A GRADIENT-BASED CONSTITUTIVE MODEL TO PREDICT SIZE EFFECTS IN THE RESPONSE OF SHAPE MEMORY ALLOYS

A Dissertation

by

MAJID TABESH

Submitted to the Office of Graduate and Professional Studies of
Texas A&M University
in partial fulfillment of the requirements for the degree of

DOCTOR OF PHILOSOPHY

Chair of Committee,	D.C. Lagoudas
Co-Chair of Committee,	J.G. Boyd
Committee Members,	A.A. Benzerga
	I. Karaman
Head of Department,	R. Bowersox

August 2015

Major Subject: Aerospace Engineering

ABSTRACT

Shape memory alloys (SMAs) show size effect in their response because the behavior of small-scale SMA structures deviates from that of the bulk material. Ni-Fe-Ga ferromagnetic SMA micropillars, for example, demonstrated a significantly increased hardening in their compressive stress-strain response as their diameter approached micron and submicron scales. This response cannot be modeled using conventional theories that lack an intrinsic length scale in their constitutive models. Constitutive models, however, are crucial for the design and simulation of SMA components at nano and micron scales as in NEMS and MEMS. Therefore, to capture such a size effect, a gradient-based thermodynamically consistent constitutive framework is established. We assume the existence of generalized surface and body forces that contribute to the free energy as work conjugates to the generalized variables of martensite volume fraction, transformation strain tensor, and their spatial gradients. The rates of evolution of the generalized variables are obtained by invoking the principal of maximum dissipation after assuming a transformation surface. This approach is compared to the theories that use a configurational force balance law. The developed constitutive model includes various energetic and dissipative length scales that can be calibrated experimentally. To demonstrate the capabilities of this model, a series of boundary value problems are solved. The boundary value problems contain the differential equation for the transformation surface as well as the equilibrium equation and are solved analytically and numerically. Example problems include pure bending of SMA beams, simple torsion of SMA cylindrical bars, and compression of SMA micro/nanopillars. The simplest version of the model, containing only the additional gradient of martensite volume fraction, predicts a response

with greater hardening for smaller structures. Also once calibrated, the model can qualitatively predict the experimentally observed response of Ni-Fe-Ga micropillars under compression.

DEDICATION

To my beloved father, Mahdi, may his soul rest in heaven for ever

and

To my dearest mother, Maryam, may her sunshine light up my life for ever

ACKNOWLEDGEMENTS

On the very outset of this dissertation, I would like to extend my sincere and heartfelt obligation towards all the personages who have helped me through this expedition. Without their active guidance, help, cooperation, and encouragement, I would not have made headway in this endeavor.

I am ineffably indebted to my advisers, Dr. Dimitris Lagoudas and Dr. James Boyd, for their conscientious guidance and encouragement, ameliorated with passion and excitement, for scientific progress over and above personal and professional betterment.

I am, also, extremely thankful and pay my gratitude to my colleagues and fellow students at the Department of Aerospace Engineering without whose genuine support the burden of this expedition would have been be unbearable.

I further acknowledge, with a deep sense of reverence, my gratitude towards my members of family, my lovely mother and my brothers, who have always supported me morally even though from across the oceans and also to all my friends who have altered the tenderness of studying abroad to enchantment.

Last but not least, recognition goes to all who directly or indirectly helped me complete my PhD studies, although not mentioned in this brief acknowledgment.

TABLE OF CONTENTS

Page

LIST OF FIGURES

LIST OF TABLES

TABLE Page

1. INTRODUCTION

Incorporation of shape memory alloys in small-scale applications cannot be successful without physical understanding and subsequent modeling of their behavior at such scales. The elastic and inelastic response of materials alters as the size of the specimen approaches the micron/nanometer region where fluctuations in the lower-scale microstructural and physical features of the material cannot be resolved by smearing their effect through averaging and homogenization. Modeling such an experimentally observed size-dependent behavior can be achieved in the realm of continuum mechanics without resort to cost-prohibitive molecular/atomisitc frameworks by using enriched or higher-order continuum theories. Section 1.1 discusses the current state-of-the-art for small-scale SMA actuators which provide a motivation for the experimental works, presented in section 1.2, aiming at understanding of the SMA behavior as the size of the specimen or its constituent microstructural features approaches the micron/nano scale. An effort is made to categorize the observed size-dependent response according to the dimensionality of the specimen. Different responses have been observed for SMA powders (0D structures), SMA wires or micropillars (denoted as 1D specimens) compared to 2D specimens such as SMA thin films or other SMA experiments involving samples with nano-sized grains (3D specimens). Finally, the prominent existing modeling techniques for the inelastic material response, falling in the category of enhanced or higher-order continuum theories, are reviewed in section 1.3. Although the majority of the research work pertains to modeling dislocation plasticity at micron scale through higher order plasticity theories, this can provide a firm ground for developing a higher-order phenomenological continuum theory to capture the response of shape memory alloys at smaller scales.

1.1 Significance of the research: small-scale actuators

The growth in the field of microsystems and its commercial applications created a great need for suitable microactuators. The low actuation frequency, with the order of 1 Hz, in bulk SMA actuators have hindered further implementation of these unique materials. Though the martensitic transformation front can, in principle, propagate with the wave speed of the material, this capability is limited by lower heat conduction rates specifically during the forward transformation stage where a cooling process is required.

Another drawback is the latent heat generated during the martensitic phase transformation. This latent heat is exothermic for forward transformation and endothermic for reverse transformation [111, 130]. It was shown by Tabesh et al. that the generation of latent heat during forward transformation and its absorption during reverse transformation decreases the actuation response [141]. Reducing the size of the actuator facilitates the dissipation of the latent heat generated.

Micro and nanoscale SMA actuators will have a higher heat transfer rate due to the high surface area to volume ratio. Therefore, it is possible to obtain much higher actuation frequencies by using small-scale actuators.

The other option for attaining higher actuation frequencies is implementation of magnetic shape memory alloys. One of the challenges in using FSMAs is their low actuation force and work output [31, 59]. Interestingly in this case too, it is possible to reach higher blocking stresses by reducing the size of the actuator.

SMAs have recently been used as high performance actuators for application in micro-electro-mechanical systems (MEMS) in the form of thin films/beams. SMA thin films, due to high actuation stress and strain and long fatigue life [55], are used in micropumps [17], microvalves [82], microgrippers [57], and microactuators

[133] (see figures 1 and 2). An NiTi thin film, for example, with dimensions of $1\,\text{mm} \times 100\,\mu\text{m} \times 5\,\mu\text{m}$ can exert a force of 0.5 N over an actuation distance of $30\,\mu\text{m}$. Also, as compared to bulk SMAs, the thin films can be thermally activated at a higher frequency due to the larger exposed free surfaces.

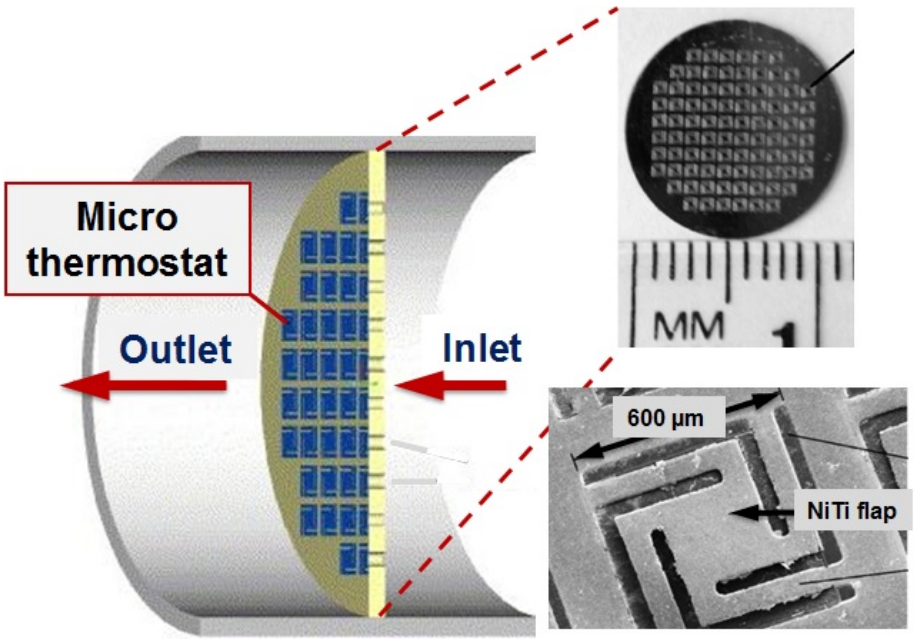


Figure 1: Schematic design of a microthermostat using NiTi microflaps [133].

The successful incorporation of SMAs in MEMS in the form of thin films/beams cannot be achieved without assessing their functionality using numerical modeling tools that can take into account the observed size effect. To this end, developing a constitutive model for shape memory alloys with the capability of capturing the size effect is chosen as the objective for this research work.

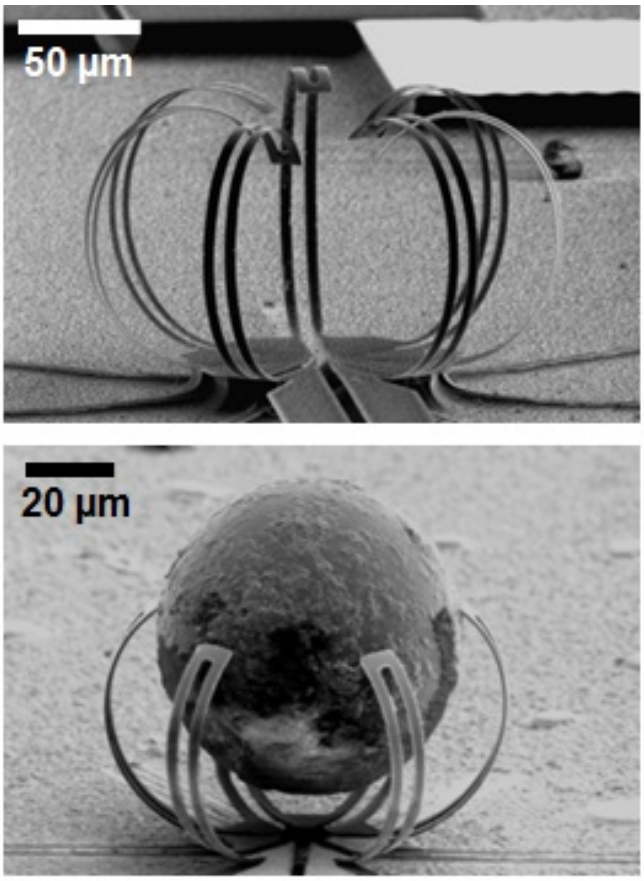


Figure 2: Schematic design of a microgripper using NiTi thin films [57].

1.2 Motivation: experimental evidence of size effect in the response of SMAs

Size effect refers to the dependence of a material property on the size of the specimen or sample. This generally occurs as a characteristic length, Ł, in the loading, boundary conditions, or geometry of the sample (such as dynamic wavelength, size of the stress concentration region, curvature of bending, or thickness of the sample) falls within the same order of magnitude as of a microstructural intrinsic length scale, ℓ,in the material (such as grain size, size of the dislocation loop in metals or martensitic twin variants in shape memory alloys). i.e. $\delta \to 1$ where

$$\delta = \frac{\ell}{L} = \frac{\text{Internal characterisitc length}}{\text{External characterisitc length}} \tag{1.1}$$

In that case, the heterogeneities in the microstructure of the specimen and hence the local state of the material become significant and cannot be averaged out to produce a larger scale, in many cases a global continuum, response [10].

The non-homogeneities in the microstrucutre of SMAs can be illustrated by considering the Ni-rich NiTi, as an example. Schryvers et al. [129] measured the lattice strain and Ni concentration near the Ni_4Ti_3 precipitates using detection tools and techniques provided by transmission electron microscopy (TEM) (as shown in figure 3). At distances about 50 nm from the precipitates with an average larger diameter of 300 nm, strains upto 2% and depletion of Ni to 49.5 at.% were detected. It is very well known, and also illustrated in figure 4, that Ni composition has a high impact on the transformation temperatures of NiTi. The strain, or stress, also shifts the transformation temperatures through the Clausius-Clapeyron effect. Therefore, the matrix contains very sharp gradients in the transformation temperature near the precipitates. This influences the response of nano or micro-sized specimens compared to larger bulk specimens.

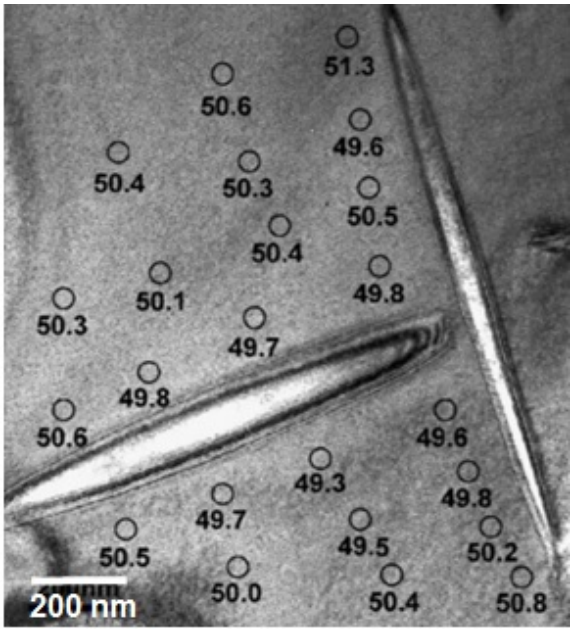


Figure 3: EDX measurements of the Ni concentration gradients near Ni_4Ti_3 precipitates in Ni-rich NiTi [129].

In addition, the size effect has been observed, for instance, for elastic properties of porous materials such as foams [92] or cortical bone tissue [26] where the flexural or torsional stiffness alters with the thickness or diameter of the specimen [16, 154]. The elastic modulus of ZnO nanowires were shown to significantly depend on the size for wires smaller than about 120 nm [23], whereas that of gold nanowires were essentially independent of the diameter [151].

A size effect is also seen in inelastic phenomena such as dislocation plasticity [48, 75, 105]. The yield stress of gold nanowires showed a 100-times strengthening when compared to the bulk material [151].

Likewise, the martensitic transformation is reported to demonstrate size effect. The material properties characterizing the unique behaviour of shape memory alloys,

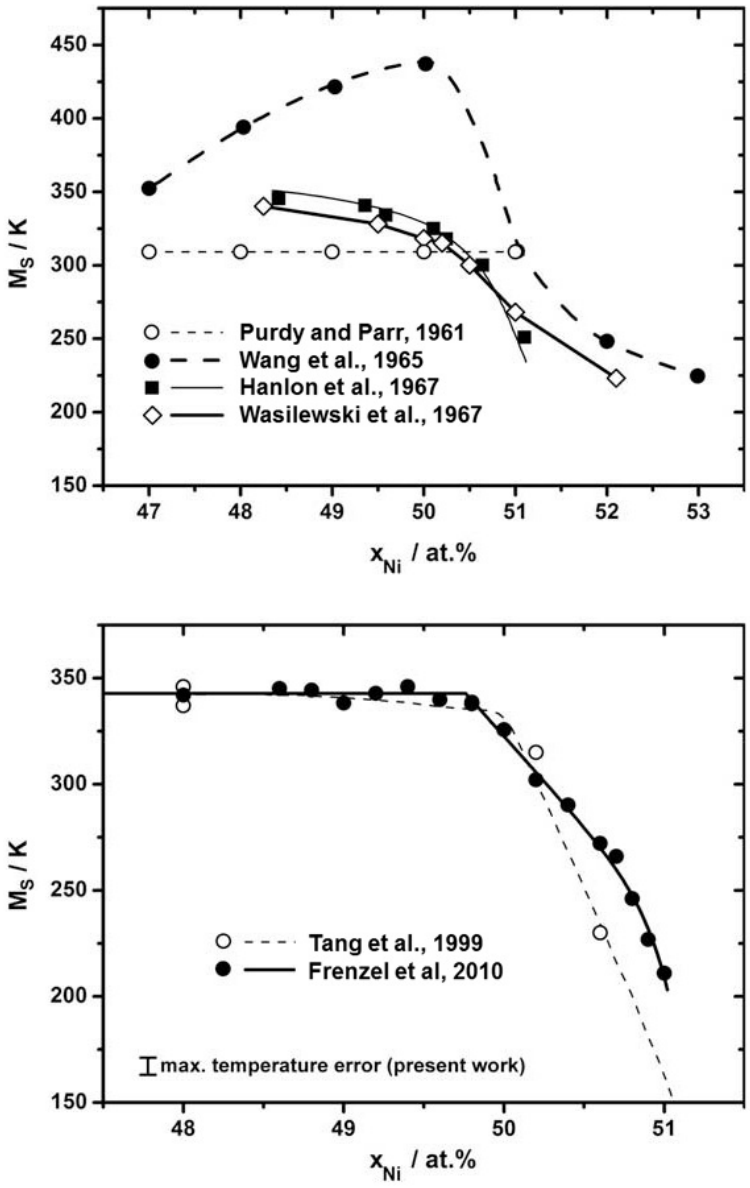


Figure 4: Effect of the Ni concentration on the martensitic transformation temperature (M_S shown here) [50].

i.e. the critical stresses to start and finish forward or reverse transformation as well as the transformation dissipation and hysteresis, are shown not to be independent

of the size of the specimen. The size effect in the SMA response has been observed in a variety of structural configurations from nanograined specimens to micropillars. In the next section, an attempt is made to categorize these phenomena based on the dimensionality of the corresponding material being tested.

Dimensionality and size effect in the reversible martensitic transformation:

The diffusionless phase transformations are prone to size effect due to the microstructural constraints (such as grain boundary incompatibility or precipitates) as well as constraining stress fields encountered at the phase propagation fronts. The experimental observations giving a clue for size effect in the behavior of shape memory alloys can be categorized based on the dimensionality of the sample or specimen being tested.

SMA nano particles represent the size effect in "0D" structures.

The size effect was demonstrated in microcrystalline particles in non-transformable solid matrix, and in free-standing powders. As observed from various experiments, by decreasing the grain size of the parent phase, the martensite transformation is fully or partially suppressed. The martensite transformation in nanoparticles within an amorphous matrix in NiTiCu alloys (produced by melt spinning technique) showed a critical size at which the martensite transformation was suppressed upon cooling [62]. Glezer et al. showed that the nanocrystals of larger than 25 nm transformed completely, the nanocrystals of 15-25 nm transformed partially, and the critical size below which no transformation occurred was determined to be 16 nm. The smaller the size of the nanocrystaline, the lower the extent of the transformation.

The martensitic transformation behavior of freestanding nanometer-sized B2-ordered AuCd particles, with an average composition of $Au_{50}Cd_{50}$ (at.%) synthesized by a wet-chemical process, were investigated by Frommen, Wilde, and Rosner [54].

The martensitic transformation start temperature, M_s, of nanometer-sized AuCd particles (with an average size of about 6 nm) was determined to be significantly lower than AuCd bulk alloys. AuCd particles with an average particle size of 46 nm transformed at ambient temperature, similar to the bulk material. Suppression of the martensite formation was attributed to the absence of pre-existing nucleation sites in the small AuCd particles. The authors also suggested that this decrease in the transformation temperatures could be linked to the fact that the free energies of the martensite and its parent phase in the nanometer-sized particles are lower than those of the bulk.

Nanocrystalline TiNi powders, prepared by the electro explosion of TiNi wires and with an average crystal size of 50 nm, has been characterized in the work by Fu and Shearwood [56]. DSC results showed the existence of martensitic transformation with a large transformation temperature hysteresis that was attributed to the altered composition of the alloy in the affected zone adjacent to the surface oxidation.

In accordance with metal plasticity, several studies have been performed on SMA micro/nanopillars and wires. Pillars and wires are essentially loaded in a one-dimensional sense representing the size effect in "1D" structures.

In-situ TEM studies performed on NiTi nanopillars showed existence of forward martensitic transformation in sample sizes of smaller than 200 nm [157]. Subsequent nucleation of martensite upon increasing the load resulted in a multi-step process. For engineering strains of more than 20%, the reverse transformation was inhibited. $Ni_{54}Fe_{19}Ga_{27}$ shape memory alloy (SMA) single crystalline micropillars (Figure 5), with diameters from 10 μm to 235 nm, were tested under compression in the work of Ozdemir et al [112]. The results revealed an increase in the critical stress for stress-induced martensitic transformation and the yield strength of martensite with decreasing pillar size (See figures 6 and 7). The stress hysteresis also increased with

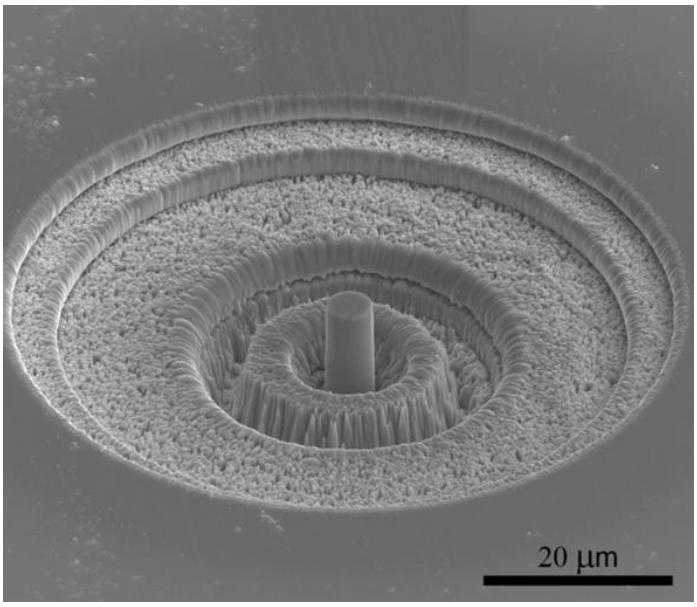


Figure 5: The SEM image of an SMA micropillar manufactured form a slab of $Ni_{54}Fe_{19}Ga_{27}$ [1 1 0]-oriented single crystal using focused ion beam (FIB) machining [112].

the reduction in the pillar size and the superelastic response started to diminish for pillar diameters below 500 nm. It was concluded that decreasing the sample size and increasing the temperature of the sample have similar effects on the superelastic response of NiFeGa SMAs. The micropillars had undergone a two-stage transformation and the results, hence, indicated that a reduction in pillar diameter decreases the transformation temperature due to the difficulty of martensite nucleation to occur on small scales. In the $1\,\mu$m and smaller pillars, showing only one-stage martensitic transformation, stress hysteresis increased significantly as the size decreased. The increase in stress-strain hysteresis and energy dissipation was attributed to the refinement of twin structures in smaller samples. Size-independent plastic deformation was observed in previous works for precipitate hardened SMA pillars. In the work

by Ozdemir et al. the SMA single crystal was free of precipitates, thus, the spacing between the precipitates was not the controlling characteristic length scale and a significant size effect for dislocation yielding of martensite was observed. The observations from this work indicate a size independent transformation-induced plasticity in the micropillars. However in accordance with the size effect in metal plasticity, the dislocation plasticity of martensite phase in the micropillars was shown to be size-dependent [112].

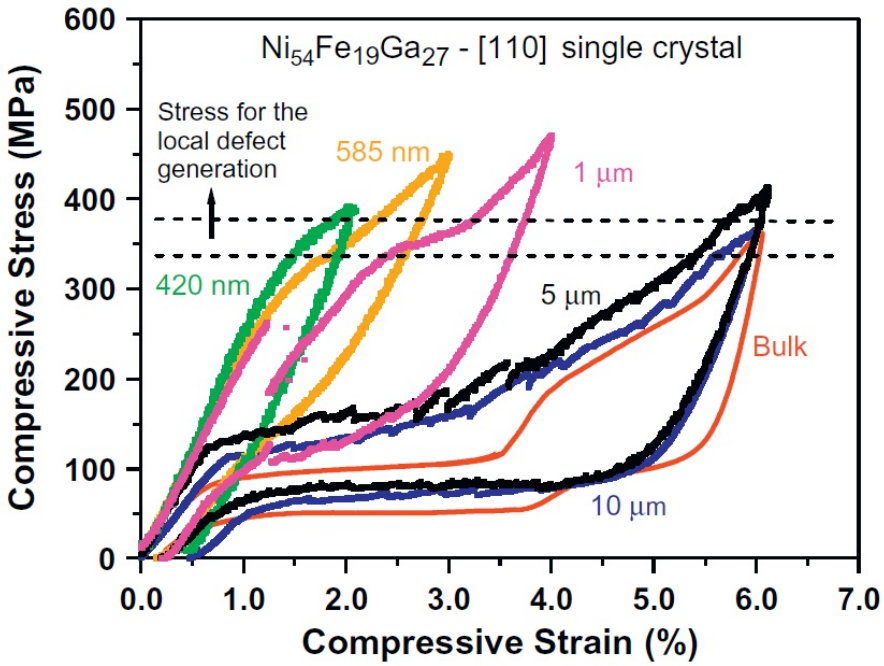


Figure 6: The superelastic stress-strain plots for $Ni_{54}Fe_{19}Ga_{27}$ SMA micropillars of various diameters [112].

The superelastic stress-induced B2 $\rightarrow$ B19$'$ transformation was investigated in

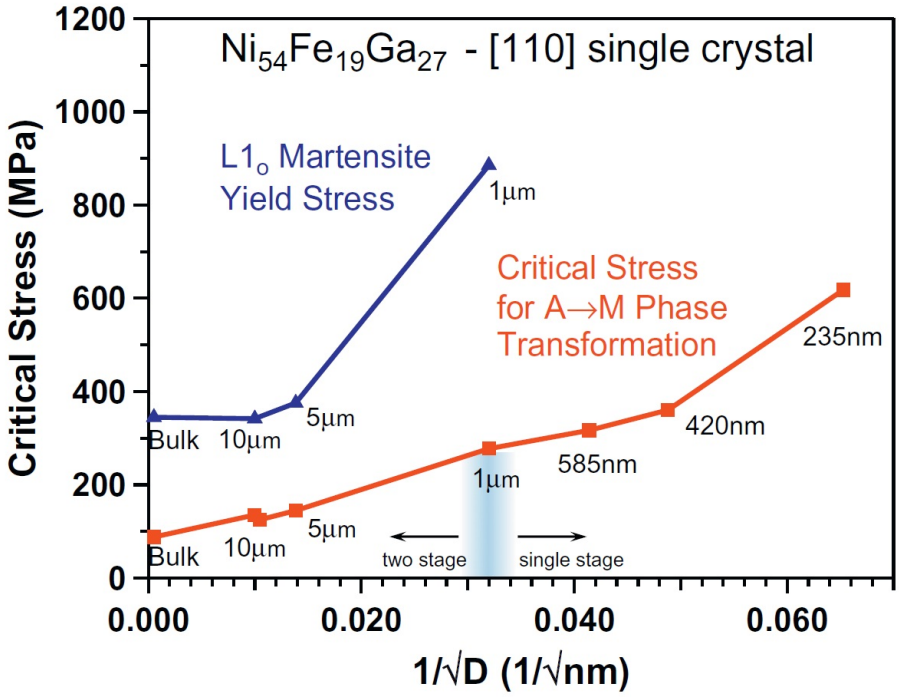


Figure 7: The critical stress for the start of martensitic transformation and martensitic plastic yield stress in an SMA micropillar as a function of the diameter, D [112].

50.7 at.% NiTi micropillars prepared by focused ion-beam (FIB) machining [110]. Microcrystal compression specimens (5 and 20 μm) were fabricated from a large [1 1 0] grain in a solutionized 50.7 at.%NiTi polycrystalline sample. The bulk [1 1 0] solutionized specimen demonstrated a ~60 MPa lower plateau stress for forward transformation compared to micropillars with similar orientation. No evidence of a complete loss in pseudoelastic behavior was found for the micropillar sizes tested and the results suggested a limited dislocation plasticity prior to the onset of the martensitic transformation.

In another work on NiTi micropillars, Frick et al. studied pillars ranging in diameter from approximately 2 μm to 200 nm [53]. The micropillars were prepared

using focused ion-beam micro-machining of aged [1 1 1] single crystal NiTi. Test results revealed size-dependent pseudoelastic behavior and size-independent dislocation yielding of the martensite phase. Pillars of aged [1 1 1] NiTi with diameters from 2 μm to 400 nm demonstrated pseudoelasticity. Decreasing the diameter led to a loss in the pseudoelastic behavior and ultimately it was suppressed for diameters less than 200 nm. The critical stress to start dislocation plasticity was shown to be independent of the pillar diameter.

Single crystal NiTi micropillars (12.6 and 9.4 μm in diameter) and bulk polycrystalline samples were studied in the work of Manjeri et al [99]. The fully reversible stress induced martensitic transformation did not show dislocation plasticity at the length scales investigated.

In their work on Cu-Al-Ni pillars, Juan et al. showed that the Cu-Al-Ni austenite and martensite phases are more stable in nanopillars than the bulk material leading to a much higher damping and energy absorption capacity [127]. Their previous work on this alloy demonstrated the existence of reversible stress-induced and temperature induced martensitic transformation in the nanometer scale [126]. The single crystal sample (produced via FIB) with a diameter of 900 nm was in austenite at room temperature. The size effect observed was attributed to geometrical aspects. Martensitic nucleation sites at the microstructural features, such as grain boundaries, dislocations or stress-concentrating surface defects, are scarce in single crystal SMA micropillars. Thus, upon loading of a SMA micropillar, there are fewer nucleation sites for martensite to from. The decrease in the critical stress for the reverse transformation was attributed to the stabilization of stress induced martensite due to the relaxation of strain energy generated during loading at the free surfaces. This was the consequence of martensite variants forming across the entire cross-section of the specimen. The two size effect phenomena observed were

1. The stabilization of austenite due to statistically reduced number of martensite nucleation sites and,

2. The improved stability of martensite due to relaxation effects at the free surface of the pillar,

leading to the forward transformation stress to be very high, the reverse transformation stress to be very low, and the load-displacement curve to exhibit a very large hysteresis. The constraining influence of gallium contamination as a result of FIB machining, which leaves an amorphous layer of less than $\sim$5 nm on the surface of the pillar, was disregarded by the authors [127].

In another recent work on single crystal Cu-Al-Ni micron and sub-micron sized pillars [81], Juan and No summarized three extrinsic size effects namely the increase in the critical stress to start the forward transformation σ^{fwd}, the decrease in the critical stress to begin the reverse transformation σ^{rev}, and the change of selection criteria for martensite variants. Nominal stress-strain results for 1.8 μm and 900 nm pillars were presented which demonstrated cyclic as well as size dependent effects compared to the bulk response. The sub-micron pillar showed a stable cyclic reversible martensitic transformation for loading and unloading with the first two extrinsic size effect features leading to an increased hysteresis and energy dissipation. On the other hand, the response of the micron sized pillar showed an increased σ^{fwd} followed by a partial recovery of the transformation. Subsequent cycling resulted in a stable forward/reverse transformation hysteresis, however, with a decreased σ^{fwd}. The dislocation generation due to the initial loading stage created ample nucleation sites for subsequent cycles of forward transformation which caused the σ^{fwd} to drop after the initial cycle. σ^{rev} was shown to be similar for both micron and sub-micron pillars, yet lower than the bulk material. As mentioned earlier, this can be attributed

to the relaxation of strain energy, otherwise conducive to the initialization of reverse transformation, at the free surfaces [81].

Frick et al. demonstrated that the behavior of NiTi compression micropillars, with diameters of 1000, 273, and 173 nm, depend on the diameter and is relatively independent of the orientation [51]. Irrespective of orientation or precipitate structure, the pseudoelastic hysteresis was diminished with diameters below approximately 400 nm, and fully inhibited with diameters below 200 nm.

Although not an experimental work, the pseudoelasticity and shape memory effect of single crystalline NiTi nanopillars were studied via molecular dynamics simulations [160]. In the bulk NiTi subjected to periodic boundary conditions, a temperature for the martensite start and finish could be determined. However in nanopillars, no phase transformation of B2 $\rightarrow$ B19$'$ was observed. The size effect, i.e. reduced phase transformation temperatures with decreasing pillar size, was attributed to the increasing role of surface atoms on phase transformation where the differences in phase energies varies between atoms at the free surface and in the bulk.

Bending of Ti-50.9 at%Ni single crystal micropillars with diameters of 1.2 μm and 200 nm were performed with in situ SEM demonstrating the existence of shape memory behavior [27]. Also, compression experiments on the nanopillars showed that the plastic deformation is independent of the size of the specimen and the precipitate microstructure; unlike the response of the bulk material.

Free standing In-21 at%Tl SMA nanowires were produced through mechanical pressure injection method (MPIM) and characterized with SEM, TEM and EDS for phase transformation [114]. It was concluded that the nanowires do not show any size effect for phase transformation ranging in diameter from 650 to 10 nm. The fact that these SMA nanowires show transformation at small scales while other configurations, though different alloys, show a size effect was attributed to the dimensionality of the

system as well as the lack of constraints in the free-standing nanowires. On the contrary, the nanograins are constrained by the matrix or adjacent grains and the thin film by the substrate or the surface oxide layer.

The blocking stress for ferromagnetic SMAs can also be increased by reducing the size of the actuator. Ganor et al. showed that, by using smaller samples (200 μm square cross section), the blocking stress achieved in a Ni_2MnGa actuators can increase up to twice the nominal 5 MPa stress [59]. Above this stress value, the strains induced by the magnetic field in a bulk sample diminish when magnetic domain rotation is energetically favorable to the variant reorientation. The modeling results also affirmed higher blocking stresses for smaller specimens due to increasing the energy barrier to magnetization rotation. In addition, domain theory calculations showed that smaller specimens favor finer twin structures.

Cu-Al-Ni microwires with diameters ranging from 500 to 20 μm were tested in the seminal work by Chen and Schuh using isothermal tensile testing and constrained thermal cycling [25]. The Oligacrystalline microwires had a bamboo type microstructure where the consecutive grains completely spanned the wire diameter. Such a microstructure alleviates the brittleness seen in the response of bulk Cu-Al-Ni SMAs that arises from intergranular fracture due to the stress concentration and incompatibility at the grain boundaries during martensitic transformation. Size effect was also observed in the response of the microwires with size-dependent transformation temperatures, critical stresses to start transformation, as well as stress hysteresis. The authors attributed the observed size dependence to:

1. The surface energy of austenite and matensite that can contribute to the size dependence of stress hysteresis,

2. The stored elastic energy associated with martensitic transformation and its

relief at the free surfaces leading to a lower reverse transformation stress, and

3. The obstacles and defects pinning the propagation of the transformation front leading to an increase in the hysteresis of the smaller microwires.

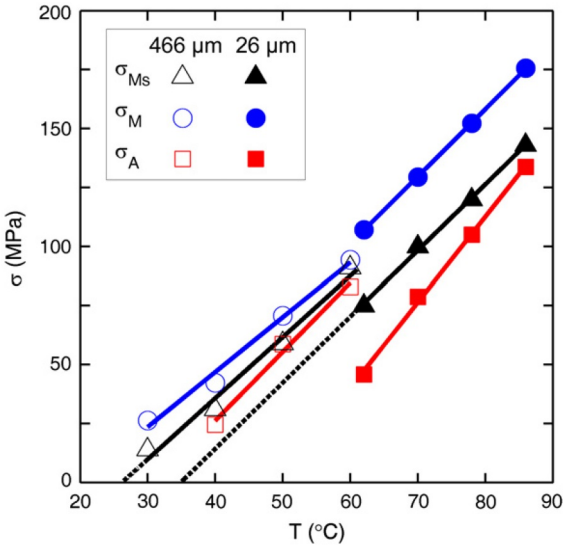


Figure 8: The martensite start stress σ_s^{fwd}, and average transformation stresses σ_M and σ_A, for 466 μm (hollow data points) and 26 μm SMA wires (solid data points) [25].

The experimental observations with regard to compression of micron and sub-micron sized pillars of shape memory alloys are summarized in table 1.1. Also included in the table, are the microstructural-based physical background giving rise to the phenomena of size effect, as discussed in the corresponding papers.

Table 1.1: Experimental observations in compression or bending of micron and sub-micron sized SMA pillars. All of the pillars were manufactured using focused ion-beam (FIB) machining.

Reported micropillar tests	**Alloy; Pillar diameter(D)**	**Microstructure**	**Transformation characteristics**	**Physical explanation for size effect**
Frick et al. †, 2007 [53]	Ti-50.9 at%Ni ; 200 nm to 2 μm	Aged (350 °C, 1.5 hr) [111] single crystal containing 10 nm Ni_4Ti_3 precipitates	Pseudoelastic behavior observed for 2 μm to 400 nm, diminishing with size reduction. No pseudoelasticity for D $\leq$ 200 nm. Transformation-induced plasticity independent of size.	(a), (e), (f)
Frick et al. †, 2008 [51] and 2010 [52]	Ti-50.9 at%Ni ; 1.03 μm, 273 nm, 173 nm	Aged (350 °C, 1.5 hr) [111],[001], and [210] single crystal containing 10 nm Ni_4Ti_3 precipitates; and solutionized (600 °C, 1.5 hr, water quenched) polycrystalline with 70 μm grain diameter (GD)	Reversible Martensitic transformation (MT) for all 1 μm pillars independent of orientation, psuedoelastic hysteresis inhibited for D $\leq$ 400 nm and suppressed for D $\leq$ 200 nm regardless of orientation and precipitation. Size independent yet orientation dependent (unlike bulk) martensitic plastic flow stress.	The transformation inhibiting mechanism size dependent but independent of orientation or precipitate microstructure, (f)

Table 1.1 Continued.

Reported micropillar tests	**Alloy; Pillar diameter(D)**	**Microstructure**	**Transformation characteristics**	**Physical explanation for size effect**
San Juan et al. †, 2009 [127]	Cu-Al-Ni ; 825 nm	[001]-oriented single-crystal, free of precipitates	Reversible MT with minimal permanent strain	(a) overruled, (b) and (c)
San Juan et al. ††, 2008 [126]	Cu-Al-Ni ; 400 nm, 1.7 μm	[001]-oriented single-crystal, free of precipitates	Reversible MT (both shape memory and psuedoelaticity via formation of martensite variants $\leq$ 25 nm in thickness) with minimal permanent strain	Size effect not investigated/observed compared to the bulk material
Clark et al. ††, 2010 [27]	Ti-50.9 at%Ni ; 1.2 μm to 200 nm for bending, 2 μm to less than 200 nm for compression	Aged (450 °C, 1.5 hr, water quenched) [111] single crystal containing 50 nm Ni_4Ti_3 precipitates	MT observed (both shape memory and psuedoelaticity) for diameters as low as 200 nm, size independent plasticity. Yielding in precipitate-free Ni single crystal micropillars was size dependent.	(a), (e)
Ye et al. ⋆, 2010 [157]	NiTi ; 140 to 200 nm	NiTi from vapor deposited thin film. Mostly single crystal pillars or with less than two grain boundaries	B2 $\leftrightarrow$ B19$'$ MT for samples $\leq$ 200 nm and loaded to $<$ 15% strain, multi-step transformation indicative of consecutive nucleation, no reverse transformation upon larger strains.	(a) overruled, (f), (g)
Manjeri et al. †, 2010 [99]	Ni-Ti (56.05 wt%Ni) ; 12.6 μm for the [101] and 9.4 μm for the [111]-oriented cylindrical micropillar.	Aged (400 °C, 1.0 hr, water quenched) [111] and [101] single crystal containing Ni_4Ti_3 precipitates	Orientation dependent reversible MT observed with minimal plastic residual strain. Orientation dependent dislocation yielding of stress-induced martensite.	Size effect not investigated/observed compared to the bulk material

Table 1.1 Continued.

Reported micropillar tests	**Alloy; Pillar diameter(D)**	**Microstructure**	**Transformation characteristics**	**Physical explanation for size effect**
Zhong et al. ∗, 2012 [160]	NiTi smooth pillars; 4.8 × 5.1 nm square cross section	compression along $< 100 >$	Irreversible twinning due to dislocation pinning of BCO twins observed at higher loads.	Nucleation stress peak and subsequent plateau with large hysteresis was observed.
Ozdemir et al. †, 2012 [112]	$Ni_{54}Fe_{19}Ga_{27}$; 235 nm to 10 μm	[110]-oriented precipitate-free single crystal	Reversible two-stage MT observed, with transformation-induced plasticity (TRIP) at higher stresses, pseudoelasticity suppressed for 235 nm pillar.	(a) overruled, (b) and (d), no size effect observed in the TRIP
San Juan and No † ⋆⋆, 2013 [81]	Cu-AL-Ni ; 900 nm, 1.8 μm	[001]-oriented single crystal	Micron scale: Increase of critical stress to start MT prior to generation of dislocations. Sub-micron scale: reversible MT, increased critical stress to start MT for forward transformation and reduced one for reverse transformation.	(b), (c)

∗ Molecular Dynamics (MD) atomistic simulations performed.

† Pillar compression, and ‡ bending tests performed using nanoindenter.

⋆ in situ TEM pillar compression as well as in situ diffraction tests performed.

⋆⋆ in situ TEM pseudoelastic tensile tests performed.

(a) Gallium contamination on the pillar surface after FIB.

(b) Stabilization of austenite due to paucity of nucleation sites leading to an increase in the critical stress to start forward transformation or a decrease in the M_s.

(c) Stabilization of stress induced martensite due to relaxation of strain energy generated during loading at the free surfaces as a result of martensite variants forming across the entire cross-section of the specimen. Hence, a decrease in the critical stress to start reverse transformation or an increase in the A_s.

(d) Increase in the energy dissipation through martensitic transformation due to the refinement of martensite twin formations.

(e) The effect of precipitates blocking dislocations with spacing smaller than the geometry of the specimen.

(f) Interaction of stress-induced martensitic transformation and dislocations generated during forward transformation (TRIP) resulting in the stabilization of martensite upon unloading.

(g) The martensitic transformation and recovery in the pillar substrate resulting in spurious hysteresis in the nominal stress-strain plots for micropillar compression.

Shape memory alloy thin films are promising candidates as small scale actuators [135]. Their response denotes size effect in "2D" structures.

NiTi thin films were prepared by Busch et al. using DC magnetron sputtering method [22]. Films were deposited up to 10 μm in thickness and were crystalized in order to develop martensitic transformation observed by DSC tests. The deposited films showed similar mechanical properties to the bulk material but showed transformation temperatures that were 100 °C lower than those of the parent material. This was attributed to the refinement of grain size in the specimen and also oxygen contamination that can alter the composition of the SMA.

In a work on Ni-Ti-Cu shape memory alloys, $Ti_{50}Ni_{50-\alpha}Cu_{\alpha}$ (α = 10, 15, 25 atomic %) thin films were tested by electric resistance for the effect of the thickness, 5-100 μm, on the characteristics of martensite transformation. It was shown that the transformation temperature decreases with the film thickness [11].

Kuninori, Sukedai, and Hashimoto investigated the size effect in the martensitic transformation of Ti-50.2%Ni thin foils [87]. Martensitic transformation was obstructed for thicknesses less than 100 nm. Thin films generally showed M_s temperatures less than that of the bulk TiNi.

The martensitic transformation in NiTi thin films on Si substrates was studied experimentally for $Ni_{50}Ti_{50}/SiO_2/Si$ film composites [125]. NiTi was sputter deposited to 1 μm with a buffer SiO_2 layer of 100 nm. The high level of residual stress state in the film changes the self-accommodation of the martensitic transformation in shape memory alloys. A cubic to tetragonal transformation of single crystalline material has been also used to model the transformation in the SMA thin films. The theory predicts considerable broadening of the temperature hysteresis of transformation as well as irreversibility of martensite transformation due to the differences in microstructures which accommodate the film/substrate and the austentite/martensite

interface constraints.

The thickness effect on the shape memory behavior of Ti-50.0 at.% Ni thin films was investigated by Ishida and Sato [76]. The films were first deposited on glass substrates in an amorphous form and then crystallized without the substrates. The transformation strain and residual strain under a constant stress were shown to be very sensitive to the film thickness when the thickness is less than the average grain size, 5 μm. As a result, the transformation strain and residual strain demonstrated a maximum around a thickness of 1-2 μm. For thicknesses of 5 μm and above, the transformation strain and residual strain were found to be almost constant. The decrease in the transformation temperature and the increase in the difference between the transformation start and finish temperatures were attributed to surface oxidation as a result of the heat treatment. Surface oxidation reduces the amount of Ti in the region beneath the surface oxide layer. This composition shift lowers the transformation temperatures and also the composition gradient across this zone increases the temperature range of the transformation. The transformation and plastic strains increased by decreasing the film thickness from 5 to 2 μm and decreased by going to 1000 and 500 nm thicknesses. There is a constrain to transformation due to the neighboring grains and its effect is in competition with the effect of the surface oxide layers. The constraint from the neighboring grains will saturate (will no longer be dependent on the thickness) as the thickness of the film becomes orders of magnitude greater than the average grain size of the SMA. The surface oxide and the affected zone beneath it lower the transformation temperatures. Also the gradients of composition occurring due to that increases the transformation hysteresis in the response [76].

In their work on SMA thin films, Fu et al. studied the effect of film thickness on phase transformation of constrained $Ti_{50.2}Ni_{49.8}$ films deposited on silicon substrates.

$Ti_{50.2}Ni_{49.8}$ films were prepared by co-sputtering of a $Ti_{55}Ni_{45}(at.\%)$ target (RF, 400 W) and a pure Ti target (70 W, DC). The actuation strain in the films were measured using the change of curvature in response to thermal cycling. For the films with a thickness below 100 nm, surface oxide and Ti-depleted adjacent regions establish constraining effect leading to local residual stresses and low recovery capabilities in the film [58].

The effect of film thickness on the phase transformations of radio-frequency (rf) sputtered TiNi shape-memory thin films were investigated by Wan and Komvopoulos using electrical resistivity (ER) measurements [149]. Thermal cycling was performed on the specimens in the temperature range of -150 to 150 °C. The films with thicknesses greater than 300 nm showed a hysteresis in their ER response. TiNi films with thicknesses lower than 300 nm had a smaller ER hysteresis and below about 50 nm no hysteresis was observed. The results indicated that constraints introduced by the film surface and film-substrate interface impose resistance on lattice distortion and twinning. The inhibition of these mechanisms, which control self-accommodation R-phase transformation, leads to the suppression and eventual disappearance of the shape memory effect for film thickness less than $\sim$100 nm.

It is possible to manufacture SMA polycrystalline samples comprised of nano-sized grains in order to attain enhanced properties. The altered behavior of such samples represents size effect in "3D" structures.

Waitz et al. used high pressure torsion to manufacture NiTi alloy specimens with a nanocrystalline microstructure. The specimens, after isothermal annealing, had grain sizes in the range of 5350 nm. Upon cooling, the nanostructures went through a partial transformation to B19 due to the pinning effect of the grain boundaries on the formation of martensite. It was concluded that with decreasing the grain size, the martensite twins were more refined, increasing the twin interfacial energy

and leading, ultimately, to an increased energy barrier. Therefore, the martensitic transformation is completely suppressed in grains smaller than 60 nm [147]. In a separate work, Waitz et al. investigated nanocrystalline NiTi samples via transmission electron microscopy (TEM) [146]. It was shown that even upon cooling to very low temperatures, grains of a size less than about 50 nm do not transform to martensite. To describe this size effect based on surface energy considerations, a micromechanical model of a spherical inclusion was used to calculate the transformation energy of the nanograins. A transformation dragging force, indicative of the barrier energy for martensitic transformation, was developed as a function of the grain diameter allowing a critical grain size (50 nm) to be found below which the martensitic transformation becomes prohibitive.

The martensitic transformation and the grain size effect in nanometer and micrometer scales were analyzed quantitatively by Malygin in the framework of the theory of diffuse martensitic transitions [98]. To that end, the phenomenological thermodynamics of phase equilibrium and kinetics of phase formation were used. The following three basic facts associated with the influence of a decreased grain size or film thickness on the parameters of the martensitic transformation in shape memory alloys were concluded:

1. A decrease in the critical transformation temperatures,

2. An increase in the transformation hysteresis, and

3. The existence of a critical grain size or film thickness below which the martensitic transformation in the SMA is inhibited.

The behavior of shape memory alloys alters by going form a single crystal to a polycrystalline form. This is due to the compatibility condition that the formation of

martensitic variants has to satisfy between neighboring grains with different crystallographic directions. Hence, the degree of symmetry in martensitic transformation becomes important where SMAs undergoing higher symmetry transformations (for example NiTi transforming from cubic to monoclinic compared to CuAlNi going form cubic to tetragonal) show higher levels of recovery and shape memory effect in polycrystalline form [18]. Also as a result of satisfying compatibility between the grains, the variant formation within a grain changes by distancing from the boundary [104]. The size of the grain imposes restrictions on the size of the twin formations too. Finer twins require more energy to form and also dissipate more energy.

The effect of the austenite grain size on the temperature to start martensitic transformation has been investigated by kinetic modeling [67,155]. Grain boundaries provide a nucleation site for martensite transformation and the size of the grain stabilizes the austenite (restricts martensitic transformation) by limiting the volume of martensite units. As a result, a lower temperature is required in finely grained austenite to start martensitic transformation, M_s.

Thermomechanical stability of martensitic phase transformation in $Ni_{49.7}Ti_{50.3}$ shape memory alloy, fabricated using equal-channel angular extrusion (ECAE), was studied under cyclic loading in the work by Kockar et al. [86]. The average grain sizes were between 100-300 nm. The experimental observations included an increase in the critical stress level for dislocation slip due to grain refinement, change in the transformation twinning mode in submicron grains, the presence of R-phase and multi-martensite variants, or a small fraction of untransforming grains due to grain boundary constraints. ECAE processing results in nano-range grains, partial R-phase stabilization, and an increase in the stress to start dislocation plasticity, thus, improved the cyclic stability of the alloy. Such small grain sizes influenced martensite morphology and helped form compound twins with twin thicknesses on the order of

a few nanometers.

Also, the effect of austenite grain size on the initiation of martensite transformation was studied by Guimares [66]. The results predicted the experimentally observed stabilization of austenite to martensite transformation in Fe-31.9%Ni-0.02%C alloy in the grain size range of 5 to 15 μm.

Ti-50.0(at.%)Ni with average grain diameter of 27 to 80 nm, fabricated via cold working (40 and 70%) and annealing, were experimentally investigated in [85]. The results of differential scanning calorimetry, transmission electron microscopy, isobaric thermal cycling, and tensile tests demonstrated that hysteresis increased with increasing the annealing temperature. This was attributed to an increase in the average grain size.

Last but not least, the energy dissipation ratio in nanoindentation of equiatomic NiTi SMA thin films with nano-sized grains was investigated by Wang et al [150]. The results demonstrated higher dissipation for thin films of smaller grain size under lower indentation depths.

Ultimately, the physical attributes that give rise to the size effect phenomena, reviewed within the context of the aforementioned SMA structures, can be summarized as below.

1. Although grain boundaries provide nucleation sites for martensite transformation, they impose constraining effect on the martensitic variant selection and also size of the twins. Hence, grain boundaries contribute to a higher energy barrier for transformation to take place. The martensitic transformation eventually suppresses for samples with grains smaller than a certain diameter.

2. The grain size in SMA thin films is smaller comparatively and gets close to the characteristic thickness of the thin film contributing to a strong size effect.

3. The constraining effect of the substrate and the surface oxide as well as the adjacent oxygen affected zone play an important role. Also, the SMA thin films have a biaxial internal residual stress state as a result of the manufacturing process that affects the transformation properties.

4. The existence of a free surface in small-scale SMA structures and its high ratio to the volume can provide relieving effect for the elastic strain energy in the sample and also accommodate formation of martensite variants otherwise less favorable to form.

5. Smaller samples have less probability of having martensite nucleation sites, i.e. precipitates, lattice defects, etc; hence a higher level of stress is required to start the martensitic transformation.

6. The heterogeneities in the SMA microstructure, resulting in spatial gradients in the evolution of martensite, become more significant in smaller scales.

1.3 Nonlocal methods in the modeling of inelastic material response

The term *nonlocal* encompasses a variety of generalized continuum theories that do not follow the principle of *local action*, as defined by Noll [109] for "simple materials". In other words, the constitutive material response for stress at a point in these models depends on the history of strain and temperature at that point as well as a neighborhood of it or the entire body. This is equivalent to the statement that the state of the material body (such as stress) at a point $\boldsymbol{X}$ at any time t, $\boldsymbol{\mathcal{S}}(\boldsymbol{X}, t)$, is a functional $\mathcal{F}$ of state variables (such as strain), $\boldsymbol{\mathcal{E}}$, at all points of the body:

$$\boldsymbol{\mathcal{S}}(\boldsymbol{X}, t) = \underset{\boldsymbol{X}^* \in V}{\mathcal{F}} [\boldsymbol{\mathcal{E}}(\boldsymbol{X}^*)] \tag{1.2}$$

In general, the functional $\mathcal{F}$ can be approximated by a series of spatial gradients of $\boldsymbol{\mathcal{E}}$ up to a desired order or by a series of volume integrals (suggesting to the integral or gradient-based approaches discussed later) [44].

The development of nonlocal continuum theories stems from the inability of classical continuum theory in describing certain experimentally observed mechanical phenomena.

The standard homogenization techniques cannot capture the dispersion of short-wavelength elastic waves in microstructurally heterogeneous or discrete crystalline media. The standard linear elastic continuum theory predicts a linear dispersion (the relation between elastic wave frequency, ω, and wave number, k) for both long and short wave-length waves, however, dispersion of elastic waves in crystalline materials are experimentally shown to deviate from such classical predictions [148,152,156]. For that purpose, strain-gradient and/or integral type elasticity models were considered [79]. The standard continuum theory, also, predicts a singular state of stress at the tip of a crack or center of a dislocation core. It is possible to eliminate these unphysical

singularities using the nonlocal theories of elasticity [40,42,93]. In addition, standard continuum mechanics theories have difficulty predicting size dependence in the elastic torsional or flexural response of certain materials [16, 23, 26, 83, 92, 151, 154].

The size effect in the elastoplastic response of micron and submicron-sized specimens, observed as increased plastic hardening of the structure, were captured using various types of implicit and explicit strain-gradient plasticity models [46, 79]. The hardness of several metals was shown to depend on the depth of indentation, being higher for smaller, nanoscale depths [63, 101, 144]. This phenomenon, called the indentation size effect (ISE), was attributed to the decreasing gradients of strain with an increasing indentation depth. Hence, strain gradient plasticity models were used for modeling purposes base on the concept of geometrically necessary dislocations that contribute to the enhanced hardening [48, 108]. The non-dimensional moment-curvature response of thin ($50-12.5\,\mu$m in thickness) Ni foils in micro bending experiments around small cylindrical mandrills demonstrated dependence on the thickness [134, 137]. Also, torsional stress-strain measurements performed on Cu wires ($15-170\,\mu$m in diameter) showed a size dependent yield stress, with smaller diameter wires being stronger [48]. The standard plasticity models, due to a lack of intrinsic material length scales, cannot capture the observed size effect.

Moreover, the phenomenon of strain localization in materials with softening behavior as a result of growth and coalescence of voids and cracks was modeled using the enriched continuum models, called localization limiters, including implicit, explicit, and integral-type nonlocal plasticity models [15, 79]. Standard plasticity or damage models with softening would lead to the loss of ellipticity in the governing differential equations and an ill-posed boundary value problem that would demonstrate itself as analytically-singular localization of the plastic deformation into a zone of zero width (a curve in 2D or a surface in 3D) or mesh-sensitivity of the discretized numerical

solution. The nonlocal material models assist the regularization of the problem by allowing the inelastic strain to concentrate in a narrow band in a continuous fashion preventing the pathological sensitivity to the discetization. The width of the localization band depends on the characteristic length of the material determined by the dominant heterogeneous features in the microstructure.

In general, if a characteristic length in the deformation field, such as the wavelength or some size in the structure, approaches a certain internal length scale in the material, second order effects become prominent and the standard continuum theory needs to be enhanced through various nonlocal approaches to account for those higher order effects. It is possible to account for such high resolution internal microstructural features explicitly by considering the spatial variation of the material properties in a lower scale (for example using an FEA model including detailed microstructural features or an atomistic/molecular modeling approach). However, this strategy requires extreme computational resources. The generalized nonlocal continuum models, with either differential or integral character, introduce one or more length scale(s) in the formulation. Any nonlocal continuum theory can be fully justified only if it produces a model that gives physically valid results in a much simpler fashion than that of a lower scale molecular theory.

