and austenite stress-strain curves. When martensite is fully twinned, it has modulus, E_m. If the stress on the twinned martenstite were to increase, detwinned martensite variants would start to form at the yield strain and stress $(\epsilon^M_{y1}, \sigma^M_{y1})$. The modulus of the martensite, when it is a combination of twinned and detwinned martensite, is E_t. At $(\epsilon^M_{y2}, \sigma^M_{y2})$, the martensite modulus changes to E_d and the material is fully detwinned martensite.

Similarly at low stresses and strains, austenite has modulus, E_a, until SIM starts forming at $(\epsilon^{Ms}, \sigma^{Ms})$. The modulus thereafter is E_{am} until $(\epsilon^{Mf}, \sigma^{Mf})$ where the material is fully SIM with a modulus of E_{sim}. The general equation for the stress-strain behaviour is given by (3.30) where

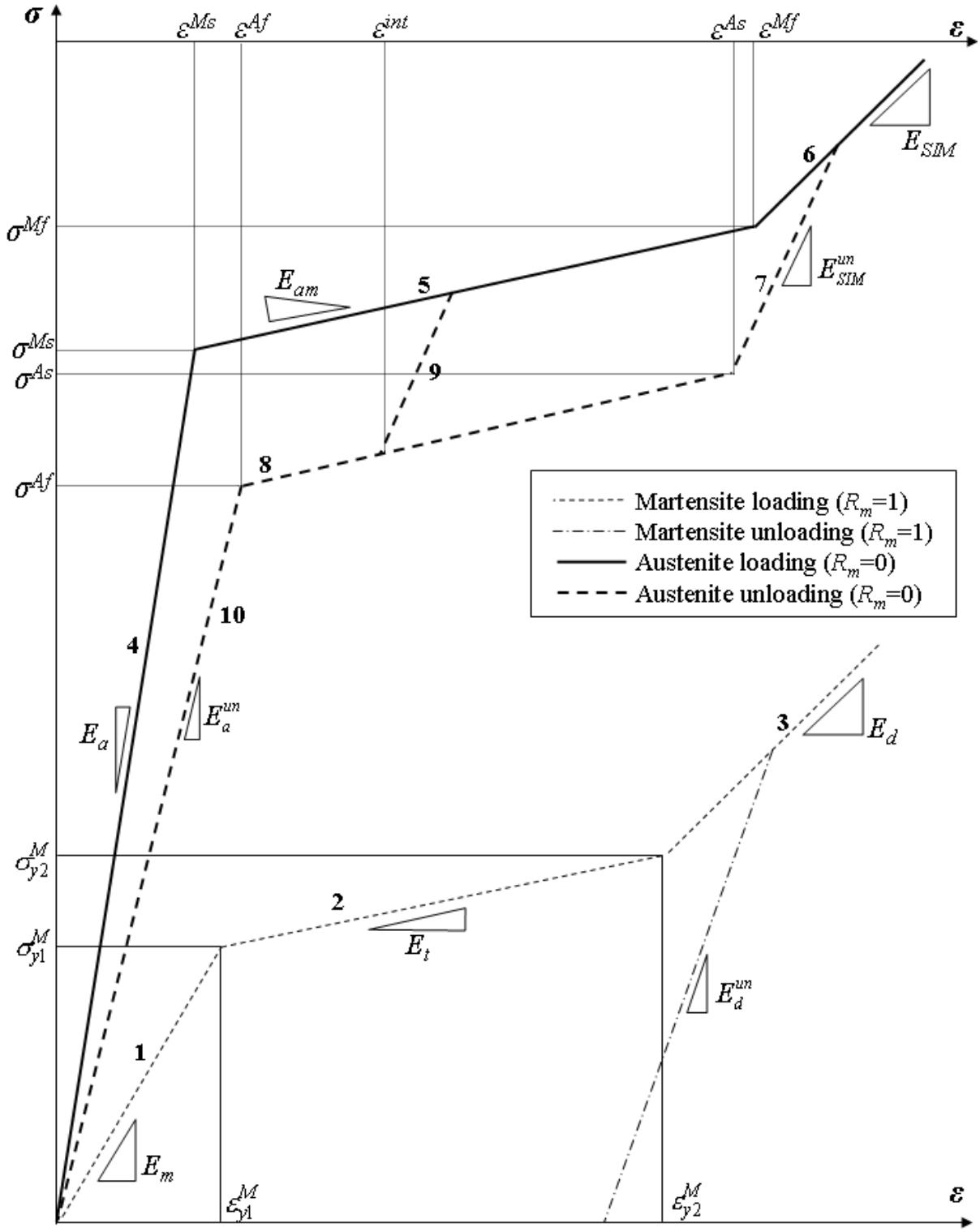

Figure 3.8: Stress-strain behaviour of martensite and austenite

E is the modulus, ϵ_{off} and σ_{off} are offsets based upon the location of the material in the stress-strain graph, whether the stress is increasing or decreasing and whether the SMA configuration is parallel or series.

$$\sigma(t) = (\epsilon(t) - \epsilon_{off})E + \sigma_{off} \tag{3.30}$$

Temperature Dependance of Austenite Parameters The austenite parameters, σ^{Af}, σ^{As}, σ^{Ms}, σ^{Mf}, E_{sim}, E_a all change with the increase of temperature above the austenite finish temperature. This results in higher yield stresses for austenite and a 'shifting' of the entire pseudoelastic loop to higher stresses with increased temperature above A_f (cf. Figure 2.15).

The next two sections decribe loading, unloading and reloading behaviour for the parallel and series mechanical models.

3.6.2 Parallel Mechanical Model

In the parallel SMA configuration, the austenite and martensite layers are parallel to each other. Due to the parallel model assumption, the strain in the martensite and austenite layers is the same while the stress is distributed amongst the different phases. Figure 3.9 shows the parallel SMA configuration showing the stress and strains in the layers.

Parallel Mechanical Model: Characteristic Equation

The stress is dependant on the moduli of the different phases and is divided according to the variable sublayer model (3.31) used by Ikuta [11]. Variables σ_m and σ_a are the martensite and austenite stresses, found by using (3.30) with the martensite and austenite moduli and offsets

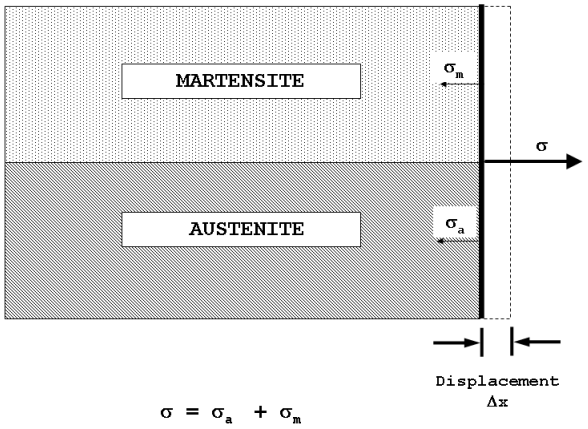

Figure 3.9: Parallel mechanical model

respectively.

$$\sigma = R_m\sigma_m + (1 - R_m)\sigma_a. \tag{3.31}$$

Isolating the strain after substituting σ_m and σ_a, both functions of strain, into (3.31) gives (3.32) where E_{mart} and E_{aust} are the martensite and austenite moduli respectively and $\sigma^M_{off}, \epsilon^M_{off}, \sigma^A_{off}, \epsilon^A_{off}$ are the martensite and austenite stress and strain offsets respectively. The values for each of these variables depend on whether the material is being loaded or unloaded and the region in the stress-strain curve of Figure 3.8. This will be made more precise in the following sections.

$$\begin{aligned} \sigma &= R_m((\epsilon - \epsilon^M_{off})E_{mart} + \sigma^M_{off}) + (1 - R_m)((\epsilon - \epsilon^A_{off})E_{aust} + \sigma^A_{off}) \\ \epsilon &= \frac{\sigma - R_m(\sigma^M_{off} - \epsilon^M_{off}E_{mart}) - (1 - R_m)(\sigma^A_{off} - \epsilon^A_{off}E_{aust})}{R_mE_{mart} + (1 - R_m)E_{aust}}. \end{aligned} \tag{3.32}$$

The loading and unloading behaviours are calculated separately for martensite and austenite and then combined according to (3.32) to compute the overall material strain.

Parallel Mechanical Model: Loading

The loading behaviour comprises Regions ***1-3*** in Figure 3.8 for martensite loading and Regions ***4-6*** for austenite loading. The modulus and offsets for each region are in (3.33) and (3.34).

$$E_{mart} = \begin{cases} E_m; & \sigma_{off}^M = 0; & \epsilon_{off}^M = 0; & \epsilon \leq \epsilon_{y1}^M \\ E_t; & \sigma_{off}^M = \sigma_{y1}^M; & \epsilon_{off}^M = \epsilon_{y1}^M; & \epsilon_{y1}^M < \epsilon < \epsilon_{y2}^M \\ E_d; & \sigma_{off}^M = \sigma_{y2}^M; & \epsilon_{off}^M = \epsilon_{y2}^M; & \epsilon \geq \epsilon_{y2}^M \end{cases} \tag{3.33}$$

$$E_{aust} = \begin{cases} E_a; & \sigma_{off}^A = 0; & \epsilon_{off}^A = 0; & \epsilon \leq \epsilon^{Ms} \\ E_{am}; & \sigma_{off}^A = \sigma^{Ms}; & \epsilon_{off}^M = \epsilon^{Ms}; & \epsilon^{Ms} < \epsilon < \epsilon^{Mf} \\ E_{sim}; & \sigma_{off}^A = \sigma^{Mf}; & \epsilon_{off}^A = \epsilon^{Mf}; & \epsilon \geq \epsilon^{Mf} \end{cases} \tag{3.34}$$

Parallel Mechanical Model: Unloading

Let the stress and strain in the material at the point of unloading be σ_{un} and ϵ_{un}. At the unloading point, σ_{un} and ϵ_{un} are distributed amongst the austenite and martensite layers according to the parallel model. Since the parallel model constrains the strains in each of the layers to be identical, the austenite and martensite unloading strains are as follows: $\epsilon_{un}^M = \epsilon_{un}^A = \epsilon_{un}$. However, the unloading stresses in the two layers, σ_{un}^M and σ_{un}^A are different and are calculated using (3.30)

giving:

$$\begin{aligned} \sigma_{un}^{M} &= (\epsilon_{un} - \epsilon_{off}^{M})E_{mart} + \sigma_{off}^{M} \\ \sigma_{un}^{A} &= (\epsilon_{un} - \epsilon_{off}^{A})E_{aust} + \sigma_{off}^{A}. \end{aligned} \tag{3.35}$$

Martensite Unloading:

As described in Section 2.6, the unloading behaviour of martensite depends on the strain at which the unloading occurs. Martensite unloading is elastic with modulus, E_m if unloaded before ϵ_{y1}^{M}. If unloaded after ϵ_{y2}^{M}, the unloading is elastic along a curve with slope, E_d^{un}. For unloading along the martensite plateau, ***2***, the slope varies continuously from E_m at $\epsilon_{un} = \epsilon_{y1}^{M}$ to E_d^{un} at $\epsilon_{un} = \epsilon_{y2}^{M}$. The unloading slope or 'effective unloading modulus' depends therefore on the fractions of twinned and detwinned martensite, R_m^T, R_m^D, present in the material at the time of unloading. R_m^T, R_m^D are defined as:

$$\begin{aligned} R_m^D &= \frac{\epsilon_{un} - \epsilon_{y1}^{M}}{\epsilon_{y2}^{M} - \epsilon_{y1}^{M}} \\ R_m^T &= 1 - R_m^D. \end{aligned} \tag{3.36}$$

Given R_m^T, R_m^D, the effective unloading modulus is calculated as a weighted combination of the twinned and detwinned unload moduli. Equation (3.37) shows the parameters for martensite unloading.

$$E_{mart} = \begin{cases} E_m & \epsilon_{un} \leq \epsilon_{y1}^{M} \\ R_m^T E_m + R_m^D E_d^{un} & \epsilon_{y1}^{M} < \epsilon_{un} < \epsilon_{y2}^{M} \quad ; \qquad \epsilon_{off}^{M} = \epsilon_{un}, \quad \sigma_{off}^{M} = \sigma_{un}^{M} \\ E_d^{un} & \epsilon_{un} \geq \epsilon_{y2}^{M} \end{cases} \tag{3.37}$$

Austenite Unloading:

As described in Section 2.6, the unloading behaviour of austenite depends on the strain at which the unloading occurs. Unloading before the first austenite yield, $(\epsilon^{Ms}, \sigma^{Ms})$, is elastic. If the unloading occurs after $(\epsilon^{Ms}, \sigma^{Ms})$, then the SMA exhibits superelasticity. The equations defining unloading behaviour are discussed below.

$\epsilon_{un} \leq \epsilon^{Ms}$: The austenite unloading when $\epsilon_{un} \leq \epsilon^{Ms}$ is elastic and follows line ***4***. The moduli and offsets are:

$$\begin{aligned} E_{aust} &= E_a \\ \sigma_{off} &= \sigma^A_{un} \\ \epsilon_{off} &= \epsilon_{un} \end{aligned} \tag{3.38}$$

$\epsilon_{un} \geq \epsilon^{Mf}$: When $\epsilon_{un} \geq \epsilon^{Mf}$, the austenite unloading is elastic with modulus E^{un}_{sim}, following line ***7***) until the strain reduces to ϵ^{As}. At this strain, SIM starts to revert back to austenite. Upon further unloading, the material unload modulus is E_{am} and follows line ***8*** until a strain of ϵ^{Af}, where the material is pure austenite. Further unloading then follows line ***10***[1] with modulus E^{un}_a. This is the superelastic or pseudoelastic effect, and the moduli and offsets are given by:

$$E_{aust} \begin{cases} E^{un}_{sim}; \quad \sigma^A_{off} = \sigma^A_{un}; \quad \epsilon^A_{off} = \epsilon_{un}; \quad \epsilon \geq \epsilon^{As} \\ E_{am}; \quad \sigma^A_{off} = \sigma^{Af}; \quad \epsilon^A_{off} = \epsilon^{Af}; \quad \epsilon^{Af} < \epsilon < \epsilon^{As} \\ E^{un}_a; \quad \sigma^A_{off} = \sigma^{Af}; \quad \epsilon^A_{off} = \epsilon^{Af}; \quad \epsilon \leq \epsilon^{Af}. \end{cases} \tag{3.39}$$

[1]Note that lines ***4*** and ***10*** are often assumed to coincide. The proposed model allows for these to differ in order to better capture observed behaviour.

$\epsilon^{Ms} \leq \epsilon_{un} \leq \epsilon^{Mf}$: For unloading along the austenite plateau the slope varies continuously from E_a at $\epsilon_{un} = \epsilon^{Ms}$ to E^{un}_{sim} at $\epsilon_{un} = \epsilon^{Mf}$. The effective unloading modulus, E^{un}_{am} depends on the fractions of pure austenite and SIM, R_a and R_{sim} respectively, present in the material at the time of unloading. R_a, R_{sim} and E^{un}_{am} are defined as:

$$\begin{aligned} R_{sim} &= \tfrac{\epsilon_{un}-\epsilon^{Ms}}{\epsilon^{Mf}-\epsilon^{Ms}} \\ R_a &= 1 - R_{sim} \\ E^{un}_{am} &= R_a E_a + R_{sim} E^{un}_{sim}. \end{aligned} \tag{3.40}$$

Figure 3.10 shows the path of unloading from ***5***. The material unloads along ***9*** with slope E^{un}_{am} until a certain strain, ϵ^{int}. Point $(\epsilon^{int}, \sigma^{int})$ is the intersection of the unloading curve originating at $(\epsilon_{un}, \sigma^A_{un})$ with slope E^{un}_{am} and the unloading plateau, ***8***, with slope E_{am}. The equation for ϵ^{int} is

$$\epsilon^{int} = \frac{\sigma^{Af} - \sigma^A_{un} - \epsilon^{Af} E_{am} + \epsilon_{un} E^{un}_{am}}{E^{un}_{am} - E_{am}}. \tag{3.41}$$

After the strain falls below ϵ^{int} further unloading proceeds along ***8*** with modulus E_{am} until a strain of ϵ^{Af}, where the material is pure austenite. The unloading then follows ***10*** with modulus E^{un}_a. The sequence for unloading along the austenite plateau is summarised in (3.42)

$$E_{aust} \begin{cases} E^{un}_{am}; \quad \sigma^A_{off} = \sigma^A_{un}; \quad \epsilon^A_{off} = \epsilon_{un}; \quad \epsilon \geq \epsilon^{int} \\ E_{am}; \quad \sigma^A_{off} = \sigma^{Af}; \quad \epsilon^A_{off} = \epsilon^{Af}; \quad \epsilon^{Af} < \epsilon < \epsilon^{int} \\ E^{un}_a; \quad \sigma^A_{off} = \sigma^{Af}; \quad \epsilon^A_{off} = \epsilon^{Af}; \quad \epsilon \leq \epsilon^{Af} \end{cases} \tag{3.42}$$

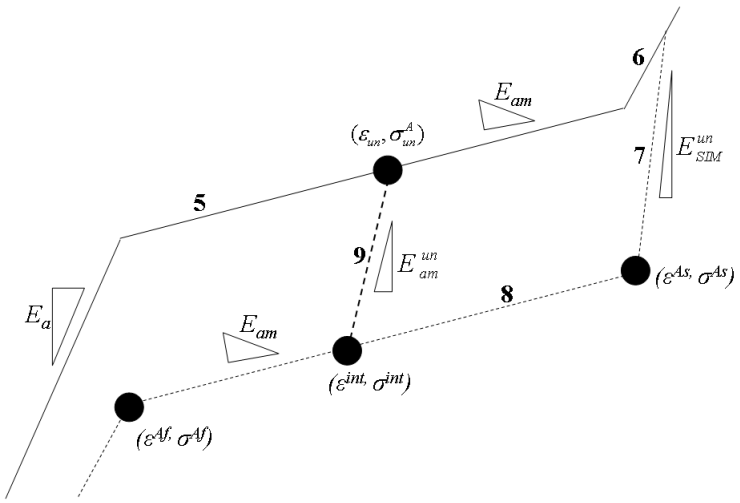

Figure 3.10: Unloading from austenite-SIM plateau, parallel model

Parallel Mechanical Model: Reloading

Reloading occurs when the stress changes direction from decreasing to increasing. It is important for the model to account for reloading so that cyclic loading and unloading can be simulated. At the point of reloading, let the strain in the material be ϵ_{rel} and the martensite and austenite stresses be σ^M_{rel} and σ^A_{rel}. As described in Section 2.6, when the strain begins to increase after a period of unloading, the reloading follows the same curve as the unloading until it rejoins the original loading curve at certain strain called the *reload limit*, ϵ_{rl}. The exception to this is shown in Figure 3.11 where unloading first occurs at $(\epsilon_{un}, \sigma^A_{un})$ and reloading occurs at $(\epsilon_{rel}, \sigma^A_{rel})$. In this case, the reloading modulus, E_{rel}, is calculated based upon the fractions of pure austenite and SIM present at the reloading point (3.43). The reload limit in this case is called ϵ'_{rl} and it is the intersection of this new loading line and segment ***5*** (see Figure 3.11).

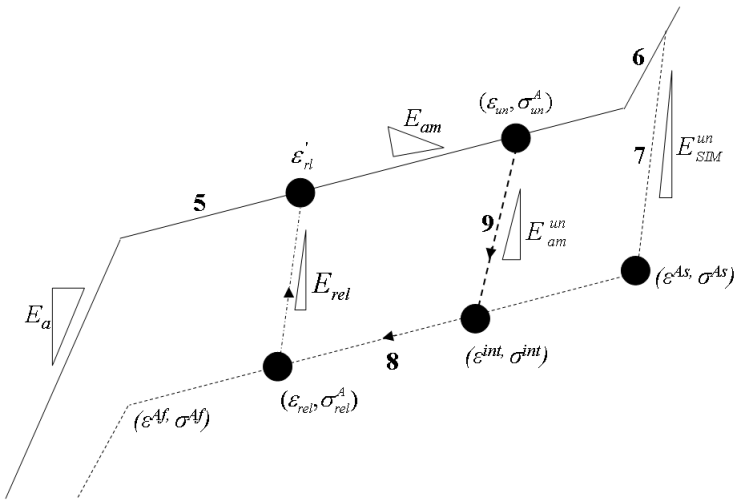

Figure 3.11: Reloading behaviour of austenite when on unloading plateau, parallel model

$$\begin{aligned} R_{sim} &= \tfrac{\epsilon_{un}-\epsilon^{Af}}{\epsilon^{As}-\epsilon^{Af}} \\ R_a &= (1 - R_{sim}) \\ E_{rel} &= R_a E_a + R_{sim} E^{un}_{sim} \\ \epsilon'_{rl} &= \tfrac{\sigma^{Ms}-\sigma^A_{rel}-\epsilon^{Ms}E_{am}+\epsilon_{rel}E_{rel}}{E_{rel}-E_{am}} \end{aligned} \tag{3.43}$$

Consider the case where the SMA is a combination of martensite and austenite and reloading were to occur at ϵ_{un} in Figure 3.12. The unload stresses are σ^M_{un} for martensite and σ^A_{un} for austenite. Martensite will unload along ***12*** and austenite along ***9*** and ***8***. Since deviation from the loading curve occurs at ϵ_{un}, the reload limit, $\epsilon_{rl} = \epsilon_{un}$. Now consider that reloading were to occur at ϵ_{rel} as shown in the figure. Martensite would continue to follow ***12*** but in the opposite

direction. Since reloading in austenite occurs along ***8***, the austenite reloading path is along ***11*** and the reload limit is ϵ'_{rl}. When the strain reaches ϵ'_{rl}, however, an examination of Figure 3.12 shows that although the austenite has rejoined the original loading curve, martensite has not. Therefore, the material is not under normal loading until the strain has increased to ϵ_{rl}. Note that ϵ'_{rl} is always less than ϵ_{rl}. The E_{aust} and E_{mart} values at the point of reloading are stored as E^{rel}_{aust} and E^{rel}_{mart} respectively. When the strain is below ϵ_{rl}, the reloading parameters are in (3.44) with the extension involving ϵ'_{rl} in (3.45)

for $\epsilon \leq \epsilon_{rl}$:

$$\begin{cases} E_{mart} = E^{rel}_{mart}; \quad \sigma^M_{off} = \sigma^M_{rel}; \quad \epsilon^M_{off} = \epsilon_{rel} \\ E_{aust} = E^{rel}_{aust}; \quad \sigma^A_{off} = \sigma^A_{rel}; \quad \epsilon^A_{off} = \epsilon_{rel} \end{cases} \tag{3.44}$$

for $\epsilon'_{rl} \leq \epsilon \leq \epsilon_{rl}$:

$$E_{aust} \begin{cases} E_{sim}; \quad \sigma^A_{off} = \sigma^{Mf}; \quad \epsilon^A_{off} = \epsilon^{Mf}; \quad \epsilon \geq \epsilon^{Mf} \\ E_{am}; \quad \sigma^A_{off} = \sigma^{Ms}; \quad \epsilon^A_{off} = \epsilon^{Ms}; \quad \epsilon < \epsilon^{Mf} \end{cases} \tag{3.45}$$

Once the stress has reached the reload limit, the loading then continues as if it were normal loading.

3.6.3 Series Mechanical Model

In the series SMA configuration, the austenite and martensite layers are in series with each other, and each phase experiences the same applied stress. Since the moduli of the phases are different,

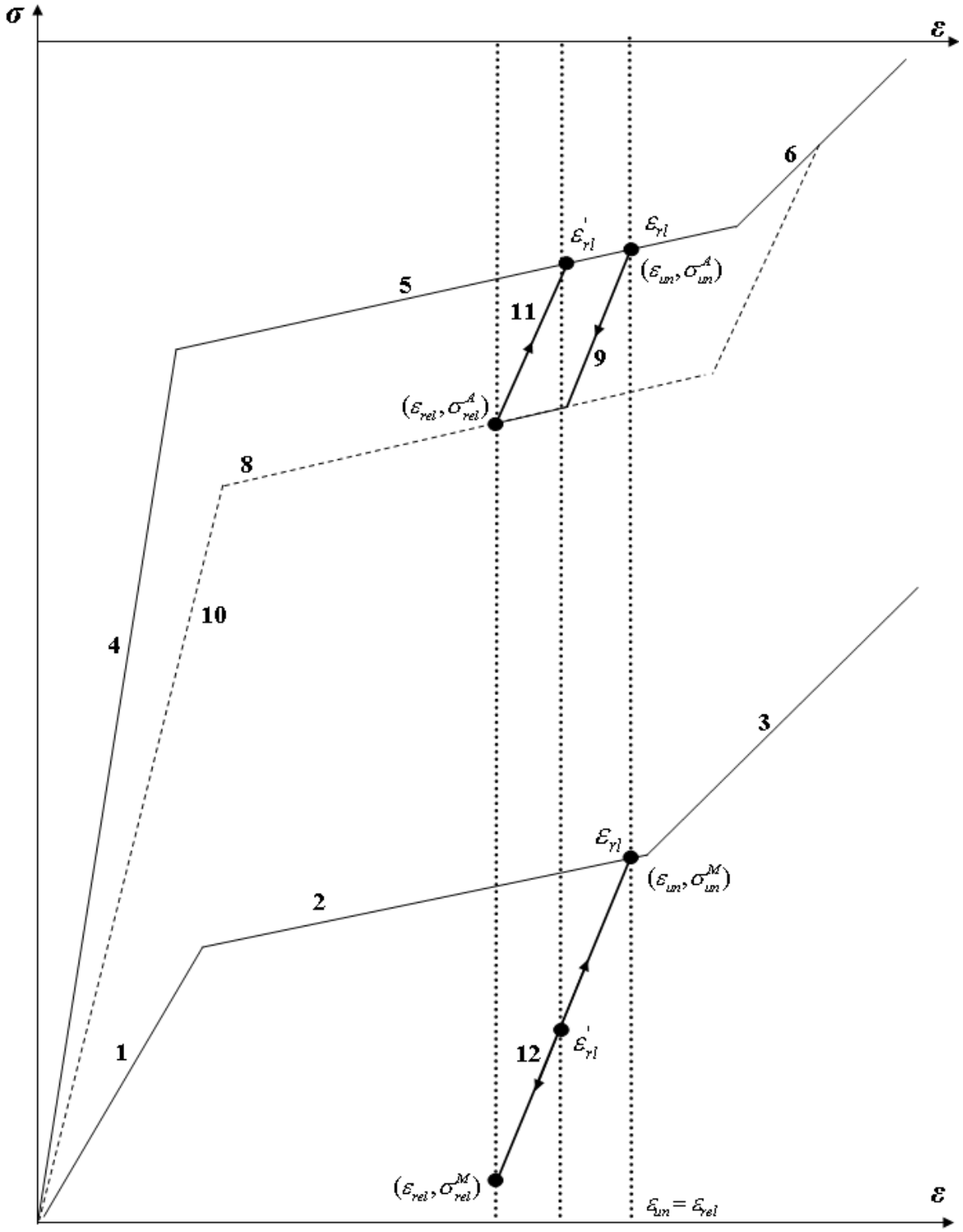

Figure 3.12: Stress-strain reloading behaviour, parallel mechanical model

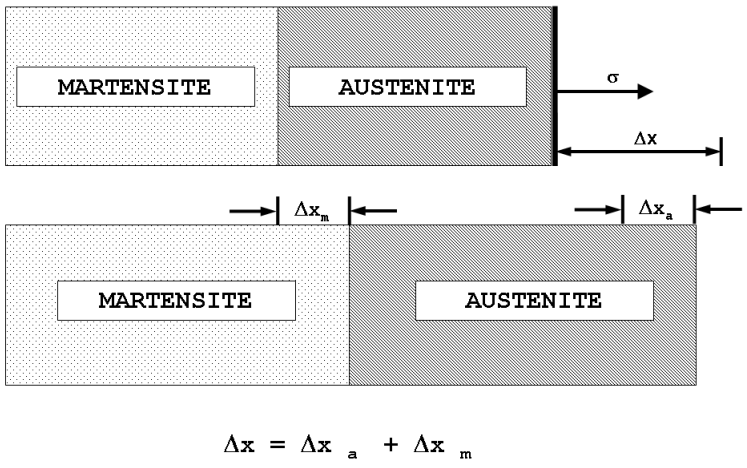

Figure 3.13: Series mechanical model

the strains in each of the phases is different and each contribute to the final strain in the SMA. Figure 3.13 shows the series SMA configuration.

Series Mechanical Model: Characteristic Equation

Within each of the phases, the strain due to the applied stress is determined based on the characteristic stress-strain curve for SMA, shown in Figure 3.8. The phase strains are then weighted by the phase fraction and combined, as shown in (3.46).

$$\begin{aligned} \epsilon &= R_m\epsilon_m + (1-R_m)\epsilon_a \\ &= R_m\left(\frac{\sigma-\sigma_{off}^M}{E_{mart}} + \epsilon_{off}^M\right) + (1-R_m)\left(\frac{\sigma-\sigma_{off}^A}{E_{aust}} + \epsilon_{off}^A\right) \end{aligned} \tag{3.46}$$

The phase moduli E_{mart} and E_{aust} as well as the offsets σ_{off} and ϵ_{off} are determined based on whether the material is being loaded or unloaded and the region in the stress-strain curve of Figure 3.8. Loading and unloading for the series model are very similar to the parallel model. The major difference is that while the parallel model distinguishes the regions of Figure 3.8 using strain, the series model uses stress.

Series Mechanical Model: Loading

Similar to the parallel case, the loading behaviour comprises regions ***1-3*** for martensite loading and regions ***4-6*** for austenite loading in Figure 3.8. The modulus and offsets for each region are in (3.47) and (3.48).

$$E_{mart} = \begin{cases} E_m; & \sigma_{off}^M = 0; & \epsilon_{off}^M = 0; & \sigma \leq \sigma_{y1}^M \\ E_t; & \sigma_{off}^M = \sigma_{y1}^M; & \epsilon_{off}^M = \epsilon_{y1}^M; & \sigma_{y1}^M < \sigma < \sigma_{y2}^M \\ E_d; & \sigma_{off}^M = \sigma_{y2}^M; & \epsilon_{off}^M = \epsilon_{y2}^M; & \sigma \geq \sigma_{y2}^M \end{cases} \tag{3.47}$$

$$E_{aust} = \begin{cases} E_a; & \sigma_{off}^A = 0; & \epsilon_{off}^A = 0; & \sigma \leq \sigma^{Ms} \\ E_{am}; & \sigma_{off}^A = \sigma^{Ms}; & \epsilon_{off}^M = \epsilon^{Ms}; & \sigma^{Ms} < \sigma < \sigma^{Mf} \\ E_{sim}; & \sigma_{off}^A = \sigma^{Mf}; & \epsilon_{off}^A = \epsilon^{Mf}; & \sigma \geq \sigma^{Mf} \end{cases} \tag{3.48}$$

Series Mechanical Model: Unloading

At the point of unloading, the stress and strain in the material is σ_{un} and ϵ_{un}. Since in the series model, the austenite and martensite stresses are equal, $\sigma_{un}^M = \sigma_{un}^A = \sigma_{un}$; and ϵ_{un}^M and ϵ_{un}^A are

calculated using (3.30) to get (3.49):

$$\begin{aligned} \epsilon_{un}^{M} &= \tfrac{\sigma_{un}-\sigma_{off}^{M}}{E_{mart}} + \epsilon_{off}^{M} \\ \epsilon_{un}^{A} &= \tfrac{\sigma_{un}-\sigma_{off}^{A}}{E_{aust}} + \epsilon_{off}^{A} \end{aligned} \tag{3.49}$$

Martensite Unloading:

As described in Section 2.6, the unloading behaviour of martensite depends on the strain at which the unloading occurs. Martensite unloading is elastic with modulus, E_m if unloaded before σ_{y1}^{M}. If unloaded after σ_{y2}^{M}, the unloading is elastic along the curve with slope, E_d^{un}. For unloading along the martensite plateau, ***2***, the slope varies continuously from E_m at $\sigma_{un} = \sigma_{y1}^{M}$ to E_d^{un} at $\sigma_{un} = \sigma_{y2}^{M}$. The unloading slope or 'effective unloading modulus' depends on the fractions of twinned and detwinned martensite, R_m^T, R_m^D, present in the material at the time of unloading. Given R_m^T and R_m^D, the effective unloading modulus is calculated as a weighted combination of the twinned and detwinned unload moduli as follows:

$$\begin{aligned} R_m^D &= \tfrac{\sigma_{un}-\sigma_{y1}^{M}}{\sigma_{y2}^{M}-\sigma_{y1}^{M}} \\ R_m^T &= 1 - R_m^D \\ E_t^{un} &= \left(\tfrac{R_m^D}{E_d^{un}} + \tfrac{R_m^T}{E_m}\right)^{-1}. \end{aligned} \tag{3.50}$$

(3.51) shows the moduli for martensite unloading. The offsets are ϵ_{un}^{M} and σ_{un}.

$$E_{mart} = \begin{cases} E_m; & \sigma_{un} \leq \sigma_{y1}^{M} \\ E_t^{un}; & \sigma_{y1}^{M} < \sigma_{un} < \sigma_{y2}^{M} \quad ; \qquad \epsilon_{off}^{M} = \epsilon_{un}^{M}, \quad \sigma_{off}^{M} = \sigma_{un} \\ E_d^{un}; & \sigma_{un} \geq \sigma_{y2}^{M} \end{cases} \tag{3.51}$$

Austenite Unloading:

As described in Section 2.6, the unloading behaviour of austenite depends on the strain at which the unloading occurs. Unloading before the first austenite yield, $(\epsilon^{Ms}, \sigma^{Ms})$, is elastic. If the unloading occurs after $(\epsilon^{Ms}, \sigma^{Ms})$, then the SMA exhibits superelasticity. The equations defining unloading behaviour are discussed below.

