A Nonlinear Thermodynamic Model for Phase Transitions in Shape Memory Alloy Wires

by

Daniel Ryan Reynolds

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE

Doctor of Philosophy

APPROVED, THESIS COMMITTEE:

Petr Klouček, Chairman
Assistant Professor of Computational and Applied Mathematics

William W. Symes
Noah Harding Professor of Computational and Applied Mathematics

Chad M. Landis
Assistant Professor of Mechanical Engineering and Materials Science

Dennis D. Cox
Professor of Statistics

Houston, Texas
May, 2003
Abstract

A Nonlinear Thermodynamic Model for Phase Transitions in Shape Memory Alloy Wires

by

Daniel Ryan Reynolds

Through a mathematical and computational model of the physical behavior of shape memory alloy wires, this study shows that localized heating and cooling of such materials provides an effective means of damping vibrational energy. The thermally induced pseudo-elastic behavior of a shape memory wire is modeled using a continuum thermodynamic description based on an improved Landau-Devonshire potential. Our construction of the potential function allows the model to account for particular alloys as well as the general solid-state phase transformation, improving over traditional potentials that idealize many of the material properties or focus only on individual processes. The material’s thermodynamic response is modeled using a nonlinear conservation of momentum and a nonlinear heat equation. The heat equation introduces an inhomogeneous version of the Fourier heat flux, thereby describing the discontinuous heat flow associated with shape memory materials more thoroughly than standard, continuous heat dissipation mechanisms do. This continuum thermodynamic model is then solved computationally to determine the resulting state of the wire in time. Continuous time Galerkin methods and affine finite elements treat the temporal and spatial discretizations of the model, respectively. Traditional methods for solution of the resulting finite-dimensional, nonlinear, nonconvex system
of equations must introduce a significant artificial dissipation to achieve existence of solutions. The solution of the discrete system here uses a novel combination of the damped Newton method and a homotopy method for minimizing the artificial dissipation. This combination, inspired by the well-known Method of Vanishing Viscosity for the solution of scalar hyperbolic conservation laws, reduces the artificial dissipation errors introduced by traditional approaches for such nonlinear, nonconvex thermomechanical models. Computational tests show that the proposed model successfully describes the relevant physical processes inherent in shape memory alloy behavior. Further computational experiments then confirm that up to 80% of an initial shock of vibrational energy can be eliminated at the onset of a thermally-induced phase transformation.
Acknowledgements

This thesis is the culmination of a long and arduous journey – one that I have only been able to make as a result of the dedicated support I have received from so many people along the way. Therefore I must share the credit of this work with my family, advisor, research associates, committee, fellow students, friends and everyone else who supported me throughout my time here at Rice.

I would first like to dedicate this work to my family: my parents Keith and Sondra, my brothers Chris and Adam, and my soon-to-be bride Kim Foster. Throughout the years, I have always been able to depend on their stability and support. Their gentle encouragement helped me to believe in myself through even the bleakest of times.

My great thanks goes to my advisor, Dr. Petr Klouček, whose continuous flow of ideas has helped me to see mathematics as a powerful and exciting tool for describing the physical world in which we live. Throughout my studies at Rice, I could always depend on Petr to have a new question that I had not yet considered – challenging my preconceived notions and forcing me to analyze everything from an alternate view.

I would also like to particularly thank Dr. Dennis Cox and Dr. Thomas Seidman, whose ideas and approaches through our outside research have opened me up to a new world of future research interests. On the same note, I also thank the rest of my thesis committee, Dr. William Symes and Dr. Chad Landis, for their continued insights and guidance into a greater understanding of partial differential equations and thermodynamic material modeling.

The last, and possibly most supportive, group of academic support that I thank is my group of fellow students, the staff, and the faculty of the CAAM department here at Rice. In particular, I would like to thank my friend and fellow student Jennifer
Wightman and my writing instructor Dr. Jan Hewitt for their (sometimes exhausting) help on proofreading this thesis. I thank my fellow students Eric Dussaud, Fernando Guevara Vasquez, Agatha Comas, Chris Husband, Zhen Wang, Shannon Scott and Heidi Thornquist for their help and comradery throughout our graduate learning process. I would like to thank the staff of the CAAM department: Daria Lawrence, Ginger Wright and Fran Moshiri, for all of their help clearing my way through the jungle of regulations, deadlines, and the general bureaucracy of research life. Finally, I express my tremendous thanks to the faculty of the CAAM department here at Rice. I knew nearly nothing about the world of applied mathematics when I walked in here five years ago. Thanks to you, I now realize that mathematics is truly a process; I look forward to finding my place in this process in the years to come.

However, life is more than academics. A second dedication of this thesis goes to my friends, who, just like my family, have supported me through all of my graduate career. I would like to thank Mike Jeffers, Winston Denny, George Irwin, Ben Archey, Blair Christian, Rudy Fink, Lori Hellkamp, Matthew Burch, Rita Siemion, Edmund Shafer Hall, Robert Kühnbach, Jonathan Smith, Nathan Day and Joe Hankins for all of their support in distracting me and keeping life normal.

Lastly, I would like to thank the following people and groups that have proved invaluable to me in the last five years. First, I would like to thank Dr. Pepper for all of their stimulating help during the research and writing process. I'd like to thank The Onion for helping me to keep a larger, more humorous perspective even when I was firmly rooted in the minute details of programming, research and writing. Finally, I would like to thank Isaac Brock, Eric Judy and Jeremiah Green for my theme song “neverending math equation”, as well as Black Francis, Kim Deal, David Lovering, Joey Santiago, Ira Kaplan, Georgia Hubley, James McNew and many others for all of their lyrical encouragement these past years.

Here's to all of you. I thank you from the bottom of my heart.
Contents

Abstract ii

Acknowledgements iv

List of Figures x

List of Tables xii

1 Introduction 1
  1.1 Dynamic Modeling of Phase Transitions ................. 1
  1.2 Current Approaches ....................................... 2
  1.3 Our Modeling Approach ..................................... 3
  1.4 Overview of Results ....................................... 4

2 Modeling Background 6
  2.1 History of Shape Memory Alloys ......................... 7
  2.2 First-Order Phase Transitions ......................... 8
    2.2.1 Hysteresis ........................................... 9
    2.2.2 Twinning ............................................ 11
    2.2.3 Shape Memory Effect ............................... 11
    2.2.4 Superelasticity ...................................... 12
    2.2.5 Latent Heat ......................................... 14
  2.3 Current and Potential Applications .................. 14
2.4 Linearized Modeling Approaches ............................................. 16
2.5 Nonlinear Modeling Approaches ............................................. 18
  2.5.1 The Search for a Model .................................................. 20
  2.5.2 Computation of Microstructure ....................................... 23
  2.5.3 Dynamics Computations ................................................. 26

3 Proposed Thermodynamic Model .............................................. 29
  3.1 First-Principles Approach ................................................ 31
  3.2 Derivation of Proposed PDE Model ..................................... 35
    3.2.1 Simplification of the Balance of Energy ......................... 37
    3.2.2 Thermoelastic Materials .............................................. 37
    3.2.3 Reduction of the Energy Equation .................................. 40
    3.2.4 The Specific Heat Capacity ......................................... 40
    3.2.5 Heat Flux ............................................................... 42
  3.3 Resulting PDE system ...................................................... 45
  3.4 Helmholtz Free Energy ..................................................... 47
    3.4.1 Landau Theory .......................................................... 48
    3.4.2 Helmholtz Requirements .............................................. 50
    3.4.3 Helmholtz Construction .............................................. 53

4 Related Theoretical Properties .............................................. 63
  4.1 Strictly Hyperbolic Systems of Equations ............................. 65
    4.1.1 General Information in One-Space Dimension .................... 65
    4.1.2 Weak Solutions ........................................................ 66
    4.1.3 Method of Vanishing Viscosity .................................... 68
  4.2 Degenerate Hyperbolic Systems .......................................... 70
    4.2.1 Quasiconvexity Approaches ........................................ 72
    4.2.2 Direct Calculations of the Young Measure ....................... 76
    4.2.3 Existence and Uniqueness for Viscous Models ................... 91
4.3 Theoretical Damping and Control Properties ........................................ 97
4.3.1 Theoretical Material Damping .................................................. 98
4.3.2 Idealized Vibrational Control .................................................. 100

5 Solution Method .................................................................................. 113
5.1 Computation of Thermodynamic Quantities ...................................... 113
5.2 Continuous Space-Time Galerkin .................................................... 114
5.2.1 Time-Discretization of $g_1^n(X^{n-1}, X^n, \varphi)$ ..................... 119
5.2.2 Time-Discretization of $g_2^n(X^{n-1}, X^n, \varphi)$ ..................... 120
5.2.3 Time-Discretization of $g_3^n(X^{n-1}, X^n, \varphi)$ ..................... 122
5.2.4 The Spatial Discretization ....................................................... 125
5.3 Numerical Difficulties ..................................................................... 126
5.4 Variable Viscosity Newton Method .............................................. 131
5.5 Implementation Details ............................................................... 134

6 Computational Results ...................................................................... 138
6.1 Visualization Techniques .............................................................. 139
6.2 Thermally-Induced Phase Transformations ..................................... 142
6.2.1 Martensite to Austenite .......................................................... 142
6.2.2 Austenite to Martensite .......................................................... 146
6.3 Stress-Induced Phase Transformation ........................................... 149
6.4 Hysteresis ..................................................................................... 154
6.5 Latent Heat .................................................................................... 156
6.6 Damping Strategy ......................................................................... 157
6.6.1 Base Case ................................................................................ 159
6.6.2 Full Phase Transformation ....................................................... 161
6.6.3 Localized Phase Transformations ............................................ 165
6.6.4 "Best Case" Damping Scheme ................................................. 170

7 Conclusions and Future Work ......................................................... 175
7.1 Accuracy of the Proposed Model .................................. 175
7.2 Improved Solution Methods ....................................... 177
7.3 Current SMA and Active Damping ............................... 178
7.4 Expansion to Plates, Solids ...................................... 179
7.5 Future Materials Research ...................................... 180

Bibliography .................................................................... 181
List of Figures

2.1 Typical Hysteresis Loop for Martensitic Transformation 10
2.2 Twinning and the Shape Memory Effect 12
2.3 Superelasticity and the Corresponding Stress-Strain Curve 13

3.1 Various Scales of Microstructure 31
3.2 The Helmholtz Free Energy and the First-Order Phase Transition 49
3.3 Strain Energy Functional $W(\gamma, \theta)$ 59

6.1 Displacement and Phase Visualization Techniques 140
6.2 Deformation Gradient and Kinetic Energy Visualization Techniques 141
6.3 Temperatures: Thermally-Induced Martensite $\rightarrow$ Austenite Transformation 143
6.4 Displacements: Thermally-Induced Martensite $\rightarrow$ Austenite Transformation 144
6.5 Phases: Thermally-Induced Martensite $\rightarrow$ Austenite Transformation 145
6.6 Temperatures: Thermally-Induced Austenite $\rightarrow$ Martensite Transformation 147
6.7 Displacements: Thermally-Induced Austenite $\rightarrow$ Martensite Transformation 148
6.8 Phases: Thermally-Induced Austenite $\rightarrow$ Martensite Transformation 149
6.9 Body Forces Used for the Stress-Induced Transformations 150
6.10 Deformation Gradients: Stress-Induced Transformation 152
6.11 Phases: Stress-Induced Transformation ................. 153
6.12 Computed Hysteresis Loop ......................... 155
6.13 Evidence for Latent Heat Effects ...................... 156
6.14 Vibrational Force Used in Damping Tests ............ 158
6.15 Kinetic Energy: Vibrations Without Heating .......... 159
6.16 Displacements: Vibrations Without Heating .......... 160
6.17 Kinetic Energy: Vibrations With Full Transformation .. 162
6.18 Temperatures: Vibrations With Full Transformation .. 163
6.19 Displacements: Vibrations With Full Transformation .. 164
6.20 Phases: Vibrations With Full Transformation .......... 165
6.21 Kinetic Energy: Vibrations with Localized Transformations .. 166
6.22 Temperatures: Vibrations with Localized Transformations .. 167
6.23 Displacements: Vibrations with Localized Transformations .. 168
6.24 Phases: Vibrations with Localized Transformations .. 169
6.25 Kinetic Energy: Vibrations with “Best Case” Control .. 170
6.26 Temperatures: Vibrations with “Best Case” Control .. 171
6.27 Displacements: Vibrations with “Best Case” Control .. 172
6.28 Phases: Vibrations with “Best Case” Control .......... 173
List of Tables

3.1  Physics Notation and Corresponding Units             32
3.2  Material Constants in the Model                     34
6.1  NiTi Material Constants                             139
Chapter 1

Introduction

1.1 Dynamic Modeling of Phase Transitions

In this thesis, we propose a new construction and computational solution of a fully-coupled nonlinear model governing the thermodynamic behavior of shape memory alloy (SMA) wires. This model is then used to examine some possible control mechanisms which can be used to implement vibration damping devices using these wires.

Shape memory alloys are characterized by their unique ability to undergo a thermally-induced crystal phase transformation, known as the first-order martensitic phase transformation. This transformation induces such materials to achieve up to 8% pseudo-elastic deformation (deformation without material damage) through small changes to their temperature. As a result of such uniquely sensitive responses to small temperature variations, these materials are currently being considered for such multi-ranging tasks as micro- and nano-machinery, biomedical implants, self-repairing shielding devices, and active damping applications.

Key to the use of these materials for such applications is a thorough understanding
of the phase transformations giving rise to their unique behaviors. Physical testing and experimentation can go only so far as to explain the static and quasi-static behaviors of these materials; however, dynamic applications involving shape memory materials require a deeper understanding than these traditional experimental methods can provide. Thus, a considerable amount of recent research on shape memory alloys has been focused on the construction of mathematical models describing the behaviors of these materials.

An impediment to the development of mathematical models for shape memory alloys is that their unique behaviors arise from nonlinear elastic material responses. Therefore, traditional approaches from linear elasticity cannot apply to the case of these phase-transforming materials. As a result, mathematical approaches for modeling these materials must draw on tools from crystallography, lattice dynamics, thermodynamics, continuum mechanics and nonlinear elasticity. Therefore, models of these materials have been approached from a multitude of areas, combining to form an increasingly rich subject area. Moreover, these approaches combine to predict a material whose free-energy potential has multiple local minima, with each minimum corresponding to a stable crystal lattice configuration. The continuum thermodynamic equations based on such a free energy, therefore, result in non-elliptic, highly degenerate systems of equations for modeling phase transitions in these materials.

1.2 Current Approaches

Recent approaches for constructing these models fall into a few major categories. The first of these categories uses traditional methods of linearization and specialization to focus models on local approximations of specific processes, disregarding the broader range of shape memory alloy behavior. These approaches prove quite successful in reproducing individual behaviors of shape memory alloys, but they are unable to
describe more than the particular behavior to which they have been fit. A second category of models for these materials strives to understand the mechanism by which the crystal lattice transforms from one phase to another. This approach examines purely static models of stress states and the corresponding stress-free configurations. Research in this area considers the limiting states resulting from minimization of non-quasiconvex variational integrals, where critical points of these functionals correspond to the microstructural formations arising in the martensitic phase transformation.

A final approach toward a mathematical description of shape memory alloy behavior considers the full nonlinear thermodynamic modeling of such materials. This nonlinear, continuum-thermodynamics approach is rooted in Landau's theory of phase transitions, in which all the relevant physics of a material system may be explained through the differential structure of its free energy potential. Through use of such a free energy, these approaches attempt to describe the full range of thermal and mechanical behaviors associated with shape memory alloys. However, nonlinear models in this area must typically rely on the introduction of significant, purely artificial dissipation mechanisms in order to guarantee a solution of the resulting models. Moreover, such nonlinear thermodynamics approaches typically treat only the phenomenological properties of general shape memory alloys, idealizing the material constants and thereby disallowing their application to particular alloy systems.

1.3 Our Modeling Approach

The current work follows this last set of approaches, proposing a new one-dimensional model for describing the nonlinear thermodynamic behavior of shape memory alloy wires. The focal points of this new approach are the form of the heat flux and the construction of the free energy density for these materials, as well as a new approach for the numerical solution to the resulting fully-coupled differential system. The heat
transfer in shape memory alloy wires is modeled using a new, inhomogeneous form of the Fourier heat flux that describes the non-smooth heat flow associated with these materials, unlike the traditional Fourier law that describes smooth heat transfer processes. Additionally, the new construction of the free energy improves over earlier forms, describing both the general phenomenological behavior of the crystal phase transformation, as well as providing control over material constants such as the elastic moduli, specific heat and equilibrium configurations of the wire. Thus, since our free energy may be adapted to describe different metallic systems, it goes beyond previous constructions of this free energy that must idealize material constants. Finally, the solution method proposed in this thesis improves upon previous methods by minimizing the amount of artificial dissipation required for the numerical solution. This is due to a proposed technique for progressively lowering the dissipation, similar to the Method of Vanishing Viscosity used in the related area of scalar hyperbolic conservation laws.

1.4 Overview of Results

Using this improved thermodynamic model and solution method, computational experiments show that the new model correctly predicts all the relevant nonlinear behaviors inherent to shape memory alloy wires. We then use our model to examine the applicability of these materials for thermally-induced vibration damping. These computational experiments show that a promising thermal control for vibration damping purposes involves the formation of localized, regularly-spaced transformation regions along the length of the wire. Using such a controlling procedure, the thermally-induced crystal phase transformation achieves damping of up to 80% of the vibrational energy.

We present the above topics in the following manner. Background on the key nonlin-
ear behaviors of shape memory alloys, some of their current and possible uses, and previous work on mathematical models for their behavior is presented in Chapter 2. With this contextual background, Chapter 3 derives the proposed nonlinear thermodynamic model for these alloys, including both the new form of the inhomogeneous heat flux and the new construction of the free energy density. Theoretical properties relating to the modeling system, including results on general systems of hyperbolic conservation laws and existence theory for systems similar to the proposed model, are presented in Chapter 4. Then, Chapter 5 describes the computational solution method used in solving the model, with special attention toward the new method for reducing errors introduced by artificial dissipation mechanisms. Finally, computational results showing first the success of this model for predicting the full range of nonlinear shape memory alloy behavior, and then the specific results on thermally-activated vibration damping are provided in Chapter 6.
Chapter 2

Modeling Background

An understanding of the historical development and the unique characteristics associated with shape memory alloys is an integral part in the development of a thermodynamic model for these materials. It is this set of key properties of shape memory alloys that allows for the bounty of opportunities for future development using these materials. Therefore, any thermodynamic model of shape memory alloys should be able to reproduce these behaviors.

However, mathematical models describing physical processes in shape memory alloys have been successful in either reproducing some of these unique material behaviors for specific alloys, or in creating phenomenological descriptions of the general nonlinear behavior. However, up to now no models for even the one dimensional case of shape memory wires have been able to succeed in both of these tasks. Approaches to providing a mathematical understanding of shape memory alloys may be categorized into two main groups, based on their treatment of the inherent nonlinear elasticity of shape memory alloys. The *linearized models* used in traditional engineering are designed to fit particular materials for an individual process of interest. Often this corresponds to linear simplifications that are valid only in small temperature and
stress regions. On the other hand, the *nonlinear approaches* of mathematical physics attempt to provide a more unified theory describing the broad range of shape memory alloy behavior. These approaches strive to retain within the model the nonlinear behavior inherent in these materials, in order to provide a more predictive description of the transformation kinetics, individual processes, or even their full thermodynamic behavior. However, these approaches have both been difficult to solve, and they have had problems properly modeling the material behavior for specific alloys.

The proposed model, derived in Chapter 3, follows the path of such nonlinear approaches, ameliorating some of these modeling and solution issues. However, before discussing the modeling approach of this work, we must provide both the material and modeling context surrounding shape memory alloys. Thus, this chapter first details some of the historical context of these materials. We then discuss some of the unique nonlinear properties of these alloys from a materials science viewpoint. Once these behaviors have been investigated, we discuss some of the possible applications of shape memory alloys. After development of this material understanding, we discuss some of the linearized modeling approaches and then more fully examine the body of prior nonlinear modeling efforts.

### 2.1 History of Shape Memory Alloys

Compared to traditional engineering materials such as steel and concrete that have been well understood for over 100 years, the understanding of shape memory alloys is quite young. The first reported findings on these materials were in 1932, when Swedish scientist A. Ölander revealed the pseudoelastic behavior of the AuCd shape memory alloy [118]. When working with the CuZn alloy in 1938, Greninger and Mooradian observed the temperature-dependent appearance of a martensitic phase [126]. Soon after, Krudjumov and Khandros (1949) and Chang and Read (1951) discovered the
thermoelastic behavior of the martensite phase, the phenomenon at the heart of the shape memory effect. The field of material understanding and use of shape memory alloys received its first boon, however, in the early 1960s. Then, Buchler and his coworker at the U.S. Naval Ordnance Laboratory discovered the shape memory effect in NiTi when a bent strip of the alloy returned to its original shape upon heating from a pipe lighter [28]. They named the alloy Nitinol (Nickel Titanium Naval Ordnance Laboratory), and since then intensive investigations have been made to understand the behavior of these shape memory materials.

2.2 First-Order Phase Transitions

At the heart of the unique behaviors of shape memory alloys is the first-order phase transition. This type of phase transformation is exhibited in other systems, such as the melting of ice and the evaporation of water. However in the solid to solid transformation of shape memory alloys, this is also referred to as the thermoelastic martensitic transformation. Because of this transition, shape memory alloys change their shape, stiffness, natural wave speed, and other properties depending on their thermal and mechanical loading. We may understand the first-order phase transition as a temperature dependent crystal phase transformation. Depending on the temperature, shape memory alloys exhibit one of two different crystal structures or phases. The transition temperatures are contingent on the specific alloy, though in general these materials exhibit distinct high and low temperature states. The higher-temperature phase is called austenite, after the British scientist Sir W.C. Roberts-Austen, who discovered such a phase in 1898 for the FeC alloy [102]. The lower-temperature phase is called martensite, named after the German scientist Adolf Martens, who in 1890 studied two alloys with the same composition but with drastically different behaviors [102].
Several properties of the austenitic and martensitic forms of these alloys are quite different. When the material is in its austenite form it is usually strong and hard. In the case of NiTi, the austenitic phase behaves similarly to titanium [126]. Here, the crystal structure of the material is rotationally symmetric, having a cubic lattice structure. When the material is in its martensite form, however, it is usually considerably more soft and ductile. Again in the case of NiTi, the martensitic form has elastic properties similar to soft pewter [126]. In this phase the crystal structure is considerably less symmetric, with crystal lattice in orthorhombic or tetragonal lattice arrangements.

Although they are distinct, for certain temperature regions both phases may coexist in a shape memory alloy. At high temperatures the austenite phase prevails, and at low temperatures, the martensite phase is dominant. However, there exists a 'middle-range' of temperatures at which either or both phases may be seen, depending on the thermomechanical load history of the material. This is indicative of sudden transformations of small regions in the crystal lattice from one structure to another as the phase transformation occurs. It is this lack of smoothness in the transition between the two states that distinguishes the first-order phase transformation from other phase transformations, and gives rise to such phenomena as hysteresis, twinning the shape memory effect, superelasticity and a release of latent heat. These phenomena are discussed in the following sections 2.2.1-2.2.5.

2.2.1 Hysteresis

The first-order phase transformation occurs in the following manner. When martensite, the less symmetric phase, is heated it begins to structurally change into austenite, the highly symmetric phase, at the austenite start temperature ($A_s$). This phase change is completed at the austenite finish temperature ($A_f$). Similarly, when austen-
Figure 2.1: Typical hysteresis loop for a martensitic transformation: the phase transformation occurs from austenite to martensite along a different pathway than the transformation from martensite to austenite. The measure of this difference is called the hysteresis.

If steel is cooled it begins to change into martensite at the martensite start temperature \( (M_s) \) and finishes the transformation at the martensite finish temperature \( (M_f) \) [28]. This process may be seen in Figure 2.1.

These transition temperatures \( (A_s, A_f, M_s, M_f) \) are typically not the same. The martensite to austenite transformation occurs at a slightly higher temperature than the reverse transformation from austenite to martensite. The difference between these temperatures is called the hysteresis. Generally, this is defined as the temperature difference between when the material is 50% transformed to austenite while heating and when it is 50% transformed to martensite while cooling. This difference can be up to \( 30^\circ C \), and provides a measure of the overall energy absorbed by the material during the phase transformations [28, 62]. The actual temperatures where this process occurs can be controlled by adjusting the chemical composition of the material itself or by using various metallurgical treatments such as smelting and rolling.
2.2.2 Twinning

A second behavior unique to shape memory alloys is a geometric phenomenon known as twinning. As the alloy is cooled from the austenitic parent phase, martensitic variants are created in a twinned pattern, in which the atoms on opposite sides of a twinning plane achieve displacements with mirror symmetry [28].

This process may be understood in the following way. The crystal lattice structures of shape memory alloys strive to achieve minimal potential energy states for a given temperature [118]. As the temperature is lowered from the high temperature region, the minimum energy state of the potential changes discontinuously between a minimum corresponding to the austenitic lattice to minima corresponding to the martensitic lattice. In the austenite phase, there is only one possible crystal lattice orientation corresponding to this minimal state. However due to the lack of symmetry in the martensitic crystal, there are many possible martensitic lattice orientations. Thus, during the cooling process, many different martensitic phases are created simultaneously from the original austenitic state, all having equal energy. If at this point the temperature is again raised, the crystal lattice will return to the highly symmetric austenitic phase. A schematic of this is seen on the left of Figure 2.2.

2.2.3 Shape Memory Effect

Another property unique to shape memory alloys, and the property for which they receive their name, is the shape memory effect. After the twinning process occurs in the thermally-induced transformation, the martensite is easily deformable (Figure 2.2, bottom left). While most materials deform by slip or dislocation motion, martensite responds to stress by simply changing the orientation of its crystal lattice through movement of the twin boundaries [126]. As these stresses are applied, martensitic crystals continue to transform into the variant producing maximum strain and min-
Figure 2.2: Twinning and the Shape Memory Effect: notice the twinning as austenite transforms into martensite. The martensitic lattice is then easily deformed under stress, though no plastic deformation occurs. When heated, the deformed martensite resumes the original austenitic configuration (adapted from [126]).

imum stress until all of the martensite has achieved this state. If at this point the stresses are released, the material will remain in its deformed position indefinitely, since the deformed lattice corresponds to an equal energy configuration as the original twinned lattice. However, once the temperature increases above the austenite finish temperature, $A_f$, the crystal lattice again transforms into the single austenitic orientation. This phenomenon, called the shape memory effect, causes the material to regain the original shape it had prior to any thermomechanical deformation [126]. A schematic of this overall process may be seen in Figure 2.2.

2.2.4 Superelasticity

Shape memory alloys permit a second means of phase transformation, which depends on stresses in the material as opposed to the temperature-induced transformation
Figure 2.3: Superelasticity and the corresponding stress-strain curve: The material begins in the stable austenitic state under zero initial load. As the load increases, the lattice deforms elastically. When the stresses exceed a stress tolerance, the lattice transforms into the metastable martensitic lattice rather than suffer plastic damage.

discussed earlier. This stress-induced phase transformation occurs in the higher temperature regime, where austenite constitutes the stable material state. In this temperature range, shape memory alloy at rest in the austenitic phase may be transformed into the martensitic phase when loaded with high stresses. This process behaves similarly to the processes occurring in the crystal lattice under loading in the martensite phase. As stresses are increased, the crystal lattice transforms from the austenitic variant into the martensitic variant corresponding to maximum strain and minimum stress, rather than allow plastic effects such as slip or dislocation motion. Upon release of these large stresses, the shape memory material transforms back into the more stable austenitic phase. This process may be seen in the following Figure 2.3.
2.2.5 Latent Heat

A final phenomenon unique to the first-order phase transition is the release of latent heat [134]. The latent heat is defined as the quantity of heat that must be added to a system to transfer one molecule reversibly from one phase to another, while keeping the temperature of the system constant [82, 31]. Most often, the latent heat is considered in more well-known first-order phase transitions such as the melting of ice and evaporation of water. However, shape memory alloys experience a release of latent heat in a solid $\rightarrow$ solid phase transformation. According to Wollants, Roos and Delaey, in a thermally-induced martensitic transformation the moving interface between the phases must perform an amount of frictional work [134]. Because no external work is involved, the only interaction of this frictional work with the specimen’s surroundings is by the exchange of the latent heat of transformation. Thus from a thermodynamic point of view, this release of latent heat is a degradation of “useful work” in that it is associated with an increase in entropy, due to the conversion of potential to kinetic energy, and the subsequent conversion of kinetic to thermal energy [134, 114, 90].

Each of these phenomena associated with shape memory alloys will be examined in further detail in Chapter 6, when we test the model to determine its applicability for modeling shape memory alloy behavior.

2.3 Current and Potential Applications

Shape memory alloys exhibit an array of unique and dramatic thermodynamic properties. Through simple changes in the temperature, these alloys dramatically change their shape and elastic properties. Furthermore, by subsequent changes to the temperature the alloys may be brought back to their initial state, without experiencing any plastic damage to the material. Clearly, such mechanical sensitivity to tempera-
ture proves highly amenable to the construction of thermally-activated switches and engines for the conversion between thermal and mechanical energy.

Shape memory alloys are currently used for numerous static applications in which only the thermally-activated shape change is necessary. Some of these applications include orthodontic braces, in order to apply differing stresses on particular teeth to improve effectiveness; eyeglass frames, cell-phone antennas and undergarment support wires, which allow easy shape repair after damage; and tubular stents for blood vessels, that expand under normal body temperature to prevent arterial blockage [126, 103]. A more recent static application of shape memory alloys is that of satellite shielding. Here, shape memory alloys are combined with other shielding materials to protect satellites from high-speed space debris. After impact by these small objects, the shielding layer may be heated to restore the original, undamaged shape [100].

Shape memory alloys may also be helpful in the design of simple, thermally-activated tools. In extreme operating conditions such as space walks in orbit and rocket launches, standard mechanical tools such as fasteners and pyrotechnic tools such as booster rocket ejectors have a low, however significant, possibility of failure. Therefore fasteners and release latches made of shape memory alloys are under consideration for such tasks.

With the current scientific emphasis on micro- and nano-machinery, researchers also imagine the use of shape memory alloys for small-scale processes. As evidence of this, Oberaigner and Fischer from the University of Leoben, Austria, are considering problems such as the minimal radius of a shape memory alloy nanotube such that it will still exhibit the shape memory and superelastic effects. With the plethora of new uses for carbon nanotubes being discovered daily, such shape memory nanotubes have nearly endless possibility for these small-scale applications.

The problem investigated in this thesis, however, considers using the dynamic pro-
cesses of shape memory alloys for the conversion of mechanical to thermal energy. Using the strong sensitivity of these materials to changes in the temperature, active control of thermally-induced mechanical processes using these alloys remains an open problem. Moreover, the first-order martensitic phase transformation has been shown to increase mechanical dissipation in shape memory alloys, to be discussed in more detail in Section 4.3.1. Therefore, such dynamic applications as the active damping of vibrational energy are promising avenues of future use of these materials.

For further information on the current and possible uses of shape memory alloy wires, Ryhänen discusses their use in various biomedical applications [126]. Duerig and Pelton also survey a wide range of possible shape memory uses, including free recovery, constrained recovery, actuators and superelastic applications [45]. A final reference for possible uses of these materials in space applications is given by Schetky [127].

However, in order to take advantage of this multitude of current and possible applications for shape memory alloys, a deeper understanding of their transformation properties and dynamics is essential. Many researchers in both the materials science and mathematics communities have attempted to provide a mathematical understanding of these processes. Some of the previous approaches are discussed in the following Sections 2.4 and 2.5.

### 2.4 Linearized Modeling Approaches

Linearized thermodynamics computations for shape memory alloys were first attempted by Bruno, Leo and Reich in 1995 [27]. This observation-based model was designed to reproduce some of the experimental behaviors in shape memory wires, such as the thermally-induced phase transition and shape memory effect. At the same time, Likhachev constructed a similar observation-based model for these wires [98].
These two models provided the beginnings of computational efforts to reproduce the thermodynamic behaviors of shape memory alloy materials; however, their approach attempted to fit observed behaviors using rate equations, and neither considered the material physics giving rise to such behaviors.

More recently, the models of Lagoudas, Bo and Qidwai [92, 1996], Bekker, Brinson and Issen [21, 1998], Oberaigner, Tanaka and Fischer [131, 117, 116, 1996-1999], and Oberaigner, Antretter, Fischer and Tanaka [115, 2002] considered linearized models of the dynamics to describe shape memory alloy behavior. These models use basic linear wave (and sometimes heat) equations having closed-form solutions to model the (thermo)dynamics of semi-infinite shape memory alloy rods. Due to these linearizations of the dynamics, such approaches incorporate the notion of a volume fraction of the martensite phase in the wire to account for some of the nonlinear behaviors in these materials. Constitutive, observation-based rate equations are added to govern the kinetics of these volume fractions. Additionally, extra kinetic and phase rate equations are added to handle further nonlinear behaviors that the simplifications models leave out, such as the transformation stress threshold and latent heats of transformation.

These approaches have proven very effective for describing particular shape memory alloy behaviors in small stress and temperature ranges, since they are both highly tunable and allow for efficient computation. However, their applicability is limited by a variety of factors. The foremost of these is that the inherently nonlinear behaviors of shape memory alloys cannot be sufficiently described in larger stress and temperature ranges through use of these models. Such approaches average out small-scale nonlinear behaviors, linearizing the behaviors of both the wave and heat equations on the macroscale. Due to this simplification, all of the nonlinear thermodynamic behaviors must then be added into these models using rate equations consisting of averaged versions of the small-scale effects. This has the unfortunate result that these models
can reproduce only a few relevant behaviors. Most often these approaches are limited to the stress-induced martensitic phase transformation, allowing them to describe the superelastic behavior but not the first-order phase transition or hysteretic behavior of shape memory alloys. A further limitation of these approaches is that the use of linear dynamics dictates that the elasticities of the austenite and martensite phases must be treated as equal, when in fact we have seen in Section 2.2 that most often the austenite state is considerably stiffer than the martensite state. The final limitation of these approaches is that the resulting dynamics are reduced to decoupled systems of ODE rate equations. Thus, interaction of the deformation and temperature is allowed only in an alternating fashion between the wave and heat equations through a linear dependence on the time-rate-of-change of the volume fraction.

However, the major benefit of these models remains. Due to the simplified form of such models, their solution is available through standard means such as Green’s function solutions and methods of characteristics. Furthermore, due to their specialization toward individual processes associated with shape memory alloys, they accurately reproduce these processes through adjustments to a wide range of fitting parameters. Therefore, current research approaches into the design of applications using shape memory alloy wires, as discussed in Section 2.3, exclusively use these simplified models to provide computational insight into shape memory alloy processes.

2.5 Nonlinear Modeling Approaches

Alternatively, the problem of modeling the behaviors of shape memory alloys has also been approached from a more fundamental side of mathematical physics. These approaches retain the nonlinear structure of the material, attempting to allow for a single, unified theory to provide at least a phenomenological description of multiple shape memory behaviors for a broad range of temperatures and stresses. Compared
to the linearized models discussed above, these nonlinear approaches are "simpler" in the sense that the full temperature range may be considered and extra rate equations need not be added, since the nonlinearities of the material behavior are retained in the model. The main limitations of these nonlinear approaches, however, are that, it is notably more difficult to construct a single model that encompasses the full range of shape memory behaviors, and the resulting models prove to be considerably more difficult to solve than the linearized models discussed above. As a result, these models have not typically been tunable to particular alloys, and the models have not been able to provide existence and uniqueness of solutions except through the introduction of an unphysical, artificial material viscosity (see [9, 10, 11, 59]).

At the heart of such continuum-thermodynamics approaches is a quantity called the free energy of the material. This quantity was first discussed in the framework of general phase transitions by L.D. Landau in the 1940s and A.F. Devonshire in the early 1950s [42, 43, 44, 97]. They postulated that this free energy dictates a material's stable states, its elastic and thermal properties, and lies at the center of the equations of motion in a mathematical physics description for the material. The application of these principles to a nonlinear material description of the first-order phase transition in shape memory alloys was first examined by Ericksen in the mid 1980s [46, 47, 48, 49, 50]. As a result of Ericksen's work, research on the nonlinear nature of shape memory alloys behavior has proceeded along three main paths. The first of these concentrates on the development of a complete nonlinear thermodynamic model for the first-order phase transition, as well as research examining a mathematical theory for existence and uniqueness of solutions to such models. A second area of nonlinear research has focused on the investigation of martensitic microstructure development through minimization of a combination of free and boundary energies. This field considers both the average behaviors of martensitic microstructure formation and the full space of possible microstructures for these materials using various tools of pure mathematics and probability theory such as functional analysis and density
estimation. The final approach for nonlinear research in shape memory alloys concerns itself with the development of computational approaches for solving these models, in order to both understand and possibly control the thermodynamic behavior of these unique materials.

2.5.1 The Search for a Model

The first area of nonlinear research on shape memory alloys has focused on the development of nonlinear models describing their behavior. According to the seminal work by Ericksen, and indeed even according to the historical work by Landau and Devonshire, the basis of any nonlinear continuum-thermodynamical model for phase transitions is a nonconvex free energy functional [1, 42, 43, 44, 46, 47, 48, 49, 50, 97]. In the area of shape memory alloys, researchers have thus considered the Helmholtz free energy $\Psi$ (and its related strain energy $W$) to describe the first-order phase transition. According to the Landau approach, global and local minima of this potential with respect to the deformation gradient correspond to stable and metastable states at a given temperature. Other elastic material quantities such as the stiffness of stable states, and thermal properties such as the specific heat and latent heat of transformation may also be calculated from a material's potential function.

As a result of the importance of this potential, there has been considerable effort in constructing an appropriate form of the free energy for shape memory materials. One of the first of these constructions was done by Falk in 1980 [52]. He adapted the traditional Landau-Devonshire potential (see [53]) to construct a Helmholtz free energy that expresses the first-order phase transformaiton for one-dimensional materials. This construction is of the form

$$\Psi(\gamma, \theta) = \alpha_1 \gamma^6 - \alpha_2 \gamma^4 + (\alpha_3 \theta - \alpha_4) \gamma^2 + \Psi_0(\theta),$$  \hspace{1cm} (2.1)

where $\alpha_1 \rightarrow \alpha_4$ are all positive constants, $\gamma$ denotes the order parameter describing
the material phase (the deformation gradient), and $\theta$ is the absolute temperature. Although this model indeed describes the one-dimensional first-order phase transformation as a function of temperature, the four degrees of freedom limit the application of this free energy toward individual alloys. In this case, the thermal properties such as the latent heats of transformation, transformation temperatures and hysteresis cannot be fully controlled. Dynamic properties have similar difficulties: the austenite phase typically has lower stiffness than the martensite phase, and the stable martensitic strains corresponding to particular crystalline structures cannot be properly controlled (the martensite minima move with temperature, causing thermal contraction with increased temperature). Thus, although the overall behavior of the Falk potential indeed describes a first-order phase transition, its simplicity interferes with its ability to describe individual alloys.

Niezgodka and Sprekels describe the construction of a similar one-dimensional free energy functional that satisfies further restrictions from the Falk potential [114, 1988]. They propose conditions on a generalized form of Falk's potential,

$$\Psi(\gamma, \theta) = \Psi_0(\theta) + \Psi_1(\theta)\gamma^2 + \Psi_2(\gamma),$$

where a variety of constraints on $\Psi_0$, $\Psi_1$ and $\Psi_2$ provide smoothness and boundedness conditions on the overall potential. These further constraints allow theoretical results on existence and uniqueness of weak solutions, which will be discussed in more detail in Chapter 4. However, as in the case with the Falk potential, the potential constructed here is only applicable as a qualitative description of the nonlinear material behavior and cannot be applied to particular alloys.

A drastically different statistical thermodynamics approach for consideration of the one-dimensional free energy is discussed by Huo, I. Müller and Seelecke [77, 1994]. They first consider an approach using averaged particle kinetics to describe the first-order phase transition, and then consider a simplified, piecewise quadratic form of the free energy to examine various closed-form solutions. These free energies, similarly,
do not consider applicability toward individual alloys, but instead attempt to explain
the nonconvexity of the free energy from a statistical thermodynamics viewpoint.

Three-dimensional forms of the free-energy have been proposed by Ericksen and James
[34] and Falk and Konopka [54, 1993]. Both of these approaches consider the more dif-
ficult case of a large number of crystallographically-equivalent martensitic structures.
They are able to capture the material elastic properties, but suffer from difficulties
in the thermal description (e.g. zero specific heat at nonzero temperature).

