SPRINGER BRIEFS IN APPLIED SCIENCES AND TECHNOLOGY · COMPUTATIONAL MECHANICS

Ashwin Rao
A.R. Srinivasa
J.N. Reddy

Design of Shape Memory Alloy (SMA) Actuators

Springer

SpringerBriefs in Applied Sciences and Technology

Computational Mechanics

More information about this series at http://www.springer.com/series/8886

Ashwin Rao · A.R. Srinivasa
J.N. Reddy

Design of Shape Memory Alloy (SMA) Actuators

Springer

Ashwin Rao
Department of Mechanical Engineering
Texas A&M University
College Station, TX
USA

J.N. Reddy
Department of Mechanical Engineering
Texas A&M University
College Station, TX
USA

A.R. Srinivasa
Department of Mechanical Engineering
Texas A&M University
College Station, TX
USA

ISSN 2191-530X ISSN 2191-5318 (electronic)
SpringerBriefs in Applied Sciences and Technology
ISSN 2191-5342 ISSN 2191-5350 (electronic)
SpringerBriefs in Computational Mechanics
ISBN 978-3-319-03187-3 ISBN 978-3-319-03188-0 (eBook)
DOI 10.1007/978-3-319-03188-0

Library of Congress Control Number: 2015938741

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

This book is dedicated to my parents Raghavendra Rao and Jayashree R. Rao, who have raised me, motivated me, and have guided me through all ups and downs of my life.

Ashwin Rao

This book is dedicated to my parents Ramaswamy and Saroja and to my wife Prabha and children Vishnu and Divya, who sacrificed their time so that I could finish this book.

Arun Srinivasa

This book is dedicated to my baava, Tummala Janardhan Reddy, a constant admirer of the highest evidences of the human spirit, for his friendship and for introducing me to "rational thinking."

J.N. Reddy

Preface

Shape memory alloys have been fascinating to designers, architects, and researchers in the past decade. There is something about the uncanny ability of seemingly inanimate wires suddenly reacting to external nonmechanical stimulus that evokes curiosity and childlike fascination in everyone. However, commercial applications (other than in the medical field) has been slow. Part of the reason is the lack of accessible explanations that allow people with only basic exposure to such materials to carry out designs that are viable.

Too often, papers and books written (many by the authors themselves) about arcane aspects of SMA behavior are not meant for designers. This leads to exasperation from a designer who wants us to "tell me how do I design with this?"

This book seeks to provide an accessible account of SMA behavior together with examples of preliminary design methodology to students with a basic undergraduate background. The aim is to provide an "on ramp" to explore the unique properties of these devices, and so the book only deals with the "bare necessities" and ignores many nuances including important issues of functional fatigue. Rather the design recommendations are based on being conservative and making design decisions that will eliminate the need for considering such issues at the expense of not being optimal. Our philosophy in designing with SMA is "robust, repeatable, guaranteed behavior" over "optimal" response.

This monograph is by no means extensive but just an introduction and an invitation to the readers to explore the behavior of these materials. It grew out of a National Science Foundation Grant to develop "strength of materials-like" approaches to shape memory wires and springs. While the current applications at the cutting edge have moved on to tubes, plates, and so on, the commercial availability of wires and springs for common devices has grown quite a bit and this is what we wish to emphasize.

If a reader gains a qualitative understanding of SMA response together with the ability to do a "first-cut" design by reading this book and is able to go on to explore SMA better, then we have succeeded in achieving our purpose of writing this book.

College Station, Texas

Ashwin Rao
A.R. Srinivasa
J.N. Reddy

Acknowledgments

The authors acknowledge the help of a large group of people in creating this book. In particular, the authors acknowledge the invaluable help, advice, discussions, and guidance of Dr. S. Chandrasekaran (SMA specialist and visiting Scientist, IIT Madras), Prof. Sivakumar Srinivasan (Department of Mechanical Engineering, IIT Madras), and Dr. Vidya Shankar Buravalla (GE Global Research) in the writing of this book. The authors also gratefully acknowledge the support of the research by the National Science Foundation through research Grant No. CMMI-1000790 (with A.R. Srinivasa and J.N. Reddy as co-PIs) to Texas A&M University, which inspired them to undertake the writing of this book.

Ashwin Rao thanks his colleagues at Evident Thermoelectrics for their support and encouragement during the completion of this book.

Contents

Chapter 1
Introduction to Shape Memory Alloys

1.1 Smart Materials–An Overview

Classical materials like metals and alloys have played a significant role as structural materials for many centuries [1]. Engineers have designed components and selected alloys by employing the classical engineering approach of understanding the macroscopic properties of the material and selecting the appropriate one to match the desired functionality based on the application [2]. With advancements in material science and with increasing space and logistical limitations, scientists have been constantly developing high performing materials for various applications [2]. The everlasting goal for engineers in many cases is to improve product efficiency and reduce its weight without comprising on either its cost or performance. To achieve this goal, replacing multi-component and multi-material systems with fewer multi-functional light weight, high performing materials has been an attractive alternative [2, 3]. Such advanced materials have played a leading role in the development of many engineering innovations and achievements like the Airbus A380, Boeing 787 Dreamliner, reliable fuel efficient cars, superior drug delivery devices, and so on, to list a few. The ingenious commercial products across various engineering disciplines are meeting all requirements by encompassing many of the latest technologies and meeting the challenges of tomorrow's needs.

In pursuit of this, material scientists over the last few decades have focused on the possibility of tailoring the microstructure of the material to generate the required functionality for different applications [2, 4]. Such an effort has resulted in an entire new area of **active or multifunctional materials** that posses more than one desirable property [5, 6]. With the introduction of such materials, researchers are now focusing on how the combined microstructural changes of such materials are able to perform multiple functions. The integration of multiple functions like actuation, sensing, and control into a single structure using one or more material constituent is seen as a possibility [2, 3]. Mamoda in her recent review of future materials discusses some application ideas with such materials like: a smart solar panel that can change its orientation automatically during the day depending on sun's position; a smart shock

A. Rao et al., *Design of Shape Memory Alloy (SMA) Actuators*,
SpringerBriefs in Computational Mechanics, DOI 10.1007/978-3-319-03188-0_1

Table 1.1 Multiphysics coupling

Input energy	Output energy			
	Mechanical	Magnetic	Thermal	Electrical
Mechanical		Magnetostriction		Piezoelectric, piezoresistive
Magnetic	Magnetostriction		Magnetocaloric	Magnetoresistance
Thermal	thermomechanical		Pyroelectric	Thermoelectric
Electrical	Electrostrictive, piezoelectric, ER fluid		Thermoelectric	

absorber that can alter its damping ability based on the road profile; morphable wings and blades for aircraft maneuvering during flight; a coating that changes color on demand, to list a few examples [3, 6].

Smart materials are a subgroup of such active/multifunctional materials that show a capability of recognizing a non-mechanical external stimuli from its surrounding environment and reversibly respond to the same [2]. Such materials can judge the magnitude of this external stimuli (signal) and react with an optimal response by either changing its physical or mechanical properties (generally macroscopic shape change) [2, 3, 7]. The non-mechanical stimuli could be in the form of changes in temperature, magnetic field, electric potential, light intensity, moisture, changes in pH (chemical stimuli), and others [8]. Table 1.1 and Fig. 1.1 illustrate the various possible couplings between the the input signal (i.e. magnitude of external stimuli) that the material is exposed to and its corresponding physical response for the same. For example, magnetostrictive phenomenon involves changes in magnetic fields that result corresponding changes in mechanical properties. Table 1.2 provides a summary of some commonly used smart materials with their corresponding input signals that show such multiphysics coupling.

1.2 Smart Structures—System Level Response

The combination of one or more smart materials can lead to a system level design that can combine multiple functions in a single material system. With such a design one can reduce system complexity and thus lowering system weight/volume [3]. To illustrate this further, two examples are considered in Fig. 1.2. Lord Corporation has developed controllable Magnetorheological (MR) dampers that can smartly adapt to the terrain on which the vehicle is travelling with the goal of improving the overall passenger ride comfort [11, 12]. This is made possible with the use of a MR fluid[1] that can change their viscosity under the action of external magnetic fields

[1] A typical MR fluid consists of carbonyl iron particles in oil, petroleum products [13].

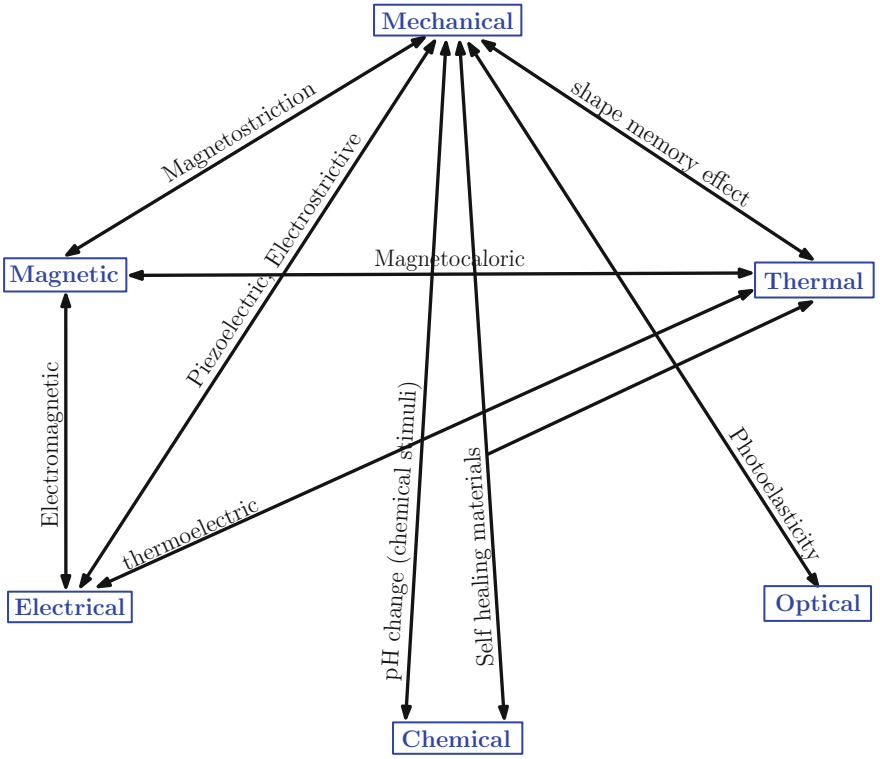

Fig. 1.1 Smart materials can involve multi-physics coupling based on the external stimuli it is subjected to which results in changes of physical/mechanical properties, that is, generally macroscopic shape change (adapted from [9, 10])

[4, 14]. The ability of such materials to change their viscosities real time based on the terrain makes them better alternatives compared to conventional dampers employing standard SAE fluids. Such MR dampers have replaced the classic passive dampers in many off-road and high end vehicles and are even finding applications as seismic dampers in earthquake resistant structures [11, 15].

Biodegradable shape memory polymers (SMP)[2] have found many biomedical applications in the form smart stents or sutures. The sutures used are programmed to shrink when heated above their *glass transition temperature* T_g. In order to heal a wound, a SMP fiber is loosely tied across the wound and then heated (above their T_g) for tightening the knot with an optimal force. Since the SMPs are programmed to shrink upon heating, they remain as a knot till the wound is fully healed.

[2]Thermally responsive SMP's have shown the ability to recover large deformations when subjected to external thermal stimuli. Above the glass transition temperature T_g, there is an onset of long range molecular motion i.e. it changes from a glassy solid phase to an unordered rubbery phase and vice versa [2]. During this process, SMP's can exhibit large changes in shape and moduli which can used for many applications [2].

Table 1.2 Commonly used smart materials with their corresponding input signals (i.e., the external stimuli it is exposed to)

Smart material	Input signal
Shape memory alloys (SMA)	Temperature/magnetic field
Shape memory polymers (SMP)	Temperature/moisture/light/pH
Ionic polymer metal composite (IPMC)	Electric field
Mangnetorheological fluids (MR fluids)	Magnetic field
Electrorheological fluids (ER fluids)	Electric field
Dielectric elastomer (DE)	Electric field

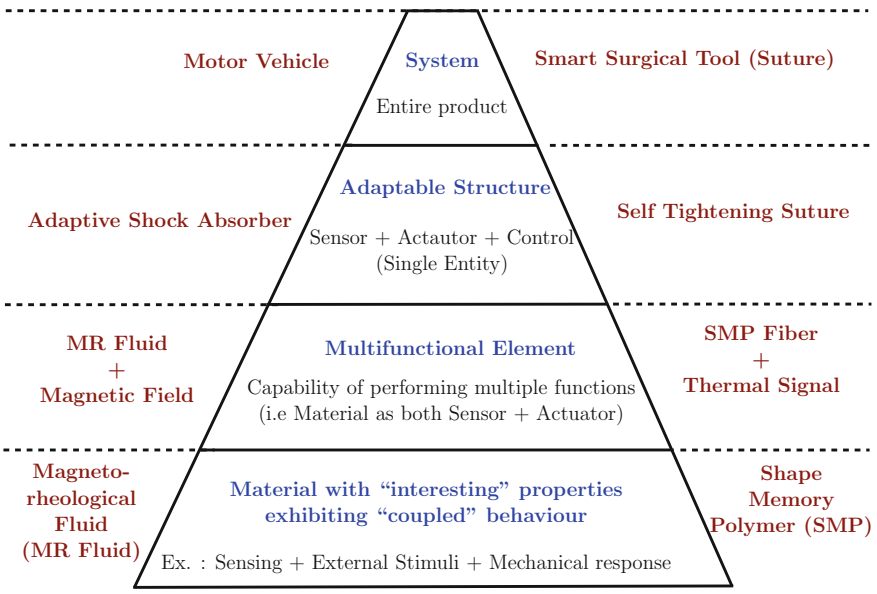

Fig. 1.2 Integration of multiple functions like actuation, sensing, and control into a single structure using one or more smart material constituents (adapted from [4])

In both of these examples, either the MR fluid or biodegradable SMP acts both as a sensor and actuator with its ability to recognize non-mechanical stimulus (like changes in temperature, pH, moisture, magnetic or electric fields etc.) from its surrounding environment and subsequently change its mechanical properties as desired for the application [2]. This integration of multiple functions like actuation, sensing and control into a single structure has provided a system level design with the use of a single material system. These materials particularly have the potential to completely revolutionize the design of a wide variety of devices in applications ranging in areas from biomedical, automotive, aerospace, civil engineering to energy harvesting [16].

1.3 Shape Memory Alloys: Temperature Induced Phase Transformations

Shape memory alloys (SMAs) are a subset of a broad class of smart materials where the functionalities arise from their underlying microstructural changes when subjected to external non-mechanical stimuli like temperature or magnetic field changes [17, 18]. In thermally responsive SMAs, the reversible solid-solid, diffusionless[3] thermoelastic phase transformations between a stable high temperature austenitic phase and low temperature martensitic phase are responsible for them to demonstrate interesting phenomenon like **shape memory effect (SME)** and **superelasticity (SE)**. Their capability to return to a predetermined shape on heating is referred to as the shape memory effect (SME). Their ability to recover large strains ($\sim$8 %) and associated large stress–strain hysteresis due to mechanical loading–unloading under isothermal conditions is referred to as superelasticity/pseudoelasticity (SE) [18].

The phase transformations are a result of shear lattice distortions (twinning) rather than long range diffusion of atoms [1, 19]. The austenitic phase in SMA has a ordered B2 cubic crystal structure as compared to the martensitic phase which has either tetragonal, orthorhombic or monoclinic structures [20]. Due to its non-cubic structure, the martensitic phase (B19$'$ crystal structure) can have different orientations (variants)[4] [1, 19]:

- "self-accommodated martensite variants" (also referred to as twinned martensite) or
- "single variant martensite" (also referred to as detwinnned or reoriented martensite) with a specific variant being dominant.

Figure 1.3 shows the cubic structure of austenite and different variants of martensite. Under stress free conditions, the martensite in SMA's are usually self-accommodating in nature [22]. All of these phase transformations occur over characteristic transformation temperatures namely martensitic start (M_s), martensitic finish (M_f), austenitic start (A_s) and austenitic finish (A_f). ASTM standard F2004-05R10 discusses the details of measuring these transformation temperatures using a "Differential Scanning Calorimetry (DSC)" test for a Ni-Ti SMA [24]. The test uses small, stress-free annealed SMA samples along with a reference sample (empty Aluminium pans in some cases) in a DSC setup that is automatically capable

[3]The word diffusionless suggests that the atoms do not leave their lattice positions and there is no "long range diffusion" of atoms or species. Some literature on SMA commonly refer the twinning events in phase transformations as "short range diffusion" where the atomic displacements are less than the average interatomic distances.

[4]Phase transformations in SMA (responsible for their functionality) under both SE and SME are between austenite and single variant martensite variants. The "self-accommodated martensite" is generally a combination of single variant martensite species [21]. For example, under tension and compression loading cases, two unique single variant martensite species (say M^t and M^c for tension and compression loading cases) can exist. In such a case, the "self-accommodated martensite" version would be 50 % M^t and 50 % M^c.

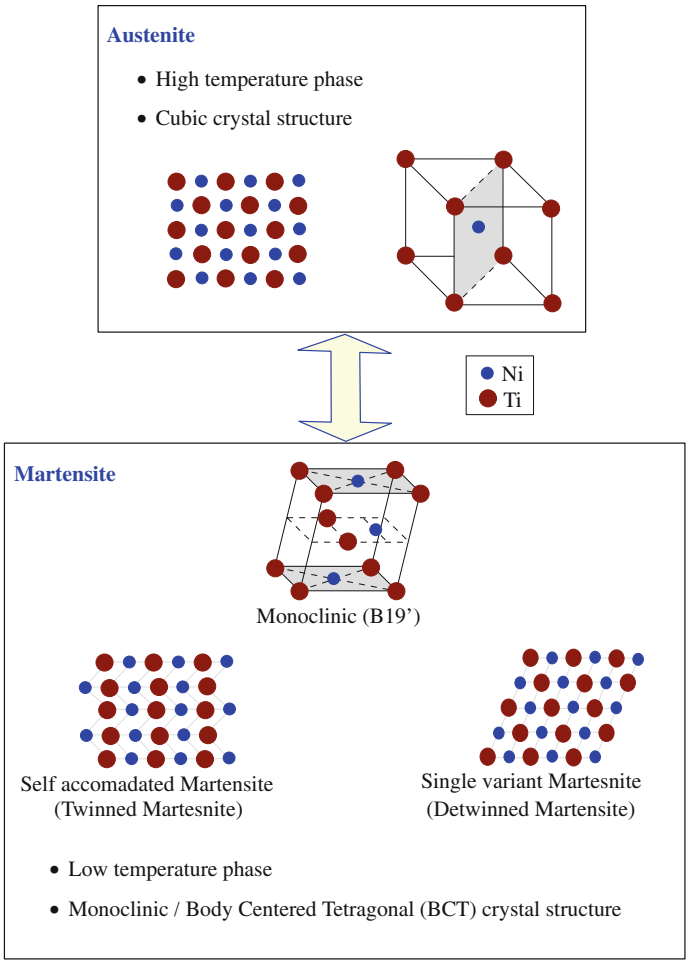

Fig. 1.3 The underlying microstructural changes in shape memory alloys during thermoelastic phase transformations occur between a stable high temperature austenitic phase and low temperature martensitic phase. The non-cubic martensite phase can exist in different orientations (variants). (adapted from [1, 23])

of heating and cooling the samples up to 10 °C/min [24]. The DSC setup automatically records the differential energy inputs between the specimen and reference by monitoring the heat flow rates. The DSC setup automatically generates a heat flow $\left(\frac{J}{gK}\right)$ versus temperature (in °C) [24]. Figure 1.4 shows an ideal DSC plot with the tangents drawn on the heating and cooling spikes to determine the four characteristic transformation temperatures. The peaks observed in the plot are the latent heat of transformations due to austenite–martensite phase transformations in SMA. These endothermic and exothermic peaks are characteristic to any first order

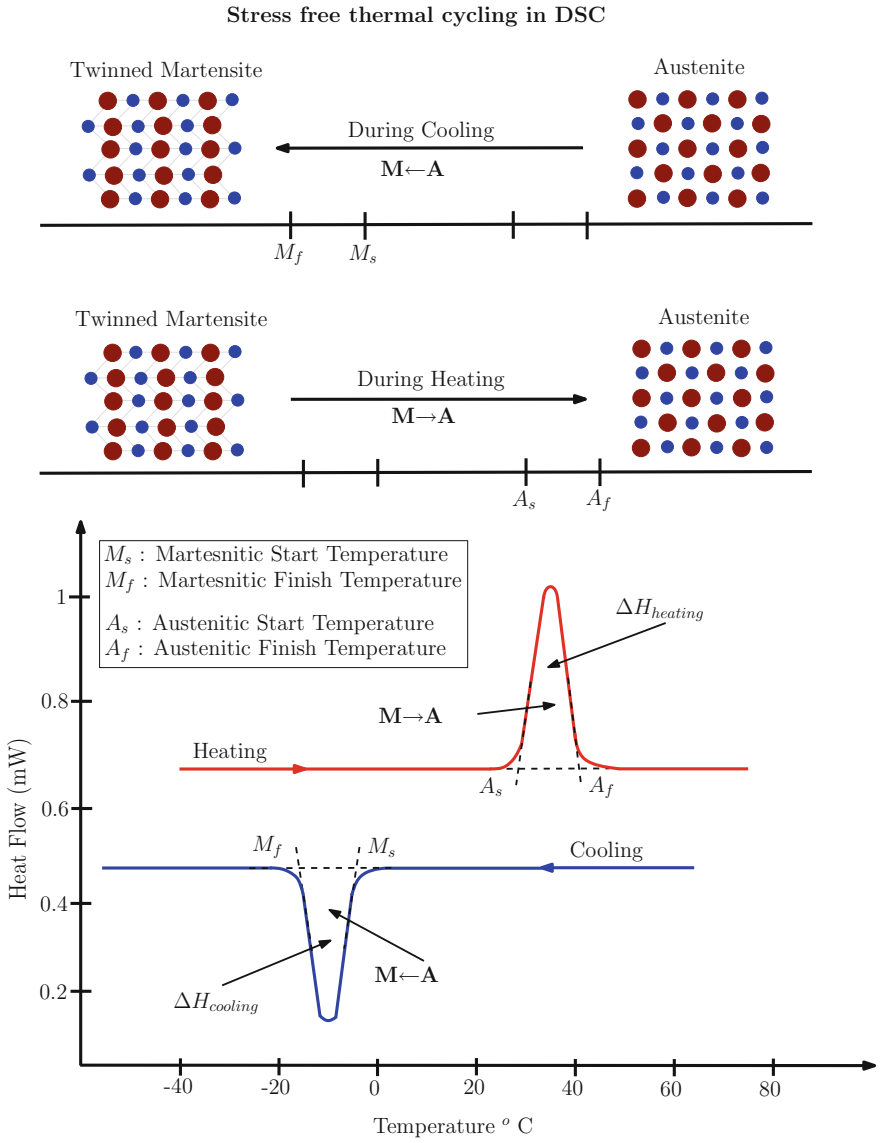

Fig. 1.4 Crystal structures of self-accommodated martensite and austenite along with associated forward and backward transformation cases. Transformation temperatures martensitic start (M_s), martensitic finish (M_f), austenitic start (A_s) and austenitic finish (A_f) are obtained from a standard DSC test as per ASTM standard F2004-05R10 (adapted from [1, 24])

phase transformations like melting of ice (solid–liquid), boiling of water (liquid–gas) etc. [23]. However, in case of SMA's, these phase transformations are solid–solid (austenite–martensite) in nature.

1.4 Shape Memory Effect and Superelasticity/Pseudoelasticity

The ability of SMA to return to a predetermined shape on heating above the characteristic transformation temperature A_f is referred to as the **shape memory effect (SME)**. A typical SME effect in SMA is depicted as path ①–⑥ on stress–strain–temperature diagram as shown in Fig. 1.5. The various paths in SME are described below.

- Path ①–②: Under stress-free conditions, the self accommodated martensite variant favourably align themselves resulting in deformation at lower stresses without any change in volume. This part of the response looks like a classic elastic deformation of metal or alloy specimen.
- Path ②–③: With higher loads, the self accommodated martensitic variants twins into stress preferred single variant martensite (also called detwinned martensite) typically associated with large (∼6–8 %) macroscopic strains and shape changes [22]. These macroscopic strains are also referred to as martensite plateaus or the twinning response of SMA purely due to mechanical loading. If deformed further then the single variant martensite under higher stresses can undergo permanent plastic deformation due to slip which cant be recovered.
- Path ③–④: Unloading from point ③ results in elastic unloading of detwinned martensite as shown in point ④ [25]. It must be noted that the material does not completely return back to its undeformed state ① and there is a clear residual strain of about 4–6 % depending on the extent to which it was loaded. This residual strains seems like classic permanent unrecoverable strains observed when metals or alloys which are plastically deformed when beyond their yield point. However, in case of SMA, this residual deformation can be recovered back upon heating and thus making them a unique compared to classical materials.
- Path ④–⑤–⑥: Upon heating above A_f, residual strains are completely recovered as the low symmetric martensitic phase (M) is transformed back to stable austenitic phase (A). This complete strain recovery upon heating is referred to as shape memory effect [22]. During the heating event, as the temperature reaches A_s (point ⑤), the SMA starts transforming back to austenite and the transformation is complete when the temperature reaches or exceeds A_f, that is, point ⑥.
- Path ⑥–①: At point ⑥, if the SMA is cooled below M_f in absence of external loads, austenite transforms back to self-accommodating martensite twins, that is, point ① without any volume change. This event indicates temperature induced phase transformations in absence of any external loads. This event is commonly observed as the cooling peak (latent heat of transformation) in a DSC test. It must be highlight that this particular stress free transformation in quite rare in SMA applications as the SMA's are subjected to some form of external loads at all times.

A simple illustration of SMA wire under external load demonstrating SME is shown in Fig. 1.6. In this illustration, SMA wire is in its martensitic state A is

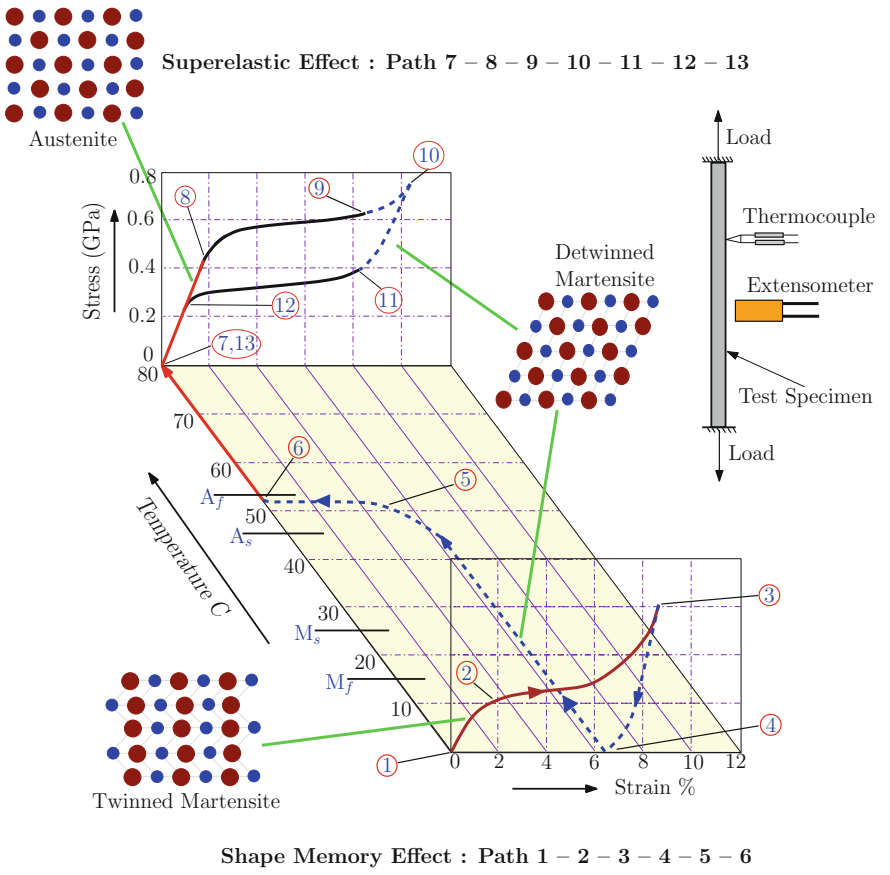

Fig. 1.5 A three-dimensional stress–strain–temperature plot describing the thermomechanical experiment showing both shape memory effect (path 1–6) and superelastic effects (path 7–13). A description of these states are discussed in Sect. 1.4. These reversible effects are manifestations of solid–solid phase transformations between a stable high temperature austenitic phase and low temperature martensitic phase (adapted from [23])

deformed to state B under external load. Upon heating above temperatures A_f, the SMA wire contracts back to its austenitic state denoted by state C or hot state position. Upon cooling, the SMA wire reaches state D. Cycle A↔D is termed as shape memory effect (SME).

On the other hand, these martensitic transformations can be induced purely due to mechanical loading–unloading in the austenitic phase with SMA. The ability of SMA to recover large strains (∼8 %) with associated stress–strain hysteresis due to mechanical loading–unloading under isothermal conditions is referred to as **superelastic/pseudoelastic effect (SE)** [22]. These effects are observed at temperatures

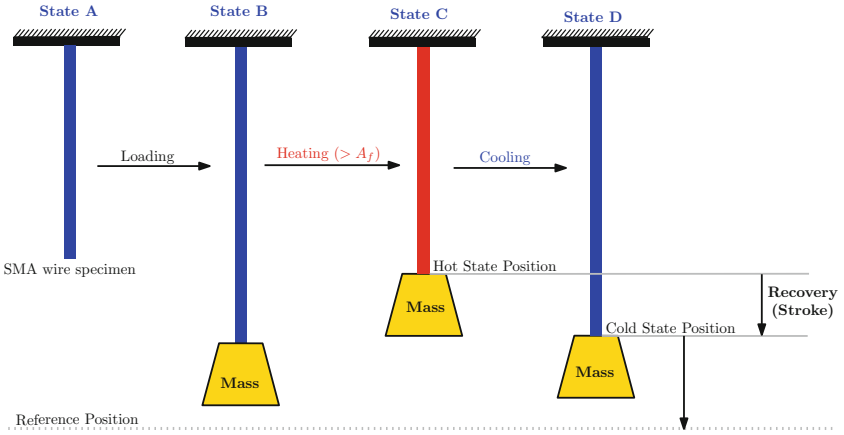

Fig. 1.6 A simple example of SMA wire in its martensitic state A is deformed to state B under external load. Upon heating above temperatures A_f, the SMA wire contracts back to its austenitic state denoted by state C or hot state position. Upon cooling, the SMA wire reaches state D. Complete cycle A↔D is termed as shape memory effect (SME)

greater than A_f as shown by path ⑦–⑬ in Fig. 1.5. The various paths are described below.

- Path ⑦–⑧: Above A_f, under moderate external loads, the austenite elastically deforms which is represented by path ⑦–⑧. This looks like a classic elastic deformation of metal or alloy specimen.
- Path ⑧–⑨: Upon continued loading, at some critical stress, it is thermodynamically stable for austenite to start transforming to a single variant martensite (also commonly refereed to stress induced martensite (SIM) to differentiate this martensite variant from that formed due to twinning in SME). This forward transformation from a cubic to monoclinic crystal structure results in large macroscopic deformation of about 6–8 % strains. In a classic superelastic response under displacement controlled experiment, this event occurs at relatively constant stress and both austenite and martensite phases coexist during this event. Towards the end of the transformation, most of the austenite has transformed to SIM. However, in reality some pockets of austenite do not transform to martensite.
- Path ⑨–⑩: Upon complete transformation (i.e., the end of flat plateau region), if loaded further, this results in elastic deformation of SIM. However, higher loads can cause plastic deformation of SIM and they cannot be recovered back completely. In real world applications, this part of the response is rarely used in designing SMA components due to poor fatigue life and thus the maximum strains in SMA's are restricted to about 6 % strains (i.e., up to point ⑧).
- Path ⑩–⑪: Upon unloading from point ⑪ with the assumption that the stress level at ⑪ did not induce any permanent plastic deformation, the SIM unloads

until point ⑩ at which it is thermodynamically stable for it to transform back to austenite.

- Path ⑪–⑫: This represents the reverse transformation from SIM to austenite. Similar to path ⑧–⑨, it is accompanied by large macroscopic strains (in this case its shortening) until it transforms completely back to austenite by time it reaches ⑫. This flat plateau is particular important for deigning many biomedical applications that use SMA (details available in Sect. 1.6.1)
- Path ⑫–⑬: In an ideal superelastic response, this path overlaps on path ⑦–⑧ and it represents the elastic unloading of the austenite. However, in most superelastic responses, there is generally some residual strains and points ⑬ and ⑦ do not generally overlap. The residual strains could be attributed to pockets of austenite or SIM that do not transform back during either the forward or backward transformations. During transformation regions ⑧–⑨ or ⑪–⑫, the deformation is highly nonhomogeneous as both the phases coexist which might results some pockets of SMA from not transforming from austenite to SIM or vice-versa. Further, higher stress levels during the path ⑨–⑩ can also results in residual strains due to possible plastic deformation of SIM.

An ASTM standard F2516-07ϵ2 [26] discusses a standard tension test methodology on NiTi superelastic materials and explains the salient features of such superelastic effects.[5] Some important salient features worth noting in a typical superelastic response (see Fig. 1.7) are outlined here.

- **Upper plateau strength (UPS)**: The stress at 3 % strain during loading of the sample.
- **Lower plateau strength (LPS)**: The stress at 2.5 % strain during unloading of sample after loading to 6 %.
- **Residual Elongation ($\mathbf{El}_r$)**: The difference between strains at a stress of 7 MPa during loading and unloading operations.
- **Uniform Elongation**: Elongation determined at maximum force sustained by specimen prior to necking or fracture or both.

The phase transformations in the superelastic response of SMA results in nonlinear hysteretic response which make them excellent candidates as damping materials

[5]Commercially, the NiTi SMA's are available as actuator wires or superelastic wires. In case of actuator wires, the SMA is martensitic at room temperature, that is, the M_f is slightly above the room temperature and the A_f may be around 60 to 100 °C depending on the alloy composition and other material processing conditions. Such actuator wires under external mechanical loads at room temperature causes the self accommodated twins to detwin into more stress preferred martensite variants and thus demonstrating macroscopic shape change. Such wires would show superelastic behaviour above A_f. On the other hand, the superelastic wires are austenitic at room temperature, that is, its A_f is below room temperature and its M_f is far below sub zero temperatures (around -60 to 100 °C again depending on the alloy composition and other material processing conditions). Such superelastic wires will show SE behaviour at room temperature and demonstrate martensite detwinning at temperatures below zero (i.e., below M_f). A more detailed discussion on the effect of composition and other material processing conditions will be taken up in future chapters of this book.

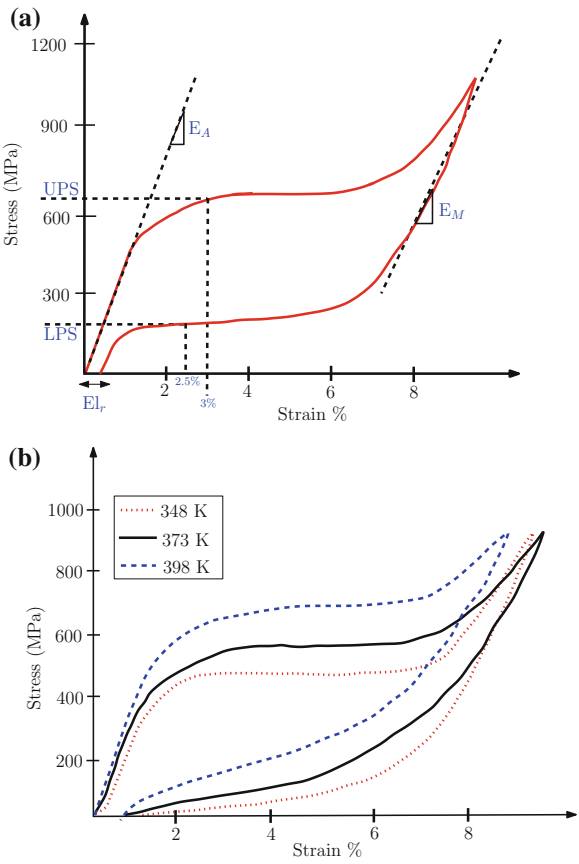

Fig. 1.7 **a** Shows a typical superelastic response with internal loops of a NiTi SMA with some salient features like upper plateau strength (UPS), lower plateau strength (LPS), residual elongation (El_r) highlighted as reported in ASTM standard F2516-07ϵ [26]. In addition, the austenitic and martensitic moduli E_A and E_M are shown. **b** Shows the temperature dependence of the superelastic response at three different temperatures

due to their superior energy dissipation capabilities [27]. The ability of SMA to deliver large plateau strains over relatively constant forces/stresses makes them good candidates as actuators [28]. Due to different crystal structures of austenite and martensite, SMA's have different moduli over various parts of the superelastic response [see Fig. 1.7a]. Further, the superelastic responses are also sensitive to the external stimuli, that is, operating temperature as shown in Fig. 1.7b. With increasing temperatures, the critical stresses for transformation get higher and thus making the material response sensitive to temperature fluctuations.

1.5 Commonly Used Shape Memory Alloys

The martensitic phase transformations were initially observed in many steels (in a Iron–carbon phase diagram) during the early 20th century. However, the reversible martensitc phase transformations were underutilized until the discovery of SMA in

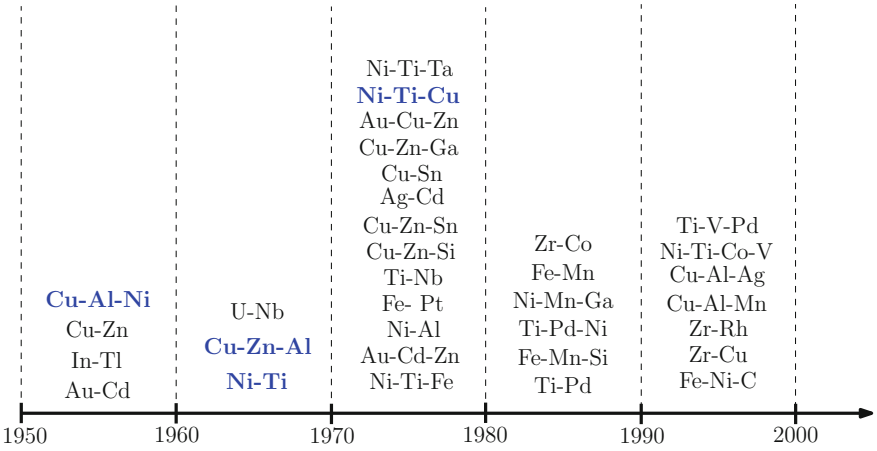

Fig. 1.8 Various shape memory alloys discovered over the second half of 20th century. The most important commercially used alloys are highlighted in *blue*. (adapted from Wayman and Duerig [29] and Shaw et al. [23])

early 1900s [1]. The SME in Au-Cd alloy was first reported in 1950s by Chang and Read [29]. Since then, several other alloy systems (intermetallics) demonstrating SME were discovered and many of them are summarized in Fig. 1.8 [23, 29]. Nearly equiatomic Ni-Ti, Cu-Al-Ni, Cu-Zn-Al and Ni-Ti-Cu alloys are some of the commercially exploited SMA in different applications. The most popular among these is the Ni-Ti alloy that was discovered by William J. Buehler and coworkers from the US Naval Ordnance Laboratory in White Oak, Maryland in 1963 and is often referred to as NiTiNOL in its honour [1, 29]. A comparison of some important properties of three commercially available SMA's with 316 stainless steel is tabulated in Table 1.3. In this book, most of the discussion will be restricted to NiTiNOL and Cu-Zn-Al alloys.

Some High temperature SMAs (HTSMAs) like Pt-Ti, Ni-Ti-Pt, Ni-Ti-Hf, Ni-Ti-Pd, Ni-Mn-Pt, Ni-Mn-Al, Ti-Ni-Zr, Cu-Al-Ni-Mn-Ti, Ru-Ta, Ru-Nb, and so on are also under development where the transformation temperatures vary from 200 °C to as high as 700 °C [35, 36]. Some ferromagnetic and metamagnetic SMAs like Fe-Ni-Co-Ti, Fe-Ni-Co-Al, Fe-Ni-Mn-Al, Co-Ni-Al, Co-Ni-Ga, Ni-Mn-Ga, Ni-Mn-In, Ni-Mn-Co, Ni-Mn-Sn, Ni-Fe-Ga, etc. are some alloys under investigation due to their ability to work at higher frequencies [36]. However, further work needs to be done for better understanding of underlying microstructural changes, operational limits, creep and fatigue studies for HTSMAs, ferromagnetic and metamagnetic SMAs [35, 36].