$\sigma_{un} \leq \sigma^{Ms}$: The austenite unloading when $\sigma_{un} \leq \sigma^{Ms}$ is elastic and follows line ***4***. The moduli and offsets are:

$$\begin{aligned} E_{aust} &= E_a \\ \sigma_{off}^{A} &= \sigma_{un} \\ \epsilon_{off}^{A} &= \epsilon_{un}^{A}. \end{aligned} \tag{3.52}$$

$\sigma_{un} \geq \sigma^{Mf}$: When $\sigma_{un} \geq \sigma^{Mf}$, the austenite unloading is elastic with modulus E_{sim}^{un} and follows line ***7*** until the stress reduces to σ^{As}. At this stress, SIM starts to revert back to austenite. Upon further unloading, the material unload modulus is E_{am} and follows line ***8*** until a stress of σ^{Af}, where the material is pure austenite. Further unloading then follows line ***10*** with modulus

E_a^{un}. This is the superelastic or pseudoelastic effect, and the moduli and offsets are given by:

$$E_{aust}\begin{cases} E_{sim}^{un}; \quad \sigma_{off}^{A} = \sigma_{un}; \quad \epsilon_{off}^{A} = \epsilon_{un}^{A}; \quad \sigma \geq \sigma^{As} \\ E_{am}; \quad \sigma_{off}^{A} = \sigma^{Af}; \quad \epsilon_{off}^{A} = \epsilon^{Af}; \quad \sigma^{Af} < \sigma < \sigma^{As} \\ E_{a}^{un}; \quad \sigma_{off}^{A} = \sigma^{Af}; \quad \epsilon_{off}^{A} = \epsilon^{Af}; \quad \sigma \leq \sigma^{Af}. \end{cases} \tag{3.53}$$

$\sigma^{Ms} \leq \sigma_{un} \leq \sigma^{Mf}$: For unloading along the austenite plateau the slope varies continuously from E_a at $\sigma_{un} = \sigma^{Ms}$ to E_{sim}^{un} at $\sigma_{un} = \sigma_{Mf}$. The effective unloading modulus, E_{am}^{un} depends on the fractions of pure austenite and SIM, R_a and R_{sim} respectively, present in the material at the time of unloading.

$$\begin{aligned} R_{sim} &= \tfrac{\sigma_{un}-\sigma^{Ms}}{\sigma^{Mf}-\sigma^{Ms}} \\ R_a &= 1 - R_{sim} \\ E_{am}^{un} &= \left(\tfrac{R_a}{E_a} + \tfrac{R_{sim}}{E_{sim}^{un}}\right)^{-1} \end{aligned} \tag{3.54}$$

Figure 3.14 shows the path of unloading from ***5***. The material unloads along ***9*** with slope E_{am}^{un} until a certain stress, σ^{int}. Point $(\epsilon^{int}, \sigma^{int})$ is the intersection of the unloading curve originating at $(\epsilon_{un}^{A}, \sigma_{un})$ with slope E_{am}^{un} and the unloading plateau, ***8***, with slope E_{am}. The equation for σ_{int} is in (3.55)

$$\sigma^{int} = \frac{\frac{\sigma_{un}}{E_{am}^{un}} - \epsilon_{un}^{A} - \frac{\sigma_{Af}}{E_{am}} + \epsilon^{Af}}{\frac{1}{E_{am}^{un}} - \frac{1}{E_{am}}}. \tag{3.55}$$

After the stress falls below σ^{int} further unloading proceeds along ***8*** with modulus E_{am} until a stress of σ^{Af}, where the material is pure austenite. The unloading then follows ***10*** with modulus

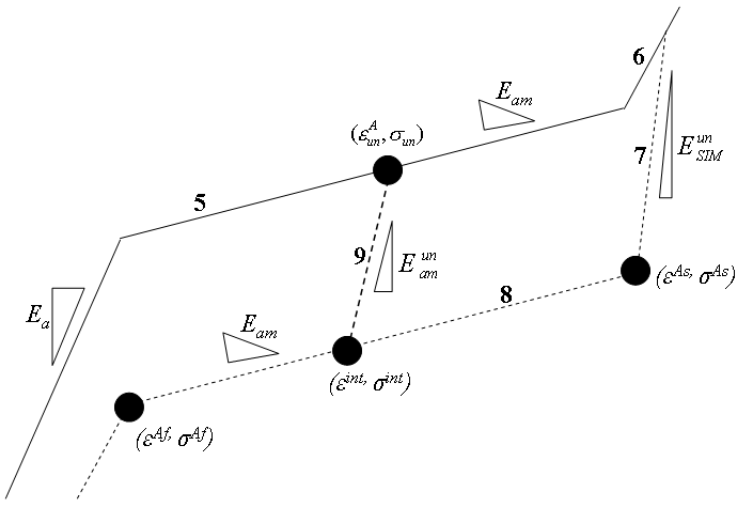

Figure 3.14: Unloading from austenite-SIM plateau, series model

E^{un}_a. The sequence for unloading along the austenite plateau is summarised in (3.56)

$$E_{aust} \begin{cases} E^{un}_{am}; \quad \sigma^A_{off} = \sigma_{un}; \quad \epsilon^A_{off} = \epsilon^A_{un}; \quad \sigma \geq \sigma^{int} \\ E_{am}; \quad \sigma^A_{off} = \sigma^{Af}; \quad \epsilon^A_{off} = \epsilon^{Af}; \quad \sigma^{Af} < \sigma < \sigma^{int} \\ E^{un}_a; \quad \sigma^A_{off} = \sigma^{Af}; \quad \epsilon^A_{off} = \epsilon^{Af}; \quad \sigma \leq \sigma^{Af} \end{cases} \tag{3.56}$$

Series Mechanical Model: Reloading

Reloading occurs when the stress changes direction from decreasing to increasing. At the point of reloading, the stress in the material is σ_{rel} and the martensite and austenite strains are ϵ^M_{rel} and ϵ^A_{rel}, found by substituting σ_{rel} and either E_{mart} or E_{aust} into (3.30). As described in Section 2.6, when the stress begins to increase after a period of unloading, the reloading follows the same

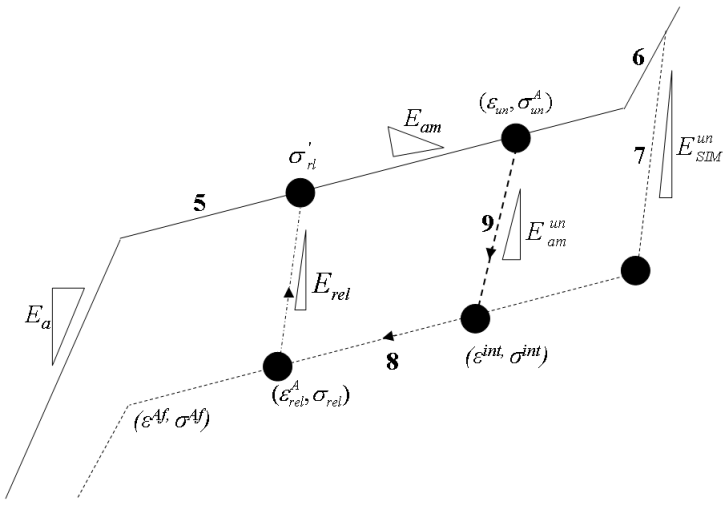

Figure 3.15: Reloading behaviour of austenite when on unloading plateau, series model

curve as the unloading until it rejoins the original loading curve at certain strain called the *reload limit*, σ_{rl}. The exception to this is shown in Figure 3.15 where unloading first occurs at $(\sigma^A_{un}, \sigma_{un})$ and reloading occurs at $(\epsilon^A_{rel}, \sigma_{rel})$. In this case, the reloading modulus, E_{rel}, is calculated based upon the fractions of pure austenite and SIM present at the reloading point (3.57). The reload limit in this case is called σ'_{rl} and it is the intersection of this new loading line and segment ***5*** (see Figure 3.15).

$$\begin{aligned} R_{sim} &= \tfrac{\sigma_{un}-\sigma^{Af}}{\sigma^{As}-\sigma^{Af}} \\ R_a &= (1-R_{sim}) \\ E_{rel} &= \left(\tfrac{R_a}{E_a}+\tfrac{R_{sim}}{E^{un}_{sim}}\right)^{-1} \\ \sigma'_{rl} &= \tfrac{\frac{\sigma_{rel}}{E_{rel}}-\epsilon^A_{rel}-\frac{\sigma_{Ms}}{E_{am}}+\epsilon^{Ms}}{\frac{1}{E_{rel}}-\frac{1}{E_{am}}} \end{aligned} \tag{3.57}$$

Consider the case where the SMA is a combination of martensite and austenite and reloading were to occur at σ_{un} in Figure 3.16. The unload strains are ϵ^M_{un} for martensite and ϵ^A_{un} for austenite. Martensite will unload along ***12*** and austenite along ***9*** and ***8***. Since deviation from the loading curve occurs at ϵ_{un}, the reload limit, $\sigma_{rl} = \sigma_{un}$. Now consider that reloading were to occur at σ_{rel} as shown in the figure. Martensite would continue to follow ***12*** but in the opposite direction. Since reloading in austenite occurs along ***8***, the austenite reloading path is along ***11*** and the reload limit is σ'_{rl}. When the strain reaches σ'_{rl}, however, an examination of Figure 3.16 shows that although the austenite has rejoined the original loading curve, martensite has not. Therefore, the material is not under normal loading until the stress has increased to σ_{rl}. Note that σ'_{rl} is always less than σ_{rl}. The E_{aust} and E_{mart} values at the point of reloading are stored as E^{rel}_{aust} and E^{rel}_{mart} respectively. When the stress is below σ_{rl}, the reloading parameters are in (3.58) with the extension involving σ'_{rl} in (3.59)

for $\sigma \leq \sigma_{rl}$:

$$\begin{cases} E_{mart} = E^{rel}_{mart}; \quad \sigma^M_{off} = \sigma_{rel}; \quad \epsilon^M_{off} = \epsilon^M_{rel} \\ E_{aust} = E^{rel}_{aust}; \quad \sigma^A_{off} = \sigma_{rel}; \quad \epsilon^A_{off} = \epsilon^A_{rel} \end{cases} \tag{3.58}$$

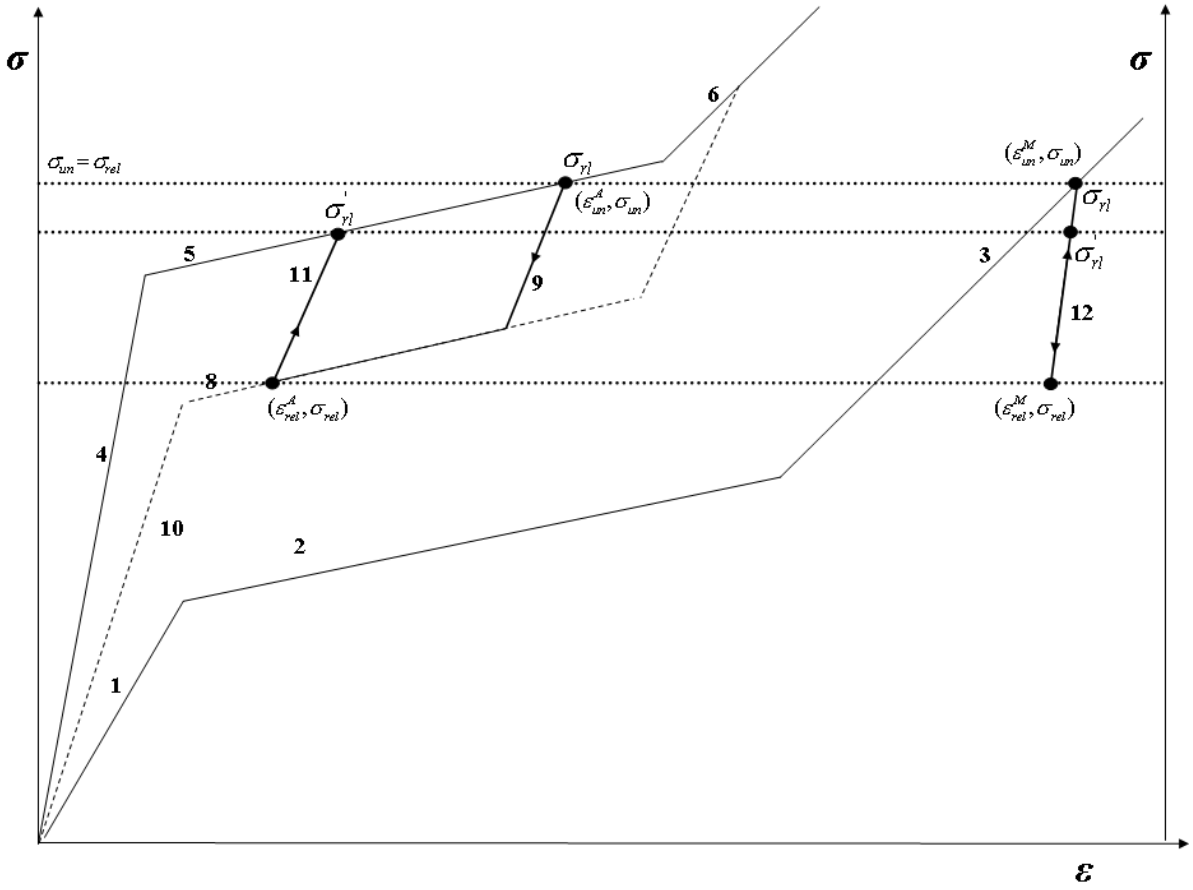

Figure 3.16: Stress-strain reloading behaviour, series mechanical model

$$\text{for } \sigma'_{rl} \leq \sigma \leq \sigma_{rl}\text{:}$$

$$E_{aust} \begin{cases} E_{sim}; \quad \sigma^A_{off} = \sigma^{Mf}; \quad \epsilon^A_{off} = \epsilon^{Mf}; \quad \epsilon \geq \epsilon^{Mf} \\ E_{am}; \quad \sigma^A_{off} = \sigma^{Ms}; \quad \epsilon^A_{off} = \epsilon^{Ms}; \quad \epsilon < \epsilon^{Mf} \end{cases} \tag{3.59}$$

Once the stress has reached the reload limit, the loading then continues as if it were normal loading.

3.7 Inverse Mechanical Model

The inverse mechanical model uses stress and strain information to calculate the martensite fraction, R_m. Using the forward model equations for either the series or parallel configurations (3.33-3.58), and knowledge of the strain variation and previous state, the strains ϵ_{Mart} and ϵ_{Aust} can be determined from σ in the series configuration and the stresses σ_{Mart} and σ_{Aust} can be determined from ϵ in the parallel configuration. R_m is then calculated by inverting the model equation. Equations (3.60) and (3.61) show the calculation of R_m for the parallel and series mechanical model:

Parallel Mechanical Model:

$$R_m = \frac{\sigma - \sigma_{aust} - E_{aust}\epsilon}{E_{mart}\epsilon - E_{Aust}\epsilon + \sigma_{mart} - \sigma_{aust}} \tag{3.60}$$

Series Mechanical Model:

$$R_m = \frac{\epsilon - \epsilon_{aust}}{\epsilon_{mart} - \epsilon_{aust}} \tag{3.61}$$

3.8 Spring-biased SMA Actuator

This section derives the characteristic stress-strain behaviour of a spring-biased SMA actuator. It is effectively a specific application of the general mechanical model of Section 3.6, where stress is constrained by the spring force. The spring-biased SMA actuator is a common actuator configuration discussed previously in Section 2.7. The following points summarise the basic actuation mechanism:

- The total strain in the configuration is set at the start of the experiment and is fixed for the duration of the experiment.
- Individual strains in the spring and SMA vary based on temperature of the SMA.
- The actuator is quasi-static, therefore the forces in the spring and the SMA are equal.

Equation 3.62 presents the linear model of the spring, where the stress, σ, strain, ϵ_s and modulus, E_s, of the spring are effective values.

Spring Equation:

$$\epsilon_s = \frac{\sigma}{E_s} \tag{3.62}$$

The strains in the SMA are shown in 3.63 and 3.64 for both the parallel and series models.

SMA Equation, Parallel Model: From (3.32)

$$\epsilon_{SMA} = \frac{\overbrace{\sigma - R_m(\sigma^M_{off} - \epsilon^M_{off}E_{mart}) - (1-R_m)(\sigma^A_{off} - \epsilon^A_{off}E_{aust})}^{\sigma^R}}{\underbrace{R_m E_{mart} + (1-R_m)E_{aust}}_{E_{SMA}}} \tag{3.63}$$

SMA Equation, Series Model: From (3.46)

$$\begin{aligned}\epsilon_{SMA} &= R_m\epsilon_m + (1-R_m)\epsilon_a \\ &= R_m\left(\frac{\sigma-\sigma_{off}^M}{E_{mart}}+\epsilon_{off}^M\right)+(1-R_m)\left(\frac{\sigma-\sigma_{off}^A}{E_{aust}}+\epsilon_{off}^A\right)\end{aligned} \tag{3.64}$$

The spring imposes a time varying stress onto the SMA wire. The stress is based upon the total strain in the system and temperature of SMA, which in turn determines the modulus and strain in the SMA. The stress is calculated in (3.66) and (3.67) for the parallel and series mechanical SMA models respectively. ϵ_{tot} is the total strain in the system.

Parallel Model:

$$\begin{aligned}\epsilon_s + \epsilon_{SMA} &= \epsilon_{tot} \\ \frac{\sigma}{E_s}+\frac{\sigma+\sigma^R}{E_{SMA}} &= \epsilon_{tot} \\ \sigma\left(\frac{1}{E_s}+\frac{1}{E_{SMA}}\right) &= \epsilon_{tot}-\frac{\sigma^R}{E_{SMA}} \\ \sigma &= \frac{\epsilon_{tot}-\frac{\sigma^R}{E_{SMA}}}{\frac{1}{E_s}+\frac{1}{E_{SMA}}}\end{aligned} \tag{3.65}$$

Series Model:

$$\begin{aligned} \epsilon_s + \epsilon_{SMA} &= \epsilon_{tot} \qquad (3.66) \\ \sigma\left(\frac{1}{E_s} + \frac{1-R_m}{E_{aust}} + \frac{R_m}{E_{mart}}\right) &= \epsilon_{tot} + \underbrace{R_m\left(\frac{\sigma^M_{off}}{E_{mart}} - \epsilon^M_{off}\right) + (1-R_m)\left(\frac{\sigma^A_{off}}{E_{aust}} - \epsilon^A_{off}\right)}_{\epsilon^R} \\ \sigma &= \frac{\epsilon_{tot} + \epsilon^R}{\frac{1}{E_s} + \frac{1-R_m}{E_{aust}} + \frac{R_m}{E_{mart}}} \end{aligned}$$

The path of travel in a stress-strain plot for a spring-biased SMA actuator is shown in Figure 3.17 with different total strains. The slope of the travel path is the spring modulus in pascals. The figure shows that the spring applies a time varying stress to the SMA.

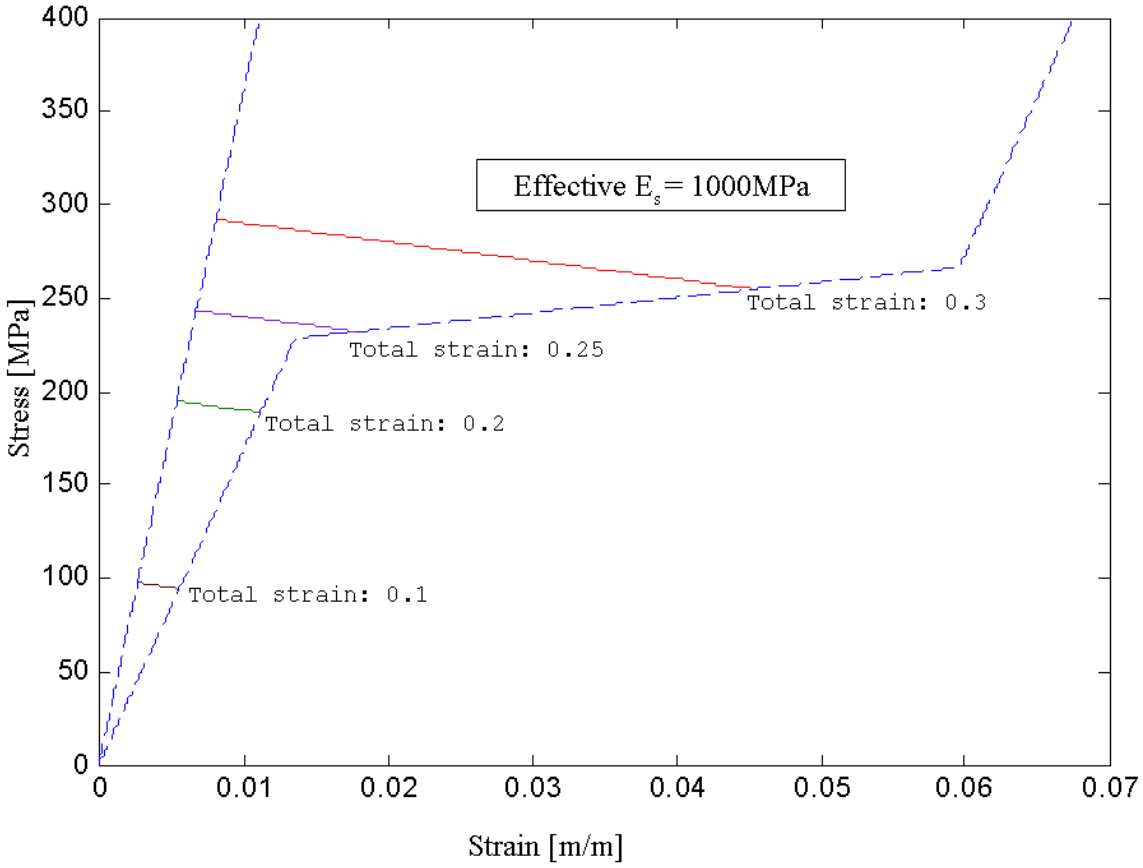

Figure 3.17: Stress-strain path of spring-biased SMA actuator for different total strains

Chapter 4

Qualitative Model Verification Via Simulation

This chapter presents qualitative simulations to validate that the model developed in Chapter 3 is able to reproduce observed SMA behaviour described in Chapter 2. Section 4.1 presents the model parameters that were used for the simulations. The simulation results are presented in Section 4.2 for the phase kinetics model and in Section 4.3 for the mechanical model.

4.1 Model Parameters

Tables 4.1 and 4.2 show the parameters for the phase kinetics and mechanical model simulations. The parameters in Table 4.1 were taken from the SMA wire data sheet [30] from the wire supplier Dynalloy Incorporated and from [10].

The parameters for the mechanical model were extracted by imposing varying stress pro-

Table 4.1: Phase kinetics model simulation parameters

Parameter	*Value*	*Units*
D_{SMA}	250.00	μm
L_{SMA}	0.1	m
k_m^C	0.053	^{o}C
k_m^H	0.07	^{o}C
h	70.00	W/m^{2}K
R(linear)	20.00	Ω/m
C	320.00	J/kg^{o}C
ρ	6450.00	kg/m^3
c_m	0.23	^{o}C/MPa
M_s	72.00	^{o}C
M_f	62.00	^{o}C
A_s	88.00	^{o}C
A_f	98.00	^{o}C
T_∞	22.00	^{o}C

files onto the SMA wire at different temperatures in order to reproduce loading, unloading and reloading behaviour. The wire was pulled at both low and high temperature beyond the second yield points and unloaded. From these tests, the moduli, $E_m, E_t, E_d, E_d^{un}, E_a, E_{am}, E_{sim}, E_a^{un}$ and E_{sim}^{un} were extracted by fitting linear lines through each of the regions on the two stress-strain curves. The yield stresses and strains, $\sigma_{y1}^M, \sigma_{y2}^M, \sigma^{Ms}, \sigma^{Mf}, \sigma^{Af}, \epsilon_{y1}^M, \epsilon_{y2}^M, \epsilon^{Ms}, \epsilon^{Mf}, \epsilon^{As}$ and ϵ^{Af} were found by extracting the co-ordinates of the intersection points of the linear fits to the stress-strain data.

Table 4.2: Mechanical model simulation parameters

Parameter	*Value*	*Units*
E_a	50.00	GPa
E_{am}	1.17	GPa
E_{sim}	12.00	GPa
E_{sim}^{un}	26.80	GPa
E_m	4.50	GPa
E_t	330.52	MPa
E_d	1.65	GPa
E_{un}^{d}	8.20	GPa
E_{un}^{a}	25.00	GPa
σ_{y1}^{M}	27.00	MPa
σ_{y2}^{M}	44.02	MPa
σ^{Ms}	465.00	MPa
σ^{Mf}	525.49	MPa
σ^{Af}	160.00	MPa
ϵ_{y1}^{M}	0.006	m/m
ϵ_{y2}^{M}	0.0575	m/m
ϵ^{Ms}	0.0093	m/m
ϵ^{Mf}	0.0610	m/m
ϵ^{Af}	0.0076	m/m
ϵ^{As}	0.0509	m/m

4.2 Phase Kinetics Model Simulation Results

This section validates in simulation the qualitative behaviour of the following model key properties:

- The relationship between the transformation temperatures, M_s, M_f, A_s, A_f and the hysteresis loop.
- The formation of major loops by complete temperature reversal.
- The formation of minor loop branches by partial temperature reversal.
- The shifting of the hysteresis loops due to changes in stress.
- The formation of closed hysteresis loops.
- Immediate output reversal on input reversal.

The simulation results for the phase kinetics model with time varying stress are shown in Figures 4.1 to 4.3. Figure 4.1 is the $R_m(\theta)$ curve for a decaying sinusoidal temperature input and constant stress. The decaying temperature causes the SMA to go through a number of minor hysteresis loops as observed in Figure 4.1. Figure 4.2 is the $R_m(\theta)$ curve for for a decaying sinusoidal temperature input and a ramp stress. When a ramp stress is applied, the resulting increase in transformation temperatures causes minor loops of Figure 4.1 to shift to higher temperatures.

Figure 4.3a shows four $R_m(\theta)$ hysteresis curves, each for a triangular temperature profile and for different constant stress values of 0, 150, 300 and 450 MPa. The temperature profile is chosen to ensure that the entire major hysteresis loop is generated. It is observed that the hysteresis curve shifts towards higher temperatures with increasing stress. This is due to the effect of increasing

stress on transformation temperatures [11]. In Figure 4.3b, the SMA is subjected to a triangular stress profile while maintaining the triangular temperature profile from before. The stress and θ profiles are shown in the inset plot of Figure 4.3b. The input stress, σ, is twice θ. Therefore, when $\theta = 75^o$, $\sigma = 150MPa$ and from the figure is observed that at $\theta = 75^o$, the bold curve intersects with the hysteresis curve for a constant stress of 150MPa. Further examination of Figure 4.3b shows that this observation is consistent with ($\theta = 150^o$, $\sigma = 300MPa$) and ($\theta = 225^o$, $\sigma = 450MPa$) thereby confirming the ability of the extended model to smoothly account for time-varying stress in computing the martensite fraction, R_m.

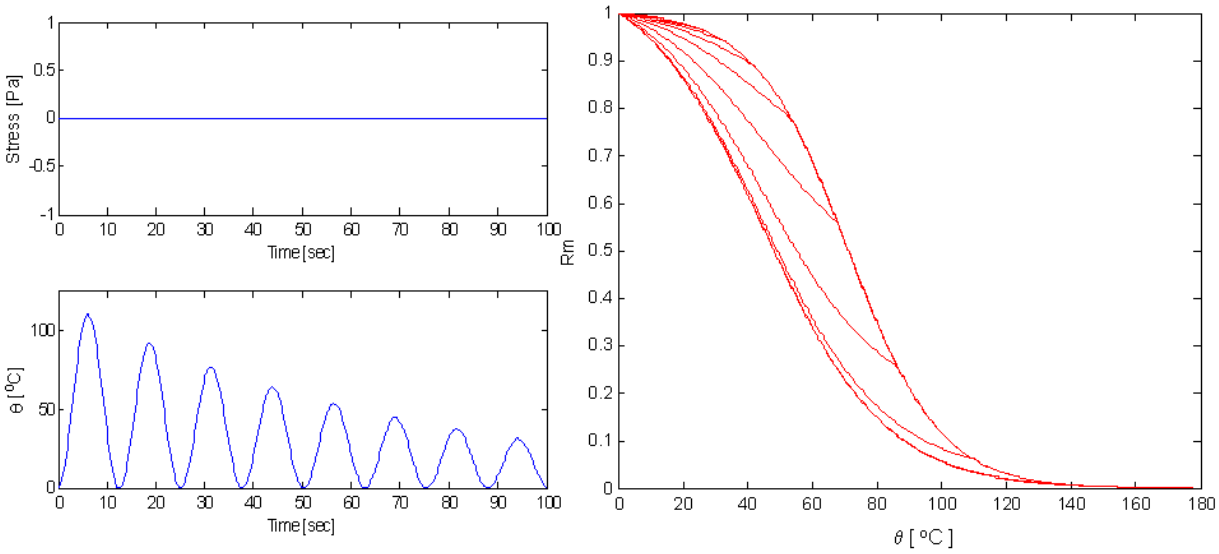

Figure 4.1: R_m vs θ with decaying sinusoidal θ profile and constant stress to show minor loop behaviour.

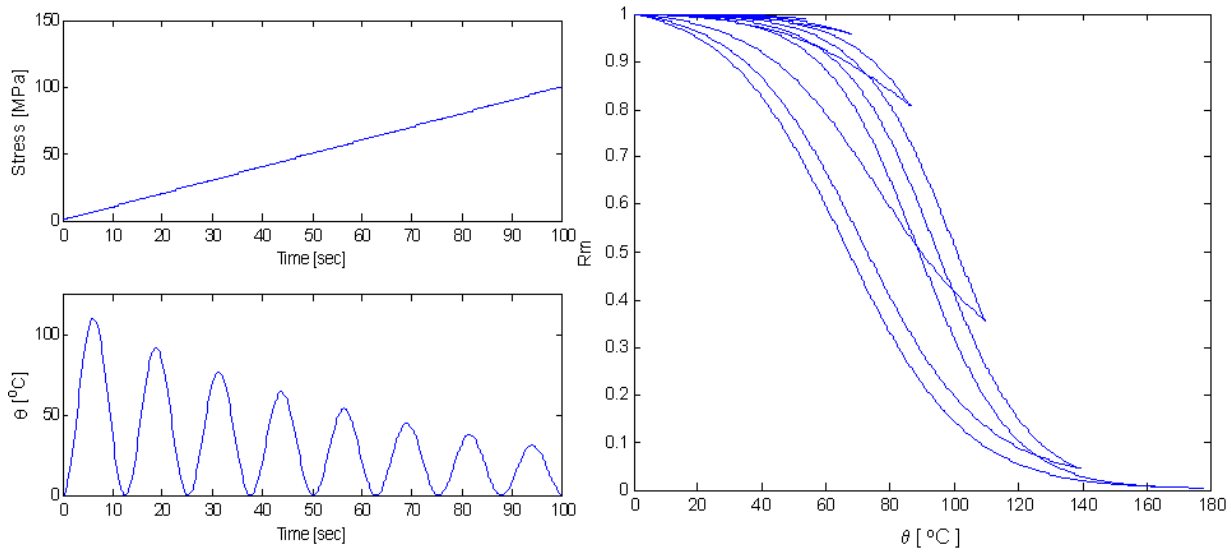

Figure 4.2: R_m vs θ with decaying sinusoidal θ profile and ramp stress to show minor loop behaviour.

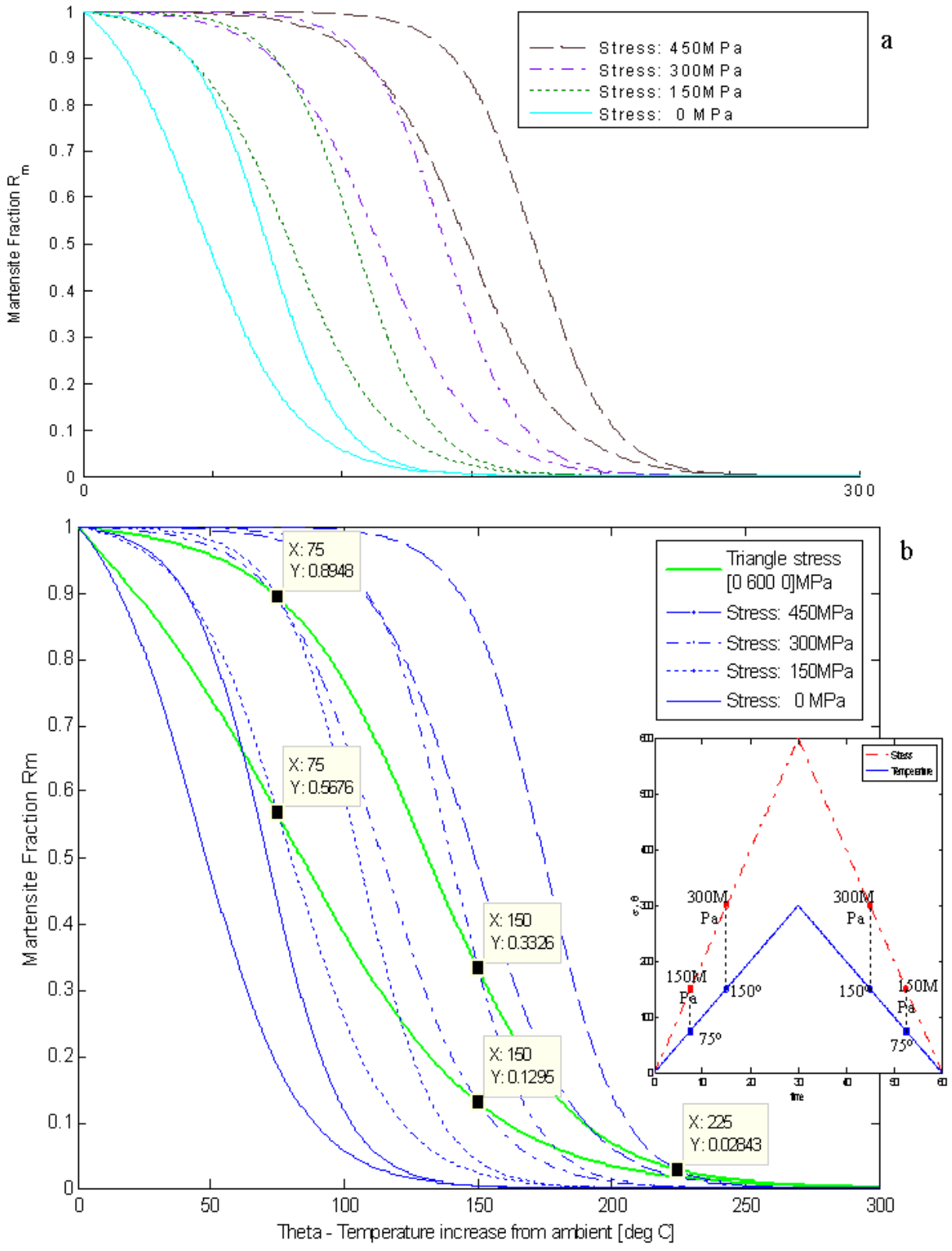

Figure 4.3: R_m vs θ with triangular θ profile and (a) constant stresses, (b) triangular stress.

4.3 Mechanical Model Simulation Results

This section presents the simulation results for both the parallel and series mechanical models for loading, unloading and reloading when in the martensite ($R_m = 1$) and austenite ($R_m = 0$) phases. Additionally, the simulation results for loading and unloading for values of R_m different from 1 and 0 are presented. The purpose of these simulations is to qualitatively confirm following key features of SMA mechanical behaviour for both the series and parallel mechanical models:

- The loading behaviour of martensite and austenite based upon percentages of martensite and austenite present in the material.
- The unloading and reloading behaviour of martensite based upon location of unloading in the stress-strain plane.
- The unloading and reloading behaviour of pure austenite.
- The unloading and reloading behaviour when the SMA is a combination of austenite and SIM. This includes superelasticity and the formation of minor superelastic loops on cyclic loading and unloading.
- The shifting of superelastic loops on increase of temperature above the austenite finish temperature, A_f.

4.3.1 Loading

Figure 4.4 shows the simulation results for loading when the SMA is in the austenite phase (solid curve) and martensite phase (dotted curve). Notice that the shape of the loading curves are the

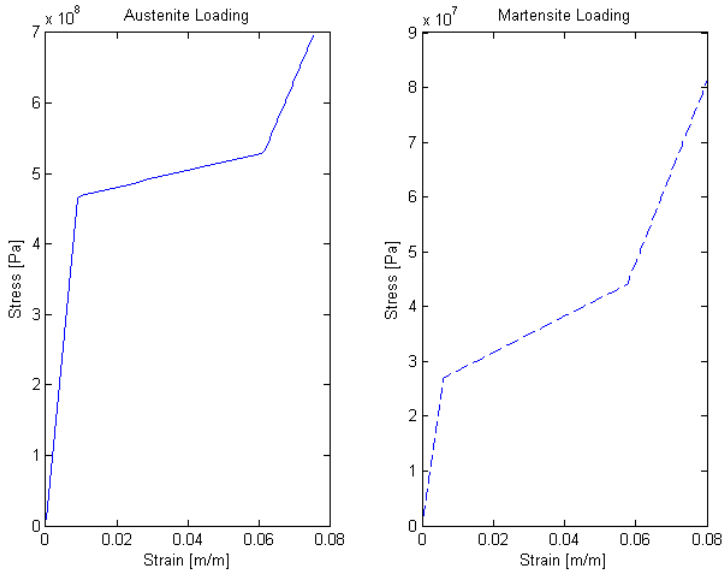

Figure 4.4: Simulation results for loading in austenite (solid) and martensite (dotted) phases

same for both phases but the Young's moduli and yield strengths are different. Note that when the material is composed of only one phase, either the series or parallel models will, as expected, produce the same results[1].

4.3.2 Unloading

Figure 4.5 shows the simulation results when the SMA is in its martensite phase. The upper plot shows the elastic behaviour of twinned martensite when the stress is increased to 20 MPa and then decreased to 0 MPa. The middle plot shows the unloading behaviour when the material is composed of a mixture of twinned and detwinned martensite when the stress in increased to

[1]The simulations were done for both parallel and series models to confirm this. The results for the parallel and series case are, as expected, identical, therefore only the parallel case is shown in the figure.

40 MPa and then decreased to 0 MPa. The arrangement of the twinned and detwinned layers affects the unloading modulus. This is seen by a difference in the unloading results using the series and parallel models. The difference, however, is not significant. The bottom plot shows the unloading behaviour for detwinned martensite when the stress is increased to 70 MPa and then decreased to 0 MPa. Unloading from either the martensite plateau or after complete detwinning results in residual strain.

Figure 4.6 shows the simulation results for unloading when the SMA is in its austenite phase. The top most plot shows the elastic behaviour of austenite below its first yield point when the stress is increased to 400 MPa and then decreased to 0 MPa. The middle plot shows the unloading behaviour when the material is composed of a mixture of austenite and stress induced martensite when the stress is increased to 500 MPa and then decreased to 0 MPa. The arrangement of the austenite and SIM layers affects the unloading modulus. This is seen by a difference in the unloading results using the series and parallel models. Similar to martensite unloading in the plateau, the difference is not significant. The bottom plot shows the unloading behaviour for SIM when the stress is increased to 580 MPa and then decreased to 0 MPa. The unloading behavior in the SIM plateau and for pure SIM exhibits the pseudoelastic behaviour.

4.3.3 Reloading

Figure 4.7 shows the simulation results for reloading when the SMA is in the martensite phase. The bottom plot shows the results of applying the stress profile shown in the top plot. The stress profile first increases from 0 MPa to 20 MPa and then back down to 0 MPa. Since 20 MPa is below the twinned martensite yield point, σ_{y1}^{M}, the material behaves elastically. Next, the stress increases to 28 MPa causing some of the martensite to detwin. Unloading back to 0 MPa causes

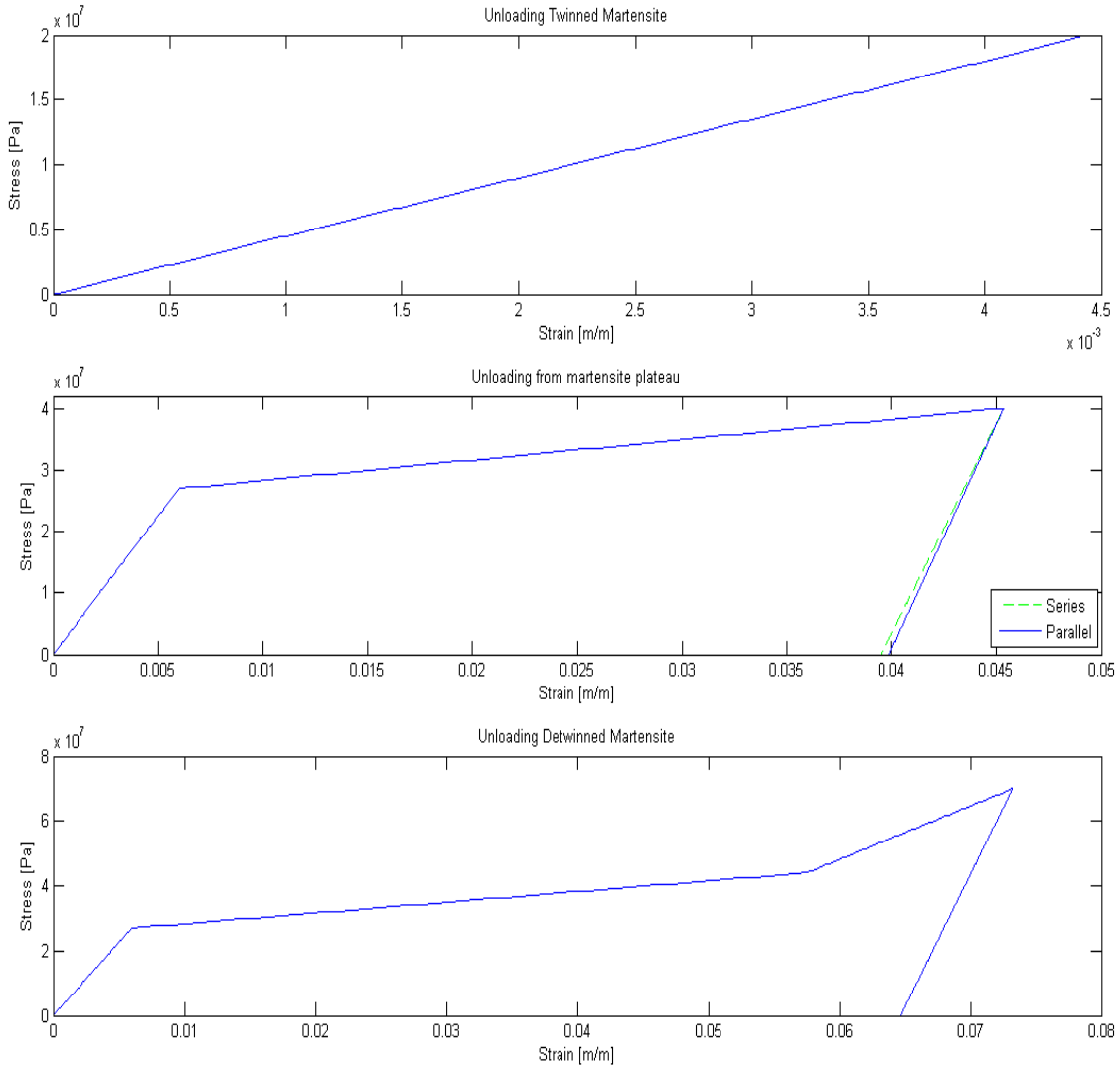

Figure 4.5: Simulation results for unloading in the martensite phase

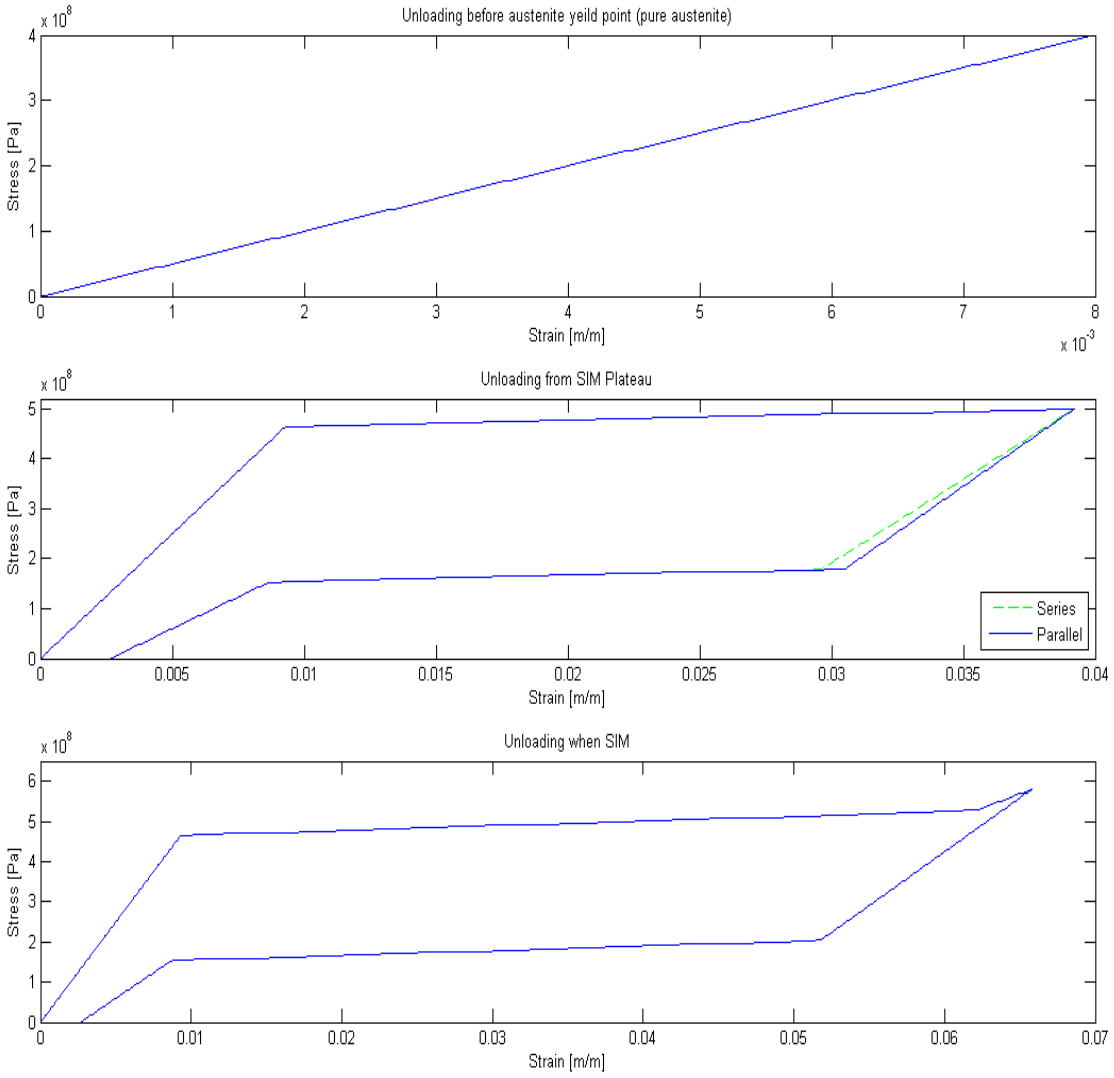

Figure 4.6: Simulation results for unloading in the austenite phase

a residual strain. In response to a stress increase to 35 MPa from 4 seconds to 5 seconds, the material follows the same path as it did when unloading until it rejoins the original loading curve at 28 MPa. Further increase of stress to 35 MPa causes the material to proceed as if it were under regular loading. The progression from 5 seconds to the end proceeds in a similar pattern of loading, unloading and reloading.