However, even with the difficulties of previous free energies in fully describing various
material properties, all of these energies provide at least a phenomenological descrip-
tion for the dynamic behavior of the first-order phase transition in shape memory
alloys. Significant research has therefore been done on the solvability of models using
such nonconvex potentials. Frieselcke and Dolzmann established global existence of
solutions to the viscoelastic equation [59, 1997]

\[ u_{tt} = \text{Div} \left( \frac{\partial \Psi}{\partial (Du)} (Du) + (Du)_t \right), \]

where \( u \) constitutes the deformation, \( Du \) is the corresponding deformation gradient,
and \( \Psi = \Psi(\nabla u) \) provides the nonconvex Helmholtz free energy.

Niezgodka and Sprekels proved local-in-time existence of weak solutions to the wave-
heat system [114, 1988]:

\[ \rho u_{tt} - (\mu \gamma_t)_x - (\rho \Psi)_x - \rho f = 0 \]

\[-\rho \theta (\Psi_\theta)_t - \kappa \theta_{xx} - \alpha \kappa \theta_{xx} - \rho r - \mu \rho (\gamma_t)^2 = 0, \]

where \( u \) is the deformation, \( \gamma = u_x \) is the deformation gradient, \( \theta \) is the temperature,
\( \Psi = \Psi(\gamma, \theta) \) is the Helmholtz free energy, \( f \) provides the external body force, \( r \)
provides the external heat supply, \( \rho \) and \( \kappa \) are material constants, and the positive
constants \( \mu \) and \( \alpha \) provide dynamic and thermal viscosities.

Hoffmann and Showalter then expanded on these results to provide local-in-time
existence results for the beam-wave-heat system [72, 2000]:

\[
\begin{align*}
\text{(beam)} & \quad v_{tt} - (\mu_1 v_{xt} + T(\theta, u_x, u_{xt})v_x - Rv_{xxx})_x = f_1(x, t) \\
\text{(wave)} & \quad u_{tt} - (\Psi_{u_x}(\theta, u_x) + \mu_2 u_{xt})_x = f_2(x, t) \\
\text{(heat)} & \quad \theta(\Psi_{\theta}(\theta, u_x))_t - k(\alpha \theta_{xt} + \theta_x)_x = \mu_1 v_{xt}^2 + \mu_2 u_{xt}^2 + Rv_{xxx}^2 + f_3(x, t),
\end{align*}
\]

where the constants $\mu_1, \mu_2, \alpha, k, R$ are all positive, $v$ is the vertical displacement, $u$ is the horizontal displacement, $\theta$ is the absolute temperature, and $\Psi = \Psi(u_x, \theta)$ is the Helmholtz free energy. However, as will be discussed further in chapters 4 and 5, the results for each of these three models (Friesecke and Dolzmann, Niezgodka and Sprekels, Hoffman and Showalter) are valid only for significant, though unphysical, viscosity and capillarity terms $\mu, \alpha, \mu_1, \mu_2$, and $R$.

Thus, up to this point, models describing the nonlinear thermodynamics of shape memory alloy wires have been able to provide only a phenomenological description of such alloys. Furthermore, theoretical results pertaining to the resulting models using such potentials have been able to show existence and uniqueness of solutions only in the cases of large viscosity and capillarity terms.

### 2.5.2 Computation of Microstructure

A second branch of mathematical research into the nonlinear behavior of shape memory alloys has been in the understanding of the development of martensitic microstructures. Unlike the work discussed in Section 2.5.1, these models do not consider the actual modeling of shape memory alloys. Instead they examine the behavior of sequences that minimize the nonconvex free energies associated with these materials. Due to the nonconvexity of the free energy, along with various boundary and interface conditions, these minimizing sequences are associated with the development of martensitic microstructures. An understanding of the formation of these microstructures is an important key to discovery of macroscopic values such as martensitic
dynamic behavior that depend on various mixtures of microscale states. It is through this area of research that quantities such as the Young modulus, which may serve as a bridge between these multiple scales, can be calculated according to physical principles.

The first researchers to deal with this area of study were Ball and James in 1987 [14, 15, 16, 17, 18, 19, 20]. They proposed a method for understanding the development of martensitic microstructure in which they examined the convergence behavior of minimizing sequences \( \{u_n\}_{n=1}^{\infty} \subset W^{1,\infty}(\Omega) \) for the "energy" functional

\[
J[u] = \int_{\Omega} \mathcal{W}(\nabla u) \, du
\]

\[
u = g \text{ on } \partial\Omega,
\]

where \( \mathcal{W} \) has the non-quasiconvex, multi-well structure associated with the martensitic phase of shape memory alloys, as discussed in Section 2.5.1. In their studies, Ball and James found that the problem of determining all possible energy-minimizing deformations resulted in finer and finer oscillations of minimizing sequences at the austenite-martensite interface. This is a result of the lack of weak lower semicontinuity in the functional \( J \) at this interface. Thus, strong convergence of the iterates \( u_n \) is not possible, and only statistical convergence of the behavior of the iterates can be expected.

This statistical convergence has been further examined by Ball [14, 15, 16, 17], Dacorogna [39], Dacorogna and Marcellini [40] and others. They consider a quantity called the \textit{Young measure}, a statistical measure of the volume fraction for each material phase, to consider the convergence properties of the sequence \( \{\gamma_n\}_{n=1}^{\infty} \subset L^\infty(\Omega) \) to the function \( \gamma \in L^\infty(\Omega) \) (note: \( \gamma = \nabla u \) for \( u \in W^{1,\infty}(\Omega), \Omega \subset R^N \)). This Young measure is a Radon probability measure representing the spatial composition of the expected value \( \gamma \),

\[
\gamma(x) = \int_{R^{N\times N}} y \, d\mu_{x,\gamma}(y).
\]
However, due to the lack of weak lower semicontinuity in $J$ at the austenite-martensite interface, the sequence $\{\gamma_n\}_{n=1}^{\infty}$ converges only weakly to $\gamma$ as $n \to \infty$. On the other hand, the Young measures $\{\mu_{x,n}\}_{n=1}^{\infty}$ associated with the minimizing sequence converge strongly to the Young measure $\mu_x$ as $n \to \infty$. These concepts of Young measure convergence for the description of microstructure have been generalized to multiscale problems by Alberti and S. Müller [7, 2001]. Furthermore, as mentioned when discussing the macroscale linearized models in Section 2.4, some of the large-scale nonlinear behavior can be captured using this volume fraction of the various martensitic phases.

Due to the potential applicability of such volume fractions, this area of research has attempted to reproduce such Young moduli. One approach has been to consider minimization of the surrogate functional

$$
\mathcal{J}[u] = \int_\Omega \mathcal{W}(\nabla u) \, dx
$$

$$
u = g \text{ on } \partial \Omega,
$$

where $\mathcal{W}$ is a quasiconvexified form of the original multi-well strain energy $\mathcal{W}$ (see Pipkin [124], Ball [14, 15, 16, 17], Bhattacharya and Dolzmann [22, 23], and Milton [107]). This revised form (2.6) has the same average behavior as the original problem (2.5), without the difficulties arising from the lack of weak lower semicontinuity.

A separate approach to the production of these Young measures has been through direct calculation. Luskin was the first to attempt this in 1996 [99], but the aforementioned convergence problems with the problem (2.5) prohibited the convergence of such calculations. Further work in this direct calculation approach has been performed by Cox, Klouček and Reynolds [35, 36, 37, 38]. In this work on a one-dimensional analog of (2.5), extra energy terms are added to the functional $J[u]$ in order to force the minimization sequences to proceed past local minima by incorporating weak white noise properties into the sequence $\{u_n\}_{n=1}^{\infty}$. This approach of adding such white noise properties has, at least in one dimension, been proven to allow the
direct computation of the desired Young moduli.

These microstructure approaches have been able to consider the quasi-static formation of martensitic microstructures, which may be used in some slow-scale calculations of shape memory alloys. Yet these approaches have been unhelpful toward computations of the considerably faster thermodynamic behaviors for these materials.

2.5.3 Dynamics Computations

The third area of nonlinear research into shape memory alloys has attempted to combine the above nonlinear models from Section 2.5.1 with computational techniques for their solution. These computational efforts consider either the isothermal wave equation alone (similar to (2.2)) or the full system of nonlinear partial differential equations (as in (2.3)) to model the thermodynamic behavior of shape memory alloys. Indeed, many of the microstructure approaches from Section 2.5.2 were abstractions of the isothermal model having a wave equation with nonconvex free energy. Two types of models, however, have been examined to determine at least some of the nonlinear thermodynamical behavior of shape memory alloys and the first-order phase transformation.

The work by Abeyaratne and Knowles [3, 4, 5, 6] and Abeyaratne, Bhattacharya and Knowles [1] has primarily been concerned with the dynamics of the interface between austenitic and martensitic regions in a transforming material. They consider a trilinear stress-strain approximation of an ideal shape memory alloy wire in the isothermal setting of stress and impact-induced loading. They then use piecewise closed-form solutions of the equations of motion to consider the jump dynamics across the phase interface. With these phase interface approximations, they then analyze the movement of shock waves, predict the large damping capabilities known for these materials, and attempt to construct a physically-based constitutive function providing
a continuum-level characterization of the microscale phase transformation.

The second of these approaches to computational modeling of shape memory alloy dynamics attempts to solve dynamic models such as (2.2) and thermodynamic models such as (2.3). One of the first of these nonlinear dynamics computations solved the three-dimensional form of the nonlinear wave equation, done by Klouček and Luskin in 1994 [86]. This model used the Ericksen-James form of the free energy mentioned earlier in Section 2.5.1, and incorporated a strong viscosity into the solution to model the development of twinned martensitic structures in the In-Tl shape memory alloy. A similar three-dimensional model by Arndt, Griebel and Roubíček employs the Falk-Konopka free energy along with a strong dynamic viscosity to treat the isothermal case for a variety of shape memory alloys [12].

The three-dimensional approach was expanded to further account for thermal effects by Klouček [84, 83]; however, as mentioned in Section 2.5.1, the form of the Ericksen-James free energy limited the applicability of the model to a small region around the transformation temperature. The model again incorporated a strong viscosity to allow for computations, slowing the computed dynamics, but still providing the correct phenomenological behavior for the three-dimensional In-Tl shape memory alloy.

Melnik and Roberts have attempted the computation of three-dimensional dynamics for general classes of shape memory alloys [104, 105, 105]. Their approach is based on the Falk-Konopka representation of the free energy function. Additionally, they assume nonlocality of the constitutive variables, allowing nonlocal effects, and further incorporate both strong viscosity and capillarity effects into the model. These assumptions seem to be driven by the need for computational feasibility of the model, and thus distort the resulting computed dynamics of the alloy. Of particular interest in their approach, however, is their use of a computational method called the Centre Manifold Technique for model reduction and the eventual model solution. This tech-
nique may prove a beneficial approach to such nonlinear problems, though it relies heavily on computer algebra systems (such as Mathematica and Maple), and so may not provide the scalability necessary for large-scale design.

Although there have been various approaches toward the computational modeling of the nonlinear dynamics of shape memory alloys in one and three dimensions, all of these have been unable to fully handle the quantitative behavior of these materials. The models that these computations have been based on have been unable to allow for the small or zero material viscosity and capillarity levels seen in computational experiment, to be further discussed in Section 5.3. Often these models are based on free-energies that cannot fully describe both the dynamic and thermal nonlinear behaviors associated with shape memory alloys. However even with these quantitative difficulties, these nonlinear thermodynamics models have been able to cleanly account for the full nonlinear behavior of these remarkable materials.
Chapter 3

Proposed Thermodynamic Model

The proposed model is an expanded version of those discussed in Section 2.5.1, in that we build a nonlinear, continuum physics based thermodynamic model of shape memory alloy wires. This model improves upon the aforementioned models in a number of ways. First, the proposed model does not allow linear simplifications of the temperature response as in (2.1). Such simplifications limit the applicability of those models to a small temperature region around the transition temperature and decouple the resulting system of partial differential equations. Second, as opposed to the traditional Fourier heat flux, the proposed model incorporates an inhomogeneous form of the heat flux that accounts for phase-dependent changes in the heat transfer. A further improvement is that the model is based on a modified version of the Landau-Devonshire potential. The proposed potential surmounts the difficulties of the Falk and Landau-Devonshire potentials, in that it allows for control over the material constants for particular shape memory alloys, while still satisfying the theoretical requirements necessary for traditional existence theory. A final unique contribution of the proposed model is that it solves the resulting nonlinear system of partial differential equations in a way that minimizes the artificial dissipation required by previous solution methods, however we wait to discuss this topic until Chapter 5.
This chapter first examines the construction of the general, nonlinear thermodynamic model, which incorporates the inhomogeneous heat flux into the fully-coupled nonlinear model. The chapter then details our construction of the potential energy function at the heart of the model.

The spatial scale is the first consideration in constructing any model of material behavior. From Chapter 2 we know that the first-order phase transformation itself occurs as an instantaneous change on the molecular level (\(\sim 10^{-9}\) m). However, molecular models encounter the major difficulty that computations must remain limited to a very small region, since each particle must be accounted for separately. Furthermore, on the molecular level such averaged quantities as temperature and entropy are not defined, and instead are replaced with particle kinetics. Macroscopic models (\(\sim 10^0\) m) are similarly unhelpful, since the most interesting behaviors of these materials derive from the first-order phase transformation, which cannot be described by physical equations at the macroscopic level.

Therefore the proposed model is on a scale between these two spatial levels. This *mesoscopic* scale (\(\sim 10^{-5}\) m) consists of basic elements called *lattice particles*, which may be thought of as small parts of the continuum metallic lattice forming the body (see Huo, I. Mürler and Seelecke [77]). Mesoscale models are based on the assumption that the usually negligible fluctuations of the molecular particles can be ignored (see Lifshitz and Pitaevskii [97]). A schematic of these different scales is given in Figure 3.1.

The choice of modeling on this mesoscopic scale has a number of consequences. First, at this scale the material can be thought of on a continuum level, where such thermodynamic quantities as the entropy and temperature are still defined. Moreover, this mesoscale model assumes that lattice particles at this scale transform between phases as whole units. In other words, at this level the microscopic crystal lattice deformations correspond to mesoscopic strains in the lattice particle. Therefore the
Figure 3.1: The various scales of microstructure. On the macroscale, the shape memory properties are seen through volume averages of the deformation. On the microscale the phase transformations are witnessed, but the scale is too small for larger-scale computations. The mesoscale, however, may be considered to consist of lattice particles which both satisfy continuum thermodynamics and experience phase transformations as a unit (figure adapted from Tanaka et al. [131]).

Mesoscopic strain may be considered as an order parameter. Thus, measurements of the mesoscopic strain permit recognition of individual material phases for each lattice particle (see Niezgódka and Sprekels [114]).

3.1 First-Principles Approach

In order to achieve the most general model for the thermodynamic behavior of shape memory alloy wires possible, we begin with the standard physical laws for solid materials. Then using as few assumptions and linearizations as possible we combine these physical laws to derive a coherent mathematical model for the thermodynamic behavior of shape memory alloy wires.

We first define the quantities to be used throughout the derivation in the following Table 3.1. Some of these have been discussed in Chapter 2, however we include them all here for completeness.

Remark 3.1.1. We make the following remarks on the variables and notation used throughout the derivation of the thermodynamic model.
<table>
<thead>
<tr>
<th>Variable</th>
<th>Range</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$u$</td>
<td>$\mathbb{R}^N$</td>
<td>m</td>
<td>deformation</td>
</tr>
<tr>
<td>$\theta$</td>
<td>$\mathbb{R}^+ \setminus {0}$</td>
<td>K</td>
<td>absolute temperature</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>$\mathbb{R}^{N \times N}$</td>
<td></td>
<td>deformation gradient, $\gamma = \nabla u$</td>
</tr>
<tr>
<td>$b$</td>
<td>$\mathbb{R}^N$</td>
<td>N kg$^{-1}$</td>
<td>external body force</td>
</tr>
<tr>
<td>$r$</td>
<td>$\mathbb{R}$</td>
<td>W kg$^{-1}$</td>
<td>external heat supply</td>
</tr>
<tr>
<td>$q$</td>
<td>$\mathbb{R}^N$</td>
<td>W m$^{-2}$</td>
<td>referential heat flux</td>
</tr>
<tr>
<td>$\sigma$</td>
<td>$\mathbb{R}^{N \times N}$</td>
<td>Pa</td>
<td>first Piola-Kirchoff stress</td>
</tr>
<tr>
<td>$\eta$</td>
<td>$\mathbb{R}$</td>
<td>J kg$^{-1}$ K$^{-1}$</td>
<td>entropy density</td>
</tr>
<tr>
<td>$\varepsilon$</td>
<td>$\mathbb{R}$</td>
<td>J kg$^{-1}$</td>
<td>internal energy density</td>
</tr>
<tr>
<td>$\Psi$</td>
<td>$\mathbb{R}$</td>
<td>J kg$^{-1}$</td>
<td>Helmholtz free energy density</td>
</tr>
</tbody>
</table>

Table 3.1: Variables and their units used throughout the model derivation. All quantities are defined for points $(x,t)$ in the domain $\Omega \times \mathbb{R}^+$. 

1. The reference configuration of the body is denoted $\Omega \subset \mathbb{R}^N$, and is taken to be unstressed austenite.

2. Since we consider the deformation as opposed to the displacement, the austenitic reference configuration is therefore given by $u(x) = x$.

3. The model is derived using an Lagrangian description of the physics, i.e. we treat all quantities in regards to the reference configuration unless specifically mentioned otherwise. As such, $\text{Div} (\cdot)$ refers to the divergence operator and $\nabla (\cdot)$ refers to the gradient, both taken with respect to the reference configuration.

4. Due to the multiple, distinct stable phases of shape memory alloys, the deformation has limited regularity, $u \in H^1(\mathbb{R}^+; H^1(\Omega))$. Thus the deformation gradient $\gamma(x,t) = \nabla u(x,t)$ is only defined almost everywhere $x \in \mathbb{R}^N$.

5. The Helmholtz free energy is defined according to Landau as $\Psi = \varepsilon - \theta \eta$ [97].

6. Derivatives adhere to the following shorthand:

   - Time derivatives: $\frac{\partial}{\partial t} w(x,t) = \dot{w}(x,t)$;
• Partial derivatives: \( \frac{\partial}{\partial x} w(x, t) = \partial_x w(x, t) \) and \( \frac{\partial^2}{\partial x \partial y} w(x, y) = \partial_{xy}^2 w(x, y) \);

• Total derivatives: \( \frac{d}{dt} w(x) = dw(x) \) and \( \frac{d^2}{dt^2} w(x) = d^2 w(x) \).

Moreover, the various material constants used throughout the model are summarized in Table 3.2. We proceed with the standard conservation laws of continuum physics, that will be used in the derivation of our model in Section 3.2:

**Referential conservation of mass:** The principle of conservation of mass states that the time-derivative of the referential density does not change (see Dafermos [41]). Denoting \( \rho_0 \) for the referential density, this may be written as

\[
\dot{\rho}_0 = 0. \tag{3.1}
\]

**Referential conservation of linear momentum:** The principle of conservation of linear momentum states that momentum is conserved in the moving body. Under the external body force per unit mass \( b \), the referential field equations may be posed mathematically as

\[
\partial_t (\rho_0 \dot{u}) = \text{Div}(\sigma) + \rho_0 b, \tag{3.2}
\]

where \( \dot{u} \) is the deformation and \( \sigma \) is the first Piola-Kirchoff stress tensor per unit area (see Dafermos [41] and Klouček [84]).

**Referential conservation of angular momentum:** The principle of conservation of angular momentum similarly states that angular momentum is conserved in the moving body. This may be written as the material response condition (see Dafermos [41])

\[
\partial_t (u \wedge \rho_0 \dot{u}) = \text{Div}(\dot{u} \wedge \sigma) + \dot{u} \wedge \rho_0 b, \tag{3.3}
\]

which translates into the following symmetry condition on the referential stress,

\[
\sigma \gamma^T = \gamma \sigma^T,
\]
<table>
<thead>
<tr>
<th>Constant</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_0$</td>
<td>kg m$^{-3}$</td>
<td>referential density</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>W m$^{-1}$ K$^{-1}$</td>
<td>thermal conductivity</td>
</tr>
<tr>
<td>$c_p$</td>
<td>J kg$^{-1}$ K$^{-1}$</td>
<td>specific heat capacity</td>
</tr>
<tr>
<td>$E_a$</td>
<td>GPa</td>
<td>elastic modulus of austenite phase</td>
</tr>
<tr>
<td>$E_m$</td>
<td>GPa</td>
<td>elastic modulus of martensite phase</td>
</tr>
<tr>
<td>$\theta_A$</td>
<td>K</td>
<td>austenite finish temperature</td>
</tr>
<tr>
<td>$\theta_C$</td>
<td>K</td>
<td>critical transition temperature</td>
</tr>
<tr>
<td>$\theta_M$</td>
<td>K</td>
<td>martensite finish temperature</td>
</tr>
</tbody>
</table>

Table 3.2: Material Constants used in the model, along with their physical units.

where $\gamma$ is the referential deformation gradient.

**Referential conservation of energy (First Law of Thermodynamics):** The principle of conservation of energy states that the time rate of change of the overall energy in any part of the body must equal the sum of the rate of work of the body force, heat supply, stress and heat flux. Denoting the external heat supply per unit mass as $r$, the balance of energy may be written as

$$\partial_t \left( \rho_0 \varepsilon + \frac{1}{2} \rho_0 \dot{u} \cdot \dot{u} \right) = \text{Div} (\dot{u} \cdot \sigma - q) + \rho_0 \dot{u} \cdot b + \rho_0 r,$$

(3.4)

where $\varepsilon$ denotes the internal energy density per unit referential volume, and $q$ provides the referential heat flux (see Carlson [29], Dafermos [41] and Klouček [84]).

**Clausius-Duhem Inequality (Second law of Thermodynamics):** The conservation laws are complemented by the Second Law of Thermodynamics, first discovered by R. Clausius in 1865 [97]. This law postulates that the growth of the entropy in any part of the body must equal at least the sum of the entropy flux across the boundary and the entropy production in the interior of that part. This may be posed mathematically in the referential form as

$$\partial_t (\rho_0 \eta) \geq - \text{Div} \left( \frac{q}{\theta} \right) + \rho_0 \left( \frac{r}{\theta} \right),$$

(3.5)
where the entropy density per unit referential volume is given by $\eta$ (see Dafermos [41]).

### 3.2 Derivation of Proposed PDE Model

In addition to the above laws of physics, we add various constitutive material laws to derive a fully coupled, nonlinear model describing the thermodynamic behavior of shape memory alloy wires. As with any model based on physical phenomena, we must make certain constitutive assumptions on our material to allow for a meaningful and tractable mathematical model. We state these assumptions here in order to clearly define the abilities and limitations of our model for describing the thermodynamics of shape memory alloy wires.

First, we assume some level of regularity on the functional components of the model. These assumptions allow the application of standard tools from multivariable calculus and partial differential equations for the construction and analysis of the model.

- The Helmholtz free energy $\Psi$ is at least a $C^2$-function of its arguments.

- The internal energy $\varepsilon$ is at least a $C^1$-function of its arguments.

- The entropy $\eta$ is at least a $C^1$-function of its arguments.

- The temperature $\theta$, which may be defined according to $\partial \eta / \partial \varepsilon = 1/\theta$, (see Lifshitz and Pitaevskii [97]) is strictly positive.

- The velocity $\dot{u}$ is Lipschitz continuous with respect to $x$ and $t$.

The following assumptions are used to reduce the complexity of the proposed thermodynamic model, thereby limiting the applicability of the model toward some physical systems.
- The material is assumed to be thermoelastic, as described in further detail in Section 3.2.2.

- The specific heat can be represented by a constant with respect to the deformation gradient and temperature, as discussed in further detail in Section 3.2.4.

- The material system is considered to be under constant pressure, with possibly changing volume.

- The wire is modeled in one dimension instead of three, though this condition will not be imposed until the end of the derivation.

Finally, we make the further constitutive assumption that any possible viscous stresses in the material may be incorporated into the stress through adding the term $A\dot{\gamma}$, where $A \in \mathbb{R}^{3 \times 3}$ is the symmetric viscous anisotropy matrix having units kg m$^{-1}$ s$^{-1}$. Therefore, the inclusion of viscous stresses into the above laws results in a modified balance of linear momentum,

$$\partial_t (\rho_0 \dot{u}) = \text{Div} (\sigma + A\dot{\gamma}) + \rho_0 b,$$

and in a modified conservation of energy

$$\partial_t \left( \rho_0 \varepsilon + \frac{1}{2} \rho_0 \dot{u} \cdot \dot{u} \right) = \text{Div} (\dot{u} \cdot \sigma + \dot{\gamma} \cdot A\dot{u} - q) + \rho_0 \dot{u} \cdot b + \rho_0 r.$$

**Remark 3.2.1.** The use of such viscous stresses has no physical justification, as described in further detail in both Chapter 4 and Chapter 5. Rather, viscous effects are included here in the derivation of the thermoelastic model to provide a natural means for numerical regularization of the resulting nonlinear thermodynamics. Therefore, the true case of a thermoelastic material corresponds to a symmetric viscous anisotropy matrix $A = 0$. The attempt at resolving such zero-viscosity states forms a central numerical issue of the discussion in Chapter 5.
3.2.1 Simplification of the Balance of Energy

With these assumptions in mind, we rewrite the referential balance of energy (3.7) using the linearity of the differentiation operator to expand the time derivative and the divergence terms to obtain

\[ \rho_0 \dot{\varepsilon} + \rho_0 \ddot{u} \cdot \dot{u} = \text{Div}(\dot{u} \cdot \sigma) - \text{Div}(q) + \text{Div}(\dot{\gamma} \cdot A \dot{u}) + \rho_0 \ddot{u} \cdot b + \rho_0 \dot{r}. \] (3.8)

Similarly, application of the conservation of mass and the linearity of the divergence operator to the referential balance of linear momentum (3.6) yields

\[ \rho_0 \ddot{u} = \text{Div}(\sigma) + \text{Div}(A \dot{\gamma}) + \rho_0 b. \] (3.9)

Now since \( \dot{u}(x, t) \) is Lipschitz continuous in both of its arguments, the product rule for differentiation, along with equation (3.9) allow us to write

\[ \text{Div}(\dot{u} \cdot \sigma) = \text{Div}(\sigma) \cdot \dot{u} + \text{Div}(\dot{u}) \cdot \sigma \]
\[ = (\rho_0 \ddot{u} - \rho_0 \dot{b} - \text{Div}(A \dot{\gamma})) \cdot \dot{u} + \sigma \cdot \dot{\gamma}. \] (3.10)

We now use the product rule to expand the divergence term Div(\( \dot{\gamma} \cdot A \dot{u}) \) from equation (3.8) to

\[ \text{Div}(\dot{\gamma} \cdot A \dot{u}) = \text{Div}(A \dot{\gamma}) \cdot \dot{u} + \text{Tr}(\dot{\gamma}^T A \dot{\gamma}). \] (3.11)

Thus combining (3.8), (3.10) and (3.11), the balance of energy may be simplified to

\[ \rho_0 \dot{\varepsilon} = \sigma \cdot \dot{\gamma} - \text{Div}(q) + \text{Tr}(\dot{\gamma}^T A \dot{\gamma}) + \rho_0 \dot{r}. \] (3.12)

3.2.2 Thermoelastic Materials

Following the work by Carlson and Dafermos [29, 41], we now examine the assumption that shape memory alloys may be considered as thermoelastic materials. Thermoelastic materials are characterized by the constitutive assumption that the material quantities \( \Psi, \sigma, \eta \) and \( q \) depend only on the deformation gradient \( \gamma \), temperature \( \theta \),
and temperature gradient $\nabla \theta$ (and implicitly on $x$ and $t$ through these quantities), i.e.

$$
\Psi = \Psi(\gamma, \theta, \nabla \theta), \quad \sigma = \sigma(\gamma, \theta, \nabla \theta), \quad \eta = \eta(\gamma, \theta, \nabla \theta) \quad \text{and} \quad q = q(\gamma, \theta, \nabla \theta).
$$

According to the principle of equipresence, the constitutive assumption of thermoelectricity must be applied toward all of the intensive (i.e. independent of the material volume) thermodynamic quantities shown above. That is, we must first assume that these quantities all depend only on $\gamma$, $\theta$ and $\nabla \theta$. A fundamental premise from continuum physics, however, holds that the Clausius-Duhem inequality may be regarded as a condition on the material response, thereby inducing restrictions on these constitutive relations and eliminating some of the assumed dependencies above. To this end, we consider the Clausius-Duhem inequality (3.5):

$$
\partial_t (\rho_0 \eta) \geq - \text{Div} \left( \frac{q}{\theta} \right) + \rho_0 \left( \frac{\tau}{\theta} \right).
$$

Using the definition of the Helmholtz free energy, $\Psi = \varepsilon - \theta \eta$, and the product rule for differentiation, we have the following relations

$$
\dot{\Psi} = \dot{\varepsilon} - \partial \varepsilon \partial \eta - \partial \eta \partial \varepsilon,
$$

$$
\theta \text{Div} \left( \frac{q}{\theta} \right) = \text{Div}(q) - \frac{q}{\theta} \cdot \nabla \theta.
$$

We combine these with the Clausius-Duhem inequality (3.5) and the simplified form of the balance of energy (3.12) to obtain the inequality

$$
\rho_0 \dot{\Psi} + \rho_0 \eta \dot{\varepsilon} - \sigma \cdot \dot{\gamma} + \frac{1}{\theta} q \cdot \nabla \theta - \text{Tr} \left( \dot{\gamma}^T A \dot{\gamma} \right) \leq 0.
$$

Using the chain rule for the time differentiation of the free energy

$$
\dot{\Psi}(\gamma, \theta, \nabla \theta) = \partial_t \Psi(\gamma, \theta, \nabla \theta) \dot{\gamma} + \partial_\gamma \Psi(\gamma, \theta, \nabla \theta) \dot{\theta} + \partial_\theta \Psi(\gamma, \theta, \nabla \theta) \dot{\theta}.
$$
we rewrite the inequality (3.13) as (dropping the arguments $\gamma, \theta$ and $\nabla \theta$)

\[
0 \geq \rho_0 \partial_{\tau} \Psi \cdot \dot{\gamma} + \rho_0 \partial_{\theta} \Psi \dot{\theta} + \rho_0 \partial_{\nabla \theta} \Psi \cdot \nabla \dot{\theta} + \rho_0 \eta \dot{\theta} - \sigma \cdot \dot{\gamma} + \frac{1}{\dot{\theta}} q \cdot \nabla \theta - \text{Tr} \left( \dot{\gamma}^T A \dot{\gamma} \right) \]

\[\Leftrightarrow\]

\[
0 \geq (\rho_0 \partial_{\tau} \Psi - \sigma) \cdot \dot{\gamma} + \rho_0 (\partial_{\theta} \Psi + \eta) \dot{\theta} + \rho_0 \partial_{\nabla \theta} \Psi \cdot \nabla \dot{\theta} + \frac{1}{\dot{\theta}} q \cdot \nabla \theta - \text{Tr} \left( \dot{\gamma}^T A \dot{\gamma} \right). \]

This inequality must hold true for all $\dot{\gamma} \in \mathbb{R}^{3 \times 3}, \dot{\theta} \in \mathbb{R}$ and $\nabla \dot{\theta} \in \mathbb{R}^3$, and all allowable $\gamma, \theta$ and $\nabla \theta$. Therefore, since the time derivatives $\dot{\gamma}, \dot{\theta}$ and $\nabla \dot{\theta}$ may independently be positive, negative or zero, the above condition (3.14) can only be satisfied if the internal physical quantities $\Psi, \sigma, \eta, q$ and $A$ meet the following conditions

\[
\Psi \neq \Psi(\nabla \theta) \Rightarrow \Psi = \Psi(\gamma, \theta), \quad (3.15)
\]

\[
\sigma \neq \sigma(\nabla \theta) \Rightarrow \sigma = \sigma(\gamma, \theta), \quad (3.16)
\]

\[
\eta \neq \eta(\nabla \theta) \Rightarrow \eta = \eta(\gamma, \theta), \quad (3.17)
\]

\[
q = q(\gamma, \theta, \nabla \theta), \quad (3.18)
\]

\[
\sigma = \rho_0 \partial_{\tau} \Psi \quad (3.19)
\]

\[
\eta = -\partial_{\theta} \Psi \quad (3.20)
\]

\[
q \cdot \nabla \theta \leq 0, \quad (3.21)
\]

\[
\text{Tr} \left( \dot{\gamma}^T A \dot{\gamma} \right) \geq 0. \quad (3.22)
\]

That is, in order for the Clausius-Duhem inequality (3.5) to hold for thermoelastic materials, the Helmholtz free energy $\Psi$, the stress $\sigma$, and the entropy $\eta$ must all be independent of the temperature gradient $\nabla \theta$ (conditions (3.15)-(3.17)); the heat flux $q$ is the only quantity that depends on this quantity (condition (3.18)). Furthermore, the Clausius-Duhem inequality implies that both the stress $\sigma$ and the entropy density $\eta$ may be written as functions of the Helmholtz free energy $\Psi$ (conditions (3.19) and (3.20)). Also, the condition (3.21) affirms the standard thought that heat flows from
areas of high temperature to areas of low temperature. Lastly, the condition \((3.22)\) places the condition that the symmetric viscous anisotropy matrix \(A\) must be positive semidefinite.

### 3.2.3 Reduction of the Energy Equation

Using the above conditions \((3.15)-(3.22)\), the balance of energy \((3.12)\) may be further simplified. From the conditions \((3.15)-(3.17), (3.19)\) and \((3.20)\), as well as the definition of the Helmholtz free energy \(\Psi = \varepsilon - \theta \eta\), the time derivative of the free energy may be equivalently written in the following two forms:

\[
\hat{\Psi}(\gamma, \theta) = \partial_\gamma \Psi(\gamma, \theta) \cdot \dot{\gamma} + \partial_\theta \Psi(\gamma, \theta) \dot{\theta} = \frac{1}{\rho_0} \sigma(\gamma, \theta) \cdot \dot{\gamma} - \dot{\theta} \eta(\gamma, \theta),
\]

and

\[
\hat{\Psi}(\gamma, \theta) = \dot{\varepsilon}(\gamma, \theta) - \dot{\theta} \eta(\gamma, \theta) - \theta \dot{\eta}(\gamma, \theta).
\]

Setting these two forms equal and canceling common terms, the time derivative of the internal energy density may be written as

\[
\rho_0 \dot{\varepsilon}(\gamma, \theta) = \sigma(\gamma, \theta) \cdot \dot{\gamma} + \rho_0 \theta \dot{\eta}(\gamma, \theta). \tag{3.23}
\]

Inserting this into the energy equation \((3.12)\) and canceling terms, the balance of energy may be written in the entropy conservation form

\[
\rho_0 \theta \dot{\eta}(\gamma, \theta) = - \text{Div} \left( g(\gamma, \theta, \nabla \theta) \right) + \text{Tr} \left( \dot{\gamma}^T A \dot{\gamma} \right) + \rho_0 r. \tag{3.24}
\]

### 3.2.4 The Specific Heat Capacity

The energy equation \((3.24)\) may be further reduced through use of the specific heat capacity. The specific heat capacity is generally defined as the quantity of heat which must be gained by a material element in order to raise the temperature of that element
by one unit. According to Landau and Lifshitz [97] and Dafermos [41], at constant pressure this is given by

\[ c_p = \theta \partial_\theta \eta(\gamma, \theta) > 0. \]

**Remark 3.2.2.** We note that from the definition of, and smoothness assumptions on, the Helmholtz free energy \( \Psi \), the specific heat may be equivalently written as

\[
c_p(\gamma, \theta) = \theta \partial_\theta \eta(\gamma, \theta) = \partial_\theta \Psi(\gamma, \theta) + \eta(\gamma, \theta) + \theta \partial_\theta \eta(\gamma, \theta) = \partial_\theta \varepsilon(\gamma, \theta),
\]

which agrees with alternate definitions of the specific heat (see Carlson [29] and Klouček [83, 84]).

Thus since the entropy is assumed continuous with respect to both \( t \) and \( \theta \), and since \( \Psi \) is smooth in \( \gamma \) and \( \theta \), application of the chain rule results in the relation

\[
\rho_0 \theta \dot{\eta}(\gamma, \theta) = \rho_0 \theta \partial_\theta \eta(\gamma, \theta) \dot{\theta} + \rho_0 \theta \partial_\gamma \eta(\gamma, \theta) \cdot \dot{\gamma}
\]

\[
= \rho_0 c_p(\gamma, \theta) \dot{\theta} - \rho_0 \theta \partial^2_{\gamma \theta} \Psi(\gamma, \theta) \cdot \dot{\gamma},
\] (3.25)

We combine this with the entropy conservation form of the energy equation (3.24) to obtain the following nonlinear heat equation

\[
\rho_0 c_p(\gamma, \theta) \dot{\theta} = \rho_0 \theta \partial^2_{\gamma \theta} \Psi(\gamma, \theta) \cdot \dot{\gamma} - \text{Div}(q(\gamma, \theta, \nabla \theta)) + \text{Tr}(\dot{\gamma}^T A \dot{\gamma}) + \rho_0 r.
\] (3.26)

We now enforce the assumption that we may approximate the specific heat \( c_p(\gamma, \theta) \) by a constant. The necessity of this assumption arises from the data provided in the materials science literature, which furnishes only a single constant for the measured heat capacity of these materials [21, 24, 26, 78, 45, 92, 94, 115, 116, 117]. Our proposed model, therefore, considers the following simplified form of the nonlinear heat equation,

\[
\rho_0 c_p \dot{\theta} = \rho_0 \theta \partial^2_{\gamma \theta} \Psi(\gamma, \theta) \cdot \dot{\gamma} - \text{Div}(q(\gamma, \theta, \nabla \theta)) + \text{Tr}(\dot{\gamma}^T A \dot{\gamma}) + \rho_0 r.
\] (3.27)
3.2.5 Heat Flux

The final ingredient needed in the derivation of our general nonlinear thermodynamic model is a proper form for the referential heat flux $q(\gamma, \theta, \nabla \theta)$ in the energy equation (3.27). Traditionally, thermodynamic models incorporate the standard Fourier heat flux

$$q(\nabla \theta) = -K \nabla \theta,$$

where $K$ provides the heat conductivity tensor (of rank two). However, in the case of materials such as shape memory alloys that experience solid-state phase transformations, the standard Fourier heat flux provides an insufficient description for the heat transfer properties of the material.

According to Abeyaratne and Knowles [3], "in the theory of [thermodynamically] irreversible processes, it is customary to postulate a constitutive relationship in which the present value of [the heat] flux is a function of the present value of the [material state]." In their thermodynamic theory on the dynamics of phase boundaries in solid-state phase transformations, Abeyaratne and Knowles make the assumption that the heat flux is only piecewise continuous in the spatial variable. The necessity of such an assumption is explained by R.D. James and P. Klouček [83]. Experimental evidence witnesses the movement of strain discontinuities (phase boundaries) during the thermoelastic martensitic transformation. This movement is attributed to a non-smooth velocity field across the phase boundary. On both sides of this phase boundary, however, the material resides in a stress-free deformation. Therefore the movement of the boundary must result from varying rates of temperature change in each of these phases. This varying temperature change alludes to a jump in the heat flux across the phase boundary. Since the deformation gradient $\gamma$ serves as the order parameter characterizing the different material phases, varying flow of heat requires an inhomogeneous heat flux $q$ depending on the deformation gradient $\gamma$ as well as the temperature $\theta$ and temperature gradient $\nabla \theta$. 
Thus, the proposed model incorporates such an inhomogeneous heat flux \( q(\gamma, \theta, \nabla \theta) \), which we describe in the remainder of this section. From the Clausius-Duhem requirement (3.21), the heat flux must satisfy

\[
q(\gamma, \theta, \nabla \theta) \cdot \nabla \theta \leq 0. \tag{3.28}
\]

The referential heat flux must further satisfy the *principle of frame indifference*,

\[
q(Q\gamma, \theta, \nabla \theta) = q(\gamma, \theta, \nabla \theta), \tag{3.29}
\]

for all possible states \( \gamma, \theta, \nabla \theta \) and all proper orthogonal matrices \( Q \in SO(3) \) (see Ciarlet [32], Dafermos [41] or Gurtin [68] for further details). Moreover, the spatial heat flux \( \hat{q} \), related to the referential heat flux \( q \) by \( \hat{q}(\gamma, \theta, \nabla \theta) = \det(\gamma)^{-1} \gamma q(\gamma, \theta, \nabla \theta) \), must satisfy the material symmetry condition

\[
\hat{q}(\gamma R, \theta, R^T \nabla \theta) = \hat{q}(\gamma, \theta, \nabla \theta)
\]

for all matrices \( R \) in the material symmetry group of the crystal lattice [32, 41, 68]. This translates into the following material symmetry condition on the referential heat flux:

\[
Rq(\gamma R, \theta, R^T \nabla \theta) = q(\gamma, \theta, \nabla \theta). \tag{3.30}
\]

Two of the possible choices for a referential heat flux satisfying (3.28), (3.29) and (3.30) are

\[
q(\gamma, \theta, \nabla \theta) = -\kappa \nabla \theta, \quad \text{and} \tag{3.31}
\]

\[
q(\gamma, \theta, \nabla \theta) = -\kappa \det(\gamma) \nabla \theta, \tag{3.32}
\]

where \( \kappa \in \mathbb{R}^+ \setminus \{0\} \) is the coefficient of thermal conductivity. The first of these choices corresponds to the standard Fourier heat flux. However, the second incorporates information on the material state (encoded in the deformation gradient \( \gamma \)) into an inhomogeneous version of the material heat flux. We choose this second form of the referential heat flux (3.32). Therefore, the nonlinear heat equation (3.27) becomes

\[
\rho_0 c_p \dot{\theta} = \rho_0 \theta \partial_{\gamma \theta}^2 \Psi(\gamma, \theta) \cdot \dot{\gamma} + \text{Tr} (\dot{\gamma}^T A \dot{\gamma}) + \text{Div}(\kappa \det(\gamma) \nabla \theta) + \rho_0 r. \tag{3.33}
\]
Remark 3.2.3. The proposed heat flux (3.32) satisfies the conditions (3.28), (3.29) and (3.30).

