1

Introduction

Christoph Haberland, Mahmoud Kadkhodaei and Mohammad H. Elahinia

This chapter is on introductory materials on shape memory alloys (SMA) behavior. Shape memory effect, and superelasticity will be covered. In this context, the benefits of SMAs in actuation will be highlighted. Phase transformation as the underlying phenomenon for the unique properties of these alloys will be presented and discussed. Different actuation mechanisms and designs will be presented and compared. Example of aerospace, automotive, industrial, and biomedical applications of SMA actuation will be used to discuss the benefits and limitations of actuations using these alloys. Particular attention will be on rotary SMA actuators. This type of actuators will be used as a continuous example throughout the book.

1.1 Shape memory alloys

SMAs are distinguished from conventional metallic materials by their ability to restore their shape after large deformations, which can significantly exceed the actual elastic deformability of the material. This is referred to as
Shape memory effect (SME) characteristic and was first observed in 1932 in a gold–cadmium\(^1\) alloy following a thermally induced change in the crystal structure [1, 2]. Nearly 20 years later, Chang and Read [3] identified the fundamental mechanisms in the crystal lattice and attributed this phenomenon to a thermoelastic behavior of the martensitic phase. In the following years, the SME was observed in other alloys, more than 25 binary, ternary, or quaternary alloys and alloy systems are now known to show shape memory properties [4]. In contrast to nickel–titanium (NiTi), majority of these systems however have only been considered in principle and as such have not yet achieved any practical technological importance [5]. In NiTi, the SME was observed for the first time by Buehler et al. at the US Naval Ordnance Laboratory (NOL, White Oak, Maryland) in the 1960s [6, 7]. Because of the place of discovery, besides NiTi or TiNi, the term nitinol is also commonly used for this alloy. The application of SMAs spans a wide range of length scales, and these alloys are now used in multiscale devices ranging from nanoactuators used in nanoelectromechanical systems to very large devices used in civil engineering applications. SMA devices range from simple parts like cell phone antennas or eyeglass frames to complicated devices in mechanical [8–10], biomechanical [11–13], aerospace [14], and civil engineering [15].

Today, more than 90% of all commercial shape memory applications are based on binary NiTi or ternary NiTi-Cu and NiTi-Nb alloys [5]. This is despite the relatively high world market prices for high-purity nickel and especially for high-purity titanium. It should be noted that the price of Fe- or Cu-based SMAs is lower. Additionally, as explained in Chapter 6, the manufacturing processes of NiTi are complex and challenging, which adds to the production costs. The main reason for the dominance of NiTi-based SMAs is due to their excellent structural and functional properties. The SME in NiTi allows for relatively large reversible deformations of up to 8%, characterized by good functional stability [5, 16–18]. In addition, NiTi has good wear and corrosion resistance and biocompatible properties, making it an attractive candidate for various medical applications such as surgical tools, stents, or orthodontic wires [19–22]. Furthermore, the low stiffness of NiTi attracts interest for use in bone implant applications and in regenerative medicine [23]. For actuation and motion control applications, this alloy can be easily heated by passing an electrical current while offering several advantages for system miniaturization such as high power-to-mass ratio,

\(^1\) The mechanism of SME in AuCd is related to aging and is considered to be different from that of NiTi, which is due to detwinning.
maintainability, reliability, and clean and silent actuation. Due to its outstanding predominant role amongst other SMAs, in this book we mainly focus on NiTi.

The fundamental reason for the unique behavior of these alloys is due to the martensitic phase transformation. Originally, this term referred to the crystallographic phase transformation, which results in rapid cooling to a specific crystallographic phase in the Fe–C structure. This is also the basic mechanism in the hardening of steels. With increasing scientific understanding of the underlying mechanisms, the term martensitic phase transformation has been extended to a variety of other alloys (e.g., Fe–Ni or Cu–Zn) and even other material systems (e.g., some specific polymers or ceramics). In general, the martensitic phase transformation is a specific type of a crystallographic phase transformation in the solid state. When cooling the material from the high-temperature phase (β-phase, austenite), the material transforms into a low-temperature phase (α-phase, martensite). This transformation is diffusionless and therefore can occur at very low temperatures. Since no diffusion processes take place, the local concentration of the chemical composition is not affected; only the crystal structure changes. Usually this change in crystal structure is driven by a shear process, which can be described by a coordinated, cooperative movement of atoms in the crystal lattice. This results in the formation of plate, lenticular, or acicular martensite crystals.

In polycrystalline materials, the martensitic transformation is typically a heterogeneous nucleation process that starts at favored nucleation sites such as phase or grain boundaries, precipitates, or crystal defects. It should be emphasized that alloys that demonstrate martensitic transformations do not necessarily have the shape memory properties. This is due to the fact that the martensitic transformation induces high mechanical stresses. The compensation of these stresses is usually associated with irreversible processes, such as dislocation slip. In alloy systems, which have shape memory properties, the martensitic transformation is to a large extent reversible. In these alloys, the transformation stresses are compensated by twinning processes and self-accommodation of favored martensite variants. Due to the fact that no additional lattice defects are created in these reversible processes, this type of martensitic transformation is also called thermoelastic martensitic phase transformation [24, 25].

Generally, the SME can occur in three related phenomena. Two of these are the thermally induced one-way and two-way shape memory (pseudoplastic) effects. The third effect is the pseudoelasticity, which is also known as superelasticity and mechanical memory. Several factors define which of these effects takes place. The most important factor is the alloy composition.
Others include the thermomechanical treatment, the microstructure of the material, and the ambient temperature [26, 27]. Generally, the binary alloy nickel–titanium is known to show all three effects depending on these factors.

Shape memory effect, shown in Figure 1.1b, as the ability of these alloys to recover a certain amount of unrecovered strain upon heating, takes place when the material is loaded such that the structure reaches the detwinned martensite phase and then unloaded while the temperature is below the austenite start temperature ($A_s$). Heating the material at this stage to austenite will lead to strain recovery, and the material will regain its original shape. The combined stress–strain-temperature diagram as shown in Figure 1.1b can better explain this phenomenon. During SME, as the result of cooling, the twinned variants directly transform from the austenitic phase (see Figure 1.1a, transformation $\beta \rightarrow \alpha^+ / \alpha^- $). In loading the material transforms to detwinned martensite resulting in large deformation before reaching the yield stress and therefore without dislocation, beyond which dislocation plasticity starts (transformation $\alpha^+ / \alpha^- \rightarrow \alpha^+ / \alpha^- $). During the heating transformation to austenite and macroscopic initial shape recovery takes place (transformation $\alpha^+ \rightarrow \beta $). To complete the cycle by cooling, the austenite transforms into self accommodating twinned martensite, without an apparent shape change (transformation $\beta \rightarrow \alpha^+ / \alpha^- $).

Starting from point A, as shown in Figure 1.1b, the material is initially in the austenite phase. Cooling the alloy to a temperature below its martensite

---

**Figure 1.1** Shape memory effect path in stress–strain–temperature space
finish temperature ($M_f$) will result in the twinned martensite crystal, point B. This forward transformation starts when the temperature reaches martensite start ($M_s$). Loading the alloy at a constant temperature from point B results in elastic deformation of the martensite phase. At the critical stress levels $\sigma_s^e$ and $\sigma_f^e$, the transformation to the detwinned martensite phase starts and finishes. This part of the loading cycle induces a large strain with minimal increase in the stress to point C. As loading continues, the detwinned martensite is elastically deformed beyond point C. Further loading will induce plastic deformation at the detwinned martensite phase.