Table 1.3 Comparison of some important properties of three commercially available SMA's with 316 stainless steel (data complied from [30–34]

Properties	**NiTi**	**CuZnAl**	**CuAlNi**	**316 Stainless steel**
Specific heat (J/Kg °C)	450–620	390–400	373–574	500
Thermal conductivity (20 °C) $\left(\frac{W}{mK}\right)$	8.6–18	84–120	30–75	15–22
Density $\left(\frac{g}{m^3}\right)$	6.4–6.5	7.54–8	7.1–7.2	7.8–8
Latent heat (kJ/Kg)	19–32	7–9	70–9	26
Electrical resistivity (10^{-6} Ωm)	0.5–1.1	0.07–0.12	0.1–0.14	0.74
Thermal expansion coefficient $\left(\frac{10^{-6}}{K}\right)$	6.6–11	17	17	17.3
Maximum recovery stress (MPa)	500–900	400–700	300–600	- na -
Normal working stress (MPa)	100–130	40	70	- na -
Maximum transformation strain (%)	6–8	4–6	5–6	2 (elastic)
Young's Modulus (GPa) (M and A)	28–83	70–100	80–100	197
Transformation temperature range (°C)	(–200) to 200	(–200) to 150	(–200) to 200	- na -
Maximum overheating temperature (°C)	400	150	300	700
Melting, casting and composition control	Difficult	Fair	Fair	Easy
Forming (rolling, extrusion)	Difficult	Easy	Difficult	Easy
Cold-working	Fair	Restricted	Difficult	Easy
Machinability	Difficult	Very good	Good	Easy

1.6 SMA Applications: Overview

The ability of SMA to reversibly respond to external temperature changes and change their physical/mechanical properties has enabled them to find many applications. In a thermomechanical system, SMAs can be used as combined sensors and actuators where they can sense the changes in external stimuli and monitor certain desired functions [37]. The unique characteristics of SME and SE discussed earlier have made SMAs the material system of choice in applications ranging from sensing and control, vibration damping, biomedical, automotive and aerospace areas [16, 37–40]. A review of many SMA devices in use across many engineering applications is detailed in the following sections.

1.6.1 Biomedical Applications

In particular, NiTi SMAs have found many biomedical applications due to their excellent biocompatibility.[6] Studies have shown that Nickel on its own is quite toxic and any contact with nickel can lead to various medical complications [41]. However, in case of intermetallic NiTi alloys, the bonding between Ni and Ti is quite strong (like in ceramic materials) as compared to the Nickel bonding in steel and other materials [43]. Further, the commercial NiTi alloys uses a passive TiO_2 (titanium oxide) layer coating on its outer surface that prevents any nickel leakage as it acts as a physical and chemical barrier in preventing Ni oxidation [43, 44]. The TiO_2 layer is harmless to human body and provides high resistance towards corrosion of Ti alloys [41]. Many clinical studies have shown minimal Nickel contamination due to use of NiTi SMAs and thus showing good biocompatibility [44].

In addition, NiTi SMA's show excellent MRI compatibility, kink resistance, corrosion resistance and substantial moduli differences between austenitic and martensitic phases. All of these properties make SMAs a good choice for many biomedical applications like: drug delivery systems, self-expanding stents, stent delivery systems, implantable devices, catheters, guide-wires, atrial occlusion devices, and thrombectomy devices [35, 41, 43–48]. In most of these biomedical applications, transformation temperatures of the SMA are programmed such that the A_f is below the body temperature [35]. The superelastic SMA components are cooled to their martensitic state (i.e., below M_f) and deformed to a temporary shape for easy insertion. Upon

[6]Biocompatibility is dependent on the allergic reactions between the foreign material and the host [41]. Material characteristics, patients health's and several other factors play an important role for assessing biocompatibility. In case of SMAs, several clinical studies were performed to establish Ni–Ti alloys biocompatibility and FDA® has cleared many SMA products for medical use [41]. FDA® is a registered trademark of "The Food and Drug Administration (FDA or USFDA)" affiliated with the United States Department of Health and Human Services (an United States' Federal executive department) [42].

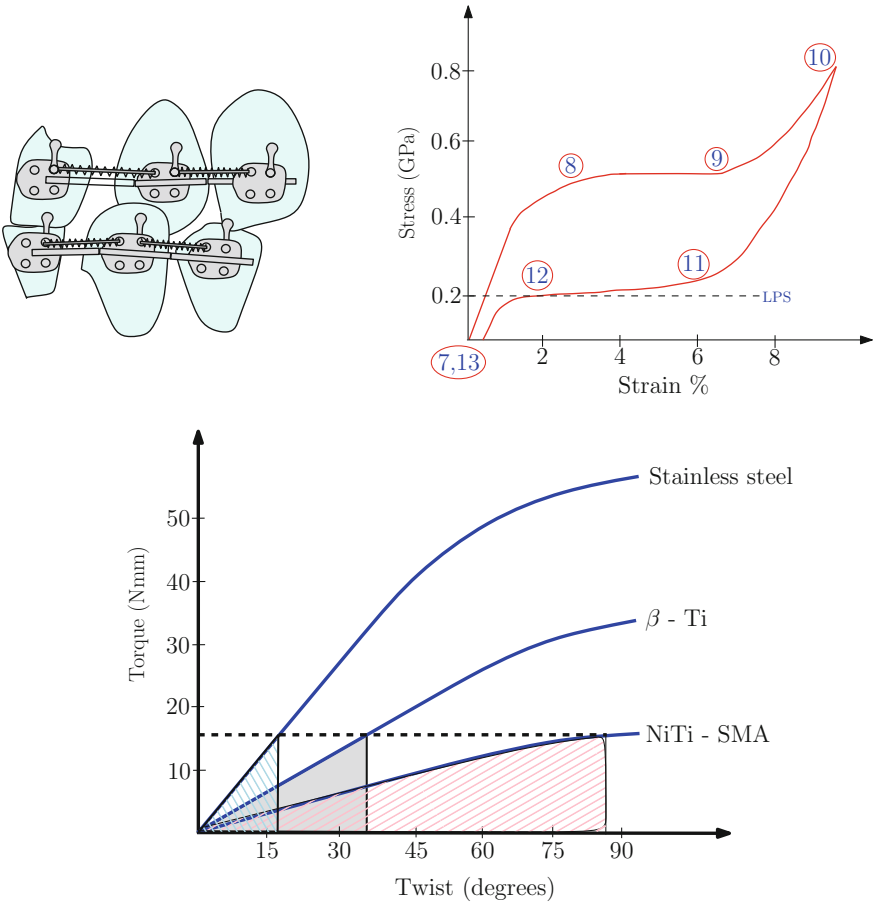

Fig. 1.9 A SMA torsional spring used for tooth movement, space closure. The idea is to design the spring at the LPS so that it can deliver constant forces over large strokes (Figure adapted from [58]. Comparison graph adapted from [54])

deployment at the right location in the body, the SMA device is heated above A_f (i.e., austenitic state) where the NiTi component recovers back to its original shape and performs the necessary function as desired. Some specific examples are discussed below to illustrate this point.

SMA springs, wires and braces in many orthodontic applications are designed in the LPS region so that they can deliver relatively constant forces over large activation strokes as shown in Fig. 1.9. Their ability to deliver constant forces are commonly employed for space closure and tooth movement in many orthodontic applications [49–51]. Further, based on various test results, researchers have suggested that SMAs provide superior spring-back properties, large recoverable strains thus making them better alternatives when compared against its counterparts like stainless steel,

β-Ti, Co-Cr for medical applications [52–54]. Though the stainless steel counterparts can deliver higher forces, however their force delivery rapidly decays over time as compared to SMA springs for relatively long activation ranges as shown in Fig. 1.9 [55, 56]. In many of these dental applications, by choosing suitable wire diameters of SMA, the required force can be varied based on application [57].

SMAs have found applications in many "minimal invasive surgery applications" as they can easily pass through convoluted paths and reach various parts of human body [43]. NiTi wires, tubes that demonstrate good kink resistance[7] are used to maneuver complex paths and still remain in its original shape at the time of delivery [43]. Similarly, several occlusion systems uses catheters that generally consist of NiTi wire loops on small umbrellas that support "microporous polyurethene webs" at its ends [44]. The use of SMA here provides optimal occlusion in areas of limited surgical exposure and access as the SMA loops spring back to their original shape upon release of external loads. Several scissors, biopsy forceps, graspers and other endoscopic instruments use NiTi rods for actuation as they can undergo large bending deformations without any permanent deformation (no kinks or buckling) [44].

SMAs are used in the neurosurgical field as coils, stents or microguidewires mainly to treat cerebral aneurysms [59]. To prevent aneurysm rupture, coils are positioned into the aneurysm to facilitate "clotting or thrombotic reaction within the aneurysm" [59]. The superelastic effects of SMA are utilized here which allows large deformations and prevent crushing of the coil. Microguidewires made of NiTi are employed for stent positioning due to superior kink resistance and flexibility [44, 59]. Similarly, with the goal of trapping blood clots and dissolving them, many cardiovascular devices have been developed to tackle pulmonary embolism and one of the first SMA filter called "Simon Filter" was developed [41, 60].

Self expanding stents named in the honour of dentist C.T. Stent find applications in cardiovascular applications with the goal of preventing collapse of blood vessels [41, 45]. As shown in Fig. 1.10, NiTinol Stents are shape set in their deployed configuration (generally expanded diameter in its austenitic state) and then compressed into a catheter at lower temperatures below M_f. Stents employed in arteries may be subjected to continuously varying external pressures and collapsing or crushing of deployed stents could result in serious medical complications [44, 46]. SMA stents with their superior flexibility and spring back properties prevent the stents from collapsing when compared to its counterparts [44]. The stents are designed to work in the LPS region such that even higher external pressures does not allow them to transform to their martensitic phase as the difference between plateau stresses are quite significant. This ensures that the stents are in their austenitic state without undergoing any permanent deformation under constant external pressures at all times during their deployment.

[7]These applications indicate the use of SMA like a "Metallic rubber band" without the component loosing its shape. Superelastic SMA devices in their austenitic state provide high resistance to deformation/kinks and keeping them in desired shape without any permanent deformation even under considerable loads.

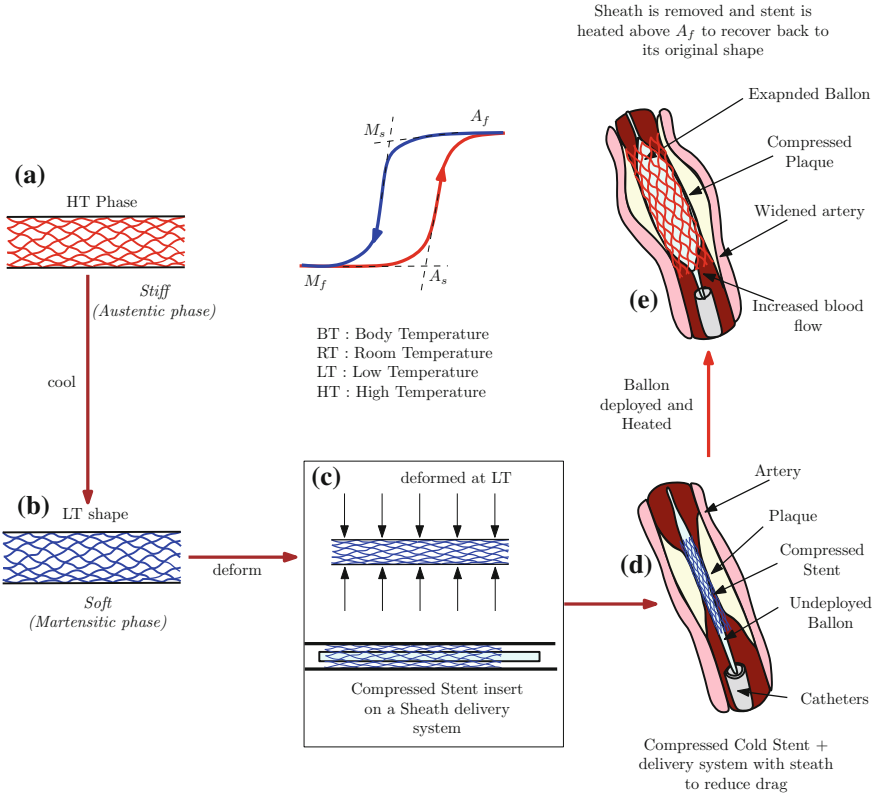

Fig. 1.10 Smart Nitinol stents are deployed to improve blood flow in blocked arteries. **a** SMA stent shape at high temperature (≈ID of artery); **b** SMA stent at room temperature (martensitic state) under no external force; **c** SMA stent deformed at room temperature to a temporary shape that is easy for insertion with the delivery system (typically balloon deployment; **d** Deformed SMA stent inserted into required position using catheters; **e** SMA stent heated to high temperature (i.e., above A_f) for it to recover back to its original shape and unblock the artery (figure adapted from [61])

SMA spacers have found applications in many orthopaedic applications with the intention of applying constant forces on fractured bones to accelerate the bone healing process [41]. As shown in Fig. 1.11, spacers (sometimes addressed as staples) in its opened (deformed) shape are deployed at the site and then heated so that the SMA in its austentic shape can apply the required compressive forces to facilitate rebuilding fracture bones [41, 47, 61].

All of these unique features makes the use of SMA components a feasible option in many biomedical application with many of them already having FDA® approvals.

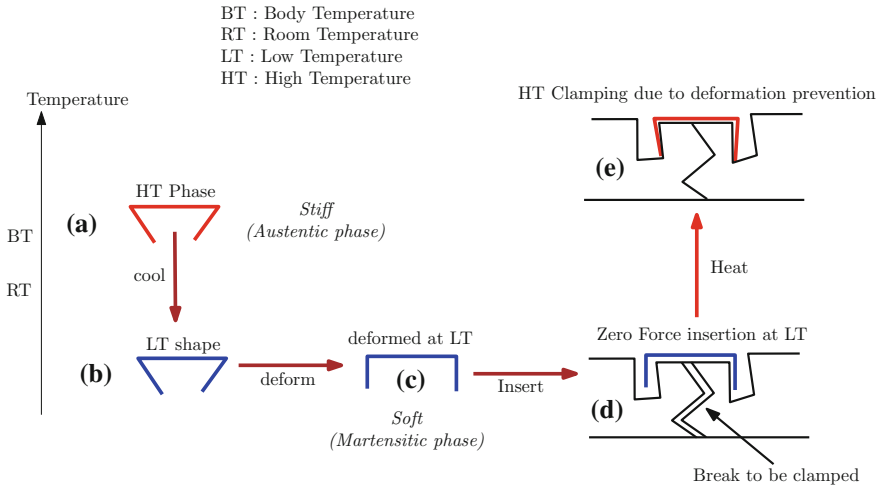

Fig. 1.11 A smart bone clamping technique using smart SMA clamps (staples) that provide enough compressive forces to clamp small separations. **a** SMA clamp shape at high temperature; **b** SMA clamp at room temperature (martensitic state) under no external force; **c** SMA clamp deformed at room temperature to a temporary shape that is easy for insertion; **d** Deformed SMA clamp inserted into required position; **e** SMA clamp heated to high temperature (i.e., above A_f) for it to recover back to its original shape and clamp due to deformation prevention (figure adapted from [61])

1.6.2 Civil Engineering Applications

In many civil engineering applications, large SMA wires, ropes, springs and beams are being used as damping elements in bridges, buildings and also in seismic resisting systems due to their excellent energy dissipation and recentering capabilities [39, 62–65]. Figures 1.12 and 1.13 show different SMA reinforcements commonly used for energy dissipation in civil engineering structures.

In seismically active areas, buildings and bridges can be prone to damage due to lateral displacements during an earthquake event [66]. Building earthquake resistant structures have been an intense area of interest lately and study of SMA rebars as possible reinforcements in this pursuit has received significant attention (see Fig. 1.13) [64, 66]. The use of SMA components as reinforcements in reinforced concrete (RC) have performed better in confining lateral column displacements even at higher amplitudes of dynamic loading (i.e., simulating large earthquake events) compared with their counterparts like steel-RC columns [66]. With superior damping capacity and ability to deliver large plateau strains at relatively constant stresses, SMAs make excellent candidates for re-centering applications in the form of bracings or dampers [66]. Many different types of SMA braces for framed structures have been designed for energy dissipation devices [39] (see Fig. 1.13).

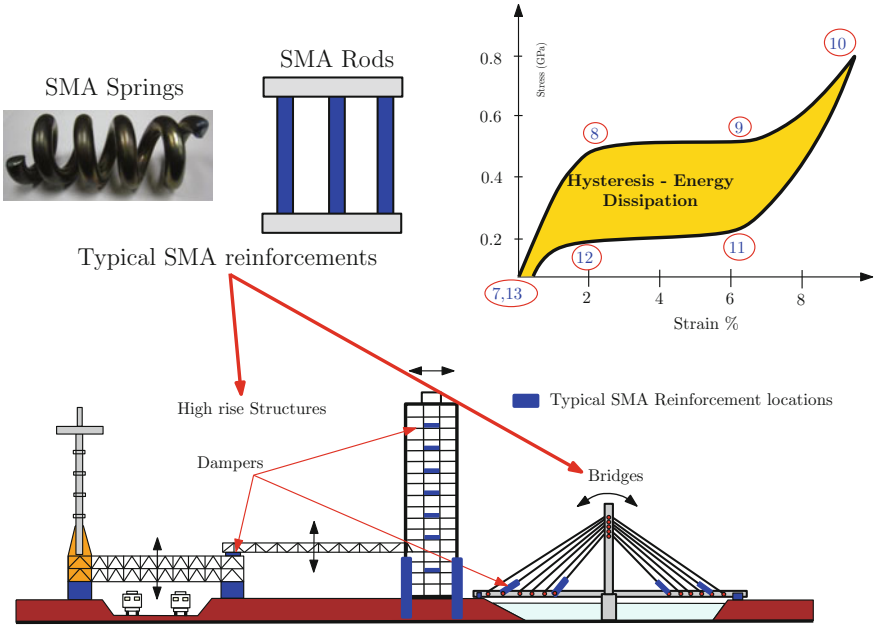

Fig. 1.12 Common SMA reinforcements in many civil structures for recentering and damping applications. The large loading–unloading hysteresis (shaded region under the entire stress–strain curve) associated with superelastic responses in SMA makes them a very efficient energy dissipation system (photographs reproduced with permission from Elsevier [39, 67–69])

In many cases, use of bundles of post-tensioned superelastic SMA wires as reinforcements have been employed for re-centering braces [39].

In addition, SMA's have also been effective contenders as beam–column or column–foundation joints in many retrofit structures as the large plateau strains in SE can be used to provide desired clamping forces to hold the members together [66]. This helps in controlling relative hinge displacements when used as restrainers [66]. By connecting the bearings and bridge deck with superelastic bars, the "position stability" of simply supported bridges could be improved [68, 70] (see Fig. 1.13).

If a SMA reinforcement in its martensitic form is deformed (also referred to as pre-tensioning) and embedded in concrete and then electrically activated such that it reaches temperatures above A_f, sufficient constraining forces can be generated as the SMA tries to recover back to its original shape [66, 70]. The extent of prestressing can be increased or decreased by controlling the amount of initial deformation of SMA reinforcement [66]. To illustrate this point, Mo and co-workers looked at developing a "smart concrete structure" that can potentially "self heal" after a damaging earthquake event [71]. In their study, as shown in Fig. 1.14, a concrete block with many reinforced SMA ropes was considered [71]. The SMA reinforcements were initially pre-deformed (pre-tensioning) in their martensitic state (i.e., below their M_f) and reinforced in the concrete block with an initial crack as shown in Fig. 1.14. On a test

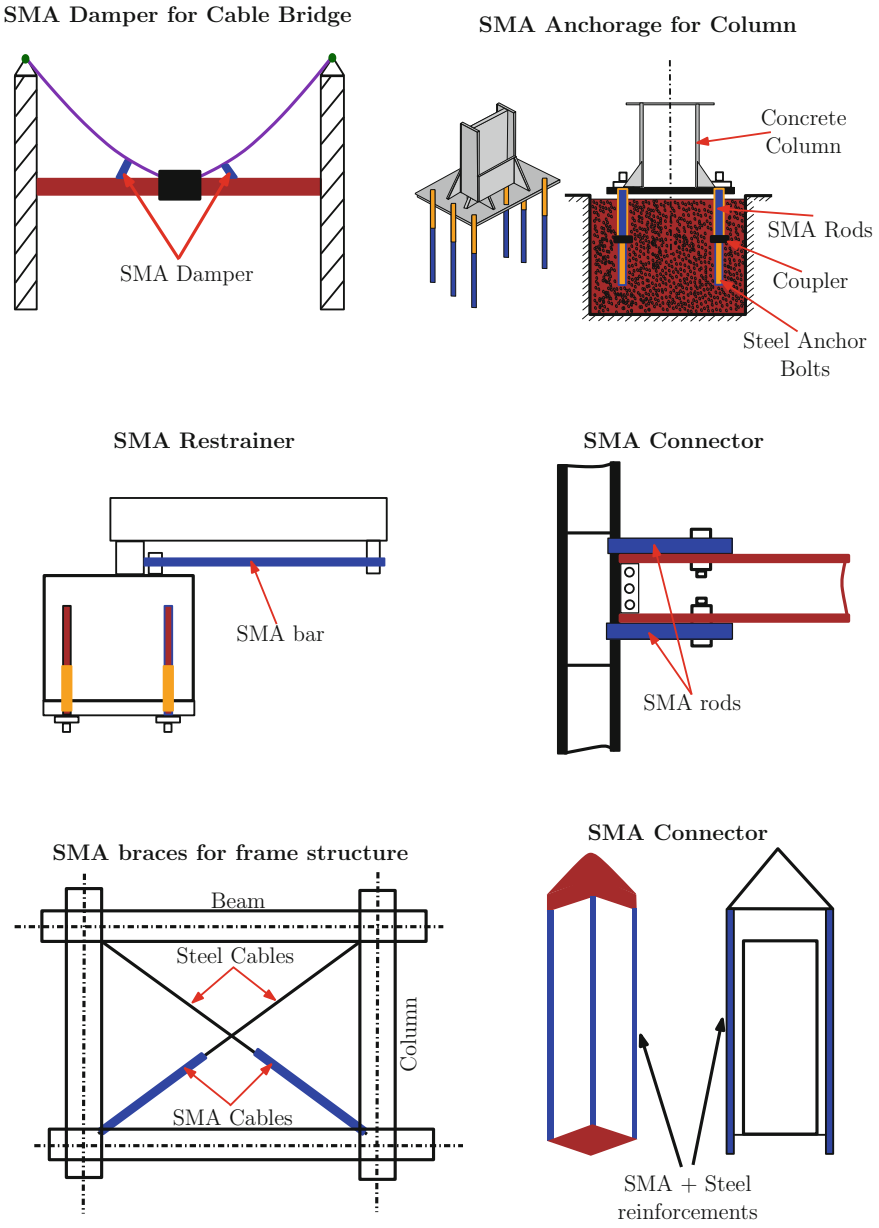

Fig. 1.13 SMA components are used as energy dissipation devices as dampers for simply supported cable bridges, anchorages and connectors for columns, braces for framed structures (photographs reproduced with permission from Elsevier [39])

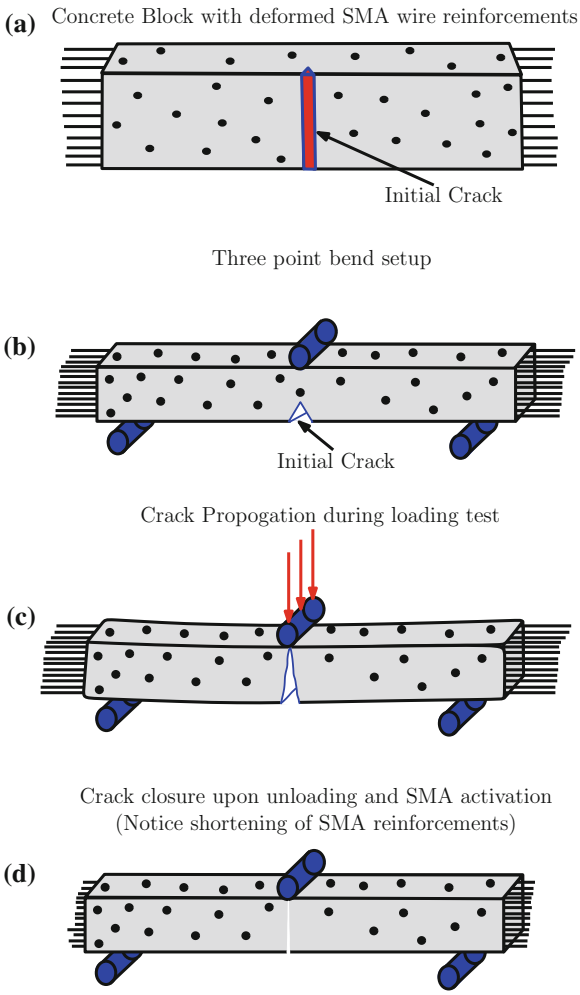

Fig. 1.14 **a** Some pretensioned SMA reinforcements in a concrete block with an initial crack. **b** Shows a classic 3-point bend setup. **c** Shows the crack propagation during loading. **d** Shows the crack closure upon unloading and activating (heating) SMA reinforcements. The SMA wires shorten causing the crack to close (figures adapted from [39, 71, 72])

bench, dynamic 3-point bend tests were conducted which resulted in crack propagation with increasing loads/frequencies mimicking a damaging earthquake event. The cracked structure is then unloaded and the SMA ropes were activated by heating them above their A_f. Upon heating, the SMA ropes transform back to austenite and shorten in length back to its original shape. This results in large recovery forces that results in crack closure (crack healing) as the concrete structures are pulled inwards and thus minimizing the chances of building collapse after an earthquake event. Many such "self-restoration" options have been proposed with almost complete crack closure upon re-centering [39]. On similar lines, many active confinement techniques where concrete columns are helically wrapped with SMA bands that can provide necessary tensioning upon heating have also been proposed [68].

In one of the first reported retrofit applications using SMA devices, the San Giorgio chruch in Italy used bundles of SMA wires to refurbish the church structure after an earthquake event [64, 69]. As shown in Fig. 1.12, SMA devices were arranged in series in order to limit the forces to the masonry under the required limits and further seismic analysis were performed to analyze the performance of SMA devices [64, 69].

However, given the size of civil engineering structures, the use of NiTi SMA components have been limited due to their high initial material and processing costs when compared to biomedical or automotive/aerospace applications [68]. Going forward, developing more cheaper copper or iron based SMAs that could replace NiTi SMA components would greatly benefit the civil engineering community which can potentially use many more *smart structures with active control abilities*. Developing such copper or iron based SMAs and further understanding their underlying microstructural changes that influence their functionality is still an active area of research [35].

1.6.3 Aerospace and Automotive Applications

SMA components are used as thermal actuators in different temperature regimes depending on the kind of applications in automotive industry [16]. The use of SMA actuators are finding applications in many pre-commercialized concepts like smart automotive lighting systems, fuel management, climate control, mirrors adjustments, locking systems, suspension adjustments, and so on [16, 73, 74]. Similarly, SMA cables, beams, torque tubes are used in tailoring inlet geometry and orientation, increase/decrease fan nozzle in various aerospace applications [37, 75–77].

In many European cars, air intake manifolds along with tumple flaps are mounted on the crankshaft. The crank reciprocates between two angular positions to open and close series of tumble flaps upon demand. The traditional arrangements uses electromagnetic or pneumatic actuators for achieving this reciprocating motion [74]. In one proposed case study, sets of SMA springs were used to generate the required actuation forces (push and pull out forces) to switch the tumple flap's angular positions against existing traditional electromagnetic or pneumatic devices [74]. Battery power supplies were used to actuate the SMA springs upon demand.

In some aircraft applications, smart SMA couplings have shown the possibility of substitute socket welds or compression fittings [78]. As shown in Fig. 1.15, at temperatures below M_f, the couplings are expanded to a larger diameter such that it can easily slide through the pipes to be fastened [78]. Once the SMA coupling transforms back to austenite (i.e., above A_f which is generally programmed to be lower than room temperature), it shrinks to its original dimension and thus securely holding the pipe ends as the SMA coupling enforces sufficient compressive forces [78].

Chopra [76] reviews many possible applications of using smart materials in aerospace systems. For example, adjusting the "trailing edge tab of a helicopter

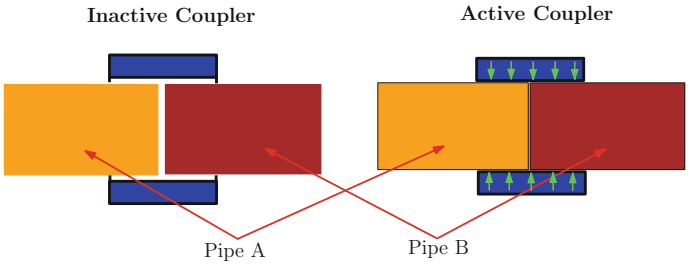

Fig. 1.15 SMAs can be used as couplers that can substitute socket welds or compression fittings in aerospace, civil, oil exploration industries. The idea is to deform the couplers in its martensitic state to larger diameters for easy sliding along the pipes and then heated to temperatures above A_f to hold the ends of pipes firmly with appropriate compressive forces (figure adapted from [78, 79])

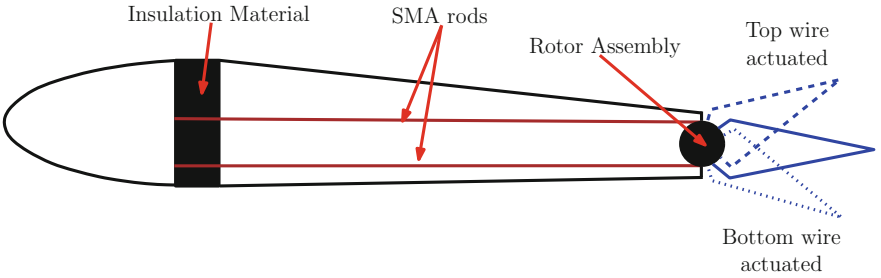

Fig. 1.16 The use of pretwisted SMA rods to torsionally actuate blade tabs between two angular positions is possibility by differentially heating of top and lower SMA rods. This way the angular positioning of "trailing edge tab of a helicopter rotor" can be adjusted and thus allowing for user in-flight tracking of its position (figure reproduced from [76] with permission from AHS International)

rotor for in-flight tracking" using SMA components is one of the key proposed ideas [76]. Pretwisted SMA rods employed to torsionally actuate blade tabs between two angular positions is possibility by differentially actuating the SMA rods as shown in Fig. 1.16. By heating the upper rod, the SMA rod shortens and causes the rotor assembly to actuate in one direction. By heating the lower rod, the SMA rod enables the rotor assembly to actuate in a different direction. The actuation of the rotor assembly in different directions causes the trailing edge to occupy multiple angular positions during flight maneuver. The SMA rods clearly perform both sensing and actuating functions and thereby reducing the use of separate working elements for sensors and actuators to perform the same task.

On similar lines, "Chevrons" in aircraft engines are employed for flow mixing and their configurations can be altered by using SMA beams based on changes in flow temperature (normally temperature changes with different altitudes) [37]. At low altitudes and speeds, with higher engine temperatures, the chevrons fold inward for improved gas mixing and thus lowering noise [37]. These chevrons relax to a different configuration at lower temperatures (i.e., at higher altitudes and speeds) to improve engine performance [37].

Japanese bullet trains came up with some smart automatic oil adjustment devices that use SMA coil springs along with a bias spring arrangement with the intention of preventing temperature rise in gearbox (see Fig. 1.3 in [80]). The oil temperature can be used to actuate the SMA spring and open and close bypass holes [80].

1.6.4 Miscellaneous Applications

The oil and gas industry uses many hydraulic and electro-hydraulic systems along with additional signal conditioning components which makes the overall system quite bulky [79]. Many SMA components along with bias systems have been considered to replace these traditional hydraulic and electro-hydraulic systems [79]. As shown in Fig. 1.17, a combination of SMA and steel bias springs were employed to actuate pistons in "blowout preventer's" that either block or unblock drill pipes

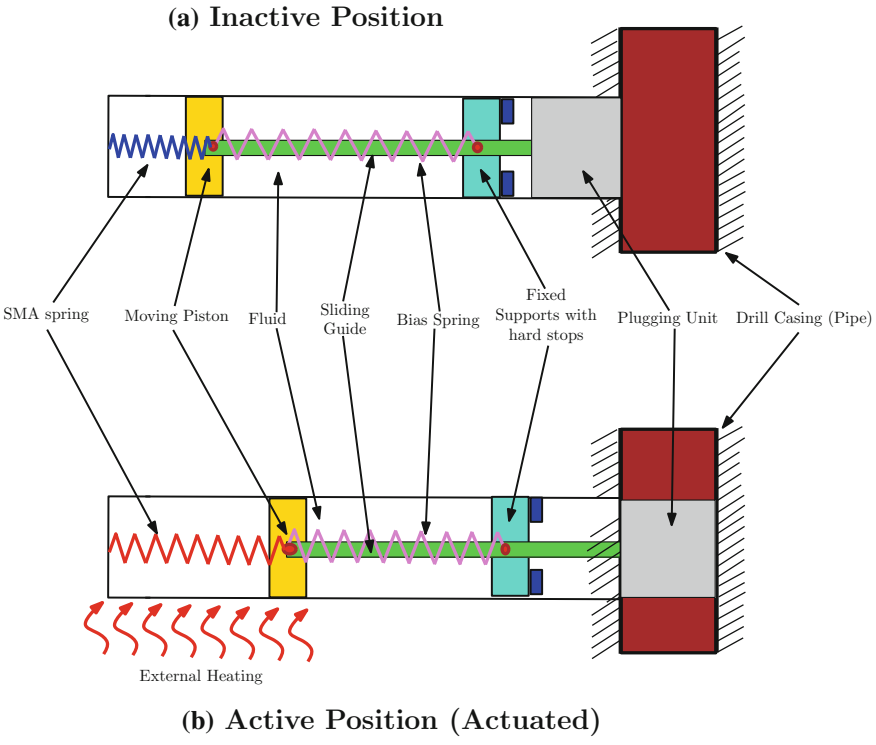

Fig. 1.17 A combination of SMA springs and bias springs are employed to actuate pistons in "blowout preventer's" that either block or unblock drill pipes. **a** Shows the the inactive position of blocking piston where the bias springs keeps the SMA spring compressed at lower temperatures (i.e., below M_f). **b** Shows the SMA spring in its austenitic state (at higher temperatures, i.e. above A_f) overpowers the bias spring forces and actuates the blocking piston to prevent any flow in drill pipe to prevent any accidents (figure adapted from [79])

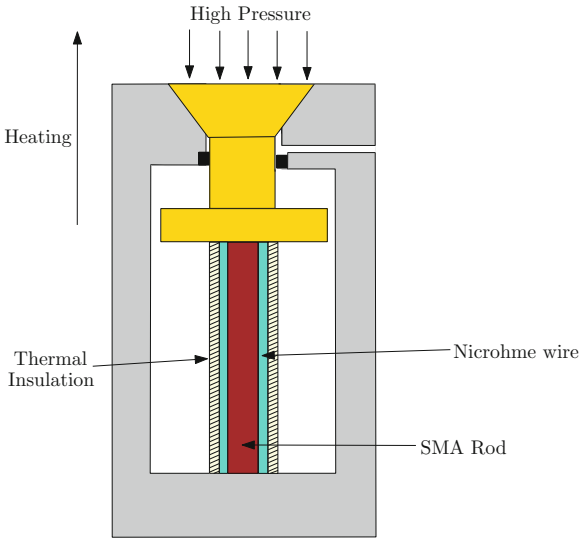

Fig. 1.18 SMA rods used as smart valves to close/open valves. SMA rods in the core are surrounded by Nichrome wires for resistive heating. At lower temperatures, SMA rods are inactive and upon heating of SMA rods results in enough recovery force generation to close/open valves depending on design

[79]. With reference to Fig. 1.17, position (a) indicates the inactive position of the blocking piston as at lower temperatures (i.e., below M_f), the bias spring keeps the SMA spring compressed and thus allowing free flow through the drill pipes during operation. At higher temperatures (i.e., above A_f), the SMA spring in its austenitic state overpowers the bias spring forces and actuates the blocking piston to block any flow in drill pipe to prevent any blowout situation. On similar lines, another example of a smart valve is discussed in Fig. 1.18 that can open or close the valve when the SMA is actuated by changes in temperature.

Due to NiTi alloys excellent corrosion resistance properties, many underwater couplers/connector applications are considered [79]. SMA springs are cryogenically cooled to temperatures below M_f so that they can easily be deformed into a temporary shape for deployment (generally expanded to larger diameter for easy deployment) [79]. On exposure to actual operating temperatures (above A_f) they can revert back to austenitic phase and tightly hold the pipe ends [79]. The couplers in their austenitic state can provide the required compressive forces to hold the pipe ends in place. On similar lines, some smart couplings have also been tested for submarine applications and underwater gas pipelines as shown in Fig. 1.15 earlier. Such couplings are easy to work with as it does not require any initial pipe surface preparation or subsequent welding or safety hazards [78].

In some fire protection devices, the ability of SMA to reversibly respond to changes in temperature have enabled them to find applications like safety valves [35]. Use of Cu-Zn-Al actuators are responsible for shutting off toxic flammable gases [35]. The idea here is that when the high temperature toxic flammable gases are sensed by the safety valves, they transform to their austenitic state which forces the safety valve to shift to its original shape that shuts the flow of toxic gases and force them through another exit.

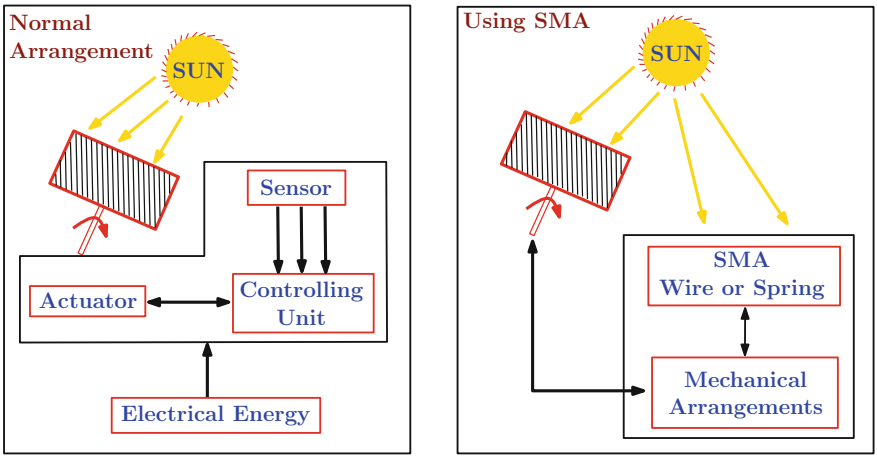

Fig. 1.19 Comparing the number of working elements with existing sun tracking system (*left*) versus number of working elements for a proposed smart sun tracking system (*right*) using SMA components that play the role of actuator and sensor simultaneously (figure adapted from [81])

On a similar note, SMA compression springs along with bias steel springs are finding applications in sprinkler systems that can automatically turn on or off based on the water temperature [35]. Ni-Ti-Cu springs are finding applications in thermostats, electrical circuit breakers where, for example, door positions in an self-cleaning oven are monitored using SMA components [35]. NiTi discs located between electrical contacts in Lithium ion batteries can be used to break the circuits at critical temperatures [35].

Ganesh and co-workers [81] have attempted to design a simple "sun tracking mechanism (SSTM)" to illustrate the dual sensing and actuating capability of SMA components. The solar receptor position is automatically changed over the course of the day by using the solar radiation to actuate the SMA springs [81]. As described in Fig. 1.19, the use of such smart SMA springs considerably reduces the number of working elements as compared against a convectional sun tracking system that employs separate sensor and actuator for providing system level response [81].

1.7 Chapter Summary

The following are the main points of this chapter.

- Smart materials are a subgroup of advanced materials that can recognize non-mechanical external stimuli like changes in temperature, magnetic field, electric field, pH changes, and so on, and reversibly respond to the same by either changing its physical or mechanical properties.

