Experimental Techniques for Characterizing the Thermo-Electro-Mechanical Shakedown Response of SMA Wires and Tubes

by

Christopher B. Churchill

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Aerospace Engineering)
in The University of Michigan
2010

Doctoral Committee:

Associate Professor John A. Shaw, Chair
Professor Nicolas Triantafyllidis
Assistant Professor Samantha Daly
Nilesh Mankame, General Motors Corporation

Acknowledgements

I would like to thank all who had a hand in my work: my wife Celia and entire family for their emotional support, my colleagues in Ann Arbor, Ben Reedlunn and Petros Michailidis, for their consistent and helpful input; my partners at General Motors for their insight and funding. I appreciate John Shaw and Mark Iadicola for letting me present experimental data from their experiments on SMA wires and Jeff Tyber for arranging X-ray diffraction on my NiTi tubes. The UM technical staff: Tom, Dave, Terry, Eric, and Chris provided invaluable support, experience and advice. Finally, I am in debt to my advisor John Shaw for his guidance and patience.

Table of Contents

Acknowledgements . . . . . **ii**

List of Figures . . . . . **vii**

List of Tables . . . . . **xix**

Chapter

I. Introduction 1

1.1 Background 1
1.2 Motivation 7
1.3 Organization 11

II. Fundamental Characterization of SMA Wires Through Differential Scanning Calorimetry and Uniaxial Tension Experiments 15

2.1 Differential Scanning Calorimetry 16
2.1.1 Two DSC Experiments 17
2.1.2 Some Recommendations on DSC Measurements 25
2.2 Basic Thermomechanical Behavior 26
2.2.1 Shape Memory & Superelasticity 27
2.2.2 Two Thermomechanical Experiments 30
2.3 Uniaxial Experimental Methods 34
2.3.1 Loading Method 34
2.3.2 Strain Measurement & Control 37
2.3.3 Temperature Control & Measurement 42

2.4 Isothermal Mechanical Responses 46
2.4.1 Experimental Setup 47
2.4.2 Specimen Materials 50
2.4.3 Shape Memory Wire Responses 51
2.4.4 Superelastic Wire Responses 60
2.5 Material Instability and Transformation Fronts 61
2.6 Nucleation Experimental Setup 69
2.7 Transformation Fronts 73
2.8 A Nucleation Experiment 78
2.9 Quasi-Phase Diagram 85
2.10 Summary 88

III. Thermo-electro-mechanical Shakedown Response of Conditioned SMA Wires 91

3.1 Conditioned NiTi Wires 93
3.1.1 R-phase and Resistivity 97
3.1.2 Measuring Resistivity 99
3.2 Experimental Setup 104
3.3 Resistivity Experiments, 19 to 478 MPa 106
3.3.1 Experiment 1: 19 MPa (2 N Load) 107
3.3.2 Experiment 4: 287 MPa (30 N Load) 111
3.3.3 Summary of Resistivity Experiments 112
3.4 Constant Load Shakedown Experiments 116
3.4.1 Example shakedown experiment: 379 MPa (40 N) . . . 116
3.4.2 Strain Shakedown Responses 122
3.4.3 Resistivity Shakedown Responses 127
3.5 Shakedown Experiments Under Alternating Constant Loads 134
3.5.1 Counter-clockwise Experiment 134
3.5.2 Clockwise experiment 137
3.6 Strain Evolution With Dead Load Cycles 142
3.7 Conclusions 145

IV. A Reduced-Order Thermomechanical Model for an SMA Bias Spring Actuator 148

4.1 Thermomechanical Problem of an SMA Actuator 150
4.1.1 The Full 1-D Constitutive Model 152
4.1.2 Governing PDEs 155
4.1.3 Simplified Model 156
4.1.4 Simplified Governing Equations 159
4.1.5 Dimensionless Parameters & Equations 163
4.2 Analytical Solution 167
4.2.1 Mechanical Equilibrium Paths 167
4.2.2 Solution Without Phase Transformation 168
4.2.3 Solution During Phase Transformation 170
4.3 Numerical Example 175
4.3.1 Time-dependent Response 177
4.3.2 Comparison to Finite Element Simulation 178
4.3.3 Actuation times 182
4.4 Design Aspects 184
4.4.1 Actuation Stroke 185
4.4.2 Sizing of Springs 186
4.4.3 Actuation Times 190
4.4.4 Minimum Power Requirements 191
4.4.5 Energy Usage 193
4.4.6 Energy Efficiency 196
4.4.7 Potential Model Refinements 199
4.5 Summary & Conclusions 206

V. Localized Phenomena and Rate Effects in Tensile and Compressive NiTi Tubes 209

5.1 Background 209
5.2 Tension Experiments 214
5.2.1 Tension Experimental Setup 214

5.2.2 Infrared Image Post-Processing 218
5.2.3 Isothermal Characterization 219
5.2.4 Loading Rate Study 225
5.3 Compression Experiments 251
5.3.1 Compression Experimental Setup 251
5.3.2 Isothermal Characterization 252
5.3.3 Loading Rate Study 254
5.4 Tension-Compression Asymmetry 256
5.5 Bending Experiments 261
5.5.1 Bending Experimental Setup 262
5.5.2 Bending Isothermal Responses and Loading Rate Study 266

VI. Summary and Conclusions 269

6.1 Summary 269
6.2 Future Work 272
6.2.1 Thermo-electro-mechanical Shakedown 273
6.2.2 Reduced-order SMA actuator model 274
6.2.3 NiTi Tubes 275

Bibliography **277**

List of Figures

Figure

1.1 (a) The shape memory effect, (b) the superelastic effect 2

1.2 (a) The Frangibolt™ is fitted around a notched bolt, fracturing it when activated by heating. (b) The Cryofit™ coupler contracts when warmed to room temperature, joining fuel lines without dangerous welding or brazing 3

1.3 Medical devices which use NiTi: (a) Orthodontic archwire (b) compressive staples to repair bone fractures (c) arterial stent (d) eyeglases 5

1.4 History of the discovery of important shape memory alloys. [82] 6

1.5 Superelastic bending response of NiTi wires under load control. Data during transformation is sparse due to the choice of load rather than displacement control. 9

2.1 Schematics of crystal structures and microstructures in Nitinol. 18

2.2 Differential scanning calorimetry (DSC) thermograms of two Nitinol alloys: (a) shape memory wire [86], (b) superelastic wire [15]. 20

2.3 Photographs of (a) sample pan and lid, (b) wire specimens, and (c) crimped assembly for DSC. 26

2.4 Transmission electron micrographs of martensite in a NiTi alloy at residual strains: (a) 0 %, (b) 6.4 %, (c) 11.3 %, and (d) 24 %. Images taken from Nishda, *et. al.* [71] (reprinted with permission from Elsevier, http://www.sciencedirect.com/science/journal/13596462). 29

2.5 Thermomechanical experiment on shape memory Nitinol, showing shape memory effect (0 to 5) and superelasticity (5 to 10). Data taken from [84]. 33

2.6 Thermomechanical experiment on superelastic Nitinol: shape memory effect (0 to 5) and superelasticity (5 to 10). Data taken from [15]. 35

2.7 Setup for experiments on SM wire in a temperature controlled bath. . . . 48

2.8 Setup for experiments on SE wire in an air chamber. 50

2.9 Isothermal mechanical responses of SM wire at selected temperatures. Data taken from [84, 86]. 52

2.10 Comparision of liquid nitrogen-quenched material (thin line) and as-received material (thick line) responses of SM wire at intermediate temperatures. Data taken from [86]. 55

2.11 Example mechanical responses showing how characteristic plateau stresses and strains are extracted: (a) 10 °C experiment, (b) 70 °C experiment. . . 57

2.12 Characteristic transformation stresses of SM wire: initially quenched in liquid nitrogen (left), as-received (right). Respective heating and cooling DSC thermograms also shown along the stress-free axis for reference. . . 58

2.13 Characteristic transformation strains of SM wire (as-received). 60

2.14 Isothermal mechanical responses of as-received SE wire at selected temperatures. Data taken from [47] and [15]. 62

2.15 Characteristic transformation stresses of SE wire: initially quenched in liquid nitrogen (left), as-received (right). Respective heating and cooling DSC thermograms also shown along the stress-free axis for reference. . . 63

2.16 Characteristic transformation strains of SE wire (as-received). 63

2.17 Axial stress vs. local-strain responses (left) and measured local vs. global strain responses (right) in SM wire during two experiments: (a) 30 °C, (b) 70 °C. 65

2.18 Schematics of (a) local stress-strain response, and (b) global force-elongation ("structural") response with specimen configurations before, during, and after $A \rightarrow M^+$ transformation. 67

2.19 Temperature control apparatus for nucleation peak measurements. 70

2.20 Experiment at 24.4 °C at $\dot{\delta}/L = \pm 1 \times 10^{-3}$ s^{-1} using the setup of Figure 2.19 with passive thermoelectrics: (a) mechanical response, (b) streak plot from a sequence of optical images, and (c) IR measured temperature history. 77

2.21 Nucleation peak experiment at 24 °C at $\dot{\delta}/L = \pm 1 \times 10^{-4}$ s^{-1} using the setup of Figure 2.19, showing the onset of $A \rightarrow M^+$ during loading (a-b) and $M^+ \rightarrow A$ during unloading (c-d): (a) mechanical response, (b) IR measured temperature profiles during and just after nucleation events, (c) streak plot from a sequence of optical images, with lines overlaid for clarity. 79

2.22 Propagation stresses (σ_{P}, closed circles) and nucleation stresses (σ_{N}, open circles) during superelastic responses at several temperature. 87

3.1 90C Flexinol: (a) differential scanning calorimetry, (a) isothermal mechanical responses. 94

3.2 Transformation (plateau) stresses and DSC. 96

3.3 Constant load experiment at 19 MPa under temperature control, showing strain (ε_{e}), relative resistivity ($\bar{\rho}_{\mathrm{e}}$) and DSC output ($\dot{q}/|\dot{T}|$) vs. temperature (T). 100

3.4 Circuit diagram for four point measurement of wire resistance. The current supplied by the power supply I is calculated from the voltage drop V_1 across a calibrated shunt resistor, then used to determine the unknown gauge length resistance. 102

3.5 A custom fastener to measure resistance and strain from NiTi wires: (a) top view digram (b) photograph, as installed 103

3.6 Resistance response at 574 MPa of 30 mm long 90C Flexinol. 104

3.7 Schematic of experimental setup: (A) front view, (B) top view. 106

3.8 19 MPa dead load experiment showing: (a) temperature (T), strain (ε_{e}), and relative resistivity ($\bar{\rho}_{\mathrm{e}}$) histories during a single temperature cycle; (b) strain response; (c) relative resistivity response; (d) resistivity vs. strain . 108

3.9 Constant Load Experiment 4 (287 MPa dead load) showing: (a) temperature (T), strain (ε_{e}), and relative resistivity ($\bar{\rho}_{\mathrm{e}}$) histories during a single thermal cycle, (b) strain response; (c) relative resistivity response; (d) resistivity vs. strain 110

3.10 Idealized construction lines overlay responses from the 287 MPa dead load experiment in Figure 3.9. (a) The idealized ε_{e} - T response is flat for "pure" M^+ and A. (b) The idealized $\bar{\rho}_{\mathrm{e}}$ - T response at 287 MPa is linear and diverging for the pure phases. (a) This causes a non-proportional relationship between ε_{e} and $\bar{\rho}_{\mathrm{e}}$ outside the transformation region. 113

3.11 Summary of constant load experiment at six loads from 19 to 478 MPa, showing: (a) strain (ε_{e}) vs. temperature (T), (b) relative resistivity vs. temperature, and (c) relative resistivity vs. strain 114

3.12 Idealized T - $\bar{\rho}_{\mathrm{e}}$ transformation paths for low load thermal cycles (left dashed box) and high load thermal cycles (right dashed box) T - $\bar{\rho}_{\mathrm{e}}$ experimental data from 478 and 19 MPa experiments are in the background. 115

3.13 Initial heating and first cycle of shakedown experiment at 379 MPa, with histories of temperature T, stress P/A_0, strain ε_{e}, and relative resistivity $\bar{\rho}_{\mathrm{e}}$.117

3.14 Mechanical response of the experiment in Figure 3.13 ($532 < t < 540\,\mathrm{s}$), solid line data. The response is fit (dashed line) to determine the stress-free reference length $L_{\mathrm{e}0}$. 119

3.15 Strain response during 150 cycle shakedown at 379 MPa: (a) every 10th cycle, (b) first cycle shifted. 120

3.16 Relative resistivity $\bar{\rho}_{\mathrm{e}}$ response during shakedown at 379 MPa (every tenth cycle shown). 123

3.17 Strain responses for six constant load shakedown experiments at 30 cycle intervals. As the material is cycled, the area within each loop grows smaller, especially at loads above 191 MPa. 124

3.18 Cyclic evolution of characteristic strains: (a) max & min strains, (b) strokes, (c) residual strains. Shakedown is minimal at 96 and 191 MPa. . . 125

3.19 Evolution of residual strains (open squares) and minimum strain (solid line) during 379 MPa experiment, and minimum strains shifted to account for elastic deformation (dotted line). 128

3.20 Relative resistivity $\bar{\rho}_{\mathrm{e}}$ responses for all constant load shakedown experiments at 30 cycle intervals. 129

3.21 Relative resistivity $\bar{\rho}_{\mathrm{e}}$ maximums and minimums for six constant load shakedown experiments. 130

3.22 Relative resistivity $\bar{\rho}_{\mathrm{e}}$ vs. strain ε_{e} for the 379 MPa experiment, showing every tenth cycle. 131

3.23 Relative resistivity $\bar{\rho}_{\mathrm{e}}$ vs. strain ε_{e} responses for six constant load shakedown experiments at 30 cycle intervals 132

3.24 Two experiments traversing similar square-wave stress-temperature load paths: (a) CCW (b) CW 135

3.25 Responses during 150 counter-clockwise temperature-stress cycles (every tenth cycle shown):a) strain - temperature, b) relative resistivity - temperature, c) relative resistivity - strain. 136

3.26 Responses during 150 CW temperature-stress cycles (every tenth cycle shown):a) strain - temperature, b) relative resistivity - temperature, c) relative resistivity - strain. 138

3.27 Definition of $\varepsilon_{\mathrm{e}}^{\max}$ and $\varepsilon_{\mathrm{e}}^{\min}$ for experiments CW and CCW involving alternating loads. 140

3.28 Comparison of strain extremes between three cyclic load paths: a) $\varepsilon_{\mathrm{e}}^{\max}$ (bold lines) and $\varepsilon_{\mathrm{e}}^{\min}$ vs. cycle number, b) stroke $\Delta\varepsilon_{\mathrm{e}}$ vs. cycle number. . . 141

3.29 (A) maximum and (B) minimum strain responses (bold lines) with two-term Prony series fits (dashed lines) and their asymptotes. 144

3.30 The values for coefficients a_1 and a_2 from Table 3.1 have roughly linear dependence on the stress σ 144

4.1 Prototype uniaxial SMA/bias spring actuator working against an external spring as electrical power is applied to the SMA element: (a) reference configurations for SMA element, bias spring, and external spring, (b) after cooling and assembly, (c) contraction of SMA element during Joule heating, (d) reset configuration when cold. 151

4.2 Admissible phase fraction space: (a) superelastic path in tension, (b) shape memory effect path. 153

4.3 The simplified constitutive model: a) isothermal responses, b) transformation stresses, c) shape memory cycle ($\varepsilon - T$ plane) and superelastic response ($\sigma - \varepsilon$ plane). 160

4.4 Kinetics of phase transformation (P.T.). Transformation rate ($\dot{\xi}$) is zero when $|\mu| < \mu_c$ and infinite when $|\mu| = \mu_c$. 162

4.5 Example actuator response during heating in SMA stress vs. strain space: (a) response with bias spring only (dashed line), (b) response with the addition of the external spring (bold line). 168

4.6 The real-valued part of the Lambert $W(z)$ function. 172

4.7 Dimensionless SMA actuator response: (a) the applied power history (orange line), (b) temperature (red line), stress (green line), and strain (blue line), and (c) M^+ phase fraction (brown line) and normalized driving force (salmon line). Shaded blue and red area indicate time regimes for pure M^+ and A phases, respectively. 179

4.8 Dimensional SMA actuator response in stress-strain-temperature space along with planar projections. 180

4.9 Comparison of simulation histories between lumped actuator model (solid line) and finite element (FE) simulation (dashed line) of Chang, *et. al.* [15]: (a) SMA strain (global average for FE case), (b) temperature (mid-length for FE case), (c) external spring force. 181

4.10 Times to start, t_1 (A_s, blue line), and finish, t_2 (A_f, red line), $M^+ \rightarrow A$ transformation: (a) as a function of convective film coefficient, h, (b) as a function of electrical power input, $\mathcal{P}_\mathrm{e}$. All other parameters are fixed as given in Table 4.2. Open and closed circles correspond to the numerical example. 184

4.11 Minimum bias spring mismatch ($\bar{\Delta}_\mathrm{min}$) as a function of dimensionless bias spring stiffness (η_B) to ensure initial complete reorientation of martensite (M^+) during initial actuator assembly (cold). The gray region indicates feasible designs for values of β and $\bar{\mu}_\mathrm{c}$ of Table 4.3. 187

4.12 Design stresses as a function of bias spring (η_B) and external spring (η_E) stiffnesses: (a) "non-optimal design" using $\bar{\Delta} = 0.0948$ (solid circle is the example of Section 4.3), (b) "optimal design" using $\bar{\Delta}_\mathrm{min}$. 189

4.13 Dimensionless maximum stress ($\bar{\sigma}^\mathrm{best}_\mathrm{max}$) as a function of external spring stiffness (η_E) for dead loading bias ($\eta_\mathrm{B} = 0$, $\bar{\sigma}_\mathrm{o} = \bar{\sigma}_M$). 190

4.14 Dimensionless times (τ) during $M^+ \rightarrow A$ transformation as a function of dimensionless power ($\bar{\mathcal{P}}_\mathrm{e}$): start time, τ_1 (to reach A_s), time period during $M^+ \rightarrow A$ transformation, $\Delta\tau_{12}$, and finish time, τ_2 (to reach A_f); 191

4.15 Minimum power requirements to start (blue line) and finish (red line) $M^+ \rightarrow A$ transformation as a function of film coefficient. 192

4.16 Minimum dimensionless power curves to start $M^+ \rightarrow A$ transformation, $\bar{\mathcal{P}}_\mathrm{e}(A_\mathrm{s})$, in the design space of dimensionless bias (η_B) and external ($\bar{\Delta}$) spring constants. 193

4.17 Minimum dimensionless power curves to finish $M^+ \rightarrow A$ transformation, $\bar{\mathcal{P}}_\mathrm{e}(A_\mathrm{f})$, in the design space of dimensionless bias (η_B) and external (η_E) spring stiffnesses: (a) with fixed $\bar{\Delta} = 0.0948$, (b) with $\bar{\Delta} = \bar{\Delta}_\mathrm{min}$ from eq. (4.4.4). 194

4.18 Input energies as a function of applied power to reach A_f with shaded regions indicated sensible heat and latent heat/spring energy portions. Numerical example values are shown by open and closed circles. (a) Dimensionless plot, holding all SMA parameters and ambient temperature constant, (b) Dimensional plot, holding all parameter values except input power constant. 196

4.19 Energy efficiency ($\mathcal{E}_{\rm out}/\mathcal{E}_{\rm in}$) as a function of dimensionless power ($\bar{\mathcal{P}}_{\rm e}$) for the design of Table 4.3 (bold line) and with minimum bias spring mismatch (thin line). 199

4.20 Contours of energy efficiency ($\mathcal{E}_{\rm out}/\mathcal{E}_{\rm in}$) as a function of dimensionless power ($\bar{\mathcal{P}}_{\rm e}$) and stress ($\bar{\sigma}_{\rm o}$) for a dead load actuator. 200

4.21 (a) Low temperature isothermal experiments on NiTi wire (same material as used in Chang, *et. al.* [15]) with large solid dots showing the termination points of the loading plateaus. The dotted gray line shows the predicted value for martensite reorientation from the constitutive model ($\sigma_M = 78.5$ MPa), which agrees with the plateau stress of the experiment at -50 °C. (b) Dimensionless stress-strain paths of the actuator, $(\bar{\Delta}, \eta_{\rm B}) = (0.0948, 0.05)$, for bias spring only (thick dashed line) and with external spring, $\eta_{\rm E} = 0.025$ (thick solid line). The M^+-fit curve is a fit of the loading plateau termination points, according to eq. (4.4.27), which represents the prestress $\bar{\sigma}_{\rm o}$ of the actuator at "full" M^+. 202

4.22 Contours of constant ε_0 (a) and β (b) as a function of $(\eta_{\rm B}, \bar{\Delta})$. The open circles show the points $(\eta_{\rm B}, \bar{\Delta}) = (0.05, 0.0948)$. 205

4.23 Contours of constant dimensionless actuator stroke, $\Delta\varepsilon$ in the space of dimensionless spring constants $(\eta_{\rm B}, \eta_{\rm E})$ for $\bar{\Delta} = 0.0948$. The open circle shows the point $(\eta_{\rm B}, \eta_{\rm E}) = (0.05, 0.025)$. 205

5.1 Evolution of $A \rightarrow M^+$ transformation fronts in $12.7 \times 2.5 \times 0.25$ mm superelastic NiTi strips, reprinted with permission from Shaw *et. al.* [85]. (a) Strip at constant time intervals during loading, (b) detail view of a criss-crossing transformation front 211

5.2 A derivation of $A \rightarrow M^+$ front inclination angle α. If deformations are isotropic, volume preserving, and extension along the front axis is constrained to be 0, Mohr's circle of strain allows a simple geometric solution. 212

5.3 The two primary propagating front geometries in SMA tubes: (a) helical front with photo and schematic [94], (b) ringed front, with cartoons and photos during loading [31]. Reprinted with permission. 213

5.4 Equal-area texture projection for a NiTi tube specimen– the [111] direction is aligned with the tube axis. 215

5.5 Tube tension setup geometry 215

5.6 Tube tension setup photograph 217

5.7 (a) bare and (b) painted 3.175 mm diameter NiTi tubes 217

5.8 Crown temperature history of a NiTi tube with 60 °C fluid flowing through the center, captured every 0.66 s, at $\dot{\delta}/L = 5 \times 10^{-4}\ \mathrm{s}^{-1}$. Engineering stress P/A_0 is overlaid in black, and numbered circles correspond to full-field IR snapshots in Figure 5.9 221

5.9 Selected infrared images of the experiment in Figure 5.8. In each frame, three separate images of the cylindrical surface have been processed to form a single rectangular surface temperature map. 222

5.10 Mechanical response of the experiment in Figure 5.8. The number circles corresponding to full-field temperature snapshots in Figure 5.9. 224

5.11 Summary of isothermal tube tension experiments, from 10-70 °C: (a) Mechanical responses, (b) propagation stresses vs. temperature, $A \rightarrow M^+$ (red) and $M^+ \rightarrow A$ (blue). 224

5.12 Superelastic mechanical responses of NiTi tubes at various rates 226

5.13 Crown temperature history of a NiTi tube in room temperature air, captured every 1 s, at $\dot{\delta}/L = 1 \times 10^{-4}\ \mathrm{s}^{-1}$. Engineering stress P/A_0 is overlaid in black, with numbered circles corresponding to full-field IR snapshots in Figure 5.14. 228

5.14 Selected infrared images of the experiment in Figure 5.13. In each frame, three separate images of the cylindrical surface have been processed to form a single rectangular surface temperature map. 229

5.15 Mechanical response of the experiment in Figure 5.13. Dotted blue lines represent momentary slopes. 230

5.16 Detail view of a nucleation of A within M^+ during the experiment in Figure 5.13. (a) Mechanical response, (b) Thermal field. 230

5.17 Crown temperature history of a NiTi tube in room temperature air, captured every 1/3 s, at $\dot{\delta}/L = 8 \times 10^{-4}\,\mathrm{s}^{-1}$. Engineering stress P/A_0 is overlaid in black, with numbered circles corresponding to full-field IR snapshots in Figure 5.18. 234

5.18 Selected infrared images of the experiment in Figure 5.17. In each frame, three separate images of the cylindrical surface have been processed to form a single rectangular surface temperature map. 235

5.19 Mechanical response of the experiment in Figure 5.17. 236

5.20 Detail view of nucleations of M^+ (loading) and A (unloading) in the experiment from Figure 5.17. The detail mechanical response in (a) and (b) contain square tags referring to the infrared images in (c) and (d) 237

5.21 Idealized diagram of an M^+ (blue) nucleation within A (red) in an unwrapped tube. A minimum of two of helix rotations are necessary to achieve a uniform distribution of M^+ along each vertical slice, to satisfy axial strain compatibility. 239

5.22 Crown temperature history of a NiTi tube in room temperature air, captured every 1/3 s, at $\dot{\delta}/L = 8 \times 10^{-4}\,\mathrm{s}^{-1}$. Engineering stress P/A_0 is overlaid in black, with numbered circles corresponding to full-field IR snapshots in Figure 5.22. 241

5.23 Selected infrared images of the experiment in Figure 5.22. In each frame, three separate images of the cylindrical surface have been processed to form a single rectangular surface temperature map. 242

5.24 Mechanical response of the experiment in Figure 5.22. 243

5.25 Detail view of nucleations of (a) M^+ (loading) and (b) A (unloading) in the experiment from Figure 5.22. The detail mechanical response in (a) and (b) contain square tags referring to the infrared images in (c) and (d) . 244

5.26 Nucleation of two helical martensite inclusions: At the top are are temperature fields ((a$_1$) and (a$_2$)) from the experiment at $\dot{\delta}/L = 8 \times 10^{-4}$ s^{-1} with a plot of the material phase along the helical axis s_t described in Figure 5.27. The independent axis is the normalized grip displacement, offset from the point of nucleation: $\Delta\delta/L = (\delta - \delta_{nuc})/L$. Dark lines represent the position of the 2 tips of the helix. At the bottom are two temperature fields ((b$_1$) and (b$_2$)) from the experiment at $\dot{\delta}/L = 4 \times 10^{-3}$ s^{-1} with the corresponding plot of helix tip position. 246

5.27 Helical coordinate system: s_n is normal to the helical inclusion, s_t is tangent to the inclusion. The origin is at the initial nucleation site. 247

5.28 Cross-section of tube compression fixture (not to scale), with detail view on the right. The stainless steel grips suppress buckling and pass fluid through the tube center, but still allow the tube to slide freely. 252

5.29 Mechanical responses of NiTi tubes in isothermal compression ($\dot{\delta}/L = 2.6 \times 10^{-4}$ s^{-1}) 254

5.30 Mechanical responses of NiTi tubes in compression at various loading rates. Because of the large thermal mass of the buckling-arresting grips, loading rate had a minimal effect on the response. 256

5.31 Stress and crown temperature history of a NiTi tube in uniaxial compression. Self-heating and cooling is smooth during transformation, free of any localized phenomena. 257

5.32 Propagation stresses for all isothermal tube experiments: $A \rightarrow M$ (red) and $M \rightarrow A$ (blue). Blue shading represents regimes of stable M and red shading regimes of stable A. Purple regions represent overlap where phase is history dependent. 258

5.33 Comparison of tension and compression superelastic mechanical responses. Both experiments were performed at slow rates in room temperature air: $\dot{\delta}/L = 1 \times 10^{-4}$ s^{-1} and $\dot{\delta}/L = 5.9 \times 10^{-5}$ s^{-1} for tension and compression, respectively. 260

5.34 Four-point bending machine, front view 263

5.35 Rotation-curvature relationship for the machine in Figure 5.34. Pure bending between the loading wheels is assumed. 264

5.36 Isothermal bending response at 60 °C. In spite of a symmetric machine geometry, there is still some difference between the left wheel rotation θ_1 and right wheel rotation θ_2 265

5.37 Moment-curvature responses for (a) isothermal at $\dot{\theta} = 6.66 \times 10^{-3}\mathrm{rad/s}$, (b) in 19 °C air at various rates. The curvature estimate is normalized by the tube radius r and assumes pure bending behavior. 267

List of Tables

Table

1.1 Various actuator energy densities [54] 6

2.1 Working fluid for SM wire experiments 48

2.2 Transformation temperatures as measured by DSC, with uncertainty of ± 2 °C [89]. 51

2.3 Specific heats (c_o ± 0.02 J/(g·K)) and latent heats of transformation (Λ ± 1 J/g) as measured by DSC [89]. 51

3.1 Fit parameters for ε_e^{min} and ε_e^{max} 143

4.1 SMA parameters. 159

4.2 Spring parameters, environment constants, and electrical power. 176

4.3 Dimensionless parameters. 176

4.4 Dimensionless actuator transformation temperatures. 176

4.5 Numerical results: values at significant times. 182

5.1 Comparison tension and compression response in RT air (from Figure 5.33) using several common metrics. 260

Chapter I

Introduction

1.1 Background

Shape memory alloys (SMAs) are a fascinating group of metals that have two remarkable properties, the *shape memory effect* and *superelasticity* [24, 34, 74]. *Shape memory* refers to the recovery of shape (i.e., strain) after apparent "permanent" deformation (induced at relatively cold temperatures) by heating above a characteristic transformation temperature (often near room temperature). Both the shape memory effect and superelasticity arise from *martensitic transformations* in the SMA material. Martensitic transformations are solid-to-solid phase transformations that occur without diffusion or plasticity, potentially making them reversible. They involve changes in the crystalline structure that can be induced by changes in either temperature or stress. The high temperature, stress-free phase is called *austenite*, which has a high symmetry crystal structure usually based on a cubic lattice. The low temperature, stress-free phase is called *martensite*, which has a crystal structure with lower symmetry, such as tetragonal, rhombohdral, orthorhombic, monoclinic, or triclinic (see any standard introductory materials science text, such as [51], if these are unfamiliar), depending on the particular alloy.

Figure 1.1(a) shows the full shape memory cycle: at the top, the material exists in the symmetric austenite (A) phase. When cooled (right), it takes the asymmetric martensite (M) phase, consisting of an *twinned* crystal structure that preserves the same global shape. This structure can be easily stretched (bottom), *detwinning* the structure to create an apparent permanent deformation (left) but without plastic dislocations. When heated, the material reverts to the symmetric A phase, recovering the deformation.

Superelasticity (also known as pseudoelasticity) refers to the isothermal recovery of relatively large strains during a mechanical load-unload cycle that occurs at temperatures above a characteristic transformation temperature. Figure 1.1(b) shows the superelastic cycle, which is carried out at a high temperature. Starting again with a high temperature A phase (left), stress is applied to the alloy. Under enough stress, the material changes to its M phase, taking advantage of the asymmetric structure to stretch in the direction of the applied load. When the load is removed, the material reverts to its symmetric A phase. These two effects, or some combination, are being exploited in an increasing number of useful devices.

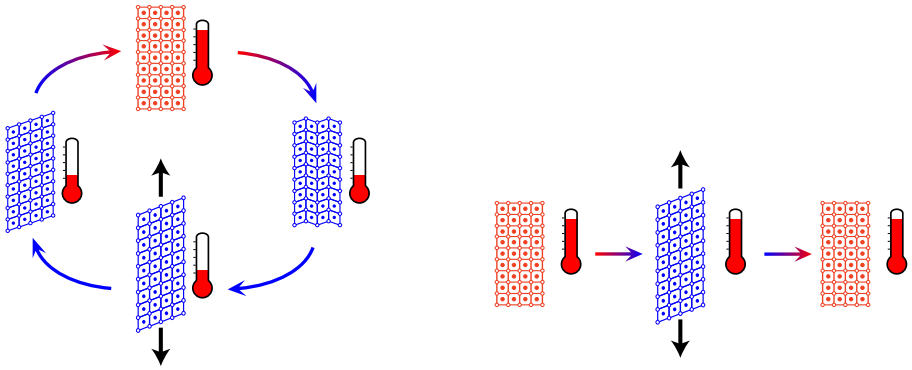
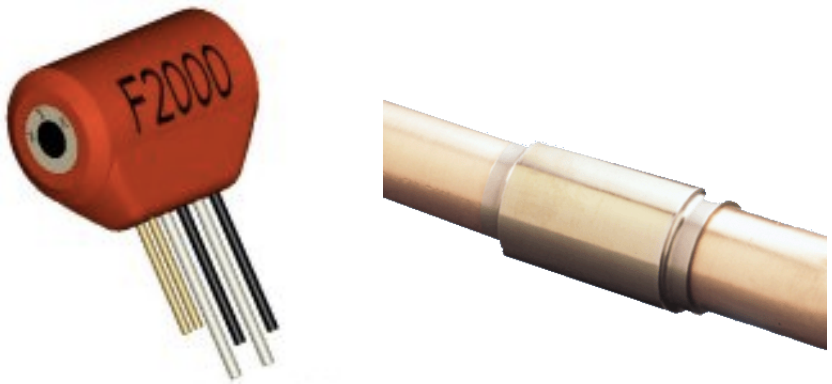


Figure 1.1: (a) The shape memory effect, (b) the superelastic effect



(a) (b)

Figure 1.2: (a) The Frangibolt™ is fitted around a notched bolt, fracturing it when activated by heating. (b) The Cryofit™ coupler contracts when warmed to room temperature, joining fuel lines without dangerous welding or brazing

SMA devices can be usually be divided into several categories: sensor, actuators, dampers, and fatigue resistance. Sensors provide a mechanical response and a change in electrical resistivity (ER) in response to a change in temperature. Actuators respond in much the same manner, but the heat energy (from a temperature change) is converted to mechanical work. Shape memory actuators can be one-time, one-way, like the Frangibolt™ explosive bolt replacement and Cryofit™ fuel line coupler [53]. The Frangibolt in Figure 1.2(a) consists of a notched bolt and NiTi collar, which has been expanded from its reference configuration (left side of Figure 1.1(a)) to be larger than the bolt. When heated by the surrounding collar (top of Figure 1.1(a), the NiTi collar contracts, fracturing the bolt at the notch. The Cryofit coupler in Figure 1.2(b) is used in much the same way, except that its transformation temperature is much lower than the Frangibolt. The coupler is stored in liquid nitrogen in its expanded form, then (while still cold) loosely fit between two lengths

of pipe. As the coupler warms to room temperature (just like the heated Frangibolt), it contracts, sealing the two sections together. Actuators can also undergo (up to) millions of two-way cycles with the addition of an external bias load.

SMAs can also use the superelastic effect to either provide excellent fatigue resistance, energy dissipation, or both. Many cellular phone antennas are made of NiTi wire because of its fatigue resistance. Many medical devices, such as vascular stents [9], joint pins [103], orthodontic wires [2], and eyeglass frames [25] (see Figure 1.3, take advantage of fatigue resistance, and to a limited extent, the shape memory effect. The ability to use SMAs as simultaneous structural elements and energy dissipators (dampers) is of special interest to civil engineers [16, 97].

A large number of shape memory alloys have been discovered since the mid to late 1900s, and the list continues to grow [82]. Many of these alloys, while scientifically interesting, consist of precious metals or only exhibit useful properties as single crystals, which do not lend them to practical use in commercial applications. A few alloys, however, have emerged as commercially viable for novel devices. These include certain copper alloys (CuAlZn) and nickel-titanium-based alloys, such as near equiatomic NiTi (50 to 50.8 at % Ni), known as Nitinol (first discovered in the early 1960s [53]) and some ternary alloys such as NiTiCu and NiTiNb (see Fig. 1.4). To date, it is fair to say that NiTi-based SMAs have the best shape memory and superelasticity properties of all the known polycrystalline SMAs. The NiTi family of alloys can withstand large stresses and can recover strains near 8 % for low cycle use or up to about 2.5 % strain for high cycle use. Their energy density is also very large, orders of magnitude above other common active materials like

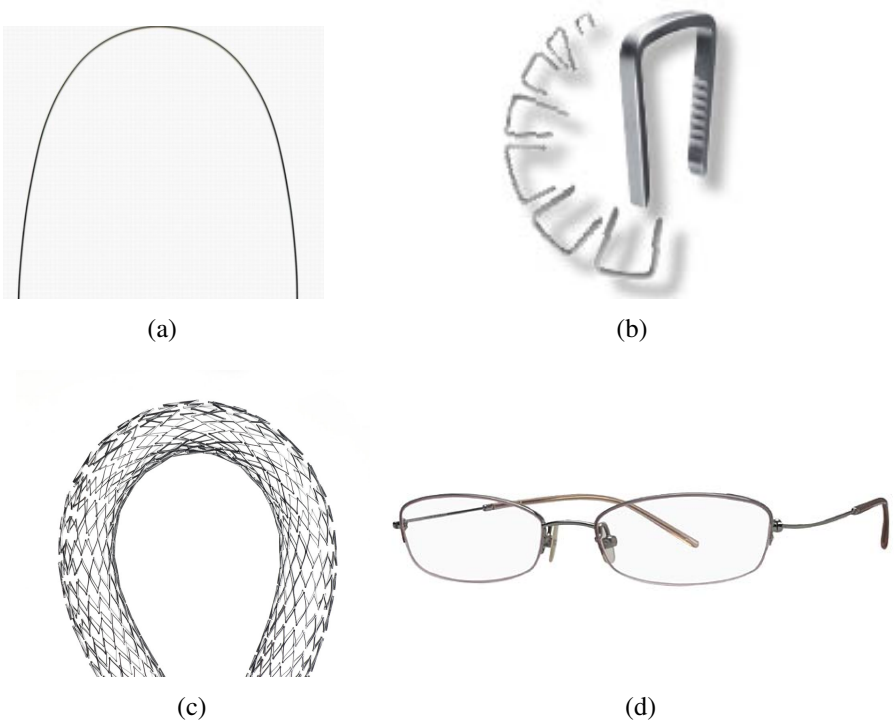


Figure 1.3: Medical devices which use NiTi: (a) Orthodontic archwire (b) compressive staples to repair bone fractures (c) arterial stent (d) eyeglases

piezoelectrics and shape memory polymers, as Table 1.1 shows. NiTi SMAs have further advantages in terms of corrosion resistance, fatigue resistance, and biocompatibility [80], thereby making them the preferred material system for most shape memory applications being considered today.

The materials science and mechanics literature regarding SMAs is vast, and a complete review will not be attempted here (see [73] for a recent review). The field remains an active area of research, and the understanding of the mechanisms involved at all scales from the crystalline lattice to the macroscopic scale has progressed significantly, even during the past decade.

SMAs exhibit some rather surprising phenomena as well as extreme sensitivities to test-

Actuator type	$W(J/m^3)$	*Description*
NiTi SMA	2.5×10^7	one shot, 5% strain, 500 MPa
	6.0×10^6	thousands of cycles, 2% strain, 300 MPa
Solid-liquid phase change	4.7×10^6	water (2.2 GPa bulk modulus)
		acetamide (8% volume change)
Thermo-pneumatic	1.2×10^6	measured, 20 N, $50\ \mu$m displacement
Thermal expansion	4.6×10^5	ideal, Ni on Si, 200 °C
Electro-magnetic	4.0×10^5	ideal, variable reluctance
	2.8×10^4	measured, variable reluctance
	1.6×10^3	measured torque, external field
Electro-static	1.8×10^5	ideal, 100 V, $0.5\ \mu$m gap
	3.4×10^3	measured, comb drive
	7.0×10^2	measured, integrated force array
Piezoelectric	1.2×10^5	calculated, PZT, 40kV/cm
	1.8×10^2	calculated, ZnO, 40kV/cm
Muscle	1.8×10^4	measured, 350kPa, 10% strain
Microbubble	3.4×10^2	measured, 71 μm bubble

Table 1.1: Various actuator energy densities [54]

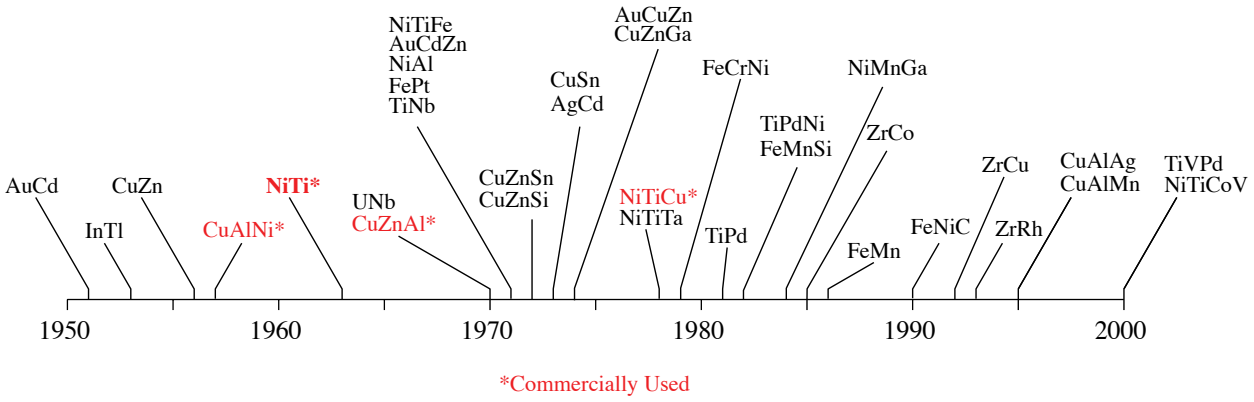


Figure 1.4: History of the discovery of important shape memory alloys. [82]

ing conditions, which can create pitfalls in material testing and interpretation for someone who is new to SMAs. Nevertheless, the need for high quality data is increasingly common in industry, especially since small changes in chemistry or processing can result in both quantitative and qualitative differences in the material behavior and new experimental alloys are continually being developed. Unfortunately, the testing of SMAs is not yet standardized (although there is work in this direction at ASTM), and unlike conventional alloys, material property tables either are not available or provide incomplete, or even incorrect, information for the user. This is, perhaps, not surprising since SMA behavior is nonlinear, hysteretic, and extremely temperature dependent, requiring more properties to be known than is usual for conventional alloys. Since each SMA is different, the user is often faced with testing SMAs in their own laboratory to obtain a satisfactory characterization of the material at hand. Even under simple tensile loading, performing meaningful experiments on SMA wire is not a trivial matter, and good practice is rarely explained sufficiently in the journal literature.

1.2 Motivation

Advances in materials processing have resulted in production of Nitinol SMAs with good quality control, reproducible properties, and in relatively large quantities. Nitinol wire, in particular, is being produced with excellent properties and is relatively inexpensive compared to most other forms. This trend has widened their domain from their beginnings in the medical and aerospace industries. Though they remain relatively expensive compared to structural metals, their cost has reached a level ($x < \$3/\mathrm{m}$ for 0.375 mm diameter

NiTi wire) that it is no longer an impediment in even const-sensitive fields, such as automotive [13] and civil engineering systems [68]. Adoption is broadening, yet SMAs remain a difficult group of materials to use, so there remain many gaps in the literature. These gaps are widest for the engineer new to the field, unfamiliar with specialized techniques for characterizing SMA wires.

For example, a common method of performing constant load experiments is to suspend a weight from a wire. The temperature is roughly controlled through resistive (joule) heating. Temperature is measured by a thermocouple either spot-welded or glued to the wire. In order to obtain an accurate temperature measurement, the current through the wire is changed slowly. However, it is difficult to hold a steady temperature during joule heating. In addition to unsteady convective currents, during transformation the wire's electrical resistance can drop over 20%. Measurement of temperature via thermocouple is problematic during joule heating, since changing the current through the wire too quickly can induce a current in the thermocouple, altering the measurement. Unless special care is taken, joule heating can cause ground loop current leakage through the thermocouple. Methods for alleviating these kinds of problems will be discussed in Chapter III.

Another example of a common pitfall is performing superelastic experiments under load control. For an experimentalist familiar with testing traditional alloys, it may seem like performing an isothermal experiment on superelastic NiTi wires would be a satisfactory method. However, this control method can overlook a major portion of the material response, where the tangent modulus is quite flat. As an example, Figure 1.5 from Berg [6] shows a bending experiment on NiTi wires performed under load control. Because of the

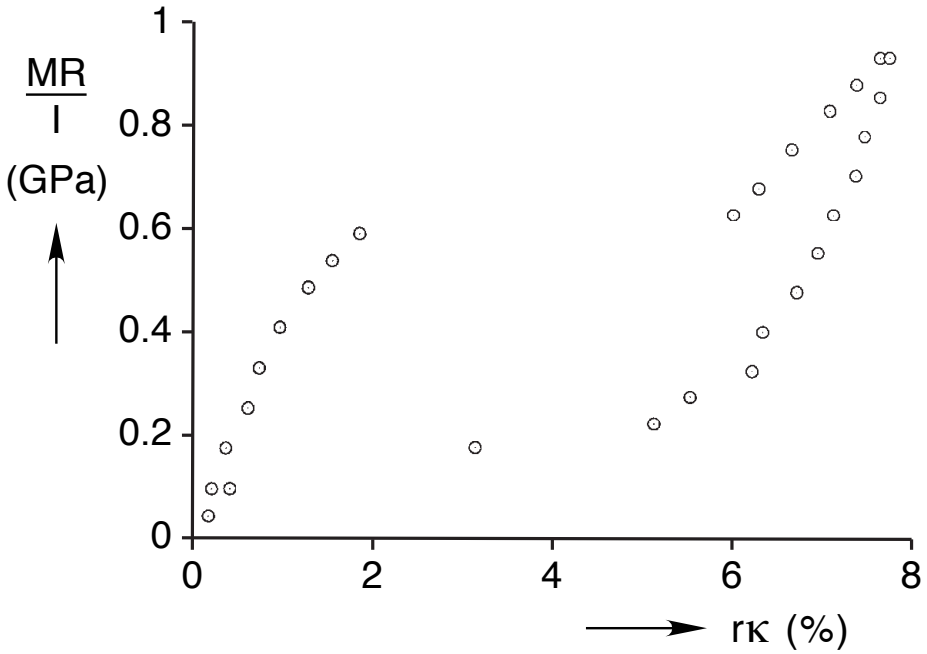


Figure 1.5: Superelastic bending response of NiTi wires under load control. Data during transformation is sparse due to the choice of load rather than displacement control.

nearly flat tangent modulus of the wire during transformation, data is sparse for that portion of the response . The strain rate during this period is unknown, and could not be captured unless the experiment were redone using rotation control. The effect of strain rate on SMA response will be explored in depth in Chapter V.

This thesis aims to fill the gap in the literature by simultaneously explaining common experimental pitfalls, methods to avoid them, and some basic SMA phenomena using SMA wires and tubes. At the same time, it investigates SMA wire behavior in increasing detail, dealing first with the measurement of electrical properties, and then with the evolution of the thermo-electro-mechanical response over multiple thermomechanical cycles.

In fact, a universal problem for any SMA designer is that of shakedown, which is the evolution of material properties as a result of thermomechanical cycling [93]. A device

that cycles an SMA multiple times must, in its design, account for any subsequent shakedown in SMA properties. All SMAs experience shakedown on some timescale, though depending on the conditions it may only be noticable over tens to thousands of cycles. Thus, either the application must account for changing properties, or the material must be pre-conditioned to achieve a repeatable behavior. On the other hand, the designer can use shakedown as an advantage, to achieve new material behaviors like the two-way shape memory effect [42, 62, 76], or to eliminate the flat plateau in super-elastic responses which proved problematic in the experiment from Figure 1.5. While a number of models have been proposed to predict shakedown behavior [10, 102], and many shakedown experiments have been performed [3, 41, 79, 98], there remains a need for higher quality experimental data suitable for thermo-electro-mechanical constitutive modeling that can then be used for performance and lifetime predictions in actuator applications.

Consider also the needs of that actuator designer to model the behavior of their device. The trend in modeling SMAs has tended towards increasingly complex and accurate models (see [15, 44, 52, 78] for a few recent examples). At the same time, the solution of even simple boundary value problems is not a straight-forward matter, requiring the solution of coupled, partial differential equations (PDEs) governing equilibrium, heat transfer, and phase transformation kinetics. Usually this must be accomplished numerically to capture the detailed time-dependent nonuniform distributions of the state variables. The time and resources required for an engineer to take advantage of these advancements are considerable, especially for those who are not experts in the field.

For design and optimization studies, a more tractable setting is desirable, so this thesis

introduces a simple model of a prototype SMA wire/bias spring actuator that admits analytical solutions for the time-dependent behavior. One of the few examples in the literature coming closest to a closed-form SMA actuator solution was provided in Wu, *et. al.* [101], but unlike our current case, the SMA constitutive model used resulted in a linear PDE that was solved by separation of variables in terms of an infinite series solution in a somewhat different boundary value problem context. The simplicity of our analytical solution allows this model for a spring-based actuator to be implemented in a matter of days rather than weeks or months.

1.3 Organization

This thesis is organized as follows. Chapter II uses historical data from experiments performed by John Shaw and Mark Iadicola (both published and unpublished) to outline a set of experimental techniques for characterizing SMA wires. First, Section 2.1 presents a pair of differential scanning calorimetry (DSC) experiments and shows how to interpret them to obtain material transformation temperatures and enthalpies of transformation. After a review of basic thermomechanical SMA behavior in Section 2.2, a set of techniques are presented for performing uniaxial constant-temperature experiments. Special attention is given to the accurate measurement of strain and temperature. Isothermal experiments are presented and interpreted for the two different NiTi alloys. Section 2.5 uses additional techniques for measuring local strains and temperatures to capture the nucleation and propagation of localized transformations in uniaxial NiTi wires. It is important for an experimentalist to understand how these localizations occur and how to recognize them, if

only to learn how to avoid misinterpreting experimental data.

Next, the techniques from Chapter II are refined to include electrical resistivity (ER) measurements during constant-load thermal cycling. The experiments in Chapter III were performed by the author on conditioned SMA wires (Flexinol™ from Dynalloy, Inc.), rather than the unconditioned "virgin" wires used in Chapter II. Keeping with the theme of speaking to the device designer, conditioned wires were chosen because they are representative of the material used in automotive and aerospace applications, which must operate over a broad range of vehicle temperatures (-50 to 50 °C). Electrical measurements were included because SMA systems sometimes use ER measurements as feedback on the state of the device, since they cost less to implement than strain or temperature measurements. In fact, ER measurements are often used in place of DSC to obtain stress-free transformation temperatures. The chapter starts in Section 3.1 with isothermal characterization at numerous temperatures of the Flexinol wire, and highlights the differences between it and virgin wires (Nitinol) from Chapter II. One of those differences is the presence of a significant two-way shape memory effect (TWSME) in the Flexinol, which results in extension upon cooling even in stress-free conditions. After Section 3.1.2 introduces a novel method for measuring ER and strain at the same local gauge length, the ER response and its relationship with intermediate R-phase (somewhat less well-known in NiTi) is examined with a series of constant load experiments in Section 3.3. Section 3.4 then presents a series of six 150-cycle constant-load shakedown experiments at loads progressing from 19 to 478 MPa.

Now that a designer knows how to characterize SMA wires through either isothermal or isobaric experiments, they may need a model to describe how their device will func-

tion. Chapter IV presents a simple, yet functional, model of an SMA wire and bias spring subjected to electric joule heating. Section 4.1 defines the thermomechanical problem of interest for our prototype actuator system and shows how the full 1-D constitutive model of Chang, *et. al.* [15] can be reduced to a "lumped" model, having a minimum set of degrees of freedom, by neglecting the details of nonuniform strain and temperature fields. The resulting model has similarities to many others, including Brinson [14], but with simpler kinetics. With some simplifying assumptions, the set of governing PDEs is reduced to a single nonlinear, ordinary differential equation (ODE) in time. Dimensionless parameters and state variables are defined in order to study the scaling of the solution over a large range of pre-strain, spring properties, environmental conditions, and power inputs. Section 4.2 shows how the governing ODE during phase transformation can be integrated exactly in terms of a known special function, the *Lambert* function, which can be used piecewise in time as a constant electrical power is toggled. This is the first such analytical solution for a thermo-mechanically coupled SMA/bias-spring model that we know of. Section 4.3 provides a numerical example of the time-dependent behavior of a prototype actuator and discusses sensitivities of various parameters. Section 4.4 discusses a systematic approach to the design and optimization of such as actuator, considering the sizing of springs, actuation stroke, actuation times, minimum power requirements, and energy usage and efficiency.

Chapter V uses many of the techniques introduced in Chapter II, but takes a departure from simple 1-D wire to characterize thin-walled material-structure interactions in superelastic Nitinol tubes. These tubes allow thermomechanical characterization in not only tension, but compression and bending, providing a wealth of constitutive and struc-

tural information. Chapter V characterizes the response of SMA tubes in three loading modes: tension, compression, and bending. Isothermal responses are presented for each mode, followed by a study investigating the influence of loading rate on the response in each mode, demonstrating some phenomena not reported before. In Section 5.2, special attention is paid to the nucleation and propagation of transformation fronts in tension, and the important effect of loading rate on front morphology.

Chapter II

Fundamental Characterization of SMA Wires Through Differential Scanning Calorimetry and Uniaxial Tension Experiments

Differential scanning calorimetry (DSC) is often the first step in the characterization of SMA wires, establishing transformation temperatures and latent heats of transformation. In this chapter, DSC thermograms are shown on two commercially-available Nitinol wire alloys of slightly different compositions. One exhibits shape memory relative to room temperature (RT), designated SM wire; the other exhibits super-elasticity at RT, designated SE wire. Section 2.2 then provides a description of the basic thermomechanical behavior of uniaxial SMA wire, and thermomechanical data on these two alloys are presented to demonstrate the shape memory effect and superelasticity. Section 2.1.2 summarizes recommendations and conclusions for obtaining high quality data of SMA wire at a basic level of thermomechanical characterization.

This chapter then discusses techniques and good practice for isothermal temperature control and strain measurement. The characterization of the same two alloys is continued through a series of isothermal experiments to show the dramatic range of tensile responses in a temperature window spanning the respective stress-free transformation temperatures.

Characteristic transformation stresses and strains are mapped to quantify the thermomechanical sensitivities in the material.

Difficulties in performing experiments on shape memory alloys are often rooted in one (or a combination) of two phenomena. The first is material-level instabilities, and the second is thermomechanical coupling associated with the latent heat of transformation. Both create particular, and somewhat subtle, issues for the SMA experimentalist. The chapter concludes with a discussion of the experimental issues associated with the first one, namely localization of strain and temperature fields. These are often overlooked by the novice, yet they play an important role in the kinetics of stress-induced phase transformation, and in turn, exacerbate the material response's sensitivities to loading rate and ambient media. Special experimental techniques are introduced to enable (otherwise) difficult to measure features in the underlying material response. The focus is on the superelastic response of the SE NiTi wire that was used in Sections 2.1 and 2.3.

2.1 Differential Scanning Calorimetry

In Nitinol (near equiatomic NiTi), the austenite has an ordered B2 crystal structure, which can be viewed as two interpenetrating simple cubic lattices of Ni and Ti, respectively[a]. The martensite has an ordered B19′ crystal structure (CsCl prototype), where B19 denotes an orthorhombic structure (AuCd prototype) resulting from unequal normal strains relative to the $<110>$ directions of the austenite structure, and the prime ($'$) indicates that it has additionally been distorted by a shear strain, resulting in monoclinic symmetry. An-

[a]It is often incorrectly called a body centered cubic (BCC) structure, which would only be accurate if the material was monoatomic.

other intermediate phase that sometimes appears is called the R-phase, which is a rhombohedral distortion of the B2 structure (see Figure 2.1)[b], we deonte as B2′. The gray planes shown in the figure's crystal structures indicate lattice correspondence planes between the standard unit cells shown, i.e. $(110)_{B2} \leftrightarrow (110)_{B2'} \leftrightarrow (001)_{B19'}$.

The first step in characterizing an SMA material is to determine the characteristic transformation temperature. Actually, the material is hysteretic and there are several transformation temperatures to speak of, including the austenite start temperature (A_{s}) and the austenite finish temperature (A_{f}) during heating, and the martensite start temperature (M_{s}) and the martensite finish temperature (M_{f}) during cooling. Additionally, an intermediate phase (R) often appears during cooling, having its own start temperature (R_{s}) and finish temperature, (R_{f}), before the transformation proceeds to martensite at lower temperatures. Under stress-free conditions, these are commonly measured by differential scanning calorimetry (DSC) thermograms, either provided by the material supplier or obtained by the user. Alternate methods exist to measure transformation temperatures, such as electrical resistivity scanning, but while potentially convenient (in case a DSC instrument is not available), they are more difficult to interpret and do not provide any information about latent heats of transformation or specific heats.

2.1.1 Two DSC Experiments

This chapter focuses on two different Nitinol alloys obtained from Memry Corp., one that has a stress-free transformation above room temperature (termed "shape memory wire")

[b] In the interest of simplicity, the figure ignores the complex stacking structures that actually occur in the superlattice. See [74] for details.

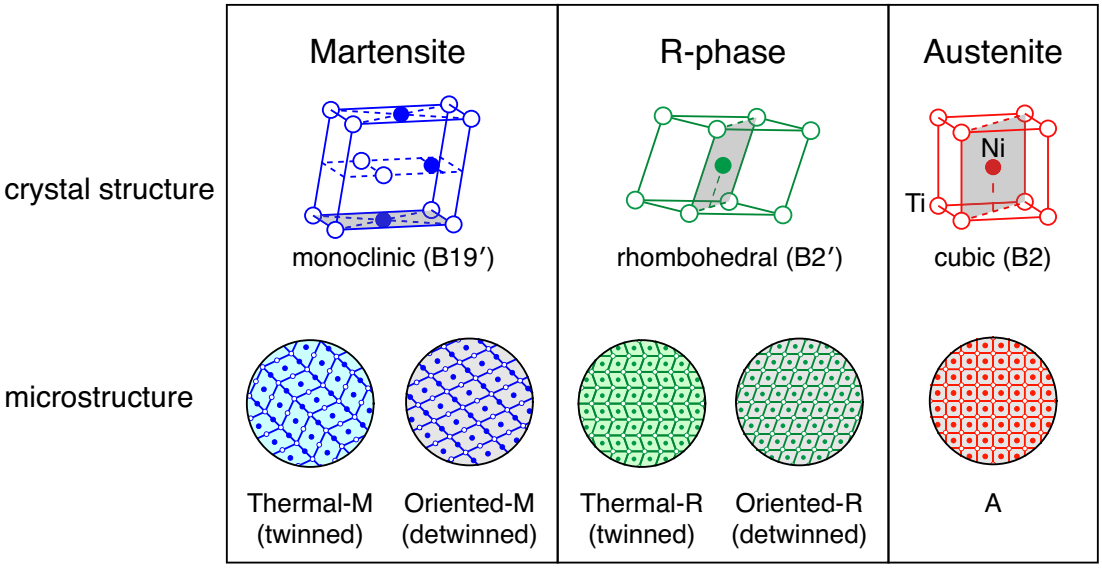


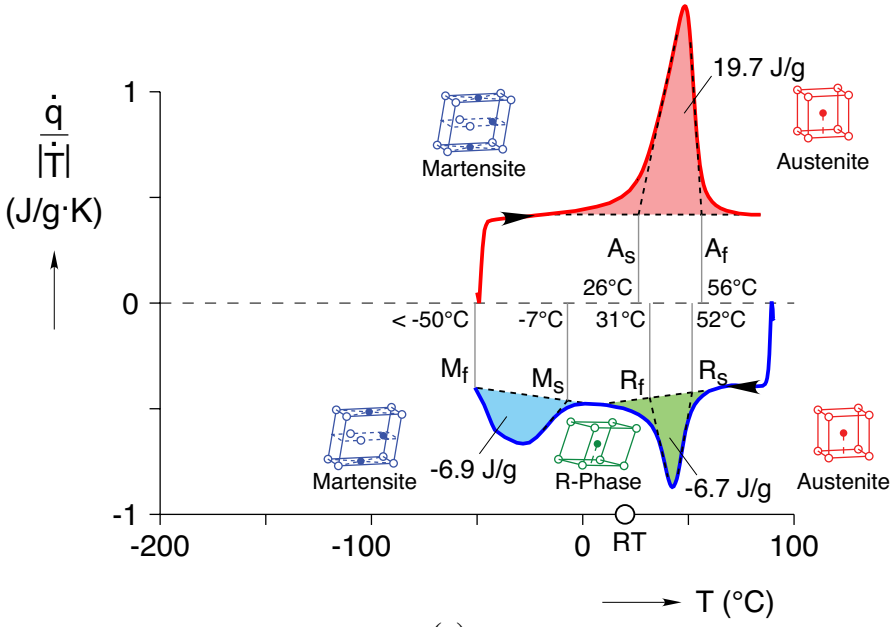
Figure 2.1: Schematics of crystal structures and microstructures in Nitinol.

and one that has a stress-free transformation below room temperature (termed "superelastic wire"). It is well-known that transformation temperatures of a Nitinol alloy can be tailored by the supplier anywhere from cryogenic temperatures to as high as about 100 °C by small changes in chemistry, by aging heat treatments in the range 350 °C to 500 °C, and by thermomechanical processing (cold work developed during wire drawing, and/or cyclic loading performed by certain suppliers). The first order effect on transformation temperatures is alloy chemistry. Just 1 % excess Ni above stoichiometric NiTi (i.e, 50 at % Ni and Ti) can suppress transformation temperatures by over 100 °C! While not measured directly, this shape memory wire likely has a composition near 50.2 at % Ni, while the superelastic wire has a composition near 50.6 at % Ni (with the balance being Ti, except for trace amounts of other interstitial elements like C, N, and O). These alloys were probably aged by the manufacturer near 500 °C for 10-15 minutes after cold working. They were not subjected to any further thermomechanical cycling, so the as-received state is referred to as so-called

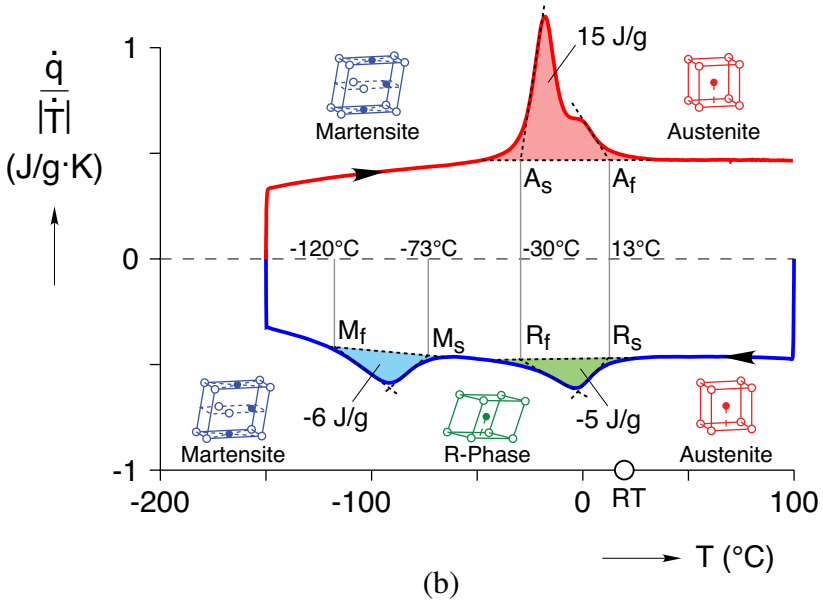
"virgin Nitinol".

DSC thermograms for the two selected NiTi alloys are shown in Figure 2.2. A small sample of material (typically 50 mg to 100 mg) was placed in a sample pan and then installed in a differential scanning calorimeter (Perkin-Elmer Pyris 1 DSC). The DSC is programmed by the user to scan the sample at a constant temperature rate while monitoring the heat flow input, $\dot{Q}$, to the sample pan compared to another empty pan (hence the "differential" part of the name). The vertical scales in the figure have been converted to specific heat-like units, J/(g·K), as described below. The upright peak and two upside-down peaks seen in each thermogram occur due to the respective endothermic (positive) and exothermic (negative) latent heats of transformation, respectively. The peaks are bell-shaped, so the start and finish temperatures of a transformation are typically extracted by a straight line constructions fitted to the steepest sides of the peak. The particular transformation temperature is read off the intersection of this line with a "baseline" that cuts off the peak. One can appreciate that this procedure has some uncertainty associated with it, since it depends on the shape of the enthalpy peak and the baseline that is chosen.

The latent heat peaks separate temperature regimes where pure (or nearly so) solid-state phases exist. At sufficiently high temperature the material is *austenite*, and at sufficiently low temperature the material is *thermal-martensite*. At intermediate temperatures upon cooling the material is *thermal R-phase*. As the temperature traverses a given latent heat peak the material has a progressive mixture of the two phases from either side of the peak. Each peak represents "excess" heating or cooling needed to maintain the temperature rate, and it is caused by the extra energy addition or subtraction needed to transform the crystal



(a)



(b)

Figure 2.2: Differential scanning calorimetry (DSC) thermograms of two Nitinol alloys: (a) shape memory wire [86], (b) superelastic wire [15].

structure. All first-order phase transformations are associated with latent heat of transformation, regardless of the form of the phases.

As a familiar example of a phase change, latent heating/cooling is required to cause ice to melt or reform, although that is a solid-liquid transformation. The current case, by contrast, is a solid-solid transformation. The latent heat is smaller, 15 J/g to 20 J/g, compared to ice-water 334 J/g. Also, the transformation temperature is path dependent, while melting or freezing of water-ice at atmospheric pressure occurs at 0 °C with little hysteresis. Unlike water-ice, martensitic transformations usually occur under non-equilibrium thermodynamic conditions, thereby exhibiting overall temperature hysteresis, a finite width in the latent heat peaks, and possible rate-dependence (discussed later) in the DSC measurements.

The material state, therefore, is quite history dependent, and transformations occur in a multi-step manner, as is typical for Ni-rich NiTi alloys that have been heat-treated by the supplier. Note in particular that between M_{s} and A_{s} the material state can be either thermal-R or thermal-M, depending on the prior temperature history. The SM wire, for example, can exist in either phase at room temperature (RT) as seen in Figure 2.2a. Thus, the as-received wire is presumably thermal-R, since the prior history involved heat treatment by the supplier at much higher temperatures. During cooling, separate peaks are evident for the sequential $A \rightarrow R$ and $R \rightarrow M$ transformations in Figure 2.2a and 2.2b. Upon heating, however, the austenite peak has a single large peak in Figure 2.2a, but has a strange double hump shape in Figure 2.2b. The latter reveals that separate, but overlapping, $M \rightarrow R$ and $R \rightarrow A$ martensitic transformations actually occurred during heating. Often these two peaks overlap to an extent that they cannot be distinguished, resulting in one large peak (as

in Figure 2.2a).

The specific latent heat Λ (per unit mass) for a given transformation is obtained by integrating the shaded area of the peak by

$$\Lambda = \int_{t_1}^{t_2} \dot{q}_{\mathrm{L}}(t)\ \mathrm{d}t, \tag{2.1.1}$$

where $q \equiv Q/m$ is the heat energy per unit mass, m is the mass of the sample, and t_1 and t_2, are the respective initiation and ending times of the transformation on the baseline. The specific latent heat power history $\dot{q}_{\mathrm{L}}(t)$ is

$$\dot{q}_{\mathrm{L}}(t) = \dot{q}(t) - \dot{q}_{\mathrm{C}}, \tag{2.1.2}$$

where $\dot{q}(t)$ is the specific power history measured by the DSC, and $\dot{q}_{\mathrm{C}}$ is the specific sensible heat power (associated with the specific heat of the material) estimated along the baseline (nearly a constant for NiTi). Alternatively, the latent heat can be written as

$$\Lambda = \int_{T_1}^{T_2} \frac{dq_{\mathrm{L}}}{dT}\ \mathrm{d}T, \tag{2.1.3}$$

where $T_1 = T(t_1)$ and $T_2 = T(t_2)$ are the respective initiation and ending temperatures chosen along the baseline, and $dq_{\mathrm{L}}/dT = \dot{q}_{\mathrm{L}}(t)/\dot{T}(t)$ where $\dot{T}$ is the temperature scan rate. The raw power measured by the DSC has been converted to $\dot{q}/|\dot{T}|$ in Figure 2.2, where $\dot{q}$ is the power per unit mass, to make a clear graphical interpretation of the signed shaded areas of the thermogram as latent heats. The remaining area for each heating/cooling direction

between the baseline and zero is associated with the sensible heat changes. Note that T_1 and T_2 should encompass the entire temperature range of the transformation and will be slightly different than the "start" and "finish" temperatures obtained by the dotted construction lines. The latent heats for the $A \rightarrow R$ transformation are about -6.7 J/g and -5 J/g for the respective alloys. The latent heats for the $R \rightarrow M$ transformation are about -6.9 J/g and -6 J/g for the respective alloys, although this one has a larger uncertainty due to the slope and curvature of the baseline cooling curve. The lumped latent heats during $M \rightarrow R \rightarrow A$ transformation during heating are about 19.7 J/g and 15 J/g, respectively. Note that if the specimen had only been cooled to -50 °C, or so, only one peak would exist on cooling ($A \rightarrow R$) and the peak on heating would have a smaller latent heat, since it would only include the $R \rightarrow A$ contribution. Such a thermogram often creates confusion if one does not recognize that the martensite transformation has been cut off. This is a common pitfall when dealing with room temperature superelastic NiTi where the martensite transformation occurs at very low temperatures, requiring a DSC capable of liquid nitrogen cooling.

Typically the heating and cooling transformations between A and R exhibit little hysteresis, less than about 2 °C. The transformations between A and M, however, are quite hysteretic, exhibiting 80 °C and 75 °C temperature offsets between the heating and cooling peaks for the respective alloys. Digressing for a moment, both transformations, the "large one" $A \leftrightarrow M$ and the "smaller one" $A \leftrightarrow R$, are technically considered "thermoelastic martensitic transformations" (not to be confused with the names of the phases) in the sense that they are diffusionless and reversible (yet hysteretic). Both have a shape memory effect relative to austenite that can be used in applications, yet the former is better known since

it has a larger strain change ($\approx$ 5 %) compared to the latter ($\approx$ 0.5 to 1 %). The large hysteresis of the $A \leftrightarrow M$ transformation, however, has an adverse effect on the cyclic stability and ultimate fatigue resistance under transformation cycling. Conversely, the $A \leftrightarrow R$ transformation has better cyclic repeatability, so it might be useful, depending on the intended application, if the smaller strain change and smaller stress level are acceptable. The reason for the dramatic difference in hysteresis between the two transformations is the fact that the rhombohedral R-phase crystal structure is kinematically compatible with that of A; whereas, the monoclinic structure of M is not compatible with A, requiring internal twinning of M to produce approximately compatible habit planes between M and A with consequent plasticity near the interface (see [8] for a more complete discussion of kinematics and compatibility requirements in NiTi crystals). In any event, the large hysteresis between the onsets of transformation (M_{s} versus A_{s}) is generally attributed to nucleation energy barriers, associated with compatibility requirements at the microscale, that must be overcome to initiate transformation. The width of the latent heat peak ($A_{\mathrm{f}} - A_{\mathrm{s}}$, for example) is due to resistance to growth (frictional kinetics) of the daughter phase within the parent phase. These features arise due to the strong role compatibility plays in the nucleation and kinetics of martensitic phase transformations. By comparison, these energy barriers are more easily surmounted during diffusional phase transformations or liquid-solid phase transformations where thermal activation is dominant (perhaps, more familiar to typical metallurgists and thermodynamicists).

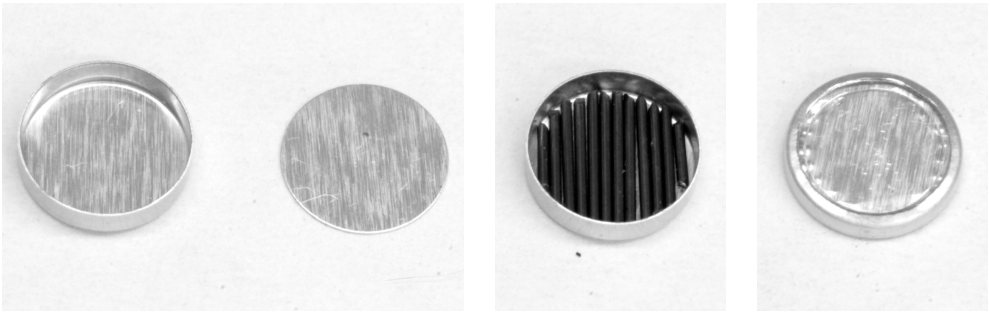
As can be seen in Figure 2.2, the specific heat of the material is about 0.45 J/(g·K) to 0.50 J/(g·K) in the austenite range. The curvature of the baseline seems to be sensitive to

the tuning/calibration of the DSC machine, especially at very low temperature where liquid nitrogen cooling is required. Based on experience the specific heat in the martensite range seems to be about the same as for austenite. Many material supplier property tables have reported a value of 0.87 J/(g·K) for the specific heat of Nitinol, but in all likelihood this value was incorrectly extracted from the maximum $\dot{q}$ of a latent heat peak, meaning that it included both specific heat and latent heat contributions.

2.1.2 Some Recommendations on DSC Measurements

Specimen preparation can affect the quality of the results. It is important to ensure good thermal contact between the sample pan and the material specimen, which will ensure that the temperature read by the DSC thermocouple accurately measures the temperature of the sample. This is generally optimized by having thin, flat specimens. In this case, wire was cut into a number of short lengths that were placed side-by-side (see Figure 2.3), creating a single layer of material that filled the sample pan (to approximate the shape of a wafer). To avoid plastic deformation, which introduces residual stresses that could skew results, the wire was cut using a low-speed, water-cooled diamond cutoff saw. The lid was placed over the sample and crimped to ensure good thermal contact. Care was taken to keep the specimen free of contaminants, including moisture, to avoid introducing unwanted artifacts into the thermogram.

Despite the fact that the transformations are generally characterized as *athermal*, meaning that they are not thermally activated nor rate-dependent according to the usual Arrhenius law (which would apply to diffusional transformations), the consistency of the DSC



(a) (b) (c)

Figure 2.3: Photographs of (a) sample pan and lid, (b) wire specimens, and (c) crimped assembly for DSC.

measurements can be affected by the heating/cooling rate specified by the user. Commonly, a rate of ± 10 °C/min is used (as used in Figure 2.2), which is a compromise between trade-offs. Higher rates tend to give sharper enthalpy peaks, which can be integrated more accurately, but may not give accurate transformation temperatures due to thermal lag. Lower temperature rates will reduce thermal lag, but excessively slow rates can make the enthalpy peaks rather indistinct.

2.2 Basic Thermomechanical Behavior

This section briefly describes the underlying micro-mechanical mechanisms responsible for the shape memory behavior and superelasticity of SMAs. Two thermomechanical experiments are then presented to demonstrate these features for the two alloys previously addressed.

2.2.1 Shape Memory & Superelasticity

Under stress-free conditions, the low symmetry martensite lattice can exist in several lattice correspondence variants (12, in the case of Nitinol), which are reflections or rotations of each other. Consequently, martensite can exist in many different microstructures, depending on the thermomechanical history (see [8] for details on the connections between lattice parameters and compatible microstructures). Thermal-martensite is internally twinned, where twin-related variants create a coherent mirror image of the lattice across each twin boundary (special crystallographic planes). Some requirements for shape memory behavior to exist are (1) the transformation between austenite and martensite occurs with little volume change, and (2) the distortional strains relative to austenite are relatively small, typically on the order of 10 %. In other words, the structure change can occur by small, coordinated shifts of the atomic positions without diffusion or plasticity. The shape memory effect arises from the fact that martensite can arrange itself into a self-accomodating, finely twinned (heterogeneous) structure with little, or no, macroscopic strain relative on average to austenite. Hence, upon cooling from austenite to martensite little, if any, strain (or shape change) is usually observed (unless the material has been heavily processed to have a so-called two way shape memory effect). This self-accommodated form is termed *thermal-martensite*.

If the martensite becomes sufficiently stressed at cold temperatures it exchanges certain variants by the motion of twin boundaries to others more preferentially aligned with the stress, which is termmed *oriented-martensite* or *tensile-martensite*[c]. The resulting macro-

[c]Historically, thermal-martensite and tensile-martensite have often been called twinned-martensite and

scopic strain remains upon removal of the stress, since all variants are again energetically equivalent. Figure 2.4 shows transmission electron micrographs, for example, taken by Nishida, *et. al.* [71] of equiatomic NiTi specimens in the martensite phase before and after tensile testing taken to several different residual strains after unloading. Thermal-M is shown in Figure 2.4a before loading with a mixture of many twins in a self-accomodating arrangement. Tensile-M is shown in Figure 2.4b at 6.4 % strain with relatively few twins remaining (and some roughening due to plasticity), while Figure 2.4c and Figure 2.4d show further roughening due to plasticity, yet some twins remain even at these extreme strains (11.3 and 24), which is thought to be partly responsible for the material's excellent ductility by providing an alternate, and reversible, accommodation mechanism to plastic slip. In this case with Ni-rich Nitinol, the aging treatment results in a fine dispersion of Ni_4Ti_3 precipitates that greatly improves resistance to slip compared to equiatomic NiTi [33], which promotes good superelasticity at high stresses. Upon stress-free heating, the structure reverts back to the cubic austenite and the strain is recovered, provided little plasticity has occurred during loading. The initial shape of the specimen is recovered, since the required atomic shifts are relatively small and the high symmetry cubic structure is unique relative to the martensite variants. This sequence is the mechanism of the shape memory effect.

At high temperatures (above A_{f}) the material is stable in the austenite phase under stress-free conditions. Adding sufficient stress (isothermally), however, can destabilize austenite in favor of one, or more, martensite variants (called *stress-induced martensite*, similar in microstructure to tensile martensite discussed previously, but arrived at by a dif-

detwinned-martensite, respectively, but this is a bit misleading, since martensite in most SMA polycrystals cannot be completely detwinned to a single variant without concurrent plastic slip.