1. Checking the heat flow condition $q(\gamma, \theta, \nabla \theta) \cdot \nabla \theta \leq 0$:

$$q(\gamma, \theta, \nabla \theta) \cdot \nabla \theta = -\kappa \det(\gamma) \nabla \theta \cdot \nabla \theta$$

$$= -\kappa \det(\gamma) \|\nabla \theta\|_2^2$$

$$\leq 0,$$

since $\kappa > 0$ and by definition the deformation gradient satisfies the orientation-preserving condition [32]

$$\det(\gamma(x)) > 0 \quad \forall x \in \Omega.$$

2. Checking the principle of frame indifference $q(Q\gamma, \theta, \nabla \theta) = q(\gamma, \theta, \nabla \theta)$ for all $Q \in SO(3)$:

$$q(Q\gamma, \theta, \nabla \theta) = -\kappa \det(Q\gamma) \nabla \theta$$

$$= -\kappa \det(Q) \det(\gamma) \nabla \theta \quad [\text{property of determinant}]$$

$$= -\kappa \det(\gamma) \nabla \theta \quad [Q \text{ is orthogonal}]$$

$$= q(\gamma, \theta, \nabla \theta).$$

3. Checking invariance under material symmetry $Rq(\gamma R, \theta, R^T \nabla \theta) = q(\gamma, \theta, \nabla \theta)$, for all matrices $R$ in the material symmetry group of the crystal lattice:

$$Rq(\gamma R, \theta, R^T \nabla \theta) = -R\kappa \det(\gamma R) R^T \nabla \theta$$

$$= -RR^T \kappa \det(\gamma) \det(R) \nabla \theta \quad [\kappa \det(\gamma R) \in IR]$$

$$= -\kappa \det(\gamma) \nabla \theta \quad [\text{see } \dagger]$$

$$= q(\gamma, \theta, \nabla \theta).$$

\dagger Since $R$ is in the material symmetry group, which is a subset of $SO(3)$.

Remark 3.2.4. The proposed heat flux (3.32) is not entirely different from the standard Fourier heat flux (3.31). When the material is at rest in the undeformed reference configuration (austenite), $\det(\gamma) = 1$ and the two heat fluxes are equivalent. When in the martensitic state, however, $|\det(\gamma) - 1| \leq 0.08$; allowing for the required discontinuities in the flux across phase boundaries, while remaining “close” to the standard
*Fourier flux. However, our choice of the heat flux allows different conductivities for the two martensitic variants, which may be unrealistic. Therefore a form of the inhomogeneous heat flux that allows a single conductivity for both martensitic forms may prove more true to the physics of shape memory alloys.*

### 3.3 Resulting PDE system

The proposed thermodynamic model considers the fully-coupled nonlinear system of equations given by the balance of linear momentum (3.6) and the conservation of energy (3.33) to describe the thermodynamic behavior of a broad class of materials. Often, the solutions to these equations are considered separately, where one alternates between solving one equation and then the other. However, those approaches are flawed due to the dependence of both equations on the thermodynamic quantities $u$ and $\theta$. Therefore, we instead consider the fully coupled, nonlinear system of partial differential equations given by

\[
\begin{align*}
\rho_0\ddot{u} &= \text{Div}(\rho_0 \partial_\gamma \Psi + A\dot{\gamma}) + \rho_0 b \\
\rho_0 c_p \dot{\theta} &= \rho_0 \partial_{\gamma \theta} \dot{\gamma} + \text{Tr}(\gamma^T A\dot{\gamma}) + \text{Div}(\kappa \det(\gamma)\nabla \theta) + \rho_0 r.
\end{align*}
\]

We consider the following experimental conditions for calculations of the thermodynamics in the material system:

- The material is shaped like a wire of length $L$, with small cross-sectional area given by $\beta$, which will be modeled in $\mathbb{R}$ instead of $\mathbb{R}^3$. In this case, deformations $u(x, t) > x$ imply extension of the wire, and deformations $u(x, t) < x$ imply compression of the wire.

- Body forces are experienced as compressive vibrational stresses acting on the material elements at the $x = L$ end of the wire only; this end is also considered to be fully thermally insulated.
• At the $x = 0$ end of the wire the temperature is fixed in a thermal bath at $\theta_{DC}$ and its deformation is held fixed at zero.

• The body begins with initial deformation $u_0(x)$, initial velocity $v_0(x)$ and initial temperature distribution $\theta_0(x)$.

• The one-dimensional symmetric viscous anisotropy matrix $A$ is represented by the constant $\alpha$.

Under these experimental conditions, we reduce the four-dimensional nonlinear PDE system (3.34) to the following two-dimensional nonlinear PDE system

$$
\rho_0 \ddot{u} = \rho_0 \partial_{xx}^2 \Psi + \alpha \partial_{xx}^2 \dot{u} + \rho_0 b
$$

$$
\rho_0 c_p \dot{\theta} = \rho_0 \theta \partial_{\gamma\theta}^2 \Psi \dot{\gamma} + \alpha \dot{\gamma}^2 + \kappa \partial_x (\gamma \partial_x \theta) + \rho_0 r;
$$

with the experimental conditions dictating the initial and boundary conditions

$$
u(x, t_0) = u_0, \quad \dot{u}(x, t_0) = v_0, \quad \theta(x, t_0) = \theta_0$$

$$
u(0, t) = 0, \quad \dot{u}(0, t) = 0, \quad \theta(0, t) = \theta_{DC}(t), \quad \partial_x \theta(L, t) = 0.$$

Furthermore, the behavior of the SMA wire is controlled through the heat supply term $r(x, t)$. The vibrational forces are implemented in the body force term $b(x, t)$, and affect the material at the $x = L$ end of the wire. These forces are implemented as a force acting on the material elements in a small amount of the material in the spatial interval $(L - \delta x, L)$.

The system (3.35) is solved to find the thermodynamic state ($u$ and $\theta$) in the shape memory alloy wire at any point $(x, t) \in [0, L] \times [0, \infty)$. However, in order to apply this general thermodynamic model to specific alloys, we must construct the Helmholtz free energy, which lies at the core of the thermodynamic system (3.35).
3.4 Helmholtz Free Energy

We propose a new construction of the Helmholtz free energy, \( \Psi(\gamma, \theta) \), which provides the basic quantity by which all of the physics of a shape memory material may be derived. The general system of equations (3.35) provides a self-consistent system of partial differential equations describing the behavior of thermoelastic wires on a mesoscopic level. In order to apply this thermoelastic model to shape memory alloys, the key behaviors discussed in Chapter 2 and the material constants found in Table 3.2 must be incorporated into the Helmholtz free energy. Thus a proper construction of this free energy is extremely important in the reliable modeling of specific types of shape memory alloy wires.

For our thermodynamic model, we construct a new one-dimensional form of the Helmholtz free energy \( \Psi(\gamma, \theta) \) that describes both the austenitic and martensitic phases as well as the transformation pathway between the two, while still providing a reasonable control over the material constants. This free energy is an improvement over other forms of the free energy used in a variety of models discussed in Section 2.5 ([1]-[6], [34], [46]-[54], [59], [72], [77], [83]-[87], [99], [104]-[106], [114]). As mentioned in Section 2.5.1, most of these nonlinear models may be characterized as those that are either based on a three-dimensional form of \( \Psi \) that is valid in only a very small temperature range, or those that are based on one-dimensional forms of \( \Psi \) that reproduce only a few of the phenomenological properties and leave out material properties such as the specific heat and elastic modulus. Our proposed form of the free energy is, therefore, similar to these in that it provides a phenomenologically-based, one-dimensional, nonlinear model for shape memory wires. However, it is materially tunable, allowing for control over both the nonlinear physics inherent to shape memory alloys, as well as the material constants intrinsic to particular materials.
3.4.1 Landau Theory

In the plenary work by Landau and Lifshitz [97], the transition from a molecular
dynamics description of materials to a continuum description is made using the tools
of statistical thermodynamics. As seen in the model derivation, Section 3.2, at the
heart of of this continuum material description is the Helmholtz free energy $\Psi$. Up
until the construction of this free energy, the thermodynamical model (3.34) applies
as a mesoscopic model for a broad range of thermoelastic materials. It is, thus, in the
Helmholtz free energy that we must include the general properties of the first-order
phase transition as well as the material properties which allow the specification of the
particular shape memory alloy under consideration.

Of particular importance in Landau's theory is the differential structure of the free
energy [97]. As discussed in Section 2.2, the molecular structure of the crystal lattice
depends on the minimum energy states of the potential energy. In a state of
thermal equilibrium any mechanical equilibrium configuration of the lattice particle
corresponds to a minimum of the free energy functional $\Psi$ with respect to the deforma-
tion gradient $\gamma$. In other words, we may mathematically understand the stable
configurations of a lattice particle at a given temperature as global minima of the
free energy with respect to $\gamma$ at that temperature. Moreover, metastable states at a
particular temperature correspond to local minima of the free energy. The fact that
the first-order phase transformation occurs as a sudden lattice transition between dif-
ferent states demands a non-convexity of this potential energy, such that there exists
an energy barrier between various stable and metastable states (see Huo, I. Müller
and Seelecke [77]).

With this framework in mind, we see that the phenomenology of the first-order phase
transition for one-dimensional materials can be characterized by a free-energy having
isothermal energy profiles as seen in Figure 3.2. Here, for large temperatures ($\theta >$
Figure 3.2: Sketch of a first-order phase transition. Here, $\theta_C$ is the transition temperature at which $\Psi$ takes on its minimum at three values. For temperatures above this, the global minimum is at $\gamma = 1$, with relative minima developing at $\gamma = 1 \pm \gamma_0$ from $\theta_A \searrow \theta_C$. These minima become absolute and the relative minimum at $\gamma = 1$ disappears from $\theta_C \searrow \theta_M$. The curves have been displaced according to temperature to make the picture clearer (picture adapted from Ericksen [50]).

$\theta_A$) there is only one stable lattice configuration (austenite) corresponding to the reference configuration having deformation gradient $\gamma = 1$. When the temperature is lowered, the overall free energy loses convexity, forming two metastable lattice configurations (martensite variants) corresponding to the deformation gradients $\gamma = 1 \pm \gamma_0$. At the critical temperature $\theta_C$, the two martensitic variants become stable, and all three lattice configurations coexist as stable structures. Below this temperature, for $\theta \in (\theta_M, \theta_C)$, the martensitic configurations remain stable while the austenitic configuration becomes metastable. The austenitic state eventually becomes unstable for $\theta < \theta_M$. As the temperature is reduced in this process, the absolute minima change discontinuously. This abrupt change in the minima gives rise to the sudden crystal phase transformation and causes the release of latent heat.
Remark 3.4.1. We note that the above phenomenological description adheres to the required properties on the free energy of both observer invariance and material frame indifference [1]. The one-dimensional reduction of the principle of observer invariance corresponds to a condition of symmetry around the austenitic reference configuration $\gamma = 1$. Moreover, in this case, material frame indifference is equivalent to the free energy having equal energies at the martensitic wells of $\gamma = 1 \pm \gamma_0$.

3.4.2 Helmholtz Requirements

In keeping with Landau's theory, we construct our new Helmholtz free energy to satisfy a variety of constraints. First, we must match the material constants for particular shape memory alloys, to allow for evaluation and modeling of different materials. These constants for the NiTi alloy are found in Table 3.2, though we could include the material constants for any other shape memory alloy that experiences the same first-order phase transition. In particular, the Helmholtz free energy must match the following material constraints relating to the stable crystal configurations:

- the stable austenitic deformation gradient must be given by $\gamma = 1$,

$$\gamma = 1$$

- the stable martensitic deformation gradients must be given by $\gamma = 1 \pm \gamma_0$;

$$\gamma = 1 \pm \gamma_0$$

the transition temperatures:

- the critical transformation temperature is $\theta_C$,

$$\theta_C$$

- the martensite finish temperature is $\theta_M$, $\theta_M < \theta_C$,

$$\theta_M < \theta_C$$

- the austenite finish temperature is $\theta_A$, $\theta_A > \theta_C$;

$$\theta_A > \theta_C$$
the elasticity constants:

- the linear elastic modulus at \((1 \pm \gamma_0, \theta_M)\) is \(E_m\)  
  \[ E_m = \partial_\gamma \sigma(1 \pm \gamma_0, \theta_M), \]  
  (3.41)
- the linear elastic modulus at \((1, \theta_A)\) is \(E_a\)  
  \[ E_a = \partial_\gamma \sigma(1, \theta_A); \]  
  (3.42)

and the thermal properties:

- the specific heat constant is met at \(\theta_C\), i.e. \(\theta_C \partial_\theta \eta(\gamma, \theta_C) = c_p\),  
  (3.43)
- the specific heat is positive and close to \(c_p\) for \(\theta \neq \theta_C\)  
  \[ \theta \partial_\theta \eta(\gamma, \theta) > 0, \text{ and } \theta \partial_\theta \eta(\gamma, \theta) \approx c_p. \]  
  (3.44)

Further discussion on the specific heat requirements (3.43) and (3.44) is given in Section 3.4.3.

As discussed in Section 3.4.1, the first-order phase transition may be encoded into the Helmholtz free energy through the location of local and global minima with respect to \(\gamma\) at various temperatures \(\theta\). These differential constraints for a one-dimensional shape memory wire require that at the martensitic temperature \(\theta_M\),

\[ \partial_\gamma \Psi(1 \pm \gamma_0, \theta_M) = 0, \quad \partial_{\gamma \gamma} \Psi(1 \pm \gamma_0, \theta_M) > 0, \]  
  and \(\Psi(1 \pm \gamma_0, \theta_M) < \Psi(1, \theta_M); \)  
  (3.45)

at the transition temperature \(\theta_C\),

\[ \partial_\gamma \Psi(1 \pm \gamma_0, \theta_C) = 0, \quad \partial_\gamma \Psi(1, \theta_C) = 0, \]
\[ \partial_{\gamma \gamma} \Psi(1 \pm \gamma_0, \theta_C) > 0, \quad \partial_{\gamma \gamma} \Psi(1, \theta_C) > 0, \]  
  and \(\Psi(1 \pm \gamma_0, \theta_C) = \Psi(1, \theta_C); \)  
  (3.46)

and at the austenitic temperature \(\theta_A\),

\[ \partial_\gamma \Psi(1, \theta_A) = 0, \quad \partial_{\gamma \gamma} \Psi(1, \theta_A) > 0, \]  
  and \(\Psi(1 \pm \gamma_0, \theta_A) > \Psi(1, \theta_A), \)  
  (3.47)
which define the essential structure of the Helmholtz free energy necessary to provide the first-order phase transition.

The derivation of the general thermodynamic model given in Section 3.2 also places further requirements on the functional form of the Helmholtz free energy. One of these additional requirements is that the Helmholtz free energy must satisfy the smoothness constraint

$$\Psi(\gamma, \theta) \in C^2(\mathbb{R}^+, \mathbb{R}^+) \quad \forall \gamma, \theta,$$

which from (3.19) and (3.20) trivially satisfies the smoothness constraints on $\varepsilon$ and $\eta$. We further require that the second-derivative with respect to the deformation gradient $\gamma$ of the free energy remain bounded as $\gamma \to \infty$, i.e.

$$|\partial^2_{\gamma\gamma} \Psi(\gamma, \theta)| \leq C < +\infty \quad \forall \theta \text{ as } \gamma \to \infty,$$

which is necessary for the existence theory discussed in Section 4.2.3. Last, the assumption in Section 3.2.2 that shape memory alloys are thermoelastic further constrains the Helmholtz free energy as seen from the relations (3.19) and (3.20). These requirements, along with the definition of the Helmholtz free energy, provide the constraints

$$\Psi(\gamma, \theta) = \varepsilon(\gamma, \theta) - \theta \eta(\gamma, \theta),$$

$$\sigma(\gamma, \theta) = \rho \partial_\gamma \Psi(\gamma, \theta),$$

$$\eta(\gamma, \theta) = -\partial_\theta \Psi(\gamma, \theta),$$

$$\varepsilon(\gamma, \theta) = \Psi(\gamma, \theta) - \theta \partial_\theta \Psi(\gamma, \theta).$$

Upon satisfying the constraints (3.36)-(3.53), the resulting potential energy functional will be able to provide the characteristics of the Landau-Devonshire potential discussed by Niezgódka and Sprekels in [114]:

- existence of a Curie point $(\sigma_A, \theta_A)$, a critical point in the half-plane $\mathbb{R} \times \mathbb{R}^+$ which is the end of the domain of coexisting phases;
• possibility of temperature-induced phase transformations for $|\sigma| < \sigma_A$;

• occurrence of stress-induced phase transformations $A \Leftrightarrow M_\pm$ for $\theta \in (\theta_M, \theta_A)$;

• occurrence of stress-induced phase transformations $M_- \Leftrightarrow M_+$ for $\theta < \theta_M$;

• shape memory effects.

The computational results in Chapter 6 show that indeed the current model succeeds in these criteria. Thus, according to the theory proposed by Niezgódka and Sprekels, the potential satisfies the phenomenological requirements for modeling the physical behavior of shape memory alloys. Moreover, the proposed model goes further than Landau-Devonshire based models in that by satisfying the constraints (3.36)-(3.44), the potential also allows for the specification of material constants for individual shape memory alloys.

### 3.4.3 Helmholtz Construction

In order to satisfy the requirements on the Helmholtz free energy from Section 3.4.2, we propose the following new approach to the construction of $\Psi(\gamma, \theta)$. This approach allows both the general behavior of shape memory alloys, as well as control over the material constants. The construction consists of two phases:

1. First construct an elastic strain energy $\mathcal{W}(\gamma, \theta)$, which satisfies the material constant constraints (3.36)-(3.42), the requirements of the first-order phase transition (3.45)-(3.47), and the smoothness and boundedness constraints (3.48) and (3.49). We enforce the additional constraints that $\partial_{\theta\theta}^2 \mathcal{W}$ be small for all $(\gamma, \theta)$, and that at the transition temperature $\theta = \theta_C$, the strain energy satisfies $\partial_{\theta\theta} \mathcal{W}(\gamma, \theta_C) = 0$. These relate to the specific heat constraints (3.43) and (3.44), to be discussed further at the end of the chapter.
2. Then write the general form of the Helmholtz free energy as \( \Psi(\gamma, \theta) = \mathcal{W}(\gamma, \theta) + \hat{\Psi}(\theta) \), whereupon we construct the function \( \hat{\Psi}(\theta) \) to satisfy the remaining constraints (3.43)-(3.44), (3.48), and (3.50)-(3.53).

The proposed elastic strain energy density \( \mathcal{W} \) is constructed using the following form:

\[
\mathcal{W}(\gamma, \theta) = \mathcal{W}_M(\gamma)C_M(\theta) + \mathcal{W}_C(\gamma)C_C(\theta) + \mathcal{W}_A(\gamma)C_A(\theta). \tag{3.54}
\]

The functions \( \mathcal{W}_M(\gamma) \), \( \mathcal{W}_C(\gamma) \) and \( \mathcal{W}_A(\gamma) \) provide control over the isothermal energy profiles at the temperatures \( \theta_M \), \( \theta_C \) and \( \theta_A \), respectively. The coefficient functions \( C_M(\theta) \), \( C_C(\theta) \) and \( C_A(\theta) \) connect these isothermal profiles to provide a smooth transition through the range of allowable temperatures. This construction is significantly different from the traditional approaches of the Landau-Devonshire and Falk potentials as discussed in Section 2.5.1, which connect sextic and quadratic function functions of the stretch \( \gamma \) using a linear function of the temperature,

\[
\Psi(\gamma, \theta) = \alpha_1 \gamma^6 - \alpha_2 \gamma^4 + (\alpha_3 \theta - \alpha_4) \gamma^2 + \Psi_0(\theta).
\]

Such constructions do not allow complete control over the full range of requirements (3.36)-(3.53). Furthermore, the linear temperature dependence of the Landau-Devonshire and Falk potentials limits their applicability to a small range of temperatures around the transformation region.

In building the strain energy density \( \mathcal{W} \) in the form (3.54), we first consider the construction of the isotherms \( \mathcal{W}_M \), \( \mathcal{W}_C \) and \( \mathcal{W}_A \), and then discuss the construction of the coefficient functions \( C_M \), \( C_C \) and \( C_A \):

### 3.4.3.1 Martensite Isotherm \( \mathcal{W}_M \)

At temperatures below the martensite finish temperature \( \theta_M \), the isothermal elastic strain energy is provided by the function \( \mathcal{W}_M(\gamma) \). The constraints placed on \( \mathcal{W}(\cdot, \theta) \)
for this temperature range ((3.37), (3.39) (3.41), (3.45), (3.48) and (3.49)) dictate the following criteria on the isotherm $\mathcal{W}_M$:

- $\mathcal{W}_M(\gamma)$ has two global minima of equal energy at the deformation gradient values $1 \pm \gamma_0$,
- $\mathcal{W}_M(1)$ does not constitute a minimum of any kind,
- the elastic moduli at the stable deformation gradients are met: $\rho_0 d^2 \mathcal{W}_M(1 \pm \gamma_0) = E_m$,
- the second derivative remains bounded: $d^2 \mathcal{W}_M(\gamma) < C$ as $\gamma \to \infty$.

These constraints are satisfied if the isotherm $\mathcal{W}_M$ meets the following guidelines:

- $\mathcal{W}_M(\gamma)$ is a polynomial of even degree at least four for $\gamma \in [1 - \gamma_0, 1 + \gamma_0]$,
- $\mathcal{W}_M(\gamma)$ is quadratic for $\gamma < 1 - \gamma_0$ and $\gamma > 1 + \gamma_0$,
- the stable deformation gradients have equal energy: $\mathcal{W}_M(1 \pm \gamma_0) = 0$,
- the stable deformation gradients are minima: $d \mathcal{W}_M(1 \pm \gamma_0) = 0$ and $d^2 \mathcal{W}_M(1 \pm \gamma_0) > 0$,
- the elastic moduli at the stable deformation gradients are met: $\rho_0 d^2 \mathcal{W}_M(1 \pm \gamma_0) = E_m$.

We construct the martensitic strain energy profile $\mathcal{W}_m(\gamma)$ according to the above guidelines in the following manner:

$$
\mathcal{W}_M(\gamma) = \begin{cases} 
\frac{E_m}{2\rho_0} (\gamma - 1 - \gamma_0)^2, & \gamma < 1 - \gamma_0, \\
\frac{E_m}{8\rho_0 \gamma_0^2} ((\gamma - 1)^4 - 2\gamma_0^2 (\gamma - 1)^2 + \gamma_0^4), & \gamma \in [1 - \gamma_0, 1 + \gamma_0], \\
\frac{E_m}{2\rho_0} (\gamma - 1 + \gamma_0)^2, & \gamma > 1 + \gamma_0.
\end{cases}
$$

(3.55)
3.4.3.2 Transition Isotherm $\mathcal{W}_C$

Similarly at the transition temperature $\theta_C$, the elastic strain energy is provided by $\mathcal{W}_C(\gamma)$. At this temperature the relevant constraints on $\mathcal{W}(\cdot, \theta_C)$ are (3.36), (3.37), (3.38), (3.46), (3.48) and (3.49), which enforce the following requirements on $\mathcal{W}_C$:

- $\mathcal{W}_C(\gamma)$ must have three global minima of equal energy at the deformation gradient values $1$, $1 - \gamma_0$, and $1 + \gamma_0$,
- the second derivative remains bounded: $d^2\mathcal{W}_C(\gamma) < C$ as $\gamma \to \infty$.

Again, these constraints are satisfied by the following guidelines on the construction of the isotherm $\mathcal{W}_C$:

- $\mathcal{W}_C(\gamma)$ is a polynomial of even degree at least six for $\gamma \in [1 - \gamma_0, 1 + \gamma_0]$,
- $\mathcal{W}_C(\gamma)$ is quadratic for $\gamma < 1 - \gamma_0$ and $\gamma > 1 + \gamma_0$,
- the stable deformation gradients have equal energy: $\mathcal{W}_C(1) = \mathcal{W}_C(1 - \gamma_0) = \mathcal{W}_C(1 + \gamma_0) = 0$,
- the second-derivative remains bounded: $d\mathcal{W}_C(\gamma) = 0$ for $\gamma \in \{1, 1 - \gamma_0, 1 + \gamma_0\}$,
- the isotherm is locally convex on the set of stable deformation gradients: $d^2\mathcal{W}_C(\gamma) > 0$ for $\gamma \in \{1, 1 - \gamma_0, 1 + \gamma_0\}$.

We satisfy these guidelines with the following choice of the transition strain energy profile

$$
\mathcal{W}_C(\gamma) = \begin{cases}
\frac{E_a + E_m}{4\rho_0} (\gamma - 1 - \gamma_0)^2, & \gamma < 1 - \gamma_0, \\
\frac{E_a + E_m}{16\rho_0} \left( (\gamma - 1)^6 - 2\gamma_0^2(\gamma - 1)^4 + \gamma_0^4(\gamma - 1)^2 \right), & \gamma \in [1 - \gamma_0, 1 + \gamma_0], \\
\frac{E_a + E_m}{4\rho_0} (\gamma - 1 + \gamma_0)^2, & \gamma > 1 + \gamma_0.
\end{cases}
$$

(3.56)
3.4.3.3 Austenite Isotherm $\mathcal{W}_A$

The final strain energy isotherm $\mathcal{W}_A(\gamma)$ applies to temperatures above the austenite finish temperature, $\theta_A$. In this temperature region, the constraints (3.36), (3.40), (3.42), (3.47), (3.48) and (3.49) on $\mathcal{W}(\cdot, \theta)$ translate into the following constraints on $\mathcal{W}_A$:

- $\mathcal{W}_A(\gamma)$ has exactly one global minimum at the deformation gradient value $\gamma = 1$,
- $\mathcal{W}_A(1 \pm \gamma_0)$ do not constitute minima of any kind,
- the elastic modulus at the reference configuration is met: $\rho_0 \, d^2\mathcal{W}_A(1) = E_a$,
- the second derivative remains bounded: $d^2\mathcal{W}_A(\gamma) < C$ as $\gamma \to \infty$.

These constraints are fulfilled by functions $\mathcal{W}_A$ constructed according to the following requirements:

- $\mathcal{W}_A(\gamma)$ is a polynomial of even degree at least two for $\gamma \in [1 - \gamma_0, 1 + \gamma_0]$,
- $\mathcal{W}_A(\gamma)$ is quadratic for $\gamma < 1 - \gamma_0$ and $\gamma > 1 + \gamma_0$,
- the austenitic deformation gradient corresponds to zero elastic stress: $\mathcal{W}_A(1) = 0$,
- the austenitic deformation gradient is a minimum: $d\mathcal{W}_A(1) = 0$,
- the martensitic deformation gradients are not minima: $d\mathcal{W}_A(1 \pm \gamma_0) \neq 0$, and
- the austenitic elastic modulus is met: $\rho_0 \, d^2\mathcal{W}_A(1) = E_a$.

These requirements are easily satisfied with the austenitic strain energy profile

$$\mathcal{W}_A(\gamma) = \frac{E_a}{2\rho_0} (\gamma - 1)^2.$$  \hspace{1cm} (3.57)
3.4.3.4 Coefficient Functions $C_M$, $C_C$ and $C_A$

At this point, the only remaining components of the strain energy density are the coefficient functions $C_M(\theta)$, $C_C(\theta)$ and $C_A(\theta)$ used to piece the isotherms $W_M$, $W_C$ and $W_A$ together. As mentioned previously, these coefficient functions are chosen to provide a smooth transition in the strain energy $W(\gamma, \theta)$ as the temperature changes, to ensure that $\partial_{\theta\theta}^2 W$ is small, and to satisfy the requirement that $\partial_{\theta\theta}^2 W(\gamma, \theta_C) = 0$. The following properties on the coefficients $C_*(\theta)$ ensure these requirements:

- $C_A(\theta) = 1$ for all $\theta \geq \theta_A$,
- $C_A(\theta) = 0$ for all $\theta \leq \theta_C$,
- $C_M(\theta) = 1$ for all $\theta \leq \theta_M$,
- $C_M(\theta) = 0$ for all $\theta \geq \theta_C$,
- $C_C(\theta) = 1$ for $\theta = \theta_C$,
- $C_C(\theta) = 0$ for all $\theta \geq \theta_A$, and $\theta \leq \theta_M$,
- $C_A(\theta) + C_M(\theta) + C_C(\theta) = 1$ for all $\theta$,
- $C_M(\theta)$, $C_C(\theta)$ and $C_A(\theta)$ are all $C^2$-functions of temperature, and
- $d^2 C_M(\theta)$, $d^2 C_C(\theta)$ and $d^2 C_A(\theta)$ are “small”; all are equal to 0 at $\theta_C$.

Piecewise cubic representations of these coefficient functions are chosen to satisfy the above properties. These are constructed using natural cubic splines on a regular temperature mesh containing points at the set of temperatures $\{\theta_M, \theta_C, \theta_A\}$. This choice is based on the fact that the natural cubic spline interpolant provides the curve with minimal approximate curvature intersecting the desired values [81]. This allows for the coefficient functions $C_*(\theta)$ to match the desired nodal values at various
temperatures, ensure that $C_*(\theta) \in C^2(\mathbb{R}^+)$, and at the same time minimize and match the second-derivative values as required above.

A picture of the resulting strain energy functional

$$W(\gamma, \theta) = C_M(\theta)W_M(\gamma) + C_C(\theta)W_C(\gamma) + C_A(\theta)W_A(\gamma)$$

is given in the following Figure 3.3.

![Strain Energy Density](image)

Figure 3.3: Strain energy functional $W(\gamma, \theta)$. Note the two global minima at $\gamma = 1 \pm \gamma_0$ for low temperatures, and the first-order phase transition to a single global minimum at $\gamma = 1$ for high temperatures. The strain energy density extends through the range of all possible $\theta$ and $\gamma$ values.

3.4.3.5 Remainder of the Helmholtz Free Energy

As discussed in the beginning of this Section 3.4.3, given the strain energy density $W(\gamma, \theta)$, we proceed to write the full form of the Helmholtz free energy as

$$\Psi(\gamma, \theta) = W(\gamma, \theta) + \bar{\Psi}(\theta),$$

(3.58)

so that we may examine the proper form of $\bar{\Psi}$ to satisfy the constraints (3.43)-(3.44), (3.48), and (3.50)-(3.53). From the thermodynamic constraints (3.50)-(3.53), the
internal energy and the entropy may be calculated from the Helmholtz free energy according to

\[ \varepsilon(\gamma, \theta) = \Psi(\gamma, \theta) - \theta \partial_\theta \Psi(\gamma, \theta), \]
\[ \eta(\gamma, \theta) = -\partial_\theta \Psi(\gamma, \theta). \]

Incorporating the expansion (3.58) into the above relations, the internal energy and entropy may be written

\[ \varepsilon(\gamma, \theta) = \mathcal{W}(\gamma, \theta) + \dot\Psi(\theta) - \theta(\partial_\theta \mathcal{W}(\gamma, \theta) + \partial_\theta \dot\Psi(\theta)), \]
\[ \eta(\gamma, \theta) = -\partial_\theta \mathcal{W}(\gamma, \theta) - \partial_\theta \dot\Psi(\theta). \] (3.59)

Traditional materials science literature ([21, 24, 26, 78, 45, 92, 94, 115, 116, 117]) dictates only one constant for the description of the specific heat. However the physics literature defines the specific heat by

\[ c_p(\gamma, \theta) = \theta \partial_\theta \eta(\gamma, \theta). \] (3.60)

Clearly this cannot be satisfied by a constant function at all points \((\gamma, \theta)\), if only for the simple fact that for positive temperatures the specific heat must be positive, while at a temperature of absolute zero the specific heat must also be zero [97]. Furthermore, since the Helmholtz free energy changes differently with temperature for various values of the deformation gradient, the specific heat cannot be independent of \(\gamma\) either. To this end, the constraints (3.43) and (3.44) have been formulated as chosen, where the specific heat constant is satisfied at the transition temperature,

\[ \theta_C \partial_\theta \eta(\gamma, \theta_C) = c_p, \] (3.61)

and the specific heat constant is only approximately satisfied elsewhere,

\[ \theta \partial_\theta \eta(\gamma, \theta) \approx c_p \quad \text{for} \quad \theta \neq \theta_C. \] (3.62)

This is also the cause for the previously discussed constraints on \(\partial^2_{\theta\theta} \mathcal{W}(\gamma, \theta)\), in which

\[ \partial^2_{\theta\theta} \mathcal{W}(\gamma, \theta_C) = 0 \quad \text{and} \quad \partial^2_{\theta\theta} \mathcal{W}(\gamma, \theta) \text{ is } "small" \text{ for } \theta \neq \theta_C.
With the above requirements on the strain energy $\mathcal{W}$, the specific heat conditions (3.61) and (3.62) may be met using the Helmholtz term $\hat{\Psi}$ having

$$-\theta \partial_{\theta \theta}^2 \hat{\Psi} (\gamma, \theta) = c_p.$$ 

Therefore, the remainder of the Helmholtz free energy is constructed as

$$\hat{\Psi}(\theta) = c_p(\theta - \theta \ln \theta) + D\theta + E.$$  \hspace{1cm} (3.63)

With this choice of the Helmholtz term $\hat{\Psi}$ along with (3.59), the internal energy and entropy are given by

$$\varepsilon(\gamma, \theta) = \mathcal{W}(\gamma, \theta) - \theta \partial_\theta \mathcal{W}(\gamma, \theta) + c_p \theta + E,$$

$$\eta(\gamma, \theta) = -\partial_\theta \mathcal{W}(\gamma, \theta) + c_p \ln \theta - D.$$ \hspace{1cm} (3.64)

Finally, the shifting constants $D$ and $E$ are chosen to satisfy Nernst's Theorem, which states that at a temperature of absolute zero, both the internal energy and entropy must be zero for stable states [97].

From the above construction, the full form of the proposed Helmholtz free energy is given by

$$\Psi(\gamma, \theta) = C_A(\theta) \mathcal{W}_A(\gamma) + C_M(\theta) \mathcal{W}_M(\gamma) + C_C(\theta) \mathcal{W}_C(\gamma)$$

$$+ c_p(\theta - \theta \ln \theta) + D\theta + E,$$ \hspace{1cm} (3.65)

where the functions $\mathcal{W}_*(\gamma)$ are given by equations (3.55), (3.56) and (3.57); the coefficient functions $C_*(\theta)$ are described above; the scalar $D$ is used to shift the entropy; and the scalar $E$ is used to shift the internal energy.

By construction, the free energy $\Psi$ satisfies the requirements set forth in (3.36)-(3.53), thereby dictating the physical behavior of shape memory alloy wires. In so doing, the free energy accounts for the phenomenological behavior of shape memory alloys through the first-order phase transformation. Moreover, the form (3.65) for the free energy is rich enough to allow for control over the material constants of particular
shape memory alloys. Therefore the general thermodynamic model derived in Section 3.2, with the Helmholtz free energy (3.65) at its core, should fully describe the nonlinear thermodynamic behavior of shape memory alloys described in Chapter 2. Chapter 6 provides computational results showing that indeed the proposed model fulfills this task. However, in order for this model to be usable, we must first understand the mathematical behavior of such nonlinear systems of partial differential equations (discussed in Chapter 4), and then treat their numerical solution (in Chapter 5).
Chapter 4

Related Theoretical Properties

This chapter provides a theoretical context for the solution and control of the proposed thermodynamic model (3.35), centered around the nonconvex Helmholtz free energy (3.65). Although there is presently no theory that can provide existence and uniqueness of solutions for our model of shape memory alloy wires, the body of literature surrounding similar dynamic systems is extensive. Therefore the current chapter first provides a characterization of the role our proposed model plays in the broader theory and then details aspects of that theory particularly relevant to the solution and control of nonlinear thermodynamic systems. Within these areas, this chapter serves as both a digest of external work and also as a venue for presenting our previous work relevant to the current modeling problem.

In Section 4.1, we discuss some of the background on general hyperbolic systems of conservation laws. We begin by defining the key characteristics of these systems. From there, we move to the topic of weak solutions to systems of conservation laws. Of key importance in this discussion is the multiplicity of possible weak solutions to these problems, and the corresponding use of entropy-entropy flux pairs as a selection principle to discern the proper weak entropy solutions to such systems. We end
this section by considering the standard Method of Vanishing Viscosity for scalar conservation laws, which serves as a methodical selection principle for attainment of weak entropy solutions. This method, though not proven to be applicable to the proposed modeling system, gives credence to the computational methods introduced in Chapter 5.

Section 4.2 then discusses some of the current theory specifically applicable to our proposed modeling problem. We first explain why the standard methods for solution and analysis of hyperbolic conservation laws fail for our model of shape memory alloys. This failure is centered around the degeneracy of our system of partial differential equations (3.35), where the degeneracy is a direct result of the non-convexity of the free energy functional $\Psi$. We then discuss the direct method for solving nonlinear partial differential equations. Specifically, this discussion examines the requirement of quasiconvexity on the system, necessary for the success of the direct method. Again, our shape memory alloy model fails this requirement, though in doing so it leads to the idea of finding measure-valued solutions to these problems, known as Young measures, which we have calculated in a few circumstances. This section then concludes with a discussion of the current theory most applicable to the proposed thermodynamic model. These theoretical results prove both existence and uniqueness to perturbed models that incorporate a significant level of artificial viscosity into the nonlinear thermodynamic model. Although the theory does not specifically discuss it, these viscosity solutions closely resemble the weak viscosity solutions used in the Method of Vanishing Viscosity. This close semblance directly leads to the computational solution method that we develop for our system of partial differential equations in Chapter 5.

The remaining section in this chapter deals with theoretical results pertaining to mechanical damping properties and thermal controllability of shape memory alloy wires. We first relate some recent work by Rohan Abeyaratne and James Knowles
on the theoretical damping properties arising from the formation and movement of phase boundaries between austenitic and martensitic regions in shape memory alloys. We then detail some of our current work in construction of simplified models and generalized controls for thermally-induced damping in shape memory alloy wires.

4.1 Strictly Hyperbolic Systems of Equations

The proposed thermodynamic model derived in Chapter 3 provides a nonlinear system of conservation laws. Although the nonlinearity of our model precludes the use of many standard methods for the solution of hyperbolic systems of equations, the computational approach we use to solve the model follows the general approach of these methods. Thus, in this section we discuss some of the properties of hyperbolic systems of equations, and a general approach to finding weak solutions to these systems.

4.1.1 General Information in One-Space Dimension

As discussed by Dafermos [41], Málek et al. [101] and Visintin [132], we first consider the space \( \Omega \times \mathbb{R}^+ = \mathbb{R} \times \mathbb{R}^+ \) containing typical points \((x,t)\), and the set of solutions \( U(x,t) \) to the canonical semilinear system of balance laws in one space dimension

\[
\partial_t U(x,t) + \partial_x F(U(x,t)) = 0,
\]

where \( U \in \mathcal{O}, \mathcal{O} \) is an open subset of \( \mathbb{R}^n \), and \( F \) is a given smooth map from \( \mathcal{O} \) to \( \mathbb{R}^n \); along with the initial conditions

\[
U(x, 0) = U_0(x), \quad x \in \mathbb{R}^n,
\]

and \( U_0 \) is a given bounded measurable function from \( \mathbb{R} \) to \( \mathcal{O} \).
Remark 4.1.1. We note that everything to be discussed in relation to the problem (4.1), (4.2) is extendible to the general inhomogeneous case [41, Chapter 4].