Unloading by removing the applied stress at a constant temperature from point C results in a linear strain recovery to point D. During this step, the material remains in the detwinned martensite phase. At point D, the remaining strain is called the residual strain $\epsilon_{SME}$. This strain can be recovered by increasing the temperature. By heating the alloy above the austenite start temperature ($A_s$) at point E, the transformation from the detwinned martensite to the austenite phase starts. At point F when the alloy passes the austenite finish temperature ($A_f$), the material is fully austenite and recovery of the residual strain completes.

In contrast to the thermally activated effects, in superelasticity, no temperature change is required. Instead, an external mechanical load is used to induce phase transformation. In this phenomenon, the detwinned and favorably oriented martensite variants directly transform from the austenitic phase (see Figure 1.2, transformation $\beta \rightarrow \alpha^\prime$). To complete the cycle, the material is unloaded before reaching the yield stress of the detwinned martensite beyond which dislocation plasticity starts. During the unloading, the martensite transforms back into austenite. This results in macroscopically recovering the initial shape. Due to this behavior, this phenomenon is sometimes called “rubberlike material behavior,” but more common are the terms superelasticity and pseudoelasticity. It is worth noting that from the materials science point of view, this behavior actually is not elastic.

Despite the fact that no thermal activation is required to observe pseudoelasticity, it can only be activated in the temperature range $A_f < T < M_d$ where $A_f$ refers to the austenite finish temperature (where the transformation into austenite is completed) and $M_d$ refers to the martensite dead temperature (or martensite destruct temperature) above which no stress-induced martensite can be formed because the high-temperature phase (austenite) is stable. As Figure 1.2b shows, during loading and unloading, the material shows a hysteresis because of different stress levels for the actual transformation. It is also worth pointing out that this hysteresis corresponds well to the mechanical behavior of human tissues under a mechanical load.
The superelastic behavior takes place at temperatures above the austenite finish temperature ($A_{f}$) where the material is fully austenitic. As shown in Figure 1.2b, loading the material from point A initially induces elastic deformation of the austenite phase. Further loading leads to stress-induced formation of the detwinned martensite crystal and the macroscopically elastic-like deformation of this structure to point B. The transformation strain generated during this forward transformation from austenite to martensite is fully recovered in the reverse transformation. This takes place during the unloading from point B, which involves an initial elastic recovery of martensite followed by the transformation to austenite (recovering seemingly plastic $\epsilon_{PE}$) and finally the elastic recovery of austenite to point A. At this point, the strain is completely recovered.

### 1.2 Metallurgy of NiTi

For the nickel–titanium alloy, the high-temperature phase austenite has a body-centered cubic structure (B2). This structure transforms during cooling into the monoclinic lattice structure of the martensite (B19), whereby each lattice atom retains its nearest neighbor atoms. In binary NiTi alloys, the thermally induced SME and the pseudoelasticity only occur in a narrow range of
the chemical composition. Only around the stoichiometric composition, the intermetallic phase NiTi (B2) can exist in the absence of other phases under equilibrium conditions (see gray area of the binary phase diagram shown in Figure 1.3). In addition, from the phase diagram in Figure 1.3, it is evident that the intermetallic phase NiTi has a limited solubility for titanium of less than 51 at.%. This solubility limit is almost independent of the temperature [29]. In these Ti-rich compositions and in almost equiatomic balances, the NiTi phase is stable even at low temperatures. The maximum nickel content of the NiTi phase cannot exceed 57 at.%. Below 1118°C, the solubility limit of nickel in the phase NiTi decreases with decreasing temperature. At lower temperatures, this overstoichiometric balance of the B2 phase does not exist in absence of the other phases. However, by rapid quenching from this overstoichiometric balance, the Ni-rich composition of the B2 phase can be “frozen,” and the material approaches a metastable state without secondary phases.

However, if this metastable state is subjected to heat treatment, it will approach a state of thermodynamic equilibrium by complex diffusion and precipitation processes [30]. Depending on the temperature and the aging time, a two-phase state of the phase NiTi and Ni-rich precipitations will be formed. This results in a depletion of nickel in the B2 phase due to precipitation of Ni-rich phases.

Figure 1.4a shows a section of the binary phase diagram; in Figure 1.4b, the precipitation kinetics for Ni-rich NiTi (Ni₅₂Ti₄₈) are shown. In the beginning,

![Figure 1.3](image_url)  
**Figure 1.3** Phase diagram of the system nickel–titanium. The single-phase NiTi (B2) is shaded; important temperatures are highlighted. Reproduced with permission from Ref. [28], ASM International
at low temperatures and short aging, metastable precipitates of type \( \text{Ni}_4\text{Ti}_3 \) form. With longer durations and higher temperatures, metastable \( \text{Ni}_3\text{Ti}_2 \) precipitates are formed. Precipitation of stable \( \text{Ni}_3\text{Ti} \) only occurs at very long aging treatments. The \( \text{Ni}_4\text{Ti}_3 \) phase has a significant influence on the martensitic phase transformation. With the formation of these precipitates, stress fields arise in the crystal structure, which can lead to a multistage phase transformation. In this case, during cooling from the B2 (austenite) phase, the trigonal R-phase is formed as a premartensitic state so that the transformation sequence becomes \( \text{B2} \rightarrow \text{R} \rightarrow \text{B19}' \) \cite{33, 34}. \( \text{Ni}_4\text{Ti}_3 \) precipitates also act as nucleation sites for the martensite, and therefore, its existence reduces the critical stress to start the transformation. Moreover, the presence of the \( \text{Ni}_4\text{Ti}_3 \) particles hinders the dislocation movement, and their precipitation hardening effect leads to an increase in yield stress. As a result, the irreversible processes, which are usually associated with deformation, are reduced. Additionally, these effects contribute to an increase in the cyclic stability of the pseudoelastic effect. This in turn leads to a significant reduction in functional fatigue \cite{35–37}.

In this metastable state, the phase transformation temperatures of Ni-rich NiTi strongly depend on the nickel–titanium ratio of the B2 matrix \cite{33, 38, 39}. This relationship is shown in Figure 1.5 for both \( M_s \), the martensite start temperature, and \( T_0 \), temperature of the thermodynamic equilibrium of the B2 phase and the B19' phase. Binary NiTi SMAs with a substoichiometric nickel content, where NiTi and Ti-rich phases (Ti_2Ni, i.e., Ti_4NiO_2X) are in

---

**Figure 1.4** Section of the phase diagram of the system nickel–titanium (a). Isothermal transformation diagram of a \( \text{Ni}_{52}\text{Ti}_{48} \) alloy (b). Reproduced with permission from Refs. [30–32], Elsevier

---

Shape Memory Alloy Actuators
the thermodynamic equilibrium, show no significant effect (see range 1 in Figure 1.5). On the other hand, for nickel contents higher than 49.7 at.\%, both temperatures continuously decrease with increasing nickel content (see range 2 in Figure 1.5). This dependence can be used to adjust the phase transformation temperatures by the nickel–titanium ratio. Alloys can thus be produced, which show either a thermal or pseudoelastic memory effect in the ambient temperature. It should however be noted that during aging of Ni-rich NiTi, the previously described precipitation processes again affect the nickel–titanium balance of the B2 matrix. During the formation of Ni-rich phases, the B2 phase is depleted of nickel, which results in an increase of the transformation temperatures [33, 39].