- The use of smart materials can provide a system level design by integrating multiple functions like actuation, sensing and control into a single structure with the use of one or more smart materials that can replace separate sensing and actuator components.
- Shape memory alloys (SMAs) is one such smart material where the functionalities arise from their underlying microstructural changes when subjected to external non-mechanical stimuli like temperature or magnetic field changes.
- In thermally responsive SMAs, the reversible solid-solid, diffusionless thermoelastic phase transformations between a stable high temperature austenitic phase and low temperature martensitic phase are responsible for them to demonstrate interesting phenomenon like **shape memory effect (SME)** and **superelasticity (SE)**.
- Phase transformations occur over characteristic transformation temperatures namely martensitic start (M_s), martensitic finish (M_f), austenitic start (A_s) and austenitic finish (A_f).
- SME is the ability of SMA to return to a predetermined shape upon heating above the characteristic transformation temperature A_f; SE is the ability of SMA to recover large strains ($\sim$8 %) with associated stress-strain hysteresis due to mechanical loading-unloading under isothermal conditions above A_f.
- The unique characteristics SME and SE enabled SMAs to find applications across many areas like vibration damping, biomedical, automotive and aerospace areas. Smart structures with active control abilities is now a possibility with the use of SMA.

References

1. Lagoudas D (2008) Shape memory alloys: modeling and engineering applications. Springer, New York
2. Ghosh P, Rao A, Srinivasa AR (2013) Design of multi-state and smart-bias components using shape memory alloy and shape memory polymer composites. Mater Des 44:164–171
3. Momoda LA (2004) The future of engineering materials: multifunction for performance-tailored structures, vol 34. National Academy of Engineering, The Bridge
4. Janocha H (1999) Adaptronics and smart structures: basics, materials, design, and applications. Springer, Heidelberg
5. Matic P (2003) Overview of multifunctional materials. In: Proceedings of SPIE, vol 5053, p 61
6. Ghosh P, Rao A, Srinivasa A (2012) Multifunctional smart material system (msms) using shape memory alloys and shape memory polymers. In: SPIE smart structures and materials+nondestructive evaluation and health monitoring. International society for Optics and Photonics, pp 83411O–83411O
7. Okano T (1993) Molecular design of temperature-responsive polymers as intelligent materials. Responsive Gels Vol Transit II 110:179–197
8. Leng J, Lan X, Liub Y, Dua S (2011) Shape-memory polymers and their composites: stimulus methods and applications. Prog Mater Sci 56(7):1077–1135
9. Noor AK (1999) Computational structures technology: leap frogging into the twenty-first century. Comput Struct 73(1):1–31
10. Naraghi M (2012) Multifunctional materials—aero 606 course notes, Fall 2012

11. Corporation L (2012) Company website, 2012. http://www.lord.com
12. Leventon W (1993) Fluid damper may make truck seats safer. Des News 4:185–186
13. Carlson JD (2002) What makes a good Mr fluid? J Intell Mater Syst Struct 13(7–8):431–435
14. Carlson JD (2005) Mr fluids and devices in the real world. Int J Mod Phys B 19(07n09):1463–1470
15. Carlson J, Catanzarite D, Clair KS (1996) Commercial magneto-rheological fluid devices. Int J Mod Phys B 10(23n24):2857–2865
16. Stoeckel D (1990) Shape memory actuators for automotive applications. Mater Des 11(6):302–307
17. Huang W, Ding Z, Wang C, Wei J, Zhao Y, Purnawali H (2010) Shape memory materials. Mater Today 13(7–8):54–61
18. Doraiswamy S, Rao A, Srinivasa A (2011) Combining thermodynamic principles with preisach models for superelastic shape memory alloy wires. Smart Mater Struct 20(8):085032
19. Van Humbeeck J (2001) Shape memory alloys: a material and a technology. Adv Eng Mater 3(11):837–850
20. McNichols J, Cory J (1987) Thermodynamics of nitinol. J Appl Phys 61(3):972–984
21. Rao A (2013) Modeling bending response of shape memory alloy wires/beams under superelastic conditions-a two species thermodynamic preisach approach. Int J Struct Changes Solids 5:1–26
22. Wayman C (1992) Shape memory and related phenomena. Prog Mater Sci (UK) 36:203–224
23. Shaw J, Churchill C, Iadicola M (2008) Tips and tricks for characterizing shape memory alloy wire: part 1 differential scanning calorimetry and basic phenomena. Exp Techn 32(5):55–62
24. A. S. F2004–05R10 (2010) Standard test method for transformation temperature of nickel-titanium alloys by thermal analysis, ASTM International, West Conshohocken, PA
25. Shaw JA, Kyriakides S (1995) Thermomechanical aspects of niti. J Mechan Phys Solids 43(8):1243–1281
26. A. S. F. 2007e2 (2007) Standard test method for tension testing of nickel-titanium superelastic materials, ASTM International, West Conshohocken, PA
27. Rao A, Srinivasa AR (2014) A three-species model for simulating torsional response of shape memory alloy components using thermodynamic principles and discrete preisach models. Math Mech Solids doi:10.1177/1081286514545917
28. Rao A, Srinivasa A (2013) A two species thermodynamic preisach model for the torsional response of shape memory alloy wires and springs under superelastic conditions. Int J Solids Struct 50(6):887–898
29. Wayman C, Duerig T (1990) An introduction to martensite and shape memory. Butterworth-Heinemann, Engineering Aspects of Shape Memory Alloys (UK), 1990, pp 3–20
30. Otsuka K, Wayman CM (1999) Shape memory materials. Cambridge University Press, Cambridge
31. Duerig T, Pelton A (1994) Ti-ni shape memory alloys. Materials properties handbook: titanium alloys, pp 1035–1048
32. Huang W (2002) On the selection of shape memory alloys for actuators. Mater Des 23(1):11–19
33. Baumann MA (2004) Nickel-titanium: options and challenges. Dent Clin North Am 48(1):55–68
34. Leng J, Lan X, Liu Y, Du S (2011) Shape-memory polymers and their composites: stimulus methods and applications. Prog Mater Sci 56(7):1077–1135
35. Bogue R (2009) Shape-memory materials: a review of technology and applications. Assembly Autom 29(3):214–219
36. Ma J, Karaman I, Noebe RD (2010) High temperature shape memory alloys. Int Mater Rev 55(5):257–315
37. Hartl D, Lagoudas D (2007) Aerospace applications of shape memory alloys. Proc Inst Mech Eng Part G J Aerosp Eng 221(4):535
38. Sutapun B, Tabib-Azar M, Huff M (1998) Applications of shape memory alloys in optics. Appl Opt 37(28):6811–6815

39. Song G, Ma N, Li H (2006) Applications of shape memory alloys in civil structures. Eng Struct 28(9):1266–1274
40. Van Humbeeck J (1999) Non-medical applications of shape memory alloys. Mater Sci Eng A 273:134–148
41. Machado L, Savi M (2003) Medical applications of shape memory alloys. Braz J Med Biol Res 36(6):683–691
42. FDA (2012) United states food and drug administartion, 2012. http://www.fda.gov
43. Lagoudas D, Rediniotis OK, Khan MM (2000) Applications of shape memory alloys to bio-engineering and biomedical technology. In Proceedings of the 4th international workshop on scattering theory and biomedical applications, Perdika, Greece, pp 195–207 Oct 1999
44. Duerig T, Pelton A, Stöckel D (1999) An overview of nitinol medical applications. Mater Sci Eng A 273:149–160
45. Mantovani D (2000) Shape memory alloys: properties and biomedical applications. JOM J Miner Met Mater Soc 52(10):36–44
46. Morgan N (2004) Medical shape memory alloy applicationsthe market and its products. Mater Sci Eng A 378(1):16–23
47. El Feninat F, Laroche G, Fiset M, Mantovani D (2002) Shape memory materials for biomedical applications. Adv Eng Mater 4(3):91–104
48. Gil F, Manero J, Planell J (1996) Relevant aspects in the clinical applications of niti shape memory alloys. J Mater Sci Mater Med 7(7):403–406
49. Manhartsberger C, Seidenbusch W (1996) Force delivery of ni-ti coil springs. Am J Orthod Dentofac Orthop 109(1):8–21
50. Miura F, Mogi M, Ohura Y, Karibe M (1988) The super-elastic japanese niti alloy wire for use in orthodontics part iii. studies on the japanese niti alloy coil springs. Am J Orthod Dentofac Orthop 94(2):89–96
51. Spinella I, Dragoni E, Stortiero F (2010) Modeling, prototyping, and testing of helical shape memory compression springs with hollow cross section. J Mech Des 132:061008
52. Drake S, Wayne D, Powers J, Asgar K (1982) Mechanical properties of orthodontic wires in tension, bending, and torsion. Am J Orthod 82(3):206–210
53. Miura F, Mogi M, Ohura Y, Hamanaka H (1986) The super-elastic property of the japanese niti alloy wire for use in orthodontics. Am J Orthod Dentofac Orthop 90(1):1–10
54. Kapila S, Sachdeva R (1989) Mechanical properties and clinical applications of orthodontic wires. Am J Orthod Dentofac Orthop 96(2):100–109
55. Von Fraunhofer J, Bonds P, Johnson B (1993) Force generation by orthodontic coil springs. Angle Orthod 63(2):145
56. Schneevoigt R, Haase A, Eckardt V, Harzer W, Bourauel C (1999) Laboratory analysis of superelastic niti compression springs. Med Eng Phys 21(2):119–125
57. Thompson S (2000) An overview of nickel-titanium alloys used in dentistry. Int Endod J 33(4):297–310
58. Maganzini AL, Wong AM, Ahmed MK (2010) Forces of various nickel titanium closed coil springs. Angle Orthod 80(1):182–187
59. Petrini L, Migliavacca F (2011) Biomedical applications of shape memory alloys. J Metal 2011:15, Article ID 501483. doi:10.1155/2011/501483
60. Ryhänen J (1999) Biocompatibility evolution of nickel-titanium shape memory alloy. Academic Dissertation, Faculty of Medicine, University of Oulu, Oulu, Finland.[Links]
61. Poncet PP (2000) Nitinol medical device design considerations. Strain 2(4):6
62. Speicher M, Hodgson D, DesRoches R, Leon R (2009) Shape memory alloy tension/compression device for seismic retrofit of buildings. J Mater Eng Perform 18(5):746–753
63. Saadat S, Salichs J, Noori M, Hou Z, Davoodi H, Bar-On I, Suzuki Y, Masuda A (2002) An overview of vibration and seismic applications of niti shape memory alloy. Smart Mater Struct 11:218
64. DesRoches R, Smith B (2004) Shape memory alloys in seismic resistant design and retrofit: a critical review of their potential and limitations. J Earthq Eng 8(3):415–429

65. Wilson J, Wesolowsky M (2005) Shape memory alloys for seismic response modification: a state-of-the-art review. Earthq Spectra 21:569
66. Alam M, Youssef M, Nehdi M (2007) Utilizing shape memory alloys to enhance the performance and safety of civil infrastructure: a review. Can J Civ Eng 34(9):1075–1086
67. Mirzaeifar R, DesRoches R, Yavari A (2011) A combined analytical, numerical, and experimental study of shape-memory-alloy helical springs. Int J Solids Struct 48(3–4):611–624
68. Janke L, Czaderski C, Motavalli M, Ruth J (2005) Applications of shape memory alloys in civil engineering structuresoverview, limits and new ideas. Mater Struct 38(5):578–592
69. Indirli M, Castellano MG, Clemente P, Martelli A (2001) Demo-application of shape memory alloy devices: the rehabilitation of the s. giorgio church bell tower. In: SPIE's 8th annual international symposium on smart structures and materials, International Society for Optics and Photonics, pp 262–272
70. DesRoches R, Delemont M (2002) Seismic retrofit of simply supported bridges using shape memory alloys. Eng Struct 24(3):325–332
71. Mo Y, Song G, Otero K (2004) Development and testing of a proof-of-concept smart concrete structure. In: Proceeding of smart structures technologies and earthquake engineering
72. Tamai H, Kitagawa Y (2002) Pseudoelastic behavior of shape memory alloy wire and its application to seismic resistance member for building. Comput Mater Sci 25(1):218–227
73. Williams K, Chiu G, Bernhard R (2002) Adaptive-passive absorbers using shape-memory alloys. J Sound Vib 249(5):835–848
74. Bellini A, Colli M, Dragoni E (2009) Mechatronic design of a shape memory alloy actuator for automotive tumble flaps: a case study. IEEE Trans Ind Electron 56(7):2644–2656
75. Loewy R (1997) Recent developments in smart structures with aeronautical applications. Smart Mater Struct 6:R11
76. Chopra I (2000) Status of application of smart structures technology to rotorcraft systems. J Am Helicopter Soc 45(4):228–252
77. Razov A, Cherniavsky A (1999) Applications of shape memory alloys in space engineering: past and future. Eur Space Agency Publ ESA SP 438:141–146
78. Brook G (1983) Applications of titanium-nickel shape memory alloys. Mater Des 4(4):835–840
79. Song G, Patil D, Kocurek C, Bartos J (2010) Applications of shape memory alloys in offshore oil and gas industry: a review. In: Proceedings of earth and space 2010: engineering, science, construction, and operations in challenging environments, vol 366, Honolulu, HI, USA, pp 14–17 March 2010
80. Otsuka K, Kakeshita T (2002) Science and technology of shape-memory alloys: new developments. MRS Bull 27(02):91–100
81. Ganesh N, Maniprakash S, Chandrasekaran L, Srinivasan S, Srinivasa A (2011) Design and development of a sun tracking mechanism using the direct sma actuation. J Mech Des 133:075001

Chapter 2
Need and Functionality Analysis

Generally the word "smart" is used synonymously with "intelligent," but in reality the dictionary definition of "smartness" is "marked by often sharp forceful activity or vigorous strength," that is, it is active. Thus, a "smartphone" can only wake you up with an alarm, but in reality it is smart only if it can "slap" you if you don't wake up! Fig. 2.1 shows the hierarchy of systems ranging from "passive" to "intelligent".

In fact there is a fundamental divide between being intelligent and being smart. Your computer or your cell phone is intelligent. You do what it tells you to do. Think about this for a minute. Your computer is your "boss". You wake up and look to it for instructions, (email, calendar, etc.) you do what it says and then you report back to it. If that is not a definition of a "boss," we don't know what is? On the other hand, your car is like your valet or butler: you tell it what to do and it actually carries out the actions. These are "actuators" that perform actions (typically mechanical) based on our command. It is the latter category that we need to focus on.

The capabilities of your car came from the assembly or configuration of its parts, just like the first "valves" were made by assembling parts. Today, most actuators (think solenoids, electric motors, etc.) are in this stage of development. Over time, the valves gave way to transistors (where the configuration is microscopic) and finally to very large scale integrated circuits (VLSI) where the functionality is from its "microscopic" configuration and no further assembly is required. We are gradually heading to this stage in actuators too and the challenges are immense due to the need for motion. Just as information technology is based on the ability to store, manipulate and retrieve information (bits), actuator technology is based on the ability to store, retrieve and manipulate shapes or geometry. For cars, and electric motors, this ability to change geometry is based on the assembly of parts. Shape memory alloys (SMA) have this ability because of their "microstructure." They are the equivalent of transistors.

In order to have the ability to store shapes and retrieve them on demand, we need to supply two things: Energy to transform the shape and a means for transmitting a "signal" to initiate the shape change. For a typical SMA, the energy is in the form

A. Rao et al., *Design of Shape Memory Alloy (SMA) Actuators*,
SpringerBriefs in Computational Mechanics, DOI 10.1007/978-3-319-03188-0_2

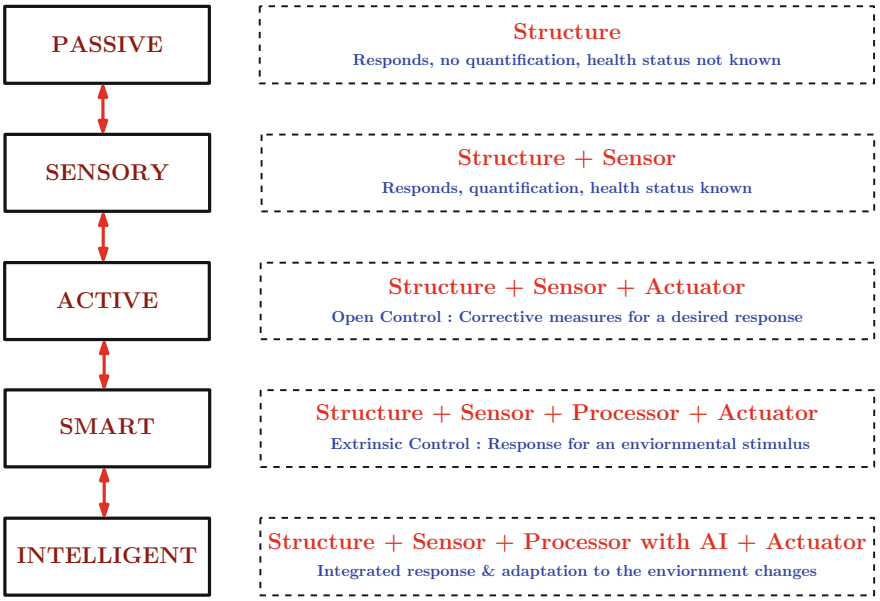

Fig. 2.1 Smartness in systems—a hierarchy

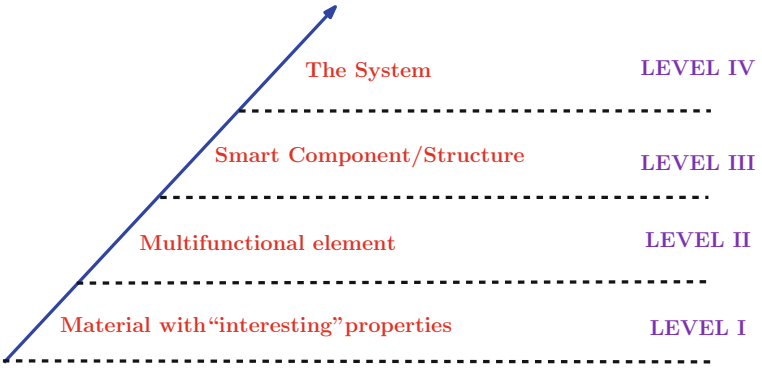

Fig. 2.2 Smart system: an integrated use of components/structures, actuators, sensors and control systems that can reversibly respond to changes in external conditions

of heat and the signal is a temperature signal. The core idea here is to find or create materials with "interesting" properties and create such systems (see Fig. 2.2). The question is how best to design components that can exploit this ability? We turn to this in the next section.

2.1 The System Design Process

One of the key activities that distinguishes humans from other animals is our ability to *design*, that is, create or synthesize something that we can use that wasn't there before. This activity of designing is something that we take so much for granted that we automatically do it without paying much attention.

So what is this activity called designing and how do we do it? *Design is the act of identifying societal needs (who is the end user? and what do they need?) and constraints; quantifying them and using a combination of scientific principles and experience to synthesize products and processes to satisfy the needs.* This combination of products or devices and processes that they can undergo is called a "system." Thus we have the electric grid system, the sanitation system, the transport system, the computer network system, etc.

The above quote shows the essential difference between a scientist and an engineer: a scientist typically creates "explanations". An engineer creates products and processes. In other words, when you are explaining how nature works, you are undertaking a scientific activity. On the other hand, when you are building a product or a process, you are undertaking an engineering activity.

Thus all designs start with the the idea of "I/we want...," "I/we need...," or "wouldn't it be nice if we could..." and ends with a specific product (cell phone) or process (wireless transmission). For a design example, if one looks at a cell phone, it is a product that allows to communicate and record information (need) in a portable way (constraint). However, the task of designing is not a simple linear process starting from a need and ending in a product. It is fundamentally *iterative*: We don't quite know what we want when we start but as we begin building we discover things and improve our design. There are thus three major activities that occur in a cycle (see Fig. 2.3) involved in designing things:

1. identification of needs and constraints,
2. embodiment or synthesis of product or process, and
3. evaluation of the product or process to see if it meets the need or constraints.

Typically, if you are a novice designer, (remember your high school project), you will do this iterative process at a subconscious level, without thinking deeply about it. In particular, you probably built and rebuilt your product until it worked (probably many times). However, this is not a sensible way to carry out large projects. Imagine building a bridge this way: You "eye ball" the river and begin building; at the end, if your bridge doesn't reach the other end, or if the first truck that goes across it plunges into the river, you say "oops ... let us rebuild it," and start over. This would not be a viable process. Similarly, if you are building a rocket going to the moon, you don't say, "oops ... missed again," and then rebuild the rocket. The lesson to learn from these admittedly extreme examples is that *avoid a simple trial and error approach when designing.*

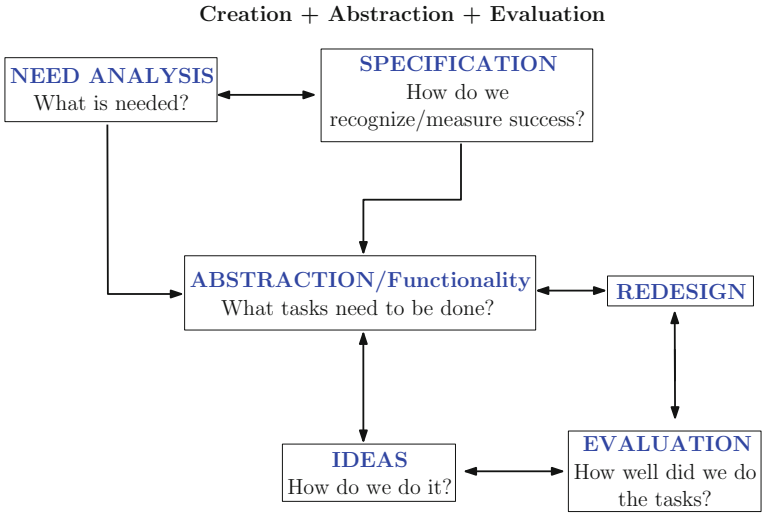

Fig. 2.3 Design methodology involves creation, abstraction and evaluation: structure and guidelines

A much better way to design something is to follow a systematic design process. There are many ways to organize this process and one such way is shown in the Fig. 2.3. Note the number of cycles or iterative loops in the process. The essential process is described in the subsection.

2.1.1 Design Methodology: Structure and Guidelines

The list below (see Fig. 2.3) is a very abbreviated/bowlderized version of Nam Suh's *Axiomatic Systems Design Approach* [1].

1. Know your clients, not only users but all those who are affected by your design.
2. Develop a need statement. A precise, abstract, qualitative high level statement that incorporates the principal need(s) and the principal constraint(s).
3. Develop a function hierarchy or function structure. Each function should be an action that has to be performed on an object producing a measurable outcome. The hierarchy is arranged in a tree so that asking "why should this function be performed?" leads up the hierarchy into the root of the tree. Asking "how this function is to be performed?" leads down the hierarchy into the "leaves."
4. Using the outcomes of the function hierarchy, develop specifications that will be used to measure the degree of success of the design. This is typically done using a "Quality Function Deployment Technique" or "house of quality" [2].

5. Develop at least three embodiments that can be used to satisfy the functions. The guidelines for developing embodiments are given below:
    a. Separate functions. Don't have more than one optimization criteria for a single component.
    b. Provide short and direct paths. Don't make complicated conduits; Don't make too many transformations of energy etc.
    c. Exploit symmetry. This is the idea of "balance" and "harmony" in art and design.
    d. Avoid sharp gradients. Don't make abrupt changes in properties or geometry and try not to connect dissimilar components together.
    e. Minimize interfaces. Pay close attention to transitions or connections (most designs fail at the connections)
    f. Do not over constrain.
    g. Minimize information content. More parts, more items to specify the less elegant is the design
6. For each function, evaluate how well it meets the need and specifications.

Each of the above items requires an extensive discussion which is outside the scope of this book. In many cases, the specifications will change as we design (since we will know more about the product than when we started, and people will gradually realize what they actually wanted) hence this is an iterative process. Most design books suggest that about 50–80 % of the cost of the product will be set at the time when we are designing. In other words, it is well worth your time and effort to go through many "paper" iterations before actually building the product. This is a hard thing to do, since it appears that we have made no progress, but in reality, it will considerably shorten the total design time since it will eliminate many costly mistakes or misunderstandings

It is clear that if we could *predict* the way your product or process is going to behave *before* you build it then you wouldn't have to actually build the product to see if it works. You could save time, money, and even the lives of many people, if you could *mathematically* simulate the process or product under a variety of conditions and test its performance virtually. Use science and mathematics.

One can predict the behavior and evaluate its performance, only if we have scientific (reproducible) explanations for its behavior and mathematical quantification of its performance. For example, if one wants to predict whether a bridge or a building is strong, one needs a scientific explanation of how bridges or buildings fail and a quantification of how much load this particular bridge or building has to withstand. Hence the essence of engineering is *quantitative decision making*, that is, measure or calculate something and then make decisions on whether it is sufficient or not.

This point is so important that we will restate it.

> The purpose of all engineering calculations is to make decisions in other words, for EVALUATING options. Thus, engineering is the art of quantitative decision making,

In your statics and dynamics class, when you solve truss problems and were asked to find forces in a truss, why did you need to do that calculation? You need

to do it in order to decide whether the truss is safe (i.e., whether it will collapse or cease to function) or not. In the same manner, EVERY calculation that we make in engineering has a decision attached to it. This is one of the reasons that engineers are in such demand: they have been taught the process of quantitative decision making that is essential for so many disciplines. The purpose of the course on mechanics of materials is to help you master this capability of decision making in structural engineering.

2.2 The Five Major Subsystems

In the previous section we introduced the notion of the design of a "system" to satisfy a societal need. In order to gain a better understanding of such systems, let us consider an example system such as a car (which is itself a part of a transport system).

Can you think of what societal need a car satisfies? it is a system to transport by land, a small group of individuals and objects (luggage) safely and at a reasonable speed between any two chosen locations. This is what a car *does*. The above description of what a car does is called a *need statement*. Notice how it is an abstract statement (i.e. it does not directly reference what a car looks like (its embodiment) or even how it does what it does. It describes what happens to the user when they use a car, somewhat like a "before" and "after" statement.

Another example that might give you some insight into a need statement is that for a pen. What societal need does a pen satisfy? Before you read further, just pause and write down what you think is an answer. As a hint, imagine why you wanted to use a pen and what happens when you finish using a pen. A pen is a device that records information in a visually retrievable form on a two dimensional surface. You might have had something like "a pen writes on paper." What a need statement does is asks, "give me a physics description of "write" and "paper." In the need statement here, we have converted "write" into "record information in a visually retrievable form" and "paper" as "two dimensional surface." You might quibble with this but that is quite acceptable; your definition may be at variance with ours.

The point about a need statement is that it allows for other possible solutions than just the one that we are currently utilizing. For example, we now have electronic versions of this, namely the iPad® and other tablet computers that can do the same thing. Thus a pen, a pencil, and a tablet PC are all different "embodiments" that satisfy the same need statement in different ways. Seen in this way, a car is a specific embodiment. A scooter, a bicycle, and (if you are in India) an auto rickshaw are other embodiments that offer personal land transportation.

In spite of the different embodiments looking completely different, all embodiments share some common features from a scientific point of view. They are all systems that can transmit, convert store or dissipate some basic physical quantities.

Different embodiments differ in the extent to which they carry out this activity. Specifically, every embodiment can have the following features:

1. *Transmit and store momentum*. Systems whose primary task is to transmit momentum are called structures. If a system can also store momentum then it is said to have inertia. Classical physics tells us that we cannot convert momentum.
2. *Transmit, store and convert mass*. Systems that are primarily designed to transmit mass are called conduits pipelines, etc. Mass of one kind can be converted through chemical reactions to other kinds of mass and systems that can do this are called chemical plants or reactors.
3. *Transmit store and convert motion*. Systems that are primarily designed to transmit and convert motion are called mechanisms. By motion conversion, we mean converting rotational to translational motion, etc. springs and solenoids and SMAs are designed to "store" and convert motion.
4. *Transmit, store and convert power* (*specially mechanical and electrical*). Systems that are primarily devoted to transmit power are called power transmissions or simply "transmission." We emphasize that you cannot transmit mechanical power without simultaneously transmitting force and motion. Power conversion devices are generally referred to as "transducers." If a system converts chemical or thermal power to mechanical power it is called a "heat engine" or simply engine. If it converts electrical power it is called a "motor." Power storage devices are generally referred to as "batteries."
5. *Transmit, store and convert signals/information*. Systems that are primarily devoted to transmission, storage and conversion of signals are called information systems. If they are also used in conjunction with the other systems they are referred to as "control systems."

How does this classification come into play with a car? Some comments are in order.

1. The body of the car and its suspension are primarily involved in the transmission of forces and momentum and so they are the structural system of the car.
2. The engine and the AC system is naturally the power system or transducer in the car.
3. The fuel and electrical conduits, the heating and AC conduits and the doors and windows are the system involved with transmission of mass.
4. The sensors, speedometer, indicators, steering wheel, brake[1] and so on, form the information system in the car.

[1] You might be thinking, "wait, the brake transmits forces. If you think about it a little bit, you will realize that the brake only transmits the braking signal to the hydraulic system that actually supplies the braking force." You cannot stop the car with the force generated by your foot.

2.3 How Do We Identify Need and Functionality for SMAs

A SMA actuator is indicated when one needs motion with a large strokes and forces in a confined area and with a small footprint. SMA is like a solid-state micro scale hydraulic actuator. Due to its very inefficient energy conversion capabilities (less than 5 % conversion efficiency) it is not yet competitive with piezoelectric vibrators and motors for rapid repeatable action. However, due to its large stroke and wire-like form factor, it has its niche (which is growing rapidly).

NiTi SMA wires being very corrosion resistant and thus find suitable applications for use in harsh environments or for deployment in situations where maintenance is not easy. For example, it could be used for various applications in solar plants where heat is readily available and slow large motion is required. Since, SMAs performance can be affected by operating temperature, loading condition, loading rate, component geometry, material composition and thermomechanical history, it is essential to design a proper feedback control system if precision motion is required. However, in many applications, it is sufficient to develop a "two state" or two configuration system, that is, a simple "on-off" type mechanism such as for an emergency relief valve. In these cases, we can obtain guaranteed motion by incorporating a "slot" or a "cam" in the system so that motion of the wire is guided and the "on" and "off" positions are precisely located.

In general, the following questions need to be asked before deciding on the use of an SMA actuator:

1. *Is the motion needed primarily linear* "*back and forth*" *rather than cyclic* "*rotary*?" If it is a full rotatory motion, it might be better to use an electric motor.
2. *Is the intended motion fast or slow, i.e., what is the cycle duration*? If the frequency is in the range of 1–10 Hz then it would be reasonable to use an SMA. Above that, there might not be sufficient cooling time for the SMA to retract and complex multiplexing or cooling systems may be necessary.
3. *Is it a two state system*? SMAs perform very well in such situations where precise location is required only at the start and the end.
4. *Is it going to be used primarily for power conversion* (i.e. *as a heat engine*)? If so, then other technologies may be more competitive due to their higher efficiency. SMAs could be useful only in cases where there is limited space.
5. *What kind of heat source is available*? Electrically heated systems perform best due to the relative ease of heating, however cooling is always a problem. If direct heating is not possible, then circulating a heated fluid may be very effective (e.g., an SMA tube with internal heating/cooling fluid flow may be feasible). There have been designs where an external Nichrome wire was wrapped round the SMA to heat it by passing current to the Nichrome wire.
6. *What is the operating temperature*? *Are there temperature fluctuations*? Typical SMAs work very well between room temperature and about 150 °C. While there are high temperature SMAs they tend to have brittleness and stability issues and may be very expensive. Also ambient temperature excursions may cause either

premature or delayed actuation and hence it is better if SMAs are in enclosures that are reasonably well protected unless great precision is not needed.

7. *How many cycles should it be able to run for*? SMAs are remarkably tolerant to corrosive environments but repeated cycling causes "memory loss" (a phenomenon that is common to laptop batteries and in some sense similar to it) and eventual degradation. Thus, if a large number of thermomechanical cycles are needed, it is best to restrict strains to about 2–4 % even though it may be able to be deformed and recover to nearly 8 % strains. Commercial vendors typically suggest about 4 % strain

Once these questions are answered, the feasibility of using a SMA may be determined.

References

1. Website W (2014) Nam suhs axiomatic systems design approach. http://en.wikipedia.org/wiki/Axiomatic_design
2. H. of Quality Matrix (2014) Company website. http://www.webducate.net/qfd/qfd-hoq-tutorial.swf

Chapter 3
Manufacturing and Post Treatment of SMA Components

We would like to briefly review the different manufacturing techniques and post-treatment techniques employed in the SMA community. This is by no means an exhaustive discussion but a broad overview of some important items. As designers, it is important to having an understanding of these details especially as many of these techniques/treatments discussed here influence the performance of the component immensely.

We start with reviewing some of the major techniques used in manufacturing NiTi SMAs and later discuss some key issues about post processing these manufactured SMAs.

3.1 Different Manufacturing Techniques

Much of the SMA manufacturing and the processing literature is propriety and only a small part of it is published. A broad classification of many production techniques are summarized in Fig. 3.1. The two primary manufacturing techniques can be classified under casting or powder metallurgy techniques.

The casting techniques are popular with large scale production of NiTi alloys that generally involves melting followed by hot working and final machining [1]. The most popular commercial casting techniques are using Vacuum induction melting (VIM), Vacuum Arc remelting (VAR) and Electronic Beam Melting (EBM) techniques.

Several powder metallurgy or additive manufacturing techniques have been proposed for small scale NiTi component production. These techniques are mainly used by the biomedical community that involves manufacturing of porous NiTi components or intricate components due to their expensive initial costs.

A. Rao et al., *Design of Shape Memory Alloy (SMA) Actuators*,
SpringerBriefs in Computational Mechanics, DOI 10.1007/978-3-319-03188-0_3

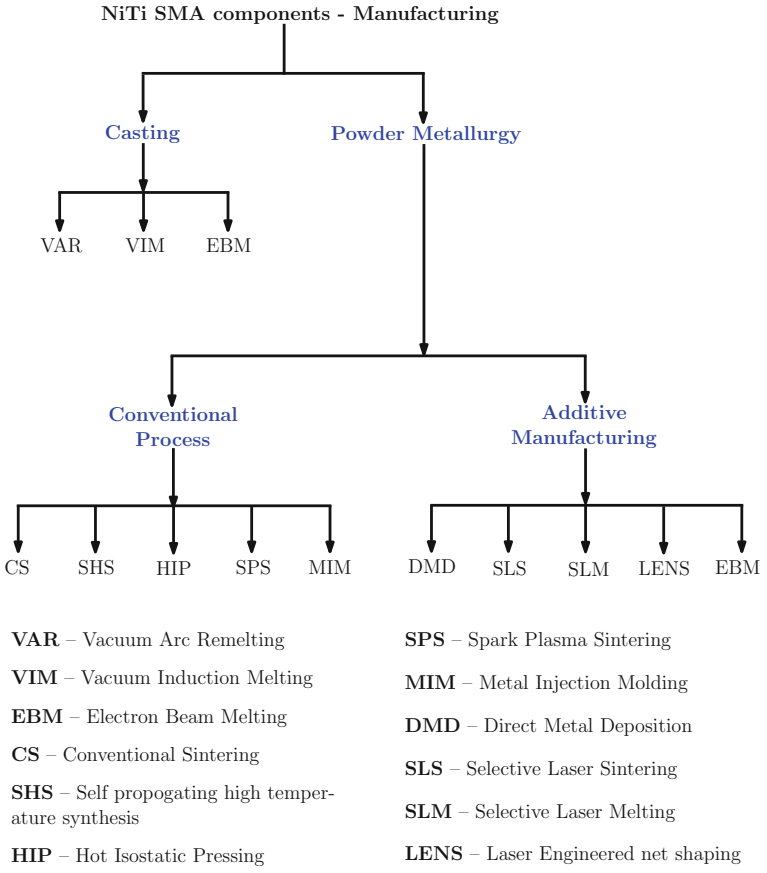

Fig. 3.1 Different manufacturing techniques employed for NiTi production—a broad classification (figure reproduced from [1] with permission from Elsevier)

3.1.1 Vacuum Induction Melting (VIM) Technique

The VIM process involves melting of metals/alloys in inert gas or vacuum conditions [2]. This technique is particularly beneficial with metals that show a strong affinity for oxidation [2]. By using eddy currents produced by electromagnetic induction, the metallic charge consisting of Ni pellets and Ti rods housed in a graphite crucible are melted and the electrodynamic forces provide a good homogeneous mixing of the melt [1, 2] (see Fig. 3.2). Due to different melting points of Ni and Ti, the Ni pellets with lower melting point melt first followed by Ti rods that slowly dissolve to form a homogeneous melt over time [2]. The use of additional Titanium cladding elements is to prevent direct contact between Nickel and graphite crucible. The molten charge is then poured into a "shaped steel bowl with Y_2O_3 coating" [2]. Frenzel and co-workers [2] discuss the use of graphite crucibles along with VIM

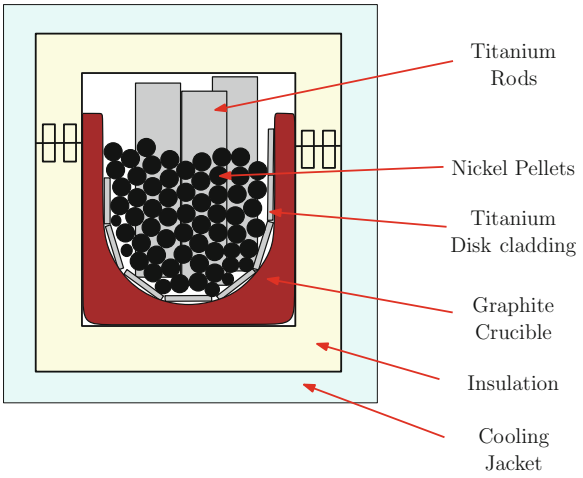

Fig. 3.2 A schematic of a vacuum induction melting setup for processing of NiTi alloys. The use of additional Titanium cladding elements is to prevent direct contact between Nickel and graphite crucible (figure adapted from [1, 2])

due to various characteristics like better electrical properties, good resistance against thermal cracking, good chemical homogeneity, minimal carbon dissolution into melt along with lower costs and easy handling. However, presence of large amounts of carbon in NiTi alloys can form TiC which can increase Ni concentration and thus affecting transformation temperatures [1, 2]. Hence, the carbon content in molten NiTi charges must be under permissible limits (between 300–700 ppm) so that they do not affect any shape memory characteristics of fatigue life [1, 2]. The use of low porosity graphite and using crucibles with lower contact area can significantly reduce extent of carbon contamination [2, 3]. VIM also allows good control on the resulting chemical composition, allows easy process control and degassing options of the final alloy [1, 3].

3.1.2 Vacuum Arc Remelting (VAR) Technique

In case of VAR, the procedure is quite similar to VIM except that this process doesn't use a crucible but uses either consumable or non-consumable electrodes made of materials desired to be melted [1]. In the non-consumable version, the raw materials are sued in a copper mold using a tungsten rod for irradiation by argon arc [1]. In the consumable version, the consumable electrodes themselves are made of the desired raw materials that serves both as heater and raw material which result in higher yields compared to the non-consumable versions [1]. Again, using an argon arc, the electrode is heated and the molten contents are collected in cylindrical ingot [1]. VAR technique can result in pure NiTi alloys if the "melting cycles, material impurity, ingot homogeneity and oxidation" issues are carefully handled [1]. The density difference between the two melts and their extreme reactivity of the melt constituents can be a major drawback in VAR and VIM techniques.

3.1.3 Electronic Beam Melting (EBM) Technique

Some conventional casting techniques like electronic beam melting (EBM) have been employed to produce NiTi since their successful use in refining refractory metals such as Mo, Ta, Nb, and W with high melting points since the 1950s [1, 3, 4]. In such techniques, the use of high vacuum environments and water cooled copper crucibles results in almost zero carbon contamination or oxidation depending on oxygen levels in the initial raw material charges [1, 4]. The Ni and Ti billets are continuously fed under a electron beam path that melts the charge and the molten mixture drops into a "copper extractor" that is water cooled externally [1] (see Fig. 3.3).

However, the control of chemical composition under high vacuum conditions is quite difficult as compared with VIM or VAR techniques and might also result in component evaporation (especially, Ni which has higher vapor pressure than Ti) resulting in lower yields [1, 3, 4].