Based on the definition by Rogula [124], any physical theory with a fundamental form given by an (possibly nonlinear) operator, $\mathcal{A}$, acting on a set of generalized degrees of freedom, χ, under the generalized forces, J, as in

$$\mathcal{A}\left[\chi(\boldsymbol{X})\right] = J(\boldsymbol{X}) \tag{1.3}$$

is local if it bears no direct relation between $\chi(\boldsymbol{X})$ and $J(\boldsymbol{Y})$; $\boldsymbol{X} \neq \boldsymbol{Y}$ and $\boldsymbol{X}, \boldsymbol{Y} \in \mathbb{R}^3$.

Another definition for locality is given based on the absence of characteristic length scales in the model. If the constants that characterize the constitutive model can *not* be combined into a parameter with the dimension of length, then the constitutive form is *invariant* to an arbitrary scaling of the physical space, hence it is local [124].

Any constitutive relation that can be categorized in the sense of both of the above definitions is called *strictly local*. The models that are local according to the first definition but contain a characteristic length scale are *weakly nonlocal*. Strain gradient elasticity or strain gradient plasticity models are weakly nonlocal as these theories are not form-invariant under the scaling of physical space. Finally, *strongly nonlocal* theories, such as nonlocal integral or implicit gradient theories, are nonlocal in the sense of both of the above definitions.

The micropolar theory of elasticity developed upon the early works of the Cosserat brothers [30,33]. In a micropolar continuum, in addition to the classical displacement degrees of freedom, each material point has three independent rotational degrees of freedom [39]. The isotropic theory for a micropolar continuum consists of six elastic constants (compared to two for the classical theory for a linear isotropic elastic materiel) making the parameter identification and model calibration complicated. Inspired by this, many nonlocal continuum theories were developed including Mindlin's theory of elasticity with microstructure [102] or the couple stress theory in which it is assumed that surfaces transmit both force and moment traction vectors [142]. The couple stress theory introduces both force and moment equilibrium equations as well as a non-symmetric Cauchy stress tensor and a moment stress tensor.

Another branch of nonlocal elasticity theories includes gradient elasticity models. In this theory, a.k.a Toupin-Mindlin theory, no new independent kinematic field is introduced. However, the constitutive equations are assumed to depend in addition

to the first, on the second and higher order gradients of displacement [64, 103]. The equilibrium equation is modified in this case to include the extra generalized stresses.

The nonlocal integral elasticity models consider the response functions for the material to depend on a volume integral of the dependent variables over the entire body [34, 35, 115]. As mentioned earlier, such models were intended for cases where the lower-scale atomistic/molecular approaches were too costly and classical local theories, on the other hand, fail to predict the observed experimental phenomena. The weighting function used for the spatial integration cannot be uniquely determined for a finite material body. Also, the weighting function has a decaying character as the distance with the corresponding material point increases. Therefore, its definition for the points near the material boundary becomes problematic leading to a boundary layer type behavior even in the presence of a homogeneous deformation [15].

The classical associated flow theory of plasticity with isotropic hardening (or softening) can be given through the following set of equations.

$$\mathrm{d}\sigma_{ij} = C_{ijkl}\left(\mathrm{d}\varepsilon_{kl} - \mathrm{d}\varepsilon^{p}_{kl}\right) \quad , \quad \mathrm{d}\varepsilon^{p}_{ij} = \mathrm{d}\lambda\,\frac{\partial f}{\partial \sigma_{ij}} \tag{1.4}$$

where λ is the plastic multiplier determining the norm of increment of the plastic strain ε^{p}_{ij} with respect to the yield surface f given by

$$f\left(\sigma_{ij}, \varepsilon^{p}\right) = F\left(\sigma_{ij}\right) - \sigma^{Y}\left(\varepsilon^{p}\right) \quad , \quad \sigma^{Y}\left(\varepsilon^{p}\right) = \sigma^{Y}_{0} + h\left(\varepsilon^{p}\right) \tag{1.5}$$

$F\left(\sigma_{ij}\right)$ is the equivalent stress which is equal to the Von-Mises stress $F\left(\sigma_{ij}\right) = \sigma^{eqv} = \sqrt{\frac{3}{2}\acute{\sigma}_{ij}\acute{\sigma}_{ij}}$ for the case of J_2 plasticity. σ^{Y} is the yield stress which is a function of the accumulated plastic strain, ε^{p}, representing the isotropic hardening/softening through h. The accumulated plastic strain that, here, plays the role of harden-

ing/softening variable is given via

$$\varepsilon^p = \int \mathrm{d}\varepsilon^p = \int \sqrt{\frac{2}{3}\dot{\varepsilon^p}_{ij}\dot{\varepsilon^p}_{ij}}\ \mathrm{d}t \tag{1.6}$$

The rate of change of the plastic multiplier during loading and unloading can be found via the Kuhn-Tucker conditions

$$\mathrm{d}\lambda \geq 0 \quad , \quad f\left(\sigma_{ij}, \varepsilon^p\right) \leq 0 \quad , \quad \mathrm{d}\lambda\, f\left(\sigma_{ij}, \varepsilon^p\right) = 0 \tag{1.7}$$

and the consistency condition $\dot{f} = 0$.

Such a model for a softening material behavior will lead to an ill-posed boundary value problem (because of the loss of ellipticity of the governing differential equations) that does not have a unique solution and ceases to continuously depend on the input data. As mentioned before, its incorporation in a finite element framework will lead to mesh sensitivity of the results. Nonlocal models help regularize the problem as localization limiters.

The inelastic response of materials can be described by enhanced or generalized nonlocal continuum models. For example, it is possible to generalize the classical flow theory to a nonlocal integral type [15] by redefining the yield stress in equation (1.5) to include the nonlocal accumulated plastic strain $\bar{\varepsilon}^p$.

$$f\left(\sigma_{ij},\ \varepsilon^p,\ \bar{\varepsilon}^p\right) = F\left(\sigma_{ij}\right) - \sigma^Y\left(\varepsilon^p,\ \bar{\varepsilon}^p\right) \quad , \quad \bar{\varepsilon}^p\left(\boldsymbol{X}\right) = \int_V \alpha\left(\boldsymbol{X}, \boldsymbol{Y}\right) \bar{\varepsilon}^p\left(\boldsymbol{Y}\right)\ \mathrm{d}\boldsymbol{Y} \tag{1.8}$$

where α is the selected nonlocal weighting function that can be related to the influence of the state of the material in a neighborhood of the point $\boldsymbol{X}$ and decays with increasing distance. It is commonly assumed to be the Gaussian distribution

function within which the material intrinsic length scale can be introduced. The weighted averaging for $\bar{\varepsilon}^p$ is performed over the entire material domain V.

The nonlocal integral model for plasticity, introduced by Eringen [41], was used to describe the interaction of dislocations and also stress distribution at the crack tips. The strain space plasticity framework similar to the classical plasticity theory was used, however, the stress (or the elastic strain in other words) was determined by integral averaging of the local stress on the material domain. The stress-space, associated J_2 plasticity version of Eringen's model was also introduced [43].

Nonlocal integral plasticity models that incorporate a nonlocal plastic strain tensor [14] or a nonlocal softening variable [20, 145] were also developed.

The thermodynamics-based associative isotropic hardening/softening plasticity model by Borino et al. was based on a nonlocal postulate of maximum plastic dissipation [20]. To that end, it was assumed that the plastic dissipation affects not only the point at which it occurs but also other material points in the body. Hence, the first and second laws of thermodynamics must be enforced in a global fashion and obtaining their local form for a point accompanies a nonlocality residual term. The state variables, except the internal variable representing softening/hardening, retained their local form. The state equations were obtained using a nonlocal Clausius-Duhem inequality followed by yielding laws for a nonlocal associative plasticity model. The plastic flow and yielding laws were shown to be the necessary and sufficient conditions for a nonlocal principle of maximum dissipation.

Weakly nonlocal constitutive models for inelastic material behavior include gradients of the state variables (hence also called *explicit* gradient models) that capture the effect of an infinitesimally small neighborhood around a material point. This can be the higher gradients of the displacement field (or the gradients of total strain) [45] or can be exclusive to the gradients of the internal variables such as the accumulated

plastic strain [2, 6, 47] or the martensite volume fraction [119, 140].

The early models of Aifantis [2, 3] and Coleman and Hodgdon [29] incorporated the second gradient of the plastic strain in the yield function representing the interaction of trapped and free dislocations as the kinetic nature for evolution of the plastic deformation [4]. The motivation was capturing the structure and evolution of the plastic flow localization in a small finite-sized region (the shear band) especially in the post-localization regime. The response of the models in the elastic regime was considered to be local. Higher order gradients of plastic strain (second and fourth order) were also considered in the yield function and constitutive equation for the flow stress to regularize the shear band localization [159]. The addition of intrinsic length scales through these models provides a size for the localization of the shear band (See [80] for a description of the regularization capability in such models using a one-dimensional example). The generalized yield surface in the explicit gradient theory based on the original Aifantis's model can be written as

$$f(\sigma_{ij},\ \varepsilon^p) = \sigma^{eqv} - \sigma^Y(\varepsilon^p) + C_1(\varepsilon^p)\,(\nabla\varepsilon^p \cdot \nabla\varepsilon^p)^m + C_2(\varepsilon^p)\,\nabla^2\varepsilon^p = 0 \tag{1.9}$$

The aforementioned gradient theory with the Laplacian of equivalent plastic strain was shown to be a generalization of Mindlin's theory of elasticity with microstructure (or the micromorphic theory) when the independent micro-deformation field variable is forced to coincide with the accumulated plastic strain as an additional constraint [49].

The previous strain-gradient theory of plasticity was also developed through variational principles in the work of Mühlhaus and Aifantis [106]. The variation of a potential functional was assumed to vanish upon equilibrium. The functional de-

pended on the displacement field and also the field of the equivalent plastic strain. It comprised of the elastic energy, the external work and the dissipated energy. The Euler-Lagrange equations of the functional resulted in the equilibrium equation and standard continuum boundary conditions as well as an additional differential equation for the effective plastic strain and its corresponding non-standard boundary condition. The method, also, gave a framework for numerical implementation of the model using finite element approach. The non-standard boundary condition, for the case of second-gradient of the equivalent plastic strain, can be of the Neumann or Dirichlet forms:

$$\frac{\partial \dot{\varepsilon}^p}{\partial n} \equiv \boldsymbol{n} \cdot \nabla \dot{\varepsilon}^p = 0 \quad \text{or} \quad \dot{\varepsilon}^p = 0 \tag{1.10}$$

which has to be satisfied at the material boundary and/or the a priori *unknown* elastic-plastic interface with a unit normal $\boldsymbol{n}$.

The problem of size effect in the torsion and bending of metallic bars was also investigated using the strain gradient models. As the thickness or diameter of the specimens become comparable to the internal length scale of the material, the specimen shows a size effect through an increased apparent hardening. Other phenomena explored are the formation of a boundary layer in shearing of a thin metal between two rigid plates (observed in and motivated by dislocation dynamics simulations) as well as indentation size effect in the nano/micro-indentation of metals and also the grain-size dependence in the mechanical response of ultrafine grain and nanocrystalline polycrystals [6–8, 143]. The coefficients of the gradient terms can be phenomenologically derived based on observation of shear band width or calibration with the size effect response plots. Also, a micromechanical approach in determination of the coefficients were used (utilizing a self consistent scheme along with a volume integral

definition for the average effective plastic strain over the considered RVE). The Taylor expansion for the average strain would lead to an expression for the coefficient of the second gradient of plastic strain in (1.9) based on the characteristic geometric features of the underlying heterogeneity [5].

Another class of strain gradient plasticity models were developed based on the concept of geometrically necessary dislocations (GND). Dislocations, as line defects in crystalline lattices, are kinematically divided to *statistically stored* dislocations (SSD) and *geometrically necessary* dislocations (GND). SSDs result from random entanglement of mobile dislocations leading to a homogeneous strain field with a vanishing net Berger's vector. GNDs, on the other hand, have a polarized net dislocation density and result in a nonuniform plastic strain field. GNDs are required to produce plastic strain gradients. It is then argued that both SSDs and GNDs contribute to the plastic strain hardening and the observed size effect is a result of the dominance of the contribution form GNDs in the case of a strain field with high gradients. The flow stress, τ, is then assumed to be proportional to

$$\tau \propto G\, b \sqrt{\rho_s + \rho_G} \tag{1.11}$$

where ρ_G and ρ_s are the densities of GNDs and SSDs, respectively. b is the magnitude of the Burger's vector in the system considered and G the shear modulus. ρ_G can be kinematically related to the gradient of accumulated plastic strain.

The three major experimental observations of size effect, i.e. torsion of thin wires, bending of thin foils and micro/nano-indentation tests, represent such strain fields.

The phenomenological strain gradient model by Fleck and Hutchinson [45, 48] is an extension of the couple stress gradient theory to the deformation theory of plasticity. In the couple stress theory it is assumed that surfaces transmit both force

and couple vectors. The elastic energy comprises, as work conjugates, contributions from Cauchy stress tensor σ_{ij} and the rate of strain tensor $\varepsilon_{ij} = \frac{1}{2}(u_{i,j} + u_{j,i})$ as well as couple stress tensor μ_{ij} and the rate of curvature tensor $\chi_{ij} \equiv \theta_{i,j}$ (the gradient of rotation vector associated with the displacement field $\theta_i = \frac{1}{2}\epsilon_{ijk}u_{k,j}$). Application of the principle of virtual power leads to force equilibrium and moment equilibrium equations with the corresponding force and moment traction boundary conditions. For extension to plasticity, the strain and curvature tensors are additively decomposed into elastic and plastic parts. The norm of the plastic curvature tensor is defined as the scalar measure of the density of GNDs. A generalized measure of an equivalent strain is considered that contains the effective strain and also the effective curvature with a length scale multiplier. The generalized equivalent stress (as the addition of effective stress and effective couple stress) is assumed to be related to the equivalent strain through the derivative a power-law strain energy function. The flow theory version is also derived based on assuming a yield surface as a function of the equivalent stress. The rates of plastic strain and curvature tensors are given based on the derivatives of the yield surface. The stress and elastic strain tensor and couple stress and elastic curvature tensors are considered proportional using tensorial elastic constants. The gradient-based nonlocality in the aforementioned model exists for both the elastic and plastic regimes, making the model complicated with several material constants as well as additional length scales.

In a subsequent work, Fleck and Hutchinson [46] presented their strain-gradient theory based on the Toupin-Mindlin gradient elasticity theory [102, 103, 142]. The Toupin-Mindlin theory can be connected with the couple stress theory by only considering those components of the strain gradient that correspond to the gradients of rotation, i.e. a couple stress theory were the rotations are not independent and constrained to the curvature resulting from the displacement field. The symmetric

strain tensor, ε_{ij}, and the gradient of displacement, $\eta_{ijk} \equiv u_{k,ij}$, are considered to be the kinematic variables. Cauchy stress tensor σ_{ij} is the strain energy work conjugate to ε_{ij} and the higher order stress τ_{ijk} to η_{ijk}. A linear elastic isotropic material using this model has 5 more material constants in addition to the classical Lame constants. The principle of virtual power provides the generalized equilibrium equation ($\sigma_{ij,j} + \tau_{ijk,jk} + b_i = 0$) and two traction boundary conditions. For generalization to plasticity, the flow theory of strain gradient plasticity was derived by decomposing η_{ijk} to elastic and plastic parts. As before, a generalized effective stress, Σ, was defined consisting of norms of both the Cauchy and the higher order stresses. Also, a generalized effective plastic strain, $\mathcal{E}^p$, was considered whose rate is given as the derivative of the yield surface with respect to Σ. $\mathcal{E}^p$ combined the invariant of plastic strain and three invariants of the plastic strain gradient with corresponding coefficient length scales. In addition, the elastic parts of the strain and strain gradient tensor were related to the stress and higher order stress through derivatives of the elastic strain energy.

$$\begin{aligned}
\varepsilon_{ij} &= \varepsilon^e_{ij} + \varepsilon^p_{ij} \quad , \quad \eta_{ijk} = \eta^e_{ijk} + \eta^p_{ijk} \\
\dot{\varepsilon}^e_{ij} &= M_{ijkl}\,\dot{\sigma}_{kl} \quad , \quad \dot{\eta}^e_{ijk} = K_{ijklmn}\,\dot{\tau}_{lmn} \\
\Sigma^2 = \frac{3}{2}\acute{\sigma}_{ij}\acute{\sigma}_{ij} + \sum_{I=1}^{3}\left[\frac{1}{\ell_I^2}\acute{\tau}_{ijk}\acute{\tau}_{ijk}\right] &\quad , \quad \Phi\left(\Sigma,\, \sigma^Y\right) = \Sigma\left(\sigma_{ij},\, \tau_{ijk}\right) - \sigma^Y\left(\mathcal{E}^p\right) = 0 \\
\dot{\varepsilon}^p_{ij} = \frac{3}{2h}\frac{\acute{\sigma}_{ij}}{\Sigma}\dot{\Sigma} &\quad , \quad \dot{\eta}^p_{ijk} = \frac{\dot{\Sigma}}{h\Sigma}\sum_{I=1}^{3}\left[\frac{1}{\ell_I^2}\acute{\tau}_{ijk}\right]
\end{aligned} \tag{1.12}$$

A single crystal plasticity version of the theory was also given. In addition to the previously mentioned size effect observations, the role of strain gradients in macroscopic strengthening of metal matrix composites containing rigid particle inclusions,

the growth of micro-voids, and the evolution of the stress field at the tip of mode I and II cracks were studied with this model [24, 153]. The macroscopic strength of particle-reinforced metal-matrix composites was shown to not only depend on the volume fraction of the particles but also on the particle diameter. An increased strength is observed for smaller particles in the range form 0.1 to 10 μm as a result of sharper gradients of plastic strain, hence GNDs, around the smaller particles. Such a phenomena was also observed for two-phase alloys.

A generalization of the Aifantis gradient plasticity model was presented in [47] by Fleck and Hutchinson. The elastic response in the new model was assumed to follow the classical local theory. The third order plastic strain gradient tensor was used to define a generalized effective plastic strain measure. Three length scales were introduced, to that end, representing the rotation as well as stretch rate gradients (corresponding to three quadratic invariants of the strain gradient tensor $\varepsilon^p_{ij,k}$ considering the incompressibility of ε^p_{ij}). A generalized form of the principle of virtual power was invoked were the displacement, plastic strain and gradient of plastic strain contributed to the internal virtual power through their stress-like conjugates. In addition to the conventional traction vector, the conjugate vector to plastic strain was also assumed to contribute to the external work applied on the boundary. Based on this idea, a potential functional depending on the displacement and generalized effective strain gradient field was developed the minimization of which provides the rate form of the differential equation for the yield surface and the corresponding non-standard boundary conditions on the elastic-plastic interface (similar to (1.9) and (1.10) including only the Laplacian term).

The mechanism-based strain gradient theory, pioneered by Huang and Gao [60,74, 120,121,158,158], is a multiscale framework that extends the concept of Taylor model in dislocation mechanics to the gradient plasticity. Two scales, namely microscale and

mesoscale, were considered in an RVE. The stress $\tilde{\boldsymbol{\sigma}}_{ij}$ and strain $\tilde{\varepsilon}_{ij}$ at the microscale were assumed to vary according to a local dislocation plasticity definition based on the Taylor model where the effect of the GNDs were accounted in the microscale yield stress $\tilde{\boldsymbol{\sigma}}^{Y}$ through the effective strain gradient, η, from the mesoscale; $\eta = \frac{1}{2} b\, \rho_G$.

$$(\tilde{\boldsymbol{\sigma}}^{Y})^2 = \sigma_0^2\, f^2(\tilde{\varepsilon}^p) + 18\, \alpha^2 G^2 b\, \eta \tag{1.13}$$

The intrinsic material length scale introduced was proportional to $\ell \equiv 18\, \alpha^2\, (\frac{G}{\sigma_0})^2\, b$ with $\alpha = 0.3$ being the constant of Taylor dislocation model. The zeroth-order average (volume integration over the RVE) of the microscale stress and strain provided the mesoscale stress and strain, while the first-order average of them accounted for the mesoscale higher order stress and the gradient of strain. The zeroth and first order averages of a microscale variable $\tilde{\xi}$ are defined below.

$$\xi^{(0)} = \frac{1}{V_{RVE}} \int_V \tilde{\xi}\, \mathrm{d}V \quad , \quad \xi^{(1)} = \frac{1}{V_{RVE}} \int_V \tilde{\xi}\, \tilde{\boldsymbol{x}}\, \mathrm{d}V \tag{1.14}$$

The generalized equilibrium equation, including the higher order stress and both stress traction and higher order stress traction, was considered with non-standard boundary conditions. The flow and deformation versions of this theory were developed which become identical for the case of an incompressible solid under proportional loading. The flow theory of mechanism based gradient plasticity was numerically implemented in a finite element framework. The theory was used to investigate the stress field around a stationary and steady state quasi-statically propagating mode-I crack tip. It was shown that the stress level predicted by the mechanism-based gradient plasticity is much larger than that of the classical plasticity due the dominance of GNDs, however both theories predict a similar plastic zone size [78]. The problem of microindentation in crystalline materials was also investigated using

this model to predict and characterize both the size effects form the indention depth as well as the spherical indenter radius [118, 121]. Also, the experimentally observed particle size effect in metal-matrix composites (A356 aluminum alloy reinforced by 15 vol.% SiC particles of 7.5 or 16 μm diameters [95]), attributed to the hardening effect of GNDs [122], and the problem of shear band localization [131, 132] was studied using the mechanism-based gradient plasticity model.

Gudmundson [65] presented a general thermodynamic framework that can be used to derive many of the existing strain gradient plasticity models. A generalized principle of virtual power was assumed that considers the contributions from elastic strain ε^e_{ij}, plastic strain ε^p_{ij}, and the plastic strain gradient $\varepsilon^p_{ij,k}$ along with their respective work conjugates of Cauchy stress σ_{ij}, microstress q_{ij}, and moment stress m_{ijk}.

$$
\begin{aligned}
\delta W_{int} &= \int_V \left[\sigma_{ij}\,\delta\varepsilon^e_{ij} + q_{ij}\,\delta\varepsilon^p_{ij} + m_{ijk}\,\delta\varepsilon^p_{ij,k}\right]\,\mathrm{d}V \\
&= \int_V \left[\sigma_{ij}\,\delta\varepsilon_{ij} + (q_{ij} - \acute{\sigma}_{ij})\,\delta\varepsilon^p_{ij} + m_{ijk}\,\delta\varepsilon^p_{ij,k}\right]\,\mathrm{d}V \\
\delta W_{ext} &= \int_S \left[t_i\,\delta u_i + M_{ij}\,\delta\varepsilon^p_{ij}\right]\,\mathrm{d}S \\
\delta W_{int} &= \delta W_{ext} \\
\sigma_{ij,j} = 0 \quad &, \quad m_{ij,k} + \acute{\sigma}_{ij} - q_{ij} = 0 \quad , \quad t_i = \sigma_{ij} n_j \quad , \quad M_{ij} = m_{ijk} n_j
\end{aligned}
\tag{1.15}
$$

Hence the principle of virtual power led to the balance equations and corresponding standard and non-standard boundary conditions with respect to force and moment force tractions. The balance law regarding microstresses is often denoted as the microforce balance after Gurtin [69]. The microforce balance acts as a generalized flow rule in this context. The rates of the plastic strain and the gradient plastic

strain were obtained based on satisfaction of the second law of thermodynamics (or the dissipation inequality) through derivatives of a yield function. The yield function was defined based on a generalized accumulated stress and strain measures that accommodated three different intrinsic material length scales. Connections between this theory and strain-gradient theories of Aifantis and Fleck-Hutchinson were also demonstrated. The presence of moment tractions, M_{ij}, can be used in modeling of the interface between two different plastically deforming phases. Example problems for bending of thin firms, torsion of thin wires and spherical void growth under far-field hydrostatic tension were solved.

Gurtin and Anand [70, 94] developed a theory of strain gradient viscoplasticity based on a generalized principle of virtual power consisting of power expenditures by the rates of elastic strain, plastic strain, and plastic strain gradient. This resulted in the macroscopic force balance law and the corresponding traction boundary conditions and also the microscopic balance law (microforce balance) supplemented with microtraction boundary conditions. The free energy for the material was assumed to depend on the elastic strain and also on the curl of the plastic strain (defined as the Burger's tensor). This led to the introduction of two energetic and dissipative material length scales. It is possible to introduce an additional length scale related to strain hardening due to the accumulation of GNDs. Assuming constitutive equations for the microstresses (work conjugates to the plastic strain and its gradient) consistent with the dissipation inequality (the isothermal form of the second law of thermodynamics), the microforce balance equation provides a nonlocal flow rule in terms of a tensorial differential equation for the plastic strain. The variational form of this equation can be used for implementation in finite element algorithms.

In another work by Gurtin and Anand [71], the thermodynamics consistency of the aforementioned models proposed by Aifantis and Fleck-Hutchinson was discussed.

The gradient theories were derived using the generalized principle of virtual power. It is shown that the generalized principle of virtual power based on the plastic strain tensors and its third-order-tensor gradient (similar to the one used by [65, 70] as in equation (1.15) can be reduced to the one using the accumulated plastic strain and its gradient (first order tensor, used by [47, 106]) assuming that the direction of the evolution of plastic strain is governed by the deviatoric part of stress. The general thermodynamic framework was based on considering the microsrtess, π, and micromoment, ζ_i, as work conjugates to the accumulated plastic strain, γ^p, and its gradient, $\nabla\gamma^p$, respectively. Hence, the first and second law of thermodynamics (considering an isothermal process) led to

$$\sigma_{ij}\dot{\varepsilon}_{ij} + \pi\dot{\gamma}^p_{ij} + \zeta_i\nabla\dot{\gamma}^p_i - \dot{\psi} \geq 0 \tag{1.16}$$

and assuming elastic and plastic (representing the defects) parts in the Helmholtz free energy, the dissipation became

$$\begin{aligned} \psi &= \psi^e(\varepsilon^e) + \psi^p(\gamma^p,\, \nabla\gamma^p_i) \\ D^p &= \left(\pi - \frac{\partial\psi^p}{\partial\gamma^p}\right)\dot{\gamma}^p_{ij} + \left(\zeta_i - \frac{\partial\psi^p}{\partial\nabla\gamma^p_i}\right)\nabla\dot{\gamma}^p_i \geq 0 \\ \sigma^{eqv} &= \pi - \zeta_{i,i} \quad , \quad \pi^{dis} \equiv \pi - \frac{\partial\psi^p}{\partial\gamma^p} \quad , \quad \zeta^{dis}_i \equiv \zeta_i - \frac{\partial\psi^p}{\partial\nabla\gamma^p_i} \end{aligned} \tag{1.17}$$

The microforce balance, in (1.17)c, represents the flow rule in this theory. The gradient plasticity models by Aifantis were obtained by assuming constitutive relations for π^{dis}, ζ^{dis}_i, and ψ^p. However, the Fletch-Huntchinson model was shown not to be compatible with this thermodynamics-based framework.

The thermodynamic framework introduced by Polizzotto and Borino [117] was

used to present a unified method for obtaining integral nonlocal as well as explicit gradient elasticity and plasticity theories [116]. The nonlocal first and second laws of thermodynamics in integral form can be localized for a single material point including a nonlocality residual as an additional state variable. Gradient elasticity and explicit gradient plasticity models could, therefore, be derived from this approach. The concept of nonlocality energy residual was introduced in the works of Edelen, Laws, and Eringen [34,35,41]. For the gradient plasticity model, the Helmholtz free energy ψ (and also the internal energy) was assumed to depend on the elastic strain, the effective plastic strain k, and its spatial gradients ∇k and $\nabla^2 k$.

$$
\begin{aligned}
&\psi = \psi^e\left(\varepsilon_{ij} - \varepsilon_{ij}^p\right) + \psi^p\left(k,\, \nabla k,\, \nabla^2 k\right) \\
&D^p = \sigma_{ij}\dot{\varepsilon}_{ij} - \dot{\psi} + R \geq 0 \text{ in } V \quad , \quad \int_{V^*} R \, \mathrm{d}V = 0 \quad \forall V^* :\ V^d \subseteq V^* \subseteq V \\
&D^p = \sigma_{ij}\dot{\varepsilon}_{ij}^p - \left(\frac{\partial \psi^p}{\partial k}\right)\dot{k} - \left(\frac{\partial \psi^p}{\partial \nabla k}\right)\nabla \dot{k} - \left(\frac{\partial \psi^p}{\partial \nabla^2 k}\right)\nabla^2 \dot{k} + R = \sigma_{ij}\dot{\varepsilon}_{ij}^p + X\dot{k} \geq 0
\end{aligned}
\tag{1.18}
$$

V^d was defined as the domain where the material points are under the nonlocal long-range effect of each other. The vanishing integral for R (called the insulation condition) provided the constitutive condition for X plus the corresponding nonstandard boundary conditions. It was shown that the dissipation inequality $D^p \geq 0$ can be satisfied through the principle of maximum dissipation leading to the following yielding and plastic flow equations.

$$
\begin{aligned}
f = f\left(\sigma_{ij},\, X\right) \leq 0\,, \quad \dot{\lambda} \geq 0\,, \quad \dot{\lambda} f = 0 \\
\dot{\varepsilon}_{ij}^p = \dot{\lambda}\frac{\partial f}{\partial \sigma_{ij}} \quad , \quad \dot{k} = -\dot{\lambda}\frac{\partial f}{\partial X}
\end{aligned}
\tag{1.19}
$$

An explicit gradient elasticity model was also obtained using this thermodynamic approach.

The explicit strain-gradient plasticity models can be divided to lower order and higher order ones. The higher order gradient theories, such as the ones by Aifantis and Fleck and Hutchinson, include higher order stresses and non-standard boundary conditions. On the other hand, the lower order gradient theories maintain the features of classical local flow theory of plasticity except for the incorporation of the strain gradient effect in the incremental tangential modulus. For that reason, such models are extremely amenable to implementation in numerical frameworks, specifically finite element method since they do not involve higher order stresses or non-standard boundary conditions. The lower order models, yet, are not able to predict the boundary layer phenomena related to the interface between grains or the pinning effect of the plastic-rigid interfaces. Using such models, the uniform stretching of a thin film on a non-metallic substrate, for example, will not show any size effect. Acharya and Bassani [1, 13] introduced one of the first versions of the lower order strain gradient theories for dislocation plasticity.

$$\dot{\sigma}_{ij} = C_{ijkl}\,(\dot{\varepsilon}_{kl} - \dot{\varepsilon}^{p}_{kl}) \quad , \quad \dot{\varepsilon}^{p}_{ij} = \left(\frac{\dot{\varepsilon}^{p}}{\sigma^{eqv}}\right)\acute{\sigma}_{ij} \quad , \quad \dot{\sigma}^{eqv} = h(\varepsilon^{p},\, \nabla^{2}\varepsilon^{p})\,\dot{\varepsilon}^{p} \tag{1.20}$$

in which the gradient dependence was entered in the instantaneous plasticity hardening function, $h(\varepsilon^{p},\, \nabla^{2}\varepsilon^{p})$, rather than the yield function. The problem of size effect in torsion of thin wires and hardening of metal-matrix composites (small hard particles in a single crystal matrix) were studied using this lower order gradient theory [13]. The lower order theory for the mechanism-based strain gradient plasticity was also developed in [74].

Niordson and Hutchinson [107] used a lower order gradient theory to investigate the deformation of two finite and infinite metallic layers under a pure shear loading. The results showed certain discontinuities (and mesh-size dependence) introduced by the model (vertex-type shear localization).

In contrast to the explicit gradient, the *implicit* gradient models belong to the category of the strongly nonlocal models. The explicit gradient models directly include the spatial gradients of the intended variable in the constitutive equations. The implicit gradient models, on the other hand, include an integral nonlocal variable in the constitutive equations that can be obtained via solution of a differential equation over the material body domain.

The nonlocal accumulated plastic strain given by equation (1.8)b can be approximated through a Taylor expansion to obtain

$$\bar{\varepsilon}^p(\boldsymbol{X}) = \varepsilon^p(\boldsymbol{X}) + C_1\nabla^2\varepsilon^p(\boldsymbol{X}) + C_2\nabla^4\varepsilon^p(\boldsymbol{X}) + \cdots \tag{1.21}$$

where the constants depend on the weight function α used (which includes the intrinsic material length scale) and odd-order derivatives vanish due to the isotropy of α. On the other hand, it is possible to find $\bar{\varepsilon}^p$ exactly through the Green's function method by wisely choosing $\alpha\left(\boldsymbol{X},\boldsymbol{Y}\right) = G\left(\boldsymbol{X},\boldsymbol{Y}\right)$ to be the Green's function of the Helmholtz equation. Therefore

$$\bar{\varepsilon}^p(\boldsymbol{X}) - C\,\nabla^2\bar{\varepsilon}^p(\boldsymbol{X}) = \varepsilon^p(\boldsymbol{X}) \quad , \quad \frac{\partial\bar{\varepsilon}^p}{\partial n} \equiv \boldsymbol{n}\cdot\nabla\bar{\varepsilon}^p = 0 \quad \text{on} \quad \partial V \tag{1.22}$$

A Helmholtz partial differential equation is, hence, obtained for the nonlocal variable with Neumann boundary conditions applied on the external boundary of the

domain (and not on the evolving elastic-plastic boundary) and solved in a coupled fashion with the equilibrium equation over the entire domain. The boundary condition, also, guarantees the equality in volume averages of the local and corresponding nonlocal variable over the domain $\int_V \bar{\varepsilon}^p \, \mathrm{d}V = \int_V \varepsilon^p \, \mathrm{d}V$.

Engelen et al. [38] proposed an implicit gradient plasticity model in which the nonlocal yield function was

$$f\left(\sigma_{ij},\ \varepsilon^p,\ \bar{\varepsilon}^p\right) = F\left(\sigma_{ij}\right) - \sigma^Y\left(\varepsilon^p,\ \bar{\varepsilon}^p\right) = \sigma^{eqv} - \left[1 - \omega^p(\bar{\varepsilon}^p)\right] \sigma^Y(\varepsilon^p) \tag{1.23}$$

The rate of the plastic strain tensor was given through an associative flow rule; $\dot{\varepsilon}^p_{ij} = \dot{\varepsilon}^p \frac{\partial f}{\partial \sigma_{ij}}$. The point of departure from classical plasticity was the dependence of the yield function on the nonlocal variable through the damage variable ω^p that increases from 0 to 1 upon complete failure of the material. The softening material response and subsequent shear band localization was studied after implementation of the model in a finite element framework. They demonstrated the effectiveness of the implicit gradient model in contrast to the explicit models that are not able to reach a stress-free state upon complete material failure. The finite strain version of this model was also developed in [61].

Engelen et al. presented a comparison between the higher-order gradient plasticity theories as for their ability in capturing the problem of size effect as well as shear localization [37]. The pure beam bending of thin foils and one-dimensional problem of localization in a bar in tension were studied using the strain gradient theories of Fleck-Hutchinson 1997 [46], Fleck-Hutchinson 2001 [47], and the implicit gradient theory of Engelen et al. [38]. Both the theories of Fleck-Hutchinson, unlike the implicit nonlocal model, were able to predict strengthening effects comparable to the

experimental results. The harmonic incremental solutions for a bar under uniform tensile strain were considered using the previous models to examine their localized behavior as bifurcation form a homogeneous state. The nonlocal implicit model, due to the ability in capturing the nonlocal long range interactions in the plastic and elastic regions, predicted the transition from hardening to softening regime and allowed for a shear band with finite size. The localization predicted by Fleck-Hutchinson 1997 did not vanish at complete failure. But, it could predict finite-sized shear bands due to the existence of gradient effects in both the elastic and plastic response. The conclusions by Engelen et al. are presented in table 1.2 with ++ denoting the best, + a satisfactory, and – a performance with concerns or shortcomings.

Table 1.2: The performance of higher-order strain gradient models in predicting size effect and localization in the elastic-plastic material response [37].

Gradient theory	Size effect	localization
Fleck-Hutchinson 1997 [46]	++	+
Fleck-Hutchinson 2001 [47]	++	–
Nonlocal implicit Engelen et al. [38]	–	++

Phenomenological modeling of the inelastic, history-dependent, thermomechanical response of shape memory alloys through internal variable approach has been under development for the past decade [88]. Although the martensitic transformation underlying the unique behavior of SMAs is diffusionless and shear-driven, the similarities in the phenomenological modeling approaches with that of dislocation-based plasticity is undeniable. To that sense, the higher-order nonlocal continuum models developed to capture some of the experimentally observed phenomena in the response of SMAs, such as transformation front localization and propagation or macroscopic

size effect, follow the footsteps of gradient-enriched plasticity models [12, 32, 119].

Duval et al. [32] provided the nonlocal extension of an existing phonomenological SMA constitutive model [113]. This nonlocal implicit gradient model was inspired by the work of Engelen et al. [38]. In addition to the conventional internal variable of martensite volume fraction (MVF) f, the integral average of it was also considered as the nonlocal MVF $\bar{f}$. Similar to the implicit gradient plasticity models, $\bar{f}$ was obtained from an additional PDE with a Neumann-type boundary condition on the domain boundary ensuring the equality of local and nonlocal MVF averaged over the entire domain. The developed nonlocal model was implemented in the finite element software package ABAQUS through the user element (UEL) feature. The experimentally observed softening behavior (stress peak of nucleation and subsequent stress plateau for propagation) in the SMAs was captured by gradually decreasing the critical force for martensitic transformation through the nonlocal MVF resulting in lower stresses required for propagation of martensite compared to its nucleation. Therefore, the nonlocal model performed as a localization limiter improving the otherwise pathological behaviors of the local models. The nucleation and propagation of martensite transformation front in tensile loading of an SMA plate with a hole was simulated with results showing the dependence of the localization width and the stress peak of nucleation on the intrinsic length scale introduced in the model through the nonlocal variable.

Another nonlocal implicit gradient SMA model was presented in the work of Badnava et al. [12]. The model was based on a 3D extension of Brinson's SMA model [21] and was intended to capture the unstable softening and localization behavior of SMAs upon nucleation and propagation of martensite transformation. The effectiveness of the model as a localization limiter was demonstrated through FEA simulation of SMA structures after implementation in ABAQUS UEL.

The explicit gradient-based SMA model of Qiao et al. [119] was inspired by the strain gradient plasticity work of Gurtin and Anand [70]. This isothermal one-dimensional model was intended to capture the size-dependent superelastic response of SMA micro/nanopillars under compression tests. The Helmholtz free energy for the material contained the gradient of martensite volume fraction through which an energetic length scale was introduced. A generalized form of the principle of virtual power was also used where the internal power expenditure was attributed to the work of stress and elastic strain, as well as the microstresses conjugate to the martnesite volume fraction and its gradient. The external power supply, in addition to the stress traction, came form the work of microstress traction on the boundary conjugate to the martensite volume fraction. This resulted in the stress equilibrium equation and a balance equation regarding the microstresses (equivalent to the microforce balance) as well as standard traction and non-standard microtraction boundary conditions. Constitutive relations were assumed for the microstresses in which a second dissipative length scale was introduced. Therefore, the microforce balance played the role of the transformation partial differential equation giving the rate of martensite volume fraction. The variational form of this equation was implemented in a 1D finite element framework to simulate size effect in the response of Cu-Al-Ni SMA micropillars under compression.

In the work by Sun and He [138], a multiscale continuum phenomenological strain-gradient model was developed to study the effect of grain-size in the response of polycrystalline SMA specimens. The model was based on a nonlocal non-convex strain energy function. The characteristic length scales introduced were the specimen size L, grain size l, and the intrinsic material length scale ℓ related to the width of austenite-martensite interface. The results demonstrated that the energy dissipation during phase transformation is governed by the ratios of $\frac{L}{l}$ and $\frac{l}{\ell}$. The martensitic

transformation, hence dissipation, was diminished in the case of an SMA with large grains close to the size of the specimen (denoting a single crystal behavior) or in the case where the grains are of the order of nanometer in size close to the intrinsic length scale (denoting the transformation inhibition observed in ultrafine grain SMAs). The microscopic domain nucleation, front propagation, and domain vanishing for a 2D quasi-static stretching of a polycrystalline SMA strip was captured by implementation of a Cahn-Hilliard phase field SMA model in an FE framework. The results also showed the nucleation peak stress and multiple nucleation and vanishing events as serrated stress-strain response plots as well as the effect of the governing length scale ratios on the dissipation and stress hysteresis.

1.4 Goals

Any attempt to implement shape memory alloys in small size actuation or sensing applications requires fundamental understanding of their behavior in smaller scales. Similar to other inelastic phenomena, martensitivc transformation shows a size dependent response where such key properties as critical stresses to start or finish the transformation or the transformation hysteresis alter with the specimen size at scales of below $100\,\mu$m. The conventional constitutive models are not able to predict the size effect due to a lack of characteristic length scale.

The purpose of this research is to develop a gradient-based continuum nonlocal constitutive model for shape memory alloys based on a generalized thermodynamics-based framework and also the thermodynamics of internal variables. In addition to the conventional internal variables of martensite volume fraction and transformation strain, the model using the internal variable theory contains the spatial gradient of maretensite volume fraction and the transformation strain as an independent internal variable. This enables introduction of energetic and dissipative length scales that

contribute to a defined *nonlocal parameter*.

The various length scales combined into a nonlocal parameter can be calibrated based on experimental observations constituting a non-homogeneous state of stress and hence martensitic transformation. Compression of SMA micropillars, torsion of SMA wires, and bending of SMA thin films are shown, herein, to provide a calibration/prediction ground for the developed nonlocal model. The rest of the SMA material properties can be measured, as in the classical way, from uniaxial stress-strain and calorimetry experiments.

The developed gradient-based model will be used to study the response of three SMA structural problems. Compression of SMA micropillars, bending of SMA thin films, and torsion of SMA bars. The experimental observations for micropillar compression of SMAs exist and show a significant size effect; i.e. increase in the critical stress to start martensitic transformation in smaller micropillars. To the best knowledge of the authors, experimental evidence for bending of SMA thin films has not yet been published.

2. BRIEF REVIEW OF CONTINUUM MECHANICS

In this chapter, the fundamentals of continuum mechanics, namely the kinematics, balance laws and constitutive equations, are reviewed. The reference textbooks, [68,91,97] , are recommended for more extensive discussions. A generalization is made to the first law of thermodynamics by considering a set of generalized state variables that are capable of performing work through generalized body and surface forces.

The material is assumed to be controlled by changing the state of stress, $\boldsymbol{\sigma}$, and absolute temperature, T. A set of generalized variables $\boldsymbol{\Upsilon}$, related to the microstrucutre of the material, is also assumed in the current modeling framework as independent variables. $\boldsymbol{\sigma}$, T, and $\boldsymbol{\Upsilon}$ can vary independently. The definition of $\boldsymbol{\Upsilon}$ as internal variables or internal degrees of freedom is discussed later in this chapter. The dependent state variables that change in response to the changes in the independent state variables are assumed to be

$$\begin{aligned} \boldsymbol{\varepsilon} = \hat{\boldsymbol{\varepsilon}}\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right) \quad , \quad \boldsymbol{q} = \hat{\boldsymbol{q}}\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right) \quad , \quad s = \hat{s}\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right) \\ \boldsymbol{\mu}^s = \hat{\boldsymbol{\mu}}^s\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right) \quad , \quad \boldsymbol{\mu}^b = \hat{\boldsymbol{\mu}}^b\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right) \quad , \quad G = \hat{G}\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right) \end{aligned} \tag{2.1}$$

The thermodynamic state of the material is completely characterized by the above response functions [1]. $\boldsymbol{\varepsilon}$ is the infinitesimal strain, $\boldsymbol{q}$ the heat flux vector, s is the specific entropy and G the specific Gibb's free energy. $\boldsymbol{\mu}^s$ and $\boldsymbol{\mu}^b$ are the generalized forces that perform as energy conjugates to $\boldsymbol{\Upsilon}$ through surface or body action, respectively.

[1] Coleman and Gurtin showed that it is necessary and sufficient to satisfy the 2nd law of thermodynamics if the response functions depend on the temperature only and not its spatial gradients (of any order) and also if the Fourier's constitutive law is used [28].

The specific internal energy, u, appearing in the first law of thermodynamics is a *themodynamic potential* that characterizes the thermodymaic state of the material. However, *Helmholtz free energy*, ψ, and *Gibbs free energy*, G are commonly used as the potential functions for derivation of the material constitutive relations. Helmholtz free energy can be defined as the portion of the internal energy available for doing work at a constant temperature and the Gibbs free energy as the portion of the Helmholtz free energy that can be released as heat at constant applied stress. ψ and G can be related to u through the following Legendre Transformations.

$$\psi = u - sT \quad , \quad G = u - \frac{1}{\rho}\boldsymbol{\sigma} : \boldsymbol{\varepsilon} - sT \tag{2.2}$$

Tensors of any arbitrary rank can constitute the set of generalized state variables $\boldsymbol{\Upsilon} = \{\boldsymbol{\Upsilon}_1, \boldsymbol{\Upsilon}_2, \boldsymbol{\Upsilon}_3, \cdots\}$ and their corresponding generalized forces $\boldsymbol{\mu}^s = \{\boldsymbol{\mu}^s_1, \boldsymbol{\mu}^s_2, \boldsymbol{\mu}^s_3, \cdots\}$ and $\boldsymbol{\mu}^b = \{\boldsymbol{\mu}^b_1, \boldsymbol{\mu}^b_2, \boldsymbol{\mu}^b_3, \cdots\}$ [2]. It is possible to relate the variables $\boldsymbol{\Upsilon}_i$ to different physical phenomena underlying the mechanism of microstructural rearrangement being studied. For example, plastic deformation as a result of the movement of dislocations, twinning of crystals, or slip of grain boundaries or transformation strain due to the stress/temperature-induced martensitic transformation. This relation between any mentioned mechanism and the chosen generalized state variable is point-wise but is based on an average measure of those physical mechanisms. In the case of SMAs, the volume average extent of transformation from the parent phase to the product phase is considered without worrying about any crystallographic details such as twin boundaries or habit planes. Their effect, on the other hand, is considered in the terms related to mixture and interfacial energies within the free energy definition.

For the current study of an SMA gradient theory, martensitic transformation is

[2]It is assumed here that if $\boldsymbol{\Upsilon}_i \in \mathcal{T}^m$ then $\boldsymbol{\mu}^s_i \in \mathcal{T}^{m+1}$ and $\boldsymbol{\mu}^b_i \in \mathcal{T}^m$.

considered as the mechanism underlying the microstructural changes in the material. Therefore, $\boldsymbol{\Upsilon}$ is taken to represent the extent of martensitic transformation, or martensite volume fraction at a point, ξ, and its spatial gradients, $\nabla\xi, \cdots, \nabla^p\xi$, as well as the martensitic transformation strain tensor, $\boldsymbol{\varepsilon}^{tr}$, and its spatial gradients, $\nabla\boldsymbol{\varepsilon}^{tr}, \cdots, \nabla^q\boldsymbol{\varepsilon}^{tr}$. Hence, the set of generalized variables may include

$$\begin{aligned}
\boldsymbol{\Upsilon} &= \{\Upsilon_1, \boldsymbol{\Upsilon}_2, \cdots, \boldsymbol{\Upsilon}_p, \boldsymbol{\Upsilon}_{p+1}, \cdots, \boldsymbol{\Upsilon}_{p+q}\} \\
\Upsilon_1 = \xi\ , &\quad \boldsymbol{\Upsilon}_2 = \nabla\xi\ , \quad \boldsymbol{\Upsilon}_3 = \nabla\nabla\xi\ , \cdots,\ \boldsymbol{\Upsilon}_p = \nabla^p\xi \\
\boldsymbol{\Upsilon}_{p+1} = \boldsymbol{\varepsilon}^{tr}\ , &\quad \boldsymbol{\Upsilon}_{p+2} = \nabla\boldsymbol{\varepsilon}^{tr}\ , \quad \boldsymbol{\Upsilon}_{p+3} = \nabla\nabla\boldsymbol{\varepsilon}^{tr}\ , \cdots,\ \boldsymbol{\Upsilon}_{p+q} = \nabla^q\boldsymbol{\varepsilon}^{tr}
\end{aligned} \tag{2.3}$$

or

$$\boldsymbol{\Upsilon} = \{\xi, \nabla\xi, \nabla\nabla\xi, \cdots, \nabla^p\xi, \boldsymbol{\varepsilon}^{tr}, \nabla\boldsymbol{\varepsilon}^{tr}, \cdots, \nabla^q\boldsymbol{\varepsilon}^{tr}\} \tag{2.4}$$

By considering the spatial gradients of martensite volume fraction, it is possible to extract information from the microscopic non-homogeneous microstructure of the SMA material to enhance the continuum macroscopic representation of its behavior.

A vital assumption here is the independence of $\boldsymbol{\Upsilon}_i$; whether or not they are related through spatial differentiation.

2.1 Kinematics

The material body is assumed to occupy a region , Ω, of the 3D Euclidean space, $\mathcal{E}$ (closure of an open and simply connected region, to be precise), with a boundary $\partial\Omega$. A configuration of this material body, $\mathcal{B}$, can be considered the reference configuration with points in the body $\mathbf{p} \in \mathcal{B}$. A deformed configuration of this body is defined through a one-to-one and smooth mapping, $\mathbf{f}$, such that:

$$\mathbf{f}_t(\mathcal{B}) = \{\mathbf{x} \in \mathcal{E} |\ \mathbf{x} = \mathbf{f}(\mathbf{X}, t), \quad \mathbf{X} \in \mathcal{B}\} \tag{2.5}$$

Hence, the displacement of a material point can be defined via

$$\mathbf{u}(\mathbf{X}, t) = \mathbf{f}(\mathbf{X}, t) - \mathbf{X} \tag{2.6}$$

For infinitesimal gradients of displacement, the difference between the reference configuration and the deformed configuration becomes negligible. Based on this, the infinitesimal strain tensor, $\boldsymbol{\varepsilon}$, can be defined.

$$\boldsymbol{\varepsilon} = \frac{1}{2}\left[(\nabla\mathbf{u}) + (\nabla\mathbf{u})^T\right] \tag{2.7}$$

where the gradient can interchangeably be considered over the deformed or reference configurations.

2.2 Balance (Conservation) laws

The basic conservation laws of continuum mechanics are presented in this section. First, an integral form is introduced and then the local form (as a differential equation) is presented. The *localization* procedure, in going form the integral to the PDE form, assumes a continuous and smooth function for all of the fields integrated.

2.2.1 Principle of conservation of mass

The total mass of a continuum body cannot change with time or deformation.

$$\frac{D}{Dt}\left(\int_{\Omega}\rho \mathrm{d}V\right)=0 \tag{2.8}$$

Here $\frac{D}{Dt}$ denotes material (or Lagrangian) time derivative and V is the volume of the body with density ρ. The local form of this principle is given by:

$$\frac{\partial \rho}{\partial t}+\nabla\cdot(\rho\mathbf{v})=0 \tag{2.9}$$

where $\mathbf{v}=\dot{\mathbf{u}}$ is the velocity of the material point with $(\dot{})$ representing material time derivative.

2.2.2 Principle of conservation of linear momentum

The rate of change of linear momentum for a continuum body is equal to the total sum of the surface and body forces applied to it.

$$\frac{D}{Dt}\left(\int_{\Omega}\rho\mathbf{v}\mathrm{d}V\right)=\int_{\partial\Omega}\boldsymbol{t}\mathrm{d}S+\int_{\Omega}\boldsymbol{b}\mathrm{d}V \tag{2.10}$$

in which $\boldsymbol{b}$ is the body force per unit volume. The surface forces are applied through surface traction $\boldsymbol{t}$. The local form of this relation is given by:

$$\nabla\cdot\boldsymbol{\sigma}+\boldsymbol{b}=\rho\dot{\mathbf{v}} \tag{2.11}$$

$\boldsymbol{\sigma}$ is the symmetric Cauchy stress tensor which is related to surface traction operating on a surface with a unit normal $\boldsymbol{n}$ as $\boldsymbol{\sigma n}=\boldsymbol{t}$.

2.2.3 Principle of conservation of angular momentum

Similarly, the rate of change of angular momentum in a continuum body should be equal to sum of the moments applied to it through the surface traction and body forces.

$$\frac{D}{Dt}\left(\int_{\Omega} \boldsymbol{r} \times \rho \mathbf{v}\ \mathrm{d}V\right) = \int_{\partial\Omega} \boldsymbol{r} \times \boldsymbol{t}\mathrm{d}S + \int_{\Omega} \boldsymbol{r} \times \boldsymbol{b}\mathrm{d}V \tag{2.12}$$

In the local form, this principle gives the symmetry of the the stress tensor, $\boldsymbol{\sigma} = \boldsymbol{\sigma}^T$.

2.2.4 Generalized principle of conservation of energy

This principle, the first law of thermodynamics, states that the rate of change of kinetic and internal energy for a continuum body is equal to the rate at which external mechanical work is exerted on that body as well as external thermal energy added/subtracted to it. The external work is performed partly due to the contact and body forces. In addition, it is assumed that the the generalized state variables are capable of performing work through generalized surface and body forces. No physical justification is given for it here, however, one can argue that since the generalized state variables contribute to the internal energy, an avenue must be opened for them through external work in the balance of energy.

$$\begin{aligned}\frac{D}{Dt}\left(\int_{\Omega} \frac{1}{2}\rho \mathbf{v} \cdot \mathbf{v}\ \mathrm{d}V + \int_{\Omega} \rho \mathrm{u}\ \mathrm{d}V\right) = &\int_{\partial\Omega} \boldsymbol{t} \cdot \mathbf{v}\mathrm{d}S + \int_{\Omega} \boldsymbol{b} \cdot \mathbf{v}\mathrm{d}V + \int_{\partial\Omega} -\boldsymbol{q} \cdot \boldsymbol{n}\mathrm{d}S + \int_{\Omega} \rho r \mathrm{d}V \\ &+ \int_{\partial\Omega} (\boldsymbol{\mu}_i^s \cdot \boldsymbol{n}) \cdot \dot{\boldsymbol{\Upsilon}}_i\ \mathrm{d}S + \int_{\Omega} \boldsymbol{\mu}_i^b \cdot \dot{\boldsymbol{\Upsilon}}_i\ \mathrm{d}V\end{aligned} \tag{2.13}$$

in this equation u is the specific internal energy, $\boldsymbol{q}$ is the heat flux vector acting

on the surface of the body and r is a body heat source. $\boldsymbol{\mu}_i^s$ and $\boldsymbol{\mu}_i^b$ are the generalized surface and body forces, respectively.

In order to obtain the local form of this equation, the equilibrium equation (2.11) and the following are taken into account.

$$\begin{aligned} &\int_{\partial\Omega} \boldsymbol{q}\cdot\boldsymbol{n}\mathrm{d}S = \int_{\Omega} \mathrm{Div}\left(\boldsymbol{q}\right)\mathrm{d}V \\ &\int_{\partial\Omega} \left(\boldsymbol{\mu}_i^s\cdot\boldsymbol{n}\right)\cdot\dot{\boldsymbol{\Upsilon}}_i\,\mathrm{d}S = \int_{\partial\Omega}\left(\boldsymbol{\mu}_i^{s*}\cdot\dot{\boldsymbol{\Upsilon}}_i\right)\cdot\boldsymbol{n}\,\mathrm{d}S = \int_{\Omega}\mathrm{Div}\left(\boldsymbol{\mu}_i^{s*}\cdot\dot{\boldsymbol{\Upsilon}}_i\right)\mathrm{d}V \end{aligned} \tag{2.14}$$

in which the divergence theorem is used assuming the required continuity and smoothness for the functions. $\boldsymbol{\mu}_i^{s*}$ is the adjoint of $\boldsymbol{\mu}_i^s$ that satisfies $\left(\boldsymbol{\mu}_i^{s*}\cdot\dot{\boldsymbol{\Upsilon}}_i\right)\cdot\boldsymbol{n} = \left(\boldsymbol{\mu}_i^s\cdot\boldsymbol{n}\right)\cdot\dot{\boldsymbol{\Upsilon}}_i$. Notice here that $\boldsymbol{\mu}_i^{s*}\cdot\dot{\boldsymbol{\Upsilon}}_i$ results in vectorial identities and hence $\mathrm{Div}\left(\boldsymbol{\mu}_i^{s*}\cdot\dot{\boldsymbol{\Upsilon}}_i\right)$ will be scalar quantities.