The effect of impurity related phases should also be considered. During high-temperature processing, the pickup of impurities, for example, carbon and/or oxygen, can result in the formation of Ti-rich phases since the B2 phase has a low solubility for both elements [41]. Carbon forms carbides of type TiC [31, 42–44], while oxygen is dissolved in the Ti2Ni phase and forms a stable phase of type Ti4Ni2Ox [41, 45, 46]. In addition to the degradation of functional and structural properties due to the impurity pickup, the formation of these Ti-rich phases also results in a shift of the nickel–titanium balance in favor of the nickel content. In contrast to precipitation of Ni-rich phases, this causes a decrease in transformation temperatures. These effects

**Figure 1.5** Influence of nominal nickel concentration on martensite start and thermodynamic equilibrium temperature. Reproduced with permission from Ref. [40], Elsevier
must therefore be given high attention in manufacturing and processing of NiTi alloys.

1.3 Thermomechanical Behaviors

The crystalline structure of SMAs undergoes a solid–solid phase transformation when cooled from its stiff, high-temperature austenite (A) phase to its softer, low-temperature martensite (M) structure. The stress–temperature–transformation (phase) plot is a schematic representation of the transformation regions for SMAs. Usually, a stress–temperature–transformation plot shows the temperature along the abscissa and stress along the ordinate. A widely accepted stress–temperature–transformation plot of SMA materials is shown in Figure 1.6 [47]. As described earlier, SMAs as the result of the transformation between two phases can exhibit SME and pseudoelasticity. The lines in the plot show the phase boundaries that separate the two solid phases of an alloy. The stress–temperature–crystalline structure pattern during shape memory effect is depicted on the

![Figure 1.6](image)

*Figure 1.6* Stress–temperature–transformation (phase) plot of a shape memory material depicts the stable area for each crystalline structure. Crystal transformation takes place as the result of variation in stress and temperature (s refers to transformation start and f to transformation finish, respectively)
transformation diagram (Figure 1.7), to complement the thermomechanical and crystalline structure changes shown in Figure 1.1 for more clarity. Superelasticity as the ability of SMAs to recover a large amount of strain through mechanical loading/unloading is shown in Figure 1.2. The stress–temperature–crystalline structure cycle of the superelastic behavior is shown on the transformation plot in Figure 1.8.

As shown in Figures 1.6, 1.7, and 1.8, the four transformation temperatures that define the behavior of SMAs are stress dependent. The two parameters

\[ C_A : \frac{1}{C_A} = \frac{dA_s}{d\sigma} = \frac{dA_f}{d\sigma} \quad \text{and} \quad C_M : \frac{1}{C_M} = \frac{dM_s}{d\sigma} = \frac{dM_f}{d\sigma} \]

quantify the effect of stress on austenite and martensite transformation temperatures, respectively.

These temperatures are also affected by the thermomechanical history of the alloys. In single crystals and polycrystals of CuAlZnMn, for example, the reverse martensitic transformation is influenced by the history of the forward transformation [48]. The transformation path from austenite to martensite, the deformation in martensite, and the annealing of martensite under stress affect the transformation temperatures. The incomplete and complete

![Diagram of Shape Memory Effect and Crystalline Changes](image-url)

**Figure 1.7** Shape memory effect and the associated crystalline changes presented in a stress–temperature–crystalline structure transformation diagram
transformation cycling, that is, thermal cycling between the forward and the reverse transformation finish temperatures, in NiTi leads to variations of the transformation temperatures. Referred to as the “thermal arrest memory effect,” the alloy remembers the arrested temperature in incomplete heating–cooling cycles by changing its transformation temperature in the consequent heating–cooling cycles [49–51]. When a reverse transformation is stopped (arrested) between $A_s$ and $A_f$ followed by cooling below $M_f$ or in other words a complete martensitic transformation after an interrupted heating cycle, a memory of the arrest is induced in the alloy [52–55]. In a consequent heating, this memory of the arrested temperature affects the thermomechanical behavior of the alloy at an intermediate point where the transformation shows an apparent completion. If after arresting between $A_s$ and $A_f$ during reverse transformation the temperature is reduced below $A_s$ but above $M_s$, the subsequent reverse transformation in the next heating cycle does not start at $A_s$. Instead, the transformation starts at the arrested temperature as shown in Figure 1.9 [56]. Similarly, for any interruption during the forward transformation, change in the martensite start temperature is observed if the specimen is subjected to heating to a temperature below $A_f$ followed by a cooling below $M_s$. Moreover, if the latter heating cycle continues to a temperature above $A_{sr}$, the reverse transformation does not start at $A_s$ but at a greater temperature.

![Figure 1.8](image-url) Superelasticity and the associated crystalline changes presented in a stress–temperature–crystalline structure transformation diagram.
1.4 Actuation

Recovering large strains as a result of heating is the actuation principle of SMAs. These alloys have a very high energy density; therefore, actuators that implement these alloys are compact and lightweight alternatives to other types of actuators such as DC motors and solenoids. In other words, SMA actuation is an effective way to reduce weight and to minimize the complexity of systems. The benefits of these alloys over other smart materials such as piezoelectric materials, electrostrictive materials, and magnetostrictive materials include the high force-to-weight ratio and large displacement capability. The disadvantages are slow actuation cycle due to longer cooling time, low energy efficiency due to conversion of heat to mechanical energy, and challenging motion control due to hysteresis, nonlinearities, parameter uncertainties, and difficulties in measuring state variables such as temperature.

Large mechanical stress is produced when SMA elements are heated beyond the austenite start temperature. NiTi actuators, for example, can apply up to 600 MPa stress. As a result, a 0.1 mm diameter NiTi wire can apply a force of 18.8 N, which is enough to lift 100 000 times its own weight. SMA actuators therefore offer the potential of significantly reducing the weight of active structures.

Figure 1.9  Shifts in $A_s$ of NiTi after an incomplete cooling cycle. Reproduced with permission from Ref. [56], Elsevier
An essential step in creating the shape memory actuation is shape setting. After the material is produced and is exposed to cold work and annealing, the desired memorized shape of the actuator is instilled during shape setting. For this step, the cold-worked material is constrained in a mandrel and heat-treated with a specific duration and a predetermined temperature, which is followed by a rapid cooling. The duration is in the order of about 10 min and the temperature is around 500°C. This process is essential for both shape memory actuators as well as superelastic devices. In addition to creating the desired shape, this final heat treatment defines mechanical properties of the part. A higher temperature results in a lower tensile strength. The shape setting for actuators can also be performed continuously as a part of the fabrication process for wires and tubes, which usually result in straight memorized forms. In most cases, however, the required heat treatment for an actuator with a more complex memorized form is a separate procedure. Heat treatments and shape setting are described in detail in chapter 6.