Figure 4.8 shows the simulation results for reloading when the SMA is in its austenite phase. The bottom plot shows the results of applying the stress profile shown in the top plot. The first two seconds show elastic behaviour as the material loads to 400 MPa and unloads back to 0 MPa. From 2 seconds to 4 seconds, the SIM starts to form and then the material undergoes unloading from the SIM plateau at 3 seconds from 470 MPa to 380 MPa and then is reloaded back to 500 MPa. Unloading at 5 seconds causes the material to display the pseudoelastic effect. Reloading from 6 seconds to 7 seconds causes the material to follow a reloading modulus that is equal to its most recent unloading modulus path until the material rejoins the loading curve. Further loading causes a strain of approximately 5.6%. When the material is unloaded from this point, the material displays the pseudoelastic effect once again. Reloading from 8 - 9 seconds causes the material to transform completely to SIM. Unloading and reloading from 9 - 11 seconds accounts for the deviation from the SIM curve at about 6.7%.

4.3.4 Pseudoelastic Shifting

Figure 4.9 shows the simulation results for shifting of pseudoelastic loops to higher stresses with increasing temperature. The plots on the left side show the stress and temperature profile used to generate stress-strain curves.

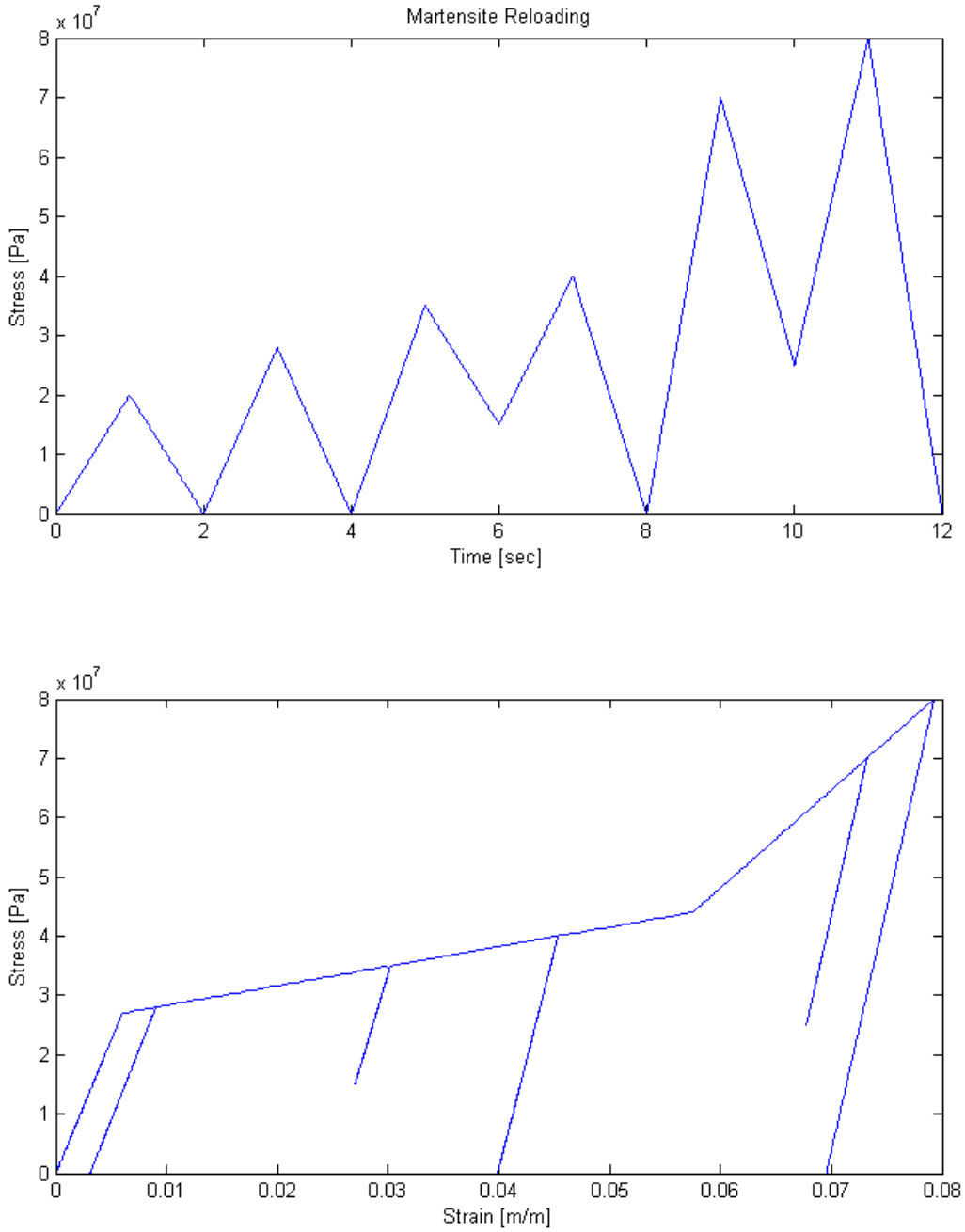

Figure 4.7: Simulation results for reloading in the martensite phase

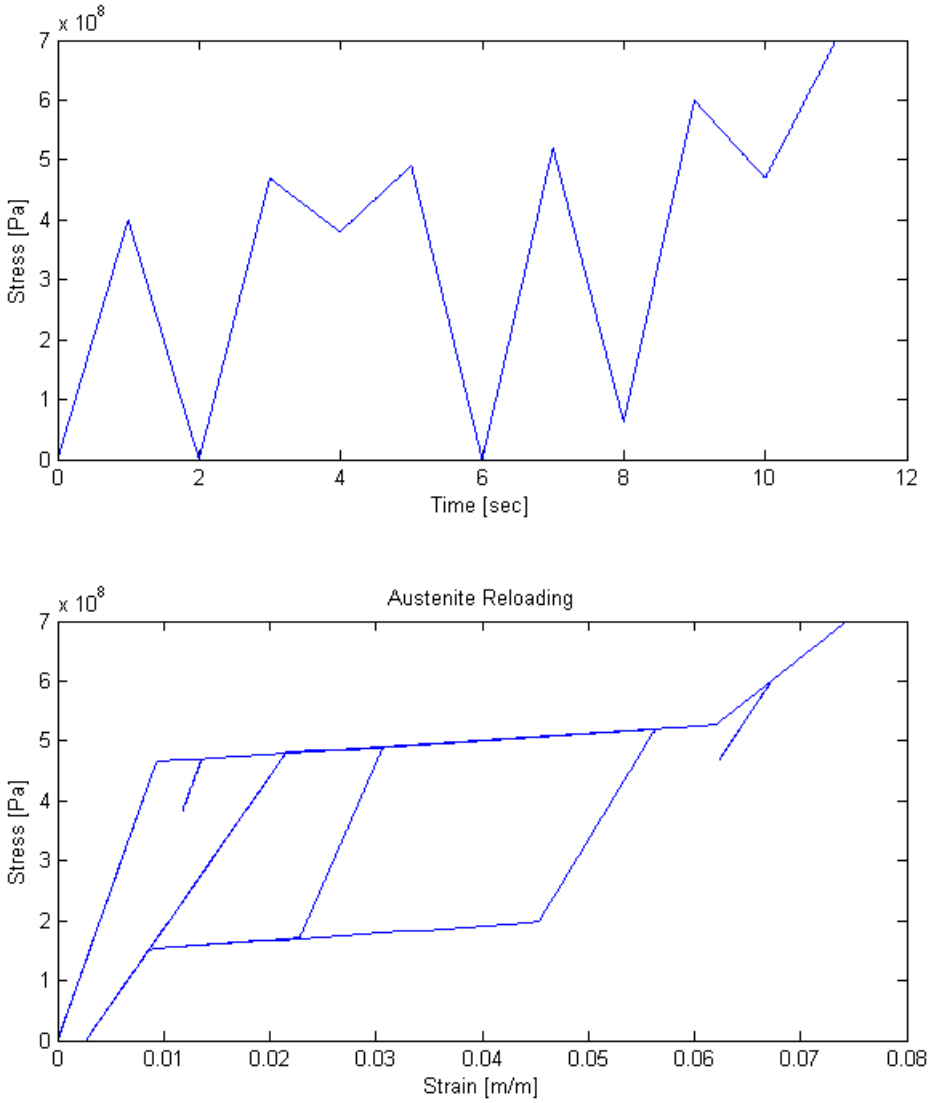

Figure 4.8: Simulation results for reloading in the austenite phase

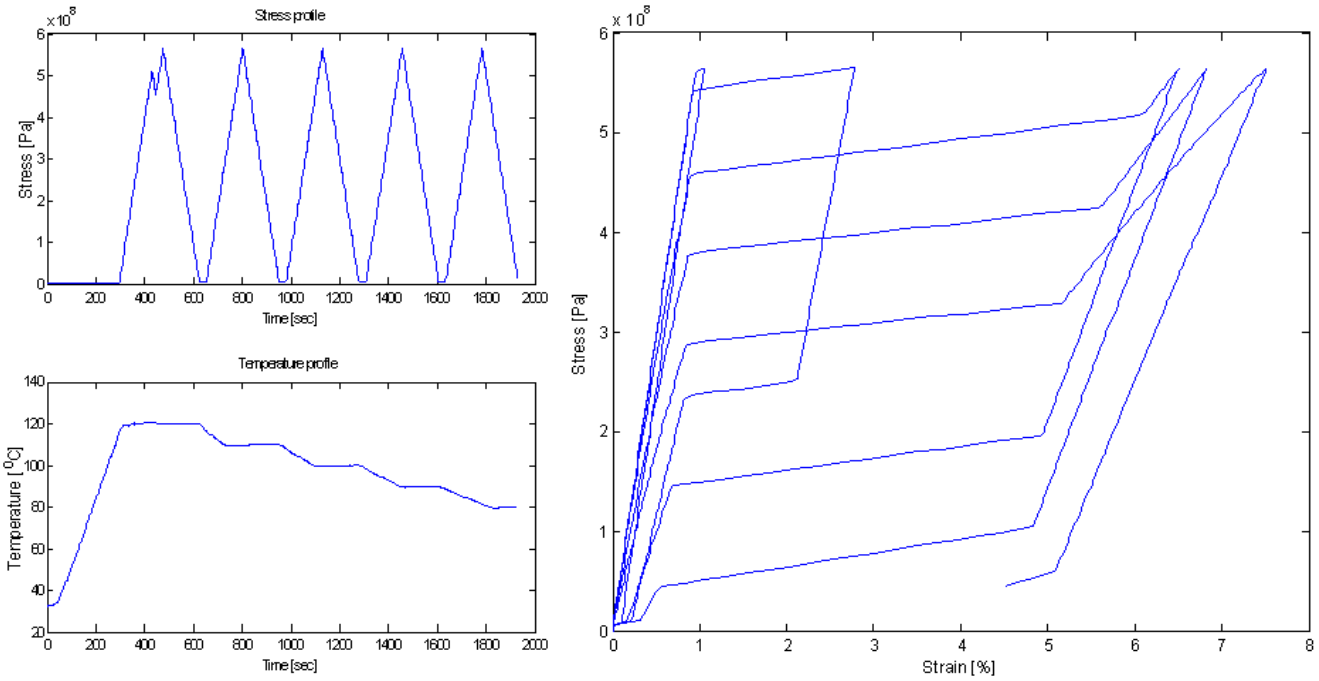

Figure 4.9: Simulation results for pseudoelastic shifting showing stress and temperature profiles

4.3.5 Loading and Unloading for Different Values of R_m

Figures 4.10 and 4.11 show a series of loading and unloading curves as the value of R_m changes from 1 to 0. From the figures is can be observed that the stress-strain behaviour for values of R_m other than 1 and 0, the series and parallel models predict different results. In order to decide between the series and parallel models, it is necessary to perform experimental analysis. The results of this analysis are presented in Chapter 5.

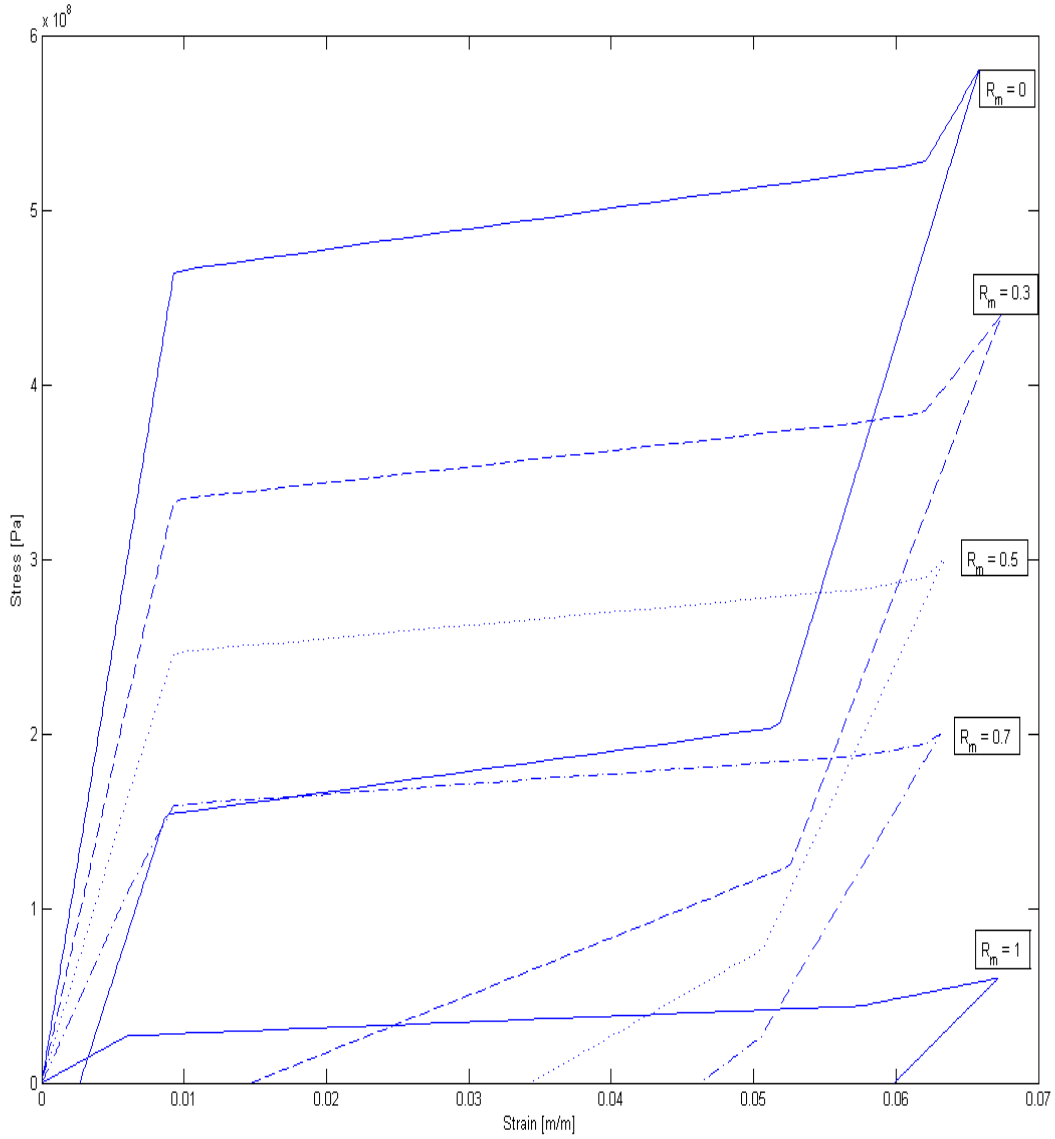

Figure 4.10: Simulation results for loading and unloading using the parallel model

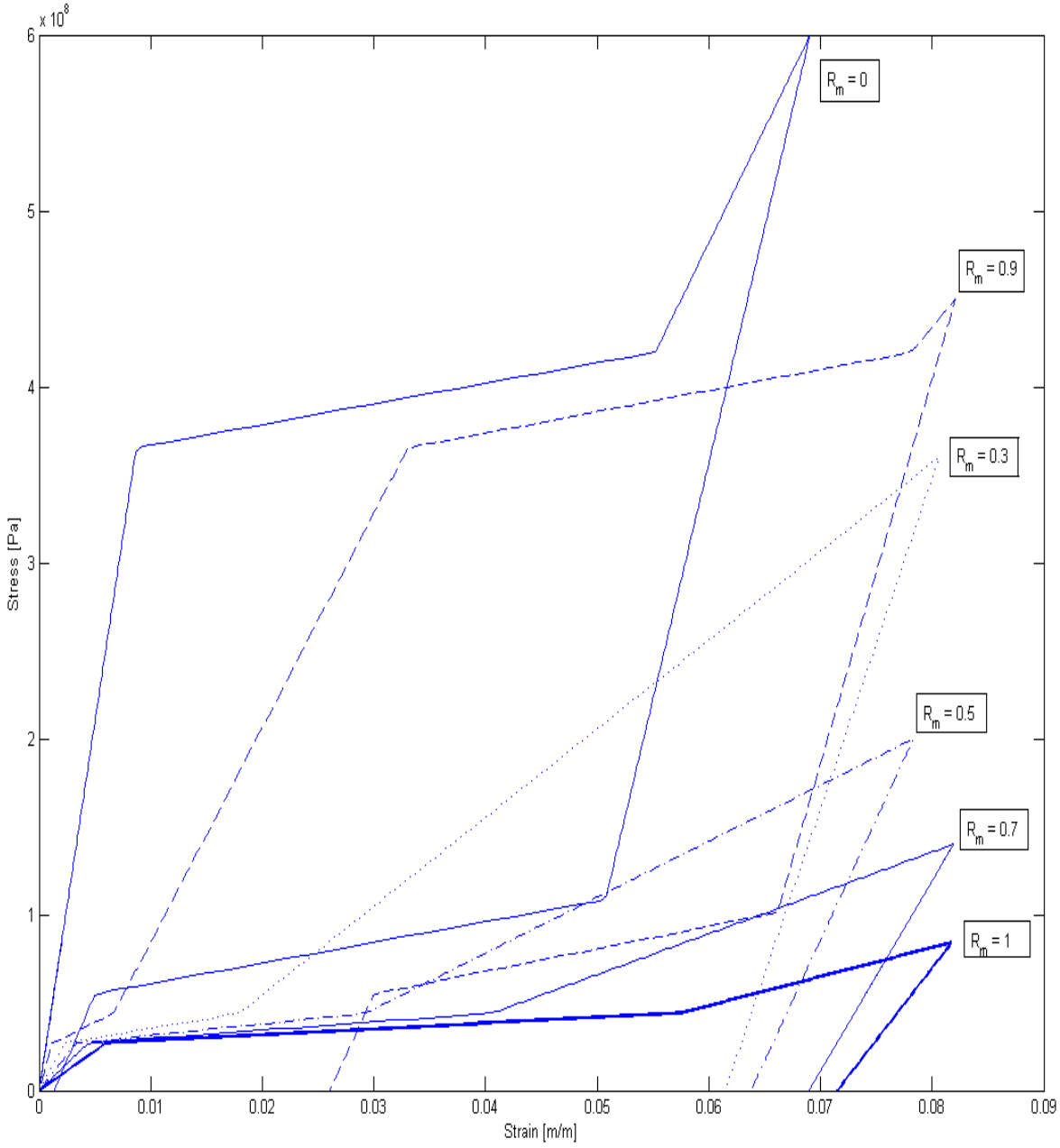

Figure 4.11: Simulation results for loading and unloading using the series model

Chapter 5

Experimental Model Verification

This chapter presents experimental results to validate the SMA wire model presented in Chapter 3. The experimental setup is first presented followed by the experimental procedure. The identification of model parameters is defined. Experimental results on individual components of the model are presented with model predictions.

5.1 Experimental Setup

Figure 5.1 shows a photograph of the experimental set up. For clarity, a top view schematic of the experimental setup is shown in Figure 5.2.

The experiment consists of a voice-coil type actuator (shown on the left) that is capable of applying approximately 7 newtons of force per amp of current, through a movable shaft. The experiment saturates the force actuator current at 4 amps (28 newtons) in order to prevent overheating of the actuator coil. The actuator is coupled to the rest of the experiment via an adjustable shaft

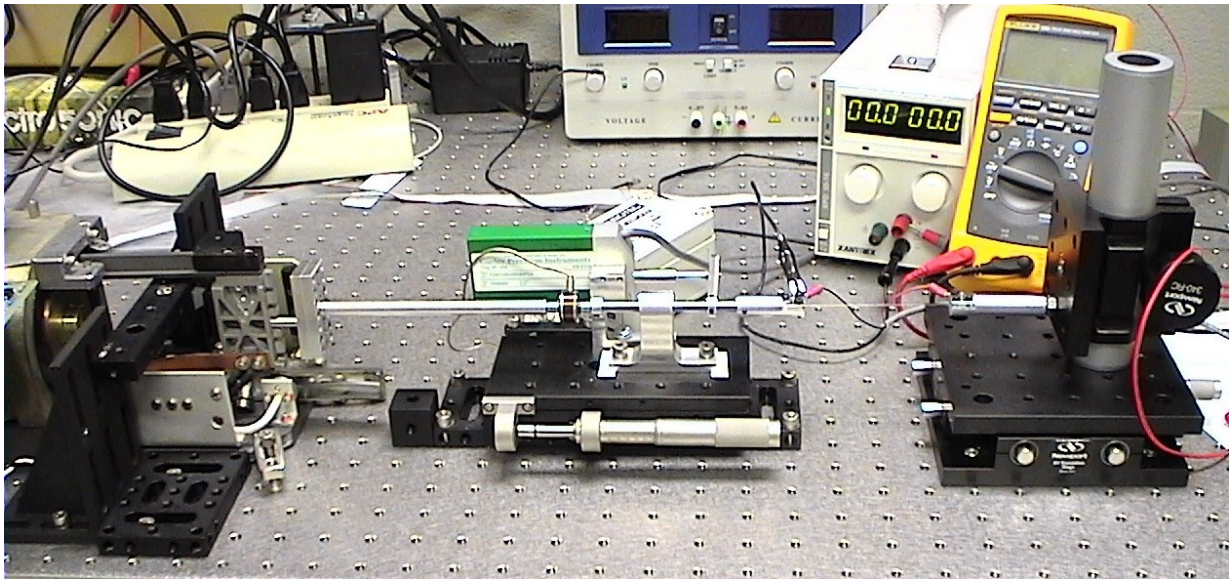

Figure 5.1: Experimental setup

coupler which facilitates the alignment of the forces with the moving components. Continuing to the right in Figure 5.1, an in-line force cell is mounted onto the shaft for force measurement. The output of the strain-gauge-based sensor is amplified by an in-line amplifier. Displacement measurements are made with a linear incremental encoder that is mounted parallel to the direction of motion. The plunger of the encoder is coupled to the rest of the experiment via a rigid connector. One end of the SMA wire is fastened to the movable shaft via a clamp and the other end of the wire is fastened via a similar clamp to an x-y-z translation stage to adjust the tension in the wire as well as ensure that the wire is straight and leveled. The SMA clamp consists of a centering post around which the wire is wound, and three screws to hold the wire in place (Figure 5.3). Wire current is supplied by a power supply, configured as a voltage-controlled current supply with a gain of 1.6A/V, connected in series with an ammeter and the SMA wire. The experiment

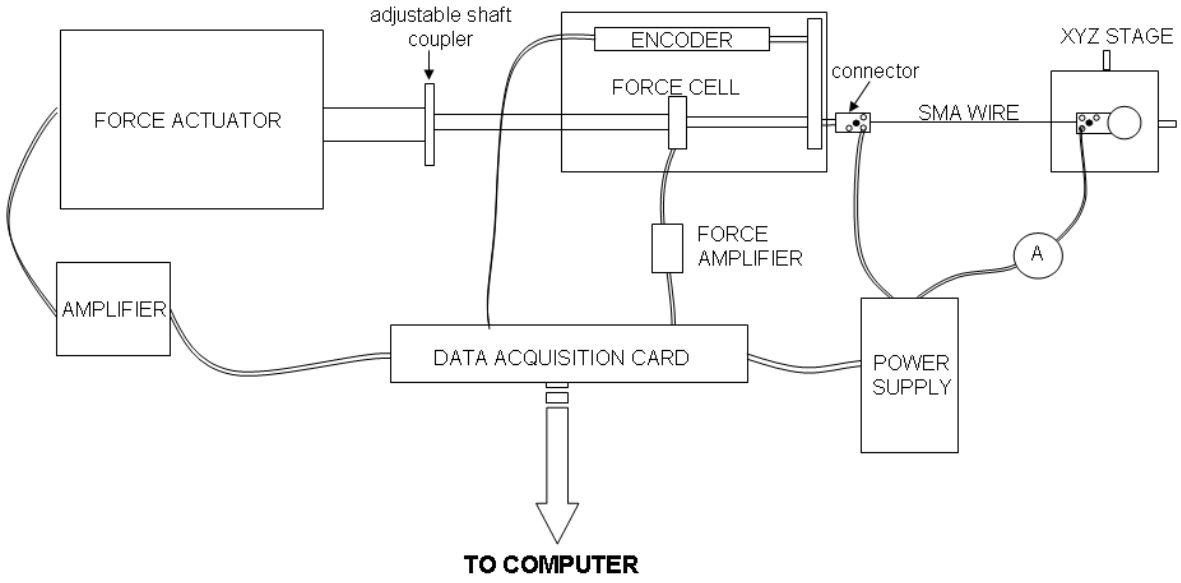

Figure 5.2: Top view of experimental setup

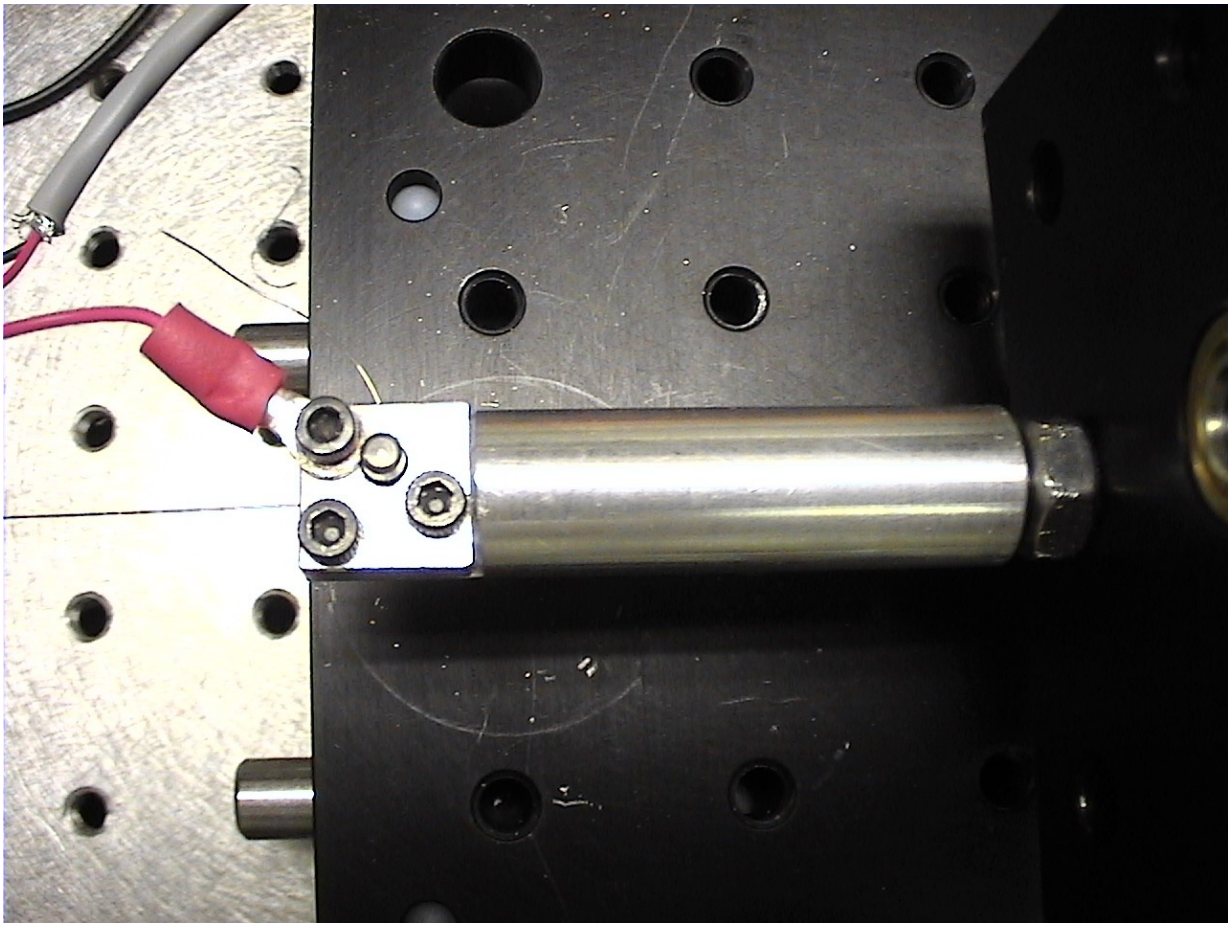

Figure 5.3: Clamp with central post and three screws to fasten SMA wire

is controlled by a computer using MATLAB and SIMULINK's Real Time Windows Target and a data acquisition card. The data acquisiton card outputs signals to control the experiment via its analog output ports and facilitates data collection via the analog and encoder input ports. The experiment uses two analog outputs: the first is connected to the current amplifier that drives the force actuator with a gain of 0.8A/V. The second output is connected to the power supply to control the SMA wire current. The experiment uses two analog inputs. The first is the force cell reading to measure wire force, the second is a current feedback signal from the power supply to read the actual current applied to the SMA wire. The experiment also uses an encoder input to collect the encoder data to measure displacement of the wire. The experiment is performed on an active damping vibration table. The makes and models of the equipment are listed below:

1. Force actuator: salvaged from a Data General 6070 10Mb disk drive (circa 1978).

2. Current amplifier: EG & G Torque Systems model CO502-001.

3. Force cell: Honeywell Sensotec model AL311AR, 10 lb maximum.

4. Force cell amplifier: Sensotec In-Line amplifier model 1023976.

5. Encoder: Gurley Precision Instruments model P20010A05B040PSA, 1 μm resolution.

6. Data acquisition card: Quanser MultiQ3.

7. Ammeter: Fluke 189 multimeter.

8. Power supply: Xantrex power supply model XPD 33-16.

9. x-y-z stage: Newport 443 Series linear stage.

10. Vibration table: Newport RS4000.

Figure 5.4 shows the mechanical and electrical schematics of the wire when in the experimental setup. The mechanical schematic shows that there is a net force applied onto the wire due to the actuator force. As the wire is pulled, the plunger of the encoder is depressed by the same amount as the displacement in the wire. The encoder pulses multiplied by the 1 μm resolution of the encoder gives the displacement in the wire. There are frictional forces present in the setup but they are considered negligible because of the use of an oiled linear bearing on the shaft between the force cell and the wire. The force cell measures the force on the wire, F_{SMA}. The electrical schematic shows the SMA wire in series with an ammeter and a voltage controlled current supply. The voltage, V_{ref} for the supply is from the data acquisition card. The current applied to the wire is $1.6V_{ref}$.

Figure 5.5 shows the software setup used in the experiment. The software portion provides the control and data-collection interface needed to apply arbitrary current and force profiles to the wire and to collect strain data. During the experiment, the strain in the wire is found as follows: the encoder pulses read from the encoder input are multiplied by the encoder's resolution of 1 μm/pulse to find the wire displacement, which is then divided by the length of the wire to find the strain. In order to raise the temperature of the wire for experiments other than room temperature experiments, the power supply acting as a voltage-controlled current-supply is connected to a second analog output port on the data acquisition card. The desired current is multiplied by a factor of 1/1.6 V/A to offset the power supply's internal amplification gain of 1.6 A/V. While current control is performed internally by the Xantrex supply, application of controlled forces requires the design of a closed loop controller. The desiged force control loop (Figure 5.5) is implemented in SIMULINK with a sample time of 0.01 seconds and a fixed-step

Dormand-Prince solver. The wire force is measured via the force cell, which was calibrated with a digital fish scale in order to convert voltage readings in volts, V_{cell} to force readings in newtons. The calibration is $F_{SMA} = 8.84V_{cell} - 0.076$. The error between the reference force and the actual measured force is applied to a PI controller, whose output is multiplied by the force actuator calibration gain of 0.142 A/N. The force actuator was calibrated by experiments conducted to ascertain the relationship between current input to the force actuator in amps and force output from the actuator in newtons as measured by a digital fish scale. The signal is then saturated at a maximum of 4 A to protect the actuator windings. Before being fed into the analog output of the data acquisition card, the signal is multiplied by a factor of 1.25 V/A in order to account for the 0.8 A/V gain of the force actuator current amplifier.

5.2 Experimental Procedure

For all the experiments performed, after being fastened in the wire connectors, the wire is heated under zero load conditions to bring it to its austenite length, and the pre-strain adjusted by manually moving the x-y-z stage in order to zero the force sensor to compensate for low-level offsets in the sensor. If performing a low temperature experiment, the wire is left to cool by convection to ambient air for about 30 seconds thereafter. In order to determine the hysteresis and stress-strain characteristics of the wire and extract parameters for the model of Chapter 3, different current and stress and profiles are applied to the wire via SIMULINK. The current profile is applied directly to the analog output port which controls the power supply.