On the other hand, according to the definitions and identities given in the appendix B.3,

$$\mathrm{Div}\left(\boldsymbol{\mu}_i^{s*}\cdot\dot{\boldsymbol{\Upsilon}}_i\right) = \boldsymbol{\mu}_i^s\cdot\nabla\dot{\boldsymbol{\Upsilon}}_i + \mathrm{Div}\left(\boldsymbol{\mu}_i^s\right)\cdot\dot{\boldsymbol{\Upsilon}}_i \tag{2.15}$$

Therefore in the local form, the generalized principle of conservation of energy reduces to:

$$\rho\dot{u} = \boldsymbol{\sigma}:\dot{\boldsymbol{\varepsilon}} - \nabla\cdot\boldsymbol{q} + \rho r + \boldsymbol{\mu}_i^s\cdot\nabla\dot{\boldsymbol{\Upsilon}}_i + \left[\mathrm{Div}\left(\boldsymbol{\mu}_i^s\right) + \boldsymbol{\mu}_i^b\right]\cdot\dot{\boldsymbol{\Upsilon}}_i \tag{2.16}$$

For the special case given in equation (2.4) with $p = q = 1$, or

$$\begin{aligned}
\boldsymbol{\Upsilon} &\equiv \{\xi,\, \boldsymbol{\varepsilon}^{tr},\, \nabla\xi,\, \nabla\boldsymbol{\varepsilon}^{tr}\} \\
\boldsymbol{\mu}^s &\equiv \{\boldsymbol{\mu}_1^s,\, \boldsymbol{\mu}_2^s,\, \boldsymbol{\mu}_3^s,\, \boldsymbol{\mu}_4^s\} \\
\boldsymbol{\mu}^b &\equiv \{\mu_1^b,\, \boldsymbol{\mu}_2^b,\, \boldsymbol{\mu}_3^b,\, \boldsymbol{\mu}_4^b\}
\end{aligned} \tag{2.17}$$

the generalized principle of conservation of energy will be

$$\begin{aligned}
&\frac{D}{Dt}\left(\int_\Omega \frac{1}{2}\rho\mathbf{v}\cdot\mathbf{v}\,\mathrm{d}V + \int_\Omega \rho\mathrm{u}\,\mathrm{d}V\right) = \int_{\partial\Omega} \boldsymbol{t}\cdot\mathbf{v}\mathrm{d}S + \int_\Omega \boldsymbol{b}\cdot\mathbf{v}\mathrm{d}V + \int_{\partial\Omega} -\boldsymbol{q}\cdot\boldsymbol{n}\mathrm{d}S + \int_\Omega \rho r\mathrm{d}V \\
&+ \int_{\partial\Omega} (\boldsymbol{\mu}_1^s\cdot\boldsymbol{n})\,\dot{\xi} + (\boldsymbol{\mu}_2^s\cdot\boldsymbol{n}) : \dot{\boldsymbol{\varepsilon}}^{tr} + (\boldsymbol{\mu}_3^s\boldsymbol{n})\cdot\nabla\dot{\xi} + (\boldsymbol{\mu}_4^s\cdot\boldsymbol{n})\cdot\nabla\dot{\boldsymbol{\varepsilon}}^{tr}\,\mathrm{d}S \\
&+ \int_\Omega \mu_1^b\,\dot{\xi} + \boldsymbol{\mu}_2^b : \dot{\boldsymbol{\varepsilon}}^{tr} + \boldsymbol{\mu}_3^b\cdot\nabla\dot{\xi} + \boldsymbol{\mu}_4^b\cdot\nabla\dot{\boldsymbol{\varepsilon}}^{tr}\,\mathrm{d}V
\end{aligned} \tag{2.18}$$

in the integral form and

$$\begin{aligned}
&\rho\dot{u} = \boldsymbol{\sigma} : \dot{\boldsymbol{\varepsilon}} - \nabla\cdot\boldsymbol{q} + \rho r \\
&+ \left[\mathrm{Div}\,(\boldsymbol{\mu}_1^s) + \boldsymbol{\mu}_1^b\right]\dot{\xi} + \left[\mathrm{Div}\,(\boldsymbol{\mu}_2^s) + \boldsymbol{\mu}_2^b\right] : \dot{\boldsymbol{\varepsilon}}^{tr} + \left[\mathrm{Div}\,(\boldsymbol{\mu}_3^s) + \boldsymbol{\mu}_3^b + \boldsymbol{\mu}_1^s\right]\cdot\nabla\dot{\xi} \\
&+ \left[\mathrm{Div}\,(\boldsymbol{\mu}_4^s) + \boldsymbol{\mu}_4^b + \boldsymbol{\mu}_2^s\right]\cdot\nabla\dot{\boldsymbol{\varepsilon}}^{tr} + \boldsymbol{\mu}_3^s : \nabla\nabla\dot{\xi} + \boldsymbol{\mu}_4^s\cdot\nabla\nabla\dot{\boldsymbol{\varepsilon}}^{tr}
\end{aligned} \tag{2.19}$$

in the local form.

2.2.5 Entropy inequality

This inequality principle represents the second law of thermodynamics as the rate of entropy production in a continuum body is always greater than or equal to the external entropy supplied to it. The second law of thermodynamics in the Clausius-Duhem form is given as:

$$\frac{D}{Dt}\left(\int_{\Omega}\rho s\mathrm{d}V\right) \geq \int_{\partial\Omega}\frac{-\boldsymbol{q}}{T}\cdot\boldsymbol{n}\mathrm{d}S + \int_{\Omega}\frac{\rho r}{T}\mathrm{d}V \tag{2.20}$$

with s being the specific entropy. In its local form, the Clausius-Duhem inequality is given as:

$$\rho\dot{s} \geq -\nabla\cdot\left(\frac{\boldsymbol{q}}{T}\right) + \frac{\rho r}{T} \tag{2.21}$$

The second law of thermodynamics, in principle, places restrictions on the nature of energy conversion, making certain states inaccessible through any thermomechanical process. The constitutive relations developed between the state variables must not violate the second law of thermodynamics.

2.3 Constitutive equations

The problem of solving for the field variables for a continuum body is indeterminate without considering constitutive equations. They define a relation between the dependent field variables, or response functions, and the independent ones; equations (2.1). Thermodynamic sate variables are the variables that completely characterize the sate of a material body. If the variable is *measurable* and *controllable*, for example through a surface action such as an imposed traction, it is called an *external* sate variable. Internal variable of state (ISV), on the other hand, is *identifiable* and *measurable*, but it can not be linked to any external-force variable that, similar to a body force or surface traction, may provide a means of control. Therefore, these new ISVs do not perform work, i.e. they do not appear a priori in the mechanical work statement in the first law of thermodynamics [100]. The internal variables representing such microstructural or microscopic changes as martensitic transformation or crystal dislocations can be *observed* through proper experimental

apparatus such as in-situ TEM, however, cannot be independently controlled. The constitutive equations, hence, give a relationship between the value of the dependent state variables at time t, and the independent external state variables and the ISVs. Such constitutive equations must be accompanied by *evolution equation* laws that describe the evolution of the internal variables.

The generalized state variables $\boldsymbol{\Upsilon}$ defined at the beginning of this chapter contribute to the internal energy thorough their surface and body work conjugates. According to the aforementioned definition, though, $\boldsymbol{\Upsilon}_i$ cannot be denoted as internal state variables; they are rather the *internal degrees of freedom* related to the microstructure.

In order to derive the constitutive equations, it is assumed that all the material points in the continuum body at any time, t, go through admissible thermodynamic processes that obey the 2$^{\text{nd}}$ law of thermodynamics.

The first and second laws of thermodynamics in equations 2.16 and 2.21 can be combined. Using the Legendre transform, 2.2, for G it is possible to obtain:

$$\begin{aligned} -\rho\dot{G} - \rho s\dot{T} - \boldsymbol{\varepsilon} : \dot{\boldsymbol{\sigma}} + \boldsymbol{\mu}_i^s \cdot \nabla\dot{\boldsymbol{\Upsilon}}_i + \left[\mathrm{Div}\left(\boldsymbol{\mu}_i^s\right) + \boldsymbol{\mu}_i^b\right] \cdot \dot{\boldsymbol{\Upsilon}}_i \geq \\ \nabla \cdot \boldsymbol{q} - T\nabla \cdot \left(\frac{\boldsymbol{q}}{T}\right) = \frac{1}{T}\, \boldsymbol{q} \cdot \nabla T \end{aligned} \tag{2.22}$$

If the classical linear theory of heat conduction, or the Fourier's law, is considered for the thermal constitutive response of the material body

$$\boldsymbol{q} = -\boldsymbol{K}\nabla T \tag{2.23}$$

where $\boldsymbol{K} = \boldsymbol{K}(T)$ is the positive-definite symmetric heat conduction tensor (T is also assumed to be positive), then equation (2.22) reduces to the Clausius-Plank

inequality, as another form for the second law of thermodynamics.

$$
\begin{aligned}
-\rho\dot{G} - \rho s\dot{T} - \boldsymbol{\varepsilon} : \dot{\boldsymbol{\sigma}} + \boldsymbol{\mu}_i^s \cdot \nabla\dot{\boldsymbol{\Upsilon}}_i + \left[\mathrm{Div}\left(\boldsymbol{\mu}_i^s\right) + \boldsymbol{\mu}_i^b\right] \cdot \dot{\boldsymbol{\Upsilon}}_i \geq 0 \\
-\left(\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} + \boldsymbol{\varepsilon}\right) : \dot{\boldsymbol{\sigma}} - \rho\left(\frac{\partial G}{\partial T} + s\right)\dot{T} + \boldsymbol{\mu}_i^s \cdot \nabla\dot{\boldsymbol{\Upsilon}}_i + \left[\mathrm{Div}\left(\boldsymbol{\mu}_i^s\right) + \boldsymbol{\mu}_i^b - \rho\frac{\partial G}{\partial \boldsymbol{\Upsilon}_i}\right] \cdot \dot{\boldsymbol{\Upsilon}}_i \geq 0
\end{aligned}
\tag{2.24}
$$

This equation shows that by assigning surface work to some generalized state variables, $\boldsymbol{\Upsilon}$, the gradients $\nabla\boldsymbol{\Upsilon}$ appear in the rate of dissipation of energy.

If the set of generalized variables, $\boldsymbol{\Upsilon}$ as given in equation (2.17) is used, then we will have

$$
\begin{aligned}
&-\left(\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} + \boldsymbol{\varepsilon}\right) : \dot{\boldsymbol{\sigma}} - \rho\left(\frac{\partial G}{\partial T} + s\right)\dot{T} \\
&+ \left[\mathrm{Div}\left(\boldsymbol{\mu}_1^s\right) + \mu_1^b - \rho\frac{\partial G}{\partial \xi}\right]\dot{\xi} + \left[\mathrm{Div}\left(\boldsymbol{\mu}_2^s\right) + \boldsymbol{\mu}_2^b - \rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}}\right] : \dot{\boldsymbol{\varepsilon}}^{tr} \\
&+ \left[\mathrm{Div}\left(\boldsymbol{\mu}_3^s\right) + \boldsymbol{\mu}_3^b + \boldsymbol{\mu}_1^s - \rho\frac{\partial G}{\partial \nabla\xi}\right] \cdot \nabla\dot{\xi} + \left[\mathrm{Div}\left(\boldsymbol{\mu}_4^s\right) + \boldsymbol{\mu}_4^b + \boldsymbol{\mu}_2^s - \rho\frac{\partial G}{\partial \nabla\boldsymbol{\varepsilon}^{tr}}\right] \cdot \nabla\dot{\boldsymbol{\varepsilon}^{tr}} \\
&+ \boldsymbol{\mu}_3^s : \nabla\nabla\dot{\xi} + \boldsymbol{\mu}_4^s \cdot \nabla\nabla\dot{\boldsymbol{\varepsilon}^{tr}} \geq 0
\end{aligned}
\tag{2.25}
$$

The second law of thermodynamics, or the Clausius-Plank inequality (2.25), is sufficiently satisfied by considering the following constitutive equations.

$$
\boldsymbol{\varepsilon} = -\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} \quad , \quad s = -\frac{\partial G}{\partial T} \quad , \quad D \geq 0 \tag{2.26}
$$

where

$$
\begin{aligned}
D = & \\
& \left[\mathrm{Div}\left(\boldsymbol{\mu}_1^s\right) + \mu_1^b - \rho\frac{\partial G}{\partial \xi}\right]\dot{\xi} + \left[\mathrm{Div}\left(\boldsymbol{\mu}_2^s\right) + \boldsymbol{\mu}_2^b - \rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}}\right] : \dot{\boldsymbol{\varepsilon}}^{tr} \\
& + \left[\mathrm{Div}\left(\boldsymbol{\mu}_3^s\right) + \boldsymbol{\mu}_3^b + \boldsymbol{\mu}_1^s - \rho\frac{\partial G}{\partial \nabla\xi}\right]\cdot\nabla\dot{\xi} + \left[\mathrm{Div}\left(\boldsymbol{\mu}_4^s\right) + \boldsymbol{\mu}_4^b + \boldsymbol{\mu}_2^s - \rho\frac{\partial G}{\partial \nabla\boldsymbol{\varepsilon}^{tr}}\right]\cdot\nabla\dot{\boldsymbol{\varepsilon}}^{tr} \\
& + \boldsymbol{\mu}_3^s : \nabla\nabla\dot{\xi} + \boldsymbol{\mu}_4^s \cdot \nabla\nabla\dot{\boldsymbol{\varepsilon}}^{tr}
\end{aligned}
\tag{2.27}
$$

Whether internal degrees of freedom or internal state variables, the above equations have to be complemented by evolution equations for the state variables $\dot{\boldsymbol{\Upsilon}} = \{\dot{\boldsymbol{\Upsilon}}_i\}$, as well as constitutive equations for $\boldsymbol{\mu}^s = \{\boldsymbol{\mu}_i^s\}$ and $\boldsymbol{\mu}^b = \{\boldsymbol{\mu}_i^b\}$, listed in (2.17).

In any case, it is possible to use the concept of generalized thermodynamic forces and fluxes. The contribution to the rate of dissipation of energy, D or dissipation for short, is considered to originate in the dissipative work of generalized thermodynamic forces, $\boldsymbol{\Gamma}$, and fluxes, $\dot{\boldsymbol{\Pi}}$.

In relation to equation (2.24)b, after including (2.26)a and b, one can write

$$
\begin{aligned}
& \boldsymbol{\Gamma} \equiv \{\left[\mathrm{Div}\left(\boldsymbol{\mu}_i^s\right) + \boldsymbol{\mu}_i^b - \rho\frac{\partial G}{\partial \boldsymbol{\Upsilon}_i}\right], \boldsymbol{\mu}_i^s\} \quad , \quad \boldsymbol{\Pi} \equiv \{\boldsymbol{\Upsilon}_i, \nabla\boldsymbol{\Upsilon}_i\} \\
& D = \boldsymbol{\Gamma}_1 \cdot \dot{\boldsymbol{\Pi}}_1 + \boldsymbol{\Gamma}_2 \cdot \dot{\boldsymbol{\Pi}}_2 = \boldsymbol{\Gamma} \cdot \dot{\boldsymbol{\Pi}} \geq 0
\end{aligned}
\tag{2.28}
$$

For the case of SMA response, by redefining the variables in equation (2.27), we will have

$$
\begin{aligned}
&\boldsymbol{\Gamma}_1 \equiv \mathrm{Div}\left(\boldsymbol{\mu}_1^s\right) + \mu_1^b - \rho\frac{\partial G}{\partial \xi} \quad , \quad \boldsymbol{\Pi}_1 \equiv \xi \\
&\boldsymbol{\Gamma}_2 \equiv \mathrm{Div}\left(\boldsymbol{\mu}_2^s\right) + \boldsymbol{\mu}_2^b - \rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} \quad , \quad \boldsymbol{\Pi}_2 \equiv \boldsymbol{\varepsilon}^{tr} \\
&\boldsymbol{\Gamma}_3 \equiv \mathrm{Div}\left(\boldsymbol{\mu}_3^s\right) + \boldsymbol{\mu}_3^b + \boldsymbol{\mu}_1^s - \rho\frac{\partial G}{\partial \nabla\xi} \quad , \quad \boldsymbol{\Pi}_3 \equiv \nabla\xi \\
&\boldsymbol{\Gamma}_4 \equiv \mathrm{Div}\left(\boldsymbol{\mu}_4^s\right) + \boldsymbol{\mu}_4^b + \boldsymbol{\mu}_2^s - \rho\frac{\partial G}{\partial \nabla\boldsymbol{\varepsilon}^{tr}} \quad , \quad \boldsymbol{\Pi}_4 \equiv \nabla\boldsymbol{\varepsilon}^{tr} \\
&\boldsymbol{\Gamma}_5 \equiv \boldsymbol{\mu}_3^s \quad , \quad \boldsymbol{\Pi}_5 \equiv \nabla\nabla\xi \\
&\boldsymbol{\Gamma}_6 \equiv \boldsymbol{\mu}_4^s \quad , \quad \boldsymbol{\Pi}_6 \equiv \nabla\nabla\boldsymbol{\varepsilon}^{tr} \\
&\boldsymbol{\Gamma} \equiv \{\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \boldsymbol{\Gamma}_3, \boldsymbol{\Gamma}_4, \boldsymbol{\Gamma}_5, \boldsymbol{\Gamma}_6\} \quad , \quad \boldsymbol{\Pi} \equiv \{\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2, \boldsymbol{\Pi}_3, \boldsymbol{\Pi}_4, \boldsymbol{\Pi}_5, \boldsymbol{\Pi}_6\} \\
&D = \boldsymbol{\Gamma}_i \cdot \dot{\boldsymbol{\Pi}}_i \geq 0
\end{aligned}
\tag{2.29}
$$

A common approach to achieve a rate-independent response is to define a threshold for $\boldsymbol{\Gamma}$ at which the dissipative mechanism activates. This can be done by assuming the existence of a convex set K such that

$$
\forall\, \boldsymbol{\Gamma}^* \in K \; : \quad D = \left(\boldsymbol{\Gamma}_i - \boldsymbol{\Gamma}_i^*\right) \cdot \dot{\boldsymbol{\Pi}}_i \geq 0 \tag{2.30}
$$

No dissipation occurs if $\boldsymbol{\Gamma}$ is inside K. Equation 2.30 is denoted as *Hill-Mandel's principle of maximum dissipation*. A hypersurface

$$
\Phi\left(\boldsymbol{\Gamma}, \boldsymbol{\Pi}\right) = \phi\left(\boldsymbol{\Gamma}\right) - g\left(\boldsymbol{\Pi}\right) \tag{2.31}
$$

can be defined in the space of generalized forces that defines the convex boundary of the convex set K. In the case of dislocation plasticity, $\Phi = 0$ is the well-known yield surface. Φ can also be associated to the dissipation potential the result of which is the normality in the space of generalized forces. To that end, the state of

generalized variables is considered to be the one that maximizes the dissipation D; or

$$\text{Max} \left\{ D = \boldsymbol{\Gamma}_i \cdot \dot{\boldsymbol{\Pi}}_i \mid \boldsymbol{\Gamma} \ , \ \Phi(\boldsymbol{\Gamma}, \boldsymbol{\Pi}) \leq 0 \right\} \tag{2.32}$$

Upon defining a Lagrangian, L, the minimization problem here is solved through the Lagrange method of multipliers.

$$L = -\boldsymbol{\Gamma}_i \cdot \dot{\boldsymbol{\Pi}}_i + \dot{\lambda} \, \Phi(\boldsymbol{\Gamma}, \boldsymbol{\Pi}) \tag{2.33}$$

It is shown that the problem in (2.32) is equivalent to finding the minimum of L or

$$\frac{\partial L}{\partial \boldsymbol{\Gamma}_i} = -\dot{\boldsymbol{\Pi}}_i + \dot{\lambda} \frac{\partial \Phi}{\partial \boldsymbol{\Gamma}_i} = 0 \tag{2.34}$$

on the condition that the following Kuhn-Tucker conditions are satisfied:

1. $\Phi(\boldsymbol{\Gamma}, \boldsymbol{\Pi})$ is convex in $\boldsymbol{\Gamma}$ and $\Phi \leq 0$
2. $\dot{\lambda} \geq 0$
3. $\dot{\lambda} \Phi(\boldsymbol{\Gamma}) = 0$

Therefore, the rates of the generalized fluxes can be found via

$$\dot{\boldsymbol{\Pi}}_i = \dot{\lambda} \frac{\partial \Phi}{\partial \boldsymbol{\Gamma}_i} = \dot{\lambda} \frac{\partial \phi}{\partial \boldsymbol{\Gamma}_i} \tag{2.35}$$

demonstrating the associativity of the response in the space of generalized forces. The existence of time derivative on both sides does not imply any rate-dependent

response as it can be viewed as differential with respect to an evolution or loading parameter. A discussion is, also, provided in the Appendix A.2 for various coupling options possible in forming the surface Φ.

The rate of the Lagrange multiplier can be determined from the consistency condition.

$$\dot{\Phi} = \dot{\phi} - \dot{g} = 0 \quad , \quad \frac{\partial \phi}{\partial \mathbf{\Gamma}_i} \cdot \dot{\mathbf{\Gamma}}_i + \dot{\lambda} \left(\frac{\partial g}{\partial \mathbf{\Pi}_j} \cdot \frac{\partial \phi}{\partial \mathbf{\Gamma}_j} \right) = 0 \tag{2.36}$$

It is customary for $\phi(\mathbf{\Gamma})$ to be assumed a homogeneous function of degree k. In that case the dissipation, (2.29c) satisfying the second law of thermodynamics, can be obtained from:

$$D = \mathbf{\Gamma} \cdot \dot{\mathbf{\Pi}} = \dot{\lambda} \left(\mathbf{\Gamma}_i \cdot \frac{\partial \phi}{\partial \mathbf{\Gamma}_i} \right) = k \dot{\lambda} \phi = k \dot{\lambda} g \geq 0 \tag{2.37}$$

The form of the function Φ along with the presumed generalized state variables can result in various types of partial differential equations needed to be solved to obtain the generalized state variables.

2.3.1 Connection with the gradient theories based on the microforce balance law

As discussed in section 1.3, several nonlocal gradient-based theories of plasticity postulate an additional balance law that incorporates the *configurational* or *microscopic* forces associated with the effective or accumulated plastic strain [71, 94]. This *microscopic force balance* or *microforce balance* law [69] can be obtained form a generalized form of the principle of virtual power. In that, it is assumed that *microtractions* act as work conjugates to the generalized variable of the accumulated plastic strain and contribute to the external power. Also, internal power contribution is assumed for the accumulated plastic strain and its gradient.

As an example, equation (3.9) from Gurtin and Anand [71] is rewritten here.

$$\begin{aligned} \mathcal{W}_{ext} &= \int_{\partial P} (\boldsymbol{t}(\boldsymbol{n}) \cdot \dot{\boldsymbol{u}} + \chi(\boldsymbol{n})\dot{\gamma}^p)\, \mathrm{d}A + \int_P \boldsymbol{b} \cdot \dot{\boldsymbol{u}}\, \mathrm{d}V \\ \mathcal{W}_{int} &= \int_P \left(\boldsymbol{T} : \dot{\boldsymbol{E}}^{el} + \pi\dot{\gamma}^p + \boldsymbol{\xi} \cdot \nabla\dot{\gamma}^p \right) \mathrm{d}V \end{aligned} \tag{2.38}$$

Following Gurtin and Anand's notation, the scalar microstress π is the power-conjugate to the accumulated plastic strain $\dot{\gamma}^p = \|\dot{\boldsymbol{E}}^p\|$ and $\boldsymbol{\xi}$ a vector microstress conjugate to $\nabla\dot{\gamma}^p$. $\chi(\boldsymbol{n})$ is the surface microtraction. It is, then, assumed that $\mathcal{W}_{int} = \mathcal{W}_{ext}$ for any arbitrary subregion P of the body and for any consistent virtual velocities $\mathcal{V} = \left(\dot{\boldsymbol{u}}, \dot{\boldsymbol{E}}^{el}, \dot{\gamma}^p \right)$.

The generalized principle of virtual power results in the conventional and micro-force balance laws

$$\mathrm{Div}(\boldsymbol{T}) + \boldsymbol{b} = \boldsymbol{0} \quad , \quad \tau - \pi + \mathrm{Div}(\boldsymbol{\xi}) = 0 \tag{2.39}$$

as well as standard traction and microtraction boundary conditions

$$\boldsymbol{t}(\boldsymbol{n}) = \boldsymbol{T}\boldsymbol{n} \quad , \quad \chi(\boldsymbol{n}) = \boldsymbol{\xi} \cdot \boldsymbol{n} \tag{2.40}$$

τ is the (Von-Mises) equivalent stress. Furthermore, the dissipation inequality as given in equation (6.13) of [71] is

$$D = \pi_{dis}\dot{\gamma}^p + \boldsymbol{\xi}_{dis} \cdot \nabla\dot{\gamma}^p \geq 0 \quad , \quad \pi_{dis} = \pi - \frac{\partial\psi}{\partial\gamma^p} \quad , \quad \boldsymbol{\xi}_{dis} = \boldsymbol{\xi} - \frac{\partial\psi}{\partial\nabla\gamma^p} \tag{2.41}$$

where $\psi = \hat{\psi}\left(\boldsymbol{E}^{el},\, \gamma^p,\, \nabla\gamma^p\right)$ is Helmholtz free energy. By replacing π and $\boldsymbol{\xi}$ in

the microforce balance (2.39)b, we will have

$$\tau - \pi_{dis} - \frac{\partial \psi}{\partial \gamma^p} + \text{Div}\left(\boldsymbol{\xi}_{dis} + \frac{\partial \psi}{\partial \nabla \gamma^p}\right) = 0 \tag{2.42}$$

which is equation (6.15) in [71]. The postulated microforce balance law in conjunction with constitutive equations for π_{dis} and $\boldsymbol{\xi}_{dis}$ effectively act as the flow rule for the gradient plasticity theory.

It will be shown here that this equation can be obtained following the general model established in the previous section. This, however, is going to be performed for the shape memory response; i.e. by replacing the accumulated plastic strain with the martensite volume fraction.

To that end, it is assumed that the generalized variables are

$$\begin{aligned} \boldsymbol{\Upsilon} &\equiv \{\xi,\, \boldsymbol{\varepsilon}^{tr},\, \nabla\xi\} \\ \boldsymbol{\mu}^s &\equiv \{\boldsymbol{\mu}_1^s\} \\ \boldsymbol{\mu}^b &\equiv \{\mu_1^b,\, \boldsymbol{\mu}_3^b\} \end{aligned} \tag{2.43}$$

$\boldsymbol{\mu}_1^s$, is the surface power conjugate to ξ. Also, μ_1^b and $\boldsymbol{\mu}_3^b$ are, respectively, the body power conjugates to ξ and $\nabla\xi$ in the generalized first law of thermodynamics (2.18). The dissipation in equation (2.27) reduces to

$$D = \left[\text{Div}\,(\boldsymbol{\mu}_1^s) + \mu_1^b - \rho\frac{\partial G}{\partial \xi}\right]\dot{\xi} + \left[-\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}}\right] : \dot{\boldsymbol{\varepsilon}}^{tr} + \left[\boldsymbol{\mu}_3^b + \boldsymbol{\mu}_1^s - \rho\frac{\partial G}{\partial \nabla\xi}\right]\cdot \nabla\dot{\xi} \geq 0 \tag{2.44}$$

Hence, the list of generalized forces and fluxes turn into

$$
\begin{aligned}
&\boldsymbol{\Gamma}_1 \equiv \mathrm{Div}\left(\boldsymbol{\mu}_1^s\right) + \mu_1^b - \rho\frac{\partial G}{\partial \xi} \quad , \quad \boldsymbol{\Pi}_1 \equiv \xi \\
&\boldsymbol{\Gamma}_2 \equiv -\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} \quad , \quad \boldsymbol{\Pi}_2 \equiv \boldsymbol{\varepsilon}^{tr} \\
&\boldsymbol{\Gamma}_3 \equiv \boldsymbol{\mu}_3^b + \boldsymbol{\mu}_1^s - \rho\frac{\partial G}{\partial \nabla\xi} \quad , \quad \boldsymbol{\Pi}_3 \equiv \nabla\xi \\
&\boldsymbol{\Gamma} \equiv \{\boldsymbol{\Gamma}_1, \boldsymbol{\Gamma}_2, \boldsymbol{\Gamma}_3\} \quad , \quad \boldsymbol{\Pi} \equiv \{\boldsymbol{\Pi}_1, \boldsymbol{\Pi}_2, \boldsymbol{\Pi}_3\} \\
&D = \boldsymbol{\Gamma}_i \cdot \dot{\boldsymbol{\Pi}}_i \geq 0
\end{aligned}
\tag{2.45}
$$

At this point, constitutive assumptions are considered for the generalized forces and fluxes. Firstly, it is assumed that $\boldsymbol{\Gamma}_3 \equiv \mathbf{0}$, therefore

$$
\boldsymbol{\mu}_1^s = \rho\frac{\partial G}{\partial \nabla\xi} - \boldsymbol{\mu}_3^b \tag{2.46}
$$

In addition, the transformation surface in equation (2.31) is considered to have the form

$$
\Phi\left(\boldsymbol{\Gamma}, \boldsymbol{\Pi}\right) = \phi\left(\boldsymbol{\Gamma}\right) - g\left(\boldsymbol{\Pi}\right) = \sqrt{\boldsymbol{\Gamma}_2 : \boldsymbol{\Gamma}_2} + \boldsymbol{\Gamma}_1 - Y = 0 \tag{2.47}
$$

or

$$
\Phi = \sqrt{\boldsymbol{\Gamma}_2 : \boldsymbol{\Gamma}_2} + \mathrm{Div}\left(\rho\frac{\partial G}{\partial \nabla\xi} - \boldsymbol{\mu}_3^b\right) + \mu_1^b - \rho\frac{\partial G}{\partial \xi} - Y = 0 \tag{2.48}
$$

Now, if the form of the Gibbs free energy is such that

$$
\boldsymbol{\Gamma}_2 \equiv -\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} = \boldsymbol{\sigma} \tag{2.49}
$$

then

$$\Phi = \tau + \text{Div}\left(\rho\frac{\partial G}{\partial \nabla\xi} - \boldsymbol{\mu}_3^b\right) - \rho\frac{\partial G}{\partial \xi} + \mu_1^b - Y = 0 \tag{2.50}$$

equivalent to the microforce balance in equation (2.42) augmented with the constitutive relations $\pi_{dis} \equiv Y - \mu_1^b$ and $\boldsymbol{\xi}_{dis} \equiv -\boldsymbol{\mu}_3^b$.

2.3.2 Reformulation of the gradient theory with internal variables

In order to capture the microstructural changes in an SMA material body due to the solid-state martensitic phase transformation, the approach of thermodynamics with internal variables is chosen. According to the discussion in section 2.3, the internal variables based on their definition can not perform work through surface or body forces. The internal variable theory is recovered by dismissing the generalized forces, $\boldsymbol{\mu}^s$ and $\boldsymbol{\mu}^b$, or *equivalently* by setting them to zero in all of the previous relations such as in equation (2.24).

$$\boldsymbol{\varepsilon} = -\rho\frac{\partial G}{\partial \boldsymbol{\sigma}}\,, \quad s = -\frac{\partial G}{\partial T}\,, \quad \boldsymbol{\mu}^s \equiv \mathbf{0}\,, \quad \boldsymbol{\mu}^b \equiv \mathbf{0}\,, \quad D = -\rho\frac{\partial G}{\partial \boldsymbol{\Upsilon}} \cdot \dot{\boldsymbol{\Upsilon}} \geq 0 \tag{2.51}$$

In this sense, observing

$$\boldsymbol{\Upsilon} \equiv \{\xi,\, \boldsymbol{\varepsilon}^{tr},\, \nabla\xi,\, \nabla\boldsymbol{\varepsilon}^{tr}\} \tag{2.52}$$

equation (2.27) reduces to

$$D = \left[-\rho\frac{\partial G}{\partial \xi}\right]\dot{\xi} + \left[-\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}}\right] : \dot{\boldsymbol{\varepsilon}}^{tr} + \left[-\rho\frac{\partial G}{\partial \nabla\xi}\right] \cdot \nabla\dot{\xi} + \left[-\rho\frac{\partial G}{\partial \nabla\boldsymbol{\varepsilon}^{tr}}\right] \cdot \nabla\dot{\boldsymbol{\varepsilon}}^{tr} \geq 0 \tag{2.53}$$

It is acknowledged that the equivalence of "non-existence" versus "being present with zero value" is a point of controversy [123]; nonetheless the final outcome is equivalent for the purpose of the current constitutive formulation.

The formulation of gradient-based constitutive models using the internal variable approach will be discussed in details in the ensuing chapters.

3. GRADIENT-BASED RATE-DEPENDENT SMA CONSTITUTIVE MODELING

In order to develop a rate-dependent nonlocal model for the response of shape memory alloys, the formulation for modeling of linear viscolestic solids is adopted. The thermodynamics with internal variables is followed in this section with $\boldsymbol{\Upsilon} = \{\boldsymbol{\varepsilon}^{tr}, \xi, \nabla\xi, \nabla\boldsymbol{\varepsilon}^{tr}\}$. ξ denotes the martensite volume fraction which is a measure of the extent of transformation from austenite to martensite phase; $0 \leq \xi \leq 1$. $\boldsymbol{\varepsilon}^{tr}$ is the transformation strain tensor which is the *apparent* inelastic strain as a result of that martensitic transformation. The spatial gradients of both these internal variables, $\nabla\xi$ and $\nabla\boldsymbol{\varepsilon}^{tr}$, are also introduced *independently* as internal variables. The constitutive equations (or complementary laws) are obtained as linear relationship between generalized thermodynamic forces and fluxes. To that end, the Clausius-Plank inequality in equation (2.24) is rewritten here, ignoring the generalized body and surface forces.

$$\begin{aligned} -\rho\dot{G} - \rho s\dot{T} - \boldsymbol{\varepsilon} : \dot{\boldsymbol{\sigma}} \geq 0 \\ -\left(\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} + \boldsymbol{\varepsilon}\right) : \dot{\boldsymbol{\sigma}} - \rho\left(\frac{\partial G}{\partial T} + s\right)\dot{T} \\ -\rho\left(\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} : \dot{\boldsymbol{\varepsilon}}^{tr} + \frac{\partial G}{\partial \xi}\dot{\xi} + \frac{\partial G}{\partial \nabla\xi}\cdot\nabla\dot{\xi} + \frac{\partial G}{\partial \nabla\boldsymbol{\varepsilon}^{tr}}\cdot\nabla\dot{\boldsymbol{\varepsilon}}^{tr}\right) \geq 0 \end{aligned} \tag{3.1}$$

which is sufficiently satisfied by the following constitutive equations.

$$\begin{aligned} \boldsymbol{\varepsilon} &= -\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} \quad , \quad s = -\frac{\partial G}{\partial T} \\ D^{tr} &= -\rho\left(\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} : \dot{\boldsymbol{\varepsilon}}^{tr} + \frac{\partial G}{\partial \xi}\dot{\xi} + \frac{\partial G}{\partial \nabla\xi}\cdot\nabla\dot{\xi} + \frac{\partial G}{\partial \nabla\boldsymbol{\varepsilon}^{tr}}\cdot\nabla\dot{\boldsymbol{\varepsilon}^{tr}}\right) \geq 0 \end{aligned} \tag{3.2}$$

D^{tr} in relation (3.2)c is the rate of energy dissipation due to the microstructural changes subsequent to martensitic transformation. This can arise from formation of martensite twinning and variant boundaries and/or the propagation of the transformation front. Based on this, it is possible to define the rate of evolution of the internal variables, $\dot{\boldsymbol{\Upsilon}}$, as the generalized thermodynamic fluxes and attribute corresponding thermodynamic forces, $\boldsymbol{\Gamma}$, to them. Hence,

$$\dot{\boldsymbol{\Upsilon}} = \{\dot{\boldsymbol{\varepsilon}}^{tr},\ \dot{\xi},\ \nabla\dot{\xi},\ \nabla\dot{\boldsymbol{\varepsilon}^{tr}}\} \quad , \quad \boldsymbol{\Gamma} = \{\boldsymbol{\varsigma}^{D},\ \pi,\ \bar{\boldsymbol{\pi}},\ \boldsymbol{\tau}\} \quad , \quad \boldsymbol{\Gamma} = -\rho\frac{\partial G}{\partial \boldsymbol{\Upsilon}} \tag{3.3}$$

or

$$\boldsymbol{\varsigma}^{D} = -\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} \quad , \quad \pi = -\rho\frac{\partial G}{\partial \xi} \quad , \quad \bar{\boldsymbol{\pi}} = -\rho\frac{\partial G}{\partial \nabla\xi} \quad , \quad \boldsymbol{\tau} = -\rho\frac{\partial G}{\partial \nabla\boldsymbol{\varepsilon}^{tr}} \tag{3.4}$$

To obtain the constitutive (phenomenological) equations, the generalized thermodynamic forces, in the neighborhood of equilibrium, are assumed to be a linear function of the rate of evolution of the internal variables or the fluxes.

$$D^{tr} = \boldsymbol{\Gamma}\cdot\dot{\boldsymbol{\Upsilon}} \quad , \quad \boldsymbol{\Gamma} = \boldsymbol{\mathcal{L}}\,\dot{\boldsymbol{\Upsilon}} \quad \text{or} \quad \Gamma_i = \mathcal{L}_{ij}\,\dot{\Upsilon}_j \quad , \quad \dot{\boldsymbol{\Upsilon}} = \boldsymbol{\mathcal{A}}\,\boldsymbol{\Gamma} \quad \text{or} \quad \dot{\Upsilon}_i = \mathcal{A}_{ij}\,\Gamma_j \tag{3.5}$$

with Onsager reciprocal relations [97]

$$\boldsymbol{\mathcal{L}} = \boldsymbol{\mathcal{L}}^T \quad \text{or} \quad \mathcal{L}_{ij} = \mathcal{L}_{ji} \quad , \quad \boldsymbol{\mathcal{A}} = \boldsymbol{\mathcal{A}}^T \quad \text{or} \quad \mathcal{A}_{ij} = \mathcal{A}_{ji} \tag{3.6}$$

It is shown that the Onsager reciprocal relations can be derived as a result of the principle of maximum dissipation (see section A.1).

Equation (3.5b) can be rewritten in terms of the individual thermodynamic forces.

$$\begin{aligned}
\boldsymbol{\varsigma}^D &= \boldsymbol{A}\dot{\boldsymbol{\varepsilon}}^{tr} + \boldsymbol{B}\dot{\xi} + \boldsymbol{C}\nabla\dot{\xi} + \boldsymbol{D}\nabla\dot{\boldsymbol{\varepsilon}}^{tr} \\
\pi &= \boldsymbol{E} : \dot{\boldsymbol{\varepsilon}}^{tr} + a\dot{\xi} + \boldsymbol{b}\cdot\nabla\dot{\xi} + \boldsymbol{F}\cdot\nabla\dot{\boldsymbol{\varepsilon}}^{tr} \\
\bar{\boldsymbol{\pi}} &= \boldsymbol{H}\dot{\boldsymbol{\varepsilon}}^{tr} + \boldsymbol{l}\dot{\xi} + \boldsymbol{M}\nabla\dot{\xi} + \boldsymbol{N}\nabla\dot{\boldsymbol{\varepsilon}}^{tr} \\
\boldsymbol{\tau} &= \boldsymbol{O}\dot{\boldsymbol{\varepsilon}}^{tr} + \boldsymbol{Q}\dot{\xi} + \boldsymbol{S}\nabla\dot{\xi} + \boldsymbol{V}\nabla\dot{\boldsymbol{\varepsilon}}^{tr}
\end{aligned}$$

$$\begin{aligned}
\varsigma^D_{ij} &= A_{ijkl}\,\dot{\varepsilon^{tr}}_{kl} + B_{ij}\,\dot{\xi} + C_{ijk}\,\dot{\xi}_{,k} + D_{ijklm}\,\dot{\varepsilon^{tr}}_{kl,m} \\
\pi &= E_{ij}\,\dot{\varepsilon^{tr}}_{ij} + a\,\dot{\xi} + b_i\,\dot{\xi}_{,i} + F_{ijk}\,\dot{\varepsilon^{tr}}_{ij,k} \\
\bar{\pi}_i &= H_{ijk}\,\dot{\varepsilon^{tr}}_{jk} + l_i\,\dot{\xi} + M_{ij}\,\dot{\xi}_{,j} + N_{ijkl}\,\dot{\varepsilon^{tr}}_{jk,l} \\
\tau_{ijk} &= O_{ijklm}\,\dot{\varepsilon^{tr}}_{lm} + Q_{ijk}\,\dot{\xi} + S_{ijkl}\,\dot{\xi}_{,l} + V_{ijklmn}\,\dot{\varepsilon^{tr}}_{lm,n}
\end{aligned} \tag{3.7}$$

The tensor for transformation strain, $\boldsymbol{\varepsilon}^{tr}$, is assumed to be symmetric (later it can be proved from (3.2)a and the decomposition of symmetric total strain tensor into elastic and transformation parts). This means the conjugate force, $\boldsymbol{\varsigma}^D$, to transformation strain must be a symmetric second order tensor as it is the derivative of the free energy with respect to $\boldsymbol{\varepsilon}^{tr}$. This is also the case for the force conjugate to the gradient of the transformation strain; $\tau_{ijk} = \tau_{jik}$. Therefore, the tonsorial constants in equation (3.7) must have the following symmetries.

$$
\begin{aligned}
A_{ijkl} = A_{jikl} = A_{ijlk}\ ,\ B_{ij} = B_{ji}\ ,\ C_{ijk} = C_{jik} \\
D_{ijklm} = D_{jiklm} = D_{ijlkm}\ ,\ E_{ij} = E_{ji}\ ,\ F_{ijk} = F_{jik} \\
H_{ijk} = H_{ikj}\ ,\ N_{ijkl} = N_{ikjl}\ ,\ Q_{ijk} = Q_{jik}\ ,\ S_{ijkl} = S_{jikl} \\
O_{ijklm} = O_{ijkml} = O_{jiklm}\ ,\ V_{ijklmn} = V_{ijkmln} = V_{jiklmn}
\end{aligned}
\tag{3.8}
$$

In addition, due to the Onsager reciprocity property stated in equation (3.6), the subsequent symmetries ought to be taken into account.

$$
\begin{aligned}
B_{ij} = E_{ij}\ ,\ C_{ijk} = H_{ijk}\ ,\ D_{ijklm} = O_{ijklm} \\
l_i = b_i\ ,\ F_{ijk} = Q_{ijk}\ ,\ N_{ijkl} = S_{ijkl}
\end{aligned}
\tag{3.9}
$$

The symmetry observed in the material response places another restriction on the phenomenological equations in (3.7). That is, for instance and without loss of generality, the conjugate force to the transformation strain, $\boldsymbol{\varsigma}^D$, must remain the same if the direction of transformation strain changes according to the symmetry of the material.

$$
\begin{aligned}
&\text{for} \quad \varsigma^D_{mn} = F\left(\varepsilon^{tr}_{ij},\, \xi_{,i},\, \varepsilon^{tr}_{ij,k}\right)\ : \\
\forall\, \boldsymbol{Q} \in \mathcal{S} \quad &, \quad Q_{im}\, Q_{jn}\, \varsigma^D_{ij} = F\left(Q_{ki}\, Q_{lj}\, \varepsilon^{tr}_{kl},\ Q_{li}\, \xi_{,l},\ Q_{pi}\, Q_{qj}\, Q_{rk}\, \varepsilon^{tr}_{pq,r}\right)
\end{aligned}
\tag{3.10}
$$

where $\mathcal{S}$ is the symmetry group of the material and $\boldsymbol{Q}$ is an orthogonal transformation with $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{I}$ and $\det(\boldsymbol{Q}) = +1$. For the linear relationship given in (3.7)a, we will have:

$$\text{for} \quad \varsigma^D_{mn} = A_{mnij}\,\varepsilon^{tr}_{ij} \;:$$

$$Q_{im}\,Q_{jn}\,\varsigma^D_{ij} = A_{mnij}\,Q_{ki}\,Q_{lj}\,\varepsilon^{tr}_{kl} \Rightarrow \varsigma^D_{mn} = Q_{pm}\,Q_{pn}\,Q_{ri}\,Q_{sj}\,A_{pqrs}\,\varepsilon^{tr}_{ij} \tag{3.11}$$

$$A_{mnij} = Q_{pm}\,Q_{pn}\,Q_{ri}\,Q_{sj}\,A_{pqrs}$$

which means that A_{mnij} must be invariant under any orthogonal transformation belonging to the symmetry group of the material.

In order to proceed the development in this section, an isotropic response is assumed for the shape memory alloy modeled. This means that $\mathcal{S} = \mathrm{Orth}^+$, the group of proper orthogonal transformations with $\boldsymbol{Q} \in \mathrm{Orth}^+ \rightarrow \det(\boldsymbol{Q}) = +1$.

Therefore, the tensorial constants of equation (3.7) have to be isotropic tensors. According to the results of B.1, the odd-ranked tensors should be eliminated for the case of isotropy. Therefore, equation (3.7) can be reduced to:

$$\begin{aligned}
\varsigma^D_{ij} &= A_{ijkl}\,\dot{\varepsilon^{tr}}_{kl} + B_{ij}\,\dot{\xi} \\
\pi &= B_{ij}\,\dot{\varepsilon^{tr}}_{ij} + a\,\dot{\xi} \\
\bar{\pi}_i &= M_{ij}\,\dot{\xi}_{,j} + N_{ijkl}\,\dot{\varepsilon^{tr}}_{jk,l} \\
\tau_{ijk} &= N_{ijkl}\,\dot{\xi}_{,l} + V_{ijklmn}\,\dot{\varepsilon^{tr}}_{lm,n}
\end{aligned} \tag{3.12}$$

Isotropic tensors of up to rank 6 are given in table B.2, based on which, the tensor constants in equation (3.12) can be written. Also the symmetry requirements given in equations (3.8) and (3.9) have to be considered. The results are given in the following:

$$
\begin{aligned}
B_{ij} &= B\,\delta_{ij} \quad , \quad E_{ij} = E\,\delta_{ij} \quad , \quad M_{ij} = M\,\delta_{ij} \\
A_{ijkl} &= A_1\,\delta_{ij}\delta_{kl} + A_2\,(\delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}) \quad , \quad N_{ijkl} = N_1\,(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl}) + N_2\,\delta_{il}\delta_{jk} \\
S_{ijkl} &= S_1\,(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl}) + S_2\,\delta_{il}\delta_{jk} \\
V_{ijklmn} &= V_1\,(\delta_{ij}\delta_{kl}\delta_{mn} + \delta_{jk}\delta_{in}\delta_{lm} + \delta_{ik}\delta_{jn}\delta_{lm} + \delta_{ij}\delta_{km}\delta_{nl}) + V_2\,(\delta_{ij}\delta_{kn}\delta_{lm}) \\
&\quad + V_3\,(\delta_{jk}\delta_{il}\delta_{mn} + \delta_{jk}\delta_{im}\delta_{nl} + \delta_{ik}\delta_{jl}\delta_{mn} + \delta_{ik}\delta_{jm}\delta_{nl}) \\
&\quad + V_4\,(\delta_{il}\delta_{jm}\delta_{kn} + \delta_{jl}\delta_{kn}\delta_{im}) \\
&\quad + V_5\,(\delta_{jl}\delta_{km}\delta_{in} + \delta_{kl}\delta_{im}\delta_{jn} + \delta_{il}\delta_{jn}\delta_{km} + \delta_{kl}\delta_{in}\delta_{jm})
\end{aligned}
\tag{3.13}
$$

as well as

$$
B = E\ ,\ S_1 = N_1\ ,\ S_2 = N_2 \tag{3.14}
$$

where 11 scalar material constants are introduced.

Considering the reduced form of the material constant tensors in equation (3.13), the generalized forces listed in (3.12) can be rewritten as

$$
\begin{aligned}
\varsigma^D_{ij} =& A_1\delta_{ij}\,\dot{\varepsilon^{tr}}_{kk} + A_2\left(\dot{\varepsilon^{tr}}_{ij} + \dot{\varepsilon^{tr}}_{ji}\right) + B\delta_{ij}\,\dot{\xi} = A\,\dot{\varepsilon^{tr}}_{ij} + B\,\dot{\xi}\delta_{ij} \\
\pi =& E\,\delta_{ij}\,\dot{\varepsilon^{tr}}_{ij} + a\,\dot{\xi} = a\,\dot{\xi} \\
\bar{\pi}_i =& M\delta_{ij}\,\dot{\xi}_{,j} + N_1\left(\dot{\varepsilon^{tr}}_{il,l} + \dot{\varepsilon^{tr}}_{li,l}\right) + N_2\dot{\varepsilon^{tr}}_{kk,i} = M\dot{\xi}_{,i} + 2N_1\dot{\varepsilon^{tr}}_{ij,j} \\
\tau_{ijk} =& S_1\left(\delta_{ij}\dot{\xi}_{,k} + \delta_{ik}\dot{\xi}_{,j}\right) + S_2\,\delta_{jk}\dot{\xi}_{,i} + V_1\left(\delta_{ij}\dot{\varepsilon^{tr}}_{kn,n} + \delta_{jk}\dot{\varepsilon^{tr}}_{ll,i} + \delta_{ik}\dot{\varepsilon^{tr}}_{nn,j} + \delta_{ij}\dot{\varepsilon^{tr}}_{kl,l}\right) \\
&+ V_2\left(\delta_{ij}\dot{\varepsilon^{tr}}_{ll,k}\right) + V_3\left(\delta_{jk}\dot{\varepsilon^{tr}}_{in,n} + \delta_{jk}\dot{\varepsilon^{tr}}_{li,l} + \delta_{ik}\dot{\varepsilon^{tr}}_{lj,l}\right) + V_4\left(\dot{\varepsilon^{tr}}_{ij,k} + \dot{\varepsilon^{tr}}_{ji,k}\right) \\
&+ 2V_5\left(\dot{\varepsilon^{tr}}_{jk,i} + \dot{\varepsilon^{tr}}_{ki,j}\right) \\
=& N_1\left(\delta_{ij}\dot{\xi}_{,k} + \delta_{ik}\dot{\xi}_{,j}\right) + N_2\,\delta_{jk}\dot{\xi}_{,i} + 2V_1\delta_{ij}\,\dot{\varepsilon^{tr}}_{kn,n} + 2V_3\delta_{jk}\,\dot{\varepsilon^{tr}}_{in,n} + V_3\delta_{ik}\,\dot{\varepsilon^{tr}}_{jn,n} \\
&+ 2V_4\dot{\varepsilon^{tr}}_{ij,k} + 2V_5\left(\dot{\varepsilon^{tr}}_{kj,i} + \dot{\varepsilon^{tr}}_{ki,j}\right)
\end{aligned}
\tag{3.15}
$$

It is assumed for the above derivation that the transformation strain tensor is traceless $\mathrm{tr}\,(\boldsymbol{\varepsilon}^{tr}) = 0$. The dissipation due to the martensitic transformation can be calculated using the above results.

$$
\begin{aligned}
D =& A\,\dot{\varepsilon^{tr}}_{ij}\,\dot{\varepsilon^{tr}}_{ij} + a\dot{\xi}^2 + M\dot{\xi}_{,i}\,\dot{\xi}_{,i} + (2N_1 + N_2)\,\dot{\varepsilon^{tr}}_{ij,j}\,\dot{\xi}_{,i} \\
&+ 3V_3\dot{\varepsilon^{tr}}_{ij,j}\,\dot{\varepsilon^{tr}}_{ik,k} + 2V_4\dot{\varepsilon^{tr}}_{ij,k}\,\dot{\varepsilon^{tr}}_{ij,k} + 4V_5\dot{\varepsilon^{tr}}_{ij,k}\,\dot{\varepsilon^{tr}}_{ik,j} \geq 0
\end{aligned}
\tag{3.16}
$$

Here, it was also assumed that the trace of the transformation strain vanishes. As a result of this assumption, the terms with coefficients of V_1 and B will not yield any contribution to the dissipation and hence their corresponding terms in the formulation of generalized forces are ignored. Furthermore, the generalized stress ${\varsigma^D}_{ij}$ must be replaced with its deviatoric part $\varsigma^{\acute{D}}$. Based on this, the set of equations

relating the generalized forces and fluxes can be rewritten.

$$
\begin{aligned}
\acute{\varsigma}^{D}{}_{ij} =& A\,\dot{\varepsilon^{tr}}_{ij} \\
\pi =& a\,\dot{\xi} \\
\bar{\pi}_i =& M\,\dot{\xi}_{,i} + 2N_1\,\dot{\varepsilon^{tr}}_{ij,j} \\
\tau_{ijk} =& N_1\left(\delta_{ij}\dot{\xi}_{,k} + \delta_{ik}\dot{\xi}_{,j}\right) + N_2\,\delta_{jk}\dot{\xi}_{,i} + 2V_3\delta_{jk}\,\dot{\varepsilon^{tr}}_{in,n} + V_3\delta_{ik}\,\dot{\varepsilon^{tr}}_{jn,n} \\
& + 2V_4\dot{\varepsilon^{tr}}_{ij,k} + 2V_5\left(\dot{\varepsilon^{tr}}_{kj,i} + \dot{\varepsilon^{tr}}_{ki,j}\right)
\end{aligned}
\tag{3.17}
$$

As it can be seen, the relation between the forces and fluxes is almost decoupled in the evolution equations of (3.17). This result is due to assuming a linear relationship between the generalized forces and the rate of internal variables as well as considering an isotropic and isochoric response for the SMA. The constitutive formulation is concluded by considering a form for the Gibb's free energy

$$
G = \hat{G}\left(\sigma_{ij},\, T,\, \dot{\varepsilon^{tr}}_{ij},\, \xi,\, \xi_{,i},\, \dot{\varepsilon^{tr}}_{ij,k}\right) \tag{3.18}
$$

based on which the strain and the generalized forces can be obtained via (3.2) and (3.4). The effect of the degree of the polynomial form assumed for the free energy and the corresponding coupling terms are discussed in more details in the subsequent chapters.

4. GRADIENT-BASED RATE-INDEPENDENT SMA CONSTITUTIVE MODELING

The development of a thermodynamically consistent gradient-based model for the SMAs, using the internal variable approach (*thermodynamics with internal variables, T.I.V*), is discussed in details in this chapter.

The development, to that end, begins with forming a specific Gibbs free energy for the material as a function of the state variables; the *controllable* external state variables, stress and temperature, as well as the internal state variables [100]. The strain and entropy are chosen, in accordance with the second law of thermodynamics, to be the derivatives of the Gibbs free energy with respect to stress and entropy. Ultimately, a dissipation potential is formulated in order to obtain rate-independent evolution equations for the internal variables.

To begin, the dependent and independent state variables are introduced.

$$\boldsymbol{\varepsilon} = \hat{\boldsymbol{\varepsilon}}\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right) \quad , \quad \boldsymbol{q} = \hat{\boldsymbol{q}}\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right) \quad , \quad s = \hat{s}\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right) \quad , \quad G = \hat{G}\left(\boldsymbol{\sigma}, T, \boldsymbol{\Upsilon}\right)$$

$$\boldsymbol{\Upsilon} \equiv \{\xi, \{\nabla\xi, \nabla\nabla\xi, \cdots, \nabla^n\xi\}, \{\nabla\boldsymbol{\varepsilon}^{tr}, \nabla\nabla\boldsymbol{\varepsilon}^{tr}, \cdots, \nabla^m\boldsymbol{\varepsilon}^{tr}\}\} \tag{4.1}$$

A small-strain framework is assumed here with infinitesimal gradients of displacement. The set of internal variables $\boldsymbol{\Upsilon}$ includes $\boldsymbol{\varepsilon}^{tr}$, the tensor for transformation strain, and ξ, the volume fraction of martensite phase in the SMA. Also included are the spatial gradients of the martensite volume fraction and the transformation strain tensor up to orders n and m. Without loss of generality, it is assumed henceforth that $m = n$.