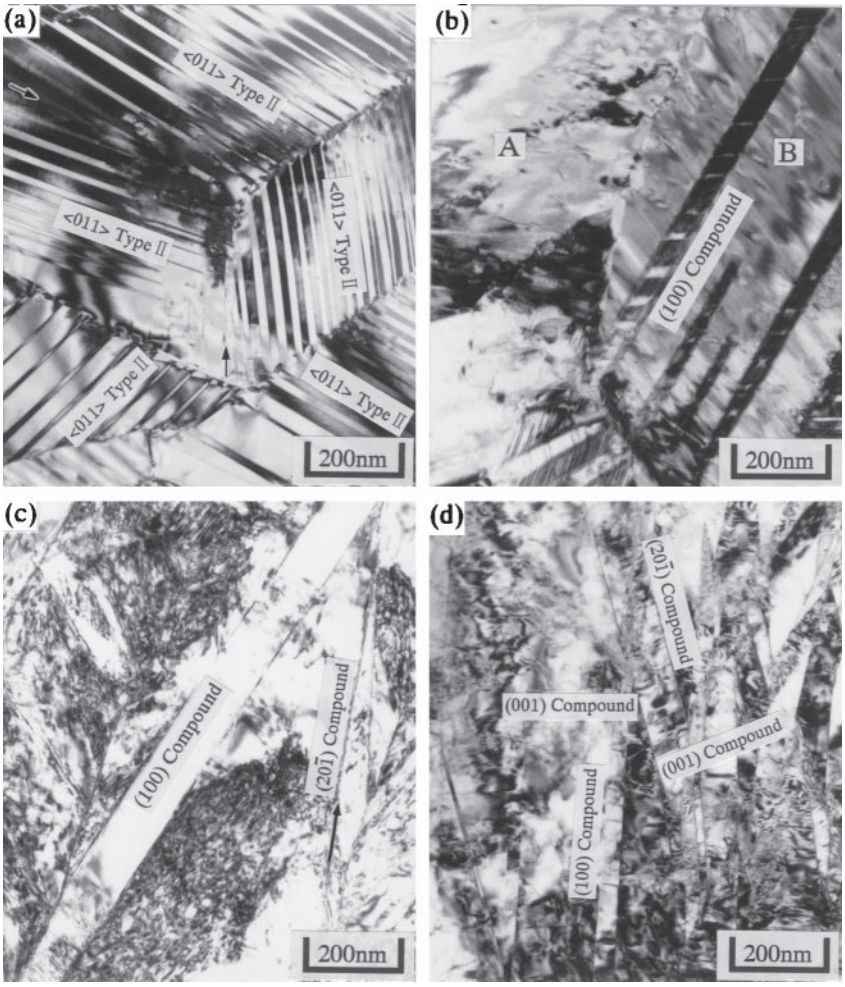


Figure 2.4: Transmission electron micrographs of martensite in a NiTi alloy at residual strains: (a) 0 %, (b) 6.4 %, (c) 11.3 %, and (d) 24 %. Images taken from Nishda, *et. al.* [71] (reprinted with permission from Elsevier, http://www.sciencedirect.com/science/journal/13596462).

ferent process) and a macroscopic strain occurs. During subsequent removal of the stress, another lower critical stress is reached at which martensite is no longer thermodynamically stable and the material reverts to austenite and the strain is recovered upon complete unloading. This isothermal, yet hysteretic (in stress), sequence is the mechanism of superelasticity.

2.2.2 Two Thermomechanical Experiments

Figure 2.5 shows experimental results (data taken from [84]) on the shape memory wire in a 3D plot against stress (vertical axis), strain (horizontal axis), and temperature (oblique axis). The stress is measured by P/A_0, axial load over initial cross-sectional area. Strain is measured as $\varepsilon_{\mathrm{e}} = \delta_{\mathrm{e}}/L_{\mathrm{e}}$ based on the elongation (δ_{e}) and gage length (L_{e}) of a miniature, waterproof extensometer. Temperature (T) is measured by a 76 μm exposed junction K-type thermocouple attached to the specimen. The diameter and free length of the specimen were $d = 1.02$ mm and length $L = 63.5$ mm, respectively. The specimen had been quenched in liquid nitrogen (77 K) to ensure that it started in the thermal-M phase. The experiment was performed in a conventional electromechanical testing machine (Instron model 5585) with the specimen immersed in a special water bath to control the temperature. Initially, the specimen was allowed to warm from 77 K to only 10 °C (state ⓪) and then subjected to a slow isothermal displacement-controlled (δ) load-unload cycle (⓪ to ②) at $\dot{\delta}/L = 4 \times 10^{-4}\ \mathrm{s}^{-1}$. This process converted the material from *thermal-martensite* (denoted M) to *tensile-martensite* (denoted M^+), leaving a 5 % residual strain. The testing mode was then switched to load control to maintain a small positive stress, just

to avoid buckling of the wire, and the temperature was raised to 70 °C while monitoring the strain. The strain was relatively static until the temperature reached above 45 °C, at which point the strain decreased rapidly toward zero (③ to ④) while the material transformed from tensile-martensite (M^+) to austenite (A), then it remained static at a small value until state ⑤ . This sequence ⓪ to ⑤ in the experiment demonstrates quantitatively the shape memory effect.

Next, an isothermal load-unload cycle was imposed under displacement control again at 70 °C (⑤ to ⑩ in the figure), which demonstrates superelasticity in the same specimen. Between states ⑤ and ⑥ the austenite response is nearly linearly elastic ($E = \approx 70\mathrm{GPa}$), like a conventional metal, but at ⑥ the apparent tangent modulus suddenly switched to near zero with elongation continuing at constant stress. The cross-head motion was reversed at ⑦ where the strain was a bit larger than 6 %, at which point a large fraction of the specimen was M^+. Unloading proceeded from ⑦ to ⑧ along a nonlinear path with a lower tangent modulus than previously for $A \rightarrow M^+$ (from ⑤ to ⑥). At ⑧ the stress reached another plateau, along which the material reverted to A. This occurs since M^+ (*stress-induced martensite*) is not thermodynamically stable at this elevated temperature without sufficient stress. Between ⑧ and ⑨ the strain decreased at constant stress until the path nearly met the initial elastic response of A at ⑨ . The final segment between ⑨ and ⑩ was elastic unloading of A, and the strain was very nearly recovered. Note that the recovery of strain occurred through a large hysteresis, indicating that a significant portion of work energy was dissipated in the material during this cycle[d].

[d] In this regard, "superelasticity" is a bit of a misnomer, since while the strain is in fact recovered the process is not "elastic" in the sense normally used in mechanics (which requires path-independence during

One should note that in Figure 2.5 martensite reorientation (⓪ to ①) is associated with a relatively soft and nonlinear response, and unloading exhibits a large hysteresis. The initial response of austenite in Figure 2.5 (⑤ to ⑥), is stiff by comparison[e]. The overall thermomechanical response is quite nonlinear and involves significant hysteresis, or path dependence, in stress-strain-temperature, so it is impossible to develop a simple function that relates stress, strain, and temperature in an algebraic way. Consequently, many constitutive models have been proposed that only treat certain aspects of the behavior, or are developed according to time-dependent formulations, making it a challenging and ongoing research topic still today.

Figure 2.6 shows the results of a similar experiment performed on superelastic wire. Qualitatively, the features described above are the same, except that the transformation temperature is lower so that the material is superelastic at room temperature, and the shape memory effect is observed by chilling the material to subambient temperatures, then loading/unloading, and then allowing it to warm to room temperature. Interestingly, for this alloy the stress required to orient martensite (⓪ and ①) is less than for the shape memory wire. This experiment was performed in a temperature-controlled air chamber rather than in a water bath due to the low temperatures involved. As a result, the temperature is not precisely controlled, and the path between ③ and ④ , during the shape memory recovery, is somewhat jagged compared to the previous case in water. The strain measurement was

loading and unloading.) Accordingly, some prefer to call this response "pseudoelasticity".

[e]Incidentally, this is opposite to the well known case of carbon steel, where austenite is the soft phase and martensite is the hard phase. The austenite in steel has a cubic structure and martensite has a tetragonal structure. A shape memory effect is not possible in steel, since the transformation involves a significant volume change, requiring a large energy barrier to be surmounted (so called non-thermoelastic martensitic transformation).

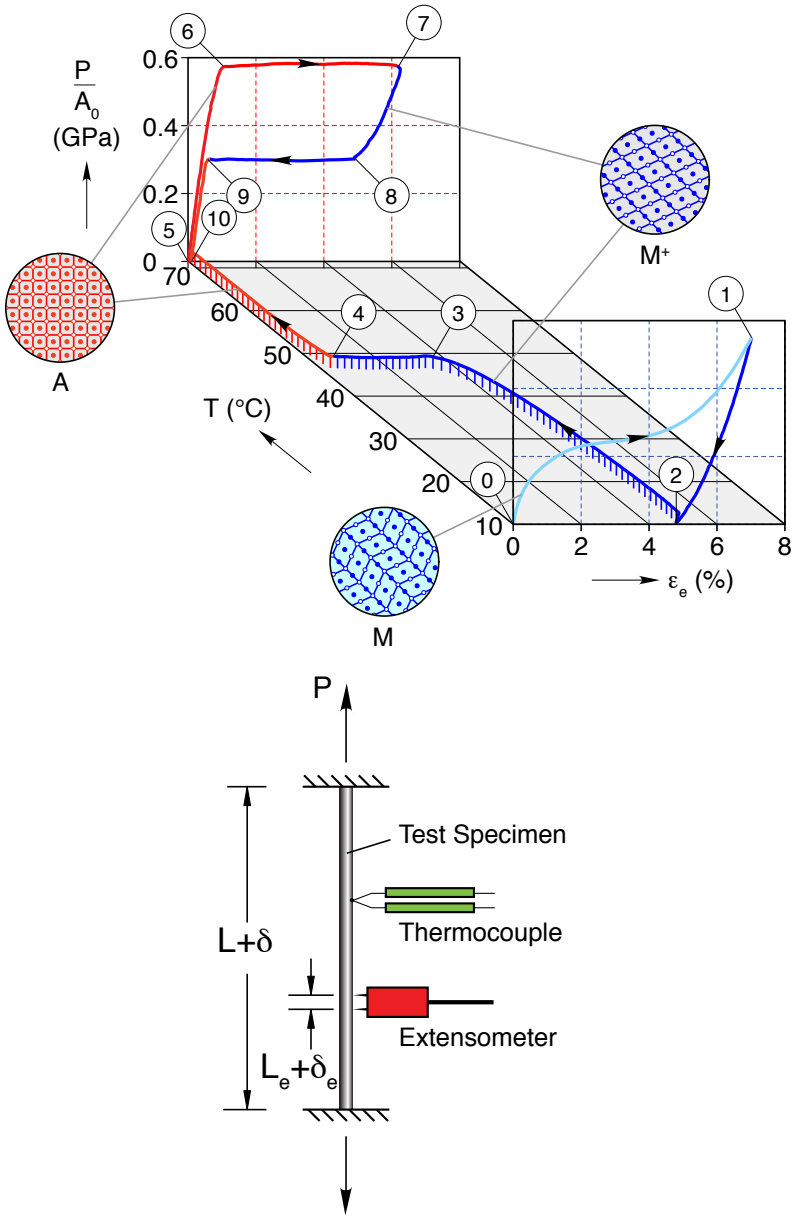


Figure 2.5: Thermomechanical experiment on shape memory Nitinol, showing shape memory effect (0 to 5) and superelasticity (5 to 10). Data taken from [84].

obtained by a laser extensometer system (Electronic Instruments Research, model EIR-05) exterior to the chamber that transmitted a planar laser sheet through the viewing glass of the chamber and received a signal reflected from retro-reflective tags[f] attached to the wire specimen.

2.3 Uniaxial Experimental Methods

Before presenting experimental data, this section provides an overview of different methods for thermomechanical testing of SMA wire. The section discusses the tradeoffs associated with the choice of loading method, strain control/measurement, and temperature control/measurement.

2.3.1 Loading Method

Once DSC scans have been performed to establish the relevant transformation temperatures (as discussed previously in Part 1 [89]), it is necessary to obtain an overview of the material behavior with respect to axial load-elongation-temperature space. A clear picture of the thermomechanical behavior of an SMA requires multiple experiments over a wide range of temperatures. Mechanical data at one or even two temperatures is not generally sufficient due to the material's extreme temperature sensitivity, as will be demonstrated in Section 2.4. Thermomechanical data is customarily obtained in one of two methods: either (1) holding constant axial load, sweeping temperature up and down, while measuring the

[f] Similar material is used as reflective decals on sports shoes, and is made of a polymer embedded with glass micro-beads.

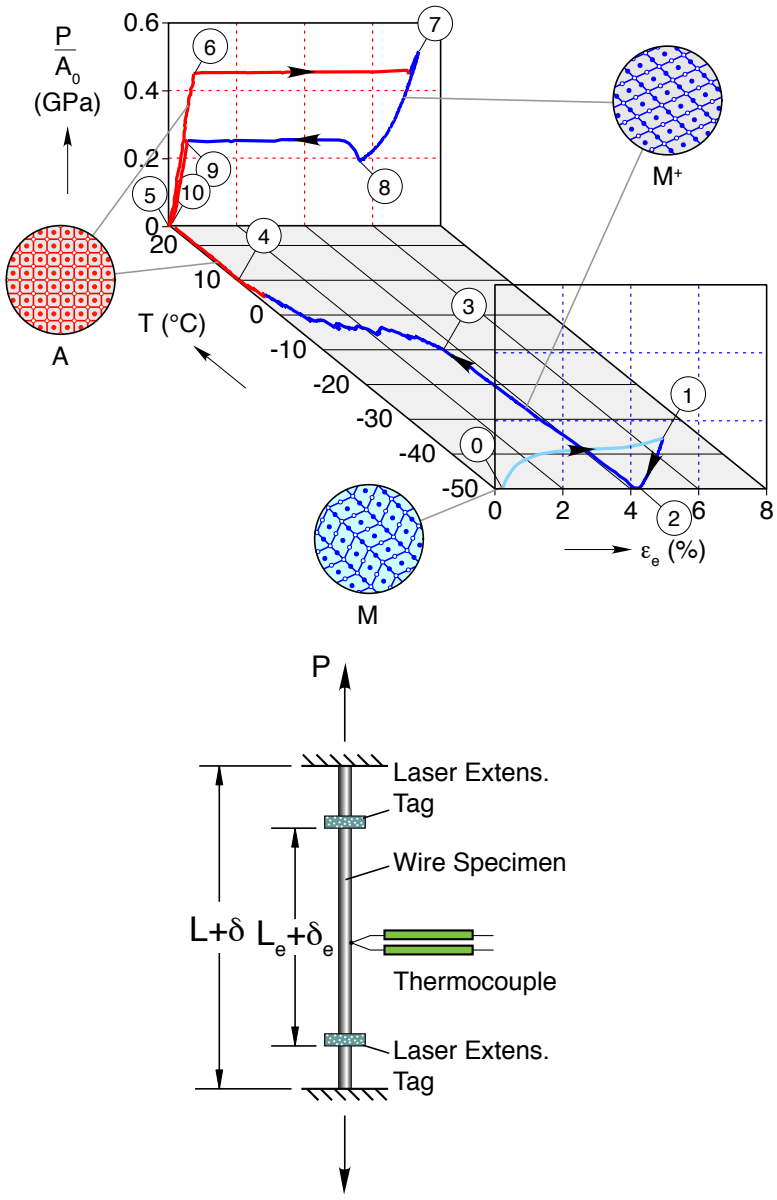


Figure 2.6: Thermomechanical experiment on superelastic Nitinol: shape memory effect (0 to 5) and superelasticity (5 to 10). Data taken from [15].

strain response, or (2) holding constant temperature, sweeping the elongation up and down, while measuring the axial load response.

The first method can be executed by hanging a weight on the wire or using a conventional testing machine in load-control, a so-called *soft-loading device*. A servo-hydaulic testing machine most easily accomplishes the task, although an electro-mechanical testing machine may be used if a suitable feedback control is available to simulate load-control. Several temperature scans at different load levels are required to obtain a view of the phase transformation "surfaces" in force-elongation-temperature space. Depending on the setup, it may be difficult to precisely control the temperature rate, and thermal inertia of the setup may limit it to inconveniently slow rates.

The second method is most easily executed with a screw-type, electro-mechanical testing machine in displacement-control, a so-called *stiff-loading device*, although it can also be accomplished using elongation-control feedback. The control loop in the machine must be carefully tuned to the stiffness of the test specimen being used. We generally prefer the stiff-loading device (isothermal mode) for SMA constitutive experiments, since it tends to be more well-behaved and reveals more detailed behavior if instabilities occur in the material, as discussed further in Section 2.5. The material behavior is history dependent (hysteretic), and for now, the focus is on the monotonic load-unload behavior across the strain range where obvious phase transformation occurs, i.e. the so-called "outer loops" of the behavior.

2.3.2 Strain Measurement & Control

The next issue the experimentalist faces is the choice of strain measurement and control, a choice made especially important by peculiarities in SMA behavior. Here are some common choices progressing from simple to complex.

Global Strain Measurement

The simplest strain measurement can be derived from the displacement of the testing machine's cross-head (δ), thereby giving a "global", or average, strain measure δ/L, where L is the free length of the wire specimen between the grips. This works well, provided the stackup of fixtures between the base and cross-head is much stiffer than the test specimen. While grip displacement provides the preferred control method (as discussed later), it may not give the best strain measurement due to stress/strain concentrations induced at the grips and possible grip slippage.

Local Strain Measurement

It is of course good practice to measure the "local strain" in the gage section of a test specimen for accurate strain measurement, and SMA testing is no exception. The need is obvious for standard "dog-bone" specimens, which are thicker at the ends and taper to a uniform middle gage section, since strain gradients across the tapered region and strains measured from grip displacement will not match the average strain in the gage section. The need may be less obvious for straight wire specimens, but it is still necessary to avoid grip

artifacts in the strain measurement and to capture possible strain localization, as explained in Section 2.5.

- *Strain gages*: While strain gaging is commonly used to accurately measure local strains in gage section of conventional specimens, it is usually not a practical option for SMA wire. Attaching strain gages to very thin wire specimens is difficult, and most strain gages are incapable of withstanding strains to ≈ 8 % and back without suffering plastic deformation.

- *Mechanical extensometer*: Alternatively, a clip-on extensometer can be easily installed and can follow axial strains of up to 20 % or larger, depending on the model. They provide accuracy near to that of strain gages. One must ensure, however, that the weight of the extensometer, which is attached at two knife edges and cantilevered from the specimen, does not induce any significant bending of the wire specimen. For very thin wire specimens, it may be necessary to counter-balance the extensometer, so that it "floats" aside the specimen.

- *Video extensometer*: A non-contacting extensometer is another instrument that can be used to virtually eliminate any loading artifacts that might arise from mechanical extensometry or strain gaging, which is especially useful for testing very compliant specimens. An optical video extensometer is a non-contacting measurement system that uses targets illuminated by visible light. A high resolution digital camera takes images of the specimen, while a computer processes the images in real time, noting the distance between two or more visible markers. These may be tags affixed to the specimen or painted markers. Fairly inexpensive systems (about $10,000) can be

used when the strains are quite large, such as when testing elastomeric specimens. These have not been typically used for stiff metallic specimens where the strains are relatively small, but more expensive systems ($25,000 and up) are now available that provide sufficient accuracy at small strains (resolutions down to 0.50 μm) and high frame rates (above 200 kHz).

- *Laser extensometer*: Another non-contacting version of the above concept is a laser extensometer system (such as Electronics Research Corp.'s, model EIC-05, used throughout this thesis), where thin retro-reflective tags are affixed to the specimen and a laser sheet is shined across the tags. The system in turn detects the reflected laser signal from the tags and gives an accurate measurement of the distance between the edges of the tags (L_{e}). The resolution of the system used here is about 2 μm, so the accuracy of the length measurement is mostly dependent on the quality of the tag attachment and alignment of the laser sheet with the specimen axis. Overall, the strain accuracy is nearly as good as a mechanical extensometer, and it can be used to very large strains (up to several hundred percent). It also circumvents the need to attach any bulky instrumentation to the specimen, which might prevent other types of imaging, and adds essentially no additional loads or moments, which is especially useful for testing very thin wire. This requires a line of sight between the laser system and the specimen, and if the laser sheet is planar it can be used through a window of an air chamber without any parallax concerns. A laser extensometer system can be used at rates up to 100 Hz, but the signal can be noisy unless oversampled in time, effectively limiting the measurement rate to about 20 Hz. This is sufficient for slow

to moderate strain rates of typical wire SMA specimens.

Full-field Strain Measurement

The above methods provide strain measurements of specific regions of interest in a specimen defined by the locations and spacing of tags/markers/knife edges. The methods are discrete in nature, giving strain measurements necessarily averaged across a region of interest. Following are two full-field strain measurement techniques that give a broader view of the strain field in the specimen.

- *Brittle Coatings*: A classical technique for full-field strain measurement is to coat the specimen with a brittle coating that changes its reflectivity as it is strained and to then take photographs of the specimen during testing. The optical characteristics of the coating can be calibrated to known strains, but the accuracy is usually limited. The information achieved is full-field in nature, but is fairly qualitative and can only be used for one loading cycle without reapplication, since the brittle coating is usually permanently deformed at typical SMA strains. In fact, certain Nitinol alloys have a naturally occurring dark oxide layer that can sometimes be used in this way to track non-uniform strain fields [87].

- *Digital Image Correlation (DIC)*: A more quantitative technique involves digital image correlation (DIC), where digital images are taken during testing and then post-processed to give the displacement and strain fields along the specimen [95]. The resolution and accuracy of the measurement is dependent on the details of the specimen surface specularity, which may require the specimen surface to be treated/painted to

give good results (requiring some art). This is the most elaborate technique mentioned here and it is included only for completeness, yet if well calibrated it can provide a wealth of deformation information, especially when strain fields are non-uniform and deformations are multiaxial [23].

Gripping Issues

A few words should be said about gripping wire specimens. The setups to be presented here used flat, hardened steel plates each having a shallow v-groove to aid in specimen installation and alignment. In setups where the specimen was immersed in a fluid bath the plates were simply clamped together by capscrews. This was adequate in most cases, but high torques were required to prevent excessive slippage at the highest loads. In setups in an air chamber pneumatic grips were used to maintain constant pressure between the grip plates. This worked better since the clamping force was held constant even as the specimen strained and thinned across the diameter. Nevertheless, some amount of grip slippage is nearly unavoidable when clamping wire in this manner for additional reasons (which will be discussed in Section 2.5). Spool grips are also available from various testing machine manufacturers, which nearly eliminate any stress concentrations at the ends. This is a useful technique if one simply wants to find the breaking strength of the wire, where stress concentration could otherwise cause premature failure. It is not recommend for careful experiments for constitutive characterization, since the effective free length of the specimen remains uncertain due to friction effects on the spool.

Strain Control

In the experiments to be presented the specimen elongation, δ, was controlled at a prescribed rate ($\dot{\delta}$) during loading ($\dot{\delta} > 0$) and unloading ($\dot{\delta} < 0$), chosen in the range of global strain rates $\dot{\delta}/L = \pm 5 \times 10^{-5}\,\mathrm{s}^{-1}$ to $\dot{\delta}/L = \pm 4 \times 10^{-4}\,\mathrm{s}^{-1}$. These strain rates might seem excessively slow to those familiar with mechanical testing of conventional metal (about $\dot{\delta}/L = \pm 1 \times 10^{-3}\,\mathrm{s}^{-1}$), but are in fact necessary to achieve near isothermal conditions for SMA wire.

As a final caution, one must resist the temptation to use a "strain-control" mode via a local strain measurement (such as extensometer) feedback, due to possible strain localization in SMAs as discussed further in Section 2.5. When using a stiff-loading device, like a screw-type electromechanical testing machine, the imposed end-displacement could wander while straining occurs outside the extensometer region. More dangerously still, when using a soft loading device, like a servo-hydraulic testing machine, this control mode could cause the machine to jump suddenly if no straining is detected within the extensometer region. Instead, grip-displacement (global strain-control) should be used as the control mode.

2.3.3 Temperature Control & Measurement

The next consideration is the choice of method for temperature control and measurement of the specimen, and several are detailed below. Unlike conventional metals, where moderate changes in temperature can be tolerated without affecting the quality of the results, a few degrees change of specimen temperature can significantly influence the results

of an SMA experiment, so more care is typically needed.

- *Joule heating*: Electrical ohmic (Joule) heating is a common method for changing the temperature of an SMA wire during testing, and it is often used for actuation in devices. The method is simple and convenient, but has several limitations. It is typically only used for shape memory (SM) wire (i.e., those with sufficiently high transformation temperatures) for testing in room temperature air, since the ambient temperature must be sufficiently low compared to the material's transformation temperatures. The electrical power and convective environment must be carefully controlled, since the material's electrical resistivity is somewhat phase dependent [18] and stray air currents can cause undesireable fluctuations in the test specimen's temperature. Even seasonal changes and heating/air conditioning systems in the laboratory can affect these results, so a suitable enclosure is often desireable. The electrical attachment method can also be problematic, since any contact resistance may be of the same order as the specimen resistance. Furthermore, temperature measurement is difficult, since thermocouples attached to the SMA wire specimen give unreliable measurements when large electrical current is flowing through the specimen. A non-contacting infrared temperature probe can potentially be used, but this takes some care to get accurate temperature measurements from thin wire specimens.

- *Environmental chamber*: Another common method is to use a temperature-controlled air chamber fitted within a mechanical testing frame. If it is available with a cooling system, such as liquid CO_2 or liquid N_2, the temperature range for testing is considerably broadened into subambient temperatures, accommodating testing of SE SMA

wire across its range of transformation temperatures. Thermocouples can be attached directly to the test specimen to monitor the local temperature in the gage section, but the thermocouple wire should be relatively fine compared to the SMA wire diameter to achieve accurate temperature measurement. Good thermal data can be gathered from 0.076 mm (0.003 in.) diameter exposed junction K-type thermocouples attached to the SMA specimen by small clips. A daub of thermally conductive paste (such as Omegatherm 201) ensures good thermal contact between the thermocouple junction and the specimen surface. It is wise to use additional thermocouples to monitor the air temperature in proximity to the specimen, as well as the temperature of grips which are typically massive compared to the test specimen and have significant thermal lag. Infrared temperature measurement is usually not practical because typical window glass is opaque to infrared radiation. Overall, the advantage of the air chamber method is the flexibility in ambient temperature for testing, but stray air currents within typical commercial chambers can still cause temperature fluctuations, especially at subambient temperatures where forced-air cooling is used (typically ± 5 °C). Attaching a simple flow diffuser to the back of the chamber at the forced-air exit significantly improved the temperature spatial and temporal uniformity to about ± 1 °C.

- *Liquid bath*: One of the most precise temperature control methods involves immersing the test specimen in a liquid bath, such as water, that is temperature-controlled by an external fluid circulator. This advantage comes with some additional complexity and typically requires a custom-made experimental setup. The temperature range for

testing is limited by the freezing/boiling points of the working fluid, but many suitable fluids can be selected based on the known transformation temperatures of the SMA specimen to be tested. Thermocouples can still be used for measurement of fluid temperature and specimen temperature, provided they have been water-proofed (or are compatible with the chosen working fluid.) Mechanical extensometers for strain measurement are typically not water-proof, so some suitable modification or coating addition is required to electrically insulate them. Infrared temperature measurement is not possible, since most liquids are opaque to infrared radiation.

- *Conduction contact*: The last temperature control method to mention is to use a temperature-controlled heat sink in thermal contact with the back side of the SMA wire specimen [49]. Potentially, this can result in an even more precise control of the SMA specimen temperature, provided the specimen diameter is relatively small so that temperature gradients across the diameter can be reasonably neglected (which is often the case). In this scheme, a thermal grease can be used to ensure good thermal contact while allowing the specimen to slide freely in its axial direction. Of course, careful alignment of the experimental stackup is required to make this work well. Also, inserting thermocouple wafers between the heat sink and the specimen allows fine control of the temperature. Using multiple wafers along the specimen length can even allow control of non-uniform temperature distributions (that are useful in some cases, as discussed in Section 2.5. Another important advantage of this scheme is that it leaves the entire front surface of the specimen free for full-field imaging, both optical (strain) and infrared (temperature) measurements.

- *Infrared thermography*: If available, a digital infrared radiometer (thermal imaging system) can give full-field surface temperature information that is especially useful in interpreting experimental results. Typical research radiometers detect temperature changes on the order of 0.1 °C and have array sizes between 256×256 and 1024×1024. If the wire has a sufficiently small diameter, the measured surface temperature is nearly the same as the temperature at the core of the wire (which can be verified by calculating the Biot number, see [50]). The most accurate temperature results with infrared thermography are obtained when the specimen surface has an emissivity near unity (ideal blackbody) and the surface is flat. Nitinol, however, can be obtained in a variety of surface conditions from heavily oxidized (black) to moderately oxidized (coppery) to highly polished (shiny), i.e., with a large range of emissivities. Since accurate temperature measurement requires a sufficiently high emissivity (low reflectivity) of the wire surface, it may be necessary to paint the wire to raise its emissivity above 0.8, generally a good goal. The curved surface of the wire can also create inaccurate temperature measurements if insufficient pixel resolution is available across the wire diameter. Some post-processing of the results may be necessary in this case.

2.4 Isothermal Mechanical Responses

Our laboratory have used most of the above methods at various times, depending on the goals of the specific measurements. For now, however, the focus is on isothermal mechanical responses (over relevant temperature windows) of the same two commercial Niti-

nol wire alloys (Memry Corp) that were previously introduced in Section 2.1. As before, the alloy with stress-free austenite transformation temperatures above room temperature is termed "shape memory (SM) wire" (same material as used in [84, 86]), while the alloy with transformations below room temperature is termed "superelastic (SE) wire" (same material as used in [15].)

2.4.1 Experimental Setup

The experiments to be presented on the SM and SE wires were performed in a liquid bath and an air chamber, respectively, as shown in Figure 2.7 and Figure 2.8. The custom-built liquid setup of Figure 2.7 for SM wire experiments had a stainless steel specimen bath with a glass front and an o-ring seal in its base. This allowed the bath to be lowered for specimen installation into the grips and then raised, immersing the specimen and its instrumentation, for testing within an electro-mechanical testing machine. The long grip extension was made of stainless steel and was designed to be several orders of magnitude stiffer than the test specimens. The bath temperature was regulated by an external circulator (Neslab RTE-110) that reached temperatures in the range -20 °C to 120 °C, depending on the working fluid listed in Table 2.1. This range spanned the window of transformation temperatures for the SM wire. The fluid temperature was preset at the external circulator, accounting for any heat losses/gains along the fluid piping by measuring the specimen bath temperature by a temperature probe.

The load frame was operated under displacement control at the rather slow elongation rate of $\dot{\delta}/L = \pm 4 \times 10^{-4}$ s^{-1} to maintain nearly isothermal conditions. The nominal axial

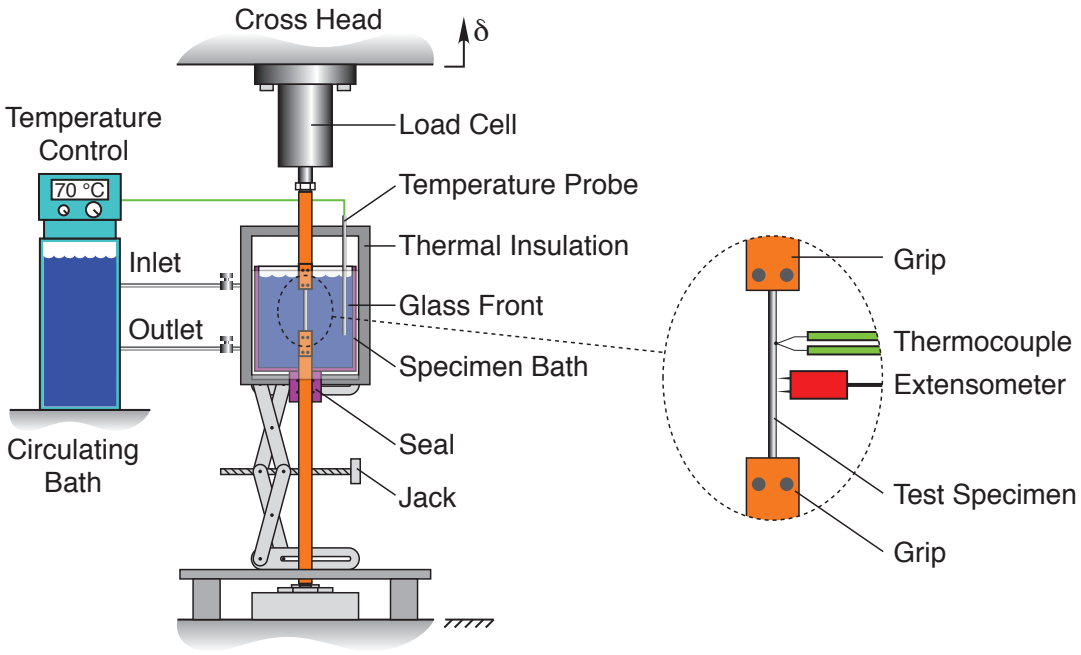


Figure 2.7: Setup for experiments on SM wire in a temperature controlled bath.

Table 2.1: Working fluid for SM wire experiments

Working Fluid	Temperature Range
glycol/water	-18 to 10 °C
water	10 to 80 °C
mineral oil	80 to 100 °C

stress was measured as the ratio (P/A_0) based on the measured tensile force (P) and the original cross-sectional area (A_0). Local strain measurement was obtained by a custom-built, miniature, water-proof extensometer having a gage length between the knife edges of 2.54 mm (0.1 in) (see [84, 86] for details). The strain measured by the extensometer was the ratio (ε_e) of the elongation in the extensometer δ_e to the gage length of the extensometer at zero load L_e. Local temperature measurement was obtained by one or more small thermocouples attached to the specimen.

The SE wire required yet lower temperatures to span its transformation temperatures, so a commercial air chamber with liquid nitrogen cooling was used instead as shown in Figure 2.8. The top/bottom openings of the air chamber were sealed with thermal insulation to be as airtight as possible without constraining the motion of the cross head. Strain was measured through the front observation window by a laser extensometer, with a gage length near 60 mm. Because air is more thermally insulating than liquid, the strain rate was reduced to a very slow $\dot{\delta}/L = \pm 5 \times 10^{-5}$ s^{-1} to avoid self-heating/cooling during phase transformation.

In both cases, the axial load on the specimen was measured by a high quality 5 kN (1000 lb) load cell that was balanced to account for the weight of fixturing above the specimen. The load cell was shielded on the underside to protect it from overheating or condensation/frost buildup due to convection and conduction from within the chamber. A small fan kept room temperature air blowing across it to minimize temperature changes in the load cell that would otherwise result in load measurement drift.

For all experiments presented, the uncertainty of the measured load and strain is on the order of the thickness of the thickest line in the figure. For data derived from construction lines or estimated from response curves, the uncertainty is about the size of the diameter of circles shown in the plot, with a worst case error of about three circle diameters in the few cases where points were estimate by "eye" (explicit error bars are omitted for clarity).

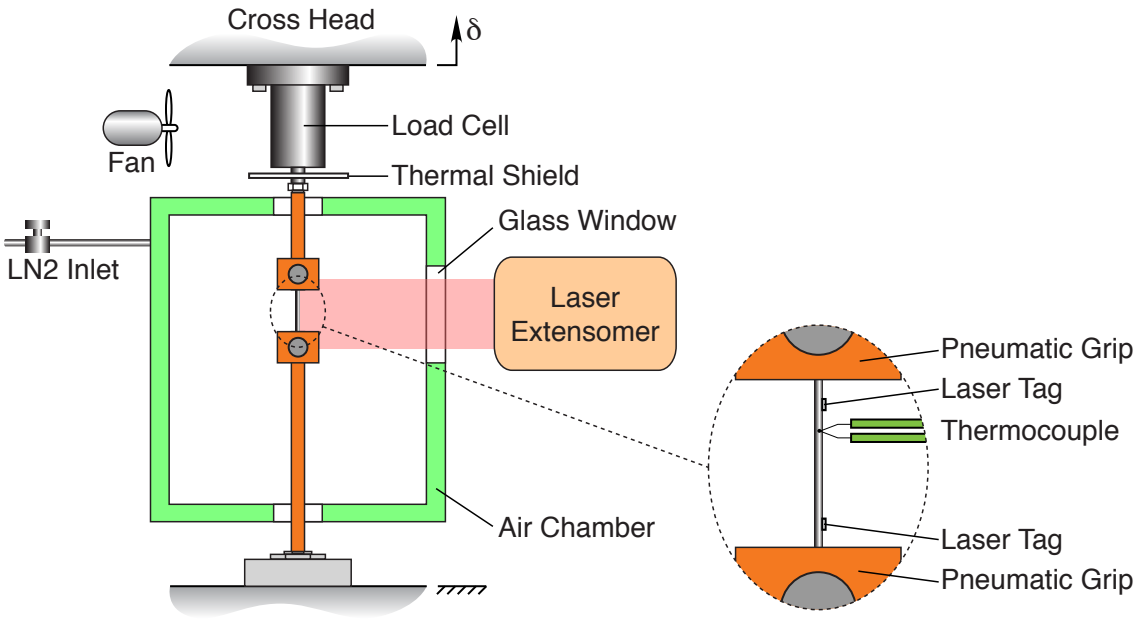


Figure 2.8: Setup for experiments on SE wire in an air chamber.

2.4.2 Specimen Materials

Reviewing information gathered from the DSC thermograms in Part 1, Table 2.2 shows stress-free transformation temperatures between the three phases: austenite (A), thermal-R-phase (R), and thermal-martensite (M). Specific heat and latent heat of transformation values for the three transformations $M \rightarrow A$ (Λ_{MA}), $A \rightarrow R$ (Λ_{AR}), and $R \rightarrow M$ (Λ_{RM}) are provided in Table 2.3. The SM wire had a diameter of 1.016 mm (0.040 in.), while the SE wire had a diameter of 0.762 mm (0.030 in.). The wire length between the grips (L) was about 69 mm and 127 mm for the respective alloys, yet each was measured accurately (by gage blocks and small crosshead motion after specimen installation) at the start of each experiment at a very small tensile load, just enough to remove any slack in the wire. All experiments shown below were performed on specimens from the same spool of wire,

and each experiment was performed on a different wire specimen to avoid introducing any unwanted history effects.

Table 2.2: Transformation temperatures as measured by DSC, with uncertainty of ± 2 °C [89].

Alloy	M_{f} (°C)	M_{s} (°C)	R_{f} (°C)	R_{s} (°C)	A_{s} (°C)	A_{f} (°C)
SM wire	< -50	-7	31	52	26	56
SE wire	-120	-73	-30	13	-30	13

Table 2.3: Specific heats (c_{o} ± 0.02 J/(g·K)) and latent heats of transformation (Λ ± 1 J/g) as measured by DSC [89].

Alloy	c_{o} (J/(g·K))	Λ_{MA} (J/g)	Λ_{AR} (J/g)	Λ_{RM} (J/g)
SM wire	0.45	19.7	-6.7	≈ -7
SE wire	0.45	15	-5.0	≈ -6

2.4.3 Shape Memory Wire Responses

A series of isothermal mechanical responses of the SM wire, one every 10 °C, is shown in Figure 2.9. The material was first tested in its "as received" state, so below 0 °C the initial material phase was a mixture of thermal-R-phase (R) and thermal-martensite (M). At these low temperatures, the response during loading is soft and nonlinear. The initial knee in the stress-strain curve before 1 % strain is due to reorientation of thermal-R to tensile-R (R^+). Subsequent to this initial knee the response briefly stiffens, but then flattens into a stress plateau as the material is transformed to M^+. The initial load-unload cycle was taken to a strain just beyond the end of the stress plateau, near 4 % to 5 %, leaving a residual strain upon unloading that could potentially be recovered (not shown) upon stress-free heating by the shape memory effect. For temperatures below and including 40 °C a

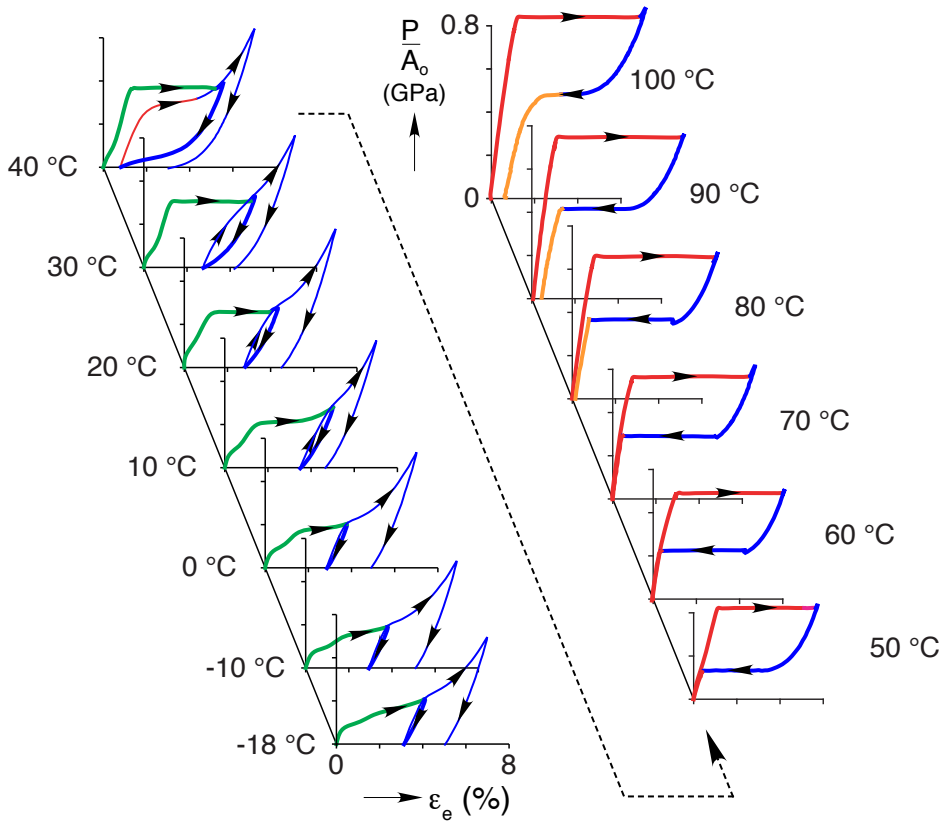


Figure 2.9: Isothermal mechanical responses of SM wire at selected temperatures. Data taken from [84, 86].

second load-unload cycle was taken to a maximum strain of about 7 % before unloading, showing the post-plateau response of M^+. As the temperature was raised in the sequence of experiments in Figure 2.9, one sees the loading plateaus ($\dot{\delta} > 0$) trend upward in stress, and the unloading responses ($\dot{\delta} < 0$) becoming more nonlinear. Note that above 20 °C the loading responses develop distinct stress plateaus, and these occur at progressively higher stresses as the temperature of the experiment is raised. A dramatic change in the unloading response occurs as the temperature reaches 40 °C, in which the residual strain is less than 1 % after the first load-unload cycle. This temperature happens to be near the average of A_s and A_f (see again Table 2.2).

At higher temperatures, 50 °C and above, the responses are superelastic. These consist of an initially stiff loading response associated with elastic loading of A, followed by a sudden tangent modulus change to zero, during which $A \rightarrow M^+$ transformation occurs along a stress plateau. A subsequent sharp upturn in the stress response occurs at the end of the plateau (see, for example, the 70 °C response), and elongation rate was reversed just beyond this point. Upon subsequent unloading, a nonlinear path is taken initially, but then another critical stress is reached where the elongation decreases along another stress plateau, during which the reverse $M^+ \rightarrow A$ transformation occurs. Once the elongation intersects the original austenite loading path, the stress takes a sharp downturn and unloading finishes along the austenite elastic curve. For all of the superelastic responses shown, unloading commenced just after exhaustion of the loading plateau to avoid introducing additional plasticity effects that would further complicate the interpretation of the response. Even so, one can see that perfect superelasticity no longer occurs for responses at 80 °C,

and above, where a progressively larger residual strain remains after unloading as the temperature is increased. In fact, the extent of the unloading stress plateau has nearly vanished at 100 °C.

For conventional metals the critical stress at the onset of significant nonlinearity would be thought of as a plastic *yield point*. Increasing the temperature normally softens the yield point of a conventional metal (although the effect is rather gradual until the temperature gets quite high). By contrast, the apparent "yield point" of an SMA increases dramatically as the temperature is raised. According to this view, the SMA material seems to get significantly stronger as the temperature is raised, contrary to conventional materials, which tends to catch the SMA novice by surprise. The underlying reason is that it is not a yield point associated with actual plasticity, but rather a transformation stress where the parent phase becomes unstable in favor of the daughter phase. As the temperature rises away from the cold temperatures where M^+ is naturally stable (stress-free), it requires progressively larger stresses to destabilize A, i.e. to stress-induce M^+. Furthermore, if the temperature continues to rise, the transformation stress will eventually exceed that which causes actual plastic slip.

Notice again that for temperatures between 50 °C and 70 °C perfect superelasticity occurs, but a small and then progressively larger amount of residual strain exists upon unloading at higher temperatures. This is caused by some amount of plastic deformation accompanying $A \rightarrow M^+$ transformation for transformation stresses between 650 MPa to 850 MPa in the experiments at 80 °C to 100 °C. Although not shown, stress-free heating can recover some of this residual strain, but not all, indicating actual unrecoverable

plasticity has occurred. This is not surprising, considering the stress reached high values during loading, where the required transformation stress has exceeded the true yield stress of the material. As the temperature increases further (not shown) progressively smaller portions of the residual strain at the end of unloading would be recovered by further stress-free heating. At excessively high temperatures, the unloading plateau vanishes altogether (becoming just nonlinear, akin to the 100 °C unloading response shown) and the permanent strain grows dramatically. Consequently, there is a higher temperature (well above 100 °C), called A_{d}, above which little stress-induced martensite is recovered upon unloading, and the material has a mechanical reponse similar to that of a conventional, ductile elasto-plastic material.

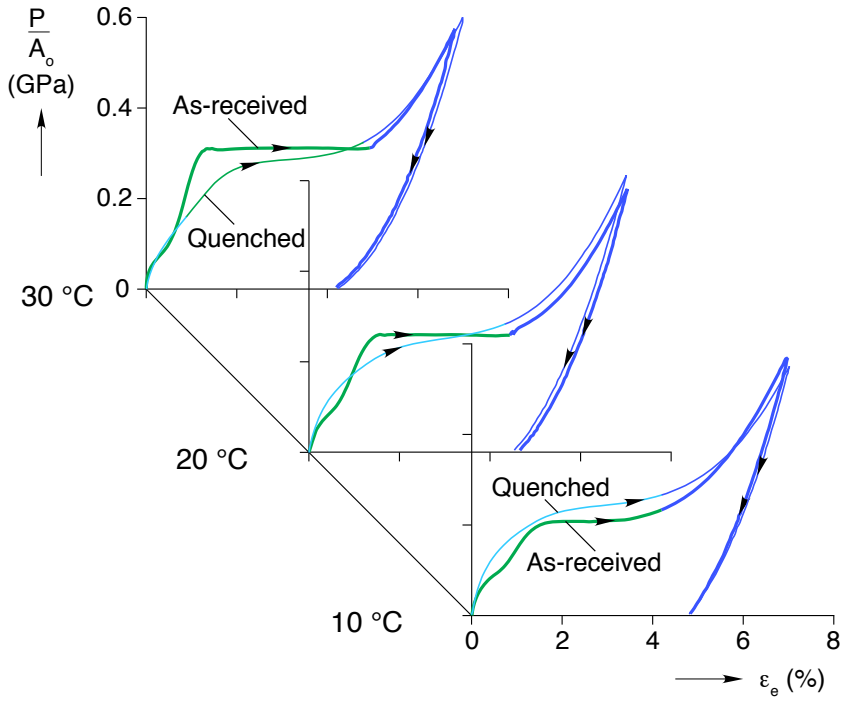


Figure 2.10: Comparision of liquid nitrogen-quenched material (thin line) and as-received material (thick line) responses of SM wire at intermediate temperatures. Data taken from [86].

As indicated by DSC, the stress-free phase of the material is not uniquely identified at certain intermediate temperatures. For example, Figure 2.10 shows three experiments between 10 °C and 30 °C that each compare mechanical responses for as-received specimens and specimens that were quenched in liquid nitrogen (-196 °C) before the start of the experiment. The as-received specimens were predominately thermal-R when the experiment was started (since they presumably were cooled to room temperature after the wire was processed at much high temperatures), whereas the quenched specimens were predominately thermal-M after being allowed to warm to their test temperature from -196 °C. A double knee is apparent in the first 1.5 % strain of the responses for the initially thermal-R material (as-received) as it undergoes a multi-step transformation $R \rightarrow R^+ \rightarrow M^+$, first reorienting thermal-R to tensile-R then transforming to tensile-M. The quenched material, on the other hand, exhibits a single knee associated with reorientation of martensite $M \rightarrow M^+$.

The different responses, even for experiments under identical thermomechanical conditions, again underscore the importance of knowing the prior material history before interpreting them. Nevertheless, once the specimen is taken beyond about 4 % strain the phase is largely M^+ and subsequent loading and unloading responses are quite similar between the as-received and quenched materials. Although not shown in Figure 2.10, additional experiments were performed on quenched wire over the entire temperature range considered, but little change in the response was observed at temperatures above 40 °C from the as-received responses shown in Figure 2.9.

A useful map of the material behavior is obtained by plotting the stress of the plateaus and their extents in strain as a function of temperature. Figure 2.11 depicts how plateau

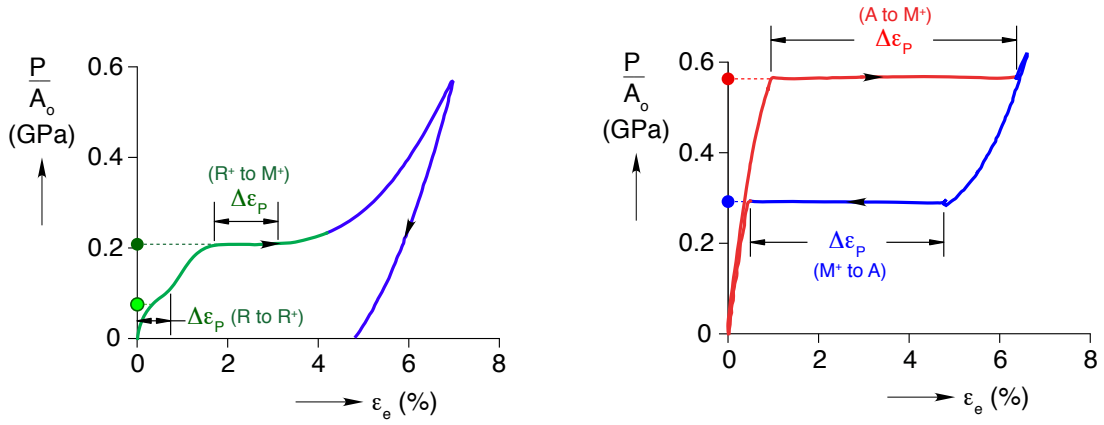


Figure 2.11: Example mechanical responses showing how characteristic plateau stresses and strains are extracted: (a) 10 °C experiment, (b) 70 °C experiment.

stresses (σ_{P}) and strain jumps ($\Delta\varepsilon_{\mathrm{P}}$) are extracted from the curves in Figure 2.10. Characteristic stresses are found from either a visual estimate of the knee of the curve or wherever distinct stress plateaus occur. Transformation strains are likewise obtained from the associated strain change that occurs during changes in microstructure (reorientation) or phase (length of stress plateaus). Strictly speaking, phase transformation does not occur exclusively during stress plateaus, i.e., some microstructural changes occur just before and after the plateaus and are responsible for most of the nonlinearity observed at either ends of the plateau. However, the strain change due to strict elasticity versus transformation in these portions of the response is difficult to distinguish precisely by macroscopic measurements, so only the transformation strain across the stress plateau will be reported.

Figure 2.12 plots characteristic stresses versus temperature for the SM wire starting in the as-received and quenched states. The DSC heating and cooling thermograms are shown along the stress-free axis for reference. While these transformations are hysteretic,

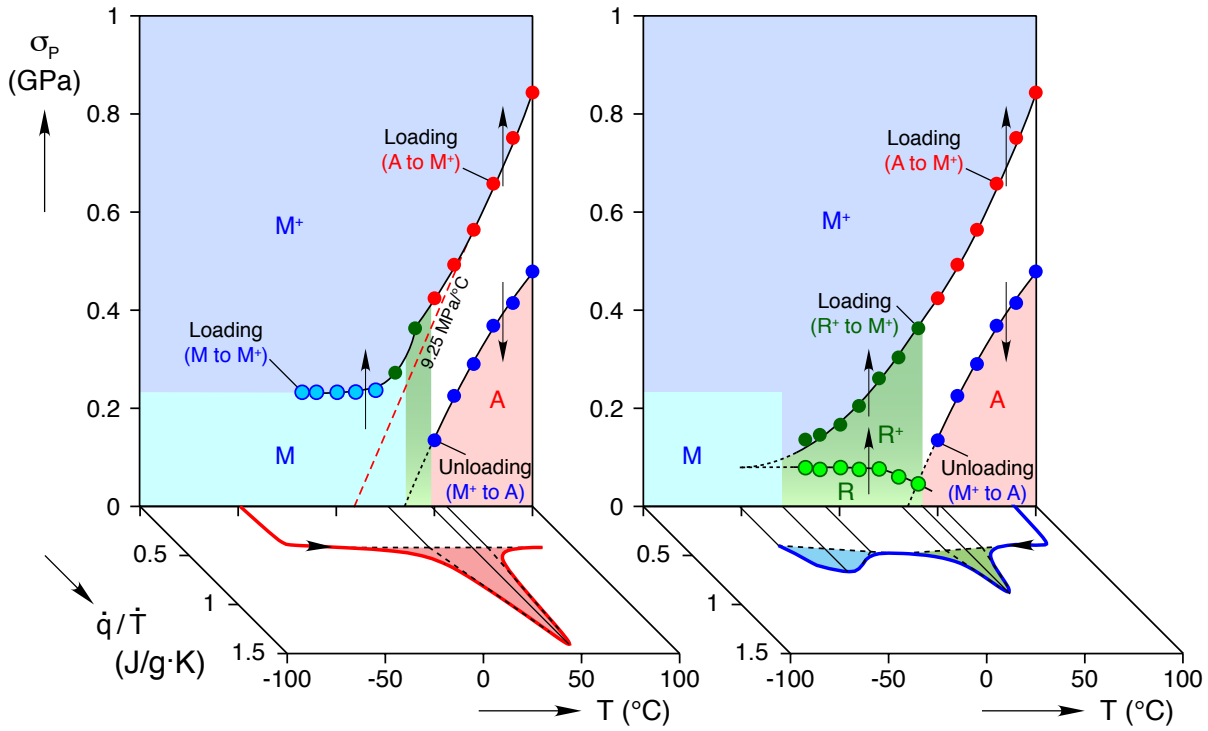


Figure 2.12: Characteristic transformation stresses of SM wire: initially quenched in liquid nitrogen (left), as-received (right). Respective heating and cooling DSC thermograms also shown along the stress-free axis for reference.

the plot can still be interpreted as a quasi-phase diagram in temperature-stress space. The plot on the left shows quenched wire that starts as M at temperatures below about 0 °C, while the right plot shows as-received wire starting as R at similarly low temperatures. The shaded areas depict the stable phase. In areas between the lines, the phase is dependent on loading history, so the vertical arrows show the direction that a phase boundary is crossed. Quenching has little to no effect in the high temperature, superelastic range, but does have an effect at intermediate to low temperatures. The plots also indicate that reorientation of $R \rightarrow R^+$ or $M \rightarrow M^+$ is relatively independent of temperature, at least for the temperature range shown. The transformation stresses for $R^+ \rightarrow M^+$ and $A \rightarrow M^+$, however, are extremely temperature dependent, rising from about 150 MPa to 850 MPa (as-received material), nearly a factor of six, over the 120 °C temperature range of the data shown. The slope for $A \rightarrow M^+$ transformation is about 9.25 MPa / °C at the high end of the temperature range, for both quenched and as-received wires.

A summary of transformation strains are presented in a similar manner in Figure 2.13. The trend for each type of transformation is quite nonlinear, tracing a quadratic-like curve in each case. The $R \rightarrow R^+$ strain is the exception, once again being relatively insensitive to temperature (although these were estimated by "eye" from the response curves). Both $M^+ \rightarrow A$ and $A \rightarrow M^+$ strains increase with temperature until they reach maxima at 70 °C and 90 °C, respectively. At the highest temperature (100 °C), increasing plasticity causes the $M^+ \rightarrow A$ plateau strain to fall below 2 %, while $A \rightarrow M^+$ just begins to decrease from its maximum of about 5.8 % strain.

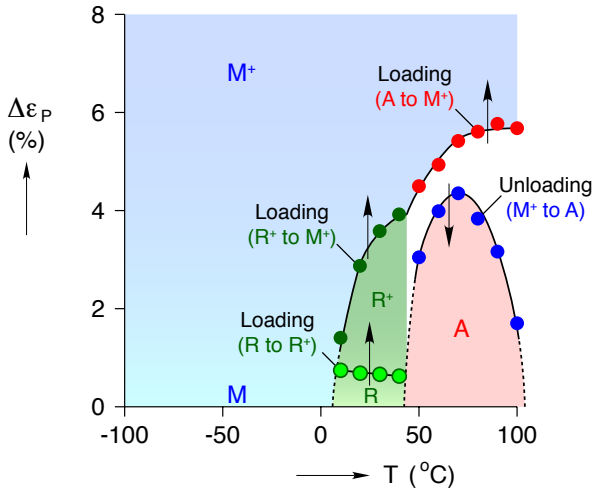


Figure 2.13: Characteristic transformation strains of SM wire (as-received).

2.4.4 Superelastic Wire Responses

For comparison, the same type of data presented above is now provided for the SE wire. Isothermal mechanical responses of the as-received SE wire are shown in Figure 2.14, but the experiments were performed in an air chamber over a temperature range of -50 °C to 70 °C. The responses have similar trends to that of the SM wire of Figure 2.9, but shifted to lower temperatures. Similar to the SM wire, the lowest temperature responses of the SE wire have low plateau stresses and a residual strain upon unloading. Superelasticity is exhibited in the responses above 0 °C, but imperfect superelasticity with a residual strain after unloading occurs in responses at 40 °C and above. Plots of transformation strains and stresses are shown in Figs. 2.15 and 2.16, respectively (including data from a few experiments not shown in Figure 2.14), which bring out a few quantitative differences with the SM wire. Most noticeable is the lower $M \rightarrow M^+$ transformation stress for the SE wire, less than half that of the SM wire in Figure 2.12. The slope of the $A \rightarrow M^+$ transformation

for the SE wire is about 6.7 MPa/°C as shown in Figure 2.15, less than the corresponding 9.25 MPa/°C of the SM wire. Figure 2.16 shows the SE wire $A \rightarrow M^+$ transformation strain continuing to rise in the temperature range tested, while the corresponding transformation strain of the SM wire reaches a peak value and begins decreasing at the highest temperature in Figure 2.13. No data points are given for the $M^+ \rightarrow A$ transformation for temperatures above 60 °C in Figs. 2.15 and 2.16, since the unloading plateau (see Figure 2.14) was too nonlinear to extract distinct values for stress or transformation strain. Also, the maximum plateau strain measured in the SE wire is about 7.4 %, significantly larger than the SM wire maximum value of about 5.8 %.

2.5 Material Instability and Transformation Fronts

During uniaxial stress-induced transformation, many NiTi alloys have (now) well known material level instabilities [86, 87] that give rise to localized deformation and propagating transformation fronts. As an example, Figure 2.17 shows detailed measurements at two different temperatures (70 °C and 30 °C) for the SM wire. On the left, the mechanical responses are shown with axial strain measured by a miniature mechanical extensometer (ε_e), and on the right, the extensometer response (local strain) is compared to the grip motion, normalized as δ/L (global strain). During testing of conventional stable materials, one would expect the two strain measurements (local and global) to agree, at least approximately, and simply trace a 45° line in the plot (dashed line). Focusing first on the 30 °C response in Figure 2.17a, this is what occurs initially during loading between states ⓪ and ① during $R \rightarrow R^+$. During the stress plateau between ① and ② ($R^+ \rightarrow M^+$), how-

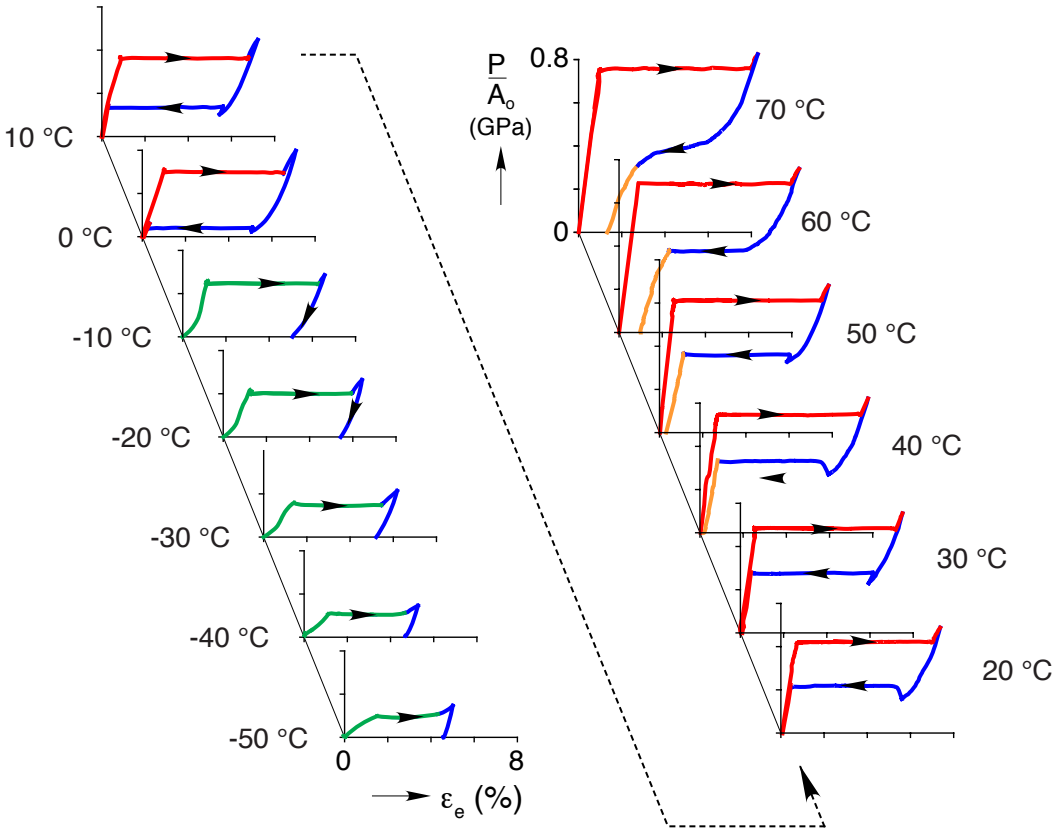


Figure 2.14: Isothermal mechanical responses of as-received SE wire at selected temperatures. Data taken from [47] and [15].

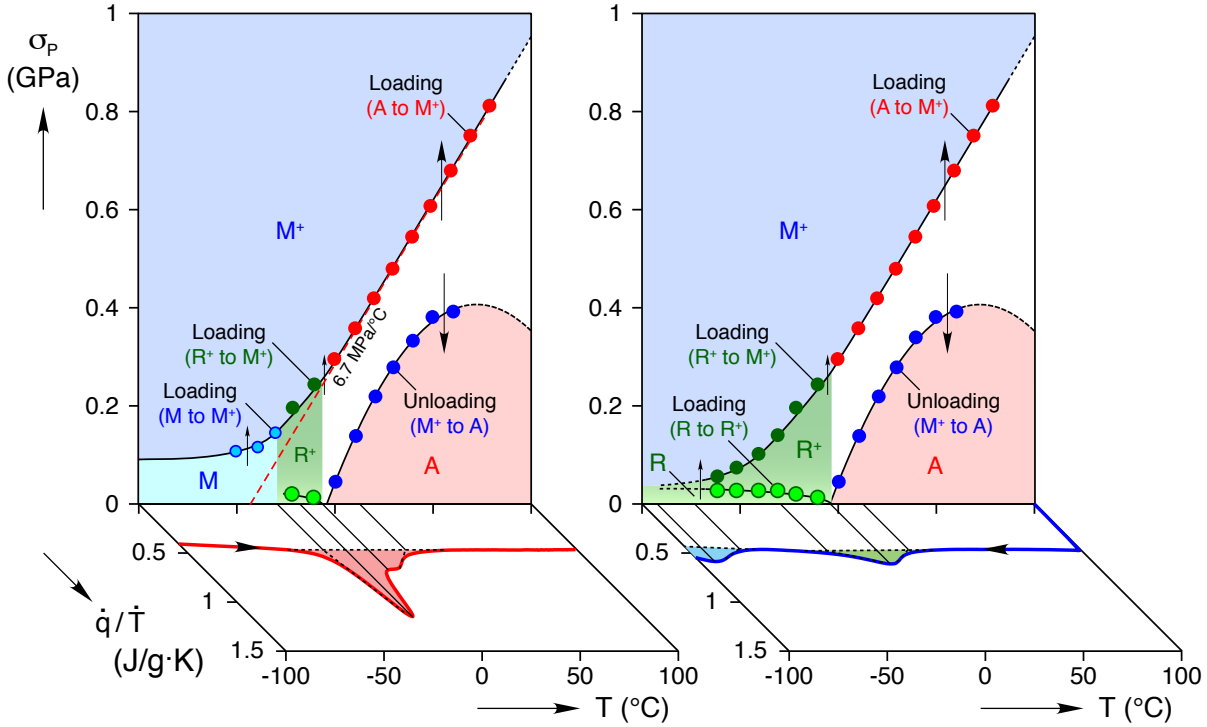


Figure 2.15: Characteristic transformation stresses of SE wire: initially quenched in liquid nitrogen (left), as-received (right). Respective heating and cooling DSC thermograms also shown along the stress-free axis for reference.

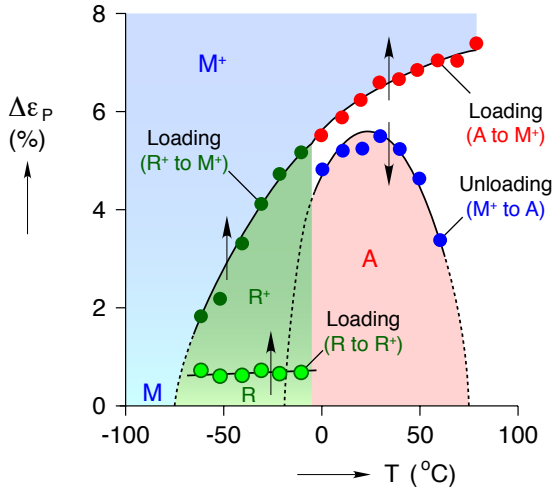


Figure 2.16: Characteristic transformation strains of SE wire (as-received).

ever, the extensometer records a more rapid increase in strain (ε_e) compared to the global strain (δ/L), but then between (2) and (3) the local strain stays static while the global strain increases (and the point does not move in the force vs. local strain response). Loading continues briefly between (3) and (4), where an upturn in the force response is measured and the local vs. global strain trace again along a 45° (but slightly shifted) line. At (4) the direction of the crosshead is reversed and the specimen unloads along a nonlinear path in the force response, and the local vs. global strain traces backward along the 45° line until the force is zero at point (5). One also notices the local strain diverges somewhat from the global strain near point (1), but this is an artifact of the experiment where some grip slippage has occurred (hence, the offset between the two 45° dashed lines). This is caused by stress concentrations and localized axial straining (and thus lateral thinning of the wire) at the grips, creating a small amount of "inch-worming" of the wire out of the grips (about 0.5 % global strain) that is difficult to avoid in this setup. Thus, a global force-elongation response (not shown) would erroneously show a larger permanent strain than actually occurs in the material.

The stress plateau during $R^+ \rightarrow M^+$ has a zero tangent modulus and is associated with local/global strain responses that diverge at the start of transformation (1) then re-converge at the end of transformation (3). In this case, transformation first localizes within the gage section of the extensometer (and a rapid increase in local strain is measured) and then deformation progresses outside the extensometer's gage section (and no change is recorded by the extensometer) while overall elongation continues until the specimen has been completely consumed by the new phase (M^+). Conversely, one observes that wherever the

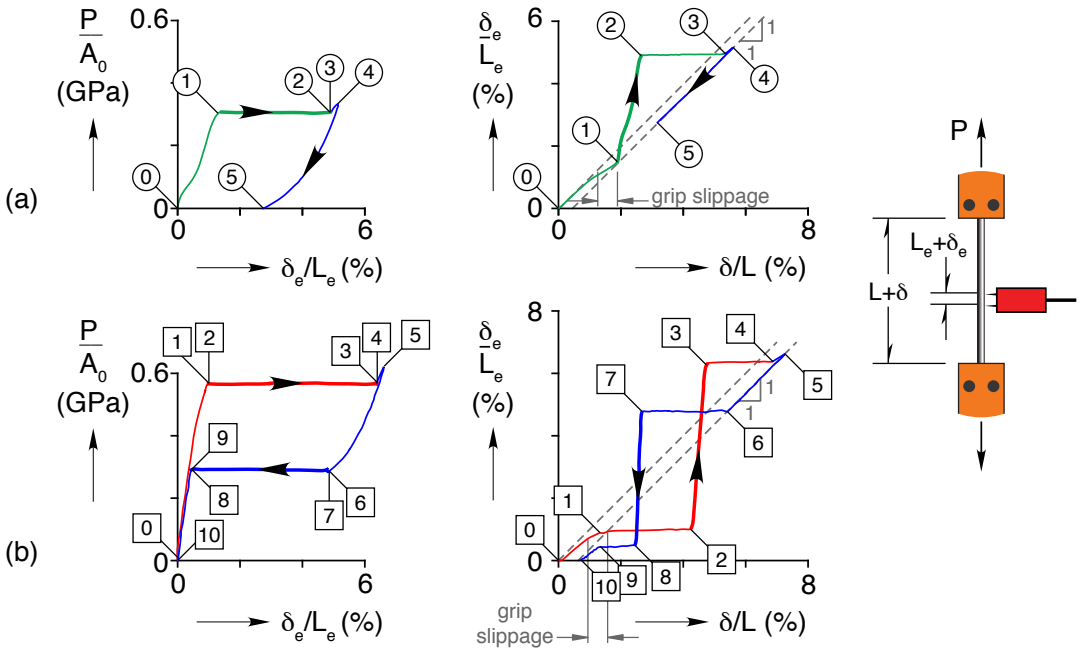


Figure 2.17: Axial stress vs. local-strain responses (left) and measured local vs. global strain responses (right) in SM wire during two experiments: (a) 30 °C, (b) 70 °C.

force-elongation response maintains a positive tangent modulus the local and global strain measurements follow a 45° line, meaning the whole specimen undergoes uniform deformation. Accordingly, reorienting of $R \rightarrow R^+$ during loading and re-twinning of M^+ during unloading (responsible for the nonlinear path) are stable processes in this alloy.

Figure 2.17b shows similar phenomena in the superelastic response of the SM wire at 70 °C, but now localization occurs during both loading ($A \rightarrow M^+$ between 1 and 4) and unloading ($M^+ \rightarrow A$ between 6 and 9). During both transformations, the local strain first stays static, then jumps, then stays static again. This is caused by transformation occurring by the propagation of one or more boundaries separating nearly uniform high strain and low strain regions. The axial extent of such a *transformation front* is only a few wire radii, where a steep gradient in strain exists, thus it can be considered a propagating

neck in the specimen. The extensometer, therefore, will only record a change in strain when a transformation front passes through its measurement region. The specific times at which the strain jumps are recorded depend on the location of the extensometer and the particular motion of transformation fronts.

Such localization and propagation phenomena, while uncommon in materials testing, certainly do exist in other materials but for different microscopic reasons. Examples include mild steel that exhibits Lüders bands during uniaxial testing [1, 47, 58] and high density polyethylene [19, 99] that exhibits large strain propagating necks during tensile cold-drawing. Figure 2.18 shows a schematic of the common underlying cause of these phenomena. The material constitutive law (local stress-strain) has an up-down-up character (Fig 2.18a), which gives rise to the force-elongation, or "structural", response (Fig 2.18b) under displacement-control, consisting of the following sequence: (1) an initial rising curve, (2) a nucleation peak initiated at σ_{N} with a drop in load to σ_{P} at constant elongation, (3) a force plateau at σ_{P} with increasing elongation and propagating fronts, (4) another drop in load again at constant elongation as the strain delocalizes, and (5) a rising curve after the plateau elongation has been exhausted and specimen has been completely converted to the high strain phase. The negative slope in the local constitutive law represents an unstable segment that is energetically unfavorable unless the material is otherwise constrained, as in the neck region by material on either side. At the onset of instability a finite portion of the wire transforms in a dynamic manner (2), creating a high strain inclusion somewhere in the wire length. Since the total elongation is being controlled, i.e. fixed during this instant, the remaining material must unload, so the load drops. During subsequent propagation of

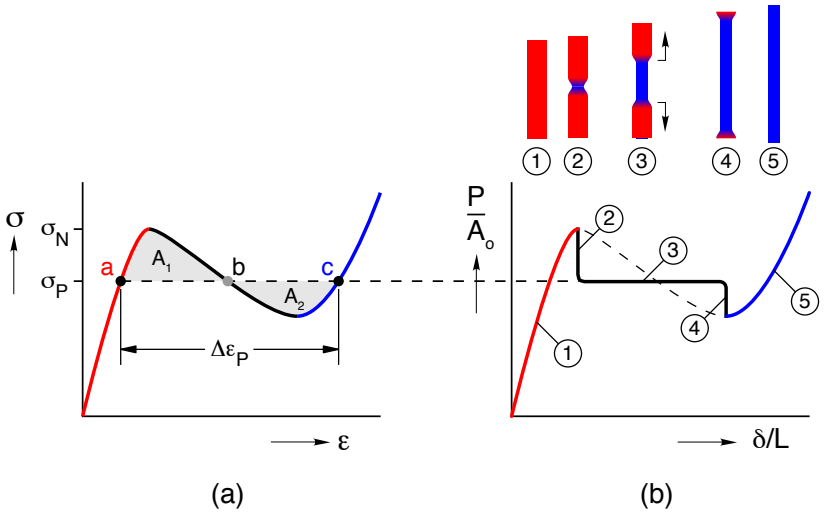


Figure 2.18: Schematics of (a) local stress-strain response, and (b) global force-elongation ("structural") response with specimen configurations before, during, and after $A \rightarrow M^+$ transformation.

fronts away from the nucleation site, equilibrium requires that the force be constant along the length, so the material can exist in states (a), (b), or (c) as shown in the figure, but only states (a) and (c) are stable. As such, the unstable portion of the constitutive law is difficult to measure in practice, and must be inferred from the global structural response as follows. The plateau stress (σ_{P}) and plateau strain ($\Delta\varepsilon_{\mathrm{P}}$) can be found by the construction shown in the figure where the shaded areas are equal ($\mathrm{A}_1 = \mathrm{A}_2$). The theory is based on nonlinear elasticity [28] but gives a reasonable approximation even when dissipation/hysteresis is involved. In the hysteretic case, one imagines two underlying local stress-strain curves, an upper one for loading ($\dot{\delta} > 0$) and a lower one for unloading ($\dot{\delta} < 0$), each having an up-down-up character, thereby giving rise to two stress plateaus [15].

Nucleation peaks, however, are absent in the experiments of Figure 2.17 at the onset of transformation during loading. Unless the experiment is carefully designed, stress con-

centrations at the grips will usually suppress a portion of, or the entire, load peak that accompanies the nucleation of a M^+ region, thereby giving the appearance of an up-flat-up response during loading. The measured nucleation stress σ_{N} is imperfection sensitive, i.e. will be reduced by geometric or loading imperfections from a theoretical perfect upper bound. In this case, the multiaxial stress concentrations at the grip entrances act as starters for $R^+ \rightarrow M^+$ or $A \rightarrow M^+$ transformation and are large enough to effectively mask the nucleation stress peak. If one looks closely, however, at the unloading responses at 60 °C to 80 °C in Figure 2.9 and the unloading responses between 0 °C and 50 °C in Figure 2.14, nucleation peaks can be seen at the onset of the $M^+ \rightarrow A$ transformation plateaus. The nucleation peaks upon unloading are upside down, similar to state (4) of Figure 2.18, since the strain is decreasing from a large value. These unloading peaks are apparent, since the onset of $M^+ \rightarrow A$ transformation began somewhere away from the grips, since stress concentrations do not favor this reverse transformation.

Interestingly, analogous localization and propagation phenomena also exist in certain structural problems, including for example, crushing of honeycombs and foams with localized densification bands [37, 75], external pressure of long pipes with propagating buckles [26], and inflation of rubber tubes with propagating bulges [57]. The common feature is that the local structural response becomes unstable, but then stiffens at larger strains due to material stiffening or internal contact mechanisms, i.e. the "local" response (cell or cross-section) has an up-down-up character [55]. In these cases material points are locally stable, but instability occurs by geometric softening. By contrast, SMA wire instabilities are the result of both geometric effects (wire thinning) and local material instability (phase

transformation at the microscale).

2.6 Nucleation Experimental Setup

Much of the game in developing a good understanding of SMA behavior is to understand the kinetics of transformation between the various phases. Since uniaxial tension of NiTi wire often exhibits phase transformation through localized transformation and then quasi-static propagation of one, or more, distinct phase boundaries (fronts) as the specimen elongates, the axial strain and temperature fields that are rather discontinuous both temporally and spatially which presents some unique challenges for testing and interpretation of results. On the other hand, such inhomogeneous fields can be usefully exploited as markers to track where transformation is actively occurring, using a suitable experimental setup like the one described below.