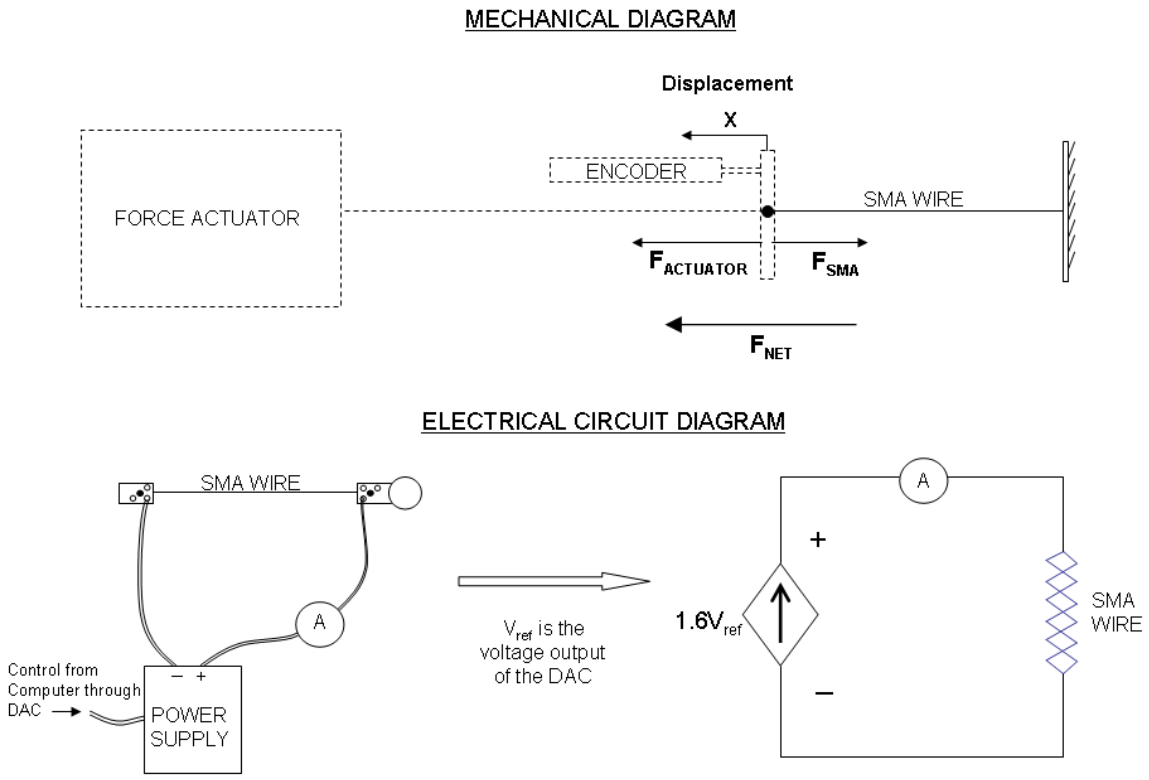

Figure 5.4: Mechanical and electrical diagrams of SMA wire in experimental setup

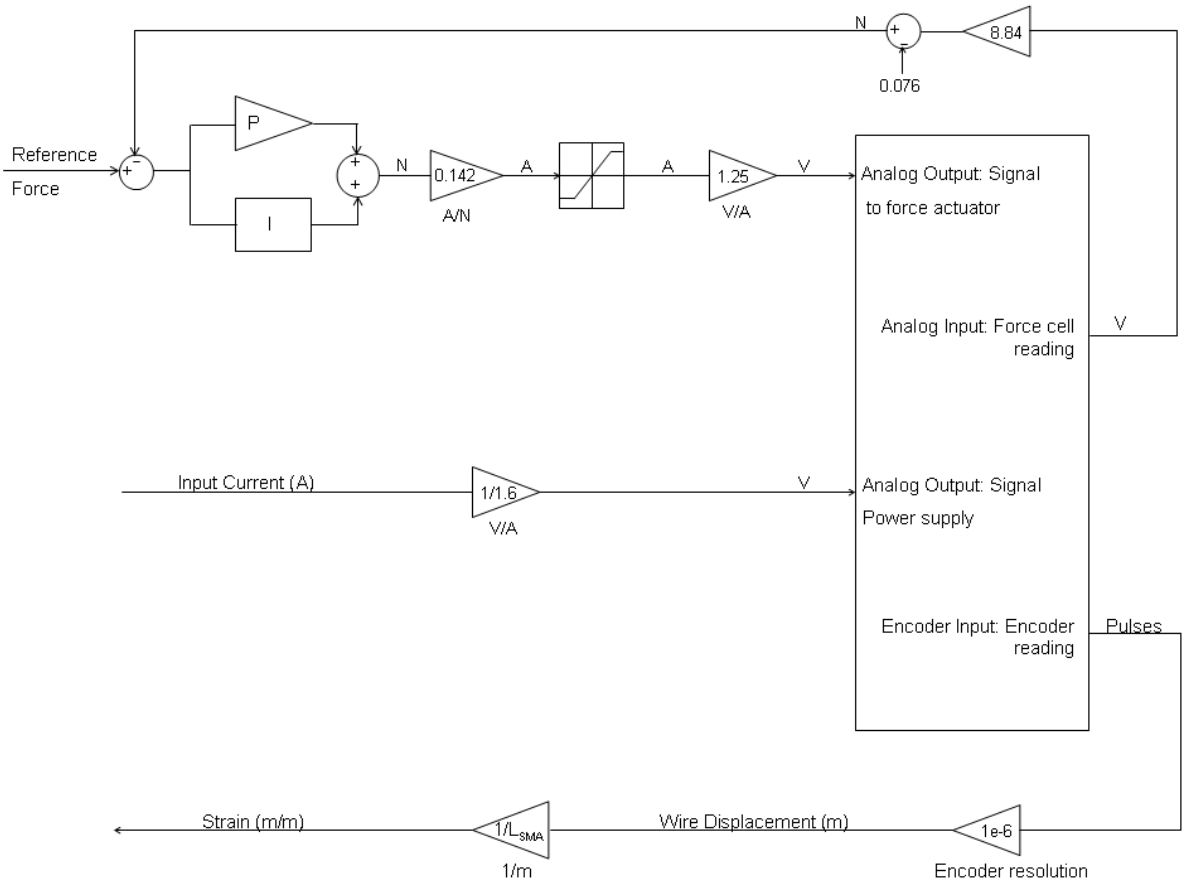

Figure 5.5: Force control loop, current output and strain measurement setup

5.3 Phase Kinetics Model Experimental Model Verification

This section presents the experimental results for tests reproducing the key features of the phase kinetics model. It must be noted that the actual temperature of the wire cannot be measured or controlled with the existing equipment. An attempt is made to validate the model using quasi-static currents and assuming that if the current in the wire is changing slowly, the temperature of the wire is closely follows the predicted temperature found by using the equation for the heating model in Chapter 3, and shown again in 5.1. For the experiments whose results are presented in this section, the current rate of change was ≈1 mA/s (1 A in 15 minutes).

$$CV\rho\frac{dT}{dt} = I(t)^2R - hA(T(t) - T_\infty) \tag{5.1}$$

Since R_m is also not measurable, strain versus predicted temperature curves are converted by scaling the y-axis such that the maximum value is 1, corresponding to low temperature and therefore martensite and the minimum is 0, corresponding to high temperature and therefore austenite.

5.3.1 Major Hysteresis Loop and Parameter Identification

Figure 5.6 shows the current profile applied to the SMA wire and the corresponding predicted temperature. Figure 5.7 shows the resulting hysteresis loops. Since the current is cycled between the minimum and maximum allowable current for two cycles, two major loops are formed[1].

[1]The 'roughness' of the hysteresis curves are due to the fact that the experiment was performed with the wire exposed to the environment, therefore any air currents formed, for example by movement around the experiment, would cause cooling due to convection. Conduction of the experiment in a closed chamber to negate the these convection effects would produce smoother hysteresis curves

Before the start of the temperature profile, the wire is pulled to 244 MPa (12 N) to stretch the wire at room temperature. For the duration of the experiment, the stress is maintained at 244 MPa. In order to extract the k_m values, the slope of the heating and cooling curves (S^H and S^C) is found by applying a linear fit to the R_m data from $R_m = 0.1$ to $R_m = 0.8$ for each of the major loops separately and then taking the average. The values of k_m^H and k_m^C are calculated by using $k_m^H = -4S^H$ and $k_m^C = -4S^C$[2]. The values of the transformation temperatures are found by linear fits to the beginning and end portions of the hysteresis loop. Recall that the hysteresis experiment was carried out at a load of 244 MPa. Therefore, the extracted transformation temperatures are higher than typical transformation temperatures which are extracted from hysteresis loops generated at no load. The extracted parameters as well as the physical characteristics of the wire are in Table 5.1.

Table 5.1: Phase kinetics model experimentally extracted parameters

Parameter	*Value*	*Units*
D_{SMA}	250.00	μm
L_{SMA}	0.08	m
h	70.00	W/m^{2}K
R(linear)	20.00	Ω/m
C	320.00	J/kg^{o}C
ρ	6450.00	kg/m^3
k_m^C	0.072	^{o}C
k_m^H	0.0984	^{o}C
M_s	106.00	^{o}C
M_f	82.00	^{o}C
A_s	110.00	^{o}C
A_f	150.00	^{o}C
T_∞	22.00	^{o}C

The model is run with the extracted parameters and the resulting prediction is superimposed

[2]Equation 3.5.

on the experimental data as shown in in Figure 5.7. The parameters in Table 5.1 are used to make predictions for the rest of the phase kinetics experiments presented in this section.

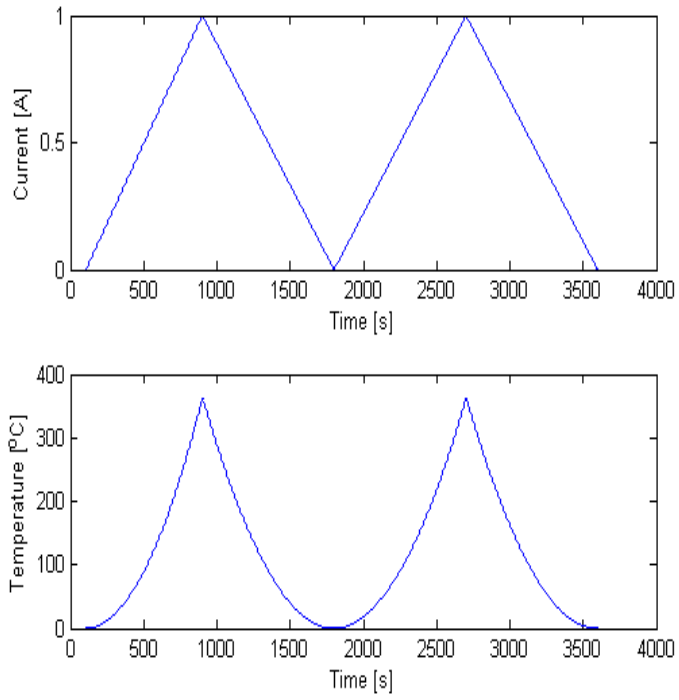

Figure 5.6: Current profile and predicted temperature profile for major hysteresis loops

5.3.2 Minor Hysteresis Loop

Figure 5.8 shows the current profile applied to the SMA wire and the corresponding predicted temperature. Figure 5.9 shows the resulting hysteresis loops. At the start of the experiment, the wire is pulled to 244 MPa in 100 seconds to stretch the wire at room temperature and maintained at that force until the end of the experiment. The current is started at 105 seconds and until 3600

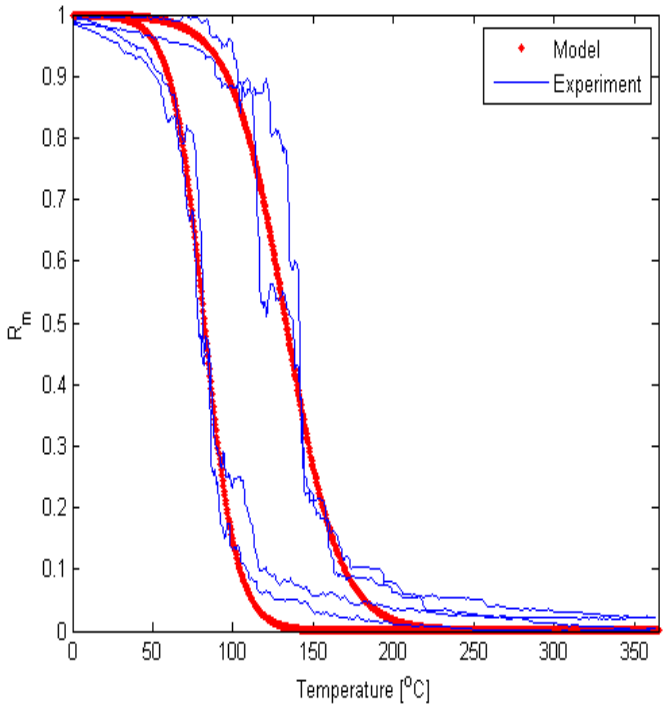

Figure 5.7: Experimental hysteresis major loops with model prediction

seconds, the two triangular currents varying from 0 amps to 1 amp are applied to the wire. This causes the formation of two major loops as in the previous experiment. From 3600 seconds to 4770 seconds, the current is increased to 0.65 amps and then reduced to 0 amps. This causes the formation of a minor loop cooling branch. The current is then increased to 1 amp. Similar to the major loop experiment, the 'roughness' of the hysteresis curves are due to the effects of convection cooling on the exposed wire. Figure 5.10 shows the model prediction is superimposed on the experimental data. The experimental data is plotted only from 3600 seconds to the end of the profile in order to clearly see the minor loop and the model fit.

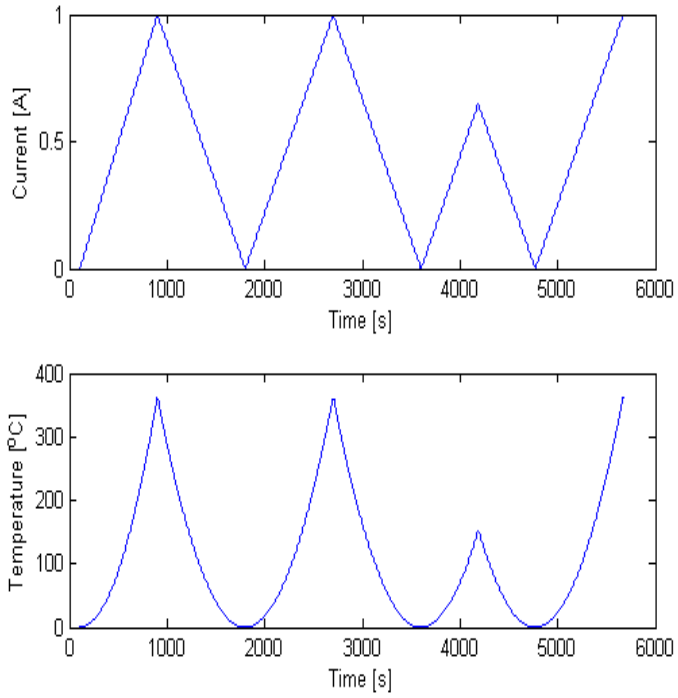

Figure 5.8: Current profile and predicted temperature profile for major and minor hysteresis loops

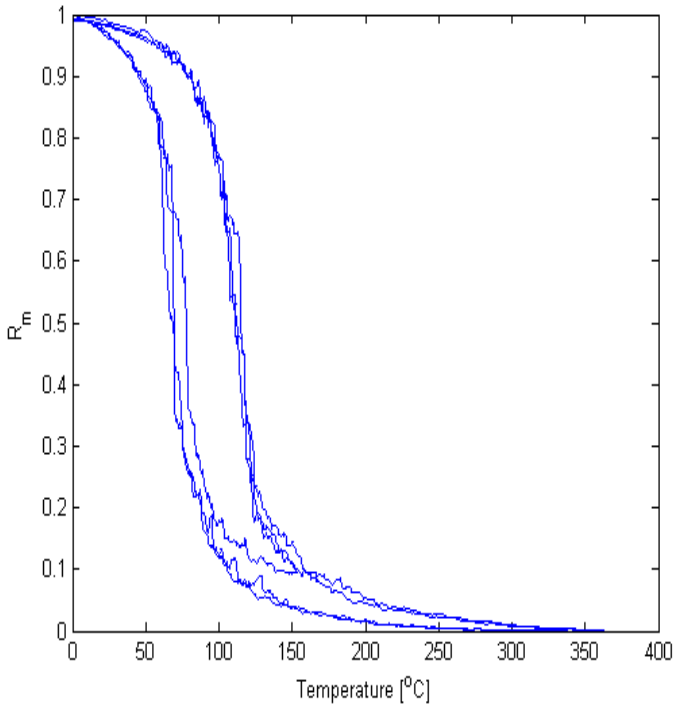

Figure 5.9: Experimental hysteresis major and minor loops

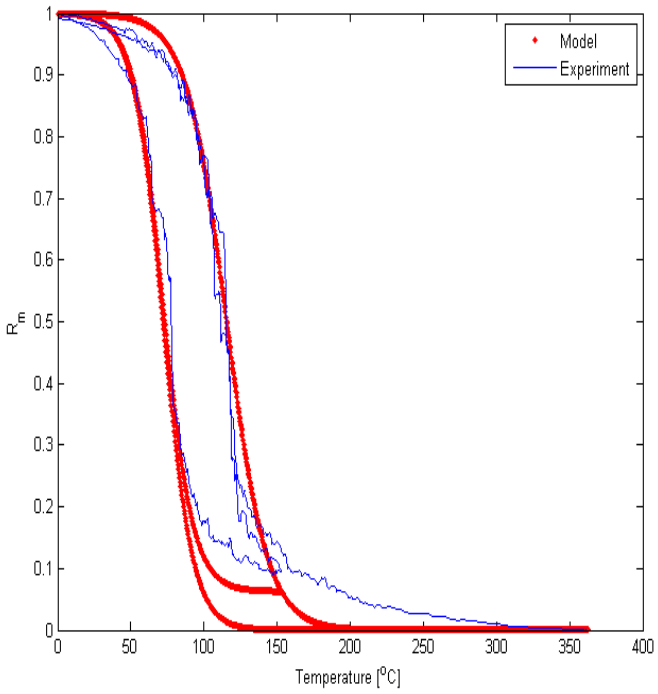

Figure 5.10: Experimental hysteresis minor loop with model prediction

Figure 5.12 shows the result of applying the current profile of Figure 5.11. At the start of the experiment, the wire is pulled to 244 MPa in 100 seconds to stretch the wire at room temperature and maintained at that force until the end of the experiment. This experiment is similar to the previous experiment where the first 3600 seconds generate two major loops, the current is increased to 0.65 amps from 3600 seconds to 4700 seconds. In this experiment, instead of letting the current fall from 0.65 amps to 0 amps, the current is reduced to 0.45 amps until 4770 seconds causing a minor loop cooling branch. Subsequently, the current is increased from 0.45 amps to 1 amp causing a minor loop heating branch that originates on the previous minor loop cooling branch. The model prediction from 3600 seconds is superimposed on the experimental data. The model shows qualitatively validation of the hysteresis behaviour through the ability to reproduce experimentally observed trends in the SMA wire when exposed to a current profile as shown in Figure 5.11.

5.3.3 Hysteresis Loop Shifting

Figure 5.13 shows hysteresis loops generated at different constant stress. During each of the constant stress durations, the current was ramped up from 0 to 1 amp in 12 minutes and then ramped down to 0 amps in 12 minutes. The hysteresis curves shown in the figure are plotted as strain versus the predicted temperature for four different constant stresses. The figure shows that the increase in the stress causes a shifting of the hysteresis loops, as expected. The strain reading on the y-axis for this figure is not converted to R_m in order to show that the hysteresis loops when plotted as strain versus temperature shift vertically as well as horizontally with increasing stress. The vertical shift is due to the fact that higher stresses cause higher strains in the wire.

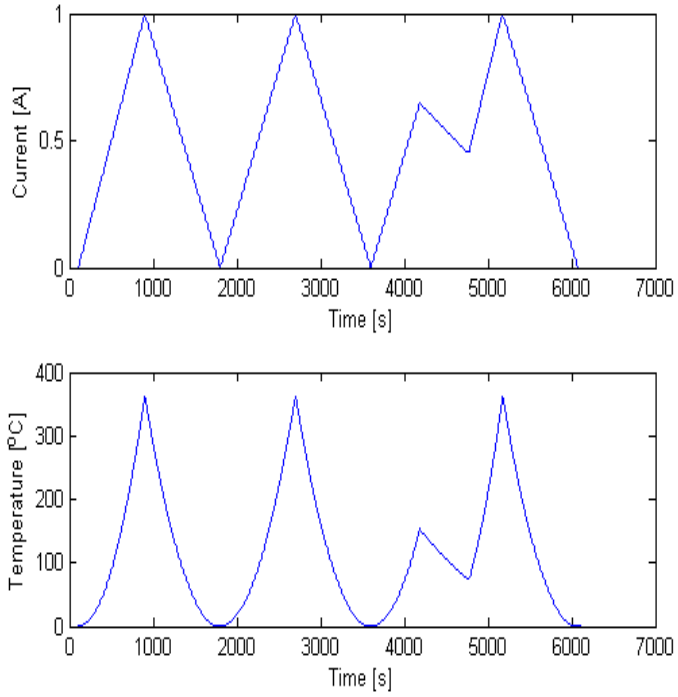

Figure 5.11: Current profile and predicted temperature profile minor hysteresis loops

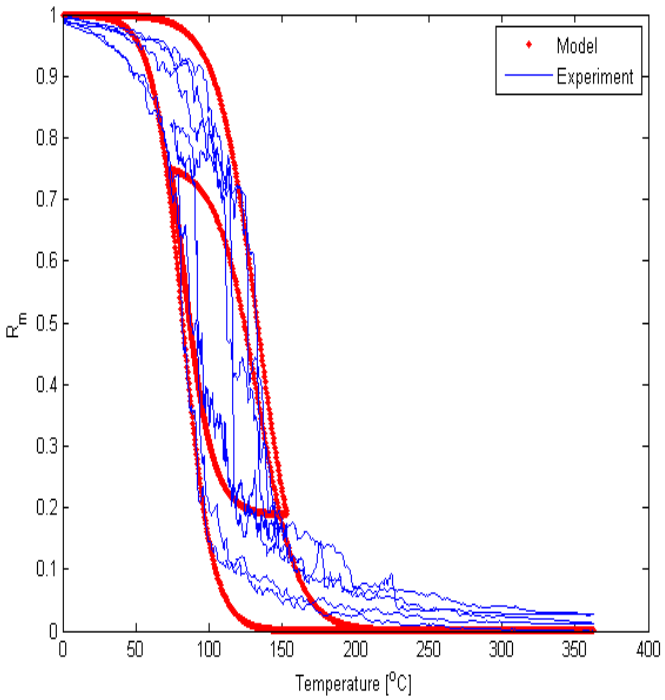

Figure 5.12: Experimental hysteresis curves showing minor loop branching with model prediction

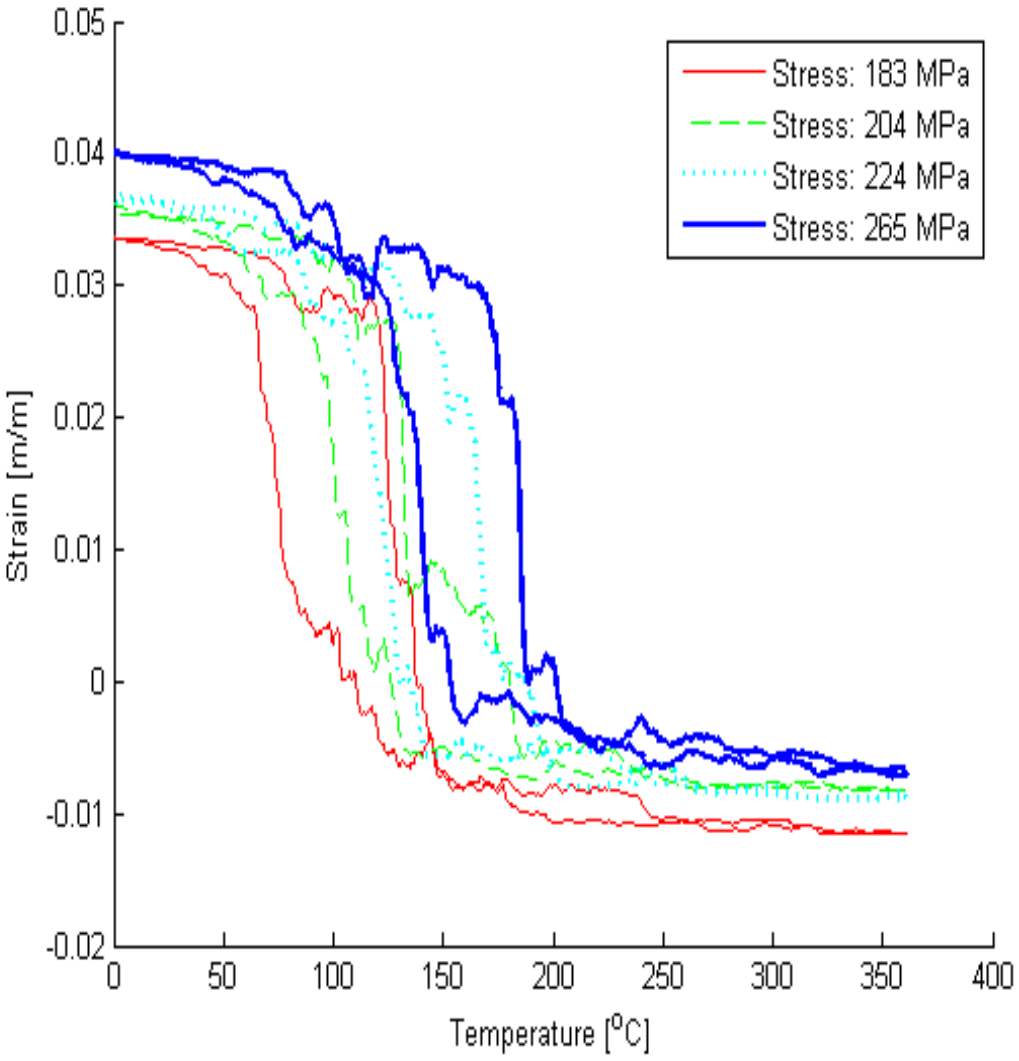

Figure 5.13: Experimental hysteresis major loops showing shifting with increasing stress

5.4 Mechanical Model Experimental Model Verification

This section presents the experimental results for tests reproducing the mechanical behaviour of SMAs. The experimental curves are overlaid with model predictions for comparison. The results are presented in stress strain curves for low temperature and high temperature.

5.4.1 Model Parameters

The parameters for the mechanical model were extracted as follows: At room temperature, the SMA wire was pulled past its second yield point and then completely unloaded. From this test, the moduli, E_m, E_t, E_d, E_d^{un} were extracted by fitting linear lines through each of the regions on the two stress-strain curves. The yield stresses and strains, $\sigma_{y1}^M, \sigma_{y2}^M, \epsilon_{y1}^M, \epsilon_{y2}^M$ were found by extracting the co-ordinates of the intersection points of the linear fits to the stress-strain data. A similar pull is conducted at high temperature. When the material was unloaded, it displayed pseudoelasticity. The moduli and yield stresses and strains $E_a, E_{am}, E_{sim}, E_{sim}^{un}, \sigma^{Ms}, \sigma^{Mf}, \sigma^{Af}, \epsilon^{Ms}, \epsilon^{Mf}, \epsilon^{As}$ and ϵ^{Af} are also found by linear fits. The extracted parameters (Table 5.2) are used to predict the behaviour for subsequent experiments.

5.4.2 Low Temperature Stress Strain Curves

Figure 5.14 shows two stress profiles applied to a 250 μm diameter SMA wire (Dynalloy Flexinol 250 90^{o}C) at room temperature[1]. The first is a ramp stress from 0-122 MPa in 250 seconds to show the loading behaviour of martensite. The second stress profile is a sequence of triangular

[1]The stress profiles shown are found by dividing the measured force from the force cell by the cross-sectional area of the SMA wire

Table 5.2: Mechanical model experimentally extracted parameters

Parameter	*Value*	*Units*
E_a	40.00	GPa
E_{am}	1.17	GPa
E_{sim}	14.00	GPa
E_{sim}^{un}	24.00	GPa
E_m	4.50	GPa
E_t	330.52	MPa
E_d	1.65	GPa
E_{un}^{d}	8.20	GPa
E_{un}^{a}	15.00	GPa
σ_{y1}^{M}	27.00	MPa
σ_{y2}^{M}	44.02	MPa
σ^{Ms}	336.00	MPa
σ^{Mf}	383.50	MPa
σ^{Af}	45.60	MPa
ϵ_{y1}^{M}	0.006	m/m
ϵ_{y2}^{M}	0.0575	m/m
ϵ^{Ms}	0.0084	m/m
ϵ^{Mf}	0.049	m/m
ϵ^{Af}	0.005	m/m
ϵ^{As}	0.0441	m/m

stresses to show loading, unloading and reloading behaviour in martensite. Figure 5.15 shows the resulting stress-strain profile with the model prediction superimposed on the experimental data. The figure shows good agreement between the model and simulation.

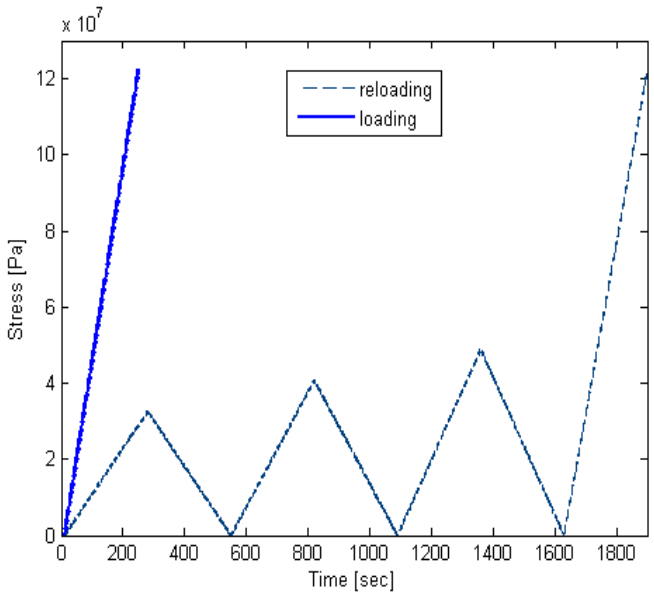

Figure 5.14: Stress profile for martensite (room temperature) pull

Figure 5.16 shows the stress profile applied to a 250 μm diameter SMA wire with 0.1 Amps of current in the wire. Due to the current in the wire, the martensite fraction is less than 1 ($R_m < 1$). The stress profile is a number of loading and unloading ramps to show loading, unloading and reloading behaviour in the SMA. Figure 5.17 shows the superimposed plots for reloading at $R_m = 1$ and $R_m < 1$ for comparison. From the figure, it is observed that as temperature increases, the yield stresses increase but the yield strains remain relatively the same. This behaviour is predicted by the parallel model (see Figure 3.24). One can conclude, therefore, that the parallel mechanical model is better suited to model SMA behaviour than the series

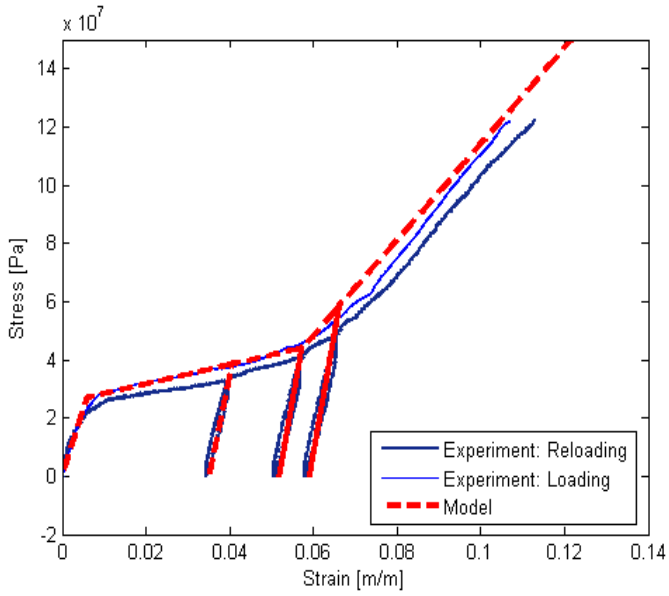

Figure 5.15: Experimental stress - strain profile for martensite (room temperature) pull

mechanical model.

Figure 5.18 shows the resulting stress-strain profile with the model prediction superimposed on the experimental data. The figure shows good agreement between the model and simulation. The R_m value used for the model is $R_m = 0.97$. The value of R_m was calculated using 5.2, an equation relating slopes of austenite and martensite for the parallel model, with E_m as the slope of the twinned martensite portion of Figure 5.15, E_a as the slope of a high temperature pull, presented in the next section and E as the slope of the curve in Figure 5.18 before the first yield point. The calculated value was $R_m = 0.97$. Following the calculation, different values of R_m around the calculated value were simulated. The best fit was found to remain at $R_m = 0.97$.

$$E = R_m E_m + (1 - R_m) E_a \tag{5.2}$$

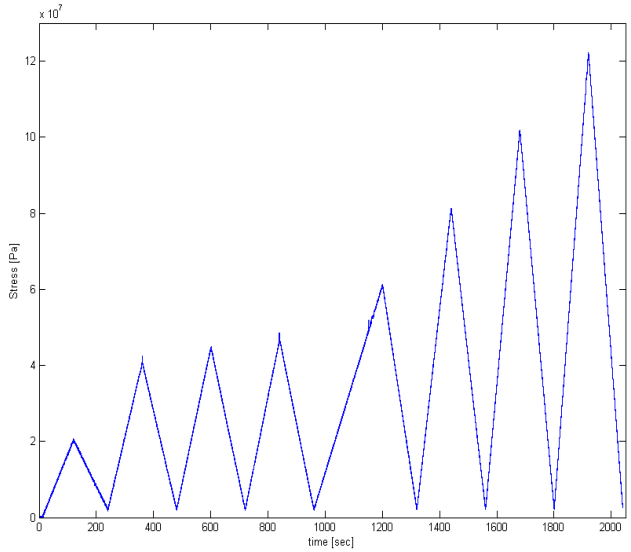

Figure 5.16: Stress profile for low temperature pull for $R_m < 1$

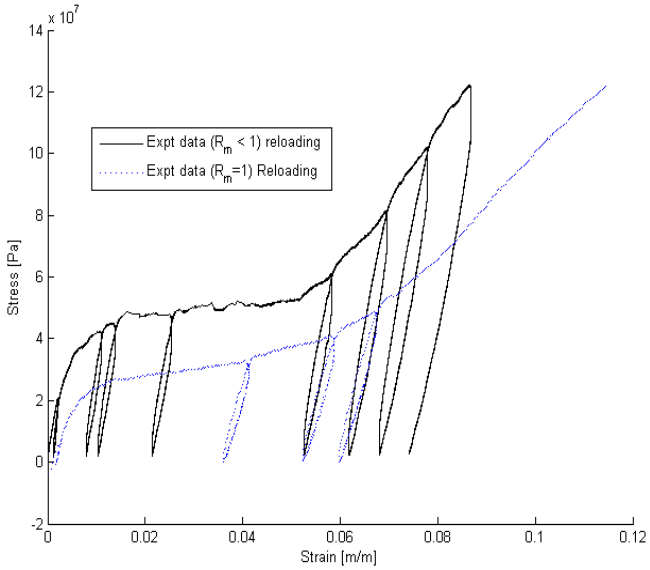

Figure 5.17: Experimental stress - strain profile for $R_m = 1$ and $R_m < 1$

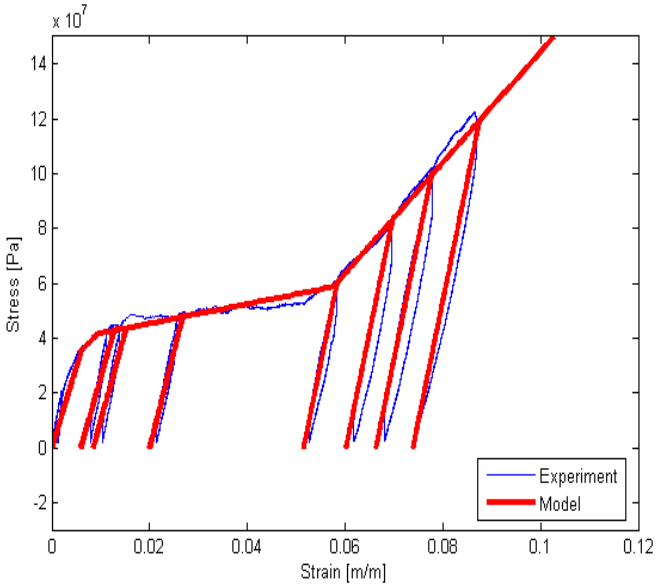

Figure 5.18: Experimental stress - strain profile for low temperature pull $R_m < 1$

5.4.3 High Temperature Stress Strain Curves

The high temperature experiments were carried out at the Canadian Space Agency facility in Montreal, Quebec. The equipment consisted of a Texas Instrument Dynamic Mechanical Analyser (DMA), model 2980. The SMA wire is fastened between two grippers and enclosed in a cavity in order to control the temperature of the wire. The wire is heated by heating the air in the cavity. The experiments were performed without an active cooling mechanism[1]. Figure 5.19 presents the results of a experiment run at 120^{o}C showing a partial pseudoelastic loop with the model prediction overlaid on the experimental result.

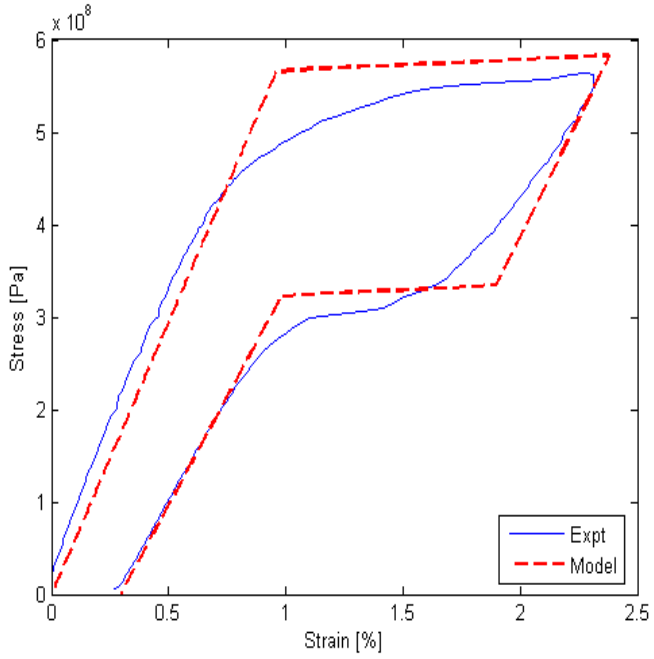

Figure 5.19: Experimental stress - strain profile for SMA wire at 120^oC

[1]Although the DMA is capable of active cooling, at the time of the experiments, the liquid nitrogen required for active cooling was not available.

Figure 5.21 shows the results of applying the stress and temperature profile shown in Figure 5.20 to the SMA wire. In order to focus on the main parts of the profile, the first 350 seconds have been omitted from the figures. During this first 350 seconds, the temperature of the SMA is brought to 120^oC from room temperature. The stress is at 0 Pa. Once the wire is at 120 ^{o}C, it is loaded to a stress of 565MPa and unloaded to 10MPa[2]. Thereafter the wire is allowed to cool to 110 ^{o}C and subsequently to 100 ^{o}C, 90 ^{o}C and 80 ^{o}C. At each of these temperatures, the stress profile for loading and unloading is repeated.

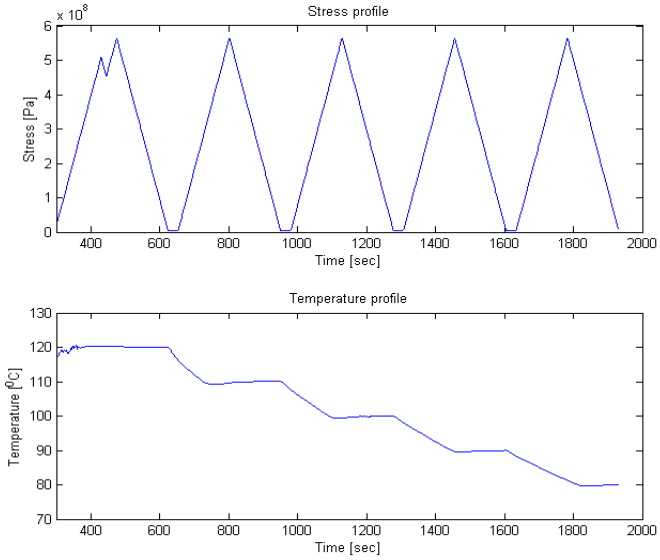

Figure 5.20: Stress and temperature profile for high temperature SMA pulls

Figure 5.21 also shows the the model predictions overlaid on the resultant experimental stress-strain profile. The figure shows the presence of distinct pseudoelastic loops from 90^oC and a vertical shift of these loops with higher temperatures. For clarity, Figures 5.22 - 5.25 show

[2]The wire is not unloaded to 0 MPa to ensure that the wire is always in tension.

the stress - strain curves at each temperature with the model overlaid over each curve. Note that $R_m = 0$ for $T \geq 90^oC$. The model prediction for the stress-strain curve at 80^oC is generated for $R_m = 0.16$. The value of R_m was first calculated using (5.2), yielding $R_m = 0.20$ and then different values of R_m around the calculated value were simulated. The best fit was found at $R_m = 0.16$. This value is reasonable due to the fact that the temperature is less than the austenite finish temperature, A_f, therefore there could be some martensite in the material, further shown by the presence of significant residual strain.