$\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}^{tr}$ belong to the space of symmetric second order tensors. $\nabla^n \xi$ and $\nabla^n \boldsymbol{\varepsilon}^{tr}$, also, have the following symmetry with respect to the coordinates within an orthonormal basis:

$$\sigma_{ij} = \sigma_{ji}\ , \quad \varepsilon^{tr}_{ij} = \varepsilon^{tr}_{ji}\ , \quad \varepsilon^{tr}_{ij,\,k_1\cdots k_n} = \varepsilon^{tr}_{ji,\,k_1\cdots k_n} \tag{4.2}$$

$$\xi_{,\,i_1\cdots i_n} = \xi_{,\,(i_1\cdots i_n)} \quad , \quad \varepsilon^{tr}_{ij,\,k_1\cdots k_n} = \varepsilon^{tr}_{ji,\,(k_1\cdots k_m)} \tag{4.3}$$

Identities in (4.2) are due to the principle of conservation of linear momentum (see section 2.2.3), and the symmetry of the infinitesimal strain tensor, equation (2.7). Equation (4.3), on the other hand, comes form the continuity of the field being differentiated. By $(k_1 \cdots k_n)$ any combination of symbols form k_1 to k_n is implied.

The constitutive are derived such that the second law of thermodynamics is satisfied through any admissible thermodynamic process. Therefore, the conventional first and second laws of thermodynamics are invoked, assuming the Fourier's law (2.23) for heat conduction, in order to reach the Clausius-Plank inequality:

$$\begin{aligned} &-\rho\dot{G} - \rho s\dot{T} - \boldsymbol{\varepsilon} : \dot{\boldsymbol{\sigma}} \geq 0 \\ &-\left(\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} + \boldsymbol{\varepsilon}\right) : \dot{\boldsymbol{\sigma}} - \rho\left(\frac{\partial G}{\partial T} + s\right)\dot{T} \\ &-\rho\left[\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} : \dot{\boldsymbol{\varepsilon}}^{tr} + \frac{\partial G}{\partial \xi}\dot{\xi} + \sum_{k=1}^{n}\left(\frac{\partial G}{\partial \nabla^k \xi}\cdot\nabla^k\dot{\xi} + \frac{\partial G}{\partial \nabla^k \boldsymbol{\varepsilon}^{tr}}\cdot\nabla^k\dot{\boldsymbol{\varepsilon}}^{tr}\right)\right] \geq 0 \end{aligned} \tag{4.4}$$

Assuming that all the constitutive variables and their rates are independent, the Clausius-Plank inequality, and hence the second law of thermodynamics as previously mentioned in section 2.3, are satisfied given the following constitutive equations:

$$\boldsymbol{\varepsilon} = -\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} \quad , \quad s = -\frac{\partial G}{\partial T} \tag{4.5}$$

Therefore the dissipation associated with the phase transformation and also the generalized thermodynamic forces conjugate to the dissipative internal variables will be established.

$$D^{tr} = \boldsymbol{\Gamma} \cdot \dot{\boldsymbol{\Upsilon}} = \boldsymbol{\varsigma}^D : \dot{\boldsymbol{\varepsilon}}^{tr} + \pi\dot{\xi} + \sum_{k=1}^{n} \bar{\boldsymbol{\pi}}^k \cdot \nabla^k \dot{\xi} + \sum_{k=1}^{n} \boldsymbol{\tau}^k \cdot \nabla^k \dot{\boldsymbol{\varepsilon}}^{tr} \geq 0 \tag{4.6}$$

where $\dot{\boldsymbol{\Upsilon}}$ denotes the set of generalized thermodynamic fluxes with corresponding thermodynamic forces, $\boldsymbol{\Gamma}$, given by

$$\boldsymbol{\Gamma} = -\rho\frac{\partial G}{\partial \boldsymbol{\Upsilon}} \tag{4.7}$$

that is

$$\boldsymbol{\varsigma}^D = -\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} \quad , \quad \pi = -\rho\frac{\partial G}{\partial \xi} \quad , \quad \bar{\boldsymbol{\pi}}^k = -\rho\frac{\partial G}{\partial \nabla^k \xi} \quad , \quad \boldsymbol{\tau}^k = -\rho\frac{\partial G}{\partial \nabla^k \boldsymbol{\varepsilon}^{tr}} \tag{4.8}$$

It is common to consider the martensitic transformation a volume preserving process; i.e. $\mathrm{tr}\,(\boldsymbol{\varepsilon}^{tr}) = 0$. Therefore, the hydrostatic part of the generalized stress, $\boldsymbol{\varsigma}^D$, does not contribute to the rate of energy dissipated, D^{tr}, hence only its deviatoric part will be considered for determining the evolution of the internal variables.

$$\acute{\varsigma}_{ij} = \varsigma_{ij} - \frac{1}{3}\mathrm{tr}\left(\boldsymbol{\varsigma}^{D}\right)\delta_{ij} \quad , \quad \mathrm{tr}\left(\acute{\boldsymbol{\varsigma}}^{D}\right) = \acute{\varsigma}_{ii} = 0 \tag{4.9}$$

It is assumed here that a transformation surface exists such that it defines the region for the thermoelastic state of the SMA material. The transformation surface is a function of the generalized forces and any admissible state of the material must satisfy the condition imposed by it.

$$\Phi\left(\boldsymbol{\Gamma}\right) = \Phi\left(\acute{\boldsymbol{\varsigma}}^{D},\ \pi,\ \bar{\boldsymbol{\pi}}^{1}, \cdots, \bar{\boldsymbol{\pi}}^{n},\ \boldsymbol{\tau}^{1}, \cdots, \boldsymbol{\tau}^{n}\right) \leq 0 \tag{4.10}$$

The rate of the internal variables, or the generalized thermodynamic fluxes, are commonly assumed to be linear or nonlinear functions of the generalized thermodynamic forces. For a rate-independent response, the case considered here, the fluxes are associated to the forces through derivatives of a transformation potential [96]. One approach uses the *principle of maximum dissipation*, PMD, such that the fluxes will be the normals to the transformation surface. Normality of the fluxes in the space of the generalized forces and also convexity of the transformation surface are the results of the PMD and satisfaction of the 2nd law of thermodynamics. According to PMD, the transformation state of the SMA material, belonging to the set of admissible states, is the one that maximizes the dissipation D^{tr}; or

$$\mathrm{Max}\left\{D^{tr} = \boldsymbol{\Gamma}\cdot\dot{\boldsymbol{\Upsilon}} \mid \boldsymbol{\Gamma}\ ,\ \Phi(\boldsymbol{\Gamma}) \leq 0\right\} \tag{4.11}$$

This is a minimization programming which is solved through the Lagrange method

of multipliers. A Lagrangian, L^{tr}, is defined as

$$L^{tr} = -\mathbf{\Gamma} \cdot \dot{\mathbf{\Upsilon}} + \lambda\, \Phi(\mathbf{\Gamma}) \tag{4.12}$$

with the multiplier, λ, introduced and it is shown that the problem in (4.11) is equivalent to finding the minimum of L^{tr} or

$$\frac{\partial L^{tr}}{\partial \mathbf{\Gamma}} = -\dot{\mathbf{\Upsilon}} + \lambda \frac{\partial \Phi}{\partial \mathbf{\Gamma}} = 0 \tag{4.13}$$

provided that the following Kuhn-Tucker conditions are satisfied:

1. $\Phi(\mathbf{\Gamma})$ is convex and $\Phi \leq 0$
2. $\lambda \geq 0$
3. $\lambda \Phi(\mathbf{\Gamma}) = 0$

Therefore the rates of the internal variables can be found via

$$\dot{\boldsymbol{\varepsilon}}^{tr} = \lambda \frac{\partial \Phi}{\partial \boldsymbol{\varsigma}^{D}} \quad , \quad \dot{\xi} = \lambda \frac{\partial \Phi}{\partial \pi} \quad , \quad \nabla^k \dot{\xi} = \lambda \frac{\partial \Phi}{\partial \bar{\boldsymbol{\pi}}^k} \quad , \quad \nabla^k \dot{\boldsymbol{\varepsilon}}^{tr} = \lambda \frac{\partial \Phi}{\partial \boldsymbol{\tau}^k} \tag{4.14}$$

subjected to the aforementioned Kuhn-Tucker conditions. Hence, the response is associative in the space of generalized thermodynamic forces.

Next is to assume a form for the transformation surface $\Phi(\mathbf{\Gamma})$. For Φ to be a general anisotropic convex function of $\boldsymbol{\varsigma}^{'D}$, π, $\bar{\boldsymbol{\pi}}^1, \cdots, \bar{\boldsymbol{\pi}}^n$, and $\boldsymbol{\tau}^1, \cdots \boldsymbol{\tau}^n$ satisfying the second law of thermodynamics as the for forward and reverse martensitic transformations, it is sufficient to have

$$\begin{aligned}\Phi^{\text{fwd}}\left(\boldsymbol{\varsigma}^{\acute{D}},\ \pi,\ \bar{\boldsymbol{\pi}}^1,\cdots,\bar{\boldsymbol{\pi}}^n,\ \boldsymbol{\tau}^1,\cdots,\boldsymbol{\tau}^n\right)&=\phi-Y\\ \Phi^{\text{rev}}\left(\boldsymbol{\varsigma}^{\acute{D}},\ \pi,\ \bar{\boldsymbol{\pi}}^1,\cdots,\bar{\boldsymbol{\pi}}^n,\ \boldsymbol{\tau}^1,\cdots,\boldsymbol{\tau}^n\right)&=\phi+Y\end{aligned}\tag{4.15}$$

$$\phi=\widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}})+\check{\varphi}(\pi)+\sum_{k=1}^{n}\bar{\varphi}^k(\bar{\boldsymbol{\pi}}^k)+\sum_{k=1}^{n}\hat{\varphi}^k(\boldsymbol{\tau}^k)$$

where various φ are homogeneous of degree 1, convex functions of their respective variables. A general quadratic form is commonly assumed such that:

$$\begin{aligned}\widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}})&=\left(\Lambda^I_{ijrs}\,\acute{\varsigma}_{ij}\,\acute{\varsigma}_{rs}\right)^{\frac{1}{2}}\quad,\quad\check{\varphi}(\pi)=\pi\\ \bar{\varphi}^k(\bar{\boldsymbol{\pi}}^k)&=-\boldsymbol{\Xi}^{II_k}\cdot\bar{\boldsymbol{\pi}}^k-\left(\bar{\boldsymbol{\pi}}^k\cdot\boldsymbol{\Lambda}^{II_k}\bar{\boldsymbol{\pi}}^k\right)^{\frac{1}{2}}\\ \hat{\varphi}(\boldsymbol{\tau})&=-\boldsymbol{\Xi}^{III_k}\cdot\boldsymbol{\tau}^k-\left(\boldsymbol{\tau}^k\cdot\boldsymbol{\Lambda}^{III_k}\boldsymbol{\tau}^k\right)^{\frac{1}{2}}\end{aligned}\tag{4.16}$$

with no summation over k. Here, notice that if $\bar{\boldsymbol{\pi}}^k$ or $\boldsymbol{\tau}^k\in\mathcal{T}^m$ then $\boldsymbol{\Xi}^{II_k}$ or $\boldsymbol{\Xi}^{III_k}\in\mathcal{T}^m$ and $\boldsymbol{\Lambda}^{II_k}$ or $\boldsymbol{\Lambda}^{III_k}\in\mathcal{T}^{2m}$. The tensors $\boldsymbol{\Lambda}$ and $\boldsymbol{\Xi}$ can be assigned the required symmetry in order to achieve the directional dependence of the yield surface and flow observed in the response of the SMA. Y is a constant that is equal to one-half of the amount of energy dissipation in a full transformation path.

ϕ given by equation (4.15) can result in multiple forms of differential equations. For example, consider $\bar{\boldsymbol{\pi}}^1=-\rho\frac{\partial G}{\partial\nabla\xi}=a\nabla\xi$ and $\bar{\boldsymbol{\pi}}^2=-\rho\frac{\partial G}{\partial\nabla^2\xi}=b\nabla^2\xi$ which can be part of an isotropic SMA constitutive model including ξ, $\nabla\xi$, and $\nabla^2\xi$ as internal variables. Hence, one can have first order and second order differential terms, including the Laplacian $\xi_{,ii}$, within the definition of the partial differential equation for the transformation surface.

$$\begin{aligned} \bar{\pi}_i^1 &= a\,\xi_{,i} \qquad \bar{\varphi}^1(\bar{\boldsymbol{\pi}}^1) = -A\sqrt{\xi_{,i}\,\xi_{,i}} \\ \bar{\pi}_{ij}^2 &= b\,\xi_{,ij} \qquad \bar{\varphi}^2(\bar{\boldsymbol{\pi}}^2) = -B\,\xi_{,ii} - \sqrt{C\,\xi_{,ii}^2 + D\,\xi_{,ij}\,\xi_{,ij}} \end{aligned} \tag{4.17}$$

It is possible to eliminate λ from equations (4.14) such that the rate of evolution of the internal variables will be proportional to $\dot{\xi}$ since $\dot{\xi} = \lambda\frac{\partial\Phi}{\partial\pi} = \lambda$. This implies that $\dot{\xi}$ plays the role of the plastic multiplier as in the analysis of dissipation surfaces and flow in plasticity. $\dot{\xi}$ can, furthermore, be determined via the consistency condition.

According to the Kuhn-Tucker condition (iii) whilst $\dot{\xi} = \lambda \neq 0 \rightarrow \Phi = 0$, thus the consistency condition during forward or reverse transformation (fwd/rev) can be obtained considering the transformation surface in (4.15):

$$\begin{aligned} \dot{\Phi}^{\text{fwd/rev}} &= 0 \\ \frac{\partial\Phi}{\partial\boldsymbol{\sigma}}^{\text{fwd/rev}} : \dot{\boldsymbol{\sigma}} + \frac{\partial\Phi}{\partial T}^{\text{fwd/rev}} \dot{T} + \frac{\partial\Phi}{\partial\boldsymbol{\varepsilon}^{tr}}^{\text{fwd/rev}} : \dot{\boldsymbol{\varepsilon}}^{tr} + \frac{\partial\Phi}{\partial\xi}^{\text{fwd/rev}} \dot{\xi} \\ + \sum_{k=1}^{n} \frac{\partial\Phi}{\partial\nabla^k\xi}^{\text{fwd/rev}} \cdot \nabla^k\dot{\xi} + \sum_{k=1}^{n} \frac{\partial\Phi}{\partial\nabla^k\boldsymbol{\varepsilon}^{tr}}^{\text{fwd/rev}} \cdot \nabla^k\dot{\boldsymbol{\varepsilon}}^{tr} = 0 \end{aligned} \tag{4.18}$$

which leads to the following for the evolution of martensite volume fraction.

$$\dot{\xi} = -\frac{\frac{\partial\phi}{\partial\boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} + \frac{\partial\phi}{\partial T}\dot{T}}{\frac{\partial\phi}{\partial\xi} + \frac{\partial\phi}{\partial\boldsymbol{\varepsilon}^{tr}} : \frac{\partial\phi}{\partial\boldsymbol{\varsigma}^D} + \sum_{k=1}^{n} \frac{\partial\phi}{\partial\nabla^k\xi} \cdot \frac{\partial\phi}{\partial\bar{\boldsymbol{\pi}}^k} + \sum_{k=1}^{n} \frac{\partial\phi}{\partial\nabla^k\boldsymbol{\varepsilon}^{tr}} \cdot \frac{\partial\phi}{\partial\boldsymbol{\tau}^k}} \tag{4.19}$$

Once the evolution of internal variables are given by the derivatives of the trans-

formation surface as in equation (4.14), the dissipation due to martensitic transformation can be obtained.

$$D^{tr} = \lambda \left(\frac{\partial \Phi}{\partial \mathbf{\Gamma}} \cdot \mathbf{\Gamma} \right) = \lambda \left(\frac{\partial \phi}{\partial \mathbf{\Gamma}} \cdot \mathbf{\Gamma} \right) = \lambda \, \phi \left(\mathbf{\Gamma} \right) = \begin{cases} \quad Y\dot{\xi}, & \dot{\xi} > 0 \\ - \quad Y\dot{\xi}, & \dot{\xi} < 0 \end{cases} \tag{4.20}$$

In which Euler's homogeneous function theorem is used as $\phi(\mathbf{\Gamma})$ is continuously differentiable and homogeneous of degree one. The satisfaction of the second law of thermodynamics is guaranteed through equation (4.20). This equation requires careful attention.

Firstly, the rates of internal variables ξ and $\nabla^k \xi$ and also $\boldsymbol{\varepsilon}^{tr}$ and $\nabla^k \boldsymbol{\varepsilon}^{tr}$ are not totally independent as they are related by a spatial gradient. This imposes a certain restriction on the transformation surface as below.

$$\begin{aligned} \nabla \left(\lambda \frac{\partial \Phi}{\partial \pi} \right) &= \lambda \frac{\partial \Phi}{\partial \bar{\boldsymbol{\pi}}^1} \quad , \quad \nabla \left(\lambda \frac{\partial \Phi}{\partial \boldsymbol{\varsigma}^D} \right) = \lambda \frac{\partial \Phi}{\partial \boldsymbol{\tau}^1} \\ \nabla \left(\lambda \frac{\partial \Phi}{\partial \bar{\boldsymbol{\pi}}^k} \right) &= \lambda \frac{\partial \Phi}{\partial \bar{\boldsymbol{\pi}}^{k+1}} \quad , \quad \nabla \left(\lambda \frac{\partial \Phi}{\partial \boldsymbol{\tau}^k} \right) = \lambda \frac{\partial \Phi}{\partial \boldsymbol{\tau}^{k+1}} \end{aligned} \tag{4.21}$$

The first one of the above constraints is trivially satisfied in light of $\dot{\xi} = \lambda$ and (4.14)c. A discussion about such restrictions on the transformation surface can also be found in [128].

Secondly, as observed in equation (4.20), the inelastic energy dissipation does not explicitly depend on the rate of the gradient of martensite volume fraction and transformation strain. This is a consequence of the specific form assumed for the transformation surface, (4.15), and also the rates of ξ and $\nabla \xi$ as well as $\boldsymbol{\varepsilon}^{tr}$ and $\nabla \boldsymbol{\varepsilon}^{tr}$. However, the nonhomogeneous distributions of the martensitic volume fraction

and the transformation strain in the material contribute to the dissipation through changing the evolution of ξ as a result of their influence on the transformation surface.

4.1 The most general *anisotropic* 3rd-degree SMA gradient model

In this section, an anisotropic gradient-based constitutive model for shape memory alloys is presented, with $\boldsymbol{\Upsilon}$ considered as internal variables.

$$\boldsymbol{\mathcal{X}} \equiv \{\boldsymbol{\sigma},\, T,\, \boldsymbol{\Upsilon}\} \quad , \quad \boldsymbol{\Upsilon} \equiv \{\xi,\, \nabla\xi,\, \boldsymbol{\varepsilon}^{tr},\, \nabla\boldsymbol{\varepsilon}^{tr}\} \tag{4.22}$$

The derivation of the constitutive response follows the procedure laid out so far from the beginning of this chapter. Only the key equations will be discussed in this section. The form for the Gibbs free energy

$$G = \hat{G}\left(\boldsymbol{\mathcal{X}}\right) = \hat{G}\left(\boldsymbol{\sigma},\, T,\, \xi,\, \nabla\xi,\, \boldsymbol{\varepsilon}^{tr},\, \nabla\boldsymbol{\varepsilon}^{tr}\right) = \hat{G}\left(\sigma_{ij},\, T,\, \xi,\, \xi_{,i},\, \varepsilon^{tr}_{ij},\, \varepsilon^{tr}_{ij,k}\right) \tag{4.23}$$

is commonly selected to be a polynomial in terms of the variables considered to determine the state of the material. Depending on the level of coupling between variables and on the physical phenomena sought to be captured, the polynomial function can be truncated to within a certain degree. This form can be influenced by the material symmetry considered as well [19, 136].

The Gibbs free energy per unit volume can be given by ρG. It is possible to define a reference state, $\boldsymbol{\mathcal{X}}_0$, for the free energy as below.

$$G_0 = \hat{G}\left(\boldsymbol{\mathcal{X}}_0\right) = \hat{G}\left(\mathbf{0},\, T_0,\, 0,\, \mathbf{0},\, \mathbf{0},\, \mathbf{0}\right) \tag{4.24}$$

It is possible to establish the Taylor's expansion of the free energy with respect to this reference state in order to achieve a polynomial form for the free energy function.

$$\begin{aligned}\hat{G}\left(\boldsymbol{\mathcal{X}}\right) = \hat{G}\left(\boldsymbol{\mathcal{X}}_0 + \left(\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{X}}_0\right)\right) = G_0 &+ \mathrm{D}_{\boldsymbol{\mathcal{X}}} G(\boldsymbol{\mathcal{X}}_0)\left[\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{X}}_0\right] \\ &+ \frac{1}{2!}\mathrm{D}^2_{\boldsymbol{\mathcal{X}}} G(\boldsymbol{\mathcal{X}}_0)\left[\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{X}}_0, \boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{X}}_0\right] \\ &+ \frac{1}{3!}\mathrm{D}^3_{\boldsymbol{\mathcal{X}}} G(\boldsymbol{\mathcal{X}}_0)\left[\boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{X}}_0, \boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{X}}_0, \boldsymbol{\mathcal{X}} - \boldsymbol{\mathcal{X}}_0\right] \\ &+ \mathcal{O}(\boldsymbol{\mathcal{X}}^4)\end{aligned} \tag{4.25}$$

Implicit in this equation is the assumption of continuity and differentiability of the Gibbs free energy up to the desired order.

Based on this expansion, the most general anisotropic Gibbs free energy with a polynomial form including terms upto 3^{rd}-degree can be obtained.

The free energy includes all possible couplings between the state variables with their corresponding coefficient tonsorial material constants. The material constants are the derivatives of the free energy calculated at the reference state. For example, the compliance tensor is obtained from

$$A_{ijkl} = \frac{1}{2}\frac{\partial^2 G}{\partial \sigma_{ij}\, \partial \sigma_{kl}} \tag{4.26}$$

This leads to the following major and minor symmetries.

$$A_{ijkl} = A_{klij} \quad , \quad A_{ijkl} = A_{ijlk} = A_{jikl} \tag{4.27}$$

The dimensions of the material constants follow the physics involved.

$$\rho\, G\left(\boldsymbol{\sigma},\, T,\, \xi,\, \boldsymbol{\varepsilon}^{tr},\, \nabla\xi,\, \nabla\boldsymbol{\varepsilon}^{tr}\right) = \rho\, G_0 \tag{4.28}$$

$$+ \Big(\xi P_{ijkl} + T Q_{ijkl} + F_{ijklmn}\,\sigma_{mn} + N_{ijklmn}\,\boldsymbol{\varepsilon}^{tr}_{mn} + O_{ijklm}\,\xi_{,m} + \mathbb{A}_{ijklmno}\,\varepsilon^{tr}_{mn,o} + A_{ijkl}\Big)\sigma_{ij}\,\sigma_{kl}$$

$$+ \Big(\xi \bar{A}_{ijkl} + T\bar{B}_{ijkl} + S_{ijklmn}\,\boldsymbol{\varepsilon}^{tr}_{mn} + Z_{ijklm}\,\xi_{,m} + \mathbb{F}_{ijklmno}\,\varepsilon^{tr}_{mn,o} + C_{ijkl}\Big)\sigma_{ij}\,\boldsymbol{\varepsilon}^{tr}_{kl}$$

$$+ \Big(\mathbb{B}_{ijklmno}\,\varepsilon^{tr}_{mn,o} + \xi\, V_{ijkl} + T\, U_{ijkl} + K_{ijklmn}\,\boldsymbol{\varepsilon}^{tr}_{mn} + \tau_{ijklm}\,\xi_{,m} + B_{ijkl}\Big)\boldsymbol{\varepsilon}^{tr}_{ij}\,\boldsymbol{\varepsilon}^{tr}_{kl}$$

$$+ \Big(\mathbb{C}_{ijlmn}\,\varepsilon^{tr}_{lm,n} + L_{ij}\,\xi + M_{ij}\,T + W_{klij}\,\sigma_{kl} + Y_{klij}\,\boldsymbol{\varepsilon}^{tr}_{kl} + L_{ijk}\,\xi_{,k} + C_{ij}\Big)\xi_{,i}\,\xi_{,j}$$

$$+ \Big(\mathbb{H}_{ijklmn}\,\sigma_{ij} + \mathbb{K}_{ijklmn}\,\boldsymbol{\varepsilon}^{tr}_{ij} + \mathbb{U}_{klmn}\,\xi + \mathbb{V}_{klmn}\,T\Big)\xi_{,k}\,\varepsilon^{tr}_{lm,n}$$

$$+ \Big(\widetilde{N}_{ijklmnopq}\,\varepsilon^{tr}_{op,q} + \widetilde{H}_{ijklmnpq}\,\sigma_{pq} + \widetilde{J}_{ijklmnpq}\,\boldsymbol{\varepsilon}^{tr}_{pq} + \widetilde{K}_{ijklmnp}\,\xi_{,p} + \widetilde{L}_{ijklmn}\,\xi + \widetilde{M}_{ijklmn}\,T\Big)\varepsilon^{tr}_{ij,k}\,\varepsilon^{tr}_{lm,n}$$

$$+ \Big(\mathbb{I}_{ijlmn}\,\xi\,\varepsilon^{tr}_{lm,n} + \mathbb{J}_{ijlmn}\,T\,\varepsilon^{tr}_{lm,n} + D_{ij}\,\xi + \bar{C}_{ijk}\,\xi\,\xi_{,k} + \bar{D}_{ijk}\,T\,\xi_{,k} + N_{ij}\,\xi^2 + E_{ij}\,T + P_{ij}\,\xi T + Q_{ij}\,T^2 + \alpha_{ij}\Big)\sigma_{ij}$$

$$+ \Big(\mathbb{L}_{ijlmn}\,\xi\,\varepsilon^{tr}_{lm,n} + \mathbb{M}_{ijlmn}\,T\,\varepsilon^{tr}_{lm,n} + F_{ij}\,\xi + O_{ij}\,\xi^2 + K_{ij}\,T + \bar{K}_{ij}\,T\xi + S_{ij}\,T^2 + \bar{E}_{ijk}\,\xi\,\xi_{,k} + \bar{F}_{ijk}\,T\,\xi_{,k} + \beta_{ij}\Big)\boldsymbol{\varepsilon}^{tr}_{ij}$$

$$+ \Big(D_{mni}\,\sigma_{mn} + E_{mni}\,\boldsymbol{\varepsilon}^{tr}_{mn} + g_i\,T + f_i\,\xi + r_i\,T\xi + m_i\,\xi^2 + o_i\,T^2 + a_i\Big)\,\xi_{,i}$$

$$+ \Big(\widetilde{A}_{ijklmn}\,\varepsilon^{tr}_{lm,n} + \widetilde{B}_{ijklm}\,\sigma_{lm} + \widetilde{C}_{ijklm}\,\boldsymbol{\varepsilon}^{tr}_{lm} + \widetilde{D}_{ijkl}\,\xi_{,l} + \widetilde{E}_{ijk}\,\xi + \widetilde{F}_{ijk}\,T + \mathbb{O}_{ijk}\,T\xi + \mathbb{D}_{ijk}\,\xi^2 + \mathbb{E}_{ijk}\,T^2 + \widetilde{\alpha}_{ijk}\Big)\varepsilon^{tr}_{ij,k}$$

$$+ \Big(l\,T^2 + d\,T + b\Big)T + \Big(k\,\xi^2 + c\,\xi + a\Big)\xi + \Big(p\,T + n\,\xi + f\Big)T\xi$$

The generalized forces, according to equation (4.14), as well as the transformation surface given in (4.15) and (4.16) can be rewritten for this case.

$$\boldsymbol{\varsigma}^D = -\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} \quad , \quad \pi = -\rho\frac{\partial G}{\partial \xi} \quad , \quad \bar{\boldsymbol{\pi}} = -\rho\frac{\partial G}{\partial \nabla\xi} \quad , \quad \boldsymbol{\tau} = -\rho\frac{\partial G}{\partial \nabla\boldsymbol{\varepsilon}^{tr}} \tag{4.29}$$

$$\dot{\boldsymbol{\varepsilon}}^{tr} = \lambda\frac{\partial \Phi}{\partial \boldsymbol{\varsigma}^D} \quad , \quad \dot{\xi} = \lambda\frac{\partial \Phi}{\partial \pi} \quad , \quad \nabla\dot{\xi} = \lambda\frac{\partial \Phi}{\partial \bar{\boldsymbol{\pi}}} \quad , \quad \nabla\dot{\boldsymbol{\varepsilon}}^{tr} = \lambda\frac{\partial \Phi}{\partial \boldsymbol{\tau}} \tag{4.30}$$

$$\Phi^{\text{fwd}}\left(\boldsymbol{\varsigma}^{\acute{D}},\ \pi,\ \bar{\boldsymbol{\pi}},\ \boldsymbol{\tau}\right) = \phi - Y \quad , \quad \Phi^{\text{rev}}\left(\boldsymbol{\varsigma}^{\acute{D}},\ \pi,\ \bar{\boldsymbol{\pi}},\ \boldsymbol{\tau}\right) = -\phi - Y$$

$$\phi = \widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}}) + \breve{\varphi}(\pi) + \bar{\varphi}(\bar{\boldsymbol{\pi}}) + \hat{\varphi}(\boldsymbol{\tau}) \tag{4.31}$$

where

$$\begin{aligned} \widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}}) &= \left(\Lambda^{I}_{ijkl}\,\acute{\varsigma}_{ij}\,\acute{\varsigma}_{kl}\right)^{\frac{1}{2}} \quad , \quad \breve{\varphi}(\pi) = \pi \\ \bar{\varphi}(\bar{\boldsymbol{\pi}}) &= -\Xi^{II}_{i}\bar{\pi}_i - \left(\Lambda^{II}_{ij}\,\bar{\pi}_i\,\bar{\pi}_j\right)^{\frac{1}{2}} \\ \hat{\varphi}(\boldsymbol{\tau}) &= -\Xi^{III}_{ijk}\bar{\tau}_{ijk} - \left(\Lambda^{III}_{ijklmn}\,\tau_{ijk}\,\tau_{lmn}\right)^{\frac{1}{2}} \end{aligned} \tag{4.32}$$

4.2 The most general *isotropic* 3rd-degree SMA gradient model

Material symmetry applies certain restrictions to the constant tensors included in the Gibbs free energy, equation(4.28). The response of the material for the case of isotropy is obtained in the following.

The response functions in equation (4.1) are isotropic if and only if for any orthogonal transformation $\boldsymbol{Q}$

$$
\begin{aligned}
&\forall\, \boldsymbol{Q} \in \mathrm{Orth}^{+} \ : \\
Q_{ki}\, Q_{lj}\, \hat{\varepsilon}_{kl}\left(\sigma_{ij}, T, \xi, \xi_{,i}, \varepsilon^{tr}_{ij}, \varepsilon^{tr}_{ij,k}\right) &= \\
\hat{\varepsilon}_{ij}\left(Q_{ki}\, Q_{lj}\, \sigma_{kl}, T, \xi, Q_{ji}\, \xi_{,j}, Q_{ki}\, Q_{lj}\, \varepsilon^{tr}_{kl}, Q_{mi}\, Q_{nj}\, Q_{lk} \varepsilon^{tr}_{mn,l}\right) & \\
Q_{ji}\, \hat{q}_{j}\left(\sigma_{ij}, T, \xi, \xi_{,i}, \varepsilon^{tr}_{ij}, \varepsilon^{tr}_{ij,k}\right) &= \\
\hat{q}_{i}\left(Q_{ki}\, Q_{lj}\, \sigma_{kl}, T, \xi, Q_{ji}\, \xi_{,j}, Q_{ki}\, Q_{lj}\, \varepsilon^{tr}_{kl}, Q_{mi}\, Q_{nj}\, Q_{lk}\, \varepsilon^{tr}_{mn,l}\right) & \qquad (4.33) \\
\hat{s}\left(\sigma_{ij}, T, \xi, \xi_{,i}, \varepsilon^{tr}_{ij}, \varepsilon^{tr}_{ij,k}\right) &= \\
\hat{s}\left(Q_{ki}\, Q_{lj}\, \sigma_{kl}, T, \xi, Q_{ji}\, \xi_{,j}, Q_{ki}\, Q_{lj}\, \varepsilon^{tr}_{kl}, Q_{mi}\, Q_{nj}\, Q_{lk}\, \varepsilon^{tr}_{mn,l}\right) & \\
\hat{G}\left(\sigma_{ij}, T, \xi, \xi_{,i}, \varepsilon^{tr}_{ij}, \varepsilon^{tr}_{ij,k}\right) &= \\
\hat{G}\left(Q_{ki}\, Q_{lj}\, \sigma_{kl}, T, \xi, Q_{ji}\, \xi_{,j}, Q_{ki}\, Q_{lj}\, \varepsilon^{tr}_{kl}, Q_{mi}\, Q_{nj}\, Q_{lk}\, \varepsilon^{tr}_{mn,l}\right) &
\end{aligned}
$$

For $\hat{G}$ to be isotropic, the coefficient tensors in equation(4.28) must be isotropic. For example, selecting only the term related to the elastic energy and ignoring other contributions, without loss of generality, we can have :

$$
\begin{aligned}
G^{el} = \hat{G}^{el}\left(\sigma_{ij}\right) = A_{ijkl}\,\sigma_{ij}\,\sigma_{kl} \;\; : \\
\hat{G}^{el}\left(Q_{ki}\,Q_{lj}\,\sigma_{kl}\right) = A_{ijkl}\,Q_{mi}\,Q_{nj}\,\sigma_{mn}\,Q_{rk}\,Q_{sl}\,\sigma_{rs} = Q_{im}\,Q_{jn}\,Q_{kr}\,Q_{ls}\,A_{mnrs}\,\sigma_{ij}\,\sigma_{kl} \\
\hat{G}^{el}\left(\sigma_{ij}\right) = \hat{G}^{el}\left(Q_{ki}\,Q_{lj}\,\sigma_{kl}\right) \\
A_{ijkl} = Q_{im}\,Q_{jn}\,Q_{kr}\,Q_{ls}\,A_{mnrs}
\end{aligned}
\tag{4.34}
$$

which means that the components A_{ijkl} must not change under any orthogonal transformation belonging to the group of proper orthogonal transformations; (4.34)d. Isotropic tensors of rank up to 6 are discussed in the Appendix B.1.

The isotropic tensors of even rank can be expressed as linear combinations of Kronecker deltas, δ_{ij}. The isotropic odd-ranked tensors (except rank one) are given by linear combinations of Kronecker deltas and permutation tensors, ϵ_{ijk}. There is only one isotropic tensor of rank 2 and 3, that being the δ_{ij} and ϵ_{ijk} respectively. There are three independent isotropic tensors of rank 4, fifteen for rank 6, and 91 for rank 8 [84]. Hence, the material constant tensors can be expanded in terms of such fundamental tonsorial combinations. For example:

$$
A_{ijkl} = A_1\delta_{ij}\delta_{kl} + A_2\delta_{ik}\delta_{jl} + A_3\delta_{il}\delta_{jk} \tag{4.35}
$$

Furthermore, the material constants are the derivatives of the Gibbs free energy with respect to their corresponding state variables. Hence, they follow certain symmetries as the order of differentiation is irrelevant under the assumption of a sufficiently continuous free energy function. For instance:

$$\mathbb{H}_{ijklmn} = \frac{1}{2}\frac{\partial^3 G}{\partial\sigma_{ij}\,\partial\xi_{,k}\,\partial\varepsilon^{tr}_{lm,n}}$$
$$\mathbb{H}_{ijklmn} = \mathbb{H}_{kijlmn} = \mathbb{H}_{ijlmnk} = \mathbb{H}_{lmnijk} = \mathbb{H}_{lmnkij} = \mathbb{H}_{klmnij} \tag{4.36}$$

This type of symmetry requirement eliminates all of the isotropic odd-ranked tensors as they include linear combinations of the permutation tensor. In the case of coupling between the gradient of martensite volume fraction and stress:

$$D_{ijk} = \frac{\partial^2 G}{\partial\sigma_{ij}\,\partial\xi_{,k}} = \lambda\,\epsilon_{ijk}$$
$$D_{ijk} = D_{kij} \;\Rightarrow\; \lambda\,\epsilon_{ijk} = -\lambda\,\epsilon_{ijk} \;\Rightarrow\; \lambda = 0 \tag{4.37}$$

Additional symmetries must be considered due to the fact that $\boldsymbol{\sigma}$ and $\boldsymbol{\varepsilon}^{tr}$ are symmetric second order tensors plus $\varepsilon^{tr}_{ij,k} = \varepsilon^{tr}_{ji,k}$. Therefore for $\mathbb{H}$ considered in (4.36):

$$\mathbb{H}_{ijklmn} = \mathbb{H}_{jiklmn} = \mathbb{H}_{ijkmln} \tag{4.38}$$

The 15 independent constants included in each of $\mathbb{H}, \mathbb{K}, \mathbb{U}$ or $\mathbb{V}$, as isotropic rank-six tensors, must reduce to only one due to the restrcitions imposed by the type of equations given in (4.36) and (4.38).

The constraint due to the isochoric martensitic transformation, furthermore, must be addressed. The terms involving ε^{tr}_{ii} or $\varepsilon^{tr}_{ii,j}$ do not contribute to the Gibbs free energy and hence the corresponding material constants become indeterminate.

The Gibbs free energy for the case of isotropic material response can be obtained by rewriting the constant tensors included in (4.28) in terms of their isotropic invari-

ants, albeit considering the aforementioned restrictions. Thus we will have:

$$
\begin{aligned}
&\rho\, G\left(\boldsymbol{\sigma},\, T,\, \xi,\, \boldsymbol{\varepsilon}^{tr},\, \nabla\xi,\, \nabla\boldsymbol{\varepsilon}^{tr}\right) = \rho\, G_0 \qquad (4.39)\\
&+ \Big(\alpha + \xi D_1 + \xi^2 N_6 + T E_1 + T\xi P_3 + T^2 Q_3\Big)\mathrm{tr}\left(\boldsymbol{\sigma}\right)\\
&+ \Big(A_1 + \xi P_1 + T Q_1 + \mathrm{tr}\left(\boldsymbol{\sigma}\right) F_1\Big)\mathrm{tr}\left(\boldsymbol{\sigma}\right)^2 + 2\Big(A_2 + \xi P_2 + T Q_2 + 3\mathrm{tr}\left(\boldsymbol{\sigma}\right) F_2\Big)\mathrm{tr}\left(\boldsymbol{\sigma}^2\right)\\
&+ 8F_5\,\mathrm{tr}\left(\boldsymbol{\sigma}^3\right) + 2\Big(C_2 + \xi\bar{A}_2 + T\bar{B}_2 + 2\mathrm{tr}\left(\boldsymbol{\sigma}\right) N_2\Big)\mathrm{tr}\left(\boldsymbol{\sigma}\boldsymbol{\varepsilon}^{tr}\right) + 8N_5\,\mathrm{tr}\left(\boldsymbol{\sigma}^2\boldsymbol{\varepsilon}^{tr}\right) + 8S_5\,\mathrm{tr}\left(\boldsymbol{\sigma}\boldsymbol{\varepsilon}^{tr^2}\right)\\
&+ 2\Big(B_2 + \xi V_2 + T U_2 + \mathrm{tr}\left(\boldsymbol{\sigma}\right) S_2\Big)\mathrm{tr}\left(\boldsymbol{\varepsilon}^{tr^2}\right) + 8K_5\,\mathrm{tr}\left(\boldsymbol{\varepsilon}^{tr^3}\right)\\
&+ \Big[\Big(c_1 + \xi L_1 + T M_1 + \mathrm{tr}\left(\boldsymbol{\sigma}\right) W_1\Big)\delta_{ij} + 2W_1\sigma_{ij} + 2Y_1\varepsilon^{tr}_{ij}\Big]\;\xi_{,i}\xi_{,j}\\
&+ \Big(\widetilde{H}_{ijklmnpq}\,\sigma_{pq} + \widetilde{J}_{ijklmnpq}\,\boldsymbol{\varepsilon}^{tr}_{pq}\Big)\varepsilon^{tr}_{ij,k}\,\varepsilon^{tr}_{lm,n}\\
&+ 4\Big(\xi\widetilde{L}_5 + T\widetilde{M}_5 + \widetilde{A}_5\Big)\varepsilon^{tr}_{ij,j}\,\varepsilon^{tr}_{ik,k} + 2\Big(\xi\widetilde{L}_{10} + T\widetilde{M}_{10} + \widetilde{A}_{10}\Big)\varepsilon^{tr}_{ij,k}\,\varepsilon^{tr}_{ij,k}\\
&+ 4\Big(\xi\widetilde{L}_{11} + T\widetilde{M}_{11} + \widetilde{A}_{11}\Big)\varepsilon^{tr}_{ij,j}\,\varepsilon^{tr}_{ik,j} + 2\Big(\widetilde{D}_1 + \xi\widetilde{U}_1 + T\widetilde{V}_1 + \mathrm{tr}\left(\boldsymbol{\sigma}\right)\widetilde{H}_1\Big)\xi_i\,\varepsilon^{tr}_{ij,j}\\
&+ 2\left[\Big(\xi_j\,\varepsilon^{tr}_{ik,k} + \xi_i\,\varepsilon^{tr}_{jk,k}\Big) + \Big(\varepsilon^{tr}_{ij,n} + \varepsilon^{tr}_{ni,j} + \varepsilon^{tr}_{nj,i}\Big)\xi_n\right]\Big(\widetilde{H}_1\,\sigma_{ij} + \widetilde{K}_1\,\varepsilon^{tr}_{ij}\Big)\\
&+ \Big(l\,T^2 + d\,T + b\Big)T + \Big(k\,\xi^2 + c\,\xi + a\Big)\xi + \Big(p\,T + n\,\xi + f\Big)T\xi
\end{aligned}
$$

In this equation, the notation for the scalar constants are chosen based on their corresponding tensor. The rank-8 tensors, as listed, are not expanded due to the prohibitive large number (91) of scalar constants involved. As described before, four conditions are taken into account for the coefficient tonsorial material constants in order to derive equation (4.39) from equation (4.28).

1. isotropy,

2. Major symmetry as a result of interchangeability in differentiation of G with respect to state variables,

3. Minor symmetry as a result of symmetry in σ, $\boldsymbol{\varepsilon}^{tr}$, and $\nabla\boldsymbol{\varepsilon}^{tr}$,

4. $\mathrm{tr}\,(\boldsymbol{\varepsilon}^{tr}) = 0$ due to the volume-preserving assumption for the transformation strain.

For example:

$$W_{ijkl} = \frac{1}{2}\frac{\partial^3 G}{\partial\sigma_{ij}\,\partial\xi_{,k}\,\partial\xi_{,l}} = W_1\Big(\delta_{ij}\delta_{kl} + \delta_{ik}\delta_{jl} + \delta_{il}\delta_{jk}\Big) \quad , \quad C_{ij} = \frac{1}{2}\frac{\partial^2 G}{\partial\xi_{,i}\,\partial\xi_{,j}} = c_1\Big(\delta_{ij}\Big)$$

The constitutive variables and generalized forces are similarly given by equations (4.5) and (4.8). The response of the SMA in this section is supposed to be isotropic; i.e. any directional dependence in the thermoelastic response, transformation surface, and also evolution of martensitic volume fraction and thus transformation strain is neglected.

Hence in equation 4.16, the tensors Ξ, being of odd rank, vanish and the tensors Λ can be rewritten in terms of the fundamental isotropic tensors.

$$\widetilde{\varphi}(\boldsymbol{\varsigma}^{'D}) = H\left(\frac{3}{2}\varsigma_{ij}^{'}\varsigma_{ij}^{'}\right)^{\frac{1}{2}} \quad , \quad \bar{\varphi}(\bar{\boldsymbol{\pi}}) = -\left(\frac{\bar{\boldsymbol{\pi}}\cdot\bar{\boldsymbol{\pi}}}{\ell_{d_1}^2}\right)^{\frac{1}{2}} \tag{4.40}$$

$$\hat{\varphi}(\boldsymbol{\tau}) = -\left(\frac{1}{\ell_{d_2}^2}\tau_{ijk}\tau_{ijk} + \frac{1}{\ell_{d_3}^2}\tau_{ijk}\tau_{kji} + \frac{1}{\ell_{d_4}^2}\tau_{jji}\tau_{kki} + \frac{1}{\ell_{d_5}^2}\tau_{iik}\tau_{kjj} + \frac{1}{\ell_{d_6}^2}\tau_{kii}\tau_{kjj}\right)^{\frac{1}{2}}$$

The six constants ℓ_d can be regarded as *dissipative* length scales. Hence the forward and reverse transformation surface will be

$$\Phi^{\mathrm{fwd}}\left(\boldsymbol{\varsigma}^{'D},\ \pi,\ \bar{\boldsymbol{\pi}},\ \boldsymbol{\tau}\right) = \phi - Y \quad , \quad \Phi^{\mathrm{rev}}\left(\boldsymbol{\varsigma}^{'D},\ \pi,\ \bar{\boldsymbol{\pi}},\ \boldsymbol{\tau}\right) = -\phi - Y$$

$$\phi = \widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}}) + \pi + \bar{\varphi}(\bar{\boldsymbol{\pi}}) + \hat{\varphi}(\boldsymbol{\tau}) \tag{4.41}$$

As a result, the rates of internal variables given in (4.14) are simplified to:

$$\begin{aligned}
\dot{\xi} &= \lambda \quad , \quad \nabla\dot{\xi} = -\lambda \, \frac{1}{|\ell_{d_1}|} \frac{\bar{\boldsymbol{\pi}}}{\sqrt{\bar{\boldsymbol{\pi}} \cdot \bar{\boldsymbol{\pi}}}} \\
\dot{\boldsymbol{\varepsilon}}^{tr} &= \lambda \, \frac{3}{2} H \frac{\boldsymbol{\varsigma}^{\acute{D}}}{\sqrt{\frac{3}{2}\boldsymbol{\varsigma}^{\acute{D}} : \boldsymbol{\varsigma}^{\acute{D}}}} \\
\dot{\nabla\boldsymbol{\varepsilon}^{tr}} &= -\frac{1}{2}\lambda \, \frac{\frac{2}{\ell_{d_2}^2}\tau_{ijk} + \frac{1}{\ell_{d_3}^2}\left(\tau_{kji} + \tau_{kij}\right) + \frac{1}{\ell_{d_5}^2}\delta_{ii}\tau_{knn} + \frac{1}{\ell_{d_6}^2}\left(\delta_{ik}\tau_{jnn} + \delta_{jk}\tau_{inn}\right)}{\left(\frac{1}{\ell_{d_2}^2}\tau_{ijk}\tau_{ijk} + \frac{1}{\ell_{d_3}^2}\tau_{ijk}\tau_{kji} + \frac{1}{\ell_{d_4}^2}\tau_{jji}\tau_{kki} + \frac{1}{\ell_{d_5}^2}\tau_{iik}\tau_{kjj} + \frac{1}{\ell_{d_6}^2}\tau_{kii}\tau_{kjj}\right)^{\frac{1}{2}}}
\end{aligned} \tag{4.42}$$

4.3 The simplified SMA gradient model including $\nabla\xi$ and $\nabla\boldsymbol{\varepsilon}^{tr}$: model I

In this section, a simplified version of the general isotropic SMA model introduced in section 4.2 is considered. The free energy is decomposed into local and nonlocal parts with no coupling between the local state variables and the nonlocal internal state variable. For the nonlocal part of the free energy, only three quadratic terms including ξ_i and $\varepsilon^{tr}_{ij,k}$ are retained. The Gibbs free energy in equation (4.39) is simplified to:

$$G(\boldsymbol{\sigma}, T, \boldsymbol{\varepsilon^{tr}}, \xi, \nabla\xi, \nabla\boldsymbol{\varepsilon}^{tr}) = G^{\text{local}}(\boldsymbol{\sigma}, T, \boldsymbol{\varepsilon^{tr}}, \xi) + G^{\text{nonlocal}}(\nabla\xi, \nabla\boldsymbol{\varepsilon}^{tr}) \tag{4.43}$$

The local and nonlocal parts are given as below.

$$
\begin{aligned}
G^{\text{local}}(\boldsymbol{\sigma}, T, \boldsymbol{\varepsilon}^{tr}, \xi) = & -\frac{1}{2\rho}\boldsymbol{\sigma} : \boldsymbol{S}(\xi)\boldsymbol{\sigma} - \frac{1}{\rho}\boldsymbol{\sigma} : \left[\boldsymbol{\alpha}(\xi)(T - T_0) + \boldsymbol{\varepsilon}^{tr}\right] \\
& + c(\xi)\left[(T - T_0) - T\ln\left(\frac{T}{T_0}\right)\right] - s_0(\xi)T + u_0(\xi) + \frac{1}{\rho}f(\xi) \\
G^{\text{nonlocal}}(\nabla\xi, \nabla\boldsymbol{\varepsilon}^{tr}) = & \frac{1}{2\rho}\Big(a_1\ell_1^2\ \xi_i\ \xi_i + a_2\ell_2^2\ \varepsilon^{tr}_{ij,k}\ \varepsilon^{tr}_{ij,k} + a_3\ell_3^2\ \xi_{,i}\ \varepsilon^{tr}_{ij,j}\Big)
\end{aligned}
\tag{4.44}
$$

$\boldsymbol{S}(\xi)$ is the phase-dependant isotropic fourth-order compliance tensor defined linearly (rule of mixtures) with respect to the compliance of the austenite $\boldsymbol{S}^{\text{A}}$ and martensite $\boldsymbol{S}^{\text{M}}$ phases. $\boldsymbol{\alpha}(\xi)$ is the phase dependent effective coefficient of thermal expansion. The material parameters c, s_0, and u_0 are the effective specific heat, effective specific entropy at a reference state, and the effective specific internal energy at the reference state, respectively. They, too, follow a rule of mixture with respect to their corresponding value for austenite and martensite phases; i.e.

$$
\begin{aligned}
\boldsymbol{S} &= \boldsymbol{S}^{\text{A}} + \left(\boldsymbol{S}^{\text{M}} - \boldsymbol{S}^{\text{A}}\right)\xi \quad , \\
\boldsymbol{\alpha} &= \boldsymbol{\alpha}^{\text{A}} + \left(\boldsymbol{\alpha}^{\text{M}} - \boldsymbol{\alpha}^{\text{A}}\right)\xi \quad , \quad s_0 = s_0^{\text{A}} + \left(s_0^{\text{M}} - s_0^{\text{A}}\right)\xi \\
u_0 &= u_0^{\text{A}} + \left(u_0^{\text{M}} - u_0^{\text{A}}\right)\xi \quad , \quad c = c^{\text{A}} + \left(c^{\text{M}} - c^{\text{A}}\right)\xi
\end{aligned}
\tag{4.45}
$$

In addition, $f(\xi)$ is the hardening function attributable to the obstacles inhibiting the propagation of the transformation phase front, affecting the phenomenological stress-strain response of the SMA. $f(\xi)$ can take various forms, form linear to smooth hardening, as discussed in [89] and [90]. The *energetic* length scale ℓ_1 to ℓ_3 are introduced here with corresponding constant coefficients a_1 to a_3 (with dimensions of energy per volume). The material constants included in the local part of the free

energy can be related to the common SMA properties such as the transformation temperatures and the Clausius-Clapeyron slopes on the stress-temperature phase diagram. This will be further discussed in the upcoming sections.

The response of shape memory alloys can be directional or anisotropic in the thermoelastic behaviour, including an anisotropic elastic response or heat conduction and thermal expansion. Also the transformation surface as well as the transformation strain can show anisotropy. As mentioned earlier, an *isotropic* response is assumed for the thermoelastic as well as the trasnformation behaviour of the SMAs in this section.

The constitutive variables and thermodynamic forces can be determined via (4.5) and (4.8) using the free energy form given in (4.44).

$$
\begin{aligned}
\boldsymbol{\varepsilon} &= -\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} = \boldsymbol{S}(\xi)\boldsymbol{\sigma} + \boldsymbol{\alpha}(\xi)\,(T-T_0) + \boldsymbol{\varepsilon}^{tr} \\
s &= -\frac{\partial G}{\partial T} = \frac{1}{\rho}\boldsymbol{\sigma} : \boldsymbol{\alpha} + c\ln\left(\frac{T}{T_0}\right) + s_0 \\
\boldsymbol{\varsigma}^D &= -\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} = \boldsymbol{\sigma} \\
\pi &= -\rho\frac{\partial G}{\partial \xi} = \frac{1}{2}\boldsymbol{\sigma} : \left(\boldsymbol{S}^M - \boldsymbol{S}^A\right)\boldsymbol{\sigma} + \boldsymbol{\sigma} : \left(\boldsymbol{\alpha}^M - \boldsymbol{\alpha}^A\right)(T-T_0) \\
&\quad - \rho(c^M - c^A)\left[(T-T_0) - T\ln\left(\frac{T}{T_0}\right)\right] + \rho\Delta s_0 T - \rho\Delta u_0 - \frac{\partial f}{\partial \xi} \\
\bar{\pi}_i &= -\rho\frac{\partial G}{\partial \xi_i} = -a_1\ell_1^2\,\xi_i - \frac{1}{2}a_3\ell_3^2\,\varepsilon^{tr}_{ij,j} \\
\tau_{ij,k} &= -\rho\frac{\partial G}{\partial \varepsilon^{tr}_{ij,k}} = -a_2\ell_2^2\,\varepsilon^{tr}_{ij,k} - \frac{1}{4}a_3\ell_3^2\;(\delta_{ij}\,\xi_k + \delta_{ik}\,\xi_j + \delta_{jk}\,\xi_i)
\end{aligned}
\tag{4.46}
$$

Also, the transformation surfaces in equation (4.40) are simplified to:

$$\Phi^{\text{fwd}} = \phi - Y \quad , \quad \Phi^{\text{rev}} = -\phi - Y \quad , \quad \phi = \widetilde{\varphi}(\boldsymbol{\varsigma}^{'D}) + \breve{\varphi}(\pi) + \bar{\varphi}(\bar{\boldsymbol{\pi}}) + \hat{\varphi}(\boldsymbol{\tau})$$

$$\widetilde{\varphi}(\boldsymbol{\varsigma}^{'D}) = H\left(\frac{3}{2}\varsigma_{ij}^{'}\varsigma_{ij}^{'}\right)^{\frac{1}{2}} \quad , \quad \bar{\varphi}(\bar{\boldsymbol{\pi}}) = -\left(\frac{1}{\ell_{d_1}^2}\bar{\pi}_i\bar{\pi}_i\right)^{\frac{1}{2}} \quad , \quad \hat{\varphi}(\boldsymbol{\tau}) = -\left(\frac{1}{\ell_{d_2}^2}\tau_{ijk}\tau_{ijk}\right)^{\frac{1}{2}} \tag{4.47}$$

which can be rewritten by substituting for the generalized forces determined in (4.46).

$$\begin{aligned}
\widetilde{\varphi}(\boldsymbol{\varsigma}^{'D}) &= H\left(\frac{3}{2}\sigma_{ij}^{'}\sigma_{ij}^{'}\right)^{\frac{1}{2}} \\
\bar{\varphi}(\bar{\boldsymbol{\pi}}) &= -\frac{1}{\ell_{d_1}}\left[\left(a_1\ell_1^2\right)^2\ \xi_i\xi_i + \frac{1}{4}\left(a_3\ell_3^2\right)^2\ \varepsilon_{ij,j}^{tr}\varepsilon_{ik,k}^{tr} + a_1\ell_1^2 a_3\ell_3^2\,\xi_i\,\varepsilon_{ij,j}^{tr}\right]^{\frac{1}{2}} \\
\hat{\varphi}(\boldsymbol{\tau}) &= -\frac{1}{\ell_{d_2}}\left[\left(a_2\ell_2^2\right)^2\ \varepsilon_{ij,k}^{tr}\varepsilon_{ij,k}^{tr} + \frac{15}{16}\left(a_3\ell_3^2\right)^2\ \xi_i\xi_i + a_2\ell_2^2 a_3\ell_3^2\,\xi_i\,\varepsilon_{ij,j}^{tr}\right]^{\frac{1}{2}}
\end{aligned} \tag{4.48}$$

In which $\sigma_{ij}^{'}$ are the components of the deviatoric stress and the newly introduced dissipative length scales, ℓ_{d_1} and ℓ_{d_2}, are assumed to be positive. As for the nonlocal constants, this model has three ones introduced through the Gibbs free energy and two through the transformation surface. It can be seen that the incorporation of $\varepsilon_{ij,k}^{tr}$ in the Gibbs free energy and subsequently its corresponding thermodynamic force, $\tau_{ij,k}$, in the transformation surface results in identical contributions to the transformation surface, however, through different material constants (see equation 4.48 a and b).