Figure 2.19 shows a special experimental setup designed to capture localization in SMA wire during uniaxial tension experiments [49]. The wire is held at each end between two hardened steel plates, kept in alignment by a small groove machined in one of the plates. The plates are held by pneumatic grips (to maintain constant gripping force) in an electromechanical load frame, which provides displacement control of the upper grip and load measurement (Instron 5 kN load cell). Although not used in the experiments presented below, two reflective targets could be epoxied to the wire specimen, delimiting a gauge length for strain measurement. The distance between targets would be measured by a noncontacting laser extensometer (model EIC-05 from Electronics Research Corp., Irwin, PA), having a resolution of 2 microns). When used, this method avoided introducing artificial

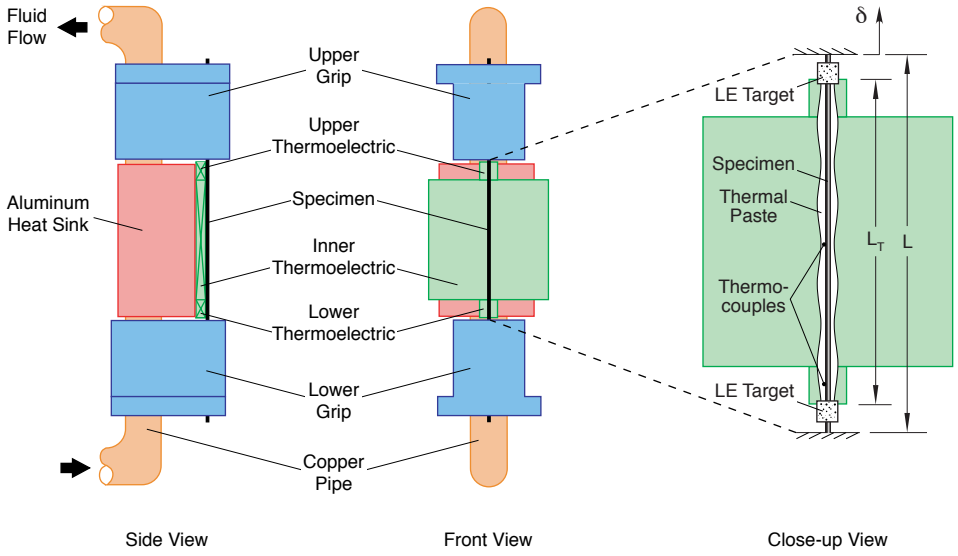


Figure 2.19: Temperature control apparatus for nucleation peak measurements.

loads or bending on the specimen that a mechanical extensometer might cause and minimized obscuring the front surface of the specimen for imaging. A relatively large aluminum heat sink is suspended behind the wire. Fluid from a temperature-controlled bath (NESLAB model RTE-140) flows through both the heat sink and the two grips, regulating both at a user prescribed temperature.

Fluid from a thermal bath is pumped through both the wire grips and an aluminum heat sink placed behind the wire, holding them at a prescribed temperature. Between the heat sink and the backside of the wire specimen sit three thermoelectric (TE) wafers: a single $40 \times 40 \times 3.3$ mm inner TE (Melcor UT8-12-40-F1-W6, Laird Technologies, St. Louis, MO) and two $6 \times 6 \times 3.4$ mm top and lower TEs (Melcor CP0.8-7-06L) that are controlled from two independent electrical circuits. These TE modules are solid-state devices that use the Peltier effect of semiconductor elements within them [32] to pump heat to, or from, the heat sink contacting their back faces. By applying positive (or negative) DC voltage to their

terminals a temperature gradient is created across the thickness, heating (or cooling) their front faces, which are in thermal contact with the specimen. By controlling the voltage of each TE one can control the temperature field along the wire, independently of the grip temperature. In order to facilitate heat transfer between the TEs and wire, a small amount of thermally conductive paste (Omegatherm 201, $k = 2.26$ W/K/m, from Omega Engineering, Stamford, CT) is applied between them, which allows the specimen to slide freely during mechanical straining. Besides allowing precise temperature control of the specimen, this arrangement also leaves the front surface of the specimen free for optical or infrared imaging.

If the specimen is relatively thin and thermally conductive, which is usually the case, temperature gradients across the wire diameter can be reasonably neglected, and the outer surface temperature can be assumed to be representative of the interior temperature within the cross-section. This assumption can be confirmed by calculating the Biot number (Bi) across the wire diameter, the dimensionless ratio of the conductive thermal resistance to the convective thermal resistance, and verifying $Bi \ll 1$. In this case, it is a quite good assumption, since $Bi = h\,d/k = 5.4 \times 10^{-4}$, where h is the convective film coefficient of air (about $120\mathrm{W/m^2K}$ [g]), d is the wire diameter ($7.6 \times 10^{-4}\mathrm{m}$), and k is the thermal conductivity of austenite, reported by the manufacturer as $18\ \mathrm{W/(m\,K)}$.

The full temperature field was captured using an Inframetrics (now FLIR, Boston, MA) model SC1000 digital infrared radiometer with a PtSi 256×256 detector array. The accu-

[g]The film coefficient h is a linear approximation of the convective heat transfer between the wire and ambient media, which is influenced by surface coatings, the air velocity, and even horizontal or vertical orientation of the wire.

racy of the temperature depends on the emissivity ($0 \leq \epsilon \leq 1$) of the specimen surface and the transmissivity of the medium between the target surface and the detector. The best results are obtained for an ideal black-body surface with $\epsilon = 1$ in a vacuum. A few inches of ambient air, however, is still quite transmissive in the IR range. Note that glass and most fluids are not transmissive, so IR imaging cannot be done for specimens submerged in a fluid bath or behind the glass window of an air chamber. Nitinol can be obtained with various surface finishes from heavily oxidized (with high emissivity) to shiny (with low emissivity), so it must be calibrated for each wire finish. In this case, emissivity was measured at $\epsilon = 0.66$ for the lightly oxidized (bluish tint) wire. Since the reflectivity of a non-transmissive IR surface is $r = 1 - \epsilon$, this means that background radiation will reflect off the specimen surface and can give false readings. In particular, it is a good practice to shield the specimen from the operator's body heat, or anyone else in the room, which can be noticeably detected by the IR camera. A sheet of cardboard is usually sufficient to act as an IR shield. In addition, one must be careful using a red laser extensometer shined directly on the specimen, since this may also be detected by the IR radiometer if it is sensitive to a broad range of IR frequencies that includes some of the low frequency (red) visible spectrum. In this case, the laser sheet can be offset from the crown of the wire, yet still catch laser targets if they extend laterally at bit beyond the specimen axis, to avoid interfering with IR imaging.

A confirming measurement of temperature is obtained from two small (0.08 mm diameter) type K exposed-junction thermocouples placed against the wire, similarly immersed in thermal paste. These were connected via thermocouple signal conditioning modules (Fluke

80 TK) to the data acquisition system and each thermocouple/module combination was calibrated to ASTM certified glass thermometers at two temperatures spanning the range of interest.

There are a few other issues to be aware of to make this setup work well. The uniformity of the temperature across the thermoelectric's front face is dependent on uniform thermal contact with the block behind it. Initially, a liquid adhesive was used, which required some skill to achieve a uniform fit (and was used in the experiment of Section 2.8). More recent efforts use a conductive adhesive film (Thermattach T412 by Chomerics, Woburn, MA), which is easier to use. Another issue is that the heatsink assembly must be precisely aligned with the wire, not only to avoid bending or kinking the wire but to ensure a uniform contact between wire and TE across the entire gauge length. This can be facilitated by mounting the entire heat sink/TE assembly on an adjustable tilt/translation stages (like the Newport 426).

2.7 Transformation Fronts

Figure 2.20 shows a load-unload experiment at $\dot{\delta}/L = \pm 1 \times 10^{-3}$ s^{-1} on the SE wire with the heat sink and grips held at 24.4 °C, and the TEs left in open circuit (passive). The specimen material used is the same superelastic (SE) 0.762 mm Nitinol wire (Memry Corp., Bethel, CT) used in Parts 1 and 2 [17, 89]. This rate was chosen such that small, yet measurable, latent heat effects could be used to track phase transformation in the gauge length of the specimen. This experiment will be used as a baseline comparison with another experiment in the next section that uses the TEs in an active mode. The mechanical

response to a single load-unload cycle is shown in Figure 2.20(a) as axial stress vs. global strain (δ/L, derived from the motion of the upper grip) with numbered tags to enable synchronization with optical and IR images as shown in Figure 2.20(b) and 2.20(c).

Figure 2.20(b) shows two streak plots, one during loading ($\dot{\delta} > 0$) and one during unloading ($\dot{\delta} < 0$). Each plot was obtained by capturing optical images of the specimen front with a digital imaging camera (Princeton Instruments RTE/CCD-1300-Y/HS 1300×1030 detector array, with a MicroMax 5 MHz ST-133 controller) at 1.3 s intervals and laying them side-by-side in time. Each vertical column of pixels corresponds to a narrow grayscale image of the specimen, where the variation of surface reflectivity has been processed to span black to white to augment the reflectivity variation. These were normalized by the initial free length of the wire (x/L) and scaled to the times they were taken on the horizontal axis. The uppermost dotted line shows the motion of the upper grip in time, so loading began at time 0 and ended somewhat after 80 s near point (6), and then time was reset to 0 at the start of unloading (horizontal axis on the right). The edges of the TEs are depicted by the inner dotted lines, which show that small air gaps existed at either end (to enable assembly) and the TEs followed the upper grip during the experiment, first upward then downward. The discontinuities in streaks with time (and the bright pixels at the interface) show the motion of transformation fronts. In fact, transformation fronts are propagating necks separating low strain regions from high strain regions and can even be seen with the naked eye (albeit with acute vision and favorable lighting conditions) as slight glints along the wire axis. The left plot shows that during loading two $A \rightarrow M^+$ fronts start at the gripped ends and propagate towards each other, coalescing

about two-thirds up the specimen length just before (6) where a sharp upturn is seen in the mechanical response of Figure 2.20(a) at the end of the loading plateau. During unloading, conversely, an $M^+ \rightarrow A$ nucleation event occurs between (7) and (8), creating two reverse transformation fronts that propagated away from each other until they reach the top and bottom ends.

At the same time IR images were taken at 1.6 s intervals during the experiment and similar contour maps, but in temperature, are shown in Figure 2.20(c) for loading and unloading. These are made up of a large number of IR images, scaled to x/L and time, with the color calibrated according to the temperature legend on the right. Note that the apparently cold ends (above and below the central TE region) were caused by IR reflection off the inside surface of the grips, which had a different emissivity that the wire specimen, therefore the measured temperature of those regions was not accurate and should be disregarded. In this case fronts show up as traveling hot spots during loading, due to the exothermic release of latent heat of the $A \rightarrow M^+$ transformation, then fronts show up as cold spots during unloading, due to the endothermic latent heat absorption of the reverse $M^+ \rightarrow A$ transformation. Recalling the differential scanning thermograms of Part 1 (Figure 3(b)), latent heats of transformation were measured during temperature-induced transformation [89]. Here, there are similar thermal exchanges with the environment during stress-induced transformation.

During loading, the hot spots are initially about 3.5 °C above ambient temperature as they cross the air gaps at the ends, which results in the load peak seen in Figure 2.20(a) between (2) and (3). As the fronts enter the TE section they settle to slightly less than 2 °C

above ambient temperature, but then become hotter as the two fronts near each other, to about 3.5 °C above ambient. During unloading, the first nucleation of $M^+ \rightarrow A$ shows up as a sudden cold spot by about -4.5 °C at the material location where fronts had met during loading (this is often the case, probably due to the complex stress state caused by dynamic coalescence event that leaves some residual M^+). Then fronts quickly settle to about -2.5 °C cold spots for most of their traversal in the specimen. As the fronts traverse the air gaps near the ends of the specimen they become colder yet, which is responsible for the dip in the unloading plateau after (10) in Figure 2.20(a).

The mechanical response is generally similar to those of Figure 2.14, but with some differences stemming from the different loading rate and thermal environment. The plateaus are not as flat and a residual strain of about 0.55 % is apparent after unloading. The plateaus are wavy due to the temperature fluctuations just mentioned, since the SMA transformation stresses are extremely temperature sensitive. The two bumps just after (1) are due to the entry of $A \rightarrow M^+$ from the grips (where significant stress concentrations exist due to clamping of the wire) into the free length. Thus, some grip slippage occurred here, which is responsible for most of the residual strain measured, and it should not be misinterpreted as damage in the specimen. As mentioned before in Section 2.3.2, it is important to use both local (extensometer) and global (grip displacement) strain measurement to remove grip slippage artifacts inherent in the global strain measurement.

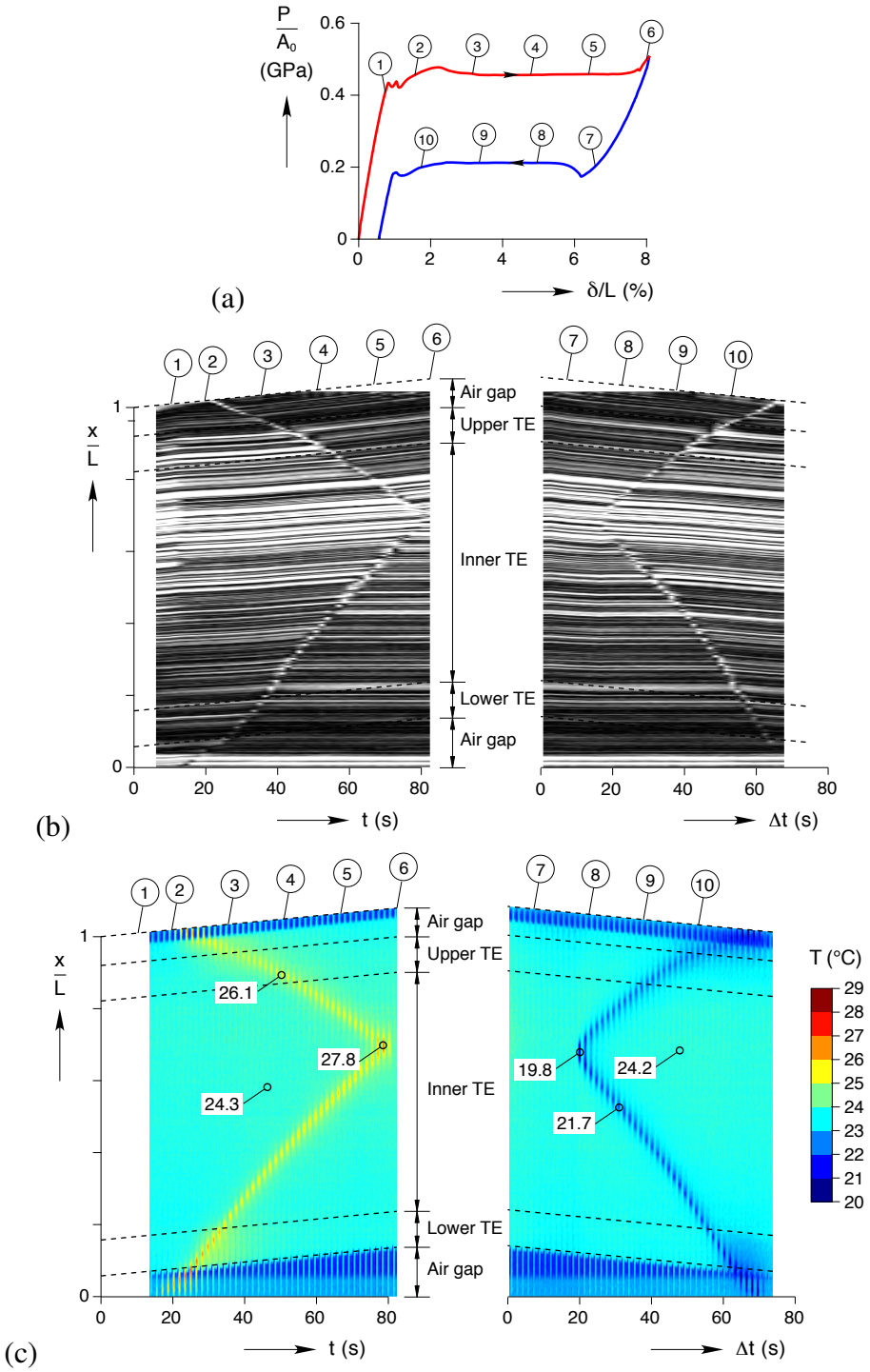


Figure 2.20: Experiment at 24.4 °C at $\dot{\delta}/L = \pm 1 \times 10^{-3}$ s^{-1} using the setup of Figure 2.19 with passive thermoelectrics: (a) mechanical response, (b) streak plot from a sequence of optical images, and (c) IR measured temperature history.

2.8 A Nucleation Experiment

The following experiment was designed to create favorable conditions for nucleation within the gauge section for both $A \rightarrow M^+$ and $M^+ \rightarrow A$ during a single load-unload cycle, and to measure the stress peaks caused by each. Figure 2.21 shows the results of the experiment, performed with the upper grip moving under displacement control at the slow rate of $\dot{\delta}/L = \pm 1 \times 10^{-4}\ \mathrm{s}^{-1}$. Figure 2.21(a) shows the axial stress response vs. global elongation (δ/L). Figure 2.21(b) shows the temperature profile (horizontal axis, T) along the wire's normalized length (vertical axis, x/L) at selected times during the experiment. Figure 2.21(c) shows a streak plot taken from a sequence of photographs of the wire specimen during loading and unloading transformation, where again, kink discontinuities in time indicate the passage of transformation fronts. Thin lines have been overlaid for clarity.

Initially, the center and outer thermoelectrics held the gauge length of the wire at 24 °C while the grips were held at 63 °C. See the bowed temperature profile (b) in Figure 2.21(b), which was essentially the same as at the start of the experiment. Since the ends of the wire were much hotter (by 39 °C), the local $A \rightarrow M^+$ transformation stress was higher by about 260 MPa ($= 6.7$ MPa/°C $\times$ 39 °C), according to Figure 9 of [17], suppressing $A \rightarrow M^+$ transformation there. The center TE section was cooler, so the nucleation stress would be reached there first, despite the extra stress concentrations at the grips. The inner TE section of the gauge length was held at a relatively uniform temperature, and the wire cross-section was quite uniform, thereby minimizing imperfections that would otherwise cause early nucleation. This trick effectively created a "dogbone" specimen with a "weaker" yet

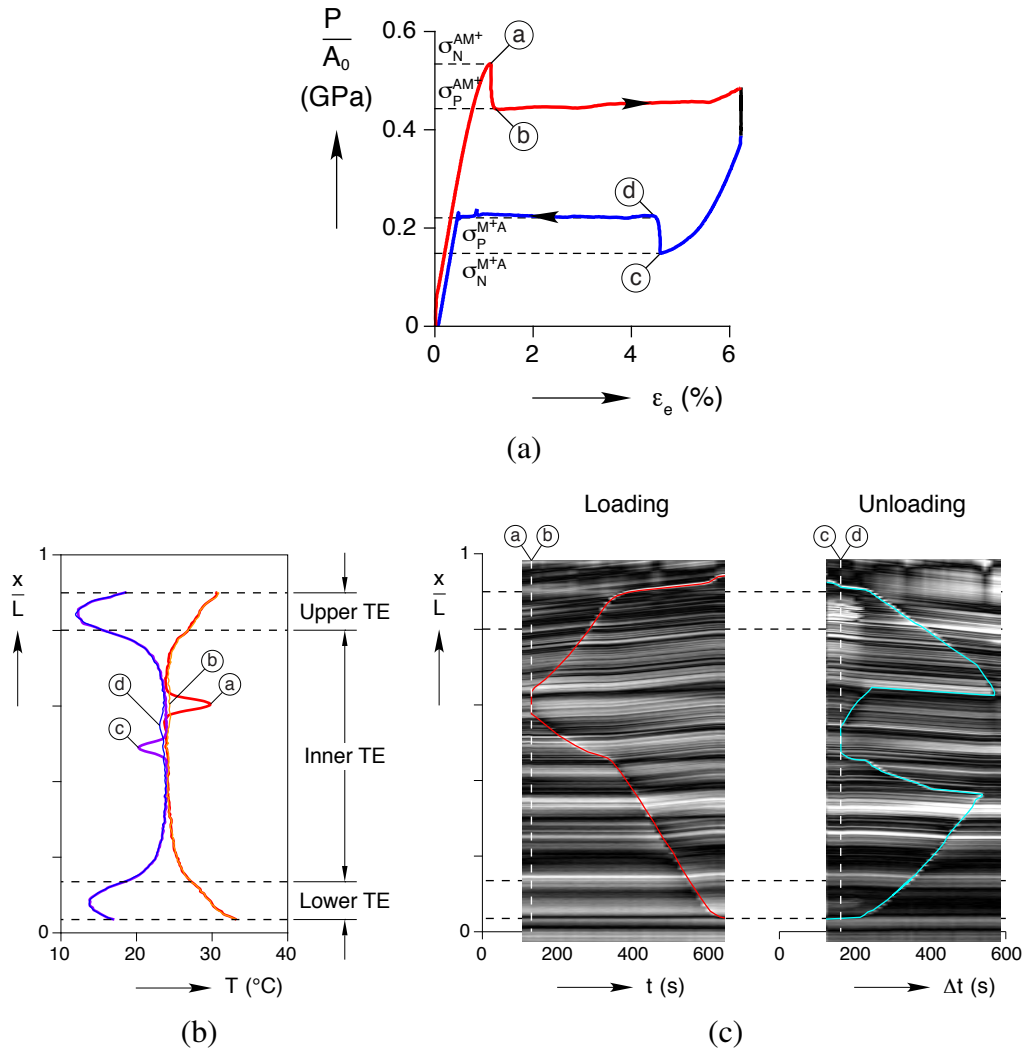


Figure 2.21: Nucleation peak experiment at 24 °C at $\dot{\delta}/L = \pm 1 \times 10^{-4}\ \mathrm{s}^{-1}$ using the setup of Figure 2.19, showing the onset of $A \rightarrow M^+$ during loading (a-b) and $M^+ \rightarrow A$ during unloading (c-d): (a) mechanical response, (b) IR measured temperature profiles during and just after nucleation events, (c) streak plot from a sequence of optical images, with lines overlaid for clarity.

uniform gage length by exploiting the temperature dependence of the transformation stress.

From the start of the experiment in Figure 2.21(a) the wire stress increases monotonically but has a growing nonlinearity (indicating some early uniform $A \rightarrow M^+$ transformation above about 450 MPa) until (a). At this point the cooler center region reaches its nucleation stress and a small pocket of martensite begins to form near $x/L = 0.61$ with an axial extent of a few wire diameters. The nucleation of M^+ within the (mostly) A-phase is a dynamic event with a sudden load drop between times (a) and (b) in Figure 2.21(a), even under elongation-controlled loading. The symmetry of uniformly deformed wire is broken, indicating a bifurcation of the equilibrium path. Since the pocket forms in a region of the gage section with uniform stress and temperature (so-called homogeneous nucleation), the change of axial strain in the pocket is accommodated by a reduction of elastic strain in the remainder of the wire specimen, since the global elongation has not changed appreciably during this small time interval (less than 1 s). In addition, equilibrium of the axial force along the wire is reached at (b). Under end-displacement control, axial compatibility of the pre-transformed and post-nucleated states causes the stress to drop suddenly, and can even be heard as an audible "pop" from the specimen as nucleation occurs. Additionally, the formation of M^+ within A during loading is accompanied by an exothermic release of latent heat, resulting in a momentary local temperature rise of about 6 °C as seen in profile (a) of Figure 2.21(b). By (b), the temperature profile has already settled to its original one due to the slow loading rate and the rapid heat transfer to the TE surfaces.

The reason for this dynamic event is the fact that the transformation is nearly isochoric (volume preserving), causing high strained regions to be thinner than low strained regions.

A lateral incompatibility develops across the wire cross-section at the boundary of the two phases, thus a nucleation cannot occur as a sharp region of infinitesimal axial extent. Abrupt changes in wire diameter are not energetically favorable in a polycrystalline wire, so a finite-length neck (hourglass-like profile), with strong strain gradients at either end, forms instead. This lateral strain compatibility creates an energy barrier that must be surmounted, so the axial force overshoots the final equilibrium stress associated with the fully-developed nucleation region. Consequently, $A \rightarrow M^+$ nucleation during axial loading starts at a high axial stress, $\sigma_n^{\mathrm{AM}^+} = 535$ MPa, then drops until the response reaches the propagation stress, $\sigma_p^{\mathrm{AM}^+} = 441$ MPa (a nucleation peak of 94 MPa in this case). It is reasonable to expect that the dynamic nature of the nucleation event also causes some local damage, or true plastic slip, which likely contributes to the small, 0.05 % residual strain measured after complete unloading.

This nucleation peak is similar to the upper yield point/lower yield point instability observed at the onset of Lüders band formation in mild steels [1, 58, 88] and is akin to certain structural problems with propagating buckles [4, 56]. The size of this load drop is strongly dependent on the quality of the experimental setup, boundary conditions, and geometry of the specimen, so any imperfections will cause an "early" (so-called heterogeneous) nucleation, masking the true response. In particular, stress concentrations at grips often chop off, or dramatically reduce, the measured nucleation peak for $A \rightarrow M^+$ initiation. Consequently, nucleation peaks are not often observed in SMA superelastic responses unless special care is taken to minimize such experimental artifacts.

After (b), the two fronts spawned by the nucleation event propagate away from one

another as seen in the streak plot of Figure 2.21(c) during the $A \rightarrow M^+$ transformation. As $A \rightarrow M^+$ fronts enter the temperature gradient regions, the loading plateau slightly increases in slope past about $\delta/L = 3\%$ in Figure 2.21(a) then takes a somewhat sharper upturn at $\delta/L = 5.6\%$ as the fronts reach the outer edges of the upper and lower TEs. At this point the crosshead motion stops to prepare for unloading.

Before unloading, a higher voltage is applied to the outer two TEs, lowering their temperature by about -11 °C, resulting in a temperature profile similar to (d) in Figure 2.21(b) and the load settling straight downward as seen in Figure 2.21(a). The crosshead motion is then reversed to unload the specimen and the axial stress response decreases in a nonlinear manner until about $\delta/L = 4.6\%$ at point (c). During this time the two fronts reverse direction, now $M^+ \rightarrow A$ fronts, but quickly reach the colder outer TE and stop. Since colder regions require a lower stress (than the current stress) for this reverse transformation to continue, conditions for propagation become unfavorable, and the existing fronts are "frozen" in place, although some $M^+ \rightarrow A$ transformation likely occurrs in a uniform way in the gauge section, judging from the nonlinear stress response. From (c) to (d) a pocket of A appears within the (largely) M^+ phase near the midspan of the specimen, a sudden *jump up* in stress occurs from $\sigma_n^{\mathrm{M^+A}} = 148$ MPa to $\sigma_p^{\mathrm{M^+A}} = 225$ MPa in Figure 2.21(a), and self-cooling occurs locally (by about -3.5 °C) as seen in profile (c) of Figure 2.21(b). Unlike the nucleation during loading, the temperature *drop* is caused by the *endothermic* latent heat absorption, and the nucleation of pocket of A (low strain) causes elastic *reloading* of the specimen.

Thus, the nucleation response during unloading is up-side-down compared to the one

during loading. The temperature field is manipulated to create a "reverse dog-bone" specimen to cause $M^+ \rightarrow A$ nucleation in the gauge length. This nucleation phenomenon during unloading is often observed to some degree in SMA superelastic experiments. For example, see the superelastic responses of Figure 2.14, where, even when the temperature is not so carefully controlled, grip stress concentrations are not usually favorable for $M^+ \rightarrow A$ transformation, thereby causing reverse nucleation to occur somewhere in the gauge section. However, the $M^+ \rightarrow A$ stress jump here (77 MPa) is larger than is normally measured, indicating again that this setup and procedure more effectively minimized heterogeneities and more closely measured the ideal, perfect specimen response. It is important to note, however, that this arrangement was intended to measure the nucleation peaks, but not the strain jump across the entire plateau, since all the material was not transformed in the free length of the wire specimen. For that, one can rely instead on the previous experiments, such as those of Figure 2.14, for accurate measurements of the transformation strains along the stress plateaus.

Returning to the description of the experiment, just after the $M^+ \rightarrow A$ nucleation the temperature returns to the profile (d) of Figure 2.21(b) where the local latent heat peak has already dissipated. Soon thereafter, the voltage in the outer two TEs is adjusted back to its original level and the temperature slowly returns to the profile like (b). While the temperature is evolving four $M^+ \rightarrow A$ fronts travel in the specimen as seen in Figure 2.21(c), the original two converging from the ends and the two new ones diverging from the nucleation site. Later, however, once the temperature reaches the original bowed shape that was hotter at the ends, the two outer fronts speed up, and the inner two fronts (first the upper

middle one then the lower middle one) stop moving. During this time the strain continues to decrease while the stress remains relatively constant at the propagation stress from Figure 2.21(a), although some minor waviness is seen as the temperature is adjusted. The lowermost front eventually reaches the next (now stationary) front above it, and the uppermost front reaches the next (also stationary) front below it, as shown in Figure 2.21(c). At these times momentary up-ticks in the stress response at $\delta/L = 0.85\%$ and 0.46% are measured as front pairs near each other suddenly (dynamically) coalesce together. At these times the necks vanish pairwise, and the strain de-localizes. Final unloading of A occurs elastically, and a minor amount of residual strain remains upon full unloading.

During tension of PVC and HDPE polymers it has been shown that while the nominal (engineering) stress-strain response is locally up-down-up, leading to neck formation and propagation, the true (Cauchy) stress-logarithmic strain maintains a positive tangent modulus (slope) [19, 39]. An argument is sometimes made that the local behavior of SMA wire stems solely from the change in cross-sectional area after transformation as it does in polymers, but this is not the entire story for SMAs.

In SMA wire tension, the geometric effect arising from the reduction in cross-section diameter is a contributing factor, but does not account for the stress peak measured. In this case, assuming volume-preserving, axisymmetric transformation, a transformation strain of $\Delta\varepsilon_{\mathrm{P}} = 0.06$ (Figure 10 of [17]) corresponding radial (lateral) strain of $\Delta\varepsilon_{\mathrm{r}} = -\Delta\varepsilon_{\mathrm{P}}/2 = -0.03$. The relative reduction of area is, therefore, $(A_{M^+} - A_A)/A_A = -2\Delta\varepsilon_{\mathrm{r}} = -0.06$, or -6%. With a nucleation stress of 535 MPa, the drop in nominal stress that can be attributed to the area change is only 32.1 MPa, yet a stress drop of 94 MPa was measured, a factor of

almost 3 higher. Likewise, the strain jump during reverse transformation is, conservatively, -0.055, so the stress jump at a nucleation stress of 148 MPa due to cross-sectional area increase is 8.14 MPa, yet 77 MPa was measured, a factor of 9.5 higher! Since this is an imperfection sensitive problem and one cannot claim to have eliminated *all* geometric and temperature field heterogenieties, the ideal, perfect stress peaks are likely to be larger yet. Thus, true material instability, not just geometric softening, is at play here, and the SMA wire would exhibit a local up-down-up response even if the true local stress-strain response could be measured. Once again, Figure 2.18 shows pieces of the mechanical response just measured near the nucleation peaks (thick lines) and the dotted lines are speculative extrapolations of the local stress-strain response inherent in the material. Unfortunately, one cannot hope to capture the unstable (negative slope) portions of the local response, so one is left to construct a hypothetical local response that matches the stable portions of the measured response and capture the propagation stress by the Maxwell construction (equal area rule) for each direction of transformation (see again Figure 12 of [17]). It is true that the Maxwell construction assumes reversible equilibrium thermodynamics, which is not the case here as evidenced by the hysteretic behavior of the material, but it serves as a starting approximation in the absence of better information.

2.9 Quasi-Phase Diagram

This same type of experiment was repeated across a range of temperatures by varying the TE voltages and circulating fluid temperature. At each temperature, four relevant stresses can be extracted, $\sigma_n^{\mathrm{AM}^+}$, $\sigma_p^{\mathrm{AM}^+}$, $\sigma_n^{\mathrm{M}^+\mathrm{A}}$, and $\sigma_p^{\mathrm{M}^+\mathrm{A}}$, and plotted against the temper-

ature of the inner TE as shown in Figure 2.22. The plot is a more complete version of the quasi-phase diagrams from Part 2 [17], having added the nucleation stresses, (open circles) to the previously measured propagation stresses (filled circles). Below 0 °C, the $M^+ \rightarrow A$ transformation is not spontaneous even at zero load, so only show the superelastic range of temperatures is shown. The solid lines show fits for nucleation stress σ_n, and the dotted lines for propagation stress σ_p. The trends of both $A \rightarrow M^+$ stresses are nearly linear, with the difference between $\sigma_n^{\mathrm{AM}^+}$ and $\sigma_p^{\mathrm{AM}^+}$ growing as the temperature is increased. Above 50 °C, permanent deformation sustained during loading suppresses $\sigma_n^{\mathrm{M}^+\mathrm{A}}$ and lowers $\sigma_p^{\mathrm{M}^+\mathrm{A}}$, leading to the nonlinear dependence on temperature. Since the $\sigma_n^{\mathrm{M}^+\mathrm{A}}$ and $\sigma_p^{\mathrm{M}^+\mathrm{A}}$ stresses are sensitive to the prior history and the point of unloading, the measured values were obtained by unloading from a point just beyond the loading stress plateau in each case.

It is important to reiterate that these nucleations are rather unusual local phenomena, and capturing the maximum stress peak is highly dependent on the quality of the experiment. This requires a specialized experimental setup before they can be studied in a controlled manner. Additionally, the measurements shown here have only been performed on one alloy of one diameter, and there may be differences in the exact shape of the mechanical response, the measured nucleation peaks and propagation stresses and their hysteresis due to the geometry of the specimen (gauge length and wire diameter). Thus, these aspects have not been broadly studied, so further measurements of the particular wire alloy at hand should be performed.

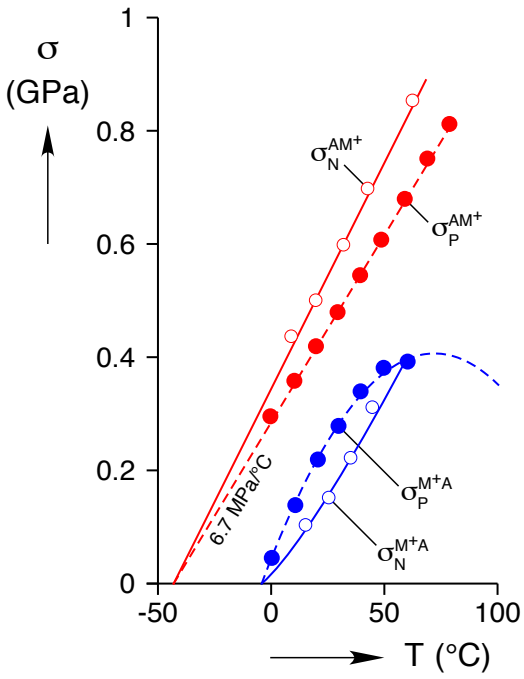


Figure 2.22: Propagation stresses (σ_{P}, closed circles) and nucleation stresses (σ_{N}, open circles) during superelastic responses at several temperature.

2.10 Summary

This chapter started by presenting basic material behavior through the techniques of DSC and isothermal tension experiments. Next, the discussion focused on mechanical instabilities in SMA wires during superelastic behavior. After reviewing how these instabilities give rise to localized, propagating transformation fronts, we presented a novel experimental setup and procedure designed to capture the nucleation of these fronts. By using multiple thermoelectric wafers in contact with the wire, the wire's axial temperature field was manipulated to create an effective "dogbone" specimen. Upon unloading, the temperature field was inverted to create an opposite, "reverse dogbone" effect in the same specimen.

An experiment performed with this new setup showed that at 24 °C a nucleation of M^+ was accompanied by a stress drop of -94 MPa from $\sigma_n^{\mathrm{AM}^+}$ to $\sigma_p^{\mathrm{AM}^+}$, and a nucleation of A by a stress rise of 77 MPa from $\sigma_n^{\mathrm{M}^+\mathrm{A}}$ to $\sigma_p^{\mathrm{M}^+\mathrm{A}}$. In experiments without a specialized setup such as this, these dynamic localization events are either diminished or even fully suppressed, and the true material behavior is masked. By repeating the same experiment across a range of superelastic temperatures, we were able to create a more detailed phase diagram showing how nucleation stresses are strongly coupled to temperature, just like the propagation stresses. These phenomena lead to interesting, and hypersensitive, loading rate and ambient media dependence of the response of SMA wire.

Some final comments are in order regarding why the applications engineer should take interest. While we find these phenomena fascinating, we recognize that localized transformations are more specialized (dare we say esoteric) phenomena than those introduced in

this chapter. It is true that transformation fronts occur only in relatively virgin NiTi under uniaxial loading, not in other SMAs nor in NiTi subjected to more complex stress states or even uniaxial compression. Many device designers get by quite happily by ignoring their existence, and if one is testing conditioned material having undergone sufficient prior thermomechanical training cycles this *may* be reasonable (as will be shown in Chapter III). We have also, however, met engineers who were unaware of their existence and fell into one or more of the following pitfalls:

- permanent strain was misinterpreted as a material effect, which actually arose from grip slippage exacerbated by transformation fronts;
- inaccurate data was obtained at loading rates too high, unaware of latent heat induced self-heating/cooling effects;
- closed loop instabilities occurred during load control material testing due to sudden jumps in displacement upon nucleation and/or monitoring strain in an inactive region of the specimen;
- temperature or strain was monitored in an actuation device at one discrete location, sending a discontinuous feedback signal to a controller, resulting in unpredictable behavior.

Due to these difficulties, some have even abandoned SMA device implementation as being too complex or finicky, which is a pity. Consequently, we feel that understanding localized transformation is essential for accurate thermo-mechanical characterization and modeling

of SMAs, and a mature awareness improves the chances for successful implementation, especially for early-cycle, uniaxial NiTi applications where localized transformation behavior is likely. Furthermore, researchers wanting to undertake more advanced SMA experimentation, involving training, fatigue testing, rate studies, or micro-mechanical research, should be interested in the stress, strain, and temperature inhomogeneities resulting from localized transformation.

Chapter III

Thermo-electro-mechanical Shakedown Response of Conditioned SMA Wires

Until this point, experiments have been performed on two different "virgin" NiTi wires. This chapter departs somewhat from that theme by focusing on the response of conditioned NiTi wires. After introducing typical differences between virgin and conditioned wires, the particular wire chosen for this study is characrized: 90C Flexinol™ from Dynalloy Corp. These wires are subjected to a number of thermomechanical cycles at several dead-loads, and the evolution of the response is examined. In addition to the properties measured in the previous chapter (strain, temperature, stress) this separate study measures electrical resistivity.

The electrical properties of SMA wires change with both the temperature and phase, making them another significant material property that in many cases can be measured with less expensive equipment. Electrical resistivity (ER) is of interest as both a method for position feedback control [96] and a way to identify phase changes [38, 46]. After an introduction of effective methods of resistivity measurement, a series of constant-load experiments is presented, exploring the resistivity response of 90C Flexinol.

The natural progression of the techniques introduced in previous chapters is to apply

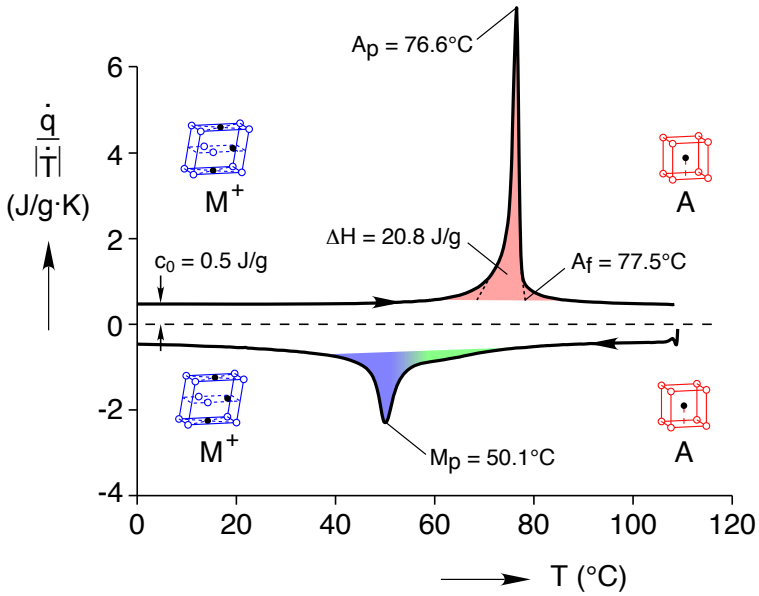
them to wires undergoing multiple thermo-mechanical cycles. The specific term *shake-down* is used to describe this evolution of SMA properties as the material is cycled early on in its "life". Two types of cycling are typical for experiments. The easier of the two is constant temperature cycling, which requires no additional expertise beyond the techniques in Section 2.3, provided a quality load frame is used for the displacement control. The second mode involves thermal cycling, either under a constant load or against a spring. More complex cycling regimes are certainly possible, and one of them, involving various alternating constant loads, will be treated at the end of this chapter.

Temperature control, which was used in all of this chapter's experiments, is inherently difficult because of the time scales involved in achieving large (100 °C or more) changes in environmental temperature can cause excessively long experiment times. For example, a thermal chamber, which can take hours to reach thermal equilibrium for a single cycle, is suitable for only a few cycles. For $10^3 - 10^5$ cycles, most experimenters employ joule heating of the SMA wire in a convective medium, either air [11, 79] or fluid [7, 60]. However, both of these methods both have the previously mentioned drawbacks, preventing accurate measurements of wire temperature. In this chapter, an alternate method is used which allows moderate rates of temperature change, while still maintaining uniform temperature (< 1.5 ºC) and accurate temperature measurements. It is in fact quite similar to the thermoelectric-controlled setup described in section 3.2. Since temperature control via thermoelectrics can achieve a thermal ramp of 1 °C/s, a reasonable experiment length of 18-24 hours allowed 150 temperature cycles to be performed.

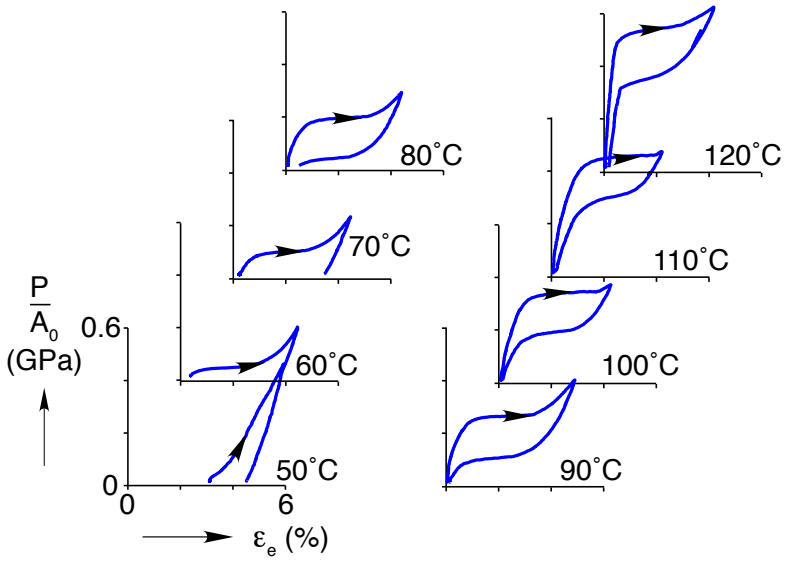
3.1 Conditioned NiTi Wires

While "virgin" equiatomic NiTi wire is often used in scientific studies, it has complications that hinder its use in commercial applications. Often, its austenite finish temperature (A_f) is below 60 °C, making it difficult to build robust actuators that are functional over operating temperatures for typical automotive and aerospace applications. Also, the response of "virgin" NiTi may evolve rapidly with thermomechanical cycling during the first tens to hundreds of cycles, making it problematic for repeatable actuation. Conditioned SMA wire, such as the one chosen for this study (0.365 mm diameter 90C Flexinol) has relatively high transformation temperatures, has good fatigue life, and has been conditioned to minimize most shakedown effects at moderate tensile loads.

The processing and conditioning by the manufacturer results in some important differences from typical virgin NiTi wire. Figure 3.1(a) shows a thermogram obtained from a Perkin-Elmer Pyris 1 differential scanning calorimeter (DSC), using a scan rate of ± 10 °C per minute. The measured differential power $\dot{q} = dq/dt$ to the specimen (per unit mass) has been converted to $\dot{q}/|\dot{T}|$ to give it specific heat-like units. Several properties can be extracted from this experiment. The specific heat c_{o} is half the difference in $\dot{q}/|\dot{T}|$ between the heating and cooling curves, 0.5 $\mathrm{J}/(\mathrm{g} \cdot \mathrm{K})$. The austenite finish temperature $A_{\mathrm{f}} = 77.5$ °C is estimated from the trailing edge of the $M^+ \rightarrow A$ peak (top curve), and the latent heat ΔH, the shaded area under the peak, is 20.8 J/g. Transition temperatures at peak values, $A_{\mathrm{p}} = 76.6$ °C (heating) and $M_{\mathrm{p}} = 50.1$ °C (cooling), are also shown. The hysteresis between these two peaks is only 26.5 °C, much less than the 60 to 80 °C typical of virgin NiTi (see Figure 2.2(b)). During cooling (lower curve), the $A \rightarrow R$ and $R \rightarrow M^+$ peaks,



(a)



(b)

Figure 3.1: 90C Flexinol: (a) differential scanning calorimetry, (a) isothermal mechanical responses.

typically separate for virgin NiTi (slightly nickel-rich), are essentially merged, with the $A \rightarrow R$ peak only showing up faintly as a precursor to the larger peak. During heating (top curve), the $M^+ \rightarrow A$ peak is tall and thin, spanning only 15 °C, compared to the typical 20 to 40 °C for virgin NiTi. The $M^+ \rightarrow A$ peak is also centered at least 10 °C above the $A \rightarrow R$ peak, while the offset is usually less than 2 °C for virgin NiTi. This narrow and high temperature $M^+ \rightarrow A$ peak in the DSC thermogram is typical of a specimen with a significant two-way shape memory effect [41]. Experimental and modeling efforts usually neglect the effects of the R-phase due to its low transformation strain (about 1%) and its absence above moderate loads. Yet, as will be seen, the R-phase has a large effect on electrical resistivity, and thus plays a significant role in interpreting the low-load resistivity response.

As an initial baseline thermomechanical characterization, a series of isothermal tension experiments, performed at 10 °C increments from 50 to 120 °C, are shown in Figure 3.1(b). The supplier provided the SMA wire pre-strained to about 4 to 5% to facilitate its subsequent application in thermal actuators. For consistency, austenite is used as the reference configuration to establish a consistent reference length. In each experiment, the specimen was assembled in an setup similar to the one described later in Section 3.2 and held at a stress of about 20 MPa, just enough to keep it straight. It was then heated to 140 °C, well into the A regime, at which point the reference gauge length $L_{\mathrm{e}0}$ was identified as the zero strain configuration. Next, the temperature was brought to the selected experiment temperature and loaded-unloaded under displacement control at a slow rate. The mechanical responses are reported as axial stress P/A_0 (axial force/reference cross-sectional area) vs.

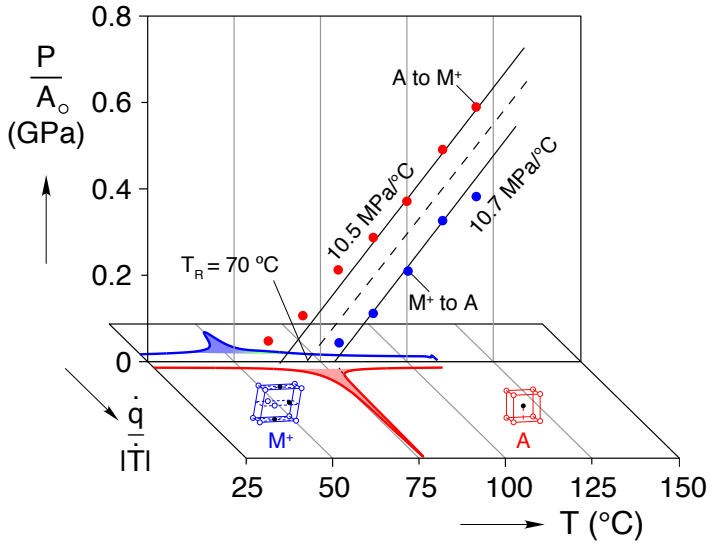


Figure 3.2: Transformation (plateau) stresses and DSC.

gauge strain ε_e (gauge elongation/reference gauge length from a non-contacting laser extensometer). At temperatures above 90 °C the mechanical responses are superelastic, similar to that of NiTi wire, with force plateaus during $A \rightarrow M^+$ (austenite to tensile-martensite, loading) and $M^+ \rightarrow A$ (unloading) that rise progressively with increasing temperatures. The plateaus are somewhat less distinct and have less extent in strain compared to that of typical virgin NiTi wire, having a plateau strain $\Delta\varepsilon$ of about 3.5%. A small residual strain is apparent after the load-unload cycle in all the superelastic responses, with a minimum value of about 0.1% at 90 °C and then increasing to about 0.18% at 120 °C. At temperatures below about 50 °C, the loading plateau disappears below the zero load axis, and loading starts at about 3% strain, as a consequence of the large *two-way* shape memory effect [42, 62] in this material. This means that under load-free conditions, a specimen undergoing cooling transforms from austenite to oriented martensite ($A \rightarrow M^+$) (rather than the typical austenite to thermal martensite, $A \rightarrow M^{+/-}$), resulting in a strain change

on both heating and cooling. Thus, at 50 °C, the specimen starts as largely M^+ (tensile-oriented martensite), so the response is nonlinear but rather stiff, with only an additional 1.5% residual strain after loading to 450 MPa. Similar responses were obtained at lower temperatures (not shown).

Figure 3.2 shows a plot of the loading and unloading stress plateaus against temperature, which are fitted with straight lines, along with a comparison to the stress-free DSC thermogram. The plateau stresses increase at temperature rates of and 7.9 MPa/°C for loading and unloading, respectively. The average of these fits (dotted line), projected to zero stress, gives a reference temperature of $T_{\mathrm{R}} = 70$ °C, which is a useful estimate of the theoretical, equilibrium (hysteresis-free) transformation temperature between austenite and martensite.

3.1.1 R-phase and Resistivity

At zero load, a NiTi crystal can take one, or a mixture, of three phases: a low-symmetry (monoclinic) martensite phase (in various microstructural arrangements [8]) at low temperatures, a high-symmetry (cubic) austenite phase at high temperatures, and an R-phase (rhombohedral) at intermediate temperatures. The R-phase appears most prominently upon cooling from austenite, and its effect is an initial nonlinear knee in the mechanical response during the first 0.5 to 1.5% strain, or so. While it is often ignored in constitutive models and experiments, it does however make a large contribution to the resistivity and thus cannot be ignored if a model is to make accurate predictions of actuator power and performance.

Studies of NiTi SMA resistivity are numerous but are by no means exhaustive. They

have been performed for a single load, single temperature cycle, but with no temperature measurement [79], single load, single temperature cycle with temperature measurement [3], many temperature cycles, but limited data resolution and temperature control [98], several loads, but under 80 MPa, with no temperature cycling [41], and some interesting recent experiments performed to identify the R-phase using *in situ* neutron diffraction and electrical resistivity measurements [92]. Modeling efforts have for the most part neglected this property in favor of purely thermomechanical properties (see [27] for a few models using R).

Figure 3.3 shows a dead-load experiment at a small stress (19 MPa) as the temperature was increased and decreased between room temperature and about 160 °C (solid lines – heating, dotted lines – cooling). The strain response (ε_{e}, black line) shows a hysteretic loop with low strains at high temperature (A) and strains near 3.8% at low temperature (M^{+}). It shows a large TWSME, stretching out even at the small stress. We suspect this two-way effect is driven by residual martensite, locked in place by plastic dislocations during the conditioning process. The resulting internal stresses are in an orientation that produces tensile martensite even at zero global stress.

The relative resistivity response ($\bar{\rho}_{\mathrm{e}} = (\rho_{\mathrm{e}} - \rho_{\mathrm{e0}})/\rho_{\mathrm{e0}}$, blue line) shows an unusual (for NiTi) hysteresis loop with local maxima occurring during both heating and cooling. Also, overlaid on the plot is the DSC thermogram with vertical construction lines (red) showing the peak temperatures, M_p and A_p, for reference. The temperature A_p during heating corresponds well with the temperature at which both $\varepsilon_{\mathrm{e}}$and $\bar{\rho}_{\mathrm{e}}$ drop sharply. The temperature M_p during cooling (increasing strain) coincides exactly with the onset of the drop

in $\bar{\rho}_{\mathrm{e}}$. The blue construction lines indicate temperatures where $\bar{\rho}_{\mathrm{e}}$ has a local extremum. During heating, $\bar{\rho}_{\mathrm{e}}$ has a peak that occurs as the strain (near 3.5%) begins to decrease more dramatically and the DSC thermogram begins its upturn into the austenite peak. During cooling, the minimum $\bar{\rho}_{\mathrm{e}}$ occurs as the strain (near 0.2%) begins increasing and where the DSC thermogram begins decreasing towards the martensite peak. The maximum $\bar{\rho}_{\mathrm{e}}$ line during cooling is associated with the sudden plateau in $\bar{\rho}_{\mathrm{e}}$, which seems to be where the DSC thermogram enters the martensite peak more sharply. It is also associated with a strain of 2%, seeming to indicate the exhaustion of transformation to R-phase.

3.1.2 Measuring Resistivity

A common method used to measure wire resistivity is called the *four point method*. The four point method is preferred because a moderate contact resistance (several Ohms) between the measurement leads and the wire can be tolerated with no loss in precision. Figure 3.4 shows the circuit, in which a relatively constant current (as much as possible without causing significant resistive heating) is passed through both the electrically isolated wire and a calibrated shunt resistor (R_s). The voltage across both the shunt resistor (V_1) and gauge length (V_2) are measured, and the unknown resistance R_{e} of the gauge section can be calculated from Ohm's law as $R_{\mathrm{e}} = R_s \frac{V_2}{V_1}$. The gauge length resistance R_{e} is related to the material's resistivity (ρ_{e}, $\omega\mathrm{m}$) by $R_{\mathrm{e}} = \frac{\rho_{\mathrm{e}} L_{\mathrm{e}}}{A}$. If elongation and resistance are measured from the same gauge length, the geometric effects of wire lengthening and thinning can be removed from the resistance measurement. Assuming uniform isochoric deformation ($AL_{\mathrm{e}} = A_0 L_{\mathrm{e}0}$) and a reference area A_0, the electrical resistivity is

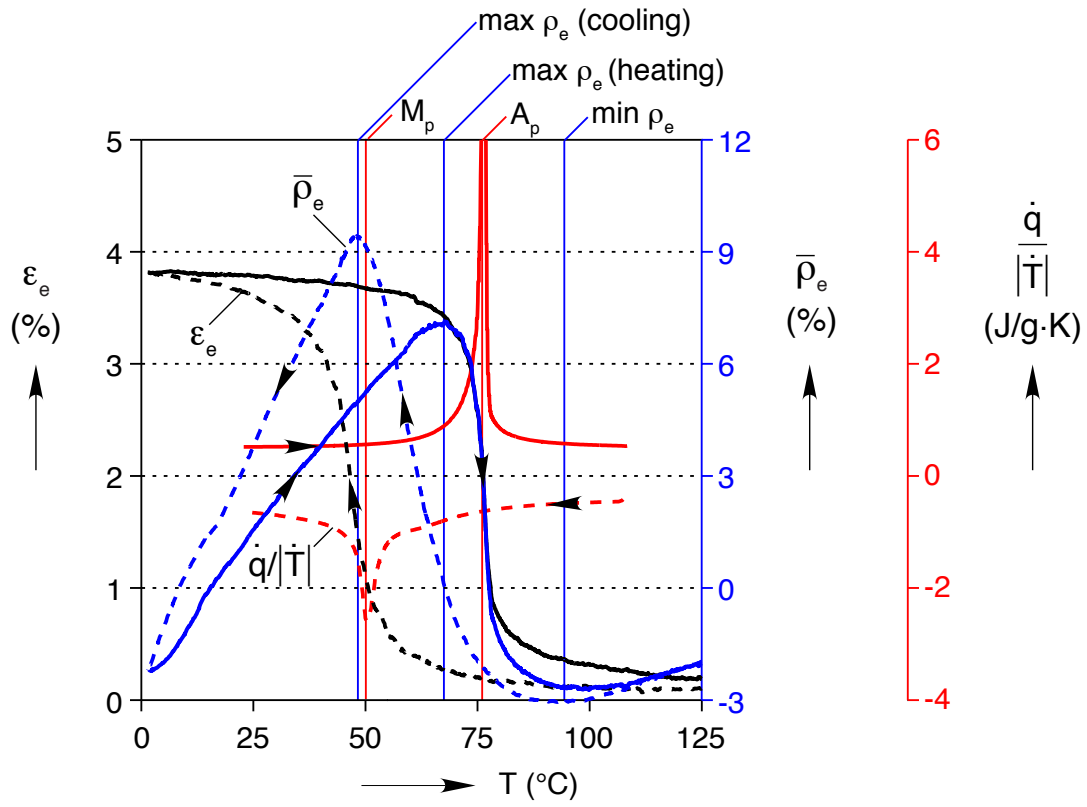


Figure 3.3: Constant load experiment at 19 MPa under temperature control, showing strain (ε_e), relative resistivity ($\bar{\rho}_e$) and DSC output ($\dot{q}/|\dot{T}|$) vs. temperature (T).

$\rho_e = R_e A_0 L_{e0}/(L_e)^2$. Although resistivity is already independent of specimen length, the relative value $\bar{\rho}_e = (\rho_e - \rho_{e0})/\rho_{e0}$ helped remove some variability in the electrical attachment and the relative position of the laser extensometer tags. Although resistivity is already independent of specimen length, the use of the relative value $\bar{\rho}_e = (\rho_e - \rho_{e0})/\rho_{e0}$ helps to remove any remaining variability in the electrical attachment and the relative position of the laser extensometer tags.

While the method is simple on its face, several factors make an accurate SMA resistivity measurement problematic, especially with gauge sections less than 100 mm. First of all, measurement from the grips, as mentioned in section 2.3.2, is difficult for resistance for the same reason it is for strain: slippage is inevitable. Unless the wire is very long, it is best to measure both resistance and strain from the same local gauge length.

To obtain the most precise measurement possible of the voltage drop across the gauge length, electrical leads must maintain contact with the same material point for the entire experiment. Especially during experiments at high loads, the leadwire must then remain firmly attached while the specimen experiences axial and radial strains up to 15% and -6.8%, respectively. This mismatch is sufficient to induce slippage or fracture of adhesives or solder joints, resulting in noise and an artificial increase in measured resistance. As the wire slips through the connection, it causes considerable noise in the resistance measurement.

After several iterations for attachment methods, it was found that the optimum method required a strong clamp around the wire, as shown in Figure 3.5, that performed the dual tasks of connecting a resistance measurement leadwire to the specimen and holding a re-

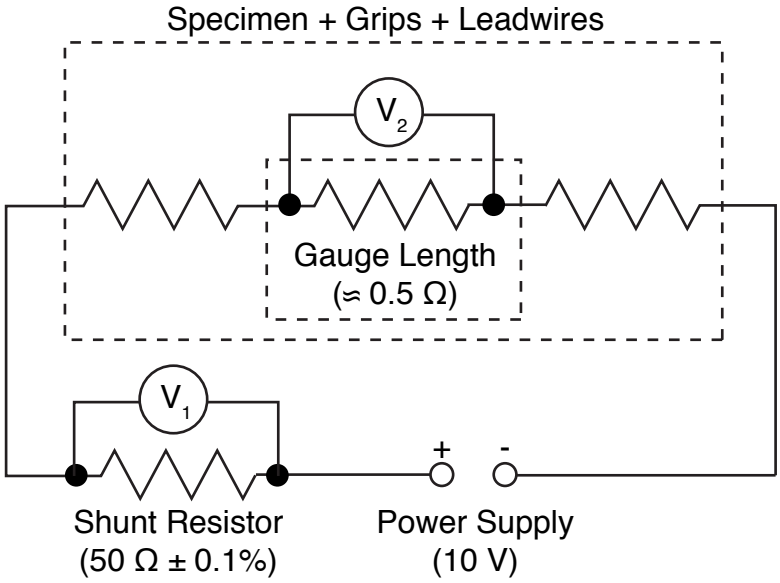
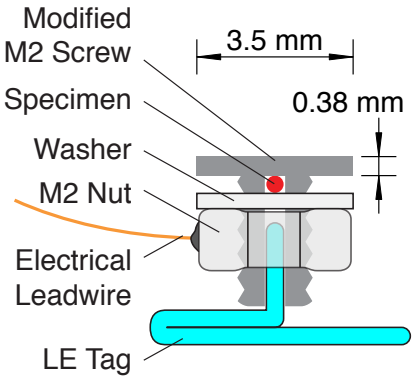
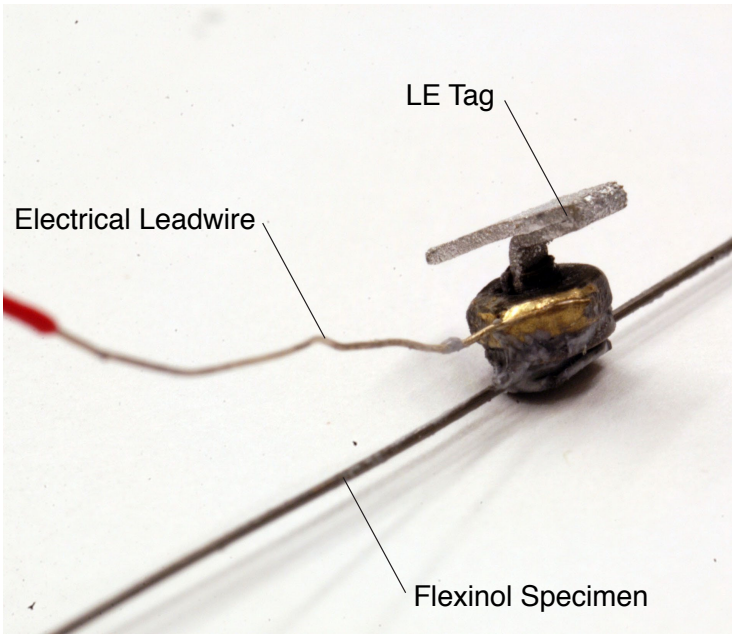


Figure 3.4: Circuit diagram for four point measurement of wire resistance. The current supplied by the power supply I is calculated from the voltage drop V_1 across a calibrated shunt resistor, then used to determine the unknown gauge length resistance.

flective LE tag for strain measurement. As the schematic shows, a small M2 socket head cap screw was machined until the head was only 0.38 mm thick, to allow the wire to be held in close thermal contact with the thermoelectric plate. The specimen was held in place in a 0.38 mm slot cut through the screw's threads by a washer and nut. The resistance leadwire was soldered directly to the nut. A high-temperature aluminum LE tag (ARO Measurement Systems HTT-01) was held in the same slot as the wire. As the nut tightened on the wire, it pinched the two sides of the screw thread together, clamping the tag in place. Figure 3.6 shows an example experiment of a specimen held at the high stress of 574 MPa, undergoing thermal cycling between 39 and 194 °C. Even at this relatively extreme load, the resistance measurement remained clean during 20 cycles.



(a)



(b)

Figure 3.5: A custom fastener to measure resistance and strain from NiTi wires: (a) top view digram (b) photograph, as installed

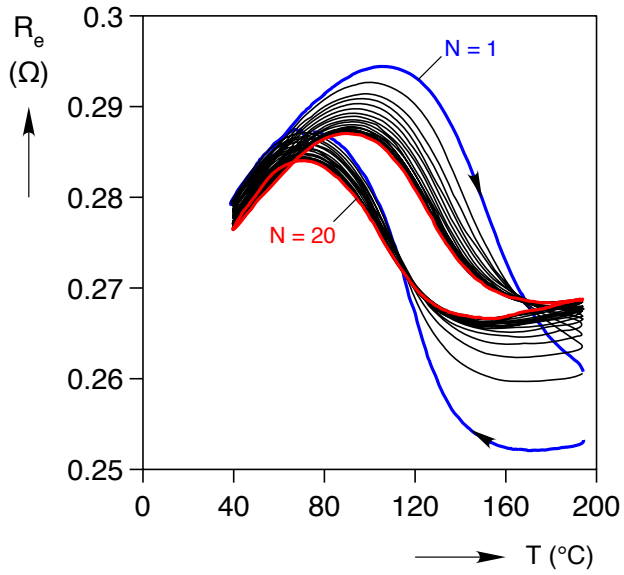


Figure 3.6: Resistance response at 574 MPa of 30 mm long 90C Flexinol.

3.2 Experimental Setup

The experimental setup described below evolved from similar ones previously used in our laboratory [18, 49], and was designed to specific experimental goals: (1) accurate temperature control and measurement of the specimen, (2) load control (i.e. dead-load) or displacement control experiments, (3) access to one face of the specimen for optical and infrared imaging, if necessary, (4) non-contacting, local strain measurement, and (5) electrical resistance R_e measured across the same local gauge length, away from the specimen grips. Under dead-loading, the temperature window for each experiment was greater than 160 °C. We desired a temperature rate of 1 °C/s controlled to $<$ 5 °C of the prescribed temperature and a temperature uniformity of $<$ 1.5 °C along the gauge length. These goals precluded the use of fluid baths (temperature range too narrow), air chambers (too slow), or joule heating (too inaccurate). We chose instead to use conduction contact of the specimen

with thermoelectric modules attached to a temperature-controlled heat sink that fit within a general purpose testing machine.

The thermo-mechanical setup is shown in Figure 3.7. The specimen was held by flat-plate grips in a servo-hydraulic load frame (MTS model 585) that could be operated in either displacement or load control. The local gauge length L_{e} was measured via a non-contacting laser extensometer (EIR model LE-05) as the distance between two reflective LE tags affixed to the specimen (Figure 3.7A). The engineering strain was calculated from this length as $\varepsilon_{\mathrm{e}} = (L_{\mathrm{e}} - L_{\mathrm{e0}})/L_{\mathrm{e0}}$, where L_{e0} was the measured reference length. The load P was measured by a 500 N load cell (not shown). A $50 \times 50 \times 90$ mm aluminum heat sink, held at a user-selected, constant temperature by a recirculating bath, was placed behind the specimen. Two 40 mm square thermoelectric (TE) wafers (making a 40×80 mm surface) sandwiched between the specimen and heat sink provided up to 200 W of active heating or cooling to the specimen via an external PID controller (Figure 3.7B). The TEs' back surface was affixed to the heat sink using adhesive tape (Thermattach T412) to assure uniform heat transfer, and a 1.6 mm thick copper plate was placed between the TE front surface and the back side of the specimen to improve the axial uniformity of the temperature field. The wire was gripped about 0.4 mm in front of the copper plate, and thermally conductive and electrically insulating paste (OmegaTherm™201, $2.3\ \mathrm{W/(m \cdot K)}$) filled the gap to ensure good thermal contact, while allowing the specimen to slide freely during testing. Using an infrared radiometer, the temperature variation across the copper plate was verified to be < 1.5 °C, while being held at 100 °C above the heat sink temperature. The temperature of the plate was measured during experiments at two places by 0.076 mm

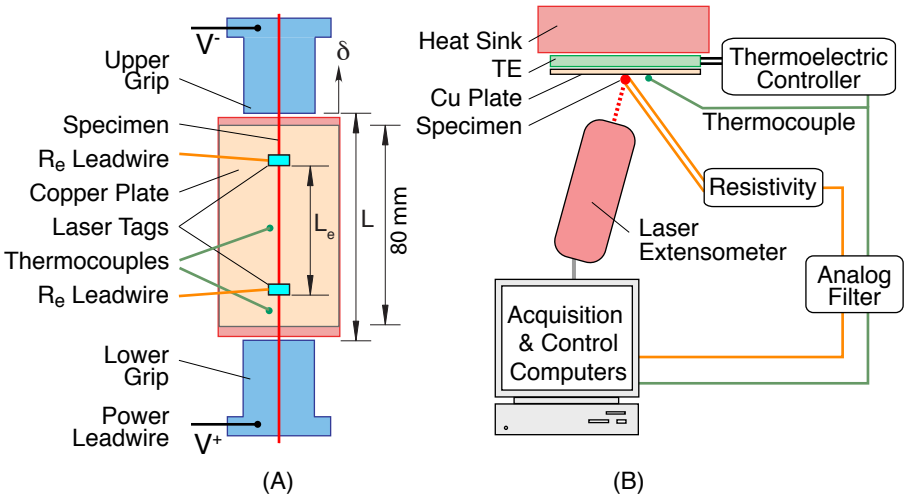


Figure 3.7: Schematic of experimental setup: (A) front view, (B) top view.

K-type thermocouples immersed in thermal paste adjacent to the wire, one at the center of the plate and one about 10 mm from the bottom edge.

The electrical resistance of the gauge section of the specimen was measured by the four-point probe described in Figure 3.4. The specimen was electrically isolated from the rest of the setup, both by painting the copper plate with insulating paint (Rustoleum High-Heat Enamel) and by using nylon sleeves between the grips and the load frame. Using a shunt resistor in series with the specimen, an external power supply (Agilent™ E3633A) provided a constant electrical current of $I_e = 200$ mA through the specimen. The temperature rise in the specimen due to Joule heating was predicted to be only about 0.1 °C.

3.3 Resistivity Experiments, 19 to 478 MPa

Following are two dead-load thermal cycles, presented in detail, as part of a larger series of experiments. The first was at the lowest load (19 MPa) that could maintain a

straight wire. The second is at a moderately high load of 287 MPa, high enough that transformations to R-phase should be suppressed. This section ends with an overview of six similar experiments, spanning dead loads from 19 MPa to 479 MPa.

3.3.1 Experiment 1: 19 MPa (2 N Load)

Figure 3.8(a) shows the time (t) histories of temperature (T, black line), strain ε_{e}, blue line), and resistivity change($\bar{\rho}_{\mathrm{e}}$, green line) during a single thermal cycle. The rate of temperature change was held near 1 °C/s, except for times when phase transformations were expected and the rate was lowered to 0.5 °C/s. In this and all other experiments using this alloy, the reference length $L_{\mathrm{e}0}$ and $\rho_{\mathrm{e}0}$ were measured when the wire is stress-free at 160 °C. As was mentioned in section 3.1, the material exhibits a strong two-way shape memory effect (TWSME) and thus ε_{e} increases during cooling to a maximum of 3.83% even at this small load. In contrast, the resistivity response is non-monotonic. During cooling, it first reached a local minimum at $t = 64$ s, then ramped up to a very rapid change to a local maximum at 131 s. Note again, just as in Figure 3.3, that the rapid rise in resistivity $\bar{\rho}_{\mathrm{e}}$ between these local extremes only accompanies a 0.3% rise in ε_{e}. In fact, the rise in ε_{e} to its maximum coincided with a drop in $\bar{\rho}_{\mathrm{e}}$ from its peak of 9.4% to below 5%. On heating, the two changed almost simultaneously, with ε_{e} dropping to a minimum of 0.1% and resistivity from its local maximum of 7.1% to -2.7%, due to $M^{+} \rightarrow A$ transformation.

Figures 3.8(b)–(c) show the evolution of ε_{e} and $\bar{\rho}_{\mathrm{e}}$ as the temperature was cycled between 160 and 1 °C. Figure 3.8(b) shows that the magnitude of the TWSME was 3.77%, and that ε_{e} changes very little at the high and low temperature extremes, since it did not

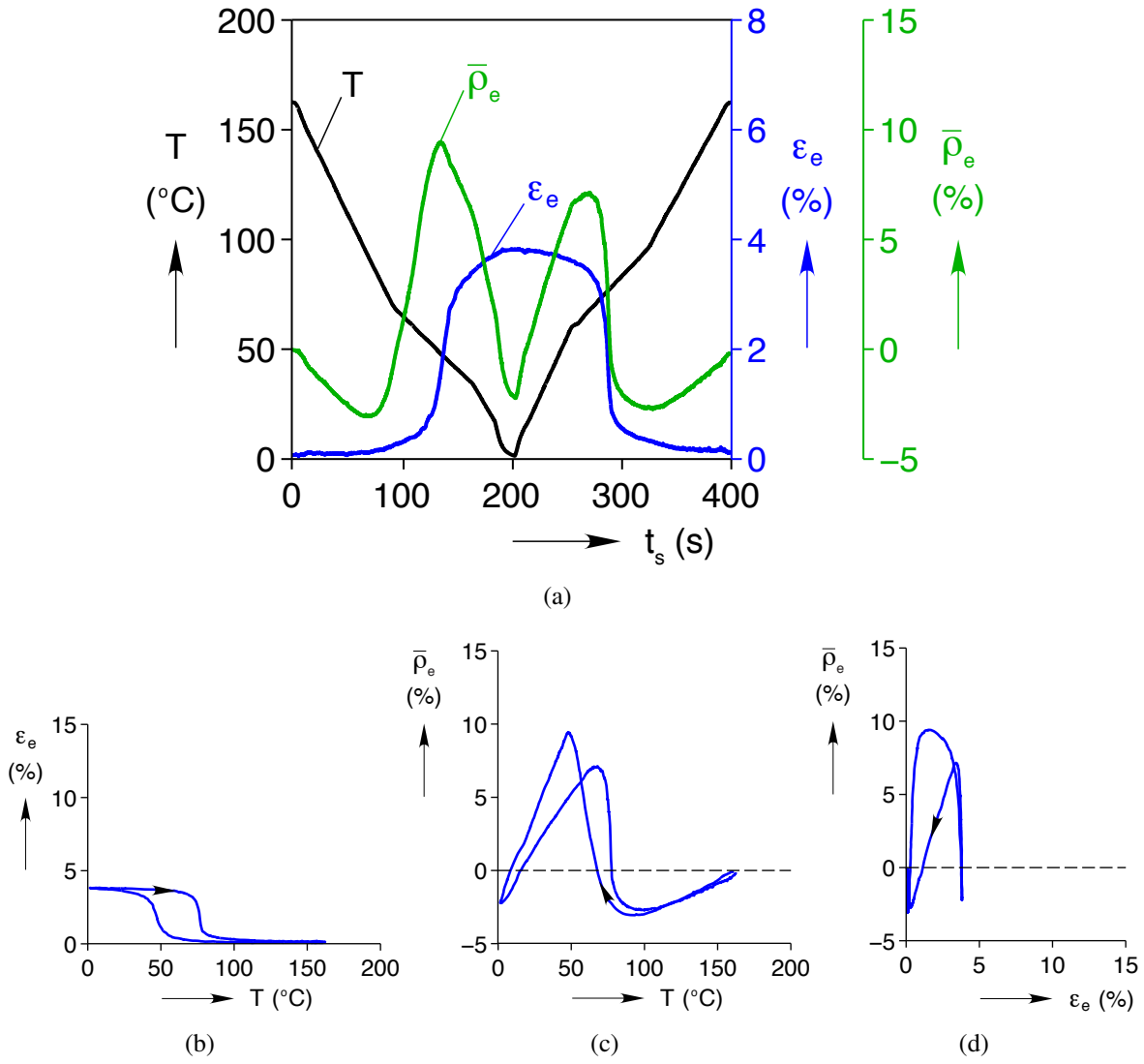


Figure 3.8: 19 MPa dead load experiment showing: (a) temperature (T), strain (ε_e), and relative resistivity ($\bar{\rho}_e$) histories during a single temperature cycle; (b) strain response; (c) relative resistivity response; (d) resistivity vs. strain

undergo phase change. The temperature hysteresis based on ε_{e} is about 30 °C.

In contrast, Figure 3.8(c) also shows that at high temperatures, when transformation is complete, $\bar{\rho}_{\mathrm{e}}$ had a linear dependence on temperature of about 0.049%/°C in A. On cooling, $\bar{\rho}_{\mathrm{e}}$ reached a local maximum at $T = 48$ °C. This sharp peak is associated with the production of R. At low temperatures, the $\bar{\rho}_{\mathrm{e}}$ heating and cooling paths did not coincide. This was due to experimental limitations, as the specimen did not reach a sufficiently low temperature to complete the transformation (likely $R \rightarrow M^{+/-}$) associated with the subsequent drop in ρ_{e}. Upon heating, $\bar{\rho}_{\mathrm{e}}$ rose again, peaking at a temperature of 65 °C at a local maximum of 7.25%. This heating peak is absent in typical "virgin" wire under stress-free conditions, and it is typically only associated with wire being cycled under load [98]. Because of the early transformation during cooling, the temperature hysteresis based on $\bar{\rho}_{\mathrm{e}}$ was only about 17 °C.

Of particular relevance to actuators and sensors is the relationship between strain and resistivity shown in Figure 3.8(d). During the portion of decreasing strain associated with heating in the transformation regime, there is a nearly linear relationship between the two. During the cooling transformation regime (increasing strain), the nonlinearity is another manifestation of the asynchronous action of ρ_{e} and ε_{e} shown previously in Figure 3.3. At high strains (low temperatures), the steep "tail" represents the change in resistivity while the strain remains relatively constant.