3.1.4 Conventional/Normal Sintering Technique

NiTi alloys have been manufactured using the popular powder metallurgy techniques either by conventional sintering or combustion sintering [6]. In conventional sintering,

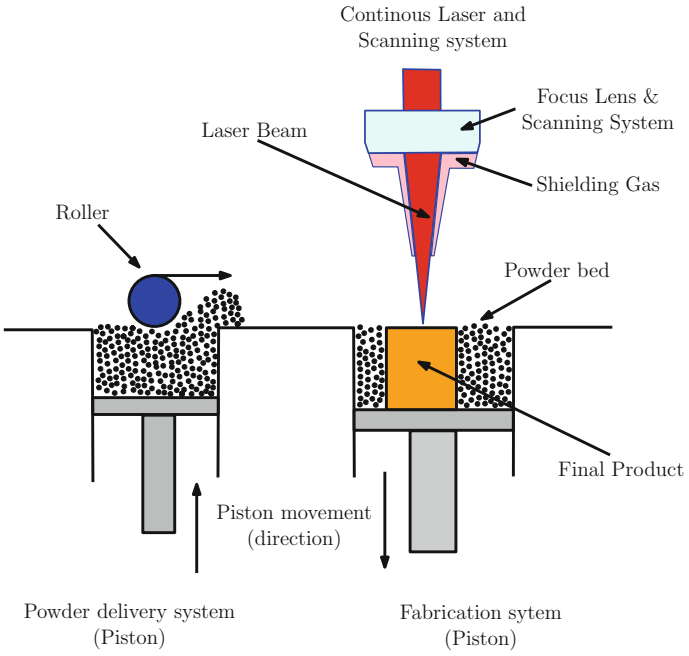

Fig. 3.3 Selective laser melting (figure reproduced from [1] with permission from Elsevier) [5]

the "solid state interdiffusion" of the two constituents is responsible for them to develop intermetallic phases at the interface [6]. The extreme reactivity of the constituents results in formation of melt pool with temperatures reaching 1300 °C in an inert vacuum environment [6]. In case of combustion sintering, the initial compacted mixture is combusted which results in violent combustion of the constituent powders to result in NiTi alloy [6]. The reactions generally occurs at the interface of connecting particles around the eutectic temperature for desired alloy properties [6]. Generally these techniques are time consuming with issues of dimensional stabilities [6].

3.1.5 Selective Laser Sintering (SLS)

Gureev and co-workers have presented a technique of Selective Laser Sintering (SLS) of powder compositions [7]. The setup uses YAG:Nd laser with a focusing lens arrangement to melt the powdered mixture in argon environment [7]. This technique uses the SLS process in cognizance with SHS technique that avoids sudden heating of individual constituents and thus resulting in low exothermic synthesis of NiTi alloys [7]. With such techniques, the surface finish and porosity is uncontrollable [7, 8].

3.1.6 Hot Isotactic Pressing (HIP)

HIP is one of the powder metallurgy techniques for NiTi alloy production from their respective elemental powders [9]. The elemental powders in required proportion are packed in cylindrical gas-tight stainless steel containers subjected to simultaneous hydraulic pressures at elevated temperatures [9]. In some cases, an inert atmosphere like Argon gas is employed in the absence of gas-tight containers which might result in trapped Argon gas bubbles. The use of higher temperatures (generally joule heating) and pressures causes better diffusion of constituents due to partial melting of components [9, 10]. A follow up sintering procedure on the resultant product at reduced pressures can force the trapped Air/Argon bubbles thus resulting in near spherical pores [11]. HIP technique allows superior control over surface finish, final geometry, microstructure homogeneity and porosity of the NiTi samples compared to conventional cintering techniques [9]. However, $NiTi_2$ or Ni_3Ti precipitates in the resultant product cannot be avoided with such a technique [1, 11].

3.1.7 Spark Plasma Sintering (SPS)

SPS similar to HIP uses elemental constituent powders of Cu, Zn, and Al in a graphite mold for manufacturing copper based SMA [10]. As shown in Fig. 3.4, the mixture

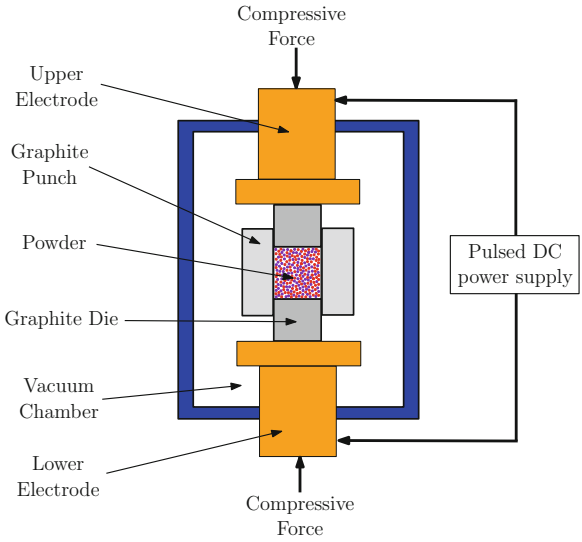

Fig. 3.4 Selective plasma sintering (figure reproduced from [1] with permission from Elsevier)

is heated by repeated DC pulses between electrodes positioned at the top and bottom of die punch. The entire process was conducted in Argon atmosphere with relatively lower mechanical pressures compared to HIP [10].

3.1.8 Selective Laser Melting (SLM)

Shishkovsky et al. [12] looked at using direct SLM of NiTi powder to prepare non-porous samples for medical applications. Nitinol powder containing 99 wt% intermetallic NiTi phase was use to prepare samples with 55 % Ni – 45 % Ti samples [12]. Upon granulometric analysis of particles, the powders were dried upon heating up to 80 °C for 12 h and then cooled in a cooling chamber [12]. The processing was carried out in a SLM machine that employs continuous wave Ytterbium fiber laser in a nitrogen controlled environment [12]. A 50 W power was focused on a 70 µm of the powdered layer and using scan velocities of 100–160 mm/s resulted in SLM samples with 97 % bulk density [12].

In some cases, powdered mixture of individual Ni and Ti powders are milled together with stainless balls in an argon atmosphere under various rotational speeds [13, 14]. As shown in Fig. 3.5, a CO_2 laser beam is scanned over rectangular areas. Using multiple scan and traveling speeds, the width and the length of the scan regions can be achieved [13, 15].

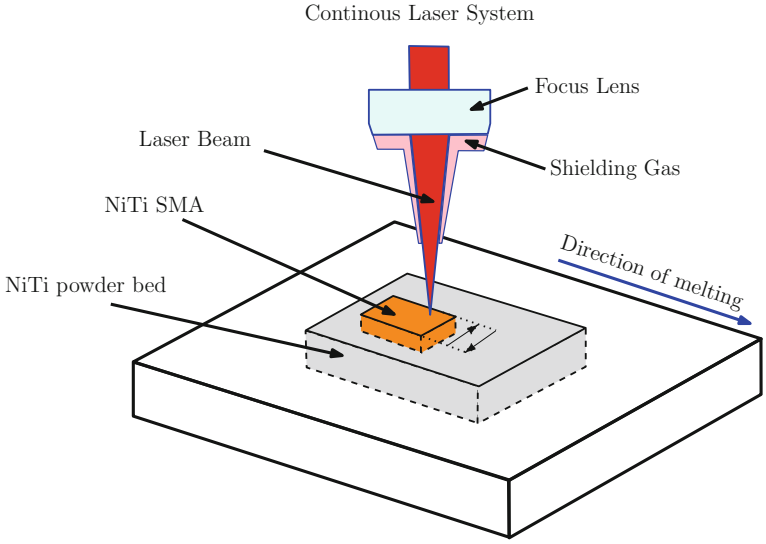

Fig. 3.5 Selective laser melting (figure adapted from [13])

3.1.9 Metal Injection Molding (MLM)

Injection molding has been a popular manufacturing process in plastic industry that injects materials into pre-designed molds. Such an injection molding technique has been used manufacture some "near net shape NiTi products" mostly from powdered constituents [1]. As shown in Fig. 3.6, MIM consists of four major steps that starts with binding of NiTi feedstock, followed by injection molding, then debonding and followed by a sintering process [1]. Such techniques are mainly utilized for cost effective large scale production of smaller parts for the micro-manufacturing industries as they allow for high degree of geometric precision [1].

3.2 Post Treatment of SMAs

NiTi microstructure and surface properties in as cast/manufactured conditions using any of the above conditions is not acceptable to many industries especially the biomedical fraternity [1]. Numerous post-treatment options like cold and hot working, joining, heat treatments, surface treatments are considered to tailor the properties of NiTi as per requirements. The NiTi phase diagram shown in Fig. 3.7 clearly shows that temperatures can greatly influence formation of various precipitates and can greatly influence thermomechanical treatments [1].

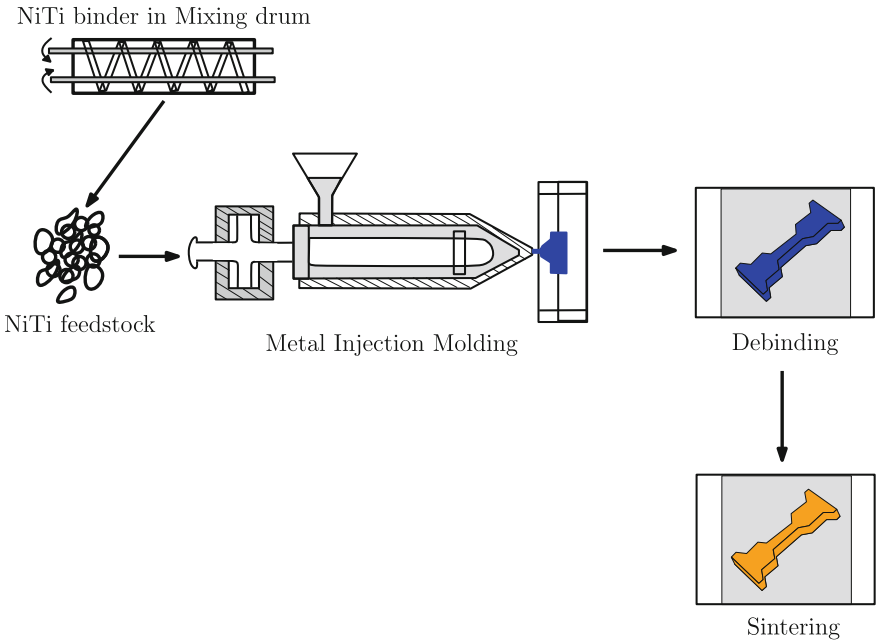

Fig. 3.6 Metal injection molding [16] (figure reproduced from [1] with permission from Elsevier)

3.2.1 Machining of SMA Components

Components such as wires, tubes and sheets are commonly processed by cold working as the ductility of NiTi allows up to 30–50 % of cold work [1]. However, studies have shown that machining of SMAs is quite difficult due to its unique properties especially superior resistance to deformation. Tool wear is a major concern while using conventional machining techniques, hence, techniques like laser/plasma cutting, electric discharge machining (EDM), water jet machining and other abrasive methods like girding and abrasive saws are preferred [1, 18]. Laser cutting techniques like Nd:YAG laser with CNC controls are preffered due to good precision/control over heat input, smaller energy density, minimal/no mechanical stresses, insignificant heat dissipation, no tool wear and good resolutions [1, 18].

3.2.2 Surface Treatment of SMA Components

For many biomedical applications, in order to prevent any Nickel leakage, a pure titanium dioxide (TiO_2) surface coating is added to the component surface. This is achieved by thermal oxidation in low oxygen environments to avoid any Nickel reaction/oxidation [1]. This TiO_2 coating prevent release of any Nickel ions in actual use.

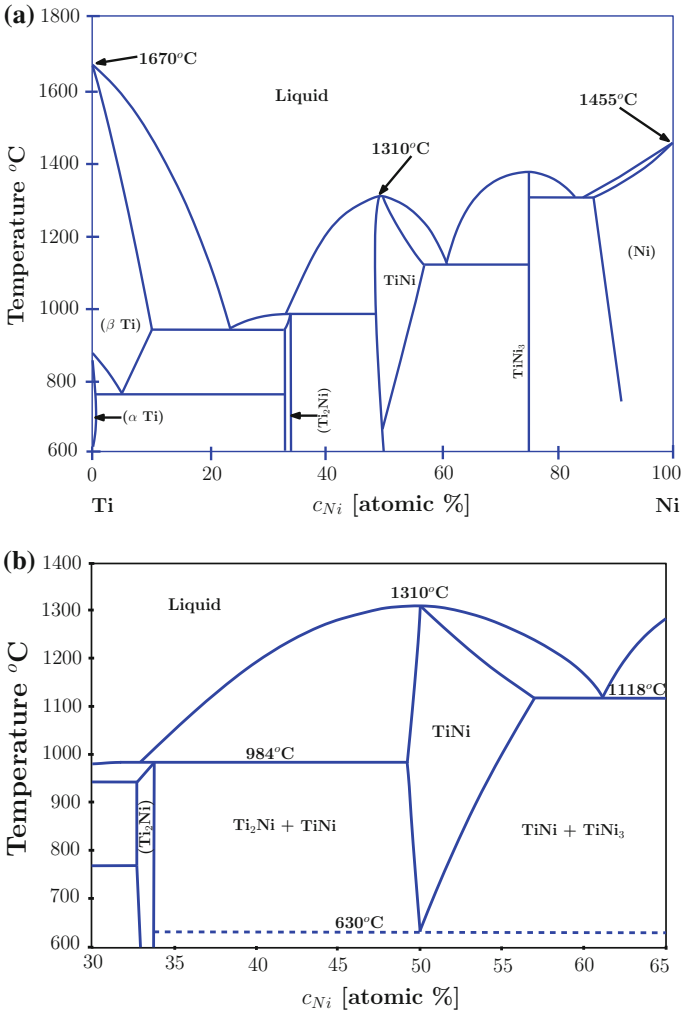

Fig. 3.7 **a** Shows a simplified NiTi phase diagram with the melting points of individual constituents Nickel (Ni) and Titanium (Ti). **b** Shows the central portion of Ni-Ti phase diagram that is of importance to metallurgists dealing with nearly equiatomic NiTi SMA's (both figures are adapted from [2, 17])

However, this oxide layer can cause difficulty in joining especially to dissimilar metals [1] (to be discussed in Sect. 3.2.4 later). Also, fatigue studies have shown that as the bulk SMA undergoes large strains (order of 1–4 % depending on the applied external load), the strain incompatibility between the bulk SMA and outer TiO_2 layer (generally brittle) is quite significant [19]. With repeated thermomechanical cycling, the surface cracks reaching a critical length could lead to final failure [19].

Depending on the application chosen, Nitinol surface finish can vary and the choice of finish can significantly influence fatigue, corrosion and other properties [20]. Many Nitinol surface conditions techniques have been looked at and some important ones are briefly highlighted below [20]:

1. *Oxide surface.* Oxide surface with different colors are achieved by annealing in controlled environments that vary from light blue to dark black coatings.
2. *Mechanically polishing.* Use of abrasive grinding tools to polish surfaces for a aesthetic look [20]. Typically employed to provide a shiny surface on tubes [20].
3. *Chemical Etching.* Use of acid baths provides improved coating adhesion but rougher finish compared to pure oxide surfaces [20]. Such techniques are employed in etching shape set devices after heat treatment process (to be discussed later in Sect. 3.2.5 [20]).
4. *Sandblasted surface.* Use of sandblasting apparatus prior to straightening results in better adhesion properties due to rough finish [20].
5. *Electropolishing.* Many biomedical devices like filters, stents using such polishing techniques to provide notably smooth oxide surfaces that enhances corrosion resistance and biocompatibility [20].
6. *Coatings.* Many polymer coatings like teflon, polyurethane, Polytetrafluoroethylene [20] (PTFE) are available but significantly affect the superelastic responses at higher temperatures [20].
7. *Platings.* Many metals like Copper, Nickel, Silver, Gold, Cobalt, etc. have been plated with Nitinol [20]. However, at higher strains their performance is questionable and issues of hydrogen embrittlement exist during plating operation [20].

3.2.3 Annealing and Coldworking of SMA

One way to change the material properties of the SMA including the yield strengths and the A_f temperature is through aging or annealing of the SMA subsequent to the shape setting operation. In general, annealing will tend to make the A_f temperature approach that of the original ingot A_f. For "high Nickel" alloys (55.8 to 56 % Nickel), annealing at around 300 to 500 °C will cause the A_f to rise (longer the duration higher it rises) but will then decrease above 500 °C. For low Nickel alloys, annealing above 500 °C will cause the A_f to rise towards the ingot A_f. Once, the material is shape set annealed, it demonstrate superelastic and shape memory properties [20].

Commercial Nitnol components are typically subjected to 30–40 % cold work reduction during the last drawing steps [20]. Amount of cold work also influences the ultimate strength of the material. Most commercial SMA components like wires, tubes are shape set annealed condition.

3.2.4 Joining of SMA to Itself and Other Materials Like Stainless Steel

With growing applications of SMAs, the issue of "connecting" SMA components to itself and other materials is of significant importance for designers and application developers. Studies have shown that Nitinol is very difficult to join/attach with other common dissimilar materials with most of the reliable options being mechanical in nature like crimping, swaging or staking [21]. Researchers over the past two decades have focused on the use of more traditional/reliable joining techniques like welding, brazing, soldering or adhesive bonding with SMA materials [22].

Given the ubiquitous use of NiTi among all the commercially available SMAs, NiTi has been the focus of attention for joining research with itself and other materials. Joining NiTi to itself using CO_2 laser, Nd:YAG laser, tungsten inert gas (TIG), and resistance and friction welding under inert atmospheres have been successful with their unique properties being intact even almost after joining.

However, connecting NiTi to dissimilar materials like Stainless steel (SS) has been a challenge using traditional joining techniques like welding or brazing. Such traditional joining techniques have localized effects on the characteristic properties of SMA due to different mechanical and thermal effects [22].

Laser welding is one of the popular welding techniques as they can provide good precision/control over heat input and smaller energy density that results in smaller heat affected zones resulting in minimal distortion and residual stresses [18]. Laser welding is generally accompanied by extremely high heating and cooling rates and this disparity in cooling rates also influences the homogenization process of chemical composition in weld pool that results in non-uniform microstructure and distribution of elements [23].

The disparity of physical and chemical properties between NiTi and SS makes welding challenging [24]. A direct fusion welding of Nitinol to stainless steel results in the formation NiTi, $TiFe_2$, $NiTi_2$, $TiNi_3$ and γ-Fe phases with many brittle intermetallic phases between Ti and Fe namely TiFe and $TiFe_2$ [22, 24]. These phases makes the weld very brittle and often cracks during welding process or shows poor tensile strength and brittle fracture at the weld interface [22, 25]. Prior studies on welding of Ti alloys with other metals such as Fe, Cu, Ni have shown formation of many brittle intermetallics that make Ti extremely unforgiving when it come to welding/brazing as a preferred joining technique [26]. The presence of Ti in NiTi makes chemically very active and readily react with commonly used metals/alloys [26].

To overcome these issues, researchers have tried different filler materials that can potentially change the weld chemistry using different alloy additions and this has been an active area of interest in joining SMA components [25]. The common approach of deciding appropriate filler materials is by studying binary/ternary phase diagrams of Ni, Ti and filler materials [26, 27]. This provides a good indication of element interactions at elevated temperatures and solubility levels that can provide important information on the formation of "brittle intermetallic" phases during the actual welding process [27]. The fact that other common materials don't exhibit

unique characteristics like shape memory effect and superelasticity makes it very challenging to find the appropriate filer material. It is important for the weld fusion zone to have good mechanical strength close to parent NiTi material and good ductility properties that reduce internal stresses during thermal cycling to replicate shape memory effects.

The first choice was to use to Ni filler material to suppress formation of "brittle intermetallic" phases during welding of Nitinol to stainless steel. Li and co-workers [23] suggested an increase of Ni interlayer thickness can suppress the formation of intermetallic phases with Ti like $TiFe_2$ and $TiCr_2$ [23]. Study of weld microstructures contained γ-Fe, B2, B19, $TiFe_2$, $TiCr_2$, $TiNi_3$ and Ti_2Ni phases with dendrite and needle like structures [23]. An increase in Ni content causes increase in γ-Fe, $TiNi_3$ and B19 phases with suppression of brittle intermetallic phases (like $TiFe_2$ and $TiCr_2$) and reduction of gas pores and shrinkage cavities [23]. Generally the concentration of Fe and Cr decreases from the stainless steel side to NiTi side and the concentration of Ni and Ti increased correspondingly as expected [23]. The increase in Ni content at the weld zone results in superior tensile strength of the joint. An increase in γ-Fe also results in superior tensile strength and deformation characteristics [23]. Addition of Ni filler material of the order 30 wt% ensured the weld joint were crack free in spite of extensive plastic deformation during tensile tests [25]. However, large Ni concentration (>60 wt%) can have reverse effect of forming more $TiNi_3$, TiFe and $TiFe_2$ brittle intermetallic phases, pores and shrinkage cavities. Hence an appropriate choice of Ni is important [23, 25]. Failures are generally observed at Nitinol fusion line for the welds with the use of Ni filler material.

Li and co-workers [24] have looked at investigating the effect of Copper filler material on the chemical composition, microstructure and mechanical properties of a dissimilar NiTi alloy and SS laser welded joint [24]. The choice of Copper is because it has lower melting point, lower yield strength with improved ductility that could potentially improve the weld pool and brittleness [24]. Further, the binary alloy phase diagrams indicate that Cu doesn't form intermetallic phases with Fe, Cr and Ni elements. However, Cu tends to form intermetallics with Ti which can be brittle in nature but are not as brittle as those formed with Fe, Cr and Ni [24]. The increase of Cu metal thickness results in higher Cu content at the weld interface that renders the soft Cu solid solutions to be coarser and further results in the suppression of brittle intermetallic phases like $TiFe_2$, $TiNi_3$ etc. and improving the tensile strength of the joints [24]. There is however an upper limit of Cu addition (5.1 %) as too much Cu content could result in additional Cu-Ti intermetallic phases at the weld interface that along with existing intermetallic brittle phases like $TiFe_2$, $TiNi_3$ could reduce the tensile strength of the joints [24]. Microstructural investigations also revealed that the cracks mostly originate from the weld interface which is the weak zone due to presence of brittle intermetallic compound and propagate during tensile testing [24].

Li and co-workers [27] have looked at another filler material possibility of using Cobalt (Co) by studying binary phase diagrams Fe–Co, Cr–Co, Ni–Co and Ti–Co. Phase diagrams give an indication of interactions and solubility of Co with Fe, Cr, Ni, Ti in liquid and solid phases. This information can give an indication of possible compounds that can results in the HAZ. Ti has good solubility in Co that can result

in superior mechanical properties of the joint. Further, Co has a linear expansion between NiTi and SS with good ductility (deformation) and corrosion resistance capabilities. Microstructure investigations of the fusion zone indicates use of Co filler material results in a ductile fracture (combination of cleavage and dimples) against a brittle fracture observed in other NiTi SS welding cases. This can be attributed to the decrease in brittle intermetallic compounds like $TiFe_2$, $TiCr_2$, etc. with Co addition. Too much Co addition can result in Co-Ti intermetallics that can reduce joint strengths.

Though the use of these filler materials have shown improved joint strength and suppressing the formation of brittle intermetaalics with Ti, the important concern is the performance of these joints under superelastic and shape memory conditions [25]. Fatigue life of these joints would also be of grave concern. Since parent NiTi has excellent biocompatibility and corrosion resistance, it is imperative to understand the joint performance under these conditions [25]. A more in depth investigation on these issues and study of the microstructure of the joint (weld metallurgy, phase identification, and so on) is required and these issues are an open area of research with the SMA community [25].

3.2.5 Shape Setting of Nitinol

As we have seen in the previous chapters, Nitinol actuators need to be "formed" or shape set into desired shapes for actuation. While commercial vendors supply wires, springs, tubes and other simple shapes, sometimes it is necessary to buy wires/tubes/sheets and form the appropriate shape yourself. Be warned that "shape setting" of Nitinol is somewhat of a dark art and vendors are secretive about the process.

3.2.5.1 How to Shape Set Nitinol? Some Guidelines

In this section, we describe briefly, how Nitinol can be shape set. The recommendations are based on our own experience with successes and failures in shape setting of springs and available information in the literature. The basic idea is relatively simple. Using a jig or fixture, "lock" the Nitinol into the desired shape at room temperature.

Typical jigs ranges from mandrels for winding SMA springs from wires to pin and plate forms and other such ideas for the shape setting (see some shape setting fixture examples in Figs. 3.8 and 3.9). The jig is placed in a furnace whose stabilized temperature is about 400–500 °C for about 30 min, then remove and water or air cool it. The Nitinol is now shape set or so the theory says. In practice, there are a number of variables to consider.

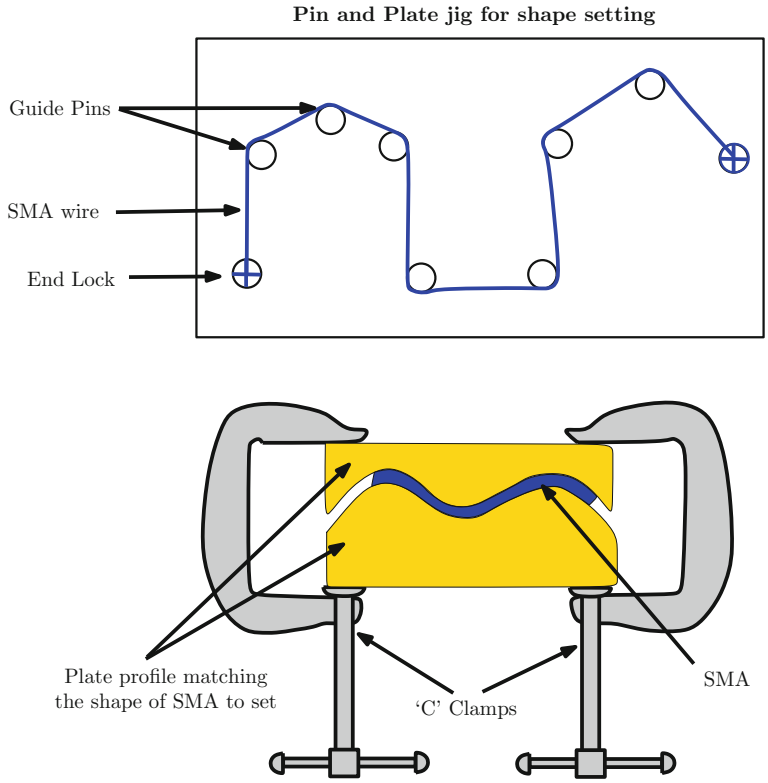

Fig. 3.8 Pin and Plate test rigs to shape set SMA components in desired shapes. Notice the importance of constraining the wire ends to prevent spring back of "martensitic wire" when introduced to an oven (figure adapted from [28])

1. *Condition of the starting sample of Nitinol* Actuator wires (martensitic at room temperature) are commercially available in "straight annealed" or "as drawn" conditions [29][1]. Further, heat treatment may cause the transformation temperatures and other characteristics to be significantly affected. So a trial and error method for shape setting may become necessary (as we found in our experiments).
2. *Note on shape setting fixtures* As shown in Figs. 3.8 and 3.9, fixtures are designed to confine SMA wires into desired shapes. It is important to constrain the ends of

[1] Straight annealed implies that the wire is preset into a straight length as it finishes the drawing operation [29]. They are annealed in a tube furnace so that they are straight. Subsequently they may be cold worked to "adjust" their transformation temperatures typically, cold working raises the transformation temperature. Studies have shown that any heat treatment after cold-working nullifies the effects of cold working by restoring shape memory response in SMAs [1]. Thus a good strategy might be to first shape set at 500 °C and then use annealing to obtain the right A_f. This is still an art and there are no clear guidelines that will guarantee good results. On the other hand "as drawn" wires are directly coiled onto rolls from the final drawing operation [29].

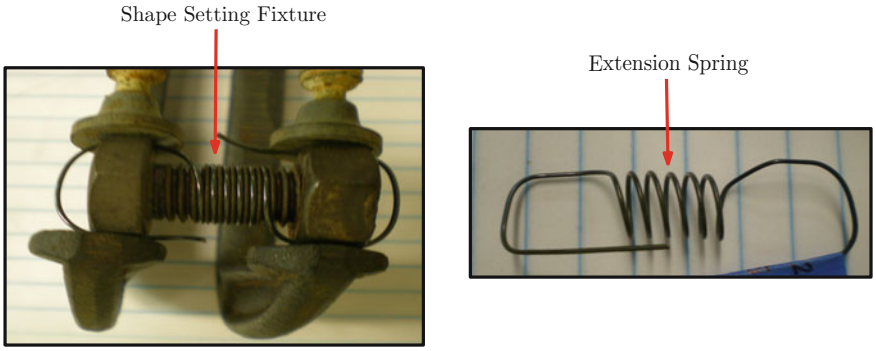

Fig. 3.9 A simple jig used to shape set actuator SMA Nitinol wires into extension springs. Notice the use of heavy metal clamps to constrain the wire ends to prevent unwinding back to initial state when placed in oven

wires (especially actuator wires that are martensitic at room temperature) as they tend to return back to their austenitic state (which could be straight if straight annealed) when introduced in oven maintained at high temperature. It is also advised to avoid sharp spools (i.e., avoid sharp gradients) of SMA wires and restrict the amount of strain to $<3\,\%$ as they lead to pockets of material that do not transform in actual use (for example hooks in extension springs) [1]. This can also improve the fatigue life for shape set samples [1].

3. *Shape setting temperature and soaking times* Generally the temperature range for shape setting/heat treatment is chosen between 400 and 550 °C [29]. The soaking times in ovens vary between 5 and 30 min. Longer soaking times results in increase in transformation temperatures (especially A_f) [20]. The literature recommends steels such as 304, 440, D2, H13, and 17-4PH for the shape setting. Also locking the wire ends surely is crucial. We have found that the use of heavy metal clamps for locking the springs works well (see Fig. 3.9). Temperatures and times involved with the heat treatment/shape setting conditions significantly influence the final microstructure of the shape set product and its mechanical response, transformation temperatures, energy dissipation capabilities and plateau strains [29]. Liu and co-workers [30] have looked at studying the effects of temperatures and times in shape setting and their results showed the decrease of plateau stresses with aging time and temperature. For example, Fig. 3.10 shows the surface plots of variation of upper plateau stresses and ultimate tensile strengths with different heat treatment temperatures and soaking times [29]. It is evident from these plots that heat treatment temperature and time have significant effect on the mechanical response of Nitinol being shape set [29].
   The choice of times and temperature vary depending upon the wire diameter and the mass of the jig being utilized for shape setting. There is unfortunately no reliable guidance for choosing ideal shape setting treatment protocol to ensure the shape set product's reproducible and predictable repeatable mechanical

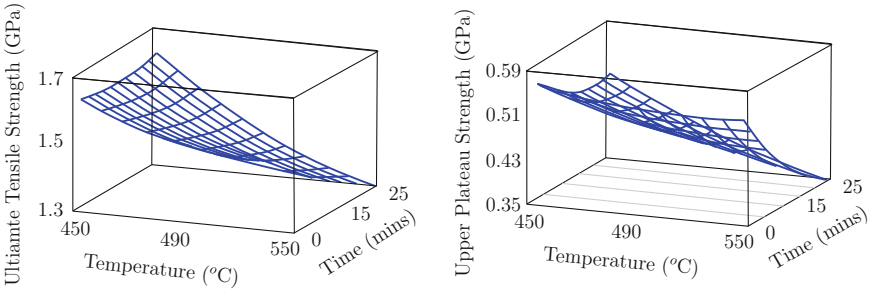

Fig. 3.10 Variation of upper plateau stresses and ultimate tensile strengths with different heat treatment temperatures and soaking times. Graphs adapted from [29]

performance. With appropriate experimentation, the heat treatment temperature/time is finalized for a given application [29]. For larger fixtures, use of thermocouples in contact with material and fixture is recommended to get a better understanding of temperature distribution [20].

4. *Type of furnace* When the Nitinol is shape set, oxide layers may be formed on the surface. To eliminate this, vacuum furnaces (flushed with inert gases as necessary) may have to be used especially when cooling. In this regard, water quenching may be a suitable option. Also care must be taken to ensure uniform heating since otherwise different sections may get different properties. Smith and Hodgson [31] have looked at a variety of different furnaces and fixtures/mandrels for heat treatment purposes that include box and tube furnaces, induction heaters, bath furnaces—fluidized and salt baths, heated presses and vacuum furnaces [1, 31]. The growing conclusion was the oxidation of NiTi in a non-protected environment [1, 31].
5. *Quenching technique* Upon heat treatment in the oven, the entire jig can be either oil or water quenched. Rapid cooling is preferred to avoid aging effects. Studies suggest that the SMA properties are relatively unaffected with the quenching technique followed as long as rapid cooling is ensured.
6. *Functional Fatigue* One of the more vexing problems with the literature on shape setting is that too often the shape set wire is not subjected to repeated cyclic loading. We have found that shape set pieces "forget" their set shape when subject to thousands of thermomechanical cycles unless they have been properly done. So a major task is to ensure that the shape setting is done properly and particularly minimize sharp gradients and large strains during shape setting.

References

1. Elahinia MH, Hashemi M, Tabesh M, Bhaduri SB (2012) Manufacturing and processing of niti implants: a review. Prog Mater Sci 57(5):911–946
2. Frenzel J, Zhang Z, Neuking K, Eggeler G (2004) High quality vacuum induction melting of small quantities of niti shape memory alloys in graphite crucibles. J Alloy Compd 385(1):214–223
3. Otubo J, Rigo O, Neto CM, Mei P (2006) The effects of vacuum induction melting and electron beam melting techniques on the purity of niti shape memory alloys. Mater Sci Eng A 438:679–682
4. Otubo J, Rigo O, Moura Neto C, Kaufman M, Mei P (2003) Scale up of niti shape memory alloy production by ebm. Journal de Physique IV (Proceedings) 112:873–876
5. Gibson I, Rosen DW, Stucker B (2010) Additive manufacturing technologies: rapid prototyping to direct digital manufacturing. Springer, New York
6. Zhang N, Babayan Khosrovabadi P, Lindenhovius J, Kolster B (1992) Tini shape memory alloys prepared by normal sintering. Mater Sci Eng A 150(2):263–270
7. Gureev D, Petrov AL, Shishkovsky IV (1999) Formation of intermetallic phases under laser sintering of powdered shs compositions. In: 6th International conference on industrial lasers and laser applications' 98, pp 237–242. International Society for Optics and Photonics
8. Yang Y, Huang Y, Wu W (2008) One-step shaping of niti biomaterial by selective laser melting. Proc SPIE 6825:68250C–1
9. McNeese MD, Lagoudas DC, Pollock TC (2000) Processing of tini from elemental powders by hot isostatic pressing. Mater Sci Eng A 280(2):334–348
10. Park NJ, Lee SJ, Lee IS, Cho KS, Kim SJ (2004) Manufacturing of cu-15.0 zn-8.1 al shape memory alloy using spark plasma sintering. Mater Sci Forum 449:1109–1112 (Trans Tech Publ)
11. Lagoudas DC, Vandygriff EL (2002) Processing and characterization of niti porous sma by elevated pressure sintering. J Intell Mater Syst Struct 13(12):837–850
12. Shishkovsky I, Yadroitsev I, Smurov I (2012) Direct selective laser melting of nitinol powder. Phys Procedia 39:447–454
13. Kyogoku H, Ramos J, Bourell D (2002) Laser melting of Ti-Ni shape memory alloy. Proceedings SFF symposium, Austin, pp 668–675
14. Bormann T, Schumacher R, Müller B, de Wild M (2012) From powder to complex-shaped niti structures by selective laser melting. Euro PM2012 Proc 1:193–197
15. Bormann T, Schumacher R, Müller B, Mertmann M, de Wild M (2012) Tailoring selective laser melting process parameters for niti implants. J Mater Eng Perform 21(12):2519–2524
16. Krone L, Mentz J, Bram M, Buchkremer H-P, Stöver D, Wagner M, Eggeler G, Christ D, Reese S, Bogdanski D et al (2005) The potential of powder metallurgy for the fabrication of biomaterials on the basis of nickel-titanium: a case study with a staple showing shape memory behaviour. Adv Eng Mater 7(7):613–619
17. Duerig T, Pelton A (1994) Ti-Ni shape memory alloys. Materials properties handbook: titanium alloys, pp 1035–1048
18. Pouquet J, Miranda R, Quintino L, Williams S (2012) Dissimilar laser welding of niti to stainless steel. Inter J Adv Manuf Technol 61(1–4):205–212
19. Ramaiah K, Saikrishna C, Ranganath V, Buravalla V, Bhaumik S (2011) Fracture of thermally activated niti shape memory alloy wires. Mater Sci Eng A 528(16):5502–5510
20. NiTinol (2014) Nitinol university—nitinol devices & components, inc. http://www.nitinol.com/nitinol-university/nitinol-facts
21. Poncet PP (2000) Nitinol medical device design considerations. Strain 2(4):6
22. Gugel H, Theisen W (2009) Microstructural investigations of laser welded dissimilar nickel-titanium-steel joints. In: Proceedings 8th European symposium on martensitic transformations, pp 7–11
23. Li H, Sun D, Cai X, Dong P, Wang W (2012) Laser welding of tini shape memory alloy and stainless steel using ni interlayer. Mater Des 39:285–293

24. Li H, Sun D, Gu X, Dong P, Lv Z (2013) Effects of the thickness of cu filler metal on the microstructure and properties of laser-welded tini alloy and stainless steel joint. Mater Des 50:342–350
25. Hall PC (2002) Laser welding nitinol to stainless steel. In: SMST-2003: Proceedings of the international conference on shape memory and superelastic technologies, pp 5–8
26. Masunaga S, Murata K, Nakamura M, Shiroyama K (1994) Joined parts of ni-ti alloys with different metals and joining method therefore. 25 Oct 1994. US Patent 5,358,796
27. Li H, Sun D, Cai X, Dong P, Gu X (2013) Laser welding of tini shape memory alloy and stainless steel using co filler metal. Opt Laser Technol 45:453–460
28. Case L, Kreiner Z, Trease B (2004) Shape memory alloy shape training tutorial—a teacher's guide to teaching sma shape training. http://www-personal.umich.edu/btrease/share/SMA-Shape-Training-Tutorial.pdf
29. Morgan N, Broadley M (2004) Taking the art out of smart!-forming processes and durability issues for the application of niti shape memory alloys in medical devices. In: Medical device materials: proceedings from the materials & processes for medical devices conference 2003, Anaheim, California, p 247, ASM International, 8–10 Sept 2003
30. Liu X, Wang Y, Yang D, Qi M (2008) The effect of ageing treatment on shape-setting and superelasticity of a nitinol stent. Mater Charact 59(4):402–406
31. Smith S, Hodgson D (2004) Shape setting nitinol. In: Medical device materials: proceedings from the materials & processes for medical devices conference 2003, Anaheim, California, p 266, American Society for Metals, 8–10 Sept 2003

Chapter 4
Basic SMA Component Geometries and Responses

SMA components must allow for a large surface area compared to their volume in order for them to be able to be cooled rapidly and repeated use at reasonable actuation frequencies. Thus SMAs are commonly used in the form of wires/rods, springs, tubes or beams under different loading conditions (tension, torsion or bending) for exploiting their unique characteristics in many practical applications. All these geometries are governed by their high surface to volume ratios. To begin with, a review of different SMA geometries commonly used in different applications and their complex responses are discussed below.

4.1 SMA Wire Response—Tensile Loading

SMA wires under tension are used widely in civil, aerospace and biomedical applications. The phase transformations in the superelastic response of SMA results in non-linear hysteretic response which make them excellent candidates as damping materials due to their superior energy dissipation capabilities [1].

As discussed previously in Sect. 1.4, ASTM standard F2516-07ϵ2 [2] discusses a standard tension test methodology on NiTi superelastic materials and explains the salient features of such superelastic effects. Some important salient features worth noting in a typical superelastic response (see Fig. 1.7) are listed here.

- **Upper plateau strength (UPS)**: Stress at 3 % strain during loading of the sample.
- **Lower plateau strength (LPS)**: Stress at 2.5 % strain during unloading of sample after loading to 6 %.
- **Residual Elongation ($\mathbf{El}_r$)**: The difference between strains at a stress of 7 MPa during loading and unloading operations.
- **Uniform Elongation**: Elongation determined at maximum force sustained by specimen prior to necking or fracture or both.