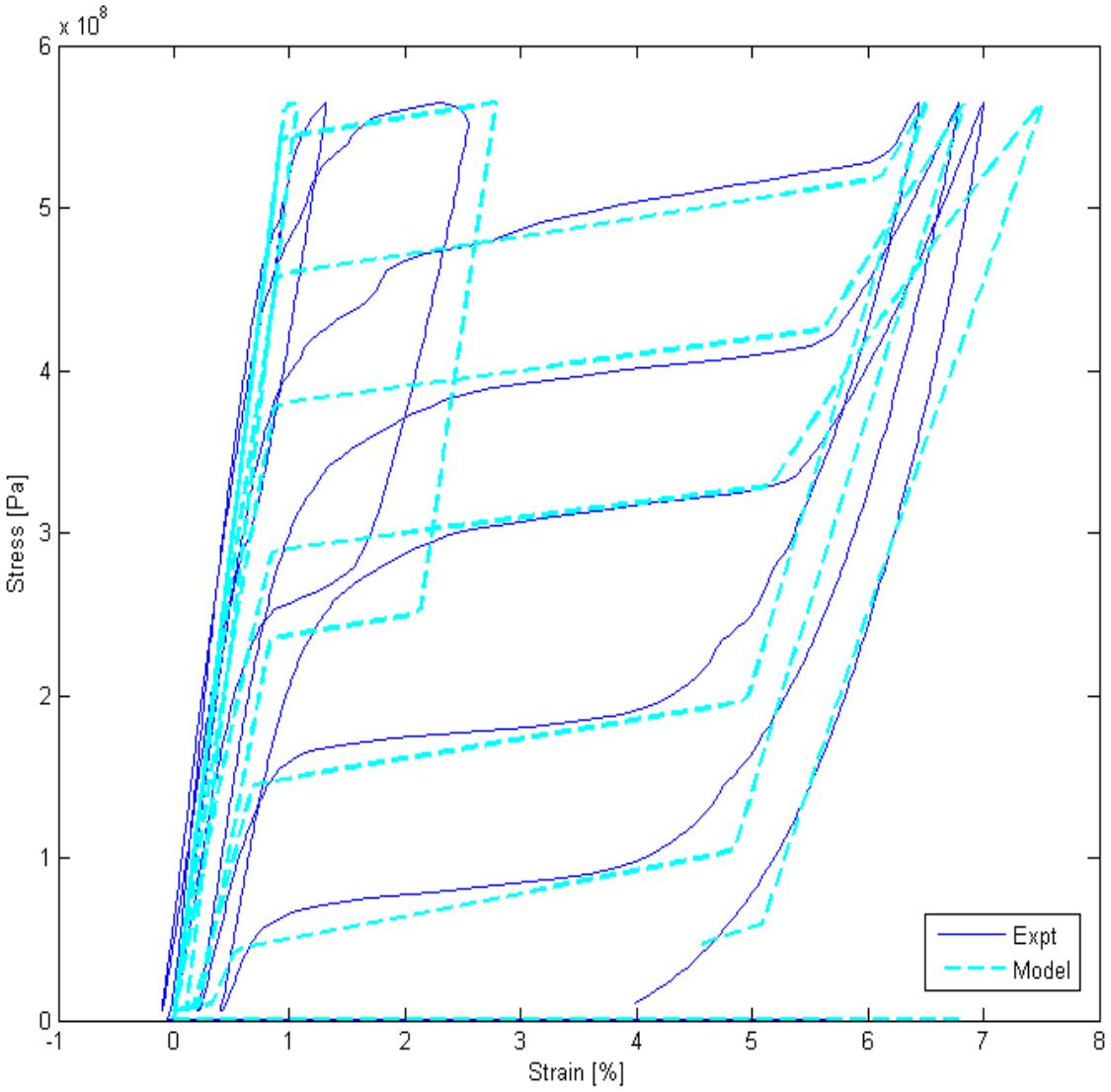

Figure 5.21: Experimental stress - strain profile for high temperature SMA pulls

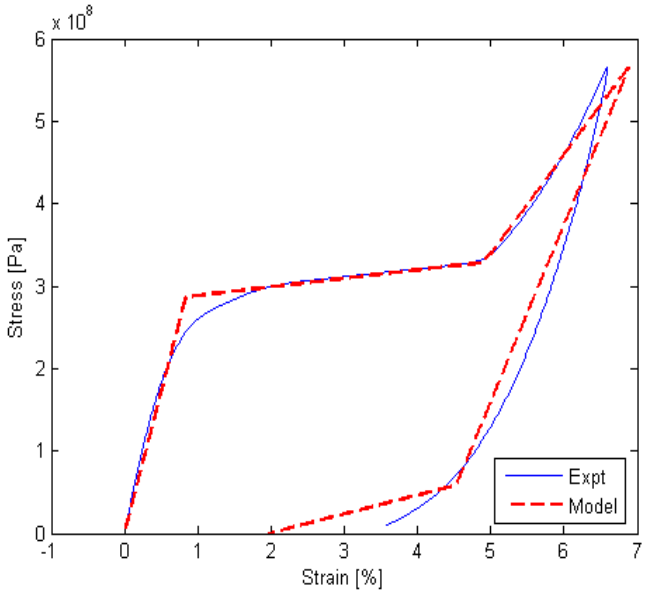

Figure 5.22: Experimental stress - strain profile for SMA wire at 80^oC

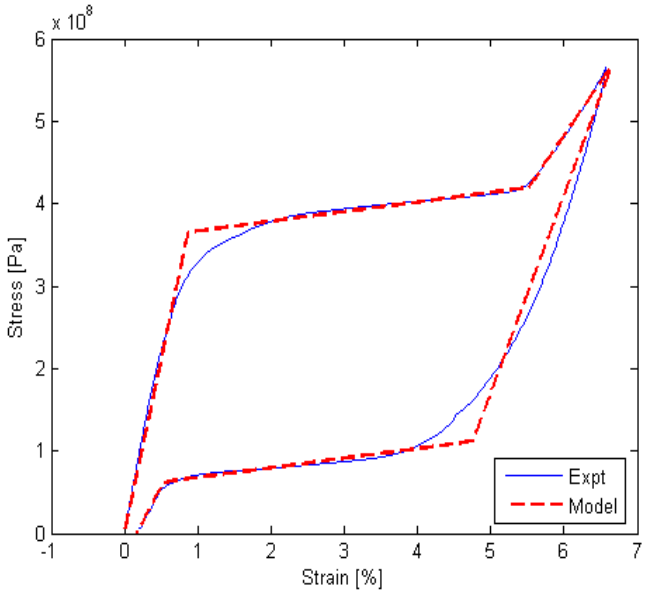

Figure 5.23: Experimental stress - strain profile for SMA wire at 90^oC

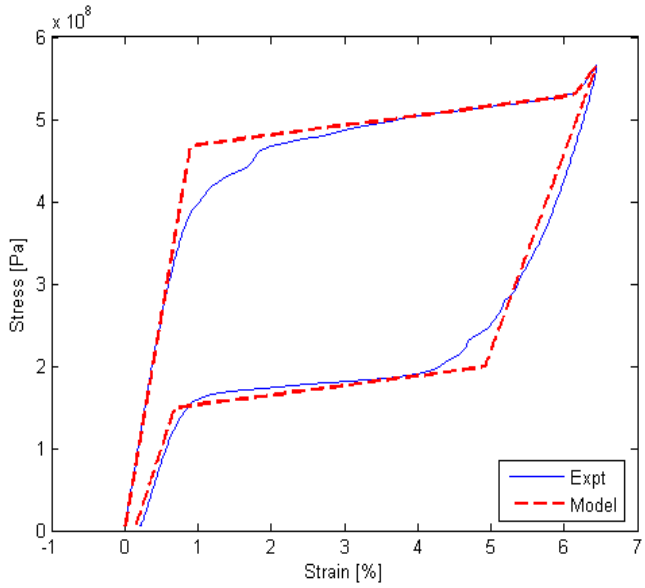

Figure 5.24: Experimental stress - strain profile for SMA wire at 100^oC

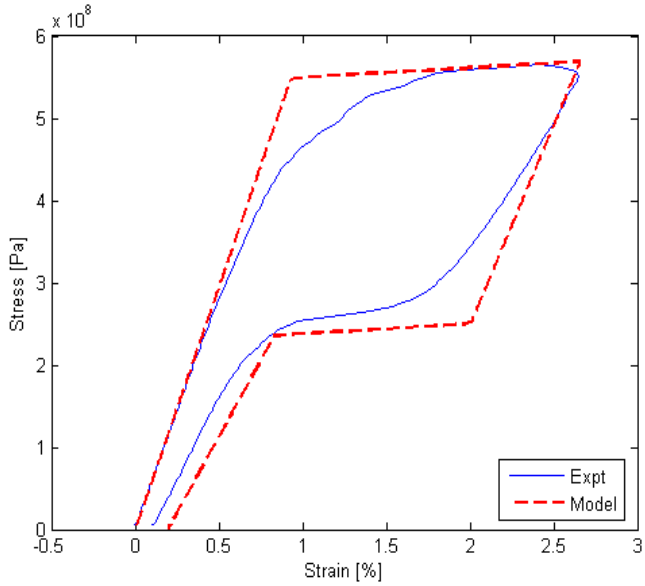

Figure 5.25: Experimental stress - strain profile for SMA wire at 110^oC

Chapter 6

Control

This section discusses the control of shape memory alloy acutators, beginning with a review of past work. Section 6.1 presents the literature review on past SMA control techniques. Sections 6.3, 6.4, 6.5 present the control of constant force, spring-biased and variable force SMA actuators. Section 6.6 presents the results for disturbance rejection tests on a constant force SMA actuator.

6.1 Literature Review on Control of SMA

Researchers in the past decade have proposed a number of different types of control techniques for shape memory alloys. Various control architectures have been proposed such as feedforward, feedback, adaptive control to mention a few.

Classical proprtional-integral-derivative (PID) feedback control is used by Reynaerts and Van Brussel [31], Ikuta [11], Hashimoto et al. [32] who use a PD controller and Troisfontaine et al.

[33] who use a PI controller with temperature feedback. Arai et al. [34] use feedback linearization to control the SMA.

Feedback control has been augmented with feedforward compensation in an attempt to linearize the SMA hysteresis. Briggs and Ostrowski [35] use a PID controller in feedback with model predictive hysteresis control in feedforward. They assume that strain is not observable and use temperature measurements for feedback. In a later experiment, they employ a state estimator to estimate strain. Ma, Song and Lee [29] perform position control using a feedforward neural network and a PID controller using resistance measurements in feedback. Song et al. [36] use a neural network in feedforward with a sliding mode based robust controller in feedback. Majima et al. [37] employ a PID feedback controller in conjunction with a feedforward controller that is dervied by modifying the Preisach hysteresis model.

Adaptive control has been used by Webb and Lagoudas [38] who use an adaptive hysteresis model for on-line identification as well as closed loop compensation. Tao et al. [39] study an adaptive control system using a piecewise linear model as an inverse hysteresis model. Dickinson and Wen [40] use adaptive feedback control to control an SMA attached to a beam.

Variable structure control was studied by Grant et al. [41] and Elahinia et al. [42] on an SMA actuated manipulator. Eren et al. [43] use a hybrid combination of gain scheduling, B-spline approximation, variable structure control and integral control.

Other feedback control schemes have been employed by Song and Quinn [44] who use a sliding mode based robust controller, Hughes and Wen [45] who use the Preisach model to control an SMA wire actuator, Kumagai et al. [46] who use a neural fuzzy controller, Gorbet and Morris [47] who use dissipativity, Selden et al. [48] who use a control scheme known as segmented control where the SMA is divided into segments and control is achieved by exploiting hysteresis

in each segment.

It must be noted though that since the most common SMA actuators are spring biased, most of the control efforts presented above deal specifically with spring-biased actuators. The model and inverse that were presented in Chapter 3 account for arbitrary stress profiles applied to the SMA actuator and therefore expands the scope of modelling and control of SMA actuators to beyond just the SMA spring biased actuator. The controller structure investigated in this chapter is model-inverse feedforward with PI feedback to compensate for model inaccuracies. Performance in sinusoidal and step strain tracking will be compared with classic PI control. The controller is also able to provide robustness through the rejection of stress disturbances, and this is investigated as well.

6.2 Tracking Control

The strain tracking control of SMA actuators is performed with two control architectures, namely proportional integral (PI) control and feedforward model-inverse control with PI feedback control where the inverse controller is used to partially linearize the system. Two types of actuators are considered: an SMA wire actuator with constant force (Section 6.3) and a spring-biased SMA actuator, which applies time varying forces to the SMA wire (Section 6.4). The reference strains considered are:

1. Step signal

2. Sinusoidal signal with period of 5 seconds

The wire used in these experiments is Flexinol 250HT, a 250 μm diameter wire with a nominal austenite transformation temperature, A_f, around 90 C.

The ability of the controller to track a sinusoidal signal with a varying stress profile different from the spring stress is shown in Section 6.5. Additonally, Section 6.6 presents the results for disturbance rejection in an SMA wire actuator with constant force.

The parameters[1] used for the inverse controller are as shown in Tables 6.1 and Tables 6.2

Table 6.1: Phase kinetics model parameters for inverse controller

Parameter	*Value*	*Units*
D_{SMA}	250.00	μm
L_{SMA}	0.08	m
k_m^C	0.15	^{o}C
k_m^H	0.12	^{o}C
h	70.00	W/m^{2}K
R(linear)	20.00	Ω/m
C	320.00	J/kg^{o}C
ρ	6450.00	kg/m^3
c_m	0.015	^{o}C/MPa
M_s	72.00	^{o}C
M_f	62.00	^{o}C
A_s	88.00	^{o}C
A_f	98.00	^{o}C
T_∞	22.00	^{o}C

Due to the lack of equipment to directly measure temperature, the values of k_m^H and k_m^C extracted from quasi-static current experiments in Section 5.3 are not accurate values. The values of k_m^H and k_m^C used in the control experiments in this section were found by taking the average of the slopes of the hysteresis curves for Flexinol in [1], [35] and [49].

[1]The variables in the table where extracted from stress-strain pulls for the wire that was used for the control experiments. This wire was taken from a different spool than the wire that was used for the experimental verification and hence low temperature parameters were reidentified

Table 6.2: Mechanical model parameters for inverse controller

Parameter	*Value*	*Units*
E_a	40.00	GPa
E_{am}	1.17	GPa
E_{sim}	14.00	GPa
E_{sim}^{un}	24.00	GPa
E_m	20.40	GPa
E_t	904.00	MPa
E_d	15.7	GPa
E_{un}^{d}	39.20	GPa
E_{un}^{a}	15.00	GPa
σ_{y1}^{M}	67.32	MPa
σ_{y2}^{M}	106.64	MPa
σ^{Ms}	336.00	MPa
σ^{Mf}	383.5	MPa
σ^{Af}	45.6	MPa
ϵ_{y1}^{M}	0.033	m/m
ϵ_{y2}^{M}	0.0468	m/m
ϵ^{Ms}	0.0084	m/m
ϵ^{Mf}	0.0490	m/m
ϵ^{Af}	0.0050	m/m
ϵ^{As}	0.0441	m/m

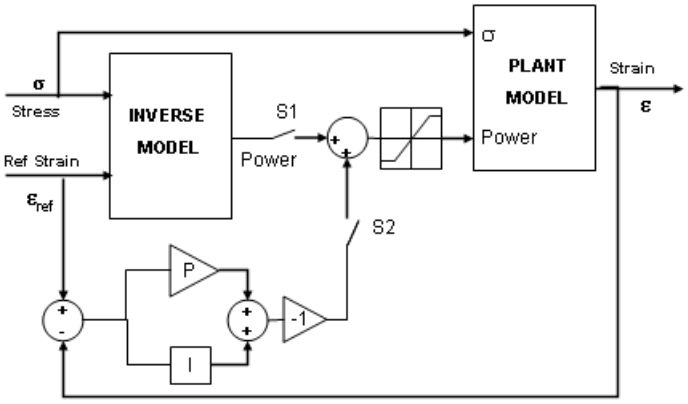

Figure 6.1: Control architecture, constant force actuator

6.3 Control of Constant Force SMA Actuator

In these experiments, the force actuator applies constant forces on the SMA wire. The force is controlled using a force control feedback loop with proportional integral (PI) control. The wire is first pulled when it is in its martensite state (low temperature) to its initial strain. This is the 'loading period'. The control action begins after a 150 second loading period after which the force applied to the wire is constant for the duration of the experiment. The general control configuration is shown in Figure 6.1. The two architectures compared are PI feedback (S1 open, S2 closed), and feedforward plus PI feedback (both switches closed).

Power input to the plant is limited by the 1 A maximum safe current for the wire, and negative power values are saturated at zero to model the absence of an active cooling mechanism. PI gains were tuned manually to avoid oscillation due to the nonlinear power saturation, and to provide good tracking error. Note that with the control error defined in the traditional way ($\epsilon_{ref} - \epsilon$) and

positive PI gains K_p and K_i, the output of the PI controller needs to be inverted. This is because a negative error means $\epsilon > \epsilon_{ref}$, requiring the wire to be heated and hence a positive power input to the plant. With the additional model inverse feedforward, PI feedback is combined with the feedforward signal from the model-inverse controller, and applied to the plant. The PI output is used to adjust the power signal predicted by the plant inverse, in order to compensate for plant perturbations and disturbances. In the following sections, controllers are qualitatively compared based on plots of the strain error, and quantitatively using the 2-norm of the strain error.

6.3.1 Step Reference Strain

The first reference used for ϵ_{ref} is a step signal from strain of $\approx$2.5% to 0.5%[1]. The control signal is started at 160 seconds and the reference step change occurs at 170 seconds. Figures 6.2 and 6.3 show the step tracking results using the PI controller and the PI + Inverse controller respectively. Figure 6.4 shows the error of using the two control architectures.

The gains for the controllers were tuned so as to give the best performance. The proportional gains for both controllers was 850. The integral gain was 0.01 for the PI + Inverse controller and 15 for the PI controller. The presence of integral control typically reduces steady state error. In this case, however, higher integral gains were found to be detrimental to the tracking control because of the accumulation of error from the cooling segments. The error in the PI + Inverse controller at 166 seconds can be explained as follows: when control action starts at 160 seconds, the actual strain is higher than the desired strain due to modelling errors. The controller action causes heating in the wire to reduce the strain but overcompensates for the error

[1]Note that the inital loading sequence is a ramp from 0N to 9 N. The strain at the end of this loading is taken as the inital strain and the final strain is set to 0.5%. The inital strain will not be exactly the same for every experiment because loading is carried out using force control and not strain control.

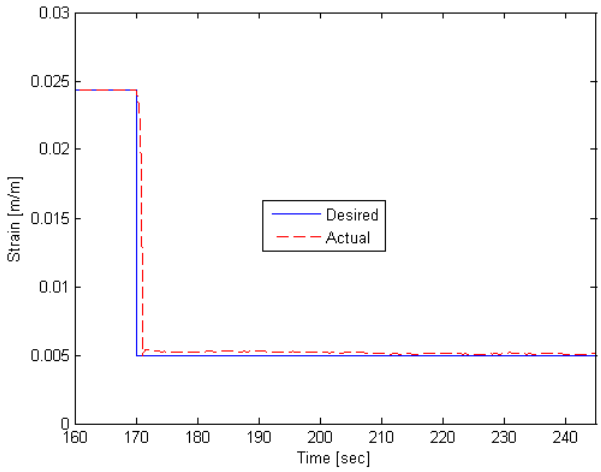

Figure 6.2: Step tracking with PI control, constant force actuator

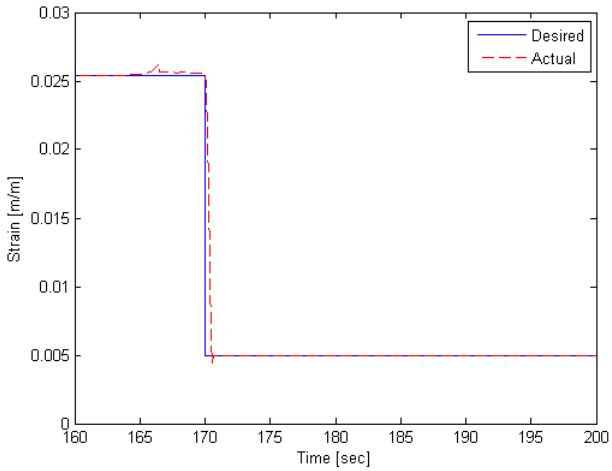

Figure 6.3: Step tracking with PI + Inverse control, constant force actuator

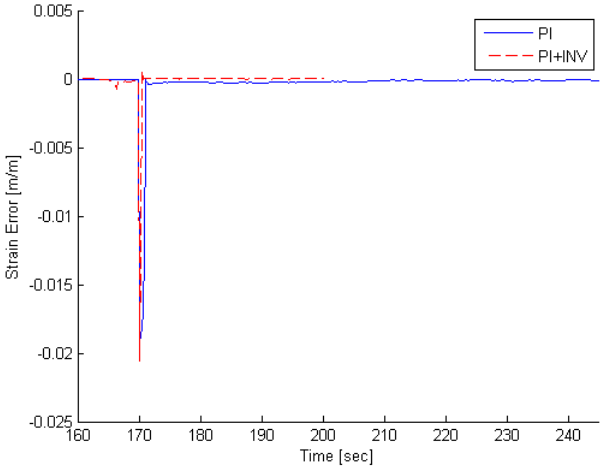

Figure 6.4: Tracking error for step reference, PI and PI + Inverse control, constant force actuator

requiring the wire to cool at 162 seconds. The system lacks an active cooling mechanism leading to the absence of accurate control of cooling. Therefore, convection cooling causes the strain in the wire to increase beyond the desired strain. Controller action then, once again, causes the input of current into the wire to bring it to the desired strain. Figure 6.5 shows the spikes in the measured force on the wire during the step transition. These spikes contribute to the undershoot in both controllers after the step transition. The error in the step tracking at 170 - 171 seconds is due to saturation of the current signal output by the controller. Figure 6.6 shows the control currents using the PI and PI+Inverse controllers. The maximum allowable current for 250 μm wire is 1 A. In order to protect the wire, therefore, the control current signal passes through a saturation block before it is applied to the SMA. Note that saturation is not accounted for in the inverse model, therefore generation of control signals that drive the system to saturation are not desirable. However, saturation results when faced with high frequency reference signals like

pure step changes. The ability of the controller to recover from saturation contributes to the degree of robustness of the system. Figure 6.6 shows that the control current for the PI controller exceeds the 1 A saturation limit for a much longer duration than the current for the PI + Inverse controller. The maximum current specified by the PI + Inverse controller (23 A) is higher than the PI controller (16.7 A). However, the presence of a saturation block ties the tracking error not to the magnitude of the current but the duration of the saturation. Note that during saturation, the integrator is accumulating error, contributing to an undershoot after the step change at 171 seconds and for the PI controller, a higher settling time. In application, it may be desirable to limit the slew rate of any step references to avoid saturation. Despite the saturation, the PI + Inverse controller appears to recover faster and provide better tracking and faster time response, suggesting it is more robust to the effects of saturation. The error norms for the control schemes are given below:

Table 6.3: Error two-norms, step reference, constant force SMA actuator

controller	*error norm*
PI	0.1702
PI+Inv	0.0983

The PI + Inverse controller provides better tracking performance for a step signal than using traditional PI control. The use of the PI + Inverse controller greatly reduces the steady-state error present when using only PI control.

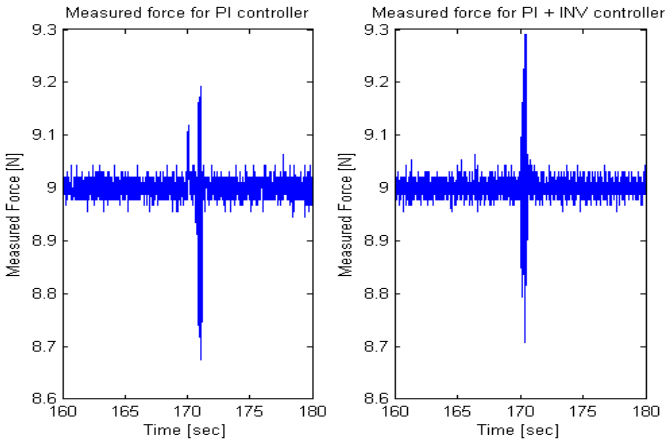

Figure 6.5: Measured force for step reference, PI and PI + Inverse control, constant force actuator

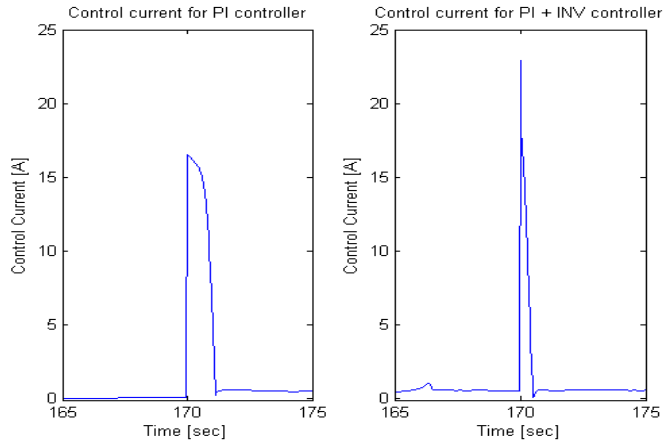

Figure 6.6: Current for step reference, PI and PI + Inverse control, constant force actuator

6.3.2 Sinusoidal Reference Strain

The next references used for ϵ_{ref} are sinusoidal signals with a period of 5 seconds. The wire is pulled from no load and 0% strain to a load of 10 N corresponding to a strain of approximately 3% in 50 seconds[1]. The strain at 50 seconds, ϵ_o and a minumum strain value of 1.3% is used to generate the reference sinusoidal strain, R_{str} in (6.1)

$$\begin{aligned} A &= \frac{\epsilon_o}{2} - \frac{0.013}{2} \\ \omega &= \frac{2\pi}{prd} \\ \phi &= \frac{\pi}{2} - \frac{2}{pi} \times \frac{50}{5} \\ B &= \frac{\epsilon_o}{2} + \frac{0.013}{2} \\ R_{str} &= A(sin(\omega\, t + \phi)) + B \end{aligned} \tag{6.1}$$

where A [m/m] is the amplitude of the wave, ω [rad/s] is the frequency, ϕ [rad] is the phase to ensure that the sinusoidal signal starts at 50 seconds, and B [m/m] is the bias to ensure that the strain at 50 seconds is the maximum strain. Figures 6.7 and 6.8 show the sinusoidal tracking results using the two controllers with the tracking error in Figure 6.9[1].

The largest errors are observed in the cooling portions of the reference strain. This is ex-

[1]The strain at the end of this loading is not exactly the same for every experiment because loading is carried out using force control and not strain control.

[1]Note that the time scales of the figures is from 51 seconds to 102 seconds since control action begins at 51 seconds.

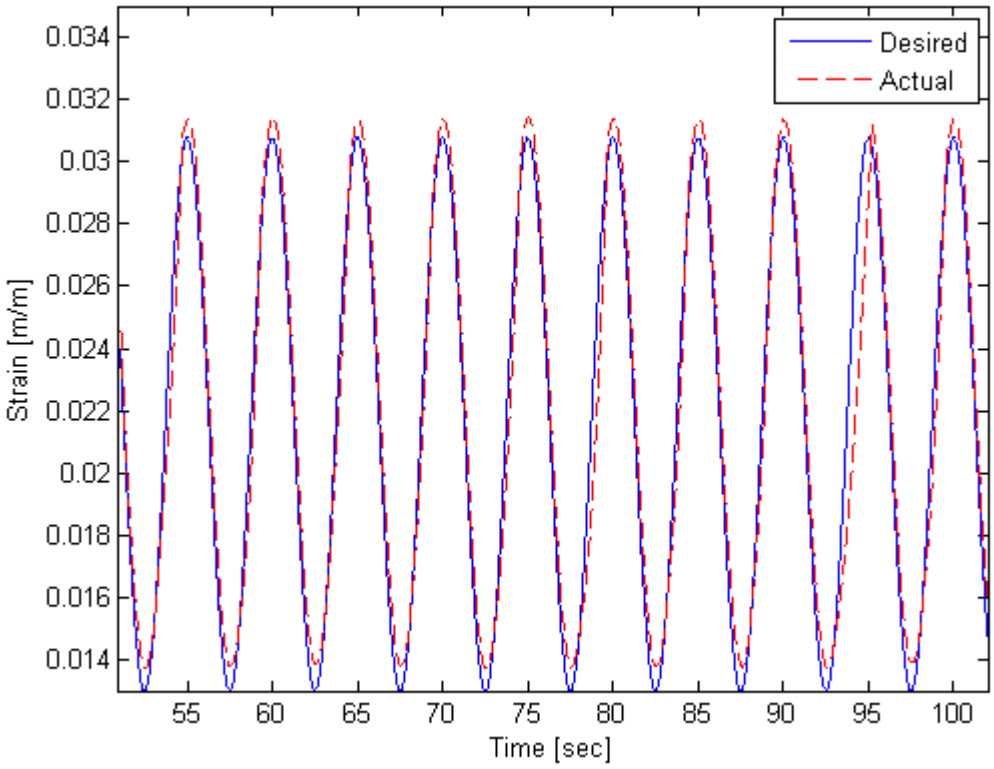

Figure 6.7: Sinusoidal tracking (5 second period), with P controller, constant force SMA actuator

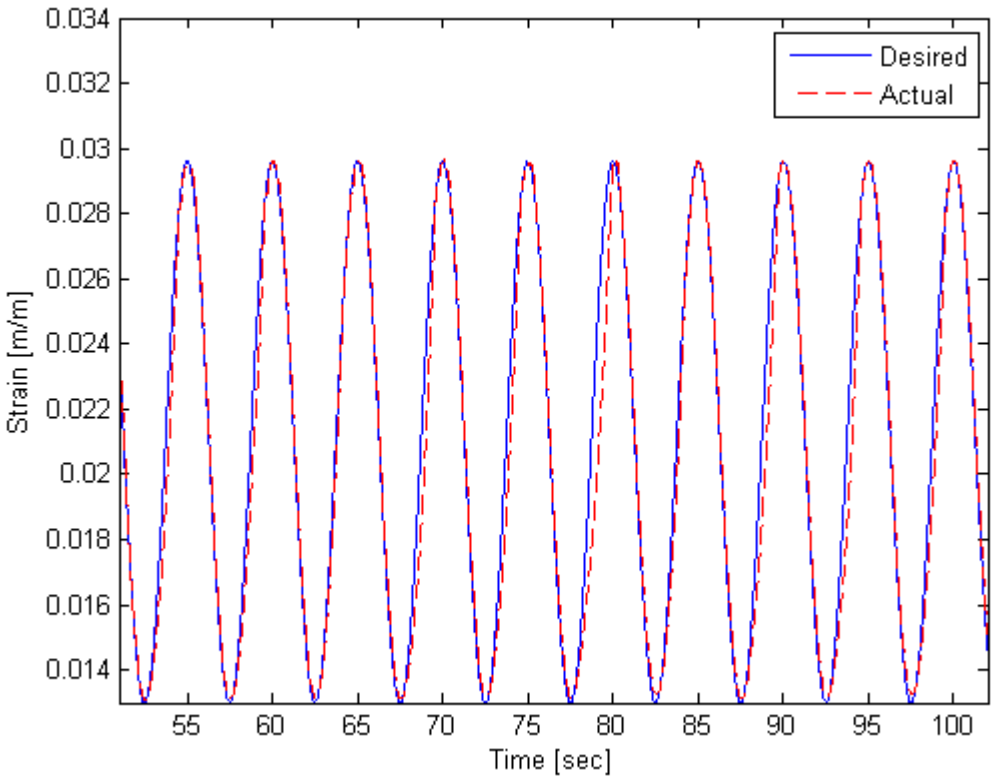

Figure 6.8: Sinusoidal tracking (5 second period), with PI + Inverse controller, constant force SMA actuator

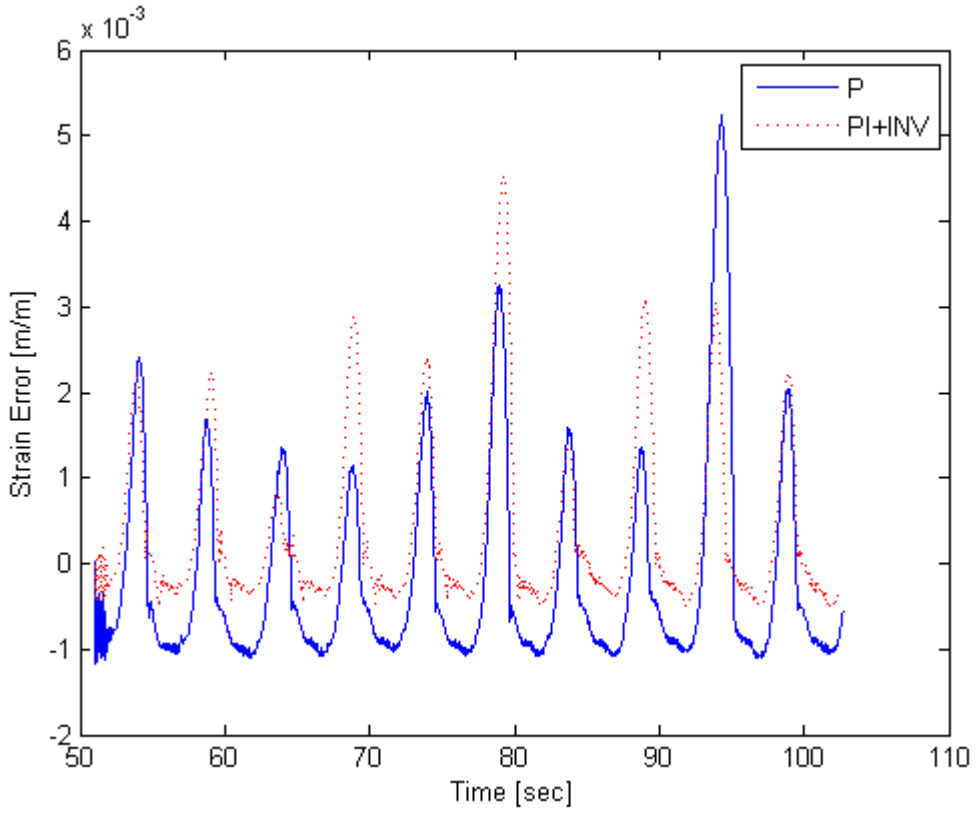

Figure 6.9: Tracking error for sinusoidal reference, constant force SMA actuator

pected because of the lack of an active cooling mechanism and the frequency of the reference strain. Additionally, the performance is the worst around the maximum and the minimum of the reference when the temperature profile has to switch directions. The PI gains for both controllers were tuned separately for the best performance. The proportional gain for both controllers was 850. The integral gain was 0.01 for the PI + INV controller and 0 for the PI controller, reducing the controller to a pure proportional controller and therefore sustaining a steady state tracking error. The presence of integral control typically reduces steady state error. However, any amount of integral gain was found to be detrimental to the tracking control because the integrator accumulates the error from the cooling segments. For future control experiments, the steady state error could be improved when using PI control could be improved by 'turning off' the integral action during the cooling segments in order to prevent the accumulation of error. The error norms for the control schemes are given below:

Table 6.4: Error two-norms, sinusoidal reference, constant force SMA actuator

controller	*error norm*
P	0.0869
PI+Inv	0.0785

The results presented in Figures 6.7 to 6.9 and Table 6.4 show that while both controllers have large errors in the cooling segments, mainly due to the frequency of the reference, the PI + Inverse controller performs better when compared with the P controller especially when the reference strain changes direction.

6.4 Control of Spring-biased SMA Actuator

In these experiments, the force actuator simulates the behavior of a spring in series with an SMA wire. The spring-biased SMA actuator is a common actuator configuration discussed in Section 2.7. The following points summarise the basic actuation mechanism:

- The total strain in the configuration is set at the start of the experiment and is fixed for the duration of the experiment.
- Individual strains in the spring and SMA vary based on temperature of the SMA.
- The actuator does not include dynamics such as inertia on the spring, therefore the forces in the spring and the SMA are equal.

The force that is exerted by a spring, F_{spring} is given by (6.2) where k_{spr} is the spring constant (in newtons)[1] and ϵ is the strain in the spring. When the spring is acting in its elastic region, the force - strain curve is a straight line through the origin with slope of k_{spr}.

$$F_{spring} = k_{spr}\epsilon \tag{6.2}$$

The change in spring strain as the SMA contracts causes a change in F_{spring} according to (6.2). Consequently, the wire is subject to a varying force during the SMA phase transformation. Figure 6.10 shows the force - strain curves of martensite and austenite showing the 'path of travel' of the spring-biased SMA configuration as the SMA wire heats up. When the SMA is cold, the force balance places the operating point on the martensite curve at point 1. As the SMA wire heats up,

[1]Traditionally, the spring constant is given in newtons/meter and the spring force uses the change of length in meters instead of strain.

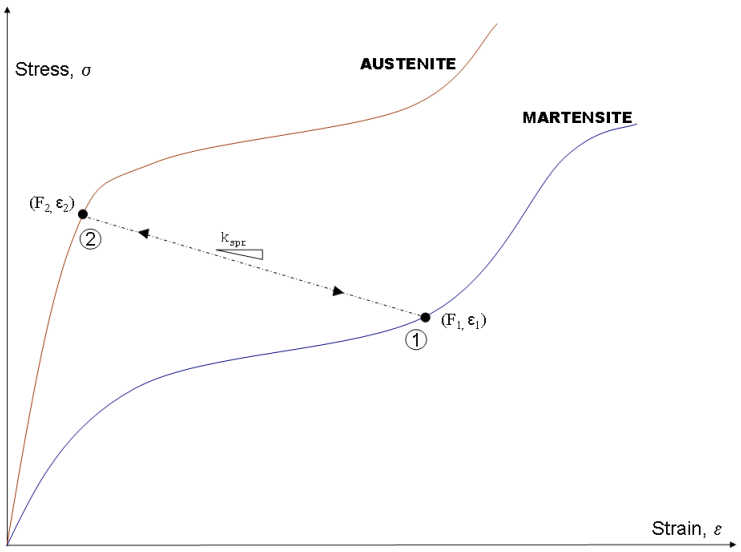

Figure 6.10: Spring-biased SMA actuator, stress - strain diagram

motion in the stress-strain plane is constrained by the spring to the dotted line with slope k_{spr} in Figure 6.10. The spring constant, k_{spr} is determined by considering the start and end strains of the spring-biased SMA actuator (points 1 and 2 in Figure 6.10) as follows:

$$k_{spr} = \frac{F_1 - F_2}{\epsilon_1 - \epsilon_2} \tag{6.3}$$

To implement a spring-biased actuator model using the voice-coil force actuator, the desired force is set based on (6.4). The actuator is sent signals to make it behave as though it were a spring. This is done as follows: For every time step, the current value of the strain as measured by the encoder is used to calculate the value of the force at the next time step using:

$$F_{spr}[k+1] = k_{spr}(\epsilon[k]) - \epsilon[0]) + F_{spr}[0]. \tag{6.4}$$

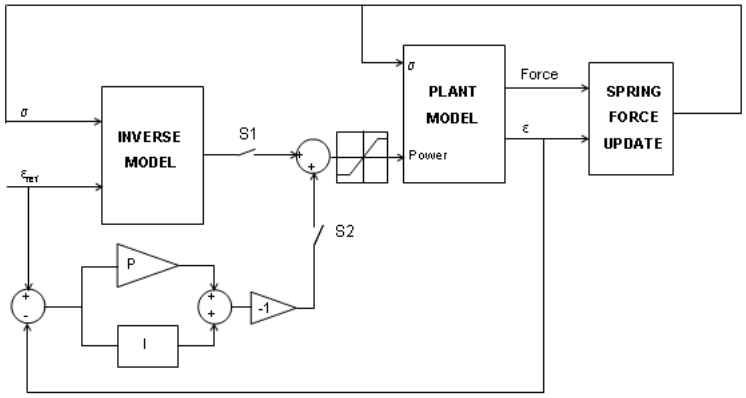

Figure 6.11: Control architecture for spring-biased SMA actuator

Variables $\epsilon[0]$ and $F_{spr}[0]$ are the initial strain and initial force, ϵ_1 and F_1 in Figure 6.10), $\epsilon[k]$ is the strain at the current time step and $F_{spr}[k+1]$ is the force at the next time step.