One-way SME refers to the ability of an SMA that is deformed at a low temperature to recover the deformation when heated to a higher temperature. To create a repeatable actuation, an external bias force is used. This force always opposes the SMA actuator and during cooling resets the actuator for the next cycle. In the less common two-way SMA actuators, the SMA elements can exhibit repeatable shape changes without a bias mechanical load. Instead, the SMA alternates between memorized shapes when subjected to a cyclic thermal load. In this case, the SMA memorizes a martensite shape that is different from the austenite shape. This behavior is called the two-way SME. In this mode, when the SMA is cooled from austenite to martensite, instead of adapting to a self-accommodated structure, some variants of the martensite are favored, and the martensite adopts a shape different from that of the self-accommodated structure. This memory is usually the result of training in the form of cyclic thermomechanical loadings or aging for precipitation under stress and/or under constraint. In addition to traditional heat treatment methods, laser scanning has been used for creating the two-way memory effect [57]. It is important to note that two-way SME provide very limited strains and therefore, it cannot be used in most applications.

Multiple SME is a well-known phenomenon in shape memory polymers (SMPs) in which a deformed SMP can recover from a temporary shape back to the original shape through a number of intermediate shapes in a step-by-step manner. While this behavior is ordinarily not observed in SMAs, it is possible to induce multiple SME in these alloys [58]. To this end, a laser beam is used to locally heat-treat NiTi. This treatment causes precipitation and local evaporation of nickel, which results in different local \( A_f \), which is higher than that of the base material. By adjusting the duration and number of pulses of
laser, it is possible to increase the transformation temperatures. If this NiTi element is deformed at low temperature (<\(A_s\)) and consequently heated, initially, the untreated sections of the alloy transform to austenite followed by the heat-treated sections, resulting in intermediate memorized configurations. Alternatively, a series of local cooling–loading–cooling cycles can be used to induce multiple SME [59].

SMAs have been studied and deployed to a wide range of actuation mechanisms. To this end, various forms of SMAs have been fabricated from mostly NiTi and NiTiCu alloys. In terms of actuation, the main difference between these two alloys is in a smaller temperature hysteresis in NiTiCu. Wires are the most common form for actuation. Other forms that have been used are springs, thin films, sheets, tubes, rods, and more recently three-dimensional fabricated shapes.

Thin film actuators offer miniaturization and high bandwidth. The higher actuation speed is achieved due to much faster convection cooling in films. NiTi films are produced by MEMS fabrication methods in vacuum deposition, electron irradiation, and magnetron sputtering [60]. The applications include microactuation, microswitching, and micropumping [61]. Similarly, sheets are formed to 3D actuators for folding origami, microactuation, and robotic applications [62, 63]. To this end, laser cutting and shape setting are used to achieve the final shape and functionality. It should be noted that methods for fabricating NiTi are described in detail in chapter 6. Figures 1.10 and 1.11 show a sheet-based NiTi actuator that move a series of cylinder-and-sleeve joints, which allow the elements to rotate relative to one another during actuation [64, 65]. The SMA sheets are prestrained at a low temperature (\(T < A_s\)) to approximately one-half

![Figure 1.10](image-url)  
**Figure 1.10** Schematic of the morphing wing structure. Reproduced with permission from Ref. [64], SPIE
of their maximum recoverable strain. Heating an SMA sheet causes a contraction that results in bending of the structure. During this process, the opposite sheet undergoes a tensile strain. The curvature of the structure is reversed by consequent heating of the opposing SMA sheet. The SMA elements are heated by resistance heating elements wound helically around each sheet. Heating all the SMA sheets on one side of the structure results in a uniform curvature. Individual SMA actuators may be selectively heated to achieve various shapes.

Rod SMA actuators have been developed for breaking concrete and boulder [66]. In this actuator, a series of mechanically parallel rods, upon heat activation, push two steel wedges apart by generating up to 900 kN of axial force. Structural control applications such as morphing aerospace systems require higher level of force and torque. To achieve this, a larger cross-sectional area
of the actuator is necessary. Beam actuators have been used to modify the outlet geometry of jet engines [67]. These beams are shape set to a curved geometry. By activating these NiTi beams, the attached elastic laminated structure bends causing change in the engine outlet area. This reduced engine outlet area leads to creating a mixture of jet engine fan flow with core flow. The result is a noise reduction in takeoff and landing. The outlet returns to normal for cruise flight regime.

As another aerospace application, Boeing, NASA, AFRL, and DARPA have developed NiTi tube rotary actuators also called “torque tubes”. Through cyclic thermomechanical loadings, these SMA tubes are trained as a heat-activated actuator with actuation torque of 17 N·m and angular displacement of 60°, with repeated actuation of over 10 000 cycles [68]. Such an actuator can be trained for two-way actuation and is envisioned for deploying and retracting tabs on rotor blades for reducing noise and for morphing blades [69]. In 2013, an SMA rotary actuator was integrated into the hinge line of a small flap on the trailing edge of a Boeing 737–800 wing and tested in flight in a collaboration between Boeing and the FAA [70, 71]. As Figure 1.12 shows, the actuator is made of two tubes and provides 45° of bidirectional trailing edge flap motion under realistic aerodynamic loads. This type of actuator can be placed at various locations on the wing. Rotary actuation of the edge flap enhances the performance in various flight regimes by improving lift and optimizing span loading. At small angles, better fuel economy and lower emissions are possible during high-speed cruise. On the other hand, with larger angles, lift is increased and noise is decreased during takeoff and approach.

Wires are the most common form of SMA actuators. Joule heating is mostly used as an effective and simple way for actuating SMA wires. Miga Motor
Company offers commercial wire-based linear and rotary actuators with integrated sensing and an on–off controller. As shown in Figure 1.13, to achieve a higher level of linear stroke, wire actuators are arranged in a cascade form. By using various length and diameter of wires, these actuators deliver a range of motion, force, and speed. A lifetime of over one million cycles at a cyclic actuation time of less than a second is achieved.

Standardization and consistent performance are needed for SMA actuators to find widespread acceptance in automotive applications. Figure 1.14 shows one such attempt as an SMA actuator which is optimized for automotive applications with long fatigue life. Using a superelastic antagonistic wire, an adaptive resetting mechanism is integrated in the actuator. This mechanism offers close-to-constant bias force which in turn allows for a larger useful stroke. The other benefit of this bias mechanism is that in reaction to the change of the ambient temperature, the superelastic wire adjusts its resistive force. When used for fuel door unlocking, this actuator reduced the weight of the actuation system by more than 90% [72].

One part of the actuation market is in individualized actuation and for replacing conventional actuators in current systems. In an attempt to address this need, as Figure 1.15 shows, an SMA spring actuator is designed with a 3D-printed polymer housing to create a thermal control valve actuator [73]. During the printing process, the seal elements, resetting (bias) springs, and SMA actuator springs are intended for automatic insertion by the handling systems. The packing material for the SMA and bias spring is ABS-based.
**Figure 1.14** Standardized SMA actuator with adaptive resetting. From Ref. [72]

**Figure 1.15** Production steps of an SMA valve actuator by 3D-printed housing. From Ref. [73]
material in the wire form which is removed by rinsing at the end of the printing process.

General Motors has developed shape memory wire-based actuators to replace the heavier conventional actuators. The Active Hatch Vent, as one of these devices, is shown in Figure 1.16 which consists of a U-shaped SMA wire, a bias spring, and an assembly to create the rotary motion necessary

![Diagram of Active Hatch Vent](image)

**Figure 1.16** A wire-based shape memory alloy actuator developed by General Motors to open a vent when the wire is heated. Courtesy of General Motors, Warren, Michigan
to open a vent. This vent actuator is incorporated in the 2014 Chevrolet Corvette as shown in Figure 1.17 and has reduced the weight of the conventional actuation systems by 0.6 kg.