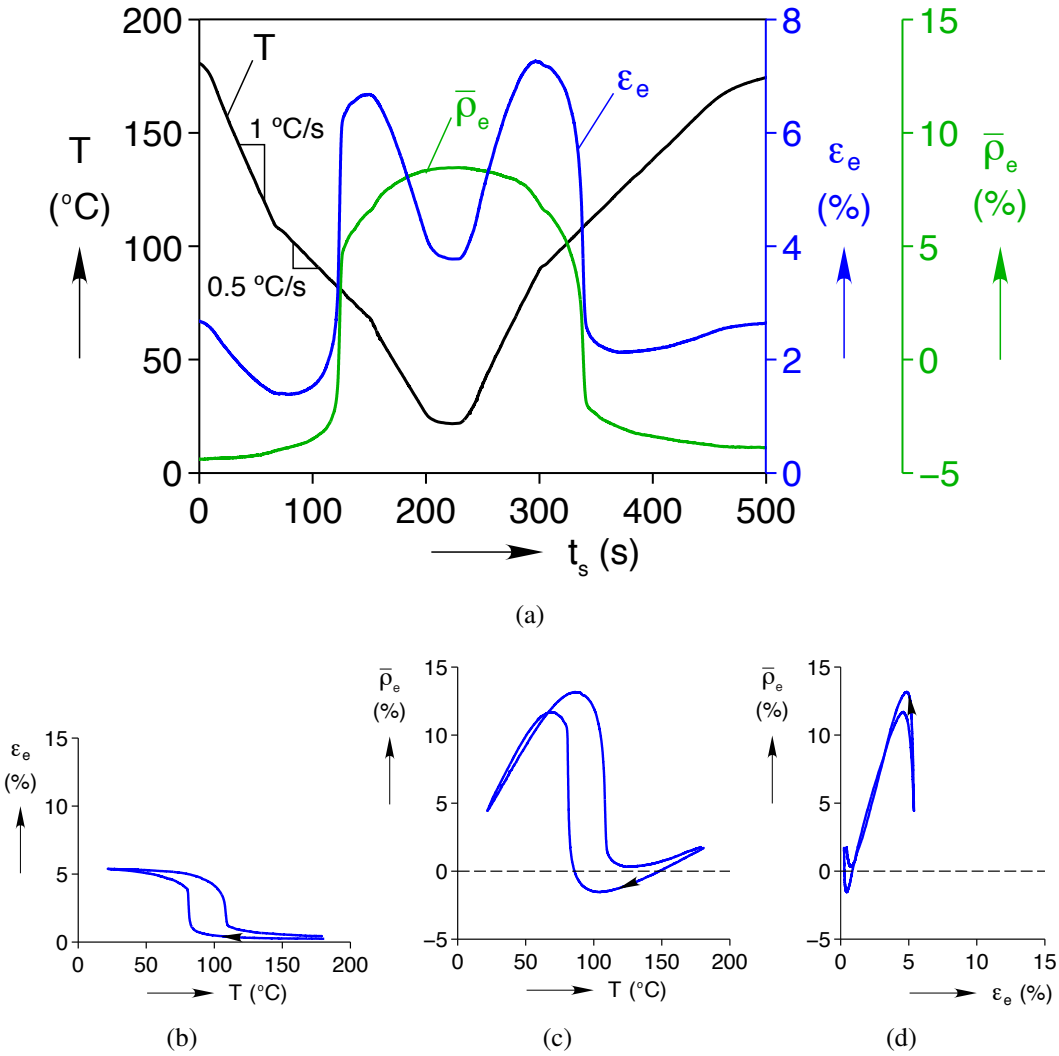


Figure 3.9: Constant Load Experiment 4 (287 MPa dead load) showing: (a) temperature (T), strain (ε_e), and relative resistivity ($\bar{\rho}_e$) histories during a single thermal cycle, (b) strain response; (c) relative resistivity response; (d) resistivity vs. strain

3.3.2 Experiment 4: 287 MPa (30 N Load)

The previous experiment is now compared to a similar one at a higher constant load of 287 MPa (30 N load). At this load level (see Figure 3.9), the temperature cycled between 25 °C and 184 °C, reflecting the higher temperatures needed to transform from M^+ to A. Due to limitations in heating efficiency, the linear temperature rise could not be maintained following $t = 400$ s, but the maximum temperature was eventually reached. Here, there was a predictably larger recoverable strain of 5.13%, from a minimum of 0.25% to the maximum of 5.38%. Unlike Figure 3.8, both ε_{e} and $\bar{\rho}_{\mathrm{e}}$ rose simultaneously at $t = 124$ s during cooling and fell at $t = 338$ s during heating.

Figure 3.9(c) shows the resistivity response against temperature. On cooling, $\bar{\rho}_{\mathrm{e}}$ reached a local minimum at $T = 104$°C, and then rose (in tandem with ε_{e}) at $T = 82$°C to a local maximum of 11.7%. Compared to Figure 3.8(c), this peak was not as sharp, and quickly joined the non-hysteretic path of the heating $\bar{\rho}_{\mathrm{e}}$ response. On heating, $\bar{\rho}_{\mathrm{e}}$ reached another local maximum of 13.2% at $T = 86.5$ °C, though unlike Figure 3.8(c) the heating peak was the larger of the two. The reason behind this change is the absence of R at this higher load. Without the rise in $\bar{\rho}_{\mathrm{e}}$ associated with R, the response largely followed the linear dependence on temperature at each (hot and cold) temperature extreme. In this case, a higher peak upon heating was only a sign that the heating transformation from $M^+ \rightarrow A$ took place at a higher temperature than the cooling transformation from $A \rightarrow M^+$.

This absence of R is most obvious in the plot of $\bar{\rho}_{\mathrm{e}}$ vs. ε_{e} in Figure 3.9(d). Here, the relationship was almost perfectly proportional during transformation, when $0.8\% < \varepsilon_{\mathrm{e}}$ 4.5%. There remains a similar nonlinear relationship between the two outside this range

due to the mismatch in temperature coefficients between $\bar{\rho}_{\mathrm{e}}$and ε_{e} (5.4×10^{-4}/ °C vs.$-8.4\times$ 10^{-6}/ °C for A, respectively.

These effects are easier to visualize with the idealized responses in Figure 3.10, which overlay responses from Figure 3.9. First, the ε_{e} - T response in Figure 3.10(a) has a flat response in the "pure" regions of M^{+} and A, indicated by solid blue and red lines. Next, the $\bar{\rho}_{\mathrm{e}}$ response in Figure 3.10(b), which takes the response from Figure 3.9(c) and overlays straight construction lines. The two linear responses of "pure" M^{+} and A have slopes of about 0.15 and 0.05 %/°C, respectively, diverging with increasing temperature. Therefore, the curve that stays on the linear (pure) path longest will create the more extreme peak. When these two responses are combined in Figure 3.10(c), the result is a relationship between ε_{e} and $\bar{\rho}_{\mathrm{e}}$ that is not hysteretic, but only proportional during transformation.

3.3.3 Summary of Resistivity Experiments

Relative resistivity $\bar{\rho}_{\mathrm{e}}$ and average strain ε_{e} are now summarized for all six the constant loads performed. The trends noted in sections 3.3.1 and 3.3.2 are verified with four additional loads: 96, 191, 379, and 478 MPa.

For reference, Figure 3.11(b) shows strain responses for each load. It was as expected; as the load increased, so did the transformation temperatures, maximum strains, and minimum strains. At the highest load of 478 MPa, ε_{e} reached maximum and minimums of 7.22 and 1.85%, respectively. Temperature hysteresis between transformation also increased with load, from 29.5 to 39.5 °C. Relative resistivity responses in Figure 3.11(a) show much of the same story, with temperature hysteresis as well as maximum $\bar{\rho}_{\mathrm{e}}$ increasing with load.

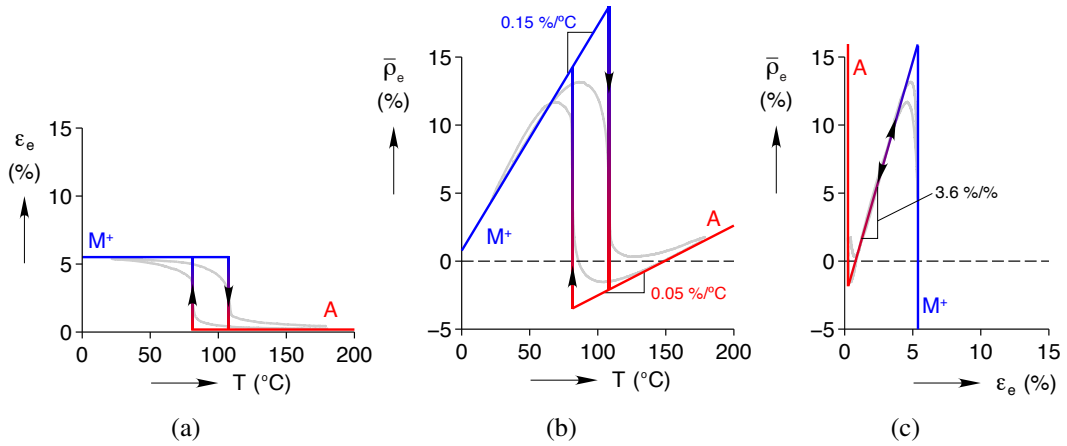


Figure 3.10: Idealized construction lines overlay responses from the 287 MPa dead load experiment in Figure 3.9. (a) The idealized ε_e - T response is flat for "pure" M^+ and A. (b) The idealized $\bar{\rho}_e$ - T response at 287 MPa is linear and diverging for the pure phases. (a) This causes a non-proportional relationship between ε_e and $\bar{\rho}_e$ outside the transformation region.

The one exception is at the lowest loads, where the distinct R-phase peak was prominent at 19 MPa, less pronounced at 96 MPa, and nearly absent by 191 MPa. The influence of R-Phase on the response is even easier to see in Figure 3.11(c), where the gap between the heating (decreasing ε_e) and cooling (increasing ε_e) curves decreased with load.

Another idealized response created from the experiments in Figure 3.11(a) is shown in Figure 3.12. The bold green, blue, and red lines represent presumed linear $\bar{\rho}_e$ - T responses for pure R, M^+, and A, respectively. Transformation paths are shown as shaded vertical lines. At a high stress like 478 MPa (experimental data is in the background), transformation is limited to the paths within the dashed box on the right. The path is similar to that of Figure 3.10(b), with $\bar{\rho}_e$ moving along the two linear paths, jumping between them at critical transformation temperatures, higher ones in this case due to the higher constant load. At a low stress (19 MPa), the path during cooling is multi-step, moving from the red pure A line to the green, higher R region, and finally to the blue M^+ response. During heating, the R

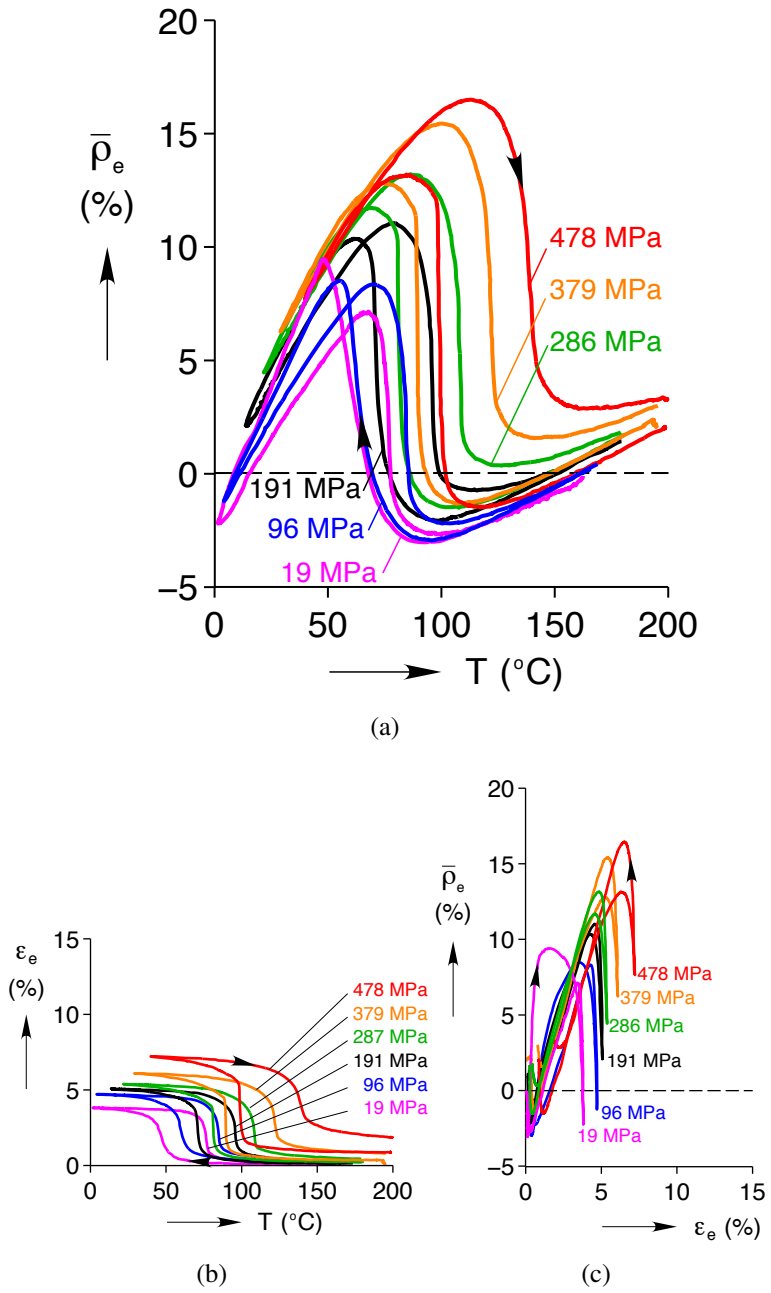


Figure 3.11: Summary of constant load experiment at six loads from 19 to 478 MPa, showing: (a) strain (ε_e) vs. temperature (T), (b) relative resistivity vs. temperature, and (c) relative resistivity vs. strain

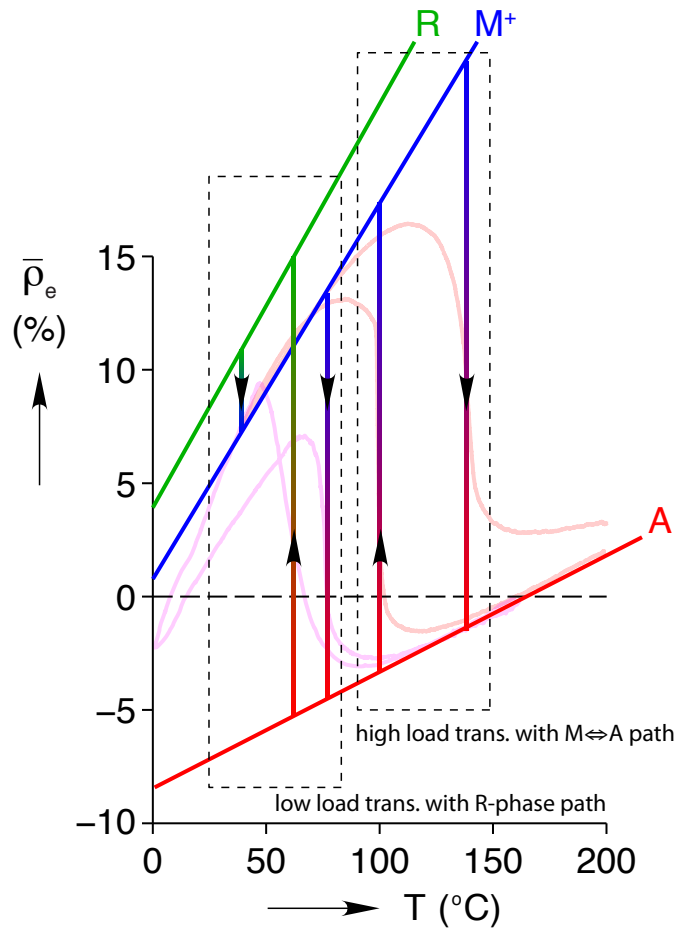


Figure 3.12: Idealized T - $\bar{\rho}_e$ transformation paths for low load thermal cycles (left dashed box) and high load thermal cycles (right dashed box) T - $\bar{\rho}_e$ experimental data from 478 and 19 MPa experiments are in the background.

is not part of the transformation, which is directly between M^+ and A. One shortcoming of these assumed paths is that none of the transformations are "sharp", as material properties have likely been widely distributed by internal damage from the manufacturer's conditioning. As a result, the transformation $A \rightarrow R$ likely never reaches pure R before parts of the specimen begin the $R \rightarrow M^+$ transformation. During low load heating, $\bar{\rho}_{\mathrm{e}}$ is likely suppressed from the pure M^+ path by premature transformation to A. Some success has been reported in the literature in modeling the mechanical response of conditioned SMAs with distributions of material properties [43], and it would seem to be a good way to model the $\bar{\rho}_{\mathrm{e}}$ response as well.

3.4 Constant Load Shakedown Experiments

A series of six shakedown experiments was performed at constant loads between 19 and 478 MPa. The as-received material from the manufacturer is initially curved on a spool, so during experiments a minimum tensile stress of about 20 MPa was required to keep the wire straight. Complete results from an experiment at 379 MPa are presented below. For brevity, only selected results are presented for the remaining experiments. As before, the resistivity change is normalized using the measured stress-free resistivity at 160 °C, $\rho_{\mathrm{e}0} \equiv R_{\mathrm{e}0} A_0 / L_{\mathrm{e}0}$.

3.4.1 Example shakedown experiment: 379 MPa (40 N)

To demonstrate the shakedown and reference measurement procedures, data histories leading up to, and including the first deadload cycle ($N = 1$) at 379 MPa are shown

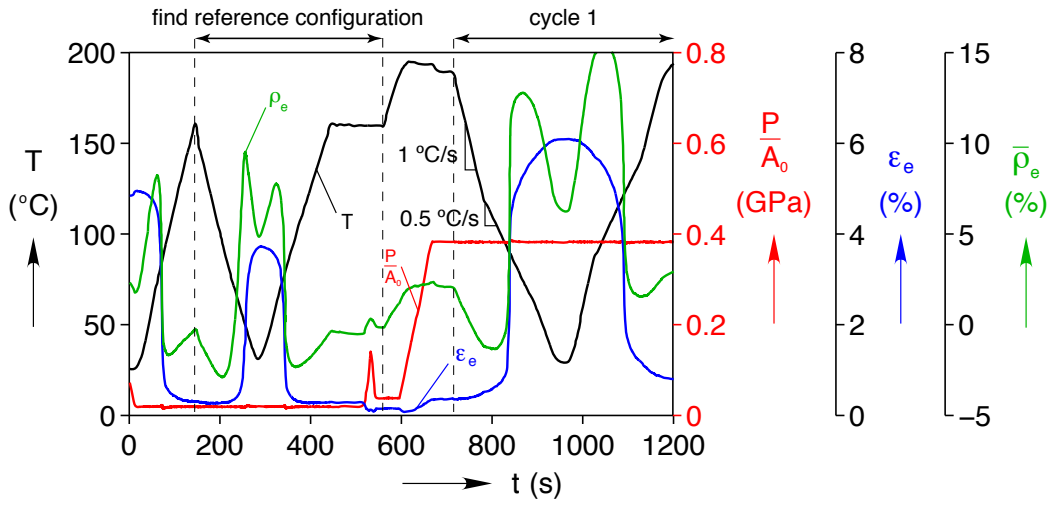


Figure 3.13: Initial heating and first cycle of shakedown experiment at 379 MPa, with histories of temperature T, stress P/A_0, strain ε_e, and relative resistivity $\bar{\rho}_\mathrm{e}$.

in Figure 3.13. The multiple vertical axes show temperature, stress, strain, and relative resistivity. The experiment starts with the specimen at 25 °C at 75 MPa stress (to assist in assembly). For the first 13 s, the load was lowered to 19 MPa, just enough to keep the specimen straight. Under load control, the temperature was raised to 160 °C, to provide a uniform starting point (first vertical dashed line). From 150 to 450 s, the temperature was lowered to 30 °C and then back to 160 °C at a rate of 1 °C/s. This low-load cycle ensured that the "zero strain" measurement occurred at the same point in the hysteresis loop.

Next, the temperature was held at 160 °C, and at $t = 510$ s the specimen was loaded to 140 MPa, then unloaded to 37 MPa in displacement control at $\dot{\delta}/L = 3 \times 10^{-4}$ s^{-1}. Notice that as loading started, ε_e went down before it went up. This was an anomaly caused by the further straightening of the wire and rotation of the reflective tags. To correct for this, a linear fit was applied to the part of the elastic response above 80 MPa. Figure 3.14 shows how that fit was extrapolated to zero stress, in this case giving $L_{\mathrm{e}0} = 43.11\mathrm{mm}$. The

reference resistivity was also measured in an identical fashion to be $\rho_{e0} = 8.21 \times 10^{-4}\ \Omega\mathrm{m}$ (in these experiments, ρ_{e0}ranges from 8.08 to $8.35 \times 10^{-4}\ \Omega\mathrm{m}$).

Following the L_{e0} and ρ_{e0} measurements, the temperature rose to the start point of the prescribed thermal cycle, 195 °C, which was followed by a displacement-controlled, isothermal, stress rise to the experiment's cycling stress, 379 MPa. The second dashed line marks the start of the first shakedown cycle. First, the temperature was decreased at a rate of -1 °C/s until it reached 115 °C, at which point the rate was slowed to -0.5 °C/s. The slower rate was used to improve accuracy during transformation while measurements were expected to be changing rapidly. Up to $t = 830$ s, the wire remained in its A state, with ε_e holding constant at 0.35% and $\bar{\rho}_e$ dropped nearly linearly with temperature from a maximum of 2.8%. From $t = 830$ to 850 s, the specimen transformed to M^+, and both ε_e and $\bar{\rho}_e$ increased rapidly. Following the relatively rapid transformation, strain continued to increase as the temperature dropped. By contrast, the $\bar{\rho}_e$ response was nonmonotonic, increasing first, then decreasing with temperature as it did before transformation. The temperature reached its lowest value of 30 °C at $t = 960$ s, at which point the temperature was increased. Initially, $\bar{\rho}_e$ rose with the temperature while ε_e remained constant, until transformation started at $t = 1050$ s. As the specimen transformed back into A, ε_e and $\bar{\rho}_e$ dropped simultaneously. Following transformation, $\bar{\rho}_e$ resumed its increase with temperature but at a slower rate, while ε_e slowly dropped until it reached its minimum of 0.8%, and the temperature returned to its starting value of 195 °C.

This same heating and cooling shakedown cycle was repeated to a total of 150 times. Additionally, after every tenth cycle a 19 MPa thermal cycle was performed, identical to the

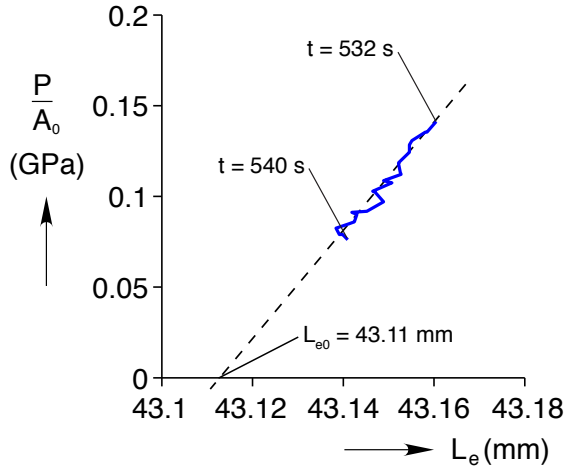


Figure 3.14: Mechanical response of the experiment in Figure 3.13 ($532 < t < 540$ s), solid line data. The response is fit (dashed line) to determine the stress-free reference length L_{e0}.

"find reference configuration" method in Figure 3.13, to measure the accumulated stress-free residual strain. The strain responses from every tenth shakedown cycle are overlaid in Figure 3.15(a), with the first and last cycles highlighted in bold blue and red, respectively. At this high load, the shakedown evolution was relatively rapid, with the minimum strain ε_e^{min} rising from 0.8 to 2.1% during the first 10 cycles. By cycle $N = 150$, the strain evolution had slowed considerably but not stopped, rising from 5.08 to 5.16% during the last 10 cycles. Overall, the specimen's stroke $\Delta\varepsilon_e \equiv \varepsilon_e^{max} - \varepsilon_e^{min}$ dropped monotonically from 5.87 to 4.14%.

Besides the changes in maxima and minima, the strain loop shape evolved over the 150 cycles from having sharp transitions to a more diffuse behavior. For example, during the cooling portion of the cycle, the maximum slope $d\varepsilon_e/dT$ fell from -2.95%/°C ($N = 1$) to -0.17 %/°C ($N = 150$). This had the effect of broadening the temperature span of "active" transformation, mostly in the direction of increasing temperature. One can imagine that this is consistent with the broadening of the distribution of the transformation temperatures

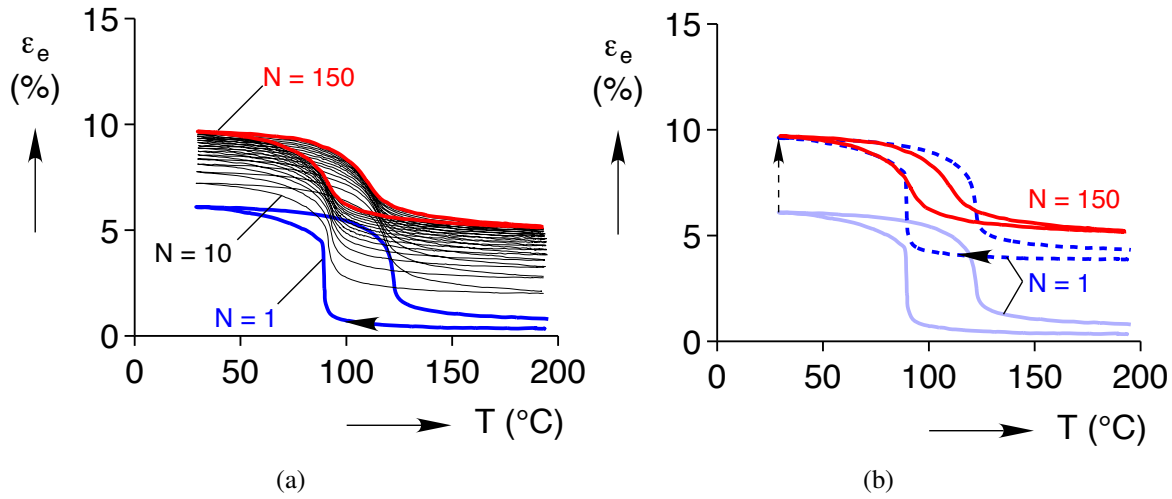


Figure 3.15: Strain response during 150 cycle shakedown at 379 MPa: (a) every 10th cycle, (b) first cycle shifted.

associated with a more heterogeneous spatial distribution of the local residual stress fields in the microstructure. The temperature hysteresis also dropped from 32.4 to 18 °C, with most of the shift happening on the heating side of the transformation ($M^+ \rightarrow A$). This is easier to see in Figure 3.15(b), which shows the first and last cycle strain responses overlaid, where the first cycle's response has been shifted up to meet the maximum strain (low temperature point) of the last cycle. The shape of the strain responses below 90 °C (M^+ regime) was almost identical between all cycles. This suggests that damage occurs primarily during the $M^+ \rightarrow A$ transformation (heating), accumulating in the A phase with relatively little effect on the M^+ response (except to shift the curve upward).

Figure 3.16 shows every tenth cycle of the shakedown response of the relative resistivity $\bar{\rho}_e$. Outside the transformation region NiTi resistivity has a somewhat linear temperature dependence [66]. During the first cycle, the temperature coefficient, $\bar{\alpha}_e \equiv d\bar{\rho}_e/dT$, of the initial section of the cooling curve near $T = 180$°C (A phase) is measured as 0.052%/°C,

while the initial section of the heating curve near $T = 50$ °C (M^+ phase) is measured as 0.168%/°C. Starting from a high temperature (A phase), the resistivity dropped linearly with temperature during cooling until the onset of the $A \rightarrow M^+$ transformation near 105 °C. From there, the resistivity rose rapidly until it "joined" the decreasing linear regime from the M^+ phase, now with a steeper slope. On heating, the reverse happens, and $\bar{\rho}_{\mathrm{e}}$ follows a similar increasing linear path until the onset of $M^+ \rightarrow A$ transformation at 100 °C when the response becomes noticeably nonlinear. Since the $M^+ \rightarrow A$ transformation takes place at a higher temperature and M^+ has a larger temperature coefficient than A by a factor of 3.2, $\bar{\rho}_{\mathrm{e}}$ rises to a higher value on heating than its maximum on cooling. Transformation then occurs rapidly, and $\bar{\rho}_{\mathrm{e}}$ dropped until it nearly rejoined the cooling curve, at which point it rose gently again.

From $N = 1$ to 150, the temperature hysteresis in $\bar{\rho}_{\mathrm{e}}$ dropped from 32.4 to 19.8 °C, similar to the change in temperature hysteresis in the ε_{e} response. During shakedown, there was both a drop in maximum ($\bar{\rho}_{\mathrm{e}}^{\max}$) and a rise in minimum ($\bar{\rho}_{\mathrm{e}}^{\min}$), which seem to be results of different effects. The drop in $\bar{\rho}_{\mathrm{e}}^{\max}$ can be explained by Figure 3.15(b), which shows the evolution of the $M^+ \rightarrow A$ transformation (heating) to lower temperatures. Since the transformation starts earlier, it leaves the steeper M^+ curve earlier (departing downward) and attains a lower maximum. On the cooling side during cycling, the initial $\bar{\rho}_{\mathrm{e}}$ value at 194 °C shifted upward by 4.1%, and $\bar{\rho}_{\mathrm{e}}^{\min}$ rose by 5.4%, which we suspect is due to defect accumulation in the austenite, possibly including pockets of residual martensite. Again, it seems that most shakedown effects occurred during heating, above the onset of $M^+ \rightarrow A$ transformation. Apparently, "defect accumulation" does not significantly influence the M^+

side of the response at low temperatures, perhaps due to the already high "defect density" associated with the complex microstructures of the low-symmetry martensite crystal.

The slope of the $\bar{\rho}_{\mathrm{e}}$-T responses outside the transformation region, however, changed only slightly during cycling. Using the initial section of the cooling curve near $T = 180$°C (A phase), the temperature coefficient $\bar{\alpha}_{\mathrm{e}}$ changed from 0.052 to 0.055%/°C during the 150 cycles. Using the initial section of the heating curve near $T = 50$ °C, the change was from 0.168 to 0.152%/°C (see the dotted construction lines in Figure 3.16).

3.4.2 Strain Shakedown Responses

Strain responses for every 30^{th} cycle of all six experiments are shown in Figure 3.17. The first cycle responses in Figure 3.17 shift upward in strain and to the right in temperature as the stress level increases (as expected), and a significant strain increase is seen during cooling, even at the lowest stress of 19 MPa, consistent with the two-way shape memory effect. At load levels above 191 MPa (most evident at 478 MPa) the first cycle loops do not close due to considerable strain ratcheting. One can compare the first and last cycles for each stress level and see that the strain responses do not change much for the 191 MPa case, which we note corresponds to the 20 N max force recommended by the material supplier. At lower load levels (most evident at 19 MPa) the strain response shifts downward somewhat, which is perhaps surprising (discussed further below). At higher load levels the curves shift upward significantly and transformations become more diffuse with less area within the hysteresis loop (seen most dramatically at the highest load 478 MPa with maximum strain approaching 13%), which is not surprising.

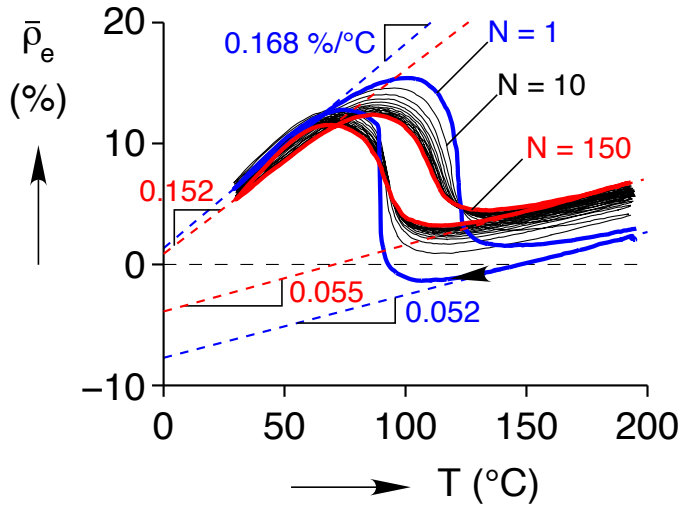


Figure 3.16: Relative resistivity $\bar{\rho}_e$ response during shakedown at 379 MPa (every tenth cycle shown).

The figure also shows that the temperature width of the hysteresis loop during the first cycle was initially larger at the higher loads, 28 °C at 19 MPa, and 40 °C at 478 MPa. By $N = 150$, however, the trend was reversed with the temperature hysteresis becoming smaller at higher loads, 29 °C at 19 MPa and only 16.5 °C at 478 MPa. At the two highest loads, temperature width of the hysteresis loop dropped, but the temperature range over which $M^+ \rightarrow A$ transformation occurred became noticeably larger.

Figure 3.18(a) shows the progression of maximum (ε_e^{max}) and minimum (ε_e^{min}) strains with temperature cycles, and Figure 3.18(b) shows the evolution in stroke defined as $\Delta\varepsilon_e \equiv \varepsilon_e^{max} - \varepsilon_e^{min}$. In general, the greater the difference in stress from 191 MPa, the more accelerated the shakedown. For experiments below 191 MPa, Figure 3.18(a) shows both ε_e^{max} and ε_e^{min} decreased with increasing N, in effect reversing the conditioning applied by the manufacturer. For the lowest load, ε_e^{max} fell faster than ε_e^{min}, leading to a 1% decrease in stroke (recoverable strain) in Figure 3.18(b). At the intermediate loads of 96 and 191 MPa,

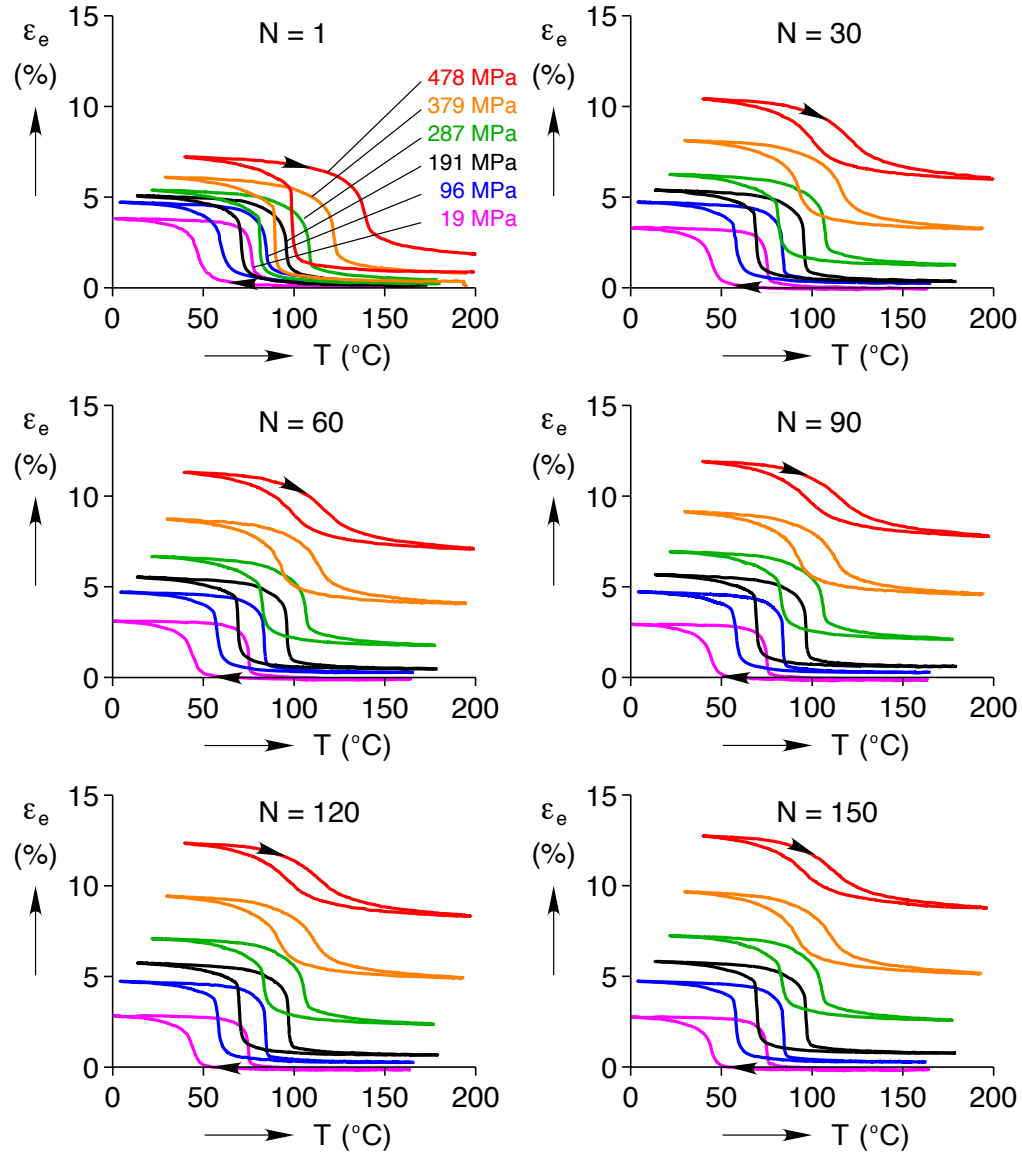


Figure 3.17: Strain responses for six constant load shakedown experiments at 30 cycle intervals. As the material is cycled, the area within each loop grows smaller, especially at loads above 191 MPa.

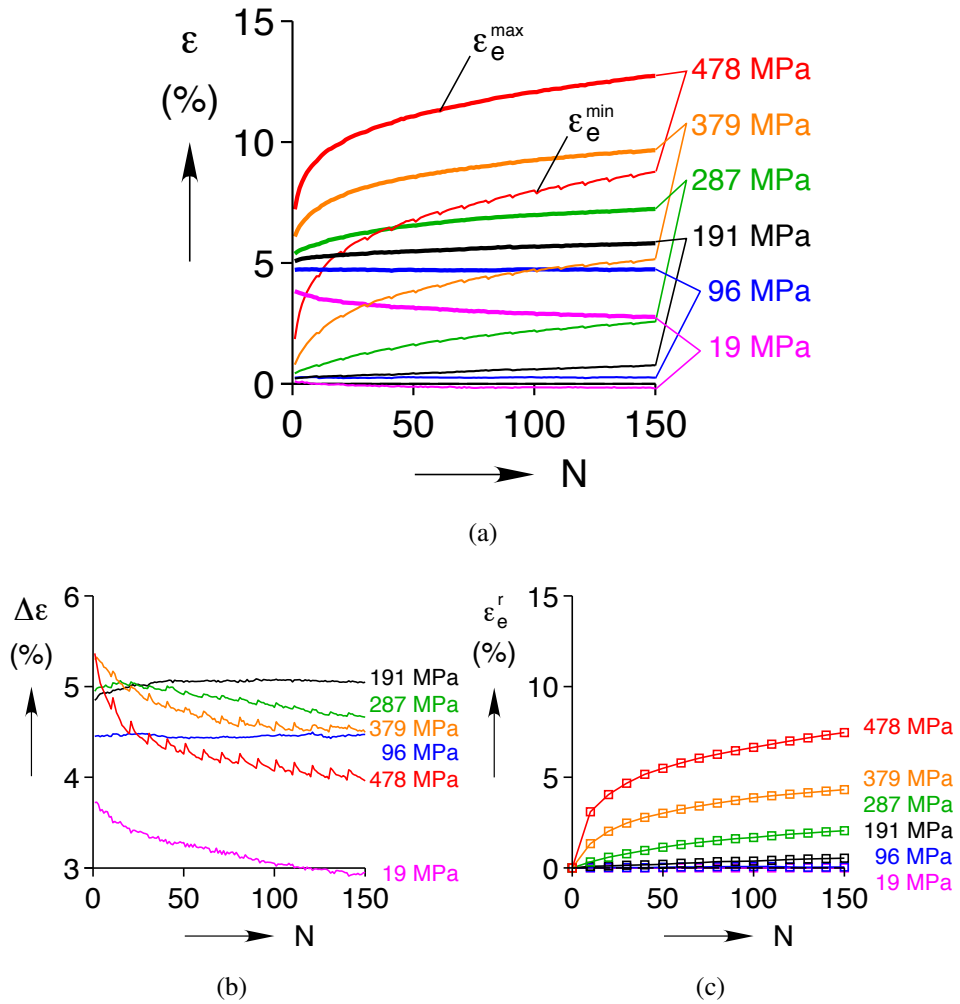


Figure 3.18: Cyclic evolution of characteristic strains: (a) max & min strains, (b) strokes, (c) residual strains. Shakedown is minimal at 96 and 191 MPa.

the stroke actually increased by 0.1 and 0.2%, respectively. At the highest loads (379 and 478 MPa), shakedown was the most severe with both $\varepsilon_{\mathrm{e}}^{\max}$ and $\varepsilon_{\mathrm{e}}^{\min}$ increasing with N, so much so that by $N = 15$, $\Delta\varepsilon_{\mathrm{e}}$ was lower than $\Delta\varepsilon_{\mathrm{e}}$ at 191 MPa experiment. At $N = 150$, $\Delta\varepsilon_{\mathrm{e}}$ at 478 MPa dropped the most, from 5.9 to 4.1%.

Note that in Figure 3.18(b) a "sawtooth" pattern is evident in the stroke at loads above 191 MPa. Its origin comes from a similar inverted pattern in $\varepsilon_{\mathrm{e}}^{\min}$, barely discernible in Figure 3.18(a). Recall that every 10 cycles, a full thermal cycle was performed to 19 MPa for the purpose of measuring residual strain. For each of the 3 highest loads, there is a brief drop in $\varepsilon_{\mathrm{e}}^{\min}$ (or rise in $\Delta\varepsilon_{\mathrm{e}}$) following each 19 MPa cycle, at $N = 11$, $N = 21$, $N = 31$, etc. About 5 cycles after each 19 MPa cycle, the strain response rejoined the larger trend. This phenomenon suggests that the internal processes responsible for shakedown can be reversed, at least momentarily, to a degree. Note, however, the near absence of the sawtooth pattern in every $\varepsilon_{\mathrm{e}}^{\max}$ response (only barely discernible by magnifying the response). This is connected to the aforementioned pattern of the cooling $A \rightarrow M^{+}$ transformation (where ε_{e} rises to its maximum) is less susceptible to reverse shakedown than the heating $M^{+} \rightarrow A$ transformation (where ε_{e} falls to its minimum).

The residual strain $\varepsilon_{\mathrm{e}}^{\mathrm{r}}$ measured after every tenth cycle is shown in Figure 3.18(c). It grew during cycling at loads above 191 MPa (getting worse as the load level was increased), yet actually decreased with cycling at loads below 191 MPa. This effect at the two lowest loads is small, only -0.27%. Interestingly, the residual strain tracks nicely with the evolution of $\varepsilon_{\mathrm{e}}^{\min}$, after accounting for the elastic strain. As an example, Figure 3.19 shows the evolution of 379 MPa residual strains $\varepsilon_{\mathrm{e}}^{\mathrm{r}}$, marked by open squares, and with $\varepsilon_{\mathrm{e}}^{\max}$, marked

by the solid line. It is reasonable to expect these two measurement to follow the same trend, since each was measured at a high temperature, far enough above the transformation regime to expect the specimen to be nearly complete A. This is confirmed by shifting the minimum strain values down by 0.81% to account for mostly elastic deformation (dotted line), which overlays the $\varepsilon_{\mathrm{e}}^{\mathrm{r}}$ data nearly perfectly. Thus, the minimum strain values can be used to reasonably track the evolution of the zero-stress residual strain in future experiments, and no interruption by a low-load cycle is necessary.

3.4.3 Resistivity Shakedown Responses

We can make a similar evaluation of resistivity shakedown responses. Figure 3.20 shows $\bar{\rho}_{\mathrm{e}}$ vs. T at $N = 1$ for all six loads at 30 cycle intervals. All seemed to follow similar temperature coefficients (slopes) outside of the transformation regime. In the intermediate temperatures during transition, $\bar{\rho}_{\mathrm{e}}$ generally has a higher peak during heating, as explained previously. The exception to this rule was the two lowest load experiments (see the 19 MPa case in the figure). As Section 3.1.1 explained, the stress was low enough that an R-phase transformation upon cooling caused a dominant rise in the $\bar{\rho}_{\mathrm{e}}$ response, resulting in a *higher* peak during cooling. The effect was most pronounced at 19 MPa, but it was still present enough at 96 MPa to make the two peaks equal in size.

Over 150 cycles, temperature hysteresis changes in $\bar{\rho}_{\mathrm{e}}$ were similar to those in ε_{e} as shown in Figure 3.17. Hysteresis was initially larger at the higher loads, but after 150 cycles, grew smaller with increasing load. The lowest stress experiment, once again, was the exception due to the appearance of the R-phase during cooling, and the temperature

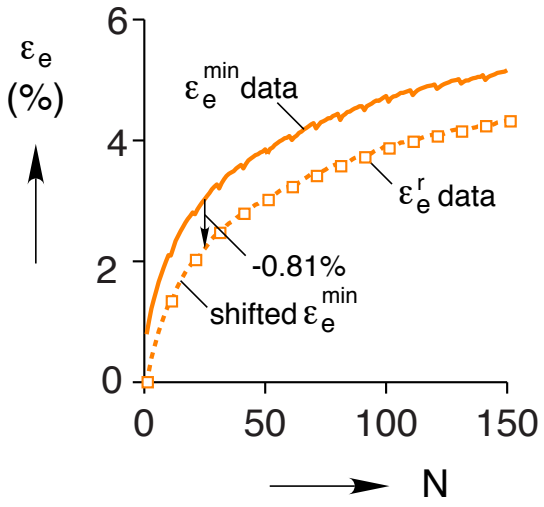


Figure 3.19: Evolution of residual strains (open squares) and minimum strain (solid line) during 379 MPa experiment, and minimum strains shifted to account for elastic deformation (dotted line).

hysteresis remained the smallest for that experiment. At the same time, at 19 MPa the cooling peak in $\bar{\rho}_e$ actually grew larger after cycling. This is indicative of the $A \rightarrow R$ transformation thermally separating itself from the $R \rightarrow M^+$ transformation [67], which led to a larger fraction of R being formed before M^+ appeared.

Similar to the strain response limits in Figure 3.18, Figure 3.21 shows a plot of the minimum, maximum, and difference in $\bar{\rho}_e$ for each cycle at all six load levels. At 191 MPa the maximum and minimum values remained relatively static over all 150 cycles. At 379 MPa and 478 MPa, generally the minimum rose, the maximum fell, and the difference also fell (and was most dramatic at the highest load), with momentary serrations due to the near zero load cycle performed every tenth cycle (though not counted in the cycle count N). Also, for both of these higher loads, while the evolution slowed significantly towards the later cycles, the responses to not appear to have stabilized completely even after 150 cycles.

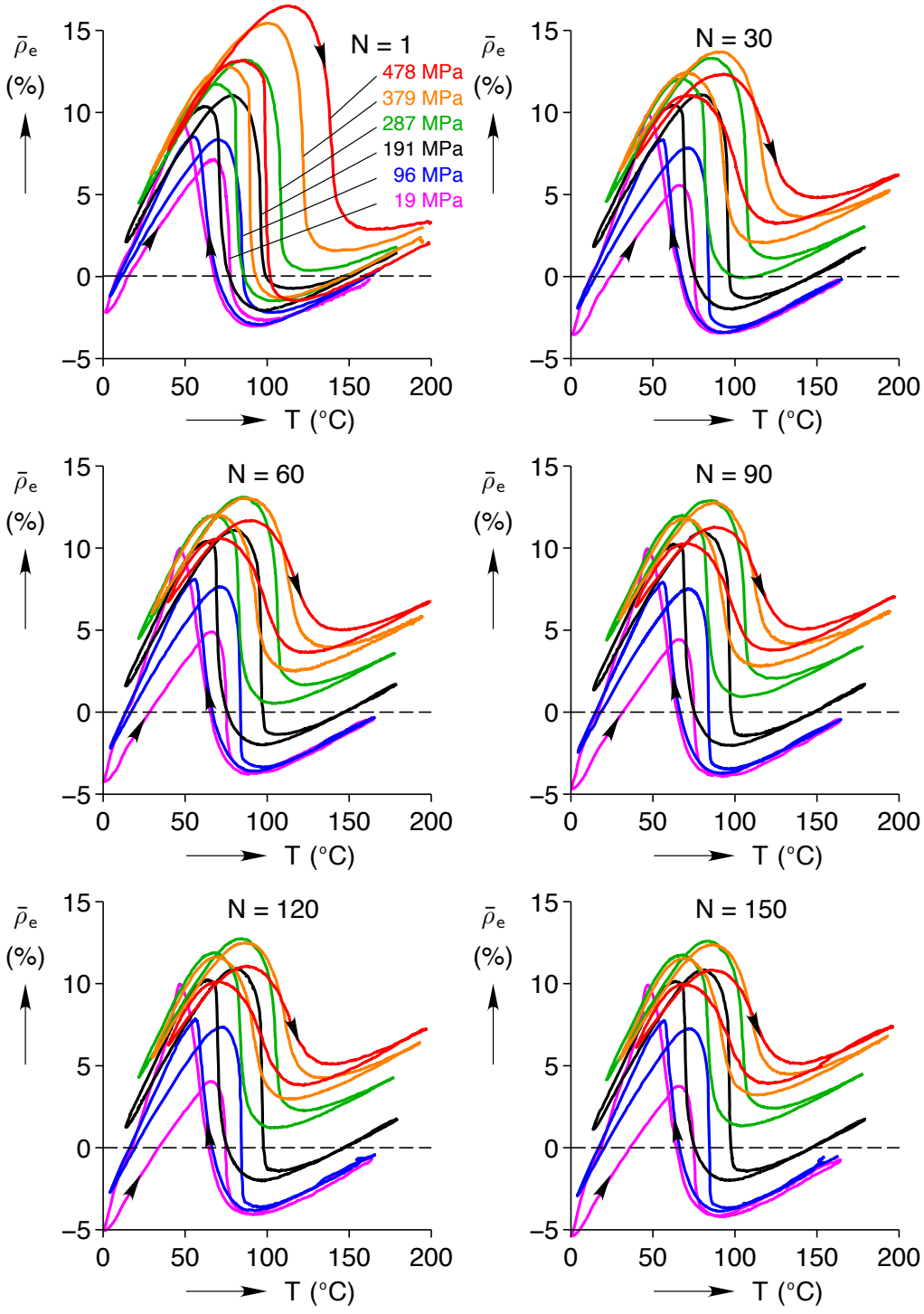


Figure 3.20: Relative resistivity $\bar{\rho}_e$ responses for all constant load shakedown experiments at 30 cycle intervals.

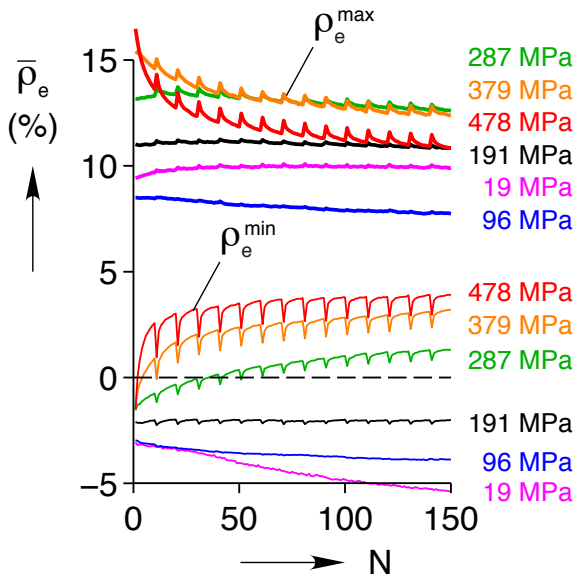


Figure 3.21: Relative resistivity $\bar{\rho}_e$ maximums and minimums for six constant load shakedown experiments.

The resistivity responses were also plotted against strain. Figure 3.22 shows the details of the resistivity evolution for every 10 cycles for the 379 MPa experiment. During the first cycle, relatively little hysteresis is apparent, except outside the transformation region where the single phase temperature coefficient dominates the response while the strain saturates. In fact, for this high load level the response is linear and nearly path independent within the transformation region, i.e., between about 1.5% and 4% strain. During cycling, however, the loop maintains the same qualitative shape but progresses toward the right (strain ratcheting) and the minimum $\bar{\rho}_e$ rises while the maximum $\bar{\rho}_e$ falls, resulting in a smaller overall change during the last cycle.

Figure 3.23 shows the $\bar{\rho}_e$ vs. ε_e responses at 30 cycles intervals for all six constant loads. Comparing the different load responses during the first cycle ($N = 1$), the two highest load responses have a similar shape and a restricted strain region where they nearly overlap. The lowest load, 19 MPa, response is quite different and has a considerably larger hysteresis

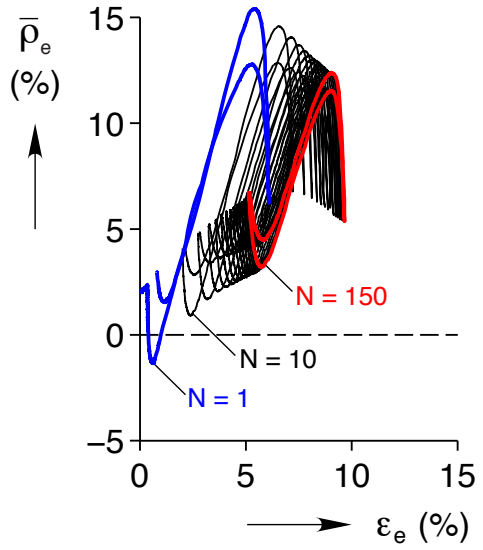


Figure 3.22: Relative resistivity $\bar{\rho}_e$ vs. strain ε_e for the 379 MPa experiment, showing every tenth cycle.

loop, again to the effect of the R-phase during cooling (corresponding to increasing strain in the figure). Considering the evolution with increasing cycles, at 191 MPa, the first and last cycle responses are rather similar with only a small downward shift in the loop. During cycling, the 19 MPa response gets "squashed" toward the left in the figure, mostly due to the reduction in the two-way shape memory effect. Generally, higher loads above 191 MPa have increasingly greater residual strains, so responses shift to the right on the strain axis. The result is that the higher the load, the greater the shift from the overlapping linear relationship during transformation. However, despite the response shift towards increasing ε_e, the slope $d\bar{\rho}_e/d\varepsilon_e$ during transformation remained the same.

The relationship between ε_e and $\bar{\rho}_e$ is often of particular interest to engineers wishing to use electrical resistivity to sense strain, i.e. to measure the position of an SMA actuator. Recall from previous plots that ε_e increased during cooling and decreased during heating. Figure 3.23 shows a roughly linear relationship during transformation in a restricted strain

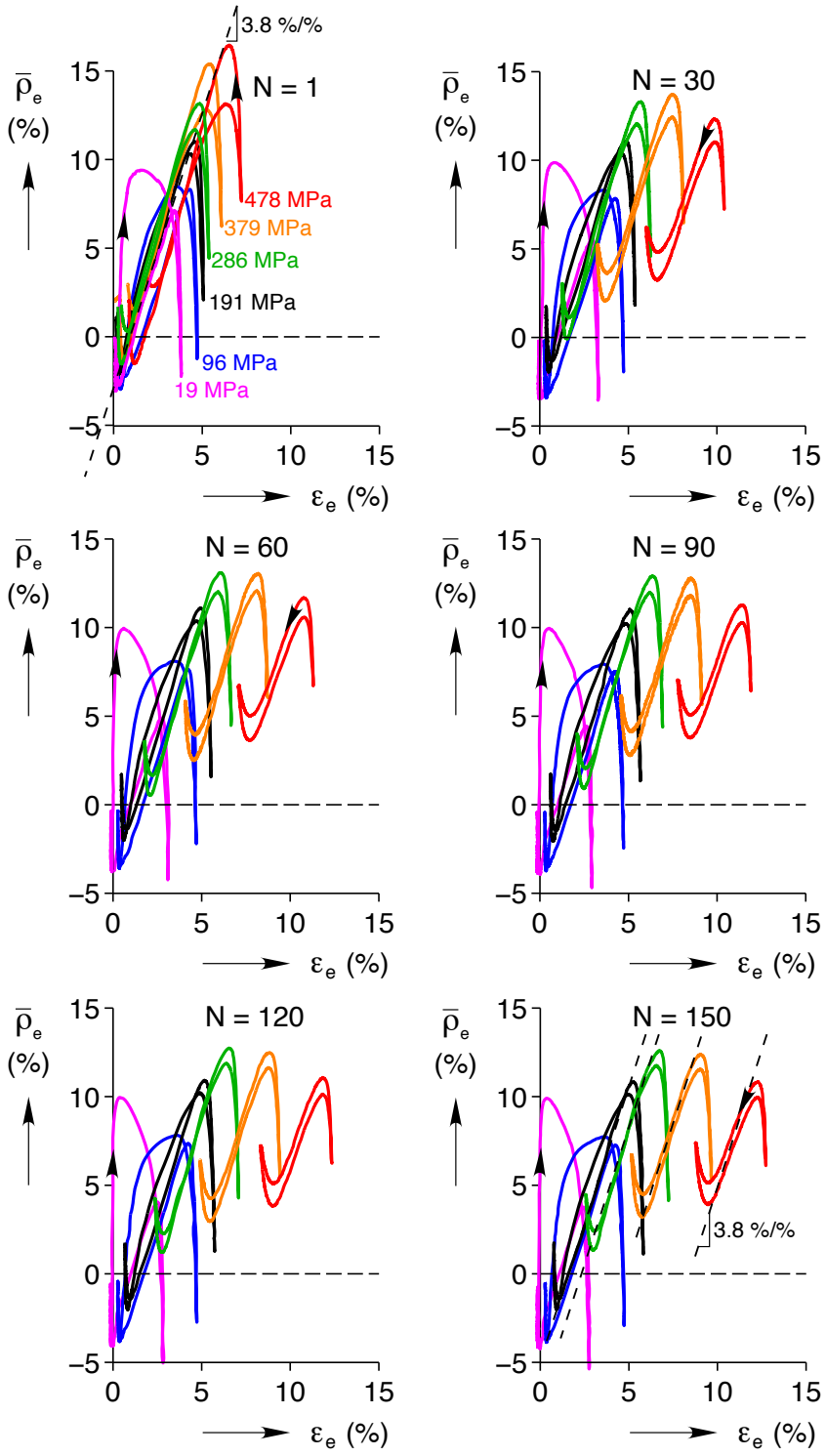


Figure 3.23: Relative resistivity $\bar{\rho}_e$ vs. strain ε_e responses for six constant load shakedown experiments at 30 cycle intervals

range ($1\% < \varepsilon_{\mathrm{e}} < 5\%$), at least for the higher loads. The slope of this line is about $d\bar{\rho}_{\mathrm{e}}/d\varepsilon_{\mathrm{e}} = 3.8$ (dotted lines in Figure 3.23, $\mathrm{N} = 1$).

There were of course several exceptions to the aforementioned relationship. First, there was the strong temperature dependence of $\bar{\rho}_{\mathrm{e}}$. Unlike the small effect of the coefficient of thermal expansion on strain, the temperature coefficient for resistivity $\bar{\alpha}_{\mathrm{e}}$ was definitely not small in the range of normal SMA operation. As Figure 3.20 showed, for temperatures above and below the transformation region, $\bar{\rho}_{\mathrm{e}}$ continued to rise linearly with temperature. In this temperature region, ε_{e} also continued to evolve, but at a much lower rate. The effect in Figure 3.23 is that the ends of each curve have a "tail" where $\bar{\rho}_{\mathrm{e}}$ continued to evolve while ε_{e} remained relatively stationary, thus eliminating any possibility of a one-to-one mapping of resistivity to strain at low strains or high strains. Secondly, at low loads where one might want to operate an actuator to minimize shakedown effects, the R-phase complicates matters. At 19 MPa during cooling, ε_{e} and $\bar{\rho}_{\mathrm{e}}$ were uncorrelated with the two measurements changing rapidly at different times, first $\bar{\rho}_{\mathrm{e}}$, followed by ε_{e}. As the load rose to 96 MPa, the effect of the R-phase still caused $\bar{\rho}_{\mathrm{e}}$ and ε_{e} to be uncorrelated, though about half as much as at 19 MPa. At 191 MPa, the effect was smaller yet, and above that load it was gone, since R is only stable at low stresses. During heating, this "problem" was absent and ε_{e} and $\bar{\rho}_{\mathrm{e}}$ were better correlated, just as it was on the heating side of the thermogram in Figure 3.1(a) where the multistep transformation was absent, so the relationship was linear during transformation for every experiment. For each case, there was a simultaneous drop in $\bar{\rho}_{\mathrm{e}}$and ε_{e}, marking a rapid $M^{+} \rightarrow A$ transformation.

3.5 Shakedown Experiments Under Alternating Constant Loads

The experiments presented in section 3.4 repeatedly suggested that there were differences in shakedown behavior between heating and cooling. Specifically, phase transformation during heating, $M^+ \rightarrow A$, had the greatest effect on shakedown. In order to investigate this premise in more detail, a pair of experiments were performed with alternating loads on heating and cooling. The first load path, referred to as "CCW," is shown in Figure 3.24(a), where each cycle forms a counter-clockwise square in stress-temperature space, with a load of 379 MPa during cooling and 96 MPa during heating. Experiment "CW," follows the clockwise path in Figure 3.24(b). As Figure 3.24 shows, the SMA wire in each experiment (assuming equal stroke) undergoes the same magnitude of mechanical work, though of opposite sign. Behind each of the load paths (red lines) is the idealized phase diagram from Figure 3.2, with expected transformation points marked with black circles. The hysteresis in temperature T_{hyst} is expected to be much larger with the CW experiment.

3.5.1 Counter-clockwise Experiment

The strain-temperature response of the CCW experiment is plotted in Figure 3.25(a) for 150 cycles (every 10^{th} shown). The first cycle (in blue) started at a high temperature, as usual, on the top right corner of the curve, at a stress of 379 MPa. It was then cooled under a constant load, and underwent $A \rightarrow M^+$ transformation at about 91 °C, rising from 0.62 to 6.38%. When it reached its low temperature of 5 °C, the stress was lowered to 96 MPa, causing ε_{e} to drop accordingly to 5.61%. Upon heating, the reverse transformation was centered about 89 °C as ε_{e} recovered to 0.26%.

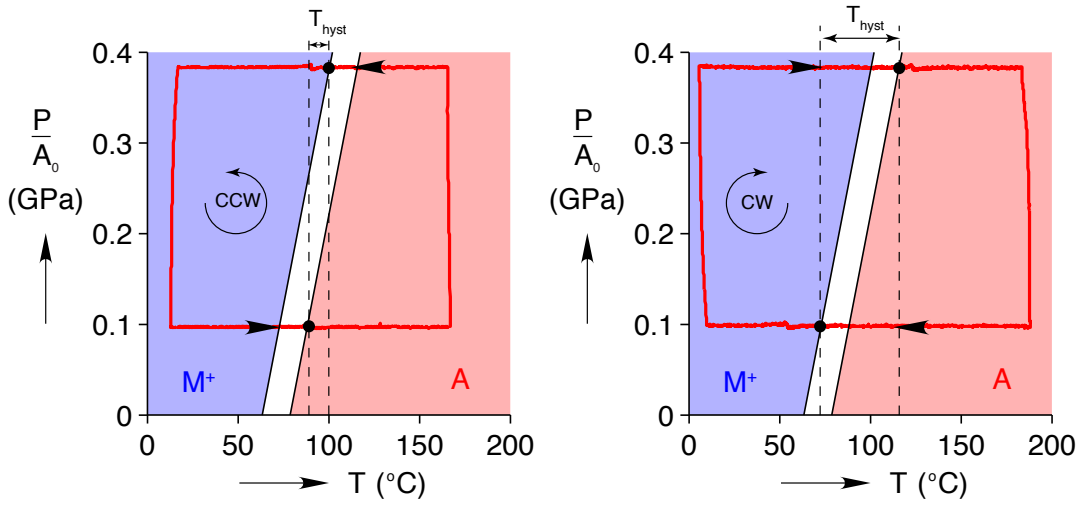


Figure 3.24: Two experiments traversing similar square-wave stress-temperature load paths: (a) CCW (b) CW

Recall from Figure 3.2 that transformation temperatures rise with stress along a Clausius-Clapeyron slope of about 10%/°C. Since the higher stress of 379 MPa is applied during cooling, this will tend to raise the $A \rightarrow M^+$ transformation temperature, reducing the temperature hysteresis. In this case, the difference in stress is 283 MPa. We would therefore expect a drop in the temperature hysteresis of 28 °C, versus the constant load case. This is just the case, in fact the cooling transformation during the first cycle was at a *higher* (albeit 1 °C) temperature than on heating. Another effect was that the strain response proceeded in a CCW fashion, contrary to constant-load behavior. Over 150 cycles, there is relatively little shakedown in the shape of the ε_e response. There is a very small increase in the heating transformation temperature, about 2 °C, as well as a slightly larger 5 °C increase in the cooling, $A \rightarrow M^+$ temperature. Changes in the strain extremes will be addressed in Figure 3.28(a).

The relative resistivity response $\bar{\rho}_e$ in Figure 3.25(b) followed much the same pattern

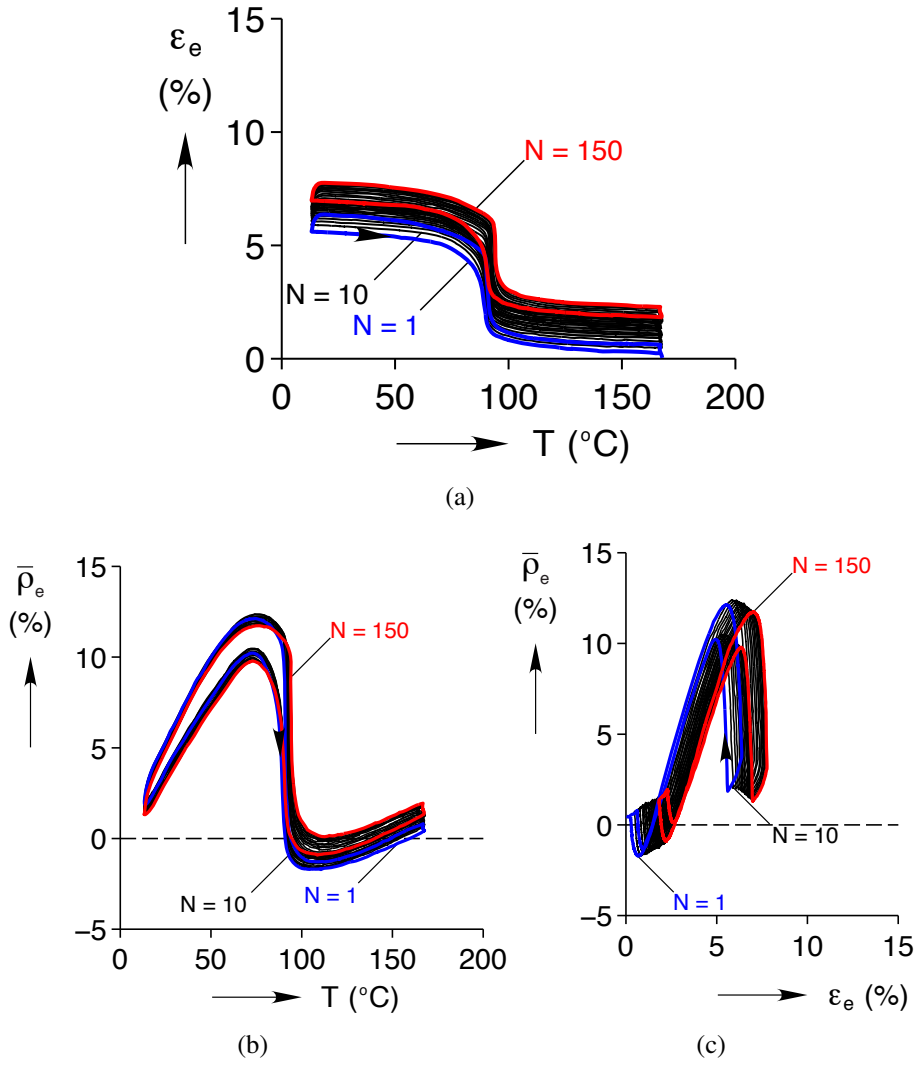


Figure 3.25: Responses during 150 counter-clockwise temperature-stress cycles (every tenth cycle shown):a) strain - temperature, b) relative resistivity - temperature, c) relative resistivity - strain.

as with ε_e, following a counter-clockwise path in $\bar{\rho}_e$-T space and having little temperature hysteresis between the heating and cooling curves. Figure 3.25(c) confirms this, as there was a linear relationship, both during heating and cooling, between ε_e and $\bar{\rho}_e$ during transformation. Also similar to ε_e, there was very little change in the shape of the $\bar{\rho}_e$ response over 150 cycles, just a slight (identical) shift in the apparent transformation temperatures (Figure 3.25(b)) and some shifting to larger strains (Figure 3.25(c)).

3.5.2 Clockwise experiment

The second experiment, as outlined in Figure 3.24(b), was performed under the same loads, but in a clockwise direction in stress-temperature space. The strain response in Figure 3.26(a) is remarkably different from the previous experiment. Looking only at the first cycle (in blue), the wire started at 190 °C at a stress of 96 MPa. ε_e stayed near its minimum of 0.6% until it rose with a point of inflection at 57 °C to a maximum of 4.46% at its minimum temperature of 8 °C. The stress was then increased until it reached 379 MPa, meanwhile ε_e increased to 5.88%. Assuming an elastic modulus of 28 GPa[a], one would expect a corresponding strain increase of only 1% instead of the observed increase of 1.42%. The remaining 0.42% strain was likely the result of permanent deformation (plasticity) or additional detwinning of martensite. On heating, the reverse transformation was centered around 123 °C and ε_e decreased to 0.9%. Note that reversing the load path had the opposite effect on temperature hysteresis, increasing it to 66 °C, based on the temperature of steepest $\partial\varepsilon_e/\partial T$. Also, the path in ε_e-T space was a clockwise loop. It is plain from the large sep-

[a]Unloading modulus from the 70 °C experiment in Figure 3.1(b)

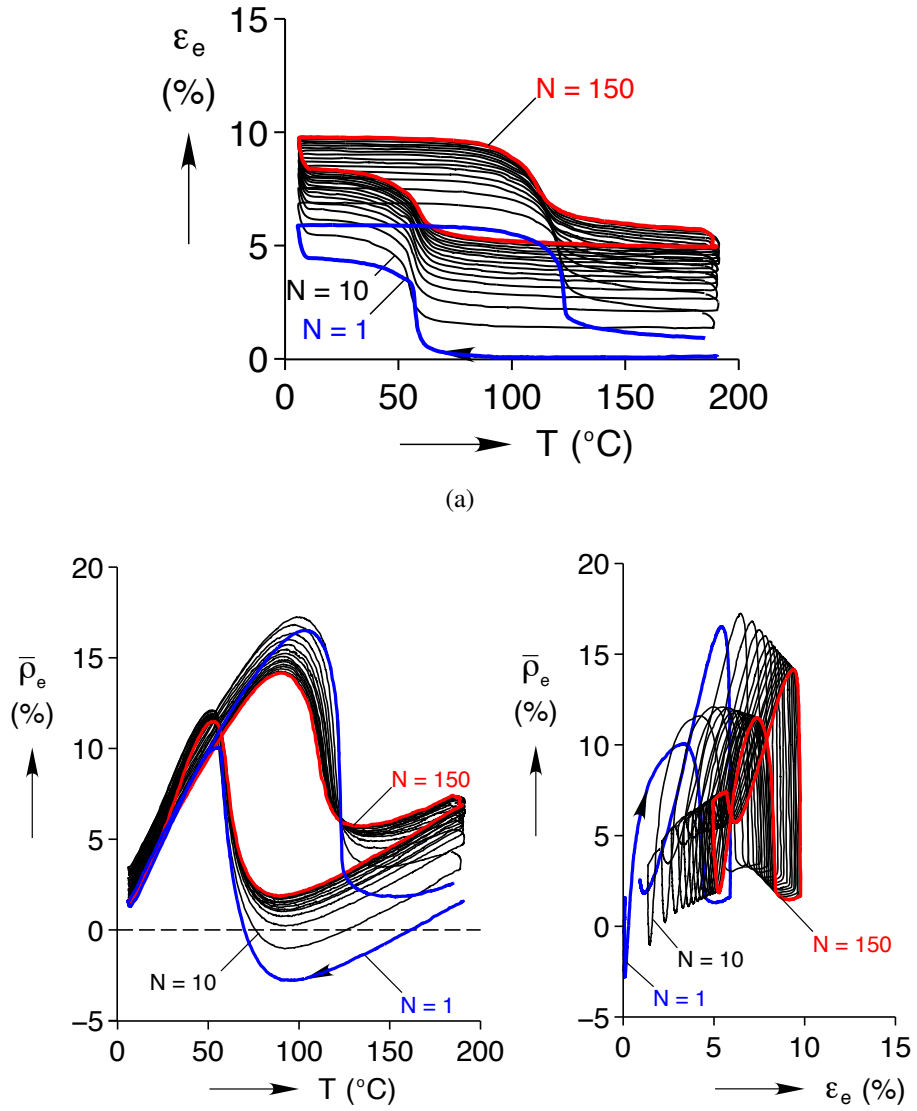


Figure 3.26: Responses during 150 CW temperature-stress cycles (every tenth cycle shown):a) strain - temperature, b) relative resistivity - temperature, c) relative resistivity - strain.

aration between the first and tenth cycle ε_{e} responses that shakedown progressed rapidly, much faster than for the CCW experiment. In the heating response especially, the transformation became broader and the onset took place at progressively lower temperatures with cycling.

The CW $\bar{\rho}_{\mathrm{e}}$ response in Figure 3.26(b) has several important features. First, since the load during cooling was only 96 MPa, there was a sharp peak in $\bar{\rho}_{\mathrm{e}}$ at 57 °C, a byproduct of the intermediate R-phase in the $A \rightarrow R \rightarrow M^{+}$ transformation (see section 3.1.1). As the material was cycled, this peak actually grew higher, from 10.0% to 11.5%, and the response following the peak (the colder side) began to separate from the adjacent heating $\bar{\rho}_{\mathrm{e}}$ response. As with other experiments involving rapid shakedown, such as Figure 3.16, $\bar{\rho}_{\mathrm{e}}$ increased more at the hot end (5.3%) vs. the cold end (0.16%). The rapid shakedown in strain response, coupled with the influence of R during cooling, results in an weirdly nonlinear relationship between $\bar{\rho}_{\mathrm{e}}$ and ε_{e}, show in in Figure 3.26(c). The only linear portion occurs during transformation upon heating.

After this initial look, these two experiments suggest that besides the role of stress and temperature, transformation direction has a dramatic effect on shakedown. The difference in response between the CW and CCW experiments suggest that the $M^{+} \rightarrow A$ transformation is responsible for a greater portion of the shakedown response than the reverse transformation, either $A \rightarrow M^{+}$ or $A \rightarrow R \rightarrow M^{+}$. The difference is also apparent in the maximum and minimum strains are compared. So that a fair comparison could be made to constant-load experiments, Figure 3.27 shows how both $\varepsilon_{\mathrm{e}}^{\max}$ and $\varepsilon_{\mathrm{e}}^{\min}$ are defined. The strain extremes are measured at 379 MPa only, disregarding extremes measured at the

alternate load of 96 MPa.

Figure 3.28 compares those strain extremes, the most striking of which is that there is a very close match in $\varepsilon_{\mathrm{e}}^{\max}$ between the constant-load experiment and CW experiment. The match is nearly as good between the two in $\varepsilon_{\mathrm{e}}^{\min}$, as both have the same shape, though the CW $\varepsilon_{\mathrm{e}}^{\min}$ increases slightly faster, ending about 0.6% higher. The CCW experiment did show some increases in $\varepsilon_{\mathrm{e}}^{\max}$ and $\varepsilon_{\mathrm{e}}^{\min}$but they were smaller, on the order of 1.5% (versus 5% for the CW and constant-load experiments). Another feature of note is the sawtooth pattern in $\varepsilon_{\mathrm{e}}^{\min}$ resulting from the 19 MPa thermal cycle after every 10 "normal" cycles. It is apparent in the CCW and constant-load experiments but *not* in the CW experiment. This is a curious finding, since the CW experiment had the higher load during $A \rightarrow M^{+}$ transformation (where the sawtooth pattern appears) but actually experienced less reverse shakedown. Additional experiments would be necessary to fully understand this phenomenon.

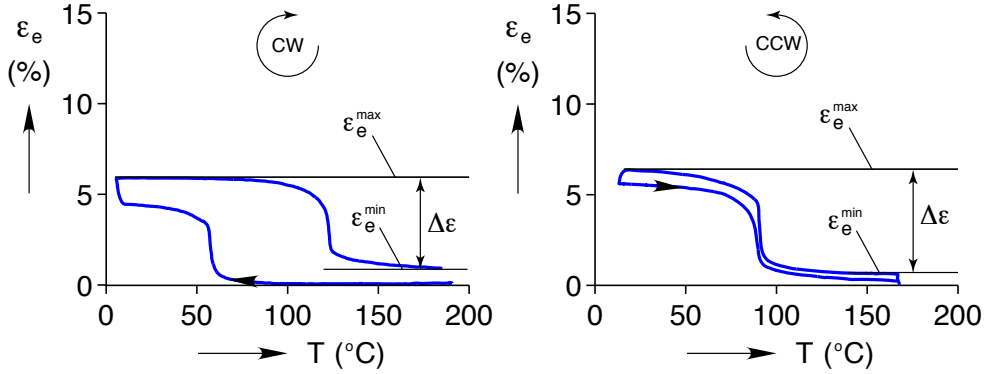
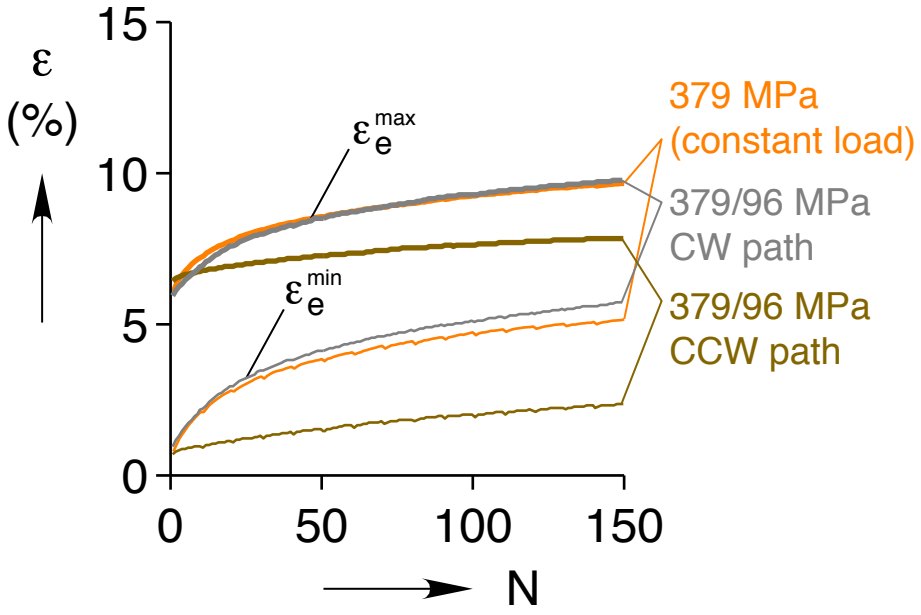
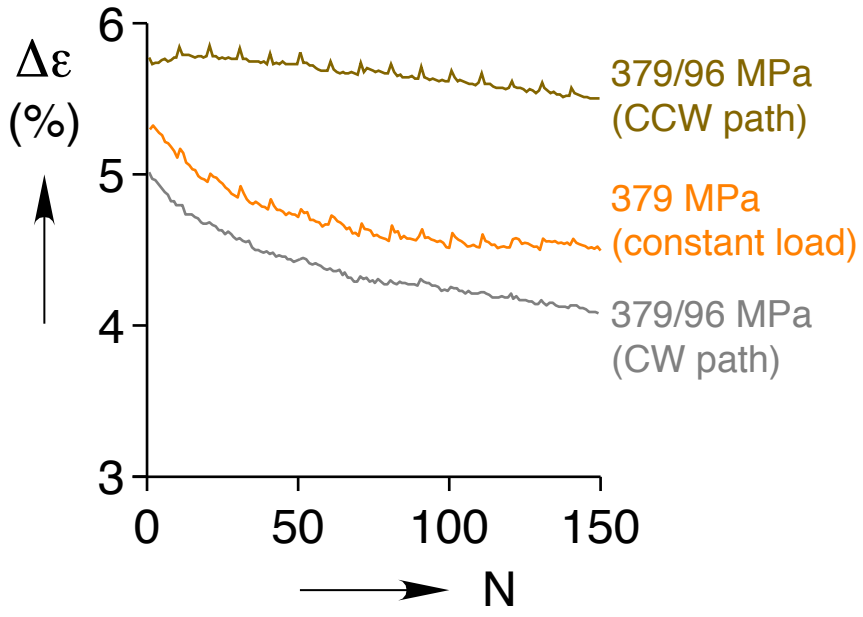


Figure 3.27: Definition of $\varepsilon_{\mathrm{e}}^{\max}$ and $\varepsilon_{\mathrm{e}}^{\min}$ for experiments CW and CCW involving alternating loads.