Definition 4.1.2. The semilinear system of balance laws (4.1) is called hyperbolic if for each function $U \in \mathcal{O}$ the $n \times n$ Jacobian matrix $DF(U)$ has real eigenvalues (characteristic speeds) $\lambda_1(U) \leq \cdots \leq \lambda_n(U)$ and is diagonalizable (has $n$ linearly independent eigenvectors $R_1(U), \cdots, R_n(U)$). The eigenvalue $\lambda_i$ of $DF(U)$ is called the $i$-characteristic speed of the system (4.1).

Waves traveling through the material move along these characteristics. Specifically, weak fronts propagate along these characteristics. Therefore when the matrix $DF$ has non-simple eigenvalues, solutions to the system (4.1) become extremely complex due to resonance of these weak fronts. For this reason, we further consider systems having simple eigenvalues only.

Definition 4.1.3. The system (4.1) is strictly hyperbolic if it is hyperbolic and the eigenvalues $\lambda_i(U)$ are distinct, meaning that for each $U \in \mathcal{O}$ the characteristics may be ordered so that $\lambda_1(U) < \cdots < \lambda_n(U)$.

4.1.2 Weak Solutions

Classical solutions to systems of conservation laws, i.e. bounded, locally Lipschitz functions $U$, defined on $\mathbb{R} \times [0, T)$, taking values in $\mathcal{O}$, and satisfying the problem (4.1), (4.2) almost everywhere $x \in \mathbb{R}$, typically only exist for strictly hyperbolic systems. General hyperbolic initial value problems cannot usually have a classical solution. This is a result of the fact that the classical solution must be constant along the characteristics. However, for general hyperbolic systems these characteristics can intersect in finite time for suitably smooth initial data $U_0$, causing discontinuities at the points of intersection [101]. We thus consider weak solutions to the general hyperbolic system (4.1). A weak solution $U$ to this problem on the interval $[0, T)$ is
a bounded measurable function from $\mathcal{H} \times [0, T)$ to $\mathcal{O}$ that satisfies (4.1), (4.2) in the sense of distributions, i.e.

$$
\int_0^T \int_{\mathcal{H}} [\partial_t \phi U + \partial_x \phi F(U)] \, dx \, dt + \int_{\mathcal{H}} \phi(x, 0) U_0(x) \, dx = 0,
$$

(4.3)

for every Lipschitz test function $\phi$ with compact support in $\mathcal{H} \times [0, T)$. A difficulty, however, with the consideration of these solutions is that there are usually multiple weak solutions satisfying the equation (4.3). Often, only one of these solutions is physically relevant [41, Section 4.2]. A means of overcoming such difficulties is through *entropy-entropy flux pairs*, which serve as a selection principle for screening such weak solutions.

If there exists a smooth function $\eta$ such that for all functions $U \in \mathcal{O}$ there is a nontrivial smooth function $q$ satisfying

$$
Dq(U) = D\eta(U) DF(U), \quad \text{and}
$$

$$
D^2 \eta(U) DF(U) = DF(U)^T D^2 \eta(U),
$$

(4.4)

then $\eta$ is called an *entropy* for the original system (4.1), and $q$ is called the *entropy flux* associated with $\eta$. The pair $(\eta, q)$ is known as an *entropy-entropy flux pair* for the system of conservation laws (4.1), and provide the companion balance law for the solution $U$,

$$
\partial_t \eta(U(x, t)) + \partial_x q(U(x, t)) = 0.
$$

(4.5)

Furthermore, any classical solution to the original system (4.1) is also a classical solution to the companion balance law (4.5) [41, Section 1.5].

**Remark 4.1.4.** The companion balance law (4.5) is related to the Second Law of thermodynamics (see Section 3.1), hence the function $\eta$ is referred to as an 'entropy' for the system (4.1). This quantity is not to be confused with the actual entropy density $\eta(\gamma, \theta)$, discussed in the derivation of the thermodynamic model (Chapter 3), which depends on the material and not on the system of balance laws under consideration.
Remark 4.1.5 (From Dafermos [41], Sections 1.5, 3.2 and 7.4). The entropy-entropy flux requirements (4.4) impose \(\frac{1}{2}n(n - 1)\) conditions on the entropy function \(\eta\). Therefore the problem of determining a nontrivial entropy-entropy flux pair for the problem (4.1) is overdetermined. Nontrivial companion balance laws are in fact only possible if the system (4.1) is symmetrizable, or equivalently, the entropy \(\eta(U)\) is uniformly convex:

\[
R_j(U)^T D^2 \eta(U) R_j(U) > 0, \quad j = 1, \ldots, n,
\]

where \(R_j\) is the eigenvector associated with the characteristic \(\lambda_j\).

In general, weak solutions \(U\) to the system (4.3) do not typically satisfy the companion balance law (4.5) (unless \(U\) is also a strong solution). Instead, a less restrictive form of the companion balance law is used to screen weak solutions. A weak solution \(U\) of the system (4.3) on the interval \([0, T]\) is considered admissible if it satisfies the inequality

\[
\int_0^T \int_R \left[ \partial_t \psi \eta(U(x,t)) + \partial_x \psi q(U(x,t)) \right] \, dx \, dt + \int_R \psi(x,0) \eta(U_0(x)) \, dx \geq 0, \quad (4.6)
\]

for all nonnegative Lipschitz continuous test functions \(\psi\) having compact support in \(R \times [0, T]\). Admissible weak solutions to (4.6) are often called entropy solutions to the original hyperbolic system (4.1).

4.1.3 Method of Vanishing Viscosity

As mentioned above, the hyperbolic system (4.1) must be equipped with a convex entropy in order to provide the companion balance law necessary for selection of admissible weak solutions to the problem (4.3). However, the requirements (4.4) for finding an entropy-entropy flux pair \((\eta, q)\) having uniformly convex \(\eta\) often fail in the case of nonlinear and semilinear hyperbolic systems (4.1), where the systems are
typically not strictly hyperbolic. For scalar conservation laws of this type, the Method of Vanishing Viscosity, first introduced by Hopf in 1950 [74] and Lax in 1954 [93], proves useful for finding admissible weak solutions (4.6).

Within this method, one examines the parabolic perturbation

$$\partial_t u(x,t) + \partial_x f(u(x,t)) = \mu u_{xx}(x,t),$$

(4.7)

where $u : \mathbb{R} \times [0,T) \to \mathbb{R}$ and $\mu > 0$, to the scalar hyperbolic equation

$$\partial_t u(x,t) + \partial_x f(u(x,t)) = 0.$$  \hspace{1cm} (4.8)

For $\mu > 0$ sufficiently large, the differential system (4.7) is parabolic. Therefore by standard theory for parabolic differential equations, there always exists a unique, smooth solution to the initial value problem (4.7), (4.2) for $t > 0$, so long as the functional $f \in C^1(\mathbb{R})$ and the initial data satisfies $U_0 \in L^\infty(\mathbb{R})$ (Ladyzhenskaya et al. [91], Málek et al. [101]). Such solutions $u_\mu$ to the initial value problem (4.7), (4.2) are called weak viscosity solutions. The impetus behind the Vanishing Viscosity method is therefore to consider the $\mu \searrow 0$ limiting solutions of the sequence $\{u_\mu\}_{\mu > 0}$.

Under the minimal assumptions $u_0 \in L^\infty(\mathbb{R})$ and $f \in C^1(\mathbb{R})$, Málek et al. have proven that indeed there exists a function $u \in L^\infty(\mathbb{R} \times [0,T))$ and a subsequence, still denoted $u_\mu$, such that [101, Theorem 4.62]

$$u_\mu \to u \text{ a.e. in } L^\infty(\mathbb{R} \times [0,T)), \text{ as } \mu \to 0_+.$$  

This limiting function $u$ is an entropy solution to the original scalar hyperbolic conservation law (4.8), (4.2) (see [101, Theorem 4.71]); furthermore the entropy solution $u$ is unique (see [101, Theorem 5.1]).

As a final remark on scalar conservation laws, Otto has shown that a similar approach using the Method of Vanishing Viscosity, along with a similarly-defined boundary entropy-entropy flux pair $(\eta, q)$, allows one to prove both existence and uniqueness
(in the sense of distributions) of weak entropy solutions \( u \in L^\infty(\Omega \times [0, T]) \) to the initial-boundary value problem

\[
\begin{align*}
\partial_t u(x, t) + \text{Div} f(u(x, t)) &= 0 \quad \text{in } \Omega \times (0, T), \\
\quad u(0, \cdot) &= u_0 \quad \text{in } \Omega, \\
\quad u(\cdot, t) &= u_D \quad \text{on } \partial \Omega \times [0, T),
\end{align*}
\]

where \( u_0 \in L^\infty(\Omega), \ u_D \in L^\infty(\partial \Omega), \ \Omega \subset \mathbb{R}^d \) is bounded and smooth, and with \( \partial \Omega \) the boundary of \( \Omega \) (see Otto [119, 120], Málek et al. [101, Theorem 7.28]).

We thus see that, at least in the case of scalar conservation laws, the Method of Vanishing Viscosity provides a clear, methodical approach for attaining the desired weak entropy solutions. In this way, the Method of Vanishing Viscosity serves as a selection principle in the same manner as the entropy-entropy flux pairs, essentially "weeding out" extraneous weak solutions to the original hyperbolic conservation law. As we see in the next Section 4.2, this method cannot be directly applied to the proposed model for the thermodynamic behavior of shape memory alloy wires; however, it gives credence to the computational approach that we introduce in Section 5.4 for solving our proposed model of shape memory alloy wires.

### 4.2 Degenerate Hyperbolic Systems

As discussed in Section 4.1.2 above, a key characteristic necessary for the standard theory to apply to a system is that there exists a corresponding entropy-entropy flux pair \((\eta, q)\) satisfying (4.4) that provides a companion balance law (4.5) for the original system of equations. However, in the case of the proposed system of partial differential equations modeling shape memory alloys,

\[
\begin{align*}
\rho_0 \ddot{u} &= \rho_0 \partial_{\gamma} \Psi (\gamma, \theta) + \alpha \partial_{xx}^2 \dot{u} + \rho_0 b \\
\rho_0 c_p \dot{\theta} &= \rho_0 \theta \partial_{\gamma}^2 \Psi \dot{\gamma} + \alpha \dot{\gamma}^2 + \kappa \partial_x (\gamma \partial_x \theta) + \rho_0 r,
\end{align*}
\]

(4.9)
the non-monotone structure of the stress $\sigma = \rho_0 \partial_\gamma \Psi$ disallows such entropies $\eta$ for small viscosity parameters $\alpha$. Because of this non-monotonicity, the characteristics $\lambda_i$ associated with the above system are no longer distinct, allowing for nonlinear resonance of wave and shock fronts.

The set of equilibria to the viscoelastic equation (from (4.9)) alone is typically geometrically very complicated: it forms a noncompact, infinite-dimensional, non-smooth, multiply-connected set. The associated dynamics are not smooth either: the dynamics of the linearized equation and the nonlinear equation are topologically different. The set of stable equilibria to the isothermal problem,

$$\{ u \in W^{1,\infty} : D\Psi(\nabla u(x)) = 0, D^2\Psi(\nabla u(x)) > 0 \text{ a.e. in } \Omega \},$$

and the set of unstable equilibria (which differs from the stable set by dropping the requirement of positive definitiveness of the elastic module $D^2\Psi$) are noncompact in the $W^{1,p}(\Omega)$ topology. Moreover, the weak-$*$ closure in $W^{1,\infty}$ of either set contains the unit ball [60]. In the language of dynamical systems, the lack of compactness and regularity in the equilibrium problem prevents the equations from possessing finite-dimensional attracting sets or inertial manifolds.

A further difficulty is that the equilibrium states are not isolated, but form multiply-connected sets in $W^{1,p}$ [60]. Even if precompactness of an orbit is known, it is not clear whether the orbit stabilizes to a unique limiting pattern. Using once more the language of dynamical systems, it would not be clear whether the $\omega$-limit sets are singletons for given initial data. A recent result by Rybka and Hoffmann indicates that this is true if we take capillarity into account (i.e. adding in $\Delta^2 u$ to the wave equation in (4.9)) [125]. However, analysis shows that since the second variation of the energy degenerates at the equilibria, the theory of Hale-Raugel (developed for proving convergence even if manifolds of equilibria are present) is not applicable to this class of problems [71].
Moreover, as time goes to infinity, typical solutions may be expected to converge exponentially fast to a unique traveling wave. This is true for dynamical systems whose equilibria form a smooth hyperbolic manifold (see [69, 70, 71, 122, 123]). Again, these results are not applicable in the present case because the dynamics of the first-order phase transition are not differentiable: we note that the mapping $\nabla u \mapsto D\Psi(\nabla u)$ is not Fréchet differentiable as a map from $L^p$ to $L^p$ with $1 \leq p < \infty$ when $D\Psi(\cdot)$ is not affine.

Thus we see that the thermodynamic system (4.9) poses unique challenges in both its understanding and solution. The standard approaches used for analysis of the linear wave equation, for example, cannot apply to the proposed model because of its degeneracy between stable elastic regimes. Hence, we next consider alternate methods for the solution of our thermodynamic system (4.9).

4.2.1 Quasiconvexity Approaches

We next consider the following general approach to the analysis and solution of nonlinear partial differential equations. The so-called direct method of the calculus of variations is a standard approach that attempts to solve the differential system through minimization of a related functional. In this approach, the nonlinear differential equation under consideration is first written as the Euler-Lagrange system associated with a potential function, say $\Phi(x, u, Du)$. That is to say, there exists a potential function $\Phi(x, u, Du)$ such that weak solutions of the original nonlinear differential system exist as critical points of the functional

$$I(u) = \int_\Omega \Phi(x, u, Du) \, dX,$$

where $u \in \mathcal{A} = \{v \in W^{1,1}(\Omega) : v|_{\partial\Omega} = g\}; g$ provides the boundary conditions; and $\Omega \subset \mathbb{R}^n$ is the bounded, open, connected domain having Lipschitz boundary $\partial\Omega$ for the original differential system. Then if the functional $I$ is
1. coercive: there exist constants $c_0 > 0$, $p > 1$ and $c \in \mathbb{R}$ such that for all $x \in \Omega$, $u \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times n}$, $\Phi$ satisfies

$$\Phi(x, u, A) \geq c_0 |A|^p - c_1,$$

and

2. bounded from below: there exists a constant $l > -\infty$ such that $l = \inf_A I$,

one may attempt to solve the original system through minimization of $I$ (see Dacorogna [39, 40] and Visintin [132]).

We consider a minimizing sequence $\{u^j\}_{j=1}^\infty \subset \mathcal{A}$ of the functional $I$, i.e.

$$\lim_{j \to \infty} I(u^j) = l.$$

If $\{u^j\}_{j=1}^\infty$ is a bounded sequence in the reflexive Banach space $W^{1,p}(\Omega)$, and if the function $g$ is also bounded, then there exists a weakly convergent subsequence $\{u^{j_k}\}_{j_k}$ such that $u^{j_k} \rightharpoonup u$ in $W^{1,p}(\Omega)$, and $u \in \mathcal{A}$. It turns out that the salient feature of such problems to guarantee that the limit $u$ is indeed a minimizer of $I$ in $\mathcal{A}$ (i.e. $I(u) = l$), is the condition that the functional $I$ must be sequentially weakly lower semicontinuous in the space $W^{1,p}(\Omega)$ (see Ball [14, 15, 16, 17], Dacorogna [39, 40], Milton [107], Pedregal [121]). That is to say, $I$ must satisfy the following implication:

$$u^j \rightharpoonup u \text{ in } W^{1,p}(\Omega) \implies I(u) \leq \lim_{j \to \infty} \inf I(u^j). \quad (4.11)$$

In his work that gave rise to this entire field, Morrey proved that sequential weak lower semicontinuity of the functional $I$ rests solely on the requirement that the potential $\Phi(x, u, Du)$ must be quasiconvex for each $x$, $u$ [109, 110]. Quasiconvexity of the potential $\Phi$ is defined in the following manner.

**Definition 4.2.1.** Let $1 \leq p \leq \infty$. Then $\Phi$ is $W^{1,p}$ quasiconvex at the matrix $A \in \mathbb{R}^{m \times n}$ if for some bounded open set $E \subset \mathbb{R}^n$ with $\mathcal{L}^n(\partial E) = 0$,

$$\Phi(x, u, A) \leq \frac{1}{\mathcal{L}^n(E)} \int_E \Phi(A + D\phi(x)) \, dx$$
for all $\phi \in W^{1,p}_0(E; \mathbb{R}^m)$, whenever the integral exists. The functional $\Phi$ is $W^{1,p}$ quasiconvex if $\Phi$ is $W^{1,p}$ quasiconvex at every $A$.

Thus quasiconvexity (a multi-dimensional analog of standard convexity in one dimension) may be seen as the key characteristic of a potential $\Phi$ to guarantee solutions of general nonlinear partial differential equations using the direct method.

**Remark 4.2.2.** Quasiconvexity is in fact the appropriate level of convexity required of the entropy $\eta$ in entropy-entropy flux pairs for general systems of hyperbolic conservation laws (see Dafermos [41]).

**Remark 4.2.3.** In the current situation of modeling solid-state phase transitions for shape memory alloys, the convexity properties of the nonlinear problem match those of the Helmholtz free energy $\Psi$. However, in the case of these phase transitions, the free energy is not even rank-one convex, and for one-dimensional materials the free energy is entirely nonconvex [19, 20, 84, 83, 99, 107]. We note the one way implications: convex $\Rightarrow$ polyconvex $\Rightarrow$ quasiconvex $\Rightarrow$ rank-one convex (see Dacorogna and Marcellini [40]).

This lack of quasiconvexity then results in approximating sequences $\{u^j\}$ that converge weakly but not strongly to the function $u$. However, although such strong convergence fails, the Young Measures associated with gradients of the minimizing sequence indeed converge to the Young Measure associated with the limiting state $u$. This may be more precisely stated as

**Theorem 4.2.4.** Let $\{Du^j\} \subset \mathbb{R}^{m \times n}$ be a sequence of gradients bounded in $L^p$, $1 \leq p \leq \infty$. Then there exists a family $\{\mu_x\}_{x \in \Omega}$ of probability measures on $\mathbb{R}^{m \times n}$, depending measurably on $x$, such that for a subsequence $\{Du^{j_k}\}$,

$$\Phi(Du^{j_k}) \rightharpoonup \int_{\mathbb{R}^{m \times n}} \Phi(A) d\mu_x(A) \quad \text{in} \quad L^1(E)$$
for any continuous $\Phi : \mathbb{R}^{m \times n} \to \mathbb{R}$ and any open $E \subset \Omega$, whenever this weak limit exists. The family $\{\mu_x\}_{x \in \Omega}$ is called a $W^{1,p}$ gradient Young Measure (see Ball [14], Pedregal [121, Theorem 2.2]).

Moreover, although these Young Measures cannot sufficiently describe the small-scale dynamics as would be possible through a minimizing solution to the original nonlinear partial differential equation, they do sufficiently describe the dynamics of such systems on larger spatial scales (see Alberti and S. Müller [7]).

These results then lead to an approach currently under intense investigation. This approach attempts quasiconvexification of the potential function $\Phi$, i.e. construction of a quasiconvex potential function $\tilde{\Phi}$ having the same minima as $\Phi$ but without the various energy barriers inhibiting quasiconvexity (see Alberti and S. Müller [7], Ball [14], Bhattacharya and Dolzmann [22, 23], Milton [107], Pipkin [124]). There are a number of limitations to the use of these approaches for the problem (4.9):

1. It is still an open problem to find a general method for the quasiconvexification of a given potential function $\Phi$. Although means have been found for a few such rank-one potential functions, general methods often over-convexify the potential, rendering any semblance to the small-scale dynamics null (see Ball [17], Bhattacharya and Dolzmann [22, 23] and Milton [107]).

2. At the present time, these approaches are applied to only the degenerate wave equation

$$u_{tt} = \text{Div} \sigma(Du),$$

(4.12)

where $\sigma = \partial_{Du} \Phi(Du)$ is non-monotone due to the multi-well structure of the potential $\Phi$.

3. The large-scale information provided through quasiconvexification of the thermodynamic model (4.9) would omit many of the relevant nonlinear properties
required from the thermodynamic model. Namely, such behaviors as the hysteresis and latent heat of transformation would be lost by quasiconvexification.

Therefore quasiconvexification of the Helmholtz free energy $\Psi$ in the proposed model (4.9) does not provide for an alternative approach to the desired computation of the nonlinear thermodynamic behavior of shape memory alloys.

4.2.2 Direct Calculations of the Young Measure

An alternate approach than quasiconvexification for the computation of these Young measures has been in the search for a perturbation of the original free energy that bypasses the difficulties associated with non-quasiconvexity. Ball et al. [18] and Friesecke and McLeod [60, 61] have shown that this lack of quasiconvexity, and the resulting lack of weak lower semicontinuity, associated with the multi-well free energy $\Psi$ result in the failure of the direct method for the computation of Young Measures. This is because the potential function used in this direct method has a large number of local minimizers, which lead the corresponding minimization dynamics toward unphysical asymptotic states. Moreover, the resulting asymptotic state directly depends on the initial guess used in the minimization.

Therefore, methods for the direct calculation of such Young measure solutions consider the construction of alternate potential functions that avoid local minima [33, 35, 36, 37]. These methods look to find an perturbed potential function $\Phi(x, u, Du)$, whose critical points still satisfy the original nonlinear differential system, but do not have local minima in the class of weakly-differentiable functions.

In this section we present some of our work on the search for such a perturbation in the setting of an idealized, one-dimensional model of the martensitic structures at an austenite-martensite interface in a shape memory plate. The results shown
here are largely taken from our paper "The non-local relaxation of non-attainable
differential inclusions using a subgrid projection method: one-dimensional theory
and applications" [37]. This paper deals with the computation of Young measures
in the one-dimensional case having normalized martensitic displacement gradients at
the values \( \gamma \in \{-1, 1\} \), austenitic displacement gradient at \( \gamma = 0 \), and macroscopic
shape of the austenitic interface given by \( f \).

4.2.2.1 First Order Constrained Partial Differential Inclusions

We build stochastic processes by which we can approach solution(s) of the following
problem.

Problem 4.2.5 (First Order Constrained Partial Differential Inclusion). Let
\( f \in W^{1, \infty}(0, 1) \). Find a function(s) \( u \in W^{1, \infty}(0, 1) \) such that

\[
\begin{align*}
  u'(x) & \in \mathcal{A} = \{-1, 1\}, \quad \text{for almost all } x \in (0, 1), \\
  u(x) & = f(x), \quad \text{for almost all } x \in (0, 1).
\end{align*}
\] (4.13)

Solutions to this problem can be determined by investigation of the infimum of the functional

\[
I(u) \overset{\text{def}}{=} \int_0^1 \mathcal{W}(u, u') \, dx,
\] (4.14)

where the potential \( \mathcal{W} \) is minimized by functions having \( u = f \) and \( u' \in \{-1, 1\}, \)

\[
\mathcal{W}(u, u') = \frac{1}{4} \left( (u')^2 - 1 \right)^2 + (u - f)^2.
\] (4.15)

For certain combinations of functions \( f \) and sets of allowable gradients (here given by
\( \{-1, 1\} \)), this problem has the non-attainment structure discussed in Section 4.2.1:
uniformly bounded weak derivatives of any minimizing sequence \( \{u_n\}_{n \in \mathbb{N}} \) of the func-
tional \( I(u) \) converge weakly--* in \( L^\infty(0, 1) \) to its expected value (weak--* limit) \( f' \), but
they do not converge strongly to any function. The weak-* convergence is a convergence in the sense of averages. Namely,
\[ \int_{a}^{b} u_n'(x) \, dx \to \int_{a}^{b} f'(x) \, dx, \quad \text{for any } a, b \in (0, 1). \]
The composition of the expected value \( f' \) can be recovered by the structure of the Young measure \( \mu_{x, u_n'} \). Namely,
\[ f'(x) = \lim_{n \to \infty} \int_{\mathbb{R}} y \, d\mu_{x, u_n'}(y). \]
The main result of this Section 4.2.2, proven in Theorem 4.2.18, shows that the application of a Steepest Descent Algorithm to the functional
\[ I_{\tau_3}(u, \Delta) \overset{\text{def}}{=} \frac{1}{2} \int_{0}^{1} |u(x)|^2 \, dx + \frac{1}{4} \int_{0}^{1} |u'(x)^2 - 1|^2 \, dx + \Delta \sum_{Q \in \tau_3} \int_{Q} \frac{1}{2} |u'(x) - \bar{u}'(x)|^2 \, dx \]
(4.16)
yields sequences \( \{u_n\}_{n \in \mathbb{N}} \) relaxing to the infimum of the functional (4.14) as \( \text{card } \tau_3 \to +\infty. \)

**Remark 4.2.6.** For the set of allowable gradients \( A = \{-1, 1\} \) currently under consideration, the Problem 4.2.5 has a solution \( u \in W^{1,\infty}(0, 1) \) if there exists a function \( v \in C^0(0, 1) \) such that \( f(x) = v(x) \) and \( |v'(x)| = 1 \) a.e. for \( x \in (0, 1) \).

The previous Remark 4.2.6 indicates that we can solve Problem 4.2.5 if \( f'(x) \in A \) almost everywhere in \( (0, 1) \). Next, we turn to the case in which the function \( f \) has its derivative strictly inside the convex hull of \( A \) (i.e. \( (-1, 1) \) in the present case). This turns out to be a more delicate situation.

**Theorem 4.2.7.** Let \( f \in W^{1,\infty}(0, 1) \), and let us assume that
\[ |f'(x)| < 1, \quad \text{almost everywhere in } (0, 1). \]
(4.17)
Then the Problem 4.2.5 has no solution, but for any \( \epsilon > 0 \) there exists \( u_\epsilon \in W^{1,\infty}(0, 1) \) such that
\[ u_\epsilon' \in \{-1, 1\}, \quad \text{a.e. in } (0, 1), \quad \text{and} \]
\[ \|u_\epsilon - f\|_{L^\infty(0, 1)} < \epsilon. \]
(4.18)
Moreover, for any continuous function $h = h(x)$ such that $h(u'_\epsilon) \rightharpoonup g$, weakly in $L^1(0, 1)$, as $\epsilon \to 0_+$, we have
\[ g(x) = \int_{-\infty}^{+\infty} h(y) d\mu_{x, u'_\epsilon}(y), \] (4.19)
where $\mu_{x, u'_\epsilon} = \lambda(x)\delta_{-1} + (1 - \lambda(x))\delta_{+1}$; $\delta_{\pm 1}$ denotes the Dirac measures on $\mathbb{R}$ giving unit mass to the points $\pm 1$; and $\lambda(x) = \frac{1}{2}(1 - f'(x))$ a.e. in the interval $[0, 1]$.

Proof. We split the proof into three parts.

The Problem 4.2.5 has no solution:
Suppose that the density $\mathcal{W}(u, u') = 0$ a.e. in $(0, 1)$, for $\mathcal{W}$ from (4.15). Then
\[ ((u')^2 - 1) = 0 \quad \text{and} \quad (u - f)^2 = 0 \quad \text{a.e. in } (0, 1). \]
Thus $u = f$ a.e. in $(0, 1)$, and since they agree pointwise, $u' = f'$ a.e. in $(0, 1)$. Thus $|u'| < 1$ a.e. in $(0, 1)$, which fails the supposition that $((u')^2 - 1)^2 = 0$. Therefore $\mathcal{W}(u, u') \neq 0$ a.e. for any function $u \in W^{1, \infty}(0, 1)$.

Approximability:
Choosing $\epsilon > 0$, there exists a continuous piecewise affine function $\phi(x)$ defined on $(0, 1)$ such that
\[ ||\phi - f||_{W^{1, \infty}(0, 1)} < \frac{\epsilon}{2}, \] [39]. Thus, in particular, $|\phi'| \leq 1$ almost everywhere in $(0, 1)$. Since $\phi$ is piecewise affine, partition $(0, 1)$ into disjoint intervals $Q_j = (a_j, b_j)$ such that $\phi$ is affine on each interval $Q_j$. On each $Q_j$, we have from [40] that there exists $u^{j}_\epsilon \in W^{1, \infty}(Q_j)$ such that
\[ ||u^{j}_\epsilon - \phi||_{W^{1, \infty}(Q_j, b_j)} \leq \frac{\epsilon}{2}, \]
\[ u^{j}_\epsilon \in \{-1, 1\}, \quad \text{a.e. in } (a_j, b_j) \quad \text{and} \]
\[ u^{j}_\epsilon(a_j) = \phi(a_j), \quad u^{j}_\epsilon(b_j) = \phi(b_j). \] (4.20)
Let $u_\epsilon = u^j_\epsilon$ on each interval $Q_j$. Note $u_\epsilon$ is continuous, since $\phi(b_j) = \phi(a_{j+1})$ and $|u_\epsilon'| = 1$ a.e. in $(0, 1)$, thus $u_\epsilon \in W^{1, \infty}(0, 1)$. Furthermore
\[ ||u_\epsilon - f||_{W^{1, \infty}(0, 1)} = ||u_\epsilon - \phi + \phi - f||_{W^{1, \infty}(0, 1)} \]
\[ \leq ||u_\epsilon - \phi||_{W^{1, \infty}(0, 1)} + ||\phi - f||_{W^{1, \infty}(0, 1)} < \epsilon. \] (4.21)
Weak-* Convergence:

In order to verify (4.19) we apply [121, Theorem 2.2]. First, we have for any \( a, b \in (0, 1) \)
\[
\left| \int_a^b u'_\varepsilon(x) - f'(x) \, dx \right| \leq 2 \| u_\varepsilon - f \|_{L^\infty(0, 1)} < 2\varepsilon. \tag{4.22}
\]
Hence from [121, Proposition 2.10],
\[
u'_\varepsilon(x) \rightharpoonup f'(x), \quad \text{weakly-* in } L^\infty(0, 1). \tag{4.23}
\]

Thus, [[121], Theorem 2.2] guarantees that for any sequence \( \varepsilon_n \to 0 \), there is a subsequence \( \varepsilon_{n_k} \) such that there exists a Young measure \( \mu_{x,u'_{\varepsilon_{n_k}}} \) satisfying (4.19) for any
continuous function \( h \). We first identify the Young measure and show it is independent
of the original sequence \( \varepsilon_n \). It then follows that (4.19) holds in general by a
standard argument. In order to identify the measure, we observe that \( u'_\varepsilon \in \{-1, 1\} \)
almost everywhere. Hence
\[
supp \mu_{x,u'_\varepsilon} \subseteq \{-1, 1\}. \tag{4.24}
\]

Also, \( \mu_{x,u'_\varepsilon} \) depends measurably on \( x \). Thus there exists a function \( \lambda = \lambda(x) \in L^\infty(0, 1) \) such that
\[
\mu_{x,u'_\varepsilon} = \lambda(x) \delta_{-1} + (1 - \lambda(x)) \delta_{+1}. \tag{4.25}
\]
The representation (4.19) of the weak limit yields for \( h(x) = x \) and for any \( a, b \in (0, 1) \)
\[
\left| \int_a^b f'(x) \, dx \right| = \left| \int_a^b \int_{\mathbb{R}} y \, d\mu_{x,u'_\varepsilon}(y) \, dx \right|
\]
\[
= \left| \int_a^b \lambda(x) \int_{\mathbb{R}} y \, d\delta_{-1}(y) \, dx \right| + \left| \int_a^b (1 - \lambda(x)) \int_{\mathbb{R}} y \, d\delta_{+1}(y) \, dx \right| \tag{4.26}
\]
\[
= \left| \int_a^b (1 - 2\lambda(x)) \, dx \right|.
\]
Therefore \( \lambda(x) = \frac{1}{2}(1 - f'(x)) \) follows from (4.26) and Lebesgue's Differentiation
Theorem. \( \square \)
Remark 4.2.8 ([121], Page 15). The Young measure $\mu_{x,u^e}$ in (4.19) can be reconstructed from $u^e$ by post-processing. Namely,

$$\mu_{x,u^e}((\infty,y]) = \lim_{R \to 0^+} \lim_{\epsilon \to 0^+} \frac{\text{meas}(\{z \in B_R(x) \mid u^e(z) \leq y\})}{\text{meas}(B_R(x))}. \tag{4.27}$$

In particular, we can compute the coefficients $\lambda(x)$ and $1 - \lambda(x)$ as follows

$$\lambda(x) = \lim_{r \to 0^+} \lim_{R \to 0^+, \epsilon \to 0^+} \frac{\text{meas}(\{z \in B_R(x) \mid |u^e(z) + 1| \leq r\})}{\text{meas}(B_R(x))},$$

$$1 - \lambda(x) = \lim_{r \to 0^+} \lim_{R \to 0^+, \epsilon \to 0^+} \frac{\text{meas}(\{z \in B_R(x) \mid |u^e(z) - 1| \leq r\})}{\text{meas}(B_R(x))}. \tag{4.28}$$

The probabilities $\lambda(x)$ and $1 - \lambda(x)$ represent local meso-scopic volume fractions. These quantities are important for constitutive modeling of martensitic thermodynamics, as discussed in Sections 2.4 and 4.2.1.

4.2.2.2 Subgrid Projection Method

As alluded to in the introduction of Section 4.2.2, we wish to find a perturbation to the potential for guiding these minimizing sequences $u^e$ in the computation of the proper Young measures $\mu_x$. For this purpose, we introduce the following Subgrid Projection Method.

We assume that the functional $\mathcal{W}$ corresponding to the equilibrium state of the system described in Problem 4.2.5 has the form [2, 79, 80, 85]

$$\mathcal{W}(u,u^e) = \mathcal{W}_{\text{macro}}(u) + \mathcal{W}_{\text{micro}}(u^e) + \mathcal{W}_{\text{meso}}(u^e). \tag{4.29}$$

We define the mesoscaling term $\mathcal{W}_{\text{meso}}$ as follows. First, we construct $\bar{u}(x)$ defined by

$$\bar{u}(x) = u \left( a_i + \frac{\delta_i}{2} \right) + \delta_i u \left( \frac{x - a_i}{\delta_i} \right), \quad x \in Q_i = (a_i, a_i + \delta_i). \tag{4.30}$$

The points $\{a_i\}$ and the partition $\{\delta_i\}$ form a decomposition $\tau_0$ of $(0,1)$. We denote

$$\delta = \max_i \delta_i. \quad \text{We assume that } \min \delta_i > 0, \ i = 1, 2, \ldots, N, \text{ and we assume that}$$
\( (0, 1) = \bigcup_{i=1}^{N} Q_i, Q_i \in \tau_\delta, Q_i \cap Q_j = \emptyset \text{ if } i \neq j. \) The function \( \bar{u} \) is not continuous at \( x = a_i. \) The scaled projection \( \bar{u}(x) \) copies the entire function \( u(x) \) onto each interval \( Q_i. \) Given a function \( u \in C^1(0, 1), \) we define the mesoscaling term pointwise by
\[
W_{\text{meso}}(u')(x) \overset{\text{def}}{=} \frac{1}{2} |u'(x) - \bar{u}'(x)|^2, \quad x \neq a_i.
\] (4.31)

Consequently, the “total” mesoscale energy is given by
\[
\sum_{Q \in \tau_\delta} \int_Q \frac{1}{2} |u'(x) - \bar{u}'(x)|^2 \, dx
\] (4.32)

We presume that a desirable property of the \( \int W_{\text{meso}} \, dx \) contribution should be that its infimum is attained on “white noise” processes. Translating this into the presented framework,
\[
\inf_{\tau_\delta} \inf_{V \in \mathcal{V}} \sum_{Q \in \tau_\delta} \int_Q |u'(x) - \bar{u}'(x)|^2 \, dx
\] (4.33)

should provide a variational principle for the “white noise”. The first infimum is taken over a countable collection of partitions \( \tau_\delta \) of \( (0, 1), \) and \( V \) is a suitable admissible space determined by the original problem. By “white noise”, we mean a sequence of discrete time stochastic processes \( Z_{n,i}, n = 1, 2, \ldots, i = 1, 2, \ldots, n, \) satisfying
\[
\mathcal{E}(Z_{n,i}) = 0, \quad \text{and} \quad \mathcal{E}(Z_{n,i}Z_{n,j}) = \begin{cases} 1, & \text{if } j = i, \\ 0, & \text{if } j \neq i, \end{cases} \text{ for all } n, i.
\] (4.34)

where \( \mathcal{E}(\cdot) \) is the abstract Expectation operator, i.e. integration over the probability space \( (\Omega, \mathcal{F}, P) \) where \( \Omega \) is the underlying set, \( \mathcal{F} \) is the \( \sigma \)-field of subsets of \( \Omega, \) and \( P \) is the probability. We assume all random variables \( Z_{n,i} \) to be Borel or Lebesgue measurable mappings defined on \( \Omega, \) and then
\[
\mathcal{E}(Z_{n,i}) = \int_{\Omega} Z_{n,i} \, dP.
\]

As is customary in probability, we suppress the argument of the random variable. For each \( n, \) we consider the stochastic process with continuous time index given by
\[
z_n(x) = Z_{n,\lfloor nx \rfloor}, \quad n = 1, 2, \ldots,
\] (4.35)
that is, on each interval \([(i-1)/n, i/n)\), \(z_n(x) = Z_{n,i}\). It is easy to verify then that

\[
\lim_{n \to \infty} \mathcal{E}(z(x)) = 0, \quad \text{and} \quad \lim_{n \to \infty} \mathcal{E}(z(x)z(x+\tau)) = \begin{cases} 1, & \text{if } \tau = 0, \\
0, & \text{if } \tau \neq 0, \end{cases} \tag{4.36}
\]

We further assume that the sequence of processes \(z_n\) considered here satisfies the Ergodic property, which implies that integration over the probability space can be replaced with averaging over "time," which is represented in our framework by the space variable \(x\). We note that the Ergodic property implicitly assumes only "short distance" interactions. The application of the Ergodic hypothesis allows for an exchange of the operator

\[
\mathcal{E}(z(x)) = \int_{\Omega} z(x) dP(x) \tag{4.37}
\]

with the averaging

\[
E_{(a,b)}[z(x)] = \frac{1}{b-a} \int_{a}^{b} z(x) dx, \quad a, b \in (0, 1). \tag{4.38}
\]

The hypothesis of ergodicity tells us that \(\lim_{n \to \infty} E_{(a,b)}[z_n(x)] = 0\) almost surely (abbreviated a.s., and meaning except on a set of probability 0), and similarly for \(E_{(a,b)}[z_n(x)z_n(x+\tau)]\). This motivates the following definition, which provides a finer characterization of weakly convergent sequences.

**Definition 4.2.9.** We say that a sequence \(\{u_\epsilon\}_{\epsilon > 0} \subset W^{1,\infty}(0,1), u'_\epsilon \in (-1,1)\) almost everywhere in \((0,1)\), that converges weakly-* in \(W^{1,\infty}(0,1)\) to a given Lipschitz continuous function \(f\), becomes asymptotically weak white noise if the following three conditions hold true. Let

\[
z_\epsilon(x) \overset{\text{def}}{=} \frac{u'_\epsilon(x) - f'(x)}{\sqrt{4\lambda(x)(1-\lambda(x))}}, \quad \lambda(x) = \frac{1}{2} (1 - f'(x)) \tag{4.39}
\]

The first condition is that for all \(a, b \in (0,1)\),

\[
E_{(a,b)}[z_\epsilon(x)] = \frac{1}{b-a} \int_{a}^{b} z_\epsilon(x) dx \to 0, \quad \text{as } \epsilon \to 0+.. \tag{4.40}
\]
The second and third conditions are that for almost all $\tau \in \mathbb{R}$ the autocorrelation function $E[\cdot]$ has the following two properties

$$E[z_{\varepsilon}(x) z_{\varepsilon}(x + \tau)] = \frac{1}{b - a} \int_a^b z_{\varepsilon}(x) z_{\varepsilon}(x + \tau) \, dx \to \begin{cases} 1, & \text{if } \tau = 0, \\
0, & \text{if } \tau \neq 0, \end{cases}$$ \hspace{1cm} (4.41)

for all $a, b \in (0, 1)$, as $\varepsilon \to 0$. We assume that the functions $u'_\varepsilon$ are extended periodically onto $\mathbb{R}$, in order for $u(x + \tau)$ to be defined for $x + \tau \notin (0, 1)$.