A. Rao et al., *Design of Shape Memory Alloy (SMA) Actuators*,
SpringerBriefs in Computational Mechanics, DOI 10.1007/978-3-319-03188-0_4

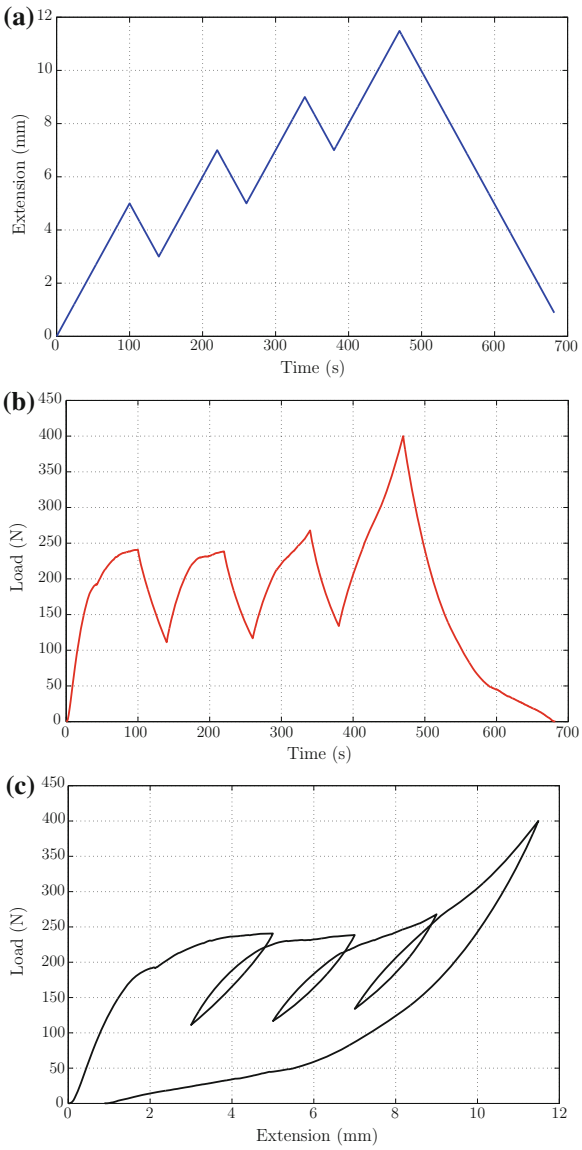

Fig. 4.1 **a** The tensile extension–time, **b** load–time, and **c** load–extension responses for a superelastic response with three loading internal loops. The nonlinear nature of the material response in observed in the transformation regimes (i.e. during loading and unloading stages of the response). (Graphs reproduced from [5] with permission from IOP ©)

Superelastic responses discussed thus far (as shown in Fig. 1.7 earlier) and the ASTM standard F2516-07ϵ2 [2] focus purely on the outer loop response. However, this is not the case in many real world scenarios. If the SMA undergoes intermediate loading and unloading prior to complete transformation, then it can result in smaller hysteretic responses referred commonly as "internal loops" which closely mimic the outer fully transformed loop. A typical SMA wire response is shown in Fig. 4.1. In applications, understanding SMA component responses under partial or fully

transformed cases with internal loops is of particular importance as the entire response might not be utilized always and only a portion of the entire response (internal loop) might be of significance to designers [3, 4]. The area of these smaller hysteretic internal loops depends on the extent of the loading and unloading levels that the component is subjected to during service within the transformation region (plateau region) [4]. A few examples will be discussed in the following sections to highlight these features and illustrate them better.

Consider a SMA wire subjected to a loading sequence described in Fig. 4.1a. The three steps observed here correspond to intermediate loading and unloading sequences that mimic partially transformed (internal loops) responses. In this case, the wire 120 mm long was loaded up to 5 mm extension (∼4.1 % strain) and then unloaded back to 3 mm extension that results in the first triangular step (internal loop 1). Further the sample is loaded up to 7 mm extension and unloaded back to 5 mm extension to observe the second triangular step (internal loop 2). Similarly, sample loaded up to 9 mm extension and unloading up to 7 mm extension results in the third triangular step (internal loop 3). The final loading up to 11 mm extension and unloading back to initial state completes the superelastic response. The corresponding load–time response for the triangular forcing function discussed above is shown in Fig. 4.1b with the same time stamp. The nonlinear nature of the material response is clearly visible during the internal loops (stepped regions) and the unloading part of the response where it is phase transforming between the two phases.

The "outer" loop corresponds to the stress–strain response when the wire is started from a purely austenitic state and is fully transformed to a martensitic phase and is then completely unloaded back to the austenitic phase. The shape of the outer loop is of great importance in the response of the wire. The "inner loop" is the result of partially transforming the SMA wire and returning it back to the austenitic state.

The corresponding load–extension and stress–strain plots are shown in Fig. 4.2. Two important aspects to be noted in this plot is the concept of *return point memory* (*RPM*), that is, the tendency of SMA to return back to the point of unloading on the outer loop and the other is *residual elongation*, i.e. the elongation remaining upon complete removal of loads (representing some untransformed martensite).

RPM provides important information on the ability of SMA components to return back to its original unloading point upon completion of a smaller hysteretic loop (partial transformation) and is of particular importance to designers. SMA components showing good RPM is a desirable feature that indicates minimal residual/irreversible deformations after repeated complete or partial transformations. Figure 4.2b shows the RPM aspect of this wire and it is evident that the smaller hysteretic loop does not completely return back to its original unloading point which eventually results in a residual elongation (inability to completely recover strains). Designers must account for RPM characteristics before using them in actual component design.

Figure 4.3 shows typical response of SMA wires under tension at different temperature regimes. Starting from the bottom and moving upwards, the first curve represents the detwinning of SMA wires at temperatures below M_f followed by unloading to stress free condition and heating to temperature above A_f to complete

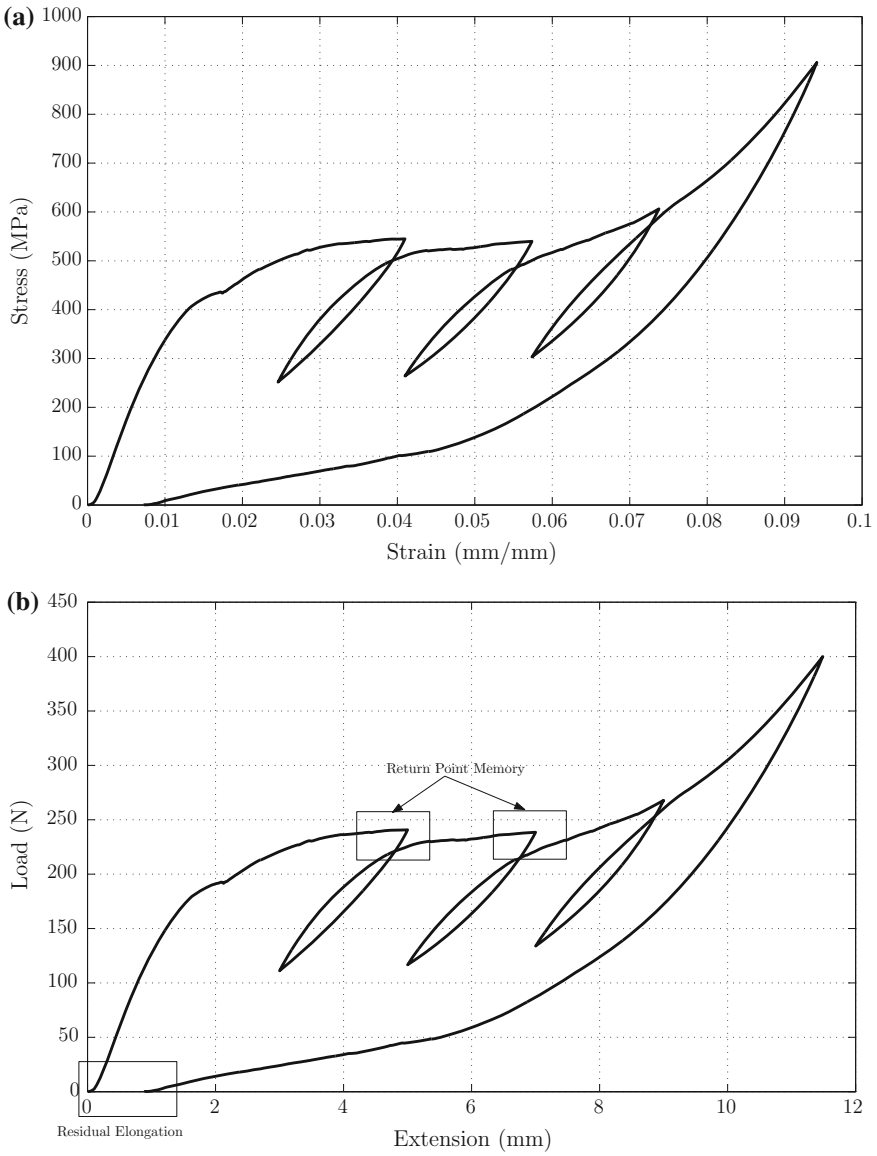

Fig. 4.2 Figures shows the corresponding stress–strain plot for data discussed in Fig. 4.1. (Experimental data reproduced from [5] with permission from IOP ©). **a** Stress–strain plot. **b** Return point memory and residual elongation

the shape memory effect. The second curve represents SMA wires responses at temperatures $M_f < T < A_f$, that is, closer to A_s. There is only partial recovery of strains. The third curve represents a classic superelastic response above A_f. The fourth curve represents temperature dependence of superelastic response. With increase

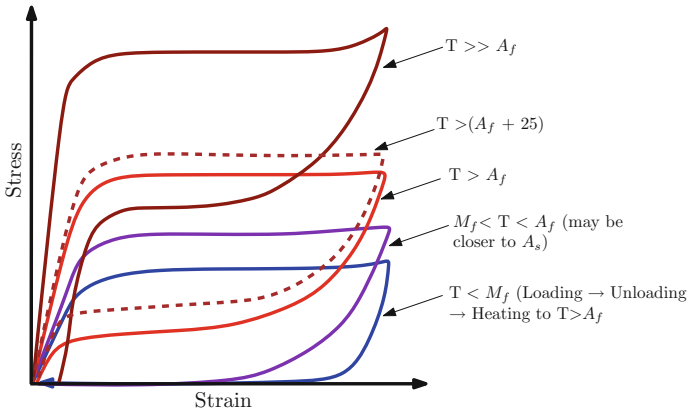

Fig. 4.3 Typical response of SMA wires under tension at different temperature regimes. Starting from the *bottom—First curve* represents the detwinning pf SMA wires at temperatures below M_f followed by unloading to stress free condition and heating to temperature above A_f to complete the shape memory effect. *Second curve* represents SMA wires responses at temperatures $M_f < T < A_f$, that is, closer to A_s. *Third curve* represents a classic superelastic response above A_f. *Fourth curve* represents temperature dependence of superelastic response. *Final curve* represents responses at temperatures far greater than A_f (figure adapted from [6])

in temperature, the critical stresses for phase transformation increase causing shift of response upwards. Generally, shift in plateau stresses is accompanied by stiffer austenitic moduli. If one were to separate the thermoelastic and dissipative part of the response at different temperatures above A_f, then the purely dissipative part of the response overlap over each other as they would roughly be the same. This suggests that the temperature dependence in superelastic response is solely due to the thermoelastic effects and dissipative part of the response is "athermal." The final curve represents responses at temperatures far greater than A_f. One can observe residual strains at these high temperatures due to plastic deformation of martensitic phase and at very high stress levels and some residual pockets of austensite untransformed (more details on this aspect will be discussed in Sect. 5.4 later).

4.2 SMA Wire Response—Torsional Loading

The study of SMA characteristics under torsional loading is of great importance to the design of SMA wires/tubes/springs since these are primarily loaded in torsion. We have carried out torsional experiments on SMA wires at room temperature. The procedure to obtain the response of a typical superelastic torsional loading and unloading of a superelastic SMA wire under isothermal conditions (say, room temperature) consists of twisting the wire up to 1800° twist (loading cycle) and untwisting back to initial state of 0° twist (unloading cycle) as shown in Fig. 4.4a.

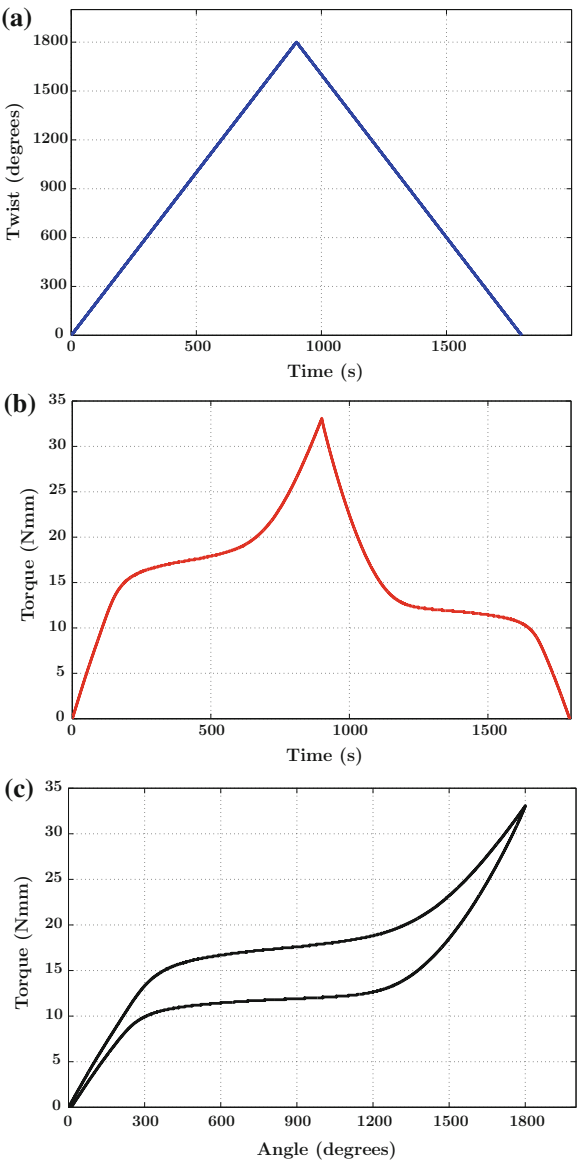

Fig. 4.4 Figures shows the twist–time, torque–time, and torque–twist responses for a simple loading and unloading test up to 1800° maximum twist. The nonlinear nature of the material response in observed in the transformation regime (experimental data reproduced from [4] with permission from Elsevier ©). **a** Twist–time plot. **b** Torque–time plot. **c** Torque–twist plot

The standard commercial torsional test rigs (microtorsion test rigs) are equipped to continuously record torque values for the triangular forcing functions programmed in by the user. The corresponding torque versus time plot is shown on Fig. 4.4b which clearly exhibits a non-linear behavior especially during the phase transformation event. The complete response (i.e., torque versus twist) is shown in Fig. 4.4c. The

near perfect pseudoelastic response indicates minimum residual deformation upon unloading with well defined plateau stresses (UPS and LPS).

SMA components like torque tubes, wires could be subjected to partial twisting–untwisting scenarios similar to tension loading case discussed earlier. Figure 4.5 shows an example of internal loops during the loading and unloading stages of the superelastic response. As shown in Fig. 4.5a, a SMA wire is twisted up to 900° twist (1st loading cycle) and then untwisted back to 400° twist (1st unloading cycle) that leads to a small hysteretic response (internal loop during the loading cycle). The wire sample is further loaded up to 1800° twist (2nd loading cycle) and then unloaded (untwisted) up to 800° twist (2nd unloading cycle) which resembles a classic superelastic response of flat plateaus and large strains/twists as seen previously in Fig. 1.7 [4]. The sample is then loaded up to 1300° twist (3rd loading cycle) and unloaded (untwisted) back to 0° twist (3rd unloading cycle) or the initial state. The nonlinear nature of the response of the material response for the triangular force function (see Fig. 4.5a) during the phase transformation event can be clearly observed in torque–time response as shown in Fig. 4.5b [4]. A much more complex situation is shown in Fig. 4.6 where complex loading–unloading scenarios can be observed that result more complex internal loops of smaller hysteretic responses [4].

4.3 SMA Spring Response—Torsional Loading

The primary role of a spring is to store energy. Springs made of conventional materials like steel, Copper, etc. obey Hooke's law and demonstrate a linear force deflection relationship. With large deformations, conventional springs can undergo plastic deformation that can result in permanent strains. However, SMA springs under torsion have shown the ability to deliver near constant forces over large working strokes thus enabling them to find various engineering applications. In addition, SMA springs that dissipative energy are used to perform useful work. Their response to temperature make them ideal candidates for thermal actuators. Compression or extension SMA springs are the most commonly used in applications.

If a SMA spring is subjected to an axial load P, then the torque applied on any cross section of the spring wire can be computed using the expression

$$T = P\left(\frac{D_m}{2}\right) \tag{4.1}$$

where D_m is the mean diameter of the spring with wire diameter d. If one assumed the angle of twist is uniform over the entire length of the active coils and further ignoring curvature effects, the angle of twist per unit length (ϕ) can be evaluated in terms of the spring axial displacement as

$$\phi = \frac{2\delta}{\pi D_m^2 n} \tag{4.2}$$

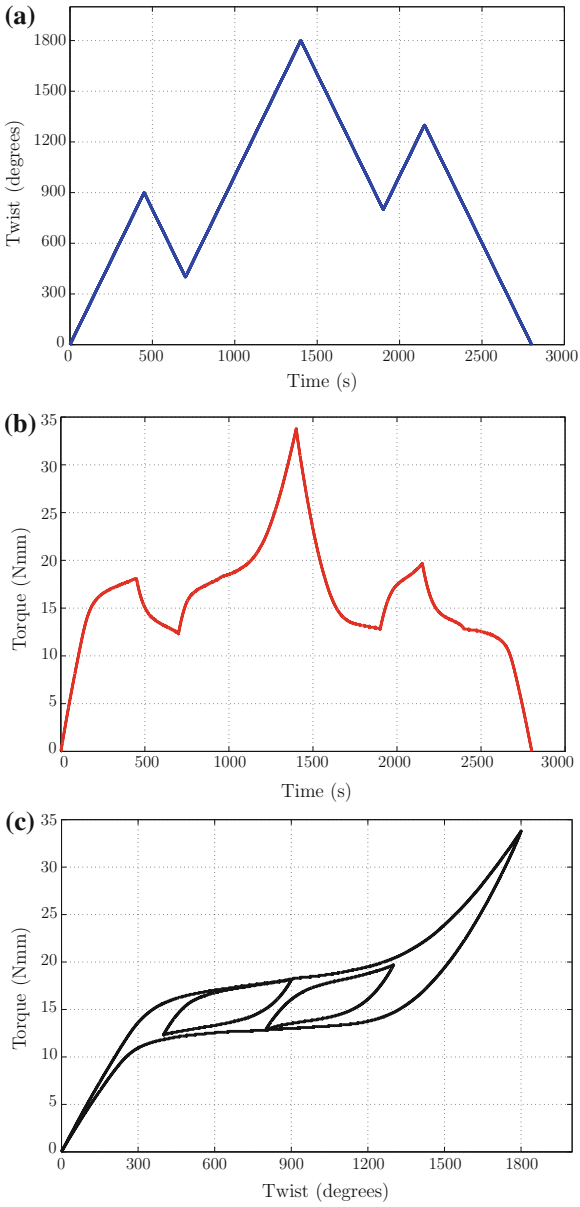

Fig. 4.5 Figures show the twist–time, torque–time, and torque–twist responses for a SMA wire that show internal loops under loading and unloading stages of the response. The overall superelastic response indicates a 1800° maximum twist (experimental data reproduced from [4] with permission from Elsevier ©).
a Twist–time plot.
b Torque–time plot.
c Torque–twist plot

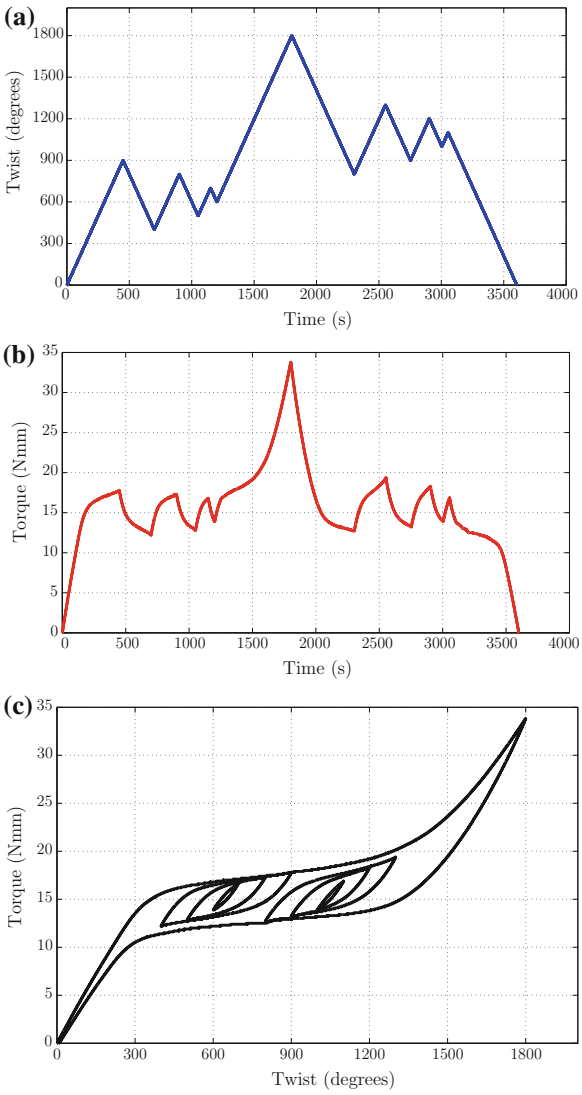

Fig. 4.6 Figures shows the twist–time, torque–time, and torque–twist responses for test up to 1800° maximum twist with complex internal loops during both loading and unloading stages of the response (experimental data reproduced from [4] with permission from Elsevier ©).
a Twist–time plot.
b Torque–time plot.
c Torque–twist plot

The two results [i.e., Eqs. (4.1) and (4.2)] are independent of each other with the former being purely kinetic and the latter being purely geometrical in nature and are based on elementary strength of materials considerations (i.e., that the cross sections rotate as rigid planes around the central line of the wire). No constitutive theory relating the torque to the angle of twist is necessary for these results to be established.

Figure 4.7 shows experiments performed at different temperatures to mimic detwinning and superelastic responses. Tests performed at room temperature (298 K)

to simulate shape memory effect and three other temperatures above A_f viz. 348, 373, and 398 K to observe superelastic behavior of SMA springs.

The force-extension results as obtained from the test setup are shown in Fig. 4.7a with 75 mm being the maximum axial displacement of the spring. The corresponding torque–twist plot using Eqs. (4.1) and (4.2) respectively. The high temperature trials (348, 373, and 398 K) demonstrating near perfect pseudo-elasticity are partially transformed cases where the springs are partially transformed to stress induced martensite (SIM) and the extent of transformation from austenite to SIM is different in each temperature trial. This is due to the fact that the higher the temperature above A_f, the higher is the critical stress required for transformation and thus lesser is the transformation from austenite to stress induced martensite (SIM) when compared against the same maximum deformation (75 mm) of the spring. This is also evident in Fig. 4.7a where the stiffness increases and the hysteresis area decreases with the increase in the working temperature above A_f. The low temperature trail (298 K) discussed in Fig. 4.7a shows a residual elongation of 50 mm which can be completely recovered upon heating to temperatures above A_s.

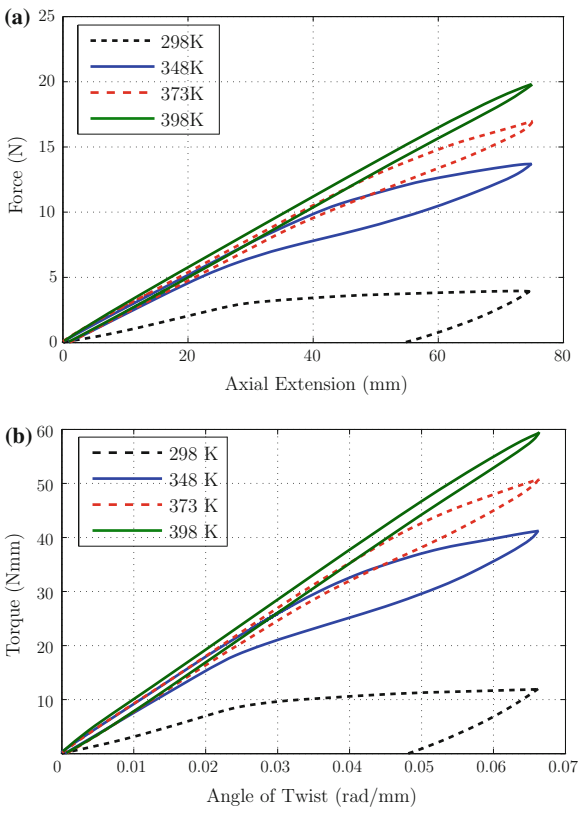

Fig. 4.7 NiTi SMA springs tested at temperatures 298, 348, 373, and 398 K to mimic detwinning and superelastic responses at different temperatures. The extent of transformation in each high temperature trial is different when compared against the same maximum spring elongation (75 mm) or corresponding twist chosen for each trial (experimental data reproduced from [3] with permission from Elsevier ©). **a** Isothermal force versus axial extension results as measured in the test rig. **b** Force versus axial extension of spring converted to torque versus angle of twist using Eqs. (4.1) and (4.2).

References

1. Rao A, Srinivasa AR (2014) A three-species model for simulating torsional response of shape memory alloy components using thermodynamic principles and discrete preisach models. Math Mech Solids. doi:10.1177/1081286514545917
2. A. S. F. 2007e2 (2007) Standard test method for tension testing of nickel-titanium superelastic materials. ASTM International, West Conshohocken
3. Rao A, Srinivasa A (2013) A two species thermodynamic preisach model for the torsional response of shape memory alloy wires and springs under superelastic conditions. Int J Solids Struct 50(6):887–898
4. Rao A, Ruimi A, Srinivasa AR (2014) Internal loops in superelastic shape memory alloy wires under torsion-experiments and simulations/predictions. Int J Solids Struct 51(25):4554–4571
5. Doraiswamy S, Rao A, Srinivasa A (2011) Combining thermodynamic principles with preisach models for superelastic shape memory alloy wires. Smart Mater Struct 20(8):085032
6. Stoeckel D (1995) The shape memory effect—phenomenon, alloys and applications. In: Shape memory alloys for power systems EPR. Nitinol Devices and Components Inc, 47533 Westinghouse Drive Fremont, California 94539, pp 1–13

Chapter 5
Factors Influencing Design of SMA Actuators

Given the complex nonlinear nature of SMA responses, understanding their coupled thermomechanical responses of different components has been of significant interest for both researchers and application developers [1]. SMA component responses in different applications can be significantly affected by several factors like material compositions, component geometries, operating temperature, loading conditions, loading rates, thermomechanical history, processing conditions etc. and mostly combination of several of these parameters in many cases. Since these materials demonstrate complex non-linear responses due to phase transformations resulting in microstructural changes, it is imperative to understand their influence on the overall mechanical response of SMA components. The factors listed below and their affect on the mechanical response is taken up in this chapter.

- Geometry factors
  - Tube/Solid wire
  - Wire diameter
  - Spring index
  - Number of coils
- Material/Alloy composition
- Spring shape setting conditions
- Spring response at different operating temperatures
- Loading rate
- Wire training/Hysteresis stabilization

5.1 Geometry Factors

Commercial NiTi SMA wires are available in solid or hollow cross sections with diameters ranging from 0.1 to 6 mm. When a SMA wire is twisted under torsional

A. Rao et al., *Design of Shape Memory Alloy (SMA) Actuators*,
SpringerBriefs in Computational Mechanics, DOI 10.1007/978-3-319-03188-0_5

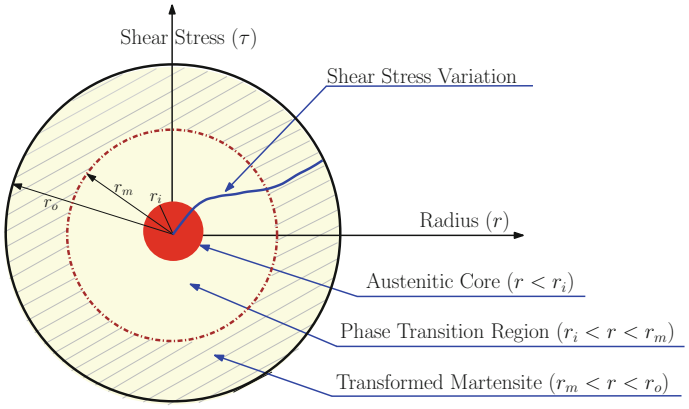

Fig. 5.1 Nonlinear shear distribution across the wire cross section under torsional loading with an inner austenitic core, a phase transition region and an outer transformed martensitic layer (figure reproduced from [4] with permission from SAGE ©)

loading, the phase transformation front gradually moves from the outer fibers towards the neutral axis. The shear strain tends to zero at the core of the specimen cross-section as the wire twists [2]. This implies that the possibility of having a fully transformed case can only be possible if the angle of twist asymptotically reaches infinity [3]. Figure 5.1 is typical partially transformed case seen in wires/springs under torsion. One would observe three zones—an inner austenitic core that remains untransformed due to really small shear strains at the core; a transition region that is undergoing phase transformation from austenite to martensite; a outer martensitic layer that is already transformed from austenite to martensite. The shear stress distribution is highly non-linear across the wire cross section as shown in Fig. 5.1.

The geometry of the wire (solid versus hollow) would play an important role affecting the overall mechanical response. If one compares the response of a hollow and a solid wire with "equal polar moment of inertia" to the same maximum torque, the hollow wire can undergo complete transformation as an untransformed austenitic core (due to really small shear strains at the core) is no longer an issue at higher torque levels. However this is not the case in solid wires which might have an untransformed core at higher torque levels leading to partially transformed cases.

The diameter of the wire also influences the mechanical response. Especially in case of springs, where spring index (ratio of mean coil diameter to wire diameter) can affect the mechanical response greatly. Figure 5.2 shows the effect of Spring Index on the response of SMA springs. Figure to the left shows the response of springs with different spring indexes in its austenitic state (i.e., above A_f). Figure to the right shows detwinning response of springs with same spring indexes (i.e., above M_f). Springs with lower spring index show a stiffer response compared to those with higher spring index as expected. The effect of spring is quite evident from these responses and they play an important role in spring design especially as the

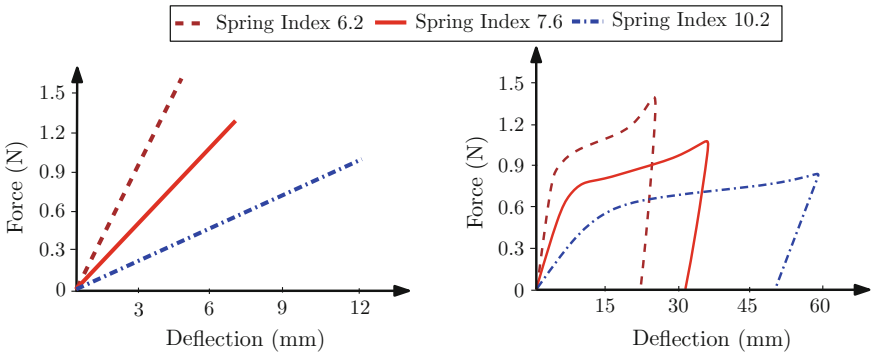

Fig. 5.2 Effect of Spring Index on the response of SMA springs. Figure to the *left* shows the response of springs with different spring indexes in its austenitic state (i.e., above A_f). Figure to the *right* shows detwinning response of springs with same spring indexes (i.e., above M_f). Experimental data reproduced from [5] with permission from IOP Publishing ©

commercial availability of SMA wire in different diameters is quite limited. Hence, a wide range of spring indexes are chosen in SMA spring design (discussed in greater detail in Sect. 7.4 later).

5.2 Effect of Alloy Composition

The most widely used SMA is NiTi due to its superior mechanical properties, corrosion resistance and biocompatibility. The commonly used equiatomic NiTi alloy is martensitic at room temperature with an A_f of around 120 °C [6, 7]. The Nickel content in NiTi SMA's plays a very significant role in determining the characteristic transformation temperatures and slight fluctuations can result in large changes [8]. An increase in atomic percentage of Ni above the equiatomic composition results in a decrease of transformation temperature [6, 7, 9]. Studies have shown that the changes in transformation temperatures with variation in Ni composition is quite sensitive with as much as 10 °C decrease with increase of 0.1 atomic percent from its equiatomic composition [7, 9]. In addition, the presence of elements like Oxygen, Carbon and Nitrogen can affect hysteresis loops, strength and transformation temperatures [8]. Generally decrease of Ni percentage below the equiatomic composition doesn't affect the transformation temperatures significantly [6]. Prior knowledge of alloy composition thus becomes an important parameter for designers as the transformation temperatures play an important role while designing SMA actuators. Engineers designing thermal actuators are mostly concerned with M_f and A_f transformation temperatures for design purposes.

5.3 Effect of Shape Setting Conditions for Custom Shape (Making SMA Springs)

As discussed in Sect. 3.2.5, SMAs can be shape set into particular shapes by certain heat treatment conditions. Temperatures and times involved with the heat treatment/shape setting conditions significantly influence the final microstructure of the shape set product and its mechanical response, transformation temperatures, energy dissipation capabilities and plateau strains [9]. Hence, the initial condition and thermomechanical treatments of SMA components influence the design process and must be accounted for.

5.4 Effect of Operating Temperature on Mechanical Response

SMA components are also sensitive to the working temperature at which they operate. With increasing temperature the plateau stresses keep increasing as seen in Fig. 5.3. This is quite contrary to classical materials like steel, aluminum where one would expect a decrease in yield stress as the dislocations become much more mobile at higher temperatures. However in case of SMAs, the critical stresses for forward and backward transformations increase linearly with increasing temperature. This can be explained by Clausius–Clapeyron equation [see Eq. (5.1)] where $\sigma^{A\leftrightarrow M}$ are the critical stresses during loading cycle, θ being the absolute temperature, $\Delta\eta$ being the entropy change, ΔH being the enthalpy change, υ_o is the molar volume and ϵ_{max} being the transformation strain which is roughly around 6–8 % for NiTi alloys [10] (see Fig. 5.3).

$$\frac{\Delta\sigma^{A\leftrightarrow M}}{\Delta\theta} = \frac{\Delta\eta}{\upsilon_o\epsilon_{max}} = \frac{-\Delta H}{\epsilon_o\theta} \tag{5.1}$$

It must be highlighted that there is a limited range of temperatures above A_f where superelasticity can be observed. This is related to the thermodynamic stability of the two phases under different temperatures and the ease of conventional dislocation glide at higher loads and temperatures. As shown in Fig. 5.4, superelastic effects occur above A_f and the plateau stresses must be less than stresses required to induce slip. Hence superelastic effects are observed only under limited range of temperatures above A_f (highlighted triangular region in the figure) [12]. This can be clearly observed in discussion we had earlier in Sect. 4.1, Fig. 4.3, where at higher temperatures above A_f the critical stress for transformation exceeds the stress levels inducing slip and thus the resultant remnant plastic strain (irreversible permanent changes) at the end the loading–unloading cycle.

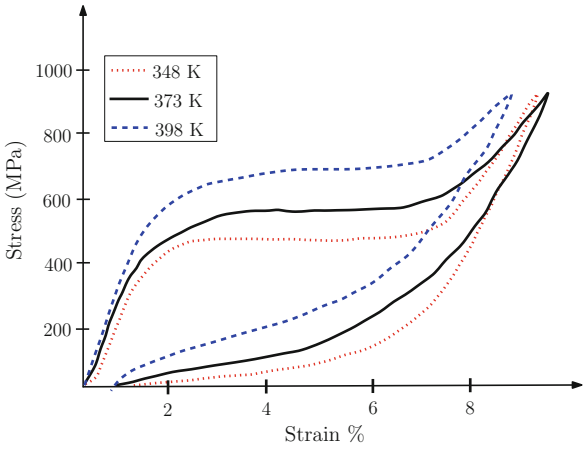

Fig. 5.3 Figure shows the effect of operating temperature on the upper plateau stresses in SMAs. Data is for an SMA wire under tension loading conditions [11]

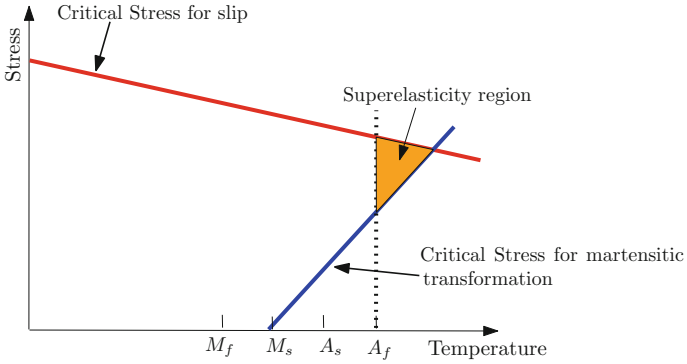

Fig. 5.4 Superelastic effects occur above A_f and the plateau stresses must be less than stresses required to induce slip. Hence, superelastic effects are observed only under limited range of temperatures above A_f, as highlighted by the triangular region (figure adapted from [12])

5.5 Effect of Loading Rates

At lower strain rates, the plateau stresses are relatively flat as there is larger time for martensitic bands to nucleate and propagate during the phase transformation event at lower loading rates [13, 14]. Further, the latent heat at different strain rates induces macroscopic and localized temperature changes with a significant effect of plateau stresses for phase transformation. Figure 5.5 shows the effect of loading rates (order of magnitude effect) on the plateau stresses, overall energy dissipation (hysteresis) and amount of plastic strain accumulation (over cycles).

At faster loading rates, the phase transformation changes from a diffusionless or shear-like mechanism to a more classical plasticity-like dislocation slip [13, 14]. This means an increase in plateau stresses with an increase in loading rate, and the response looks similar to a hardening response as observed with other classical materials. At

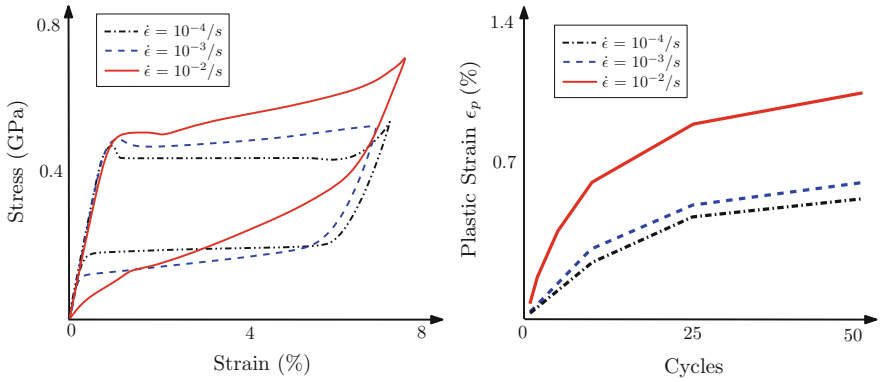

Fig. 5.5 Effect of loading rate on plateau stresses and accumulation of plastic strains over cycles with increase in strain rates. Experimental data adapted from [13]

higher strain rates, more martensitic bands nucleate and propagate simultaneously leading to higher thermal energies and plastic strains due to mismatch of these bands as seen in Fig. 5.5 [13, 14]. Also, studies have shown that with an increase in number of cycles at higher loading rates, the net dissipated energy (hysteresis) significantly decreases [13, 15]. The plateau stresses and dissipation energy show a stronger sensitivity towards higher temperature than higher strain rates [15].

5.6 Wire Training/Hysteresis Stabilization

The martensitic phase transformations seen in SMAs result in a hysteretic loading and unloading response. The hysteresis is not a result of dislocation glide seen in plastic deformation of classical materials but is due to the twinning deformation of martensite that occurs as a result of nucleation, rotation, and shrinking of individual martensite variants (24 of them) [16]. All of these variants transform back to a single parent austenite phase upon reverse transformation. During repeated mechanical or thermal cycling of SMAs, one can see changes in macroscopic material response due to these changes in microstructures [6]. This is generally observed as remnant plastic strain (irreversible permanent changes) at the end of each loading and unloading that can accumulate over several loading cycles (also referred to as ratcheting response in plasticity literature[1]) and also in "non-closure of hysteresis loops" [6, 17]. The material response is considered stable if the hysteresis for several subsequent load cycles are stable (i.e., show no drift) as shown in Fig. 5.6. This "stability"

[1] In classical plasticity literature, ratcheting refers to failure to close/stabilize a hysteretic loop with measured strains creeping (or ratcheting) to the right with every cycle. If one were to plot maximum strain values against the number of cycles, it would represent an increasing trend. This is generally due to accumulation of plastic strains as the dislocation density increases.