Maximum recoverable strain defines the stroke of SMA actuators when the actuator applies no force. There is an inherent compromise between the actuation force or torque and the deliverable linear or rotary motion: the larger the actuation force, the smaller the maximum stroke. It is however possible to overcome the limitation in stroke, as well as the compromise between the output torque and angular displacement of SMA actuators through the use of wire-on-drum modular actuators combined with unidirectional coupling mechanisms as shown in Figure 1.18 [74, 75]. The resetting bias force is provided by a beam spring which generates a nearly constant force tangential to the drum. The SMA wire is activated through Joule heating which results in the rotation of the drum. The rotation of the drum is transferred through an overrunning clutch in one direction. When the drum is rotated backward, the SMA wire is recoiled to prepare for the next step of the actuation, while the output shaft does not rotate. A second overrunning clutch links the shaft to the frame to stop accidental unwanted rotations of the output shaft.

Figure 1.17  The shape memory alloy actuator shown in Figure 1.16 is applied to develop the Active Hatch Vent to open the vent in the trunk of a luxury sedan vehicle. Courtesy of General Motors, Warren, Michigan
Figure 1.18 (Top) Schematic of a wire-on-drum SMA rotary actuator combined with an overrunning clutch for creating a module with unlimited angular motion; (middle) the modular combined actuator; (bottom) the modules can be combined to increase the output torque or speed in series or parallel arrangement [74, 75]. Courtesy of Dragoni and Scirè Mammano, University of Modena and Reggio Emilia, Italy.
In a similar use of rectifying mechanisms, one of the early applications of NiTi wire actuators was in heat engines to harness energy from sources of heat, such as internal combustion engines. An example of this type of actuators is shown in Figure 1.19 [76]. This reciprocating ratchet-type heat engine uses an SMA wire, which is actuated by alternating flow of hot/cold water through automatic valves and thereby harvesting energy from heat. A spring provides the bias force for repeated actuation. The SMA wire is connected to a chain and sprocket, which only transmits the rotation in one direction and therefore creates a one-directional rotary motion at the output shaft.

### Figure 1.19
(a) Reciprocated heat engine employing an SMA wire. (b) Action mechanism of the heat engine. Reproduced with permission from Ref. [76], the Japan Society of Mechanical Engineers
Wire-based NiTi actuators have been developed for robotic [77] and positioning applications. An example is an automotive side mirror actuator [78]. Wire actuators are also embedded in polymers to create composite actuators in the form of beams or plates [79]. The off-center wires are installed on or embedded in both sides of the structure to create bending moment in the structure. The elastic substrates store the strain energy, which is released for repeated actuation. This type of actuators has been designed for biomimetic fins, morphing wings, flapping wings, stiffness variation, adaptive structures, and minirobotic applications [80–82].

There has long been an interest in three-dimensional SMA actuator elements including monolithic and porous structures [83]. Figure 1.20 depicts an SMA porous actuator that is additively manufactured using selective laser melting. This one-way Ti-rich NiTi actuator functions without the need to shape setting; the memorized configuration is the printing shape. In addition to the simplicity of manufacturing, the porous structure provides a very lightweight actuator with adjustable stiffness [84].

Figure 1.20 A shape memory porous beam produced in additive manufacturing using powder-bed selective laser melting [84]
1.5 Modeling and Simulation

Conducting experiments on SMAs is expensive and time consuming. Therefore, in the design of SMA actuators, it is essential to have reliable modeling platforms to avoid unnecessary trials and errors. There have been many attempts to mathematically model the SMA features over the past three decades. The resulting models can be divided into two categories: micromodels and phenomenological macromodels. In general, micromechanical-based models utilize information about the microstructure of the SMA to predict the macroscopic response. Micromechanical models are useful in understanding the fundamental phenomena of SMA, although they may not be easily deployed for engineering applications. On the other hand, phenomenological models use the principles of continuum thermodynamics to describe the material response. They are calibrated by a limited number of parameters measured at the macroscopic scale through experimental observations, and thus, they are more computationally efficient.

In a general case, the behavior of an SMA-based device could be demonstrated through the following interconnected submodels: constitutive model, phase transformation kinetics, heat transfer model, and kinematics and dynamics of the device (Figure 1.21). In this section, these submodels are discussed briefly for a representative rotary actuator. Chapter 2 presents a more fundamental and comprehensive treatment of the modeling.

A one-degree-of-freedom rotary SMA actuator is shown in Figure 1.22. This system is used as an example throughout the book. This SMA system, despite its simplicity, presents several interesting and unique properties,

![Diagram of SMA systems](image)

**Figure 1.21** Modeling of the SMA systems; four submodels are interconnected. From Ref. [85]
which make it fitting as an interesting example. To allow a more comprehensive learning experience, a complete model of the system has been made available at http://smartsys.eng.utoledo.edu. The accompanying MATLAB/Simulink code allows for simulation of the system behavior in both open- and closed-loop forms. This system and the associated model have been the subject of in a series of publications in the area of modeling and control of SMA actuators [77, 86–93].

The design parameters of the rotary SMA actuator consist of the length and diameter of the NiTi wire. The length of the wire is defined through the kinematic relationship between the angular rotation and strain of the wire:

\[ \varepsilon = -\frac{2r\theta}{l_0} \]  

(1.1)

Maximum strain takes place at the lower position of \(-45^\circ\) where the wire is assumed to be mostly in the detwinned martensite phase. At \(+90^\circ\) angular position, the actuator wire is assumed to be mostly in the austenite phase and at the minimum strain state. As shown in Chapter 5, in order to maximize the fatigue life of the SMA wire, it is recommended to limit the maximum
strain to 4%. Using this strain and Equation 1.1, the initial (short) length of the wire can be calculated, as shown in Table 1.1.

In Chapter 3, it will be shown that the stress of the SMA wire in this actuator varies as a function of the angular position of the arm. This nonlinear behavior is due to the torques generated by the bias spring and the payload being nonlinear functions of the angular position. Additionally, the relationship between the stress of the wire and angular position of the arm is modified due to angular acceleration as captured in the dynamic model:

\[ l\theta + c\dot{\theta} + \tau_{\text{spring}} + \tau_{\text{load}} = \tau_{\text{SMA}} \]  

(1.2)