(a)



(b)

Figure 3.28: Comparison of strain extremes between three cyclic load paths: a) ε_e^{max} (bold lines) and ε_e^{min} vs. cycle number, b) stroke $\Delta\varepsilon_e$ vs. cycle number.

3.6 Strain Evolution With Dead Load Cycles

Observing the exponential-like trends in the responses in Figure 3.18(a), a simple first attempt to model shakedown kinetics is to fit the data to a Prony series of the form

$$f(N,\sigma) = a_0(\sigma) + \sum_{i=1}^{n} a_i(\sigma)(1 - e^{-N/\tau_i}), \tag{3.6.1}$$

where N is the number of thermal cycles, τ_i are the characteristic time-like constants, and a_i are stress dependent constant coefficients. A solution of this form arises naturally from a superposition of independent linear first-order processes each of which is Maxwellian. A similar method has been used to model relaxation and damage in polymers (see, for example, [91]) and to analyze shakedown in NiTi wires with a strain-limiting hard stop [93]. A remarkably good fit is obtained (dashed lines) using only $n = 2$ for the evolution of both $\varepsilon_{\mathrm{e}}^{\min}$ and $\varepsilon_{\mathrm{e}}^{\max}$ (solid lines) for the 3 highest loads (to reduce R-phase effects) as shown in Figure 3.29, using the respective values in Table 3.1. The fits were obtained by manually choosing time constants τ_i (from a series of trial fitting attempts), then using a least-squares algorithm to find the ideal coefficients a_i.

The limiting, asymptotic values for a large number of cycles were then estimated from these fits as the sum of the coefficients a_i.

$$f_\infty(\sigma) = \lim_{N\to\infty} f(N) = \sum_{i=0}^{n} a_i(\sigma), \tag{3.6.2}$$

which are also shown in Figure 3.29 as dotted lines. Unfortunately, not enough experimen-

tal cycles were performed for the data to actually reach the asymptotes. Using the time constants in Table 3.1, for exmple, it would take at least 300 cycles for the fit to reach 99 % of the asymptote $f_\infty(478\ \mathrm{MPa})$. The lack of sufficient cycles leaves open the possibility that the material will never reach an constant asymptote, instead ratcheting until failure like some of the NiTiCu specimens in [60], or that a third time constant $\tau_3 > 200$ could be necessary to model the long-term shakedown of the material.

Interestingly, the coefficients a_i seem to have a rough dependence on the constant stress σ. Figure 3.30 plots each of the four series of coefficients against σ, revealing that the a_1 has a slope of 11.6 %/GPa for both $\varepsilon_{\mathrm{e}}^{\min}$ and $\varepsilon_{\mathrm{e}}^{\max}$. If these simple relationships hold for additional loads, it could simplify modeling of shakedown. The quality of the fit, using only two rate constants one decade apart, suggests two different damage (relaxation) processes are in play: one that saturates early, on the order of 30 cycles, and the other that saturates later, at 300 cycles or more. Possible mechanisms for this damage, in order of length scale, are damage (i.e. plastic slip) near $\mathrm{Ni_4Ti_3}$ precipitates, damage at grain boundaries, and damage at habit plane boundaries.

Table 3.1: Fit parameters for $\varepsilon_{\mathrm{e}}^{\min}$ and $\varepsilon_{\mathrm{e}}^{\max}$

f (%)	Stress (MPa)	a_0 (%)	a_1 (%)	a_2 (%)	τ_1	τ_2	f_∞ (%)
$\varepsilon_{\mathrm{e}}^{\min}$	286	0.507	2.39	0.31	95	9.2	3.2
$\varepsilon_{\mathrm{e}}^{\min}$	379	0.580	3.55	1.81	95	9.2	5.94
$\varepsilon_{\mathrm{e}}^{\min}$	478	1.870	4.62	3.14	95	9.2	9.63
$\varepsilon_{\mathrm{e}}^{\max}$	286	5.657	1.88	0.50	95	9.2	8.04
$\varepsilon_{\mathrm{e}}^{\max}$	379	6.015	2.84	1.38	95	9.2	10.24
$\varepsilon_{\mathrm{e}}^{\max}$	478	7.386	4.11	2.02	95	9.2	13.52

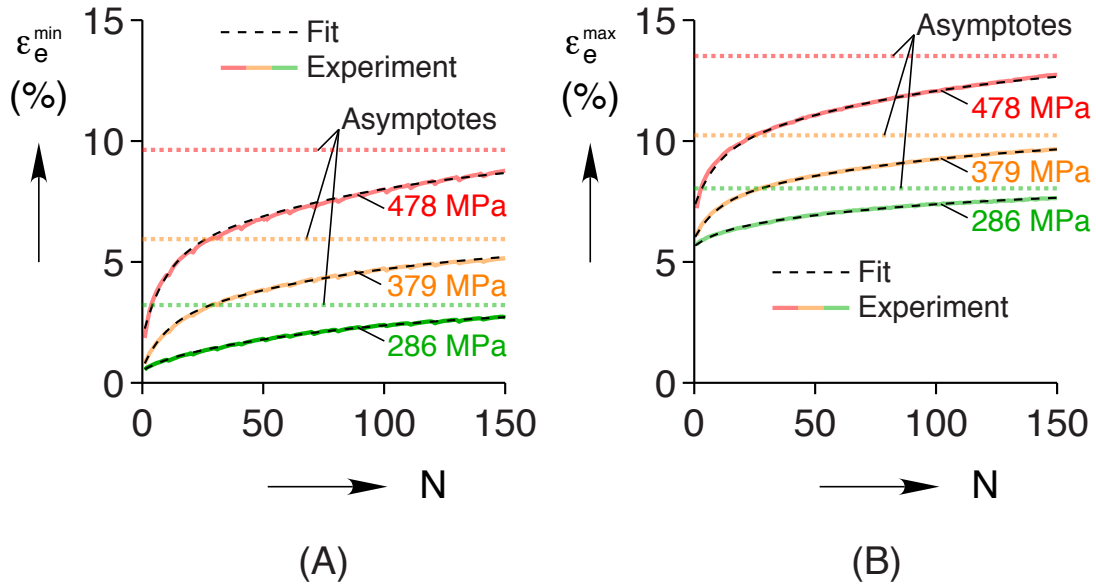


Figure 3.29: (A) maximum and (B) minimum strain responses (bold lines) with two-term Prony series fits (dashed lines) and their asymptotes.

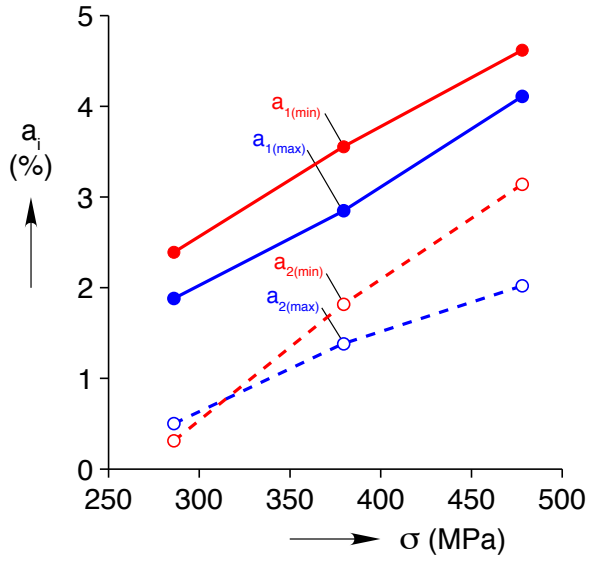


Figure 3.30: The values for coefficients a_1 and a_2 from Table 3.1 have roughly linear dependence on the stress σ

3.7 Conclusions

A systematic dead-load shakedown study was performed on preconditioned, commercially-available NiTi-based wire (90C Flexinol™from Dynalloy, Inc.). Relationships between temperature, strain, electrical resistivity, and the number of shakedown cycles were examined.

At each constant load, a wire specimen was subjected to 150 thermal cycles. A specialized experimental setup enabled precise, uniform temperature control without sacrificing the ability to perform reasonably rapid (± 1 $^{\circ}\mathrm{C}/\mathrm{s}$) thermal cycling across a 180°C temperature range. The setup was used to measure both resistance and elongation in the same axial gauge length. After every tenth cycle, a single low-stress cycle was performed at 19 MPa with a momentary load-unload to measure the evolution of residual strain. The least amount of shakedown occurred at the load level of 191 MPa that corresponds to the recommended maximum design load of the material supplier (20 N). Shakedown effects were evident both at lower loads (reverse sense) and higher loads (strain ratcheting), although they were significantly more severe at the higher loads. A simple 2-term Prony series fit quite well the cyclic evolution of minimum and maximum strain responses at the three highest loads, using the same cyclic rate constants (95 and 9.2 cycles) across multiple loads. Two of the three constant coefficients (a_1 and a_2) exhibited a rough linear relationship with the constant load.

It was shown that low-load cycling exhibited reverse shakedown, in effect undoing some of the conditioning imposed on the specimen by the manufacturer. For example, at 19 MPa, the magnitude of the inherent two-way shape memory effect was reduced over

150 cycles from 3.7 to 2.9% and the residual strain was -0.2%. Thus, it is important to note that preconditioned wires do not necessarily increase in length for every cyclic load path, and a degradation in stroke can occur even at very low loads.

In several cases, shakedown measurements strongly suggested that specimen properties changed predominantly at high temperatures during transformation to the austenite phase. First, temperature hysteresis in both ε_{e} and $\bar{\rho}_{\mathrm{e}}$, for loads above 191 MPa, changed through a downward shift of the $M^{+} \rightarrow A$ transformation temperature (the hot side), while the $A \rightarrow M^{+}$ temperature remained unchanged. Second, the strain vs. temperature responses showed how the first cycle, after being offset to account for residual strains, matched the response at cycle $N = 150$ at low temperatures. Finally, after each low-stress cycle, there was a "sawtooth" effect on the high-temperature $\varepsilon_{\mathrm{e}}^{\min}$ but no effect on $\varepsilon_{\mathrm{e}}^{\max}$, measured at the lowest temperature.

These features were further investigated through a pair of similar experiments using counter-clockwise (CCW) and clockwise (CW) load paths in stress-temperature space. The CW experiment, with a high load only during heating, experienced very similar shakedown in $\varepsilon_{\mathrm{e}}^{\max}$ and $\varepsilon_{\mathrm{e}}^{\min}$ to the shakedown of the 379 MPa constant load experiment. Somewhat suprisingly, shakedown in the CCW experiment was much less sever, and experienced less than half the change in $\varepsilon_{\mathrm{e}}^{\max}$ and $\varepsilon_{\mathrm{e}}^{\min}$and relatively little change in the overall shape of the response. Electrical resistivity responses were examined as well. The strong temperature dependence of $\bar{\rho}_{\mathrm{e}}$, and the existence of R-phase at low loads, caused the response to be quite non-monotonic and path dependent. Significant changes in electrical resistivity responses were also measured during shakedown cycling.

While it seems that one could use resistivity as an indicator of strain for actuator application, the approach is fraught with complications and should only be attempted under a restrictive set of conditions:

- The temperature should not rise too high above or below the transformation region.
- The stress should be kept high enough to avoid the appearance of R-phase which de-corellates the resistivity and strain signals during cooling.
- The stress should be kept low enough to minimize shakedown effects and strain ratcheting.

Chapter IV

A Reduced-Order Thermomechanical Model for an SMA Bias Spring Actuator

The previous chapter focused on two important SMA phenomena: electrical resistivity response and shakedown. As important as these phenomena are, our ability to model them accurately is at best limited, and always complicated. At the beginning stages of device design, the engineer needs fast answers to more simple questions, like "how much power will my device use," and "how long will it take to actuate?" This thesis will now take a step in the opposite direction and seek to simplify an SMA model as much as possible, removing all features except those required to model the simplest prototype actuator. The result is a thermomechanical model which has an analytical solution, eliminating altogether the need for a computationally and intellectually expensive finite element implementation.

4.1 Nomenclature

Independent variables.

t	time
τ	dimensionless time
x	reference axial coordinate

SMA phases.

A	Austenite
M	Martensite
M^+, M^-	tensile–M, compressive-M
$M^{+/-}$	thermal-M

SMA state variables.

ε	nominal axial strain
T	absolute wire temperature
ξ_1, ξ_2, ξ_3	mass fractions M^+, M^-, A

SMA thermodynamic variables.

ϕ	specific Helmholtz free energy
σ,	nominal axial stress
s	specific entropy
μ, μ_1	thermo-driving force $A \to M^+$
μ_2	thermo-driving force $A \to M^-$
q_x	axial heat flux
A_s, A_f	A start, finish temperatures
M_s, M_f	M^+ start, finish temperatures

SMA physical, mechanical, & electrical constants.

L, d	reference wire length, diameter
A	reference cross–sectional area
ρ	mass density
E	elastic modulus
β	transformation strain
c_I	phase interaction parameter
ρ_e	electrical resistivity

SMA thermodynamic constants.

T_R	reference transformation temperature
Δs	equilibrium $A \to M$ entropy jump
H	enthalpy
c_o	specific heat
v_o, μ_c	kinetic law stiffness, critical driving force
σ_M	$M \to M^+$reorientation stress
K	reference thermal conductivity

Spring constants.

k_B, k_E	bias, external spring stiffnesses
Δ	bias spring-SMA length mismatch

Actuator, control & environmental variables.

F	SMA axial force
δ_o, ε_o, σ_o	SMA pre-elong., pre-strain, pre-stress
F_B, F_E	bias, external spring forces
δ_B	bias spring displacement
T_a, h	ambient temperature, film coefficient
$\mathcal{P}_\mathrm{e}$	electrical power input
t^*	characteristic time
$\mathcal{E}_\mathrm{in}$, $\mathcal{E}_\mathrm{out}$	energy input, external work output

Dimensionless actuator variables.

θ	SMA temperature
θ_a	ambient temperature
θ^ss	SMA steady-state temperature
$\bar{\sigma}$	SMA axial stress
f	SMA $A \to M^+$ driving force
$\bar{A}_\mathrm{s}$, $\bar{A}_\mathrm{f}$	A start, finish temperatures
$\bar{M}_\mathrm{s}$, $\bar{M}_\mathrm{f}$	M^+ start, finish temperatures
$\bar{\mathcal{P}}_\mathrm{e}$	electrical power
a_n, b_n	time interval constants during transformation
W	Lambert function
$\bar{\mathcal{E}}_\mathrm{in}$, $\bar{\mathcal{E}}_\mathrm{out}$	energy input, external work output

Dimensionless actuator constants.

λ, $\bar{c}$	SMA latent heat, specific heat
$\bar{\mu}_\mathrm{c}$	$A \to M^+$ critical driving force
$\bar{\sigma}_M$	$M \to M^+$reorientation stress
$\bar{\sigma}_\mathrm{o}$	SMA pre-stress
$\bar{\sigma}_\mathrm{max}$	maximum SMA stress
η_B, η_E	bias, external spring stiffnesses
η	spring ratio
$\bar{\Delta}$	bias spring mismatch
δ_E	external spring stroke

4.2 Thermomechanical Problem of an SMA Actuator

As defined in Chang, *et. al.* [15] the thermomechanical problem of interest involves a thin SMA wire of initial length, L, and diameter, d, under uniaxial tensile loading (see Figure 4.1). The reference configuration chosen for the SMA wire is unstressed austenite at a sufficiently high temperature (above the stress-free austenite finish temperature, A_{f}). We define a reference transformation temperature for the material, T_{R}, at a lower temperature where martensite and austenite are in thermodynamic equilibrium, near the average of the stress-free martensite start temperature, M_{s}, and austenite start temperature, A_{s}. The SMA wire is immersed in a thermal fluid bath (gas or liquid) with convective film coefficient, h, at an ambient temperature, T_a, that is sufficiently below T_{R} to cause the initial stress-free state of the SMA to be thermal martensite. The left end of the wire in Figure 4.1 is held fixed, and a bias spring of stiffness k_{B} and initial length mismatch to the SMA element of Δ (Figure 4.1a) is force-fit into place. The SMA prestress, in mechanical equilibrium with the bias spring force, results in an elongation of the SMA by an amount δ_{o} that is assumed to produce tensile martensite (a check for validity of this assumption is provided later). After assembly of the SMA/bias spring actuator, an external spring of stiffness k_{E}, representing the stiffness of whatever external object is to be actuated against, is attached in its natural (force-free) configuration (Figure 4.1b). Electrical power, $\mathcal{P}_{\mathrm{e}}$, can then be applied to resistively heat the SMA wire, causing it to revert to austenite and to contract against the action of both springs to some lesser displacement δ (Figure 4.1c). Subsequently, switching off the electrical power causes the SMA temperature to return to ambient temperature and elongate during transformation back to δ_{o}, tensile martensite M^+ (Figure 4.1d).

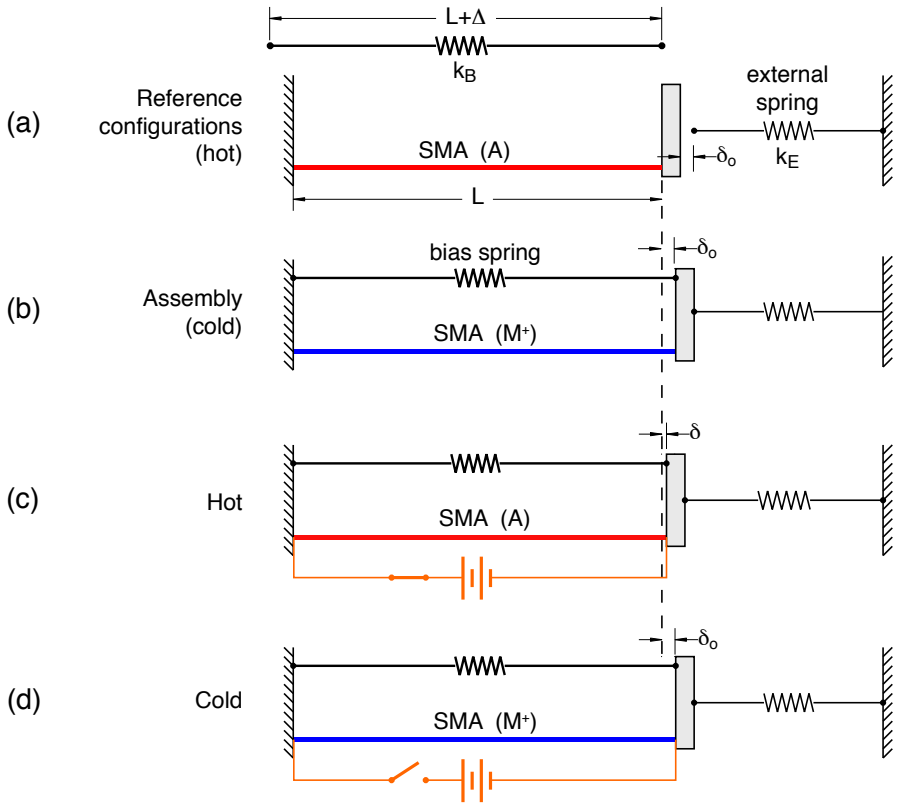


Figure 4.1: Prototype uniaxial SMA/bias spring actuator working against an external spring as electrical power is applied to the SMA element: (a) reference configurations for SMA element, bias spring, and external spring, (b) after cooling and assembly, (c) contraction of SMA element during Joule heating, (d) reset configuration when cold.

The following subsections present (1) a review of the full uniaxial thermodynamic model of Chang, *et. al.* [15], (2) some simplifying assumptions, and (3) the resulting lumped form of the governing equations.

4.2.1 The Full 1-D Constitutive Model

The constitutive model of Chang, *et. al.* [15] is the starting point for the derivation of the reduced-order, lumped SMA model, so it is briefly reviewed here for completeness. For clarity, we have already omitted strain gradient effects associated with transformation fronts, which are ignored in our current context. The state of each point in the SMA wire is determined by the strain field, $\varepsilon(x,t)$, the temperature field, $T(x,t)$, and the internal phase field vector, $\boldsymbol{\xi}(x,t) \equiv \{\xi_1, \xi_2\}$, for tensile ($M^+$) and compressive ($M^-$) variants of martensite, respectively. Constitutive relations are derived from the specific Helmholtz free energy,

$$\begin{aligned}\phi(\varepsilon, T, \boldsymbol{\xi}) =& \frac{E_{\mathrm{A}} + (\xi_1 + \xi_2)\Delta E}{2\rho} \left[\varepsilon - (\xi_1 - \xi_2)\beta\right]^2 \\ & - (T - T_{\mathrm{R}})(\xi_1 + \xi_2)\Delta s \\ & + c_I \left(1 - \xi_1 - \xi_2\right)\left(\xi_1 + \xi_2\right) \\ & + (c_{\mathrm{o}} - s_{\mathrm{o}})(T - T_{\mathrm{R}}) - c_{\mathrm{o}}\, T \ln(T/T_{\mathrm{R}}),\end{aligned} \tag{4.2.1}$$

where the separate lines on the right hand side are respective contributions from elastic energy, phase-dependent entropy, energy of mixing, and phase-independent thermal energy. The mass density is ρ, the elastic modulus of pure austenite is E_{A}, the difference between the effective martensite and austenite elastic moduli is $\Delta E \equiv E_{\mathrm{M}} - E_{\mathrm{A}}$ is (usually a neg-

ative constant), and the stress–free transformation strain is β (a positive material constant). Thermal expansion is neglected in the elastic energy, since its effect is usually small compared to the transformation strain. The specific entropy change from austenite to martensite is $\Delta s \equiv s_M - s_A$ (a negative constant), which is related to the latent heat of transformation. The free energy of mixing has a material constant, c_I, that affects the slope of the pseudoelastic transformation path. Lastly, $s_{\rm o}$ and $c_{\rm o}$ are material constants representing the reference specific entropy and specific heat, respectively.

The non–negativity requirement of the local entropy production leads to the Gibb's relations

$$
\begin{aligned}
s &= -\frac{\partial \phi}{\partial T} = s_{\rm o} + (\xi_1 + \xi_2)\,\Delta s + c_{\rm o} \ln(T/T_{\rm R}), \\
\sigma &= \rho \frac{\partial \phi}{\partial \varepsilon} = \left[E_{\rm A} + (\xi_1 + \xi_2)\Delta E\right]\left[\varepsilon - (\xi_1 - \xi_2)\beta\right], \qquad (4.2.2)
\end{aligned}
$$

where σ is the nominal axial stress.

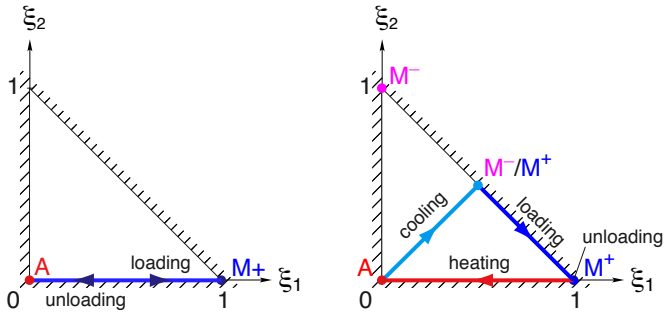


Figure 4.2: Admissible phase fraction space: (a) superelastic path in tension, (b) shape memory effect path.

Entropy production (consistent with the Coleman/Noll viewpoint) arises from the heat flux and phase transformation terms, and sufficient conditions to satisfy the second law of thermodynamics are

$$-q_x \frac{\partial T}{\partial x} \geq 0, \tag{4.2.3}$$

$$\boldsymbol{\mu} \cdot \frac{\partial \boldsymbol{\xi}}{\partial t} \geq 0, \tag{4.2.4}$$

where q_x is the axial heat flux, and $\boldsymbol{\mu} \equiv -\phi_{,\boldsymbol{\xi}}$ is the thermodynamic driving force for phase transformation conjugate to the rate of change of $\boldsymbol{\xi} \equiv (\xi_1, \xi_2)$. The inequality in (4.1.3) is easily satisfied by assuming the conventional axial heat conduction law $q_x = -K\,\partial T/\partial x$, where K is a positive thermal conductivity for the material. Written explicitly, the thermodynamic driving forces are

$$\mu_1 = -\frac{\partial \phi}{\partial \xi_1} = \frac{\beta\sigma}{\rho} - \frac{\Delta E}{2\rho}\left[\frac{\sigma}{E_{\mathrm{A}} + (\xi_1+\xi_2)\Delta E}\right]^2 + (T - T_{\mathrm{R}})\,\Delta s - c_I\,(1 - 2\xi_1 - 2\xi_2)\,, \tag{4.2.5}$$

$$\mu_2 = -\frac{\partial \phi}{\partial \xi_2} = -\frac{\beta\sigma}{\rho} - \frac{\Delta E}{2\rho}\left[\frac{\sigma}{E_{\mathrm{A}} + (\xi_1+\xi_2)\Delta E}\right]^2 + (T - T_{\mathrm{R}})\,\Delta s - c_I\,(1 - 2\xi_1 - 2\xi_2)\,. \tag{4.2.6}$$

The second inequality in (4.1.4) is satisfied if the "projection" of the driving force to its conjugate phase fraction rate ($\boldsymbol{\mu} \cdot \boldsymbol{\xi}$) is non-negative. A simple, piecewise linear, kinetic relation is chosen, according to

$$\frac{\partial \boldsymbol{\xi}}{\partial t} = v_{\mathrm{o}} < \boldsymbol{\mu} \cdot \boldsymbol{m} - \mu_{\mathrm{c}} > \boldsymbol{m} \tag{4.2.7}$$

where v_{o} is the "stiffness" coefficient of the kinetic law, $< x >$ is the Macauley discontinuity function ($< x >= x$ for $x > 0$ and $< x >= 0$ for $x <= 0$), and $\boldsymbol{m}$ is the unit vector in the direction of phase transformation in (ξ_1, ξ_2) space. The parameter, μ_{c}, is a positive constant that represents the critical thermodynamic driving force necessary for hysteretic phase transformation. Within the phase fraction space described above, the direction of phase transformation is assumed to be a unit vector colinear with the thermodynamic driving force ($\boldsymbol{m} = \boldsymbol{\mu}/\sqrt{\boldsymbol{\mu} \cdot \boldsymbol{\mu}}$), except at a boundary of the admissible phase region ($\xi_1 = 0$, $\xi_2 = 0$, or $\xi_1 + \xi_2 = 1$) where it is a unit vector tangent to the boundary to keep the phase fractions within the admissible region (see Figure 4.2).

4.2.2 Governing PDEs

Assuming traction–free lateral boundary conditions and no body forces, mechanical equilibrium requires that the axial force be constant along the length, or

$$\frac{\partial F(x,t)}{\partial x} = 0, \tag{4.2.8}$$

where the axial tensile force is simply $F \equiv \sigma A$. The heat equation, accounting for sensible and latent heat, axial heat conduction, lateral convection, and Joule heating, is

$$\rho A \left[c_{\mathrm{o}} \frac{\partial T}{\partial t} - (\boldsymbol{\mu} - T\boldsymbol{\mu}_{,T}) \cdot \frac{\partial \boldsymbol{\xi}}{\partial t} \right] = \frac{\partial}{\partial x} \left(KA \frac{\partial T}{\partial x} \right) - h\pi d(T - T_a) + \frac{\mathcal{P}_{\mathrm{e}}}{L}, \tag{4.2.9}$$

where h is the Fourier law coefficient (or convective film coefficient) for lateral heat transfer from the wire. The governing equations for the thermomechanical boundary value problem

are the field equations (4.1.8) and (4.1.9), subject to the kinetic law (4.1.7) at each point $x \in (0, L)$, along with appropriate mechanical and thermal boundary conditions at $x = 0, L$.

4.2.3 Simplified Model

In addition to the assumptions already inherent in Chang, *et. al.* [15] (no inertia effects, no R-phase, no thermal expansion, and no plasticity or shakedown effects), the following assumptions are made to simplify and reduce the system of partial differential equations to ordinary differential equations.

1. Spatial gradients ($\frac{\partial}{\partial x}(\cdot)$) of strain ($\varepsilon$), temperature ($T$), and phase fraction ($\boldsymbol{\xi}$) are neglected. This means that transformation is assumed to be uniform (no transformation fronts or boundary effects), and axial heat conduction is ignored. It is a reasonable assumption for wire with constant cross-sectional area (A) that has already been conditioned (trained) to have repeatable cyclic behavior, and for SMA wire that is thermally insulated at its ends. This is verified in Section 4.3.2 by comparison of the simplified model to a finite element simulation.

2. The SMA element is always under sufficient tension (after assembly) to avoid thermal martensite. In other words, $\mu_2 < 0$, is sufficiently negative to cause $\xi_2 = 0$ for all time $t > 0$. This assumption is valid if the bias spring constant and prestrain are sufficiently large, or the material has a sufficient two way shape memory effect (TWSME). A condition will be given later to check the validity of this assumption.

3. The elastic moduli of austenite and martensite are the same ($\Delta E = 0$). This assumption poses no significant restriction if unloading of the martensite are avoided during assembly and operation.

4. The mixing energy between austenite and martensite is neglected ($c_I = 0$). This assumption neglects any hardening or softening during superelastic transformation, which may be reasonable, or not, depending on the desired accuracy of the simulation.

5. The characteristic speed of phase transformation is fast compared to heat transfer rate and mechanical loading rate (i.e., we take the limiting case $v_{\mathrm{o}} \to \infty$). This assumption is reasonable for slow to moderate loading rates where displacement rates are small compared to the inherent velocity of martensitic transformations. Roughly speaking, inherent martensitic transformation propagation velocities are only an order of magnitude slower than elastic wave velocities, as measured in high rate impact experiments [29]. For Nitinol this should not be a serious restriction for moderate displacement rates, say $\dot{\delta} < 10^2$ m/s.

6. The electrical power $\mathcal{P}_{\mathrm{e}}$ and ambient temperature T_a are piecewise constant functions of time.

The state of the wire is now minimially defined by three time-dependent variables, the strain $\varepsilon(t)$, the temperature $T(t)$, and the single phase fraction $\xi(t)$ for tensile martensite ($\xi \in [0, 1]$). Under these assumptions, the specific Helmholtz free energy of the SMA

material reduces to

$$\phi(\varepsilon, T, \xi) = \frac{E}{2\rho}\left[\varepsilon - \xi\beta\right]^2 - (T - T_{\mathrm{R}})\xi\Delta s + (c_{\mathrm{o}} - s_{\mathrm{o}})(T - T_{\mathrm{R}}) - c_{\mathrm{o}}\, T \ln(T/T_{\mathrm{R}}), \quad (4.2.10)$$

and the following simplified constitutive relations for nominal stress (σ), entropy (s), and transformation driving force ($\mu = \mu_1$) are obtained.

$$\sigma = E\left[\varepsilon - \xi\beta\right] \quad (4.2.11)$$

$$s = s_{\mathrm{o}} + \xi\Delta s + c_{\mathrm{o}} \ln(T/T_{\mathrm{R}}) \quad (4.2.12)$$

$$\mu = \frac{\beta\sigma}{\rho} + (T - T_{\mathrm{R}})\Delta s \quad (4.2.13)$$

Figure 4.3(a) shows isothermal constitutive responses in stress-strain-temperature space predicted by the simplified model using typical SMA properties of Table 4.1, which were calibrated to the SE wire used in Chapter II (guideBB-30 from Memry Corp.). This simplified model has constant stress transformation paths which could also have been derived from a simple Gibb's mixture rule of the phases,

$$g(\sigma, T, \xi) = (1 - \xi)g_A(\sigma, T) + \xi g_M(\sigma, T). \quad (4.2.14)$$

Figure 4.3(b) shows a quasi-phase diagram in stress-temperature space, and Figure 4.3(c) shows a stress-free shape memory cycle along with an isothermal superelastic response. In fact, Figure 4.3(c) shows a two-way shape memory effect, since the simplified two-phase (A and M^+) model cannot capture a zero strain martensite that would be necessary to cap-

ture a one-way shape memory effect. The behavior of the simplified model, consequently, is more representative of a highly conditioned SMA that has two-way shape memory. If the actual SMA being used has a one-way shape memory, the model should not be used at certain low-stress/low-temperature regimes. A suitable check will be provided later in this paper to ensure the model's validity in the latter case.

4.2.4 Simplified Governing Equations

For this simplified model, the governing equations for the actuator's operation are algebraic equations for equilibrium and transformation kinetics and an ordinary differential equation in time for heat transfer. Neglecting inertial effects, equilibrium requires the axial force in the SMA wire to be uniform. It is balanced by the forces in the bias spring and external spring

$$F(t) = F_{\mathrm{B}}(t) + F_{\mathrm{E}}(t). \tag{4.2.15}$$

Accounting for the natural configurations of the springs and the SMA constitutive equation for stress (4.1.11), leads to the equilibrium equation for $t \geq 0$ (after assembly),

$$EA\left[\varepsilon - \beta\xi\right] = k_{\mathrm{B}}\left[\Delta - L\varepsilon\right] + k_{\mathrm{E}}L\left[\varepsilon_{\mathrm{o}} - \varepsilon\right], \tag{4.2.16}$$

Table 4.1: SMA parameters.

Mechanical	Thermal	Physical, Kinetic, Electrical
L=60 mm	$T_{\mathrm{R}} = 249.3$ K	$\rho = 6.5 \times 10^6$ g/m^3
$d = 0.762$ mm	$T_{\mathrm{R}}\Delta s = -16.74$ J/g	$\mu_{\mathrm{c}} = 1.01$ J/g
$\beta = 0.0591$	$c_{\mathrm{o}} = 0.5$ J/g·K	$\rho_{\mathrm{e}} = 1 \times 10^{-6}$ Ω·m
$E = 70$ GPa		

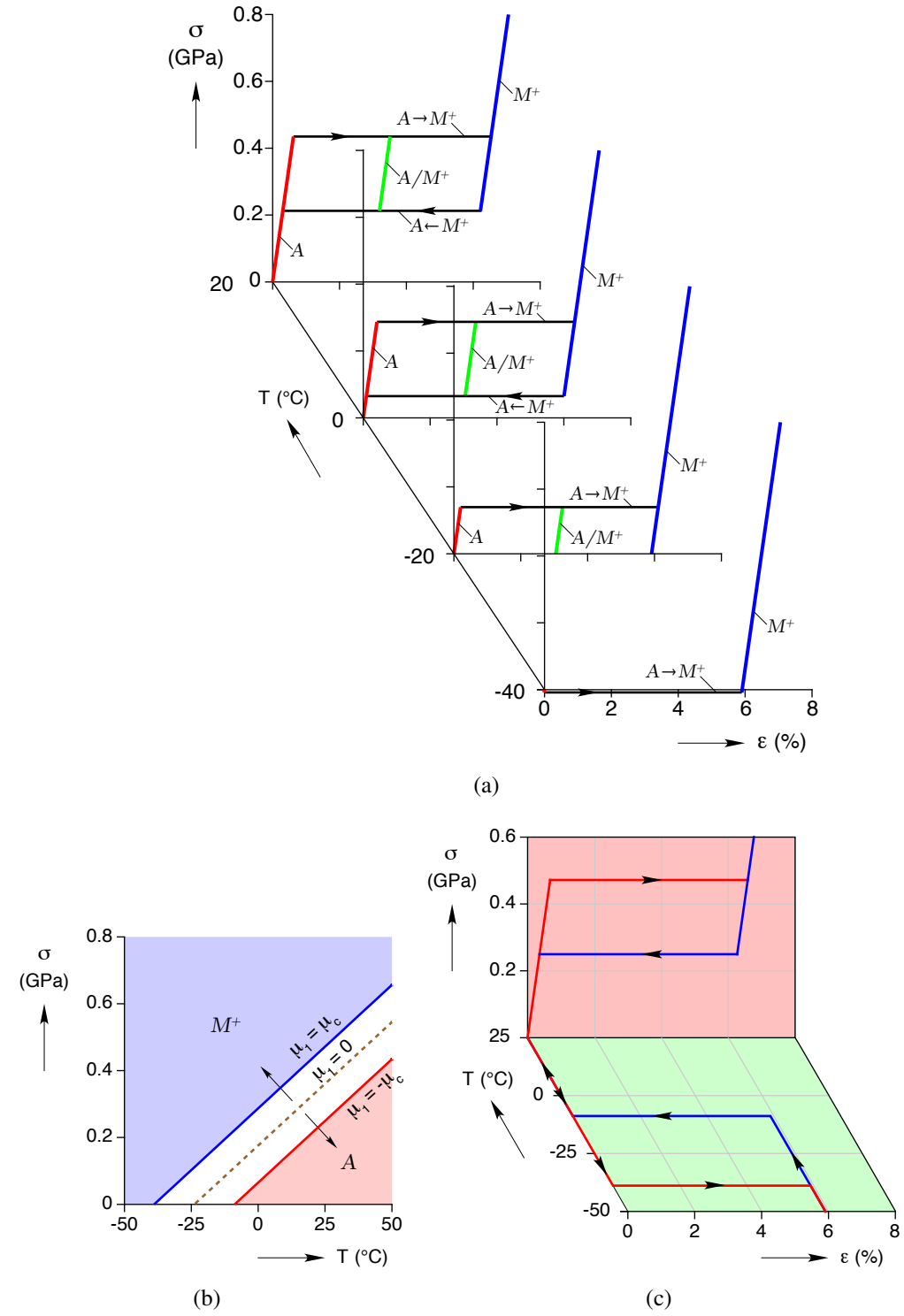


Figure 4.3: The simplified constitutive model: a) isothermal responses, b) transformation stresses, c) shape memory cycle ($\varepsilon - T$ plane) and superelastic response ($\sigma - \varepsilon$ plane).

where $\varepsilon_{\mathrm{o}} \equiv \delta_{\mathrm{o}}/L$ is the prestrain in the SMA wire caused during cold assembly. Equation (4.1.16) is an equation relating the unknowns $\varepsilon(t)$ and $\xi(t)$, where the explicit time dependence has been omitted for brevity. At $t = 0$, just after assembly, the external spring is not yet loaded and the initial condition is defined by

$$EA\left[\varepsilon_{\mathrm{o}} - \xi_o\beta\right] = k_{\mathrm{B}}\left[\Delta - \varepsilon_{\mathrm{o}}L\right], \tag{4.2.17}$$

from which the prestrain, ε_{o}, can be found.

From an analysis of martensite reorientation, using the full model [15], the stress required for reorientation from thermal martensite to tensile martensite is $\sigma_M = \rho\mu_{\mathrm{c}}/(\beta\sqrt{2})$, provided no TWSME is present. Using the initial condition for equilibrium (4.1.17) leads to the following condition that must be satisfied for the bias spring mismatch and stiffness to cause full reorientation of martensite ($\xi_o = 1$),

$$\frac{\Delta}{L} \geq \beta + \frac{\sigma_M}{E}\left[1 + \frac{EA}{k_{\mathrm{B}}L}\right]. \tag{4.2.18}$$

Considering now kinetics, Figure 4.4 illustrates how forward transformation ($A \rightarrow M^+$) occurs when the thermodynamic driving force, μ, reaches the critical value, μ_{c}, and the amount of tensile martensite has not yet reached saturation ($0 \leq \xi < 1$). Conversely, reverse transformation ($M^+ \rightarrow A$) occurs when the thermodynamic driving force reaches the critical value, -μ_{c}, and the amount of austenite has not yet reached saturation ($0 < \xi \leq 1$). Otherwise, no phase transformation occurs ($\dot{\xi} = 0$, where a dot above a variable denotes a derivative with respect to time, t), and the kinetics equation is not active. During phase

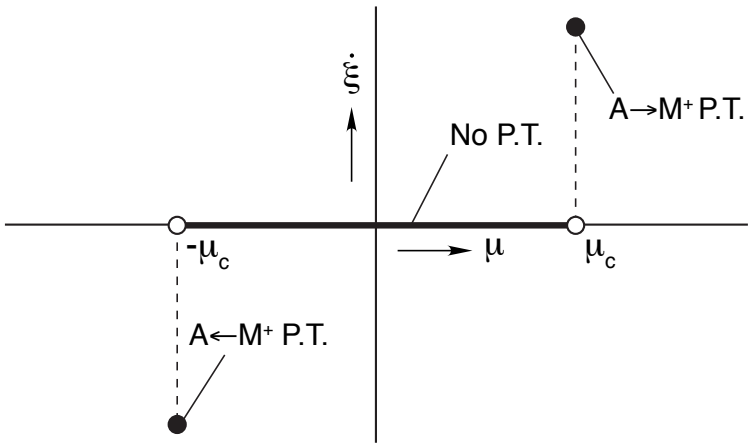


Figure 4.4: Kinetics of phase transformation (P.T.). Transformation rate ($\dot{\xi}$) is zero when $|\mu| < \mu_{\rm c}$ and infinite when $|\mu| = \mu_{\rm c}$.

transformation, however, the SMA stress and temperature are coupled through one of the following conditions

$$\frac{\beta\sigma}{\rho} + (T - T_{\rm R})\,\Delta s = \begin{cases} \mu_{\rm c} & \text{if } \xi \in [0,1) \text{ and } \dot{\xi} > 0 \ (A \rightarrow M^{+}), \\ -\mu_{\rm c} & \text{if } \xi \in (0,1] \text{ and } \dot{\xi} < 0 \ (M^{+} \rightarrow A). \end{cases} \tag{4.2.19}$$

It is apparent from these equations that forward and reverse isothermal superelastic transformations are predicted to occur along stress plateaus with stress hysteresis $2\rho\mu_{\rm c}/\beta$ centered on the stress $(T - T_{\rm R})\,(-\rho\Delta s)/\beta$. These equations represent implicit relations between the state variables $\varepsilon(t)$, $T(t)$ and $\xi(t)$, which can be rewritten using the constitutive equation for stress (4.1.11) as

$$E\,(\varepsilon - \xi\beta) + \frac{\rho}{\beta}\,(T - T_{\rm R})\,\Delta s = \begin{cases} \frac{\rho\mu_{\rm c}}{\beta} & \text{if } \xi \in [0,1) \text{ and } \dot{\xi} > 0, \\ -\frac{\rho\mu_{\rm c}}{\beta} & \text{if } \xi \in (0,1] \text{ and } \dot{\xi} < 0. \end{cases} \tag{4.2.20}$$

Finally, the heat equation (4.1.9) now simplifies to

$$\rho A L c_{\mathrm{o}} \dot{T} = \rho A L \left[\mathrm{sgn}(\dot{\xi}) \mu_{\mathrm{c}} - T \Delta s \right] \dot{\xi} - h \pi d L (T - T_a) + \mathcal{P}_{\mathrm{e}}, \tag{4.2.21}$$

where the left hand side is the rate of energy associated with sensible heat, the first term on the right hand side is the latent heat source rate (sgn is the *signum* function), the second term is the heat loss rate to the environment, and $\mathcal{P}_{\mathrm{e}}$ is the applied Joule heat rate.

4.2.5 Dimensionless Parameters & Equations

To simplify the equations and clarify the scaling of various quantities we define the following dimensionless variables for time, stress, and temperature

$$\tau \equiv \frac{t}{t^*}, \tag{4.2.22}$$

$$\bar{\sigma} \equiv \frac{\sigma}{E}, \tag{4.2.23}$$

$$\theta \equiv (\frac{T}{T_{\mathrm{R}}} - 1)\lambda, \tag{4.2.24}$$

where

$$t^* \equiv \frac{\rho c_{\mathrm{o}} d}{4h}, \tag{4.2.25}$$

is a characteristic time for the rate of sensible heat change, and

$$\lambda \equiv -\frac{\rho \Delta s T_{\mathrm{R}}}{\beta E} \tag{4.2.26}$$

is a dimensionless (positive) reference latent heat (associated with the reversible case $\mu_{\mathrm{c}} = 0$), which incidentally, is also related to the Clausius-Clapeyron slope of stress-induced transformation. The following dimensionless parameters (constants) are also helpful to define,

$$\theta_{\mathrm{a}} \equiv (\frac{T_a}{T_{\mathrm{R}}} - 1)\lambda, \tag{4.2.27}$$

$$\bar{c} \equiv -\frac{c_{\mathrm{o}}}{\Delta s}, \tag{4.2.28}$$

$$\bar{\mu}_{\mathrm{c}} \equiv \frac{\rho \mu_{\mathrm{c}}}{\beta E}, \tag{4.2.29}$$

$$\eta_{\mathrm{B}} \equiv \frac{k_{\mathrm{B}} L}{EA}, \tag{4.2.30}$$

$$\eta_{\mathrm{E}} \equiv \frac{k_{\mathrm{E}} L}{EA}, \tag{4.2.31}$$

$$\eta \equiv \frac{\eta_{\mathrm{B}} + \eta_{\mathrm{E}}}{1 + \eta_{\mathrm{B}} + \eta_{\mathrm{E}}}, \tag{4.2.32}$$

$$\bar{\Delta} \equiv \frac{\Delta}{L}, \tag{4.2.33}$$

$$\bar{\mathcal{P}}_{\mathrm{e}} \equiv \frac{\mathcal{P}_{\mathrm{e}} \lambda}{\pi d L T_{\mathrm{R}} h}, \tag{4.2.34}$$

where θ_{a} is a dimensionless ambient temperature, $\bar{c}$ is a dimensionless specific heat, $\bar{\mu}_{\mathrm{c}}$ is a dimensionless critical driving force, η_{B} is a dimensionless bias spring constant, η_{E} is a dimensionless external spring constant, $\eta \in [0, 1)$ is a dimensionless ratio of spring constants, $\bar{\Delta}$ is a dimensionless bias spring/SMA length mismatch, and $\bar{\mathcal{P}}_{\mathrm{e}}$ is a dimensionless

electrical power.

Time derivatives in the heat equation (4.1.21) can be converted to dimensionless time derivatives, denoted by $(\cdot)' \equiv \partial(\cdot)/\partial\tau$, by the chain rule $\partial(\cdot)/\partial t = \partial(\cdot)/\partial\tau/t^*$. Now the respective governing equations for equilibrium, kinetics, and heat can be written in dimensionless form as

$$\varepsilon(\tau) - \varepsilon_{\mathrm{o}} + \beta\,(1-\eta)\,[1-\xi(\tau)] = 0, \tag{4.2.35}$$

$$\varepsilon(\tau) - \beta\,\xi(\tau) - \theta(\tau) - \mathrm{sgn}(\xi')\bar{\mu}_{\mathrm{c}} = 0, \tag{4.2.36}$$

$$\bar{c}\left[\theta(\tau) + \theta'(\tau) - \theta_{\mathrm{a}} - \bar{\mathcal{P}}_{\mathrm{e}}\right] - \left[\varepsilon(\tau) - \beta\,\xi(\tau) + \lambda\right]\xi'(\tau) = 0. \tag{4.2.37}$$

Equation (4.1.35) assumes the condition for full initial reorientation of martensite (4.1.18) is satisfied (discussed further in Section 4.4), rewritten in dimensionless form as

$$\bar{\Delta} \geq \beta + \bar{\sigma}_M \left[1 + \frac{1}{\eta_{\mathrm{B}}}\right], \tag{4.2.38}$$

where $\bar{\sigma}_M = \bar{\mu}_{\mathrm{c}}/\sqrt{2}$ is the dimensionless stress to orient martensite ($M \rightarrow M^+$). The respective prestrain and dimensionless prestress in the SMA element (related by $\varepsilon_{\mathrm{o}} = \beta + \bar{\sigma}_{\mathrm{o}}$) are

$$\varepsilon_{\mathrm{o}} = \frac{\beta + \eta_{\mathrm{B}}\bar{\Delta}}{1+\eta_{\mathrm{B}}}, \tag{4.2.39}$$

$$\bar{\sigma}_{\mathrm{o}} = \frac{(\bar{\Delta} - \beta)\eta_{\mathrm{B}}}{1+\eta_{\mathrm{B}}}. \tag{4.2.40}$$

According to eqs. (4.1.35) and (4.1.36), the dimensionless (stress-dependent) start and fin-

ish temperatures for the actuator are dependent on the prestress and spring stiffnesses, given by

$$\begin{aligned}
\bar{A}_{\mathrm{s}} &= \bar{\sigma}_{\mathrm{o}} + \bar{\mu}_{\mathrm{c}}, \\
\bar{A}_{\mathrm{f}} &= \bar{\sigma}_{\mathrm{o}} + \bar{\mu}_{\mathrm{c}} + \beta\eta, \\
\bar{M}_{\mathrm{s}} &= \bar{\sigma}_{\mathrm{o}} - \bar{\mu}_{\mathrm{c}} + \beta\eta, \\
\bar{M}_{\mathrm{f}} &= \bar{\sigma}_{\mathrm{o}} - \bar{\mu}_{\mathrm{c}}.
\end{aligned} \tag{4.2.41}$$

A dimensionless steady state temperature, based on the current power level and ambient conditions, is defined as

$$\theta^{\mathrm{ss}} \equiv \theta_{\mathrm{a}} + \bar{\mathcal{P}}_{\mathrm{e}}, \tag{4.2.42}$$

which is shown later to be the asymptotic temperature for long times ($\tau \to \infty$). A comparison between it and the appropriate dimensionless transformation temperature of eq. (4.1.41) can quickly assess whether transformation will start and/or complete. Furthermore, it is convenient to define a normalized, time-dependent, driving force function (for strictly $\mu_{\mathrm{c}} > 0$) as

$$f(\tau) \equiv \frac{\mu\left(\varepsilon(\tau), \theta(\tau), \xi(\tau)\right)}{\mu_{\mathrm{c}}} = \frac{\varepsilon(\tau) - \beta\,\xi(\tau) - \theta(\tau)}{\bar{\mu}_{\mathrm{c}}}, \tag{4.2.43}$$

which can be monitored to determine whether or not phase transformation occurs. This allows eq. (4.1.36), when active, to be written simply as

$$f = \mathrm{sgn}(\xi'). \tag{4.2.44}$$

4.3 Analytical Solution

In this section analytical solutions to the system of equations (4.1.35)-(4.1.37) are derived for $\varepsilon(\tau)$, $\theta(\tau)$, and $\xi(\tau)$ for piecewise constant electrical power input. Assuming that condition (4.1.38) is satisfied, the initial conditions at $\tau = 0$ are $\varepsilon(0) = \varepsilon_{\rm o}$, $\theta(0) = \theta_{\rm a}$, and $\xi(0) = 1$. Presuming sufficient initial prestress exists to create M^+ and sufficient electrical power is supplied to completely transform the SMA wire to austenite, the stress-strain path is uniquely determined by the equilibrium equation (4.1.35).

4.3.1 Mechanical Equilibrium Paths

An example actuator response, based on equilibrium considerations alone, is shown in Figure 4.5. The initial equilibrium point between M^+ (dotted line) and the bias spring is shown by the solid circle. Open circles show the end points on the pure austenite elastic curve (dotted line). Note that, using the equilibrium eq. (4.1.35) and the constitutive eq. (4.1.11), one can show that the slopes of the respective dimensionless paths are

$$\frac{d\bar{\sigma}}{d\varepsilon} = \begin{cases} -\eta_{\rm B}, & \text{bias spring only (path (a))} \\ -\left(\eta_{\rm B} + \eta_{\rm E}\right), & \text{bias plus external springs(path (b)).} \end{cases} \tag{4.3.1}$$

Our immediate concern is to quantify the transient response of the actuator. The character of the time-dependent solution depends on whether or not phase transformation occurs during the time interval, so the general solution can be derived piecewise in time starting from an initial condition for each time interval $\tau_n \leq \tau \leq \tau_{n+1}$, where $n = 0, 1, 2, \ldots$. The

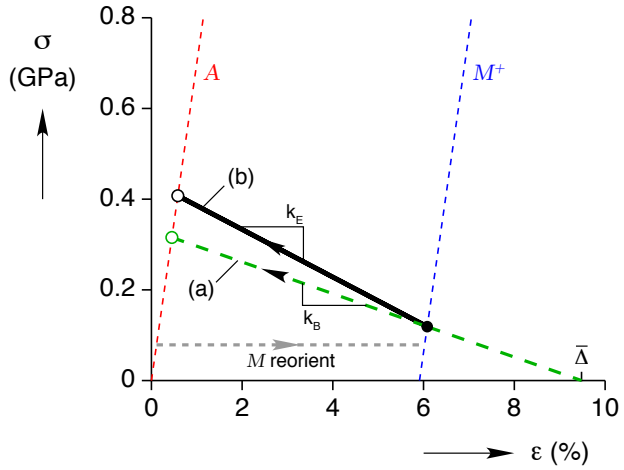


Figure 4.5: Example actuator response during heating in SMA stress vs. strain space: (a) response with bias spring only (dashed line), (b) response with the addition of the external spring (bold line).

power input and ambient temperature for each time interval are then denoted as $\bar{P}_{e,n}$ and $\theta_{a,n}$, respectively. Each τ_n for $n = 1, 2, \ldots$ is a time instant that defines either the onset or termination of phase transformation or a change in power level or environment. The normalized driving force function, $f(\tau)$ of eq. 4.1.43, is monitored during the evolution to determine whether or not phase transformation occurs and to determine the time, τ_n, for the onset of transformation.

4.3.2 Solution Without Phase Transformation

If no phase transformation occurs during some interval $\tau_n \leq \tau \leq \tau_{n+1}$, the phase fraction remains constant at its initial value, $\xi(\tau) = \xi(\tau_n) \equiv \xi_n$. By equilibrium (4.1.35) the strain $\varepsilon(\tau) = \varepsilon(\tau_n) \equiv \varepsilon_n$ also remains constant at its initial value according to

$$\varepsilon_n = \varepsilon_o - \beta\,(1-\eta)\,(1-\xi_n)\,. \tag{4.3.2}$$

Temperature, $\theta(\tau)$, is the only variable that evolves with time. Substituting $\xi'(\tau) = 0$ into (4.1.37), results in the classical sensible heat equation

$$\theta'(\tau) + \theta(\tau) = \theta_{a,n} + \bar{P}_{e,n}, \tag{4.3.3}$$

which has the general solution

$$\theta(\tau) = \theta_n^{\text{ss}} + (\theta_n^{\text{ss}} - \theta_n)e^{-(\tau-\tau_n)}, \quad \tau \in [\tau_n, \tau_{n+1}] \tag{4.3.4}$$

where $\theta_n \equiv \theta(\tau_n)$ is the initial temperature for the time interval, and

$$\theta_n^{\text{ss}} \equiv \theta_{a,n} + \bar{P}_{e,n} \tag{4.3.5}$$

is the dimensionless steady state temperature for long times ($\tau \to \infty$) for the current power and environment.

This solution is valid until phase transformation is detected at time τ_{n+1}, for which

$$\text{Start} \begin{cases} A \to M^+: & f(\tau_{n+1}) = 1, \quad \text{and } f'(\tau_{n+1}) > 0, \quad \text{and } \xi < 1, \\ & \text{or} \\ M^+ \to A: & f(\tau_{n+1}) = -1, \quad \text{and } f'(\tau_{n+1}) < 0, \quad \text{and } \xi > 0. \end{cases} \tag{4.3.6}$$

Since temperature is the only evolving variable, the direction of transformation can be

determined equivalently by

$$f'(\tau_{n+1}) > 0 \implies \theta'(\tau_{n+1}) < 0, \quad \text{and} \quad f'(\tau_{n+1}) < 0 \implies \theta'(\tau_{n+1}) > 0. \tag{4.3.7}$$

Defining the starting stress for the interval as $\bar{\sigma}_n = \varepsilon_n - \beta\xi_n$, onset of transformation is detected when

$$\tau_{n+1} = \tau_n + \ln\left(\frac{\theta_n^{\text{ss}} - \theta_n}{\theta_n^{\text{ss}} - [\bar{\sigma}_n + \text{sgn}(\theta')\bar{\mu}_{\text{c}}]}\right), \tag{4.3.8}$$

provided, of course, that the logarithm is positive.

4.3.3 Solution During Phase Transformation

If, on the other hand, transformation occurs during some time interval $\tau_n \leq \tau \leq \tau_{n+1}$, all three state variables, $\varepsilon(\tau)$, $\theta(\tau)$, and $\xi(\tau)$ evolve from their starting values when phase transformation begins at time τ_n. During the time interval, $\varepsilon(\tau)$ and $\theta(\tau)$ can be solved in terms of $\xi(\tau)$ using eqs. (4.1.35) and (4.1.36), which when substituted into (4.1.37) gives the following nonlinear, first-order ODE for $\xi(\tau)$.

$$\bar{c}\,[\beta\eta + \bar{\sigma}_{\text{o}} - \text{sgn}(\xi')\bar{\mu}_{\text{c}} - \theta_n^{\text{ss}}] - [(1+\bar{c})\,\beta\eta + \bar{\sigma}_{\text{o}} + \lambda]\,\xi'(\tau) + \beta\eta\,\xi(\tau)\,[\xi'(\tau) - \bar{c}] = 0. \tag{4.3.9}$$

Using the initial condition $\xi(\tau_n) = \xi_n$ for the interval, the solution is found to be

$$\xi(\tau) = \xi_n + \frac{1}{b_n}\left[a_n - W\left(a_n\, e^{[a_n - b_n\bar{c}(\tau-\tau_n)]}\right)\right], \quad \tau \in [\tau_n, \tau_{n+1}] \tag{4.3.10}$$

where the following dimensionless constants are defined for the time interval

$$a_n \equiv \frac{\beta\eta\,(1-\xi_n) + \bar{\sigma}_{\rm o} - [\theta_n^{\rm ss} + {\rm sgn}(\xi')\bar{\mu}_{\rm c}]}{\beta\eta\bar{c} + \lambda + \theta_n^{\rm ss} + {\rm sgn}(\xi')\bar{\mu}_{\rm c}},$$
$$b_n \equiv \frac{\beta\eta}{\beta\eta\bar{c} + \lambda + \theta_n^{\rm ss} + {\rm sgn}(\xi')\bar{\mu}_{\rm c}}, \tag{4.3.11}$$

and where $W(z)$ is the Lambert function [20], also known as the *Product-Logarithm*. This is a special function defined as the implicit solution to $W\,e^W = z$. It is available as a built-in function in symbolic manipulation software, such as *Mathematica*®, or can be calculated iteratively as shown in Boyd [12]. The Lambert function $W(z)$ is real valued for $z \geq -e^{-1}$, below which a branch cut exists. It has unique real values for $z \geq 0$ and has two real values for $-e^{-1} < z < 0$ as shown in Figure 4.6. In our case, only the principle branch, $W(z) > -1$, is relevant, so W can be alternately defined by $W(y\,e^y) = y$ for $y \geq -1$. From this definition, we see that

$$\lim_{\tau\to\tau_n} W\left(a_n\,e^{[a_n - b_n\bar{c}(\tau-\tau_n)]}\right) = a_n, \tag{4.3.12}$$

making it apparent that the initial condition is recovered in eq. (4.2.10). While eq. (4.2.9) is nonlinear, it is only first order, and the solution (4.2.10) was discovered with the help of *Mathematica*® once suitable simplifying constants were defined by eq. (4.2.11). The reader can readily verify the solution by substituting eq. (4.2.10) into (4.2.9).

The remaining state variables follow from the solution above. Using eqs. (4.2.10),

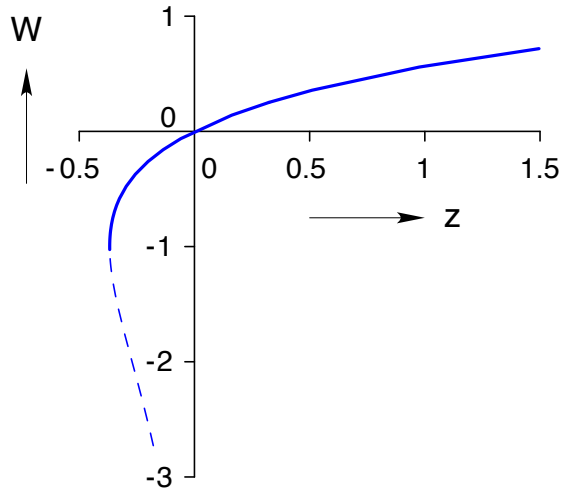


Figure 4.6: The real-valued part of the Lambert $W(z)$ function.

(4.1.35), and (4.1.36), it can be shown that

$$\theta(\tau) = \theta_n - \frac{\beta\eta}{b_n}\left[a_n - W\left(a_n\, e^{[a_n - b_n\bar{c}(\tau-\tau_n)]}\right)\right], \tag{4.3.13}$$

$$\varepsilon(\tau) = \varepsilon_n + \frac{\beta\,(1-\eta)}{b_n}\left[a_n - W\left(a_n\, e^{[a_n - b_n\bar{c}(\tau-\tau_n)]}\right)\right], \tag{4.3.14}$$

where the initial values are uniquely identified with ξ_n by

$$\theta_n = \beta\eta\,(1-\xi_n) + \bar{\sigma}_{\mathrm{o}} - \mathrm{sgn}(\xi')\bar{\mu}_{\mathrm{c}} \tag{4.3.15}$$

$$\varepsilon_n = \beta\,[\eta + (1-\eta)\xi_n] + \bar{\sigma}_{\mathrm{o}}. \tag{4.3.16}$$

From (4.1.11) the evolution of the dimensionless stress is then

$$\bar{\sigma}(\tau) = \sigma_n - \frac{\beta\eta}{b_n}\left[a_n - W\left(a_n\, e^{[a_n - b_n\bar{c}(\tau-\tau_n)]}\right)\right], \tag{4.3.17}$$

Phase transformation terminates when either the driving force is no longer sufficient

to sustain the transformation or when the phase fraction saturates, i.e.,

$$\text{End} \begin{cases} A \to M^+ : \quad f(\tau_{n+1}) = 1, \quad \text{and } \left(f'(\tau_{n+1}) < 0 \text{ or } \xi = 1\right), \\ \qquad\qquad\qquad \text{or} \\ A \leftarrow M^+ : \quad f(\tau_{n+1}) = -1, \quad \text{and } \left(f'(\tau_{n+1}) > 0 \text{ or } \xi = 0\right). \end{cases} \tag{4.3.18}$$

Alternatively, the termination condition can be found as follow. It can be shown according to eq. (4.2.13), that the long time, asymptotic temperature is still θ_n^{ss} as defined in eq. (4.2.5). Note that the temperature evolution is always monotonic within each time interval. If θ_n^{ss} does not exceed $\bar{A}_{\text{f}}$ during heating, or drop below $\bar{M}_{\text{f}}$ during cooling, phase transformation does not reach saturation and the given solutions are valid until the next power change or ambient condition change. In this case, the steady state value of the phase fraction for long times, $\xi^{\text{ss}} \equiv \xi(\infty)$, is found from eq. (4.2.9) by setting $\xi'(\infty) = 0$ and solving as

$$\xi^{\text{ss}} = \frac{\bar{\sigma}_{\text{o}} - \text{sgn}(\xi')\bar{\mu}_{\text{c}} - \theta_n^{\text{ss}}}{\beta\eta}. \tag{4.3.19}$$

If, conversely, one of the threshold temperatures is predicted to be exceeded by θ_n^{ss} the transformation saturates at one of following final phase fraction values

$$\xi_{\text{f}} = \begin{cases} 0, & \xi_n' < 0, \\ 1, & \xi_n' > 0, \end{cases} \tag{4.3.20}$$

In this case, the time interval is to be subdivided, and a new time (τ_{n+1}) is calculated for the termination of transformation. The time (τ_{n+1}) at which this happens is found by solving eq. (4.2.10) for $\xi(\tau_{n+1}) = \xi_{\text{f}}$. While at first glance it might appear that an iterative

numerical solution is necessary, it can be solved analytically as follows. First, define the intermediate variable

$$y_n \equiv W\left(a_n\, e^{[a_n - b_n \bar{c}(\tau_{n+1}-\tau_n)]}\right), \tag{4.3.21}$$

which must satisfy

$$y_n = a_n + b_n\left(\xi_n - \xi_{\mathrm{f}}\right). \tag{4.3.22}$$

By the property $y = W(ye^y)$ equality between the two can be rewritten as

$$y_n e^{y_n} = a_n\, e^{[a_n - b_n \bar{c}(\tau_{n+1}-\tau_n)]}, \tag{4.3.23}$$

which can then be solved for τ_{n+1} (with y_n replaced) as

$$\tau_{n+1} = \tau_n + \frac{1}{\bar{c}}\left[\xi_{\mathrm{f}} - \xi_n - \frac{1}{b_n}\ln\left(1 - \frac{b_n}{a_n}(\xi_{\mathrm{f}} - \xi_n)\right)\right]. \tag{4.3.24}$$

In fact, this is just a special instance of the implicit form of eq. (4.2.10) where time can be considered a function of the phase fraction

$$\tau(\xi) = \tau_n + \frac{1}{\bar{c}}\left[\xi - \xi_n - \frac{1}{b_n}\ln\left(1 - \frac{b_n}{a_n}(\xi - \xi_n)\right)\right]. \tag{4.3.25}$$

4.4 Numerical Example

As a numerical example, the simplified model is compared to one of the finite element calculations of Chang, *et. al.* [15], with thermally insulated ends. The finite element model was implemented using an in-house research code. Weak forms of the equilibrium and heat equations were solved through staggered Newton-Raphson iteration between thermal and mechanical degrees of freedom. The kinetic law was implemented using an explicit 4^{th} order Runge-Kutta algorithm. The results presented here used 200 elements along the specimen length, with displacement and its gradient given by Hermite cubic shape functions and temperature by linear shape functions.

The SMA material parameters are taken from Table 4.1, and Table 4.2 lists the spring parameters, ambient environment constants (-50 °C stagnant air), and electrical power history similar to that used in [15]. The finite element simulation used thermally insulated ends, for which the temperature field was relatively uniform (as opposed to constant temperature boundary conditions), to provide a fair comparison with our lumped model simulation. The full set of dimensionless constants is provided in Table 4.3. (The subset of mechanical parameters from this list were used to construct Figure 4.5.) These parameters result in a steady state temperature rise of $\Delta T_\infty = 100.6$ °C ($\Delta\theta^{\text{ss}} = \bar{\mathcal{P}}_{\text{e}} = 0.01061$) above the ambient temperature when power is applied, a characteristic time of $t^* = 6.88$ s for sensible heating, and an initial pre-stress of $\sigma_{\text{o}} = 119$ MPa ($\bar{\sigma}_{\text{o}} = 0.00170$). The specific dimensionless (stress-dependent) temperatures at which transformation occurs in the actuator during heat and cooling are provided in Table 4.4.

Table 4.2: Spring parameters, environment constants, and electrical power.

Springs	Ambient	Power
$k_{\mathrm{B}} = 26.60$ N/mm	$T_a = 223.15$ K	$\mathcal{P}_{\mathrm{e}} = 1.301$ W
$k_{\mathrm{E}} = 13.30$ N/mm	$h = 90$ W/m^{2}K	
$\Delta = 5.688$ mm		

Table 4.3: Dimensionless parameters.

SMA	Springs	Ambient	Power
$\beta = 0.0591$	$\eta = 0.06977$	$\theta_{\mathrm{a}} = -0.002758$	$\bar{\mathcal{P}}_{\mathrm{e}} = 0.01061$
$\lambda = 0.02630$	$\eta_{\mathrm{B}} = 0.05$		
$\bar{c} = 7.446$	$\eta_{\mathrm{E}} = 0.025$		
$\bar{\mu}_{\mathrm{c}} = 0.001587$	$\bar{\Delta} = 0.0948$		

Table 4.4: Dimensionless actuator transformation temperatures.

$\bar{M}_{\mathrm{f}}$	$\bar{M}_{\mathrm{s}}$	$\bar{A}_{\mathrm{s}}$	$\bar{A}_{\mathrm{f}}$
0.00011	0.00424	0.00329	0.00741

4.4.1 Time-dependent Response

Simulation results for a complete heating/cooling cycle are shown graphically in Figure 4.7 for dimensionless quantities. A column of plots are shown against a common time axis for (a) the applied power ($\bar{\mathcal{P}}_e$), (b) evolution of temperature (θ), stress ($\bar{\sigma}$), and strain (ε), and (c) M^+ phase fraction (ξ) and normalized driving force (f). Electrical power (1.3 W) is applied for about 7.268 dimensionless time units (50 s), and then turned off for the same time period. Initially, the phase fraction, strain, and stress remain constant, and the temperature rises rapidly according to an exponential curve during sensible heating. This is interrupted once $M^+ \rightarrow A$ phase transformation starts, at which point the phase fraction and strain decrease while the stress increases. The temperature continues to rise by another exponential-like curve, but at a slower rate, since both sensible heat and latent heat must be supplied to transform the SMA material. Once the transformation is complete ($\xi = 0$), the phase fraction, strain, and stress again remain constant while the temperature rises at a more rapid rate according to sensible heating only.

When the power is turned off, a similar, but reverse direction behavior, occurs. The phase fraction, strain, and stress stay constant initially while the temperature decreases rapidly according to sensible cooling. At the onset of $A \rightarrow M^+$ transformation the temperature decrease is arrested somewhat at latent heat is extracted from the surroundings. Once the transformation saturates ($\xi = 1$), the actuator has reset to its initial phase fraction, strain, and stress, and the temperature decreases asymptotically toward the ambient temperature according to sensible cooling again. Overall, the start and termination of phase transformation are accompanied by noticeable discontinuities in the rates of change of all

state variables. Finally, the response path taken for the entire cycle is plotted in stress-strain-temperature space in Figure 4.8, showing the complete hysteresis loop (with time suppressed).

4.4.2 Comparison to Finite Element Simulation

To explore the reasonableness of the lumped model solution we now compare the above simulation to that of the previous finite element numerical simulation (see Figures 27 and 28 of [15]). Figure 4.9(a) shows a comparison of the average strain histories. Figure 4.9(b) shows a comparison of the temperature histories (mid-span temperature for the FE case). The open and closed circles show the onset and termination of transformations of the lumped model. The slight delay in the onset of transformation of the FE case was due to the input power history that was slightly different, i.e., ramped up linearly over 2.5 s, rather than instantaneously, to the final constant value (and then ramped down over the same time interval when switched off). Figure 4.9(c) shows force histories in the external spring. Overall, the responses shown in Figure 4.9 are reasonably similar, and other minor differences, such as the jaggedness in the FE temperature history in Figure 4.9(b) can be attributed to the spatial non-uniformities due to mechanical instabilities and localization effects (a focus of that study) captured by the finite element simulation but not by our lumped model. As a convenient design-level simulation model, we consider the results of the lumped model to be quite satisfactory.

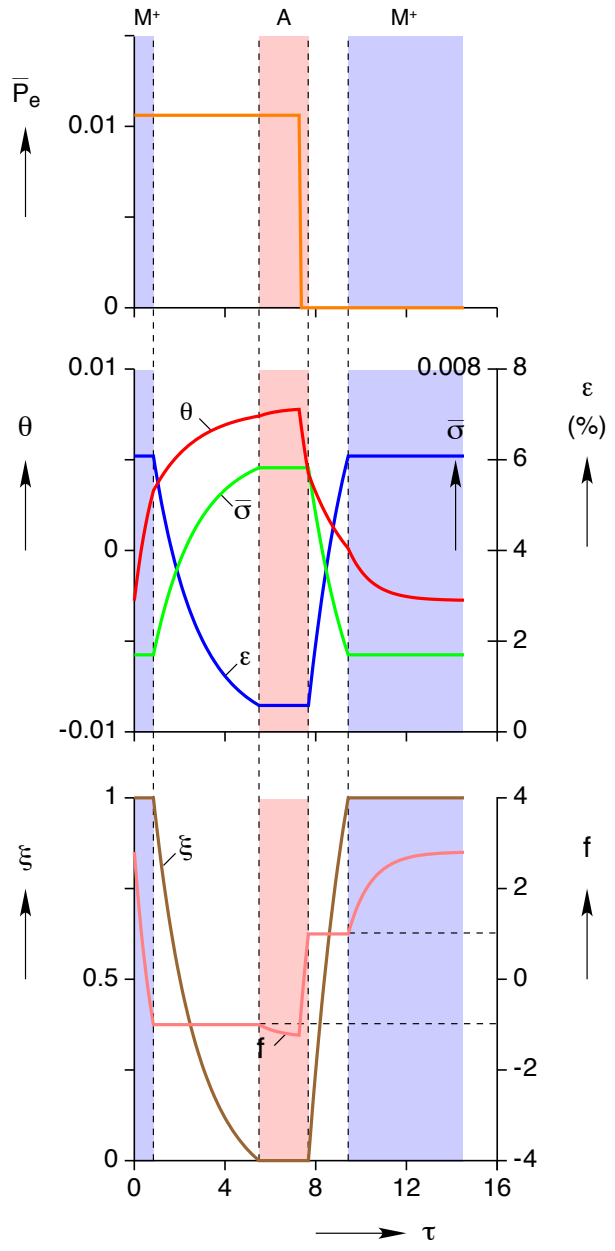


Figure 4.7: Dimensionless SMA actuator response: (a) the applied power history (orange line), (b) temperature (red line), stress (green line), and strain (blue line), and (c) M^+ phase fraction (brown line) and normalized driving force (salmon line). Shaded blue and red area indicate time regimes for pure M^+ and A phases, respectively.

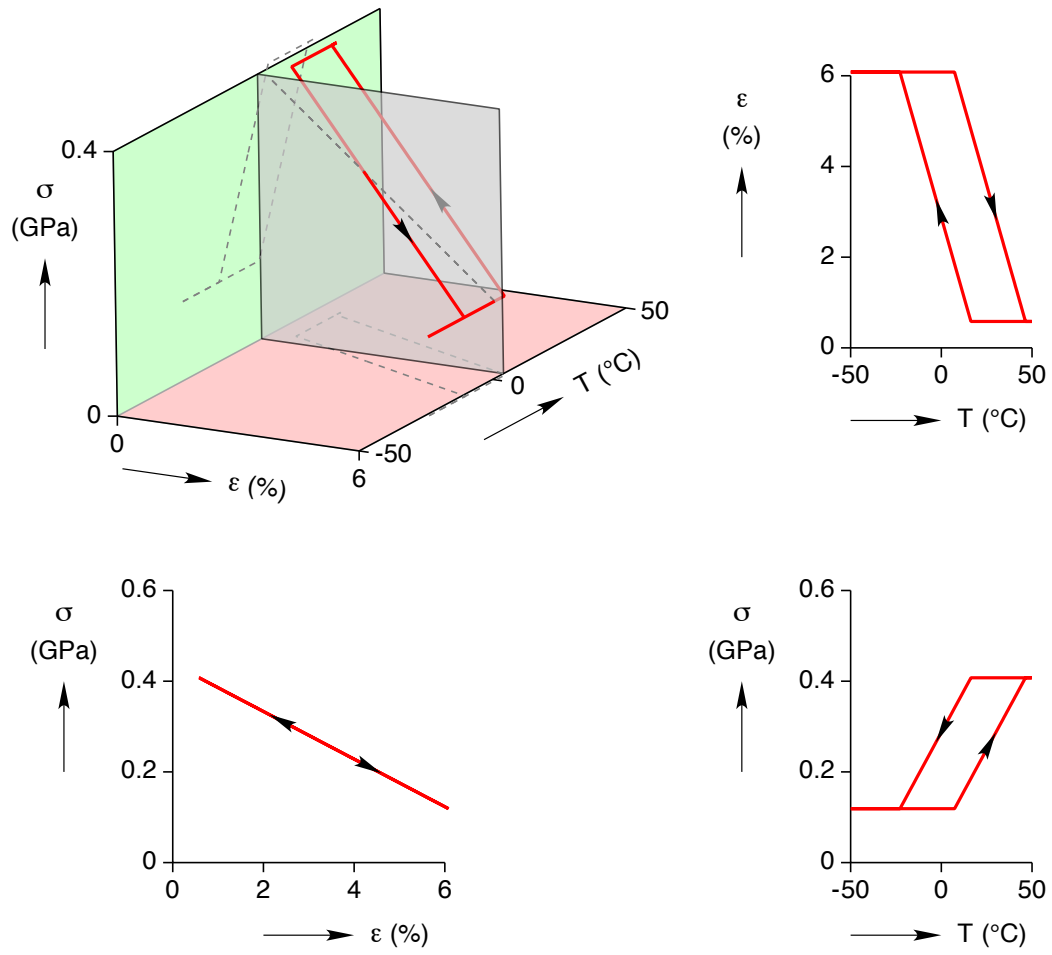


Figure 4.8: Dimensional SMA actuator response in stress-strain-temperature space along with planar projections.