The rates of the internal variables can thus be obtained.

$$
\begin{aligned}
\dot{\varepsilon^{tr}_{ij}} &= \frac{3}{2} H \frac{\acute{\sigma_{ij}}}{\sqrt{\frac{3}{2}\acute{\sigma_{ij}}\acute{\sigma_{ij}}}}\,\dot{\xi} \\
\dot{\xi}_{,i} &= \frac{1}{\ell_{d_1}} \frac{a_1 \ell_1^2\, \xi_i + \frac{1}{2} a_3 \ell_3^2\, \varepsilon^{tr}_{ij,j}}{\sqrt{\overline{\pi}_i \overline{\pi}_i}}\,\dot{\xi} \\
\dot{\varepsilon^{tr}_{ij,k}} &= \frac{1}{\ell_{d_2}} \frac{a_2 \ell_2^2\, \varepsilon^{tr}_{ij,k} + \frac{1}{4} a_3 \ell_3^2\, \left(\delta_{ij}\, \xi_k + \delta_{ik}\, \xi_j + \delta_{jk}\, \xi_i\right)}{\sqrt{\tau_{ij,k} \tau_{ij,k}}}\,\dot{\xi}
\end{aligned}
\tag{4.49}
$$

In order to further investigate the contribution form the gradient of transformation, the model is reduced for a one-dimensional problem were only a single component of stress is nonzero; $\sigma_{ij} = \delta_{1i}\delta_{1j}$. In this case, it is possible to directly integrate the equation for evolution of transformation strain, (4.49)a , such that $\varepsilon^{tr}_{11} = H\xi,\ \varepsilon^{tr}_{22} = -\frac{1}{2}H\xi,\ \varepsilon^{tr}_{33} = \frac{1}{2}H\xi$. Hence the surfaces for forward and reverse transformations will become:

$$
H\left|\sigma_{11}\left(x\right)\right| + \pi - \mathcal{M}\left|\frac{\mathrm{d}\xi\left(x\right)}{\mathrm{d}x}\right| = \pm Y \tag{4.50}
$$

where

$$
\mathcal{M} = \frac{a_1 \ell_1^2 + \frac{1}{2} H a_3 \ell_3^2}{\ell_{d_1}} + \frac{1}{\ell_{d_2}} \sqrt{\frac{3}{2}\left(a_2 \ell_2^2\right)^2 H^2 + \frac{15}{16}\left(a_3 \ell_3^2\right)^2 + a_2 \ell_2^2 a_3 \ell_3^2 H^2} \tag{4.51}
$$

This shows that the 1D transformation surface, as a result of the nonlocal terms in this model, identifies as a first order differential equation and the nonlocal material constants can be grouped in a *nonlocal parameter*, $\mathcal{M}$. Specifically, due to the fact that the transformation strain evolves proportional to the martensite volume fraction,

including $\nabla\boldsymbol{\varepsilon}^{tr}$ results in additional restrictions and constants without changing the form of the transformation surface that needs to be finally solved in conjunction with the equilibrium equation.

Inspired by this outcome, $\nabla\boldsymbol{\varepsilon}^{tr}$ will be excluded form the next simplified SMA gradient model and the focus of the study will be given to the coupling of $\nabla\xi$ with other state variables.

4.4 The simplified SMA gradient model including only terms with $\nabla\xi$: model II

The general isotropic SMA model introduced in section 4.2 is simplified in this section to only include the terms with $\nabla\xi$ without considering $\nabla\boldsymbol{\varepsilon}^{tr}$. The Gibbs free energy cannot be decomposed as before, as the cubic terms provide coupling between the local and nonlocal state variables.

$$\begin{aligned} G(\boldsymbol{\sigma}, T, \boldsymbol{\varepsilon^{tr}}, \xi, \nabla\xi, \nabla\boldsymbol{\varepsilon}^{tr}) = &-\frac{1}{2\rho}\boldsymbol{\sigma} : \boldsymbol{S}(\xi)\boldsymbol{\sigma} - \frac{1}{\rho}\boldsymbol{\sigma} : \left[\boldsymbol{\alpha}(\xi)(T-T_0) + \boldsymbol{\varepsilon}^{tr}\right] \\ &+ c(\xi)\left[(T-T_0) - T\ln\left(\frac{T}{T_0}\right)\right] - s_0(\xi)T + u_0(\xi) + \frac{1}{\rho}f(\xi) \\ &+ \frac{1}{2\rho}\Big[a_1\ell_1^2\,\nabla\xi\cdot\nabla\xi + a_2\ell_2^2\,\nabla\xi\cdot\boldsymbol{\sigma}\nabla\xi + a_3\ell_3^2\,T\,\nabla\xi\cdot\nabla\xi \\ &+ a_4\ell_4^2\,\xi\,\nabla\xi\cdot\nabla\xi + a_5\ell_5^2\,\nabla\xi\cdot\boldsymbol{\varepsilon}^{tr}\nabla\xi + a_6\ell_6^2\,\mathrm{tr}\,(\boldsymbol{\sigma})\;\nabla\xi\cdot\nabla\xi\Big] \end{aligned} \tag{4.52}$$

Six energetic length scales are introduced in this model that attribute to the coupling between $\nabla\xi$ and the rest of the state variables [1]. The material constants a_1, a_4 and a_5 have dimensions of energy per volume, while a_3 has dimension of energy per volume per temperature, and a_2 and a_6 are dimensionless. Because of being originated from the same tensorial constant, there exists a relation between

[1] Notice that the nonlocal length scales here are independent of the ones in (4.44).

the constants coupling $\boldsymbol{\sigma}$ and $\nabla\xi$; i.e. $a_2\ell_2^2 = 2\,a_6\ell_6^2$. The material constants, listed in equation (4.45), follow a similar linear rule of mixture.

Identically, the constitutive variables and thermodynamic forces can be determined via (4.5) and (4.8) using the free energy form given in (4.52).

$$
\begin{aligned}
\boldsymbol{\varepsilon} =& -\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} = \boldsymbol{S}(\xi)\boldsymbol{\sigma} + \boldsymbol{\alpha}(\xi)\,(T-T_0) + \boldsymbol{\varepsilon}^{tr} \\
& - \frac{1}{2}\Big[a_2\ell_2^2\ (\nabla\xi\otimes\nabla\xi) + a_6\ell_6^2\ (\nabla\xi\cdot\nabla\xi)\ \mathbf{I}\Big] \\
s =& -\frac{\partial G}{\partial T} = \frac{1}{\rho}\boldsymbol{\sigma} : \boldsymbol{\alpha} + c\ln\left(\frac{T}{T_0}\right) - \frac{1}{2\rho}a_3\ell_3^2\ (\nabla\xi\cdot\nabla\xi) + s_0 \\
\boldsymbol{\varsigma}^D =& -\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} = \boldsymbol{\sigma} - \frac{1}{2}a_5\ell_5^2\ (\nabla\xi\otimes\nabla\xi) \\
\pi =& -\rho\frac{\partial G}{\partial \xi} = \frac{1}{2}\boldsymbol{\sigma} : \left(\boldsymbol{S}^M - \boldsymbol{S}^A\right)\boldsymbol{\sigma} + \boldsymbol{\sigma} : \left(\boldsymbol{\alpha}^M - \boldsymbol{\alpha}^A\right)(T-T_0) \\
& - \rho(c^M - c^A)\left[(T-T_0) - T\ln\left(\frac{T}{T_0}\right)\right] + \rho\Delta s_0 T - \rho\Delta u_0 - \frac{\partial f}{\partial \xi} \\
& - \frac{1}{2}a_4\ell_4^2\ (\nabla\xi\cdot\nabla\xi) \\
\bar{\pi}_i =& -\rho\frac{\partial G}{\partial \xi_i} = -a_1\ell_1^2\,\nabla\xi - a_2\ell_2^2\,\boldsymbol{\sigma}\nabla\xi - a_3\ell_3^2\,T\,\nabla\xi \\
& - a_4\ell_4^2\,\xi\,\nabla\xi - a_5\ell_5^2\,\boldsymbol{\varepsilon}^{tr}\nabla\xi - a_6\ell_6^2\,\mathrm{tr}\,(\boldsymbol{\sigma})\ \nabla\xi
\end{aligned}
\tag{4.53}
$$

The transformation surfaces will be:

$$
\begin{aligned}
&\Phi^{\mathrm{fwd}} = \phi - Y \quad , \quad \Phi^{\mathrm{rev}} = -\phi - Y \quad , \quad \phi = \widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}}) + \check{\varphi}(\pi) + \bar{\varphi}(\bar{\boldsymbol{\pi}}) \\
&\widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}}) = H\left(\frac{3}{2}\acute{\varsigma}_{ij}\acute{\varsigma}_{ij}\right)^{\frac{1}{2}} \quad , \quad \bar{\varphi}(\bar{\boldsymbol{\pi}}) = -\left(\frac{1}{\ell_{d_1}^2}\bar{\pi}_i\bar{\pi}_i\right)^{\frac{1}{2}}
\end{aligned}
\tag{4.54}
$$

which can be rewritten by substituting for the generalized forces determined in (4.53).

The rates of the internal variables can thus be obtained.

$$\dot{\boldsymbol{\varepsilon}}^{tr} = \frac{3}{2} H \frac{\boldsymbol{\varsigma}^{\acute{D}}}{\sqrt{\frac{3}{2}\boldsymbol{\varsigma}^{\acute{D}} : \boldsymbol{\varsigma}^{\acute{D}}}}\, \dot{\xi} \quad , \quad \nabla\dot{\xi} = -\frac{1}{\ell_{d_1}} \frac{\bar{\boldsymbol{\pi}}}{\sqrt{\bar{\boldsymbol{\pi}} \cdot \bar{\boldsymbol{\pi}}}}\, \dot{\xi} \tag{4.55}$$

in which

$$\begin{aligned} \boldsymbol{\varsigma}^{\acute{D}} &= \boldsymbol{\acute{\sigma}} - \frac{1}{6} a_5 \ell_5^2 \Big[3 \left(\nabla\xi \otimes \nabla\xi \right) - \left(\nabla\xi \cdot \nabla\xi \right) \mathbf{I} \Big] \\ \boldsymbol{\varsigma}^{\acute{D}} : \boldsymbol{\varsigma}^{\acute{D}} &= \boldsymbol{\acute{\sigma}} : \boldsymbol{\acute{\sigma}} - a_5 \ell_5^2 \, \nabla\xi \cdot \boldsymbol{\acute{\sigma}} \nabla\xi + \frac{1}{6} \Big[a_5 \ell_5^2 \, \left(\nabla\xi \cdot \nabla\xi \right) \Big]^2 \end{aligned} \tag{4.56}$$

The rate of martensite volume fraction is determined through the consistency condition; $\dot{\Phi} = 0$ or (4.19) repeated here.

$$\dot{\xi} = -\frac{\frac{\partial \phi}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} + \frac{\partial \phi}{\partial T} \dot{T}}{\frac{\partial \phi}{\partial \xi} + \frac{\partial \phi}{\partial \boldsymbol{\varepsilon}^{tr}} : \frac{\partial \phi}{\partial \boldsymbol{\varsigma}^{D}} + \frac{\partial \phi}{\partial \nabla\xi} \cdot \frac{\partial \phi}{\partial \bar{\boldsymbol{\pi}}} + \frac{\partial \phi}{\partial \nabla \boldsymbol{\varepsilon}^{tr}} \cdot \frac{\partial \phi}{\partial \boldsymbol{\tau}}} = -\frac{\frac{\partial \phi}{\partial \boldsymbol{\sigma}} : \dot{\boldsymbol{\sigma}} + \frac{\partial \phi}{\partial T} \dot{T}}{\mathcal{A}} \tag{4.57}$$

in which we have:

$$\begin{aligned} \frac{\partial \phi}{\partial \boldsymbol{\sigma}} =& \frac{3}{2} H \frac{\boldsymbol{\varsigma}^{\acute{D}}}{\sqrt{\frac{3}{2}\boldsymbol{\varsigma}^{\acute{D}} : \boldsymbol{\varsigma}^{\acute{D}}}} + \frac{a_2 \ell_2^2}{\ell_{d_1}} \Big[\frac{\bar{\boldsymbol{\pi}} \otimes \nabla\xi + \frac{1}{2} \left(\bar{\boldsymbol{\pi}} \cdot \nabla\xi \right) \mathbf{I}}{\sqrt{\bar{\boldsymbol{\pi}} \cdot \bar{\boldsymbol{\pi}}}} \Big] \\ & + \left(\boldsymbol{S}^M - \boldsymbol{S}^A \right) \boldsymbol{\sigma} + \left(\boldsymbol{\alpha}^M - \boldsymbol{\alpha}^A \right) \left(T - T_0 \right) \\ \frac{\partial \phi}{\partial T} =& \boldsymbol{\sigma} : \left(\boldsymbol{\alpha}^M - \boldsymbol{\alpha}^A \right) + \rho (c^M - c^A) \ln \left(\frac{T}{T_0} \right) + \rho \Delta s_0 + \frac{a_3 \ell_3^2}{\ell_{d_1}} \frac{\bar{\boldsymbol{\pi}} \cdot \nabla\xi}{\sqrt{\bar{\boldsymbol{\pi}} \cdot \bar{\boldsymbol{\pi}}}} \end{aligned} \tag{4.58}$$

and

$$\begin{aligned}\mathcal{A} =& 3H\frac{a_5\ell_5^2}{\ell_{d_1}}\,\frac{1}{\sqrt{\bar{\boldsymbol{\pi}}\cdot\bar{\boldsymbol{\pi}}}}\,\frac{1}{\sqrt{\frac{3}{2}\boldsymbol{\varsigma}^{'D}:\boldsymbol{\varsigma}^{'D}}}\,\bar{\boldsymbol{\pi}}\cdot\Big[\boldsymbol{\acute{\sigma}}\nabla\xi-\frac{1}{3}a_5\ell_5^2\,\left(\nabla\xi\cdot\nabla\xi\right)\,\nabla\xi\Big]\\ &+2\frac{a_4\ell_4^2}{\ell_{d_1}}\,\frac{1}{\sqrt{\bar{\boldsymbol{\pi}}\cdot\bar{\boldsymbol{\pi}}}}\,\left(\bar{\boldsymbol{\pi}}\cdot\nabla\xi\right)\\ &-\frac{1}{\ell_{d_1}}\Big[\left(a_1\ell_1^2+a_3\ell_3^2\,T+a_4\ell_4^2\,\xi+a_6\ell_6^2\,\mathrm{tr}\left(\boldsymbol{\sigma}\right)\right)+\frac{a_2\ell_2^2\,\boldsymbol{\sigma}+a_5\ell_5^2\,\boldsymbol{\varepsilon}^{tr}}{\bar{\boldsymbol{\pi}}\cdot\bar{\boldsymbol{\pi}}}:\left(\bar{\boldsymbol{\pi}}\otimes\bar{\boldsymbol{\pi}}\right)\Big]\\ &-\frac{\partial^2 f}{\partial\xi^2}\end{aligned}\tag{4.59}$$

It shows that the evolution of martensite volume fraction, (4.57), is effected by the heterogeneity in its distribution coupled with the temperature and also the stress tensor.

The model presented in this section contains a total of six independent nonlocal constants adding to the complexity of the problem as for as parameter identification and model calibration. In the next section, a simplified version of this model is studied by retaining only ℓ_1 and ℓ_{d_1}.

4.5 The simplest SMA gradient model including a quadratic term in $\nabla\xi$: model III

The most basic SMA gradient model, discussed in this section, includes the norm of $\nabla\xi$ in the free energy. This can be viewed as the simplification of either of the models discussed before by only retaining the first length scale.

$$G(\boldsymbol{\sigma},T,\boldsymbol{\varepsilon^{tr}},\xi,\nabla\xi)=G^{\mathrm{local}}(\boldsymbol{\sigma},T,\boldsymbol{\varepsilon^{tr}},\xi)+G^{\mathrm{nonlocal}}(\nabla\xi)\tag{4.60}$$

with the local and nonlocal parts given as:

$$
\begin{aligned}
G^{\text{local}}(\boldsymbol{\sigma}, T, \boldsymbol{\varepsilon}^{tr}, \xi) &= -\frac{1}{2\rho}\boldsymbol{\sigma} : \boldsymbol{S}(\xi)\boldsymbol{\sigma} - \frac{1}{\rho}\boldsymbol{\sigma} : \left[\boldsymbol{\alpha}(T - T_0) + \boldsymbol{\varepsilon}^{tr}\right] \\
&\quad + c\left[(T - T_0) - T\ln\left(\frac{T}{T_0}\right)\right] - s_0(\xi)T + u_0(\xi) + \frac{1}{\rho}f(\xi) \qquad (4.61) \\
G^{\text{nonlocal}}(\nabla\xi) &= \frac{1}{2\rho}a_1\ell_1^2\,\nabla\xi\cdot\nabla\xi
\end{aligned}
$$

The conventional local material constants and nonlocal terms are defined similar to the ones in sections 4.3 and 4.4 (also see equation (4.45)).

The constitutive relations and generalized thermodynamic forces will be:

$$
\begin{aligned}
\boldsymbol{\varepsilon} &= -\rho\frac{\partial G}{\partial \boldsymbol{\sigma}} = \boldsymbol{S}(\xi)\boldsymbol{\sigma} + \boldsymbol{\alpha}(\xi)\,(T - T_0) + \boldsymbol{\varepsilon}^{tr} \\
s &= -\frac{\partial G}{\partial T} = \frac{1}{\rho}\boldsymbol{\sigma} : \boldsymbol{\alpha} + c\ln\left(\frac{T}{T_0}\right) + s_0 \\
\boldsymbol{\varsigma}^D &= -\rho\frac{\partial G}{\partial \boldsymbol{\varepsilon}^{tr}} = \boldsymbol{\sigma} \\
\pi &= -\rho\frac{\partial G}{\partial \xi} = \frac{1}{2}\boldsymbol{\sigma} : \left(\boldsymbol{S}^M - \boldsymbol{S}^A\right)\boldsymbol{\sigma} + \boldsymbol{\sigma} : \left(\boldsymbol{\alpha}^M - \boldsymbol{\alpha}^A\right)(T - T_0) \\
&\quad - \rho(c^M - c^A)\left[(T - T_0) - T\ln\left(\frac{T}{T_0}\right)\right] + \rho\Delta s_0 T - \rho\Delta u_0 - \frac{\partial f}{\partial \xi} \\
\bar{\pi}_i &= -\rho\frac{\partial G}{\partial \xi_i} = -a_1\ell_1^2\,\nabla\xi
\end{aligned}
\qquad (4.62)
$$

For the purpose of model III, the transformation surfaces $\Phi(\boldsymbol{\Gamma})$ are slightly modified for the forward and reverse processes.

$$
\Phi^{\text{fwd}} = \phi^{\text{fwd}} - Y \quad , \quad \Phi^{\text{rev}} = -\phi^{\text{rev}} - Y \qquad (4.63)
$$

$$
\phi^{\text{fwd/rev}} = \widetilde{\varphi}(\boldsymbol{\varsigma}^{'D}) + \check{\varphi}(\pi) + \bar{\varphi}^{\text{fwd/rev}}(\bar{\boldsymbol{\pi}}) \qquad (4.64)
$$

$$\widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}}) = H\sqrt{3J_2}, \quad \bar{\varphi}^{\text{fwd}}(\bar{\boldsymbol{\pi}}) = -\frac{1}{\ell_2^{\text{fwd}}}\sqrt{\bar{\boldsymbol{\pi}}\cdot\bar{\boldsymbol{\pi}}}, \quad \bar{\varphi}^{\text{rev}}(\bar{\boldsymbol{\pi}}) = -\frac{1}{\ell_2^{\text{rev}}}\sqrt{\bar{\boldsymbol{\pi}}\cdot\bar{\boldsymbol{\pi}}} \tag{4.65}$$

J_2 is the second invariant of the deviatoric part of the generalized stress tensor conjugate to transformation strain, $\boldsymbol{\varsigma}^{\acute{D}} = \boldsymbol{\varsigma}^{D} - \frac{1}{3}tr(\boldsymbol{\varsigma}^{D})$ and the constant H is the maximum achievable transformation strain. Therefore, $\widetilde{\varphi}$ is a Mises type transformation surface. ℓ_2, introduced based on dimensional grounds, can be regarded as the *dissipative* length scale that has different values during the forward or reverse transformation processes; hence ℓ_2^{fwd} and ℓ_2^{rev}.

After expansion of equation (4.65) using equation (4.62), the nonlocal transformation surface for the forward and reverse transformation will be

$$\widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}}) + \pi + \bar{\varphi}^{\text{fwd}}(\bar{\boldsymbol{\pi}}) - Y = 0, \quad \dot{\xi} > 0$$

$$\begin{aligned}\Phi^{\text{fwd}} =& H\sqrt{\frac{3}{2}\acute{\boldsymbol{\sigma}} : \acute{\boldsymbol{\sigma}}} + \left[\frac{1}{2}\boldsymbol{\sigma} : \left(\boldsymbol{S}^{M} - \boldsymbol{S}^{A}\right)\boldsymbol{\sigma} + \rho\Delta s_0\, T - \rho\Delta u_0 - \frac{\partial f}{\partial \xi}\right] \\ & - \frac{a_1\ell_1^2}{\ell_2^{\text{fwd}}}\sqrt{\nabla\xi\cdot\nabla\xi} - Y = 0\end{aligned} \tag{4.66}$$

$$\widetilde{\varphi}(\boldsymbol{\varsigma}^{\acute{D}}) + \pi + \varphi^{\bar{\text{rev}}}(\bar{\boldsymbol{\pi}}) + Y = 0, \quad \dot{\xi} < 0$$

$$\begin{aligned}\Phi^{\text{rev}} =& - H\sqrt{\frac{3}{2}\acute{\boldsymbol{\sigma}} : \acute{\boldsymbol{\sigma}}} - \left[\frac{1}{2}\boldsymbol{\sigma} : \left(\boldsymbol{S}^{M} - \boldsymbol{S}^{A}\right)\boldsymbol{\sigma} + \rho\Delta s_0\, T - \rho\Delta u_0 - \frac{\partial f}{\partial \xi}\right] \\ & + \frac{a_1\ell_1^2}{\ell_2^{\text{rev}}}\sqrt{\nabla\xi\cdot\nabla\xi} - Y = 0\end{aligned} \tag{4.67}$$

The difference in coefficients of thermal expansion and specific heat between

martensite and austenite phases is ignored in deriving the above equations; i.e. $c^A = c^M$ and $\boldsymbol{\alpha}^A = \boldsymbol{\alpha}^M$. As observed, the transformation surface for the nonlocal model is the solution of a differential equation.

The energetic and dissipative intrinsic length scales, ℓ_1 and ℓ_2, are combined in a new material parameter, the *nonlocal parameter* $\mathcal{M}$. $\mathcal{M}$ has dimensions of length×energy/volume. Appearing in the transformation surface and used later on for the analysis of various SMA structures, $\mathcal{M}$ is given by:

$$\mathcal{M}^{\text{fwd/rev}} \equiv a_1 \frac{\ell_1^2}{\ell_2^{\text{fwd/rev}}} \tag{4.68}$$

A connection can be made to the nonlocal parameter given in equation (4.51) by discarding the extra length scales.

The rates of evolution for internal variables can also be obtained.

$$\dot{\boldsymbol{\varepsilon}}^{tr} = \frac{3}{2} H \frac{\acute{\boldsymbol{\sigma}}}{\sqrt{\frac{3}{2}\acute{\boldsymbol{\sigma}} : \acute{\boldsymbol{\sigma}}}} \dot{\xi} = \Lambda(\boldsymbol{\sigma})\dot{\xi} \quad , \quad \nabla\dot{\xi} = \frac{1}{\ell_2} \frac{\nabla\xi}{\sqrt{\nabla\xi \cdot \nabla\xi}} \dot{\xi} \tag{4.69}$$

Based on this, the gradient of martensite volume fraction evolves only in the direction of the current gradient in ξ. This is the outcome of the isotropic forms and the extent of coupling between variables assumed for the free energy, (4.60)-(4.61), as well as the transformation surface, (4.65). It can be shown that in case of a rate-dependent formulation with linear constitutive equations relating the generalized thermodynamic forces to the fluxes, the restriction imposed from isotropy lead to a similar relationship for the rate of martensite volume fraction (see section 3). However, notice that this limitation is not the case for the model in section 4.4 due to the coupling terms with the temperature and stress tensor.

Also, the consistency condition during forward or reverse transformation (fwd/rev) can be shown to result in:

$$\begin{aligned}\left[\Lambda(\boldsymbol{\sigma}) + \left(\boldsymbol{S}^M - \boldsymbol{S}^A\right)\boldsymbol{\sigma}\right] : \dot{\boldsymbol{\sigma}} + \rho\Delta s_0\,\dot{T} \\ - \left[\frac{\partial^2 f^{\text{fwd/rev}}}{\partial\xi^2} + a_1\left(\frac{\ell_1}{\ell_2^{\text{fwd/rev}}}\right)^2\right]\dot{\xi}^{\text{fwd/rev}} = 0\end{aligned} \tag{4.70}$$

As before, the gradient of martensite volume fraction does not explicitly appear in the consistency condition. In order to describe the effect of $\nabla\xi$ in the present constitutive model, equations (4.69) and (4.70) are re-written below:

$$\begin{aligned}\dot{\xi}^{\text{fwd/rev}} &= \frac{\left[\Lambda(\boldsymbol{\sigma}) + \left(\boldsymbol{S}^M - \boldsymbol{S}^A\right)\boldsymbol{\sigma}\right] : \dot{\boldsymbol{\sigma}} + \rho\Delta s_0\,\dot{T}}{\frac{\partial^2 f^{\text{fwd/rev}}}{\partial\xi^2} + \frac{\mathcal{M}^{\text{fwd/rev}}}{\ell_2^{\text{fwd/rev}}}} \\ \nabla\dot{\xi}^{\text{fwd/rev}} &= \frac{1}{\ell_2^{\text{fwd/rev}}}\frac{\nabla\xi}{\sqrt{\nabla\xi\cdot\nabla\xi}}\,\dot{\xi}\end{aligned} \tag{4.71}$$

Both of the above equations have to be simultaneously satisfied when the martensite volume fraction ξ is evolved as a field variable. The value for the gradient, $\nabla\xi$ though, directly influences the transformation surfaces; as stated in (4.66) and (4.67).

To consummate the modeling section, it is necessary to introduce the hardening function. A linear form can be chosen for the hardening function $f(\xi)$ with the relevant material constants b^M, b^A, μ_1, and μ_2 :

$$f(\xi) = \begin{cases} \frac{1}{2}\rho b^M\xi^2 & + (\mu_1 + \mu_2)\,\xi, \quad \dot{\xi} > 0 \\ \\ \frac{1}{2}\rho b^A\xi^2 & + (\mu_1 - \mu_2)\,\xi, \quad \dot{\xi} < 0 \end{cases} \tag{4.72}$$

In addition, a smooth hardening function can be introduced that results in a continuous transition between the linear elastic and transforming regimes, for example, in a uniaxial stress-strain loading plot. It's differentiated form is given by:

$$\frac{\partial f(\xi)}{\partial \xi} = \begin{cases} \frac{1}{2} b_1 \Big[1 + \xi^{n_1} - (1-\xi)^{n_2} \Big] + b_3, & \dot{\xi} > 0 \\ \\ \frac{1}{2} b_2 \Big[1 + \xi^{n_3} - (1-\xi)^{n_4} \Big] - b_3, & \dot{\xi} < 0 \end{cases} \tag{4.73}$$

where n_1 to n_4 and b_1 to b_3 are constants.

Finally, it is interesting, as for implementation of the nonlocal model in computational frameworks, to determine the continuum mechanical and thermal tangent moduli. The continuum tangent stiffness tensor, $\mathcal{L}$, and continuum tangent thermal modulus, Θ, are defined as:

$$\mathcal{L} \equiv \frac{\mathrm{d}\boldsymbol{\sigma}}{\mathrm{d}\boldsymbol{\varepsilon}} \quad , \quad \Theta \equiv \frac{\mathrm{d}\boldsymbol{\sigma}}{\mathrm{d}T} \tag{4.74}$$

In order to proceed, the Hook's law in equation (4.62)a is written in an incremental form (again ignoring the difference in coefficients of thermal expansion).

$$\mathrm{d}\boldsymbol{\varepsilon} = \boldsymbol{S}(\xi)\, \mathrm{d}\boldsymbol{\sigma} + \boldsymbol{\alpha}\, \mathrm{d}T + \left[\Lambda(\boldsymbol{\sigma}) + \left(\boldsymbol{S}^M - \boldsymbol{S}^A \right) \boldsymbol{\sigma} \right] \mathrm{d}\xi \tag{4.75}$$

During fwd/rev transformations, the consistency condition (4.70) holds. Hence

$$\mathrm{d}\boldsymbol{\varepsilon} = \left(\frac{\boldsymbol{\mathcal{A}}}{g} + \boldsymbol{S}(\xi) \right) \mathrm{d}\boldsymbol{\sigma} + \left\{ \frac{\rho \Delta s_0}{g} \left[\Lambda(\boldsymbol{\sigma}) + \left(\boldsymbol{S}^M - \boldsymbol{S}^A \right) \boldsymbol{\sigma} \right] + \boldsymbol{\alpha} \right\} \mathrm{d}T \tag{4.76}$$

$$\boldsymbol{\mathcal{A}}(\boldsymbol{\sigma}) = \left[\Lambda(\boldsymbol{\sigma}) + \left(\boldsymbol{S}^M - \boldsymbol{S}^A\right)\boldsymbol{\sigma}\right] \otimes \left[\Lambda(\boldsymbol{\sigma}) + \left(\boldsymbol{S}^M - \boldsymbol{S}^A\right)\boldsymbol{\sigma}\right]$$
$$g = \frac{\partial^2 f}{\partial \xi^2} + \frac{\mathcal{M}}{\ell_2} \tag{4.77}$$

in which $\boldsymbol{\mathcal{A}}(\boldsymbol{\sigma})$ is a fourth order tensor with major and minor symmetries and g is a new scalar variable depending on the spatial gradient of the martensite volume fraction. From this equation, though, it is possible to derive the corresponding tangent moduli as below.

$$\begin{aligned}
\mathcal{L} &= \left(\frac{\boldsymbol{\mathcal{A}}}{g} + \boldsymbol{S}(\xi)\right)^{-1} \\
\Theta &= -\left(\frac{\boldsymbol{\mathcal{A}}}{g} + \boldsymbol{S}(\xi)\right)^{-1} \left\{\frac{\rho \Delta s_0}{g}\left[\Lambda(\boldsymbol{\sigma}) + \left(\boldsymbol{S}^M - \boldsymbol{S}^A\right)\boldsymbol{\sigma}\right] + \boldsymbol{\alpha}\right\}
\end{aligned} \tag{4.78}$$

The effect of the nonlocal parameter can be realized by investigating the denominator g and the transformation surfaces (4.66) and (4.67). The hardening in the response of shape memory alloys, related to the distance between the start and finish lines for transformation on the stress-temperature phase diagram, contributes to g through the second derivative of the hardening function $f(\xi)$, which is non-negative. That is in the case of linear hardening, the larger ρb^M or ρb^A, the lower the compliance or the higher the stiffness. Nonlocality has a similar effect; i.e. by increasing the gradients in the spatial distribution of martensite volume fraction surrounding any continuum point, the response of the material at that point shows more stress to continue the transformation. This feature can be useful in dealing with the problem of transformation strain localization and loss of ellipticity encountered in numerical

modleing of SMAs showing a softening behavior.

For the purpose of analyzing various SMA structures in the subsequent sections, reduction of the current model to 1D is presented.

In case of a uniaxial sate of stress with σ_x being the only nonzero component of the stress tensor $\boldsymbol{\sigma}$, the differential equation for the transformation surfaces in equations (4.66) and (4.67) reduce to the following ordinary differential equations.

$$\begin{aligned}\Phi^{\text{fwd}} =& H\,|\sigma| + \left[\frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma^2 + \rho\Delta s_0\,T - \rho\Delta u_0 - \left(\rho b^M \xi + \mu_1 + \mu_2\right)\right] \\ & - \mathcal{M}^{\text{fwd}}\left|\frac{\mathrm{d}\xi}{\mathrm{d}x}\right| - Y = 0\end{aligned} \tag{4.79}$$

$$\begin{aligned}\Phi^{\text{rev}} =& - H\,|\sigma| - \left[\frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma^2 + \rho\Delta s_0\,T - \rho\Delta u_0 - \left(\rho b^A \xi + \mu_1 - \mu_2\right)\right] \\ & + \mathcal{M}^{\text{rev}}\left|\frac{\mathrm{d}\xi}{\mathrm{d}x}\right| - Y = 0\end{aligned} \tag{4.80}$$

with E^A, E^M,ν^A, and ν^M being the elastic moduli and Poisson ratios of austenite and martensite phases, respectively. In the above equations, the linear hardening rule, equation (4.72), is used.

If the smooth hardening rule given in equation (4.73) is used, the transformation surfaces can be written as

$$\Phi^{\mathrm{fwd}} =$$
$$H\left|\sigma\right| + \left[\frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma^2 + \rho\Delta s_0\, T - \rho\Delta u_0 - \frac{1}{2}b_1\Big[1 + \xi^{n_1} - (1-\xi)^{n_2}\Big] - b_3\right]$$
$$- \mathcal{M}^{\mathrm{fwd}}\left|\frac{\mathrm{d}\xi}{\mathrm{d}x}\right| - Y = 0 \tag{4.81}$$

$$\Phi^{\mathrm{rev}} =$$
$$- H\left|\sigma\right| - \left[\frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma^2 + \rho\Delta s_0\, T - \rho\Delta u_0 - \frac{1}{2}b_2\Big[1 + \xi^{n_3} - (1-\xi)^{n_4}\Big] + b_3\right]$$
$$+ \mathcal{M}^{\mathrm{rev}}\left|\frac{\mathrm{d}\xi}{\mathrm{d}x}\right| - Y = 0 \tag{4.82}$$

5. ANALYSIS OF SMA STRUCTURES

The proposed SMA nonlocal constitutive models are used to investigate the response of three different structures commonly encountered in microactuators i.e. wires, thin films and micropillars.

Uniaxial stretching of an SMA homogenous prismatic bar is analytically investigated only to verify that the gradient-based model recovers the results of a conventional SMA model in the case of a uniform deformation. In addition, the torsion of SMA wires is studied in which the variation of stress, and hence the martensitic volume fraction, over the diameter of the wire demonstrates the size dependent features of the gradient-based model.

The compression of an SMA micropillar is modeled as 1D uniaxial compression of an SMA bar with a variable cross section. Similar to to the thin film bending problem, an analytical closed-form solution is derived for the forward transformation during loading and the loading-unloading cycle of the micropillar compression is studied using numerical solutions.

Also, the actuation of an SMA thin film is analytically modeled for forward transformation as a beam under pure bending. The loading-unloading response of the SMA beam under pure bending, including the forward and reverse transformation, is numerically modeled. This analysis is performed for pure torsion of an SMA bar as well.

As models I to III incorporate the gradients of martensitic volume fraction and transformation strain, the response of a homogeneous SMA specimen under a homogeneous state of stress using the proposed gradient model should reduce to the conventional local response. The validity of this statement is investigated in the next

section for uniaxial loading of a 1D SMA bar.

5.1 Uniaxial stretching of an SMA prismatic bar

The purpose of this section is to show that the proposed nonlocal model yields a local response if the material undergoes an entirely homogenous state of deformation. A prismatic bar of SMA is assumed to undergo a uniaxial isothermal loading at a temperature above the austenite finish temperature, A_f. The bar is isothermally loaded from austenite so that the transformation to martensite occurs. The schematic for the 1D boundary value problem, considering the symmetry (at the left end) and boundary conditions, is shown in Figure 9. It is assumed that the state of the bar, including the internal variable martensite volume fraction, is symmetric about the left end.

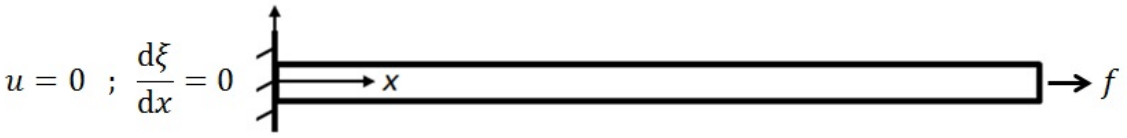


Figure 9: Schematic including boundary conditions for uniaxial loading of an SMA bar.

The kinematic and equilibrium equations, along with the equation for decomposition of strain result in the following.

$$\varepsilon_x = \frac{\partial u}{\partial x} \quad , \quad \frac{\partial \sigma_x}{\partial x} = 0 \quad , \quad \varepsilon_x = \frac{\sigma_x}{E(\xi(x))} + H\xi(x) \tag{5.1}$$

The stress does not vary through the length of the bar. The equation for transformation surface, after equation (4.79), is also rewritten as in:

$$\begin{aligned} H\left|\sigma_x\right| + \left[\frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma_x^2 + \rho\Delta s_0\, T - \rho\Delta u_0 - \left(\rho b^M \xi + \mu_1 + \mu_2\right)\right] \\ - \mathcal{M}^{\text{fwd}}\left|\frac{\mathrm{d}\xi(x)}{\mathrm{d}x}\right| - Y = 0 \end{aligned} \tag{5.2}$$

The constant terms (material properties) in the equation above can be rearranged for brevity (for their definition refer to equation (4.45) and model description in section 4.5). Also, it is assumed that $\sigma_x \geq 0$ along the SMA bar. However, no assumption is made on the sign of $\frac{\mathrm{d}\xi}{\mathrm{d}x}$. Therefore, the transformation differential equation (5.2) reduces to:

$$H\sigma_x + A\sigma_x^2 + B\xi(x) \pm \mathcal{M}^{\text{fwd}}\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} + D = 0 \quad \Rightarrow \quad B\xi(x) \pm \mathcal{M}^{\text{fwd}}\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} = F \tag{5.3}$$

This differential equation for the martensitic volume fraction as the dependent variable has the following general solution.

$$\xi(x) = \frac{F}{B} + \zeta\, \mathrm{e}^{\pm\frac{B}{\mathcal{M}^{\text{fwd}}}x} \tag{5.4}$$

The first term on the r.h.s refers to the local solution and the second term is the contribution from the nonlocal part of the model. Upon satisfying the symmetry boundary condition, $\frac{\mathrm{d}\xi}{\mathrm{d}x}|_{x=0}$ at the left end, the nonlocal part vanishes, $\zeta = 0$. Therefore the martensite volume fraction will only depend on the stress with a constant

homogenous distribution along the bar.

$$\xi\Big|_{\text{fwd}} = \frac{F}{B} = \frac{H\sigma_x + \frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma_x^2 + \rho\Delta s_0\, T - \rho\Delta u_0 - (\mu_1 + \mu_2) - Y}{\rho b^M} \tag{5.5}$$

and a similar analysis for the reverse transformation results in the following for the martensite volume fraction.

$$\xi\Big|_{\text{rev}} = \frac{H\sigma_x + \frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma_x^2 + \rho\Delta s_0\, T - \rho\Delta u_0 - (\mu_1 - \mu_2) + Y}{\rho b^A} \tag{5.6}$$

The typical pseudoelastic response of shape memory alloys, based on equations (5.5) and (5.6), is shown in Figure 10. $\sigma_{\text{s}}^{\text{fwd}}$ and $\sigma_{\text{f}}^{\text{fwd}}$ are the stresses to start and finish the forward transformation to martensite at the current temperature and $\sigma_{\text{s}}^{\text{rev}}$ and $\sigma_{\text{f}}^{\text{rev}}$ are the stresses to start and finish the reverse transformation to austenite, respectively.

The material constants used in the model can be found based on the following critical transformation start and finish conditions for a stress-free heating and cooling path.

1. At the start of forward transformation, $T = M_s$, $\xi = 0$.
2. At the end of forward transformation, $T = M_f$, $\xi = 1$.
3. At the start of reverse transformation, $T = A_s$, $\xi = 1$.
4. At the end of reverse transformation, $T = A_f$, $\xi = 0$.
5. The Gibbs free energy and thus the hardening function, $f(\xi)$, must be continuous; i.e. $f(\xi = 1)\big|_{\text{fwd}} = f(\xi = 1)\big|_{\text{rev}}$. The form of f is given in equation (4.72)

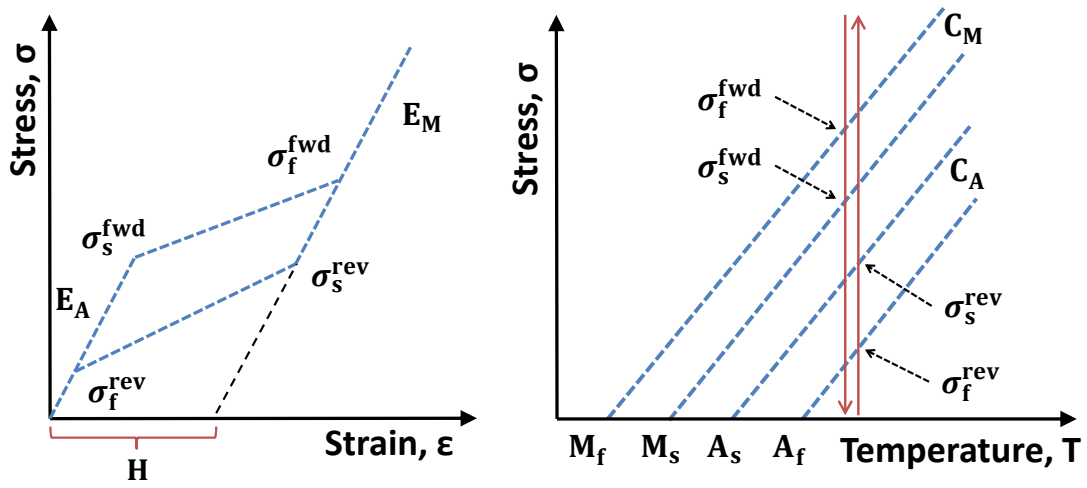


Figure 10: A typical stress-temperature phase diagram and pseudoelastic loading path for shape memory alloys.

for the linear hardening rule. Notice that this condition for $\xi = 0$ is trivially satisfied. For the smooth hardening function, $\frac{\partial f(\xi)}{\partial \xi}$ is explicitly given in (4.73). Hence the continuity condition can be written in the integral form as $\int_0^1 \frac{\partial f^{\text{fwd}}(\xi)}{\partial \xi}\, \mathrm{d}\xi = \int_0^1 \frac{\partial f^{\text{rev}}(\xi)}{\partial \xi}\, \mathrm{d}\xi$.

6. By taking the differential of (5.5) and (5.6) and letting $\mathrm{d}\xi$ to vanish, it is possible to recover the slopes of the transformaion regions on the stress-temperature phase diagram at a specific stress (for here the calibration stress $\sigma = \sigma^*$). This leads to a relation for $\rho \Delta s_0$ based on the forward and reverse transformation slopes, $C_M \equiv \frac{\mathrm{d}\sigma}{\mathrm{d}T}\big|_{\sigma=\sigma^*}^{\text{fwd}}$ and $C_A \equiv \frac{\mathrm{d}\sigma}{\mathrm{d}T}\big|_{\sigma=\sigma^*}^{\text{rev}}$.

The results of the conditions listed above are summarized in table (5.1) for the linear hardening and in table (5.2) for the smooth hardening rules.

The SMA model in this form cannot capture a response with different Clausius-

Clapeyron slopes, as per the last equation in table (5.1) or (5.2), hence $C_A = C_M$.

In addition, based on the stress-temperature phase diagram, ignoring the variation in the slopes of the forward and reverse transformation regions, such a relationship can be assumed for the case of linear hardening:

$$\begin{aligned} \sigma_{\mathrm{s}}^{\mathrm{fwd}} &= C_M \left(T - M_s\right) \quad , \quad \sigma_{\mathrm{f}}^{\mathrm{fwd}} = C_M \left(T - M_f\right) \\ \sigma_{\mathrm{s}}^{\mathrm{rev}} &= C_A \left(T - A_s\right) \quad , \quad \sigma_{\mathrm{f}}^{\mathrm{rev}} = C_A \left(T - A_f\right) \end{aligned} \tag{5.7}$$

Table 5.1: Material constants used in the local part of the SMA constitutive model using the linear hardening function.

$Y = \frac{1}{4}\Delta s_0 \left(M_s + M_f - A_s - A_f\right)$
$\rho b^M = -\rho \Delta s_0 \left(M_s - M_f\right)$
$\rho b^A = -\rho \Delta s_0 \left(A_f - A_s\right)$
$\mu_1 + \rho \Delta u_0 = \frac{1}{2}\rho \Delta s_0 \left(M_s + A_f\right)$
$\mu_2 = -\frac{1}{4}\rho \Delta s_0 \left(M_f - M_s + A_f - A_s\right)$
$\rho \Delta s_0 = -\left[H + \left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma^*\right] C_M \quad , \quad \rho \Delta s_0 = -\left[H + \left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma^*\right] C_A$

Table 5.2: Material constants used in the local part of the SMA constitutive model using the smooth hardening function.

$$Y = \tfrac{1}{2}\rho\Delta s_0\,(M_s - A_f) - b_3$$

$$b_1 = \rho\Delta s_0\,(M_f - M_s)$$

$$b_2 = \rho\Delta s_0\,(A_s - A_f)$$

$$b_3 = -\frac{b_1}{4}\left(1 + \frac{1}{n_1+1} - \frac{1}{n_2+1}\right) + \frac{b_2}{4}\left(1 + \frac{1}{n_3+1} - \frac{1}{n_4+1}\right)$$

$$\rho\Delta u_0 = \tfrac{1}{2}\rho\Delta s_0\,(A_f + M_s)$$

$$\rho\Delta s_0 = -\left[H + \left(\tfrac{1}{E^M} - \tfrac{1}{E^A}\right)\sigma^*\right] C_M \quad , \quad \rho\Delta s_0 = -\left[H + \left(\tfrac{1}{E^M} - \tfrac{1}{E^A}\right)\sigma^*\right] C_A$$

5.1.1 Experimental measurement of the SMA material properties: model III

The models developed herein contain several local and nonlocal material constants. The nonlocal model III introduces three independent length scales ℓ_1, ℓ_2^{fwd}, and ℓ_2^{rev} in addition to the SMA material constants already being used in the classical local Boyd-Lagoudas SMA model [88]. Based on the nonlocal constants, two parameters $\mathcal{M}^{\text{fwd}}$ and $\mathcal{M}^{\text{rev}}$ can be identified in the transformation differential equation. The material constants used in nonlocal modeling for shape memory alloys are categorized in table (5.3). The dissipative length scales, ℓ_2^{fwd} and ℓ_2^{rev}, are excluded from this table since they do not *independently* appear in the solution of the SMA structures studied in the next section. This is because the solutions of the transformation surfaces, (4.79) and (4.80), are considered for those one-dimensional problems rather than the rate forms in equation (4.71).

A_s and A_f denote the temperatures to start and finish reverse transformation from martensite to austenite at zero stress. Also, M_s and M_f are the temperatures

Table 5.3: SMA material properties used in the developed nonlocal model.

Thermoelastic Constants						Transformation Properties					Nonlocality	
ν^A	$[\cdots]$	E^A	[MPa]	α^A	$[\frac{1}{^\circ\mathrm{C}}]$	A_s , A_f	$[^\circ\mathrm{C}]$	C_A	$\left[\frac{\mathrm{MPa}}{^\circ\mathrm{C}}\right]$	H [%]	$\mathcal{M}^{\mathrm{rev}}$	$[\mathrm{MPa}\cdot\mathrm{mm}]$
ν_M		E^M		α^M		M_s , M_f		C_M			$\mathcal{M}^{\mathrm{fwd}}$	

at which the forward transformation from austenite to martensite starts and finishes at zero stress. Martensitic transformation is driven by stress, in addition to temperature, and hence the transformation temperatures change in the presence of stress. In this regard, a stress-temperature phase diagram is commonly associated to every shape memory alloy on which C_A and C_M are the slopes of reverse and forward transformation bands, respectively (see Figure 10, $C_A = C_M$ for the current model).

The relation between the material parameters expressed in the course of model development and the more common SMA properties listed in table (5.3) was established in this section using the problem of isothermal uniaxial stretching of an SMA bar (table (5.1)). Although trivial at the first glance, the solution to this problem verifies the fact that the nonlocal models developed herein reduce to the original classical SMA model for the cases where no structural or loading attributions exist that can give rise to the spatial gradients of martensite volume fraction or transformation strain.

Calibration of the nonlocal parameters, $\mathcal{M}^{\mathrm{fwd}}$ and $\mathcal{M}^{\mathrm{rev}}$, have to be performed utilizing specially designed experiments. To this end, bending of SMA beams with various thicknesses, torsion of SMA thin wires with various diameters, or compression of tapered SMA pillars with different diameters are promising, as they activate nonlocality through non-homogeneous loading or variation in the structure. To the best knowledge of the authors, there are no experimental results published for bend-

ing of SMA thin films with different thicknesses, although this has been done for the case of dislocation plasticity using aluminum foils [137]. Experimental observations for compression of SMA micropillars, on the other hand, are available and will be used for calibration of the current SMA nonlocal model.

5.2 Compression of an SMA micropillar

In this section, the compression of SMA pillars with micron size is investigated. The experimental response of such a structure shows a size effect which can include an increase in the critical stresses for transformation and also in energy damping through increased stress hysteresis. The compressive response of an SMA micropillar during loading is analytically derived using the developed nonlocal models II and III. To that end, the micropillar is modeled as a uniaxial bar with a taper from the top surface to the bottom (Figure 11). Again, it is assumed that the pillar is loaded at a constant temperature above the austenite finish temperature, A_f, of the SMA.

The bottom end of the micropillar, as connected to the substrate, is assumed to remain untransformed due to the pinning effect at the interface. The kinematic and equilibrium equations can be written as:

$$\varepsilon_x = \frac{\partial u(x)}{\partial x} \quad , \quad \frac{\partial \sigma_x(x)}{\partial x} + \frac{\partial \sigma_{xy}(x)}{\partial y} + \frac{\partial \sigma_{xz}(x)}{\partial z} = 0 \tag{5.8}$$

Considering the taper, $\eta = \frac{D-d}{h}$, and the variation in the area with the taper, a uniaxial state of the stress can be approximated based on the compressive force f:

$$A^{\text{pillar}} = \frac{\pi}{4}\left(D - \frac{D-d}{h}x\right)^2 \quad , \quad \sigma_x(x) = -\frac{4f}{\pi\left(D-\eta x\right)^2} \quad , \quad \sigma_{xy} = \sigma_{xz} \cong 0 \tag{5.9}$$

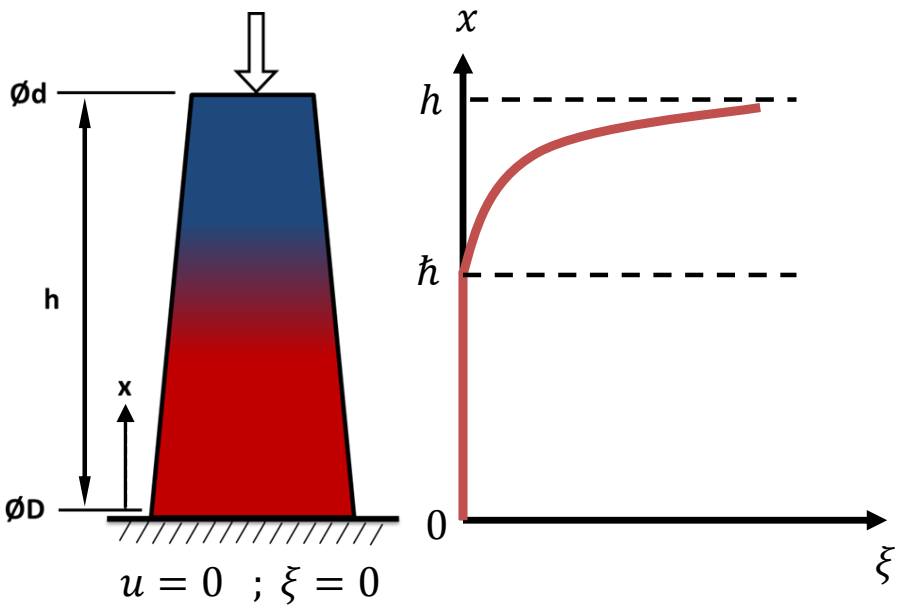


Figure 11: Schematic, including boundary conditions, for compression of a superelastic SMA micropillar with top and bottom diameters of d and D and height of h.

Notice that such a stress state is an approximation and does not exactly satisfy the equilibrium equation.

5.2.1 Solution with model II

The SMA gradient-based model II established in section 4.4 is now used to study the behavior of the SMA micropillar under compression. The decomposition of strain for $\varepsilon_x(x)$ can be obtained by rewriting equation (4.53)a for this 1D case

$$\begin{aligned}\varepsilon_x &= \frac{\sigma_x}{E(\xi)} + \alpha\,(T - T_0) + \varepsilon_x^{tr} - \frac{1}{2}\left(a_2\ell_2^2 + a_6\ell_6^2\right)\left(\frac{\partial \xi}{\partial x}\right)^2 \\ &= \frac{\sigma_x}{E(\xi(x))} + \alpha\,(T - T_0) + \varepsilon_x^{tr}(x) - \frac{3}{2}a_6\ell_6^2\left(\frac{\partial \xi}{\partial x}\right)^2\end{aligned} \tag{5.10}$$

Although normal strain components in the y and z directions exist as a result of this loading, only the relation for $\varepsilon_x(x)$ is explicitly written to serve the purpose of finding the force-contraction response of the pillar. The total strain in equation (5.10) is decomposed to elastic, thermal and transformation parts. The last term, however, is related to the effect the gradients have on the strain in the material.