### Table 1.1: Parameters of the one-degree-of-freedom SMA rotary actuator

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Description</th>
<th>Unit</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>M</td>
<td>SMA wire’s mass per unit length</td>
<td>kg</td>
<td>1.414e⁻⁴</td>
</tr>
<tr>
<td>A</td>
<td>SMA wire’s circumferential area per unit length</td>
<td>m</td>
<td>4.712e⁻⁴</td>
</tr>
<tr>
<td>C</td>
<td>Specific heat of wire</td>
<td>kcal/kg·°C</td>
<td>0.2</td>
</tr>
<tr>
<td>R</td>
<td>SMA wire’s resistance per unit length</td>
<td>Ω</td>
<td>45</td>
</tr>
<tr>
<td>T&lt;sub&gt;∞&lt;/sub&gt;</td>
<td>Ambient temperature</td>
<td>°C</td>
<td>20</td>
</tr>
<tr>
<td>H</td>
<td>Heat convection coefficient</td>
<td>J/(m²·°C·s)</td>
<td>150</td>
</tr>
<tr>
<td>E&lt;sub&gt;A&lt;/sub&gt;</td>
<td>Austenite Young’s modulus</td>
<td>GPa</td>
<td>75.0</td>
</tr>
<tr>
<td>E&lt;sub&gt;M&lt;/sub&gt;</td>
<td>Martensite Young’s modulus</td>
<td>GPa</td>
<td>28.0</td>
</tr>
<tr>
<td>θ&lt;sub&gt;T&lt;/sub&gt;</td>
<td>SMA wire’s thermal expansion factor</td>
<td>MPa/°C</td>
<td>0.55</td>
</tr>
<tr>
<td>Ω</td>
<td>Phase transformation contribution factor</td>
<td>GPa</td>
<td>-1.12</td>
</tr>
<tr>
<td>σ&lt;sub&gt;0&lt;/sub&gt;</td>
<td>SMA wire’s initial stress</td>
<td>MPa</td>
<td>75.0</td>
</tr>
<tr>
<td>ε&lt;sub&gt;0&lt;/sub&gt;</td>
<td>SMA wire’s initial strain</td>
<td>—</td>
<td>0.04</td>
</tr>
<tr>
<td>T&lt;sub&gt;0&lt;/sub&gt;</td>
<td>SMA wire’s initial temperature</td>
<td>°C</td>
<td>20</td>
</tr>
<tr>
<td>ξ&lt;sub&gt;0&lt;/sub&gt;</td>
<td>SMA wire’s initial martensite fraction</td>
<td>—</td>
<td>1.0</td>
</tr>
<tr>
<td>A&lt;sub&gt;s&lt;/sub&gt;</td>
<td>Austenite start temperature</td>
<td>°C</td>
<td>68</td>
</tr>
<tr>
<td>A&lt;sub&gt;f&lt;/sub&gt;</td>
<td>Austenite final temperature</td>
<td>°C</td>
<td>78</td>
</tr>
<tr>
<td>M&lt;sub&gt;s&lt;/sub&gt;</td>
<td>Martensite start temperature</td>
<td>°C</td>
<td>52</td>
</tr>
<tr>
<td>M&lt;sub&gt;f&lt;/sub&gt;</td>
<td>Martensite final temperature</td>
<td>°C</td>
<td>42</td>
</tr>
<tr>
<td>C&lt;sub&gt;A&lt;/sub&gt;</td>
<td>Effect of stress on austenite temperatures</td>
<td>MPa/°C</td>
<td>10.3</td>
</tr>
<tr>
<td>C&lt;sub&gt;M&lt;/sub&gt;</td>
<td>Effect of stress on martensite temperatures</td>
<td>MPa/°C</td>
<td>10.3</td>
</tr>
<tr>
<td>L&lt;sub&gt;0&lt;/sub&gt;</td>
<td>Initial length of SMA wire</td>
<td>mm</td>
<td>900</td>
</tr>
<tr>
<td>R</td>
<td>Pulley diameter</td>
<td>mm</td>
<td>8.25</td>
</tr>
<tr>
<td>m&lt;sub&gt;p&lt;/sub&gt;</td>
<td>Payload mass</td>
<td>gr</td>
<td>57.19</td>
</tr>
<tr>
<td>m&lt;sub&gt;a&lt;/sub&gt;</td>
<td>Moving link mass</td>
<td>gr</td>
<td>18.7</td>
</tr>
<tr>
<td>K</td>
<td>Bias spring stiffness</td>
<td>N/m</td>
<td>3.871</td>
</tr>
</tbody>
</table>

*Source: From Ref. [77].*
Using Equation 1.2, the maximum desirable actuation torque to overcome the maximum resistance torque of the spring and the payload and to achieve the desired angular acceleration can be calculated. This actuation torque is then used to calculate the diameter of the SMA wire, as shown in Table 1.1. The arm rotates from $-45^\circ$ to $+90^\circ$ when the NiTi wire with the calculated diameter of 150 $\mu$m is actuated by Joule heating.

While Chapter 2 includes more details on SMA modeling, in this section, a phenomenological model is presented to capture the behavior of the rotary actuator [90, 91]. This approach follows the assumption [94] that the material behavior of the alloys depends on the state variables strain ($\varepsilon$), temperature ($T$), and crystallographic phase (martensite fraction $\xi = 1$ when the alloy is 100% martensite and $\xi = 0$ when the alloy is 100% austenite), all of which are mutually dependent on each other as schematically shown in Figure 1.21. The parameters associated with this modeling approach are summarized in Table 1.1. Based on the second law of thermodynamics expressed in the Clausius–Duhem inequality form and expressing the thermodynamic potential as Helmholtz free energy, a constitutive equation captures the thermomechanical behavior of the SMA wire [95]:

$$\sigma = \sigma(\varepsilon, T, \xi)$$

$$d\sigma = E d\varepsilon + \theta_T dT + \Omega d\xi$$  \hspace{1cm} (1.3)

In Equation 1.3 which was proposed and modified by Tanaka [94], Liang and Rogers [96], and Brinson [95], $E$, $\theta_T$, and $\Omega = -\varepsilon_L E$ represent Young’s modulus, thermal expansion coefficient, and phase transformation contribution factor, respectively. $\varepsilon_L$ is the maximum recoverable (transformation) strain of the NiTi wire which is by design limited to 4% when the arm is at $-45^\circ$. For simplicity, these parameters are assumed here to be independent of the state variables.

The second part of the model represents the state of the material by calculating the martensite fraction as a function of the temperature, stress, and thermomechanical loading history. Such a model can be found by curve fitting so that the combined phase transformation–constitutive model captures macroscopic hysteretic behavior of the material. Table 1.2 summarizes the evolution of conditions for phase transformation.

For the rotary actuator, the SMA wire, in heating or mechanical unloading, transforms from detwinned martensite to austenite. In cooling, the transformation is back to detwinned martensite. The wire does not transform to twinned martensite due to the presence of the bias torque. A phase transformation model in general includes the forward and reverse transformations as
well as the detwinning effect. The associate equation to capture the reverse transformation can be written as

$$\dot{\xi} = \frac{\xi_M}{2} \cos[a_A(T - A_s) + b_A \sigma] + 1$$ (1.4)

The forward transformation takes place when the wire is cooled or mechanically loaded. Variation of the martensite phase can be modeled with the following equation:

$$\frac{\dot{T}}{C_A} > 0 \quad \frac{T - \dot{\xi}}{\frac{C_M}{A}} < 0$$

where $\xi_M$ is the minimum martensite fraction reached during heating or unloading and $\xi_A$ is the maximum martensite fraction reached during cooling or loading.