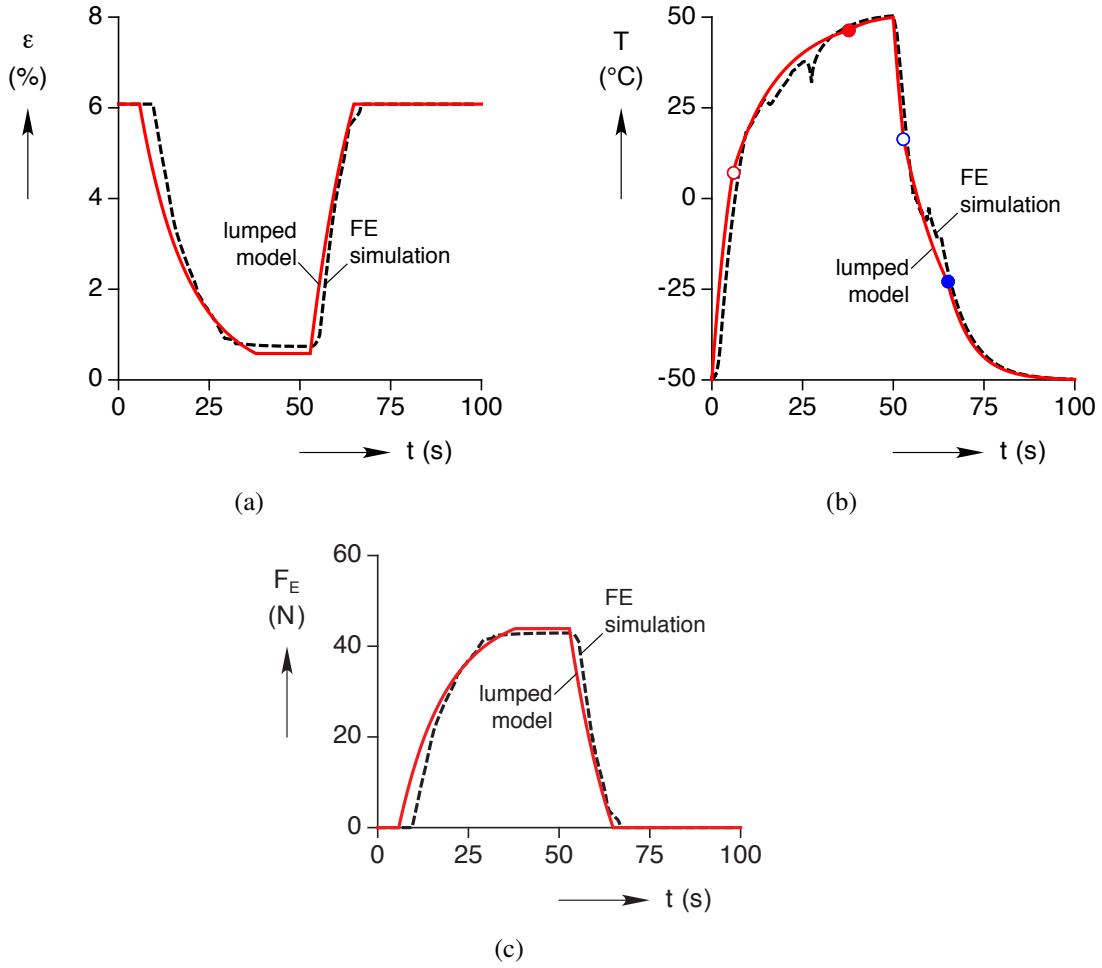


Figure 4.9: Comparison of simulation histories between lumped actuator model (solid line) and finite element (FE) simulation (dashed line) of Chang, *et. al.* [15]: (a) SMA strain (global average for FE case), (b) temperature (mid-length for FE case), (c) external spring force.

4.4.3 Actuation times

Of particular interest is the duration of time needed to complete such an actuation cycle. Table 4.5 provides both dimensionless and dimensional values at significant times during the cycle. In this example, sensible heating takes about 5.8 s before transformation starts, $M^+ \rightarrow A$ transformation takes about 32.1 s, and post transformation sensible heating occurs for another 12 s, or so (arbitrary, depending on the duration of heating). Sensible cooling takes about 2.8 s, starting from about 50 °C, and $A \rightarrow M^+$ transformation takes about 12.1 s. In this case, the time for transformation during heating is longer than that during cooling due to the heat losses to the ambient environment, but in general that need not be the case if the electrical power input during heating is large enough.

It is also interesting to study the times necessary to start and finish transformation as a function of different heat transfer environments and electrical power inputs. Setting $\theta(\tau_1) = \bar{A}_\mathrm{s}$ in eq. (4.2.8) with the appropriate initial conditions, the dimensionless time

Table 4.5: Numerical results: values at significant times.

n	τ_n	t_n (s)	θ_n	T_n (°C)	$\varepsilon_n(\%)$	ξ_n	$\bar{\sigma}_n$	σ_n (GPa)	Event
0	0	0	-0.00276	-50	6.081	1	0.0017	0.119	heat
1	0.843	5.8	0.00329	7.3	6.081	1	0.0017	0.119	A_s
2	5.504	37.9	0.00741	46.4	0.582	0	0.00582	0.408	A_f
3	7.268	50	0.00778	49.9	0.582	0	0.00582	0.408	cool
4	7.678	52.8	0.00424	16.3	0.582	0	0.00582	0.408	M_s
5	9.431	64.9	0.00011	-22.8	6.081	1	0.0017	0.119	M_f
6	14.537	100	-0.00274	-49.8	6.081	1	0.0017	0.119	—

needed to start $M^+ \rightarrow A$ transformation is

$$\tau_1 = \ln\left(\frac{\theta^{\text{ss}} - \theta_{\text{a}}}{\theta^{\text{ss}} - \bar{A}_{\text{s}}}\right). \tag{4.4.1}$$

The dimensionless time needed to finish $M^+ \rightarrow A$ transformation can be found by setting $\xi(\tau_2) = 0$ in eq. (4.2.10) with the appropriate starting conditions.

$$\tau_2 = \tau_1 + \left(1 + \frac{\theta^{\text{ss}} + \lambda - \bar{\mu}_{\text{c}}}{\beta\eta\bar{c}}\right)\ln\left(\frac{\theta^{\text{ss}} - \bar{A}_{\text{s}}}{\theta^{\text{ss}} - \bar{A}_{\text{f}}}\right) - \frac{1}{\bar{c}}. \tag{4.4.2}$$

These two are converted to dimensional times (t_1 and t_2) and are plotted in Figure 4.10 as a function of convective film coefficient (h) and electrical power input ($\mathcal{P}_{\text{e}}$), respectively, while holding all other parameters fixed. One can see in Figure 4.10(a) that the chosen electrical power ($\mathcal{P}_{\text{e}} = 1.3$ W) is insufficient to start the transformation ($t_1 \rightarrow \infty$) for convective film coefficients larger than about 158 W/m^{2}K, and it is insufficient to finish the transformation ($t_2 \rightarrow \infty$) for values larger than 93.9 W/m^{2}K. Figure 4.10(b) shows that for a fixed convection coefficient ($h = 90$ W/m^{2}K) the times to start and finish transformation are strongly dependent on the electrical power input. A modest increase in power input will dramatically decrease the time needed. For low power input they asymptotically approach infinity at particular power levels, 0.741 W to start and 1.25 W to finish the $M^+ \rightarrow A$ transformation. Accordingly, our chosen power, 1.3 W, is barely sufficient for our example actuator system.

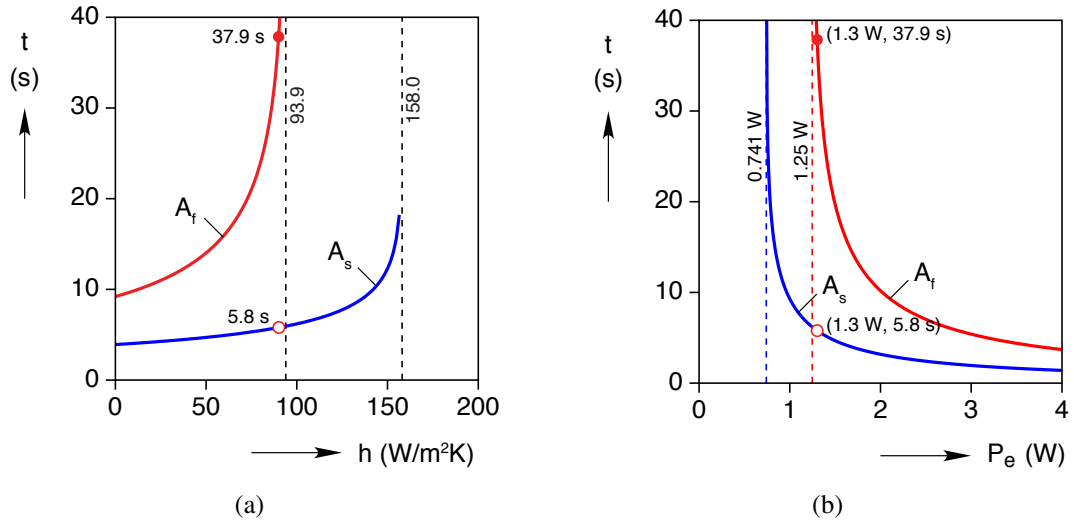


Figure 4.10: Times to start, t_1 (A_s, blue line), and finish, t_2 (A_f, red line), $M^+ \rightarrow A$ transformation: (a) as a function of convective film coefficient, h, (b) as a function of electrical power input, $\mathcal{P}_\mathrm{e}$. All other parameters are fixed as given in Table 4.2. Open and closed circles correspond to the numerical example.

4.5 Design Aspects

In this section we consider some issues related to actuator design based on our simple actuator model. The numerical example of Section 4.3 was an *ad hoc* design, and no attempt was made to optimize it. Here, we discuss a systematic design approach by considering minimum requirements for feasible actuation. We also show how the design could be optimized given SMA stress limits and constraints on power and energy. We start with the parameters of Table 4.3 and then consider modifications of parameters one at a time. Thus, parameters of Table 4.3 are assumed unless stated otherwise throughout this section. The following considers actuation stroke, sizing of springs, actuation time, minimum power requirements, and energy usage and energy efficiency. The section concludes with a discussion of potential refinements to the model to improve its accuracy and generality for design.

4.5.1 Actuation Stroke

The maximum stress clearly occurs when the SMA element is fully A phase. From eqs. (4.1.35) and (4.1.40) the maximum stress in general is

$$\bar{\sigma}_{\max} = \bar{\sigma}_{\mathrm{o}} + \beta\eta. \tag{4.5.1}$$

This can be combined with eqn. (4.1.39) to obtain the actuation stroke of the external spring, $\delta_{\mathrm{E}} = \Delta\varepsilon\, L$, which is calculated using

$$\Delta\varepsilon = \varepsilon_{\mathrm{o}} - \bar{\sigma}_{\max}, \tag{4.5.2}$$

since the dimensionless stress, $\bar{\sigma}_{\max}$ is the numerical equivalent of the minimum SMA strain during actuation. This simplifies to

$$\Delta\varepsilon = \beta(1-\eta) = \frac{\beta}{1+\eta_{\mathrm{B}}+\eta_{\mathrm{E}}}. \tag{4.5.3}$$

According this expression, the best case is $(\delta_{\mathrm{E}}/L)_{\max} = \beta$, which corresponds to complete dead loading ($\eta_{\mathrm{B}} = \eta_{\mathrm{E}} = 0$). It predicts that the stroke is independent of the starting (ambient) temperature and the initial mismatch ($\bar{\Delta}$), and that larger values of η_{B} and η_{E} reduce the actuator stroke somewhat. However, one should be wary of this conclusion, since the elastic modulus for martensite was assumed to be the same as that of austenite in our simplified SMA model. In reality, the martensite mechanical response is quite nonlinear at large strains, having a tangent modulus much lower than the austenite

modulus. We consider this to be the most severe limitation of our model, since β is treated as a material constant. Changing the ambient temperature θ_{a} and/or the prestress $\bar{\sigma}_{\mathrm{o}}$ may give less accurate stroke predictions. A possible refinement for generalizing the model to more accurately predict the actuator stroke under general conditions is discussed later in Section 4.4.7. Nevertheless, the actuator model as presented is simple and convenient, so we continue to explore its design implications in the next several sections.

4.5.2 Sizing of Springs

As discussed in Section 4.1.5, eq. (4.1.38) is the design condition to ensure that full initial tensile martensite ($\xi_0 = 1$) exists upon assembly of the actuator. The lower bound of feasible designs is

$$\bar{\Delta}_{\min} = \beta + \bar{\sigma}_M \left[1 + \frac{1}{\eta_{\mathrm{B}}}\right], \tag{4.5.4}$$

recalling that $\bar{\sigma}_M \equiv \bar{\mu}_{\mathrm{c}}/\sqrt{2}$ is the dimensionless stress needed to ensure M^+ exists when cold. This is plotted in Figure 4.11 (solid black line). In our design example this corresponds to the minimum pre-stress condition $\sigma_{\mathrm{o}} \geq \sigma_M (= 78.5 \text{ MPa})$. The filled black circle in the figure is our chosen design ($\bar{\Delta} = 0.0948$), while the open circle is the minimum feasible value ($\bar{\Delta}_{\min} = 0.0827$) for the same bias spring stiffness. This latter case can be considered an "optimal" choice, since it minimizes the stress in the SMA element. For reference, the other lines in the figure indicate lower bounds to other feasible designs for SMA wire with other values of hysteresis ($\bar{\mu}_{\mathrm{c}}$).

Throughout this study we have assumed that no cyclic shakedown, or ratcheting, of the SMA element occurs during operation of the actuator. This is true only if the SMA element

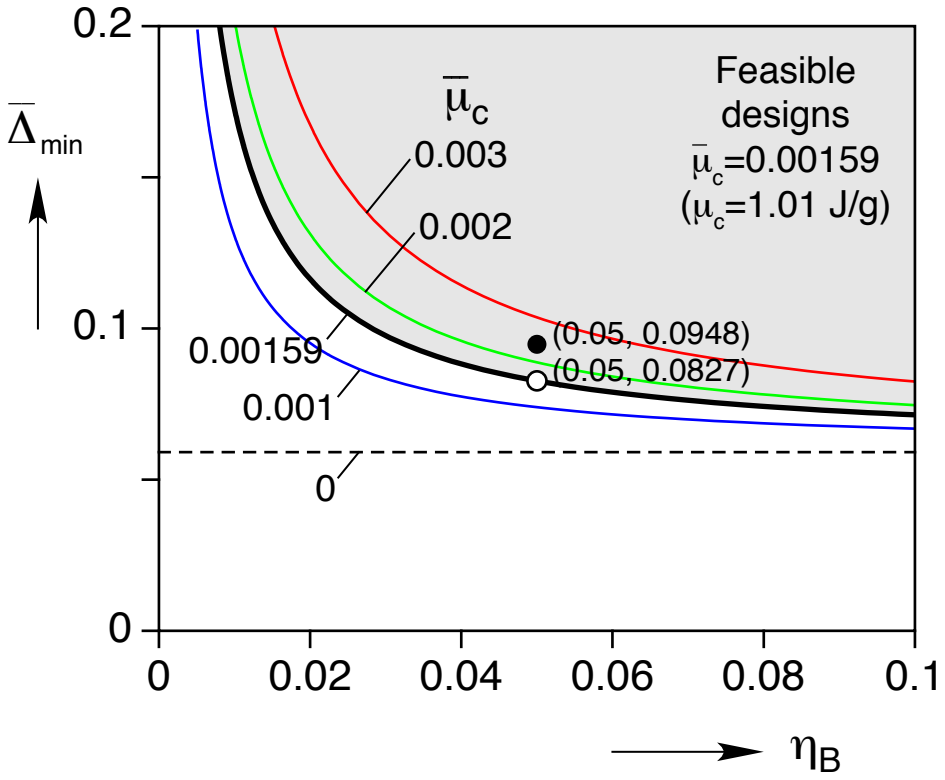


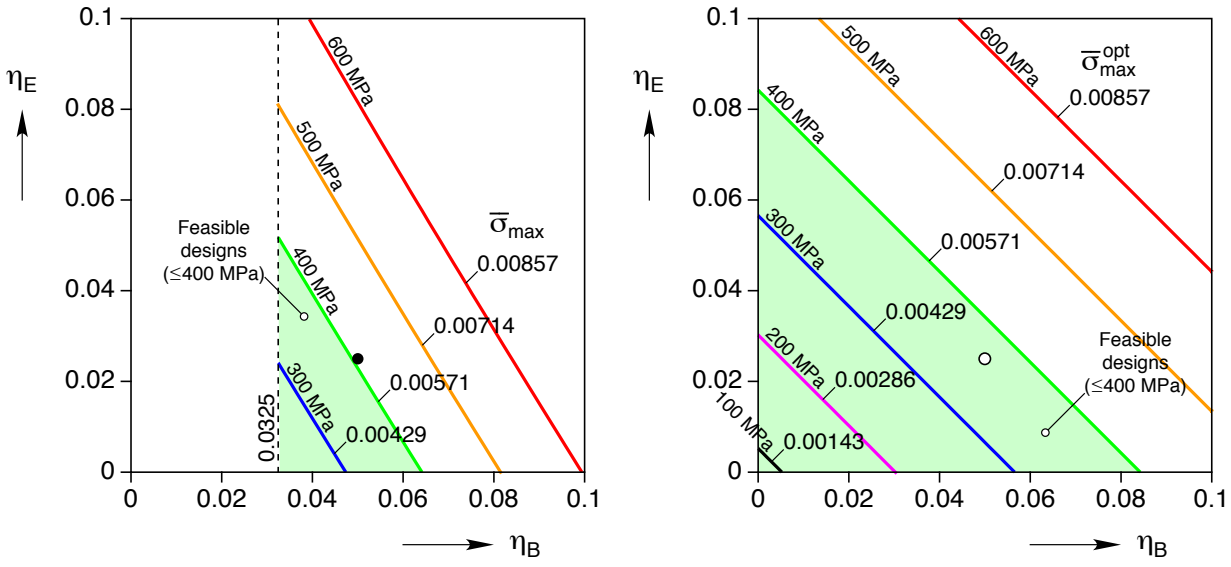
Figure 4.11: Minimum bias spring mismatch ($\bar{\Delta}_{\min}$) as a function of dimensionless bias spring stiffness (η_{B}) to ensure initial complete reorientation of martensite (M^+) during initial actuator assembly (cold). The gray region indicates feasible designs for values of β and $\bar{\mu}_{\mathrm{c}}$ of Table 4.3.

has been conditioned (or trained) to achieve repeatable cyclic response, or if the stress level is kept sufficiently low. In either case the maximum SMA stress must be maintained acceptably low during operation. With the maximum stress from eq. (4.4.1), the maximum stress for the "optimal design" is

$$\bar{\sigma}_{\max}^{\text{opt}} = \bar{\sigma}_M + \beta\eta. \tag{4.5.5}$$

These constant maximum stress lines are plotted, respectively, in Figure 4.12(a) and 4.12(b) in the space of bias spring and external spring stiffnesses. The space in Figure 4.12(a) is truncated by a dotted line, which is the minimum bias spring stiffness ($\eta_{B,\min} = 0.0325$) for the chosen bias spring mismatch ($\bar{\Delta} = 0.0948$) according to eq. (4.1.38). The solid black circle indicates the maximum stress of 408 MPa for the chosen values of η_{B} and η_{E} in Table 4.3. Figure 4.12(b) has no such restriction, since $\bar{\Delta} = \bar{\Delta}_{\min}$ has been substituted into eq. (4.4.5). In this case, the open circle indicates a lower maximum stress of 367 MPa, since the optimal bias spring mismatch has been used. The shaded regions of Figure 4.12 show regimes of η_{B} and η_{E} to keep the maximum stress under a hypothetical design value of 400 MPa.

Another possible step to minimize the worst case stress is to use a very compliant bias spring with a suitable value of $\bar{\Delta}$ to maintain $\bar{\sigma}_{\text{o}} = \bar{\sigma}_M$. The designer should use as compliant a bias spring as is reasonable to meet actuator size requirements. In the limit as $\eta_{\text{B}} \to 0$ the bias spring mismatch gets very large ($\bar{\Delta} \to \infty$), which corresponds essentially to dead loading by the bias spring. While this may be difficult to achieve in practice, we consider it as a limiting ideal case. Substituting $\eta_{\text{B}} = 0$ in eq. (4.4.5) gives a best case



(a) (b)

Figure 4.12: Design stresses as a function of bias spring (η_{B}) and external spring (η_{E}) stiffnesses: (a) "non-optimal design" using $\bar{\Delta} = 0.0948$ (solid circle is the example of Section 4.3), (b) "optimal design" using $\bar{\Delta}_{\min}$.

maximum stress of

$$\bar{\sigma}^{\mathrm{best}}_{\mathrm{max}} = \bar{\sigma}_M + \frac{\beta\eta_{\mathrm{E}}}{1+\eta_{\mathrm{E}}}, \tag{4.5.6}$$

which is plotted in Figure 4.13. For the chosen value of $\eta_{\mathrm{E}} = 0.025$ this give a maximum stress of only 0.00256 (179 MPa) as shown by the open circle in the figure.

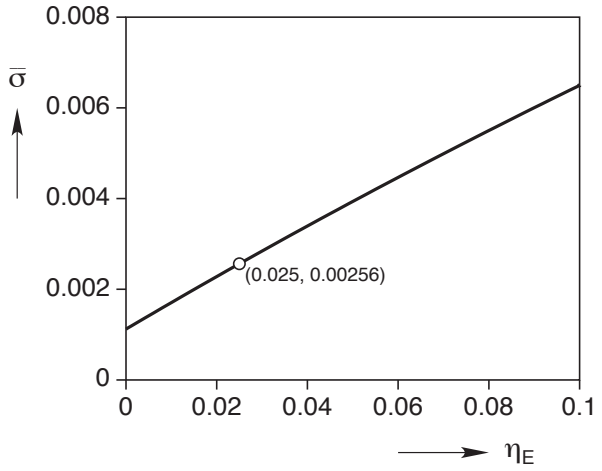


Figure 4.13: Dimensionless maximum stress ($\bar{\sigma}_{\max}^{\text{best}}$) as a function of external spring stiffness (η_{E}) for dead loading bias ($\eta_{\text{B}} = 0$, $\bar{\sigma}_{\text{o}} = \bar{\sigma}_M$).

4.5.3 Actuation Times

Here we revisit the issue of actuation times to start and finish $M^+ \rightarrow A$ transformation. Equations (4.3.1) and (4.3.2) can be rewritten explicitly in terms of $\bar{\mathcal{P}}_{\text{e}}$ as

$$\tau_1 = \ln\left(\frac{\bar{\mathcal{P}}_{\text{e}}}{\bar{\mathcal{P}}_{\text{e}} + \theta_{\text{a}} - \bar{A}_{\text{s}}}\right), \tag{4.5.7}$$

$$\tau_2 = \tau_1 + \left(1 + \frac{\bar{\mathcal{P}}_{\text{e}} + \theta_{\text{a}} + \lambda - \bar{\mu}_{\text{c}}}{\beta\eta\bar{c}}\right) \ln\left(\frac{\bar{\mathcal{P}}_{\text{e}} + \theta_{\text{a}} - \bar{A}_{\text{s}}}{\bar{\mathcal{P}}_{\text{e}} + \theta_{\text{a}} - \bar{A}_{\text{f}}}\right) - \frac{1}{\bar{c}}. \tag{4.5.8}$$

Figure 4.14a shows how Figs 4.10(a) and 4.10(b) collapse to a single plot on a dimensionless basis, since the definition of $\bar{\mathcal{P}}_{\text{e}}$ includes both the input power and the film coefficient. Other parameters in Table 4.3 are held fixed. Open and closed circles correspond to the example of Section 4.3. Dotted lines indicate minimum dimensionless powers to start (0.00604) and finish (0.0102) the transformation, i.e., where $\tau_1 \rightarrow \infty$ and $\tau_2 \rightarrow \infty$, respectively. The figure also shows the time interval ($\Delta\tau_{12} = \tau_2 - \tau_1$) during which transformation occurs.

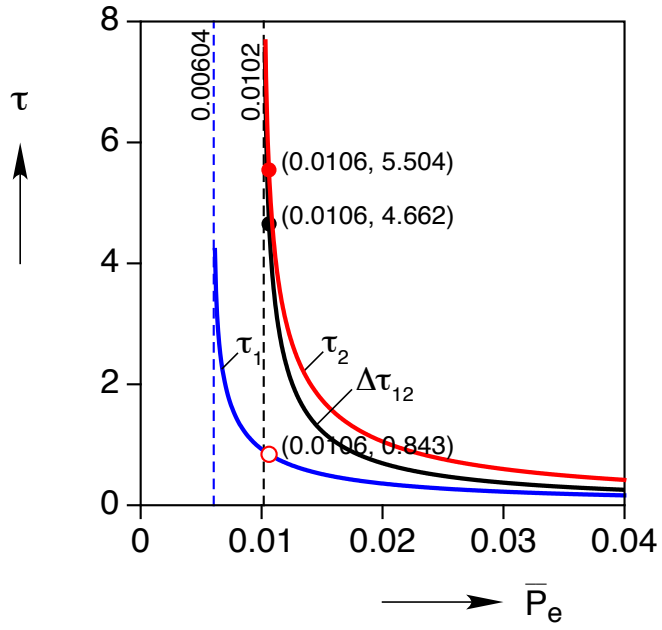


Figure 4.14: Dimensionless times (τ) during $M^+ \rightarrow A$ transformation as a function of dimensionless power ($\bar{\mathcal{P}}_{\rm e}$): start time, τ_1 (to reach $A_{\rm s}$), time period during $M^+ \rightarrow A$ transformation, $\Delta\tau_{12}$, and finish time, τ_2 (to reach $A_{\rm f}$); .

4.5.4 Minimum Power Requirements

Minimum dimensionless powers to start and finish the $M^+ \rightarrow A$ transformation are the singular points of eqs. (4.4.7) and (4.4.8), given by

$$\left(\bar{\mathcal{P}}_{\rm e}\right)_{\rm min} = \begin{cases} \bar{A}_{\rm s} - \theta_{\rm a}, & \text{start} \\ \\ \bar{A}_{\rm f} - \theta_{\rm a}, & \text{finish.} \end{cases} \tag{4.5.9}$$

These can be converted to dimensional power requirements using definition (4.1.34) as simple linear functions of the film coefficient as plotted in Figure 4.15, keeping all other parameters fixed in Table 4.3. The gray region shows feasible powers to achieve complete actuation. The solid circle is the example of Section 4.3.

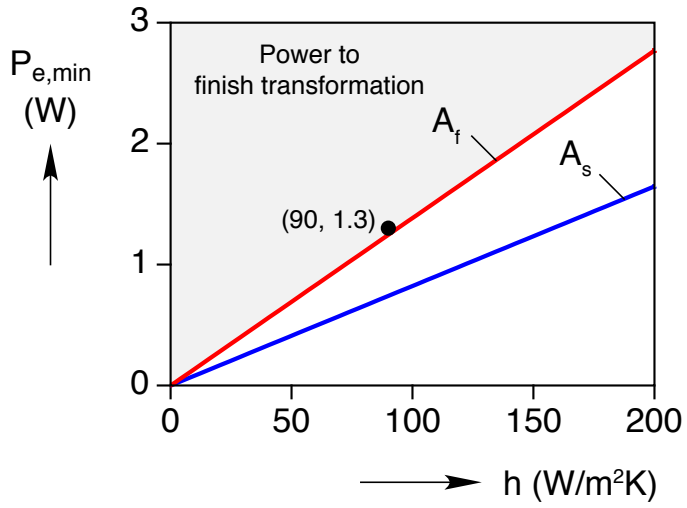


Figure 4.15: Minimum power requirements to start (blue line) and finish (red line) $M^+ \rightarrow A$ transformation as a function of film coefficient.

Again using eq. (4.4.9) we now consider relaxing other parameters to study the minimum power requirements for alternate designs. Figure 4.16 shows lines of constant dimensionless power needed to achieve A_s in the design space of dimensionless bias and external spring constant (η_B) and bias spring mismatch ($\bar{\Delta}$). The lower bound (dotted black line) is the minimum design curve consistent with the solid black line of Figure 4.11, which requires a minimum power of $\bar{\mathcal{P}}_e = 0.00547$. The bold dotted black line corresponds to the power ($\bar{\mathcal{P}}_e = 0.0106$) used in the example of Section 4.3. The gray region is the feasible design space for (η_B, $\bar{\Delta}$) for this power level. The solid black circle is the corresponding coordinates (η_B, $\bar{\Delta}$) of the example, and the open circles indicate the minimum and maximum values of $\bar{\Delta}$ for the same η_B.

Figure 4.17(a) then shows lines of constant dimensionless power needed to achieve A_f in the design space of dimensionless bias spring and external spring stiffnesses (η_B, η_E) while holding $\bar{\Delta} = 0.0948$ of the example. Clearly, larger values of η_B and η_E require larger power levels, since they raise $\bar{\sigma}_{max}$ and thus raise the corresponding transformation

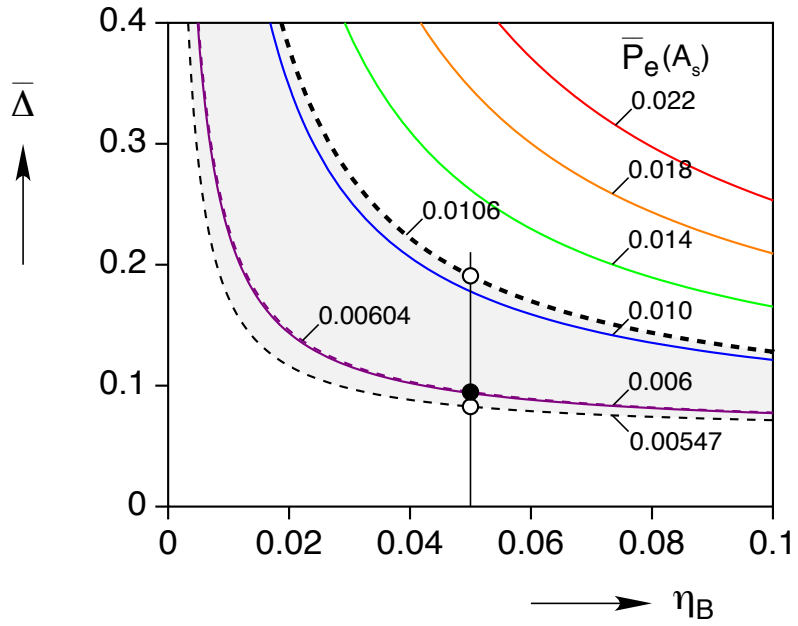


Figure 4.16: Minimum dimensionless power curves to start $M^+ \rightarrow A$ transformation, $\bar{\mathcal{P}}_e(A_s)$, in the design space of dimensionless bias (η_B) and external ($\bar{\Delta}$) spring constants.

temperature, A_f. The gray region is the feasible design space for (η_B, η_E) for the power level $\bar{\mathcal{P}}_e = 0.0106$. The solid black circle is the corresponding coordinates $(\eta_B, \eta_E) = (0.05, 0.025)$ of the example, and the open circles indicate the minimum and maximum values of η_E for the same η_B. Figure 4.17(b) shows a similar design plot where an optimal value of $\bar{\Delta}_{min}$ is chosen according to eq. (4.4.4). This better choice expands the design space to larger permissible values of η_E for a given power level, or allows a lower power level to be used for a given (η_B, η_E) design.

4.5.5 Energy Usage

Another consideration beside power requirements is the amount of input energy necessary to start and finish actuation. This could be important if the power source has a finite

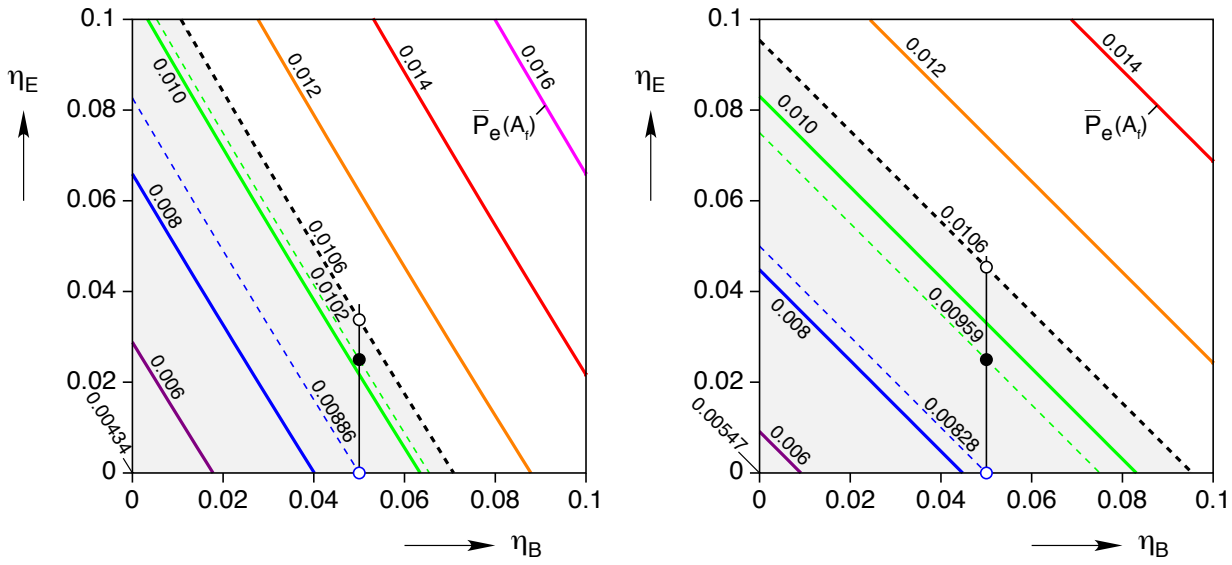


Figure 4.17: Minimum dimensionless power curves to finish $M^+ \rightarrow A$ transformation, $\bar{\mathcal{P}}_{\mathrm{e}}(A_{\mathrm{f}})$, in the design space of dimensionless bias (η_{B}) and external (η_{E}) spring stiffnesses: (a) with fixed $\bar{\Delta} = 0.0948$, (b) with $\bar{\Delta} = \bar{\Delta}_{\min}$ from eq. (4.4.4).

energy storage, such as a battery or capacitor. Since the applied power is assumed constant during the actuation portion of the cycle, the input energy is simply $\mathcal{E}_{\mathrm{in}} = \mathcal{P}_{\mathrm{e}}\Delta t$. Defining a dimensionless input energy as

$$\bar{\mathcal{E}}_{\mathrm{in}} \equiv \frac{\mathcal{E}_{\mathrm{in}}}{\mathcal{E}^*}, \tag{4.5.10}$$

where a characteristic energy has been defined as $\mathcal{E}^* \equiv EAL\beta$. The dimensionless input energy necessary to reach A_{f} from T_a is

$$\bar{\mathcal{E}}_{\mathrm{in}} = \bar{c}\,\bar{\mathcal{P}}_{\mathrm{e}}\,\tau_2, \tag{4.5.11}$$

which, using eqs. (4.4.7) and (4.4.8), works out to be

$$\bar{\mathcal{E}}_{\rm in} = \bar{\mathcal{P}}_{\rm e}\left[\bar{c}\,\ln\left(\frac{\bar{\mathcal{P}}_{\rm e}}{\bar{\mathcal{P}}_{\rm e}+\theta_{\rm a}-\bar{A}_{\rm f}}\right)+\left(\frac{\bar{\mathcal{P}}_{\rm e}+\theta_{\rm a}+\lambda-\bar{\mu}_{\rm c}}{\beta\eta}\right)\ln\left(\frac{\bar{\mathcal{P}}_{\rm e}+\theta_{\rm a}-\bar{A}_{\rm s}}{\bar{\mathcal{P}}_{\rm e}+\theta_{\rm a}-\bar{A}_{\rm f}}\right)-1\right]. \tag{4.5.12}$$

The total input energy can be partitioned as

$$\bar{\mathcal{E}}_{\rm in,S} = \bar{c}\,\bar{\mathcal{P}}_{\rm e}\ln\left(\frac{\bar{\mathcal{P}}_{\rm e}}{\bar{\mathcal{P}}_{\rm e}+\theta_{\rm a}-\bar{A}_{\rm f}}\right), \tag{4.5.13}$$

$$\bar{\mathcal{E}}_{\rm in,T} = \bar{\mathcal{P}}_{\rm e}\left[\left(\frac{\bar{\mathcal{P}}_{\rm e}+\theta_{\rm a}+\lambda-\bar{\mu}_{\rm c}}{\beta\eta}\right)\ln\left(\frac{\bar{\mathcal{P}}_{\rm e}+\theta_{\rm a}-\bar{A}_{\rm s}}{\bar{\mathcal{P}}_{\rm e}+\theta_{\rm a}-\bar{A}_{\rm f}}\right)-1\right]. \tag{4.5.14}$$

where $\bar{\mathcal{E}}_{\rm in,S}$ is the total sensible heating contribution to reach $A_{\rm f}$, and $\bar{\mathcal{E}}_{\rm in,T}$ is the latent heat and spring energy during transformation from $A_{\rm s}$ to $A_{\rm f}$. Equations (4.4.13) and (4.4.12) are plotted against $\bar{\mathcal{P}}_{\rm e}$ in Figure 4.18(a) using the remaining parameters of Table 4.3. One can see that the required input energy drops dramatically as the power is increased, since the time for ambient heat loss is reduced. The shaded areas show the individual contributions of the sensible heat and latent heat/spring energy contributions. Figure 4.18(b) also shows a corresponding dimensional plot, holding all parameters fixed except the input power, to give a sense of the energy and power magnitudes for the numerical example.

In the limit as $\bar{\mathcal{P}}_{\rm e}\to\infty$ the actuation is adiabatic (no ambient heat loss) and the input energy reaches a finite lower bound. The infinite limit can be found by a change of variables to a zero limit and then an application of *L'Hôpital's* rule as

$$\lim_{x\to\infty} x\ln\left(\frac{x+a}{x+b}\right) = \lim_{y\to 0}\frac{\ln\left(\frac{1+ay}{1+by}\right)}{y} = \lim_{y\to 0}\frac{a-b}{(1+ay)\,(1+by)} = a-b, \tag{4.5.15}$$

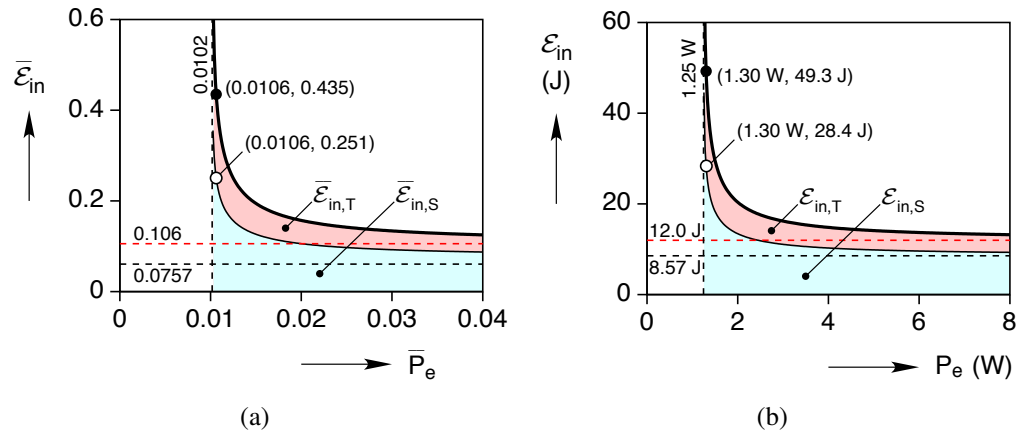


Figure 4.18: Input energies as a function of applied power to reach A_{f} with shaded regions indicated sensible heat and latent heat/spring energy portions. Numerical example values are shown by open and closed circles. (a) Dimensionless plot, holding all SMA parameters and ambient temperature constant, (b) Dimensional plot, holding all parameter values except input power constant.

Accordingly, eqs. (4.4.13), (4.4.14), (4.4.12) become

$$\bar{\mathcal{E}}_{\mathrm{in,S}}^{\infty} = \bar{c}\,\left(\bar{A}_{\mathrm{f}} - \theta_{\mathrm{a}}\right), \tag{4.5.16}$$

$$\bar{\mathcal{E}}_{\mathrm{in,T}}^{\infty} = \frac{1}{2}\beta\eta + \lambda + \bar{\sigma}_{\mathrm{o}}, \tag{4.5.17}$$

$$\bar{\mathcal{E}}_{\mathrm{in}}^{\infty} = \frac{1}{2}\beta\eta + \lambda + \bar{\sigma}_{\mathrm{o}} + \bar{c}\,\left(\bar{A}_{\mathrm{f}} - \theta_{\mathrm{a}}\right), \tag{4.5.18}$$

as $\bar{\mathcal{P}}_{\mathrm{e}} \to \infty$. The lower bound values $\bar{\mathcal{E}}_{\mathrm{in,S}}^{\infty}$ and $\bar{\mathcal{E}}_{\mathrm{in}}^{\infty}$ are shown by dashed lines in Figure 4.18(a).

4.5.6 Energy Efficiency

The numerical example of Section 4.3 applied power longer than was necessary to achieve full actuation, just for illustration purposes. In practice, a lower power could be used to maintain the actuator position, or better yet, a latching mechanism could be

employed. Inducing temperatures above A_f does no additional work against the external spring, so is wasted energy. In this section, we consider the efficiency of the actuator on an energy basis. We compare the input energy (now denoted $\mathcal{E}_\mathrm{in}$, dropping the subscript 2) to the output energy ($\mathcal{E}_\mathrm{out}$) of the actuator. The output energy is simply the work done against the external spring, which is

$$\bar{\mathcal{E}}_\mathrm{out} = \frac{\beta\,\eta_\mathrm{E}}{2\left(1+\eta_\mathrm{B}+\eta_\mathrm{E}\right)^2}, \tag{4.5.19}$$

using a similar dimensionless energy ($\bar{\mathcal{E}}_\mathrm{out} = \mathcal{E}_\mathrm{out}/\mathcal{E}^*$). The energy efficiency is the ratio $\mathcal{E}_\mathrm{out}/\mathcal{E}_\mathrm{in}$ (or $\bar{\mathcal{E}}_\mathrm{out}/\bar{\mathcal{E}}_\mathrm{in}$),

$$\frac{\mathcal{E}_\mathrm{out}}{\mathcal{E}_\mathrm{in}} = \frac{\beta\,\eta_\mathrm{E}}{2\left(1+\eta_\mathrm{B}+\eta_\mathrm{E}\right)^2}\left\{\left(\theta^\mathrm{ss}-\theta_\mathrm{a}\right)\left[\bar{c}\,\ln\left(\frac{\theta^\mathrm{ss}-\theta_\mathrm{a}}{\theta^\mathrm{ss}-\bar{A}_\mathrm{f}}\right)+\left(\frac{\theta^\mathrm{ss}+\lambda-\bar{\mu}_\mathrm{c}}{\beta\eta}\right)\ln\left(\frac{\theta^\mathrm{ss}-\bar{A}_\mathrm{s}}{\theta^\mathrm{ss}-\bar{A}_\mathrm{f}}\right)-1\right]\right\}^{-1}. \tag{4.5.20}$$

The best efficiency in the limit of infinite power input is

$$\left(\frac{\mathcal{E}_\mathrm{out}}{\mathcal{E}_\mathrm{in}}\right)^\infty = \frac{\beta\,\eta_\mathrm{E}}{\left(1+\eta_\mathrm{B}+\eta_\mathrm{E}\right)^2}\left[\frac{1}{\beta\eta+2\left(\lambda+\bar{\sigma}_\mathrm{o}+\bar{c}\,\left[\bar{A}_\mathrm{f}-\theta_\mathrm{a}\right]\right)}\right]. \tag{4.5.21}$$

The efficiency curve, eq. (4.4.20), is plotted in Figure 4.19 (bold line) against the dimensionless power input using $\theta^\mathrm{ss} \equiv \bar{\mathcal{P}}_\mathrm{e} + \theta_\mathrm{a}$ with the remaining parameters taken from Table 4.3. The efficiency of our numerical example is quite low (0.147%), while the best case efficiency is 0.605% for infinite power input. The case of optimal bias spring mismatch ($\bar{\Delta}_\mathrm{min}$) is also shown for comparison (thin line), having slightly better efficiency (0.206%) at the chosen input power and upper bound efficiency (0.634%). While the energy density

of SMA's is several orders of magnitude larger than other typical adaptive materials [45], the energy usage can be quite lossy for a given work output, which is a well-known issue for typical SMA actuators. As a final case, it is also interesting to study the best possible energy efficiency of dead loading ($\eta = 0$). For this case, the constants in eq. (4.2.11) simplify to

$$a_1 = \frac{\bar{\sigma}_{\mathrm{o}} - \left(\bar{\mathcal{P}}_{\mathrm{e}} + \theta_{\mathrm{a}} + \bar{\mu}_{\mathrm{c}}\right)}{\lambda + \bar{\mathcal{P}}_{\mathrm{e}} + \theta_{\mathrm{a}} + \bar{\mu}_{\mathrm{c}}},$$
$$b_1 = 0, \tag{4.5.22}$$

where $\bar{\sigma} = \bar{\sigma}_{\mathrm{o}}$ is the dead load stress. The governing equation, (4.2.9), simplifies to

$$\bar{c}\, a_1 - (1 + a_1)\, \xi'(\tau) = 0, \tag{4.5.23}$$

which has the simple linear solution in time

$$\xi(\tau) = (\tau - \tau_1)\, \frac{\bar{c}\, a_1}{1 + a_1}. \tag{4.5.24}$$

Proceeding as before, accounting for the time to reach A_{s} and the time to transform the SMA element, the dead load energy efficiency is

$$\frac{\mathcal{E}_{\mathrm{out}}}{\mathcal{E}_{\mathrm{in}}} = \bar{\sigma}_{\mathrm{o}} \left\{ \bar{c}\, \bar{\mathcal{P}}_{\mathrm{e}} \ln \left[\frac{\bar{\mathcal{P}}_{\mathrm{e}}}{\bar{\mathcal{P}}_{\mathrm{e}} + \theta_{\mathrm{a}} - (\bar{\mu}_{\mathrm{c}} + \bar{\sigma}_{\mathrm{o}})} \right] + \frac{\bar{\mathcal{P}}_{\mathrm{e}}\ (\lambda + \bar{\sigma}_{\mathrm{o}})}{\bar{\mathcal{P}}_{\mathrm{e}} + \theta_{\mathrm{a}} - (\bar{\mu}_{\mathrm{c}} + \bar{\sigma}_{\mathrm{o}})} \right\}^{-1}, \tag{4.5.25}$$

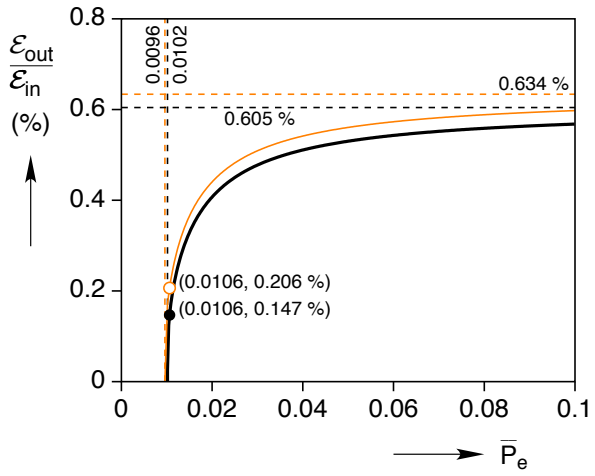


Figure 4.19: Energy efficiency ($\mathcal{E}_{\text{out}}/\mathcal{E}_{\text{in}}$) as a function of dimensionless power ($\bar{\mathcal{P}}_{\text{e}}$) for the design of Table 4.3 (bold line) and with minimum bias spring mismatch (thin line).

Note that $\bar{A}_{\text{s}} = \bar{A}_{\text{f}} = \bar{\mu}_{\text{c}} + \bar{\sigma}_{\text{o}}$ in this case. The limiting case of $\bar{\mathcal{P}}_{\text{e}} \to \infty$ is

$$\left(\frac{\mathcal{E}_{\text{out}}}{\mathcal{E}_{\text{in}}}\right)^{\infty} = \frac{\bar{\sigma}_{\text{o}}}{\bar{\sigma}_{\text{o}} + \lambda + \bar{c}\,[\bar{\mu}_{\text{c}} + \bar{\sigma}_{\text{o}} - \theta_{\text{a}}]}, \tag{4.5.26}$$

Figure 4.20 shows contours of energy efficiency, eq. (4.4.25), in the space of $\bar{\sigma}_{\text{o}}$ and $\bar{\mathcal{P}}_{\text{e}}$, leaving the SMA parameters and ambient temperature fixed according to Table 4.3. The efficiency generally improves as $\bar{\sigma}_{\text{o}}$ increases, due to the greater output work done, except near minimum powers (slanted dashed line) where time for actuation becomes long. Here, one can see that energy efficiencies of a few percent can be achieved for moderate dimensionless stress and power levels.

4.5.7 Potential Model Refinements

The actuator model presented herein was developed primarily with simplicity and ease of use in mind. Once the parameters $(\beta, \eta_{\text{B}}, \bar{\Delta})$ have been calibrated to give an accurate

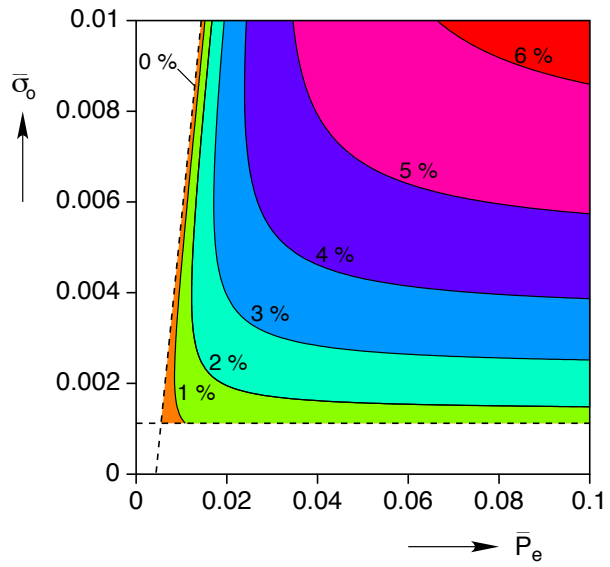


Figure 4.20: Contours of energy efficiency ($\mathcal{E}_{\rm out}/\mathcal{E}_{\rm in}$) as a function of dimensionless power ($\bar{\mathcal{P}}_{\rm e}$) and stress ($\bar{\sigma}_{\rm o}$) for a dead load actuator.

prestress $\bar{\sigma}_{\rm o}$, we expect the model to give reasonably accurate performance predictions. However, as noted in Section 4.4.1, allowing $\eta_{\rm B}$ and/or $\bar{\Delta}$ to change for fixed β would produced only rough approximations of the prestress, since the low temperature martensite isothermal responses are actually quite nonlinear and temperature dependent.

Here, we discuss how the model could be generalized somewhat for design studies, focussing on improving the predictions of $\bar{\sigma}_{\rm o}$, thereby giving more accurate calculations of the actuator stroke over the design space of $(\eta_{\rm B}, \bar{\Delta})$. There are many ways this could be accomplished, for example using a more sophisticated SMA constitutive model, but instead we outline here an incremental extension of the current model that seems to be a pragmatic alternative that retains the useful analytical solutions. For design purposes, the limitation of the current actuator model stems from treating β as a material constant, independent of the actuator design and ambient temperature. We now relax this interpretation and use it as a fitting parameter for a range of different bias spring designs in order to more accurately

capture the initial stress, $\bar{\sigma}_{\rm o}$.

Figure 4.21(a) shows eight isothermal, displacement-controlled, load-unload experiments on NiTi wire specimens (as used in Chang, *et. al.* [15]) at relatively low temperatures between -60 °C and 10 °C. Each as-received specimen was cooled monotonically from room temperature before the experiment, so many started in the R-phase initially and were transformed to tensile martensite during the loading process, hence the appearance of a double "knee" at low strains for the lowest temperature experiments. (Note that the appearance of the R-phase has no bearing on the quality of the actuator predictions, since the SMA operates at much higher stresses and temperatures at low strains, where the R-phase never appears.) The experiments at the highest temperatures actually show the superelastic behavior of the material. Points at the termination of the loading plateaus are shown by large dots. The overall envelope of these points was fitted by a three-parameter curve (M^+-fit) of the form (see Figure 4.21(b))

$$\bar{\sigma}(\varepsilon_0) = a + b\, e^{c\,\varepsilon_0}, \tag{4.5.27}$$

with fitting constants

$$\{a, b, c\} = \{7.9561 \times 10^{-4}, 1.7118 \times 10^{-5}, 80.622\}. \tag{4.5.28}$$

The curve represents the loci of the initial states of the actuation cycle for the whole class of designs. Again, we must check that the assembly takes the SMA to a point somewhere beyond the plateau of the isothermal (cold) response, i.e. prestress must intersect the M^+-

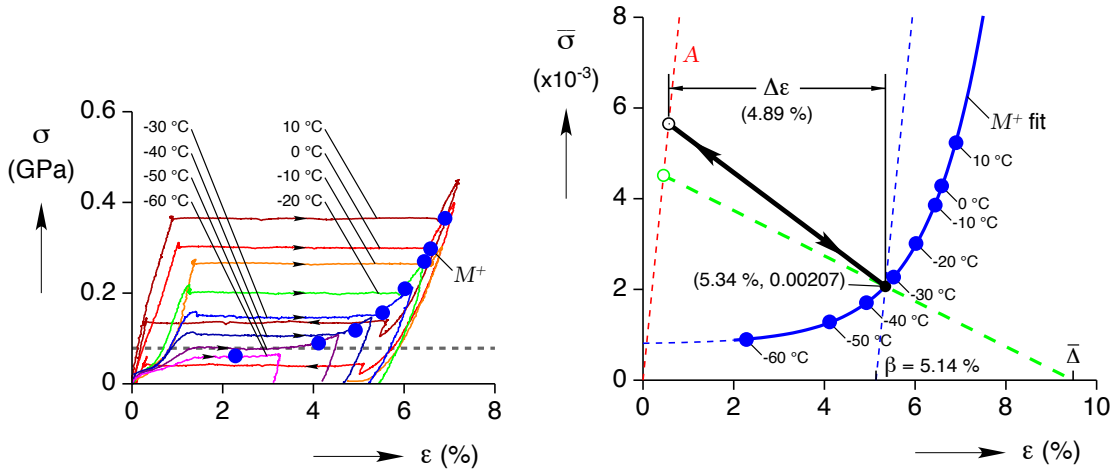


Figure 4.21: (a) Low temperature isothermal experiments on NiTi wire (same material as used in Chang, *et. al.* [15]) with large solid dots showing the termination points of the loading plateaus. The dotted gray line shows the predicted value for martensite reorientation from the constitutive model ($\sigma_M = 78.5$ MPa), which agrees with the plateau stress of the experiment at -50 °C. (b) Dimensionless stress-strain paths of the actuator, $(\bar{\Delta}, \eta_{\mathrm{B}}) = (0.0948, 0.05)$, for bias spring only (thick dashed line) and with external spring, $\eta_{\mathrm{E}} = 0.025$ (thick solid line). The M^+-fit curve is a fit of the loading plateau termination points, according to eq. (4.4.27), which represents the prestress $\bar{\sigma}_{\mathrm{o}}$ of the actuator at "full" M^+.

fit curve above the ambient temperature indicated on the curve. (This criterion is satisfied, for example, by the intersection point shown in Figure 4.21(b), which corresponds to the plateau stress for an isothermal response near -33 °C, well above the point corresponding to the ambient temperature -50 °C.) Each design, i.e. combination of $(\bar{\Delta}, \eta_{\mathrm{B}})$, can be calibrated to a particular value of β, that once fixed, is associated with M^+ ($\xi_0 = 1$). For our previous values $(\bar{\Delta}, \eta_{\mathrm{B}}) = (0.0948, 0.05)$ a new value of transformation strain was calculated, $\beta = 0.0514$ (as shown in Figure 4.21(b)), somewhat smaller than used before (0.0591), but probably more accurate in light of these particular experiments. The actuator path in SMA stress-strain space is shown in Figure 4.21(b) for the cases with a bias spring only (dashed line) and with an external spring (thick solid line) for which the predicted dimensionless stroke is 4.89 %.

The calculations of fitted transformation strain (β) and the cold assembly strain (ε_0) are accomplished as follows. The intersection point (small solid dot) in Figure 4.21(b) is defined by equating the dimensionless stresses

$$\varepsilon_0 - \beta = a + b\, e^{c\,\varepsilon_0}, \tag{4.5.29}$$

but now ε_0 is a function of β and $\bar{\Delta}$ according to eq. (4.1.39). The solution is

$$\varepsilon_0 = 1 - \frac{1}{\bar{\Delta}} \left[\frac{a}{\eta_{\mathrm{B}}} + \frac{1}{c} W \left(\frac{b\,c}{\eta_{\mathrm{B}}} e^{c\left(\bar{\Delta} - \frac{a}{\eta_{\mathrm{B}}}\right)} \right) \right], \tag{4.5.30}$$

where W is again the Lambert function (!), and $\{a, b, c\}$ are the fitting constants of the M^+ envelope. Accordingly, the transformation strain is found to be

$$\beta = 1 - \frac{1+\eta_{\mathrm{B}}}{\bar{\Delta}} \left[\frac{a}{\eta_{\mathrm{B}}} + \frac{1}{c} W \left(\frac{b\,c}{\eta_{\mathrm{B}}} e^{c\left(\bar{\Delta} - \frac{a}{\eta_{\mathrm{B}}}\right)} \right) \right]. \tag{4.5.31}$$

Contours of constant ε_0 and β in the design space of $(\eta_{\mathrm{B}}, \bar{\Delta})$ are shown in Figure 4.22(a) and Figure 4.22(b), respectively. The open circles in the figure are the values at $(\eta_{\mathrm{B}}, \bar{\Delta}) = (0.05, 0.0948)$ associated with $(\varepsilon_0, \beta) = (5.34\ \%,\ 5.14\ \%)$. These show monotonically increasing values of ε_0 and β as either value of $(\eta_{\mathrm{B}}, \bar{\Delta})$ is increased.

The dimensionless actuator stroke can be calculated from eqs. (4.4.2), (4.1.35), and (4.4.30) as

$$\Delta\varepsilon = \frac{1}{1+\eta_{\mathrm{B}}+\eta_{\mathrm{E}}} \left\{ \bar{\Delta} - (1+\eta_{\mathrm{B}}) \left[\frac{a}{\eta_{\mathrm{B}}} + \frac{1}{c} W \left(\frac{b\,c}{\eta_{\mathrm{B}}} e^{c\left(\bar{\Delta} - \frac{a}{\eta_{\mathrm{B}}}\right)} \right) \right] \right\}. \tag{4.5.32}$$

A contour plot of constant $\Delta\varepsilon$ in the space of $(\eta_{\mathrm{B}}, \eta_{\mathrm{E}})$ for $\bar{\Delta} = 0.0948$ is shown in Figure 4.23. One can see that for constant η_{B}, increasing η_{E}, the actuator stroke is monotonically reduced, yet for constant η_{E}, increasing η_{B}, the actuator stroke increases then decreases. For our chosen parameters, $(\eta_{\mathrm{B}}, \eta_{\mathrm{E}}) = (0.05, 0.025)$, the actuator stroke is $4.78\ \%$ as shown by the open circle in the plot. The stroke without an external spring is slightly larger, $4.89\ \%$. Additionally, the dotted line in the plot shows that the stroke changes in a non-monotonic manner (increasing then decreasing) as $\eta_{\mathrm{B}} + \eta_{\mathrm{E}}$ is increased through this point (or most others for that matter), contrary to the overly simplistic eq. (4.4.3).

As a final comment, the original actuator model was simple, making it especially

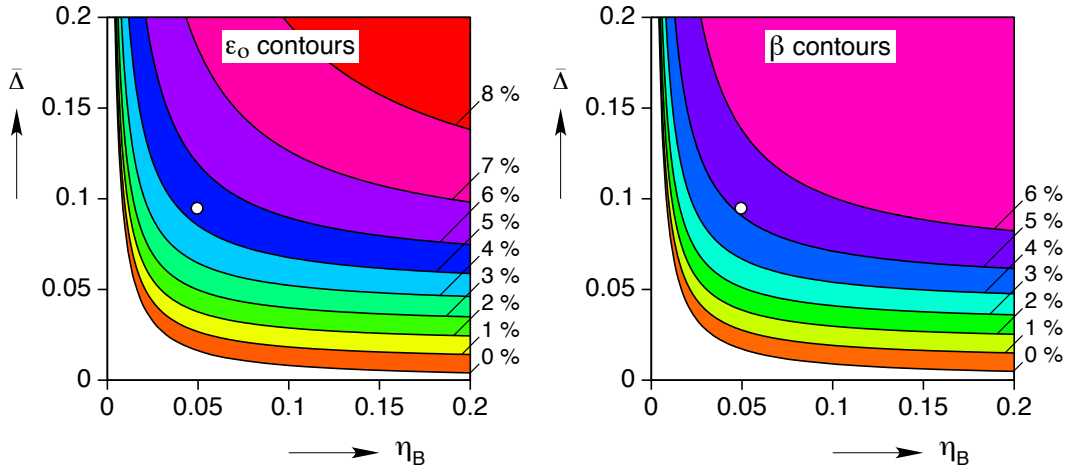


Figure 4.22: Contours of constant ε_0 (a) and β (b) as a function of $(\eta_{\mathrm{B}}, \bar{\Delta})$. The open circles show the points $(\eta_{\mathrm{B}}, \bar{\Delta}) = (0.05, 0.0948)$.

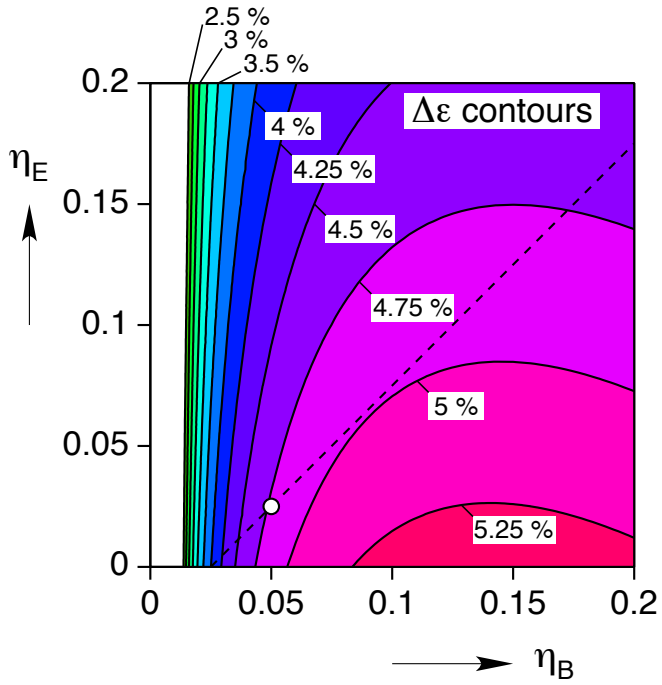


Figure 4.23: Contours of constant dimensionless actuator stroke, $\Delta\varepsilon$ in the space of dimensionless spring constants $(\eta_{\mathrm{B}}, \eta_{\mathrm{E}})$ for $\bar{\Delta} = 0.0948$. The open circle shows the point $(\eta_{\mathrm{B}}, \eta_{\mathrm{E}}) = (0.05, 0.025)$.

convenient for first-order design calculations. The purpose of this section was to suggest a possible refinement, reusing as much of the previous actuator model as possible, yet this was achieved at the expense of introducing some additional complexity. We have not yet investigated all the implications on the design guidelines presented in earlier parts of this section using this more refined model. That is left for future work, where we also intend to perform relevant experiments on SMA bias spring actuators using the current SMA material for direct comparison, in order to validate predictions over a wide range of conditions and design parameters.

4.6 Summary & Conclusions

A reduced-order thermodynamic model was presented for a prototype uniaxial SMA/bias-spring actuator immersed in a thermal bath and actuated by Joule heating. The constitutive model, a derivative of a more complex, experimentally validated model Chang, *et. al.* [15], was used to solve mechanical equilibrium, heat transfer, and kinetic laws, but in lumped form. A dimensionless set of equations was derived, resulting in a minimum set of dimensionless parameters governing the transient actuator response.

The actuator behavior was governing by a first order, nonlinear ordinary differential equation (ODE) during phase transformation, rather than the usual set of partial differential equations (PDEs). Analytical solutions were then derived for the time-dependent actuator response, during both sensible (only) heating/cooling and the combination of sensible/latent heating/cooling for martensitic transformation. To our knowledge, the latter is the first such analytical solution available for an SMA actuator system. During transforma-

tion, the evolution of strain, stress, temperature, and tensile martensite phase fraction were found in terms of the special *Lambert* function, given a prescribed ambient temperature and input power. The analytical solution compared favorably to a finite-element simulation using the parent model [15].

The analytical nature of the transient solution to the governing ODE allowed us to study several important actuator performance indicators in terms of the governing non-dimensional parameters. Actuation time, spring size, maximum stress, minimum power, and energy efficiency were derived in closed form. By contrast, countless simulations using traditional numerical techniques to solve PDEs would be required to encompass the entire range of potential actuators. The simpler approach herein gives the designer a quick and convenient model to assess the various measures of actuator performance.

While the model presented gives a good first approximation of actuator performance, its most limiting assumptions are (1) uniform temperature field, (2) uniform strain field, and (3) the rather crude fit of the low temperature martensite reorientation response by the simple constitutive model. Assumption (1) is reasonably satisfied if the ends of the SMA wire are thermally insulated. Assumption (2) is reasonably satisfied if the SMA wire has been conditioned (pre-cycled) to eliminate the occurrence of strain localization associated with transformation fronts, yet the simple actuator model still gave satisfying results compared to the finite element simulation (that included transformation fronts), at least in terms of the gross behavior of the actuator. Assumption (3) was not an issue, provided the transformation strain parameter (β) gave a good estimate of the prestress in the SMA element. This prestress also needed to be sufficiently large to ensure tensile

martensite existed in the actuator's cold state. To achieve accurate stroke predictions for a range of design parameters and ambient conditions it was necessary to reinterpret β as a design parameter, rather than a material constant, by fitting it to the post-plateau tensile martensite responses.

We expect this simplified setting will be a useful first-order calculational tool for sizing SMA actuators, predicting their performance, and for use in design optimization software to effectively integrate such actuators into complex systems. We hope it will be a convenient tool for engineers to perform initial actuator design studies in a systematic way.

Chapter V

Localized Phenomena and Rate Effects in Tensile and Compressive NiTi Tubes

This chapter aims to combine a discussion of rate effects on uniaxial SMAs with a series of interesting NiTi tube tension, compression, and bending experiments. NiTi Tubes are not widely used in industry in their as-received form, and are commonly cut into more complex structures, especially as arterial stents. Though there have been limited efforts to perform biaxial loading of NiTi strips [72], tubes remain the most fruitful SMA structure for biaxial loading because of the limited need for complex buckling suppression systems [5, 69, 94, 100].

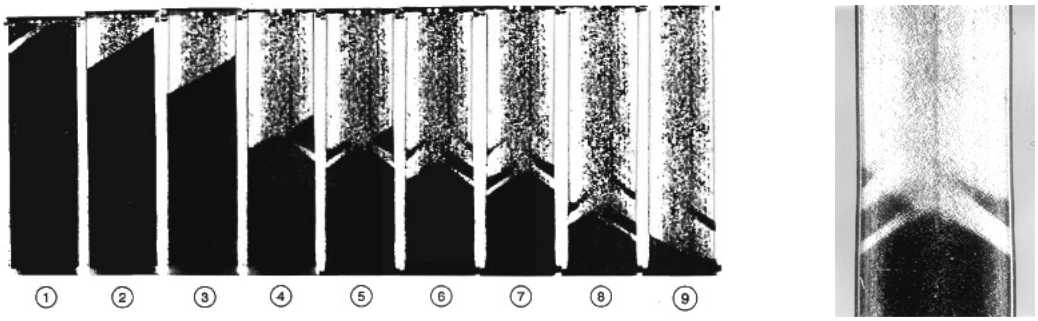
5.1 Background

Chapter II has already investigated the concept of transformation fronts in NiTi wires, a 1D SMA structure. The effects of loading rate and thermomechanical coupling on NiTi wires have been thoroughly explored by Shaw et al. [50, 87]. A natural extension of this effort is to look at NiTi strips in tension, a 2D structure. In both NiTi and mild steel strips, Shaw and Kyrikides [88] described Lüders-like bands of localized deformation, inclined at an angle to the axis of loading. Figure 5.1 shows just one example of the evolution

of these fronts during displacement-controlled loading ($A \rightarrow M^+$transformation). They showed that in a 1D sense, fronts tended to propagate at the same velocity, provided the thermal boundary conditions are uniform (i.e. no ambient temperature or film coefficient gradients). Compatibility and equilibrium then require the front velocity c to be related to the number of fronts n and global strain rate $\dot{\delta}$ by

$$c = \frac{\dot{\delta}}{n\Delta\varepsilon}. \tag{5.1.1}$$

In Figure 5.1, the angle of the fronts to the loading axis always occur at 50-60°. In this case, the use of a shortened strip increased the bending moment in the strip due to the inclined front, causing the front to change directions, to create a "crisscross" effect. Neglecting all bending moments and elastic strains, the origin of this angle is a simple compatibility argument. With those assumptions, consider the infinite plane of A and M^+ in Figure 5.2, the two phases separated by a straight boundary, inclined from the loading axis (vertical) an the angle α. Since the deformation is volume-preserving, the M^+ phase has a lateral strain of $\Delta\varepsilon_2 = -\frac{1}{2}\Delta\varepsilon_1$, giving a Mohr's circle centered at $\Delta\varepsilon = \frac{1}{4}\Delta\varepsilon_1$, and a radius of $\frac{3}{4}\Delta\varepsilon_1$, shown on the right of Figure 5.2. Compatibility requires no jump in extensional strain along the boundary, neglecting the small incompatibility in the thickness direction. In order to satisfy this condition the boundary must be rotated around Mohr's circle by the angle 2α ($\alpha = \sin^{-1}\sqrt{\frac{2}{3}} = 54.7°$). A more complex analysis of this problem has been performed by Li and Sun [61]. Their model included elastic effects and treated an inclusion of M^+ within A, rather than a single boundary [61], but still came within 1° to



(a) (b)

Figure 5.1: Evolution of $A \rightarrow M^+$ transformation fronts in $12.7 \times 2.5 \times 0.25$ mm superelastic NiTi strips, reprinted with permission from Shaw *et. al.* [85]. (a) Strip at constant time intervals during loading, (b) detail view of a criss-crossing transformation front

this simple derivation with an estimate of $\pm 55.7^{\text{o}}$.

Several researchers have already looked extensively at localization phenomena in tension and torsion of NiTi tubes. One recent effort of note is that of Sun et. al. [31,40,61,70], who used surface profilimetry to measure localized transformation fronts in superelastic NiTi microtubes. They found that, outside of the special case of converging fronts, two morphologies dominated the localized front were prevelant: (1) a finely branched, akin to Figure 5.1(b), but overall ring-shaped front and (2) a helical inclusion consisting of 2 nearly parallel helical fronts. The helical inclusion in Figure 5.3(a) is inclined at an angle near the theoretical value of 54.7^{o}, and curls around the tube in the hoop (θ) direction. The ringed front in Figure 5.3(b) consists of a number of small, identical inclusions repeating in a ring around the tube's axis. Notice that though the overall front is normal to the tube axis, the actual interface is inclined between 63 and 72^{o} to the loading axis. As the overall "neck" propagates along the tube axis, the tips of these inclined branches actually rotate around the tube, each following its own helical path. The mere existence of the ringed front, as

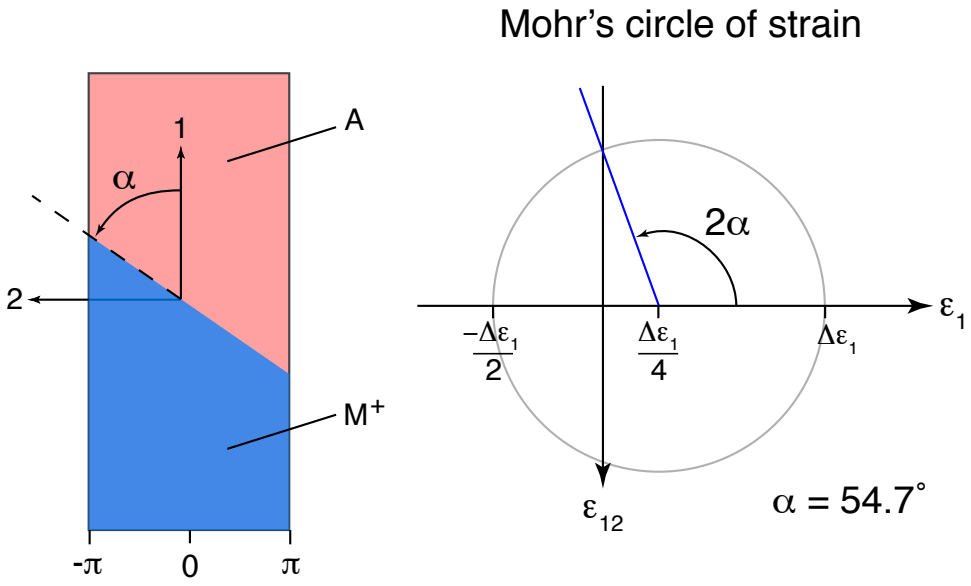
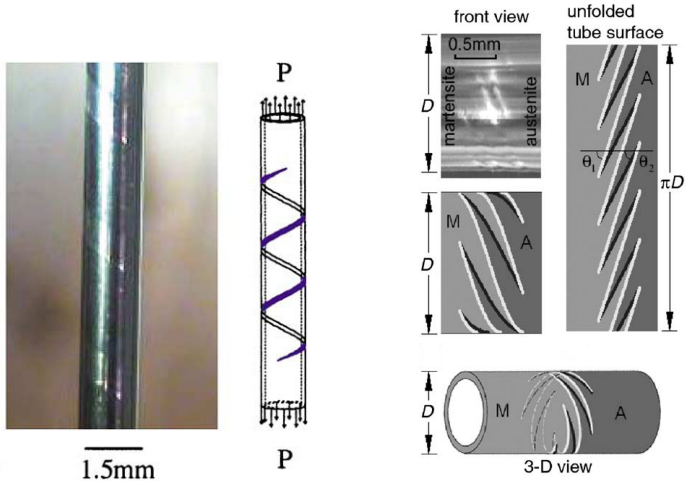


Figure 5.2: A derivation of $A \rightarrow M^+$ front inclination angle α. If deformations are isotropic, volume preserving, and extension along the front axis is constrained to be 0, Mohr's circle of strain allows a simple geometric solution.

opposed a single ring normal to the axis, shows how critical the angle of front inclination is to minimizing the tube's elastic strain energy arising from compatibility effects. In the following section, the temperature data used to track fronts is not able to resolve these individual branches, but they likely still exist.

Favier et al. (2007) have performed several experiments on superelastic NiTi tubes, taking full-field temperature and strain data via digital image correlation (DIC) [30]. They have also developed a novel method of processing temperature data to remove effects of convective cooling and conductive heat transfer, producing a full-field map of heat sources [81]. However, they have not yet published a study comparing responses across multiple loading rates.

Tension-compression asymmetry in the response of NiTi SMA material has also been well documented in the literature, in wires [77], strips [72], and single crystals [36]. Some



(a) (b)

Figure 5.3: The two primary propagating front geometries in SMA tubes: (a) helical front with photo and schematic [94], (b) ringed front, with cartoons and photos during loading [31]. Reprinted with permission.

amount of asymmetry has been shown to occur regardless of texture [72]. The mechanism behind the asymmetry is largely attributed to the activation of different martensite habit plane variants (HPVs) depending on the stress state. Different HPVs in tension and compression have been both observed [63] and modeled [35]. Still, many SMA models still assume that isotropic elastic properties are sufficient to model bending responses [83] and experimentalists often extend this assumption to assume a static neutral axis [6], usually because a typical characterization of an SMA material only includes a tension response. For this reason, there is still a need to call attention to tension-compression asymmetry in NiTi polycrystals, as even the simplest assumption about the asymmetry (half $\Delta\varepsilon$, double $\partial\sigma_p/\partial T$) could greatly increase the accuracy of bending models.

5.2 Tension Experiments

All tubes used in this chapter were cold-drawn, nearly equiatomic room temperature (RT) superelastic NiTi tubes from SAES Memry, with a quoted A_s temperature of -15 °C. They have an outer diameter $D = 3.175\text{mm}$ and wall thickness $h = 0.3175\text{mm}$. To determine the dominant texture, a tube sample was subjected to X-ray diffraction at the Georgia Institute of Technology. The results in Figure 5.4 show that the tube has a slight [111] texture, a common property of drawn SMA material. Some attempts were made to generate a thermogram via DSC, but the geometry of the sample caused some difficulty obtaining high quality results.

5.2.1 Tension Experimental Setup

Experiments were performed in an electro-mechanical load frame (Instron model 5585) . Figure 5.5 shows how each end of a specimen was held in a wedge grip between two hardened steel toothed plates. The lower grip was held stationary while the upper grip was attached inline with a load cell and moved under displacement control. It was expected that some slippage would occur at the grip-specimen interface, so average strain was additionally measured along a local gauge length by a laser extensometer (LE). Similar to experiments in Chapter III, the LE measured the distance between two self-adhesive retro-reflective tags affixed to the tube. A small section of tube protruded from the top and bottom of each grip, providing a convenient mounting point for a flexible silicone tube to flow fluid from a temperature-controlled bath through the center of the specimen. In this manner, isothermal experiments (reasonably neglecting the temperature gradient across the

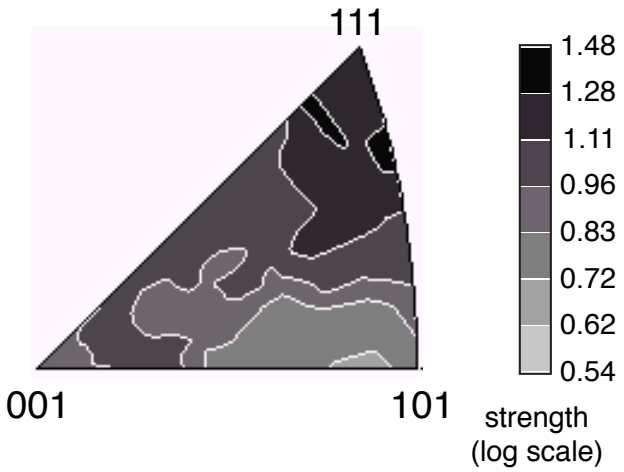


Figure 5.4: Equal-area texture projection for a NiTi tube specimen– the [111] direction is aligned with the tube axis.