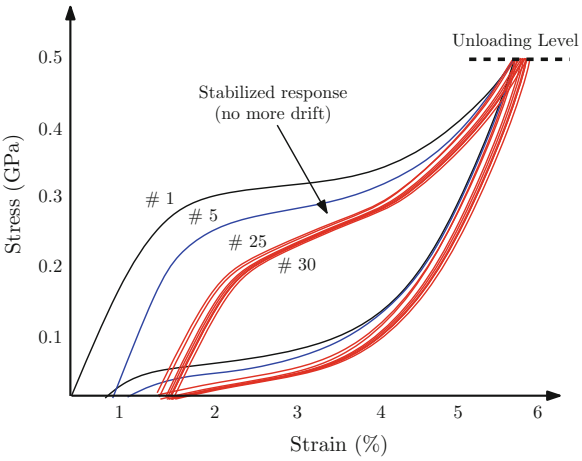

Fig. 5.6 A typical stabilization process observed in superelastic responses of SMAs. One can clearly observe accumulation of plastic strains over initial cycles and reaches a "stabilized state" where there is no more drift in hysteresis and accumulation of plastic strains. Notice the same unloading level for each loading cycle [18]

in response is critical from an application standpoint and deliberate introduction of defects/dislocations during the thermomechanical processing helps in stabilized superelastic responses [17]. Stabilization can be affected by various factors such as processing history, effect of loading conditions, and number of service/loading cycles [17]. Studies have shown that stability is achieved if a material is repeatedly loaded to same maximum load levels. This is due to inherent history dependency of the material response, that is, the crytallographic changes from austenite to different martensite variants over cycles and then eventually "stabilizing" [17] (see Fig. 5.6). If one were to change the loading level (say use higher loads), then additional cycles are required to achieve stabilized hysteretic responses at these load levels.

Training of SMA component (typically in actuator components that are martensitic at room temperature) refers to repeated loading and unloading in a well defined thermomechanical cycle that results in saturated hysteretic response [6]. Generally SMA wires or springs are subjected to known thermal cycles under constant dead loads that allows the material to stabilize and saturate the amount of permanent strains [6]. This process of SMA components undergoing repeated thermomechancial cycling in often termed as *two-way shape memory effect (TWSME)* [6]. This stabilized hysteretic response is often used to design actuators to avoid any drift during its operation and is often interpreted as "imperfect memory" requiring multiple cycles to become "perfect" [17].

References

1. Shaw JA, Kyriakides S (1995) Thermomechanical aspects of niti. J Mech Phys Solids 43(8):1243–1281
2. Aguiar R, Savi M, Pacheco P (2010) Experimental and numerical investigations of shape memory alloy helical springs. Smart Mater Struct 19:025008

3. Doaré O, Sbarra A, Touzé C, Moussa M, Moumni Z (2012) Experimental analysis of the quasi-static and dynamic torsional behaviour of shape memory alloys. Int J Solids Struct 49(1):32–42
4. Rao A, Srinivasa AR (2014) A three-species model for simulating torsional response of shape memory alloy components using thermodynamic principles and discrete preisach models. Math Mech Solids. doi:10.1177/1081286514545917
5. An S-M, Ryu J, Cho M, Cho K-J (2012) Engineering design framework for a shape memory alloy coil spring actuator using a static two-state model. Smart Mater Struct 21(5):055009
6. Lagoudas D (2008) Shape memory alloys: modeling and engineering applications. Springer, Berlin
7. Buehler WJ, Gilfrich J, Wiley R (1963) Effect of low-temperature phase changes on the mechanical properties of alloys near composition tini. J Appl Phys 34(5):1475–1477
8. Elahinia MH, Hashemi M, Tabesh M, Bhaduri SB (2012) Manufacturing and processing of niti implants: a review. Prog Mater Sci 57(5):911–946
9. Morgan N, Broadley M (2004) Taking the art out of smart!-forming processes and durability issues for the application of niti shape memory alloys in medical devices. In: Medical device materials: proceedings from the materials and processes for medical devices conference 2003, ASM International, Anaheim, California, p 247, 8–10 Sept 2003
10. Van Humbeeck J (2003) Damping capacity of thermoelastic martensite in shape memory alloys. J Alloy Compd 355(1):58–64
11. Doraiswamy S, Rao A, Srinivasa A (2011) Combining thermodynamic principles with preisach models for superelastic shape memory alloy wires. Smart Mater Struct 20(8):085032
12. University of Cambridge (2014) Dissemination of it for the promotion of materials science (doitpoms). http://www.doitpoms.ac.uk/tlplib/superelasticity/limits.php
13. Kim K, Daly S (2011) Experimental studies of phase transformation in shape memory alloys. In: Mechanics of time-dependent materials and processes in conventional and multifunctional materials, vol 3, Springer, New York, pp 81–87
14. Pieczyska E, Gadaj S, Nowacki W, Hoshio K, Makino Y, Tobushi H (2005) Characteristics of energy storage and dissipation in tini shape memory alloy. Sci Technol Adv Mater 6(8):889–894
15. Nemat-Nasser S, Guo W-G (2006) Superelastic and cyclic response of niti sma at various strain rates and temperatures. Mech Mater 38(5):463–474
16. Graesser E, Cozzarelli F (1991) Shape-memory alloys as new materials for aseismic isolation. J Eng Mech 117(11):2590–2608
17. Khandelwal A, Buravalla V (2009) Models for shape memory alloy behavior: an overview of modeling approaches. Int J Struct Changes Solids 1(1):111–148
18. Eggeler G, Hornbogen E, Yawny A, Heckmann A, Wagner M (2004) Structural and functional fatigue of niti shape memory alloys. Mater Sci Eng A 378(1):24–33

Chapter 6
Graphical Description of Temperature Controlled Actuation of SMA Wires

In this chapter, we will demonstrate how the SMA wires and springs are actuated and the way a given stroke is achieved by means of temperature changes. We wish to emphasize that the approach presented here is purely graphical and provides a good insight into how the SMA actuator operates. The aim here is to give a the reader a "feel" for how the actuation occurs. In subsequent chapters we will discuss quantitative examples in more detail. Without the understanding provided in this chapter it will be hard to follow calculations.

6.1 SMA Wire + Bias Spring Arrangement

To illustrate this graphical design approach, let us consider a simple case study using a combination of SMA wire and bias spring arrangement as shown in Fig. 6.1. In order to design such a setup, it involves finalizing the following important items:

1. SMA wire geometry—wire diameter; wire length; number of wires
2. Bias spring selection
3. Actuation temperature—control system

We will be discussing the selection of SMA wire geometry in greater detail in the next chapter. In this chapter, we will focus on the effect of temperature on the SMA response and how can one estimate the strokes graphically.

6.1.1 SMA Wire Selection

Commercial SMA wires are available in specific diameters with limited thermomechanical processing and surface finish options due to manufacturing and processing difficulties. Dynalloy®, one of the leading suppliers of SMA actuator wires, provide a rough estimate/guideline of the recovery force estimates against different electrical currents [1]. Dynalloy claims that the mechanical load and the temperature heavily

A. Rao et al., *Design of Shape Memory Alloy (SMA) Actuators*,
SpringerBriefs in Computational Mechanics, DOI 10.1007/978-3-319-03188-0_6

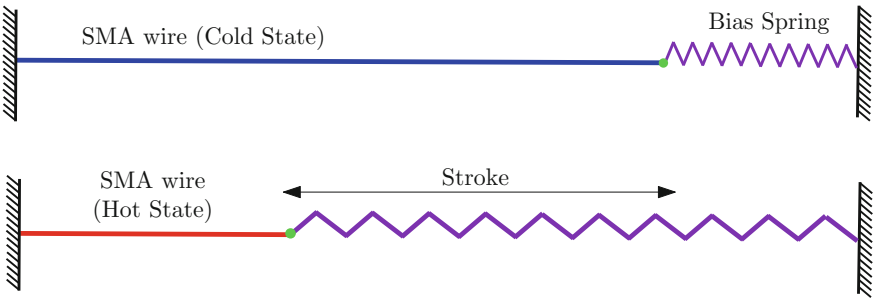

Fig. 6.1 Figure shows an arrangement consisting of SMA wire in series with a steel bias spring. SMA wire here acts as an sensor and actuator with changes in temperature. Bias spring resets the mechanism back to its original position for next cycling

influence the permanent set in the wire and hence influencing the material influence greatly [1]. Table 6.1 provides a reasonable estimate of recovery/pull force for different wire diameters. According to the web site, the heating pull forces were based on constant "safe stress" of 172 MPa and similarly the cooling pull force based on constant "safe stress" of 70 MPa.

A useful plot of heating pull force versus wire diameter is shown in Fig. 6.2. One can get a reasonable estimate of either the wire diameter or corresponding pull force if one of them is fixed based on the end application. SMA wires can recover 6 or 8 % strains under ideal conditions but we limit it to under 3–4 % strains due to fatigue constraints.

Table 6.1 SMA wire selection—wire diameter versus pull force estimate. Data reproduced from [1]

Diameter (mm)	Resistance $\left(\frac{\Omega}{m}\right)$	Heating pull force (N)	Current estimate for 1s (mA)	Cooling pull force (N)
0.025	1425	0.09	45	0.04
0.038	890	0.2	55	0.08
0.05	500	0.35	85	0.14
0.076	232	0.78	150	0.31
0.1	126	1.4	200	0.56
0.13	75	2.186	320	0.87
0.15	55	3.14	410	1.26
0.2	29	5.58	660	2.24
0.25	18.5	8.73	1050	3.49
0.31	12.2	12.55	1500	5.02
0.38	8.3	22	2250	8.83
0.51	4.3	34.91	4000	13.96

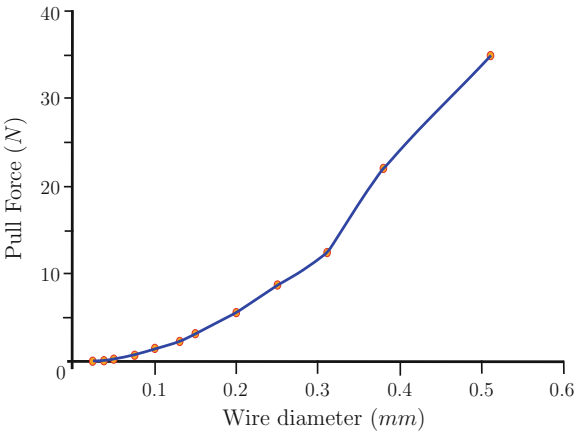

Fig. 6.2 Graph showing pull force (heating cycle) versus wire diameter. Data plotted from Table 6.1 [1]

6.1.2 Operating Temperature of SMA Wire

The operating temperature of SMA wire also plays an important role as SMA components tend to show a temperature dependence in their mechanical responses. As shown in Figs. 1.7 or 5.3, above temperatures of A_f, the plateau stresses increase linearly with temperature. If one were to consider the UPS using the ASTM standard discussed in Fig. 1.7 at different temperatures, a linearized version of the material response at different temperatures can be constructed. In subsequent sections, the effect of operating temperature on stroke estimation will shown graphically under different case studies.

6.2 Graphical Design Approach for Stroke Estimation

A typical SMA wire response in its martensitic or austenitic state showing the detwinning (shape memory effect, that is, at temperatures $< M_f$) and superelastic responses (at temperatures $> A_f$) is shown in top part of Fig. 6.3. Now, using the ASTM standard F2516-07ϵ [2] discussed earlier in Fig. 1.7 some "salient features" of the superelastic response can be extracted for design studies. For a linearized approximation limited to loading part of the response, one needs the following information

- Austenitic moduli E_A that represents the initial slope of the superelastic response
- The critical stress at which the transformation occurs in order to account of flat plateaus in superelastic responses.
- Martensitic moduli E_M that represents the final slope of the superelastic response equivalent to elastic deformation of detwinned martensite or stress induced martensite (SIM).

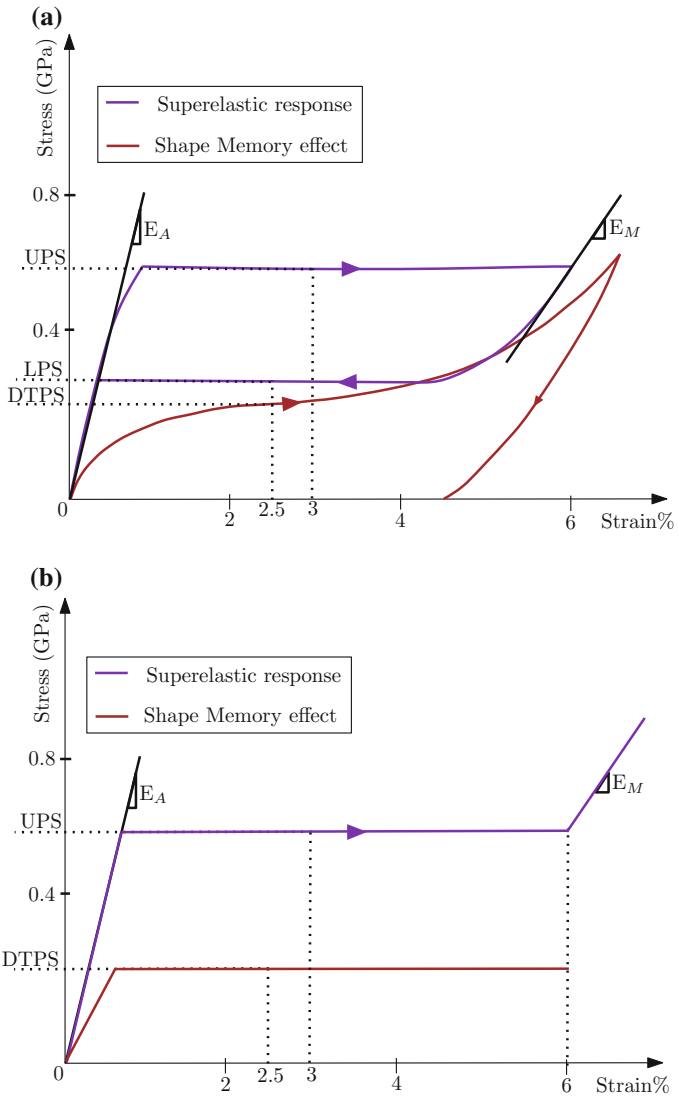

Fig. 6.3 **a** The *upper* part of the figure shows SMA wire response at temperatures $< M_f$ showing the shape memory effect and superelastic responses at temperatures $> A_f$. **b** The lower part of the figure shows the linearized approximation of experimental responses (considering only the loading part) based on ASTM standard F2516-07ϵ [2] DTPS is Detwinning Plateau stress corresponding to 2.5 % Loading. Also refer Fig. 1.7 for more details. Experimental data is taken from [3]

ASTM standard F2516-07ϵ [2] provides information in measuring all these three parameters described above.

- E_A is the slope of the initial response of the superelastic response starting from the origin.
- The critical stress or the **Upper plateau strength (UPS)** is calculated using the stress level corresponding to 3 % strain during loading response. These stress levels serves as the linear approximation for the start of flat plateaus (loading level).
- Martensitic moduli E_M is the final slope of the superelastic response (after 6 % flat plateaus). This is generally not important as the strains are restricted to <(3–4 %) due to poor fatigue life.

Unfortunately, there is no ASTM standard for the detwinning/shape memory effect (SME) responses. We propose using stress levels corresponding to 2.5 % strain during loading as a linear approximation measure to determine the corresponding martesnitic plateau as shown in the bottom half of Fig. 6.3. The martensitic moduli is the initial slope of the detwinning response similar to the superelastic response. Using this information, the linearized approximate plots for the superelastic (limited to loading response) and the detwinning response (shape memory effect) are described in the bottom half of Fig. 6.3.

6.2.1 Graphical Design Approach for Stroke Estimation—Load and Displacement Controlled Tests

SMA wire can be used under constant displacement or constant load or in a simultaneous load–displacement case. Figure 6.4 discusses linearized plots for a constant load or constant displacement situation. The linearized approximation here represents only the loading cycle (single switch event between austenite and martensite) and not the entire hysteretic response that typically includes both loading and unloading aspects.

The top half of Fig. 6.4 shows SMA wire loading responses (linearized approximations) at three different temperatures above A_f and a detwinning response at temperature below M_f. If a SMA wire under constant external load (F) is heated from its martensitic state, the SMA wire contracts back to its austenitic state and delivering a large stroke. The net stroke under the same external load can be varied by changing the temperature up to which the wire is heated. The circular points on the responses between the cold martensitic state and hot austenitic state illustrates this point of variable stroke effects with changes in operating temperature. The net stroke (strain) is the difference between the cold and high states (temperatures),

$$\Delta\epsilon = \mid \epsilon_c - \epsilon_h \mid \tag{6.1}$$

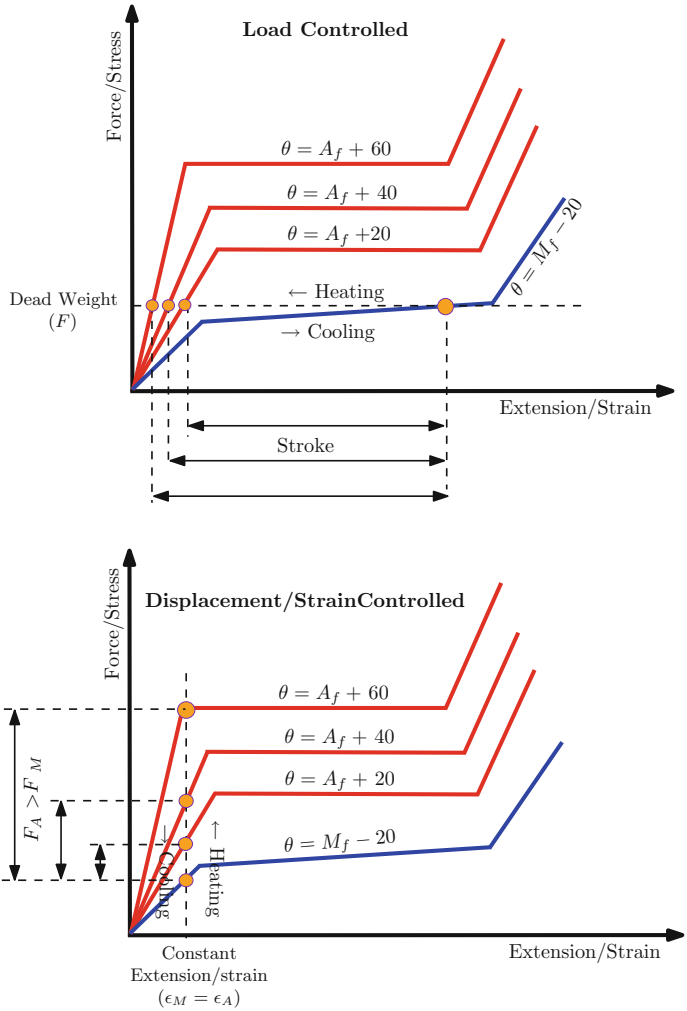

Fig. 6.4 Figure shows the linearized version of SMA wire response in its austenite and martensite state. The *red lines* depict the SMA wire response at different temperature above A_f. The *blue line* depicts the martensite detwinning response at temperature below M_f. By changing temperature, one can achieve large strokes or recovery/pull forces depending on the loading state (i.e., constant load or displacement state)

Similarly, in a displacement controlled setup, a SMA wire under constant strain/deformation can provide a differential recovery/pull force under different operating temperatures. As shown in Fig. 6.4, the net recovery (pull) force between the martensitic and austenitic state can be varied by changing the temperature up to which the wire is heated.

6.2.2 *SMA Wire + Bias Spring: Graphical Design Approach for Stroke Estimation Using Linearized Loading Response only*

SMA wires are rarely used in isolation and are generally accompanied by other passive structural members to reset the SMA wire back to its original state when cooled. For example, a simple arrangement of SMA wire and bias spring combination is shown in Fig. 6.1 where the SMA wire must overcome the bias spring stiffness for the system to actuate.

State **A** in Fig. 6.5 shows the initial state with SMA wire in its cold state (i.e., at temperatures $< M_f$) with the bias spring partially deformed to balance the martensitic SMA wire stiffness. The martensitic SMA wire stiffness can be evaluated using the relation $k = \left(\frac{E_M A}{L}\right)$, where k is the martensitic stiffness, E_M is the martensitic moduli around 30–40 GPa, A is the wire cross section, and L is the length of SMA wire.

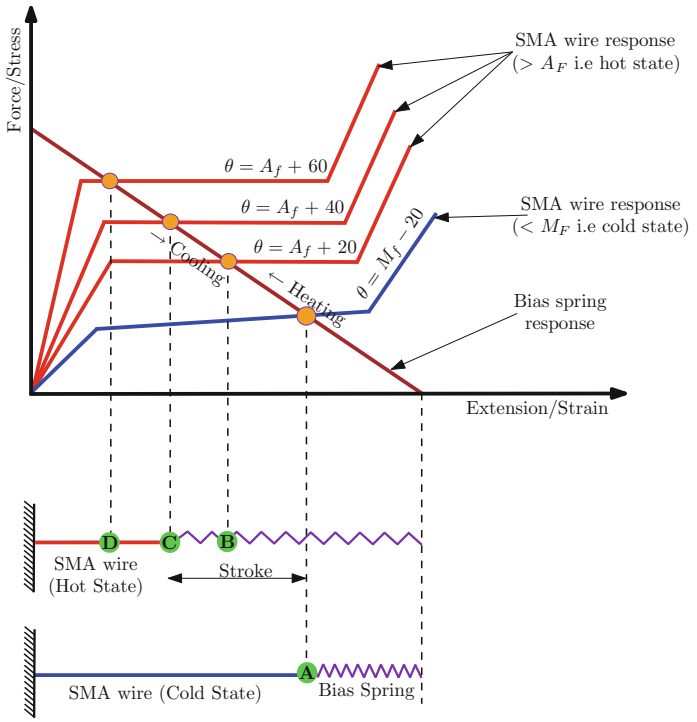

Fig. 6.5 Figure shows an arrangement consisting of SMA wire and steel bias spring with their response overlapped on a force/stress—extension/strain plot. The linearized SMA wires responses at both hot and cold states along with the bias spring response are overlapped for stroke estimation

When the temperature of the system is raised above A_f, the system will acquire a position **B** where the SMA wire is in its austenitic state causing a recovery pull force which forces the extension of bias spring. State **B** can either be **C** or **D** depending on the temperature to which the SMA wire is heated to above A_f. Upon cooling to temperatures below M_f, the system returns back to position **B** (i.e. original state) where the extended bias spring overpowers the stiffness of the martensitic SMA wire. The response of the bias spring is linear (obeying Hooke's law) and depicted by the brown slanting line that intersects the SMA wire responses at different points. The stiffness of the bias spring determines the slope of brown line and the starting point of the line on the deformation/strain axis. The net stroke is difference between state **A** and state **B/C/D** depending on the temperature of the SMA wire above A_f. Such a graphical representation is a convenient way of reasonably estimating the stroke in such (SMA wire + bias spring) arrangements.

6.3 Case Study 2: Linear to Rotary Arrangement Using a SMA Wire + Bias Spring Arrangement Using Linearized Loading Response Only

Using the information discussed in the above sections, a simple linear to rotary mechanism using SMA components is discussed in this section. The goal here is to rotate the link by 90° to demonstrate a simple linear to rotary motion.

This design would involve identifying the following steps:

1. Step 1 (Component Design/Selection): Identifying mechanism(s) for linear to rotary motion; appropriate SMA wire selection; Bias spring selection
2. Step 2 (Configuration): Finalize mechanism using choices from step 1
3. Step 3 (Geometry): Fixing geometry—SMA wire details (i.e., wire diameter, length and number of wires); Bias spring design; Mechanism parts design

Several different linear to rotary mechanisms are available and in this case the simplest linear–rotary crank mechanism that one can start with is shown in Fig. 6.6. It consists of a SMA wire and bias spring opposing each other connected to one end of the link. The other end of the link is connected to the spindle/shaft of the device that needs rotation (for example, rotating a door spindle by 90°).

The given data would be the torque required to rotate the spindle/shaft. Lets consider 160 N mm as the motor torque rating used for this application. This torque value will be used as a guideline for designing link, SMA wire diameter and bias spring stiffness. As shown in Table 6.2, different combinations of link lengths and force (i.e. the torque arm and force) delivering the required torque can be picked. The net stroke $S = \mathrm{r} \times$ rotation depends on the initial selection of link. The goal would be to minimize the stroke for packaging restrictions. The advantage of SMAs in general is that they can provide relatively large recovery/pull forces. However, for superior performance and fatigue/degradation considerations, SMA are designed

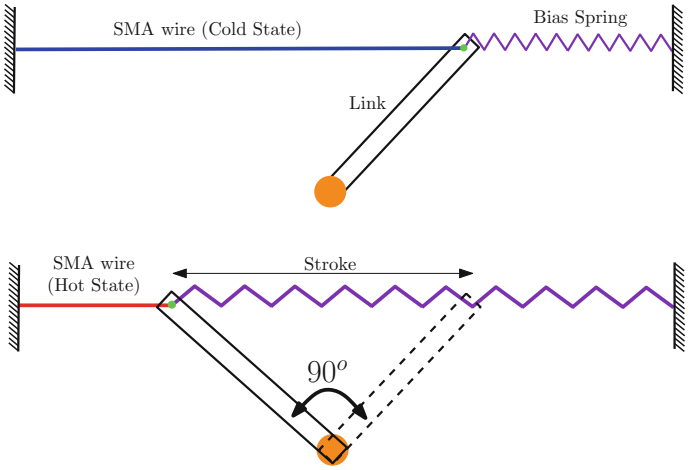

Fig. 6.6 An arrangement consisting of SMA wire(s) and steel bias spring along with the link form a simple linear to rotary crank mechanism. With changes in temperature, the recovery force generated by SMA wire causes the link to rotate by 90°. Bias spring here again resets the mechanism back to its original position for next cycling

Table 6.2 Choice of link dimensions, force and required stroke for fixing SMA wire lengths

Torque (N mm)	Link length (r in mm)	Force (N)	Stroke mm ($S = \text{r} \times \text{rotation}$)
160	2	80	3.14
160	4	40	6.28
160	*6*	*26.67*	*9.42*
160	8	20	12.57
160	10	16	15.71
160	12	13.33	18.85
160	14	11.43	21.99
160	16	10	25.13

with 3 % strains as limitations. Given the packaging and fatigue limitations, values italicized in Table 6.2 are selected for fixing the link length (torque arm of 6 mm), force of 27 N and a net stroke of 9.5 mm.

Fixing these, one would now have to select a bias spring or a torsion spring for returning back the (SMA wire + link) arrangement to its low temperature position. Considering a linear bias spring as the return arrangement, one needs to select the spring deformation and stiffness for this arrangement. With the torque/recovery force values and the link dimensions fixed, one can assume different spring deformations and pick an appropriate spring stiffness. Since the bias spring is a commodity, different deformation–stiffness combinations can be selected for the same torque/recovery force values. In this case, as shown in the italicized values of Table 6.3, the bias spring

Table 6.3 Bias spring design—stiffness estimation

Torque (N mm)	Link length (r in mm)	Force (N)	Deformation (assumed) mm	Stiffness $\left(\frac{N}{mm}\right)$
60	2	80	2	40
160	4	40	4	10
160	*6*	*26.67*	*6*	*4.44*
160	8	20	8	2.5
160	10	16	10	1.6
160	12	13.33	10	1.6
160	14	11.43	12	1.11
160	16	10	14	0.82

deformation was fixed at 6 mm and hence the appropriate stiffness is 4.44 N/mm. Instead of using a linear spring, one can also also select a torsion spring of the required stiffness calculated using $\left(\frac{2\times T}{\pi}\right)$.

Once the stroke and bias spring stiffness are selected an appropriate SMA wire(s) and length meeting these force–stroke requirements must be picked. Using Fig. 6.7, depending on the combined recovery pull force and spring stiffness, a suitable wire diameter is selected. The combined recovery pull force and bias spring stiffness is around 55 N. For the SMA wire to deliver 55 N pull force, one would require a higher diameter SMA wire (outside the scope of Fig. 6.7). Due to manufacturing difficulties, SMA wire diameters are available in limited quantity. With higher wire diameters, one would need higher power to heat the wire to desired temperatures (i.e., above A_f). Considering these factors, a wire diameter of 0.4 mm corresponding to 26 N is selected. This would mean that *two wires* of this diameter are to be used in parallel for this mechanism to function.

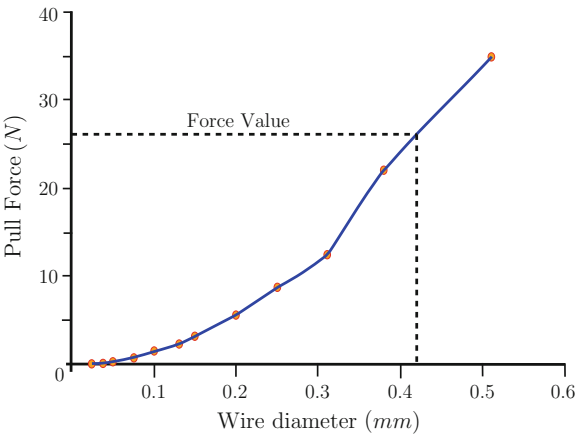

Fig. 6.7 Graph showing pull force (heating cycle) versus wire diameter. *Dotted black line* represents the force value selected for Linear to rotary arrangement using a SMA wire + bias spring arrangement. Corresponding wire diameter can be selected from this plot. Data plotted from Table 6.1 [1]

Once the wire diameter is fixed, the length of the SMA wire needs to be determined. Due to fatigue/degradation considerations, one limits the use of SMA applications under 3 % strains. Based on the stroke estimates for the required applications, appropriate wire lengths are selected using 3 % strain as a limiting strain guideline (i.e., a 100 mm wire can recover 3 mm repeatedly for 10^6 cycles). Using this conservative 3 % recoverable strain estimate, one can estimate length of SMA wire using

$$L = \left(\frac{\text{Stroke}}{0.03}\right) \tag{6.2}$$

This is equivalent to about 350 mm of SMA wire. Considering the packaging constraints in accommodating long wires, the SMA wires along with the bias spring can be connected in different ways as shown in Fig. 6.8.[1] This completes fixing all the variables of the system (SMA component and bias springs) using a graphical technique. Such a technique can be extended to analyze more complex scenarios of linear to rotary arrangement like a scissor mechanism shown in Fig. 6.9. By connecting the SMA wires opposing each other (antagonistic) and heating them can result in a rotary motion. Such a mechanism also allows an automatic return back to initial state due to the antagonistic nature and further allows better packaging. This is also an improvement upon the spooling options discussed in Fig. 6.8 as it can avoid complications arising due to sharp bends around guide pins that can result in localized plastic strains and stroke degradation over thermal cycling.

In most commercial applications, external DC power supplies are employed for joule heating of the wires. Given the intermetallic nature and metallurgical compositions of commercial SMA's, they generally have high resistivity which suggests that passing direct current could subsequently internally heat such alloys. Depending on the power supply rating and the selected wire diameter, the corresponding wire resistance and required current estimates for resistive heating of these wires can be

[1] Figure 6.8 shows two ways of arranging the 350 mm SMA wire. Top option looks at using a long SMA wire that makes the mechanism unusually long and also aesthetically/practically inconvenient from an application standpoint. Another option shown in bottom half involves bending wires around guide pins to better package the long wires (referred commonly as "Spooling" technique). Based on the application and packaging constraints, one can choose either of the two options or use shorter wires with lever arrangements for stroke amplifications. One can find several mechanisms for stroke/force amplifications in any "Mechanism design" texts that could potentially be very useful in designing SMA actuators especially from a packaging perspective (more on this in Chap. 8 later). However, it must be highlighted that such *spooling* technique can result in localized pockets of untransformed regions at the locations where the wires is bent (especially over guide pins). These sharp bends can result in accumulation of local plastic/residual strains over thermal cycling (phase transformation event) that can result in stroke degradation over repeated thermal cycling. It is thus advised to avoid sharp gradients especially during spooling around guide pins to minimize these pockets of untransformed regions.

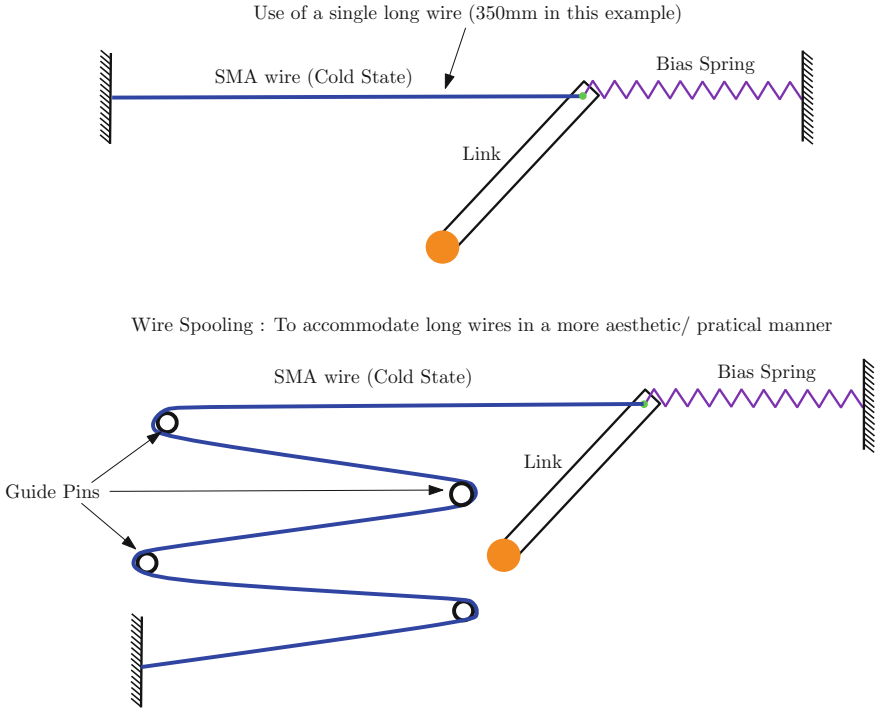

Fig. 6.8 Example showing the arrangements consisting of SMA wire(s) and steel bias spring along with the link form a simple linear to rotary crank mechanism as shown in Fig. 6.6 earlier. Two possible examples of arranging SMA wires are discussed here

obtained from Table 6.1. The following equations could help estimate current and power estimates. The corresponding resistance and current values for different wire diameters are tabulated in Table 6.1.

$$\text{Resistance of the wire} \quad R = \frac{\rho L}{A} \tag{6.3}$$

where ρ is the resistivity, L is the SMA wire length, and A is the area of cross-sectional area.

$$\text{Power Requirement} \quad W = I \times V = I^2 \times R \tag{6.4}$$

where I is the current, V is the voltage, and R is the resistance. Once the wire geometry (i.e., diameter, length and number) details are fixed, then the wire is repeatedly cycled for 50–100 cycles to obtain stabilized hysteretic responses of the material.

It must be highlighted that the resistance/current estimates from these tables are quite conservative and approximate in nature. Studies have shown that the wire resistance changes during repeated cycling due to microscopic changes with repeated phase transformations [4]. However, due to limited literature on this aspect, one

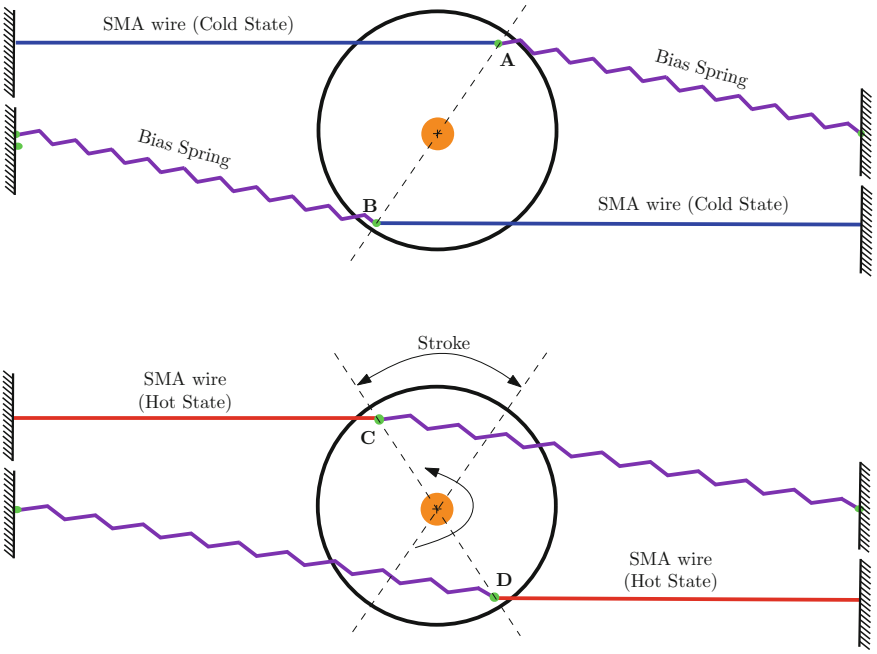

Fig. 6.9 An antagonistic scissor mechanism with two SMA wires and bias springs attached opposing each other. By heating the SMA wires, one can achieve rotary motion

might have to perform experimental trial runs using these estimated values to finalize the required current values. We will show how to estimate current values and heating/cooling rates for resistive/joule heating of SMA wires of known resistivity and dimensions later in Sect. 7.7.

6.4 Case Study 3: SMA Wire + Bias Spring Arrangement Using Linearized Loading—Unloading Response

Until now, the stroke estimation was based on the linearized loading response of SMA wire. This was the simplest approximation that one could consider to estimate strokes. However, in reality SMA wire responses are much more complex in nature which results in hysteretic loading–unloading responses due to inherent phase transformations. Top part of Fig. 6.10 shows a typical SMA wire response at temperatures $< M_f$ depicting the shape memory effect and superelastic responses at temperatures $> A_f$. The superelastic response is clipped at 6 % strain due to poor fatigue life's at large strokes.

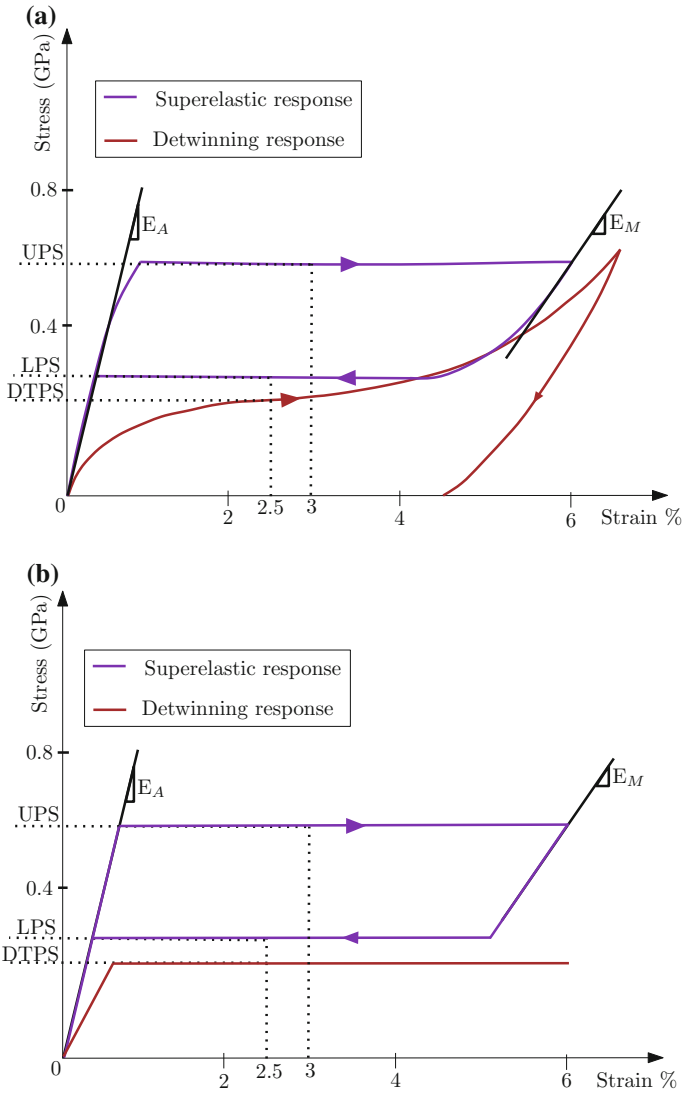

Fig. 6.10 **a** The *upper* part of the figure shows SMA wire response at temperatures $< M_f$ showing the shape memory effect and superelastic responses at temperatures $> A_f$. **b** The *lower* part of the figure shows the linearized approximation of experimental responses based on ASTM standard F2516-07ϵ [2]. Refer Fig. 1.7 for more details. Experimental data from [3]

In order to approximate the complex nonlinear hysteretic response of SMAs, one can invoke to the ASTM standard F2516-07ϵ [2] to estimate some "salient features" of the superelastic response that can serve as a premise for linear approximation. In Sect. 6.2, **Upper plateau strength (UPS)** was employed to approximate the loading plateau and the two moduli's E_A and E_M. Now, in order to approximate the unloading part of the response, ASTM standard F2516-07ϵ [2] provides information on **Lower**

plateau strength (LPS), which corresponds to stress levels at 2.5 % strain during unloading stage of the superelastic response. These stress levels serve as inputs for the linear approximations of loading and unloading plateaus as shown in the bottom half of Fig. 6.10.