In the more common bias type SMA actuation the material transforms to detwinned martensite during cooling. In certain SMA actuators, unlike the rotary actuation of this example, there is a possibility of transforming to twinned martensite when the actuation element is cooled under $M_s$ at low level of stress below $\sigma_s^\tau$. This element, when consequently heated, transforms to austenite. As an example, this phenomenon takes place, as shown in Figure 1.23, when two SMA wires form an antagonistic actuation–bias mechanism [97]. Figure 1.24 shows a similar antagonistic microactuator consisting of a bistable curved beam, two SMA active elements, and laser sources [98]. The laser beam is a contactless energy transfer source to activate the SMA elements. In each direction, one SMA active element actuates the bistable curved

<table>
<thead>
<tr>
<th>Author</th>
<th>Martensite to austenite transformation</th>
<th>Austenite to martensite transformation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tanaka [94]</td>
<td>$\dot{\xi} &lt; 0$ \quad $\sigma \leq - (T - A_s) A_2 / B_a$</td>
<td>$\dot{\xi} &gt; 0$ \quad $\sigma \geq (T - M_s) A_m / B_m$</td>
</tr>
<tr>
<td>Liang and Rogers [96]</td>
<td>$\dot{T} &gt; 0$ \quad $A_t + \frac{\sigma}{C_A} \geq T \geq A_s + \sigma / C_A$</td>
<td>$\dot{T} &lt; 0$ \quad $M_m + \frac{\sigma}{C_M} \geq T \geq M_f + \sigma / C_M$</td>
</tr>
<tr>
<td>Brinson [95]</td>
<td>$\dot{\xi} &lt; 0$ \quad $C_A (T - A_s) \geq \sigma \geq C_A (T - A_t)$</td>
<td>$\dot{\xi} &gt; 0$ \quad $C_M (T - M_s) \geq \sigma \geq C_M (T - M_f)$</td>
</tr>
<tr>
<td>Elahinia [85]</td>
<td>$\dot{T} - \frac{\sigma}{C_A} &gt; 0$ \quad $A_t + \sigma / C_A \geq T \geq A_s + \sigma / C_A$</td>
<td>$\dot{T} - \frac{\sigma}{C_M} &lt; 0$ \quad $M_m + \sigma / C_M \geq T \geq M_f + \sigma / C_M$</td>
</tr>
</tbody>
</table>

Source: From Ref. [91].
beam between its two stable positions. The memorized shape of each of the two SMA elements is the predetermined flat shape, which is achieved upon heating (Figure 1.24b). During the actuation, the passive antagonistic SMA element is also deformed and transformed to stress-induced martensite to prepare for the next step of actuation. The reverse switching process is obtained by heating the second SMA active element as shown in Figure 1.24c. At the end of each actuation cycle, when each active element is cooled, it transforms to twinned martensite. Other examples of such actuators as antagonistic shape memory–superelastic systems are highlighted in Chapter 3.

To capture the behavior of antagonistic SMA actuators, Brinson divided martensite volume fraction into two portions: temperature induced (twinned) and stress induced (detwinned) [95]. These fractions together form the total martensite fraction

$$\xi = \xi_T + \xi_S$$  (1.6)
and the constitutive equation is simplified as

$$\sigma = E\varepsilon + \Omega \xi_S = E(\varepsilon - \varepsilon_L \xi_S)$$

To capture various actuation possibilities, a phase transformation plot can be divided into 9 regions as shown in Figure 1.25. In the system of Figure 1.23, before the actuation and when the actuator is in the middle neutral position, the two wires are assumed to be 50% temperature-induced and 50% stress-induced martensite. Upon complete actuation of each of the two wires, the actuating wire is transformed to austenite (region 4 in Figure 1.25), while the opposing wire is transformed to 100% stress-induced martensite (region 0 in Figure 1.25). In this antagonistic actuator, when the active wire is cooled at the end of actuation, it transforms to twinned martensite (region 8 in Figure 1.25) as long as the following two conditions are simultaneously satisfied:

$$\dot{T} - \frac{\sigma}{C_M} < 0 \text{ and } M_f + \frac{\sigma}{C_M} < T < M_s + \frac{\sigma}{C_M}$$

The transformation of the stress- and temperature-induced martensite variants can be described as
where $\xi_{AT}$ is the twinned martensite fraction prior to this transformation. When this wire is consequently stretched (toward region 0 in Figure 1.25) by actuation of the antagonistic wire, the following opposing transformations take place:

$$\xi_s = \frac{1 - \xi_{AS}}{2} \cos \left[ \frac{\pi}{\sigma_s - \sigma_f} \left( \sigma - \sigma_l - C_M (T - M) \right) \right] + \frac{1 + \xi_{AS}}{2}$$

$$\xi_T = -\xi_s$$

when the following four conditions are simultaneously satisfied:

$$\hat{\sigma} > 0, \sigma > \sigma_s^{cr}, \sigma < \sigma_f^{cr}, \text{ and } M + \frac{\sigma}{C_M} > T$$

In Equation 1.10, $\xi_{AS}$ is the stress-induced martensite fraction prior to this transformation.

To capture other possible behaviors of the SMA element, this section summarizes the behavior in all nine zones as depicted in Figure 1.25 [99]. Within each region, the stress-induced martensite $\xi_s$ and the temperature-induced martensite $\xi_T$ are defined on the basis of the tensile stress $\sigma$ and the temperature $T$, in addition to the initial stress-induced martensite $\xi_{S0}$ and the initial temperature-induced martensite $\xi_{T0}$. It is worth noting that the values $\xi_{S0}$ and $\xi_{T0}$ have to be updated every time the boundary between two regions is crossed. Depending on the thermomechanical loading history, it is possible that the material may exist in a region with or without a phase transformation.

**Region 0**: Pure stress-induced martensite

$$\xi_s = 1$$

$$\xi_T = 0$$

**Region 1**: Transformation from austenite or from twinned martensite to detwinned martensite

$$\xi_s = \frac{1 - \xi_{S0}}{2} \cos \left[ \frac{\pi}{\sigma_s^{cr} - \sigma_f^{cr}} \left( \sigma - \sigma_s^{cr} - C_M (T - M) \right) \right] + \frac{1 + \xi_{S0}}{2}$$

Shape Memory Alloy Actuators
\[ \xi_T = \xi_{T0} - \frac{\xi_{T0}}{1 - \xi_{S0}}(\xi_S - \xi_{S0}) \]  

(1.14)

**Region 2:** Mixture of twinned and detwinned martensite and austenite, depending on the loading/temperature history; no transformation occurs in this area

\[ \xi_S = \xi_{S0} \]  

(1.15)

\[ \xi_T = \xi_{T0} \]  

(1.16)

**Region 3:** Transformation from martensite to austenite

\[ \xi = \frac{\xi_0}{2} \left[ \cos \left( a_A \left( T - A_s - \frac{\sigma}{C_A} \right) \right) + 1 \right] \]  

(1.17)

\[ \xi_S = \xi_{S0} - \frac{\xi_{S0}}{\xi_0} (\xi_0 - \xi) \]  

(1.18)

\[ \xi_T = \xi_{T0} - \frac{\xi_{T0}}{\xi_0} (\xi_0 - \xi) \]  

(1.19)

**Region 4:** Pure austenite, no transformation

\[ \xi_S = 0 \]  

(1.20)

\[ \xi_T = 0 \]  

(1.21)

**Region 5:** Transformation from austenite or twinned martensite to detwinned martensite and transformation from austenite to twinned martensite

\[ \xi_S = \frac{1 - \xi_{S0}}{2} \cos \left( \frac{\pi}{\sigma_{S}^{\text{tr}} - \sigma_{T}^{\text{tr}}} (\sigma - \sigma_{T}^{\text{tr}}) \right) + \frac{1 + \xi_{S0}}{2} \]  