Figure 6.11 shows the control architecture for the spring - SMA plant. The two architectures compared are PI feedback (S1 open, S2 closed), and feedforward plus PI feedback (both switches closed).

The wire is first pulled when it is in its martensite state (low temperature). This is the 'loading period'. After the loading period, the force actuator reference signal is defined by the simulated spring force (6.4) for the duration of the experiment. The force is controlled by the force control feedback loop described in Section 5.1. As before, the current is saturated at 1A and the PI gains are tuned manually. In the following sections, the tracking performance of a PI controller and a PI controller with the inverse model in feedforward are compared with step and sinusoidal

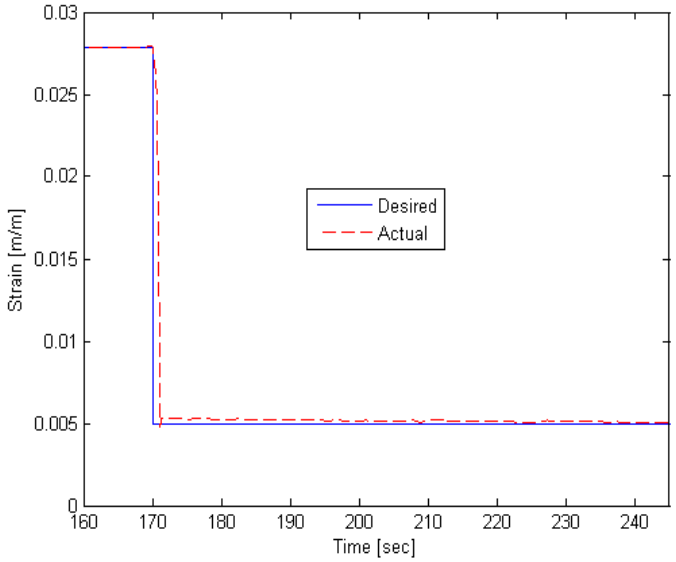

Figure 6.12: Step tracking with PI control for spring-biased SMA actuator

reference strains.

6.4.1 Step Reference Strain

The first reference used for ϵ_{ref} is a step signal from strain of ≈3% to 0.5%.[1] The control signal is started at 160 seconds with the step change occuring at 170 seconds. The controller gains are the same as those used in the constant actuator case i.e. P = 850, I for the PI + Inverse controller is 0.01 and I for the PI controller is 15. Figures 6.12 and 6.13 show the step tracking results using the PI controller and the PI + Inverse controller respectively. Figure 6.14 shows the errors using the two control architectures.

[1]Note that the inital loading sequence is a ramp from 0N to 9N. The strain at the end of this loading is taken as the inital strain and the final strain is set to 0.5%. The inital strain will not be exactly the same for every experiment because loading is carried out using force control and not strain control.

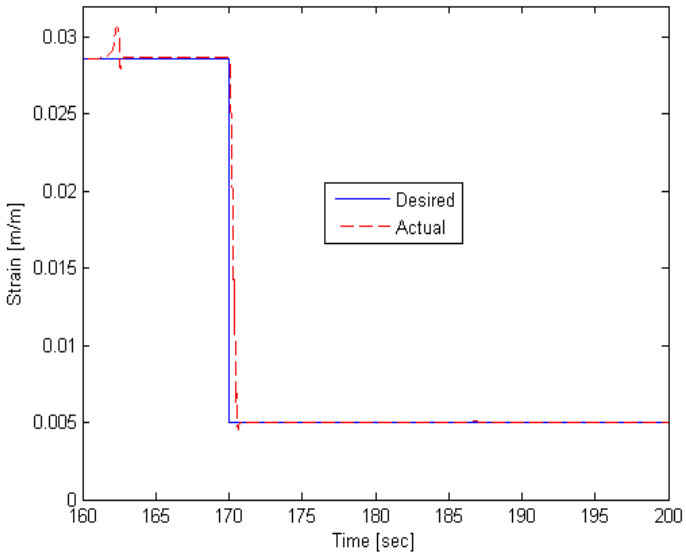

Figure 6.13: Step tracking with PI + Inverse control for spring-biased SMA actuator

At around 162 seconds, the PI + Inverse controller has an error due to a disturbance in the force profile as seen in Figure 6.15. The undershoot in both controllers is, similar to the constant actuator case, due to the response of the force control loop. The error in the step tracking at 170 - 171 seconds is due to saturation in the current signal output by the controller. Figure 6.16 shows the control currents using the PI and PI+Inverse controllers. Similar to the constant force actuator case, the current for the PI controller saturates for a much longer duration than the current for the PI + Inverse controller. Note that maximum control current specified by the PI + Inverse controller (25 A) is higher than the 19 A specified by PI controller. However, only 1A is applied to the wire via a current saturation block in order to protect the wire[1]. Therefore, it is not the magnitude of the current but the duration of the saturation that plays a part in the

[1]The maximum allowable current for 250 μm wire) is 1 A

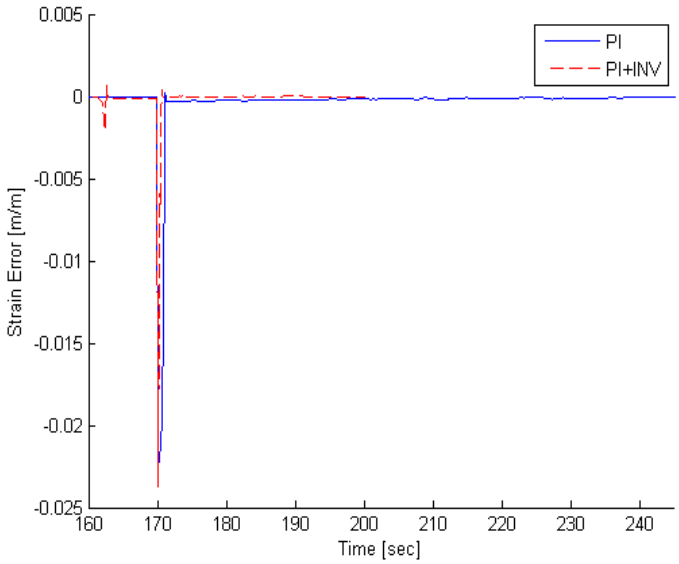

Figure 6.14: Tracking error for step reference with PI and PI + Inverse control for spring-biased SMA actuator

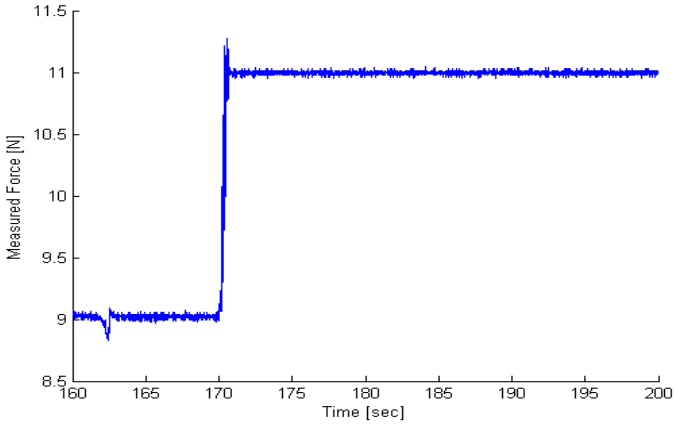

Figure 6.15: Measured force for step reference, PI + Inverse control, spring-biased SMA actuator

tracking error. The effect of the saturation contributes to the undershoot at 171 seconds and higher settling times when using PI control. The PI + Inverse controller also reduces the steady state error present when using only a PI controller. The error norms for the control schemes are given below:

Table 6.5: Error two-norms, step reference, spring-biased SMA actuator

controller	*error norm*
PI	0.1961
PI+Inv	0.1110

The PI + Inverse controller provides better tracking performance for a step signal than using traditional PI control for a spring-biased SMA actuator.

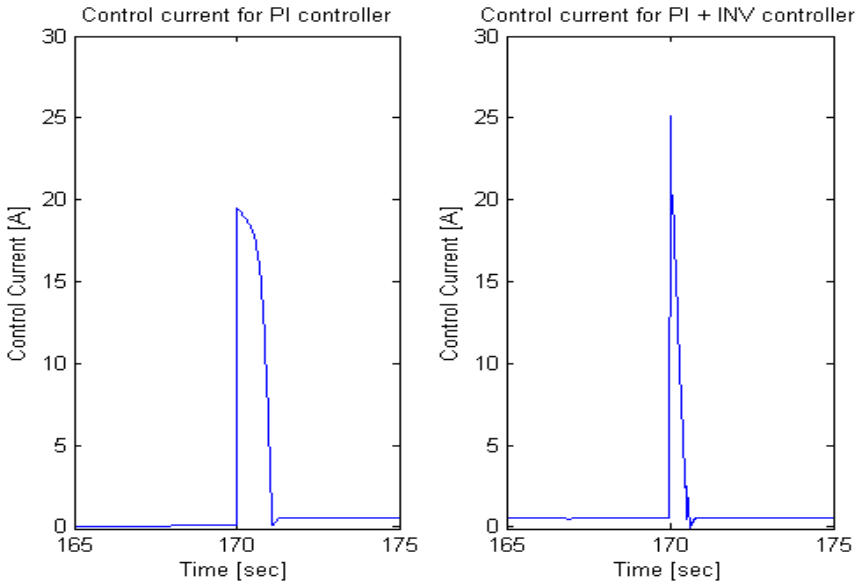

Figure 6.16: Controller output current for step reference with PI and PI + Inverse control in spring-biased SMA actuator

6.4.2 Sinusoidal Reference Strain

The next reference used for ϵ_{ref} is a sinusoidal signal with a period of 5 seconds. As before, the reference is generated using (6.1). Figures 6.17 and 6.18 show the tracking results for a PI controller and PI + Inverse controller both with P gains of 900 and I gains of 0.1. Figure 6.19 shows error of using the two control architectures. Note that the time scale of the figures is from 51 seconds to 102.5 seconds since controller action starts at 51 seconds.

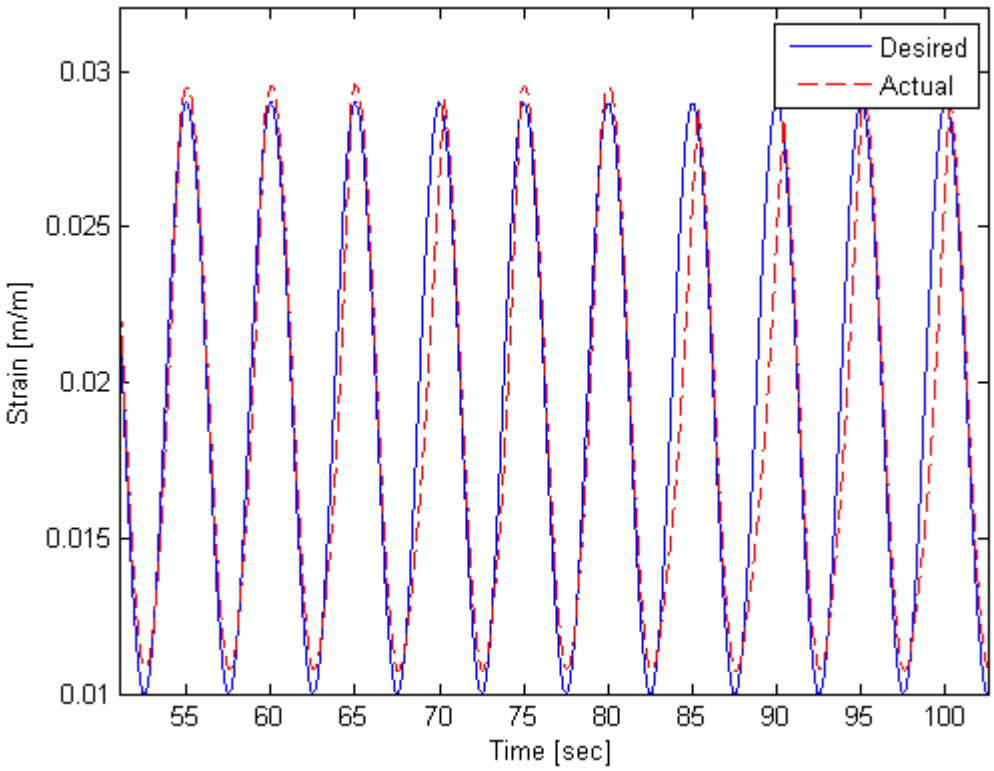

Figure 6.17: Sinusoidal tracking (5 second period) with PI control, spring-biased SMA actuator

The results presented in Figures 6.17 to 6.19 show that similar to the constant force actuator

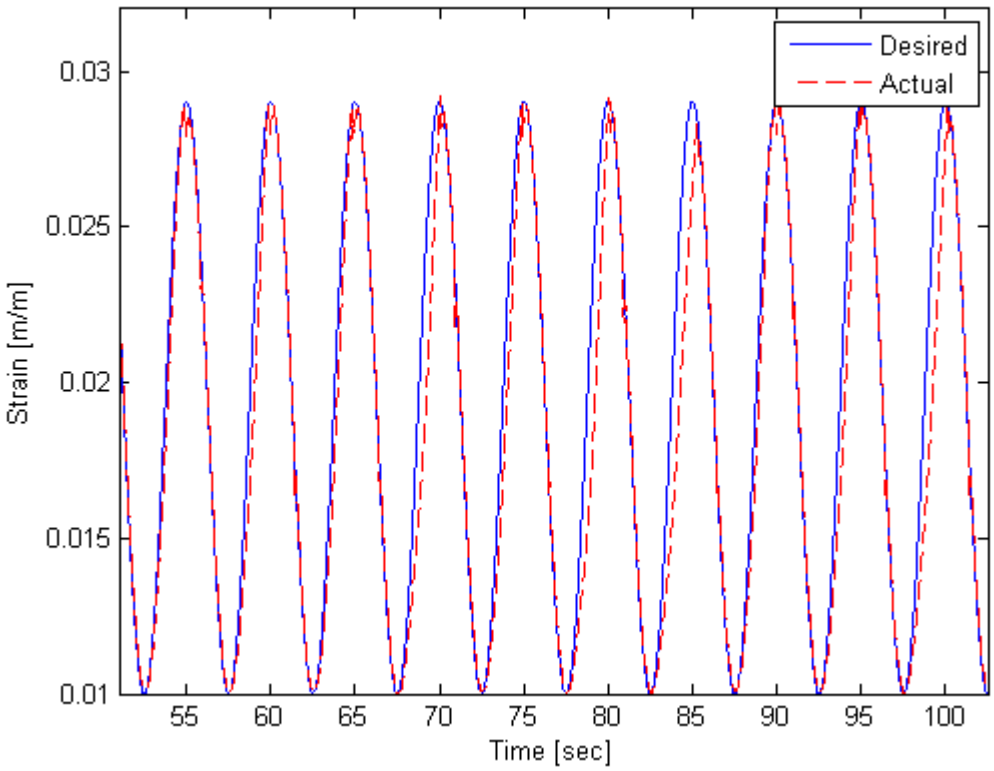

Figure 6.18: Sinusoidal tracking (5 second period) with PI + Inverse control, spring-biased SMA actuator

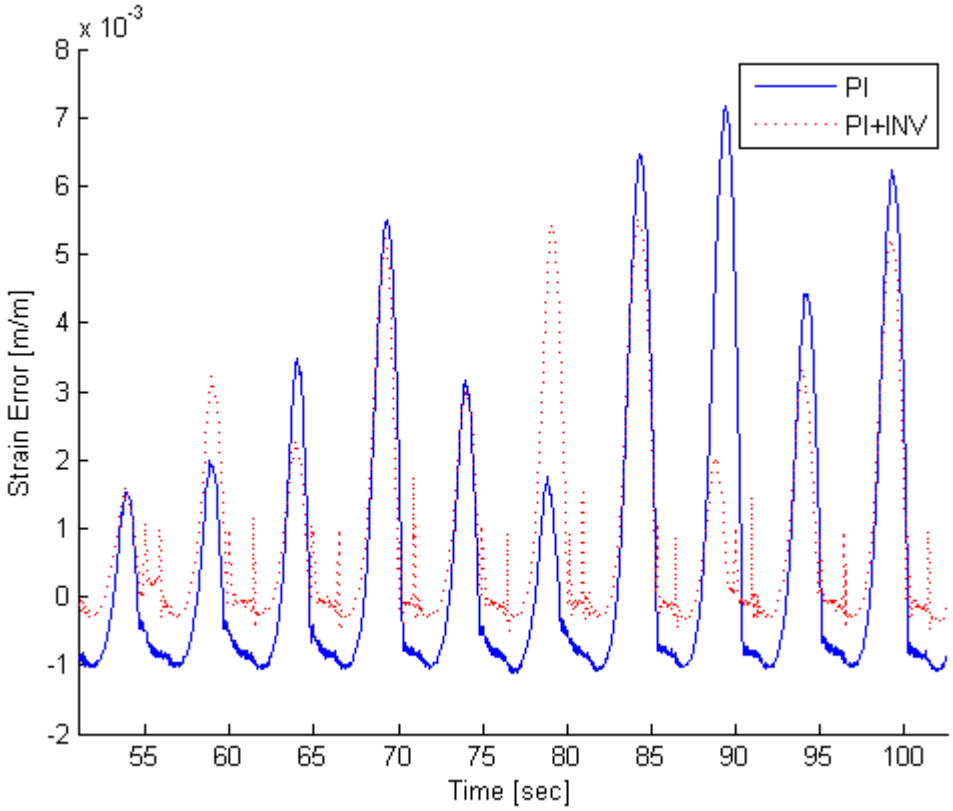

Figure 6.19: Tracking error for sinusoidal reference (5 second period) for spring-biased SMA actuator

case, the largest error is in the cooling segments of the sinusoidal signal. This is due to the lack of an active cooling mechanism and the high frequency of the signal. The low integral gain of 0.1 is necessary because any larger gains cause degradation of performance due to the accumulation of error in the cooling sections. Similar to the constant force actuator case, strain tracking using the PI + Inverse controller performs better than the PI controller especially when the reference strain is changing direction. The error norms are shown below:

Table 6.6: Error two-norms, sinusoidal reference, spring-biased SMA actuator

controller	*error norm*
PI	0.1961
PI+Inv	0.1110

6.5 Control of Variable Force SMA Actuator

Webb, Lagoudas and Kurdila [38] describe an adaptive controller to deal with time-varying stresses different from the stress imposed on the wire by a bias spring. The results in this section present a comparison between the performance of the PI + Inverse model controller and the adaptive controller presented in [38]. The experimental set up in [38] uses a weighted plate suspended on an untrained nitinol wire with diameter of 0.58 mm and undeformed length of 30.48 cm. The displacement of the wire is measured with an LVDT. A varying stress is applied to the wire by adding or removing weights on the plate. The resulting stress profile is shown in Figure 6.20. A similar stress profile is applied to our SMA wire as shown in Figure 6.21. It must be noted though, that the step transitions in our experiment are much faster than the experiment in [38] and the amplitudes of the stresses are higher by 60 MPa. The reference displacement profile

from [38] is a sinusoidal signal with period of 60 seconds and amplitude of approximately 7 mm for two periods after which the amplitude of the reference is reduced to about 5.5 mm with the minimum at 1 mm. The displacement is converted to strain and applied to our experiment. The reference strain does not reduce its amplitude, however, because a reduction in amplitude will allow for better performance by the controller. Keeping the reference strain at the same amplitude as the beginning of the experiment provides a more challenging tracking problem. Additionally, the minimum strain is 0.5%. Figures 6.22 and 6.23 show the experimental tracking results for the adaptive controller and the PI + Inverse controller. The tracking error for the PI + Inverse controller is shown in the top plot of Figure 6.24. The error is worst at the transition time between applied stresses. Additionally, the spikes at 220 seconds, 330 seconds, 450 seconds and 570 seconds are due to the errors in the stress tracking response (bottom plot of Figure 6.24) due to the respose to a step signal and therefore are not caused entirely by the controllers. The error norm for the controller is 0.0681. The length of the wire used in [38] is 304.4 mm, therefore for the first two periods, the strain varies from 0% to 2.29% and then from 2.2% to 0.33% (peak-to-peak strain 1.87%) for the rest of the experiment. A quantitative comparision of the tracking response between the adaptive and PI + Inverse controllers cannot be made due to lack of tracking error data in [38]. However, qualitatively, the PI + Inverse controller tracks a more agressive reference with a peak-to-peak amplitude of 2.25% and 60 MPa higher stress. Nevertheless, other than errors due to the force control due to fast step transitions, the PI + Inverse controller tracks reasonable well and seems to perform better than the adaptive controller.

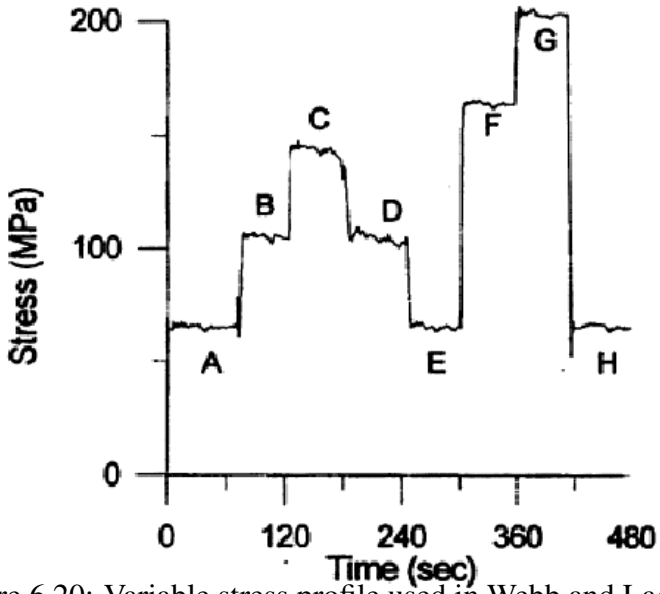

Figure 6.20: Variable stress profile used in Webb and Lagoudas [38]

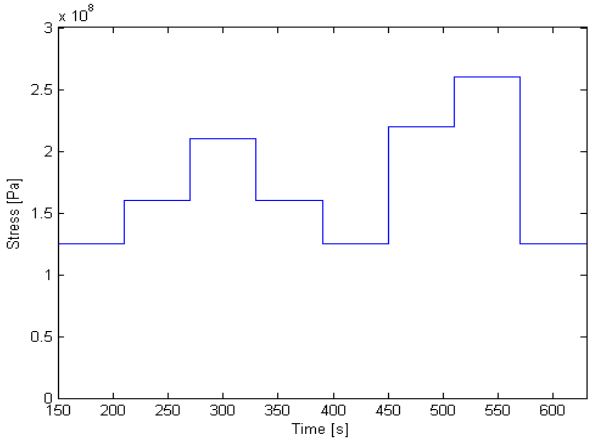

Figure 6.21: Variable stress profile applied to SMA wire

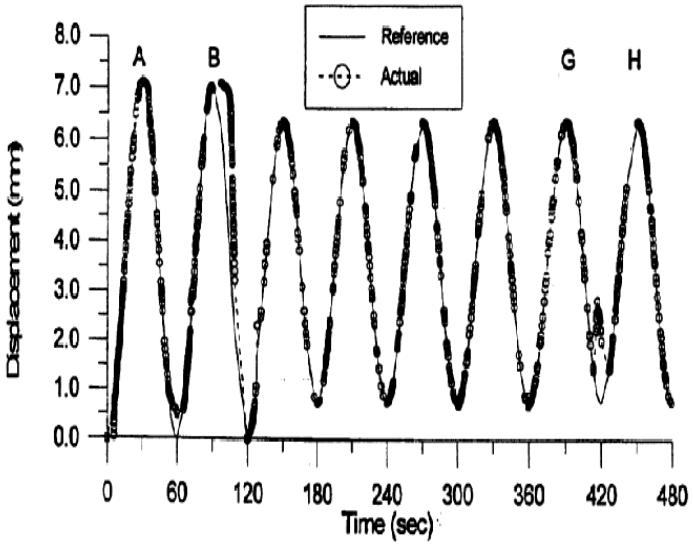

Figure 6.22: Experimental results with variable stress profile, adaptive control [38]

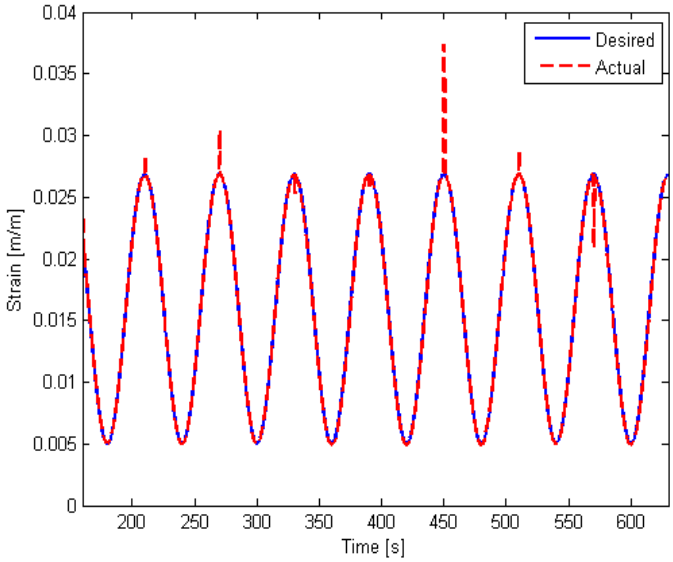

Figure 6.23: Experimental results with variable stress profile, PI + Inverse control

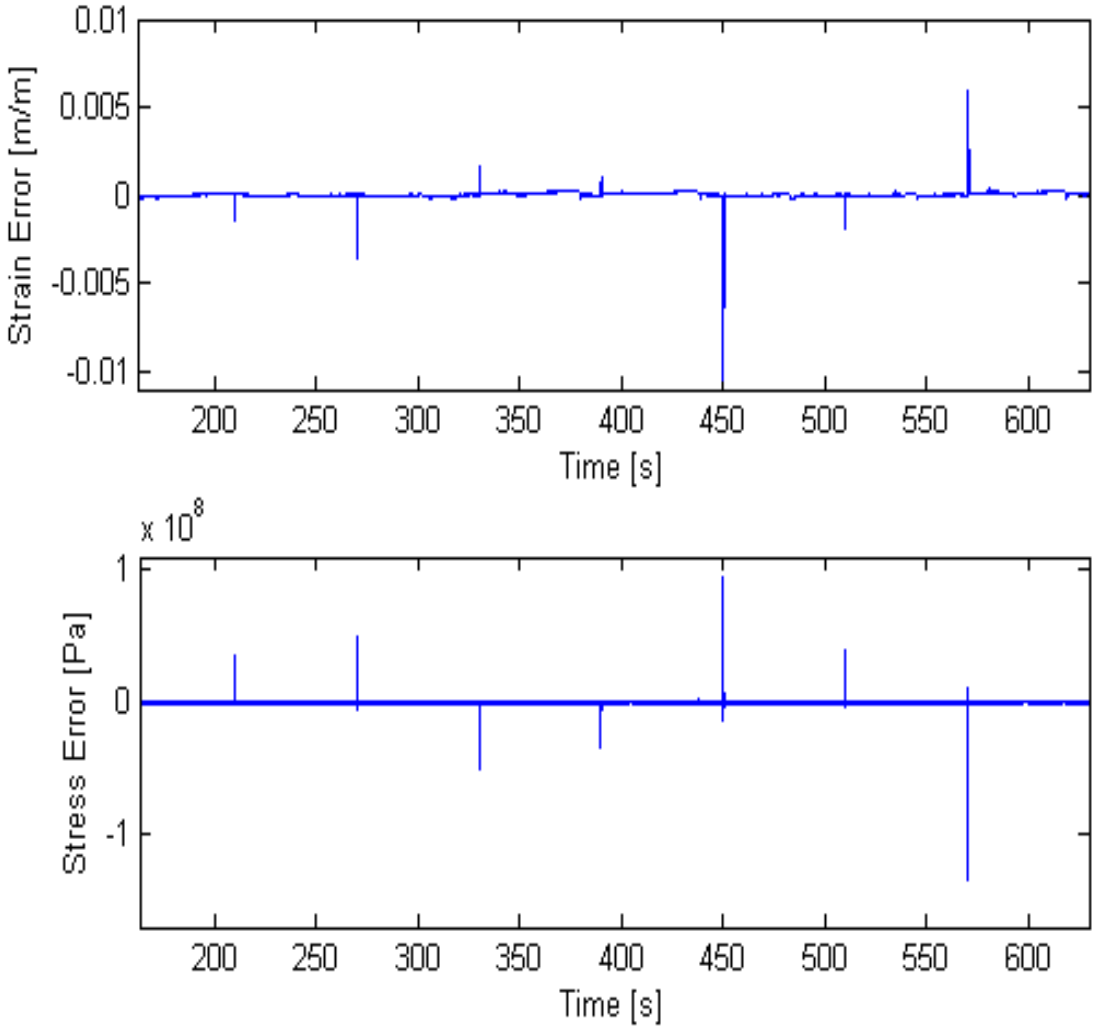

Figure 6.24: Strain tracking error and force error for SMA wire with variable stress profile

6.6 Disturbance Rejection

In this section, we examine the ability of the two closed-loop controllers, PI and PI with Inverse control, to reject transient and sinusoidal stress disturbances applied to a constant force SMA actuator during sinusoidal tracking. Disturbance rejection is investigated in two scenarios. The first investigation assumes that there is full knowledge of the stress profile with the disturbance via force sensor data. The second investigation assumes that there is no force sensor and therefore the system has no knowledge of the varying stress profiles applied to the actuator. The disturbance profiles investigated are impulse and sinusoidal disturbances.

6.6.1 Disturbance Rejection with Force Sensor

In order to simulate the disturbance of a system with a force sensor and quantitatively compare the PI + Inverse and PI controllers, a disturbance profile is applied to both the force control and strain control loops (Figure 6.25).

Impulse Disturbance

Figure 6.26 shows the commanded stress profile. The constant load on the actuator is 9N, or approximately 184 MPa. The "impulse" stress is a 100 MPa spike with a duration of 1 second. The first spike occurs in the heating portion of the sinusoidal tracking, at 56 seconds, and the second in the cooling portion of the tracking at 59 seconds. The figure also shows good agreement between the commanded stress and the actual stress profile as sensed by the force sensor.

Figures 6.27 and 6.28 show the tracking response of the system to a sinusoidal strain of period 5 seconds. The control acts from 51 seconds to 62.5 seconds. The vertical lines from 56 to 57

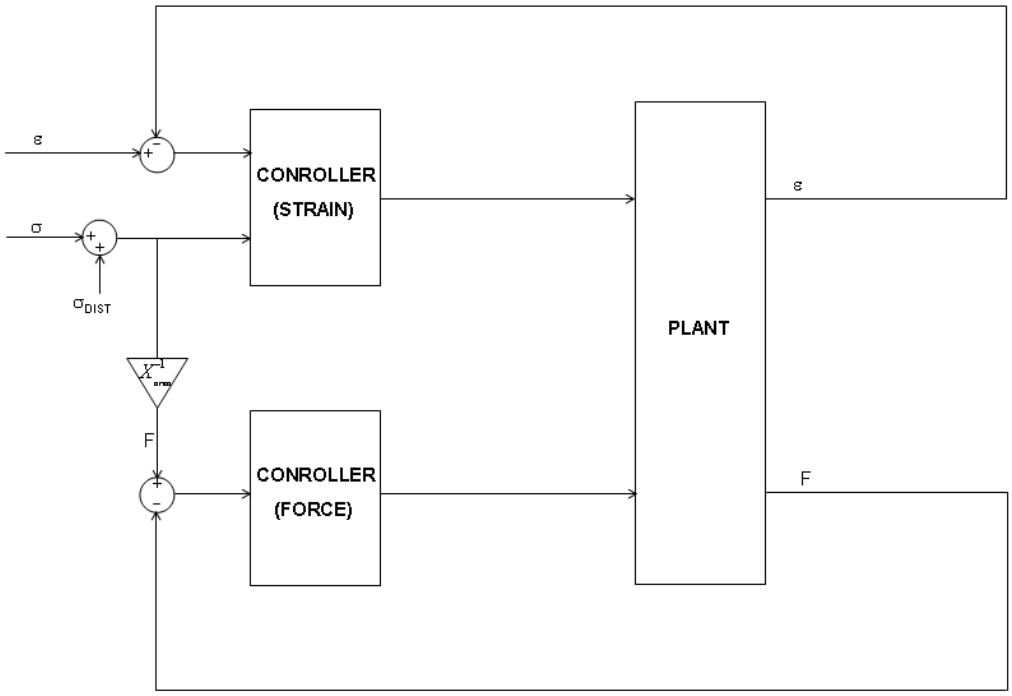

Figure 6.25: Control loop configuration to simulate disturbance with force sensor

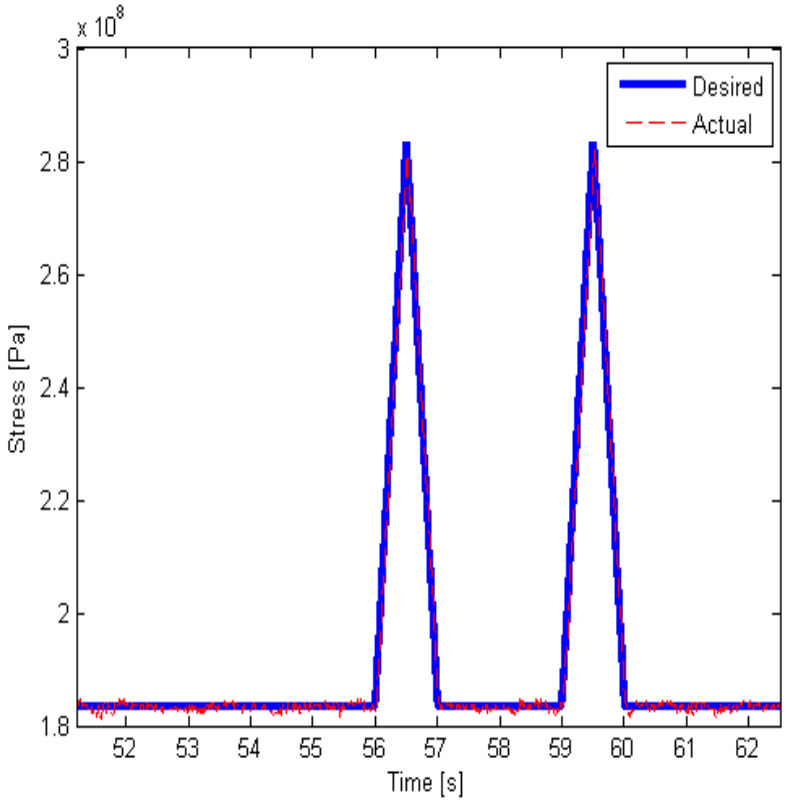

Figure 6.26: Impulse stress disturbance

seconds and 59 to 60 seconds mark the time that the disturbance occurs.

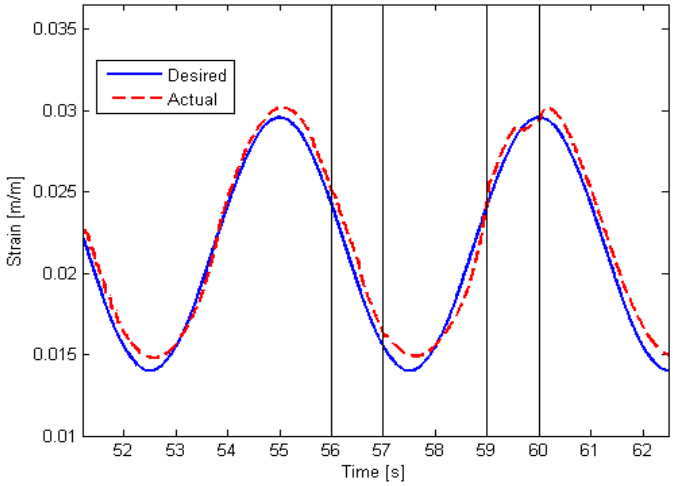

Figure 6.27: Tracking response with impulse stress disturbance and PI control, with force sensor

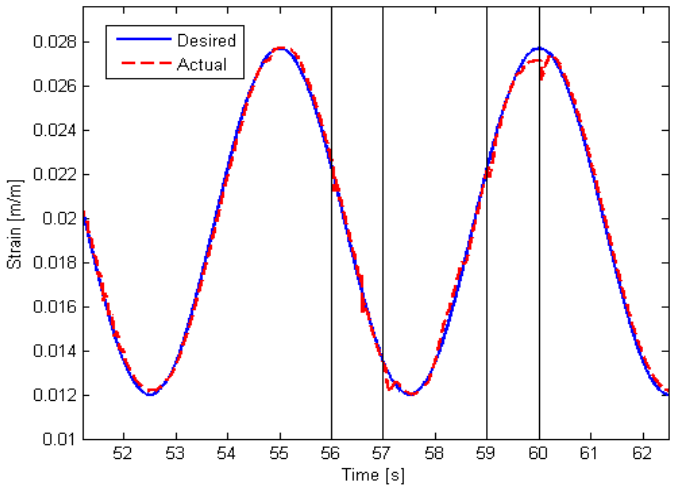

Figure 6.28: Tracking response with impulse stress disturbance and PI + Inverse control, with force sensor

Figure 6.29 shows the error of the two control architectures. Figures 6.30 and 6.31 show the strain error around the time of the disturbance. The vertical lines from 56 to 57 seconds and 59 to 60 seconds mark the time that the disturbance occurs. As expected, impulse disturbances cause error spikes. However, the PI controller has larger spikes due to errors in strain tracking during the cooling segments of the sinusoidal reference. In order investigate disturbance rejection, it is necessary to remove the effect of error due to tracking before computing error norms. The effect of tracking errors is removed as follows: From the start of controller action at 51 seconds to the first disturbance at 56 seconds, the strain error present is purely tracking error and represents a full period of input. The error norm for the entire signal from 51 seconds to 56 seconds is assumed to be representative of the tracking error in the system and is subtracted from the error norm of the entire signal from 51 seconds to 62.5 seconds in order to isolate errors due to the disturbances. The resulting adjusted error norms are shown in Table 6.7. The table shows that the disturbance errors are worst when the disturbance occurs in the cooling section. This is expected because the controllers perform poorly in the cooling section contributing not only to higher tracking errors but also poorer disturbance rejection capabilities.

Table 6.7: Error two-norms, impulse disturbance, with force sensor

controller	*error norm (51-62.5 s)*	*error norm (55.5-56.5 s)*	*error norm (59.5-60.5 s)*
PI	0.0134	0.0112	0.0120
PI+Inv	0.0045	0.0024	0.0036

Figure 6.29 and Table 6.7 shows, however, that the PI + Inverse controller is more robust than the PI controller.