Unfortunately, there is no ASTM standard for the detwinning/shape memory effect (SME) responses. As discussed in Sect. 6.2, we propose using stress levels corresponding to 2.5 % strain during loading as a linear approximation measure to determine the corresponding martensitic plateau (referred to as DTPS i.e. Detwinning Plateau stress corresponding to 2.5 % Loading). The martensitic moduli is the initial slope of the detwinning response similar to the superelastic response as shown in the bottom half of Fig. 6.10. Figure 6.11 uses this salient points of linearized loading–unloading response discussed above along with the bias spring information. As discussed in the previous sections, one estimate the stroke by marking the intersection points (shown in orange) denoted by points **A**, **B** and **C**. **AC** would be the maximum stroke at the high temperature as shown in Fig. 6.11.

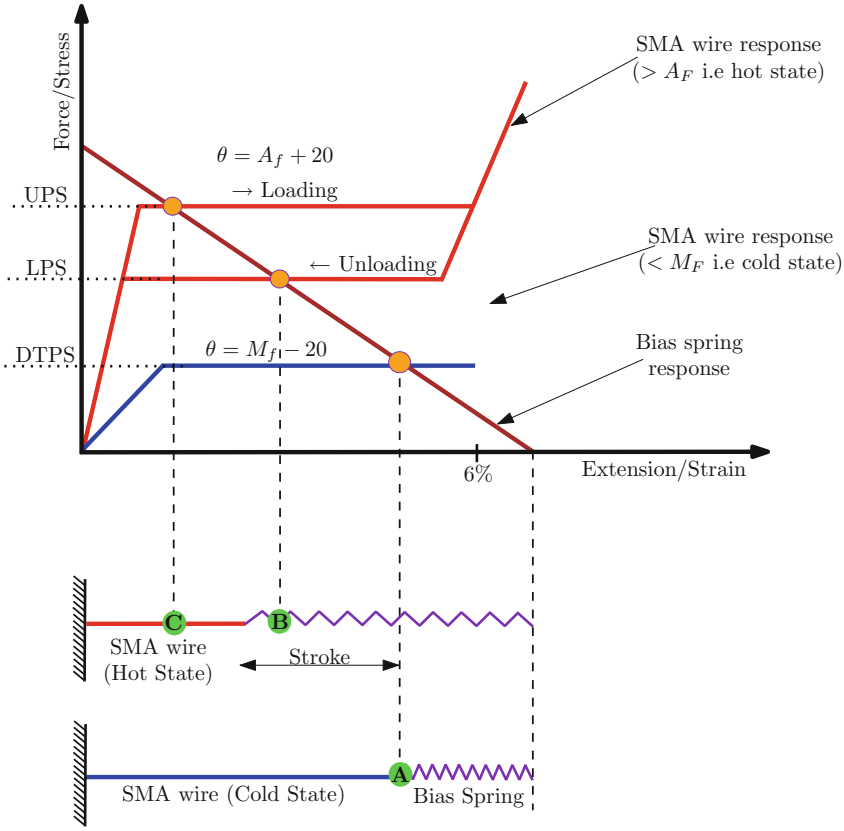

Fig. 6.11 Figure shows linearized SMA loading–unloading responses at temperatures above A_f and ones below M_f. The bias spring response is overlapped to establish the stroke again using this information

6.5 Case Study 4: SMA Wire + Bias Spring Arrangement Using Complete Hysteretic Loading—Unloading Response

So far, in the previous three case studies discussed in this chapter, we have discussed linearized options from the simplest case of loading response only to a linearized loading–unloading responses. Now, we can look at a case that considers a complete hysteretic response with the bias spring response as shown in the Fig. 6.12. This case

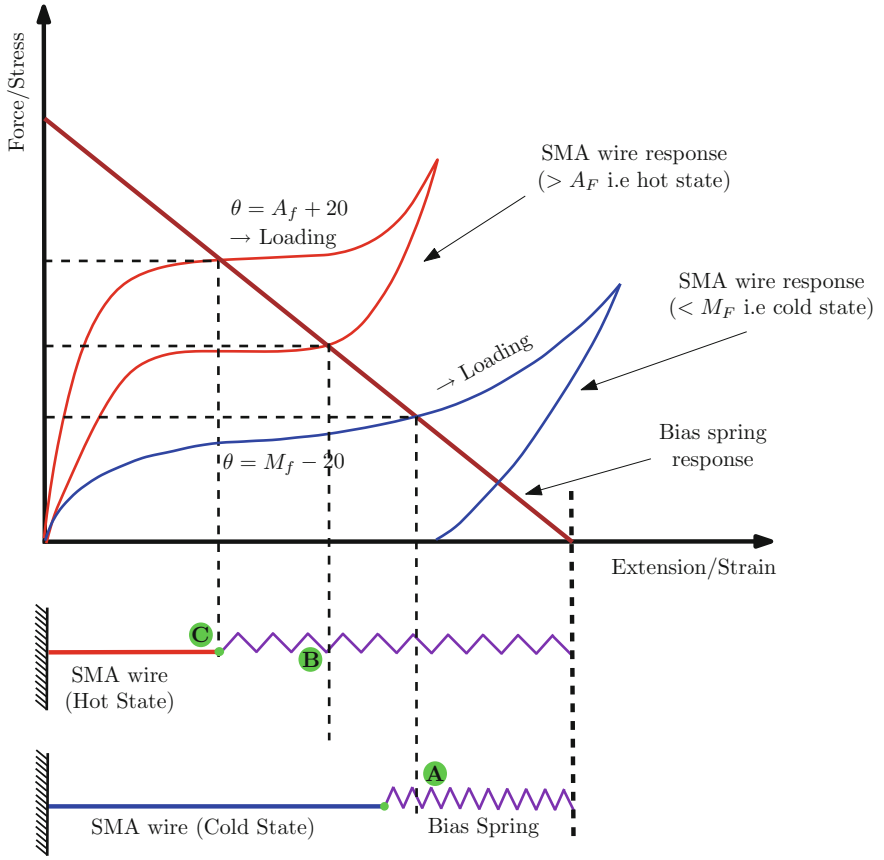

Fig. 6.12 Figure shows SMA wire responses at two different temperatures along with the bias spring response overlapped for stroke estimation

is very similar to the example discussed in case study 3 (see Fig. 6.11) but for the complete response considered here against the linearized version in case study 3. One can expect more accurate/refined stroke estimates compared to case study 3.

In the next chapter we will show how we use "stroke or strain limiters" to guarantee that:

1. the required stroke is obtained and
2. the low temperature stress does not exceed 70 MPa.

References

1. Website DIC (2014) Dynalloy—makers of dyanmic alloys. http://www.dynalloy.com/tech_data_wire.php
2. 2007e2 ASF (2007) Standard test method for tension testing of nickel-titanium superelastic materials. ASTM International, West Conshohocken, PA
3. Shaw JA, Kyriakides S (1995) Thermomechanical aspects of niti. J Mech Phys Solid 43(8):1243–1281
4. Ramaiah K, Saikrishna C, Ranganath V, Buravalla V, Bhaumik S (2011) Fracture of thermally activated niti shape memory alloy wires. Mater Sci Eng A 528(16):5502–5510

Chapter 7
Case Studies in the Preliminary Design of SMA Actuators

In spite of the rather complex and hysteretic behavior of SMA components (wires and springs), and the complex changes that occur in their microstructures, the preliminary design of SMA wires and springs turns out to be reasonably simple *provided we utilize their shape memory characteristics in specific ways*. Designing SMA actuators requires defining actuation intervals (temperature profile), load levels and/or displacement outputs over its intended designed life. The alloy composition dictates the transformation temperatures and the component's operating range. SMA components could be functional under load controlled or displacement controlled or simultaneous load and displacement controlled setups. Broadly, if we follow the two key principles explained below, we will be able to design a "two-state" actuator rather simply.

1. We need to ensure that at the high temperature, the loaded SMA actuator is entirely austenitic, that is, we do not load it with sufficient load to cause stress induced martensite (SIM) transformation. Said differently, we need to heat the SMA wire high enough so that the upper plateau stress is more than the maximum stress that is applied to the wire. This will ensure that complete transformation occurs during the heating process. Also this will considerable ease computations. As discussed in Chap. 6, Dynalloy® [1] one of the leading suppliers of SMA actuator wires recommends the safe maximum stress for their wires at the high temperature is 172 MPa and at lower temperatures is 70 MPa.
2. We use a "stroke limiter" or "hard stop" for the wire or spring so that at low temperature, the elongation of the wire is limited by us and is NOT based on the low temperature SMA response to the load. The hard stop is chosen such that the maximum strain in the wire (in the cold condition) is about 3–4 % and for springs the maximum shear strain is limited to about 1.5 %. Again, this will ensure that the SMA does not degrade with cycling and also will considerably simplify the device.

A. Rao et al., *Design of Shape Memory Alloy (SMA) Actuators*,
SpringerBriefs in Computational Mechanics, DOI 10.1007/978-3-319-03188-0_7

These two conditions will ensure that the required geometrical dimensions of the wire are easily computed for a preliminary design and that the wire will last millions of cycles.

The core elements of decision making when using a SMA wire are as described next.

1. *Find the stroke needed.* As a rule of thumb, a typical SMA wire will provide a stroke of about 3–4 mm per 10 cm of wire (strain of about 3–4 %) for long term reliable performance. So, if you want a stroke of 10 mm and you choose to use an SMA wire directly, you will need approximately 30 cm of wire. You may not have space to accommodate such a long wire. You have two options :

   a. "Winding" the wire around a system of pulleys i.e. the spooling system discussed in Chap. 6 to package long length of wires or
   b. Use a shorter wire but connect it to a "lever" mechanisms or pulley that will amplify the stroke.

   For example, you may choose to use a 10 cm wire connected to a 1:3 lever arm and amplify the stroke. However, as you know by now, "there is no free lunch." If you amplify the stroke, you will actually decrease the force exerted by the same ratio. That is, if the 10 cm wire is capable of exerting a force of 10 N, then the 1:3 stroke amplification will cause the force exerted to drop to about 3.33 N. We will discuss stroke amplification and other matters in the next chapter.
2. *Find the diameter of the wire to use.* The previous calculation gives the maximum force that the wire is capable of generating for the required stroke. The force scales as the cross-sectional area (and hence scales with the square of the diameter of the wire). Using recommendations from Dynalloy® [1], safe stresses of 172 and 70 MPa can be selected for heating and cooling pull forces. Dynalloy offers different wire diameters and as discussed in Chap. 6, plots of heating pull force v/s wire diameters is useful information to start with (see Fig. 6.2 and Table 6.1). One can get a reasonable estimate of either the wire diameter or corresponding pull force if one of them is fixed based on the end application. SMA wires can recover 6 or 8 % strains under ideal conditions but we limit it to under 3–4 % strains due to fatigue constraints.
   Roughly, a stress of about 170 MPa =24.66 ksi can be sustained by the wire when it is heated, so that the load that can be sustained by the wire during heating is about $133.5d^2$ N or about $19364d^2$ lbs. This will provide a initial estimate of the wire diameter needed, of the maximum force that needs to be lifted is known. For example, to lift a weight of 5 lb through a distance of 1″, we will need approximately 40″ of wire of diameter 0.016″. If the wire diameter needed exceeds the available wires, then we need to design a set of parallel wires to share the load.

The above calculations is an indication of whether an SMA wire is even feasible to use. In this chapter, we will go a little bit deeper to demonstrate how to make design decisions regarding SMA wires and springs and how to estimate their geometry as well as how much current is needed if we have to electrically heat it (joule heating).

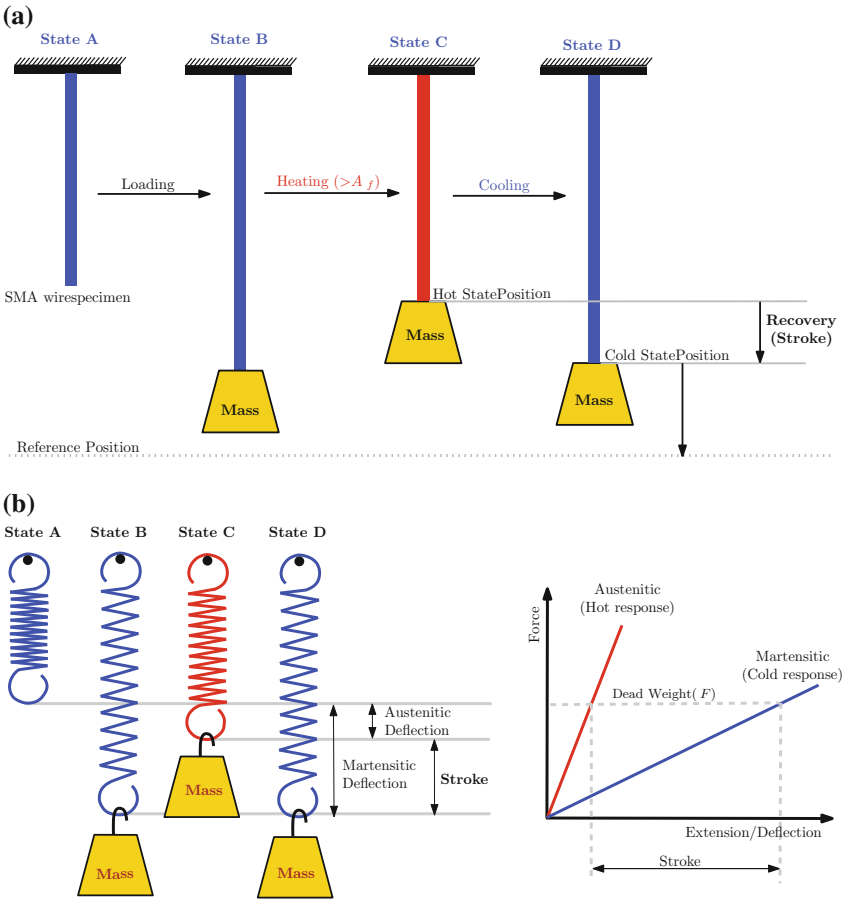

Fig. 7.1 SMA wire/spring in its martensitc state A is deformed to state B under external load. Upon heating above temperatures A_f, the SMA wire/spring contracts back to its austenitic state denoted by state C or hot state position. Upon cooling, the SMA wire/spring reaches state D. Removal of external load in state D results in SMA wire/spring returning back to state A in an ideal case. Complete cycle A↔D is termed as shape memory effect (SME). **a** SMA wire-shape memory effect. **b** SMA spring-shape memory effect (figure adapted from [2])

7.1 Different Modes of Operation

We begin by considering the common modes of operation of SMA actuators (see Fig. 7.1).

7.1.1 Constant Force Mode

Here a shape memory wire/spring lifts, pulls or pushes against a constant force and the length of the spring varies with temperature. Under external loads, the spring in

its martensitic state ($T < M_f$) deforms to a temporary shape as shown in state Ⓑ of Fig. 7.1b. Once the temperature is raised to ($T > A_f$), the wire/spring returns back to a intermediate configuration between the fully compressed length under no load (initial configuration of state Ⓐ) and the free length under external load in its martensistc state Ⓑ. During this event, the SMA wire/spring lifts the weight and behave like a thermal actuator as show in state Ⓓ. Once cooled, the SMA spring returns back to state Ⓓ. Notice that the actual stroke obtained depends upon the difference in the elongations between the cold state and the hot state. In most practical situations, the Cold state elongation (Ⓑ of Fig. 7.1b) is externally limited by a Hard stop or a stroke limiter.

If the force is applied by a weight W to an martensitic SMA wire/spring in its high temperature state (i.e., at $T > A_f$) and the spring temperature is brought down to below M_f, the weight will deform the wire/spring until motion is prevented by the hard stop and the load is transferred from the wire to the hard stop. For any shape memory spring the transformation temperatures increase in a linear fashion with the applied force and the slope of such a line is called the "stress rate." For Ni–Ti SMA springs it is of the order of 0.145 °C/MPa. It also follows that the stroke obtained decreases with an increase in applied force.

7.1.2 Constant Deflection Mode

Here the spring is constrained to a constant length and the force output from the spring varies with temperature. Consider a martensitic wire/SMA spring constrained at constant deformation/length subjected to temperature cycling. With increase in temperature above A_f, the wire/spring tries to return back to its original shape and results in a recovery pull force due to external constraints. Under constant deflection mode, the recovery force can be tweaked by changing the operating temperature above A_f. In this configuration, the internal stresses are continuously changing and thus spreading the transformation over a "range" of temperatures. The recovery pull force is responsible for it to perform as a smart actuator in different applications. Constant deflection modes are commonly used in "locking mechanisms" that are of interest to the bio-medical community (e.g., a three step bone stapling scheme utilizing a body temperature SMA, called biotinol as shown in Fig. 1.11). These applications are not related to the use of SMAs as actuators but only as thermally activated clamps and so we will not discuss them further in this book.

7.1.3 Simultaneous Force-Deflection Mode

In many real world cases, applications typically requires variation in force and displacement simultaneously with temperature. In such cases it is common practice to design an assembly consisting of an SMA spring working against a steel "bias" spring. At low temperatures ($< M_f$), the bias spring compresses the SMA spring to the hard stop (a coil bound state) and as temperature increases ($> A_f$), the stiffness of

the SMA spring increases where it can exerts enough recovery force to compress the steel spring (bias spring) giving a net "stroke". SMA springs are generally used with bias springs where the bias springs overpowers the SMA springs at lower temperatures and the SMA spring take over in their austenitic state, that is, at temperatures $T > A_f$.

7.2 Design of SMA Wires Under Constant Force

Our first example is to design a SMA wire actuator to carry a load of 100 N and provide a stroke of 10 mm between the hot and cold states. This is an extremely simple example and serves to illustrate the process. We will assume that the wire is capable of withstanding 172 MPa at high temperature and that the strain in the wire cannot exceed 3 % for long life. Further, we will use a "stroke limiter" to ensure that the strain in the wire does not exceed 3 %. This strain limiter has an added advantage: we don't have to worry about the low temperature response of the SMA at all—we simply have to ensure that the load is high enough to stretch the wire so that it reaches the stroke limiter, that is, as long as the stress in the wire is above about 100 MPa, the wire is guaranteed to reach the stroke limiting condition (see Fig. 7.2). Under these constraints, we consider the steps described next.

1. *Choice of number of wires and wire diameter.* The wire diameter is chosen based on the maximum allowable stress of 172 MPa. Then, since the load is 100 N, the required wire cross sectional area for the wires is

$$A = \frac{F}{\sigma_{max}} = \frac{100}{172} = 0.58 \text{ mm}^2 \tag{7.1}$$

   If we go for a single wire, we will need a wire diameter of $d = \sqrt{4A/\pi} = 0.86$ mm. Large diameter wires might be difficult to procure commercially and would requires higher currents for heating due to larger resistances. The requirements of higher currents for heating and associated longer cooling times with larger diameter wires, we might be better off opting for multiple smaller diameter wires (for example, 3 wires of diameter 0.45/0.5 mm, which gives a total cross sectional area of $\approx 0.59\,\text{mm}^2$ or something similar). After the preliminary design is carried out, we will have to calculate its heating and cooling rates to ensure that the wire(s) operate(s) under suitable conditions. The choice of wire diameter and number of wires is a balance between competing requirements (trade-off). The need for a simple design implies that we use a single wire, but then the cooling rate will be very low (as we shall see later) and it will take a longer time for the wire to complete an actuation cycle (thereby decreasing the actuation frequency for use in multiple cycles repeatedly).
2. *Choice of the wire length.* We now have to compute the length of the wire needed. This dimension is controlled by the required stroke. We compute the length as

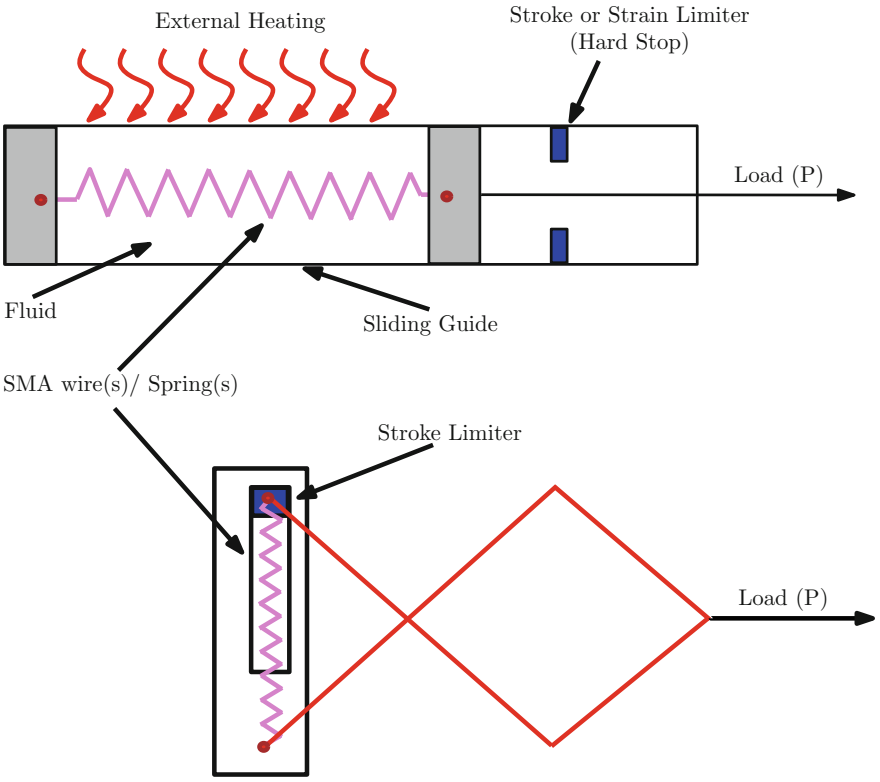

Fig. 7.2 Stroke limiter system to limit maximum strain on SMA component (wire/spring) and also to limit the stress on the SMA at low temperatures

follows. First, we calculate the strain in the wire at high temperature (using a Young's modulus of $E = 75$ GPa) as

$$\epsilon_H = \frac{F}{AE} = \frac{100}{(0.59 \times 75000)} = 0.2\,\% \tag{7.2}$$

Now we will require that the maximum low temperature strain should be equal to $\epsilon_L = 4\,\%$ so that the strain change due to actuation is $\epsilon_L - \epsilon_H = 3.8\,\%$ This should provide a stroke $S = 10$ mm. Hence the wire length L_f is calculated as

$$L_f = \frac{S}{\epsilon_L - \epsilon_H} = 265 \text{ mm} \tag{7.3}$$

We note that in order to obtain a 10 mm stroke, we need a wire of length 26.5 cm! This is a fairly long wire and may be need to be spooled or wound around a set of pulleys. Its hot length will be $L_f \times (1 + \epsilon_H) = 265.5$ mm and its cold length

will be $L_f \times (1 + \epsilon_L) = 275.6$ mm giving the required stroke of about 10 mm. *Note that the since the load is constant and the axial stress in the wire is about 171 MPa which is well above the "plateau stress" of the low temperature phase, we will need a "hard stop" or "stroke limiter" that will stop the stretching of the wire and placed at a distance of 275.6* mm from the grip of the wire to stop the wire from stretching further at the low temperature. In other words, our design will always be such that the wire stress is above the plateau stress for the "soft martensite," and we will use a "stroke limiter" to get the required stroke.

We note that in general we will need quite long wires to get a decent stroke in an SMA. If we want to shorten the wire length, we have two options:

1. Choose a spring configuration or
2. Select some sort of leverage to obtain the required mechanical advantage.

The second option implies that the load will be multiplied while the stroke decreases, but we may be able to tolerate that in a design.

7.3 Case Study II: Design Procedure for Ti-Ni (SMA) Springs[1]

The first step in the design of SMA actuator springs is to define the actuation temperature interval, force and/or motion output expected and the number of operational cycles required. The transformation temperatures necessary to realize the required temperature interval of operation usually dictate the alloy composition and the processing route to be chosen.

7.4 Spring Design Case Study

Unlike the design of the wire, spring design may require several iterations, since the stresses are a function of the spring index which are in turn dependent upon wire dimensions.

7.4.1 Design Model and Assumptions

The following approach provides a "quick and dirty" estimation of the SMA design. The approach is based on the work of Waram [3]. Several factors are noteworthy here:

1. The design is based on replacing the actual stress response with simple linear approximation at the high temperature. This works well for springs since the

[1]We thank Dr. S. Chandrasekaran, SMA specialist at IIT Madras and formerly at Qinetiq Corporation, for this chapter.

hysteresis in the torque versus angle of twist or the shear stress versus shear strain curve is very small in the range of use.
2. For compression springs, the design is based on using the "coil bound" condition as a stroke limiter. For extension springs, stroke limiters are to be used to prevent excessive deformations.
3. In cases where the load varies so that there is not enough force at low temperature to cause the necessary stroke, a "bias spring" is may be necessary to increase the low temperature load to a sufficient value to create the needed stroke.
4. The formulae are based on linear elasticity and can be found in spring design books such as Shigley [4].
5. Actual results may vary but this first cut allows for a preliminary design.

When the preliminary design is complete and the spring is made, it is perhaps a good practice to obtain plots of a number of isothermal shear stress–shear strain curves at various temperatures by measuring the force–deflection behavior of a test spring at several temperatures and then convert to stress–strain data (see Fig. 4.7 discussed earlier). It must be highlighted that this approach is a simple linearized approximation of SMA spring responses as only the initial linear part of the material response is considered. At higher force/torque—stroke/twist levels, the response enters the nonlinear regime i.e. at higher extents of phase transformation regime. For now, to avoid mathematical complexities, the normal design is to restrict the strains to values below the onset of nonlinearity. Under such an assumption of linear response, the equation for shear stress reduces to the standard linear elastic expression used for conventional spring design,

$$\text{Shear Stress} = \tau = \frac{8PD}{\pi D^3} \tag{7.4}$$

The linearity assumption greatly simplifies the design process; however, it brings two important limitations due to this design simplification. They are listed below.

1. The ability to predict the resetting forces for the design of the bias springs are somewhat limited as the nonlinear phase transformation region is not captured
2. The above formula is based on isothermal stress-strain plots obtained at constant temperatures. In reality, however, an SMA spring operates under changing temperature conditions. Since this model ignores any dynamic effects of cooling and heating the inaccuracies that result from this limitation are more pronounced at lower temperatures and affect the resetting behavior on cooling. The above limitation is rectified by

    a. first testing the prototype springs and measuring the resetting forces under dynamic cooling conditions, and then
    b. manufacturing/selecting the required bias spring to either extend (for tension springs) or compress (for compression spring) the spring at the low temperature so that we can obtain the necessary stroke.

Table 7.1 Symbols and their meaning

Symbol	Meaning
δ	Deflection
d	Wire diameter
n	Active coils
c	Spring index
OD	Spring outer diameter
D	Spring average diameter
τ	Shear stress
γ	Shear strain
τ_c	Corrected shear stress
P	Applied force
W	Wahl correction factor
K	Spring Rate
G	Shear modulus

7.4.2 Terms Used in Design of SMA Springs

For a quick reference, we list the symbols (see Table 7.1) and classical design equations and for springs (directly from spring design methodology [4])

$$D = \text{OD} - d \tag{7.5}$$

$$c = \frac{D}{d} \tag{7.6}$$

$$\tau = \frac{8PD}{\pi d^3} = \frac{8Pc}{\pi d^2} \tag{7.7}$$

$$W = \frac{(4c-1)}{(4c-4)} + \frac{0.615}{c} \tag{7.8}$$

$$\tau_c = \frac{8WPc}{\pi d^2} \tag{7.9}$$

$$K = \frac{Gd^4}{8nD^3} \tag{7.10}$$

$$\gamma = \frac{d\delta}{\pi n D^2} \tag{7.11}$$

Armed with these formulae we will now consider an example application which goes through the full design estimation for a particular application.

7.5 Example: Design of a Remote Controlled Flow Control Valve Using an SMA Compression Spring

7.5.1 Statement of Requirement

Petroleum crude is extracted from oil bearing sub-soil rock formations by positioning hollow shafts in the oil-bed and forcing the oil to rise through them using pressurised steam. In order to regulate the flow of oil from the oil bed into the hollow shaft as required, a sliding steel sleeve valve is used. This valve has entry holes on its sides matching the entry holes on the side of the hollow vertical tube shaft over which it can slide. This assembly is buried well below the earth surface and stands immersed in an oil bed. The petroleum fluid surrounding the assembly is extracted out through the hollow shaft under normal circumstances by positioning the sleeve in such a way that the holes in the annular shaft and on the sliding sleeve are in exact alignment. However, once the petroleum fluid is exhausted water from surrounding regions begins to rush in. This needs to be arrested by shifting the position of the sleeve such that it slides on the hollow shaft and shuts the entry holes completely.

In the specific application taken up for design the sleeve was to be lifted to a height of 100 mm by remote operation. The lifting action should be triggered only above 55 °C and should be achievable within 5 min. It was proposed to design a compression SMA spring for this purpose. The sleeve (see Table 7.2 for dimensions) would sit on the spring in such a way that the entry holes are in perfect alignment to allow oil flow under normal circumstances. When oil bed is exhausted and water ingress is suspected the SMA spring is operated remotely by sending a signal that will trigger electrical heating and cause the spring to lift the sleeve by 100 mm shutting the entry holes.

The mass of this sleeve is $\left(\frac{\pi(13^2-9^2)}{4}\right) \times 15 \times 7.8 = 8086\,\text{g} \approx 8.1$ kg (taken as 10 kg for convenience). The force exerted by the 10 kg mass is 10 kgf $\approx$ 100 N. Thus, our design requirements are as follows.

1. required load = 100 N dead load (weight of the sleeve);
2. required stroke (i.e., motion when the wire is heated) = 100 mm;
3. available space = 130 mm diameter (since the outer diameter of the sleeve is 130 mm).

Table 7.2 Dimensions of sleeve

OD of sleeve	130 mm
Wall thickness	20 mm
Length	150 mm
Density	$7.8\frac{g}{cc}$

For the shape memory alloy, the material constraints are that we have only 7 mm diameter wire (from which to make the spring) and we will restrict the shear strain in the wire to $<1.5\,\%$. In spite of the fact that the SMA is reported to provide strains of $8\,\%$, we will almost never use the full transformation since there will be rapid stroke degradation of the wire as the cycles increase. The SMA taken up for the design was a Ti–Ni alloy with an $A_s = 55\,^{\circ}\mathrm{C}$. The effective high temperature austenite shear modulus G_H of the spring made of this alloy was taken as 20.7 GPa while the effective shear modulus G_L for springs in the low temperature martensitic state was taken as 2.75 GPa (see page 52 of [3]).[2]

The spring we are about to design is a compression spring. Such springs tend to have a longer life and further they have the advantage that they can be made to be "coil bound," that is, in a state where the coils are touching. Further deflection is prevented and therefore there is no question of an overload.

Thus, we will design the spring so that when it is in the cold, soft coil bound condition, the shear strain is $1.5\,\%$. and when it is fully actuated, that is, when it is in the hot, stiff deployed condition, its stress does not exceed 140 MPa. This makes the design really easy since in the hot condition the response of the spring is purely elastic so that the spring can be designed like a traditional elastic spring in the hot condition. Then everything else is actually only a matter of kinematics. For such springs, we have to consider the elastic response of only the high temperature state (i.e., the austenite) and ensure that as long as the martensite is soft enough there will be no further deformation since the spring will be coil bound. Given this initial data, we will proceed with the design.

1. *The spring index and the choice of wire diameter d.* For a given value of force P to be exerted by the SMA spring when it expands with rising temperature, the first step is to calculate the required diameter d of the SMA wire and the spring major diameter D. The wire diameter is determined from the maximum allowed shear stress, which results in the following equation

$$d = \sqrt{\frac{8WPc}{\pi \tau_c}} \tag{7.12}$$

where P is the force to be exerted by the spring, c is the spring index, and W is the Wahl correction factor.

We use an iterative scheme for the geometry design as follows: (1) pick a spring index "c" and choose a suitable working stress τ_c; (2) Find a suitable wire diameter from the equation above; and (3) find the coil mean diameter D using the wire diameter and the assumed spring index. If it is not satisfactory go back to (1). For most of the SMA springs it is usual to choose a value for the spring index c between 6 and 10. However, there seems to be no stringent requirement that the maximum value should not exceed 10. In the case of SMA wires one

[2] In reality, this is the just the slope of the stress strain curve for the twinning deformation and is not really an elastic modulus.

is sometimes forced to choose higher values of c because of the difficulty in commercially procuring appropriate wire diameters.

The value of the shear stress τ_c is kept relatively low to ensure a low level of functional fatigue on cycling and it is usually taken to be between 100 and 140 MPa for Ti–Ni SMA. In the present case the memory spring the mean diameter is taken as 130 mm (to match the OD of the sleeve). Only 7 mm diameter wire was available. Thus the spring index c was estimated as $c = \left(\frac{130-7}{7}\right) = 18.1 \approx 18$. Assuming a critical shear stress τ_c of 140 MPa for austenite SMA, the wire diameter d for the spring is recalculated with these updated numbers as explained next.

a. First, calculate the Wahl correction factor,

$$W = \left[\frac{(4 \times 18 - 1)}{(4 \times 18 - 4)}\right] + \left[\frac{0.615}{18}\right] = 1.078 \tag{7.13}$$

b. Next, estimate the wire diameter,

$$d = \sqrt{\frac{8PWc}{\pi \tau_c}} = \sqrt{\frac{8 \times 100 \times 1.078 \times 18}{\pi \times 140}} \approx 6.0 \text{ mm} \tag{7.14}$$

It turns out to be not very far away from the available diameter $d = 7$ mm of the SMA wire. So, if we choose a 7 mm diameter wire and a spring index of 18, we will end up with a coil diameter of 126 mm.

2. *Calculation of active number of coils n.* Determination of the number of active coils in the spring is carried out through imposing a constraint on the shear strain γ. Since a linear stress–strain behavior is assumed, at at high temperature we have $\gamma_H = \tau_c / G_H$ so that the high temperature shear strain γ_H is then calculated as

$$\gamma_H = \frac{140}{20700} = 0.0068 = 0.68\,\% \tag{7.15}$$

For a good cyclic life it is customary to limit the low temperature shear strain $\gamma_L = 0.015 = 1.5\,\%$ We will prevent further deformation in spring by ensuring that this shear strain will cause the spring to become coil bound.

The difference between the low temperature and high temperature shear strains is

$$\Delta\gamma = \gamma_L - \gamma_H = 0.015 - 0.0068 = 0.0082 \tag{7.16}$$

This shear strain difference must deliver the required stroke.

We can now calculate he number of active coils n using the expression

$$n = \frac{(d \times S)}{\pi D^2(\Delta\gamma)} \tag{7.17}$$

where S is the stroke. This gives

$$n = \frac{(7.0 \times 100)}{\pi \times 126^2 \times 0.0082} \approx 1.6 \approx 2 \text{ coils} \tag{7.18}$$

Thus we are looking for a wide shallow spring with only two active coils. Since it is a compression spring, assuming that the spring has unground, closed ends the total number of coils N is given by the expression $N = n + 2$, so in this case the number of coils is $2 + 2 = 4$.

3. *Calculation of total length of coil.* The first step is to determine the high temperature spring rates K_H using the formula

$$K_H = \frac{G_H d^4}{8nD^3} = \frac{(20700 \times 7^4)}{(8 \times 4 \times 126^3)} \approx 0.758 \text{ N/mm} \tag{7.19}$$

Thus the high temperature deflection $\delta_H = P/K_H = 100/0.758 = 131$ mm in compression, and hence the low temperature deflection is[3] $\delta_L = \delta_H + S = 131 + 100 = 231$ mm. Now, the length of the spring L_L when it is fully compressed at low temperature is given by

$$L_L = d(n + 3) = 7 \times (4 + 3) = 49 \text{ mm} \tag{7.20}$$

It follows that the length of the spring when it is at high temperature is given by $L_H = L_L + S = 49 + 100 = 149$ mm . Finally, the manufactured length of the spring, also called the free length L_F, is

$$L_F = L_H + \delta_H = 149 + 131 = 280 \text{ mm} \tag{7.21}$$

We finally have to ensure that with the 100 N force, the spring will become coil bound at the low temperature, For this reason, we make an approximate estimate of the slope of the martensitic twinning transformation line at low temperature and obtain the result that for the spring to deflect 231 mm.

We first calculate the spring rate at low temperature as

$$K_L = \frac{G_L d^4}{8nD^3} = \frac{(2750 \times 7^4)}{(8 \times 4 \times 126^3)} \approx 0.1 \text{ N/mm} \tag{7.22}$$

[3] Note we do not use the spring constant at low temperature because the spring will be "plastic" at low temperature and we are designing it to be coil bound.

Thus the force required to compress by 231 mm is $R_F = (K_L \times \delta_L) = 0.1 \times 231 = 23.1$ N. It can easily be seen that this spring would go coil-bound under the action of the constant force of 100 N acting on it.

The key point here is that the we cleverly use the "hard stop" provided by the coil bound condition to no only guarantee that there will not be excessive deformation but also dramatically simplify the calculations of the spring dimensions based on only the high temperature elastic response.

7.6 Extensional Spring Design

The design of an SMA extensional springs uses exactly the same formulae as were used in the design of an SMA compression spring for the wire diameter, spring diameter, and active number of coils , provided we make use of a slot or hard stop to limit the stretch at low temperature. In other words, we need to design the spring with a "hard stop." Unlike the compression spring where this is provided by the coil bound condition, for tension springs we need to add hard stops as shown in Fig. 7.2. The only difference in the design is the determination of the various spring lengths L_H, L_L and L_F.

First, the spring is "shape set" such that it is close coiled (coil-bound) at high temperature with no force applied (as can be seen in commercial SMA springs). The length of the spring in this condition is called the body length of the spring.

The body length L_B of the close coiled section of the extension spring (i.e., excluding hook lengths) is given by $L_B = d(n + 1)$, where d is the wire diameter and n is the number of active coils. If it is assumed that end hooks are added on both ends that are each equal in length to the coil inner diameter (ID), then, for the example above, the free length, L_F of the spring is given by

$$L_F = L_B + 2\,(\text{ID}) = d(n + 1) + 2\,(\text{ID}) \tag{7.23}$$

$$= 7(4 + 1) + 2(127 - 2 \times 7) = 261 \text{ mm} \tag{7.24}$$

Now, the high temperature length L_H is equal to the free length L_F plus the high temperature deflection δ_H

$$L_H = L_F + \delta_H = 261 + 131 = 392 \text{ mm} \tag{7.25}$$

and the low temperature length $L_L = L_H + S = 392 + 100 = 492$ mm. We need to provide a hard stop at this length so that we get repeatable response and so that the stroke is appropriate.

Thus, there are three essential features of the spring design:

1. The spring diameter is obtained by limiting the high temperature elastic response stress to be about 140 MPa with an assumed spring index. This is an iterative process to obtain a satisfactory spring index and inner and outer diameters.
2. The low temperature strain is fixed by considerations of life.

3. The strain difference combined with the required stroke will give the number of coils.

We will ensure that the low temperature deflection is limited by a hard stop (either externally for tension springs or with the coil bound condition for tensile springs).

7.7 Heating and Cooling of Shape Memory Wires

In order for the material to be activated, it must be heated and cooled. Of these, the cooling is the more problematic issue since, while it is easy to heat the wire, cooling depends upon ambient conditions. In general there are two means of heating an SMA wire or spring:

1. Internal heating through the passage of an electric current or
2. Contact heating through ambient.

Since the resistivity of SMA is rather high (approximately $10\,\mu\Omega$ mm $= 1 \times 10^{-6}\,\Omega$ cm), it is quite simple to heat the wire by passing an electric current, I.