(1.22)

\[ \xi_T = \Delta T \xi - \frac{\Delta T \xi}{1 - \xi_{S0}} (\xi_S - \xi_{S0}) \]  

(1.23)

where if \( T < T_0 \),

\[ \Delta T \xi = \frac{1 - \xi_{S0} - \xi_{T0}}{2} \cos \left( \frac{\pi}{M_s - M_t} (T - M_t) \right) + \frac{1 - \xi_{S0} + \xi_{T0}}{2} \]  

(1.24)
otherwise,

\[ \Delta T^e = \xi^e T \]  

(1.25)

**Region 6**: Transformation from austenite to twinned martensite

\[ \xi^e S = \xi^e S_0 \]  

(1.26)

\[ \xi^e T = \frac{1 - \xi^e 0}{2} \cos \left[ a_M(T - M_t) \right] + \frac{1 + \xi^e 0}{2} - \xi^e S_0 \]  

(1.27)

**Region 7**: Transformation from twinned martensite to detwinned martensite

\[ \xi^e S = \frac{1 - \xi^e 0}{2} \cos \left[ \frac{\pi}{\sigma_{CT} - \sigma_{CT}^f} (\sigma - \sigma_{CT}^f) \right] + \frac{1 + \xi^e 0}{2} \]  

(1.28)

\[ \xi^e T = \xi^e T_0 - \frac{\xi^e T_0}{1 - \xi^e S_0} (\xi^e S - \xi^e S_0) \]  

(1.29)

**Region 8**: Twinned and detwinned martensite are present, depending on the loading/temperature history; no transformation occurs

\[ \xi^e S = \xi^e S_0 \]  

(1.30)

\[ \xi^e T = \xi^e T_0 \]  

(1.31)

1.5.1 Examples

Using the model presented in this section, it is possible to compute the martensite fraction for various thermomechanical \( \sigma - T \) paths. To allow a more comprehensive learning experience, a complete Wolfram Mathematica implementation of the model is available at http://smartsys.eng.utoledo.edu. The accompanying code allows for simulation of various \( \sigma - T \) paths.

As the first example, we consider the superelastic effect, described by the path A–B–C–D–E–F–A in Figure 1.26. The point A is in the region 4, where \( \xi^e S = \xi^e T = 0 \) (Eqs. 1.20 and 1.21), so the alloy is completely in the austenite phase. By imposing a stress (tensile for a wire), the alloy crosses zone 4, zone 3 and zone 2, where no transformation occurs and therefore, stress-induced martensite and temperature-induced martensite are still both 0.
If the stress is further increased, it is possible to arrive at point B in zone 1, where the austenite phase starts to transform to detwinned martensite (Eqs. 1.13 and 1.14). The transformation occurs within the whole region up to point C, where \( \xi_S = 1, \xi_T = 0 \). These values of martensite fractions are considered as the initial values \( (\xi_{S0} = 1, \xi_{T0} = 0) \) in the transformation in the following zone (zone 0, Eqs. 1.11 and 1.12). After point C, the stress can be increased and no transformation occurs up to the martensite yield stress, where the plastic flow begins (not considered here). It is worth noting that the yield stress can occur only for stress induced martensite.

When the stress decreases (D–E), the zones 1 and 2 are crossed again, and no transformation occurs as \( \xi_{S0} = 1, \xi_{T0} = 0 \) (Eqs. 1.13–1.16). By further decreasing the stress, E–F, region 3 is crossed and the inverse transformation from martensite to austenite occurs (Eqs. 1.17–1.19) up to point F, where \( \xi_S = 0, \xi_T = 0 \). Finally, zone 4 is reached. Here, it is necessary to reupdate the initial values of martensite fractions to \( \xi_{S0} = 0, \xi_{T0} = 0 \) which remains constant in the zone, and the starting condition is restored (Eqs. 1.20 and 1.21).

As another example, we consider the SME, depicted by the path A–B–C–D–E–F–G–H–A in Figure 1.27. The point A is in the region 8 where, if previously no stress (for example tension for wire) is applied to the alloy, only twinned martensite is present (Eqs. 1.30 and 1.31, \( \xi_S = 0, \xi_T = 1 \)). By applying tension, point B is reached where the martensite detwinning transformation begins.
During the path B–C, the alloy is in the zone 7 where the stress-induced martensite increases, and the temperature-induced martensite decreases as the tension rises (Eqs. 1.28 and 1.29). Once point C is reached, only stress-induced martensite is present (Eqs. 1.11 and 1.12, $\xi_s = 1$, $\xi_T = 0$). If the stress is still increased, no transformation will occur up to the martensite yield stress, where the plastic flow begins. These martensite fraction values will be considered as the initial values for the transformation in zone 0 ($\xi_{S0} = 1$, $\xi_{T0} = 0$, path C–D) where they remain constant. By removing the stress, zone 7 and zone 8 are crossed where no transformation occurs (Eqs. 1.28–1.31). At the end of this first part of the path (A–B–C–D–E), martensite fractions are computed in zone 8, point E, with $\xi_{S0} = 1$, $\xi_{T0} = 0$.

Subsequently, the alloy is heated (A–F–G–H). Up to point F, zone 6 and zone 2 are crossed and no transformation occurs since $\xi_{S0} = 1$, $\xi_{T0} = 0$ (Eqs. 1.15 and 1.16 and Eqs. 1.26 and 1.27). Then zone 3 is crossed (F–G) and the inverse transformation occurs (from martensite to austenite; Eqs. 1.17–1.19) in zone 3. Once point G is reached, martensite has completely transformed to austenite and the martensite fraction values have to be considered as the initial values for the following zone (zone 4, $\xi_{S0} = 0$, $\xi_{T0} = 0$). Finally, the alloy is cooled following the path H–G–F–I–J–A. Up to point I, no transformation occurs because in zones 3 and 2 martensite fractions remain constant if $\xi_{S0} = 0$, $\xi_{T0} = 0$ (Eqs. 1.15–1.19). Then zone 6 is crossed where the
austenite completely transforms to twinned martensite (Eqs. 1.26 and 1.27)
and the initial condition is restored in zone \(8 (\xi_{S} = 0, \xi_{T} = 1)\).

1.6 Summary

This chapter introduced the unique properties of SMAs such as the SME and
superelasticity. The underlying microstructural phenomena corresponding
to such behaviors were briefly discussed, and the phase diagram was intro-
duced as a beneficial tool to schematically represent the associated phase
transformations. As mentioned, although NiTi is the most common SMA,
other shape memory materials are also available for specific applications.
A well-known phenomenological modeling approach was presented for the
design and analysis of the SMA actuators.

References

[1] Ölander A. The crystal structure of AuCd. Zeitschrift Fur Kristallographie
1932;\textbf{83}:145–8.
Journal of the American Chemical Society 1932;\textbf{54}:3819–33.
[3] Chang L, Read T. Plastic Deformation And Diffusionless Phase Changes In
Metals—the Gold–Cadmium Beta-phase. Transactions of the American Institute
of Mining and Metallurgical Engineers 1951;\textbf{191}:47–52.
the Mechanical Properties of Alloys Near Composition TiNi. Journal of Applied


Introduction


Introduction