**Remark 4.2.10.** (i) We note that $z_{\varepsilon}(x)$ is (approximately) the standardized form of the sequence $\{u'_\varepsilon - f'\}_{\varepsilon > 0}$, where $f'$ is the weak-* limit of the sequence $\{u_\varepsilon\}_{\varepsilon > 0}$. Standardization in probability theory refers to subtracting off the mean and dividing by the standard deviation. The asymptotic mean of $u'_\varepsilon$ is $f'$ and the asymptotic standard deviation as given by the Young measure appears in the denominator of (4.39).

(ii) The Definition 4.2.9 can be immediately generalized for sequences of functions with derivatives other than $+1$ or $-1$ by changing the standardization in (4.39) appropriately.

(iii) The approximation sequences $\{u_\varepsilon\}_{\varepsilon > 0}$ which become asymptotically weak white noise must locally have very irregular structure. In other words, the approximating functions $u_\varepsilon$ which become asymptotically weak white noise in the sense of Definition 4.2.9 cannot periodically change the sign of their derivatives.

The following Lemma provides a lower bound for the mesoscaling term.

**Lemma 4.2.11.** Let $u_\varepsilon = u_\varepsilon(x)$ be a continuous piecewise affine approximation of a positive function $f$, $|f'(x)| < 1$, for a. a. $x \in (0, 1)$, $f(0) = f(1) = 0$, such that $u'_\varepsilon \in \{-1, 1\}$. Then for every $\varepsilon > 0$ there exists a partition $\tau_\varepsilon$ such that

$$\sum_{Q \in \tau_\varepsilon} \int_Q \left| \frac{1}{2} u'_\varepsilon(x) - \overline{u'_\varepsilon}(x) \right|^2 \, dx \geq 1 - 2 \|u_\varepsilon - f\|_{L^\infty(0,1)} - \sum_{Q \in \tau_\varepsilon} \int_Q f'(y) \bar{f}'(y) \, dy,$$ \hspace{1cm} (4.42)
where the scaled projections $\overline{u}_\epsilon$, $\overline{f}$ are computed with respect to the decomposition $\tau_\delta$.

Proof. Details of the proof are provided in [37, Lemma 3.3]. □

The next Theorem shows that (4.33) does not establish in general a variational principle for weak white noise if $\overline{u}_\epsilon$ are computed with respect to $\tau_\delta$, such that $\lim_{\epsilon \to 0^+} \delta_\epsilon > 0$, $\delta_\epsilon = \max_i \delta_i^\epsilon$.

Theorem 4.2.12. Given any countable collection of non-vanishing partitions $\mathcal{T}$, there exists a sequence of functions $u_\epsilon$ satisfying the following properties:

(i) $u_\epsilon \rightharpoonup 0$, weakly-* in $W^{1,\infty}(0,1)$;

(ii) $u_\epsilon' \in \{-1,1\}$, a.e.;

(iii) For each $\tau_\delta \in \mathcal{T}$ and any $(a,b) \subset [0,1]$,

$$\int_a^b u_\epsilon'(x)\overline{u}_\epsilon(x) \, dx \to 0,$$

where $\overline{u}_\epsilon$ is computed with respect to $\tau_\delta$.

(iv) $\overline{u}_\epsilon$ is not a weak white noise.

Proof. We again defer the proof to our paper [37, Theorem 3.4]. □

The following Corollary shows that weak white noise yields the infimum of the subgrid projection term.

Corollary 4.2.13. Let $u_\epsilon = u_\epsilon(x)$ be the continuous piecewise affine approximation of a given function $f$ such that $u_\epsilon' \in \{-1,1\}$ almost everywhere in $(0,1)$ guaranteed by Theorem 4.2.7. Let us assume that the sequence $\{u_\epsilon'\}_{\epsilon > 0}$ becomes weak white noise in the sense of Definition 4.2.9 as $\epsilon \to 0^+$. Then

$$\liminf_{\epsilon \to 0^+} \inf_{\tau_\delta \in \mathcal{T}} \sum_{Q \in \tau_\delta} \int_Q \frac{1}{2} |u_\epsilon'(x) - \overline{u}_\epsilon'(x)|^2 \, dx = 1,$$

(4.43)
where $\mathcal{T}$ is an arbitrary countable collections of partitions of $(0, 1)$. The scaled projections are computed with respect to $\tau_5$.

Proof. The proof follows immediately from the construction used in the previous proof of Theorem 4.2.12 and from Lemma 4.2.11. \hfill \square

Remark 4.2.14. Theorem 4.2.12 and Corollary 4.2.13 indicate that the Subgrid Projection method is necessary but not sufficient to establish a variational principle for the weak white Noise. We address this deficiency in [35].

4.2.2.3 Convergence of the Descent Algorithm

We now show in this section that the following Descent Algorithm applied to minimization of the functional having density (4.29) yields correct microstructures (Young measures).

Definition 4.2.15 (Descent Algorithm). Let us consider the following functional

$$\int_0^1 \Phi(u'(x)) \, dx + \int_0^1 (u(x) - f(x))^2 \, dx,$$

(4.44)

where the density $\Phi$ is positive, at least twice continuously differentiable with respect to its argument, and satisfies

$$\Phi(s) \geq C \left( |s|^p - 1 \right), \quad C > 0, \quad 2 \leq p < +\infty.$$

(4.45)

We define the Descent Algorithm associated with the functional (4.44) as follows.

Let $u_0 \in W^{1,p}(0, 1)$, $u_0(0) = f(0)$, $u_0(1) = f(1)$ be given. Let $\tilde{g}(u_n) \in W^{1,q}_0(0, 1)$, $1/p + 1/q = 1$, be the weak solution of

$$-\tilde{g}''(u_n)(x) = \partial_x D\Phi(u'_n(x)).$$

(4.46)

The steepest (w.r.t. $W^{1,p}$-topology) descent direction $g(u_n) \in W^{1,p}(0, 1)$ is given by

$$g'(u_n)(x) \overset{\text{def}}{=} \frac{\tilde{g}'(u_n)(x)}{|\tilde{g}'(u_n)(x)|^{1-\frac{1}{p}}}, \quad \text{a.e. in } (0, 1).$$

(4.47)
Let the step length be given by

\[ \alpha_n = \operatorname{Argmin}_{\alpha \in \mathbb{R}^+} \int_0^1 \Phi(u_n - \alpha g(u_n))(x) \, dx. \] (4.48)

The descent update is then

\[ u_{n+1} = u_n - \alpha_n g(u_n). \] (4.49)

We make the following remark on the Descent Algorithm 4.2.15.

**Remark 4.2.16.** We note that \( u_0 \in W^{1,6+\epsilon}_0(0,1), \epsilon > 0, \) guarantees for the Descent sequence \( \{u_n\}_{n \in \mathbb{N}} \) existence of a constant \( C, \) independent of \( n, \) such that

\[ \|u_n\|_{W^{1,6+\epsilon}_0(0,1)} \leq C < +\infty \] (4.50)

provided \( \sum_{i=0}^n \alpha_i < +\infty. \)

We need the following Lemma 4.2.17 in the proof of the Section's main result Theorem 4.2.18.

**Lemma 4.2.17.** Let us assume that the sequence \( \{u_n\}_{n \in \mathbb{N}} \subset W^{1,2}(0,1) \) has a weak limit denoted by \( u, u_n(0) = u(0), u_n(1) = u(1). \) Then

\[ \lim_{n \to \infty} \sum_{Q \in \mathcal{T}_n} \int_Q (u_n' - \bar{u}_n')(x) (\varphi' - \bar{\varphi}')(x) \, dx = 2 \int_0^1 u'(x)\varphi'(x) \, dx \] (4.51)

for all \( \varphi \in W^{1,2}_0(0,1). \) Thus, in particular, the left-hand side of (4.51) vanishes in the limit, so long as the weak limit is a harmonic function.

**Proof.** We refer to our paper [37, Lemma 10.1] for the proof of the lemma. \( \square \)

This brings us to the final result, indicating that the Subgrid Projection term indeed provides the correct energy perturbation necessary for the direct minimization, allowing the Descent Algorithm, Definition 4.2.15, to succeed in approximation of the desired Young measure.
Theorem 4.2.18. Let us assume that the sequence \( \{u_n\}_{n \in \mathbb{N}} \subset W_0^{1,6+\epsilon}(0,1), \epsilon > 0, \) is generated by the Descent Algorithm, Definition 4.2.15, applied to minimization of the functional

\[
I_{\tau_k}(u, \Delta) \overset{\text{def}}{=} \int_0^1 \mathcal{W}_{\text{macro}}(u(x)) + \mathcal{W}_{\text{micro}}(u'(x)) \, dx + \sum_{Q \in \tau_k} \int_Q \mathcal{W}_{\text{meso}}(u'(x)) \, dx. \tag{4.52}
\]

We assume that this sequence is uniformly bounded with respect to \( n \) in \( W_0^{1,6+\epsilon}(0,1), \) c.f. Remark 4.2.16. Here,

\[
\mathcal{W}_{\text{macro}}(u(x)) \overset{\text{def}}{=} \frac{1}{2} |u(x) - f(x)|^2,
\]

\[
\mathcal{W}_{\text{micro}}(u'(x)) \overset{\text{def}}{=} \frac{1}{4} |u'(x)^2 - 1|^2,
\]

\[
\mathcal{W}_{\text{meso}}(u'(x)) \overset{\text{def}}{=} \Delta \frac{1}{2} |u'(x) - \bar{u}(x)'|^2, \quad \text{for a.a. } x \in (0,1).
\] \tag{4.53}

Let us further assume for simplicity that the target function \( f \) is identically zero.

Then for any initial iterate \( u_0 \in W_0^{1,6+\epsilon}(0,1), \epsilon > 0, \) such that

\[
\int_0^1 \mathcal{W}_{\text{macro}}(u_0(x)) + \mathcal{W}_{\text{micro}}(u_0'(x)) \, dx \leq \delta_0, \quad 0 < \delta_0 \ll 1, \tag{4.54}
\]

there exists a sequence of partitions \( \tau_{\delta_n}, \text{card}(\tau_{\delta_n}) \to \infty \) as \( n \to \infty, \) such that

\[
\lim_{n \to +\infty} \int_0^1 \mathcal{W}_{\text{macro}}(u_n(x)) + \mathcal{W}_{\text{micro}}(u_n'(x)) \, dx = 0. \tag{4.55}
\]

Proof. The minimizing sequence \( \{u_n\}_{n \in \mathbb{N}} \) generated by the Descent algorithm is uniformly (w.r.t. \( n \)) bounded in the Sobolev space \( W^{1,4}(0,1). \) Let us denote its weak limit by \( u. \) Also, let us denote the weak limit of the sequence \( ((u_n')^2 - 1) u_n' \) by \( w. \)

Applying Lemma 4.2.17, we obtain from the weak form of (4.46) the relation

\[
\int_0^1 w(x)\varphi'(x) \, dx + \int_0^1 u(x)\varphi(x) \, dx + 2\Delta \int_0^1 u'(x)\varphi'(x) \, dx = 0, \tag{4.56}
\]

for all \( \varphi \in W_0^{1,4}(0,1). \) The solution \( u \) of (4.56) with the right-hand side generated by \( w \) is given by

\[
u(x) = \frac{1}{2\sqrt{2\Delta}} \left( \exp \frac{x}{\sqrt{2\Delta}} - \exp - \frac{x}{\sqrt{2\Delta}} \right) \int_x^1 \exp \frac{y}{\sqrt{2\Delta}} w'(y) \, dy. \tag{4.57}\]
Comparing (4.57) and (4.56) we find that \( w' \equiv 0 \). Otherwise, \( \lim_{\Delta \to 0^+} u = +\infty \) in \((0,1)\), which would contradict the uniform boundedness of \( u \) with respect to \( \Delta \), since the weak limit is contained in the closer of the convex hull of the minimizing sequence. Consequently, the weak limit \( u \) must be zero. Thus in regards to equation (4.56), this yields \( w \equiv 0 \). Hence

\[
w_n = \left( (u_n')^2 - 1 \right) u_n' \rightharpoonup 0 \quad \text{weakly in } L^{4/3}(0,1). \tag{4.58}\]

Now we show that the weak convergence in (4.58) is in fact a strong one. In order to obtain this result we define the sequence of functionals \( G_n \in (W^{1,2}(0,1))^* \) by

\[
\langle G_n, \varphi \rangle \overset{\text{def}}{=} \int_0^1 w_n(x)\varphi'(x) \, dx + K \int_0^1 \varphi(x) \, dx, \quad K > 0. \tag{4.59}\]

From Remark 4.2.16, we note that \( u_n \in W^{1,6+\epsilon}(0,1), \epsilon > 0 \), which implies that \( w_n \in L^p(0,1), p > 2 \). Since the weak limit \( u \) is equal to zero, we have for some \( K > 0 \),

\[
\int_0^1 u_n(x)\varphi(x) \, dx + \sum_{Q \in \mathcal{T}_n} \int_Q (u_n' - \overline{u}_n')(x) (\varphi' - \overline{\varphi}')(x) \, dx \leq K \int_0^1 \varphi(x) \, dx, \tag{4.60}\]

for all \( \varphi \in \mathcal{D}^+(0,1) \), where \( \mathcal{D}^+(0,1) \) denotes positive and smooth functions with compact support. Thus we obtain from (4.56) the inequality

\[
\langle G_n, \varphi \rangle \geq 0, \quad \text{for all } \varphi \in \mathcal{D}^+(0,1). \tag{4.61}\]

Denoting

\[
\langle G_n, \varphi \rangle = K \int_0^1 \varphi(x) \, dx,
\]

it follows from the continuous embedding of \( W^{1,p}(0,1) \) into \( W^{1,2}(0,1) \) for any \( p > 2 \) that \( G_n \rightharpoonup G \) weakly in \((W^{1,2}(0,1))^*\). Applying the compactness criteria of Murat [111], c.f. also [113, Theorem 4.1], we conclude that the functionals \( G_n \) converge strongly to \( G \) in \((W^{1,p}(0,1))^*, p > 2 \). Now we define

\[
v_n(x) \overset{\text{def}}{=} \int_0^x w_n(y) \, dy.
\]
Since $v_n \in W^{1,p}(0,1)$, $p > 2$, and since $\{v_n\}_{n \in \mathbb{N}}$ is uniformly bounded, we have
\[
\|w_n\|_{L^2(0,1)}^2 = \langle G_n - G, v_n \rangle \longrightarrow 0. \quad (4.62)
\]
Thus the sequence $\{u_n\}_n$ cannot converge strongly to its weak limit, since otherwise the algorithm would converge to a local maximum, which is not possible. Hence it follows from (4.62) that $(v'_n)^2 - 1)^2$ converges to zero in $L^1(0,1)$. The result of the proof follows.

We finish with the following result justifying the convergence of approximate probabilities.

**Corollary 4.2.19.** Let us assume that the assumptions of Theorem 4.2.18 are satisfied. Let the sequence $\{u_n\}_{n \in \mathbb{N}} \subset W^{1,4}(0,1)$ be given by the Descent Algorithm, Definition 4.2.15, applied to the minimization of the functional $I_{\mathbb{R}}(u, \Delta)$, (4.52). Then the Young measure, representing the structure of the weak limit of the minimizing sequence $\{u_n\}_{n \in \mathbb{N}} \subset W^{1,4}(0,1)$, is given by $\mu_{x,u_n} = \lambda(x)\delta_{-1} + (1 - \lambda(x))\delta_{+1}$, where
\[
\lambda(x) = \lim_{r \to 0^+} \lim_{n \to +\infty} \frac{\text{meas} \left( \{ z \in B_R(x) \mid |u'_n(x) - 1| \leq r \} \right)}{\text{meas} (B_R(x))}. \quad (4.63)
\]

**Proof.** We refer to [37, Corollary 6.6] for the proof.

From these theoretical results, we see that indeed in the simplified case of harmonic macroscopic austenitic displacements of a shape memory alloy plate, the Subgrid Projection Method provides a method for the direct computation of the probabilistic behavior of the martensitic structure at the austenite-martensite interface. Therefore, at least in these special cases, isothermal dynamics calculations of shape memory alloy wires are possible through incorporation of such perturbations to the strain energy function $\mathcal{W}(\gamma)$. Furthermore, some of our more recent work [35, 38] has focused on alternate means of providing such perturbations to the free-energy, based on the weak white noise requirement discussed in Section 4.2.2.2. This work has
attempted to construct perturbations that allow for direct computations of measure-valued solutions for more general cases than that discussed in this section.

However, the development of perturbations for temperature dependent free energies used in thermodynamics calculations is far from complete. Thus, our proposed model describing the thermodynamics of shape memory alloy wires (Chapter 3) cannot benefit from these free-energy perturbations at the current time. Further work on the development of such stochastic perturbations for the temperature-dependent case, though, may someday provide the key to both microstructure and thermodynamics computations for materials having general, nonconvex free energies. Since these stochastic perturbations await future development, in the next section we present the current theoretical results most closely related to our proposed thermodynamic model for shape memory alloys.

4.2.3 Existence and Uniqueness for Viscous Models

Results up to now on the existence of solutions to systems resembling our thermodynamic model (4.9) have traditionally relied on significant viscosity and/or capillarity terms to stabilize the dynamical system. As discussed in Section 2.5 and to be further discussed in Chapter 5, such viscosity effects are unphysical and may further violate physical principles such as material frame indifference [9, 11, 10, 128, 129]. However, as in the case of traditional, scalar hyperbolic conservation laws, such artificial viscosity is required to guarantee the existence of weak solutions.

One of the key works in this area is by Friesecke and Dolzmann [59], where they consider the semilinear viscoelastic system over the open bounded domain $\Omega \subset \mathbb{R}^n$
given by

\[ u_{tt} = \text{Div}(\sigma(Du) + Du_t) \quad \text{in } \Omega \times (0, \infty), \]

\[ u = g \quad \text{on } \partial \Omega \times [0, \infty), \]

\[ u = u_0 \quad \text{in } \Omega \times \{0\}, \tag{4.64} \]

\[ u_t = v_0 \quad \text{in } \Omega \times \{0\}, \]

where \( u : \Omega \times [0, \infty) \rightarrow \mathbb{R}^N, g \in W^{1,p}(\Omega, \mathbb{R}^N), u_0 \in W^{1,p}(\Omega, \mathbb{R}^N) \) and \( v_0 \in L^2(\Omega, \mathbb{R}^N) \) are given functions, \( \sigma(F) = \partial \Psi(F) / \partial F \). The potential \( \Psi : \mathbb{R}^{N \times n} \rightarrow \mathbb{R} \) is assumed to satisfy the following hypotheses:

(H1) \( \Psi \in C^2(\mathbb{R}^{N \times n}) \);

(H2) there exist constants \( c > 0, C > 0, \) and \( p \geq 2 \) such that

\[ c |F|^p - C \leq \Psi(F) \leq C (|F|^p + 1), \quad |\sigma(F)| \leq C (|F|^p-1 + 1); \]

(H3) there exists \( K > 0 \) such that the Andrews-Ball condition (see [8]) is satisfied:

\[ (\sigma(\bar{F}) - \sigma(F)) \cdot (\bar{F} - F) \geq 0 \quad \forall \bar{|F|}, |F| \geq R \quad \text{and some } R > 0. \]

Through discretization of this system in space and time, Friesecke and Dolzmann examine the existence of weak solutions \( u \) to the problem (4.64). The key feature of their analysis resides in examination of the discretized damping term \( Du_t \). Essentially, denoting the time discretization as \( h \), the weak form of the original evolution system may be satisfied at critical points of the Euler-Lagrange equation

\[ J^h[u] = \int_{\Omega} \left( \tilde{\Psi}(Du) + \mathcal{C}(u, Du) \right) \, dx, \tag{4.65} \]

where \( \mathcal{C}(u, Du) \) is uniformly convex in both \( u \) and \( Du \), and

\[ \tilde{\Psi}(F) \overset{\text{def}}{=} \Psi(F) + \frac{1}{2h} |F|^2. \]

Clearly, for some \( h \leq h_0 \), the potential \( \tilde{\Psi} \) (and therefore the overall functional \( J \)) overcomes the non-convexity of \( \Psi \) and is convex. As a result, Friesecke and Dolzmann are able to prove the following Theorem 4.2.20...
Theorem 4.2.20 ([59], Theorem 4.1). Let $\Omega \subset \mathbb{R}^N$ be open and bounded, assume $\Psi$ satisfies (H1), (H2), and (H3); let $g \in W^{1,p}(\Omega, \mathbb{R}^N)$, and let the initial data $u_0 \in \mathcal{A} = \{u \in W^{1,p}(\Omega, \mathbb{R}^N) : u - g \in W^{1,p}_0(\Omega, \mathbb{R}^N)\}$ and $v_0 \in L^2(\Omega, \mathbb{R}^N)$. Then there exists

$$u \in L^\infty(0, \infty; \mathcal{A})$$
$$\cap W^{1,\infty}(0, \infty; L^2(\Omega, \mathbb{R}^N))$$
$$\cap W^{1,2}_{loc}([0, \infty); W^{1,2}(\Omega, \mathbb{R}^N))$$
$$\cap W^{2,2}_{loc}([0, \infty); W^{-1,q}(\Omega, \mathbb{R}^N)), \quad 1/p + 1/q = 1,$$

which is a weak solution of (4.64), i.e. $u|_{t=0} = u_0, u_t|_{t=0} = v_0$, and

$$\int_0^\infty \int_\Omega \left[ (\sigma(Du) + Du_t) \cdot D\xi - u_t \cdot \xi_t \right] \, dx \, dt = 0 \quad \forall \xi \in C^\infty_0(\Omega \times (0, \infty), \mathbb{R}^N).$$

These results indicate that, at least in a case of the single non-convex partial differential equation governing viscoelastic motion in shape memory alloys, addtion of artificial viscosity to the system guarantees existence of solutions.

Similar developments have been shown in the case of thermoviscoelastic systems resembling the proposed model (4.9). Niezgódka and Sprekels investigate the behavior of the following thermomechanical model of shape memory alloy wires in one-space dimension.

**Problem 4.2.21.** Find a pair of functions $(u, \theta)$ satisfying the following:

(i) Differential equations in $\Omega \times (0, T)$:

$$\rho_0 u_{tt} - (\sigma_d)_x - (\sigma_\eta)_x - \rho_0 f = 0,$$

$$\rho \theta_t s_t - \kappa \theta_{xx} - \alpha \kappa \theta_{xxt} - \rho r - \sigma_d \gamma_t = 0; \quad (4.66)$$

(ii) Initial conditions in $\Omega$:

$$\theta(\cdot, 0) = \theta_0,$$

$$u(\cdot, 0) = u_0, \quad u_t(\cdot, 0) = v_0; \quad (4.67)$$
(iii) Boundary conditions on \( \partial \Omega \times (0, T) \):

\[
\kappa \frac{\partial \theta}{\partial x} = \theta_T, \\
u = 0;
\]  
(4.68)

(iv) Constitutive relations:

\[
s = -\frac{\partial \Psi}{\partial \theta}, \\
\sigma_q = \rho_0 \frac{\partial \gamma}{\partial \gamma}, \\
\sigma_d = \mu \rho_0 \gamma_t, \\
\gamma = u_x,
\]  
(4.69)

with Helmholtz free energy given by

\[
\Psi(\gamma, \theta) = \Psi_0(\theta) + \Psi_1(\theta) \gamma^2 + \Psi_2(\gamma).
\]  
(4.70)

Niezgódka and Sprekels constrain this problem using the following set of assumptions on the structure of the free energy, the external heat supply, the external body force, and the initial and boundary conditions.

(A1) \( \Psi_0 \in C^3(\mathbb{R}^+) \) has the properties:

(i) there exists a minimal temperature \( \theta_s > 0 \) such that \( \theta_t(t) > \theta_s \) for all \( t \geq 0 \) and \( \Psi_0(\theta) \equiv 0 \) for \( \theta \leq \theta_s \);

(ii) \( \Psi_0''(\theta) \leq 0 \) for all \( \theta \in \mathbb{R}^+ \), \( \Psi_0(\theta) - \theta \Psi_0'(\theta) \geq -C > -\infty \) for all \( \theta \in \mathbb{R}^+ \) and some constant \( C > 0 \);

(iii) \( \left| \int_0^\infty \xi^2 \Psi_0''(\xi) \, d\xi \right| \leq C < +\infty \).

(A2) \( \Psi_1 \in C^3(\mathbb{R}^+) \) satisfies the conditions:

(i) \( \Psi_1(\theta) \equiv \Psi_{1s} < 0 \) for \( \theta \leq \theta_s \);
(ii) \(|\theta^2 \Psi_1'(\theta)| \leq C < +\infty\) for all \(\theta \in \mathbb{R}^+\);

(iii) \(\int_0^\infty \xi^2 \Psi''_1(\xi) \, d\xi \leq C < +\infty\);

(iv) there exists some temperature \(\theta_1 > \theta_s\) such that

\[
\Psi_1(\theta_1) = 0, \quad \Psi_1''(\theta_1) = 0, \quad \Psi_1''(\theta) \begin{cases} > 0, & \text{for } \theta \in (\theta_s, \theta_1), \\ < 0, & \text{for } \theta > \theta_1; \end{cases}
\]

(v) there exists some displacement gradient \(\gamma_0 > 0\) such that

\[-\theta \Psi''_0(\theta) - \theta \Psi'_1(\theta) \gamma^2 \geq 0\] for every \((\theta, \gamma) \in \mathbb{R}^+ \times [0, \gamma_0]\).

(A3) \(\Psi_2 \in C^3(\mathbb{R})\) satisfies the properties:

(i) it is of the form

\[
\Psi_2(\gamma) = \begin{cases} -a_4 \gamma^4 + a_6 \gamma^6, & \text{for } |\gamma| \leq \gamma_M, \\ \varphi_0(\gamma), & \text{for } |\gamma| > \gamma_M, \end{cases}
\]

with constants \(a_4, a_6, \gamma_M > 0\) and a convex function \(\varphi_0(\gamma)\) which satisfies the following asymptotic condition for \(a_2, a'_2 > 0\)

\[0 < a_2 \leq \varphi''_0(\gamma) \leq a'_2 < +\infty\] for \(|\gamma| > 2\gamma_M\);

(ii) for every \((\theta, \gamma) \in \mathbb{R}^+ \times \mathbb{R}\) there holds

\[
(\Psi_1(\theta) - \theta \Psi_1'(\theta)) \gamma^2 + \Psi_2(\gamma) \geq C_1 \gamma^2 - C_2,
\]

with positive constants \(C_1, C_2\).

(A4) \(r \in L^2_{\text{loc}}(0, \infty; L^2(\Omega))\) is non-negative.

(A5) \(f \in L^2_{\text{loc}}(0, \infty; L^2(\Omega))\) and \(\theta_T \in H^1_{\text{loc}}(0, \infty)\) satisfies \((\theta_T)_t \geq 0\) on \((0, \infty)\).

(A6) the initial data satisfies:
(i) \( u_0 \in C^2(\Omega) \) satisfies \( u_0|_{\Gamma} = u_0''|_{\Gamma} = 0 \) with \( \| u_0\|_{H^2(\Omega)} \overset{\text{def}}{=} u_{2,0} \);

(ii) \( v_0 \in H^1_0(\Omega) \);

(iii) \( \theta_0 \in H^2(\Omega) \) satisfies \( \theta_0(x) > \theta_\ast \) for all \( x \in \Omega \);

(iv) \( -\Psi''_0(\theta_0(x)) - \Psi'_1(\theta_0(x))(u'_0(x))^2 \geq D_1 > 0 \) for every \( x \in \Omega \).

(A7) \( \rho_0 \) is constant, and is normalized so that \( \rho_0 \equiv 1 \).

Under this significant list of assumptions, Niezgódka and Sprekels are able to show that for some time \( T > 0 \), there exists at least one solution \((u, \theta)\) on \( \Omega \times [0, T) \) that satisfies the following appropriate weak formulation of the Problem 4.2.21:

**Problem 4.2.22.** Find a pair \((u, \theta)\) and a time interval \([0, T]\), \( T > 0 \), such that

(i) \( u \in H^2(0, T; L^2(\Omega)) \cap H^1(0, T; H^2(\Omega)) \cap H^1_0(\Omega) \),

(ii) \( \theta \in L^2(0, T; H^1(\Omega)) \cap H^1(0, T; L^2(\Omega)) \),

(iii) the boundary function \((\theta_t)_t\) is well-defined in \( L^2(0, T) \),

(iv) the initial conditions (4.67) hold,

(v) the constitutive relations (4.69) hold almost everywhere on \( \Omega \times (0, T) \),

(vi) \( \theta(x, t) \geq \theta_\ast > 0 \) almost everywhere in \( \Omega \times (0, T) \),

(vii) for any time \( t \in (0, T] \) the following variational inequalities are satisfied:

\[
\int_0^T \int_\Omega [u_t \phi - (\sigma_\theta)_{xx} - \mu(\sigma_\theta)_x - f] \phi \, dx \, dt \quad \forall \phi \in L^2(0, T; L^2(\Omega))
\]

\[
\int_0^T \int_\Omega \left[ \theta \partial_t \psi + \mu \psi_{xx} \psi + \mu \psi_{\partial_x} \psi_{xx} + \kappa \theta_x \psi_x + \alpha \kappa \theta_{xx} \psi_x - r \psi \right] \, dx \, dt
\]

\[
+ \int_0^t \left[ \partial_t + \alpha(\theta_t)_t \right] \psi(\partial \Omega, t) \, dt = 0 \quad \forall \psi \in L^2(0, T; H^1(\Omega)).
\]
Therefore the nonlinear viscoelastic model in Problem 4.2.21 has been proven to obtain at least one weak solution, satisfying the requirements laid out in Problem 4.2.22. This result is essentially based on the smoothness of the free energy $\Psi$ and the existence of the positive viscosity coefficients $\mu, \alpha$, that regularize the behavior of the differential system [114]. Moreover, Hoffmann and Songmu [73] and Hoffmann and Showalter [72] have expanded on the above theory laid down by Niezgódka and Sprekels, relaxing some of the smoothness conditions (A1)-(A3) and allowing the existence to rely on either a significant mechanical ($\mu$) or a thermal ($\alpha$) viscosity.

Thus the proposed model (4.9) has been constructed to allow for such mechanical viscosity and is based on an alternate form of the Helmholtz free energy $\Psi$ that satisfies the desired regularity and growth properties (see Section 3.4.3). Furthermore, the solution method presented in Chapter 5 employs an approach similar to the Method of Vanishing Viscosity (Section 4.1.3) in order to lower the artificial dissipation introduced through the viscous terms required from the theoretical results discussed in this section [59, 72, 73, 114]. Moreover, the success of the Method of Vanishing Viscosity as a selection principle for standard scalar conservation laws suggests that the solution methods in Chapter 5 may work for the nonlinear differential system (4.9); although there is currently no theory to support this conjecture.

### 4.3 Theoretical Damping and Control Properties

We now turn to the proposed application of our thermodynamic model for shape memory alloys—thermally-activated vibration damping. There are theoretical results in two key areas dealing with this topic in the case of shape memory alloys. The first of these deals with the theoretical damping ability of these alloys. Specifically, work in this area has concentrated on the jump dynamics associated with the austenite-martensite interface in a transforming piece of shape memory material. Some of these
results are summarized in the following Section 4.3.1. The second of these research areas on thermally-controlled vibration damping considers the construction of reduced models and generalized controls for such processes. We present in Section 4.3.2 some of our work in this area, taken from our paper "On thermodynamic active control of shape memory alloy wires" [88].

4.3.1 Theoretical Material Damping

One of the key reasons for the current abundance of research on the thermodynamics of shape memory alloys is the large theoretical damping potential of these materials. Due to their high sensitivity to temperature fluctuations (see Chapter 2), as well as their unique solid-state first-order phase transformation, shape memory alloys exhibit a large potential for active vibration damping.

R. Abeyaratne and J. Knowles have done a considerable amount of research into these theoretical damping properties through their study of the interface dynamics between austenitic and martensitic regions in a transforming shape memory alloy [3, 4, 5]. Within these alloys, Abeyaratne and Knowles find that the movement of the austenite-martensite phase boundary acts as an active mechanism for conversion of mechanical to thermal energy (vibrations → heat). They consider the processes of shock- and stress-induced phase transformations in a purely mechanical model of the phase transformation. Within this framework, they analyze these processes using a combination of nonlinear elasticity and jump dynamics at the phase boundary.

To understand their analysis, we first consider the austenite-martensite phase boundary as a dynamic discontinuity, where the position of the interface is given by the
function \( s = s(t) \). Moreover, we define for any function \( g = g(x, t) \) the jump function

\[
\llbracket g(t) \rrbracket = g^+(t) - g^-(t), \quad \text{where}
\]

\[
g^+(t) = \lim_{x \to s(t)+} g(x, t),
\]

\[
g^-(t) = \lim_{x \to s(t)-} g(x, t);
\]

the value \( \llbracket g(t) \rrbracket \) provides a measure of the discontinuity in \( g \) at the point \((s(t), t)\). Due to the regularity of the quantities of interest in the dynamic problem, the deformation \( u \) is continuous in \( x \), while the deformation gradient \( \gamma \) is discontinuous across the phase discontinuity,

\[
\llbracket u(t) \rrbracket = 0, \quad \llbracket \gamma(t) \rrbracket \neq 0.
\]

Combining the Clausius-Duhem inequality (3.5) with the assumption that the entropy of a material element cannot decrease as it moves across strain discontinuity results in the relation [5]

\[
f(t) \dot{s}(t) \geq 0.
\]

Here, the function \( f = f(t) \) denotes the driving force acting on the discontinuity at \( s(t) \), given by

\[
f(t) = \dot{f}(\gamma^-, \gamma^+) \overset{\text{def}}{=} \int_{\gamma^-}^{\gamma^+} \sigma(z) \, dz - \frac{1}{2} \left( \sigma(\gamma^+) + \sigma(\gamma^-) \right) (\gamma^+ - \gamma^-)
\]

We then define the dissipation \( D(t) \) of mechanical energy on a small interval \((x_1, x_2) \subset (0, L)\) as the difference between the change in external work and the rate of change of total energy,

\[
D(t) \overset{\text{def}}{=} (\sigma u_t) \bigg|_{x_1}^{x_2} - \frac{d}{dt} E(t),
\]

where \( E(t) \) denotes the total energy (kinetic + strain) in the interval \((x_1, x_2)\) at time \( t \). The field and jump equations then dictate that at points where \( \gamma, v \) are smooth,

\[
\sigma'(\gamma) \gamma_x - \rho v_t = 0, \\
v_x - \gamma_t = 0;
\]
and at points where either \( \gamma \) or \( v \) jump across the curve \( x = s(t) \), (4.76) becomes [5]

\[
[\sigma] + \rho \dot{s} [v] = 0, \\
[\gamma] \dot{s} + [v] = 0.
\] (4.77)

The equations (4.76) and (4.77), along with the definition (4.75) and the assumption (4.73), imply the following relation on the dissipation,

\[
D(t) = f(t) \dot{s}(t) \geq 0.
\] (4.78)

Thus the dissipation may be thought of as the amount of work done on the bar by the moving discontinuity. As seen from (4.75), as \( D \) increases the total energy of the system decreases. Furthermore, since the driving force \( f \) is nonzero at the discontinuity \( x = s(t) \) and the interface \( s(t) \) is assumed to move, \( f(t) \dot{s}(t) \neq 0 \), and the overall energy of the system decreases proportionally to the speed of the phase boundary.

### 4.3.2 Idealized Vibrational Control

Given the damping potential of shape memory alloys shown, we move next to the topic of controllability for such damping processes. Clearly, with the theoretical damping possibilities of these alloys, the question of finding suitable controls for these systems has received intense study ([89, 115, 117, 131] and others). The great majority of these approaches, however, are founded on highly linearized models of shape memory alloy behavior, as discussed in Section 2.4. This fact is, at least in part, due to the significant difficulties involved in the solution of nonlinear models for shape memory alloy behavior (Section 4.2). However, some of our recent work has dealt with the controllability of nonlinear models for these alloys [88]. In the remainder of this chapter, we present some of the results from this work. However, in order to avoid many of the solution difficulties inherent in the fully nonlinear thermodynamic model,
we also consider a simplified control model for analysis and construction of generalized controls.

4.3.2.1 A Linearized model

The first major simplification of the shape memory alloy model (4.9) in this section is that we consider only two possible temperatures, $\theta_0$ and $\theta_1$, respectively corresponding to the austenite and martensite phases, and then take the pointwise choice between these temperatures as a control, ignoring the dynamics of heat transfer through the wire as in (4.9) and dissipation to the environment, etc. Thus, the control $\chi = \chi(x, t) \in \{0, 1\}$ identifies a choice of temperature $\theta(x, t) \in \{\theta_0, \theta_1\}$. This simplification provides us with our first step in the control process; however, in doing so, the thermodynamics of the system and consideration of the relevant time scale for the temperature variation are lost. An implicit assumption of this control choice is that the time scale and control mechanism permits rapid temperature fluctuations, allowing us to (fairly arbitrarily) “choose” $\chi = \chi(t, x)$ pointwise. Furthermore, the discussion of interface motion leading to (4.78) treats a setting in which such irregular temperature variations as considered here would not be allowed. All of this is perhaps plausible in spirit, but it is certainly false in detail, leaving us with a possibly overly optimistic picture of the possible control. Nevertheless, we do anticipate that the new model should still provide a rough approximation to certain possible controlled system behaviors.

In this section, we further consider the displacement as opposed to the deformation for the model. Thus since the deformation gradient and displacement gradient only differ by 1, the austenitic reference configuration is given by the displacement gradient $\gamma = 0$, and the martensitic variants are given by displacement gradients $\pm \gamma_0$.

Under these simplifications, the strain energy density becomes, pointwise, $\mathcal{W} =$
\[ W(\cdot, \chi) = W_{\chi}(\cdot, \chi), \text{ so} \]
\[ W(\gamma, \chi) = \chi W_1(\gamma) + (1 - \chi) W_0(\gamma), \quad W = W_0 + \chi [W_1 - W_0], \quad (4.79) \]

where
\[ W_0(\gamma) = W(\gamma, \theta_0) \approx \frac{1}{2} \gamma^2, \quad \text{for the austenite phase}, \quad (4.80a) \]
\[ W_1(\gamma) = W(\gamma, \theta_1) \quad \text{for the martensite phase.} \quad (4.80b) \]

The density \( W_1 \) remains a double-well potential, e.g., \((\gamma^2 - \tilde{\gamma}^2)^2\). Here \( \chi \in \{0, 1\} \) specifies pointwise the choice of temperature \( \theta \in \{\theta_0, \theta_1\} \), and thus the choice of the energy density function. Consequently, we can write the viscoelastic system (4.9) in the form
\[ u_{tt}(x, t) = -\partial_x (\sigma(x, t)) \quad \text{with} \quad \sigma(x, t) \overset{\text{def}}{=} -\partial_\gamma W(\gamma(x, t), \chi(x, t)) = u_{xx}(x, t) + \partial_x (\chi(x, t) \partial_\gamma (W_1 - W_0)(\gamma(x, t))), \quad \text{for} \ x \in (0, 1). \quad (4.81) \]

In order to model the austenite and martensite phases, we take \( W_0(\gamma) \overset{\text{def}}{=} \frac{1}{2} \gamma^2 \) and \( W_1(\gamma) \overset{\text{def}}{=} (\gamma^2 - \tilde{\gamma}_0^2)^2 \).

At present, we cannot treat existence of solutions of (4.81) for the step function \( \chi \). We do note that if the problem were regularized and \( \chi \) would be a twice differentiable function, then there would necessarily exist a global weak solution of (4.81). This result is stated in the following Lemma and is an immediate consequence of the work by Friesecke and Dolzmann, shown in Theorem 4.2.20.

**Lemma 4.3.1.** Let us assume that \( \chi \in C^2((0, 1) \times (0, T)) \), let \( u_0, v_0 \in W^{1,4}(0, 1) \).

Then there exists a unique global weak solution of (4.81) having the properties
\[ u \in L^\infty((0, \infty); \{v \in W^{1,p}(\Omega) : v - g \in W_0^{1,p}(\Omega)\}) \]
\[ \cap W^{1,\infty}((0, \infty); L^2(\Omega)) \]
\[ \cap W^{1,2}_{\text{loc}}([0, \infty); W^{1,2}(\Omega)) \]
\[ \cap W^{2,2}_{\text{loc}}([0, \infty); W^{-1,q}(\Omega)), \quad (4.82) \]

with \( 1/p + 1/q = 1 \).
Proof. In order to use Theorem 4.2.20 we must verify the conditions

\[ W \in C^2(R^{N \times N}), \]
\[ C_1 |s|^p - 1 \leq W(s) \leq C_2 (|s|^p + 1) , \quad C_1, C_2 > 0 , \quad p \geq 2 , \quad (4.83) \]
\[ (D W(s_1) - D W(s_2)) \cdot (s_1 - s_2) \geq -K |s_1 - s_2| , \quad K > 0 , \]

for the one-dimensional strain energy density \( W \) from equation (4.79), along with the proper initial and boundary conditions.