The rest of the constitutive variables, used for derivation of the model, are also obtained from equation (4.53).

$$\varsigma_x^D = \sigma_x - \frac{1}{2}a_5\ell_5^2\left(\frac{\partial \xi}{\partial x}\right)^2 \quad , \quad \boldsymbol{\varsigma}^D : \boldsymbol{\varsigma}^D = \frac{2}{3}\left[\sigma_x - \frac{1}{2}a_5\ell_5^2\left(\frac{\partial \xi}{\partial x}\right)^2\right]^2 \tag{5.11}$$

The components of $\boldsymbol{\varsigma}^D$ other than ς_x^D are zero.

$$\pi = \frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma_x^2 + \rho\Delta s_0\,T - \rho\Delta u_0 - \frac{\partial f^{fwd}}{\partial \xi} - \frac{1}{2}a_4\ell_4^2\left(\frac{\partial \xi}{\partial x}\right)^2 \tag{5.12}$$

In order to find a relation for the components of the vector $\bar{\boldsymbol{\pi}}$, one must know

the components of the transformation strain, $\boldsymbol{\varepsilon}^{tr}$. According to equation (4.55)a, the rate of transformation strain can be obtained from

$$\dot{\varepsilon_x^{tr}} = H\,\dot{\xi}\,\mathrm{sgn}(\bar{\sigma}_x) \quad , \quad \dot{\varepsilon_y^{tr}} = \dot{\varepsilon_z^{tr}} = -\frac{1}{2}\dot{\varepsilon_x^{tr}} \tag{5.13}$$

where $\bar{\sigma}_x = \sigma_x - \frac{1}{2}a_5\ell_5^2\left(\frac{\partial\xi}{\partial x}\right)^2$. This can be readily integrated, by assuming $\bar{\sigma}_x(x) \leq 0$ a priori to the solution, such that

$$\varepsilon_x^{tr} = -H\,\xi \quad , \quad \varepsilon_y^{tr} = \varepsilon_z^{tr} = -\frac{1}{2}\varepsilon_x^{tr} \tag{5.14}$$

Therefore, we will have

$$\begin{aligned}
\bar{\pi}_x &= -\left[a_1\ell_1^2 + a_2\ell_2^2\,\sigma_x + a_3\ell_3^2\,T + a_4\ell_4^2\,\xi + a_5\ell_5^2\,\varepsilon_x^{tr} + a_6\ell_6^2\,\mathrm{tr}\,(\boldsymbol{\sigma})\right]\frac{\partial\xi}{\partial x} \\
&= -\left[a_1\ell_1^2 + a_3\ell_3^2\,T + a_4\ell_4^2\,\xi + a_5\ell_5^2\,\varepsilon_x^{tr} + 3\,a_6\ell_6^2\,\sigma_x\right]\frac{\partial\xi}{\partial x} \\
\bar{\pi}_y &= \bar{\pi}_z = 0
\end{aligned} \tag{5.15}$$

So the transformation surfaces finally take the form

$$\begin{aligned}
\widetilde{\varphi}(\boldsymbol{\varsigma}^{'D}) &= H\left(\frac{3}{2}\varsigma_{ij}^{'}\varsigma_{ij}^{'}\right)^{\frac{1}{2}} = H\left|\sigma_x - \frac{1}{2}a_5\ell_5^2\left(\frac{\partial\xi}{\partial x}\right)^2\right| \\
\bar{\varphi}(\bar{\boldsymbol{\pi}}) &= -\left(\frac{1}{\ell_d^2}\bar{\pi}_i\bar{\pi}_i\right)^{\frac{1}{2}} = -\frac{1}{\ell_d}\left|a_1\ell_1^2 + a_3\ell_3^2\,T + a_4\ell_4^2\,\xi - a_5\ell_5^2\,H\xi + 3\,a_6\ell_6^2\,\sigma_x\right|\left|\frac{\partial\xi}{\partial x}\right|
\end{aligned} \tag{5.16}$$

$$\begin{aligned}
\Phi^{\mathrm{fwd}} &= \widetilde{\varphi}(\boldsymbol{\varsigma}^{'D}) + \pi + \bar{\varphi}(\bar{\boldsymbol{\pi}}) - Y = 0 \\
\Phi^{\mathrm{rev}} &= -\widetilde{\varphi}(\boldsymbol{\varsigma}^{'D}) - \pi - \bar{\varphi}(\bar{\boldsymbol{\pi}}) - Y = 0
\end{aligned} \tag{5.17}$$

For the current problem , the stress σ_x is known from equation (5.9)b and hence $\bar{\sigma}_x \leq 0$. Also, $\frac{\partial \xi}{\partial x} \geq 0$. Therefore to find the distribution of martensite volume fraction under a certain stress, the differential equations for the transformation surface must be solved considering the boundary condition shown in Figure 11.

For the forward transformation

$$
\begin{aligned}
\Phi^{\text{fwd}} = & \\
& - H \left[\sigma_x - \frac{1}{2} a_5 \ell_5^2 \left(\frac{\partial \xi}{\partial x} \right)^2 \right] + \frac{1}{2} \left(\frac{1}{E^M} - \frac{1}{E^A} \right) \sigma_x^2 + \rho \Delta s_0 \, T - \rho \Delta u_0 \\
& - \frac{\partial f^{fwd}}{\partial \xi} - \frac{1}{2} a_4 \ell_4^2 \left(\frac{\partial \xi}{\partial x} \right)^2 \\
& - \frac{1}{\ell_d} \left| a_1 \ell_1^2 + a_3 \ell_3^2 \, T + a_4 \ell_4^2 \, \xi - a_5 \ell_5^2 \, H \xi + 3 \, a_6 \ell_6^2 \, \sigma_x \right| \frac{\partial \xi}{\partial x} - Y = 0
\end{aligned}
\tag{5.18}
$$

The boundary conditions required to solve this differential equation can be identified as

$$
\xi(\hbar) = 0 \quad , \quad \left. \frac{\mathrm{d}\xi}{\mathrm{d}x} \right|_{\hbar} = 0 \tag{5.19}
$$

which ensures the continuity of $\xi(x)$ and $\frac{\mathrm{d}\xi(x)}{\mathrm{d}x}$ across the interface, $x = \hbar$, between the top region where forward transformation is occurring and the bottom region where it is still loading elastically (Figure 11). They also enable a solution to determine $\hbar$.

For the reverse transformation

$$
\begin{aligned}
\Phi^{\text{rev}} = & \\
& H\left[\sigma_x - \frac{1}{2}a_5\ell_5^2\left(\frac{\partial\xi}{\partial x}\right)^2\right] - \frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma_x^2 - \rho\Delta s_0\, T + \rho\Delta u_0 \\
& + \frac{\partial f^{rev}}{\partial\xi} + \frac{1}{2}a_4\ell_4^2\left(\frac{\partial\xi}{\partial x}\right)^2 \\
& + \frac{1}{\ell_d}\left|a_1\ell_1^2 + a_3\ell_3^2\, T + a_4\ell_4^2\,\xi - a_5\ell_5^2\, H\xi + 3\, a_6\ell_6^2\,\sigma_x\right|\frac{\partial\xi}{\partial x} - Y = 0
\end{aligned}
\tag{5.20}
$$

Similarly, the boundary condition for the above is the continuity of $\xi(x)$ and $\frac{\mathrm{d}\xi(x)}{\mathrm{d}x}$ across the interface, $x = \acute{h}$, where the transition form the reverse transformation to elastic unloading takes place. The two boundary conditions can be used for solving the above differential equation and also determining $x = \acute{h}$.

ξ obtained form the above equations can be replaced in equation (5.14) to find the transformation strain along the pillar axis. Therefore, the total strain is given by

$$
\varepsilon_x = \frac{\sigma_x}{E(\xi(x))} + \alpha\,(T - T_0) - H\xi(x) - \frac{3}{2}a_6\ell_6^2\left(\frac{\partial\xi}{\partial x}\right)^2 \tag{5.21}
$$

Ultimately, the deformation at the top of the micropillar can be determined form the relation below:

$$
u(x) = \int_0^x \varepsilon_x(\tau)\,\mathrm{d}\tau + u(0) \quad , \quad u(0) = 0 \tag{5.22}
$$

5.2.2 Solution with model III

The decomposition of strain using model III, as in equation (4.46)a, is more straightforward.

$$\varepsilon_x = \frac{\sigma_x}{E(\xi(x))} - H\xi(x) \tag{5.23}$$

Noting that $\sigma_x \leq 0$, the forward transformation differential equation becomes:

$$\begin{aligned} -H\sigma_x(x) + \left[\frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma_x(x)^2 + \rho\Delta s_0\, T - \rho\Delta u_0 - \frac{\partial f^{fwd}}{\partial \xi}\right] \\ - \mathcal{M}^{\text{fwd}}\left|\frac{\mathrm{d}\xi(x)}{\mathrm{d}x}\right| - Y = 0 \end{aligned} \tag{5.24}$$

In order to establish a closed-form integration for the above, the linear hardening function

$$\frac{\partial f^{fwd}}{\partial \xi} = \rho b^M \xi(x) + \mu_1 + \mu_2 \tag{5.25}$$

as well as $E = E^A = E^M$ must be assumed.

The stress is maximum at the top and therefore the transformation starts and propagates from the top to the bottom of the pillar, $\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} \geq 0$, with the interface between the transforming-elastic regions at $x = \hbar$. Hence, the transformation differential equation can be rewritten in terms of the volume fraction of martensite as:

$$B\xi(x) + \mathcal{M}^{\text{fwd}}\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} = -\left[N + \frac{4fH}{\pi\left(D - \eta x\right)^2}\right] \quad , \quad \xi(\hbar) = 0 \quad , \quad \left.\frac{\mathrm{d}\xi}{\mathrm{d}x}\right|_{\hbar} = 0 \tag{5.26}$$

Here it is assumed that ξ and $\nabla\xi$ are continuous at the forward transformation front. This means that the martensite volume fraction and its derivative along x are zero at the elastic-transforming interface. Also

$$\begin{aligned} B =& -\rho b^M = -H\left(\sigma_{\mathrm{f}}^{\mathrm{fwd}} - \sigma_{\mathrm{s}}^{\mathrm{fwd}}\right) \\ N =& \rho\Delta s_0\, T - \rho\Delta u_0 - (\mu_1+\mu_2) - Y = -H\sigma_{\mathrm{s}}^{\mathrm{fwd}} \end{aligned} \tag{5.27}$$

The general solution to the differential equation in (5.26) is:

$$\begin{aligned} \xi(x) &= -\frac{N}{B} + \zeta\, \mathrm{e}^{-\frac{B}{\mathcal{M}^{\mathrm{fwd}}}x} - \frac{4fH}{\pi\mathcal{M}^{\mathrm{fwd}}}\left[\frac{1}{\eta\,(D-\eta x)} + \frac{B}{\mathcal{M}^{\mathrm{fwd}}}\frac{1}{\eta^2}\,Ei(-z)\,\mathrm{e}^{z}\right] \\ z &= \frac{B}{\eta\mathcal{M}^{\mathrm{fwd}}}\,(D-\eta x) \end{aligned} \tag{5.28}$$

where $Ei(z)$ is the exponential integral function with the following definition and differentiation properties.

$$\begin{aligned} Ei(z) = -\int_{-z}^{\infty}\frac{\mathrm{e}^{-\tau}}{\tau}\,\mathrm{d}\tau \quad \text{or} \quad Ei(z) = \int_{-\infty}^{z}\frac{\mathrm{e}^{\tau}}{\tau}\,\mathrm{d}\tau \\ g(t) = -\mathrm{e}^{t}\,Ei(-t) \quad \Rightarrow \quad \frac{\mathrm{d}g}{\mathrm{d}t} = g(t) - \frac{1}{t} \end{aligned} \tag{5.29}$$

Using the prescribed boundary condition at the interface between the top transforming and the bottom elastic austenitic regions, the integration constant ζ and the location of the front $\hbar$ can be determined to complete the solution.

$$\begin{aligned}
\zeta &= \left\{ \frac{N}{B} + \frac{4fH}{\pi \mathcal{M}^{\mathrm{fwd}}} \left[\frac{1}{\eta \left(D - \eta \hbar \right)} + \frac{B}{\mathcal{M}^{\mathrm{fwd}}} \frac{1}{\eta^2} \, Ei(-\bar{z}) \, \mathrm{e}^{\bar{z}} \right] \right\} \mathrm{e}^{\frac{B}{\mathcal{M}^{\mathrm{fwd}}} \hbar} \\
\bar{z} &= \frac{B}{\eta \mathcal{M}^{\mathrm{fwd}}} \left(D - \eta \hbar \right) \quad , \quad \hbar = \frac{1}{\eta} \left(D - \sqrt{\frac{4f}{\pi \sigma_{\mathrm{s}}^{\mathrm{fwd}}}} \right)
\end{aligned} \tag{5.30}$$

Having determined the distribution of martensite volume fraction, the total strain can be obtained using equation (5.23) and the displacement via:

$$u(x) = \int_0^x \varepsilon_x(\tau) \, \mathrm{d}\tau + u(0) = - \int_0^x \left[\frac{4f}{\pi E \left(D - \eta \tau \right)^2} + H \xi(\tau) \right] \, \mathrm{d}\tau \tag{5.31}$$

Additionally, the following nondimensional number is introduced for the problem of compression of the micropillar which will later be used for presentation of the results:

$$\varpi = \frac{\mathcal{M}^{\mathrm{fwd}}}{BD} \eta \tag{5.32}$$

For the purpose of comparison, it is also helpful to present the local solution to the problem of uniaxial SMA micropillar under a compressive force as defined in the aforementioned boundary value problem. The conventional local SMA constitutive model is used to that end [88].

$$\xi^{\text{local}}(x) = -\frac{1}{B}\left[N + \frac{4fH}{\pi\left(D-\eta x\right)^2}\right] \quad , \quad \varepsilon_x^{\text{local}}(x) = -\frac{4fH}{E\pi\left(D-\eta x\right)^2} - H\xi^{\text{local}}(x) \tag{5.33}$$

It can be shown for the micropillar compression that the reverse transformation begins at the top surface and penetrates towards the base of the pillar upon unloading. The interface between the region of reverse transformation and the region where unloading takes place elastically is denoted by $\acute{h}$ and the transformation finishes when $\acute{h}$ reaches $\hbar$. A closed-from analytical solution is, as well, possible to derive for the unloading step using the reverse transformation differential equation

$$\begin{aligned} H\sigma_x(x) - \left[\frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma_x(x)^2 + \rho\Delta s_0\, T - \rho\Delta u_0 - \frac{\partial f^{rev}}{\partial \xi}\right] \\ + \mathcal{M}^{\text{rev}}\frac{\mathrm{d}\xi(x)}{\mathrm{d}x} - Y = 0 \end{aligned} \tag{5.34}$$

and by taking into account the required continuity boundary conditions across the reverse transformation front, i.e. the continuity of ξ and $\frac{\mathrm{d}\xi(x)}{\mathrm{d}x}$ at $x = \acute{h}$. The loading-unloading solution is, nonetheless, obtained numerically and presented in chapter 6.

5.3 Simple torsion of an SMA bar: model II

The solution for the torsion of an SMA bar using the nonlocal model II is obtained in this section. It is assumed that the bar has a circular cross section representing a wire with a diameter of D. A opposite and equal torque T is applied to the ends of the wire (Figure 12). Also, it is assumed that the wire is initially in a fully austenitic state at a temperature above A_f.

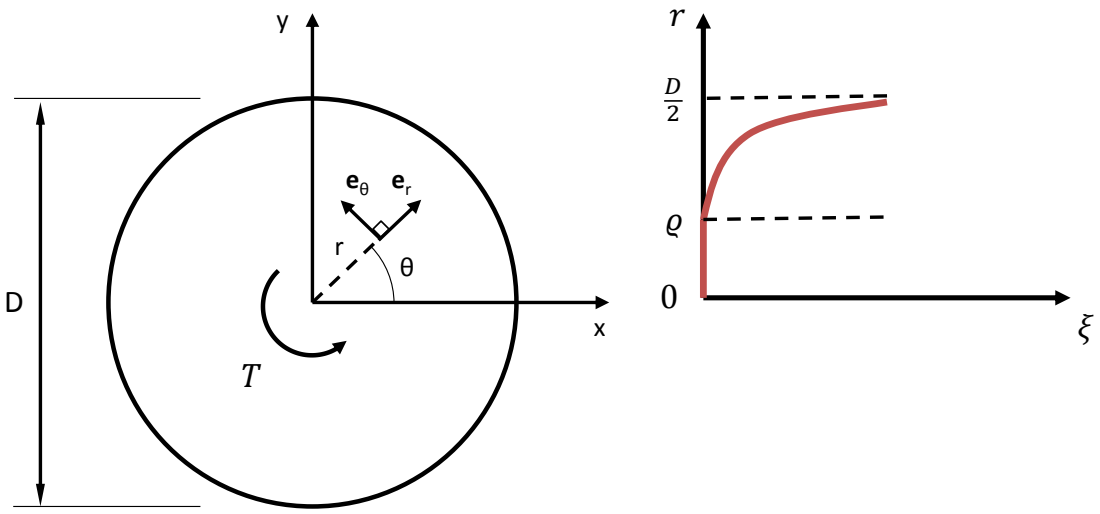


Figure 12: Simple torsion of an SMA bar with a circular cross section showing, to the right side, a typical distribution for martensite volume fraction during forward transformation.

The twist per unit length of the circular bar is denoted by Θ, a constant. Due to the symmetry existing in the problem, the polar coordinate system will be used for the analysis where none of the variables are a function of θ. For this problem, it is

assumed that the radii on the cross section remain straight. The shear strain in the plane perpendicular to the axis is, hence, given by

$$u_\theta = r\Theta z \quad , \quad \varepsilon_{\theta z} = \frac{1}{2}\left(\frac{\partial u_\theta}{\partial z} + \frac{1}{r}\frac{\partial u_z}{\partial \theta}\right) = \frac{1}{2} r\Theta \tag{5.35}$$

The state of stress in the SMA bar, satisfying the boundary conditions for end moments and traction-free side surfaces, is considered to be

$$\sigma_{ij} = 0 \quad \setminus \quad \sigma_{\theta z} \tag{5.36}$$

Therefore, the equilibrium equation reduces to

$$\frac{\partial \sigma_{\theta z}}{\partial z} = 0 \quad , \quad \frac{1}{r}\frac{\partial \sigma_{\theta z}}{\partial \theta} = 0 \tag{5.37}$$

which are all satisfied in the case of $\sigma_{\theta z} = \hat{\sigma}_{\theta z}(r)$. This implies that $\xi = \hat{\xi}(r)$. Also, for the regions in the material undergoing elastic loading $\sigma_{\theta z} = Gr\Theta$ where $G = \frac{E}{2(1+\nu)}$.

The relation for total strain, according to model II, is given in equation (4.53)a. For the case of this problem, after considering the fact that martensite volume fraction changes only with radial location, $\xi = \hat{\xi}(r)$, the θz component of strain tensor can be written as

$$\varepsilon_{\theta z} = \frac{\sigma_{\theta z}}{2G(\xi)} + \varepsilon^{tr}_{\theta z} - \frac{1}{2} a_6 \ell_6^2 \left(\frac{\partial \xi}{\partial r}\right)^2 \tag{5.38}$$

The generalized stress tensor, $\boldsymbol{\varsigma}^D$, conjugate to transformation strain can be determined by simplifying equation (4.53)c which finally leads to

$$\boldsymbol{\varsigma}^{'D} : \boldsymbol{\varsigma}^{'D} = 2\sigma_{\theta z}^2 + \frac{2}{3}\left[\frac{1}{2}a_5\ell_5^2\left(\frac{\partial\xi}{\partial r}\right)^2\right]^2 \tag{5.39}$$

In this sense then

$$\widetilde{\varphi}(\boldsymbol{\varsigma}^{'D}) = H\sqrt{3\sigma_{\theta z}^2 + \left[\frac{1}{2}a_5\ell_5^2\left(\frac{\partial\xi}{\partial r}\right)^2\right]^2} \tag{5.40}$$

which means, after equation (4.55)a, the rate of transformation strain can be given by

$$\dot{\boldsymbol{\varepsilon}}^{tr} = \frac{3}{2}H\frac{\dot{\xi}}{\sqrt{3\sigma_{\theta z}^2 + \left[\frac{1}{2}a_5\ell_5^2\left(\frac{\partial\xi}{\partial r}\right)^2\right]^2}}\,\boldsymbol{\varsigma}^{'D} \tag{5.41}$$

or in terms of components

$$\begin{aligned}
\dot{\varepsilon}_{rr}^{tr} &= -2\dot{\varepsilon}_{\theta\theta}^{tr} = -2\dot{\varepsilon}_{zz}^{tr} = H\dot{\xi}\frac{-\frac{1}{2}a_5\ell_5^2}{\sqrt{3\sigma_{\theta z}^2 + \left[\frac{1}{2}a_5\ell_5^2\left(\frac{\partial\xi}{\partial r}\right)^2\right]^2}}\left(\frac{\partial\xi}{\partial r}\right)^2 \\
\dot{\varepsilon}_{\theta z}^{tr} &= \frac{3}{2}H\frac{\dot{\xi}}{\sqrt{3\sigma_{\theta z}^2 + \left[\frac{1}{2}a_5\ell_5^2\left(\frac{\partial\xi}{\partial r}\right)^2\right]^2}}\,\sigma_{\theta z} \quad , \quad \dot{\varepsilon}_{r\theta}^{tr} = \dot{\varepsilon}_{rz}^{tr} = 0
\end{aligned} \tag{5.42}$$

As observed, in this simple torsion problem in addition to transformation shear strains, model II results in transformation strain in normal directions as well. Therefore by the start and propagation of transformation in the bar, the cross section may undergo swelling or warping. This point can be used for identification and calibration of the nonlocal parameter $a_5\ell_5^2$.

Furthermore, (4.53)d reduces to

$$\pi = \frac{1}{2}\left(\frac{1}{G^M} - \frac{1}{G^A}\right)\sigma_{\theta z}^2 + \rho\Delta s_0\, T - \rho\Delta u_0 - \frac{\partial f^{fwd/rev}}{\partial \xi} - \frac{1}{2}a_4\ell_4^2\left(\frac{\partial \xi}{\partial r}\right)^2 \tag{5.43}$$

The generalized force $\bar{\boldsymbol{\pi}}$ in component form can be given by

$$\begin{aligned}
\bar{\pi}_r &= -\left[a_1\ell_1^2 + a_3\ell_3^2\, T + a_4\ell_4^2\,\xi + a_5\ell_5^2\,\varepsilon_{rr}^{tr}\right]\frac{\partial \xi}{\partial r} \\
\bar{\pi}_\theta &= -a_5\ell_5^2\,\frac{\partial \xi}{\partial r}\,\varepsilon_{r\theta}^{tr} \\
\bar{\pi}_z &= -a_5\ell_5^2\,\frac{\partial \xi}{\partial r}\,\varepsilon_{rz}^{tr}
\end{aligned} \tag{5.44}$$

The last two components vanish in light of the relations given for the rate of transformation strain components, equation (5.42)c. Hence, we will have

$$\bar{\varphi}(\bar{\boldsymbol{\pi}}) = -\left(\frac{1}{\ell_d^2}\bar{\pi}_i\bar{\pi}_i\right)^{\frac{1}{2}} = -\frac{1}{\ell_d}\left|a_1\ell_1^2 + a_3\ell_3^2\, T + a_4\ell_4^2\,\xi + a_5\ell_5^2\,\varepsilon_{rr}^{tr}\right|\left|\frac{\partial \xi}{\partial r}\right| \tag{5.45}$$

Without loss of generality, it is reasonable to assume that $\sigma_{\theta z} \geq 0$ and $\frac{\partial \xi}{\partial r} \geq 0$. Therefore, the forward transformation surface will become

$$\begin{aligned}
\Phi^{\text{fwd}} = \\
&H\sqrt{3\sigma_{\theta z}^2 + \left[\frac{1}{2}a_5\ell_5^2\left(\frac{\partial \xi}{\partial r}\right)^2\right]^2} + \frac{1}{2}\left(\frac{1}{G^M} - \frac{1}{G^A}\right)\sigma_{\theta z}^2 + \rho\Delta s_0\, T - \rho\Delta u_0 - \frac{\partial f^{fwd}}{\partial \xi} \\
&- \frac{1}{2}a_4\ell_4^2\left(\frac{\partial \xi}{\partial r}\right)^2 - \frac{1}{\ell_d}\left|a_1\ell_1^2 + a_3\ell_3^2\, T + a_4\ell_4^2\,\xi + a_5\ell_5^2\,\varepsilon_{rr}^{tr}\right|\frac{\partial \xi}{\partial r} - Y = 0
\end{aligned} \tag{5.46}$$

The boundary conditions required to solve this differential equation is determined by considering the interface between the elastic and transforming regions in the circular bar. The forward transformation begins at the outer surface while the inner core is still loading elastically. The schematic for the distribution of martensite volume fraction in this case is shown in Figure 12 where ϱ denotes the radial position of the interface. Hence

$$\xi(\varrho) = 0 \quad , \quad \frac{\mathrm{d}\xi}{\mathrm{d}r}\Big|_{\varrho} = 0 \tag{5.47}$$

due to the zeroth and first-order continuity of $\xi(r)$. (5.46) and (5.47) enable one to find ϱ.

For the reverse transformation

$$\begin{aligned}\Phi^{\mathrm{rev}} = & \\ & - H\sqrt{3\sigma_{\theta z}^2 + \left[\frac{1}{2}a_5\ell_5^2\left(\frac{\partial\xi}{\partial r}\right)^2\right]^2} - \frac{1}{2}\left(\frac{1}{G^M} - \frac{1}{G^A}\right)\sigma_{\theta z}^2 - \rho\Delta s_0\, T + \rho\Delta u_0 + \frac{\partial f^{rev}}{\partial\xi} \\ & + \frac{1}{2}a_4\ell_4^2\left(\frac{\partial\xi}{\partial r}\right)^2 + \frac{1}{\ell_d}\left|a_1\ell_1^2 + a_3\ell_3^2\, T + a_4\ell_4^2\,\xi + a_5\ell_5^2\,\varepsilon_{rr}^{tr}\,\right|\frac{\partial\xi}{\partial r} - Y = 0\end{aligned} \tag{5.48}$$

Similarly, the boundary conditions for the above is the continuity of $\xi(x)$ and $\frac{\mathrm{d}\xi(r)}{\mathrm{d}r}$ across the interface, $x = \acute{\varrho}$, where the transition form the reverse transformation to elastic unloading takes place.

For this problem the differential equations for the transformation surfaces, (5.46) or (5.48) , have to be solved in a coupled fashion with the equations for the rate of transformation strain, (5.42), and the equation for the decomposition of strain,

(5.38), in order to find the distribution of $\xi(r)$ and $\sigma_{\theta z}(r)$.

Ultimately, it is possible to find the relation between the twist per length in the bar Θ, as the loading parameter, and the applied torque.

$$T(\Theta) = 2\pi \int_0^{\frac{D}{2}} \sigma_{\theta z}(r)\, r \, \mathrm{d}r \tag{5.49}$$

The complexity of the problem here prohibits one from obtaining any closed-form solution. In the next section, pure bending of an SMA beam is considered and its solution for model III is developed.

5.4 Pure bending of an SMA beam: model III

The deflection of SMA thin film actuators is investigated in this section. It is considered that a slender prismatic SMA beam (with a width of W and height of $2h$) undergoes a state of pure bending (Figure 13) such that the Euler-Bernoulli assumptions hold, i.e. the planar sections remain planar and perpendicular to the neutral axis and rotate with reasonably small slopes, hence:

$$u(y) = -y\tan(\theta) = -y\frac{\partial v(x)}{\partial x} \quad , \quad \varepsilon_x = -y\frac{\partial^2 v}{\partial x^2} = \kappa y \tag{5.50}$$

Where κ is the curvature of bending deformation. We assume that the only nonzero component of stress is the normal stress on the cross section, σ_x.

$$\frac{\partial \sigma_x(y)}{\partial x} = 0 \tag{5.51}$$

Moreover, the decomposition of strain in the axial x direction using the SMA

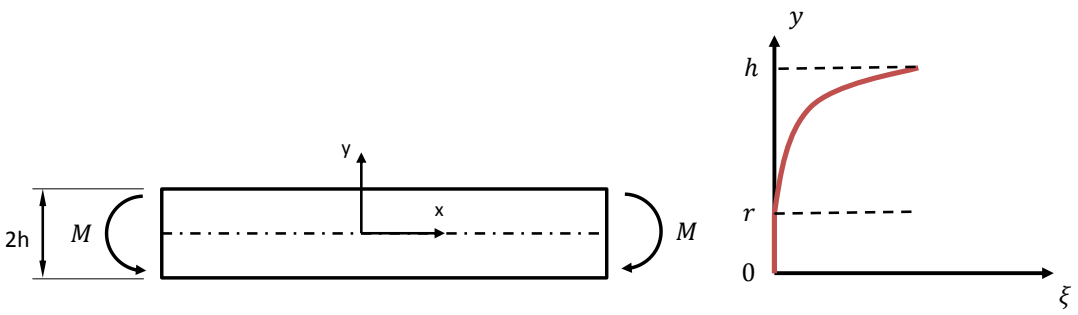


Figure 13: Pure bending of an SMA beam showing, to the right side, a typical distribution for martensite volume fraction through the (half) thickness during forward transformation.

gradient-based model II results in

$$\varepsilon_x = \frac{\sigma_x}{E(\xi(x))} + H\xi(x) \tag{5.52}$$

Firstly, the loading of the beam, that causes the start of forward transformation on the outermost layers, is investigated with the purpose of finding an analytical non-dimensional response. The state of the beam including stress, strain and martensitic volume fraction reduce to a one-dimensional problem with variation along the thickness of the beam or y direction; $\sigma_x(y)$, $\varepsilon_x(y)$, $\xi(y)$. Normal stress in the x direction is the only nonzero component of the stress tensor, therefore $J_2 = \frac{1}{3}\sigma_x^2$. The deformation is isothermal at a temperature above A_f thus the loading begins while the beam is entirely in an austenite state. The loading causes the transformation to martensite to begin at the outer layer and propagate towards the neutral axis. The interface between transforming outer layers and the elastic austenitic inner layers of the beam is denoted by r on the y axis, $\xi(y = r) = 0$. Due to the 1D nature of the

stress state, the forward transformation equation is going to be as given in (4.79) for the SMA gradient-based model III.

$$H\left|\sigma_x(y)\right| + \left[\frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right)\sigma_x(y)^2 + \rho\Delta s_0^{\text{fwd}}\,T - \rho\Delta u_0 - \left(\rho b^M \xi(y) + \mu_1 + \mu_2\right)\right] - \mathcal{M}^{\text{fwd}}\left|\frac{\mathrm{d}\xi(y)}{\mathrm{d}y}\right| - Y = 0 \tag{5.53}$$

in which a linear hardening function is used. Considering the symmetry about the neural axis, only the top portion of the beam is studied hence $\sigma_x \geq 0$ and, therefore, $\frac{\mathrm{d}\xi}{\mathrm{d}y} \geq 0$. The material constants in the above equation can be collected together to reach:

$$H\sigma_x(y) + A\sigma_x(y)^2 + B\xi(y) - \mathcal{M}^{\text{fwd}}\frac{\mathrm{d}\xi(y)}{\mathrm{d}y} + D = 0 \tag{5.54}$$

with the new introduced constants, of forward transformation with linear hardening, as of:

$$\begin{aligned} A &= \frac{1}{2}\left(\frac{1}{E^M} - \frac{1}{E^A}\right) \quad , \quad B = -\rho b^M = -H\left(\sigma_{\text{f}}^{\text{fwd}} - \sigma_{\text{s}}^{\text{fwd}}\right) \\ D &= \rho\Delta s_0^{\text{fwd}}\,T - \rho\Delta u_0 - (\mu_1 + \mu_2) - Y = -H\sigma_{\text{s}}^{\text{fwd}} \end{aligned} \tag{5.55}$$

Equation (5.54) is a differential equation that requires a boundary condition for its solution. Obviously, the condition at the boundary of the elastic-transforming regions or, $\xi(r) = 0$, must be satisfied. But this relation gives the constant of integration for (5.54) as a function of r which is still an unknown parameter of

loading. Implicit in the analysis of the nonlocal constitutive model developed here is the fact that the forward or reverse transformation front entails a smooth transition form the elastic to the transforming region; i.e. all the state variables, including $\nabla\xi(\boldsymbol{X})$, must be continuous across the interface. In a poly-crystalline SMA sample, the transformation front progresses by the martensitic transformation first occurring in the grains with favorable crystalline direction to the stress direction and then gradually transitioning to the elastic behavior in the transformed regions. This can be simulated by having a smooth scalar field variable, ξ. Thus, the condition at the elastic-transforming boundary cannot be short of:

$$\xi\Big|_{y=r} = 0 \quad , \quad \frac{\mathrm{d}\xi}{\mathrm{d}y}\Big|_{y=r} = 0 \tag{5.56}$$

Satisfaction of the boundary conditions in (5.56) gives a relation for both the constant of integration in the ODE of (5.54) as well as the location of the elastic-transforming boundary r with respect to the current curvature, κ, as the loading parameter. It is worth mentioning that the transition to the fully martensitic region is not going to be smooth in terms of ξ, i.e. $\nabla\xi$ will not be continuous across the finish transformation front.

Interestingly, using (5.56) in the equation for forward transformation surface (5.54), we will have

$$\left[H\sigma_x(y) + D\right]_{y=r} = 0 \quad \Rightarrow \quad \sigma_x(r) = \sigma_{\mathrm{s}}^{\mathrm{fwd}} \quad \Rightarrow \quad r = \frac{\sigma_{\mathrm{s}}^{\mathrm{fwd}}}{E^A\kappa} \tag{5.57}$$

The loading here is proportional. So in light of equation (4.69)a, it is possible to directly integrate the rate of evolution of transformation strains

$$\dot{\varepsilon_x^{tr}} = H\,\dot{\xi}\,\mathrm{sgn}(\sigma_x) \quad , \quad \dot{\varepsilon_y^{tr}} = \dot{\varepsilon_z^{tr}} = -\frac{1}{2}\dot{\varepsilon_x^{tr}} \tag{5.58}$$

to obtain:

$$\varepsilon_x^{tr}(y) = H\xi(y) \quad , \quad \sigma_x(y) = E(\xi)\,(\kappa y - H\xi(y)) \tag{5.59}$$

In order for a closed from solution to be obtainable, the difference in the elastic modulus of austenite and martensite is ignored ; i.e. $E = E^M = E^A$ which means $A = 0$. The differential equation in equation (5.54) (with $A = 0$) is linear first order with a general solution given by

$$\xi(y) = -\frac{\frac{\mathfrak{M}^{\mathrm{fwd}}}{EH}}{\left(\frac{B}{EH} - H\right)^2}\kappa - \frac{\frac{D}{EH} + \kappa y}{\frac{B}{EH} - H} + \zeta\,\mathrm{e}^{\left(\frac{\frac{B}{EH} - H}{\frac{\mathfrak{M}^{\mathrm{fwd}}}{EH}}y\right)} \tag{5.60}$$

Which after applying the boundary condition in equation (5.56) and replacing ζ, yields

$$\xi(y) = -\frac{D + HE\kappa y}{B - EH^2} - \frac{HE\mathfrak{M}^{\mathrm{fwd}}}{(B - EH^2)^2}\kappa\left(1 - \mathrm{e}^{\frac{B - EH^2}{\mathfrak{M}^{\mathrm{fwd}}}(y-r)}\right) \tag{5.61}$$

The first term on the r.h.s of the above equation pertains to the local part of the model and gives a bilinear distribution in ξ, while the second term is the exponential contribution from the nonlocal part. The moment, M, applied to the beam section corresponding to the current curvature, κ, can be readily calculated from:

$$M(\kappa) = \int_{-h}^{h} W\sigma_x(y)\,y\,\mathrm{d}y = 2W\int_0^h \sigma_x(y)\,y\,\mathrm{d}y \tag{5.62}$$

using equation (5.59) and (5.61) in the integration of (5.62) results in:

$$M^* = \kappa^* + \frac{3}{\beta-1}\left\{\frac{1}{2}\frac{\mathcal{M}^{\text{fwd}*}}{\beta-1}\left(\kappa^* - \frac{1}{\kappa^*}\right) + \frac{1}{3}\kappa^* + \frac{1}{6}\left(\frac{1}{\kappa^*}\right)^2 - \frac{1}{2} - \left(\frac{\mathcal{M}^{\text{fwd}*}}{\beta-1}\right)^2 \kappa^* \left[\left(1 - \frac{\mathcal{M}^{\text{fwd}*}}{\beta-1}\right) \mathrm{e}^{\frac{\beta-1}{\mathcal{M}^{\text{fwd}*}}\left(1-\frac{1}{\kappa^*}\right)} - \left(\frac{1}{\kappa^*} + \frac{\mathcal{M}^{\text{fwd}*}}{\beta-1}\right)\right]\right\} \tag{5.63}$$

which is the equation of Moment-Curvature in a non-dimensional form for pure bending of an SMA beam.

The non-dimensional parameters introduced in (5.63) are defined. The non-dimensional moment, M^*, and non-dimensional curvature are obtained based on a normalization using the critical moment and curvature to start the forward transformation, $M_{\text{s}}^{\text{fwd}}$ and $\kappa_{\text{s}}^{\text{fwd}}$.

$$M^* = \frac{M}{M_{\text{s}}^{\text{fwd}}} = \frac{M}{\frac{2}{3}Wh^2\sigma_{\text{s}}^{\text{fwd}}} \quad , \quad \kappa^* = \frac{\kappa}{\kappa_{\text{s}}^{\text{fwd}}} = \frac{Eh}{\sigma_{\text{s}}^{\text{fwd}}}\kappa \tag{5.64}$$

In addition, $\mathcal{M}^{\text{fwd}*}$ can be defined as the non-dimensional form of the nonlocal parameter as well as the non-dimensional hardening or β.

$$\beta = \frac{B}{EH^2} \quad , \quad \mathcal{M}^{\text{fwd}*} = \frac{1}{EH^2}\frac{\mathcal{M}^{\text{fwd}}}{h} \quad \text{or} \quad \frac{\mathcal{M}^{\text{fwd}*}}{\beta-1} = \frac{1}{h}\frac{\mathcal{M}^{\text{fwd}}}{B-EH^2} \tag{5.65}$$

It is also helpful to write down the local non-dimensional response of the SMA beam or:

$$M^* = \kappa^* + \frac{3}{\beta - 1}\left\{\frac{1}{3}\kappa^* + \frac{1}{6}\left(\frac{1}{\kappa^*}\right)^2 - \frac{1}{2}\right\} \tag{5.66}$$

Comparing equations (5.63) and (5.66) reveals the size effect in the response for pure bending of SMA beams. The first term on the r.h.s of both equations pertains, clearly, to the elastic portion of the SMA behaviour; i.e. $M^* = \kappa^*$. This portion is size independent as the nonlocality is only considered for the *inelastic* part of the response. Nonetheless, the rest of equation (5.66) is also size independent, as there is no effect from the thickness of the beam in the hardening parameter β. Therefore, the moment-curvature response using the conventional local model cannot principally capture the size effect in the pure bending of SMAs because it can be fully normalized with respect to the thickness h. The response according to the local constitutive model, on the other hand, is able to capture the size effect, all thanks to the non-dimensional local parameter $\mathcal{M}^{\text{fwd}^*}$ which *depends* on the thickness; $\mathcal{M}^{\text{fwd}^*} = \frac{1}{EH^2}\frac{\mathcal{M}^{\text{fwd}}}{h}$. Hence by changing $\mathcal{M}^{\text{fwd}^*}$ either through $\mathcal{M}^{\text{fwd}}$ or h, the moment-curvature response of the beam changes.

The analysis of the reverse transformation can be performed by following similar steps to the above. Upon unloading, the reverse transformation first begins where the differential equation, (4.80) re-written below, is first satisfied.

$$-H\sigma_x(y) - A\sigma_x(y)^2 + B'\xi(y) + \mathcal{M}^{\text{rev}}\frac{\mathrm{d}\xi(y)}{\mathrm{d}y} + D' = 0 \tag{5.67}$$

$$
\begin{aligned}
B' &= \rho b^A = H\left(\sigma_{\mathrm{s}}^{\mathrm{rev}} - \sigma_{\mathrm{f}}^{\mathrm{rev}}\right) \\
D' &= -\rho\Delta s_0^{\mathrm{rev}}\, T + \rho\Delta u_0 + (\mu_1 - \mu_2) - Y = H\sigma_{\mathrm{f}}^{\mathrm{rev}}
\end{aligned}
\tag{5.68}
$$

It can be shown that the reverse transformation begins at the outer layer and propagates towards the neutral axis with progression of unloading. The interface between the region of reverse transformation and the region where unloading takes place elastically is denoted by r'. The reverse transformation finishes when r' reaches r. Although extremely lengthy, it is possible to develop a closed-from analytical response for unloading step as well. Equation (5.67) can be integrated considering the continuity of ξ and $\frac{\mathrm{d}\xi(y)}{\mathrm{d}y}$ at $y = r'$.

However, a numerical algorithm is developed to solve both of the ODEs in (5.54) and (5.67) for forward and reverse transformations using the finite difference method.

The results of this analysis are presented in chapter 6.

6. RESULTS AND DISCUSSIONS

The results of the developed solutions for the response of SMA structures, studied in chapter 5, are presented in this section. Specifically, the loading and unloading of the SMA micropillar and SMA beam are analyzed using the solutions developed with the SMA gradient-based model III. Numerical solution for the system of equations expressed in sections 5.2.2 and 5.4 are considered for the purpose of this chapter.

First, the effect of the nonlocal parameter defined in the gradient-based model III is investigated within a parametric analysis of the SMA micropillar compression as well as beam bending problems.

Second, an attempt is made to calibrate the introduced nonlocal parameter using the experiential results for the compression of Ni-Fe-Ga SMA micropillars expressed in the work of Ozdemir et al. [112]. The model can qualitatively predict the hardening observed in the results as the diameter of the pillars decreases.

6.1 Parametric study for the effect of the nonlocal parameter

A parametric study of the effect of the nonlocal parameter introduced in the gradient-based SMA model III is performed in this section. The linear hardening function is considered here for the response of the local part of the model with the calibration stress $\sigma^* = 0$. Where nondimensional numbers could not be realized, the SMA material properties listed in table (6.1) are used for the local part of the SMA constitutive model. They are selected from the properties of a typical NiTi shape memory alloy.

First, the effect of changing the geometry of the micropillar on its compressive stress-strain response is investigated. Modeling is performed both using the conventional local SMA constitutive model and also the nonlocal gradient-based model III

Table 6.1: SMA material properties used in the presentation of the modeling results.

Thermoelastic Constants		Transformation Properties					
$E^A = 85000$	[MPa]	$A_s = 27$	$A_f = 47$	[°C]	$C_A = 10$	$[\frac{\text{MPa}}{\text{°C}}]$	$H = 5.5$ [%]
$E^M = 75000$		$M_s = -53$	$M_f = -73$		$C_M = 10$		$T = 67$ [°C]

[1]. The pillars studied have various dimensions listed in table (6.2). In addition, they have the material properties as given in table (6.1) and are loaded under a compressive force applied at the top surface while at a constant temperature $T = 67°C > A_f$.

In compression of a micropillar, the nominal stress S^* is defined as the compressive force divided by the mid-height cross sectional area of the pillar. Also, the nominal strain ε^* is defined as the displacement at the top of the pillar normalized by its height. Due to having higher stresses, the transformation starts at the top of the pillar and propagates to the bottom by further loading.

The nominal stress-strain responses are plotted in figures 14 to 16.

In Figure 14, the pillars are proportionally increased in size while maintaining the similarity in the geometry; as listed in the first row of table (6.2). The response predicted by the local conventional model is essentially the same for all of the micropillars. However, the nonlocal model with $\mathcal{M}^{\text{rev}} = \mathcal{M}^{\text{fwd}} = 1\,\text{mm.MPa}$ demonstrates an increase in the hardening as the pillar becomes smaller.

Figure 15 demonstrates the response of three micropillars that have similar heights and tapers. A higher stress is required to begin the transformation as the pillar diameter increases which is evident in the results of the local model. Therefore, it is difficult to associate any observed hardening in the response to the nonlocality and size effects.

[1]Please refer to section 5.2.2 for description of notation and model parameters.

Table 6.2: The dimensions of the pillars used for the parametric studies in figures 14 to 16.

Case		Top Diameter d μm	Bottom Diameter D μm	Height h μm	Taper Angle $\theta = \tan^{-1}\left(\frac{D-d}{2h}\right)$
Figure 14	(a)	1	5	10	11.31°
	(b)	2	10	20	11.31°
	(c)	4	20	40	11.31°
Figure 15	(a)	1	5	10	11.31°
	(b)	6	10	10	11.31°
	(c)	11	15	10	11.31°
Figure 16	(a)	1	5	10	11.31°
	(b)	3	5	10	5.71°
	(c)	5	5	10	0.0°

The comparison between the local conventional and nonlocal modeling results for pillars with similar bottom diameters and heights is shown in Figure 16. For the case of the straight pillar, due to a lack of taper and hence nonhomogenity in the stress field, the gradient-based model reduces to the local conventional model. Hence the stress-strain response from two models does not differ.

As noted, the first case represented by pillars with similar geometries is the most suitable to investigate the size effect since it allows for separation of the material response form the structural response (Figure 16). The nondimensional parameter studied for this case is $\varpi = \frac{\mathcal{M}^{\text{fwd}}}{BD}\eta$ where the taper is defined as $\eta = \frac{D-d}{h}$ with D, d, and h being the bottom and top diameters and the height of the pillar. It is assumed for the parametric study in this section that the nonlocal parameters take the same value for the forward and reverse transformations $\mathcal{M}^{\text{rev}} = \mathcal{M}^{\text{fwd}}$. ϖ combines the nonlocal parameter with the geometric aspects of the pillar that contribute to the nonhomogeneity in the distribution of stress and martensite volume fraction.

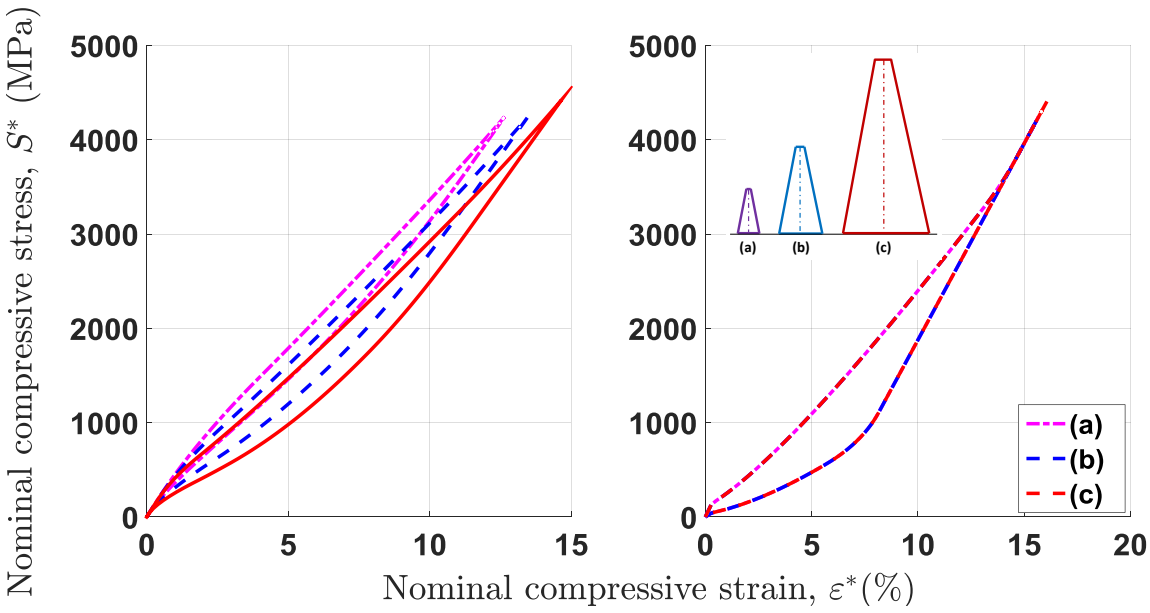


Figure 14: Effect of the geometry on the nominal compressive stress-strain response of the SMA micropillars shown in the inset as for modeling with (left) $\mathcal{M}^{\text{rev}} = \mathcal{M}^{\text{fwd}} = 1\,\text{mm.MPa}$ and (right) $\mathcal{M}^{\text{rev}} = \mathcal{M}^{\text{fwd}} = 0$ (local model).

Figure 17 illustrates the nominal stress-strain results for various choices of the parameter ϖ.

On one hand, if either the taper in the pillar increases while keeping the same material, or if materials with higher nonlocal parameters $\mathcal{M}^{\text{fwd}}$ are used in pillars of same geometry, the response of the pillar would demonstrate a hardening effect. On the other hand, if the dimensions of the pillar increase proportionally without changing the aspect ratios and the taper, its response becomes closer to that of the classical nonlocal model, thus capturing the size effect observed in the SMA micro/nanopillars.

This phenomenon is additionally explained by the plots in Figure 18. While the conventional local model predicts a completed transformation for the top portion of

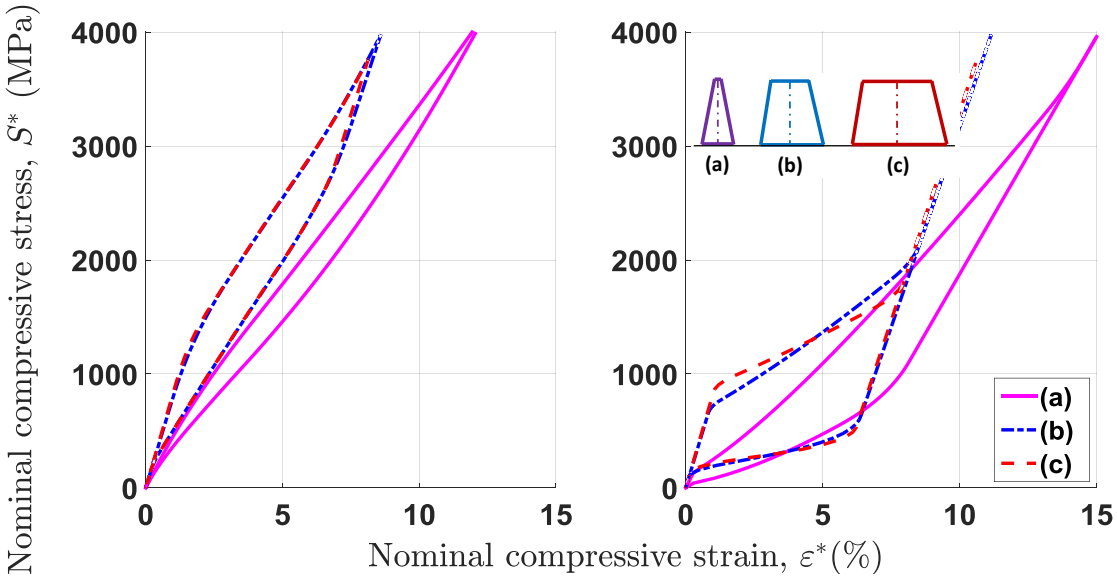


Figure 15: Effect of the geometry on the nominal compressive stress-strain response of the SMA micropillars shown in the inset as for modeling with (left) $\mathcal{M}^{\text{rev}} = \mathcal{M}^{\text{fwd}} = 1\,\text{mm.MPa}$ and (right) $\mathcal{M}^{\text{rev}} = \mathcal{M}^{\text{fwd}} = 0$ (local model).

the pillar and a narrow transforming region, the nonlocal model, however, results in lower amounts of transformation and a wider transforming region. It is worth mentioning that the nonlocal aspect of the model is tied to the existence of nonhomogeneity in the stress state of the structure. This can arise from the heterogeneity in the material properties. It can also be due to loading and boundary conditions or the geometry being considered resulting a nonhomogeneous state of stress.

Furthermore, the effect of the nonlocal parameter $\mathcal{M}^{\text{fwd}}$ (mm.MPa) on the pure bending of SMA beams is investigated[2]. For an SMA beam with $W = 1$ (mm), the plot of the nondimensional moment vs. nondimensional curvature is illustrated in Figure 19. It is, again, assumed that $\mathcal{M}^{\text{rev}} = \mathcal{M}^{\text{fwd}}$.

[2]Please refer to section 5.4 for description of notation and model parameters.

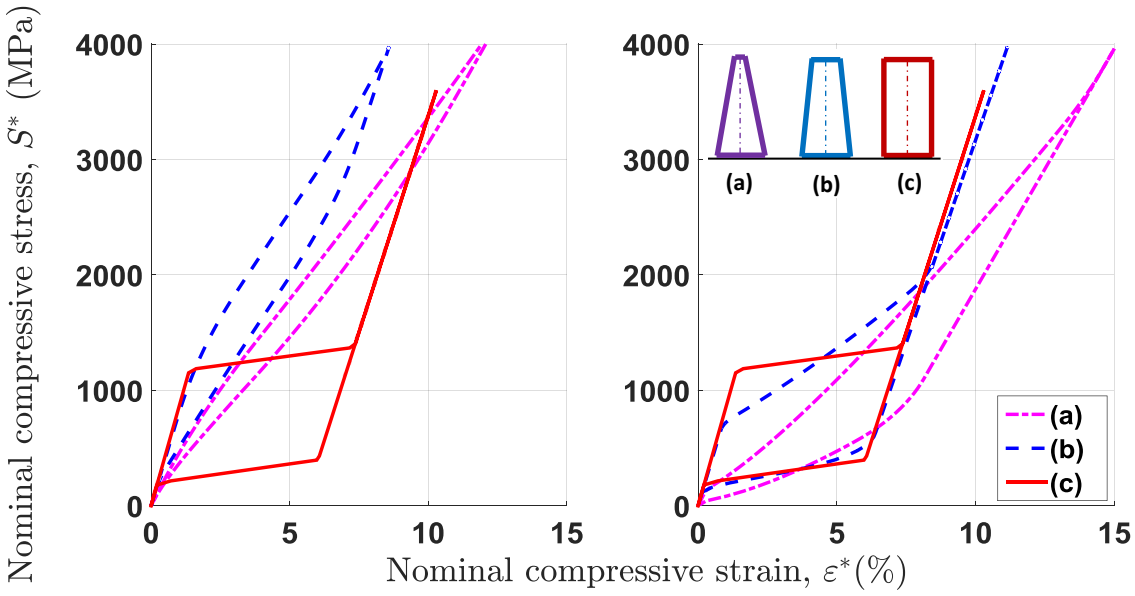


Figure 16: Effect of the geometry on the nominal compressive stress-strain response of the SMA micropillars shown in the inset as for modeling with (left) $\mathcal{M}^{\text{rev}} = \mathcal{M}^{\text{fwd}} = 1\,\text{mm.MPa}$ and (right) $\mathcal{M}^{\text{rev}} = \mathcal{M}^{\text{fwd}} = 0$ (local model).

The case where the non-dimensional nonlocal parameter vanishes, $\mathcal{M}^{\text{fwd}^*} = 0$, pertains to the local conventional model . Inclusion of the nonlocal effect in the model results in a macroscopic hardening effect. More energy is dissipated due to the existence of the nonlocal term and therefore more flexural work is required to achieve a certain level of curvature.