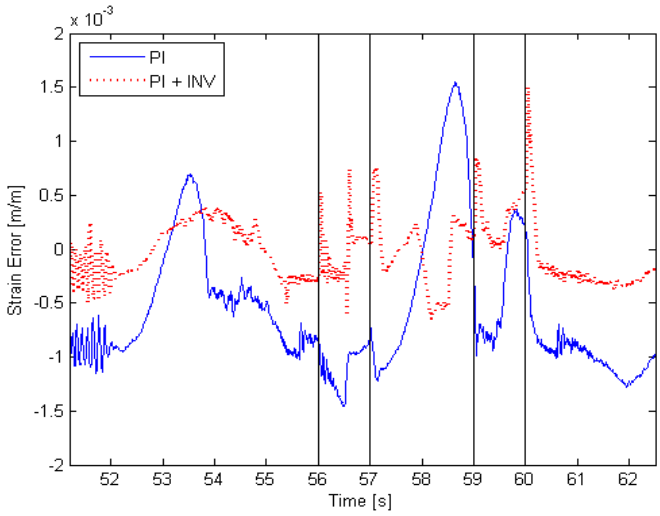

Figure 6.29: Tracking error with impulse stress disturbance, with force sensor

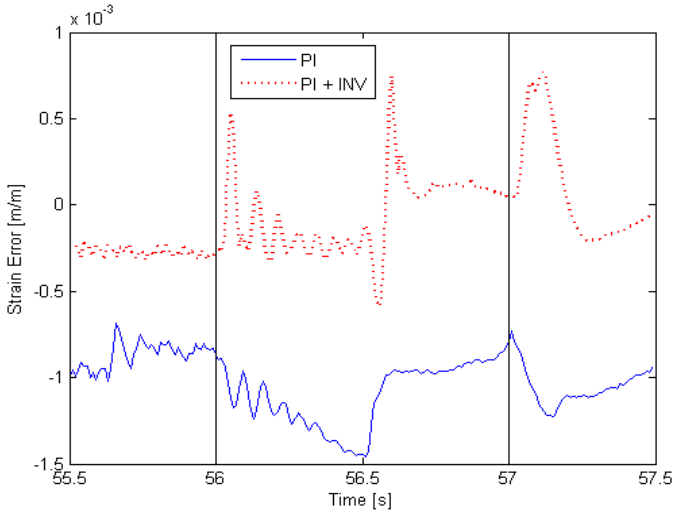

Figure 6.30: Tracking error with impulse stress disturbance during duration of first impulse

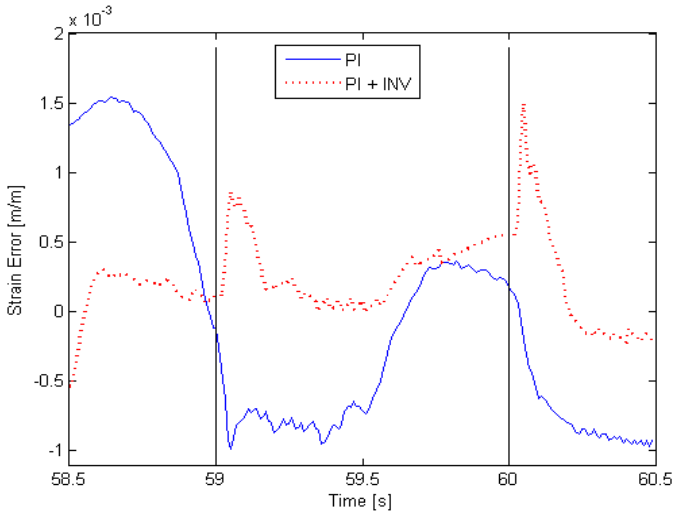

Figure 6.31: Tracking error with impulse stress disturbance during duration of second impulse

Sinusoidal Disturbance

In this section a sinusoidal signal is superimposed on the stress profile to produce a sinusoidal disturbance. Low frequency sinusoidal disturbances may be encountered when a controlled structure resonates at its natural frequency. The sinusoidal disturbance has frequency of 0.02 Hz and is applied for 2 periods as shown in Figure 6.32. The figure also shows good agreement between the desired and actual force signals.

Figures 6.33 and 6.34 show the tracking response of the system with a PI controller and a PI + Inverse controller respectively. The control acts from 51 seconds to 150 seconds.

Figure 6.37 shows the errors in the two control architectures with the disturbance. The errors for the PI controller are seen to be worse than the PI + Inverse controller. Figure 6.35 shows that when the stress disturbance is increasing, the actual signal undershoots the reference at the maximum and minimum strains. The reason for this is that when the disturbance stress is increasing

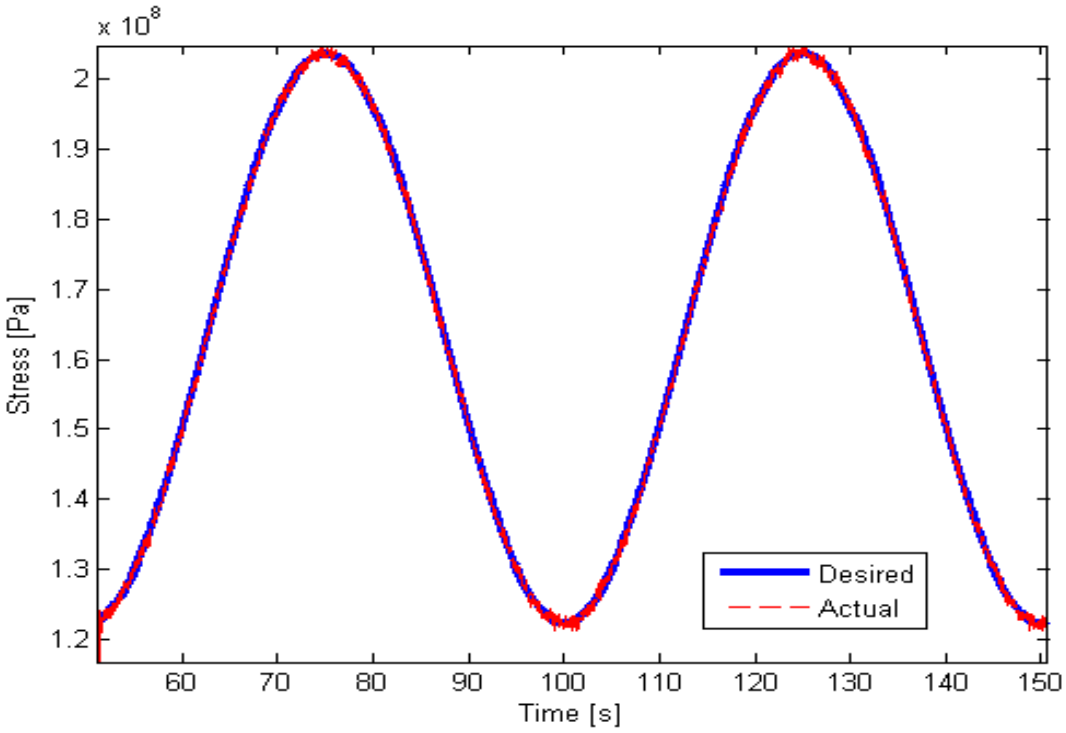

Figure 6.32: Sinusoidal stress disturbance

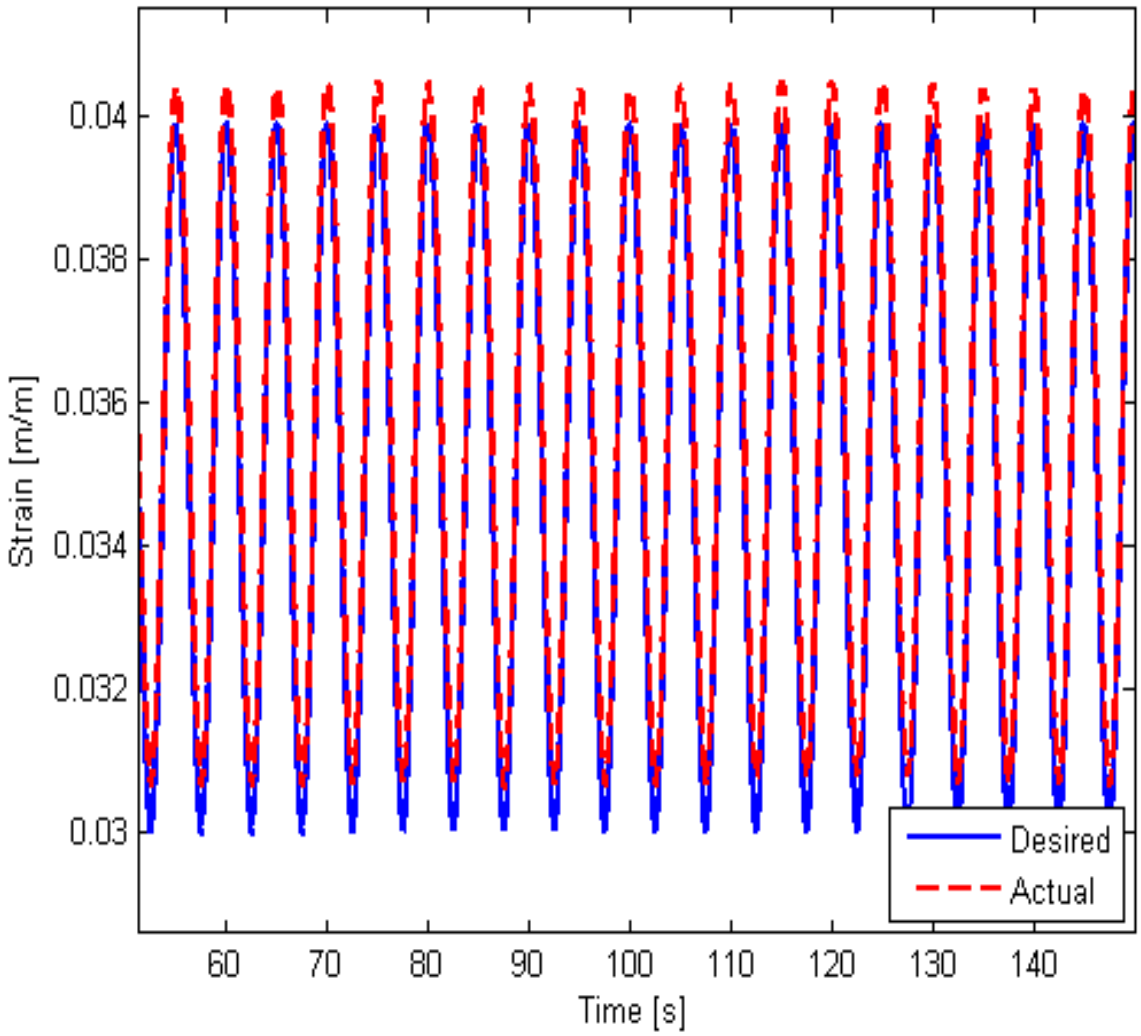

Figure 6.33: Tracking response with sinusoidal stress disturbance and PI control, with force sensor

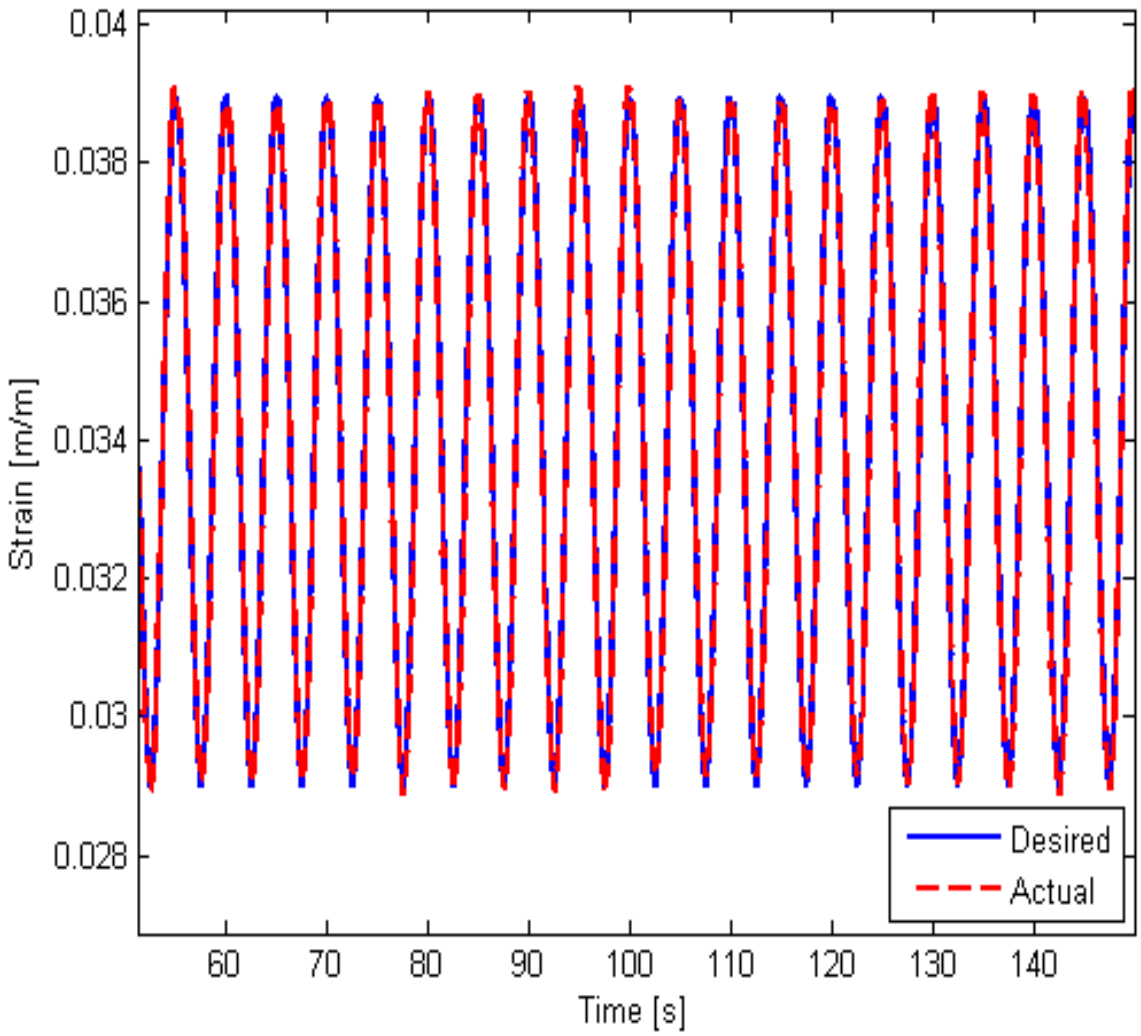

Figure 6.34: Tracking response with sinusoidal stress disturbance and PI + Inverse control, with force sensor

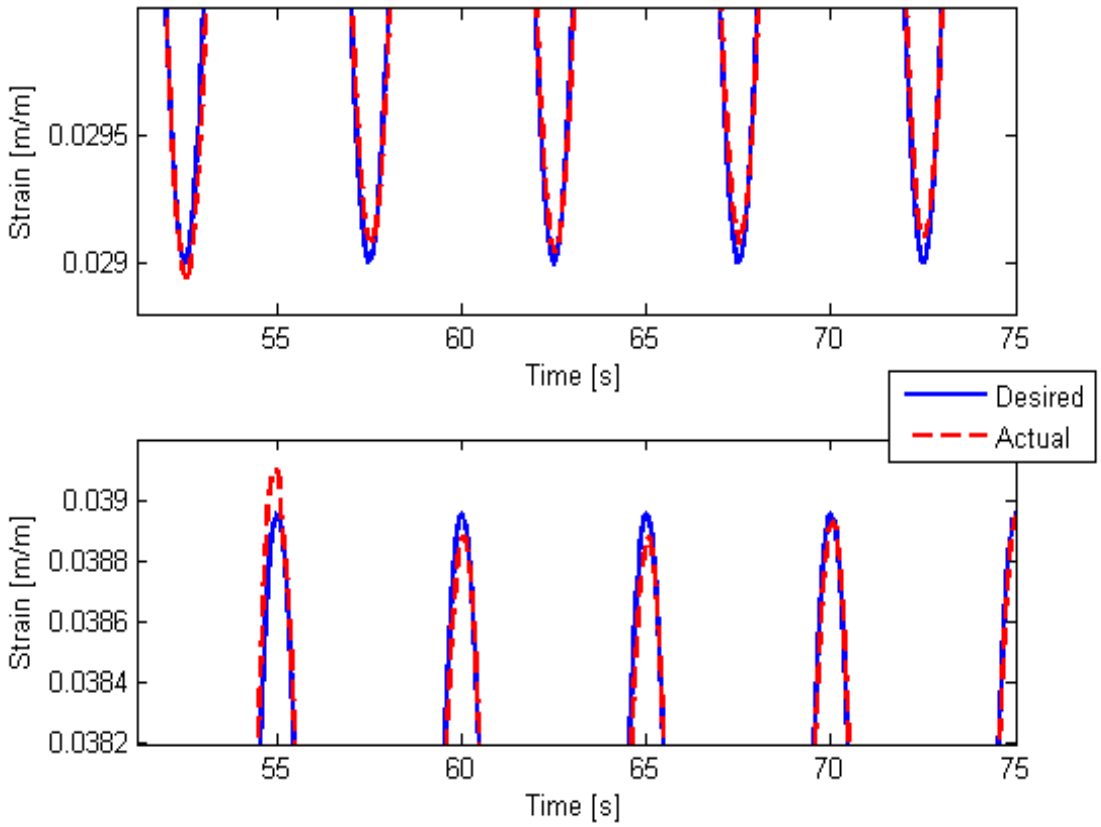

Figure 6.35: Tracking response with sinusoidal stress disturbance and PI + Inverse control, with force sensor from 51 to 75 seconds

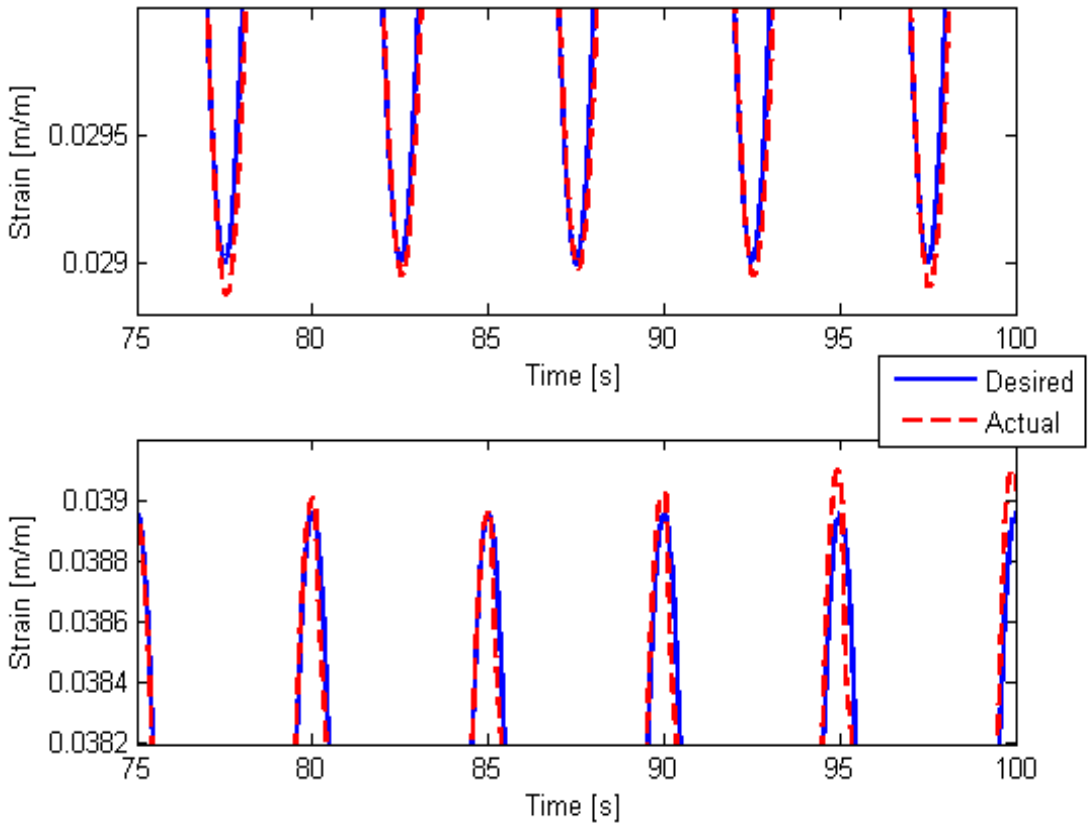

Figure 6.36: Tracking response with sinusoidal stress disturbance and PI + Inverse control, with force sensor from 75 to 100 seconds

and the reference strain is decreasing, the strain increases too and the wire needs more heat to contract to same length. The hysteresis curve has a lower slope because the increasing stress increases the transformation temperatures and the wire takes longer to contract. The undershoots during the cooling segments are consistent with tracking errors because of the lack of an active cooling mechanism. Figure 6.36 shows the tracking response around the mininmum and maximum strains when the disturbance stress is decreasing. In this case, the wire does not need to be heated as much and the actual strain overshoots the reference at the minimums. On the cooling segments, the lower energy input into the wire in the heating segment translates to faster cooling times and therefore at the maximum of the reference, there is either good tracking or overshoot.

Figure 6.37 shows that the error for the PI + Inverse controller seems displays patterns that appear correlated with increasing or decreasing stress periods. When the stress disturbance is increasing, the error is similar to just the PI controller. When the stress is decreasing, the error is more erratic but more centered around 0. Figure 6.38 shows that the inverse controller is not contributing significantly to the total control current signal when the stress is increasing. The reason for this is that the inverse controller considers the increasing stress as contributing positively to the tracking. The PI controller is doing most of the work to compensate for the errors. When the stress is decreasing, however, the inverse controller commands high currents and high cooling in order to compensate for the lower stress. Nevertheless, the figures shows the PI + Inverse controller is more robust than the PI controller for sinusoidal disturbances. This is further shown by the error norms for the two control architectures (Table 6.8).

Table 6.8: Error two-norms, sinusoidal disturbance, with force sensor

controller	*Error norm*
PI	0.0613
PI+Inv	0.0195

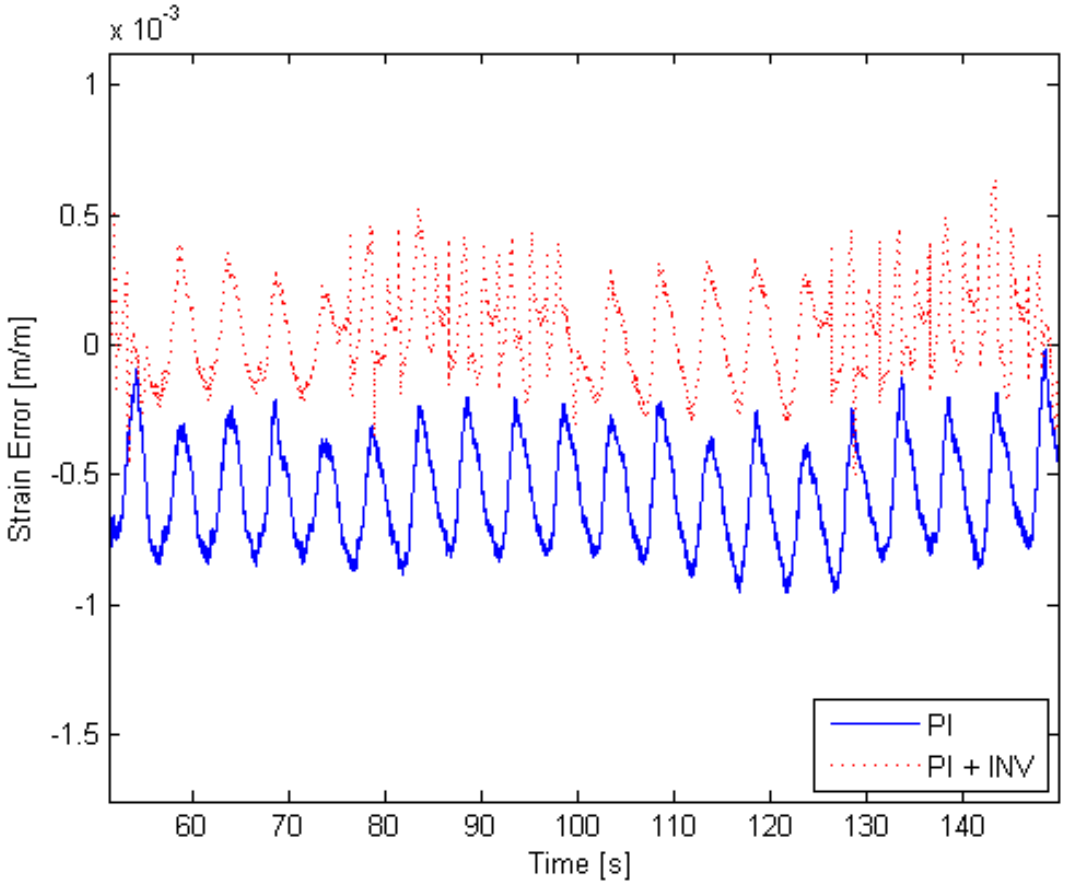

Figure 6.37: Tracking error for sinusoidal stress disturbance, with force sensor

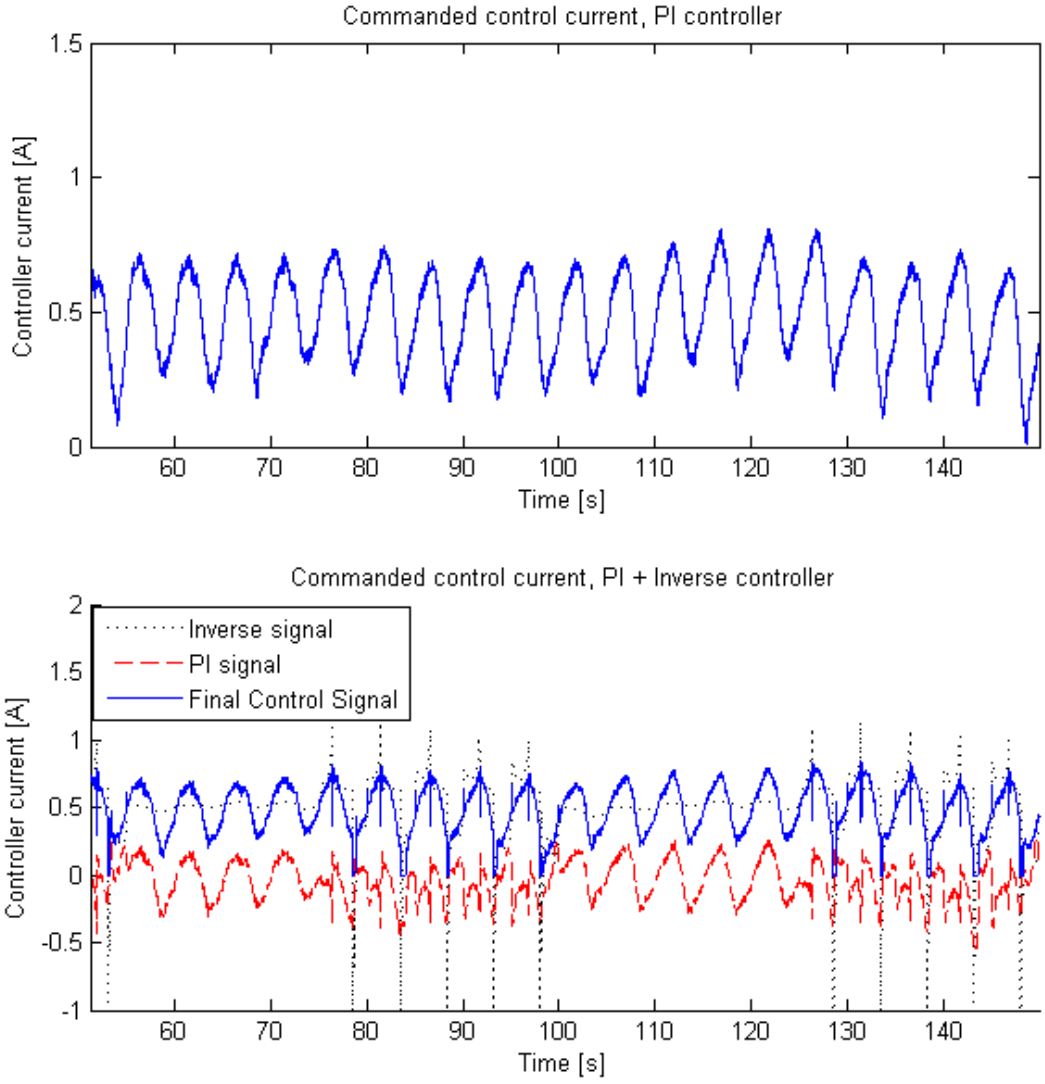

Figure 6.38: Control signal for sinusoidal stress disturbance, with force sensor

6.6.2 Disturbance Rejection without Force Sensor

In order to simulate the disturbance of a system without a force sensor and quantitatively compare the PI + Inverse and PI controllers, a disturbance profile is applied only to the force control loop but not the strain control loop (Figure 6.39). The disturbance profiles investigated are impulse disturbances and sinusoidal disturbances.

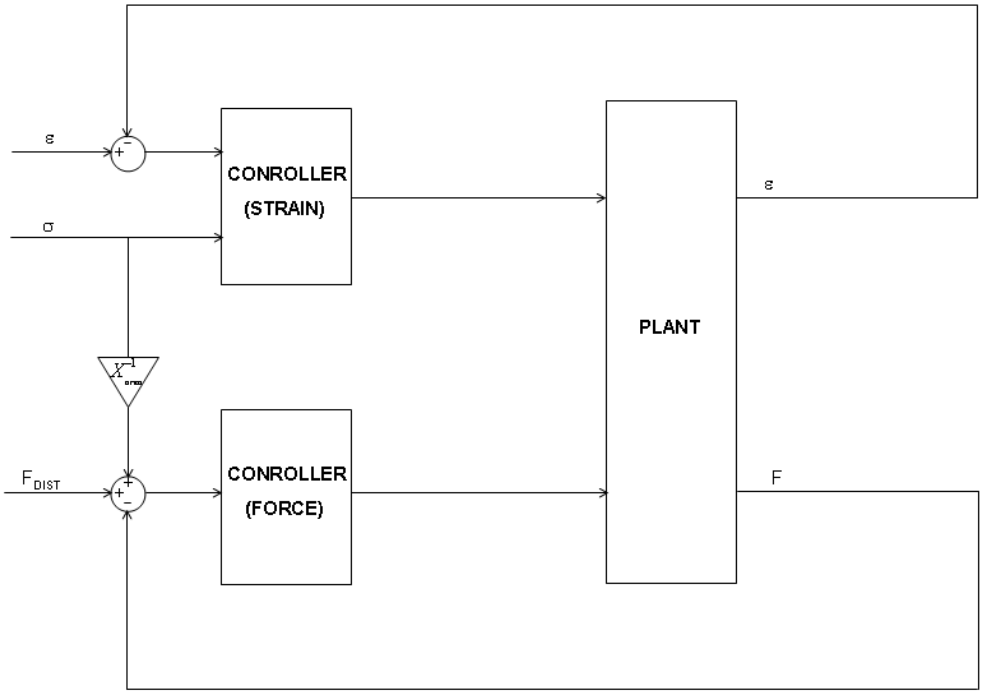

Figure 6.39: Control loop configuration to simulate disturbance without force sensor

Impulse Disturbance

The "impulse" stress is a 100 MPa spike with a duration of 1 second. The first one occurs in the heating portion of the tracking at 56 seconds and the second in the cooling portion of the tracking at 59 seconds. Figures 6.40 and 6.41 show the tracking response of system to a sinusoidal stress of period 5 seconds with vertical lines from 56 to 57 seconds and 59 to 60 seconds marking the duration of the disturbance. The control acts from 51 seconds to 62.5 seconds.

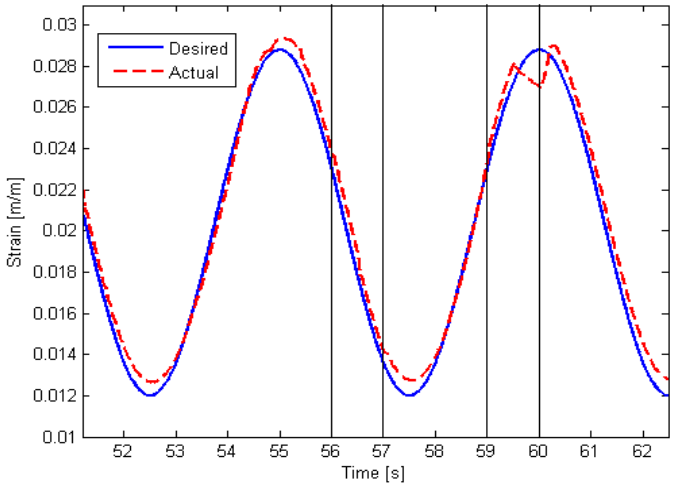

Figure 6.40: Tracking response with impulse stress disturbance and PI control, without force sensor

Figure 6.29 shows the error of the two control architectures. Figures 6.43 and 6.44 show the strain error around the time of the disturbance. The figures show that around the time of the disturbance, the pattern of the error is similar between the PI and PI + Inverse controllers. This is expected because the PI + Inverse controller has no prior knowledge of the disturbance, therefore will react in a similar manner to the disturbance as the PI controller. Figure 6.45 shows

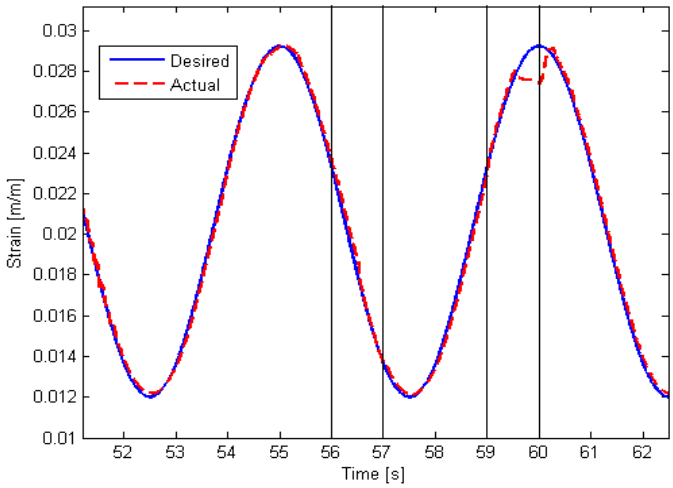

Figure 6.41: Tracking response with impulse stress disturbance and PI + Inverse control, without force sensor

the current signals sent to the wire from the controllers. The PI + Inverse controller commands a higher current due to the added signal from the inverse controller. This additional current accounts for the offset that exists in error plots between the two controllers. Around the first disturbance, however, the commanded signal saturates therefore losing some of the benefits of having the inverse controller. Table 6.9 shows the error norms for the controllers. As before, the effect of tracking errors is subtracted from the error norms of the signal to give a clearer idea of the disturbance rejection capabilities of the controllers.

Table 6.9: Error two-norms, impulse disturbance, without force sensor

controller	*error norm (51-62.5 s)*	*error norm (55.5-56.5 s)*	*error norm (59.5-60.5 s)*
PI	0.0114	0.0092	0.0095
PI+Inv	0.0085	0.0044	0.0087

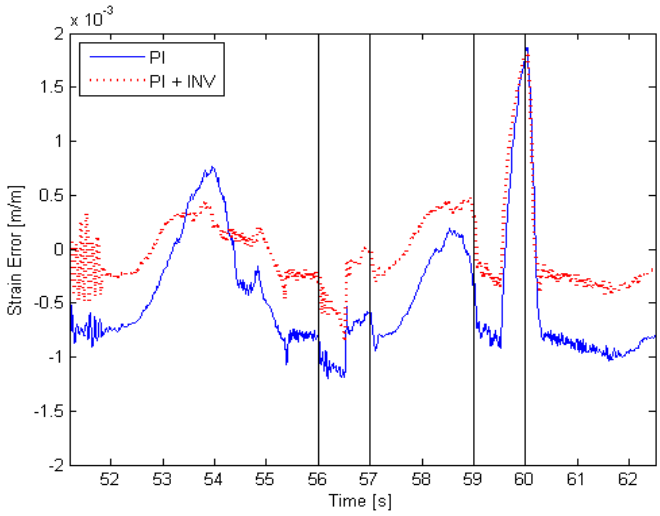

Figure 6.42: Tracking error with impulse stress disturbance, without force sensor

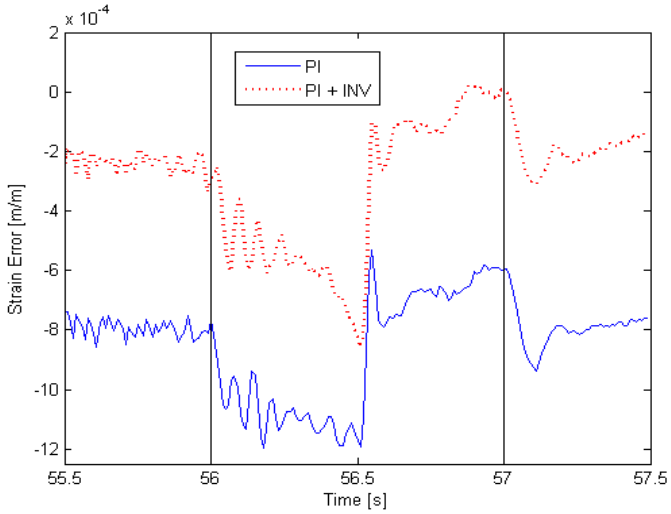

Figure 6.43: Tracking error with impulse stress disturbance during duration of first impulse

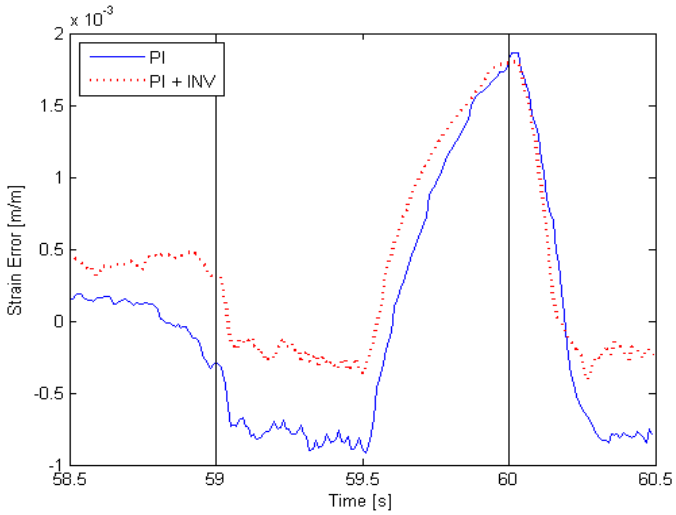

Figure 6.44: Tracking error with impulse stress disturbance during duration of second impulse

Sinusoidal Disturbance

In this section a sinusoidal signal is superimposed on the stress profile to produce a sinusoidal disturbance. The sinusoidal disturbance has a frequency of 0.02 Hz and is applied for two periods.

Figures 6.46 and 6.47 show the tracking response of the system with a PI controller and a PI + Inverse controller respectively. The control acts from 51 seconds to 150 seconds.