The initial and boundary conditions from Theorem 4.2.20 hold by assumption, using the choice \( p = 4 \). Further, in our consideration of (4.81) we may safely assume the following bounds for use within the proof of the lemma

\[ \gamma(x, t) \in (-1, 1) , \quad \gamma_0 \in (0, 1) , \quad \chi(x, t) \in [0, 1] . \quad (4.84) \]

Now, \( W \in C^2 \) since \( \chi \in C^2((0, 1) \times (0, T)) \), by assumption, and

\[ \frac{\partial^2}{\partial \gamma^2} W(\gamma, \chi) = \chi(x, t) \left( 12 \gamma^2 - 4 \gamma_0 - 1 \right) + 1 \in C^0 . \quad (4.85) \]

Using the bounds (4.84), we have (with \( p = 4 \) as mentioned above) that

\[ C_1 |\gamma|^p - 1 \leq W(\gamma, \chi) \leq C_2 (|\gamma|^p + 1) \]

for all allowable \( \gamma, \gamma_0 \) and \( \chi \) with \( C_1 = 1 = C_2 \).

Finally, the third requirement from (4.83) for the problem (4.81) becomes, for some \( K > 0 \),

\[ (D W(\gamma_1) - D W(\gamma_2)) \cdot (\gamma_1 - \gamma_2) \geq -K |\gamma_1 - \gamma_2|^2 \]
\[ \iff \]
\[ \left( 4 \chi(x, t) (\gamma_1^2 - \gamma_2^2) - 4 \chi(x, t) \gamma_0^2 + 1 - \chi(x, t) \right) (\gamma_2 - \gamma_1) \geq -K (\gamma_1 - \gamma_2)^2 \]
\[ \iff \]
\[ 4 \chi(x, t) (\gamma_1^2 - \gamma_2^2) - 4 \chi(x, t) \gamma_0^2 + 1 - \chi(x, t) \geq -K . \]
Note: if $(\gamma_1 - \gamma_2)^2 = 0$, then the second inequality is trivially satisfied, and we may consider only the final inequality from (4.87). Again considering the bounds on acceptable values for $\gamma$, $\gamma_1$ and $\chi$ from (4.84), we see that the above inequality is satisfied with $K = 8$.

Thus the energy density $\mathcal{W}$ for problem (4.81) satisfies the requirements of the Theorem 4.2.20, and the Lemma is complete. \hfill \Box

The second major simplification of our control model introduces a seemingly linear simplification of the stress-strain relation. We introduce a function $\phi$ by

$$
\phi(x, t) = \chi(x, t) \Gamma(\gamma(x, t)) \quad \text{with} \quad (4.88a)
$$

$$
\Gamma(\gamma) = \frac{\partial}{\partial \gamma} \mathcal{W}_1(\gamma) - \frac{\partial}{\partial \gamma} \mathcal{W}_0(\gamma). \quad (4.88b)
$$

Using this notation, the equation (4.81) becomes an apparently linear controlled wave equation

$$
\begin{align*}
&u_{tt}(x, t) = u_{xx}(x, t) + \partial_x \phi(x, t), \quad (x, t) \in (0, 1) \times (0, T), \quad T > 0, \\
&u(x, 0) = u_0(x), \quad x \in (0, 1), \\
&w(x, 0) = v_0(x), \quad x \in (0, 1).
\end{align*} \quad (4.89)
$$

For definiteness we consider the homogeneous Dirichlet conditions

$$
u(0, t) = u(1, t) = 0, \quad 0 \leq t \leq T. \quad (4.90)$$

The function $\phi = \phi(x, t)$ may now be interpreted as a linear control in its appearance in (4.89). Of course, this linearity is only apparent. In view of (4.79), the choice of the function $\phi$ is here subject to an extremely strong constraint. We must have at any point $(x, t)$

$$
\frac{\phi(x, t)}{\Gamma(\gamma(x, t))} \in \{0, 1\}. \quad (4.91)
$$

Our strategy for controllability of (4.89), summarized in Theorem 4.3.3, consists in taking "arbitrary" $\phi$ and in the interpretation of (4.89) as a linear problem. Given
an integrable function \( \phi = \phi(x, t) \), the existence of a solution to (4.89) is then not a problem. It can be represented by the formula

\[ u(x, t) = \frac{1}{2} \left( u_0(x + t) + u_0(x - t) + \int_{x-t}^{x+t} u_1(\xi) \, d\xi \right) + \int_0^t \phi(\xi, \tau) \left. \right|_{\xi=x-t}^{\xi=x+t} \, d\tau \quad (4.92) \]

where, for the boundary conditions (4.90), we formally extend \( u_0(\xi), u_1(\xi), \) and \( \phi(\xi, \tau) \) as odd functions of \( \xi \) across 0 and 1. The issues of concern are thus regularity and the constraint (4.91). Certainly (4.92) gives an \( L^2(0,1) \) solution \( u \) if the data is \( L^2(0,1) \). However, we can expect to have \( \phi \in L^2 \) from (4.88a) only if \( u_x \in L^2 \) (essentially, for solutions with finite strain energy \( W \)), and from (4.92) we can expect only \( H^{-1}(0,1) \) regularity for \( u_x \) corresponding to \( \phi \in L^2 \). On the other hand, we could not get \( \phi \) smoother than that (e.g., \( \phi \in H^1(0,1) \)) would be needed to get \( \phi_x \in L^2(0,1) \) yielding \( u_x \in L^2(0,1) \), even for very smooth \( u_x \), since we have taken \( \chi \) as pointwise discontinuous. This is a consequence of our simplifying consideration of only two temperatures.

### 4.3.2.2 Generalized Solutions

Our next step is to consider a suitably generalized notion of solutions which correspond in some sense to the "two-temperature" setting. What we want is to permit replacing \( \chi \) as our control function by a smoother function \( \bar{\chi} \) allowed to take more general values, and so relaxing the constraint (4.91) on \( \phi \) in such a way as to obtain consistency with (4.92).

Note that when we switch values for the temperature at some material point, the energy density jumps by \( \pm \Gamma(u_x) \) at that point. Thus, our assumption that we can change the temperature arbitrarily can be expected to defeat any suggestion of an intrinsic mechanism for bounding the energy. However, we do note that rapidly oscillating temperature fluctuations (rapid alternation of these temperature switchings) would tend to cancel in their resultant effect on energy; compare the classical
Riemann-Lebesgue Lemma of real analysis: $\int_a^b f(x) \sin n x \, dx \to 0$ as $n \to \infty$ for any integrable function $f$. Thus, the choice of a rapidly oscillating control $\chi$ may plausibly produce a finite energy solution, even a smooth solution, in an appropriate limit. Let us describe the nature of such a limit, which we will then consider as a generalized solution of the model (4.81).

As usual, we may reasonably consider a modeling domain at a material point $x$ having a characteristic length scale $h$, and we then let $\bar{\chi}_h = \bar{\chi}_h(x, t)$ be the fraction of this domain, with respect to length in the reference coordinates, for which the temperature is $\theta_1$. Thus $1 - \bar{\chi}_h$ is the complementary fraction for which the temperature is $\theta_0$, i.e., $\bar{\chi}_h$ is the local average of $\chi$. Namely,

$$\bar{\chi}_h(x, t) = \frac{1}{h} \int_x^{x+h} \chi(\xi, t) \, d\xi. \quad (4.93)$$

In this sense we are considering a small "modeling domain" $[x, x+h]$ with $h$ small enough that we can visualize the situation as consisting of a large number of "slices", alternating austenite and martensite phases, with $u$ piecewise smooth. Indeed, we assume at this scale that $u$ is piecewise linear and that the austenite slices all have $u_x \approx \gamma_A$ while the martensite slices, within this domain, all have $u_x \approx \gamma^M$. We then obtain

$$(u(x+h, t) - u(x, t)) = \int_x^{x+h} u_x(\xi, t) \, d\xi = h \left( \bar{\chi}(x, t) \gamma^M + (1 - \bar{\chi}(x, t)) \gamma_A \right), \quad (4.94)$$

where $\bar{\chi}(x, t)$ and $(1 - \bar{\chi}(x, t))$, respectively, give the local fractions of martensite and austenite. Thus $\gamma(x, t)$ is equal to the locally averaged value of $u_x$ at $(x, t)$, i.e.,

$$\gamma(x, t) \overset{\text{def}}{=} \langle u_x(x, t) \rangle = \bar{\chi}(x, t) \gamma^M + (1 - \bar{\chi}(x, t)) \gamma_A. \quad (4.95)$$

We also assume continuity of the forces across the interfaces. We assume that there is no significant gross motion of the interfaces relative to each other. Hence, the force, which is effectively constant within the domain, is given by

$$\sigma \overset{\text{def}}{=} -\frac{\partial}{\partial \gamma} W_0(\gamma^A) = -\frac{\partial}{\partial \gamma} W_1(\gamma^M), \quad \text{i.e.,} \quad \sigma^A(\gamma^A) = \sigma^M(\gamma^M). \quad (4.96)$$
We are faced, at this point, with the problem:

Given \( \overline{\chi} \) and \( \gamma \), determine \( \sigma = \sigma^*(\gamma, \overline{\chi}) \).

Presumably we can find the function \( \sigma^* \) by first determining \( \gamma^A \) and \( \gamma^M \) from (4.95) and (4.96). For our present purposes we need not investigate this fully. We may simply assume that we can identify \( \sigma^*(\gamma, \overline{\chi}) \) locally where needed. We remark however that the determination of \( \sigma^* \) is not likely to be unique globally since \( \sigma^M(\cdot) \) is not injective, and that this non-uniqueness is responsible for the hysteresis observed in shape memory alloys.

Then the partial differential equation (4.81) becomes, in this setting,

\[
  u_{tt}(x, t) = -\partial_x (\sigma^*(\gamma(x, t), \overline{\chi}(x, t))) \quad \text{with} \quad \gamma(x, t) \overset{\text{def}}{=} \langle u_x \rangle(x, t). \tag{4.97}
\]

The averaging operator \( \langle \cdot \rangle \) is given by (4.95). We expect the locally averaged version of the solutions of (4.81) with highly oscillatory gradients to satisfy (4.97) as the oscillatory controls \( \chi(\gamma \rightarrow) \) converge to \( \overline{\chi} \).

What is important is a situation in which we have a solution \( u \) of \( u_{tt} = u_{xx} + \phi_x \) for some \( \phi = \phi(x, t) \) and we wish to interpret it as a solution of the nonlinear equation (4.89), i.e., with \( \phi = \chi(x, t) \Gamma(u_x(x, t)) \). Note that we are taking \( \mathcal{W}_0(\gamma) = \frac{1}{2} \gamma^2 \) for austenite. Thus

\[
  \sigma^A(\gamma) \overset{\text{def}}{=} -\partial_y \mathcal{W}_0 = -\gamma.
\]

Whence the term \( u_{xx} \) is just \( -\partial_x (\sigma^A(\gamma)) \) with \( \gamma = u_x \). Consequently, we want to interpret \( \phi \) as \( \sigma^*(\gamma, \overline{\chi}) - \gamma \). Hence our problem, instead of determining \( \sigma \) from \( \gamma, \overline{\chi} \) as before, becomes

Given \( \gamma \) and \( \phi \), determine \( \overline{\chi} \) so that \( \sigma^*(\gamma, \overline{\chi}) = \phi + \gamma \).
This problem is much easier, since we have\footnote{We do note that \((\sigma^M)^{-1}\) would, in general, be set-valued for a double-well potential \(W_1\), but assume that we will be working locally in an interval where \(\sigma^M\) is increasing, above the upper inflection point in \(W_1\), so this can be taken as simply defined.}

\[ \gamma^A = (\sigma^A)^{-1} (\phi + \gamma), \quad \gamma^M = (\sigma^M)^{-1} (\phi + \gamma) \quad (4.98) \]

With \(\gamma^A, \gamma^M\) now determined by (4.98) from \(\sigma\), we have from (4.95) that

\[ \overline{\chi} = \frac{\gamma - \gamma^A}{\gamma^M - \gamma^A} \quad (4.99) \]

so we see that \(\overline{\chi} \in [0, 1]\), as desired, if and only if

\[ \gamma^A \leq \gamma \leq \gamma^M. \quad (4.100) \]

Our basic assumptions here are that we can consider (a sequence of) approximating control functions for which the fraction \(\overline{\chi}_h\) remains fairly constant, and that this approximate constancy remains true as the domains become small enough compared to observable regions of the wire: \(h \ll 1\) so that we are approximating a well-defined “limit control” \(\overline{\chi}(x, t)\). Note that if \(\overline{\chi}\) is to take values other than 0,1, we must vary the approximating controls \(\chi\) within these small domains, i.e., \(\chi\) must rapidly alternate so we may guarantee that the relevant energies will remain bounded as \(\chi \to \overline{\chi}\).

If \(\{u_x\}\) remains \(L^2\)-bounded, then (for a subsequence) we obtain weak convergence \(u_x \rightharpoonup \overline{u}_x\) as \(t \to T\). Then \(\overline{u}_x\) represents the expected value of \(u_x\). Assuming ergodicity of the control process, we can write the expected value in the form of the local average given by (4.95), i.e.,

\[ \overline{u}_x(x) = \overline{\chi}(x, t)\gamma^M + (1 - \overline{\chi}(x, t))\gamma^A. \quad (4.101) \]

We obtain \(\overline{\chi}\) as the measure-valued limit of \(\chi\). In view of the above, we introduce the following
**Definition 4.3.2.** We say that a solution \( u \) of (4.89) is a generalized solution of (4.81) corresponding to the generalized control \( \overline{x} \) if \( \overline{x} \) takes values in \([0, 1]\) and, pointwise with \( \gamma = u_x \), one has

\[
\sigma^*(\gamma, \overline{x}) - \gamma = \phi.
\]

We note that the earlier constraint (4.91) has not entirely disappeared, but has been relaxed to ensure that \( \overline{x} \in [0, 1] \) when we compute \( \overline{x} \) from \( u, \phi \) by taking \( \gamma \overset{\text{def}}{=} u_x \) pointwise and applying (4.99) with (4.98): we now have the requirement that, pointwise,

\[
(\sigma^A)^{-1}(\phi + \gamma) \leq \gamma \leq (\sigma^M)^{-1}(\phi + \gamma).
\]  \hspace{1cm} (4.102)

With \( \overline{x} \) (and the generalized solution \( u \)) already defined, we may hope to obtain approximating (highly oscillatory) controls \( \chi \) for the original problem (4.81) for which the discussion above correctly describes the behavior. Of course, any such function \( \overline{x} \) is obtainable as a limit of \( \{0, 1\} \)-valued functions but, at present, we still have no rigorous proof of even the existence of corresponding solutions of (4.81), much less hard estimates permitting a convergence argument.

In order to obtain an approximating control for the original problem (4.81) satisfying the strong constraint (4.91) from our generalized solution \( \overline{x} \), we may then construct any highly oscillating control \( \chi(x, t) \in \{0, 1\} \) such that

\[
\overline{x}(x, t) = \lim_{h \to 0} \int_x^{x+h} \chi(\tau, t) \, d\tau.
\]  \hspace{1cm} (4.103)

In view of our discussion, this is plausible since (4.102) gives \( \overline{x} \) taking values in \([0, 1]\), which is obtainable as a limit of \( \{0, 1\} \)-valued functions as in (4.93). Note that if \( \phi \) is smooth, then the solution (4.92) will be smooth and we would expect \( \overline{x} = \phi / \Gamma(u_x) \) to give \( \overline{x} \) smooth. Conversely, if \( \overline{x} \) is smooth and small (and \( \Gamma \) has bounded derivatives), then a fixed point argument gives solvability of (4.81) in the sense above.
4.3.2.3 A Controllability Result

In this section we consider a situation of the above form, for which it will be convenient to modify the martensitic energy function $\mathcal{W}_1$. As indicated in Figure 3.2, $\mathcal{W}_1$ is a double-well potential approximately symmetric about $\gamma = 0$; we let $\hat{\gamma}$ be the positive minimum, corresponding to equilibrium (at the temperature $\theta_1$) for a mildly extended wire, i.e., of length $(1 + \hat{\gamma})\ell$ if the reference configuration, at austenitic equilibrium, gives length $\ell$. We have taken $\mathcal{W}_0(\gamma) = \frac{1}{2}\gamma^2$, although this can be expected to be only an approximation which loses validity as one gets farther away from $\gamma = 0$. Although this is typically taken as a quartic (just to have a double-well potential), we will now use the quadratic density $\mathcal{W}_1(\gamma) = \frac{1}{2}(\gamma - \hat{\gamma})^2$ since we will here be concerned with $\mathcal{W}_1(\gamma)$ only for $0 < \gamma < \hat{\gamma}$. This is a perfectly reasonable approximation near $\hat{\gamma}$, but (especially as we must differentiate!) it would be quite artificial if $\gamma$ were to approach 0. From this we get

$$
\frac{\partial}{\partial \gamma} \mathcal{W}(\gamma, \chi) = \begin{cases} 
\gamma = -\sigma^A(\gamma) & \text{in austenite (}\chi = 0), \\
\gamma(x, t) - \hat{\gamma} = -\sigma^M(\gamma) & \text{in martensite (}\chi = 1). 
\end{cases} \tag{4.104}
$$

This easily enables us to compute $\sigma^*(\gamma, \bar{\chi})$ explicitly for $\bar{\chi} \in [0, 1]$; when $\sigma^A(\gamma_A) = \sigma^M(\gamma_M)$ we have here $\gamma^M - \gamma^A = \hat{\gamma}$, independent of $\gamma$, so

$$
\gamma^A = \gamma - \bar{\chi}\hat{\gamma}, \quad \gamma^M = \gamma + (1 - \bar{\chi})\hat{\gamma}, \quad \sigma = \sigma^*(\gamma, \bar{\chi}) = \gamma - \bar{\chi}\hat{\gamma}. \tag{4.105}
$$

As it happens, we here have

$$
\sigma^*(\cdot, \bar{\chi}) = -\frac{\partial \bar{\chi} \mathcal{W}_1 + (1 - \bar{\chi})\mathcal{W}_0}{\partial \gamma}(\gamma).
$$

Now choose some intermediate value $0 < \gamma_* < \hat{\gamma}$, preferably taking $\gamma_*$ moderately close to $\hat{\gamma}$ so our quadratic form of $\mathcal{W}_1$ may plausibly be a good approximation near $\gamma_*$. Setting $\chi_* \overset{\text{def}}{=} \gamma_*/\hat{\gamma}$, we then easily see that the material as given by $U = U(x) \overset{\text{def}}{=} \gamma_*x$ will be in equilibrium in the presence of the generalized control $\bar{\chi} \equiv \chi_*$. This means
not only that \( u(\cdot, t) \equiv U \) is a solution of (4.89) with \( \phi \) given by (4.88) so using \( \bar{x} \equiv x \) makes \( \phi \) a constant and \( \phi_x \equiv 0 \), but also that there is no tension in the wire. We will now take this new state as providing the reference configuration and then work with deviations from this state. We consider this situation for \( 0 \leq x \leq 1 \) and \( 0 \leq t \leq T \), where we will assume \( T > 2 \) for simplicity.

**Theorem 4.3.3.** Let the initial state \([u_0, u_1]\) and the target state \([u_0^T, u_1^T]\) each be close to \([U, 0]\) in \( H^1(0, 1) \times L^2(0, 1) \). Then there exists a generalized control \( \bar{x} \) near \( x \), and a corresponding generalized solution \( u \) of (4.81) in the sense of Definition 4.3.3 such that one attains the target at time \( t = T \): i.e., \( u(\cdot, T) = u_0^T \), \( u(\cdot, T) = u_1^T \).

**Proof.** Let \( v \overset{\text{def}}{=} u - U \) and \( f \overset{\text{def}}{=} -\gamma(\bar{x} - x)_x \). Then we wish \( v \) to satisfy

\[
v_{tt} = v_{xx} + f \quad \text{with} \quad v = 0 \text{ at } x = 0, 1, \tag{4.106}
\]

with

\[
v(\cdot, 0) = v_0 \overset{\text{def}}{=} u_0 - U, \quad v_t(\cdot, 0) = v_1 \overset{\text{def}}{=} u_1
\]

and attaining the target:

\[
v(\cdot, T) = u_0^T \overset{\text{def}}{=} u_0^T - U, \quad v_t(\cdot, T) = u_1^T \overset{\text{def}}{=} u_1^T.
\]

It is well known that one has exact controllability in, e.g., \( H^1 \times L^2 \) for the linear controlled wave equation (4.106) with controls \( f \) in \( C([0, T]; L^2(0, 1)) \). [The assumption that \( T > 2 \) would enable us to do this even with restrictions on the support of \( f \).] Further, one has an available bound of the form \( \|f(\cdot, t)\| \leq K\|\text{data}\| \) for, e.g., the minimum-norm control. If, as assumed, the data (for \( v \)) is small, then

\[
\hat{\phi}(x, t) \overset{\text{def}}{=} \int_0^x f(\xi, t) \, d\xi
\]

will be uniformly small so that, since \( 0 < x \overset{\gamma}{\sim} \gamma < \hat{\gamma} \), one sees that \( \phi \overset{\text{def}}{=} \hat{\phi} - x \overset{\gamma}{\sim} \hat{\gamma} \) will remain between \( \Gamma = -\hat{\gamma} \) and 0. Thus, \( f = \phi_x \) with \( \phi \) satisfying the constraint
(4.102) of Definition 4.3.2. It follows that

$$\bar{\chi} \overset{\text{def}}{=} \hat{\phi}/(-\hat{\gamma}) + \chi_*$$

takes values in $[0, 1]$ as desired and so is admissible as the generalized control of the Theorem. Note that in computing $\bar{\chi}$ we do not need $\gamma^A, \gamma^M \in [0, \hat{\gamma}]$. \hfill \Box

These generalized controls, along with the theoretical damping properties discussed in Section 4.3.1 provide an indication that, indeed, thermally-induced vibration damping using shape memory wires is a worthy and possible goal. However, just as with the control results from other linearized models, the results using the simplified model in Section 4.3.2 provide only an approximation of possible controlled system behaviors. Even then, the computational results in Chapter 6 using the proposed model (4.9) indicate that, indeed, spatially-oscillatory temperature controls are very successful in damping vibrational energy from shape memory alloy wires.

Before presenting these computational results using the proposed model, however, we must discuss the methods used for approximating weak solutions to the system (4.9). Given the existence theory discussed in Section 4.2.3, specifically the results by Friescke and Dolzmann [59] and Niezgódka and Sprekels [114], along with the spirit of the Method of Vanishing Viscosity from Section 4.1.3, we discuss the proposed computational solution method in the next chapter.
Chapter 5

Solution Method

5.1 Computation of Thermodynamic Quantities

In the following solution and evaluation of the proposed model, we must calculate some thermodynamic quantities of the overall body. The thermodynamic densities $\varepsilon$ and $\eta$ provide the respective internal energy and internal entropy per referential unit volume; however, macroscale experimental observations are made on the overall energy and overall entropy of the body. With this in mind, we calculate the total internal energy (in Joules) in the wire at any given time as

$$\mathcal{E}(t) = \int_{\Omega} \rho_0 \varepsilon(x, t) \, dx = \rho_0 \beta \int_{0}^{L} \varepsilon(x, t) \, dx,$$

(5.1)

where $x$ denotes the spatial variable in three-dimensions and $x$ denotes the corresponding one-dimensional spatial variable. Note that the cross-sectional area $\beta$ must be included in the integral conversion from $\Omega$ to $[0, L]$. Similarly, we may calculate the overall entropy (in Joules/Kelvin) in the wire at the time $t$ according to

$$S(t) = \int_{\Omega} \rho_0 \eta(x, t) \, dx = \rho_0 \beta \int_{0}^{L} \eta(x, t) \, dx.$$

(5.2)

Since the proposed application of the model is in the thermally-activated damping of
vibrational energy, we must also compute the overall kinetic energy (in Joules) in the wire. This is given at the time \( t \) as

\[
\mathcal{K}(t) = \int_\Omega \rho_0 v(x, t)^2 \, dx = \rho_0 \beta \int_0^L v(x, t)^2 \, dx.
\]  

(5.3)

The final macroscale thermodynamic quantity of interest is that of the overall enthalpy in the wire at a given time. This is because the latent heat of transformation from a state \( P_1 \) to the state \( P_2 \), as discussed in Section 2.2.5, may be calculated as the change in enthalpy between the two states, \( L_{P_1 \to P_2} = H_{P_2} - H_{P_1} \) [31, 82, 134]. At constant pressure (as assumed in Section 3.2), the total enthalpy (in Joules) of the wire at the time \( t \) is given by

\[
H(t) = \int_\Omega \rho_0 \theta(x, t) \eta(x, t) \, dx = \rho_0 \beta \int_0^L \theta(x, t) \eta(x, t) \, dx.
\]  

(5.4)

## 5.2 Continuous Space-Time Galerkin

We now consider the computational solution of the proposed model. As discussed in Chapter 3, equation (3.35), we wish to find approximate solutions to the following system of nonlinear partial differential equations

\[
\rho_0 \ddot{u} = \rho_0 \partial_{xx}^2 \Psi + \alpha \partial_{xx} \dot{u} + \rho_0 b \\
\rho_0 c_p \dot{\theta} = \rho_0 \theta \partial_{xx}^2 \Psi \dot{\gamma} + \alpha \dot{\gamma}^2 + \kappa \partial_x (\gamma \partial_x \theta) + \rho_0 r,
\]  

(5.5)

where \( u = u(x, t) \) is the one-dimensional deformation, \( \theta = \theta(x, t) \) is the absolute temperature, \( \gamma = \gamma(x, t) = \partial_x u(x, t) \) is the deformation gradient (referential strain), \( \Psi = \Psi(\gamma, \theta) \) is the Helmholtz free energy, \( b = b(x, t) \) is the external body force per unit mass, and \( r = r(x, t) \) is the external heat supply per unit mass. In order to represent the case of a vibrating shape memory alloy wire having one end fixed in space and temperature and the other end free to move, in the computational simulations we
subject the this differential system to the following initial and boundary conditions:

\[ u(x, 0) = u_0, \quad \dot{u}(x, 0) = v_0, \quad \theta(x, 0) = \theta_0 \]
\[ u(0, t) = 0, \quad \dot{u}(0, t) = 0, \quad \theta(0, t) = \theta_{DC}(t), \quad \partial_x \theta(L, t) = 0. \]

Of particular importance in the numerical solution of such hyperbolic systems of conservation laws is the necessity for the discrete scheme to satisfy similar discrete versions of those conservation laws. This means that the discrete scheme used to compute solutions to the continuous Lyapunov system (5.5) must in turn provide a discrete Lyapunov system. This discrete analog of conservation of energy assures the long-term stability of the approximate solution and provides a discrete total thermo-mechanical energy which will remain unconditionally conservative or even dissipative in time.

For this purpose we employ Continuous Time Galerkin (CTG) Methods for the time-discretization of our PDE system. CTG methods were first introduced in the context of ordinary differential equations by Hulme [75, 76], and further expanded by French and Jensen [55, 56], French and Peterson [57], and French and Schaeffer [58]. According to these studies, CTG methods satisfy the required Lyapunov stability in time. This occurs because such methods retain continuity in time, so discrete versions of the energy conservation principles remain applicable to the resulting discrete systems. It is known that other numerical methods, such as the method of lines and various integral methods, also satisfy this discrete Lyapunov property. However, the further benefit to CTG schemes is that, when combined with finite element approximations of the spatial variables, they provide for a unified treatment of both spatial and temporal discretizations. Furthermore, CTG methods allow for a systematic method of proving existence of discrete solutions for long time scales, through examination of the Lyapunov nature of the total energy in time. Continuous Time Galerkin methods have thus been used successfully in the approximation of solutions to nonlinear wave equations similar to that used in the current model (see [57, 63, 64, 83, 84, 86, 130]
for further details).

We therefore choose to use piecewise linear approximations for our state variables \( u \) and \( \theta \) in both space and time, as dictated by the underlying regularity of our PDE system. We first consider the weak formulation of our system (5.5) in the context of Continuous Time Galerkin methods. We then consider the time discretization of the PDE system, along with the resulting nonlinear in space, implicit in time updates for our discrete solutions.

To use the Continuous Time Galerkin Method, we us first define the function spaces \( \mathcal{V} = C([0, \infty)) \), \( \mathcal{S} = \{ \chi \in C([0, L]) : \chi(0) = 0 \} \), and their product space \( \mathcal{A} = \mathcal{S} \otimes \mathcal{V} \). We define the following inner product notation, for use in expressing the weak formulation of our system (5.5),

\[
\langle u, v \rangle = \int_{\Omega} \int_{0}^{\infty} u(x, t) v(x, t) \, dt \, dx.
\]

Since the application of the Continuous Time Galerkin method to our system of equations (5.5) results in the linear momentum equation becoming a second-order-in-time ordinary differential equation in \( u \), we introduce the weak velocity \( v \in \mathcal{A} \) as the weak solution to the equation

\[
\langle v, \varphi \rangle = \langle \dot{u}, \varphi \rangle \quad \forall \varphi \in L^2(0, T; \mathcal{S}).
\]

With the above definitions and notation, we then wish to find functions \( (u, v, \theta) \in \mathcal{A} \times \mathcal{A} \times \mathcal{A} \) as a weak solution to the first-order-in-time thermodynamic system

\[
\begin{align*}
\langle \ddot{u} - v, \varphi \rangle &= 0 \\
\langle (\rho_0 \dot{\theta} - \rho_0 \partial_{xx} \Psi - \alpha \partial_{xx}^2 v - \rho_0 b), \varphi \rangle &= 0 \\
\langle \left( \rho_0 c_p \dot{\theta} - \rho_0 \theta \partial_{xx} \Psi \dot{\gamma} - \alpha \dot{\gamma}^2 - \kappa \partial_x (\gamma \partial_x \theta) - \rho_0 \gamma, \varphi \right) &= 0
\end{align*}
\]

for all test functions \( \varphi \in L^2(0, T; \mathcal{S}) \). We denote this nonlinear differential system as

\[
g(X, \varphi) = 0, \quad \text{where} \quad X(x, t) = [u(x, t), v(x, t), \theta(x, t)]^T \in \mathcal{A} \times \mathcal{A} \times \mathcal{A}.
\]
We now consider the following uniform space and time discretizations of our computational domain $[0, L] \times [0, \infty)$,

\begin{align*}
0 &= x_0 < x_1 < \ldots < x_N = L, \quad I_j = (x_{j-1}, x_j), \quad x_j - x_{j-1} = \Delta x \\
0 &= t_0 < t_1 < t_2 < \ldots, \quad J_n = (t_{n-1}, t_n), \quad t_n - t_{n-1} = \Delta t.
\end{align*}

Associated with these discretizations, we have the finite element approximation spaces

\begin{align*}
S_{\Delta x} &= \{ \chi \in S : \chi |_{I_j} \in P_s(I_j), \ j = 1, \ldots, N \} \subset C([0, L]) \\
V_{\Delta t} &= \{ \tau \in V : \tau |_{J_n} \in P_r(J_n), \ n = 1, \ldots, \infty \} \subset C([0, L]),
\end{align*}

and define the overall approximation space $\tilde{A} = S_{\Delta x} \otimes V_{\Delta t}$. Note: the spaces $P_s(I_j)$ and $P_r(J_n)$ denote the spaces of piecewise $s$-th order polynomials in space and piecewise $r$-th order polynomials in time, respectively. Also associated with the time discretization, we define the inner-products

\begin{align*}
\langle \langle u, v \rangle \rangle_n &= \int_0^L \int_{t_{n-1}}^{t_n} u(x, t) v(x, t) \, dt \, dx, \\
\langle u, v \rangle &= \int_0^L u(x) v(x) \, dx.
\end{align*}

Note that due to the linear integration over $[0, L]$ as opposed to three-dimensional integration over the body $\Omega$, we must multiply the one-dimensional integral by the cross-sectional area $\beta$, resulting in

\begin{equation}
\langle \langle u, v \rangle \rangle = \beta \sum_{n=1}^{\infty} \langle \langle u, v \rangle \rangle_n.
\end{equation}

Our goal is then to move the approximate finite element solution $(u, v, \theta) \in \tilde{A} \times \tilde{A} \times \tilde{A}$ through time by satisfying the weak system (5.6) on each interval $[0, L] \times J_n$. This process may be represented as finding the solution $X(x, t) |_{J_n}$ on each interval $J_n$ satisfying the vector-valued function

\begin{equation}
g^n(X, \varphi) = \begin{pmatrix}
\langle \langle \dot{u} - v, \varphi \rangle \rangle_n \\
\langle \langle \rho_0 \dot{v} - \rho_0 \partial_x^2 \Psi - \alpha \partial_x^2 \dot{u} - \rho_0 \dot{b}, \varphi \rangle \rangle_n \\
\langle \langle \rho_0 c_p \dot{b} - \rho_0 \theta \partial_\theta^2 \Psi \dot{\gamma} - \alpha \dot{\gamma}^2 - \kappa \partial_z (\gamma \partial_x \theta) - \rho_0 \dot{r}, \varphi \rangle \rangle_n
\end{pmatrix} = 0 \quad (5.8)
\end{equation}
for all test functions \( \varphi \in L^2(J_n; S_{\Delta x}) \).

Continuous Galerkin schemes have been evaluated for systems similar to the one proposed here by Aziz and Monk [13] in the case of the heat equation, and by R. Winther [133] for first-order hyperbolic systems of conservation laws. Both of these researchers found that not only did continuous-time Galerkin methods retain the Lyapunov nature in their respective discretized systems, but that both problems showed \( O(\Delta t^{2r}) \) convergence rates when using \( r \)-point Gaussian quadrature methods.

In the case of shape memory materials that undergo sudden phase transformations, although the thermodynamic quantities \( u, v \) and \( \theta \) may be considered to be continuous in time, their derivatives may suffer jumps (Note: \( v \) is the weak velocity associated with \( u \), and so therefore experiences higher regularity than \( \dot{u} \)). Since these thermodynamic quantities do not have sufficient regularity to take advantage of high-order methods in time, we take the finite element approximation space \( \mathcal{A} \) to be piecewise linear in time. The state variables \( u, v \) and \( \theta \) may therefore be represented in the time interval \( J_n \) according to the linear approximations

\[
\begin{align*}
    u(x,t) &= u(x,t_{n-1}) + \frac{t - t_{n-1}}{\Delta t} (u(x,t_n) - u(x,t_{n-1})) , \\
    v(x,t) &= v(x,t_{n-1}) + \frac{t - t_{n-1}}{\Delta t} (v(x,t_n) - v(x,t_{n-1})) , \\
    \theta(x,t) &= \theta(x,t_{n-1}) + \frac{t - t_{n-1}}{\Delta t} (\theta(x,t_n) - \theta(x,t_{n-1})) .
\end{align*}
\]

(5.9)

Furthermore, for computational simplicity we make the assumption that we may represent the quantities \( \gamma(x,t), r(x,t), b(x,t), \partial_x \Psi(x,t) \) and \( \partial^2_{\gamma,\theta} \Psi(x,t) \) linearly in time as well. We assume that errors introduced through these linear approximations are of the same magnitude as those introduced through the piecewise linear temporal discretization of \( u, v \) and \( \theta \). For notational simplification, we also denote the time-level states \( u(x,t_n) = u^n(x) \), and similarly for the other quantities mentioned above.

Using these approximations, the time evolution of our thermodynamic system (5.7) occurs through finding approximate weak solutions to the equation \( \mathbf{g}^n(X, \varphi) = 0 \). The
piecewise linear approximations of the state variables described above then allows us to denote this nonlinear time evolution method as

\begin{equation}
   \mathbf{g}^n(X^{n-1}, X^n, \varphi) = 0, \quad \text{where}
   X^n(x, t) = [u(x, t_n), v(x, t_n), \theta(x, t_n)]^T.
\end{equation}

Solution of the nonlinear system (5.10) for the updated time level states \(u(x, t_n), v(x, t_n)\) and \(\theta(x, t_n)\) is accomplished by using a variant on Newton’s method, discussed further in Section 5.4. This method requires both evaluation of the function \(\mathbf{g}^n(X^{n-1}, X^n, \varphi)\), and construction of the directional (or Gâteaux) derivatives \(D\mathbf{g}^n(X^{n-1}, X^n, \varphi)(\mathbf{X})\) (discussed in further detail in Sections 5.2.1 and 5.4) within particular time intervals \(J_n\). We consider the conversion of these terms into discrete time levels in the following Sections 5.2.1-5.2.3.

### 5.2.1 Time-Discretization of \(g^n_1(X^{n-1}, X^n, \varphi)\)

Beginning with the simplest of the three time-discretized nonlinear equations, representing the weak equivalence between \(\dot{u}\) and \(\dot{v}\), we now consider

\begin{equation}
   g^n_1(X^{n-1}, X^n, \varphi) = \langle \dot{u} - v, \varphi \rangle_n.
\end{equation}

We further simplify the notation by dropping all reference to the variables \(x\) and \(t\) within the thermodynamic quantities. Those functions continuous in time are denoted without a superscript (i.e. \(u(x, t) = u\)), and those spatial states constant at each time level will have a superscript (i.e. \(u(x, t_n) = u^n\)). Now, inserting the linear time approximations for \(u\) and \(v\) from (5.9), the equation (5.11) expands to become

\begin{align*}
   g^n_1(X^{n-1}, X^n, \varphi) &= \int_\Omega \int_{t_{n-1}}^{t_n} \left( \frac{u^n - u^{n-1}}{\Delta t} - v^{n-1} - \frac{t-t_{n-1}}{\Delta t} (v^n - v^{n-1}) \right) \varphi \, dt \, dx \\
   &= \int_\Omega \left( u^n - u^{n-1} - \frac{\Delta t}{2} (v^n + v^{n-1}) \right) \varphi \, dx \\
   &= \langle u^n - u^{n-1} - \frac{\Delta t}{2} (v^n + v^{n-1}) , \varphi \rangle
\end{align*}
We wish to solve \( g_1^n(X^{n-1}, X^n, \varphi) = 0 \). In order to limit roundoff errors, we divide the entire equation by \(-\Delta t\). This leaves the following set of inner products relating the weak velocity \( v \) to the deformation \( u \),

\[
g_1^n(X^{n-1}, X^n, \varphi) = \langle \frac{1}{2}(v^n + v^{n-1}) - \frac{1}{\Delta t}(u^n - u^{n-1}), \varphi \rangle. \tag{5.12}
\]

Variational methods for the solution of the nonlinear system (5.10) require the three directional (Gâteaux) derivatives \( Dg_1^n(X^{n-1}, X^n, \varphi)(\bar{X}) \), \( Dg_2^n(X^{n-1}, X^n, \varphi)(\bar{X}) \), and \( Dg_3^n(X^{n-1}, X^n, \varphi)(\bar{X}) \); the Gâteaux derivative is calculated as

\[
Df(X, \varphi)(\bar{X}) \overset{\text{def}}{=} \lim_{\tau \to 0} \frac{\partial}{\partial \tau} f (X + \tau \bar{X}, \varphi),
\]

and the vector-valued function \( \bar{X} = [\bar{u}, \bar{v}, \bar{\theta}]^T \) represents the direction in which the derivative for the functional (5.10) is taken [32]. The first Gâteaux derivative of the function \( g_1^n(X^{n-1}, X^n, \varphi) \) is thus given by

\[
Dg_1^n(X^{n-1}, X^n, \varphi)(\bar{X}) = \lim_{\tau \to 0} \frac{\partial}{\partial \tau} g_1^n(X^{n-1}, X^n + \tau \bar{X}, \varphi)
\]

\[
= \lim_{\tau \to 0} \frac{\partial}{\partial \tau} \langle \frac{1}{2}(v^n + \tau \bar{v} + v^{n-1}) - \frac{1}{\Delta t}(u^n + \tau \bar{u} - u^{n-1}), \varphi \rangle
\]

\[
= \langle \frac{1}{2}\bar{v} - \frac{1}{\Delta t}\bar{u}, \varphi \rangle. \tag{5.13}
\]

The use of these quantities is explored in further detail in the discussion of the Newton system in Section 5.4, which deals with solving the problem \( g_1^n(X^{n-1}, X^n, \varphi) = 0 \), \( \forall \varphi \in \mathcal{L}^2(J_n, S_{\Delta t}) \), for the updated time level \( X^n \).