This plot can act as the characterization point for the nonlocal parameter. Given an SMA with known local properties (as listed in table 6.1 for example), it is possible to determine the value of the nonlocal parameter by calibrating with respect to the nondimensional flexural response of SMA samples with varying thicknesses. Figure 20 shows the variation in the distribution of martensite volume fraction for this SMA beam under a curvature of $\kappa^* = 3.3$ during loading and $\kappa^* = 1.6$ during

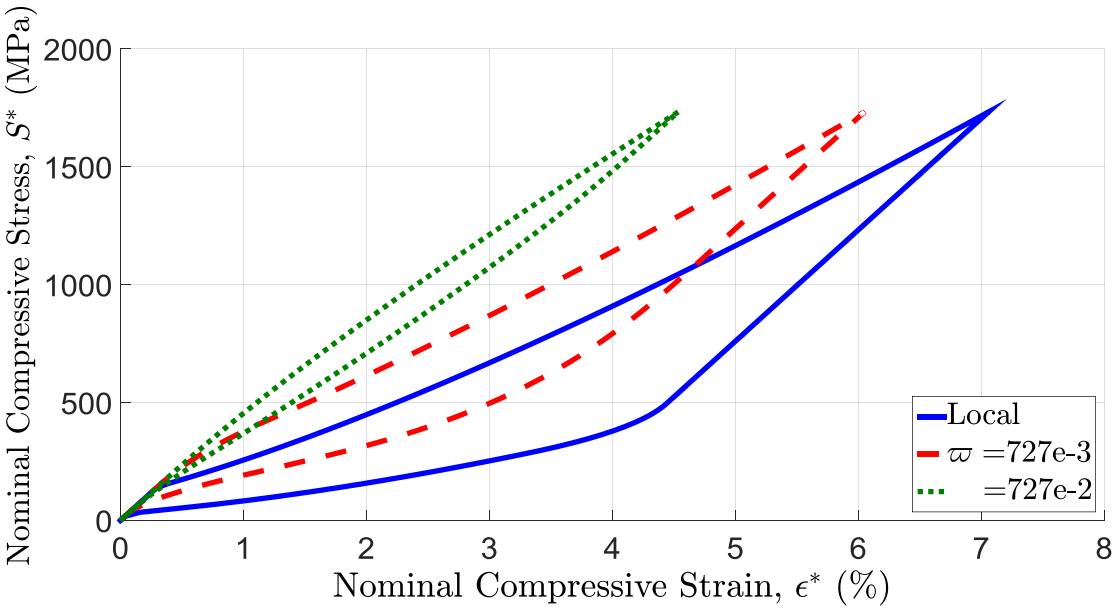
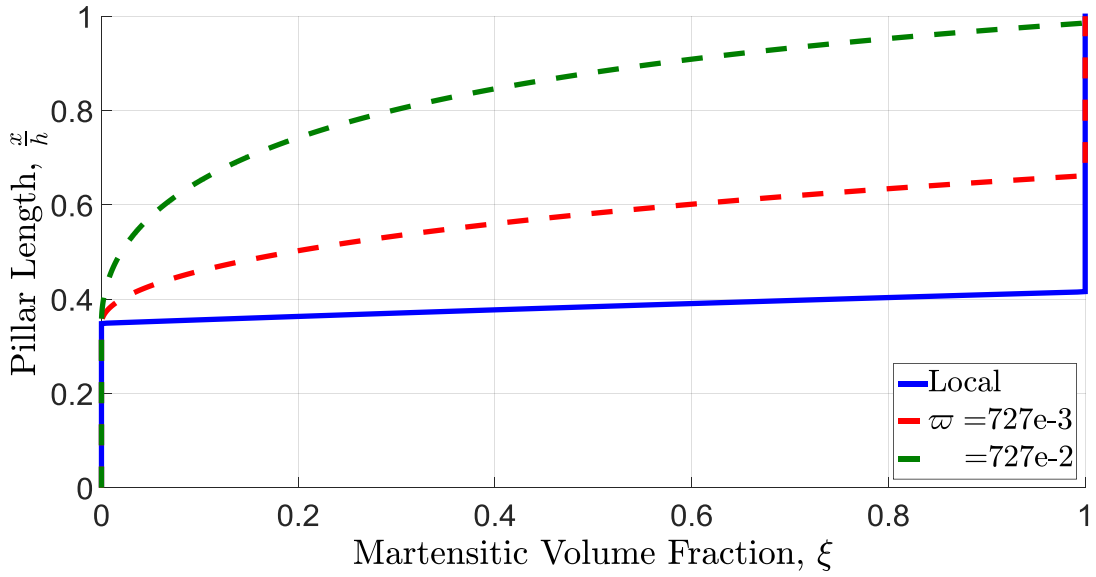
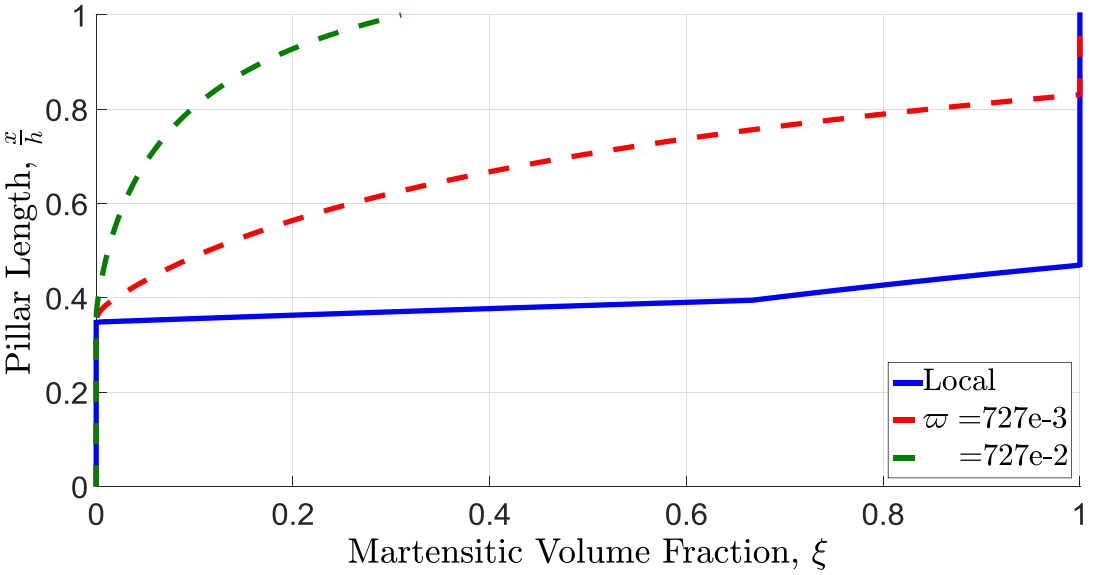


Figure 17: Effect of the nonlocal parameter $\varpi = \frac{\mathcal{M}^{\text{fwd}}}{BD}\eta$ on the nominal stress-strain response of an SMA micropillar under compression.

unloading. Through the nonlocal model, the inelastic energy generated contributes both to the martensitic transformation and also its gradient. The higher the nonlocal effect, the lower the total amount of transformation. Note that the conventional local model, ($\mathcal{M}^{\text{fwd}^*} = 0$), results in a linear distribution of ξ, as expected from the linear choice for the hardening function. Also as expected, note that the model does not capture any size effect in the elastic regime.



(a) During loading at $S^* = 1730$MPa



(b) During unloading at $S^* = 433$MPa

Figure 18: Effect of the nonlocal parameter $\varpi = \frac{\mathcal{M}^{\text{fwd}}}{BD}\eta$ on the distribution of martensite volume fraction in SMA pillar of figure 17.

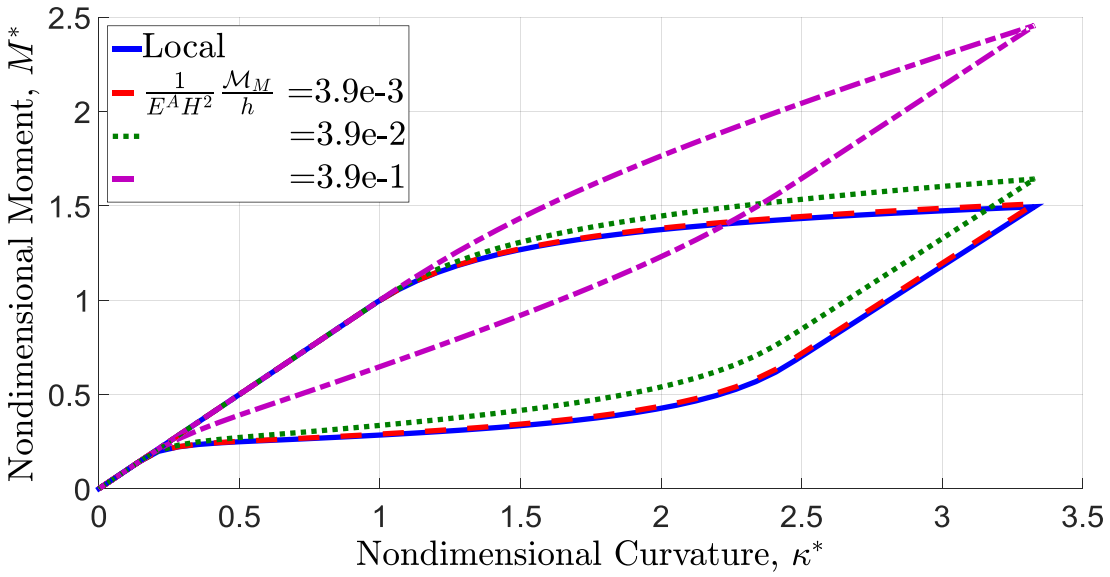
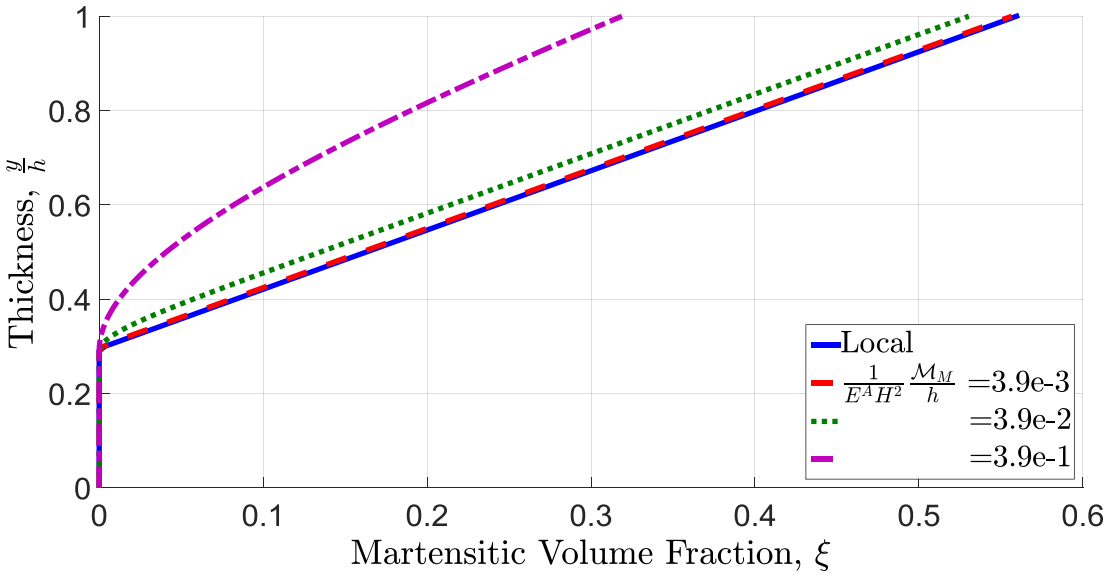
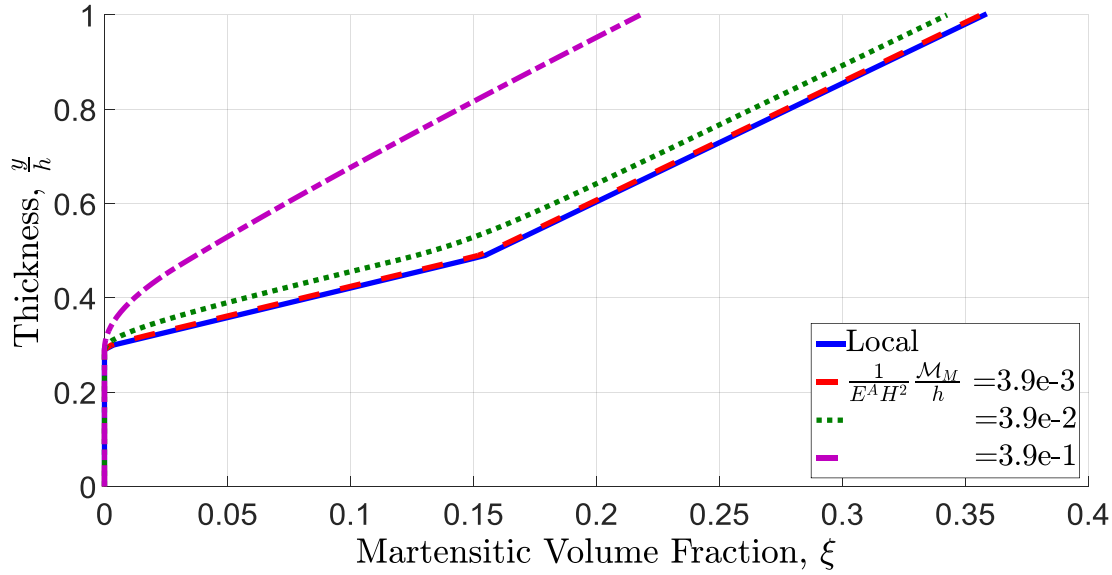


Figure 19: Effect of the nonlocal parameter $\mathcal{M}^{\text{fwd}*} = \frac{1}{E^A H^2}\frac{\mathcal{M}^{\text{fwd}}}{h}$ on the nondimensional moment-curvature response of an SMA beam with $W = 1$.



(a) During loading at $\kappa^* = 3.3$



(b) During unloading at $\kappa^* = 1.6$

Figure 20: Effect of the nonlocal parameter $\mathcal{M}^{\text{fwd}*} = \frac{1}{E^A H^2}\frac{\mathcal{M}^{\text{fwd}}}{h}$ on the distribution of martensite volume fraction in SMA beam of figure 19.

6.2 Qualitative prediction of the SMA micropillar experiemtal resulst

The capability of the proposed nonlocal model in capturing size effect in the compression of SMA micropillars and bending of SMA thin films was qualitatively shown so far. Pillars of various diameters and beams of various thicknesses from the same shape memory alloy show different responses. As the diameter of the SMA pillar or thickness of the SMA beam increases, its response becomes closer to that of the bulk material. This is supported by the fact that by moving away from small scales, the inelastic (either phase transformation or dislocation plasticity) behavior of the materials become independent of the size and can be nondimensionalized based on the mode and complexity of the deformation. SMA pillars of smaller diameters and beams of smaller thicknesses, according to Figures 19 and 17, show a hardening effect demonstrating the experimentally established "smaller is stronger".

In section 1.2, an overview of the experimental observations for size effect in the response of shape memory alloys was presented. Figure 7, obtained form the work of Ozdemir et. al. [112], is demonstrated again in this section (Figure 22). It shows the nominal stress-strain response for compression tests on $Ni_{54}Fe_{19}Ga_{27}$ shape memory alloy (SMA) single crystalline micropillars with various diameters D. Due to the existing taper in the geometry, D is taken to be the average diameter. The scanning electron microscopy images of 10 and 5 μm micropillars before and after the compression test are shown, as an example, in Figure 21. The red arrows point to the twin marks on the surface of the pillars. From these images, the corresponding sizes of the pillars, i.e. the top diameter d, the bottom diameter D, and the height h are obtained and listed in table 6.3.

The capability of the nonlocal SMA constitutive model in predicting the observed size effect in the response of the micropillars is investigated. To that end, model III

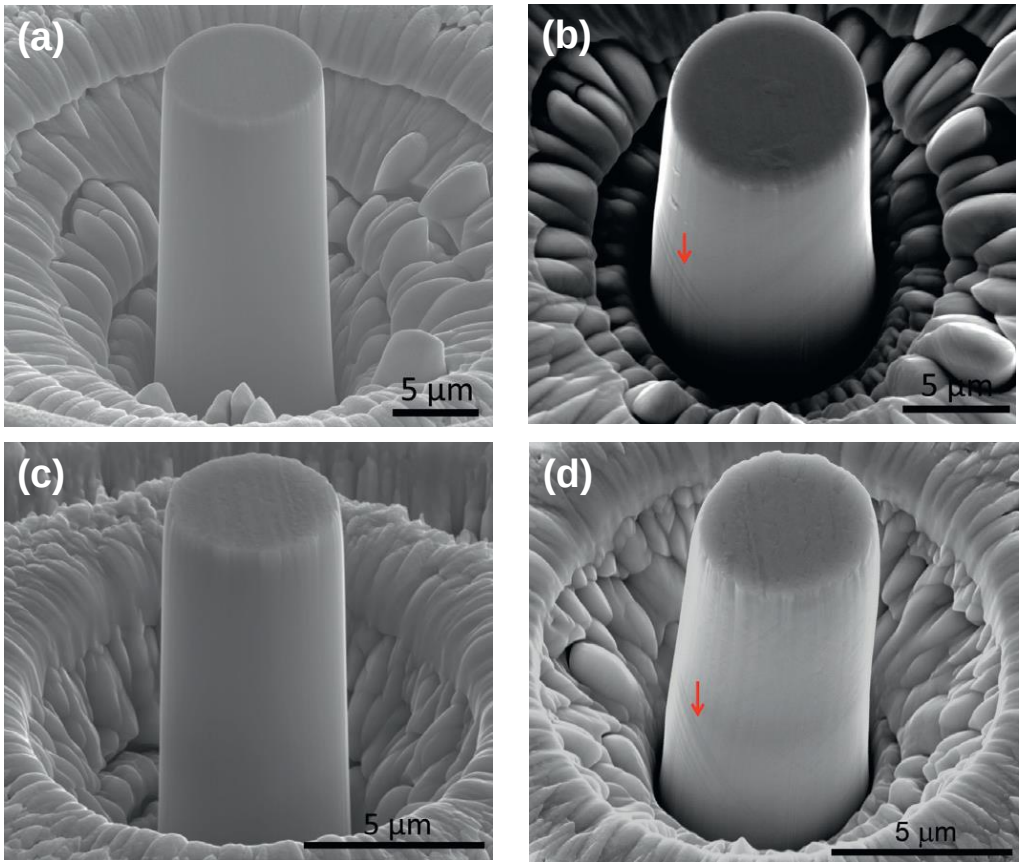


Figure 21: SEM images of the $Ni_{54}Fe_{19}Ga_{27}$ micropilars before and after deformation. (a)-(b) the 10 μm and (c)-(d) the 5 μm micropillars. Red arrow shows the twinning marks on the surface [112].

as developed and described in section 4.5, is used as it contains fewer numbers of nonlocal material constants that need to be identified and calibrated. Figure 22 contains the stress-stress responses of the pillars with average diameters form 10 μm to 420 nm as well as that of the bulk material. As verified in section 5.1, for a homogeneous prismatic bar of SMA, the stress-strain outcome of the nonlocal models is reduced to that of the conventional local model. Therefore, it is possible to calibrate the SMA material properties used in the local part of the model using only the data

Table 6.3: The dimensions of the pillars used for the calibration/prediction of the experimental results in [112] for compression of $Ni_{54}Fe_{19}Ga_{27}$ SMA micropillars.

Micropillar's Nominal Diameter	Top Diameter d μm	Bottom Diameter D μm	Height h μm	Taper Angle $\theta = \tan^{-1}\left(\frac{D-d}{2h}\right)$
10 μm	9.11	10.49	18.13	2.18°
5 μm	4.66	5.35	9.10	2.16°
1 μm	0.79	1.09	2.00	4.30°
585 nm	0.46	0.64	1.17	4.29°
420 nm	0.33	0.46	0.84	4.29°

available in [112] for the response of the bulk material at two different temperatures of $T = 22°C$ and $30°C$. This is shown in Figure 23 with the calibrated material constants presented in table 6.4. To that end, the smooth hardening function with properties given in table 5.2 is used.

Table 6.4: SMA material properties (local) calibrated using the plots shown in Figure 23 for the stress-strain response of the bulk $Ni_{54}Fe_{19}Ga_{27}$ SMA.

Thermoelastic Constants		Transformation Properties					
$E^A = 12$ $E^M = 25$	[GPa]	$A_s = -8$ $M_s = 0$	$A_f = 14$ $M_f = -83$	[°C]	$C_A = 4.7$ $C_M = 4.7$	$\left[\frac{\text{MPa}}{°\text{C}}\right]$	$H = 6.15$ [%]
		$\sigma^* = 400$ MPa		$n_1 = 0.96$	$n_2 = 0.31$	$n_3 = 0.50$	$n_4 = 0.15$

Next step is to determine the values of the nonlocal parameters $\mathcal{M}^{\text{fwd}}$ and $\mathcal{M}^{\text{rev}}$ as described in section 5.1.1. For that purpose, the behavior of SMA structures with non-uniform deformation, due to either the structure itself or to a nonhomogenous

loading, must be investigated such that the nonlocal part of the model is activated. For the current problem, the taper in the pillar provides a nonhomogneous state of stress, hence a non-homogenous distribution of martensite volume fraction, along the axis that involves the nonlocal part of the SMA constitutive model. An schematic is presented in figure 24 that illustrates the geometry of the pillars to the scale. As seen form the schematic and also the last column of table 6.3, the taper in the pillars is small ($\sim 3^\circ$).

This can be further analyzed by modeling the response of the micropillars under compression using only the conventional local model with the SMA material properties (table 5.2) obtained via calibration of the bulk material response (Figure 23). The results are shown in Figure 25. The local conventional model obtains an almost identical response for all of the micropillars with no size effect. This requires the use of the SMA nonlocal constitutive model.

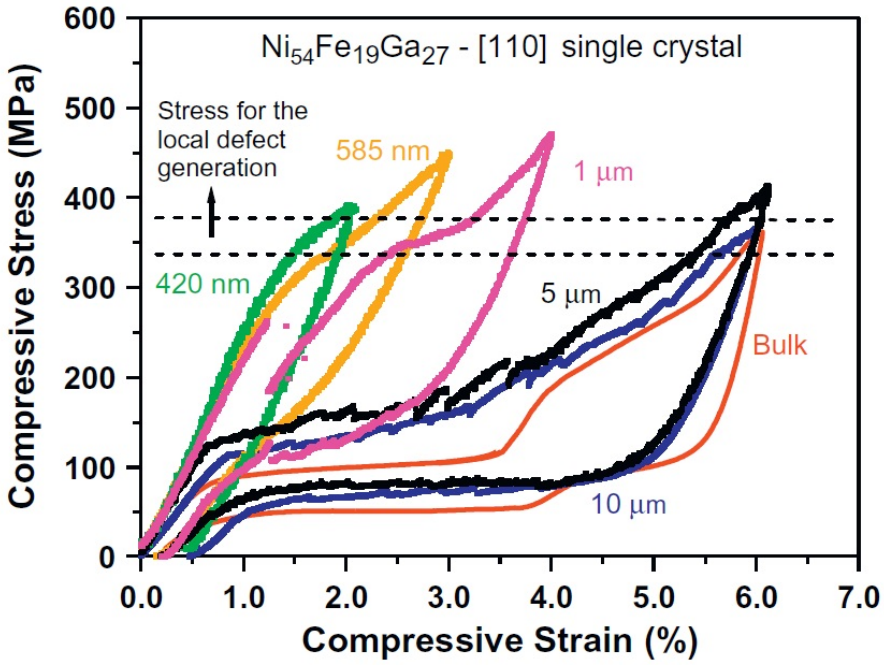
With the conventional SMA properties already calibrated using the bulk material response, the nonlocal pararmeters, $\mathcal{M}^{\text{fwd}}$ and $\mathcal{M}^{\text{rev}}$, can be determined by calibrating the analysis results of section 5.2.2 with the experimental responses of the $Ni_{54}Fe_{19}Ga_{27}$ micropillars. This is performed in Figure 26 were the compressive stress-strain plot of the $1\,\mu$m micropillar was used in order to obtain

$$\mathcal{M}^{\text{fwd}} = 0.09 \quad , \quad \mathcal{M}^{\text{rev}} = 0.02 \quad \text{mm.MPa} \tag{6.1}$$

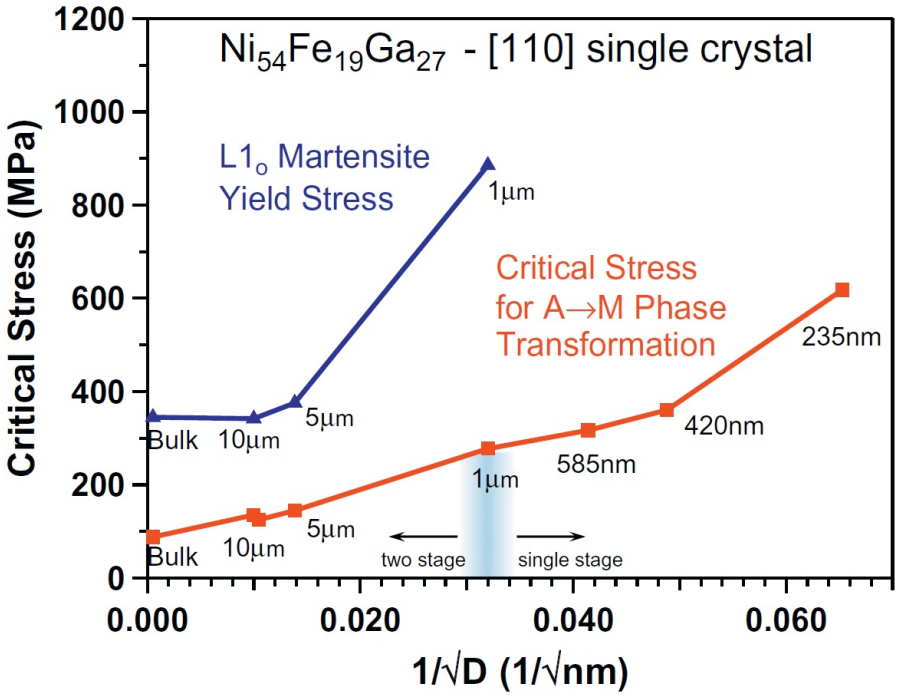
It is, hence, possible to predict the response of the other micropillars as shown in Figures 27 and 28. The model is capable of predicting the size effect and hardening observed in a qualitative fashion which is acceptable considering only having the two additional nonlocal parameters at one's avail. A more accurate prediction calls for considering other aspects of the response of the micropillars. As mentioned by

Ozdemir et al. [112] and observed in Figure 22, the $Ni_{54}Fe_{19}Ga_{27}$ SMAs undergo a two-stage transformation when tested in bulk that ceases from happening in the small scales. Previous works on the superelastic behavior of Ni-Ga-Fe ferromagnetic single-crystal shape memory alloys reported a two-stage transformation from the parent austenite phase (Cubic $L2_1$) to a $10M/14M$ modulated martensite (Monoclinic) and to a $L1_0$ martensite (Tetragonal) under an isothermal stress-strain response [36, 73, 139]. As the temperature of the loading changes, the sequence of inter-martensitic transformation also changes (Figure 29). Hamilton et al. [73] mentioned that by increasing the temperature, the transformation steps evolve form $A \rightleftarrows 10M \rightleftarrows 14M \rightleftarrows L1_0$ to $A \rightleftarrows 14M \rightleftarrows L1_0$ in tension and form $A \rightleftarrows 10M \rightleftarrows 14M$ to $A \rightleftarrows 14M \rightleftarrows L1_0$ to $A \rightleftarrows L1_0$ in compression.

Such a phenomenon also occurs while reducing the size of the specimen to the micron and submicron region. The current model can take into account only a single stage of transformation, as the response in Figure 23 was captured by averaging the two-stage step through the hardening of the single-stage transformation. Effective modeling and prediction of the response of $Ni_{54}Fe_{19}Ga_{27}$ micropillars cannot be achieved without considering this phenomenon.



(a) The superelastic stress-strain response



(b) The critical stresses for the start of martensitic transformation as well as martensitic plastic yield as a function of the micropillar diameter.

Figure 22: The size effect in compression of $Ni_{54}Fe_{19}Ga_{27}$ SMA micropillars with various diameters D [112].

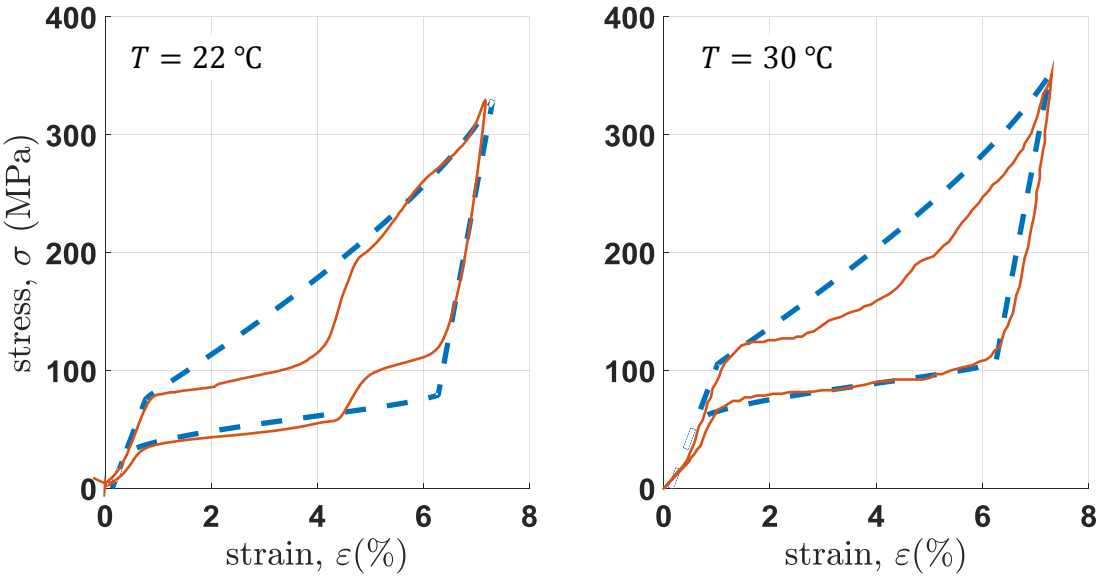


Figure 23: Calibration of the local SMA material properties using the experimental compression response of the bulk $Ni_{54}Fe_{19}Ga_{27}$ SMA at two different temperatures. (solid line) Experiments [112], (dashed line) SMA constitutive model.

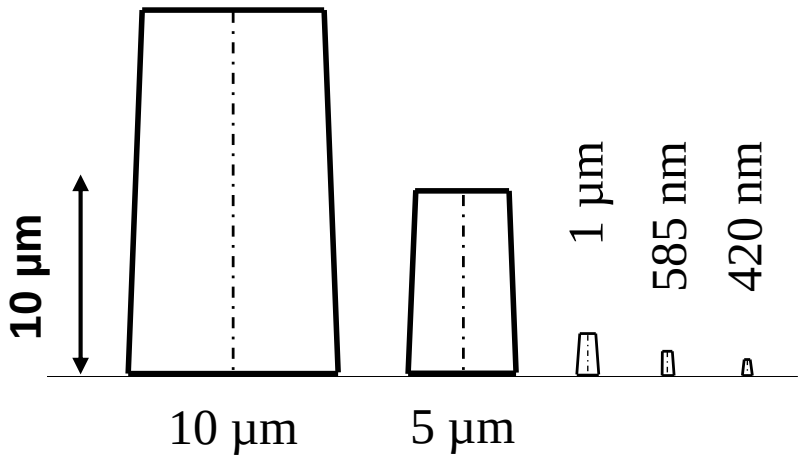


Figure 24: The schematics of the micropillars studied with dimensions given in table 6.3.

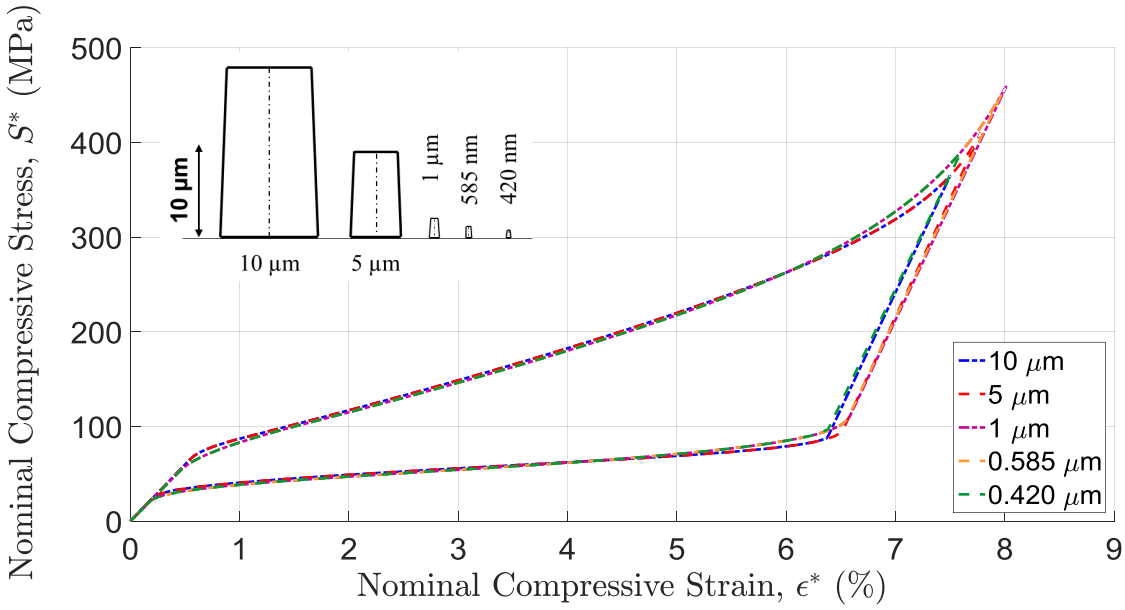


Figure 25: The isothermal room temperature compressive stress-strain response of the SMA micropillars, with the geometries shown in the inset, using the local SMA constitutive model.

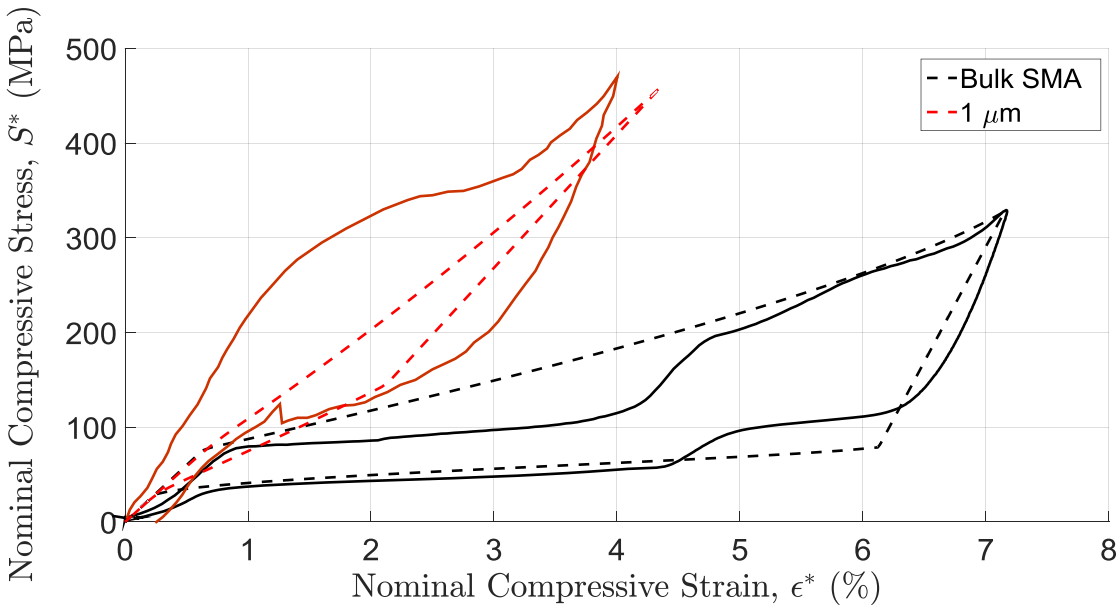


Figure 26: The isothermal room temperature compressive stress-strain response of the bulk $Ni_{54}Fe_{19}Ga_{27}$ shape memory alloy and the $1\,\mu$m micropillar used for the calibration of the SMA nonlocal constitutive model. (solid line) Experiments [112], (dashed line) SMA constitutive model.

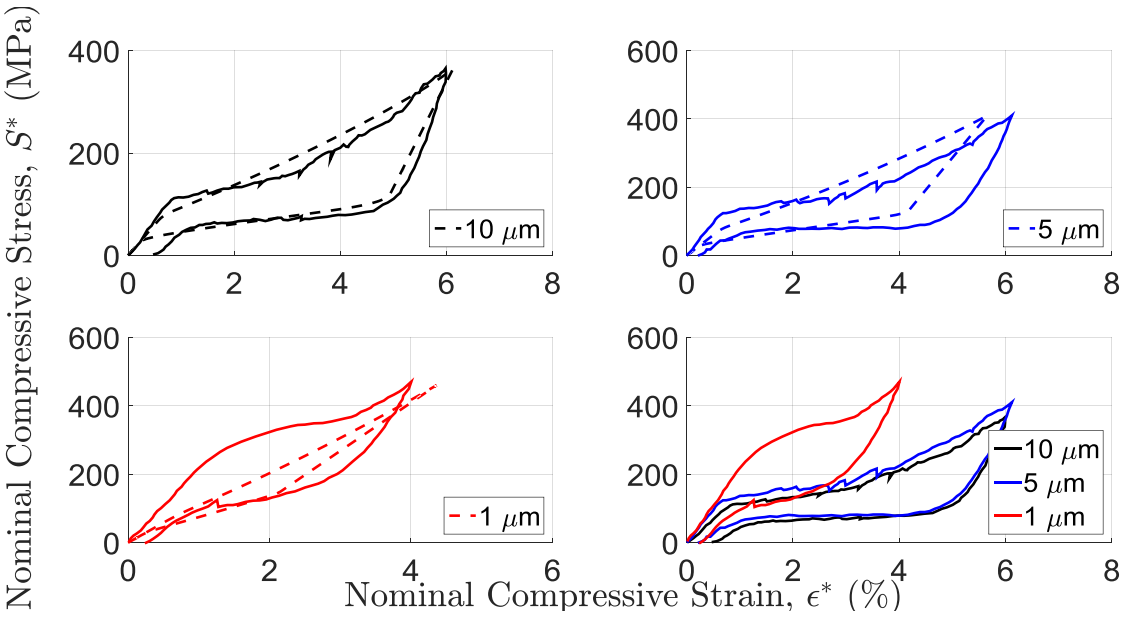


Figure 27: The isothermal room temperature compressive stress-strain response of $Ni_{54}Fe_{19}Ga_{27}$ shape memory alloy micropillars. (solid line) Experiments [112], (dashed line) SMA nonlocal model prediction.

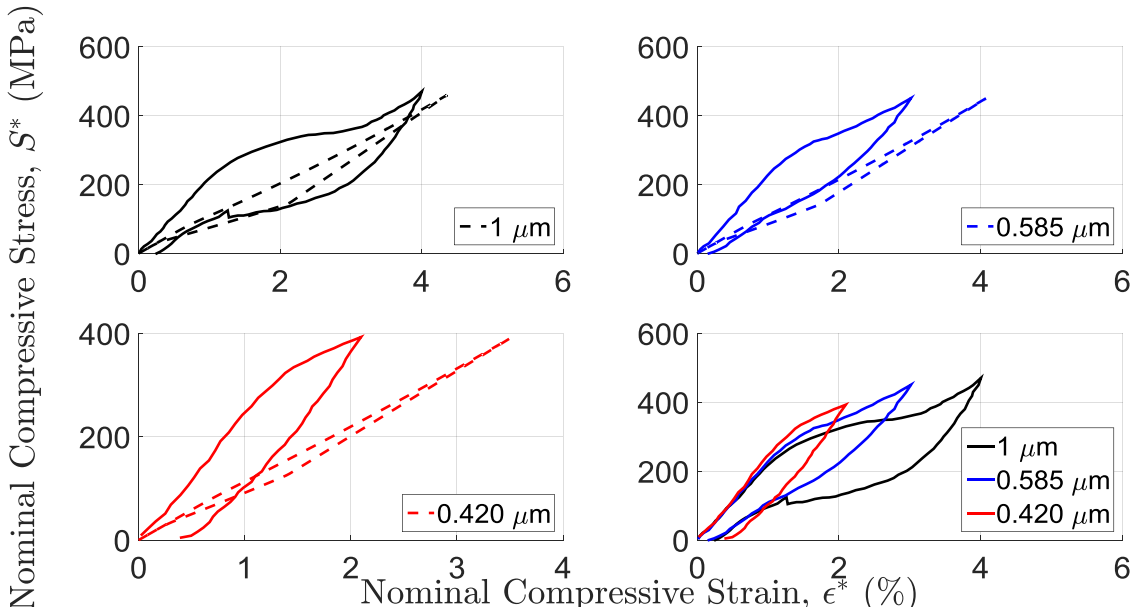


Figure 28: The isothermal room temperature compressive stress-strain response of $Ni_{54}Fe_{19}Ga_{27}$ shape memory alloy micropillars. (solid line) Experiments [112], (dashed line) SMA nonlocal model prediction.

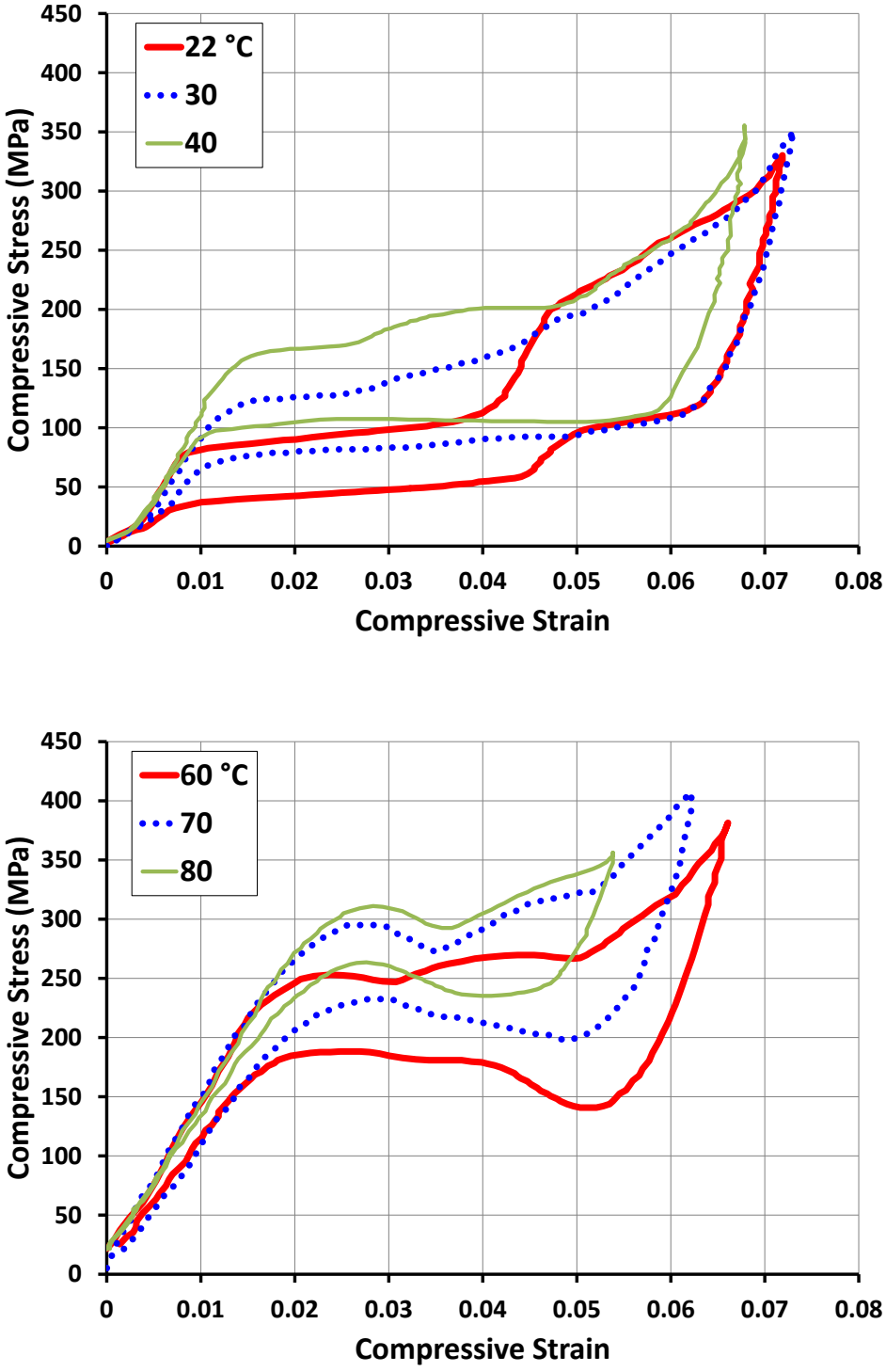


Figure 29: The superelastic stress-strain response of the bulk $Ni_{54}Fe_{19}Ga_{27}$ SMA at different temperatures (Data obtained form Figure 11 in [112]).

7. CONCLUSIONS AND FUTURE WORK

The response of shape memory alloy (SMA) structures in small scales show a deviation from that of the bulk material. $Ni_{54}Fe_{19}Ga_{27}$ ferromagnetic SMA micropillars, for instance, demonstrated a significantly increased hardening in their compressive stress-strain response as the diameter of the micropillars approached the micron and submicron scales [112]. A review of the experimentally observed size effect in shape memory alloy structures was presented with a categorization according to the dimensionality of the specimens tested. From nano particles and powders to 3D specimens with ultrafine grains, reducing the size of an external characteristic dimension as compared to the internal or intrinsic characteristic lengths dominating the martensitic transformation and its underlying mechanisms, contributed to the size effect. For the compression of SMA micro/nanopillars, as summarized in table 1.1, the scarcity of martensite nucleation sites with reduction in the diameter, refinement of twin formations, and relaxation effect form the free surfaces, are given as some of the physical grounds for the observed enhanced hardening as well as the change in the dissipation in the stress-strain response.

The conventional constitutive models, due to their lack of intrinsic length scales, are not able to capture such an experimentally observed size effect. A variety of nonlocal generalized continuum theories have been developed in order to model this behavior. These models include nonlocal implicit and explicit strain gradient theories for dislocation plasticity which were used to model the torsional response of thin wires, bending of thin films, growth of micron sized voids, and indentation size effect among many other observed phenomena.

In the current work, a generalized gradient-based, thermodynamically consistent

constitutive framework is established aimed at modeling the response of shape memory alloys. In this framework, the existence of some generalized surface and body forces are presumed that contribute to the free energy as work conjugates to the generalized variables of martensite volume fraction, transformation strain tensor, and their spatial gradients. The rate of evolution of the generalized variables can be obtained by invoking the principal of maximum dissipation after assuming a generalized transformation surface. The generalized transformation surface, in this gradient-based theory, is a differential equation in terms of the generalized thermodynamic forces. For conventional constitutive theories, the equation for the transformation surface is an algebraic one. The connection between this framework and the theories that use a configurational force (microforce) balance law is established by showing the latter to be a specific case of a generalized transformation surface.

In addition, gradient-based SMA constitutive models based on the theory of internal variables were deduced from the developed generalized constitutive modeling framework. The constitutive models, in this case, comprise of the conventional internal variables of martensitic volume fraction and transformation strain as well as their spatial gradients. Three different versions for such gradient based constitutive models were developed which included various energetic and dissipative length scales that can be calibrated experimentally. The length scales contribute to additional hardening in the structural SMA response.

The developed gradient-based internal variable constitutive models were simplified for 1D. A boundary value problem, defined by such models, contains the differential equation for the transformation surface including the gradients of internal variables as well as the equilibrium equation. Such a boundary value problem is solved analytically and, where impossible, numerically for pure bending of SMA beams, simple torsion of SMA cylindrical bars, and compression of SMA micro/nanopillars.

The most simplified version of the gradient-based SMA constitutive model, containing only the additional gradient of martensite volume fraction, can capture the size effect in the response of these structures demonstrating a stronger response for smaller sizes. Also, the calibrated model is shown to be able to qualitatively predict the experimentally observed response of Ni-Fe-Ga micropillars under compression, as presented in Ozdemir et al [112].

Future Work

The energetic and dissipative length scales were included in the current internal variable constitutive modeling framework in a phenomenological fashion. The nonlocal parameter, $\mathcal{M}$, in the gradient-based model III for example, represents the lower-scale microstructural effects in a continuum scale model. As proposed herein, it is possible to phenomenologically calibrate the nonlocal parameter as a material constant using, for example, nondimensional results from pure bending of SMA thin films.

The contribution from the nonlocal part of the model is closely coupled with the existence of nonhomogeneity in the stress state of the structure and hence the distribution of martensite volume fraction. This may be originated from either geometric or material heterogeneities. Nonetheless, if the state of the stress and martensite volume fraction are uniform, for example in uniaxial loading of an SMA wire, the model prediction reduces to the conventional one. Thus, the nonlocal model loses its capability in predicting any size effect for that case. The model can be enhanced by inclusion of additional terms in the free energy related to the surface energy of the SMA structure that can contain information from length scales such as, for the case of the wire, the surface to volume ratio of the structure.

The size effect was also observed in more complicated experiments such as nanoin-

dentation of shape memory alloys. The effective hardness was shown to be dependent on the depth of indentation being larger for lower depths [9]. Modeling such a behavior deems near impossible using analytical techniques. Therefore, implementation of the proposed gradient based constitutive model in a finite element numerical framework constitutes one of the necessary next steps in this work.

The response of $Ni_{54}Fe_{19}Ga_{27}$ single crystal micropillars was modeled using certain assumptions in order to simplify the analysis. Such a ferromagnetic SMA, in its bulk form, demonstrates a two-stage inter-martensitic transformation [36, 73, 139] that is inhibited when the size is reduced to the micron scale. The gradient-based SMA constitutive model as well as the corresponding conventional local SMA constitutive model are not capable of capturing such a transformation. In the current micropillar analysis, the two-stage transformation was approximated by a single-stage counterpart with an averaged hardening. In order to achieve a more quantitative prediction for the compressive response of the micropillars, it is necessary to develop a gradient-based constitutive model that incorporates the various phases of martensite as internal variables in its reduced local or conventional form. This contribution also remains to be implemented as a future work.

REFERENCES

[1] A. Acharya and J.L. Bassani. On non-local flow theories that preserve the classical structure of incremental boundary value problems. In *IUTAM Symposium on Micromechanics of Plasticity and Damage of Multiphase Materials*, pages 3–9. Springer, 1996.

[2] E.C. Aifantis. On the microstructural origin of certain inelastic models. *Journal of Engineering Materials and Technology*, 106(4):326–330, 1984.

[3] E.C. Aifantis. The physics of plastic deformation. *International Journal of Plasticity*, 3(3):211–247, 1987.

[4] E.C. Aifantis. On the role of gradients in the localization of deformation and fracture. *International Journal of Engineering Science*, 30(10):1279–1299, 1992.

[5] E.C. Aifantis. Pattern formation in plasticity. *International Journal of Engineering Science*, 33(15):2161–2178, 1995.

[6] E.C. Aifantis. Strain gradient interpretation of size effects. *International Journal of Fracture*, 95(1-4):299–314, 1999.

[7] E.C. Aifantis. Update on a class of gradient theories. *Mechanics of Materials*, 35(3):259–280, 2003.

[8] E.C. Aifantis. On the gradient approach–relation to eringens nonlocal theory. *International Journal of Engineering Science*, 49(12):1367–1377, 2011.

[9] A. Amini, W. Yan, and Q. Sun. Depth dependency of indentation hardness during solid-state phase transition of shape memory alloys. *Applied Physics Letters*, 99(2):021901, 2011.

[10] E.l. Arzt. Size effects in materials due to microstructural and dimensional constraints: a comparative review. *Acta Materialia*, 46(16):5611–5626, 1998.

[11] M.B. Babanly, V.A. Lobodyuk, and N.M. Matveeva. Size effect in martensite transformation in tinicu alloys. *Fizika Metallov i Metallovedenie*, 75(5):89–95, 1993. cited By (since 1996)6.

[12] H. Badnava, M. Kadkhodaei, and M. Mashayekhi. A non-local implicit gradient-enhanced model for unstable behaviors of pseudoelastic shape memory alloys in tensile loading. *International Journal of Solids and Structures*, 51(23):4015–4025, 2014.

[13] J.L. Bassani. Incompatibility and a simple gradient theory of plasticity. *Journal of the Mechanics and Physics of Solids*, 49(9):1983–1996, 2001.

[14] Z. Bažant and F.B. Lin. Non-local yield limit degradation. *International Journal for Numerical Methods in Engineering*, 26(8):1805–1823, 1988.

[15] Z.P. Bazant and M. Jirásek. Nonlocal integral formulations of plasticity and damage: survey of progress. *Journal of Engineering Mechanics*, 128(11):1119–1149, 2002.

[16] Z.P. Bazant and J. Planas. *Fracture and Size Effect in Concrete and Other Quasibrittle Materials*. New Directions in Civil Engineering. Taylor & Francis, 1997.

[17] W.L. Benard, H. Kahn, A.H. Heuer, and M.A. Huff. Thin-film shape-memory alloy actuated micropumps. *Microelectromechanical Systems, Journal of*, 7(2):245–251, 1998.

[18] K. Bhattacharya. *Microstructure of Martensite: Why it forms and how it gives rise to the shape-memory effect*. Oxford Series on Materials Modelling. OUP

Oxford, 2003.

[19] J.P. Boehler. Representations for isotropic and anisotropic non-polynomial tensor functions. In J.P. Boehler, editor, *Applications of Tensor Functions in Solid Mechanics*, volume 292 of *International Centre for Mechanical Sciences*, pages 31–53. Springer Vienna, 1987.

[20] G. Borino, P. Fuschi, and C. Polizzotto. A thermodynamic approach to nonlocal plasticity and related variational principles. *Journal of Applied Mechanics*, 66(4):952–963, 1999.

[21] L.C. Brinson. One-dimensional constitutive behavior of shape memory alloys: thermomechanical derivation with non-constant material functions and redefined martensite internal variable. *Journal of Intelligent Material Systems and Structures*, 4(2):229–242, 1993.

[22] J.D. Busch, A.D. Johnson, C.H. Lee, and D.A. Stevenson. Shapememory properties in niti sputterdeposited film. *Journal of Applied Physics*, 68(12), 1990.

[23] C.Q. Chen, Y. Shi, Y.S. Zhang, J. Zhu, and Y.J. Yan. Size dependence of young's modulus in zno nanowires. *Physical Review Letters*, 96:075505, Feb 2006.

[24] J.Y. Chen, Y. Wei, Y. Huang, J.W. Hutchinson, and K.C. Hwang. The crack tip fields in strain gradient plasticity: the asymptotic and numerical analyses. *Engineering Fracture Mechanics*, 64(5):625–648, 1999.

[25] Y. Chen and C.A. Schuh. Size effects in shape memory alloy microwires. *Acta Materialia*, 59(2):537–553, 2011.

[26] K. Choi, J.L. Kuhn, M.J. Ciarelli, and S.A. Goldstein. The elastic moduli of human subchondral, trabecular, and cortical bone tissue and the size-dependency

of cortical bone modulus. *Journal of Biomechanics*, 23(11):1103 – 1113, 1990.

[27] B.G. Clark, D.S. Gianola, O. Kraft, and C.P. Frick. Size independent shape memory behavior of nickel titanium. *Advanced Engineering Materials*, 12(8):808–815, 2010.

[28] B.D. Coleman and M.E. Gurtin. Equipresence and constitutive equations for rigid heat conductors. *Zeitschrift für angewandte Mathematik und Physik ZAMP*, 18(2):199–208, 1967.

[29] B.D. Coleman and M.L. Hodgdon. On shear bands in ductile materials. In *Analysis and Thermomechanics*, pages 227–255. Springer, 1987.

[30] E. Cosserat and F. Cosserat. Théorie des corps déformables. *Paris*, 1909.

[31] D.C. Dunand and P. Mllner. Size effects on magnetic actuation in ni-mn-ga shape-memory alloys. *Advanced Materials*, 23(2):216–232, 2011.

[32] A. Duval, M. Haboussi, and T. Ben Zineb. Modelling of localization and propagation of phase transformation in superelastic sma by a gradient nonlocal approach. *International Journal of Solids and Structures*, 48(13):1879–1893, 2011.

[33] J. Dyszlewicz. *Micropolar Theory of Elasticity*, volume 15. Springer Science & Business, 2012.

[34] D.G.B. Edelen, A.E. Green, and N. Laws. Nonlocal continuum mechanics. *Archive for Rational Mechanics and Analysis*, 43(1):36–44, 1971.

[35] D.G.B. Edelen and N. Laws. On the thermodynamics of systems with nonlocality. *Archive for Rational Mechanics and Analysis*, 43(1):24–35, 1971.

[36] C. Efstathiou, H. Sehitoglu, J. Carroll, J. Lambros, and H.J. Maier. Full-field strain evolution during intermartensitic transformations in single-crystal nifega. *Acta Materialia*, 56(15):3791–3799, 2008.

[37] R. Engelen, N.A. Fleck, R. Peerlings, and M. Geers. An evaluation of higher-order plasticity theories for predicting size effects and localisation. *International Journal of Solids and Structures*, 43(7):1857–1877, 2006.

[38] R. Engelen, M. Geers, and F. Baaijens. Nonlocal implicit gradient-enhanced elasto-plasticity for the modelling of softening behaviour. *International Journal of Plasticity*, 19(4):403–433, 2003.

[39] V.A. Eremeyev, L.P. Lebedev, and H. Altenbach. *Foundations of Micropolar Mechanics*. SpringerBriefs in Applied Sciences and Technology. Springer, 2012.

[40] A.C. Eringen. Line crack subject to shear. *International Journal of Fracture*, 14(4):367–379, 1978.

[41] A.C. Eringen. On nonlocal plasticity. *International Journal of Engineering Science*, 19(12):1461–1474, 1981.

[42] A.C. Eringen. On differential equations of nonlocal elasticity and solutions of screw dislocation and surface waves. *Journal of Applied Physics*, 54(9):4703–4710, 1983.

[43] A.C. Eringen. Theories of nonlocal plasticity. *International Journal of Engineering Science*, 21(7):741–751, 1983.

[44] A.C. Eringen. Vistas of nonlocal continuum physics. *International Journal of Engineering Science*, 30(10):1551 – 1565, 1992.

[45] N.A. Fleck and J.W. Hutchinson. A phenomenological theory for strain gradient effects in plasticity. *Journal of the Mechanics and Physics of Solids*, 41(12):1825–1857, 1993.

[46] N.A. Fleck and J.W. Hutchinson. Strain gradient plasticity. *Advances in Applied Mechanics*, 33:295–361, 1997.