Figure 6.48 shows the error of the two control architectures with the disturbance. The errors for the PI controller are shown to be worse than the PI + Inverse controller. The figures show that similar to the impulse disturbance presented previously, the pattern of the error is similar between the PI and PI + Inverse controllers. This is expected because the PI + Inverse controller has no prior knowledge of the disturbance, therefore will react in a similar manner to the disturbance as the PI controller. Additionally, the offset in the error and consequently better performance of

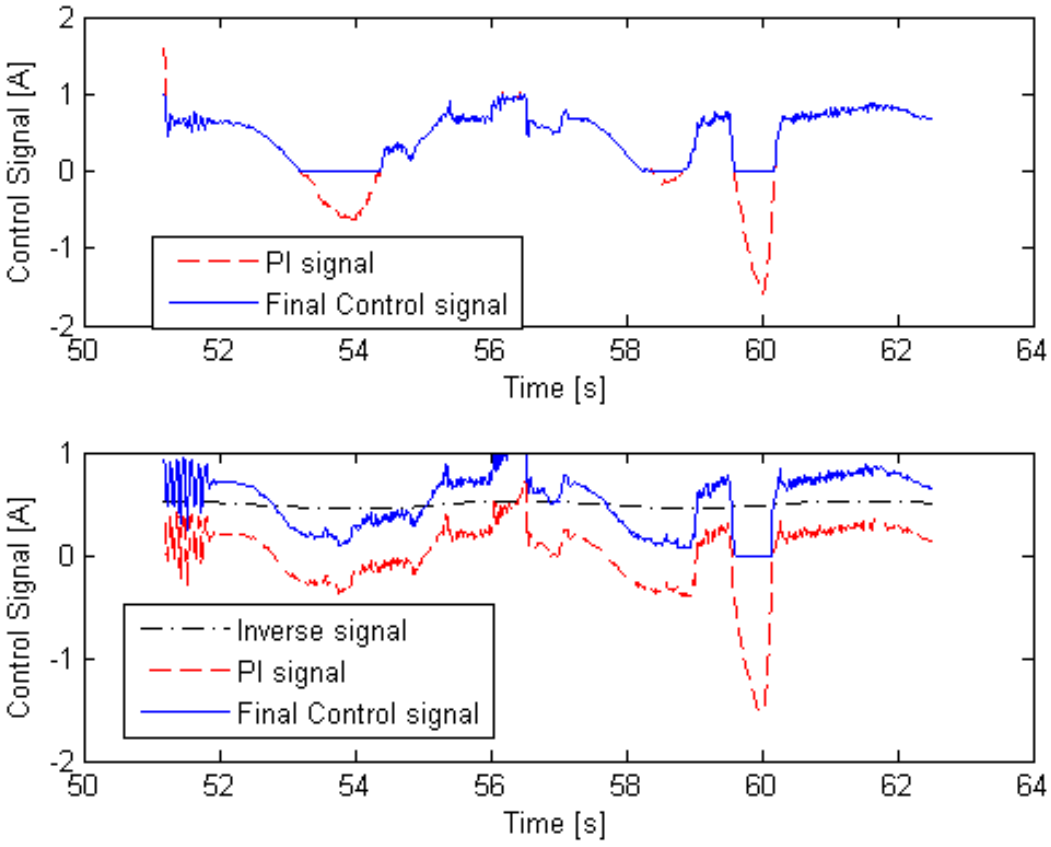

Figure 6.45: Control signal for impulse disturbance without force sensor

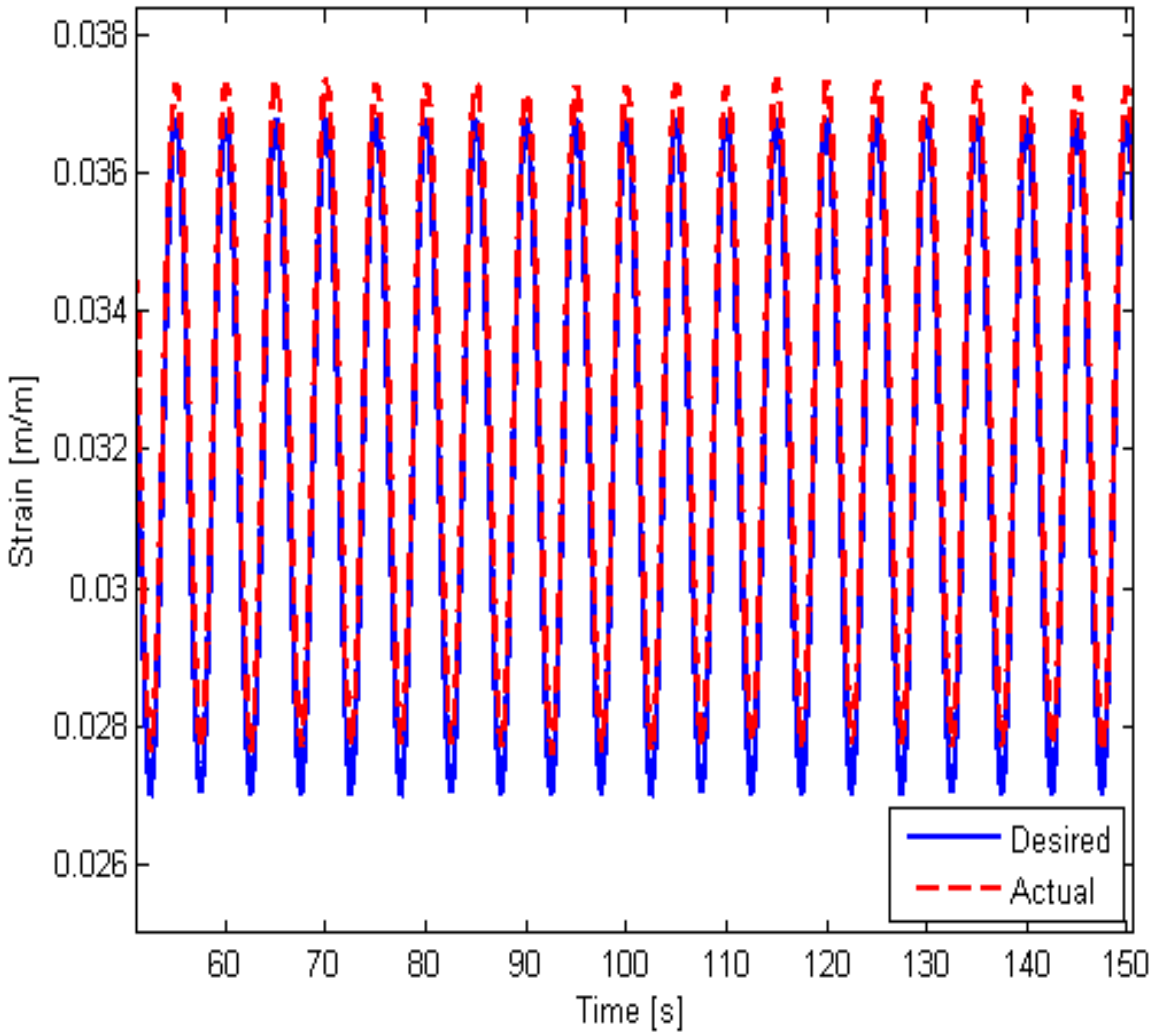

Figure 6.46: Tracking response with sinusoidal stress disturbance and PI control, without force sensor

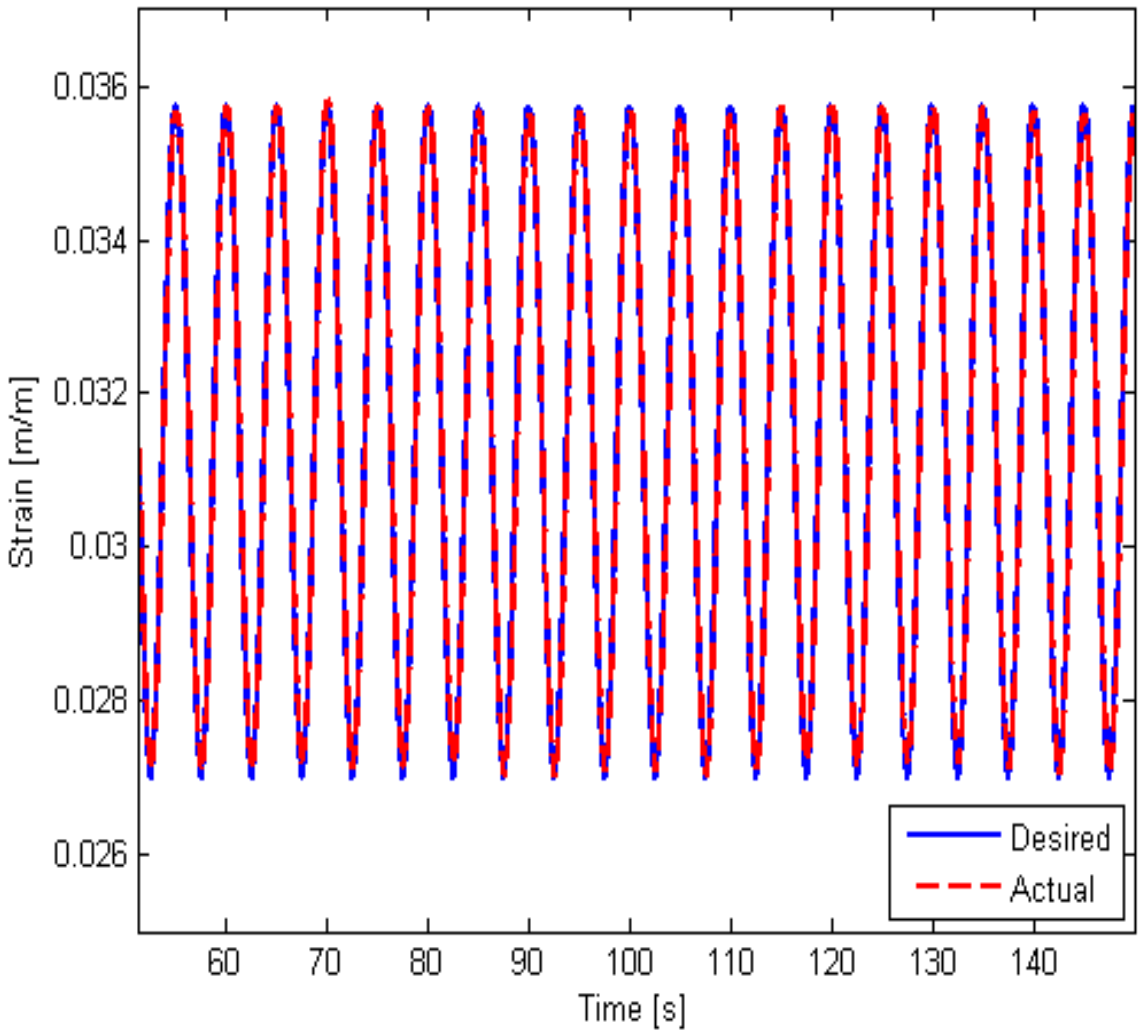

Figure 6.47: Tracking response with sinusoidal stress disturbance and PI + Inverse control, without force sensor

the PI + Inverse controller is due to the additional contribution of the inverse controller as shown by the commanded currents in Figure 6.49. Nevertheless, the figures shows the PI + Inverse controller is more robust than the PI controller for sinusoidal disturbances. This is further shown by the error norms for the two control architectures (Table 6.10).

Table 6.10: Error two-norms, sinusoidal disturbance, without force sensor

controller	*Error norm*
PI	0.0623
PI+Inv	0.0196

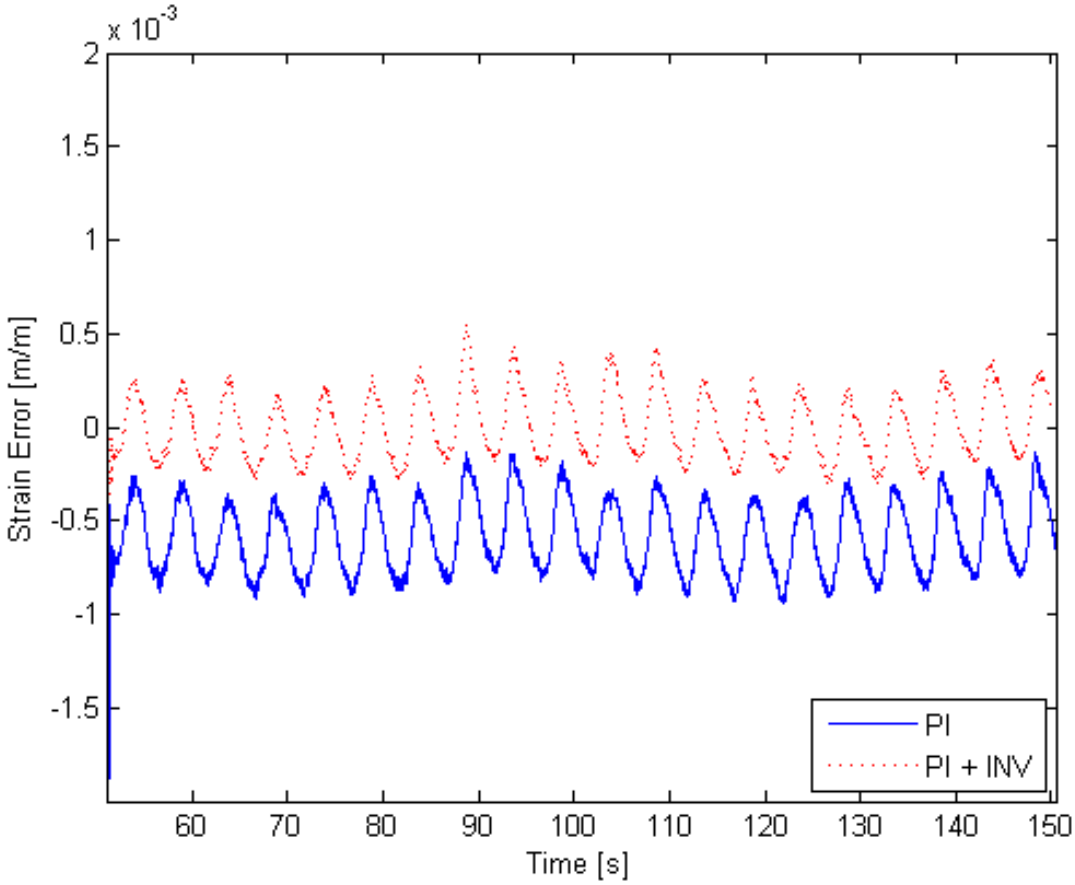

Figure 6.48: Tracking error for sinusoidal stress disturbance, without force sensor

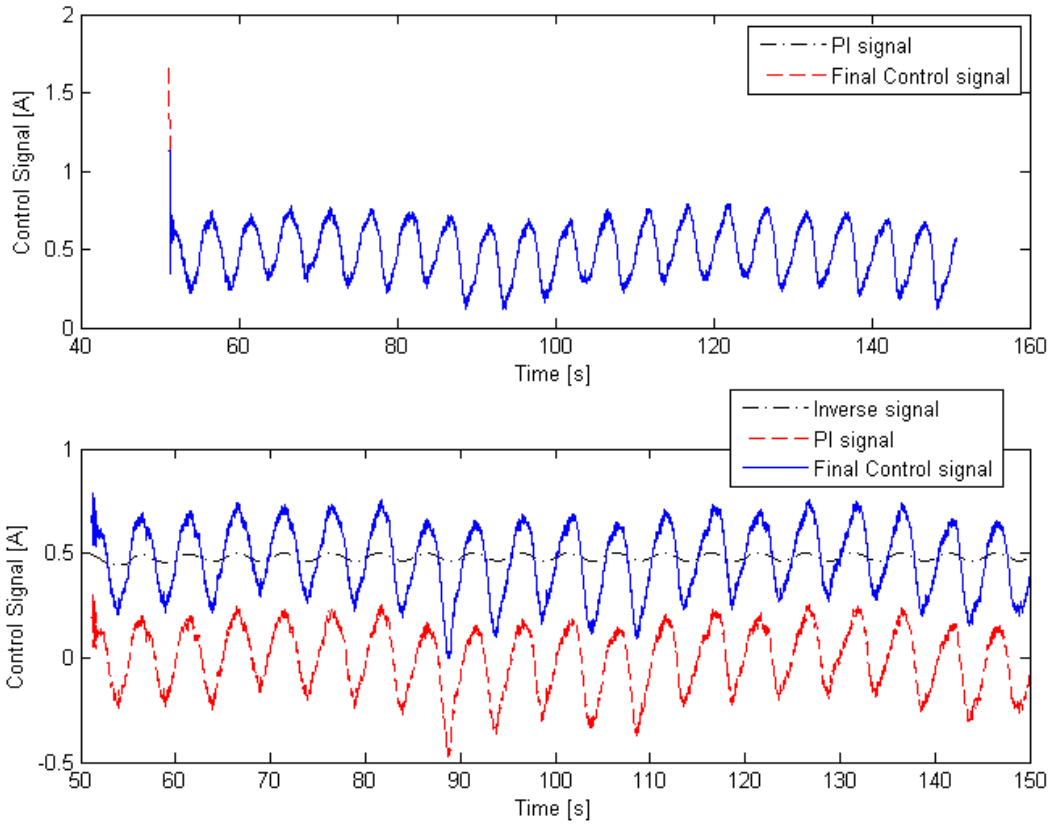

Figure 6.49: Control signal for sinusoidal stress disturbance, without force sensor

6.6.3 Summary of Results

Tables 6.11 and 6.12 show a summary of the error two-norms for impulse and sinusoidal disturbances with the different control architectures and controllers. In all cases it is observed that the PI + Inverse controller performs at least as well as, or better than the PI controller. The performance of the PI controllers with or without a force sensor is very similar. This is expected since PI control depends only on the desired and measured strain, therefore additional knowledge about stress disturbance does not improve the performance of the system. On the other hand, prior knowledge of impulse disturbance signals improves the tracking performance when using PI + Inverse control. This is expected because the inverse controller can cancel out the disturbances on the system. For the sinusoidal disturbances, however, the performance of the PI + Inverse controllers are similar for both systems with and without a force sensor. The inverse controller in particular performs differently with prior knowledge especially when the disturbance is decreasing, but it appears that the contribution of the inverse controller does not affect the performance of the system by a large amount. One can conclude therefore that in the presence of sinusoidal disturbances, a force sensor does not add to the robustness of the system.

Table 6.11: Error two-norms for impulse disturbance, summary:

	Controller	51-62.5 s	55.5-56.5 s	59.5-60.5 s
Without prior disturbance knowledge	PI	0.0114	0.0092	0.0095
	PI+INV	0.0085	0.0044	0.0087
With prior disturbance knowledge	PI	0.0134	0.0112	0.0120
	PI+Inv	0.0024	0.0024	0.0036

Table 6.12: Error norms for sinusoidal disturbance, summary:

	Controller	0.02Hz
Without prior disturbance knowledge	PI	0.0623
	PI+INV	0.0196
With prior disturbance knowledge	PI	0.0613
	PI+INV	0.0195

Chapter 7

Conclusions and Future Work

7.1 Conclusions

Shape memory alloy materials display characteristics that make them promising choices for use as actuators due to properties such as high energy density, smooth and silent actuation, reduced part counts compared to traditional alternatives, and scalability down to the micromechanical level.

The use of shape memory alloys as actuators, for example in positioning systems, requires the development of good models to predict and control the highly non-linear behaviour inherent in the materials.

In this work, a new phenomenological model was developed that is able to predict phase kinetics and mechanical behaviour of SMA. The model is invertible and gives good accuracy and robustness when used for control in conjunction with a PI controller. Novel contributions include the extension of Madill's model to include time varying stresses in the phase kinetics

portion of the model and the development of a mechanical model to describe SMA stress-strain behaviour. The SMA is modelled in literature as either having parallel layers or series layers. The mechanical model presented here predicts the stress and strain behaviour of the SMA including loading, unloading, reloading and cycling. Both the parallel and series SMA models are developed and compared. Additionally, experimental results were shown to validate the model. Experimental validation shows that the parallel model is more accurate in describing SMA behaviour than the series model.

The model is then inverted in order to use it in feedforward model-based control. Although model inversion is presented for both the parallel and series mechanical models, the parallel model is used in the control experiments because it better describes observed SMA behavior. The control portion of this work investigates the performance of using classical control with a PI feedback controller and using the inverse model in feedforward in conjunction with PI feedback. The reference strain signals used are a step signal and sinusoidal signals. The control experiments are performed on two different plants: an SMA actuator with constant force and a spring-biased SMA actuator. The ability of the inverse model + PI controller to deal with time-varying stresses different from spring stresses is also demonstrated. Additionally, disturbance rejection to sinusoidal and impulse stress disturbances is investigated.

The results show that for tracking control the PI+Inverse controller performs better than the PI controller in all cases. The worst errors are observed in the cooling sections due to the lack of an active cooling mechanism and, in the case of sinusoidal tracking, the high frequency of the reference signal. The errors in the cooling sections are accumulated by the integrator, therefore integrator gains have to be low. The cost of having low integrator gains is steady-state tracking error. This error is worse when using only the PI controller because

the additional contribution of the model-inverse controller when using a PI + Inverse controller compensates for the error. The system is more robust in the presence of stress disturbances when using a PI + Inverse controller as opposed to only a PI controller. In the presence of a force sensor, the performance of the PI control is not affected, as expected. However, when using PI + Inverse control on a system exposed to transient disturbances like impulses, the presence of a force sensor improves the tracking performance of the system which is expected since the inverse controller has prior knowledge of the stress profile with the disturbance and is therefore able to compensate for it. However, when the system is exposed to sinusoidal disturbances, the presence of a force cell does not improve tracking performance. Additionally, the benefits of having a feedforward model-inverse controller can be lost due to saturation of the control signal in order to protect the wire especially when the system is exposed to transient stress disturbances in the absence of a force sensor. One can conclude that the PI + Inverse controller gives better tracking performance and improves robustness of the system. However, in situations where a high degree of precision is not required, the PI controller performs relatively well such that the added complexity of an inverse controller may not be justified.

7.2 Future Work

The next steps for continuation of the model and control of SMA can be summarised as follows:

- Refine the heating model to include varying parameters such as the resistance of the wire, the convection coefficient and the heat capacity.
- With accurate temperature measurement, for example an oven, a collection of high tem-

perature pulls at various temperatures in order to have a complete set of experimental data for stress-strain behaviour of SMAs. Furthermore, temperature measurement will allow for extraction of accurate parameters for the phase kinetics model and validation of the inverse phase kinetics and inverse mechanical models.

- Inclusion of an active-cooling mechanism into the experiment to improve tracking.
- Analysis of stability of the closed-loop system with the inverse-model based controller.
- Analysis of other possible control architectures with inverse-model based control to improve tracking performance and robustness.

Chapter 8

List of Symbols

Symbol	*Description*	*Units*
ϵ	strain	m/m
σ	stress	Pa
σ_s	stress at point of temperature switch from heating to cooling or vice versa	Pa
σ_s^+	stress right after temperature switch	Pa
σ_s^-	stress right before temperature switch	Pa
t	time	sec
t_s	time at which temperature switches from heating to cooling or vice versa	sec
T	temperature	^oC
T_∞	ambient temperature	^oC
θ	temperature difference from ambient	^oC
θ_s	θ at point of temperature switch from heating to cooling or vice versa	^oC
θ_s^+	θ right after temperature switch	^oC
θ_s^-	θ right before temperature switch	^oC
R_m	martensite fraction ($0 \leq R_m \leq 1$)	unitless
$R_m(\theta, \sigma)$	R_m at a given temperature and stress	unitless
R_m^H	martensite fraction during heating	unitless
R_m^C	martensite fraction during cooling	unitless
P	power	W

Symbol	*Description*	*Units*
D_{SMA}	diameter of wire	m
L_{SMA}	length of wire	mm
k_m^C	cooling constant, slope of cooling portion of hysteresis	^oC
k_m^H	heating constant, slope of heating portion of hysteresis	^oC
h	convection heat transfer coefficient	W/m^2K
R	resistance	Ω
C	specific heat capacity	$J/kg.^oC$
A	wire surface area	m^2
V	wire volume	m^2
ρ	material density	kg/m^3
c_m	stress shifting factor, inverse of stress-temperature slope	$^oC/MPa$
M_s	martensite start temperature	^oC
M_f	martensite finish temperature	^oC
A_s	austenite start temperature	^oC
A_f	austenite finish temperature	^oC
β^H	effect of stress on heating hysteresis curve	^oC
β^C	effect of stress on cooling hysteresis curve	^oC
α^H	scaling factor of heating hysteresis curve	unitless
α^C	scaling factor of cooling hysteresis curve	unitless
α_p	scaling factor of hysteresis curve, stress independent portion of α	unitless
γ^H	vertical translation factor of heating hysteresis curve	unitless
γ^C	vertical translation factor of cooling hysteresis curve	unitless
α_{inv}^H	scaling factor of heating hysteresis curve (inverse model)	unitless
α_{inv}^C	scaling factor of cooling hysteresis curve (inverse model)	unitless
α_p^{inv}	stress independent portion of α_{inv}	unitless
γ_{inv}^H	vertical translation factor of heating hysteresis curve(inverse model)	unitless
γ_{inv}^C	vertical translation factor of cooling hysteresis curve(inverse model)	unitless
E_a	austenite elastic modulus	Pa
E_{am}	austenite plateau modulus	Pa
E_{sim}	stress induced martensite modulus	Pa
E_m	twinned martensite modulus	Pa
E_t	martensite plateau modulus	Pa
E_d	detwinned martensite modulus	Pa
E_a^{un}	austenite unload modulus	Pa
E_{sim}^{un}	stress induced martensite unload modulus	Pa
E_d^{un}	detwinned martensite unload modulus	Pa

Symbol	*Description*	*Units*
E_{mart}	martensite modulus	m/m
E_{aust}	austenite modulus	m/m
E_{rel}	reloading modulus	m/m
σ_{y1}^{M}	first yield stress of martensite	Pa
σ_{y2}^{M}	second yield stress of martensite	Pa
ϵ_{y1}^{M}	first yield strain of martensite	m/m
ϵ_{y2}^{M}	second yield strain of martensite	m/m
σ^{Ms}	first yield stress of austenite	Pa
σ^{Mf}	second yield stress of austenite	Pa
σ^{Af}	stress at which SMA is fully austenite after pseudoelastic unloading	Pa
σ^{As}	stress at which austenite begins forming after unloading from SIM	m/m
ϵ^{Ms}	first yield stress of austenite	m/m
ϵ^{Mf}	second yield stress of austenite	m/m
ϵ^{Af}	strain at which SMA is fully austenite after pseudoelastic unloading	m/m
ϵ^{As}	strain at which austenite begins forming after unloading from SIM	m/m
ϵ_{off}	strain offset(indicator of location in stress-strain curve)	m/m
σ_{off}	stress offset(indicator of location in stress-strain curve)	m/m
ϵ_{off}^{M}	martensite strain offset	m/m
σ_{off}^{M}	martensite stress offset	m/m
ϵ_{off}^{A}	austenite strain offset	m/m
σ_{off}^{A}	austenite stress offset	m/m
σ_m	martensite stress in parallel model	m/m
σ_a	austenite stress in parallel model	m/m
ϵ_m	martensite stress in series model	m/m
ϵ_a	austenite stress in series model	m/m
ϵ_{un}	strain at point of unloading	m/m
σ_{un}	stress at point of unloading	m/m
ϵ_{un}^{M}	martensite strain at point of unloading	m/m
σ_{un}^{M}	martensite stress at point of unloading	m/m
ϵ_{un}^{A}	austenite strain at point of unloading	m/m
σ_{un}^{A}	austenite stress at point of unloading	m/m
R_m^D	twinned martensite fraction	unitless
R_m^D	detwinned martensite fraction	unitless
R_{SIM}	stress induced martensite fraction	unitless
R_a	pure austenite fraction	unitless

Symbol	*Description*	*Units*
ϵ_{int}	strain at which unloading cuve joins pseudoelastic unloading loop	m/m
σ_{int}	strain at which unloading cuve joins pseudoelastic unloading loop	m/m
ϵ_{rel}	strain at point of reloading	m/m
σ_{rel}	stress at point of reloading	m/m
ϵ_{rel}^{M}	martensite strain at point of reloading	m/m
σ_{rel}^{M}	martensite stress at point of reloading	m/m
ϵ_{rel}^{A}	austenite strain at point of reloading	m/m
σ_{rel}^{A}	austenite stress at point of reloading	m/m
ϵ_{rl}	strain at which point the reloading curves rejoin loading curves	m/m
σ_{rl}	stress at which point the reloading curves rejoin loading curves	m/m
ϵ_{rl}'	strain reload limit for reloading withing pseudoelastic loop	m/m
σ_{rl}'	stress reload limit for reloading withing pseudoelastic loop	m/m
ϵ_{mart}	martensite strain from inverse model	m/m
ϵ_{aust}	austenite strain from inverse model	m/m
σ_{mart}	martensite stress from inverse model	m/m
σ_{aust}	austenite stress from inverse model	m/m
ϵ_{SMA}	strain in SMA in spring-biased SMA actuator	m/m
ϵ_{s}	strain in spring in spring-biased SMA actuator	m/m
ϵ_{tot}	total strain or prestrain in spring-biased SMA actuator	m/m
E_{SMA}	modulus of SMA in spring-biased SMA actuator	Pa
E_{s}	modulus of spring in spring-biased SMA actuator	Pa
ϵ_{ref}	reference strain	m/m
F_{SMA}	force exerted by an SMA	N
F_{spring}	force exerted by a spring	N
k_{spr}	spring constant	N

Bibliography

[1] A. J. Zak, M. P. Cartmell, W. M. Ostachowicz, and M. Wiercigroch, "One-dimensional shape memory alloy models for use with reinforced concrete structures," *Smart Materials and Structures*, vol. 12, pp. 338–346, 2003.

[2] S. Seelecke and I. Muller, "Shape memory alloy actuators in smart structures: Modeling and simulation," *Applied Mechanics Review*, vol. 57, no. 1, pp. 23–46, 2004.

[3] L. C. Brinson and M. S. Huang, "Simplifications and comparisons of shape memory alloys constitutive models," *Journal of Intelligent Materials Systems and Structures*, vol. 7, pp. 108–114, 1996.

[4] D. R. Madill, "Modelling and stability of a shape memory alloy position control system," Master's thesis, University of Waterloo, 2004.

[5] K. Otsuka and C. Wayman, *Shape Memory Materials*. Cambridge Univeristy Press, 1998.

[6] J. A. Shaw and S. Kyriakides, "Thermomechanical aspects of NiTi," *Journal of Mechanics and Physics of Solids*, vol. 43, no. 8, pp. 1243–1281, 1995.

[7] DynalloyInc, "Introduction to nitinol," Available at http://www.dynalloy.com/AboutNitinol.html, April 2007 [date accessed].

[8] T. Deurig, K. Melton, D. Stöckel, and C. Wayman, *Engineering Aspects of Shape Memory Alloys*. Butterworth-Heinemann, 1990.

[9] L. M. Toews, "The development of a monolithic shape memory alloy actuator," Master's thesis, University of Waterloo, 2004.

[10] D. Madill and D. Wang, "Modelling and L2-stability of a shape memory alloy position control system," *IEEE Trans. Control Systems Technology*, vol. 6, no. 4, pp. 473–481, 1998.

[11] K. Ikuta, M. Tsukamoto, and S. Hirose, "Mathematical model and experimental verification of shape memory alloy for designing micro actuator," in *Proceedings of IEEE MEMS*, Nara, Japan, 1991, pp. 103–108.

[12] Mueller, Musolff, and Sahota, "Analysis of deformation, temperature, load and resistivity," Available at http://www.smaterial.com//SMA/research/analysis/analysis.html, February 2007 [date accessed].

[13] R. Velazquez, M. Hafez, J. Szewczyk, and E. Pissaloux, "Experimental and computational thermomechanical study of a shape memory alloy micro actuator: aspects of antagonist-type behaviour," in *3rd MIT conference on Computational Fluid and Solid Mechanics*, Boston, USA, May 2005, pp. 103–108.

[14] L. C. Brinson, "One-dimensional constitutive behaviour of shape memory alloys: thermomechanical derivation with non-constant material functions and redefined martensite inter-

nal variable," *Journal of Intelligent Materials Systems and Structures*, vol. 4, no. 2, pp. 229–242, 1993.

[15] D. Wolons, F. Gandhi, and B. Malovrh, "Experimental investigation of the pseudoelastic hysteresis damping characteristics of shape memory alloy wires," *Journal of Intelligent Materials Systems and Structures*, vol. 9, no. 2, pp. 116–126, 1998.

[16] J. A. Shaw, "A thermomechanical model for a 1-d shape memory alloy wire with propagating instabilities," *International Journal of Solids and Structures*, vol. 39, pp. 1275–1305, 2002.

[17] M. Achenbach, "A model for an alloy with shape memory," *International Journal of Plasticity*, vol. 5, no. 4, pp. 371–395, 1989.

[18] J. Boyd and D. Lagoudas, "A thermodynamical constitutive model for shape memory materials. part 1. the monolithic shape memory alloy," *International Journal of Plasticity*, vol. 12, no. 6, pp. 805–842, 1996.

[19] Y. Ivshin and T. J. Pence, "Thermomechanical model for a one variant shape memory material," *Journal of Intelligent Materials Systems and Structures*, vol. 5, pp. 455–473, 1994.

[20] B. Goo and C. Lexcellent, "Micromechanics-based modeling of two-way memory effect of a single crystalline shape memory alloy," *Acta Materialia*, vol. 45, no. 2, pp. 727–737, 1997.

[21] L. Brinson and R. Lammerung, "Finite element analysis of the behavior of shape memory alloys and their applications," *International Journal of Solids and Structures*, vol. 30, no. 23, pp. 3261–3280, 1993.

[22] J. J. Amalraj, A. Bhattacharyya, 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, no. 5, pp. 622–631, 2000.

[23] F. Preisach, "Uber die magnetische Nachwirkung," *Zeitschrift fuer Physik*, vol. 94, pp. 277–302, 1935.

[24] K. Tanaka, S. Kobayashi, and Y. Sato, "Thermomechanics of transformation, pseudoelasticity and shape memory effect in alloys," *International Journal of Plasticity*, vol. 2, no. 11, pp. 59–72, 1986.

[25] C. A. Rogers and C. Liang, "One-dimensional thermomechanical constitutive relations for shape memory materials," *Journal of Intelligent Material Systems and Structures*, vol. 8, no. 4, pp. 285–302, 1997.

[26] S. Sukhatme, *A Textbook on Heat Transfer*. University Press, 1996.

[27] J. McNichols and J. Cory, "Thermodynamics of nitinol," *Journal of Applied Physics*, vol. 61, no. 3, pp. 972–984, 1987.

[28] S. Dutta and F. Ghorbel, "Differential hysteresis modeling of a shape memory alloy wire actuator," *IEEE/ASME Transactions on Mechatronics*, vol. 10, pp. 189–197, 2005.

[29] G. Song, N. Ma, and H. Lee, "Position control of shape memory alloy actuators with internal electrical resistance feedback using neural networks," *Smart Materials and Structures*, vol. 13, pp. 777–783, 2004.

[30] Mondotronics, "Muscle wires," Available at http://www.robotstore.com, January 2006 [date accessed].

[31] D. Reynaerts and H. Van Brussel, "Shape memory alloy based electrical actuation for robotic applications," in *SMST-94: The First International Conference on Shape Memory and Superelastic Technologies*, Pacific Grove, CA; USA., 1995, pp. 271–276.

[32] M. Hashimoto, M. Takeda, H. Sagawa, I. Chiba, and K. Sato, "Application of shape memory alloy to robotic actuators," *Journal of Robotic Systems*, vol. 2, no. 1, pp. 3–25, 1985.

[33] N. Troisfontaine, P. Bidaud, and M. Larnicol, "Optimal design of micro-actuators based on sma wires," *Smart Materials and Structures*, vol. 8, pp. 197–203, 1999.

[34] K. Arai, S. Ararnaki, and K. Yanagisawa, "Feedback linearization for SMA (shape memory alloy)," in *SICE*, Sapporo, 1995, pp. 1383–1386.

[35] J. P. Briggs and J. P. Ostrowski, "Experimental feedforward and feedback control of a one-dimensional SMA composite," *Smart Materials and Structures*, vol. 11, no. 1, pp. 9–23, 2002.

[36] G. Song, V. Chaudhury, and C. Batur, "Precision tracking control of shape memory alloy actuators using neural networks and sliding-mode based robust controller," *Smart Materials and Structures*, vol. 12, pp. 223–231, 2003.

[37] S. Majima, K. Kodama, and T. Hasegawa, "Modeling of shape memory alloy actuator and tracking control system with the model," *IEEE Transactions Control Systems Technology*, vol. 9, no. 1, pp. 54–59, 2001.

[38] G. V. Webb, D. C. Lagoudas, and A. J. Kurdila, "Hysteresis modeling of sma actuators for control applications," *Journal Of Intelligent Material Systems And Structures*, vol. 9, no. 6, pp. 432–448, 1998.

[39] G. Tao and P. V. Kokotovic, "Adaptive control of plants with unknown hysteresis," *IEEE Transactions Automatic Control*, vol. 40, pp. 200–212, 1995.

[40] C. A. Dickinson and J. T. Wen, "Feedback control using shape memory alloy actuators," *Journal Of Intelligent Material Systems And Structures*, vol. 9, no. 4, pp. 242–250, 1998.

[41] D. Grant and V. Hayward, "Variable structure control of shape memory alloy actuators," *IEEE Control Systems Magazine*, vol. 17, no. 3, pp. 80–88, 1997.

[42] M. Elahinia and H. Ashrafiuon, "Nonlinear control of a shape memory alloy actuated manipulator," *ASME, Journal of Vibration and Acoustics*, vol. 124, pp. 566–557, 2002.

[43] Y. Eren, C. Mavroidis, and J. Nikitczuk, "B-Spline based adaptive control of shape memory alloy actuated robotic systems," in *ASME International Mechanical Engineering Congress, Exposition*, Louisiana, 2002, pp. 1–8.

[44] G. Song and D. Quinn, "Robust tracking control of a shape memory alloy wire actuator," in *Control of vibration and noise: New millenium; Proceedings of the Symposium from the 2000 International Mechanical Congress and Exposition*, Orlando, FL; UNITED STATES, 2000, pp. 9–15.

[45] D. Hughes and J. Wen, "Preisach modeling and compensation for smart material hysteresis," in *Proceedings of the SPIE - The International Society for Optical Engineering*, vol. 2427, College Station, TX, USA.

[46] A. Kumagai, T.-I. Liu, and P. Hozian, "Control of shape memory alloy actuators with a neuro-fuzzy feedforward model element," *Journal of Intelligent Manufacturing*, vol. 17, no. 1, pp. 45–56, 2006.

[47] R. B. Gorbet and K. A. Morris, "Closed-loop position control of Preisach hystereses," *Journal Of Intelligent Material Systems And Structures*, vol. 14, no. 8, pp. 483–495, 2003.

[48] B. Selden, C. Kyujin, and H. Asada, "Segmented shape memory alloy actuators using hysteresis loop control," *Smart Materials and Structures*, vol. 15, no. 2, pp. 642–652, 2006.

[49] P. L. Potapov and E. P. Da Silva, "Time response of shape memory alloy actuators," *Journal Of Intelligent Material Systems And Structures*, vol. 11, pp. 125–134, 2000.