### 5.2.2 Time-Discretization of \( g_2^n(X^{n-1}, X^n, \varphi) \)

The second part of the nonlinear system (5.10), corresponding to the balance of linear momentum in the model, is

\[
g_2^n(X^{n-1}, X^n, \varphi) = \langle (\rho_0 \dot{v} - \rho_0 \partial_{xx}^2 \Psi - \alpha \partial_{xx}^2 \dot{v} - \rho_0 b), \varphi \rangle_n. \tag{5.14}
\]
The various pieces of the integral (5.14) separate into the following time-discretized inner-products:

\[
\rho_0 \langle \hat{v}, \varphi \rangle_n = \rho_0 \int_{t_{n-1}}^{t_n} \int_{\Omega} \frac{1}{\Delta t} (v^n - v^{n-1}) \varphi \, dt \, dx
\]
\[
= \rho_0 \langle v^n - v^{n-1}, \varphi \rangle,
\]

\[-\alpha \langle \partial_{xx}^2 v, \varphi \rangle_n = \alpha \langle \partial_x v, \partial_x \varphi \rangle_n
\]
\[
= \alpha \int_{t_{n-1}}^{t_n} \int_{\Omega} \left( \partial_x v^{n-1} + \frac{t-t_{n-1}}{\Delta t} (\partial_x v^n - \partial_x v^{n-1}) \right) \partial_x \varphi \, dt \, dx
\]
\[
= \alpha \frac{\Delta t}{2} \langle \partial_x v^n + \partial_x v^{n-1}, \varphi \rangle,
\]

\[-\rho_0 \langle b, \varphi \rangle_n = -\rho_0 \int_{t_{n-1}}^{t_n} \int_{\Omega} \left( b^{n-1} + \frac{t-t_{n-1}}{\Delta t} (b^n - b^{n-1}) \right) \varphi \, dt \, dx
\]
\[
= -\rho_0 \frac{\Delta t}{2} \langle b^{n-1} + b^n, \varphi \rangle.
\]

Due to the assumption that the stress \( \Psi_{\gamma}(\gamma, \theta) \) may be approximated piecewise linearly in time, the remaining term from (5.14) becomes

\[-\rho_0 \langle \partial_{\gamma \gamma}^2 \Psi, \varphi \rangle_n = \rho_0 \langle \partial_{\gamma} \Psi, \partial_{\gamma} \varphi \rangle_n
\]
\[
= \rho_0 \int_{t_{n-1}}^{t_n} \int_{\Omega} \left( \partial_{\gamma} \Psi^{n-1} + \frac{t-t_{n-1}}{\Delta t} (\partial_{\gamma} \Psi^n - \partial_{\gamma} \Psi^{n-1}) \right) \partial_{\gamma} \varphi \, dt \, dx
\]
\[
= \rho_0 \frac{\Delta t}{2} \langle \partial_{\gamma} \Psi^n + \partial_{\gamma} \Psi^{n-1}, \partial_{\gamma} \varphi \rangle.
\]

Combining these terms together, the viscoelastic portion of the nonlinear system (5.10) corresponding to the balance of linear momentum reduces to

\[
g^n_2(X^{n-1}, X^n, \varphi) = \rho_0 \langle v^n - v^{n-1}, \varphi \rangle + \alpha \frac{\Delta t}{2} \langle \partial_x v^n + \partial_x v^{n-1}, \partial_x \varphi \rangle
\]
\[
+ \rho_0 \frac{\Delta t}{2} \langle \partial_{\gamma} \Psi^n + \partial_{\gamma} \Psi^{n-1}, \partial_{\gamma} \varphi \rangle - \rho_0 \frac{\Delta t}{2} \langle b^{n-1} + b^n, \varphi \rangle.
\]
As with the weak velocity equation from Section 5.2.1, the first Gâteaux derivative of the function \( g_2^n(X^{n-1}, X^n, \varphi) \) becomes

\[
Dg_2^n(X^{n-1}, X^n, \varphi)(\tilde{X}) = \lim_{\tau \to 0} \frac{\partial}{\partial \tau} g_2^n(X^{n-1}, X^n + \tau \tilde{X}, \varphi) \\
= \lim_{\tau \to 0} \frac{\partial}{\partial \tau} \left[ \rho_0 \left\langle v^n + \tau \overline{v} - v^{n-1}, \varphi \right\rangle \\
+ \alpha \frac{\Delta t}{2} \left\langle \partial_x v^n + \tau \partial_x \overline{v} + \partial_x v^{n-1}, \varphi_x \right\rangle \\
+ \frac{\rho_0 \Delta t}{2} \left\langle \partial_{\gamma} \Psi^n (\gamma + \tau \overline{\gamma}, \theta + \tau \overline{\theta}) + \partial_{\gamma} \Psi^{n-1}, \partial_x \varphi \right\rangle \\
- \frac{\rho_0 \Delta t}{2} \left\langle \overline{b}^{n-1} + b^n, \varphi \right\rangle \right] \\
= \rho_0 \left\langle \overline{v}, \varphi \right\rangle + \alpha \frac{\Delta t}{2} \left\langle \partial_x \overline{v}, \partial_x \varphi \right\rangle \\
+ \frac{\rho_0 \Delta t}{2} \left\langle \gamma \partial_{\gamma}^2 \Psi^n + \overline{\theta} \partial_{\theta}^2 \Psi^n, \partial_x \varphi \right\rangle.
\]
(5.16)

Again, these terms are used in the variant on Newton's method described in Section 5.4 for the solution of the nonlinear equation (5.10).

### 5.2.3 Time-Discretization of \( g_3^n(X^{n-1}, X^n, \varphi) \)

The last of the quantities comprising the nonlinear system (5.10), corresponding to the conservation of energy in the model, is

\[
g_3^n(X^{n-1}, X^n, \varphi) = \left\langle \rho_0 c_p \tilde{\theta} - \alpha (\tilde{\gamma})^2 - \kappa \partial_x (\gamma \partial_x \theta) - \rho_0 \theta \partial_{\gamma \theta}^2 \Psi \tilde{\gamma} - \rho_0 r, \varphi \right\rangle_n. \tag{5.17}
\]

Again the integral (5.17) is split into its various components. We first have the simplification

\[
\rho_0 c_p \left\langle \tilde{\theta}, \varphi \right\rangle_n = \rho_0 c_p \int_\Omega \int_{t_{n-1}}^{t_n} \frac{1}{\Delta t} (\theta^n - \theta^{n-1}) \varphi \, dt \, dx \\
= \rho_0 c_p \left\langle \theta^n - \theta^{n-1}, \varphi \right\rangle.
\]
Similarly, we reduce the following simple terms from (5.17):

\[- \alpha \left\langle (\dot{\gamma})^2, \varphi \right\rangle_n \]

\[= -\alpha \int_\Omega \int_{t_{n-1}}^{t_n} \left( \frac{1}{\Delta t} \right)^2 (\gamma^n - \gamma^{n-1})^2 \varphi \, dt \, dx \]

\[= -\alpha \frac{\Delta t}{\Delta t} \left\langle (\gamma^n - \gamma^{n-1})^2, \varphi \right\rangle, \]

\[- \kappa \left\langle \partial_x (\gamma \partial_x \theta), \varphi \right\rangle_n \]

\[= \kappa \left\langle \gamma \partial_x \theta, \partial_x \varphi \right\rangle_n \]

\[= \kappa \int_\Omega \int_{t_{n-1}}^{t_n} \left[ (\gamma^n - \gamma^{n-1}) \left( \partial_x \theta^n - \frac{t-t_{n-1}}{\Delta t} (\partial_x \theta^n - \partial_x \theta^{n-1}) \right) \partial_x \varphi \right] \, dt \, dx \]

\[= \kappa \frac{\Delta t}{\delta} \left\langle \gamma^n (2 \partial_x \theta^n + \partial_x \theta^{n-1}) + \gamma^{n-1} (\partial_x \theta^n + 2 \partial_x \theta^{n-1}), \partial_x \varphi \right\rangle, \]

\[- \rho_0 \left\langle r, \varphi \right\rangle_n \]

\[= -\rho_0 \int_\Omega \int_{t_{n-1}}^{t_n} \left( r^{n-1} + \frac{t-t_{n-1}}{\Delta t} (r^n - r^{n-1}) \right) \varphi \, dt \, dx \]

\[= -\rho_0 \frac{\Delta t}{2} \left\langle r^{n-1} + r^n, \varphi \right\rangle. \]

The more complicated term from the integral (5.17) involving \(\partial_{\gamma \theta}^2 \Psi (\gamma, \theta)\) again relies on the assumption that it may be approximated linearly in time. This results in the simplification

\[- \rho_0 \left\langle \theta \partial_{\gamma \theta}^2 \Psi \dot{\gamma}, \varphi \right\rangle_n \]

\[= -\rho_0 \int_\Omega \int_{t_{n-1}}^{t_n} \left[ \frac{1}{\Delta t} (\gamma^n - \gamma^{n-1}) \left( \theta^n + \frac{t-t_{n-1}}{\Delta t} (\theta^n - \theta^{n-1}) \right) \right. \]

\[\left. \left( \partial_{\gamma \theta}^2 \Psi^{n-1} + \frac{t-t_{n-1}}{\Delta t} (\partial_{\gamma \theta}^2 \Psi^n - \partial_{\gamma \theta}^2 \Psi^{n-1}) \right) \varphi \right] \, dt \, dx \]

\[= -\rho_0 \frac{\Delta t}{\delta} \left\langle (\gamma^n - \gamma^{n-1}) (\theta^n (2 \partial_{\gamma \theta}^2 \Psi^n + \partial_{\gamma \theta}^2 \Psi^{n-1}) + \theta^n (\partial_{\gamma \theta}^2 \Psi^n + 2 \partial_{\gamma \theta}^2 \Psi^{n-1})), \varphi \right\rangle, \]
Combining these terms together, the weak form of the energy equation given by (5.17) becomes the time-discretized equation

\[ g_3^n(X^{n-1}, X^n, \varphi) = \rho_0 c_p \left( \theta^n - \theta^{n-1}, \varphi \right) \]
\[ - \frac{\alpha}{\Delta t} \left( \left( \gamma^n - \gamma^{n-1} \right)^2, \varphi \right) \]
\[ + \kappa \frac{\Delta t}{6} \left( \gamma^n (2 \partial_x \theta^n + \partial_x \theta^{n-1}) + \gamma^{n-1} (\partial_x \theta^n + 2 \partial_x \theta^{n-1}), \partial_x \varphi \right) \]
\[ - \frac{e_0}{6} \left( \theta^n (\gamma^n - \gamma^{n-1}) (2 \partial_{\gamma \theta} \Psi^n + \partial_{\gamma \theta} \Psi^{n-1}), \varphi \right) \]
\[ - \frac{e_0}{6} \left( \theta^{n-1} (\gamma^n - \gamma^{n-1}) (\partial_{\gamma \theta} \Psi^n + 2 \partial_{\gamma \theta} \Psi^{n-1}), \varphi \right) \]
\[ - \rho_0 \frac{\Delta t}{2} \left( r^{n-1} + r^n, \varphi \right) . \]

(5.18)

As with the terms \( g_1^n(X^{n-1}, X^n, \varphi) \) and \( g_2^n(X^{n-1}, X^n, \varphi) \), we calculate the Gâteaux derivative of \( g_3^n(X^{n-1}, X^n, \varphi) \) according to

\[ Dg_3^n(X^{n-1}, X^n, \varphi)(\bar{X}) \]
\[ = \lim_{\tau \to 0} \frac{\partial}{\partial \tau} g_3^n(X^{n-1}, X^n + \tau \bar{X}, \varphi) \]
\[ = \lim_{\tau \to 0} \frac{\partial}{\partial \tau} \left[ \rho_0 c_p \left( \theta^n + \tau \bar{\theta} - \theta^{n-1}, \varphi \right) \right. \]
\[ - \frac{\alpha}{\Delta t} \left( \left( \gamma^n + \tau \bar{\gamma} - \gamma^{n-1} \right)^2, \varphi \right) \]
\[ + \kappa \frac{\Delta t}{6} \left( \gamma^n (2 \partial_x \theta^n + 2 \tau \partial_x \bar{\theta} + \partial_x \theta^{n-1}), \partial_x \varphi \right) \]
\[ + \kappa \frac{\Delta t}{6} \left( \gamma^{n-1} (\partial_x \theta^n + \tau \partial_x \bar{\theta} + 2 \partial_x \theta^{n-1}), \partial_x \varphi \right) \]
\[ - \frac{e_0}{6} \left( \gamma^n + \tau \bar{\gamma} - \gamma^{n-1} (\theta^n + \tau \bar{\theta}) (2 \partial_{\gamma \theta} \Psi^n (\gamma + \tau \bar{\gamma}, \theta + \tau \bar{\theta}) + \partial_{\gamma \theta} \Psi^{n-1}), \varphi \right) \]
\[ - \frac{e_0}{6} \left( \gamma^n + \tau \bar{\gamma} - \gamma^{n-1} \theta^{n-1} (\partial_{\gamma \theta} \Psi^n (\gamma + \tau \bar{\gamma}, \theta + \tau \bar{\theta}) + 2 \partial_{\gamma \theta} \Psi^{n-1}), \varphi \right) \]
Using the chain rule and product rules for differentiation, the above calculation becomes the following set of inner-products:

\[
DG^n_3(X^{n-1}, X^n, \varphi)(X) = \rho_0 c_p \langle \bar{\theta}, \varphi \rangle \\
- \frac{2\alpha}{\Delta t} \langle \bar{\gamma} \left( \gamma^n - \gamma^{n-1} \right), \varphi \rangle \\
+ \frac{\kappa \Delta t}{6} \langle \bar{\gamma} \left( 2 \partial_x \theta^n + \partial_x \theta^{n-1} \right), \partial_x \varphi \rangle \\
+ \frac{\kappa \Delta t}{6} \langle \partial_x \bar{\theta} \left( 2 \gamma^n + \gamma^{n-1} \right), \partial_x \varphi \rangle \\
- \frac{\varepsilon_0}{6} \langle \bar{\gamma} \theta^n (2 \partial^2_{\gamma \theta} \psi^n + \partial^2_{\gamma \theta} \psi^{n-1}), \varphi \rangle \\
- \frac{\varepsilon_0}{6} \langle \bar{\gamma} \theta^{n-1} (\partial^2_{\gamma \theta} \psi^n + 2 \partial^2_{\gamma \theta} \psi^{n-1}), \varphi \rangle \\
- \frac{\varepsilon_0}{6} \langle \bar{\gamma} (\gamma^n - \gamma^{n-1}) (2 \theta^n + \theta^{n-1}) \partial^3_{\gamma \theta} \psi^n, \varphi \rangle \\
- \frac{\varepsilon_0}{6} \langle \bar{\theta} (\gamma^n - \gamma^{n-1}) (2 \partial^2_{\gamma \theta} \psi^n + \partial^2_{\gamma \theta} \psi^{n-1}), \varphi \rangle \\
- \frac{\varepsilon_0}{6} \langle \bar{\theta} (\gamma^n - \gamma^{n-1}) (2 \partial^3_{\gamma \theta} \psi^n + \theta^{n-1} \partial^3_{\gamma \theta} \psi^n), \varphi \rangle.
\]

\[\text{(5.19)}\]

\[\text{5.2.4 The Spatial Discretization}\]

It is the case with such multi-equilibrium materials as shape memory alloys that, although material deformations remain continuous, the distinct crystal lattices associated with individual material phases result in spatially-discontinuous deformation gradients. Furthermore, because of the inclusion of a discontinuous referential heat flux within the model (see Section 3.2.5), the same relationship may be seen between the temperature and temperature gradients. Therefore we may only expect the time level states of our weak solutions \(\{u^n, v^n, \theta^n\}\) to reside in the Hilbert space \(H^1((0, L))\). Thus, we choose to represent the spatially-discretized approximations of these solutions using the set of continuous, piecewise affine functions \(\{\phi_j\}_{j=1}^N\), given on the interval \(I_j\) as

\[
\phi_{j-1}(x) = \frac{x_j - x}{\Delta x}, \quad \phi_j(x) = \frac{x - x_{j-1}}{\Delta x}, \quad \phi_k(x) = 0
\]
for \( k \notin \{j-1,j\} \). We may therefore write each of our time level functions at the time \( t_n \) according to

\[
\begin{align*}
    u^n(x) &= \sum_{j=1}^{N} u^n_j \phi_j(x), \\
v^n(x) &= \sum_{j=1}^{N} v^n_j \phi_j(x), \\
\theta^n(x) &= \sum_{j=1}^{N} \theta^n_j \phi_j(x).
\end{align*}
\]

Using this spatial discretization scheme, a standard finite element approach is used to calculate the various inner-products within the terms \( g^n_1, g^n_2, g^n_3, \text{ } Dg^n_1, \text{ } Dg^n_2 \text{ and } \text{ } Dg^n_3 \). With this approach, the full nonlinear system (5.10) may be equivalently written as the following finite-dimensional nonlinear system for the coefficients \( X_{\Delta x}^n = [u^n_1, u^n_2, \ldots, u^n_N, v^n_1, \ldots, v^n_N, \theta^n_1, \ldots, \theta^n_N]^T \):

\[
g^{n,\Delta x}(X_{\Delta x}^{n-1}, X_{\Delta x}^n, \varphi) = \begin{pmatrix}
g_1^{n,\Delta x}(X_{\Delta x}^{n-1}, X_{\Delta x}^n, \varphi) \\
g_2^{n,\Delta x}(X_{\Delta x}^{n-1}, X_{\Delta x}^n, \varphi) \\
g_3^{n,\Delta x}(X_{\Delta x}^{n-1}, X_{\Delta x}^n, \varphi)
\end{pmatrix} = 0, \tag{5.20}
\]

for all \( \varphi \in L^2(J_n, S_{\Delta x}) \). We note that the subscript/superscript \( \Delta x \) refers to the spatially-discrete form of the corresponding continuous equation or variable. Using such a discretization scheme, we expect that the conforming approximations of the true time level solutions are accurate to \( O(\Delta x^2) \) (see [25, 51] for further information).

### 5.3 Numerical Difficulties

The problem of finding an appropriate weak solution of the finite-dimensional nonlinear system (5.20) may be thought of as a finite-dimensional, nonlinear, nonconvex root-finding problem for the set of coefficients \( X_{\Delta x}^{n-1} \in \mathbb{R}^{3N} \). However, due to the non-convexity of the Helmholtz free energy (Section 3.4), the numerical solution of the nonlinear system \( g^{n,\Delta x}(X_{\Delta x}^{n-1}, X_{\Delta x}^n, \varphi) = 0 \) encounters difficulties at the moment of the phase transition. As discussed in Chapter 4, theoretical results have only been able to prove existence of weak solutions for sufficiently large viscosity coefficients \( \alpha \) (see [59, 72, 114, 104]). However, viscous effects are not observed in physical tests.
on shape memory alloys (see Seelecke et al. [128, 129]). Furthermore, Antman and Seidman have shown that the addition of such linear viscous effects to material models may violate basic physical principles such as material frame indifference and the principle of observer invariance (see [9, 10, 11]). We are thus faced with the difficulty of balancing solvability of the problem, which includes an artificial numerical dissipation (represented by the inflated viscosity coefficient $\alpha$), with accuracy in the model solution (represented by the zero-viscosity solution).

These numerical instabilities of the system in the zero-viscosity case result from the general properties of the nonlinear PDE system discussed in Chapter 4. Essentially, this viscous dependence may be thought of as the transition of the PDE system between ill-posed hyperbolic and parabolic regimes. When the dynamics stay within one material phase, the solution to the hyperbolic system is relatively simple since it is behaves as a linear wave equation. However, at the moment of phase transition, the values of the deformation gradient and heat flux jump from one value to another and must traverse an energy barrier between the two stable states. Consequently, the transition of solutions to the hyperbolic system over a region of negative eigenstructure results in instabilities in the model solution, rendering the model ill-posed.

This overall instability during the phase transition manifests itself in the numerical solution through indefiniteness in the eigenstructure of the discretized problem. In the solution of this discrete root-finding problem, these points of instability correspond to local minima in the optimization surface with respect to the solution coefficients. Furthermore, since the martensitic phase transformation occurs simultaneously throughout many regions of the wire, the optimization surface will have such local minima corresponding to each transforming element. It is well-known that the root-finding properties of the standard Newton method have difficulties at local minima. This is due to the nature of the Newton method, which updates iterates as solutions to a linear fit of the root-finding surface. When iterates approach a local
minimum, however, this linear fit provides a dramatically poor approximation of the
surface since the value of such local minima are nonzero while the linear fit is nearly
flat. Because of this, Newton updates near these local minima blow up, bringing the
iterates far from the desired roots.

These unstable regions may be circumvented, however, through the addition of a
strong viscosity to the system. Such viscosity terms are required by the theory
regarding these nonconvex problems (Section 4.2.3), since the introduction of this
viscosity essentially adds a weighted Laplacian to the differential system, stabilizing
the problem by shifting the corresponding eigenstructure to a definite region, and by
removing such local minima from the discrete root-finding surface. This stabilization
of the nonlinear problem has the effect of slowing down the phase transformation
dynamics. For many applications in the modeling of shape memory alloys, the quasi-
static or long-term behavior of these wires is desired. In those cases, the introduction
of the increased viscosity and corresponding retardation of the dynamics poses no
problem to the eventual solution. However, in such dynamic applications as proposed
here, where the temporally local behavior of solutions is desired, such a slowing of
the computed dynamics is unacceptable.

In order to resolve this issue, we considered using various optimization methods from
the literature for computation of the zero-viscosity solution. In particular, we at-
ttempted this global optimization using the standard Damped Newton algorithm, the
Mountain Pass algorithm by Moré and Munson [108], the Tunneling Newton algo-
rum of Levy and Gomez [66, 95, 96] and the Terminal Repeller (TRUST) algorithm
of Cetin, Barhen and Burdick [30].

The Damped Newton algorithm provides a simple change to the original Newton algo-
rum. This change amounts to the addition of a standard line search to the Newton
updates. Although this stabilizes the Newton iterates and keeps them from blowing
up, this method is known to stagnate at local minima [95]. Due to this stagnation in
the model solution, use of the Damped Newton algorithm for the root-finding problem cannot allow for transitions between the various phases of the material.

Similarly unhelpful to this problem is the Mountain Pass algorithm. Although this algorithm works well for finding the minimal “mountain pass” between two functional values, it requires knowledge of both the starting and final values before computing the optimal pass between them. However, in the root-finding problem discussed here, the final value is unknown. Thus the benefits from the Mountain Pass algorithm, including the ability to traverse regions of zero and negative eigenstructure, may not be used in this case.

The promising methods of the Tunneling Newton and Terminal Repeller algorithms for the global optimization fail in this case as well. These methods attempt to overcome local minima by considering the root-finding on an alternative optimization surface having the same roots without the local minima. Instead of considering the original root-finding problem applied to the function $f(y) = 0$ with local minimum at $y_s$, these methods consider solution of the problems

$$
\hat{f}(y) = \frac{f(y) - f(y_s)}{\|y - y_s\|_2^2} = 0 \quad \text{(Tunneling Newton)}
$$

and

$$
\hat{f}(y) = \log \left( \frac{1}{1 + \exp (y_s - y - \alpha)} \right) = 0 \quad \text{(Terminal Repeller)}.
$$

Although these methods indeed remove the local minimum at $y_s$ and retain the same set of roots as the original function $f$, they do not work in the present case. We believe that this is due to one of two causes. The first difficulty is in the choice of the ‘tuning’ parameter $\alpha$, the choice of random updates to the initial solution guess when moving away from a local minimum, and various other tuning parameters inherent to their use. These details must be chosen very carefully to allow for successful computation, often depending on the intuition and expertise of the algorithms’ inventor. Although we attempted a variety of choices for these parameters, no appropriate combination of
parameters were found. This does not rule out the possibility of using these methods on the current problem, but it does elucidate some of the difficulties inherent in their use. The second possibility as to the failure of these methods is the large size of the optimization space. In the papers by Levy and Gomez ([66, 95, 96]) and the paper by Cetin, Barhen and Burdick ([30]), the successful examples shown have many local minima and anywhere from one to seven optimization variables. The current problem, however, has $3 \times N$ optimization variables (where $N$ is the spatial discretization size). We believe it may also be this size that impedes the reportedly global convergence of these methods in the current problem.

We invented the final two unsuccessful methods for the root-finding problem, in which we strived to use prior knowledge of the material behavior during the martensitic first-order phase transformation to aid in the optimization problem. Although the idea of a "black box" approach to the root-finding problem was more appealing, the lack of provably accurate algorithms made us believe that knowledge of the crystal behavior in shape memory materials seemed crucial for successful methods to the nonconvex problem. In the first of these methods, we attempted to surpass the difficulties from the regions of negative curvature in the energy by implementation of a surrogate energy function at the moment of phase transition. For this method, the smoothness of the free energy $\Psi$ with respect to the stretch $\gamma$ was dropped in lieu of its piecewise convex, quadratic approximation. The idea behind this approach relied on the fact that a piecewise convex quadratic approximation of the free energy does not have the local minima problems in its first derivative. This differential structure also proved to cause the failure of the method. Since such approximations have discontinuities in their first derivative, many of the meaningful thermodynamic quantities computed in the evaluation of $g^{n,\Delta x}(X_{\Delta x}^{n-1}, X_{\Delta x}^n, \varphi)$ could not be calculated correctly, leading to the inaccuracy of the approximate solutions.

The final unsuccessful method that we invented attempted to use knowledge of the
various stable crystal states to provide optimal initial guesses for the root-finding optimization method. At the moment of phase transformation, the method evaluated the current solution to determine an intelligent guess for the new root. This initial guess was then used as an input to the standard Newton method. It is well-known that the standard Newton method converges quadratically in the case where the original energy is locally convex around the solution, and the initial iterate is within the domain of attraction of this solution. Although this method was unsuccessful in producing roots to the discrete problem, the approach may still prove highly successful given a proper method for determining these initial guesses.

5.4 Variable Viscosity Newton Method

Since we were unable to resolve the difficulties inherent in the root-finding problem discussed above using traditional optimization methods and could not invent a method for taking advantage of our knowledge of the stable material states, we again returned to the existing literature on the problem of phase transitions. From considering the success of methods on the perturbed problems having high viscosity coefficients (Section 4.2.3), and due to the similarity to the well-known Method of Vanishing Viscosity used in the solution of scalar conservation laws (Section 4.1.3), we propose the following variable viscosity approach for the model solution. In this approach the moment of phase transition is detected by monitoring iterates in the damped Newton method. Once these iterates show signs of stagnation, and thus phase transition in the model solution, the viscosity coefficient $\alpha$ is increased to the level proposed by theoretical results in order to render the model well-posed. The standard Newton method is then applied to this perturbed system to obtain a perturbed solution for the current time step. From this point, the viscosity level $\alpha$ is progressively lowered, using the previous perturbed solutions as the initial guesses.
for the Newton root-finding scheme to better calculate the same time step. Denoting the original zero-viscosity root-finding problem as $\bar{g}(X, \varphi) = g^n_{\Delta x}(X^{n-1}, X^n, \varphi) = 0$ with corresponding solution $X^*$, and denoting the viscosity-perturbed problem as $\bar{g}^\alpha(X, \varphi) = 0$ with corresponding solution $X^\alpha$, this process may be written in the following manner:

**Definition 5.4.1 (Variable Viscosity Newton Method).**

0. Initialize $\alpha = 0$ and $X = X^n$.

1. Construct the linear system $D\bar{g}(X, \varphi)(\bar{X}) = -\bar{g}(X, \varphi)$, where

$$
D\bar{g}(X, \varphi)(\bar{X}) = \begin{bmatrix}
D\bar{g}_1(X, \varphi)(\bar{u}) & D\bar{g}_1(X, \varphi)(\bar{v}) & 0 \\
D\bar{g}_2(X, \varphi)(\bar{u}) & D\bar{g}_2(X, \varphi)(\bar{v}) & D\bar{g}_2(X, \varphi)(\bar{\theta}) \\
D\bar{g}_3(X, \varphi)(\bar{u}) & 0 & D\bar{g}_3(X, \varphi)(\bar{\theta})
\end{bmatrix}
$$

$$
\bar{X} = \begin{bmatrix}
\bar{u} \\
\bar{v} \\
\bar{\theta}
\end{bmatrix}^T
$$

$$
\bar{g}(X, \varphi) = [\bar{g}_1(X, \varphi) \quad \bar{g}_2(X, \varphi) \quad \bar{g}_3(X, \varphi)]^T.
$$

2. Attempt to solve the above system for $\bar{X}$. If no solution exists, then the system is stuck at a local minimum. Skip to step 5.

3. Set $\tau = 1$ and halve $\tau$ progressively until $\bar{g}(X + \tau\bar{X}, \varphi) < \bar{g}(X, \varphi)$. If this fails, or if $\tau$ decreases below a minimum tolerance, then the system is approaching a local minimum. Skip to step 5.

4. Update the Newton iterate $X = X + \tau\bar{X}$. If the stopping criteria is met, exit. Otherwise return to step 1.

5. Set $\alpha = \alpha_{\text{max}}$ and $X^\alpha = X^n$.

6. Construct the linear system

$$
D\bar{g}^\alpha(X^\alpha, \varphi)(\bar{X}) = -\bar{g}^\alpha(X^\alpha, \varphi).
$$
7. Attempt to solve the above system for $\bar{X}$. If no solution exists, return $X^\alpha$ as the final perturbed solution.

8. Update the viscous Newton iterate $X^\alpha = X^\alpha + \bar{X}$. If the Newton stopping criteria is met, halve $\alpha$. Return to step 6.

The algorithm proposed above is designed to serve two key purposes. First, it assures that the overall time-stepping scheme remains stable by reverting to the solution of the perturbed, parabolic system $\tilde{g}^\alpha(X, \varphi)$ during moments of computational difficulty. Second, the viscous Newton method shown here reduces the amount of error in the solution to the dynamic system introduced through the artificial dissipation. Typically, the method as described here achieves solutions to perturbed systems using viscosity coefficients at least an order of magnitude smaller than those required without reduction of the viscosity (i.e. [59, 122, 123]).

However, it is unlikely that the root-finding scheme discussed here is the most accurate means of solving the original PDE system. Further examination of more sophisticated optimization methods is required in order to find more accurate means of solving the hyperbolic system at the moment of phase transition. As discussed above, many of these generic optimization methods were considered for this role in the overall solution method (Terminal Repeller, Tunneling Newton, etc.), and the proposed variable viscosity method was chosen due to its provably high stability. Through this choice, the decision between accuracy, computational efficiency and stability turned in the favor of efficiency and stability, mostly because small errors in the computed dynamics were chosen to be preferable to total collapse of the computational system.

In essence, the variable viscosity method proposed here provides a sort of homotopy or continuation method for the solution of the nonlinear system at a particular time step. These methods connect the solution of an “easy problem” (large viscosity) to the solution of a “hard problem” (zero viscosity) through traditional methods for ordinary
differential equations (see Nečas [112], Zhang [135]). The proposed method achieves this through an explicit Euler-type scheme for updating the viscosity coefficient to move from the large viscosity solution to the small viscosity solution. Further research into the use of higher-order implicit techniques for ordinary differential equations within this approach may provide means of further reducing the artificial viscosity required for the final solution.

5.5 Implementation Details

Within the implementation of the Continuous Time Galerkin, finite element and variable viscosity root-finding methods, a number of numerical techniques have been used to guarantee highly accurate, robust and fast numerical procedures. We detail some of these techniques in the remainder of this section.

Numerical Quadrature

The first of these relevant implementation details regards the numerical quadrature of selected integrals. We have performed the temporal integration exactly under the previous assumption on linearity in time of our thermodynamic quantities. This corresponds to a quadrature method exact to $O(\Delta t^2)$ [13, 133, 55, 56, 57, 58]. Within the evaluation of the spatial integrals, all efforts have been made to provide exact integrals of the arguments. However, for integrals involving the Helmholtz free energy functional $\Psi (\gamma(x, t), \theta(x, t))$, exact integration is considerably more difficult, due to the nonlinear dependence on $\gamma$ and $\theta$. Thus we use the following $O(\Delta x^4)$-accurate Gaussian quadrature formula with two quadrature points on each interval $I_j$ to perform the numerical integration:

$$\int_0^1 f(x) \, dx \approx 0.5 \ast f(x_l) + 0.5 \ast f(x_r),$$

$$x_l = \frac{1}{2} - \frac{\sqrt{3}}{6}, \quad x_r = \frac{1}{2} + \frac{\sqrt{3}}{6}. $$
This scheme was chosen because it provides a high degree of accuracy through a minimal amount of computation and has been shown to work well within the context of Continuous Time Galerkin methods (see [75, 76, 81]).

**Sparse Matrix Algebra**

A further detail important to the numerical implementation of the proposed computational model is the method for storage and solution of the linear systems resulting from the variable viscosity Newton methods. These linear linear systems are dealt with using the SPARSKIT v.2 library by Youcef Saad from the University of Minnesota. SPARSKIT is a Fortran package for working with sparse matrices in various sparse storage formats. The internal storage format for the library, and the format used here, is the *compressed sparse row* storage format.

The two main routines of this library used within the present implementation were the ILUT preconditioner, and the GMRES iterative solver. The ILUT preconditioning routine performs an incomplete L-U factorization of the original nonsymmetric Newton matrix, using a robust, dual thresholding strategy for pruning elements of the matrix factors. This preconditioner is then provided to the GMRES iterative solution routine, which uses a modified Gram-Schmidt process to implement a Krylov iterative solver for non-Hermitian matrix systems. The use of these two methods proved to provide a robust and fast solver for the linear systems in the viscous Newton method. On system matrices of size 4500x4500, it required approximately three iterations to reduce the linear system residual below tolerance (2.2d-16 for absolute residual, 1d-7 for relative residual). For further details on these methods, see the texts by Golub and Van Loan [65] or Greenbaum [67].

**Unit Scaling**
The final implementation detail we discuss here involves the physical units used in the model solution. The choice of computational units for the physical problem affects the conditioning of the resulting linear Newton systems. Using the typical units \{kg, m, s, K\}, the Newton matrix typically has very a high condition number, on the order of 1d+12. We thus adjust the computational units to obtain a better-conditioned matrix. Under the new physical units, we desire the block elements in the matrix to have approximately the same magnitude. This is possible if the elements of the state vector $X^n_{\Delta x} = [u^n, v^n, \theta^n]^T$ that they multiply also have approximately the same magnitude. Using the standard units in a few test experiments, typical deformations are on the order of 1d-4 m, velocities are on the order of 0.1 m/s, and temperatures are around 335 K. This results in some blocks of the Newton matrix having magnitude 1d+6 times the magnitude of other blocks.

The scaled units in the current implementation are changed to use mm as the spatial unit, ms as the temporal unit, and g as the mass unit. This results in the deformation typically having units on the order of 0.1 mm, velocity on the order of 0.1 ms, and temperature again around 335 K. Although this still results in significant discrepancies in the unit size, the test experiments were done under relatively stable conditions. However in more extreme circumstances, the bar would move farther and faster, leaving room for the deformation and velocity to grow without detriment to the condition of the matrix. With this choice of units, the relevant material constants have the following units:

**Elastic modulus units:** $[\text{GPa}] = 10^9 \ [\text{kg} \ \text{m}^{-1} \ \text{s}^{-2}] = 10^3 \ [\text{g} \ \text{mm}^{-1} \ \text{ms}^{-2}]$;

**Density units:** $[\text{kg} \ \text{m}^{-3}] = 10^{-6} \ [\text{g} \ \text{mm}^{-3}]$;

**Heat Conductivity units:** $[\text{W} \ \text{m}^{-1} \ \text{K}^{-1}] = [\text{kg} \ \text{m} \ \text{s}^{-3} \ \text{K}^{-1}]$

$= 10^{-3} \ [\text{g} \ \text{mm} \ \text{K}^{-1} \ \text{ms}^{-3}]$;

**Specific Heat units:** $[\text{J} \ \text{kg}^{-1} \ \text{K}^{-1}] = [\text{m}^2 \ \text{s}^{-2} \ \text{K}^{-1}] = [\text{mm}^2 \ \text{ms}^{-2} \ \text{K}^{-1}]$. 
The results in Chapter 6 are given in the base units of \{g, \text{mm}, \text{ms}, \text{K}\}, except for the results detailing the overall energy in the system (potential, kinetic or total), which have been converted back into Joules.
Chapter 6

Computational Results

The computational results in the following sections are designed to serve two purposes. The first is to provide numerical evidence that the proposed model is indeed able to reproduce the range of nonlinear behaviors unique to shape memory alloys. Thus in the following sections we show the thermally-induced martensite $\rightarrow$ austenite and austenite $\rightarrow$ martensite transformations, and the stress-induced (superalastic) phase transformation; we also provide evidence that these phase changes give rise to hysteretic and latent heat effects. The second goal of this chapter is to investigate various means of vibration damping using shape memory alloy wires. To this end, we present results on how the full phase transformation and spatially-localized phase transformations affect vibrational energy in a shape memory alloy wire. We conclude the computational results by discussing a highly effective temperature control for these damping processes.

In all these numerical experiments, we consider the NiTi shape memory alloy (though given the correct set of material constants, other alloys could be tested as well). The material constants used throughout the numerical experiments are given in Table 6. However, before presenting results from the numerical simulations, we discuss some
<table>
<thead>
<tr>
<th>Constant</th>
<th>Value</th>
<th>Units</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\rho_0$</td>
<td>$6.45 \times 10^3$</td>
<td>kg m$^{-3}$</td>
<td>density</td>
</tr>
<tr>
<td>$\kappa$</td>
<td>10</td>
<td>W m$^{-1}$ K$^{-1}$</td>
<td>thermal conductivity</td>
</tr>
<tr>
<td>$c_p$</td>
<td>322</td>
<td>J kg$^{-1}$ K$^{-1}$</td>
<td>specific heat capacity</td>
</tr>
<tr>
<td>$E_a$</td>
<td>$7.5 \times 10^{10}$</td>
<td>Pa</td>
<td>austenite elastic modulus</td>
</tr>
<tr>
<td>$E_m$</td>
<td>$2.8 \times 10^{10}$</td>
<td>Pa</td>
<td>martensite elastic modulus</td>
</tr>
<tr>
<td>$1 \pm \gamma_0$</td>
<td>1 $\pm$ 0.02</td>
<td></td>
<td>stable martensitic strains</td>
</tr>
<tr>
<td>$L$</td>
<td>0.05</td>
<td>m</td>
<td>length of wire</td>
</tr>
<tr>
<td>$\beta$</td>
<td>$2 \times 10^{-8}$</td>
<td>m$^2$</td>
<td>cross-sectional area of wire</td>
</tr>
<tr>
<td>$\theta_A$</td>
<td>350</td>
<td>K</td>
<td>austenite finish temperature</td>
</tr>
<tr>
<td>$\theta_C$</td>
<td>335</td>
<td>K</td>
<td>critical transition temperature</td>
</tr>
<tr>
<td>$\theta_M$</td>
<td>320</td>
<td>K</td>
<td>martensite finish temperature</td>
</tr>
</tbody>
</table>

Table 6.1: NiTi material constants used within the calculations (from [24, 45, 78]).

of the visualization techniques used to represent output from the one-dimensional thermodynamic model.

### 6.1 Visualization Techniques

In order for the computational results in the remainder of the chapter to be easily interpreted, we discuss some of the methods used to visually represent the numerical results. The greatest difficulty in showing these computational results lies in finding meaningful representations of the deformation and material phases through the shape memory alloy wire. One-dimensional deformations constitute elongation ($u(x) > x$) and contraction ($u(x) < x$). But since length changes are difficult to observe and do not provide information as to local changes in the wire, we plot the displacement, instead of the deformation. Thus, plots of the physical shape changes in the wire will show, on the vertical axis, local distances away from the austenitic reference configuration. Therefore, positive values of displacement correspond to elongation and negative values to contraction, as seen in Figure 6.1. With this visualization
technique, austenitic regions seem "flatter" in these plots than martensitic regions, which have steeper displacement gradients. Furthermore, as a result of these steeper displacement gradients, twinned martensite appears highly oscillatory, as seen on the right side of the example displacement plot in Figure 6.1.

![Image](image)

Figure 6.1: Displacement and Phase Visualization Techniques: 1-D deformation constitutes elongation and contraction. Thus displacement (left plot) is plotted for clarity, with positive values corresponding to elongation and negative values corresponding to contraction. Phase plots (right side) use yellow as the austenite phase, and use red and blue for the martensite variants. Note the austenitic regions in the left side of both plots, and the martensitic regions in the right side of each plot.