[47] N.A. Fleck and J.W. Hutchinson. A reformulation of strain gradient plasticity. *Journal of the Mechanics and Physics of Solids*, 49(10):2245–2271, 2001.

[48] N.A. Fleck, G.M. Muller, M.F. Ashby, and J.W. Hutchinson. Strain gradient plasticity: Theory and experiment. *Acta Metallurgica et Materialia*, 42(2):475 – 487, 1994.

[49] S. Forest and E.C. Aifantis. Some links between recent gradient thermo-elasto-plasticity theories and the thermomechanics of generalized continua. *International Journal of Solids and Structures*, 47(25):3367–3376, 2010.

[50] J. Frenzel, E.P. George, A. Dlouhy, C. Somsen, M.F.X. Wagner, and G. Eggeler. Influence of ni on martensitic phase transformations in niti shape memory alloys. *Acta Materialia*, 58(9):3444–3458, 2010.

[51] C.P. Frick, B.G. Clark, S. Orso, P. Sonnweber-Ribic, and E. Arzt. Orientation-independent pseudoelasticity in small-scale niti compression pillars. *Scripta Materialia*, 59(1):7–10, 2008.

[52] C.P. Frick, B.G. Clark, A.S. Schneider, R. Maaß, S.V. Petegem, and H.V. Swygenhoven. On the plasticity of small-scale nickel–titanium shape memory alloys. *Scripta Materialia*, 62(7):492–495, 2010.

[53] C.P. Frick, S. Orso, and E. Arzt. Loss of pseudoelasticity in nickeltitanium sub-micron compression pillars. *Acta Materialia*, 55(11):3845 – 3855, 2007.

[54] C. Frommen, G. Wilde, and H. Rsner. Wet-chemical synthesis and martensitic phase transformation of aucd nanoparticles with near-equiatomic composition. *Journal of Alloys and Compounds*, 377(12):232 – 242, 2004.

[55] Y. Fu, H. Du, W. Huang, S. Zhang, and M. Hu. Tini-based thin films in mems applications: a review. *Sensors and Actuators A: Physical*, 112(2):395–408, 2004.

[56] Y. Fu and C. Shearwood. Characterization of nanocrystalline tini powder. *Scripta Materialia*, 50(3):319 – 323, 2004.

[57] Y.Q. Fu, J.K. Luo, S.E. Ong, S. Zhang, A.J. Flewitt, and W.I. Milne. A shape memory microcage of tini/dlc films for biological applications. *Journal of Micromechanics and Microengineering*, 18(3):035026, 2008.

[58] Y.Q. Fu, Sam Zhang, M.J. Wu, W.M. Huang, H.J. Du, J.K. Luo, A.J. Flewitt, and W.I. Milne. On the lower thickness boundary of sputtered tini films for shape memory application. *Thin Solid Films*, 515(1):80 – 86, 2006.

[59] Y. Ganor, D. Shilo, T.W. Shield, and R.D. James. Breaching the work output limitation of ferromagnetic shape memory alloys. *Applied Physics Letters*, 93(12), 2008.

[60] H. Gao, Y. Huang, W.D. Nix, and J.W. Hutchinson. Mechanism-based strain gradient plasticityi. theory. *Journal of the Mechanics and Physics of Solids*, 47(6):1239–1263, 1999.

[61] M. Geers, R. Ubachs, and R. Engelen. Strongly non-local gradient-enhanced finite strain elastoplasticity. *International Journal for Numerical Methods in Engineering*, 56(14):2039–2068, 2003.

[62] A.M. Glezer, E.N. Blinova, V.A. Pozdnyakov, and A.V. Shelyakov. Martensite transformation in nanoparticles and nanomaterials. *Journal of Nanoparticle Research*, 5(5-6):551–560, 2003.

[63] Y.I. Golovin. Nanoindentation and mechanical properties of solids in submicrovolumes, thin near-surface layers, and films: a review. *Physics of the solid State*, 50(12):2205–2236, 2008.

[64] A.E. Green and R.S. Rivlin. Simple force and stress multipoles. *Archive for Rational Mechanics and Analysis*, 16(5):325–353, 1964.

[65] P. Gudmundson. A unified treatment of strain gradient plasticity. *Journal of the Mechanics and Physics of Solids*, 52(6):1379–1406, 2004.

[66] J.R.C. Guimares. Excess driving force to initiate martensite transformation in fine-grained austenite. *Scripta Materialia*, 57(3):237 – 239, 2007.

[67] J.R.C. Guimares and P.R. Rios. Martensite start temperature and the austenite grain-size. *Journal of Materials Science*, 45(4):1074–1077, 2010.

[68] M.E. Gurtin. *An introduction to continuum mechanics*. Academic Press, 1982.

[69] M.E. Gurtin. Generalized ginzburg-landau and cahn-hilliard equations based on a microforce balance. *Physica D: Nonlinear Phenomena*, 92(3):178–192, 1996.

[70] M.E. Gurtin and L. Anand. A theory of strain-gradient plasticity for isotropic, plastically irrotational materials. part i: Small deformations. *Journal of the Mechanics and Physics of Solids*, 53(7):1624–1649, 2005.

[71] M.E. Gurtin and L. Anand. Thermodynamics applied to gradient theories involving the accumulated plastic strain: the theories of aifantis and fleck and

hutchinson and their generalization. *Journal of the Mechanics and Physics of Solids*, 57(3):405–421, 2009.

[72] K. Hackl and F.D. Fischer. On the relation between the principle of maximum dissipation and inelastic evolution given by dissipation potentials. *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science*, 464(2089):117–132, 2008.

[73] R.F. Hamilton, H. Sehitoglu, C. Efstathiou, and H.J. Maier. Inter-martensitic transitions in ni–fe–ga single crystals. *Acta Materialia*, 55(14):4867–4876, 2007.

[74] Y. Huang, S. Qu, K.C. Hwang, M. Li, and H. Gao. A conventional theory of mechanism-based strain gradient plasticity. *International Journal of Plasticity*, 20(4):753–782, 2004.

[75] J.W. Hutchinson. Plasticity at the micron scale. *International Journal of Solids and Structures*, 37(1):225–238, 2000.

[76] A. Ishida and M. Sato. Thickness effect on shape memory behavior of ti-50.0at.%ni thin film. *Acta Materialia*, 51(18):5571 – 5578, 2003.

[77] H. Jeffreys. On isotropic tensors. In *Mathematical Proceedings of the Cambridge Philosophical Society*, volume 73, pages 173–176. Cambridge University Press, 1973.

[78] H. Jiang, Y. Huang, Z. Zhuang, and K.C. Hwang. Fracture in mechanism-based strain gradient plasticity. *Journal of the Mechanics and Physics of Solids*, 49(5):979–993, 2001.

[79] M. Jirasek. Nonlocal theories in continuum mechanics. *Acta Polytechnica*, 44(5-6), 2004.

[80] M. Jirasek and Z.P. Bazant. Models for localization of softening and size effect. In M. Jirasek and Z.P. Bazant, editors, *Inelastic Analysis of Structures*, chapter 26, pages 515–537. John Wiley & Sons, 2002.

[81] J. San Juan and M.L. N. Superelasticity and shape memory at nano-scale: Size effects on the martensitic transformation. *Journal of Alloys and Compounds*, 577, Supplement 1:S25–S29, 2013.

[82] H. Kahn, M.A. Huff, and A.H. Heuer. The tini shape-memory alloy and its applications for mems. *Journal of Micromechanics and Microengineering*, 8(3):213, 1998.

[83] S. Kakunai, J. Masaki, R. Kuroda, K. Iwata, and R. Nagata. Measurement of apparent young's modulus in the bending of cantilever beam by heterodyne holographic interferometry. *Experimental Mechanics*, 25(4):408–412, 1985.

[84] E.A. Kearsley and J.T. Fong. Linearly independent sets of isotropic cartesian tensors of ranks up to eight. *Journal of Research of the National Bureau of Standards, Section B: Mathematical Sciences*, 79:49–58, 1975.

[85] Y.H. Kim, G.B. Cho, S.G. Hur, S.S. Jeong, and T.H. Nam. Nanocrystallization of a ti50.0ni(at.%) alloy by cold working and stress/strain behavior. *Materials Science and Engineering: A*, 438-440:531–535, 2006. Proceedings of the International Conference on Martensitic Transformations.

[86] B. Kockar, I. Karaman, J.I. Kim, Y.I. Chumlyakov, J. Sharp, and C.J. (Mike) Yu. Thermomechanical cyclic response of an ultrafine-grained niti shape memory alloy. *Acta Materialia*, 56(14):3630 – 3646, 2008.

[87] T. Kuninori, E. Sukedai, and H. Hashimoto. Martensitic transformation in thin foil specimen of a shape memory tini alloy. *Materials Transactions, JIM*,

37(7):1404–1407, jul 1996.

[88] D.C. Lagoudas, editor. *Shape Memory Alloys: Modeling and Engineering Applications*. Springer-Verlag, New York, 2008.

[89] D.C. Lagoudas, Z. Bo, and M.A. Qidwai. A unified thermodynamic constitutive model for sma and finite element analysis of active metal matrix composites. *Mechanics of Composite Materials and Structures*, 3(2):153–179, 1996.

[90] D.C. Lagoudas, D. Hartl, Y. Chemisky, L. Machado, and P. Popov. Constitutive model for the numerical analysis of phase transformation in polycrystalline shape memory alloys. *International Journal of Plasticity*, 32:155–183, 2012.

[91] W.M. Lai, D.H. Rubin, D. Rubin, and E. Krempl. *Introduction to continuum mechanics*. Butterworth-Heinemann, 2009.

[92] R.S. Lakes. Experimental microelasticity of two porous solids. *International Journal of Solids and Structures*, 22(1):55 – 63, 1986.

[93] M. Lazar and G.A. Maugin. Nonsingular stress and strain fields of dislocations and disclinations in first strain gradient elasticity. *International Journal of Engineering Science*, 43(1314):1157 – 1184, 2005.

[94] S.P. Lele and L. Anand. A small-deformation strain-gradient theory for isotropic viscoplastic materials. *Philosophical Magazine*, 88(30-32):3655–3689, 2008.

[95] D.J. Lloyd. Particle reinforced aluminium and magnesium matrix composites. *International Materials Reviews*, 39(1):1–23, 1994.

[96] J. Lubliner. *Plasticity theory*. Courier Dover Publications, 2008.

[97] L.E. Malvern. Introduction to continuum mechanics. *Prentice Hall Inc., Engle Cliffs, NJ*, 1696:301–320, 1969.

[98] G.A. Malygin. Nanoscopic size effects on martensitic transformations in shape memory alloys. *Physics of the Solid State*, 50(8):1538–1543, 2008.

[99] R.M. Manjeri, S. Qiu, N. Mara, A. Misra, and R. Vaidyanathan. Superelastic response of [111] and [101] oriented niti micropillars. *Journal of Applied Physics*, 108(2), 2010.

[100] G.A. Maugin. *The thermomechanics of nonlinear irreversible behaviors*. World Scientific, 1999.

[101] K.W. McElhaney, J.J. Vlassak, and W.D. Nix. Determination of indenter tip geometry and indentation contact area for depth-sensing indentation experiments. *Journal of Materials Research*, 13:1300–1306, 5 1998.

[102] R.D. Mindlin. Micro-structure in linear elasticity. *Archive for Rational Mechanics and Analysis*, 16(1):51–78, 1964.

[103] R.D. Mindlin. Second gradient of strain and surface-tension in linear elasticity. *International Journal of Solids and Structures*, 1(4):417–438, 1965.

[104] S. Miyazaki, T. Kawai, and K. Otsuka. On the origin of intergranular fracture in β phase shape memory alloys. *Scripta Metallurgica*, 16(4):431–436, 1982.

[105] J.L.M. Morrison. The yield of mild steel with particular reference to the effect of size of specimen. *Proceedings of the Institution of Mechanical Engineers*, 142(1):193–223, 1939.

[106] H.B. Mühlhaus and E.C. Aifantis. A variational principle for gradient plasticity. *International Journal of Solids and Structures*, 28(7):845–857, 1991.

[107] C.F. Niordson and J.W. Hutchinson. On lower order strain gradient plasticity theories. *European Journal of Mechanics-A/Solids*, 22(6):771–778, 2003.

[108] W.D. Nix and H. Gao. Indentation size effects in crystalline materials: A law for strain gradient plasticity. *Journal of the Mechanics and Physics of Solids*, 46(3):411 – 425, 1998.

[109] W. Noll. A mathematical theory of the mechanical behavior of continuous media. *Archive for Rational Mechanics and Analysis*, 2(1):197–226, 1958.

[110] D.M. Norfleet, P.M. Sarosi, S. Manchiraju, M.F.X. Wagner, M.D. Uchic, P.M. Anderson, and M.J. Mills. Transformation-induced plasticity during pseudoelastic deformation in niti microcrystals. *Acta Materialia*, 57(12):3549 – 3561, 2009.

[111] K. Otsuka and X. Ren. Physical metallurgy of tini-based shape memory alloys. *Progress in Materials Science*, 50(5):511 – 678, 2005.

[112] N. Ozdemir, I. Karaman, N.A. Mara, Y.I. Chumlyakov, and H.E. Karaca. Size effects in the superelastic response of $Ni_{54}Fe_{19}Ga_{27}$ shape memory alloy pillars with a two stage martensitic transformation. *Acta Materialia*, 60(16):5670, 2012.

[113] B. Peultier, T. Ben Zineb, and E. Patoor. Macroscopic constitutive law of shape memory alloy thermomechanical behaviour. application to structure computation by fem. *Mechanics of Materials*, 38(5):510–524, 2006.

[114] Francis R. Phillips, Dong Fang, Hongxing Zheng, and Dimitris C. Lagoudas. Phase transformation in free-standing sma nanowires. *Acta Materialia*, 59(5):1871–1880, 2011.

[115] C. Polizzotto. Nonlocal elasticity and related variational principles. *International Journal of Solids and Structures*, 38(42):7359–7380, 2001.

[116] C. Polizzotto. Unified thermodynamic framework for nonlocal gradient continuum theories. *European Journal of Mechanics-A: Solids*, 22(5):651–668, 2003.

[117] C. Polizzotto and G. Borino. A thermodynamics-based formulation of gradient-dependent plasticity. *European Journal of Mechanics-A: Solids*, 17(5):741–761, 1998.

[118] X. Qian, S. Zhang, S. Swaddiwudhipong, and L. Shen. Temperature dependence of material length scale for strain gradient plasticity and its effect on near-tip opening displacement. *Fatigue and Fracture of Engineering Materials and Structures*, 37(2):157–170, 2014.

[119] L. Qiao, J.J. Rimoli, Y. Chen, C.A. Schuh, and R. Radovitzky. Nonlocal superelastic model of size-dependent hardening and dissipation in single crystal cu-al-ni shape memory alloys. *Physical Review Letters*, 106(8):085504, 2011.

[120] X. Qiu, Y. Huang, Y. Wei, H. Gao, and K.C. Hwang. The flow theory of mechanism-based strain gradient plasticity. *Mechanics of Materials*, 35(3):245–258, 2003.

[121] S. Qu, Y. Huang, G.M. Pharr, and K.C. Hwang. The indentation size effect in the spherical indentation of iridium: A study via the conventional theory of mechanism-based strain gradient plasticity. *International Journal of Plasticity*, 22(7):1265–1286, 2006.

[122] S. Qu, T. Siegmund, Y. Huang, P.D. Wu, F. Zhang, and K.C. Hwang. A study of particle size effect and interface fracture in aluminum alloy composite via an extended conventional theory of mechanism-based strain-gradient plasticity. *Composites Science and Technology*, 65(7):1244–1253, 2005.

[123] R.S. Rivlin. On the principles of equipresence and unification. In *Collected Papers of RS Rivlin*, pages 1425–1426. Springer, 1997.

[124] D. Rogula. Introduction to nonlocal theory of material media. In *Nonlocal Theory of Material Media*, pages 123–222. Springer, 1982.

[125] A.L. Roytburd, T.S. Kim, Q. Su, J. Slutsker, and M. Wuttig. Martensitic transformation in constrained films. *Acta Materialia*, 46(14):5095 – 5107, 1998.

[126] J. San Juan, M.L. N, and C.A. Schuh. Superelasticity and shape memory in micro and nanometer-scale pillars. *Advanced Materials*, 20(2):272–278, 2008.

[127] J. San Juan, M. Nó, and C.A. Schuh. Nanoscale shape-memory alloys for ultrahigh mechanical damping. *Nature Nanotechnology*, 4(7):415–419, 2009.

[128] K. Santaoja. Gradient theory from the thermomechanics point of view. *Engineering Fracture Mechanics*, 71(4):557–566, 2004.

[129] D. Schryvers, W. Tirry, and Z.Q. Yang. Measuring strain fields and concentration gradients around Ni_4Ti_3 precipitates. *Materials Science and Engineering: A*, 438:485–488, 2006.

[130] J.A. Shaw and S. Kyriakides. Thermomechanical aspects of niti. *Journal of the Mechanics and Physics of Solids*, 43(8):1243 – 1281, 1995.

[131] M.X. Shi, Y. Huang, and K.C. Hwang. Plastic flow localization in mechanism-based strain gradient plasticity. *International Journal of Mechanical Sciences*, 42(11):2115–2131, 2000.

[132] Z. Shi, Y. Huang, J. Song, K.C. Hwang, and M. Li. Study of plastic shear localization via the flow theory of mechanism-based strain gradient plasticity. *Journal of Engineering Mechanics*, 135(3):132–138, 2009.

[133] D.D. Shin, D.G. Lee, K.P. Mohanchandra, and G.P. Carman. Thin film niti microthermostat array. *Sensors and Actuators A: Physical*, 130:37–41, 2006.

[134] P. Shrotriya, S.M. Allameh, J. Lou, T. Buchheit, and W.O. Soboyejo. On the measurement of the plasticity length scale parameter in liga nickel foils. *Mechanics of Materials*, 35(3):233–243, 2003.

[135] Y.C. Shu. Theory of sma thin films for microactuators and micropumps. In S Miyazaki, Y Q Fu, and W M Huang, editors, *Thin Film Shape Memory Alloys: Fundamentals and Device Applications*, chapter 11, pages 275–299. Cambridge University Press, 2009.

[136] A.J.M. Spencer. Isotropic polynomial invariants and tensor functions. In J.P. Boehler, editor, *Applications of Tensor Functions in Solid Mechanics*, volume 292 of *International Centre for Mechanical Sciences*, pages 141–169. Springer Vienna, 1987.

[137] J.S. Stölken and A.G. Evans. A microbend test method for measuring the plasticity length scale. *Acta Materialia*, 46(14):5109–5115, 1998.

[138] Q.P. Sun and Y.J. He. A multiscale continuum model of the grain-size dependence of the stress hysteresis in shape memory alloy polycrystals. *International Journal of Solids and Structures*, 45(13):3868–3896, 2008.

[139] Y. Sutou, N. Kamiya, T. Omori, R. Kainuma, K. Ishida, and K. Oikawa. Stress-strain characteristics in ni–ga–fe ferromagnetic shape memory alloys. *Applied Physics Letters*, 84(8):1275–1277, 2004.

[140] M. Tabesh, J.G. Boyd, and D.C. Lagoudas. Modeling size effect in the sma response: a gradient theory. volume 9058, pages 905803–905803–11, 2014.

[141] M. Tabesh, B. Lester, D. Hartl, and D.C. Lagoudas. Influence of the latent heat of transformation and thermomechanical coupling on the performance of shape memory alloy actuators. In *ASME 2012 Conference on Smart Materials, Adaptive Structures and Intelligent Systems*, pages 237–248. American Society of Mechanical Engineers, 2012.

[142] R.A. Toupin. Elastic materials with couple-stresses. *Archive for Rational Mechanics and Analysis*, 11(1):385–414, 1962.

[143] I. Tsagrakis, A. Konstantinidis, and E.C. Aifantis. Strain gradient and wavelet interpretation of size effects in yield and strength. *Mechanics of Materials*, 35(8):733–745, 2003.

[144] N.I. Tymiak, D.E. Kramer, D.F. Bahr, T.J. Wyrobek, and W.W. Gerberich. Plastic strain and strain gradients at very small indentation depths. *Acta Materialia*, 49(6):1021 – 1034, 2001.

[145] P.A. Vermeer and R.B.J. Brinkgreve. A new effective non-local strain measure for softening plasticity. *Localisation and Bifurcation Theory for Soils and Rocks*, pages 89–100, 1994.

[146] T. Waitz, T. Antretter, F.D. Fischer, N.K. Simha, and H.P. Karnthaler. Size effects on the martensitic phase transformation of niti nanograins. *Journal of the Mechanics and Physics of Solids*, 55(2):419 – 444, 2007.

[147] T. Waitz, V. Kazykhanov, and H.P. Karnthaler. Martensitic phase transformations in nanocrystalline niti studied by tem. *Acta Materialia*, 52(1):137 – 147, 2004.

[148] R.F. Wallis. *Lattice Dynamics: Proceedings of the International Conference Held at Copenhagen, Denmark, August 5–9, 1963*. Elsevier Science, 2013.

[149] D. Wan and K. Komvopoulos. Thickness effect on thermally induced phase transformations in sputtered titanium-nickel shape-memory films. *Journal of Materials Research*, 20:1606–1612, 6 2005.

[150] F. Wang, P. Huang, W. Chen, and K. Xu. Size effects of superelasticity in nanocrystalline niti shape memory alloy. In *Nanoelectronics Conference (INEC), 2010 3rd International*, pages 989–990, Jan 2010.

[151] B. Wu, A. Heidelberg, and J.J. Boland. Mechanical properties of ultrahigh-strength gold nanowires. *Nature Materials*, 4:525–529, 2005.

[152] M. Wuttig, R. Franchy, and H. Ibach. The rayleigh phonon dispersion curve on cu(100) in the x direction. *Solid State Communications*, 57(6):445 – 447, 1986.

[153] Z.C. Xia and J.W. Hutchinson. Crack tip fields in strain gradient plasticity. *Journal of the Mechanics and Physics of Solids*, 44(10):1621–1648, 1996.

[154] F. Yang. Size-dependent effective modulus of elastic composite materials: Spherical nanocavities at dilute concentrations. *Journal of Applied Physics*, 95(7):3516–3520, 2004.

[155] H.S. Yang and H.K.D.H. Bhadeshia. Austenite grain size and the martensite-start temperature. *Scripta Materialia*, 60(7):493 – 495, 2009.

[156] J.L. Yarnell, J.L. Warren, R.G. Wenzel, and S.H. Koenig. Phonon dispersion curves in bismuth. *IBM Journal of Research and Development*, 8(3):234–240, 1964.

[157] J. Ye, R.K. Mishra, A.R. Pelton, and A.M. Minor. Direct observation of the niti martensitic phase transformation in nanoscale volumes. *Acta Materialia*, 58(2):490 – 498, 2010.

[158] G. Yun, K.C. Hwang, Y. Huang, and P.D. Wu. A reformulation of mechanism-based strain gradient plasticity. *Philosophical Magazine*, 85(33-35):4011–4029, 2005.

[159] H.M. Zbib and E.C. Aifantis. A gradient-dependent flow theory of plasticity: application to metal and soil instabilities. *Applied Mechanics Reviews*, 42(11S):S295–S304, 1989.

[160] Y. Zhong, K. Gall, and T. Zhu. Atomistic characterization of pseudoelasticity and shape memory in niti nanopillars. *Acta Materialia*, 60(18):6301 – 6311, 2012.

APPENDIX A

ON SOME PROPERTIES OF THE DISSIPATION POTENTIAL

A.1 Principle of maximum dissipation and Onsager reciprocal relations

In the theory of irreversible thermodynamics it is assumed that in the neighborhood of equilibrium, the constitutive equations are given by:

$$J_i = \mathcal{L}_{ij}\,\chi_j \quad , \quad \chi_i = \mathcal{A}_{ij}\,J_j \tag{A.1}$$

in which J is the vector of generalized thermodynamic fluxes and χ the vector of generalized thermodynamic forces. The dissipation due to the inelastic process considered is:

$$D \equiv \boldsymbol{\chi} \cdot \boldsymbol{J} = \chi_m J_m \tag{A.2}$$

The principle of maximum dissipation states that from all admissible fluxes J (that may be constrained by internal constrains, boundary conditions, or conservation laws), the ones hold that maximize the dissipation D according to the equivalence relation $D(\boldsymbol{J}, \boldsymbol{\chi}) \equiv \boldsymbol{\chi} \cdot \boldsymbol{J}$ [72].

$$\boldsymbol{\chi} = \mathrm{Max}\{D \mid \boldsymbol{\chi}\,; \quad D = \boldsymbol{\chi} \cdot \boldsymbol{J}\} \tag{A.3}$$

Using the method of Lagrange multipliers, a Lagrangian is formed.

$$\boldsymbol{L}_\chi = D + \lambda\,(D - \boldsymbol{\chi} \cdot \boldsymbol{J})\,, \quad \frac{\partial \boldsymbol{L}_\chi}{\partial \boldsymbol{\chi}} = 0 \quad \Rightarrow \quad (1+\lambda)\,\frac{\partial D}{\partial \boldsymbol{\chi}} - \lambda = 0\boldsymbol{J} \tag{A.4}$$

Multiplying (inner product) both sides of equation (A.4) by $\boldsymbol{\chi}$ we will have:

$$(1+\lambda)\,\frac{\partial D}{\partial \boldsymbol{\chi}}\cdot\boldsymbol{\chi}=\lambda\boldsymbol{\chi}\cdot\boldsymbol{J}=\lambda D \quad \Rightarrow \quad \frac{1+\lambda}{\lambda}=\frac{D}{\frac{\partial D}{\partial \boldsymbol{\chi}}\cdot\boldsymbol{\chi}} \tag{A.5}$$

Furthermore, rearranging (A.4) using equation (A.5) and $D=\boldsymbol{\chi}\cdot\boldsymbol{\mathcal{L}}\boldsymbol{\chi}$, it is possible to obtain:

$$\boldsymbol{J}=\frac{\boldsymbol{\chi}\cdot\boldsymbol{\mathcal{L}}\boldsymbol{\chi}}{\boldsymbol{\chi}\cdot\left(\boldsymbol{\mathcal{L}}+\boldsymbol{\mathcal{L}}^{T}\right)\boldsymbol{\chi}}\left(\boldsymbol{\mathcal{L}}+\boldsymbol{\mathcal{L}}^{T}\right)\boldsymbol{\chi} \tag{A.6}$$

On the other hand, $\boldsymbol{J}=\boldsymbol{\mathcal{L}}\boldsymbol{\chi}$ after (A.1). Hence, both equations (A.1) and (A.6) are satisfied iff $\boldsymbol{\mathcal{L}}=\frac{\boldsymbol{\mathcal{L}}+\boldsymbol{\mathcal{L}}^{T}}{2}$; that is the tensor $\boldsymbol{\mathcal{L}}$ must be symmetric.

A.2 State of coupling in the gradient-based rate-independent constitutive modeling with internal variables

The method used in this dissertation for the constitutive modeling of inelastic material response using a gradient based internal variable approach includes assuming the existence of one yield/transformation function that acts as a potential to obtain the rates of internal variables. Assume the internal variable α and its spatial gradient $\nabla\boldsymbol{\alpha}$ with respective conjugate thermodynamic forces γ_1 and $\boldsymbol{\gamma}_2$. The dissipation, therefore can be written as

$$\gamma_1=-\rho\frac{\partial G}{\partial \alpha} \quad , \quad \boldsymbol{\gamma}_2=-\rho\frac{\partial G}{\partial \nabla\boldsymbol{\alpha}} \quad , \quad D=\alpha\,\gamma_1+\nabla\boldsymbol{\alpha}\cdot\boldsymbol{\gamma}_2\geq 0 \tag{A.7}$$

It is possible to relate the rates $\dot{\alpha}$ and $\nabla\dot{\boldsymbol{\alpha}}$ to one or two yield/transformation surfaces after invoking the principle of maximum dissipation. Depending on the existence of coupling between internal variables in the free energy G, three different cases may follow.

1. Explicit Coupling

In this case, one yield/transformation surface $\Phi(\gamma_1, \boldsymbol{\gamma}_2)$ is assumed. Also, it is assumed that a coupling between α and $\nabla\boldsymbol{\alpha}$ exists in the free energy function, G. The rates of internal variables, therefore, can be given by

$$\begin{aligned} \dot{\alpha} &= \lambda \frac{\partial \Phi(\gamma_1, \boldsymbol{\gamma}_2)}{\partial \gamma_1} \quad , \quad \gamma_1 = \hat{\gamma}_1(\alpha, \nabla\boldsymbol{\alpha}) \\ \nabla\dot{\boldsymbol{\alpha}} &= \lambda \frac{\partial \Phi(\gamma_1, \boldsymbol{\gamma}_2)}{\partial \boldsymbol{\gamma}_2} \quad , \quad \boldsymbol{\gamma}_2 = \hat{\boldsymbol{\gamma}}_2(\alpha, \nabla\boldsymbol{\alpha}) \end{aligned} \tag{A.8}$$

2. Implicit Coupling

In this case two surfaces $\Phi_1(\gamma_1)$ and $\Phi_2(\boldsymbol{\gamma}_2)$ are defined with two respective Lagrange multipliers. A coupling is assumed to exist between α and $\nabla\boldsymbol{\alpha}$ in the free energy function, G. Therefore

$$\begin{aligned} \dot{\alpha} &= \lambda_1 \frac{\partial \Phi_1(\gamma_1)}{\partial \gamma_1} \quad , \quad \gamma_1 = \hat{\gamma}_1(\alpha, \nabla\boldsymbol{\alpha}) \\ \nabla\dot{\boldsymbol{\alpha}} &= \lambda_2 \frac{\partial \Phi_2(\boldsymbol{\gamma}_2)}{\partial \boldsymbol{\gamma}_2} \quad , \quad \boldsymbol{\gamma}_2 = \hat{\boldsymbol{\gamma}}_2(\alpha, \nabla\boldsymbol{\alpha}) \end{aligned} \tag{A.9}$$

3. No Coupling

Here it is assumed that that two surfaces exist. Also, no coupling between α and $\nabla\boldsymbol{\alpha}$ is considered in the free energy G. This results in

$$\begin{aligned} \dot{\alpha} &= \lambda_1 \frac{\partial \Phi_1(\gamma_1)}{\partial \gamma_1} \quad , \quad \gamma_1 = \hat{\gamma}_1(\alpha) \\ \nabla\dot{\boldsymbol{\alpha}} &= \lambda_2 \frac{\partial \Phi_2(\boldsymbol{\gamma}_2)}{\partial \boldsymbol{\gamma}_2} \quad , \quad \boldsymbol{\gamma}_2 = \hat{\boldsymbol{\gamma}}_2(\nabla\boldsymbol{\alpha}) \end{aligned} \tag{A.10}$$

which shows that α and its spatial gradient $\nabla\boldsymbol{\alpha}$ can be solved for completely separate from each other.

It is concluded that there is no need for including the gradients of internal variables, unless a coupling (either explicit or implicit) is assumed in the constitutive relations or evolution equations.

APPENDIX B

TENSOR ALGEBRA AND TENSOR CALCULUS

B.1 Isotropic tensors of rank up to 6

Tensors can, in general, be defined as multilinear transformations over a finite dimensional vector space. An orthonormal basis, $\{\boldsymbol{e}^1, \ldots, \boldsymbol{e}^n\}$, is considered for the inner product, n-dimensional vector space $\{\mathcal{V}, .\}$. The basis can be used for obtaining the components of tensors and vectors. For the purpose of this dissertation, $\mathcal{V}$ is assumed to be the three dimensional Euclidean vector space, $\mathbb{R}^3$. The space of zero-order, $\mathcal{T}^0$, and first-order tensors, $\mathcal{T}^1$, are isomorphic to the scalar field $\mathbb{R}$ and the underlying vector space $\mathbb{R}^3$, receptively. Similarly, tensor $\boldsymbol{A} \in \mathcal{T}^q$ of rank q can be defined as the multilinear transformation:

$$\boldsymbol{A}[\cdot]: \quad \mathcal{T}^s \to \mathcal{T}^r \quad , \quad q = r + s \tag{B.1}$$

The orthonormal basis $\{\boldsymbol{e}^1, \boldsymbol{e}^2, \boldsymbol{e}^3\}$ constructs a natural orthonormal basis for the space of q$^{\text{th}}$ rank tensors $\mathcal{T}^q$ and defines the components of $\boldsymbol{A} \in \mathcal{T}^q$ with respect to that basis.

$$\boldsymbol{A} = [A_{i,j,\ldots,q}] = \sum_{i,j,\ldots,q=1,2,3} A_{i,j,\cdots,q} \left(\boldsymbol{e}^i \otimes \boldsymbol{e}^j \otimes \cdots \otimes \boldsymbol{e}^q\right) \tag{B.2}$$

The components of the tensors depend on the basis and if a new basis is formed by rotation through an orthogonal transformation, the new components follow the transformation law. In other words, with respect to a new basis $\{\acute{\boldsymbol{e}}^1, \acute{\boldsymbol{e}}^2, \acute{\boldsymbol{e}}^3\}$ such that $\acute{\boldsymbol{e}}^i = \boldsymbol{Q}\,\boldsymbol{e}^i$, the new components $\acute{A}_{i,j,\ldots,q}$ are given as:

$$\acute{A}_{i,j,\ldots,q} = Q_{\hat{i}i}\, Q_{\hat{j}j} \cdots Q_{\hat{q}q}\, A_{\hat{i},\hat{j},\ldots,\hat{q}} \tag{B.3}$$

with $\boldsymbol{Q}$ being a proper orthogonal transformation $\boldsymbol{Q}\boldsymbol{Q}^T = \boldsymbol{Q}^T\boldsymbol{Q} = \boldsymbol{I}$ and $\boldsymbol{Q} \in \mathrm{Orth}^+$, $\det(\boldsymbol{Q}) = +1$.

An *isotropic* tensor is the one whose components are invariant under any proper orthogonal transformation.

$$\acute{A}_{i,j,\ldots,q} = A_{i,j,\ldots,q} \tag{B.4}$$

All scalars (tensors of rank zero) are isotropic and there is no isotropic vector (tensor of rank one). Every isotropic tensor with an even rank can be expressed as a linear combination of the products of Kronecker deltas, δ_{ij}. In that sense, isotropic second order tensors are given by $A_{ij} = \lambda\delta_{ij}$. Also, every isotropic tensor of odd rank can be expressed by linear combination of terms formed by products of Kronecker deltas and permutation tensors, ϵ_{ijk} [77, 84]. Kearsley and Fong [84] refer to such products of Kronecker deltas or Kronecker deltas and permutation tensors as fundamental isotropic Cartesian tensors (FICT). The number of dependent and independent FICT terms in tensors of up to rank 8 are shown in table B.1.

Table B.1: The number of dependent and independent FICT terms in tensors of up to rank 8 [84].

Rank q	2	3	4	5	6	7	8
Number of distinct FICTs	1	1	3	10	15	105	105
Number of linearly independent FICTs	1	1	3	6	15	36	91

The linear combination, or reduction equations, of FICT terms for tensors of

rank 2 to 6 is given in table B.2. This constitutes the forms for developing isotropic material constant tensors of various ranks. Such forms are also available in [84] for ranks 7 and 8.

Table B.2: The isotropic tensors up to rank 6 [84].

Rank	Linearly independent fundamental isotropic tensors
2	δ_{ij}
3	ϵ_{ijk}
4	$\delta_{ij}\delta_{kl}$, $\delta_{il}\delta_{jk}$, $\delta_{ik}\delta_{jl}$
5	$\epsilon_{ijk}\delta_{lm}$, $\epsilon_{ijm}\delta_{lk}$, $\epsilon_{ijl}\delta_{mk}$, $\epsilon_{ikm}\delta_{lj}$, $\epsilon_{ikl}\delta_{mj}$, $\epsilon_{iml}\delta_{jk}$
6	$\delta_{ij}\delta_{kl}\delta_{mn}$, $\delta_{ij}\delta_{km}\delta_{nl}$, $\delta_{ij}\delta_{kn}\delta_{lm}$, $\delta_{jk}\delta_{il}\delta_{mn}$, $\delta_{jk}\delta_{im}\delta_{nl}$ $\delta_{jk}\delta_{in}\delta_{lm}$, $\delta_{ik}\delta_{jl}\delta_{mn}$, $\delta_{ik}\delta_{jm}\delta_{nl}$, $\delta_{ik}\delta_{jn}\delta_{lm}$, $\delta_{il}\delta_{jm}\delta_{kn}$ $\delta_{jl}\delta_{km}\delta_{in}$, $\delta_{kl}\delta_{im}\delta_{jn}$, $\delta_{il}\delta_{jn}\delta_{km}$, $\delta_{jl}\delta_{kn}\delta_{im}$, $\delta_{kl}\delta_{in}\delta_{jm}$

The interesting point here is the existence of isotropic third and fifth rank tensors. In many cases of material constitutive modeling, the dependent or independent variables being related through tensorial material constants include symmetries (e.g. $\boldsymbol{\sigma} = \boldsymbol{\sigma}^T$) that must be imparted to the tensors for material constants. Since the isotropic odd-ranked tensors involve linear combinations of the permutation tensor, they vanish under the symmetry requirements. For example if the isotropic third order tensor $\boldsymbol{D} \in \mathcal{T}^3$ has to obey a symmetry such as $D_{ijk} = D_{ikj}$, we will have (according to table B.2):

$$D_{ijk} = \lambda\,\epsilon_{ijk} \quad , \quad \lambda\,\epsilon_{ijk} = -\lambda\,\epsilon_{ikj} = \lambda\,\epsilon_{ikj} \Rightarrow \lambda = 0 \tag{B.5}$$

Similar arguments can be developed for the fifth and seventh-rank tensors.

Therefore, material properties represented by odd-ranked tensors, such as the piezoelectric tensor, only exist in the case of anisotropic material response.

B.2 The derivative of tensor valued functions

Assume the tensor function

$$F \ : \mathcal{D} \subset \mathcal{T}^r \to \mathcal{T}^s \tag{B.6}$$

to have a domain $\mathcal{D}$ as an open subset of $\mathcal{T}^r$. The function takes values in the space of r-th rank to the space of s-th rank tensors. Also F is assumed to be continuous and differentiable in a neighborhood of $\boldsymbol{A} \in \mathcal{D}$.

The derivative of F operating on $\boldsymbol{H} \in \mathcal{T}^r$ is defined as:

$$F(\boldsymbol{A}+\boldsymbol{H}) = F(\boldsymbol{A}) + \mathrm{D}F(\boldsymbol{A})[\boldsymbol{H}] + o(\boldsymbol{H}) \tag{B.7}$$

as $\|\boldsymbol{H}\| \to 0$. The Landau, little o, denotes a function depending on $\|\boldsymbol{H}\|$, the norm of $\boldsymbol{H}$, that goes to zero faster than $\boldsymbol{H}$; i.e. $\lim_{\|\boldsymbol{H}\|\to 0} \frac{o(\boldsymbol{H})}{\boldsymbol{H}} = 0$.

The derivative, in general, can also be obtained through the following.

$$\mathrm{D}F(\boldsymbol{A})[\boldsymbol{H}] = \frac{\mathrm{d}}{\mathrm{d}\alpha} F(\boldsymbol{A}+\alpha\boldsymbol{H})\,|_{\alpha=0} = \partial_{\boldsymbol{A}} F(\boldsymbol{A}) \cdot \boldsymbol{H} \qquad \forall \boldsymbol{H} \in \mathcal{T}^r \tag{B.8}$$

The "$\cdot$" here represents a general tensorial contraction operation. If $\boldsymbol{A}$ is symmetric, the derivative must be symmetric too and thus we will have:

$$\partial_{\boldsymbol{A}} F(\boldsymbol{A}) = \mathrm{Sym}\left[\partial_{\boldsymbol{A}} F(\boldsymbol{A})\right] \tag{B.9}$$

Given F and G as two differentiable tensor functions defined on the same domain $\mathcal{D} \subset \mathcal{T}^r$,

$$F \ : \mathcal{D} \subset \mathcal{T}^r \to \mathcal{T}^p \quad , \quad G \ : \mathcal{D} \subset \mathcal{T}^r \to \mathcal{T}^q \tag{B.10}$$

It is possible to define a product function E with the assumption that it is separately linear in F and G

$$E(\boldsymbol{A}) := \hat{E}\left(F(\boldsymbol{A}), G(\boldsymbol{A})\right) \ : \mathcal{T}^p \times \mathcal{T}^q \to \mathcal{T}^s \tag{B.11}$$

The derivative of the product function E can be obtained through the *product rule* as below:

$$\mathrm{D}E(\boldsymbol{A})[\boldsymbol{H}] = \hat{E}\left(\mathrm{D}F(\boldsymbol{A})[\boldsymbol{H}], G(\boldsymbol{A})\right) + \hat{E}\left(F(\boldsymbol{A}), \mathrm{D}G(\boldsymbol{A})[\boldsymbol{H}]\right) \quad , \quad \forall\, \boldsymbol{H} \in \mathcal{T}^r \tag{B.12}$$

Moreover, assume F to be differentiable at $\boldsymbol{A} \in \mathcal{D}$ and G to be differentiable at $F(A) \in \mathcal{G}$, i.e.

$$F \ : \mathcal{D} \subset \mathcal{T}^r \to \mathcal{T}^q \quad , \quad G \ : \mathcal{G} \subset \mathcal{T}^q \to \mathcal{T}^s \tag{B.13}$$

The composite function, $E(\cdot) := G \circ F(\cdot)$, is also differentiable at $\boldsymbol{A} \in \mathcal{D}$ through the *chain rule*.

$$\mathrm{D}E(\boldsymbol{A})[\boldsymbol{H}] = \mathrm{D}G(F(\boldsymbol{A}))\left[\mathrm{D}F(\boldsymbol{A})[\boldsymbol{H}]\right] \quad , \quad \forall\, \boldsymbol{H} \in \mathcal{T}^r \tag{B.14}$$

For further development, real scalar valued functions are considered with $\boldsymbol{A} \in \mathcal{D}$ being a tensor of arbitrary order.

$$F\ : \mathcal{D} \subset \mathcal{T}^r \to \mathbb{R} \tag{B.15}$$

B.2.1 Functions of the form: $F(\boldsymbol{v}) = \boldsymbol{v} \cdot \boldsymbol{A}\boldsymbol{v}$

Where $\boldsymbol{v} \in \mathcal{T}^1$ is a vector and $\boldsymbol{A} \in \mathcal{T}^2$ is a second-order tensor. Based on the previous definition in equation (B.8):

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\alpha} F(\boldsymbol{v} + \alpha\boldsymbol{h})\,|_{\alpha=0} = \frac{\mathrm{d}}{\mathrm{d}\alpha} (\boldsymbol{v} + \alpha\boldsymbol{h}) \cdot \boldsymbol{A}(\boldsymbol{v} + \alpha\boldsymbol{h})\,|_{\alpha=0} = \\ \frac{\mathrm{d}}{\mathrm{d}\alpha} \left(\boldsymbol{v} \cdot \boldsymbol{A}\boldsymbol{v} + \alpha\boldsymbol{v} \cdot \boldsymbol{A}\boldsymbol{h} + \alpha\boldsymbol{h} \cdot \boldsymbol{A}\boldsymbol{v} + \alpha^2\boldsymbol{h} \cdot \boldsymbol{h}\right)|_{\alpha=0} = \\ \left(\boldsymbol{A}^T\boldsymbol{v} + \boldsymbol{A}\boldsymbol{v}\right) \cdot \boldsymbol{h} \end{aligned} \tag{B.16}$$

Hence

$$\partial_{\boldsymbol{v}} F(\boldsymbol{v}) = \boldsymbol{A}^T\boldsymbol{v} + \boldsymbol{A}\boldsymbol{v} \tag{B.17}$$

And if $\boldsymbol{A}$ is symmetric, $\boldsymbol{A} \in \mathrm{Sym}(\mathcal{T}^2)$:

$$\partial_{\boldsymbol{v}} F(\boldsymbol{v}) = 2\boldsymbol{A}\boldsymbol{v} \tag{B.18}$$

B.2.2 Functions of the form: $F(\boldsymbol{v}) = \boldsymbol{v} \cdot \boldsymbol{A}\boldsymbol{u}$

Where $\boldsymbol{v},\ \boldsymbol{u} \in \mathcal{T}^1$ are vectors and $\boldsymbol{A} \in \mathcal{T}^2$ is a second-order tensor. The derivative can be readily determined.

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\alpha} F(\boldsymbol{v} + \alpha\boldsymbol{h})\,|_{\alpha=0} = \frac{\mathrm{d}}{\mathrm{d}\alpha} (\boldsymbol{v} + \alpha\boldsymbol{h}) \cdot \boldsymbol{A}\boldsymbol{u}|_{\alpha=0} = \\ \boldsymbol{A}\boldsymbol{u} \cdot \boldsymbol{h} \end{aligned} \tag{B.19}$$

Therefore

$$\partial_{\boldsymbol{v}} F(\boldsymbol{v}) = \boldsymbol{A}\boldsymbol{u} \tag{B.20}$$

B.2.3 Functions of the form: $F(\boldsymbol{A}) = \mathrm{tr}(\boldsymbol{A}^n \boldsymbol{B})$

Where $\boldsymbol{A},\ \boldsymbol{B} \in \mathcal{T}^2$ are second-order tensors and $n = 1, 2, \ldots$. The derivative is calculated using the previous definition.

$$\begin{gathered}
\frac{\mathrm{d}}{\mathrm{d}\alpha} F(\boldsymbol{A} + \alpha\boldsymbol{H})\,|_{\alpha=0} = \frac{\mathrm{d}}{\mathrm{d}\alpha}\mathrm{tr}\left[(\boldsymbol{A} + \alpha\boldsymbol{H})^n \boldsymbol{B}\right]|_{\alpha=0} = \\
\frac{\mathrm{d}}{\mathrm{d}\alpha}(\boldsymbol{A} + \alpha\boldsymbol{H})^n : \boldsymbol{B}^T|_{\alpha=0} = \frac{\mathrm{d}}{\mathrm{d}\alpha}(\boldsymbol{A} + \alpha\boldsymbol{H})^n\,|_{\alpha=0} : \boldsymbol{B}^T = \\
\frac{\mathrm{d}}{\mathrm{d}\alpha}\left[\boldsymbol{A}^n + \alpha\sum_{i=0}^{n-1}\boldsymbol{A}^i\boldsymbol{H}\boldsymbol{A}^{n-1-i} + \alpha^2\sum_{i=0}^{n-2}\boldsymbol{A}^i\boldsymbol{H}\boldsymbol{A}^{n-1-i} + \cdots\right]_{\alpha=0} : \boldsymbol{B}^T = \\
\left(\sum_{i=0}^{n-1}\boldsymbol{A}^i\boldsymbol{H}\boldsymbol{A}^{n-1-i}\right) : \boldsymbol{B}^T = \left[\sum_{i=0}^{n-1}\left(\boldsymbol{A}^{n-1-i}\boldsymbol{B}\boldsymbol{A}^i\right)^T\right] : \boldsymbol{H}
\end{gathered} \tag{B.21}$$

which means

$$\partial_{\boldsymbol{A}} F(\boldsymbol{A}) = \sum_{i=0}^{n-1}\left(\boldsymbol{A}^{n-1-i}\boldsymbol{B}\boldsymbol{A}^i\right)^T \tag{B.22}$$

If $\boldsymbol{A}$ is symmetric, $\boldsymbol{A} \in \mathrm{Sym}(\mathcal{T}^2)$:

$$\partial_{\boldsymbol{A}} F(\boldsymbol{A}) = \sum_{i=0}^{n-1}\boldsymbol{A}^i\,\mathrm{Sym}(\boldsymbol{B})\,\boldsymbol{A}^{n-1-i} \tag{B.23}$$

B.2.4 Functions of the form: $F(\boldsymbol{A}) = \boldsymbol{u} \cdot \boldsymbol{A}^n \boldsymbol{B}\boldsymbol{v}$

Where $\boldsymbol{v},\ \boldsymbol{u} \in \mathcal{T}^1$ are vectors and $\boldsymbol{A},\ \boldsymbol{B} \in \mathcal{T}^2$ are second-order tensors. F can be rewritten as:

$$F(\boldsymbol{A}) = \boldsymbol{u} \cdot \boldsymbol{A}^n \boldsymbol{B}\boldsymbol{v} = \boldsymbol{A}^n : (\boldsymbol{u} \otimes \boldsymbol{B}\boldsymbol{v}) = \mathrm{tr}\,(\boldsymbol{A}^n \boldsymbol{L}) \quad , \quad \boldsymbol{L} = (\boldsymbol{u} \otimes \boldsymbol{B}\boldsymbol{v})^T \tag{B.24}$$

The results of the last section B.2.3 is used to obtain the derivative of F.

$$\partial_{\boldsymbol{A}} F(\boldsymbol{A}) = \partial_{\boldsymbol{A}} \mathrm{tr}\,(\boldsymbol{A}^n \boldsymbol{L}) = \sum_{i=0}^{n-1} \left(\boldsymbol{A}^{n-1-i} (\boldsymbol{u} \otimes \boldsymbol{B}\boldsymbol{v}) \boldsymbol{A}^i\right)^T \tag{B.25}$$

And in the case of symmetric $\boldsymbol{A}$

$$\partial_{\boldsymbol{A}} F(\boldsymbol{A}) = \sum_{i=0}^{n-1} \boldsymbol{A}^i \,\mathrm{Sym}\,(\boldsymbol{u} \otimes \boldsymbol{B}\boldsymbol{v})^T \boldsymbol{A}^{n-1-i} \tag{B.26}$$

The derivations in this section are summarized in table B.3.

Table B.3: The derivatives of tensor valued functions commonly used in material constitutive model development*.

Function $F(\boldsymbol{v})$	Derivative $\partial_{\boldsymbol{v}} F(\boldsymbol{v})$
$\boldsymbol{v}\cdot\boldsymbol{A}\boldsymbol{v}$	$2\boldsymbol{A}\boldsymbol{v}$
$\boldsymbol{v}\cdot\boldsymbol{A}\boldsymbol{u}$	$\boldsymbol{A}\boldsymbol{u}$
Function $F(\boldsymbol{A})$	**Derivative $\partial_{\boldsymbol{A}} F(\boldsymbol{A})$**
$\mathrm{tr}(\boldsymbol{A}\boldsymbol{B}) = \boldsymbol{A}:\boldsymbol{B}^T$	$\mathrm{Sym}(\boldsymbol{B})$
$\mathrm{tr}(\boldsymbol{A}^2\boldsymbol{B}) = \boldsymbol{A}^2:\boldsymbol{B}^T$	$\mathrm{Sym}(\boldsymbol{B})\,\boldsymbol{A} + \boldsymbol{A}\,\mathrm{Sym}(\boldsymbol{B})$
$\mathrm{tr}(\boldsymbol{A}^3\boldsymbol{B}) = \boldsymbol{A}^3:\boldsymbol{B}^T$	$\mathrm{Sym}(\boldsymbol{B})\,\boldsymbol{A}^2 + \boldsymbol{A}\,\mathrm{Sym}(\boldsymbol{B})\,\boldsymbol{A} + \boldsymbol{A}^2\,\mathrm{Sym}(\boldsymbol{B})$
$\boldsymbol{u}\cdot\boldsymbol{A}\boldsymbol{B}\boldsymbol{v}$	$\mathrm{Sym}(\boldsymbol{u}\otimes\boldsymbol{B}\boldsymbol{v})^T$
$\boldsymbol{u}\cdot\boldsymbol{A}^2\boldsymbol{B}\boldsymbol{v}$	$\mathrm{Sym}(\boldsymbol{u}\otimes\boldsymbol{B}\boldsymbol{v})^T\,\boldsymbol{A} + \boldsymbol{A}\,\mathrm{Sym}(\boldsymbol{u}\otimes\boldsymbol{B}\boldsymbol{v})^T$

* $\boldsymbol{A}$ in this table is assumed to be a symmetric second order tensor $\boldsymbol{A}\in\mathrm{Sym}(\mathcal{T}^2)$, $\boldsymbol{B}\in\mathcal{T}^2$ a second order tensor, and $\boldsymbol{u},\ \boldsymbol{v}\in\mathcal{T}^1$ vectors.

B.3 Some useful formulas

The *gradient* of a tensor or tensor valued function $\boldsymbol{A}(\boldsymbol{X})$ can be described according to the previous definition of tensorial derivatives in Appendix B.2.

$$\mathrm{Grad}(\boldsymbol{A}(\boldsymbol{X})) = \nabla\boldsymbol{A}(\boldsymbol{X}) \equiv \partial_{\boldsymbol{X}}\boldsymbol{A}(\boldsymbol{X}) \quad , \quad \mathrm{D}\boldsymbol{A}(\boldsymbol{X})[\boldsymbol{h}] = \nabla\boldsymbol{A}(\boldsymbol{X})\cdot\boldsymbol{h} \quad \forall\boldsymbol{h}\in\mathbb{R}^3 \tag{B.27}$$

Based on this definition, the *divergence* can also be introduced.

$$\mathrm{Div}(\boldsymbol{A}(\boldsymbol{X}))\cdot\boldsymbol{H} = \mathrm{Div}(\boldsymbol{A}^*(\boldsymbol{X})\cdot\boldsymbol{H}) \equiv \mathrm{tr}(\nabla(\boldsymbol{A}^*(\boldsymbol{X})\cdot\boldsymbol{H})) \tag{B.28}$$

For all constant tensors $\boldsymbol{H}\in\mathcal{T}^{s-1}$ if $\boldsymbol{A}\in\mathcal{T}^s$.

It is worth mentioning that for the above relation, the adjoint of $\boldsymbol{A}$ denoted by $\boldsymbol{A}^*$ is one that satisfies the following for all non-zero vectors $\boldsymbol{v} \in \mathbb{R}^3$.

$$\boldsymbol{v} \cdot (\boldsymbol{A} \cdot \boldsymbol{H}) = \boldsymbol{H} \cdot (\boldsymbol{A}^* \cdot \boldsymbol{v}) \tag{B.29}$$

Also for tensor fields $\boldsymbol{A}(\boldsymbol{X}) \in \mathcal{T}^s$ and $\boldsymbol{B}(\boldsymbol{X}) \in \mathcal{T}^{s-1}$

$$\mathrm{Div}\,(\boldsymbol{A}^* \cdot \boldsymbol{B}) = \boldsymbol{A} \cdot \nabla \boldsymbol{B} + \boldsymbol{B} \cdot \mathrm{Div}\,(\boldsymbol{A}) \tag{B.30}$$

The following formulas can be deduced for a scalar-valued field $a(\boldsymbol{X})$, vector fields $\boldsymbol{u}(\boldsymbol{X})$ and $\boldsymbol{v}(\boldsymbol{X})$, and a second-order tensor field $\boldsymbol{A}(\boldsymbol{X})$.

$$\begin{aligned}
\mathrm{Div}\,(\boldsymbol{v}) &\equiv \mathrm{tr}\,(\nabla \boldsymbol{v}) \\
\nabla\,(a\boldsymbol{v}) &= a\nabla \boldsymbol{v} + \boldsymbol{v} \otimes \nabla a \\
\mathrm{Div}\,(a\boldsymbol{v}) &= a\mathrm{Div}\,(\boldsymbol{v}) + \boldsymbol{v} \cdot \nabla a \\
\nabla\,(\boldsymbol{u} \cdot \boldsymbol{v}) &= (\nabla \boldsymbol{v})^T \boldsymbol{u} + (\nabla \boldsymbol{u})^T \boldsymbol{v} \\
\mathrm{Div}\,(\boldsymbol{u} \otimes \boldsymbol{v}) &= \boldsymbol{u}\mathrm{Div}\,(\boldsymbol{v}) + (\nabla \boldsymbol{u})\,\boldsymbol{v} \\
\mathrm{Div}\left(\boldsymbol{A}^T \boldsymbol{v}\right) &= \boldsymbol{A} \cdot \nabla \boldsymbol{v} + \boldsymbol{v} \cdot \mathrm{Div}\,(\boldsymbol{A}) \\
\mathrm{Div}\,(a\boldsymbol{A}) &= a\mathrm{Div}\,(\boldsymbol{A}) + \boldsymbol{A}\nabla a \\
\nabla\,(\mathrm{Div}\,(\boldsymbol{v})) &= \mathrm{Div}\left((\nabla \boldsymbol{v})^T\right)
\end{aligned} \tag{B.31}$$