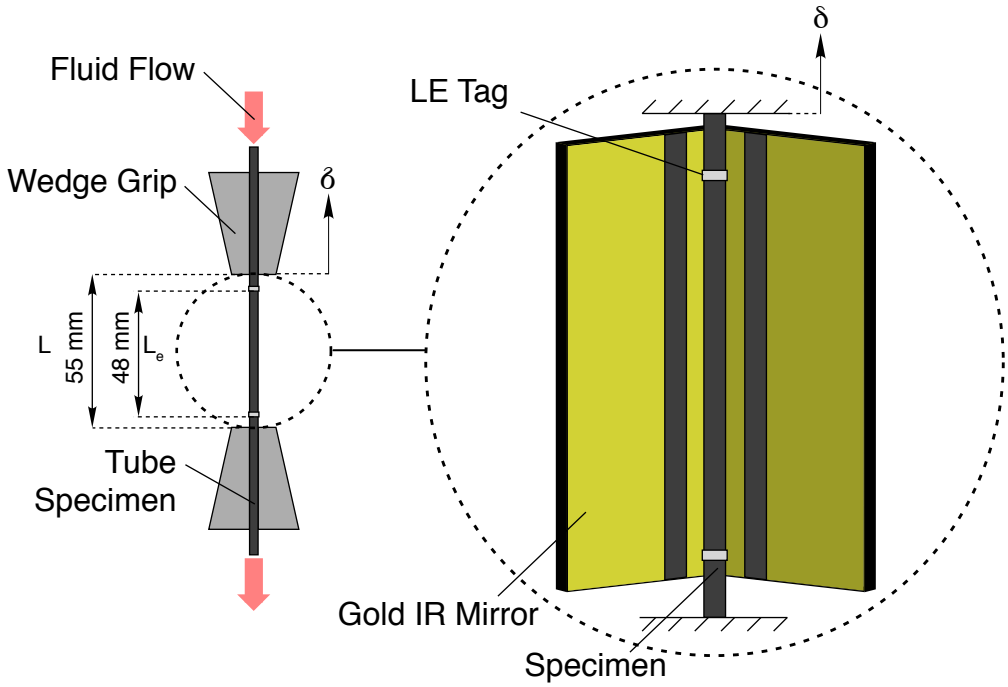


Figure 5.5: Tube tension setup geometry

tube wall) could be performed at any temperature allowed by the bath-fluid combination (-10 to 120 °C) without the need for a thermal chamber.

Temperature was measured using an infrared camera (Inframetrics SC1000) with a 256x256 CCD, a pitch (at the distance used) of about 0.35 mm/pixel, and a precision of 0.1 °C. In order to capture full-field temperature information around the entire circumference, a pair of gold-plated first-surface mirrors were held at a 120 ^{o} angle just behind the tube. The mirror reflection coefficient within camera's sensitive spectrum (near-infrared, $3-5\mu\mathrm{m}$) was in excess of 98 $\%$. Because of the short focal plane of the IR camera, the mirrors needed to be held within 5 mm of the specimen. The camera, mirrors, and specimen were positioned so that the specimen and its reflections were aligned with the pixel grid to within one pixel over the entire grip length, in order to simplify post-processing. Figure 5.6 shows a photo of the setup, without the flexible tubes used for internal fluid flow.

It was important that the specimen to have a sufficiently high emissivity, otherwise the infrared camera would measure the temperature from radiation reflecting off of the specimen, rather than from the specimen itself. A perfect black body has an emissivity $\epsilon = 1$, while a perfect mirror has $\epsilon = 0$. Figure 5.7(a) shows how the as-received tubes had been polished to a mirror finish with $\epsilon = 0.2$, so to prepare them for infrared imaging they were painted with a matte black high-temperature paint, Rustoleum™ High heat. The painted tube in Figure 5.7(b) had a much improved emissivity $\epsilon = 0.93$.

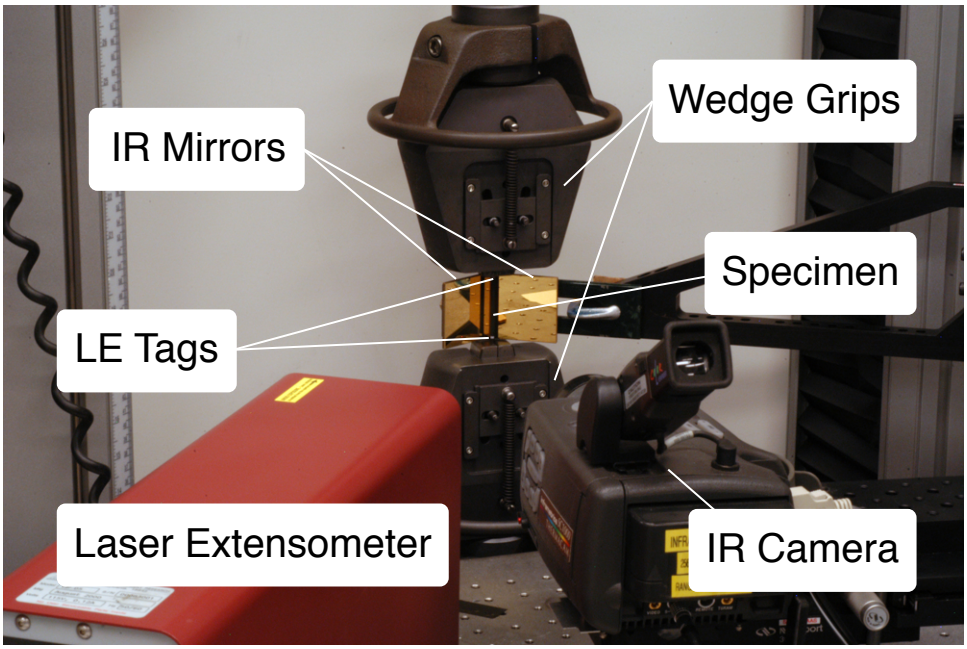


Figure 5.6: Tube tension setup photograph

(a)

(b)

Figure 5.7: (a) bare and (b) painted 3.175 mm diameter NiTi tubes

5.2.2 Infrared Image Post-Processing

Each infrared image contains three images of the specimen, shown in Figure 5.5, which were "unwrapped" and "stitched" together using a custom algorithm developed in Matlab™. Since radiation from the cylindrical surface of the specimen took a different path to reach the IR sensor, either straight-on or reflected off a mirror, several corrections were necessary to match the images. First, the mirror images were flipped along the vertical axis, since they were reflections. Next, a linear temperature correction (determined experimentally, on the order of < 1°C) was applied to account for the $< 2\ \%$ loss in radiation to the mirror. The mirror images were stretched vertically to account for the parallax created by their greater distance from the camera. Since each experiment had a slightly different mirror position (it needed to be swung away for specimen installation), this parameter was determined manually (and was on the order of a 3 % stretching). To account for the nonlinear radiation intensity distribution from the curved tube surface, another linear temperature correction was applied to the edges of the image.

Using a simple geometric argument, the angle from the crest of the tube, θ, can be related to the horizontal distance (pixel position) y through the relation $\sin\theta = \frac{y}{R}$ where R is the tube radius. This relationship was used to map the horizontal distance from the center of the tube to a circumferential position. Since the relationship is highly nonlinear towards the edge of the image ($y = R$) only the center 120° of each 180° image was used. In the final step, the three separate images were stitched together along two manually determined boundaries to form a full, unwrapped temperature map of the tube's lateral surface.

In the end, a tube infrared mirror image underwent 7 various corrections and process-

ing functions, most of them either experimentally determined through calibration experiments, or manually adjusted until temperature features were aligned between all three images. In some cases the result was quite good, especially in experiments like Figure 5.20(c) performed at temperatures near the background temperature of 22 °C, since the error, and therefore corrections, were minimal. As the tube temperature rose above the background temperature, processing was only somewhat effective, especially near the boundaries between images. In these cases, like Figure 5.9, temperatures from the tube crest was still reliable but portions near the edge can be up to 4 °C low.

5.2.3 Isothermal Characterization

A series of isothermal tension experiments were performed to characterize the NiTi tubes. A relatively constant temperature was maintained, as described in section 5.2.1, by pumping fluid, ranging from 10 to 70 °C, from a heated bath through the center of the tube. As an example, the 60 °C experiment is presented in detail. Figure 5.8 shows the engineering stress history of the 60 °C experiment, plotted over the temperature history from the crown of the tube facing the camera. The left vertical axis represents the normalized distance from the bottom grip x/L. The latent heat of transformation between A and M causes self-heating, and it is this temperature rise during loading ($A \rightarrow M^+$) and temperature drop during unloading ($M^+ \rightarrow A$) that marks localized transformation fronts. The two black streaks neat $x/L = 0.05$ and $x/L = 0.95$ represent the position of the LE tags, which read as an incorrect temperature because of their low emissivity. The black portion at the top of the image represents the motion of the upper grip, which moved at the constant

velocity $\dot{\delta}/L = 5 \times 10^{-4}$ s^{-1}. Note that since some convective cooling occurred as the fluid was flowing from top to bottom, there is a slight gradient in temperature (about 2 °C) along the tube axis. To show the front morphology, Figure 5.9 shows selected full-field infrared images, with a slight gap at the top and bottom to allow for mirror clearance. Because grip slippage was expected (and observed), the mechanical response with respect to the extensometer gage length from the LE tags is shown in Figure 5.10.

The loading response of the tube closely follows that of the superelastic wire in Figure 2.20. From 0 to 50 s, the tube responded nearly elastically to the loading with an A modulus of 70.5 GPa, reaching a local maximum $P/\mathrm{A}_0 = 647\mathrm{MPa}$. From 40 to 140 s, the stress varied somewhat but there was little sign of transformation within the LE tags. This behavior can be attributed to slippage at the grips due to premature $A \rightarrow M$ transformation triggered by stress concentrations. At ① a single front emerged from the lower grip, though a nucleation peak similar to that in Figure 2.21 was suppressed by the stress concentration at the grips. That single ringed front, which is likely the internally branched geometry of Figure 5.3(b), continued up the tube and is marked by local temperature rise of about 2 °C. A second ringed front started from the upper grip at about 180 s moving downward. Note that the bottom front moved slightly faster than the top front, about $5.7 \times 10{-}3/\mathrm{s}$ vs. $2.8 \times 10{-}3/\mathrm{s}$, which is contrary to the relationship in (5.1.1). Part of this difference (up to $1 \times 10^{-3}/\mathrm{s}$ can be attributed to a "doppler" effect. Since the IR images were taken in the current configuration, the top front was moving against the grip motion while the bottom front moved with it. The rest of the difference is likely due to the small temperature gradient along the tube axis, as the front in the colder region at the lower end

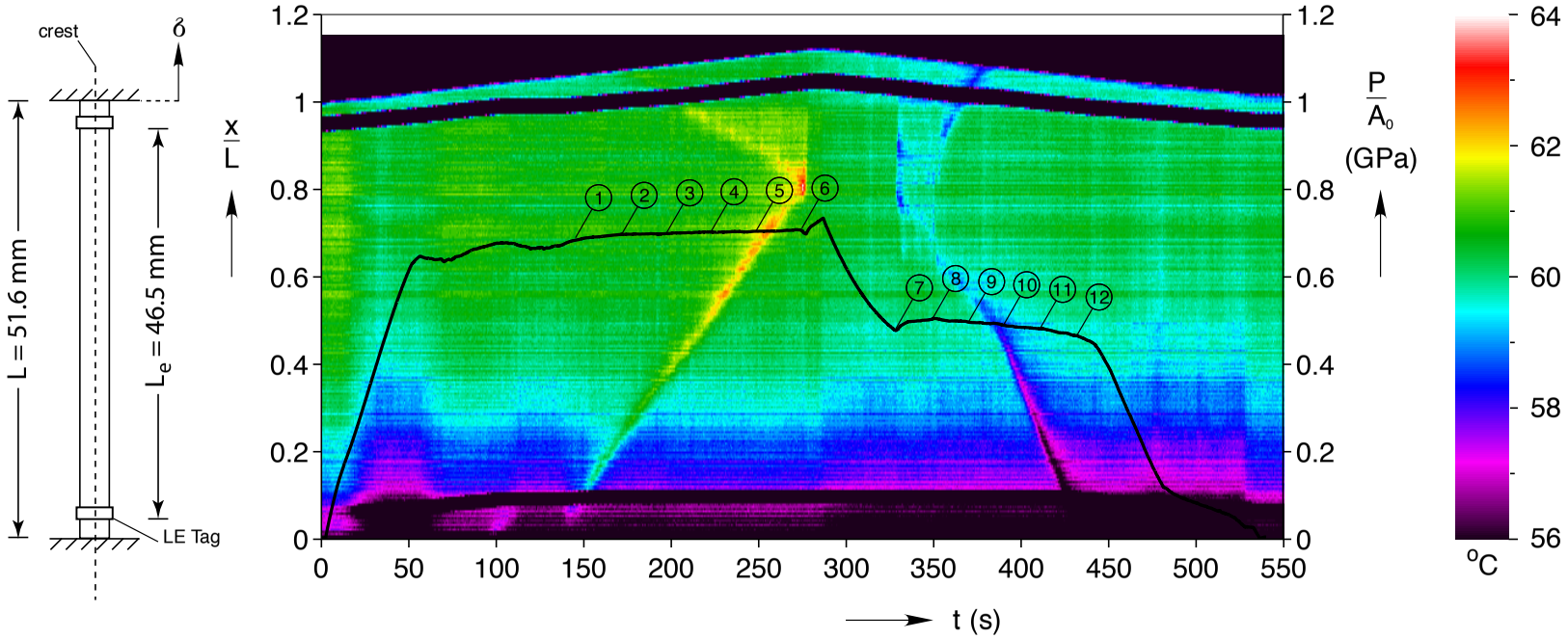


Figure 5.8: Crown temperature history of a NiTi tube with 60 °C fluid flowing through the center, captured every 0.66 s, at $\dot{\delta}/L = 5 \times 10^{-4}\ \mathrm{s}^{-1}$. Engineering stress P/A_0 is overlaid in black, and numbered circles correspond to full-field IR snapshots in Figure 5.9

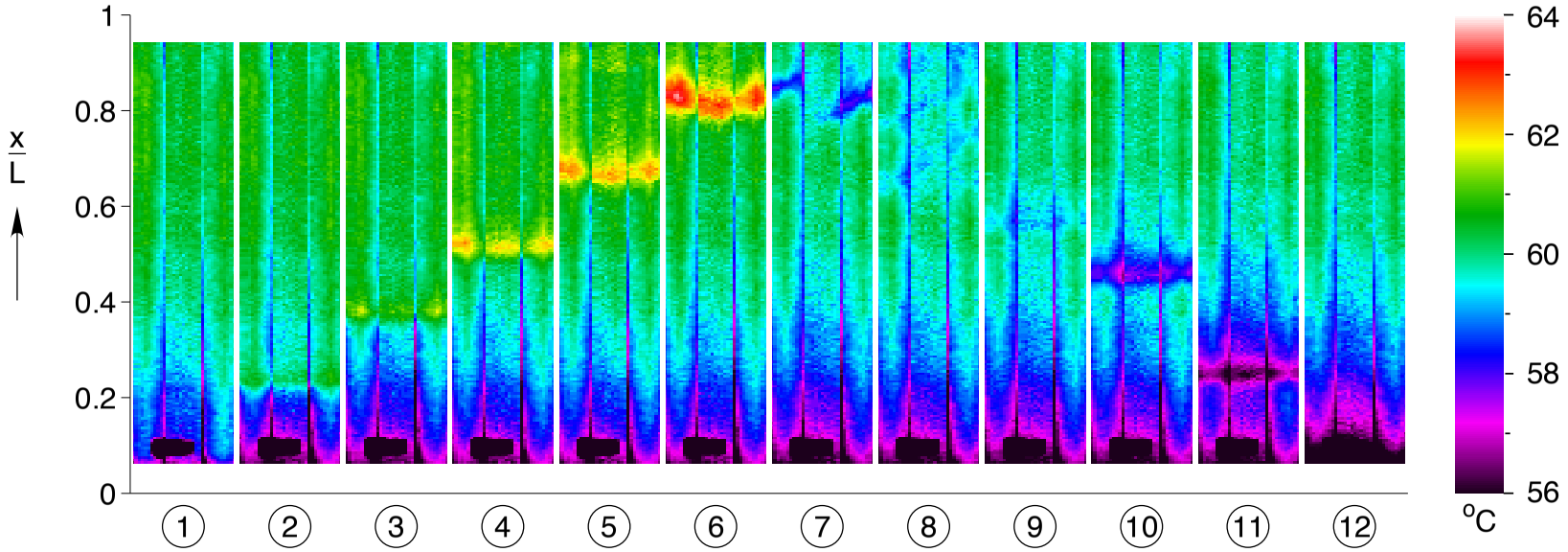


Figure 5.9: Selected infrared images of the experiment in Figure 5.8. In each frame, three separate images of the cylindrical surface have been processed to form a single rectangular surface temperature map.

is able to sustain more self-heating (and therefore a higher transformation rate) than the front in the warmer region to maintain the same axial force (engineering stress) along the tube. At (6) the two fronts coalesced, completing transformation and causing a brief rise in temperature to 64 °C as well as a slight 10 MPa drop in stress. The drop in stress is a consequence of the presumed unstable local constitutive behavior as the deformation delocalizes, explained in Figure 2.18, and the temperature rise is a superposition of the two heat sources [50].

Unloading started immediately after loading ceased, at 285 s with an initial M^+ modulus of 25 GPa. At (7) there is a nucleation of A at the same location, $x/L = 0.8$, that the loading fronts ended. However, this nucleation does not look like the ringed fronts from (3)–(6) in Figure 5.9. The infrared image for (7) in Figure 5.9 shows that a helical shaped front formed, growing from about 1 revolution in (7) to 3 revolutions in just a couple of seconds, but then stopping when it reached 3 revolutions (not shown). From shortly after (7) to (8), the helix widened until it merges with itself at (8) and then resumed as two ringed fronts. At (10), the upper front reached the top grip and the lower front doubled its velocity, in order to maintain a constant production of A as required by the constant grip velocity of $\dot{\delta}/L = 5 \times 10^{-4}$ s^{-1}. The higher velocity of the lower front also increased the self-cooling, so the lower front drops from self-cooling -1 °C locally to about -2 °C. The lower front reached the bottom grip at (12), after which the material unloaded nearly elastically.

Seven of these isothermal experiments were performed at 10 °C intervals from 10 to 70 °C. Thermal limitations of the setup (condensation on the LE tags) prevented experiments at lower temperatures that would record M reorientation. Their mechanical re-

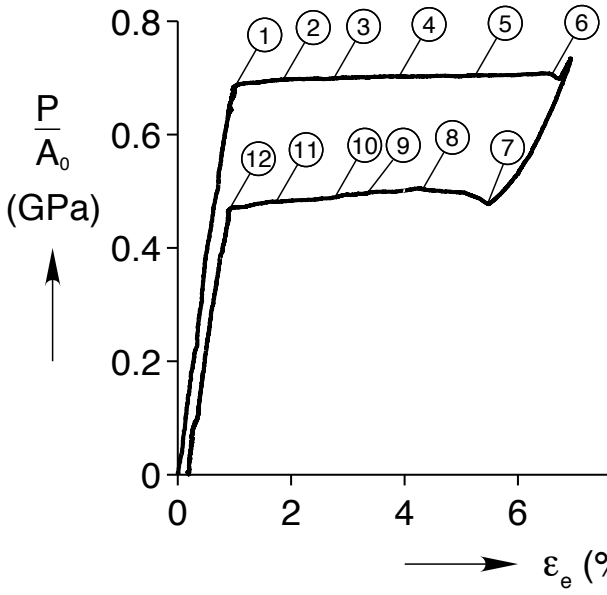
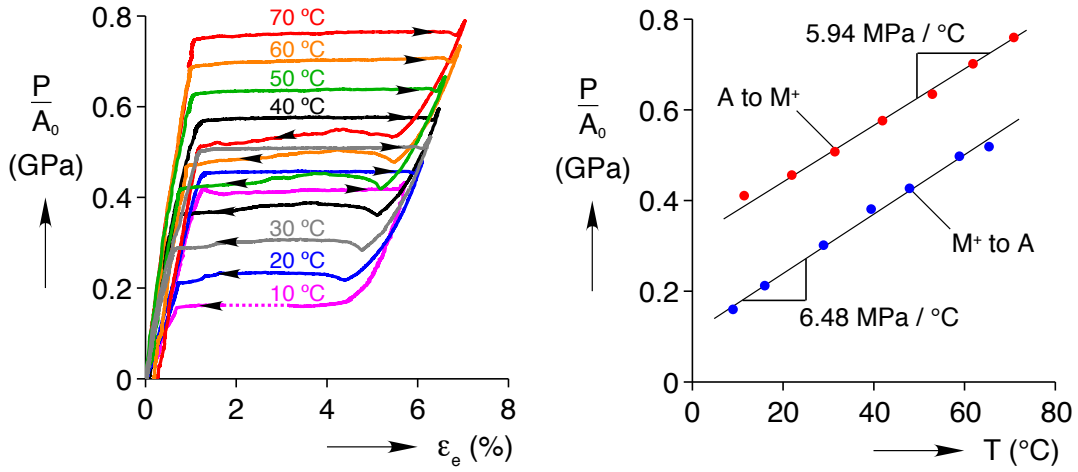


Figure 5.10: Mechanical response of the experiment in Figure 5.8. The number circles corresponding to full-field temperature snapshots in Figure 5.9.



(a) (b)

Figure 5.11: Summary of isothermal tube tension experiments, from 10-70 °C: (a) Mechanical responses, (b) propagation stresses vs. temperature, $A \rightarrow M^+$ (red) and $M^+ \rightarrow A$ (blue).

sponses are presented together in Figure 5.11(a). All of the loading stress plateaus are relatively flat, but the unloading $M^+ \rightarrow A$ plateaus vary from flat (at 20 °C) to fluctuations of 20 MPa (at 60 °C). These fluctuations are due to variations in front geometry and gradients in temperature along the length of the tube. Figure 5.11(b) shows the propagation stresses during the $A \rightarrow M^+$ and $M^+ \rightarrow A$ transformations, similar to the quasi-phase diagrams shown in Figures 2.12, and 2.15. The slopes for $A \rightarrow M^+$ and $M^+ \rightarrow A$ transformations are 5.94 and 6.48 MPa/°C, respectively. It is this relationship between transformation stress and temperature, together with the latent heat of transformation, that drives thermomechanical coupling behavior. Since (average) stress is constrained by equilibrium to be uniform along the entire tube axis, an equivalent statement would be that transformation fronts are constrained to propagate at a the same local temperature for self-similar front morphologies.

5.2.4 Loading Rate Study

Six experiments were performed at rates of $\dot{\delta}/L = 1\times10^{-4}$ s^{-1} to $\dot{\delta}/L = 4\times10^{-3}$ s^{-1} in RT air to investigate the effects of thermomechanical coupling on SMA tubes, and their combined mechanical responses from all rate experiments are shown in Figure 5.12. The lowest rate experiment $\dot{\delta}/L = 1\times10^{-4}$ s^{-1} had nearly flat loading and unloading plateaus. The next highest rate, $\dot{\delta}/L = 2\times10^{-4}$ s^{-1}, was twice as fast but the transformation plateaus were still flat. The only effect on the mechanical response was a slight increase in the stress hysteresis– a higher loading plateau and a lower unloading plateau. At rates above $\dot{\delta}/L = 8\times10^{-4}$ s^{-1} the effect was to increase the slopes of the transformation plateaus. In addition,

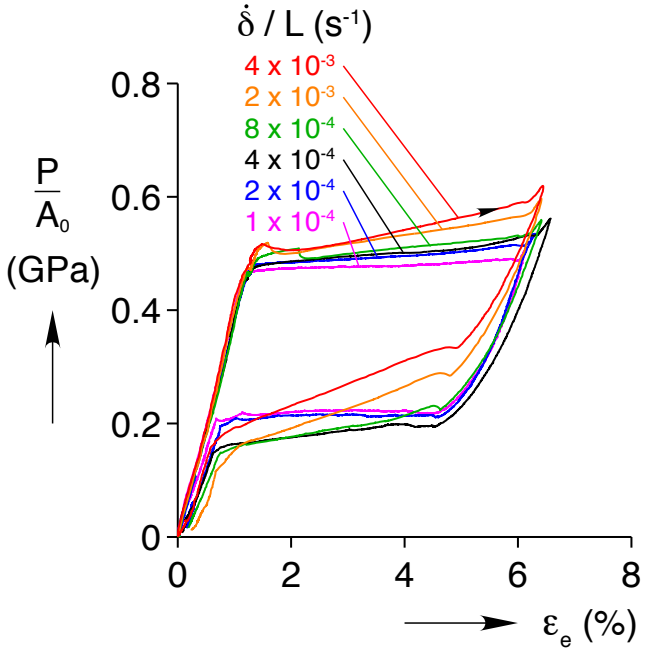


Figure 5.12: Superelastic mechanical responses of NiTi tubes at various rates

the plateau slope tended to be steeper during unloading, especially as rate increased. This was a result of the different starting temperatures for loading and unloading, since there was no pause in between.. During loading, convective cooling worked against the self-heating from $A \rightarrow M^+$ transformation, effectively lowering the slope of the mechanical response away from the adiabatic limit. However, the unloading process starts immediately after loading, when the specimen temperature is still elevated. Since convective cooling works in the same direction as self-cooling from $M^+ \rightarrow A$ transformation, it tends to increase the rate of cooling and thus the mechanical tangent modulus. Of course, the effect is absent when the rate is low and the bulk of the specimen is not significantly hotter than room temperature following loading. These features will now be explained with a detailed look at three of the experiments.

Low Rate: $\dot{\delta}/L = 1 \times 10^{-4}\ \mathbf{s}^{-1}$

The lowest rate experiment started at a more uniform temperature than the isothermal experiment in Section 5.2.3. As loading started in Figure 5.13 there was no measurable temperature gradient across the tube, which started at 21.5 °C. During loading up to 250 s, there was a slight, uniform increase in the tube temperature of about 0.5 °C. This "preheating" is likely due to a small, diffuse transformation of $A \rightarrow M^{+}$. Similar to the isothermal experiment in Figure 5.8, two ringed fronts emerged from the upper and lower grips. When they emerged they were initially quite faint, only 0.5 °C above the surrounding material, because they are near the grips which act as large heat sinks. As the two fronts moved away from the grips, the front temperature continued to increase to 24 °C by (3). By (4) the fronts were close enough that their temperature fields started to interact, and the maximum temperature rose further until the fronts coalesced at (5) at a maximum temperature of 28 °C. Note that the stress continued to increase along with the temperature. Figure 5.15 tells the same story in the mechanical response. At the beginning of the loading plateau, there was a short rise in stress as the fronts moved away from the grips and heated up. Then, from 2 to 4 % strain the stress response flattened to about 0.23 GPa as the temperature of the fronts stabilized. Above 4 % strain, the fronts approached each other and got hotter, so the stress rose again at a steeper slope of about 0.774 GPa.

Images (1) to (5) in Figure 5.14 show that both of the fronts observed during loading were of the ringed type. The front morphology on unloading is more complex than for loading. After the steep response of initial unloading, a ringed front emerged from the lower grip at 1500 s. However, unlike during loading, there was no corresponding front

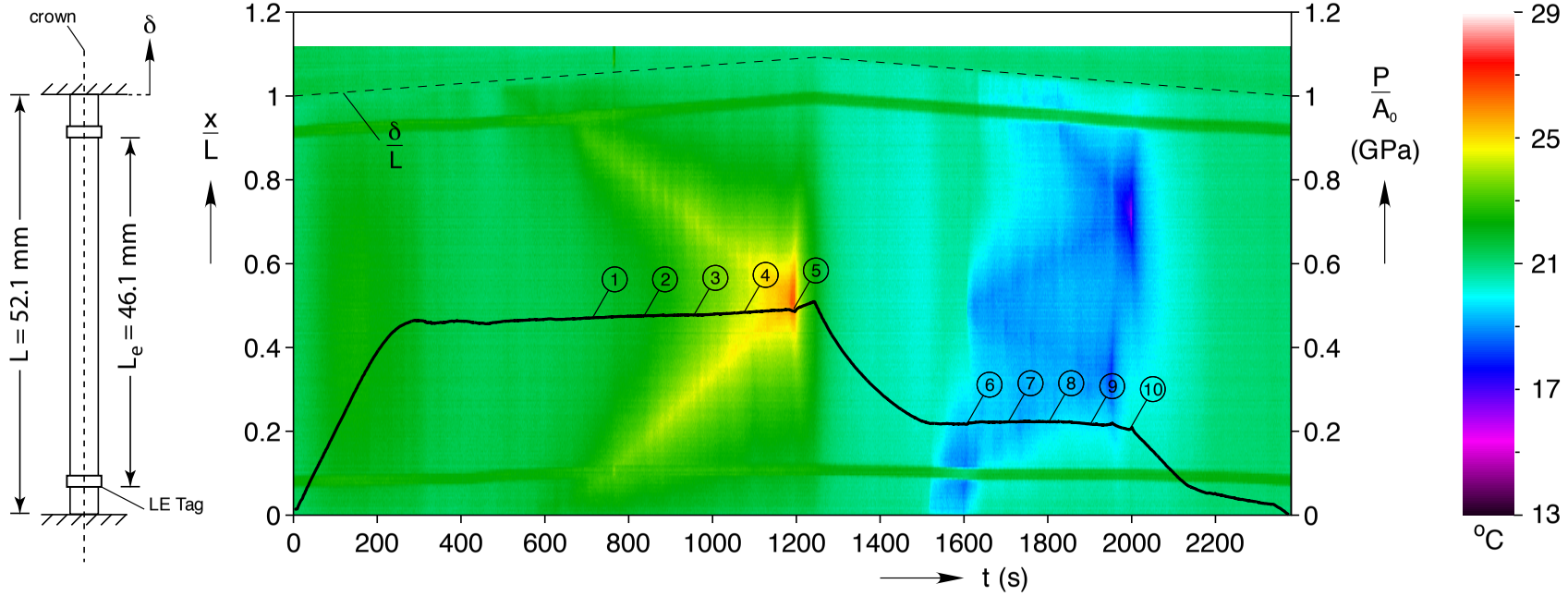


Figure 5.13: Crown temperature history of a NiTi tube in room temperature air, captured every 1 s, at $\dot{\delta}/L = 1 \times 10^{-4}$ s^{-1}. Engineering stress P/A_0 is overlaid in black, with numbered circles corresponding to full-field IR snapshots in Figure 5.14.

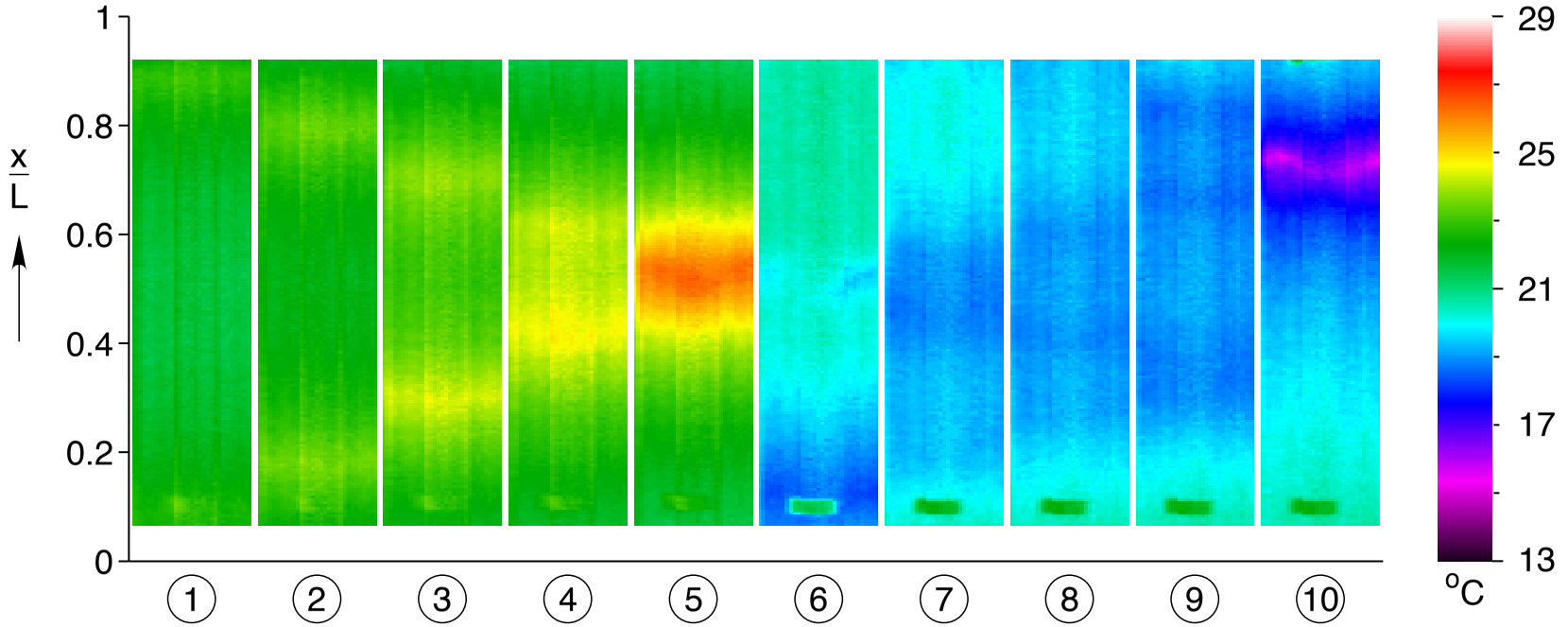


Figure 5.14: Selected infrared images of the experiment in Figure 5.13. In each frame, three separate images of the cylindrical surface have been processed to form a single rectangular surface temperature map.

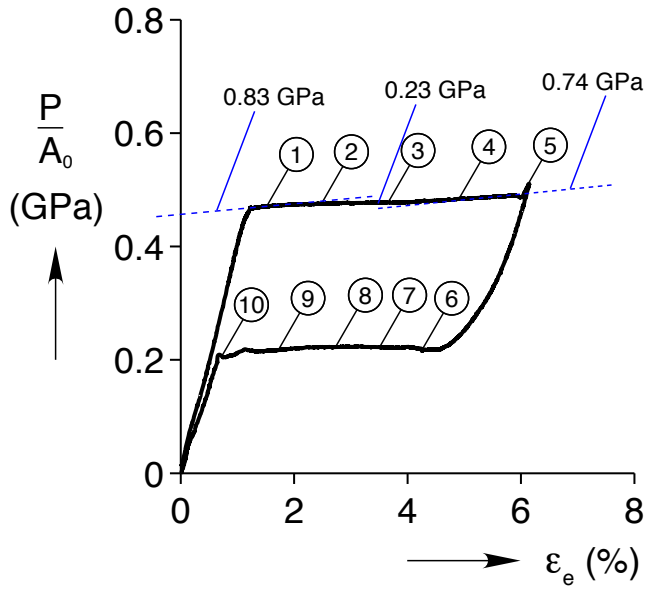


Figure 5.15: Mechanical response of the experiment in Figure 5.13. Dotted blue lines represent momentary slopes.

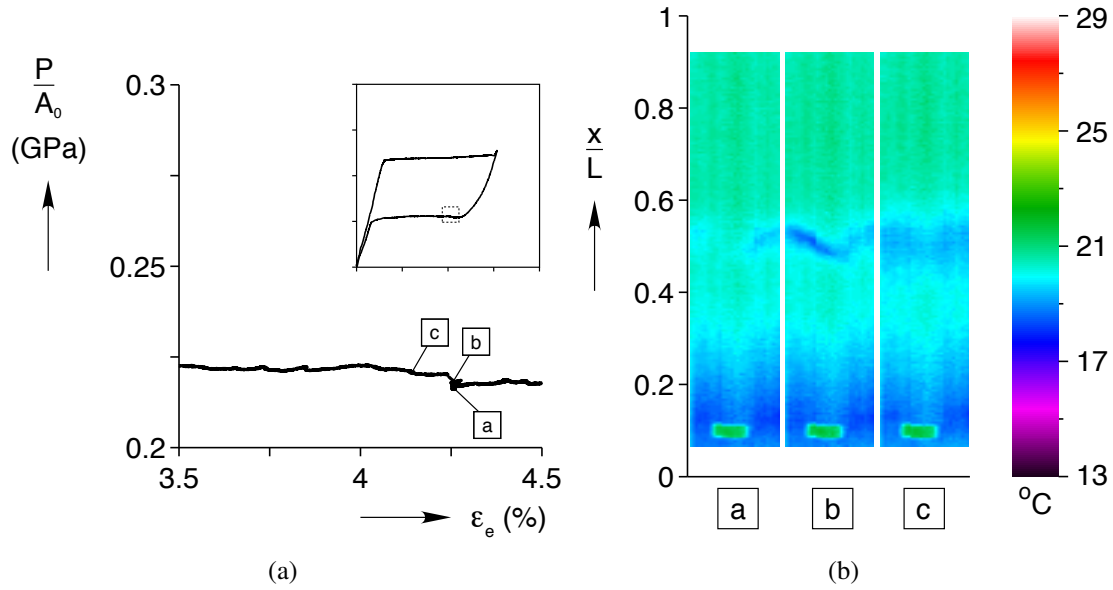


Figure 5.16: Detail view of a nucleation of A within M^+ during the experiment in Figure 5.13. (a) Mechanical response, (b) Thermal field.

coming from the upper grip. This means that the lower front propagated at twice the speed as the loading fronts, self-cooling down to 17 °C.

At (6), a new front nucleated at $x/L = 0.5$, the same place that the loading fronts coalesced. Though it looks like a ringed front in Figure 5.13, Figure 5.16 offers a closer look at the nucleation, showing a detail view of the mechanical response in Figure 5.16(a), and three select infrared snapshots in Figure 5.16(b). The image in [a] is actually the same one as (6), and shows the first frame in which the front is visible. The front forms as an inclined inclusion, not technically a helix since it only extends for about a half rotation. The inclusion grows until in [b] it takes the shape of an inclined ring. The two sides of this ring straighten out by [c], at which point they become ringed fronts and move apart.

The geometry of this nucleation is consistent with that of the final coalescence of the two $A \rightarrow M^+$ ringed fronts. Though it is faint, image (5) in Figure 5.14 shows the temperature field during the coalescence event. It has a similar oval shape, inclined to have the lowest point in the center just as in [b] in Figure 5.16. The formation of an inclined ring during merging of two ringed fronts was captured in great detail by Feng and Sun [31]. It is well known that the merging of two loading fronts creates some amount of damage in the material due to incompatibility details, possibly including some residual A, that serves as an unloading nucleation site and explains the identical pattern in [b] [48]. This idea is also supported by the only 3 MPa rise in stress during the nucleation, from [a] to [b], suggesting that the (typically larger) homogeneous nucleation barrier was assisted by pockets of untransformed material.

Shortly following the nucleation of the two ringed fronts, a fourth front emerged from

the top grip. With four fronts, the velocity of each was even lower, so the top front did not become visible until it was several tube diameters away from the heat sinking grip. The four fronts continued to propagate, each at a constant velocity and at the same temperature of about 18.5 °C until the lower two fronts met at 1980 s. The stress rose briefly as they met, in a similar but opposite phenomenon as the drop in stress at the end of loading transformation (de-localization). The two remaining fronts doubled their velocity to satisfy compatibility, so temperature (and stress) drop as a result of the increased rate of local A production. The stress rises once again when the final two fronts meet at ⑩. As they meet, they create another infrared pattern of an inclined ring, with a minimum temperature of about 15 °C. From Figure 5.15 the residual strain as measured by the laser extensometer is negligible.

Medium Rate: $\dot{\delta}/L = 8 \times 10^{-2}$ **s**$^{-1}$

The second rate experiment was performed with a grip displacement eight times faster, $\dot{\delta}/L = 8 \times 10^{-4}$ s^{-1}. Infrared images were captured every 1/3 second. Figure 5.17 shows the stress and crown temperature history, along with numbered tags corresponding to full-field snapshots in Figure 5.18. During the initial loading to 0-25 s, Figure 5.17 shows there was 3 °C of uniform heating, more than the 0.5 °C from the low rate experiment. From 25 to 75 s, there was again a period of grip slippage and transformation within the grips with no visible thermal response. One reason for this quite large amount of slippage, which is common to all the experiments presented in this section, is that the toothed wedge grips tend to bite into the specimen at an unpredictable point on the 40 mm grip face, leaving the remaining material relatively free. Note that at the end of the experiment, from 275-300 s,

there was a similar "reverse slippage" marked by the 35 MPa stress plateau, as material within the grips underwent reverse transformation.

At 80 s, two ringed fronts emerged from the grips, triggered by stress concentrations within the grips. At this moderate rate, they are visible nearly immediately, and continue to heat up as they distance themselves from the grips. This self-heating reached a maximum just before the nucleation of a new front at (1). While the ringed fronts were propagating at 31 °C, increasing the stress needed to drive transformation, the center of the tube remained at 22 °C. At (1), the stress had increased enough that it overcame the nucleation barrier at the center of the tube, so a small pocket of M^+ nucleated at $x/L = 0.5$.

The images in Figure 5.18 are on too coarse a time scale to see the nucleation of a helical inclusion, so a set of additional images are necessary. Figure 5.20(a) shows the mechanical response during nucleation, as measured by the laser extensometer and load cell. The square tags refer to full-field infrared images in Figure 5.20(c). The initial nucleation was a dynamic event, as the rapid formation of new M^+ was driven by an elastic unloading of the remainder of the tube. It was so rapid that it nucleated and grew to be more than 1 revolution long between frames (less than 1/3 s) as image [a] is the first frame in which a new front is visible. By the second frame, the helix expanded to 2 turns, but slowed its progress as the mechanical response (Figure 5.20(a)) slope fell. The helix continued to lengthen, though at a slower rate, until it was about 4 revolutions long. At that point, we assume the energy required to lengthen the helix is greater than that of the widening helix.

The total stress drop from the helix nucleation was 17.8 MPa. This is dramatically lower than the drop from a nucleation in a superelastic wire from Figure 2.21 of 77 MPa

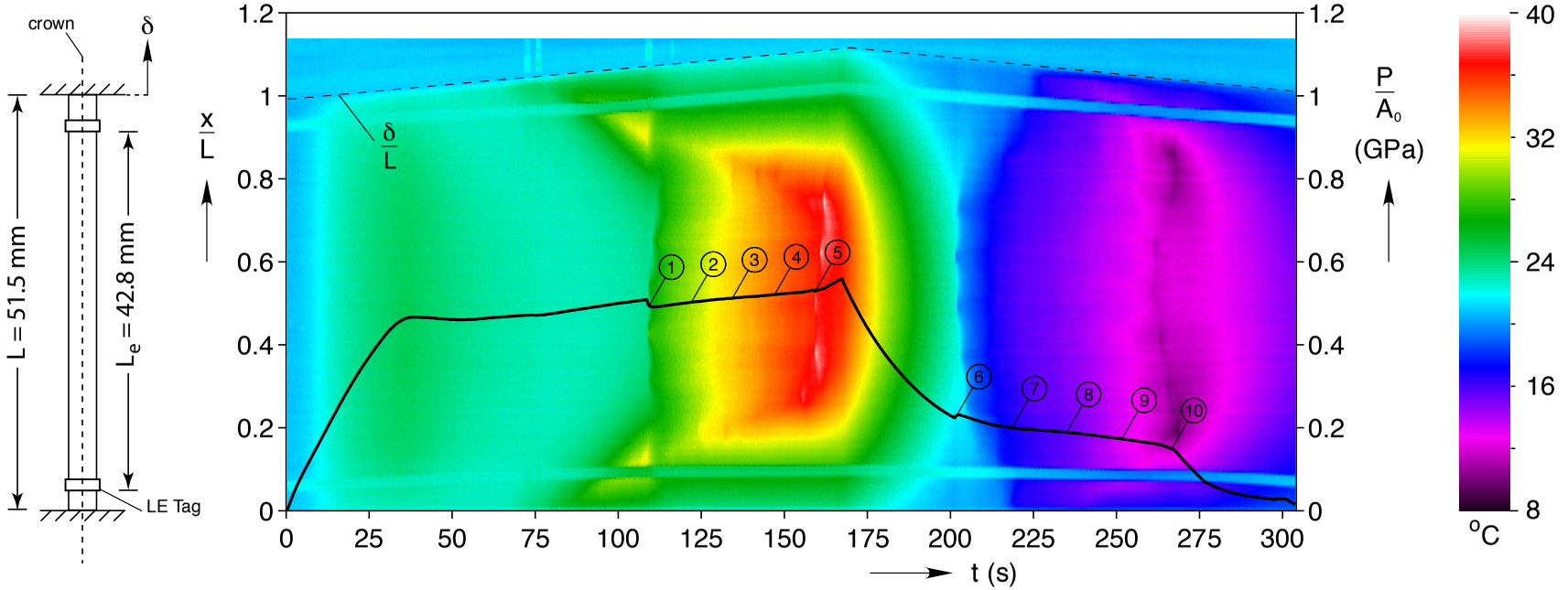


Figure 5.17: Crown temperature history of a NiTi tube in room temperature air, captured every 1/3 s, at $\dot{\delta}/L = 8\times10^{-4}$ s^{-1}. Engineering stress P/A_0 is overlaid in black, with numbered circles corresponding to full-field IR snapshots in Figure 5.18.

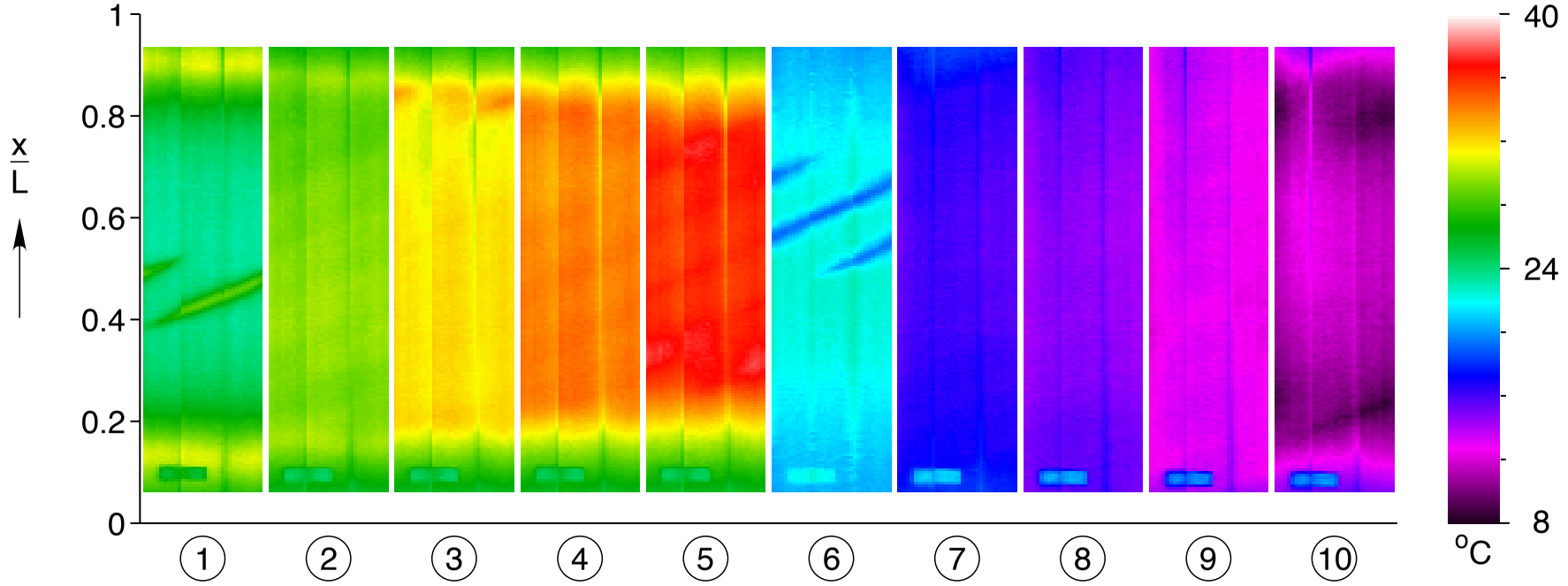


Figure 5.18: Selected infrared images of the experiment in Figure 5.17. In each frame, three separate images of the cylindrical surface have been processed to form a single rectangular surface temperature map.

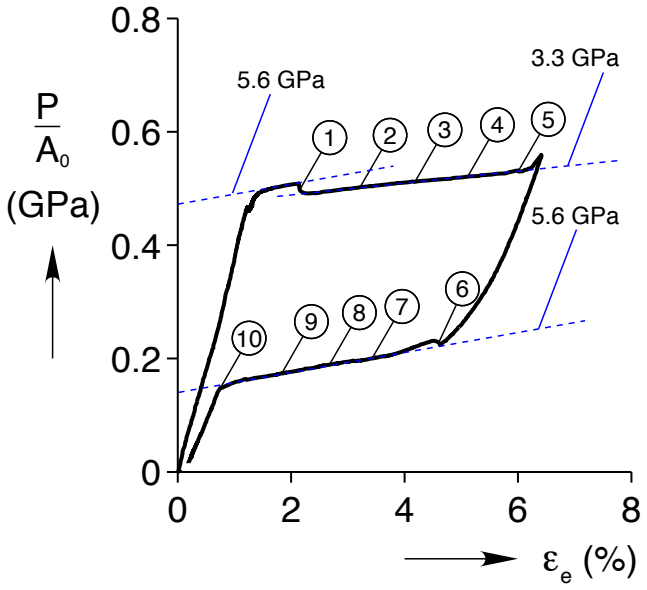


Figure 5.19: Mechanical response of the experiment in Figure 5.17.

(the tubes are actually manufactured from the same alloy, but processed differently). While 20 MPa of the reduction in the stress drop can be expected from the higher loading rate and self-heating, the remainder of the effect is likely a combination of a different material texture and the nature of the nucleation geometry (2D vs. 3D). The angle of the helix to the loading axis is difficult to measure precisely because of the low resolution of the infrared camera, but it is about 62 °, close to the angle of $\pm 54.7^{\circ}$ predicted in Section 5.1.

As soon as the helical inclusion appears, the ringed fronts effectively cease their motion, as evidenced by their cooling from [a] to [g]. Looking back again at the temperature history in Figure 5.17, the ringed fronts cool during the nucleation and lengthening of the helical inclusion at (1), but soon resume their progress, albeit at a lower temperature and propagation rate until they reach the helical front at (2) and stop. At this point, the helical geometry encompasses the remainder of the A region, and continues to widen until it coalesces with itself, in the expected helical shape, at (5). Note that once the helix is

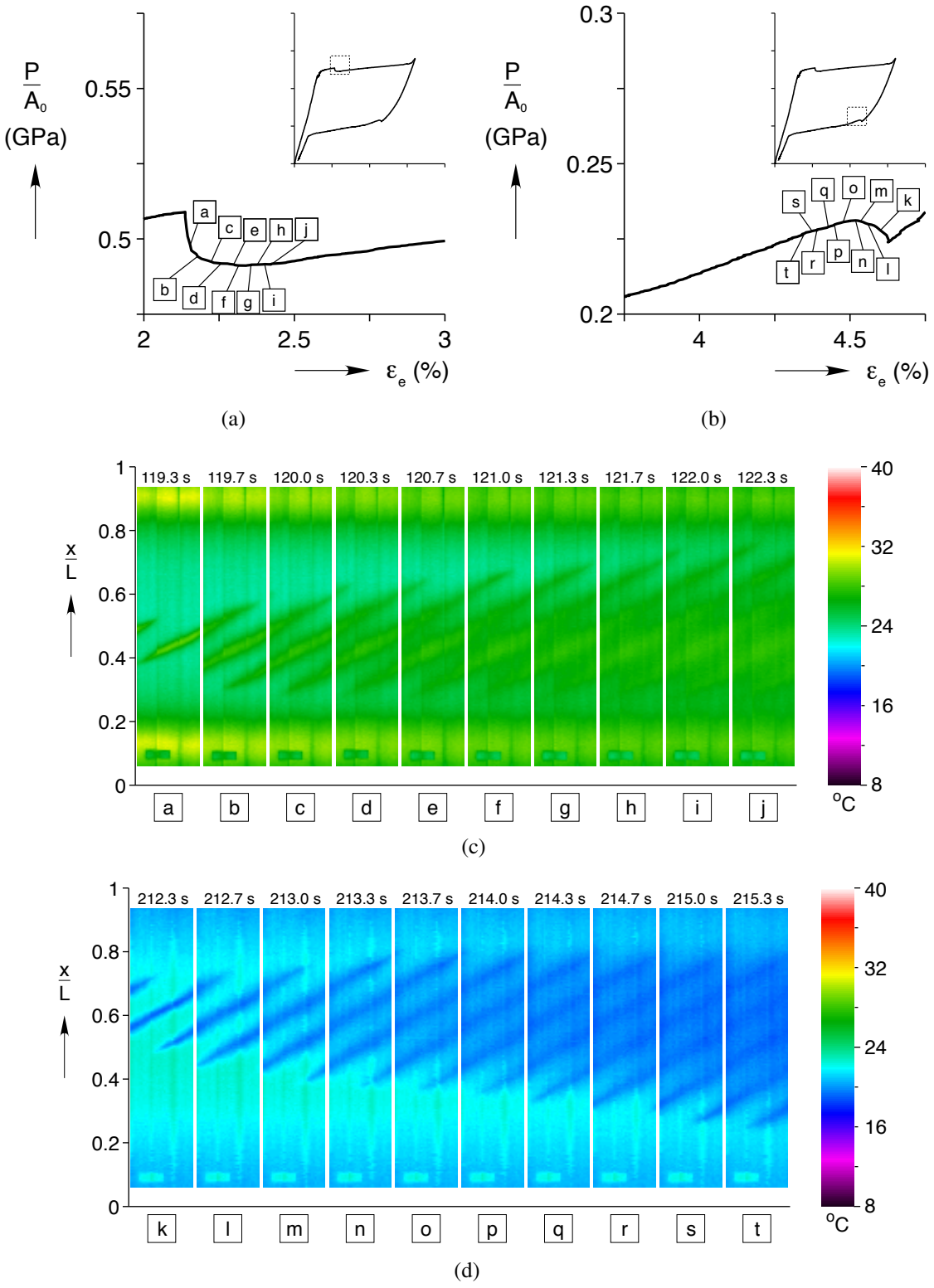


Figure 5.20: Detail view of nucleations of M^+ (loading) and A (unloading) in the experiment from Figure 5.17. The detail mechanical response in (a) and (b) contain square tags referring to the infrared images in (c) and (d)

formed, heating is nearly uniform in images (2) to (4). The uniform temperature field acts as a barrier to any additional nucleations, as they would require a colder region to repeat the effect at (1).

This is a good point to discuss a theoretical issue with the nucleation of a helical inclusion. While the front in [2] in Figure 5.20(c) propagates several turns, the question arises: what is the minimum length of a new helical nucleation? A remarkably simple argument is made in Figure 5.21 that the periodic nature of the tube surface means helix must span at least 2 full revolutions. Assuming the simplest geometry- a diamond shape, consider a new nucleation of M^+ (blue) in the diagram on the left. The new M^+ has a much higher strain, about 6 %, compared to the surround A (red). If we add up the amount of M^+ along each circumferential position, indicated by the vertical dotted lines, the pattern is a triangle centered at the nucleation site. This distribution of high-strain M^+ is not sustainable, however, because the uneven distribution will induce a significant bending strain on the tube, putting the side of the tube (where the M^+ distribution is large) under a lower stress and the opposite side (small M^+) under a higher tensile stress. This distribution would tend to drive transformation towards the higher stress region, leading to the shape in the center diagram. Even with the larger front, the distribution of M^+ on the bottom row is still nonuniform. The shortest front which can result in a uniform M along each vertical "slice" of material is one that is two turns long, shown in the diagram on the right side of Figure 5.21. How far the initial helix extends beyond this minimum of two turns (in the isothermal case) is a function of the two length scales of the tube, L/R and R/t [40].

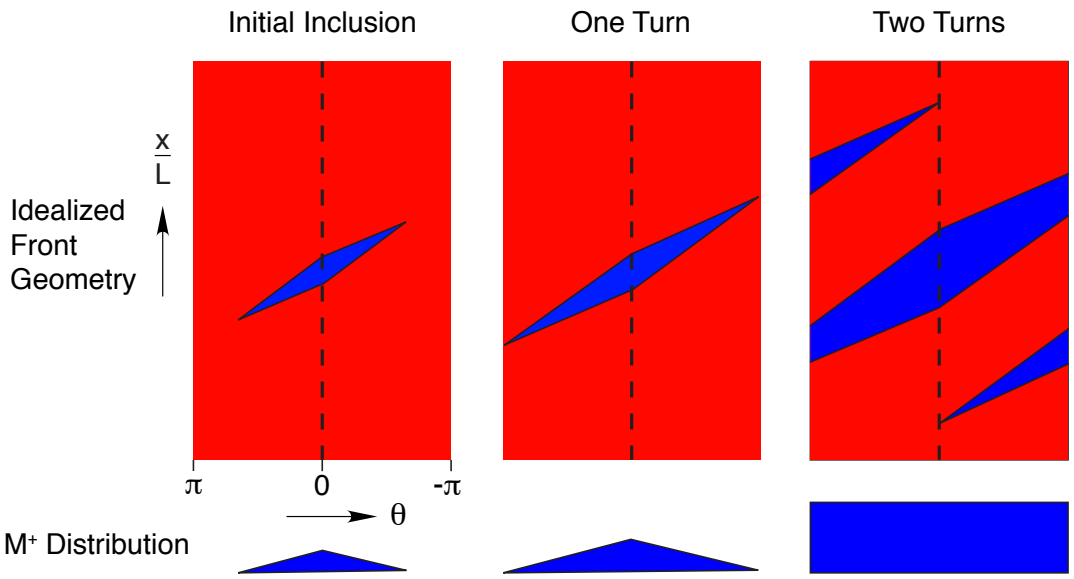


Figure 5.21: Idealized diagram of an M^+ (blue) nucleation within A (red) in an unwrapped tube. A minimum of two of helix rotations are necessary to achieve a uniform distribution of M^+ along each vertical slice, to satisfy axial strain compatibility.

The unloading front morphology is simpler than during loading. After a nearly elastic unloading from 170-200 s, there was a nucleation of A at $x/L = 0.6$. Unlike the low rate experiment, the loading $A \rightarrow M^+$ fronts coalesced in the form of a helix, so the region of favorable A nucleation follows that same helix, running most of the length of the tube. Figure 5.20(b) shows the mechanical response near the nucleation event. The stress rose from k to n as the helix grew rapidly. Looking at the temperature field in Figure 5.20(d), the helix length stops increasing as the stress starts to drop again, and self-cooling lowers the temperature further. The total stress rise during this nucleation was 7.2 MPa. Though Figure 5.20(d) does not show it, the helix continues to grow until it reaches the upper and lower grips by image (7) in Figure 5.18. Just like the nearly uniform heating during loading, the full helix provided nearly uniform cooling from (7) to (9). Between (9) and (10), coalescing fronts caused another helical "ripple" of temperatures as low as 8 °C, starting in

the center where the helix is thickest, and progressing out to the ends. After (10), the stress decrease was still non-linear because of transformation within the grips, but Figure 5.19 shows a linear response within the gauge length is linear.

High Rate: $\dot{\delta}/L = 4 \times 10^{-3}$ **s**$^{-1}$

At the highest rate, self-heating in the tube approaches the adiabatic limit. The smaller time scale reduces effects of conduction and convection. At the same time, it increases the amount of self-heating, making it easier to interpret infrared images to infer front morphology. Figure 5.22 shows the stress history overlaid on the crown infrared history. The phemomenon of heating during initial loading continues. This time, the temperature rises 5 °C in the first ten seconds. There is also (again) a period of action outside the LE tags from 8 to 19 s. Two ringed fronts just barely emerge from the upper and lower grips at $t = 17$s, but at this high rate they are 15 °C above the ambient temperature of 21 °C within 1 s of appearing. At that temperature, a new front must nucleate in the cooler center region of the tube, since $15\ ^{o}\text{C} \times 6\text{MPA}/^{o}\text{C} = 90\text{MPa}$ is sufficient to overcome the nucleation barrier[a]

The nucleation of the new helical inclusion follows in a similar manner to the front in Figure 5.20(c). Because the infrared camera captured images at a higher rate of 20 frames/sec, it was able to capture the dynamic formation of the inclusion with greater detail. The helix first appears at $x/L = 0.25$ and elongates at 63 ^{o}to the loading axis for the remainder of the tube. Figure 5.25(a) shows the mechanical response, centered around

[a]The nucleation barrier has not been quantified for these tube specimens, but the stress drop in Figure 5.19 gives us a lower bound of 17.8 MPa, and the barrier for the SE wire in Chapter II is 77 MPa.

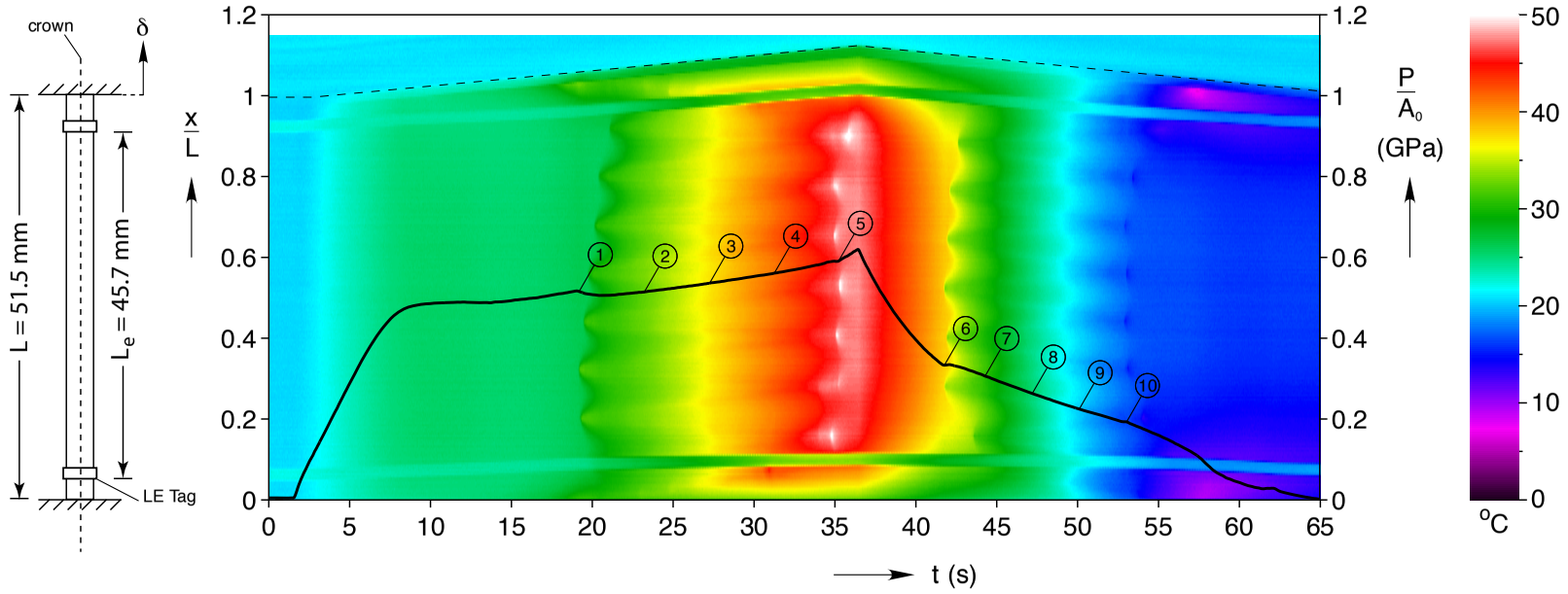


Figure 5.22: Crown temperature history of a NiTi tube in room temperature air, captured every 1/3 s, at $\dot{\delta}/L = 8\times10^{-4}$ s^{-1}. Engineering stress P/A_0 is overlaid in black, with numbered circles corresponding to full-field IR snapshots in Figure 5.22.

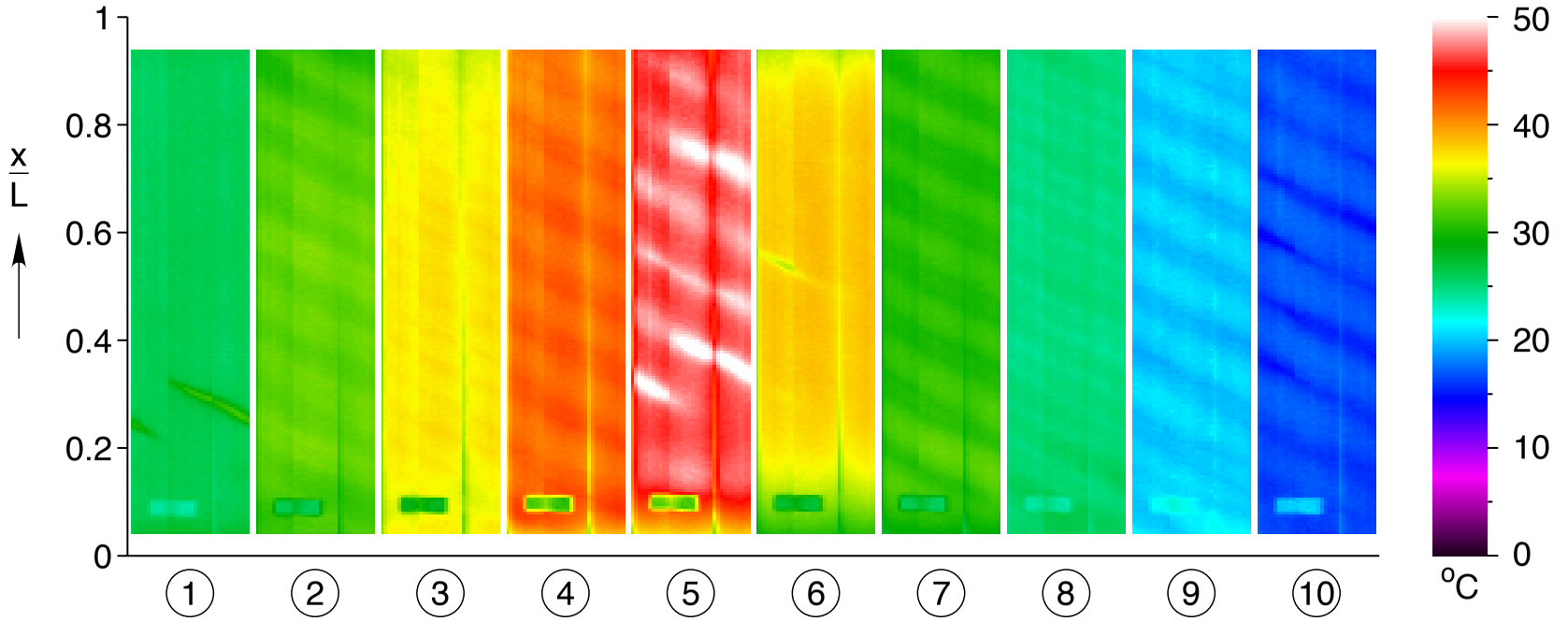


Figure 5.23: Selected infrared images of the experiment in Figure 5.22. In each frame, three separate images of the cylindrical surface have been processed to form a single rectangular surface temperature map.

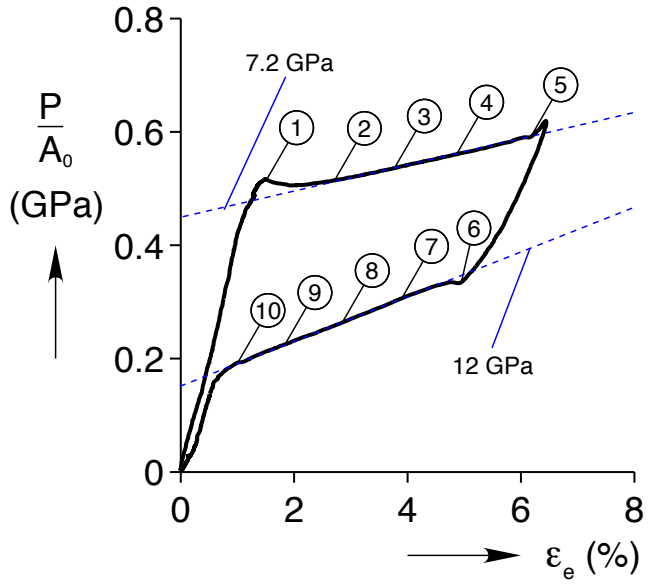


Figure 5.24: Mechanical response of the experiment in Figure 5.22.

the $A \rightarrow M^+$ stress drop of 11.5 MPa. The drop here is probably lower than the medium rate experiment (17.8 MPa) because continued self-heating never let the tube reach the fundamental isothermal response. Figure 5.25(c) shows the formation of the helix as it expands along most of the tube axis. Similar to the helical $A \rightarrow M^+$ front in the medium rate experiment (Figure 5.20), there seems to be an initial, rapid helix formation that lasts about 0.25 s, followed by a slower elongation of the helix that takes between 1 and 2 seconds longer. This is consistent with the mechanical response, where the negative slope is steeper from [a] to [e] but then more shallow for the remainder, [f] to [m].

Looking again at Figure 5.22, the helix extends the full length of the tube by $t = 22.5$s, after which its only available path is to widen. Because the rate is higher, conduction effects are not able to create as uniform temperature field as the medium rate experiment, so a helical band of self-heating is clearly visible in images (2) to (4) in Figure 5.23. When the two helical fronts coalesce, a helical band of high temperature, in excess of 50 °C (50 °C

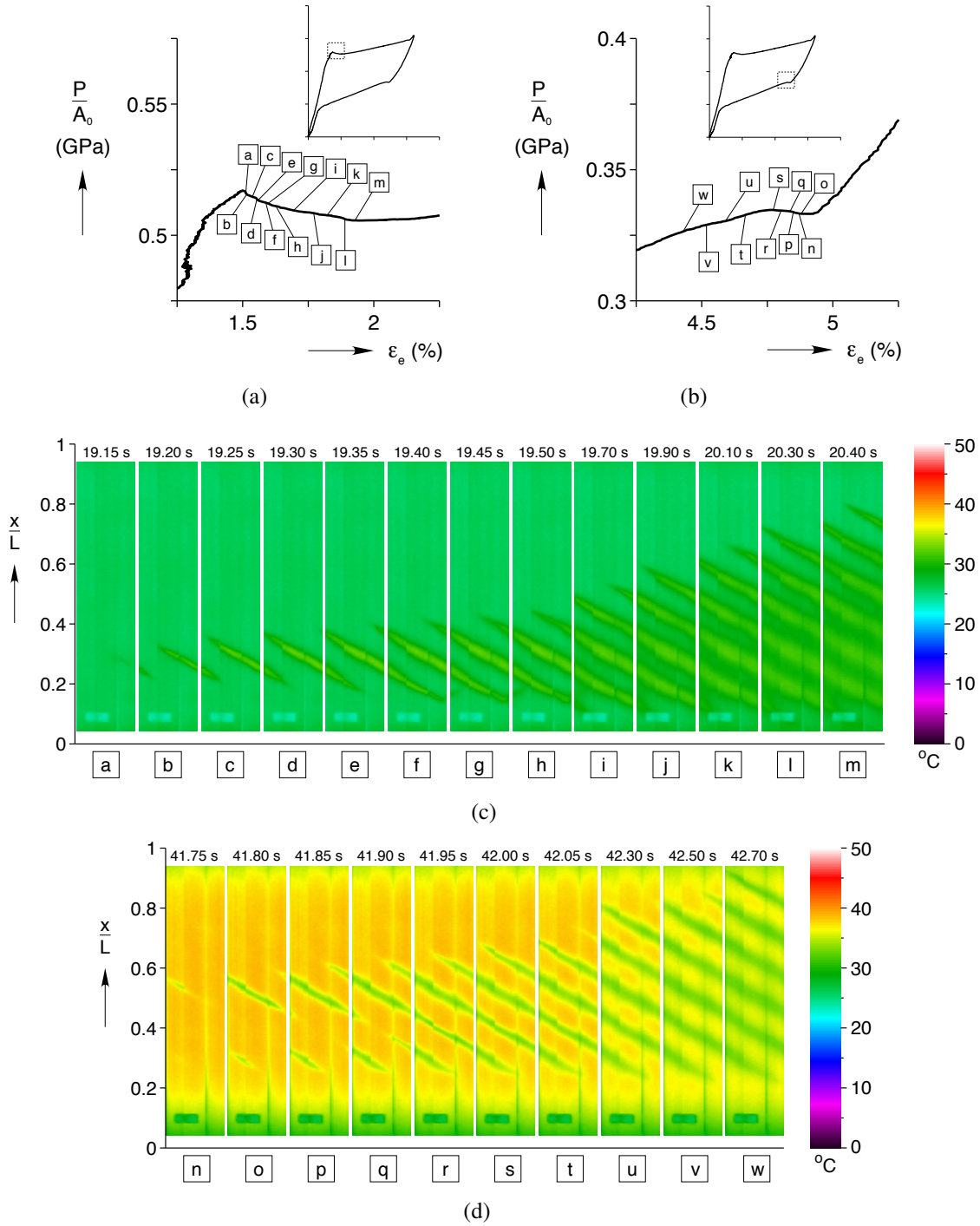


Figure 5.25: Detail view of nucleations of (a) M^+ (loading) and (b) A (unloading) in the experiment from Figure 5.22. The detail mechanical response in (a) and (b) contain square tags referring to the infrared images in (c) and (d)

was the upper limit of the selected range for the camera) is visible in image (5). The stress reaches a maximum of 592 MPa during this event, then continues to rise at a steeper slope as the superelastic plateau is exhausted.

After nearly elastic unloading of the M^+ phase between 36 and 41 s, there is a nucleation of A at (6). Images [o] to [q] in Figure 5.25(d) show that there are actually two nearly simultaneous nucleations of A, although by [r] they have joined to create a single, unbroken helix. Though two separate nucleations occurred, it is reasonable to expect them to eventually align because they both formed on the single helical material line (damage) created by the coalescence of $A \rightarrow M^+$ fronts. The rise in stress from A nucleation is only 1.5 MPa. The remainder of the $M^+ \rightarrow A$ transformation proceeds in similar fashion to loading, a wavy temperature field with a faint (colder) helical band representing the converging helical fronts. Transformation completes within the LE tags at (10) with a minimum temperature of 12 °C. Subsequenty, the load drops steeply, but at 57 s there are two additional "cold spots" of 8 °C near the top and bottom, as fronts from the grips are finally annihilated at macroscopic stresses less than 100 MPa.

Initial Nucleation and Growth of Helical M^+ Inclusions

Some additional conclusions can be drawn from an even more careful look at the initial nucleation and propagation of the two M^+ helices shown in Figures 5.20(c) and 5.25(c). Fig. 5.26 compares the front position from each helix. For each frame, the helix tip location was manually determined by carefully selecting the pixel at the tip of the temperature field. The length s_t is defined in Figure 5.27 as the axis of a 1D coordinate system oriented along a

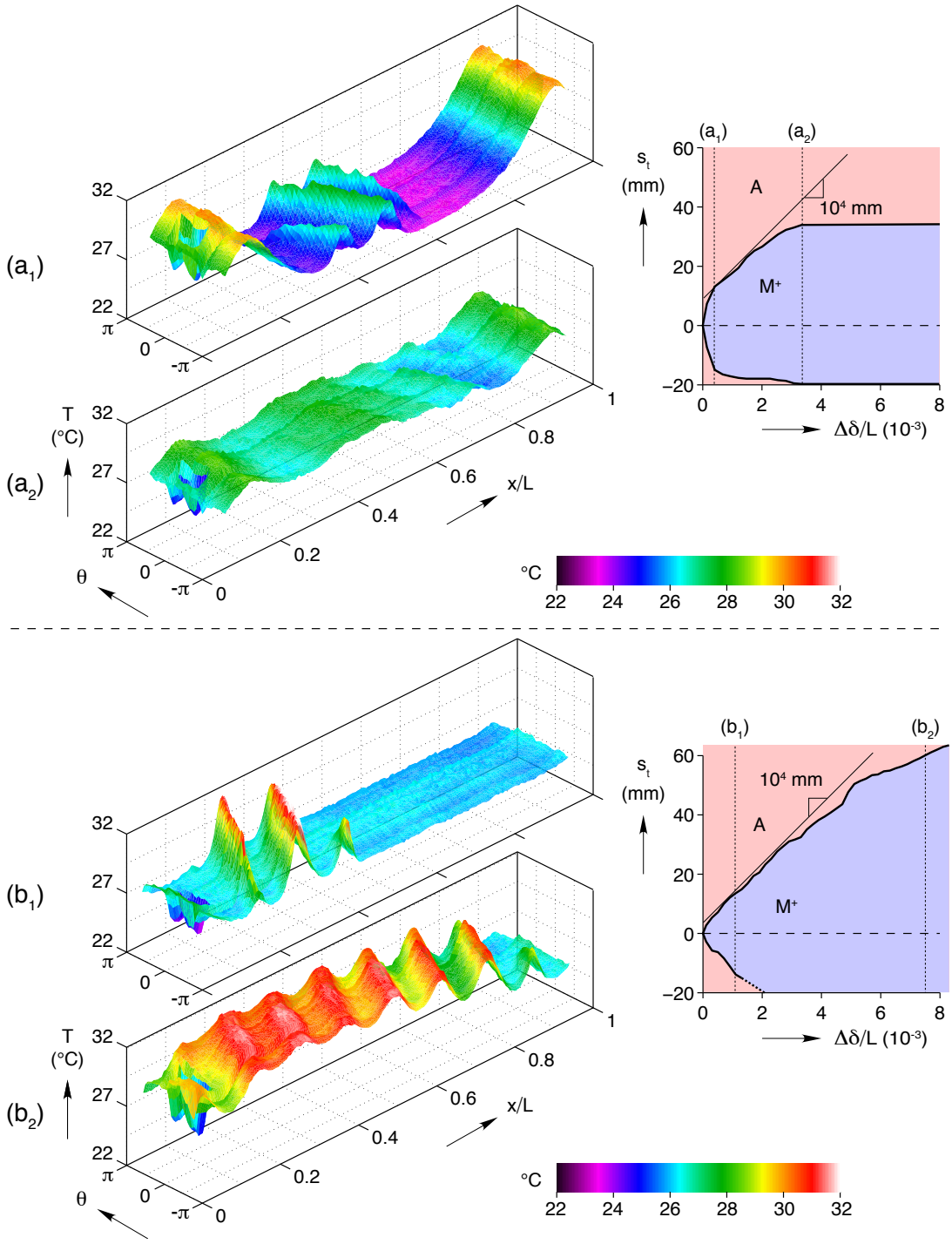


Figure 5.26: Nucleation of two helical martensite inclusions: At the top are are temperature fields ((a$_1$) and (a$_2$)) from the experiment at $\dot{\delta}/L = 8 \times 10^{-4}$ s^{-1} with a plot of the material phase along the helical axis s_t described in Figure 5.27. The independent axis is the normalized grip displacement, offset from the point of nucleation: $\Delta\delta/L = (\delta - \delta_{nuc})/L$. Dark lines represent the position of the 2 tips of the helix. At the bottom are two temperature fields ((b$_1$) and (b$_2$)) from the experiment at $\dot{\delta}/L = 4 \times 10^{-3}$ s^{-1} with the corresponding plot of helix tip position.