For such a situation, the steady state temperature that can be reached is given by the balance between the internal heat generation due to joule heating and the heat transfer to the surroundings (we will ignore the radiative heat transfer, but it would be important in situations where the wire is quite hot and in situations such as vacuum where convective effects are small). Based on this, the steady state temperature would be

$$T = T_{amb} + \frac{4\rho I^2}{\pi^2 d^3 h} \tag{7.26}$$

where T_{amb} is the ambient temperature, ρ is the resistivity ($\approx 10\,\mu\Omega$) mm, I is the current and h is the heat transfer coefficient and d is the wire diameter. Of these, the most uncertain parameter is the heat transfer coefficient and only very approximate values are available. The value depend upon:

1. Whether the wire is vertical or horizontal or
2. Whether there is free or forced convection

Correlations and calculations are available for a variety of cases and should be available in any heat transfer books.

To illustrate the vast difference that the diameter of the wire makes consider the current required to heat either a 5 mm wire or a 0.5 mm wire. Assume that the length of the wire is 200 mm, the ambient temperature is 25 °C with the target temperature to be 90 °C. We can estimate from the "thermal wizard" web site [5] that the heat transfer coefficient is 12.4 W/m^2 C. Using this value, the estimated current would be

$$I = \sqrt{\frac{(T - T_{amb})}{4\rho} \times (\pi^2 d^3 h)}\ \text{Amps} \tag{7.27}$$

To obtain a final temperature of 90 °C, we will need a current of approximately 14 Amps! One can see that heating a 5 mm diameter wire can be unfeasible. However, if one were to consider a wire of 0.5 mm diameter, then the heat transfer coefficient is 42 W/m^2 C and the current required will be

$$I = \sqrt{\frac{(T - T_{amb})}{4\rho} \times (\pi^2 d^3 h)} = 0.92 \text{ Amps} \tag{7.28}$$

If a very specific final temperature is desired, it is better to mount a thermocouple on the wire and build a feedback control system to control the current.

If the wire is very thick (or for SMA rods under compression) the current required may become excessive. In these cases, heating through an external hot body becomes feasible. In this case the small surface area becomes a disadvantage. It is better to enclose the SMA wire or spring in a tube with a conducting fluid and heat the tube. This has been shown to be very effective with solar heating where the tube can be painted black. Of course, in this case estimating the final temperature may be quite complicated .

7.7.1 Time Taken to Heat Up and Cool Down

While it is a simple matter to estimate the current required to heat up the wire, it is actually quite different to estimate how long would it take to heat up or cool it down, this is because of the fact that the wire undergoes a phase transition during heating[4] this is not correct.

A rough estimate of the time taken to heat or cool the wire is given by considering the total heat supplied by joule heating and that due to convective heat transfer and the increase in the internal energy due to changing temperature[5]. Thus, a crude approximation is based on the following equation:

$$mC_p \Delta T + mL = [-h(T_{avg} - T_{amb}) + I^2 R]\Delta t \tag{7.29}$$

where m is the mass of the wire and C_p is the heat capacity per unit mass (assumed constant), L is the latent heat per unit mass for the conversion, and T_{avg} is the average temperature at which the heat transfer takes place. In arriving at this simple expression we caution that we have made a number of simplifying assumptions: (a) We have ignored the work done by the wire. It is a reflection of the low conversion efficiency of the SMA (order of 5–10 %) that we can ignore this and still obtain reasonable answers.

[4]Beware that many papers routinely ignore this and try to estimate the heat transfer coefficient using a simplified model. The answers, not surprising show wide divergence and indicate that the heat transfer coefficient varies during the process. Actually these papers are attributing the temperature stagnation due to phase transitions to variations in the heat transfer coefficient.

[5]This would be inappropriate for situations were repeated rapid cycling takes place as in the case of a heat engine or a flapping wing system or in vibration damping applications. For these systems it is important to consider the heating due to hysteresis also.

Furthermore, we have ignored the hysteresis losses and we have approximated the heat transfer by assuming that it occurs at the average temperature rather than at the actual temperature of the wire. This is indeed a serious approximation but this is sufficient to provide an order of magnitude estimate.

For the 200 mm long 0.5 mm diameter wire for which $\Delta T = 65\,^{\circ}\text{C}$, its specific heat C_p is about 0.832 kJ/kg C, its latent heat L is about 24.2 kJ/kg, its density ρ_m is 6540 kg/m^3 and the other quantities have already been estimated. Hence, we obtain

$$\Delta t = \frac{\rho_m (C_p \Delta T + L)}{-\frac{4h(T_{avg} - T_{amb})}{d} + \frac{16 I^2 \rho}{\pi^2 d^4}} \tag{7.30}$$

When we compute the time taken using the above expression, we obtain an estimate of about 46 s for heating and cooling. Again, we caution the reader that these are very crude order of magnitude numbers and are to be used only to establish the order of magnitude of the time taken in order to establish feasibility. Due to the myriad of factors that influence the cooling, it is better to verify this with further experimentation.

References

1. D. I. C. Website (2014) Dynalloy—makers of dynamic alloys. http://www.dynalloy.com/tech_data_wire.php
2. Ganesh N, Maniprakash S, Chandrasekaran L, Srinivasan S, Srinivasa A (2011) Design and development of a sun tracking mechanism using the direct sma actuation. J Mech Des 133:075001
3. Waram T (1990) Design principles for ni-ti actuators. Butterworth-Heinemann, Engineering Aspects of Shape Memory Alloys (UK), 1990, pp 234–244
4. Shigley JE, Budynas RG, Mischke CR (2004) Mechanical engineering design. McGraw-Hill, New York
5. M. Simulation (2014) Website: thermal wizard. http://www.thermal-wizard.com/tmwiz/convect/natural/horiz-cylinder/horiz-cylinder.htm

Chapter 8
Coupling SMA Actuators with Mechanisms: Principle of Virtual Work

In Chap. 7, we covered the basics of SMA wire and spring design for dead loading conditions. We noted that if the stroke required is very large, a simple SMA wire may be unfeasible.Thus, successful design of a shape memory wire actuator depends upon the ability to combine it with a mechanism to achieve the desired force and stroke when actuated. As a rough guide, a SMA wire has characteristics similar to a hydraulic cylinder and their traits include the following:

1. Large force when compared to other components with similar size
2. Somewhat slow action compared to piezoelectric actuators
3. Inability to return to its starting shape by itself if the low temperature force is not sufficient

A challenge with the direct use of SMA wires is the fact that excessively long wires may be necessary to obtain the necessary stroke. For example, to directly lift a weight of 5 lb through a distance of $1''$, we will need approximately $40''$ of wire of diameter $0.016''$.

8.1 The Need for Mechanisms

The main difficulty is with the length of the wires. A $40''$ long wire may not be feasible and hence at this point, we need to see if we can use a simple lever mechanism to deliver the required stroke. Such lever mechanisms can range from simple levers, bell crank type mechanisms, all the way to various four bar links and other sophisticated setups. Another reason for going with mechanisms is to convert motion, that is, for heating and cooling purposes, it may not be feasible to deploy the wire directly where needed but place it somewhere more convenient and transmit the motion through a system of linkages. Alternatively, it may be desirable to generate a flapping or

A. Rao et al., *Design of Shape Memory Alloy (SMA) Actuators*,
SpringerBriefs in Computational Mechanics, DOI 10.1007/978-3-319-03188-0_8

rotational motion whereas the SMA produces only rectilinear motion. For all these reasons, it is necessary to consider linkages and mechanisms as part of design.

A critical challenge with the use of linkages is the fact that the force needed at the actuator location may vary with the configuration of the linkage even though the force exerted by the external load on the mechanism may be constant. For example, consider two configurations of the lever. In case 1:– The force at the SMA wire is equal to the applied load. However, in Case 2, the force exerted by the lever on the wire is a multiple of the load and it varies with the orientation of the arm.

For a complicated mechanism such as case 2, estimation of the force(s) required to keep the system at equilibrium is quite complicated if we simply use Euler's laws for rigid bodies, since the calculations will produce all the forces in the body and not just the forces that are needed. However, as an alternative method, dating back to the work of Archimedes, termed as the principle of virtual displacement is extremely useful for these calculations.

The principle is stated (without proof) as follows (see [1]):

> Consider a system that is in equilibrium and acted upon by several external forces, $\mathbf{F}_i^{ext}$. If we now subject the system to an arbitrarily small allowable virtual displacement $\delta\mathbf{x}^i$ at each location where the force is applied such that the displacements must be "compatible" with each other, not causing the system to be torn apart or violating other constraints), then the net work done by all the forces in moving through their respective virtual displacements is zero:
>
> $$\delta W = \sum_i^n \mathbf{F}_i^{ext} \cdot \delta\mathbf{x}^i = 0 \tag{8.1}$$

This seemingly simple result actually hides a major simplification, namely that ONLY the external forces are included, not the internal forces. In other words, *the virtual work of all the internal forces are zero*. This result dramatically simplifies the calculations of the force relationships of mechanisms and has found wide use in a variety of fields. The process for using this result in computing forces in mechanisms that are in equilibrium are as follows:

1. *First identify the degrees of freedom of the mechanism* How many ways is the mechanism allowed to move. Equivalently what is the minimum number of actuators needed to keep the mechanism locked in any particular orientation? As an illustration, consider the lever mechanism shown in Fig. 8.1, with the load on one side and the actuator on the other. Assuming that the lever arm AB is rigid, there is only one allowable motion, that is, the rotation of the arm about the fulcrum O.
2. *Choose kinematical parameters* One for each degree of freedom q^α that can be used (together with the fixed geometrical parameters) that will represent all the allowable motions of the mechanism. For the lever mechanism, we can choose the angle made by AB with the horizontal ($q^1 = \theta$) as the kinematical parameter. Notice that, this choice is by no means unique. We might as well have used the height of A or the height of B as variables. However we cannot use both the heights as variables since they cannot be *independently varied*.

   This is the key step in generating compatible displacements, since the q^α are chosen represent every allowed configuration of the body, every variation of q^α

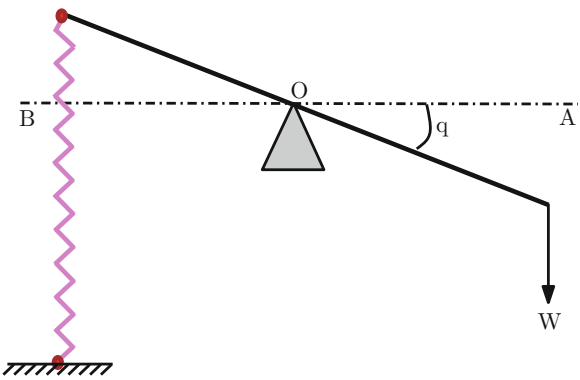

Fig. 8.1 Simple lever mechanism with the load on one side and the actuator (spring) on the other

denoted as dq^α will generate an allowed virtual displacement, without having to worry about whether the displacements are allowed or not.

3. Using basic principles of geometry and trigonometry, find the locations of the point where the external load is applied and the point of action of the SMA wire in terms of the chosen kinematical parameters. That is, we want the relations

$$\mathbf{x}^i = \mathbf{f}^i(q^\alpha),\ i = 1, \ldots Nforces,\ \alpha = 1, \ldots Ndofs \tag{8.2}$$

where $Nforces$ is the number of forces and $Ndofs$ is the number of degrees of freedom. This is the hard part and requires much ingenuity. For the simple lever problem, the relationship is

$$\mathbf{x}^A = \mathrm{OA}\,(-\cos\theta\mathbf{i} - \sin\theta\mathbf{j}),\ \ \mathbf{x}^B = \mathrm{OB}\,(\cos\theta\mathbf{i} + \sin\theta\mathbf{j}), \tag{8.3}$$

4. Using calculus find the incremental changes in the locations of the point of action of the external load in terms of the chosen kinematical variables.

$$d\mathbf{x}^i = \sum_\alpha \frac{\partial \mathbf{f}^i}{\partial q^\alpha} dq^\alpha \tag{8.4}$$

For the example in question, we have

$$d\mathbf{x}^A = \mathrm{OA}\,(\sin\theta\mathbf{i} - \cos\theta\mathbf{j})d\theta,\ \mathbf{x}^B = \mathrm{OB}\,(-\sin\theta\mathbf{i} + \cos\theta\mathbf{j})d\theta \tag{8.5}$$

5. Now compute the virtual work done by each of the forces and set the sum equal to zero

$$\sum_i \mathbf{F}_i \cdot d\mathbf{x}^i = \sum_i \sum_\alpha \mathbf{F}_i \cdot \frac{\partial \mathbf{f}^i}{\partial q^\alpha} dq^\alpha = 0 \tag{8.6}$$

For the lever, since the force are $-W\mathbf{j}$ at B and $-F_{act}\mathbf{j}$ at A, we obtain

$$-(\mathrm{OA})F\cos\theta\,d\theta-(\mathrm{OB})W\cos\theta\,d\theta=0 \tag{8.7}$$

6. Since the result is true for ALL compatible virtual displacements, that is, for all values of dq^{α}, we obtain

$$\sum_i \mathbf{F}_i\cdot\frac{\partial\mathbf{f}^i}{\partial q^{\alpha}}=0 \tag{8.8}$$

as the equilibrium equations of the body. For the lever, this condition means that

$$-(\mathrm{OA})\,F=-(\mathrm{OB})\,W\implies F=\frac{\mathrm{OB}}{\mathrm{OA}}W \tag{8.9}$$

The above seems like an awful lot of trouble to get at the principle of the lever, which is known to every high-schooler. However, the real power of the approach is seen as the mechanism gets more complicated as shown in the next example. Before proceeding with this example, we note that a key idea is to get the the virtual work and to avoid vectorial representations as much as possible. Therefore, if we choose our coordinates and what we calculate judiciously, we can use the fact that *only the component of the virtual displacement in the direction of the force actually contributes to the virtual work* and this dramatically simplify the calculations.

8.2 The Loading Curve and the SMA Response

Consider the scissor mechanism shown in Fig. 8.2, which is to be actuated by a SMA spring actuator : As with the previous example, let us assume that the load W is 100 N and the stroke S needed is 100 mm. In order to design the spring actuator the main idea is to (a) find the stroke of the spring when the load moves through a distance of 100 mm and (b) find the load on the spring at the beginning and the end of the stroke We will use this information to design the SMA actuator. If we want to compute these using a conventional Newton's laws, it will be very complicated. However, in the current instance, the use of the virtual work principle makes our life significantly simpler, as explained next.

- **DOF** First we note that the mechanism is a single degree of freedom mechanism and the most convenient parameter $q_1=\theta$ as shown in Fig. 8.2.
- **KINEMATICS** we note that the key motions are the vertical motion y of the platform and the horizontal motion x of the point P, both of which we can express in terms of the geometrical parameter L and the kinematical parameter θ. Thus, from the geometry of the mechanism, we have

$$x=2L\sin\theta\implies dx=2L\cos\theta\,d\theta \tag{8.10}$$

$$y=4L\cos\theta\implies dy=-4L\sin\theta\,d\theta \tag{8.11}$$

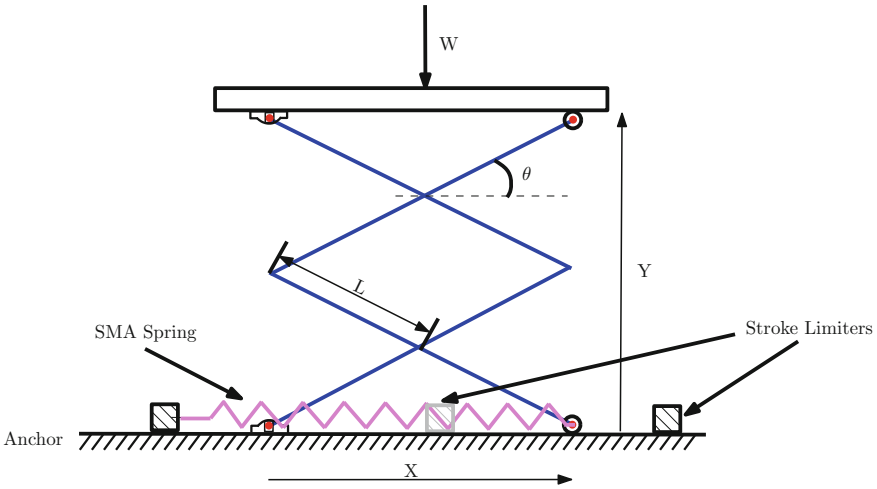

Fig. 8.2 A scissor mechanism actuated by SMA springs with hard stops

- **STROKE** We next have to decide what length L must be to obtain a desired stroke of 100 mm. Based on experience, we expect that if θ is too close to 90° (fully flattened) or 0° fully deployed, the mechanism will "jam" and require very large forces to operate. It is thus better to choose to operate around the 45° position. We will choose an angle range of 60° to provide the stroke. So we set the operating angles to range from 15 to 60°. With this choice. we see that the stroke is given by

$$S_l = \Delta y = y(\theta = 15^\circ) - y(\theta = 60^\circ) = 1.863L \tag{8.12}$$

$$S_s = \Delta x = x(\theta = 15^\circ) - x(\theta = 60^\circ) = -1.21L \tag{8.13}$$

where S_l is the stroke of the load and S_s is the stroke of the spring. Since we want a stroke of 100 mm for the load, we can therefore find the mechanism arm length L as

$$L = 100/1.863 = 53.7 \text{ mm} \implies S_s = -1.21 \times 53.7 = -65.2 \text{ mm} \tag{8.14}$$

Thus, the spring should be designed for a stroke of approximately 74.3 mm. To guarantee this, we will design a hard stop at $x = 2L \sin 15^\circ = 27.80$ mm and another stop at $2L \sin 60^\circ = 93.01$ mm. This provides a stroke of 65.2 mm.

- **FORCE** While the external applied force is constant, due to the changing geometry of the mechanism, the force at the spring will change. We need to estimate the force needed at the beginning and the end of the stroke.

  For this, we will use the principle of virtual work and note that

$$W\,dy + F_{spr}\,dx = 0 \implies -W \times 4L\cos\theta\,d\theta + F_{spr} \times 2L\sin\theta\,d\theta = 0 \tag{8.15}$$

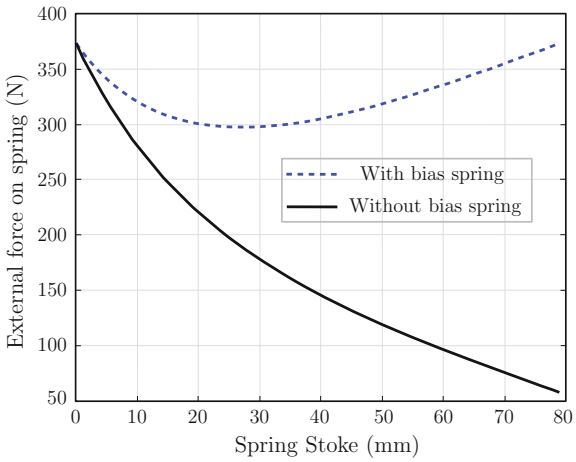

Fig. 8.3 Effect of using a bias spring results in smoother response—force versus stroke plot

Since this is true for all virtual displacements $d\theta$ we obtain

$$F_{spr} = \frac{2 \times W}{\tan\theta} \tag{8.16}$$

For $W = 100$ N, we obtain

$$F_{max} = \frac{2 \times 100}{\tan(15°)} = 746 \text{ N}, \quad F_{min} = \frac{2 \times 100}{\tan(60°)} = 115.47 \text{ N} \tag{8.17}$$

A plot of the load on the spring versus the spring stroke is shown by the solid line in Fig. 8.3. Notice from the previous calculations, that a very large force is required at the 15° angle and that the force drops sharply as the stroke progresses. We note that the maximum load on the spring increased three fold while the stroke decreased only slightly from 100 to 74 mm. This is the inefficiency as a result of not following the design rules that were listed in Chap. 2. Specifically, the design shown above does not provide a "short and direct load path", that is, the load exerted by the horizontally mounted spring is perpendicular the actual load which is vertical and is therefore not a direct path. The need to change the load direction is the cause of inefficiency. In fact, every violation of the design rules will cause new difficulties in the actual embodiment of our ideas. Nevertheless, sometimes it is not possible to satisfy all of them due to design constraints and we must be prepared for such inefficiencies.

- **SPRING DESIGN** We can now design the spring. We will use the required stroke of 65.2 mm and the max force of 746 N for the design and then verify whether the minimum force of 115 N is sufficient to cause the spring to revert back to the cold extended position. We will not detail the design procedure here but simply follow Chap. 7.

We quickly find that a single spring would be unsuitable. However, if we use two springs in parallel we get more reasonable answers. From the design rules that we have quoted in Chap. 2, this might be reasonable to maintain symmetry of the device since the scissor mechanism will have two sides that have to be actuated and we could have a spring on each side. However, it violates the design principle of not over-constraining because we are using two actuators for a one degree of freedom system. This means that we will have to be ready to deal with the problems of having to synchronize the two springs to act together. Given the highly non-linear nature of the response, we should be aware of the fact that this would not be trivial task especially for larger strokes.

Nevertheless, we will proceed by designing springs for half of the loads but with the same stroke. We thus have to design a spring to carry a maximum load of 373 N and a minimum load of 58 N with a stroke of about 65 mm. Following the approach presented in Chap. 7, we find that a preliminary design with a spring index of 4 seems suitable with a wire diameter of 6 mm an OD of 30 mm and 17 active coils with a coil bound length of 168 mm. It is also apparent that the 57.7 N force appears to be sufficient to fully extend the spring.

8.3 3-D Design

Please don't confuse the mechanism shown in Fig. 8.2 as the real design. Real designs live in a 3-D world and we have to consider a variety of factors to do this: A fully featured design is shown in Fig. 8.4.

Notice the fact that the design exploits symmetry. Note also how the SMA actuators are integrated into the base. Notice also how the "stroke limiter" is actually

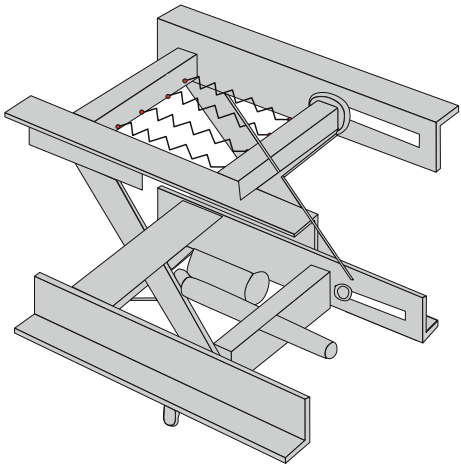

Fig. 8.4 Real world scissor mechanism

embodied in the slots. By moving the strain limiter to one platform and the SMA springs into the other platform, the design achieves separation of functions in a very efficient way.

8.4 Bias Forces

In the previous example we note that there is a huge excursion in the external force on the spring. The force starts out at 373 N when the spring is fully extended and decreases to just 57.7 N when the spring is fully retracted. This will give rise to a "jerky" motion and sometimes will cause the spring to cease contracting (due to friction) especially when it gets to the fully contracted mode. We can even this out by attaching a *bias spring* to the mechanism . The bias spring is to resist the contraction of the SMA spring. Thus, it is designed to be inactive when the SMA is fully extended and gradually increases its resistive force as the SMA spring contracts.

We choose the bias spring as follows: When the SMA spring is fully extended, the bias spring is free (i.e. coil bound) as the SMA contracts the bias spring extends. So as the SMA spring goes through its stroke of 65 mm, the bias spring is designed to exert a "compensatory force" of $373 - 58 = 315$ N thus the spring rate of the bias spring would be $315/65 = 4.85$ N/mm. When this is added, the SMA works against an almost constant force of 373 N (see the dotted line in Fig. 8.3) and its response will be much smoother. Thus, bias springs can be used to change the response and provide a reasonably "flat" or constant force against which the the SMA spring would work.

Reference

1. Reddy J (2002) Energy principles and variational methods in applied mechanics, 2nd edn. Wiley, New York

Chapter 9
Fatigue of SMAs

With growing applications of SMA components in different engineering applications, the issue of material performance over its designed life is of great concern to researchers lately. Fatigue studies in SMAs is still an unsolved puzzle and we wish to highlight some important items that affect the designers and some open questions in SMA fatigue areas.

As pointed out by Eggeler et al. [1], numerous factors can influence the fatigue life of SMAs. Factors include material, chemical composition, material impurity (i.e., particle size distribution), type of loading, applied load magnitude, extent of deformation or plateau strains, processing conditions, defect accumulation, accumulated plastic deformation, microstructure and operating temperature being a few to list [1]. In the following sections, we wish to highlight a high level classification of fatigue in SMAs and theories employed for analyzing fatigue data.

9.1 Structural and Functional Fatigue in SMAs

Eggeler et al. [1] suggest subdividing fatigue in SMAs into two categories namely "structural fatigue" and "functional fatigue" (see Fig. 9.1). SMAs failing like other engineering materials due to repeated high cyclic mechanical loads is classified under "structural fatigue" which is generally accompanied by defect accumulation, crack initiation (normally micro-structure controlled and growth–geometry dependent), crack propagation to critical lengths and final fracture [1, 2]. As seen with fatigue in classical materials, one can expect cracks to initiate from specimen surface due to surface irregularities (local surface conditions), scratches or inclusions that can act as stress raisers. Once the cracks reaches critical crack lengths, final failure is observed as the component can no longer handle external loads. In SMAs, the brittle TiO_2 layer on the NiTi bulk material acts a stress raiser in the presence of any surface irregularities or scratches that results in crack initiation and propagation [2].

A. Rao et al., *Design of Shape Memory Alloy (SMA) Actuators*,
SpringerBriefs in Computational Mechanics, DOI 10.1007/978-3-319-03188-0_9

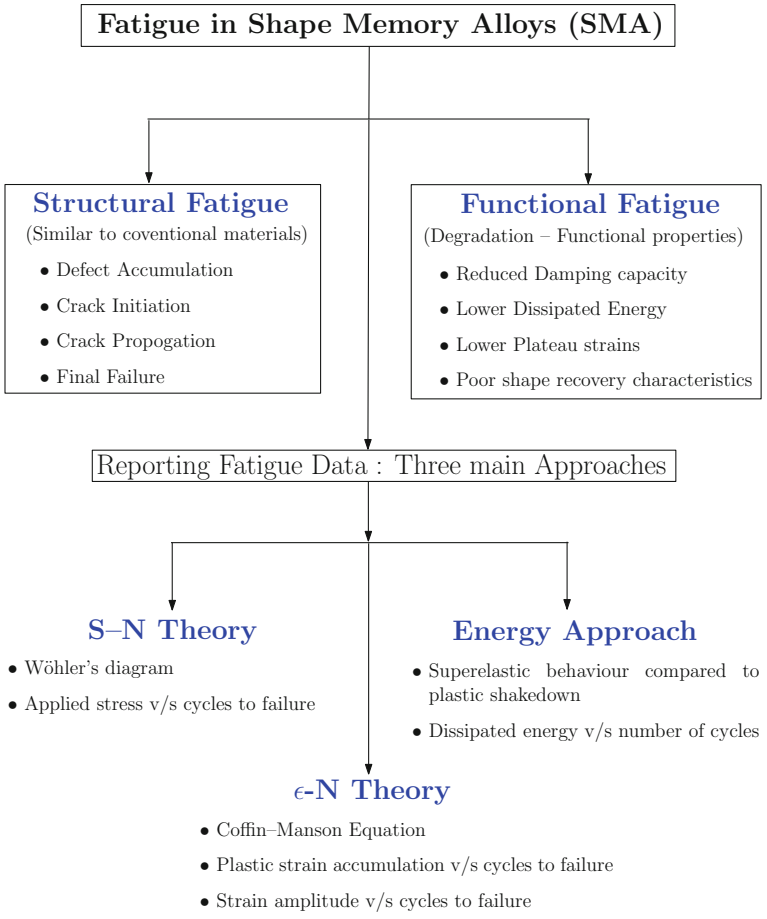

Fig. 9.1 Figure describes classification of SMA fatigue into two categories namely "structural fatigue" and "functional fatigue" by Eggeler et al. [1, 2]. It also lists the three traditional approaches (theories) of reporting fatigue data in SMA literature primarily devoted for capturing superelastic effects (mostly mechanical loading effects with temperature being an external control parameter)

However, unlike classical material systems, SMAs can also suffer a degradation of functional properties during cyclic thermomechanical loading i.e. either superelastic effect or shape memory effect. Such degradation is classified under "functional fatigue" and is predominately due to various intrinsic microstructural changes that result in changes of component transformation temperatures, loss of memory, reduction in damping capacity (i.e. area of hysteresis), lower transformation strains etc. being a few to list. [1]. In short, "functional fatigue" is the inability of the material to perform its intended functionalities.

Structural fatigue is well documented in classical fatigue theories on metals, ceramics and composites available in open literature and the same knowledge can be

extended to study structural fatigue of SMA's. Hence, here, in this section, we would like to focus more on "functional fatigue" aspects as they affect the performance of actuators immensely especially those subjected to thermal cycling and it plays a vital role from designers perspective.

Studies suggest that the degradation of these functional properties (especially in thermal cycling) is primarily attributed to an increase in dislocation density during repeated thermal cycling [1]. Initial cycles result in creation of many dislocations and their formation progressively decrease with increasing cycles. This is mainly due to "pockets/islands" of material that form stabilized martensite variants (plastically deformed martensite) that fail to transform back to parent austentic phase during thermal cycling [1, 2]. With repeated cycles, these "islands/pockets" of plastically deformed martensite variants not transforming leads to accumulation of plastic strains over cycles [1, 2]. With the diffusion processes in NiTi SMAs being practically insignificant as the the operating conditions for shape memory effect are much much lower than the melting temperature of NiTi, it makes it even harder for these "islands/pockets" to re-transform or not affect the SMA performance [1].

These "islands/pockets" are also observed in superelastic cycling where some austentite is untransformed even under higher stresses and also some martensite lattices that slips at higher load levels leading to permanent strains. Fracture studies have shown that crack growth in stress induced martensite (SIM) and these pockets/islands are much slower than that of austenite as the stress intensity factor of SIM/pockets is much higher than austenite which makes it harder for cracks to propagate [1, 2].

Ramaiah and co-workers [2] in their study of thermal cycling of SMA wires observe presence of multiple internal cracks at the core of the fracture region linking to each other. They proposed that during thermal cycling, there exists a temperature gradient across the wire cross-section (highest at the core and reducing radially) that causes the volume fraction of phases to change radially along the temperature gradient [1, 2]. This temperature gradient could be attributed to changes in local electrical resistance due to crack generation and propagation during thermal cycling. Such a temperature gradient indicates higher austenite at the core with a decrease radially to the surface and vice versa for martensite volume fraction [1, 2]. In events of forced external cooling (forced convection scenario), this could be exaggerated as the outer layers are being maintained at lower temperatures than the core and thus forcing the outer sections to be martensite rich as they are not "hot enough" to transform back to austenite instantaneously [1, 2]. This would imply that the softer martensite phase is susceptible to larger deformation compared to inner austenitic core leading to higher crack density and propensity to failure. Further, with the existence of the temperature gradient across the specimen, complete recovery of bulk material is constrained across adjacent layers through the specimen thickness. This could lead to development of shear stress/strains inhomogeneity along the longitudinal loading direction of the wire and strain inhomogeneity between adjacent layers [1, 2]. These internal stresses can act as stress risers and can cause faster crack propagations leading to premature failure internally [1, 2].

Figure 9.2 shows a description of these phases across an SMA wire crossection subjected to thermal cycling. The cross section can be treated as a composite of many phases with strain inhomogeneity between phases that lead to formation of

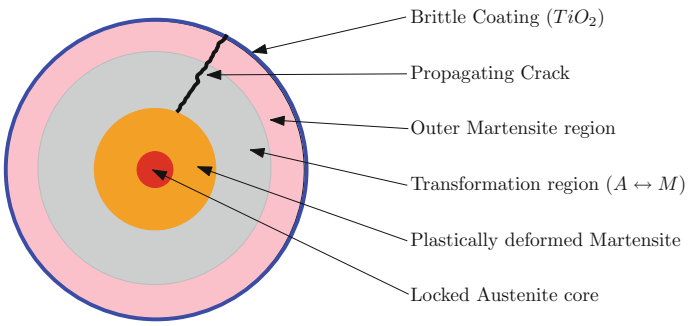

Fig. 9.2 Figure shows a cross section of actuator SMA wire/spring after several thermal cycles. The cross section can be conceived as a composite of many phases that affect the performance of the SMA actuator—loss in functionality termed as functional fatigue

these "islands/pockets" locked in the material. This eventually leads to functional degradation of its characteristic properties like loss of memory, stroke degradation, shift in transformation temperatures etc. Though, the material might not have completely fractured, such a degradation of characteristic properties of SMA can render them unproductive in their desired application. Further, the issue of crack propagation due to internal stresses and cracks originating due to surface irregularities (local surface conditions), scratches or inclusions along with the inherent sensitiveness to operating temperature fluctuations, material composition/finish/impurity, processing conditions, and so on make fatigue in SMA's a very tricky subject to understand [1, 2]. We believe better understanding of these issues experimentally and development of appropriate theories and tools for studying functional fatigue is still an open area of research.

9.2 Reporting Fatigue Data

Over the years, numerous experimental data concerning SMA fatigue under superelastic conditions and thermal cycling (mimicking shape memory effects) have been reported. Most of the focus has been in understanding fatigue behaviour of SMA wires under tension with some attention towards bending—rotating tests of SMA wire and functional fatigue of actuator springs [3–9].

The traditional approach of reporting fatigue data in most of the SMA literature (see Fig. 9.1) is using any of the traditional fatigue theories listed below :

1. **S-N theory** or Wöhler's diagram to capture "applied stress verses cycles to failure" plots [10–15].
2. **ϵ-N theory** that plot "strain amplitude verses cycles to failure" or employ "Coffin–Manson (modified Coffin–Manson approaches in some cases)" under low cycle fatigue cases to obtain "plastic strain accumulation verses cycles to failure" relationships [16–21] .

3. **Energy approach** that develop relationships between dissipated energy verses number of cycles under stabilized superelastic effects. Superelastic behaviour is compared to plastic shakedown for these analysis (see [22–27] for examples).

Such empirical theories solely focus on mechanical loading effects with temperature being a external control parameter. Generally with such theories, relationships are developed for each temperature case and are limited to analyzing superelastic effects [27, 28]. Given the fact that thermomechanical responses of SMA are sensitive to operating temperature fluctuations under superelastic conditions, such traditional theories are incapable of these effects. In addition, these theories don't throw much light in capturing functional degradation charecterisitcs discussed in the previous section under coupled thermomechanical cycling of SMA components.

To capture such complex thermomechanical responses, we believe theories considering both mechanical and thermal effects in a single framework must be employed rather than modifying classical fatigue theories developed for pure mechanical loading cycles [27, 28]. This is an open area of research and the authors are working in this regard.

References

1. Eggeler G, Hornbogen E, Yawny A, Heckmann A, Wagner M (2004) Structural and functional fatigue of niti shape memory alloys. Mater Sci Eng A 378(1):24–33
2. Ramaiah K, Saikrishna C, Ranganath V, Buravalla V, Bhaumik S (2011) Fracture of thermally activated niti shape memory alloy wires. Mater Sci Eng A 528(16):5502–5510
3. Tobushi H, Ohashi Y, Hori T, Yamamoto H (1992) Cyclic deformation of tini shape-memory alloy helical spring. Exp Mech 32(4):304–308
4. Tamura H, Mitose K, Suzuki Y (1995) Fatigue properties of ti-ni shape memory alloy springs. J Phys IV 5(8):C8–617
5. Grossmann C, Frenzel J, Sampath V, Depka T, Oppenkowski A, Somsen C, Neuking K, Theisen W, Eggeler G (2008) Processing and property assessment of niti and niticu shape memory actuator springs. Materialwissenschaft und Werkstofftechnik 39(8):499–510
6. Grossmann C, Frenzel J, Sampath V, Depka T, Eggeler G (2009) Elementary transformation and deformation processes and the cyclic stability of niti and niticu shape memory spring actuators. Metall Mater Trans A 40(11):2530–2544
7. Tobushi H, Hachisuka T, Yamada S, Lin P (1997) Rotating bending fatigue of a tini shape memory alloy wire. Mech Mater 26(1):35–42
8. de Azevedo Bahia M, Fonseca Dias R, Buono V (2006) The influence of high amplitude cyclic straining on the behaviour of superelastic niti. Int J Fatigue 28(9):1087–1091
9. Miyazaki S, Mizukoshi K, Ueki T, Sakuma T, Liu Y (1999) Fatigue life of ti-50 at.% ni and ti-40ni-10cu (at.%) shape memory alloy wires. Mater Sci Eng A 273:658–663
10. Miyazaki S (1990) Thermal and stress cycling effects and fatigue properties of ni-ti alloys. Butterworth-Heinemann, Engineering Aspects of Shape, Memory Alloys(UK), pp 394–413
11. Melton K, Mercier O (1979) Fatigue of niti thermoelastic martensites. Acta Metall 27(1): 137–144
12. Mammano G, Dragoni E (2011) Functional fatigue of shape memory wires under constant-stress and constant-strain loading conditions. Procedia Eng 10:3692–3707
13. Kang G, Kan Q, Yu C, Song D, Liu Y (2012) Whole life transformation ratchetting and fatigue of superelastic niti alloy under uniaxial stress-controlled cyclic loading. Mater Sci Eng A 535:228–234

14. Ataalla T, Leary M, Subic A (2012) Functional fatigue of shape memory alloys. Sustain Automot Technol 2012:39–43
15. Scirè Mammano G, Dragoni E (2012) Functional fatigue of ni-ti shape memory wires under various loading conditions. Int J Fatigue
16. Bertacchini O, Lagoudas D, Calkins F, Mabe J (2008) Thermomechanical cyclic loading and fatigue life characterization of nickel rich niti shape-memory alloy actuators. In: The 15th international symposium on: smart structures and materials and nondestructive evaluation and health monitoring, pp 692916–692916, International Society for Optics and Photonics
17. Bertacchini O, Lagoudas D, Patoor E (2009) Thermomechanical transformation fatigue of tinicu sma actuators under a corrosive environment-part i: experimental results. Int J Fatigue 31(10):1571–1578
18. Lagoudas D, Miller D, Rong L, Kumar P (2009) Thermomechanical fatigue of shape memory alloys. Smart Mater Struct 18(8):085021
19. Figueiredo A, Modenesi P, Buono V (2009) Low-cycle fatigue life of superelastic niti wires. Int J Fatigue 31(4):751–758
20. Runciman A, Xu D, Pelton A, Ritchie R (2011) An equivalent strain/coffin-manson approach to multiaxial fatigue and life prediction in superelastic nitinol medical devices. Biomaterials 32(22):4987–4993
21. Maletta C, Sgambitterra E, Furgiuele F, Casati R, Tuissi A (2012) Fatigue of pseudoelastic niti within the stress-induced transformation regime: a modified coffin-manson approach. Smart Mater Struct 21(11):112001
22. Moumni Z, Van Herpen A, Riberty P (2005) Fatigue analysis of shape memory alloys: energy approach. Smart Mater Struct 14(5):S287
23. Moumni Z, Zaki W, Maitournam H et al (2009) Cyclic behavior and energy approach to the fatigue of shape memory alloys. J Mech Mater Struct 4(2):395–411
24. Soul H, Isalgue A, Yawny A, Torra V, Lovey F (2010) Pseudoelastic fatigue of niti wires: frequency and size effects on damping capacity. Smart Mater Struct 19(8):085006
25. Dunand-Châtellet C, Moumni Z (2012) Experimental analysis of the fatigue of shape memory alloys through power-law statistics. Int J Fatigue 36(1):163–170
26. Gloanec A, Bilotta G, Gerland M (2013) Deformation mechanisms in a tini shape memory alloy during cyclic loading. Mater Sci Eng A 564: 351–358
27. Rao A, Srinivasa A (2013) Experiments on functional fatigue of thermally activated shape memory alloy springs and correlations with driving force intensity. In: SPIE smart structures and materials + nondestructive evaluation and health monitoring, pp 86890T–86890T, International Society for Optics and Photonics
28. Doraiswamy S, Rao A, Srinivasa A (2013) A two species thermodynamic preisach approach for simulating superelastic responses of shape memory alloys under tension and bending loading conditions. In: SPIE smart structures and materials+ nondestructive evaluation and health monitoring, pp 86890X–86890X, International Society for Optics and Photonics