We must also plot meaningful results for the various material phases at different locations along the wire. Since the material phases at equilibrium correspond to \( \gamma(x) = 1 \) for austenite and \( \gamma(x) = 1 \pm \gamma_0 \) for martensite, and also since these dynamic experiments allow for deviations away from equilibrium, we may simply choose between the phases in the following manner: at each point \( x \) in the wire,

\[
\begin{align*}
\gamma(x) \geq 1 + \gamma_0 / 2 & \Rightarrow \text{martensite}^+ \\
\gamma(x) \leq 1 - \gamma_0 / 2 & \Rightarrow \text{martensite}^- \\
|\gamma(x) - 1| < \gamma_0 / 2 & \Rightarrow \text{austenite}.
\end{align*}
\]

We then use color plots to show these phases, along with the corresponding length change in the wire, as shown in Figure 6.1, where yellow regions correspond to austenite, red regions to martensite\(^+\) and blue regions to martensite\(^-\).

Other plots that are used to represent the numerical results are the deformation gradient, the kinetic energy, the temperature and the average enthalpy. Some of these
Figure 6.2: Deformation Gradient and Kinetic Energy Visualization Techniques: Deformation gradients (left) typically cluster around the stable states $\gamma \in \{1, 1 + \gamma_0, 1 - \gamma_0\}$. Note the non-smooth jumps between these states. Kinetic energy plots (right) oscillate rapidly as a result of the exchange between kinetic and potential energy. The overall vibrational energy may be tracked by following the peaks.

are discussed along with the material property they are used to show, so we discuss only the deformation gradient and kinetic energy plots here. As may be expected by the non-convex, multi-well structure of the free energy, plots of the deformation gradient should contain values in regions near the three possible stable configurations. As seen in the example in Figure 6.2, deformation gradients along the length of the wire cluster near the martensite$^+$ and the austenite values of $1 + \gamma_0$ and 1, respectively, with a non-smooth jump between the two (note: these distinct values are connected in the plot with near-vertical lines, giving the illusion of thickness in some transforming regions).

The last type of plot that may need explanation is the kinetic energy history for a given experiment. An example plot showing the time-evolution of the kinetic energy, as calculated in Section 5.1 via equation (5.3), may be seen in Figure 6.2. Here, the highly oscillatory behavior in time is a result of the exchange between potential and kinetic energy in the wire. To obtain an idea of the time-behavior of the overall
vibrational energy in the wire, it suffices to follow the peaks of this curve in time.

With these visualization tools at our disposal, we now demonstrate the ability of this model to describe the nonlinear behaviors of shape memory alloys as laid out in Chapter 2. We begin with the thermally-induced phase transformations.

6.2 Thermally-Induced Phase Transformations

The first computational experiments that we present show the thermally-induced phase transformation in shape memory alloy wires. Since the thermally-induced first-order phase transformation, Section 2.2, is the source of many of the key nonlinear characteristics in these materials, we investigate both the martensite $\rightarrow$ austenite transformation, and the reverse austenite $\rightarrow$ martensite transformation. Both of these numerical experiments were performed under constant heating (or cooling, respectively) of the wire through the external heat supply term $r$.

6.2.1 Martensite to Austenite

We begin with the low-temperature (martensite) to high-temperature (austenite) transformation. The initial deformation begins as twinned martensite, with initial velocity set to zero and initial temperature set to the transition temperature $\theta_C$. From the beginning of the simulation, the heat supply term $r$ is held constant at $r(x, t) = 6$ kJ g$^{-1}$ ms$^{-1}$ in order to raise the temperature of the wire into the austenitic temperature range. We present time level states of the temperature distribution, displacement and phase plots at the time values $t = 0.1$ ms, 0.9764 ms, 1.1425 ms and 1.7388 ms.

The temperature plots are given in Figure 6.3. As shown in the plots, the temperature increases from the transition temperature $\theta_C = 335$ K up to near the austenite finish
Figure 6.3: Temperature Plots for the Thermally-Induced Martensite → Austenite Phase Transformation: The temperature increases due to the constant, positive external heat supply. Note the local, irregular effects of the latent heat, and the eventual smoothing of the temperature in time.

The temperature $\theta_A = 350$ K. At this point, however, distributed regions in the wire begin transforming to the austenite phase. As this transformation occurs, the temperature of the transforming regions quickly drops due to the latent heat effects. Moreover, plots of the displacement in Figure 6.4 show that the transformation occurs only in (distributionally) about half of the wire at first. Then because of the thermal drop, it takes more heat to finish the transformation to austenite throughout the remainder of the wire. Finally after the transformation is complete the heat again smoothes out through the length of the wire.
Figure 6.4: Displacement Plots for the Thermally-Induced Martensite → Austenite Phase Transformation: The wire begins as fully-twinned, relaxed martensite. The transformation begins in the left end of the wire, and proceeds in two stages. The third plot demonstrates the intermediate stage, containing mixtures of the austenite and martensite$^\pm$ phases. Note the resulting vibrations in the final austenite state, as evidenced by the non-zero displacement.

We may similarly consider the displacement plots for the thermally-induced martensite to austenite transformation in Figure 6.4. As seen in the figure, at the beginning of the simulation the wire begins in the relaxed, twinned-martensite state (hence the "thickness" of the wire – with alternating regions having gradients $\gamma = 1 + \gamma_0$ and $\gamma = 1 - \gamma_0$). As the temperature is increased, we see at $t = 0.9754$ ms the effects of the instability of the martensitic variant at this temperature, with the phase transfor-
Figure 6.5: Phase Plots for the Thermally-Induced Martensite → Austenite Phase Transformation: As the temperature increases, various regions transform from martensite\(^+\) (red) and martensite\(^-\) (blue) to austenite (yellow). Phase mixtures result in seemingly orange, green and purple hues.

Transformation beginning at the left end. As the heat increases, the first half of the transformation finishes, leaving alternating regions of austenite and some martensitic variant at the time \(t = 1.1425\) ms. The full transformation is complete by \(t = 1.7388\) ms, at which time the entire wire is near to the stable austenitic configuration. However, even though the simulation began at rest with no kinetic energy, the sudden phase transformation converted some of the potential energy to motion, causing vibrations in the wire.
The final set of plots for this transformation gives the phase distribution in the wire in time, as seen in Figure 6.5. Since the martensitic variants are given by red (martensite$^+$) and blue (martensite$^-$), the initial state of twinned martensite appears violet. As the time (and thus temperature) increases, the phase transformation begins throughout the wire. As mentioned earlier in the discussion of the temperature and displacement, the thermally-induced phase transformation occurs in two stages. This explains why in the second and third time frames, the martensitic reds and blues are intermixed with the austenitic yellow, producing regions in the wire that appear orange or green with this visualization technique. However, once the phase transformation is complete, the entire wire remains in the austenite state.

### 6.2.2 Austenite to Martensite

The second computational experiment that we show here corresponds to the thermally-induced austenite to martensite transformation. In this simulation, the wire begins with deformation as relaxed austenite, velocity set to zero, and with initial temperature at the transition temperature $\theta_C$. In this simulation, the heat supply term $r$ is again constant, but this time cools the wire at a rate of $r(x, t) = -6 \text{ kJ g}^{-1} \text{ ms}^{-1}$. We again present time level states of the temperature, displacement and phase plots in the wire, in this instance at the times $t = 0.1 \text{ ms}$, $0.9124 \text{ ms}$, $0.9204 \text{ ms}$, and $0.7288 \text{ ms}$. The corresponding temperature plots are shown in Figure 6.6, the displacement plots in Figure 6.7, and the phase plots in Figure 6.8. As in the case of the martensite $\rightarrow$ austenite transformation, the reverse transformation occurs in two bursts.

First noting the temperature plots in Figure 6.6, again we see that as the temperature changes nearly half of the wire transforms at once, with the transformations spread throughout the length. In this case, however, the latent heat of transformation results in a sharp rise in the temperature. Moreover, when the transformation is complete
Figure 6.6: Temperature Plots for the Thermally-Induced Austenite → Martensite Phase Transformation: The temperature is decreased using a negative, constant heat supply \( r(x, t) \). Note the non-smooth latent heat effects during as each lattice element changes independently, as well as the eventual smoothing of the temperature in time.

The key item of note in the plots of the displacement, Figure 6.7, is that due to the multiplicity of equally-stable martensitic variants, the resulting fully-transformed martensitic state of the wire is not uniformly twinned throughout the length of the wire. These deviations from the perfectly twinned configuration are realistic of nature, where the sudden formation of martensite upon cooling results in mostly twinned structures, along with possibly random variations from the fully-twinned state. However, as seen by the zero displacement at the rightmost end of the wire, the macro-
Figure 6.7: Displacement Plots for the Thermally-Induced Austenite → Martensite Phase Transformation: As the temperature decreases, the wire transforms, beginning at the left end. Note the final martensitic state is nearly completely twinned, with relatively few areas fully martensite$^+$ or martensite$^-$. 

Scopnic shape (length) of the wire remains the same as the initial austenitic shape, albeit with a substantially different microstructure.

Finally, the phase plots in Figure 6.8 provide the same information as seen before in the transformation from martensite to austenite. We clearly see the distributed nature of the sudden, thermally-induced phase transformation, in which regions appear to consist of intermixed variants of both austenite and martensite.
Figure 6.8: Phase Plots for the Thermally-Induced Austenite $\rightarrow$ Martensite Phase Transformation: As the temperature is decreased, the austenite transforms in a highly distributed fashion into the different martensite variants.

### 6.3 Stress-Induced Phase Transformation

As discussed in Chapters 2 and 3, a key reason for the consideration of nonlinear models (and the solution difficulties involved) as opposed to the considerably simpler linearized models, is the ability of nonlinear models to handle both the thermally-induced and stress-induced phase transformations [114]. We thus consider the following computational experiment testing the ability of the model to handle the stress-induced (superelastic) transformation, as introduced in Section 2.2.4. To summarize
Figure 6.9: Body Forces Used for the Stress-Induced Phase Transformations: The wire is first pulled under tensile force until stress-induced martensite$^+$ forms at around $t = 0.4$ ms. From $t = 0.5$ ms, the tensile force is slowly released and replaced with compressive force until the stress-induced transformation goes through austenite and finally to martensite$^-$ ($t \in (0.5, 1.5)$ ms). The compressive forces are then relaxed to the end of the simulation at $t = 2$ ms.

In this process, the wire may transform between states independently of temperature. This transformation occurs when the temperature of the material is close to the high austenite finish temperature $\theta_A$. Then, if the external stresses are large enough, they force the crystal structure of a shape memory alloy to transform to the most accommodating martensitic state (i.e. most strain, least stress), although those states are not as energetically favorable as the austenite phase at that temperature.

In this experiment, we begin with the wire in a relaxed, fully austenitic configuration, with zero initial velocity, and with initial temperature $\theta_0 = 345$ (above the transformation temperature $\theta_C = 335$, yet still below the austenite finish temperature
\( \theta_A = 350 \). The experiment runs for a simulation time of 2 ms. In the first 0.5 ms, body forces (incorporated through the term \( b(x, t) \)) linearly increase on the rightmost end of the wire from 0 to \( 1 \times 10^5 \) N kg\(^{-1} \), which provide enough tensile force to induce the transformation to martensite\(^+\). Then, at 0.5 ms, the body forces decrease linearly from \( 1 \times 10^5 \) to \( -1 \times 10^5 \) N kg\(^{-1} \) for 1 ms, in order to first relax the wire and then compress it back through the austenite state and into the martensite\(^-\) state. For the final 0.5 ms of the simulation, the body forces then linearly return from \( -1 \times 10^5 \) N kg\(^{-1} \) to relax the wire again at 0 N kg\(^{-1} \).

The first plot that we show for this experiment, in Figure 6.9, gives the body forces used to induce these transformations. As the body force first pulls on the wire, it begins to pull with increasing tension until the sudden transformation to martensite\(^+\) near \( t = 0.4 \) ms. Then, these tensile forces are decreased and eventually replaced with a compressive body force until \( t = 1.5 \) ms. These compressive forces induce the sudden transformation to the austenite and then the martensite\(^-\) states. The simulation end with the relaxation of these compressive forces.

These effects may be more readily seen in the time level states of the deformation gradients (Figure 6.10) and phases (Figure 6.11). We consider these states at the time values \( t = 0.1 \) ms, \( 0.3261 \) ms, \( 0.3861 \) ms, \( 1.2686 \) ms, \( 1.3346 \) ms and \( 1.3866 \) ms. Examining the deformation gradients in time, we see the effects of increasing tensile load in the first three plots, showing the sudden jumps in various parts of the wire from the region around the austenite phase \( \gamma = 1 \) to the region surrounding the martensite\(^+\) phase \( \gamma = 1 + \gamma_0 \). As with the temperature-induced transformation, small regions of the wire transform seemingly independently from each other, forming regions of mixed phase in the wire, until the load has increased enough to pull the entire wire into the martensite\(^+\) phase. The final three plots show the subsequent compression of the wire, in which lattice elements in the wire first begin to transform back to the austenite state \( (t = 1.2686 \) ms), and then due to the large compressive
Deformation Gradients

Figure 6.10: Deformation Gradient Plots for the Stress-Induced Phase Transformations: Note the distinct clustering near, and the sudden transformations between, the various stable states at \( \gamma \in \{1, 1 + \gamma_0, 1 - \gamma_0\} \).

force they continue the transformation into the martensite\(^-\) state.

These stress-induced phase changes are perhaps made more evident by the phase plots in Figure 6.11, where we may not only understand the local phase changes through the colors, but we may also view the macroscopic elongation and contraction in the
Figure 6.11: Phase Plots for the Stress-Induced Phase Transformations: The transformation from austenite (yellow) to martensite$^+$ (red), back to austenite, and finally to martensite$^-$ (blue). Note the observable changes in length (deformation) at the free (right) end corresponding to each state.

wire throughout the experiment. As in the deformation gradient plots in Figure 6.10, we see the austenite $\rightarrow$ martensite$^+$ transformation as the tensile forces increase.
We then see the martensite\textsuperscript{+} $\rightarrow$ austenite $\rightarrow$ martensite\textsuperscript{−} transformation as the wire comes under compressive forces. Moreover, through tracking the location of the right endpoint of the wire in time, we see the initial extension from 50 mm up to 55 mm, followed by the contraction down to 45 mm in the final plot.

### 6.4 Hysteresis

In addition to the thermally- and stress-induced phase transformations, a proper non-linear model of shape memory alloys must also account for the hysteresis associated with such phase transformations, as shown in Figure 2.1 of Chapter 2. Upon collecting simulation data from repeated thermally-induced phase transformations, we arrive at the following phase vs. temperature plot exhibiting hysteretic effects, as seen in Figure 6.12. Points on this plot represent the average state of the simulated wire at single levels within these experiments. We used data from 8 experiments, in which the heat supplies varied from $\pm 4$ to $\pm 10$ kJ g\textsuperscript{−1} ms\textsuperscript{−1}. Each of these simulations calculated 2,000 time steps, resulting in 16,000 data points on the plot, each representing the percentage of austenite in the wire as a function of the average temperature. These quantities are calculated from the state variables $u(x, t)$ and $\theta(x, t)$ at each time level $t$ using the formulas

\[
\theta_{\text{avg}}(t) = \frac{1}{L} \int_0^L \theta(x, t) \, dx
\]

\[
(\text{aust } \%) (t) = \frac{1}{L} \left| \left\{ x \in [0, L] : |u'(x, t) - 1| < \frac{\gamma_0}{2} \right\} \right|
\]

The hysteresis loop is indeed clearly reproduced. The slight negative tilt in the slope of the hysteresis plot may be understood in the following manner. First, due to the latent heats of transformation (to be further examined in Section 6.5), in which the martensite to austenite transformation is endothermic and the austenite to martensite transformation is exothermic, transformed regions of the shape memory alloy wire either lose or gain heat depending on the direction of the transformation.
Figure 6.12: Computed hysteresis loop. The plot is assembled using a number of computer experiments. The average state of the wire at any particular time is represented by one point. Note that the martensite to austenite transformation occurs at a considerably higher temperature than the reverse transformation.

Since these latent heats of transformation act opposite the heating or cooling causing the transformation, the cause the slight negative tilt. A second explanation is the fact that, due to the use of a single Helmholtz free energy, the proposed model describes an ideal wire composed of uniform shape memory crystals. However, in laboratory specimens there always exist imperfections (dislocations, grain boundaries, etc.) in the crystal lattice. These imperfections cause some parts of shape memory alloy bars used in laboratory experiments to transform at lower or higher temperatures than the pure crystal does. Thus traditional hysteresis curves for the thermally-induced transformation move more smoothly between phases, while the current model begins the transformation all at once (note the sharp corners). Thus, the differences between the computed hysteresis curve and "ideal" hysteresis curves may be easily understood,
Figure 6.13: Enthalpy Plots Showing Evidence of Latent Heat Effects: Jumps in the entropy correspond to latent heats of transformation. The thermally-induced martensite to austenite transformation on the left. The reverse transformation from austenite to martensite on the right.

though the important point remains: the model correctly predicts hysteretic effects inherent to the nonlinear, first-order phase transformation.

6.5 Latent Heat

As has been evidenced in the computational experiments of the thermally-induced phase transformation in Section 6.2, as well as in the computations providing the hysteresis loop in Section 6.4, the proposed model accounts for the effects of the latent heats of transformation. Moreover, these effects do not need to be supplied as additional terms in the model, as must be done in the linearized models discussed in Section 2.4. These effects are implicitly accounted for because the current model retains the nonlinearity of the Helmholtz free energy, thereby allowing for such energy exchanges upon transition from a metastable to a stable energy well.

The latent heat may be more directly observed through changes in the overall en-
thalpy of the wire in time. According to standard theory regarding first-order phase transformations in other media [31, 82, 134], the latent heat released on a time interval \((t_0, t_1)\) may be given by the change in enthalpy, \(\Delta H = H(t_1) - H(t_0)\), where the overall enthalpy in the wire at the time \(t\) is calculated from equation (5.4) given in Chapter 5. Figure 6.13 shows the time histories of the enthalpy from the two thermally-induced phase transformations discussed in Section 6.2. The plot on the left corresponds to the overall enthalpy resulting from the martensite to austenite transformation, and the plot on the right corresponds to the reverse transformation. The items to note in these plots are the sudden enthalpy jumps, corresponding to the latent heats of transformation. Moreover, the measured values of latent heat from the materials science literature provide the value \(k_l = 1.56 \times 10^8\) kg m\(^{-1}\) s\(^{-2}\), corresponding to the latent heat coefficient per unit volume for NiTi. Converting this number into units corresponding to the total change in enthalpy, we multiply by the overall volume of the wire. According to Table 6, this becomes

\[
\left(1.56 \times 10^8 \frac{\text{kg}}{\text{m s}}\right) (L \beta \text{m}^3) = \left(1.56 \times 10^8 \frac{\text{kg}}{\text{m s}}\right) (1 \times 10^{-9} \text{m}^3) = 0.156 \frac{\text{kg m}^2}{\text{s}^2}
\]

Thus, we see that the predicted latent heats of transformation from Figure 6.13 are within a factor of two of the measured values.

### 6.6 Damping Strategy

The remaining examples examine the applicability of the NiTi shape memory alloy for vibration damping purposes. Each of these examples begins with the wire at rest. Early into the simulations, a vibrational wave of the form

\[
b((L - \Delta x, L), t) = \text{mag} \times \cos \left(\frac{t - \text{peak}}{\text{dur}}\right) \times \exp \left(-\left(\frac{t - \text{peak}}{\text{dur}}\right)^2\right)
\]

strikes the free end of the wire, where \(\text{mag}\) provides the strength of the wave, \(\text{peak}\) provides the time of the wave peak, and \(\text{dur}\) provides the overall amount of time that
Figure 6.14: Vibrational Force Used in Damping Tests: the tensile/compressive body force $b(t)$ exerted on the rightmost end of the wire, and measured in units kN kg$^{-1}$. The axes are swapped to indicate the direction of force. Note the truncated time scale, since the full simulation lasts 2 ms.

the vibration is active. The values used of these parameters used in the following experiments were mag $= 2 \times 10^4$ N kg$^{-1}$, dur $= 0.04$ ms and peak $= 0.04$ ms. A picture of the resulting vibrational wave is given in Figure 6.14. This form of vibrational energy is used since it allows for a short but strong burst of vibrational energy to enter the wire, and from that moment we may analyze the damping behavior in time. We implement the vibrational forces as a body force, acting only within the rightmost end of the wire. This allows us to include external forces without need of surface forces on the boundary.

Each of the following four examples begin with the wire in the fully-twinned martensitic state at the transition temperature $\theta_C$, and the same vibrational wave affects each of them at the beginning of the simulation. After the initial wave hits, the examples differ in only the heating and cooling schemes used to combat the vibrational waves.
The first of these examples in Section 6.6.1 provides the base case in the absence of any temperature control. This base case will serve as the point of comparison for the subsequent examples. The second example, in Section 6.6.2, shows the effects of a full phase transformation on the vibrational energy in the wire. The third example, discussed in Section 6.6.3 shows how highly localized phase transformations perform for damping applications. The final example, in Section 6.6.4, shows a promising means for using localized phase transformations to damp vibrational energy using shape memory alloy wires.

### 6.6.1 Base Case

This example is used as a base of comparison for the various heating schemes. Therefore no heating control is applied to the wire, and the results provide the comparative case of a purely vibrating martensitic wire.

![Kinetic Energy History](image)

**Figure 6.15:** Kinetic Energy Plot for an Uncontrolled Vibrating Wire: Note the very slow, possibly indefinite damping.
Figure 6.16: Displacement Plots for the Uncontrolled Vibrating Wire: The martensitic wire vibrates under wave motion. Temperatures and stresses are too small to induce transformation.

The Figure 6.15 shows the kinetic energy history for the uncontrolled case. As expected, the vibrational energy very slowly naturally damps out of the system. The apparent random nature of the vibrational waves is likely due to the frequencies of the input wave being out of synchrony with the natural frequencies in the wire. It is expected that as time goes on, the higher frequency vibrations will slowly cancel out, while the lower frequencies will remain, steadily vibrating the shape memory wire in time.

The temperature and phase plots of this case are of little interest, since the temperature remains within 0.01 K of the initial temperature $\theta_C$ and neither the temperatures nor stresses are large enough to create the austenitic phase in the wire. Thus we limit
the snapshot plots for this example to those of the displacements in Figure 6.16, taken at the equidistant time levels \( t = 0.1 \text{ ms}, 0.7 \text{ ms}, 1.3 \text{ ms} \) and \( 1.9 \text{ ms} \). As expected, the plots show the motion of the vibrational waves through the martensitic wire in time.

Certainly, this case is of little interest in its own right. However, the remaining examples, when compared with both the kinetic energy plot and the displacement plots here, provide both unexpected and interesting results.

6.6.2 Full Phase Transformation

In this first example of thermal control for the damping of vibrations in shape memory alloy wires, we examine the effect of transforming the entire wire from martensite to austenite. To this end, we uniformly heat the wire throughout the simulation using the control \( r(x, t) = 6 \text{ kJ} \text{ g}^{-1} \text{ ms}^{-1} \). As discussed before, while investigating the thermally-induced phase transformation in Section 6.2, the phase transformation occurs in two stages. In this example, the first of these occurs at \( t = 0.8 \text{ ms} \). Then, due to the effects of latent heat, the second stage of the phase transformation occurs at \( t = 1.1 \text{ ms} \). The behavior of the kinetic energy just before, during, and after these partial phase transformations proves to be quite interesting. As seen in the time history of the kinetic energy in Figure 6.17, as the temperature increases the kinetic energy slowly decreases. Then, at the onset of the first stage in the phase transformation, the kinetic energy drops suddenly. When this first stage of the phase transformation is complete, there is a brief spike in the kinetic energy before the second stage of transformation begins. However, when this second stage is complete, the overall kinetic energy in the wire is significantly larger than when the simulation began.

We believe that all of these processes may be explained simultaneously as conver-
Figure 6.17: Kinetic Energy Plot for the Full Phase Transformation: The kinetic energy decreases as the wire is heated. At the moment of phase transformation kinetic energy drops significantly, followed by energy spikes. Note the final highly regular, larger observable vibrational energy.

sion between various kinds of energy. After the initial wave hits the wire, vibrations oscillate rapidly in a seemingly random fashion, with significant amounts in both the kinetic and potential energy states. Then, as the temperature is increased, the energetic metastability of the martensitic phase increases. This increase in metastability corresponds with a conversion of a significant amount of the kinetic energy to potential energy in the wire. As the phase transition itself occurs, the potential energy reaches its peak, causing a brief calm in the wire before the first transformation suddenly "pops" a significant amount of the wire into the austenitic state, increasing the kinetic energy dramatically. However, since the second stage of the phase transition is beginning, this kinetic energy is once again converted into potential energy. Finally, when the full phase transformation is complete, the potential energy is released. When this occurs, the resulting vibrations resonate at one of the wire's natural frequencies, causing the larger peaks and more regular structure of the kinetic energy.
Figure 6.18: Temperature Plots for the Full Phase Transformation: The temperature-induced phase transformation proceeds as normal, though begins at the leftmost end of the wire.

Plots of the temperature distribution, displacements and phases at the time levels $t = 0.25 \text{ ms}$, $0.9124 \text{ ms}$, $1.0985 \text{ ms}$ and $1.8869 \text{ ms}$ are shown in Figures 6.18, 6.19 and 6.20, respectively. As seen in these plots, the thermally-induced phase transformation occurs similarly as before, occurring first through about half of the wire and then continuing through the rest in the second stage of the transformation.

It is notable, though, that unlike the static case, the transformation begins in the vibrating wire at the fixed (left) end of the wire. We believe that this is due to the increased amount of stress experienced when waves reflect off of this end. When these
Figure 6.19: Displacement Plots for the Full Phase Transformation: the vibrating wire undergoes the two-stage phase transformation. Note that the transformation begins at the fixed end of the wire, due to increased stress at that end.

Reflections occur, the stress at the fixed end becomes maximal before sending the wave energy back through the wire. Thus at the moment of reflection, the increased stress level provides an initial starting point for the phase transformation to begin. Then as a result of the increase in the overall temperature distribution, the remainder of the wire transforms soon after.
Figure 6.20: Phase Plots for the Full Transformation: The phase transformation occurs throughout the entire length of the wire in two stages, as discussed earlier in Section 6.2.

6.6.3 Localized Phase Transformations

Since the full phase transformation resulted in an increase in the observable kinetic energy of the wire, a natural reaction is to test the effect of localized phase transformations as a control mechanism. In this example, we examine the effect of transforming highly localized areas of the wire from martensite to austenite. To this end, we unevenly heat the wire using the control

\[ r(x, t) = 10 \sin (2\pi \times (x + t/2)) \text{ kJ g}^{-1} \text{ ms}^{-1}. \]
Figure 6.21: Kinetic Energy Plot for the Highly-Localized Phase Transformation: The highly-localized phase transformation results in cancellation of the induced vibrational waves following an initial spike in the kinetic energy.

This heating control allows for both a localization of the heat control, as well as a motion of the heat in time. Due to the form of this control, there are a large number of full sine waves periodically heating and cooling the wire. Furthermore, due to the relatively large factor of 2 dividing the time (the total time $T = 2$ ms), the heating control moves slowly along the length of the wire as the simulation proceeds, allowing the temperature of the highly-localized regions to reach the required transformation temperatures. Lastly, although this three-parameter (amplitude, localization and speed) form of the heat supply amounts to a very simple form of the thermal control, it serves as an illustrative example for understanding how localization of the heat control benefits the vibrational damping.

Figure 6.21 provides the kinetic energy history of this experiment. In this plot we see that, as in the example of the full phase transformation, the kinetic energy again
Figure 6.22: Temperature Plots for the Highly-Localized Phase Transformation: The thermal control results in oscillatory temperatures through the wire. Latent heat effects decrease the peaks in transforming regions.

decreases slowly as the temperature is increased through local regions in the wire. Moreover, when the phase transformation begins in the wire at the time $t = 0.5$ ms, the kinetic energy drops suddenly. When this distributed phase transformation is complete, the kinetic energy again increases; however, the subsequent transformations cancel out most of the kinetic energy from the system.

These results may be analyzed in relation to the locality of the phase transformations. As seen in the previous example, the sudden phase transformation resulted in an abrupt burst of kinetic energy that then resonated at the characteristic speed of
Figure 6.23: Displacement Plots for the Highly-Localized Phase Transformation: A large number of very small regions transform states; however, these small regions significantly affect displacements along the wire.

the austenitic phase in the wire. Similarly, the localized phase transformations of the current example induce bursts of vibrational energy. However, in this case the distributed nature of the phase transformation regions leaves regularly-spaced regions in which these vibrations travel. Since the induced vibrations travel at the same speeds through these untransformed regions, the self-cancellation of such anti-phase vibrations results in the overall damping observed in the experiment.

The effect of the slow-moving, highly-localized heat control \( r \) on the temperature distribution may be seen in Figure 6.22. As time progresses, the temperatures in these
Figure 6.24: Phase Plots for the Highly-Localized Transformation: The small, transformed regions along the wire are seen as the appearance of thin yellow lines among the predominantly martensitic wire.

small regions grow due to the slow motion of the control. In the second and third plots, the effects of latent heat from the localized transformation can be observed. Finally, we note that due to the movement of the heating control, the temperature waves begin shrinking again as the control moves away from its initial position.

As seen in Figure 6.23, the displacement plots also show the effect of localized phase transformations on the shape memory wire. Although these transformations only occur on a relatively small amount of the overall length, they have a significant effect
on the resulting displacements in the wire. Finally, Figure 6.24 shows the phase
distributions along the shape memory wire in time. Here, the regularly-spaced regions
of yellow in the lower plots show the locations of the localized phase transformations.

6.6.4 "Best Case" Damping Scheme

The best results on vibration damping that we have yet found result from a less
localized thermal control than in the last example. In this experiment, we again

![Kinetic Energy History](image)

Figure 6.25: Kinetic Energy Plot for the "Best Case" Temperature Control: Again,
the onset of the phase transformation damps the wire, though this time the larger
phase transformation regions do not result in a spike of kinetic energy.

use a localized, slow-moving temperature control having the same overall magnitude;
however, in this case the localization is more coarse than before. The temperature
control \( r \) is given in this example by

\[
r(x, t) = 10 \sin (2\pi \times (x/10 + t/2)) \quad \text{kJ g}^{-1} \text{ms}^{-1},
\]
Figure 6.26: Temperature Plots for the “Best Case” Temperature Control: the heating control results in smoother temperature distributions, with larger transformation regions. Note that the slow movement of the thermal control in time brings the temperatures back toward the transformation temperature $\theta_C = 335$ K.

which results in only five complete thermal waves, as opposed to the fifty in the previous example. The resulting kinetic energy history is given in Figure 6.25. The general behavior of the vibrational energy is nearly identical to that of the previous example, though the temperature spike after the initial transformation no longer appears. We believe this is a result of the larger transformation regions, along with the larger distance the induced vibrations travel before encountering each other and canceling.

The resulting time level plots of the temperature are given in Figure 6.26. As seen in
Figure 6.27: Displacement Plots for the "Best Case" Temperature Control: Here there are larger transformation regions, as well as longer areas of untransformed martensite. We believe these two factors play a key role in the additional damping ability resulting from this heating scheme.

These plots, when the peaks of the thermal control reach the austenitic transformation temperature those spatial control regions transform. These transformation again cause a release of latent heat, having a pronounced effect on the resulting temperature distribution. Furthermore, as this thermal control moves to the right, the temperature variations begin to diminish.

The resulting transformation regions can also be clearly seen in the displacement plots of Figure 6.27. As seen here, very few of the localized transformation regions
Figure 6.28: Phase Plots for the “Best Case” Temperature Control: The five transformation regions are clearly visible. Note the movement of the transformation regions, most notably visible between $t = 0.6823$ ms and $t = 1.7548$ ms.

exist as purely austenite; instead these regions have mostly undergone only one stage of the thermal transformation, leaving regularly-spaced mixtures of austenitic and martensitic variants. We believe that the regular spacing of these transformations is of vital importance to the large degree of damping that results, as evidenced by Figure 6.25.

Finally, the phase plots of Figure 6.28 show two important characteristics of these localized damping regions. First, as seen by tracking the end of the wire, the macro-
scopic length of the simulated wire remains virtually unchanged, even though the phase transformations within the wire have changed the local structure. Also, this figure clearly shows the slow movement of the transformation regions to the right (toward $x = L$), as seen from the difference between the timelevel states $t = 0.6823\,\text{ms}$ and $t = 1.7548\,\text{ms}$.

Thus, we see from the computational results of this chapter that, indeed, the full range nonlinear physical behaviors of shape memory alloys are predicted by our nonlinear, continuum thermodynamic model. Moreover, this model can then be used to simulate physical experiments for various types of shape memory alloy wires under a broad range of experimental conditions. In this chapter, our model has been used to examine the possible application of partially-active vibration damping; however, the model could just as easily be applied to other applications that take advantage of thermally and stress-induced phase transformations in shape memory alloy wires.
Chapter 7

Conclusions and Future Work

This model correctly reproduces the observable physics of shape memory alloy wires. Furthermore, through the generalization of the Landau-Devonshire potential in Section 3.4, the proposed model can be used to evaluate the thermodynamic response of a variety of different shape memory materials. Finally, using this model in the context of the NiTi shape memory alloy wire, we see that the temperature-induced phase transition may be used as an effective means for damping vibrational energy. Moreover, the results from the computational experiments in Chapter 6 suggest that one possibility for achieving appreciable dissipation of the kinetic energy relies on localized, periodically-spaced heating and cooling elements as a thermal control mechanism.

7.1 Accuracy of the Proposed Model

Due to our construction of the new thermodynamic model in Chapter 3, this model accurately describes the nonlinear physics of shape memory alloy wires. Through both the incorporation of the new inhomogeneous heat flux $q$ in Section 3.2.5, as well as the careful construction of the new Helmholtz free energy in Section 3.4, this model
more accurately describes the general nonlinear thermodynamic behavior of shape memory alloys than previous nonlinear models described in Section 2.5. Moreover, the proposed construction of the free energy allows for control over important material properties such as the specific heat, elastic moduli, and martensitic free energy minima. This is in contrast with previous constructions of the free energy, such as the Landau-Devonshire potential [53], the Falk potential [52], and even the potential constructed by Niezgódka and Sprekels [114], that provide insufficient control over such material constants as discussed in Chapter 2.

Through its detailed construction of the thermodynamic system, as well as the retention of its nonlinear structure, this proposed model accurately describes the full range of nonlinear behaviors characteristic to shape memory alloys. Also, unlike the linearized models discussed in Section 2.4, this model successfully reproduces both the thermally and the stress-induced phase transformations, as evidenced through the computational results in Chapter 6.

Thus the model presented here may provide key information predicting the material response expected from various shape memory alloy wires, and without the need for detailed and costly experimentation. This information could then be used as an input to some of the more traditional linearized models discussed in Section 2.4 that are currently under development in the engineering community. The nonlinearity of these models could then be better fit, aiding in the design of macroscale devices using shape memory alloy wires. Additionally, through simulation and computational experimentation using our proposed nonlinear model, small-scale devices may be designed and controlling methods may be investigated for taking advantage of the nonlinear behaviors unique to shape memory alloys.

However, this model is not without its limitations. One of these key limitations is the one-dimensional simplification of the wire. Due to this simplification, surface energy effects at the phase boundaries may not be fully accounted for without the addition of
extra kinetic relations and observation-based terms. Although it has been postulated that these surface energy effects are minimal in comparison with the phase transformation itself, such effects may still have a significant impact in certain thermo-mechanical loading situations. Moreover, a number of improvements may still be made to the Helmholtz free energy. As shown in Section 3.4, the free energy is constructed based on material measurements, which are made at near-equilibrium configurations (energy wells). However, other non-equilibrium free energy characteristics, such as the height of energy barriers, have key effects on the overall thermodynamic behavior of these materials. Therefore, further work on the understanding and the construction of the Helmholtz free energy is possible in order to even more closely match material behavior. Lastly, as remarked in the development of our inhomogeneous heat flux $q$, further improvement of this quantity may provide a better description of the one-dimensional heat transfer properties of shape memory alloy wires. We believe that an approach incorporating two distinct thermal conductivities $\kappa_A, \kappa_M$ for the austenite and martensite phases may succeed in this task; however, further physical experimentation is necessary to provide a sufficiently accurate characterization of this inhomogeneous heat flux.

### 7.2 Improved Solution Methods

Although the solution method described in Chapter 5 succeeds in producing highly accurate solutions to the nonlinear system of partial differential equations, the moments of phase transition still pose difficulties. It is at these phase transitions that the solution requires increased viscosity levels. The proposed solution method indeed reduces the levels of artificial viscosity required by traditional theory on such systems (see Section 4.2.3) by an order of magnitude; however, the variable viscosity method (Section 5.4) does not succeed in completely removing artificial viscosity from the
system.

It would therefore be beneficial to find a solution method that could handle the zero-viscosity case. Although “black-box” optimization methods for nonlinear, non-convex optimization may someday provide the answer, possible adjustments to the variable viscosity method may prove successful as well. Namely, since the proposed variable viscosity method uses an explicit-Euler approach to reduction of the viscosity, implicit schemes for reducing this viscosity parameter may produce improved solutions.

Another means for future research on finding solutions to the nonlinear, nonconvex system of partial differential equations may be through quasiconvexification (Section 4.2.1) of the potential function. This area of analysis is quite new, and advances there may provide just the tools necessary for solution of such ill-posed dynamical systems as those described here regarding shape memory alloy behavior. However, at the moment these methods are useful for branching spatial scales, essentially averaging out much of the small-scale nonlinear information important to the damping applications we consider in this thesis. Thus future methods for balancing such quasiconvexification approaches with retention of small-scale information may prove invaluable to such multi-scale problems as microstructural phase transitions.

7.3 Current SMA and Active Damping

The results from Section 6.6 indicate that regularly-spaced transformation regions prove to be an integral tool in control schemes using the thermally-induced phase transition in shape memory alloys for vibration damping. However, in the computational experiments done here, the heat control implemented was much more powerful than anything currently practical. These excessive heating controls were used primarily because of the typically slow thermal response times associated with shape memory
alloys, as evidenced by the large specific heat capacity $c_p$ and small coefficient of thermal conductivity $\kappa$. Thus the excessive heating controls were used in order to shorten the simulation time required to observe phase transformations; however, the results pertaining to the onset of the phase transformation, as well as the suggested localization of the transformation regions, remains valid for even considerably less powerful thermal controls. Moreover, we predict that a possible "partially-active" control scheme corresponding to small, perpetually-transforming, slowly moving, distributed transformation regions along the wire will remain applicable for shape memory alloy damping devices.

However, due to these slow response times associated with changes to the thermal field in shape memory alloys, truly active damping applications appear currently out of reach. Alternatively, other materials such as piezoelectrics and ferromagnetics, which undergo similar, though less-pronounced, phase transformations as a result of changes in the electric and magnetic fields, respectively, have significantly faster response times. This is because electrical and magnetic fields are more quickly changed than thermal fields. Therefore, the application of nonlinear models similar to the one proposed here, but applied toward those materials, may prove ideal for active vibration damping devices.

### 7.4 Expansion to Plates, Solids

Another area for related work lies in the expansion of the current model to higher dimensions. We believe that the general nonlinear thermodynamic model derived in Section 3.2, and the overall solution method proposed in Chapter 5 would work similarly for the higher-dimensional case; however, the construction of applicable two-dimensional and three-dimensional forms of the Helmholtz free energy would prove much more complex. A starting point for such models would be the analy-
sis of both the Falk-Konopka [54] and Ericksen-James [83, 84, 86] free energies. A further difficulty is that shape memory films and solids allow for the possibility of much more spatially complex temperature controls. However, it is our belief that a similarly-distributed "checkerboard" pattern of heating control for the thermally-activated damping would work well in such situations.

7.5 Future Materials Research

Our final item of note for the problem of thermally-activated actuation looks toward the development of future shape memory materials. Through changes in the consistency of particular elements, as well as material processes such as rolling and smelting, the thermal and mechanical properties of these alloys may be adjusted. Therefore, it may be possible to develop new shape memory materials with much faster thermal response times than current shape memory alloys allow. This may be possible through development of thinner wires and films, or by inclusion of materials having higher conductivity in the alloy mixture.

Moreover, considerable work is currently being done on heating and cooling devices to both increase their strength and reduce their size. To our knowledge, research in these areas is focusing on the development of cooling devices using the Peltier effect and heating devices based on the Joule effect. Advances in these areas may allow for increased controllability and localization in the thermal control of shape memory materials, therefore allowing for increased command over their phase transformations and behavior.
Bibliography