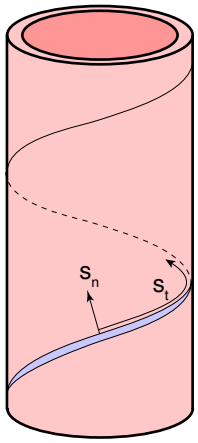


Figure 5.27: Helical coordinate system: s_n is normal to the helical inclusion, s_t is tangent to the inclusion. The origin is at the initial nucleation site.

helical path coincident with the inclusion. On the right side of Fig. 5.26, the phase fraction along that helical coordinate system is plotted against the normalized grip displacement $\bar{\delta}/L$. Each tip of the inclusion is marked by a bold line, the phase of the material between the tips is assumed to be M^+ (blue region) and the material ahead of the tip is assumed A (red region). Recall that the plot at the top right is at a strain rate of $\dot{\delta}/L = 8 \times 10^{-4}\ \mathrm{s}^{-1}$ while the plot at the bottom right is five times faster, at $\dot{\delta}/L = 4 \times 10^{-3}\ \mathrm{s}^{-1}$. At the left are four surfaces, two from each experiment, of the temperature field at select frames. Temperature is indicated by color, ranging from 22-32 °C, and the Z-axis position. The position of each temperature plot is marked by a vertical dashed line and its corresponding label (a_1)-(b_2) in the plots on the right.

For the lower rate experiment, the surfaces in (a_1) and (a_2) illustrate the effect of the two ringed fronts which originated from the grips (see Figure 5.17 to see how they start). Their high temperature essentially creates a thermal wall which the helical inclusion cannot "climb". When the lower tip of the front in (a_1) is about to hit that wall, the position of that

tip in the plot to the right stops moving soon after. The upper tip in (a_1) continues at an approximate velocity of 10^4 mm, slowing more as it gets closer to the wall. Surface (a_2) shows the temperature field as the upper front hits the high temperature region created by the upper ringed front. The nucleation of the helical inclusion has slowed M^+ production at the ringed front, so the "height" of the wall has actually fallen from 30 °C at a_1 to 28.5 °C at a_2, but it is still tall enough to stop the helical inclusion before it reaches the ringed front, assumed to be at the center of the wall.

In the higher rate experiment, there are no ringed fronts to deter helical front propagation, so the tip is free to propagate until it reaches the grips. The lower tip could not be tracked beyond a position of -15 mm because the laser tag (the low-temperature depression in (b_1) and (b_2)) masks its movement so it is assumed to have continued at a constant velocity (dashed line). The upper tip continues at an approximate velocity of 10^4 mm until it moves out of the IR field of view.

The identical tip velocities between the two experiments is likely not an coincidence, as both fronts experience similar thermal conditions. Though the rate in surfaces (b_1) and (b_2) is five times that of (a_1) and (a_2), they are still both fast enough to cause significant self-heating. In the center of the helix, self-heating in both fronts is more than 5 °C above the nominal temperature, while at the very tip the temperature remains low, since in both cases helix tip propagates faster than the heat it generates. For example, even after several turns, surface (b_2) shows the upper tip of the helix still propagating in a relatively uniform 26.5 °C field while the center of the helix remains above 31 °C. The same situation exists in (a_1), as the front tip moves through 23 °C material while the center is 28 °C. If front

propagation is limited to the tip of the helix, due to the lower temperature, it is logical that the rate will be proportional to the grip deflection, which dictates the global M^+ rate of production. As the temperature difference between helix tip and center drops, either from a barrier like (a_2) or from conduction, the tip velocity slows.

Discussion

Based on the observations presented in this section, the following conclusions can be made about the effects of loading rate and thermomechanical coupling on NiTi tubes in uniaxial tension.

- During initial loading, there is a small amount of pre-heating due to a small precursor, uniform transformation throughout the tube associated with a slight initial nonlinearity in the mechanical response. The magnitude of the temperature rise depends on the loading rate, and at the highest rate the maximum rise was 5 °C.

- The preferred isothermal front geometry, excluding nucleation events (localizations), is the ringed front, consisting of a number of small inclined "branches" repeating in a ring. It has a lower strain energy penalty (from thickness incompatibilities at the interface) due to its relatively short interface length (one circumference). When two ringed fronts meet, they form a pair of inclined rings just before they annihilate.

- The preferred geometry at nucleation is a tapered helical inclusion, which forms at an approximately 62 $^{\mathrm{o}}$ angle to the loading axis and wraps around the tube minimum of 2 revolutions. It is associated with a sudden release/absorption of latent heat and

a load drop/rise as the inclusion develops, resulting in elastic unloading/loading of the remaining material. The helical inclusion has a larger energy penalty due to its relatively large interface length, which at the minimum of 2 revolutions and including both sides of the inclusion, is at about 4 times longer than the ringed front interface.

- Under a high loading rate, the helical geometry is favored because it results in nearly uniform heating/cooling, resulting in a lower maximum temperature and thus a lower macroscopic stress P/A_0. In contrast, the ringed front causes highly localized hot (or cold) regions, resulting in greater temperature (and stress) extremes. The relationship between the two is essentially a contest between the thermally favorable helical front and the (strain) energetically favorable ringed front.

- Increased loading rate results in increased self-heating, which in turn increases the stress during transformation. Increased rate when the same number of ringed fronts are present will result in a constant offset of the stress during transformation, in effect increasing the stress hysteresis. Increased rate with a largely helical front geometry causes an increase the stress-strain slope of the transformation plateau, since steady-state (self-similar) propagation is never reached..

- The coalescence of two $A \rightarrow M^+$ fronts creates a sudden increase in temperature (or a temperature drop during unloading), as the local production of M^+ increases by about twofold[b]. The event tends to create an inhomogeneity in the shape of the coalescence, likely in the form of untransformed A or complex residual stresses. This damage/residual serves as a favorable point for nucleation of A upon unloading.

[b] see a 1-D analysis of converging point heat surces in Figure 7 of Iadicola, Shaw [50]

5.3 Compression Experiments

A significant advantage of tubular structures over wires or strips is that, due to their high bending stiffness, they can sustain much higher compressive stress without buckling. The stress required for superelastic (high temperature) $A \rightarrow M^+$ transformation can require stresses of up to 1.5 GPa, making it especially difficult to suppress the flexural buckling mode with thin wires and strips. The following isothermal characterization and rate study of NiTi tubes in compression highlights several important differences from the tension study of Section 5.2.3.

5.3.1 Compression Experimental Setup

Through trial and error, it was found that for the tube specimens used in this chapter (3.175 mm x 2.54 mm) would not buckle provided there were clamped boundary conditions on a free length of 12 mm. However, a 12 mm specimen was undesirable because the specimen aspect ratio at that length was only $L/D = 3.8$. To solve this problem, a specialized fixture was made which would allow a relatively long specimen but suppress the buckling mode over most of its length. The compression fixture used fluid flow through the center of the tube to maintain a prescribed temperature, similar to the method used for isothermal tensile isothermal experiments.

Figure 5.28 shows the compression setup. Each stainless steel grip was rigidly attached to a servo-mechanical load frame, which provided load and displacement measurements. All experiments were performed using displacement control. The grip displacement u is normalized by the total tube length L (not the free length L_f) for engineering strain

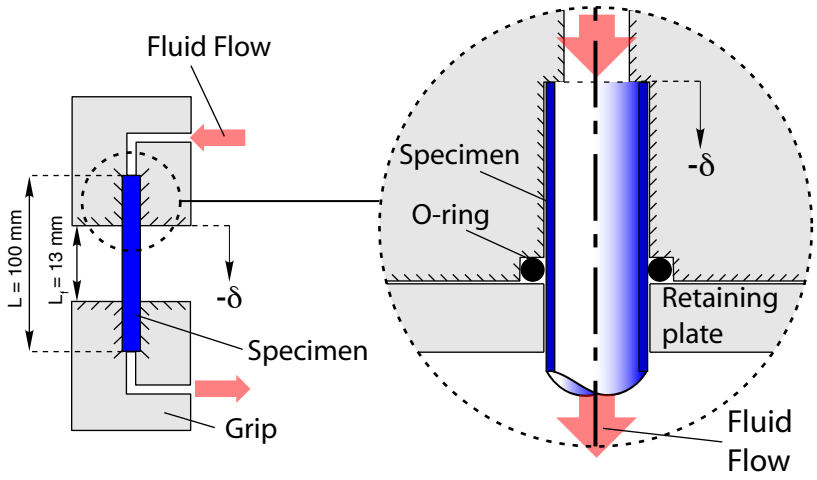


Figure 5.28: Cross-section of tube compression fixture (not to scale), with detail view on the right. The stainless steel grips suppress buckling and pass fluid through the tube center, but still allow the tube to slide freely.

ε.At the end of each grip, two holes were drilled, one 3.2 mm wide and 38 mm deep, and the second 2.5 mm wide and 50 mm deep. The "shelf" between these two holes formed a surface on which the specimen could be compressed, while the smaller hole delivered fluid for temperature control. The large hole was drilled 25μm larger than the tube diameter to provide minimal clearance (yet lateral support) and then coated with a light layer of silicone grease to reduce friction. An O-ring was positioned in a depression near the mouth of the grip and lightly held in place by a retaining plate, serving to seal the fluid circuit. Each 100 mm specimen was carefully ground at the ends to be as uniform as possible, then painted in the free section in the same manner as the tensile specimens in Figure 5.7(b).

5.3.2 Isothermal Characterization

Three isothermal experiments were performed at temperatures of 20, 40, and 60 °C. Preparation for each experiment involved running fluid through the circuit for about 30

minutes, enough for the rather massive grips to reach thermal equilibrium. All three experiments were performed under displacement control at a rate of $\dot{\delta}/L = 2.6 \times 10^{-4}$ s^{-1}, moving down ($-\delta$ direction) first.

All three isothermal mechanical responses are shown in Figure 5.29. To highlight the fact that the experiments are performed in compression, the third quadrant of axial stress-strain space is shown with increasing stress upward and tensile strain still to the right. It is immediately obvious that there is considerable tension-compression asymmetry in the response (note the large scale of 0 to 1.6 GPa). Looking just at the 20 °C experiment, there was an initial (mostly) elastic loading to a stress of about -700 MPa. The modulus was initially shallow 45 GPa, possibly due to settling of the tube ends against the "shelfs" but grew to a steeper 67.5 GPa. At -700 MPa there was a gradual knee in the curve as the $A \rightarrow M^-$ transformation started. Halfway across the plateau, a load frame error caused a momentary reversal in the crosshead velocity, but reloading was resumed resulting in the inner loop shown in Figure 5.29. The outer envelope of the response was then reached as transformation continued. The plateau has a significant slope of about 8 GPa and extended to a strain of about 3.25 % before the slope started to increase again. At this point the crosshead motion reversed and the tube unloaded with a M^- modulus of 67 GPa. This modulus is nearly three times the M^+ unloading modulus measured in tension with Figure 5.10, 25 GPa. The steep unloading response ended with another soft knee in the response as $M^- \rightarrow A$ transformation began, returning along another inclined plateau of the same slope, about 8 GPa. Transformation was exhausted with another smooth transition to unloading, starting at about -389 MPa. The final unloading had 0.1 % residual strain. The two experiments

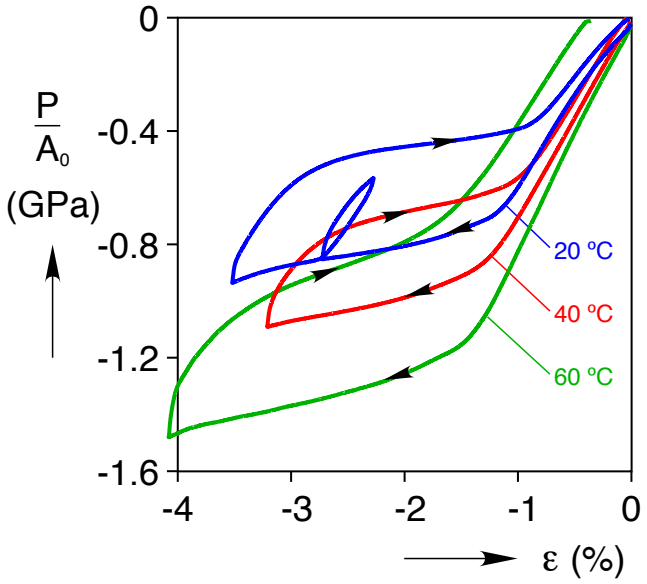


Figure 5.29: Mechanical responses of NiTi tubes in isothermal compression ($\dot{\delta}/L = 2.6 \times 10^{-4}$ s^{-1})

at higher temperatures exhibited the same general mechanical response profile raised to greater stresses, even as high as -1.5 GPa during the 60 °C experiment.

Aside from the obvious differences in stress levels and plateau strains, one important distinction from this mechanical response is the slope of the transformation plateau. The significant positive slope of the plateau ($E_t > 0$) suggests that there are no localized phenomena involved in transformation, since a localized transformation front would cause a flat isothermal tangent modulus ($E_t = 0$). The entire tube, at least at a structural level, should then exhibit uniform behavior.

5.3.3 Loading Rate Study

Performing an effective rate study using this experimental setup is problematic because of the large thermal mass of the compression grips. There exists a very different

temperature boundary condition on the 85 % portion of the tube within the fixture than the remaining 15 % exposed to air. Nevertheless, a rate study was performed, if only to verify the assertion that there was no localized behavior.

Mechanical responses from all three experiments from the rate study are shown in Figure 5.30. Three different rates were tested, each one decade apart: $\dot{\delta}/L = 6 \times 10^{-5}$ s^{-1}, $\dot{\delta}/L = 5.9 \times 10^{-4}$ s^{-1}, and $\dot{\delta}/L = 5.9 \times 10^{-3}$ s^{-1}. There is a slight increase in the transformation stress (30 MPa) on loading as the faster rate increases the rate of self-heating, and if $\partial\sigma_p/\partial T$ from Figure 5.32 is used, this rise corresponds to uniform heating of only 3 °C. Figure 5.31 shows that the heating was not uniform, and that the grips prevented most of the tube from straying from room temperature. The most notable difference is the lower stress reverse transformation plateau in the $\dot{\delta}/L = 5.9 \times 10^{-5}$ s^{-1}. This is likely a result of the 0.25 % larger strain and 30 MPa larger stress during that experiment. Chapter II showed that higher stresses during loading tend to depress the unloading plateau.

A single experiment is presented in Figure 5.31, at the rate of $\dot{\delta}/L = 5.9 \times 10^{-4}$ s^{-1}. Figure 5.31 is similar to previous figures used to describe tension experiments (like Figure 5.8) except that in this case, due to the limited free length, the IR images only capture a small portion of the tube length. The temperature history of the crown is in the background, with temperatures ranging from 12 to 28 °C. The stress history is in black, with the scale on the right of the figure.

The stress history proceeds qualitatively the same way as it did with the isothermal experiment in Figure 5.10. $A \rightarrow M^-$ transformation started gradually at about 30 s, and 735 MPa and continued until 58 s, reaching 910 MPa before the slope increases again.

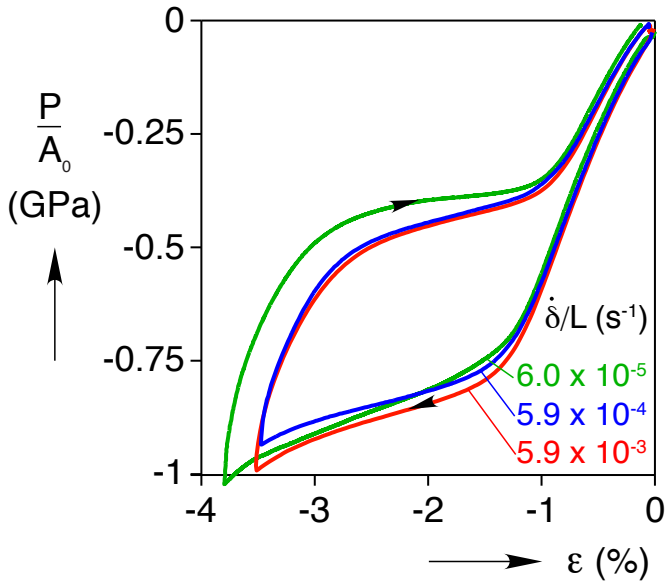


Figure 5.30: Mechanical responses of NiTi tubes in compression at various loading rates. Because of the large thermal mass of the buckling-arresting grips, loading rate had a minimal effect on the response.

During the loading plateau, there was smooth heating throughout the free length of the tube that was imaged, to a maximum of 26.7 °C. The temperature is lower towards the top and bottom of the image, since the grips acted as heat sinks to drive the temperature back to room temperature, 19 °C. The crosshead motion was reversed at 60 s, and the reverse $M^- \rightarrow A$ transformation occurred from about 75 to 100 s. The self-cooling during unloading was also smooth, with the center of the tube slightly colder because of the warmer grips. The tube reaches a minimum temperature of 12.5 °C. The thermal field shows no sign of localized behavior, with only smooth gradients.

5.4 Tension-Compression Asymmetry

In order to complete the phase diagram from Figure 5.11, the beginning and ending plateau stresses were extracted from Figure 5.29 and added to the diagram. Figure 5.32

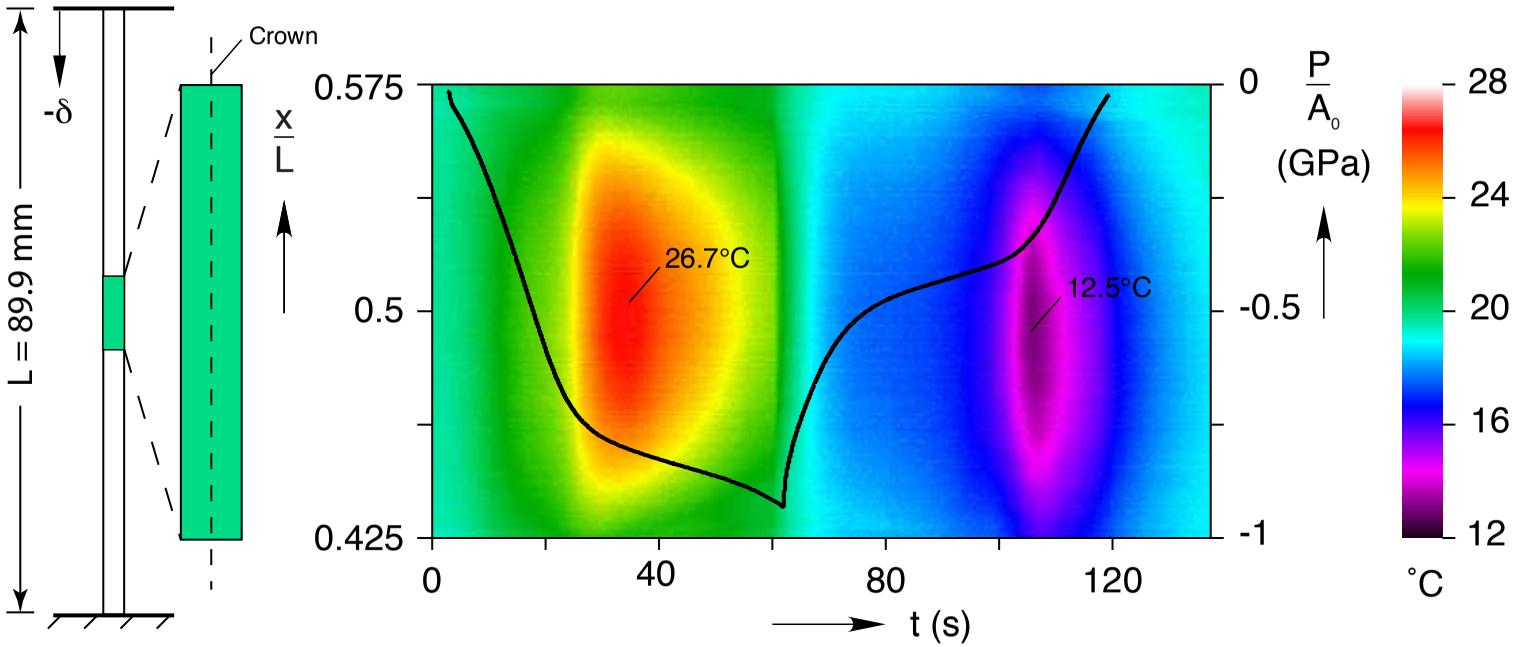


Figure 5.31: Stress and crown temperature history of a NiTi tube in uniaxial compression. Self-heating and cooling is smooth during transformation, free of any localized phenomena.

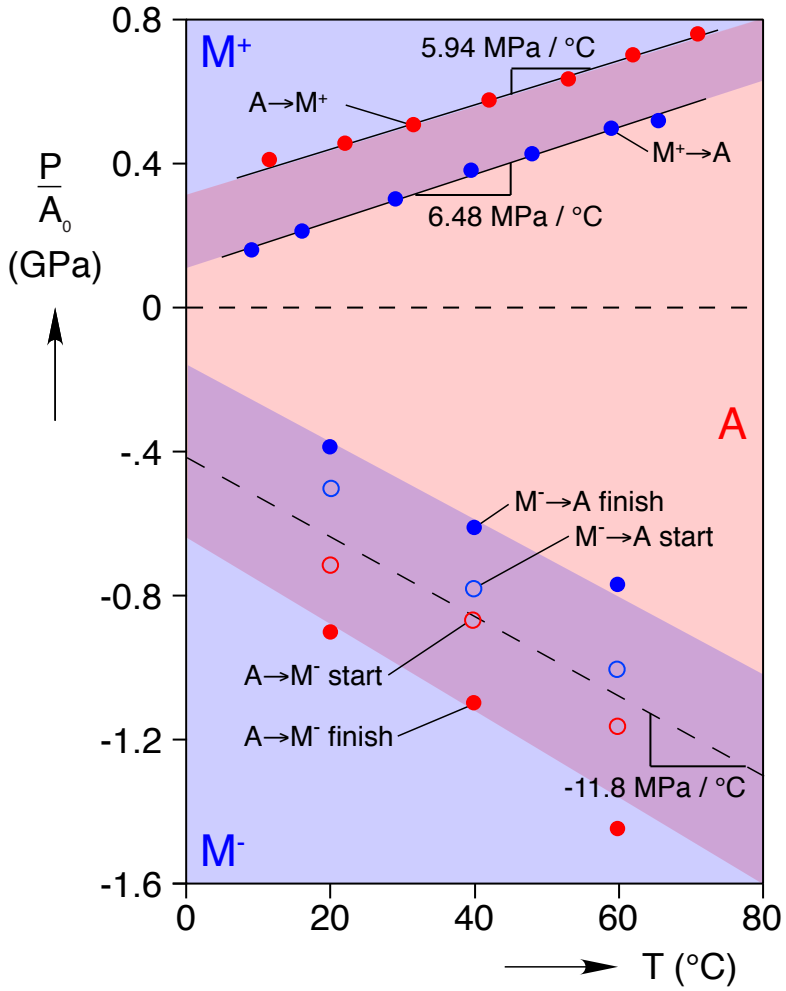


Figure 5.32: Propagation stresses for all isothermal tube experiments: $A \rightarrow M$ (red) and $M \rightarrow A$ (blue). Blue shading represents regimes of stable M and red shading regimes of stable A. Purple regions represent overlap where phase is history dependent.

shows the combined version, which is valid from 10 to 70 °C and now includes compression. In stress-temperature space, regions of unique A stability are shown in red, while regions of unique M^+ and M^- stability are in blue. Overlapping regions in purple indicate that either phase can be stable, depending on the history of the material (forward or reverse transformation). Because the transformation stress plateaus in Figure 5.29 were not flat, both the starting and ending stresses for each plateau are represented in the figure, blue for $M^- \rightarrow A$ and red for $A \rightarrow M^-$. Using a linear regression fit for all 12 points, the Clausius-Clapeyron slope $\partial\sigma_p/\partial T$ for compressive transformation is -11.8 MPa/°C, almost twice the value for tensile transformation.

This doubling of $\partial\sigma_p/\partial T$ from tension to compression makes sense from a thermodynamic modeling perspective. Using the simplified lumped model from Chapter IV, the Clausius-Clapeyron relation can be obtained by rearranging Equation 4.1.13 after assuming active transformation ($\mu = \mu_c$).

$$\frac{\partial\sigma}{\partial T} = \frac{-\rho T_R \Delta s}{T_R \Delta\varepsilon} \tag{5.4.1}$$

Because we lack a good DSC thermogram of this material, the (relatively low) value for NiTi, $T_R\Delta s = -11\mathrm{J/g}$, was generated from a fit of the available data. The value for the reference temperature $T_R = 237\mathrm{K}$ is the stress and hysteresis free transformation temperature, also known as the y-intercept for the linear fit of the plateau stresses for both $A \rightarrow M^+$ and $M^+ \rightarrow A$. Density is taken from the manufacturer as $6.5 \times 10^{-9}\mathrm{Mg/mm^3}$

Two similar experiments are overlaid in Figure 5.33, . The tension curve (black)

	Tension	Compression
Plateau Strain $\Delta\varepsilon$	5 %	-2.5 %
$A \rightarrow M$ Transformation Stress σ_p	470 MPa	-700 to -975 MPa
Clausius-Clapeyron slope $\partial\sigma_p/\partial T$	5.94 MPa/°C	11.8 MPa/°C
Latent Heat of Transformation $T_{\mathrm{R}}\Delta s$	-11 J/g	-11 J/g
Localized Behavior	Yes	No

Table 5.1: Comparison tension and compression response in RT air (from Figure 5.33) using several common metrics.

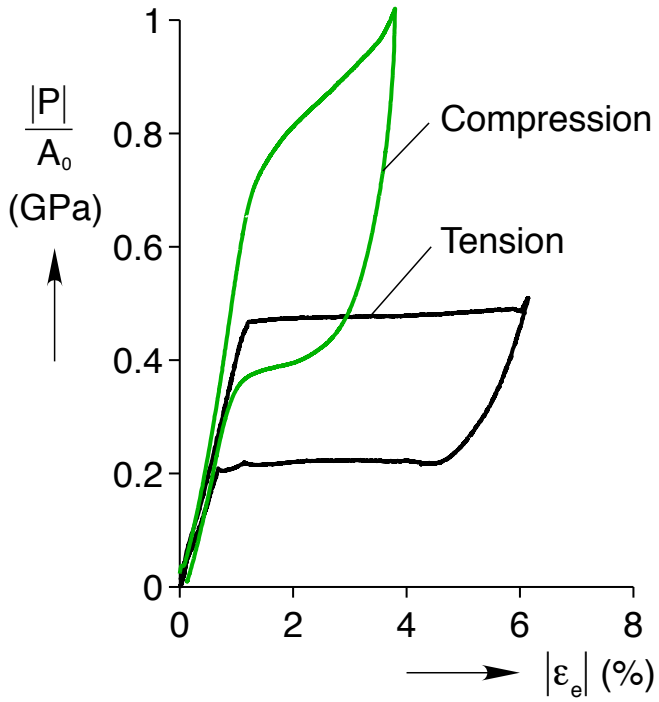


Figure 5.33: Comparison of tension and compression superelastic mechanical responses. Both experiments were performed at slow rates in room temperature air: $\dot{\delta}/L = 1 \times 10^{-4}\ \mathrm{s}^{-1}$ and $\dot{\delta}/L = 5.9 \times 10^{-5}\ \mathrm{s}^{-1}$ for tension and compression, respectively.

is from in Figure 5.15, performed in 19 °C air at the rate $\dot{\delta}/L = 1 \times 10^{-4}\,\mathrm{s}^{-1}$. The compression curve (green) is taken from Figure 5.30 and was performed in 19 °C air at $\dot{\delta}/L = 5.9 \times 10^{-5}\,\mathrm{s}^{-1}$. Using the $A \rightarrow M^{+}$ plateau strain for the tensile experiment, $\Delta\varepsilon = 0.05$, the slope predicted by Equation 5.4.1 is $\partial\sigma/\partial T = 6.03\mathrm{MPa}/^{\circ}\mathrm{C}$, a close match to the measured value of 5.94 MPa/°C. If we use the plateau strain for the compressive experiment, $\Delta\varepsilon = -0.025$, the model predicts a doubling in slope to $\partial\sigma/\partial T = -12.06\mathrm{MPa}/^{\circ}\mathrm{C}$, also close to the measured value of -11.8 MPa/°C. The model in Ch IV predicts that, assuming a constant Δs (a reasonable assumption), $\partial\sigma_p/\partial T$ is inversely related to the plateau strain, therefore it is natural to expect a higher $\partial\sigma_p/\partial T$ to correlate with a lower plateau strain.

5.5 Bending Experiments

The advantages of NiTi wires in bending were recognized early on [64], for use in orthodontics, drill bits, and stents. The literature is full of bending experiments, but most of them are qualitative, looking at bending response for direct application to a device rather than constitutive investigations. The most common testing modes are cantilevers [64] and three-point bending [65], but these bending modes produce nonuniform deformation of the specimen. NiTi wires were put in pure bending by Berg [6], but the apparatus was operated in load control, so data sparse along the transformation plateau. In addition, performing a superelastic experiment in load control removes the ability of the operator to prescribe transformation rate, introducing unknown rate effects.

5.5.1 Bending Experimental Setup

Just as with tension and compression, performing quality bending experiments with shape memory alloys is much more complex than bending of traditional alloys. The large plateau strains (5 % in tension, 2.5 % in compression) mean that the minimum radius of curvature for small wires can be prohibitively small. Assuming Euler-Bernoulli bending and isotropic properties (from tension experiments), the lower limit for radius of curvature for these 3.175 mm diameter tubes is $R_b \approx \frac{D}{2\varepsilon_{\max}} \approx 30\mathrm{mm}$, small but still manageable. A four-point bending machine was desired because it more closely provides pure, uniform bending. In addition, the setup needed to accommodate large rotations. A traditional four-point bending machine could not be used, since it cannot measure applied moments accurately at large rotations.

A scaled-down version of the four point bending machine in Corona and Kyrikides [22] was developed specifically for the tubes used in this chapter. The four loading points in Figure 5.34 ware stainless steel rollers, machined with a semi-circle relief to capture the tube in the same plane. The pegs were press-fit into miniature ball bearings so that they could rotate freely, allowing the tube to slide and removing any (total) axial tension. Each pair of bearings holding the pegs was in turn affixed to an aluminum wheel, which was then mounted to another bearing. Behind each wheel, an optical encoder measured the rotations θ_1 and θ_2 to a precision of 0.1^{o}. A single steel cable was wrapped around each wheel, one wheel at each end of the cable, and then pulled from above by a servo-mechanical load frame,. A pair of pulleys above balanced the load between the wheels. As before, fluid was pumped through the center of the tube for temperature control.

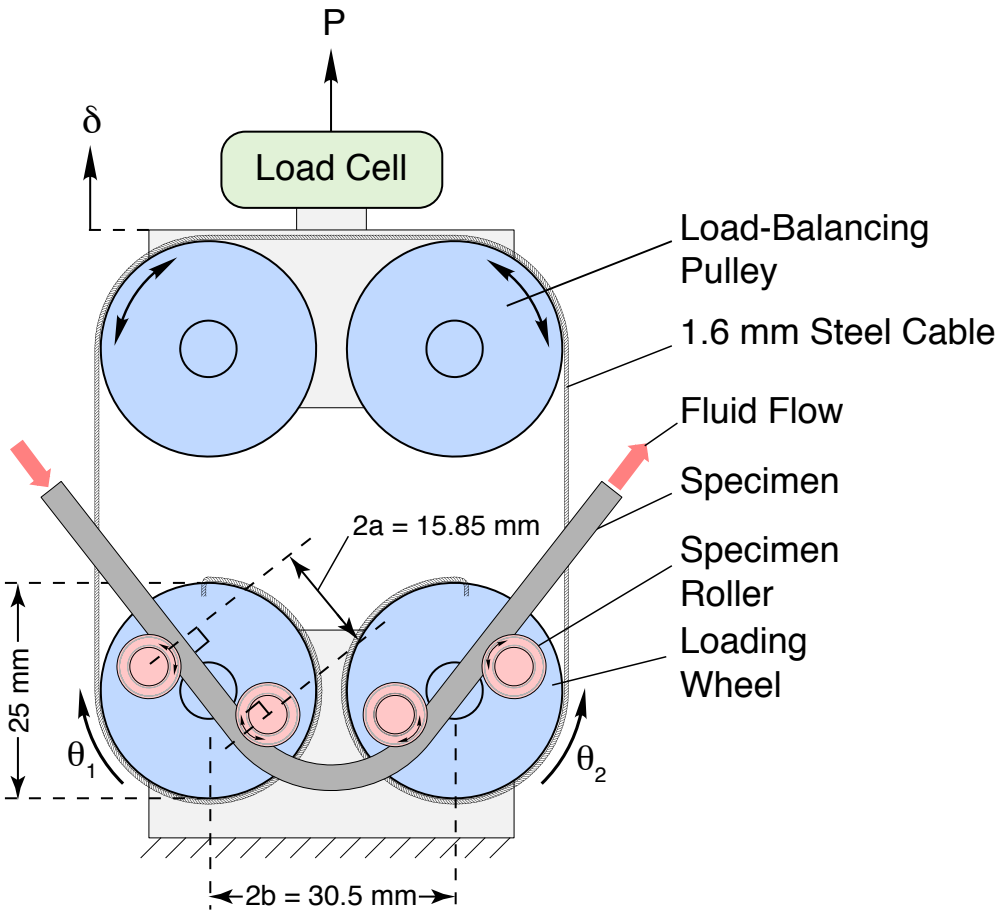


Figure 5.34: Four-point bending machine, front view

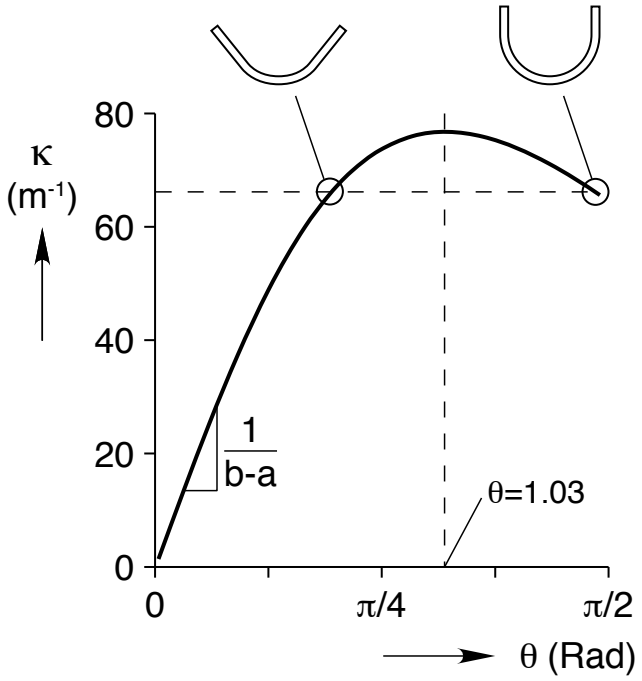


Figure 5.35: Rotation-curvature relationship for the machine in Figure 5.34. Pure bending between the loading wheels is assumed.

If pure bending is assumed, the average curvature of the tube can be calculated from the loading wheel rotation θ and the geometry of the system in Figure 5.34 as

$$\kappa = \frac{\sin\theta}{b - a\cos\theta} \tag{5.5.1}$$

where $1a$ is the distance between the contact points of the loading pegs (see Figure 5.34), and $2b$ the distance between the two loading wheel axes. As Figure 5.35 shows, this relationship is linear for small angles, but becomes very nonlinear as θ becomes large. Due to the geometry of the system, the loading wheels cannot be rotated more than $\theta = 1.03$ without a multivalued $\kappa - \theta$ relationship. This behavior is not intuitive, so Figure 5.36 shows a cartoon demonstrating an example multivalued configuration.

While the pulleys above the machine ensured an equal moment was applied to each

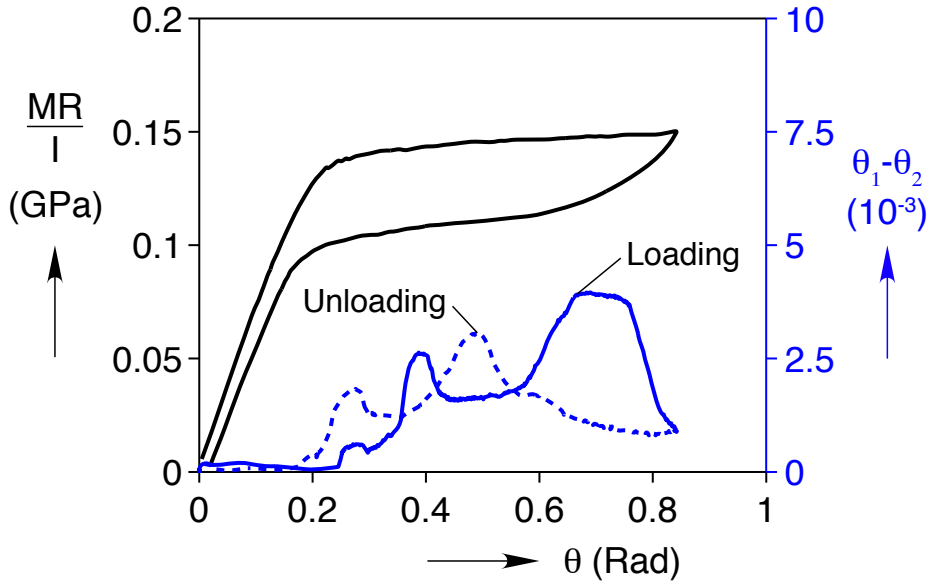


Figure 5.36: Isothermal bending response at 60 °C. In spite of a symmetric machine geometry, there is still some difference between the left wheel rotation θ_1 and right wheel rotation θ_2

side, they also raised the possibility of asymmetric (not pure) bending, should one wheel rotate more than the other. To investigate this possibility, Figure 5.36 shows a typical isothermal response at 60 °C. The moment measured at each pulley was converted to a bending-like stress measure by a the ratio R/I where R is the outer tube radius 1.59mm and I is the bending area moment of inertia, 2.95mm^4. The independent variable is the average rotation $\theta = \frac{1}{2}\theta_1 + \theta_2$ from the rotations θ_1 and θ_2 measured by each rotary encoder.

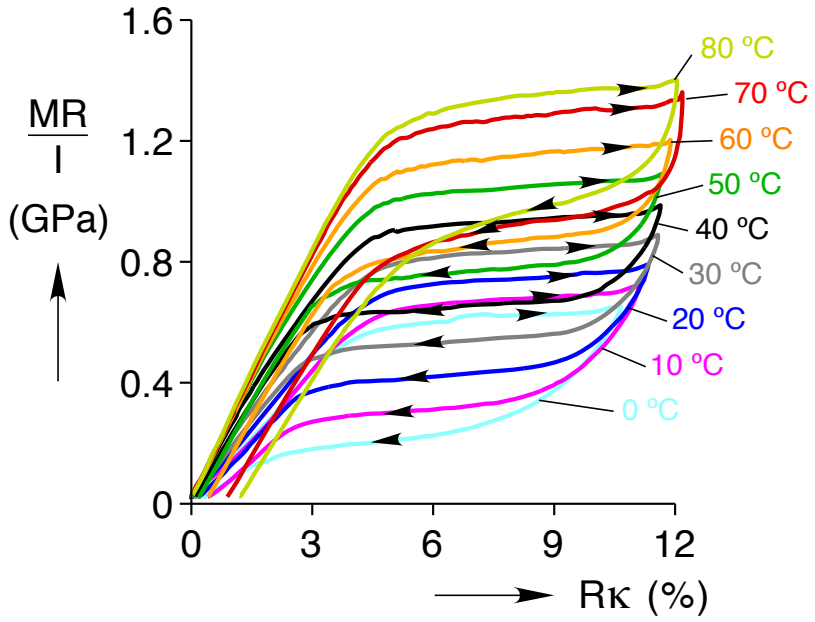
The isothermal response in Fig 5.36 was generated at a rotation rate of $\dot{\theta} = 6.66 \times 10^{-3}\text{rad/s}$. Though the moment is normalized to form a stress-like measure, it is not the maximum stress in the tube, since the stress distribution through the tube section is not linear and the neutral axis is not centered across the tube diameter. There were distinct loading and unloading plateaus, though they were flat like the isothermal tension experi-

ments. This is not surprising because the deformation has a strong radial gradient, the tube wall is curved (ring cross-section), and the response is a combination of tension (unstable) and compression (stable).

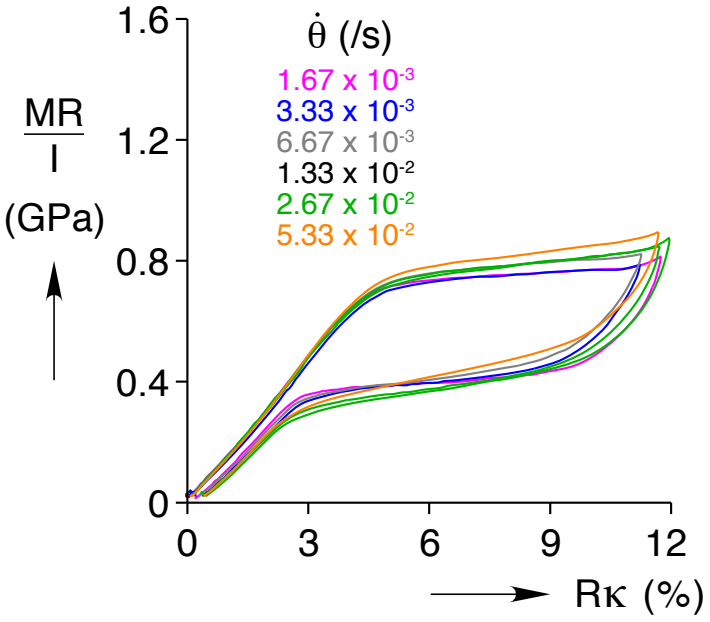
The difference $\theta_1 - \theta_2$ is in blue, and if the tube deflection were perfectly symmetric, it would be 0. However, there was a slight difference in rotation during loading transformation of about $4 \times 10^{-3}\mathrm{rad}$. The amount of asymmetry varies for each experiment, from $4 \times 10^{-3}\mathrm{rad}$ to $6 \times 10^{-2}\mathrm{rad}$ (0.23 to 3.4°). Without further testing, it is difficult to pinpoint the source of the asymmetry, which could be related to either 1) friction in the bearings of the machine, 2) asymmetric transformation fronts in the tube, or 3) an inherent structural instability, causing a bifurcation from the (symmetric) principle path. Since the difference $\theta_1 - \theta_2$ seems to deviate from zero near $\theta = 0.2\mathrm{rad}$, which coincides with the knee in the moment response, the likely reason is some asymmetry in the transformation distribution. This hypothesis is supported by a recent publication by Kyriakides *et. al.* [59] (after the date of these experiments). They demonstrated localization of curvature in pure bending of mild steel, which exhibits local instabilities (Lüders bands) similar to those in NiTi [21].

5.5.2 Bending Isothermal Responses and Loading Rate Study

All of the isothermal bending responses are shown in Figure 5.37(a), ranging from 0 to 80 °C. A lower isothermal temperature was possible with bending experiments than for the uniaxial ones because displacement was measured by rotary encoders which did not have any condensation effects (rather than a laser extensometer). For these figures, the rotation has been converted into a bending strain-like measure $\varepsilon_{\max} = R\kappa$, using the



(a)



(b)

Figure 5.37: Moment-curvature responses for (a) isothermal at $\dot{\theta} = 6.66 \times 10^{-3}\mathrm{rad/s}$, (b) in 19 °C air at various rates. The curvature estimate is normalized by the tube radius r and assumes pure bending behavior.

relationship in Equation (5.5.1). Though curvature was not uniform (so the conversion is not exact), Figure 5.35 shows some correction was necessary because the maximum rotation of 0.85 rad reached far into the nonlinear regime.

The responses follow the now familiar superelastic "flag" shape, with shallow yet positive-sloped (about 0.85 GPa) transformation plateaus. The curvature approximation predicts a maximum strain of nearly 12 %, though we know from tensile and compression experiments that strains of this magnitude are simply not possible without significant permanent deformation (there is very little in experiments below 60 °C). This suggests that the asymmetry is significant enough to cause a large error in the calculated strain, and that the assumption that plane sections remain plane may not be a good one.

Additional experiments were performed without fluid flowing through the tube to investigate the effect of loading rate on the bending response. The results are presented in Figure 5.37(b). Unfortunately, limitations of the load frame kept the maximum rate to$5.33 \times 10^{-2}\mathrm{rad/s}$ or $1.1 \times 10^{-2} \times \varepsilon_{max}/s$. This was not fast enough to cause enough self-heating to dramatically change the mechanical response. Still, there is a slight increase in plateau slope and maximum moment as the loading rate increases.

Chapter VI

Summary and Conclusions

6.1 Summary

This thesis has presented a series of experiments characterizing NiTi SMA wires and tubes. The primary goal was to document and develop a number of experimental techniques used to characterize SMA structures, especially those destined for use in devices. To further speed the development of those devices, a reduced order SMA actuator model was introduced which would be ideal for rapidly simulating SMA response in a device. Along the way, several new phenomena were observed in shakedown experiments using conditioned shape memory NiTi wires and tension/compression/bending experiments using virgin superelastic NiTi tubes.

First, two alloys of virgin NiTi wire were characterized in a series of historical experiments. The first alloy, an example of "shape memory" wire, had an austenite finish temperature $A_{\mathrm{f}} = 58$ °C, the second was "superelastic" wire with $A_{\mathrm{f}} = 13$ °C. The stress-free specific heats, transformation temperatures, and latent heats of transformation for each alloy were found using differential scanning calorimetry. There was a series of recommendations for obtaining quality DSC results, including methods for preparing NiTi wire in a

DSC sample pan and suggested temperature scan rates. Next, a series of isothermal tension experiments were carried out on each type of wire. Transformation stresses were used to create a quasi-phase diagram in stress-temperature space, and techniques for obtaining good quality temperature and displacement measurements were reviewed. Then, a specialized thermoelectric-based experimental setup was introduced for capturing the nucleation of localized regions of austenite and martensite (A and M^+) during isothermal tension.

Second, another alloy of NiTi wire was introduced, this time a conditioned wire ($A_\mathrm{f} = 78^\mathrm{o}\mathrm{C}$) which is commonly used in industry for actuator applications. The wire was characterized with a series of experiments using techniques explained in Chapter II. Because specimens had been conditioned (cycled) by the manufacturer, there was no longer any localized or unstable behavior, as there was in the virgin alloys. The conditioned wire also had a significant two-way shape memory effect, changing length even during stress-free thermal cycles. The concept of electrical resistivity (ER) measurements was introduced, along with a specialized method for measuring ER and strain along the same local gauge length. ER was shown to vary proportionally with strain under very specific conditions: only during transformation at a stress of at least 191 MPa. Apart from this specific region, a relatively large ER temperature coefficient and a strong dependence on R-phase presence made the relationship non-monotonic and hysteretic.

Shakedown is important for engineers who want to make sure their NiTi device will continue to function as desired through its entire lifecycle. Another thermoelectric-powered setup, using new techniques to ensure uniform yet rapid temperature changes, was used to study the shakedown of the conditioned NiTi wire under 150 constant load thermal cycles.

Wires were thermally cycled 150 times under loads ranging from 19 to 478 MPa. For large loads (above 191 MPa), temperature hysteresis decreased while maximum and minimum strains increased. For small loads, specimens exhibited reverse shakedown, in effect undoing some of the conditioning imposed on the specimen by the manufacturer. A pair of experiments were designed to answer the question of which transformation ($A \rightarrow M^+$ or $M^+ \rightarrow A$) was responsible for the majority of shakedown. Alternate high (379 MPa) and low (96 MPa) loads were applied during cooling and heating, with the expectation that the transformation under high load would be responsible for shakedown. It was found that the $M^+ \rightarrow A$ transformation was responsible for more than twice as much change in the response as for $A \rightarrow M^+$, an important finding that should be explored in the future.

Thirdly, a reduced-order shape memory actuator model was presented. Several simplifying assumptions were applied to the full-featured constitutive model of Chang [15], namely the lack of spacial gradients (no fronts, thermal gradients) and an actuator design that prohibited formation of thermal martensite ($M^{+/-}$). After identifying the important nondimensional parameters governing the response, an analytical solution (for the first time) to the governing equations was presented. The reduced-order model response compared favorably to that of the full-featured model for thermally insulated boundary conditions. Next, the power of the analytical solution was the exercised in a series of design studies looking at spring sizing, power requirements, response time, and energy efficiency.

Finally, a third series of experiments were presented characterizing the response of NiTi tubes in tension, compression, and bending. A custom setup and post-processing software was used to provide a full 360^o thermal field of the tubes during tension, tracking front

morphology by the latent heat exchanged during transformation. Nucleation events took the shape of a helix, inclined at an angle of about 62 °to the axis of loading. In isothermal (slow) experiments, this helix would lengthen about two revolutions before either halting and self-merging into a pair of branched fronts. At high loading rates, the helix would continue to grow until it encompassed the entire tube, a phenomenon only previously observed with $A \rightarrow M^+$ transformation [30] but (newly) captured here during $M^+ \rightarrow A$ as well. Another new phenomenon captured was the nucleation of a single helical inclusion between two active ringed $A \rightarrow M^+$ fronts. That particular experiment will provide a valuable test case to validate future thermally coupled 2D SMA models.

The tube response was shown to be highly anisotropic, as the compression response was stable (no fronts), exhibited higher transformation stresses, and had lower transformation strains. The four point bending setup was designed to directly measure the applied moment, but relied on uniform deformation to measure curvature. However, some experiments showed a significant asymmetry in the rotations between the two ends of the tube, making it difficult to draw quantitative conclusions about the bending constitutive behavior.

6.2 Future Work

Many opportunities exist for future work. Each focus of the thesis: shakedown of conditioned wire, reduced order modeling, and NiTi tubes, offers enough questions to occupy years of research.

6.2.1 Thermo-electro-mechanical Shakedown

Regarding the shakedown response of NiTi wires, the surface has only been scratched. To be thorough, it would be helpful to repeat the isothermal experiments while taking resistivity measurements, perhaps even performing multiple cycles to gather the shakedown response during superelastic cycles. Then, constant load experiments could be repeated using virgin (not conditioned) wires, preferably of the same alloy. In addition, constant load experiments could be repeated with an order of magnitude greater number of cycles. One of the time constants in the double-exponential curve fit of the shakedown response was 95 cycles, though with data for only 150 cycles it is premature to suggest this a true time constant of a processes responsible for long term shakedown. The two experiments at the end of the study, in Section 3.5, showed clockwise and counter-clockwise traverses of the same stress-temperature space rectangle. Together, they represent a significant finding that should be explored in more detail to explain the surprising differences in shakedown responses.

Anecdotal evidence from conversations with colleagues in the field has also suggested that the degree of shakedown is strongly dependent on the extremes of transformation. For example, the temperature range for the experiment in Figure 3.9(b) could be changed from 25-180 °C to 75-120 °C. Another way of limiting the amount of transformation would be to devise a control system that calculated the current slope of the R_{e}-T response and reversed the temperature ramp when a local minimum or maximum had been reached. Transformation also could be limited using a hard-stop to physically limit the amount of extension allowed [93].

Finally, though there has been significant characterization of the conditioned wire with both constant load and constant temperature experiments, additional experiments could be performed coupling two of the three possible control variables together. For example, the addition of a spring in series with the wire could couple the stress and strain in an actuator-like path. A custom load frame control algorithm could couple the applied load or strain to the temperature.

6.2.2 Reduced-order SMA actuator model

Regarding the reduced-order model, a number of refinements can be applied without losing the analytical nature of the solution. One assumption of the model is that the transformation strain β is constant for all conditions. This is far from true, as Figures 2.13 and 2.16 shows how β changes with temperature (and stress). After choosing an actuator design, one could update β to be a function of the spring parameters and maximum stress.

Since the model is only valid for constant applied power and ambient temperature, we suggest that a more powerful framework would break the solution into many pieces in time, in order to allow an arbitrary applied power and ambient temperature. At any of these intervals, most of the material properties in the model could be updated to reflect more realistic behavior. An example of a parameter that could be adjusted is the film coefficient h, which is assumed to be constant in the model. In reality, h is highly nonlinear, dependent on the air speed, wire orientation, SMA temperature, and ambient temperature. One could update the film coefficient at each interval to reflect a better estimate for the current conditions.

6.2.3 NiTi Tubes

The experiments presented in this thesis are just the beginning of what is hoped to be a better understanding of the way shape memory alloys behave in a 3D structure. In tension, the relationship between loading rate, helix length, and specimen length scales is a complex one that was only addressed by observation. Matching the experimental results to a thermally coupled finite element model would prove invaluable, and would serve as an excellent test of the model's fidelity. In bending, the experimental apparatus could be improved to remove the discrepancy between rotations at the ends of the tube. All the experiments presented are prime targets for repeating with the technique of digital image correlation (DIC), which would provide a full-field strain map of the specimen, although the cylindrical surface could present some challenges.

Following experiments in tension, compression, and bending, it is natural to perform experiments in torsion, tension-torsion, and compression-torsion. Detailed DIC would also be valuable for these experiments. The mere existence of such a range of deformation states would be invaluable to modelers.

The work of Favier et al. [81] has resulted in a post-processing technique to turn a thermal field into a heat source field, given some assumptions about material conductivity, heat capacity, and film coefficient. If this technique were applied to the tubes from this thesis, at a very high loading rate (to reduce error), it is conceivable a stress-dependent latent heat of transformation could be integrated from the results. In a sense, with enough experiments one could get the same data a thermogram offers, except under stress, something impossible to do in a DSC. This could be repeated in tension, compression, and torsion to

investigate the claim made in Section 5.4, that the latent heat of transformation is invariant between deformation modes.

Bibliography

[1] V. S. Ananthan and E. O. Hall. "Microscopic shear bands at Luders fronts in mild steel,". *Scripta Metallurgica*, 21(4):519–520, 1987.

[2] GF Andreasen and RD Barrett. "An evaluation of 55 cobalt substituted Nitinol wire for use in orthodontics,". *Journal of the American Dental Association*, 82(6):1373–5, 1971.

[3] V. Antonucci, G. Faiella, M. Giordano, F. Mennella, and L. Nicolais. "Electrical resistivity study and characterization during NiTi phase transformations,". *Thermochimica Acta*, 462(1-2):64–69, Oct 2007.

[4] C.D. Babcock. "Shell stability,". *Trans. ASME, J. Appl. Mech. (USA)*, 50(4B):935 – 40, 1983. shell stability;shell buckling;elastic postbuckling behaviour;imperfection sensitivity;plastic buckling;dynamic buckling;.

[5] B Berg. "Twist and Stretch: Combined Loading of Pseudoelastic NiTi Tubing,". In *Proceedings of the 2nd International Conference on SMA and Superelastic Technologies*, 1997.

[6] B.T. Berg. "Bending of superelastic wires, Part I: Experimental aspects,". *Journal of Applied Mechanics*, 62:459–465, 1995.

[7] Olivier W. Bertacchini, James H. Mabe, D.C. Lagoudas, and Frederick T. Calkins. "Thermomechanical cyclic loading and fatigue life characterization of nickel rich NiTi shape memory alloy actuators,". In *Proceedings of SPIE 6929*, 2008.

[8] Kaushik Bhattacharya. *Microstructure of Martensite: Why it forms and how it gives rise to the shape-memory effect*, Oxford Series on Materials Modelling, Oxford University Press, first edition, 2003.

[9] U Blum, G Voshage, J Lammer, F Beyersdor, D Tollner, G Kretschmer, G Spillner, P Polterauer, G Nagel, and T Holzenbein. "Endoluminal stent-grafts for infrarenal abdominal aortic aneurysms,". *New England Journal of Medicine*, 336:13–20, 1997.

[10] Zhonghe Bo and D. C. Lagoudas. "Thermomechanical modeling of polycrystalline SMAs under cyclic loading, Part I: theoretical derivations,". *International Journal of Engineering Science*, 37:1089–1140, 1999.

[11] Zhonghe Bo and D. C. Lagoudas. “Thermomechanical modeling of polycrystalline SMAs under cyclic loading, Part II: material characterization and experimental results for a stable transformation cycle,”. *International Journal of Engineering Science*, 37:1141–1173, 1999.

[12] J. P. Boyd. “Global Approximations to the Principal Real-Valued Branch of the Lambert W-function,”. *Applied Mathematics Letters*, 11(6):27–31, 1998.

[13] Diann Brei, Jonathan Luntz, John Shaw, Nancy L. Johnson, Alan L. Browne, Paul W. Alexander, and Nilesh D. Mankame. “General Motors and the University of Michigan smart materials and structures collaborative research laboratory,”. In L. Porter Davis, B. K. Henderson, and M. Brett McMickell, editors, *In: Proceedings of SPIE, Industrial and Commercial Applications of Smart Structures Technologies 2008*, volume 6527, page 65270U. SPIE, 2007.

[14] L. C. Brinson. “One Dimensional Constitutive Behavior of Shape Memory Alloys: thermomechanical derivation with non-constant material functions,”. *Journal of Intelligent Material Systems and Structures*, 4:229–242, 1993.

[15] B.-C. Chang, J. A. Shaw, and M. A. Iadicola. “Thermodynamics of Shape Memory Alloy Wire: Modeling, Experiments, and Application,”. *Continuum Mechanics and Thermodynamics*, 18(1-2):83–118, 2006.

[16] Eunsoo Choi, Tae-hyun Nam, Man-Cheol Kim, Jong Wan Hu, Bak-Soon Cho, and Soon-Jong Yoon. “Recovery and Residual Stress of Shape Memory Alloy Wires and its Application,”. In *Proceedings of the 2009 Conference on Smart Materials, Adaptive Structures, and Intelligent Systems*, Sep 2009.

[17] C. B. Churchill, J. A. Shaw, and M. A. Iadicola. “Tips and Tricks for Characterizing Shape Memory Alloy Wire: Part 2 – Fundamental Isothermal Responses,”. *Experimental Techniques*, 33(1):51–62, Jan-Feb 2009.

[18] Chris B. Churchill and John A. Shaw. “Shakedown Response of Conditioned Shape Memory Alloy Wire,”. In *Behavior and Mechanics of Multifunctional and Composite Materials, Proceedings of the SPIE 15th Annual International Symposium on Smart Structures and Materials*, volume 6928, 2008.

[19] P. D. Coates and I. M. Ward. “Neck Profiles in Drawn Linear Polyethylene,”. *Journal of Materials Science*, 15:2897–2914, 1980.

[20] R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. “On the Lambert W function,”. *Advanced in Computational Mathematics*, 5:329–359, 1996.

[21] E. Corona, M. Iadicola, and J. Shaw. “Buckling of Steel Bars with Lüders Bands,”. *International Journal of Solids and Structures*, 39:3313–3336, 2002.

[22] Edmundo Corona and Stelios Kyriakides. “An experimental investigation of the degradation and buckling of circular tubes under cyclic bending and external pressure,”. *Thin-Walled Structures*, 12(3):229 – 263, 1991.

[23] S. Daly, G. Ravichandran, and K. Bhattacharya. "Stress-induced martensitic phase transformation in thin sheets of Nitinol,". *Acta Materialia*, 55(10):3593–3600, 2007.

[24] T. W. Duerig, K. N. Melton, D. Stöckel, and C. M. Wayman, editors. *Engineering Aspects of Shape Memory Alloys*, Butterworth-Heinemann, Boston, MA, 1990.

[25] T. W. Duerig, A Pelton, and D Stockel. "An overview of nitinol medical applications,". *Materials Science and Engineering A*, 273-275:149–160, 1999.

[26] J. Y. Dyau and S. Kyriakides. "On the Propagation Pressure of Long Cylindrical Shells under External Pressure,". *International Journal of Mechanical Sciences*, 35(8):675–713, 1993.

[27] R. Elliott, J. A. Shaw, and N. Triantafyllidis. "Stability of Crystalline Solids—II: Application to temperature-induced martensitic phase transformations in a bi-atomic crystal,". *Journal of the Mechanics and Physics of Solids*, 54:193–232, 2006.

[28] J. L. Ericksen. "Equilibrium of Bars,". *Journal of Elasticity*, 5(3-4):191–201, 1975.

[29] J. C. Escobar, R. J. Clifton, and S.-Y. Yang. "Stress-Wave-Induced Martensitic Phase Transformations in NiTi,". In M. D. Furnish, L. C. Chhabildas, and R. S. Hixson, editors, *Shock Compression of Condensed Matter*, AIP Conference Proceedings #505, pages 267–270. American Institute of physics, 1999.

[30] D Favier, H Louche, P Schlosser, L Orgeas, P Vacher, and L Debove. "Homogeneous and heterogeneous deformation mechanisms in an austenitic polycrystalline Ti-50.8 at.Investigation via temperature and strain fields measurements,". *Acta Materialia*, 55(16):5310–5322, Jan 2007.

[31] P. Feng and Q.P. Sun. "Experimental investigation on macroscopic domain formation and evolution in polycrystalline NiTi microtubing under mechanical force,". *Journal of the Mechanics and Physics of Solids*, 54:1568–1603, 2006.

[32] J. Fraden. *Handbook of Modern Sensors*, American Institute of Physics, Woodbury, NY, 1997.

[33] Carl P. Frick, Alicia M. Ortega, Jeffrey Tyber, A. El. M. Maksound, Hans J. Maier, Yinong Liu, and Ken Gall. "Thermal processing of polycrystalline NiTi shape memory alloys,". *Materials Science and Engineering: A*, 405(1-2):34–49, 2005.

[34] H. Funakubo, editor. *Shape Memory Alloys*, Gordon and Breach Science Publishers, New York, NY, 1987.

[35] K. Gall and H. Sehitoglu. "The Role of Texture in Tension-Compression Asymmetry in Polycrystalline NiTi,". *International Journal of Plasticity*, 15:69–92, 1999.

[36] K. Gall, H. Sehitoglu, Y. Chumlyakov, and I. Kireeva. "Tension-Compression Asymmetry of the Stress-Strain Response in Aged Single Crystal and Polycrystalline NiTi,". *Acta Materialia*, 47(4):1203–1217, 1999.

[37] L. Gong and S. Kyriakides. "On the Crushing Stress of Open Cell Foams,". *Journal of Applied Mechanics*, 73:807–814, Sep. 2006.

[38] K Goubaa, L Jordan, M Masse, and G Bouquet. "Efficiency of Various Techniques in Dectecting the "R"-Phase in Ni-Ti, Ni-Ti-Cu, and Ni-Ti-Co Shape Memory Alloys,". *Scripta Metallurgica*, 26:1163–1168, 1992.

[39] C. G'sell and J. Jonas. "Determination of the plastic behaviour of solid polymers at constant true strain rate,". *Journal of Materials Science*, 14(3):583–591, 03 1979/03/21.

[40] Y.J. He and Q.P. Sun. "Effects of structural and material length scales on stress-induced martensite macro-domain patterns in tube configurations,". *International Journal of Solids and Structures*, 46:3045–3060, 2009.

[41] Z. He, K. R. Gall, and L. C. Brinson. "Use of Electrical Resistance Testing to Redefine the Transformation Kinetics and Phase Diagram for Shape Memory Alloys,". *MMT*, 37A(3):579, 2006.

[42] D. A. Hebda and S. R. White. "Effect of Training Conditions and Extended Thermal Cycling on Nitinol Two-Way Shape Memory Behavior,". *Smart Materials and Structures*, 4:298–304, 1995.

[43] O. Heintze and S. Seelecke. "A coupled thermomechanical model for shape memory alloys - From single crystal to polycrystal,". *Materials Science and Engineering A*, 481:389–394, 2008.

[44] O. Heintze and S. Seelecke. "A coupled thermomechanical model for shape memory alloys- From single crystal to polycrystal,". *Materials Science and Engineering A – Structural Materials Properties, Microstructure and Processing*, 481-482:389–394, 2008.

[45] J.E. Huber, N.A. Fleck, , and M.F. Ashby. "The selection of mechanical actuators based on performance indices,". *Proceedings of the Royal Society of London, Series A*, 453(1965):2185–2205, 1997.

[46] C.M. Hwang and C.M. Wayman. "Phase Transformations in TiNiFe, TiNiAl, and TiNi Alloys,". *Scripta Metallurgica*, 17:1345–1350, 1983.

[47] M. A. Iadicola. *On Propagating Instabilities in Nickel-Titanium and Steel Alloys*. PhD Dissertation, University of Michigan, 2002.

[48] M. A. Iadicola and J. A. Shaw. "The Effect of Uniaxial Cyclic Deformation on the Evolution of Phase Transformation Fronts in Pseudoelastic NiTi Wire,". *Journal of Intelligent Material Systems and Structures*, 13(2):143–156, 2002.

[49] M. A. Iadicola and J. A. Shaw. "An Experimental Setup for Measuring Unstable Thermo–Mechanical Behavior of Shape Memory Alloy Wire,". *Journal of Intelligent Material Systems and Structures*, 13(2):157–166, 2002.

[50] M. A. Iadicola and J. A. Shaw. "Rate and Thermal Sensitivities of Unstable Transformation Behavior in a Shape Memory Alloy,". *International Journal of Plasticity*, 20:577–605, 2004.

[51] William D. Callister Jr. *Materials Science and Engineering: An Introduction*, Wiley, 6th edition, 2003.

[52] Y. Jung, P Papadopoulos, and R.O. Ritchie. "Constitutive modelling and numerical simulation of multivariant phase transformation in superelastic shape-memory alloys,". *International Journal of Numerical Methods in Engineering*, 60:429–460, 2004.

[53] George B Kauffman and Isaac Mayo. "The Story of Nitinol: The Serendipitous Discovery of the Memory Metal and Its Applications,". *The Chemical Educator*, 2(2):1–21, 1996.

[54] P. Krulevitch, P.B. Ramsey, J.C. Trevino, J. Hamilton, and M.A. Northrup. "Thin film shape memory alloy microactuators,". *Journal of Microelectromechanical Systems*, 5(4):270–282, Dec 1996.

[55] S. Kyriakides. *Propagating Instabilities in Structures*, volume 30 of *Advances in Applied Mechanics*, pages 67–189. Academic Press, Boston, MA, Oct. 1994.

[56] S. Kyriakides. *Propagating Instabilities in Structures*, volume 30 of *Advances in Applied Mechanics*, pages 67–189. Academic Press, Boston, MA, Oct. 1994.

[57] S. Kyriakides and Y.-C. Chang. "The Initiation and Propagation of a Localized Instability in an Inflated Elastic Tube,". *International Journal of Solids and Structures*, 27(9):1085–1111, 1991.

[58] S. Kyriakides and J. E. Miller. "On the Propagation of Lüders Bands in Steel Strips,". *Journal of Applied Mechanics*, 67:645–653, Dec. 2000.

[59] S Kyriakides, A Ok, and E Corona. "Localization and propagation of curvature under pure bending in steel tubes with Lüders ...,". *IJSS*, Jan 2008.

[60] D Lagoudas, D Miller, L Rong, and P Kumar. "Thermomechanical fatigue of shape memory alloys,". *Smart Materials and Structures*, 18:1–12, Jan 2009.

[61] Z. Q. Li and Q. P. Sun. "The initiation and growth of macroscopic martensite band in nano-grained NiTi microtube under tension,". *International Journal of Plasticity*, 18(11):1481–1498, 2002.

[62] Y. Liu, Y. Liu, and J. Van Humbeeck. "Two-way Shape Memory Effect Developed by Martensite Deformation in NiTi,". *Acta Materialia*, 41(1):199–209, 1999.

[63] Y Liu, Z Xie, J Van Humbeeck, and L Delaey. "Asymmetry of stress–strain curves under tension and compression for NiTi shape memory alloys,". *Acta Materialia*, 46(12):4325–4338, 1998.

[64] I. Lopez, J. Goldberg, and C.J. Burstone. "Bending characteristics of nitinol wire,". *American Journal of Orthodontics*, 75(5):569–575, 1979.

[65] H.Y. Luo and E.W. Abel. "A comparison of methods for the training of NiTi two-way shape memory alloy,". *Smart Materials and Structures*, 16:2543–2549, 2007.

[66] Hitoshi Matsumoto. "Electrical resistivity of NiTi with high transformation temperature,". *Journal of Alloys and Compounds*, 370(1-2):244–248, 2004.

[67] Hitoshi Matsumoto. "Enhancement of a peak in electrical resistivity of NiTi alloy in early stage of transformation cycles,". *Journal of Alloys and Compounds*, 364:132–136, 2004.

[68] J. McCormick, R. DesRoches, D. Fugazza, and F. Auricchio. "Seismic vibration control using superelastic shape memory alloys,". *Trans. ASME, J. Eng. Mater. Technol. (USA)*, 128(3):294 – 301, 2006/07/.

[69] J.M. McNaney, V. Imbeni, Y. Jung, P Papadopoulos, and R.O. Ritchie. "An Experimental Study of the Superelastic Effect in a Shape-Memory Nitinol Alloy Under Biaxial Loading,". *Mechanics of Materials*, 35(10):969–986, 2003.

[70] K. L. Ng and Q. P. Sun. "Stress-induced phase transformation and detwinning in NiTi polycrystalline shape memory alloy tubes,". *Mechanics of Materials*, 38:41–56, 2006.

[71] M. Nishida, S. Ii, K. Kitamura, T. Furukawa, A. Chiba, T. Hara, and K. Hiraga. "New deformation twinning mode of B19' martensite in Ti-Ni shape memory alloy,". *Scripta Materialia*, 39(12):1749–1754, 1998.

[72] L Orgeas and D Favier. "Non-symmetric tension-compression behaviour of NiTi alloy,". *J Phys Iv*, 5(8):605–610, 1995.

[73] K. Otsuka and X. Ren. "Physical Metallurgy of Ti-Ni-based Shape Memory Alloys,". *Progress in Materials Science*, 50:511–678, 2005.

[74] K. Otsuka and C. M. Wayman, editors. *Shape Memory Materials*, Cambridge University Press, Cambridge, UK, 1998.

[75] S. Papka and S. Kyriakides. "In–plane Compressive Response of Crushing of Honeycomb,". *Journal of the Mechanics and Physics of Solids*, 42:1499–532, 1994.

[76] R.O. Perkins, J.; Sponholz. "Stress-induced martensitic transformation cycling and two-way shape memory training in Cu-Zn-Al alloys,". *Metallurgical Transactions A*, 15A:313–321, Feb 1984.

[77] R Plietsch and K Ehrlich. "Strength differential effect in pseudoelastic NiTi shape memory alloys,". *Acta Materialia*, 45(6):2417–2424, 1997.

[78] Peter Popov and D.C. Lagoudas. "A 3D constitutive model for shape memory alloys incorporating pseudoelasticity and detwinning of self-accommodated martensite,". *International Journal of Plasticity*, 23(10-11):1679–1720, 2007.

[79] M. Pozzi and G. Airoldi. "The electrical transport properties of shape memory alloys,". *Mat Sci Eng A-Struct*, 275:300–304, Jan 1999.

[80] Jorma Ryhanen. *Biocompatibility Evaluation of Nickel-Titanium Shape Memory Metal Alloy*. PhD Dissertation, Oulu University, P.O. Box 22 FIN-90221 Oulu Finland, May 1999.

[81] P Schlosser, H Louche, D Favier, and L Orgéas. "Image processing to estimate the heat sources related to phase transformations during tensile tests of NiTi tubes,". *Strain*, 43(3):260–271, 2007.

[82] M. Schwartz. *Encyclopedia of Smart Materials*, volume 1, John Wiley and Sons, New York, NY, 2002.

[83] S. Seelecke, M. Kohl, and B. Krevet. "Thermomechanical Finite Element Analysis of a Shape Memory Alloy Cantilever Beam,". In *Proceedings of SPIE 6166*, 2006.

[84] J. A. Shaw. *Material Instabilities in a Nickel-Titanium Shape Memory Alloy*. PhD dissertation, The University of Texas at Austin, Department of Aerospace Engineering, Jan. 1997.

[85] J. A. Shaw. "Simulations of Localized Thermo–Mechanical Behavior in a NiTi Shape Memory Alloy,". *International Journal of Plasticity*, 16(5):541–562, 2000.

[86] J. A. Shaw and S. Kyriakides. "Thermomechanical Aspects of NiTi,". *Journal of the Mechanics and Physics of Solids*, 43(8):1243–1281, 1995.

[87] J. A. Shaw and S. Kyriakides. "On the Nucleation and Propagation of Phase Transformation Fronts in a NiTi Alloy,". *Acta Materialia*, 45(2):683–700, 1997.

[88] J. A. Shaw and S. Kyriakides. "Initiation and Propagation of Localized Deformation in Elasto–Plastic Strips under Uniaxial Tension,". *International Journal of Plasticity*, 13(10):837–871, 1998.

[89] John A. Shaw, Chris B. Churchill, and Mark A. Iadicola. "Tips and Tricks for Characterizing Shape Memory Alloy Wire: Part 1 – Differential Scanning Calorimetry & Basic Phenomena,". *Experimental Techniques*, 32(5):55–62, Sep-Oct 2008.

[90] John A. Shaw and Christopher B. Churchill. "A Reduced-order Thermomechanical Model & Analytical Solution for Uniaxial SMA Wire Actuators,". *Smart Materials and Structures*, 18(6):1–21, 2009.

[91] John A. Shaw, Alan S. Jones, and Alan S. Wineman. "Chemorheological response of elastomers at elevated temperatures: Experiments and simulations,". *Journal of the Mechanics and Physics of Solids*, 53(12):2758 – 2793, 2005.

[92] P. Sittner, P. Sedlak, M. Landa, V. Novak, and P. Lukas. "*In situ* experimental evidence on R-phase related deformation processes in activated NiTi wires,". *Materials Science and Engineering A*, 438–440:579–584, 2006.

[93] Helen Sun, Anupam Pathak, Jonathan Luntz, Diann Brei, Paul W. Alexander, and Nancy L. Johnson. "Stabilizing Shape Memory Alloy Actuator Performance Through Cyclic Shakedown: an Empirical Study,". In *Proceedings of the SPIE 15th Annual International Symposium on Smart Structures and Materials*, page (submitted), 2008.

[94] Qing-Ping Sun and Zhi-Qi Li. "Phase transformation in superelastic NiTi polycrystalline micro-tubes under tension and torsion–from localization to homogeneous deformation,". *International Journal of Solids and Structures*, 39(13-14):3797–3809, 2002.

[95] M. A. Sutton T. C. Chu, W. F. Ranson. "Applications of digital-image-correlation techniques to experimental mechanics,". *Experimental Mechanics*, 25(3):232–244, Sep. 1985.

[96] Yuji Takeda, Hiroki Cho, Takaei Yamamoto, Toshio Sakuma, and Akihiko Suzuki. "Control Characteristics of Shape Memory Alloy Actuator Using Resistance Feedback Control Method,". *Advances in Science and Technology*, 59:178–183, 2008.

[97] Jeffrey Tyber, Jason McCormick, Ken Gall, Reginald DesRoches, Hans J. Maier, and A. El. Maksound. "Structural Engineering with NiT. I: Basic Materials Characterization,". *Journal of Engineering Mechanics*, 133(9):1009–1018, Sep 2007.

[98] J. Uchil, K. K. Mahesh, and K. Ganesh Kumara. "Electrical resistivity and strain recovery studies on the effect of thermal cycling under constant stress on R-phase in NiTi shape memory alloy,". *Physica B*, 324:419–428, Oct 2002.

[99] P. I. Vincent. "The Necking and Cold-Drawing of Rigid Plastics,". *Polymer*, 1:7–19, 1960.

[100] Y. Wang, Z.F. Yue, and J. Wang. "Experimental and numerical study of the superelastic behaviour on NiTi thin-walled tube under biaxial loading,". *Computational Materials Science*, 40(2):246–254, 2007.

[101] W. Wu, F. Gordaninejad, and R. A. Wirtz. "Modeling and Analysis of a Shape Memory Alloy–Elastomer Composite Actuator,". *Journal of Intelligent Material Systems and Structures*, 7:441–447, July 1996.

[102] X. C. Wu, D. S. Grummon, and T. J. Pence. "Modeling phase fraction shakedown during thermomechanical cycling of shape memory materials,". *Materials Science and Engineering A – Structural Materials Properties, Microstructure and Processing*, 275:245–250, Dec 15 1999.

[103] PJ Yang, YF Zhang, MZ Ge, TD CAI, JC Tao, and HP Yang. “Internal fixation with Ni-Ti shape memory alloy compressive staples in orthopedic surgery. A review of 51 cases.,”. *China Medical Journal (English)*, 100:712–714, 1987.